



André Bögelsack · Utpal Chakraborty  
Dhiraj Kumar · Elena Wolz  
Johannes Rank · Jessica Tischbirek

# SAP S/4 HANA- Systeme in Hyperscaler Clouds

Architektur, Betrieb und Setup von  
SAP S/4HANA-Systemen in Microsoft  
Azure, Amazon Web Services und  
Google Cloud

EBOOK INSIDE



Springer Vieweg

---

## SAP S/4 HANA-Systeme in Hyperscaler Clouds

---

André Bögelsack · Utpal Chakraborty ·  
Dhiraj Kumar · Elena Wolz · Johannes Rank ·  
Jessica Tischbirek

# SAP S/4 HANA-Systeme in Hyperscaler Clouds

Architektur, Betrieb und Setup von  
S/4HANA-Systemen in Microsoft Azure,  
Amazon Web Services und Google Cloud

André Bögelsack  
Accenture AG  
Bern, Schweiz

Utpal Chakraborty  
Accenture Ltd.  
Kolkata, Indien

Dhiraj Kumar  
Accenture Ltd.  
Kolkata, Indien

Elena Wolz  
Technische Universität München  
Ismaning, Deutschland

Johannes Rank  
Technische Universität München  
Irschenberg, Deutschland

Jessica Tischbierek  
Google Germany GmbH  
München, Deutschland

ISBN 978-3-658-34474-0

ISBN 978-3-658-34475-7 (eBook)

<https://doi.org/10.1007/978-3-658-34475-7>

Die Deutsche Nationalbibliothek verzeichnetet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

© Der/die Herausgeber bzw. der/die Autor(en), exklusiv lizenziert durch Springer Fachmedien Wiesbaden GmbH, ein Teil von Springer Nature 2022

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jedermann benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des jeweiligen Zeicheninhabers sind zu beachten.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag, noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Planung: Petra Steinmüller

Springer Vieweg ist ein Imprint der eingetragenen Gesellschaft Springer Fachmedien Wiesbaden GmbH und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

---

## Vorwort

Der stetige technologische Wandel und der Zwang zur Digitalisierung wurden noch nie so deutlich, wie in den vergangenen zwei Jahren, als die Corona Pandemie die Welt in Atem hielt und sich die Arbeitsgewohnheiten drastisch änderten. Spätestens zu diesem Zeitpunkt wurde aus dem Chief Information Officer (CIOs) ein Chief Digitalization Officer mit neuen Herausforderungen und neuen Anforderungen aus der allgemeinen Arbeitswelt.

Glücklich konnten sich alle die CIOs schätzen, welche bereits konsequent auf eine Digitalisierung vor Corona gesetzt hatten und den technischen Wandel nicht als Herausforderungen, sondern als Chance verstanden hatten. Die Unternehmen, welche bereits vor der Pandemie auf eine hohe Flexibilität und Skalierung gesetzt haben und die IT-Systeme bereits in die Public Cloud verschoben hatten, waren bestens gewappnet. Sie konnten das volle Potenzial der Public Cloud weiter an die Mitarbeiterinnen und Mitarbeiter geben.

Mit der gleichen Geschwindigkeit, wie die Corona-Pandemie die Adoption der Cloud vorangetrieben hat, müssen die CIOs nun auf die zweite große Herausforderung bei der IT eingehen: SAP S/4HANA. Der Wandel von den traditionellen SAP ERP Systemen hin zu einem digitalen Kern, basierend auf SAP S/4HANA, ist in vielen Unternehmen eine gewaltige Aufgabe und erfordert die enge Zusammenarbeit der Fachabteilungen und der IT-Abteilung.

Durch die Kombination der Nutzung von Public Cloud-Diensten mit der Transformation zu SAP S/4HANA, erschließen sich den Unternehmen neue Möglichkeiten. So können Geschäftsprozesse neu und flexibler gestaltet werden und die Unternehmen schaffen sich eine neue, hoch flexible technologische Plattform für den weiteren Erfolg des Unternehmens.

Dieses Buch zeigt, wie Unternehmen sich auf einfachstem Wege eine neue technische Plattform für die zukünftigen SAP S/4HANA-Systeme in der Public Cloud aufbauen können. Hierbei beschränkt sich das Buch jedoch nicht nur auf eine der großen Public Clouds, sondern beschreibt die drei wichtigsten Marktteilnehmer: Microsoft Azure, Amazon Web Services und Google Cloud. Somit erhält der Leser einen umfassenden

Überblick zu den einzelnen Public Clouds und lernt, wie diese bestmöglich für die SAP S/4HANA-Systeme eingesetzt werden können.

Das Buch ist eine Pflichtlektüre vor der Einführung für alle Unternehmen, welche noch nicht SAP S/4HANA einsetzen und sich diese technische Plattform noch schaffen werden. Es ist auch für alle die Unternehmen geeignet, welche bereits in einer Public Cloud sind, aber noch Inspirationen erlangen wollen und sehen wollen, wie andere Unternehmen diesen Wandel vollziehen.

Ich freue mich, dass die Autorinnen und Autoren das Thema so umfassend, aber prägnant in einem Buch zusammengefasst haben. Die technische Tiefe und die thematische Breite sind beachtlich und spiegeln die Wichtigkeit des Themas wider. Das Buch hilft Ihnen, liebe Leser, die wichtigen Fragestellungen bei der Bereitstellung, der Migration und dem Betrieb von SAP S/4HANA-Systemen in den Public Clouds zu beantworten. Ich wünsche Ihnen viel Freude damit!

Prof. Dr. Alexander Zeier  
1. Cloud Fellow  
CTO ASBG  
Global Managing Director  
Accenture GmbH

---

# Danksagung

Buchprojekte sind immer eine große Herausforderung für die Autoren und benötigen viel Zeit und Energie, welche nicht für andere Dinge zur Verfügung steht. Daher möchten die Autoren die Gelegenheit nutzen, sich bei den Familien, den Kolleginnen und Kollegen als auch den Projektpartner und Kunden zu bedanken.

## Individuelle Danksagungen

### *André Bögelsack*

Ich danke meiner Frau Kathrin und meinen beiden Kindern für ihr Verständnis, wenn ich an den Wochenenden oder an den Abenden eher am Laptop saß, anstatt mit ihnen Zeit zu verbringen. Ein besonderer und spezieller Dank geht an **Niklas Feil**, der mir sehr tatkräftig zur Seite stand und viel Inhalt zum Buch liefert hat. Vielen Dank hierfür!

Mein Dank geht an alle Koautoren, ohne deren Expertenwissen und deren Zeiteinsatz wir das Buch niemals geschrieben hätten. Ich bin sehr glücklich, dass wir das gemeinsam geschafft haben.

Zu guter Letzt möchte ich mich bei Debolina Banerjee für die Motivation zum Buch schreiben bedanken und bei Samiksha Munjal für das Freihalten meines Rückens, wenn ich zu sehr ins Buch schreiben abgetaucht war.

### *Utpal Chakraborty*

Zuallererst möchte ich Dr. André Bögelsack meinen Dank aussprechen, der mir die Möglichkeit gegeben hat, mein Wissen und meine Erfahrung durch dieses Buch einem breiteren Publikum zugänglich zu machen, was ohne ihn nicht möglich gewesen wäre. Ich möchte auch meiner Familie danken, die mich in dieser pandemischen Situation unterstützt hat, und ich widme meine Leistung meiner Tochter Sreejani.

Zu guter Letzt ein großes Dankeschön an meine Co-Autoren, ohne deren Unterstützung dies nicht möglich gewesen wäre.

*Dhiraj Kumar*

Mein Dank gilt meiner Frau Shikha für ihre ständige Motivation und ihr Verständnis und meinem Sohn Agastya dafür, dass er mir die Zeit für dieses Buch gelassen hat. Mein besonderer Dank gilt meinen Freunden Ravi Kumar und Ravi Surana, Kunal, Misal und Sushobhit, die mich mit ihren Wochenendanrufen am Leben erhalten (und geerdet) haben.

Ein besonderer Dank geht an Dr. André Bögelsack, der mehr an mich glaubte, als ich an mich selbst. Danke André, dass du immer für mich da bist. Dieses Buch hat seine Form durch die großartige Zusammenarbeit mit meinen Co-Autoren angenommen. Ihnen allen gilt mein ganz besonderer Dank für ihre Zeit und ihr Fachwissen.

Nicht zuletzt danke ich meinem Bruder Manoj Kumar, der mich immer inspiriert und ermutigt, was immer ich auch tue. Danke Bhaiya.

*Elena Wolz*

Meiner Familie und Freunden möchte ich danken für die Unterstützung und das Verständnis, wenn ich die Abende in Gesellschaft mit Lastenausgleichsmodulen und Verfügbarkeitsgruppen verbracht habe, anstatt mit ihnen. Auch möchte ich dem SAP UCC TU München für die Unterstützung bei dem Projekt danken, sowie meinen Kollegen am SAP UCC TU München für den immerwährenden Zuspruch. Mein Dank gilt auch den Mitautoren, die durch ihre Beiträge und dem gemeinsamen Austausch zu neuen Perspektiven verholfen haben. Ein besonderer Dank geht an Johannes Rank, ich habe die enge Zusammenarbeit sehr geschätzt.

*Johannes Rank*

Ein herzliches Dankeschön geht an meine Co-Autoren dieses Buches, insbesondere an André der es mir überhaupt ermöglicht hat meine Expertise beizusteuern und ohne dessen dieses Buch wohl nicht möglich gewesen wäre. Ganz besonders möchte ich mich auch bei meiner Familie bedanken. Bei meiner Tochter, die viele Wochenenden und Abende auf ihren Papa verzichten musste und bei meiner Frau, die mir stets den Rücken freihält.

*Jessica Tischbirek*

Zuerst möchte ich meinen Co-Autoren für den Austausch und besonders Dr. André Bögelsack für die Anfrage, Ideen und Möglichkeit an diesem Buchprojekt mitzuwirken danken.

Ein riesengroßes Dankeschön geht an alle meine Kollegen und Kolleginnen, die dieses Buch durch ihre Ideen, Motivation und vor allem ihre Durchsicht inhaltlich unterstützt haben: Hasmig Samurkashian, Thorsten Staerk, Benjamin Schuler, Alexander Schmid, Abdelkader Sellami, Alison Hettrick und an meinen Manager Ian Moria.

Schließlich geht ein ganz besonderer Dank an meinen Partner Mike für seine andauernde Unterstützung, kritische Fragen sowie sein Verständnis für das intensive Buchprojekt.

## Gemeinsame Danksagung

### *Accenture*

Das Autorenteam dankt allen, welche zu dem Erfolg dieses Buchs beigetragen haben und die das Buch bei der Entstehung begleitet haben. Dazu gehören insbesondere Tobias Regenfuss, Alan Mohr, Veronica Wolters, Bernhard Schulzki und das gesamte Führungs- team von **Accenture Cloud First**. Ohne ihren Einsatz und Unterstützung wäre dieses Buch nicht entstanden.

### *SAP*

Wir bedanken uns auch bei SAP und dem SAP University Alliances (SAP UA) Programm, welches schon seit vielen Jahren die Hochschulen weltweit unterstützt und auf diesem Wege Buchpublikationen mit ermöglicht.

### *Partner und Kunden*

Ein Buch über mehrere Hyperscaler zu schreiben, kann nur durch die Unterstützung von Microsoft, Amazon und Google gelingen. Ohne diese Unterstützung bei technischen Fragen wäre das Buch nicht zustande gekommen. Hierfür bedankt sich das gesamte Autorenteam – insbesondere bei den Alliances-Teams der Hersteller.

Um das Buch mit vielen Beispiele für den Kunden spannender zu machen, bedarf es auch der Unterstützung der jeweiligen Unternehmen. Gerne bedanken wir uns bei LANXESS und insbesondere Tobias Greskamp für seine Unterstützung.

### *Technische Universität München*

Zusätzlich bedanken wir uns bei der Technischen Universität München, dem Lehrstuhl für Wirtschaftsinformatik (i17) und dem SAP University Competence Center (SAP UCC), welche eine exzellente Basis für die Forschung und Entwicklung im Umfeld von SAP und Cloud Computing bieten. Ein besonderer Dank geht an die Leitung des SAP UCC, Dr. Holger Wittges, und an den Lehrstuhlinhaber Prof. Dr. Helmut Krcmar.

### *Google*

Wir bedanken uns bei Google Cloud und allen Teams, die im Umfeld von SAP auf Google Cloud tätig sind: Solution und Sales Specialists, Produktengineering, das Center of Excellence, das SAP Management, das SAP Partnerschaftsteam und viele mehr. Ein ganz besonderer Dank geht an das globale SAP Marketing Team und das Accenture Partner Alliances Team.

*Springer Verlag*

Nicht zuletzt bedanken wir uns bei unseren Lektorinnen und Lektoren, welche uns während der gesamten Zeit mit Rat und Tat zur Seite standen. Petra Steinmüller, Heike Jung und David Imgrund hatten immer die passenden Lösungen und Ratschläge für uns parat.

André Bögelsack  
Utpal Chakraborty  
Dhiraj Kumar  
Elena Wolz  
Johannes Rank  
Jessica Tischbierek

---

# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung und Einführung zu Hyperscaler Clouds</b>	1
1.1	Einleitung . . . . .	1
1.2	Cloud Computing allgemein . . . . .	4
1.2.1	Wichtigste Eigenschaften der Cloud . . . . .	4
1.2.2	Public, Private und Hybrid Clouds . . . . .	6
1.3	Die Clouds der wichtigsten Marktteilnehmer im Überblick . . . . .	8
1.3.1	Amazon Web Services . . . . .	8
1.3.2	Microsoft Azure . . . . .	10
1.3.3	Google Cloud . . . . .	13
1.3.4	Weitere Marktteilnehmer . . . . .	15
1.4	Die Cloud Strategie von SAP . . . . .	17
1.4.1	SAP RISE Programm . . . . .	17
1.4.2	SAP HANA Enterprise Cloud . . . . .	21
1.4.3	Zusätzliche SAP Services aus der Cloud . . . . .	24
1.5	Zusammenfassung . . . . .	27
<b>2</b>	<b>SAP S/4HANA-Systeme in den Public Clouds</b>	29
2.1	Überblick zur SAP S/4HANA Architektur . . . . .	29
2.1.1	Entwicklung zu SAP S/4HANA . . . . .	29
2.1.2	SAP S/4HANA in der Cloud . . . . .	30
2.1.3	SAP S/4HANA Architektur . . . . .	31
2.2	Anwendungsfälle für SAP S/4HANA in der Cloud . . . . .	35
2.2.1	Sizing . . . . .	35
2.2.2	Kosten . . . . .	42
2.2.3	Hochverfügbarkeit . . . . .	48
2.2.4	Desaster Recovery . . . . .	53
2.2.5	Backup und Restore . . . . .	58
2.2.6	Integration und Netzwerk . . . . .	65

2.2.7	Automation .....	71
2.2.8	Horizontale und vertikale Skalierung .....	75
2.3	Zusammenfassung .....	79
	Literatur .....	80
<b>3</b>	<b>Deployment und Migration von SAP S/4HANA-Systemen.....</b>	<b>81</b>
3.1	Rahmenbedingungen .....	81
3.1.1	Vertragliche Basis und Support .....	82
3.1.2	GxP Regulatorien und wichtige Zertifikate .....	86
3.1.3	Management der S/4 Landschaft .....	91
3.2	Auswahl eines Hyperscalers .....	94
3.2.1	Quantitative Faktoren .....	95
3.2.2	Qualitative Faktoren .....	96
3.3	Deployment und Migration .....	98
3.3.1	Notwendigkeit der Transformationen .....	98
3.3.2	Neue Ansätze durch die Cloud .....	100
3.3.3	Greenfield-Deployment .....	100
3.3.4	Brownfield-Deployment .....	101
3.3.5	Migrationsszenarien zur S/4HANA Transformation .....	103
3.4	SAP S/4HANA als SaaS out-of-the-cloud .....	104
3.4.1	Überblick .....	104
3.4.2	Unterschiede zwischen den Editionen .....	105
3.4.3	Vorteile und Nachteile .....	108
3.4.4	Anwendungsfälle .....	109
3.5	Zusammenfassung .....	109
<b>4</b>	<b>SAP S/4 on Amazon AWS – Konzepte und Architekturen .....</b>	<b>111</b>
4.1	Die Geschichte von Amazon AWS .....	111
4.1.1	Neue Infrastruktur und Innovationen .....	112
4.1.2	Vorteile von Amazon Web Services .....	114
4.1.3	Das vergangene AWS-Geschäftswachstum .....	115
4.2	Cloud-Services-Angebot .....	116
4.2.1	Computing .....	117
4.2.2	Speicher .....	117
4.2.3	Netzwerkpflege .....	119
4.2.4	Identity Management .....	121
4.2.5	Sicherheit und Compliance .....	121
4.2.6	Verwaltungswerzeuge .....	122
4.3	Regionen und Availability Zones .....	125
4.3.1	AWS-Region .....	125
4.3.2	AWS Availability Zones .....	125
4.3.3	AWS Placement Groups .....	126

4.4	AWS-Managementkonsole .....	130
4.4.1	Unterstützte Browser .....	131
4.4.2	Benutzer für die Verwaltung .....	131
4.4.3	IAM-Rolle und IAM-Benutzergruppe .....	133
4.4.4	Erste Schritte mit der AWS-Managementkonsole .....	133
4.4.5	Auswahl einer anderen Region .....	134
4.4.6	Navigation zu den Services .....	134
4.4.7	Suche nach einem AWS-Service .....	134
4.5	Integration in die Kerndienste .....	135
4.5.1	Wichtigste Komponenten .....	135
4.5.2	Interne Zugriffsarchitektur .....	138
4.5.3	Interner und begrenzter externer Zugriff .....	140
4.5.4	Interner und vollständiger externer Zugriff .....	142
4.5.5	Active Directory .....	147
4.5.6	Dynamic Host Configuration Protocol (DHCP) .....	158
4.6	Zusammenfassung .....	161
<b>5</b>	<b>SAP S/4 on Amazon AWS – Deployment .....</b>	<b>165</b>
5.1	Architekturbeispiel .....	165
5.1.1	Bereitstellungsszenarien .....	166
5.1.2	Einrichtung eines SAP-HA-Clusters .....	171
5.2	Computing .....	175
5.2.1	Verfügbare EC2-Instanzen .....	176
5.2.2	Lizenzierung der EC2-Instanzen .....	177
5.2.3	SAP-HANA-zertifizierte EC2-Instanztypen .....	177
5.2.4	VM-Bereitstellung .....	178
5.3	Datenspeicher .....	190
5.3.1	Amazon EBS .....	190
5.3.2	Amazon EC2-Instanzspeicher .....	192
5.3.3	Amazon EFS .....	192
5.3.4	Amazon S3 (Simple Storage Service) .....	192
5.3.5	Speicherlayout für S/4HANA-Systeme .....	195
5.4	PAYG vs. Commitment – Rechner .....	195
5.4.1	On-Demand-Instanzen .....	198
5.4.2	Reserved Instances .....	199
5.4.3	Zahlungsoptionen .....	200
5.4.4	AWS-Rechner .....	200
5.5	Sicherheit bei AWS .....	202
5.5.1	Amazon Virtual Private Cloud (VPC) .....	204
5.5.2	Subnetz- und Routingtabellen .....	205
5.5.3	Netzwerkzugriffskontrollliste (ACL) .....	206
5.5.4	Sicherheitsgruppen (SG) .....	206

5.5.5	Virtual Private Gateway . . . . .	206
5.5.6	Internet Gateway . . . . .	207
5.6	Sicherung und Wiederherstellung in AWS . . . . .	207
5.6.1	SAP-HANA-DB-Sicherung bzw. -Wiederherstellung mit AWS Backint Agent . . . . .	207
5.6.2	AWS-Backup . . . . .	211
5.6.3	Wiederherstellung des Snapshots von der AWS-Konsole . . . . .	215
5.7	Disaster Recovery mit AWS . . . . .	217
5.7.1	Passive DR-Architektur [RPO ist größer als 0 (null) und RTO ist höher] . . . . .	217
5.7.2	Semiaktive DR-Lösung [RPO ist nahe null und RTO ist medium] . . . . .	219
5.7.3	Aktive DR-Lösung [RPO nahe 0 (null) und RTO sehr gering] . . . . .	221
5.8	Zusammenfassung . . . . .	222
<b>6</b>	<b>SAP S/4 on Microsoft Azure – Konzepte und Architekturen . . . . .</b>	<b>223</b>
6.1	Historischer Überblick über Microsoft Azure . . . . .	224
6.2	Azure Steuerung und Abonnements . . . . .	225
6.2.1	Azure Organisationsstruktur . . . . .	226
6.2.2	Abonnement Modelle . . . . .	227
6.2.3	Abonnement Management . . . . .	228
6.3	Azure Ressourcenmanagement . . . . .	231
6.3.1	Ressourcenbereitstellung mittels Resource Manager . . . . .	232
6.3.2	Allgemeine Bereitstellungsoptionen . . . . .	235
6.3.3	Rechenleistung . . . . .	237
6.3.4	Speicher . . . . .	249
6.3.5	Netzwerk und Services . . . . .	258
6.3.6	Support und Lizensierung . . . . .	265
6.4	Azure Disaster Recovery Dienste . . . . .	266
6.4.1	Azure Backup . . . . .	266
6.4.2	Backup Vault . . . . .	266
6.4.3	Recovery Services Vault . . . . .	267
6.5	S/4 auf Azure Architektur . . . . .	269
6.5.1	Grundlegende Referenzarchitektur . . . . .	269
6.5.2	Virtuelles Privates Netzwerk Referenzarchitektur . . . . .	272
6.6	Zusammenfassung . . . . .	273
	Literatur . . . . .	273

---

<b>7 SAP S/4 on Microsoft Azure – Deployment . . . . .</b>	<b>275</b>
7.1 Azure S/4HANA Beispielarchitektur. . . . .	275
7.2 Bereitstellen einer Netzwerkbasiskonfiguration via Azure Cloud Portal . . . . .	277
7.2.1 Azure Cloud Portal . . . . .	277
7.2.2 Namenskonventionen und der Ressourcenbegriff . . . . .	278
7.2.3 Anlegen der Ressourcengruppen . . . . .	279
7.2.4 Netzwerkkonfiguration . . . . .	280
7.2.5 Anlegen der Netzwerksicherheitsgruppen . . . . .	282
7.3 Bereitstellen eines HANA HA Clusters. . . . .	284
7.3.1 HANA Cluster Architektur . . . . .	285
7.3.2 Bereitstellung der HANA Cluster Ressourcen . . . . .	286
7.3.3 Konfiguration des Pacemaker Clusters . . . . .	298
7.3.4 Einrichten des HANA HA-Clusters. . . . .	305
7.4 Bereitstellen eines S/4HANA HA Clusters . . . . .	309
7.4.1 SAP S/4HANA Clusterarchitektur . . . . .	309
7.4.2 Bereitstellen der SAP S/4HANA Cluster Ressourcen . . . . .	311
7.4.3 Hochverfügbarer NFS Speicher in Azure . . . . .	314
7.4.4 Installation ASCS und ERS . . . . .	321
7.4.5 Installation PAS und AAS . . . . .	324
7.5 Exkurs Automatisierte SAP Bereitstellung . . . . .	325
7.5.1 Ansible und Terraform in Azure . . . . .	325
7.5.2 SAP Cloud Appliance Library . . . . .	326
7.6 Zusammenfassung . . . . .	327
Literatur. . . . .	328
<b>8 SAP S/4 on Google Cloud – Konzepte und Architekturen . . . . .</b>	<b>329</b>
8.1 Überblick zu Google Cloud . . . . .	329
8.1.1 Historie von Google Cloud . . . . .	329
8.1.2 Zeitliche Entwicklung der Partnerschaft zwischen SAP und Google . . . . .	330
8.1.3 Entscheidungsgründe für SAP auf Google Cloud . . . . .	332
8.2 Google Cloud Organisationen und Ressourcen . . . . .	334
8.2.1 Google Cloud Ressourcenhierarchie . . . . .	334
8.2.2 Eigenschaften der Google Cloud Ressourcen . . . . .	336
8.3 Relevante Google Cloud Services für SAP S/4HANA . . . . .	336
8.3.1 Google Cloud Compute Engine . . . . .	337
8.3.2 Speicheroptionen . . . . .	340
8.3.3 Google Cloud VPC und Netzwerkkonzept . . . . .	345
8.3.4 Google Cloud Security . . . . .	351
8.3.5 Google Cloud Operations-Suite . . . . .	353

8.3.6	Google Cloud Customer-Care-Konzept . . . . .	355
8.3.7	Weitere relevante Google Cloud Services für SAP S/4HANA Bereitstellungen . . . . .	356
8.4	Google Cloud Frontend . . . . .	357
8.4.1	Google Cloud Frontend-Tools . . . . .	358
8.4.2	Typische Anwenderrollen einer SAP auf Google Cloud Landschaft . . . . .	359
8.5	SAP S/4HANA auf Google Cloud Architektur . . . . .	360
8.5.1	Lizenzen und Größenbestimmung . . . . .	360
8.5.2	SAP S/4HANA Architekturübersicht . . . . .	360
8.5.3	Setup für Hochverfügbarkeit . . . . .	361
8.5.4	Desaster Recovery Setup . . . . .	366
8.5.5	Erwägungen und Empfehlungen für die Speicherauswahl . . . . .	369
8.5.6	Exkurs: SAP S/4HANA Scale-Out (SAP Innovation Award 2021) . . . . .	376
8.5.7	Exkurs: SAP HANA Fast Restart und Memory Poisoning Recovery Mechanismus für SAP S/4HANA . . . . .	377
8.6	Zusammenfassung . . . . .	378
	Literatur . . . . .	379
<b>9</b>	<b>SAP S/4 on Google Cloud – Deployment . . . . .</b>	<b>383</b>
9.1	Beispielarchitektur für SAP S/4HANA auf Google Cloud . . . . .	383
9.2	Planungs- und Bereitstellungs-Checklisten für SAP auf Google Cloud . . . . .	384
9.3	Google Cloud Account, Netzwerk und Security Setup . . . . .	387
9.3.1	Setup des Google Accounts, Abrechnung und Identity und Access Management . . . . .	387
9.3.2	Setup der Shared VPC, Subnetze, Firewallregeln, Cloud NAT und Cloud DNS . . . . .	388
9.3.3	Setup von Cloud DNS . . . . .	389
9.4	Google Cloud Compute Setup . . . . .	392
9.4.1	Bereitstellung der SAP HANA Datenbank . . . . .	392
9.4.2	Bereitstellung der SAP NetWeaver Applikationsserver . . . . .	402
9.5	Preis- und Abrechnungskonzepte mit Best-Practices . . . . .	409
9.5.1	On-Demand vs. Rabatte für zugesicherte Nutzung . . . . .	410
9.5.2	Reservierungen in Google Cloud . . . . .	412
9.5.3	Google Cloud Kontingente und Budgets . . . . .	413
9.5.4	Preiskalkulation für das SAP S/4HANA auf Google Cloud Architekturbeispiel . . . . .	413
9.6	Sicherung & Wiederherstellung auf Google Cloud . . . . .	414
9.6.1	SAP NetWeaver Applikationsserver . . . . .	418
9.6.2	SAP HANA Datenbank . . . . .	419

<b>9.7</b>	<b>Scripting und Automatisierung auf Google Cloud. . . . .</b>	<b>421</b>
9.7.1	Tools für Skripterstellung und Automatisierung auf Google Cloud. . . . .	421
9.7.2	Anwendungsfall: Automatisierter Start und Stop von Instanzen . . . . .	422
9.7.3	Anwendungsfall: Autoskalierung für SAP NetWeaver Applikationsserver. . . . .	423
9.7.4	Anwendungsfall: Vereinfachung des operativen Betriebs . . . . .	423
9.7.5	SAP Landscape Manager und Google Cloud . . . . .	424
<b>9.8</b>	<b>Desaster Recovery mit Google Cloud . . . . .</b>	<b>424</b>
9.8.1	SAP HANA Desaster Recovery. . . . .	425
9.8.2	SAP NetWeaver Applikationsserver Desaster Recovery . . . . .	425
9.8.3	Empfehlungen für die Desaster Recovery Planung . . . . .	425
<b>9.9</b>	<b>Zusammenfassung . . . . .</b>	<b>426</b>
	Literatur. . . . .	426
<b>10</b>	<b>Zusammenfassung und Ausblick . . . . .</b>	<b>429</b>
10.1	Das Momentum von SAP S/4HANA. . . . .	429
10.2	Public Cloud als etablierter Trend . . . . .	430
10.3	Public Cloud als Innovationstreiber. . . . .	431
10.4	Ausblick. . . . .	432

---

## Über die Autoren



**Dr. André Bögelsack** arbeitet als Managing Director bei der Firma Accenture AG in der Schweiz und berät Kunden aller Industrien bei der Nutzung von Hyperscaler Services für den Betrieb von SAP-Systemen. Er unterstützt weltweit agierende Konzerne bei der Implementierung und dem Betrieb von neuen SAP S/4HANA-Landschaften. In seiner Laufbahn hat er mit seinem Team schon mehr als 1'000 SAP-Systeme erfolgreich migriert und kennt damit die typischen Herausforderungen bei diesen Projekten.

Vor seiner Zeit bei Accenture war André selber SAP-Basisadministrator und Unix-Administrator auf Sun Solaris 9 & 10 am SAP University Competence Center an der Technischen Universität München. Er wurde in Informatik über das Thema SAP promoviert und ist weithin in der SAP-Community durch seine Buchpublikationen und bei den Hyperscalern durch die Vielzahl der Projekte bekannt.

In seiner Freizeit läuft André Ultramarathons oder fährt Radmarathons. Wenn er nicht läuft oder fährt, verbringt er seine Zeit mit der Familie beim Wandern und dem Genießen der Schweiz – insbesondere des Schweizer Käses und Weins.



**Utpal Chakraborty** arbeitet als Manager bei der Accenture Solutions Pvt. Ltd in Indien in der Rolle eines SAP Technical Architect für Public Clouds. Er erstellt SAP-Architekturen und SAP-Designs für kosteneffiziente Lösungen zum Betrieb von Großkunden. In seiner Rolle als Delivery Manager, verantwortet er Projekte für Migrationen, Implementierungen und Upgrades von SAP HANA-Systemen. Aufgrund der Vielzahl der Projekte, besitzt er sehr viel Erfahrung bei Migrationen von SAP-Systemen (heterogene und homogene), beim Upgrade von S/4HANA-Systemen, als auch bei Greenfield Implementierungen.

Utpal besitzt einen Bachelor in Information Technology und begann seine Karriere in HCL Technologies in Indien als ein SAP-Basisadministrator. Bevor er seine Karriere bei Accenture startete, arbeitete Utpal bei IBM Private Limited in Indien als SAP-Basisadministrator.

In seiner Freizeit spielt Utpal Cricket, Tischtennis und liebt es, Serien im Internet zu schauen. Er liebt das Autofahren und nimmt seine Familie gerne auf ausgedehnte Fahrten mit.



**Dhiraj Kumar** arbeitet als Senior Manager bei Accenture Solutions Pvt. Ltd India. Er ist ein leitender technischer Architekt, der an der Planung, dem Design, dem Aufbau und der Ausführung der Implementierung und Migration von SAP-Ecosystemen in die Hyperscaler Cloud beteiligt ist. Er leitet und baut leistungsstarke Teams auf, die erfolgreich technische SAP-Projekte für verschiedene Kunden aus unterschiedlichen Branchen durchgeführt haben.

Bevor er zu Accenture kam, arbeitete Dhiraj 8 Jahre lang als SAP Basis Administrator bei IBM India Pvt. Ltd. und ein Jahr lang bei Capgemini Solutions India. Er arbeitet seit über 15 Jahren im Bereich SAP-Basis und SAP-Technologie. Dhiraj besitzt einen Bachelor in Elektronik und Kommunikationstechnik.

In seiner Freizeit kocht Dhiraj gerne und röhmt sich, darin gut zu sein. Er ist fasziniert von Geschichte und liebt es, über alte historische Orte zu lesen und sie zu erkunden.



**Elena Wolz** arbeitet als wissenschaftliche Mitarbeiterin beim SAP University Competence Center an der Technischen Universität München. Dabei verantwortet sie als SAP-Basis-administratorin unter anderem das Hosting von S/4HANA-Systemen für Bildungseinrichtungen. Im Rahmen des ECC Sunsets unterstützte sie ca. 150 Kunden bei der Migration von SAP ECC zu S/4HANA und sammelte dabei wertvolle Erfahrungen für den Betrieb von SAP-Systemen. Bei ihrer Promotion legt sie einen Forschungsschwerpunkt auf den Bereich Cloud Computing.

Davor studierte Elena Wirtschaftsinformatik an der Technischen Universität München, wo sie sich fundamentale Kenntnisse bei der Implementierung von SAP-Systemen in Hyperscalern aneignete. Bereits in ihrem Bachelorstudium legte sie durch die SAP TERP10-Beraterzertifizierung dafür den Grundstein. Zwischen dem Bachelor- und Masterstudium arbeitete sie als Sales Managerin bei einem SAP Consulting-Unternehmen.

In ihrer Freizeit geht Elena laufen oder bouldern. Daneben verbringt sie viel Zeit mit ihrer Schwester beim Gärtnern und Kochen, oder bei geselligen Abenden mit Freunden.



**Johannes Rank** ist seit 2017 als Basis Administrator am SAP UCC München tätig und inzwischen zertifizierter SAP Technology Associate für HANA (HANATEC) und Netweaver S/4HANA (TADM55). Seit 2020 leitet er die Basis des SAP UCC Münchens und beschäftigt sich dort unter anderem mit hybriden Cloud SAP-Bereitstellungen im Kontext von Microsoft Azure. Im Rahmen des ERP ECC Sunsets der University Alliance für September 2021, leitete er die technische Migration von mehr als 150 internationaler Kunden aus dem Raum APJ und EMEA. Dabei eignete er sich wertvolles Know-How im Bereich der System Conversion sowie Code- und Datenmigration an.

Vor seiner Zeit am UCC studierte Johannes Wirtschaftsinformatik an der Technischen Universität München. Dort ist er auch weiterhin als Dozent für SAP Softwareengineering auf Basis von ABAP und UI5 tätig.

Als baldiger Vater von zwei Kindern verbringt er die Freizeit meistens mit der Familie und „kindgerechten“ Hobbies wie schwimmen oder wandern. Im Winter nutzt er gerne die heimatlichen Gegebenheiten zum Skifahren oder Langlaufen.



**Jessica Tischbierek** übernimmt seit Herbst 2021 die Rolle als *SAP GTM Lead EMEA* bei Google Cloud mit Standort München. Sie hat zuvor seit 2018 bei Google Cloud im Pre-Sales Umfeld als *Customer Engineer Specialist for SAP on Google Cloud* Kunden bei ihrer Cloud Transformation beraten. Dabei arbeitet sie mit globalen Unternehmen und Partnern zusammen. Sie kennt als erfahrene Expertin die Vorteile, aber auch technischen und organisatorischen Herausforderungen von SAP Migrationen in die Public Cloud sowie die Integration in weitere Google Cloud Services wie Data Analytics, Machine Learning und Application Development.

Vor der Zeit bei Google war Jessica bei SAP und daraufhin bei DXC.Techology (vormals HPE) als SAP Technology Consultant tätig. Zu diesen Zeiten hat sie viele Projekte im Bereich der damals neuen Technologien wie SAP HANA 1.0, SAP Business Technology Platform (damals SAP HANA Cloud Platform) und SAP Fiori Entwicklung begleitet. Sie hat einen B.Sc. in Informatik und M.Sc. in Wirtschaftsinformatik.

In ihrer Freizeit lebt Jessica für den Sport und Abenteuerreisen. Neben Fitness- und Krafttraining ist sie beim Bergsteigen, auf Klettersteigen, Hochtouren, beim Skifahren oder Skitouren in den Alpen zu finden oder auf Reisen beim Entdecken, Wandern, Tauchen oder Surfen.



# Einleitung und Einführung zu Hyperscaler Clouds

1

## Zusammenfassung

Die Public Cloud existiert bereits seit mehr als zehn Jahren und bietet Unternehmen die Chance, auf unbegrenzte Kapazitäten und höchste Skalierung zurückzugreifen. Amazon Web Services, Microsoft Azure und Google Cloud sind die führenden Anbieter im Unternehmenskontext und werden durch kleinere Anbieter auf lokalen Märkten ergänzt. Die SAP verfolgt mit SAP RISE ebenfalls eine sehr strikte Cloud-Strategie und möchte das Portfolio von SAP-Services in der Cloud weiter ausbauen. Daneben wird SAP RISE als Angebot zur Transformation für Kunden positioniert, welches zu einem neuen, agilen Unternehmen basierend auf SAP S/4HANA und Cloud führen soll. Bestehende Kunden der SAP, welche aus dem Angebot der HANA Enterprise Cloud stammen, werden weiterhin durch die SAP bedient und verbleiben in der SAP HEC. Es ist derzeit noch unklar, für wie lange die SAP eine SAP HEC weiterhin anbieten wird.

## 1.1 Einleitung

Die Public Cloud existiert seit mehr als 10 Jahren und ist aus der aktuellen IT-Landschaft nicht mehr weg zu denken. Viele namenhafte Unternehmen setzen zur Verstärkung ihrer IT auf die Dienste aus der Public Cloud und so bilden die Hyperscaler mittlerweile das Rückgrat vieler Unternehmen aus den verschiedensten Industrien. Millionen von Kunden und Konsumenten der Unternehmen nutzen im alltäglichen Betrieb Webseiten oder Bestellportale, welche auf einer der großen Public Clouds betrieben werden.

SAP-Systeme stellen die neuralgischen Punkte im IT-Betrieb der Unternehmen dar. Produktionsstraßen bleiben stehen oder LKWs können die Lager nicht mehr verlassen,

wenn es zu einem Ausfall von SAP-Systemen kommt. Diese Wichtigkeit der SAP-Systeme hat kaum ein anderes IT-System in der heutigen Zeit. Dennoch bewegen immer mehr Unternehmen die SAP-Systeme in die Public Clouds. Dies wird aufgrund von Kostendruck, Druck zur Innovation oder zur simplen Erweiterung der Kapazitäten getan.

### Zielsetzung des Buchs

Mit dem aktuellen Druck bei allen SAP-Kunden, auf die neue Generation der SAP S/4HANA-Systeme zu springen, gewinnt auch das Thema der Public Clouds immer mehr Gewicht. Die Unternehmen nutzen dieses S/4-Momentum, um neben der Transformation zu SAP S/4HANA, auch die Migration in eine Public Cloud zu vollziehen.

Dieses Buch wird Ihnen mehrere Punkte der Public Clouds und der SAP S/4HANA-Systeme näherbringen. Hierbei werden neben den theoretischen Konstrukten, auch die praktischen Umsetzungen gezeigt. Das Buch befähigt zur direkten Umsetzung und der Provisionierung eines neuen SAP S/4HANA-Systems auf einer der drei großen Public Clouds: Azure, AWS und Google. Die Kapitel sind identisch aufgebaut und so können die notwendigen Schritte für den Aufbau und Betrieb von SAP S/4HANA-Systemen zwischen den Public Clouds miteinander verglichen werden.

In allen Kapiteln geben die Autorinnen und Autoren Beispiele aus der Praxis, um die Inhalte greifbar und anwendbar zu machen. Das Buch wird folgende Aspekte detailliert beschreiben:

- Die Entstehung von Public Cloud und der drei größten Hyperscaler Microsoft Azure, Amazon Web Services und Google Cloud
- Die verfügbaren Angebote von SAP zur Nutzung von Cloud-Diensten
- Die Architektur von SAP S/4HANA-Systemen
- Die wichtigsten Faktoren und Anwendungsfälle von SAP S/4HANA-Systemen auf einer Public Cloud
- Die Provisionierung und den Betrieb von SAP S/4HANA-Systemen auf den drei wichtigsten Public Clouds: Microsoft Azure, Amazon Web Services und Google Cloud

Nach dem Abschluss aller Kapitel werden Sie wissen, worauf bei einem Einsatz der Public Clouds zu achten ist und wie SAP S/4HANA-Systeme in den Public Clouds effizient betrieben werden können, ohne dass die Aufwände für den Betrieb oder aber die Kosten für die Public Clouds aus dem Rahmen laufen.

### Aufbau und Struktur

Das Buch leitet Sie durch den typischen Lebenszyklus eines SAP S/4HANA-Systems und beleuchtet hierbei die wichtigen Aspekte bei der Auswahl einer Public Cloud, der Provisionierung von SAP S/4HANA-Systemen und dem Betrieb auf den Public Clouds.

- Das Kap. 1 gibt einen Überblick zu den wichtigsten Begrifflichkeiten, wie der Private, Public und Hybrid Cloud. Es stellt auch die drei wichtigsten Public Clouds im Detail vor, umreißt aber auch den gegenwärtigen Markt der Cloudanbieter mit z. B. der IBM Cloud. Danach wirft das Kap. 1 ein Licht auf die Cloud-Strategie der SAP, dem SAP RISE Programm als auch der SAP HANA Enterprise Cloud, welche als Private Cloud im Markt präsent ist.
- Das Kap. 2 beschreibt die Architektur der SAP S/4HANA-Systeme, welche später in den nachfolgenden Kapiteln immer wieder aufgegriffen wird. Es enthält auch die wichtigen Aspekte für den Betrieb von SAP S/4HANA-Systemen auf der Public Cloud, wie das Sizing, das Backup, Hochverfügbarkeit und Automatisierung. Die nachfolgenden Kapitel zu den jeweiligen Public Clouds zeigen die Implementierungen dieser Punkte.
- Das Kap. 3 beschreibt die generellen Verfahren zum Deployment und der Migration von SAP S/4HANA-Systemen und erläutert beispielsweise Punkte, wie Greenfield und Brownfield Deployments, zeigt aber auch mögliche Migrationssszenarien für die Systeme auf. Es wird ebenfalls auf die wichtigsten Faktoren bei der Auswahl eines neuen Hyperscalers eingegangen.
- Das Kap. 4 ist der erste Teil des Buches, in dem der Einsatz von Amazon Web Services beschrieben wird. Hierzu erläutert das Kapitel zunächst die wichtigsten Begrifflichkeiten, welche AWS-spezifisch sind. Es wird auf die verfügbaren Maschinentypen eingegangen, aber auch die Grundkonzepte des Netzwerks und Storage erläutert.
- Das Kap. 5 ist der praktische Teil zu AWS und zeigt, wie ein SAP S/4HANA-System auf der AWS erstellt und gesteuert wird. Hierzu zeigt das Kapitel das zu implementierende System und beschreibt danach, mit welchen Schritten, das System auf AWS aufgebaut werden kann. Die wichtigen typischen Use Cases, wie z. B. die Wiederherstellung nach einem K-Fall, werden sukzessive beschrieben.
- Das Kap. 6 ist das erste Kapitel zu Microsoft Azure und beschreibt, wie Azure als Public Cloud-Lösung eingesetzt werden kann und welche wichtigen Grundkonzepte für den Betrieb und die Provisionierung von SAP S/4HANA-Systemen existieren.
- Das Kap. 7 setzt auf den zuvor im Kap. 6 geschaffenen Grundlagen zu Microsoft Azure auf und verdeutlicht die Provisionierung eines neuen SAP S/4HANA-Systems und zeigt die Implementierung der wichtigen Use Case aus dem Kap. 2.
- Das Kap. 8 zeigt die wichtigen Konzepte und Architekturen der Google Cloud auf und beschreibt, wie SAP S/4HANA-Systeme in der Google Cloud geplant werden.
- Das Kap. 9 zeigt anhand der vorher beschriebenen Konzepte der Google Cloud die konkrete Implementierung eines SAP S/4HANA-Systems in der Google Cloud.
- Im abschließenden Kap. 10 werden alle wichtigen Punkte aus den vorherigen Kapiteln zusammengefasst und ein Ausblick zu den Entwicklungen in den kommenden Jahren gegeben.

Nach Abschluss aller Kapitel haben Sie, lieber Leser, einen umfassenden Überblick zu allen Aspekten der Thematik «SAP S/4HANA-Systeme auf Hyperscaler Clouds» und wissen, wie Sie diese Systeme auf den Public Clouds planen, konzeptionieren, umsetzen und steuern.

---

## 1.2 Cloud Computing allgemein

Vor dem Einstieg in die Welt der SAP S/4HANA-Systeme auf der Public Cloud, werden in diesem Kapitel zunächst die Grundzüge des Cloud Computing beschrieben und die wichtigsten Begriffe eingeführt. Dies erleichtert das Verständnis des restlichen Buches.

### 1.2.1 Wichtigste Eigenschaften der Cloud

Cloud Computing ist in der Mitte der 2000'er Jahr entstanden und wurde auf Basis einer Überkapazität in Rechenzentren entwickelt. Ursprünglich ging der Ansatz darauf zurück, überschüssige Rechenkapazitäten an andere Kunden zu vergeben. Das Ziel war es, eine durchgängig hohe Auslastung über das gesamte Jahr hinweg zu erzielen.

Cloud Computing ist auf der einen Seite um eine neue Art und Weise, wie Services an die Kunden und Unternehmen herangetragen werden. So können Cloud Services und Leistungen aus der Cloud **ohne größere vorherige Vertragsverhandlungen** genutzt werden. Lediglich die Allgemeinen Geschäftsbedingungen (AGBs) sind zu akzeptieren und eine valide Zahlungsmöglichkeit zu hinterlegen und dann können die Services schon genutzt werden.

Auf der anderen Seite ist Cloud Computing ein wichtiger Faktor bei **Innovationen** und der Einführung komplett neuer Services. Von der ursprünglichen Idee überschüssige Kapazitäten zur Verfügung zu stellen, ist mittlerweile nicht mehr sehr viel übriggeblieben. Provider von Clouds stellen unentwegt neue Services zur Verfügung und regelmäßig auch neue Versionen der Services. Hiervon profitieren die Kunden, da somit neue Services nicht explizit angefragt werden müssen, sondern automatisch mit der Nutzung der Cloud mitkommen.

Neben der Einfachheit der Nutzung der Services und der sehr hohen Innovationskraft, steht ein weiterer Punkt im Fokus. Die Cloud kennt prinzipiell keine Grenzen und so können Kunden und Unternehmen von **grenzenlosen Kapazitäten** in einer Cloud ausgehen. Das unterscheidet die Clouds von den herkömmlichen Providern, da hier die Ressourcen üblicherweise beschränkt sind.

Die hohe Skalierung der Clouds, das heißt die große Anzahl der Kunden, lässt die **Economy of Scales** vollständig zutreffen. Hierbei profitieren die Kunden von der Bereitstellung der gleichen Services für andere Kunden und einer Erzielung eines sehr geringen Preises.

Werden die wichtigsten Features einer Cloud zusammengefasst, so sind die folgenden Eigenschaften, die Kernpunkte der Clouds:

- Einfachheit der Nutzung und des Zugriffs
- Hohe Innovationskraft durch neue Services
- Sehr große Kapazitäten/unendliche Kapazitäten
- Sehr hohe Preisattraktivität

Im Kontext der Hyperscaler, wie Amazon, Microsoft und Google, treffen obige Kriterien zu und Unternehmen können die Vorteile nutzen. Doch die vielen Vorteile sind auch gepaart mit einigen Nachteilen bei der Nutzung der Clouds. Hierbei sind es weniger technische, sondern eher organisatorische und prozessuale Punkte, welche als Nachteil ausgelegt werden können.

Obwohl die Nutzung der Clouds sehr einfach ist, so muss die Nutzung auch stetig und ständig gesteuert werden. Sehr oft existiert eine „Schatten-IT“ neben der eigentlichen IT, bei der sich Fachabteilungen selber IT-Services aus der Cloud bereitgestellt haben. Dies widerspricht einer zentralen IT und einer effizienten Steuerung der gesamten IT.

Obwohl die Preise der Cloud sehr gut sind, kann die Nutzung von Cloud in einem ersten Schritt teuer werden, als wenn ein Unternehmen weiterhin ohne Cloud gearbeitet hätte. Die Nutzung der Cloud muss also in einem gewissen Rahmen mit einer **Transformation** der IT einhergehen. Ein einfaches Übertragen der IT-Services in die Cloud ohne Transformation bringt meist nicht die erhofften Effekte.

Die unlimitierten Ressourcen in einer Cloud waren sehr lange Realität. Nie wurde von Engpässen berichtet oder von einer Nichtverfügbarkeit von Ressourcen. Seit der Corona-Krise jedoch ist deutlich geworden, dass Clouds auch „nur“ Rechenzentren mit herkömmlichen Kapazitäten sind. Spätestens seit die Microsoft Azure Cloud im März 2020 **keine freien Kapazitäten** mehr für bestimmte SAP-Workloads hatte, ist der Mythos der unendlichen Ressourcen widerlegt.

Unternehmen müssen bei der Nutzung der Clouds sehr viel selbst machen. Dies ist ein großer Unterschied zu der Nutzung von Services bei einem Service Provider. Hierbei erledigt der Provider sehr viele Aufgaben und übernimmt die Verantwortung für die Aufgaben und Resultate. Ein Cloud Provider wird dies nur für die Services tun, welche in seiner Verantwortung liegen. Im Gegensatz zu einem Service Provider sind dies jedoch viel weniger. Somit müssen die **Kunden Aufgaben übernehmen** oder aber einen anderen Provider damit beauftragen.

Bei einem Betrieb von SAP-Systemen auf einer Cloud, muss der Unterschied zu den Service Providern beachtet werden. In der Tab. 1.1 ist dieser verdeutlicht.

Alle dunklen Felder in der Tabelle zeigen die Verantwortlichkeiten der jeweiligen Parteien. Es wird deutlich, dass die typischen Service Provider den kompletten Stack abbilden können und in einem traditionellen Hosting alle Services erbringen können. Bei der Nutzung der Cloud, müssen die Kunden selbst Services erbringen, da der Cloud Provider dies nicht tut.

**Tab. 1.1** Unterschied in Verantwortlichkeiten

Komponente	Verantwortung Kunde	Verantwortung Cloud Provider	Verantwortung Service Provider
Rechenzentrum			
Compute / Storage / Netzwerk Ressourcen			
Sicherheit und Compliance			
Betrieb der Infrastruktur			
Design des Technologiestack			
Betrieb der Betriebssysteme und Komponenten			
Betrieb von SAP (z. B. Basis)			
Betrieb von non-SAP Applikationen			
<b>Mehr Verantwortung für Kunden</b>			

## 1.2.2 Public, Private und Hybrid Clouds

Der Markt von Cloud Computing ist sehr heterogen und von vielen Trends durchzogen. Für SAP-Workloads und einem Betrieb in einer Cloud kommen drei wichtigste Clouds zum Einsatz: die Public Cloud, die Private Cloud und die Hybrid Cloud.

Die **Public Cloud** ist eine Cloud, welche den meisten Unternehmen bekannt ist. Es handelt sich um die Art von Cloud, welche von allen Hyperscalern (also Amazon, Microsoft und Google) angeboten wird. Dabei impliziert die Namensgebung Public auch schon das wichtigste Merkmal der Cloud – sie ist öffentlich und für jeden Kunden gleichermaßen zugänglich. Das bedeutet jedoch nicht, dass alle Unternehmen alle Workloads und alle Daten aller anderer Unternehmen sehen können. Es bedeutet lediglich, dass die Cloud für alle Unternehmen gleichermaßen zur Verfügung steht und die Cloud nicht dediziert für nur ein Unternehmen aufgebaut worden ist. Die Anbieter der Public Cloud haben hierfür Mechanismen implementiert, welche den Zugang zu den Ressourcen eines Unternehmens auch nur für das eigentliche Unternehmen möglich machen.

Im Gegensatz zu einer Public Cloud ist die **Private Cloud** exklusiv für ein Unternehmen aufgebaut. Alle Ressourcen dieser Private Cloud stehen nur einem Unternehmen zur Verfügung und können auch nur durch das Unternehmen genutzt werden. Die Private Cloud hat jedoch einen entscheidenden Nachteil. Die aufgebauten Ressourcen der

Private Cloud können bis zur Erreichung der Kapazitäten genutzt werden und müssen danach erweitert werden. Die Betreiber von Private Clouds setzen solche Erweiterungen ähnlich wie die normalen Service Provider um: Hardwareressourcen werden exklusiv für den Kunden angeschafft und somit auch fakturiert. Ein weiterer Nachteil der Private Clouds besteht in der starken Fokussierung auf den Einsatzzweck. Die Anbieter der Private Clouds reichern die Clouds nicht mit neuen Services an, sondern belassen sie so, wie sie sind.

Neben der Public Cloud und der Private Cloud, existiert noch die **Hybrid Cloud**. Es handelt sich hierbei weniger um eine Cloud, sondern vielmehr um einen Zusammenschluss von einer Public Cloud mit einer Private Cloud oder mit einem bestehenden Rechenzentrum eines Unternehmens. Das zeigt auch schon, dass die Mehrzahl der Unternehmen sich direkt in einem Hybrid Cloud-Szenario befinden, sobald die ersten Services aus der Public Cloud konsumiert werden.

---

### Vom Rechenzentrum zur Hybrid Cloud

Viele Unternehmen betreiben SAP-Systeme noch in eigenen Rechenzentren. Diese Rechenzentren gelangen oftmals an Kapazitätsgrenzen oder müssen schlichtweg abgerissen werden, da sie zu alt geworden sind. Dies passierte auch bei einem großen deutschen, produzierenden Unternehmen. Die Rechenzentren auf demselben Campus sollten abgerissen werden und es sollte Platz für neue Bürogebäude entstehen.

Das Unternehmen hatte verschiedene Möglichkeiten offen. Es konnte einen externen Service Provider mit Rechenzentrumskapazitäten nutzen, konnte temporär Rechenzentrumskapazität kaufen und selber betreiben oder aber konnte die Public Cloud nutzen. Das Unternehmen hat sich strategisch für die Public Cloud entschieden. Aufgrund von vielen Projekten und Projektaktivitäten wurden verschieden viele SAP-Systeme benötigt. Als das Unternehmen begann, diese Systeme in der Public Cloud zu provisionieren, begab es sich in das Szenario der Hybrid Cloud. Es existieren noch viele Systeme in den beiden abzureißenden Rechenzentren, aber die neuen Systeme befinden sich in der Public Cloud. Dieses Szenario ist ein hybrides.

Das obige Beispiel zeigt, dass durch die simple Nutzung der ersten Services aus der Public Cloud, ein Unternehmen in das Szenario der Hybrid Cloud geht. So ergeht es allen Unternehmen, welche mit dem ersten Service in die Cloud gehen. ◀

Bei Unternehmen, welche zwei Public Clouds, also beispielsweise Google und Amazon, miteinander verbinden, kann auch eine Hybrid Cloud entstehen. So entscheiden sich die Unternehmen teilweise, SAP-Workloads in eine Public Cloud zu verschieben und die nicht-SAP-Workloads in eine andere Public Cloud zu verschieben. Auf diesem Wege werden zwei Public Clouds miteinander kombiniert und es entsteht eine große Hybrid Cloud.

## 1.3 Die Clouds der wichtigsten Marktteilnehmer im Überblick

Derzeit wird der Markt der Public Clouds durch drei Hyperscaler dominiert. Dies sind Amazon Web Services, Google Cloud und Microsoft Azure. Alle drei Hyperscaler werden in diesem Kapitel vorgestellt. Darüber hinaus werden auch die kleineren Anbieter von Clouds, wie IBM oder Oracle, beschrieben.

### 1.3.1 Amazon Web Services

#### Historie

Amazon Web Services war der erste wirkliche Cloud Anbieter auf dem Markt und ist bis heute einer der führenden Anbieter von Cloud Computing Services. Es gibt mehrere Versionen zu den Gründen, wieso Amazon Web Services gestartet ist.

Eine Version basiert auf der Situation einer Überkapazität. In den frühen 2000er Jahren entwickelte sich das Geschäft von Amazon stetig weiter und durch die starken Verkaufsaktivitäten auf der Plattform von Amazon, wurden die Bedarfe nach Rechenleistung für die Bereitstellung der Plattform und deren Dienste enorm. Insbesondere durch die saisonalen Effekte, wie zum Beispiel der Black Friday in den USA, stieg der Bedarf dramatisch an und fiel danach wieder rapide ab auf ein Normalniveau. Amazon reagierte darauf und baute die Plattform entsprechend so auf, dass die Lastspitzen durch die saisonalen Ereignisse abgedeckt wurden und die Plattform performant und stabil lief. Diese Ausrichtung auf die Maximalleistung führte zu einer stetigen Überkapazität in den Amazon Rechenzentren während des Rests des Jahres – also außerhalb der saisonalen Effekte. Diese Überkapazitäten wurden als Ressourcen für Kunden zugänglich gemacht. Im Jahre 2006 startete Amazon dann die Tochterfirma Amazon Web Services, welche über ein offen zugängliches Portal die Kapazitäten aus dem Rechenzentrum von Amazon extern verfügbar macht.

Eine andere Version zur Entstehung des Cloud Computing bei AWS ist basierend auf den internen Prozessen und Abläufen innerhalb von Amazon. Hierbei sah sich Amazon durch das stetige Wachstum seiner Plattform mit wiederkehrenden Problemen konfrontiert, bei denen die Web-Entwickler mit den Teams aus den Rechenzentren immer wieder Abstimmungen zu Netzwerk, Kapazitäten und Verfügbarkeiten vornehmen mussten. Um diese Abhängigkeit zu beseitigen, wurde eine Art von Commodity IT eingeführt (AWS), mit der die Web-Entwickler sich die notwendigen Ressourcen selber zusammenstellen konnten. Sie wurden also unabhängig von den Teams der Rechenzentren. Amazon realisierte, dass jeder Web-Entwickler solche Möglichkeiten schätzen würde und entschied sich, diese neuen Services als Amazon Web Services zur Verfügung zu stellen.

Unabhängig von der Version zur Entstehung, war der Start von Amazon Web Services ein voller Erfolg. Innerhalb von wenigen Jahren schaffte es AWS, eine große Kundschaft zu schaffen und somit eine hohe Nutzung der Plattform zu erzielen – auch, wenn es

anfangs noch relativ kompliziert war, eigene SAP-Systeme auf AWS in den Betrieb zu bekommen.

### Aktuelle Marktposition

AWS ist der Marktführer und bedient eine sehr große Anzahl von Kunden aus dem Bereich der weltweit agierenden Konzerne, aber auch aus dem Bereich der kleinen und mittelständischen Unternehmen sowie Privatanwender. Diese breite Fülle von Kunden bedient AWS durch eine sehr breite Produktpalette.

AWS zeichnet sich bei den Kunden durch die folgenden Punkte aus:

- Hohe Dichte der AWS Rechenzentren in vielen unterschiedlichen Ländern mit Verfügbarkeitszonen
- Stetige Updates der Cloud Services mit regelmäßig neuen Features
- Hohe Verfügbarkeit der Services ohne größere Auswirkungen auf eine Vielzahl von Kunden bisher
- Sehr großes Ecosystem mit vielen Drittanbietern, welche zusätzliche Services anbieten
- Einfachheit der Einführung

Die Kundschaft von AWS sieht aber auch einige negative Punkte:

- Alle Services sind entgeltlich und teilweise sind die Preise für die Cloud Services sehr hoch. Dies kann zu sehr hohen Rechnungen für die Nutzung von AWS führen.
- Strikte Verträge zwischen den Kunden und AWS ohne Flexibilität zur Individualisierung für Unternehmen
- Komplexität bei der Integration von Services innerhalb von AWS und mit Drittanbietern
- Sehr steile Lernkurve notwendig und der AWS Support kann nicht immer unterstützen – insbesondere bei Software von Drittanbietern (wie z. B. SAP!)

Generell kann AWS den ersten Platz bei den Cloud Anbietern weiter verteidigen und durch die hohe Innovationskraft und die sehr starke Kundendecke wird sich dies auf absehbare Zeit auch nicht ändern. Potenzielle Kunden sollten jedoch ein Auge auf die Kosten und die vertraglichen Regelungen haben.

### Kollaboration mit SAP

AWS begann schon früh das Potenzial von SAP auf AWS zu realisieren, denn SAP-Systeme gehören zu den wichtigsten IT-Systemen der Unternehmen. Somit werden die Systeme immer verwendet und müssen eine hohe Verfügbarkeit erzielen. Schon seit 2011 arbeiten SAP und AWS zusammen, um Kunden die Möglichkeit zu eröffnen, SAP-Systeme auf der Public Cloud zu hosten.

AWS konzentrierte sich schnell darauf, den Kunden von HANA-basierten Systemen, eine neue Plattform bereitzustellen. In dem Jahr 2016 wurden die ersten großen SAP HANA-zertifizierten Systeme zur Verfügung gestellt (z. B. r1.32xlarge). Später wurden in den Folgejahren auch immer wieder neue Workloads (Templates) für HANA-basierte Systeme bereitgestellt. AWS gab hiermit den Takt vor und andere Cloud Anbieter folgten oftmals.

AWS erweiterte das Portfolio der unterstützten SAP-Systeme und versucht nun auch die etwas älteren SAP-Systeme mit zu unterstützen. Hierbei sieht die Unterstützung von z. B. alten R/3 Systemen noch sehr eingeschränkt aus. Faktisch müssen sich alle Unternehmen aber noch immer mit R/3 oder sogar R/2 Systemen beschäftigen, da diese aus regulatorischen Gründen aufbewahrt werden müssen.

### 1.3.2 Microsoft Azure

#### Historie

Microsoft erkannte den enormen Nutzen und das enorme Potenzial von Cloud Computing sehr früh und startete eine erste Version von Microsoft Azure bereits im Jahre 2008 – also nur zwei Jahre später als AWS. Ab dem Jahre 2010 gilt die Cloud Computing Plattform von Microsoft als verfügbar, da sie vorher eingeschränkt nur für Entwickler verfügbar war. Microsoft änderte den Namen der Plattform einige Male: von der Einführung der Azure Plattform als „Windows Cloud“ über „Windows Azure“ hin zu „Microsoft Azure“ in 2014. Seit 2014 ist der Name konstant und hat sich als eine wichtige Marke in dem Cloud Computing-Markt etabliert.

Microsoft baute Azure stetig weiter aus und erweiterte das Portfolio der angebotenen Services in Azure. Parallel zur Erweiterung des Portfolios, baute Microsoft neue Regionen aus. So startete Azure zunächst in den USA und Europa und baute jedoch sehr zügig sein Angebot aus. Derzeit umfasst Azure 54 Regionen auf allen Kontinenten. Nicht jede Region bietet dieselben Services an, jedoch können die wichtigsten Infrastrukturservices in allen Regionen abgerufen werden. Auf der Webseite von Microsoft können die verfügbaren Regionen eingesehen werden (Zugriff am 20.12.2021):

<https://azure.microsoft.com/en-us/global-infrastructure/geographies/#geographies>

Microsoft orientiert sich beim Ausbau der Azure Regionen sehr stark an den wichtigsten Märkten und Ländern, welche spezifische Vorgaben und Regulatorien haben. So wurde beispielsweise in Asien die Regionen in Indien aufgebaut, um den stark wachsenden Markt zu adressieren und die rechtlichen Vorschriften nach der Datenhaltung gerecht zu werden. Mit demselben Gedanken hat Microsoft die Regionen in der Schweiz und Deutschland aufgebaut, um den Industrienormen der Schweizer Bankenindustrie und den EU-Richtlinien nach DSVGO gerecht zu werden. Die Regionen in den USA sind die am weitesten entwickelten Regionen und hier werden auch neue Services zuerst zur Verfügung gestellt.

### Aktuelle Marktposition

Microsoft Azure gilt in den aktuellen Markstudien (wie zum Beispiel von Gartner) als einer der führenden Anbieter von Cloud Computing, belegt aber meist den zweiten Platz hinter Amazon Web Services. Aus Sicht von Kunden geben beide Anbieter in den einzelnen Kategorien die Geschwindigkeit vor. So konnte Microsoft damals als erster Anbieter von Cloud Computing sehr große SAP HANA-Maschinen anbieten, während AWS dies erst kurze Zeit später konnte. Beide Anbieter wechseln sich aus Sicht der Kunden immer wieder ab.

Faktisch besitzt Microsoft mit der großen Marktdominanz im Office-Bereich aber einen entscheidenden Vorteil. Durch die weite Verbreitung von Office-Produkten hat Microsoft immer einen Startpunkt zur Verhandlung von zusätzlichen Cloud-Produkten. Darüber hinaus vollziehen Unternehmen derzeit einen starken Wandel von den traditionellen Office-Produkten hin zu Office365-Produkten, welche in der Cloud angeboten werden. Damit kann Microsoft den Anteil an Cloud Computing erweitern. Darüber hinaus wird damit die Hürde für die Nutzung weiterer Azure-Service sehr gering gehalten.

Microsoft Azure zeichnet sich bei Kunden insbesondere durch folgende Punkte aus:

- Einfache Integration von existierender Infrastruktur in die MS Azure Services (wie zum Beispiel Active Directory oder SharePoint)
- Hohe Kompatibilität mit Linux
- Sehr große Ähnlichkeit mit anderen Microsoft-Produkten, wodurch ein Umstieg erleichtert wird
- Leichtigkeit der Nutzung von neuen Services

Es gibt bei Microsoft Azure auch Punkte, welche Kunden als kritisch erachten:

- Hohe Geschwindigkeit der Neuerungen lässt die Interfaces und Dokumentationen mit den existierenden Services auseinanderlaufen
- Support für Azure vonseiten Microsofts ist verbesserungswürdig, insbesondere für spezifische Fragen und Problemstellungen
- Komplexe Bepreisung der Services und teils hohe, unerwartete, Kosten bei der Nutzung

Microsoft kann und wird den zweiten Platz verteidigen und hat die Ambitionen, AWS den ersten Platz strittig zu machen. Durch die hohe Verbreitung von Microsoft-Produkten in den Unternehmen, hat Microsoft eine hervorragende Position zur weiteren Durchdringung des Marktes.

### Kollaboration mit SAP

Microsoft und SAP arbeiten schon seit längerer Zeit zusammen und haben eine technologische Partnerschaft seit 25 Jahren. Mit der Einführung von SAP R/3 auf der Microsoft

Windows Server-Plattform startete Microsoft in die Welt des Hostings für ERP-Systeme. Bereits seit 1989 haben Microsoft und SAP die Zusammenarbeit gestartet und 1993 forcierter. Microsoft, welche mit Navision selber ein ERP-System anbieten, nutzen ebenfalls SAP als ERP-System mit über 100.000 Benutzern.

SAP und Microsoft entwickeln die Microsoft-Plattformen in Redmond und Walldorf gemeinsam weiter. Die Azure-Plattform ist seit mehreren Jahren für SAP freigegeben, wobei es auch einige Einschränkungen bei der Unterstützung von älteren Versionen von SAP gibt. Dies trifft aber auf alle Hyperscaler zu.

Microsoft und SAP haben im Jahr 2021 eine intensivierte Partnerschaft angekündigt. Hierbei soll es um die Kollaboration auf drei wichtigen Sektoren gehen: die Kollaboration zwischen den Menschen und Mitarbeiterinnern/Mitarbeitern, die Kollaboration bei der Migration zu SAP S/4HANA auf der Azure-Plattform und der Kollaboration auf der Ebene der Technologie.

- **Kollaboration zwischen den Mitarbeiterinnen und Mitarbeitern:** Hierbei planen SAP und Microsoft die sehr enge und hohe Integration von Microsoft Teams (als die zentrale Kollaborationsplattform) und den SAP Services. Vor dem Hintergrund der COVID19-Pandemie und dem Bedürfnis nach einer Änderung der Art und Weise des täglichen Arbeitens, entschieden sich beide Partner die Integration voranzutreiben. Dies soll nicht nur für die Mitarbeiterinnern/Mitarbeiter desselben Unternehmens gelten, sondern auch für die Geschäftspartner eines Unternehmens.
- **Kollaboration bei der Migration zu SAP S/4HANA auf Azure:** Beide Partner wollen die Migration von der traditionellen Infrastruktur und den nicht-HANA-Systemen zu einem neuen S/4HANA-System auf der Azure-Plattform vereinfachen. Dies soll durch einfachere Migrationsmethodiken und hoher Automatisierung erfolgen. Darüber hinaus sollen die Themen der Automatisierung, Monitoring, Provisionierung und Security weiter forciert werden.
- **Kollaboration auf der Technologieebene:** SAP und Microsoft haben jeweils große Erfahrungen im Betrieb von Applikationslandschaften und beide Partner wollen enger zusammenarbeiten, um die Themen der Integration, Verfügbarkeit, Konnektivität und Sicherheit voranzubringen. Das Ziel ist es, die SAP Business Technology Plattform (siehe hierzu das spätere Kapitel zu SAP RISE) technisch aufzuwerten.

Inwieweit die avisierte Kollaboration zwischen Microsoft und SAP für die Kunden entscheidend wird, kann schwierig abgeschätzt werden. Jedoch treten hier zwei Technologieriesen auf dem Markt auf und es ist davon auszugehen, dass sie einen positiven Einfluss auf den Markt und Kunden haben werden.

### 1.3.3 Google Cloud

#### Historie

Die Google Cloud Platform (GCP) ist Teil der größeren Google Cloud, welche Google zur Erbringung der eigenen Dienste nutzt (wie z. B. YouTube oder Google Maps) und für seine Kunden zur Verfügung stellt. Im Jahr 2008 wurde die Google Cloud Platform angekündigt, welche eine Nutzung von Cloud Computing, wie im Sinne von AWS und Azure, ermöglichen sollte. Damals wurde zunächst die „App Engine“ vorgestellt, welche primär Entwickler von Webapplikationen ansprach.

Einige Jahre später hat Google etliche neue Services hinzugefügt und so auch im Jahre 2013 die Compute Engine für die breite Öffentlichkeit zur Verfügung gestellt. Dies war der Startpunkt für die Infrastructure-as-a-Service Services in der Google Cloud, auch wenn der Cloud Storage schon im Jahre 2010 verfügbar gemacht worden ist.

Im Vergleich zu AWS und Azure gilt Google in den Bereichen von Big Data und den neueren Technologien wie Machine Learning und Artificial Intelligence als der Vorreiter. Hier kommen die Ursprünge des Unternehmens voll zum Tragen, welche nicht wie bei AWS und Azure in den Bereichen der Infrastruktur-as-a-Service, sondern vielmehr bei der Entwicklung von Cloud-nativen und hochkomplexen Applikationen zu finden ist.

Manche Kunden und Unternehmen waren vor der Einführung der Google Produkte noch unsicher, ob es sich um eine verlässliche Plattform handeln würde. Hintergrund hierzu sind oftmals Erfahrungen aus dem Privatkundenbereich, in dem sich Kunden mit abgekündigten, nicht mehr unterstützten Produkten beschäftigen müssten. Dies sollte im Umfeld von Unternehmen nicht passieren und vermieden werden. Google hat durch die personelle Stärkung des Cloudbereichs hier jedoch für eine starke Kontinuität gesorgt und den Kunden die Angst vor einer Abkündigung der Google Cloud genommen.

#### Aktuelle Marktposition

Google ist als einer der führenden Cloudanbieter weltweit etabliert, kann jedoch in den meisten Reports der Marktanalysten nur den dritten Platz hinter AWS und Azure erzielen. Die Analysten gehen von einem Marktanteil im niedrigen zweistelligen Prozentbereich aus, während der Platzhirsch AWS gut ein Drittel des Marktes beherrscht. Dennoch hat es Google geschafft, einige Leuchtturmprojekte zu gewinnen. Diese befinden sich insbesondere in den Industrien von Retail, Financial Services und Gesundheitswesen/Pharma.

Trotzdem Google keinen riesigen Marktanteil im Vergleich zu AWS und Azure besitzt, kann die Google Cloud durch einen signifikanten Fußabdruck der Rechenzentren und der Regionen weltweit punkten. Aktuell besitzt die Google Cloud 24 Regionen mit mehr als 77 Rechenzentren und ist auf allen Kontinenten der Welt vertreten. Google schafft es auch, spezifische Märkte, wie die Schweiz, Deutschland oder Indien, abzudecken.

Durch die Platzierung von der G Suite, dem Google Workplace Produkt, hat Google bereits eine breite Kundenbasis aufbauen können. Neben der Akquirierung von neuen Kunden, ist der Sprung der Kunden von einem bestehenden Vertrag mit Google Workplace, auf die volle Nutzung der Google Cloud Platform, das Ziel. Microsoft verfolgt eine ähnliche Strategie und versucht die Produkte zu bündeln.

Kunden stellen bei Google gerne folgende wichtige Vorteile heraus:

- Eine sehr hohe Integrität der Cloud Services untereinander (höher bewertet als bei AWS und Azure)
- Die Einfachheit des Google Command Line Interface, welches die Erstellung und das Management der Cloud Services stark vereinfacht und eine hohe Automatisierung zulässt.
- Die sehr fortschrittlichen und teils kosteneffizienten Services

Neben den vielen Vorteilen gibt es aber auch Punkte, welche aus Sicht der Kunden verbessierungswürdig sind:

- Das eigentliche Portal ist, im Vergleich zu Azure und AWS, noch nicht so weit entwickelt und teils noch sehr kompliziert.
- Die Gefahr der schnell stark steigenden Kosten

Google hat als dritter Hyperscaler noch viel Raum zu wachsen. Der Markt für Cloud Computing ist noch lange nicht komplett bedient. Google hat gute Chancen, seine Marktanteile weiter zu erhöhen und die Cloud Dienste weiter bei den Kunden zu platzieren. In einigen innovativen Bereichen, welche für ein Hosting von SAP S/4HANA Systemen weniger relevant sind, ist Google aber Marktführer und wird es auf absehbare Zeit auch bleiben.

### **Kollaboration mit SAP**

Im Jahre 2017 hat Google offiziell eine Partnerschaft mit SAP auf der Next'17 angekündigt und auf der später stattfindenden SAPPHIRE einen ersten Fortschritt präsentiert. Als erste Schritte in der Partnerschaft fokussierte Google die Zertifizierung der Google Cloud Platform für die SAP-Workloads. Zunächst fokussierte Google die neusten Produkte der SAP – derzeit die HANA-basierten Workloads – und ließ danach die Google Cloud Platform auch für die NetWeaver-Produkte zertifizieren. Schon im Jahr 2017 begann Google neben den einfachen Angeboten der Infrastructure-as-a-Service auch Big Data Services im Zusammenhang mit SAP anzubieten. Google wollte nicht nur als weiterer Player im Segment der virtuellen Hardware erscheinen, sondern Mehrwert für die Kunden bieten.

Über die nächsten Jahre wurde das Portfolio von Google für SAP-Systeme stetig erweitert. Im Jahr 2018 wurden neue und größere Maschinentypen angekündigt und die Integration von SAP Smart Data Access zu BigQuery von Google. Durch die Akquisition

von Stackdriver im Jahr 2014 konnte Google nun auch diese Features nutzen, um das Monitoring von SAP-Systemen auf GCP zu realisieren und eine erste Automatisierung umzusetzen. SAP setzte im Gegenzug ebenfalls auf die Google Cloud Platform und unterstützte diese aus der SAP Cloud Appliance Library heraus.

Nach und nach konnte Google die Kundenbasis erweitern und war 2019 in der Lage, erste große Erfolge mit den Kunden zusammen zu feiern. Hierzu gehört sicherlich das Kundenbeispiel von Metro, einem sehr großen Retailer. Ab diesem Jahr wurden auch die Erweiterungen für die SAP C/4HANA-Systeme in den Fokus gerückt und die Kunden bei der Entwicklung der Erweiterungen geholfen. Google schaffte auch die Integration der G Suite in die SAP-Produkte hinein und ermöglichte damit einen barrierefreien Austausch von Daten zwischen den beiden Welten.

Google und SAP haben auch bei RISE ihre Partnerschaft weiter vertieft. Kurz nach der Ankündigung durch die SAP konnte Google schon einige wichtige Kunden für sich gewinnen und SAP bot RISE auf Basis der Google Cloud an.

Nach einigen Jahren der Kollaboration und Zusammenarbeit zwischen SAP und Google ist deutlich, dass Google das Geschäft mit den Großunternehmen wichtig ist und hier mit Microsoft und Amazon Web Services auf der Ebene der Infrastructure-as-a-Service ebenbürtig ist. Schaut man hingegen auf die Themen Big Data, Analytics und Machine Learning bietet Google hier die führenden Services im Cloud Computing für SAP an.

### 1.3.4 Weitere Marktteilnehmer

Neben den drei großen Hyperscalern existieren noch weitere Anbieter von Cloud-Services, welche hier auch Erwähnung finden sollen. Diese Anbieter setzen entweder auf Private Clouds oder aber stellen Public Cloud-Services in spezifischen Ländern zur Verfügung.

#### Alibaba Cloud

Alibaba ist ein Internetkonzern aus China, welcher mit Alibaba Cloud im Jahre 2009 einen eigenen Cloud-Anbieter startete. Hierbei konzentrierte sich Alibaba Cloud zunächst auf das Geschäft in China, da es hier den wichtigsten Markt für sich sah. Da der chinesische Markt mittlerweile sehr stark reguliert ist und Unternehmen auf viele Vorgaben beim Betrieb der IT-Infrastruktur achten müssen, hat sich Alibaba Cloud in China als der wichtigste Marktteilnehmer durchgesetzt. Nach der Erschließung des chinesischen Marktes öffnet sich Alibaba immer weiter dem weltweiten Markt. Es werden neue Regionen erschlossen und aktuell betreibt Alibaba weltweit 24 Regionen, dabei einige wenige im US-amerikanischen Raum, aber vorrangig in Asien.

Alibaba Cloud bietet etliche Zertifizierungen für SAP-Workloads an. Hierbei sind die neusten Versionen der SAP S/4HANA Systeme vollständig unterstützt. Die Alibaba Cloud setzt durchgängig auf die gängige Prozessortechologie von Intel mit den neusten

Architekturen von Cascade Lake, Skylake und Broadwell. Die Unterstützung von älteren SAP-Systemen ist jedoch sehr limitiert.

### **Oracle Cloud**

Die Oracle Cloud hat eine starke Präsenz in Funk und Fernsehen durch die Promotion in Kinofilmen gefunden. Die erste Version der Oracle Public Cloud (OPC) war nicht sehr erfolgreich, bot aber alle Eigenschaften und Services einer Public Cloud an. Damals konnten bereits SAP-Workloads auf der OPC installiert und betrieben werden. Im Jahre 2018 wurde die neue Generation der Cloud als „Oracle Cloud Infrastructure“ angekündigt. Es handelte sich hierbei um eine Neuentwicklung der Cloud ohne Bezug auf die vorherige Plattform. Aktuell existieren Regionen der OCI auf allen Kontinenten.

Gleichwohl Oracle als Hersteller von Datenbanken im Umfeld der SAP-Systeme immer weiter an Gewicht verliert und sehr viele Unternehmen sich von Oracle als Datenbankplattform abwenden, gelingt es Oracle dennoch mit den Cloud-Services sein Geschäft auszubauen. Auch im Umfeld von SAP-Workloads schafft es Oracle, seine Cloud-Services zu positionieren. SAP unterstützt die Oracle Cloud weiterhin als Zielplattform, wenn auch nur eingeschränkt. Insbesondere bei SAP-Systemen, welche auf Basis von Oracle Exadata operieren, kann Oracle eine Migration in die Cloud anbieten, welche bei anderen Hyperscalern so nicht funktioniert. Die Unterstützung für SAP S/4HANA Systeme auf der Oracle Cloud ist hingegen nicht gegeben und somit können neuste SAP-Systeme nicht auf der OCI betrieben werden.

### **IBM Cloud**

Der Beginn der IBM Cloud kann auf das Jahr 2010 datiert werden. In diesem Jahr eröffnete IBM zwei Cloud Computing-Rechenzentren in den USA und in Deutschland (Ehningen bei Stuttgart). Dies war damals eher als Private Cloud zu betrachten und nicht als wirkliche Public Cloud, da sie nicht vom Internet aus erreichbar war und kein generelles Portal besaß. IBM selbst datiert den Beginn des Cloud Computings auf die 1950er Jahre, da zu diesem Zeitpunkt zum ersten Mal ein Teilen von Ressourcen der teuren Mainframes begann, was zu der Entwicklung der Virtualisierung führte. Diese ist der Grundstein für das Cloud Computing.

IBM lancierte verschiedene Cloud Services für die IBM Kunden, konnte aber lange Zeit nicht aus dem Bereich der Private Clouds ausbrechen. Mittlerweile bietet die IBM auch eine Public Cloud an, welche als IBM Cloud den Kunden eine große Auswahl von Cloud Services bietet. Die IBM Cloud besitzt eine Vielzahl von Regionen, mit einer starken Fokussierung der Märkte in den USA und Europa. In Asia ist die IBM Cloud auch vertreten, aber nur mit sehr wenigen Standorten.

SAP-Systeme der traditionellen Art (also nicht-HANA), als auch neuere SAP-Systeme basierend auf SAP HANA, wurden seit langer Zeit für die IBM Hardware zertifiziert und viele Kunden nutzen die Systeme von IBM zum Hosting von SAP. Die IBM Cloud per se ist ebenfalls für SAP-Systeme zertifiziert und somit können die neuen SAP S/4HANA Systeme auf der IBM Cloud gehostet werden.

## Private Clouds

Neben den größeren Public Cloud-Anbietern existieren noch die Anbieter von Private Clouds, als auch die traditionellen Hardwarehersteller, welche mit ihrer Hardware die Kunden zum Aufbau von Private Clouds motivieren wollen. Hierunter fallen Hersteller, wie Cisco, HPE oder auch der Softwarehersteller VMware. Alle Hersteller versuchen entweder eigene Clouds aufzubauen oder aber die spezifischen Produkte auch in den Clouds der Hyperscaler verfügbar zu machen.

---

## 1.4 Die Cloud Strategie von SAP

SAP hat sich Cloud seit dem Jahr 2021 sehr groß auf die Fahnen geschrieben, da Cloud als wichtigster Wachstumstreiber für die SAP gilt. Es kann durchaus von einer neuen SAP gesprochen werden. Obwohl die monolithischen SAP-Systeme noch immer in den Unternehmensumgebungen existieren, wird sich der Erfolg von SAP durch die Fähigkeiten der SAP-Software als „Cloud-ready“ ableiten lassen. Darüber hinaus ist die SAP auf starke Partner in der Einführung und Durchsetzung der Cloud-Strategie (Adaption, Umsetzung, Generierung der Werte) angewiesen. Dieses Kapitel zeigt die Entwicklung der SAP-Cloud-Ansätze auf.

### 1.4.1 SAP RISE Programm

Das SAP RISE Programm ist die Umsetzung der Strategie von SAP, immer mehr Umsatz durch Cloud-Services und SAP-Systemen aus der Cloud zu generieren. Dieses Kapitel beschreibt die grundlegenden Komponenten von RISE und wie das Programm durch die Kunden genutzt werden kann.

#### 1.4.1.1 Historie und Features

SAP RISE wurde im Jahr 2021 durch Christian Klein angekündigt, jedoch sind Grundfeiler von SAP RISE schon im Jahr 2020 von ihm genannt worden. Dazu gehört auch der Gedanke der Integration, welche bei SAP RISE ein zentrales Element ist. Im Jahr 2021 reflektierte Christian Klein die starke Ambition, durch die Cloud und durch die Geschäfte in der Cloud, das Unternehmen fokussierter aufzustellen und somit ein starkes Wachstum zu generieren. Das Ziel des Wachstums ist bis 2025 mit 22bUSD beziffert. Daran wird sich das Unternehmen messen lassen müssen. Dabei soll die Cloud als Faktor fungieren, der jedoch mit den folgenden wichtigsten Thematiken einhergeht:

1. Neue Geschäftsmodelle
2. Investitionen in Produkte und Plattformen
3. Fokus auf Industrien

Alle drei Thematiken sollen durch die folgenden drei Produkte erreicht werden, wobei die Produkte nicht trennscharf sind.

- SAP RISE
- SAP Business Technology Platform (BTP)
- Industry Cloud

Hierbei stellt SAP RISE also ein Produkt dar, welches durch die SAP an viele Kunden herangetragen wird und sich auch als komplexes Konstrukt darstellt. Obwohl es nur ein Vertrag zwischen der SAP und dem Kunden gibt, kann der Vertrag etliche Punkte aus SAP RISE abbilden und wird von SAP als „Business Transformation as a Service“ beworben. Die SAP bietet über RISE die folgenden Punkte an:

- **SAP S/4HANA Cloud:** Dies ist die Möglichkeit, die SAP-Systeme zu S/4HANA-Systemen zu transformieren und in einer Cloud per Subskription betreiben zu lassen.
- **Infrastruktur-Anbieter nach Wahl:** Der Kunde kann den Anbieter der Infrastruktur für die SAP S/4HANA-Systeme selbst aussuchen. Dies kann eine Private Cloud, aber auch eine Public Cloud sein.
- **Business Process Intelligence:** Die SAP hat mit der Firma Signavio einen Zukauf getätigt, welcher SAP befähigt, in dem Thema des Process Minings, ein Produkt anbieten zu können. Dies wird unter dem Namen Business Process Intelligence als Teil von SAP RISE angeboten.
- **Custom Code Analyzer, SAP Readiness Check und SAP Learning Hub:** Diese Tools, welche bereits vor SAP RISE existierten, werden als wichtige Bordmittel bei der Transformation Richtung SAP S/4HANA bei SAP RISE mitgeliefert.
- **SAP Business Network Starter Pack:** Hierbei handelt es sich um ein Netzwerk von Partner der SAP, als auch Kunden, welche mit anderen Kunden in einem von SAP gesteuerten Netzwerk zusammenarbeiten können und wollen. Ziel ist es, die Abläufe zwischen den Unternehmen so einfach, wie möglich, zu gestalten.
- **SAP Business Technology Platform:** Die SAP BTP ist faktisch das Konglomerat aller wichtigsten SAP-Technologieprodukte und stellt damit einen neuen Technologiestack dar. Kunden können aus der Vielzahl der Anwendungen und Prozesse, eine passgenaue Lösung zusammenbauen. Im nachfolgenden Abschn. 1.4.3 wird die SAP BTP noch detaillierter erklärt.

Alle obigen Punkte können von den Kunden in einem Vertrag kontrahiert werden. Der Vertrag besteht dann zwischen dem Kunden und der SAP und ihren Partnern. Es können hier mehr als ein Partner ins Spiel kommen.

Als ein wichtiges Kernelement bei SAP RISE gilt der Wandel des Lizenzmodells. Hierbei adaptiert SAP die gängige Praxis der Clouds, welche eine Abkehr von gebundenen Lizzenzen hin zu Subskriptionen macht. Damit können Kunden relativ einfach und flexibel die Nutzung bei SAP RISE bestimmen und müssen nicht durch langwierige

Prozessschritte gehen. Das ist jedoch nur für Kunden interessant, welche die SAP-Lizenzen nicht mehr in den Büchern zu stehen haben und sie nicht mehr abschreiben müssen.

Um die Konfusion bei dem neuen Produkt RISE nicht zu erhöhen, soll im Nachfolgendem gesagt werden, was RISE nicht ist:

- RISE ist keine neue Technologie oder eine komplett neue Technologieplattform
- RISE ist nicht der einzige Weg für Kunden, die SAP S/4HANA Lizenzen zu beziehen
- RISE inkludiert nicht den Wechsel der nicht-SAP-Systeme in die Public Cloud

Insbesondere der letzte Punkt erscheint nicht wirklich relevant, jedoch sind bei einem strategischen Wechsel einer SAP-Umgebung in eine Public Cloud, neben den SAP-Systemen, auch die nicht-SAP-Systeme zu beachten.

#### 1.4.1.2 Wichtigste Kernprozesse

SAP RISE wurde erst im Jahr 2021 angekündigt, allerdings existieren viele der Produkte bereits und der zentrale Punkt der Integration und des Ausbaus der Produkte, wird immer weiter vorangetrieben. SAP hat hierzu Roadmaps aufgestellt – dieses Mal jedoch nicht für die einzelnen Produkte, sondern für die Kernprozesse eines Unternehmens. Die Abkehr von Roadmaps für Produkte ist strategisch wichtig für SAP, da sie gerne die Funktionen aus der Cloud fokussieren möchte und weniger die Produkte in den Vordergrund stellen will.

Die vier definierten Kernprozesse in SAP RISE sind die folgenden:

- **Lead to Cash:** Dies ist der Prozess zur Erfassung von Kaufinteressen, über die Bestellung, bis hin zur Fakturierung und Zahlungsabwicklung.
- **Recruit to Retire:** Dies ist der Prozess für die Abwicklung aller Tätigkeiten und Schritte im Umfeld mit den Beschäftigten des Unternehmens inklusive aller wichtigsten Schritte in der Karriere.
- **Design to Operate:** Dies ist der Prozess der Entwicklung, Herstellung, Vertrieb und Lieferung der Produkte eines Unternehmens. Er ist sehr stark an der Supply Chain orientiert.
- **Source To Pay:** Dies ist der Prozess zur Beschaffung und Bestellung von Dienstleistungen und Produkten, sowie der Bezahlung.

Es ist offensichtlich, dass alle diese vier Prozesse für jedes Unternehmen der Welt zutreffen und somit jeder Kunde von SAP RISE profitieren kann. Der „Lead to Cash“-Prozess wird hierbei durch nicht nur ein Produkt, sondern zwei grundlegenden Produkten von SAP erfüllt: SAP C/4HANA für alle Aufgaben bis zur Erstellung der Bestellung und SAP S/4HANA für die Fakturierung und Rechnungsabwicklung (inklusive Zahlung). Dies zeigt schon die Komplexität in dem Produktportfolio von SAP und in SAP RISE. Die Komplexität wird sehr deutlich, wenn der Prozess „Lead to

Cash“-Prozess noch weiter detailliert wird. Dieser besteht aus den folgenden wichtigsten Prozessschritten, welche durch die dahinter aufgeführten Produkte umgesetzt wird:

1. **Contact to Lead:** SAP S/4HANA, SAP Customer Data Cloud, SAP Marketing Cloud, SAP Commerce Cloud
2. **Lead to Opportunity:** SAP Marketing Cloud, SAP Sales Cloud
3. **Opportunity to Quote:** SAP Sales Cloud, SAP Commerce Cloud, SAP CPQ
4. **Quote to Order:** SAP CPQ, SAP Qualtrics, SAP Commerce Cloud
5. **Order to Cash:** SAP Commerce Cloud, SAP S/4HANA, SAP Service Cloud

In der nachfolgenden Tabelle wird die Komplexität nochmals gezeigt. Zwischen allen den Produkten, schafft die SAP BTP die Schnittstellen und die Möglichkeiten zum Austausch der Daten (Tab. 1.2).

Die obige Tabelle zeigt, wie viele und welche Produkte aus dem SAP RISE-Portfolio für einen Kernprozess genutzt werden müssen. Dies zeigt deutlich, wie viel Aufwand in solch eine Implementierung gesteckt werden muss. Es zeigt auch, dass SAP RISE nicht nur ein Angebot zur Adoption der Public Cloud ist, sondern dahinter eine sehr komplexe

**Tab. 1.2** Nutzung der verschiedenen Produkte

SAP BTP Komponenten	Contact to Lead	Lead to Opportunity	Opportunity to Quote	Quote to Order	Order to Cash
SAP S/4HANA					
SAP Customer Data Cloud					
SAP Marketing Cloud					
SAP Commerce Cloud					
SAP Sales Cloud					
SAP CPQ					
SAP Qualtrics					
SAP Commerce Cloud					
SAP Service Cloud					
SAP RISE umfasst viele Produkte					

Transformation der Geschäftsprozesse steht. Diese Transformation gilt es bei jedem Unternehmen einzeln zu betrachten und zu evaluieren.

Für alle Kernprozess werden von der SAP mit Roadmaps zum aktuellen Stand der Implementierung, der Funktionalitäten und der Integration unterlegt. Somit erhält der Kunde einen guten Überblick zum aktuellen Stand.

#### **1.4.1.3 Aktuelle Situation**

SAP RISE wird von vielen Kunden als mögliches Szenario für einen Wechsel in die Public Cloud erachtet. Hierzu gibt es verschiedene Gründe, welche SAP RISE für Kunden attraktiv machen.

1. Einfacheres Lizensierung- und Preismodell
2. Outsourcing des SAP-Betriebs und Migration in die Public Cloud in einem Schritt
3. Einfachere Transformation zu SAP S/4HANA durch die mitgelieferten Tools
4. Flexibles und einfaches Buchen der zusätzlichen Optionen

Die SAP arbeitet mit Hochdruck daran, allen Kunden ein entsprechendes SAP RISE Angebot zu machen und somit das Ziel, bis 2025 einen Umsatz von 22MUSD mit Cloud zu erzielen, zu realisieren. Dennoch nehmen die Kunden SAP RISE noch zögerlich an. Dies liegt zum Teil daran, dass SAP RISE noch sehr frisch am Markt ist und demnach eine moderate Maturität aufweist. Die Roadmaps zur Implementierung der Kernprozesse zeigt dies deutlich.

Die Verlagerung der SAP-Systeme und die Transformation hin zu SAP S/4HANA geht für viele Kunden mit einer Migration in die Cloud einher. Jedoch bietet SAP dies nur zusammen mit einer Koppelung mit dem zukünftigen Betrieb der SAP-Systeme durch die SAP respektive deren Partner an. Kunden, welche derzeit bei einem verlässlichen Provider sind, sehen wenig Anreize, einen kompletten Anbieterwechsel zu vollziehen.

Letztlich ist SAP RISE zwar noch sehr jung, aber die SAP hat hier ein wichtiges Instrument gebaut, welches für viele Kunden in den nächsten Jahren wichtig und elementar sein wird. Es ist kein alleiniges technisches Angebot, sondern ein gesamtheitliches Angebot zur Transformation des Unternehmens.

#### **1.4.2 SAP HANA Enterprise Cloud**

Die SAP HANA Enterprise Cloud existiert schon lange und bietet Kunden der SAP eine Möglichkeit an, die SAP-Systeme auf von SAP gesteuerten und betriebenen Hardware in den Rechenzentren der SAP (oder Co-Locations) zu hosten. Dieses Kapitel wird einen Überblick zu der SAP HEC geben und beschreiben, wie Kunden die HEC nutzen können.

### 1.4.2.1 Historie

Die SAP HANA Enterprise Cloud, kurz SAP HEC, wurde von SAP als Alternative zu den übrigen Cloud-Anbietern im Jahr 2013 lanciert. Sie sollte den Kunden eine Möglichkeit bieten, schnell und einfach zu SAP HANA zu wechseln und den damit verbundenen Plattformwechsel in einem Schritt mit zu vollziehen. SAP hat hierzu mit vielen Partnern Verträge abgeschlossen, um den Kunden weltweit eine Plattform zu bieten. Damals standen die Kunden vor der Herausforderung, dass für einen Wechsel zu SAP HANA neue, zertifizierte Hardware angeschafft werden musste und diese bei den etablierten Providern nur begrenzt zur Verfügung stand. Diese Lücke füllte die SAP HEC aus.

Als die SAP die HEC lanierte, waren viele Kunden noch auf dem Weg zu SAP HANA und transformierten die SAP-Systeme hierfür. Damals gab es für SAP HANA spezielle Hardware und sehr genaue plus hohe Anforderungen an die Hardware. Nur weniger Kunden konnten diese Anforderungen in den eigenen Rechenzentren erfüllen und so sahen sich viele Kunden nach Alternativen um, damit hohe Investitionen vermieden werden.

Die SAP HEC gilt als Cloud, welche von SAP komplett gewartet und administriert wird (fully managed). Die SAP adressierte durch die Lancierung des Angebots zwei Punkte:

1. Zurverfügungstellung von neuer SAP HANA Hardware als Cloud
2. Angebot eines Fully Managed Service inklusive Patching, Security und SAP-Basis plus SAP Applikationsmanagement

Die Kombination von obigen Punkten war nicht neu auf dem Markt und wurde bereits von anderen Anbietern, wie Fujitsu oder einer IBM, auch angeboten. Neu war damals jedoch, dass die SAP als Hersteller in der einzigartigen Situation war, die von SAP erstellte Software in von SAP betriebenen Rechenzentren für Kunden zu betreiben. Die Kunden folgten dem Gedanken „Der Hersteller der Software wird am besten wissen, wie die Software zu betreiben ist“ und migrierten die SAP-Systeme in die SAP HEC hinein. Die SAP HEC konnte sich hiermit eine sehr fundierte Kundenbasis schaffen, von der sie noch lange zehrte.

Nach der Transformation zu SAP HANA steht für viele Kunden die Transformation zu SAP S/4HANA an. SAP hat die SAP HEC als die Brücke zu der neuen Transformation positioniert. SAP RISE kann als ein Nachfolgeprogramm gesehen werden, welches den Kunden auf dem Weg in die Public Cloud unterstützt.

### 1.4.2.2 Features der SAP HANA Enterprise Cloud

Die SAP HEC wurde als fully managed Cloud lanciert und war, den Kriterien folgend, keine wirkliche Cloud. Sie hat und hatte keine unlimitierten Ressourcen und war auch nicht ohne Weiteres von überall auf der Welt zugreifbar. Vielmehr handelte und handelt es sich um ein Managed Services-Ansatz, der für einige viele Kunden jedoch exakt das ist, was sie benötigen.

Die SAP HANA Enterprise Cloud bietet ihren Kunden auch heute noch einige interessante Features:

- Betrieb der SAP-Systeme durch SAP mit einer flexiblen Preisstruktur bei wachsendem Workload
- Professioneller Betrieb durch die SAP-Hosting innerhalb der Rechenzentren der SAP respektive der Rechenzentren, welche durch die SAP angemietet werden
- Marktkonforme Service Level Agreements für die Verfügbarkeit der SAP-Systeme
- Betrieb der SAP-Systeme bis hin zum Applikationsmanagement (Application Maintenance)
- Unterstützung der Industrielösungen (zum Beispiel IS-U)

Die SAP HEC ist primär für SAP-Systeme geschaffen und so ist die Unterstützung der SAP HEC für alle Systeme, welche nicht von der SAP kommen, sehr eingeschränkt. So können gängige nicht-SAP-Systeme, wie z. B. Archivsysteme, in einen Vertrag mit übernommen werden, aber die Unterstützung (Support) durch die SAP HEC ist sehr limitiert.

Die SAP HEC tritt durch die Bündelung der Services in direkte Konkurrenz zu den Hyperscalern auf der einen Seite. Auf der anderen Seite konkurriert die SAP mit den Anbietern von Managed Services um die Kunden. In beiden Fällen nutzt die SAP HEC ein flexibles Preismodell, welches sich an den jeweiligen Konkurrenten orientiert. So können Kunden entweder Preise basierend auf Volumina (also die Anzahl der Systeme) oder aber Fixpreise erwarten. Die SAP-Systeme werden jeweils als Bündel mit der Hardware zusammen angeboten und so auch abgerechnet. So erhalten die Kunden eine sehr fein granulare Aufstellung der entstandenen Kosten.

Kunden können die SAP HEC nicht einfach so nutzen, sondern werden durch die SAP durch einen unterstützten, begleiteten Prozess geführt. Am Ende dieses Prozesses steht dann die Nutzung der SAP HEC. Der Prozess besteht aus den folgenden Schritten:

1. Assessment – Durch ein Review der aktuellen Systemumgebung, als auch eine genaue Untersuchung der Systeme, gewinnt die SAP einen Überblick zu dem, was später in der SAP HEC betrieben werden soll. Die Ergebnisse aus dem initialen Assessment gehen dann in den Vertrag ein. Dieser Schritt ist im Prinzip der wichtigste, da dies die Grundlage für die spätere Kollaboration zwischen dem Kunden und der SAP bildet.
2. Onboarding – Der Prozess zum Onboarding ist das Schaffen der vertraglichen Grundlage und das Schaffen der initialen Strukturen für die Betreuung der Systeme und des Kunden. Hier werden Prozesse, Kommunikation, Kollaboration, Eskalationen etc. abgestimmt. Da die SAP mit der HEC eine hochstandardisierte Plattform geschaffen hat, können Kunden hier zwar individuelle Lösungen vereinbaren, jedoch gibt es eine klare Präferenz, die Prozesse so zu belassen, wie sie von der SAP angeboten werden. Dies entspricht auch dem Grundgedanken der Cloud.

3. Migration – Nachdem das Onboarding des Kunden abgeschlossen ist, werden die SAP- und nicht-SAP-Systeme sukzessive durch die SAP in die SAP HEC migriert. Hierbei unterstützen die Kunden natürlich und müssen teilweise die Koordination mit dem alten Provider unterstützen. Die SAP wendet für die Migration etablierte Verfahren an. Insbesondere bei den kombinierten Migrationen und Upgrades auf HANA konnte sie SAP die Vorteile voll ausspielen und Verfahren wie DMO (Data Migration Option) zum Einsatz bringen.
4. Betrieb – Nach der erfolgreichen Migration der Systeme werden diese in den Betrieb übernommen und durch die SAP mit den zugesicherten SLAs betrieben. Bei einer Erweiterung der Systeme (z. B. zusätzliche SAP-Systeme) existieren etablierte Prozesse und ein Service Katalog, aus dem die Kunden die jeweiligen Services beziehen können.

Der obige Prozess unterscheidet sich nicht grundlegend von den Prozessen anderer Hersteller und stellt kein Unterscheidungskriterium da.

Für den Weg in die SAP HEC existiert der obige Prozess und in selben Masse existiert auch der Prozess für das Verlassen der SAP HEC. Hierbei vereinbart die SAP üblicherweise einen individuellen Ansatz mit genau definierten Leistungen und Umfängen („Exit Out“).

#### **1.4.2.3 Aktuelle Situation**

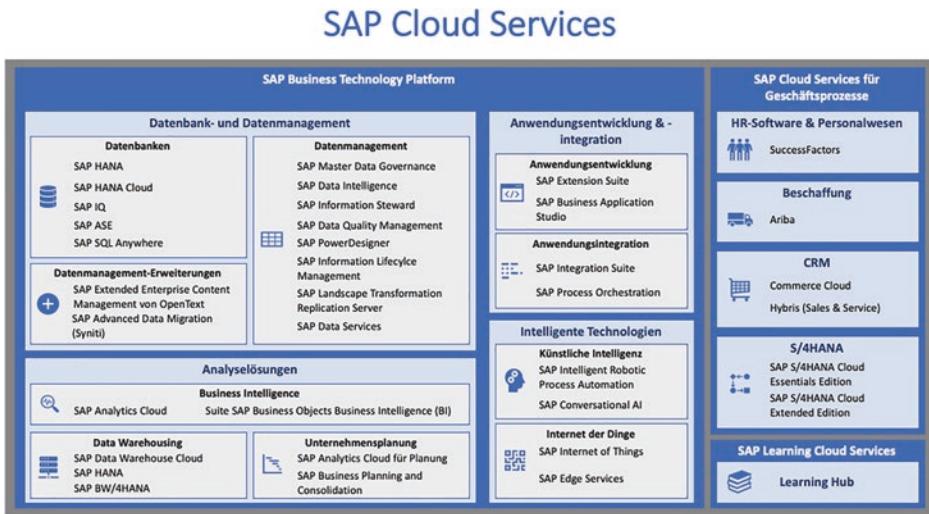
Die SAP HEC hat mit der Zeit ihre Einzigartigkeit verloren. Als wichtigstes Argument galt die Verfügbarkeit von SAP HANA-zertifizierter Hardware, welche jedoch mittlerweile in allen gängigen Hyperscaler Clouds zur Verfügung steht und auch sehr einfach zugänglich ist.

Weiterhin zeigt die SAP HEC immer noch eine sehr komplexe Preisstruktur für verschiedene Services und Optionen. Das Vertragswerk ist komplex und führt durch die Bindung an die höheren Preise zu einem Umdenken bei den Kunden, als auch bei der SAP. Die SAP hat RISE lanciert, um von dem Wachstum der Hyperscaler zu profitieren. Viele Kunden bewegen sich nun nicht mehr auf der Welle der SAP HANA-Transformation, sondern haben diese schon hinter sich gebracht und beschäftigen sich mit SAP S/4HANA.

Es ist sicherlich fraglich, wie lange die SAP HEC noch existieren wird und wann sie letztlich von SAP RISE abgelöst wird.

#### **1.4.3 Zusätzliche SAP Services aus der Cloud**

Neben den bereits genannten SAP-Lösungen besteht das Portfolio aus vielen weiteren Cloud basierten Services. Mit diesen zusätzlichen Services deckt die SAP mehrere technologische, als auch prozessuale Bereiche ab. Durch die vielfältige Kombination verschiedener Cloud Services aus dem SAP-Portfolio lässt sich eine SAP S/4HANA



**Abb. 1.1** SAP Cloud Canvas

Landschaft in eine innovative und effektive Umgebung zur Ausführung der Geschäftsprozesse transformieren. SAP bezeichnet eine optimierte und automatisierte Umgebung auch als „Intelligent Enterprise“. Der Begriff beschreibt einen ganzheitlichen Ansatz, dessen technologische Grundlage S/4HANA bildet. Darauf aufbauend werden diverse Cloud Services aus dem SAP Portfolio an die S/4HANA Architektur angebunden und dadurch in die Geschäftsprozesse integriert. Die folgende Übersicht in Abb. 1.1 stellt schematisch eine Übersicht der bekanntesten Cloud Lösungen dar.

### SAP Business Technology Platform

Ein zentraler Bestandteil des Cloud Service Angebots ist die SAP Business Technology Platform (SAP BTP). Darunter wird eine Cloud basierte Plattform verstanden, die mehrere Services auf Grundlage technischer oder fachlicher Gemeinsamkeiten gruppiert und kategorisiert. Mithilfe dieser der Business Technology Platform zugeordneten Services haben Unternehmen die Möglichkeit, die S/4HANA Landschaft zu erweitern, um die Geschäftsprozesse zu vernetzen und zu optimieren. Die Vernetzung der Prozesse kann mithilfe der Cloud-Technologie übergreifend über mehrere Unternehmensbereiche hinweg erfolgen. Die Business Technology Platform greift dabei auf eine Multicloud-Strategie zurück, wodurch der SAP-Kunde eine individuelle Auswahl über den Hyperscaler-Anbieter treffen kann. Dadurch werden die bezogenen Services der Plattform in der ausgewählten Hyperscaler-Cloud ausgeführt. Bevor jedoch die Entscheidung für einen Hyperscaler-Anbieter getroffen wird, eignet sich eine Prüfung zur Verfügbarkeit der jeweiligen Cloud Services, da vereinzelte Services lediglich bei bestimmten Cloud-Anbietern und in ausgewählten Regionen zur Verfügung stehen.

Die Cloud Services in der SAP Business Technology sind aus technologischer Perspektive in verschiedene Kategorien unterteilt. Innerhalb einer Kategorie sind die jeweiligen Services basierend auf deren technologischem Anwendungsumfeld gruppiert.

Die **Kategorie „Datenbank- und Datenmanagement“** setzt sich aus Cloud Services zusammen, die sich mit der Speicherung und Verarbeitung von Daten befassen. Bei der SAP HANA Cloud handelt es sich beispielsweise um eine In-Memory-Datenbank, die als Service in der Cloud angeboten wird. Dieses Modell wird auch als Datenbank-as-a-Service (DBaaS) bezeichnet. Als Konsument dieses Services wird eine Cloud basierte HANA-Datenbank zur Verfügung gestellt, die unter anderem Datenabfragen in Echtzeit ermöglicht.

Aufbauend auf der Speicherung von Daten können diese mithilfe der Cloud Services aus der **Kategorie „Analyselösungen“** ausgewertet werden. Eine der am häufigsten eingesetzten Services dieser Kategorie ist die SAP Analytics Cloud. Diese Software-as-a-Service (SaaS) ermöglicht umfangreiche und individuelle Auswertungen mithilfe von Business Intelligence. Die Analysen können auch zu Prognose- und Planungszwecken in einem Unternehmen hinzugezogen werden. Die SAP Analytics Cloud bietet eine Vielzahl an Funktionen, um die Reportingprozesse zu optimieren und zu automatisieren.

Der Bereich der Anwendungsentwicklung und -integration war ursprünglich unter dem Begriff „**SAP Cloud Platform**“ bekannt und wurde im Rahmen einer Portfolio-Umstrukturierung der SAP ersetzt. Die Services der früheren SAP Cloud Platform finden sich als „SAP Extension Suite“ und „SAP Integration Suite“ wieder. Beide Lösungen bilden eine Platform-as-a-Service, in der Applikationen als Erweiterung zu einer S/4HANA Instanz entwickelt und bereitgestellt werden können. Die Anwendungen können dabei auf Grundlage verschiedener Software-as-a-Services, wie beispielsweise Workflow Management, aufgebaut werden.

Die vierte Kategorie der SAP Business Technology Platform umfasst die **intelligenten Technologien**, wie künstliche Intelligenz und Internet der Dinge. Die in diesem Bereich angesiedelten Services unterstützen die Automatisierung von Geschäftsprozessen. Mithilfe der SAP Conversational AI besteht beispielsweise die Möglichkeit, einen Chatbot in das Fiori Launchpad des S/4HANA Systems zu integrieren. Darauf aufbauend können die jeweiligen Transaktionen im Hintergrund durch SAP Intelligent Robotic Process Automation automatisiert durchgeführt werden.

Zusammengefasst steht die SAP Business Technology Platform folglich für eine Sammlung von innovativen Cloud Services, die zur Erweiterung und Optimierung in eine S/4HANA Landschaft integriert werden können. Die Cloud Lösungen der SAP BTP können dabei in nahezu allen Geschäftsprozesse genutzt werden.

### **SAP Cloud Services für Geschäftsprozesse**

Neben der technologischen Perspektive besteht das Cloud Service Portfolio der SAP aus spezifischen Lösungen, die sich auf bestimmte Geschäftsprozesse oder -bereiche fokussieren. Für die Durchführung von Prozessen des Personalwesens wird dabei häufig

auf SuccessFactors zurückgegriffen. Mithilfe von SuccessFactors können Prozesse rund um Gehaltsabrechnungen, Talentmanagement, Feedback und Personalplanungen abgewickelt werden.

Mit Ariba bietet SAP einen Cloud Service zur Digitalisierung und Automatisierung des Bestellprozesses an. Der Service unterstützt Unternehmen beim Einkaufsprozess in den Bereichen Lieferantenmanagement, Management der Logistikkette, Beschaffung, sowie Verkauf und Auftragsabwicklung.

In der Kategorie „Customer-Relationship-Management“ stellt die SAP sowohl die Commerce Cloud als auch SAP Hybris bereit. Die SAP Commerce Cloud fokussiert sich dabei auf die Optimierung der E-Commerce-Prozesse. So kann mit diesem Service beispielsweise ein Online-Shop aufgebaut werden, in dem Geschäftspartner und Kunden Bestellungen tätigen können, die in der Commerce Cloud bearbeitet werden.

Übergreifend zu den spezifischen Cloud Services besteht geschäftsprozessübergreifend die Möglichkeit, SAP S/4HANA als Software-as-a-Service zu beziehen. Dieser Ansatz wird in zwei unterschiedlichen Editionen angeboten, die im Abschn. [3.4](#) genauer beschrieben werden.

### SAP Learning Cloud Services

Die dritte Kategorie des Service Portfolios enthält die SAP Learning Cloud Services. Konkret ist damit das SAP Learning Hub gemeint. Mithilfe dieser Cloud Lösung kann ein Unternehmen die Fort- und Weiterbildung seiner Mitarbeiter fördern, da es sich um eine zentrale Lernplattform handelt. Die dort verfügbaren Trainings und Lernmaterialien ermöglichen die Vertiefung des Wissens rund um SAP und unterstützen Unternehmen dabei, wie SAP bestmöglich in den täglichen Prozessen genutzt werden kann.

---

## 1.5 Zusammenfassung

Das erste Kapitel gab einen Überblick zu den Anfängen von Cloud Computing und zu den drei wichtigsten Typen von Cloud (Private, Public, Hybrid). Hierbei wurde gezeigt, wodurch sich Cloud Computing auszeichnet und wie Unternehmen von Cloud Computing profitieren können.

Die drei wichtigsten Anbieter von Public Clouds, Amazon Web Services, Microsoft Azure und Google Cloud, wurden vorgestellt und jeweils eine kurze Zusammenfassung zu der Entstehungsgeschichte zur Verfügung gestellt. Da die aktuellen Produktpportfolios sehr komplex sind, wurde bewusst auf eine Auflistung aller möglichen Services der Hyperscaler verzichtet und eher auf die Vor- und Nachteile der jeweiligen Hyperscaler aus Sicht der Kunden eingegangen.

Ein weiterer großer Abschnitt widmet sich der SAP als Anbieter der weltweit wichtigsten ERP Software und als Anbieter von neuen Methodiken zum Betrieb der

SAP-Systeme. Die SAP bietet neben der SAP HEC und RISE auch weitere Services aus der Cloud an, welche ebenfalls beschrieben worden sind.

Somit haben Sie, liebe Leser, eine sehr fundierte Basis für die weiteren Kapitel des Buches, welche auf den wichtigen Grundlagen aus dem ersten Kapitel aufbauen.



# SAP S/4HANA-Systeme in den Public Clouds

2

## Zusammenfassung

Die Architektur von SAP S/4HANA-Systemen unterscheidet sich aus der technischen Sicht heraus nicht wesentlich von der Architektur der vorherigen SAP ERP-Systeme. Dennoch ist es wichtig, die Kernkomponenten eines SAP S/4HANA-Systems zu kennen, damit die Architektur des Systems in der Public Cloud verstanden werden kann. Basierend auf der Architektur werden die wichtigsten Anwendungsfälle der SAP S/4HANA-Systeme in der Cloud besprochen: das Sizing, die Kosten für ein System in der Cloud, die Hochverfügbarkeit, die Ausfallsicherheit bei einem Desaster, die Datensicherung via Backup und Restore, die Integration in das Netzwerk des Unternehmens, sowie die Automatisierung der SAP S/4HANA-Systeme und der horizontalen plus vertikalen Skalierung.

## 2.1 Überblick zur SAP S/4HANA Architektur

Die SAP S/4HANA-Systeme basieren auf einer konsequenten Weiterentwicklung der SAP-Systeme von den R/3-Versionen zu den HANA-basierten SAP-Produkten hin. Sie unterstützen alle wichtigen Geschäftsprozesse eines Unternehmens und stellen damit das Rückgrat des Geschäftsbetriebs dar.

### 2.1.1 Entwicklung zu SAP S/4HANA

Im Jahre 1979 wurde das SAP R/2-System auf Mainframes vorgestellt und entwickelte sich damals zu einem der Standard-ERP-Systeme weltweit. Die Architektur wurde eine

lange Zeit stabil gehalten und erst im Jahre 1992 wurde durch die SAP ein signifikanter Schritt zur Evolution eingeläutet: das war das SAP R/3-System. Zum ersten Mal wurden Geschäftstransaktionen in Echtzeit (Realtime) durchgeführt und das System kombinierte alle wichtigen Geschäftsprozesse innerhalb eines großen SAP-Systems.

SAP verfolgt danach den Weg der ERP Central Components, kurz ECC. Hierbei konnten Kunden die wichtigsten Funktionen, welche für die Unternehmen wichtig waren, aktivieren und nutzen. Es gab auch spezifische Systeme, wie z. B. Supplier Relationship Management, welche zusätzlich zu den SAP ECC-Systemen installiert werden konnten. Ab dem Jahre 2004 gab es hierdurch eine Erneuerung in den Systemlandschaften der Kunden und die älteren SAP R/3-Systeme wurden sukzessive durch die SAP ECC-Systeme abgelöst. Die Dauer von 12 Jahren von R/3 zu ECC zeigt die Komplexität und Kompliziertheit bei der Transformation solcher ERP-Systeme.

Im Jahre 2010 wurde von der SAP die neue Datenbank HANA angekündigt. Die neue Plattform HANA wurde von der SAP nicht nur als neue Datenbank platziert, sondern auch als Entwicklungsplattform für OLAP-Anwendungen (Online Analytical Processing). Hierüber konnten viele Schritte in Auswertung von großen Datenmengen einfach in der HANA-Datenbank vorgenommen werden, anstatt, wie früher, innerhalb der Applikationsschicht. Ab 2013 waren dann die wichtigsten Anwendungen der SAP auf HANA verfügbar – vorher waren nur einige wenige Anwendungen für HANA freigegeben. Mit der «Suite on HANA» wurden weitere Features eingeführt, wie beispielsweise der Zugriff von mobilen Endgeräten. Dies funktionierte vorher zwar auch schon, aber wurde durch die neue Plattform signifikant vereinfacht.

In dem Jahr 2015 wurde von der SAP das S/4HANA-System vorgestellt. Hierbei handelt es sich um ein vereinfachtes (S) System für (4) HANA-basierte Systeme. Es ist im Kern ein komplett neu gestaltetes System, basiert immer noch auf der Programmiersprache ABAP, aber wurde komplett neu erstellt. Mit dieser Ankündigung vollzog die SAP auch einen Wandel weg von der Unterstützung aller gängigen, relationalen Datenbanksysteme, hin zu der exklusiven Unterstützung von der SAP-eigenen Datenbank HANA.

## 2.1.2 SAP S/4HANA in der Cloud

Mit der Entwicklung der neuen SAP S/4HANA-Plattform wurde von der SAP auch ein Hauptaugenmerk auf die Cloud gelegt. Die neuen Produkte sollten nicht mehr nur in den Rechenzentren der Kunden oder Partner betrieben werden können, sondern insbesondere auch in der Cloud. Der Grundgedanke von Software-as-a-Service trat immer mehr in den Vordergrund. Dies war sicherlich dem enormen Wachstum der Cloud-Services geschuldet, aber auch durch die Akquisitionen von z. B. SAP Ariba oder SuccessFactors, welche als reine SaaS-Lösungen vertrieben worden. Die SAP sah in dem Geschäft mit der Cloud die Zukunft und dies ist auch im Jahre 2021 noch immer eine der grundlegenden Strategien.

Als eines der ersten Produkte, basierend auf der neuen S/4HANA-Architektur, wurde die **S/4 Public Cloud Multi Tenant Edition** (MTE) im Jahre 2017 gestartet. Das Angebot ist ein SAP S/4HANA-System, welches als SaaS bereitgestellt wird. Kunden konnten und können darauf basierend die neuen S/4-Funktionalitäten nutzen. Das Angebot richtet sich prinzipiell an Kunden, welche eine sehr hohe Standardisierung anstreben und wenig Customizing benötigen. Dafür erhalten die Kunden dann ein System, welches pro Jahr regelmäßig ein Upgrade erfährt und damit immer die neuste Codebase enthält. Im Jahre 2019 wurde die Multi Tenant Edition zu **Essentials** umbenannt.

Der MTE gegenüber wurde auch eine **S/4 Public Cloud Single Tenant Edition** (STE) im Jahr 2018 – ebenfalls als reine SaaS-Lösung, gestartet. Hierüber konnten Kunden ein neues S/4HANA-System beziehen und konnten hier mehr Customizing, als bei der MTE durchführen. Darüber hinaus unterstützte die STE auch mehr Industrie-spezifische Geschäftsprozesse. Änderungen an dem eigentlichen ABAP-Code waren aber auch hier nicht möglich. STE wurde im Jahr 2019 in **Extended** umbenannt.

Im Jahre 2020 startete die SAP die **S/4 Private Cloud Edition** im Pilotbetrieb und ab dem Jahr 2021 war diese Edition auch für alle Kunden verfügbar. Hierbei handelt es sich um ein SAP S/4HANA-System, welches durch die SAP für die Kunden betrieben und unterstützt wird. Die Kunden haben hierbei die komplette S/4HANA Code Base zur Verfügung und können das System komplett anpassen (Customizing). Die SAP führt ein Mal pro Jahr ein Upgrade aus. Im Gegensatz zur Essentials Edition unterstützt die Private Cloud Edition 25 Industrie-spezifische Prozesse.

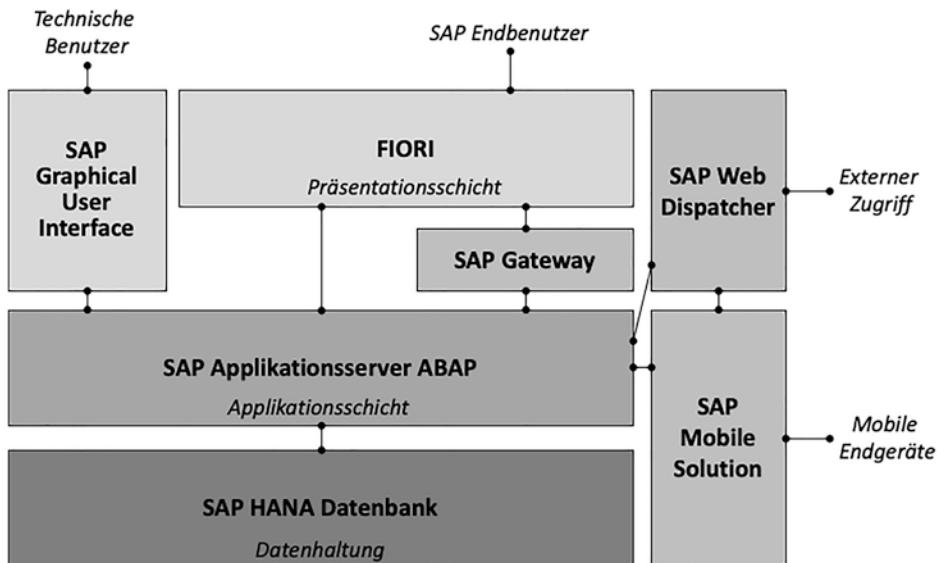
Neben den cloudbasierten Angeboten können die Kunden auch noch immer das traditionelle SAP-System in der Rechenzentrumswelt (on premise) behalten. Hierbei können die Systeme in den traditionellen Rechenzentren oder aber auch in den Hyperscaler-Clouds betrieben werden.

### 2.1.3 SAP S/4HANA Architektur

Die Architektur von SAP S/4HANA kann aus verschiedenen Perspektiven beschrieben werden und die unterschiedlichen Anwendungsfälle verschiedene Komponenten benötigen. So ist der Aufbau eines BW4HANA-Systems anders als die Architektur eines S/4HANA Systems. Dennoch gibt es die Grundkomponenten, welche in allen SAP-Systemen gleich sind.

#### Überblick

Die Architektur von SAP S/4HANA-Systemen unterscheidet sich nicht grundlegend von der Architektur der SAP ECC-Systeme. Es existieren immer noch die gleichen Komponenten, welche überblicksartig in der nachfolgenden Abbildung zusammengefasst sind (Abb. 2.1).



**Abb. 2.1** SAP S/4HANA Architektur

Die verschiedenen Schichten des SAP S/4HANA-Systems sind wie folgt:

- **Datenhaltungsschicht**: Diese Schicht wird durch die spaltenorientierte HANA Datenbank erfüllt.
- **Applikationsschicht**: Diese Schicht wird durch den Applicationserver ABAP gebildet.
- **Präsentationsschicht**: Diese Schicht wird durch die SAPGui und Fiori gebildet.
- **Kommunikationsschicht**: Diese Schicht wird durch das SAP Gateway, SAP WebDispatcher und SAP Mobile Solution gebildet.

Alle die oben genannten Schichten werden im Folgenden etwas detaillierter erläutert.

### Datenhaltung

Die Speicherung und Verarbeitung der Daten finden in der Datenbank statt. Bei allen SAP S/4HANA-Systemen handelt es sich um eine HANA-Datenbank. Die HANA-Datenbank hält alle Daten in relationalen Tabellen im Hauptspeicher, welche über Schlüsselbeziehungen miteinander verbunden sind. Diese grundlegende Funktionalität ist in allen SAP-Systemen gleich. Die HANA-Datenbank ist eine spaltenorientierte Datenbank, welche die Relationen spaltenorientiert abspeichert. Durch die Datenhaltung im Hauptspeicher, erlangt die Datenbank eine sehr hohe Geschwindigkeit gegenüber den traditionellen Datenbanken, wie zum Beispiel Microsoft SQL Server. Diese hohe Geschwindigkeit hat jedoch auch einen Preis: die HANA Datenbank benötigt sehr große und sehr starke Hardware.

Seit der Einführung von HANA wurde die Funktionalität der HANA-Schicht sukzessive erweitert. Die SAP verfolgt die Strategie, dass gewisse Schritte und Operationen in der Datenbankschicht erfolgen sollen. Hierbei geht es beispielsweise um die Aufbereitung von Daten, welche früher oft in der Applikationsschicht geschah, nun aber in der Datenbankschicht erfolgen soll. Dieses Konzept des «Code Pushdowns» wird in allen S/4HANA-Systemen gelebt und wird auch in die neuen ABAP-Programme eingearbeitet. Somit ergibt sich zu den früheren SAP-Systemen eine signifikante Änderung: die Datenbank wird als intelligente Komponente des SAP S/4HANA-Systems eingesetzt und ist umso wichtiger.

### **Applikationsschicht**

Die Applikationsschicht verarbeitet die Daten aus der HANA-Datenbank. In den SAP S/4HANA-Systemen ist das der Applikationsserver ABAP. In den vorherigen SAP-Systemen (wie z. B. ECC) gab es noch den Applikationsserver Java, welcher jedoch strategisch nicht mehr verwendet wird. Die Programmiersprache ABAP (Advanced Business Application Programming) ist eine eigene Sprache von SAP und bereits seit Urzeiten von SAP im Einsatz.

Der Applikationsserver ABAP verarbeitet nicht nur die transaktionalen Daten und Programme des S/4HANA-Systems, sondern stellt auch die Funktionalitäten für analytische Aufgabenstellungen und Suchanfragen bereit. Damit ist der Applikationsserver neben der HANA Datenbank eine der zentralen Komponenten. Der ABAP Applikationsserver besteht aus einigen sehr wichtigen Prozessen, welche in den Kapiteln zu den Implementierungen auf den Public Clouds sehr wichtig werden. Hierzu gehören folgende Prozesse:

- Dispatcher: Der Prozess verteilt alle eingehenden Anfragen auf die Work Prozesse.
- Work Prozesse: Die Work Prozesse verarbeiten die Benutzeranfragen, führen die ABAP-Programme aus und greifen auf die Daten der HANA-Datenbank zu.
- Gateway: Das Gateway bedient eingehende und ausgehende Anfragen an das SAP S/4HANA-System.
- Verbucher: Die Prozesse verbuchen die Ergebnisse von Transaktionen.
- Enqueue: Die Prozesse sperren Datensätze, um parallele Änderungen oder Änderungen im Hintergrund zu verhindern.
- Background: Die Prozesse dienen für langlaufende Programme (Jobs), welche sehr viele Daten ändern oder auswerten müssen.
- Message Server: Der Messageserver dient zur Kommunikation zwischen einzelnen Prozessen des SAP-Systems.
- Spool Prozess: Die Spool-Prozesse bedienen Druckaufträge, unter anderem auch für Massendrucke

Alle Prozesse sind in dieser Form auch in den älteren SAP-Systemen zu finden und spielen dort die gleiche wichtige Rolle.

## Präsentationsschicht

SAP hat im Laufe der Zeit verschiedenste Programme herausgebracht, welche als zentrale Schnittstelle der Benutzer mit den SAP-Systemen dienten. Das wohl bekannteste Programm ist die SAPGui, welche schon seit sehr vielen Jahren und vielen Entwicklungsschritten das zentrale Bedienelement von SAP ist. Daneben gab und gibt es aber weitere Produkte, wie z. B. für BusinessObjects.

Für die SAP S/4HANA-Systeme existiert mit Fiori eine neue Komponente, welche als zentrale Schicht die Aufgaben zur Repräsentation von Daten und die Bereitstellung von Funktionen übernimmt. Fiori ist keine optionale Komponente mehr, wie sie es vor ein paar Jahren noch war, sondern ist essenziell für die volle Nutzung der SAP S/4HANA-Systeme. So stellt Fiori auch Applikationen (apps) bereit, welche z. B. für Robotics oder Analytics genutzt werden. Der Zugriff per SAPGui ist zwar immer noch möglich, aber Endbenutzer sollten idealerweise per Fiori zugreifen.

Die drei Komponenten Fiori plus SAP HANA Datenbank plus der Applikationsserver ABAP bilden das SAP S/4HANA-System. Es existieren weitere, optionale Komponenten, welche zur Erweiterung der Funktionalitäten oder der Integration benötigt werden.

## Kommunikationsschicht

Es existieren drei wichtige technische Komponenten, welche in den SAP-Systemumgebungen installiert werden: das SAP Gateway, ein SAP WebDispatcher und die SAP Mobile Plattform.

Das SAP Gateway ist nicht mit dem normalen Gateway des Applikationsservers ABAP zu verwechseln. Es handelt sich hierbei um eine neue Komponente, welche auch als alleinstehende Komponente installiert werden kann. Das Gateway bietet SAP und nicht-SAP-Applikationen die Verbindung zum SAP S/4HANA-System an. So können auch nicht-SAP-Applikationen über die OData-Services (Open Data) auf die Daten des SAP-Systems zugreifen.

Der SAP WebDispatcher fungiert in allen Systemumgebungen als Proxy für Verbindungen in das Internet. Der WebDispatcher wird nicht parallel zum SAP-System installiert, sondern in einer DMZ (Demilitarisierten Zone), in der ein Zugriff ins Internet und aus dem Internet heraus sicher ist. Die SAP S/4HANA-Systeme sind nie direkt aus dem Internet erreichbar.

Die SAP Mobile Plattform ist eine dritte Komponente, welche den direkten Zugriff auf das SAP S/4HANA-System von mobilen Endgeräten aus ermöglicht. Auf diesem Weg können Benutzer von den Mobilgeräten mit Android oder iOS auf das SAP-System nativ zugreifen (also ohne eine Weboberfläche nutzen zu müssen). Es handelt sich bei der SAP Mobile Plattform ebenfalls um ein separates kleines SAP-System, wofür eine entsprechende Architektur definiert und implementiert werden muss.

Obwohl die Elemente der Kommunikationsschicht als optional zu sehen sind, können die Komponenten in nahezu allen Umgebungen angetroffen werden.

### Zusätzliche Komponenten

Je nach Einsatzgebiet der SAP S/4HANA-Systeme können in den Kundenumgebungen weitere Applikationen/Schnittstellen/Infrastrukturkomponenten provisioniert werden. Dies ist sehr von den Anforderungen der Unternehmen abhängig. Sehr häufig finden sich Komponenten zur Kommunikation mit anderen nicht-SAP-Systemen, wie z. B. Lagerhaltungssystemen oder Archivsysteme. Diese Komponenten sind bei den Überlegungen zur Architektur ebenfalls mit zu betrachten und zu berücksichtigen.

---

## 2.2 Anwendungsfälle für SAP S/4HANA in der Cloud

Dieses Kapitel wird die wichtigen technischen Anwendungsfälle für SAP S/4HANA-Systeme auf den Public Clouds beschreiben. Hierbei geht es noch nicht um die konkrete Implementierung, sondern um die Wichtigkeit der Anwendungsfälle, wie z. B. das korrekt Sizing oder die richtige Auswahl der Verfügbarkeit. Diese Punkte sind vor einer Provisionierung der SAP S/4HANA-Systeme zu adressieren.

### 2.2.1 Sizing

Vor der Instanzierung von neuen SAP S/4HANA Systemen ist ein fundiertes Sizing wichtig. Hierbei ist das Sizing der Prozess zur Abschätzung der zukünftigen Systemgröße und des zu erwartenden Wachstums.

#### 2.2.1.1 Wichtigkeit des Sizings

Das Sizing stellt die Grundlage für einen späteren stabilen Systembetrieb dar. Ein SAP S/4HANA System, welches zu klein dimensioniert wurde, läuft Gefahr, dass es die Benutzeranfragen nicht schnell genug abarbeiten kann. Ein SAP S/4HANA System, welches zu groß dimensioniert wurde, kann zwar die Benutzeranfragen gut abarbeiten, aber verursacht dabei hohe Kosten, welche in der Public Cloud zu Buche schlagen. Daher gilt es einen guten Zielwert für die Größe des neuen Systems/des zu migrierenden Systems zu finden. Es gilt unnötige Kosten zu vermeiden, aber auch den stabilen Systembetrieb sicherzustellen. Das Sizing umfasst neben der Auswahl der zukünftigen Größe der virtuellen Maschine/n auch den notwendigen Speicherbedarf, als auch Backup und sonstige Ressource (wie z. B. Load Balancer).

Ein SAP S/4HANA System, welches zu klein dimensioniert wird, kann verschiedene Probleme verursachen. Es gilt, diese zu vermeiden:

1. Stabilitätsprobleme – Ein Sizing, welches zu klein gewählt wird, kann zu erheblichen Problemen in der Stabilität eines SAP-Systems führen. So können eine sehr hohe Auslastung des Hauptspeichers (größer 98 %) zu gravierenden Stabilitätsproblemen führen. Neben dem Hauptspeicher kann ein zu klein dimensionierter Storage oder

eine falsche Storageanbindung zu merklichen Performanceengpässen führen, welche einen großen Einfluss auf die Benutzer und Geschäftsprozesse haben kann.

2. Zu geringe Performance – Die Performance von SAP-Systemen ist kritisch bei der Ausführung der Geschäftsprozesse der Unternehmen. So müssen viele Daten in möglichst geringer Zeit verarbeitet werden und es finden sich Branchen, wie die Retail-Branche, in denen die SAP-Systeme viele kleine Transaktionen in möglichst kurzer Zeit prozessieren müssen. Eine geringe Performance kann nicht nur Auswirkung auf die Benutzer haben, sondern ganze Lieferketten stören und zu Verzögerungen führen, welche die Kunden viel Geld kosten können.
3. Häufige Wartungen – Wenn Systeme nicht den Anforderungen entsprechend aufgebaut werden, besteht das Risiko, dass häufige Wartungen durchgeführt werden müssen. Für solche Wartungen sind oftmals Downtimes der SAP-Systeme notwendig, welche letztendlich wieder zu einer starken Auswirkung auf die Benutzer und Prozesse haben.
4. Stillstand – Der wohl schlimmste Fall, durch ein falsches Sizing, ist der Stillstand eines Systems. Dies kann durch eine Überlastung der virtuellen Maschine und/oder des Speichers erfolgen und somit zu einem Komplettabsturz führen. Insbesondere Systeme in Hochverfügbarkeitsclustern zeigen solche Muster und sollten mit viel Bedacht einem Sizing unterzogen werden.

Die aus dem Sizing hervorgehenden Daten sind die Anhaltspunkte für die spätere Dimensionierung des Systems und den Aufbau des Systems. Das nächste Kapitel beschreibt, wie der Prozess des Sizings für SAP S/4HANA Systeme durchgeführt wird.

### **2.2.1.2 Prozess zum Sizing**

Der Prozess zum Sizing eines Systems orientiert sich üblicherweise an den bekannten Rahmenparametern für ein neues System oder für die Migration eines Systems. In beiden Fällen jedoch werden die folgenden Schritte ausgeführt:

#### **Schritt 1: Transparenz zu Datenvolumen**

Die Größe eines neuen Systems leitet sich zunächst von der Größe des aktuellen Systems ab oder von der zu erwartenden Größe. Hierzu ist wichtig, dass bestehende Systeme nicht 1:1 so übernommen werden. Insbesondere bei beispielsweise Business Warehouse-Systemen gibt es oftmals viele Daten, welche nicht mehr gebraucht werden. Für SAP ERP Systeme gilt ähnliches. Oftmals werden keine Archivierungsläufe durchgeführt und somit Vergrößern sich die Systeme Tag für Tag. Es gilt also, die Systeme vorzubereiten und eine Archivierung/Selektion von zu löschen Daten vorzunehmen. Darüber hinaus sollte bereits jetzt eine Strategie festgelegt werden, wie mit zukünftigem Datenwachstum umgegangen wird. Hier können beispielsweise Near Line Storage oder auch Data Aging über SAP HANA genutzt werden.

### Schritt 2: Sizing ausführen

Zum eigentlichen Sizing bietet die SAP die Sizing Reports an, welche innerhalb eines SAP ERP Systems ausgeführt werden können. Für Brownfield Transformationen/Migrationen gilt die SAP OSS Note 18721170, um die zukünftige Größe für ein transformiertes System abzuschätzen. Nach der Ausführung der Sizing Reports im System, kann darauf basierend eine Zielgröße definiert werden. Für Greenfield Implementierungen, sollte der QuickSizer genutzt werden. Dieser kann basierend auf wenigen Inputfaktoren, ein erstes Sizing für die zukünftigen Systeme ermitteln.

### Schritt 3: Anpassen der Resultate

Nachdem ein erstes Sizing erstellt worden ist, ist eine Anpassung der Resultate wichtig. Hierfür sollten zwei Faktoren berücksichtigt werden: das zukünftige Wachstum und die speziellen Nutzungsszenarien.

- **Wachstum:** Das generelle Wachstum von SAP S/4HANA Systemen wird mit **10 % jährlich** angenommen. Dieses Wachstum bezieht sich primär auf die Daten, welche zusätzlich in einem Jahr erstellt werden. Wird von einem Sizing von 3 TB für das Jahr 1 ausgegangen, so muss im Jahr 2 mit einer Zielgröße von 3.3 TB gerechnet werden. In der Zeit, als es noch keine HANA Datenbanken gab, spielte das Wachstum eine Rolle bei der Dimensionierung des Storage. Nach der Einführung von HANA und der In-Memory-Technologie muss das Wachstum vor dem Hintergrund der Hauptspeichergröße der virtuellen Maschinen gesehen werden. Durch das Wachstum wird ein regelmäßiges Re-Sizing von Maschinen notwendig.
- **Spezielle Nutzungsszenarien:** Kurze Leistungsspitzen bei der Nutzung der SAP S/4HANA Systeme können zwar kurze Auswirkungen auf die Benutzer haben, sind aber meist verkraftbar. Bei längeren Leistungsspitzen oder sogar Überlastungen durch lange Monatsabschlüsse/Jahresabschlüsse, sollte das Sizing auf Basis der Erfahrungen der bisherigen Systeme angepasst werden.

### Schritt 4: Sizing zu Hardware übertragen

Für alle neuen Systeme oder zu migrierenden Systeme wird eine ungefähre neue Hardwaregröße vorgegeben mit der Anzahl der CPUs und RAM etc. Das CPU – RAM Verhältnis ist für SAP HANA-basierten Systemen in den Public Clouds festgeschrieben. So werden in den Hyperscaler Cloud nur die Größen von virtuellen Maschinen angeboten, welche dem Verhältnis entsprechen und somit auch von der SAP freigegeben sind. Nachdem also ein erstes Sizing erstellt worden ist, muss in der Liste der unterstützten virtuellen Maschine eine passende rausgesucht werden.

Während ein Mapping von CPU und RAM auf die verfügbaren Größen einfach ist, bestehen jedoch weitere Abhängigkeiten bei dem Storage.

### Schritt 5: Verifizierung des Sizing

Der letzte Schritt im Sizing ist die Verifizierung der Zielgröße durch eine Testinstallation oder einer Sandboxinstallation inklusive einer Migration der Daten aus dem Altsystem. Nach dem Aufbau des Systems, sollte das System in jedem Fall einem Stresstest oder Performancetest unterzogen werden. Nur durch einen Test kann die tatsächliche Performance geprüft werden.

#### 2.2.1.3 HANA Sizing

Das Sizing von HANA-basierten Systemen erfolgt anhand der Größe des Hauptspeichers. Dies erfolgt in Gigabyte oder Terrabyte. Für alle HANA-basierten Systeme existieren zertifizierte Lösungen der Hyperscaler, als auch der traditionellen Hardwarehersteller. Eine detaillierte Auflistung der zertifizierten Lösungen existiert auf der Webseite der SAP unter folgendem Link (Zugriff am 20.12.2021):

<https://www.sap.com/dmc/exp/2014-09-02-hana-hardware/enEN/-/solutions?filters=v:deCertified>

Das Sizing von HANA-Systemen ist durch das Verhältnis von CPU und Hauptspeicher (RAM) bestimmt. Dieses Verhältnis ist für OLTP und OLAP Workloads unterschiedlich. **OLTP** steht hierbei für Online Transactional Processing, während **OLAP** für Online Analytical Processing steht. OLTP Systeme sind die normalen SAP S/4HANA Systeme und BW4HANA Systeme gehören beispielsweise zu den OLAP Systemen.

Viele der Hyperscaler setzen für SAP S/4HANA-Systeme Intel-basierte Hardware ein, welche effektiv auf Intel Xeon Prozessoren basiert. Bei einem Intel Xeon E7-8890 v4 würde für beide Workloads folgende Verhältnisse gelten:

- OLTP: 1 TB je Socket
- OLAP: 0.5 TB je Socket

Für OLAP Workloads wird demnach ein niedrigeres Speicherverhältnis angenommen. Der Hintergrund hierfür die die hohe Nutzung des Speichers im OLAP-Umfeld für die Zwischenspeicherung von Daten, z. B. bei größeren Ladevorgängen in das SAP BW hinein. Diese Ladevorgänge können oftmals viel temporären Speicher benötigen. Daher wird bei einem Sizing ein geringeres Verhältnis angenommen.

Wichtig ist, dass ein Sizing für OLTP und OLAP zu Beginn nur indikativ sein kann und durch andere Nutzungsmuster (Use Cases) später nochmal verfeinert werden muss. Dennoch bleibt für ein HANA Sizing der Hauptspeicher das wichtige Kriterium, um zu einem ersten Sizing zu kommen.

#### 2.2.1.4 SAP Application Performance Standard

Der SAP Application Performance Standard Benchmark ist ein Applikationsbenchmark, der innerhalb von SAP-Systemen ausgeführt wird. Er unterscheidet sich also von synthetischen Benchmarks, wie beispielsweise iobench, welche nur einen bestimmten Aspekt der Performance testen. Der SAPS Benchmark wurde von der SAP herausgegeben

und ist immer noch der Standardbenchmark, welcher zur Bemessung und auch Zertifizierung der Leistungsfähigkeit von Hardware genutzt wird.

Der SAPS besteht primär aus einfachen Transaktionen, welche innerhalb des SAP-Systems ausgeführt werden. Diese Transaktionen sind in den neuern SAP S/4HANA Systemen auch noch verfügbar und werden so auch noch weiter bestehen bleiben. Hierzu gehören beispielsweise das Anlegen eines Materialstammsatzes via MM01 oder auch das Anlegen von Stücklisten. Der Benchmark zielt darauf ab, so viele parallele Benutzer, wie möglich, die vordefinierte Schrittfolge abarbeiten zu lassen.

Während der SAPS Benchmark früher primär als Mittel zur Bemessung der Leistungsfähigkeit von Hardware und Zertifizierung von neuen Plattformen genutzt wurde, wird er heutzutage im Sizing noch benutzt, um einen einfacheren Vergleich zu erzielen. Bei Projekten, in denen aus einem traditionellen, on-premise, Rechenzentren in die Cloud migriert werden soll, ergibt sich die Schwierigkeit der Vergleichbarkeit der Leistungsfähigkeit der Hardware. Daher gilt der SAPS Wert der alten Hardware immer als guter Anhaltspunkt für das Sizing der neuen virtuellen Maschine in der Cloud.

### Ablösung von alter Hardware durch die Cloud

Ein Kunde war im Begriff seine beiden älteren Rechenzentren auf dem Gebiet der Firma abreißen zu lassen. Die darin befindliche Hardware war bereits abgeschrieben und schon seit mehreren Jahren nicht mehr erneuert worden. Darunter befanden sich auch viele SAP-Systeme, welche noch nicht auf S/4HANA migriert worden waren.

Nach einer initialen Analyse der Situation, wurde ein erster indikativer Zeitplan für einen Datacenter Exit erstellt und das Sizing begonnen. Da die alte Hardware bereits schon mehrere Jahre alt war, wurde zunächst basierend auf historischen Daten für jeden Servertypen eine Anzahl SAPS definiert. So gab es sehr leistungsfähige Hardware, als auch weniger leistungsfähige Hardware, welche teils nur 10'000 SAPS erreichte.

Danach wurde auf den SAPS basierend die neue Zielgröße der virtuellen Maschinen, wie z. B. DS3v2, ein Mapping erstellt und Maschine für Maschine die neue VM definiert. Da die Hardware in den Rechenzentren bereits schon sehr alt war, konnte der Kunde von der sehr leistungsfähigen Hardware in der Public Cloud profitieren und so mussten nur sehr kleine virtuelle Maschinen genutzt werden. Diese kleinen VMs haben einen sehr geringen Preis (selbst im Pay-as-you-go Modell). Somit konnte der Kunde die Erneuerung der Hardware aussparen und profitierte darüber hinaus noch von einer Einsparung bei den Betriebskosten. Der Return of Invest (ROI) der eigentlichen Migration war in weniger als einem Jahr erreicht. ◀

Der SAP Benchmark bildet als ergänzende Information einen Mehrwert, insbesondere bei älterer Hardware und Systemen. Hier kann ein Mapping auf Basis von CPU und Hauptspeicher zu einem zu großen Sizing und somit Mehrkosten führen. Durch die Nutzung des SAPS Wert kann das verhindert werden.

### 2.2.1.5 I/O Sizing

Der SAP Benchmark gilt auch als ein Indikator für die Berechnung des Storagedurchsatz. Die Leistung des Storage wird durch IOPS angegeben. Hierbei steht IOPS für «Input Output Per Second» und gibt die maximale Leistung/aktuellen Durchsatz der I/O Operationen an. Während die SAPS primär die Größe von CPU und Hauptspeicher zeigen, kann der SAPS-Wert jedoch keine Aussage hinsichtlich IO treffen. Dies ist jedoch sehr wichtig, da HANA Datenbanken spezifische Anforderungen an Storage haben.

Generell kann der SAPS-Wert wie folgt in IOPS umgerechnet werden:

**1 SAPS = 0.6 IOPS (Beispiel: 200'000 SAP = 120'000 IOPS)**

In Abhängigkeit des zukünftigen SAP S/4HANA Systems und der Nutzung des Systems, kann der Wert für IOPS noch steigen. Es gibt S/4HANA Systeme, welche mit einem Verhältnis von 1 SAPS = 0.9 IOPS erstellt worden sind.

Nun ist die SAP HANA-Datenbank zwar eine In-Memory-Datenbank, jedoch müssen auch bei einer HANA-Datenbank die Daten konsistent auf Storage gespeichert werden. Hierzu werden verschiedene Datenbereiche benötigt, welche auch voneinander getrennt sein sollten.

- **Redo Log Volumes:** Alle Änderungen an der HANA-Datenbank und den Daten wird in den Redo Log Volumes protokolliert, sodass auch nach einem Ausfall der Datenbank, alle Änderungen wiederhergestellt werden können (Blockgröße 4 KB bis zu 1 MB).
- **Data Volumes:** Die Daten der HANA-Datenbank werden auf den Data Volumes gespeichert. Eine Änderung der Daten erfolgt bei den Savepoints standardmäßig alle 5 min (Blockgröße von 4 KB bis zu 16 MB, maximal 64 MB bei Super-Blocks).
- **Backup Volume:** Alle erstellten Backupdaten werden in Blöcken (Größe bis zu 64 MB) auf dem Backup Volume gespeichert.

Die jeweiligen Volumes unterscheiden sich in den Anforderungen nach IO-Performance. Die Redo Log Volumes haben die höchsten Anforderungen, da alle Transaktionen so schnell wie möglich permanent auf den Storage geschrieben werden müssen. Die Data Volumes haben ebenso hohe Anforderungen nach Performance, da bei der Erstellung eines Savepoints keine weiteren Änderungen an den Daten durchgeführt werden. Die niedrigsten Anforderungen hat das Backup Volume. Hier werden die Daten asynchron von der Datenbank gespeichert.

Entsprechend den Anforderungen kann auch der Storagetyp innerhalb der Cloud selektiert werden:

- **Redo Log Volume:** Schnellster Storage
- **Data Volumes:** Schnellster Storage
- **Backup Volume:** Normaler Storage

Die Größe der jeweiligen Volumes ist von vielen Faktoren abhängig und kann aber mit folgenden «Daumenregeln» bestimmt werden:

- **Redo Log Volume:** Die Regel besagt, dass die Größe wie folgt gewählt wird: **Redo Log Volume = 0.5 x RAM**
- **Data Volumes:** Die einfachste Regel für das Sizing eines SAP S/4HANA Systems ist: **Größe des Data Volume = 1x RAM**
- **Backup Volume:** Die Regel ist, dass die Größe des wie folgt dimensioniert wird: **Backup Volumen = 1x Data Volume + 1x Log Volume**

Diese Berechnungen gelten nur als erste Anhaltspunkt beim initialen Sizing und auch nur als Berechnung der Größe der Volumes, aber nicht für deren Layout. Die Größe der Volumes kann sehr einfach durch die angebotenen Größen der unterschiedlichen Storageklassen adressiert werden. Allerdings darf die Performance nicht vernachlässigt werden. Je grösser eine Volume wird, desto mehr IOPS werden durch die Hyperscaler angeboten. Anhand von Microsoft Azure lässt sich das einfach darstellen.

- Die kleinste Premium Disk P1 mit 4 GiB kann bis zu 120 IOPS erzielen.
- Die größte Premium Disk P80 mit 32'767 GiB kann bis zu 20'000 IOPS erzielen

Daraus folgt, dass zur Erreichung von einer hohen IOPS Zahl, mehrere Disks miteinander im Stripping verbunden werden müssen.

#### Beispiel-Sizing I/O: SAP S/4HANA System in Azure

In dem folgenden Beispiel soll ein I/O Sizing für ein neues SAP S/4HANA System in der Microsoft Azure Cloud gezeigt werden.

Durch das initiale Sizing worden folgende Eckdaten ermittelt:

- **RAM Bedarf:** 3.5 TB
- **SAPS Bedarf:** 152'000
- **IOPS Bedarf:** 91'200

Basierend auf den Eckdaten kann der folgende Instanztyp in der Microsoft Azure Cloud genutzt werden:

- **Zielinstanz:** M128ms mit 128 vCPU und 3'892 GB RAM

Dieser Instanztyp wird ohne weiteren Storage angeboten, sodass für die Data Volumes und Redo Log Volumes noch der Storage dimensioniert werden muss.

- **Größe Redo Log Volume:** 2 TB
- **Größe Data Volume:** 4 TB
- **Größe Backup Volume:** 6 TB

Die Zielgröße des Storage ist jedoch nur eine Kenngröße. Die notwendigen IOPS werden durch ein Stripping der Disks erzielt:

- **Redo Log Volume:** 2x P30 mit je 1'024 GiB und je 5'000 IOPS = **10'000 IOPS**
- **Data Volume:** 2x P40 mit je 2'048 GiB und je 7'500 IOPS = **15'000 IOPS**
- **Backup Volume:** 3x E40 mit je 2'048 GiB und je 500 IOPS = **1'500 IOPS**

Somit ergibt sich ein Zielsizing mit folgenden Daten:

#### **Compute und RAM:**

- 1x M128ms mit 128 vCPU und 3'892 GB RAM

#### **Storage:**

- 2x P30
- 2x P40
- 3x E40 ◀

### **2.2.2 Kosten**

Bei der Verwendung von Public Cloud Technologien verfolgen viele Unternehmen insbesondere das Ziel der Kostenreduzierung der IT-Landschaft. Häufig wird durch die Migration der vollständigen IT-Landschaft in die Public Cloud eine sehr hohe Kosteneinsparung angestrebt, die in einer möglichst kurzen Zeitspanne realisiert werden soll. Die Hyperscaler-Anbieter von Public Cloud Technologien bieten zu diesem Zweck eine Vielzahl von Maßnahmen zur Kostenreduzierung an. Bei der Anwendung dieser Methoden muss weiterhin die Geschäftsfähigkeit sichergestellt, sowie der Einfluss auf bestehende Prozesse im Unternehmen berücksichtigt werden. In der Praxis empfiehlt sich daher ein schrittweises Vorgehen, da die Preisoptimierungsmaßnahmen in der Regel einen Einfluss auf die IT-Architektur haben und damit entsprechende Änderungen einhergehen.

Die Kosten einer IT-Infrastruktur setzen sich aus den Betriebskosten (Operational Expenditure, kurz OpEx) und den Investitionskosten (Capital Expenditure, kurz CapEx) zusammen. Unter den Betriebskosten werden laufende Kosten für die Aufrechterhaltung und Sicherstellung der Verfügbarkeit der IT-Landschaft verstanden. Dies sind beispielsweise die Stromkosten für den Betrieb eines eigenen On-Premise-Datenzentrums, aber auch Personalkosten, die für die Wartung und die Aufrechterhaltung des Betriebs der

IT-Infrastruktur anfallen. Die Investitionskosten beziehen sich auf einmalig anfallende Kosten, die im Rahmen des Aufbaus einer IT-Architektur entstehen. Das sind zum Beispiel die Beschaffungskosten für physische Komponenten, die in der Infrastruktur der IT-Landschaft benötigt werden, wie Server und Räumlichkeiten des Datenzentrums. Die gesamten Kosten, die für eine IT-Umgebung eines Unternehmens anfallen, werden als „Total Cost of Ownership“ (kurz TCO) bezeichnet.

Die Verwendung einer Public Cloud bringt den Vorteil mit sich, dass vor allem die Investitionskosten entfallen, da der Hyperscaler-Anbieter die physischen Komponenten zur Verfügung stellt. Dadurch werden die Verantwortlichkeiten für die einzelnen Ebenen der IT-Architektur zwischen dem Anbieter und dem Unternehmen aufgeteilt. Die Public Cloud bietet folgende Modelle an, in denen die Verantwortlichkeiten unterschiedlich aufgeteilt werden: Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), Software-as-a-Service (SaaS). Je nach Modell hat ein Unternehmen die Möglichkeit, die Betriebskosten durch geringere Verantwortlichkeiten zu senken. Eine Übersicht über die Verteilung der Elemente in den unterschiedlichen Modellen bietet die sogenannte „Shared Responsibility Matrix“ in der Tab. 2.1.

Neben dem Modell der verteilten Verantwortlichkeiten für einzelne Ebenen einer IT-Architektur bieten Hyperscaler-Anbieter Preismodelle an, die Unternehmen bei der Kostenreduzierung unterstützen. Im Zusammenhang mit Public Cloud Technologien hat sich die nutzungsbasierte Abrechnung („consumption-based pricing“) etabliert. Bei diesem Preismodell stellt der Hyperscaler-Anbieter dem Unternehmen lediglich die tatsächliche Nutzungsdauer der verwendeten Ressourcen und Instanzen in Rechnung. Wenn eine virtuelle Maschine beispielsweise lediglich zu den regulären Geschäftszeiten

**Tab. 2.1** Trennung der Verantwortlichkeiten

Lokal (On-Premise)	Infrastruktur as-a-Service (IaaS)	Plattform as-a-Service (PaaS)	Software as-a-Service (SaaS)
Daten und Zugriff	Daten und Zugriff	Daten und Zugriff	Daten und Zugriff
Anwendungen	Anwendungen	Anwendungen	Anwendungen
Runtime	Runtime	Runtime	Runtime
Betriebssystem	Betriebssystem	Betriebssystem	Betriebssystem
Virtueller Computer	Virtueller Computer	Virtueller Computer	Virtueller Computer
Compute	Compute	Compute	Compute
Netzwerk	Netzwerk	Netzwerk	Netzwerk
Speicher	Speicher	Speicher	Speicher
Verantwortlichkeit des Konsumenten		Verantwortlichkeit des Cloud-Hyperscaler-Anbieters	
Verantwortlichkeiten des Kunden und das Hyperscalers			

genutzt wird und zu den übrigen Tageszeiten ausgeschaltet bleibt, wird ausschließlich der Zeitraum der Geschäftszeiten berechnet, in dem die VM eingeschaltet ist. Dieses Modell wird auch als „Pay-as-you-Go“ (kurz PAYG) bezeichnet. Sobald eine Instanz nicht kontinuierlich ( $24 \times 7$ ) benötigt wird, besteht die Möglichkeit, das nutzungsbasierte Preismodell anzuwenden.

Da in einer On-Premise-Umgebung häufig alle Systeme und Anwendungen durchgehend verfügbar sind, steht die Entscheidung über die Anwendbarkeit mit dem PAYG-Modell während der Migration der IT-Architektur in die Public Cloud in der Praxis vor einigen Herausforderungen. Da mit der Umstellung einer kontinuierlichen Verfügbarkeit auf eine beschränkte Verfügbarkeit ausgewählter Applikationen gewisser Anpassungsbedarf bei internen Prozessen einhergeht, bedarf es einer selektiven Analyse der Nutzung, bevor eine Anpassung erfolgt. Ein Beispiel für eine solche Prozessanpassung stellt die Arbeitsweise global verteilter Entwicklungsteams dar. Aufgrund der unterschiedlichen Zeitzonen muss die Verfügbarkeit eines Development-Systems auch außerhalb der einem spezifischen Land geltenden Geschäftszeiten sichergestellt werden. Da sich die tatsächliche Nutzungsdauer von Komponenten einer IT-Architektur erst im laufenden Betrieb ermitteln lässt, eignet es sich, das nutzungsbasierte Preismodell schrittweise nach der Migration in die Public Cloud zu etablieren. Alle Hyperscaler-Anbieter stellen diverse, voll integrierbare Monitoring-Möglichkeiten zur Nutzungsüberwachung bereit. Zudem kann die Nutzungsdauer einer Instanz auch nachträglich zum Deployment in der Public Cloud angepasst werden.

Im Gegensatz zur nutzungsbasierten Abrechnung unterstützen die Hyperscaler-Anbieter eine langfristige Bindung eines Unternehmens an die angebotenen Public Cloud Ressourcen. Das Abrechnungsmodell wird als Reservierung der Cloud-Instanzen („reserved instances“) angeboten. Je nach Dauer gewährt der Hyperscaler-Anbieter einen Rabatt, der bis zu 70 % betragen kann. Durch das Commitment legt sich das Unternehmen verbindlich darauf fest, die jeweilige Ressource oder den jeweiligen Service für eine bestimmte Mindestdauer zu nutzen, erhält im Gegenzug dazu jedoch einen Rabatt von dem Hyperscaler-Anbieter.

In einem praktischen Unternehmensumfeld ist die nutzungsbasierte Abrechnung in der Regel nicht auf alle Komponenten in einer IT-Architektur anwendbar. Insbesondere an die geschäftskritischen Applikationen wird üblicherweise die Anforderung einer  $24 \times 7$ -Verfügbarkeit gestellt, um essentielle Prozesse nicht zu unterbrechen und damit gegebenenfalls die Geschäftsfähigkeit zu gefährden. Für produktive Workloads eignet sich daher häufig das Modell der reservierten Instanzen, da diese zudem in der Regel längerfristig genutzt werden.

Für nicht-produktive Komponenten wird die Entscheidung folglich zwischen dem Pay-as-you-Go Modell und den reservierten Instanzen getroffen. Um diese Entscheidung möglichst kosteneffizient vorzunehmen, empfiehlt es sich, zunächst die minimal erforderliche Verfügbarkeit jeder betroffenen Instanz zu ermitteln. Mithilfe der Kalkulatoren zur Preisberechnung, die jeder Hyperscaler-Anbieter zur Verfügung stellt, lassen sich anhand des Instanztyps der monatliche Nutzungspreis für

das nutzungsbasierte Modell und für das Commitment-Modell berechnen. Als Entscheidungsgrundlage für das Abrechnungsmodell dient der Break-Even-Betrag des jeweiligen Instanztyps. Dieser kann ebenfalls mithilfe der Preiskalkulatoren ermittelt werden, indem eine Gegenüberstellung beider Abrechnungsmodelle erfolgt und eine Annäherung der Nutzungsdauer vorgenommen wird. Die Beträge der Nutzungsdauer werden abwechselnd aneinander angenähert, bis der monatliche Preis für beide Abrechnungsmodelle identisch ist. Die in diesem Fall geltende Nutzungsdauer stellt den Break-Even für den jeweiligen Instanztyp dar. Der Break-Even stellt den Wert der Nutzungsdauer dar, bei dem sowohl für das nutzungsbasierte Preismodell als auch für das Commitment identische Kosten anfallen. Ab der Break-Even-Nutzungsdauer eignet sich daher aus wirtschaftlicher Sicht das Modell der reservierten Instanzen, da die Workloads bei gleichem Preis im Gegensatz zum PAYG länger verfügbar sind. Das nachfolgende Diagramm stellt beispielhaft den Vergleich zwischen dem Verhältnis der Nutzungsdauer und den monatlichen entstehenden Kosten in Bezug zu dem jeweiligen Preismodell anhand einer virtuellen Maschine des Typs D4 v3 (4 vCPUs, 16 GB RAM, 100 GB temporärer Speicher) in der Microsoft Azure Cloud dar.

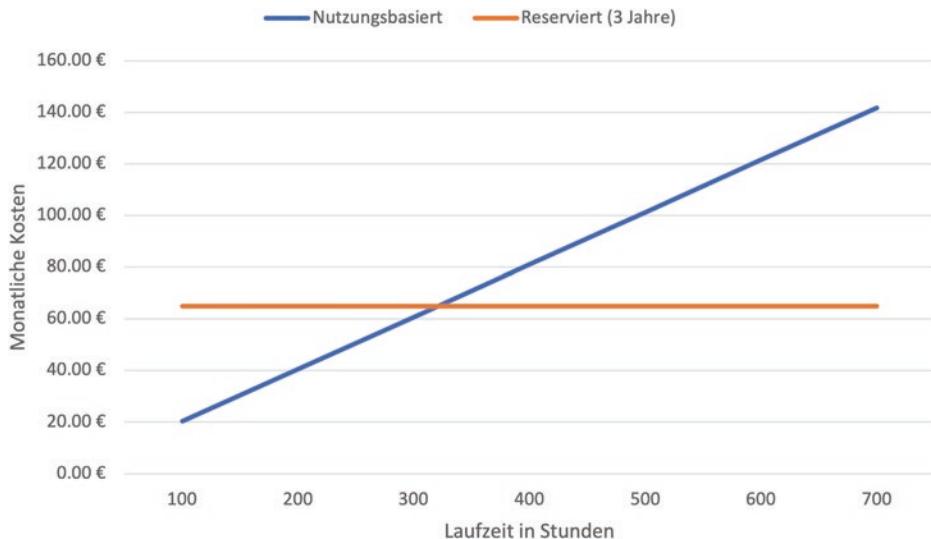
Für die Ermittlung des Break-Even-Werts wurden in dem Preisrechner von Microsoft Azure folgende Annahmen als Grundlage für die Berechnung verwendet:

- Region: West Europe
- Betriebssystem: Linux (Ubuntu)
- Tarif: Standard
- Instanz: D4 v3 (4 vCPUs, 16 GB RAM, 100 GB temporärer Speicher)
- Virtuelle Maschinen: 1
- Für den Vergleich zwischen nutzungsbasierter Bezahlung und reservierten Instanzen wurde für das langfristige Commitment eine Nutzungsdauer von 3 Jahren hinzugezogen (Abb. 2.2).

In dem Diagramm zeigt sich, dass die Kosten für die nutzungsbasierte Abrechnung ab einer operativen, monatlichen Nutzungsdauer von mehr als 320 h höher sind als bei den reservierten Instanzen. Dieser Wert entspricht dem Break-Even.

Mithilfe des Break-Even der Nutzungsdauer einer Instanz wird der Entscheidungsprozess initiiert. Der ermittelte Break-Even wird mit der tatsächlichen Nutzungsdauer verglichen. Bei diesem Vergleich werden folgende Fälle unterschieden, die als Entscheidungshilfe dienen.

1. **Break-Even ist geringer als die tatsächliche Nutzungsdauer:** In diesem Fall ist das Modell der reservierten Instanzen wirtschaftlich effizienter, da die Instanz oder der Service für einen geringeren Preis im Gegensatz zum nutzungsbasierten Abrechnungsmodell kontinuierlich ( $24*7$ ) betrieben werden kann. Zusammengefasst erhält der Kunde des Hyperscaler-Anbieters einen höheren Nutzungsgrad zu einem geringeren Preis.



**Abb. 2.2** Break-Even

2. **Break-Even entspricht der tatsächlichen Nutzungsdauer:** Die Beurteilung dieses Szenarios fällt analog zum ersten Szenario aus. Bei einer Übereinstimmung des Break-Even mit der Nutzungsdauer steht dem Anwender eine längere Nutzungsdauer der Instanz zu einem identischen Preis wie bei der nutzungsbasierten Abrechnung zur Verfügung.
3. **Break-Even ist höher als die tatsächliche Nutzungsdauer:** Wenn der Break-Even signifikant höher als die reale Nutzungsdauer der Instanz ist, eignet sich die Verwendung des nutzungsbasierten Preismodells. Eine pauschale Entscheidungsempfehlung lässt sich für diesen Fall nicht treffen, da insbesondere die Höhe der Differenz zwischen den beiden Beträgen ausschlaggebend ist. Wenn der Break-Even beispielsweise lediglich wenige Cent-Beträge höher als der Vergleichswert ist, wäre trotz der minimal höheren Kosten das Modell der reservierten Instanzen eine effizientere Option, da die Instanz in diesem Fall für einen minimalen Kostenanstieg kontinuierlich zur Verfügung stehen würde. Daher ist es empfehlenswert, für jede Instanz eine individuelle Entscheidung zu treffen.

Bezogen auf den beispielhaften Vergleich der beiden Abrechnungsmodelle hängt die Entscheidung von der geplanten, monatlichen Nutzungsdauer der virtuellen Maschine ab. Wenn diese höher als 320 h sein sollte, empfiehlt es sich aufgrund von wirtschaftlichen Gründen und der Kosteneffizienz das Modell der reservierten Instanzen zu nutzen. An dieser Stelle muss jedoch berücksichtigt werden, dass es sich um ein langfristiges Commitment handelt.

Neben den beschriebenen Preismodellen zur Reduzierung der Kosten bietet Amazon Web Services einen Mengenrabatt für ausgewählte Services an. Bei einem Nutzungsanstieg und einem damit verbundenen, höheren Speicherbedarf sinken die Kosten, wenn mehr Speicher konsumiert wird.

Eine weitere Methode zur Optimierung ist das Sizing der Cloud Infrastruktur, das bereits im vorangegangenen Kapitel ausführlicher beschrieben wurde. Mithilfe von Re-Sizing und Right Sizing werden die in der Architektur verwendeten Komponenten auf ihre tatsächliche Nutzung überprüft. Das Ziel ist die Optimierung der Ressourcen innerhalb der IT-Architektur, um beispielsweise ungenutzte Rechen- oder Speicherkapazitäten zu identifizieren. Die Anpassung der verwendeten Ressourcen an die tatsächlich genutzten Kapazitäten kann die Reduzierung der Kosten in einem Unternehmen unterstützen.

#### Beispiel-Kalkulation: SAP S/4HANA System in Azure

In dem folgenden Beispiel wird die Ermittlung des Break-Even anhand eines produktiven SAP S/4HANA Systems in der Microsoft Azure Public Cloud durchgeführt. Im ersten Schritt werden die Anforderungen und Voraussetzung definiert, die erforderlich sind, um ein S/4HANA System in der Azure Cloud zu betreiben.

- In dem produktiven SAP S/4HANA System werden die zentralen Geschäftsprozesse durchgeführt.
- Das System wird folglich kontinuierlich benötigt, um die Geschäftsfähigkeit aufrechtzuerhalten.
  1. Im ersten Berechnungsansatz wird keine Hochverfügbarkeit oder Disaster Recovery berücksichtigt.
  2. Der zweite Berechnungsansatz sieht eine Hochverfügbarkeit über die Bereitstellung einer zusätzlichen Verfügbarkeitszone vor.
- Initial soll lediglich der Aufbau der zum Betrieb erforderlichen Ressourcen berücksichtigt werden. Daher werden in der Kalkulation keine zusätzlichen Azure Native Services oder sonstige Integrationspunkte hinzugezogen.
- Das SAP S/4HANA System setzt sich aus der ERP-Applikation mit der dazugehörigen HANA-Datenbank, sowie der Fiori-Umgebung zusammen. Um eine hohe Performance und das Prinzip der Modularität zu berücksichtigen, werden beide Umgebungen in getrennten VM-Instanzen aufgebaut.

In dem Azure Preisrechner wurden folgende Eigenschaften zur Berechnung der Gesamtkosten angenommen.

- Region: West Europe
- Die Lizenzen für die Betriebssysteme sind bereits vorhanden, sodass der Azure-Hybridvorteil angewendet werden kann (Bring-Your-Own-License).

- Die Instanzen werden auf 3 Jahre reserviert.
- Als Managed Disk der virtuellen Maschinen wurde für jede Instanz eine SSD Premium Edition angenommen.

Umgebung	Workload	Type	Betriebs-system	Speicher	Instanzen	Kosten
ERP	SAP	D16s v3	Windows	650	1	384,63 €
	DB	M64ms	Linux	4096	1	2.945,36 €
Fiori	SAP	D8s v3	Windows	250	1	164,89 €
	DB	M32ts	Linux	896	1	648,87 €
<b>Kostenkalkulation für vier Workloads ohne HA</b>						

### Kosten für die Standardarchitektur:

Die monatlichen Gesamtkosten liegen in diesem Ansatz bei **4‘143,75 €** und jährlich betragen die Gesamtkosten **49‘725,00 €**. Durch das Commitment für die Reservierung der Instanzen über 3 Jahre beträgt die vollständige Summe für diesen Zeitraum bei **149‘175,00 €**.

Umgebung	Workload	Type	Betriebs-system	Speicher	Instanzen	Kosten
ERP	SAP	D16s v3	Windows	650	1	384,63 €
	DB	M64ms	Linux	4096	2	5.890,71 €
Fiori	SAP	D8s v3	Windows	250	1	164,89 €
	DB	M32ts	Linux	896	2	1.369,73 €
Azure Site Recovery					2	42,17 €
<b>Kostenkalkulation für vier Workloads mit HA</b>						

### Kosten für die Architektur mit Hochverfügbarkeit:

Die monatlichen Gesamtkosten liegen in diesem Ansatz bei **7‘852,13 €** und jährlich betragen die Gesamtkosten **94‘225,56 €**. Durch das Commitment für die Reservierung der Instanzen über 3 Jahre beträgt die vollständige Summe für diesen Zeitraum bei **282‘676,68 €**. ◀

## 2.2.3 Hochverfügbarkeit

SAP S/4HANA-Systeme müssen idealerweise 24 h und 7 Tage die Woche verfügbar sein. Das trifft auch zu, wenn es Probleme im Rechenzentrum oder in den SAP-Systemen gibt. Um die Systeme hochverfügbar zu halten, existieren mehrere Wege, welche hier beschrieben werden.

### 2.2.3.1 Hochverfügbarkeit allgemein

SAP S/4HANA Systeme bilden das Rückgrat des Geschäfts der Unternehmen und sind ein integraler Bestandteil aller Geschäftsprozesse. Daher werden SAP S/4HANA Systeme in der Architektur so ausgelegt, dass sie möglichst hochverfügbar sind und auch bei Ausfall einer Komponente weiterhin verfügbar bleiben und die Services bereitstellen.

Die Verfügbarkeit eines SAP-Systems wird in Prozent gemessen und die Architektur eines Systems richtet sich nach der Wichtigkeit/Kritikalität des Systems. So sind Systeme, wie Sandbox oder Testsystem nicht sehr kritisch und benötigen demnach keine Hochverfügbarkeit. Produktionssysteme sind jedoch sehr kritisch und werden über eine Hochverfügbarkeit vor dem Ausfall einer Komponente abgesichert. Die Unternehmen müssen bei der Architektur die Kritikalität der Systeme bewerten und die technischen Architekturen daran ausrichten.

► **Verfügbarkeit** Die Verfügbarkeit eines SAP S/4HANA Systems wird in Prozent angegeben. Als Berechnungsbasis gilt die maximale theoretische, rechnerische Verfügbarkeit des SAP-Systems im Monat in Minuten. Die maximale Verfügbarkeit beträgt 31 Tage \* 24 h \* 60 min = 44.460 min.

$$\text{Verfügbarkeir} = \frac{\text{Max. Verfügbarkeit} - \text{Nichtverfügbarkeit}}{\text{Max. Verfügbarkeit}} * 100$$

Anhand der Formel kann bei einem Ausfall eines SAP-Systems, von beispielsweise 700 min in einem Monat, von einer Verfügbarkeit von 98.43 % gesprochen werden:

$$\text{Verfügbarkeit} = \frac{44.460 - 700}{44.460} * 100 = 98.43 \%$$

Die Gesamtverfügbarkeit eines SAP-Systems ergibt sich durch die Kombination der Verfügbarkeiten der einzelnen Komponenten. In den Hyperscaler Clouds wird den Kunden z. B. eine mögliche Verfügbarkeit von bis zu 99.99 % für eine virtuelle Maschine geboten. Hierzu ist aber wichtig zu verstehen, dass die virtuelle Maschine nur eine Komponente ist. Es ist wichtig, die anderen Komponenten eines SAP-Systems für die Gesamtverfügbarkeit zu beachten. Hierzu gehören auch das Betriebssystem, das Netzwerk, der Storage und das Filesystem, als auch die einzelnen Komponenten des SAP-Systems. Wenn diese Verfügbarkeiten alle miteinander kombiniert werden, ergibt sich eine andere Gesamtverfügbarkeit eines Systems. Untenstehend ist die Gesamtverfügbarkeit als Ergebnis aus den Verfügbarkeiten der anderen Komponenten errechnet:

*Gesamtverfügbarkeit*

$$\begin{aligned} &= 99.99 \% (\text{VM}) * 99.9 \% (\text{Storage}) * 100 \% (\text{OS}) * 99.9 \% (\text{Netzwerk}) * 99.9 \% (\text{SAP}) \\ &= 99.69 \% \end{aligned}$$

Das obige Rechenbeispiel zeigt, wie die Gesamtverfügbarkeit des SAP-Systems sinkt, obwohl die Verfügbarkeit der einzelnen Komponenten sehr hoch ist. Dies gilt es bei der

Verfügbarkeitsdiskussion von SAP-Systemen zu beachten. Selbst wenn nur eine kleine Komponente in dem Gesamtverbund des S/4HANA Systems nicht verfügbar ist, kann es die Verfügbarkeit direkt beeinträchtigen. Um diese Möglichkeiten auszuschließen, können verschiedene Hochverfügbarkeitslösungen genutzt werden.

### 2.2.3.2 Hochverfügbarkeitsklassen

Die Wichtigkeit von S/4HANA Systemen für die Unternehmen lässt sich durch die Kritikalität der Systeme und einem möglichen Einfluss auf die Geschäftsprozesse ableiten. Es existieren Systeme, welche nicht wichtig für den Fortbestand der Unternehmen sind. Darüber hinaus existieren aber auch S/4HANA Systeme, welche existenziell für die Fortführung des Geschäfts sind.

Unternehmen kategorisieren die S/4HANA Systeme nach der Kritikalität und weisen jeder Kritikalität eine entsprechende notwendige Verfügbarkeit zu. Diese Verfügbarkeit sollte danach in eine entsprechende Architektur bei den SAP-Systemen übersetzt werden. Die Architektur muss die Verfügbarkeit der SAP-Systeme sicherstellen können und vor Ausfällen schützen. Eine Kategorisierung kann sehr unterschiedlich sein und kann nach z. B. Metallen benannt sein:

- **Gold-Klasse:** Systeme der höchsten Kritikalitätsstufe, welche durch einen Ausfall einen großen Einfluss auf die Geschäftsprozesse hat und wodurch ein Unternehmen nicht mehr arbeitsfähig wäre. Hierbei handelt es sich oft um die produktiven S/4HANA Systeme.
- **Silber-Klasse:** Systeme der mittleren Kritikalitätsstufe, welche durch einen Ausfall einen beschränkten Einfluss auf die Geschäftsprozesse des Unternehmens haben (z. B. auf nur einen Bereich des Unternehmens). Hierbei handelt es sich oft um die Qualitätssicherungssysteme oder aber auch produktive Business Warehouse-Systeme.
- **Bronze-Klasse:** Systeme der niedrigsten Kritikalitätsstufe, welche durch einen Ausfall einen sehr geringen Einfluss haben (z. B. eine sehr kleine Gruppe von Mitarbeitern betroffen). Hierbei handelt es sich oft um die Entwicklungs- oder auch einige Qualitätssicherungssysteme.

Neben der Benennung in Anlehnung an Metalle, kann auch einfach mit Nummern gearbeitet werden, welche auch die Kritikalität widerspiegeln. Die Tab. 2.2 gibt hierzu einen Überblick.

**Tab. 2.2** Kritikalität und Verfügbarkeiten von S/4HANA Systemen

Kategorie	Gold	Silber	Bronze
Verfügbarkeit	99,5 %	98 %	95 %
Maximaler Ausfall	4 h	14 h	37 h

Abfallende Verfügbarkeiten

Die Verfügbarkeitsklassen müssen durch eine entsprechende Architektur realisiert werden. Hierzu gibt es für die S/4HANA Systeme in den Hyperscaler Clouds unterschiedliche Möglichkeiten:

- **Cloud nativ:** Da die Hyperscaler Clouds alle auf Virtualisierung setzen, profitieren die S/4HANA Systeme auch von einer höheren Verfügbarkeit, welche durch die Verfügbarkeit der virtuellen Maschinen in der Cloud gegeben ist. Beispiel: Microsoft Azure beziffert die Zielverfügbarkeit von den einfachsten virtuellen Maschinen mit bis zu 95 %. Fällt die Verfügbarkeit niedriger aus, können bereits Service Credits an den Kunden zurückgehen.
- **Verfügbarkeitszonen:** In allen großen Hyperscalern existiert die Möglichkeit, eine virtuelle Maschine zusätzlich abzusichern. Diese zusätzliche Absicherung erfolgt durch Verfügbarkeitszonen (Availability Zones, Availability Sets, etc.) und kann die Verfügbarkeit auf bis zu 99.99 % steigern.
- **Hochverfügbarkeitscluster:** Die Verfügbarkeit eines SAP-Systems ergibt sich durch die Verfügbarkeit der Komponenten. Da die Hyperscaler Clouds nur die virtuellen Maschinen absichern, werden Hochverfügbarkeitscluster zur Sicherstellung der anderen Verfügbarkeiten genutzt.

Durch die Nutzung von Hyperscaler Clouds kann eine grundsätzliche Verfügbarkeit von S/4HANA Systemen bereits sichergestellt werden.

### 2.2.3.3 Hochverfügbarkeit der S/4 Architektur

Die SAP hat in den vergangenen Jahren kontinuierlich daran gearbeitet, die sogenannten Single Point of Failures in den SAP-Systemen durch Redundanzen zu beseitigen. Die Architektur eines SAP-Systems besteht im Wesentlichen aus den folgenden Komponenten, welche auch in der Abbildung zu sehen sind:

- **SAP Central Services:** Die SAP Central Services (SCS) stellen die wichtigsten zentralen Services des SAP-Systems bereit. Dazu gehören der Enqueue Server und der Message Server. Ohne diese beiden Komponenten können die Benutzer zwar noch weiter auf dem SAP-System arbeiten, aber nur noch mit Einschränkungen.
- **Applikationsserver:** Die Applikationsserver eines S/4HANA Systems bearbeiten die Anfragen der Benutzer und halten die disp+work Prozesse eines SAP-Systems. Da ein S/4 System mehrere Applikationsserver besitzen kann, können diese Server auch ausfallen, ohne einen Einfluss auf die Gesamtverfügbarkeit eines SAP-Systems zu haben.
- **HANA Datenbank:** Die HANA Datenbank speichert die Geschäftsdaten des SAP-Systems und ist daher als Single Point of Failure zu sehen. Wenn die Datenbank nicht verfügbar ist, ist das gesamte SAP-System nicht verfügbar.

- **File Shares:** Üblicherweise benötigen SAP S/4HANA Systeme die Dateien aus dem globalen Mountpoint und die Dateien aus dem Transportverzeichnis zugänglich für alle Komponenten des SAP-Systems und der Systemlinie (also Entwicklungs-, Qualitätssicherungs- und Produktionssystem).

Die Abb. 2.3 zeigt die Komponenten eines S/4HANA Systems, welche idealerweise hochverfügbar gehalten werden sollten. Für jede Komponente wird eine virtuelle Maschine mit angebundenem Storage genutzt. Die Abbildung zeigt, wie die einzelnen Komponenten gegen einen Ausfall abgesichert werden:

- **SAP Central Services:** Die zentralen Services des S/4HANA Systems werden durch ein Hochverfügbarkeitscluster abgesichert. Hierzu kann z. B. ein Linux Pacemaker genutzt werden, welcher die Services beobachtet (Monitor) und die Prozesse im Bedarfsfall auf dem zweiten Knoten neustartet kann.
- **Applikationsserver:** Die Applikationsserver werden im Normalfall nicht weiter gegen einen Ausfall abgesichert. Da es meist mehrere Applikationsserver in einem S/4HANA System gibt, können die Endbenutzer auch bei dem Ausfall eines Servers, weiterhin arbeiten.
- **HANA Datenbank:** Die HANA Datenbanken der S/4HANA Systeme werden durch eine HANA System Replication miteinander verbunden. Auch hier ist ein Hochverfügbarkeitscluster notwendig, um die HANA Services (wie z. B. den hdbindexserver) zu überwachen.

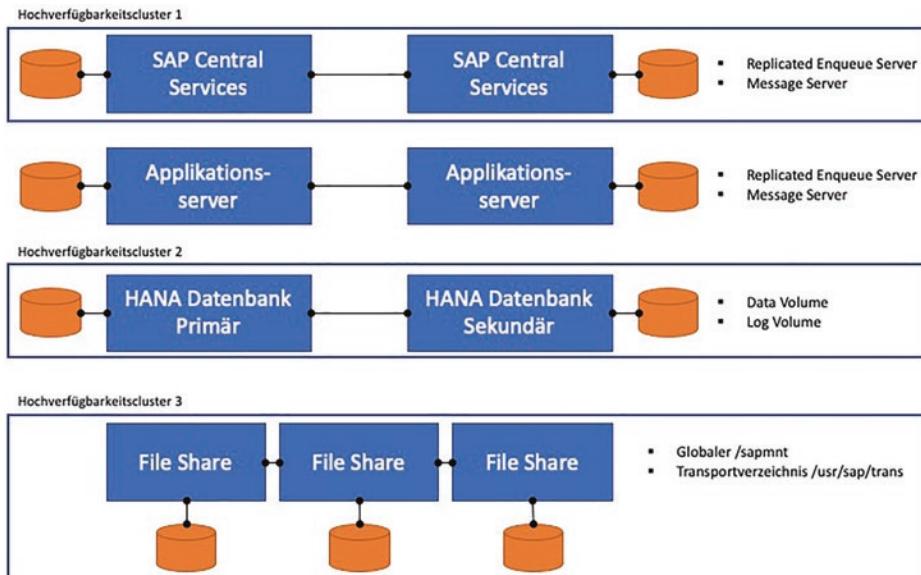


Abb. 2.3 S/4HANA Systemarchitektur

- **File Share:** Die Fileshares werden üblicherweise auch hochverfügbar gehalten. Dies kann entweder über die nativen NFS Mittel der Clouds passieren (z. B. Azure Share) oder kann durch andere Clustertechnologien erfolgen (z. B. Gluster).

Bei sehr kritische S/4HANA Systemen kann also die Architektur, welche zur Erreichung einer hohen Verfügbarkeit notwendig ist, sehr komplex werden.

#### Differenzierte Klassifizierung der SAP-Systeme

Betrachtet man sich die beispielhafte Architektur für ein hochverfügbares SAP S/4HANA Systeme, so entsteht natürlich die Frage nach den Kosten solch einer Architektur. Daher ist es wichtig, dass solche Architekturen nur für die kritischen Systeme genutzt werden, da sonst die Kostenfalle droht und ein sehr hoher Rechnungsbetrag vom Cloud-Anbieter anstünde.

Damit dies nicht passiert, hat ein beispielhafter Kunde seine SAP-Systemlandschaft wie folgt untergliedert:

- **Höchste Kritikalität:** Nur zwei der zehn produktiven SAP-Systeme wurden als so kritisch eingestuft, dass sie eine hochverfügbare Architektur erhielten. Dies waren zwei ERP-Systeme, welche für zwei Divisionen des Kunden betrieben worden. Diese beiden Systeme wurden auf Basis der gezeigten Architektur in Abb. 2.3 aufgebaut. Diese beiden Systeme wurden auch mit einem aktiven Desaster Recovery ausgestattet, welche im folgenden Teilkapitel beschrieben wird.
- **Hohe Kritikalität:** In diese Kategorie fielen alle anderen produktiven SAP-Systeme. Darunter waren beispielsweise die Business Warehouse-Systeme, die PI/PO-Systeme als auch sonstige Systeme (wie Fiori Gateway, Solution Manager und GRC). Bei einem Ausfall dieser Systeme wäre der Kunde zwar beeinträchtigt, aber das Geschäft des Kunden würde weiterlaufen können. Daher hat sich der Kunde gegen eine komplexe Architektur hierfür entschieden.
- **Geringer Kritikalität:** Alle anderen Systeme (Qualitätssicherungs-, Entwicklungs- und Sandboxsysteme) erhielten eine sehr einfache Architektur ohne weitere Absicherungen oder einer komplexen Architektur.

Der Kunde hat mit den Fachabteilungen zusammen die Kritikalität der Systeme festgelegt. Dies ist und kann keine Entscheidung der IT-Abteilung alleine sein. ◀

### 2.2.4 Desaster Recovery

#### 2.2.4.1 Desaster allgemein

Während die Hochverfügbarkeit von S/4HANA Systemen vor dem Ausfall einer Komponente schützt, kann sie jedoch nicht vor dem Ausfall eines gesamten

Rechenzentrums und somit des Gesamtsystems schützen. Der Ausfall eines Rechenzentrums kann durch verschiedene Faktoren ausgelöst werden:

- **Flugzeugabsturz:** Der berühmte Absturz eines Flugzeugs auf das Rechenzentrum kann die Katastrophe auslösen und zu einem Totalausfall des Rechenzentrums führen. Da Flugzeugabstürze jedoch sehr unwahrscheinlich sind, kann dieser Grund eher als theoretischer Grund herangezogen werden.
- **Brand:** Tatsächlich hat es in der Vergangenheit Brände in Rechenzentren oder sogar außerhalb der Rechenzentren gegeben, welche zu einem Ausfall geführt haben. Das jüngste Beispiel eines Rechenzentrumsbetreibers in Frankreich [1] zeigt das mögliche Ausmaß eines Brandes. Hierbei gingen Unmengen von Daten der Kunden verloren.
- **Ausfall Strom:** Der Ausfall der Stromversorgung wird bei Rechenzentren durch unterbrechungsfreie Stromversorgungen (USVs) oder aber durch Notstromaggregate abgesichert. Dennoch kann es zu einem Ausfall kommen, wenn beispielsweise Wartungsarbeiten falsch ausgeführt werden oder aber es ein Problem bei den USVs gibt.
- **Ausfall Kühlung:** Ein Ausfall der zentralen Kühlung ist selten möglich, kann aber zu einem sukzessiven Abschalten der Server aufgrund der Überhitzung führen.
- **Ausfall Konnektivität:** Rechenzentren sind durch mehrere, redundante Verbindungen an die Außenwelt angebunden. Dennoch kann das berühmte Beispiel des Baggers (also Erdarbeiten außerhalb des Rechenzentrums) zu einem Verlust der Verbindung führen und somit die SAP-Systeme beeinträchtigen.
- **Blitzeinschlag:** Ein Blitzeinschlag in einem Rechenzentrum sollte bei den großen Cloud-Anbietern in keinem Fall zu einem Ausfall führen, jedoch bei kleineren Cloud-Anbietern.
- **Terroristischer Akt:** Ein Bombenanschlag auf ein Rechenzentrum kann zu einem Ausfall führen, muss jedoch als sehr unwahrscheinlich eingestuft werden. Die Rechenzentren der Cloud-Anbieter sind mehrfach abgesichert und überwacht.
- **Hackerangriff:** Ein Hackerangriff ist ein sehr wahrscheinlicher Fall, der jedoch nicht das gesamte Rechenzentrum beeinträchtigen kann.
- **Notfallpatches:** Das notfallartige Patchen kann zwar nicht als Desaster gesehen werden, aber durchaus als Situation, welche zu einem Ausfall des SAP-Systems führen kann. Diese Situation kam beispielsweise nach der Entdeckung der Heartbleed Vulnerability vor.
- **Erdbeben:** Gleichwohl die Eintrittswahrscheinlichkeit eher gering ist, gelten Erdbeben als mögliches Szenario für den Ausfall eines Rechenzentrums.

Bei einem Desaster müssen Kunden entweder mit dem Ausfall des SAP S/4HANA Systems rechnen oder es gibt Vorkehrungen, um das SAP-System verfügbar zu halten. Dies ist das Ziel eines Disaster Recovery.

### 2.2.4.2 Wichtige Kenngrößen in einem Desasterfall

Der Ausfall eines Rechenzentrums ist zwar ein unwahrscheinlicher Fall, kann aber aufgrund der im vorherigen Teilkapitel genannten Gründe dennoch auftreten. Vor dem Ausfall eines Rechenzentrums schützen die Disaster Recovery Mechanismen, welche in unterschiedlicher Ausprägung implementiert werden können. Die wichtigsten zwei Kenngrößen für die Implementierung der Mechanismen sind RTO und RPO. RTO bezeichnet die Recovery Time Objective und zeichnet aus, in welchem Zeitraum nach einem Desaster ein System wieder verfügbar sein sollte. Die Recovery Point Objective beschreibt den maximalen Datenverlust nach einem Desaster. Anhand von RTO und RPO werden verschiedene Mechanismen, in Abhängigkeit der Kritikalität der SAP S/4HANA Systeme, implementiert.

► **Recovery Time Objective** Bei einem Ausfall eines Rechenzentrums oder einer ganzen Region (z. B. alle Rechenzentren in Amsterdam) spricht man von einem Desaster. Sofern SAP-Systeme für diesen Fall abgesichert sind, erfolgt automatisch ein Schwenk in die zweite Region. Die Zeit, die benötigt wird, um das SAP-System in der zweiten Region für die Endbenutzer verfügbar zu machen, wird als die **RTO – Recovery Time Objective** definiert.

► **Recovery Point Objective** Bei einem Ausfall eines Rechenzentrums oder einer ganzen Region (z. B. alle Rechenzentren in Amsterdam) spricht man von einem Desaster. SAP-Systeme sind grundsätzlich gegen ein Desaster abgesichert, jedoch existieren Systeme, bei denen kein Unterschied in den Daten zwischen der Primärregion und der Sekundärregion (oder Primärrechenzentrum und Sekundärrechenzentrum) auftreten darf. Die **RPO – Recovery Point Objective** bezeichnet den maximalen Verlust der Daten zwischen beiden Regionen und wird in Minuten/Stunden ausgedrückt.

Da nicht alle SAP S/4HANA Systeme nach einem Desaster sofort wieder einsatzbereit und verfügbar sein müssen, wird die Wichtigkeit der SAP-Systeme mit in Betracht gezogen, um zu entscheiden, welche Mechanismen implementiert werden.

Tab. 2.3 zeigt die unterschiedlichen RTO und RPO von SAP-Systemen. Für höchstkritische Systeme (Kategorie Gold) wird somit ein maximaler Datenverlust von 15 min angenommen und eine maximale Wiederherstellungszeit von 30 min. Somit müssen die kritischen Produktionssysteme nach einem Desaster unmittelbar wieder zur Verfügung stehen. Weniger kritische Systeme (Kategorie Silber) können auch später wieder angefahren werden und Systeme der Kategorie Bronze können bis zu einem Tag später verfügbar sein. Wie diese Ziele über technische Mechanismen erreicht werden können, zeigt das nachfolgende Teilkapitel.

### 2.2.4.3 Verfügbare Mechanismen

Um SAP-Systeme gegen den Ausfall eines kompletten Rechenzentrums abzusichern, können verschiedene Mechanismen genutzt werden. Hierzu wird zunächst nach der Art

**Tab. 2.3** RTO und RPO von SAP-Systemen

Kategorie	Gold	Silber	Bronze
Verfügbarkeiten	99,5 %	98 %	95 %
Maximaler Ausfall	4 h	14 h	37 h
RTO	< 30 min	< 4 h	< 24 h
RPO	< 15 min	< 1 h	< 4 h

Verfügbarkeiten erweitert um RTO und RPO

der Mechanismen unterschieden. Jeder der Mechanismen hat gewisse Vorteile und Nachteile.

- **Hot Standby** – Soll bei einem Ausfall eines SAP-Systems ein möglichst kurzer Zeitraum bis zur Wiederverfügbarkeit des Systems in dem DR-Rechenzentrum liegen, so werden Mechanismen für ein Hot Standby genutzt. Dabei werden Teile des SAP-Systems synchronisiert (zum Beispiel die Datenbank).
- **Cold Standby** – Bei einem Cold Standby kommen häufig Replikationsmechanismen zum Einsatz, welche zu einer erhöhten RTO führen.
- **Backup & Restore** – Für wenig kritische Systeme empfiehlt sich die Implementierung einer reinen Backup&Restore-Strategie für den Fall eines Desasters. Somit können Systeme in einer begrenzten Zeit wiederhergestellt werden.

Jede Komponente des SAP-Systems kann auf unterschiedliche Weise gegen ein Desaster abgesichert werden und üblicherweise ergibt sich ein Mix. Bei SAP S/4HANA Systemen werden für sehr kritische Systeme, Synchronisierungen über die Bordmittel der HANA-Datenbanken genutzt. Hierzu zählt die HANA System Replication, welche die Optionen der Replikation und der vollständigen Synchronisierung mitbringt. Für die Komponenten des Applikationsservers und der SAP Central Services können die Cloud-Bordmittel eingesetzt werden. Ein Beispiel ist die Azure Site Replication, welche die virtuellen Server der Central Services in ein DR-Rechenzentrum replizieren (Tab. 2.4).

Um S/4HANA Systeme gegen ein Desaster abzusichern, werden die Mechanismen zu Beginn des Deployments berücksichtigt und implementiert. Die Implementierung eines aktiven DR-Mechanismus, via beispielsweise HANA System Replication, bedarf einiges Aufwandes und Zeit. Darauf hinaus müssen auf der Primär- und Sekundärseite jeweils die HANA Systeme betrieben werden. Dadurch entstehen Kosten der Hyperscaler, welche signifikant ausfallen können.

Je geringer die RTO und RPO ausfallen sollen, desto höher fallen die Kosten für ein DR-Mechanismus bei der Implementierung und im Betrieb aus. Neben der initialen Implementierung fallen auch hohe Kosten für den Betrieb von synchronen und asynchronen Lösungen an, da hier die Komponenten in der Sekundärseite ständig verfügbar sein müssen.

**Tab. 2.4** Vor- und Nachteile von DR-Mechanismen

Mechanismus	Hot Standby	Cold Standby	Backup & Restore
Charakteristik	Synchron – Primär- und Sekundärseite werden synchron gehalten; es existiert kein Zeitversatz zwischen der Primär- und Sekundärseite	Asynchron – Es existiert ein kleiner Zeitversatz zwischen der Primär- und Sekundärseite	Asynchron – Es existiert ein großer Zeitversatz zwischen Primär- und Sekundärseite
RTO	< 30 min	< 4 h	< 24 h
RPO	< 15 min	< 1 h	< 4 h
Investition	Hoher Aufwand für initiales Setup	Hoher Aufwand für initiales Setup	Wenig Aufwand für initiales Setup
Betrieb	Hohe Kosten durch aktive Komponenten in der Sekundärseite	Hohe Kosten durch aktive Komponenten in der Sekundärseite	Sehr geringe Kosten durch Speicherung der Backups
Desaster Recovery Tests	Einfache DR-Tests durch simples Umschalten auf Sekundärseite	Einfache DR-Tests durch simples Umschalten auf Sekundärseite	Hoher Aufwand, da Restores durchgeführt werden müssen

#### Vor- und Nachteile von DR-Mechanismen

Insbesondere bei HANA Datenbanken können die laufenden Kosten stark ins Gewicht fallen. So kostete eine virtuelle Maschine für eine 3.8 TB HANA Datenbank in West-europa (Amsterdam) in Microsoft Azure bis zu **22.000 €** (Stand Mitte 2021) je Monat. Diese Kosten lassen sich durch Commitments etc. reduzieren, fallen aber bei synchronen und asynchronen Implementierungen monatlich an. Bei einer Lösung basierend auf Backup & Restore werden kaum Ressourcen auf der Sekundärseite benötigt und im Prinzip fallen nur die Kosten für die Speicherung der Backups an.

Neben den Kosten der Implementierung und des Betriebs, entstehen Kosten auch durch die eigentlichen DR-Tests. Hierbei gestaltet sich ein Test bei synchronen und asynchronen Lösungen eher einfach. Es erfolgt ein Umschalten von der Primärseite auf die Sekundärseite und ein abschließendes Testen:

1. Überprüfung der Synchronisierung zwischen Primär- und Sekundärseite
2. Prüfung des aktuellen Datenstandes durch einfache Tests (z. B. Anzahl der aktuell existierende Benutzer im Mandanten).
3. Wechsel von Primär- zu Sekundärseite
4. Abschalten der Primärseite
5. Überprüfung der Sekundärseite und Anpassung der Konnektivität
6. Prüfung des Datenstandes
7. Synchronisierung von Sekundärseite zu Primärseite
8. Wechsel von Sekundär- zu Primärseite und Anpassung der Konnektivität

Eine Durchführung eines DR-Tests bei Backup & Restore ist dahingehend mit mehr Aufwand verbunden. Hierzu muss zunächst neue virtuelle Infrastruktur erstellt werden und danach ein Restore durchgeführt werden. Das Durchführen dieser Tests dauert üblicherweise sehr viel länger und kann bei einem Restore basierend auf einem Datenbank-backup folgende Schritte beinhalten:

1. Überprüfung der verfügbaren Backups in der Sekundärseite
2. Provisionierung der neuen virtuellen Hardware (virtuelle Maschine, Storage, Netzwerk, Resource Groups etc.)
3. Provisionierung von Betriebssystem und Dateisystem
4. Provisionierung des leeren S/4HANA Systems (ohne Daten)
5. Installation und Konfiguration des Backup-Agenten
6. Durchführung des Restores der Datenbank
7. Testen des S/4HANA Systems nach erfolgreichem Restore
8. Prüfung des Datenstandes
9. Deprovisionierung der virtuellen Maschine

Das regelmäßige Testen von DR-Mechanismen ist wichtig für jedes Unternehmen und sollte für alle SAP-Systeme jährlich durchgeführt werden.

#### ► **Regelmäßiges Testen**

Es erscheint zwar sehr naheliegend, dass DR-Mechanismen in regelmäßigen Abständen getestet werden sollten, jedoch ist der damit verbundene Aufwand nicht zu unterschätzen. Hierbei ist noch nicht einmal der Aufwand für die eigentliche Ausführung ausschlaggebend, sondern vielmehr die Planung vor dem Test.

Ein DR-Test eines S/4HANA Systems kann nicht ohne die Abhängigkeiten der verbundenen Systeme durchgeführt werden. So sind die S/4HANA Systeme durch Schnittstellen (z. B. RFC) mit anderen Systemen verbunden und PI/PO Systeme senden und versenden stetig Daten zu den S/4 Systemen.

Bei einem DR-Test werden IP-Adressen geändert, es müssen Konfigurationen in den Umsystemen vorgenommen werden und externe Partnersysteme neu angebunden werden. Solch eine Aktivität kann sehr komplex werden.

Ein regelmäßiges Testen in einem isolierten Umfeld ist mit nur einem System ist einfach, aber meist nicht vollständig. Dennoch ist ein Test einmal pro Jahr tatsächlich wichtig und wird in Audits regelmäßig geprüft.

### **2.2.5 Backup und Restore**

Im vorherigen Kapitel wurde das Thema Desaster Recovery besprochen, was für moderne SAP S/4HANA Systeme ein elementares Thema ist. Hierbei war Backup & Restore ein Weg, um die SAP-Systeme im Falle eines Desasters wiederherzustellen.

Generell sollte jedoch Backup und Restore in keiner Umgebung fehlen, da nur auf diesem Wege, die Daten der SAP-Systeme gesichert werden und die Verfügbarkeit der Daten im Falle eines Fehlers auch sichergestellt werden kann.

### 2.2.5.1 Relevante Komponenten für ein Backup

SAP S/4HANA Systeme bestehen aus einer Reihe von Komponenten und wichtigen Teilen, welche bei einer Datensicherung einbezogen werden müssen. In Abhängigkeit der Kritikalität der SAP-Systeme werden die Daten und die Konfigurationen verschieden oft gesichert. Die folgenden Komponenten sind für eine Sicherung wichtig:

- **Virtuelle Maschine** – Bei einer Sicherung der virtuellen Maschine ist die Konfiguration der VM wichtig. Hierzu gehören der Name der VM, welche virtuellen Disks an der VM angehängt sind, welches VM Template genutzt wurde (bspw. in Azure M128) und welche virtuellen Netzwerkports konfiguriert wurden. Es ist wichtig, dass diese Informationen für einen möglichen Restore in regelmäßigen Abständen gesichert werden. Es ist aber nicht davon auszugehen, dass sich die Konfiguration zu häufig ändert.
- **Betriebssystem** – Das Betriebssystem ist die Grundlage des SAP S/4HANA Systems. Es beinhaltet die grundlegenden Konfigurationen der virtuellen Maschine, die Dateisysteme und die Konfigurationsdateien der SAP-Systeme. Da es sich nicht häufig ändert, werden die Betriebssysteme zwar in regelmäßigen, aber nicht allzu häufigen Abständen gesichert.
- **Dateisystem** – Ein SAP S/4HANA System besitzt meist nicht nur ein Dateisystem, sondern es gehören etliche Dateisysteme dazu. Darin enthalten sind alle wichtigen Konfigurationsdateien des Betriebssystems, als auch des SAP-Systems, das Kernel der Datenbank und des SAP-Systems.
- **Datenbank** – Die Datenbank beinhaltet alle wichtigen Daten des SAP-Systems und eine regelmäßige Sicherung ist unverzichtbar. Für eine Sicherung der HANA-Datenbanken sind die Sicherung der Redo Logs und der Data Volumes wichtig. Die Sicherung kann in verschiedenen Weisen (komplett und inkrementell) erfolgen.
- **SAP-spezifische Verzeichnisse** – Die SAP S/4HANA Systeme speichern verschiedene Dateien in den Dateisystemen ab. Dazu gehören der SAP-Kernel, der Datenbankkernel, die Verzeichnisse von sapmnt und das global Mount Dateisystem, sowie das Transportverzeichnis. Alle diese wichtigen Verzeichnisse sind für ein Backup wichtig.

Neben den oben aufgeführten wichtigsten Komponenten existieren oftmals noch weitere Verzeichnisse und Daten, welche für eine Sicherung wichtig sind. Hierzu zählen beispielsweise Interface- und Audit-Verzeichnisse oder Verzeichnisse für den Austausch von Dateien. Da diese jedoch sehr individuell sind, werden sie hier nicht weiter betrachtet.

### ► Sinnvolle Sicherungen

Generell gilt, dass nicht alle Komponenten eines SAP S/4HANA Systems gesichert werden müssen. Je mehr Komponenten jedoch gesichert werden, desto mehr Daten und Konfigurationen können im Bedarfsfall wiederhergestellt werden und desto schneller kann ein SAP-System wieder verfügbar gemacht werden.

Das absolute Minimum bei einer Sicherung eines SAP-Systems, ist die Sicherung der Datenbank. Hierbei kann der Aufwand durch inkrementelle Sicherungen nochmals reduziert werden. Bei kritischen SAP-Systemen müssen möglichst alle Komponenten gesichert werden, um die Zeit zu einer Wiederverfügbarkeit des Systems so gering, wie möglich zu halten.

#### 2.2.5.2 Backupklassen

Nicht alle SAP-System eines Unternehmens sind gleich wichtig oder haben die gleiche Kritikalität. Wie bei der Hochverfügbarkeit und Desaster Recovery auch, werden die SAP-Systeme in verschiedene Backupklassen eingeteilt. Jede Klasse erhält einen individuellen Sicherungsplan, welcher der Wichtigkeit der Daten entspricht. Diese Einteilung in Backupklassen wird ähnlich, wie bei Hochverfügbarkeit etc. vorgenommen.

Wenn die Klassifizierung der Backups anhand der Rolle der SAP-Systeme erfolgt, so können die Komponenten in die Sortierung wie in der nachfolgenden Tabelle gebracht werden (Tab. 2.5).

**Tab. 2.5** Klassifizierung von Backups

Komponente	Gold	Silber	Bronze
Systeme	Produktivsysteme	Qualitätssicherungssysteme	Sandbox, Training und Entwicklungssysteme
Virtuelle Maschine	Täglich	Täglich	Täglich
Betriebssystem	Alle 4 h	Alle 24 h	Alle 24 h
Dateisystem	Alle 4 h	Täglich	Täglich
Datenbank	Data Volumes: Tägliche inkrementelle Sicherung Wöchentliche Komplettsicherung Log Volumes: Nach Aufkommen oder aber alle 15 min	Data Volumes: Tägliche inkrementelle Sicherung Wöchentliche Komplettsicherung Log Volumes: Nach Aufkommen oder aber alle 60 min	Data Volumes: Tägliche Komplettsicherung Log Volumes: Nach Aufkommen oder aber alle 4 h
SAP-spezifische Verzeichnisse	Tägliche inkrementelle Sicherung Wöchentliche Komplettsicherung	Tägliche inkrementelle Sicherung Wöchentliche Komplettsicherung	Wöchentliche Komplettsicherung

Klassifizierung der Komponenten nach Kritikalität

Durch die Klassifizierung lassen sich die unterschiedlichen Komponenten anhand der Kritikalität sichern. Wichtige Komponenten werden sehr häufig gesichert (wie z. B. Log Volumes) und weniger kritische Komponenten werden nur selten gesichert.

Gleichwohl Speicherplatz in der Public Cloud sehr günstig ist, ist eine effiziente Aufbewahrung der Backups wichtig. Die gesicherten Daten können sonst zu einem erheblichen Kostenbeitrag führen.

### 2.2.5.3 Aufbewahrung

Heutige S/4HANA Systeme können leicht die 10TB-Grenze überschreiten und erreichen mittlerweile eine signifikante Größe, welche eine effiziente Aufbewahrung der Daten wichtig macht. Eine einfache Kalkulation der zu sichernden Daten kann die Wichtigkeit des Themas verdeutlichen. So wird aus einem S/4HANA System mit einer Datenbankgröße von 10 TB, ein zu sicherndes Volumen von 20TB je Monat.

#### Kalkulation der aufzubewahrenden Daten

Zur Berechnung des Gesamtvolumens der zu speichernden Daten, wird ein System von einer Größe mit folgenden Eckdaten angenommen:

Systemkomponente	Größe	Einheit
Data Volumes	10240	TB
Log Volumes	1024	TB
Betriebssystem	256	GB
Änderungsrate	10	%
Komplettsicherung	Wöchentlich	
Inkrementelle Sicherung	Täglich	
Kalkulation der Datenmenge		

Basierend auf dem obigen Beispiel, ergibt sich das folgende Volumen für die Sicherung der Daten:

- Komplettsicherung – In Summe sind für die wöchentliche Komplettsicherung 11.5 TB an Backupvolumen zu kalkulieren.
- Tägliche Sicherungen – Durch die Änderungsrate wird auch das zu sichernde Volumen bestimmt, was sich auf bis zu 2.05 TB beläuft. Diese täglichen Sicherungen ergeben je Woche ein Volumen von 14.3 TB ( $7 \times 2.05$ ).
- Wöchentliches Volumen – Basierend auf der Komplettsicherung des Systems und der Summe der täglichen Sicherungen, ergibt sich ein Gesamtvolumen für das Backup von 25.8 TB.
- Monatliches Volumen – Die wöchentlichen Sicherungen der Daten führen zu einem Gesamtvolumen von **103.5 TB**.

Dieses Backup-Volumen muss zusätzlich zu den normalen Daten gehalten werden. In den Hyperscaler Clouds müssen der belegte Storage bezahlt werden. Daher ist eine effiziente Speicherung der Backup-Daten sehr wichtig. ◀

Das obige Beispiel zeigt die Relevanz einer effizienten Speicherung der Backupdaten, da die Preise der verschiedenen Storageklassen in den Hyperscaler Clouds sehr unterschiedlich ist. Anhand der Microsoft Azure Cloud kann das einfach verdeutlicht werden:

Premium SSD Storage – Eine Premiumdisk (P30 mit 1024 GiB) kostet je Monat 125 €

Standard SSD Storage – Eine Standarddisk (E30 mit 1024 GiB) kostet je Monat 64 €

Standard HDD Storage – Eine Standarddisk (S30 mit 1024 GiB) kostet je Monat 35 €

Standard BLOB Storage – Eine Kapazität von 1024 GiB auf dem Hot Tier kostet je Monat 17 €

Standard BLOB Storage – Eine Kapazität von 1024 GiB auf dem Cold Tier kostet je Monat 8 €

Standard BLOB Storage – Eine Kapazität von 1024 GiB auf dem Archive Tier kostet je Monat 3 €

Die obigen Preisbeispiele aus Mitte 2021 zeigen, wie gravierend die Unterschiede sein können und wie wichtig eine korrekte Speicherung der Daten auf den jeweiligen Storageklassen ist. Die Hyperscaler bieten alle jeweilige Storageklassen an, welche verschiedene Bedarfe der SAP S/4HANA Systeme adressieren.

Die unterschiedlichen Storageklassen werden relevant, wenn je S/4HANA System bestimmt werden soll, worauf die normalen Daten gespeichert werden sollen und worauf das Backup gesichert werden soll:

- **Normaler Storage (SSD, HDD)** – Der normale Storagebereich der S/4HANA Systeme sollte auf den normalen Disks der Hyperscaler gespeichert sein. Die Disks erreichen die notwendige Performance für HANA-basierte Systeme, sind jedoch nicht für die Speicherung von Backups geeignet.
- **Hot Tier** – Bei einem BLOB Storage wird davon ausgegangen, dass er primär zu Sicherung von großen Datenmengen genutzt wird, welcher aber nicht mehr oft genutzt werden. Dennoch bieten Hot Tier Klassen viel Speicherplatz und Daten können auch innerhalb von kurzer Zeit wieder angesprochen und gelesen werden.
- **Cold Tier** – Bei einem BLOB Storage der Cold Tier-Klassen wird davon ausgegangen, dass die abgespeicherten Daten nur noch sehr selten gelesen werden müssen. Daher kommen hier langesame Speicher zum Einsatz, welche die Daten zwar abrufbar halten, aber keine hohen Geschwindigkeiten mehr bieten.
- **Archive** – Die Archivspeicher sind sehr langesame Speichermedien, welche aber die preiswerteste Option von Speicher darstellen. Hier kann es aber sehr lange dauern, bis Daten wieder gelesen werden können. Die Archive-Klasse ist eher für Langzeitarchivierung gedacht.

Die richtige Positionierung der Backups richtet sich nach der Dauer, für welche die Backups gespeichert werden müssen (Retention) und die Art und Weise des Zugriffs der Daten. Backuplösungen für SAP-Systeme bieten hierzu auch das Auto-Tiering an, welches die Daten nach Kriterien auf eine der verfügbaren Storageklassen verteilt. Darüber hinaus haben diese Lösungen auch weitere Eigenschaften, welche zu einer Reduktion der Backupdaten führt. Hierzu gehören die Deduplizierung und die Kompression, welche im nachfolgenden Kapitel erläutert werden.

#### 2.2.5.4 Backuptechnologie

Es existieren verschiedene Verfahren für die Sicherung der Komponenten. Alle großen Hyperscaler bieten hierzu bereits native Backuplösungen an, welche auch eine Integration in die SAP-eigene backint-Schnittstelle besitzen. Damit lassen sich die Backups auch innerhalb des SAP-Systems nachvollziehen.

Unabhängig von der gewählten Lösung existieren prinzipiell zwei verschiedene Verfahren, welche eingesetzt werden. Das sind die Snapshot-basierten Backups und die Stream-basierten Backups.

- **Snapshot-basierte Lösungen** – Die Backups werden als Snapshots (Schnappschuss) der virtuellen Maschinen und des darin befindlichen S/4HANA Systems durchgeführt. Ein Snapshot zeichnet den aktuellen Stand der Daten und des Zustands des S/4HANA Systems auf. Alle in einer VM befindlichen Daten werden damit gesichert und verbleiben zunächst auf demselben Storagebereich, wie das S/4HANA System.
- **Stream-basierte Lösungen** – Die Backups werden über einen Agenten aus der virtuellen Maschine gesichert. Der Agent liest die Daten und sichert diese Daten über Pipes in eine Zielumgebung. Die Zielumgebung kann beispielsweise ein virtuelle Tape Library (VTL) sein oder auch ein BLOB-Storage (BLOB = Binary Large Object). Es ist aber in jedem Fall ein zweiter Storagebereich, der unabhängig vom Storage des S/4HANA Systems ist.

Der wichtige Unterschied zwischen beiden Technologien ist der Verbleib der Daten. Während bei einem Stream-basierten Backup, die Daten in jedem Fall auf einem anderen Storagebereich verbleiben als das S/4HANA System, muss bei einem Snapshot-basierten Backup, diese Sicherung zunächst noch auf einen weiteren Sekundärstorage transportiert werden. Somit werden auch die üblichen Anforderungen nach einer Trennung der Backupsbereiche erfüllt.

Bei der Betrachtung der Kosten von Backups wurden die enormen Preisunterschiede zwischen den Storageklassen aufgezeigt. Da Snapshots immer auf demselben Storage liegen, wo die eigentlichen Daten liegen, kann es bei häufigen Snapshots zu einer Anhäufung von Daten kommen. Diese Anhäufung der Daten auf teurem Storage ist nicht sinnvoll und sollte verhindert werden.

Backuplösungen zeichnen sich durch eine intelligente Speicherung der Daten aus. Es werden nicht alle Daten auf dem schnellen Storage gespeichert, damit es nicht zu einer

Kostenexplosion kommt. Hierzu wird das **Auto-Tiering** verwendet, welches die Daten anhand der Zugriffsmuster verteilt. Daten, welche nicht häufig genutzt werden, werden auf einer langsamen Storageklasse gespeichert. Die Daten, welche sehr häufig genutzt werden, werden auf einer schnellen Storageklasse gespeichert. Backuplösungen können die Daten automatisch verteilen und können auch Daten nach dem erstmaligen Schreiben auf den Storage wieder umverteilen.

Neben dem Auto-Tiering gehören die **Deduplikation** und die **Kompression** mit zu den Fähigkeiten von Backuplösungen, welche auch in den Hyperscaler Clouds wichtig sind. Die Deduplikation von Daten bedeutet das einfache Abspeichern von redundanten Daten. Dies ist zum Beispiel bei den Betriebssystemen der Fall. Da die Betriebssysteme regelmäßig gesichert werden, die Daten sich aber nicht signifikant ändert und ein Großteil der Daten gleich bleibt, speichern die Backups auch nur die geänderten Daten ab. Redundante Daten (Duplikate) werden nicht vielfach gespeichert. Durch die Deduplikation kann von einer Reduktion der Backupdaten bis zu 90 % ausgegangen werden.

Die **Kompression** ermöglicht eine weitere Reduktion des Backupvolumens. Hierbei werden die Backups komprimiert abgespeichert. Die Backuplösungen können hierbei aufgrund der Struktur der Daten nicht immer dieselben Ergebnisse und Kompressionsraten erzielen. Bei SAP-Systemen jedoch kann von einer nicht unerheblichen Kompression (oftmals bis zu 50 %) ausgegangen werden.

Eine fundierte Backuplösung kann helfen, die Daten ordentlich und kosteneffizient zu sichern. Dies sollte vor dem Einsatz der Cloud-native Lösungen immer in Betracht gezogen werden, auch wenn eine Backuplösung mehr Aufwand bei dem initialen Setup benötigt.

### 2.2.5.5 Backuplösungen aus der Cloud

Für SAP S/4HANA Systeme existieren verschiedene viele Lösungen aus der Cloud und in der Cloud. Hierzu gehören die etablierten Anbieter, wie Veritas oder auch EMC, aber auch die Hyperscaler, welche zertifizierte Lösungen anbieten.

Mit Stand Mai 2021 gibt es 42 zertifizierte Backuplösungen für SAP-Systeme. Diese Lösungen unterstützen nicht alle die S/4HANA Systeme, sondern haben bestimmte Limitierungen. So gibt es Anbieter, welche sich stark auf Oracle fokussiert haben. Für alle S/4HANA Systeme gibt es 38 Lösungen, welche die Zertifizierung erlangt haben.

#### ► **Wichtig**

##### Backint Schnittstelle und Zertifizierung

Unabhängig davon, welche der verfügbaren Lösungen genutzt werden soll, ist eine Zertifizierung der Lösung von SAP wichtig. Hierzu sei auf die Backint Schnittstelle verwiesen, welche eine standardisierte Schnittstelle zur Sicherung von SAP-Systemen darstellt.

Alle Lösungen, welche für die Backint Schnittstelle zertifiziert sind, können ohne Weiteres auch für die Sicherung von S/4HANA Systemen eingesetzt werden.

Liegt keine Zertifizierung der Lösung vor, sollte eine Alternative angedacht werden. Bei Audits und Prüfungen der SAP-Systeme sind die Backups immer ein wichtiger Punkt.

Die Hyperscaler bieten mit nativen Lösungen einen guten Startpunkt für ein Backup in der Cloud. So bieten sie die folgenden Lösungen an:

- **Microsoft** – Azure Backup BackInt 1.0
- **Google** – Google Cloud Storage Backint Agent for SAP HANA
- **Amazon** – AWS Backint Agent
- **Alibaba** – Apsara HBR 1.6.0

Die Lösungen der Hyperscaler sind Snapshot-basierte Lösungen, welche die Daten auf dem primären Storage sichern. Hiermit lassen sich einfache Backupszenarien implementieren, welche den wichtigsten Anforderungen genügen. Jedoch haben die Lösungen eine Schwäche: sie erlauben keine effiziente Verwaltung der Daten und bieten keine zusätzlichen Features, wie Deduplikation und Kompression, an. Die Aufbewahrung der Snapshots wird durch die Lösungen nicht reglementiert/begrenzt, sondern die Snapshots werden so lange gespeichert, bis sie wieder aktiv gelöscht werden. Somit verbleiben die Daten auf dem Primärstorage, konsumieren Speicherplatz und verursachen hohe Kosten.

Als Alternative zu den Backupslösungen der Hyperscaler bieten die etablierten Lösungsanbieter von Backups ihre Lösungen an. Hier können zwei verschiedene Klassen unterschieden werden:

**Software-as-a-Service** – Hierbei bieten die Anbieter die Backupslösung aus der Cloud an. Somit muss sich der Kunde keine Gedanken zu einer Konfiguration von Backupservern etc. machen, sondern kann den Service konsumieren. Als Beispiel kann hier die von CommVault angebotene Lösung Metallic dienen.

**Installation auf der Cloud** – Hierbei werden die Lösungen als separate Installation auf der Cloud installiert und konfiguriert. Das bedeutet, dass Backupserver, Agenten etc. installiert und konfiguriert werden müssen. Der Kunde hat hierbei die volle Kontrolle und kann alle Einstellungen so vornehmen, wie es nötig ist.

Die wichtigsten Anbieter von zertifizierten Backupslösungen sind in Tab. 2.6 kurz beschrieben.

Es existieren derzeit noch sehr wenige Anbieter, welche eine Backup-as-a-Service Lösung anbieten. Es kann davon ausgegangen werden, dass sich diese Services in den kommenden Jahren noch verstärken werden.

## 2.2.6 Integration und Netzwerk

Cloud Computing zeichnet sich dadurch aus, dass die Dienste generell immer und von überall verfügbar sind. Die Anbieter von Cloud Computing befolgen dieses Credo

**Tab. 2.6** Backuplösungen für SAP S/4HANA Systeme

Anbieter	Lösung	Backup-as-a-Service/ Installation auf der Cloud
Actifio	<b>Actifio VDP 9.0</b>	Installation auf der Cloud
AISHU Technology Corp	<b>AnyBackup CDM 7</b>	Installation auf der Cloud
Alibaba Cloud Computing Limited	<b>Apsara HBR 1.6.0</b>	Backup-as-a-Service
Arcserve	<b>Arcserve Backup 18.0</b>	Installation auf der Cloud
Amazon.com Web Services, Inc	<b>AWS Backint Agent</b>	Backup-as-a-Service
Microsoft Corporation	<b>Azure Backup BackInt 1.0</b>	Backup-as-a-Service
Bacula Systems	<b>Bacula Enterprise Edition 12.6</b>	Installation auf der Cloud
Libelle AG	<b>BusinessShadow 6.5</b>	Installation auf der Cloud
Catalogic Software Inc	<b>Catalogic Software DPX 4.6</b>	Installation auf der Cloud
Cohesity Inc	<b>Cohesity DataProtect 6.0 for SAP HANA</b>	Installation auf der Cloud
Cohesity Inc	<b>Cohesity DataProtect 6.5 for SAP HANA</b>	Installation auf der Cloud
Commvault Systems, Inc	Commvault V11	Installation auf der Cloud
EMC Corporation	Data Domain Boost for Enterprise Applications 4.5	Installation auf der Cloud
EMC Corporation	Data Domain Boost for Enterprise Applications 4.6	Installation auf der Cloud
Dell Marketing LP	<b>Dell EMC NetWorker 19.3</b>	Installation auf der Cloud
Dell Marketing LP	<b>Dell EMC PowerProtect Application Agent 19.5</b>	Installation auf der Cloud
Linke	<b>Emory for SAP HANA 1.0</b>	Installation auf der Cloud
Google Cloud	<b>Google Cloud Storage Backint agent for SAP HANA</b>	Backup-as-a-Service
Commvault Systems, Inc	<b>Hitachi Data Protection Suite V11</b>	Installation auf der Cloud
Hewlett Packard Enterprise	<b>HPE StoreOnce Catalyst Plug-in 2.2.0 for SAP HANA</b>	Installation auf der Cloud
IBM – International Business Machines Corporation	<b>IBM InfoSphere Virtual Data Pipeline 8.1</b>	Installation auf der Cloud
IBM – International Business Machines Corporation	<b>IBM Spectrum Protect for ERP 8.1</b>	Installation auf der Cloud
Commvault Systems, Inc	<b>Metallic 1.0</b>	Backup-as-a-Service

(Fortsetzung)

**Tab. 2.6** (Fortsetzung)

Anbieter	Lösung	Backup-as-a-Service/ Installation auf der Cloud
Micro Focus	<b>Micro Focus Data Protector 10</b>	Installation auf der Cloud
Veritas Technologies LLC	<b>NetBackup</b>	Installation auf der Cloud
QSFT India Pvt. Ltd	<b>NetVault 12.3</b>	Installation auf der Cloud
Rubrik	<b>Rubrik Cloud Data Management v5.0</b>	Installation auf der Cloud
SEP AG	<b>SEP sesam 5</b>	Installation auf der Cloud
Veeam Software Group GmbH	<b>Veeam Backup &amp; Replication v10</b>	Installation auf der Cloud

Verfügbare Lösungen zum Backup von SAP S/4HANA Systemen

auch und so können alle Public Clouds über das Internet angesprochen werden. Eine Kommunikation von sensiblen Unternehmensdaten über das Internet ist jedoch nicht akzeptabel und so implementieren Unternehmen andere Wege für den Zugriff auf die Hyperscaler-Clouds.

### Zugriff via Internet

Ein Zugriff per Internet ist bei allen Hyperscaler-Clouds möglich. Dies erfolgt aus dem Unternehmens-eigenen Netzwerk heraus direkt zu den Hyperscalern. Die Portale der Public Clouds sind ganz normal per Browser erreichbar und so können die ersten Schritte ohne weitere Kosten für einen Umbau des Netzwerks durchgeführt werden.

Der Zugriff über das Internet ist als erste Möglichkeit nutzbar, aber sicherlich keine langfristige Lösung. SAP S/4HANA-Systeme, welche provisioniert werden und nur per Internet verfügbar sind, sind sehr angreifbar. Darüber hinaus müssen alle Daten, welche zwischen dem SAP S/4HANA-System und den anderen IT-Systemen des Unternehmens ausgetauscht werden, durch das Internet laufen. Dies birgt die Gefahr der Manipulation der Daten, welche gefährlich ist.

Als erster Schritt bei der Nutzung der Public Clouds ist der Zugriff via Internet in Ordnung, aber sollte so bald wie möglich abgeschaltet werden. Ein Zugriff via Internet kann als Backup-Weg weiterhin bestehen bleiben, sollte aber dann durch ein Virtual Private Network (VPN) ergänzt werden.

### Azure ExpressRoutes/AWS Direct Connect/Google Cloud Interconnect

Der übliche Weg zur Verbindung eines Unternehmensnetzwerks zu der Public Cloud sind die direkten Verbindungen aus dem Netzwerk heraus, über einen Wide Area Network-Provider (wie AT&T) zu der Public Cloud.

Über den direkten Link wird eine Verbindung aus dem Unternehmensnetzwerk zum WAN-Provider und dann weiter zur Public Cloud erstellt. Hierbei setzen die Unternehmen oft auf die bestehenden WAN-Provider, welche dann in den entsprechenden Regionen der Hyperscaler die Verbindung terminieren lassen müssen.

Solche Verbindungen haben signifikante Vorteile gegenüber einer direkten Internetverbindung:

- Es lassen sich höhere Übertragungsraten erzielen, die Latenz lässt sich hierüber verringern und die Stabilität in der Verbindung (weniger Abbrüche) erhöhen
- Es gibt eine höhere Sicherheit, da die Daten Punkt-zu-Punkt übertragen werden und die Sicherheit Ende-zu-Ende von den Unternehmen implementiert werden kann
- Bei manchen Public Cloud-Anbietern können die direkten Verbindungen zu reduzierten Kosten für den Datentransfer zwischen/aus der Cloud herausführen.

Bei der Erstellung der Verbindung gilt zu beachten, dass die WAN-Provider nicht alle in allen Regionen der Hyperscaler vertreten sind. Wählt beispielsweise ein Kunde die Region Amsterdam aus und der WAN-Provider hat dort keine Leitung gibt es zwei Möglichkeiten:

1. Der Kunde kann auf eine **andere Region** ausweichen, was sehr selten passiert, da Regionen mit Bedacht und bewusst gewählt werden (z. B. wegen regulatorischen Vorgaben)
2. Der Kunde nutzt eine andere Region zum Terminieren der Verbindung und nutzt danach den **Inter-Region-Connect** der Public Cloud-Provider. Das resultiert in höheren Hops im Netzwerk und damit in einer höheren Latenz, was bei SAP S/4HANA-Systemen sehr wichtig ist.

Da alle Public Cloud-Provider eine Verbindung zwischen ihren eignen Regionen besitzen (Backbone), weichen die Kunden oftmals auf die zweite Möglichkeit aus und verbinden die WAN-Verbindung mit einer anderen Region. Vor der Einführung einer neuen Public Cloud-Lösung sind solche Punkte zu betrachten, da sie einen Einfluss auf die Entscheidung haben.

Das Erstellen der Verbindung vom Kundennetzwerk zur Public Cloud benötigt eine gewisse Rüstzeit. Nicht selten sind hierfür mehr als 8 Woche notwendig. Da die Verbindung zur Public Cloud meist der erste Schritt in Projekten ist, kann am Anfang viel Zeit verloren gehen, wenn die Verbindung zu spät beim WAN-Provider bestellt wird.

Manche Unternehmen verfolgen eine Hybrid Cloud-Strategie. Hierbei können die Systeme auf beispielsweise zwei verschiedene Clouds verteilt werden: SAP S/4HANA-Systeme werden in der Microsoft Azure-Cloud betrieben und die nicht-SAP-Systeme werden in der Google Cloud betrieben. In solch einem Fall muss das Unternehmen zwei Verbindungen (Azure und Google) erstellen und das Routing zwischen den beiden Verbindungen einrichten. Solch ein Routing erfolgt oftmals über Devices in der

On-Premise-Umgebung von Kunden. Dies bedeutet jedoch einen hohen Aufwand und einen Verlust von Performance durch viele Hops im Netzwerk. Zusätzlich müssen diese Netzwerkconfigurationen durch die Kunden weiterhin administriert werden. Dies kann zu einer hohen Komplexität führen.

### Hybrid Cloud via Cloud Connect am Beispiel Equinix Fabric

Um die Komplexität durch viele verschiedene WAN-Anbieter zu verringern, bieten Hersteller die sogenannten Cloud Connects an. Solche Cloud Connects sind als zentrale Einstiegspunkte in die Welt der Hyperscaler zu verstehen. Hierbei profitieren die Kunden von einem zentralen Einstiegspunkt in alle Clouds hinein und muss einen Vertrag nur noch mit einem WAN-Provider halten, unabhängig von der Anzahl der Verbindungen.

Bei einem Cloud Connect wird die Verbindung vom Kundennetzwerk über den WAN-Provider zu einem Anbieter eines Cloud Connects hergestellt. Danach erfolgen Verbindungen vom zentralen Cloud Connect zu den jeweiligen Public Clouds und den jeweiligen Einwahlknoten der Public Clouds. Der Kunde muss hierfür keine weitere WAN-Verbindung nutzen, sondern kann die Backbone-Verbindung der Cloud Connect-Anbieter nutzen.

#### Cloud Connect zu Azure in Amerika und Google in Europa

Das Unternehmen ist ein global agierender Konzern mit Niederlassungen und Zweigstellen in Amerika und Europa und begann die IT sukzessive in die Public Clouds zu verschieben. Es wurde ein Hybrid Cloud-Ansatz verfolgt, um die Vorteile der jeweiligen Hyperscaler nutzen zu können und da es bereits vorher schon Geschäftsbeziehungen mit den jeweiligen Hyperscalern gab. So existierten in Amerika verschiedene lokale Lokationen, welche alle an den WAN-Provider angebunden waren. Vom WAN-Provider existierte eine Verbindung zu Azure in Amerika (ExpressRoute nach Atlanta) und eine Verbindung zur Google Cloud in Europa (Google Cloud Interconnect nach Belgien). Darüber hinaus besaß das Unternehmen noch kleiner Co-Location Rechenzentren, welche ebenfalls verbunden waren:

- Amerika: ExpressRoute zu Azure nach Atlanta mit Verbindung über WAN-Provider 1
- Amerika: 15 kleiner Lokationen am WAN-Provider 1
- Europa: Cloud Interconnect zu Google nach Belgien über WAN-Provider 2
- Europa: Co-Location Rechenzentren in Deutschland (z. B. Frankfurt und München) mit Verbindungen via WAN-Provider 2
- Europa: Altes Rechenzentrum mit zentralem Switch und mit der Routing-Infrastruktur mit Verbindung zu WAN-Provider 2

Das Unternehmen stand vor der Herausforderung, dass es zwei verschiedene WAN-Provider in Amerika und Europa hatte, mit denen das Unternehmen die Verbindung

zu den Clouds, als auch das Routing zwischen Amerika und Europa (also den beiden Public Clouds) steuern musste. Um aus diesem Engpass herauszukommen, entschied sich das Unternehmen, auf einen zentralen Cloud Connect zu setzen und die Anzahl der WAN-Provider auf eins zu reduzieren.

Als erster Schritt wurde der Cloud Connect in Europa erstellt. Hierbei wurde die Verbindung über den existierenden WAN-Provider in Europa heraus erstellt. Der Cloud Connect wurde in Belgien erstellt, da hier auch die Google Cloud bestand. Danach wurden die Verbindungen in Amerika umgebaut und die ExpressRoute wurde aufgelöst. Stattdessen kam die Backbone-Verbindung durch den Cloud Connect-Anbieter zur Nutzung. Alle Verbindungen zur Microsoft Azure-Cloud hin wurden anstatt über die ExpressRoute über den Cloud Connect gehandhabt. Des Weiteren wurde der WAN-Provider in Amerika durch den WAN-Provider aus Europa ersetzt. Die kleineren Co-Location Rechenzentren in Frankfurt und München wurden vom WAN-Provider in Europa getrennt und über die Backbone-Verbindung des Cloud Connect-Anbieters angebunden. Der Anbieter war in den Lokationen ebenfalls vertreten. So konnten die Verbindungen von den Co-Location Rechenzentren zum WAN-Provider abgelöst werden. Das alte Rechenzentrum des Unternehmens, welche die zentrale Infrastruktur des Routings beinhaltete, wurde ebenfalls dekommissioniert und das Routing über den Cloud Connect Anbieter realisiert. In Summe ergab sich ein deutlich vereinfachtes Bild:

- Amerika: Verschiedene Lokationen am neuen WAN-Provider 2
- Europa: Cloud Connect über existierenden WAN-Provider 2
- Europa und Amerika: Nutzung des Backbone-Netzwerkes des Cloud Connect-Anbieters

Durch die Konsolidierung und Nutzung des Cloud Connects konnte das Unternehmen die Komplexität stark reduzieren und musste letztlich nur noch einen WAN-Provider steuern. ◀

Die Cloud Connects sind eine gute Möglichkeit, um über einen Provider mehrere Clouds anzubieten. Als Beispiel kann das Angebot von Equinix dienen, welche über die sogenannte Equinix Fabric eine Verbindung zu den wichtigsten Hyperscalern/Public Clouds anbieten kann:

- AWS via Direct Connect
- Azure via ExpressRoute
- Google Cloud via Carrier Peering
- SAP Cloud via SAP Cloud Peering
- IBM Cloud via Direct Link
- Oracle Cloud via Fast Connect

Für Kunden mit einer maximal heterogenen Infrastruktur, verteilt über die verschiedenen Clouds, kann eine einzelne, einfache Verbindung zur Equinix Fabric schon ausreichen, um alle Clouds miteinander zu verbinden. Dies ist ein großer Vorteil.

### Vergleich der Verbindungsarten

Alle Kunden haben die Möglichkeit, aus den zuvor genannten Möglichkeiten zur Verbindung mit den Public Clouds, zu wählen. So ist es möglich, zunächst klein zu starten und mit der Verbindung zum Internet anzufangen. Als langfristige Verbindung kann dies jedoch nicht empfohlen werden. Dann sollte eine direkte Verbindung zu den Public Clouds erstellt werden oder als Alternative ein Cloud Connect genutzt werden (Tab. 2.7).

Das Internet kann zwar durch eine sehr geringe Komplexität punkten (Kunden können sofort mit der Nutzung starten), jedoch kann der Netzwerkbereich des Unternehmens nicht die Public Cloud erweitert werden. Darauf hinaus bietet das Internet keine verlässliche Sicherheit beim Zugriff, Datentransfer und dem eigentlichen Schutz der Public Cloud.

Direkte Verbindungen zu den Public Clouds sind der präferierte Weg vieler Unternehmen. Hierbei kann der Netzwerkbereich erweitert werden und die Sicherheit der SAP S/4HANA-Systeme kann hierbei gewährleistet werden. Die Komplexität kann jedoch durch eine Vielzahl von WAN-Providern, direkten Verbindungen mit den Cloud Providern und Routing sehr hoch sein.

Die Cloud Connects bieten die Vorteile der direkten Verbindungen und können durch eine Vereinfachung des Netzwerklayouts die Komplexität reduzieren. Die Sicherheit der SAP-Systeme wird hier genauso gewährleistet, wie bei den direkten Verbindungen.

### 2.2.7 Automation

Die Administration von SAP S/4HANA-Systemen kann signifikant vereinfacht werden, indem die täglichen und wiederkehrenden manuellen Schritte bei der Administration automatisiert werden. Hierbei können in der Public Cloud dieselben Automatisierungslösungen genauso eingesetzt werden, wie in den traditionellen Rechenzentren. Die Hyperscaler jedoch bieten auch Lösungen zur Automatisierung der Arbeiten auf der Public Cloud an, wie z. B. Provisionierung oder Start/Stopp. Diese Automatisierungen

**Tab. 2.7** Vor- und Nachteile der Verbindungsarten

Zugangsart	Erweiterung des Netzwerkbereichs	Sicherheit	Komplexität
Internet	Nein	Sehr gering	Sehr gering
Direkte Verbindungen	Ja	Hoch	Hoch
Cloud Connect	Ja	Hoch	Mittel

Wichtige Faktoren im Vergleich

beschränken sich jedoch meist auf die Services, welche durch die Hyperscaler angeboten werden. Hierdurch können aber auch schon einige wichtige Funktionen und Arbeitsschritte automatisiert werden. Daneben existieren auch Tools, welche durch Drittparteien angeboten werden und welche in der Cloud genutzt werden können. Hier gilt beispielsweise SAP Landscape Manager (LAMA) als ein Beispiel.

### Ziele der Automatisierung

Die komplette Automatisierung aller Aufgaben der SAP-Basisadministratoren ist sicherlich nicht möglich und ist auch nicht das Ziel einer Automatisierung. Insbesondere wenn die Automatisierung in der Cloud gestartet wird, werden folgende wichtige Ziele verfolgt:

- **Prozessautomatisierung:** Die Automatisierung von repetitiven Arbeitsschritten im Bereich der Provisionierung und des Betriebs von SAP S/4HANA-Systemen ist eines der wichtigsten Ziele.
- **Vereinheitlichung der Landschaft:** Die Systemlandschaft soll homogen gehalten werden. Durch die Automatisierung werden Systeme in derselben Art und Weise provisioniert und so vereinheitlicht.
- **Konfigurationsmanagement:** Viele Unternehmen sehen die Erfassung der IT-Landschaft und die Pflege der Configuration Items (CIs) als herausfordernd an. Die Datenbanken mit allen Konfigurationsdaten (CMDBs) sind oftmals nicht immer akkurat und gepflegt. Durch die Nutzung von Automatisierung der Hyperscaler kann dem Problem begegnet werden.
- **Landschaft auf dem aktuellen Stand halten:** Das regelmäßige Patchen der Systeme wird durch die zunehmende Anzahl von Cyberattacken auf die Unternehmen immer wichtiger. Die Hyperscaler bringen hierzu die notwendigen Tools mit, um die SAP-Infrastruktur up to date zu halten.
- **Auditing:** Audits verlangen genaue Daten und oftmals sehr viele Daten von den Unternehmen. Hierzu bieten die Hyperscaler einfache Möglichkeiten, um wichtige Reports zu generieren und an die Prüfer weiterzugeben.
- **Tagging:** Tags sind kleine Informationen zu den SAP S/4HANA-Workloads, welche in den Portalen der Hyperscaler zu den Ressourcen zugeordnet werden können. Hierüber können beispielsweise Zugehörigkeiten von Ressourcen zu bestimmten Projekten/Abteilungen/Teams kenntlich gemacht werden.
- **Zeitpläne:** Wenn SAP S/4HANA-Systeme in der Cloud eingeschaltet sind, werden Kosten generiert. Um die Kosten für Systeme zu minimieren, welche nicht immer 100 % verfügbar sein müssen, können Zeitpläne (Schedules) eingerichtet werden, welche die Ressourcen an- und abschalten.

Alle obigen Ziele können in den seltensten Fällen erreicht werden, sondern vielmehr konzentrieren sich die Unternehmen auf die sukzessive Implementierung. Des Weiteren ist die Implementierung von komplexen Automatisierungen ein sehr langfristiges Vorhaben.

## Automatisierung mit Hyperscaler

Alle Hyperscaler bieten den Kunden die Möglichkeit, über eigene, Cloud-native Tools zu automatisieren. Diese Automatisierungen können sehr leichtgewichtig ausfallen, können aber auch immer komplexer werden.

Einfache Implementierungen betreffen beispielsweise die Standardoperationen in einer Cloud:

- Starten und Stoppen von virtuellen Maschinen
- Erstellen, löschen, ändern von Ressourcen in den Clouds (wobei alle Ressourcen gemeint sind)

Die einfacheren Aufgaben können über die von den Hyperscaler angebotenen Tools, wie z. B. Azure Automation oder auch Google Composer, durchgeführt werden.

Die etwas komplexeren Aktivitäten, wie z. B. das Erstellen eines neuen SAP S/4HANA-Systems ist meist über die Cloud-nativen Tools nicht möglich. Hierzu müssen beispielsweise folgende Schritte durchgeführt werden:

1. Erstellen aller Infrastrukturkomponenten (Storage, virtuelle Maschine, Netzwerksegmente)
2. Erstellen/Installation des Betriebssystems mit dem Hostnamen und der gewünschten IP-Konfiguration
3. Installation eines leeren SAP S/4HANA-Systems
4. Befüllen des SAP S/4HANA-Systems mit den Daten

Da die Schritte untereinander koordiniert ablaufen müssen, werden zusätzliche Tools zur Automatisierung benötigt. Hierzu bieten sich die Automatisierungstools an, welche bereits schon eine breite Anwendung in den Public Clouds finden, wie z. B. Ansible oder auch Terraform.

## Drittanbieter

Automatisierungslösungen gibt und gab es schon lange, bevor die Hyperscaler den Markt betrat. Auch in den traditionellen Rechenzentren wurden die repetitiven Administrationsschritte automatisiert. Spätestens seit der Einführung der Virtualisierung (z. B. VMware) wurden viele Aufgaben durch die mitgelieferten Tools der Hersteller der Virtualisierung komplett automatisiert. Seit die Hyperscaler den Markt betreten haben, gibt es eine Vielzahl von Anbietern von Automatisierungslösungen. Zu Beginn der Public Clouds haben die Hyperscaler noch keine Lösungen zur Automatisierung angeboten, wodurch sich solche Anbieter, wie beispielsweise Ansible oder Terraform, ihre Position am Markt erarbeiten konnten.

Eine der Vorteile in der Verwendung der Drittanbieter ist die starke Fokussierung auf Skripte und Wiederverwendbarkeit. Die meisten Lösungen der Drittanbieter nutzen ein

Repository an Skripten, um die Aufgaben zu erledigen. Dieses Repository kann durch eigene Skripte, welche spezifisch angepasst sind, erweitert werden.

Durch die Speicherung der Skripte in einem Repository und die Nutzung einer Lösung eines Drittanbieters können die Kunden der Hyperscaler auch bei einem Wechsel der Plattform (also z. B. von Azure zu Google Cloud) die vorhandenen Skripte mitnehmen und in der neuen Umgebung nutzen. Nutzen die Kunden jedoch nur die eigene Automatisierung der Hyperscaler, kann die Arbeit unter Umständen nicht weiterverwendet werden.

### SAP-eigene Automatisierung

Es bietet sich für alle Kunden an, dass die SAP-eigenen Produkte für die Automatisierung genutzt werden. Das Nutzen der Lösung von SAP erscheint logisch, da die SAP die S/4HANA-Systeme am besten kennt und somit auch die passende Software hierfür bauen kann. Die SAP bietet eine Lösung für die Automatisierung der SAP-Basis-Arbeit an (SAP LAMA) und eine Lösung für die Automatisierung von repetitiven Arbeiten außerhalb der technischen Themen (SAP RPA).

SAP Landscape Management ist eine Weiterentwicklung vom SAP Landscape Virtualization Manager (SAP LVM) und dem SAP Adaptive Computer Controller (SAP ACC). Beides waren Produkte, welche durch die SAP zunächst in den eigenen Rechenzentren genutzt worden sind. Schnell erkannte die SAP den Nutzen der Lösung für z. B. Massenoperationen in den Rechenzentren und erweiterte die Lösungen zügig um einige sehr interessante Funktionen, wie z. B. das Kopieren von SAP-Systemen. Die SAP fokussiert nicht unbedingt die häufigsten Arbeiten, welche bei SAP-Basis-administratoren anfallen, sondern eher die Themen, welche sehr viel Zeit und Energie benötigen. Hierzu zählen und zählten das Kopieren von SAP-Systemen (copy, clone), als auch die sogenannten System Refreshes (also eine Kopie von PRD auf QAS mit einer Umbenennung der SID).

SAP LAMA bietet aufgrund der Historie (zunächst SAP ACC, dann SAP LVM) eine sehr gute Unterstützung von anderen Technologien, wie z. B. Hardwareherstellern. So kann SAP LAMA über eine Library mit NetApp Storage kommunizieren. Diese Unterstützung war zwingend erforderlich, als SAP ACC und SAP LVM die ersten Funktionalitäten implementierten, bei denen auch der Storage in den Rechenzentren mit angesprochen werden musste. Nach demselben Prinzip interagiert SAP LAMA auch mit den Public Clouds, welche alle eine Schnittstelle zur Administration bieten.

SAP LAMA existiert in zwei verschiedenen Versionen: einerseits können Kunden SAP LAMA direkt in dem Rechenzentrum oder der Cloud aufbauen. Andererseits können Kunden auch die SaaS-Version von SAP LAMA nutzen. Diese wird von der SAP bereitgestellt und somit müssen sich Kunden auch nicht um die Administration des SAP LAMA-Systems kümmern. Dies ist auch ein SAP-System, welche es zu administrieren gilt.

### SAP Intelligent Robotic Process Automation

Neben den rein technischen Aufgaben versuchen Unternehmen auch, weitere Aufgaben zu automatisieren. Diese sind meist wiederkehrende Arbeiten und Prozessschritte, welche auch durch nicht-humane Ressourcen ausgeführt werden können. Eine Möglichkeit, diese zu automatisieren, ist die **SAP Intelligent Robotic Process Automation** (SAP iRPA). Roboter (Bots) sind hierbei kleine Programme oder Skripte, welche exakt dieselben Prozessschritte immer und immer wieder durchführen. Über eine Parametrisierung können die Ablaufschritte auch verfeinert werden. Die Robotic-Lösungen haben immer wieder mit verschiedenen Herausforderungen zu kämpfen. Da die Roboter auf den User Interfaces (UI) arbeiten, können sich durch Änderungen am UI, große Änderungen an den Robotern ergeben, wodurch ggf. sehr viele Roboter angepasst werden müssen. Hier setzt die SAP an und erweitert die Lösung. Die Intelligenz bei iRPA kommt bei der SAP durch die Nutzung weiterer intelligenter Features: Machine Learning und Artificial Intelligence. Darüber hinaus können die Roboter nun auch mittels Schnittstellen agieren.

Die SAP hat iRPA als reine cloudbasierte Lösung aufgebaut. So werden neue Roboter in einem Cloud Studio per reinem UI (ohne Code) erzeugt. Aktuell besitzt die Lösung 200 vorkonfigurierte Roboter, welche leicht angepasst und eingesetzt werden können. Die Robots kommen in vielen Unternehmen zum Einsatz, jedoch weniger im technischen Umfeld von der Administration von SAP S/4HANA-Systemen.

### Aufwand zur Implementierung

Die Automatisierung der täglichen Arbeit von SAP-Administratoren ist durchaus sinnvoll und sehr hilfreich. Nur auf diesem Wege schaffen es die Unternehmen, die Arbeitszeit der Administratoren von den täglichen Aufgaben zu befreien und die Administratoren können sich um andere Aufgaben (wie z. B. der Einführung von neuen Features) kümmern.

Die Einführung von Automatisierungslösungen kann jedoch sehr komplex werden. Bei der Nutzung der von den Hyperscaler zur Verfügung gestellten Lösungen können Kunden sehr schnell starten und erste kleinere Aufgaben sehr zügig implementieren. Sobald jedoch die Prozesse komplexer werden oder aber mehrere Prozessschritte miteinander verbunden werden sollen, muss hierfür mehr Zeit investiert werden. Die komplette Implementierung einer neuen Lösung, wie SAP LAMA, muss als separates Projekt durchgeführt werden.

### 2.2.8 Horizontale und vertikale Skalierung

Die Skalierung der SAP S/4HANA-Systeme bezieht sich auf zwei Richtungen und wird in horizontale und vertikale Skalierung unterschieden. Die horizontale Skalierung bezieht sich bei SAP S/4HANA-Systemen auf die Erweiterung des bestehenden Systems durch weitere, zusätzliche HANA-Knoten oder weitere Applikationsserver. Die vertikale

Skalierung bezeichnet das Wachsen des Systems und die Vergrößerung des Sizings der virtuellen Maschinen und des SAP-Systems durch erhöhte Anforderungen durch das Business.

### Horizontale Skalierung

Die horizontale Skalierung beschreibt die Erweiterung des SAP-Systems durch weitere Applikationsserver oder durch weitere HANA-Knoten. Die Erweiterung des SAP-Systems durch zusätzliche Applikationsserver ist hierbei jedoch nicht neu bei den SAP S/4HANA-Systemen, sondern existiert schon seit den Anfängen der SAP R/3-Systeme. Auch damals konnten weitere Dialoginstanzen als zusätzliche Applikationsserver hinzugefügt werden.

Der horizontalen Skalierung sind prinzipiell keine Grenzen gesetzt und so können für sehr große SAP S/4HANA-Systeme durchaus bis zu acht oder mehr Applikationsserver zum Einsatz kommen. Oftmals sind die Erweiterungen durch Applikationsserver weniger den Leistungsgrenzen der einzelnen Server geschuldet, sondern vielmehr werden bestimmte Applikationsserver nur bestimmten Gruppen von Anwendern zur Verfügung gestellt. Oft verbreitet ist auch der Fall, dass Applikationsserver nur für Hintergrundjobs genutzt werden, da ein Einfluss auf die Endbenutzer durch langlaufende Jobs vermieden werden soll.

---

#### Applikationsserver je Funktionsbereich

In vielen größeren Implementierungen von SAP S/4HANA-Systemen werden mehr als ein Applikationsserver implementiert. Sehr oft nutzen einzelne Abteilungen die jeweiligen Applikationsserver.

Ein Unternehmen aus dem Pharmabereich implementierte für das wichtigste S/4HANA-System vier Applikationsserver:

1. Applikationsserver: Dieser Server war für die Finanzabteilung (FICO) vorgesehen. Durch das Prinzip der Logon-Gruppen wurde der Server nur in eine Gruppe integriert, welche auch nur der Finanzabteilung bekannt war.
2. Applikationsserver: Dieser Server wurde dem Großteil der Benutzer zugewiesen. Auch hier kam das Prinzip der Logon-Gruppen zum Einsatz und die Gruppe DEFAULT wurde hierfür genutzt.
3. Applikationsserver: Dieser Server wurde für alle eingehenden RFC-Verbindungen genutzt. Hintergrund waren einerseits Sicherheitsabwägungen, aber andererseits auch die spezifische Natur von RFC-Calls und die eingehenden Transaktionen plus IDOCs.
4. Applikationsserver: Dieser Server wurde für die langlaufenden Hintergrundjobs genutzt. Da in dem SAP S/4HANA-System sehr viele Hintergrundjobs geplant waren (oftmals mit Frequenz von 10 min), entschied sich der Kunde für einen dedizierten Applikationsserver hierfür.

Durch die Trennung der einzelnen Bereiche voneinander, konnte das Unternehmen die Last durch die Anwender besser voneinander trennen, erhöhte die Sicherheit und die gesamte Performance des SAP S/4HANA-Systems. ◀

Die horizontale Skalierung kommt aber nicht nur bei den Applikationsservern zum Einsatz, sondern auch bei den HANA Scale-Out Systemen. Diese HANA-Systeme werden für große Business Warehouse-Systeme genutzt und können dadurch sehr große Datenmengen im Hauptspeicher, jedoch verteilt auf mehrere Knoten, halten. Insbesondere bei den BW-Systemen existieren die Anforderungen nach einem starken Datenwachstum. Hintergrund des Wachstums sind die Abschlüsse, bei denen viele temporäre Daten gespeichert werden müssen, sowie große Datenladeläufe, welche ebenfalls zu vielen temporären Daten führen. Somit zeigen BW-Systeme einen signifikant größeren Datenfußabdruck. Dieses Wachstum wird durch die Provisionierung durch weitere HANA-Knoten adressiert.

Das typische Wachstum von den gesamten SAP S/4HANA-Systemen wird nicht über die vertikale horizontale Skalierung, sondern üblicherweise durch die vertikale Skalierung adressiert.

### Vertikale Skalierung

Die vertikale Skalierung beschreibt das Wachstum des SAP S/4HANA-Systems durch eine Vergrößerung der Server (HANA-Server oder Applikationsserver) oder eine Erhöhung der Ressourcen der Komponenten. Dieser Fall tritt auf, wenn mehr Ressourcen gebraucht werden, als zur Verfügung stehen. Dies kann passieren, wenn zusätzliche Benutzer auf das System Zugriff erhalten oder wenn neue Funktionalitäten bereitgestellt werden sollen.

Die vertikale Skalierung kann also auf verschiedenen Ebenen des SAP S/4HANA-Systems erfolgen:

1. **Änderung der Storageklasse:** Für einige SAP S/4HANA-Systeme kann die Änderung des Storage von langsameren SSD Speicher auf schnellen SSD-basierten Storage notwendig sein. Dies ist eine vertikale Skalierung, welche zwar mehr einen größeren Aufwand bei der Änderung bedeutet, jedoch einen enormen Geschwindigkeitszuwachs bringen kann.
2. **Änderung des VM Templates:** Für wachsende Systeme können andere Typen der virtuellen Maschinen eingesetzt werden, welche über mehr CPU und mehr Hauptspeicher verfügen. Diese Änderungen können meist nicht im laufenden Betrieb (on-the-fly) gemacht werden, sondern bedürfen eines Neustarts des Systems. Nachdem eine neue Größe für die VM genutzt wird, müssen noch das Betriebssystem und die SAP-Komponenten (also entweder die HANA-Datenbank oder die Applikationsserver) angepasst werden.
3. **Änderung der HANA Datenbank:** Die Änderung der HANA Datenbank ist einer der häufigsten Anwendungsfälle. Hierbei wird zum Beispiel die Allokation des

Hauptspeichers oder die Anzahl der CPUs geändert, um eine höhere Leistung zu erzielen.

4. **Änderung am Betriebssystem:** Nach der Vergrößerung der virtuellen Maschine oder auch der Änderung vom Storage können auch Modifikationen am Betriebssystem notwendig werden.
5. **Änderung der SAP Applikationsserver:** Die Änderung an den Puffern oder auch der Anzahl der disp+work Prozesse ist eine der häufigsten Anpassungen, um das SAP S/4HANA-System zu skalieren.

Die vertikale Skalierung erfolgt oftmals nicht nur für eine Komponente, sondern wird für mehrere der obigen Komponenten durchgeführt. Eine Vergrößerung der SAP HANA-Datenbank kann nur dann erfolgen, wenn zuvor die virtuelle Maschine und das Betriebssystem angepasst worden sind.

- **Upsizing und Downsizing** Eine vertikale Skalierung funktioniert in **beide** Richtungen: es können Ressourcen hinzugefügt werden und die Systeme vergrößert werden. Es können aber auch Ressourcen wieder entfernt werden und Systeme wieder verkleinert werden.

Im Umfeld von Cloud ist ein Up- und Downsizing sehr wichtig, da hiermit Kosten gespart werden können. Im Zuge einer effektiven Kapazitätsplanung sollten Up- und Downsizing stetig mit betrachtet werden. Hier hilft es, wenn über längere Zeiträume hinweg die Auslastung von SAP S/4HANA-Systemen und den jeweiligen Komponenten betrachtet wird.

### Verfügbarkeiten beachten

Gleichwohl die Anbieter von Public Cloud immer von einer unbegrenzten Verfügbarkeit der Ressourcen sprechen, gibt es Einschränkungen. Kunden sollten vor geplanten Kapazitätserweiterungen darauf achten, dass die entsprechenden Ziel-VM-Templates auch in den jeweiligen Regionen verfügbar sind. So kann es durchaus vorkommen, dass bestimmte VM Templates noch nicht in allen Regionen zur Verfügung stehen und die Kunden dann auf Alternativen ausweichen müssen. Hyperscaler haben üblicherweise immer ein/zwei Regionen, in denen die neuste Technologie am frühesten bereitgestellt wird. Kleinere Regionen für spezifische Märkte können oftmals nicht das komplette Portfolio der Services anbieten.

Zu Beginn der COVID Pandemie gab es auch bei den Anbietern der Public Cloud teils signifikante Ressourcenengpässe. Dies war der, enorm nach oben schnellenden, Nutzung von anderen Services der Hyperscaler geschuldet. Durch diese hohe Nutzung standen anderen Kunden keine oder nur eingeschränkt neue Ressourcen zur Verfügung. Die Hyperscaler nutzen die Public Clouds nicht nur für die Kunden, sondern auch für die eigenen Services. Zu der Zeit war es Kunden nicht möglich, einfach neue virtuelle Maschinen zu provisionieren.

Ganz konkret konnten Kunden in der Microsoft Azure Cloud keine großen HANA-Systeme (z. B. M-Class) mehr bestellen. Die darunterliegenden Maschinen wurden für andere Services genutzt. Microsoft hatte zu Beginn der Pandemie mit einem sprunghaften Anstieg der Microsoft Teams-Nutzer zu kämpfen. Da viele Arbeitsnehmer plötzlich von zuhause aus arbeiten musste, wurden die Ressourcen für MS Teams stetig erweitert. Darüber hinaus kamen neue Anforderungen für die Bekämpfung der Pandemie auf: so haben Forschungseinrichtungen neue und große Systeme auf der Cloud erstellt, um komplexe Berechnungen durchzuführen zu können. All dies führte letztlich dazu, dass der Bedarf durch andere Kunden nicht mehr gedeckt werden konnte. Microsoft reagiert darauf, indem eine Priorisierung eingeführt worden ist und Kunden den Bedarf anmelden musste. Nach solch einer Anmeldung entschied Microsoft aufgrund der Kritikalität des Bedarfs und wies die Ressourcen den Kunden zu (oder auch nicht).

Dieses Beispiel der Pandemie zeigt, dass einer Skalierung bei den Hyperscalern Grenzen gesetzt sind. Auch wenn solche Szenarien sehr unwahrscheinlich sind, kann es immer wieder zu Engpässen kommen, welche die scheinbar unbegrenzte Ressourcen dennoch limitieren.

---

## 2.3 Zusammenfassung

Dieses Kapitel hat die wichtigsten Aspekte bei der Nutzung von Public Clouds für die SAP S/4HANA-Systeme beschrieben. Das richtige Sizing der Systeme ist nicht nur aus Sicht der Anwender und der Performance wichtig, sondern insbesondere aufgrund der Kosten solcher Systeme wichtig. Ein zu groß dimensioniertes System hat in der traditionellen Rechenzentrumswelt keine höheren Kosten verursacht – in der Cloud ist dies jedoch anders und generiert hohe Kosten. Dies gilt es zu vermeiden.

Bei produktiven SAP S/4HANA-Systemen gilt es die möglichen Ausfälle der Komponenten einer Public Cloud (wie z. B. den virtuellen Maschinen) abzufangen und die Verfügbarkeit der SAP S/4HANA-Systeme weiterhin sicherzustellen. Dies wurde in dem Kapitel zur Hochverfügbarkeit erläutert und auch eine Beispielarchitektur besprochen, welche in den Kapiteln zu den jeweiligen Hyperscalern konkret implementiert wird.

Im Teilkapitel zu Desaster Recovery wurde gezeigt, wie die SAP S/4HANA-Systeme durch die Bordmittel der Hyperscaler auch bei einem Ausfall einer Region weiter verfügbar bleiben. Dies kann auch durch die Kombination von Bordmitteln und den Mechanismen der SAP-Systeme (wie z. B. HANA System Replication) erfolgen. In den jeweiligen Kapiteln zu den Hyperscalern wird die Implementierung beschrieben und Schritt für Schritt durch die Einrichtung geführt.

Die beiden wichtigen Themen der Datensicherung und der Wiederherstellung von Daten wurde beschrieben und die Möglichkeiten zur Sicherung der Daten über die Hyperscaler-eigenen Bordmittel, als auch über Produkte von Drittanbietern besprochen.

Diese unterscheiden sich in der Art und Weise, wie sie eingesetzt werden und auch in den notwendigen Kenntnissen zu den Produkten.

Die Themen der Integration und Automatisierung wurden in diesem Kapitel beschrieben und gezeigt, wie einfach mit der Nutzung der Public Cloud-Services begonnen werden kann. Durch eine geschickte Automatisierung können die initialen Aufwände bei den Clouds-Projekten reduziert werden. Insbesondere in dem alltäglichen Betrieb lässt sich durch Automatisierung viel Zeit und Aufwand sparen.

Um den Anforderungen der Fachabteilungen immer gerecht zu werden und den wachsenden Anforderungen an die SAP S/4HANA-Systeme zu begegnen, wurden die horizontale und vertikale Skalierung besprochen. Es wird gezeigt, wie diese in SAP S/4HANA-Systeme umsetzbar ist.

Das nächste Kapitel zeigt das Deployment und die Migration von SAP S/4HANA-Systeme in die Public Cloud.

---

## Literatur

1. <https://www.faz.net/aktuell/feuilleton/medien/groesstes-rechenzentrum-europas-brennt-komplett-nieder-17241629.html> (Zugriff am 20.12.2021).



# Deployment und Migration von SAP S/4HANA-Systemen

3

## Zusammenfassung

Bevor ein SAP S/4HANA-System in der Public Cloud provisioniert werden kann, müssen einige Vorbedingungen erfüllt sein und einige Rahmenbedingungen geprüft werden. Hierzu gehören neben der wichtigen vertraglichen Basis mit dem Hyperscaler auch die gesetzlichen Anforderungen, wie z. B. GxP, welche durch die Hyperscaler unterstützt wird. Die Auswahl eines Hyperscalers erfolgt oftmals basierend auf einer Liste von quantitativen und qualitativen Faktoren. Nach der Auswahl eines Hyperscalers kann ein Deployment oder eine Migration von SAP S/4HANA-Systemen direkt in der Cloud erfolgen. Bei einer Migration eines SAP ERP-Systems in die Cloud wird eine Transformation notwendig, wenn Unternehmen zu einem SAP S/4HANA-System migrieren wollen. Hierfür werden die Brownfield und Bluefield-Ansätze zur Verfügung gestellt. Als Alternative zur Provisionierung eines eigenen Systems können auch die Angebote der SAP aus der Cloud heraus genutzt werden. Hierbei sind die Vor- und Nachteile einer Nutzung durch die Unternehmen abzuwägen.

## 3.1 Rahmenbedingungen

Vor dem Einsatz einer neuen Public Cloud muss die Zielplattform ausgewählt werden. Dies geschieht nicht nur nach technischen Gesichtspunkten. Es spielen auch nicht-technische Aspekte, wie der Vertrag, die regulatorischen Anforderungen und andere qualitative Faktoren eine Rolle, welche hier beschrieben werden.

### 3.1.1 Vertragliche Basis und Support

#### Wichtige Faktoren für den Vertrag

Als in den späten 2000er Jahren die Public Clouds aufkamen, war es sehr einfach, mit der Nutzung zu beginnen. Alles, was benötigt wurde, war eine Kreditkarte und ein Account. Danach konnte es schon losgehen und Kunden konnten sich die ersten virtuellen Maschinen provisionieren. Dies funktioniert heute auch noch und so ist die Hürde für einen Einstieg in die Welt der Public Cloud sehr gering.

Im Privatkundenbereich ist dieses Vorgehen absolut richtig, da Privatkunden durch zu viele Hürden bei der Nutzung von Applikationen und Services abgeschreckt werden. Hier muss es einfach sein und durch ein paar Klicks sollten alle Punkte erledigt sein, sodass der Nutzung nichts im Wege steht. Im Umfeld der Unternehmen jedoch ist dies kein gangbarer Weg. Hier gilt es wichtige Fragestellungen vorher zu adressieren, bevor die ersten SAP S/4HANA-Systeme in der Public Cloud provisioniert werden:

1. Wie lautet die vertragliche Basis für das Hosting der SAP-Systeme?
2. Welche Service Levels (z. B. Verfügbarkeit, Erreichbarkeit, Performance) bietet der Hyperscaler an und welche Pönalen werden angeboten?
3. Welche Haftung wird vom Hyperscaler angeboten?
4. Welcher Support ist notwendig und welche Stufen im Support werden angeboten?
5. Wie sieht die Preisstruktur aus und existieren Credits für die Nutzung?

Oben sind nur einige der wichtigsten Fragen aufgeführt, welche es vor einer Nutzung zu adressieren gilt und deren Antworten sich in den Verträgen widerspiegeln muss.

In der herkömmlichen Welt der Service Provider konnten Kunden einen individuellen Vertrag aushandeln, der alle Punkte idealerweise so adressierte, wie der Kunde es benötigte. Dabei waren die Service Provider sehr flexibel und sind auf die individuellen Bedürfnisse der Kunden eingegangen. Bei der Public Cloud jedoch stellt sich dies anders dar. Die Hyperscaler verfolgen hier eine andere Strategie: **Alle Kunden erhalten denselben Vertrag.**

Aus Sicht der Hyperscaler ist dies sinnvoll, da damit eine Vielzahl von Kunden alle einen gleichlautenden Vertrag haben und der Hyperscaler keine Unterschiede machen muss. Bei der Vielzahl der Kunden ist es wichtig, denselben Vertrag zu nutzen. Aus Sicht der Kunden ist dies sicherlich gewöhnungsbedürftig. Die Punkte, welche in den Verträgen mit den Service Providern individuell geregelt waren, sind in den Verträgen mit den Hyperscalern nicht verhandelbar. Diese Vereinbarungen sind maximal inflexibel, was jedoch auch dem Naturell der Public Cloud entspricht.

**Amazon** Web Services bietet den Kunden die Vereinbarung zum Download im Internet an. Der Link lautet wie folgt (Zugriff am 20.12.2021): <https://aws.amazon.com/agreement>

Es ist wichtig zu erwähnen, dass die Nutzungsvereinbarung auch in anderen Sprachen verfügbar ist, jedoch die englische Version der Vereinbarung immer Vorrang besitzt. Bei einem rechtlichen Disput wird in jedem Fall die englische Version als Vertragsgrundlage herangezogen. Amazon regelt in der Nutzungsvereinbarung alle wichtigen Punkte für die Nutzung der AWS Services. Hierbei wird auch explizit Wert auf die Pflichten der Kunden (wie z. B. Einhaltung der Rechte etc.) gelegt.

**Microsoft** verfolgt einen anderen Weg und möchte den Kunden, die Cloud Services in dem übergreifenden Enterprise Agreement anbieten. Das Enterprise Agreement ist ein übergreifender Vertrag zwischen Microsoft und dem Kunden, welche üblicherweise ab 500 Benutzern startet. Das Agreement wird beispielsweise für die Nutzung von Office365 unterschrieben, wenn Kunden die Arbeitsplatz-Lösung von Microsoft nutzen. Das Enterprise Agreement gilt als Rahmenvertrag und kann erweitert werden. Nutzt ein Kunde bereits die Services von Office365, kann der Kunde die neuen Services zur Public Cloud durch das «Server and Cloud Enrollment» nutzen. Die einfache Erweiterung des Enterprise Agreements mit dem SCE erklärt auch, warum viele Kunden auf die Microsoft Azure-Cloud gesetzt haben. Da es bereits ein Enterprise Agreement gab, konnten die Kunden sehr leicht die vertragliche Basis für die neuen Services schaffen. Dies war mit sehr viel weniger Aufwand verbunden, als eine neue Vereinbarung mit einem neuen Provider zu unterzeichnen.

**Google** verfolgt einen ähnlichen Weg, wie Microsoft. Hierbei müssen Kunden zunächst eine Google Cloud-Rahmenvereinbarung unterzeichnen, welche die grundlegenden Punkte zwischen dem Kunden und Google regelt. Danach wird für die Nutzung der verschiedenen Services jeweils eine separate Vereinbarung unterschrieben. Das sind die Service-spezifischen Schedules. So existieren spezifische Vereinbarungen für die Nutzung von Google Workplace oder auch die Google Cloud. Der Link zu den stets aktuellen Dokumenten von Google lautet wie folgt (Zugriff am 20.12.2021): <https://cloud.google.com/terms/service-terms/index.html?hl=de>

### **Direkte Nutzung oder Nutzung über einen Partner**

Es existieren zwei Wege zur Nutzung der Cloud. Einerseits können die Kunden den direkten Vertrag mit den Hyperscalern nutzen. Andererseits können die Kunden die Services der Hyperscaler über einen Partner einkaufen (Resale). Beide Möglichkeiten haben Vor- und Nachteile.

Das Szenario, dass der Kunde direkt mit Hyperscaler kontrahiert, ist das übliche Szenario. Hierbei unterzeichnet der Kunde die Verträge und hat damit auch die volle Kontrolle über die Nutzung, die Kosten, die Architektur und alle anderen Aspekte. Der Nachteil liegt aber auch genau in diesem Aspekt: Kunden haben die volle Kontrolle und müssen über eine entsprechende Governance diese Kontrolle auch ausführen. So muss beispielsweise die Architektur der SAP-Systeme zu den Anforderungen passen und die Sicherheit der Cloud-Umgebung entsprechend den Vorgaben umgesetzt sein. Der Kunde ist für alle Aspekte der Nutzung verantwortlich und muss beispielsweise auch die

Gespräche zu Pönenaln führen, wenn ein Service in der Cloud nur eingeschränkt verfügbar war.

Das Szenario, dass der Kunde die Service von einem Partner einkauft, ist mittlerweile seltener geworden. Hierbei lagen die Vorteile des Kunden insbesondere in einer Verlagerung der Verantwortlichkeiten. Zu Beginn des Public Cloud-Trends, als die ersten Early Adopters mit den SAP-Systemen in die Public Cloud zogen, gab es nur begrenzt Erfahrungen und Wissen aufseiten der Kunden. Hierfür nutzen die Kunden die Dienstleister, um die Cloud-Umgebungen aufzubauen, diese zu verwalten und zu steuern. Somit waren die Dienstleister in der Pflicht und Verantwortung für die Umgebung. Darüber hinaus boten manche Dienstleister auch an, die Service Levels für die eigenen Services und die Services der Cloud zu kombinieren. Dies hatte den Vorteil für den Kunden, dass nur ein SLA-Regime zu steuern war (und nicht zwei: mit dem Cloud Provider und dem Dienstleister). Der große Nachteil bei diesem Szenario ist jedoch der finanzielle Mehraufwand aufseiten des Kunden. Die Dienstleister haben für die Verwaltung der Verträge, als auch die Vereinheitlichung der SLAs, einen Aufschlag auf die Services der Cloud-Anbieter aufgerechnet. Wurde beispielsweise die virtuelle Maschine für 100 EUR je Monat vom Cloud-Provider angeboten, so konnten daraus schnell 120 EUR je Monat vom Service Provider werden. Vielen Kunden waren diese Mehrkosten zu viel, sodass dieses Modell des Wiederverkaufs der Services stark abgenommen hat.

### Notwendiger Support

Alle Hyperscaler bieten unterschiedliche Supportverträge/Supportmodelle an, wobei die niedrigste Stufe des Supports für alle Kunden in den Nutzungsentgelten bereits inkludiert ist. Dieser ist üblicherweise auch kostenfrei. Das niedrige Support-Level garantiert aber, dass Kunden immer Tickets öffnen können und somit Zugriff auf die Supportorganisation bei Fehlern/Problemen haben.

Dieser grundlegende Support reicht sicherlich für Privatkunden aus, ist jedoch nicht geeignet, wenn produktive SAP S/4HANA-Systeme auf den Public Clouds betrieben werden. Dann ist ein Support mit folgenden wichtigen Eigenschaften ratsam:

- **24 x 7 Support** für die Lösung von technischen Problemen für Produktionssysteme
- **Schnelle Antwortzeiten** für geöffnete Supporttickets bei Auswirkungen auf die Produktionssysteme
- Zugriff auf einem Kunden **zugeordneten Manager**, über den langlaufende Tickets eskaliert werden können
- Support bei der **Architektur** und regelmäßige Reviews der SAP S/4HANA-Systemlandschaft
- **Schulungen/Webinare/Whitepaper und Leitfaden**

Insbesondere die ersten Punkte sind sehr wichtig bei Produktionsumgebungen. Startet ein Unternehmen zunächst mit einigen Sandbox oder Trainingssystemen, kann ein

minimaler Support schon ausreichend sein. Wenn jedoch die ersten Produktionssysteme auf die Public Cloud migriert werden, sollte der Support kontrahiert sein.

## DSVGO

Ein besonderes Augenmerk sollte durch alle Kunden auf die DSVGO gerichtet werden. Hierbei gilt es, die gültigen aktuellen rechtlichen Anforderungen nach der Datenspeicherung und Datenhaltung plus Datenverarbeitung zu berücksichtigen. Diese Punkte müssen durch alle Kunden in den Verträgen mit den Hyperscalern hinreichend gut abgedeckt und adressiert sein. Sind die Punkte aus Sicht der Rechtsabteilung eines Kunden nicht adressiert, so muss dies mit dem Hyperscaler diskutiert und angepasst werden. Im Regelfall sollte dies nicht der Fall sein, jedoch kann es in Ausnahmefällen zu Diskussionen hierüber kommen.

## Unterstützung bei Projekten

Einige der großen Hyperscaler unterstützen die Kunden bei der Nutzung der Cloud und bei der erstmaligen Migration von SAP S/4HANA-Systemen in der Cloud. Hierzu legen die Hyperscaler Programme auf, welche die Kunden motivieren sollen, in die Clouds zu migrieren und diese so lange, wie möglich, zu nutzen.

Der initiale Schritt der Migration von SAP S/4HANA-Systemen in die Cloud, kann ein sehr aufwendiger Schritt sein. Kunden müssen Architekturen festlegen, die vertragliche Basis muss geschaffen werden, viel Zeit wird für Abstimmungen und erste Gehversuche benötigt. Um diese Aufwände geringer ausfallen zu lassen, unterstützen die Hyperscaler die Kunden hierbei und liefern schon fertige, einsatzbereite Architekturen und Blueprints mit. Für die Ausführung die Migration der ersten SAP-Systeme in die Clouds unterstützen die Hyperscaler die Kunden/Partner der Kunden monetär. Auf diesem Wege wird der anfängliche Schmerz durch die Migration in die Cloud geringer gehalten.

Der Betrieb von SAP S/4HANA-Systemen kann durch die sehr hohen Anforderungen der Systeme schnell teuer werden. Insbesondere die SAP HANA-Maschinen sind Kostentreiber in den SAP-Umgebungen. Die Hyperscaler können den Kosten hier jedoch auch etwas entgegenwirken und Service Credits zur Verfügung stellen. Diese Service Credits können dann nach der Migration der Systeme auf die Cloud durch den Kunden konsumiert werden. Hier sollten Kunden darauf achten, dass die Credits auch für die migrierten Systeme genutzt werden können und nicht ausschließlich für neue SAP-Systeme gelten.

Bei solchen Angeboten durch die Hyperscaler muss klar sein, dass das Ziel der Hyperscaler darin besteht, die Kundenbeziehungen aufzubauen und die Kunden langfristig an die Clouds der Hyperscaler zu binden. Es ist offensichtlich, dass SAP-Systeme, welche bereits auf der Cloud laufen, nicht so schnell wieder aus der Cloud migriert werden.

### 3.1.2 GxP Regulatorien und wichtige Zertifikate

GxP Regulatorien sind für viele Unternehmen wichtige Rahmenbedingungen, unter denen SAP-Systeme zu betreiben sind. Dieses Teilkapitel gibt hierzu einen Überblick.

#### 3.1.2.1 GxP allgemein

Unternehmen aus den Bereichen Life Science oder auch Konsumgüterhersteller unterliegen oftmals speziellen Anforderungen hinsichtlich des Qualitätsmanagements der Produktion, aber auch der IT-Services. Diese speziellen Anforderungen für die Industrien werden in den GxP Leitsätzen festgehalten. Hierbei kann das «x» in GxP für verschiedene Industrien stehen, wie z. B. GMP für die herstellende Industrie. Neben der herstellenden Industrie existieren die GxP Grundsätze noch für die folgenden Industrien:

- GAP – Good Agricultural Practice
- GMP – Good Manufacturing Practice
- GDP – Good Distribution Practice
- GCP – Good Clinical Practice
- GCLP – Good Clinical Laboratory Practice
- GLP – Good Laboratory Practice
- GAMP – Good Automated Manufacturing Practice
- GDocP – Good Documentation Practice
- GEP – Good Engineering Practice
- GSP – Good Scientific Practice
- GVP – Good Pharmacovigilance Practice

Bei einer Migration von SAP-Systemen in die Hyperscaler Clouds, müssen die Unternehmen Sorge tragen, dass die Grundsätze auch auf der Public Cloud erfüllt werden. Jedoch existiert ein Unterschied zu der traditionellen On-Premise-Welt: Während die Unternehmen in der On-Premise-Welt die Infrastruktur der Rechenzentren und der Server unter ihrer Kontrolle hatten, haben sie die Infrastruktur der Hyperscaler nicht unter Kontrolle. Daher ist es wichtig, ob ein Hyperscaler GxP unterstützt und sich somit zu den Grundsätzen bekennt oder nicht.

#### 3.1.2.2 Trennung der Verantwortlichkeiten

Generell existiert eine natürliche Trennung der Verantwortlichkeiten bei S/4HANA Systemen auf der Hyperscaler Cloud. Dies ergibt sich dadurch, dass die Hyperscaler keinen Zugriff auf die eigentlichen Kundendaten und Kundensysteme haben. Die Hyperscaler können also nur eine Teilmenge der Komponenten kontrollieren und somit auch die GxP Compliance sicherstellen. Eine Trennung zwischen Hyperscaler und Kunde kann wie in Tab. 3.1 aufgelistet werden.

Durch die Trennung der Verantwortlichkeiten in Tab. 3.1 wird die Wichtigkeit der Kenntnisse der SAP-Systeme auf der Hyperscaler Cloud ersichtlich. Die Hyperscaler

**Tab. 3.1** Trennung der Verantwortlichkeiten bei GxP

Komponente	Inhalt	Verantwortlichkeit
Rechenzentrum	Infrastruktur der Rechenzentren (inklusive z. B. Zugangskontrollen)	Hyperscaler
Server	Serverinfrastruktur und Sicherheit der physikalischen Server	Hyperscaler
Storage	Infrastruktur der Storagekomponenten, inklusive RAID-Verbünden, Anbindung der Storagekomponenten und Redundanzen	Hyperscaler
Netzwerk	Physische Komponenten und Verkabelung der Netzwerkkomponenten (inklusive Server und Storage)	Hyperscaler
Backbones	Anbindung der Rechenzentren untereinander als auch der Regionen (z. B. Frankfurt zu London) untereinander über das Netzwerk des Hyperscalers	Hyperscaler
Betriebssystem	Setup, Konfiguration und Sicherheit des Betriebssystems	Kunde
HANA Datenbanken	Setup, Konfiguration, User und Access Management etc.	Kunde
SAP Applikation	Betrieb, Management und Kontrolle der technischen SAP Komponenten (SAP Central Services, Applikationsserver, etc.)	Kunde
SAP Daten und Programme	Daten des SAP-Systems und die ABAP/JAVA Programme des SAP-Systems	Kunde
SAP Zugänge	Benutzer, Passwörter, Rollen und Profile der Kunden	Kunde
Firewalls und Netzwerk	Alle virtuellen Komponenten auf der Hyperscaler Cloud	Kunde
Verschlüsselung	Jegliche Verschlüsselung des Netzwerksverkehrs, des Storage, der Backups und sonstige Konfigurationen	Kunde

Verantwortlichkeiten in der Public Cloud

haben auf die vom Kunden auf der Cloud betriebenen Systeme keinen Zugriff und können somit auch die Sicherheit und die Compliance der Systeme nicht sicherstellen. Dies muss durch die Kunden als Schritte der Verifikation und Validation erfolgen.

Die Hyperscaler unterstützen die Validierung und Verifikation der Kunden und ihrer Systeme nicht individuell, sondern steuern die Zertifizierungen bei. Es ist anzuraten, mit dem gewählten Hyperscaler direkt Kontakt aufzunehmen und die GxP Qualifizierung durchzusprechen. Danach kann individuell geregelt werden, welche Verantwortlichkeiten beim Hyperscaler liegen und welche beim Kunden und/oder Service Provider.

### **3.1.2.3 Weitere Zertifizierungen**

Neben der Unterstützung für GxP, haben die großen Hyperscaler weitere Zertifizierungen, welche wichtig für Kunden werden können. So müssen einige Kunden bei Audits auch zeigen können, dass die Hyperscaler Clouds entsprechende Zertifizierungen vorweisen können (Contractor).

#### **SOC 1 und 2**

Die Service and Organization Control Reports 1 und 2 wurden als Zusammenfassungen der verschiedensten Reports erstellt – zum Beispiel der SAS70, welche später durch die SOC1 Reports abgelöst wurden. Ähnliche Namen der Reports sind SSAE18 als US-amerikanische Version, ISAE3402 als europäische Version, sowie ASA3402 als australische Version der Reports. Alle Reports zeigen die effektiven Kontrollen von Maßnahmen auf den internen Prozessen. Die SOC 2 Reports fokussieren die IT-Sicherheit und sind von dem Standard AT101 abgeleitet.

Betreiben Kunden ihre S/4HANA Systeme auf den Hyperscaler Cloud, müssen sie fähig sein, diese Reports bei Audits vorzulegen. Daher ist es elementar, dass SOC1 und SOC2 als standardisierte Reports verfügbar sind und von den Hyperscalern zur Verfügung gestellt werden.

#### **HIPAA/HITECH/HITRUST**

Unternehmen, welche dem U.S. Health Insurance Portability and Accountability Act (HIPAA) aus dem Jahre 1996 unterliegen, müssen gewissen Standards bei der Prozessierung der Daten erfüllen. Dies bedingt nicht nur die Verarbeitung von Daten durch das selbige Unternehmen, sondern auch die Prozessierung der Daten durch Service Provider und die Speicherung der Daten. HIPAA umfasst die Regelungen zu den wichtigsten Daten aus dem Gesundheitswesen, wie der klinischen Pflege, Labordaten oder auch Testergebnisse. Die HIPAA Anforderungen wurden später durch die Anforderungen aus dem Health Information Technology for Economic and Clinical Health (HITECH) Act weiter ergänzt.

Alle großen Hyperscaler unterstützen die HIPAA Richtlinien, sind aber selber nicht zertifiziert, da sich die Richtlinien an den Betrieb und das Setup richten. Die Hyperscaler gelten im Sinne von HIPAA als die sogenannten «Business Associates» – die Geschäftspartner. Mit den Hyperscalern müssen entsprechende Verträge im Sinne von HIPAA geschlossen werden. Diese Verträge Business Associate Addendums Schließen die Kunden mit den Hyperscalern, um die HIPAA Compliance sicherzustellen.

Es existierte eine Zeit lang die Anforderung, dass HIPAA-unterliegende Kunden, dedizierte Hosts und Instanzen zu nutzen hätten. Es war möglich, diese Anforderungen umzusetzen, jedoch widersprach diese dem Gedanken der Public Cloud und der leichten Anpassung der Workloads. Die Anforderung nach dedizierten Hosts und Instanzen ist seit einiger Zeit zwar gefallen, jedoch wird Kunden empfohlen, sich mit den Bedingungen und Anforderungen, zur Vermeidung von Lücken, vertraut zu machen. Manche Hyperscaler, wie zum Beispiel Microsoft mit Azure, bieten auch Blueprints an,

über die Kunden, sich mit einer beispielhaften Architektur und allen notwendigen Regeln für HIPAA/HITECH vertraut machen können.

Als letztes wichtige Framework in der Gesundheitsbranche gilt das HITRUST, welches als Erweiterung von HIPAA und HITECH zu sehen ist. Es nutzt beide Acts als Grundlage und erweitert diese durch Anforderungen von anderen regulatorischen Frameworks, wie z. B. ISO27001 oder auch MARS-E. Für HITRUST gibt es derzeit noch keine breite Zertifizierung durch die Hyperscaler, da es noch nicht sehr lang existiert. Aktuell (mit Stand 2021) bietet nur Microsoft Azure die Zertifizierung für HITRUST an.

### **ISO Zertifizierungen**

Neben den industriespezifischen Zertifizierungen und die Unterstützung zu solchen Zertifizierungen, sind die ISO Zertifizierungen für Kunden sehr wichtig. Die Internationale Organisation für Normung (International Organization for Standardization=ISO) ist eine nicht-staatliche Organisation, welche durch ihre Normen und Standards eine sehr große Reichweite erlangt hat. Sie hat über 163 Gremien für die Normierung von verschiedenen Themen.

Über die Norm ISO 27001 werden die wichtigsten Standards für den Umgang und die Nutzung eines Informationssicherheitsmanagementsystems (ISMS) definiert. Dazu gehören auch Best Practices, welche Unternehmen idealerweise befolgen sollten. Die Norm definiert keine Maßnahmen, sondern den Rahmen davon und somit ist es den Unternehmen überlassen, die Maßnahmen umzusetzen.

Die drei großen Hyperscaler Amazon Web Service, Google Cloud und Microsoft Azure sind alle drei für ISO 27001 zertifiziert. Dies ist die Grundlage der Kunden, sich ebenfalls in ISO 27001 zertifizieren zu lassen.

Neben der ISO Zertifizierung 27001 existieren weitere Zertifizierungen, welche die Hyperscaler anbieten. Hier soll exemplarisch für Amazon Web Services die Fülle von Zertifizierungen aufgeführt werden:

- CSA – Cloud Security Alliance
- ISO 9001 (Qualitätsmanagement)
- ISO 27001 (Sicherheitsmanagement)
- ISO 27017 (Cloud-spezifische Kontrollen)
- ISO 27018 (Schutz personenbezogener Daten)
- PCI DSS Level 1 (Standard für Payment Card)
- SOC 1 (Prüfungskontrollbericht)
- SOC 2 (Sicherheits-, Verfügbarkeits- und Vertraulichkeitsbericht)
- SOC 3 (Allgemeiner Kontrollbericht)

Die anderen Hyperscaler bieten ähnliche Zertifizierungen. Dies sollte durch die Kunden entsprechend vorher geprüft werden. Da sich die Lage der Zertifizierungen sehr schnell ändert, wird auf eine Auflistung für Microsoft Azure und die Google Cloud verzichtet.

**Tab. 3.2** ISO 27017 Cloud Kontrollen

Nummer der Kontrolle	Kontrolle	Beschreibung
CLD 6.3.1	Geteilte Rollen und Verantwortlichkeiten	Die Rollen und Verantwortlichkeiten aufseiten der Kunden und der Cloud Anbieter bei den Sicherheitskonzepten muss dokumentiert, kommuniziert und implementiert sein
CLD 8.1.5	Entfernen von Assets der Kunden	Jegliche Assets der Kunden sollen in einem gewissen zeitlichen Rahmen nach einer Vertragsbeendigung entfernt oder an den Kunden zurückgegeben werden
CLD 9.5.1	Abschottung der Kundenumgebungen	Alle Kundenumgebungen innerhalb einer Cloud müssen voneinander geschützt und abgeschottet sein. Kunden dürfen keinen Zugriff auf andere Umgebungen erhalten
CLD 9.5.2	Härtung der virtuellen Maschinen (Hardening)	Alle virtuellen Maschinen müssen gehärtet sein, um die notwendige Sicherheit zu gewährleisten
CLD 12.1.5	Sicherheit bei der Administration	Jegliche administrativen Aktivitäten und Prozeduren auf der Cloud müssen definiert, dokumentiert und implementiert sein
CLD 12.4.5	Monitoring der Cloud Services	Kunden der Cloud Services sollten die Möglichkeit haben, die Services zu überwachen
CLD 13.1.4	Netzwerksicherheit in den virtuellen und physikalischen Netzwerken und Verbindungen	Die Konsistenz der Verbindungen, der Sicherheitsrichtlinien und der Implementierungen muss zwischen den virtuellen und physikalischen Netzwerken sichergestellt sein

#### Spezielle Cloud Kontrollen

An dieser Stelle soll noch auf die ISO 27017 Norm eingegangen werden. Sie listet zusätzliche Kontrollen für einen Einsatz von Public Cloud-Services, welche für Kunden wichtig sind. Neben den normalen ISO 27001 Kontrollen, kommen weitere Kontrollen hinzu, welche in Tab. 3.2 aufgelistet sind.

Durch die Erweiterung der ISO 27017 Kontrollen werden weitere wichtige Kontrollen für einen Betrieb von SAP S/4HANA Systemen in der Public Cloud geschaffen. Kunden sollten diesen Kontrollen und Grundsätzen idealerweise folgen, um die Sicherheit der Daten und Systeme nicht negativ zu beeinflussen.

### 3.1.3 Management der S/4 Landschaft

In allen SAP-Umgebungen existiert der SAP Solution Manager, um die SAP-Landschaften zu administrieren und wichtige Prozesse, wie beispielsweise das Release Management via ChaRM, erfolgreich einzusetzen. Der SAP Solution Manager ist so tief in die Prozesse und Abläufe integriert, dass er in jeder aktuellen Systemumgebung zu finden ist und stellt damit das Rückgrat des Application Lifecycle Managements in SAP-Umgebungen dar.

#### Wichtigste Features des Solution Managers

Der Solution Manager kann verschiedenste Bereiche im Betrieb und der Wartung von SAP-Systemlandschaften unterstützen. Hierzu zählen die folgenden wichtigsten Funktionen:

- **Maintenance Management** mit dem Maintenance Optimizer und den Systemempfehlungen
- **Template Management** mit Benachrichtigungen für Änderungen
- **Betrieb der Geschäftsprozesse** mit einem BPO Dashboard und einem Job Control und Schedule Management
- **Betrieb der technischen Prozesse** mit dem technischen Monitoring, dem End User Experience Monitoring und auch dem Data Volume Management
- **Dokumentation** der Solution mit Business Blueprints und der Reverse-Dokumentation der Geschäftsprozesse
- **Solution Implementation** mit Business Blueprint Modellen und Ende-zu-Ende Implementierungsprozessen für die Geschäftsprozesse
- **Upgrade Management** mit der Möglichkeit zur Analyse der Abhängigkeiten bei einem Upgrade
- **Incident Management** für die Applikationen mit einem Framework für automatisiertes Testen unter Nutzung von Drittanbietertools
- **Test Management** mit dem Business Process Change Analyzer
- **Change Control Management** mit CTS +

Da der Solution Manager in der traditionellen Welt der Rechenzentren entwickelt worden ist, hat die SAP mit der Version 7.2 einige wichtige Neuerungen eingeführt, welche den Solution Manager für neue Cloud-Szenarien vorbereitet:

- Der Solution Manager wurde auf **SAP HANA** portiert und es wird kein TREX Server mehr benötigt, sondern die Suchfunktionalität wird durch die HANA-Datenbank erfüllt.
- Der Solution Manager unterstützt nun vollständig **SAP HANA** und **SAP S/4HANA-Systeme**.
- Der Solution Manager unterstützt die **hybriden Cloud-Szenarien** und kann mit Cloud-Ressourcen integriert werden und diese monitoren.

## Application Operation mit dem Solution Manager

Beim Management von größeren SAP S/4HANA-Landschaften ist der SAP Solution Manager essenziell, um den Betrieb zu steuern und die Gesundheit der Systemlandschaft überwachen zu können. Der SAP Solution Manager bietet hier wichtige Funktionen an, welche den Alltag des SAP-Basisadministrators vereinfachen können.

Über ein **technisches Systemmonitoring** können die SAP-Basisadministratoren den genauen Zustand der SAP S/4HANA-System überwachen. Die Einrichtung des Monitorings unterscheidet sich nicht signifikant von den vorherigen Solution Manager Versionen. Es wird etwas Zeit benötigt, um die entsprechenden Grenzwerte zu definieren und die Alarne zu konfigurieren. Wenn dieser initiale Aufwand aber erledigt ist, können über das Systemmonitoring die S/4HANA-Systeme, die HANA-Datenbanken, als auch die darunter befindlichen virtuellen Maschinen und Betriebssysteme überwacht werden.

Als weitere Komponente wurde von der SAP das **Job Monitoring** mit eingeführt. Dies ist verglichen zu den Funktionen der Job Scheduling-Systeme (wie z. B. UC4 von Atomic) eingeschränkt, aber kann den Administratoren helfen, fehlerhafte Läufe von Hintergrundjobs zu erkennen. Dennoch nutzen Unternehmen eher die Funktionalitäten der Job Scheduling-Systeme.

Neben dem Monitoring der technischen Komponenten und der Jobs, hat SAP den Bereich des Monitorings für **Business Intelligence/Business Warehouse** stark ausgeweitet. Hierbei können Kunden zentral die SAP Business Warehouse und die SAP Business Intelligence-Lösungen monitoren. Darunter fallen alle wichtigen Services, wie System Landscape Transformation (SLT), Data Service Systeme als auch die Business Warehouse Accelerators (BWA). Somit erweitert die SAP die Möglichkeiten, über den Solution Manager neben den technischen Punkten, auch die inhaltlich wichtigen Themen zu überwachen.

Neben den rein technischen Punkten zur Überwachung, hat die SAP auch ein **Process Integration (PI) Monitoring** implementiert. Hierüber lässt sich eine zentrale Überwachung der Channels, Messages und der Systemkomponenten der PI-Systeme durchführen. Da die PI-Systeme neben den S/4HANA-Systemen zu neuralgischen Punkten in den Systemlandschaften geworden sind, bedarf es hier einer genauen Überwachung und so müssen die Administratoren über nicht gesendete Messages schnell informiert werden. Das PI-Monitoring des SAP Solution Managers kann dies adressieren.

Eine weitere wichtige Komponente für ein erfolgreiches Systemmonitoring ist die Komponente der **Root Cause Analysen**. Dies ist eine Funktion im Solution Manager, welche schon seit einigen Jahren existiert und sukzessive ausgebaut worden ist. Hierüber können Performanceprobleme in den S/4HANA-Systemen, als auch Nichtverfügbarkeiten (Downtimes) untersucht werden und die auslösende Komponente identifiziert werden. Für die erfolgreichen Analysen ist es wichtig, dass alle Komponenten korrekt eingerichtet und konfiguriert sind. Der SAP Solution Manager kann auch Produkte von Drittanbietern integrieren.

Für den professionellen IT-Betrieb der S/4HANA-Systemlandschaft bietet der SAP Solution Manager auch eine Integration in die etablierten **ITSM-Systeme**, wie z. B.

ServiceNow, an. Hierüber können für die erkannten Probleme in der SAP-Systemlandschaft, die Incidents in den ITSM-Lösungen erstellt werden. Somit muss kein manueller Transfer der Probleme zu Incidents stattfinden. Das stellt eine Alternative zu der Verwendung der eigenen ITSM-Funktionen aus dem Solution Manager dar. Viele Kunden besitzen ein übergreifendes ITSM-Tool und nutzen dieses für eine zentrale Steuerung aller ITSM-Prozesse. In solch einem Fall werden die Alarmierungen aus dem Solution Manager an das zentrale ITSM-Tool weitergegeben.

In Summe ist das Application Monitoring des Solution Managers sehr wichtig, um den Zustand der Systemumgebung holistisch zu überwachen. Durch die integrierte Sicht aller wichtigen Komponenten (technisch, als auch applikatorisch), leistet der Solution Manager einen wichtigen Beitrag zum stabilen Betrieb der SAP-Systeme.

### Nutzung von SAP Cloud ALM

Durch die weite Verbreitung von Cloud-Diensten sah sich die SAP der Herausforderung gegenübergestellt, dass einige Kunden nur noch Cloud-basierte Dienste der SAP (wie zum Beispiel Ariba) nutzen. Diese Kunden wollen für eine effiziente Überwachung keine eigene Lösung, wie den Solution Manager, aufbauen, sondern benötigen eine schlanke Lösung.

Die SAP startete mit SAP Cloud ALM ein neues Produkt, welches als Teil der SAP Business Technology Platform fungiert und den Kunden die Möglichkeit bietet, über eine cloudbasierte Lösung ein Application Lifecycle Management (ALM) durchzuführen. Um die Möglichkeiten für SAP Cloud ALM nicht zu sehr einzuschränken, bereitet die SAP das Produkt darauf vor, auch in hybriden Szenarien eingesetzt zu werden. Somit sollen die On-Premise Systeme genauso überwacht werden können, wie die SAP-Cloudlösungen. Ein klarer Fokus jedoch besteht auf SAP-Systeme generell und die Cloud-Lösungen der SAP und weniger auf die Integration von Drittanbietersoftware.

Viele Kunden stellen sich nun die Frage, welche der Lösungen (Solution Manager oder Cloud ALM) der richtige Ansatz ist. Hierzu werden im Folgenden die Funktionen der beiden Lösungen miteinander verglichen (Tab. 3.3):

SAP Cloud ALM stellt sich als gute Alternative zur Steuerung von vorrangig Cloud-basierten Systemumgebungen dar, welche auf Ariba, Fieldglass, SAP Marketing Cloud oder auch der SAP S/4HANA Cloud basieren.

Für die Nutzung von Cloud ALM gibt es mehrere Wege.

- Sofern Kunden der SAP bereits Cloud-Produkte nutzen (beispielsweise Ariba), können diese Kunden direkt als Teil des SAP Enterprise Supports nutzen.
- Sofern Kunden noch keine Cloud-Produkte nutzen, müssen die Terms & Conditions der SAP Cloud unterzeichnet werden und ein Cloud-Produkt der SAP bezogen werden. Dann erhalten diese Kunden auch den Zugriff auf Cloud ALM.

Aktuell existiert keine Version von Cloud ALM zur Evaluation oder zum Testen.

**Tab. 3.3** Vergleich Solution Manager zu Cloud ALM

Funktionalität	Solution Manager	Cloud ALM	Erläuterung
Cloud Application Management	Gering	Vollständig	
Change Management	Vollständig	Teilweise	
Projektmanagement	Vollständig	Vollständig	
Prozessanalyse und Governance	Vollständig	Sehr groß	
Key Performance Indicators	Vollständig	Sehr groß	
Kosten (Infrastruktur und Wartung)	Teilweise	Vollständig	Sehr geringe Kosten für Cloud ALM, da durch die SAP bereitgestellt
Ready-to-use	Gering	Vollständig	
Process Monitoring	Gering	Gering	
Unterstützung des Intelligent Enterprise	Gering	Vollständig	
Wartung des Produktes inklusive Upgrades	Gering	Vollständig	Wartung von Cloud ALM komplett durch die SAP
Customer Experience	Vollständig	Teilweise	
Reporting und Analytics	Sehr groß	Gering	
Unterstützung der On-Premise Umgebung	Vollständig	Gering	
Test Suite	Vollständig	Sehr groß	
Focused Build	Vollständig	nicht unterstützt	
Service Desk – ITSM	Vollständig	nicht unterstützt	

Vergleich der Funktionalitäten

## 3.2 Auswahl eines Hyperscalers

Bei der Auswahl eines zukünftigen Hyperscalers/Public Cloud gibt es einige Punkte für die Unternehmen zu beachten. Oftmals existieren bei den Kunden schon Präferenzen für einen/zwei Hyperscaler. Dies basiert aufgrund von Vorerfahrungen oder Projekten, welche sich, unabhängig von einer Gesamtstrategie, schon auf einen Hyperscaler geeinigt hatten. In solchen Fällen können die bereits existierenden Umgebungen später in den neu gewählten Hyperscaler überführt werden oder aber eine Zwei-Hyperscaler-Strategie verfolgt werden. Wird ein Hyperscaler komplett neu ausgewählt, so sollte die Entscheidung anhand von qualitativen und quantitativen Faktoren getroffen werden.

### 3.2.1 Quantitative Faktoren

Bei der generellen Auswahl eines Hyperscalers, gibt es immer wieder ein wichtigstes Kriterium, welches in allen Projekten zum Tragen kommt: die Kosten für einen zukünftigen Betrieb. Hierbei spielen die initialen Setupkosten, aber insbesondere die späteren, laufenden Kosten eine sehr wichtige Rolle.

Oftmals gibt es zwischen den Hyperscalern nur sehr geringe Unterschiede bei den Preisen und die Technologien der Hyperscaler sind, zumindest für SAP S/4HANA Systeme, vergleichbar. Dennoch spielt der zu erwartende Gesamtpreis einer SAP S/4 Umgebung in 99 % der Fälle die größte Rolle bei einer Entscheidung für/gegen einen Hyperscaler. Hierbei gibt es jedoch etliche Faktoren, welchen den Gesamtpreis beeinflussen:

- **Preis je Preisliste:** Oftmals werden Kostenabschätzungen nur auf Basis von den Standardpreisen der Hyperscaler erstellt. Die Kostenschätzungen werden auf Basis der im Internet zugänglichen Kalkulatoren erstellt. Als indikatives Bild der zu erwartenden Kosten ist das geeignet. Jedoch ergibt sich ein Gesamtpreis, der ein falsches Bild vermittelt. Wenn die SAP-Systeme/non-SAP-Systeme 1:1 aus der alten Landschaft in die neue Hyperscaler Umgebung verschoben werden würden, ergäbe sich kein Kostenvorteil. Ein so ermittelter Gesamtpreis muss also verfeinert werden.
- **Discounts/Nachklasse:** Alle Hyperscaler verstehen das aktuelle Marktmomentum, welches auch noch ein paar Jahre anhalten wird. Sie wissen um die aktuellen Bestrebungen der Kunden, in die Hyperscaler Clouds zu wechseln. Daher müssen die Hyperscaler die Kunden anlocken und den Wechsel in die jeweiligen Clouds attraktiv gestalten. Hierzu werden Discounts bereit gestellt, welche den Gesamtpreis verringern. Bei einer Betrachtung der Total Cost of Ownership ist diese Komponente sehr wichtig. Es ist wichtig, dass die Hyperscaler umso größere Discounts ermöglichen können, desto größer die Landschaften der Kunden werden können.
- **Fundings:** Neben den eigentlichen Discounts existieren weitere Programme innerhalb der Hyperscaler, welche die preislche Attraktivität erhöhen. So existieren Programme, welche sich als Ziel gesetzt haben, die Anzahl der SAP-Systeme in den jeweiligen Hyperscalern zu erhöhen. Diese Programme unterstützen die Kunden bei der Migration der SAP-Systeme in die Clouds oder unterstützen sogar bei Neu-installationen. Es existieren ähnliche Programme für weitere Systemtypen und Anwendungsgebiete (wie z. B. für Workplace). Es gilt zu beachten, dass solche Fundings desto größer ausfallen, je mehr Workloads einem Hyperscaler in Aussicht gestellt werden. Es ist demnach wenig sinnvoll, ein Funding für nur ein neues SAP S/4HANA System anzufragen. Vielmehr sollte transparent dargestellt werden, mit wie vielen neuen Systemen gerechnet wird.

- **Reservierungen:** Alle Hyperscaler bieten den Kunden die Option an, eine virtuelle Maschine für längere Zeit zu nutzen und damit eine erhebliche Preisreduktion zu erzielen. Diese Reservierungen/Commitment sind bei allen Hyperscalern zu erzielen. Sie unterscheiden sich jedoch in der Art und Weise, wie sie an die einzelnen VMs zu gewiesen werden können. So kann ein Commitment beispielsweise bei Microsoft für eine Klasse von virtuellen Maschinen erstellt werden.

Nur unter Berücksichtigung der obigen Faktoren, kann ein realer Gesamtpreis kalkuliert werden, welcher als wichtigster Entscheidungsfaktor gilt.

#### Preissenkungen

Die Hyperscaler senken die Preise auf regulärer Basis. So fallen die Preise zwar nicht signifikant, jedoch im marginalen Bereich auf jährlicher Basis. Starke Preissenkung treten bei gravierenden Ereignissen auf. So wurden die Preise von AWS und MS Azure signifikant gesenkt, als Google vor ein paar Jahren in den Markt eintrat. Beide Kontrahenten sahen sich der Situation gegenübergestellt, dass ein neuer Marktteilnehmer mit erheblichem Potenzial, das Geschäft von Amazon und Microsoft zu beeinflussen, in den Markt eintrat. Als Reaktion darauf wurden die Preise auf das Niveau von Google gesenkt, um weiterhin konkurrenzfähig zu bleiben.

Es existieren weitere Ereignisse, welche üblicherweise immer wieder zu Preisanpassungen führen. Dazu gehören die Einführungen von neuen Technologien oder neuen Berechnungsgrundlagen (wie Reservierungen/Commitments). Diese Events führen zu einer Angleichung der Preise bei den Services zwischen den Hyperscalern, auch wenn diese immer wieder bemüht sind, durch kleine Unterschiede eine Differenzierung zu erzeugen.

### 3.2.2 Qualitative Faktoren

Die Liste quantitativen Faktoren bei der Entscheidung für/gegen einen Hyperscaler sind sehr übersichtlich. Die qualitativen Faktoren hingegen können sehr vielfältig sein und im Folgenden werden die wichtigsten und gängigsten Faktoren beschrieben.

- **Strategie:** Ein wichtiger qualitativer Faktor ist die Strategie eines Unternehmens im Umgang mit der Public Cloud. In den ersten Jahren der Public Cloud wurde oftmals auf ein Hyperscaler gesetzt und dieser als der strategische Partner definiert. Hiermit einher gingen alle wesentlichen Vor- und Nachteile bei Einschränkung auf nur einen Hyperscaler. Mittlerweile ändert sich dies und der Großteil der Unternehmen setzt auf eine Dual-Cloud-Strategie. Diese hilft, einem Lock-In zu begegnen und mit beiden Hyperscalern weiterhin zu interagieren.
- **Marktpositionierung:** Die Positionierung der Hyperscaler im Markt ist sicherlich für Kunden wichtig und kann zu einer Entscheidung beitragen. Es muss jedoch angemerkt werden, dass die drei wichtigsten Hyperscaler bei allen gängigen Marktanalysten (Gartner, IDC, Forrester, PAC) immer eng beieinander liegen. Microsoft Azure und Amazon Web Services sind durch ihre lange Historie sehr gut etabliert und konnten sich bereits eine breite und stabile Kundenbasis aufbauen. Dies ermög-

licht auch die sehr gute Positionierung in den Reports der Marktanalysten. Google hingegen ist oftmals in der Position der Herausforderer (Challenger) und trägt somit zu einer Marktdynamik bei. Bei einer Auswahl des Hyperscalers sollten die gängigen Marktanalysten zu Rate gezogen werden.

- **Technische Features und Alleinstellungsmerkmale:** Alle Hyperscaler bieten sehr viele Services (IaaS, PaaS und SaaS) an, welche für Kunden generell interessant sind. Für den Betrieb von SAP S/4HANA Systemen gibt es jedoch nicht allzu viele technische Features, welche als Alleinstellungsmerkmale gelten. Die Unterstützung von sehr großen SAP-Systemen mit HANA war früher tatsächlich ein Merkmal, welche oftmals aber nur Marketingzwecken genutzt wurde. Daneben gibt es jedoch Features für Disaster Recovery, welche für Kunden von SAP S/4HANA Systemen interessant sein könnten. Da sich dies aber sehr schnell ändert, und beispielsweise Microsoft hier sukzessive nachzieht, müssen diese Faktoren sehr gründlich untersucht werden, um einen wirklichen Unterschied ausfindig zu machen.
- **Bestehende Verträge/Enterprise Agreement:** Ein sehr großer Faktor bei der Auswahl der Hyperscaler sind eventuell bestehenden Verträge. Bei einer Vielzahl von Unternehmen existieren meist schon Verträge mit den großen Hyperscalern für ein paar wenige, kleinere Services. Der Aufwand, um von diesen bereits existierenden Verträgen, zu einem erweiterten Vertrag zu kommen, welcher auch den Betrieb von SAP-Systemen abdeckt, ist überschaubar. Vergleicht man den Aufwand für die Erweiterung eines Vertrages mit dem Aufwand für einen komplett neuen Vertrag, so tendieren viele Unternehmen dazu, sich für eine Erweiterung zu entscheiden. Somit werden keine langfristigen, juristischen Prüfungen notwendig.
- **Spezifika der Industrien:** In den vorherigen Kapiteln wurden die unterschiedlichen Regulatoren für die verschiedenen Industrien (wie z. B. GxP) bereits beschrieben. Die Unterstützung solcher Vorschriften, Zertifikaten, als auch der Umsetzung sind für Kunden dieser Industrien natürlich sehr wichtig und entscheidend. Gleichwohl alle Hyperscaler hier die wichtigsten Themen unterstützen, müssen die Industrie-spezifischen Anforderungen genau geprüft werden.
- **Wertschätzung:** Jeder Kunde schätzt es, wenn er eine Bevorzugung erfährt. Kunden fühlen sich dann speziell und wertgeschätzt. Dies ist bei Unternehmen und Kunden der Hyperscaler ähnlich. Somit ist es für die Hyperscaler wichtig, den zukünftigen Kunden auch ein gutes Gefühl nicht nur über die Account-Teams zu vermitteln, sondern auch durch eine Aufmerksamkeit der höheren Führung (Senior Management) der Hyperscaler. Es gab schon Kunden, welche dies zu einem entscheidenden Faktor bei der Auswahl der Hyperscaler gemacht haben.
- **Innovation:** Alle Hyperscaler positionieren die Innovationskraft und den Ausblick bei den Innovationen als wichtige Komponente bei den Entscheidungen für/gegen einen Hyperscaler. Dies sind für viele Unternehmen oftmals wichtige Faktoren, jedoch hat sich gezeigt, dass dies keine kritischen Faktoren sind, welche tatsächlich bei einer Entscheidung «SAP S/4HANA Systeme auf Hyperscaler A/B/C» zum Tragen kommen.

**Tab. 3.4** Wichtigkeit der Faktoren

Faktor	Gewichtung bei der Entscheidung	Anmerkungen
<b>Quantitative Faktoren</b>		
Gesamtkosten (Total Cost of Ownership)	Sehr hoch	
Discounts	Hoch	Teil der Gesamtkosten
Fundings	Hoch	Teil der Gesamtkosten
Reservierungen	Mittel	Teil der Gesamtkosten
<b>Qualitative Faktoren</b>		
Strategie	Sehr hoch	
Marktpositionierung	Mittel	
Features	Mittel	
Bestehende Verträge	Hoch	
Spezifika der Industrie	Sehr hoch	Fungiert als KO-Kriterium
Beachtung	Mittel	
Innovation	Niedrig	

Unterschiedliche Gewichtungen bei einer Entscheidung

Nicht alle obig genannten Faktoren sind bei Entscheidungen gleichgewichtig. Es existieren hierbei wichtige Unterschiede (Tab. 3.4).

Generell werden die Entscheidungen zu den Hyperscalern nach reiflichen Überlegungen getroffen und werden bei vielen Unternehmen durch Scoring Modelle untermauert. Die Entscheidung für/gegen MS Azure, Amazon Web Services oder Google Cloud kann nur individuell getroffen werden, da alle Unternehmen sich in spezifischen Situationen befinden. Die obigen Faktoren können aber als Anhaltspunkte genutzt werden, um eine Entscheidung herbeizuführen.

### 3.3 Deployment und Migration

Die Public Cloud macht es sehr einfach, neue SAP-Systeme zu provisionieren oder sogar direkt aus der Cloud heraus zu konsumieren. Dieses Kapitel zeigt, wie SAP S/4HANA-Systeme in der Cloud provisioniert und migriert werden können.

#### 3.3.1 Notwendigkeit der Transformationen

In den vergangenen Jahren hat die SAP ihre Kunden durch verschiedene Erweiterungen und Erneuerung der SAP-Systeme zu Transformationen gedrängt. Es begann mit den

Enhancement Packages, ging über die Adaption von SAP HANA weiter und mündet derzeit in der Einführung von neuen S/4HANA-Systemen.

Vor der Einführung der Enhancement Packages wurden neue Funktionalitäten primär über Upgrades von Systemen bereitgestellt. Solche Upgrade-Projekte waren aber mit massiven Investitionen in Projekte verbunden und hatten eine große Auswirkung auf das Business. Aufgrund des hohen Aufwands für die Einführung von neuen Funktionen, wurde nach einer Alternative zu Upgrades gesucht. Diese fand die SAP in den Enhancement Packages, welche einfacher einzuspielen waren, aber dennoch größere Projektaufwände verursachten. Für die Enhancement Packages mussten die SAP-Systeme einem Update unterzogen werden, aber die Kunden profitierten von neuen Funktionalitäten und technischen Neuerungen. Diesen Schritt haben viele Kunden vollzogen, um die Systeme aktuell zu halten und von den Neuerungen zu profitieren.

Danach kam sehr große Bewegung in den SAP-Systemlandschaften über den Wechsel zu SAP HANA. Hierbei motivierte die SAP die Kunden zu einem Wechsel zu HANA durch etliche technische Neuerungen; jedoch war der Start für die SAP sehr schwierig. Ein sehr komplexes Lizenziertungsmodell der Features von SAP HANA und der SAP-Systeme verhinderte zu Beginn eine breite Akzeptanz und hohe Adoptionsrate von SAP HANA. Im direkten Vergleich zu ähnlichen Technologien konnte SAP mit HANA noch nicht überzeugen und hat viel Energie in die weitere Entwicklung von HANA gesteckt und die gesamte HANA-Plattform robust und ausgereift gemacht. Mittlerweile basieren hochkritische SAP-Systeme auf HANA-Plattformen. Der Wechsel zu HANA war jedoch für alle Kunden ein erheblich schwieriger Schritt und war mit hohen Aufwänden verbunden. Die Anpassungen am ABAP-Code und die Projektaufwände waren für viele Unternehmen signifikant, jedoch hat ein Großteil der SAP-Kunden die neue Plattform HANA angenommen und betreibt die SAP-Systeme darauf.

Der nächste große – und aktuell anhaltende – Schritt ist der Wechsel zu S/4HANA. Viele Unternehmen stehen derzeit vor der Herausforderung, die SAP-Systeme zu S/4HANA-Systeme zu konvertieren. Die SAP verbindet diesen technischen Schritt nicht nur mit einer Transformation der Technologie, sondern insbesondere auch der Geschäftsprozesse. Hierbei können die Unternehmen aber zwei grundsätzliche Positionen einnehmen und entscheiden, wie mit den SAP-Systemen verfahren werden soll:

1. Entweder das Unternehmen entscheidet sich für die Transformation des bestehenden Systems hin zu S/4HANA
2. Oder das Unternehmen entscheidet sich für einen Neuanfang und besitzt zwei Systemlandschaften: eine alte und eine neue basierend auf S/4HANA

Diese grundsätzliche Diskussion und die grundsätzliche Entscheidung sind abhängig von der Strategie des Unternehmens. Hierbei gibt es kein Rezept oder einen generellen Ratsschlag, der für alle Unternehmen gleichermaßen gelten kann. Jedes Unternehmen muss hierbei seine Situation analysieren und danach die grundlegende Entscheidung treffen.

Ein Großteil der Kunden agiert derzeit in einem hybriden Szenario mit einer alten und einer neuen Systemlandschaft.

### 3.3.2 Neue Ansätze durch die Cloud

Die Bereitstellung von Services in der Public Cloud erfolgt grundsätzlich nach zwei verschiedenen Ansätzen: Greenfield oder Brownfield. Die Auswahl des Ansatzes zum Deployment und der Migration hängt von den Voraussetzungen ab, die die gegebenenfalls vorhandene IT-Architektur mit sich bringt. Damit in Verbindung stehen die Anforderungen an die definierte Zielarchitektur in der Public Cloud. Bei der Auswahl eines Deploymentansatzes wird folglich der Ist-Zustand mit dem Soll-Zustand der zu migrierenden Services verglichen. Zunächst werden der Umfang und die Beschaffenheit der Services in der vorhandenen Umgebung betrachtet. Anschließend wird geprüft, ob diese Services vollständig, teilweise oder gar nicht in die Public Cloud Architektur migriert werden. In letzterem Fall werden die Services im Rahmen des Cloud-Deployments neu aufgebaut.

Die beiden Ansätze, sowie das beschriebene Vorgehen lassen sich auch auf die Bereitstellung von SAP-Umgebungen in der Public Cloud übertragen. Wie sich die Ansätze in Bezug auf SAP voneinander unterscheiden und welche Besonderheiten es bei einer Brownfield-Migration gibt, wird in den nachfolgenden Abschnitten erläutert.

### 3.3.3 Greenfield-Deployment

Unter einer Bereitstellung von Services in der IT-Architektur nach dem Greenfield-Ansatz wird der vollständige Neuaufbau der Service-Instanzen im Rahmen der Migration verstanden. Falls in der vorhandenen IT-Umgebung bereits potenzielle Services vorhanden sein sollten, werden diese nicht in die Umgebung in der Hyperscaler Public Cloud migriert. Das Greenfield-Deployment sieht vor, dass alle identifizierten Services in der Zielumgebung neu aufgebaut werden.

In der Praxis eignet sich der Greenfield-Ansatz in Bezug auf SAP-Anwendungen insbesondere für die Bereitstellung von neuen SAP-Systemen, die in der bestehenden Architektur nicht vorhanden sind. Ein konkreter Use Case kann dabei das Deployment einer SAP S/4HANA Sandbox in der Public Cloud sein. Unter der Annahme, dass die Sandbox bisher nicht im Einsatz war, erfolgt die Bereitstellung nach Greenfield. Der Greenfield-Ansatz kommt im praktischen Umfeld vergleichsweise selten vor, da sich eine Unternehmensarchitektur häufig bereits aus einer Vielzahl von vorhandenen Systemen zusammensetzt, die auch nach der Migration in die Public Cloud weiterhin verfügbar sein sollen, um die Geschäftsfähigkeit aufrecht zu erhalten.

### 3.3.4 Brownfield-Deployment

Im Gegensatz zum Greenfield-Ansatz sieht die Brownfield-Migration eine Übertragung bereits vorhandener Komponenten aus der On-Premise-Umgebung in die Zielarchitektur in der Public Cloud vor. Bei der Migration in die Hyperscaler Cloud Umgebung wird stets mindestens eine Teilmenge des Ist-Zustands der Komponenten in der IT-Architektur berücksichtigt. Während bei einem Greenfield-Deployment alle Services von Grund auf neu in der Zielarchitektur bereitgestellt werden, baut der Brownfield-Ansatz auf der vorhandenen On-Premise-Architektur und deren Komponenten auf. Verallgemeinert wird bei einem Brownfield-Deployment zwischen zwei Varianten unterschieden: homogen und heterogen.

#### Homogene Migrationen

Einerseits kann die Migration eine vollständige Eins-zu-eins-Übertragung von der On-Premise-Landschaft in die Public Cloud beinhalten. Diese Strategie wird als homogene Systemkopie bezeichnet. Bei der Migration nach diesem Vorgehen werden keine Veränderungen, weder auf Betriebssystem- noch auf Datenbankebene vorgenommen. Es wird lediglich eine Replikation der Daten nach dem „Backup-and-Restore“-Prinzip vorgenommen. Darunter wird verstanden, dass das Backup aus der On-Premise-Umgebung erstellt und anschließend in der Public Cloud Infrastruktur wiederhergestellt wird. Daher ist die homogene Migration einfach mit einer geringen Downtime umsetzbar.

Für SAP-Systeme kann die homogene Strategie lediglich unter der Voraussetzung angewendet werden, dass das bereits verwendete Betriebssystem kompatibel mit den unterstützten Betriebssystemen in der Public Cloud Umgebung ist. Bei der Microsoft Azure Cloud werden beispielsweise folgende Systeme für den Betrieb von SAP Workloads unterstützt:

- Microsoft Windows Server,
- Red Hat Enterprise Linux (RHEL),
- SUSE Enterprise Linux (SLES),
- Oracle Linux für Oracle Datenbankmanagementsysteme

Darüber hinaus stellt SAP entsprechende Notes zur Verfügung (SAP Note 1928533, SAP Note 2015553), die die Infrastructure-as-a-Service Elemente, wie Typen und Größen virtueller Maschinen, definieren, die den Betrieb von SAP-Systemen in der Microsoft Azure Cloud unterstützen. Für eine homogene Systemkopie in die Public Cloud müssen diese Eigenschaften der Infrastruktur bereits in der On-Premise-Umgebung vorhanden sein.

#### Heterogene Migrationen

Die alternative Variante zu einer homogenen Migration umfasst die Bereitstellung der identifizierten Komponenten aus der Ist-Architektur in der Public Cloud durch eine Änderung des Zustands. Diese Strategie wird als heterogene Systemkopie bezeichnet.

Bei diesem Vorgehen wird mindestens entweder das Betriebssystem oder das Datenbankmanagementsystem verändert. Im Rahmen der homogenen Migration besteht dadurch die Möglichkeit, Release-Upgrades der SAP-Anwendung durchzuführen. Die Analyse einer potenziellen Fehlerquelle ist dabei zwar komplexer, allerdings ist das Testing effizienter, da das Upgrade im Rahmen der Migration vorgenommen wurde, die folglich ebenfalls getestet wird.

Für die Durchführung einer heterogenen Systemkopie empfiehlt SAP die Nutzung des standardisierten Tools R3load. Mithilfe des Tools werden die Inhalte aus der Quelldatenbank in Dateien exportiert, die anschließend in der Zieldatenbank importiert werden. Ergänzend dazu können darüber hinaus folgende Tools des SAP Software Update Manager (SUM) eingesetzt werden:

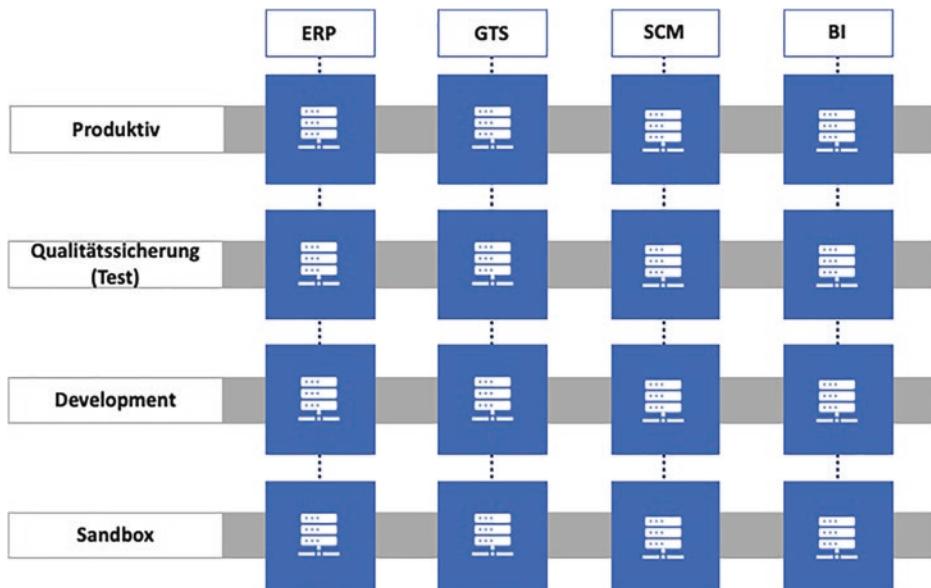
- Software Provisioning Manager (SWPM)
- Database Migration Option (DMO)

Aufgrund der Veränderung des Zustands der Komponenten des SAP-Systems durch beispielsweise Upgrades ist eine heterogene Migration komplexer und bringt eine höhere Downtime mit sich. Ein zentraler Vorteil ist jedoch die erhöhte Effizienz für Veränderungen in den jeweiligen SAP-Systemen.

### **Reihenfolge der Migration**

Bei der Bereitstellung von SAP-Systemen nach dem Brownfieldansatz sind zwei Vorgehensweise zur Festlegung der Reihenfolge definiert, in der die vollständige SAP-Landschaft in die Public Cloud migriert wird. In der Regel setzt sich eine SAP-Landschaft aus verschiedenen SAP-Anwendungen zusammen. Dies können beispielsweise SAP ERP (Enterprise Resource Planning), SAP GTS (Global Trade Services), SAP SCM (Supply Chain Management) und SAP BI (Business Intelligence) sein. Für jede dieser Anwendungen existieren in der Architektur zudem mehrere Umgebungen, die zur Aufrechterhaltung der Konsistenz im Rahmen von Entwicklungen und Systemveränderungen dienen. Nach der Best Practice Empfehlung von SAP wird für jede Anwendung folgende Umgebung eingerichtet: Sandbox, Development, Qualitäts sicherung und Produktiv. Eine SAP-Landschaft eines Unternehmens lässt sich durch folgende Abbildung beispielhaft darstellen (Abb. 3.1).

Für die Festlegung einer Migrationsreihenfolge stehen zwei Vorgehensweisen zur Verfügung. Nach dem Ansatz der horizontalen Migration wird die Reihenfolge zur Übertragung der Landschaft in die Public Cloud pro Umgebung festgelegt. Bei dem horizontalen Vorgehen werden schrittweise die einzelnen Umgebungen der SAP-Systeme gebündelt in die Cloud-Architektur migriert. In Bezug auf die beispielhafte Abbildung würden folglich zunächst die Sandbox-Umgebung und alle darin enthaltenen SAP-Instanzen übertragen werden. Der horizontale Ansatz eignet sich jedoch lediglich für die homogene Migrationsstrategie, um Inkompatibilitäten zur On-Premise-Architektur zu vermeiden.



**Abb. 3.1** Migrationsreihenfolge

Im Gegensatz zur horizontalen Vorgehensweise basiert die Reihenfolge der Migration bei dem vertikalen Ansatz auf den SAP-Anwendungen. Die Migration erfolgt dabei schrittweise pro SAP-Anwendung. Aus der oben dargestellten Abbildung würden bei einer Übertragung in die Public Cloud beispielsweise zunächst alle Instanzen der SAP ERP Anwendung migriert werden. Dieses Vorgehen eignet sich insbesondere für heterogene Migrationen, um die Kompatibilität zwischen den Instanzen einer SAP-Anwendung aufrecht zu erhalten.

### 3.3.5 Migrationsszenarien zur S/4HANA Transformation

Die Migration einer SAP-Landschaft von On-Premise in die Public Cloud Umgebung kann über vier Ansätze mit einer Konversion zu SAP S/4HANA kombiniert werden. Alle Ansätze gehen davon aus, dass in der On-Premise-Umgebung ein SAP ERP System mit einer beliebigen Datenbank gehostet wird, die nicht auf HANA basiert.

Der Ansatz mithilfe eines Greenfield-Deployments stellt die einfachste Transformation zu S/4HANA in der Public Cloud dar. Dabei wird im Rahmen der Datenbankmigration eine Übertragung der bestehenden Datenbank in eine HANA-Datenbank vorgenommen, die in der Public Cloud aufgebaut wird. Dieser Ansatz ist vergleichbar mit einer Konsolidierung des ECC-Systems mit der S/4HANA-Instanz in der Public Cloud Architektur. Die Migration zur HANA-Datenbank und die Konsolidierung finden in einem Schritt statt, wodurch die Ausfallzeit des Systems reduziert wird.

Im Gegensatz dazu kann die Transformation zu S/4HANA in der Public Cloud über verschiedene Zwischenschritte erfolgen. Im zweiten Ansatz wird dabei zunächst die HANA-Datenbank in der Public Cloud Architektur aufgebaut. Die ursprüngliche Datenbank wird in die HANA-Instanz migriert. Alle anderen Komponenten der SAP-Instanz, wie beispielsweise die ERP-Applikationsebene, werden mithilfe von Lift-and-Shift identisch von der On-Premise-Umgebung in die Public Cloud Architektur übertragen. Im ersten Schritt wird dadurch lediglich auf der Datenbankebene eine Änderung vorgenommen. Die Konversion zu S/4HANA erfolgt anschließend im nächsten Schritt.

Die alternativen beiden Ansätze zur Transformation zu S/4HANA in der Public Cloud sehen zunächst eine Übertragung der bestehenden On-Premise-Umgebung in die Hyperscaler Cloud Umgebung vor. Die Migration erfolgt entweder mithilfe von Lift-and-Shift oder Lift-and-Migrate. Das Vorgehen nach Lift-and-Shift setzt die Kompatibilität des Zustands der On-Premise SAP-Instanz mit dem Zustand der Ressourcen in der Public Cloud Architektur voraus. Falls beispielsweise das Betriebssystem voneinander abweicht, muss eine entsprechende Anpassung vorgenommen werden, indem Lift-and-Migrate angewendet wird. Nachdem die Komponenten der SAP ECC Instanz aus der On-Premise-Umgebung in die Public Cloud übertragen wurden, erfolgt im nächsten Schritt entweder die vollständige Transformation zu S/4HANA oder alternativ zunächst über eine Migration der Datenbanktechnologie zu HANA. In letzterem Fall wird die Konversion zu S/4HANA in einem separaten Schritt umgesetzt.

---

## 3.4 SAP S/4HANA als SaaS out-of-the-cloud

Nicht immer ist es zielführend für die Unternehmen, eigene SAP-Systeme zu provisionieren. Der Trend zur Nutzung von Software-as-a-Service-Angeboten setzt sich immer weiter durch und so bietet die SAP auch die Möglichkeit, SAP S/4HANA-Systeme aus der Cloud zu nutzen.

### 3.4.1 Überblick

Das Portfolio der SAP umfasst ein vielfältiges Angebot aus verschiedenen SAP-Lösungen für den Betrieb in einer On-Premise-Umgebung und den Betrieb in der Cloud. Für SAP S/4HANA bietet SAP neben der herkömmlichen Edition zum Eigenbetrieb auch eine Software-as-a-Service Variante an, in der das SAP S/4HANA System in der SAP eigenen Cloud betrieben wird. Allgemein wird bei S/4HANA zwischen On-Premise-Deployment und Cloud-Deployment unterschieden. Die On-Premise-Variante wird in einem eigenen Datenzentrum oder in einer Hyperscaler Public Cloud durch das anwendende Unternehmen selbst betrieben. Die Cloud-Option von S/4HANA wird in

der SAP eigenen Cloud gehostet und es stehen zwei unterschiedliche Editionen zur Verfügung: SAP S/4HANA Cloud Essentials Edition und SAP S/4HANA Cloud Extended Edition. Der zentrale Unterschied zwischen diesen Editionen und einem herkömmlichen On-Premise-Deployment sind die Verantwortlichkeiten im SAP Basis Betrieb. Während in einer On-Premise-Umgebung das Unternehmen selbst vollständig für den Betrieb verantwortlich ist, übernimmt in der Essentials Edition und der Extended Edition die SAP einzelne Bereiche des Basisbetriebs.

Die Essentials Edition und die Extended Edition werden vollständig in der SAP eigenen Cloud gehostet. Wenn sich ein Unternehmen für eine dieser Editionen entscheidet, sollte daher berücksichtigt werden, dass in dieser Variante kein Zugriff auf Infrastrukturkomponenten, sowie vereinzelte Bereiche des SAP Basisbetriebs möglich ist, da der Betrieb von der SAP übernommen wird. Aus der operativen Sicht eines Anwenders sind die Geschäftsprozesse in der S/4HANA SaaS-Variante grundsätzlich vergleichbar mit einem On-Premise-Deployment. Aufgrund des unterschiedlichen Entwicklungs- und Innovationsfortschritts kann es vereinzelt zu Abweichungen, beispielsweise in der Verfügbarkeit von Fiori-Apps, kommen, da bei den S/4HANA Cloud-Deployments in kürzeren Abständen neue Releases eingespielt werden. Das hängt mit der geringeren Komplexität aufgrund der Standardisierung und des zentralisierten Betriebs durch die SAP im Vergleich zu individualisierten On-Premise-Deployments zusammen.

Die Essentials Edition und die Extended Edition der Cloud-Deployment-Variante unterscheiden sich primär in der Nutzung der Ressourcen. Die Essentials Edition wurde ursprünglich als „Multi Tenant“ bezeichnet, da die Bereitstellung der S/4HANA Instanz in diesem Modell in einer geteilten Infrastruktur in der SAP Cloud erfolgt. Bei der Extended Edition wird die S/4HANA Instanz hingegen in einer dedizierten, privaten Infrastrukturumgebung für den Kunden betrieben.

### 3.4.2 Unterschiede zwischen den Editionen

Die nachfolgenden Tabellen zeigen einen Vergleich der Unterschiede zwischen einem On-Premise-Deployment, der S/4HANA Cloud Essentials Edition und der S/4HANA Cloud Extended Edition.

#### Ebene Infrastruktur

Die Ebene der Infrastruktur umfasst die Bereiche, die für die Bereitstellung der jeweiligen Deployment-Variante erforderlich sind (Tab. 3.5).

#### Ebene Implementierung

Die Ebene der Implementierung befasst sich mit dem Ansatz und der Vorgehensweise bis zur Bereitstellung einer S/4HANA Instanz in der jeweiligen Umgebung (Tab. 3.6).

**Tab. 3.5** Übersicht der Verantwortlichkeiten auf der Ebene Infrastruktur

Bereich	S/4HANA Cloud Essential Edition	S/4HANA Cloud Extended Edition	SAP S/4HANA On-Premise
Lizenzen	Subscription basiert		Bring-your-own-license (BYOL)
Updates	Automatisiert durch die SAP	Automatisiert oder manuell	Manuell
Release Zyklus	Quartalsweise	Halbjährlich	Jährlich
Anbieter des Daten-zentrums	Vordefiniert (wird durch SAP festgelegt)	Auswahl durch den SAP-Kunden	Auswahl durch den SAP-Kunden
Leistung	Beeinflusst durch die gesamte System-nutzung	Private Umgebung ermöglicht Per-formance-Optimierung durch die SAP	Private Umgebung ermöglicht Per-formance-Optimierung durch den Kunden

Ebene Infrastruktur

**Tab. 3.6** Übersicht der Szenarien bei der Implementierung

Bereich	S/4HANA Cloud Essential Edition	S/4HANA Cloud Extended Edition	SAP S/4HANA On-Premise
Zeitdauer	8–16 Wochen	Ca. 16 Wochen	Kundenspezifisch
Implementierungs-strategie	Greenfield		Greenfield oder Brownfield

Ebene Implementierung

### Ebene Prozess

In der Prozessebene werden die Funktionalitäten und Möglichkeiten in Bezug auf die operative Nutzung eines S/4HANA Systems aus der Perspektive von Endbenutzern betrachtet (Tab. 3.7).

### Konfigurationen und Erweiterungen

Die Ebene der Konfigurationen beziehen sich insbesondere auf das Customizing eines SAP-Systems, mit der auch die möglichen Erweiterungen einhergehen (Tab. 3.8).

### Sicherheit

Die Ebene der Sicherheit beschreibt den Umfang der verfügbaren Sicherheitskonzepte und wie diese in der jeweiligen S/4HANA-Instanz anwendbar sind (Tab. 3.9).

### Gesamter Vergleich

(Siehe Tab. 3.10)

**Tab. 3.7** Übersicht der Abdeckung bei den Prozessen

Bereich	S/4HANA Cloud Essential Edition	S/4HANA Cloud Extended Edition	SAP S/4HANA On- Premise
Vorkonfigurierte Geschäftsprozesse	Obligatorisch		Optional
Abdeckung von Industrie-Lösungen	Finance ERP, Professional Services, Manufacturing, Oil & Gas		Vollständige Verfügbarkeit aller 25 Industrie- Lösungen
Lokalisierung	43 Länder 25 Sprachen	64 Länder 38 Sprachen	
Benutzeroberfläche	100 % web-basiert Überwiegend Fiori	Web- und GUI-basiert Teilweise Fiori	

Ebene Prozess

**Tab. 3.8** Übersicht zur Erweiterbarkeit

Bereich	S/4HANA Cloud Essential Edition	S/4HANA Cloud Extended Edition	SAP S/4HANA On- Premise
Customizing (IMG, SPRO)	Vereinfacht (Konfiguration mit Anleitung)	Vollständig verfügbar	
In-App Erweiterbar- keit	Ja		
Klassische Erweiterung	Nein	Ja	
Veränderungen	Nein		Ja

Ebene Konfiguration

**Tab. 3.9** Übersicht zur Sicherheit

Bereich	S/4HANA Cloud Essential Edition	S/4HANA Cloud Extended Edition	SAP S/4HANA On- Premise
Infrastruktur-Sicher- heit	Durch die SAP zur Verfügung gestellte hohe Sicherheit		Kundenspezifisch, SAP Kunde ist ver- antwortlich für die Sicherheit der Infra- struktur
Risiko Datendiebstahl	Höheres Risiko durch geteilte Ressourcen		Geringes Risiko durch isolierte Ressourcen
Desaster Recovery	Geteilte Ressourcen	Isolierte Ressourcen	

Ebene Sicherheit

**Tab. 3.10** Gesamtvergleich zu den Angeboten

Bereich	S/4HANA Cloud Essential Edition	S/4HANA Cloud Extended Edition	SAP S/4HANA On-Premise
Total Cost of Ownership	Gering	Mittel	Hoch
Aufwand für organisatorisches Change Management	Sehr hoch	Hoch	Kontrolle durch den SAP Kunden
Innovation	SAP-Verantwortlichkeit, Ergänzungen durch SAP Kunden	SAP-Verantwortlichkeit, Unterstützung durch den SAP Kunden	Überwiegend kunden-spezifisch

Vergleich im Überblick

### 3.4.3 Vorteile und Nachteile

Ein zentraler Aspekt des Cloud-Deployment-Ansatzes sind die verteilten Verantwortlichkeiten. Im Gegensatz zu einer S/4HANA On-Premise-Instanz übernimmt die SAP den Betrieb in einigen Bereichen. Durch die Verlagerung eines Teils des Betriebs benötigt ein SAP-Kunde weniger personelle und finanzielle Ressourcen, um eine SAP-Instanz zur Abwicklung der Geschäftsprozesse einzusetzen. Aufgrund des geringeren Bedarfs an fachlichem Wissen zum Betrieb eines SAP-Systems sind die Kosten im Vergleich zu einem On-Premise-Deployment signifikant geringer.

Die Kosteneinsparung durch den Einsatz eines Cloud-Deployments bringt jedoch den Nachteil mit sich, dass ein SAP-Kunde einen geringeren Einfluss auf die Individualisierbarkeit der S/4HANA-Instanz hat. Wenn ein Unternehmen beispielsweise über komplexe Geschäftsprozesse verfügt und der Anpassungsbedarf des SAP-Systems hoch ist, eignet sich die Nutzung des SaaS für S/4HANA bedingt, um die Anforderungen abzudecken.

Bei einem Vergleich der Vor- und Nachteile der beiden Deployment-Ansätze ist zudem die Zugänglichkeit und Verfügbarkeit zu neuen Innovationen und Funktionalitäten zu betrachten. Die Cloud-Deployment-Editionen basieren auf einem standardisierten Aufbau und werden zentralisiert durch SAP betrieben. Aus diesem Grund ist der Release-Zyklus kürzer als bei einem On-Premise-Deployment, da neue Funktionalitäten einen geringeren Anpassungsbedarf in den standardisiert aufgebauten S/4HANA-Instanzen haben. Der Zugang zu Innovationen ist folglich schneller möglich als bei einer S/4HANA On-Premise-Variante. Bei der Bereitstellung einer neuen Funktionalität übernimmt zudem die SAP einige Basis-Konfigurationen zur Integration, sodass SAP-Kunden lediglich meist einen geringen Aufwand zur Nutzung von neuen Innovationen erbringen müssen. Aufgrund der Standardisierung in einem S/4HANA Cloud-Deployment ist jedoch zu berücksichtigen, dass die Nutzung und Integration einer Innovation in einer On-Premise-Umgebung gegebenenfalls abweichen kann.

Insgesamt ist die Auslegung der Vor- und Nachteile der jeweiligen S/4HANA Deployment-Varianten geprägt von der individuellen Strategie eines Unternehmens. Die Entscheidung für ein Modell ist abhängig von den Anforderungen, die insbesondere im Betrieb an das SAP S/4HANA System gestellt wird.

### **3.4.4 Anwendungsfälle**

Ein beispielhafter Use Case für die Nutzung eines Cloud-Deployments ist der geringe Grad an Individualisierung einer S/4HANA-Instanz. Diese Anforderung basiert auf der Annahme, dass die Geschäftsfähigkeit eines Unternehmens auf standardisierten Prozessen basiert, die in einem hohen Maß an die Geschäftsprozesse von SAP angelehnt sind. Die Standardisierung des operativen Geschäfts verringert den Bedarf an individuellen Anpassungen eines Systems, weshalb ein S/4HANA Cloud-Deployment diese Anforderung abdecken kann. Eine Voraussetzung hierfür ist die Fähigkeit der Auslagerung vereinzelter Bereiche des Basis-Betriebs. Insbesondere die unternehmensinterne Compliance muss konform zur Nutzung des SaaS-Modells der SAP sein. Unter Erfüllung dieser Annahmen und Bedingungen kann eine S/4HANA Cloud-Edition auch als produktives System genutzt werden.

Ein Cloud-Deployment einer S/4HANA-Instanz kann jedoch auch als Test-Umgebung genutzt werden, um aufgrund des kürzeren Release-Zyklus schnelleren Zugang zu neuen Innovationen zu erhalten. Durch die schnellere Verfügbarkeit neuer Funktionen in der S/4HANA Cloud-Edition kann ein Unternehmen einen Eindruck bekommen, welche Neuheiten sich beispielsweise für die produktive SAP-Umgebung eignen und welchen Mehrwert diese mit sich bringen. Durch das S/4HANA Cloud-Deployment-Modell besteht die Möglichkeit, neue Innovationen in einer „realen“ Umgebung zu testen und zu evaluieren, bevor diese für die On-Premise-Umgebung zur Verfügung gestellt werden.

---

## **3.5 Zusammenfassung**

Vor dem Einsatz einer Public Cloud für die SAP S/4HANA-Systeme, sollten Unternehmen zunächst die wichtigen Rahmenbedingungen prüfen. Die Vertragsgestaltung hat hierbei einen besonderen Stellenwert. Da diese auf Standardverträgen der Hyperscaler beruht, gibt es bezüglich der Verträge wenig Handlungsspielraum für die Kunden. Wichtiger sind jedoch die Aspekte des Bezugs der Leistungen und der sehr wichtige Support durch die Hyperscaler. Es wird auch das Funding der Hyperscaler in Kürze beschrieben.

Ein besonderes Augenmerk liegt bei vielen Kunden auf den regulatorischen Anforderungen. So müssen Unternehmen aus dem Gesundheitswesen entsprechende Vorgaben umsetzen können – auch auf der Public Cloud. Das Teilkapitel beschreibt, wie dies zu erfüllen ist und wie hierbei vorgegangen werden muss.

Die effiziente Steuerung der SAP S/4HANA-Systemlandschaft ist wichtig für jedes Unternehmen. Das Kapitel zeigt, wie der SAP Solution Manager für die Steuerung genutzt werden kann und zeigt aber auch das neue cloudbasierte Tool der SAP, Cloud ALM. Eine Gegenüberstellung beider Produkte zeigt die Vor- und Nachteile auf und erleichtert die Wahl des Tools.

Das Deployment von SAP S/4HANA-Systemen kann nach verschiedenen Ansätzen her durchgeführt werden. Hierzu zählen die Neuprovisionierung der SAP-Systeme mit dem Greenfield-Ansatz oder aber die Transformation von SAP-Systemen in der Cloud hinein mit dem Brownfield-Ansatz.

Die SAP bietet ihren Kunden die wichtigen SAP S/4HANA-Systeme auch aus der Cloud heraus an und als Cloud-Editionen. Das Kapitel hierbei auch gezeigt, welche Möglichkeiten zu einem Bezug aus einer von SAP betriebenen Cloud bestehen.

Als letzten Punkt in diesem Kapitel wurde die Auswahl eines Hyperscalers genauer erläutert. Hierzu gibt es quantitative Faktoren, wie die Kosten, aber auch qualitative Faktoren, wie die Marktpositionierung, welche zu der Auswahl eines Hyperscalers beitragen und führen.



# SAP S/4 on Amazon AWS – Konzepte und Architekturen

4

## Zusammenfassung

In diesem Kapitel stellen wir Ihnen die Amazon Web Services Hyperscaler Cloud vor. Nach einem kurzen Überblick über die Geschichte von AWS, werden die wichtigsten Cloud Services vorgestellt, welche für das Hosting, die Verwaltung und den Betrieb von SAP S/4HANA in AWS erforderlich sind. Das Kapitel zeigt auch die verschiedenen möglichen Regionen und Verfügbarkeitszonen in AWS vor. Sie erhalten darüber hinaus einen Einblick in die AWS Managementconsole, über welche Sie alle nötigen Aktivitäten für die Elemente der AWS Cloud ausführen können. Darüber hinaus werden mögliche Implementierungsszenarien für SAP S/4HANA-Systemen in AWS erläutert und die jeweiligen Integrationsaspekte abgedeckt.

## 4.1 Die Geschichte von Amazon AWS

Amazon Web Services (AWS) bietet privaten und staatlichen Unternehmen und Privatkunden als Tochtergesellschaft von Amazon.com, Inc. On-Demand-Cloud-Computing-Plattformen an. AWS ist Marktführer mit einem Marktanteil von etwa 30 % am weltweiten Cloud-Geschäft. 13 % des Gesamtumsatzes von Amazon (113,08 Mia. USD) stammen aus dem Clouddienstgeschäft. 54 % des Betriebsergebnisses von Amazon stammten von AWS.

Das uns heute als AWS bekannte Angebot wurde im Jahr 2006 als Serviceangebot und Infrastruktur von Amazon.com eingeführt. Der Hauptvorteil, den AWS im Bereich Cloud-Computing bot, war der Wegfall von Anschaffungskosten (CapEx) für den Erwerb von Infrastrukturen. Diese wurden ersetzt durch sehr geringe variable Kosten, die auf dem spezifischen Unternehmensbedarf basierten. Dadurch mussten die Unternehmen

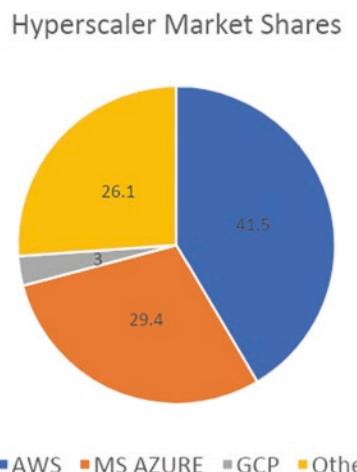
beim Erwerb von Servern oder sonstigen Infrastrukturen nicht mehr mehrere Monate im Voraus planen. Die vorhandenen Server konnten gemäß den jeweiligen Anforderungen leicht aufgerüstet, abgerüstet und nach oben oder unten skaliert werden. Dies führte zu einer völlig neuen Art des Hostings und der Verwaltung von Infrastrukturen und zu einer dauerhaften Veränderung der Welt der Technologie.

AWS betreibt SAP-Workloads bereits seit 2008, also deutlich länger als die anderen Cloud-Anbieter. AWS und SAP arbeiten bei Innovationen seit 2011 zusammen, um den Betrieb der SAP-Umgebungen der Kunden in der Cloud zu unterstützen. AWS bietet die breiteste Auswahl an SAP-zertifizierten, cloudnativen Instanztypen. Für die SAP-Kunden bedeutet dies, dass sie die Auswahl ihres Prozessors, Speichers, Netzwerks, Betriebssystems und Einkaufsmodells an ihre einzigartigen und sich verändernden Anforderungen anpassen können.

Laut einer Studie des IDC aus dem Jahr 2020 erzielen über 85 % der Kunden durch SAP in AWS Kosteneinsparungen. Mit dem Migration Acceleration Program (MAP) von AWS erhalten Sie Beratung, Training und Services für die Migration von geeigneten SAP-Workloads. Derzeit verwenden mehr als 5.000 Kunden SAP in AWS. AWS hält auf dem Hyperscaler-Markt bis heute den größten Marktanteil (Abb. 4.1).

#### 4.1.1 Neue Infrastruktur und Innovationen

AWS gründete eine Infrastrukturplattform, auf der die Entwickler neue Anwendungen entwickeln und testen können. AWS hat mehr als eine Million Benutzer, darunter Netflix und Airbnb. Aufgrund des riesigen Kundenstamms verfügt AWS über umfassende Erfahrung und über ein besseres Verständnis für die Cloud-Service-Verwendung durch die Kunden. AWS entwickelt auch weiterhin neue Infrastrukturen und Innovationen, um die Anforderungen der AWS-Kunden auf der ganzen Welt zu erfüllen.



**Abb. 4.1** Hyperscaler Marktanteile

AWS entwickelt Plattformen für Datenbanken, Entwickleranwendungen, Hilfsprogramme, Analysen und vieles mehr. Mit der von AWS entwickelten Serviceplattform, die über umfassenden Funktionalitäten verfügt, nimmt AWS als Cloud-Service-Anbieter die Spitzenposition ein. AWS bietet den Kunden schnelle und leicht skalierbare Lösungen. AWS hat die folgenden Stärken:

- Riesiger und breit gefächerter Kundenstamm
- Breites Spektrum an Übernahmen von strategischen Werkzeugen wie beispielsweise native Cloud-Anwendungen und E-Business-Hosting
- Substanzielle Technikpartnerschaften, darunter Softwareanbieter, die ihre Lösungen in AWS integrieren
- Breit angelegtes Partnernetzwerk, das Expertise in der Applikationsentwicklung mit verwalteten Services wie z. B. Rechenzentrumsmigration bietet
- Große Auswahl an IaaS- und PaaS-Funktionalitäten
- Rasche Serviceangebote und Skalierbarkeit sowie Erweiterungsfähigkeit der Lösungen

Das enorme Kapazitätsangebot von AWS ist also nützlich, um Lösungen für die geschäftlichen Herausforderungen zu finden. So können die Unternehmen beispielsweise die automatische Skalierung von AWS zur Lastverteilung nutzen, anstatt die Spitzenlasten weiterhin in traditionellen Rechenzentren zu verarbeiten. Die Unternehmen können mit der automatischen Skalierung von AWS je nach Last die Anzahl der zusätzlichen Server anpassen oder Server entfernen. Die automatische Skalierung unterstützt die Unternehmen dabei zu erkennen, wenn ein Server nicht stabil läuft, und diesen dann herunterzufahren oder durch einen anderen Server in einer stabilen Umgebung zu ersetzen. Der Vorteil von AWS ist also, den Unternehmen aus einer Hand sehr breit gefächerte Lösungen für ihre verschiedenen Anforderungen zu bieten.

AWS beherrscht aber den Markt für Cloud.-Computing wegen des Serviceangebots an die Kunden. Wettbewerber wie Microsoft, Google oder sonstige Anbieter könnten unter Umständen ein höheres Wachstum als AWS erzielen. AWS konnte aber durch ein verbessertes Infrastrukturangebot, Preissenkungen, kontinuierliche Innovationen die Marktführerschaft im Bereich Cloud-Computing bewahren.

Es existieren viele Gründe für die Nutzung von AWS und AWS ist immer noch der Marktführer im Bereich vom Cloud Computing. Die Gründe hierfür sind:

- **AWS existiert am längsten:** AWS führte sein Infrastrukturangebot im Jahr 2006 auf dem Markt ein. Seitdem baute sich AWS nicht nur mit dem Angebot großer Cloud-Computing-Plattformen einen riesigen Kundenstamm auf, sondern überarbeitete auch das Produkt. Microsoft Azure kam 2010 auf den Markt, Google App Engine (ein Vorgänger der Google Cloud) wurde 2008 eingeführt, und die Cloud von IBM wurde erst im Jahr 2009 mit der Einführung der LotusLive Collaboration Suite zugänglich.
- **Ein Jahr für kostenlose Tests:** Das Gratisangebot von AWS ermöglicht es den Benutzern, vor Cloud-Plattform von AWS vor der vollständigen Migration kostenlos in

der Praxis zu testen. Die Benutzer können zwölf Monate lang gratis auf AWS-Dienste wie Amazon EC2 (Cloud-Computing-Kapazität), Amazon S. 3 (Cloud-Speicher) und Amazon RDS (relationaler Datenbankdienst) zugreifen. Vor der endgültigen Entscheidung verfügen sie also über ein ganzes Jahr Benutzung. Von der Gratisversion zum Abonnement zu wechseln, ist sehr einfach. Die Testversion kann im laufenden Geschäftsbetrieb auf AWS mit komplettem Funktionsumfang umgestellt werden.

- **Viel größer als die Wettbewerber:** In einem im Mai in Forbes veröffentlichten Artikel wurde ein Gartner-Bericht besprochen, in dem die Rechenkapazität von AWS auf das Zehnfache der Kapazitätssumme der nächstgrößten 14 Infrastrukturbieter geschätzt wurde. AWS betreibt zahlreiche Rechenzentrum auf der ganzen Welt. Nach derzeitigem Stand verfügt AWS auf der Welt über eine ganze Reihe von Availability Zones. Für das kommende Jahr plant AWS bereits elf weitere Availability Zones in fünf weiteren geografischen Regionen. AWS kann den grundlegenden Bedarf bei der Verbesserung der IT-Infrastruktur decken und internationalen Unternehmen eine skalierbare Cloud-Lösung bieten.

#### 4.1.2 Vorteile von Amazon Web Services

Amazon Web Services zeichnet sich in der Nutzung durch einige Vorteile aus, welche wir im Folgenden kurz erläutern werden.

- **Benutzerfreundlichkeit:** Mit AWS können Anwendungsanbieter sowie unabhängige Softwarehäuser und -hersteller ihre Anwendungen schnell und sicher hosten, ganz gleich, ob es sich dabei um eine vorhandene Anwendung oder um eine neue SaaS-basierte Anwendung handelt. Der Zugriff auf die AWS-Plattform zum Hosten der Anwendungen erfolgt entweder über die Managementkonsole von AWS oder über die gut dokumentierten Webservices-APIs.
- **Flexibilität:** In AWS haben Sie die Wahl des Betriebssystems, der Datenbank, der Web-Anwendungsplattform, der Programmiersprache und der sonstigen benötigten Services. Mit AWS erhalten Sie darüber hinaus eine virtuelle Umgebung, in der Sie die Software und die Services laden können, die von Ihrer Anwendung benötigt werden. Auf diese Weise wird die Migration der bestehenden Anwendungen einfacher und zugänglicher. Gleichzeitig gibt es bei der Implementierung neuer Lösungen zahlreiche Optionen.
- **Wirtschaftlichkeit:** Bezahlte wird nur für die genutzte Rechenleistung, den genutzten Speicher und die sonstigen genutzten Ressourcen. Es müssen weder langfristige Verträge geschlossen noch Vorleistungen erbracht werden.
- **Zuverlässigkeit:** Mit AWS nutzen Sie die Vorteile einer skalierbaren, zuverlässigen und abgesicherten globalen Computing-Infrastruktur, die das virtuelle Rückgrat des milliardenschweren Online-Einzelhandelsunternehmens Amazon.com bildet und seit mehr als einem Jahrzehnt stetig weiterentwickelt wird.

- **Skalierbarkeit und extreme Leistungsfähigkeit:** Mit AWS-Werkzeugen wie der automatischen Skalierung und dem elastischen Lastenausgleich lässt sich Ihre Anwendung automatisch wie gewünscht skalieren.
- **Sicherheit:** AWS sichert und stabilisiert die Infrastruktur mit physischen, betrieblichen und softwarebasierten Maßnahmen. Hierbei wird ein End-to-End-Ansatz verfolgt. Im Kapitel Sicherheit werden wir näher darauf eingehen.

In der folgenden Abbildung sehen Sie die Amazon Web Services auf einen Blick (Abb. 4.2).

### 4.1.3 Das vergangene AWS-Geschäftswachstum

Der Cloud-Computing-Riese AWS schloss das Jahr 2020 unter der Führung des mittlerweile auch Amazon-CEOs mit einem Betriebsergebnis von 13,5 Mrd. US-Dollar ab, was gut 63 % des Jahresgewinns des gesamten Unternehmens ausmachte. Der Umsatz von AWS belief sich auf 45,3 Mrd. US-Dollar. Im Vergleich zum Vorjahr entsprach dies einer Steigerung um knapp 30 %.

AWS ist laut Amazons Finanzchef Brian Olsavsky das wohl weltweit profitabelste Großunternehmen im technischen Technologiebereich. Olsavsky machte diese Aussage in einer Telefonkonferenz mit Analysten und Investoren nach der Vorlage des Quartalsberichts für das vierte Quartal. Dabei stützte sich Olsavsky auf Jassys Erfahrung und Referenzen bei der Leitung des größeren Unternehmens.

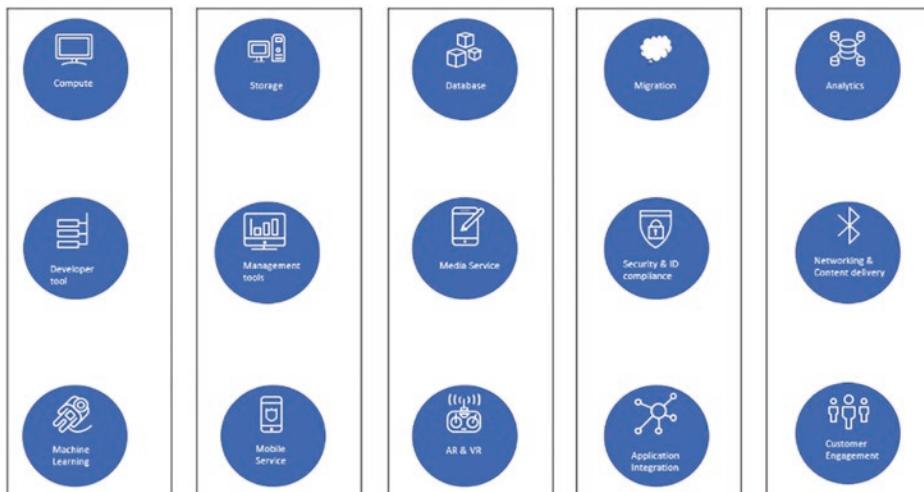


Abb. 4.2 Überblick zu den AWS Services



**Abb. 4.3** Wachstum von AWS in der Zeit von 2014–2020

Dies sind zahlreiche Vorteile, aber letztendlich maximiert Amazon kontinuierlich seine Marktposition als Marktführer im Bereich des Public-Cloud-Computing. Jassy hat seit der Einführung von AWS (Amazon Web Services) vor gut 15 Jahren ein Unternehmen im Unternehmen aufgebaut, was einer der Faktoren ist, weswegen er vom AWS-Vorstand als Nachfolger für den Amazon-Gründer und -CEO Jeff Bezos benannt wurde (Abb. 4.3).

## 4.2 Cloud-Services-Angebot

AWS ist der führende Hyperscaler-Cloud-Service-Anbieter. Das Angebot umfasst cloudbasierte On-Demand-Services unter IaaS und PaaS mit Pay-as-you-go-Preisen für Hosting und Betreiben von SAP-Anwendungen. AWS kann für das Hosting von SAP-Anwendungen seit dem Jahr 2008 genutzt werden. Seit 2011 sind AWS und SAP globale Technologiepartner. Im Laufe der Jahre sind die Angebote gereift, und mittlerweile hat AWS mehr als 5000 Kunden, welche SAP betreiben.

Derzeit werden den Kunden mehr als 200 Services angeboten. AWS bietet verfügt über eine breite Palette an SAP-zertifizierten Infrastrukturservices und -werkzeugen, die

dafür verwendet werden können, SAP-Produkte zu implementieren und zu betreiben. Die für Hosting und Betreiben von SAP-Anwendungen notwendigen Services und Werkzeuge werden im nächsten Abschnitt beschrieben:

### 4.2.1 Computing

Amazon Elastic Compute Cloud (Amazon EC2) bietet sichere und skalierbare Computing-Kapazität, die praktisch jede beliebige Workload für SAP unterstützt. Das Angebot ist breit gefächert und reicht von 2 vCPU und 3,75 GB RAM (VM mit Unterstützung für 1.995 SAPS) bis 448 vCPUs und 24,576 GB RAM (Bare Metal mit Unterstützung von 444,330 SAPS).

- ▶ **Tipp** Die aktuelle Liste der zertifizierten Amazon EC2s, die SAP-Workloads ausführen können, finden Sie in der SAP-Notiz 1.656.099 – SAP-Anwendungen in AWS: Unterstützte DB/OS- und AWS-EC2-Produkte.

Für den Betrieb von HANA-Datenbanken stellt SAP auch eine Liste der zertifizierten und unterstützten SAP-HANA-Hardware bereit. Diese ist über den folgenden Link verfügbar (Zugriff am 20.12.2021):

<https://www.sap.com/dmc/exp/2014-09-02-hana-hardware/enEN/#/solutions?filters=v:deCertified>

EC2-Instanzen werden in der Regel mit einem Amazon Machine Image (AMI) gestartet. AMIs können mit vorkonfigurierten Templates verglichen werden, die die gewünschte Konfiguration enthalten, beispielsweise das Betriebssystem und dessen Einstellungen, die Anwendungen und die Softwarekonfigurationen, die beim Start der EC2-Instanzen erwünscht sind.

SAP und AWS setzen voraus, dass die Amazon Machine Images (AMI) beim Start der EC2-Instanzen für die SAP-Workloads mit der HVM-Virtualisierung und nicht mit PV (Paravirtualisierte AMIs) verwendet werden. Dies vereinfacht die Migration zu den künftigen Instanztypen, sobald diese unterstützt werden. Weitere Erläuterungen zu HVM- und PV-AMIs finden Sie in der Amazon-Dokumentation (Zugriff am 20.12.2021): [https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/virtualization\\_types.html](https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/virtualization_types.html)

### 4.2.2 Speicher

AWS bietet für die Unterstützung von SAP-Workloads sowohl Blockspeicher als auch objektbasierten Speicher. Die wichtigsten in SAP verwendeten AWS-Speicherservices sind:

- Amazon Elastic Block Store (Amazon EBS)
- Amazon Elastic File System (Amazon EFS)
- Amazon FSx für Windows File Server
- Amazon Simple Storage Service (Amazon S. 3)

Die jeweiligen Services werden nun in den folgenden Teilkapiteln erläutert.

#### **4.2.2.1 Amazon Elastic Block Store (Amazon EBS)**

EBS sind persistente Blockspeicherlaufwerke, die EC2-Instanzen zugeordnet sind. Häufig funktionieren sie ähnlich wie ein Storage Area Network (SAN). EBS-Laufwerke sind mit unformatierten Raw-Blockgeräten vergleichbar, die auf ECS-Instanzen installiert werden können. Dasselbe Dateisystem kann dann zusätzlich zu diesen Laufwerken erstellt werden, um SAP- oder Datenbankdateien zu hosten.

Alle Dateisysteme, die keine gemeinsame Nutzung der SAP-Anwendungen und HANA-Datenbanken benötigen, sollten in EBS gehostet werden.

#### **4.2.2.2 Amazon Elastic File System (Amazon EFS)**

EFS ist ein gemeinsam genutzter Dateisystemdienst von Amazon. EFS wird verwendet, wenn in der Anwendung die Anforderung nach einem gemeinsamen Dateisystem besteht, das von allen Anwendungen genutzt wird.

- ▶ EFS wird auf Servern mit dem Windows-Betriebssystem nicht unterstützt.

In SAP kann EFS für gemeinsam genutzte Dateisysteme wie die sapmnt- und trans-Verzeichnisse verwendet werden, auf der HANA-DB-Seite für die gemeinsame Nutzung von HANA.

#### **4.2.2.3 Amazon FSx für Windows File Server**

FSx ist ein von AWS verwalteter Dateispeicher auf Basis von Microsoft Windows Server. Auch FSx wird für gemeinsame Dateisysteme verwendet und von allen auf dem Windows-Server gehosteten Anwendungen gemeinsam genutzt. Wenn Sie planen, EFS und FSx zu verwenden, werden die Gemeinkosten für die Bereitstellung, Verwaltung und Administration des Dateiservers von AWS übernommen.

#### **4.2.2.4 Amazon Simple Storage Service (Amazon S3)**

S. 3 ist ein objektbasierter Speicherservice, der hochgradige Skalierbarkeit, Zuverlässigkeit, Geschwindigkeit sowie eine kostengünstige Datenspeicherinfrastruktur bietet. Die Objekte werden redundant auf verschiedenen Geräten in unterschiedlichen Gebäuden oder Einrichtungen in einer Amazon-S. 3-Region gespeichert. Für den Betrieb von SAP, kann S. 3 in der Regel als langfristiger Backup-Speicher verwendet werden.

### 4.2.3 Netzwerkpflege

Die Netzwerkpflegeservices von Amazon bietet eine zuverlässige, sichere und hochgradig verfügbare Umgebung für das Hosting von SAP-Workloads. Die wichtigsten im SAP-Kontext verwendeten Services sind:

- Amazon Virtual Private Cloud (Amazon VPC)
- AWS Site-to-Site VPN
- AWS Direct Connect
- Amazon Route 53
- Amazon Time Sync

Im Folgenden werden wir Ihnen die jeweiligen Services erläutern.

#### 4.2.3.1 Amazon Virtual Private Cloud (Amazon VPC)

VPC gehört zu den AWS Networking Foundations und dient dazu, innerhalb der AWS-Cloud einen isolierten virtuellen Netzwerkbereich zu definieren. Dies ist in etwa mit dem dedizierten und abgesicherten Bereich in ihrem On-Premises-Rechenzentrum vergleichbar, in dem Sie SAP-Anwendungen und -Datenbanken hosten und betreiben.

Eine VPC liegt vollständig in Ihrem Eigentum, und Sie können Subnetze (in Ihrem eigenen IP-Adressenbereich), Kommunikationspfade (Routingtabellen und Netzwerk-Gateways) für das Hosting und den Zugriff auf die Anwendungen festlegen.

Die Ressourcen für Hosting und Betrieb von SAP wie EC2s, Gateways, Lastenverteiler werden in VPC bereitgestellt. Die Kommunikation wird mit Funktionalitäten wie Security Groups (SG) und Network Access Control Lists (NACLs) streng abgesichert.

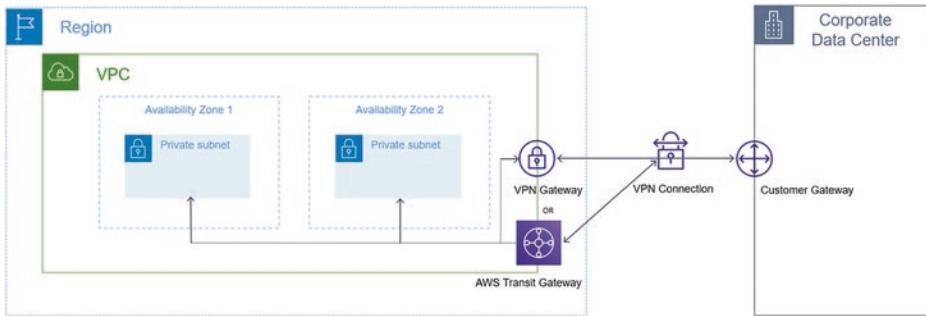
#### 4.2.3.2 AWS Site-to-Site VPN

Ihre VPC ist in AWS ein isolierter Netzwerkbereich. Für die Kommunikation zwischen Ihrem On-Premises-Netzwerk und Ihrer VPC können Sie Site-to-Site VPN verwenden.

Die Verbindung zwischen AWS VPC und dem On-Premises-Netzwerk kann mit Virtual Private Gateway oder einem Transit-Gateway als Gateway zur Amazon-Seite erleichtert werden. Auf der On-Premises-Site gibt es ein Kunden-Gateway, bei dem es sich um ein physisches Gerät oder um eine Software-Anwendung für die Site-to-Site-VPN-Verbindung handeln kann.

Die Daten fließen durch eine verschlüsselte Verbindung namens VPN-Tunnel. Jede VPN-Verbindung umfasst zwei VPN-Tunnel, die für hohe Verfügbarkeit simultan verwendet werden können (Abb. 4.4).

Mit demselben VPN-Gateway bzw. AWS-Transit-Gateway können mehrere On-Premises-Standorte verbunden werden. Mit einer solchen Konfiguration können sich auch die Benutzer im Büro mit in AWS gehosteten SAP-Anwendungen verbinden.



**Abb. 4.4** AWS Site-to-Site-VPN

#### 4.2.3.3 AWS Direct Connect

Mit Direct Connect können dedizierte private Netzwerke mit dem Rechenzentrum und mit AWS verbunden werden. Hierfür stehen zwei Verbindungsarten zur Verfügung: dediziert oder gehostet.

- Dedizierte Verbindung: Physische Ethernet-Verbindung mit genau einem Kunden.
- Gehostete Verbindung: Physische Ethernet-Verbindung, die von einem AWS Direct Connect Partner für einen Kunden bereitgestellt wird.

Für den Aufbau der Verbindung ist die Zusammenarbeit mit einem Partner im AWS Direct Connect Partner Program erforderlich. Dieser Partner bietet Unterstützung beim Aufbau der benötigten Netzwerke zwischen einem AWS-Direct-Connect-Standort und dem Rechenzentrum, der Geschäftsstelle oder der Co-Location-Umgebung.

#### 4.2.3.4 Amazon Route 53

Bei Amazon Route 53 handelt es sich um den DNS-Service (DNS=Domain Name System) von Amazon. Es ist für automatische Skalierbarkeit und hochgradige Verfügbarkeit konzipiert. Route 53 bietet drei Hauptfunktionen:

- Registrierung von Domänennamen
- DNS-Routing
- Health Check

Route 53 ermöglicht die Registrierung eines Namens für die Webseite oder Web-Anwendung. Dieser Name wird als Domänenname (domain name) bezeichnet.

In einer SAP-Installation ist es erforderlich, bei den URLs einen vollständigen Hostnamen zu verwenden (inklusive Angabe der Domäne). Die Domäne wird mit Route 53 registriert, das wiederum bei der Namen- und Adressauflösung für SAP-Systeme in AWS Unterstützung bietet.

#### 4.2.3.5 Amazon Time Sync

Amazon Time Sync ist ein hochpräziser und hochgradig zuverlässiger Zeitsynchronisationsdienst. Er wird über das Network Time Protocol (NTP) geliefert und ist in allen AWS-Regionen gratis für alle in einer VPC laufenden Instanzen verfügbar.

Um den Service zu erreichen, ist eine Verbindung zu der lokalen AWS-IP 169.254.169.123 notwendig. Der Datenverkehr verlässt das AWS-Netzwerk nicht und ist damit bei Zugriffen von Ihrem privaten Netzwerk aus vollständig abgesichert.

Wenn Sie mehrere SAP-Anwendungen in Ihrer Umgebung installieren, müssen die (Windows oder Linux betreibenden) EC2-Instanzen zeitlich synchronisiert sein. Dies ist insbesondere für verteilte SAP-Systeme empfehlenswert, in denen die SAP-Anwendungen und -Datenbanken auf unterschiedlichen EC2-Instanzen gehostet werden. Amazon Time Sync kann Unterstützung dabei bieten, die Betriebssystemzeit für alle EC2s synchron zu halten.

#### 4.2.4 Identity Management

Mit AWS Identity and Access Management (IAM) wird der abgesicherte Zugriff auf AWS Services und AWS Ressourcen abgesichert verwaltet. Es können Benutzer und Gruppen sowie die entsprechenden Berechtigungen für die Genehmigung bzw. Verweigerung des Zugriffs auf AWS-Services erstellt werden. Der IAM-Dienst ist ohne Nutzungskosten verfügbar.

Mit IAM profitieren Sie von einer sehr feinstufigen Zugriffskontrolle mit einem Sicherheitsmodell, das auf dem Prinzip der geringsten Privilegien beruht. IAM bietet eine Multi-Faktor-Authentifizierungsfunktionalität und can in den Verzeichnisdienst des Unternehmens integriert werden, beispielsweise in Microsoft Active Directory.

In einer in AWS laufenden SAP-Umgebung können die IAM-Rollen dazu genutzt werden, den Zugriff auf die AWS-Services zu genehmigen oder zu kontrollieren. Beispielsweise kann SAP-BASIS-Ressourcen die Genehmigung erteilt werden, EC2-Instanzen aufzurufen, anzuhalten und zu starten, diese aber nicht zu beenden.

#### 4.2.5 Sicherheit und Compliance

AWS bietet mehrere Sicherheitsressourcen, die zur Absicherung der Umgebung genutzt werden können, in der SAP in AWS läuft:

- **IAM** wird für die zentrale Verwaltung von Benutzern, Zugangsdaten wie Passwörtern, Zugriffsschlüsseln und Berechtigungsrichtlinien genutzt, als auch zur Kontrolle, auf welche AWS-Services und -Ressourcen die Benutzer zugreifen dürfen.
- **OS Hardening** zum Sperren der Betriebssystemkonfiguration, beispielsweise um zu vermeiden, dass ein NetWeaver-Administrator bei der Anmeldung an einer Instanz Zugangsdaten für einen Stammbenutzer erhält.

- Die nicht genutzten oder übermittelten Daten können per **Verschlüsselung** geschützt werden.
- **Security Groups** sind virtuelle Firewalls für EC2, mit denen der ein- und ausgehende Datenverkehr gesteuert wird.
- **Network Access Control** ist eine Liste, die eine weitere Sicherheitsschicht für die VPC bietet. Diese fungiert als Firewall für die Kontrolle des Datenverkehrs zu und von einem oder mehreren Subnetzen.
- **API Call Logging** – Alle Aufrufe von SAP-Anwendungen über APIs können mit AWS CloudTrail gespeichert werden. Die gespeicherten Informationen umfassen die Identität des API-Aufrufers, Datum und Uhrzeit des API-Aufrufs, die Quell-IP-Adresse des API-Aufrufers, die Abfrageparameter und die vom AWS-Service zurückgegebenen Elemente.
- **Notifications on Access** – Der Amazon Simple Notification Service (Amazon SNS) oder die Anwendungen von Drittanbietern können dazu verwendet werden, Benachrichtigungen über SSH-Anmeldungen an Ihre E-Mail-Adresse oder Ihre mobile Telefonnummer zu senden.

Darüber hinaus gibt es mehrere Standardrichtlinien und Best Practices, die bei AWS und SAP erhältlich sind, um in AWS laufende SAP-Umgebungen abzusichern und konform zu halten.

#### 4.2.6 Verwaltungswerkzeuge

Die AWS-Verwaltungswerkzeuge bieten Unterstützung bei der Verwaltung der Services und Ressourcen in der AWS-Cloud. Bei der Installation und Betrieb einer SAP-Umgebung in der AWS Hyperscaler Cloud sind mehrere AWS-Services beteiligt. Mit den AWS-Verwaltungswerkzeugen können Sie diese optimal steuern und verwalten und gleichzeitig die Compliance und die Sicherheit gewährleisten. Darüber hinaus behalten Sie die Kontrolle über die Betriebskosten. Der Zugriff auf die Verwaltungswerkzeuge von AWS ist auf zwei Arten möglich:

- AWS-Managementkonsole
- AWS Command Line Interface (AWS CLI)

##### AWS-Managementkonsole

Bei der AWS Managementkonsole handelt es sich um einen Web-Schnittstellendienst, mit dem AWS-Ressourcen bereitgestellt und administriert werden können. In der Anfangsphase des Betriebs von SAP in AWS ist es eine der ersten Anlaufstellen. Sie können damit AWS-Ressourcen für SAP bereitstellen und verwalten.

## AWS Command Line Interface (AWS CLI)

Bei CLI handelt es sich um ein Open-Source-Tool für die Interaktion mit AWS-Services. Die Befehle werden in einer Befehlszeile wie Bash, tcsh (Linux), PowerShell oder cmd (Windows) gegeben. Mit CLI können Skripte für die Automatisierung sowohl der Erstellung als auch der Ausführung von Aktivitäten von SAP-Anwendungen und -Datenbanken in AWS erstellt werden. Die wichtigsten in AWS verfügbaren Verwaltungswerkzeuge sind:

- AWS CloudFormation
- Amazon CloudWatch
- AWS CloudTrail
- AWS Config

Diese verfügbaren Verwaltungswerkzeuge können auf Basis der jeweils enthaltenen Funktionalitäten grob in drei Kategorien eingeteilt werden:

- AWS CloudFormation: Hierzu zählen alle Bereitstellungstools.
- AWS Cloudwatch: Hierzu zählen alle Überwachungstools
- AWS CloudTrail und AWS Config: Hierzu zählen alle Tools für Aufgaben der Governance

Im Folgenden werden die wichtigsten Werkzeuge aus den jeweiligen Kategorien beschrieben.

### 4.2.6.1 AWS CloudFormation

Dieser AWS-Service bietet die Möglichkeit, die Ressourcen von AWS und von Drittanbietern als Infrastructure as Code bereitzustellen und zu verwalten.

Mit dem CloudFormation-Template können Sie den gewünschten Zustand der Ressourcen sowie die entsprechenden Abhängigkeiten als Code angeben. Die Ressourcen können dann einmalig oder mehrmals zu einem Stapel gebündelt und aufgerufen und konfiguriert werden, wobei stets derselbe Zustand gewährleistet ist.

Bei der SAP-Bereitstellung ist dies sehr nützlich, wenn mehrere SAP-Anwendungen sowie Disaster-Recovery-Systeme mit denselben Einstellungen installiert werden müssen.

### 4.2.6.2 Amazon CloudWatch

Amazon Cloud Watch ist der zentrale Überwachungsservice für AWS CloudWatch sammelt Überwachungs- und Betriebsdaten zu den Ressourcen und Services, die in AWS und On Premises laufen. Mit den gesammelten Protokollen und Metriken können Sie detaillierte Informationen zum Status der Problembehebungen und sonstige Daten einsehen. Darüber hinaus können zwecks Automatisierung der Aktionen Alarne/Ereignisse

eingestellt werden. Sie können Ihre Anwendungen mithilfe von Amazon CloudWatch also reibungslos laufen lassen.

Wenn eine SAP-Umgebung in der AWS-Cloud gehostet und betrieben wird, umfasst dies unterschiedliche AWS-Services wie EC2, EBS, Netzwerkdienste etc. AWS sammelt die Metriken für jeden aktiven Service in Ihrem Konto, und auf der CloudWatch-Homepage werden alle diese Metriken an einem zentralen Ort angezeigt. Dadurch verfügen Sie mit Amazon CloudWatch über eine konsolidierte Sicht auf Zustand und Status der Ressourcen.

#### **4.2.6.3 AWS CloudTrail**

Wenn eine Anwendung in AWS Hyperscaler Cloud gehostet wird, führen die Benutzer zwangsläufig Aktionen aus und benutzen dafür die AWS-Managementkonsole, die AWS Command Line Interface oder Rollen bzw. sonstige AWS-Services via SDKs oder APIs von AWS. Um den reibungslosen Betrieb und vor allem die Umgebungssicherheit zu gewährleisten, ist es äußerst wichtig, alle im AWS-Konto ausgeführten Aktionen zu protokollieren. Der AWS-Service CloudTrail protokolliert alle Benutzeraktivitäten und die API-Verwendung in der AWS-Umgebung.

Standardmäßig wird CloudTrail bei der Erstellung des AWS-Kontos aktiviert. Alle Aktivitäten im AWS-Konto werden in CloudTrail als Ereignisse aufgezeichnet. Um die Compliance zu gewährleisten und die betrieblichen Risiken zu verringern, können diese Ereignisse analysiert werden. Wenn SAP-Systeme in AWS Hyperscaler Cloud ausgeführt werden, ist CloudTrail ein sehr wichtiges Governance-, Compliance- sowie Betriebs- und Risikoprüfungswerkzeug.

#### **4.2.6.4 AWS Config**

AWS Config ist eine Art Datenspeicher für die gesamte Konfiguration der AWS-Ressourcen im AWS-Konto. AWS Config wird auf regionaler Ebene aktiviert, und bei der Aktivierung werden sämtliche im AWS-Konto aktiven Ressourcen ermittelt und für jede ein Snapshot-Datensatz in Form von Konfigurationselementen erstellt. Dieser Datensatz umfasst die aktuellen Konfigurationszustände der Ressourcen. Darüber hinaus werden auch die Ressourcenbeziehungen detailliert erfasst. Ferner werden alle an einer AWS-Ressource vorgenommenen Konfigurationsänderungen protokolliert und historisiert.

Die Ressourcenkonfigurationen können also automatisch mit den gewünschten Konfigurationen abgeglichen werden. Wenn SAP-Anwendungen in AWS ausgeführt werden, kann die entsprechende Umgebung groß sein und sich im Zeitverlauf immer stärker erweitern. Um den Überblick über die interdependenten AWS-Ressourcen sowie deren Konfigurationsstatus zu behalten, ist das Standardtool AWS Config sehr gut geeignet.

## 4.3 Regionen und Availability Zones

### 4.3.1 AWS-Region

Unter Regionen sind die geografischen Gebiete auf der Welt zu verstehen, in denen Amazon seine Cloud-Computing-Ressourcen hostet. Die AWS-Regionen sind unabhängig und voneinander getrennt. Der Vorteil sind eine größere Fehlertoleranz sowie eine stärkere Stabilität und Ausfallsicherheit der gehosteten SAP-Anwendungen. Zum Zeitpunkt der Entstehung des vorliegenden Buches gab es 25 AWS-Regionen. Einzelheiten zu den Regionen und deren AWS-Code finden Sie hier (Zugriff am 20.12.2021): <https://docs.aws.amazon.com/general/latest/gr/rande.html#region-names-codes>

### 4.3.2 AWS Availability Zones

Innerhalb jeder Region werden die Cloud-Computing-Ressourcen in voneinander getrennten Gebieten gehostet, die als Availability Zones bezeichnet werden. Bei diesen Availability Zones handelt es sich um Cluster von Rechenzentren innerhalb einer Region, deren Aufbau so konzipiert wurde, dass Einflüsse auf die anderen Availability Zones in derselben Region ohne Auswirkung auf die betrachtete Availability Zone bleiben.

Pro Region gibt es zwischen zwei und fünf Availability Zones (Abb. 4.5).

Die Availability Zones in derselben AWS-Region sind mit vollständig redundanten und dedizierten Metro-Glasfaserleitungen miteinander verbunden. Dadurch sind beim Kommunikationsdurchsatz zwischen den Services der Availability Zones eine hohe Bandbreite und eine niedrige Latenz gewährleistet. Jede Availability Zone wird von verschiedenen Segmenten des Versorgungsanbieters gespeist und verfügt über eine redundante Verbindung zu mehreren Internet Service Providern. Im Falle von extern verursachten Störungen ist dadurch sichergestellt, dass die Serviceverfügbarkeit und die Servicekonnektivität so wenig wie möglich beeinträchtigt werden. Der gesamte Datenverkehr im Netzwerk zwischen den Availability Zones wird verschlüsselt. Bei dedizierten Metro-Glasfaserleitungen ist der Durchsatz so hoch, dass synchrone Replikationen problemlos und ohne Latenz möglich sind. Die Availability Zones befinden sich in einer sinnvollen physischen Entfernung voneinander (viele Kilometer), wobei keine Availability Zone mehr als 100 km (60 Meilen) von den anderen entfernt ist. Zur Availability Zone gibt es eine weitere Funktion namens Placement Group, die Schutz vor Hardwareausfällen oder Ausfällen aufgrund von Naturkatastrophen wie Überschwemmungen, Bränden, Stürmen etc. bietet.

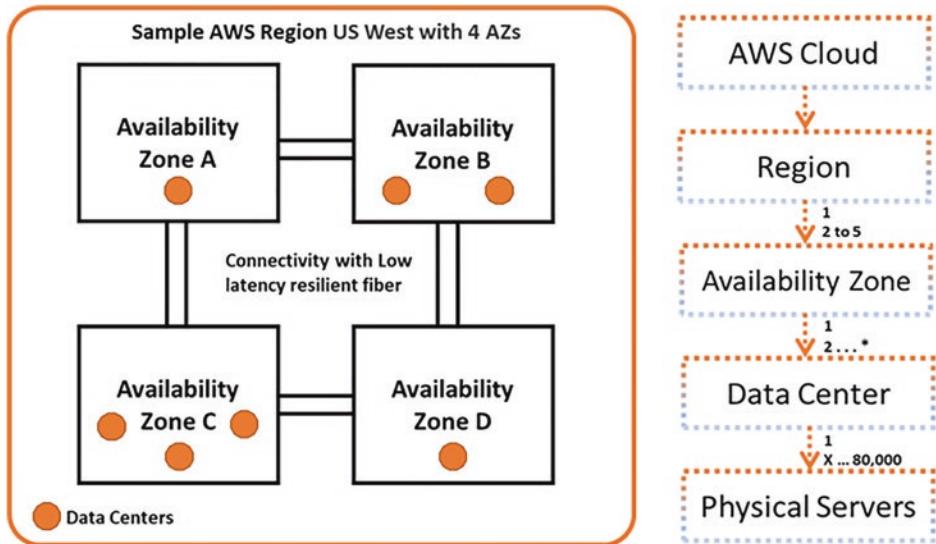


Abb. 4.5 AWS Region und Availability Zone

### 4.3.3 AWS Placement Groups

#### 4.3.3.1 Übersicht zu Placement Groups

Durch die Placement Groups ist festgelegt, wie das Hosting der SAP-Anwendung in den einzelnen Instanzen organisiert und die Instanzen in der zugrundeliegenden Hardware platziert werden. Hierbei geht es darum, die Auswirkungen von Hardwareausfällen auf ein Minimum zu begrenzen. Für die Platzierung der SAP-Anwendungsinstanzen stehen drei Placement-Group-Strategien zur Verfügung.

#### Cluster Placement Group

Bei dieser Strategie werden die Instanzen in ein und derselben Availability Zone platziert. Die räumliche Nähe der Anwendungsinstanzen ist dabei das Hauptmerkmal. Dies führt zu einer geringen Netzwerklatenz und zur bestmöglichen Kommunikation zwischen den Instanzen. Die Instanzen können in derselben VPC angesiedelt sein oder per VPC-Peering miteinander verbunden werden.

#### Partition Placement Group

Bei dieser Strategie werden die Instanzen in logischen Partitionen gebündelt. Die Bereitstellung der Instanzen in logischen Partitionen durch AWS erfolgt auf getrennter Hardware, was Schutz gegen korrelierte Hardwareausfälle bietet, weil sich die Instanzen in einer bestimmten Partition keine Hardware mit den Instanzen der anderen Partitionen teilen.

### Spread Placement Group

Bei dieser Strategie läuft jede Instanz auf einer separaten Hardware. Die Spread Placement Group kann sich auf eine oder mehrere Availability Zones erstrecken. Spread Placement Groups und Partition Placement Groups unterscheiden sich in der Bündelung der Instanzen. Während Partition Placement Groups pro Partition aus mehreren Instanzen bestehen, befindet sich in jeder Spread Placement Group lediglich eine Instanz. Die Spread Placement Group verteilt sich über unterschiedliche Hardware in den Availability Zones.

## 4.3.3.2 Planung der Implementierung eines SAP S/4HANA-Systems

### Planung der S/4HANA-Implementierung: Auswahl der Region

Bei der Planung von S/4 HANA-Systemimplementierungen ist die Auswahl der am besten geeigneten Region entscheidend. Diese Auswahl richtet sich nach dem erstellten AWS-Konto. Bei der Auswahl der AWS-Region sind die folgenden Faktoren zu berücksichtigen:

- **Regulatorische Compliance** – Überprüfung der gesetzlichen Anforderungen des jeweiligen Landes/Gebiets in Bezug auf Hosting und Speichern von Finanz- und Benutzerdaten. Auf die Auswahl der beiden Regionen für Ihre Disaster-Recovery-Einrichtung sollten Sie ein besonderes Augenmerk richten. Hierbei müssen die für die Übertragung und das Speichern von Finanz- und Benutzerdaten gültigen länderspezifischen Gesetze berücksichtigt werden.
- **Nähe zu den Benutzern** – Die AWS-Region sollte näher bei dem überwiegenden Anteil Ihrer Benutzer liegen. Dadurch lassen sich geringere Latenzen und ein höherer Durchsatz erzielen.
- **Betriebskosten** – Die Kosten sind bei der Wahl der Region ein wichtiger Faktor. Die Preise für EC2 und andere Dienste wie Egress Traffic fallen in den verschiedenen AWS-Regionen unterschiedlich aus. Es sollte also eine AWS-Region gewählt werden, in der die Hosting-Kosten für EC2-Instanzen niedrig sind und das bestmögliche Preis-Leistungs-Verhältnis herrscht.

Mit dem AWS-Preisrechner können die ungefähren Kosten ermittelt werden (Zugriff am 20.12.2021): <https://calculator.s3.amazonaws.com/index.html>

### Planung der S/4HANA-Implementierung: Auswahl der Availability Zone

Eine der Hauptanforderungen bei der Bereitstellungsplanung für SAP S/4HANA-Instanzen (SAP-Anwendungsinstanz und Datenbank) ist, dass das System kontinuierlich verfügbar sein muss. Es sollte keine ungeplanten Ausfallzeiten geben, insbesondere sollte die zugrundeliegende Hosting-Hardware keine Ausfälle verursachen.

Hochgradige Verfügbarkeit ist eine notwendige Voraussetzung für fast alle in Produktion befindlichen S/4HANA-Systeme von SAP. Bei geschäftskritischen Produktionsystemen sind ungeplante Ausfälle zu teuer. Daher ist es entscheidend, die S/4HANA-Instanzen von SAP so zu planen, dass sie hochgradig verfügbar sind. Hochgradig verfügbare Konstellationen erfordern zusätzliche Hardware mit komplexen Einstellungen. Die damit verbundenen Hardware- und Wartungskosten zählen daher zu den Einschränkungen.

Bei der Bereitstellung der Instanzen in Availability Zones ist je nach der gewählten Strategie eine hohe oder höhere Verfügbarkeit der Anwendungs- und Datenbankinstanzen gewährleistet. Gleichzeitig können die Kosten beobachtet werden.

### Auswahl der Availability Zone – Nicht produktives S/4HANA-SAP-System

Bei nicht produktiven SAP-S/4HANA-Systemen (Entwicklungs-, Qualitätssicherungs-, Test- oder Schulungssysteme) beschränkt sich die Anforderung in der Regel auf maximale Verfügbarkeit. Eine solche maximale Verfügbarkeit lässt sich mit dem folgenden Ansatz erreichen:

- SAP-System mit mehr als einer Anwendungsserverinstanz (D<nn>): Platzierung in Partition Placement Groups oder in Spread Placement Groups.
- Alle SAP-S/4HANA-Komponenten – ASCS, Anwendungsserver, Datenbank – laufen auf getrennten EC2-Instanzen.

Diese Einstellung ist für hohe RTO/RPO geeignet, d. h. mehrere Stunden oder mehrere Tage. Bei Hardwareausfällen ist die Ausfallzeit bis zur Wiederherstellung des Betriebs hoch.

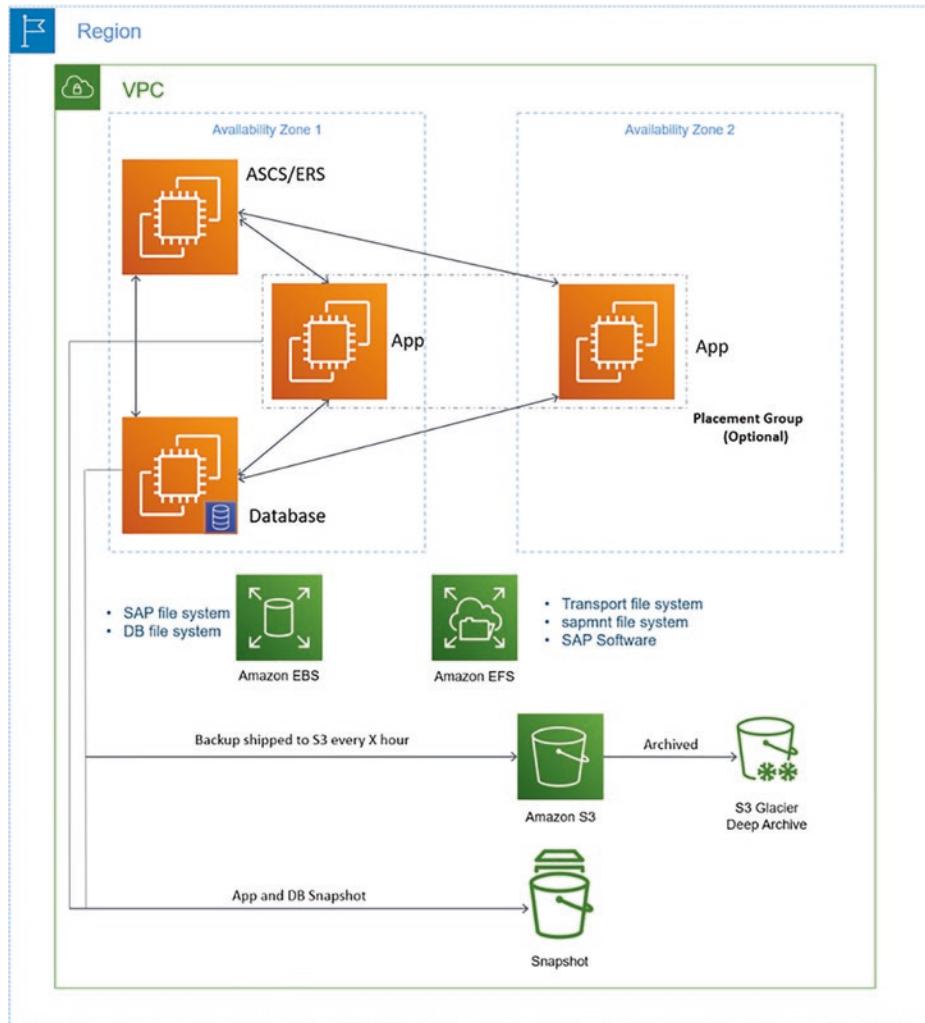
Durch die Konstellation mit einem Anwendungsserver in einer Placement Group ist sichergestellt, dass auch dann ein anderer Anwendungsserver verfügbar ist, wenn ein Server aufgrund eines Hardwareproblems ausfällt (Abb. 4.6).

### Auswahl der Availability Zone – Produktives SAP S/4HANA-SAP-System

In produktiven SAP-S/4HANA-Systemen ist die hohe Verfügbarkeit der SAP-Anwendung sowie der HANA-Datenbankinstanz entscheidend. Diese sind die wichtigsten Single-Point-Of-Failure-Komponenten (SPOF). Es empfiehlt sich der folgende Ansatz (Abb. 4.7):

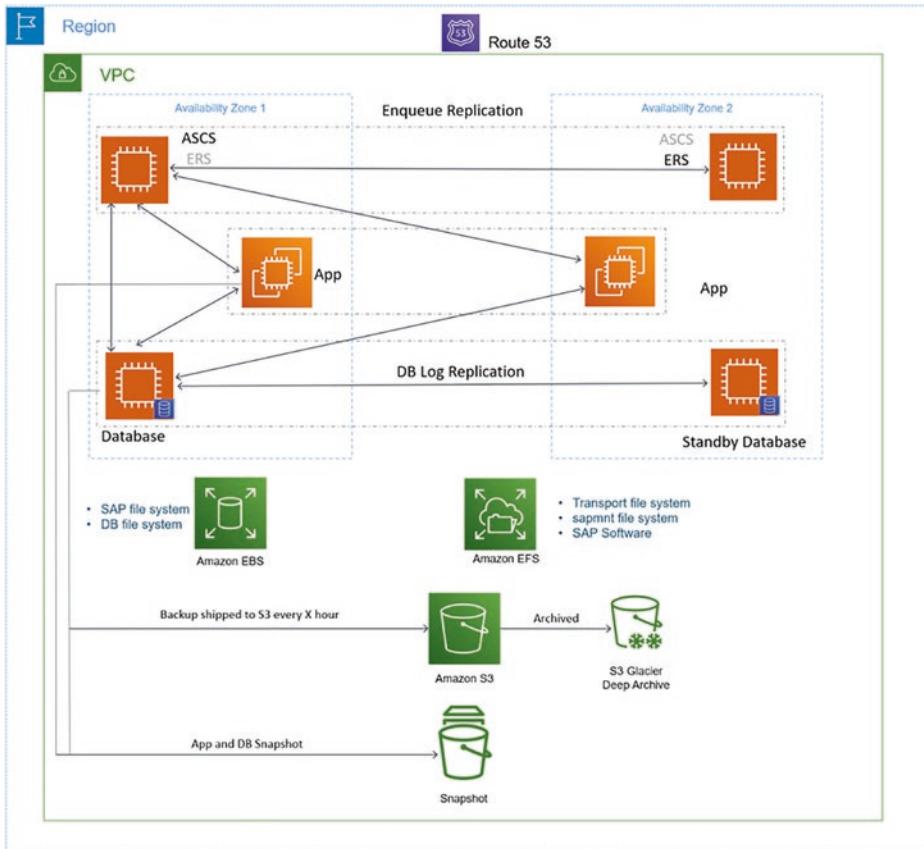
Die Konfiguration der Placement Group ist optional. Weitere Vorteile der Bereitstellung des SAP-S/4-Systems in einer Multi-Availability-Zone sind wie folgt:

- Die Backups können der Standby-HANA-Datenbankinstanz entnommen werden.
- Weniger Systemausfälle bei der Installation von Patches und Upgrades.



**Abb. 4.6** Selektion einer Availability Zone für Nichtproduktionssysteme

In AWS ist die Auswahl der Availability Zones bei der Bereitstellung von EC2-Instanzen besser steuerbar. Während der Bereitstellung einer EC2-Instanz besteht die Möglichkeit, eine Availability Zone in der Bereitstellungsregion der Instanz auszuwählen. In einem SAP S/4HANA System mit mehr als einem SAP-Anwendungsserver kann also gesteuert werden, in welcher Availability Zone die Instanz bereitgestellt wird. Dies macht die Placement Group überflüssig, da bereits Kontrolle über die Platzierung der Instanzen besteht.



**Abb. 4.7** Selektion einer Availability Zone für das Produktionssystem

## 4.4 AWS-Managementkonsole

Bei der AWS-Managementkonsole handelt es sich um ein webbasiertes Portal, mit dem die Benutzer sicher auf die Amazon Web Services zugreifen und diese verwalten können. Die Benutzer benötigen für die Anmeldung eine ID und ein Passwort. Um die Zugriffsabsicherung zu verstärken, kann in der Regel die Multi-Faktor-Authentifizierung aktiviert werden. Die Benutzer werden dann aufgefordert, den Authentifizierungscode des Geräts zu erfassen. Die AWS-Managementkonsole bietet als moderne Schnittstelle Zugriff von Mobilgeräten aus, darunter Smartphones und Tablets. Die Anwendungen sind für Android und iOS verfügbar.

#### 4.4.1 Unterstützte Browser

Die AWS-Managementkonsole unterstützt die drei neuesten Versionen von Google Chrome, Mozilla Firefox, Microsoft Edge, Apple Safari und Microsoft Internet Explorer 11. Die AWS-Managementkonsole setzt mindestens Internet Explorer, Version 11, voraus (Tab. 4.1).

#### 4.4.2 Benutzer für die Verwaltung

Es gibt zwei Arten von Benutzern, die sich an der AWS-Managementkonsole anmelden können. Es existieren die Stammbenutzer und die IAM-Benutzer (IAM=Identity and Access Management).

##### 4.4.2.1 AWS-Stammbenutzer

Die erste Anmeldung bei AWS wird als neuer Benutzer durchgeführt. Bei dieser ersten Anmeldung wird hauptsächlich das AWS-Konto erstellt. Die Benutzerkennung erhält vollständigen Zugriff auf alle AWS-Services und -Ressourcen im neu erstellten Konto. Diese Identität ist der Stammbenutzer des AWS-Kontos.

Dieser Stammbenutzer ist der mächtigste Benutzer des AWS-Kontos. Wir empfehlen daher ausdrücklich, ihn nicht für alltägliche Aufgaben zu verwenden. Nach der Einrichtung des Kontos gehen Sie wie folgt vor:

- Erstellen Sie Ihren ersten IAM-Benutzer mit Administratorberechtigungen.
  - Bewahren Sie die Zugangsdaten des Stammbenutzers an einem sicheren Ort.
  - Verwenden Sie das Stammbenutzerkonto nur für ganz bestimmte Aufgaben.
- **Tipp** Stammbenutzern kann per IAM-Richtlinien nicht der Zugriff verweigert werden.  
Zur Begrenzung der Berechtigungen des Stammbenutzers kann die Service Control Policy (SCP) von AWS Organizations verwendet werden.

**Tab. 4.1** Unterstützung für Browser

Browser	Version	Leistungen
Apple Safari	Die jüngsten 3 Versionen	Alle Services
Google Chrome	Die jüngsten 3 Versionen	Alle Services
Mozilla Firefox	Die jüngsten 3 Versionen	Alle Services
Microsoft Edge	Die jüngsten 3 Versionen	Alle Services
Microsoft Internet Explorer	11	Alle Services

Der Stammbenutzer ist für die folgenden Aufgaben nötig:

- **Änderung** an den Einstellungen des Stammbenutzerkontos
  - Sie müssen sich als Stammbenutzer anmelden, wenn Sie die Einstellungen des Stammbenutzerkontos ändern möchten. Hierzu zählen unter anderem der Kontoname, die E-Mail-Adresse, das Passwort des Stammbenutzers und die Stammbenutzerkennung.
- **Wiederherstellung** der Berechtigungen des IAM-Benutzers
  - Falls sich der IAM-Administrator versehentlich seine eigenen Berechtigungen entzieht, können Sie sich als Stammbenutzer anmelden, die Richtlinien bearbeiten und die entzogenen Berechtigungen wiederherstellen.
- **Aktivierung** des IAM-Zugriffs auf die Rechnungsstellungs- und die Kostenverwaltungskonsole
- **Schließen** des AWS-Kontos
- **Änderungen** am AWS-Supportplan bzw. Kündigung
- Konfigurierung eines Amazon-S. 3-Containers, um die MFA (Multi-Faktor-Authentifizierung) zu aktivieren
- **Bearbeitung** oder Löschen ungültiger VPC-IDs oder VPC-Endpunkt-IDs aus der S. 3-Container-Richtlinie
- **Anmeldung** an der GovCloud
- Anzeige bestimmter **Steuerrechnungen**
- **Registrierung** auf dem Reserved Instance Marketplace als Verkäufer

#### 4.4.2.2 IAM-Benutzer

Eine weitere (empfohlene) Möglichkeit, sich in der AWS-Managementkonsole anzumelden und interaktive Aufgaben auszuführen, ist die Anmeldung als IAM-Benutzer. Die erste IAM-Benutzerkennung mit Administratorberechtigung wird von dem Stammbenutzer nach der ersten Anmeldung erstellt. Alle anderen IAM-Benutzer werden mit den entsprechenden Berechtigungen danach von dem IAM-Administrator erstellt. Dieser fügt die neuen IAM-Benutzer entweder zu einer Benutzergruppe mit den entsprechenden Berechtigungsrichtlinien hinzu (empfohlen), oder er ordnet die Berechtigungsrichtlinien den neuen IAM-Benutzern direkt zu.

Die IAM-Benutzer führen die täglichen Administrationsaufgaben durch und kümmern sich um Verwaltung und Interaktion der einzelnen AWS-Services und -Ressourcen. Hinter den IAM-Benutzerkennungen können Personen oder Services stehen, die mit AWS interagieren können.

Jeder IAM-Benutzer wird in AWS durch seinen Benutzernamen und sein Passwort für die AWS-Managementkonsole identifiziert. Darauf hinaus gibt es bis zu zwei Zugangs-schlüssel, die mit der API oder der CLI verwendet werden können.

### 4.4.3 IAM-Rolle und IAM-Benutzergruppe

IAM-Rollen werden verwendet, um einem Benutzer Berechtigungen zuzuweisen, mit denen er eine Aufgabe in der AWS-Konsole ausführen kann.

Bei der IAM-Benutzergruppe handelt es sich um eine Gruppe für die Verwaltung der AWS-IAM-Benutzer. Zur Vereinfachung der Verwaltung können die IAM-Benutzer zu IAM-Gruppen gebündelt werden. Einer Gruppe von Benutzern mit identischem Rechten in der AWS-Managementkonsole Berechtigungsrichtlinien zuzuweisen und diese zu verwalten, ist auf einfache Weise möglich.

In Bezug auf die Berechtigungsrichtlinien, die den Berechtigungsinhabern die Ausführung bestimmter Aufgaben erlauben, sind sich IAM-Rollen und IAM-Benutzer ähnlich. Der einzige Unterschied besteht darin, dass IAM-Benutzer über ein individuelles Passwort oder einen individuellen Zugriffsschlüssel verfügen. Die Verwendung von IAM-Rollen empfiehlt sich, wenn eine Aktivität von einer Anwendung und nicht von einem Benutzer ausgeführt werden soll.

#### Beispiel

Eine AWS-Anwendung, die auf einer EC2 läuft, fordert einen bestimmten AWS-Service an. Anstatt einen IAM-Benutzer mit Passwort zu erstellen und diesen der Anwendung zuzuordnen oder die Anmelddaten in die Anwendung einzubetten, empfiehlt sich die Verwendung der Amazon EC2 zugewiesenen IAM-Rolle. Somit wird die auf EC2 laufende Anwendung zur Ausführung der Aufgabe berechtigt, die wiederum durch die der IAM-Rolle zugeordneten Richtlinien erlaubt ist ◀

### 4.4.4 Erste Schritte mit der AWS-Managementkonsole

Um mit der AWS Managementkonsole zu arbeiten, gehen Sie wie folgt vor:

- Öffnen Sie einen Internetbrowser.
  - Erfassen Sie die URL <https://console.aws.amazon.com/console/home> ein, um auf die AWS-Managementkonsole zuzugreifen (Zugriff am 20.12.2021).
  - Erfassen Sie die Zugangsdaten für das AWS- oder das IAM-Konto.
- **Tipp** Für AWS GovCloud (USA) lautet die URL der AWS-Managementkonsole anders.  
Öffnen Sie im Browser die Seite <https://console.amazonaws-us-gov.com> und melden Sie sich mit den Anmelddaten des IAM-Kontos an der Managementkonsole an (Zugriff am 20.12.2021).

#### 4.4.5 Auswahl einer anderen Region

Mit Ausnahme einiger weniger Services sind alle im AWS-Konto erstellten Ressourcen einer AWS-Region zugeordnet. Nach der Anmeldung an der AWS-Managementkonsole muss also die passende AWS-Region ausgewählt werden.

Manche AWS-Services werden in allen Regionen angeboten. Um zu diesen Services zu navigieren, muss keine andere Region gewählt werden.

- Identitäts- und Zugriffsverwaltung in AWS (IAM – Identity and Access Management)
- AWS-Managementkonsole
- Amazon CloudWatch

#### 4.4.6 Navigation zu den Services

Nach der Anmeldung in der AWS-Managementkonsole und nach der Auswahl der passenden Region gibt es verschiedene Möglichkeiten, einen AWS-Service zu suchen oder zu einem AWS-Service zu navigieren.

Um alle verfügbaren Services nach Kategorien gruppiert anzuzeigen, navigieren Sie wie unten beschrieben.

#### 4.4.7 Suche nach einem AWS-Service

Um nach einem AWS-Service zu suchen, gehen Sie wie im Folgenden beschrieben vor.

Erfassen Sie den Namen des gesuchten Service. Auf der Landingpage der AWS-Konsole werden die häufigsten und die zuletzt verwendeten Services angezeigt. Hier können Sie auch den gewünschten Servicenamen in das Suchfeld eintragen.

- **Tipp** Sie können ein Favoritenmenü erstellen und den gesuchten Service zu den Favoriten hinzufügen. Um einen Service zu den Favoriten hinzuzufügen, gehen Sie zu dem betreffenden Service, und klicken Sie auf den Stern links neben dem Servicenamen. Tippen Sie dann auf Start.

Über die Tastatur ist es möglich, schnell zu dem ersten Suchergebnis zu navigieren: Drücken Sie zunächst Alt+s (Windows) bzw. Option+s (macOS), um die Suchleiste aufzurufen. Beginnen Sie dann mit der Erfassung des Suchbegriffs. Wenn das gesuchte Ergebnis am Anfang der Liste erscheint, drücken Sie die Eingabetaste.

Die AWS-Sitzung bleibt nach der Anmeldung in der AWS-Managementkonsole für zwölf Stunden aktiv. Um sich erneut anzumelden und die Arbeit fortzusetzen, klicken Sie auf „Erneut anmelden“.

## Abrechnung und Kostenverwaltung

Bei der Einrichtung des AWS-Kontos müssen die Kreditkartendaten angegeben werden. Diese Karte wird mit den monatlichen Kosten für Nutzung der AWS-Services im betreffenden Konto belastet. Eine der wichtigsten Aufgaben des AWS-Administrators ist es also, die Ausgaben für AWS zu protokollieren. Hierfür muss die Nutzung überwacht werden, und die für die Nutzung der AWS-Services im betreffenden AWS-Konto anfallenden Kosten müssen analysiert und geprüft werden.

Um die Gebühreninformationen einzusehen, ist eine entsprechende Berechtigung nötig. Ihre Abrechnungsdaten finden Sie wie folgt:

1. Gehen Sie zur Navigationsleiste, und wählen Sie das Konto aus, für das Sie die Abrechnungsinformationen einsehen möchten.
2. Wählen Sie „Mein Fakturierungs-Dashboard“.
3. Das AWS-Dashboard „Fakturierungs- und Kostenverwaltung“ bietet sowohl eine Übersicht als auch eine Aufschlüsselung der monatlichen Ausgaben.

Mit der Fakturierungs- und Kostenverwaltung können Sie folgende Aufgaben ausführen:

- Schätzung und Planung der AWS-Kosten.
- Festlegung eines Kostenkontingents mit Warnungen bei Überschreitung dieses Limits.
- Übersicht über Ihre größten Investitionen in AWS-Ressourcen.
- Verwaltung mehrerer AWS-Konten an einer zentralen Stelle.

---

## 4.5 Integration in die Kerndienste

Die Planung für die Bereitstellung von SAP S/4HANA in AWS muss die Einrichtung der SAP-Umgebung umfassen. Mit der IaaS-Funktionalität von AWS wird die gesamte Hosting-Infrastruktur für die SAP-Anwendung als On-Demand-Service in AWS bereitgestellt.

### 4.5.1 Wichtigste Komponenten

Für das Hosting von AWS-Anwendungen als IaaS muss eine Virtual Private Cloud (VPC) in AWS vorbereitet werden. Die VPC ist ein logisch getrennter privater Bereich in der AWS-Cloud, in dem Sie die AWS-Ressourcen aufrufen können, die ausschließlich für Sie bestimmt sind. Dies ist mit einem für die Implementierung und Verwendung einer Anwendung vorgesehenen Bereich Ihres On-Premise-Rechenzentrums vergleichbar.

Kunden, der die Implementierung von SAP S/4HANA planen, verfügen in der Regel über ein eigenes, bereits vorhandenes Unternehmensnetzwerk. Es ist ratsam und in den meisten Fällen erforderlich, die VPC von AWS in das vorhandene Unternehmensnetzwerk zu integrieren, um die kritischen On-Premises-Funktionalitäten auch in AWS

nutzen zu können. Hierzu zählen unter anderem Active Directory, SAML-basierte Active Directory Federation Services (ADFS) für Single-Sign-On (SSO). SAP BW kann über SFTP auf den Dateiserver der Legacy-Anwendungen zugreifen.

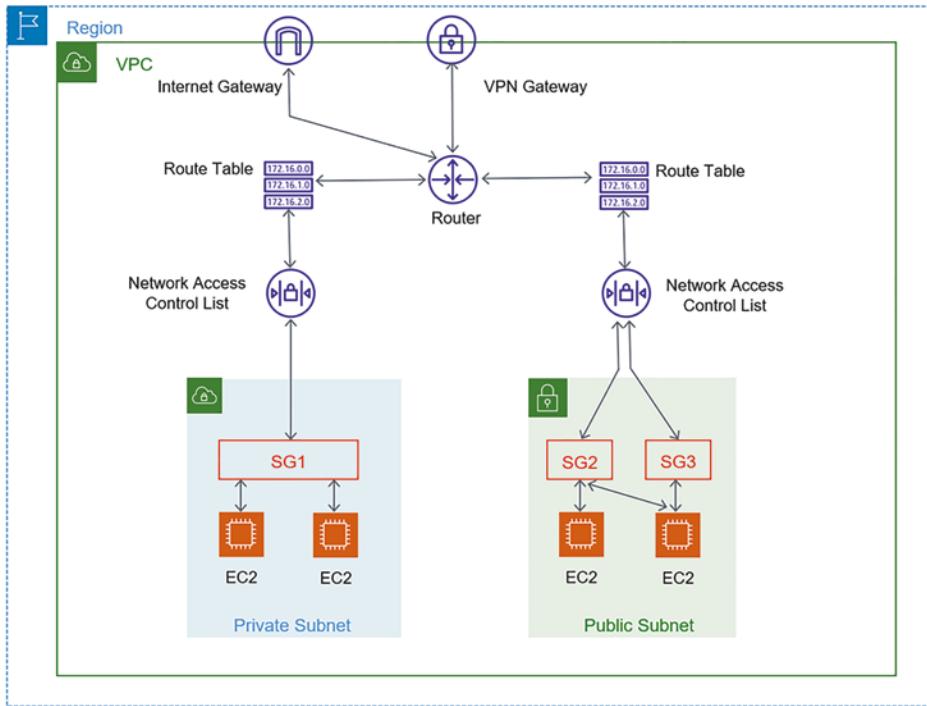
Die meisten Unternehmen verwenden für die Verwaltung ihrer Benutzer und Identitäten Active Directory mit einer On-Premises-Konfiguration. Um zu verhindern, dass Benutzerdaten mit Benutzerkennung und Passwort nicht das Unternehmensnetzwerk verlassen, ziehen es die Unternehmen vor, die Identität in AWS über das lokale Active Directory zu verwalten. In einem solchen Szenario muss die VPC von AWS in die lokalen Dienste integriert werden.

Unter Sicherheitsaspekten wird bei SAP-Bereitstellungen in der AWS-Cloud für die Komponenten stets das minimal zulässige Zugriffsmodell gewählt. Zu diesem Zweck werden die Komponenten gruppiert, und die Trennung wird durch Platzierung in verschiedenen Subnetzen erreicht. Die Kommunikation wird durch Sicherheitsgruppen (SG), Routingtabellen (RT) und Kontrolllisten für den Netzwerzugang (NACL) gesteuert.

- **Routingtabellen** sind Sätze von Netzregeln, durch die die Richtung des Datenverkehrs im Netzwerk festgelegt ist.
- **Sicherheitsgruppen** sind zustandsorientierte Firewalls auf EC2- oder NIC-Ebene, die sowohl den eingehenden als auch den ausgehenden Datenverkehr kontrollieren.
- **NACLs** sind zustandslose Firewalls ähnlich den Firewalls der lokalen Netzwerke. Sie laufen auf Subnetzebene und kontrollieren den ein- und ausgehenden Datenverkehr (Abb. 4.8).

In einer SAP-Umgebung mit SAP S/4HANA kann die Platzierung von Systemen und Komponenten sowie der zugehörige Zugriffsmechanismus wie folgt erläutert werden:

- Die Datenbanken in SAP-Systemen wie S/4HANA, BW/4HANA, GTS, GRC etc. sind die wichtigsten Vermögenswerte eines jeden Unternehmens und werden daher streng geschützt. Diese Datenbanken können in einem Subnetz platziert werden: dem Datenbank-Subnetz.
- Die Hauptkomponente von SAP-Systemen sind die Anwendungsserver. Auf sie greifen sowohl die geschäftlichen Benutzer als auch die anderen Anwendungen zu. Die Zugriffsanforderung kann von Benutzern oder Anwendungen im Unternehmensnetz (Intranet) oder von externen Benutzern oder Anwendungen (Internet) stammen. Daher werden diese Benutzer und Anwendungen in getrennten Subnetzen platziert, und der Zugang wird mit unterschiedlichen SGs und NACLs kontrolliert. So kann der Anwendungsserver mit der Datenbank und den Benutzern kommunizieren. Middleware-Komponenten von SAP wie PO. SAP-Systeme wie SAP Router, SAP Web Dispatcher (oder ein anderer Reverse Proxy) oder SAP Cloud Connector



**Abb. 4.8** Integration mit zentralen Services

benötigen Zugriff aus dem Internet. Daher sollten diese Systeme in einem getrennten Subnetz platziert werden.

- SaaS-Anwendungen wie SAP Fieldglass, Concur oder Workday können ebenfalls eine Verbindung zu der SAP-Umgebung aufbauen. Die Konnektivität wird über das oben beschriebene öffentliche Subnetz verwaltet.
- Darüber hinaus gibt es Verwaltungskomponenten wie Active Directory, DNS-Server und Überwachungslösungen wie SAP Solution Manager, die eine Verbindung zu allen Umgebungskomponenten aufbauen müssen.

Um diese Konnektivität zu gewährleisten, stehen Ihnen folgende AWS-Services zur Verfügung.

### VPN-Gateway

Um ein lokales Netzwerk über VPN mit AWS zu verbinden, ist AWS-seitig ein Virtual Private Gateway (auch „VPN-Gateway“) und kundenseitig ein Kunden-Gateway erforderlich. Ein VPN-Gateway ist immer einer VPC zugeordnet.

### NAT-Gateway und NAT-Instanz

Mit einem NAT-Gateway und einer NAT-Instanz können Instanzen in privaten Subnetzen leichter mit dem Internet, anderen VPCs und On-Premises-Netzwerken verbunden werden. Die Kommunikation über die NAT-Geräte wird so eingerichtet, dass nur ausgehender Datenverkehr erlaubt ist. Eingehende Anfragen werden abgelehnt. Die NAT-Instanz ist ein von AWS verwaltetes NAT-Gerät, während es sich beim NAT-Gateway um einen vom Kunden verwalteten und auf EC2-Instanzen erstellten NAT-Service handelt.

### Internetgateway (IGW)

Diese VPC-Komponente vereinfacht die Kommunikation zwischen der VPC und dem Internet. Ein Internetgateway erfüllt die folgenden zwei Funktionen:

- Es bietet in den VPC-Routingtabellen ein Ziel für den Datenverkehr, der über das Internet geroutet werden kann.
- Weiterhin bietet das Internetgateway eine Netzwerkadressübersetzung (NAT) für Instanzen, denen öffentliche IPv4-Adressen zugewiesen wurden.

In der Regel wird bei der Erstellung einer VPC auch ein Internetgateway angelegt.

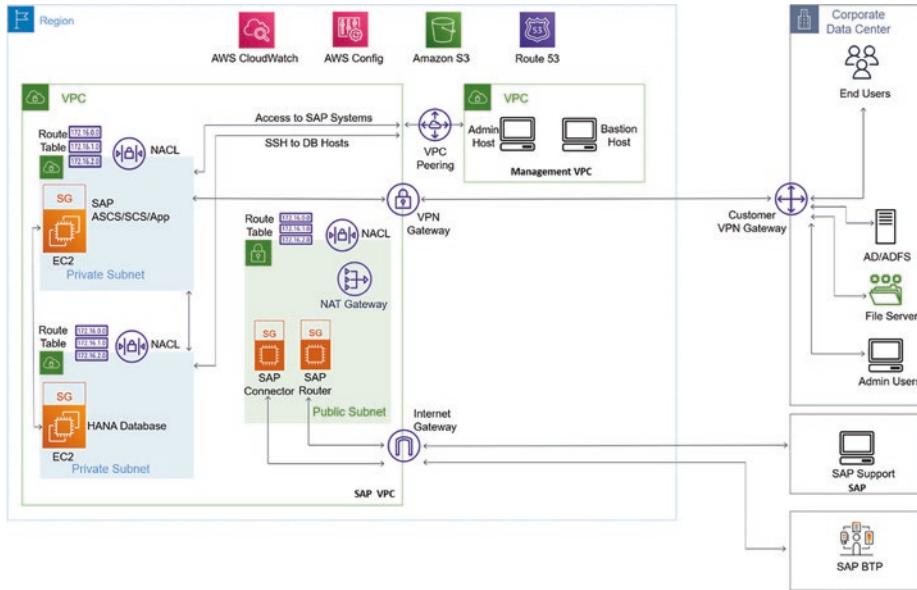
### Route 53

Route 53 ist ein hochgradig verfügbares und skalierbares DNS (Cloud Domain Name System). Es erleichtert die Übersetzung von menschenlesbaren Namen wie [www.example.com](#) in die numerischen IP-Adressen wie 172.0.10.10. Um uns mit der Integration von SAP-Systemen in die AWS-Cloud vertraut zu machen, betrachten wir einige gängige Architekturkonfigurationen.

## 4.5.2 Interne Zugriffsarchitektur

Bei dieser Konfiguration wird nur über das Unternehmensnetzwerk auf die SAP-Anwendung in AWS zugegriffen. Diese Konfiguration mit Hub-and-Spoke-Modell umfasst vier Subnetze:

- **Öffentliches Subnetz** – Dieses Subnetz enthält einen SAProuter und ein NAT-Gateway oder eine NAT-Instanz.
- **Anwendungssubnetz** – SAP-Anwendungsserver befinden sich im Anwendungssubnetz. Der Zugriff erfolgt über die SAP GUI oder über die HTML GUI mithilfe eines Browsers aus dem Unternehmensnetzwerk. Für den SAP-Support ist das Subnetz mit dem öffentlichen Subnetz verbunden, wobei das NAT-Gateway oder die NAT-Instanz die Verbindung zum SAP-Support ermöglicht.
- **Datenbanksubnetz** – Da die Datenbanken der zentrale Vermögenswert des Unternehmens sind, werden sie in diesem separaten Subnetz platziert. Eine Verbindung ist



**Abb. 4.9** Intranet-basierter Zugriff

nur aus dem Anwendungssubnetz erlaubt. Zusätzlich kann es erforderlich sein, die HANA-Datenbank über HANA Studio zu verbinden. Hierfür kann Port 3<instance\_number>15 zugelassen werden.

- **Verwaltungssubnetz** – Dabei handelt es sich um ein privates Subnetz, in dem Administrations-, Überwachungs- und Verwaltungstools wie Bastion Host (RDP Server), Microsoft System Center Configuration Manager (SCCM), HANA Studio oder BODS Database Designer gehostet werden.
- Nur spezifizierte öffentliche SAP-IPs dürfen sich mit dem SAProuter verbinden. Einzelheiten finden Sie im SAP-Hinweis 28.976, Datenblatt Remote-Verbindung (Abb. 4.9).

Die Anwendungen in diesem Subnetz werden sowohl für den Support als auch für einige Aufgaben benötigt, die die Installation von FAT-Clients erfordern, darunter SAP HANA Studio und SAP BODS Designer. Für die Verwaltungsschicht können die Hosts wie oben beschrieben in einer getrennten VPC untergebracht werden. Eine andere Möglichkeit ist, sie in einer SAP-VPC in einem getrennten Subnetz zu hosten. Das obige Szenario illustriert die Konfiguration in einer einzigen Region. Um eine hochgradige Verfügbarkeit zu gewährleisten, erstreckt sich die Konfiguration in AWS auf mehrere Availability Zones. Für die Notfallwiederherstellung erstreckt sich die Konfiguration auf mehrere AWS-Regionen.

### 4.5.3 Interner und begrenzter externer Zugriff

Neben dem Zugriff auf SAP-Anwendungen über das Unternehmensnetzwerk (Intranetzugang) muss sich das System auch mit einigen vertrauenswürdigen externen Anwendungen verbinden können, die in der SAP-Welt normalerweise als vertrauenswürdige Drittanbieteranwendungen bezeichnet werden. SAP Process Orchestration (PO) / Process Integration (PI) oder Nicht-SAP-Produkte wie MuleSoft sind die Middleware-Komponenten, welche, in einem solchen Szenario, die Brücke schlagen. Diese Konfiguration kann mit den folgenden Einstellungen umgesetzt werden. Entweder es kommt eine Virtual-Private-Network-Verbindung (VPN) sowohl für den eingehenden als auch für den ausgehenden Datenverkehr zum Einsatz. Oder es wird ein Elastic Load Balancing (ELB) für den eingehenden Datenverkehr und ein Network-Address-Translation-Gateway (NAT) für den ausgehenden Datenverkehr verwendet. Beide Optionen werden in den nachfolgenden Teilkapiteln beschrieben.

#### 4.5.3.1 Architektur mit VPN-Verbindungen

Bei dieser Konfiguration erfolgt die Verwaltung des eingehenden und des ausgehenden Datenverkehrs entweder durch:

- ein Virtual Private Gateway (VGW) in Ihrer VPC oder
- einen speziellen softwarebasierten VPN-Server. Softwarebasierte VPN-Server sind in AWS Marketplace verfügbar.

Die VPN-Komponente wird zu dem öffentlichen Subnetz hinzugefügt (Abb. 4.10).

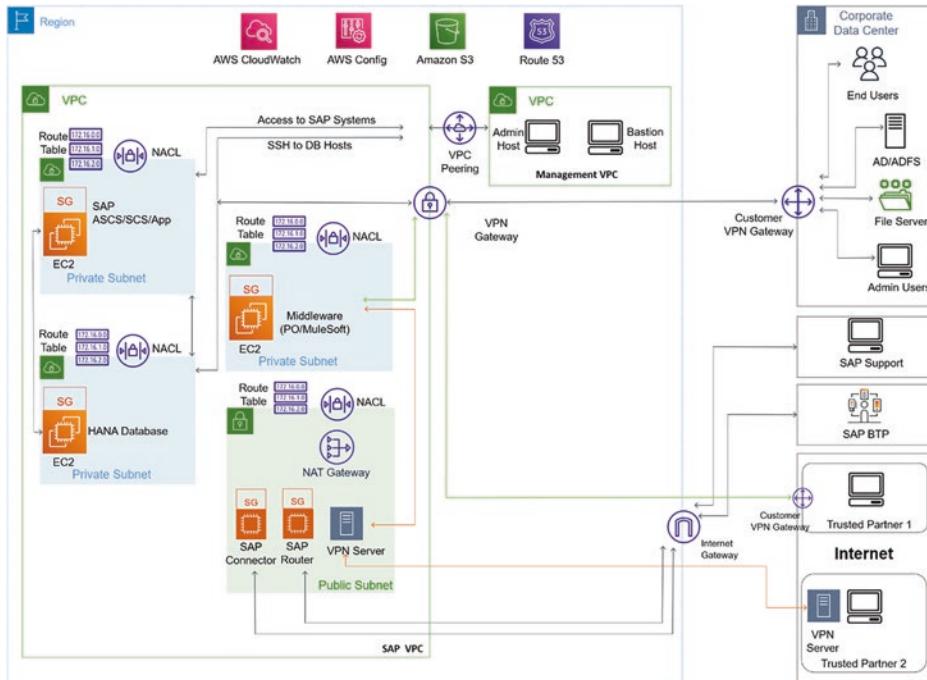
#### 4.5.3.2 Architektur mit Elastic Load Balancing (eingehender DV)/NAT-Gateway (ausgehender DV)

Es gibt zwei Arten von Lastverteilung, welche für SAP S/4HANA Systeme in AWS relevant werden: die Lastverteilung auf **Netzwerkebene** und die Lastverteilung auf **Anwendungsebene**.

Bei der Lastverteilung auf **Netzwerkebene** wird der TCP-Datenverkehr verteilt, während bei der Lastverteilung auf Anwendungsebene in der Anwendungsschicht gearbeitet wird, was für die Verteilung des HTTP- oder HTTPS-Datenverkehrs die optimale Wahl ist. In dem folgenden Architekturbeispiel wird extern aus dem Internet auf die VPC von SAP zugegriffen. Dies soll an zwei Szenarien verdeutlicht werden.

##### Szenario: Zugang zum SFTP-Server im privaten Subnetz

Hier geht es um den Zugang zu einem SFTP-Server in einem privaten Subnetz, auf dem beispielsweise ein vertrauenswürdiger Partner, wie eine Bank, Dateien ablegen kann (eingehender Datenverkehr). Eine Möglichkeit für die Gewährleistung der Konnektivität ist ein in einem öffentlichen Subnetz der VPC von SAP platziert, dem Internet zugewandter Network Load Balancer (NLB). Der Zugriff von vertrauenswürdigen

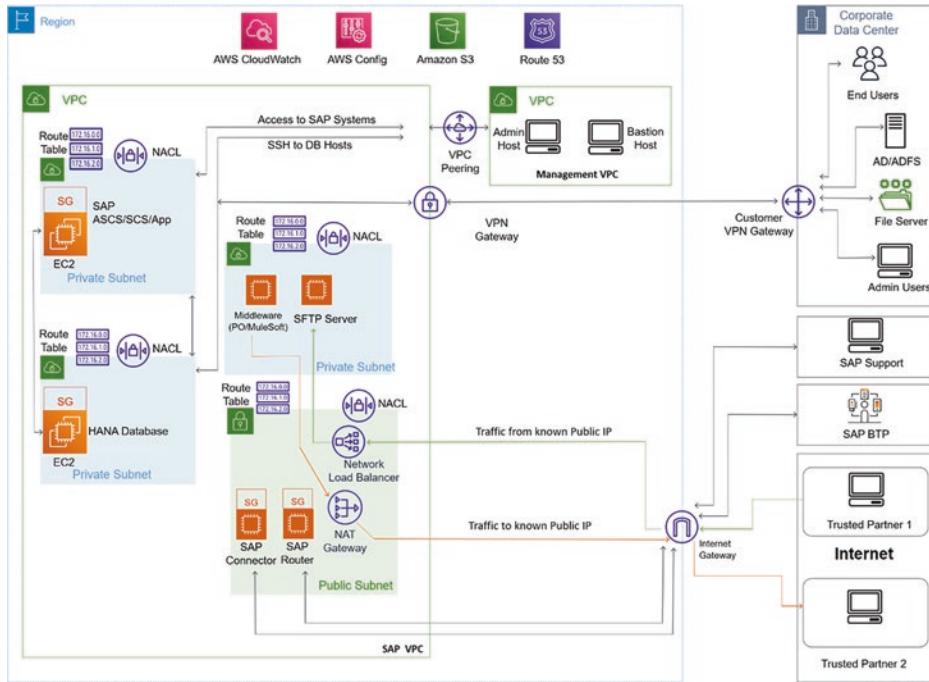


**Abb. 4.10** Interner und externer Zugriff

Partnern auf das Subnetz des SFTP-Servers wird mithilfe der Sicherheitsgruppe kontrolliert. Für den ausgehenden Datenverkehr z. B. von SAP PI/PO/MuleSoft zu den SFTP-Servern der vertrauenswürdigen externen Partner kann ein NAT-Gateway verwendet werden. Mit Amazon Route 53 kann der externe Domänennamen Ihres Unternehmens registriert werden. Weiterhin besteht die Möglichkeit, vollständig qualifizierte Domänennamen (Fully Qualified Domain Names – FQDNs) für den Load Balancer aufzulösen (Abb. 4.11).

### Szenario: HTTPS-basierte Webservice-Schnittstellen von SAP PI/PO

In diesem Szenario werden die Webservices in PI/PO von den externen Partnern mit dem HTTPS-Protokoll aus dem Internet aufgerufen, und PI/PO kann sich über das Internet mit externen vertrauenswürdigen Partnern verbinden. Der eingehende Datenverkehr (von Drittanbietern zu PI/PO) kann per anwendungsbasierten Lastverteilung abgewickelt werden. Der Zugriff bekannter IPs wird mithilfe der Sicherheitsgruppe kontrolliert, die wiederum mit dem Load Balancer verbunden ist. Der ausgehende Datenverkehr (von PI/PO zu Drittanbietern) mithilfe des NAT-Gateways aktiviert und kontrolliert werden. Amazon Route 53 kann für die Registrierung von Domänennamen und für die FQDN-Auflösung verwendet werden (Abb. 4.12).



**Abb. 4.11** Architekturbeispiel mit Elastic Load Balancing für Ingress und NAT Gateway für Egress

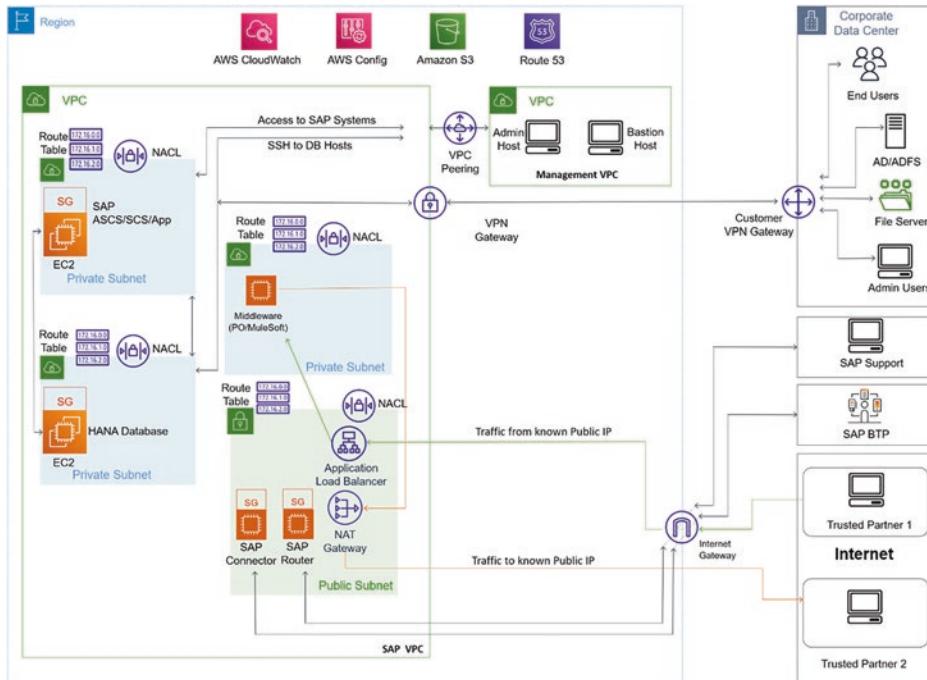
#### 4.5.4 Interner und vollständiger externer Zugriff

In diesem Szenario ist neben dem internen Zugriff auch der Zugriff über das Internet erforderlich. Typische Anwendungsszenarien wären der Zugriff auf das System über SAP Fiori oder die Existenz eines dem Internet zugewandten SAP Enterprise Portals (Abb. 4.13).

Der gesamte Internetdatenverkehr zu SAP kann über einen Application Load Balancer (ALB) im öffentlichen Subnetz abgewickelt werden. Der Datenverkehr geht vom ALB an den SAP Web Dispatcher, der in SAP S/4HANA als Reverse Proxy für SAP-Fiori-Systemanwendungen dient.

Der Schutz des ALBs kann mit den folgenden AWS-Services gewährleistet werden:

- **AWS Web Application Firewall (WAF)** für Application Load Balancers: Abwehr der häufigsten Angriffe auf die Anwendungsschicht wie SQL-Injection-Angriffen. Die AWF in Kombination mit einem ALB bietet zusätzlichen Schutz, weil der Datenverkehr von vordefinierten IP-Adressen mithilfe von Zugriffskontrolllisten (Access Control Lists – ACLs) sowie Regeln und Bedingungen kontrolliert wird. Referenzblog (Zugriff am 20.12.2021): <https://aws.amazon.com/blogs/aws/aws-web-application-firewall-waf-for-application-load-balancers/>



**Abb. 4.12** HTTPS basierte Web Service Schnittstelle mit SAP PI/PO

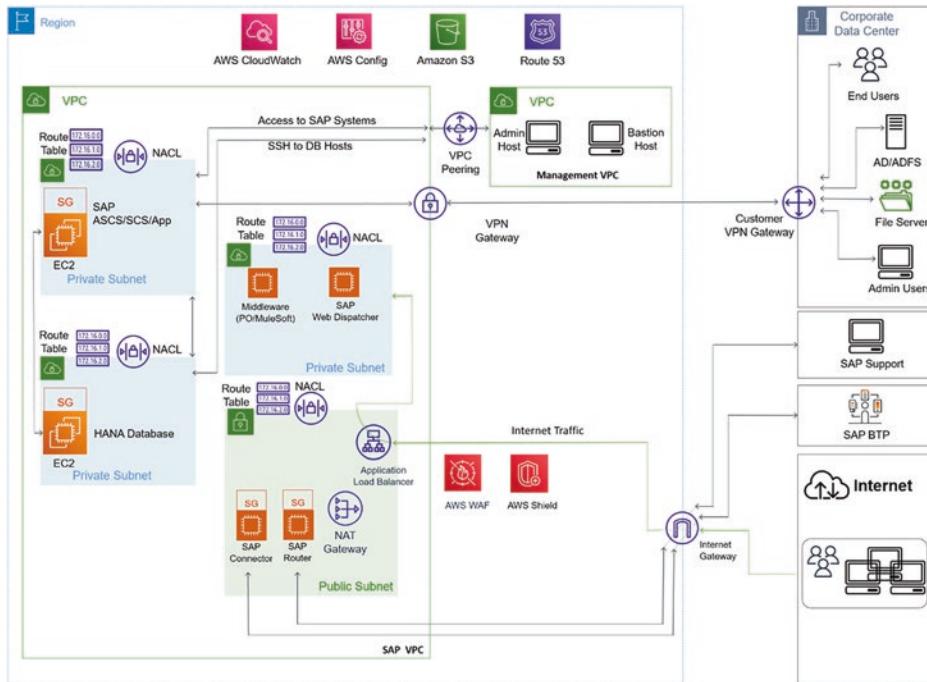
- Das **AWS Shield** ist ein verwalteter DDoS-Schutzdienst (Distributed Denial of Service) für in AWS laufende Webanwendungen. Es gibt zwei Versionen des AWS Shields: Standard und Erweitert. Für das AWS Shield in der Standardversion fallen keine zusätzlichen Kosten an.

### Umgebungseinrichtung

Als Beispiel soll die am weitesten verbreitete SAP-S/4HANA-Greenfield-Implementierung dienen. Die zu installierenden SAP-Komponenten umfassen:

- SAP S/4HANA 2021 mit eingebettetem FIORI
- SAP Web Dispatcher als Reverse Proxy
- SAP Adobe Document Service zur Unterstützung des pdf-Dienstes
- SAP BODS für die Datenmigration in SAP S/4HANA
- SAP Solution Manager als zentrales Überwachungssystem
- SAP Router für die einfachere Verbindung mit dem SAP-Support
- SAP Cloud Connector zur Verbindung mit der SAP Business Technology Platform

Je nach Bedarf lassen sich auch weitere SAP-Systeme in die Implementierung einbeziehen, da der für das erste System verfolgte Ansatz auch auf alle weiteren Systeme angewendet werden kann. Die SAP-Systeme werden in der VPC von AWS



**Abb. 4.13** Voller interner und externer Zugriff

implementiert. Die Implementierung mit zusätzlich mit den folgenden, bereits in der Kundenumgebung vorhandenen On-Premises-Komponenten verbunden:

- AD/ADFS als zentralem Benutzerspeicher
- Legacy-Dateiserver
- Außerdem stellen die Administratoren für die Administration, Verwaltung und Überwachung im lokalen Netzwerk eine Verbindung zu AWS her.

### Vorgeschlagene Architektur

Die Konfiguration kann wie folgt vorbereitet werden:

- VPC 1 für das Hosting der SAP-Systeme
- VPC 2 für das Hosting der Verwaltungssysteme

Erstellung der folgenden Subnetze in der VPC 1:

- Privates Subnetz 1 für das Hosting der Datenbanken
- Privates Subnetz 2 für das Hosting der SAP-Anwendungen
- Öffentliches Subnetz 1 für das Hosting der externen Anwendungen wie SAP Router und Cloud Connector für die Verbindung mit dem SAP Support Hub und der SAP Business Technology Platform (BTP)

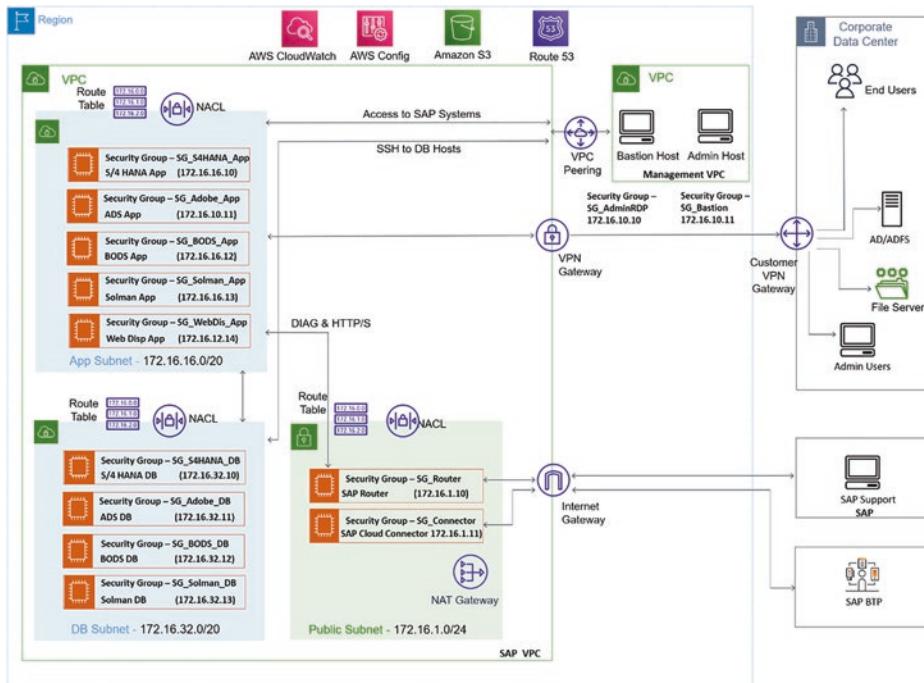
Erstellung der folgenden Subnetze in der VPC 2:

- Privates Subnetz 3 für Verwaltungsinstanzen

Die Sicherheit der einzelnen Subnetze wird über die Routingtabelle und die Network Access Control List (NACL) gewährleistet. Wie oben beschrieben umfasst die Routingtabelle eine Reihe von Regeln. Diese werden als Routes bezeichneten und legen den Datenverkehr fest. Der ein- und ausgehende Datenverkehr wird mithilfe von NACLs mit entsprechenden Autorisierungs- und Verweigerungsregeln kontrolliert. Um sicherzustellen, dass nur die ausdrücklich zugelassenen Systeme bzw. IPs mit der die Anwendung bzw. Datenbank hostenden EC2-Komponente interagieren können, kann darüber hinaus pro SAP-Anwendung und -Datenbank eine Sicherheitsgruppe erstellt werden.

Weitere für die Einrichtung der SAP-Umgebung erforderliche AWS-Services:

- NAT-Gateway – NAT-Service (Network Address Translation) für die einfachere Verbindung von Systemen in privaten Subnetzen mit externen Systemen
- Amazon CloudWatch – Sammeln und Analysieren von Überwachungsdaten
- Amazon S3 – langfristige Speicherlösung für Backups
- Amazon Route 53 – Amazon DNS-Dienst
- AWS Config – Service zur Einschätzung, Prüfung und Evaluierung der Konfigurationen der bereitgestellten AWS-Ressourcen



Die folgenden Konfigurationsempfehlungen haben unverbindlichen Charakter. Bei der Konfiguration sollten die folgenden Punkte zusätzlich berücksichtigt werden:

- In der für die Verwaltung verwendeten VPC können zwei Hosts berücksichtigt werden: einer für das Verwaltungsteam und ein weiterer für das Projektteam. Es ist aber auch möglich, sich für einen Host mit den entsprechenden Zugangskontrollen zu entscheiden.
- Die Sicherheitsgruppenkonfiguration ist für jede Anwendung getrennt vorzunehmen. Die EC2s mit ähnlichem Nutzungstyp können für identische Zugriffsmechanismen zusammengefasst und derselben Sicherheitsgruppe zugewiesen werden.
- <http://docs.aws.amazon.com/AWSEC2/latest/WindowsGuide/security-group-rules-reference.html> Weitere Informationen zu Sicherheitsgruppen sind in der AWS-Dokumentation „Sicherheitsgruppenregeln für verschiedene Anwendungsfälle“ zu finden.

<http://docs.aws.amazon.com/AWSEC2/latest/WindowsGuide/security-group-rules-reference.html>

Eine mögliche Subnetzkonfiguration ist der nachstehenden Tabelle zu entnehmen (Tab. 4.2).

Die Regeln für SG, NACL und Routingtabellen richten sich nach der für die SAP-Anwendung und HANA-DB ausgewählten Instanznummer. Detaillierte Angaben zur

**Tab. 4.2** Mögliche Subnetzkonfiguration

	CIDR	Routing-Tabelle	NACL	System	IP	Sicherheitsgruppe
<b>SAP VPC</b>	172.16.0.0/16	--	--	--	--	--
<b>VPC-Verwaltung</b>	172.16.10.0/24	--	--	--	--	--
<b>Customer DC</b>	172.16.128.0/18	--	--	--	--	--
<b>Öffentliches Subnetz</b>	172.16.1.0/24	RT_Sub_Public	NACL_Public	SAP-Router	172.16.1.10	SG_Router
				SAP Cloud Connector	172.16.1.11	SG_Connector
<b>Anwendungs-Subnetz der SAP-VPC</b>	172.16.16.0/20	RT_Sub_App_Private	NACL_App_Private	SAP S/4 HANA	172.16.16.10	SG_S4HANA_App
				SAP Adobe Service	172.16.16.11	SG_Adobe_App
				SAP BODS	172.16.16.12	SG_BODS_App
				SAP Solution Manager	172.16.16.13	SG_Solman_App
				SAP Web Dispatcher	172.16.16.14	SG_WebDisp_App
				SAP S/4 HANA DB	172.16.32.10	SG_S4HANA_DB
<b>VPC von SAP – Datenbanksubnetz</b>	172.16.32.0/20	RT_Sub_DB_Private	NACL_DB_Private	SAP Adobe Service DB	172.16.32.11	SG_Adobe_DB
				SAP BODS DB	172.16.32.12	SG_BODS_DB
				DB von SAP Solution Manager	172.16.32.13	SG_Solman_DB
				Admin-RDP-Host	172.16.10.10	SG_AdminRDP
<b>Verwaltungs-VPC – Admin-Host</b>		RT_Sub_Adm_Private	NACL_Adm_Private	Bastion-Host	172.16.10.11	SG_Bastion
<b>Verwaltungs-VPC – Bastion-Host</b>		RT_Sub_Bast_Private	NACL_Bast_Private			

Konfiguration der SAP-Ports können unter „TCP/IP Ports of All SAP Products“ eingesehen werden (Zugriff am 20.12.2021): <https://help.sap.com/viewer/ports> ◀

#### 4.5.5 Active Directory

Microsoft Active Directory ist ein von Microsoft entwickelter Verzeichnisdienst. Für die Verwaltung verzeichnisbasierter Dienste ist es eines der gängigsten Produkte und wird von den meisten Unternehmen verwendet. Zu solchen verzeichnisbasierten Diensten zählen:

- Domänedienste
- Verzeichnisdienst Light Weight
- Zertifikatsdienst
- Rechteverwaltungsdienste
- Verbunddienste

Active Directory wurde im Jahr 1999 allgemein verfügbar gemacht. Im Laufe der Jahre wurden mehrere Funktionen und Funktionalitäten hinzugefügt. Der für die Domänenverwaltung genutzte Teil von Active Directory wurde in Active Directory Domain Service (AD DS) umbenannt. Darüber hinaus wurde Active Directory zu einem Überbegriff für eine breitere Palette verzeichnisbasierter Dienste.

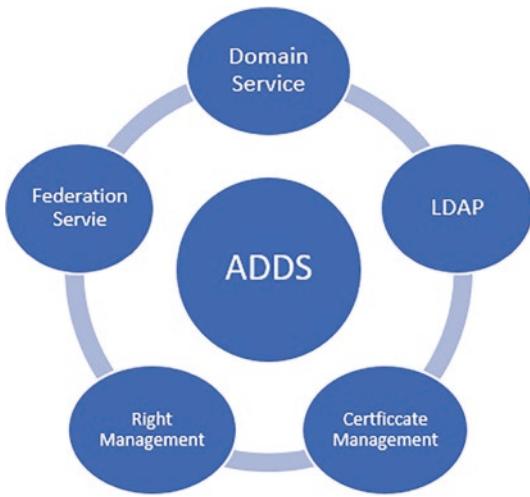
AD DS wird für die Verwaltung der Unternehmensdomäne verwendet. Es speichert Informationen über Benutzer, Computer, Dateisysteme, Drucker und sonstige Ressourcen und bietet einen Authentifizierungs- und Autorisierungsdienst. AD DS wird außerdem für die Verwaltung der Gruppenrichtlinien, für die Verschlüsselung der in der Umgebung vorhandenen Systeme, für Domänennamendienste, Remote-Desktop-Dienste, Austauschdienste und für die Verwaltung des SharePoint-Servers verwendet (Abb. 4.14).

In der AWS-Cloud gibt es für die Verwendung der Active Directory Domain Services drei Optionen:

- **Option 1:** AWS-verwalteter Microsoft Active Directory Service
- **Option 2:** Selbst verwaltete AD-DS-Bereitstellung in der AWS-Cloud
- **Option 3:** Erweiterung des lokalen AD DS in AWS

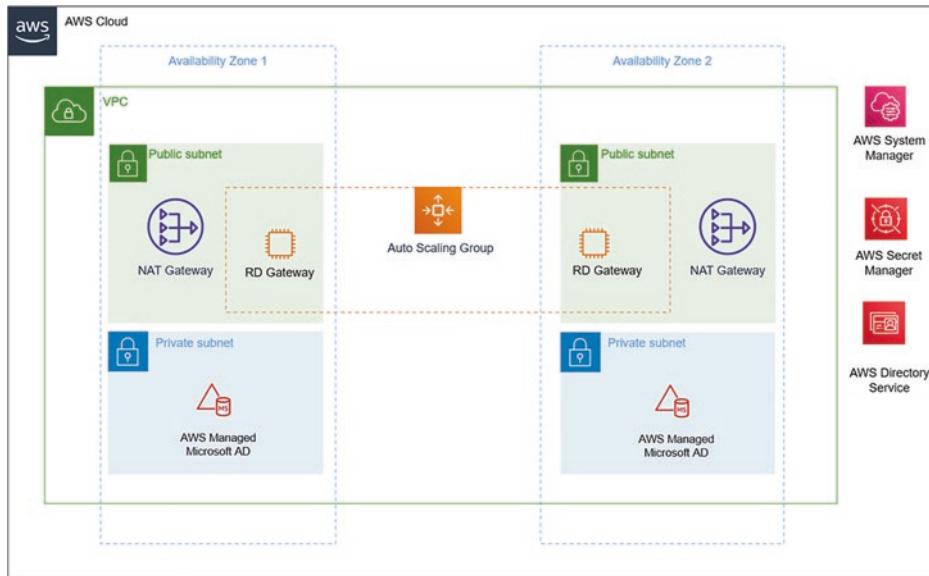
AWS empfiehlt, ein getrenntes Konto für Identitätsdienste wie AD einzurichten und den Zugriff auf dieses Konto auf eine kleine Gruppe von Administratoren zu beschränken. Um eine dieser drei Lösungen zu implementieren, muss sichergestellt sein, dass AWS AD in einer VPC bereitgestellt wird, auf die von allen VPCs mit von Active Directory abhängigen Workloads aus zugegriffen werden kann. Auch in diesem Szenario kann die VPC auf Basis des Hub-and-Spoke-Modells bereitgestellt werden. Die VPC von AD fungiert dabei als Hub, und alle VPCs der Directory-Benutzer werden mittels

**Abb. 4.14** Active Directory Domain Service



VPC-Peering mit dieser AD-VPC verbunden. Die am weitesten verbreitete Architektur wird im folgenden Abschnitt beschrieben.

#### 4.5.5.1 Option 1: AWS-verwalteter Microsoft Active Directory Service (Siehe Abb. 4.15)



**Abb. 4.15** Managed Microsoft Active Directory Service in der AWS Cloud

Bei dieser Lösung wird die Active-Directory-Instanz von AWS gewartet. Das Setup für den DS von AD umfasst die folgenden Komponenten:

- **VPC**
- Öffentliche und private Subnetze in zwei Availability Zones für hohe Verfügbarkeit
- Die öffentlichen Subnetze in den Availability Zones verfügen über ein **NAT-Gateway** und über ein **RD-Gateway**.
  - Mit den NAT-Gateways kann dem ausgehenden Datenverkehr Zugriff auf die Ressourcen in den privaten Subnetzen gewährt werden.
  - Die RD-Gateway-Instanzen befinden sich in einer Auto-Scaling-Gruppe und sichern den Remote-Zugriff auf die Instanzen in den privaten Subnetzen ab.
- AWS Systems Manager Automation für Einrichtung und Konfiguration des DS von AD und des in AD integriertem DNS
- AWS Secrets Manager zum Speichern der Passwörter
- AWS Directory Service zur Bereitstellung und Verwaltung des DS von AD in den privaten Subnetzen

#### 4.5.5.2 Option 2: Selbst verwaltete AD-DS-Bereitstellung in der AWS-Cloud

(Siehe Abb. 4.16)

Bei dieser Lösung werden die Active-Directory-Instanzen vom IT-Team des Kunden installiert und verwaltet. Die Konfiguration für den DS von AD umfasst die folgenden Komponenten:



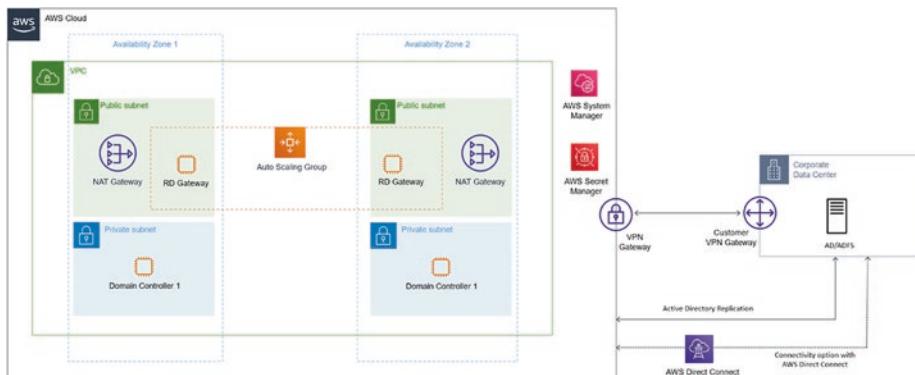
**Abb. 4.16** Eigene Instanz von Active Directory Domain Domain Service in AWS

- Eine für hohe Verfügbarkeit mit öffentlichen und privaten Subnetzen in zwei Availability Zones konfigurierte **VPC**
- Öffentliche Subnetze in den AZs mit einer **NAT-Gateway**- und einer **RD-Gateway**-Instanz
  - Mit den NAT-Gateways kann dem ausgehenden Datenverkehr Zugriff auf die Ressourcen in den privaten Subnetzen gewährt werden.
  - Die RD-Gateway-Instanzen befinden sich in einer Auto-Scaling-Gruppe und sichern den Remote-Zugriff auf die Instanzen in den privaten Subnetzen ab.
- **Private Subnetze** in den AZs mit:
  - einem Windows-Server-Park und Domänenfunktionen mit Sicherheitsgruppen und Regeln für den Datenverkehr zwischen den Instanzen
- AWS-Systems-Manager-Automation-Dokumente für Einrichtung und Konfiguration des DS von AD und des in AD integriertem DNS
- **AWS Secrets Manager** zum Speichern der Passwörter

#### 4.5.5.3 Option 3: Erweiterung des lokalen Active-Directory-DS in AWS

(Siehe Abb. 4.17)

Bei dieser Lösung fungiert die Active-Directory-Instanz im Unternehmensrechenzentrum als primärer Domain Controller. Die im privaten Subnetz eingerichteten Windows-Instanzen werden an das lokale AD angebunden und als Domain Controller bekannt gegeben. Die abgesicherte Konnektivität von AWS zum Unternehmensrechenzentrum wird durch ein VPN-Gateway gewährleistet, das mit dem VPN-Gateway des Kunden verbunden ist. Für die Verbindung kann auch AWD Direct Connect zurückgegriffen werden.



**Abb. 4.17** Erweiterung eines On-Premise AD zu AWS AD

Die Konfiguration für den DS von AD umfasst die folgenden Komponenten:

- Eine für hohe Verfügbarkeit mit öffentlichen und privaten Subnetzen in zwei Availability Zones konfigurierte **VPC**
- Öffentliche Subnetze in den AZs mit einer **NAT-Gateway**- und einer **RD-Gateway**-Instanz
  - Mit den NAT-Gateways kann dem ausgehenden Datenverkehr Zugriff auf die Ressourcen in den privaten Subnetzen gewährt werden.
  - Die RD-Gateway-Instanzen befinden sich in einer Auto-Scaling-Gruppe und sichern den Remote-Zugriff auf die Instanzen in den privaten Subnetzen ab.
- **Private Subnetze** in den AZs mit:
  - Windows-Server-Park und Domänenfunktionen mit Sicherheitsgruppen und Regeln für den Datenverkehr zwischen den Instanzen
- AWS-Systems-Manager-Automation-Dokumenten für Einrichtung und Konfiguration des DS von AD und des in AD integriertem DNS
- **AWS Secrets Manager** zum Speichern der Passwörter

Details zur Bereitstellung finden Sie in der AWS-Dokumentation (Zugriff am 20.12.2021): <https://aws-quickstart.github.io/quickstart-microsoft-activedirectory/>

### **AWS-verwalteter Microsoft Active Directory Service**

AWS bietet Microsoft Active Directory Service als SaaS-Lösung an. Bekannt ist diese Lösung unter dem Namen AWS Managed Microsoft Active Directory (AD). Dieser AD-Service kann von den Unternehmen in AWS genutzt werden. Er schafft eine Verbindung zwischen allen EC2-Instanzen, Datenbankinstanzen, Benutzern, Gruppen und Geräten wie z. B. dem herkömmlichen DS von Microsoft AD.

AWS Managed Microsoft Active Directory Service bietet auch die anderen oben beschriebenen AD-DS-Domain-Controller-Services, wie z. B. die Verwaltung von Gruppen- und Passwortrichtlinien, Kerberos-basiertes Single Sign-on etc. Für die Verwaltung von AWS Managed AD ist kein spezielles Tool erforderlich. Die Funktionsweise entspricht dem On-Premises-AD-DS von Microsoft.

AWS Managed Microsoft Active Directory ist eine SaaS-Lösung, in die die hochgradige Verfügbarkeit in der AWS-Infrastruktur bereits integriert ist. Der Domain Controller wird intern von AWS überwacht. Wenn ein Fehler auftritt und es zu einem Ausfall kommt, wird für Ersatz in derselben Availability Zone und mit derselben IP-Adresse gesorgt, sodass es nicht zu Unterbrechungen kommt. AWS wickelt das gesamte Patching ab und hält die Software auf dem neuesten Stand. Darüber hinaus umfasst die Konfiguration die Datenreplikation und einen täglichen automatischen Snapshot. Dies erspart den Unternehmen das Hosting, die Verwaltung, die Überwachung und der Aktualisierung von AD DS und damit eine Reihe von komplexen Aufgaben.

#### 4.5.5.4 Implementierungsaspekte

##### 4.5.5.4.1 AWS Managed Microsoft Active Directory als zentrales AD DS

Bei der Implementierung von AWS Managed Microsoft Active Directory können die vorhandenen On-Premises-Benutzer und -Ressourcen nahtlos mit dem zentralen Verzeichnisdienst verbunden und verwaltet werden. Um geschäftlichen Benutzern wie z. B. den SAP-FIORI-Benutzern eine nahtlose Single-Sign-on-Erfahrung zu bieten, kann der Dienst in Active Directory Federation Service (AD FS) und in AWS Single Sign-on integriert werden.

Auch der Microsoft Network Policy Server (NPS) kann mit Amazon Managed AD verbunden werden. Bei der Einrichtung von AWS Managed AD wird ein Administratorkonto erstellt. Es verfügt über delegierte Verwaltungsrechte für die Sicherheitsgruppe Remote Access Service (RAS) und Internet Authentication Service (IAS). Damit können NPS bei AWS Managed Microsoft AD registriert und Netzwerzkugriffsrichtlinien für Konten in der Domäne verwaltet werden.

AWS Managed Microsoft AD ist in den Produktversionen **Standard** und **Enterprise** enthalten.

- **Standard Edition:** Die Standard Edition eignet sich für kleine und mittelständische Unternehmen mit bis zu 5.000 Mitarbeitern. Sie bietet Speicherkapazität für bis zu 30.000 Verzeichnisobjekte, darunter Benutzer, Gruppen und Computer.
  - **Enterprise Edition:** Die Enterprise Edition eignet sich für Großunternehmen mit bis zu etwa 500.000 Verzeichnisobjekten.
- Bei dieser Obergrenze handelt es sich um eine Annäherung. Je nach Größe der Objekte und Verhalten und den Performancebedarf der Anwendungen kann das Verzeichnis eine höhere oder niedrigere Anzahl von Verzeichnisobjekten unterstützen.

##### Aspekte zu einem von AWS verwalteten AD mit SAP S/4HANA DR

Mit AWS Managed Microsoft AD ist es möglich, eine mehrere Regionen umfassende Replikation einzurichten. Mit dieser Konfiguration kann bei einem Ausfall in einer Region die Notfallwiederherstellung für Ihre SAP-Anwendung einschließlich S/4HANA abgerufen werden. Die Verzeichnis-ID (directory\_id) ist in der gesamten DR-Region identisch und wird in demselben AWS-Konto bereitgestellt wie die primäre Region.

Bei der Einrichtung der AD-Replikation stellt AWS Domain Controller in der DR-Region bereit und repliziert alle AD-Verzeichnisdaten der Primärregion, einschließlich Benutzer-, Gruppen-, Gruppenrichtlinienobjekt- (GPO) und Schemadaten. Das Verzeichnis zielt also in der DR-Region darauf ab, im Katastrophenfall die gewünschten Dienste anbieten zu können.

#### 4.5.5.4.2 Erstellung eines AWS Managed Microsoft Active Directory

Bei der Konfiguration eines AWS Managed Microsoft Active Directorys wird ein hochverfügbares, mit der Virtual Privat Cloud (VPC) verbundenes Domain-Controller-Paar erstellt. Die EC2-Instanzen für die Domain Controller (DC) laufen in separaten Availability Zones in Ihrer Region.

Die beiden EC2-Instanzen, auf denen die DC laufen, werden von AWS verwaltet. Dies bedeutet, dass Sie auf die DC-Betriebssysteme nicht zugreifen können. Jede EC2-Instanz verfügt über zwei elastische Netzwerkschnittstellen (ENI): ETH0 und ETH1. ETH0 ist der Verwaltungsadapter für AWS. Sie können nicht über Ihr Konto darauf zugreifen. ETH1 wird in Ihrem Konto erstellt.

- ▶ Der IP-Bereich für die Verwaltung des ETH0-Netzwerks des Verzeichnisses lautet 198.18.0.0/15. Elastic IP kann nicht mit diesen ENI verknüpft werden.

Für die Verzeichniserstellung werden folgende Angaben benötigt:

- Die von AWS verwaltete Active-Directory-Edition: Standard Edition oder Enterprise Edition von AWS Managed Microsoft AD. Zur Auswahl siehe AWS Managed Microsoft AD, Abschnitt Editionen.
  - VPC, die von AD verwaltet werden soll
  - Zwei Subnetze, in denen Active-Directory-Windows-EC2s eingesetzt werden sollen und die sich in unterschiedlichen Availability Zones befinden
- 
- ▶ Von AWS Directory Service wird die Verwendung von Network Address Translation (NAT) bei Active Directory nicht unterstützt. Die Verwendung von NAT kann zu Replikationsfehlern führen.

Die Erstellung eines solchen aktiven Verzeichnisses dauert etwa 20 bis 40 min und wird in folgenden Schritten erledigt.

1. Melden Sie sich an der AWS-Managementkonsole an.
2. Wählen Sie im Navigationsbereich „Directory Service“, und wählen Sie „Verzeichnis einrichten“.
3. Wählen Sie auf dem nächsten Bildschirm „AWS Managed Microsoft AD“, und klicken Sie auf „Weiter“.
4. Wählen Sie in dem Bereich „Verzeichnis-Informationen“
  - „Edition“, Standard Edition oder Enterprise Edition
  - Geben Sie als DNS-Name den vollständigen Verzeichnisnamen ein.
  - Geben Sie dann als NetBios-Name des Verzeichnisses einen Kurznamen für das Verzeichnis ein, z. B. CORP.

- Im Bereich „Verzeichnisbeschreibung“ kann eine Beschreibung für das Verzeichnis erfasst werden.
  - Das Administratorpasswort ist das Passwort des Verzeichnisadministrators „Admin“. Es wird automatisch erstellt.
5. Klicken Sie auf Weiter.
6. Erfassen Sie auf der Seite „VPC und Subnetze auswählen“ die Angaben über die VPC, die mit AD verwaltet werden soll.
- „VPC“: Die VPC für das Verzeichnis.
  - „Subnetze“: Wählen Sie die Subnetze für die Domain Controller aus. Die beiden Subnetze müssen sich in unterschiedlichen Availability Zones befinden.
7. Klicken Sie auf Weiter.
8. Überprüfen Sie auf der Seite „Überprüfen & Erstellen“ die Verzeichnisangaben. Wählen Sie „Verzeichnis erstellen“.

Die Bereitstellung kann bis zu 40 min dauern. Nach der Erstellung ändert sich der Status in „Aktiv“. Sie erhalten ein hochverfügbares AWS-verwaltetes AD-Paar mit einer eigenen Sicherheitsgruppe (SG). Die SG verfügt über vordefinierte Sicherheitsregeln für den ein- und ausgehenden Datenverkehr. Diese Sicherheitsregeln gewährleisten den sicheren Datenverkehr innerhalb der VPC, von anderen gepeerten VPCs und von Netzwerken, die über AWS Direct Connect, AWS Transit Gateway oder Virtual Private Network verbunden sind. Von der Änderung der Standardregeln wird dringend abgeraten, da sonst die Konnektivität zu AD verloren gehen kann.

#### 4.5.5.4.3 DNS-Namensauflösung mit Route 53 Resolver

Mit einer Hyperscaler Cloud können die Unternehmen ihre Ressourcen über mehrere Konten und VPCs verteilen. In der Regel können auch On-Premises-Server und/oder -Systeme einbezogen werden. Beim Einsatz von Servern in der Cloud entsteht ein hybrides Netzwerk mit mehreren Domänen und privat gehosteten Bereichen. Dies macht eine einheitliche DNS-Konfiguration erforderlich. Außerdem ist es wichtig, die Latenzzeiten gering zu halten. Daher liegt es nahe, die Namensauflösungsdienste in demselben lokalen Netz bereitzustellen wie die Ressourcen. On-Premises gibt es in der Regel einen DNS-Server, und in AWS wird der DNS-Service von Route 53 bereitgestellt.

Bei einer Hybridkonfiguration ist eine bidirektionale Abfrageweiterleitung erforderlich, damit die On-Premises-Abfrage nach AWS-Ressourcen in Route 53 aufgelöst werden kann und die Abfrage der AWS VPC nach vor Ort befindlichen Ressourcen das On-Premises-DNS erreichen kann. Gibt es mehrere Konten mit mehreren VPCs und mehreren Domänen und ist ein einheitlicher DNS-Dienst erforderlich, ist die Konfiguration noch komplexer. Für ein integriertes DNS kann Route 53 Resolver for Hybrid Clouds verwendet werden. Es bietet bidirektionale Abfragen zwischen On-Premises und AWS, indem es zur DNS-Namensauflösung private Verbindungen aufbaut.

Für die Unterstützung von DNS-Abfragen in Hybridumgebungen bietet Route 53 Resolver zwei Funktionen:

- Eingehende Abfragen: Route-53-Resolver-Endpunkte
- Ausgehende Abfragen: Bedingte Weiterleitungsregeln

## Konfiguration einer integrierten DNS-Ansicht für Amazon VPCs und lokale Netzwerke

Eine beispielhafte Konfiguration könnte aus den folgenden Elementen bestehen:

- On-Premises-DNS
- VPCs zu mehreren AWS-Konten, jeweils mit eigenen Route 53 Resolvern

## Vorgeschlagene Architektur

Zur Konfiguration gehen Sie wie folgt vor:

1. Verbinden Sie mit AWS Transit Gateway alle Spoke-VPCs mit der Hub-VPC.
  2. Verbinden Sie sich On Premises über VPN oder AWS Direct Connect mit der Hub-VPC.
  3. Konfigurieren Sie die ein- und ausgehenden Endpunkte von Route 53 Resolver.
    - Eingehende Endpunkte entsprechen den IP-Adressen, an die die DNS-Resolver in Ihrem Netzwerk die DNS-Abfragen weiterleiten sollen.
    - Um die DNS-Abfragen von Ihren VPCs an Ihr Netzwerk weiterzuleiten, erstellen Sie einen ausgehenden Endpunkt. Ein ausgehender Endpunkt gibt die IP-Adressen an, von denen die Abfragen stammen.
- Um bei den Endpunkten eine hohe Verfügbarkeit zu gewährleisten, empfiehlt sich die Angabe von IP-Adressen in mindestens zwei Subnetzen, die sich wiederum in zwei Availability Zones befinden.

Konfigurieren Sie die Route 53 Resolver mit privat gehosteten Bereichen in der Hub-VPC, und verknüpfen Sie diese. Jede VPC muss ihre privat gehosteten Bereiche allen anderen VPCs zuordnen. Die Zuordnung der privat gehosteten VPC-Bereiche zu den Spoke-VPCs muss also vollständig abgeschlossen werden.

## Erstellung von Weiterleitungsregeln

Eine Weiterleitungsregel erstellen Sie wie folgt:

- Durch bedingte Weiterleitungsregeln sind die On-Premises-Domänenamen der DNS-Abfragen festgelegt, die Sie an die On-Premises-DNS-Resolver weiterleiten möchten.
- Sie müssen für jeden Domänenamen eine Weiterleitungsregel erstellen und den Namen der Domäne angeben, für die Sie Abfragen weiterleiten möchten.

- Nach der Erstellung einer Weiterleitungsregel muss diese mit einer oder mehreren VPCs verknüpft werden. Nachdem eine VPC mit einer Regel verknüpft ist, leitet Resolver die DNS-Abfragen für den in der Regel angegebenen Domänennamen an die in der Regel angegebenen DNS-Resolver weiter. Die Abfragen laufen über den bei der Regelerstellung angegebenen ausgehenden Endpunkt.

#### 4.5.5.4.4 Verarbeitung von Anfragen

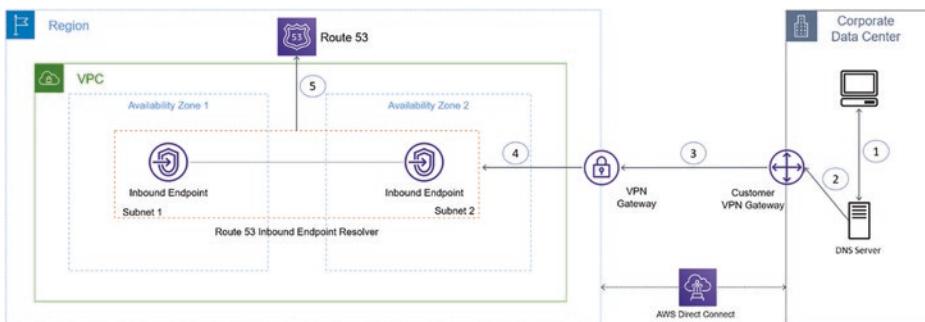
##### Abfrage von On-Premises an AWS Route 53

In diesem Szenario wird On-Premises eine DNS-Namensauflösungsabfrage für eine Ressource in der AWS-Cloud gestellt (Abb. 4.18).

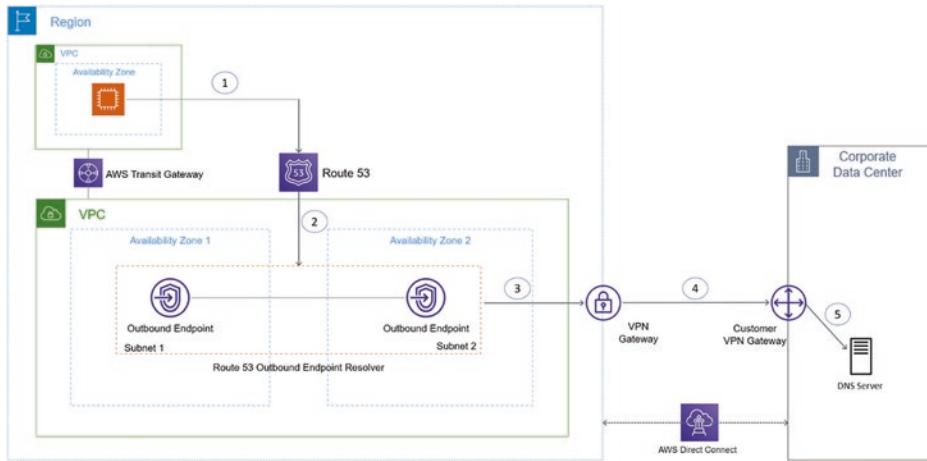
- Ein Client im On-Premises-Netzwerk initiiert eine DNS-Anfrage, und diese geht an den On-Premises-DNS-Server. Die Konfiguration auf dem On-Premises-DNS-Server erkennt, dass die Anfrage für eine AWS-Domäne bestimmt ist. Der On-Premises-DNS-Server leitet die Abfrage an Route 53 Inbound Endpoint Resolver weiter.
  - Die Anfrage wird an das Customer-VPN-Gateway weitergeleitet.
  - Weiter geht es zum AWS-VPN-Gateway und zu Route 53 Inbound Endpoint Resolver.
  - Zur Namensauflösung geht die Abfrage vom Inbound Endpoint Resolver an Route 53 in AWS. Route 53 löst den Namen auf, und die Antwort wird an den On-Premises-Client zurückgesendet.
- Die Verbindung zwischen der On-Premises-Umgebung und der AWS-Cloud kann über VPN und/oder AWS Direct Connect erfolgen.

##### Abfrage von AWS-Ressourcen für On-Premises-DNS

In diesem Szenario geht eine Abfrage von einer in der Spoke-VPC befindlichen EC2-Instanz aus. Die abgefragte Domäne stammt aus dem On-Premises-DNS. Im Folgenden ist der Datenverkehr durch das Netzwerk dargestellt (Abb. 4.19).



**Abb. 4.18** Query von On-Premise zu AWS Route 53



**Abb. 4.19** Abfrage von AWS Ressourcen für On-Premise DNS

1. Eine EC2-Instanz in einer Spoke-VPC initiiert einen DNS-Lookup. Diese Spoke-VPC ist mit der Hub-VPC verbunden. Diese wiederum ist über das Transit Gateway mit Endpunkten ausgestattet. Route 53 empfängt die Abfrage und stellt anhand der Auflösungsregeln fest, dass sie für eine On-Premises-Domain bestimmt ist. Wenn die Auflösungsregel greift, wird die Abfrage an das On-Premises-DNS weitergeleitet.
2. Die Abfrage wird dort an Outbound Endpoint weitergeleitet.
3. Outbound Endpoint sendet die Abfrage an das AWS-VPN-Gateway.
4. Das AWS-VPN-Gateway leitet die Abfrage an das Customer-VPN-Gateway weiter.
5. Das Customer-VPN-Gateway sendet die Abfrage an den On-Premises-DNS-Server.

Dieser löst den Namen auf und sendet die Antwort an die EC2-Instanz zurück, die die Abfrage gestellt hat. Bei der Konfiguration lassen sich wie folgt Verbesserungen erzielen:

- Das DNS sollte sich in der Hub-VPC befinden, damit alle verbundenen Spoke-VPCs nahtlos mit dieser zentralen Stelle verbunden werden können. Der Datenverkehr kann mithilfe einer bedingten Weiterleitung verwaltet werden.
- Stellen Sie sicher, dass alle Active-Directory-DNS-Domänen von allen Clients aufgelöst werden können, denn die Clients verwenden die Domänen, um Active-Directory-Dienste zu finden und ihre DNS-Namen mithilfe dynamischer Updates zu registrieren.
- Die Amazon-EC2-Instanzen sollten für die DNS-Auflösung .2 verwendet. Damit erhält jede elastische Netzwerkschnittstelle maximal 1.024 Datenpakete pro Sekunde.
- Stellen Sie im DHCP die Domain Name Server so ein, dass sie auf AmazonProvidedDNS und nicht auf die Resolver-Endpunkte verweisen. Um sicherzustellen, dass AmazonProvidedDNS über die von Ihnen benötigte DNS-Ansicht verfügt, verwenden Sie Weiterleitungsregeln.

- Für geringere Latenzen halten Sie die DNS-Namensauflösung lokal in der AWS-Region.
- Verwenden Sie den Amazon-DNS-Server als Weiterleitungsdienst für alle sonstigen DNS-Domänen, die für Ihre DNS-Server auf AD-Domain-Controllern nicht maßgeblich sind. Mit dieser Konfiguration können Ihre DC Datensätze im privaten Bereich von Amazon Route 53 rekursiv auflösen und die bedingten Weiterleitungen von Route 53 Resolver verwenden.

AWS-Referenzdokument (Zugriff am 20.12.2021): <https://docs.aws.amazon.com/Route53/latest/DeveloperGuide/resolver.html>

#### 4.5.5.5 Überwachung von AWS Managed AD DS

Da AD DS zu den für das Unternehmen wichtigsten Diensten zählt, müssen die Ressourcen überwacht werden. Zur Überwachung von AWS Managed AD sind in Amazon mehrere Optionen verfügbar.

- Aktive Überwachung der Verzeichnis-Sicherheitsprotokolle mit den in Amazon CloudWatch enthaltenen Protokollen
  - Benachrichtigung über Verzeichnisänderungen per Einbindung des Simple Notification Service (SNS)
  - Verbindung zu AWS CloudTrail, um die Protokolle für Audits zu speichern
- AWS lässt die Installation von Überwachungsagenten auf AWS Managed Microsoft AD Domain Controllern nicht zu.

#### 4.5.6 Dynamic Host Configuration Protocol (DHCP)

Wenn in der AWS-Cloud eine Anwendung gehostet werden soll und dafür eine Infrastruktur eingerichtet wird, umfasst diese mehrere Komponenten. Diese erhalten bei der Einbindung in das Netzwerk mittels jeweils einer IP-Adresse ihre eindeutige Identität. Neben der IP-Adresse müssen den Komponenten auch andere Kommunikationsparameter für die TCP/IP-Kommunikation zugewiesen werden, darunter die Subnetzmaske, das Standard-Gateway, die DNS-Adresse, der Nameserver und der Zeitserver.

Im Zuge der Netzwerkerweiterungen und der immer zahlreicherem Netzwerkkomponenten in der Umgebung werden manuelle Zuweisung (oder Rücknahme), Verwaltung und Verfolgung zu einem mühsamen Prozess. Hier kommt der Dynamic Host Configuration Protocol Server ins Spiel (DHCP), ein Netzwerkverwaltungs-Protokollserver, der die automatische Zuweisung von IPs und sonstigen Netzwerkparametern zentral unterstützt.

DHCP basiert auf dem Server-Client-Modell. Die Umgebung umfasst einen zentralen DHCP-Server und DHCP-fähige Clients, die die Parameter vom DHCP-Server

empfangen können. Normalerweise ist bei Windows-basierten Client-Betriebssystemen der DHCP-Client integriert. Bei einem Linux-Betriebssystem können Sie entweder mit dhclient einen DHCP-Client einrichten oder mit der Datei configue/etc./network/interfaces DHCP nutzen.

Wenn ein Gerät wie z. B. eine virtuelle Maschine eine Verbindung zu einem Netzwerk aufbaut, was in der Regel nur während des Bootvorgangs der Fall ist, sendet die mitgelieferte DHCP-Client-Software eine DHCP-Broadcast-Anfrage nach den benötigten Netzwerkdaten. Der DHCP-Server des Netzwerks gibt die angeforderten Daten zurück. Welche Daten angegeben werden müssen, kann von einem Administrator konfiguriert werden.

Der DHCP-Server verwaltet einen Datensatz mit allen den Netzwerknoten zugewiesenen IP-Adressen. Diese einzelnen Netzwerkkomponenten werden über die eindeutige MAC (Media Access Control) erkannt. Doppelte IP-Zuweisungen werden dadurch verhindert.

Bei der Erstellung der VPC legt AWS automatisch einen Datensatz mit den DHCP-TCP/IP-Standardkommunikationsparametern an und ordnet der VPC die betreffenden Parameter zu. Jeder in der VPC erstellte AWS-Service erbt diese Parameter. Die Standard-DHCP-Optionen umfassen auch die beiden Optionen Domänenname und Domänennamenserver.

```
domain.name = <domain_name_ihrer_region> Beispiel: sa-east-1.compute.internal
domain-Namen-Server = AmazonProvidedDNS
```

Auch das mit der VPC verbundene Standard-AmazonProvidedDNS ist bekannt, und auf einer der reservierten IPs der VPC läuft Route 53 Resolver. Bei IPV4 entspricht diese IP dem um 2 erhöhten Netzwerkbereich der VPC. Wurde der VPC beispielsweise der CIDR-Bereich 172.16.0.0/16 zugewiesen, lautet die IP des DNS-Servers der VPC 172.16.0.2. Bei VPCs mit mehreren IPv4-CIDR-Blöcken befindet sich die IP-Adresse des DNS-Servers im primären CIDR-Block. Der DNS-Server befindet sich nicht innerhalb eines bestimmten Subnetzes oder einer Availability Zone in einer VPC.

Standardmäßig erhalten allen in der VPC erstellten Diensten einen nicht auflösbarren Host, wie in der Tab. 4.3 aufgezeigt.

**Tab. 4.3** Öffentlicher und privater Hostname

Region	Privater Hostname	Öffentlicher Hostname
US East (N. Virginia) us-ost-1	IP-<Private_IP>.ec2.internal	ec2-<Private_IP>.compute-1.amazonaw.com
Sonstige Regionen	IP-<Private_IP>.compute.internal	ec2-<Private_IP>.compute.amazonaw.com

Für Ihre VPC müssen Sie Ihre eigenen DHCP-Optionen konfigurieren. Bei der Einrichtung der benutzerdefinierten DHCP-Optionen sind die folgenden Überlegungen anzustellen:

- Die DHCP-Optionen können nachträglich nicht mehr geändert werden.
- Einer VPC kann jeweils nur ein DHCP-Optionssatz zugeordnet werden. Es können aber auch mehrere solche Sätze konfiguriert werden.
- Wenn bei einer VPC die DHCP-Optionen geändert werden müssen, kann also ein neuer DHCP-Optionssatz erstellt und der VPC zugeordnet werden.
- Wird einer aktuellen VPC ein neuer Satz von DHCP-Optionen zugeordnet, verwenden alle vorhandenen Instanzen und zukünftigen Komponenten automatisch die neuen Optionen. Ein Neustart ist nicht erforderlich. Allerdings kann die Aktualisierung einige Zeit in Anspruch nehmen. Die Aktualisierung kann auch manuell erzwungen werden.

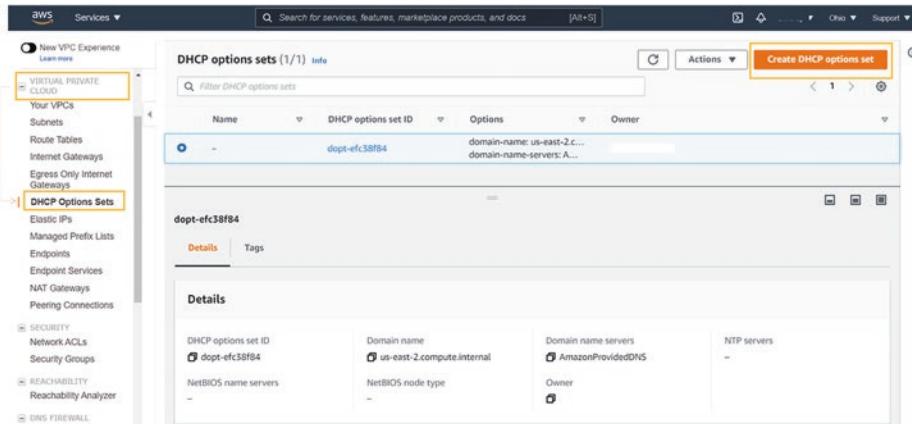
Nachstehend finden Sie die DHCP-Optionen, die an individuelle Werte angepasst werden können:

- domain-name-servers – Name des DNS-Servers, der für die Namensauflösung verwendet werden soll
- domain-name – Benutzerdefinierter Domänenname der Instanzen
- ntp-servers – Name der Network-Time Protocol-Server (NTP), um die Uhrzeit in allen Instanzen synchron zu halten
- netbios-name-servers – Name der NetBIOS-Nameserver
- netbios-node-type – NetBIOS-Knotentyp (1, 2, 4 oder 8). Es wird empfohlen NetBIOS-Knotentyp 2 anzugeben (point-to-point oder P-node). Derzeit werden Broadcast und Multicast in AWS nicht unterstützt.

## DHCP-Optionen

Die Arbeit mit DHCP-Optionen ist eine der ersten Aufgaben, die nach der Einrichtung der ersten VPC anfallen. Wie bereits erwähnt, weist AWS der VPC bei der Bereitstellung ein Standard-DHCP mit eigenen Optionen zu. Dieses muss in der Folge an die Unternehmensanforderungen angepasst werden. Auf einen neuen Satz von Option umzustellen, kann auch im Dauerbetrieb erforderlich sein. Die Aufgaben in Zusammenhang mit den Optionen kann wie folgt kategorisiert werden:

- DHCP-Optionen erstellen
- DHCP-Optionen löschen
- Den der VPC zugeordneten DHCP-Optionen ändern
- Den der VPC zugeordneten DHCP-Optionen löschen (Abb. 4.20 und 4.21)



**Abb. 4.20** DHCP Optionen

Hinweis: Wenn in der VPC ein Internet-Gateway verwendet werden soll, muss als „Domain Name Server“ entweder der Amazon-DNS-Server (AmazonProvidedDNS) oder der unternehmenseigene DNS-Server angegeben werden. Andernfalls kann die Instanz nicht auf das Internet zugreifen (Abb. 4.22).

## 4.6 Zusammenfassung

In den vorangegangenen Kapiteln haben wir die Konzepte und Architekturen von AWS plus die Vorteile bei der Nutzung von AWS für den Betrieb von SAP-Umgebungen beleuchtet. Hierzu wurden zunächst die wichtigsten Konzepte von AWS zur generellen Architektur und die von AWS notwendigen Dienste erläutert. Dazu gehören elementare Punkte, wie Regionen, VPC, Verfügbarkeitszonen und Placement Groups. Über die AWS Managementkonsole können Sie die Dienste alle miteinander verbinden und somit die Grundlage für hochverfügbare und resiliente SAP-Systeme schaffen.

Eine abgeschottete SAP-Umgebung ist sehr selten und daher sind wir auch auf die Integration von einer neuen AWS-Umgebung mit den wichtigsten Kerndiensten, wie DHCP, Active Directory oder auch DNS eingegangen.

AWS bietet Ihnen ein großes Sammelsurium von Diensten für eine SAP-Umgebung zur Verfügung. Nicht alle von diesen Diensten müssen genutzt werden. Die Implementierung der wichtigsten Dienste, wie Compute, Storage und Backup&Restore werden im folgenden Kapitel anhand einer Beispielarchitektur gezeigt. Hier wird eine Umgebung neu implementiert.

VPC > DHCP options sets > Create DHCP options set

## Create DHCP options set Info

Dynamic Host Configuration Protocol (DHCP) provides a standard for passing configuration information to hosts on a TCP/IP network. The options field of a DHCP message contains configuration parameters.

**Tag settings**

DHCP options set name - *optional*

**DHCP options**  
Specify at least one configuration parameter.

Domain name Info

Domain name servers Info  
  
Enter up to four IP addresses, separated by commas.

NTP servers  
  
Enter up to four IP addresses, separated by commas.

NetBIOS name servers  
  
Enter up to four IP addresses, separated by commas.

NetBIOS node type  
  
We recommend that you select point-to-point (2 - P-node). Broadcast and multicast are not currently supported.

▶ AWS Command Line Interface command

**Tags**  
A tag is a label that you assign to an AWS resource. Each tag consists of a key and an optional value. You can use tags to search and filter your resources or track your AWS costs.

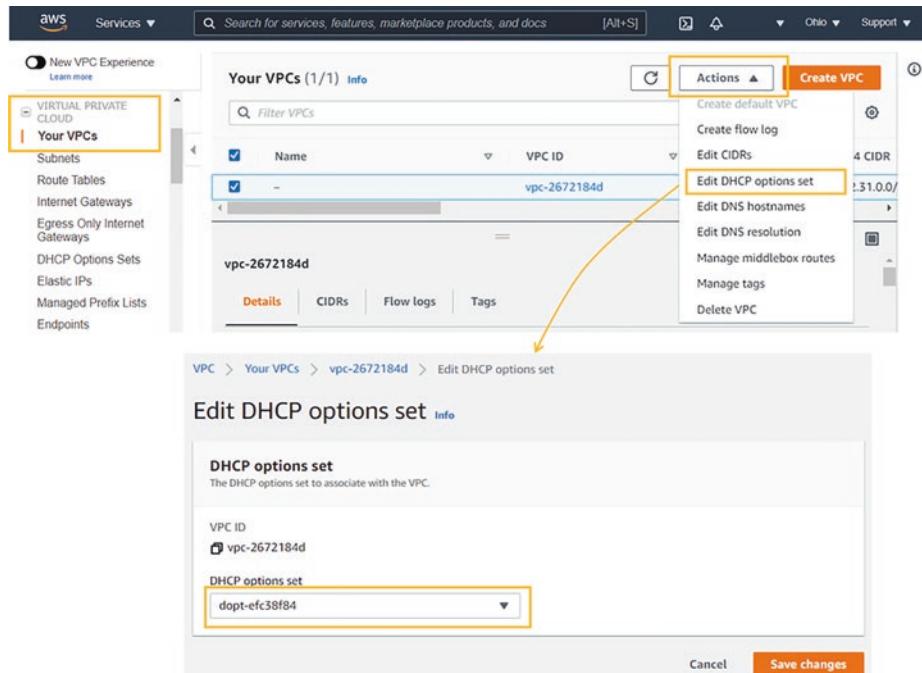
No tags associated with the resource.

Add new tag

You can add 50 more tags.

Cancel Create DHCP options set

**Abb. 4.21** DHCP-Optionen erstellen



**Abb. 4.22** DHCP-Optionssatz bearbeiten



# SAP S/4 on Amazon AWS – Deployment

5

## Zusammenfassung

Dieses Kapitel zeigt Ihnen die konkrete Implementierung eines SAP S/4HANA-Systems in der AWS Cloud. Hierzu wird zunächst eine beispielhafte Architektur erläutert, welche als spätere Architektur für das zu implementierende System genutzt werden wird. Für die Implementierung werden Ihnen alle notwendigen Komponenten vorgestellt. Dazu erläutern wir Ihnen die Compute-Komponenten, als auch alle wichtigen Komponenten zum Storage. Um eine möglichst kostengünstige Variante des SAP S/4HANA-Systems zu implementieren, zeigt Ihnen das Kapitel auch die Möglichkeit, sich vorher die Transparenz zu Pay-As-You-Go und Commitments und den daraus resultierenden Kosten zu schaffen. Die Sicherheit des SAP S/4HANA-Systems wird hier ebenfalls diskutiert und wir zeigen Ihnen auch, wie Sie die Daten des SAP S/4HANA-Systems in AWS geeignet sichern können. Zuletzt gehen wir auch auf das Thema vom Disaster Recovery, also dem kompletten Ausfall einer AWS Region, ein und zeigen die Mechanismen, welche zum Schutz implementiert werden können.

## 5.1 Architekturbeispiel

SAP S/4HANA kann mit dem Infrastructure-as-a-Service-Modell von AWS hochverfügbar, fehlertolerant und kostengünstig bereitgestellt werden. In diesem Kapitel erläutern wir, wie unterschiedlich aufgebaute SAP-S/4HANA-Systeme mit spezifischen Verfügbarkeits-SLAs in AWS bereitgestellt werden können.

### 5.1.1 Bereitstellungsszenarien

SAP S/4HANA-Systeme können in verschiedenen Konfiguration und mit verschiedenen Verfügbarkeiten bereitgestellt werden. Wir werden in diesem Kapitel die folgenden Szenarien sukzessive beschreiben:

- Eine Verfügbarkeitszone und ein Knoten
- Eine Verfügbarkeitszone und mehrere Knoten
- Mehrere Verfügbarkeitszonen und mehrere Knoten mit Hochverfügbarkeit

Dies erlaubt Ihnen, sich ein detailliertes Bild für die üblichen SAP-Architekturen in AWS zu machen.

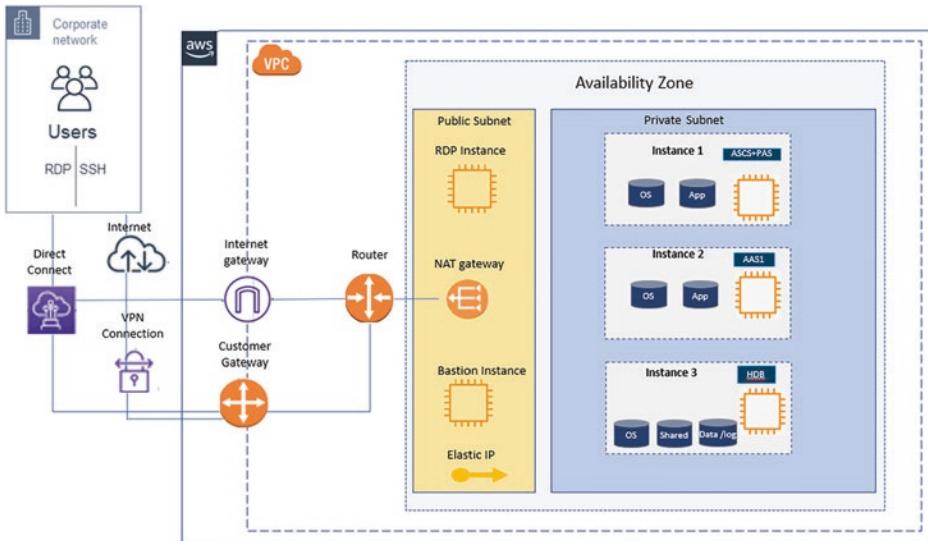
#### 5.1.1.1 Eine Availability Zone und Single-Node-Architektur

Dies ist die standardmäßige und kostenminimierende Architektur bei SAP-Systembereitstellungen in AWS. Das SAP-System wird dabei in einer einzigen Availability Zone mit einem SAP-HANA-Datenbankserver und einem oder mehreren SAP-S4-Anwendungsservern bereitgestellt. Im VM-Verfügbarkeits-SLA sind 99,99 % vereinbart, was bedeutet, dass die EC2-Instanz für maximal 4,32 min pro Monat unverfügbar sein darf. Hinweis: Es handelt sich dabei um ein VM-Verfügbarkeits-SLA und nicht um das Anwendungsverfügbarkeits-SLA von SAP S4/HANA. Eine solche Architektur wird auch als „Self-Healing-Architektur“ bezeichnet, weil von AWS bei Problemen mit der zugrunde liegenden EC2-Instanz-Hardware automatisch ein Neustart der EC2-VM durchgeführt wird, durch den für dieselbe Einrichtung und Konfiguration eine andere, gesunde Hardware zugrunde gelegt wird. Sie müssen sich also nur um den Anwendungsteil kümmern.

Für die Anwendungsdaten übernimmt AWS keine Haftung. Sollten bei dem Neustart Daten beschädigt werden, gibt es nur die Möglichkeit, die Dienste per Datenwiederherstellung in den Normalzustand zurückzubringen. Wie erwähnt handelt es sich bei dieser Architektur um die kostenminimierende Standardarchitektur. Sie eignet sich am besten für nicht produktive Workloads und kann auch für nicht kritische produktive Workloads verwendet werden (Abb. 5.1).

Dies ist die einfache Standardarchitektur für SAP-Systeme, welche keine Hochverfügbarkeit benötigen. Die SAP-Anwendung und die HANA-DB werden in derselben VPC, derselben Availability Zone und demselben Subnetz bereitgestellt. Die Netzsegmentierung hängt vollständig von den Kundenanforderungen ab. Hier können verschiedene Szenarien umgesetzt werden:

- **Mikrosegmentierung** – Jede SAP SID in einem anderen Subnetz.
- **Segmentierung nach produktiv/nicht produktiv** – Die produktiven Anwendungen und Datenbanken werden jeweils in einem anderen Subnetz, die nicht produktiven Anwendungen und Datenbanken ebenfalls jeweils in einem anderen Subnetz bereitgestellt.



**Abb. 5.1** SAP S/4HANA-System mit einem Knoten

- ▶ Gemäß den Best Practice sollten sich SAP und DB in verschiedenen Subnetzen befinden, um die direkten Datenbankzugriffe zu beschränken.

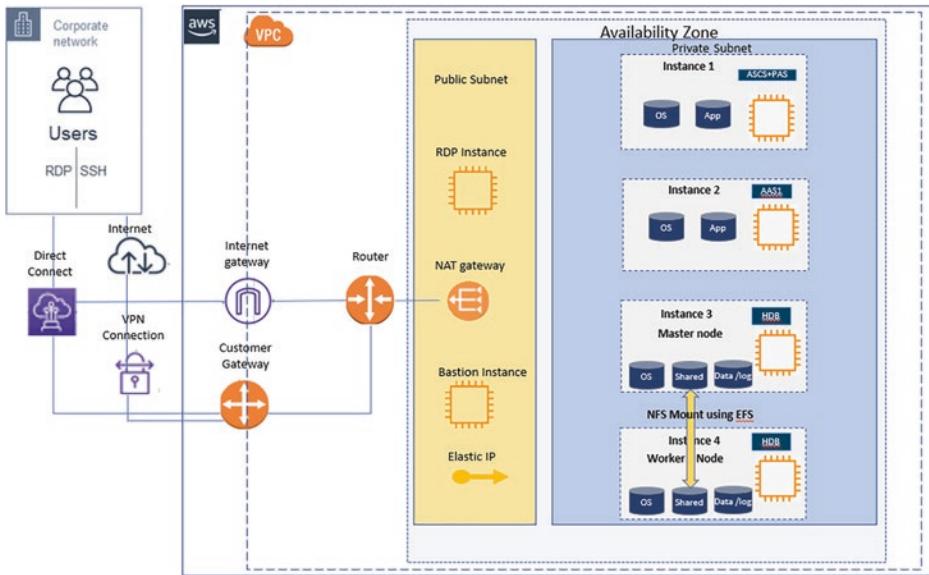
Die SAP-Anwendungen (PAS und AAS) und die HANA-Datenbank werden gemäß der SAP-Standardinstallationsanleitung installiert. Die Anforderungen an VM und Betriebssystem bei der SAP-Bereitstellung werden im nächsten Kapitel ausführlich erläutert.

### 5.1.1.2 Eine Availability Zone und Multi-Node-Architektur

Es handelt sich um eine SAP-S/4-Bereitstellung in einer HANA-Scale-out-Umgebung. Pro Availability Zone mit Multi-Node-Option können in AWS standardmäßig bis zu fünf Knoten bereitgestellt werden. Zur Bereitstellung einer HANA-Scale-out-Umgebung mit mehr als fünf Knoten müssen Sie sich an AWS wenden. Um die Compliance zur besten Sicherheitspraxis zu gewährleisten, müssen alle HANA-Knoten unabhängig von ihrer Funktion in demselben Subnetz bereitgestellt werden (Abb. 5.2).

Bei der Installation von SAP S/4 und der HANA-Datenbank muss der SAP-Standardleitfaden beachtet werden. Die Anforderungen an VM und Betriebssystem bei der SAP-Bereitstellung werden im nächsten Kapitel ausführlich erläutert.

- ▶ **Hinweis** Die SAP-HANA-Scale-out-Hosts benötigen für die Kommunikation zwischen den Knoten verfügen ein separates Netzwerk (eine zusätzliche NIC). Der Parameter „listeninterface“ im Abschnitt [communication] muss auf „internal“ gesetzt sein oder eine CIDR-Netzmaske enthalten. Bitte beachten Sie **SAP KBA 2183363**.



**Abb. 5.2** SAP S/4HANA-System mit mehreren Knoten

### 5.1.1.3 Mehrere AZs, Single-Node-SAP-S/4 mit Hochverfügbarkeitsarchitektur

Die Bereitstellung von AWS in mehreren Availability Zones bietet die Möglichkeit, geschäftskritische SAP-Workloads mit hoher Resilienz und starker Fehlertoleranz in einer hochverfügbaren Umgebung auszuführen. Bei Anwendungsausfällen, unerwarteten Hardwarefehlern oder Nichtverfügbarkeit einer der Availability Zones sind nur mit sehr geringer Wahrscheinlichkeit Auswirkungen auf das Geschäft zu erwarten.

Bei der SAP-HA-Einrichtung muss zunächst ein privates Subnet innerhalb einer VPC erstellt werden, um das Netzwerk von den anderen virtuellen Netzwerken in AWS zu isolieren. Dann wird der Datenverkehr zu den in der VPC bereitgestellten Instanzen umgeleitet. Der SAP-S/4-ASCS/ERS-Service und die HANA-Datenbank müssen in beiden Availability Zones mit einem bestimmten privaten Subnetzbereich (CIDR-Bereich) bereitgestellt werden. Subnetze können in AWS nicht zonenübergreifend überbrückt werden. Um den Datenverkehr zu aktiven ASCS- oder aktiven HANA-DB-Knoten umzuleiten, muss eine Overlay IP (OIP) verwendet werden.

Es gibt zwei Möglichkeiten für die Einrichtung einer SAP-HA-Umgebung in AWS mit einer OIP, welche wie folgt aussehen (Abb. 5.3).

#### Szenario 1: OIP mit AWS-Transit-Gateway

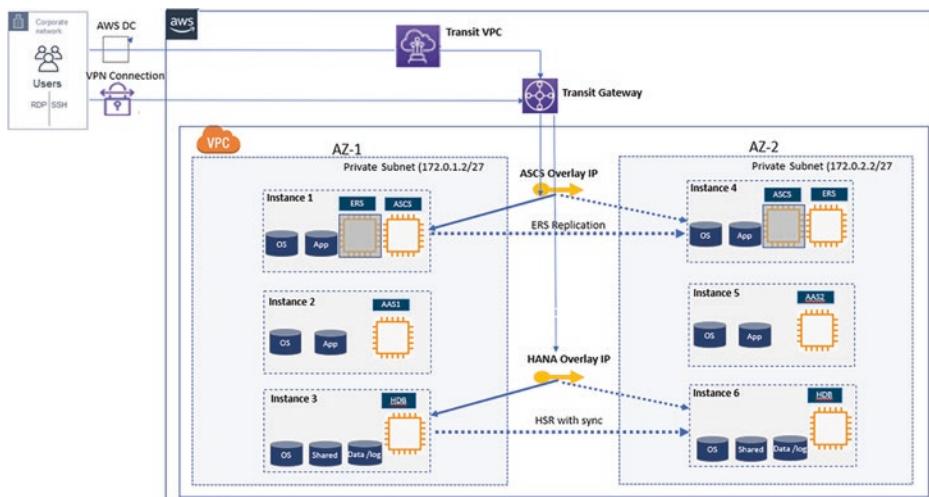
Die obige Architektur entspricht der Hochverfügbarkeitskonfiguration mit AWS Transit Gateway. Mit dem Transit Gateway können die AWS-Routingtabellenregeln verwendet werden. Diese ermöglichen es der OIP-Adresse, ohne zusätzliche Komponenten wie

dem Network Load Balancer mit der SAP-Instanz zu kommunizieren. Die Verbindung zur OIP ist auf einfache Weise von einer anderen VPC, einem anderen Subnetz oder per Direct Connect vom Unternehmensnetzwerk aus möglich.

AWS Transit Gateway fungiert als Hub, der den speichenförmigen Datenverkehr zwischen allen angeschlossenen Netzwerken steuert. Transit Gateway leitet die Datenpakete gemäß den Transit-Gateway-Routingtabellen von der Quelle zum Ziel. Die Routingtabellen können so konfiguriert werden, dass sie die Routen aus den Routingtabellen für die angeschlossenen VPCs und VPN-Verbindungen abrufen. Auch statische Routen können zu den Transit-Gateway-Routingtabellen hinzugefügt werden. Die Overlay-IP-Adresse oder der CIDR-Adressbereich können als eine statische Route in die Routingtabelle des Transit-Gateways hinzugefügt werden. Das Ziel ist dabei die VPC, in der die EC2-Instanzen des SAP-Clusters laufen. Auf diese Weise wird der gesamte Datenverkehr an die Overlay-IP-Adressen weitergeleitet.

In Abb. 3.6.1.3 wird für die Isolierung des Netzwerks von den anderen virtuellen AWS-Netzwerken eine VPC mit den beiden Availability Zones AZ1 und AZ2 verwenden. Als nächstes muss ein privates Subnetz mit einem bestimmten CIDR-Bereich in AZ1 (172.0.1.2/16) und AZ2 (172.0.2.2/16) erstellt werden. Das private Subnetz kann sich nicht über beide Availability Zones erstrecken. Nun wird S/4 ASCS in AZ1 installiert (in der Abbildung: Instanz 1). Der Standalone-Enqueue-Server2 (ENSA2) wird in AZ2 installiert (In der Abbildung: Instanz 4). ENSA2 ist jetzt eine Standardinstallation von S/4 1809 und die Idee ist, mit dieser Funktionalität die Lock-Tabelle reibungslos und effizient zu verwalten, was wiederum in ABAP-Systemen für Datenkonsistenz sorgt.

Der ASC- und der ERS- bzw. ENSA2-Service sind nur mit einem virtuellen Host zu installieren, dem eine als Overlay-IP bekannte virtuelle IP zugeordnet ist. Die



**Abb. 5.3** SAP S/4HANA-System mit mehreren Verfügbarkeitszonen und Transit Gateway

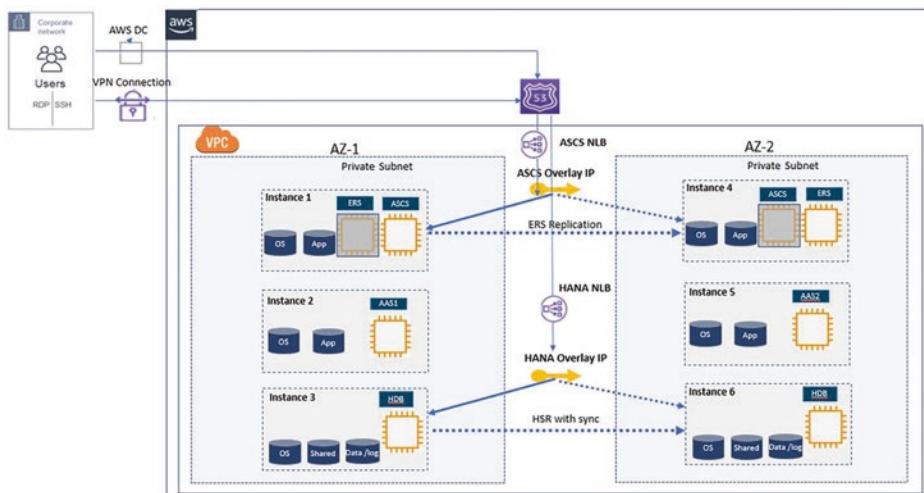
S/4-Anwendungsinstanz ist in beiden AZs (Instanz 2, Instanz 3) zu installieren, um bei einer Störung in einer Zone den Ausfall zu vermeiden. Nun muss die HANA-DB in beiden AZs mit dem virtuellen Host installiert werden, dem die Overlay-IP zugeordnet ist (Instanz 3, Instanz 4). Zwischen den beiden HANA-Datenbanken ist die HANA-Systemreplikation mit dem synchronen Replikationsmodus zu konfigurieren.

Die Overlay-IP zeigt immer auf den aktiven ASCS- oder HANA-DB-Knoten. Bei einem Knotenausfall oder bei einer Serviceunterbrechung wechselt ASCS oder die HANA zu Availability Zone 2, und die Overlay-IP wechselt zu den aktiven Knoten von Availability Zone 2. Die Endbenutzer werden durch den Ausfall eines Knotens nicht beeinträchtigt, da sie stets eine Verbindung zur Overlay-IP herstellen und die Overlay-IP immer auf die aktiven Knoten verweist.

Bei der Konfiguration des AWS Transit Gateways sollte der AWS-Standardleitfaden zugrunde gelegt werden. Auf die SAP-HA-Cluster-Einrichtung in anderen Betriebssystemen wird weiter unten in diesem Kapitel eingegangen (Zugriff am 20.12.2021): <https://docs.aws.amazon.com/vpc/latest/tgw/tgw-getting-started.html>

### Szenario 2: Overlay-IP-Routing mit Lastverteilung im Netzwerk

Falls AWS Transit Gateway nicht verwendet werden kann, ist es möglich, mit Netzwerklastverteilern (Network Load Balancer – NLB) vom externen Netzwerk aus auf die Overlay-IP-Adressen zuzugreifen. NLB arbeiten in der vierten Schicht des OSI-Modells und sind in der Lage, Millionen von Anfragen pro Sekunde anzunehmen. Die Anfrage geht vom externen Netzwerk aus ein. Das Ziel wird gemäß Netzwerklastverteiler-Zielgruppe gewählt. Die Anfrage wird dann zur angegebenen Overlay-IP-Adresse geroutet. Die SAP-S/4- und -HANA-DB-Installation entspricht dem im vorherigen Abschnitt erläuterten Konzept.



**Abb. 5.4** SAP S/4HANA-System mit mehreren Verfügbarkeitszonen und NLB Architektur

Bei der Konfiguration des Netzwerklastverteilers sollte der AWS-Standartleitfaden zugrunde gelegt werden (Zugriff am 20.12.2021): <https://docs.aws.amazon.com/elasticloadbalancing/latest/network/create-network-load-balancer.html> (Abb. 5.4).

## 5.1.2 Einrichtung eines SAP-HA-Clusters

In diesem Kapitel beschreiben wir, wie ein Cluster für die Hochverfügbarkeit von SAP S/4HANA-Systemen auf Windows und Linux erstellt werden kann. Darüber hinaus werden auch die Testszenarien für solch einen Cluster beschrieben.

### 5.1.2.1 SAP-NetWeaver-/SAP-HANA-Hochverfügbarkeitseinrichtung auf RHEL und SUSE

In diesem Abschnitt erläutern wir die Einrichtung von Clusterlösungen in den Betriebssystemen RHEL und SUSE. Das Ziel dabei ist eine hochgradig verfügbare SAP-Konfiguration. Die Einrichtung basiert auf dem Standalone Enqueue Server 2, der ab SAP S/4HANA 1809 der Standardinstallation entspricht. Für die Einrichtung von SAP S/4 HA und SAP HANA HA unter SUSE oder RHEL in der AWS Public Cloud ist ein Standardleitfaden verfügbar. Am Ende des vorliegenden Kapitels werden die bei der Clustereinrichtung zu beachtenden Knowledge Base Artikel (KBA) angegeben. Es ist wichtig zu verstehen, wie Cluster funktionieren.

Die Clusterlösung für RHEL oder SUSE basiert auf Pacemaker, der automatischen Ausfallsicherung bei Ausfällen zwischen zwei Knoten. Beide Server entsprechen Paketen, die das RHEL-/SUSE-Pacemaker-Cluster beinhalten. Pacemaker funktioniert für S/4-ASCS- und für HANA-Cluster sowohl in der RHEL- als auch in der SUSE-Variante identisch. Der Ausfall des aktiven ASCS-Knotens wird automatisch vom Pacemaker-Cluster erkannt. Die ASCS-Dienste (Single Point of Failure) wechseln automatisch zu dem zweiten Knoten, auf dem der Enqueue Replication Server (ERS) läuft. Darüber hinaus wird die virtuelle IP (Overlay-IP) von dem primären Knoten auf den sekundären Knoten umgestellt, womit sichergestellt werden soll, dass alle Verbindungen immer nur zu dem aktiven Knoten gehen und es nicht zur Serviceunterbrechung kommt.

**Bedeutung von Fencing:** Im Normalbetrieb umfasst die Clusterlösung keinen Fencing-Agent. Die Cluster der beiden Knoten kommunizieren über Corosync und tauschen kontinuierlich Informationen über den Knotenzustand aus. Problematisch wird es, wenn einer der Knoten nicht mehr reagiert. Es können sich dann zwei Szenarien ergeben.

**Szenario 1:** Der primäre Knoten ist aufgrund eines Absturzes oder eines Neustarts nicht verfügbar.

**Szenario 2:** Der primäre Server ist verfügbar, aber aufgrund eines Netzwerkproblems nicht erreichbar.

Im ersten Fall wird das Cluster auf den funktionsfähigen sekundären Knoten wechseln, da der primäre Knoten nicht mehr verfügbar ist. Aber im zweiten Fall darf kein

Wechsel erfolgen, da beide Knoten aktiv oder verfügbar sind. Diese Situation wird als SPLIT BRAIN oder Dual Primary bezeichnet und ist im HANA-Cluster sehr gefährlich. Die HANA-Datenbank läuft auf beiden Knoten unabhängig, und wenn das Cluster den Wechsel zwischen den Knoten vollziehen kann, können beide Knoten leicht als primärer Knoten fungieren. Stellen wir uns vor, eine Transaktion wird in Knoten 1 aufgezeichnet, und später wird eine Transaktion in Knoten 2 aufgezeichnet. Nun verursacht Knoten 2 den Verlust von Geschäftsdaten. Das Pacemaker-Cluster führt nun einen Fencing-Mechanismus namens STONITH ein (Shoot-The-Other-Node-In-The-Head). Bei Kommunikationsproblemen zwischen den Knoten wird der gesunde Knoten den anderen fehlerhaften Knoten killen, damit nicht beide Knoten gleichzeitig aktiv sind und eine sogenannte Split-Brain-Situation entsteht.

Nachfolgend finden Sie die wichtigsten Schritte zur Installation von SAP S/4 oder SAP HANA in einer HA-Umgebung basierend auf einer RHEL- oder SUSE-Plattform (Tab. 5.1).

1. **Zugrunde liegende AWS-Clouds erstellen** – Design und die Implementierung der zugrunde liegenden AWS-Clouds sind die ersten Schritte vor dem Beginn eines jeden AWS-Cloud-Projekts. Hierunter fallen die Festlegung der Region, des Netzwerkdesigns (VPC, Subnetze, Firewalls, Transit Gateway oder Netzwerklastverteiler, Route 53), der Sicherheits- und Ressourcenrichtlinien, der Konnektivität zwischen On-Premises und AWS, der AD-Authentifizierung etc. Manche dieser Punkte wurden bereits im vorherigen Abschnitt erläutert, auf andere wird noch eingegangen, z. B. weiter hinten in diesem Kapitel auf das Thema Sicherheit.
2. **EC2-Instanz bereitstellen** – Bereitstellung der SAP-NetWeaver-zertifizierten und der SAP-HANA-zertifizierten EC2-Instanz. Diese Bereitstellung kann manuell oder automatisiert mit den von SAP empfohlenen Betriebssystemeinstellungen für die jeweilige Betriebssystemversion (RHEL oder SLES) durchgeführt werden. Die VM-Bereitstellung wird weiter unten im vorliegenden Kapitel im Abschnitt „Compute“ ausführlich erläutert.
3. **Aufbau des Dateisystems festlegen** – Um SAP und HANA hochgradig verfügbar zu machen, benötigen sie die folgenden Verzeichnisse:
  - Lokales Dateisystem auf dem ASCS-Server: /usr/sap/<SID>/ASCS<nn>
  - Lokales Dateisystem auf dem ERS-Server: /usr/sap/<SID>/ERS<nn>
  - Gemeinsam genutztes Dateisystem: /sapmnt , /usr/sap/trans, /usr/SAP/<SID>/SYS
  - Lokales Dateisystem auf dem Anwendungsserver: /usr/SAP/<SID>/D<nn>
4. **File Shares** – Für die gemeinsam genutzten Verzeichnisse, die als NFS auf verschiedenen Servern installiert werden sollen, sollte Amazon Elastic File Share verwendet werden.
5. **Filesysteme** – Für die beiden HANA-Knoten sollten die folgenden Dateisysteme mit jeweils derselben Größe eingerichtet werden, wobei der Dateityp von/hana/\* XFS sein und die anderen Dateitypen EXT4 sein sollten. /usr/sap, /hana/shared, /hana/data/<SID>, /hana/log/<SID>, /hana/backup

**Tab. 5.1** Testfälle für Hochverfügbarkeit

Testszenario	Erwartetes Ergebnis
<b>Anwendung – SAP-Cluster</b>	
ASCS vom primären auf den sekundären Knoten umschalten	ASCS sollte auf den sekundären Knoten wechseln, die Benutzersitzung sollte intakt bleiben
ASCS vom sekundären auf den primären Knoten umschalten	ASCS sollte auf den primären Knoten wechseln, die Benutzersitzung sollte intakt bleiben
Message-Server im primären Knoten killen	ASCS sollte auf dem sekundären Knoten starten, die Benutzersitzung sollte intakt bleiben
Message-Server im sekundären Knoten killen	ASCS sollte auf dem primären Knoten starten, die Benutzersitzung sollte intakt bleiben
ASCS sapstartsrv im primären Knoten killen	ASCS sollte auf den sekundären Knoten wechseln, die Benutzersitzung sollte intakt bleiben
ASCS sapstartsrv im sekundären Knoten killen	ASCS sollte auf den primären Knoten wechseln, die Benutzersitzung sollte intakt bleiben
Neustart des primären ASCS-Knotens	ASCS sollte auf den primären Knoten wechseln, die Benutzersitzung sollte intakt bleiben
Neustart des sekundären ASCS-Knotens	ASCS sollte auf den primären Knoten wechseln, die Benutzersitzung sollte intakt bleiben
Isolierung des Speichers eines Knotens (A/B)	Der ASCS-Dienst sollte automatisch auf dem anderen Knoten starten
<b>Anwendung – DB-Cluster</b>	
Manuellen Wechsel vom primären auf den sekundären Knoten ausführen	HDB sollte auf dem sekundären Knoten laufen, die Benutzersitzung sollte intakt bleiben
Wechsel vom sekundären auf den primären Knoten ausführen	HDB sollte auf dem primären Knoten laufen, die Benutzersitzung sollte intakt bleiben
DB auf dem primären Knoten anhalten	HDB sollte auf dem sekundären Knoten laufen, die Benutzersitzung sollte intakt bleiben
DB auf dem sekundären Knoten anhalten	HDB sollte auf dem primären Knoten laufen, die Benutzersitzung sollte intakt bleiben
Hdbindexserver im primären Knoten killen	HDB sollte auf dem sekundären Knoten laufen, die Benutzersitzung sollte intakt bleiben
Hdbindexserver im primären Knoten killen	HDB sollte auf dem primären Knoten laufen, die Benutzersitzung sollte intakt bleiben
Sonstige DB-Dienste im primären Knoten killen	HDB sollte auf dem sekundären Knoten laufen, die Benutzersitzung sollte intakt bleiben
Sonstige DB-Dienste im sekundären Knoten killen	HDB sollte auf dem primären Knoten laufen, die Benutzersitzung sollte intakt bleiben
Neustart des primären DB-Knotens	HDB sollte auf dem sekundären Knoten laufen, die Benutzersitzung sollte intakt bleiben
Neustart des sekundären DB-Knotens	HDB sollte auf dem primären Knoten laufen, die Benutzersitzung sollte intakt bleiben

6. **Overlay-IP-Adressen erstellen** – Erstellung der virtuellen IP (Overlay-IP) in AWS sowie Zuordnung zu allen HA-Knoten.
7. **S/4 ASCS und ERS installieren** – Installation von ASCS in einem Knoten, Installation von ERS in einem anderen Knoten. Der virtuelle Host und die Overlay-IP sind gemäß SAP-HA-Standardleitfaden zu setzen.
8. **Clustereinrichtung konfigurieren** – Einrichtung der Pacemaker-Konfiguration im ASCS- und im ERS-Knoten.
9. **HANA DB installieren** – Installation der eigenständigen HANA DB in beiden DB-Knoten mit virtuellem Hostnamen und Overlay-IP gemäß SAP-Standardleitfaden.
10. **HANA-Systemreplikation konfigurieren** – Konfiguration von HSR mit synchronem Betriebsmodus und Logreplay-Replikationsmodus.
11. **Clustereinrichtung konfigurieren** – Einrichtung der Pacemaker-Konfiguration in beiden HANA-Knoten.
12. **S/4-Anwendungsserver installieren** – Installation der S/4-Anwendungsserver in beiden Availability Zones. Die Anwendungsserver gehören nicht zum Cluster.
13. **Cluster testen** – Durchführung strenger Clustertests wie nachstehend beschrieben.

Für die Einrichtung der Hochverfügbarkeitslösung auf einer RHEL- oder SUSE-Plattform sind die folgenden Standard-KBA für S/4 in AWS zu befolgen.

- Hochverfügbarkeit S/4 für **SUSE** (Zugriff am 20.12.2021): [https://documentation.suse.com/sbp/all/html/SAP\\_HA740\\_SetupGuide\\_AWS/index.html](https://documentation.suse.com/sbp/all/html/SAP_HA740_SetupGuide_AWS/index.html)
- Hochverfügbares SAP HANA für **SUSE** (Zugriff am 20.12.2021): [https://documentation.suse.com/sbp/all/html/SLES4SAP-hana-sr-guide-PerfOpt-15\\_AWS/index.html](https://documentation.suse.com/sbp/all/html/SLES4SAP-hana-sr-guide-PerfOpt-15_AWS/index.html)
- Hochverfügbares S/4 für **RHEL** (Zugriff am 20.12.2021): [https://documentation.suse.com/sbp/all/html/SAP\\_HA740\\_SetupGuide\\_AWS/index.html](https://documentation.suse.com/sbp/all/html/SAP_HA740_SetupGuide_AWS/index.html)  
<https://access.redhat.com/articles/3916511>
- Hochverfügbares SAP HANA für **RHEL** (Zugriff am 20.12.2021): <https://access.redhat.com/articles/3569621>

### **5.1.2.2 Verfügbarkeitseinrichtung von SAP NetWeaver auf einer Windows-Plattform**

Auf einer Windows-Plattform lässt sich SAP mit hochgradig verfügbaren ASCS- und ERS-Services auf zwei Arten einrichten:

- Windows Server Failover Clustering (WSFC) mit **Amazon FSx** als Dateifreigabelösung
- Windows Server Failover Clustering (WSFC) mit der **SIOS Protection Suite** in AWS

## Windows Server Failover Clustering (WSFC) mit Amazon FSx als Dateifreigabelösung

Für hochgradige SAP-Verfügbarkeit in Windows müssen WSFC-Cluster verwendet werden. Bei dieser Lösung werden Amazon-FSx-Dateisysteme für die gemeinsam genutzten Verzeichnisse verwendet. Wie wir bereits wissen, benötigen wir gemeinsam genutzte Dateisysteme, um ASCS und ERS hochverfügbar zu konfigurieren. Amazon FSx ist ein vollständig verwaltetes AWS-Dateisystem (wie EFS), das über Availability Zones hinweg gemeinsam genutzt werden kann. Für die gemeinsame Nutzung von Dateien in Windows muss Amazon FSx für mehrere Availability Zones eingerichtet werden. Weitere Informationen zur Einrichtung von hochverfüglichen SAP-S/4-Systemen unter Windows mit Amazon FSx finden Sie in dem folgenden AWS-Leitfaden (Zugriff am 20.12.2021):

<https://aws.amazon.com/blogs/awsforsap/how-to-setup-sap-netweaver-on-windows-mscs-for-sap-ascss-ers-on-aws-using-amazon-fsx/>

## Windows Server Failover Clustering (WSFC) mit der SIOS Protection Suite in AWS

Um beim Ausfall einer Zone oder bei einer Serviceunterbrechung eine reibungslose Ausfallsicherung gewährleisten zu können, ist die SIOS Protection Suite sehr stark in das WSFC-Cluster integriert. Der SIOS DataKeeper ist Teil der SIOS Protection Suite. Der Data Keeper ist eine serverbasierte Replikationslösung (ähnlich NFS oder GFS), die zur Verwaltung der Windows-HA-Lösungen eine Datenreplikation auf Blockebene über die Availability Zones hinweg durchführt. Es wird empfohlen, in beiden Zonen getrennte VMs für das gemeinsam genutzte Dateisystem bereitzustellen. Diese sollten synchron repliziert und auf ASCS-/ERS-Servern installiert werden, damit es bei Zonenausfällen nicht zu Serviceunterbrechungen kommt.

/sapmnt und/usr/sap/trans müssen von einer zentralen Dateifreigabelösung (Amazon FSx oder SIOS DataKeeper) verwaltet und von den VMs gemeinsam genutzt werden.

Weitere Informationen zur Einrichtung von SAP-S/4-HA-Systemen unter Windows mit der SIOS Protection Suite finden Sie in dem folgenden AWS-Leitfaden (Zugriff am 20.12.2021):

<https://aws.amazon.com/blogs/awsforsap/deploying-highly-available-sap-systems-using-sios-protection-suite-on-aws/>

---

## 5.2 Computing

In diesem Kapitel erläutern wir die verschiedenen Möglichkeiten, virtuelle Maschinen für das S/4-System bereitzustellen. Wir beschreiben die verfügbaren Templates und gehen darauf ein, welche dieser Templates für HANA-Datenbanken freigegeben und welche allgemein für SAP-Workloads freigegeben sind. Dies soll Unterstützung bei der Auswahl des am besten geeigneten VM-Templates bieten. Darüber hinaus werden

die zur Bereitstellung des VM-Templates benötigten Build-Daten beschrieben. Wir beschreiben zwei mögliche Vorgehensweisen: die manuelle und die skriptbasierte. Weiterhin erläutern wir im vorliegenden Kapitel, wie VM-Templates geändert werden können, um sie an die umfangreicheren Workloads eines S/4-Systems anzupassen. Ein weiteres Thema ist die Erstellung von Templates auf der Grundlage bestehender Systeme.

Der Begriff „Computing“ bezieht sich auf die für die Workload-Verarbeitung benötigte Rechenleistung. Als Architekt sind Sie für die korrekte Dimensionierung der Rechenleistung verantwortlich. Diese Dimensionierung muss sich nach der Workload-Größe richten (groß, mittelgroß, klein) und darf die Anforderungen weder übererfüllen noch hinter ihnen zurückbleiben. Amazon EC2 (Elastic Compute Cloud) bietet die Möglichkeit, virtuelle Maschinen namens EC2-Instanzen zu starten und sich auf diese Weise an die benötigte Rechenleistung anzupassen. Es müssen nicht im Vorhinein große Instanzen eingeplant werden, sondern die Größe der VM kann je nach Auslastung auf einfache Weise (nach oben und nach unten) angepasst werden. Bei Amazon ist eine Vielzahl verschiedener Instanztypen verfügbar. Sie werden hauptsächlich nach CPU, Arbeits- und Datenspeicher sowie Netzwerkperformance unterschieden. Wenn ein leistungsstärkerer Arbeitsspeicher benötigt wird, können CPU-optimierte EC2-Instanztypen verwendet werden. Wenn ein umfangreicherer Arbeitsspeicher benötigt wird (z. B. für die HANA-DB), können speicheroptimierte EC2-Instanzen verwendet werden etc.

Alle EC2-Instanzen sind mit dedizierten Ressourcen wie CPU, Arbeitsspeicher, Instanzspeicher sowie mit gemeinsam genutzten Ressourcen des Host-Rechners verbunden, darunter beispielsweise das Netzwerk. Alle Instanztypen bieten IOPS aus gemeinsam genutzten Ressourcen, wobei den performanteren Instanztypen mehr gemeinsam genutzte Ressourcen zugewiesen werden.

### 5.2.1 Verfügbare EC2-Instanzen

Zur Unterstützung aller Arten von Workloads bietet Amazon eine breite Palette an EC2-Instanztypen.

C\* (C4, C5, C5a, C6g etc.): Die in Bezug auf die **Rechenleistung optimierten** Instanztypen eignen sich am besten für Workloads, für die ein Hochleistungsprozessor benötigt wird.

D\*, H1, I3, I3en: Die in Bezug auf den **Datenspeicher optimierten** Instanztypen sind zu verwenden, wenn lokal umfangreiche sequenzielle Lese- und Schreibvorgänge mit einem großen Datenvolumen benötigt werden. Diese Instanztypen bieten niedrige Latenzzeiten und eine sehr hohe Anzahl von IOPS.

M\*, T\* (M4, M5, M5a, T2, T3, T3a etc.): **Allzweckinstanztypen** mit einem ausgewogenen Verhältnis zwischen CPU und Arbeitsspeicher, die sich ideal für kleine bis mittelgroße Datenbanken und Workloads eignen.

R\* (r4, r5, r5a, r5ad and etc.): In Bezug auf den **Arbeitsspeicher optimierte** Instanztypen. Diese eignen sich ideal für Workloads, bei denen große Datenmengen im Arbeitsspeicher verarbeitet werden müssen. Es wird empfohlen diese Instanztypen in In-Memory-Datenbanken wie HANA zu verwenden.

Amazon bietet zur Unterstützung von Workloads mit sehr hohem Speicherbedarf auch einige Bare-Metal-Lösungen wie u6tb1.metal, u-9tb1.metal, u-12tb1.metal, u-18tb1.metal und u-24tb1.metal. Diese Lösungen sind als Bare-Metal-Instanz auf einem dedizierten EC2-Host konzipiert.

### 5.2.2 Lizenzierung der EC2-Instanzen

Für SAP-Workloads werden am häufigsten die folgenden drei Lizenztypen verwendet:

**On-Demand-Instanz:** Bei On-Demand-Instanzen wird je nach den betriebenen Instanzen und je nach Nutzung entweder stunden- oder minutenweise abgerechnet. Es müssen weder langfristige Verträge abgeschlossen noch Vorauszahlungen getätigt werden. Die Rechenkapazität kann je nach Anwendungsbedarf flexibel gesteigert und reduziert werden, und für die genutzten Instanzen fällt lediglich der Stundensatz an. Dies ist insbesondere dann sehr vorteilhaft, wenn die Workloads nur zu bestimmten Zeiten anfallen.

**Reserved Instances:** Wenn die Kapazitäten in einer bestimmten Availability Zone benötigt werden, können Sie mit Reserved Instances reserviert werden. Im Vergleich zu den Preisen für On-Demand-Instanzen sind Amazon EC2 Reserved Instances (RI) in einem solchen Szenario stark vorteilhaft (bis zu 72 % Abschlag). Reserved Instances empfehlen sich für alle Workloads in Produktion, die rund um die Uhr laufen sollen.

**Dedicated Hosts:** Ein dedizierter Host ist ein physischer EC2-Server, der für eine bestimmte Nutzung reserviert wird. Mit Dedicated Hosts lassen sich Kosten sparen, indem die vorhandenen serverbasierten Softwarelizenzen genutzt werden, darunter Windows Server, SQL Server und SUSE Linux Enterprise Server (je nach den betreffenden Lizenzbedingungen). Dedicated Hosts können je nach Bedarf On-Demand oder als Reserved Instances (mit 70 % Rabatt).

### 5.2.3 SAP-HANA-zertifizierte EC2-Instanztypen

AWS arbeitet bei der Zertifizierung von EC2-Aufrechnungsinstanzen für SAP- und HANA-Workloads eng mit SAP zusammen. Dem nachstehenden Link ist die aktuellen SAP-zertifizierten EC2-Instanztypen der neuesten Generation zu entnehmen (Zugriff am 20.12.2021):

[https://aws.amazon.com/sap\(instance-types/](https://aws.amazon.com/sap(instance-types/)

Für produktive und nicht-produktive HANA-Workloads werden die folgenden Instanztypen unterstützt:

- u-6tb1.metal, u-9tb1.metal, u-12tb1.metal, u-18tb1.metal, u-24tb1.metal-u-6tb1.56xlarge, u-6tb1.112xlarge, u-9tb1.112xlarge, und u-12tb1.112xlarge
- x1.16xlarge, x1.32xlarge und x1e.32xlarge
- r3.8xlarge
- r4.8xlarge und r4.16xlarge
- r5.8xlarge, r5.12xlarge, r5.16xlarge, r5.24xlarge und r5.metal
- r5b.8xlarge, r5b.12xlarge, r5b.16xlarge, r5b.24xlarge und r5b.metal

Die folgenden Instanztypen werden nur für nicht produktive HANA-Workloads unterstützt:

- x1e.xlarge, x1e.2xlarge, x1e.4xlarge
- r3.2xlarge, r3.4xlarge
- r4.2xlarge, r4.4xlarge
- r5.2xlarge, r5.4xlarge
- r5b.2xlarge, r5b.4xlarge, r5b.4xlarge

Dem nachstehenden SAP-Link ist die aktuelle Liste der zertifizierten HANA-VMs in AWS zu entnehmen (Zugriff am 20.12.2021):

<https://www.sap.com/dmc/exp/2014-09-02-hana-hardware/enEN/#/solutions?filters=ias;ve:23>

## 5.2.4 VM-Bereitstellung

In diesem Abschnitt beschreiben wir die für die Bereitstellung von virtuellen Maschinen in AWS erforderlichen Komponenten und Services sowie die wichtigsten Schritte bei der Bereitstellung der VM. Zunächst müssen die Kundenanforderungen bewertet und erfasst werden, um die zugrunde liegenden Clouds und die Landing Zone zu erstellen.

Zu berücksichtigen sind folgende Punkte:

- **Region** – Wählen Sie die primäre und die DR-Region.
- **Availability Zones** – Wählen Sie die Availability Zones sowie die Art der Bereitstellung aus (eine oder mehrere AZs).
- **Netzwerk** – Entwerfen Sie die Netzwerksegmentierung.
  - Virtual Private Cloud (VPC): Anzahl der für die Trennung von anderen virtuellen Netzwerken erforderlichen VPCs, seien es verwaltete VPC oder SAP VPC
  - Privates Subnetz: Entwerfen Sie ein privates Subnetz für die SAP-Workloads, z. B. ein Subnetz für jede SAP SID oder ein Subnetz für nicht-produktive Workloads und eines für produktive Workloads oder jeweils verschiedene Subnetze für die SAP-S/4-Anwendungen und die HANA-DB-Schichten.

- Öffentliches Subnetz: Bastion-Host in einem öffentlichen Subnetz mit einer Elastic-IP-Adresse für den Zugriff auf die EC2-Instanzen.
- Internet-Gateway für den Zugang zum Internet
- NAT-Gateway: Das NAT-Gateway ermöglicht den Zugang zum Internet (ausgehender Datenverkehr) für private Subnetzressourcen.
- **AWS-Konnektivität** – Sie können das Unternehmensnetzwerk auf zwei Weisen mit der AWS VPC verbinden:
  - VPN-Verbindung – Verschlüsselte IPSec-Verbindung zwischen On-Premises und AWS VPC. Sie können mehrere schnell und leicht konfigurierbare VPN-Verbindungen zu einer VPC erstellen.
  - Direct Connect: Mit Direct Connect können Sie eine Verbindung zur AWS VPC über Ihr Rechenzentrum herstellen. Direct Connect bietet eine hohe Bandbreite und einen hohen Durchsatz.
- **Sicherheit** – Geben Sie zur Beschränkung des ein- und ausgehenden Datenverkehrs eine Sicherheitsrichtlinie vor.

Um Ressourcen auf AWS bereitzustellen zu können, muss ein AWS-Konto erstellt werden. Nach der Bereitstellung der AWS Landing Zone kann die Bereitstellung von S/4- und HANA-VMs beginnen.

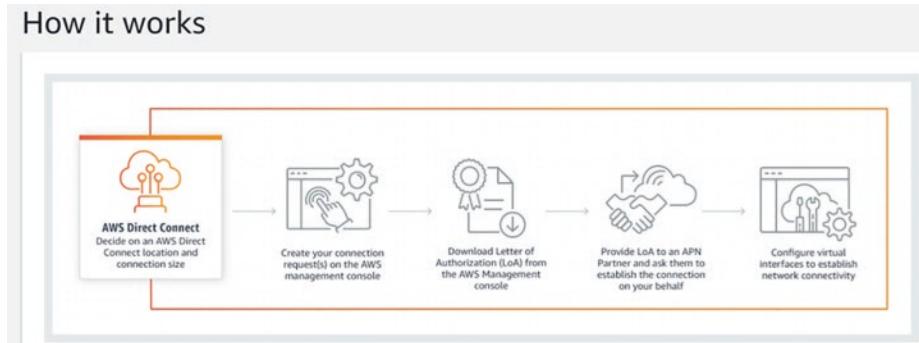
Für die Bereitstellung von S/4-Anwendungen und HANA-DB-Servern in AWS sind die folgenden Informationen notwendig:

1. Konto-ID
2. IAM-Rollen/-Richtlinien
3. VPC/Subnetz/Region
4. Sicherheitsgruppe
5. ggf. S3-Container
6. Speicherkapazität und Struktur des Dateisystems – Installationspunkt, Laufwerk (physisch und logisch), Größe, Speichertyp (EBS/EFS), Dateisystemtyp
7. Instanzgröße
8. Plattform (Windows/RHEL/OEL)
9. Physischer und logischer Hostname
10. UID/GID-Daten
11. Die Betriebssystempakete sollten gemäß der SAP-Empfehlung installiert werden
12. Server-Tagging-Informationen

Nun können die zugrundeliegenden Clouds und die EC2-Instanzen manuell oder vollautomatisch mit einem Skript bereitgestellt werden.

### **Manuelle Bereitstellung mit der AWS-Konsole**

Im folgenden Beispiel werden Erscheinungsbild und Funktionalitäten der AWS-Konsole sowie die Ressourcenerstellung auf AWS erläutert.



**Abb. 5.5** Manuelle Erstellung einer VM – Schritt 1

Erstellen Sie eine Direct-Connect-Verbindung (Abb. 5.5):

- Suchen Sie in der AWS-Konsole nach dem Service
- Oder gehen Sie auf „Alle Services“, „Netzwerk & Bereitstellung von Inhalten“, „Direct Connect“.

Klicken Sie auf „Erstellen einer Verbindung“. Nach Angabe aller Informationen wie Name, Standort, an dem sich Ihre Verbindung befindet, Port-Geschwindigkeit, Service-Anbieter erstellt Billto-Tag die Verbindung (Abb. 5.6 und 5.7).

Erstellen Sie VPC, Subnetz, NAT-Gateway (Abb. 5.8)

- Erstellen Sie die elastische IP-Adresse, bevor Sie die VPC erstellen.

Die zugrundeliegende Cloud-Umgebung ist nun erstellt, und die VM für SAP S/4 und HANA können wie nachstehend beschrieben bereitgestellt werden. Im Beispiel wählen wir SLES für SAP 15 SP1 als AWS-Maschinenimage mit Pay-As-You-Go-Abonnement (PAYG). Es kann auch Bring Your Own Subscription verwendet werden (Abb. 5.9, 5.10, 5.11, 5.12, 5.13, 5.14, 5.15, 5.16, 5.17, 5.18, 5.19, 5.20, 5.21).

### Automatisierte Bereitstellung mithilfe von Skripten

Wenn die Server zahlreich sind, wäre es sehr mühsam und hektisch, sie manuell einzurichten zu müssen. Um den manuellen Aufwand zu reduzieren, können VM und SAP-Anwendung vollautomatisch eingerichtet werden. Für die automatisierte Bereitstellung müssen Terraform und ein Ansible-Skript verwendet werden.

Terraform verwendet zur Bereitstellung von Cloud-Ressourcen wie z. B. EC2-Instanzen und Festplatten sowie für die Einrichtung von Lastenverteilern etc. API-Aufrufe. Ansible-Skripte hingegen werden für die Konfiguration der Betriebssystemeinstellungen verwendet. Beispiele hierfür sind Festplatteninstallation, Dateisystemstruktur, Erstellung

## Connection settings

### Name

A name to help you identify the connection.

Name must contain no more than 100 characters. Valid characters are a-z, 0-9, and – (hyphen)

### Location

The location in which your connection is located.



### Port speed

Desired bandwidth for the new connection.

 1Gbps 10Gbps

### On-premises

 Connect through an AWS Direct Connect partner.

**Abb. 5.6** Manuelle Erstellung einer VM – Schritt 2

### Service provider

Service provider providing connectivity for your connection at this location.



### ► Additional settings

**Abb. 5.7** Manuelle Erstellung einer VM – Schritt 3

von Benutzergruppen, Konfiguration der Betriebssystemparameter usw. Das folgende Diagramm zeigt die Funktionsweise (Abb. 5.22).

Sie müssen wie folgt vorgehen:

1. Erstellt einen Klon des Terraform Git Repositorys auf den Laptop.
2. Startet Terraform vom Laptop aus, um die zugrunde liegende Cloud-Umgebung wie VPC, Subnetz, Route 53, NLB, Transit-Gateway usw. bereitzustellen.
3. In AWS muss ein Toolserver bereitgestellt werden, von dem aus Terraform und Ansible gestartet werden können, um die VM und SAP bereitzustellen.

### Step 1: Select a VPC Configuration

VPC with a Single Public Subnet

**VPC with Public and Private Subnets**

VPC with Public and Private Subnets and Hardware VPN Access

VPC with a Private Subnet Only and Hardware VPN Access

In addition to containing a public subnet, this configuration adds a private subnet whose instances are not addressable from the Internet. Instances in the private subnet can establish outbound connections to the Internet via the public subnet using Network Address Translation (NAT).  
Creates:

A /16 network with two /24 subnets. Public subnet instances use Elastic IPs to access the Internet. Private subnet instances access the Internet via Network Address Translation (NAT). (Hourly charges for NAT devices apply.)

**Important:**  
If you are using a Local Zone with your VPC [follow this link](#) to create your VPC.

Select

**Abb. 5.8** Manuelle Erstellung einer VM – Schritt 4

### Step 2: VPC with Public and Private Subnets

IPv4 CIDR block:\* 10.0.0.0/16 (65531 IP addresses available)

IPv6 CIDR block:  No IPv6 CIDR Block  
 Amazon provided IPv6 CIDR block  
 IPv6 CIDR block owned by me

VPC name: sapvpc

Public subnet's IPv4 CIDR:\* 10.0.0.0/24 (251 IP addresses available)

Availability Zone:\* us-east-1a

Public subnet name: sn\_sappublic

Private subnet's IPv4 CIDR:\* 10.0.1.0/24 (251 IP addresses available)

Availability Zone:\* us-east-1a

Private subnet name: sn\_S4app\_S4D

You can add more subnets after Amazon Web Services creates the VPC.

Specify the details of your NAT gateway ([NAT gateway rates apply](#)).

Elastic IP Allocation ID:\* eipalloc-0ece118a276ca2f3d

**Abb. 5.9** Manuelle Erstellung einer VM – Schritt 5

4. Erstellt auf dem Toolserver einen Klon des Terraform- und des Ansible-Git-Repositorys.
5. Startet Terraform vom Toolserver aus, um die SAP-Infrastruktur bereitzustellen.
6. Startet Ansible vom Toolserver aus, um das Betriebssystem zu konfigurieren und S4 und HANA bereitzustellen.

Public subnet's IPv4 CIDR\*: 10.0.0.0/24 (251 IP addresses available)

Availability Zone\*: us-east-1a

Public subnet name: sn\_sappublic

Private subnet's IPv4 CIDR\*: 10.0.1.0/24 (251 IP addresses available)

Availability Zone\*: us-east-1a

Private subnet name: sn\_S4app\_S4D

You can add more subnets after Amazon Web Services creates your VPC.

Specify the details of your NAT gateway (NAT gateway rates apply).

Elastic IP Allocation ID\*: eipalloc-0ece118a276ca2f3d

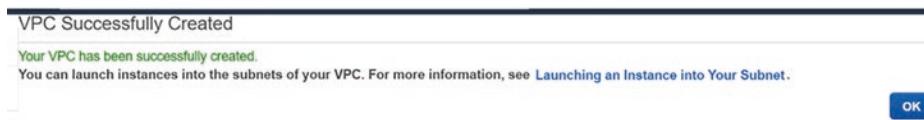
Service endpoints

Add Endpoint

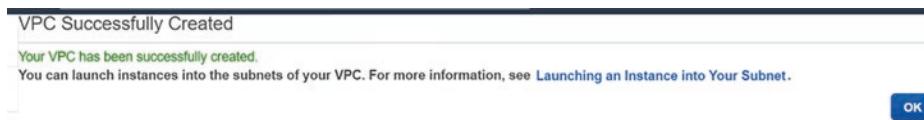
Enable DNS hostnames\*: Yes  No

Hardware tenancy\*: Default

Waiting for NAT Gateway status to become available (This may take a few minutes)



**Abb. 5.10** Manuelle Erstellung einer VM – Schritt 6



**Abb. 5.11** Manuelle Erstellung einer VM – Schritt 7



**Abb. 5.12** Manuelle Erstellung einer VM – Schritt 8

**Terraform** bietet die Möglichkeit, die SAP- und DB-Infrastruktur automatisch mit Terraform-Konfigurationsdateien zu erstellen. Die Terraform-Workflows benötigen die Befehle „plan“, „apply“ und „destroy“, und um diese verwenden zu können, muss ein

**SUSE Linux Enterprise Server for SAP Applications 15 SP1**

**SUSE**

**SUSE Linux Enterprise Server for SAP Applications 15 SP1**

SUSE Linux Enterprise Server (SLES) for SAP Applications is the leading Linux platform for SAP HANA, SAP NetWeaver, SAP S/4HANA and SAP Business Applications providing optimized performance and reduced downtime as well as faster SAP system deployments.

Reduce complexity of managing SAP landscapes with SUSE's Expanded Service Pack Overlap ...

[More info](#)

[View Additional Details in AWS Marketplace](#)

**Product Details**

By Amazon Web Services

**Customer Rating** ★★★★☆ (1)

I select Marvin... v20210704

Instance Type	Software	EC2	Total
t2.large	\$0.43	\$0.093	<b>\$0.523/hr</b>
t3.micro	\$0.213	\$0.01	<b>\$0.223/hr</b>
t3.small	\$0.213	\$0.021	<b>\$0.234/hr</b>
c3.large	\$0.43	\$0.105	<b>\$0.535/hr</b>
c3.xlarge	\$0.43	\$0.21	<b>\$0.64/hr</b>
c3.2xlarge	\$0.51	\$0.42	<b>\$0.93/hr</b>
c3.4xlarge	\$0.51	\$0.84	<b>\$1.35/hr</b>
c3.8xlarge	\$0.51	\$1.68	<b>\$2.19/hr</b>
c4.large	\$0.43	\$0.10	<b>\$0.53/hr</b>
c4.xlarge	\$0.43	\$0.199	<b>\$0.629/hr</b>
c4.2xlarge	\$0.51	\$0.398	<b>\$0.908/hr</b>

You will not be charged until you launch this instance.

[Cancel](#) [Continue](#)

Abb. 5.13 Manuelle Erstellung einer VM – Schritt 9

1. Choose AMI   2. Choose Instance Type   3. Configure Instance   4. Add Storage   5. Add Tags   6. Configure Security Group   7. Review

**Step 2: Choose an Instance Type**

	r5	r5.large	2	16	EBS only	Yes	Up to 10 Gigabit	Yes
<input checked="" type="checkbox"/>	r5	r5.xlarge	4	32	EBS only	Yes	Up to 10 Gigabit	Yes
<input type="checkbox"/>	r5	r5.2xlarge	8	64	EBS only	Yes	Up to 10 Gigabit	Yes
<input type="checkbox"/>	r5	r5.4xlarge	16	128	EBS only	Yes	Up to 10 Gigabit	Yes
<input type="checkbox"/>	r5	r5.8xlarge	32	256	EBS only	Yes	10 Gigabit	Yes
<input type="checkbox"/>	r5	r5.12xlarge	48	384	EBS only	Yes	10 Gigabit	Yes
<input type="checkbox"/>	r5	r5.16xlarge	64	512	EBS only	Yes	20 Gigabit	Yes
<input type="checkbox"/>	r5	r5.24xlarge	96	768	EBS only	Yes	25 Gigabit	Yes
<input type="checkbox"/>	r5	r5.32xlarge	128	1024	EBS only	Yes	30 Gigabit	Yes

[Cancel](#) [Previous](#) [Review and Launch](#) [Next: Configure Instance Details](#)

Abb. 5.14 Manuelle Erstellung einer VM – Schritt 10

initialisiertes Arbeitsverzeichnis vorhanden sein. Terraform wird in dem Arbeitsverzeichnis ausgeführt, in dem alle betreffenden Konfigurationsdateien gespeichert sind.

Der Befehl „terraform init“ wird verwendet, um das Arbeitsverzeichnis zu initialisieren, das Terraform verwenden soll.

```
terraform init
```

Der Terraform-Befehl „Plan“ wird verwendet, um die gewünschte Konfiguration zu evaluieren und mit der bestehenden Zielausstruktur abzugleichen. Bei Abweichungen

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Add Tags 6. Configure Security Group 7. Review

### Step 3: Configure Instance Details

Configure the instance to suit your requirements. You can launch multiple instances from the same AMI, request Spot instances to take advantage of the lower pricing, assign an access management role to the instance, and more.

Number of instances	<input type="text" value="1"/>	Launch into Auto Scaling Group
Purchasing option	<input type="checkbox"/> Request Spot instances	
Network	vpc-0b9de29ab49a4d694   sapvpc	<input type="button" value="Create new VPC"/>
Subnet	subnet-0194d7c00c6a024d   sn_S4app_S4D   us-east-1	<input type="button" value="Create new subnet"/> 251 IP Addresses available
Auto-assign Public IP	<input type="checkbox"/> Use subnet setting (Disable)	
Placement group	<input type="checkbox"/> Add instance to placement group	
Capacity Reservation	<input type="button" value="Open"/>	
Domain join directory	No directory	<input type="button" value="Create new directory"/>
IAM role	<input type="checkbox"/> None	
CPU options	<input type="checkbox"/> Specify CPU options	
Shutdown behavior	<input type="button" value="Stop"/>	
Stop - Hibernate behavior	<input type="checkbox"/> Enable hibernation as an additional stop behavior	
Enable termination protection	<input type="checkbox"/> Protect against accidental termination	
Monitoring	<input type="checkbox"/> Enable CloudWatch detailed monitoring <small>Additional charges apply.</small>	
EBS-optimized instance	<input type="checkbox"/> Launch as EBS-optimized instance	

**Abb. 5.15** Manuelle Erstellung einer VM – Schritt 11

### Step 3: Configure Instance Details

Tenancy	<input type="checkbox"/> Shared - Run a shared hardware instance	<small>Additional charges will apply for dedicated tenancy.</small>																
Elastic Inference	<input type="checkbox"/> Add an Elastic Inference accelerator																	
<small>Additional charges apply.</small>																		
File systems	<input type="button" value="Add file system"/>	<input type="button" value="Create new file system"/>																
<b>▼ Network interfaces</b> <table border="1"> <thead> <tr> <th>Device</th> <th>Network Interface</th> <th>Subnet</th> <th>Primary IP</th> <th>Secondary</th> <th>IPv6 IPs</th> <th>IPv4 Prefixes</th> <th>IPv6 Prefix</th> </tr> </thead> <tbody> <tr> <td>eth0</td> <td>New network interface</td> <td>subnet-0194d7c5</td> <td>Auto-assign</td> <td>Add IP</td> <td>The selected subnet does not support IPv6 because it does not have an IPv6 CIDR.</td> <td>None</td> <td></td> </tr> </tbody> </table>			Device	Network Interface	Subnet	Primary IP	Secondary	IPv6 IPs	IPv4 Prefixes	IPv6 Prefix	eth0	New network interface	subnet-0194d7c5	Auto-assign	Add IP	The selected subnet does not support IPv6 because it does not have an IPv6 CIDR.	None	
Device	Network Interface	Subnet	Primary IP	Secondary	IPv6 IPs	IPv4 Prefixes	IPv6 Prefix											
eth0	New network interface	subnet-0194d7c5	Auto-assign	Add IP	The selected subnet does not support IPv6 because it does not have an IPv6 CIDR.	None												

1. Choose AMI 2. Choose Instance Type 3. Configure Instance 4. Add Storage 5. Add Tags 6. Configure Security Group 7. Review

### Step 4: Add Storage

Volume Type	Device	Snapshot	Size (GiB)	Volume Type	IOPS	Throughput (MB/s)	Delete on Termination	Encryption
Root	/dev/sda1	snapshot-0b1abedb626e899ec	10	General Purpose SSD (gp2)	100 / 3000	N/A	<input checked="" type="checkbox"/>	alias/aws/ebs
EBS	/dev/sdb	Search (case-insensitive)	25	General Purpose SSD (gp2)	100 / 3000	N/A	<input type="checkbox"/>	alias/aws/ebs

**Add New Volume**

Free tier eligible customers can get up to 30 GB of EBS General Purpose (SSD) or Magnetic storage. [Learn more](#) about free usage tier eligibility and usage restrictions.

**Abb. 5.16** Manuelle Erstellung einer VM – Schritt 12

**Step 5: Add Tags**

A tag consists of a case-sensitive key-value pair. For example, you could define a tag with key = Name and value = Webserver. A copy of a tag can be applied to volumes, instances or both. Tags will be applied to all instances and volumes. [Learn more](#) about tagging your Amazon EC2 resources.

Key	(128 characters maximum)	Value	(256 characters maximum)	Instances	Volumes	Network Interfaces
SAP S/4 S4D		600099		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

Add another tag (Up to 50 tags maximum)

**Step 6: Configure Security Group**

A security group is a set of network rules that control the traffic for your instance. On this page, you can add rules to allow specific traffic to reach your instance. For example, if you want to set up a web server and allow Internet traffic to reach your instance, add rules that allow unrestricted access to the HTTP and HTTPS ports. You can create a new security group or select from an existing one below. [Learn more](#) about Amazon EC2 security groups.

Assign a security group:  Create a new security group  Select an existing security group

Security group name: SUSE Linux Enterprise Server for SAP Applications 15 SP1-v20210304-Autogen

Description: This security group was generated by AWS Marketplace and is based on recom

Type	Protocol	Port Range	Source	Description
SSH	TCP	22	Anywhere	e.g. SSH for Admin Desktop
HTTP	TCP	80	Anywhere	e.g. SSH for Admin Desktop

Add Rule

**Abb. 5.17** Manuelle Erstellung einer VM – Schritt 13

wird eine Beschreibung der Änderungen geliefert, die für die Erreichung derselben Struktur notwendig sind. Diese Änderungen werden lediglich beschrieben, aber nicht durchgeführt.

`terraform plan`

Der Befehl „`terraform apply -var-file=<<input_file_name>>`“ entspricht dem Befehl „`plan`“, aber zusätzlich werden die für die Erreichung des gewünschten Zustands notwendigen Änderungen durchgeführt. Vor der Durchführung der Änderungen werden Sie um eine Bestätigung gebeten, sofern dies nicht ausdrücklich als unerwünscht angegeben wurde.

`terraform destroy -var-file=<<input_file_name>>`

Mit dem Befehl „`terraform destroy -var-file=<<input_file_name>>`“ werden alle im aktuellen Arbeitsverzeichnis verwalteten Ressourcen außer Betrieb genommen.

## Step 7: Review Instance Launch

**AMI Details**

SUSE Linux Enterprise Server for SAP Applications 15 SP1  
SUSE Linux Enterprise Server for SAP Applications 15 SP1 (HVM, 64-bit, SSD-Backed)

Root Device Type: ebs Virtualization type: hvm

Hourly Software Fees: \$0.43 per hour on r5.xlarge instance. Additional taxes or fees may apply.  
Software charges will begin once you launch this AMI and continue until you terminate the instance.

Annual Subscriptions are available for this product, which can save you up to 70% when compared to hourly prices.

To purchase an Annual Subscription go to the [Your Software](#) page after launching the instance.

If you have an existing license entitlement to use this software, then you can launch this software without creating a new subscription. If you do not have an existing entitlement, then by launching this software, you will be subscribed to this software and agree that your use of this software is subject to the pricing

**Instance Type** [Edit instance type](#)

Instance Type	ECUs	vCPUs	Memory (GiB)	Instance Storage (GB)	EBS-Optimized Available	Network Performance
r5.xlarge	-	4	32	EBS only	Yes	Up to 10 Gigabit

**Security Groups** [Edit security groups](#)

**Security group name** SUSE Linux Enterprise Server for SAP Applications 15 SP1-v20210304-AutogenByAWSMP-  
**Description** This security group was generated by AWS Marketplace and is based on recommended settings for SUSE Linux Enterprise Server for SAP Applications 15 SP1 version v20210304 provided by Amazon Web Services

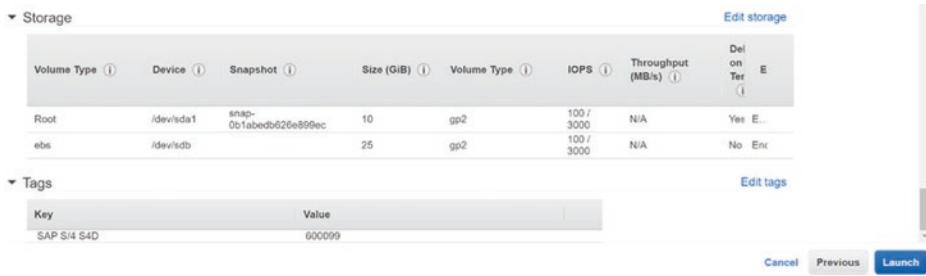
Type	Protocol	Port Range	Source	Description
SSH	TCP	22	0.0.0.0/0	

**Abb. 5.18** Manuelle Erstellung einer VM – Schritt 14

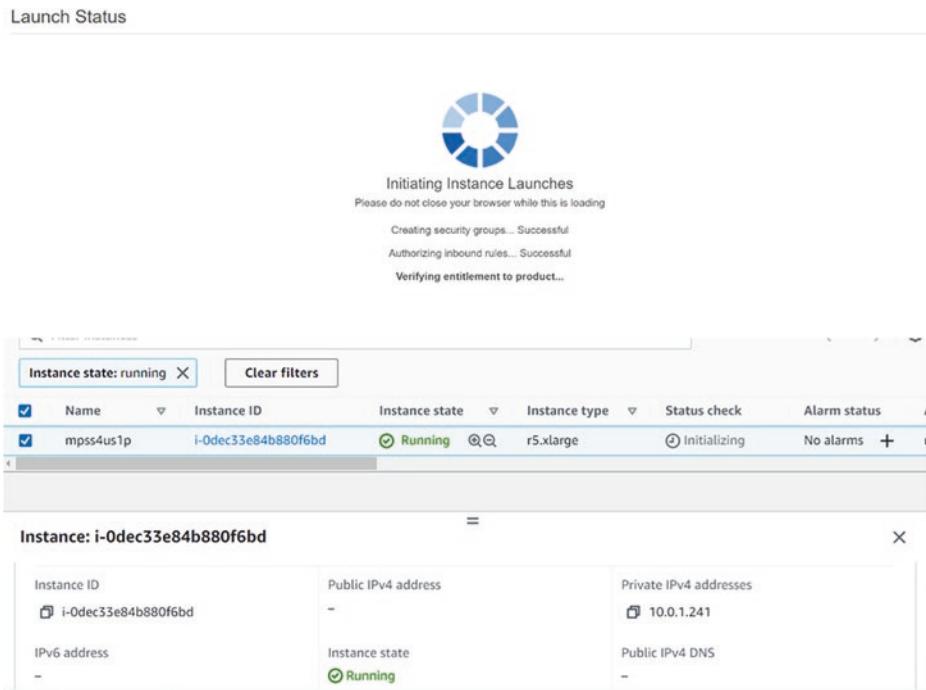
**Instance Details**

Number of Instances	1	Purchasing option	On demand			
Network	vpc-0b9de29ab49a4d694					
Subnet	subnet-0194d7c90cf6a024d					
EBS-optimized	Yes					
Monitoring	No					
Termination protection	No					
Shutdown behavior	Stop					
Stop - Hibernate behavior	Disabled					
Capacity Reservation	open					
IAM role	None					
Domain join directory	None					
Tenancy	default					
Credit specification						
Host ID						
Host resource group name						
Affinity	Off					
Kernel ID	Use default					
RAM disk ID	Use default					
Enclave	false					
Metadata accessible	Enabled					
Metadata version	V1 and V2 (token optional)					
Metadata token response hop limit	1					
User data						
Assign Public IP	Use subnet setting (Disable)					
Assign IPv6 IP	Use subnet setting (Disable)					
Network interfaces						
Device	Network Interface	Subnet	Primary IP	Secondary IP Addresses	IPv4 Prefixes	IPv6 Prefixes
eth0	New network interface	subnet-0194d7c90cf6a024d	Auto-assign		None	None

**Abb. 5.19** Manuelle Erstellung einer VM – Schritt 15



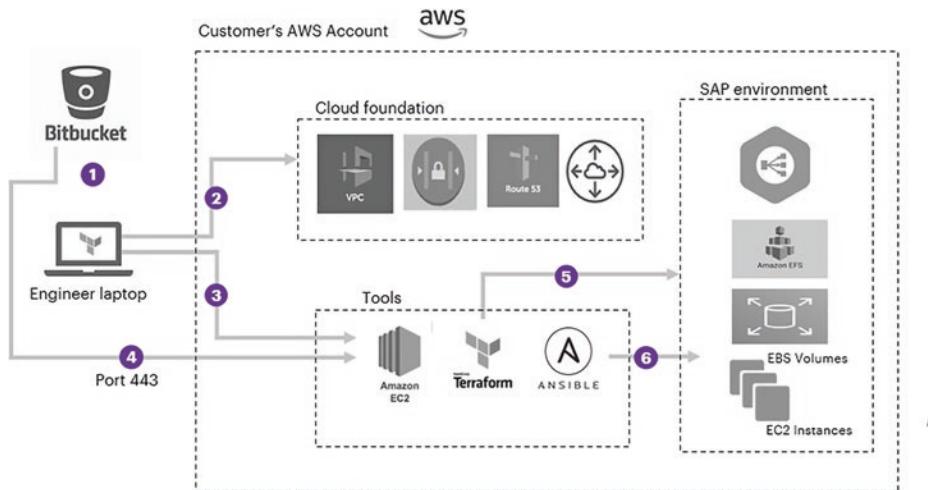
**Abb. 5.20** Manuelle Erstellung einer VM – Schritt 16



**Abb. 5.21** Manuelle Erstellung einer VM – Schritt 17

Terraform.tfvars ist die wichtigste Variablendatei, die entsprechend der jeweiligen Anforderungen aktualisiert werden muss. Main.tf ist das Hauptfunktionsmodul. Es ruft verschiedene andere Module aufruft, um die Infrastruktur aufzubauen.

Um Terraform zu installieren, öffnen Sie <https://www.terraform.io/downloads.html> (Zugriff am 20.12.2021), und wählen Sie das entsprechende Linux-Paket. Entpacken Sie die Datei, und kopieren Sie sie auf den Toolserver, auf dem Sie das Terraform-Skript ausführen möchten.



**Abb. 5.22** Automatische VM- und SAP-Bereitstellung

Öffnen Sie das Github-Repository. Gehen Sie die Terraform-Beispielskripte und -module durch, und führen Sie sie aus, um Ihre erste SAP-Infrastruktur in AWS automatisch zu erstellen (Zugriff am 20.12.2021).

<https://github.com/aws-samples/terraform-aws-sap-netweaver-on-hana>

**Ansible** wird für die Automatisierung der Betriebssystemkonfiguration und der Bereitstellung von SAP- und DB-Anwendungen verwendet. Wie Terraform, muss auch das Ansible-Skript auf dem Toolserver ausgeführt werden, um die Anwendung auf den verwalteten Servern bereitzustellen. Die yml-Datei (\*.yml) ist die Hauptkonfigurationsdatei in Ansible. Sie enthält sämtliche Aufgaben in Zusammenhang mit der Konfiguration der Infrastruktur und der Bereitstellung der S/4- und HANA-Anwendungen.

```
ansible-playbook -l<host group name>
```

Mit dem Befehl „hana-deploy.yml“ werden die Ansible-Aufgaben ausgeführt. Mit „ansible-playbook“ wird die yml-Datei aufgerufen. Der Hostgruppenname enthält die verwalteten Hosts, auf denen das Betriebssystem konfiguriert und die Anwendung bereitgestellt werden sollen. Die Hostgruppe muss mit den in der Ansible-Hostdatei enthaltenen Zielhosts aktualisiert werden. Die Hostgruppe kann mehrere Hosts enthalten, und die Konfiguration wird auf allen Hosts parallel bereitgestellt.

Für die Batchbereitstellung in der Public Cloud müssen wir nun Terraform und Ansible verwenden. Dies erspart uns viel Zeit spart, die wir sonst für manuelle Eingaben in die AWS-Konsole, die Betriebssystemkonfiguration und die SWPM/Hbdinst-Maske aufwenden müssten. Darauf hinaus können SAP/DB-Anwendungen auf mehreren Hosts parallel erstellt werden, was den Aufwand im Vergleich zur manuellen Installation erheblich reduziert.

Um Ansible zu installieren, gehen Sie auf den nachstehenden Link, und wählen Sie entsprechend Ihrer Toolserverversion RHEL oder SUSE aus. Führen Sie den Befehl wie erwähnt im Toolserver aus. Ansible wird dann installiert und ist einsatzbereit (Zugriff am 20.12.2021). <https://software.opensuse.org/download.html?project=systemsmanagement&package=ansible>

---

## 5.3 Datenspeicher

In diesem Kapitel werden die in Amazon Elastic Block Store (EBS) verfügbaren Speichertypen näher erläutert. Ausgehend von den Anforderungen des S/4-Systems werden Hinweise zum allgemeinen Speicherlayout (Stripping) sowie zu der möglichen Dimensionierung für S/4-Systeme gegeben (Datenbank und Anwendungsserver). Ferner werden kostenoptimierte Lösungen für den Betrieb nicht kritischer SAP-Systeme diskutiert, worunter beispielsweise Archivsysteme fallen, die nur für den Lesezugriff bereitstehen. Gemäß den DR-Anforderungen müssen für manche dieser Speicher bestimmte Einstellungen ausgewählt werden (Georeplikation). Auch dies wird in diesem Kapitel behandelt.

AWS bietet verschiedene Arten von Speicheroptionen, die flexibel, kostenoptimiert und einfach zu verwenden sind und sich in Haltbarkeit und Performance unterscheiden. Im vorliegenden Abschnitt gehen wir auf die unterschiedlichen, von EC2 unterstützten Speichertypen ein. Die Auswahl richtet sich nach den geschäftlichen Anforderungen und nach dem für die S/4- und HANA-Workload verwendeten Speichertypen.

Amazon Web Service bietet die folgenden Speicheroptionen:

- Elastic Block Store (EBS)
- Instance Store
- Elastic File System (EFS)
- Simple Storage Service (S3)

### 5.3.1 Amazon EBS

Elastic Block Store (EBS) bietet haltbare Blockspeichervolumes, die bestimmten EC2-Instanzen zugeordnet werden können. EBS-Volumes können in einer bestimmten Availability Zone erstellt werden. Eine zonenübergreifende Nutzung ist nicht möglich. Um die Daten im Falle von Komponentenausfällen zu schützen, repliziert EBS automatisch die Daten in der betreffenden Availability Zone. Jeder EC2-Instanz können mehrere EBS-Volumes zugeordnet werden. Umgekehrt ist es jedoch nicht möglich, dasselbe EBS-Volume mehreren EC2-Instanzen zuzuordnen. EBS-Volumes sollten eingesetzt werden, wenn die Daten schnell verfügbar sein müssen und eine langfristige Persistenz erforderlich ist. Hauptsächlich werden EBS-Volumes für Datenbank- und

Anwendungsdateisysteme verwendet. EBS hängt nicht von der Haltbarkeit der EC2-Instanz ab, sondern das EBS-Volume kann nach Ablauf der Haltbarkeit der EC2-Instanz in Funktion bleiben. EBS kann nicht über die Availability Zones hinaus erweitert werden. Es ist jedoch möglich, einen Snapshot des vorhandenen EBS zu erstellen und diesen Snapshot in eine andere Zone zu übertragen und wiederherzustellen. Auf diese Weise können Sie die Daten des EBS-Volumes in mehreren Zonen nutzen. AWS bietet verschiedene Arten von EBS-Volumes:

- **Allzweck-SSDs:** Allzweck-SSDs (gp2 und gp3) bietet einen kostengünstigen Speicher mit hoher Leistung und ist am besten für kleine bis mittelgroße Workloads geeignet. Allzweck-SSD ist von 1 GB bis 16 TB verfügbar und bietet drei IOPS pro GB, d. h. für die Bereitstellung eines 1-TB-Volumes erhalten Sie 3000 IOPS. Allzweck-SSDs werden auf der Grundlage des zugeordneten Volumens und damit unabhängig von dem verwendeten Datenvolumen abgerechnet.
- **Bereitstellung von IOPS-SSDs:** Die Bereitstellung von IOPS-SSDs (io1, io2) bietet unter allen EBS-Volumes den höchsten Durchsatz und die beste Performance. IOPS-SSDs werden eingesetzt, wenn eine durchgängig konsistente IOPS-Rate benötigt wird. Dies ist der teuerste EBS-Speichertyp. Die Volume-Größe, reichen von 4 GB bis 16 TB betragen. Die Preisgestaltung ist ähnlich wie bei Allzweck-SSDs: Bezahlte wird für das angeschlossene Volume und die bereitgestellten IOPS. IOPS-SSDs werden hauptsächlich für große Datenbank-Workloads wie HANA verwendet.
- **Magnetfestplatten-Volumes:** EBS-Magnetfestplatten (st1, sc1) sind unter allen EBS-Volumes die kostengünstigsten, allerdings bieten sie auch die schlechteste Leistung. Sie werden verwendet, wenn nur selten auf die Daten zugegriffen muss und der Zugriff sequenziell erfolgt.

EBS-Volumes können per Amazon-Verschlüsselung verschlüsselt werden. Bei dieser Art der Verschlüsselung wird bei der Erstellung des verschlüsselten Volumes und der Snapshots ein AWS-KMS-Schlüssel verwendet. Um die Datensicherheit im Ruhezustand und bei der Übertragung zwischen Host und EBS-Volume zu gewährleisten, erfolgt die Verschlüsselung auf Host-Ebene. Folgende Datentypen werden verschlüsselt:

- Daten im Ruhezustand, die sich auf dem Volume befinden
- Daten, die zwischen Volume und Instanz übermittelt werden
- Snapshots, die von dem Volume erstellt werden
- Alle aus diesen Snapshots erstellten Volumes

Zur Verschlüsselung des Datenvolumes verwendet EBS den Standardalgorithmus AES-256. EBS verschlüsselt den Datenschlüssel mit dem KMS-Schlüssel und speichert diesen auf der Festplatte. Ihr Datenschlüssel wird nie als Klartext gespeichert, sondern immer in verschlüsselter Form.

Es ist ebenfalls möglich, ein verschlüsseltes EBS-Volumen aus einem unverschlüsselten Volumen zu erstellen. Hier muss ein Snapshot eines vorhandenen unverschlüsselten EBS-Volumes erstellt werden. Dieser Snapshot ist dann zu verschlüsseln und in einem verschlüsselten EBS-Volumen wiederherzustellen. Weitere Informationen zur EBS-Verschlüsselung finden Sie in dem folgenden AWS-Standardleitfaden (Zugriff am 20.12.2021):

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/EBSEncryption.html>

### 5.3.2 Amazon EC2-Instanzspeicher

Der EC2-Instanzspeicher ist ein EBS-ähnlicher Blockspeicher, in dem die Daten aber nur vorübergehend gespeichert werden. Der Instanzspeicher ist physisch mit dem EC2-Host verbunden. Dies bedeutet, dass der gespeicherten Daten verloren gehen, wenn die EC2-Instanz beendet wird. Der EC2-Instanzspeicher eignet sich ideal für temporären Speicherbedarf wie Puffer, Cache oder Daten, die in den einzelnen Instanzen repliziert werden. Instanzspeicher kann im Swap Space verwendet werden. Nicht verwendbar hingegen ist er in SAP-Workload-Dateisystemen.

### 5.3.3 Amazon EFS

Das Amazon Elastic File System (EFS) ist ein Dateispeicher, der in EC2-Instanzen verwendet werden kann. Es handelt sich um einen gängigen Speichertyp. Sie können in EFS ein Dateisystem erstellen und dieses in mehreren EC2-Instanzen installieren. EFS kann für die gängigen Dateisysteme in SAP-HA-Umgebungen verwendet werden, darunter/ `/usr/sap/<SID>/ASCS##`, `/usr/sap/<SID>/ERS##` etc. In HANA-Scale-out-Umgebungen kann es für/hana/shared und/hana/backup verwendet werden. In dem folgenden AWS-Standardleitfaden wird die Erstellung von EFS-Dateisystemen Schritt für Schritt erläutert (Zugriff am 20.12.2021):

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/EBSEncryption.html>

### 5.3.4 Amazon S3 (Simple Storage Service)

Amazon S3 ist ein Objektspeicher, der zum Speichern und Abrufen beliebiger Datentypen und -volumen von jedem beliebigen Ort aus über das Internet nutzbar ist. Bei S3 wird nur der tatsächlich genutzte Speicher abgerechnet. S3 wird üblicherweise in folgenden Fällen verwendet:

- Speichern von Backups (DB, VM-Snapshot usw.) und Archivdaten
- Disaster Recovery

- Speichern von Inhalten, Medien und Software
- Hosting von statischen Webseiten
- Hosting von Webanwendungen etc.

S3 unterstützt unterschiedliche Anwendungsfälle. Hierfür werden verschiedene Speicherklassen angeboten, beispielsweise für Allzwecknutzung, seltenen Zugriff und Archivnutzung. Im Folgenden wird kurz auf die S3-Komponenten eingegangen.

### **Container**

S3-Container halten in S3 gespeicherte Objekte. Bei einem Container handelt es sich um eine allgemeine, also nicht nur auf das AWS-Konto bezogene Einheit. Es empfiehlt sich also, einen für alle AWS-Konten eindeutigen Containernamen zu wählen. Obwohl der Container-Namespace allgemein ist, können Sie einen auf Ihre spezifische Region begrenzten S3-Container. Dies minimiert die Latenzen und sorgt für eine schnellere Dateübertragung. Für eine bessere Reaktion bei Katastrophenszenarien können Sie auch einen Container in einer anderen Region als der primären erstellen.

### **Objekte**

Objekte sind die Dateien, die in S3-Containern gespeichert werden. Objekte können jeden beliebigen Datentyp in jedem beliebigen Format speichern. Es ist möglich, bis zu 5 TB Daten in einem einzigen Container zu speichern, wobei die Anzahl der Objekte unbegrenzt ist. Jedes Objekt besteht aus (tatsächlichen) Daten und Metadaten (Informationen über Daten).

### **Schlüssel**

Der Schlüssel ist ein eindeutiger Bezeichner für alle im S3-Container gespeicherten Objekte. Der Schlüssel kann ein Dateiname sein. Dieser sollte innerhalb des betreffenden Containers eindeutig sein und kann in mehreren Containern verwendet werden.

### **Objekt-URL**

Auf Amazon S3-Objekte kann mit einer spezifischen URL zugegriffen werden, die aus Serviceendpunkt, Containername und Objektschlüssel gebildet wird (Zugriff am 20.12.2021):

<http://testbucket.s3.amazonaws.com/test.doc>

wobei „testbucket“ der Containername und „test.doc“ Ihr Schlüssel bzw. Dateiname ist.

### **Haltbarkeit und Verfügbarkeit**

Amazon S3 bietet eine sehr hohe Haltbarkeit und eine sehr hohe Verfügbarkeit. Der Amazon S3-Standardspeicher ist für eine Haltbarkeit von 99,999.999.999 % und für eine Verfügbarkeit von 99,99 % ausgelegt. S3 erreicht diese Haltbarkeit durch die automatische Datenreplikation an allen Standorten in einer bestimmten Region. Wenn für die

nicht kritischen Daten keine hohe Haltbarkeit benötigt wird, können Sie Kosten sparen, indem Sie Reduce Redundant Storage (RRS) mit einer Haltbarkeit von 99,99 % verwenden.

### Datenkonsistenz

Amazon S3 garantiert die Datenkonsistenz, denn die Daten werden standardmäßig auf verschiedenen Servern oder Standorten innerhalb der Region repliziert. Die Ausführung von Änderungen kann etwas Zeit beanspruchen. Dadurch können veraltete Daten übermittelt werden. Insbesondere gilt dies, wenn Sie neue Daten in ein S3-Objekt schreiben, der nachfolgende GET-Befehl aber noch die alten Daten zurückgibt. Der GET-Vorgang ist aber immer konsistent und liefert entweder nur alte oder nur neue Daten.

### Zugangskontrolle

Für den Zugriff auf S3-Container oder die Objektebene verwendet Amazon S3 die Amazon Access Control List (ACL) und IAM. WRITE-, READ-, und FULL-CONTROL-Rechte können auf Container- oder Objektebene gewährt werden.

### Speicherklasse

Um eine kosteneffiziente Strategie für die Datensicherung und -archivierung festzulegen, müssen Sie die verschiedenen in S3 verfügbaren Speicherklassen kennen. Grob gesprochen sind drei Speicherklassen verfügbar:

- Amazon S3 Standard: Der Amazon S3 Standard bietet die höchste Haltbarkeit, die höchste Verfügbarkeit, geringe Latenzzeiten und eine hohe Performance bei der Datenübertragung zwischen EC2 und S3. Amazon S3 Standard ist sehr nützlich für kurz- oder langfristig gespeicherte häufig genutzte Daten.
- Amazon S3 Standard IA: Standard Infrequent Access (IA) bietet in Bezug auf Haltbarkeit, Verfügbarkeit und Latenzzeiten dasselbe wie die Standardspeicherklasse. Diese Speicherklasse wird für Daten verwendet, auf die weniger häufig zugegriffen wird. Die Daten müssen mindestens 30 Tage lang gespeichert werden. Standard Infrequent Access eignet sich also für langfristig gespeicherte Daten, auf die weniger häufig zugegriffen werden muss. Die Kosten sind deutlich geringer als bei der Standard-Speicherklasse.
- Amazon Glacier: Die Amazon-Glacier-Speicherklasse bietet die kostengünstigste Möglichkeit, Archivdaten oder langfristige Backups zu speichern. Darüber hinaus bietet Amazon Glacier Haltbarkeit und Datensicherheit. Das Abrufen von Daten aus Glacier kann mehrere Stunden beanspruchen, kann aber gegen Aufpreis beschleunigt werden. Auf Basis der oben genannten drei Speicherklassen können Sie nun Ihren Cloud-Speicher planen. Die wichtigsten Kriterien sind dabei die benötigte Zugriffs-frequenz, die Aufbewahrungsdauer und selbstverständlich die Kosten.

Einzelheiten finden Sie in der folgenden AWS KBA (Zugriff am 20.12.2021):

<https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/AmazonS3.html>

### 5.3.5 Speicherlayout für S/4HANA-Systeme

SAP HANA ist eine In-Memory-Datenbank. Dies bedeutet, dass die meisten Daten im Speicher gespeichert und verarbeitet werden. Um die Daten zu schützen, müssen sie in einem Persistenzspeicher abgelegt werden. Für einen optimalen Datendurchsatz und eine optimale Performance muss das Volume der Hana-Daten- und -Protokolle mithilfe der SAP-zertifizierten Speicheroptionen für HANA-Workloads konfiguriert werden. Für SAP-HANA-Workloads sollten SAP-zertifizierte EBS-Volumes (gp2, gp3) gewählt und IOPS-SSD (io1, io2) bereitgestellt werden.

Für nicht kritische oder mittelgroße geschäftskritische Systeme können gp2 und gp3 verwendet werden, da Allzweck-SSDs ein ausgewogenes Preis-Leistungs-Verhältnis bieten. io1 und io2 bieten die stärkste Performance und werden für die meisten geschäftskritischen HANA-Workloads empfohlen. Die folgende Übersicht illustriert die Festlegung der SAP-HANA-Daten- und Protokoll-Volumes für die einzelnen Instanztypen. gp2 oder gp3 können für/hana/shared volume verwendet werden. Amazon EFS (Elastic File System) ist für/hana/shared und/hana/backup in der HANA-Scale-out-Umgebung zu verwenden.

AWS hat vor kurzem ein EBS-Volumen vom Typ gp3 auf den Markt gebracht. Es wird nun empfohlen, gp3 zu verwenden, da gp3 im Vergleich zu EBS-Volumes vom Typ gp2 einen besseren Durchsatz bzw. eine bessere Performance und Kosteneffizienz bietet.

gp2- und gp3-Volumes für **produktive HANA-Workloads** (Tab. 5.2)

gp2- und gp3-Volumes lediglich für **nicht produktive Workloads** (Tab. 5.3)

IOPS-Volume-Bereitstellung (io1, io2) für **produktive Workloads** (Tab. 5.4)

IOPS-Volume-Bereitstellung (io1, io2) lediglich für **nicht produktive Workloads** (Tab. 5.5)

Siehe nachstehenden AWS-Standardleitfaden (Zugriff am 20.12.2021):

<https://docs.aws.amazon.com/sap/latest/sap-hana/hana-ops-storage-config.html>

---

## 5.4 PAYG vs. Commitment – Rechner

Obwohl die Public Cloud ein großes Einsparpotenzial bietet, bleibt für viele Kunden die Beantwortung der Frage schwierig, ob sie sich für eine gewisse Zeit an eine Workload binden oder ob sie flexibel bleiben möchten. Diese Entscheidung hat starke finanzielle Auswirkungen und wird in diesem Kapitel erörtert. Wir werden auch den Rechner erläutern, mit dem sich ausrechnen lässt, ab wann die jeweilige Variante vorteilhaft ist. In AWS können EC2-Instanzen je nach Bedarf in der Kategorie „On-Demand“ oder in der Kategorie „Reserve Instance“ erworben werden.

**Tab. 5.2** Instanztypen für produktive SAP-Workloads

Instanztyp	Speicher (GB)	vCPU	GP SSD (gp2) für/ hana/data (mit LVM- Striping)	gp2- Speicher für /hana/log (mit LVM- Striping)	GP SSD (gp3) für/ hana/data (mit LVM- Striping)	gp3- Speicher für /hana/log (mit LVM- Striping)
u-24tb1.metal	24.576	448	6*4800 (GB)	2*300 (GB)	2*14400 (GB)	1*512 (GB)
u-18tb1.metal	18.432	448	6*3600 (GB)	2*300 (GB)	2*10800 (GB)	1*512 (GB)
u-12tb1.112xlarge	12.288	448	6*2400 (GB)	2*300 (GB)	2*7200 (GB)	1*512 (GB)
u-12tb1.metal	12.288	448	6*2400 (GB)	2*300 (GB)	2*7200 (GB)	1*512 (GB)
u-9tb1.112xlarge	9216	448	6*1800 (GB)	2*300 (GB)	2*5400 (GB)	1*512 (GB)
u-9tb1.metal	9216	448	6*1800 (GB)	2*300 (GB)	2*5400 (GB)	1*512 (GB)
u-6tb1.112xlarge	6144	448	6*1200 (GB)	2*300 (GB)	2*3600 (GB)	1*512 (GB)
u-6tb1.56xlarge	6144	224	6*1200 (GB)	2*300 (GB)	2*3600 (GB)	1*512 (GB)
u-6tb1.metal	6144	448	6*1200 (GB)	2*300 (GB)	2*3600 (GB)	1*512 (GB)
x1e.32xlarge	3902	128	3*1600 (GB)	2*300 (GB)	2*2400 (GB)	1*512 (GB)
x1.32xlarge	1952	128	3*800 (GB)	2*300 (GB)	2*1200 (GB)	1*512 (GB)
x1.16xlarge	976	64	3*400 (GB)	2*300 (GB)	2*1200 (GB)	1*512 (GB)
r5.metal	768	96	3*400 (GB)	2*300 (GB)	1*920 (GB)	1*512 (GB)
r5b.metal	768	96	3*400 (GB)	2*300 (GB)	1*920 (GB)	1*512 (GB)
r5.24xlarge	768	96	3*400 (GB)	2*300 (GB)	1*920 (GB)	1*512 (GB)
r5b.24xlarge	768	96	3*400 (GB)	2*300 (GB)	1*920 (GB)	1*512 (GB)
r5.16xlarge	512	64	3*225 (GB)	2*300 (GB)	1*615 (GB)	1*512 (GB)

(Fortsetzung)

**Tab. 5.2** (Fortsetzung)

Instanztyp	Speicher (GB)	vCPU	GP SSD (gp2) für/ hana/data (mit LVM- Striping)	gp2- Speicher für /hana/log (mit LVM- Striping)	GP SSD (gp3) für/ hana/data (mit LVM- Striping)	gp3- Speicher für /hana/log (mit LVM- Striping)
r5b.16xlarge	512	64	3*225 (GB)	2*300 (GB)	1*615 (GB)	1*256 (GB)
r4.16xlarge	488	64	3*225 (GB)	2*300 (GB)	1*585 (GB)	1*256 (GB)
r5.12xlarge	384	48	3*225 (GB)	2*300 (GB)	1*460 (GB)	1*192 (GB)
r5b.12xlarge	384	48	3*225 (GB)	2*300 (GB)	1*460 (GB)	1*192 (GB)
r5.8xlarge	256	32	3*225 (GB)	2*300 (GB)	1*320 (GB)	1*128 (GB)
r5b.8xlarge	256	32	3*225 (GB)	2*300 (GB)	1*320 (GB)	1*128 (GB)
r4.8xlarge	244	32	3*225 (GB)	2*300 (GB)	1*300 (GB)	1*128 (GB)

**Tab. 5.3** Instanztypen für nicht-produktive SAP-Workloads

Instanztyp	Speicher (GB)	vCPU	GP SSD (gp2) für/ hana/data (mit LVM- Striping)	gp2-Speicher für /hana/log (mit LVM- Striping)	GP SSD (gp3) für/ hana/data (mit LVM- Striping)	gp3- Speicher für /hana/log (mit LVM- Striping)
x1e.4xlarge	488	16	3*225 (GB)	2*175 (GB)	1*585 (GB)	1*245 (GB)
x1e.2xlarge	244	8	3*225 (GB)	2*175 (GB)	1*295 (GB)	1*125 (GB)
x1e.xlarge	122	4	3*225 (GB)	2*175 (GB)	1*150 (GB)	1*64 (GB)
r5.4xlarge	128	16	3*225 (GB)	2*175 (GB)	1*150 (GB)	1*64 (GB)
r5b.4xlarge	128	16	3*225 (GB)	2*175 (GB)	1*150 (GB)	1*64 (GB)
r5.2xlarge	64	8	3*225 (GB)	2*175 (GB)	1*80 (GB)	1*32 (GB)
r5b.2xlarge	64	8	3*225 (GB)	2*175 (GB)	1*80 (GB)	1*32 (GB)
r4.4xlarge	122	16	3*225 (GB)	2*175 (GB)	1*150 (GB)	1*64 (GB)
r4.2xlarge	61	8	3*225 (GB)	2*175 (GB)	1*80 (GB)	1*32 (GB)

**Tab. 5.4** IO Volumen für produktive SAP-Workloads

Instanztyp	Speicher (GB)	vCPU	Hana-Daten (mit LVM-Striping)	/hana/log (mit LVM-Striping)
u-24tb1.metal	24.576	448	6*4800 (GB)	1*525 (GB)
u-18tb1.metal	18.432	448	6*3600 (GB)	1*525 (GB)
u-12tb1.112xlarge	12.288	448	6*2400 (GB)	1*525 (GB)
u-12tb1.metal	12.288	448	6*2400 (GB)	1*525 (GB)
u-9tb1.112xlarge	9.216	448	6*1800 (GB)	1*525 (GB)
u-9tb1.metal	9.216	448	6*1800 (GB)	1*525 (GB)
u-6tb1.112xlarge	6.144	448	6*1200 (GB)	1*525 (GB)
u-6tb1.56xlarge	6.144	224	6*1200 (GB)	1*525 (GB)
u-6tb1.metal	6.144	448	6*1200 (GB)	1*525 (GB)
x1e.32xlarge	3.902	128	3*1600 (GB)	1*525 (GB)
x1.32xlarge	1.952	128	3*800 (GB)	1*525 (GB)
x1.16xlarge	976	64	1*1200 (GB)	1*525 (GB)
r5.metal	768	96	1*1200 (GB)	1*525 (GB)
r5b.metal	768	96	1*1200 (GB)	1*525 (GB)
r5.24xlarge	768	96	1*1200 (GB)	1*525 (GB)
r5b.24xlarge	768	96	1*1200 (GB)	1*525 (GB)
r5.16xlarge	512	64	1*600 (GB)	1*260 (GB)
r5b.16xlarge	512	64	1*600 (GB)	1*260 (GB)
r4.16xlarge	488	64	1*600 (GB)	1*260 (GB)
r5.12xlarge	384	48	1*600 (GB)	1*260 (GB)
r5b.12xlarge	384	48	1*600 (GB)	1*260 (GB)
r5.8xlarge	256	32	1*300 (GB)	1*260 (GB)
r5b.8xlarge	256	32	1*300 (GB)	1*260 (GB)
r4.8xlarge	244	32	1*300 (GB)	1*260 (GB)

#### 5.4.1 On-Demand-Instanzen

Bei On-Demand-Instanzen wird nur die Zeit, in der die EC2-Instanzen tatsächlich laufen. Wenn die EC2-Instanzen angehalten werden, muss nicht für die Rechenleistung bezahlt werden, sondern nur für den zu den Instanzen gehörenden Speicher. Wenn Instanzen wie Sandbox, Training, nicht produktive Instanzen etc. nur begrenzt genutzt werden, ist dies sehr nützlich und kosteneffizient. Wenn die Nutzungszeiten bekannt sind, können Start und Stopp automatisch eingeplant werden. Zum Beispiel kann die Instanz automatisch vor Beginn der Geschäftszeit gestartet und nach Ende der

**Tab. 5.5** IO Volumen für nicht-produktive SAP-Workloads

Instanztyp	Speicher (GB)	vCPU	Hana-Daten (mit LVM-Striping)	/hana/log (mit LVM-Striping)
x1e.4xlarge	488	16	1*600 (GB)	1*260 (GB)
x1e.2xlarge	244	8	1*300 (GB)	1*260 (GB)
x1e.xlarge	122	4	1*300 (GB)	1*260 (GB)
r5.4xlarge	128	16	1*300 (GB)	1*260 (GB)
r5b.4xlarge	128	16	1*300 (GB)	1*260 (GB)
r5.2xlarge	64	8	1*300 (GB)	1*260 (GB)
r5b.2xlarge	64	8	1*300 (GB)	1*260 (GB)
r4.4xlarge	122	16	1*300 (GB)	1*260 (GB)
r4.2xlarge	61	8	1*300 (GB)	1*260 (GB)

Geschäftszeit gestoppt werden. On-Demand-Instanzen sind sehr nützlich und kostengünstig, wenn keine langfristigen Verpflichtungen bestehen.

Die Anzahl der laufenden On-Demand-Instanzen ist pro AWS-Konto und -Region beschränkt. Das Limit ist abhängig von der Anzahl der vCPUs. Das Limit für die vCPUs pro AWS-Region und -Konto kann erhöht werden. Wenden Sie sich hierfür an das AWS-Supportportal.

### 5.4.2 Reserved Instances

Im Vergleich zu den On-Demand-Instanzkosten sind Reserved Instances deutlich billiger. Allerdings ist eine langfristige Verpflichtung notwendig, die zwischen einem und drei Jahren liegen kann. Bei einer dreijährigen Bindung ist der Rabatt höher. Reserved Instances werden nach dem Ablauf nicht automatisch verlängert, sondern werden in On-Demand-Instanzen umgewandelt. Sollen die Reserved Instances weiterverwendet werden, müssen sie erneut erworben werden. Diese Option ist sehr nützlich und kosten-effizient für produktive Workloads, bei denen die EC2-Instanz das gesamte Jahr über rund um die Uhr laufen muss.

Der Preis für Reserved Instances richtet sich nach vier Merkmalen: Instanztyp (klein oder groß), Kaufregion, Tenancy (Standard oder dediziert) und Plattform (Unix oder Windows).

Reserved Instances sind in zwei Ausführungen verfügbar: **Standard** und **Convertible**. Standard Reserved Instances sind wesentlich preisgünstiger, können aber nicht ausgetauscht, sondern lediglich geändert werden. Convertible Reserved Instances sind im Vergleich zu Standard Reserved Instances weniger preisgünstig, bieten dafür aber den Vorteil, gegen andere Convertible Reserved Instances ausgetauscht und geändert werden zu können.

Eine **Reserved Instance ändern** zu können bedeutet, dass der Instanztyp innerhalb derselben Familie geändert werden kann. Der Instanztyp kann von m4\* auf m5\* geändert werden. Dies setzt jedoch voraus, dass die Größe des Instanz-Footprints identisch bleibt. Der Instanztyp m4\* kann also nicht zu c\* geändert werden. Nur die Linux- bzw. Unix-Plattform kann geändert werden.

Eine **Convertible Reserved Instance auszutauschen** bedeutet, dass der Instanztyp gegen einen Instanztyp einer Familie mit demselben oder mit einem höheren Footprint als der ursprüngliche Instanztyp ausgetauscht werden kann.

### 5.4.3 Zahlungsoptionen

Bei Reserved Instances fallen unabhängig von der Nutzung Zahlungen an. Es sind drei Zahlungsoptionen verfügbar.

#### Vollständige Vorauszahlung

Der gesamte Betrag ist vor dem Betrieb der Instanz fällig. Während der Laufzeit fallen keine weiteren Zahlungen an. Mit dieser Zahlungsoption erhalten Sie den höchstmöglichen Rabatt.

#### Teilweise Vorauszahlung

Ein Teil des Gesamtbetrags ist vor dem Beginn der Laufzeit zu entrichten, auf den Restbetrag wird der diskontierte Stundensatz angewendet.

#### Keine Vorauszahlung

Keine Vorauszahlung, in Rechnung gestellt wird der diskontierte Stundensatz.

### 5.4.4 AWS-Rechner

Die Kosten können anhand der Stückliste des jeweiligen Projekts mit dem AWS-Rechner berechnet werden. Hier erläutern wie, wie der AWS-Rechner im normalen Geschäft eingesetzt werden kann.

Öffnen Sie die URL (Zugriff am 20.12.2021) <https://calculator.s3.amazonaws.com/index.html>.

Auf der linken Seite sind alle AWS-Services für die Kostenberechnung aufgeführt. Oben sind die beiden Registerkarten „Services“ und „Estimate of your Monthly Bill“ zu sehen. Im vorliegenden Beispiel schätzen wir die Kosten für eine EC2-Instanz und ein On-Demand-EBS-Volume mit Reserved Instances mit einer Laufzeit von drei Jahren (Abb. 5.23).

In dem Beispiel gibt es vier Instanzen aus der r\*-Familie und ein EBS-Volume mit 2.048 GB Provisioned IOPS SSD (io2).

**Simple Monthly Calculator depreciation update:** We appreciate your continuous feedback regarding the [AWS Pricing Calculator](#). The Simple Monthly Calculator's depreciation date is delayed to ensure the features requested from our customers are available in the AWS Pricing Calculator. We will continue to add services to the AWS Pricing Calculator to guarantee parity with the Simple Monthly Calculator. If you have any feedback, contact us by using the [Feedback](#) link in the AWS Pricing Calculator.

**FREE TIER:** New Customers get free usage tier for first 12 months

**Services** Estimate of your Monthly Bill (\$ 0.00)

Choose region: US East (N. Virginia) Inbound Data Transfer is Free and Outbound Data Transfer is 1 GB free per region per month

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. It is designed to make web-scale computing easier for developers. Amazon Elastic Block Store (EBS) provides persistent storage to Amazon EC2 Instances. [Clear Form](#)

Newer versions of the EC2 calculators are available: [Amazon EC2](#), [EC2 Dedicated Host](#), [Elastic Graphics](#), [Elastic IP](#)

**Compute: Amazon EC2 Instances:**

Description	Instances	Usage	Type	Billing Option	Monthly Cost
S4/HANA EC2	2	100 % Utilized/Month	Red Hat Enterprise Linux on r5.8xlarge	On-Demand (No Cor)	\$ 3141.76

[Add New Row](#)

**Compute: Amazon EC2 Dedicated Hosts:**

Description	Number of Hosts/Usage	Type	Billing Option

[Add New Row](#)

**Common Customer Samples**

- Free Website on AWS
- AWS Elastic Beanstalk Default
- Marketing Web Site
- Large Web Application (All On-Demand)
- Media Application

**Abb. 5.23** AWS Calculator – Schritt 1

**FREE TIER:** New Customers get free usage tier for first 12 months

**Services** Estimate of your Monthly Bill (\$ 6796.59)

Choose region: US East (N. Virginia) Inbound Data Transfer is Free and Outbound Data Transfer is 1 GB free per region per month

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. It is designed to make web-scale computing easier for developers. Amazon Elastic Block Store (EBS) provides persistent storage to Amazon EC2 Instances. [Clear Form](#)

Newer versions of the EC2 calculators are available: [Amazon EC2](#), [EC2 Dedicated Host](#), [Elastic Graphics](#), [Elastic IP](#)

**Compute: Amazon EC2 Instances:**

Description	Instances	Usage	Type	Billing Option	Monthly Cost
S4/HANA EC2	2	100 % Utilized/Month	Red Hat Enterprise Linux on r5.8xlarge	On-Demand (No Cor)	\$ 3141.76
S4/HANA EC2	2	100 % Utilized/Month	Red Hat Enterprise Linux on r5.8xlarge	On-Demand (No Cor)	\$ 3141.76

[Add New Row](#)

**Compute: Amazon EC2 Dedicated Hosts:**

Description	Number of Hosts/Usage	Type	Billing Option

[Add New Row](#)

**Storage: Amazon EBS Volumes:**

Description	Volumes	Volume Type	Storage	IOPS	Baseline Throughput	Snapshot Storage
S4/HANA EBS	4	Provisioned IOPS SSD (io2)	2048 GB	50	12.5 MB/sec	10000 GB-month of Storage

[Add New Row](#)

**Compute: Amazon Elastic Graphics:**

Description	Number of Elastic Graphics/Usage	Elastic Graphics Size and Memory

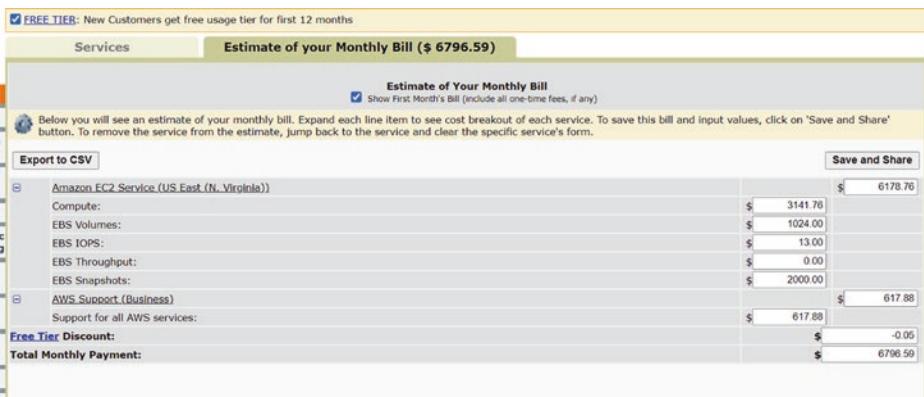
[Add New Row](#)

**Abb. 5.24** AWS Calculator – Schritt 2

### Kosten der On-Demand-Instanz

Klicken Sie im Bereich „Compute: Amazon EC2 Instances“ (Berechnung der EC2-Instanzen) auf „Add new row“ (neue Zeile hinzufügen). Erfassen Sie die Angaben gemäß Ihren Anforderungen. Erfassen Sie analog dazu die Angaben in „Storage: Amazon EBS Volumes“ (Speicherbedarf der EBS-Volumes) (Abb. 5.24).

Die monatliche Rechnung für zwei r5.8xlarge-Instanzen und vier IOPS SSD (io2) Volumes mit 2.048 GB wird auf \$ 6.796,59 geschätzt. In der Registerkarte „Estimate of



**Abb. 5.25** AWS Calculator – Schritt 3

The screenshot shows the AWS Calculator interface. At the top, it says "Choose region: US East (N. Virginia)" and "Inbound Data Transfer is Free and Outbound Data Transfer is 1 GB free per region per month". It also mentions "Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. It is designed to make web-scale computing easier for developers. Amazon Elastic Block Store (EBS) provides persistent storage to Amazon EC2 instances." There is a "Clear Form" button. Below that, it says "Newer versions of the EC2 calculators are available: [Amazon EC2](#), [EC2 Dedicated Host](#), [Elastic Graphics](#), [Elastic IP](#)".

**Compute: Amazon EC2 Instances:**

Description	Instances	Usage	Type	Billing Option	Monthly Cost
S4HANA EC2	2	100 % Utilized/Month	Red Hat Enterprise Linux on r5.8xlarge	3 Yr All Upfront Rese	\$ 0.00

**Add New Row**

**Compute: Amazon EC2 Dedicated Hosts:**

Description	Number of Hosts	Usage	Type	Billing Option
Add New Row				

**Storage: Amazon EBS Volumes:**

Description	Volumes	Volume Type	Storage	IOPS	Baseline Throughput	Snapshot Storage
S4HANA EBS	4	Provisioned IOPS SSD (io2)	2048 GB	50	12.5 MBs/sec	10000 GB-month of Storage

**Add New Row**

**Compute: Amazon Elastic Graphics:**

**Abb. 5.26** AWS Calculator – Schritte 4

your Monthly Bill“ (Schätzung der monatlichen Kosten) sind die Kosten aufgeschlüsselt (Abb. 5.25, 5.26, 5.27, 5.28).

Hier haben wir Reserved Instances mit dreijähriger Laufzeit und mit vollständiger Vorauszahlung gewählt. Die monatlichen Kosten reduzieren sich im Vergleich zu On-Demand-Instanzen von \$ 6.796 auf \$ 3.340. Auf diese Weise können Sie eine an die jeweiligen Anforderungen angepasste Kostenschätzung erhalten.

## 5.5 Sicherheit bei AWS

Wenn S/4-Systeme in AWS platziert werden, müssen auch Sicherheitsmaßnahmen ergriffen werden. Das reicht von der Netzwerksicherung über die Sicherheitsgruppen bis hin zum Sicherheits-Hub. Im vorliegenden Kapitel befassen wir uns mit der Einführung und Nutzung von Sicherheitsmaßnahmen, die in der Cloud zur Verfügung stehen.

**Services**      **Estimate of your Monthly Bill (\$ 3340.65)**

**Estimate of Your Monthly Bill**  
 Show First Month's Bill (include all one-time fees, if any)

Below you will see an estimate of your monthly bill. Expand each line item to see cost breakout of each service. To save this bill and input values, click on "Save and Share" button. To remove the service from the estimate, jump back to the service and clear the specific service's form.

**Export to CSV**      **Save and Share**

Amazon EC2 Service (US East (N. Virginia))	\$ 49711.00
Compute:	\$ 0.00
EBS Volumes:	\$ 1024.00
EBS IOPS:	\$ 13.00
EBS Throughput:	\$ 0.00
EBS Snapshots:	\$ 2000.00
Reserved Instances (one-time fee):	\$ 46674.00
AWS Support (Business)	\$ 3779.78
Support for all AWS services:	\$ 303.70
Support for Reserved Instances (one-time fee):	\$ 3476.08
<b>Free Tier Discount:</b>	\$ -0.05
<b>Total One-Time Payment:</b>	\$ 50150.08
<b>Total Monthly Payment:</b>	\$ 3340.65

**Abb. 5.27** AWS Calculator – Schritt 5

**FREE TIER:** New Customers get free usage tier for first 12 months

**Services**      **Estimate of your Monthly Bill (\$ 3340.65)**

**Choose region:** US East (N. Virginia)      Inbound Data Transfer is Free and Outbound Data Transfer is 1 GB per region per month

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. It is designed to make web-scale computing easier for developers. Amazon Elastic Block Store (EBS) provides persistent storage to Amazon EC2 Instances. [Clear Form](#)

Newer versions of the EC2 calculators are available: [Amazon EC2](#), [EC2 Dedicated Host](#), [Elastic Graphics](#), [Elastic IP](#)

**Compute: Amazon EC2 Instances:**

Description	Instances	Usage	Type	Billing Option	Monthly Cost
S4HANA EC2	2	100 % Utilized/Month	Red Hat Enterprise Linux on r5.8xlarge	3 Yr All Upfront Resc.	\$ 0.00
<a href="#">Add New Row</a>					

**Compute: Amazon EC2 Dedicated Hosts:**

Description	Number of Hosts	Usage	Type	Billing Option
<a href="#">Add New Row</a>				

**Abb. 5.28** AWS Calculator – Schritt 6

Sicherheit spielt in der IT-Branche eine zentrale Rolle. Sie steht beim Betrieb eines jeden SAP-Systems in AWS im Mittelpunkt der Überlegungen, des Designs und der Implementierung. Die Cloud-Sicherheit von AWS ist mehr oder weniger mit On-Premises-Sicherheit zu vergleichen. Der größte Unterschied ist, dass die zugrundeliegende Hardware und Netzwerksicherheit bei Cloud-Sicherheit für die Kunden keine Rolle spielen, denn diese werden von AWS verwaltet und auch nicht in Rechnung gestellt. AWS Security arbeitet mit einem gemeinsam genutzten Modell, in dem sich AWS um die Sicherheit der zugrundeliegenden Infrastrukturen wie Computing, Storage, globales Netzwerk usw. kümmert und der Kunde für die in der Cloud bereitgestellten Elemente kümmert, also beispielsweise um das Betriebssystem, die SAP-Anwendung, das Netzwerk, die Einrichtung der Firewall etc.

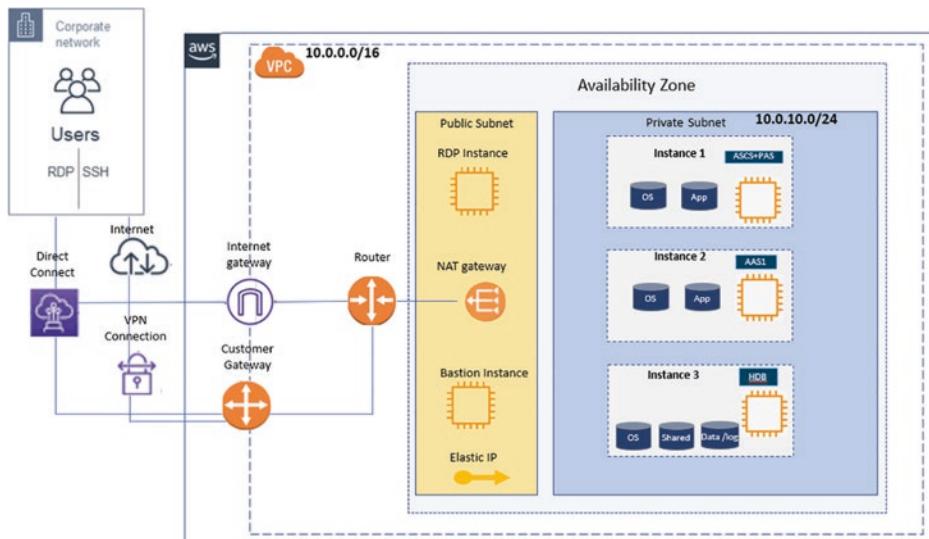
Die AWS-Konten und -Ressourcen können mit verschiedenen von AWS bereitgestellten Sicherheitsfunktionen geschützt werden, darunter AWS-Anmelde Daten, AWS

Identity und Zugriffsverwaltung (IAM), verschlüsselte Datenübertragung mit HTTPS-Endpunkten.

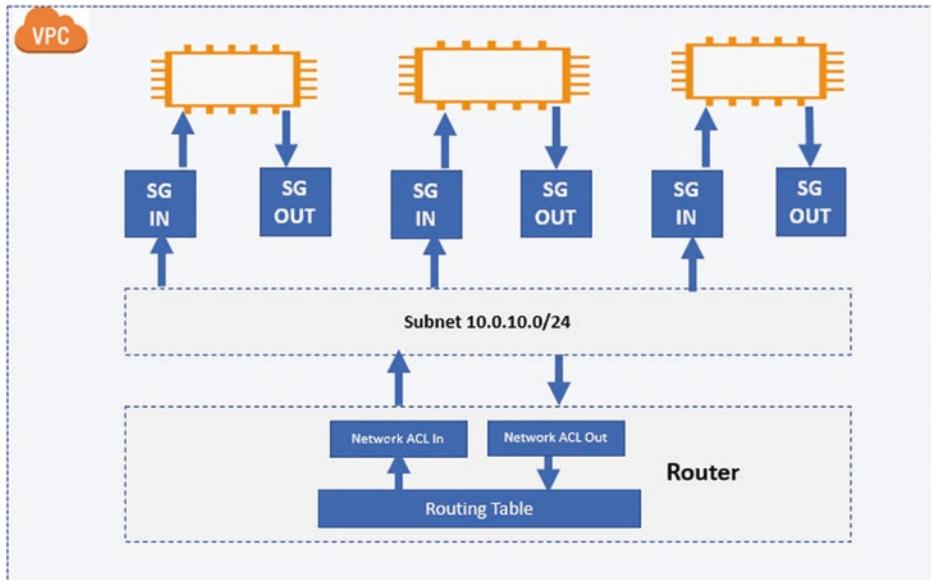
Sie können verschiedene Arten von Anmeldedaten verwenden, um Ihre AWS-Ressourcen vor unbefugtem Zugriff zu schützen, z. B. Passwörter, digitale Signaturen und Zertifikate, Schlüsselpaare und Multi-Faktor-Authentifizierung (MFA). AWS bietet nicht nur Kontosicherheit, sondern auch integrierte Sicherheit in jeder Infrastrukturschicht. Um die Unternehmensdaten vor Bedrohungen zu schützen, stehen auch auf Serviceebene Sicherheitsfunktionen zur Verfügung. Dies umfasst die EC2-Instanzen, Gastbetriebssysteme, EBS-Volumes, den Speicher (S. 3, EFS, FSx), das Netzwerk etc. Die Netzwerksicherheit ist ein sehr wichtiger Aspekt beim Design der SAP-Architektur in AWS. Hier wird erörtert, wie Sie Ihr Netzwerk vor unerwünschtem Datenverkehr schützen können.

### 5.5.1 Amazon Virtual Private Cloud (VPC)

Bei der Bereitstellung von Instanzen in AWS muss eine VPC erstellt werden. Die VPC gehört zu den zugrundeliegenden Clouds. VPCs bieten eine isolierte Umgebung in AWS, die von der Außenwelt, den anderen Unternehmen in AWS. Sofern VPC-Peering nicht aktiviert ist, sind VPCs sogar von den anderen VPCs innerhalb desselben AWS-Kontos isoliert. VPCs ist ein IP-Bereich zugewiesen, und S/4-EC2-Instanzen können mit privaten IP-Adressen gestartet werden, die in dem jeweiligen VPC-IP-Bereich liegen. Zur Kontrolle des ein- und des ausgehenden Datenverkehrs müssen nun ein Subnetz, eine Sicherheitsgruppe und eine Netzwerkzugriffskontrollliste (Network Access Control List – ACL) erstellt werden (Abb. 5.29 und 5.30).



**Abb. 5.29** Netzwerkarchitektur Amazon VPC



**Abb. 5.30** Netzwerksicherheitsarchitektur

In den obigen Diagrammen ist der Datenverkehr im Netzwerk dargestellt. Beim Start einer EC2-Instanz in AWS starten wird diese mit einer öffentlichen IP-Adresse verknüpft. Mit VPCs kann eine innerhalb von AWS isolierte Umgebung mit privaten IP-Adressen eingerichtet werden, die wiederum den EC2-Instanzen zugeordnet sind. Um die Umgebung abzusichern, müssen sowohl für den eingehenden als auch für den ausgehenden Datenverkehr an und von EC2-Instanzen entsprechende Firewalls für die VPCs eingerichtet werden. Sie können den Datenverkehr auf IP-Adressen und Ports beschränken. Die Firewalls werden nicht vom Gastbetriebssystem verwaltet, sondern können nur über die VPC-APIs geändert werden. Das Routing, die Sicherheitsgruppen und sonstige Funktionen können über API-Aufrufe nur mit dem Sicherheitsschlüssel erstellt, gelöscht oder geändert werden. Ohne Sicherheitsschlüssel können keine API-Aufrufe durchgeführt werden. Es empfiehlt sich, stets SSL-verschlüsselte API-Endpunkte zu verwenden.

## 5.5.2 Subnetz- und Routingtabellen

Die nächste Sicherheitsebene ist das Subnetz. Um die EC2-Instanzen in demselben CIDR-Block zu starten, können in derselben VPC ein oder mehrere Subnetze erstellt werden. Beispielsweise können ein Subnetz für den produktiven und ein Subnetz für nicht produktiven Datenverkehr erstellt werden. Alle produktiven EC2-Instanzen werden dann im produktiven Subnetz, alle nicht produktiven EC2-Instanzen im nicht produktiven Subnetz bereitgestellt. Auf diese Weise kann der CIDR-Block separiert

werden, und es ist möglich, den Datenverkehr von und zu den produktiven und nicht produktiven Systemen detaillierter zu kontrollieren.

### 5.5.3 Netzwerkzugriffskontrollliste (ACL)

Die Netzwerk-ACL ist die nächste Sicherheitsebene in der VPC. Die Netzwerk-ACL kann so konfiguriert werden, dass sie den ein- und ausgehenden Datenverkehr aus dem Subnetz kontrolliert. Die ACL ist genauso zu behandeln wie eine Firewall mit Quell-/Ziel-IPS, Serviceports und IP-Protokoll. Netzwerk-ACL arbeiten auf Subnetzebene und sind zustandslos, d. h. standardmäßig ist keine Antwort auf die Anfragen erlaubt, sondern dies muss explizit in den Regeln für den ein- und den ausgehenden Datenverkehr festgelegt werden. Bei der Genehmigung oder Ablehnung der Anfrage hat die Regel mit der niedrigsten Nummer die höchste Priorität. Beispiel: Wenn dieselbe Anfrage in Regel 100 genehmigt und in Regel 200 abgelehnt wird, genehmigt die Netzwerk-ACL die Anfrage.

### 5.5.4 Sicherheitsgruppen (SG)

Sicherheitsgruppen fungieren als virtuelle Firewall auf EC2-Instanzebene, um den ein- und ausgehenden Datenverkehr zu und von der Instanz zu kontrollieren. Da dies nicht auf Subnetzebene abläuft, können für verschiedene Instanzen in demselben Subnetz und in derselben VPC unterschiedliche Sicherheitsregeln konfiguriert werden. Wird keine Sicherheitsgruppe auf Instanzebene angegeben, wird die Standardsicherheitsgruppe für VPC verwendet.

Die Sicherheitsgruppe ist zustandsbehaftet. Wenn eine ausgehende Anfrage gesendet wird, gelangt die Antwort auf diese Anfrage also unabhängig von der Eingangsregel zu der Instanz. Bei eingehenden Anfragen erfolgt die Antwort unabhängig von den in der Sicherheitsgruppe festgelegten Ausgangsregeln. Jede erstellte SG umfasst standardmäßig eine Ausgangsregel, die den gesamten ausgehenden Datenverkehr zulässt. Die Ausgangsregeln können geändert werden, und es ist auch möglich, keine Eingangsregeln festzulegen. Der gesamte eingehende Datenverkehr wird in diesem Fall abgelehnt. Mit SG können also keine Instanzen in demselben Subnetz miteinander kommunizieren. Um Datenverkehr zuzulassen, eingehenden Datenverkehr erstellt werden. SGs sind mit Netzwerkschnittstellen (NICs) verbunden.

### 5.5.5 Virtual Private Gateway

Mit Virtual Private Gateway kann eine sichere Verbindung zwischen der Amazon VPC und einem anderen Netzwerk hergestellt werden. Von lokalen Gateway-Geräten aus kann eine Netzwerkverbindung zum virtuellen Private Gateway hergestellt werden.

### 5.5.6 Internet Gateway

Internet Gateway ist mit der VPC verbunden und ermöglicht eine direkte Verbindung zu Amazon S3, sonstigen Diensten und dem Internet. Um diese Art von Anfragen zuzulassen, müssen eine elastische IP-Adresse oder ein NAT-Gateway im öffentlichen Subnetz eingerichtet werden. Jede dem Internet zugewandte Anfrage sollte entweder über eine elastische IP-Adresse oder ein NAT-Gateway (Network Address Translation) abgewickelt werden.

Sie konnten nun eine Vorstellung davon entwickeln, wie Sie in Ihrer SAP-Umgebung für Netzwerksicherheit sorgen können. Weitere Details finden Sie in dem folgenden AWS-Netzwerkleitfaden (Zugriff am 20.12.2021):

<https://docs.aws.amazon.com/vpc/latest/userguide/what-is-amazon-vpc.html>

---

## 5.6 Sicherung und Wiederherstellung in AWS

In diesem Kapitel wird ein Überblick über die Möglichkeiten gegeben, ein Backup von AWS zu erstellen, entweder mit AWS backint oder mit der Lösung eines Drittanbieters oder mit den integrierten Lösungen für die jeweiligen Datenbanken (Dump to Disk).

Hier wird die native AWS-Sicherungslösung erörtert. Diese kann für SAP HANA DB und EC2-Instanzen sowie für Dateisystem angepasst werden, die mit dem richtigen Lebenszyklus im S3-Speicher gehalten werden sollen. Früher wurden HANA-DB- und Log-Backups auf Festplattenebene mithilfe von Hana Studio oder Hana Cockpit erstellt. Die Backup-Dateien wurden dann per Skript in S3 übertragen. Um diesen betrieblichen Aufwand und die zusätzlichen Speicherkosten einzusparen, hat AWS kürzlich eine eigene Sicherungslösung eingeführt. Mit ihr kann die Sicherung/Wiederherstellung der HANA-DB direkt auf/von dem S3-Speicher konfiguriert werden.

### 5.6.1 SAP-HANA-DB-Sicherung bzw. -Wiederherstellung mit AWS Backint Agent

AWS Backint Agent funktioniert ähnlich wie andere Backint-Agents. Zunächst muss AWS Backint über die AWS-Systemmanagerkonsole installiert werden. Bei der Installation des Backint Agent müssen Angaben zu dem S3-Container gemacht werden, in dem die Daten gesichert werden sollen. Nur nach Mai 2019 erstellte S3-Container sind mit dem AWS Backint Agent kompatibel.

Zurzeit unterstützt AWS Backint Agent nur die Speicherklassen S3 Standard, S3 Standard IA und S3 One Zone-IA. S3 Glacier wird von Backint Agent nicht unterstützt. Nach der Installation und Konfiguration des Agenten müssen die backint-Parameter in der HANA DB (global.ini Konfigurationsdatei) aktualisiert werden, um die DB- und Log-Sicherung von HANA Studio oder HANA Cockpit mit der Backint-Option anstelle

**Tab. 5.6** Übersicht zu den Backup-Typen

Backup-Typ	Sicherungsduer	Speicherklasse	S3-Lebenszyklus
Tägliche DB- und Log-Backups	14 Tage	S3 Standard	Nach 14 Tagen aus S3 Standard löschen
Wöchentliche vollständige Backups	30 Tage	S3 Standard	Nach 30 Tagen aus S3 Standard löschen
Monatliche vollständige Backups	12 Monate	S3 Standard IA	Das Backup wird nach 30 Tagen in S3 Glacier verschoben und nach einem Jahr aus Glacier gelöscht
Jährliche vollständige Backups	11 Jahre	S3 Standard IA	Das Backup wird nach 30 Tagen in S3 Glacier verschoben und nach elf Jahren aus Glacier gelöscht

des Dateisystems zu konfigurieren. Standardmäßig wird das HANA-Protokoll alle 15 min gesichert, die Häufigkeit kann aber entsprechend dem definierten RPO angepasst werden. Hier muss der S3-Lebenszyklus konfiguriert werden, um die Sicherungsdateien möglichst wirtschaftlich zu speichern. Hier ein Beispiel für die kostenoptimierte Backup-Sicherung in verschiedenen S3-Speicherklassen in AWS S3 und gemäß Backup-Richtlinie (Tab. 5.6).

Zur Installation und Konfiguration des AWS Backint Agent für die (vollständige oder inkrementelle) HANA DB und zur Protokollsicherung direkt im S3-Speicher empfehlen wir die Befolgung der standardmäßigen AWS KBA (Zugriff am 20.12.2021).

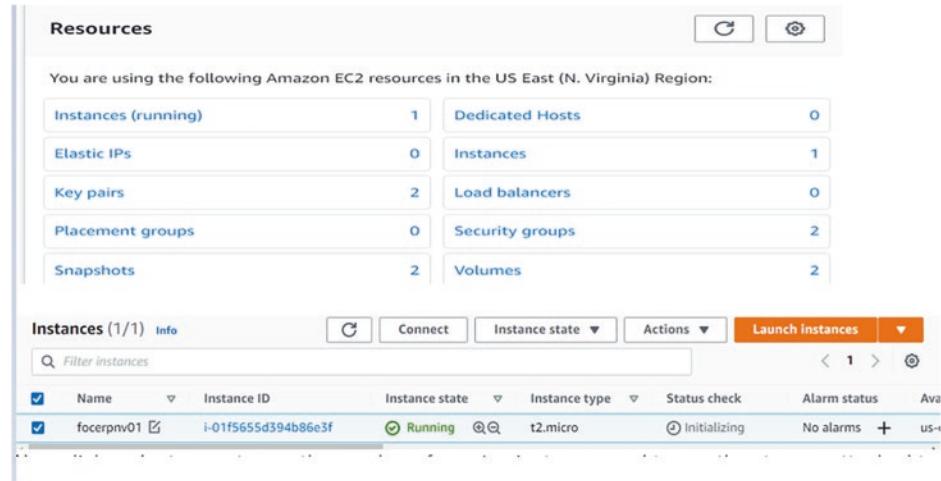
<https://docs.aws.amazon.com/sap/latest/sap-hana/aws-backint-agent-prerequisites.html>

### 5.6.1.1 Backup von EC2-Instanzen und EBS-Volumes in S3

Um das SAP-System vor Ausfällen zu schützen, können Sie Snapshot-Backups von EC2-Instanzen und EBS-Volumes erstellen. Wenn EBS-Volumes einzeln als Snapshot auf S3 gesichert werden sollen und für die Wiederherstellung eine EC2-Instanz benötigt wird, muss der Snapshot im Server verfügbar gemacht werden. Wenn der betreffenden Installationspunkt bereits vorhanden ist, muss er deinstalliert werden. Der EBS-Snapshot kann wiederhergestellt und entweder in dieselbe EC2-Instanz oder eine andere EC2-Instanz eingebunden werden.

In S3 können auch EC2-Snapshot-Backups erstellt werden. Ein EC2-Snapshot enthält ein AMI (Betriebssystemkonfiguration) und alle damit verbundenen EBS-Volumes, sodass ein Replikat oder einen Klon des vorhandenen Systems erstellt werden kann. Im Falle eines Serverabsturzes kann die EC2-Instanz mit der gesamten Konfiguration innerhalb weniger Minuten wiederhergestellt werden.

Wenn also ein bestimmtes Dateisystem wiederhergestellt werden soll, muss der EBS-Snapshot verwendet werden. Wenn ein anderes System mit demselben AMI wiederhergestellt werden soll, wird der EC2-Snapshot verwendet. Hinweis: Der Snapshot kann



**Abb. 5.31** EC2\_Instance\_and\_EBS\_volume\_backup\_to\_S3\_1

nicht in andere S3-Speicherklassen verschoben werden. Der Snapshot ist ein von AWS verwalteter Service und wird immer im S3-Standardspeicher gespeichert. Im Anschluss zeigen wir, wie ein EC2- und EBS-Snapshot erstellt werden können.

Melden Sie sich bei Ihrem AWS-Konto an, und gehen Sie zum EC2-Dashboard, wo alle in Ihrer AWS-Region bereitgestellten EC2-Ressourcen angezeigt werden (Abb. 5.31).

Klicken Sie nun auf „Instanzen“, um die Anzahl der laufenden Instanzen zu sehen und um den an EC2 angeschlossenen Speicher zu sehen. Klicken Sie auf die Instanz-ID und gehen Sie zu der Option „Speicher“. Hier sehen Sie beispielhaft eine Instanz (focerpnv01), die mit einem Root-Volume und einem EBS-Volume läuft.

### 5.6.1.2 Snapshot-Backup über die AWS-Konsole

Wir erläutern nun, wie ein Backup des EBS-Volumes und des Snapshots der EC-Instanz erstellt werden kann. Klicken Sie im EC2-Dashboard auf die Option „Schnappschüsse“, und erstellen Sie einen Snapshot (Abb. 5.32).

Wenn Sie einen bestimmten EBS-Snapshot erstellen möchten, wählen Sie „Volume“. Wählen Sie andernfalls „Instance“, um einen Snapshot für alle an die Instance angeschlossenen EBS-Volumes zu erstellen (Abb. 5.33).

Hier haben wir „Instance“ ausgewählt, denn das gesamte Volume soll gewählt werden. Nun muss die richtige Instanz-ID ausgewählt werden. Hier kann das Root-Volume ausgeschlossen werden. Dann werden nur die EBS-Volumes gesichert. Fügen Sie die Tags gemäß den von Ihnen festgelegten Richtlinien hinzu (Abb. 5.34).

Ihr Snapshot wird erstellt, sobald Sie auf die Schaltfläche „Snapshot erstellen“ klicken.



The screenshot shows a table of snapshots. There are two rows:

Name	Snapshot ID	Size	Description	Status	Started
snap-0a4181894b1...	8 GiB	Created by CreateImage(0e2099eed0a6804ce) for ami-0c86e...	completed	August 18	
snap-0dc0d0703d7...	8 GiB	Created by CreateImage(0e2099eed0a6804ce) for ami-0c86e...	completed	August 18	

### Create Snapshot

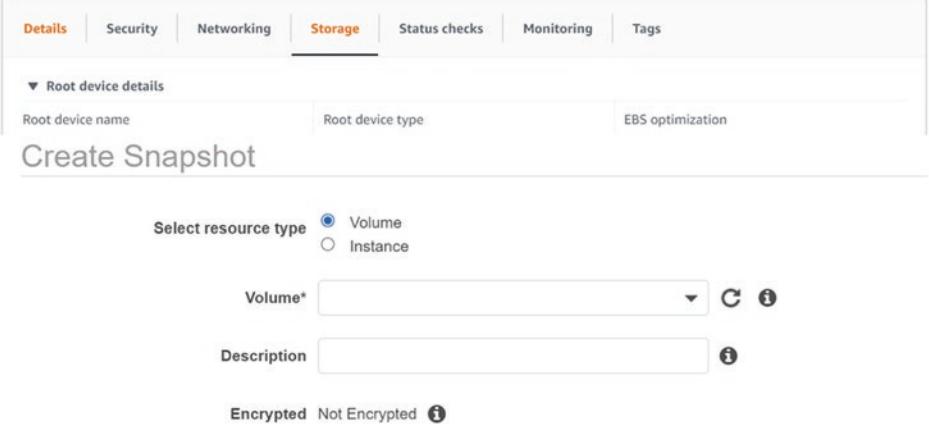
Select resource type  Volume  Instance

Volume\*  C ⓘ

Description  ⓘ

Encrypted Not Encrypted ⓘ

**Abb. 5.32** Backup mit einem Snapshot – Schritt 1



The screenshot shows the EC2 instance details page with the Storage tab selected. Below the storage section, there is a 'Create Snapshot' button.

Select resource type  Volume  Instance

Volume\*  C ⓘ

Description  ⓘ

Encrypted Not Encrypted ⓘ

**Abb. 5.33** Backup mit einem Snapshot – Schritt 2

Wenn Sie nun im EC2-Dashboard die Option „Snapshot“ wählen, sehen Sie zwei sogenannte testsnapshots, einen für das Root-Volume und einen für das EBS-Volumen (Abb. 5.35).

#### 5.6.1.3 EC2-Image-Backup

Wir führen Sie nun durch die Schritte zur Erstellung eines EC2-Image-Backups (AMI), mit dem die gesamte EC2-Instanz innerhalb weniger Minuten wiederhergestellt werden kann.

## Create Snapshot

Select resource type  Volume  Instance

Volume\*  C i

Description  i

Encrypted Not Encrypted i

Volume ID	Volume Type	Encryption
vol-042aa78d53ab27800	Root	Not Encrypted
vol-029ea6e1d6afbf081	EBS	Not Encrypted

**Abb. 5.34** Backup mit einem Snapshot – Schritt 3

Create Snapshot Actions ▾						
Owned By Me		Filter by tags and attributes or search by keyword				Actions
Name	Snapshot ID	Size	Description	Status	Started	
	snap-0a4181894b1...	8 GiB	Created by CreateImage(i-0e2099eed0a6804ce) for ami-0c86e...	<span style="color: green;">● completed</span>	August 18	
	snap-0db1e4124cb...	10 GiB	testsnapshot	<span style="color: green;">● completed</span>	August 18	
	snap-0dcod0703d7...	8 GiB	Created by CreateImage(i-0e2099eed0a6804ce) for ami-0c86e...	<span style="color: green;">● completed</span>	August 18	
	snap-0f9074a99889...	8 GiB	testsnapshot	<span style="color: green;">● completed</span>	August 18	

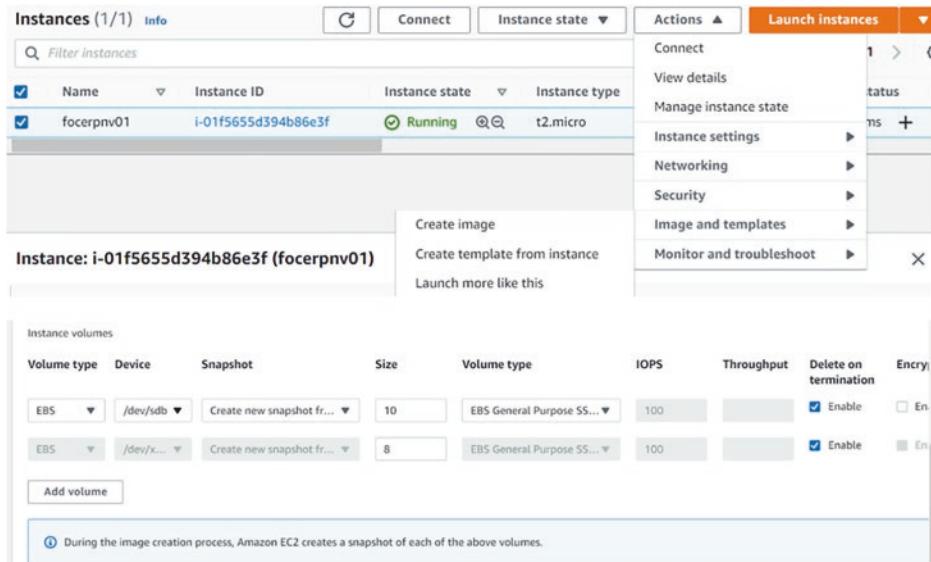
**Abb. 5.35** Backup mit einem Snapshot – Schritt 4

Wählen Sie die Instanz aus, deren Image erstellt werden soll. Gehen Sie zu Aktion Image und Template, Image erstellen, erfassen Sie den Image-Namen und die sonstigen erforderlichen Angaben, und klicken Sie auf die Schaltfläche „Image erstellen“ (Abb. 5.36).

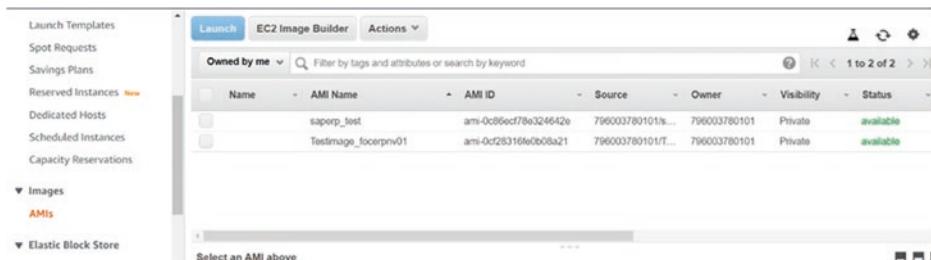
Jetzt wird das das Instanz-Image in Images AMI angezeigt (Abb. 5.37).

## 5.6.2 AWS-Backup

AWS hat jüngst den verwalteten AWS-Service „AWS Backup“ eingeführt. Es handelt sich um eine zentralisierte Backup-Lösung für die automatisierte Sicherung von AWS-Ressourcen wie EC2-Instanz-AMI, EBS-Volume, EFS-Dateifreigabe etc. Es können die Backup-Richtlinien über den Backup-Plan erstellen, einschließlich Zeitplanung, Aufbewahrungszeitraum und Backup-Überwachung geplant werden. Hierfür empfiehlt sich



**Abb. 5.36** Backup einer EC2 Instanz und EBS Volumen

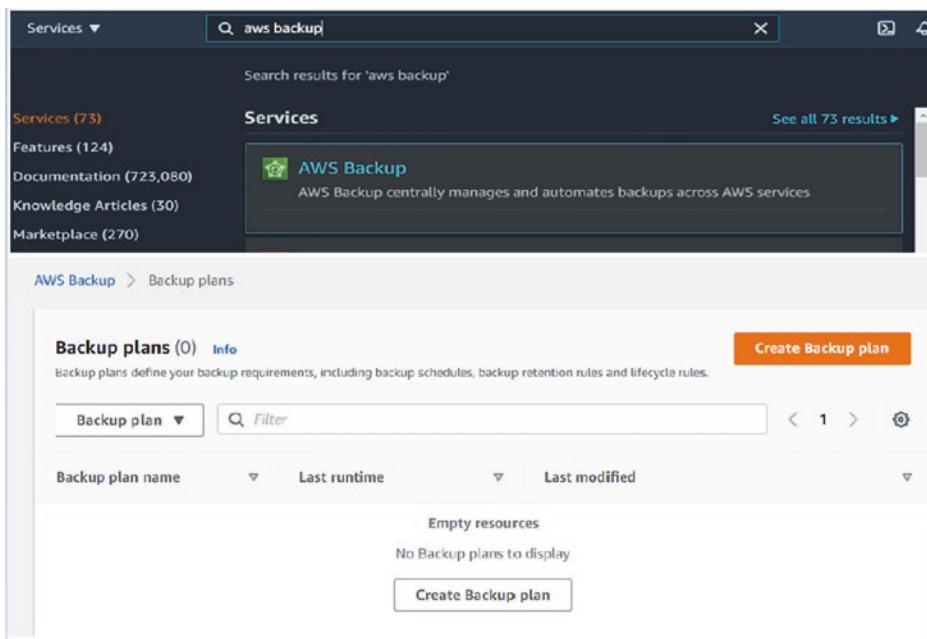


**Abb. 5.37** Backup einer EC2-Instanz als Image

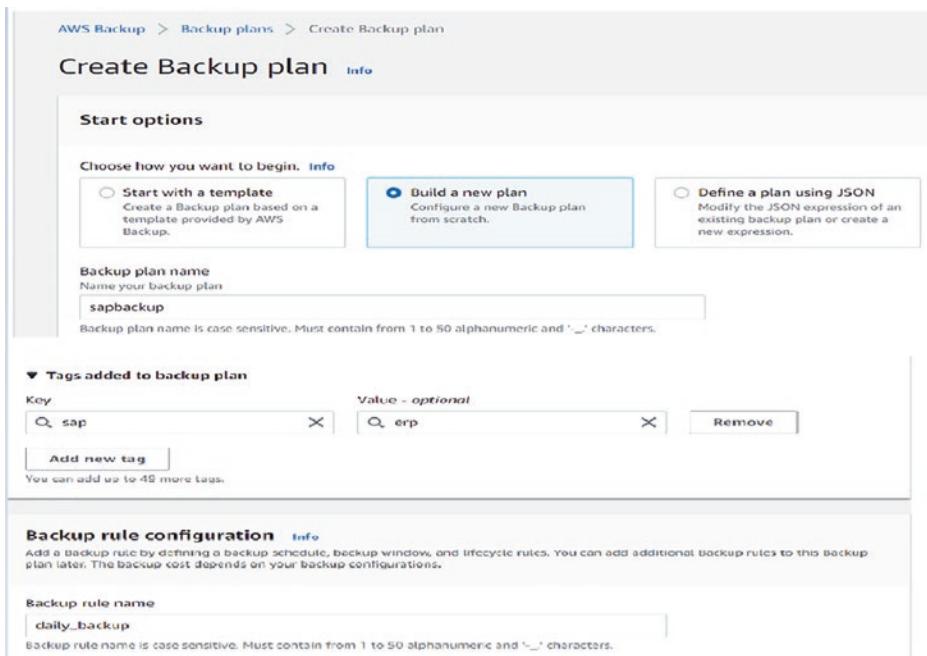
der Service „AWS Backup“. Sie können das AMI-Backup für verschiedene Regionen direkt konfigurieren und die (als DR-Region gekennzeichneten) AMI automatisch in die DR-Region kopieren. Im Katastrophenfall kann das System in der DR-Region mit dem AMI des Primärservers schnell wiederhergestellt werden. Der Service „AWS Backup“ unterstützt lediglich die Ressourcen EBS, EC2, EFS, RDS, DynamoDB, Storage Gateway und Amazon FSx.

### Backup der Konfigurationsschritte aus der AWS-Konsole

Suchen Sie im Bereich Service nach AWS Backup, und klicken Sie darauf. Erstellen Sie einen Backup-Plan gemäß den Backup-Richtlinien Ihres Kunden (Abb. 5.38, 5.39, 5.40, 5.41).



**Abb. 5.38** Erstellen eines AWS Backups – Schritt 1



**Abb. 5.39** Erstellen eines AWS Backups – Schritt 2

Backup vault [Info](#)

Default [▼](#) Create new Backup vault

Backup frequency [Info](#)

Daily [▼](#)

Enable continuous backups for supported resources [Info](#)

Backup window

Use backup window defaults - recommended [Info](#)  
5 AM UTC, starts within 8 hours.

Customize backup window

Transition to cold storage [Info](#)

Never [▼](#)

Retention period [Info](#)

Days [▼](#) 14 [Remove](#)

Copy to destination - optional [Info](#)

Europe (Stockholm) [▼](#) Remove

Copy to another account's vault

Destination Backup vault

The vault to which your backup copy will be made.

Default [▼](#) Create new Backup vault

▼ Advanced settings

Transition to cold storage [Info](#)

Never [▼](#)

Retention period [Info](#)

Always [▼](#)

Add copy

**Abb. 5.40** Erstellen eines AWS Backups – Schritt 3

▼ Advanced backup settings

Application-consistent backup [Info](#)

Enable application-consistent snapshots for the selected third-party software running on EC2.

Windows VSS

**ⓘ You can assign resources to this Backup plan after the plan has been created.**

**ⓘ You can add more rules to this Backup plan after the plan has been created.**

[Cancel](#) [Create plan](#)

**Abb. 5.41** Erstellen eines AWS Backups – Schritt 4

Resource assignments (1)			Delete	Assign resources
Resource assignments specify which resources will be backed up by this Backup plan.			< 1 ... >	⚙️
Name	IAM role ARN			
saperp	arn:aws:iam::796003780101:role/service-role/AWSBackupDefaultServiceRole			

**Abb. 5.42** Erstellen eines AWS Backups – Schritt 5

Klicken Sie auf „Plan erstellen“. Der Backup-Plan wird erstellt. Nun müssen die zu sichernden Ressourcen zugewiesen werden. Ressourcen können auf zwei Arten zugewiesen werden: per Tag oder per Ressourcen-ID. Bei der Auswahl „Tag“ werden alle Ressourcen mit demselben Tag in die Richtlinie aufgenommen. Bei der Auswahl „Ressourcen-ID“ müssen die Ressourcen zugewiesen werden, die in den Backup-Plan aufgenommen werden sollen. In diesem Beispiel wurde die Ressourcen-ID ausgewählt (Abb. 5.42).

In Ihrem Backup-Plan werden nun die Ressourcen angezeigt, die gemäß der von Ihnen festgelegten Richtlinie gesichert und aufbewahrt werden sollen. Erstellen Sie ein AWS-Konto, und erforschen Sie so viele Optionen wie möglich.

### 5.6.3 Wiederherstellung des Snapshots von der AWS-Konsole

Zur Einhaltung des RPO/TRO muss der Wiederherstellungsprozess für nicht kritische Umgebungen mindestens einmal im Quartal getestet werden. Die EC2-Instanz kann hierfür mit der gesamten Konfiguration des Quellsystems wiederhergestellt werden. Ebenfalls ist es möglich, nur das EBS-Volume wiederherzustellen und es in eine beliebige EC2-Instanz einzubinden. Hier zeigen wir die Wiederherstellung der EC2 Instanz AMI.

Gehen Sie zu AMI Images. Dort finden Sie alle Images, die in Ihrer AWS-Region erstellt wurden. Wählen Sie das wiederherzustellende AMI aus, und rufen Sie es auf (Abb. 5.43).

Wählen Sie den Typ der Zielinstanz. Hier wurde t2.micro gewählt. Vervollständigen Sie den Startassistenten mit allen Details, überprüfen Sie Ihre Eingaben, und starten Sie dann die Instanz (Abb. 5.44).

In der Option „Storage“ sehen Sie ein Root- und ein EBS-Volume, die in das AMI eingebunden wurden. Sie können je nach Bedarf weitere EBS-Volumes hinzufügen oder vorhandene EBS-Laufwerke entfernen. Stammlaufwerke können nicht entfernt werden (Abb. 5.45).

Name	AMI Name	AMI ID	Source	Owner	Visibility	Status	C
	saperp_test	ami-0c86ecf78e324642e	796003780101/s...	796003780101	Private	available	A
<input checked="" type="checkbox"/>	Testimage_focerpnv01	ami-0cf28316fe0b08a21	796003780101/T...	796003780101	Private	available	A

**Abb. 5.43** Wiederherstellung eines Snapshots – Schritt 1

Step 3: Configure Instance Details

Configure the instance to suit your requirements. You can launch multiple instances from the same AMI; request Spots instances to take advantage of the lower pricing; assign an access management role to the instance, and more.

Number of instances: 1      Launch into Auto Scaling Group:

Purchasing option:  Request Spot instances

Network: vpc-2905775b (default)       Create new VPC

Subnet: No preference (default subnet in any Availability Zone)       Create new subnet

Auto-assign Public IP: Use subnet setting (Enable)     

Placement group:  Add instance to placement group

Cancel Previous Review and Launch Next: Add Storage

Step 4: Add Storage

Your instance will be launched with the following storage device settings. You can attach additional EBS volumes and instance store volumes to your instance, or edit the settings of the root volume. You can also attach additional EBS volumes after launching an instance, but not instance store volumes. [Learn more about storage options in Amazon EC2.](#)

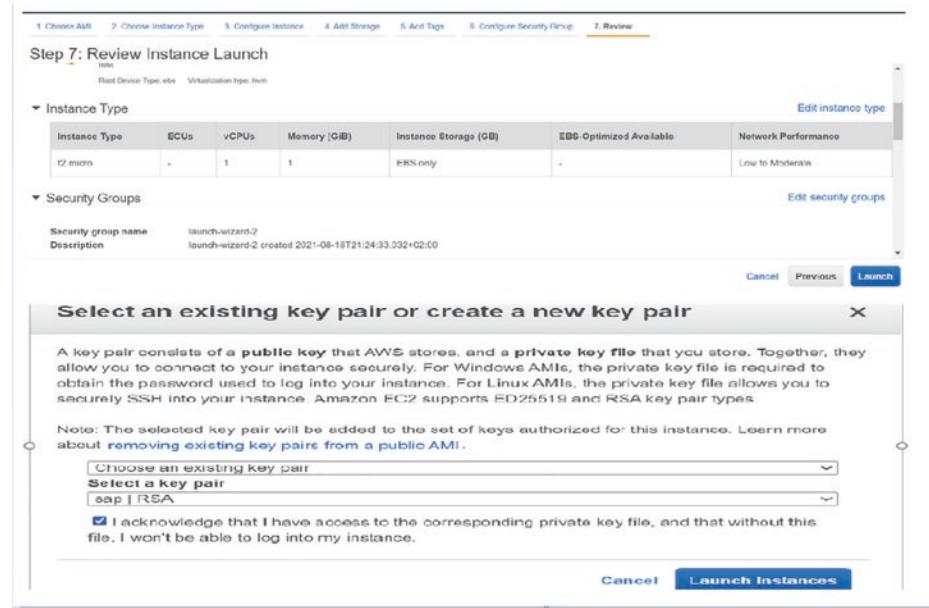
Volume Type	Device	Snapshot	Size (GiB)	Volume Type	IOPS	Throughput (MB/s)	Delete on Termination	Encryption
Root	/dev/xvda	snap-09761a96e303e337d	8	General Purpose SSD (gp2)	100 / 3000	N/A	<input checked="" type="checkbox"/>	Not Encrypted
EBS	/dev/sdb	snap-0532a9de922e03	10	General Purpose SSD (gp2)	100 / 3000	N/A	<input checked="" type="checkbox"/>	Not Encrypted

Add New Volume      Cancel Previous Review and Launch Next: Add Tags

**Abb. 5.44** Wiederherstellung eines Snapshots – Schritt 2

Voraussetzung ist das Schlüsselpaar für die Anmeldung an der Instanz nach dem Start. Hier wurde ein bestehendes Schlüsselpaar gewählt. Um die Instanz zum ersten Mal zu starten, wählen Sie die Option „Schlüsselpaar erstellen“.

Ihre neue EC2-Instanz mit dem Image der bestehenden Instanz ist bereit. Dies ist der grundlegende Sicherungs- und Wiederherstellungsprozess über die AWS-Konsole. Die ist über die AWS-Befehlszeilenschnittstelle möglich und kann automatisiert werden. Erstellen Sie dafür lediglich ein AWS-Konto, und erforschen Sie so viele Optionen wie möglich.



**Abb. 5.45** Wiederherstellung eines Snapshots – Schritt 3

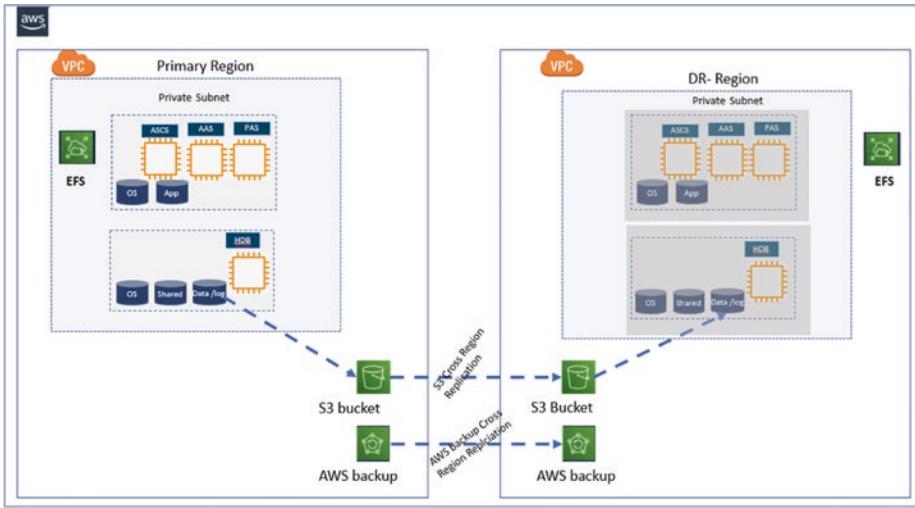
## 5.7 Disaster Recovery mit AWS

In diesem Kapitel erläutern wir die Möglichkeiten zur Implementierung einer geeigneten DR-Lösung auf der Grundlage von AWS-Mechanismen oder datenbankbasierten DR-Mechanismen. Das Kapitel wird auch dazu beitragen, die RTO/RPO für jede Lösung festzulegen.

Disaster Recovery ist sowohl für die SAP-Umgebung also für die Geschäftskontinuität ein sehr wichtiger Aspekt. In AWS kann die DR-Lösung für produktive Workloads für einzelne Regionen mit mehreren Availability Zones oder für mehrere Regionen je nach Kundenanforderung geplant und gestaltet werden. Hier zeigen wir, wie unterschiedliche Arten von DR-Lösungen basierend auf definierten RPO (Recovery Point Objective) und RTO (Recovery Time Objective) entworfen werden können.

### 5.7.1 Passive DR-Architektur [RPO ist größer als 0 (null) und RTO ist höher]

Diese DR-Lösung ist am kostengünstigsten. Die SAP-DR-Umgebung muss nicht im Voraus eingerichtet werden, sodass keine laufenden Kosten für EC2-Instanzen und



**Abb. 5.46** Passive DR-Lösung

Speicher anfallen. Der nachstehenden Architektur ist zu entnehmen, wie eine passive DR in der Cross-Region erreicht werden kann (Abb. 5.46).

Wenn es in der primären Region zu einer Katastrophe kommt, ist es sehr einfach, die Systeme in einer anderen Region einzusetzen. Hierfür wird die Konfiguration der Datenbank, der EC2-Instanzen und ggf. des EFS-Speichers benötigt.

**DB-Backup** – Das DB-Backup muss mit dem AWS-Backint-Agenten für HANA oder mit einem Tool eines Drittanbieters wie Commvault konfiguriert sein. Die Datenbank und die Protokolle müssen im AWS-S3-Container gespeichert sein. Wenn sich die DR-Umgebung in einer anderen Region befindet, muss explizit Cross Region Replication (CRR) für den S3-Container eingestellt werden. Die Daten können so aus der primären Region in die DR-Region repliziert werden. Wenn sich die DR-Umgebung in derselben Region, aber dort in einer anderen Availability Zone befindet, repliziert S3 standardmäßig Daten in alle Availability Zones, die sich in derselben Region befinden.

**AWS-Backup** – AWS-Backups können auch in anderen Regionen wiederhergestellt werden. Hierfür müssen die AWS-Backups für die AMIs der SAP- und DB-Instanzen und für die Speicher konfiguriert sein. Die AWS-Sicherung wird in eine andere Region repliziert. Im Katastrophenfall kann sehr schnell das AMI mit derselben Konfiguration wie das primäre AMI in der DR-Region wieder, und dasselbe EFS-Volume (/sapmnt, /usr/sap/trans, /interface) kann in der DR-Umgebung installiert werden. Sobald der Server bereit ist, muss die HANA-DB aus dem S3-Container wiederhergestellt werden.

## Schritte zur Konfiguration der passiven DR

- Konfigurieren Sie das Backup der HANA-DB-Daten und der Protokolle mit AWS Backint oder mit einem beliebigen Backup-Werkzeug eines Drittanbieters, und speichern Sie die Sicherung im S3-Speicher.
- Aktivieren Sie S3 Cross Region Replication (CRR)
- Erstellen einen Backup-Plan, und Konfigurieren Sie das AWS-Backup für die EC2-Instanz-AMIs (S/4 und HANA DB AMI) und den EFS-Speicher
- Das AWS-Backup repliziert das AMI in die DR-Region.

## Schritte zum Aufrufen der passiven DR

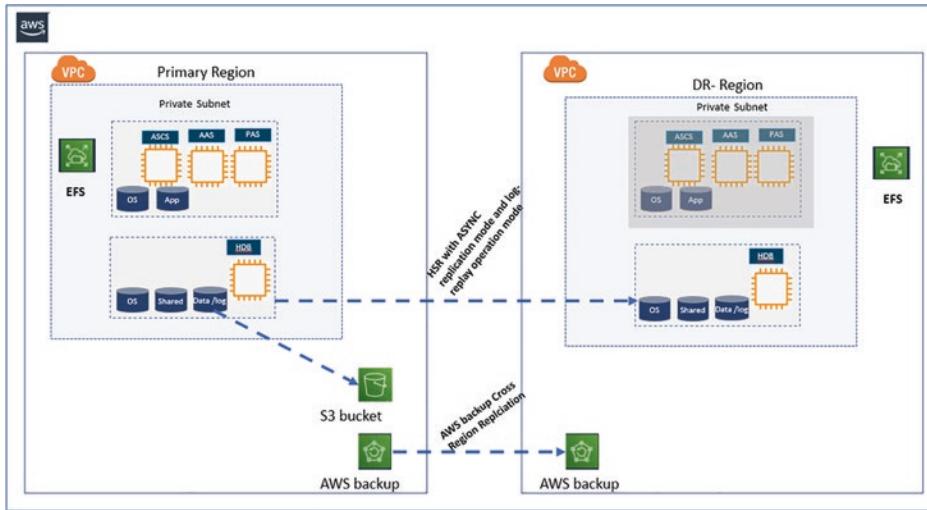
- Stellen Sie die S/4-Anwendung und den HANA-DB-Server von einem per AWS-Backup gesicherten AMI wieder her.
  - Stellen Sie die HANA-Datenbank aus der S3-Sicherung wieder her.
  - Installieren Sie das EFS-Dateisystem, das in die DR-Region kopiert wurde.
  - Starten Sie die HANA-Datenbank und die S/4-Anwendung.
- Wenn Sie die DR in derselben Region, aber in einer anderen Availability Zone konfigurieren, müssen Cross Region Replication und AWS Backup nicht konfiguriert werden.

### 5.7.2 Semiaktive DR-Lösung [RPO ist nahe null und RTO ist medium]

Bei dieser DR-Lösung werden RPO als nahe null und RTO als etwa 12 bis 15 h festgelegt. Dies ist eine kostengünstige DR-Lösung. Die Größe des in der DR-Region bereitgestellten HANA-DB-Systems muss dabei geringer sein als in der primären Region. Das Speicher-volumen des DR-HANA-DB-Servers muss der Mindestgröße der DB-Zeilenspeicher-tabelle + 60 GB entsprechen. Um die Daten kontinuierlich vom primären zum DR-Standort zu replizieren, muss die HANA-Systemreplikation mit dem asynchronen Replikations-modus und dem Betriebsmodus Log-Reply oder Delta-Replay konfiguriert werden. Um die laufenden DR-Kosten zu minimieren, entscheiden wir uns hier für das kleinere DB-DR-Server-Computing-Volumen. Die Tabellenoption „Preload“ muss daher während der HSR-Konfiguration zwischen Primär- und Sekundärserver deaktiviert werden.

```
global.ini/[system_replication] -> preload_column_tables=false
```

Das AWS-Backup für die S/4-Anwendung, die EC2-Instanz-AMIs und den EFS-Speicher muss zur Wiederherstellung der SAP-Anwendung in der DR-Region konfiguriert werden (Abb. 5.47).



**Abb. 5.47** Semiaktive DR-Lösung

### Schritte zur Konfiguration der semiaktiven DR

- Schätzen Sie die Größe des DR-DB-Servers – (Größe der Zeilenspeichertabelle + 60) GB. Für die Installation von HANA DB sind mindestens 64 GB Speicher erforderlich.
- Berechnen Sie die Größe des Zeilenspeichers:
- select host, round (sum(page\_size\*USED\_BLOCK\_COUNT)/1024/1024/1024,2) as „RowStore Size GB“ from m\_data\_volume\_page\_statistics where page\_sizeclass = ‚16k-RowStore‘ group by host;
- Stellen Sie die HANA-DB-VM bereit, und installieren Sie HANA-Software.
- Konfigurieren Sie die Hana-Systemreplikation mit asynchronem Replikationsmodus und Log-Replay-Betriebsmodus.
- Deaktivieren Sie die Option zur Vorbelegung der Tabelle.
- Konfigurieren Sie das AWS-Backup für die SAP-S/4-Instanzen und den EFS-Speicher.

### Schritte zum Aufrufen der DR

- Führen Sie am DR-Standort ein HANA-Takeover durch.
- Halten Sie die DR-Hana-Datenbank an.
- Dimensionieren Sie die HANA VM wie den primären Server.
- Starten Sie die HANA-Datenbank.
- Stellen Sie die SAP-S/4-Anwendungsinstanzen aus einem in AWS gesicherten AMI wieder her.
- Installieren Sie das EFS-Dateisystems auf den Anwendungsservern.
- Starten Sie die S/4-Anwendung.

- Wenn Sie die DR in derselben Region, aber in einer anderen Availability Zone konfigurieren, können Sie zwecks Kostensenkung einen Qualitätsserver als DR-Server verwenden. Stoppen Sie im Katastrophenfall das Qualitätssystem, und ändern Sie die Größe der VM, um die DR aufrufen zu können. Es ist nicht erforderlich, das AWS-Backup in derselben Region zu konfigurieren.

### 5.7.3 Aktive DR-Lösung [RPO nahe 0 (null) und RTO sehr gering]

Diese DR-Lösung ist teuer, bietet aber ein geringes RTO. Die DR-Umgebung kann innerhalb von zwei Stunden in Betrieb genommen werden. Das aktive DR-Design ist dasselbe wie das semiaktive. Der einzige Unterschied besteht darin, dass der HANA-DB-Server identisch zum primären Server bereitgestellt wird.

#### Schritte zur Konfiguration der aktiven DR

- Stellen Sie HANA-DB-Server in der DR-Region bereit, die dieselben Computing- und Speicherkapazitäten aufweisen wie der primäre Server.
- Installieren Sie die HANA-DB-Software.
- Konfigurieren Sie die Hana-Systemreplikation mit asynchronem Replikationsmodus und Log-Replay-Betriebsmodus. Aktivieren Sie dabei die Vorbelegungsoption.
- Konfigurieren Sie das AWS-Backup für das EC2-AMI der SAP-S/4-Anwendung für die DR-Region.
- Konfigurieren Sie das AWS-Backup des EFS-Speichers für die DR-Region, damit Sie dieselben Dateisysteme auf den DR-Servern installieren können.

#### Schritte zum Aufrufen der DR

- Führen Sie die HANA-DB-Übernahme durch.
- Stellen Sie die SAP-S/4-Anwendungsinstanzen aus einem in AWS gesicherten AMI wieder her.
- Installieren Sie die EFS-Dateisysteme auf den DR-Servern.
- Starten Sie die SAP-S/4-Anwendung.

Nach dem Aufruf der DR-Umgebung im Katastrophenfall muss die DNS aktualisiert werden. Der Hostname oder Alias Ihres SAP S/4-Anwendungsservers muss dem des Primärservers entsprechen. Die IP-Adresse ist aber eine andere. Nachdem der Datenverkehr auf die DR-Server umgestellt ist, muss die IP-Adresse des Primärserver im DNS-Server durch die IP-Adresse des DR-Servers ersetzt werden. Aus Sicht der Endbenutzer und der Schnittstelle sind daher im Katastrophenfall keine Änderungen erforderlich.

Wir haben wir mehrere Optionen zur DR-Konfiguration in verschiedenen Regionen gezeigt. Es ist auch möglich, die DR in derselben Region für verschiedene Availability

Zones zu konfigurieren. In diesem Fall muss das AWS-Backup für die EC2-Instances nicht konfiguriert werden. Auch ist es nicht notwendig, die CRR für den S3-Speicher zu aktivieren. Die EC2-Instanz kann direkt aus S3 in einer anderen Availability Zone wiederhergestellt werden, da S3 standardmäßig die Daten anderer Zonen in derselben Region repliziert.

Ein RPO von null und ein sehr geringes RTO sind möglich, wenn die DR-Umgebung in derselben Region für eine andere Availability Zone konfiguriert wird, um die HANA-Systemreplikation mit dem synchronen Replikationsmodus und dem Log-Replay-Betriebsmodus durchzuführen.

Konfigurationsleitfaden für SAP HANA HSR (Zugriff am 20.12.2021):

[https://help.sap.com/doc/c81e9406d08046c0a118c8bef71f6bdc/2.0.04/en-US/SAP\\_HANA\\_System\\_Replication\\_Guide\\_en.pdf](https://help.sap.com/doc/c81e9406d08046c0a118c8bef71f6bdc/2.0.04/en-US/SAP_HANA_System_Replication_Guide_en.pdf)

---

## 5.8 Zusammenfassung

Dieses Kapitel hat Ihnen den konkreten Aufbau und die jeweiligen Schritte für die Implementierung von einem neuen SAP S/4HANA-System in der AWS Cloud gezeigt. Dazu wurde zunächst die beispielhafte SAP-Systemarchitektur erläutert und die wichtigsten Komponenten zur Implementierung identifiziert. Danach wurde gezeigt, wie Compute und Storage-Komponenten bereitgestellt werden.

Um das neue SAP-System möglichst kostengünstig zu provisionieren, haben wir Ihnen gezeigt, wie Sie die Kosten eines neuen SAP S/4HANA-Systems für die beiden Optionen Pay-As-You-Go und Commitment berechnen können.

Als weiteren wichtigen Faktor sind wir auf die Sicherheit des neuen SAP-Systems in AWS eingegangen und haben die wichtigsten Werkzeuge, wie ACLs und SGs gezeigt. Die Einrichtung solcher Sicherheitsmaßnahmen mag manchmal etwas arbeitsam erscheinen, aber ist absolut erforderlich. Nur durch den Schutz der Umgebung können Angreifer abgehalten werden.

Die Sicherung und die Wiederherstellung von Daten aus einem SAP S/4HANA-System kann auf verschiedene Wege erfolgen. So können die eigenen AWS-Werkzeuge genutzt werden, oder aber die auf AWS erhältlichen Tools von Drittanbietern. Die AWS-Bordmittel bieten aber bereits einen guten Schutz vor einem potenziellen Datenverlust.

Durch die Verteilung der Daten eines SAP S/4HANA-Systems auf mehr als eine Region oder Verfügbarkeitszone können die SAP-Systeme auch in einem Katastrophenfall weiterhin verfügbar gehalten werden. Hier gibt es mehrere Möglichkeiten, welche in Abhängigkeit der RTO und RPO gewählt werden können.



# SAP S/4 on Microsoft Azure – Konzepte und Architekturen

6

## Zusammenfassung

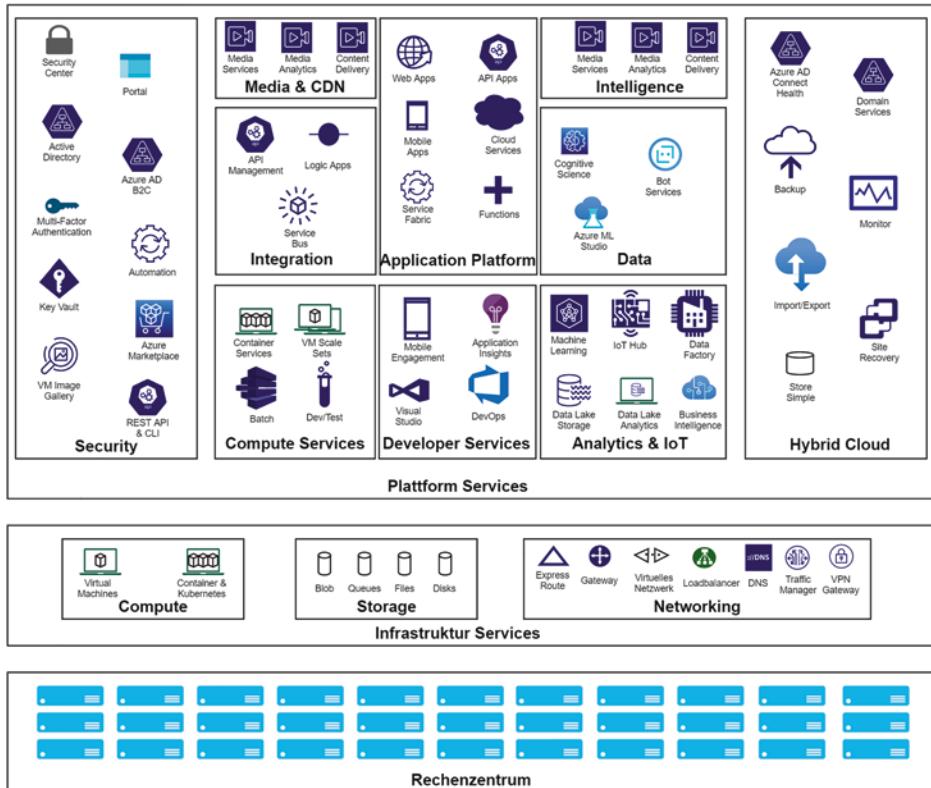
Dieses Kapitel bietet einen Einstieg in die grundlegenden Konzepte und Funktionsweisen der Microsoft Azure Cloud. Ziel des Kapitels ist es, eine detaillierte Einführung in das Cloud Hosting zu geben und insbesondere das theoretische Verständnis für die verschiedenen Komponenten und Begrifflichkeiten zu schaffen. Es bildet somit die Basis für die praktische Implementierung der beispielhaften Systemarchitektur in Kap. 7. Nach einem kurzen Einstieg in die Geschichte von Microsoft Azure werden zunächst die Governance-Strukturen erläutert, welche die Abbildung komplexer Unternehmensstrukturen sowie Verantwortlichkeiten und Genehmigungsprozesse in Azure ermöglichen. Aspekte, welche hierbei beleuchtet werden betreffen die Accountverwaltung mittels Abonnements und Verwaltungsgruppen, sowie Rollenkonzepte und Kontingente. Der Hauptteil des Kapitels befasst sich schließlich mit dem eigentlichen Ressourcenmanagement. Zu diesem Zweck werden alle für das SAP Hosting relevanten Terminologien und Bereitstellungsoptionen erläutert. Der Fokus wird hierbei auf die Bereitstellung mittels virtueller Computer gelegt, wobei alternative Angebote wie z. B. Hana Large Instances ebenfalls eingeführt werden. Den Abschluss des Kapitels bildet schließlich die Vorstellung einer S/4HANA Referenzarchitektur nach dem Vorbild der offiziellen Microsoft Dokumentation. Hierbei wird insbesondere auf das Zusammenspiel der eingangs erläuterter Komponenten eingegangen.

## 6.1 Historischer Überblick über Microsoft Azure

Microsoft begann vor einigen Jahrzehnten am Cloud Markt, als die bedeutenden Auswirkungen ersichtlich wurden, die die öffentliche Cloud haben würde. Schon bald gab Microsoft zusammen mit AWS das Tempo vor. Das Kapitel zeichnet ein Bild der Geschichte von Azure und der engen Zusammenarbeit zwischen Microsoft und SAP beim Hosting von SAP-Systemen.

Microsoft Azure startete 2010 am Cloud Markt und erweitert seitdem kontinuierlich das Serviceangebot der Cloud Plattform. Auch wenn die Azure Cloud Plattform von Microsoft angeboten wird, ist es neben der Nutzung von Windows-Servern auch möglich Linux-Server zu verwenden. Dabei werden Produkte in den Bereichen Infrastructure-as-a-Service, Platform-as-a-Service and Software-as-a-Service angeboten. Dadurch wird eine große Flexibilität bei der Betriebssystem-Auswahl geboten: SUSE, Debian, Ubuntu und Windows. Ebenfalls bei der Implementierung von Anwendungen namhafter Hersteller bietet sich ein großes Spektrum, da eine große Auswahl an Anwendungen von SAP, Oracle und IBM in Azure implementierbar ist. Durch die hohe Anzahl an Rechenzentren, die weltweit lokalisiert sind, kann eine hohe Performanz durch kurze Latenzzeiten geboten werden. Dabei besteht für Kunden die Möglichkeit zu definieren, in welchem Rechenzentrum Ressourcen bezogen werden. Damit können auch strenge Datenschutzvorgaben eingehalten werden. Zusätzlich bietet sich die Möglichkeit eine Hybrid Cloud zu betreiben, wobei lediglich einzelne Anwendungen in Microsoft Azure betrieben werden und in die eigens betriebene Systemlandschaft integriert sind. Abb. 6.1 zeigt einen Überblick der angebotenen Services für die jeweiligen Bereiche.

Seit 25 Jahren besteht zwischen SAP und Microsoft eine bewährte Partnerschaft. Im Januar 2021 gaben Microsoft Azure und SAP eine noch engere Kooperation bekannt, insbesondere in Bezug auf die Einführung und Nutzung von S/4HANA in der Cloud. Damit einhergehend wurde Azure zertifizierter Anbieter von Cloudplattformen für SAP-Anwendungen. Die folgenden SAP Lösungen können in Azure genutzt werden: SAP HANA, SAP S/4HANA, SAP BW/4HANA, SAP Business Suite, SAP HEC, SAP Business One, SAP Hybris, SAP Cloud Platform. Von anderen Cloud Anbietern kann sich Microsoft Azure vor allem anhand der Bereitstellungsoptionen in Bezug auf SAP HANA absetzen. Hierbei kann Microsoft Azure Bare-Metal Hardware anbieten, weshalb selbst hoher Ressourcenbedarf von bis zu 24–120 TB Arbeitsspeicher abgedeckt werden kann. Insgesamt können SAP-Anwendungen in Microsoft Azure im Rahmen von Entwicklungs-, Test- und Produktionsszenarien mit vollständigem Support betrieben werden. Azure Kunden erhalten dadurch die Möglichkeit, SAP-Legacyinfrastrukturen in Azure mit umfänglicher Zertifizierung zu migrieren. Neue Optionen der Cloud-Automatisierung sowie die Vereinfachung der Migration von On-Premise-Strukturen in die Azure Cloud sind dadurch möglich geworden. Bereitgestellte Referenzarchitekturen sowie ausführliche Dokumentationen unterstützen das Vorhaben des Kunden, in die Azure Cloud zu migrieren. Die Migration der SAP-Systeme in Microsoft Azure kann in fünf Schritte untergliedert werden: initiale Sondierung, Vorbereitung, Migration, Betrieb,



**Abb. 6.1** Serviceportfolio der Microsoft Azure Plattform

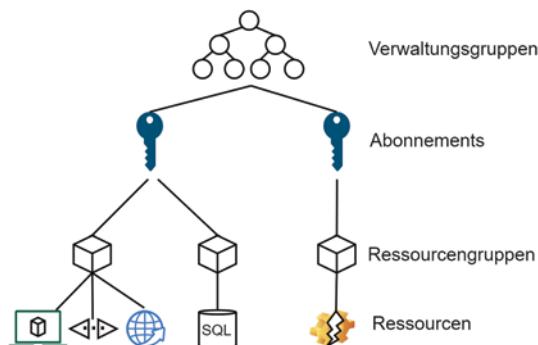
Innovation. Die Dokumentationen unterstützen bei der initialen Sondierung wie die vorliegende SAP Landschaft erfasst werden kann und Ziele für die Umsetzung definiert werden. Die Vorbereitungphase wird anhand von Dokumentationen und Referenzarchitekturen bei der Planung und der richtigen Dimensionierung begleitet. Auch bei der Migration und Implementierung unterstützt SAP und Microsoft Azure das Vorhaben und gibt Vorschläge zur Optimierung der Landschaftsarchitektur. Über das Azure Active Directory wird ein sicherer Single Sign-On Zugang zur SAP Anwendung in Azure ermöglicht. Zusätzlich wird eine Integration von SAP-Anwendungen mit Microsoft Teams und Microsoft Office angestrebt, was die Partnerschaft weiterhin unterstreicht.

## 6.2 Azure Steuerung und Abonnements

In diesem Kapitel wird das Konzept der Verwaltungsgruppen und die Beziehung zwischen diesen und den Abonnements erläutert. Hierbei gehen wir insbesondere auf die verschiedenen Abonnement Modelle ein und erläutern wie komplexe Governance Strukturen eines Unternehmens in Azure abgebildet werden können.

### **6.2.1 Azure Organisationsstruktur**

Wie in Abb. 6.2 dargestellt weist die Azure Organisationsstruktur insgesamt vier Ebenen auf: Verwaltungsgruppen, Abonnements, Ressourcengruppen und Ressourcen. Im Wesentlichen definieren **Abonnements** das Benutzerkonzept, sowie Grenzwerte und Kontingente für die darin enthaltenen Ressourcen. Auf diese Weise können Unternehmen eine logische Trennung beispielsweise für unterschiedliche Business Units oder Projekte vornehmen. Eine **Verwaltungsgruppe** (Management group) ist dabei den Abonnements übergeordnet und bietet die Möglichkeit diese zu gruppieren. Einige Konzepte, welche auf Ebene von Abonnements definiert werden können, wie Role-Based-Access-Control (RBAC) Berechtigungen oder Azure-Policies lassen sich auch auf Ebene der Verwaltungsgruppe definieren. Auf diese Weise können allgemeingültige Richtlinien und Zugriffe definieren werden, welche von allen untergeordneten Abonnements automatisch geerbt werden. Hierbei ist zu beachten, dass auch der Besitzer des Abonnements Vorgaben welche auf Verwaltungsgruppenebene definiert wurden, nicht überschreiben kann. Durch das Anlegen mehrere Verwaltungsgruppen und untergeordneter Abonnements kann so die Organisationsstruktur abgebildet werden, beispielsweise indem pro Abteilung eine Verwaltungsgruppe bestimmt wird. Auch kann dadurch ein feingranulares Kosten-Reporting etabliert werden. Am Ende der Hierarchie befinden sich die eigentlichen **Ressourcen**. Dabei handelt es sich um alle buchbaren Dienste, wie virtuelle Maschinen, Speicher oder virtuelle Netzwerke. Zur vereinfachten Verwaltung werden diese in **Ressourcengruppen** zusammengefasst. Hierbei handelt es sich um ein vom Kunden definiertes Containerobjekt, welches alle „verwandten“ Ressourcen gruppiert. In der Regel sollten alle Ressourcen, welche für einen Dienst benötigt werden und/oder dem gleichen Softwarelebenszyklus unterliegen zur gleichen Ressourcengruppe hinzugefügt werden. Im Folgenden wird das Konzept der Abonnements näher beschrieben.



**Abb. 6.2** Organisationsstruktur in Microsoft Azure

## 6.2.2 Abonnement Modelle

Nachdem ein Azure Konto erstellt wurde, können noch nicht direkt Ressourcen angelegt werden. Zunächst muss innerhalb des Kontos ein Azure Abonnement erstellt werden. Das Azure Abonnement ist der zentrale Baustein einer jeden Azure Landschaft. Alle Ressourcen müssen genau einem Abonnement zugeordnet werden. Über das Abonnement erfolgen dann die meisten Verwaltungsvorgänge wie z. B. Kosten- und Rechnungsmanagement, die Definition rollenbasierter Zugriffskonzepte oder das Festlegen von Grenzwerten und Kontingente. Durch dieses Modell können Unternehmen beispielsweise eigenständige Abonnements für Projekte oder Teams nutzen. Jedes Abonnement kann dabei nur einem Azure Konto zugeordnet sein. Allerdings können beliebig viele Abonnements pro Account erstellt werden, um beispielsweise die Abrechnung aufzuteilen. Als Praxisempfehlung hat sich eine Trennung von SAP- und nicht-SAP-Workload, sowie eine Trennung von SAP Entwicklungsumgebungen und SAP Produktivumgebungen in eigenständige Abonnements bewährt.

Azure unterscheidet zwischen drei Abonnement Modellen, welche sich auf die Kostenabrechnung/Preispolitik, sowie auf die angebotenen Ressourcen und Dienstleistungen auswirken:

**Kostenlose Testversion** Ein Abonnement als Testversion kann kostenlos beantragt werden und beinhaltet einen 30-tägigen Probezugang mit 200 USD Budget (dieses wird bei Abschluss mit dem aktuellen Umrechnungskurs in die Landeswährung umgerechnet). Das kostenlose Abonnement steht allen Kunden zur Verfügung welche Azure bisher nicht kostenpflichtig verwendet haben bzw. auch noch keinen Testzugang abgeschlossen haben. Im Rahmen der Registrierung muss eine Kreditkarte angegeben werden. Dieses Angebot richtet sich vorrangig an Azure Neulinge die sich mit der Plattform vertraut machen möchten und eine Sandbox-Umgebung aufbauen möchten. Hierbei ist zu beachten, dass auch der Zugriff auf Dienstleistungen und Ressourcen stark eingeschränkt ist. Derzeit unterstützt der kostenlose Testzugang beispielsweise keine Instanzen mit mehr als 32 GB RAM, wodurch bestenfalls eine HANA Express Installation möglich ist. Das kostenlose Abonnement kann jederzeit in ein nutzungsbasiertes Modell (PAYG) umgewandelt werden (Upgrade). Unabhängig vom Upgrade verfällt das verbleibende Testbudget immer nach spätestens 30 Tagen.

**Pay-as-you-go (PAYG)** Beim PAYG-Modell fallen Kosten für Azure Services und Ressourcen nutzungsbasiert an und werden monatlich abgerechnet. Virtuelle Maschinen welche beispielsweise heruntergefahren und nicht länger allokiert sind verursachen somit keine aktiven „Computing Kosten“. Jedoch sollte beachtet werden, dass manche Dienste wie etwa eine Firewall oder Ressourcen wie z. B. verwaltete Datenträger weiterhin vorgehalten werden und somit aktiv Kosten verursachen. Auch das nutzungsbasierte Modell erfordert bei der Standard-Onlineregistrierung die Angabe einer Kreditkarte.

Eine Bezahlung per Lastschrift ist grundsätzlich ebenfalls möglich, erfordert jedoch eine Begründung und eine Anfrage beim Azure Support. Bezahlung auf Rechnung wird offiziell nicht unterstützt.

**Partner Angebote** Microsoft bietet eine Reihe attraktiver Angebote mit hohen Rabatten, welche Kostenvorteile gegenüber dem PAYG-Modell bieten können. Die Liste aller Angebote wird kontinuierlich aktualisiert und kann jederzeit unter folgender Seite eingesehen werden (Zugriff am 20.12.2021): <https://azure.microsoft.com/de-de/support/legal/offer-details/>. Wie bei der kostenlosen Testversion können Angebote mit Einschränkungen bezüglich der buchbaren Ressourcen und Dienste einhergehen.

### 6.2.3 Abonnement Management

Im Folgenden werden die Grundlagen des Abonnement Managements vermittelt. Dabei werden die Konzepte des Berechtigungsmanagements erläutert, die sich in verschiedene Administratorrollen kategorisieren lassen, sowie Grenzwerte, Budgets und Ausgabekräfte. Abschließend werden die Umsetzungsmöglichkeiten von Policies über das Abonnement Management beschrieben.

#### 6.2.3.1 Berechtigungsmanagement

Damit die Organisationsstruktur des Unternehmens auch in Azure abgebildet werden kann, bietet Microsoft ein umfangreiches Berechtigungskonzept um die Zugriffe auf die Abonnements, sowie deren zugehörigen Ressourcen zu steuern. Hierbei kann zwischen zwei Berechtigungsmodellen „Administratorrollen für klassische Abonnements“ und „Azure-Rollen“ unterschieden werden.

**Administratorrollen für klassische Abonnements** Hierbei handelt es sich um das ursprüngliche Berechtigungskonzept, welches eine rudimentäre Steuerung ermöglicht. Das Konzept besteht aus nur drei Rollen. Diese sind der Kontoadministrator, der Dienstadministrator, sowie einer beliebigen Anzahl an Co-Administratoren (maximal 200 pro Abonnement). Mit diesen drei Rollen ist die Zugriffssteuerung einfach umsetzbar und leicht verständlich, jedoch wenig flexibel und damit nicht für jedes Unternehmen geeignet. Diese Rollen richten sich somit vorrangig an kleine Organisationsstrukturen oder Azure Accounts, die für nicht-produktive Anwendungsszenarien verwendet werden. Nachfolgend werden die Rollen einzeln vorgestellt:

- **Kontoadministrator** (1 pro Azure Konto): Der Kontoadministrator ist die umfassendste Rolle. Sie berechtigt das Erstellen von neuen Abonnements, das Verwalten aller Abrechnungen, sowie das Zuweisen von Dienstadministrator-Rollen für jedes Abonnement. Pro Azure Konto kann es nur einen Kontoadministrator geben. Der Ersteller des Azure Kontos wird automatisch zum Kontoadministrator.

- **Dienstadministrator** (1 pro Azure Abonnement): Jedes Abonnement besitzt genau einen Dienstadministrator, welcher zum Verwalten des zugehörigen Abonnements berechtigt ist. Der Dienstadministrator hat Vollzugriff auf das Azure-Portal und kann deren Dienste verwalten, aber auch das Abonnement kündigen sowie Co-Administratoren für das Abonnement ernennen. Das Konto, welches zum Registrieren von Azure verwendet wird, wird standardmäßig sowohl Kontoadministrator als auch Dienstadministrator. Der Kontoadministrator kann den Dienstadministrator jedoch jederzeit ändern.
- **Co-Administrator** (200 pro Abonnement): Co-Administratoren werden ebenfalls auf Abonnement-Ebene definiert und besitzen identische Zugriffsrechte wie der Dienstadministrator. Diese können jedoch nicht die Zuordnung von Abonnements zu Azure Verzeichnissen ändern. Co-Administratoren können weitere Co-Administratoren ernennen, jedoch nicht den Dienstadministrator ändern.

**Rollenbasierte Zugriffssteuerung (Azure RBAC)** Um eine granulare Berechtigungsverwaltung zu ermöglichen, bietet Azure die rollenbasierte Zugriffssteuerung (Azure RBAC) an, welche auf dem Azure Resource Manager (ARM) basiert. Darin sind zahlreiche „Spezialrollen“ (aktuell mehr als 200) enthalten, welche sich an unterschiedliche Ressourcen richten und für diese Ressourcen vordefinierte Berechtigungen gewähren. Eine Auflistung aller Rollen und deren enthaltene Zugriffe findet sich in den Microsoft Docs [1]. Unabhängig von diesen granularen Rollen gibt es noch drei Basis Rollen, welche für alle Ressourcentypen gelten:

- **Besitzer:** Als Besitzer erhält man den Vollzugriff auf alle Ressourcen und kann gleichzeitig den Zugriff an andere Benutzer delegieren. Dienstadministratoren und Co-Administratoren (klassische Abonnement Administratorrollen) erhalten beispielsweise die Besitzerrolle für ihren Abonnement-Bereich.
- **Mitwirkender:** Mitwirkende können ebenfalls alle Arten von Azure Ressourcen erstellen und verwalten. Jedoch können sie diesen Zugriff nicht weiter delegieren.
- **Leser:** Leser sind lediglich zum Anzeigen aller Azure Ressourcen berechtigt.

Mit diesen Basis-Rollen ist die Berechtigungsverwaltung nicht wesentlich flexibler als das klassische Modell. Eine fein-granulare Steuerung wird erst durch den zusätzlichen Einsatz der Spezialrollen ermöglicht. Außerdem sollte noch die Rolle „Benutzerzugriffsadministrator“ erwähnt werden. Diese ist notwendig um den Benutzerzugriff auf Azure Ressourcen zu verwalten.

### 6.2.3.2 Grenzwerte, Budgets und Ausgabenlimits

Azure bietet zur Kostenkontrolle zahlreiche Limits und Grenzwerte an, welche auf Abonnement-Ebene definiert werden. Dadurch können sowohl die Nutzung korrekter Ressourcen als auch die anfallenden Kosten zentral gesteuert werden.

**Azure-Ausgabenlimit** Das Ausgabenlimit ist für alle Abonnementtypen aktiviert, welche über ein Guthaben verfügen, wie beispielsweise das kostenlose Testabonnement mit 200 US\$. Das Limit fungiert als „Hard Limit“ und ist immer identisch zum jeweiligen Guthaben. Es kann weder erhöht noch gesenkt werden (eine Entfernung des Limits ist jedoch möglich). Sobald die angefallenen Kosten den Wert des Ausgabelimits erreichen, werden alle Ressourcen und Dienste gestoppt, sodass keine weiteren Kosten mehr anfallen. Das bedeutet, dass laufende VMs heruntergefahren und deren Zuordnung aufgehoben werden. Dadurch wird die VM nicht gelöscht, aber die Reservierung auf der tatsächlichen Hardware wird freigegeben. Außerdem ist der Zugriff auf Speicherkonten nur noch schreibgeschützt möglich. Ein derartige „harte Abschaltungstrategie“ eignet sich nicht für produktive Einsatzzwecke. Ausgabenlimits können jedoch auch nicht für reguläre nutzungsbasierte Abonnements aktiviert werden.

**Azure-Budgets** Ähnlich zum Ausgabenlimit definiert ein Budget einen Grenzwert für anfallende Kosten. Hierbei handelt es sich jedoch um ein „Soft Limit“, denn Budgets können frei definiert werden und haben standardmäßig lediglich eine Alert-Funktion. D. h. im Unterschied zum Ausgabenlimit werden bei Überschreitung des Budgets keine automatischen Stop-Mechanismen gestartet. In der Regel werden Budgets in Kombination mit Kostenschwellenwerten definiert (z. B. 80 %), welche eine Benachrichtigung senden sobald die angefallenen Kosten den Schwellenwert des Budgets erreichen. Mittels der „Azure Monitor-Aktionsgruppen“ können auch weiterführend Aktionen orchestriert werden die beispielsweise ein geordnetes Herunterfahren bestimmter Ressourcengruppen verursachen. Eine Schritt-für-Schritt-Anleitung zur Kopplung von Schwellenwerten mit einem Azure Automation-Runbook, welches VMs automatisch beendet, findet sich in den Microsoft Docs [2].

### Ressourcen Grenzwerte und Kontingente

Kontingente werden auf Ebene des Abonnements definiert und bestimmen Obergrenze zur Nutzung von Ressourcen. Sie können sowohl für bestimmte SKUs wie z. B. „vCPUs der ESv3-Familie“, als auch für bestimmte Ressourcenarten z. B. Anzahl virtueller Computer Kontingente definieren. Im Rahmen der „Überprüfen+Erstellen“ Phase während der Bereitstellung werden dann die von Ihnen angeforderten Ressourcen mit den Kontingenten abgeglichen und im Falle eines Überschreitens die Anforderung abgelehnt. Das Anheben eines Kontingents erfolgt mittels Support Anforderung. Sie müssen jedoch hierfür kein manuelles Ticket anlegen, sondern lediglich eine Erhöhung des Kontingents anfordern. Die Überprüfung durch den Support erfolgt im Hintergrund automatisch.

#### 6.2.3.3 Policies

Azure Policies sind ein nützliches Instrument, um Organisationsstandards zu etablieren. Insbesondere in Azure Umgebungen mit mehreren Teams oder Abonnements kann es

notwendig werden Standards zu definieren. Diese beinhalten unter anderem Regeln, wie Ressourcen konfiguriert und verwendet werden dürfen, beispielsweise um Sicherheit zu gewährleisten oder Compliance Anforderungen gerecht zu werden. Um die Notwendigkeit für Azure Policies zu verstehen, lohnt es sich, zunächst den traditionellen Governance Ansatz zu betrachten. Der naheliegendste Weg, welcher von vielen Kunden verwendet wird, ist es dedizierte Personen als „Azure Verwalter“ einzusetzen. Diesen obliegt dann die vollständige Verwaltung von Azure inklusiver aller Ressourcen und Abonnements, sowie der alleinige Zugriff auf das Azure Portal. Der Zugang auf die jeweiligen Ressourcen wird dann durch den Verwalter an die Endanwender wie Entwickler und Administratoren bereitgestellt. Im SAP On-Premise Hosting sind diese klassischen Aufteilungen von Verantwortlichkeiten zwischen Fachabteilungen (z. B. Basis und Entwicklung) sehr vorbereitet und werden direkt als Governance Struktur für die Cloud übernommen. Auf diese Weise verliert der Kunde jedoch viel Innovationspotential, da der eigentliche Nutzerkreis von den Leistungen der Azure Cloud abgeschirmt sind. Insbesondere im Kontext von Cloud-native und DevOps-Szenarien wird dies immer wichtiger. Policies sollten nicht mit Berechtigungskonzepten wie RBAC verwechselt werden. RBAC definiert welche Benutzer bzw. Rollen Zugriff auf welche Ressourcen haben. Policies hingegen steuern welche Eigenschaften und Konfiguration Ressourcen haben müssen. Eine Policy kann z. B. definieren, dass nur der VM-Instanztyp D provisioniert werden darf. Eine Policy besteht dabei immer aus der Definition welche das gewünschte Verhalten beschreibt (Einschränkung auf VM-Instanttypen) und einer Parametrisierung (Instanz D). Je nachdem auf welchen Bereich die Policy angewandt wird kann diese auch mehrfach mit unterschiedlichen Parametrisierungen verwendet werden (z. B. im Abonnement „SAP HANA Sandbox Systems“ dürfen nur Instanzen des Typs M provisioniert werden). Policies werden im JSON-Format definiert und nutzen eine eigene Domänsprache, welche im Wesentlichen auf der Kombination von Logikoperatoren (and, not,...), Feldern (Tag, Location, Name) und Bedingungen (equal, contains, exist) basiert. Es müssen jedoch nicht von Grund auf eigene Policies definiert werden. Microsoft Docs bietet eine große Auswahl an vordefinierten Beispielen zu allen Ressourcen an. Darin sind auch Policy Sets zur Einhaltung von internationalen Standards wie ISO 27001 (IT Sicherheitsmanagement) enthalten.

---

## 6.3 Azure Ressourcenmanagement

In diesem Kapitel wird das Ressourcenmanagement in Microsoft Azure erläutert. Dabei wird auf die verschiedenen Ressourcenbereitstellungsmodelle eingegangen sowie mögliche Bereitstellungsoptionen. Darauf aufbauend werden die Grundlagen für die Bereitstellung von Rechenleistung, Speicher sowie Netzwerk- und weitere Serviceleistungen gegeben. Abschließend werden Hinweise zum Support und der Lizenzierung vorgestellt.

### 6.3.1 Ressourcenbereitstellung mittels Resource Manager

Für Microsoft Azure existieren zwei Arten für Bereitstellungsmodelle von VM-Ressourcen: die klassische Bereitstellung und die Bereitstellung über den Azure Resource Manager (ARM).

#### 6.3.1.1 Klassisches vs. ARM Ressourcen Management

Ursprünglich konnten Ressourcen nur über das klassische Bereitstellungsmodell, dem sogenannten Azure Service Manager, erstellt und verwaltet werden, wobei eine Gruppierung von Ressourcen nicht möglich war. Als Resultat müssen bei diesem Bereitstellungsmodell diejenigen Ressourcen, welche zu einer Anwendung gehören, durch geeignete Benennung oder eine selbsterstellte Übersicht zusammengefasst werden. Nur so können alle Beteiligten manuell die Zusammengehörigkeit von Ressourcen zur Bereitstellung einer Anwendung nachvollziehen. Zusätzlich ist eine Erstellung von Ressourcen lediglich einzeln über das Azure Portal möglich. Eine Automatisierung der Schritte kann nur über die Entwicklung und Nutzung von Skripten erreicht werden. Die Zuordnung von Tags zu verschiedenen Ressourcen ist nur anhand des ARMs möglich, wodurch eine bessere Transparenz der genutzten Ressourcen ermöglicht wird.

Seit 2014 kann der ARM genutzt werden, womit Rechen-, Netzwerk- und Speicherressourcen verwaltet werden können. Einhergehend damit wurden Ressourcengruppen eingeführt, welche zugehörige Ressourcen zusammenfassen. Die Nutzung des Resource Managers wird von Microsoft empfohlen, da sie eine einfachere Verwaltung und Bereitstellung von Ressourcen ermöglicht.

Anhand des Resource Managers können Sie zusammengehörige Ressourcen gemeinsam erstellen und verwalten. Damit ist die Ressourcenbereitstellung wesentlich effizienter und bietet die Möglichkeit der homogenen Ressourcenerstellung. Weiterhin kann die Zugriffssteuerung auf alle Ressourcen einer Gruppe angewendet und Abhängigkeiten zwischen diesen Ressourcen definiert werden. Über die Erstellung einer ARM-Vorlage mithilfe der JavaScript Object Notation (JSON) kann eine Vorlage für die Ressourcen-Infrastruktur einer bestimmten Lösung definiert werden. Die Zuordnung von Tags zu Ressourcen, wurde Kunden im Rahmen des Resource Managers nun ermöglicht.

Seit Einführung des ARMs werden alle Ressourcen zu Standardressourcengruppen zugeordnet. Das ist auch für Ressourcen der Fall, welche über die klassische Bereitstellung erstellt wurden oder werden. Diese werden dann in eine Standardressourcengruppe des jeweiligen Dienstes eingefügt, auch wenn keine Ressourcengruppe bei der Erstellung definiert wurde. Sie sollten jedoch beachten, dass Ressourcen, welche über das klassische Bereitstellungsmodell erstellt wurden nicht automatisch in das Resource Manager Modell migriert werden. Das ist auch nicht der Fall, wenn diese Ressourcen sich in einer Ressourcengruppe befinden.

Generell gilt, dass Ressourcen (Compute-, Netzwerk- und Speicherressourcen), die über die klassische Bereitstellung erstellt wurden im Rahmen des klassischen Modells verwaltet werden. Das gleiche gilt für die Erstellung von Ressourcen über den ARM:

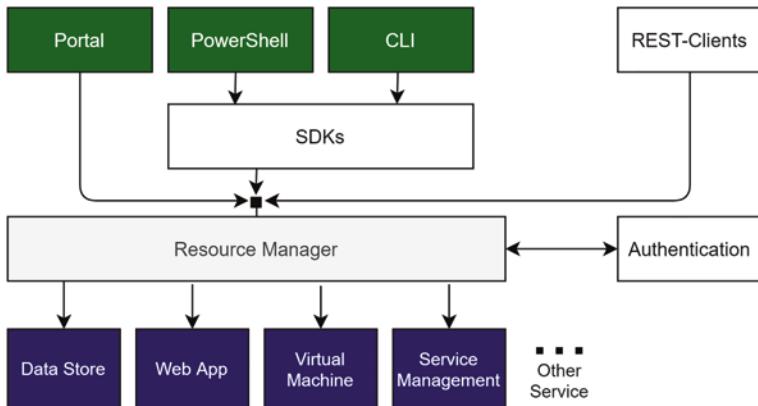
werden Ressourcen über dieses Modell erstellt, muss die Verwaltung auch über das Modell des Resource Managers erfolgen. Die Komplexität der Ressourcenverwaltung steigt für den Fall, dass Ressourcen teilweise über das klassische Bereitstellungsmodell erstellt wurden und teilweise über den Resource Manager. Das sollten Sie bei der Umsetzung berücksichtigen, da beide Bereitstellungsmodelle unterschiedliche Verwaltungs-Vorgänge unterstützen. Als Beispiel können Ressourcen, die über die klassische Bereitstellung erstellt wurden, mithilfe eines Resource Manager-Befehls in andere Ressourcengruppen verschoben werden. Das bedeutet jedoch nicht, dass mit dieser Ressource andere ARM-Prozesse möglich sind.

### 6.3.1.2 Migration klassischer Ressourcen zu ARM

Seit dem 28. Februar 2020 hat Microsoft Azure begonnen, den Azure Service Manager über einen Zeitraum bis zum 1. März 2023 abzuschalten. Dabei ist es wichtig zu beachten, dass nach abschalten des Azure Service Managers alle VMs, die über das klassische Bereitstellungsmodell erstellt wurden, abgeschaltet werden und nicht mehr nutzbar sind. Seit Februar 2020 können deshalb VMs nicht mehr über den Azure Service Manager erstellt werden. Ab dem 1. März 2023 können die VMs, welche über den Microsoft Azure Service Manager bereitgestellt wurden, nicht mehr gestartet werden. Noch laufende virtuelle Instanzen werden beendet und deren Zuordnungen werden gelöst. Die geplante Löschung dieser VMs wird an die Nutzer weitergegeben, um den Nutzern eine Migration der VMs in das verbleibende Bereitstellungsmodell des ARMs zu ermöglichen. Davon nicht betroffen sind die folgenden Services:

- Azure Cloud Services (klassisch): stellen Platform-as-a-Service Produkte von Microsoft Azure dar, die über das klassische Bereitstellungsmodell erstellt wurden
- Speicherkonten, die nicht von Infrastructure-as-a-Service VMs (klassisches Bereitstellungsmodell) genutzt werden
- Virtuelle Netzwerke, die nicht von Infrastructure-as-a-Service VMs (klassisches Bereitstellungsmodell) genutzt werden
- Sonstige Ressourcen (klassisches Bereitstellungsmodell)

Bei der Migration der klassisch bereitgestellten VMs in den Azure Resource Manager müssen Sie folgendes berücksichtigen. Da die Datenstruktur beim Azure Service Manager und Azure Resource Manager gleich ist, sich jedoch die Darstellung der Daten bei beiden Bereitstellungsmodellen unterscheidet, muss die Daten-Darstellung der klassischen VMs bei der Migration modifiziert werden. Das bedeutet, dass neue APIs, SDKs und Tools für die Migration erforderlich sind. Deshalb sollten Sie sicherstellen, dass Ressourcen, die Sie migrieren möchten, keine nicht unterstützten Einstellungen und Funktionen enthalten. Durch eine Prüfung kann Microsoft Azure diese Fehler erkennen und Sie informieren. Abb. 6.3 stellt das Zusammenspiel von ARM mit den bereitgestellten Tools, APIs und SDKs dar. Aufgrund der geplanten Einstellung des Azure Service Managers, fokussiert sich das Buch auf die Nutzung des Azure Resource Managers.



**Abb. 6.3** Zusammenspiel von ARM mit Tools, APIs und SDKs

### 6.3.1.3 Bereitstellung mit dem Azure Resource Manager

ARM fungiert als zentraler Bereitstellungsdiensst um Ressourcen zu erstellen, zu ändern oder zu löschen. Es unterstützt dabei vier APIs um Benutzeranfragen entgegenzunehmen.

**Das Azure Portal** ist der gängige Weg um interaktiv über ein grafisches Front End Ressourcen zu verwalten. Anfragen werden mittels HTTP(s) als REST Anfragen an den ARM übermittelt.

**Die REST API** des ARM kann jedoch unabhängig auch direkt angesprochen werden ohne das Azure Portal zu nutzen. So können etwa eigene web-basierte Automatisierungs-lösungen zur Verwaltung der Azure Landschaft entwickelt werden.

Als **Kommandozeilentools** unterstützt ARM zudem die **Azure Powershell** sowie die **Azure CLI**. Beide Tools sind vom Funktionsumfang vergleichbar, wobei die Azure CLI als plattformunabhängiges Tool (Mac/Linux/Windows) entwickelt wurde, welches neben Bash auch PowerShell und Windows Command Prompt als Shell Umgebung unterstützt. Sie können die Azure CLI auf Windows, Linux und MacOS installieren. Befolgen Sie dazu ganz einfach die Anweisungen in den Microsoft Docs [3].

Anstelle die Azure CLI lokal z. B. auf einer VM zu installieren, bietet Microsoft auch die **Azure Cloud Shell** an. Diese kann entweder direkt über den Browser aufgerufen werden [4] oder über einen Button im Azure Portal. Zudem bietet die Cloudshell sowohl eine Powershell als auch eine Bash Darstellung. Sollten Sie über das Portal darauf zugreifen ist es nicht notwendig, dass Sie sich zunächst authentifizieren. Zur Nutzung der Cloudshell ist jedoch zwingend ein Speicherkonto notwendig. Mehr zur Nutzung der Cloudshell erfahren Sie in Abschn. 7.3.3.4.

ARM arbeitet mit einer deklarativen Syntax. Es wird also lediglich spezifiziert was erstellt werden soll ohne die hierfür notwendigen Schritte zu beschreiben. Dadurch unterstützt ARM insbesondere die **Vorlagen-basierte Bereitstellung**. Ein ARM Vorlage ist typischerweise ein JSON File welche die Komponentenzusammenstellung vollständig beschreibt. Durch Separierung der „Parameter Sektion“ von der „Ressourcen“

Sektion“ können ARM Vorlagen wiederholt verwendet werden um schnell und bequem Ressourcenzusammenstellung mit unterschiedlichen Ausprägungen bereitzustellen. Mehr zur Bereitstellung mittels ARM Vorlagen erfahren Sie in Abschn. 7.3.2.

### 6.3.2 Allgemeine Bereitstellungsoptionen

Die Grundlagen der allgemeinen Bereitstellungsoptionen lassen sich in verschiedene Kategorien untergliedern: Regionen und Verfügbarkeit, Azure Blueprints, Verfügbarkeitsgruppen, Näherungsplatzierungsgруппen und Skalierungsgruppen. Diese werden in den nächsten Abschnitten näher umrissen.

#### 6.3.2.1 Regionen und Verfügbarkeit

Azure bietet weltweit verteilte Rechenzentren auf allen fünf Kontinenten. Auch wenn die tatsächliche physische Lokation eines Dienstes für den Kunden transparent ist, so muss zumindest die Hosting Region z. B. „Europa Westen (Niederlande)“ oder „Deutschland Norden (Berlin)“ angegeben werden. Dies kann nicht nur aus Compliance Gründen relevant sein, sondern beeinflusst auch die Hosting Kosten, da regional erhöhte Betriebskosten (z. B. Strom) an den Kunden weitergegeben werden. Auch sind bestimmte Dienste nicht in allen Regionen verfügbar. Eine der wichtigsten Einschränkungen für den SAP-Betrieb ist hierbei die Option der **Verfügbarkeitszonen**. Bei Verfügbarkeitszonen handelt es sich um einzelne Rechenzentren innerhalb einer Region. Diese zeichnen sich durch eigenständige Infrastrukturversorgung (Strom, Netzwerk, Kühlung) sowie eine räumliche Trennung voneinander aus. Redundanz- und Hochverfügbarkeitslösungen welche über mehrere Verfügbarkeitszonen repliziert sind, sind somit vor Ausfällen innerhalb einer Zone bzw. einem Rechenzentrum geschützt. Nicht alle Regionen verfügen jedoch über eigenständige Verfügbarkeitszonen (Beispielsweise verfügt „Deutschland, Westen-Mitte“ über mehrere Verfügbarkeitszonen, nicht jedoch *Deutschland Norden*). Bei dieser Redundanzoption sollte daher geprüft werden, ob die gewünschte Region Verfügbarkeitszonen unterstützt. Eine laufend aktualisierte Übersicht aller Regionen findet sich unter (Zugriff am 20.12.2021): <https://docs.microsoft.com/de-de/azure/availability-zones/az-region>.

#### 6.3.2.2 Azure Blueprints

Azure Blueprints ermöglichen einfache Bereitstellungen, indem Artefakte deklarativ über eine Blaupause definiert werden. Diese Artefakte können neben ARM-Vorlagen auch Richtlinien und rollenbasierte Zugriffsbeschränkungen beinhalten. Azure Blueprints ersetzen damit nicht die bestehenden ARM-Vorlagen sondern stellen vielmehr eine Erweiterung dar. Der Fokus der Blaupausen liegt unter anderem anderem auf der Orchestrierung komplexer Bereitstellung, sodass die individuellen Vorgänge darin überprüft und nachvollzogen werden können. Des Weiteren ermöglichen sie das Zusammenstellen umfangreicher Ressourcenkonfigurationen welche insbesondere das komplexe

Zusammenspiel zwischen IT-Governance und IT-Applikation erleichtern indem die Blaupause die Logik für Kontingente, Rollen und Zugriffsbeschränkungen kapselt. Das zentrale Cloud Architekturteam kann auf diese Weise z. B. Blaupausen für verschiedene Rollen z. B. (DEV/QAS) oder verschiedene Business Units erstellen. Blaupausen werden bei Erstellung einem Abonnement oder einer Verwaltungsgruppe zugeordnet und werden darauf gespeichert. Parameter für ein Artefakt, welches einer Blaupause zugeordnet wird, können entweder benutzerspezifisch (in der Blaupause nicht definiert) oder fest vorgegeben sein. Auf diese Weise wird sichergestellt das bestimmte Werte wie z. B. die Region nicht überschrieben werden kann.

### **6.3.2.3 Verfügbarkeitsgruppen**

Anhand von Verfügbarkeitsgruppen können virtuelle Maschineninstanzen auf mehrere Hardware-Endpunkte verteilt werden, wodurch die Ausfallsicherheit höher ist. Hardware in einem Rechenzentrum kann in Update- und Fehlerdomänen unterteilt werden. Die Updatedomäne stellt eine Gruppe von Maschineninstanzen dar, welche die gleiche Hardware nutzen und somit gleichzeitig gestartet und beendet werden können. Fehlerdomänen umfassen virtuelle Maschinen, welche den gleichen Speicher, Strom und Netzwerkswitch teilen. Um eine hohe Verfügbarkeit zu generieren, ist es sinnvoll, mindestens zwei Maschinen in einer Verfügbarkeitsgruppe zu erstellen. Nachdem die virtuelle Maschine erstellt wurde, kann die Verfügbarkeitsgruppe nicht mehr geändert werden, da die Verteilung der Hardwarekomponenten nicht mehr sinnvoll erfolgt. Um die Verfügbarkeitsgruppe zu ändern muss somit die Instanz gelöscht und neu erstellt werden. Es gilt zu beachten, dass bei der Bereitstellung von zwei Maschineninstanzen einer Verfügbarkeitsgruppe die gleiche Fehlerdomäne erhalten können. Um das zu vermeiden, sollten die erstellten Maschineninstanzen während der Erstellung der anderen virtuellen Maschinen nicht beendet oder freigegeben werden. Bevor die virtuellen Maschinen erstellt werden, können die möglichen VM-Größen in der Verfügbarkeitsgruppe eingesehen werden. Die möglichen VM-Größen hängen von der genutzten Hardware der Verfügbarkeitsgruppe ab. Das ist anhand des Befehls Get-AZVMSIZE möglich. Es ist auch möglich Verfügbarkeitsgruppen innerhalb von Verfügbarkeitszonen bereitzustellen. Dies erfordert jedoch, dass zusätzlich Näherungsplatzierungsgruppen verwendet werden.

Im Kontext von SAP sollten alle Maschinen innerhalb einer Verfügbarkeitsgruppe die gleiche Rolle aufweisen z. B. eine eigene Verfügbarkeitsgruppe für die Applikationsserver und eine eigene Verfügbarkeitsgruppe für die ASCS und ERS.

### **6.3.2.4 Näherungsplatzierungsgruppen**

Eine Näherungsplatzierungsgruppe (Proximity placement Group) stellt sicher, dass Ressourcen, welche sich innerhalb dieser Gruppe befinden, in „räumlicher Nähe“ zueinander bereitgestellt werden. Dadurch werden Computer z. B. im gleichen Rechenzentrum platziert und sichergestellt, dass die Latenz zwischen den Komponenten

geringgehalten wird. Ein typisches Anwendungsszenario für Näherungsplatzierungsgruppen ist es alle Instanzen einer S/4HANA Installation (PAS, ASCS, AAS), sowie die Datenbank in eine Gruppe zu setzen. Näherungsplatzierungsgruppen können auch mit Verfügbarkeitsgruppen kombiniert werden. Da hiermit jedoch der Bereitstellungsprozess mit zusätzlichen Einschränkungen versehen wird, kann dieser leichter fehlschlagen, falls die Anforderungen nicht erfüllt werden können.

### 6.3.2.5 Skalierungsgruppe

Eine VM-Skalierungsgruppe ist ein Containerobjekt zur Zusammenstellung mehrerer virtueller Computer in Kombination mit einer Skalierungsrichtlinie, einem Lastenausgleich sowie diverser Automatisierungsfeatures. Basierend auf einer fixen Ausgangsanzahl und Instanzen kann die Gruppe hochskaliert werden um höheren Lastanforderungen gerecht zu werden. Die Skalierungsgruppe unterscheidet dabei zwischen der Skalierungsrichtlinie „manuell“, bei der die Skalierung händisch vom Benutzer durchgeführt wird und „Benutzerdefiniert“ welche eine Autoskalierung basierend auf einer Metrik ermöglicht. Auf diese Weise kann ein CPU-Schwellenwert z. B. 90 % in Kombination mit einer Zeitdauer z. B. 15 min gesetzt werden (erlaubte Werte sind 5–60 min). Sobald der Schwellenwert innerhalb der angegebenen Zeitdauer beibehalten wird, wird die Skalierungsgruppe um eine definierte Anzahl an virtuellen Computern erweitert, wobei eine maximale Anzahl an Instanzen definiert werden kann. Auch eine Metrik für das Abskalieren kann auf identische Weise gesetzt werden. Die Skalierungsgruppe selbst ist unentgeltlich, lediglich für das Bereitstellen der virtuellen Instanzen fallen Kosten an. Beachten Sie, dass Sie anstelle einer „Skalierungsgruppe“ auch manuelle mehrere VMs gruppieren können und mittels weiterer Komponenten wie einem Lastenausgleich Scale-out Architekturen umsetzen könne. Die Stärke der Skalierungsgruppe liegt in ihrem Fokus auf einheitlichem Automatismus wie z. B. die Integrations von „Continous Delivery“ pipelines. Aus diesem Grund eignet sich die Skalierungsgruppe auch eher für Container basierte Workloads und weniger für „hand gepflegte“ VMs wie diese häufig bei SAP-Systemen zum Einsatz kommen.

## 6.3.3 Rechenleistung

Das Kapitel behandelt mögliche Sizing-Optionen und die verfügbaren Maschinentypen für HANA- und Nicht-HANA-Systeme. Da sich die Maschinentypen häufig ändern, konzentriert sich das Kapitel eher auf die VM-Vorlagenklassen als auf spezifische Maschinen. Für die Bereitstellung einer VM für die Nutzung von SAP-Software sind mehrere Aspekte zu berücksichtigen. Schlussendlich müssen alle Voraussetzungen für die Installation der SAP-Software auf VM-Instanzen in Microsoft Azure erfüllt werden, um eine reibungslose Nutzung der Software-Komponenten sicherzustellen. Diese Voraussetzungen und Aspekte werden im Folgenden näher beschrieben.

### 6.3.3.1 Bereitstellungsmodelle

Generell existieren verschiedene Möglichkeiten zur Bereitstellung von S/4HANA Systemen in Microsoft Azure. In den folgenden Unterkapiteln werden die Bereitstellungen über Azure VMs, über Azure große Instanzen und über die SAP Cloud Appliance Library (CAL) näher beschrieben.

#### 6.3.3.1.1 Azure VMs

Traditionell können Sie zur Implementierung von SAP Software in Microsoft Azure virtuelle Computerinstanzen nutzen. Diese stellen die notwendigen Ressourcen zur Verfügung um die entsprechenden Anwendungen zu betreiben, wobei einige Tools von Azure zusätzlich genutzt werden können um Hochverfügbarkeit und Skalierbarkeit zu ermöglichen. Dabei können die Computerinstanzen über das klassische Bereitstellungsmodell oder über den Resource Manager erstellt und verwaltet werden. Allerdings existieren Einschränkungen bezüglich des Betriebssystems und verwendeten Datenbanksystems. Außerdem können nicht alle Systemarchitekturen, die On-Premise unterstützt werden, auch in Azure realisiert werden. In der folgenden Tabelle wird ein Überblick für nicht unterstützte SAP-Systemarchitekturen in Azure gegeben. Um solche Systemarchitekturen in Azure aufzubauen, muss auf einen Drittanbieter zurückgegriffen werden (Tab. 6.1).

Um eine virtuelle Computer-Instanz bereitzustellen, werden mehrere Komponenten benötigt. Die folgende Tabelle gibt dir einen Überblick dieser Komponenten (Tab. 6.2).

Die Nutzung von Azure Site Recovery um die virtuellen Instanzen der Datenbankebene zu replizieren wurde bisher nicht getestet, weshalb noch keine Erfahrungswerte diesbezüglich existieren.

#### 6.3.3.1.2 Azure große Instanzen für SAP HANA

Neben der Nutzung von VMs zur Bereitstellung von SAP Software bietet Azure auch ein dediziertes Angebot namens „Azure large instances (HLI)“. Bei diesem für HANA zertifizierten tailored data center integration (TDI) Ansatz kann der Kunde eine SAP HANA auf einem Bare-Metal-Server bereitstellen. Die hierdurch bereitgestellten Größen und Sizings unterscheiden sich von den klassischen VM-Vorlagen und bieten größere Flexibilität von 36 Kernen mit 768 GB Arbeitsspeicher (kleinste Einheit) bis zu 480 Kernen und 24 TB Arbeitsspeicher (größte Einheit). Auf diese Weise ist es möglich HANA Scale-Out Bereitstellungen mit bis zu 120 TB RAM umzusetzen. Außerdem wird durch den Wegfall der Virtualisierungsschicht in HLI eine noch höhere Performance erzielt als dies bei Azure VMs der Fall ist. HLI SKUs starten typischerweise mit dem Prefix S\*. Da mit steigender Datenkapazität auch typischerweise die HANA Arbeitsspeicher-Anforderungen steigen, bietet HLI ähnlich zu Azure VMs an, zuerst mit kleineren Instanzen zu starten, um später auf größere Versionen zu wechseln. Grundsätzlich differenziert Azure bei den HLI Instanzen zwischen den Klassen I und II, wobei I die Einstiegsklasse bildet und Klasse II für fortgeschrittene Szenarien geeignet sind.

**Tab. 6.1** Nicht-unterstützte Systemarchitekturen

Szenario	Hinweise
Speichersoftwareappliance	In Azure werden verschiedene Softwareappliances angeboten. Die Unterstützung von Bereitstellungen im Zusammenhang mit SAP-Software muss vom Drittanbieter gewährleistet werden. (Siehe auch SAP Note 2015553)
Hochverfügbarkeitsframeworks	Nur Pacemaker und Windows Server-Failovercluster werden für SAP-Systeme in Azure unterstützt. Andere Produkte, wie beispielsweise Datakeeper von SIOS müssen über den jeweiligen Anbieter unterstützt werden
Cluster mit gemeinsamen Datenträgern für Datenbankdateien	Lediglich MaxDB wird bei diesem Szenario unterstützt. Bei anderen Datenbanken können nur separate Speicherorte verwendet werden um ein Hochverfügbarkeitsszenario zu erreichen, das von Azure unterstützt ist
Szenarien, in denen es zu großer Netzwerklatenz zwischen der SAP-Anwendungsschicht und der SAP-DBMS-Ebene kommt	Beispiele hierfür sind: <ul style="list-style-type: none"> <li>Bereitstellung eines S/4HANA Systems, wobei eine der Ebenen On-Premise und die andere Ebene in Azure bereitgestellt wird</li> <li>Bereitstellung der Anwendungsschicht und der Datenbankschicht eines Systems in verschiedenen Regionen in Azure</li> <li>Bereitstellung virtueller Netzwerkappliances zwischen der Anwendungs- und Datenbankschicht</li> <li>Speichernutzung für die Datenbank oder das Transportverzeichnis, der in parallel zu Azure verwendeten Rechenzentren bereitgestellt wird</li> </ul> Nutzung von zwei verschiedenen Cloud-Anbietern für die Implementierung der Anwendungs- und Datenbankschicht
HANA-Pacemaker-Cluster	HANA-Pacemaker-Cluster mit mehreren Instanzen
Windows-Cluster	Windows-Cluster mit gemeinsamen Datenträgern über Azure NetApp Files für unterstützte SAP-Datenbanken in Windows. Besser ist die Nutzung einer Hochverfügbarkeitsreplikation der Datenbanken
SAP-Datenbanken unter Linux mit NFS-Freigaben	SAP-Datenbanken mit NFS-Freigaben zusätzlich zu Azure NetApp Files. Das gilt nicht für SAP HANA, Oracle auf Oracle Linux und Db2 auf SUSE und Red Hat
Oracle DB weder auf Windows noch auf Oracle Linux	Implementierung von Oracle Datenbanken auf einem anderen Betriebssystem als Windows oder Oracle Linux. Siehe auch SAP Note 2039619

HLI ist ausschließlich zur Provisionierung von SAP HANA Systemen entworfen worden und kann nicht für andere Hostings genutzt werden. Aus diesem Grund ist auch die Wahl der verfügbaren Betriebssysteme auf Suse Linux Enterprise und Redhat Linux eingeschränkt. Bereits bei der Provisionierung müssen SAP-spezifische Angaben wie z. B. die SID Microsoft mitgeteilt werden. Die Instanzen sind direkt an das Azure Backbone angebunden und bieten dadurch geringe Latenz zu den virtuellen Maschinen

**Tab. 6.2** Komponentenbeschreibung für virtuelle Computer-Instanzen in Microsoft Azure

Komponente	Beschreibung
Virtuelles Netzwerk	Das virtuelle Netzwerk wird über den Resource Manager verfügbar gemacht
Speicherkonto	Anhand des Speicherkontos können die Datenträger im BLOB-Speicher langfristig gesichert werden
Verfügbarkeitsgruppen	Verfügbarkeitsgruppen verteilen die erstellten virtuellen Maschinen auf unterschiedliche Hardware-Endpunkte, um somit die Fehleranfälligkeit zu reduzieren
Lastenausgleich	Lastenausgleich kann über den Azure Load Balancer ermöglicht werden. Dieser verteilt den eingehenden Datenverkehr auf die virtuellen Maschinen
Virtuelle IP-Adresse	Es besteht die Möglichkeit, die öffentliche IP-Adresse statisch oder dynamisch zu definieren, wobei dynamische IP-Adressen mit einem Lastenausgleichsmodul versehen werden können. Mithilfe von Sicherheitsgruppen können öffentliche IP-Adressen geschützt werden
DNS-Name	DNS-Namen können bei öffentlichen IP-Adressen definiert werden. Der FQDN lautet: <domainlabel>.<region>.cloudapp.azure.com
Netzwerkschnittstellen	Netzwerkschnittstellen verweisen auf die IP-Adresse der zugeordneten VM, das Subnetz des virtuellen Netzwerks der VM und optional auf die Netzwerksicherheitsgruppe der VM

in denen die SAP Anwendungsserver bereitgestellt werden, wobei die typische RTT innerhalb eines virtuellen Netzwerks i. d. R. geringer ist. Als Gateway Lösung muss UltraPerformance Gateway gewählt werden.

Bei der Buchung von HLI Instanzen muss darüber hinaus keine eigenständige Speicherprovisionierung vorgenommen werden. HLI Instanzen werden mit festem Plattspeicher gemäß den Empfehlungen von SAP bereitgestellt. Als Daumenregel beträgt die Größe der Datenträger i. d. R. das Vierfache des Arbeitsspeichers (gilt nicht für Typ-II Instanzen). Weitere Kapazität lässt sich mittels separater Datenträger in 1 TB Schritten zur Instanz hinzufügen.

### 6.3.3.1.3 SAP Cloud Appliance Library (CAL)

Die SAP Cloud Appliance Library (SAP CAL) bietet ein Online-Bibliothek mit konfigurierten SAP-Lösungen, die sofort in der Cloud genutzt werden können. Dies ermöglicht es innerhalb kürzester Zeit Test-, Demo-, und Entwicklungssysteme bereitzustellen.

Die SAP Cloud Appliance Library, auch CAL abgekürzt, ermöglicht es in Microsoft Azure ein SAP-System mit wenig Aufwand zu erstellen. Dabei gibt es verschiedene Lösungen, die über die CAL in Azure bereitgestellt werden können: Für SAP S/4HANA gibt es beispielsweise über die *SAP S/4HANA 2020 FPS02 Fully-Activated Appliance* die Möglichkeit, die Version 2020 (FPS02) mit bereits durchgeführten Best Practices

Konfigurationen und bereitgestellten Kernfunktionen zu nutzen. Hierbei kann der Zugriff auf das System über SAP Fiori, SAP GUI, SAP HANA Studio, RDP oder auf Betriebssystemebene erfolgen. Es kann aber auch eine Standardinstallation über *S/4HANA 2020 (FPS02)* gewählt werden, wo lediglich SAP Fiori im Client 100 aktiviert ist und Zugriff initial nur über RDP ermöglicht wird.

Um S/4HANA über die CAL bereitzustellen müssen Sie mehrere Schritte durchführen. Zuerst melden Sie sich bei SAP CAL über Ihren SAP Benutzer an. Als nächstes eröffnen Sie ein SAP CAL-Konto, welches für die Bereitstellung des S/4HANA-Systems genutzt wird. Dabei müssen Sie für die Bereitstellung in Microsoft Azure das Bereitstellungsmodell entsprechend auswählen. Hierbei können Sie den Resource Manager oder das klassische Bereitstellungsmodell definieren. Der Resource Manager ist hierbei empfehlenswert, da das klassische Bereitstellungsmodell auslaufen wird. Weiterhin geben Sie das Azure-Abonnement an, wobei zu berücksichtigen ist, dass das CAL-Konto nur einem Abonnement zugewiesen werden kann. Auch muss die Berechtigung zur Bereitstellung von Azure Ressourcen über die CAL für das Azure-Abonnement gegeben werden. Um die SAP CAL zu verwenden, müssen Sie ein Abonnement abschließen. Preislich liegt das Abonnement aktuell bei 750 € pro Monat bei einer Laufzeit von 12 Monaten. Nach Ablauf der Laufzeit wird das Abonnement automatisch um die ursprüngliche Laufzeit verlängert. Sie können das Abonnement kündigen, indem Sie das Abonnement mindestens 30 Tage vor Ablauf der Laufzeit zum Ende der Laufzeit hin beenden. Über eine kostenfreie Testversion können Sie erste Einblicke in die SAP CAL erhalten und sich damit vertraut machen.

- ▶ Bitte beachten Sie, dass in verschiedenen Systemen der SAP CAL virtuelle Instanzen verwendet werden, die gegebenenfalls in Ihren Azure-Kontingent nicht enthalten sind. Beispielsweise werden für SAP HANA Datenbanken teilweise VMs der M-Serie benötigt. Hierbei wird dann eine Erhöhung des Azure-Kontingents nötig. Wie eine Erhöhung des Azure-Kontingents durchgeführt wird, können Sie im Kap. 7 nachlesen.
- ▶ Bei der Bereitstellung von SAP-Systemen in Azure über die SAP CAL Testversion steht eine beschränkte Auswahl an Azure Regionen zur Verfügung. Möchten Sie das System in einer Azure-Region implementieren, welches nicht zur Auswahl steht, so müssen Sie bei SAP ein kostenpflichtiges CAL-Abonnement nutzen. Gegebenenfalls ist es auch nötig, eine Anfrage an SAP zu stellen, um die Bereitstellung des Systems in der speziellen Region freizuschalten.

### 6.3.3.2 Virtuelle Computer

Im Folgenden werden die Grundlagen der virtuellen Computing-Instanzen in Microsoft Azure vermittelt. Dabei wird der Fokus auf VM-Instanztypen, VM-Generationen, VM Zuständen gelegt. Weiterhin werden die Beziehungen zwischen den Ressourcen einer Instanz sowie die verschiedenen Nutzungsmodelle beschrieben. Auch werden Spot-Instanzen und vCPU Kontingente erläutert.

### 6.3.3.2.1 Beziehungen zwischen den Ressourcen

Bei der Erstellung einer VM ist es wichtig, dass Sie die folgenden Beziehungen zwischen Ressourcen beachten. Innerhalb einer Ressourcengruppe werden alle Ressourcen verfügbar gemacht. Der VM wird ein Speicherkonto zugewiesen, damit die Datenträger im Blobspeicher gesichert werden können. Die VM weist auf eine Netzwerkschnittstellenkarte und eine Verfügbarkeitsgruppe. Die Netzwerkschnittstellenkarte wiederum auf die IP-Adresse der VM sowie deren Subnetz und gegebenenfalls die zugehörige Netzwerksicherheitsgruppe, sofern vorhanden. Das Subnetz im virtuellen Netzwerk kann ebenfalls auf eine Netzwerksicherheitsgruppe hinweisen. Der Lastenausgleich leitet Anfragen von einer öffentlichen oder privaten IP-Adresse zu IP-Adressen die innerhalb einer Netzwerkschnittstellenkarte eines VMs definiert.

### 6.3.3.2.2 PAYG versus reservierte Instanzen

Microsoft bietet Kunden an, sich für einen bestimmten Zeitraum an eine Instanz zu binden, um Geld für den Cloud-Betrieb zu sparen. Abhängig von der Größe der Maschine und dem Zweck der Maschine kann es jedoch für einige Maschinen- und Systemtypen kosteneffizienter werden, Pay-As-You-Go-Optionen zu nutzen. Dies wird in diesem Kapitel erläutert.

Bei der Nutzung von Ressourcen auf Pay-as-you-go-Basis werden die Ressourcen in Rechnung gestellt, welche wirklich genutzt wurden. Bei Compute-Ressourcen geschieht das beispielsweise auf Basis der Laufzeit, in welcher die VM-Instanz im gestarteten Status war. Über den Peiskalkulator von Microsoft Azure können die Kosten kalkuliert werden, indem eine VM-Instanz ausgewählt wird, sowie die voraussichtliche Laufzeit. Zusätzlich ergeben sich unterschiedliche Preise je nachdem welche Region und welches Betriebssystem gewählt wird. Zusätzlich besteht die Möglichkeit Instanzen für einen bestimmten Zeitraum zu reservieren, das heißt diese Instanz wird für diesen Zeitraum im Voraus gekauft. Die Zeiträume können auf ein oder drei Jahre festgelegt werden und ermöglichen einen Rabatt von bis zu 72 % gegenüber der nutzungsbasierten Bezahlung für Azure-Dienste. Falls ein Workload mit langer Laufzeit und hohem Ressourcenbedarf in Azure abgedeckt werden soll, so macht es Sinn eine reservierte Instanz zu nutzen, um Kosten zu sparen. Stattdessen sollte bei saisonalen Workloads, die kurzfristig anfallen, das nutzungsbasierte Zahlmodell gewählt werden. Eine Kalkulation der Kosten beider Zahlmodelle ist für den jeweiligen Fall sinnvoll, um die kostengünstigere Alternative zu selektieren. Eine beispielhafte Rechnung wurde im Kap. 2 bereits durchgeführt. Grundsätzlich fallen für einige Azure-Dienste unabhängig von der nutzungsbasierten Abrechnung Fixkosten an. Als Beispiel kann der Service NetApp Files erst ab einer Größe von mindestens 4 TB genutzt werden und muss an die Nutzungslaufzeit der virtuellen Maschineninstanzen gekoppelt werden.

### 6.3.3.2.3 VM Instanz-Typen

Bei der Bereitstellung einer VM-basierten Lösungen ist die wichtigste Entscheidung zunächst den VM-Typ zu wählen. Azure bietet verschiedene VM-Serien an welche

über den Prefix der VM identifiziert wird (A, D, M, ...). Jede Serie bietet wiederum mehrere Größen an welche als VM-Typ bezeichnet werden (z. B. Standard\_D8\_v3). Jede VM-Serie bietet unterschiedliche Hardwareausstattung und richtet sich somit an einen anderen Einsatzzweck. So sind VMs der Serie M Arbeitsspeicheroptimiert und bieten daher ein hohes Verhältnis von Arbeitsspeicher zu CPU Kernen. Im Folgenden werden exemplarisch einige Serien erläutert, die für das SAP Hosting geeignet sind (Stand 2022):

#### **A-Serie:** Universelle VM-Größen

Diese Instanz-Typen bestehen aus einer ausbalancierten CPU- und Arbeitsspeichergröße und sind somit für Test- und Entwicklungsumgebungen geeignet, sowie Webserver mit höchstens mittlerer Auslastung. Für SAP HANA Datenbanken sind sie weniger geeignet, da sie lediglich kleine bis mittlere Datenbanken unterstützen.

#### **D-Serie:** Compute-optimiert

Compute-optimierte VM-Instanzen sind für Webservices, Netzwerkappliances und Batchvorgänge sowie Anwendungsserver mittlerer Auslastung geeignet, da sie über ein hohes Verhältnis zwischen CPU und Arbeitsspeicher verfügen.

#### **E-Serie:** Arbeitsspeicheroptimiert

Die E-Serie eignet sich sehr gut für relationale Datenbankserver und In-Memory-Analysen durch ein hohes Verhältnis von Speicher zu CPU.

#### **FX-Serie:** Compute-optimiert

Die FX-Serie eignet sich gut für Anwendungsserver mit mittlerer Auslastung durch viel CPU im Verhältnis zum Arbeitsspeicher.

#### **M-Serie:** Arbeitsspeicheroptimiert

Arbeitsspeicheroptimierte VM-Größen bestehen aus einem hohen Arbeitsspeicher-zu-CPU-Verhältnis. Deshalb eignen sie sich gut für SAP HANA Datenbanken, da sie mittelgroße bis große Caches und In-Memory-Analysen ermöglichen.

Weiterhin gibt es noch Instanz-Serien, welche Speicheroptimiert, GPU-optimiert und FPGA-optimiert sind sowie Serien für High Performance Computing:

#### **Speicheroptimiert**

Diese Instanz-Typen ermöglichen effiziente Datenträgerdurchsätze sowie E/A, weshalb sie sich gut für Datawarehousing und Big Data eignen.

#### **GPU – beschleunigte Compute-Prozesse**

Diese VM-Größen sind auf rechenintensive, grafikintensive und visualisierungsorientierte Prozesse spezialisiert.

### FPGA: beschleunigte Compute-Prozesse

Diese virtuellen Maschinen ermöglichen die effiziente Abarbeitung von rechenintensiven Prozessen.

### High Performance Computing

Diese virtuellen Instanzen sind speziell für High Performance Computing ausgelegt und ermöglichen hohe Skalierbarkeit, Rechenleistung und Kosteneffizienz.

Die meisten VM Serien werden für das **non-HANA** SAP Hosting unterstützt und können somit nach Preis-/Leistungskriterien frei gewählt werden. Zu den unterstützten VMs zählen Typen der Serien A, D, Das, Dds, DS, Eas, Es, GS, M, Mds. Voraussetzung ist jedoch, dass das SAP-System als 2-tier oder 3-tier Konfiguration implementiert wird und jedes SAP-RDBMS, welches nicht auf Basis der A- oder D-Serie bereitgestellt wurde Premium SSD Speicher (siehe Abschn. 7.3) zur Persistierung verwendet. Weitere Informationen finden sich in SAP Note 1928533 [5]. Für das SAP HANA Hosting ist die Auswahl an VMs wesentlich eingeschränkter, da nur ausgewählte VM-Typen eine Zertifizierung nach HANA-IAAS aufweisen. Zu den HANA zertifizierten Angeboten zählen Stand 2021 VM-Typen der Serie M (ab M32\*, 192 GB – 3892 GB) sowie einige Vertreter der Edv4 Serie (ab E20ds, 160 GB – 504 GB), sowie die Typen GS5 (448 GB RAM) und DS14v2 (112 GB RAM). Für eine aktuelle und genau Auflistung der HANA zertifizierten Azure Instanzen sollte das Hardware Verzeichnis [6] berücksichtigt werden. Insbesondere bei der Wahl der Edv4-Serie ist auf die genaue Bezeichnung zu achten, da alle Serien der E\* Familie als arbeitsspeicheroptimiert gelten, jedoch die Eav4 Serie auf AMD EPYC basiert und damit keinerlei SAP Unterstützung bietet.

#### 6.3.3.2.4 VM Generationen

Microsoft Azure bietet virtuelle Maschinen mit zwei unterschiedlichen Generationen an: Generation 1 und Generation 2. Virtuelle Instanzen der Generation 2 unterstützen zusätzlich zu den Funktionen der Generation 1 mehr Speicher, sowie virtualisierten persistenten Speicher, und Intel Software Guard Extensions (Intel SGX). Außerdem wird die Nutzung des UEFI-Boots mit VMs der zweiten Generation ermöglicht, bei der ersten Generation wird der BIOS-Boot genutzt. Dadurch weisen virtuelle Computer der zweiten Generation effizientere Start- und Installationsprozesse auf. VMs der ersten Generation werden von allen VM-Größen außer der Mv2-Serie unterstützt, wohingegen bei der zweiten Generation weitere Einschränkungen vorliegen: B-Serie, DCsv2-Serie, DSv2-Serie, Dsv3-Serie, Dsv4-Serie, Dasv4-Serie, Ddsv4-Serie, Esv3-Serie, Esv4-Serie, Easv4-Serie, Edsv4-Serie, Fsv2-Serie, GS-Serie, HB-Serie, HC-Serie, Ls-Serie, Lsv2-Serie, M-Serie, Mv2-Serie1, NCv2-Serie, NCv3-Serie, ND-Serie, NVv3-Serie, NVv4-Serie, NCasT4\_v3-Serie. Bitte beachten Sie, dass die Generation nach Erstellung des virtuellen Computers nicht mehr geändert werden kann. Aufgrund dessen ist eine vorausschauende Planung nötig.

### 6.3.3.2.5 VM Zustände

Die Kosten einer Azure VM werden stündlich abgerechnet, jedoch nur, wenn sich diese im Status „**Wird ausgeführt**“ befindet. Wenn eine VM nicht aktiv benötigt wird kann diese jederzeit heruntergefahren werden. Sollten Sie die Maschine manuell über das Betriebssystem herunterfahren, wechselt diese in den Zustand „**Beendet**“. Da eine beendete Maschine weiterhin die zugeordneten Computerressourcen reserviert, fallen auch weiterhin die Kosten der VM an, obwohl diese keine Rechenleistung mehr konsumiert. Unabhängig davon ob sich die VM im Status „Beendet“ oder „Wird ausgeführt“ befindet können Sie auch jederzeit über das Portal den Button „**Beenden**“ betätigen, um die VM in den Status „**Beendet (Zuordnung aufgehoben)**“ zu versetzen. Hierbei werden dann alle reservierten Ressourcen wie z. B. auch die Zugeordnete dynamische IP aufgehoben. VMs die sich in diesem Zustand befinden verursachen keine aktiven Kosten mehr, wobei berücksichtigt werden muss, dass weiterhin Kosten für reservierten Speicher, Sicherungen, etc. anfallen.

### 6.3.3.2.6 Spot-Instanzen

Azure Spot-Instanzen nutzen Ressourcen, die in Microsoft Azure aktuell ungenutzt sind. Dadurch kann Microsoft Azure sein Angebot effizienter gestalten, indem ungenutzte Ressourcen nutzbar gemacht werden. Diese Azure Spot-Instanzen sind preislich günstiger als anderweitige Instanzen. Allerdings sollten Sie sich bewusstmachen, dass die Azure-Spot Instanzen wieder entfernt werden, sobald die Ressourcen wieder anderweitig benötigt werden und keine Service Level Agreements für diese Instanzen existieren. Es besteht kein Anspruch auf die Nutzung der Azure-Spot-Instanzen, die verfügbaren Ressourcen können sich abhängig von verschiedenen Aspekten wie bspw. die Region und Tageszeit unterscheiden. Deshalb eignen sich Spot-VMs am besten für Prozesse bei welchen Unterbrechungen nicht kritisch sind, wie beispielsweise Development- und Test-Umgebungen. Da bei SAP S4/HANA Unterbrechungen kritisch sind, ist die Verwendung von Spot-Instanzen nicht zu empfehlen. Für Spot-Instanzen lassen sich auch Entfernungsrichtlinien definieren. Diese bestimmen, ab welchem Kapazitätswert oder Preis eine VM entfernt werden soll. Dabei kann die Entfernungsrichtlinie die Aufhebung der Zuordnung der Instanz in Gang setzen, oder die Löschung der virtuellen Maschine. Als Standardeinstellung ist die Aufhebung der Zuordnung der VM vorgesehen. Die Aufhebung der Zuordnung der virtuellen Maschine beendet zwar lediglich die Instanz und hebt die Zuordnung auf, allerdings wird keine Garantie auf eine erfolgreiche weitere Zuordnung bei erneuter Bereitstellung gegeben. Auch werden weiterhin Kosten für Speicherservices in Rechnung gestellt, die den virtuellen Maschinen zugeordnet sind. Beim Löschen der Instanzen werden auch die dazugehörigen Speicherservices gelöscht, weshalb dafür keine weiteren Kosten anfallen werden.

### 6.3.3.2.7 vCPU Kontingente

Microsoft Azure hat das Konzept der vCPU Kontingente eingeführt. Für virtuelle Instanzen mit nutzungsbasierter Abrechnung und reservierte VM-Instanzen werden

sogenannte vCPU Standardkontingente vergeben, Spot-VMs erhalten eine Spot-vCPU-Kontingent. Dabei kann das vCPU Kontingent in zwei Ebenen betrachtet werden: Regionale vCPUs und vCPUs der VM-Typfamilien. Zugeordnet werden diese Kontingente jedem Abonnement und jeder Region. Die vCPU Kontingente beschränken die VM-Bereitstellung insofern, dass die vCPUs der virtuellen Instanz das Kontingent für die VM-Typfamilie sowie das regionale vCPU-Kontingent nicht überschreiten. Die Anzahl der virtuellen Computer-Instanzen ist ebenfalls regional limitiert durch ein Kontingent. Die Kontingente können im Azure-Portal auf der Abonnement-Ebene unter Nutzung und Kontingente eingesehen werden. Wird eine höhere Anzahl an VMs oder vCPUs benötigt, so kann eine Erhöhung des Kontingents angefordert werden. Sollten Sie eine Erhöhung des vCPU-Kontingents für einen VM-Instanztyp anfordern, so erhöht sich automatisch auch das Kontingent der regionalen vCPUs. Eine Erhöhung des Kontingents kann auf zwei Arten angestoßen werden: Entweder über *Hilfe und Support* im Azure Portal, indem eine Support Anfrage gestellt wird oder über die Erhöhung der Kontingente über *Nutzung und Kontingente* im Bereich *Abonnements*.

### 6.3.3.3 Dedizierte Azure-Hosts

Microsoft Azure bietet zusätzlich die Möglichkeit einen dedizierten Host zu reservieren, welcher einem physischen Server zugeordnet ist. Darüber können Sie dann virtuelle Instanzen hosten, die flexibel an die jeweiligen Anforderungen angepasst werden können. Dedizierte Hosts können in einer Fehlerdomäne, Verfügbarkeitszone und einer Region bereitgestellt werden. Durch die Nutzung eines dedizierten Hosts können keine Hosts anderer Abonnements und Nutzer auf dem Host platziert werden. Somit können keine unvorhergesehnen Bottlenecks bei Lastspitzen mehrerer virtueller Instanzen entstehen. Außerdem besteht mehr Kontrolle bei Wartungsarbeiten, da bei der Buchung dedizierter Hosts ein spezifisches Wartungsfenster angemeldet werden kann.

Um einen dedizierten Host zu reservieren, erstellen Sie eine Hostgruppe in einer Verfügbarkeitsgruppe und einer Region. Nach Erstellung der Hostgruppe können Sie den dedizierten Host hinzufügen.

Host	Ein Host stellt eine Ressource dar, die einem physischen Server zugeordnet ist. Auf jedem Host können verschiedene virtuelle Maschinen gehostet werden, die derselben Größenserie angehören. Die unterstützte Größenserie ist anhand der SKU definiert.
SKU	Die SKU, sogenannte stock keeping unit, kann als Tarif beschrieben werden.
Hostgruppe	Eine Hostgruppe stellt eine Sammlung dedizierter Hosts dar.

Um Hochverfügbarkeit Ihres SAP-Systems sicherzustellen, sollten Sie mehrere virtuelle Maschinen nutzen, welche über mehrere Hosts verteilt sind. Anhand von dedizierten Hosts können Sie die Fehleranfälligkeit besser kontrollieren, indem Sie Bottlenecks bei Lastspitzen vermeiden und durch die Verteilung der dedizierten Hosts auf unterschiedliche Fehlerdomänen und Verfügbarkeitszonen Single Point of Failure vermeiden.

Da eine Hostgruppe in einer Verfügbarkeitszone erstellt wird, teilen sich alle virtuelle Maschinen dieser Hostgruppe die Rechenzentren dieser Zone, und somit auch dieselbe Stromversorgung, Kühlung und Netzwerke. Um eine bessere Verfügbarkeit zu erreichen, sollten Sie somit mehrere Hostgruppen über verschiedene Verfügbarkeitszonen hinweg erstellen und die Hosts somit über diese unterschiedlichen Verfügbarkeitszonen verteilen.

Durch die Verteilung von dedizierten Hosts in unterschiedliche Fehlerdomänen können die virtuellen Maschinen in verschiedene physische Racks innerhalb des Rechenzentrums platziert werden. Somit kann die Verfügbarkeit auch nach Ausfall eines physischen Racks im Rechenzentrum sichergestellt werden. Bei der Erstellung der Hostgruppe wird die Anzahl der Fehlerdomänen definiert, eine Bestimmung der Fehlerdomäne auf Ebene der virtuellen Maschine ist nicht nötig. Wichtig ist hierbei zu beachten, dass keine Antiaffinität der Fehlerdomänen zwischen zwei verschiedenen Hostgruppen angenommen werden kann. Dies ist nur garantiert, falls sich die beiden Hostgruppen in verschiedenen Verfügbarkeitszonen befinden. Zusätzlich kann nicht davon ausgegangen werden, dass Hosts mit derselben Fehlerdomäne im Rechenzentrum physisch nahe beieinanderliegen.

- ▶ Antiaffinität bedeutet in diesem Kontext, dass sich nicht zwei oder mehrere Hosts beider Hostgruppen in einer Fehlerdomäne befinden.

#### **6.3.3.4 Anforderungen für SAP und Microsoft Azure Support (VM)**

Um vollen Support von SAP und Microsoft für die Softwarearchitektur zu erhalten, sollten Sie die veröffentlichten Anforderungen einhalten. Dazu hat SAP eine Liste der zertifizierten Hardware für die Nutzung von SAP HANA veröffentlicht [6]. Hierbei können Sie durch die Filtersetzung eine Liste an VM-Instanzen generieren, die von SAP zur Installation von SAP HANA unterstützt werden.

Die SAP Note „2015553 – SAP on Microsoft Azure: Support prerequisites“ listet generelle Anforderungen für das Bereitstellen von SAP Anwendungen in Microsoft Azure auf, um vollen Support von SAP und Microsoft zu erhalten. Daraus geht hervor, dass nur VMs aus dem Standard-Tier unterstützt werden, da diese gewährleisten, dass keine Überbeanspruchung von Ressourcen, das heißt Ressourcenknappheit, möglich ist. Die SAP Note „1928533 – SAP Applications on Azure: Supported Products and Azure VM types“, listet diese unterstützten VM Typen auf. Hierbei solltest Sie auf die Aktualität der SAP Note achten, da die unterstützten VM Typen regelmäßig angepasst werden. Generell können Maschinen Instanzen der im folgenden beschriebenen Instanzklassen verwendet werden (Tab. 6.3).

Microsoft Azure bietet auch VM Instanzen mit limitierten vCPUs an. Dabei können die vCPUs auf die Hälfte oder ein Viertel der Original-Größe limitiert werden. Diese Instanzen können Sie ebenfalls für SAP Anwendungen nutzen. Hierbei sollten Sie allerdings eine Kalkulation der SAPS für das geplante System anhand der folgenden Formel anstellen:

**Tab. 6.3** Unterstützte Instanzklassen für SAP Anwendungen in Microsoft Azure

Instanzklasse	Größenspektrum	SAPS	Instanz-Typ
A-Serie	2–16 vCPU, 14–112 GB	1,500–22,000	Universelle VM-Größen
D-Serie	2–16 vCPU, 14–112 GB	2,325–18,600	Universelle VM-Größen
DS-Serie	2–16 vCPU, 14–112 GB	2,325–18,600	
DSv2-Serie	2–20 vCPU, 14–140 GB	3,530–30,430	Universelle VM-Größen (schneller als D-Serie)
DSv3-Serie	2–64 vCPU, 8–256 GB	2,178–69,680	Universelle VM-Größen (verfügen über Hyper-Threading-Technologie)
Easv4-Serie	2–96 vCPU, 16–672 GB	3,022–135,080	Arbeitsspeicheroptimiert
Dasv4-Serie	2–96 vCPU, 8–384 GB	3,022–135,080	Universelle VM-Größen (mehr CPUs und Arbeitsspeicher als DSv2)
Esv3-Serie	2–64 vCPU, 16–432 GB	2,178–70,050	Arbeitsspeicheroptimiert
Ddsv4-Serie	2–64 vCPU, 8–256 GB	3,142–100,550	Universelle VM-Größen (mehr lokaler Speicher & bessere IOPS auf lokalen Datenträgern für Lese- und Schreibvorgänge im Vergleich zu Dsv3)
Edsv4-Serie	2–64 vCPU, 16–504 GB	3,142–100,550	Arbeitsspeicheroptimiert (lokaler Hochgeschwindigkeitsspeicher, geringe Latenz)
GS-Serie	2–32 vCPU, 28–448 GB	3,580–41,670	Arbeitsspeicher- und massenspeicheroptimierte virtuelle Computer
M-Serie	8–128 vCPU, 219–3892 GB	8,616–134,630	Arbeitsspeicheroptimiert (bis zu 128 vCPU-Anzahl und bis zu 3,8 TB Arbeitsspeicher)
Mv2-Serie	208–416 vCPU, 2,85–11,4 TB	259,950–488,230	Arbeitsspeicheroptimiert (hohes Arbeitsspeicher/CPU-Verhältnis, Hyperthreading)
Msv2-Serie	32–192 vCPU, 875–4096 GB	42,711–256,750	Arbeitsspeicheroptimiert (nur als Generation 2 verfügbar)
Mdsv2-Serie	32–192 vCPU, 875–4096 GB	42,711–256,750	Arbeitsspeicheroptimiert (nur als Generation 2 verfügbar)

► SAPS der limitierten VM = (SAPS der unlimitierten VM \* Anzahl vCPUs der limitierten VM) / Anzahl vCPUs der unlimitierten VM.

Als Fazit, können Sie grundsätzlich alle von Microsoft Azure angebotenen Instanztypen für das Bereitstellen nutzen, sofern das Sizing der Ressourcen korrekt ist. Bei VMs die laut der entsprechenden SAP Note nicht unterstützt werden, wird allerdings kein Ressourcenscheduling garantiert. Das bedeutet, dass bei einer VM-übergreifenden hohen Nachfrage an Ressourcen, die maximale Ressourcennutzung der VM nicht gewährleistet werden kann (noisy neighbour/overcommitment).

Auf Betriebssystemebene gelten die folgenden Einschränkungen. Da die Nutzung der Azure Dienste auf x86-64- bzw. x64-Hardware beschränkt ist, muss das System auf einem der folgenden Betriebssysteme aufgebaut werden:

- Windows Server 64 Bit für die x86-64-Plattform
- SUSE Linux 64 Bit für die x86-64-Plattform
- Red Hat Linux 64 Bit für die x86-64-Plattform
- Oracle Linux 64 Bit für die x86-64-Plattform

Bei der Migration des Systems auf Microsoft Azure müssen Sie somit je nach bestehendem Betriebssystem einen Wechsel auf eines der oben genannten Betriebssysteme durchführen.

Weiterhin sind die benötigten NetWeaver/ABAP- oder Java-Stacks sowie Kernel-Mindestversionen abhängig von dem zu installierenden S/4HANA- und HANA-System. Das gilt es zu beachten, um den Support zu gewährleisten. In diesem Zusammenhang muss bei der Umstellung auf Azure geprüft werden, ob eine Aktualisierung der SAP-Kernel im Zuge der Migration notwendig wird. Entsprechende Hinweise sind in der SAP Note „1928533 – SAP Applications on Azure: Supported Products and Azure VM types“ enthalten.

### 6.3.4 Speicher

Microsoft bietet eine große Bandbreite an verschiedenen Speichertypen an. Diese unterscheiden sich nicht nur in ihrer Performance und deren Hostingkosten, sondern auch in ihrer Kompatibilität hinsichtlich der verschiedenen SAP workload Typen, sowie in den administrativen Verwaltungsmöglichkeiten. Dieses Kapitel soll einen Überblick über die verschiedenen Storageoptionen bieten und deren Vorteile und Einsatzmöglichkeiten erläutern.

#### 6.3.4.1 Leistungsmerkmale von Speichertechnologie

Bevor auf die einzelnen Speichertechnologien eingegangen und miteinander verglichen werden, wollen wir zunächst die wesentlichen Leistungskriterien erläutern. Hierfür

sind drei Metriken von besonderem Interesse: Disk-Größe (in GiB), Input/Output Operationen (IOPS) und Durchsatz (MB/sec).

**Disk Größe** Die meisten Disks müssen in fixen Größenstufen gebucht und gezahlt werden, auch wenn der tatsächlich genutzte Speicherbedarf geringer ist. Eine Ausnahme bilden hierbei nicht-verwaltete Datenträger (siehe nächste Sektion). Die zur Auswahl stehenden Größen unterscheiden sich häufig je Datenträgertyp. Während Premium Storage SSD Datenträger z. B. auch in als 32 GB Version gebucht werden kann, so beträgt die minimale Größe von Azure NetApp Files (ANF) 4 TB. Auch nachträgliches Resizing der initialen Größe wird nicht von allen Datenträger Typen unterstützt. Bei der Familie der verwalteten Datenträger (siehe Kap. 7), welche hauptsächlich im Kontext des SAP Hostings verwendet wird beträgt die maximale Größe eines Datenträgers in der Regel 32 TB (Ausnahme UltraDisk mit 64 TB). Insgesamt spielt die Disk Größe bei der Wahl des richtigen Speichertyps eine untergeordnete Rolle da je nach VM Typ bis zu 64 Datenträger angebunden werden können und somit die Flexibilität besteht auch Speicher im Petabyte Bereich bereitzustellen.

**IOPS** Die IOPS bestimmen die Anzahl der Lese- und Schreibzugriffe pro Sekunde die auf dem Datenträger durchgeführt werden. Alle Azure Datenträger Typen weisen vordefinierte IOPS Limitierungen auf. Sollte die Anwendung mehr IOPS benötigen als die zugrundeliegenden Speicher zur Verfügung stellt, so tritt automatisch ein Performance Verlust ein, da die IOPS Limitierung zum Flaschenhals wird. Die IOPS des Speichers werden von drei Variablen beeinflusst. Der erste Einflussfaktor ist der gewählte Speichertyp. Höherwertige Speicher wie z. B. „Premium SSD“ weisen höhere IOPS Limitierungen auf als niedrigerwertige Speicher wie etwa „Standard SSD“. Allerdings beeinflusst auch die Größe des Datenträgers die IOPS Anzahl. Hierbei wächst das IOPS Limit i. d. R. linear mit der Größe des Datenträgers. Außerdem kann das Limit auch erhöht werden indem mehrere Datenträger gleichzeitig verwendet werden. Hierbei fungiert das IOPS Limit kumulativ. D. h. 2× Disks mit jeweils 10.000 IOPS ergeben ein IOPS Limit von 20.000. Dies kommt natürlich nur dann zur Geltung, wenn die beiden Disks im Verbund arbeiten, also die Anwendung beide Datenträger gleichzeitig beschreibt. Bei den IOPS Limits kann außerdem zwischen „cached IOPS“ und „uncached IOPS“ unterschieden werden. Die „cached IOPS“ Grenzwerte sind höher als die „uncached IOPS“, werden aber nur dann berücksichtigt, wenn die Host Zwischenspeicherung für den Datenträger aktiviert wird. Wurde Host Zwischenspeicherung mit der Strategie „Nur Lesezugriffe“ aktiviert, so erhält der entsprechende Datenträger einen dedizierten Cache. Wann immer ein Lesezugriff vom Cache direkt beantwortet werden kann, so wird der I/O Zugriff nur für das „cached IOPS Limit“ angerechnet. Sollte sich das angeforderte Datum noch nicht im Cache besitzen oder falls ein Schreibzugriff durchgeführt wurde, wird der Zugriff für den cached und uncached IOPS Grenzwert angerechnet. Falls das IOPS Limit „Lese- und Schreibzugriffe“ aktiviert wurde, so werden Schreibzugriffe immer zunächst in den cache geschrieben und die Änderungen

werden asynchron und automatisch in einem Hintergrundjob auf Disk geschrieben. Für Lesezugriff ergibt sich dabei keine Änderung. Der Schreibzugriff selbst zählt dann lediglich für das cached IOPS Limit.

Wichtig ist außerdem, dass neben dem Speicher auch die VM selbst über eine IOPS Limitierung verfügt. Es wäre somit nicht lohnenswert, eine günstige VM, z. B. Standard\_A2\_v2 (2000 IOPS) mit einem hochpreisigen Speicher, z. B. Ultra Disk 64 GB (19.200 IOPS) zu kombinieren da in diesem Fall die VM selbst zum Flaschenhals wird. Eine genaue Planung der IOPS Kapazitäten ist somit unerlässlich, um das beste Preis-Leistungsverhältnis zu erzielen.

**Durchsatz** Der Durchsatz ist meist in Megabyte pro Sekunde (MB/s oder Mbps) angegeben und bestimmt die maximale Bandbreite beim Lesen oder Schreiben von Daten. Insbesondere bei sequentiellen Lese- oder Schreibvorgänge wie z. B. beim Kopieren großer Dateien oder Fortschreiben des Datenbanklogs ist dieser Wert von besonderer Bedeutung. Ähnlich wie bei den IOPS ist der Durchsatz abhängig vom Datenträger Typ, der Größe des Datenträgers und der Anzahl. Somit bedingen hohe IOPS automatisch auch einen hohen Durchsatz. Zu beachten ist jedoch, dass nicht alle Speicher Typen ihren theoretischen Durchsatz auch praktisch erbringen. Insbesondere bei günstigen Standard HDD Speicher kann die tatsächliche Performance erheblich von der theoretisch möglichen abweichen. Lediglich teurere Datenträger Typen wie z. B. „Ultra Disk“ bieten auch Performance Garantien.

#### 6.3.4.2 Verwaltete vs. nicht-verwaltete Datenträger

Grundsätzlich unterscheidet MS Azure bei der Wahl des Speichers zwischen verwalteten Datenträgern (engl. „managed“) und nicht verwalteten Datenträgern (engl. „unmanaged“). Verwaltete Datenträger werden vollständig von Microsoft gesteuert und vereinfachen die Administration erheblich. Auch der Funktionsumfang insbesondere hinsichtlich Skalierbarkeit, Zuverlässigkeit und Sicherheit ist größer als bei nicht verwalteter Disk. So beinhalten verwaltete Disks standardmäßig ein rollenbasiertes Zugriffskonzept mit Verschlüsselung, sowie höhere Verfügbarkeit und Fehlertoleranz durch Replizierung der Daten auf zwei weitere Disks (RAID Spiegelung). Nicht verwaltete Datenträger werden hingegen vollständig vom Kunden gesteuert. Hierzu muss zunächst ein sogenanntes Speicherkonto (Storage Account) angelegt werden. Dieses stellt im Grunde einen Namespace bereit über welchen alle Speicher-Objekte die mit dem Konto assoziiert sind identifiziert werden können. Datenträger können anschließend als VHD Dateien im Konto angelegt werden. Hierbei handelt es sich dann um nicht-verwalteten Speicher, da der Nutzer nun selbstständig Konzepte implementieren muss um Verschlüsselung oder Datenwiederherstellung sicherzustellen. Aus technischer Sicht basieren VHD Dateien aufseiten Blobs. Neben diesen können auch andere Speicherobjekte wie block Blobs/append blobs (für unstrukturierte Daten), Tabellen (NoSQL strukturierte Daten), Queues (Nachrichtenaustausch zwischen Anwendungs-komponenten) und Dateien (z. B. Netzwerkspeicher basierend auf NFS) provisioniert

werden. Während nicht-verwaltete Datenträger inzwischen als veraltet gelten und anstelle verwalteter Datenträger verwendet werden sollten, kommt der der Nutzung der anderen Speicherobjekte wie Blobs oder Dateien noch immer eine hohe Bedeutung zu. Diese spielen jedoch im Kontext des SAP Hostings eine untergeordnete Rolle. Sowohl SAP als auch Microsoft empfehlen für den SAP-Betrieb die Nutzung von verwalteten Datenträgern gegenüber nicht-verwalteten. Hinsichtlich der Preisstrategie gibt es laut Microsoft keine Unterschiede zwischen verwalteten und nicht-verwalteten Datenträgern. Es sollte jedoch darauf hingewiesen werden, dass verwaltete Datenträger grundsätzlich in festen Größen provisioniert werden und immer mit ihrer vollen Kapazität (in GiB) bepreist werden. Die Größe von nicht-verwaltetem Standardspeicher kann hingegen flexibel gewählt werden, wodurch sich ein geringes Einsparungspotential ergeben kann, da nur der tatsächlich genutzte Speicher bepreist wird. Dies ist jedoch nur für Standardspeicher der Fall. Bei Premiumspeicher ergeben sich bei der Bepreisung zwischen verwalteten und nicht-verwalteten Speicher keine Unterschiede. Die folgende Tabelle gibt einen kurzen Überblick über die beiden Datenträger Varianten (Tab. 6.4).

**Tab. 6.4** Vergleich verwaltete und nicht verwaltete Disks

Verwaltung und Steuerung	Zugehöriges Speicher Konto muss vom Kunden zuvor angelegt werden. Disk kann anschließend manuell als VHD Datei angelegt. Volle Kontrolle durch den Kunden, aber sämtlich zahlreiche Einschränkungen welche vom Nutzer berücksichtigt werden müssen	Zugehöriges Speicher Konto wird vom Azure Ressourcen Manager automatisch angelegt und verwaltet. Dadurch einfacheres Design und Nutzung für den Kunden
Größe	Benutzerspezifische Größe kann gewählt werden insofern „Standard“ als Performanceklasse gewählt wurde	Feste Größen abhängig von der Speicherklasse. Nachträgliche Anpassung der Größe ist aber möglich
Skalierbarkeit	Lediglich 250 Storage Accounts pro Abonnement/Region möglich. Bis zu 40 VMs pro Speicherkonto möglich. Erfordert ggf. die Verteilung auf mehrere Speicherkonten	Bis zu 50.000 Datenträger pro Abonnement/Region
Sicherheit	Mehrere VHD Files befinden sich in einem Storage Account. Keine standardmäßige Verschlüsselung oder rollenbasiertes Zugriffskonzept	Disks sind voneinander isoliert. Standardmäßige Verschlüsselung und rollen-basiertes Zugriffskonzept
Performance		
Verfügbarkeit		Disk Spiegelung auf zwei weitere physikalisch Datenträger standardmäßig vorhanden. Umfangreiche Unterstützung von Verfügbarkeitszonen und Verfügbarkeitssets

Im Folgenden werden wir uns auf verwaltete Disks fokussieren, da diese im Rahmen des SAP-Hostings hauptsächlich verwendet werden. Außerdem werden wir Azure NetApp Files (ANF) betrachten. Hierbei handelt es sich um einen hochperformanten Fileserver Dienst (NFS/SMB), welchem insbesondere im Kontext Hochverfügbarkeits-szenarien eine besondere Bedeutung zukommt.

#### 6.3.4.3 Datenträgerrollen

Abhängig vom Einsatzzweck des Datenträgers kann zwischen den drei Rollen Betriebssystem-Datenträger, Temporärer Datenträger und regulärer Datenträger (Datenträger für Daten) unterschieden werden. Diese Trennung ist wichtig da nicht alle **Datenträgertypen** (siehe nächste Sektion) für alle Rollen verwendet werden können.

**Betriebssystem Datenträger** Auf diesem Datenträger wird das Betriebssystem im Rahmen der VM-Erstellung provisioniert. Die Größe des Datenträgers ist auf maximal 4 TB begrenzt. Auch die Wahl der Disk Typen ist eingeschränkt. So kann etwa die schnellst verfügbare Storagelösung (Ultra disk) nicht als Betriebssystem-Datenträger verwendet werden.

**Temporärer Datenträger** Viele VMs beinhalten standardmäßig neben dem Betriebssystem-Datenträger einen zusätzlichen temporären Datenträger. Dieser dient vor allem als kurzfristiger Speicher z. B. als Auslagerungsspeicher für swapping oder paging. Anders als der Name es vermuten lässt gehen diese Daten jedoch nicht im Rahmen eines Neustarts verloren. Allerdings gibt es auch keine Garantieansprüche. Vor allem bei Wartungsmaßnahmen, Resizing oder einer Neuallokierung der VM können die Daten verloren gehen. Aus diesem Grund sollte der Temporäre Datenträger nicht als persistenter Speicher und insbesondere nicht für SAP Anwendungen verwendet werden.

**(Regulärer) Datenträger** Hierbei handelt es sich um klassischen Laufwerke welche zur langfristigen Persistierung von Anwendungsdaten verwendet werden. Im Rahmen des SAP Hostings werden i. d. R. eigenständige Datenträger für die verschiedenen SAP Verzeichnisse wie z. B. /hana/data oder /hana/log angelegt und eingehängt. Diese offenbaren sich dem Betriebssystem als SCSI Laufwerke und können maximal eine Kapazität von 32 TB (pro Disk) aufweisen. Die Anzahl an Datenträgern die eingehängt werden können ist abhängig von der gewählten VM Größe.

**Snapshots und Images** Bei Snapshots und Images handelt es sich nicht um eigenständige Datenträger im klassischen Sinne. Ein Snapshot ist lediglich eine Kopie/Momentaufnahme eines bestehenden Datenträgers und wird in erster Linie für Sicherungszwecke verwendet. Der Snapshot existiert unabhängig vom Quelldatenträger. D. h. der ursprüngliche Datenträger kann jederzeit geändert oder sogar gelöscht werden ohne dass dies Auswirkungen auf den Snapshot hat. Im Gegensatz zu regulären Datenträgern werden Snapshots nur mit ihrer tatsächlichen Größe berechnet d. h. ein

Snapshot einer 100 GB Disk auf welcher nur 20 GB in Verwendung sind, wird nur zu 20 GB bepreist. Eine erweiterte Form des Snapshots ist das sogenannte Image. Während der Snapshot lediglich eine Momentaufnahme eines Datenträgers darstellt, umfasst ein Image die vollständige VM inklusive aller zugeordneten Datenträger. Auf diese Weise kann auch das Zusammenspiel mehrerer Datenträger, beispielsweise im Rahmen von RAID Verbünden, mitgesichert werden.

#### **6.3.4.4 Redundanzoptionen**

Damit gespeicherte Daten im Falle eines physischen Datenträgerverlustes nicht verloren gehen bietet Microsoft vier verschiedene Redundanzoptionen an:

Lokal Redundanter Speicher (LRS): Es werden drei Kopien innerhalb eines Azure Rechenzentrums vorgehalten. Der Schreibzugriff erfolgt immer synchron, d. h. alle drei Kopien werden zeitlich beschrieben. Die Daten sind somit vor Ausfall einzelner Laufwerke sogar eines ganzen Serverracks geschützt. Ein Katastrophenfall der das ganze Rechenzentrum betrifft (Feuer/Wasser) könnte jedoch zum Datenverlust führen. LRS wird von allen Azure Storage-Diensten unterstützt und ist gleichzeitig die einzige Redundanzoption welche von verwalteten-Datenträgern unterstützt wird.

Zonenredundanter Speicher (ZRS): Bei ZRS werden die drei synchronen Kopien auf drei Verfügbarkeitszonen innerhalb der Region verteilt, d. h. die Daten bleiben in der primären Region z. B. „US-West“ befinden sich jedoch in verschiedenen Rechenzentren welche über unabhängige Stromversorgung, Kühlung, etc. verfügen. Ein Ausfall müsste somit dauerhaft alle drei Zonen betreffen damit ein potentieller Datenverlust auftritt. ZRS sollte daher für hochverfügbare Lösungen verwendet werden.

Georedundanter Speicher (GRS): Bei GRS werden Daten über zwei Regionen z. B. „Deutschland-West“ und „US-Ost“ verteilt. Innerhalb einer Region werden die Daten als LRS vorgehalten. Schreibvorgänge erfolgen in der primären Region zuerst und werden dann an die sekundäre Region asynchron repliziert. Sollte die primäre Region im Katastrophenfall nicht mehr erreichbar sein muss ein Failover durchgeführt werden in welcher die sekundäre Region zur primären Region wird, ansonsten sind die Daten nicht verfügbar.

Geozonenredundanter Speicher (GZRS): GZRS ist eine Sonderform des GRS in welcher zusätzlich zur asynchronen Replikation der Daten über zwei Zonen hinweg, auch innerhalb einer Zone ZRS verwendet wird. D. h. die Daten werden sowohl innerhalb der primären als auch der sekundären Region über drei Verfügbarkeitszonen hinweg gespeichert (insgesamt sechs). GZRS ist die höchste Redundanzoption, jedoch wird diese Option nicht für alle Regionen angeboten.

### 6.3.4.5 Datenträgertypen

Azure unterscheidet zwischen den vier Datenträgertypen „HDD Standard“, „SSD Standard“, „SSD Premium“ und „Ultra Disk“. Diese unterscheiden sich sowohl hinsichtlich ihrer Leistungsmerkmale (Kap. 7), ihren möglichen Datenträgerrollen (Kap. 7), sowie natürlich preislich.

**Ultra Disk** Ultra Disks bieten die beste Performance mit Übertragungsraten von bis zu 2000 MB/s und maximalen IOPS von 160.000 (was jedoch nur von wenigen VM Größen wie z. B. M128s unterstützt wird). Eine besondere Eigenschaft der Ultra Disk ist außerdem die Möglichkeit die Datenträgerleistung dynamisch (ohne Neustart der VM) an den Workload anpassen zu können.

Ultra Disk kommen jedoch auch mit einige Einschränkungen. Beispielsweise ist nur die Rolle als regulärer Datenträger unterstützt, nicht jedoch z. B. das Betriebssystem Datenträger. Die Nutzung von Snapshots ist ebenfalls nicht möglich. Als Hochverfügbarkeitsangebot werden lediglich Verfügbarkeitszonen (engl. „availability zones“) angeboten. Dies ist jedoch zusätzlich auf einige Regionen und VM Familien eingeschränkt z. B. unterstützen in der Region „Deutschland, Westen-Mitte“ nur einzelne VMs „Ultra Disk“ überhaupt. Dafür werden ausnahmslos alle Anwendungsszenarien im Rahmen des SAP Hostings unterstützt, darunter auch latenzkritische Pfade wie z. B. /hana/log.

**SSD Premium** SSD Premium bietet ebenfalls hohe Performance kommt jedoch nicht an die Leistung der Ultra Disk heran (maximal 900 MB/s Durchsatz und 20.000 IOPS pro Datenträger). Anders als die Ultra Disk unterstützt SSD Premium jedoch auch Betriebssystem Datenträger. Voraussetzung zur Nutzung für SSD Premium ist jedoch, dass die gewählte VM „Storage Premium“ unterstützt. Dies ist z. B. für alle M-Instanzen (Memory Optimiert) der Fall, nicht jedoch für günstiger VMs wie z. B. die A- oder D-Serie. SSD Premium bietet darüber hinaus garantierte Übertragungsraten. Für kleinere Datenträgergrößen bis 512 GB (P20) wird außerdem das „Bursting“ Feature angeboten. Dabei erhält jeder Datenträger voll automatisch ein Kontingent welches ermöglicht Durchsatz und IOPS über die garantierten Grenzen hinaus zu steigern. Solange der Workload geringer ist als die garantierten Übertragungsraten füllt sich das Guthaben auf. Bei Lastspitzen werden die Übertragungsraten erhöht, womit jedoch das Guthaben aufgebraucht wird. Vor allem bei kleineren Größen kann dadurch ein erheblicher Performancevorteil erzielt werden. Beispielsweise hat ein P4 Datenträger eigentlich nur einen Durchsatz von 25 MB/s. Mit Bursting kann dieser jedoch auf bis zu 170 MB/s angehoben werden. Bursting ist standardmäßig aktiviert und muss nicht manuell aktiviert werden. SSD Premium sollte idealerweise als Betriebssystemdatenträger zum Einsatz kommen, da schnellere Varianten wie Ultra disk oder ANF nicht verwendet werden können. Außerdem wird die Nutzung von SSD-Premium für nahezu alle SAP Anwendungsszenarien empfohlen. Lediglich als HANA Log Volume (/hana/log) ist die Nutzbarkeit eingeschränkt da hierbei nicht die für HANA erforderlichen I/O Latenzen erzielt werden können. Aus diesem Grund sind lediglich die VM Familien M und Mv2

unterstützt unter der Prämisse, dass die Azure Schreibbeschleunigung aktiviert ist. Hierbei handelt es sich um ein exklusives Feature der M-Serie welche in Kombination mit SSD Premium verwendet werden kann. Die Schreibbeschleunigung ist optimiert für das Schreiben von Log-Dateien. Es sollte daher ausschließlich in diesem Kontext aktiviert werden, nicht aber für klassische Datenvolumes z. B. /hana/data. Die Schreibbeschleunigung ist außerdem mit einigen Restriktionen verbunden. Mit Aktivierung der Beschleunigung können fortan keine Snapshots vom Datenträger gemacht werden. Außerdem muss die die Datenträgerzwischenspeicherung ausgeschaltet oder im Modus „Schreibgeschützt“ sein.

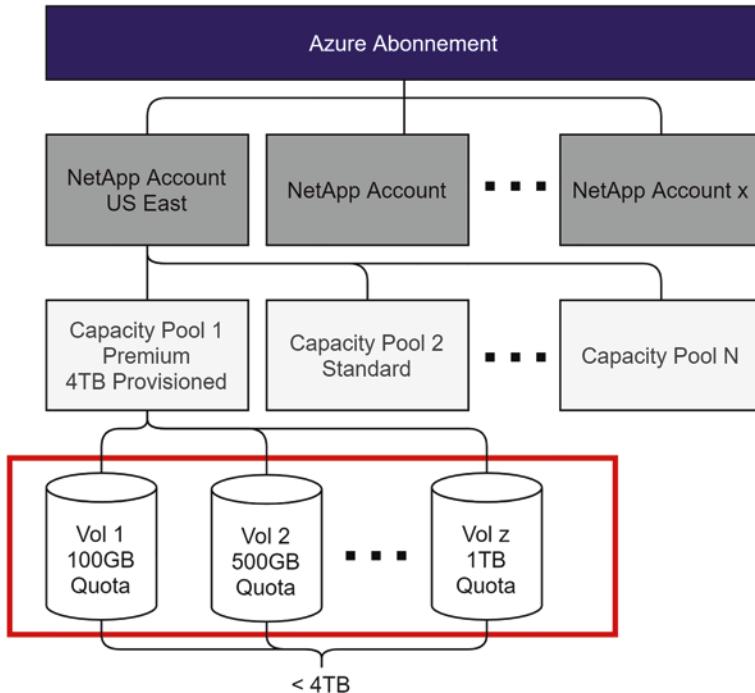
**SSD Standard** SSD Standard ist die günstigere Variante zu SSD Premium. Die theoretische Performance ist ähnlich zu Premium allerdings kann diese größeren Schwankungen ausgesetzt sein. Auch SSD-Standard unterstützt Bursting, allerdings ist hier der Burst-Durchsatz mit 150 MB/s i. d. R. 20 MB/s geringer als bei SSD Premium. SSD-Standard wird zudem durchgehend von allen VM Instanzen unterstützt. Für das SAP Hosting eignet sich SSD Standard nur eingeschränkt. Für nicht produktive Instanzen sollte SSD Standard lediglich für das SAP-System-Mount-Verzeichnis/sapmnt, sowie als Betriebssystemdatenträger verwendet werden. Andere Anwendungszwecke wie z. B. als globales Transportverzeichnis werden nicht länger unterstützt.

**HDD Standard** HDD Standard ist der kostengünstigste aber zugleich auch langsamste Datenträgertyp. Der theoretische Durchsatz bis 4 TB (S4 bis S50) ist mit 60 MB/s nicht einmal halb so groß wie die von Standard SSD. Auch wird kein Bursting angeboten. HDD eignet sich somit nur als Massespeicher, wenn Performance eine untergeordnete Rolle spielt oder das Netzwerk den Flaschenhals darstellt (z. B. Archiv oder Fileserver). In anderen Fällen lohnt es sich fast immer den Aufpreis für SSD zu investieren. Für das SAP Hosting kann HDD Standard lediglich als Betriebssystemdatenträger sowie für das SAP-System-Mount-Verzeichnis oder Backupablage verwendet werden, wobei selbst diese Anwendungsszenarien nicht empfohlen sind. Abb. 6.4 zeigt die hierarchische Zuordnung von Speicher zu Abonnements, Accounts und Pools.

### Azure NetApp Files (ANF)

Bei ANF handelt es sich um einen Azure NFS Dienst welcher hochperformanten Netzwerkspeicher bereitstellen kann. Wie in Abb. 6.4 gezeigt, muss dazu zunächst ein NetApp Account erstellt werden. Dieser ist regionsabhängig und kann somit nur in der Region verwendet werden in welcher dieser erstellt wurde.

Auf Basis dieses Accounts können dann Kapazitätspools erstellt werden welche wiederum die einzelnen Datenträger/Kontingente bereitstellen. Diese werden als Netzwerkspeicher als NFS oder SMB Datenträger eingebunden. Stand 2021 ist ANF das einzige geteilte Filesystem welches SAP HANA zertifiziert ist. ANF eignet sich



**Abb. 6.4** Speicherressourcen-Zuordnung in Microsoft Azure

somit insbesondere, wenn große Speicherkapazitäten in Kombination mit hoher Performance erforderlich sind. Die Mindestgröße bei Erstellung eines ANF Kapazitätspools beträgt 4 TB und kann in 1 TB Schritten auf bis zu 400 TB vergrößert werden. ANF bietet damit unter allen Speicherlösungen die höchste Skalierbarkeit. Pro Datenträger muss ein minimales Kontingent von 100 GiB bis maximal 100 TB zugewiesen werden. KOSTEN!!! Bei Verwendung von SAP HANA wird NFS lediglich in der Protokollversion 4.1 unterstützt. ANF bietet drei Dienstebenen: Standard (16 MB/s), Premium (64 MB/s) und Ultra (128 MB/s). Diese Durchsätze gelten kumulativ pro TiB Kontingent. D. h. je größer das Kontingent desto höher der Durchsatz z. B. bietet ein 4 TiB Premium Speicher  $4 \times 64 \text{ MB/s} = 256 \text{ MB/s}$  und wäre damit performanter als 1 TiB Ultra Speicher. Zu berücksichtigen ist hierbei, dass ANF Durchsatz immer auf die Netzwerkanbandbreite angerechnet wird und nicht auf die Speicheranbandbreite. ANF eignet sich für alle Arten von SAP Volumes die im Kontext von S/4HANA auftreten. Es gilt lediglich die Einschränkung, dass sowohl /hana/data als auch /hana/log auf ANF bereitgestellt wird. ANF eignet sich besonders im Kontext von HA-Szenarien als Lösung um hochverfügbarer geteilten NFS Speicher z. B. für /sapmnt oder das Transportverzeichnis /usr/sap/trans bereitzustellen.

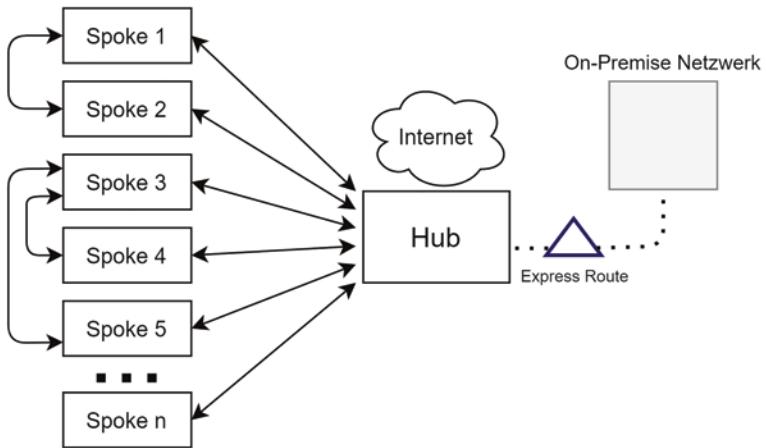
## 6.3.5 Netzwerk und Services

In diesem Kapitel wird das Konzept der Hub-Spoke-Topologie beschrieben sowie weitere relevante Services und Tools für die Netzwerkkonfiguration erklärt.

### 6.3.5.1 Hub- Spoke Topologie und Netzwerk Peering

**Das Nabe-/Speiche Netzwerk (Hub-Spoke) Topologie** ist ein Standardkonzept welches in fast allen Azure Netzwerkarchitekturen Anwendung findet. Die **Nabe (Hub)** dient dabei als zentrale Anlaufstelle aus dem on-premise Netzwerk und umfasst mehrere Subnetze. Wie in Grafik 6.5 dargestellt, bildet das Gateway Subnetz das Tor zum Unternehmensnetz. Das Gateway Subnetz ist ein dedizierter Netzwerktyp welches nicht für normale VM Bereitstellungen genutzt werden kann, sondern ausschließlich für die Bereitstellung einer Gateway Komponente genutzt werden kann. Die Verbindung zum Unternehmensnetz wird über eine von zwei Methoden hergestellt. Ein Site-to-site (S2S) VPN ermöglicht es eine VPN Verbindung zu etablieren indem ein dediziertes VPN Gerät aus dem on-premise Netzwerk mit einem Azure VPN Gateway gekoppelt wird. Dem gegenüber steht die Express Route Variante. Hierbei kann in Kooperation mit dem Internet Service Provider eine private Verbindung mit vereinbarten Bandbreiten zwischen dem Unternehmensnetz und Azure eingerichtet werden. Das Gateway Netz mündet häufig in einem DMZ-Netz, welche zentrale Dienste, sowie ggf. eine Firewall hosted. Das Management Subnetz stellt schließlich eine oder mehrere Jumpboxes auch Bastion-Hosts genannt bereit. Hierbei handelt es sich um Remotesysteme die der Administration der SAP Landschaft dienen und ein gehärtetes Betriebssystem aufweisen. Außerdem werden entsprechende Clienttools bereitgestellt und vorkonfiguriert wie etwa SAP Logon, HANA Studio oder HANA Cockpit. Zusätzlich ist das Nabe Netzwerk so eingerichtet, dass für administrative Zugriffe wie SSH, nur von dort aus eine Weiterverbindung in die angrenzenden Speichernetzwerke ermöglicht wird. Dies wird durch entsprechende Einstellungen in den Netzwerksicherheitsgruppen (NSG) ermöglicht. Man verbindet sich somit nicht direkt auf die Ziel-Systeme, sondern baut eine Verbindung zum Jumpserver auf, von welchem man sich dann auf die jeweilige Zielmaschine weiterverbinden kann. Grundsätzlich kann jede VM als Jumpserver konfiguriert werden. Azure bietet alternativ den hauseigenen „Azure Bastion“ PAS Dienst an, welcher unter anderem bereits mit RDP und SSH vorkonfiguriert ist und außerdem direkt über das Azure Portal aufgerufen werden kann (Abb. 6.5).

Die Speiche (spoke network) hostet schließlich den eigentlichen SAP Workload. Hierbei wird typischerweise eine Aufteilung unterschiedlicher Workloadtypen (z. B. Produktiv vs. Entwicklung) auf unterschiedliche Speiche-Netzwerke vorgenommen. Der Benutzerzugriff wird ebenfalls entsprechend eingerichtet, sodass Entwickler und Endbenutzer für unterschiedliche Speiche Netzwerke freigeschaltet sind, wobei ein umfangreicher Managementzugriff nur aus dem Managementsubnetz, der Nabe ermöglicht wird.



**Abb. 6.5** Hub- und Spoke-Topologie

**VNet peering** Um das Konzept des Nabe- Speiche Netzwerks umzusetzen bedarf es dem Konzept des VNet Peerings. Jedes virtuelle Netzwerk in Azure definiert standardmäßig einen privaten Adressraum. Damit Instanzen eines VNets auf die Instanzen eines anderen VNets zuzugreifen ohne dass hierfür über eine öffentliche IP kommuniziert werden muss, bedarf es dem VNet peering. Gepeerte Netzwerke erscheinen als ob all VMs Teil des gleichen Netzes wären. Voraussetzung für die Umsetzung ist, dass beide VNets über nicht-überlappende IP-ranges verfügen. Auch VNets aus unterschiedlichen Regionen oder Abonnements können miteinander gepeert werden. Aus einer Performance Sicht erhalten gepeerte Netzwerke den gleichen Durchsatz als ob sie sich in einem gemeinsamen Netzwerk befinden würden, jedoch nur solange diese sich in der gleichen Region befinden. Auch in diesem Fall erfolgt der Datentransfer Microsoft intern und wird nicht über das öffentliche Netz geroutet, jedoch fallen bei gepeerten Netzen Transferkosten an welche insbesondere im SAP Hosting hohe Datenvolumen und somit auch Kosten verursachen können. Aus diesem Grund empfiehlt es sich meistens Anwendungsserver und Datenbank einer SAP-Installation nicht in VNets zu isolieren, sondern Subnetze zu verwenden.

**SAP Supporteinschränkungen** Die Aufteilung des SAP Workloads auf die jeweiligen virtuellen Netze sollte sorgfältig geplant und unter Berücksichtigung der SAP Empfehlungen aus SNOTE 2015553 [7] vorgenommen werden. Grundsätzlich empfiehlt SAP Anwendungslayer und Datenbanklayer nicht mittels unterschiedlicher VNETs zu trennen, sondern stattdessen eigene Subnetze zu definieren. Eine Trennung auf virtueller Netzebene ist jedoch supported solange diese gepeert werden. Ein hybride-Cloud Schichtentrennung, bei der beispielsweise nur der Anwendungsserver- oder Datenbanklayer in der Cloud gehostet wird und die jeweilige andere Schicht on-premise gehostet wird ist explizit nicht supported. In allen Fällen sollte auch die Latenz zwischen der Azure Region und dem Unternehmensnetz berücksichtigt werden. Insbesondere für

produktiven Workload sollte eine sorgfältige Evaluierung der ExpressRoute oder Site-to-Site VPN durchgeführt werden.

### 6.3.5.2 ExpressRoute vs. S2S VPN

Die Implementierung eines Hybridnetzwerks welches ein bestehendes Unternehmensnetz mit dem virtuellen Azure Netzwerk koppelt ist im Kontext des SAP Hostings der Regelfall. Wie eingehend in Abschn. 6.3.4.1 erläutert stehen hierzu zwei Möglichkeiten zur Verfügung welche folgend näher erläutert werden.

**S2S VPN** Eine Site-to-Site VPN Verbindung lässt sich schnell implementieren und ist einfach zu konfigurieren. Hierbei wird der Datenverkehr zwischen dem Azure VPN Gateway und dem on-premise Gateway Device mittels IPsec verschlüsselt und regulär über das öffentliche Netz geroutet. Hierbei kann jedoch keine Aussage zur Latenz und zur maximalen Bandbreite getroffen werden. Aufgrund der Kritikalität von SAP Workload, lautet daher die offizielle Empfehlung von Microsoft eine Expressroute zu verwenden um bessere Performance (Durchsatz und Latenz), sowie Sicherheit und Zuverlässigkeit zu gewährleisten. In der Praxis ist eine S2S VPN Verbindung jedoch für viele Anwendungsfälle (einschließlich SAP) ausreichend. SAP selbst supported beide Verbindungsszenarien (SNOTE 2015553 [7])

1. Zur Implementierung muss zunächst im Azure Gateway Subnetz ein VPN Gateway erstellt werden. Diese Komponente kommt grundsätzlich in sechs verschiedenen Preisklassen (Basic+VpnGw1 – VpnGw5), sowie einigen AZ Subvarianten welche zusätzliche Zonenredundanz bieten. Die Bandbreite reicht je nach Preisklasse von 100 Mbit/s bis 10 Gbit/s. Hierbei werden die Basiskosten pro Stunde berechnet (z. B. 3,65 USD/std). In diesem Preis sind bis zu 10 Tunnel inkludiert. Weitere Tunnel werden separat berechnet (ebenfalls pro Stunde und für die Dauer der Nutzung) wobei ein Gateway maximal 30 Tunnel insgesamt unterstützt. Zu beachten ist, dass ein einzelner Tunnel maximal 1 Gbit/s Bandbreite bietet. Die Preisklasse des Gateways kann i. d. R. dynamisch geändert werden, insofern nicht die Basic Preisklasse gewählt wurde, außerdem ist ein Wechsel eines Gen1 Gateways zu Gen2 ist nicht möglich.
2. Anschließend muss das on-premise VPN Gerät eingerichtet werden. Neben der Konfiguration eines geteilten Schlüssels und der öffentlichen IP des Azure VPN Gateways sollten auch die empfohlenen Konfigurationseinstellungen des Geräteherstellers bzw. von Microsoft berücksichtigt werden. Microsoft unterstützt nicht alle VPN Geräte, jedoch eine große Anzahl welche unter folgender Seite zu finden sind (Zugriff am 20.12.2021): <https://docs.microsoft.com/azure/vpn-gateway/vpn-gateway-about-vpn-devices>. Für Geräte von Cisco, Juniper und Ubiquiti stehen darüber hinaus Konfigurationsskripte zur Verfügung welche die Einrichtung vereinfachen und im Azure Portal heruntergeladen werden können [8]. Sollte kein Hardware VPN zur Verfügung stehen können auch software basierte VPNs genutzt werden wie beispielsweise OpenSWAN (Linux) oder MS Routing and Remote Access Service (Windows).

**Express Route** Bei der Express Route handelt es sich um eine private Verbindung zwischen dem Unternehmensnetz und Azure welche von einem externen Netzwerk-anbieter bereitgestellt wird. Pakete welche über die ExpressRoute verschickt werden, werden nicht über das öffentliche Netz geroutet dadurch können zum einen höhere Performance Leistungen erzielt werden, als auch eine höhere Sicherheit und Zuverlässigkeit gewährleistet werden. ist somit aufwendiger in der Umsetzung bietet jedoch höhere Performance. Bei Einrichtung einer Express Route kann zwischen vier Modis unterschieden werden:

- Cloud Exchange Zusammenstellung: Der Cloud Exchange Modus eignet sich insbesondere wenn das Unternehmen bereits einen Cloud-Exchange nutzt. Bei einem Cloud-Exchange-Anbieter (i. d. R. bekannte Carrier wie AT&T oder Verizon) handelt es sich um einen Service Provider der sein Netzwerk mit gängigen Cloud Providern verbunden hat und diesen Zugang mittels eines WAN Zugangs an das Unternehmen weitergibt. Hierbei werden die Pakete direkt an den Cloud Exchange übergeben ohne dass dies über das öffentliche Netzwerk geroutet werden. In diesem Fall kann einfach eine weitere Querverbindung zur MS Cloud eingerichtet werden. In der Regel handelt es sich hierbei um Layer 2 oder Layer 3 Verbindungen.
- Point-to-Point Ethernet Verbindung: Auch PPPoE wird von Azure angeboten, bedarf jedoch die Unterstützung vom Internet Service Provider. Wie auch bei Cloud Exchange können auf diese Weise sowohl Layer 2 Verbindungen als auch verwaltete Layer 3 Verbindungen eingerichtet werden.
- Any to Any IPVPN Verbindung (MPLS): Any to Any IPVPN Verbindungen basieren i. d. R. auf MPLS VPN. In diesem Fall wird das private WAN eines Carriers genutzt. Dies setzt voraus, dass der Carrier als ExpressRoute Anbieter registriert ist. In dieser Konstellation wird eine verwaltete Layer 3 Verbindung genutzt.
- ExpressRoute Direct: Schließlich bietet Microsoft weltweit verteilte Peeringstandorte welche mittels Express Route Direct verbunden werden können. Hierbei können Spitzenübertragungen von bis zu 100 Gbit/s realisiert werden.

(EWo) **Global Reach und FastPath:** Weiterhin bietet Microsoft Azure über ExpressRoute Global Reach und Express Route FastPath eine Möglichkeit, die Latenz zu verringern. Sollten Sie zwei oder mehr ExpressRoute-Verbindungen in Ihrer Landschaft nutzen, so stellt ExpressRoute Global Reach eine mögliche Option zur Verringerung der Latenz dar. Da hierbei die Border Gateway Protocol-Route (BGP-Route) zwischen zwei ExpressRoute-Routingdomänen überbrückt wird, kann eine Reduzierung der Latenz erst erreicht werden, wenn mehrere ExpressRoute-Leitungen durchlaufen werden. Aktuell kann ExpressRoute Global Reach nur für privates Peering verwendet werden und bietet keine Möglichkeiten den Netzwerk-Zugriff zu ändern. Aufgrund dessen sollten Sie eine lokale Netzwerkdatenfilterung implementieren, um den Ressourcenzugriff geeignet zu limitieren. ExpressRoute FastPath ist standardmäßig bei neu erstellten ExpressRoute-Verbindungen implementiert. Bei bereits vorhandenen Verbindungen kann FastPath

über den Azure Support aktiviert werden. ExpressRoute FastPath wird auch als Microsoft Edge Exchange v2 definiert und verringert Netzwerkops für viele Datenpakete. Allerdings sollte beachtet werden, dass mit FastPath kein VNET-Peering kombiniert werden kann. Sollten für andere VNETs ein Peering implementiert sein, müssen Sie alle VNETs direkt mit der spezifischen ExpressRoute-Leitung verknüpfen. Sonst geht der Datenverkehr an das VNET-Gateway statt an die ExpressRoute.

### 6.3.5.3 Netzwerksicherheitsgruppen und die Azure Firewall

**Netzwerksicherheitsgruppen (NSGs)** steuern welcher Datenverkehr erlaubt ist und können sowohl auf Ebene eines virtuellen (Sub-) Netzes, als auch auf Ebene eines VM Interfaces definiert werden. Das Konzept der NSGs ist vergleichbar mit Paketfiltern wie IPtables. Eine NSG besteht dabei aus einer beliebigen Anzahl an einzelnen Regeln welche auf Basis der 5-Tuple Information Quell-IP, Quellport, Ziel-IP, Zielport und Protokoll eine Aktion (Allow oder Deny) definiert. Hierbei ist jede Regel mit einer Zahl zwischen 100 und 4096 priorisiert, wobei Regeln mit niedrigerer Priorität zuerst angewandt werden. Jedes eingehende oder ausgehende Paket wird mit der Regelsammlung gemäß deren Priorität verglichen und die erste Übereinstimmung gemäß der 5-Tuple Information wird angewandt. Weitere potentiell gültige Regeln mit höherer Priorität werde hingegen nicht weiter berücksichtigt. Im Kontext von SAP empfiehlt sich meistens ein Whitelisting Ansatz. D. h. zunächst werden alle notwendigen Zugriffe als Regel mit der Aktion „Allow“ definiert z. B. ein SSH Zugriff aus dem Management Netz. Abschließend wird dann eine „Deny All“ Auffangregel mit höchster Priorität gesetzt. Auf diese Weise wird sichergestellt, dass kein Zugriff möglich ist, außer dieser ist explizit gestattet. Eine NSG wird als eigenständige Ressource verwaltet kann auch definiert werden ohne dass diese einem Netz oder Interface zugeordnet wird (solange ist die NSG jedoch nicht aktiv). NSGs können außerdem wiederverwendet werden, beispielsweise kann eine NSG mehreren Ressourcen zugeordnet werden.

**Azure Firewall und Network Virtual Appliances (NVA)** Während die Konfiguration von geeigneten NSGs verpflichtend ist, kann optional auch die Bereitstellung einer Firewall lohnenswert sein. Diese werden sowohl von Drittparteien als „Network Virtual Appliance“ (NVA) oder von Microsoft als Azure Firewall angeboten. Die Unterschiede zwischen den jeweiligen Angeboten müssen individuell verglichen werden. Der zentrale Vorteil der Nutzung einer Firewall gegenüber NSGs ist, dass diese im Rahmen einer Hub-Spoke Topologie zentral in das DMZ Subnetz bereitgestellt werden kann und somit alle eigenen VNETs der Azure Landschaft an dieser Stelle absichert. Auch ist das Funktionsspektrum NVAs weit größer als die der NSGs. In der Regel arbeiten NVAs stateful und bieten fortgeschrittenere Regelkonfigurationen und Analysefunktionen wie Intrusion Detection an. Die Nutzung einer Firewall sollte immer ergänzend zu NSGs, nicht aber anstelle verwendet werden.

#### 6.3.5.4 Netzwerkschnittstellen

Allen VMs einer Azure Landschaft liegt eine gemeinschaftliche Software definiertes Netzwerk zu Grunde. Dadurch geht auch einheitliche Netzwerkfabric einher weshalb es aus Performance Sicht keinen Mehrwert bringt mehrere Netzwerkschnittstellen (NICs) für eine VM zu definieren. Standardmäßig unterstützt eine Azure NIC auch mehrere IP-Adressen wodurch auch die Nutzung von virtuellen Hostnamen ermöglicht wird (zur Nutzung siehe SNOTE 962955). Das manuelle Erstellen von zusätzlichen NICs eignet sich somit lediglich, wenn der Datenverkehr aufgeteilt werden soll, z. B. aus Monitoring Gründen.

#### 6.3.5.5 Netzleistung

Die Bandbreite eines virtuellen Computers wird immer auf Basis des von der gesamten Maschine ausgehenden Netzwerkverkehrs berechnet. Eingehende Daten werden nicht auf die Bandbreite angerechnet. D. h. dass **unabhängig** vom Kommunikationsziel (auch VMs im gleichen Subnetz) und unabhängig von der Anzahl der Netzwerkschnittstellen der kumulative Übertragungswert mit der zur Verfügung stehenden Bandbreite z. B. 1 Gbit/s verrechnet wird und im Falle eines Überschreitens auf die maximale Bandbreite gedrosselt wird.

Für virtuelle Maschinen kann darüber hinaus der beschleunigte Netzwerkbetrieb aktiviert werden. Dies basiert auf der SR-IOV Technologie welche den virtuellen Switch der Virtualisierungsschicht umgeht und dabei einen Großteil der Netzwerklast von der CPU wegnimmt und in die FPGA-beschleunigten Netzwerkinterfaces verschiebt. Diese Mehrleistung kommt insbesondere dann zur Geltung wenn auch andere VMs im virtuellen Netzwerk dieses Feature aktiviert haben. Im Kontext des SAP Hostings empfiehlt es sich das Feature sowohl für die VMs der Datenbankschicht als auch für die der Applikationsschicht zu aktivieren. Hierzu muss die VM gestoppt und die Zuordnung aufgehoben sein.

#### 6.3.5.6 Lastenausgleich

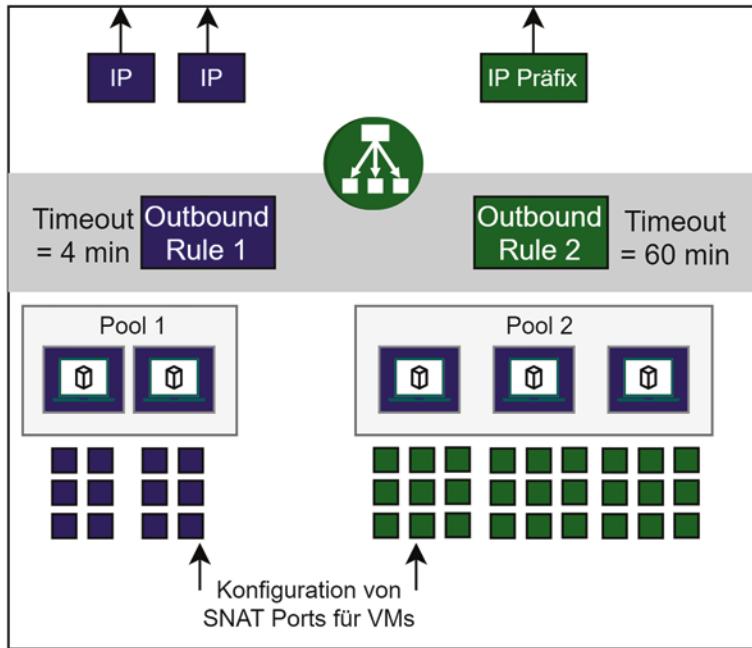
Ein Lastenausgleich kann jegliche TCP oder UDP basierte Anfragen entgegennehmen und verteilt diese auf Basis eines konfigurierten Regelwerks an die jeweiligen Backend Systeme. Für das SAP Hosting ist ein Lastenausgleich daher ein essentieller Baustein und kommt insbesondere in Hochverfügbarkeitsszenarien zum Einsatz indem es eine feste virtuelle IP geben muss welche dynamisch an die primäre Instanz weitergeleitet werden soll. Im Wesentlichen besteht ein Lastenausgleich aus vier Komponenten:

1. **Frontend-End-IP(s):** Definiert eine oder mehrere IPs (IPv4 und IPv6 unterstützt) über die der Lastenausgleich angesprochen wird. Beim Anlegen eines Lastenausgleichs muss definiert werden ob es sich um einen internen oder öffentlichen Lastenausgleich handeln soll. Ein öffentlicher Lastenausgleich kann mit öffentlichen Front-end-IPs versehen werden und somit von außerhalb angesprochen werden. Ein typisches Beispiel hierfür wäre ein Server-Cluster für ein Kunden Servicedesk.

Bei einem internen Lastenausgleich können hingegen lediglich private IP Adressen angegeben werden, beispielsweise um eine hochverfügbare HANA Cluster an die S/4HANA Applikationsserver anzubinden.

2. **Back-End-Pool(s):** Ein Back-End-Pool ist eine Sammlung von Zielsystemen an welche der Loadbalancer die eingehenden Anfragen weiterleiten soll. In einen Pool können entweder Maschinen, oder eine VM-Skalierungsgruppe aufgenommen werden. Es können außerdem mehrere Pools definiert werden.
3. **Integritätstest(s):** Ein Integritätstest bestimmt welche Instanzen eines Backend Pools als Ziel des Lastenausgleichs infrage kommen. Dazu definiert der Test eine Art Heartbeatmechanismus bestimmt durch ein Protokoll (TCP/UDP), einen Zielport, einem Intervall (in Sekunden), sowie einen Fehlerschwellenwert. Der Integritätstest überprüft daraufhin regelmäßig die Konnektivität zum Zielport des Backend Systems. Sollte dieser Test öfter fehlschlagen als der Fehlerschwellenwert erlaubt wird die Instanz nicht länger als mögliches Ziel des Lastenausgleichs berücksichtigt. Auf Seite der Backend Instanze erfordert dies, dass ein Service auf dem angegebenen Port horcht (z. B. netcat, socat oder azure-lb).
4. **Lastenausgleichsregel:** Die Lastenausgleichsregel verknüpft schließlich Frontend-IP, Back-End-Pool und Integritätstest zu einer Regel. Über die Zuweisung einer Frontend-IP wird zunächst definiert für welche IP Adresse die Lastenausgleichsregel angewendet werden soll, so können unterschiedliche Regeln für auf IP Ebene definiert werden. Port und Protokoll (UDP/TCP) bestimmen für welche Arten von Anfragen die Regel gültig ist. Back-End-Port und Backend-Pool definieren schließlich das Ziel an welche der Lastenausgleich die Anfrage weitergegeben werden soll. Außerdem muss der Integritätstest festgelegt werden der zur Prüfung der geeigneten Back-End Systeme angewendet werden soll. Dadurch ist es möglich für jede Regel eine eigene Prüfung zu hinterlegen. Durch Definition einer **Sitzungspersistenz** kann außerdem konfiguriert werden ob für einen Client der Datenverkehr einer Sitzung immer zum gleichen Back-End-System weitergeleitet werden soll („Client-IP“ oder „Client-IP und Protokoll“) oder diese von beliebigen VMs erfolgen kann („Keine“). Der **Leerlauftimeout** bestimmt schließlich wie lange eine TCP- bzw. HTTP Verbindung offen gehalten werden soll, auch wenn keine entsprechende Meldung durch den Client erfolgt. Die folgende Abbildung zeigt wie über Ausgangsregeln das Verhalten von ausgehenden Quell-Netzwerkadressenübersetzungs-Verbindungen konfiguriert werden können (Abb. 6.6):

**Basic vs. Standard** Der Lastenausgleich ist sowohl als kostenlose „Basic“ Variante, als auch als kostenpflichtige „Standard“ Variante verfügbar. Grundsätzlich lassen sich die meisten Use Cases wie z. B. SAP-HA auch mit einem Basic Lastenausgleich umsetzen. Die Standard Variante bietet jedoch ein erweitertes Funktionsspektrum wie z. B. die Nutzung von HA-ports welche ein 1:1 Mapping von Front-end-ports auf Back-end-ports ermöglicht, sowie größere Back-end-pools mit bis zu 1000 Instanzen. Für weitere Unterschiede zwischen den beiden SKUs siehe die offizielle Dokumentation [9].



**Abb. 6.6** Konzept des Lastenausgleichs

### 6.3.6 Support und Lizensierung

Die Lizensierung von SAP Produkten muss vom Kunden durchgeführt werden. Bei Bestellungen über den Azure Marketplace ist stets angegeben wie die Lizensierung abgewickelt wird. Bei der Wahl des Betriebssystemimages gibt es beispielsweise eigene „Bring your own subscription“ (BYOS) Vorlagen.

Beim Weg in die Azure Cloud sollte außerdem neben dem klassischen SAP Support die Buchung zusätzlicher Supportleistungen in Erwägung gezogen werden. Microsoft bietet eine Reihe von Supportplänen an. Der Basic Plan ist standardmäßig für alle Kunden kostenlos inbegriffen und ermöglicht es grundsätzlich Supporttickets zu erstellen. Für nicht produktive SAP Landschaften kann der Standard Supportplan in Erwägung gezogen werden. Für 84,33 €/Monat ist hiermit bereits 4 h \* 7 Support mit Reaktionszeiten von unter einer Stunde bei schwerwiegenden Beeinträchtigungen eingeschlossen. SAP empfiehlt als Minimum für alle Azure Kunden die Buchung des Professional Direct Supports (ca. 843,3 €/Monat) welches sich an Unternehmen mit geschäftskritischen Landschaften richtet. Neben den in Standard abgedeckten Supportleistungen sind darin zusätzliche Angebote wie Webinare sowie die programmgesteuerte Verwaltung von Supporttickets über die Microsoft Support-API inbegriffen. Im Falle von SAP Hosting auf Basis von Windows Server oder MS SQL Server empfiehlt Microsoft

darüber hinaus die Buchung von Premium Support. Beim Betrieb von Linux (SLES/RHEL) ist ein gültiger Support-Vertrag Voraussetzung. Auch hierbei werden unterschiedliche Leistungsklassen von Support angeboten z. B. Standard vs. Priority Support bei SUSE. Mit Hinblick auf Azure ergeben sich jedoch hier keine Sonderbedingungen gegenüber dem on-premise Hosting.

---

## 6.4 Azure Disaster Recovery Dienste

In diesem Kapitel werden die in Azur verfügbaren Disaster Recovery Dienste vorgestellt. Zunächst wird hierfür der Dienst „Azure Backup“ erläutert der als Oberbegriff für die jeweiligen Sicherungsdienste steht. Anschließend werden die zwei Arten von Tresoren (engl. vaults) vorgestellt, welche die zentralen Bausteine des Azure Sicherungskonzepts darstellen. Wir werden außerdem die Sicherung von Datenträgern, virtuellen Maschinen und SAP HANA Datenbanken erläutern. Abschließend wird der Dienst „Azure Site Recovery“ vorgestellt welcher neben regionsübergreifender Replizierung auch Migrationsszenarien unterstützt.

### 6.4.1 Azure Backup

Azure Backup ist ein SaaS Dienst der es erlaubt Sicherungen für verschiedenste Ressourcen wie virtuelle Maschinen, Anwendungen (z. B. MS Sharepoint oder Datenbanken) und sogar on-premise Systeme durchzuführen. Unter Anderem bietet der Dienst auch eine nahtlose Unterstützung von SAP HANA Datenbanken an. Sicherungen werden auf Basis einer **Sicherungsrichtlinie** erstellt. Diese definiert die Planungsparameter wie z. B. die Häufigkeit der automatischen Sicherungen, sowie Aufbewahrungszeiten (Retention Policy). Zur Speicherung verwendet Azure sogenannte Tresore. Bei dieser Ressource wird zwischen zwei Typen unterschieden, den sogenannten „Recovery Services vaults“ und den Sicherungstresoren (Backup vaults). Zentraler Einstiegspunkt für das Verwalten von Sicherungen und das Erstellen von Tresoren ist der Dienst „**Microsoft Backup Center**“. Dieser wird verwendet um Tresore und Richtlinien zu erstellen und Sicherungsvorgänge zu überwachen.

### 6.4.2 Backup Vault

Der Sicherungstresor (Backup vault) ist eine Speicherentität und wird benötigt um Backups und Snapshots für Azure-Blobs, Azure-Datenträger oder Azure Datenbanken für PostgreSQL-Server durchzuführen. Im Kontext des SAP Hostings ist die Sicherung von verwalteten Datenträgern ein gängiges Einsatzszenario. Hierfür kann zunächst ein Datenträger ausgewählt und eine Sicherungshäufigkeit definiert werden. Diese kann

zwischen 4 h bis maximal 24 h gewählt werden. Außerdem muss ein Aufbewahrungszeitraum angegeben werden welcher zwischen 1 bis 30 Tage betragen kann. Bei der Wahl ist jedoch zu beachten, dass die maximale Anzahl an Momentaufnahmen limitiert ist. Pro Datenträger können theoretisch bis zu 200 inkrementelle Sicherungsvorgänge vor gehalten werden. Azure beschränkt diese Sicherungen jedoch auf 180 um Kapazität für on-Demand Sicherungen vorzuhalten. Auf Basis dieser Planungswerte wird Azure nun regelmäßige und automatisch inkrementelle Snapshots durchführen. Diese werden im Sicherungstresor gespeichert. Die Daten des Backup vaults können sowohl georedundant (Standardeinstellung) als auch lokal redundant gespeichert werden. Die Konfiguration der Redundanz muss jedoch vor der ersten Sicherung erfolgen, da diese anschließend nicht mehr geändert werden können. Azure Backup funktioniert dabei ohne Agent Installation und führt zu keiner Verlangsamung des Systems. Kosten bei der Nutzung von Sicherungen fallen nur für die Speicherung der inkrementellen Daten an. Bepreist wird also nur das Delta zwischen den beiden letzten Sicherungen.

### 6.4.3 Recovery Services Vault

Ähnlich zum Sicherungstresor ist auch der Recovery Services-Tresor eine Speicherentität. Dieser wird jedoch zur Sicherung anderer Komponenten benötigt z. B. virtueller Maschinen, Datenbanken oder Azure Files. Außerdem wird der Recovery Services Tresor zur Notfallwiederherstellung mittels Azure Site Recovery benötigt.

#### 6.4.3.1 Sicherung von Azure VMs

Für die regelmäßige Sicherung einer VM kann ähnlich zur Datenträgersicherung eine Häufigkeit (täglich oder wöchentlich+Uhrzeit), sowie eine Aufbewahrungsdauer (7 bis 9999 Tage) definiert werden. Neben der Aufbewahrung der täglichen Sicherungen können zudem auch die Vorhaltung wöchentlicher, monatlicher und jährlicher Sicherungspunkte konfiguriert werden. Standardmäßig wird bei der Durchführung einer VM Sicherung ein Snapshot aller Datenträger mit Ausnahme der temporären Disk durchgeführt. Mit der Option „**Nur Sicherung des Betriebssystemdatenträgers**“ kann angegeben werden, dass lediglich der OS Datenträger gesichert werden soll. Das Hinzufügen einer VM zur regelmäßigen Sicherung funktioniert ohne manuellen Aufwand. Beim erstmaligen Konfigurieren muss lediglich die VM ausgewählt und mit einer Sicherungsrichtlinie welche die genannten Planungsparameter definiert verknüpft werden. Im Hintergrund installiert Azure automatisch die hierfür notwendige Erweiterung auf dem Betriebssystem der VM. Mit Einrichtung des Sicherungsprozesses kann nun auch jederzeit manuell eine on Demand Sicherung durchgeführt werden. Beachten Sie, dass die Sicherung einer VM durchaus mehrere Stunden in Anspruch nehmen kann.

#### 6.4.3.2 Sicherung von HANA DBs

Auf Basis des Recovery Services-Tresor kann auch die Sicherung von SAP HANA Datenbank auf Basis der Backint Schnittstelle durchgeführt werden. Azure Backup verspricht

hierbei eine RPO von 15 min und bietet damit eine großartige Alternative, wenn keine eigene Backup-Infrastruktur aufgebaut werden soll. Um eine HANA mittels Azure Backup zu verwalten, müssen zwei Phasen durchgeführt werden. Im Rahmen der „Discovery“ (Phase1) werden zunächst die Datenbank auf den verfügbaren VMs ermittelt. Nachdem die Discovery gestartet wurde, werden alle verfügbaren virtuellen Computer gelistet. Diese müssen sich in der gleichen Region wie der Tresor befinden. Anschließend wird Ihnen ein Skript zum Download angeboten, welches Sie auf allen VMs mit HANA DB ausführen müssen. Das Skript erfordert Root Rechte und benötigt die Angabe eines „System Schlüssels“. Bei dem Schlüssel handelt es sich um einen hdbuserstore Key für den SYSTEM Benutzer. Auf diese Weise erhält Azure Backup die erforderlichen Rechte, um auf die HANA Datenbank zugreifen zu können. Anschließend können die jeweiligen VMs ausgewählt werden und die Discovery Phase damit abgeschlossen werden. Nun können die eigentlichen Backups konfiguriert werden. Für jede einzelne (System-) Datenbank bzw. Tenant können Sie eine eigene Backup Policy hinterlegen. Hierbei können Sie, ähnlich zur Sicherung von virtuellen Maschinen, zwischen der Aufbewahrung von täglichen, wöchentlichen monatlichen und Jährlichen Backups unterscheiden. Außerdem werden Backup Policies nach Backupart angelegt, d. h. es können eigene Policies für Voll-, Differentielle- und Inkrementelle Sicherungen angelegt werden.

#### **6.4.3.3 Azure Site Recovery**

Site Recovery ist ein Azure Dienst zur Notfallwiederherstellung welcher ebenfalls auf dem Recovery Services-Tresor basiert. Es unterstützt nicht nur die Replikation ganzer Azure Landschaften (z. B. Virtuelle Maschinen, Datenträger, virtuelle Netze, Ressourcengruppe, ...) sondern auch on-premise Systeme basierend auf VMware und Hyper-V. Auf diese Weise lässt sich Azure Site Recovery nicht nur für die Notfallwiederherstellung nutzen, sondern auch zur Migration von Workloads in die Cloud. Bei der Nutzung von Site Recovery für Azure VMs muss sich der Tresor in einer anderen Region als die Quelle befinden, ansonsten ist lediglich eine Replizierung im Rahmen von Verfügbarkeitszonen (innerhalb einer Region) möglich. Nachdem eine VM hinzugefügt wurde, werden alle Metadaten der VM (z. B. Größe, verbunden Netzwerkkomponenten, etc.) im Tresor gespeichert. Im Rahmen eines Failovers ist es ausreichend diese Ressourcen auf Basis der Metadatendefinition in der Zielregion anzulegen. Die Datenträger werden hingegen tatsächlich repliziert. Datenmodifikationen (nach der initialen Replikation) werden zunächst in den Speicherkonten Cache geschrieben und anschließend in die Zielregion repliziert. Auf diese Weise wird sichergestellt, dass der Produktivworkload nicht beeinträchtigt wird. Die Replikation selbst lässt sich individuell konfigurieren, so können etwa einzelne Datenträger ausgenommen und das Abonnement der replizierten Ressourcen geändert werden. Für das Failover der Ressourcen ist es möglich Wiederherstellungspläne zu definieren. In diesen lässt sich die Notfallwiederherstellung orchestrieren, um Abhängigkeiten zwischen Ressourcen abzubilden. Beispielsweise könnte für eine S/4HANA 3-Tier Architektur die Reihenfolge definiert werden, dass zunächst die VMs der HANA Datenbank, anschließend die ASCS Instanz

und abschließend die VMs der Anwendungsserver hochgefahren werden. Wiederherstellungspläne bieten für Testzwecke auch ein Testfailover an. In diesem Fall werden die replizierten Maschinen hochgefahren ohne dass der Primärstandort abgeschaltet wird.

---

## 6.5 S/4 auf Azure Architektur

In diesem Kapitel wird eine vollständige Architektur vorgestellt, die den Service Level Agreements der Azure-Umgebung und der angebotenen Verfügbarkeit der Azure-Komponenten entspricht. Diese Architektur umfasst eine Diskussion über Hochverfügbarkeit, Disaster Recovery, Sicherung und Wiederherstellung sowie die Sicherheitsmaßnahmen, die zur Sicherung der Umgebung ergriffen werden müssen. Das Kapitel behandelt auch Scale-Out- und Scale-Up-Szenarien.

### 6.5.1 Grundlegende Referenzarchitektur

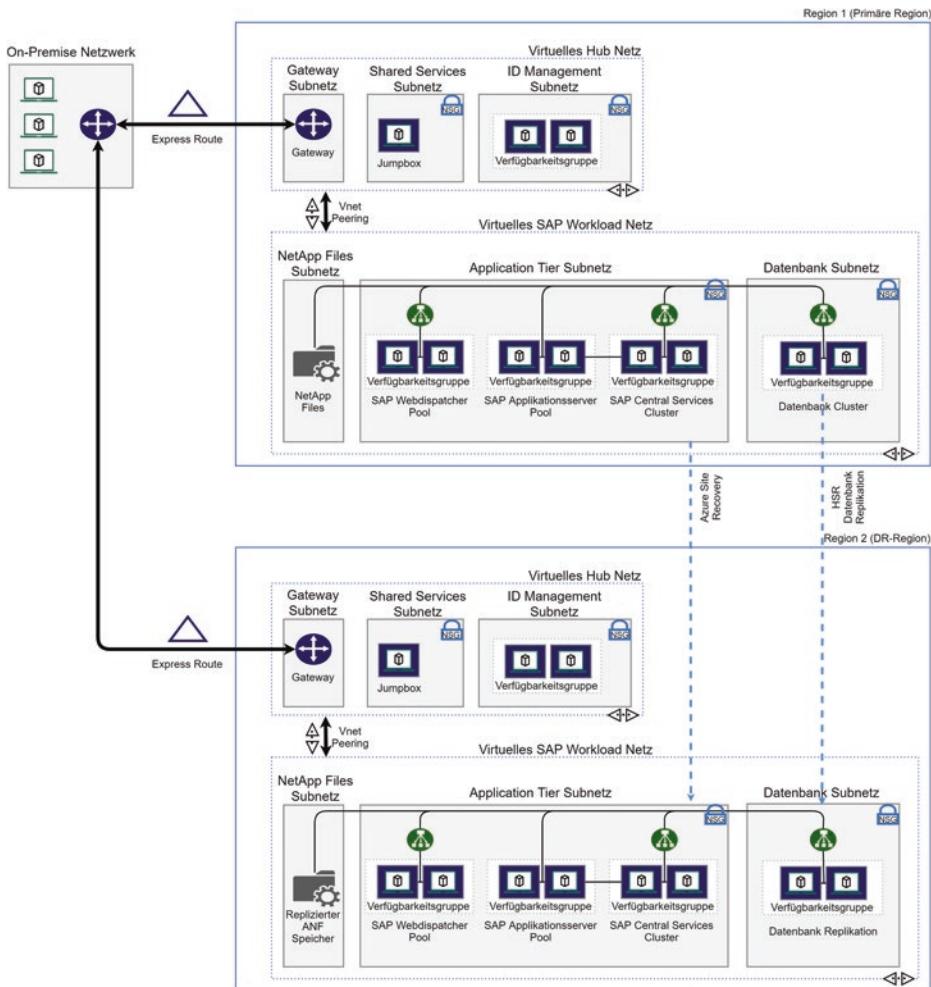
Die hier vorgestellte Referenzarchitektur besteht aus virtuellen Maschinen mit installiertem Linux-Betriebssystem. Sie umfasst Hochverfügbarkeitsstrategien im Umfeld von S/4HANA Systemen und unterstützt die Wiederherstellung bei Störfällen in der Microsoft Azure Umgebung. Dabei kann diese Architektur als standardisiertes Grundgerüst gesehen werden, das je nach Kontext und Unternehmensanforderungen angepasst werden kann.

Die Architektur ist in Abb. 6.7 visualisiert und umfasst die im Folgenden beschriebenen Komponenten.

Das **Azure Virtual Network (VNET)** stellt den Ausgangspunkt dar. Es verbindet die benötigten Azure-Ressourcen und verhindert unerlaubten Zugriff darauf. In dieser Architektur können die Ressourcen über das Gateway von außen erreicht werden. Dabei stellt das Gateway ein Hub und das VNET die Speiche des in Abschn. 6.3.3.1 beschriebenen Nabe-Speiche-Modells dar.

Über das **Gateway** werden die Netzwerke von außerhalb des VNETs verknüpft, es repräsentiert die Schnittstelle zwischen den Netzwerken. Dabei können Sie aus verschiedenen Optionen wählen, um sich über das Gateway mit den Azure Ressourcen zu verbinden. ExpressRoute wird zur Erstellung von sicheren Verbindungen von Azure empfohlen. Dabei sollten Sie bei der Verbindung zwischen dem VNET und dem außenliegenden Netzwerk die Latenzzeit beachten. Diese ist je nach gewählter Region der Azure Ressourcen unterschiedlich groß. Um die Latenz zu reduzieren, können Sie die Dienste ExpressRoute Global Reach und ExpressRoute FastPath verwenden. Alternativ kann auch ein virtuelles privates Netzwerk genutzt werden.

Das Spoke-Netzwerk im VNET enthält die virtuellen Instanzen mit den SAP-Anwendungen sowie zugehöriger Dienste. Über **VNET Peering** wird das Hub-Netzwerk mit dem Spoke-Netzwerk verknüpft. Das ermöglicht die Isolation verschiedener



**Abb. 6.7** Referenzarchitektur SAP S/4HANA auf Azure

Azure-Ressourcen voneinander. Durch das Peering können diese isolierten Ressourcen ohne Leistungsverringerung verknüpft werden, solange sie in der gleichen Region liegen. Im Hub- sowie im Spoke-Netzwerk werden weitere Subnetze realisiert, um die einzelnen Anwendungen voneinander zu trennen. Im Hub-Netzwerk existieren drei Subnetze, wie auch im Spoke-Netzwerk. In den Subnetzen befinden sich die virtuellen Recheninstanzen.

Die **SAP HANA** bildet das Fundament der S/4HANA Bereitstellung und wird anhand von zwei Linux-Instanzen hochverfügbar implementiert. Dadurch, dass die beiden Instanzen geclustert sind, können Sie die Datenbank auch zentral skalieren und ein automatisiertes Failover bei Störfällen realisieren. Für Hochverfügbarkeit ist zu beachten,

dass die beiden Instanzen in unterschiedlichen **Verfügbarkeitsgruppen** bereitgestellt werden. Dieses Vorgehen verhindert Systemausfälle bei Ausfällen einzelner Azure Ressourcen. Sollte aufgrund von Systemfehlern oder Wartungen eine Datenbankinstanz ausfallen, bleibt die andere Datenbank trotzdem funktionsfähig. Über Systemreplikation wird Redundanz zwischen den beiden HANA Datenbanken sichergestellt. Hierbei muss beachtet werden, dass ein Mechanismus implementiert werden muss, der das ausgefallene System herunterfahrt, um den sogenannten Split-Brain-Zustand zu vermeiden, welcher Dateninkonsistenzen ermöglicht.

Darauf aufbauend werden mehrere **Front-End-Server** aufgesetzt, zwei **SAP Web Dispatcher**, zwei **Anwendungsserver** und zwei Instanzen für die **SAP Central Services**. Dabei werden die beiden Instanzen der jeweiligen SAP Komponente in jeweils verschiedene **Verfügbarkeitsgruppen** lokalisiert, um eine bessere Verfügbarkeit zu realisieren.

Im dritten Subnetz des Spoke-Netzwerks befindet sich der **Netzwerkdateifreigabe-Dienst**. Um den Hochverfügbarkeitsanforderungen zu entsprechen, gibt es mehrere Möglichkeiten: Azure NetApp Files, ein NFS-Clusterserver oder der SIOS DataKeeper stellen die dafür nötigen Eigenschaften zur Verfügung. Die Verwendung des Pacemaker-Clusters ermöglicht Hochverfügbarkeit bei freigegebenen Azure-Datenträgern.

- ▶ Bei Verwendung von Red Hat Enterprise Linux können Sie GlusterFS für eine hochverfügbare Dateifreigabe nutzen.

Die Jumpbox ist im Hub-Netzwerk verortet und wird genutzt, um eine sichere Verbindung mit den anderen virtuellen Instanzen im VNET herzustellen. Man „springt“ sozusagen vom äußeren Netzwerk zuerst auf die Jumpbox, um darüber Zugriff auf die anderen virtuellen Computer zu erhalten. Auf der Jumpbox werden beispielsweise Dienste wie SAP HANA Studio, SAPGUI, SWPM betrieben, die zur Installation und Verwaltung der SAP Landschaft benötigt werden. Der Dienst Azure Bastion stellt eine gute Möglichkeit als Jumpbox dar, sollten Sie nur Remotedesktop- oder Secure Shell-Dienste für die Implementierung nutzen.

Über Netzwerksicherheitsgruppen (NSGs) können zusätzlich subnetinterne sowie aus- und eingehende Datentransaktionen eingeschränkt werden. In der beispielhaften Referenzarchitektur sind in allen Subnetzen NSGs implementiert außer für das Gateway und den Dateifreigabedienst.

Lastenausgleichsmodule sind in dieser Architektur vor den SAP Webdispatcher-Pool, den Central Services Cluster sowie den beiden HANA Datenbanken geschaltet. Damit kann der Datenverkehr auf die Instanzen verteilt werden. Dabei sollten Sie jedoch beachten, dass die virtuellen Instanzen hinter den Lastenausgleichsmodulen über standardmäßig keine ausgehende Verbindung haben.

Anhand von Näherungsplatzierungsgruppen kann optional noch die Zusammenlegung von virtuellen Computern im gleichen Rechenzentrum ermöglicht werden.

Über Anwendungssicherheitsgruppen können Instanzen gruppiert werden und für Anwendungen nur gefilterten Datenverkehr erlauben.

## 6.5.2 Virtuelles Privates Netzwerk Referenzarchitektur

Für die Implementierung eines virtuellen privaten Netzwerks wird ein lokales Netzwerk sowie ein Hub-and-Spoke-Netzwerk erstellt. Die beiden Netzwerke werden durch unterschiedliche Ressourcengruppen voneinander getrennt. Über AVN-Gateways werden die beiden Netzwerke miteinander verknüpft. Diese Art der Verknüpfung ähnelt technisch der Verbindung des Netzwerks außerhalb von Azure mit den Netzwerken innerhalb von Microsoft Azure. Über das Microsoft Azure Portal kann die Bereitstellung automatisiert erfolgen (ca. 45 min Bereitstellungsdauer). Abb. 6.8 veranschaulicht die Systemarchitektur.

Die Architektur umfasst die folgenden Komponenten. Das VPN-Gateway des Azure-Stack Netzwerk sendet Daten verschlüsselt über eine öffentliche Verbindung. Es besteht aus einem speziellen Gatewaysubnetz, einem Netzwerkgateway und einem Site-to-Site-Tunnel. Im virtuellen Netzwerk sind die Anwendungen und das Gateway sowie der Azure Bastion lokalisiert. Der Azure Bastion ermöglicht eine SSH- oder RDP-Verbindung auf die virtuellen Instanzen ohne dass die Instanzen über die Internetverbindung frei verfügbar sind. Im Falle einer Störung der VPN-Verbindung sind somit die Instanzen über die SSH- oder RDP-Verbindung immer noch ansteuerbar. Bei der Erstellung einer solchen Architektur sollten Sie beachten den Adressraum des virtuellen Netzwerks entsprechend zu definieren, da er sich nicht mit dem des lokalen Netzwerks überschneiden darf. Auch sollten Sie das Gateway-Subnet am oberen Ende des Adressraums des virtuellen Netzwerks zu platzieren. Bitte beachten Sie keine virtuellen Instanzen im Gatewaysubnetz zu lokalisieren und keine Netzwerksicherheitsgruppe dort zu definieren, da das Gateway sonst nicht mehr funktionstüchtig bleibt.

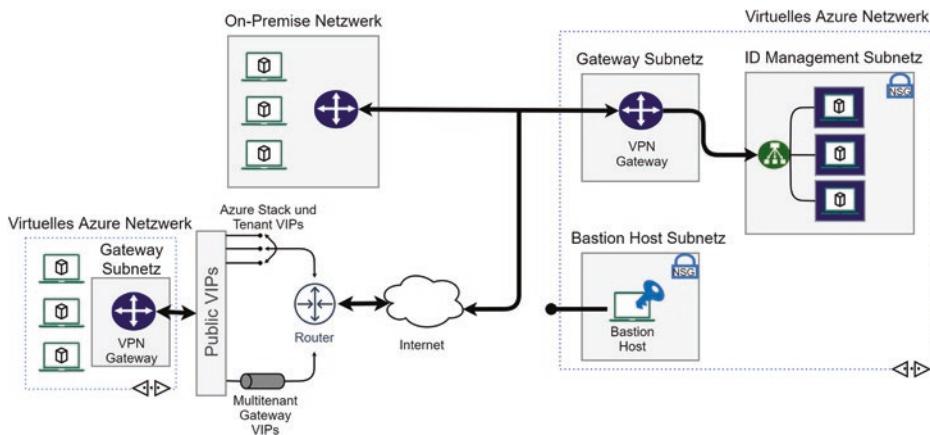


Abb. 6.8 Referenzarchitektur eines virtuellen privaten Netzwerks in Microsoft Azure

## 6.6 Zusammenfassung

Microsoft Azure gehört zu den größten Hyperscaler-Anbietern und viele Kunden betreiben die SAP-Umgebungen basierend auf Azure. Dieses Kapitel hat Ihnen gezeigt, welche wichtige Komponenten für solche SAP-Umgebungen wichtig sind. Ausgehend von einem kurzen historischen Abriss zu MS Azure, wurden die wichtigsten Punkte vor dem Beginn einer Bereitstellung beschrieben. Hierzu gehören die Punkte der Abonnements und der wichtigsten Benutzergruppen.

Das Kapitel zeigte Ihnen die Kernkomponenten von MS Azure für den Aufbau eines SAP S/4HANA-Systems, wie Compute, Storage und Netzwerkkomponenten, als auch deren Integration in die unternehmenseigenen Dienste.

Zur Absicherung der SAP-Systeme bietet Ihnen MS Azure wichtige grundlegende Funktionen, wie Azure DR zum Disaster Recovery an und kann auch mit einem grundlegenden Feature zur Sicherung und Wiederherstellung der Daten unterstützen.

Für produktive SAP S/4HANA-Systeme bietet Microsoft Premium Support an, welcher in jedem Fall genutzt werden sollte. Darüber erhalten Kunden wichtige Unterstützung und zeitnahen Support im Fehlerfall. Dies ist für produktive SAP S/4HANA-Systeme elementar.

Im folgenden Kapitel werden wir die konkrete Implementierung zeigen und die Schritte zur Bereitstellung eines neuen SAP S/4HANA-Systems in der MS Azure Cloud.

---

## Literatur

1. <https://docs.microsoft.com/de-de/azure/role-based-access-control/built-in-roles> (Zugriff am 20.12.2021).
2. <https://docs.microsoft.com/de-de/azure/cost-management-billing/manage/cost-management-budget-scenario> (Zugriff am 20.12.2021).
3. <https://docs.microsoft.com/en-us/cli/azure/install-azure-cli> (Zugriff am 20.12.2021).
4. <https://shell.azure.com> (Zugriff am 20.12.2021).
5. <https://launchpad.support.sap.com/#/notes/1928533> (Zugriff am 20.12.2021).
6. <https://www.sap.com/dmc/exp/2014-09-02-hana-hardware/enEN/#/solutions?filters=iaas;ve:24> (Zugriff am 20.12.2021).
7. <https://launchpad.support.sap.com/#/notes/2015553> (Zugriff am 20.12.2021).
8. <https://docs.microsoft.com/en-us/azure/vpn-gateway/vpn-gateway-download-vpndevicescript> (Zugriff am 20.12.2021).
9. <https://docs.microsoft.com/de-de/azure/load-balancer/skus> (Zugriff am 20.12.2021).



# SAP S/4 on Microsoft Azure – Deployment

7

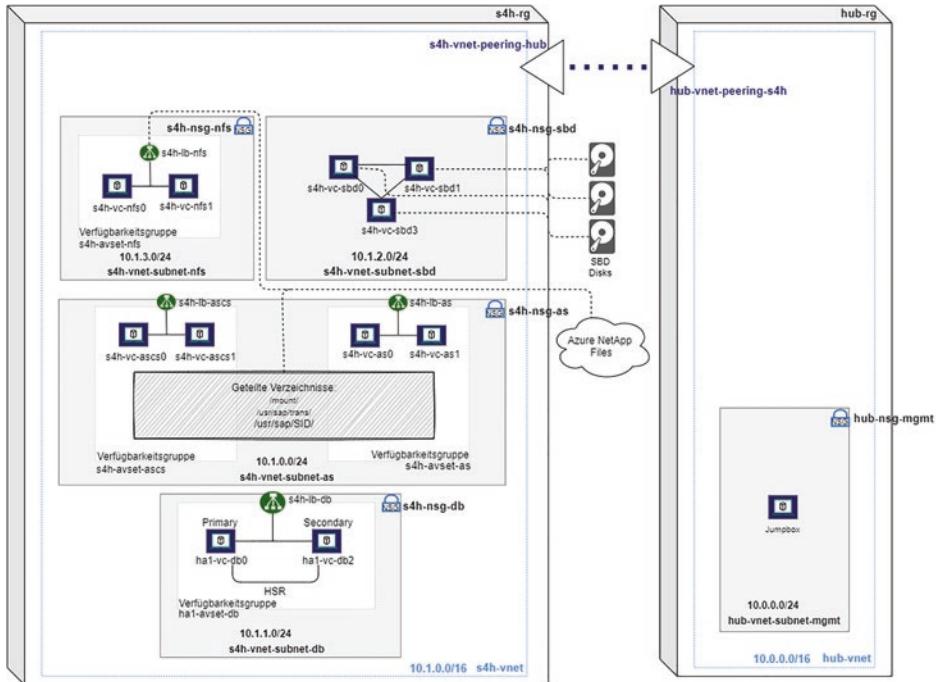
## Zusammenfassung

Dieses Kapitel wird die theoretischen Grundlagen aus Kap. 6 praxisorientiert vertiefen, indem wir Schritt für Schritt ein S/4HANA System auf Basis von Azure VM bereitstellen. Zu diesem Zweck werden wir zunächst die Beispielarchitektur vorstellen. Hierbei handelt es sich um die Systemressourcen welche Sie nach Ausführung aller Schritte als Endresultat erhalten werden. Anschließend werden wir der Reihe nach die grundlegende Netzwerkarchitektur, das HANA HA-Cluster und das S/4HANA HA Cluster bereitstellen. Wir werden hierbei sowohl zeigen wie Sie Ressourcen manuell bereitstellen, als auch wie Sie benutzerdefinierte ARM-Vorlagen erstellen und nutzen. Abschließend werden wir in einem kurzen Exkurs alternative Bereitstellungsautomatisierungen auf Basis der SAP CAL, sowie Ansible und Terraform beleuchten.

## 7.1 Azure S/4HANA Beispielarchitektur

Wie in Abb. 7.1 dargestellt werden wir in diesem Kapitel eine leichtgewichtige S/4 HANA Hochverfügbarkeitsarchitektur mit eigenem Managementnetz (Nabennetzwerk), sowie einem Workloadnetzwerk (Speiche) bereitstellen. Hierfür werden wir unterschiedliche Alternativen beleuchten. Das Beispiel sollte somit als Übungsszenario betrachtet werden und nicht als Best Practice Empfehlung für eine Produktivimplementierung. Die Konfigurationsschritte auf Betriebssystemebene basieren auf SUSE Linux Enterprise 12 SP5. Anleitungen für Windows und RedHat Linux werden an entsprechender Stelle verlinkt.

Für die Beispieldurchführung verwenden wir die folgenden Angaben (Tab. 7.1):



**Abb. 7.1** Zielarchitektur für SAP S/4HANA on Azure

**Tab. 7.1** Namenskonventionen

Komponente	Bezeichnung
<b>SID – HANA</b>	<b>HA1</b>
HANADB (HA1) – Instanznummer	03
<b>SID – Netweaver</b>	<b>S4H</b>
ASCS – Instanznummer	00
ERS – Instanznummer	02
PAS – Instanznummer	01
AAS – Instanznummer	04

**Hochverfügbarkeitsaspekte** Für das Fencing bietet Azure zwei Alternativen an basierend auf SBD und dem Azure Fencing Agent. Wir werden SBD für das Fencing aller Pacemaker Cluster nutzen, jedoch im Rahmen eines Exkurses auf die Erstellung eines Fence Agents inklusive STONITH Konfiguration eingehen. Als Hochverfügbaren geteilten Speicher kann entweder ein eigenes HA-NFS Cluster bereitgestellt werden oder auf den SaaS Dienst Azure Netapp Files zurückgegriffen werden. Wir verwenden das NFS Cluster als SAP Mount Verzeichnis/sapmnt und nutzen ANF für das Transportverzeichnis/usr/sap/trans. Es empfiehlt sich in der Praxis bei nur einer Variante zu bleiben.

## 7.2 Bereitstellen einer Netzwerkbasiskonfiguration via Azure Cloud Portal

Wir beginnen die Bereitstellung unserer Zielarchitektur zunächst mit einer Grundkonfiguration. Wie in Abb. 7.2 dargestellt beinhaltet diese neben der Ressourcengruppe auch die Hub-Spoke Netzwerkarchitektur inklusive Peering, Subnetzkonfiguration und Netzwerksicherheitsrichtlinien. Die Bereitstellung und Konfiguration wird im Folgenden am Beispiel des Azure Cloud Portals erläutert.

### 7.2.1 Azure Cloud Portal

Das Azure Cloud Portal ist das zentrale Werkzeug für sämtliche Verwaltungsvorgänge und ermöglicht unter anderem auch die Erstellung von Ressourcen mittels einer grafischen Oberfläche. Das Portal ist dabei lediglich eine von insgesamt vier Schnittstellen. Insbesondere für Bereitstellungen mittels ARM Vorlagen oder auch für automatische Ressourcenverwaltungsvorgänge nehmen die Azure CLI, die Azure Powershell und die REST API einen hohen Stellenwert ein. Als Einstieg ist das Portal jedoch besonders geeignet da es erstens, einen Einblick in das umfangreiche Angebot der Azure Plattform bietet, wie z. B. den Azure Marketplace, welches auch Angebote von dritten beinhaltet, und zweitens Eingabehilfen und ergänzende Beschreibungen enthält (Abb. 7.3).

**Navigation** Das Azure Portal bietet eine Hashtag Navigation wodurch einfach mittels Ergänzung des jeweiligen Hashtags in der URL in den entsprechenden Azure Dienst navigiert werden kann. Die URL (Zugriff am 20.12.2021) <https://portal.azure.com>.

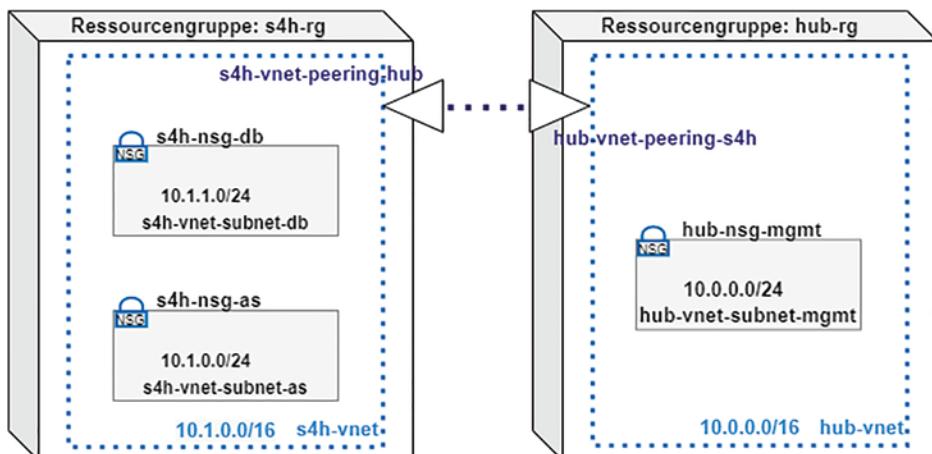
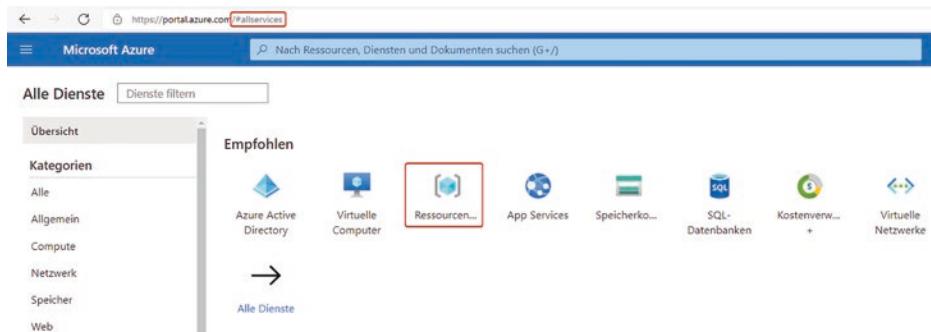


Abb. 7.2 Grundkonfiguration



**Abb. 7.3** Konfiguration

[com/#home](#) zeigt beispielsweise auf das Homeverzeichnis. Im Rahmen dieses Kapitels werden wir an entsprechender Stelle den Zielhashtag angeben, um die Navigation zu erleichtern. Sollten Sie den entsprechenden Dienst nicht finden können sie auch jederzeit die zentrale Suche verwenden. Nützliche Hashtags sind beispielsweise die folgenden:

[#home](#) – Verweist auf das Homeverzeichnis. Von hier kann direkt auf kürzlich verwendete Ressourcen, zentrale Azure Dienste und Informationsangebote abgesprungen werden.

[#allservices](#) – Bietet eine kategorisierte Übersicht über alle Azure Dienste. Von dieser Seite kann mit wenigen Klicks in alle relevanten Azure Dienste abgesprungen werden.

[#create/hub](#) – Direktlink auf den Azure Marketplace. Von diesem Punkt aus kann das breite Angebot an Ressourcen und Dienstleistungen durchstöbert werden und direkt in die Erstellung von Ressourcen abgesprungen werden.

[#create/Microsoft.VirtualMachine-ARM](#) – Navigiert direkt in die Erstellungsmaske eines virtuellen Computers auf Basis des Azure Ressourcen Managers. Sie werden im Verlauf dieses Kapitels einige virtuelle Computer anlegen.

[#create/Microsoft.LoadBalancer-ARM](#) – Navigiert direkt in die Erstellungsmaske zum Anlegen eines Lastenausgleichs über den Azure Ressourcen Manager. Sie werden im Verlauf dieses Kapitels mindestens zwei Lastenausgleichs Komponenten anlegen.

## 7.2.2 Namenskonventionen und der Ressourcenbegriff

Alle Objekte, welche Sie in Azure anlegen wie virtuelle Maschinen, Netzwerke oder Disks werden unter dem Sammelbegriff „Ressourcen“ geführt. Viele dieser Objekte sind

im on-premise Hosting lediglich ein Teil einer Konfiguration, werden in Azure jedoch explizit als Ressource verwaltet. Beispielsweise würden Sie on-premise lediglich eine öffentliche IP oder Netzwerkeinschränkungen auf Basis von iptables konfigurieren. In Azure gibt es hierfür separate Ressourcen wie z. B. „öffentliche IP-Adressen“ oder „Netzwerksicherheitsgruppen“. Als Resultat werden Sie in Azure mit einer Vielzahl an Ressourcen arbeiten welche über eindeutige Namen identifiziert werden müssen. Aufgrund der Vielzahl dieser Objekte bietet es sich an im Voraus ein Namensschema zu etablieren, um Ressourcen und deren Beziehungen zueinander eindeutig beschreiben zu können. Um Flexibilität in der Benennung zu ermöglichen, sollte zumindest ein Präfix-Schema definiert werden. Das Schema, welches wir im Rahmen dieses Buches verwenden ist das Folgende (Namen in [] sind optionale Bestandteile):

```
<sid>-<ressourcen typ>-[<rolle>[nr]]-[<kind ressourcen typ>[nr]-[<bezeichnung>]]
```

Die minimale Bezeichnung einer Ressource besteht damit aus <sid> und <ressourcen typ>. Die <sid> wird vom künftigen SAP-System abgeleitet, in unserem Beispiel also „s4h“ bzw. „ha1“ für die Datenbank. Der Ressourcentyp ist eine Abkürzung für die Art der jeweiligen Ressource. Für eine Verfügbarkeitsgruppe (engl. availability set) könnte dies z. B. „avset“ sein, im Falle eines Lastenausgleichs (engl. Load Balancer) „lb“. Die <rolle> spezifiziert optional den Einsatzzweck der Ressource, falls es mehrere Ressourcen des gleichen typs geben kann. Sollte z. B. eine Verfügbarkeitsgruppe für die Datenbank verwendet werden könnte die Rolle z. B. „db“ sein, während „asc“ für ein ASCS/ERS Cluster verwendet wird. Ressourcen, welche Bestandteil eines anderen Objektes sind und zu diesem in einer 1:1 Beziehung stehen wie z. B. ein Subnetz oder ein Datenträger, haben als Präfix den Ressourcennamen des übergeordneten Objektes sowie ihren eigenen Ressourcentyp. Ein verwalteter Datenträger des virtuellen Computers h41-vc-db0, könnte z. B. die Bezeichnung h41-vc-db0-disk-data0 aufweisen, um darauf hinzuweisen, dass diese unter/hana/data eingebunden ist.

### 7.2.3 Anlegen der Ressourcengruppen

Der erste Schritt im Rahmen dieser Bereitstellung ist das Anlegen einer Ressourcengruppe. Alle für das S/4HANA-System benötigten Ressourcen wie die virtuellen Maschinen und Lastenausgleiche werden später dieser Gruppe hinzugefügt und unterliegen dem gleichen Lebenszyklus. Es ist auch möglich mehrere Ressourcengruppen zu verwenden, um z. B. die Applikationsserver und Datenbank in getrennten Gruppen zu verwalten.

Navigieren Sie nach **#allservices** und wählen Sie die Kategorie „Allgemein“ und anschließend „Ressourcengruppen“. Sie befinden sich nun in der Ansicht des Dienstes „Ressourcengruppen“ [1]. Klicken Sie auf den Button „Erstellen“ um eine neue Gruppe

hinzuzufügen. Wählen Sie das Abonnement welches für die Abrechnung verwendet werden soll, sowie eine Region. Legen Sie außerdem einen Namen für die Ressourcengruppe fest wie z. B. **s4h-rg**.

Klicken Sie anschließend auf „Überprüfen+Erstellen“. Azure prüft nun ob die von Ihnen angeforderten Ressourcen bereitgestellt werden können. Zum einen werden diese mit den von Ihnen festgelegten Kontingenzen abgeglichen. Zum anderen wird auch die Verfügbarkeit in der Zielregion überprüft. Die Überprüfung der Ressourcengruppe sollte jedoch nur wenige Sekunden dauern da es in diesem Fall nur wenige Abhängigkeiten gibt. Klicken Sie abschließend auf „Erstellen“. Wählen Sie nun die im vorherigen Schritt erstellte Ressourcengruppe aus. Machen Sie sich zunächst mit der linken Navigationsleiste vertraut. Mit einem Klick auf eine der Funktionen öffnet sich eine neue Detailsicht welche die entsprechenden Informationen anzeigt. Die folgenden Sichten möchten wir kurz vorstellen:

- Übersicht: Die Übersicht zeigt Ihnen alle in der Gruppe enthaltenen Ressourcen an. Außerdem werden dort die meisten Verwaltungsvorgänge angestoßen wie z. B. das Erstellen oder verschieben einer Ressource sowie das Löschen einer Gruppe
- Aktivitätsprotokoll: Im Aktivitätsprotokoll werden alle Vorgänge innerhalb der Ressourcengruppe protokolliert. Dies beinhaltet auch Vorgänge wie das stoppen oder löschen einer virtuellen Maschine sowie den Initiator des Vorgangs
- Zugriffssteuerung: Hier können aktuelle Rollenzuweisungen eingesehen werden sowie neue Zugriffe hinzugefügt werden.
- Ressourcenschnellansicht: Wenn Sie Ressourcen in der Übersicht selektieren können Sie über die Ressourcenschnellansicht einen Graphen erzeugen welcher Ihnen die Ressourcen sowie deren Abhängigkeiten in einer grafischen Baumdarstellung anzeigt.
- Bereitstellungen: Ähnlich zum Aktivitätsprotokoll werden hier die Bereitstellungen neuer Ressourcen gelistet jedoch keine anderweitigen Verwaltungsvorgänge. Hierbei handelt es sich jedoch nicht um ein schlichtes Protokoll. Die ARM-Vorlage aller vergangenen Bereitstellungen werden hier vorgehalten und können jederzeit erneut bereitgestellt werden.

Wiederholen Sie die Schritte für eine weitere Ressourcengruppe „**hub-rg**“.

## 7.2.4 Netzwerkkonfiguration

Basierend auf der Ressourcengruppe wollen wir als nächstes die virtuellen Netzwerke erstellen. Grundsätzlich müssen Sie das virtuelle Netzwerk nicht separat erstellen, sondern können es auch direkt mit dem Anlegen der ersten VM hinzufügen lassen. Zur besseren Planung führen wir die Netzwerkkonfiguration als eigenständigen Schritt durch. Navigieren Sie dazu zunächst in die Übersicht der Ressourcengruppe „**hub-rg**“ und klicken Sie auf den Button „Erstellen“. Geben Sie in die Suche „virtual network“

ein (zum Zeitpunkt dieses Buches muss der englische Begriff für die Suche verwendet werden). Klicken Sie zunächst auf den Objektnamen. Sie erhalten nun eine Detailsicht mit einer kurzen Beschreibung der Ressource. Sie sollten einen Hinweis auf die Bereitstellung mittels des Ressource Manager sehen. Betätigen Sie nun den Button „Erstellen“.

Sie befinden Sich nun in der Konfigurationsmaske zum Anlegen eines virtuellen Netzwerks. Achten Sie darauf, dass die oberste Beschreibung der Maske „Virtuelles Netzwerk erstellen“ lautet und nicht das Suffix „(klassisch)“ aufweist. Andernfalls würde das Objekt als Teil der klassischen Ressourcenbereitstellung provisioniert werden und nicht über den Azure Ressource Manager.

Abgeleitet von Ihrer Ressourcengruppe sollten bereits die meisten Felder vorausgefüllt sein. Prüfen Sie zunächst in den Grundeinstellungen die Region und das ausgewählte Abonnement. Legen Sie anschließend „hub-vnet“ als Netzwerknamen fest. Klicken Sie auf „Weiter: IP-Adressen“ und legen Sie die IP-Range auf **10.0.0.0/16** fest. Fügen Sie nun ein Subnetz mit der Bezeichnung „hub-vnet-subnet-mgmt“ und dem IP Adressraum **10.0.0.0/24** hinzu (kein NAT-Gateway). Navigieren Sie anschließen weiter zum Reiter „Sicherheit“. Hier haben Sie die Möglichkeit einen BastionHost, sowie eine Firewall und DDoS Protection in die Bereitstellung mitaufzunehmen. Beim Bastion Host handelt es sich um ein PaaS Angebot zur Verwaltung der virtuellen Maschinen in der Azure Landschaft. Der Bastion Host definiert eine öffentliche IP mittels derer Sie sich auf die jeweiligen Zielmaschinen auf Basis von SSH oder RDP weiterverbinden können. Im Rahmen unserer Hub- / Spoke- Netzwerkarchitektur ist dies ein klassischer Service den man in den Hub bereitstellen könnte. Alternativ können Sie auch klassisch einen virtuellen Computer als „Jumpbox“ definieren. Beachten Sie, dass es sich in allen drei Fällen um gebührenpflichtige Dienste handelt. Schließen Sie anschließend die Bereitstellung ab.

Navigieren Sie nun in ihre Ressourcengruppe „s4h-rg“ und wiederholen Sie dort die Bereitstellung für das virtuelle Netzwerk „**s4h-vnet**

- **s4h-vnet-subnet-as** [10.1.0.0/24] # Subnetz für die Applikations Server und Central Instances
- **s4h-vnet-subnet-db** [10.1.1.0/24] # Subnetz für die Datenbank Server

Schließen Sie die Bereitstellung ab und wechseln Sie zur neu angelegten VNET Ressource. Wie bereits von den Ressourcengruppen bekannt finden Sie am linken Bildschirmrand eine Navigationsleiste mit zahlreichen Einstellungen. Über die Menüpunkte „Adressbereich“ und „Subnetze“ können Sie jederzeit die IP-Ranges anpassen und Subnetze verwalten. Wählen Sie „Peerings“ um in die Peering Konfigurationssicht zu wechseln.

Das Nabennetzwerk „**s4h-hub**“ stellt den zentralen Zugang zu den Speiche Netzwerken bereit. Zum jetzigen Zeitpunkt handelt es sich jedoch in beiden Fällen, um interne Netzen zwischen denen kein Routing möglich ist. Um Datenverkehr zu ermöglichen, erstellen wir nun ein Peering zwischen den beiden Netzen. Klicken Sie auf den Button „*Hinzufügen*“. Die Peering Ressource wird zweimal angelegt. Einmal als Knoten des Netzwerks „**s4h-vnet**“ und einmal als Knoten des Netzwerks „**hub-vnet**“. Sie müssen daher zwei Namensbezeichnungen festlegen. Wählen sie „**s4h-vnet-peering-hub**“ und „**hub-vnet-peering-s4h**“. Belassen sie die Default Einstellungen und wählen Sie als Zielnetzwerk „**hub-vnet**“ aus. Klicken Sie anschließend auf „*Hinzufügen*“.

## 7.2.5 Anlegen der Netzwerksicherheitsgruppen

Nun da wir die Netzwerkbereiche erstellt haben wird es Zeit die Portregeln festzulegen. Zu diesem Zweck werden wir jedem Subnetz eine eigene Netzwerksicherheitsgruppe (NSG) zuordnen. Navigieren Sie erneut zu ihrer Ressourcengruppe (via #allservices und anschließend „Ressourcengruppen“) und wählen Sie „**hub-rg**“ aus. Such Sie im Marketplace nach „Netzwerksicherheitsgruppe“ und erstellen Sie diese mit dem Namen „**hub-nsg-mgmt**“.

Wiederholen Sie die Schritte zunächst für die NSG „**s4h-nsg-as**“ in der Ressourcengruppe „**s4h-rg**“. Nachdem Sie diese erstellt haben, warten Sie bis die Bereitstellung abgeschlossen ist und sie sich in der Ansicht „Bereitstellung“ befinden (alternativ navigieren Sie über die Ressourcengruppe „**s4h-rg**“ und wählen Sie „Bereitstellungen“).

Um die verbleibende NSG „**s4h-nsg-db**“ hinzuzufügen werden wir nun über die Bereitstellungsvorlage navigieren. Klicken Sie auf den Button „*Erneut bereitstellen*“. Sie erhalten nun die bekannte Eingabemaske, jedoch sind diesmal die Felder mit den ursprünglichen Parameterwerten vorausgefüllt. Beachten Sie jedoch, dass für einige Felder keine Eingabehilfen (Dropdown) mehr zur Verfügung stehen, sondern lediglich ein Freitext angeboten wird. Vorlagenbereitstellungen haben so gut wie keine Eingabeverifikation und sollten aus diesem Grund immer sorgfältig editiert werden. Überprüfen Sie die vorausgefüllten Werte und legen Sie die Bezeichnung für die Datenbank NSG fest. Auch wenn die Aufwandseinsparung für das Anlegen einer NSG mittels Vorlage gering ist, kann dieses Prinzip grundsätzlich für jede Ressource durchgeführt werden.

Nun wird es Zeit, die NSGs den jeweiligen Subnetzen zuzuordnen, sowie die Portregeln zu konfigurieren. Navigieren Sie zunächst zur Ressource „**s4h-nsg-db**“ und wählen Sie in der linken Navigationsleiste den Menüpunkt „Subnetze“. Klicken Sie auf Zuordnen und wählen Sie zunächst das VNET „**s4h-vnet**“ und anschließend das Subnetz „**s4h-vnet-subnet-db**“.

Wählen Sie nun den Menüpunkt „*Eingangsregeln*“ in der linken Navigationsleiste. Sicherheitsregeln bestehen im Grunde aus einer 4-Tupel Information. Die **Priorität** spezifiziert die Reihenfolge in welcher die Regeln angewendet werden, wobei der niedrigste Wert zuerst ausgeführt wird. Der **Name** definiert lediglich eine Bezeichnung

für die Regel. Die Kombination aus **Port**, **Protokoll**, **Quelle** und **Zieladresse** muss gemeinsam erfüllt sein damit die **Aktion** der Regel (Allow oder Deny) ausgeführt wird. Klassischerweise werden NSGs nach dem Whitelisting-Prinzip konfiguriert, d. h. alle Arten von Verbindungen sind untersagt, außer diese sind explizit erlaubt. Die Standardeinstellungen nach Erstellung einer neuen NSG spiegeln dieses Prinzip wieder. Health-probes des Azure Load Balancer sowie Datenverkehr mit dem Service Tag „VirtualNetwork“ ist grundsätzlich auf allen Ports erlaubt. Andere Verbindungsanfragen wie z. B. Anfragen aus dem öffentlichen Netz sind untersagt. Beachten Sie, dass Sie die Standardregeln nicht entfernen können, jedoch durch Regeln mit niedrigerem Prioritätswert überschreiben werden können.

**Welche Regeln für welche NSG?** Wie frei Sie den Zugang zu ihrer Landschaft gestalten wollen ist grundsätzlich unternehmensabhängig. Im Rahmen einer Hub-Spoke Architektur empfiehlt es sich jedoch die Freischaltungen für jegliche SAP Workloadnetze (Spokes) restriktiv zu gestalten. Umfangreiche Zugänge sollten lediglich aus dem Management Subnet (Teil des Hubs) möglich sein. Folgende Überlegungen könnten hierbei von Relevanz sein:

**s4h-nsg-db:** Das Datenbanknetz sollte i. d. R. besonders restriktiv gestaltet sein, insbesondere falls die darin enthaltenen HANA Plattformen lediglich als Datenbank für S/4HANA Instanzen verwendet werden. Folgende Regeln könnten hierfür relevant sein (Tab. 7.2):

**s4h-nsg-as:** Die Applikationsserver sollten lediglich aus dem Unternehmensnetz via Sap GUI oder Fiori verfügbar sein (Tab. 7.3).

Zum Spezifizieren von Ziel oder Quelle unterstützen NSGs neben IP-Adressen auch sogenannte Diensttags (engl. Service Tags) sowie Anwendungssicherheitsgruppen. Bei den Diensttags handelt es sich um vordefinierte Adressbereiche welche für einen bestimmten Anwendungszweck gelten und von Microsoft gepflegt werden. Der Diensttag „Virtual Network“ deckt beispielsweise alle IPs aus dem verbundenen on-premise

**Tab. 7.2** NSG-Regeln Datenbanknetz

Priorität	Quell-IP	Quell-Port	Ziel-IP	Ziel-Ports	Protokoll	Aktion	Beschreibung
100	subnet-as	*	Any	3<nr>13, 3<nr>15, 3<nr>40, 3<nr>41, 3<nr>42	Any	Allow	SAP-HANA Ports
101	subnet-mgmt	*	Any	22,3389	Any	Allow	SSH, RDP
102	subnet-mgmt	*	Any	7630	Any	Allow	Hawk (Pacemaker)

**Tab. 7.3** NSG-Regeln Applikationsserversubnetz

Priorität	Quell-IP	Quell-Port	Ziel-IP	Ziel-Ports	Protokoll	Aktion	Beschreibung
100	Diensttag=Virtual Network	*	Any	32<nr>, 33<nr>, 80<nr>, 443<nr>	Any	Allow	SAP Dispatcher, SAP Gateway, ICM HTTP, ICM HTTPs
101	subnet-mgmt	*	Any	22,3389	Any	Allow	SSH, RDP
102	subnet-mgmt	*	Any	4237	Any	Allow	SWPM
103	subnet-mgmt	*	Any	7630	Any	Allow	Hawk (Pacemaker)

Adressbereich, sowie gepeerte virtuelle Netze und Netze welche an einem Netzwerk Gateway angeschlossen sind ab. Eine Liste der verfügbaren Diensttags findet sich in der Azure Dokumentation [2].

Anwendungssicherheitsgruppen [3] (ASGs) ermöglichen es einzelne Netzwerkschnittstellen zu einer Gruppe hinzuzufügen und NSG Regelungen auf dieser Basis festzulegen. Grundsätzlich kann eine vergleichbare Steuerung auch auf Basis von IP-Adressen festgelegt werden, ASGs ermöglichen jedoch eine größere Flexibilität, da sich diese im Gegensatz zur IP-Adresse i. d. R. nicht ändern.

- ▶ **Hinweis:** Für die meisten produktiven SAP-Systeme empfiehlt es sich, dass diese lediglich über das Intranet, nicht jedoch über das öffentliche Netzwerk verfügbar sind. Damit eine VM jedoch Zugang zu öffentlichen Endpunkten, u. A. benötigt für Repositories, den Azure Fencing Agent oder Azure Backup erhält, ist es nötig, dass diese entweder über eine öffentliche IP-Adresse verfügen oder ein Internetzugang mittels eines Lastenausgleichs für ausgehende Kommunikation oder einer Azure Firewall bereitgestellt wird. Lesen Sie hierzu die Hinweise [4] von Microsoft zur Realisierung eines solchen Szenario. Aus IT-Sicherheits-Perspektive sind die letzteren beiden Alternativen zu bevorzugen. Im Rahmen dieser Beispielarchitektur werden wir zur Vereinfachung mit öffentlichen IP-Adressen bei jeder VM arbeiten.

### 7.3 Bereitstellen eines HANA HA Clusters

Im Folgenden wird auf die Bereitstellung eines HANA HA Clusters näher eingegangen. Dabei wird zuerst die HANA Cluster Architektur umrissen und danach die Bereitstellung der notwendigen HANA Cluster Ressourcen thematisiert. Abschließend wird die Konfiguration des Pacemaker Cluster sowie die Einrichtung des HANA HA-Clusters näher beschrieben.

### 7.3.1 HANA Cluster Architektur

Nachdem die Grundkonfiguration abgeschlossen ist, wollen wir nun ein HANA Hochverfügbarkeitscluster bereitstellen. Zu diesem Zweck erweitern wir unsere Architektur um die Komponenten, welche in Abb. 7.4 illustriert sind:

**Virtuelle Computer und Verfügbarkeitsgruppen** In das Datenbank Subnet „s4h-vnet-subnet-db“ stellen wir zwei virtuelle Computer „ha1-vc-db0“ und „ha1-vc-db2“ bereit. Beide Instanzen sind Teil der Verfügbarkeitsgruppe „ha1-avset-db“. Zwischen

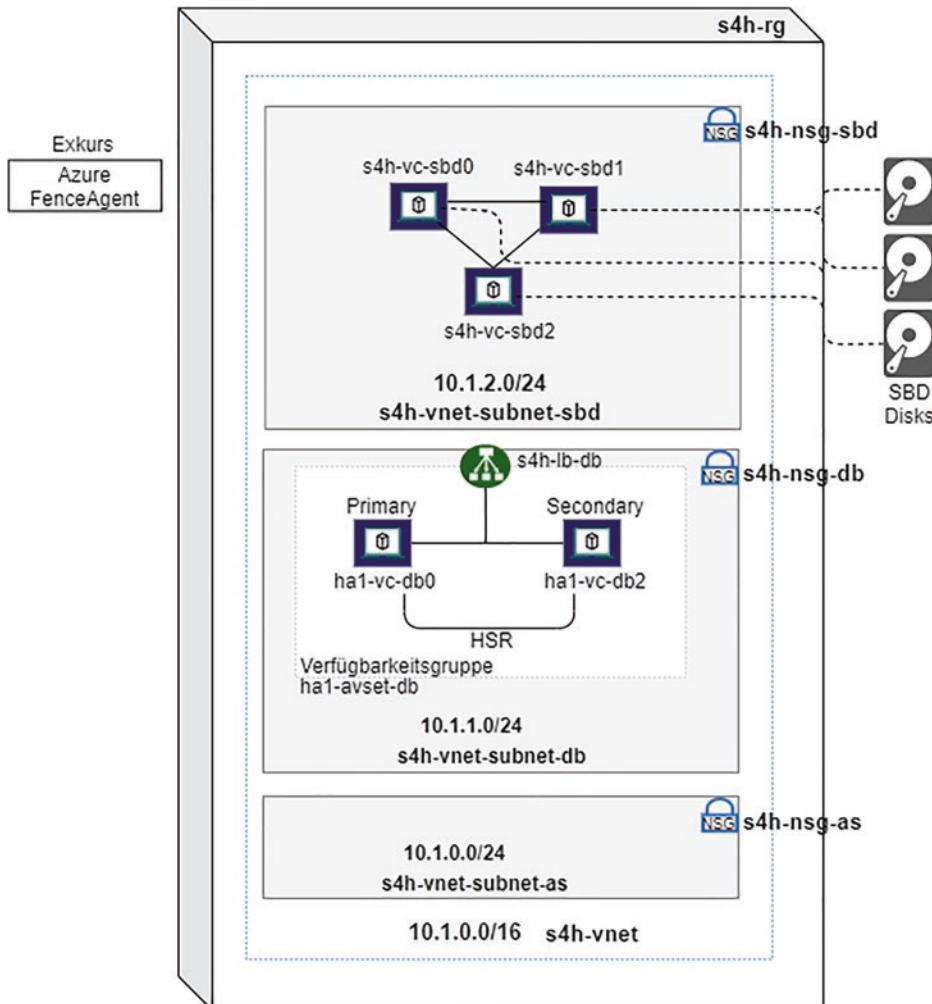


Abb. 7.4 HANA Cluster Architektur

den beiden HANA Systemen wird außerdem eine Systemreplikation eingerichtet, damit der secondary Knoten die Änderungen des primary Knotens ebenfalls erhält.

**Lastenausgleich** Damit beim Ausfall eines Knotens die Anwendungsserver davon nicht betroffen sind wird ein interner Lastenausgleich vorgeschaltet. Der Lastenausgleich bietet eine statische IP Adresse nach außen welche sich selbst bei Ausfall eines Knotens nicht ändert. Die Anwendungsserver kommunizieren mit der IP des Lastenausgleichs, während der Lastenausgleich die Anfragen intern an den entsprechenden primary HANA Knoten, auf Basis eines Integritätstests weiterleitet.

**iSCSI Target VMs** Um ein split-brain Szenario im Pacemaker Cluster zu vermeiden wird ein Fencing Mechanismus benötigt. Eine Möglichkeit dies umzusetzen ist mittels des Azure Fencing Agents welchen wir in Kapitel 7.3.3.4 näher beleuchten werden. Der klassische Weg ist die Nutzung von Stonith Block Devices (SBD). Zu diesem Zweck muss ein geteilter Datenträger auf den Knoten des Clusters verfügbar gemacht werden. Im Gegensatz zum on-premise Bereich wo ein solches Szenario meist auf Basis eines SANs konfiguriert werden kann, nutzen wir in Azure das iSCSI Protokoll. Hierbei können SCSI Kommandos basierend auf IP Netzwerken ausgetauscht um Datentransfer zwischen verteilten Systemen zu ermöglichen. iSCSI basiert auf einer Client- / Serverarchitektur bei der der Server (das sogenannte „target“) SBD Geräte bereitstellt und die Clients (die sogenannten „initiators“) mittels IP darauf zugreifen. Für die clients erscheinen die iSCS Geräte als wären diese lokal verbunden. Die iSCSI Target VM ist ein eigenständiger virtueller Computer und kann von mehreren Pacemaker Cluster geteilt verwendet werden. Aus diesem Grund stellen wir die iSCSI VMs in einem eigenständigen virtuellen Subnetz bereit. Für eine Grundkonfiguration ist eine iSCSI VM ausreichend. Dies erfordert jedoch, dass das SBD Gerät immer verfügbar ist. Für den Produktivbetrieb empfiehlt es sich daher mit drei iSCSI targets (drei SBD Geräte) zu arbeiten. In diesem Fall kann wartungsbedingt auch eine target VM heruntergefahren/gestoppt werden. Beachten Sie, dass Ihnen zwei SBD Geräte keine zusätzliche Absicherung ermöglicht, da in diesem Fall bei Ausfall eines SBD Geräts Pacemaker nicht in der Lage ist ein automatisches Fencing durchzuführen.

### 7.3.2 Bereitstellung der HANA Cluster Ressourcen

Um den manuellen Aufwand so gering wie möglich zu halten, nutzen wir für die Bereitstellung der HANA Datenbank eine sogenannte Schnellstartvorlage. Azure bietet zahlreiche ARM Vorlagen für verschiedenste Einsatzszenarien. Eine Übersicht aller Vorlagen findet sich auf der Microsoft Website [5]. Suchen Sie nach dem Stichwort „sap“ um alle SAP spezifischen Bereitstellungsvorlagen anzuzeigen. Wir werden anschließend die manuelle Bereitstellung eines virtuellen Computers auf Basis der iSCSI Target VM näher beleuchten.

### 7.3.2.1 Struktur einer Schnellvorlage

Im Folgenden werden wir zunächst die Struktur einer Schnellvorlage erläutern, bevor wir anschließend die Vorlage für unseren eigenen Bedarf anpassen. Auf diese Weise lernen Sie am besten den Ablauf und die Funktionsweise von ARM Vorlagen kennen.

Als Vorbereitung zur Bereitstellung benötigen wir zunächst die ID des von uns angelegten Datenbank Subnetzes, sowie unserer Netzwerksicherheitsgruppe. Navigieren Sie in die Ressourcengruppe „**s4h-rg**“, klicken Sie auf die Ressource „**s4h-vnet**“ und wählen sie im Bereich „Übersicht“ den Button „JSON-Ansicht“ am oberen rechten Bildschirmrand. Die JSON-Ansicht beschreibt vollständig die Konfiguration des Objektes inklusive aller ID-Referenzen auf andere Ressourcen. Suchen Sie nach dem Subnetznamen „**s4h-vnet-subnet-db**“ und notieren Sie sich die darunter stehende ID. Wiederholen Sie diesen Schritt für die Netzwerksicherheitsgruppe „**s4h-nsg-db**“. Anstelle der JSON-Ansicht können Sie auch über die Navigationsleiste die Ansicht „Eigenschaften“ öffnen und sich anschließend die Ressourcen-ID anzeigen lassen. Wechseln Sie nun zurück in die Ressourcengruppe „**s4h-rg**“ und erstellen Sie die Ressource „Vorlagenbereitstellung (mit benutzerdefinierten Vorlagen bereitstellen)“. Selektieren Sie als Vorlagenquelle „Schnellstartvorlage“ und wählen Sie „sap-3-tier-marketplace-image-multi-sid-db-md“.

Wir wollen nun ein tieferes Verständnis für die Funktionsweise von ARM Vorlagen aufbauen. Klicken Sie daher zunächst auf „*Vorlage bearbeiten*“.

Sie befinden Sich nun in einem Editiermodus. Azure Vorlagen basieren auf der JSON-Syntax und bestehen im Wesentlichen aus den drei Bereichen Parameter, Variablen und Ressourcen:

- **Parameter:** Hierbei handelt es sich um die Felder welche dem Anwender im Rahmen einer Bereitstellung als Formular zum Ausfüllen angezeigt werden. Die von uns gewählte Vorlage besitzt insgesamt zwölf Parameter. Klicken Sie auf den Parameter „osType“ um mehr über diesen zu erfahren. In der rechten JSON Ansicht sehen Sie, dass dieses Feld eine Enumeration ist welche lediglich einen von acht erlaubten Werten akzeptiert. Als Defaultwert ist „Windows Server“ voreingestellt. Notieren Sie sich die erlaubten Werte z. B. „SLES 12 BYOS“. Navigieren Sie nun zum Parameter „sapSystemSize“ und notieren Sie sich die dort erlaubten Werte (z. B. Demo, Small). Wie Sie anhand dieses Beispiels sehen, haben Parameter i. d. R. eine Benutzersicht und kapseln nicht die technischen Detailinformationen. Welche VM Instanz mit dem Wert „Small“ erstellt wird können wir auf dieser Basis noch nicht identifizieren.
- **Variablen:** Variablen beinhalten die tatsächlichen (technischen) Informationen welche zur Bereitstellung verwendet werden. Die vom Benutzer festgelegten Parameter verweisen meistens lediglich auf Hintergrundvariablen. Wählen Sie nun die Variable „images“ aus und prüfen Sie die Werte im JSON Baum. Hier sind erneut die Werte gelistet welche wir bereits für den Parameter „osType“ notiert haben. Prüfen Sie nun welche „sku“ z. B. für „SLES 12 BYOS“ definiert ist. In unserem Fall lautet diese

„12-SP3“, was eine veraltete SLES Version ist. Es steht Ihnen frei hier die ersten Anpassungen vorzunehmen. In unserem Fall ändern wir die SKU auf „12-SP5“. Achten Sie bei der Kombination „offer“ und „sku“ darauf, dass diese tatsächlich existiert und in ihrer Version verfügbar ist. Andernfalls erhalten Sie im Rahmen des Überprüfungs-schritts eine entsprechende Fehlermeldung. Wie Sie die Verfügbarkeit von Images für Ihre Region prüfen ist in folgendem Artikel beschrieben [6]. Wählen Sie anschließend die Variable „sizes“. Die Struktur ist hier etwas komplexer. Unter „Demo“ sollte zunächst der Wert „HANA“ auftauchen bevor die tatsächliche „vmSize“ als „standard\_e8s\_v3“ angegeben ist. Wenn Sie weiter herunter scrollen werden Sie den Wert „SQL“ finden. Das liegt daran, dass die Vorlage unterschiedliche VM Größen für HANA und Non-HANA Datenbanken verwendet. Die Größe „standard\_e8s\_v3“ für ca. 0,64 \$/Stunde verwendet 8 vCPUs sowie 64 GB RAM und eignet sich daher bestenfalls für eine blanke Netweaver Installation. Wenn Sie der Vorlage treu bleiben möchten und ein echtes S/4HANA System installieren wollen, müssen Sie daher zumindest die Größe „Small“ (Standard\_E32s\_v3 mit 32 vCPUs und 256 GB RAM) wählen welche jedoch das Vierfache kostet (ca 2,56 \$/Stunde). Alternativ können Sie die „vmSize“ auch gerne selbst anpassen. Aus unserer Sicht bedienen die folgenden VM Größe gut das preisliche Mittelfeld zwischen „Demo“ und „Small“, wobei die FX-Serie ein besonders gutes Preis-/Leistungsverhältnis auf Kosten weniger Kerne anbietet. Beachten Sie jedoch, dass Sie in diesem Fall auch „offer“ und „sku“ im Knoten „imageReference“ der Bereitstellungsvorlage anpassen müssen (Tab. 7.4):

- **Ressourcen:** Die Ressourcen definieren die Objekte welcher im Rahmen der Vorlage bereitgestellt werden sollen. In dieser Datei wird dann die Kombination aus Parametern und Variablen angewendet um die Ressourcen eindeutig zu spezifizieren. Wählen Sie zunächst den Ressourcenknoten der virtuellen Maschine aus.

Im Bereich des „storageProfiles“ wird das Zusammenspiel zwischen Parameter und Variablen ersichtlich (siehe Abb. 7.5). Die Ressource identifiziert die „sku“ indem der vom Benutzer eingegebene Parameter (referenziert über das Feld „osType“) und die Variable „images“ als Key übergeben wird und der damit assoziierte Wert zurückgegeben wird. Haben Sie bereits Anpassungen der VM Größe vorgenommen

**Tab. 7.4** Geeignete VM-Größen

VM-Größe	vCPUs	RAM	Kosten (West Europe)	Verfügbare Image Offers (SUSE)	Verfügbare SKUs
Standard_E16s_v3	16	128	1,28 \$/Stunde	sles-sap-12-sp5-byos	gen1/gen2
Standard_E20s_v3	20	160	1,6 \$/Stunde	sles-sap-12-sp5-byos	gen1/gen2
Standard_FX4mds	4	80	0,45 \$/Stunde	sles-sap-12-sp5-byos	gen2
standard_FX12mds	12	256	1,35 \$/Stunde	sles-sap-12-sp5-byos	gen2



**Abb. 7.5** Zusammenspiel von Parametern und Variablen

müssen Sie sicherstellen, dass „sku“ und „offer“ für die von Ihnen gewählte VM-Größe verfügbar ist.

- **Hinweis:** Die Semantik von **sku** und **offer** wird nicht von allen Anbietern in Azure immer konsistent angewendet. Für das Offer „SLES-SAP“ bietet SUSE z. B. die Skus „12-sp4“ und „12-sp4-gen2“. Gen2 steht hierbei für die Hypervisor Version. Manche VM Größen wie z. B. die FX-Serie werden beispielsweise nur als Gen2 Variante angeboten. In diesem Fall könnte die Sku „12-sp4“ nicht verwendet werden. SUSE bietet jedoch auch das Offer „sles-12-sp5“ mit den Skus „gen1“ und „gen2“ an. Beachten Sie daher beim Erstellen einer Vorlage wie die genaue Kombination von **sku** und **offer** lauten muss um eine gültige Bereitstellung zu erhalten.

### **7.3.2.2 Anpassen der DB Schnellstartvorlage**

Die Vorlage entspricht derzeit noch nicht ganz den Anforderungen unserer Beispielarchitektur. Zunächst wollen wir die Namen der Ressourcen soweit anpassen, dass diese unserem Namensschema entsprechen.

- Hinweis:** Das Editieren einer Vorlage kann leicht zu Problemen bei der Bereitstellung führen. Neben semantischen Fehlern (ungültiges JSON File) kann es insbesondere zu Inkonsistenzen bei Versionsabhängigkeiten oder Komponenten kommen. Wir empfehlen Ihnen, dass Sie dennoch die Schritte dieses Kapitels befolgen, um ein Verständnis für Vorlagenbereitstellungen aufzubauen. Sollte es später zu Problemen kommen können Sie jederzeit die ursprüngliche Vorlage bereitstellen oder die notwendigen Schritte manuell durchführen.

Wählen Sie den Ressourcenknoten der virtuellen Maschine aus und entfernen Sie das Trennzeichen zwischen Variablename und `copyIndex()`. Ändern Sie dazu den Wert „`name`“ auf `"[concat(variables('vmNameDB'), copyIndex())]"`, wie in Abb. 7.6 demonstriert.

```

"copy": {
  "name": "dbVMLoop",
  "count": "[variables('dbvmCount')]"
}

```

```

"copy": {
  "name": "dbVMLoop",
  "count": "[variables('dbvmCount')]"
}

```

**Abb. 7.6** Modifikation des Ressourcenknotens der virtuellen Maschine

Wiederholen Sie diesen Schritt für die Netzwerkschnittstelle, sowie für die öffentliche IP Adresse. Öffnen Sie anschließend die Variable „vmNameDB“ und ändern Sie dessen Wert auf.

```
"vmNameDB": "[concat(variables('sidlower'), '-vc-db')]"
```

Die Namensgenerierung für die Verfügbarkeitsgruppe, sowie den Lastenausgleich muss nicht angepasst werden. Die von der Vorlage angelegten Ressourcen beinhalten derzeit acht Objekte, darunter ein virtuelles Netzwerk und eine NSG. Beide Ressourcen wurden bereits im Rahmen unserer Grundkonfiguration angelegt und sollen daher nicht erneut erstellt werden. Daher passen wir im nächsten Schritt die Vorlage so an, dass die Bereitstellung in unser existierendes Subnetz erfolgt und bestehende NSGs wiederverwendet.

Wechseln Sie auf den Knoten „Ressourcen (8)“ und löschen Sie in der JSON Ansicht die Objekte „networkSecurityGroups“ sowie „virtualNetworks“. Achten Sie beim Löschen auf die korrekte Klammerung des JSON Files wie in Abb. 7.7 dargestellt:

Der Ressourcenknoten in der linken Navigationsleiste sollte sich daraufhin auf (6) aktualisieren. Für die Angabe des Subnetzes existiert bereits der Parameter „subnetId“. Für die Angabe der NSG definieren wir einen neuen Parameter „nsgId“. Wählen Sie in der Navigationsleiste den Knoten „subnetId“ aus und kopieren Sie dessen Parameterdefinition. Ändern Sie anschließend die ID auf „nsgId“ ab sowie die Beschreibung auf „The id of the nsg you want to use.“. Abb. 7.8 veranschaulicht das Ergebnis dieses Schrittes.

Wir müssen außerdem den Parameter „\_artifactsLocation“ anpassen. Hierbei handelt es sich um eine URI welche es ermöglicht weitere externe Dateien in die Bereitstellung miteinzubeziehen. Im Kontext dieser Vorlage werden z. B. Onlineskripte eingebunden welche die erstellten Datenträger der VM automatisch partitionieren und einbinden. Die Skripte stehen im zugehörigen Git Repository der Vorlage bereit. Wir tragen als „defaultValue“ die URL zum Repository ein. Achten Sie darauf die „raw“ und nicht die HTML Darstellung der Ressource referenzieren. Die URL muss außerdem mit „/“ enden. Tragen Sie die URL „<https://raw.githubusercontent.com/Azure/azure-quickstart-templates/master/application-workloads/sap/sap-3-tier-marketplace-image-multi-sid-db-md/>“ und vergleichen Sie ihr Ergebnis mit Abb. 7.9.

Aufgrund möglicher Abhängigkeiten zwischen einzelnen Ressourcenobjekten, z. B. benötigt eine Netzwerkschnittstelle eine Öffentliche IP, können diese auch im Rahmen der Vorlagenbereitstellung definiert werden. Dadurch wird sichergestellt, dass die Bereitstellung für die nachfolgende Ressource abgebrochen wird, wenn dessen Voraussetzung

Home > Ressourcengruppen > s4h-rg > Ressource erstellen > Marketplace > Vorlagenbereitstellung (mit benutzerdefinierten Vorlagen bereitstellen) > Benutzerdefinierte Berei

### Vorlage bearbeiten ...

Azure Resource Manager-Vorlage bearbeiten

```

    "resources": [
      {
        "type": "Microsoft.Network/networkSecurityGroups",
        "name": "[concat(variables('nsName'))]",
        "apiVersion": "2020-07-01",
        "location": "[parameters('location')]",
        "condition": "[equals(length(parameters('subnetId')), 0)]",
        "properties": {
          "securityRules": "[variables('selectedSecurityRules')]"
        }
      },
      {
        "type": "Microsoft.Network/virtualNetworks",
        "name": "[variables('vnetName')]",
        "apiVersion": "2020-07-01",
        "location": "[parameters('location')]",
        "dependsOn": [
          "[resourceId('Microsoft.Network/networkSecurityGroups', variables('nsName'))]"
        ],
        "condition": "[equals(length(parameters('subnetId')), 0)]",
        "properties": {
          "addressSpace": {
            "addressPrefixes": [
              "10.0.0.0/16"
            ]
          },
          "subnets": [
            {
              "name": "[variables('subnetName')]",
              "properties": {
                "addressPrefix": "10.0.0.0/24",
                "networkSecurityGroup": {
                  "id": "[resourceId('Microsoft.Network/networkSecurityGroups', variables('nsName'))]"
                }
              }
            }
          ]
        }
      }
    ],
  ],
  "variables": [
    ...
  ],
  "outputs": [
    ...
  ]
}
  
```

**Abb. 7.7** Modifikation des Ressourcenknotens

```

    "description": "Password or ssh key for the Virtual Machine."
  },
  "subnetId": [
    {
      "type": "string",
      "defaultValue": "",
      "metadata": {
        "description": "The id of the subnet you want to use."
      }
    },
    {
      "type": "string",
      "defaultValue": "",
      "metadata": {
        "description": "The id of the nsg you want to use."
      }
    }
  ],
  "location": [
    ...
  ]
}
  
```

**Abb. 7.8** Modifikation des Subnetzes des Ressourcenknotens

```

110     },
111     "_artifactsLocation": {
112       "type": "string",
113       "metadata": {
114         "description": "The base URI where artifacts required by this template are located. When the template is deployed using the accompanying scripts, a private location in the subscription will be used and this value will be automatically generated."
115       },
116       "defaultValue": "https://raw.githubusercontent.com/Azure/azure-quickstart-templates/master/application-workloads/sap/"
117     },
118     "_artifactsLocationSasToken": {

```

Change

**Abb. 7.9** Modifikation des Parameters „\_artifactsLocation“

nicht erfüllt ist bzw. nicht bereitgestellt werden konnte. Da wir das virtuelle Netzwerk sowie die NSG in der Vorlage gelöscht haben, müssen wir die entsprechenden Abhängigkeitsdefinitionen in der Ressourcendefinition ebenfalls anpassen. Löschen Sie dazu die „dependsOn“ Einträge für die Ressourcentypen „publicIPAddresses“, und „loadBalancers“. Beim Ressourcentyp „networkInterfaces“ ist es ausreichend nur das Item für das virtuelle Netzwerk aus dem Array zu entfernen. Abb. 7.10 stellt diese Änderung grafisch dar.

Für die Ressource „networkInterface“ definieren wir außerdem im Bereich „properties“ eine Referenz auf den neuen Parameter „nsgId“ um die Netzwerksicherheitsgruppe zu definieren. Da wir die Namensbezeichnung für öffentliche IP-Adressen abgeändert haben müssen wir zudem den Bindestrich in „publicIPAddress“ entfernen. Abb. 7.11 fasst diese Änderungen zusammen.

Nutzen Sie abschließend die Browsersuche um die verbleibenden vier Vorkommnisse von ‘-‘, zu entfernen. Für die Ressource „virtualMachines“ muss der Bindestrich dreimal entfernt werden in den Bereichen „osProfile“, „networkInterfaces“ und „osDisk“. Außerdem muss der Namen für die Ressource „virtualMaschines/extensions“ angepasst werden.

Damit sind unsere Änderungen an der Vorlage abgeschlossen. Klicken Sie auf den Button „Herunterladen“ um die von Ihnen angepasste Version der Vorlage lokal auf Ihrem PC zu speichern. Über „Datei laden“ haben Sie später jederzeit die Möglichkeit die Vorlage wiederherzustellen. Klicken Sie abschließend auf den Button „Speichern“ um die Änderungen zu übernehmen und in die Bereitstellungsmaske zu wechseln.

### 7.3.2.3 Benutzerdefinierte Vorlagenbereitstellung

In der Ansicht zur „Benutzerdefinierten Bereitstellung“ erhalten Sie wie gewohnt die Auswahl des Abonnements, sowie der Ressourcengruppe. Stellen Sie zunächst sicher, dass die hier eingetragenen Werte korrekt sind. Die unter Instanzdetails gelisteten Parameter entsprechen den Werten welche wir bereits im Rahmen der Voralgenbearbeitung gesehen haben. Unter dem Eintrag „Subnet Id“ sollte Außerdem der von uns neu definierte Parameter „Nsg Id“ inklusive Tooltip gelistet sein. Befüllen Sie lediglich den Parameter „SAP-System Id“ mit dem Wert „HA1“ und klicken Sie anschließend auf den Button „Visualisieren“ um sich einen Überblick über die zu erstellenden Ressourcen zu verschaffen. In der Ressourcenschnellansicht sollte nun ein Graph mit fünf Ressourcenknoten angezeigt werden, die Pfeile stehen dabei für die „dependsOn“ Abhängigkeiten.

```
{
  "type": "Microsoft.Network/publicIPAddresses",
  "name": "[concat(variables('publicIpNameDB'), '-', copyIndex())]",
  "apiVersion": "2020-07-01",
  "condition": "[equals(length(parameters('subnetId')), 0)]",
  "dependsOn": [
    "[resourceId('Microsoft.Network/virtualNetworks/', variables('vnetName'))]"
  ],
  "location": "[parameters('location')]",
  "copy": {
    "name": "dbpipLoop",
    "count": "[variables('dbvmCount')]"
  },
  "properties": {
    "publicIPAllocationMethod": "Dynamic"
  }
},
{
  "type": "Microsoft.Network/loadBalancers",
  "name": "[variables('loadBalancerNameDB')]",
  "apiVersion": "2020-07-01",
  "location": "[parameters('location')]",
  "condition": "[greater(variables('dbvmCount') , 1)]",
  "dependsOn": [
    "[resourceId('Microsoft.Network/virtualNetworks/', variables('vnetName'))]"
  ],
  "properties": {
    "frontendIPConfigurations": "[variables('lbFrontendConfigs')[parameters('dbtype')][variables('internalOSType')]]",
    "backendAddressPools": "[variables('lbBackendPools')[parameters('dbtype')][variables('internalOSType')]]",
    "loadBalancingRules": "[variables('lbRules')[parameters('dbtype')][variables('internalOSType')]]",
    "probes": "[variables('lbProbes')[parameters('dbtype')][variables('internalOSType')]]"
  }
},
{
  "type": "Microsoft.Network/networkInterfaces",
  "name": "[concat(variables('nicNameDB'), '-', copyIndex())]",
  "apiVersion": "2020-07-01",
  "copy": {
    "name": "dbNICLoop",
    "count": "[variables('dbvmCount')]"
  },
  "dependsOn": [
    "dbpipLoop",
    "[resourceId('Microsoft.Network/virtualNetworks/', variables('vnetName'))]",
    "[resourceId('Microsoft.Network/loadBalancers/', variables('loadBalancerNameDB'))]"
  ]
}
```

**Abb. 7.10** Modifikation der Netzwerksicherheitsgruppe

```

{
  "type": "Microsoft.Network/networkInterfaces",
  "name": "[concat(variables('nicNameDB'), copyIndex())]",
  "apiVersion": "2020-07-01",
  "copy": {
    "name": "dbNICLoop",
    "count": "[variables('dbvmCount')]"
  },
  "dependsOn": [
    "dbPIPLoop",
    "[resourceId('Microsoft.Network/loadBalancers/', variables('loadBalancerNameDB'))]"
  ],
  "location": "[parameters('location')]",
  "properties": {
    "ipConfigurations": [
      {
        "name": "ipconfig1",
        "properties": {
          "privateIPAllocationMethod": "Dynamic",
          "publicIPAddress": "[if(equals(length(parameters('subnetId')), 0), json(concat('{"id": \'", resourceId('Microsoft.Network/publicIPAddresses', concat(variables('publicIpNameDB'), copyIndex()), '\")), json('null')))]",
          "subnet": {
            "id": "[variables('selectedSubnetId')]"
          },
          "loadBalancerBackendAddressPools": "[if(greater(variables('dbvmCount'), 1), variables('nicBackAddressPools')[parameters('dbType')][variables('internalOSType')], json('null'))]"
        }
      }
    ],
    "networkSecurityGroup": {
      "id": "[parameters('nsgId')]"
    },
    "enableAcceleratedNetworking": "[variables('sizes')[parameters('sapSystemSize')][parameters('dbType')].useFastNetwork]"
  }
}

```

New

Delete

**Abb. 7.11** Modifikation der Ressource „networkInterface“ und „publicIPAdress“

Dass es sich bei der sechsten „Ressource“ um ein „VM extensions Objekt“ handelt welches nicht eigenständig angelegt wird, sondern lediglich Nachbearbeitungsschritte definiert wird dieses auch nicht im Rahmen der Visualisierung angezeigt. Achten Sie auch auf die Ressourcenbezeichnung um sicherzustellen, dass die Namenskonventionen eingehalten wurden.

Schließen Sie die Schnellansicht wieder und selektieren Sie in der Eingabemaske für den Parameter „System Availability“ den Wert „HA“. Klicken Sie nun erneut auf den Button „Visualisieren“. Es sollten nun drei weitere Ressourcen bestehend aus einer weiteren virtuellen Maschine, einer Netzwerkschnittstelle und einer öffentlichen IP-Adresse angezeigt werden. Schließen Sie nun die Ansicht und legen Sie die folgenden Eingabeparameter fest:

- Region: <Ihre Region>
- SID: HA1
- Os Type: <HANA unterstützte Linux Version wie z. B. SLES12 (BYOS = Bring your own Subscription)>
- Dbtype: HANA
- Sap System Size:
- System Availability: HA
- Admin Username: <Name für administratoruser z. B. cloudadm>
- Pw oder Key

- Subnet Id: <Notierte ID Ihres virtuellen Subnetzes>
- Nsg Id: <Notierte ID Ihrer NSG>

Für die verbleibenden Felder können die Defaultwerte übernommen werden. Klicken Sie nun auf „Überprüfen+erstellen“ und warten Sie bis die Bereitstellung abgeschlossen ist.

- **Hinweis:** Sollte es zu Fehlern in der Bereitstellung kommen klicken Sie auf „Details“ um nähere Informationen zum Problem zu erhalten. Falls die Fehler auf Probleme in der Vorlage zurückzuführen sind (z. B. „invalid JSON“), klicken Sie auf „erneut Bereitstellen“ und laden Sie ihre Vorlage erneut hoch. Machen Sie anschließend die entsprechenden Anpassungen. Denken Sie daran bereits erstellte Ressourcen zu löschen bevor Sie eine neue Bereitstellung versuchen. Alternativ können Sie auch die ursprüngliche Vorlage bereitstellen (in diesem Fall müssen Sie ggf. die Grundkonfiguration aus 7.1 anpassen oder löschen) oder die Ressourcen manuell bereitstellen. Die manuelle Bereitstellung einer VM werden wir im nächsten Schritt im Rahmen der iSCSI target VM erläutern.

Als Ergebnis sollten sich in Ihrer Ressourcengruppe nun 24 neue Objekte bestehend aus 18 × Datenträgern (2 × OS-Datenträger + 16 HANA-Datenträger), 2 × virtuellen Computern, 2 × Netzwerk Schnittstellen, sowie einer Verfügbarkeitsgruppe und einem Lastenausgleich befinden. Die meisten Objekte wie der Lastenausgleich sind bereits vorkonfiguriert und müssen nicht weiter angepasst werden insofern die HANA Installation mit Instanz Nummer 03 vorgenommen wird, passen Sie andernfalls die Ports der Lastenausgleichsregeln an.

**IP-Konfiguration anpassen** Prüfen Sie zunächst die IP-Einstellungen ihrer Ressourcen (virtuelle Computer + Lastenausgleich). Alle IP-Adressen müssen auf statisch geändert werden. Wir haben die dynamische Setzung in der Vorlage belassen damit automatisch gültige Werte aus dem Pool der nicht verwendeten IP Adressen gewählt werden. Für beide virtuelle Computer wollen wir außerdem eine öffentliche IP-Adresse definiert um Zugriff auf die Repositories zu haben. Wählen Sie dazu die Netzwerkschnittstelle ihrer VM in der Ressourcenansicht aus und klicken Sie auf „IP-Konfigurationen“ in der linken Navigationsleiste. Wählen Sie nun die IP-Konfiguration aus und ordnen Sie eine öffentliche IP-Adresse zu indem Sie auf „Neu erstellen“ klicken. Wählen Sie als Namen „**s4h-pip-db0**“, als SKU „Basic“ und als Zuordnung „Statisch“ aus. Ändern Sie anschließend die Zuordnung der privaten IP auf „Statisch“ und speichern Sie. Falls Sie auf die VM mittels SSH aus dem öffentlichen Netz zugreifen wollen müssen Sie zuerst die Netzwerksicherheitsgruppe für SSH Zugriffe aus dem Internet freischalten. Beachten Sie bitte, dass dies ein erhebliches Sicherheitsrisiko darstellt und keinesfalls als dauerhafte Lösung angesehen werden sollte. Verwenden Sie stattdessen lieber einen Jumpserver oder einen Bastion-Host im Rahmen ihrer „**hub-vnet**“.

**Skript Nachbearbeitung prüfen** Um sicherzustellen, dass die Postprocessing Skripte korrekt ausgeführt wurden, verbinden Sie sich nun auf ihren virtuellen Computer. Prüfen Sie die Speicherplatzverfügbarkeit mit „df -h“. Wie Sie sehen wurden die 8 × HANA-Datenträger bereits unter/hana/data, /hana/log und/hana/backup eingegebunden. Zu diesem Zweck wurden drei LVM volume-groups bestehend aus jeweils zwei Datenträger konfiguriert. Die Skriptbasierten Nacharbeiten wurden also korrekt durchgeführt.

### 7.3.2.4 Manuelle Bereitstellung der iSCSI Zielserver

Bevor wir mit der Erstellung der virtuellen Maschine beginnen benötigen wir ein neues Subnetz in welche die iSCSI Target VM(s) bereitgestellt werden. Wählen Sie dazu das virtuelle Netzwerk „s4h-vnet“ aus und fügen Sie unter „Subnetze“ ein neues hinzu. Als Bezeichnung wählen wir „s4h-vnet-subnet-sbd“ und als Adressbereich 10.1.2.0/24. Erstellen Sie ggf. eine neue Netzwerksicherheitsgruppe „s4h-nsg-sbd“ mit Portfreigaben für 860 und 3260 (iSCSI).

- **Hinweis:** Die Schritte zu Erstellung eines iSCSI target Servers basieren auf der Empfehlung von Microsoft und werden im Folgenden gekürzt dargestellt. Für eine ausführliche Anleitung siehe (Zugriff am 20.12.2021): <https://docs.microsoft.com/de-de/azure/virtual-machines/workloads/sap/high-availability-guide-suse-pacemaker>

#### 7.3.2.4.1 Erstellen der virtuellen Maschine

Navigieren Sie zu ihrer Ressourcengruppe „s4h-rg“ und erstellen Sie einen neuen virtuellen Computer (#create/Microsoft.VirtualMachine-ARM). Treffen Sie in den **Grundeinstellungen** folgende Angaben:

- Name: s4h-vc-sbd0
- Region: <Ihre Region>
- Verfügbarkeitsoption: Verfügbarkeitsgruppe
- Verfügbarkeitsgruppe: Neu erstellen -> „s4h-avset-sbd“, Fehlerdomäne = 2, Update-domäne = 20
- Image: SLES Linux for SAP (BYOS)
- Keine Spot-Instanz
- Größe: kleine VM ausreichend für unsere Beispielarchitektur z. B. „Standard\_DS1\_v2“

Navigieren Sie weiter zum Reiter **Datenträger** und wählen Sie SSD Premium mit Standard Verschlüsselung als Betriebssystemdatenträger.

Treffen Sie unter Netzwerk die folgenden Angaben und schließen die Bereitstellung ab:

- Virtuelles Netzwerk: s4h-vnet
- Subnetz: s4h-vnet-subnet-sbd
- Öffentliche IP: Neu erstellen (basic, statisch)
- Netzwerksicherheitsgruppe: s4h-nsg-sbd

### 7.3.2.4.2 iSCSI Zielserver Konfiguration

Wir werden nun die Konfigurationsschritte zum Erstellen eines iSCSI basierten SBD Geräts am Beispiel der HANA Datenbank HA1 vorstellen. Der iSCSI Ziel Server kann jedoch grundsätzliche SBD Geräte für weitere Pacemaker Cluster wie z. B. ASCS/ERS und NFS bereitstellen. Wenn Sie dies umsetzen möchten, wiederholen Sie die entsprechenden Schritte. Ersetzen Sie in diesem Fall alle Vorkommnisse von „**ha1**“ mit der entsprechenden Bezeichnung (z. B. **nfs** oder **s4h**). Verbinden Sie sich zunächst via SSH auf Ihren neuen Zielserver, lizenziieren Sie Ihr Betriebssystem im Falle von BYOS und installieren Sie den iSCSI Target Dienst.

```
# Einmalig auszuführen
sudo SUSEConnect --url=https://scc.suse.com -e <user> -r <key>
sudo zypper update
sudo zypper install targetcli-fb dbus-1-python
sudo systemctl enable targetcli
sudo systemctl start targetcli
```

Legen Sie nun das Verzeichnis/sbd an und erstellen Sie darin die sogenannten „**Backstores**“. Hierbei handelt es sich um den lokalen Speicherbereich auf dem Zielserver welcher für die iSCSI Exporte verwendet werden soll. Die Backstore Objekte sollen Dateibasiert verwendet werden (FILEIO). Für SBD muss außerdem der Filesystem Cache deaktiviert werden indem „write\_back=false“ gesetzt wird. Anschließend definieren wir für jedes Pacemaker Cluster, welches wir mit einem SBD Device bedienen wollen, einen sogenannten „**iSCSI Qualified Name (IQN)**“ über welchen die SBD Geräte im Netzwerk eindeutig identifiziert werden können.

```
sudo mkdir /sbd
# Wiederholen Sie die folgenden Befehle für zusätzliche Cluster
sudo targetcli backstores/fileio create sbdha1 /sbd/sbdha1 50M write_
back=false
sudo targetcli iscsi/ create iqn.2006-04.ha1.local:ha1
```

Durch anlegen der IQN wird automatisch eine sogenannte „**Target Portal Group (TPG)**“ mit Bezeichnung **tpg1** erstellt. Das Portal nutzt standardmäßig den Port 3260 und horcht auf alle dort eingehenden IPv4 Anfragen. Weitere TPGs können zusätzlich erstellt werden, dies ist jedoch in den meisten Fällen nicht nötig. Die backstores wollen wir als SCSI Logical Units (LUNS) exportieren. Zu diesem Zweck verlinken wir im ersten Schritt die backstore Objekte mit den jeweiligen IQNs. Für jeden initiator/client muss anschließend noch der Zugriff erlaubt werden indem wir die ACL entsprechend konfigurieren. Hierfür müssen wir jeden Host einzeln hinzufügen. Wir machen dies am Beispiel der HANA Hosts „**ha1-vc-db0**“ und „**ha1-vc-db1**“.

```
# Wiederholen Sie den folgenden Befehl für zusätzliche Cluster
sudo targetcli iscsi/iqn.2006-04.hal.local:hal/tpg1/luns/ create /
backstores/fileio/sbdhal

# Tauschen Sie hal-vc-db0 durch die entsprechenden Hostnamen aus
# Beispiel: s4h-vc-ascs0 oder s4h-vc-nfs0
sudo targetcli iscsi/iqn.2006-04.hal.local:hal/tpg1/acls/ create
iqn.2006-04.hal-vc-db0.local:hal-vc-db0

sudo targetcli iscsi/iqn.2006-04.hal.local:hal/tpg1/acls/ create
iqn.2006-04.hal-vc-db1.local:hal-vc-db1
```

Speichern Sie abschließend Ihre Änderungen. Damit ist der iSCSI Zielserver eingerichtet.

```
sudo targetcli saveconfig
```

### 7.3.3 Konfiguration des Pacemaker Clusters

In diesem Kapitel behandeln wir die notwendigen Schritte um ein grundlegendes Pacemaker Cluster basierend auf SBD und SLES 12 SP5 für die künftigen HANA VMs „hal-vc-db0“ und „hal-vc-db1“ aufzusetzen. Sollten Sie eine SLES Version <= 12 SP4 [7] verwenden oder RedHat Linux [8] empfehlen wir Ihnen die Anweisungen gemäß der offiziellen Microsoft Dokumentation zu befolgen.

Manche der folgenden Schritte müssen auf allen Maschinen [A], bzw. nur auf „hal-vc-db0“ [db0] oder „hal-vc-db1“ [db1]. Wir verwenden die Bezeichnungen [A], [db0] und [db1] um zu kennzeichnen auch welchem Host die darauffolgenden Schritte durchgeführt werden sollen.

#### 7.3.3.1 Einbinden der iSCSI Disk

[A] Zunächst müssen wir das iSCSI basierte SBD Gerät unseres Zielservers lokal verfügbar machen um Fencing zu ermöglichen. Verbinden Sie sich via SSH auf beide Maschinen und machen Sie die Dienste „iscsi“, „iscsid“ und „sbd“ verfügbar. Konfigurieren Sie außerdem den Initiator Name in „/etc/iscsi/“.

```
# ssh <hal-vc-db0 / hal-vc-db1>
sudo systemctl enable iscsid
sudo systemctl enable iscsi
sudo systemctl enable sbd
sudo vi /etc/iscsi/initiatorname.iscsi
# [db0] Ändern Sie den Wert auf
    InitiatorName=iqn.2006-04.hal-vc-db0.local:hal-vc-db0
# [db1] Ändern Sie den Wert auf
    InitiatorName=iqn.2006-04.hal-vc-db1.local:hal-vc-db1
```

[A] Nach einem Neustart des iSCSI Dienstes können die SBD Geräte eingebunden werden. Hierzu muss zunächst eine discovery mittels sendtarget (st) durchgeführt werden. Hierbei wird eine Liste der verfügbaren Targets an den initiator gesendet. Als Target kann die interne IP des ISCSI-Servers verwendet werden z. B. 10.1.2.0. Der Standardport des TPG ist 3260.

```
sudo systemctl restart iscsid
sudo systemctl restart iscsi
sudo iscsadm -m discovery --type=st --portal=<IP of s4h-vc-sbd0>:3260
```

[A] Anschließend führen wir einen Login beim Portal (TPG) unter Angabe des IQN durch. Damit wir nach einem Neustart nicht manuell den Login erneut durchführen müssen, konfigurieren wir außerdem einen automatischen Login. Prüfen Sie abschließend, ob das neue SCSI Gerät vom Initiator eingebunden wurde!

```
sudo iscsadm -m node -T iqn.2006-04.ha1.local:ha1 --login
--portal=<IP of s4h-vc-sbd0>:3260
sudo iscsadm -m node -p=<IP of s4h-vc-sbd0>:3260 -T iqn.2006-04.ha1.local:ha1 --op=update --name=node.startup --value=automatic
# Prüfen ob die Disk "sbdha1" vorhanden ist
lsscsi
```

### 7.3.3.2 Verwenden der iSCSI Disk als SBD Gerät

[A] Wir wollen nun das SCSI Gerät als SBD verwenden. Hierfür müssen wir zunächst die ID der Disk identifizieren. Nutzen Sie „lsscsi“ um zu prüfen als welches logische Gerät die iSCSI-Disk bereitgestellt wurde. Wenn Sie die Vorlagenbereitstellung verwendet haben ist das i. d. R. /dev/**sdk**. Um die ID zu identifizieren, führen Sie anschließend den folgenden Befehl aus:

```
ls -l/dev/disk/by-id/scsi-* | grep sdk
```

[A] Die Disk ist über drei Pfade verfügbar. Azure empfiehlt, dass Sie den Pfad kopieren welcher als ID mit dem Präfix „/dev/disk/by-id/scsi-3“ startet. Definieren Sie anschließend das SBD Device unter Angabe des Pfades und mit einem „msgwait“ timeout von 120 s.

```
sudo sdb -d <device-ID> 60 -4 120 create
```

Nun da das SBD Gerät erstellt wurde müssen Sie dieses in die Pacemaker Konfiguration mitaufnehmen. Sie können bis zu drei SBD Geräte als Liste getrennt mit Semikolon angeben.

```
sudo vi/etc/sysconfig/sbd
# Setzen Sie die folgenden Werte:
SBD_DEVICE="<device-ID1>;<device-ID2>;<device-ID3>"
SBD_PACEMAKER="yes"
SBD_STARTMODE="always"
```

Als watchdog wird in diesem Setup lediglich der „softdog“ unterstützt. Aktivieren Sie das Modul auf beiden Datenbank VMs.

```
echo softdog | sudo tee /etc/modules-load.d/softdog.conf
sudo modprobe -v softdog
```

### 7.3.3.3 Initialisieren des Pacemaker Clusters

Damit der Azure Lastenausgleich später identifizieren kann, welche HANA VM als primary agiert und damit alle Datenbankanfragen empfangen soll, setzt SUSE auf das Tool „socat“. Hierbei handelt es sich um eine weiterentwickeltes netcat Programm, welches lediglich auf dem primary Knoten gestartet wird. Der Lastenausgleich verwendet einen Integritätstest, welcher prüft ob auf den VMs im Backendpool ein definierter Port (socat) erreichbar ist. Indem wir sicherstellen, dass socat nur auf dem primary Knoten bzw. bei einem Failover auf dem ehemaligen secondary Knoten läuft, wird sichergestellt, dass der Lastenausgleich Anfragen immer an die richtige HANA weiterleitet. Abhängig von der Version ihres SUSE Systems kann auch azure-lb anstelle von socat verwendet werden. Neben socat benötigen wir außerdem noch den sogenannten Resource Agent. Hierbei handelt es sich um Skriptsammlungen, welche das Starten und Stoppen der Instanzen ermöglichen sowie Rückgabewerte definieren auf welche Pacemaker reagieren kann.

```
sudo zypper install -y socat
sudo zypper in -y resource-agents
```

Microsoft empfiehlt außerdem die Standard Ressourcenlimitierung „DefaultTasksMax“ zu erhöhen. Dieser Zahl spezifiziert die maximale Anzahl an Prozessen oder Threads, die ein Service erzeugen kann. Der Standardwert von 512 kann in manchen Fällen von Pacemaker überschritten werden was zu Problemen beim Heartbeat Mechanismus führt. Außerdem sollte bei Nutzung von SLES12/11 die Größe des Änderungscache reduziert werden, um einem bekannten Performancebug vorzubeugen.

```
sudo vi/etc/systemd/system.conf
# Ändern Sie DefaultTasksMax von 512 zu 4096
DefaultTasksMax=4096
sudo systemctl daemon-reload
sudo vi/etc/sysctl.conf
# Setzen Sie die folgenden Werte:
```

```
vm.dirty_bytes = 629145600
vm.dirty_background_bytes = 314572800
```

[A] Bevor wir die Cluster Initialisierung anstoßen können, müssen wir zunächst passwortlose SSH Kommunikation innerhalb des Clusters ermöglichen, sowie das Azure Python SDK installieren. Hierfür müssen wir das SLES Public Cloud Modul auf allen Hosts aktivieren. Achten Sie hierbei auf die korrekte Angabe ihrer SUSE Version (12/15). Außerdem konfigurieren wir die interne Namenslösung mittels/etc/hosts.

```
[A]
sudo SUSEConnect -p sle-module-public-cloud/12/x86_64
sudo zypper in -y python-azure-mgmt-compute
sudo vi /etc/hosts
# Fügen Sie beide VMs in die Hosts Datei hinzu -> <priv IP> <Hostname>
    10.1.1.1 db0
    10.1.1.2 db1
sudo ssh-keygen
# 3x Enter (Standard Pfad, kein Passwort, kein Passwort)
sudo cat/root/.ssh/id_rsa.pub
# Kopieren Sie die öffentlichen Schlüssel für [db0-pubkey] und [db1-pubkey]
[db0]
echo "[db1-pubkey]" >> vi /root/.ssh/authorized_keys
[db1]
echo "[db0-pubkey]" >> vi /root/.ssh/authorized_keys
# Fügen Sie den Key von
```

[db0] Nun können wir die Initialisierung des Clusters starten. Sollte danach gefragt werden den privaten Schlüssel/root/.ssh/id\_rsa zu überschreiben, verneinen sie mit „n“, ansonsten akzeptieren sie die Standardvorschläge, wobei wir keine Administration IP für Hawk konfigurieren.

```
sudo ha-cluster-init -u
# ! NTP is not configured to start at system boot.
# Do you want to continue anyway (y/n)? y
# /root/.ssh/id_rsa already exists - overwrite (y/n)? n
# Address for ring0 <IP-[db0]> Press ENTER
# Port for ring0 [5405] Press ENTER
# SBD is already configured to use <SBD Device> overwrite? n
#WARNING: Not configuring SBD - STONITH will be disabled.
# Do you wish to configure an administration IP (y/n)? n
```

[db1] Nachdem das Cluster initialisiert wurde, können wir nun den zweiten Knoten hinzufügen.

```
sudo ha-cluster-join
# IP address of hostname of existing node (e.g.: 192.168.1.1) [] <IP-[db0]>
```

[A] Auf beiden Knoten sind nun einige Nachbearbeitungsschritte notwendig. Zunächst passen wir die Corosync Einstellungen an. Die Defaultwerte sind für on-premise Installationen optimiert und spezifizieren einen quorum timeout (token) von 5 s. Für Azure ist dieser Wert zu niedrig. Aufgrund von Wartungsarbeiten kommt es z. B. vor, dass VMs zwischen Servern migriert werden, wodurch es zu kurzen timeout Szenarien kommt. Erfahrungswerte haben hier gezeigt, dass der Default Timeout zu klein gesetzt ist und es dadurch „häufiger“ zu Szenarien kommt, wo ein Knoten als fehlerhaft markiert wird, obwohl dieser innerhalb weniger Sekunden wieder verfügbar ist. Das Cluster wird dadurch instabil, weshalb inzwischen ein Wert von 30 s empfohlen ist. Durch Anpassen des Tokens müssen wir außerdem den „consensus“ Wert ändern. Dieser sollte immer das 1,2-fache von token sein. Außerdem setzen wir ein Passwort für den im Rahmen der Cluster Initialisierung neu erstellten Benutzer „hacluster“.

```
passwd hacluster
vi /etc/corosync/corosync.conf
    token: 30000
    consensus: 36000
sudo service corosync restart
```

[db0] Abschließend konfigurieren wir auf dem primary Knoten die von Azure empfohlenen Standardwerte bei Verwendung von Pacemaker auf Basis von SBD. Wir legen den Timeout für STONITH Aktivitäten auf 144 s fest (default 60) und setzen das Fencing auf 15 s.

```
sudocrm configure property stonith-timeout=144
sudocrm configure property stonith-enabled=true
sudocrm resource stop stonith-sbd
sudocrm configure delete stonith-sbd
sudocrm configure primitive stonith-sbd stonith:external/sbd \
    params pcmk_delay_max="15" \
    op monitor interval="15" timeout="15"
```

[db0] Optional empfiehlt es sich, noch Azure Events für Pacemaker zu konfigurieren. Azure Dienste sind nicht frei von geplanten Wartungereignissen. Damit der Kunde auf derartige Vorgänge reagieren kann, stellt Azure einen Metadatendienst namens „Scheduled Events“ bereit, welcher über anstehende Wartungereignisse informiert. Der Resource Agent „azure-events“ kann diese Ereignisse überwachen und entsprechend darauf reagieren um z. B. Ressourcen auf einen anderen Knoten zu migrieren bevor das entsprechende Ereignis eintritt. Dies geschieht indem der Knoten welcher vom

Wartungssereignis betroffen ist in den Standby Modus versetzt wird. Der „azure-events“ Agent sollte bereits mit Installation des Pakets „resource-agents“ installiert worden sein. Auf dem Hauptknoten (ha1-vc-db0) müssen sie lediglich die entsprechenden Pacemaker Ressourcen konfigurieren. Hierzu versetzen wir zunächst das Cluster in den „Wartungsmodus“ und konfigurieren anschließend die Ressource.

```
sudocrm configure property maintenance-mode=true  
sudocrm configure primitive rsc_azure-events ocf:heartbeat:azure-events op monitor interval=10s  
sudocrm configure clone cln_azure-events rsc_azure-events  
#Wartungsmodus beenden -> Warnungen können ignoriert werden  
sudocrm configure property maintenance-mode=false
```

#### 7.3.3.4 Exkurs Azure Fencing Agent

Anstelle des SBD Setups auf Basis von iSCSI Zielservern, kann alternativ auch Fencing über den Azure Fencing Agent ermöglicht werden. Das Einrichten des Pacemaker Clusters erfolgt analog zu 7.3.3.3. beantworten Sie jedoch die Clusterinitialisierung (ha-cluster-init -u) die Frage zur Nutzung von SBD mit „n“. Installieren Sie außerdem das „fence-agents“ Paket.

```
sudozypper install fence-agents"
```

Anschließend erstellen wir unseren Fence Agent. Hierfür werden wir die Azure Cloud Shell verwenden auf welche wir über das Portal zugreifen. Alternativ können Sie auch die Azure CLI nutzen indem Sie die notwendigen Pakete auf ihrem Jumpserver installieren [9]. Suchen Sie zunächst nach dem Dienst „Azure Active Directory“. Öffnen Sie nun in der linken Navigationsleiste den Menüpunkt „Eigenschaften“ und kopieren Sie ihre „**Mandanten-ID**“. Wählen Sie anschließend den Menüpunkt „App-Registrierung“ und klicken Sie auf den Button „**Neue Registrierung**“. Wählen Sie einen Namen z. B. „**S4HFenceAgent**“ und selektieren Sie in den unterstützten Kontotypen „Nur Konten in diesem Organisationsverzeichnis“. Geben Sie außerdem eine Umleitungs-URI vom Typen „Web“ und mit URL <http://localhost.com> an (die URL wird nicht tatsächlich genutzt). Wählen Sie nun den erstellten „S4HFenceAgent“ aus. Im Bereich „Übersicht“ kopieren Sie die „**Anwendungs-ID**“ und klicken Sie anschließend in der linken Navigationsleiste auf „**Zertifikate & Geheimnisse**“. Erstellen Sie darin einen neuen geheimen Schlüssel „**ClientSecret0**“ mit einer Gültigkeit von 24 Monaten. Sobald dieser angelegt wurde kopieren Sie die Zelle in der Spalte „**Wert**“ dieser wird als Passwort für den Service Principal benötigt. Navigieren Sie abschließend zum Dienst „Abonnements“ und kopieren Sie die angezeigte **Abonnement-ID**.

Öffnen Sie nun die „Cloud Shell“ indem Sie auf das Konsolen-Icon rechts neben der Suchleiste im Azure Portal klicken. Die Cloud Shell benötigt eine Azure-Dateifreigabe. Wenn Sie noch über kein Speicherkonto verfügen, erhalten Sie eine Meldung, ob Sie einen neuen Speicher erstellen wollen. Beachten Sie, dass durch ein Speicherkonto

zusätzliche Gebühren anfallen. Die Dateifreigabe hat zwar nur eine Größe von etwa 6 GiB, jedoch wird der Dateidienst mit einer Freigabekapazität von 5 TiB angelegt. Wir werden die Cloud Shell im Folgenden nutzen, um eine benutzerspezifische Rolle für den Azure Fencing Agent anzulegen. Diese soll es erlauben virtuelle Maschinen zu starten/ stoppen sowie Informationen zu lesen.

```
vi customRoleFenceAgent.json
# Fügen Sie die folgenden Zeilen hinzu
{
    "Name": "Custom Role Fence Agent",
    "description": "Allows to power-off and start virtual machines",
    "assignableScopes": [
        "/subscriptions/<Abonnement-ID>"
    ],
    "actions": [
        "Microsoft.Compute/*/read",
        "Microsoft.Compute/virtualMachines/powerOff/action",
        "Microsoft.Compute/virtualMachines/start/action"
    ],
    "notActions": [],
    "dataActions": [],
    "notDataActions": []
}
```

Erstellen Sie nun die neue Rolle mit dem Aufruf:

```
New-AzRoleDefinition -InputFile "customRoleFenceAgent.json"
```

Wir haben nun die neue Rolle definiert. Nun müssen wir in der Zugriffssteuerung der jeweiligen virtuellen Maschinen (für welche der Fence Agent zuständig ist) den Fence Agent mit der Rolle verknüpfen. Navigieren Sie dazu zum virtuellen Computer z. B. „**s4h-vc-db0**“ und wählen Sie in der linken Navigationsleiste den Punkt „Zugriffssteuerung (IAM)“ aus. Klicken Sie anschließend auf den Button „Hinzufügen“ und „Rollenzuweisung hinzufügen“. Treffen Sie die Auswahl in der Maske wie folgt:

- Rolle: Custom Role Fence Agent
- Zugriff zuweisen zu: Benutzer, Gruppe oder Dienstprinzipal
- Auswählen: S4HfenceAgent

Klicken sie anschließend auf „Speichern“ und wiederholen Sie die Schritte für weitere virtuelle Computer. Damit der FencingAgent als STONITH Device verwendet wird müssen wir abschließend noch die Clusterkonfiguration erweitern:

```
sudocrm configure property stonith-enabled=true
sudocrm configure property concurrent-fencing=true
sudocrm configure primitive rsc_st_azure stonith:fence_azure_arm \
    params subscriptionId="<Abonnement-ID>" resourceGroup="<Ressourcen
gruppe z.B. s4h-rg>" tenantId="<Mandanten-ID>" login="Anwendungs-ID" 
passwd="<Passwort Wert>" \
    pcmk_monitor_retries=4 pcmk_action_limit=3 power_timeout=240 pcmk_
reboot_timeout=900 \
    op monitor interval=3600 timeout=120
sudocrm configure property stonith-timeout=900
```

### 7.3.4 Einrichten des HANA HA-Clusters

Nachdem das Pacemaker Cluster eingerichtet wurde, wollen wir in diesem Kapitel das HANA Cluster einrichten. Dazu werden wir zunächst zwei HANA Datenbanken installieren und die HANA Systemrepplikation zwischen beiden Knoten einrichten. Anschließend werden wir Pacemaker so einrichten, dass ein automatisches Failover zwischen den Knoten möglich ist.

#### 7.3.4.1 Installation der HANA

Falls Sie die Schnellstartvorlage verwendet haben, sollten 10 Datenträger + optional weitere SBD Geräte einhängt sein. Außerdem sollten bereits eigenständige Mountpoints für die Pfade/mnt, /usr/sap, /hana/shared, /hana/backup, /hana/data und /hana/log eingestellt sein.

Für data, backup und log wurde jeweils eine eigenständige LVM Volume Group mit jeweils 2 × jeweils 2 × physischen Disks eingerichtet (insgesamt 6x). Die restlichen vier Disks werden für root, /hana/shared, /mnt und /usr/sap.

Sollten Sie ihr Filesystem Layout manuell konfigurieren wollen verwenden Sie entweder die grafische Partitionierungshilfe via „yast -> partitioner“ oder die entsprechenden Terminal Befehle (pvcreate, vgcreate, lvcreate, mkfs.xfs, mkdir, mount, sudo vi/etc/fdisk). Stellen Sie dabei sicher, dass die von Ihnen genutzten Platten den Performanceanforderungen von HANA entspricht:

- **/hana/log** mind. 250 MB/sec read/write für 1 MB I/O
- **/hana/data** mind. 400 MB/sec read und 250 MB/sec write für 16 MB und 64 MB I/O

Für HANA System Replication ist es notwendig, dass beide Hana Systeme mit identischer SID (HA1) und Instanznummer (03) installiert werden. Machen Sie die HANA Installationsdateien verfügbar und führen Sie „hdblcm“ aus um die HANA Installation durchzuführen. Für unser Setup sind keine speziellen Komponenten erforderlich. In diesem Beispiel nutzen wir lediglich „server“, „client“ und „afl“ zur Serverinstallation.

### 7.3.4.2 Systemreplikation einrichten

[A] Damit der Datenbestand der Primary HANA an den Secondary übermittelt werden, müssen wir als nächsten Schritt die HANA Systemreplikation einrichten. Zunächst erstellen wir eine Datenbank für das künftige S4HANA System. Für den Tenantnamen verwenden wir die SID des künftigen S/4 Systems. Anschließend müssen wir die Systemdatenbank sowie alle Tenants des HANA Systems sichern. Dies ist eine Voraussetzung, um die Systemreplikation einzurichten.

**Hinweis** Der Standardpfad für backups lautet/hana/shared/<SID>/HDB<Instanznr.>/backup/data/

Der Mountpoint/hana/backup wird also nicht standardmäßig verwendet.

```
su - haladm
hdbsql -u SYSTEM -p "<Passwort>" -i 03 -d SYSTEMDB 'CREATE DATABASE
S4H SYSTEM USER PASSWORD "<Passwort>"'
hdbsql -d SYSTEMDB -u SYSTEM -p "<Passwort>" -i 03 "BACKUP DATA USING
FILE ('initialSYS')"
hdbsql -d HA1 -u SYSTEM -p "<Passwort>" -i 03 "BACKUP DATA USING FILE
('initialHA1')"
hdbsql -d S4H -u SYSTEM -p "<Passwort>" -i 03 "BACKUP DATA USING FILE
('initialS4H')"
```

[db0] Da die Systemreplikation verschlüsselt erfolgt müssen wir außerdem die PKI Systemdateien des primary Knotens auf den Secondary übertragen.

```
su - haladm
scp /usr/sap/HA1/SYS/global/security/rsecssfs/data/SSFS_HA1.DAT    hal-
vc-db1:/usr/sap/HA1/SYS/global/security/rsecssfs/data/
scp /usr/sap/HA1/SYS/global/security/rsecssfs/key/SSFS_HA1.KEY hal-vc-
db1:/usr/sap/HA1/SYS/global/security/rsecssfs/key/
```

Als nächstes definieren wir den Standortnamen. Hierbei sollten nicht Begriffe wie „primary/secondary“ oder „master / slave“ verwendet werden, da diese Rolle sich jederzeit ändern können. Wir verwenden „DB0“ und „DB1“ als Bezeichnung (gemäß den Hostnamen unserer Systeme). Als erstes aktivieren wir die Systemreplikation für db0 und registrieren anschließend db1 beim primary Knoten.

```
[db0]
hdbnsutil -sr_enable --name=DB0
[db1]
sapcontrol -nr 03 -function StopWait 600 10
hdbnsutil -sr_register --remoteHost=hal-vc-db0 --remoteInstance=03
--replicationMode=sync --name=DB1
```

[db0] Prüfen Sie abschließend den Status der Replikation. DB0 sollte als „primary“ und „source system“ erkannt werden. Außerdem sollte der Replizierungsmodus von DB1 mit „sync“ bezeichnet sein. Optional können Sie sich weitere Details zur Replizierung mit dem Python Skript „systemReplicationStatus.py“ anzeigen.

```
hdbnsutil -sr_state  
HDBSettings.sh systemReplicationStatus.py
```

[A] Ab HANA2.0 empfiehlt es sich außerdem den Python Systempelikationshook zu konfigurieren, welcher für eine bessere Integration von HANA in das Pacemaker Cluster ermöglicht. Dazu erstellen wir zunächst ein neues Verzeichnis unter/hana/shared, wo wir das Pythonskript speichern, und ermöglichen den sudo Zugriff für den Benutzer <sid>adm. Anschließend machen wir einen Eintrag in der „global.ini“ zur Nutzung des Hooks und starten abschließend beide HANA Systeme neu.

```
mkdir - p/hana/shared/myHooks  
cp /usr/share/SAPHanaSR/SAPHanaSR.py /hana/shared/myHooks  
chown -R haladm:sapsys /hana/shared/myHooks  
sapcontrol -nr 03 -function StopSystem  
cat << EOF > /etc/sudoers.d/20-saphana  
haladm ALL=(ALL) NOPASSWD: /usr/sbin/crm_attribute -n hana_hal_site_  
srHook_*  
EOF  
Vi /hana/shared/HAN1/global/hdb/custom/config/global.ini  
# Fügen Sie folgende Zeilen hinzu  
[ha_dr_provider_SAPHanaSR]  
provider = SAPHanaSR  
path = /hana/shared/myHooks  
execution_order = 1  
[trace]  
ha_dr_saphanasr = info  
sapcontrol -nr 03 -function StartSystem
```

[db0] Prüfen Sie abschließend die korrekte Installation der HANA.

```
cdtrace  
awk '/ha_dr_SAPHanaSR.*crm_attribute/ \  
{ printf "%s %s %s %s\n",$2,$3,$5,$16 }' nameserver_*
```

### 7.3.4.3 HANA Cluster Ressourcen in Pacemaker definieren

[db0] Alle folgenden Schritte dieses Kapitels können auf einer beliebigen VM ausgeführt werden solange diese Teil des Clusters sind. Zunächst muss die HANA-Topologie als Cluster Ressource angelegt werden. Dazu müssen wir das Cluster in den Wartungsmodus setzen.

```

sudo crm configure property maintenance-mode=true
sudo crm configure primitive rsc_SAPHanaTopology_HA1_HDB03
ocf:suse:SAPHanaTopology \
operations \${id="rsc_sap2_HA1_HDB03-operations" \
op monitor interval="10" timeout="600" \
op start interval="0" timeout="600" \
op stop interval="0" timeout="300" \
params SID="HA1" InstanceNumber="03"

sudo crm configure clone cln_SAPHanaTopology_HA1_HDB03 rsc_
SAPHanaTopology_HA1_HDB03 \
meta clone-node-max="1" target-role="Started" interleave="true"

```

Wir erstellen nun die HANA-Ressourcen. Passen Sie ggf. die SID und Instanznummer entsprechend ihrer Konfiguration an. Achten Sie darauf, dass der Azure Loadbalancer Port sich aus der Instanznummer ableitet (in diesem Beispiel 62.503 und dass Sie die Frontend IP des Lastenausgleichs korrekt angeben.

```

sudo crm configure primitive rsc_SAPHana_HA1_HDB03 ocf:suse:SAPHana \
operations \${id="rsc_sap_HA1_HDB03-operations" \
op start interval="0" timeout="3600" \
op stop interval="0" timeout="3600" \
op promote interval="0" timeout="3600" \
op monitor interval="60" role="Master" timeout="700" \
op monitor interval="61" role="Slave" timeout="700" \
params SID="HA1" InstanceNumber="03" PREFER_SITE_TAKEOVER="true" \
DUPLICATE_PRIMARY_TIMEOUT="7200" AUTOMATED_REGISTER="false"
sudo crm configure ms msl_SAPHana_HA1_HDB03 rsc_SAPHana_HA1_HDB03 \
meta notify="true" clone-max="2" clone-node-max="1" \
target-role="Started" interleave="true"
sudo crm configure primitive rsc_ip_HA1_HDB03 ocf:heartbeat:IPaddr2 \
meta target-role="Started" \
operations \${id="rsc_ip_HA1_HDB03-operations" \
op monitor interval="10s" timeout="20s" \
params ip="<IP Lastenausgleich>"
```

Wie bereits erläutert führt der Integritätstest des Azure Lastenausgleichs TCP Anfragen an einen definierten Port (z. B. 62503) durch um zu überprüfen ob die Maschine als Lastenausgleichsziel geeignet ist. Abhängig von Ihrer SUSE Version sollten Sie entweder „socat“ oder den neueren „azure-lb“ als Dienst verwenden. In diesem Beispiel arbeiten wir mit „socat“ und fügen diese als Clusterressource hinzu. Pacemaker muss sicherstellen, dass der Dienst immer nur auf dem primary Knoten aktiv ist. Die restliche Clusterkonfiguration folgt den Empfehlungen von SUSE und Microsoft. Schließen Sie die Konfiguration ab, indem Sie den Status der Clusterressourcen überprüfen.

```
sudo crm configure primitive rsc_nc_HA1_HDB03 anything \
    params binfile="/usr/bin/socat" cmdline_options="-U TCP-LISTEN:625
03,backlog=10,fork,reuseaddr/dev/null" \
    op monitor timeout=20s interval=10
sudo crm configure primitive rsc_nc_HA1_HDB03 azure-lb port=62503 \
    meta resource-stickiness=0
sudo crm configure group g_ip_HA1_HDB03 rsc_ip_HA1_HDB03 rsc_nc_HA1_
HDB03
sudo crm configure colocation col_saphana_ip_HA1_HDB03 4000: g_ip_HA1_
HDB03:Started \
    ms1_SAPHana_HA1_HDB03:Master
sudo crm configure order ord_SAPHana_HA1_HDB03 Optional: cln_
SAPHanaTopology_HA1_HDB03 \
    ms1_SAPHana_HA1_HDB03
sudo crm resource cleanup rsc_SAPHana_HA1_HDB03
sudo crm configure property maintenance-mode=false
sudo crm configure rsc_defaults resource-stickiness=1000
sudo crm configure rsc_defaults migration-threshold=5000
sudo crm_mon -r
```

Damit ist die Konfiguration abgeschlossen. Wir haben nun ein hochverfügbares HANA Cluster und können mit der Bereitstellung des S/4HANA Systems fortfahren.

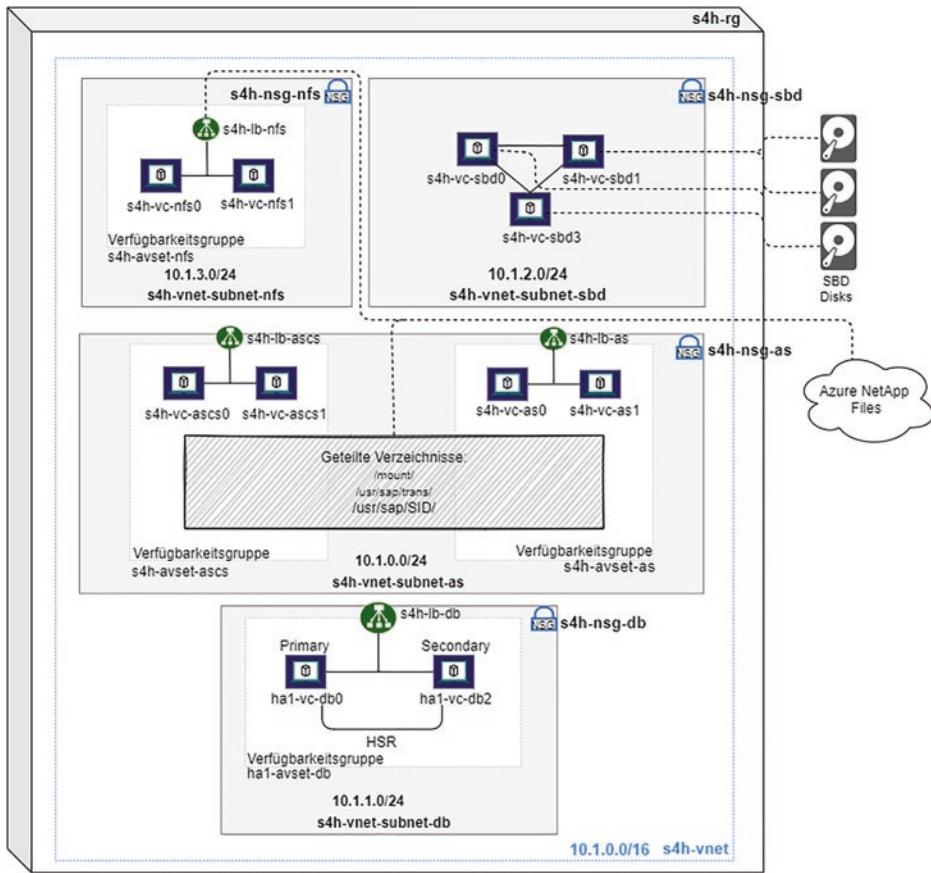
---

## 7.4 Bereitstellen eines S/4HANA HA Clusters

In diesem Kapitel schließen wir die Bereitstellung der Beispielarchitektur ab, indem wir ein hochverfügbares S/4HANA System aufsetzen. Wir werden zunächst die hierfür noch ausstehenden Ressourcen bereitstellen und dabei tiefer auf die Funktionsweise des Azure Lastenausgleichs eingehen. Anschließend werfen wir einen näheren Blick auf hochverfügbare geteilte Speichermöglichkeiten in Azure indem wir zunächst ein NFS-Cluster bereitstellen und anschließend Azure Netapp Files als Alternative dazu vorstellen. Darauf aufbauend erstellen wir das ASCS Cluster unter der Nutzung des Enqueue Replication Server 2. Den Abschluss bildet schließlich die Installation der Anwendungsserver.

### 7.4.1 SAP S/4HANA Clusterarchitektur

Wie in Abb. 7.12 dargestellt wird das „**s4h-vnet**“ um ein weiteres Subnet „**s4h-vnet-subnet-nfs**“ erweitert in welchem wir zwei NFS Server in einem Pacemaker Cluster inklusive Verfügbarkeitsgruppe bereitstellen. Das Cluster nutzt die bestehenden iSCSI-Zielserver für das SBD Fencing. Wir benötigen NFS als hochverfügbare Lösung um sicherzustellen, dass auch bei einem Knotenausfall die geteilten Verzeichnisse/usr/



**Abb. 7.12** SAP S/4HANA Clusterarchitektur

sap/<SID>, /usr/sap/trans, sowie/sapmnt verfügbar bleiben und es nicht zu einer Beeinträchtigung des S/4HANA Clusters kommt. Wir werden außerdem ANF als hochverfügbares SaaS Angebot vorstellen. Es steht Ihnen frei alle Verzeichnisse über eine NFS oder ANF bereitzustellen. Die ASCS und ERS Instanz werden wir schließlich im bestehenden „s4h-vnet-subnet-as“ bereitstellen inklusive der weiteren Anwendungsserver. Im Unterschied zur ASCS und ERS Instanz benötigen die Anwendungsserver (AAS/PAS) kein Pacemaker Cluster, wir werden diese jedoch ebenfalls als Teil einer Verfügbarkeitsgruppe bereitstellen, um einen gleichzeitigen Ausfall beider Instanzen vorzubeugen.

## 7.4.2 Bereitstellen der SAP S/4HANA Cluster Ressourcen

Im Gegensatz zur HANA Bereitstellung mittels Schnellstartvorlage in Kapitel 7.3, werden wir in diesem Kapitel die Azure Ressourcen manuell erstellen und konfigurieren. Wenn Sie lieber erneut eine Vorlage verwenden möchten, können Sie z. B. die „application-workloads/sap/sap-3-tier-marketplace-image-converged“ verwenden. Diese Vorlage ist ähnlich zur Datenbank Vorlage und beinhaltet einen Lastenausgleich, eine Verfügbarkeitsgruppe, sowie virtuelle Maschinen für ASCS, ERS und Anwendungsserver.

### 7.4.2.1 Virtuellen Computer anlegen

**Verfügbarkeitsgruppe** Legen Sie zunächst zwei neue Verfügbarkeitsgruppen für die ASCS und ERS Instanzen, sowie für die Anwendungsserver PAS und AAS an. Suchen Sie dazu nach „Verfügbarkeitsgruppe“ und verwenden Sie die folgenden Angaben:

- Ressourcengruppe: s4h-rg
- Name: s4h-avset-ascsc/s4h-avset-as
- Fehlerdomänen: 2
- Updatedomänen: 20
- Verwaltete Datenträger: Ja (ausgerichtet)

**Virtuelle Computer (ASCS/ERS)** Erstellen Sie nun zwei virtuelle Computer für das Central Service Cluster (#create/Microsoft.VirtualMachine-ARM). Verwenden Sie die folgenden Parameter (für Detailschritte siehe Sektion 7.3.2.4.1):

- Ressourcengruppe: s4h-rg
- Name: s4h-vc-ascsc0/s4h-vc-ascsc1
- Verfügbarkeitsoption: Verfügbarkeitsgruppe
- Verfügbarkeitsgruppe: s4h-avset-ascsc
- Image: z. B. SLES for SAP 12 SP5-BYOS
- Größe: z. B. Standard\_D2s\_v3 (2 vcpus, 8 GB RAM)
- OS-Datenträger: SSD Premium
- Zusätzlicher-Datenträger: keine
- Virtuelles Netzwerk: s4h-vnet
- Subnetz: s4h-vnet-subnet-as
- IP: Statische interne IP (z. B. 10.1.0.1/10.1.0.2)+Statische Öffentliche IP (basic)
- Netzwerksicherheitsgruppe: s4h-nsg-as
- Beschleunigter Netzwerkbetrieb: Ja
- Lastenausgleich platzieren: Nein

**Virtuelle Computer (PAS/AAS)** Erstellen Sie außerdem zwei virtuelle Computer für den Anwendungsserverpool (für Detailschritte siehe Sektion 7.3.2.4.1):

- Ressourcengruppe: s4h-rg
- Name: s4h-vc-as0/s4h-vc-as1
- Verfügbarkeitsoption: Verfügbarkeitsgruppe
- Verfügbarkeitsgruppe: s4h-avset-as
- Image: z. B. SLES for SAP 12 SP5-BYOS
- Größe: z. B. Standard\_DS4\_v2 (8 vcpus, 28 GB RAM)
- OS-Datenträger: SSD Premium
- Zusätzlicher-Datenträger: 128 GB SSD Premium (P10)
- Virtuelles Netzwerk: s4h-vnet
- Subnetz: s4h-vnet-subnet-as
- IP: Statische interne IP (z. B. 10.1.0.3/10.1.0.4) + Statische Öffentliche IP (basic)
- Netzwerksicherheitsgruppe: s4h-nsg-as
- Beschleunigter Netzwerkbetrieb: Ja
- Lastenausgleich platzieren: Nein

### 7.4.2.2 Lastenausgleich erstellen und konfigurieren

**Lastenausgleich** Der Lastenausgleich ist in zwei SKUs verfügbar „Standard“ und „Basic“. Der Basic Loadbalancer ist kostenlos unterliegt jedoch einigen Einschränkungen wie z. B., dass „nur“ bis zu 300 Instanzen angebunden werden können und dass sich alle Backend Systeme in einer einzigen Verfügbarkeitsgruppe oder Skalierungsgruppe befinden müssen. Für das von uns verwendete Beispiel ist daher die Nutzung des Basic Lastenausgleichs möglich. Um die Konfiguration zu erleichtern, nutzen wir jedoch die Standard Version, welche auch von Microsoft empfohlen ist.

#### 7.4.2.2.1 ASCS/ERS Lastenausgleich

Erstellen Sie einen Lastenausgleich für das ASCS/ERS-Cluster (#create/Microsoft.LoadBalancer-ARM). Treffen Sie folgenden Werte für die Grundeinstellungen:

- Ressourcengruppe: s4h-rg
- Name: s4h-lb-asc
- Typ: Intern
- SKU: Standard
- Tarif: Regional

Fügen Sie anschließend zwei Front-End-IP-Adressen hinzu über welche der Lastenausgleich Anfragen entgegennimmt:

- Name: s4h-lb-ascs-fip-ascs/s4h-lb-ascs-fip-ers
- Virtuelles Netzwerk: s4h-vnet
- Subnetz: s4h-vnet-subnet-as
- Zuordnung: Statisch
- IP-Adresse: z. B. 10.1.0.5/10.1.0.6
- Verfügbarkeitszone: Zonenredundant

Definieren Sie anschließend zwei Backendpools. Diese beinhalten jeweils die gleichen virtuellen Computer (ASCS und ERS Instanz).

- Name: s4h-lb-ascs-backend-ascs/s4h-lb-ascs-backend-ers
- Virtuelles Netzwerk: s4h-vnet
- Konfiguration des Back-End-Pools: Netzwerkschnittstelle
- IP-Version: IPv4
- Virtuelle Computer [s4h-vc-ascs0,s4h-vc-ascs1] / [s4h-vc-ascs0,s4h-vc-ascs1]

Abschließend konfigurieren wir die eigentlichen Lastenausgleichsregeln. Falls Sie den Lastenausgleich in der SKU „Standard“ verwendet haben, ist es ausreichend „HA-Ports“ zu wählen. Im Falle von „Basic“ müssen die jeweiligen Ports für die ASCS und ERS basierend auf der Instanz Nummer (IN) manuell konfigurieren. Dies Lauten für die ASCS Instanz (32<IN>, 36<IN>, 39<IN>, 81<IN>, 5<IN>13, 5<IN>14, 5<IN>16) und für die ERS Instanz (32<IN>, 33<IN>, 5<IN>13, 5<IN>14, 5<IN>16).

- Name: s4h-lb-ascs-inbrule-ascs
- IP-Version: IPv4
- Front-End-IP-Adresse: s4h-lb-ascs-fip-ascs/s4h-lb-ascs-fip-ers
- HA-Ports: Ja
- Back-End-Pool: s4h-lb-ascs-backend-ascs/s4h-lb-ascs-backend-ers
- Integritätstest ASCS: Neu erstellen -> s4h-lb-ascs-probe-ascs, Port=621<InstNr. ASCS> (62.100), Intervall=5, Fehlerschwellenwert=2
- Integritätstest ERS: Neu erstellen -> s4h-lb-ascs-probe-ers, Port=621<InstNr. ERS> (62.102), Intervall=5, Fehlerschwellenwert=2
- Sitzungspersistenz: Keine
- Leerlauftimeout (Minuten): 30
- TCP-Zurücksetzung: Deaktiviert
- Floating IP: Aktiviert

#### 7.4.2.2 Lastenausgleich für den Anwendungsserverpool

Optional können Sie außerdem einen Lastenausgleich für den Anwendungsserverpool erstellen. Für ein Hochverfügbarkeitsszenario auf Basis von SAPGUI ist dies nicht notwendig, jedoch wenn Ihre Anwendungsserver auch Webdienste wie z. B. einen embedded Fiori Frontendserver verwenden. Beachten Sie, dass es sich hierbei um

eine Aktiv/Aktiv Konfiguration handelt (im Gegensatz zum HANA Aktiv/Passiv Szenario), d. h. dass der Integritätstest für alle Fiori Frontendserver erfolgreich sein muss (Pacemaker socat Steuerung). Berücksichtigen Sie außerdem die Nutzung einer Sitzungspersistenz. Die Einstellung „Client IP“ sollte für die meisten Szenarien ausreichend sein. Dadurch werden alle Anfragen einer IP an denselben Backend Server weitergeleitet.

### 7.4.3 Hochverfügbarer NFS Speicher in Azure

Der klassische Weg, um hochverfügbaren geteilten Netzwerkspeicher für ein SAP-System bereitzustellen, ist die Nutzung eines NFS Clusters (bzw. SMB Failover-Cluster in Windows). Azure bietet jedoch auch den SaaS Dienst ANF, welcher erlaubt hochverfügbare SMB oder NFS Shares zu erstellen ohne dass Sie die hierfür notwendige Infrastruktur administrieren müssen. Im Folgenden werden wir beide Wege kurz beleuchten um Ihnen bei der Entscheidungsfindung zu unterstützen.

#### 7.4.3.1 Alternative1: NFS-Server Cluster

Für das NFS Cluster werden wir zwei virtuelle Maschinen erstellen. Manche der folgenden Konfigurationsschritte müssen entweder auf allen oder nur auf einem der beiden Knoten durchgeführt werden. Wir kennzeichnen dies mit den Bezeichnungen [nfs0] für „s4h-vc-nfs0“, [nfs1] für „s4h-vc-nfs1“ und [A] für alle.

##### 7.4.3.1.1 Azure Ressourcen anlegen

Erstellen Sie zunächst ein neues Subnetz „s4h-vnet-subnet-nfs“ [10.1.3.0/24] und platzieren Sie zwei neue virtuelle Computer (**s4h-vc-nfs0**, **s4h-vc-nfs1**), sowie eine Verfügbarkeitsgruppe (**s4h-avset-nfs**) darin. Orientieren Sie sich bei den Parametern an Kapitel 7.4.2. Da wir die virtuellen Computer ausschließlich als NFS für unser Demosystem verwenden ist eine kleine SKU wie z. B. „Standard\_DS1\_v2“ (1 vcpu, 3.5 GiB RAM) ausreichend. Binden Sie außerdem einen weiteren Datenträger ein z. B. 128 GB SSD Premium (P10).

Erstellen Sie einen **internen** Lastenausgleich „s4h-lb-nfs“. Für unsere Beispielarchitektur ist die SKU „basic“ ausreichend. Erstellen Sie eine frontendIP (**s4h-lb-ascs-fip-nfs**) sowie einen Back-End-Pool (**s4h-lb-nfs-backend**) mit den zwei eben angelegten virtuellen Computern. Für den Integritätstest (**s4h-lb-nfs-probe**) verwenden wir einen TCP Probe auf den Port 61000 mit einem Intervall von 5 s und einen Fehlerschwellenwert von 2. Erstellen Sie abschließend die Lastenausgleichsregeln. Wir erstellen die Regel „s4h-lb-nfs-inbrule-tcp“ für den Port 2049 als TCP Verbindung unter Verwendung der zuvor konfigurierten Einstellungen (IP, Pool und Probe) und mit einem Leerlauftimeout von 30 min, sowie aktiverter Floating IP. Da alte Protokollversionen (insb. NFSv2) weiterhin UDP als Protokoll verwenden erstellen wir außerdem die Regel „s4h-lb-nfs-inbrule-udp“ identisch zur vorherigen Regel, aber mit UDP als Protokoll.

#### 7.4.3.1.2 NFS Cluster konfigurieren

Nachdem die benötigten Ressourcen erstellt wurden, erstellen wir zunächst die grundlegende Pacemaker Cluster Konfiguration. Konfigurieren Sie dazu zunächst ein neues SBD Gerät auf Ihren iSCSI Zielserver (siehe Kapitel 7.3.2.4.2) und initialisieren Sie abschließend das Cluster gemäß Kapitel 7.3.3.

**[A] Filesystem Layout** Unser NFS Dienst wird Dateien über das Verzeichnis/srv/nfs/bereitstellen. Als Hintergundspeicher werden wir ein LVM über den per SCSI eingebunden zusätzlichen Datenträger erstellen. Als volumegroup verwenden wir „vg-NFS“ und als logical volume „lv-S4H“.

```
mkdir /srv/nfs
sudo sh -c 'echo srv/nfs/ *\\(rw,no_root_squash,fsid=0\\)>/etc/exports'
sudo sh -c 'echo -e "n\\n\\n\\n\\n\\nw\\n" | fdisk/dev/disk/azure/scsi1/lun0'
sudo vgcreate vg-NFS /dev/disk/azure/scsi1/lun0-part1
sudo lvcreate -l 100%FREE -n lv-S4H vg-NFS
```

Zur Replizierung der Dateien zwischen den beiden NFS Servern wird „drbd“ verwendet. Installieren Sie die notwendigen Pakete und treffen Sie anschließend die folgenden Konfigurationseinstellungen.

```
sudo zypper install drbd drbd-kmp-default drbd-utils
sudo vi /etc/drbd.conf
    # Stellen Sie sicher das die folgenden Zeilen enthalten sind
    include "drbd.d/global_common.conf";
    include "drbd.d/*.res"
sudo vi/etc/drbd.d/global_common.conf
    # Ändern Sie die „handlers“ sektion wie folgt ab
    global {
        usage-count no;
    }
    common {
        handlers {
            fence-peer "/usr/lib/drbd/crm-fence-peer.sh";
            after-resync-target    "/usr/lib/drbd/crm-unfence-peer.
sh";
            split-brain "/usr/lib/drbd/notify-split-brain.sh root";
            pri-lost-after-sb "/usr/lib/drbd/notify-pri-lost-after-
sb.sh; /usr/lib/drbd/notify-emergency-reboot.sh; echo b > /proc/sysrq-
trigger ; reboot -f";
        }
        startup {
            wfc-timeout 0;
    }
```

```

options {
}
disk {
    md-flushes yes;
    disk-flushes yes;
    c-plan-ahead 1;
    c-min-rate 100M;
    c-fill-target 20M;
    c-max-rate 4G;
}
net {
    after-sb-0pri discard-younger-primary;
    after-sb-1pri discard-secondary;
    after-sb-2pri call-pri-lost-after-sb;
    protocol C;
    tcp-cork yes;
    max-buffers 20000;
    max-epoch-size 20000;
    sndbuf-size 0;
    rcvbuf-size 0;
}
}
}

```

Wir erstellen nun das drbd Gerät /dev/drbd0 welches als Datenträger das logische Volume lv-S4H verwendet.

```

sudo vi /etc/drbd.d/S4H-nfs.res
#
resource S4H-nfs {
    protocol C;
    disk {
        on-io-error detach;
    }
    on s4h-vc-nfs0 {
        address 10.1.3.0:7790;
        device /dev/drbd0;
        disk /dev/vg-NFS/lv-S4H;
        meta-disk internal;
    }
    on s4h-vc-nfs1 {
        address 10.1.3.1:7790;
        device /dev/drbd0;
        disk /dev/vg-NFS/lv-S4H;
        meta-disk internal;
    }
}

```

Wir können nun die drbd Ressource online nehmen, zuvor müssen wir jedoch den Metadaten Speicher initialisieren.

```
sudo drbdadm create-md S4H-nfs
sudo drbdadm up S4H-nfs
```

**[nfs0]** Wir verkürzen nun die initiale resynchronization und setzen den Primary. Anschließend warten wir, bis die drbd Geräte synchronisiert sind und erstellen anschließend das Filesystemlayout.

```
drbdadm new-current-uuid --clear-bitmap S4H-nfs
drbdadm primary --force S4H-nfs
drbdsetup wait-sync-resource S4H-nfs
# Warten Sie bis die Synchronisierung abgeschlossen ist
sudo mkfs.xfs /dev/drbd0
sudo mkdir /srv/nfs/S4H
sudo chattr +i /srv/nfs/S4H
sudo mount -t xfs /dev/drbd0 /srv/nfs/S4H
sudo mkdir /srv/nfs/S4H/sidsys
sudo mkdir /srv/nfs/S4H/sapmntsid
sudo mkdir /srv/nfs/S4H/trans
sudo mkdir /srv/nfs/S4H/ASCS
sudo mkdir /srv/nfs/S4H/ASCSERS
sudo umount /srv/nfs/S4H
```

**[nfs0]** Abschließend nehmen wir die NFS drbd Geräte in die Cluster Konfiguration mit auf. Achten Sie auf die korrekte Angabe der Lastenausgleichs IP Adresse, sowie des Ports für den Integritätstest.

```
sudo crm configure rsc_defaults resource-stickiness="200"
sudo crm configure property maintenance-mode=true
sudo crm configure primitive drbd_S4H_nfs \
    ocf:linbit:drbd \
    params drbd_resource="S4H-nfs" \
    op monitor interval="15" role="Master" \
    op monitor interval="30" role="Slave"
sudo crm configure ms ms-drbd_S4H_nfs drbd_S4H_nfs \
    meta master-max="1" master-node-max="1" clone-max="2" \
    clone-node-max="1" notify="true" interleave="true"
sudo crm configure primitive fs_S4H_sapmnt \
    ocf:heartbeat:Filesystem \
    params device=/dev/drbd0 \
    directory=/srv/nfs/S4H \
    fstype=xfs \
```

```

op monitor interval="10s"
sudo crm configure primitive nfsserver systemd:nfs-server \
    op monitor interval="30s"
sudo crm configure clone cl-nfsserver nfsserver
sudo crm configure primitive exportfs_S4H \
    ocf:heartbeat:exportfs \
    params directory="/srv/nfs/S4H" \
    options="rw,no_root_squash,crossmnt" clientspec="*" fsid=1 wait_ \
for_leasetime_on_stop=true op monitor interval="30s"

sudo crm configure primitive vip_S4H_nfs \
    IPaddr2 \
    params ip=10.1.3.2 cidr_netmask=24 op monitor interval=10 \
timeout=20
sudo crm configure primitive nc_S4H_nfs azure-lb port=61000
sudo crm configure group g-S4H_nfs \
    fs_S4H_sapmnt exportfs_S4H nc_S4H_nfs vip_S4H_nfs
sudo crm configure order o-S4H_drbd_before_nfs inf: \
    ms-drbd_S4H_nfs:promote g-S4H_nfs:start
sudo crm configure colocation col-S4H_nfs_on_drbd inf: \
    g-S4H_nfs ms-drbd_S4H_nfs:Master
sudo crm configure property maintenance-mode=false

sudo crm configure property maintenance-mode=false

```

#### **7.4.3.2 Alternative2: Azure Netapp Files**

Anders als die meisten anderen Dienste kann ANF nicht standardmäßig aus dem Azure Portal gebucht werden. Um Zugriff auf ANF zu erhalten, muss zunächst eine sogenannte Wartelistenanforderung gesendet werden. Dieser Schritt ist notwendig da Azure sehr strenge SLA Anforderungen an den ANF Dienst stellt und daher zunächst Überprüfungen auf Basis ihrer Planungswerte durchführen muss. Erst wenn sie darauf eine Emailbestätigung erhalten, können sie ANF über das Portal buchen. Davor können Sie zwar ANF auswählen aber bei der Erstellung des NetApp-Kontos sind alle Optionen ausgegraut.

##### **7.4.3.2.1 Für ANF registrieren**

Füllen Sie zunächst die Wartelistenanforderung aus. Die Fragen umfassen die üblichen Kundenkontaktdaten, sowie ANF Nutzungsangaben. Die folgenden sechs Angaben möchten wir kurz hervorheben, da diese ggf. zusätzliche Vorausplanungen bzw. Abstimmungen erfordern:

- **Workload für ANF:** Hier wird z. B. auch zwischen SAP File Share (z. B. /sapmnt) und SAP HANA (z. B. /hana/data) unterschieden. Wählen Sie die Optionen die Ihrem Setup am nächsten kommt. Sie können auch mehrere Anwendungsszenarien gleichzeitig wählen.
- **Nutzung von HANA auf ANF:** Diese Ja/Nein Frage ist eine Erweiterung zur allgemeinen Workload Frage. Falls Sie diese Frage mit „Ja“ beantworten werden Sie zusätzlich aufgefordert ein weiteres Formular auszufüllen in welchem Sie unter anderem die erwarteten CPU Kerne, sowie die VM Größe der HANA spezifizieren sollen.
- **Anwendungsfälle/Use Cases:** Hier haben Sie die Wahl zwischen „Produktivumgebung“, „Entwicklungs-/Testumgebung“, „Disaster Recovery“ und „Synchronisierung“. Mehrfachantworten sind hier ebenfalls möglich.
- **Welche Regionen für ANF:** Hier müssen Sie die von Ihnen verwendete Region auswählen. Da ANF Subscriptions regionsabhängig sind muss Microsoft sicherstellen, dass die von Ihnen gesendeten Planungsparameter in der Zielregion umsetzbar sind
- **Subscription:** Tragen Sie hier die GUID ihres Abonnements ein. Es kann außerdem eine zusätzliche ID eingetragen werden falls Sie ANF für zwei Abonnements freischalten wollen. Für weitere Abos müssen Sie das Formular erneut ausfüllen.
- **Subscription Environment:** Wählen Sie hier „Azure Commercial“, wenn Sie keine US Behörde sind die besonderen Sicherheitsanforderungen unterliegt.

Nachdem Sie das Formular übermittelt haben warten Sie auf eine Mailbestätigung. Erstellen Sie anschließend einen neuen NetApp Account. Erstellen Sie dazu eine neue Ressource und Suchen Sie nach „Azure NetApp Files“ im Marketplace. Geben Sie hier einen Namen, das Abonnement welches zugeordnet werden soll, sowie die Region und Ressourcen Gruppe an.

#### 7.4.3.2.2 Kapazitätspool anlegen

Nachdem der ANF Account erstellt wurde können Sie nun einen Kapazitätspool anlegen. Hierbei handelt es sich lediglich um ein Containerobjekt welches später die eigentlichen Volumes, auch Quota genannt, bereitstellen wird. Folgende Einstellungen müssen sie beim Anlegen des Pools vornehmen:

**Name:** Verwenden Sie einen eindeutigen Namen für Ihren Kapazitätspool.

**Dienstebene:** Hier wählen Sie die Performanceschicht aus welche für alle Volumes des Pools angewandt wird. Sie können wählen zwischen Standard (16 MiBs/TB), Premium (64 MiBs/TB) und Ultra (128 MiBs/TB). Beachten Sie, dass für/hana/data ein Lesezugriff von mindestens 400 MB/sec Lesezugriff (bzw. 250 MB/s Schreibzugriff) empfohlen ist. Microsoft empfiehlt für alle Arten von SAP Workload die Performancestufe Ultra zu verwenden. Jedoch kann durch Wahl eines größeren Quotas die Performanceanforderungen auch mit einer niedrigeren Dienstebene erreicht werden.

Um also auf die Zielperformance von 400 MB/sec zu kommen können Sie alternativ 3,125 TiB Ultra-Speicher, 6,25 TiB Premium Speicher oder 25 TiB Standardspeicher bereitstellen.

**Größe (in TB):** Die Poolgröße beträgt mindestens 4 TiB und ist regional auf 25 TiB pro Abonnement limitiert. Das Limit von 25 TiB kann durch stellen einer Supportanfrage weiter vergrößert werden. Ein unterschreiten der 4 TiB Grenze ist jedoch nicht möglich. Beachten Sie, dass die Kosten nach Poolgröße berechnet werden, nicht für tatsächlich genutzten Speicher. Die Preise werden Pro GB abgerechnet und liegen bei Erstellung dieses Buches bei etwa 0,124 € für Standard, 0,248 € Premium und 0,331 € für Ultra-Speicher. Nachdem der Pool erstellt wurde, können Sie nun einzelne Volumes erstellen.

#### 7.4.3.2.3 ANF Volumes anlegen und einbinden

**Basis Konfiguration** Neben der Höhe der Quota müssen Sie hier auch ein virtuelles Netzwerk angeben. Da ANF Volumes als NFS oder Samba Share eingebunden werden sollten Sie hier das Netzwerk der künftigen Clients z. B. der ASCS oder HANA Instanzen auswählen. Auf diese Weise bleibt der Traffic innerhalb eines Netzwerks was sowohl aus Kosten- als auch Performancesicht lohnenswert ist. Geben Sie anschließend ein existierendes Subnetz an oder erstellen Sie ein eigenes. Beachten Sie, dass sie für ein ANF Netzwerk keine Netzwerksicherheitsgruppen erstellen können.

**Protokoll** Hier legen Sie fest ob das Volume als SMB oder NFS Share eingebunden werden soll. Außerdem legen Sie die Protokollversion fest. Sollte das Volume für HANA verwendet werden, beachten Sie, dass sie NFS v4.1 nutzen müssen um/hana/data und /hana/log über ANF bereitzustellen. Für/hana/shared können Sie NFS v3 verwenden. Für das Transportverzeichnis können Sie sowohl SMB als auch NFS verwenden. Für das Säpmount Verzeichnis empfiehlt es sich SMB für Windows Systeme bzw. NFS für Linux Systeme zu verwenden. Für NFS können Sie zusätzlich im Bereich „Export Policy“ den Zugriff auf bestimmte IP Adressen einschränken wobei die Standardeinstellungen alle Zugriffe erlaubt. Bei Verwendung von SMB müssen Sie zusätzlich eine Active Directory Verbindung sowie eine Bezeichnung für das Share anlegen (Sie können die Pflege der Active Directory Verbindungen im Hauptverzeichnis ihres ANF Accounts vornehmen).

Nachdem das Volume erstellt wurde können Sie dieses auswählen und unter „Einbindungsanweisungen“ Instruktionen zum Einbinden des shares betrachten. Das ANF Volume verhält sich wie ein reguläres Samba bzw. NFS Share. Im Falle von NFS werden in den Einbindungshinweisen auch die vorausgesetzten Pakete wie z. B. nfs-common/nfs-utils gelistet. Über den Windows NFS-Client können Sie auch NFS Shares unter Windows einbinden. Dafür müssen Sie jedoch das Share zunächst unter Linux einbinden und mittels chmod 777 oder 775 uneingeschränkte Lese- und Schreibzugriffe ermöglichen.

## 7.4.4 Installation ASCS und ERS

Wir verwenden im Folgenden die Bezeichnung [A] für Schritte die auf allen Knoten ausgeführt werden müssen und nutzen [ASCS] für „s4h-vc-asc0“ bzw. [ERS] für „s4h-vc-asc1“ falls ein Vorgang nur auf einem bestimmten virtuellen Computer durchgeführt werden sollen.

### 7.4.4.1 Filesystem vorbereiten und NFS einbinden

[A] Erstellen Sie zunächst einen Eintrag für das NFS Cluster in der/etc/hosts Datei. Verwenden Sie die Frontend-IP des NFS Lastenausgleichs „**s4h-lb-nfs**“. Fügen Sie anschließend die Frontend IPs des Lastenausgleichs s4h-lb-asc0 hinzu.

```
vi /etc/hosts
# Fügen Sie folgenden Eintrag hinzu
10.1.3.2 s4h-nfs
10.1.0.5 s4h-asc0
10.1.0.6 s4h-ers
```

Falls Sie Azure NetappFiles bzw. NFS in der Version 4.1 verwenden, stellen sie sicher, dass die Domänenkonfiguration in der Datei/etc/idmapd.conf identisch zu ANF ist. Ansonsten werden alle Dateien mit der ownership **nobody:nobody** gesetzt. Außerdem muss der Parameter/sys/module/nfs/parameters/nfs4\_disable\_idmapping auf „Y“ gesetzt sein.

Erstellen Sie nun die Verzeichnisse

```
sudo mkdir -p /sapmnt/S4H && chattr +i/sapmnt/S4H
sudo mkdir -p /usr/sap/trans && chattr +i/usr/sap/trans
sudo mkdir -p /usr/sap/S4H/SYS && chattr +i/usr/sap/S4H/SYS
sudo mkdir -p /usr/sap/S4H/ASC00 && chattr +i/usr/sap/S4H/ASC00
sudo mkdir -p /usr/sap/S4H/ERS02 && +i/usr/sap/S4H/ERS02
```

Anstelle von statischen mounts via fstab verwenden wir autofs zum dynamischen Einhängen von NFS Dateien. Fügen Sie die NFS Einträge hinzu und starten Sie anschließend den autofs service. Wir konfigurieren außerdem den Azure Linux Agent zur Nutzung des SWAPs.

```
sudo vi /etc/auto.master
# Fügen Sie die folgenden zwei Einträge hinzu
+auto.master
/- /etc/auto.direct
sudo vi/etc/auto.direct
# Fügen Sie folgende Einträge hinzu
/sapmnt/S4H -nfsvers=4,nosymlink,sync s4h-nfs:/S4H/sapmntsid
/usr/sap/trans -nfsvers=4,nosymlink,sync s4h-nfs:/S4H/trans
```

```
/usr/sap/S4H/SYS - nfsvers=4, nosymlink, sync s4h-nfs:/S4H/sidsys
sudo systemctl enable autofs
sudo service autofs restart
sudo vi /etc/waagent.conf
    # Treffen Sie folgende Einstellungen
    ResourceDisk.EnableSwap=y
    ResourceDisk.SwapSizeMB=2000
sudo service waagent restart
```

#### 7.4.4.2 Installation ASCS Cluster

[ASCS] Zunächst installieren wir den SUSE Cluster Connector. Hierbei handelt es sich um ein Bindeglied zwischen „sapcontrol“ zum Starten und Stoppen der SAP Instanzen und Pacemaker. Anschließend nehmen wir die ASCS Instanz in die Pacemaker Konfiguration mit auf.

```
sudo zypper install sap-suse-cluster-connector
sudo crm node standby s4h-vc-ascs1
sudo crm configure primitive fs_S4H_ASCS Filesystem device='s4h-nfs:/S4H/ASCS' directory='/usr/sap/S4H/ASCS00' fstype='nfs4' \
    op start timeout=60s interval=0 \
    op stop timeout=60s interval=0 \
    op monitor interval=20s timeout=40s
sudo crm configure primitive vip_S4H_ASCS IPaddr2 \
    params ip=10.1.0.5 cidr_netmask=24 \
    op monitor interval=10 timeout=20
sudo crm configure primitive nc_S4H_ASCS azure-lb port=62000
sudo crm configure group g-S4H_ASCS fs_S4H_ASCS nc_S4H_ASCS vip_S4H_ASCS \
    meta resource-stickiness=3000
```

Nun können wir die ASCS Installation starten. Starten Sie den SWPM mit der „SAPINST\_USE\_HOSTNAME“ Option unter Angabe des virtuellen Hostnamens „s4h-ascs“.

```
sudo <SWPM Verzeichnis> sapinst SAPINST_USE_HOSTNAME=s4h-ascs
# Wählen Sie im SWPM die Hochverfügbarkeitsoption und anschließend
ASCS
```

#### 7.4.4.3 Installation ERS (ENSA2) Cluster

[ERS] Erstellen Sie nun die Clusterkonfiguration für die ERS Instanz und installieren Sie dort ebenfalls den SUSE Cloud Connector.

```
sudo zypper install sap-suse-cluster-connector
sudo crm node online s4h-vc-ascsl
sudo crm node standby s4h-vc-ascso
sudo crm configure primitive fs_S4H_ERS Filesystem device='s4h-nfs:/S4H/ASCSERS' directory='/usr/sap/S4H/ERS02' fstype='nfs4' \
    op start timeout=60s interval=0 \
    op stop timeout=60s interval=0 \
    op monitor interval=20s timeout=40s
sudo crm configure primitive vip_S4H_ERS IPaddr2 \
    params ip=10.1.0.6 cidr_netmask=24 \
    op monitor interval=10 timeout=20
sudo crm configure primitive nc_S4H_ERS azure-lb port=62102
sudo crm configure group g-S4H_ERS fs_S4H_ERS nc_S4H_ERS vip_S4H_ERS
```

Installieren Sie nun die ERS Instanz mit dem SWPM unter Verwendung des virtuellen Hostnames „**s4h-ers**“.

```
sudo <SWPM Verzeichnis>/sapinst SAPINST_USE_HOSTNAME=s4h-ers
# Wählen Sie im SWPM die Hochverfügbarkeitsoption und anschließend
```

#### 7.4.4.4 Abschließende Konfiguration

[ASCS0] Ändern Sie die Einstellungen im Profil/sapmnt/S4H/profile/S4H\_ASCS00\_s4h-asc der ASCS Instanz.

```
# Ersetzen Sie „Restart_Program_01 = local ${_EN} pf=${_PF}“ durch
Start_Program_01 = local ${_EN} pf=${_PF}
# Fügen Sie außerdem die folgenden Einträge hinzu
service/halib = ${DIR_CT_RUN}/saphascriptco.so
service/halib_cluster_connector = /usr/bin/sap_suse_cluster_connector
```

[A] Setzen Sie auf beiden Knoten die Keepalive Einstellungen gemäß SAP Note 1410736 und fügen Sie s4hadm zur haclient Gruppe hinzu.

```
sysctl -w net.ipv4.tcp_keepalive_time = 300
sysctl -w net.ipv4.tcp_keepalive_intvl = 75
sysctl -w net.ipv4.tcp_keepalive_probes = 9
sudo usermod -aG haclient s4hadm
```

Stellen Sie sicher das in der/usr/sap/sapservices Datei auf beiden Knoten sowohl ein Eintrag für die ASCS als auch für die ERS enthalten sind.

```
cat /usr/sap/sapservices | grep ASCS00 | sudo ssh s4h-vc-ascs1 "cat >>/usr/sap/sapservices"
sudo ssh s4h-vc-ascs1 "cat/usr/sap/sapservices" | grep ERS02 | sudo tee -a/usr/sap/sapservices
```

Wir fügen abschließend die Pacemaker Konfiguration für ENSA2 hinzu.

```
sudo crm configure property maintenance-mode="true"
sudo crm configure primitive rsc_sap_S4H_ASCS00 SAPIstance \
operations \${id=rsc_sap_S4H_ASCS00-operations \
op monitor interval=11 timeout=60 on-fail=restart \
params InstanceName=S4H_ASCS00_s4h-ascs START_PROFILE="/sapmnt/
S4H/profile/S4h_ASCS00_s4h-ascs" \
AUTOMATIC_RECOVER=false \
meta resource-stickiness=5000

sudo crm configure primitive rsc_sap_S4H_ERS02 SAPIstance \
operations \${id=rsc_sap_S4H_ERS02-operations \
op monitor interval=11 timeout=60 on-fail=restart \
params InstanceName=S4H_ERS02_s4h-ers START_PROFILE="/sapmnt/S4H/
profile/S4H_ERS02_s4h-ers" AUTOMATIC_RECOVER=false IS_ERS=true

sudo crm configure modgroup g-S4h_ASCS add rsc_sap_S4H_ASCS00
sudo crm configure modgroup g-S4H_ERS add rsc_sap_S4H_ERS02
sudo crm configure colocation col_sap_S4H_no_both -5000: g-S4H_ERS
g-S4H_ASCS
sudo crm configure order ord_sap_S4H_first_start_ascs Optional:
rsc_sap_S4H_ASCS00:start rsc_sap_S4H_ERS02:stop symmetrical=false
sudo crm node online s4h-vc-ascs0
sudo crm configure property maintenance-mode="false"
```

#### 7.4.5 Installation PAS und AAS

Passen Sie zunächst ihre/etc/hosts auf allen Applikationsserver Instanzen an. Verwenden Sie für „s4h-nfs“, „s4h-ascs“, „s4h-ers“ und „s4h-db“ die Frontend-IP des jeweiligen Lastenausgleichs.

```
10.1.3.2 s4h-nfs
10.1.0.5 s4h-ascs
10.1.0.6 s4h-ers
10.0.0.2 s4h-db
10.1.0.3 s4h-vc-as0
10.1.0.4 s4h-vc-as1
```

Binden Sie anschließend die NFS Verzeichnisse/sapmnt/S4h und/usr/sap/trans ein. Gehen Sie dazu wie in Kapitel 7.4.4.1.1 beschrieben vor.

Installieren Sie nun den Primary Application Server auf „**s4h-vc-as0**“ sowie den Additional Application Server auf „**s4h-vc-as1**“. Sollten Sie einen embedded Fiori Frontend Server verwenden können Sie den Lastenausgleich „**s4h-lb-as**“ als virtuellen Hostnamen verwenden.

```
sudo <SWPM Verzeichnis>/sapinst SAPINST_USE_HOSTNAME=virtual_hostname
```

Damit ist die Installation der Beispielarchitektur abgeschlossen. Bitte beachten Sie, dass für ein produktives Setup das Testen aller HA-Komponenten unerlässlich ist. Mehr Informationen dazu finden Sie in der offiziellen Microsoft Dokumentation [10] und den SUSE ER2 HA Setup Guide [11].

---

## 7.5 Exkurs Automatisierte SAP Bereitstellung

Bisher haben wir unsere Systeme größtenteils manuell bereitgestellt und konfiguriert. Vorlagenbasierte Bereitstellungen auf Basis von ARM-Templates bieten zwar ein Mindestmaß an Automatisierung, jedoch deckt dies lediglich die Ressourcenbereitstellung ab. In diesem Kapitel werden wir einen kurzen Blick auf weitere Automatisierungsmöglichkeiten im Kontext des Azure Cloud Deployments werfen.

### 7.5.1 Ansible und Terraform in Azure

Ansible und Terraform sind weit verbreitete Cloud Automatisierungswerzeuge deren Zusammenspiel es erlaubt eine „Fullstack-Bereitstellung“ (Ressourcen+Anwendung) durchzuführen. Ein besonderer Vorteil hierbei ist, dass diese Form der Automatisierung Cloud-agnostisch ist, d. h. Sie kann für alle Hyperscaler ohne Anpassungen wiederverwendet werden. Außerdem bietet diese Form der Bereitstellung das höchste Maß an Benutzerkonfiguration.

Die Aufgabe des **Terraform Moduls** ist es die eigentlichen Infrastrukturkomponenten wie z. B. virtuelle Netzwerke, VMs und Datenträger ähnliche einer ARM-Vorlage bereitzustellen. Nach erfolgreicher Bereitstellung ruft das Modul zudem das jeweilige **Ansible Playbook** auf welches wiederum verschiedene **Ansible Roles** nutzt um Software Installationen sowie Konfigurationen durchzuführen. SAP-Systeme können damit mit nur wenigen Kommandozeilen bereitgestellt werden. Terraform Module sind in HCL oder optional in JSON geschrieben. Für Azure bietet Microsoft ein Git Repository [12] welches zahlreiche Bereitstellungsszenarien für SAP abdeckt. Klonen Sie dazu das Git Repository mittels:

```
git clone https://github.com/Azure/sap-hana.git
```

Für die Erstellung der Ressourcen ist ein Service Principal notwendig. Legen Sie diesen zunächst an [13] und erstellen Sie außerdem eine „Key Vault Ressource“. Folgen Sie anschließen den Instruktionen zum Bereitstellen eines Beispielsystems [14].

### 7.5.2 SAP Cloud Appliance Library

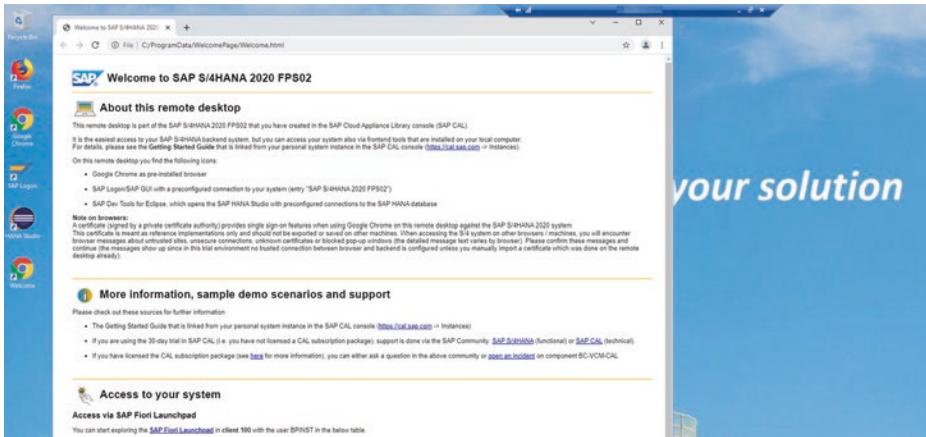
SAP bietet mit seiner SAP Cloud Appliance Library (SAP CAL) ein Repository mit vorkonfigurierten SAP Lösungen welche direkt auf GCP, AWS und natürlich auch MS Azure bereitgestellt werden kann. Eine große Stärke der SAP CAL ist, dass diese u. A. vollkonfigurierte Appliance Lösungen beinhaltet welche hervorragend geeignet sind um die neuesten SAP Innovationen zu evaluieren. Zudem gibt es einige Lösungen welche CAL-exklusiv sind und nicht im Rahmen des klassischen SAP Download Portals verfügbar sind. Ein Beispiel für eine solche Lösung ist die SAP Model Company (MC). Hierbei handelt es sich um maßgeschneiderte Branchenlösungen z. B. „Model Company Core Retail“ in welche SAP ihre über 20-jährige Branchenerfahrung einfließen hat lassen. Diese Systeme fungieren somit als eine Art Sammlung von branchenspezifischen best-practice Prozessen welche Kunden auf Basis der MC evaluieren und für ihr Produktivsystem adaptieren können.

Die Bereitstellung einer Lösung mittels CAL erfolgt direkt über die SAP Website (Zugriff am 20.12.2021) <https://cal.sap.com/> auf Basis des eigenen S-Users. Nach dem Login finden Sie im Bereich „Lösungen“ eine Liste der möglichen Systeme, sowie deren Verfügbarkeit. Wählen Sie eine Lösung aus klicken Sie auf „Instanz anlegen“. Akzeptieren Sie anschließende die allgemeinen Geschäftsbedingungen. Wählen Sie als Cloud-Anbieter „Microsoft Azure“ aus und geben Sie Ihre Abonnement-ID ein. Belassen Sie den Berechtigungstyp auf „Standard Authorization“.

Klicken Sie nun auf Berechtigen, um die Verbindung zu Ihrem Azure Account herzustellen. Im rechten oberen Bildschirmbereich wird Ihnen nun eine Kostenschätzung angeboten. Sie können jetzt außerdem die Instanzdetails konfigurieren.

Bei der Bereitstellung haben Sie die Möglichkeit zwischen dem **Standardmodus** und dem **erweiterten Modus** zu wechseln. Der Standardmodus benötigt als Eingabewerte lediglich Ihre Abonnement-ID, sowie SID, Region und Kennwort. Alle weiteren Parameter werden mit Standardwerten besetzt, z. B. wird standardmäßig ein neues virtuelles Netzwerk erstellt.

Über den erweiterten Modus haben Sie die Möglichkeit die Bereitstellung weiter zu konfigurieren. So können Sie etwa die Instanz in eine bestehendes (Sub-)Netzwerk bereitstellen, eine öffentliche statische IP-Adresse vergeben, sowie die Standardwerte für VM Größe und Speichergröße ändern. Treffen Sie die entsprechenden Angaben und klicken Sie abschließend auf „Anlegen“.



**Abb. 7.13** SAP CAL Landing Page

Speichern Sie den Private Key in der SAP Cloud Appliance Library oder laden Sie diesen optional herunter. Beachten Sie die Warnung zur Bereitstellung. Diese dauert in etwa 60 min und sie sollten während der Aktivierung keine Verbindungsversuche auf die Instanz versuchen. Sie können den Fortschritt der Bereitstellung in der CAL im Reiter „Instanzen“ überwachen. Sobald der Status auf „Aktiv“ wechselt, steht die Instanz bereit. Sie haben nun die Möglichkeit direkt über die CAL mit dem Button „Verbinden“ auf die Instanz zuzugreifen. Im Falle der Lösung „S/4HANA 2020 FPS 02“ steht beispielsweise eine RDP Verbindung zur Verfügung. Verbinden Sie sich nun auf die Instanz, verwenden Sie als Passwort das im Rahmen der Erstellung angegebene Masterkennwort. Wie in Abb. 7.13 dargestellt erhalten Sie, sobald sie sich auf die Instanz verbunden haben, eine Willkommensnachricht mit weiteren Infos über die bereitgestellte Lösung sowie deren Zugangskennungen.

## 7.6 Zusammenfassung

Dieses Kapitel hat Ihnen die Erstellung eines neuen SAP S/4HANA-Systems auf der MS Azure Cloud gezeigt. Hierzu wurde zunächst die beispielhafte Architektur erläutert, welche es zu implementieren gilt. Um die Provisionierung zu starten, ist es notwendig, zunächst die grundlegende Netzwerkkonfiguration zu erstellen. Danach kann mit der Bereitstellung der Ressourcengruppen und weiterer Komponenten begonnen werden.

Zur Absicherung von produktiven SAP S/4HANA-Systemen werden Hochverfügbarkeitscluster eingesetzt. Hierzu zeigte Ihnen das Kapitel die Schritte, wie solche Cluster mit SAP S/4HANA-Systemen eingerichtet und genutzt werden. Somit kennen Sie alle Schritte, um SAP S/4HANA-Systeme auf der Azure Cloud bereitzustellen.

Das Kapitel hat auch einen Ausblick zu den automatisierten Verfahren, wie Ansible und Terraform, sowie der SAP Cloud Appliance Library gegeben. Ebenso, wie in den Kapitel zu den anderen Hyperscalern, ergibt sich für Sie die Möglichkeit, die Provisionierung zu standardisieren und signifikant zu beschleunigen.

---

## Literatur

1. <https://portal.azure.com/#blade/HubsExtension/BrowseResourceGroups> (Zugriff am 20.12.2021).
2. <https://docs.microsoft.com/en-us/azure/virtual-network/service-tags-overview> (Zugriff am 20.12.2021).
3. <https://docs.microsoft.com/de-de/azure/virtual-network/application-security-groups> (Zugriff am 20.12.2021).
4. <https://docs.microsoft.com/en-us/azure/virtual-machines/workloads/sap/high-availability-guide-standard-load-balancer-outbound-connections> (Zugriff am 20.12.2021).
5. <https://azure.microsoft.com/de-de/resources/templates/?term=sap> (Zugriff am 20.12.2021).
6. <https://docs.microsoft.com/en-us/azure/virtual-machines/windows/cli-ps-findimage>.
7. <https://docs.microsoft.com/de-de/azure/virtual-machines/workloads/sap/high-availability-guide-suse-pacemaker>.
8. <https://docs.microsoft.com/de-de/azure/virtual-machines/workloads/sap/high-availability-guide-rhel-pacemaker>.
9. <https://docs.microsoft.com/en-us/cli/azure/install-azure-cli>
10. <https://docs.microsoft.com/en-us/azure/virtual-machines/workloads/sap/high-availability-guide-suse>
11. [https://documentation.suse.com/sbp/all/pdf/SAP\\_S4HA10\\_SetupGuide-SLE15\\_color\\_en.pdf](https://documentation.suse.com/sbp/all/pdf/SAP_S4HA10_SetupGuide-SLE15_color_en.pdf)
12. <https://github.com/Azure/sap-hana>
13. [https://github.com/Azure/sap-hana/blob/master/documentation/SAP\\_Automation\\_on\\_Azure\\_Process\\_Documentation/readme.md](https://github.com/Azure/sap-hana/blob/master/documentation/SAP_Automation_on_Azure_Process_Documentation/readme.md)
14. [https://github.com/Azure/sap-hana/blob/master/documentation/SAP\\_Automation\\_on\\_Azure/Process\\_Documentation/Getting\\_started\\_with\\_the\\_SAP\\_Deployment\\_Automation\\_cloudshell.md](https://github.com/Azure/sap-hana/blob/master/documentation/SAP_Automation_on_Azure/Process_Documentation/Getting_started_with_the_SAP_Deployment_Automation_cloudshell.md)



# SAP S/4 on Google Cloud – Konzepte und Architekturen

8

## Zusammenfassung

Ziel des Kapitels ist es, einen detaillierten Überblick über die ersten Schritte und Planung bei der Bereitstellung und Konfiguration eines SAP S/4HANA Systems auf Google Cloud zu bekommen. Die Zielgruppe für dieses Kapitel stellen technische SAP Architekten, Cloud Architekten und Berater dar, die nach einer technischen Vertiefung für jedes der Themen suchen. Nach einer kurzen Beschreibung der Historie von Google Cloud und der Partnerschaft mit SAP, werden die Entscheidungsgründe für Google Cloud erläutert. Daraufhin werden die relevanten Google Cloud Services und grundlegenden Konzepte für SAP S/4HANA auf Google Cloud eingeführt. Dabei werden die Architekturkomponenten und Mechanismen für Hochverfügbarkeit, Desaster Recovery und Datenmanagement mithilfe der zuvor beschriebenen Google Cloud Services im Detail erklärt und die einzelnen Entscheidungskriterien aufgezeigt.

## 8.1 Überblick zu Google Cloud

### 8.1.1 Historie von Google Cloud

Google startete die Google Cloud Platform mit dem ersten Service namens App Engine im April 2008 in Preview. Zwei Jahre später wurde Google Cloud Storage als nächster Service angeboten. Die Entwicklungen und Veröffentlichungen von Services verliefen kontinuierlich über die Jahre weiter, mit unter anderem Google Compute Engine in Preview in 2012 und Google Kubernetes als vollständig verwalteter Service auf der Google Cloud in Alpha in 2014. Damals wurde auch Kubernetes als Open Source Projekt von Google auf dem Markt freigegeben.

Viele Google Cloud Services folgten und mit Stand September 2021 bietet Google Cloud mehr als 100 verschiedene Cloud Services und Produkte an, welche von IaaS über PaaS bis zur SaaS Kategorie reichen. Die gesamte Liste kann auf der Website [1] gefunden werden. Google ist bekannt für die vielen Veröffentlichungen und Beiträge in die Open Source Gemeinschaft in den letzten zwei Jahrzehnten. Die Whitepaper von Google beeinflussten Technologien wie Beam, Hadoop und viele weitere Frameworks und Technologien, die heute in der IT als Industriestandards etabliert sind. Tatsächlich hat Google mehr als 2000 Open Source Projekte beigesteuert. Mit zu den größten Projekten auf GitHub gehören Kubernetes und TensorFlow, die beide ebenfalls von Google erfunden wurden.

Analysten wie Forrester und Gartner bestätigen seit mehreren Jahren in Folge, dass Google Cloud mit vielen Services der Cloud Plattform auf dem Markt führend ist, unter anderem mit den folgenden Services:

- Google a Leader in Magic Quadrant for Cloud Infrastructure and Platform Services (Gartner, 2021)
- Google BigQuery a Leader in The Forrester Wave™: Cloud Data Warehouse (Q1 2021)
- Google Cloud a Leader in Magic Quadrant for Cloud Database Management Systems (DBMS) (Gartner, 2020)
- Google Cloud a Leader in The Forrester Wave™: Infrastructure as a Service (IaaS) Platform Native Security (Q4 2020)
- Google (Apigee) a Leader in the Magic Quadrant for Full Life Cycle API Management (Gartner, 2020)
- Google Apigee a Leader in The Forrester Wave™: API Management Solutions (Q3 2020) zum vierten Mal in Folge
- Google Cloud a Leader in The Forrester Wave™: Unstructured Data Security Platforms, (Q2 2021)
- und viele mehr...

### **8.1.2 Zeitliche Entwicklung der Partnerschaft zwischen SAP und Google**

Im März 2017 wurde die große Partnerschaft zwischen SAP und Google offiziell verkündet. SAP HANA war ab diesem Tag generell auf Google Compute Engine VMs verfügbar und von SAP zertifiziert [2]. Außerdem wurde ab diesem Zeitpunkt die SAP HANA Express Edition über den Google Cloud Marketplace angeboten.

Viele weitere gemeinsame Aktivitäten und Erfolge dieser Partnerschaft wurden in den folgenden Jahren verkündet und veröffentlicht, welche die Partnerschaft zwischen SAP und Google stärken und viele hilfreiche Möglichkeiten und Services für Unternehmenskunden bieten. Die Highlights sind folgend chronologisch beschrieben:

- **Juni 2018** – Das neue Produkt SAP Data Custodian war auf der Google Cloud als erste Public Cloud verfügbar [3].
- **Mai 2019** – Um Erweiterungen für SAP C/4HANA zu entwickeln hat SAP das Projekt Kyma veröffentlicht, ebenfalls SAP Cloud Platform Kyma Runtime genannt. Dieses basiert auf Knative und wurde direkt zu Beginn auf der Google Cloud angeboten [4].
- **September 2019** – Google Cloud war der erste Hyperscaler, der 6 TB und 12 TB virtuelle Maschinen zertifiziert von SAP und voll virtualisiert (kein Bare Metal) angeboten hat [5]. Zu dieser Zeit hatten andere Hyperscaler diese Größen nur als Bare Metal Service verfügbar.
- **Juni 2020** – Das allererste “SAP Data Center powered by Google Cloud”, operativ in Frankfurt, Deutschland, wurde verkündet. Es ist ein Service, der nur für SAP nutzbar ist [6]. Außerdem wurden während der SAPPHIRE 2020 ein Co-Innovationsprojekt zwischen Google Cloud Visual Inspection und SAP Digital Manufacturing Cloud mit zwei Referenzkunden präsentiert: Kaeser Kompressoren und AES.
- **Januar 2021** – Offizielle Ankündigung des neuen SAP Programms namens “RISE with SAP” in Partnerschaft mit Google Cloud und den anderen großen Hyperscalern.
- **April 2021** – Alphabet, Google Cloud’s Mutterfirma, migrierte von Oracle ERP auf den gesamten SAP S/4HANA Stack und verkündete den erfolgreichen Go-Live [7]. Die öffentlichen Ankündigungen zu dieser Strategie und diesem Projekt erfolgten schon zuvor in 2019 [8].
- **Juli 2021** – Es erfolgte eine aktualisierte Mitteilung zur “RISE with SAP” Partnerschaft mit SAP, welche seither nur zwischen Google [9] und SAP [10] so detailliert wird. SAP und Google Cloud arbeiten in dieser erweiterten strategischen Partnerschaft zusammen, um die Geschäftsprozess- und Cloudmigrationen ihrer Kunden zu beschleunigen. Dazu sollen Google Cloud’s zuverlässiges, skalierbares und hoch performantes Netzwerk und Echtzeit-Integrationen in Google Cloud’s AI und ML Services dienen. Tiefe Integrationen werden ebenfalls weitergeführt und SAP Lösungen wie SAP Analytics Cloud, SAP Data Warehouse Cloud und SAP Business Technology Platform (BTP) werden auf Google Cloud laufen. Zusammen haben SAP und Google Cloud sehr innovative Projekte etabliert wie beispielsweise SAP HANA Fast Restart [11] mit den Memory Poisoning Recovery [12] Möglichkeiten der Google Cloud für SAP S/4HANA Landschaften (siehe Exkurs Abschn. 8.5.7).

Während dieser wenigen Jahre von 2017 bis 2021 hat Google Cloud eine große weltweite Kundenbasis in allen Hauptindustrien gewonnen und etabliert: Handel, Maschinenbau, Banken, Finanzservices, Gesundheitswesen, Pharma- und Life-Sciences-Industrie und viele mehr. Einige dieser Kunden sind offizielle Referenzen und können auf der Website eingesehen werden, eine Auswahl von SAP auf Google Cloud Kunden im Oktober 2021 ist: Vodafone, Schlumberger, Siemens Energy, Deutsche Börse, Metro, MediaMarktSaturn, PayPal, Otto Group, Kaeser Kompressoren, The Home Depot, McKesson, Carrefour, Cardinal Health, Loblaw, und viele mehr.

### 8.1.3 Entscheidungsgründe für SAP auf Google Cloud

Es gibt viele Gründe weshalb sich Kunden entscheiden sollten Ihre SAP Landschaft auf Google Cloud zu migrieren und zu betreiben. Dazu gehören zusammenfassend fünf Hauptbereiche, die im Folgenden erläutert werden: Innovation, Risiko- und Ausfallminimierung, Flexibilität durch vereinfachte Bereitstellung und Nachhaltigkeit.

#### Innovation

Wie zu Beginn des Kapitels erläutert ist Google bekannt für die Leistungen und Services in den Bereichen von Datenanalysen, Big Data und Maschinellem Lernen, was auch durch die vielen Open Source Projekte und Analystenbewertungen bestätigt wird.

Für SAP Landschaften ist die Modernisierung der Weg in die Zukunft. Vor allem die Maximierung der Erkenntnisse, die Sie aus Ihren SAP-Daten mit KI, ML und Advanced Analytics von Google gewinnen können, ist ein wichtiger zukünftiger Baustein, um wettbewerbsfähig zu bleiben. Integrieren Sie ganz einfach Google's ML-Dienste für beispielsweise Vision, Übersetzung und Text-to-Speech und nutzen Sie gleichzeitig intelligente Entscheidungsfindung, um Prozesse zu automatisieren, Vorhersagen zu treffen und Geschäftsprozesse und -abläufe zu optimieren. Sie können SAP Prozesse mit von Google bereitgestellten Datensätzen, cloud-nativen und container-basierten Services und Erweiterungen und dem API-Management von Apigee erweitern.

#### Risiko- und Ausfallminimierung

Die Risikosenkung und Ausfallminimierung durch Mechanismen für erhöhte granulare Sicherheitsvorkehrungen und Hochverfügbarkeit werden mit Google Cloud Diensten vollstens unterstützt und auf den zukunftsweisenden und modernsten Stand der Technik gebracht. Erhöhen Sie die Sicherheit und Leistung mit dem globalen Premiumnetzwerk von Google, bei dem Ihre Daten nicht über öffentliche Netzwerke übertragen werden und standardmäßig im Ruhezustand (at-rest) und während der Übertragung (in-transit) verschlüsselt werden. Google besitzt eines der größten privaten Netzwerke der Welt mit mehr als 100 Points of Presence (PoPs). Dabei funktioniert das Netzwerk wie ein globales Netzwerk und es muss kein VPN oder Peering zwischen den Regionen oder Zonen aufgebaut werden.

Deutlich reduzierte Ausfallzeiten erhalten Sie für Ihre SAP-Anwendungen durch Google's globale Verteilung mit Zonen- und Regionenkonzept sowie der nativen Live-Migration, die hardwareseitige Konfigurationsänderungen ohne Verzögerung und kontinuierliche Hardwarewartung ohne benötigte geplante Ausfälle und Neustarts ermöglicht.

#### Flexibilität und Kostenoptimierung durch vereinfachte Bereitstellung

Nutzen Sie sowohl für traditionelle Projekte als auch für "RISE mit SAP" die Google Cloud Migrationsunterstützung, um doppelte Infrastrukturkosten während der Migration zu minimieren. Stellen Sie OLTP- und OLAP-Umgebungen mit den größten VM-Größen der Branche bereit, die sogar für benutzerdefinierte VM-Konfigurationen

(custom machines) zertifiziert sind. Verwalten Sie einfach Tausende von VMs mit dem VM Manager und optimieren Sie Speicher Kosten und Leistung mit den verschiedenen Speichertypen der Google Cloud. Mit Anthos werden Multi-Cloud- und Hybrid-Cloud-Landschaften vereinfacht und nicht nur deren Flexibilität erhöht, sondern auch deren Sicherheit. Automatisierung als Grundbaustein moderner Cloud-Bereitstellungen werden durch verschiedenste Google Cloud Services ermöglicht und vereinfacht.

Im IDC Business Value Report (August 2020) wurde bestätigt, dass Bereitstellungen von SAP auf Google Cloud zu folgenden Verbesserungen führen:

- 65 % weniger Zeitaufwand für die Bereitstellung/Migration
- 98 % Reduzierung der Bereitstellungszeit
- 56 % effizientere IT Teams durch:
  - 66 % effizientere IT-Infrastruktur-Teams
  - 39 % effizientere Datenbankadministrations-Teams
  - 60 % effizientere IT-Security-Teams

Ferner wurde im Forrester TEI Report (August 2020) bestätigt, dass SAP Upgrades um 35 % effizienter sind und dadurch die Agilität und Flexibilität für das Unternehmen verbessert werden [13].

Eine weitere Studie von Forrester in 2020 bestätigt [14], dass ein Betrieb von SAP auf Google Cloud 160 % Kapitalrentabilität (ROI) generiert und eine Amortisierungszeit von nur sechs Monaten oder weniger benötigt. Das lässt sich aufgrund der Einsparungen von veralteten Infrastrukturkosten, Verhinderung von Ausfällen (Downtimes) und Performance- und Produktivitätsverbesserungen erklären.

## Nachhaltigkeit

Reduzieren Sie sofort Ihre IT-Emissionen, indem Sie SAP-Anwendungen in Google's effiziente und intelligente Rechenzentren verlagern. Mit datengetriebenen Innovationen und der Nutzung von Google Daten und Services können neue Geschäftsmodelle entwickelt und Ihre Nachhaltigkeitsziele vorangetrieben werden.

Seit 2007 ist Google vollständig klimaneutral ( $\text{CO}_2$ -neutral), in 2017 war Google das erste Unternehmen seiner Größe, das 100 % seines weltweiten jährlichen Stromverbrauchs mit erneuerbarer Energie ausglich. Heute ist Google auf Jahresbasis der weltweit größte gewerbliche Abnehmer von erneuerbarer Energie und eine der "saubersten" Clouds der Branche. Das Ziel von Google ist es, bis 2030 komplett auf kohlenstofffreie Energie umzusteigen und diese überall und rund um die Uhr zu nutzen [15].

Folgende Services von Google Cloud werden unter anderen angeboten, um die Kunden im Umfeld von Nachhaltigkeit zu unterstützen: Asset Inventory & Machine Config Benchmark, Cloud Value Tool  $\text{CO}_2$  Calculator, Data Center Carbon Footprint Estimate, Environmental Insights Explorer und Sustainable Value Chain Assessment.

## 8.2 Google Cloud Organisationen und Ressourcen

In diesem Unterkapitel wird der Aufbau der Google Cloud und der Services erläutert. Dabei spielen Ressourcen die grundlegende Rolle. Zuerst wird die Ressourcenhierarchie beschrieben und daraufhin die Eigenschaften von Ressourcen.

### 8.2.1 Google Cloud Ressourcenhierarchie

Google Cloud Organisationen sind der Startpunkt für alle Aktivitäten, die auf der Google Cloud ausgeführt werden, und werden in diesem Kapitel beschrieben. Alle Ressourcen in der Google Cloud hängen in einer Ressourcenhierarchie, welche aus dem oberen Knoten der **Organisation** als Stammknoten, daraufhin einer oder mehrerer Ebenen **Ordnern** und darunter den **Projekten** besteht. Die **Ressourcen** von Google Cloud Diensten hängen direkt an Projekten und machen somit die unterste Ebene der Hierarchie aus. Die Vorteile solch einer Hierarchie sind, dass es eine Zugehörigkeit und Bindung zwischen den Elementen und ihren übergeordneten Elementen gibt und sie somit auch mit dem Lebenszyklus der jeweils oberen Hierarchie zusammenhängen. Dadurch erben die Ebenen und Ressourcen alle Zugriffssteuerungen und Organisationsrichtlinien von ihrer oberen Ebene, also von oben nach unten.

Der Organisationsknoten repräsentiert eine Organisation, beispielsweise ein Unternehmen, und ist der oberste Knoten. Dieser besitzt eine Organisations-ID und ist direkt verknüpft mit dem Cloud Identity Account. Dieser Account kann beispielsweise die E-Mail-Adresse des IT-Leiters sein. Es kann nur eine Organisations-ID für einen Cloud Identity Account geben. Der Organisationsknoten enthält Ordner-Ressourcen und/oder Projekt-Ressourcen.

Vorteile der Organisationsressource:

- Ressourcen (wie Projekte) gehören zu einer Organisation, wenn der/die Ersteller(in) der Ressource das Unternehmen verlässt, wird die Ressource nicht gelöscht.
- Administratoren der Organisation haben Zugriff auf alle Ressourcen dieser Organisation.
- Wird eine Rolle auf der Organisationsebene vergeben, so erben alle Ressourcen, die darunter hängen, diese Rolle. Dies vereinfacht das gesamte Zugriffsmanagement.

Die Ordner-Ressource bietet als Ebene einen Gruppierungsmechanismus und kann somit verschiedene Projekte voneinander isolieren. Ordner können als Sub-Organisationen innerhalb der Organisation angesehen werden, also ähnlich wie Unternehmensbereiche, Abteilungen, Fachbereiche und/oder Teams. Ein Ordner kann sowohl weitere Ordner-ebenen oder auch direkt Projekte enthalten.

Projektressourcen sind die unterste Ebene der Hierarchie und enthalten die Google Cloud Service Ressourcen. Ein Projekt wird immer benötigt, um Google Cloud Services

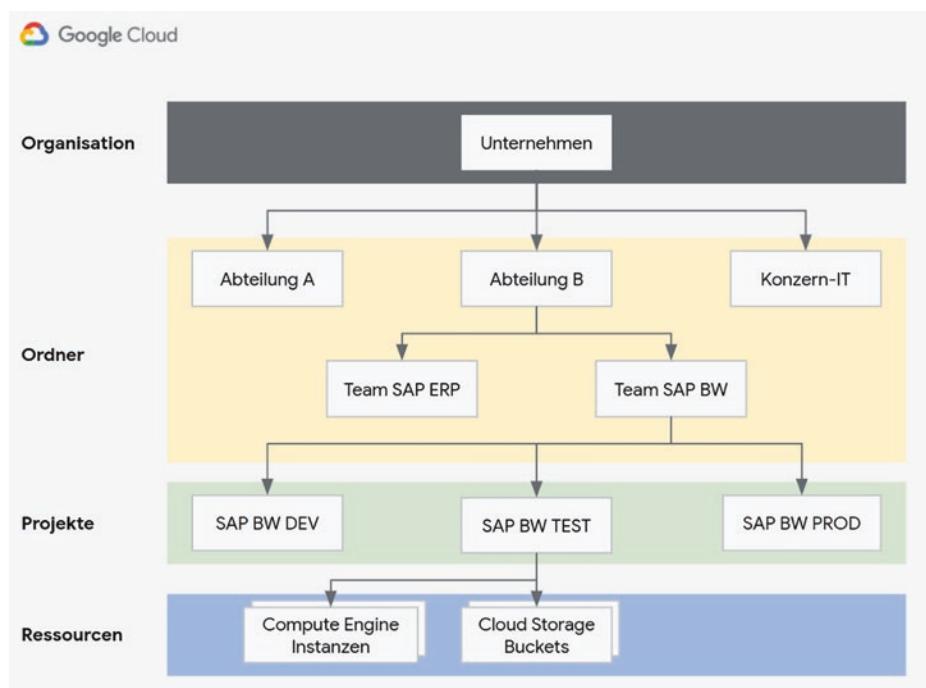
zu nutzen und Ressourcen erstellen zu können. Projekte sind voneinander isoliert. Die Netzwerkkommunikation zwischen ihnen findet standardmäßig nicht statt, kann aber ermöglicht werden (siehe Abschn. 8.3.3).

Alle Projekte enthalten die folgenden Eigenschaften [16]:

- Zwei Identifizierer (Kennzeichen):
  - Projekt-ID, welche eine eindeutige Kennzeichnung des Projektes ist
  - Projektnummer, welche automatisch zugewiesen wird, wenn das Projekt erstellt wird und nicht geändert werden kann (schreibgeschützt)
- Ein Anzeigename, der änderbar ist
- Einen Lebenszyklusstatus des Projektes, bspw. *ACTIVE* oder *DELETE\_REQUESTED*
- Eine Sammlung von Labels, die zum Filtern von Projekten verwendet werden können
- Erstellungszeit des Projektes

Zur Abrechnung der Google Cloud Ressourcen werden Cloud Billing-Konten benötigt, diese sind mit einem oder mehreren Projekten verknüpft.

Die folgende Grafik veranschaulicht eine beispielhafte Ressourcenhierarchie eines Unternehmens (Abb. 8.1).



**Abb. 8.1** Google Cloud Ressourcenhierarchie

Die empfohlenen “Best Practices” zur Erstellung von Ressourcenhierarchien sollten vorab auf der Google Cloud Website [17] eingesehen werden.

### 8.2.2 Eigenschaften der Google Cloud Ressourcen

Ressourcen können in verschiedenen Google Cloud Regionen angelegt werden. Eine **Region** ist ein unabhängiger geografischer Standort und besteht aus mindestens drei verschiedenen Zonen. Ein Beispiel für eine Region ist *europe-west4* (Google Cloud Region in den Niederlanden). Mit Stand September 2021 gibt es 27 weltweite Google Cloud Regionen, die stetig erweitert werden [18].

Eine **Zone** ist ein Bereitstellungsbereich für Ressourcen in einer Region. Zonen sind “Single Failure Domains”. Eine Zone in der Region *europe-west4* ist beispielsweise *europe-west4-a*. Um ausfallsichere Applikationen und Systeme mit Hochverfügbarkeitskonzept aufzusetzen, sollten die Ressourcen und Applikationen über verschiedene Zonen verteilt werden. Um abgesichert für den Ausfall einer gesamten Region aufgrund von Desaster- und Katastrophenfällen zu sein, sollte ein Desaster Recovery Plan erstellt werden (siehe Abschn. 8.5.4).

Ressourcen in Google Cloud können zonal, regional, multi-regional und sogar global sein und werden wie folgt definiert [19]:

- **Globale** Ressourcen sind global repliziert, Beispiele sind vorkonfigurierte Festplatten-Images, Festplatten Snapshots und Netzwerke, aber auch der Load-Balancing Service und die globale Cloud Management Oberfläche (Google Cloud Console).
- **Regionale** Ressourcen sind redundant über verschiedene Zonen einer Region bereitgestellt, zum Beispiel App Engine Applikationen, regionale Managed-Instance-Groups oder statische externe IP-Adressen.
- **Zonale** Ressourcen laufen in nur einer Zone, das bedeutet zonale Ausfälle können manche oder alle Ressourcen in dieser Zone betreffen. Zonale Ressourcen sind beispielsweise Compute Engine – Virtuelle Maschinen (VMs) und ihre Festplatten.
- **Multi-regionale** Ressourcen sind redundant und verteilt in den Regionen und über mehrere Regionen hinweg, um Verfügbarkeit, Performance und Ressourceneffizienz zu optimieren. Beispiele dafür sind der Google Cloud Storage, Cloud Key Management Service und BigQuery.

---

### 8.3 Relevante Google Cloud Services für SAP S/4HANA

Dieser Abschnitt gibt eine detaillierte Übersicht zu den möglichen Services der Google Cloud, die in einer SAP S/4HANA Bereitstellung relevant sind und vor einer Provisionierung im Detail bekannt sein sollten. Das Kapitel gibt ein Verständnis darüber, welche Komponenten für SAP S/4HANA Systeme benötigt werden und wie diese

zusammen agieren. Alle beschriebenen Services sind die Basis für die später folgenden Kapitel zur Architektur und Bereitstellung.

### 8.3.1 Google Cloud Compute Engine

Compute Engine ist Google's flexibler Compute Service mit welchem Virtuelle Maschinen (VMs) auf Google's Infrastruktur in den verschiedenen Regionen erstellt und verwaltet werden können.

Mit Compute Engine ist es möglich vordefinierte Maschinentypen mit einer vordefinierten Anzahl an vCPU (virtual central processing unit) und RAM (random access memory) zu nutzen. Es ist jedoch auch möglich benutzerdefinierte Maschinentypen anzulegen, die aus einer selbst gewählten Anzahl von vCPU und RAM bestehen. Außerdem gibt es sogenannte präemptive Maschinentypen (Instanzen auf Abruf), welche zu einem geringeren Preis als die normalen Instanzen genutzt werden können. Diese können jedoch zu jeder Zeit von Compute Engine heruntergefahren werden und eignen sich somit nur für Batch-Jobs oder fehlertolerante Applikationen. Sie werden daher nicht für SAP-Systeme empfohlen und sind auch nicht SAP-zertifiziert.

Die Compute Engine Maschinentypen werden in verschiedene Gruppen, den Maschinenfamilien, zusammengefasst, wobei nicht alle davon SAP-zertifiziert sind (Stand September 2021):

- Maschinenfamilien für allgemeine Zwecke – **E2, N2, N2D, N1** – SAP-zertifiziert (außer E2)
- Arbeitsspeicheroptimierter Maschinentyp – **M1, M2** – SAP-zertifiziert
- Computing-optimierter Maschinentyp – **C2** – zertifiziert für SAP Anwendungen
- Horizontal skalierbarer Maschinentyp – **T2D** – nicht SAP-zertifiziert
- Beschleunigungsoptimierter Maschinentyp – **A2** – nicht SAP-zertifiziert
- Bare Metal Maschinentyp – **O2** – SAP-zertifiziert

Alle Größen bis zu 12 TB RAM sämtlicher Maschinentypen außer die o2-Instanzen sind komplett virtualisiert. Die zertifizierten Maschinentypen für SAP HANA basieren auf der Intel CPU Platform, manche sind Intel Broadwell und Skylake, die meisten jedoch auf Basis von Intel Cascade Lake (n2-, m2- und o2-Instanzen). Die zertifizierten Maschinentypen für SAP Applikationen sind ebenfalls basierend auf den zuvor genannten Intel Generationen oder basieren mindestens auf AMD EPYC Rome (n2d-Instanzen).

Die Google Cloud Bare Metal Maschinentypen (o2) sind für spezielle Arbeitslasten und Systeme gedacht, zum einen für extra-große, hochperformante SAP HANA Datenbanken, die größer als 12 TB sind und heutzutage in der Industrie noch nicht als VMs angeboten werden. Als weiterer Anwendungsfall dienen diese Maschinentypen auch für SAP-Systeme, die aufgrund von Lizenzanforderungen und weiteren Gründen auf

nicht-virtualisierten Maschinen bereitgestellt werden müssen. Diese o2-Maschinentypen basieren auf Intel Cascade Lake und werden bis 24 TB RAM angeboten (Stand September 2021). Es handelt sich hierbei um einen komplett verwalteten Dienst mit Hardware, Speicher und Netzwerk, der in den Google Cloud Support und das Cloud Billing integriert ist und ein SLA garantiert. Unterschiede zu den Compute Engine VMs sind ferner, dass die Maschinen dediziert (single-tenant) sind und in einer Co-Location mit sehr geringer Latenz in die Google Cloud Rechenzentren angebunden werden [20].

In einer SAP S/4HANA Bereitstellung werden die SAP Datenbank, die SAP Applikationsserver, der SAP Web Dispatcher und weitere Jump Hosts auf Google Compute Engine installiert und bereitgestellt. Die SAP-zertifizierten Maschinentypen sind dokumentiert auf der Google Cloud Website unter den [Zertifizierungen für SAP HANA](#) [21] und den [Zertifizierungen für SAP Anwendungen](#) [22] und außerdem auch im SAP HANA Hardware Directory und der folgenden SAP Note zu finden:

- [SAP HANA Hardware Directory](#) [23]
- [2456432](#) – SAP Applications on Google Cloud: Supported Products and GCP VM types [24]

Google Cloud ist mit aktuellem Stand der einzige Hyperscaler, der benutzerdefinierte Maschinentypen zertifiziert von SAP anbietet. Diese unterliegen den in der folgenden Tabelle formulierten Größenregeln und Konditionen, um von SAP supported zu werden (siehe auch SAP Note im vorherigen Abschnitt und Zertifizierungen für SAP Anwendungen) (Tab. 8.1).

**Tab. 8.1** Größenregelung für benutzerdefinierte Compute Engine Maschinentypen

Maschinentyp	vCPU Anforderung	Option mit Standard-Speicher	Option mit großem Speicher
n1	1 oder eine gerade Zahl bis 96	3,75 GB pro vCPU	6,5 GB pro vCPU Bis zu insgesamt 624 GB pro VM
n2	Jede beliebige gerade Zahl bis 32 Nach 32 bis zu 80 vCPU muss die Zahl teilbar durch 4 sein	4 GB oder mehr pro vCPU	8 GB oder mehr pro vCPU Bis zu maximal 640 GB pro VM
n2d	2 bis maximal 96 vCPU in folgenden Inkrementen: 2, 4, 8, oder 16 vCPUs. Nach 16 vCPUs muss im 16 Inkrement gesteigert werden	Zwischen 1 GB und 16 GB pro vCPU	Bis zu 768 GB pro VM

- Benutzerdefinierte Maschinentypen müssen zuerst von SAP über ein SAP Support Ticket analysiert werden, bevor SAP diese für Ihre Systeme unterstützt [25].

Die in Compute Engine eingebaute, auf maschinellem Lernen basierende Funktionalität “Rightsizing Recommendation” bietet automatisiert Größenempfehlungen in der Google Cloud Console zu Maschinentypen. Diese Empfehlungen basieren auf der Auslastung, dem Nutzungsverhalten und -statistiken der Maschinen. Nutzer können dadurch die VM-Instanzen anpassen und Ressourcen der Instanzen effizienter und oftmals kostengünstiger nutzen.

### **Business Continuity Funktionalitäten**

Ein einzigartiger Mechanismus im Vergleich der Hyperscaler ist Google’s Live-Migration Funktion. Außerdem bietet Compute Engine die Möglichkeit für automatischen Neustart von Instanzen. Beide Mechanismen sind für Hochverfügbarkeit relevant und werden in Abschn. 8.5.3 erläutert.

### **Service Level Agreements (SLA) für Compute Engine**

Wenn Instanzen in einem Hochverfügbarkeitssetup, das bedeutet über mindestens zwei Zonen einer Region, aufgesetzt wurden und durch einen Load Balancer unterstützt werden, dann bietet Compute Engine eine monatliche Betriebszeit von  $\geq 99,99\%$ . Der Load-Balancing Service bietet ebenfalls eine SLA von  $\geq 99,99\%$ . Das Standard-SLA für einzelne Google Compute Engine Instanzen ist  $\geq 99,5\%$ , das bedeutet die Instanz ist in einer einzigen Zone ohne Hochverfügbarkeit aufgesetzt (Stand September 2021) [26].

### **Betriebssysteme auf Compute Engine**

Fast alle der zuvor genannten Komponenten einer SAP S/4HANA Architektur benötigen ein installiertes Betriebssystem. Zertifiziert für SAP HANA sind SUSE Linux Enterprise Server (SLES) und Red Hat Enterprise Linux (RHEL). Sie können auf Compute Engine das Betriebssystem entweder mit einem eigenen Betriebssystem-Image (Bring-your-own-image, BYOI) und einer eigenen Lizenz (Bring-your-own-license, BYOL) verwenden oder direkt über die Google Cloud Images in Compute Engine nutzen und abrechnen. Die Betriebssystem-Images von Google Cloud bieten Hochverfügbarkeitsfunktionen spezifisch konfiguriert für die Google Cloud. Empfohlen wird zuvor die Dokumentation sowohl für SAP HANA [27] als auch für SAP NetWeaver [28] zu prüfen.

Die folgenden Betriebssysteme werden für SAP HANA von Google Cloud Compute Engine und von SAP supported (September 2021):

- RHEL 8: 8.1 für SAP und ausstehend: 8.2, 8.4
- RHEL 7: 7.4, 7.6 und 7.7 für SAP
- SLES 15: SP1, SP2 für SAP und ausstehend: SP3
- SLES 12: SP3, SP4, SP5 für SAP

Siehe hierzu die SAP-Dokumentation in SAP Note:

- SAP HANA Supported OS – [2235581 \[29\]](#)

Die folgenden Betriebssysteme werden für SAP NetWeaver von Google Cloud Compute Engine und von SAP supported (September 2021):

- RHEL 8: 8.1, 8.2 und 8.4 für SAP
- RHEL 7: 7.4, 7.5, 7.6 und 7.7 für SAP
- SLES 15: SP1, SP2 und SP3 für SAP
- SLES 12: SP3, SP4 und SP5 für SAP
- Windows Server 2019, 2016 und 2012 R2

Siehe SAP Dokumentation und SAP Note:

- SAP Applications on Google Cloud Platform: Supported Products and Google VM types – [2456432 \[30\]](#)
- [Product Availability Matrix \[31\]](#)

Neben SAP HANA sind weitere zertifizierte Datenbanken für SAP Landschaften ebenfalls verfügbar und supported auf Google Cloud (siehe SAP Note wie oben – 2456432), unter anderem:

- IBM DB2
- Microsoft SQL Server
- SAP ASE
- SAP MaxDB

### **8.3.2 Speicheroptionen**

SAP-Systeme benötigen verschiedene Typen von Speicher: Festplatten, verteilte Dateisysteme und Speicher für die Sicherungen. Die im Folgenden beschriebenen Google Cloud Dienste decken dabei alle benötigten Komponenten ab.

#### **8.3.2.1 Google Cloud Blockspeicher**

Nichtflüchtiger Speicher (Laufwerke, Persistent Disks) ist ein dauerhafter und hochperformanter Blockspeicher für Compute Engine Instanzen. Dieser Speicher entspricht in On-Premise Landschaften den Festplatten. Dieser ist in drei Klassen verfügbar, und zwar Standard-, SSD- und Extrem-Format, wobei der SSD-Typ aufgeteilt wird in SSD und ausgeglichener (balanced) Blockspeicher. Die drei Klassen unterscheiden sich in ihren Leistungsmerkmalen [\[32\]](#) vs. den Preisen.

- Nichtflüchtiger Standardspeicher (pd-hdd): Speicher mit niedrigen Kosten
- SSD-basierte Speicher (pd-ssd und pd-balanced): Schneller Blockspeicher mit hohen IOPS und geringer Latenz
- Extrem nichtflüchtiger Speicher: für High-End-Datenbankarbeitslasten mit höchstem Durchsatz

Der Blockspeicher kann entweder zonal oder regional erstellt werden, wobei zonale Laufwerke nur in einer Zone verfügbar sind und regionale Laufwerke synchron über zwei Zonen einer Region repliziert werden. Dadurch bieten regionale nichtflüchtige Speicher ein sehr geringes Risiko eines Datenverlustes und eine hohe Verfügbarkeit. Sie können zur Notfallwiederherstellung verwendet werden. Andererseits bieten zonale Laufwerke höhere Performance.

Außerdem gibt es lokale SSDs, die physisch mit der Hardware verbunden sind, auf der die VM-Instanz läuft. Diese bieten zwar einen höheren Durchsatz und geringere Latenz, jedoch gibt es keine Garantie, dass die VM-Instanz immer auf derselben Hardware läuft. Daher kann die Verbindung der lokalen SSD-Platte an die VM gelöscht werden, wenn die VM auf eine andere Hardware migriert wird. Das kann der Fall sein, wenn die VM-Instanz neu gestartet, oder eine Live-Migration durchgeführt wird. Lokale SSDs sollten daher also nur für temporäre Daten verwendet werden, und sind für SAP-Systeme kaum relevant.

Von den Laufwerken können Snapshots und aus diesen wiederum neue Laufwerke erstellt werden. Wenn eine Instanz terminiert wird, behält der nichtflüchtige Speicher die Daten und kann an eine andere Instanz zugewiesen werden. Mehr zu Snapshots im Umfeld von Sicherungen wird in Abschn. 8.5.5 erläutert.

Auf Laufwerken werden die unterschiedlichen Dateiverzeichnisse und Unterordner des SAP-Systems gespeichert. Die folgende Tabelle gibt Empfehlungen zu den Speichertypen pro SAP Dateiverzeichnistyp [33] (Tab. 8.2):

**Tab. 8.2** Empfehlungen zu den Speichertypen für SAP Strukturen

SAP HANA Dateiverzeichnis	Empfohlener Speichertyp
/usr/sap	SSD PD
/hana/data	SSD PD, zonal [34]
/hana/log	SSD PD, zonal
/hana/shared	SSD PD, zonal
/hanabackup	HDD Standard PD
ABAP Dateiverzeichnis	Empfohlener Speichertyp
/sapmnt	HDD Standard PD
/usr/sap/	HDD Standard PD
Boot and exe files	HDD Standard PD oder Balanced PD

- Für SAP HANA sind derzeit für die Verzeichnisse/hana/data und/hana/log nur SSD-Festplatten zertifiziert.

Die Performanz von SSD-basierten Laufwerken erhöht sich je größer der Speicher und je größer die Anzahl an vCPUs sind. Google Cloud gibt eine Empfehlung zur Mindestplattengröße pro Kategorie und pro zertifizierten Maschinentypen für SAP HANA in der Dokumentation [35]. Dabei werden die folgenden Formeln für SAP HANA Verzeichnisse empfohlen, wobei der hier genannte “Speicher” der Arbeitsspeicher der Compute Engine Instanz (in GB) ist. Die folgenden Regeln gelten für die vertikale und nicht die horizontale Nutzung:

- /hana/data:  $1,2 \times$  Speicher
- /hana/log: entweder  $0,5 \times$  Speicher (jedoch so, dass es ein Vielfaches von 64 ist) oder 512 GB, je nachdem, welcher Wert kleiner ist
- /hana/shared: entweder  $1 \times$  Arbeitsspeicher oder 1.024 GB, je nachdem, welcher Wert kleiner ist
- /usr/sap: 32 GB
- /hanabackup:  $2 \times$  Speicher, optionale Zuordnung

Es wird außerdem empfohlen, weniger und dafür größere Festplatten anzulegen und logisch im Betriebssystem in mehrere Dateisysteme zu partitionieren. Bei SAP HANA wird empfohlen/hana/data, /hana/log, /usr/sap und/hana/shared in ein Compute Engine Laufwerk zu mappen. Die Nutzung von größeren und insgesamt weniger Festplatten erlaubt zudem einfachere Größenanpassungen und vereinfachtes Management und Betrieb sowie eine höhere Performance.

### **8.3.2.2 Dateifreigabelösungen (File Sharing) für SAP auf Google Cloud**

Google Cloud unterstützt verschiedene Dateifreigabelösungen für SAP Bereitstellungen, die Auswahl hängt von den benötigten Anforderungen an Regionen und Zonen sowie der Performanz ab. Die folgenden Lösungen werden derzeit empfohlen (Stand September 2021) [36]:

1. Google Filestore: Google Cloud's leistungsstarker und vollständig verwalteter Dateispeicher
2. NetApp Cloud Volumes Service (CVS) Performance (Standard und Extrem) für Google Cloud: NetApps leistungsstarker und vollständig verwalteter Dateispeicher, welcher direkt in der Google Cloud Console bereitgestellt, konfiguriert und abgerechnet werden kann
3. NetApp Cloud Volumes ONTAP: NetApps leistungsstarker Dateispeicher, welcher vom Kunden auf Compute Engine installiert und verwaltet werden kann

**Tab. 8.3** Vergleich der Dateifreigabe (File Sharing) Lösungen für SAP auf Google Cloud

Funktion	Google Filestore	NetApp CVS Performance	NetApp Cloud Volumes ONTAP
Managed Service	Ja	Ja	Nein
SLA	99,5 %	99,99 %	99,99 % jedoch nur für Google Cloud Compute Engine
Hochverfügbar	Nein	99,99 % SLA	HA-Lösung mit mehreren Zonen
Desaster Recovery	Manuell	Multi-regionale Replikation	Automatisiert mit Snapmirror
Lokale Snapshots	Ja	Ja	Ja
Regionale Replikation	Nein	Ja	Mit Snapmirror
Regionale Verfügbarkeit	Alle Regionen	Siehe NetApps Regionen [37]	Alle Regionen
RPO	N/A	In nur 15 min	In nur 15 min
RTO	N/A	In nur 30 min	In nur 30 min
Protokolle	NFSv3	NFSv3, v4.1, SMB	NFSv3, v4.1, SMB iSCSI
Mindestgröße	1 TB	1 TB	100 GB Volume mit 638 GB Systemlaufwerk
Supportanbieter	Google Cloud	Google Cloud	NetApp für die Software und Google Cloud für die Infrastruktur
Durchsatzleistung	100 MB/s R/W	128 MB/s R/W	Abhängig von der Konfiguration [38]

Die Unterschiede zwischen diesen drei Lösungen sind in der folgenden Tabelle ersichtlich (Tab. 8.3):

In welchem Anwendungsfällen welche der drei Lösungen empfohlen wird, lässt sich unter den folgenden Kategorien ermitteln:

1. Schnittstellenverzeichnisse
  - Allgemeiner Speicherort für SAP und andere Systeme zum Übertragen von Dateien zwischen den Servern
  - NetApp Cloud Volumes ONTAP
  - NetApp CVS-Performance, Extrem
2. SAP-Transportverzeichnis
  - Als Speicherort von SAP-Systemen, um gemeinsam genutzte Anwendungsdateien in verteilten Landschaften oder bei Hochverfügbarkeitssetup zu speichern, oder um

SAP-Dateien und -Updates zwischen den verschiedenen operativen Umgebungen zu übertragen.

- NetApp Cloud Volumes ONTAP
- NetApp CVS-Performance, Standard

### 3. Sicherungsverzeichnis

- Zur Verwendung als zentralen Speicherort für Sicherungen
- NetApp Cloud Volumes ONTAP
- NetApp CVS-Performance, Extrem

### 4. SAP HANA-System mit horizontaler Skalierung (Scale-Out)

- Zur Verwendung in einer einzelnen Zone durch SAP HANA-Systeme mit horizontaler Skalierung, um Dateien zwischen den SAP HANA-Knoten zu teilen.
- Google Filestore
- NetApp Cloud Volumes ONTAP
- NetApp CVS-Performance, Extrem

#### **8.3.2.3 Google Cloud Storage**

Der Objektspeicher von Google Cloud namens Cloud Storage ist ein vollständig verwalteter Service zur Speicherung von Objekten, das bedeutet Dateien eines jeden Formats. Die Objekte werden in Containern gespeichert, die sich Storage Buckets nennen und welche zu einem Projekt gehören. In On-Premise Landschaften entspricht Google Cloud Storage unter anderem der Sicherungsanwendung und -speicher.

Google Cloud Storage besitzt die folgenden Eigenschaften:

- Unlimitierter Speicher ohne Mindestobjektgröße
- Hohe jährliche Speicherlanglebigkeit von 99,99999999 %
- Weltweiter Zugriff und globale Speicherorte
- Geringe Latenz (Zeit bis zum ersten Byte üblicherweise wenige Millisekunden)
- Automatische Georedundanz, wenn der multi-regionale oder dual-regionale Speicher-typ gewählt wird

Google Cloud Storage teilt sich in unterschiedliche Speicherklassen [39], welche Unterschiede in den benötigten Verfügbarkeiten und in den Preisen mit sich bringen. Die vier Klassen sind:

- Standard Storage – häufig genutzte Daten – keine Mindestspeicherzeit
- Nearline Storage – für einmal oder weniger im Monat abgerufene Daten – Mindestspeicherzeit 30 Tage
- Coldline Storage – für weniger genutzte Daten, auf die einmal oder weniger im Quartal (90 Tage) zugegriffen wird
- Archive Storage – für weniger als einmal im Jahr aufgerufene Dateien – Mindestspeicherzeit 365 Tage

Cloud Storage kann lokal (regional), dual-regional oder multi-regional aufgesetzt werden. Die Auswahl hängt von den Anforderungen an Datenspeicherort und Einschränkungen, Latenzen für Sicherungen und die Wiederherstellung, sowie Anforderungen an regionale Ausfallsicherheit ab. Wählen Sie dual- oder multi-regionale Buckets in denselben Regionen oder nahe an den Regionen in welchen die Compute Engine Instanzen (SAP-Systeme) laufen. Die Storage Bucket Standorte können der Dokumentation [40] entnommen werden.

Die Verfügbarkeits-SLAs unterscheiden sich ebenfalls nach Speicherklasse und Standorttyp. Für den Standard Storage sehen die SLAs wie folgt aus:

- Multi-regional – 99,95 % SLA – >99,99 % typische monatliche Verfügbarkeit
- Dual-regional – 99,95 % SLA – >99,99 % typische monatliche Verfügbarkeit
- Regional (lokal) - 99,9 % SLA – 99,99 % typische monatliche Verfügbarkeit

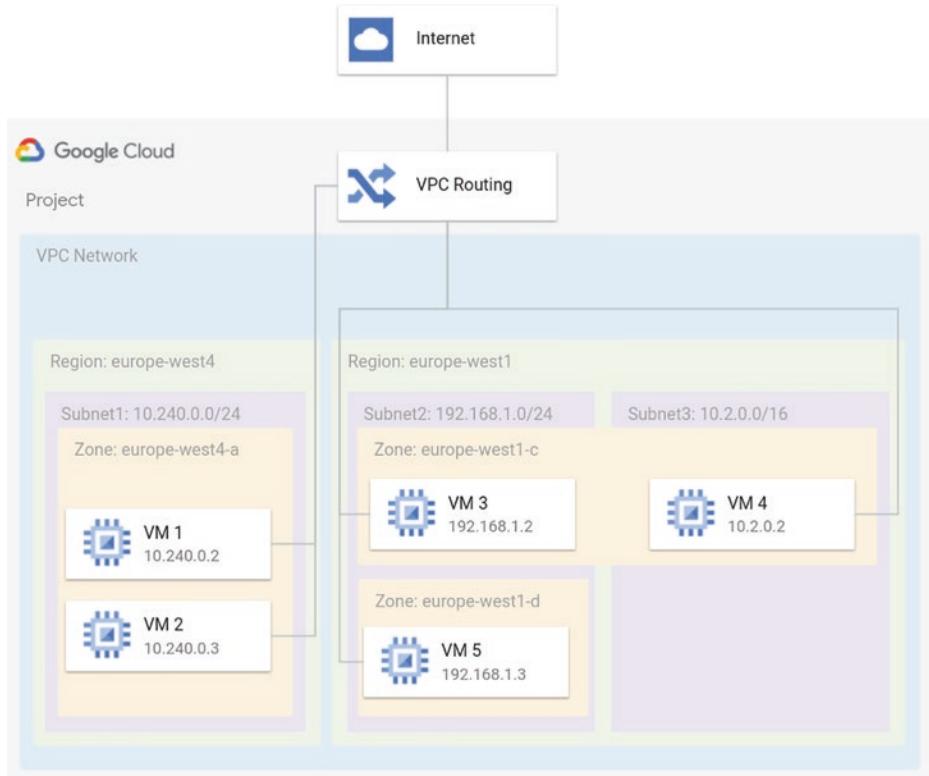
In SAP auf Google Cloud Landschaften ist Cloud Storage vor allem relevant für das Speichern der Festplatten-Snapshots und SAP HANA Sicherungen, denn Cloud Storage ist SAP HANA Backint zertifiziert. Das bedeutet, dass eine SAP HANA Datenbank Konfiguration mit Cloud Storage als Speicherort für die SAP HANA Sicherungen möglich ist. Dazu empfiehlt sich die Dokumentation: Cloud Storage Backint-Agent für SAP HANA [41] und [SAP Note 2031547](#) [42]. Ferner gibt es zu Google Cloud Storage generell noch eine detaillierte Best-Practices Übersicht [43].

### 8.3.3 Google Cloud VPC und Netzwerkkonzept

Dieser Abschnitt gibt einen Überblick auf die Services, die benötigt werden, um in einer SAP auf Google Cloud Bereitstellung alle Netzwerkkomponenten aufzubauen. Das Netzwerkkonzept hat eine große Auswirkung auf Verfügbarkeit, Performanz und Resilienz der zu aufbauenden SAP Landschaft. Die meisten Netzwerkdienste der Google Cloud sind virtuell (Software-defined Networking), was das Design von komplexen Landschaften für Administratoren vereinfacht. Trotzdem können dadurch organisatorische und sicherheitsrechtliche Anforderungen erfüllt werden.

#### 8.3.3.1 Google Cloud VPC Setup

Ein Virtual Private Cloud (VPC) Netzwerk [44] ist eine virtuelle Version eines physischen Netzwerks, und in Google Cloud nur virtuell vorhanden. Dieses basiert auf Google's selbst implementierten und privaten Andromeda Netzwerk. Die VPC verbindet alle Google Cloud Dienste und die verschiedenen Regionen (Standorte) ohne dabei über das Internet zu kommunizieren, sondern verbleibt in den privaten Netzwerkkabeln von Google. VPC-Netzwerke sind globale Ressourcen und bestehen aus regionalen virtuellen Subnetzwerken (Subnets), die über ein globales Wide Area Network (WAN) verbunden sind. Subnetze, die zu einer VPC gehören verbinden die Zonen einer Region



**Abb. 8.2** Google Cloud VPC Konzept

miteinander. Die VPC-Netzwerke sind logisch voneinander getrennt und gehören zu Projekten (Abb. 8.2).

Ein Projekt kann mehrere verschiedene VPCs besitzen und ein VPC-Netzwerk kann auch zwischen mehreren Projekten einer Organisation geteilt werden. Dieses Konzept wird freigegebenes (Shared) VPC genannt. VPCs von verschiedenen Projekten können über VPC-Netzwerk-Peering miteinander verbunden werden.

VPC-Netzwerke bieten außerdem TCP/UDP und HTTP(S) Load-Balancing und verbinden die Google Cloud zu On-Premise Landschaften via Cloud VPN oder Cloud Interconnect (siehe nächster Kapitelabschnitt).

Google Cloud's privates Netzwerk bietet Netzwerkbandbreiten von 32 Gbps und niedrige Latenzen für Instanzen in derselben Zone. Für SAP Landschaften wird empfohlen, die Datenbankinstanzen und Applikationsserver in derselben Zone bereitzustellen, um eine maximale Performance zu erreichen.

Als Best Practices für ein Shared VPC Setup im SAP Umfeld können zwei Szenarien betrachtet werden [45]:

- Bereitstellung der SAP Landschaft in einer einzigen Shared VPC
  - Vereinfachung und Reduzierung von administrativen Mehraufwand
  - Die Shared VPC dient als Netzwerkhub
  - Möglichkeit die Landschaften (bspw. DEV, QA, PROD) mithilfe von Projekten und Subnetzen zu isolieren
  - Netzwerkinspektion: feingranulare Zugriffssteuerung mithilfe der Nutzung von Firewalls, Netzwerktags und Service Accounts (siehe Abschn. 8.3.5)
- Mehrere Shared VPCs für die Bereitstellung der SAP Landschaft
  - Eine Shared VPC pro Landschaft, um die Netzwerkinspektion zu erhöhen
  - Peering wird zwischen den Shared VPCs benötigt, um eine Verbindung aufzubauen
  - Die Shared VPCs sind komplett voneinander isoliert und die Kommunikation ist nur durch das Peering via Firewall und den geöffneten Ports möglich
  - Richtlinien und der administrative Aufwand werden erhöht, da die Einstellungen und Konfigurationen pro Shared VPC erfolgen

### 8.3.3.2 Cloud VPN, Partner oder Dedicated Interconnect

Die Verbindung zwischen dem On-Premise und Google Netzwerk kann über das Internet oder über ein privates Netzwerk hergestellt werden. Da ein privates Netzwerk mehr Sicherheit bietet, kommt für die meisten SAP Kunden nur dieses infrage. Für die Anbindung von On-Premise Unternehmensnetzwerken und -landschaften (Hybrid-Cloud-Konzept) sowie anderen Hyperscalern (Multi-Cloud-Konzept) an die Google Cloud gibt es drei Möglichkeiten: Cloud VPN, Partner Interconnect, Dedicated Interconnect [46]. Diese werden im Folgenden mit ihren Vor- und Nachteilen genauer beschrieben:

#### Cloud VPN [47]

- Verbindet das Peer-Netzwerk an die Google Cloud VPC über eine IPsec-VPN-Verbindung (Tunnel) über das Internet. Die Bandbreite ist durch die Internet-Verbindung und die Anzahl der Tunnel (siehe unten) limitiert. Der Vorteil des Cloud VPN gegenüber einer direkten Internetverbindung ist nicht die Bandbreite, sondern die Verschlüsselung und die Möglichkeit, private IP Adressen zu verwenden.
- Bi-direktonaler Tunnel
- Dynamisches Routing
- Datenverkehr wird verschlüsselt durch das erste VPN Gateway und durch das andere VPN Gateway wieder entschlüsselt
- Jeder Cloud VPN Tunnel unterstützt 1,5 Gbps bis zu 3 Gbps für eine Summe von eingehendem (ingress) und ausgehendem (egress) Datenverkehr
- Es gibt zwei Typen von Cloud VPN Gateways: HA VPN und Klassisches VPN
  - HA (Hochverfügbarkeit) VPN – SLA 99,99 %
  - Klassisches VPN – SLA 99,9 % – Achtung: teilweise werden bestimmte Funktionen am 31.10.2021 eingestellt

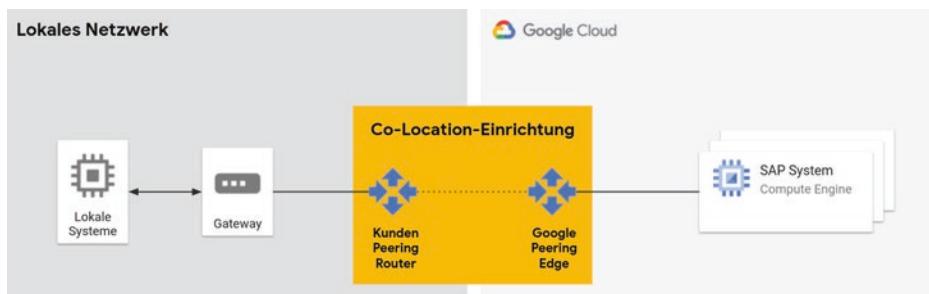
- Vorteile:
  - Leichtes und schnelles Setup
  - Keine weiteren Verträge mit externen Dienstleistern oder Hardware und Konfigurationen benötigt
- Nachteile:
  - Limitierte Bandbreite
  - Tunnel über das Internet
- Empfehlung im SAP-Umfeld nur für Proof-of-Concepts, Piloten oder kurzzeitige Verbindungen

### Dedicated Interconnect [48]

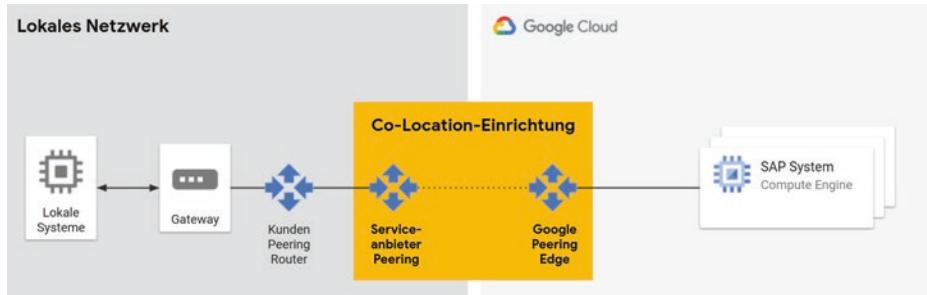
- Eine direkte physische Verbindung zwischen dem lokalen Netzwerk und dem Google-Netzwerk
- Die Verbindung zum Google-Netzwerk wird in einer Co-Location-Einrichtung hergestellt, in welcher der Kunde die eigene Routingausrüstung bereitstellt
- Verbindungen mit 10 Gbps, 100 Gbps oder mehrere dieser Leitungen gebündelt sind möglich
- SLA von 99,9 % und 99,99 % (empfohlen, über Redundanzen) möglich (Abb. 8.3)

### Partner Interconnect [49]

- Ebenfalls eine physische Verbindung zwischen dem lokalen Netzwerk und dem Google-Netzwerk, jedoch durch einen offiziellen und unterstützten Serviceanbieter
- Verbindungen von 50 Mbps bis 10 Gbps möglich
- Empfehlung diese Möglichkeit zu nutzen, falls keine eigene Routingausrüstung in der Co-Location-Einrichtung bereitgestellt werden kann bzw. die Einrichtung nicht erreicht werden kann oder, wenn der Kunde keine 10-Gbps-Verbindung benötigt
- SLA von 99,9 % und 99,99 % (empfohlen, über Redundanzen) möglich (Abb. 8.4)



**Abb. 8.3** Google Cloud Dedicated Interconnect



**Abb. 8.4** Google Cloud Partner Interconnect

### 8.3.3.3 Weitere Netzwerkkomponenten

#### Eingehender (Ingress) und ausgehender (Egress) Datenverkehr

Bei Google Cloud ist eingehender Netzwerkverkehr (Ingress) kostenlos, während ausgehender Netzwerkverkehr (Egress) Kosten verursacht. Diese Kosten sind von der Übertragungsmenge, der Quellservices- und -region, Zielservices und -region und der Art der Verbindung abhängig.

Die unterschiedlichen Verbindungsarten sind:

- Datenverkehr innerhalb von Google Cloud über interne IP Adressen
- Datenverkehr von Google Cloud zu weiteren Google Diensten (wie YouTube und Maps)
- Datenverkehr von Google Cloud in das Internet
- Datenverkehr von Google Cloud in ein Cloud VPN
- Datenverkehr von Google Cloud zu einem (Partner oder Dedicated) Interconnect
- und weitere

Die Dokumentation liefert einen genauen Überblick und weitere Details, beispielsweise die Preise [50].

#### Routen und Weiterleitungsregeln (Forwarding Rules)

Routen legen für die VM-Instanzen und das VPC-Netzwerk fest, wie der Datenverkehr von einer Instanz zum definierten Ziel innerhalb oder außerhalb der Google Cloud fließen soll. Jedes VPC-Netzwerk hat automatisch generierte Routen um den Datenverkehr zwischen den Subnetzen zu definieren und von Instanzen weiterzuleiten, beispielsweise ins Internet.

Im Vergleich zu Routen, die den Datenverkehr definieren, der eine Instanz verlässt, definieren die Weiterleitungsregeln wie Datenverkehr von außen in eine Google Cloud Ressource im VPC-Netzwerk fließen soll (basierend auf IP-Adresse, Protokoll und Port). Weiterleitungsregeln können Datenverkehr entweder von außerhalb oder innerhalb

dieselben Netzwerks zum Ziel leiten. Ziele können Instanzen, Load-Balancer oder Cloud VPN Gateways sein.

### **Firewalls**

Jedes VPC-Netzwerk implementiert eine verteilte virtuelle Firewall, die beliebig konfiguriert werden kann. Firewall Regeln kontrollieren, welche Datenpakete und Kommunikationen zu welchen Destinationen im VPC-Netzwerk erlaubt sind. Jedes VPC-Netzwerk hat zwei direkt automatisch und standardmäßig inbegriffene Firewall Regeln, die erste blockt alle eingehenden Verbindungen und die zweite erlaubt alle ausgehenden Verbindungen.

### **Cloud DNS**

Cloud DNS ist ein vollständig verwalteter, hoch-performerter, resilenter und globaler Domain Name Service (DNS) für die Übersetzung von Domänennamen in IP-Adressen. Die DNS Weiterleitungsregeln können eingehend oder ausgehend sein. Die Best-Practices [51] für Cloud DNS sollten ebenfalls eingesehen werden.

### **Cloud NAT**

In produktiven Landschaften und im Unternehmensumfeld besitzen die VM-Instanzen üblicherweise keine öffentlichen externen IP-Adressen mit einer direkten Verbindung in das Internet aufgrund der Sicherheitsrisiken. Deshalb werden Dienste für Netzwerkadressübersetzung (NAT) genutzt. Mit Cloud NAT bietet Google einen verteilten, Software-definierten und vollständig verwalteten Service mit welchem VM-Instanzen ausgehende Verbindungen in das Internet aufbauen können und darüber auch eingehende Rückmeldungen entgegennehmen können. Eine Alternative wäre eine Compute Engine VM-Instanz manuell als NAT aufzubauen, dann ist es jedoch kein verwalteter Service.

In einer SAP Landschaft werden Internetverbindungen über Cloud NAT aus verschiedenen Gründen genutzt, ein häufiger Grund ist die Registrierung und Aktivierung der Betriebssystemlizenzen bei SUSE oder Red Hat.

### **Cloud Load-Balancing**

Der Google Cloud Load-Balancing Dienst ist ein vollständig verwalteter, Software-definierter, hochperformanter und skalierbarer Load-Balancer. Mit Cloud Load-Balancer können Nutzerzugriffe und Datenverkehr auf mehrere Instanzen der Anwendung automatisch skalierend verteilt werden. Es gibt globales und regionales Load-Balancing, wenn sich die Anwendungen über mehrere Regionen verteilen, so sollte globales und ansonsten das regionale Load-Balancing gewählt werden. Der Cloud Load-Balancer (LB) ist nicht instanz- oder hardwarebasiert, daher müssen sich die Kunden nicht um Hochverfügbarkeit, Skalierung oder Verwaltung kümmern.

Die folgenden Optionen sind auf Google Cloud verfügbar und können je nach Anforderung Ihrer Landschaft anhand eines Entscheidungsbaums in der Dokumentation ausgewählt werden [52]:

- Globales externes Load-Balancing mit HTTP(S), SSL Proxy oder TCP Proxy LB
- Regionales externes Netzwerk LB
- Regionales internes HTTP(S), TCP, UDP LB

Der Google Cloud Load-Balancer ist integriert in den Cloud DNS und garantiert eine SLA von 99,99 %. In SAP auf Google Cloud Landschaften werden Load-Balancer typischerweise eingesetzt, um eine einzelne IP-Adresse für redundante Web Dispatcher darzustellen.

### 8.3.4 Google Cloud Security

Google Cloud ist eine durch das Design sichere Infrastruktur, die auf einem Google-eigenen, privaten Netzwerk (Google Backbone) aufbaut [53]. Durch standardmäßige Verschlüsselung bei Speicherung (at rest) und Übertragung (in transit) werden personenbezogene, Applikations-, Service- und andere Daten sowie Instanzen und Systeme geschützt.

Eine große Liste [54] von Sicherheits- und IT-Security Services und Funktionen sind auf der Google Cloud verfügbar. Eine Auswahl an relevanten Diensten für SAP Landschaften sind die folgenden, diese werden hier jedoch zum großen Teil nicht im Detail betrachtet und können der Dokumentation entnommen werden:

#### Security Produkte

Access Transparency, Cloud Intrusion Detection System, Cloud Key Management, Firewalls, Secret Manager, Security Command Center, VPC Service Controls uvm.

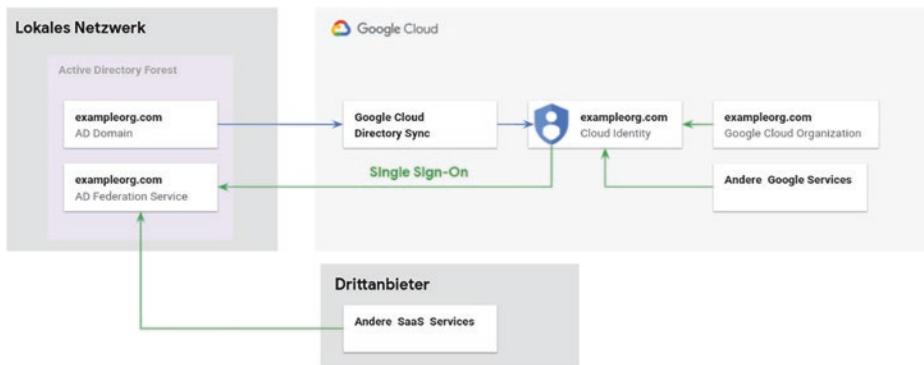
#### Identity und Access Produkte

Certificate Authority Service, Cloud Identity, Identity und Access Management, Managed Service für Microsoft Active Directory, Policy Intelligence, Resource Manager uvm.

##### 8.3.4.1 Identity Provider und Autoritative Quelle

Während des Einstiegs in die Google Cloud muss der Kunde entscheiden, welches System der Identity Provider (IdP) wird, und welches die autoritative Quelle. Die folgenden Optionen stehen hier zur Auswahl [55]:

- Google als IdP und autoritative Quelle
- Google als IdP und ein Personalsystem (Human Resources Information System, HRIS) wie SAP SuccessFactors oder Workday als autoritative Quelle
- Microsoft Active Directory als IdP und autoritative Quelle
- Azure Active Directory (AD) als IdP und Active Directory als autoritative Quelle
- Externes IDaaS (Identity as a Service, Drittanbieterlösung) als IdP und autoritative Quelle



**Abb. 8.5** Verbund mit Active Directory als IdP für Google Cloud

Mit Active Directory als IdP und autoritative Quelle kann die Verteilung umgesetzt werden, wie in der folgenden Grafik [56] veranschaulicht (Abb. 8.5):

Google Cloud Directory Sync ist ein kostenloser, von Google bereitgestellter Dienst, der die Nutzerdaten und Gruppen von Active Directory in Google Cloud Identity oder Google Workspace repliziert. Die Synchronisation erfolgt dabei immer nur in eine Richtung, sodass Active Directory die autoritative Quelle bleibt. Das Active Directory kann dabei entweder in der On-Premise Landschaft sein, wie in der Abbildung veranschaulicht. Kunden können alternativ ebenfalls den vollständig verwalteten Active Directory Dienst in der Google Cloud verwenden. Für die Einmalanmeldung (Single-Sign-On) verwendet Google Cloud Identity die Active Directory Verbunddienste (Federation Services). Auch weitere lokale Unternehmensanwendungen oder andere Drittanbieter-Cloudlösungen nutzen die Active Directory Federation Services als IdP.

### 8.3.4.2 Google Cloud Identity und Access Management (IAM)

Mit Google Cloud Identity und Access Management (IAM) kann die Zugriffssteuerung zu allen Ressourcen durch sogenannte **Richtlinien** (*Policies*) verwaltet werden. Diese sind wie folgt definiert.

- Eine Definition **wer** (eine Identität, wie ein Nutzer, eine Google Gruppe oder ein Service Account) **welche Art** von Zugriff (Rollen wie Viewer, Editor oder Owner) auf **welche** Ressourcen hat [57].

Ein **Google Account** repräsentiert eine Person, die mit Google Cloud interagiert, z. B. ein Administrator oder Entwickler. Der Google Account ist mit einer Mail-Adresse als Identität assoziiert, die zu jeder Domain gehören kann.

Ein **Service Account** wird für eine Anwendung anstelle eines Individuums verwendet. Service Accounts repräsentieren die unterschiedlichen logischen Komponenten einer Applikation.

Eine **Google Gruppe** ist eine Sammlung von Google Accounts und Service Accounts.

Eine **Cloud Identity Domain** ist eine virtuelle Gruppe von allen Google Accounts in einer Organisation welche die Internet Domain dieser Organisation repräsentiert.

Eine **Ressource** ist ein Bestandteil von Google Cloud, der benötigt wird, um einen Dienst zu erfüllen, und auf welchen Berechtigungen gewährt werden können. Eine Ressource kann beispielsweise ein Google Cloud Storage Bucket sein, auf welchen eine VM-Instanz (repräsentiert durch einen Service Account) schreibend zugreifen kann, um eine Sicherung zu speichern.

Eine **Berechtigung** bezieht sich auf eine Handlung, die mit einer Ressource verbunden ist, z. B. braucht man für das Anlegen einer Sicherung auf Google Cloud Storage die Berechtigung *storage.objects.create*.

Eine **Rolle** ist eine Sammlung von Berechtigungen, da diese nicht direkt an Benutzer oder Service Accounts vergeben, sondern davor zu Rollen gruppiert werden sollten. Wenn eine Rolle einem Benutzer zugewiesen wird, so werden alle Berechtigungen in dieser Rolle ebenfalls zugewiesen. Diese können unterteilt werden in:

- Basis Rollen: Owner, Editor, Viewer → nicht empfohlen für produktive Landschaften!
- Vordefinierte Rollen: Diese erlauben eine noch feinere Steuerung als die Basis Rollen
- Benutzerdefinierte Rollen: Diese Rollen können vom Benutzer erstellt werden, um gemäß dem Least-Privilege-Prinzip nur die nötigen Berechtigungen zuweisen zu können und um somit den Anforderungen des Unternehmens gerecht zu werden.

Es gibt viele Security Best-Practices und Empfehlungen für Google Cloud [58] und es ist empfohlen, diese im Detail zu lesen. Eine kurze Zusammenfassung einiger wichtiger Punkte ist:

- Das Least-Privilege-Prinzip sollte immer befolgt und nur der kleinstmögliche Zugriffsbereich vergeben werden
- Service Accounts und Service Account Schlüssel sollten genutzt und regelmäßig rotiert werden
- Alle Zugriffe, Logs, Änderungen und mehr sollten mit Google Cloud Audit Logs auditiert werden
- IAM Richtlinien sollten genutzt und auf der Organisationsebene gesetzt werden
- Rollen sollten an Gruppen vergeben werden und nicht an individuelle Nutzer

### 8.3.5 Google Cloud Operations-Suite

Mit Google Cloud Operations-Suite (früherer Produktnamen Stackdriver) bietet Google eine funktionsreiche Plattform für IT-Operations als völlig verwalteten Service, mit welchem sich die Landschaften und Applikationen bewachen, Fehler suchen und beheben, sowie die Performance der Cloud-Landschaft optimieren lassen [59].

### 8.3.5.1 Generelle Google Cloud Operations Funktionen

#### Monitoring

- Mit den Funktionen für Monitoring können grafische Übersichten und Dashboards zum Einblick in Performance, Zustand und Status aller Systeme beobachtet und analysiert werden
- Sammlung von Metriken, Events, Metadaten von allen Google Cloud Diensten

#### Logging

- Mit den Logging Funktionen können alle gesammelten Logs analysiert und die Behebung von Fehlern und Problemen in Applikationen und Bereitstellungen beschleunigt werden
- Außerdem können Logs von Applikationen und Systemen von vielen anderen Quellsystemen außerhalb von Google Cloud hinzugefügt werden

#### Alerting

- Automatisierte Warnmeldung bei unvorhergesehenen Events über E-Mail, SMS und Auslösen von weiteren Aktionen mithilfe von Automatisierung

#### Application Performance Management (APM)

- Für das Management und die Optimierung der Performance stehen die Produkte Cloud Trace, Cloud Debugger und Cloud Profiler zur Verfügung

### 8.3.5.2 Google Cloud Monitoring und Logging Agent für SAP HANA und SAP NetWeaver

Der Cloud Logging Agent kann als Lösung genutzt werden, um Logs über die Aktivitäten und den Status der gesamten Instanzen und Betriebssysteme von den SAP HANA und SAP NetWeaver Systemen zu sammeln und zu analysieren. Die Nutzung des Cloud Logging Agents ist optional, aber eine durchaus empfohlene Komponente.

Der Monitoring Agent für SAP HANA (V2.0) [60] sammelt Metriken von SAP HANA mithilfe von SQL Abfragen. Im Vergleich zur ersten Version des Monitoring Agenten ist die neue Version eine refaktorierte Version mit Änderungen in Metriken, Installationsmethode, Konfigurationen und Verzeichnissen. Vorformulierte Abfragen können genutzt werden, es gibt jedoch auch die Möglichkeit, eigene Abfragen zu erstellen. Die SQL Syntax und die Systemviews werden von SAP definiert in der Dokumentation vorgegeben [61]. Hauptanwendungsfälle für Monitoring-Metriken und -Dashboards sind beispielsweise:

- Kapazitätsplanung mithilfe der SAP HANA Speicher Auslastung
- Management der Speicher-basierten SAP HANA Lizenzierung
- Identifizierung von Nutzungs- und Performancetrends von zusammen korrelierenden SAP HANA und Compute Engine Metriken
- Erstellung von benutzerdefinierten Abfragen zur Analyse weiterer Metriken, wie beispielsweise SAP HANA Installationsmetriken
- Erstellung von Dashboards zur Visualisierung von SAP HANA Metriken mit Schwellwert-basierten Alarmmöglichkeiten (Alerts)

Um Monitoring Daten von SAP HANA in Google Cloud Monitoring Dashboards anzeigen zu lassen kann nach *sap\_hana* im Metrikfeld gesucht werden.

Der Google Cloud Monitoring Agent für SAP NetWeaver ist zwingend notwendig für den SAP Support von SAP NetWeaver auf Google Cloud [62]. Wenn der Agent auf einer Compute Engine Instanz installiert wird, kombiniert der Monitoring Agent die Daten des Systems und der Compute Engine APIs und liefert diese Daten zum SAP Host Agent. Dieser Agent wird ebenfalls benötigt, wenn SAP NetWeaver auf Google Bare Metal Instanzen bereitgestellt werden soll.

### 8.3.6 Google Cloud Customer-Care-Konzept

Google Cloud bietet verschiedene Kundenbetreuungsangebote (Customer Care [63]), welche alle eine integrierte Nutzererfahrung bieten. Es gibt vier unterschiedliche Support-Klassen, und teilweise können noch zusätzliche Services hinzugebucht werden:

- **Basis Support**
  - Immer inkludiert
  - Support nur für Fragen zur Abrechnung (Billing) → für produktive Unternehmensapplikationen nicht empfohlen
- **Standard Support**
  - SLO von 4 h (bei P2-Fällen)
  - 8 h/5 Tage (Mo-Fr) Support für schwerwiegende Probleme
- **Erweiterter Support**
  - SLO von 1 h (bei P1-Fällen)
  - 24/7 Support für schwerwiegende und kritische Probleme
- **Premium Support**
  - SLO von 15 min (bei P1-Fällen)
  - 24/7 Support für schwerwiegende und kritische Probleme
  - Einen dediziert zugewiesenen Technical Account Manager (TAM)
  - Support für Drittanbietertechnologien

- Trainingsangebote und -guthaben
- Kosten werden basierend auf die monatlichen Nettoausgaben gerechnet und werden mit dem offiziellen Google Cloud Kalkulator oder mit dem Google Cloud Vertrieb berechnet

Für SAP Landschaften auf Google Cloud, die in fast allen Fällen unternehmenskritisch sind, ist es empfehlenswert den Erweiterten oder den Premium Support zu wählen, denn die anderen beiden Google Supportklassen werden von der SAP in einem Supportfall nicht akzeptiert. Diese und weitere Hinweise zum Support sind im SAP-Hinweis 2456406 [64] und in der Google Dokumentation [65, 66] beschrieben. SAP und Google Cloud arbeiten im Rahmen des Supports eng zusammen.

Ferner bietet Google Cloud viele weitere Services, beispielsweise Mission Critical Services (Premium Support) oder Technical Account Advisor Service (Erweiterter Support), oder mit Google Consulting Services auch Beratungspakete für die SAP Migration (beispielsweise Cloud Sprint for SAP 67).

### **8.3.7 Weitere relevante Google Cloud Services für SAP S/4HANA Bereitstellungen**

Vor allem mit Blick auf Innovations- und Erweiterungsprojekte sowie Implementierung von neuen SAP-nahen Lösungen und Applikationen können die folgenden Google Cloud Dienste herangezogen werden:

- **Google BigQuery**
  - Google's Enterprise Data Warehouse, welches vollständig verwaltet und serverlos, zudem hochskalierbar sowie sehr kosteneffizient ist
  - Besitzt Funktionen wie BigQuery Machine Learning (ML), BigQuery Omni (für Multi-Cloud Setups), BigQuery BI Engine, BigQuery GIS (Geographic Information Systems)
  - Integration von SAP-Systemen in BigQuery ist auf drei Arten möglich: über SAP-Integrationslösungen (wie SAP Data Services, SAP SLT, SAP Data Intelligence uvm.), Google Cloud Lösungen (wie Cloud Data Fusion) oder über Partnerlösungen (wie Qlik, Informatica uvm.)
- **Google Vertex AI Plattform**
  - Vollständig verwaltete Auswahl an AI-Lösungen mit vortrainierten und benutzerdefinierten Tools zur Erstellung, Bereitstellung und Skalierung von ML-Modellen
  - Die Erweiterung von SAP Geschäftsprozessen ist damit möglich und wurde schon von Referenzkunden von Google Cloud umgesetzt, beispielsweise für visuelle Inspektion

- **Google Kubernetes Engine (GKE)**
  - Google's vollständig verwaltete Kubernetes Container Plattform
  - SAP Data Intelligence (ehemals SAP Data Hub) und SAP Hybris werden von SAP auf GKE supported → es gibt auch schon Referenzkunden, die diese SAP Lösungen auf GKE betreiben
- **Google Apigee API Management**
  - Google's vollständig verwaltete API Management Plattform mit einer großen Auswahl an Funktionalitäten für den API-Lebenszyklus
  - Für Design, Standardisierung, Absicherung, Analyse, Monetarisierung und Skalierung von APIs
- **Google Industrie Lösungen [68]**
  - Relevant für verschiedene Industrien
    - Advanced Marketing Analytics
    - Contact Center AI
    - Remote productivity and collaboration
    - Procurement DocAI
    - Demand Forecasting
  - Einzelhandel
    - Recommendation AI
  - Verbrauchsgüter
    - Vision API Produktsuche
  - Fertigung
    - Production quality control with Visual Inspection AI
    - Connected Vehicle Telematics
  - Automobilbranche
    - Connected Car Telemetry Platform
    - Connected vehicle solution
    - Infotainment solutions by Google (Android Automotive)
  - Finanzdienstleistungen
    - Open Banking API
    - Lending DocAI
  - Gesundheitswesen und Biowissenschaften
    - Cloud Healthcare API

---

## 8.4 Google Cloud Frontend

Google Cloud bietet verschiedene Möglichkeiten, um mit den Services, Diensten und Ressourcen zu interagieren, je nachdem welche Rolle der Nutzer hat. In diesem Kapitel werden zum einen die Frontend-Tools von Google Cloud vorgestellt und daraufhin die Benutzung dieser Tools nach typischen Nutzerrollen.

### 8.4.1 Google Cloud Frontend-Tools

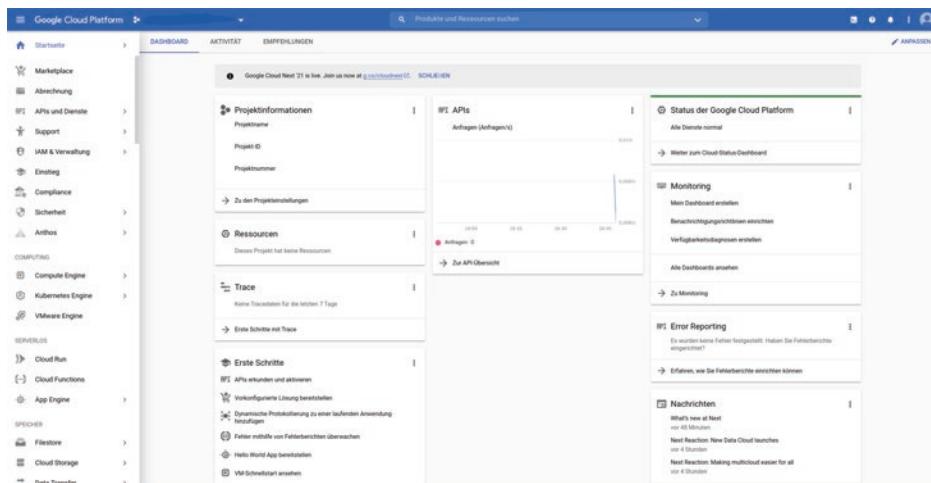
Die **Google Cloud Console** ist eine Weboberfläche im Browser mit sehr guter Suchfunktionalität, in der Nutzer die Hauptaufgaben rund um die typischen Google Cloud Services durchführen können, wie Compute Engine VM-Instanzen anlegen, Cloud DNS Einträge erstellen und vieles mehr (Abb. 8.6).

Mit dem Befehlszeilenspielzeug **gcloud**, welches Teil des Cloud SDK ist, können viele der Plattformaufgaben auf Google Cloud durchgeführt werden, entweder direkt über die Kommandozeile oder über Automatisierungsskripte. Es lassen sich beispielsweise Compute Engine VM-Instanzen und andere Ressourcen erstellen, Cluster, Netzwerke und Subnetze, Cloud DNS verwaltete Zonen und Datensätze anlegen und bearbeiten und vieles mehr. Es gibt zwei Wege das Befehlszeilenspielzeug zu verwenden: entweder über Cloud SDK oder über Cloud Shell:

Das **Cloud SDK** wird auf dem Benutzersystem heruntergeladen und installiert. Vom Computer bzw. System kann dann über ein Terminalfenster das gcloud-Tool genutzt werden, um die Google Cloud Ressourcen über Befehle zu managen.

Die **Cloud Shell** ist eine Admin-Maschine in der Cloud. Diese ist direkt in der Google Cloud Console im Browser verfügbar und läuft auf einer immer sofort verfügbaren Linux Maschine, auf welcher das Cloud SDK direkt vorkonfiguriert ist. Cloud Shell wird verwendet, um schnell Aufgaben über die Befehlszeile zu erledigen. Es sind viele Admin-Tools vorhanden (Abb. 8.7).

Das **Cloud SDK** enthält verschiedene Client-Bibliotheken welche eine einfache Erstellung und das Management von Ressourcen ermöglichen. Google Cloud Client-Bibliotheken bieten APIs für zwei Anwendungsfälle: Admin-APIs für das Ressourcenmanagement und die -erstellung und App-APIs, um Servicezugriffe zu gewähren.



**Abb. 8.6** Google Cloud Console

A screenshot of a terminal window titled "Cloud Shell". The window shows a command being run: "curl -H Metadata-Flavor: Google" metadata/computeMetadata/v1/instance/zone | cut -f1 -f4". The terminal interface includes tabs, a status bar, and a scroll bar.

**Abb. 8.7** Google Cloud Befehlszeilentool gcloud (in der Cloud Shell)

### 8.4.2 Typische Anwenderrollen einer SAP auf Google Cloud Landschaft

In SAP-Umgebungen arbeiten verschiedene Nutzer- und Anwenderrollen auf verschiedenen Ebenen des Systems, und haben dadurch auch unterschiedliche Zugriffe auf die SAP-Systeme. Alle Benutzer, Gruppen, Rollen, Zuständigkeiten und Richtlinien können mithilfe von Google Cloud Identity und Access Management Funktionen konfiguriert und festgelegt werden, wie zuvor im Abschn. 8.3.5 beschrieben:

- **Infrastruktur-Administratoren**

- Arbeiten hauptsächlich mit der Google Cloud Console in den Bereichen Compute Engine, Netzwerk, Identity und Access Management sowie die Kommandozeile gcloud über Cloud Shell und ebenfalls mit Automatisierungsskripten und Bash-Skripten
- Nutzen je nach Bereitstellung und Landschaft ebenfalls Google Operations zu Monitoring- und Loggingzwecken, sowie Cloud Build, Cloud Scheduler und Cloud Functions (siehe Abschn. 9.7)

- **Betriebssystem-Administratoren**

- Arbeiten hauptsächlich mit der Google Cloud über die Kommandozeile gcloud oder über SSH auf der Compute Engine Instanz direkt im System
- Nutzen je nach Bereitstellung und Landschaft ebenfalls Google Operations zu Monitoring- und Loggingzwecken

- **Netzwerkadministratoren**

- Arbeiten hauptsächlich mit der Google Cloud Console im Bereich Netzwerkdienste oder über das Kommandozeilentool gcloud
- Nutzen je nach Bereitstellung und Landschaft ebenfalls Google Operations zu Monitoring- und Loggingzwecken

- **SAP-Applikationsentwickler und SAP-Customizing**

- Arbeiten hauptsächlich am SAP-System selbst, nämlich über die SAP GUI, über das SAP HANA Studio oder über Erweiterungsplattformen und Zusatzservices von SAP oder von Google Cloud

- **Endanwender**

- Arbeiten hauptsächlich auf dem SAP-System direkt, das bedeutet entweder mit SAP Transaktionen über die SAP GUI oder auf dem SAP Fiori Launchpad (im SAP S/4HANA Umfeld meistens der Fall)

## 8.5 SAP S/4HANA auf Google Cloud Architektur

Alle Komponenten für die Bereitstellung einer SAP S/4HANA Architektur auf Google Cloud werden in den nachfolgenden Unterkapiteln beschrieben. Dabei wird zuerst Bezug auf die Lizenzen und die Größenbestimmung genommen und daraufhin die gesamte Architektur erläutert. Auf Basis dieser Architektur werden dann die Planung und die zu treffenden Entscheidungen für das Setup von Hochverfügbarkeit, das Desaster Recovery Setup und die Kriterien für die Auswahl der Speicherlösungen betrachtet. Zuletzt gibt es noch zwei Exkurskapitel, zum einen zu SAP S/4HANA Scale-Out Bereitstellungen und ferner zum Thema SAP HANA Fast Restart und Memory Poisoning Recovery Mechanismus.

### 8.5.1 Lizenzen und Größenbestimmung

SAP Kunden können für SAP auf der Google Cloud ihre vorhandenen SAP Business Suite bzw. S/4HANA Lizenzen im Rahmen des BYOL-Modell (Bring-Your-Own-License-Modell) verwenden und bereitstellen. Die Betriebssystem Lizenzen können entweder über die Google Cloud Compute Engine bereitgestellt und abgerechnet werden, oder Kunden können ihr eigenes Betriebssystem-Image und eigene Lizenzen mitbringen (siehe Abschn. 8.3.2).

Die benötigte Größenbestimmung der SAP Landschaft wird bestmöglich mit dem SAP Quick Sizer Programm durchgeführt. Es ist wichtig zu beachten, dass die SAP Komponenten, hauptsächlich die Datenbank und die Applikationsserver, die auch im nachfolgenden Kapitel vorgestellt werden, auf SAP-zertifizierten Compute Engine Instanzen installiert werden. Ansonsten kann ein Support von SAP nicht garantiert werden. Die relevante Dokumentation zu den Zertifizierungen kann wie zuvor beschrieben in Abschn. 8.3.2 eingesehen werden.

### 8.5.2 SAP S/4HANA Architekturübersicht

Die folgenden Komponenten sind Teil einer SAP S/4HANA Architektur in einer verteilten und nicht zentralisierten Bereitstellung (eine zentralisierte Bereitstellung, das bedeutet SAP S/4HANA Applikation und SAP HANA Datenbank auf einer Compute Engine Instanz, wird nicht für Produktionsumgebungen, sondern nur für Sandbox- und Entwicklungsumgebungen empfohlen und deshalb hier nicht weiter betrachtet, kann jedoch der Dokumentation entnommen werden) [69]:

- SAP HANA Datenbank
- PAS – SAP S/4HANA Primärer Applikationsserver
- AAS – SAP S/4HANA Zusätzliche(r) Applikationsserver

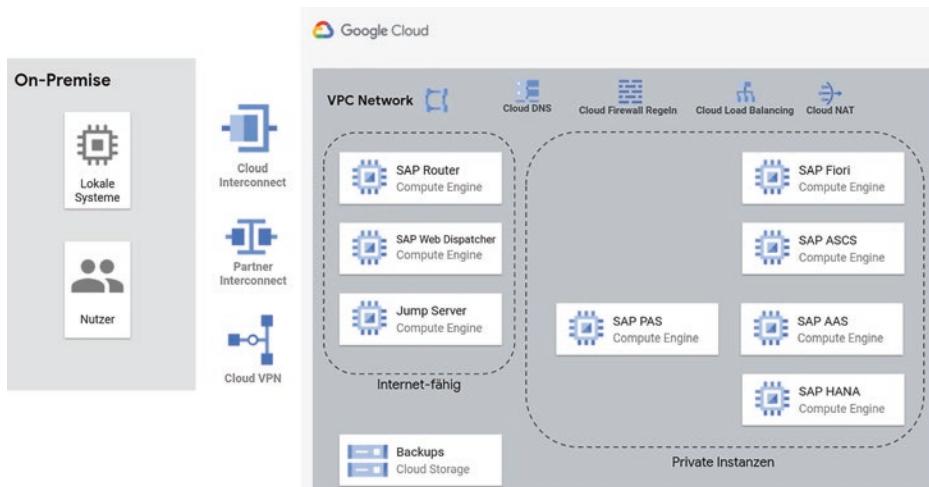
- SAP S/4HANA ASCS
- SAP Fiori Frontend Server
- SAP Web Dispatcher oder Load-Balancer
- SAP Router
- Jump Server/Bastion Host

Ein Load-Balancer ist in einer verteilten Bereitstellung zwingend nötig. Diese Komponenten werden in der folgenden Architekturübersicht auf hoher Ebene veranschaulicht. Mit abgebildet werden einige weitere Google Cloud Dienste, die zuvor im Abschn. 8.3 beschrieben wurden. Dazu gehören unter anderem Google Cloud Storage, Cloud DNS, Cloud NAT, Firewallregeln und das VPC-Netzwerk (Abb. 8.8).

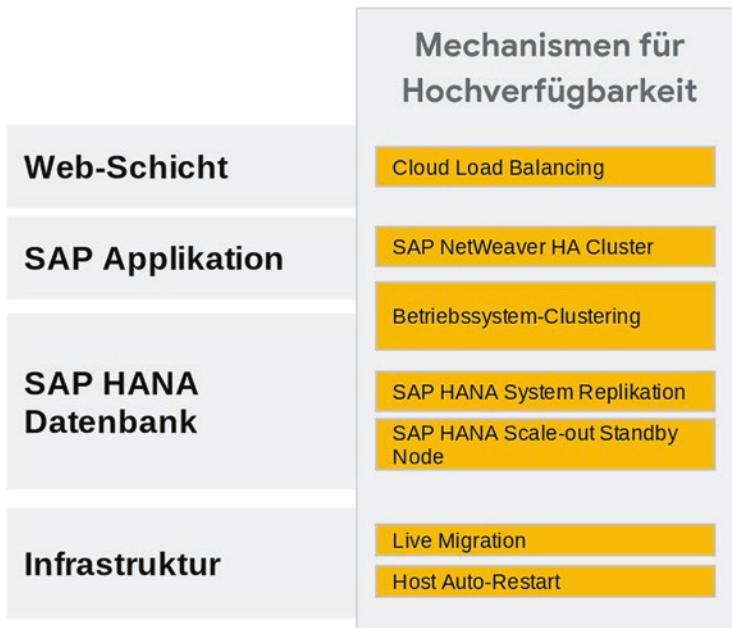
### 8.5.3 Setup für Hochverfügbarkeit

Die Google Cloud bietet verschiedene Mechanismen, um Hochverfügbarkeit in SAP Landschaften umsetzen und sicherstellen zu können. Dabei wird in vier unterschiedliche Ebenen aufgeteilt, welche in der folgenden Grafik veranschaulicht werden. Dazu gehören Mechanismen auf der Infrastruktur-, Datenbank- Applikations- und Webebene [70] (Abb. 8.9).

Es ist wichtig darauf hinzuweisen, dass die beschriebenen Mechanismen der Infrastruktur wie Live-Migration und automatischer Neustart nicht für die installierten Komponenten auf den VM-Instanzen gelten. Dazu gehören das Betriebssystem, die SAP Datenbank und die Applikationsebene. Um hier eine Hochverfügbarkeitsarchitektur



**Abb. 8.8** SAP S/4HANA auf Google Cloud Übersicht der Architekturkomponenten



**Abb. 8.9** Mechanismen für Hochverfügbarkeit für SAP auf Google Cloud

aufzubauen, wird Betriebssystem-Clustering zwischen den SAP Komponenten genutzt: auf SAP HANA Datenbankebene kann die SAP HANA System Replikation und auf der SAP NetWeaver Applikationsebene der Linux Pacemaker Cluster verwendet werden. Dabei wird die für Ihre Landschaft als primär ausgewählte Region der Google Cloud genutzt und das Cluster über zwei Zonen dieser Region verteilt. In Scale-Out Szenarien, beispielsweise für SAP Business Warehouse oder sehr große SAP S/4HANA Systeme mit Scale-Out Konfiguration, können die SAP HANA Scale-Out Standby Nodes für die Umsetzung von Hochverfügbarkeit genutzt werden.

Die oberste Ebene nutzt den Google Cloud Load Balancer, um den Datenverkehr zwischen den Zonen zu leiten und zu verteilen, wenn eine Zone ausfallen sollte.

### 8.5.3.1 Infrastrukturebene

Auf der untersten Ebene, der Infrastrukturebene, ist Google Cloud schon von Design aus hochverfügbar aufgrund der redundanten Infrastruktur mit global verteilten Rechenzentren. Diese Regionen enthalten jeweils mindestens drei Zonen, welche ebenfalls so aufgesetzt sind, dass sie unabhängig voneinander sind. Jede Zone hat eine eigene Strom-, Kühlungs- und Netzwerkversorgung und ist isoliert von den anderen Zonen. Um ein hochverfügbares Cluster für eine SAP-Landschaft aufzubauen, sollten mindestens zwei verschiedene Zonen der primär gewählten Region genutzt und die VM-Instanzen auf diesen verteilt werden. Hinzu kommt, dass Google Compute Engine VMs zwei weitere

wichtige eingebaute Mechanismen liefern, die für Hochverfügbarkeit relevant sind: Live-Migration und Automatischer Neustart (Host Auto-Restart).

### Live-Migration

Mit Live-Migration von Compute Engine laufen die VM-Instanzen weiter, sogar in dem Fall, wenn eine Aktivität wie angekündigte Hardware-, Sicherheits- oder Softwarewartung stattfindet. Live-Migration ist eine inkludierte Funktionalität ohne weitere Zusatzkosten. Live-Migration wird wie folgt definiert:

„Die Live-Migration hält Ihre Instanzen während folgender Ereignisse am Laufen:

- Regelmäßige Infrastrukturwartung und -upgrades
- Netzwerk- und Stromnetzwartung in den Rechenzentren
- Ausfälle der Hardware, z. B. Speicher, CPU, Netzwerkkarten, Festplatten, Netzteil usw. Dies erfolgt auf Best-Effort-Basis. Wenn Hardware vollständig ausfällt oder anderweitig die Live-Migration verhindert, stürzt die VM ab und startet automatisch wieder neu. Dabei wird ein hostError protokolliert.
- Host-Betriebssystem- und BIOS-Upgrades
- Sicherheitsrelevante Aktualisierungen, die schnell umgesetzt werden müssen.
- Änderungen der Systemkonfiguration, einschließlich der Änderung der Größe der Host-Root-Partition, zur Speicherung von Host-Image und Paketen.“ [71]

Live-Migration ändert keine Attribute, Konfigurationen oder Eigenschaften (wie IP-Adressen, Netzwerk, Blockspeicher wie Festplatten usw.) der VM und ist beim Starten einer VM-Instanz automatisch aktiviert. Es ist möglich diese Funktionalität auszuschalten, jedoch wird das generell nicht empfohlen. Während eines Live-Migration Events laufen die installierten Applikationen und Datenbanken weiter und benötigen keine manuelle Aktivität. Eine einzigartige Eigenschaft von Google Cloud Live-Migration ist, dass Live-Migration für alle Compute Engine Maschinentypen funktioniert, somit auch für die großen, SAP HANA relevanten und zertifizierten Maschinenklassen (m1 und m2). Dies bringt große Vorteile für die verfügbare Betriebszeit (uptime) und betriebliche Kontinuität der kritischen SAP S/4HANA Systeme.

### Automatischer Neustart von Instanzen (Host Auto-Restart)

Eine weitere Funktion ist der automatische Neustart von Instanzen (host auto-restart). Wenn dieser für eine Compute Engine Instanz aktiviert ist, so wird die Instanz im Fall das diese durch ein unvorhergesehenes Event runtergefahren wurde wieder automatisch gestartet. Die Applikationen auf dieser Instanz sollten so eingerichtet sein, dass sie beim Hochfahren der VM-Instanz automatisch starten (mit Startup-Skripten etc.). Die Funktionalität automatischer Neustart benötigt keine zusätzliche Einrichtung und wird direkt als eingebaute Funktion der Compute Engine ohne Zusatzkosten zur Verfügung gestellt.

### 8.5.3.2 Hochverfügbarkeit der SAP HANA Datenbankebene

Hochverfügbarkeit für die SAP HANA Datenbank funktioniert mit dem SAP HANA Datenbank nativen Mechanismus namens SAP HANA System Replikation (HSR) und dem Linux-Betriebssystem-Clustering Pacemaker mit dem STONITH-Fencing-Mechanismus. Die Daten werden kontinuierlich und synchron von dem primären System, welches in einer Zone liegt, zu dem sekundären System, welches in einer anderen Zone liegt, repliziert. Das sekundäre System läuft und die Daten sind komplett im Arbeitsspeicher vorgeladen. SQL Befehle zur Bearbeitung von Daten (DML, wie insert, update, delete) können nur im primären System durchgeführt werden. Im Falle eines zonalen Ausfalls kommt es durch die synchrone Datenreplikation somit zu keinem Datenverlust und nur einer minimalen Ausfallzeit. Da jede SQL Transaktion der primären Datenbank erst vollständig verarbeitet ist, wenn diese ebenfalls im sekundären System vollständig verarbeitet ist, kann hier ein Recovery-Point-Objective (RPO) von null erreicht werden (das bedeutet kein Datenverlust). Beide Systeme liegen in derselben Google Cloud Region, jedoch in zwei verschiedenen Zonen, um die SLA von 99,99 % für die Infrastrukturebene mit Compute Engine zu erreichen. In der Bereitstellung wird außerdem die Konfiguration des internen TCP/UDP Load-Balancers und die VIP (Virtuelle IP-Adresse) für das Routen des Datenverkehrs im Fall von Ausfällen durchgeführt. Statische VIP oder Alias IP Bereitstellungen werden von Google Cloud nicht mehr empfohlen.

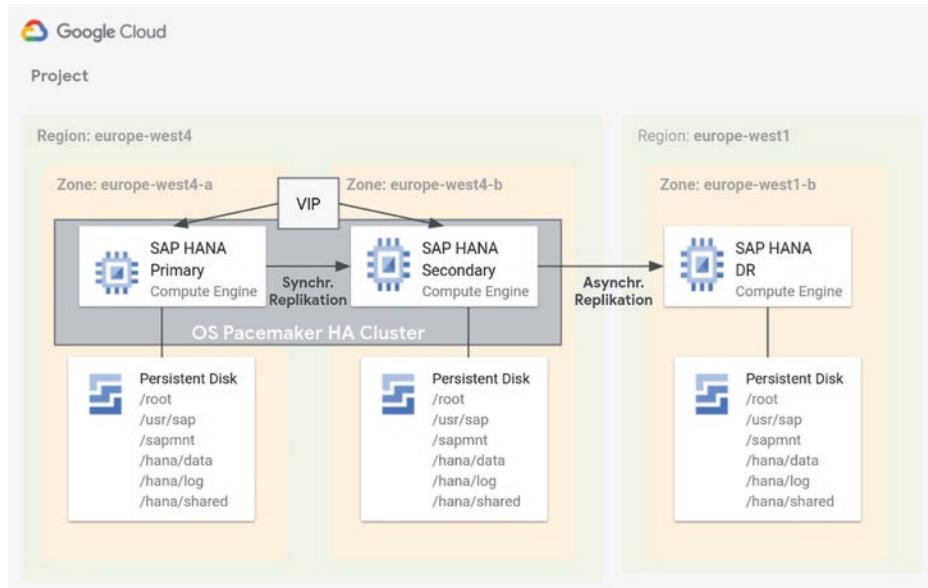
Mehr Informationen können der Planungsdokumentation entnommen werden [72]. Die folgende Architektur [73] zeigt die SAP HANA Ebene mit SAP HANA System Replikation und dem Betriebssystem-Cluster (Abb. 8.10).

### 8.5.3.3 SAP HANA Scale-Out Node Host Auto-Failover

SAP HANA bietet ebenfalls einen nativen Mechanismus für Fehlerbehebung für Scale-Out Bereitstellungen, die von der Google Cloud Infrastruktur unterstützt werden. Der Host Auto-Failover Mechanismus nutzt einen oder mehrere Standby Hosts, welche eingeschaltet sind, um in dem Fall zu übernehmen, wenn ein Primär- oder Worker-Knoten aufgrund eines unvorhergesehenen Fehlers heruntergefahren wird. Wenn der Standby Knoten übernimmt, so wird die Wiederherstellungszeit durch die Größe der SAP HANA Daten des heruntergefahrenen Knoten bestimmt, da diese erstmal in den Arbeitsspeicher geladen werden müssen. Nachdem der Failover durchgeführt wurde, wird der zuvor heruntergefahrenen Knoten wieder neugestartet und übernimmt als neuer Standby Host (siehe Architektur unten) [74].

SAP unterstützt bis zu drei Standby Knoten auf Google Cloud in einem Scale-Out System. Diese Standby Knoten zählen nicht zur maximalen Summe der 16 möglichen aktiven Knoten, sondern zusätzlich dazu.

Wichtig zu merken ist, dass SAP HANA Host Auto-Failover nicht vor einem zonalen Fehler oder Ausfällen schützt, da alle Knoten in einer Zone angelegt werden. In Google Cloud kann anstatt von Host-Auto-Failover ebenfalls die Google Cloud Compute Engine Funktion Auto-Restart genutzt werden. Das hat zwar den Nachteil, dass das



**Abb. 8.10** SAP HANA Hochverfügbarkeit auf Google Cloud

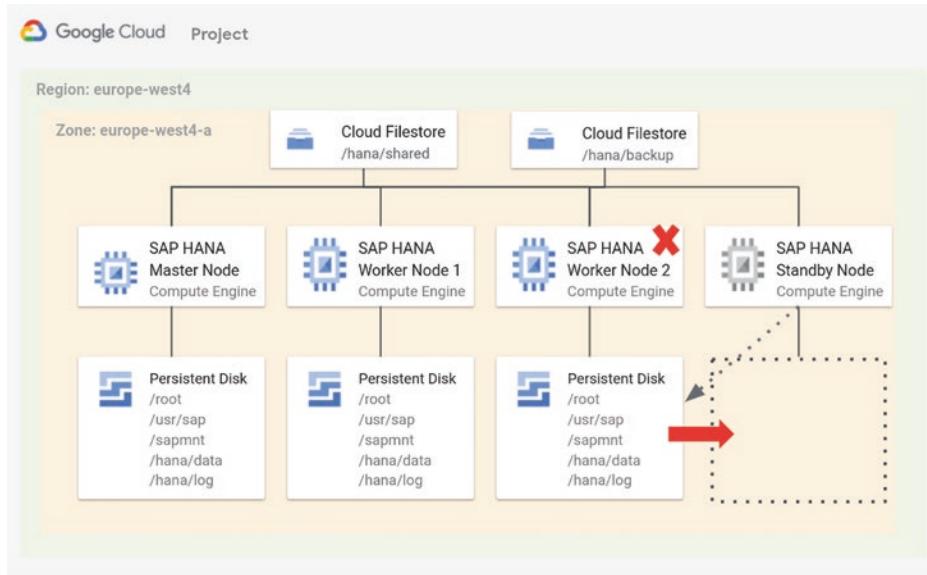
Betriebssystem und SAP HANA gestartet werden muss, bevor der Knoten seine Daten wieder laden kann, jedoch den großen Vorteil, dass nicht dauerhaft eine VM-Instanz in der Größe eines Workers bezahlt werden muss (Abb. 8.11).

#### 8.5.3.4 Hochverfügbarkeit der SAP NetWeaver Applikationsserver

Um Hochverfügbarkeit auf der Ebene der SAP NetWeaver Applikationsserver zu garantieren, kann das Setup mit mindestens vier VM-Instanzen gewählt werden [75]: eine Instanz mit aktivierten ABAP SAP Central Services (ASCS) in der primären Zone der primären Region und eine Instanz mit aktiviertem Standalone Enqueue Replication Server (ERS) in der sekundären Zone derselben primären Region. Die Applikationsserver werden über die zwei Zonen verteilt. Das Cluster für die SAP ASCS und ERS Komponenten enthält das Linux-Betriebssystem-Hochverfügbarkeitscluster (basierend auf Pacemaker) und einen STONITH-Fencing Mechanismus. Das Cluster wird mit einem automatischen Neustart (Host Auto-Restart) für alle Systeme als neue sekundäre Instanzen konfiguriert.

In der Bereitstellung wird außerdem die Konfiguration des internen TCP/UDP Load-Balancers und die VIP (Virtuelle IP-Adresse) für das Routen des Datenverkehrs im Fall von Ausfällen durchgeführt. Statische VIP oder Alias IP Bereitstellungen werden von Google Cloud nicht mehr empfohlen.

Das SAP NetWeaver globale Dateisystem muss für alle SAP NetWeaver-Instanzen in Ihrer Systemlandschaft verfügbar sein und ist daher ein potentieller Single Point of Failure. Um die Verfügbarkeit des globalen Dateisystems in Google Cloud



**Abb. 8.11** SAP HANA Scale-Out Hochverfügbarkeit auf Google Cloud

sicherzustellen, verwenden Sie entweder hochverfügbare Dateifreigabespeicher oder replizierte, zonale nichtflüchtige Speicher (Festplatten) wie in Abschn. 8.3.2 beschrieben wurde.

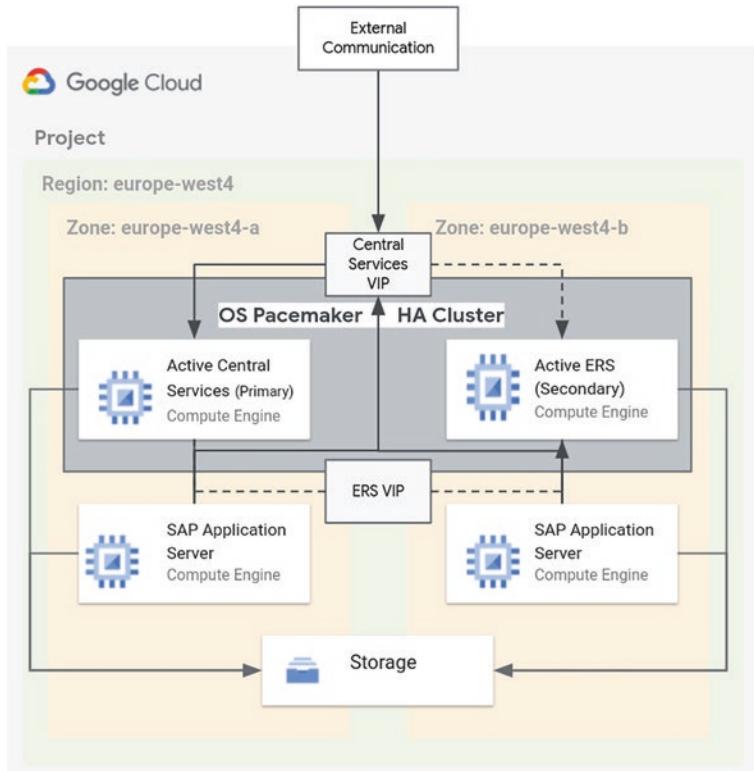
Die folgende Grafik zeigt die Abhängigkeiten zwischen den verschiedenen Komponenten der SAP Applikationsebene (Abb. 8.12):

### 8.5.4 Desaster Recovery Setup

Ein Desaster Recovery (DR) Setup benötigt mehrere Entscheidungen und kann auf viele verschiedene Arten aufgesetzt werden. Mehr generelle Informationen dazu wurden schon im Abschn. 2.2.4 beschrieben.

Die erste Entscheidung, die getroffen werden sollte, ist die Regionenauswahl. Ein Desaster Recovery Szenario auf Google Cloud basiert auf einer primären und einer anderen, sekundären Region, die mit dem Stand der Regionen von September 2021 sehr viele Kilometer voneinander entfernt sind. Ein Regionenpaar in Europa kann beispielsweise *europe-west3* (Frankfurt, Deutschland) und *europe-west1* (St. Ghislain, Belgien) sein.

Die zweite Entscheidung, die getroffen werden sollte, ist die Festlegung der benötigten Recovery-Point-Objective (RPO) und Recovery-Time-Objective (RTO). Generell gibt es fünf verschiedene Strategien ein Desaster Recovery Szenario für SAP-Systeme auf Google Cloud aufzubauen, welche jeweils zu unterschiedlichen RPOs und



**Abb. 8.12** SAP NetWeaver Applikationsserver Hochverfügbarkeit auf Google Cloud

RTOs führen [76]. Die folgende Grafik gibt einen Überblick auf diese fünf Optionen (Abb. 8.13):

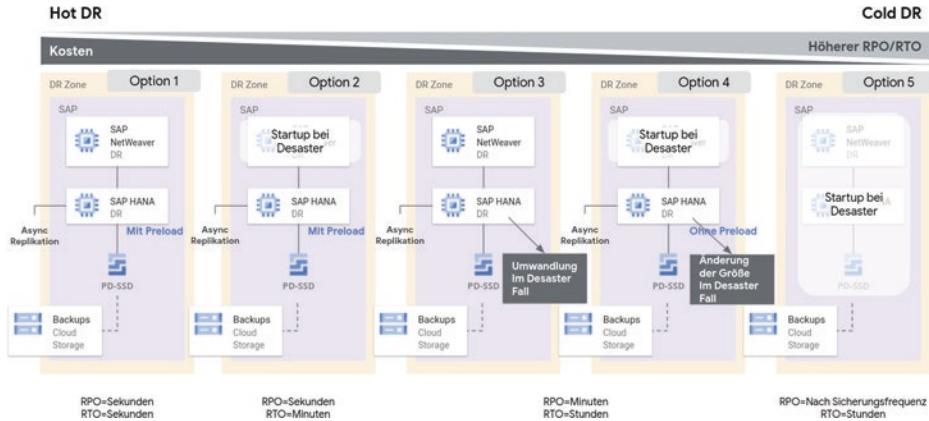
Die Unterschiede der fünf Optionen werden wie folgt definiert:

### 1. Option: Volles DR Setup = Hot DR

- Alle DR Systeme sind vollständig konfiguriert und laufen in einem vollen asynchronen Replikationsmodus
- **Vorteile:** die geringste mögliche RPO und RTO von jeweils nahezu 0, die Desaster Systeme sind verfügbar in wenigen Sekunden
- **Nachteile:** die höchsten Kosten, da die DR Systeme immer am Laufen sind
- Empfehlung: das Umschalten von einer primären Region im Desaster Fall zur sekundären Region sollte manuell für die Nutzer durchgeführt werden, dadurch kann es zu einer RTO im Minutenbereich kommen

### 2. Option: Volles DR Setup nur für die Datenbank

- Die Datenbankschicht ist vollständig konfiguriert und am Laufen → niedrigster RPO von nahezu 0



**Abb. 8.13** Desaster Recovery Optionen für SAP auf Google Cloud

- Die Applikationsserver sind ausgeschaltet und werden aus Snapshotsicherungen mit Automatisierung direkt neu bereitgestellt – aufgrund der kleinen Größe benötigen diese jedoch nicht lange → RTO ist im Minutenumfeld
- **Vorteile:** die laufenden Instanz-Kosten für die Applikationsserver können in der sekundären Region eingespart werden und werden dann erst im Desasterfall abgerechnet
- **Nachteile:** Die RTO erhöht sich im Vergleich zu Option 1 (Hot DR), zwar nicht um Stunden, aber um mehrere Minuten – außerdem sollten die neu bereitgestellten Applikationsserver und die Bereitstellung aus der Sicherung getestet werden

### 3. Option: Umwandlung im Desaster Fall

- Das QA/Test-System liegt bei dieser Option generell in der sekundären Region und wird im Desaster Fall umfunktioniert, das heißt die Delta-Logs seit der letzten Sicherung werden nachgeladen
- Hier liegt der RPO bei Minuten- und der RTO im Stundenumfeld
- **Vorteile:** hohe Kostensparnis, da keine extra DR-Systeme gezahlt werden müssen
- **Nachteile:** QA/Test-System ist während der Desaster Phase nicht verfügbar; es fallen erhöhte Datenverkehrskosten (egress) aufgrund der ständigen Kommunikation zwischen den Regionen an; erhöhte RTO und RPO, da keine kontinuierliche asynchrone Systemreplikation stattfindet, sondern Delta-Logs im Desaster Fall geladen werden

### 4. Option: Datenbank mit halber Größe

- Die Datenbank in der sekundären Region ist nur halb so groß wie die produktive Datenbank und alle Änderungen werden kontinuierlich asynchron über die SAP HANA Systemreplikation geladen

- In einem Desaster Fall startet das DR System neu und wird auf die Größe der Quelldatenbanken hochskaliert; daraufhin müssen die Daten jedoch in den Arbeitsspeicher geladen werden
- **Vorteile:** Kostensparnis und minimaler Datenverlust (RPO) im Minutenbereich, der RTO ist im Stundenbereich
- **Nachteil:** Neustart der SAP HANA Datenbank benötigt und daher hohe RTO, da die Daten in den Speicher geladen werden müssen; es ist zudem nicht garantiert, dass die tatsächlich benötigte Größe in der sekundären Region verfügbar ist falls keine Reservierung der Maschinen vorliegt

## 5. Option: Neue Systembereitstellung im Desaster Fall = Cold DR

- Im Desaster Fall werden die SAP-Systeme in der sekundären Region mit den SAP HANA und Applikationsserver Sicherungen aus dem Google Cloud Storage komplett neu aufgesetzt → höchste RPO und RTO
- Instanzen sollten mithilfe von Automatisierung (Skripte) aufgesetzt werden, um die RTO so gering wie möglich zu halten
- **Vorteile:** die höchsten Kostensparnisse sind mit dieser Methode möglich, jedoch ist die Verfügbarkeit der Compute Engine Instanzen im Desaster Fall nicht garantiert, dafür sind dann Reservierungen nötig, die jedoch wiederum zu keiner Kostensparnis führen (siehe Abschn. 9.5.2)
- **Nachteile:** Die Ladezeit der Daten in den Arbeitsspeicher beträgt pro TB der SAP HANA Größe mehrere Minuten – dies beeinflusst die RTO, welche dadurch in den meisten Fällen im Stundenbereich liegt

### 8.5.5 Erwägungen und Empfehlungen für die Speicherauswahl

Jede Schicht der SAP Architektur hat andere Anforderungen an den benötigten Speicher [77]. Die folgende Tabelle gibt eine Übersicht der Ebenen und die jeweilige Empfehlung für den primären und sekundären Speicher. Die sekundäre Speicherempfehlung entspricht dem Speicher für die Sicherungen (Tab. 8.4).

Für die Sicherungen, also die sekundäre Speicherempfehlung, werden die verschiedenen möglichen Optionen für einzelne Verzeichnisse in der folgenden Grafik veranschaulicht (Abb. 8.14):

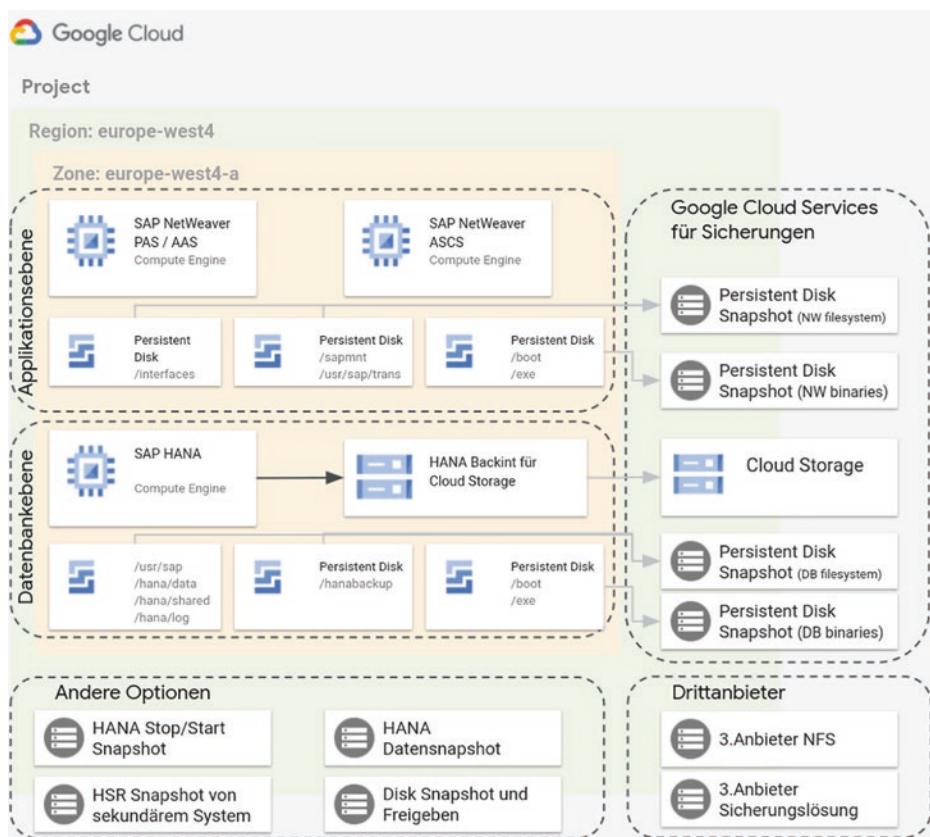
Eine Empfehlung, welche dieser Optionen welche Vor- und Nachteile mit sich bringen und in welchen Anwendungsfällen diese genutzt werden sollten (beispielsweise für produktive Systeme oder nicht-produktive), wird in der folgenden Tabelle gegenübergestellt [78] (Tab. 8.5):

#### 8.5.5.1 Snapshots der persistenten Festplatten

Mit Snapshots lassen sich die Daten von nichtflüchtigem Speicher, also persistente Festplatten wie in Abschn. 8.3.2 beschrieben, schrittweise sichern. Snapshots sind inkrementell und werden automatisch komprimiert, was dadurch eine Kostenoptimierung

**Tab. 8.4** Speicher Erwägungen und Empfehlungen

SAP-System Ebene	Primäre Speicherempfehlung	Sekundäre Speicherempfehlung
Boot-Festplatte und Applikationsverzeichnisse	Festplatten	Festplatten Snapshots
SAP HANA Datenbank	Festplatten	Festplatten Snapshots, Google Cloud Storage Buckets mit SAP HANA Backint
SAP NetWeaver Applikationsserver	Festplatten, Managed NFS (bspw. Google Filestore oder NetApp CloudVolumes...)	Festplatten Snapshots, Google Filestore, Google Cloud Storage Buckets

**Abb. 8.14** Übersicht der Sicherungslösungen für SAP auf Google Cloud

**Tab. 8.5** Gegenüberstellung Vor- und Nachteile der Speicherlösungen und Empfehlung

Speicher-/Sicherungsmethode	Funktionalitäten	Bedenken	Eignung für Landschaft	Empfehlung
Festplatten Snapshot (ganze Festplatte)	Einfaches Tooling, ausfallbeständig, globale Ressource	Konsistenz von Datei- und Applikationslevel nicht gewährleistet	Produktive und nichtproduktive	Für Systemklone, ergänzt die anderen Methoden
Volume Snapshots (Datenbank Sicherungsvolume)	Einfaches Tooling, ausfallbeständig, schnelle Performance, globale Ressource	Nicht anwendungsbezogen, stilllegen des Dateisystems für korrekte Snapshots, kein regelbasiertes Management des Lebenszyklus	Produktive und nichtproduktive	Für produktive Landschaften
SAP HANA Backint Agent für Cloud Storage	Keine Egress-Kosten, Applikations-konsistent, kostenlos, Streaming mit Prüfung der Konsistenz	Kompression ist nicht sehr effektiv, zusätzliche Schritte benötigt für Systemkopien	Produktive und nichtproduktive	Nur für SAP HANA Datenbanken
Drittanbieter NFS	Multi-Zonale Verfügbarkeit	Zusätzliche Kosten, langsamer als nativer Speicher, Egress-Kosten	Produktive und nichtproduktive	Für SAP NetWeaver Hochverfügbarkeitssetup
Drittanbieter Sicherungsagent und Managed Service	Unternehmenslösung, zentralisiertes Management, Erfüllung von industriespezifische Retention und Compliance Anforderung	Zusätzliche Kosten, Egress-Kosten	Produktive	Für spezifische Kunden und Industrien mit langen Retention Perioden
SAP HANA Datensnapshot	SQL-basiert, Applikations-konsistent, Datenbank Tooling	Nicht unterstützt für multi-tenant Datenbanken, extra Konfiguration benötigt	Nur nichtproduktive Landschaften	Nur für SAP HANA Datenbanken
“Stop-und-Start” Snapshots	Einfacher Weg um Konsistenz sicherzustellen	VM Neustart ist nötig	Nur nichtproduktive Landschaften	Schnelle nichtproduktive Systemklone
Snapshot und freigeben	Speicher wird nur für die Snapshot-erstellung hinzugefügt	Scripting ist nötig	Nur nichtproduktive Landschaften	Anwendungsfälle mit hohem Kostendruck

mit sich bringt. Dabei funktionieren sie wie folgt: Der erste erfolgreiche Snapshot eines nichtflüchtigen Speichers ist ein vollständiger Snapshot, der alle Daten des nichtflüchtigen Speichers enthält. Der zweite Snapshot enthält nur neue Daten oder Daten, die seit dem ersten Snapshot geändert wurden. Daten, die seit dem ersten Snapshot nicht geändert wurden, werden nicht inkludiert. Stattdessen enthält der zweite Snapshot bei den unveränderten Datenblöcken einfach Verweise auf den ersten Snapshot. Der dritte Snapshot enthält dann die Daten, die seit dem zweiten Snapshot neu hinzugefügt oder geändert wurden, enthält aber ebenfalls keine unveränderten Daten aus dem ersten und zweiten Snapshot, sondern Verweise auf diese [79].

Snapshots sind globale Ressourcen und können somit über Projekte hinweg geteilt und genutzt werden [80]. Da Snapshots ebenfalls im Google Cloud Storage gespeichert werden, können die Storage Buckets regional oder multi-regional sein (dabei ist die nächste Region die Standardeinstellung).

Als Best-Practice sollten Sie bei der Erstellung der Snapshots die Stilllegung der Festplatte in Betracht ziehen, das bedeutet ein Einfrieren der Dateisysteme oder die anwendungsbezogene Schreibunterbrechung. Außerdem empfiehlt es sich Snapshots zu automatisieren und direkt mit Zeitplänen zu versehen [81].

Die Abwägungen zur Verwendung von Snapshots für Sicherungen in SAP-Landschaften sind:

- **Vorteile**
  - Schnell und kosteneffizient
  - Es bedarf keiner weiteren Replikation zwischen Regionen, da Google Cloud Storage global adressierbar ist und die Replikationen (je nach Speicherklasse) direkt übernimmt
  - Absturzkonsistent
- **Nachteile**
  - Keine Konsistenz auf Datei- oder Anwendungsebene gewährleistet, nur Konsistenz auf Blockebene

Im SAP Applikationsserver Umfeld können die folgenden Festplatten und Binaries mit Hilfe von Snapshots gesichert werden [82]:

- SAP NetWeaver Binär- und Konfigurationsdateien
- Geteilte Dateien oder Transportdateien in einer SAP Landschaft
- Ordner, welche in hochverfügaren Systemen benötigt werden (locking table bei Failover-Clustern)
- Ordner mit Schnittstellen für andere Systeme

Im SAP HANA Umfeld sind Snapshots von Festplatten zwar ausfallbeständig (crash-consistent), sie geben jedoch keine Konsistenz auf Applikationsebene oder Dateiebene und sind deshalb nicht im Produktivumfeld für die SAP HANA Daten in/hana/data

empfohlen. Stattdessen wird empfohlen die SAP HANA Backup Volumes zu sichern. Für die folgenden Volumes werden Snapshots im SAP HANA Umfeld empfohlen:

- Applikationsbinärdateien und -verzeichnisse
- Datenbankkonfigurationsdateien
- Verzeichnisse mit Daten oder Datenbanktransaktionen
- Geteilte Dateien wie Binärdateien, Logs und Konfigurationen
- Redo-Logs
- Dateien und Dateisysteme mit Datenbank, Dateischnittstellen und Applikationskonfigurationssicherungen

Es wird empfohlen mit Maschinen-Images im Umfeld von Snapshots zu arbeiten, da die meisten SAP-Systeme aus mehreren Festplatten bestehen. Diese Images speichern alle Konfigurationen, Metadaten, Berechtigungen und Daten von einem oder mehreren Festplatten, die für eine vollständig funktionierende VM-Instanz benötigt werden. Das Image wird ebenfalls im Google Cloud Storage in einem regionalen oder multi-regionalen Bucket gespeichert. Daher ist hier keine Replikation nötig. Die IP-Adressen und die Daten im Arbeitsspeicher werden nicht gespeichert [83].

Maschinen-Images werden für folgende Anwendungsfälle genutzt:

- Erstellung von Instanzen
- Clonen von Instanzen
- Sicherung von Instanzen
- Teilen von Instanzen
- Desaster Recovery

### 8.5.5.2 Google Cloud Storage Buckets mit SAP HANA Backint

Google Cloud Storage bietet einen kostenlosen und SAP-zertifizierten Backint Agent für SAP HANA [84]. Wenn SAP HANA Backint für die Sicherungen genutzt wird, so werden die nativen SAP HANA Datenbanktools genutzt, um Sicherungen und die Eigenschaften zu konfigurieren. SAP HANA Backint kann in Scale-Up und in Scale-Out Landschaften genutzt werden und muss auf allen Knoten installiert und konfiguriert werden. Der Backint-Agent bietet einen hohen Durchsatzdurchschnitt und ist zudem Multi-Stream-fähig, weshalb große Datenbanken ebenfalls schnell gesichert werden können. SAP HANA Backint kann in drei verschiedenen Sicherungsmodi agieren: voll, inkrementell und differentiell.

- **Vorteile:**
  - Eliminierung von Festplatten für das Sicherungsvolume
  - Konsistenz der Applikation gesichert
- **Nachteile:**
  - Kompression ist nicht effektiv
  - Nur spezifisch und ausschließlich für SAP HANA Datenbanken

Weitere Sicherungslösungen, die genutzt werden können, sind Actifio, NetApp, Commvault und viele weitere Partner- und Drittanbieterlösungen.

### 8.5.5.3 Actifio

Actifio gehört zu Google Cloud und bietet Funktionalitäten für anwendungskonsistente und inkrementelle Sicherungen. Die Vorteile sind sehr geringe RPO aufgrund der häufigeren kleineren Sicherungsfenster sowie ein RTO, welcher von Stunden zu Minuten minimiert werden kann. Das wird durch die einzigartige Funktionalität von Actifio ermöglicht, indem Multi-TB Datenbanken sofort wiederhergestellt werden können. Actifio speichert Sicherungen im nativen Format. Laufwerke können dann direkt von der Backup-Appliance gemountet werden. Unterstützt werden Datenbanken wie SAP HANA, Oracle, Microsoft SQL Server, PostgreSQL und MongoDB. Actifio kann Sicherungen auf allen unterschiedlichen Applikationsebenen durchführen sowie in On-Premise Landschaften und in verschiedensten Private und Public Clouds: Google Cloud, Azure, AWS, IBM, Oracle und mehr [85].

Das neueste Produkt Actifio GO für Google Cloud ist eine vollständig verwaltete Lösung (SaaS), die für Unternehmenssicherungen, Disaster Recovery Bereitstellungen, schnelles Klonen von Datenbanken und mehr dient. Dabei besteht Actifio GO für Google Cloud aus zwei Hauptkomponenten:

**Actifio Global Manager (AGM):** Diese Managementkonsole wird automatisch in der Google Cloud provisioniert. AGM dient der Konfiguration von Sicherungs-SLAs, Wiederherstellung von Dateien, Ordnern und VM-Instanzen sowie der Disaster Recovery Bereitstellung auf Google Cloud. AGM wird außerdem genutzt um einen oder mehrere Actifio Sky Data Mover (“Appliance”) zu administrieren und zu monitoren.

**Actifio Sky Data Mover:** Der Sky Data Mover ist eine VM-Instanz, welche im On-Premise Rechenzentrum oder in der Google Cloud installiert wird, um Ihre Systeme zu schützen. Sky Data Mover ist die Komponente, welche Aufgaben wie inkrementelle ewige Sicherungen, sofortige Wiederherstellung und sofortiges Klonen von Datenbanken durchführt.

Weitere Eigenschaften von Actifio:

- Durch die ***incremental forever*** Technologie können mehrere Datenstände (Snapshots) vorgehalten werden, ohne den Speicherverbrauch über eine Gebühr zu erhöhen.
- Datenstände können im read/write Modus direkt von der Appliance gemountet und mit SAP HANA verwendet werden. Dadurch kann im Fall eines Datenverlusts der Betrieb wieder aufgenommen werden, ohne dass ein Restore notwendig wäre. Der Restore-Schritt kann je nach Datenbankgröße sehr zeitaufwändig sein. Restore-Zeiten von acht Stunden und mehr sind bei großen Datenbanken keine Seltenheit. Und gerade große Datenbanken sind typischerweise besonders unternehmenskritisch.
- Die Möglichkeit, Snapshots zu mounten, ermöglicht das Eingrenzen eines logischen Fehlers. Wenn zum Beispiel eine Tabelle versehentlich gelöscht wurde, ist es möglich,

auf einem nicht produktiven System mehrere Snapshots zu mounten, um zu sehen, zwischen welchen Snapshots diese Tabelle gelöscht wurde. Dadurch wird auch eine Reparatur der Datenbank denkbar.

- Durch die Möglichkeit, Snapshots read/write zu mounten, können auch produktive Datenbankumgebungen über Anonymisierung und SID-rename in Pre-produktive, Qualitätssicherungs-, Test- und Entwicklungs-Umgebungen überführt werden.

## Vorteile

- Unternehmenslösung, welche viele Plattformen und Setups wie Hybrid- und Multi-Cloud unterstützt
- Erfüllung von industrie-spezifischen Aufbewahrungs- und Compliance-Anforderungen
- Wiederherstellung zu bestimmten Zeitpunkten (point-in-time recovery)
- Sofortige Wiederherstellungsfunktion
- Virtuelle Kopien ohne extra Speicher

## Nachteile

- Abwägung der zusätzlichen Service Kosten vs. Vorteile des vollständig verwalteten Services
- Bei komplexen Setups können im Sonderfall minimale Kosten für ausgehenden Datenverkehr anfallen

### 8.5.5.4 NetApp

NetApp Cloud Volumes (CVS) wurden schon im Abschn. 8.3.2 beschrieben und sind als vollständig verwalteter Datenspeicherservice integriert mit Google Cloud und der Console.

Sie bieten verteilte Dateisysteme, die im SAP-Umfeld typischerweise benötigt werden für:

- Schnittstellen, also Dateien, die von einer Applikation geschrieben und von einer oder mehreren anderen gelesen werden
- Das Transportverzeichnis zum Propagieren von Änderungen über Umgebungsgrenzen hinweg (typischerweise/usr/sap/trans)
- Das sogenannte SAPMNT Verzeichnis, auf das mehrere Applicationsserver Zugriff haben
- Die über mehrere HANA Scale-out Knoten geteilten ausführbaren Dateien, Profile und Traces, typischerweise in/hana/shared

NetApp Cloud Volumes bieten außerdem viele Möglichkeiten für die Migration und die Bereitstellung von SAP Landschaften:

- Snapshotscheduling
- Speicherung von Snapshots in Google Cloud Storage Buckets
- SAP NetWeaver Hochverfügbarkeitsetups
- und mehr

Derzeit werden nur Multi-Zonen-Verfügbarkeiten angeboten, jedoch keine Multi-Regionen Desaster Recovery Funktionen. Für die Persistenz von SAP HANA Installationen (/hana/data und /hana/log) sollte CVS nicht verwendet werden (Stand September 2021).

#### **8.5.5.5 CommVault**

Eine weitere Sicherungslösung, die nativ mit Google Cloud integriert ist, ist CommVault. Es ist eine bekannte Lösung, die schon bei vielen Unternehmen und Branchen im Einsatz ist, in denen die Richtlinien und Vorschriften zur Datenaufbewahrung streng sind. Es wird keine Installation eines Agenten benötigt und CommVault bietet auch einige sehr spezielle Sicherheitsfunktionalitäten und Funktionen, die die langfristige Aufbewahrung der Backup-Daten unterstützen. Ein Merkmal sind optimierte, selbstanpassende Sicherungen mithilfe von Maschinellem Lernen, das die Leistung für Sicherungs erstellung verbessert.

#### **8.5.6 Exkurs: SAP S/4HANA Scale-Out (SAP Innovation Award 2021)**

Dieser Exkurs gibt eine Übersicht zum Thema SAP S/4HANA Scale-Out Architektur auf Google Cloud, welches Google gemeinsam mit SAP in einem Co-Innovationsprojekt mit dem Kunden PayPal durchgeführt hat [86].

Die Ziele des Projektes waren die Erprobung der Skalierbarkeit und Performance der Google Cloud für SAP S/4HANA Financial Products Subledger (FPLS). Mit SAP HANA ist es möglich, gemischte Arbeitslasten zu optimieren und parallel verarbeiten zu lassen – SAP HANA unterstützt sowohl OLTP als auch OLAP Transaktionen in einer Plattform. Die ersten Tests wurden mit einem Cluster mit 8 Knoten (Nodes) mit jeweils 4 TB (Maschinentyp m1-ultramem-160), also 32 TB DRAM durchgeführt. Die zweite Phase des Projektes wurde mit 8 Knoten à 12 TB (Maschinentyp m2-ultramem-416), also 96 TB DRAM durchgeführt und getestet.

Fakten über die Daten, die im SAP-System genutzt wurden:

- Das Projekt startete mit 27 TB realen Finanzdaten
- 10 Mio. Konten und 40 Mio. tägliche Zahlungstransaktionen
- Insgesamt 12 Mrd. Zahlungen und mehr als 200 Mrd. Nebenbucheinträge für die Jahresendabrechnung

Im Folgenden werden die wichtigen Erkenntnisse und Erfolge, die mit diesem Projekt erreicht wurden, zusammengefasst [87]:

- Die **Bearbeitungsdauer** von 200 Mrd. Einträgen für Jahressalden (year to date balances) wurde **reduziert von 20 min auf nur 30 s**, verglichen mit dem Performancelevel für Scale-Up Systeme von ähnlicher Größe
- **40-fache Beschleunigung Abfragelaufzeitgeschwindigkeit durch Parallelisierung**
- **Laufzeitreduzierung von 2,7-fach – also von zwei Stunden auf 44 min**, verglichen mit dem Performancelevel von On-Premise Scale-Up Systemen
- Die Lasten wurden fortschreitend von 40 Mio. auf 160 Mio. erhöht, und somit die **Anzahl der Einzelposten von 500 Mio. zu 2 Mrd. pro Tag vervierfacht**
- Implementierung und Nutzung von neuen Optimierungsmöglichkeiten wie mit SAP HANA Data Tiering brachte eine **78%-ige Speicherbedarfsreduzierung**
- Ingenieure konnten eine konsistente Geschwindigkeit von mehr als 5 GB/Sekunde für Sicherungen (Backups) in einem multi-regionalen Cloud Storage Bucket messen (16,97 TB mit 5,39 GB/Sekunde = 54 min)
- Die Zeit für das Applikationsmanagement (Wartungsfenster) zur Umverteilung von SAP HANA Scale-Out Knoten (von 8 zu 10 Knoten) konnte **um 50 % gesenkt werden**

Mit diesem Projekt haben die beteiligten Unternehmen PayPal und Google Cloud den SAP Innovation Award 2021 [88,89] gewonnen.

### 8.5.7 Exkurs: SAP HANA Fast Restart und Memory Poisoning Recovery Mechanismus für SAP S/4HANA

Eine ganz neue Funktionalität der Google Cloud, veröffentlicht im September 2021, ist Memory Poisoning Recovery (MPR) [90] in Kombination mit SAP HANA Fast Restart [91]. Der Hintergrund dazu ist, dass unkorrigierbare Speicherfehler eine der häufigsten Ursachen für Hardwareausfälle sind. Traditionell führten diese zu einem Neustart der Hardware. Dieser kann, je nach Datenbankgröße, eine lange Zeit in Anspruch nehmen, da die Daten wieder zurück in den Speicher geladen werden müssen.

Wie funktioniert MPR und wie hilft es diese Herausforderung anzugehen? Die Lösung basiert auf verschiedene Ebenen, zum einen die Google Cloud Live-Migration Funktion und dazu die Möglichkeiten auf CPU Ebene (von Intel) und die Ebene mit der Datenbank SAP HANA. Zwei Hauptprozesse werden durchgeführt:

- **Isolierung des Speicherfehlers (memory error)**
  - Die gehärtete VM Technologie von Compute Engine fängt Fehler vom System ab und analysiert diese. Die signalisierten Regionen des Speicher-DIMMs werden als “befallen” markiert.

- Verschiedene Prozesse werden gestartet, welche den beschädigten Bereichen nachgehen.
- **Wiederherstellung des Speicherfehlers (memory error)**
  - Notifizierung des Gastbetriebssystems & der Memory-Check-Exception-fähigen Applikationen, dass ein Fehler verzeichnet wurde. SAP HANA reagiert darauf mit einem Neustart. Sofern SAP HANA Fast Restart eingestellt ist, muss dazu nur der befallene Teil der Daten von der Festplatte geladen werden.
  - Es erfolgt eine direkte Kommunikation an Compute Engine, um ein Live-Migration-Event zu starten und den betroffenen Host zu evakuieren und die Hardware-Wartung durchzuführen.

Nach aktuellem Stand wird MPR im letzten Quartal von 2021 für die arbeitsspeicher-optimierten Compute Engine Maschinen (m1- und m2-Maschinentypen) verfügbar sein und in der Zukunft kontinuierlich auf weitere Google Cloud Maschinentypen ausgerollt werden.

---

## 8.6 Zusammenfassung

Dieser Abschnitt gibt eine Zusammenfassung der vorherigen Kapitel mit allen Hauptpunkten und Überlegungen, die bei der Planung einer SAP S/4HANA Landschaft auf Google Cloud im Hinblick auf die Architektur und Konzepte bedacht werden sollten.

Nach einer Übersicht der Geschichte von Google Cloud und der zeitlichen Entwicklung der Partnerschaft mit SAP, welche seit 2017 besteht, werden die Entscheidungsgründe für SAP auf Google Cloud näher beleuchtet. Dabei sind die Hauptgründe unter anderem die Innovation, die Risiko- und Ausfallminimierung, die Flexibilität durch vereinfachte Bereitstellung und die Nachhaltigkeit der Google Cloud.

Daraufhin werden die generellen Konzepte und relevanten Google Cloud Services für eine SAP S/4HANA Bereitstellung erläutert. Die Google Cloud Organisationshierarchie mit Ordnern, Projekten und Ressourcen bildet dabei die Basis. Danach werden die Dienste Compute Engine, die relevanten Speicherdiene wie Blockspeicher, Dateifreigabespeicher und Google Cloud Storage erklärt. Dann folgt das Netzwerkkonzept, welches aus dem VPC-Netzwerk, den Subnetzwerken, Routen und Weiterleitungsregeln, Firewallregeln, dem Cloud Load Balancer, Cloud DNS und Cloud NAT besteht. Die Netzwerkanbindung erfolgt durch ein Cloud VPN, Partner Interconnect oder ein Dedicated Interconnect. Dann werden die Grundlagen zu Google Cloud Security und dem Identity und Access Management gegeben. Letztendlich folgen jeweils ein Überblick zur Google Cloud Operations-Suite mit Monitoring und Logging Services sowie dem Google Cloud Support und die genutzten Google Cloud Frontend Tools wie die Console und gcloud Kommandozeile.

In den weiteren Kapiteln werden die Architekturkomponenten von SAP S/4HANA auf Google Cloud und die Überlegungen und Entscheidungen zur Bereitstellung für

Hochverfügbarkeit, Desaster Recovery und Datenmanagement auf Basis der zuvor beschriebenen Google Cloud Konzepte und Services erklärt.

Zuletzt werden zwei spezielle Exkurs-Kapitel im Umfeld von Co-Innovationen aufgezeigt: zum einen die Ergebnisse eines Projektes für ein 96 TB großes SAP S/4HANA Scale-Out Projekt gemeinsam mit der SAP und dem Referenzkunden PayPal, welches mit dem SAP Innovation Award 2021 ausgezeichnet wurde. Zum anderen werden die Vorteile von SAP HANA Fast Restart mit Google Memory Poisoning Recovery erläutert. Diese beiden Projekte zeigen unter anderem, dass Google Cloud hier im Vergleich zu anderen Anbietern einzigartige Innovationen bietet.

Zusammenfassend bietet Google Cloud eine moderne Infrastruktur für SAP-Systeme. Virtuelle Maschinen werden als Google Compute Engine Instanzen abgebildet, die eine ähnliche Steuerbarkeit bei höherer Flexibilität bieten. Die Steuerbarkeit erstreckt sich im Unterschied zu anderen Instanztypen über das gesamte Betriebssystem. Die Infrastruktur von Google Cloud bringt Hochverfügbarkeitsmechanismen mit und für Sicherungen und Speicher gibt es viele Services zur Auswahl.

Neu im Vergleich zu On-Premise Landschaften ist die Elastizität. Es macht keinen Unterschied wie viele Maschinen ein Benutzer provisioniert. Ferner kommt die Flexibilität hinzu: statt wochenlanger Beschaffungszeiten können Instanzen in Minuten bereitgestellt und sekundengenau abgerechnet werden. Nutzer können sich aber auch für andere Abrechnungsmodelle entscheiden, mehr dazu im nachfolgenden Kapitel.

---

## Literatur

1. <https://cloud.google.com/products> (Zugriff am 20.12.2021)
2. <https://news.sap.com/2017/03/just-google-it-cloud-has-arrived/> (Zugriff am 20.12.2021)
3. <https://news.sap.com/2018/06/general-availability-sap-data-custodian-google-cloud-platform/> (Zugriff am 20.12.2021)
4. <https://blogs.sap.com/2019/05/20/getting-started-with-the-sap-cloud-platform-extension-factory/> (Zugriff am 20.12.2021)
5. <https://cloud.google.com/blog/products/sap-google-cloud/announcing-the-general-availability-of-6-and-12tb-vms-for-sap-hana-instances-on-gcp> (Zugriff am 20.12.2021)
6. <https://www.linkedin.com/pulse/sap-google-cloud-partnership-our-joint-journey-continues-gary-slater/?trackingId=KOfUM2yTQbSWZnmaReCBBg%3D%3D> (Zugriff am 20.12.2021)
7. <https://www.heise.de/news/Google-migriert-Finanzsoftware-von-Oracle-zu-SAP-6005704.html> (Zugriff am 20.12.2021)
8. <https://www.asug.com/events/alphabets-sap-s-4hana-journey-and-partnering-for-innovation> (Zugriff am 20.12.2021)
9. <https://cloud.google.com/blog/products/sap-google-cloud/sap-and-google-cloud-expand-their-partnership> (Zugriff am 20.12.2021)
10. <https://news.sap.com/2021/07/google-cloud-and-sap-accelerate-business-transformations-cloud/> (Zugriff am 20.12.2021)
11. <https://cloud.google.com/blog/products/sap-google-cloud/protect-hana-uptime-with-fast-restart-on-google-cloud> (Zugriff am 20.12.2021)

12. <https://cloud.google.com/blog/products/sap-google-cloud/mitigating-memory-errors-for-your-sap-environment> (Zugriff am 20.12.2021)
13. <https://cloud.google.com/blog/products/sap-google-cloud/reports-examine-business-value-of-running-sap-on-google-cloud> (Zugriff am 20.12.2021)
14. [https://services.google.com/fh/files/misc/forrester\\_tei\\_for\\_sap\\_on\\_google\\_cloud\\_infographic.pdf](https://services.google.com/fh/files/misc/forrester_tei_for_sap_on_google_cloud_infographic.pdf) (Zugriff am 20.12.2021)
15. <https://sustainability.google/intl/de/> (Zugriff am 20.12.2021)
16. <https://cloud.google.com/resource-manager/docs/cloud-platform-resource-hierarchy#projects> (Zugriff am 20.12.2021)
17. [https://cloud.google.com/iam/docs/resource-hierarchy-access-control#best\\_practices](https://cloud.google.com/iam/docs/resource-hierarchy-access-control#best_practices) (Zugriff am 20.12.2021)
18. <https://cloud.google.com/about/locations> (Zugriff am 20.12.2021)
19. <https://cloud.google.com/compute/docs/regions-zones/global-regional-zonal-resources> (Zugriff am 20.12.2021)
20. <https://cloud.google.com/solutions/sap/docs/sap-hana-on-bms-planning> (Zugriff am 20.12.2021)
21. <https://cloud.google.com/solutions/sap/docs/certifications-sap-hana> (Zugriff am 20.12.2021)
22. <https://cloud.google.com/solutions/sap/docs/certifications-sap-apps> (Zugriff am 20.12.2021)
23. <https://www.sap.com/dmc/exp/2014-09-02-hana-hardware/enEN/#/solutions?filters=iaas;ve:29> (Zugriff am 20.12.2021)
24. <https://launchpad.support.sap.com/#/notes/2456432> (Zugriff am 20.12.2021)
25. <https://cloud.google.com/solutions/sap/docs/certifications-sap-apps#sap-certified-vms-custom> (Zugriff am 20.12.2021)
26. <https://cloud.google.com/compute/sla> (Zugriff am 20.12.2021)
27. <https://cloud.google.com/solutions/sap/docs/sap-hana-os-support> (Zugriff am 20.12.2021)
28. <https://cloud.google.com/solutions/sap/docs/netweaver-os-support> (Zugriff am 20.12.2021)
29. <https://launchpad.support.sap.com/#/notes/2235581> (Zugriff am 20.12.2021)
30. <https://launchpad.support.sap.com/#/notes/2456432> (Zugriff am 20.12.2021)
31. [https://userapps.support.sap.com/sap\(bD1lb1ZjPTAwMQ==\)/support/pam/pam.html#ts=0](https://userapps.support.sap.com/sap(bD1lb1ZjPTAwMQ==)/support/pam/pam.html#ts=0) (Zugriff am 20.12.2021)
32. <https://cloud.google.com/compute/docs/disks#introduction> (Zugriff am 20.12.2021)
33. [https://cloud.google.com/solutions/sap/docs/architectures/sap-s4hana-on-gcp#recommended\\_linux\\_directory\\_structure\\_for\\_sap\\_hana](https://cloud.google.com/solutions/sap/docs/architectures/sap-s4hana-on-gcp#recommended_linux_directory_structure_for_sap_hana) (Zugriff am 20.12.2021)
34. <https://cloud.google.com/solutions/sap/docs/checklist-sap-hana> (Zugriff am 20.12.2021)
35. <https://cloud.google.com/solutions/sap/docs/sap-hana-planning-guide#hana-minimum-pd-sizes-ssd-balanced> (Zugriff am 20.12.2021)
36. <https://cloud.google.com/solutions/sap/docs/filers-for-sap> (Zugriff am 20.12.2021)
37. <https://cloud.netapp.com/cloud-volumes-global-regions> (Zugriff am 20.12.2021)
38. <https://www.netapp.com/pdf.html?item=/media/9090-tr4816pdf.pdf> (Zugriff am 20.12.2021)
39. [https://cloud.google.com/storage/docs/storage-classes#available\\_storage\\_classes](https://cloud.google.com/storage/docs/storage-classes#available_storage_classes) (Zugriff am 20.12.2021)
40. <https://cloud.google.com/storage/docs/bucket-locations#location-r> (Zugriff am 20.12.2021)
41. <https://cloud.google.com/solutions/sap/docs/sap-hana-backint-overview#:~:text=Backint%20for%20SAP%20HANA%20certification> (Zugriff am 20.12.2021)
42. <https://launchpad.support.sap.com/#/notes/2031547> (Zugriff am 20.12.2021)
43. <https://cloud.google.com/storage/docs/best-practices> (Zugriff am 20.12.2021)
44. <https://cloud.google.com/vpc/docs/overview> (Zugriff am 20.12.2021)
45. <https://cloud.google.com/blog/products/sap-google-cloud/set-up-your-network-for-a-seamless-sap-cloud-deployment> (Zugriff am 20.12.2021)

- 
- 46. <https://cloud.google.com/network-connectivity/docs/how-to/choose-product> (Zugriff am 20.12.2021)
  - 47. <https://cloud.google.com/network-connectivity/docs/vpn/concepts/overview> (Zugriff am 20.12.2021)
  - 48. <https://cloud.google.com/network-connectivity/docs/interconnect/concepts/dedicated-overview> (Zugriff am 20.12.2021)
  - 49. <https://cloud.google.com/network-connectivity/docs/interconnect/concepts/partner-overview> (Zugriff am 20.12.2021)
  - 50. <https://cloud.google.com/vpc/network-pricing> (Zugriff am 20.12.2021)
  - 51. <https://cloud.google.com/dns/docs/best-practices> (Zugriff am 20.12.2021)
  - 52. [https://cloud.google.com/load-balancing/docs/choosing-load-balancer#flow\\_chart](https://cloud.google.com/load-balancing/docs/choosing-load-balancer#flow_chart) (Zugriff am 20.12.2021)
  - 53. <https://cloud.google.com/infrastructure> (Zugriff am 20.12.2021)
  - 54. <https://cloud.google.com/products/security-and-identity> (Zugriff am 20.12.2021)
  - 55. <https://cloud.google.com/architecture/identity/reference-architectures> (Zugriff am 20.12.2021)
  - 56. [https://cloud.google.com/architecture/identity/reference-architectures#external\\_idaaS\\_as\\_idP\\_and\\_authoritative\\_source](https://cloud.google.com/architecture/identity/reference-architectures#external_idaaS_as_idP_and_authoritative_source) (Zugriff am 20.12.2021)
  - 57. <https://cloud.google.com/iam/docs/overview> (Zugriff am 20.12.2021)
  - 58. <https://cloud.google.com/iam/docs/using-iam-securely> (Zugriff am 20.12.2021)
  - 59. <https://cloud.google.com/products/operations> (Zugriff am 20.12.2021)
  - 60. <https://cloud.google.com/solutions/sap/docs/sap-hana-monitoring-agent-planning-guide> (Zugriff am 20.12.2021)
  - 61. <https://help.sap.com/viewer/4fe29514fd584807ac9f2a04f6754767/2.0.02> (Zugriff am 20.12.2021)
  - 62. <https://cloud.google.com/solutions/sap/docs/netweaver-operations-guide#the-monitoring-agent-for-sap-netweaver> (Zugriff am 20.12.2021)
  - 63. <https://cloud.google.com/support> (Zugriff am 20.12.2021)
  - 64. <https://launchpad.support.sap.com/#/notes/2456406> (Zugriff am 20.12.2021)
  - 65. <https://cloud.google.com/solutions/sap/docs/getting-support> (Zugriff am 20.12.2021)
  - 66. <https://cloud.google.com/solutions/sap/docs/getting-support-from-sap> (Zugriff am 20.12.2021)
  - 67. [https://services.google.com/fh/files/misc/gc\\_consulting\\_cloud\\_sprint\\_sap.pdf](https://services.google.com/fh/files/misc/gc_consulting_cloud_sprint_sap.pdf) (Zugriff am 20.12.2021)
  - 68. <https://cloud.google.com/solutions?hl=de#section-2> (Zugriff am 20.12.2021)
  - 69. <https://cloud.google.com/solutions/sap/docs/architectures/sap-s4hana-on-gcp> (Zugriff am 20.12.2021)
  - 70. <https://cloud.google.com/solutions/sap-on-google-cloud-high-availability.pdf> (Zugriff am 20.12.2021)
  - 71. <https://cloud.google.com/compute/docs/instances/live-migration> (Zugriff am 20.12.2021)
  - 72. <https://cloud.google.com/solutions/sap/docs/sap-hana-ha-planning-guide> (Zugriff am 20.12.2021)
  - 73. [https://cloud.google.com/solutions/sap/docs/architectures/sap-s4hana-on-gcp#high\\_availability\\_and\\_disaster\\_recovery](https://cloud.google.com/solutions/sap/docs/architectures/sap-s4hana-on-gcp#high_availability_and_disaster_recovery) (Zugriff am 20.12.2021)
  - 74. [https://cloud.google.com/solutions/sap/docs/sap-hana-ha-planning-guide#architecture\\_of\\_an\\_sap\\_hana\\_system\\_with\\_host\\_auto-failover](https://cloud.google.com/solutions/sap/docs/sap-hana-ha-planning-guide#architecture_of_an_sap_hana_system_with_host_auto-failover) (Zugriff am 20.12.2021)
  - 75. <https://cloud.google.com/solutions/sap/docs/netweaver-ha-planning-guide> (Zugriff am 20.12.2021)
  - 76. <https://cloud.google.com/solutions/sap-on-google-cloud-disaster-recovery-strategies.pdf> (Zugriff am 20.12.2021)
  - 77. <https://cloud.google.com/solutions/sap-on-google-cloud-backup-strategies.pdf> (Zugriff am 20.12.2021)

78. <https://cloud.google.com/solutions/sap-on-google-cloud-backup-strategies.pdf> (Zugriff am 20.12.2021) auf Seite 17
79. <https://cloud.google.com/compute/docs/disks/snapshots> (Zugriff am 20.12.2021)
80. <https://cloud.google.com/compute/docs/images/sharing-images-across-projects> (Zugriff am 20.12.2021)
81. <https://cloud.google.com/compute/docs/disks/scheduled-snapshots> (Zugriff am 20.12.2021)
82. <https://cloud.google.com/solutions/sap-on-google-cloud-backup-strategies.pdf> (Zugriff am 20.12.2021), Seite 10
83. <https://cloud.google.com/compute/docs/machine-images> (Zugriff am 20.12.2021)
84. <https://cloud.google.com/solutions/sap/docs/sap-hana-backint-overview> (Zugriff am 20.12.2021)
85. <https://www.actifio.com/> (Zugriff am 20.12.2021)
86. <https://www.sap.com/documents/2020/11/86745fb6-c67d-0010-87a3-c30de2ffd8ff.html> (Zugriff am 20.12.2021)
87. <https://cloud.google.com/blog/products/sap-google-cloud/google-cloud-and-sap-demonstrate-scalability-for-financial-services-customers> (Zugriff am 20.12.2021)
88. [https://www.sap.com/idea-place/sap-innovation-awards/submission-details-2021.html?idea\\_id=2376](https://www.sap.com/idea-place/sap-innovation-awards/submission-details-2021.html?idea_id=2376) (Zugriff am 20.12.2021)
89. <https://www.sap.com/bin/sapdxc/inm/attachment.11189/pitch-deck.pdf> (Zugriff am 20.12.2021)
90. <https://cloud.google.com/blog/products/sap-google-cloud/mitigating-memory-errors-for-your-sap-environment> (Zugriff am 20.12.2021)
91. <https://cloud.google.com/blog/products/sap-google-cloud/protect-hana-uptime-with-fast-restart-on-google-cloud> (Zugriff am 20.12.2021)



# SAP S/4 on Google Cloud – Deployment

9

## Zusammenfassung

Das Ziel des Kapitels ist eine detaillierte Übersicht der unterschiedlichen Schritte zu geben, wie ein SAP S/4HANA System auf Google Cloud aufgesetzt wird und welche Konfigurationen möglich sind. Die Zielgruppe für dieses Kapitel sind SAP Architekten und Administratoren, die einen technischen Einblick in die Bereitstellung und das Setup von SAP S/4HANA auf Google Cloud haben möchten. Dabei baut dieses Kapitel auf den Google Cloud Services und Konzepten des vorherigen Kapitels auf und stützt sich auf eine Beispielarchitektur, die zu Beginn definiert wird. Nach der Beschreibung der Bereitstellung aller Komponenten werden zum Ende des Kapitels die Preis- und Abrechnungskonzepte der Google Cloud, die Konfiguration von Sicherungen und Wiederherstellung, die Möglichkeiten für Scripting und Automatisierung und die Best-Practices zu Desaster Recovery aufgezeigt.

## 9.1 Beispielarchitektur für SAP S/4HANA auf Google Cloud

Dieses Kapitel führt eine beispielhafte SAP S/4HANA Architektur auf Google Cloud ein, welche für die folgenden Kapitel herangezogen wird. Mit dieser Architektur werden die Hauptkomponenten aufgezeigt und benannt (bspw. Größe der Applikationsserver und Festplatten, Regionenauswahl, Netzwerkname usw.).

Die Bereitstellung wird ein vollständiges Hochverfügbarkeitscluster und ein volles Desaster Recovery Setup für jeweils Datenbank und Applikationsserver enthalten. Das hier gewählte Betriebssystem ist SLES für SAP 15 SP2, dennoch wäre auch RHEL eine mögliche Alternative und wird ebenfalls vollständig von SAP und Google Cloud unterstützt, zertifiziert und ebenfalls empfohlen.

In der Architektur finden sich alle Komponenten der Architektur, beginnend mit der Auswahl der zwei Regionen für die Desaster Recovery Strategie, nämlich europe-west1 (Belgien) als primäre Region und europe-west4 (Niederlande) als sekundäre Region, dann die Auswahl der zwei Zonen für die Hochverfügbarkeitsverteilung, und zwar Zone c (europe-west1-c, Belgien) als primäre Zone und Zone d (europe-west1-d, Belgien) als sekundäre Zone.

Das VPC heißt *demonetwork* und die Subnetze sind *subnet-europe-west1* und *subnet-europe-west4*.

Die gewählten Compute Engine Maschinen sind drei n1-highmem-32 (mit 32 vCPU und 208 GB RAM) für die SAP HANA Datenbanken mit den Namen: *primaryhana*, *secondaryhana*, *disasterhana*. Jede SAP HANA Compute Engine Instanz bekommt jeweils die folgenden Festplattenspeicher zugewiesen:

- SSD-Festplatte (pd-ssd) mit 834 GB für/hana/data, /hana/log, /hana/shared und/usr/sap
- Standard-Festplatte (pd-hdd) mit 30 GB als Boot-Disk
- Standard-Festplatte (pd-hdd) mit 416 GB für/hana/backup

Für die SAP NetWeaver Applikationsserver werden für die ASCS/ERS Komponente mit dem Linux Cluster drei n2-standard-4 Instanzen (mit 4 vCPU und 16 GB RAM) mit den Namen *nw-ha-vm-1*, *nw-ha-vm-2* und *nw-ha-vm-3* aufgesetzt. Die Applikationsserver selbst sind drei n2-standard-16 (mit 16 vCPU und 64 GB RAM) mit den Namen *nw-app-srv-vm-1*, *nw-app-srv-vm-2*, *nw-app-srv-vm-3*. Jede Applikationsserver-Instanz bekommt folgenden Festplattenspeicher zugewiesen:

- Standard-Festplatte (pd-hdd) mit 54 GB

Eine weitere VM-Instanz kommt als RDP Client für das SAP HANA Studio hinzu (n1-standard-4, mit 4 vCPU und 15 GB RAM) und eine weitere als Jumpbox/SAProuter (n1-standard-4, mit 4 vCPU und 15 GB RAM).

Die Architektur wird im folgenden Architekturschaubild mit allen dazugehörigen Komponenten dargestellt (Abb. 9.1).

---

## 9.2 Planungs- und Bereitstellungs-Checklisten für SAP auf Google Cloud

Vor der Bereitstellung muss sichergestellt werden, dass die sogenannte *Cloud Foundation* bzw. *Landing Zone* aufgesetzt wurde, das bedeutet, dass das Setup auf der Google Cloud in Bezug auf Sicherheit, Compliance und weiterer Eigenschaften bereit für produktive Arbeitslasten ist. Die 25 folgenden Schritte sind generelle Empfehlungen und Hauptaufgaben, die vor und während einer SAP S/4HANA Bereitstellung erledigt

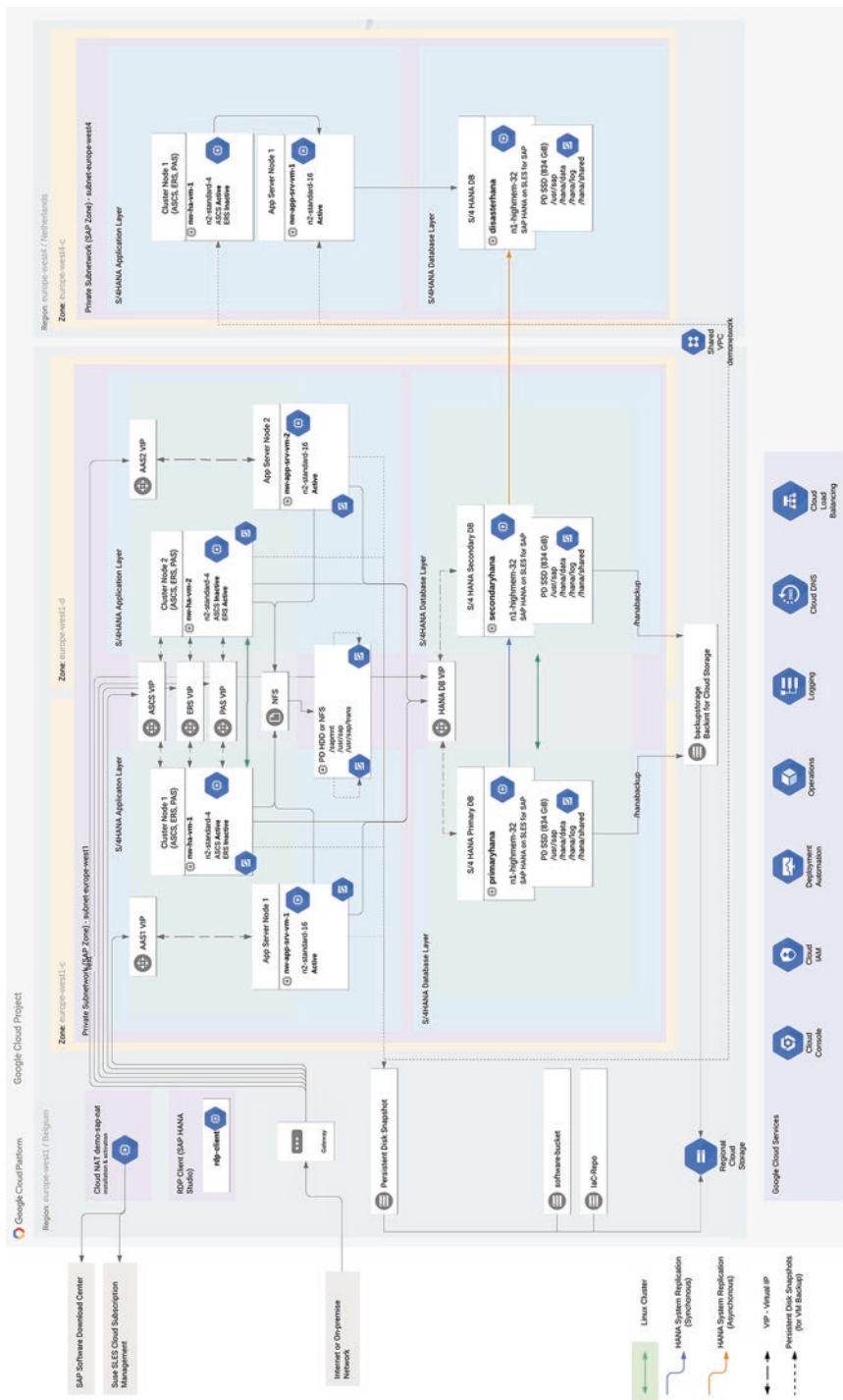


Abb. 9.1 Beispielarchitektur für SAP S/4HANA auf Google Cloud

werden sollten, stellen jedoch nicht in allen Fällen eine vollständige Liste dar. Manche der folgenden Schritte können auch auf unterschiedliche Art und Weise und in manchen Fällen in einer unterschiedlichen Reihenfolge durchgeführt werden.

Generelle Vorbereitung und Checkliste [1] für das Setup in Unternehmen (nicht SAP-spezifisch):

1. Cloud Identity-Konto einrichten und bestätigen
2. Benutzer und Gruppen zum Cloud Identity-Konto hinzufügen
3. Administratorzugriff für die Organisation einrichten
4. Cloud Billing (Abrechnung) aufsetzen
5. Ressourcenhierarchie aufsetzen (siehe Abschn. 8.2)
6. Zugriffe für die Ressourcenhierarchie aufsetzen
7. Support einrichten
8. Netzwerkkonfiguration einrichten
9. Logging und Monitoring aufsetzen
10. Sicherheitskonfiguration für die Organisation einrichten

Nachdem die vorherigen grundlegenden Schritte für die erste Einrichtung der Google Cloud durchgeführt wurden, können die SAP-spezifischen Entscheidungen, Planungen und Konfigurationen für SAP HANA [2] und SAP NetWeaver [3] durchgeführt werden. Zusammengefasst sehen die Schritte wie folgt aus:

11. SAP HANA und SAP Applikationsserver Landschaft und Architektur planen (Versionen etc.)
12. Migrationsempfehlungen befolgen und Migrationskonzept planen
13. Regionen und Zonen für die Architektur wählen
14. Bereitstellungsmodell für die SAP-Systeme wählen (manuelle Installation vs. Automatisierung)
15. Geeignete und zertifizierte Maschinentypen und Betriebssysteme auswählen
16. Netzwerkanforderungen planen und designen
17. Sicherheitskonzept planen und designen
18. Speicher, Festplattenspeicher, Dateisystem und Konfigurationen auswählen
19. Empfohlen: SAP HANA Fast Restart implementieren
20. Sicherungsstrategie für die SAP Landschaft wählen
21. Hochverfügbarkeitskonzept für die SAP Landschaft planen
22. Desaster Recovery Konzept für die SAP Landschaft planen
23. Monitoring und Alerting für die SAP Landschaft verwenden
24. Integrationsschnittstellen für SAP NetWeaver aktivieren
25. Load-Balancing und Skalierung aufsetzen

## 9.3 Google Cloud Account, Netzwerk und Security Setup

Bevor die SAP Komponenten der geplanten Landschaft in der Google Cloud aufgesetzt werden können, müssen die grundlegenden Basiskonfigurationen wie der Google Cloud Account, die Organisationshierarchie, die Projekte und das Netzwerkkonstrukt mit allen benötigten Sicherheitsrichtlinien (Identity und Access Management) eingerichtet werden. Die Schritte für ein beispielhaftes Setup werden in den folgenden Unterkapiteln erläutert.

### 9.3.1 Setup des Google Accounts, Abrechnung und Identity und Access Management

#### 1. Setzen Sie einen Google Account (Identity Account) auf:

Für die Erstellung des Identity Account wird eine E-Mail-Adresse, ein Nutzerkonto für den ersten Super-Admin-Nutzer und die Domain des Unternehmens benötigt. Nach der Anmeldung erstellt Google Cloud den Stammknoten der Ressourcenhierarchie, auch Organisationsressource genannt.

#### 2. Fügen Sie Nutzer und Gruppen dem Identity Account hinzu:

Hier erstellen Sie die Nutzer und Gruppen, die an den folgenden Aktivitäten der Checkliste und an der Bereitstellung der Systeme beteiligt sein werden.

#### 3. Setzen Sie den Administratorzugriff für die Organisation auf:

Hier wird der Administratorzugriff eingerichtet, um dem Administrator eine zentrale Übersicht und Kontrolle über die Cloud-Ressourcen der Organisation zu geben.

#### 4. Konfigurieren Sie die Abrechnung (Billing):

Stellen Sie sicher, dass die Abrechnung für Ihr Projekt aktiviert ist. Über die Navigation in der Console kommen Sie direkt in “Abrechnung”. Wenn Sie nur einen Cloud Billing Account haben, dann kommen Sie direkt auf die Übersichtsseite. Bei mehreren Cloud Billing Accounts öffnet sich ein Fenster mit dem Text, welcher Account verknüpft ist. Wenn Sie keinen Account haben, bekommen Sie die Meldung, um einen Account zu verknüpfen.

#### 5. Setzen Sie die Ressourcenhierarchie auf:

Hier wird eine Ressourcenhierarchie für Ihre Organisation mit Ordnern und Projekten angelegt (siehe Abschn. 8.2). Die für die SAP Landschaft benötigten Projekte und die Ordnerstruktur sollten hier direkt angelegt werden.

#### 6. Setzen Sie Sicherheitsrichtlinien und Zugriffe für die Ressourcenhierarchie auf:

Daraufhin wird die Zugriffssteuerung für die IAM-Ressourcenhierarchie angelegt und die Rollen und Richtlinien an die jeweiligen Nutzer und Gruppen zugewiesen.

#### 7. Konfigurieren Sie den Support:

Hier sollten Sie eine Option für den Support für Ihre Google Cloud Projekte auswählen (siehe Abschn. 8.3.6) und einrichten.

Bevor die nachfolgenden Schritte der nächsten Kapitel durchgeführt werden, sollten Sie immer sicherstellen, dass alle Konfigurationen und Bereitstellungen in Ihrem für die SAP Landschaft definierten Zielprojekt umgesetzt werden.

### 9.3.2 Setup der Shared VPC, Subnetze, Firewallregeln, Cloud NAT und Cloud DNS

Im Folgenden werden alle Netzwerkkomponenten mit Shared VPC, Subnetzen, Firewallregeln, dem Cloud NAT und dem Cloud DNS konfiguriert.

#### 9.3.2.1 Setup des Shared VPC und Subnetze

Zuerst wird ein Shared VPC Netzwerk erstellt, entweder über die Google Cloud Console oder über die Befehlszeile [4] mit den folgenden Befehlen:

1. Erstellen Sie das Netzwerk **demonetwork** mit folgendem Befehl:  
*gcloud compute networks create demonetwork --subnet-mode=custom*

2. Erstellen Sie das Subnetzwerk in der primären Region namens **subnet-europe-west1** und eines in der sekundären Region namens **subnet-europe-west4**.

Beispielbefehl für das Subnetzwerk der primären Region:

*gcloud compute networks subnets create subnet-europe-west1 --network=demonetwork --region=europe-west1 --range=1012800/20*

#### 9.3.2.2 Setup der Firewallregeln

Erstellen Sie Firewallregeln, die nur die Zugriffe erlauben, die benötigt werden, und jeden externen Zugriff, der nicht benötigt wird, blockieren. Ein Beispiel für Firewallregeln, die mit folgenden Befehlen in der Cloud Shell angelegt werden können, ist:

---

##### Beispiel

1. *gcloud compute firewall-rules create icmp --network=demonetwork --action=allow --target-tags=icmp,sap-ports --source-ranges=10.128.0.0/20 --rules=tcp,icmp,udp*
2. *gcloud compute firewall-rules create rdp --network=demonetwork --action=allow --target-tags=rdp --source-ranges=<ip.range.ihrer.admins> --rules=tcp:3389*
3. *gcloud compute firewall-rules create sap-ssh --network=demonetwork --action=allow --target-tags=sap-ssh --source-ranges=<ip.range.ihrer.admins> --rules=tcp:22* ◀

In einer produktiven Unternehmenslandschaft werden wesentlich mehr Firewallregeln benötigt und diese werden durch die Sicherheitsrichtlinien Ihrer Organisation definiert. Für SAP Landschaften gibt es ebenfalls einige Empfehlungen, die der Dokumentation entnommen werden können [5].

Spezifizieren Sie Netzwerk-Tags für die VM-Instanzen. Damit können Firewallregeln und Routen auf bestimmte VM-Instanzen angewendet werden. Wenn Deployment Manager Templates (siehe Abschn. 9.7) genutzt werden, dann können diese wie folgt definiert werden: networkTag: [TAG] [6].

Wenn Sie VMs ohne externe IP-Adresse erstellen, dann spezifizieren Sie in Deployment Manager Templates den Parameter publicIP: No.

### 9.3.2.3 Setup von Cloud NAT

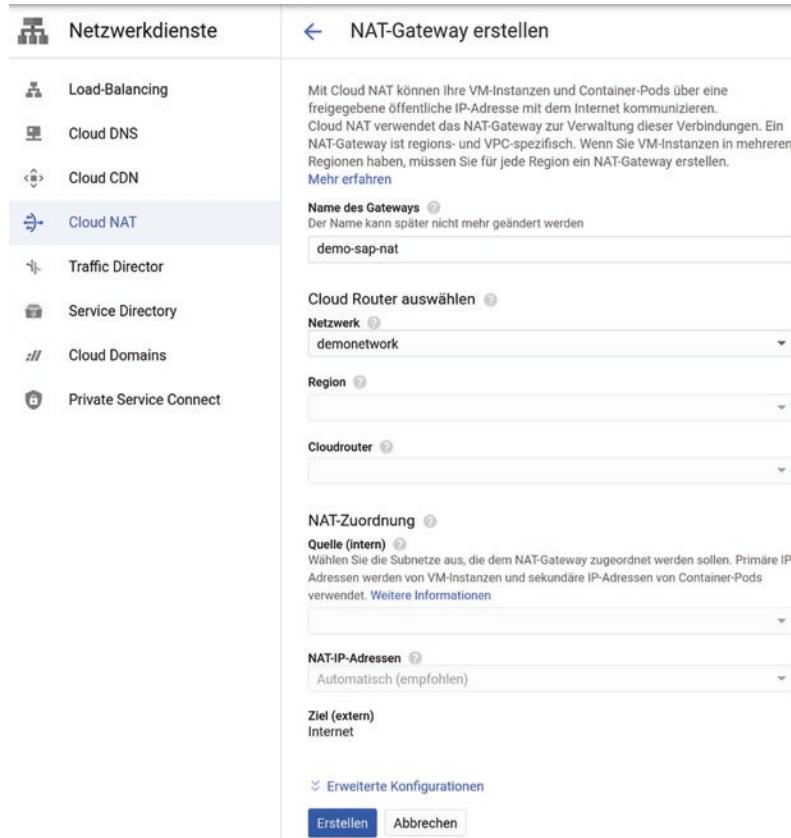
Es ist empfohlen ein Cloud NAT Gateway zu nutzen, um VM-Instanzen nicht dem Internet zu exponieren und ohne externe IP-Adressen aufzusetzen, aber trotzdem die Möglichkeit für Software Updates oder Betriebssystem Registrierungen über das Internet zu erhalten. Folgen Sie diesen Schritten [7], um Cloud NAT für Ihre Landschaft aufzusetzen:

1. Öffnen Sie die Cloud Console und links das Navigationsmenü und wählen Netzwerk Services → Cloud NAT.
2. Klicken Sie auf “NAT Gateway erstellen”.
3. Geben Sie den gewünschten NAT Gateway Namen ein.
4. Wählen Sie das definierte VPC-Netzwerk *demonetwork* aus.
5. Setzen Sie die Google Cloud Region für das Gateway.
6. Wählen oder erstellen Sie einen Cloud Router in der Region (Abb. 9.2).
7. Unter “NAT-Zuordnung” werden die IP-Adressen der Subnetze zugeordnet, die auf öffentliche IP-Adressen zugreifen möchten, um die NAT-IP-Adressen zu binden.
8. Klicken Sie auf “Erweiterte Konfigurationen”.
9. Wählen Sie “Übersetzung und Fehler” im Bereich “Stackdriver Logging”.
10. Definieren Sie die “Mindestanzahl an Ports pro VM-Instanz” und “Zeitlimits für Protokollverbindungen” falls benötigt.
11. Klicken Sie auf “Erstellen”.

### 9.3.3 Setup von Cloud DNS

In diesem Abschnitt wird der Cloud DNS Eintrag für die Virtuelle IP-Adresse (VIP) der SAP Landschaft konfiguriert, sodass die Systeme jeweils über den DNS-Namen angesprochen werden können.

1. Bevor der Eintrag erstellt wird, muss noch ein Metadaten Eintrag im Projekt erstellt werden, sodass das DNS global repliziert wird. Der Befehl für gcloud ist: *gcloud compute project-info add-metadata --metadata=VmDnsSetting=GlobalOnly*
2. Öffnen Sie das Navigationsmenü in der Console und wählen Netzwerkdienste → Cloud DNS.
3. Klicken Sie im angezeigten Fenster namens Zonen auf “Zone erstellen” um eine neue DNS-Zone für Ihr Projekt zu erstellen.



**Abb. 9.2** Konsolenscreenshot: Cloud NAT-Gateway erstellen

4. Tragen Sie die folgende Konfiguration ein und klicken Sie auf “Erstellen” (Abb. 9.3):
5. Im nachfolgenden Fenster klicken Sie auf “Datensatz hinzufügen”.
6. Setzen Sie die folgenden Konfigurationswerte ein (Abb. 9.4):
7. Klicken Sie auf “Erstellen”.

Nachdem das Cloud DNS fertig konfiguriert ist können Sie die Floating-IP reservieren, damit Sie vermeiden, dass diese während der Bereitstellung einer anderen VM-Instanz zugewiesen wird. Der Befehl für die gcloud ist:

```
gcloud compute addresses create alias01 --project=[YOUR PROJECT ID]
--subnet=subnet-europe-west1 --region=europe-west1
--addresses=10.128.0.35
```

Netzwerkdienste

**DNS-Zone erstellen**

Load-Balancing  
Cloud DNS  
Cloud CDN  
Cloud NAT  
Traffic Director  
Service Directory  
Cloud Domains  
Private Service Connect

Eine DNS-Zone ist ein Container mit DNS-Einträgen für dasselbe Suffix des DNS-Namens. In Cloud DNS werden alle Datensätze innerhalb einer verwalteten Zone auf denselben von Google verwalteten autoritativen Nameservern gehostet. [Weitere Informationen](#)

Wenn Sie noch keine Domain haben, erwerben Sie eine über [Cloud Domains](#).

Zonentyp  Privat  Öffentlich

Zonenname \* demo-zone

DNS-Name \* demo.sap.gcp

Beschreibung

Optionen \* Standard (privat)

Netzwerke demonetwork

Ihre private Zone ist in den ausgewählten Netzwerken sichtbar.

Nachdem Sie die Zone erstellt haben, können Sie Ressourceneinträge hinzufügen und festlegen, in welchen Netzwerken die Zone sichtbar sein soll.

**ERSTELLEN** **ABBRECHEN**

BEFEHLSZEILENAQUIVALENT

**Abb. 9.3** Konsolenscreenshot: DNS-Zone erstellen

Netzwerkdienste

**Datensatz erstellen**

Load-Balancing  
Cloud DNS  
Cloud CDN  
Cloud NAT  
Traffic Director  
Service Directory  
Cloud Domains  
Private Service Connect

DNS-Name hana.demo.sap.gcp.

Ressourceneintragstyp A TTL \* 5 TTL-Ein... Minuten

Routingrichtlinie

Standardeintragstyp  Gewichtetes Round Robin VORSCHAU  Geobasiert VORSCHAU

IPv4-Adresse ?

IPv4-Adresse 1 \* 10.128.0.35 Z. B. 192.0.2.91

+ ELEMENT HINZUFÜGEN

**ERSTELLEN** **ABBRECHEN**

**Abb. 9.4** Konsolenscreenshot: DNS Datensatz erstellen

## 9.4 Google Cloud Compute Setup

Jede Funktion und Komponente des SAP-Systems wird auf einer eigenen VM-Instanz installiert, wie schon im Abschn. 8.5.2 in der dezentralisierten Landschaftsübersicht erläutert. Das bedeutet, jeder Applikationsserver, die Datenbank und das SAP HANA Studio so wie der Jump-Server laufen jeweils auf einer anderen VM.

Um eine VM-Instanz und die Konfiguration von Betriebssystem, Hochverfügbarkeitscluster und mehr zu erstellen, kann entweder der automatisierte Weg über Skripte (Deployment Manager Templates, Terraform usw.) gewählt werden. Bei diesem Weg werden auch Installationen der Software und weitere Schritte automatisiert. Alternativ kann auch die Google Cloud Console oder gcloud über die Cloud Shell bzw. das Cloud SDK genutzt werden. Daraufhin erfolgen zur Konfiguration und Installation der Software jedoch manuelle Schritte. Beide Möglichkeiten werden in den folgenden Kapiteln aufgezeigt. Dabei werden die Kapitel in die Bereitstellung von SAP HANA Datenbank und SAP NetWeaver Applikationsserver aufgeteilt. Jedes Unterkapitel wiederum enthält dann Abschnitte über die Voraussetzungen, die automatisierte Bereitstellung, die manuelle Bereitstellung, die Verifizierung der Bereitstellung und spezifische Folgeaufgaben.

### 9.4.1 Bereitstellung der SAP HANA Datenbank

Bevor die SAP HANA Datenbank bereitgestellt und installiert wird, muss ein zertifizierter Maschinentyp in der Dokumentation bzw. dem SAP HANA Hardware Directory gewählt werden, wie zuvor in Abschn. 8.3.1 beschrieben wurde. Im Architekturbeispiel von Abschn. 9.1 wurde für die SAP HANA Instanzen der Maschinentyp n1-highmem-32 gewählt.

#### 9.4.1.1 Voraussetzungen für die Bereitstellung des SAP HANA Hochverfügbarkeitssystems

Prüfen Sie vor der Bereitstellung der VM-Instanzen, ob die folgenden Konfigurationen getroffen wurden:

- Prüfen Sie, dass die Kontingente (Quotas) für das Projekt und für die Größe der geplanten Bereitstellung der gesamten Landschaft groß genug sind (siehe Abschn. 9.5.3).
- Der Google Identity Account, die Organisationshierarchie und die Projekte sind angelegt und konfiguriert, wie in Abschn. 9.3.1 beschrieben.
- Das Shared VPC-Netzwerk, Subnetze, Cloud NAT und Cloud DNS wurden aufgesetzt (siehe Abschn. 9.3.2).
- Die Firewallregeln wurden erstellt (siehe Abschn. 9.3.2).
- Laden Sie die SAP Mediendateien für die Installation der Software (SAP HANA und weitere) herunter und laden diese in einen Google Cloud Storage Bucket mit den

benötigten Berechtigungen hoch. In Abschn. 9.1 wurde dafür der Name *sap-software-bucket* festgelegt.

- Gehen Sie dafür im Menü der Console unter Speicher auf “Cloud Storage”.
- Klicken Sie im Cloud Storage Browser auf “Bucket erstellen” und geben Sie hier den Bucketnamen (*sap-software-bucket*) ein.
- Legen Sie den Speicherort als multi-regional in europäischen Regionen fest wie in der Abbildung veranschaulicht (Abb. 9.5).
- Legen Sie dann die Speicherklasse (siehe Abschn. 8.3.2) und die Zugriffe fest.
- Klicken Sie auf “Erstellen”.
- Sobald das Bucket erstellt wurde, können die Mediendateien hochgeladen werden.
- Erstellen Sie eine VM mit installierten SAP Tools wie SAP HANA Studio für den möglichen Zugriff auf die Datenbank nach der Installation. Dafür können Sie folgenden gcloud-Befehl verwenden, welcher ein von Ihnen zuvor vorkonfiguriertes Image nutzen kann, das die Installationsmedien für das SAP HANA Studio schon bereitstellt:

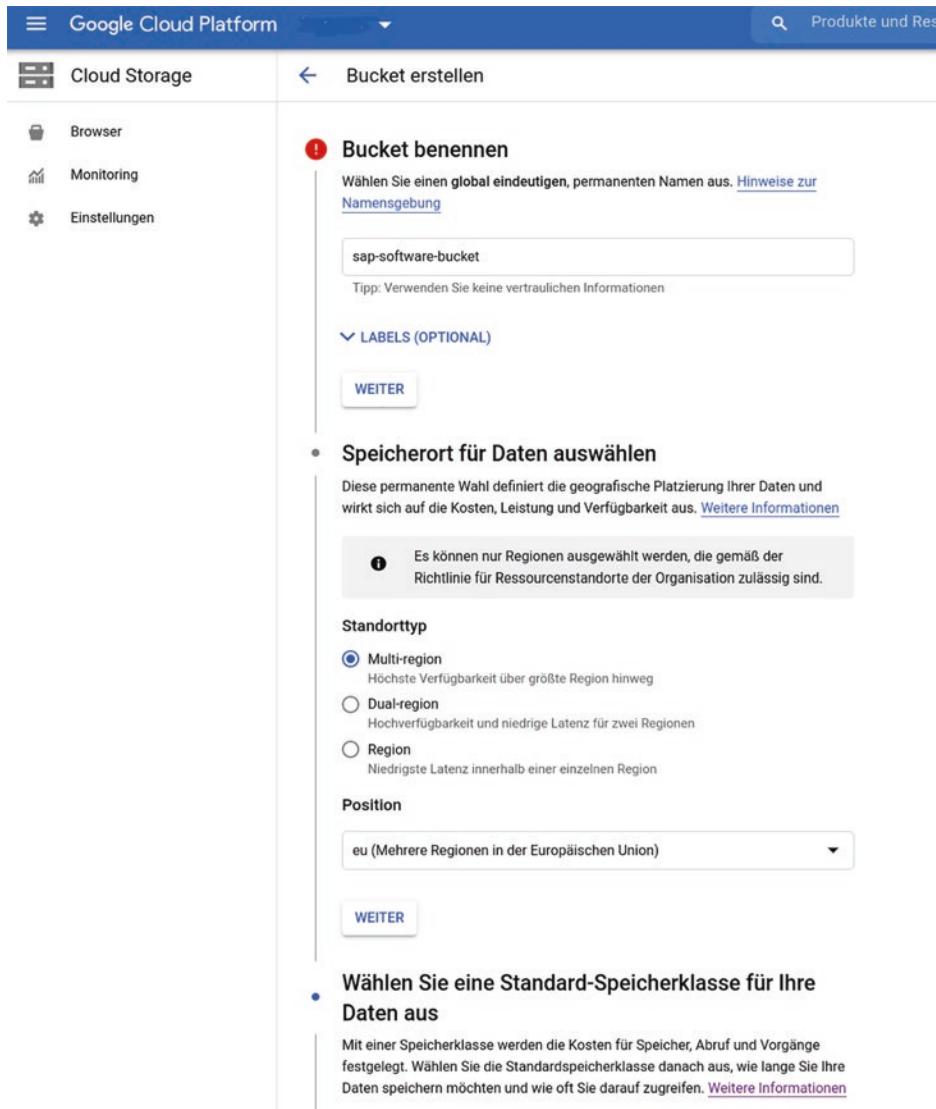
```
gcloud compute instances create rdp-client --zone=$PRIMARY_ZONE  
--machine-type=n1-standard-4 --image-project=<ihr.projekt> --  
image=<ihr.sap.rdp.image> --network=demonetwork --subnet=subnet-$PRIMARY_REGION --tags=rdp,http-server,https-server --boot-disk-type=pd-ssd
```

Nach einer erfolgreichen Bereitstellung der VM-Instanz kann eine RDP Verbindung auf diese Compute Engine Instanz aus der Console erfolgen, indem Sie das Windows Passwort dieser VM neu setzen, irgendwo sicher abspeichern und sich dann mit dem Benutzernamen und Passwort anmelden.

- Erstellen Sie eine weitere VM als SAProuter bzw. Jumpbox, entweder über den manuellen Weg über Compute Engine Menü in der Console oder über gcloud. Mehr Informationen zur Installation und Setup des SAP Support Channels können der Dokumentation entnommen werden [8].

#### **9.4.1.2 Automatisierte Bereitstellung der SAP HANA Hochverfügbarkeitssysteme**

Die SAP HANA Hochverfügbarkeitssysteme können mithilfe der bereitgestellten Templates für Google Cloud Deployment Manager oder Terraform (siehe Abschn. 9.7) aufgesetzt werden. Die automatisierte Bereitstellung gehört zu den Best-Practices von SAP und Google Cloud und spart Zeit und Aufwand. Die gespeicherten Skripte können zu jeder Zeit wiederverwendet werden, was für hochverfügbare und ausfallsichere Landschaften notwendig ist, um schnelle Wiederherstellungen zu ermöglichen. In den folgenden Schritten werden die folgenden Komponenten der Landschaft automatisiert aufgesetzt und konfiguriert:



**Abb. 9.5** Konsolenscreenshot: Google Cloud Storage Bucket erstellen

- VM-Instanzen mit den benötigten Festplatten für jede SAP HANA Instanz
- SAP HANA Installation auf den VM-Instanzen mit einer aktivierten Konfiguration für synchrone System Replikation und Arbeitsspeichervorladung (Preload)
- Automatisches Failover-Cluster
- Automatischer Neustart ist aktiviert (Host Auto-Restart)
- Eine Reservierung der VIP, die von Ihnen spezifiziert wird

- Failover-Support mithilfe von internem TCP/UDP Load-Balancing
- Firewallregeln, die Compute Engine Health-Checks erlauben, sodass Monitoring der VM Instanzen im Cluster möglich ist
- Pacemaker Hochverfügbarkeitscluster Ressourcenmanager
- Google Cloud Fencing Mechanismus

Die folgenden Schritte müssen für dieses Setup [9] durchgeführt werden:

- Öffnen Sie die gcloud Kommandozeile
- Laden Sie das YAML-Template in Ihr Arbeitsverzeichnis:

```
wget  
https://storage.googleapis.com/cloudsapdeploy/deploymentmanager/  
latest/dm-templates/sap_hana_ha_ilb/template.yaml
```

- Editieren Sie das Template mit Ihren spezifischen Einstellungen und Namensgebungen und ändern Sie den Dateinamen, wenn gewünscht. Das Beispiel hier zeigt die SAP HANA Instanzen mit den Maschinengrößen und Regionen, so ausgewählt wie zuvor in Abschn. 9.1 eingeführt:

```
imports:  
- path: https://storage.googleapis.com/sapdeploy/dm-templates/sap_hana_  
ha/sap_hana_ha.py  
resources:  
  
- name: sap_hana_ha  
  type: https://storage.googleapis.com/sapdeploy/dm-templates/sap_hana_  
ha/sap_hana_ha.py  
  properties:  
    primaryInstanceName: primaryhana  
    secondaryInstanceName: secondaryhana  
    primaryZone: europe-west1-c  
    secondaryZone: europe-west1-d  
    instanceType: n1-highmem-32  
    subnetwork: subnet-europe-west1  
    linuxImage: family/sles-15-sp2-sap  
    linuxImageProject: suse-sap-cloud  
    sap_hana_deployment_bucket: sap-software-bucket  
    sap_hana_sid: GCP  
    sap_hana_instance_number: 00  
    sap_hana_sidadm_password: testpw  
    sap_hana_system_password: testpw  
    sap_vip: 10.128.0.35  
    networkTag: icmp,sap-ports,sap-ssh
```

- Nachdem Sie Ihre Anpassungen der Konfigurationsdatei gespeichert haben, können Sie das Bereitstellungsskript ausführen lassen:

```
gcloud deployment-manager deployments create hanaha --config hana_ha_na.yaml
```

Nun dauert es um die 20 bis 30 min bis die automatisierte Bereitstellung, die Installation der SAP HANA Datenbanken und die Konfiguration des Hochverfügbarkeitsclusters durchgeführt wurden.

Befolgen Sie daraufhin die Schritte im Unterkapitel „Verifizierung der Bereitstellung der SAP HANA Hochverfügbarkeitssysteme“.

#### **9.4.1.3 Manuelle Bereitstellung des SAP HANA Hochverfügbarkeitssystems**

Wenn Sie sich dafür entscheiden haben, nicht den empfohlenen Weg über automatisierte Bereitstellungsskripte zu befolgen, können Sie eine manuelle Bereitstellung in der Google Cloud Console und mit gcloud-Befehlszeile durchführen. Dafür müssen die folgenden Schritte durchgeführt werden, die hier nicht im Detail erläutert jedoch der Dokumentation [10] vollständig entnommen werden können:

- Erstellen Sie die benötigten Compute Engine VM-Instanzen
- Installieren Sie SAP HANA
- Verifizieren Sie die Bereitstellung im SAP HANA System (siehe Abschnitt Verifizierung)
- Installieren Sie den Google Cloud Monitoring Agent v2.0
- Deaktivieren Sie SAP HANA Autostart (denn Pacemaker wird SAP HANA starten)
- Erlauben Sie eine SSH Verbindung zwischen der primären und sekundären SAP HANA Maschine
- Aktivieren und konfigurieren Sie SAP HANA System Replikation (HSR)
- Konfigurieren Sie das Google Cloud Load Balancer Failover
- Setzen Sie das Pacemaker Cluster auf
- Testen Sie ein Failover

#### **9.4.1.4 Verifizierung der Bereitstellung der SAP HANA Hochverfügbarkeitssysteme**

Nach einem erfolgreichen manuellen oder automatisierten Setup muss das SAP HANA HA-Cluster in Cloud Logging und auf den VM-Instanzen in SAP HANA selbst überprüft werden. Dafür sollten die folgenden Schritte durchgeführt werden:

## 1. Überprüfen des Cloud Logging

Navigieren Sie über das Menü zu Operations und hier in Cloud Logging unter Logging. Im Log-Explorer wählen Sie als Ressourcentyp “Global” und klicken auf “Abfrage ausführen”.

Sobald in den Logs die Zeile “*INSTANCE DEPLOYMENT COMPLETE*” für beide Instanzen, *primaryhana* und *secondaryhana*, zu sehen ist, ist das Setup und die Konfiguration von SAP HANA auf diesen Compute Engine Instanzen erfolgreich durchgeführt und beendet (siehe Abbildung) (Abb. 9.6).

## 2. Überprüfen der VM-Konfiguration und der SAP HANA Installation

Sie können nun über das Menü in der Console zu Ihren Compute Engine Instanzen navigieren und sich per SSH auf die *primaryhana* anmelden. Geben Sie nun den folgenden Befehl in der SSH Console ein: *top*.

Überprüfen Sie nun, ob der “hdbindexserver” sichtbar ist. Sobald dieser Eintrag auftaucht, ist die Installation beendet und das Clustering wird aufgesetzt (Abb. 9.7).

## 3. Überprüfen Sie nun den Load-Balancer und die Instanzgruppen

Navigieren Sie über das Menü der Console zu den Netzwerkdiensten und dort zum Load-Balancing. Es sollte ein TCP-Load-Balancer für Ihr HA-Cluster konfiguriert sein. Überprüfen Sie in den Instanzgruppen die VMs und ob deren Status als fehlerfrei angezeigt wird. Eine Instanzgruppe zeigt 1/1 an und die derzeit sekundäre (Failover) Instanzgruppe eine 0/1 an.

## 4. Überprüfen Sie die SAP HANA Datenbank mithilfe des SAP HANA Studios

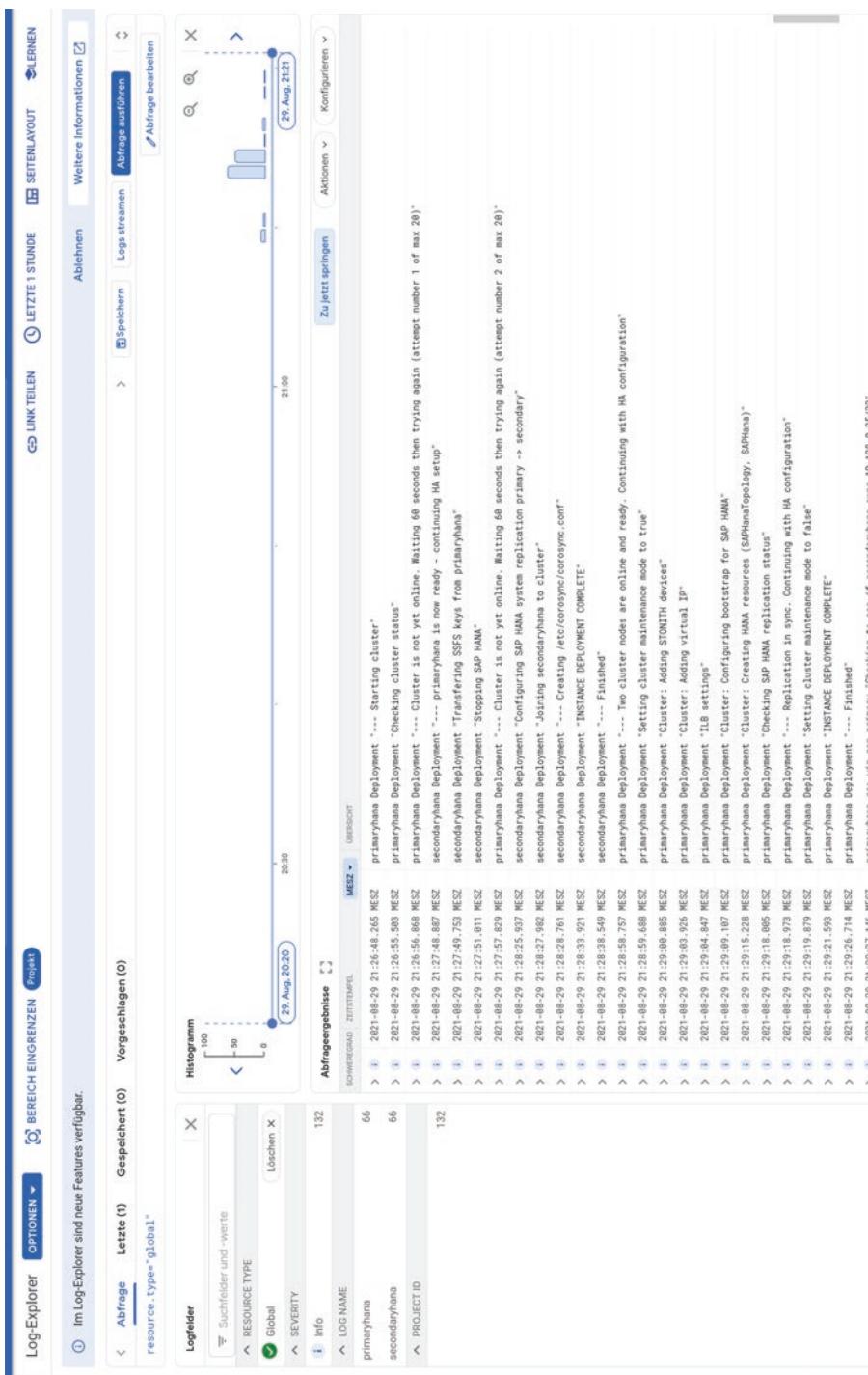
Melden Sie sich über RDP auf der VM-Instanz *rdp-client* an. Zur Bedienung des SAP HANA Studios und Überprüfen der SAP HANA Datenbank kann die Dokumentation [11] befolgt werden.

## 5. Durchführung eines Failover-Test

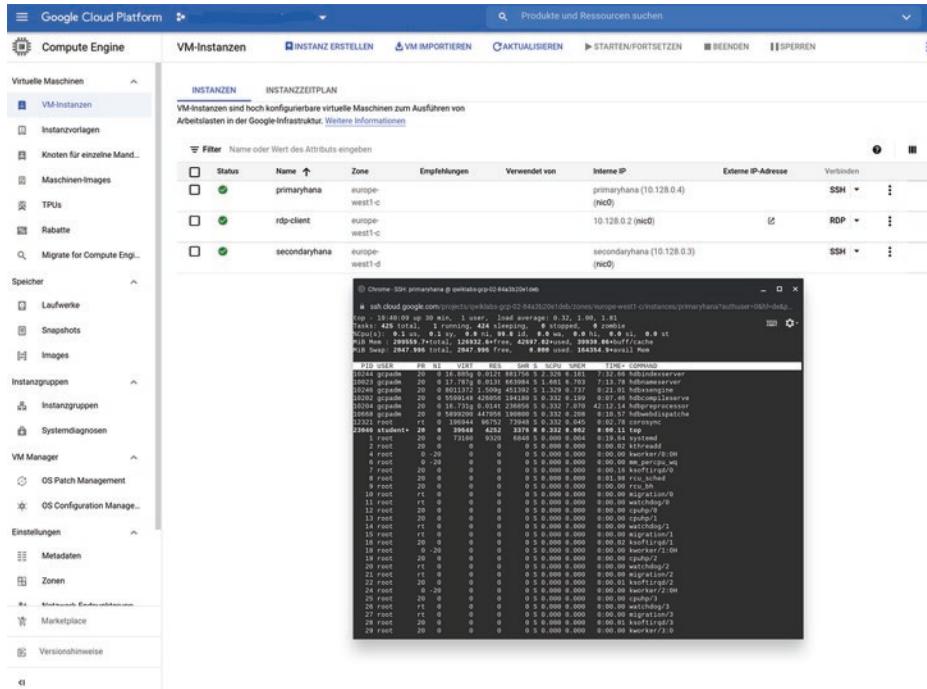
Mit diesem Befehl können Sie den Status des SUSE Pacemaker Clusters überprüfen, sowie welcher Server der primäre und welcher der sekundäre ist: *crm status* (Abb. 9.8).

Das folgende Kommando zeigt, dass die Bereitstellung eine standardmäßige Festplattenkonfiguration für SAP HANA benötigt. Jede Festplattengröße basiert auf der Arbeitsspeichermenge, die der Maschine bei der Installation zugewiesen wurde: *df -h* (Abb. 9.9).

Wählen Sie sich nun per SSH in die primäre VM-Instanz ein und führen Sie die folgenden Befehle durch:



**Abb. 9.6** Konsolenscreenshot: Google Cloud Logging



**Abb. 9.7** Konsolenscreenshot: Überprüfen des SAP HANA Status auf Google Cloud Compute Engine

```

sudo su -
zypper install net-tools-deprecated
ifconfig eth0 down

```

Navigieren Sie dann im Menü zu Google Cloud Logging. Im Cloud Log Explorer sollten Sie nun sehen können, wie die *secondaryhana* Instanz als primäre Instanz übernimmt (Abb. 9.10):

Sind diese Schritte zur Verifizierung erfolgreich durchgeführt worden, dann hat die Bereitstellung fehlerfrei funktioniert.

#### 9.4.1.5 Folgeaufgaben nach der Installation der SAP HANA Systeme

Bevor die SAP HANA Instanzen verwendet werden, sollten nach der Bereitstellung die folgenden Aufgaben erledigt werden [12]:

- Temporäre Passwörter ändern
- SAP HANA Software und Patches aktualisieren
- Bei Bedarf weitere Komponenten installieren oder Konfigurationen durchführen
- Sicherung und Wiederherstellung konfigurieren (siehe Abschn. 9.6)
- SAProuter für den Supportkanal mit SAP konfigurieren [13]

```
primaryhana:~ # crm status
Stack: corosync
Current DC: primaryhana (version 1.1.15-23.9.1-e174ec8) - partition with quorum
Last updated: Thu Jun 18 00:17:09 2020
Last change: Thu Jun 18 00:17:03 2020 by root via crm_attribute on primaryhana

2 nodes configured
8 resources configured

Online: [ primaryhana secondaryhana ]

Full list of resources:

  STONITH-primaryhana    (stonith:external/gcpstonith):  Started secondaryhana
  STONITH-secondaryhana  (stonith:external/gcpstonith):  Started primaryhana
  Resource Group: g-primary
    rsc_vip_int-primary      (ocf::heartbeat:IPAddr2):      Started primaryhana
    rsc_vip_gcp-primary      (ocf::gcp:alias):      Started primaryhana
  Clone Set: cln_SAPHanaTopology_GCP_HDB00 [rsc_SAPHanaTopology_GCP_HDB00]
    Started: [ primaryhana secondaryhana ]
  Master/Slave Set: msl_SAPHana_GCP_HDB00 [rsc_SAPHana_GCP_HDB00]
    Masters: [ primaryhana ]
    Slaves: [ secondaryhana ]

primaryhana:~ #
```

**Abb. 9.8** Konsolenscreenshot: Überprüfen des SAP HANA Hochverfügbarkeitscluster

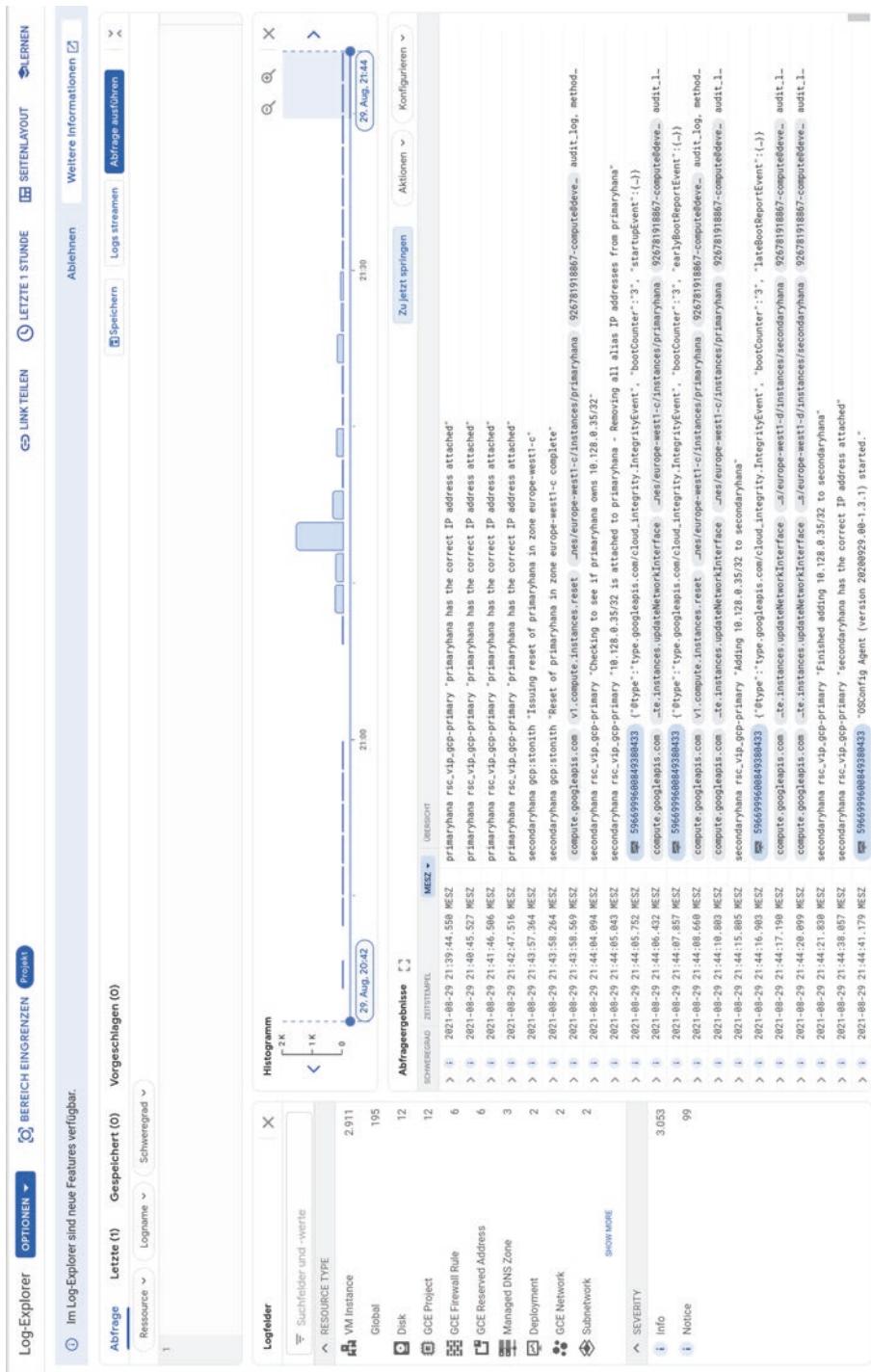
```
Online: [ primaryhana secondaryhana ]

Full list of resources:

  STONITH-primaryhana    (stonith:external/gcpstonith):  Started secondaryhana
  STONITH-secondaryhana  (stonith:external/gcpstonith):  Started primaryhana
  Resource Group: g-primary
    rsc_vip_int-primary      (ocf::heartbeat:IPAddr2):      Started primaryhana
    rsc_vip_gcp-primary      (ocf::gcp:alias):      Started primaryhana
  Clone Set: cln_SAPHanaTopology_GCP_HDB00 [rsc_SAPHanaTopology_GCP_HDB00]
    Started: [ primaryhana secondaryhana ]
  Clone Set: msl_SAPHana_GCP_HDB00 [rsc_SAPHana_GCP_HDB00] (promotable)
    Masters: [ primaryhana ]
    Slaves: [ secondaryhana ]

primaryhana:~ # df -h
Filesystem              Size  Used Avail Use% Mounted on
/devtmpfs                103G   8.0K  103G  1% /dev
tmpfs                     154G   54M  154G  1% /dev/shm
tmpfs                     103G   18M  103G  1% /run
tmpfs                     103G     0  103G  0% /sys/fs/cgroup
/dev/sda3                  30G   5.5G  25G  19% /
/dev/sda2                 20M   3.6M  17M  18% /boot/efi
/dev/mapper/vg_hana-shared  204G   35G  170G  18% /hana/shared
/dev/mapper/vg_hana-sap     32G   264M  32G  1% /usr/sap
/dev/mapper/vg_hana-data    496G   6.7G  490G  2% /hana/data
/dev/mapper/vg_hana-log     102G   5.3G  97G  6% /hana/log
/dev/mapper/vg_hanabackup-backup 416G   5.7G  411G  2% /hanabackup
tmpfs                      21G     0   21G  0% /run/user/900
tmpfs                      21G     0   21G  0% /run/user/472
tmpfs                      21G     0   21G  0% /run/user/1001
primaryhana:~ #
```

**Abb. 9.9** Konsolenscreenshot: Überprüfen der Festplatten- und Dateisystemkonfiguration von SAP HANA



**Abb. 9.10** Konsolenscreenshot: Überprüfen des Failover im Cloud Logging

## 9.4.2 Bereitstellung der SAP NetWeaver Applikationsserver

Für die SAP NetWeaver Applikationsserver müssen die benötigten Größen ebenfalls anhand der zertifizierten Größen überprüft werden. Der folgende Abschnitt fokussiert das Setup von SAP NetWeaver Systemen auf Linux [14] (die Windows Installationen sind der offiziellen Dokumentation [15] zu entnehmen).

### 9.4.2.1 Voraussetzungen für die Bereitstellung der SAP NetWeaver Applikationsserver

- Prüfen Sie, dass die Kontingente (Quotas) für das Projekt und für die Größe der geplanten Bereitstellung der gesamten Landschaft groß genug sind.
- Der Google-Identity Account, die Organisationshierarchie und die Projekte sind angelegt und konfiguriert wie in Abschn. 9.3.1 beschrieben.
- Das Shared VPC-Netzwerk, Subnetze, Cloud NAT und Cloud DNS wurden aufgesetzt (siehe Abschn. 9.3.2).
- Die Firewallregeln wurden erstellt (siehe Abschn. 9.3.2).
- Laden Sie die SAP Mediadateien für die Installation der Software in einen Google Cloud Storage Bucket mit den benötigten Berechtigungen hoch.

### 9.4.2.2 Automatisierte Bereitstellung der SAP NetWeaver Applikationsserver

Die SAP NetWeaver Applikationsserver können mithilfe der bereitgestellten Templates für Google Cloud Deployment Manager oder Terraform (siehe Abschn. 9.7) aufgesetzt werden. Die automatisierte Bereitstellung gehört zu den Best-Practices von SAP und Google Cloud und spart Zeit und Aufwand. Die gespeicherten Skripte können zu jeder Zeit wiederverwendet werden, was für hochverfügbare und ausfallsichere Landschaften notwendig ist, um schnelle Wiederherstellungen zu ermöglichen. In den folgenden Schritten werden die folgenden Komponenten der Landschaft automatisiert aufgesetzt und konfiguriert:

- Zwei Host-VMs für die SAP NetWeaver Applikationsserver
- Zwei Host-VMs, eine mit einem aktiven SAP Central Services (ASCS) und eine mit einem aktiven Standalone Enqueue Server (ERS)
- Clusterressourcen-Manager von Pacemaker für Hochverfügbarkeit
- STONITH-Fencing-Mechanismus
- Automatischer Neustart der fehlgeschlagenen Instanz als neue sekundäre Instanz

Die folgenden Schritte müssen für dieses Setup durchgeführt werden:

- Öffnen Sie die gcloud Befehlszeile
- Laden Sie das YAML-Template in Ihr Arbeitsverzeichnis herunter

```
wget  
https://storage.googleapis.com/cloudsapdeploy/deploymentmanager/  
latest/dm-templates/sap_nw/template.yaml
```

- Editieren Sie das Template mit Ihren spezifischen Einstellungen [16] und Namensgebungen und ändern Sie den Dateinamen, wenn gewünscht. Das Beispiel hier zeigt die ASCS und ERS Instanzen mit den Maschinengrößen und Regionen wie zuvor in Abschn. 9.1 eingeführt:

```
resources:  
- name: sap_nw_node_1  
  type:  
    https://storage.googleapis.com/cloudsapdeploy/deploymentmanager/latest/  
    dm-templates/sap_nw/sap_nw.py  
  properties:  
    instanceName: nw-ha-vm-1  
    instanceType: n2-standard-4  
    zone: europe-west1-c  
    subnetwork: subnet-europe-west1  
    linuxImage: family/sles-15-sp2-sap  
    linuxImageProject: suse-sap-cloud  
    usrsapSize: 15  
    sapmntSize: 15  
    swapSize: 24  
    networkTag: cluster-ntwk-tag,allow-health-check  
    serviceAccount: limited-roles@example-project-123456.iam.  
gserviceaccount.com  
- name: sap_nw_node_2  
  type:  
    https://storage.googleapis.com/cloudsapdeploy/deploymentmanager/latest/  
    dm-templates/sap_nw/sap_nw.py  
  properties:  
    instanceName: nw-ha-vm-2  
    instanceType: n2-standard-4  
    zone: europe-west1-d  
    subnetwork: subnet-europe-west1  
    linuxImage: family/sles-15-sp2-sap  
    linuxImageProject: suse-sap-cloud  
    usrsapSize: 15  
    sapmntSize: 15  
    swapSize: 24  
    networkTag: cluster-ntwk-tag,allow-health-check  
    serviceAccount: limited-roles@example-project-123456.iam.  
gserviceaccount.com
```

Wie im Skript ersichtlich ist werden hier zwei VMs vom Maschinentyp n2-standard-4 mit jeweils 4 vCPU und 16 GB RAM mit dem Betriebssystem SLES 15 SP2 erstellt.

- Nun erfolgt die automatisierte Bereitstellung der beiden Applikationsserver selbst, auch diese muss zweimal hintereinander definiert werden, hier ist das Beispiel für *nw-app-srv-vm-1*.

```
resources:
- name: sap_nw_app
  type: https://storage.googleapis.com/cloudsapdeploy/deploymentmanager/
latest/dm-templates/sap_nw/sap_nw.py
#
# By default, this configuration file uses the latest release of the
deployment
# scripts for SAP on Google Cloud. To fix your deployments to a
specific release
# of the scripts, comment out the type property above and uncomment
the type property below.
#
# type: https://storage.googleapis.com/cloudsapdeploy/
deploymentmanager/202103310846/dm-templates/sap_nw/sap_nw.py
properties:
instanceName: nw-app-srv-vm-1
instanceType: n2-standard-16
zone: europe-west1-c
subnetwork: subnet-europe-west1
linuxImage: family/sles-15-sp2-sap
linuxImageProject: suse-sap-cloud
usrsapSize: 15
sapmntSize: 15
swapSize: 24
```

- Erstellen Sie beide Deployments mit den beiden Templates über folgenden Befehl:

```
gcloud deployment-manager deployments create [DEPLOYMENT_NAME] --config
sap_nw_app.yaml
```

Führen Sie danach die Schritte der Unterkapitel „Verifizierung der Bereitstellung der SAP NetWeaver Applikationsserver“ und „Konfiguration des Hochverfügbarkeits-clusters“ durch.

#### 9.4.2.3 Manuelle Bereitstellung der SAP NetWeaver Applikationsserver

Falls Sie sich nicht für den empfohlenen, automatisierten Bereitstellungsweg entscheiden, können Sie auch eine manuelle Bereitstellung der VM-Instanzen durchführen. Dies erfolgt in der Console oder in gcloud:

In der Google Cloud Console navigieren Sie im Menü zu Compute Engine:

1. Klicken Sie auf „Image“ in der linken Menüleiste
2. Wählen Sie hier das Image (SLES für SAP 15 SP2 in diesem Architekturbeispiel) und klicken Sie auf „Instanz erstellen“, danach geben Sie folgende Werte ein:
  - Name der VM-Instanz
  - Wählen Sie die Region und Zone
  - Wählen Sie den Maschinentypen, in diesem Beispiel n2-standard-4
  - Konfigurieren Sie eine Boot Festplatte (hier mit 20 GB) und dem Betriebssystem SLES 15 SP2
  - Wählen Sie den Service Account und die Zugriffsbereiche aus (Wenn Sie ein benutzerdefiniertes Dienstkontakt verwenden, das den Zugriff auf Google Cloud-Ressourcen einschränkt, wählen Sie den vollen Zugriff auf alle APIs) (Abb. 9.11)

**← Instanz erstellen**

Wählen Sie eine der Optionen aus, um eine VM-Instanz zu erstellen:

- + Neue VM-Instanz**  
Einzelne VM-Instanz neu erstellen
- [+] Neue VM-Instanz aus Vorlage erstellen**  
Einzelne VM-Instanz aus einer vorhandenen Vorlage erstellen
- [+] Neue VM-Instanz von Maschinen-Image**  
Einzelne VM-Instanz aus einem vorhandenen Maschinen-Image erstellen
- Marketplace**  
Sofort einsatzbereite Lösung auf VM-Instanz bereitstellen

**Name \***  
nw-hana-vm-1

**Labels** [?](#)  
[+ ADD LABELS](#)

**Region \***  
europe-west1 (Belgien) [?](#)  
Die Region kann später nicht mehr geändert werden

**Zone \***  
europe-west1-c [?](#)  
Die Zone kann später nicht mehr geändert werden

**Maschinenkonfiguration**

**Maschinenfamilie**

**ALLGEMEINER ZWECK** **COMPUTING-OPTIMIERT**

Maschinentypen für gängige Arbeitslasten, optimiert für Kosten und hohe Flexibilität

**Reihe**  
N2

Mit den CPU-Plattformen Intel Cascade Lake und Ice Lake

**Maschinentyp**  
n2-standard-4 (4 vCPU, 16 GB Arbeitsspeicher)

vCPU	Memory
4	16 GB

**CPU-PLATTFORM UND GPU**

**Anzeigegerät**  
Aktivieren Sie diese Option, um Bildschirmaufnahme- und Aufzeichnungstools zu verwenden.

Anzeigegerät aktivieren

**Vertraulicher VM-Dienst** [?](#)

Confidential Computing-Dienst auf dieser VM-Instanz aktivieren.

**Abb. 9.11** Konsolenscreenshot: Instanz erstellen

3. Expandieren Sie den Konfigurationsabschnitt für „Netzwerk, Laufwerke, Sicherheit, Verwaltung, Einzelne Maschinen“
4. Im Abschnitt „Netzwerk“ wählen Sie das Netzwerk und Subnetzwerk
5. Wenn ein NAT-Gateway genutzt wird, fügen Sie unter „Netzwerk“ → „Netzwerk-Tags“ das Tag hinzu, das beim Einrichten der Route für den Traffic über das Gateway definiert wurde
6. In Verwaltung → „Verfügbarkeitsrichtlinie“ wählen Sie für:
  - Abrufbarkeit → Aus (empfohlen)
  - Bei Hostwartung → VM-Instanz migrieren (empfohlen)
  - Automatischer Neustart → An (empfohlen)
7. Im Abschnitt „Laufwerke“ wählen Sie unter „Zusätzliche Laufwerke“ → „Neues Laufwerk Hinzufügen“
  - Fügen Sie SAP NetWeaver Binaries und das Swap-Laufwerk hinzu (siehe Abschn. 8.3.2)
  - Geben Sie den Namen, den Laufwerkstyp (als Nichtflüchtiger Standardspeicher), die Quelle (als Leeres Laufwerk) und die Größe ein
8. Klicken Sie auf „Erstellen“ um die Instanz zu erstellen

Alternativ kann in der Console rechts neben dem „Erstellen“-Button das Befehlszeilenäquivalent für die gcloud Kommandozeile angezeigt und kopiert werden. Für das hier beschriebene System würde das Kommando wie folgt aussehen:

```
gcloud compute instances create nw-ha-vm-1 --project=<ihr.project.name> --zone=europe-west1-c --machine-type=n2-standard-4 --network-interface=network-tier=PREMIUM,subnet=subnet-europe-west1 --maintenance-policy=MIGRATE --service-account=<ihr.automatischer.service.account>-compute@developer.gserviceaccount.com --scopes=https://www.googleapis.com/auth/devstorage.read_only,https://www.googleapis.com/auth/logging.write,https://www.googleapis.com/auth/monitoring.write,https://www.googleapis.com/auth/servicecontrol,https://www.googleapis.com/auth/service.management.readonly,https://www.googleapis.com/auth/trace.append --create-disk=auto-delete=yes,boot=yes,device-name=nw-ha-vm-1,image=projects/suse-sap-cloud/global/images/sles-15-sp2-sap-v20210604,mode=rw,size=20,type=projects/<ihr.project.name>/zones/europe-west1-c/diskTypes/pd-balanced --create-disk=device-name=nw-disk,mode=rw,name=nw-disk,size=54,type=projects/<ihr.project.name>/zones/europe-west1-c/diskTypes/pd-balanced --no-shielded-secure-boot --shielded-vtpm --shielded-integrity-monitoring --reservation-affinity=any
```

Nachdem die VM-Instanz bereitgestellt wurde, melden Sie sich über SSH auf der VM an. Nun müssen die Festplatten und nichtflüchtiger Speicher konfiguriert und gemounted werden. Die folgenden Befehle müssen für alle Festplatten, die angelegt werden sollen, wiederholt werden [17]:

1. Melden Sie sich als Super-User an

```
sudo su -
```

2. Erstellen Sie ein physisches Volumen für die Festplatte

```
pvcreate /dev/disk/by-id/google-[DISK]
```

3. Erstellen Sie eine Volumengruppe für die Festplatte

```
vgcreate vg_usrsap /dev/disk/by-id/google-[DISK]
```

4. Erstellen Sie ein logisches Volumen für die Festplatte

```
lvcreate -l 100%FREE -n vol vg_usrsap
```

5. Formatieren Sie die Festplatte mit dem Dateisystem, z. B. single xfs

```
mkfs -t xfs /dev/vg_usrsap/vol
```

6. Aktualisieren Sie die Dateisystemtabelle (fstab)

```
echo "/dev/vg_usrsap/vol/usr/sap xfs defaults,discard,nofail 0 2" >>/etc/fstab
```

7. Erstellen Sie den Mount-Punkt

```
mkdir -p /usr/sap
```

8. Mounten Sie die Festplatte an die VM

```
mount -a
```

9. Prüfen Sie, dass die Festplatten ordnungsgemäß angehängt wurden

```
df -h
```

Formatieren und mounten der Swap-Festplatte:

10. Erstellen Sie ein physisches Volumen für die Swap-Festplatte

```
pvcreate /dev/disk/by-id/google-[DISK]
```

11. Erstellen Sie eine Volumengruppe

```
vgcreate vg_swap/dev/disk/by-id/google-[DISK]
```

12. Erstellen Sie ein logisches Volumen für die Swap-Festplatte

```
lvcreate -l 100%FREE -n vol vg_swap
```

13. Formatieren Sie die Festplatte

```
mkswap/dev/vg_swap/vol
```

14. Aktualisieren Sie fstab

```
echo "/dev/vg_swap/vol none swap defaults,nofail 0 2" >>/etc/fstab
```

15. Mounten Sie die Festplatte an die VM

```
swapon/dev/vg_swap/vol
```

Zuletzt bereiten Sie das Betriebssystem nach der SAP Note Dokumentation [18] vor.

#### **9.4.2.4 Verifizierung der Bereitstellung der SAP NetWeaver Applikationsserver**

Verifizieren Sie die Bereitstellung in Cloud Logging → Log Explorer unter Ressourcentyp → Global und verbinden Sie sich per SSH auf die VM-Instanz.

- Prüfen Sie die Dateisysteme/-verzeichnisse mit dem Befehl:

```
df -h
```

- Prüfen Sie, ob das Auslagerungsverzeichnis ebenfalls erstellt wurde:

```
cat/proc/meminfo | grep Swap
```

- Installieren Sie den Cloud Logging Agent und den Monitoring Agent für SAP NetWeaver.
- Installieren Sie daraufhin SAP NetWeaver – mehr Informationen dazu siehe SAP Help Portal [19] und im SAP NetWeaver Master Guide [20].

#### **9.4.2.5 Konfiguration des Hochverfügbarkeitsclusters**

Wenn die SAP NetWeaver Applikationsserver in einem Hochverfügbarkeitscluster aufgebaut werden sollen, ist das Setup komplexer und benötigt weitere Schritte zur Konfiguration des Linux Pacemaker Cluster, des Load-Balancing und mehr. Im

Folgenden werden die Schritte [21] nach der Erstellung der VM-Instanzen im vorherigen Abschnitt grob erläutert, die Details können der Dokumentation entnommen werden.

1. Firewallregeln erstellen, die den Zugriff auf Host-VMs zulassen, beispielsweise zwischen allen Cluster-VMs und ebenfalls zwischen diesen und dem Jump-Server
2. Aktivieren Sie die Kommunikation der Load-Balancer zwischen den VMs
  - Loggen Sie sich auf jeder VM des Clusters über SSH an, wechseln Sie zum Root-User und führen den folgenden Befehl durch, um lokales Routing zu aktivieren:  
`echo net.ipv4.conf.eth0.accept_local=1 >> /etc/sysctl.conf  
sysctl -p`
  - Erstellen Sie auf jeder VM ein Startskript für die Kommunikation zwischen den VMs, siehe Beispielskript in der Dokumentation [22]
3. Konfigurieren Sie die SSH Schlüssel zwischen den Hosts, also der primären und sekundären VM-Instanz
4. Richten Sie die Dateifreigabelösung (NFS) ein und konfigurieren Sie die zwischen den VMs freigegebenen Verzeichnisse
5. Konfigurieren Sie die Failover-Unterstützung für Cloud Load Balancing

Der interne TCP/UDP-Load-Balancer mit Failover-Unterstützung leitet den Datenverkehr von ASCS- und ERS-Systemen an die aktiven Instanzen (Applikationsserver) in einem SAP NetWeaver-Cluster weiter. Das interne TCP/UDP-Load-Balancing verwendet dazu virtuelle IP-Adressen (VIP), Back-End-Dienste, Instanzgruppen und Systemdiagnosen, um diesen Datenverkehr entsprechend weiterzuleiten.

6. Einrichtung von Pacemaker
7. Konfiguration der Clusterressourcen des Hochverfügbarkeitsclusters für die Infrastruktur
8. ASCS und ERS installieren
9. SAP-Dienste konfigurieren
10. Clusterressourcen für ASCS und ERS konfigurieren
11. Cluster testen

---

## 9.5 Preis- und Abrechnungskonzepte mit Best-Practices

Dieses Kapitel behandelt die Preis- und Abrechnungskonzepte auf Abruf (On-demand bzw. Pay-as-you-go) und mit Rabatten für zugesicherte Nutzung (Committed Use Discount). Die Möglichkeiten von Reservierungen und Kontingenten werden ebenfalls erläutert. Schließlich wird anhand des offiziellen Google Cloud Preiskalkulators das Architekturbeispiel von Kapitel 9.1 gerechnet.

### 9.5.1 On-Demand vs. Rabatte für zugesicherte Nutzung

Google Cloud bietet die Möglichkeit die Nutzung von speziellen Diensten und Arbeitslasten zuzusichern und somit Google eine Zusicherung des Bedarfes zu bestätigen. Diese Zusicherung der Nutzung führt zu signifikanten Kostenreduzierungen für den Kunden anhand von Rabatten [23]. Rabatte für zugesicherte Nutzung gibt es unter anderem auf Compute Engine Ressourcen wie vCPU, RAM, lokale SSDs. Diese Rabatte betragen für die meisten Ressourcenarten bis zu 57 %, bei speicheroptimierten Maschinen (m1- und m2-Maschinentypen) sogar bis zu 70 %. Die Entscheidung wird auf Basis der Compute Engine Maschinentypen und der Nutzungsart getroffen. Dieser Abschnitt hilft bei der Entscheidung welche Systeme mit zugesicherter Nutzung versehen werden sollten.

Rabatte für zugesicherte Nutzung werden in der Google Cloud Console im Menü links unter “Compute Engine → Rabatte → Zusicherung kaufen” eingespeichert und zugesichert. Diese müssen für jeden Maschinentypen extra angelegt werden (n1, n2, n2d und arbeitsspeicheroptimierter Maschinentyp). Die totale Summe der vCPU und RAM, die mit diesem Maschinentypen planmäßig laufen soll, wird hier ebenfalls mitgegeben, wie in der Abbildung veranschaulicht (Abb. 9.12).

Für alle Systeme, die 24/7 betrieben werden sollen, was beispielsweise meistens auf produktive SAP-Systeme zutrifft, wird definitiv empfohlen entweder eine 1-Jahres oder, um bessere Rabatte zu erhalten, eine 3-Jahres zugesicherte Nutzung abzuschließen.

Wenn Sie Sandbox oder Entwicklungssysteme in der Landschaft haben, die nicht 24/7 laufen müssen, dann kann es sinnvoll sein, dass diese mit der Standard On-Demand-Rate auf Abruf und ohne zugesicherte Nutzung bezahlt werden. Die Stunden pro Monat bei welchen die Rentabilitätsschwelle überschritten wird variiert pro Maschinentyp und –größe. In vielen Fällen liegt diese bei um die 200–250 h im Monat. Empfohlen wird, dieses Preisverhältnis für Ihre Maschinen im Preiskalkulator (siehe Abschn. 9.5.4) zu überprüfen. Eine Maschine die wesentlich weniger als 200–250 h pro Monat laufen soll, ist definitiv günstiger mit den Preisen auf Abruf und dadurch auch viel flexibler, da es keine Verpflichtung auf eine bestimmte Jahresdauer gibt. Dasselbe gilt für auch für Maschinen, die nur für eine kurze Zeit benötigt werden, also nur für wenige Wochen oder Monate. Das trifft beispielsweise bei Migrationsmaschinen zu. Diese Möglichkeiten sollten ebenfalls mit dem Preiskalkulator für Ihr Szenario geprüft werden und 1-Jahres oder 3-Jahres Verpflichtungen erst ab der Rentabilitätsschwelle angelegt werden.

Auch bei Betriebssystemlizenzen wie SLES for SAP gibt es die Möglichkeit, Rabatte für zugesicherte Nutzung zu kaufen und dadurch eine Kostenersparnis von bis zu 57 % zu erreichen. Das Speichern und Kaufen der Rabatte erfolgt wie bei den Maschinentypen im Menü für Compute Engine, jedoch wie der Abbildung zu entnehmen ist über „Neue Lizenz mit Rabatt für zugesicherte Nutzung“.

Eine weitere Empfehlung ist, dass der Kauf der Rabatte im selben Projekt und in derselben Region konfiguriert wird, in welchen diese genutzt werden sollen. Somit können doppelte Kosten vermieden werden. Rabatte können zwar zwischen Projekten geteilt werden, nicht jedoch zwischen den Regionen. Außerdem sollten Sie bedenken,

The screenshot shows the Google Cloud Platform interface with the following details:

- Header:** Google Cloud Platform, search bar (Produkte und Ressourcen suchen), navigation icons.
- Breadcrumbs:** Rabatt für zugesicherte Nutzung kaufen.
- Section:** Wählen Sie eine der Optionen, um einen Rabatt für zugesicherte Nutzung zu erwerben.
- Option 1: Neuer Rabatt für zugesicherte Nutzung für Hardware**
  - Rabatt für zugesicherte Nutzung für Hardware erwerben
- Option 2: Neue Lizenz mit Rabatt für zugesicherte Nutzung**
  - Eine Lizenz mit Rabatt für zugesicherte Nutzung erwerben
- Offer Summary:**

Name: sap-commitment	Geschätzter monatlicher Gesamtbetrag: 842,22 \$ ~730 Stunden pro Monat	Stundensatz: 1.154 \$
----------------------	---	-----------------------
- Offer Details:**
  - Region:** europe-west1
  - Zusicherungsart:** Allgemeiner Zweck N2
  - Dauer:** 1 Jahr (selected)
  - WCPUs:** 48
  - Speicher:** 192 GB
  - Lokale SSDs:** + Lokales SSD hinzufügen
- Note:** Rabatte für zugesicherte Nutzung gelten nicht für VM-Instanzen auf Abort, f1-micro- und g1-smal-Maschinentypen mit gemeinsam genutztem Kern oder erweiterten Speicher.
- Text:** Sie erklären sich damit einverstanden, die anfallenden monatlichen Gebühren für die oben ausgewählte Laufzeit zu zahlen. Die Laufzeit wird nicht automatisch verlängert. Diese monatlichen Gebühren können nicht geändert werden und fallen unabhängig von der tatsächlichen Nutzung an. Mehr erfahren
- Text:** Lesen Sie sich die dienstspezifischen Nutzungsbedingungen der Google Cloud Platform durch, bevor Sie fortfahren.
- Buttons:** Kaufen, Abbrechen.

**Abb. 9.12** Konsolenscreenshot: Rabatt für zugesicherte Nutzung kaufen

dass gekaufte Rabatte nicht geändert oder gelöscht werden können. Es können lediglich zusätzliche, das bedeutet inkrementelle Rabatte gekauft werden.

Rabatte für die zugesicherte Nutzung sind sehr flexibel in der jeweiligen Maschinentyp-Gruppe und diese Rabatte sind nicht vergleichbar mit dem Konzept der Reservierungen. Rabatte für die zugesicherte Nutzung werden auf einer totalen Anzahl von vCPU und RAM pro Maschinentyp gekauft und können dann nach Belieben in die möglichen VM-Größen aufgeteilt werden. Es findet somit keine Zusicherung auf bestimmte Maschinengrößen-Kombinationen statt.

Um immer einen Überblick über die Kosten zu behalten, werden im Cloud Billing viele vorgefertigte Berichte zu Kostenverlauf, aktuellen Kostentrends und prognostizierte Kosten von Google Cloud automatisch bereitgestellt [24].

## 9.5.2 Reservierungen in Google Cloud

Reservierungen [25] in Google Cloud bieten die Möglichkeit, Ressourcen zu reservieren, sodass Sie für Ihre Google Cloud Projekte in der Zukunft bereitstehen und reserviert sind. Dies wird empfohlen, wenn Sie zukünftigen Wachstum, Migrationen oder (un) geplante Spitzen, Disaster Recovery, Wachstum oder anderes planen. Reservierungen in der Google Cloud sind somit nicht vergleichbar mit Reservierungen auf anderen Hyperscalern. Sobald die Reservierung erstellt ist, wird diese reservierte Instanz in Rechnung gestellt. Die Instanz ist so lange verfügbar und für Sie reserviert, bis Sie die Reservierung löschen. Das bedeutet, Reservierungen können zu jeder Zeit gelöscht werden, ungleich der Commitments (zugesicherte Nutzung) aus Abschn. 9.5.1, welche nicht vor Ende der Commitment-Zeit gelöscht werden können. Reservierungen werden entweder über die Preise auf Abruf (on-demand), oder, wenn Rabatte für zugesicherte Nutzung erstellt wurden, über diese Rabatte abgerechnet.

### Vorteile von Reservierungen

- Eine zugesicherte Verfügbarkeit von Ressourcen, wenn diese benötigt werden.
- Keine Zeitbeschränkung, das bedeutet Reservierungen können jederzeit angelegt und gelöscht werden.
- Reservierungen nutzen alle bestehenden Rabatte (zugesicherte Nutzung usw.).
- Reservierungen können zwischen verschiedenen Projekten geteilt werden, um Flexibilität in der Landschaftsplanung und -bereitstellung zu ermöglichen.

### Limitierungen von Reservierungen

- Reservierungen sind nicht auf alle Google Cloud Dienste anwendbar, sondern nur für Compute Engine, Google Kubernetes Engine und Dataproc.
- Nur maximal 1000 VM-Instanzen können pro Reservierung gewählt werden.

- Die Kontingente müssen zu Genüge vorhanden sein.
- Wenn Sie Instanzen aus Ihren Reservierungen verwenden, können Sie die Reihenfolge, in der die Reservierungen genutzt werden, nur eingeschränkt bestimmen.

### 9.5.3 Google Cloud Kontingente und Budgets

**Kontingente** (Quotas) und Budgets ermöglichen Kostenkontrolle auf der Google Cloud. Kontingente limitieren wie viel von einer bestimmten Ressource genutzt werden kann. Ein Kontingent ist eine genau quantifizierbare (zählbare) Ressource, beispielsweise die maximale Anzahl von vCPUs oder Anzahl der Load-Balancer, die im Google Cloud Projekt genutzt werden. Es gibt zwei verschiedene Typen von Kontingenten: **Ratenkontingente** (rate quota), um eine Anzahl von Anfragen zu limitieren, z. B. Anzahl der API-Anfragen pro Tag, und **Zuteilungskontingente** (allocation quota), die Nutzung von Ressourcen limitieren, z. B. Anzahl von VMs zur selben Zeit im Projekt. Manche Kontingente sind global und manche regional oder zonal.

Google Cloud setzt Kontingente als Regel durch, um die gesamte Nutzergemeinschaft von Google Cloud zu schützen und unvorhersehbare Lastspitzen zu verhindern. Kontingente helfen zudem auch beim Management der Ressourcen, indem unerwartete Kosten und Abrechnungen für Nutzer vermieden werden.

Kontingente müssen verwaltet und modifiziert werden, je nachdem, wie viele Ressourcen ein Nutzer benötigt. Ansonsten wird das Kontingent aufgebraucht und der Nutzer erhält eine Fehlermeldung während der Durchführung einer Aufgabe wie beispielsweise eine neue VM oder Projekt anlegen oder eine API aufrufen. Normalerweise werden Kontingenterhöhungen automatisch evaluiert und genehmigt, jedoch kann es Ausnahmen bei ungewöhnlichen Anfragen geben. Diese können bis zu 2–3 Werktagen in Anspruch nehmen. Daher empfiehlt es sich die Kontingenterhöhungen so früh wie möglich planbar durchzuführen. Mehr Informationen zum Management, Monitoring und Erhöhung von Kontingenten auf der Website [26].

**Budgets** ermöglichen die Erfassung der aktuellen Google Cloud Ausgaben im Vergleich zu geplanten Ausgaben, um Überraschungen in der Höhe der Kosten zu vermeiden. Mit Budget-Warnungen können automatisch E-Mail-Benachrichtigungen und weitere Benachrichtigungen gesendet werden. Diese können dann automatisierten Kostenkontrollen und Aktionen dienen. Die Möglichkeiten werden in der Dokumentation erläutert [27].

### 9.5.4 Preiskalkulation für das SAP S/4HANA auf Google Cloud Architekturbeispiel

Um eine geplante Cloud Landschaft zu bepreisen, stellt Google Cloud einen Preiskalkulator auf der Website [28] bereit. Hier soll beispielhaft anhand der Architektur von

Kapitel 9.1, die in den vorherigen Kapiteln aufgebaut wurde, der Preiskalkulator erklärt und ein Preis der Landschaft ermittelt werden.

Die meisten Google Cloud Services können selbstständig im Preiskalkulator berechnet werden. Um die Komplexität im Rahmen zu halten, wird Im folgenden Beispiel nur auf die Kostenkalkulation der Compute Engine Instanzen für die SAP HANA und SAP Applikationsserver in beiden Regionen, sowie die dazugehörigen benötigten Festplatten und die SLES für SAP Betriebssystemlizenzen fokussiert.

In einer realistischen Preiskalkulation müssen ebenfalls noch Google Cloud Storage für die Sicherungen, der Cloud Load Balancer, IP-Adressennutzung, Netzwerkegress, Dedicated oder Partner Interconnect bzw. Cloud VPN, Cloud NAT, Cloud DNS, Cloud Operations, Support und je nach Nutzung weitere Services für NFS oder Sicherungstools von Google oder von Drittanbietern hinzugerechnet werden.

Die folgenden Systeme der Architektur wurden im Preiskalkulator hinzugefügt:

- 2x n1-standard-4 für SAP HANA Studio und JumpBox, mit SLES
- 3x n2-standard-4 mit je 54 GB HDD, mit SLES für SAP
- 3x n2-standard-16 mit je 54 GB HDD, mit SLES für SAP
- 3x n1-higmem-32 mit je 843 GB SSD und 446 GB HDD, mit SLES für SAP
- Alle mit 3-Jahres zugesicherte Nutzung (CUD)

Das Ergebnis ist ein Listenpreis (ohne Unternehmensrabatte oder ähnliches – dafür nehmen Sie über die Website Kontakt mit Ihrem Google Cloud Vertriebsverantwortlichen auf). Für die gewählte Architektur betragen die Kosten:

**USD 5344,11 [29] pro Monat** (Stand 22-September-2021)

Diese Kalkulation kann entweder als URL-Link gespeichert oder per E-Mail versendet werden und zu zukünftigen Zeitpunkten weiter angepasst und abgeändert werden [30]. Die folgenden Abbildungen zeigen die vollständige Preiskalkulation im Kalkulator mit den benötigten VM-Instanzen in den Regionen, den Betriebssystemlizenzen und den Festplatten pro Speicherklasse und Region an (Abb. 9.13, 9.14 und 9.15).

---

## 9.6 Sicherung & Wiederherstellung auf Google Cloud

Es gibt verschiedene Möglichkeiten, Sicherungen und Wiederherstellungen auf Google Cloud zu konfigurieren. Diese wurden in Abschn. 8.5.5 erläutert. In diesem Kapitel wird die Konfiguration der empfohlenen, nativen Lösungen betrachtet.

**Abb. 9.13** Google Cloud Preiskalkulator Beispielberechnung – Teil 1

Estimate		
Compute Engine		
2 x SAP NetWeaver App Servers		
Region: Belgium		
1,460 total hours per month		
Commitment term: 3 Years		
VM class: regular		
Instance type: n2-standard-16 Committed Use Discount applied	USD 561.48	
Operating System / Software: Paid Committed Use Discount applied	USD 215.50	
<b>Estimated Component Cost: USD 776.98 per 1 month</b>		
2 x SAP ASCS and ERS		
Region: Belgium		
1,460 total hours per month		
Commitment term: 3 Years		
VM class: regular		
Instance type: n2-standard-4 Committed Use Discount applied	USD 140.37	
Operating System / Software: Paid Committed Use Discount applied	USD 178.85	
<b>Estimated Component Cost: USD 319.22 per 1 month</b>		
2 x SAP HANA		
Region: Belgium		
1,460 total hours per month		
Commitment term: 3 Years		
VM class: regular		
Instance type: n1-highmem-32 Committed Use Discount applied	USD 1,367.89	
Operating System / Software: Paid Committed Use Discount applied	USD 215.50	
<b>Estimated Component Cost: USD 1,583.39 per 1 month</b>		

**Abb. 9.14** Google Cloud Preiskalkulator Beispielberechnung – Teil 2

1 x SAP NetWeaver App Server DR			
Region: Netherlands			
730 total hours per month			
Commitment term: 3 Years			
VM class: regular			
Instance type: n2-standard-16 Committed Use Discount applied	USD 280.99		
Operating System / Software: Paid Committed Use Discount applied	USD 107.75		
<b>Estimated Component Cost: USD 388.73 per 1 month</b>			
1 x SAP ASCS and ERS			
Region: Netherlands			
730 total hours per month			
Commitment term: 3 Years			
VM class: regular			
Instance type: n2-standard-4 Committed Use Discount applied	USD 70.25		
Operating System / Software: Paid Committed Use Discount applied	USD 89.43		
<b>Estimated Component Cost: USD 159.67 per 1 month</b>			
1 x SAP HANA DR			
Region: Netherlands			
730 total hours per month			
Commitment term: 3 Years			
VM class: regular			
Instance type: n1-highmem-32 Committed Use Discount applied	USD 684.55		
Operating System / Software: Paid Committed Use Discount applied	USD 107.75		
<b>Estimated Component Cost: USD 792.30 per 1 month</b>			
2 x 1) Jumpbox + SAP Router and 2) SAP HANA Studio etc			
Region: Belgium			
1,460 total hours per month			
Commitment term: 3 Years			
VM class: regular			

2 x 1) Jumpbox + SAP Router and  
2) SAP HANA Studio etc

Region: Belgium

1,460 total hours per month

Commitment term: 3 Years

VM class: regular

Instance type: n1-standard-4	USD 137.31
Committed Use Discount applied	
Operating System / Software: Paid	USD 160.60

**Estimated Component Cost: USD 297.91 per 1 month**

Persistent Disk

Belgium

Regional Standard PD: 1,108 GiB	USD 88.64
Regional SSD PD: 1,686 GiB	USD 573.24

**USD 661.88**

Netherlands

Regional Standard PD: 554 GiB	USD 48.75
Regional SSD PD: 843 GiB	USD 315.28

**USD 364.03**

**Total Estimated Cost: USD 5,344.11 per 1 month**

Estimate Currency

USD - US Dollar ▾

**EMAIL ESTIMATE**

**SAVE ESTIMATE**

**Abb. 9.15** Google Cloud Preiskalkulator Beispielberechnung – Teil 3

### 9.6.1 SAP NetWeaver Applikationsserver

Für SAP NetWeaver Applikationsserver sind Snapshots die Standardoption um Sicherungen zu erstellen. Google Cloud Compute Engine bietet dazu native automatisierte Snapshot Funktionalitäten für die angehängten Festplatten. Empfohlen wird die SAP NetWeaver Dateisysteme wie/interfaces, /sapmnt und/usr/sap/trans und die Binaries wie/exe und/boot mit Snapshots zu sichern.

Die Bedienung [31] über die Google Cloud Console sieht wie folgt aus:

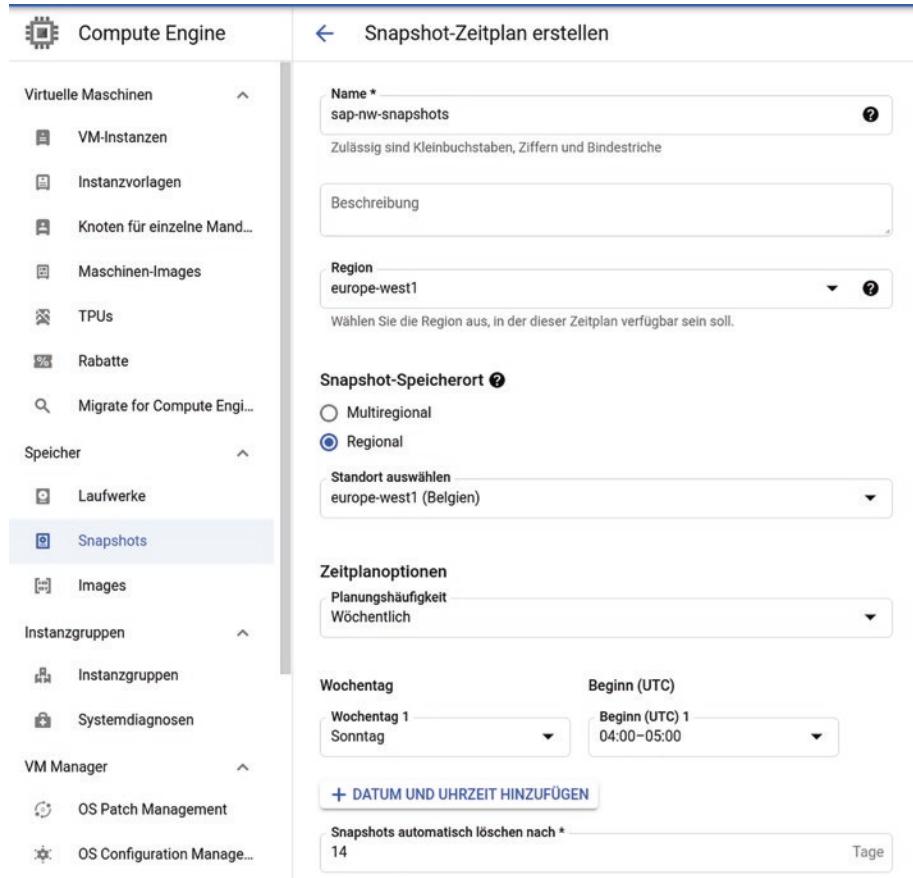
1. Öffnen Sie in der Google Cloud Console die Seite Compute Engine → Snapshots → Snapshot erstellen.
2. Geben Sie einen Namen für den Snapshot ein und optional eine Beschreibung des Snapshots.
3. Wählen Sie im Drop-down-Menü das Quelllaufwerk aus.
4. Legen Sie den Google Cloud Storage Speicherort für den Snapshot fest. Sie können den Standardspeicherort oder einen benutzerdefinierten Speicherort verwenden.
  - Wählen Sie unter Speicherort aus, ob Sie Ihren Snapshot an einem multiregionalen oder einem regionalen Speicherort speichern möchten.
  - Wählen Sie die gewünschte Region oder multiregionalen Speicherort aus. Wenn Sie die Region bzw. den multiregionalen Speicherort verwenden möchten, der dem Quelllaufwerk am nächsten liegt, wählen Sie je nach Standort des Laufwerks (Standard) aus.
5. Klicken Sie auf Erstellen, um den Snapshot zu erstellen.

Für Snapshots kann ein Zeitplan erstellt werden [32], sodass diese automatisch stündlich, täglich oder wöchentlich usw. angelegt werden. Dazu wird ebenfalls der Speicherort in Cloud Storage konfiguriert, und Löschregeln sowie die Anwendungskonsistenz können definiert werden. Die Konfiguration ist im folgenden Screenshot ersichtlich (Abb. 9.16).

Nach der Erstellung des Snapshot-Zeitplans können diese den einzelnen Laufwerken jeweils zugeordnet werden. Snapshot-Zeitpläne können sogar auch direkt während einer Laufwerkerstellung ausgewählt werden. Snapshots, die mit Laufwerken verknüpft sind, erstellen zudem kontinuierlich Systemereignisse, welche über Cloud Monitoring und Audit-Logs analysiert werden können. Generell können Snapshots außerdem projektübergreifend freigegeben werden.

Zur Wiederherstellung dienen Snapshots nach ihrer Erstellung entweder indem bei der Laufwerkerstellung als Quelle der Snapshot ausgewählt wird und daraufhin an VM-Instanzen gehängt wird. Oder aber, sie dienen direkt der Instanzneuerstellung, bei welcher die Instanz anhand des Snapshots neu erstellt wird.

Die “Best Practices” zur Erstellung von anwendungskonsistenten Snapshots können der Dokumentation entnommen werden [33]. Snapshots sollten wann immer möglich zu Zeiten erstellt werden, wenn das System nicht unter Spitzenlast steht.



**Abb. 9.16** Screenshot: Snapshot-Zeitplan erstellen

## 9.6.2 SAP HANA Datenbank

Wie in Abschn. 8.5.5 aufgezeigt, sind Snapshots für die SAP HANA Datenbank auf Google Cloud nur eine Möglichkeit zur Sicherung, die jedoch nicht zusätzlich von SAP zertifiziert ist. Der SAP HANA Backint Agent für Google Cloud Storage als weitere Möglichkeit für Sicherungen ist SAP-zertifiziert und wird im zweiten Teil dieses Unterkapitels beschrieben.

### 9.6.2.1 Setup von Snapshot-basierten Sicherungen für SAP HANA

Im SAP HANA Umfeld sind Snapshots von Festplatten zwar ausfallbeständig (crash-consistent), sie geben jedoch keine Konsistenz auf Applikationsebene oder Dateiebene und sind deshalb nicht im Produktivumfeld für die SAP HANA Daten in/hana/data empfohlen. Stattdessen wird empfohlen die SAP HANA Backup Volumes zu sichern.

Bei der Installation der SAP HANA Datenbanken wie in Abschn. 9.4.1 aufgezeigt, wird im Deployment Manager Template der nichtflüchtige Speicher als Standardsicherungsverzeichnis mit/hanabackup/data/SID konfiguriert. Eine erste Sicherung der SAP HANA Datenbank wird mit den folgenden Befehlen [34] durchgeführt:

1. Wechseln Sie auf den primären Host und melden sich via SSH an

```
sudo -i -u sid adm
```

2. Erstellen Sie eine Datenbanksicherung für das SAP HANA System

```
hdbsql -t -u system -p system-password -i inst-num \  
"backup data using file ('full')"
```

Konfigurieren Sie die Snapshots für jedes Verzeichnis, das gesichert werden soll. Die Konfiguration für Snapshot Erstellungen mit Zeitplänen wird, wie in Abschn. 9.6.1 beschrieben, durchgeführt. Damit Snapshots konsistent sind, sollten Schreibvorgänge in das Dateisystem angehalten werden und Laufwerkzwischenspeicher geleert werden, damit der Snapshot konsistent ist [35].

### **9.6.2.2 Setup von SAP HANA Backint mit dem Cloud Storage Backint Agent**

Wenn der SAP HANA Backint Agent für Google Cloud Storage verwendet wird benötigt die SAP HANA Datenbank keine Festplatte für das/hanabackup, sondern die Sicherungen werden direkt in Google Cloud Storage gespeichert.

Vor der Installation des SAP HANA Backint Agent und der Konfiguration von Sicherungen sollten Sie folgendes beachten:

- Nutzen Sie die SAP HANA Datenbank Administrationstools, um Sicherungen zu konfigurieren und einzuplanen.
- Nutzen Sie native Betriebssystemkompression, da Sicherungen nicht unbedingt effizient komprimiert werden, sondern auf Kosten der CPU-Zyklen und Durchsatz.
- Eine Deduplizierung ist nicht möglich, obwohl es ein nativer Agent ist, der in das SAP HANA Datenbanktool integriert ist. Die Applikationskonsistenz ist jedoch trotzdem besser als bei Volume-Snapshots.
- Es müssen zusätzliche manuelle Schritte durchgeführt werden, um Systemkopien auf Basis von Sicherungen zu erstellen.

Die folgenden Schritte müssen bei einer Installation und Konfiguration von Sicherungen mit SAP HANA Backint durchgeführt werden:

1. Die SAP HANA Datenbank, die gesichert werden soll, muss installiert sein.
2. Erstellen Sie einen Google Cloud Storage Bucket für Ihre Sicherungen.
  - a. Wählen Sie die Bucket Lokation und die Speicherklasse nach der Beschreibung in Abschn. 8.3.2
3. Installieren Sie den Backint Agent auf dem SAP HANA Host (im Falle eines Hochverfügbarkeitsclusters oder eines Systems mit horizontaler Skalierung muss der Agent auf jedem SAP HANA System installiert werden).
4. Konfigurieren Sie den Backint Agenten und SAP HANA in der parameters.txt Datei (die Konfigurationsoptionen können auf der Website eingesehen werden [36]) und geben Sie die folgenden Parameter ein:
  - a. Cloud Storage Bucket Namen, Service Account und weitere Parameter wie parallel\_factor, rate\_limit, threads und mehr.
5. Testen Sie die Konfiguration und Funktionsweise der Sicherungen über SAP HANA Backint.
6. Für Fehlerbehebung und Supportunterstützung gibt es Empfehlungen in der offiziellen Dokumentation [37].

---

## 9.7 Scripting und Automatisierung auf Google Cloud

Das folgende Kapitel gibt einen Überblick auf die Skript- und Automatisierungstools, welche im Bereich von SAP-Systemen auf Google Cloud genutzt werden können. Außerdem werden drei verschiedene Anwendungsfälle für Automatisierungsskripte im Detail aufgezeigt: das automatisierte Starten und Stoppen von Instanzen und die Auto-skalierung von SAP Applikationsservern. Zuletzt wird der SAP Landscape Manager Connector von Google Cloud beschrieben.

### 9.7.1 Tools für Skripterstellung und Automatisierung auf Google Cloud

Es gibt unterschiedliche Möglichkeiten und Tools für Automatisierung in Cloud Umgebungen. Häufig genutzte Möglichkeiten sind entweder mit Google Cloud Deployment Manager oder mit anderen Tools und Open Source Projekten wie beispielsweise Terraform, Ansible, Python und mehr.

Die Vorteile von Infrastruktur-als-Code (IaC) bzw. Skripterstellung ist die automatisierte Erstellung von Google Cloud Ressourcen wie VM-Instanzen, Festplatten und so weiter. Diese werden einmal geskriptet und dann für dieselbe oder ähnliche Landschaften wiederverwendet. Dieser Bereitstellungsweg bringt eine große Zeiter sparnis im Betrieb mit sich und ist zudem weniger fehleranfällig und risikobehaftet.

Der Google Cloud Deployment Manager Service dient zur Bereitstellung von Infrastruktur. Über diesen werden das Erstellen und Verwalten von Google Cloud-Ressourcen

automatisiert. Flexible Vorlagen und Konfigurationsdateien können erstellt werden, in denen eine Reihe von Google Cloud-Diensten zum Einsatz kommen, z. B. Cloud Storage, Compute Engine und Cloud SQL.

Diese IaC-Skripte werden in Google Cloud Source Repositories verwaltet, welches vollständig verwaltet und skalierbar ist und ein privates Git-Repository darstellt. Quellcode kann entwickelt und getestet, sowie über Versionsverwaltung verwaltet und synchronisiert werden. Es gibt die Möglichkeit mit weiteren Google Cloud-Tools wie Cloud Build, App Engine, Pub/Sub und Produkten für das IT Operations Management wie Cloud Monitoring und Cloud Logging zu integrieren.

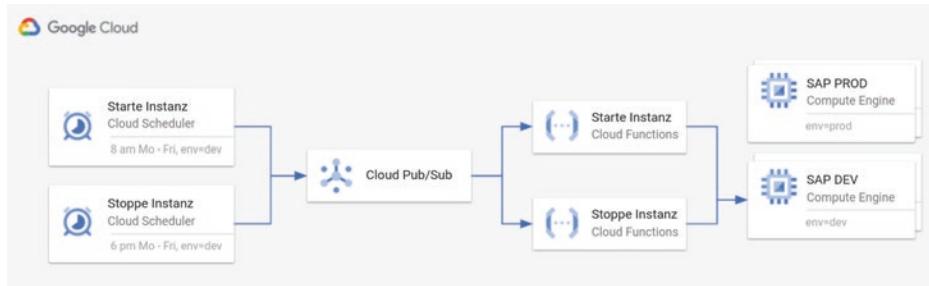
Skriptbeispiele und -vorlagen sowie Empfehlungen und Vorgehen zur Fehlerbehebung für den Deployment Manager werden auf der Google Website [38] und für Terraform auf GitHub [39] zur Verfügung gestellt.

### **9.7.2 Anwendungsfall: Automatisierter Start und Stop von Instanzen**

Mit Google Cloud Scheduler können viele Aktivitäten für die Automatisierung terminiert werden, beispielsweise das Starten und Stoppen von Systemen. Cloud Scheduler ist ein vollständig verwalteter Cron-Job-Scheduler, mit dem praktisch jeder Job geplant werden kann, einschließlich Batch-Jobs und Cloud-Infrastrukturvorgänge. Sie können alles automatisieren, auch Wiederholungsversuche während eines Fehlers, um manuelle Arbeit und Eingriffe zu reduzieren. Cloud Scheduler dient als einzige Konsolenoberfläche, mit der alle Automatisierungsaufgaben in einer Oberfläche verwaltet werden können. Um Sicherungsaufgaben zu planen, kann ebenfalls Cloud Scheduler verwendet werden.

In SAP Landschaften kann dies für Sandbox Systeme, oder sogar für Entwicklungs- und Testsysteme sinnvoll sein, wenn diese während der Nacht und während Wochenenden ausgeschaltet werden können, um Kosten zu sparen.

Für solch eine Konfiguration [40] werden die Dienste Google Cloud Scheduler, Cloud Pub/Sub und Cloud Functions benötigt. Der Cloud Scheduler definiert den genauen Zeitplan mit der Zeitangabe und dem Tag- oder Wiederholungsmodus (bspw. Montag-Freitag). Diese von Cloud Scheduler ausgelösten Events werden von Cloud Pub/Sub empfangen, Google's Messaging-orientierter Middleware Lösung. Cloud Pub/Sub löst daraufhin Cloud Functions aus, welche Code-Ausschnitte ausführt, zum Beispiel das Starten von Instanzen und deren angehängten Startup-Skripten. Die Abfolge wird in der folgenden Abbildung veranschaulicht (Abb. 9.17).



**Abb. 9.17** Automatisierung von Start und Stop von Instanzen

### 9.7.3 Anwendungsfall: Autoskalierung für SAP NetWeaver Applikationsserver

In den meisten SAP-Landschaften gibt es eine variierende aber vorhersehbare Nutzung von Applikationsservern. Die Vorhersehbarkeit macht diese Systeme zu guten Kandidaten die Elastizität der Cloud zu nutzen, beispielsweise mit automatischer Skalierung, also dem Verkleinern und Vergrößern bzw. Hinzufügen und Entfernen von VM-Instanzen. Die Autoskalierung für SAP-Applikationsserver kann auf Google Cloud einfach implementiert und automatisiert werden. Vorteile von automatisch skalierenden Applikationsservern sind vor allem Ressourcen- und Kostensparnisse. Dabei können zwei Varianten von Autoskalierung aufgesetzt werden:

#### 1. Nutzungsbasierte Autoskalierung

- Basierend auf Schwellwerten von CPU, RAM oder weiterer Ressourcennutzung
- Wird umgesetzt mit der Hilfe von Instanzvorlagen (oder Maschinen-Images), verwaltete Instanzgruppen (Managed Instance Groups, MIG) und Google Operations

#### 2. Zeitplanbasierte Autoskalierung

- Basierend auf einem spezifischen Nutzungszeitplan mit Tages- und Uhrzeiten
- Wird umgesetzt mit der Hilfe von Google Cloud Scheduler

Mehr Informationen, Automatisierungsbefehle und Code Beispiele können der Dokumentation entnommen werden [41].

### 9.7.4 Anwendungsfall: Vereinfachung des operativen Betriebs

Durch die Google Cloud und den cloud-nativen Möglichkeiten vereinfacht sich der operative Betrieb für SAP-Teams und Systemadministratoren. Anfragen von

verschiedenen Projekten und Teams für dringliche Sandbox- oder Projektssysteme können schnell, mit wenig Aufwand und kostengünstig erfüllt werden. Mit Hilfe von Instanzvorlagen (Maschinenimages) können Systeme schnell kopiert und in einer neuen VM-Instanz bereitgestellt werden. Die On-Demand-Abrechnung gibt hier ebenfalls weitere Flexibilität, denn die Systeme können nach wenigen Wochen oder Monaten wieder gelöscht werden und verursachen danach keine Kosten mehr. Das Klonen ist sowohl für die SAP HANA als auch SAP NetWeaver Ebene möglich [42].

### 9.7.5 SAP Landscape Manager und Google Cloud

Eine weitere Möglichkeit für die Automatisierung und Verwaltung der SAP Landschaft ist der SAP Landscape Manager (SAP LaMa) mit dem Google Cloud Connector. Google stellt diesen Connector von der Google Cloud Seite aus kostenlos zur Verfügung, die SAP LaMa Lizenz muss bei SAP erworben werden. Die folgenden Optionen für die Verwaltung der SAP Landschaft über SAP LaMa werden angeboten:

- SAP-System-/Instanzerstellung
- SAP-System-/Instanzverschiebung
- SAP-System Start und Stop, auch inklusive virtuellen Hosts
- Speicher-basierte Kopierprozeduren, welche auf der Snapshot-Technologie von Google Cloud Festplatten basieren
  - SAP-Systemklone und Systemkopien
  - SAP-Systemaktualisierung und Datenbankaktualisierung
  - Erstellung von SAP HANA Systemreplikationsstufen

Die Systemanforderungen, Anforderungen an Authentifizierung und Berechtigungen, die möglichen Szenarien [43] und die Installationsanleitung [44] können der Dokumentation entnommen werden.

---

## 9.8 Desaster Recovery mit Google Cloud

Um eine Desaster Recovery Systemlandschaft bereitzustellen muss die gewünschte Desaster Recovery Option anhand der Anforderungen an RPO und RTO gewählt werden. Diese wurden im Abschn. 8.5.4 im Detail erläutert. In diesem Beispieldaten umgebung der Architektur aus Abschn. 9.1 wurde die volle, also heiße (hot) DR Option gewählt. Dies bedeutet, dass auf allen Ebenen des Systems die Instanzen asynchron repliziert werden.

### 9.8.1 SAP HANA Desaster Recovery

Auf der SAP HANA Datenbankebene wird das Desaster Recovery ebenfalls wie in Abschn. 9.4.1 mit der SAP HANA Systemreplikation (HSR) implementiert. Dazu muss eine dritte SAP HANA Instanz bereitgestellt werden, wozu ebenfalls entweder eine manuelle oder eine automatisierte Bereitstellung gewählt werden kann. Diese wird mit einer asynchronen Systemreplikation ausgehend von der sekundären SAP HANA Instanz der primären Region konfiguriert. Es wird empfohlen in einem Desaster Fall das Takeover-Verfahren der SAP HANA Instanzen manuell nach der Dokumentation auszulösen [45].

### 9.8.2 SAP NetWeaver Applikationsserver Desaster Recovery

Auf der SAP Applikationsserverebene wird das Desaster Recovery ebenfalls wie in Abschn. 9.4.2 mit zwei VM-Instanzen bereitgestellt: eine Instanz für ASCS/ERS und eine Instanz für den SAP NetWeaver Applikationsserver. Auch im Bereich der Applikationsserver wird empfohlen in einem Desaster Fall das Takeover-Verfahren manuell auszulösen.

### 9.8.3 Empfehlungen für die Desaster Recovery Planung

Je nach Wahl der Desaster Recovery Methode anhand der gewünschten RPO und RTO sollten die folgenden Punkte für alle Optionen beachtet werden:

1. Kapazitätsplanung
  - Falls keine volle Desaster Recovery Option gewählt wird, sondern eine kostenoptimierte Variante, so sollten die Kapazitäten mit einer Kapazitätsplanung für den Katastrophenfall gesichert werden.
2. Automatisierung
  - Normalerweise wird ein Desaster Recovery Konzept manuell initiiert. Dieses sollte in der Wiederherstellung aus den Sicherungen und dem Start der Instanzen soweit vollständig wie möglich automatisiert werden, um eine schnelle (geringstmögliche RTO) und fehlerfreie Wiederherstellung zu gewährleisten.
3. Dokumentation und Test
  - Dokumentieren Sie Ihren Desaster Recovery Plan vollständig und stellen Sie sicher, dass die Dokumentation im Katastrophenfall zur Verfügung steht. Testen Sie Ihren Plan regelmäßig und stellen Sie sicher, dass Systemadministratoren und das Unternehmen diesen Plan kennen und in diesem Verfahren geschult sind. Testen Sie diesen Plan regelmäßig mit der Hilfe von Replika Landschaften, die schnell kopiert und wieder runtergefahren werden können.

## 9.9 Zusammenfassung

Dieses Kapitel gibt nach der Einführung einer vollständigen Beispielarchitektur für SAP S/4HANA auf Google Cloud eine Übersicht auf die generellen Planungs- und Bereitstellungslisten, die für jegliche SAP Bereitstellungen verwendet werden können. Insgesamt werden 25 Schritte gelistet, die aufgeteilt werden in generelle Google Cloud Konfigurationen für Unternehmen und SAP-spezifische Schritte auf der Google Cloud. Diese Schritte werden dann in den Folgekapiteln anhand der eingeführten Architektur beispielhaft durchgeführt.

Nach dem Google Identity Account, Security und Netzwerksetup folgt die Bereitstellung der Compute Engine VM-Instanzen für die gesamte SAP S/4HANA Landschaft, also die SAP HANA Instanzen und die SAP NetWeaver Applikationsserver Instanzen im Hochverfügbarkeitsmodus. Hier wird jeweils auf die manuelle und die automatische Bereitstellung eingegangen und beide Wege aufgezeigt. Empfohlen wird eine Bereitstellung basierend auf Automatisierung mit Skripten.

Daraufhin folgen Kapitel zum Preis- und Abrechnungskonzept der Google Cloud. Der Unterschied zwischen Abrechnungen auf Abruf (On-Demand) und Rabatte für die zugesicherte Nutzung sowie Reservierungen, Kontingente und Budgets werden aufgezeigt. Das Kapitel schließt mit einer exemplarischen Berechnung anhand der SAP S/4HANA auf Google Cloud Beispielarchitektur, die zu Beginn in Kapitel 9.1 eingeführt wurde. Diese Kalkulation basiert auf dem öffentlichen Preiskalkulator der Google Cloud.

Dann folgen Kapitel zu den Konfigurationen für Sicherungen und Wiederherstellung mit Snapshots und SAP HANA Backint für Google Cloud Storage.

Das vorletzte Kapitel gibt einen Überblick zu Scripting und Automatisierung und die möglichen Tools sowie Anwendungsfälle im Bereich von automatischem Start und Stop von Instanzen, Autoskalierung für die Applikationsserver und die Vereinfachung des operativen Betriebs.

Zuletzt wird das Setup für Desaster Recovery für diese Beispielarchitektur erläutert und ein kurzer Überblick der Best-Practices Empfehlungen gegeben.

---

## Literatur

1. <https://cloud.google.com/docs/enterprise/setup-checklist> (Zugriff am 20.12.2021)
2. <https://cloud.google.com/solutions/sap/docs/checklist-sap-hana> (Zugriff am 20.12.2021)
3. <https://cloud.google.com/solutions/sap/docs/checklist-sap-netweaver> (Zugriff am 20.12.2021)
4. <https://cloud.google.com/solutions/sap/docs/netweaver-deployment-linux-dm#creating-a-network> (Zugriff am 20.12.2021)
5. [https://cloud.google.com/solutions/sap/docs/netweaver-deployment-linux-dm#adding\\_firewall\\_rules](https://cloud.google.com/solutions/sap/docs/netweaver-deployment-linux-dm#adding_firewall_rules) (Zugriff am 20.12.2021)
6. <https://cloud.google.com/vpc/docs/add-remove-network-tags> (Zugriff am 20.12.2021)
7. <https://cloud.google.com/nat/docs/using-nat> (Zugriff am 20.12.2021)

8. [https://cloud.google.com/solutions/sap/docs/sap-hana-operations-guide#setting\\_up\\_your\\_sap\\_support\\_channel\\_with\\_saprouter](https://cloud.google.com/solutions/sap/docs/sap-hana-operations-guide#setting_up_your_sap_support_channel_with_saprouter) (Zugriff am 20.12.2021)
9. [https://cloud.google.com/solutions/sap/docs/sap-hana-ha-dm-deployment#creating\\_a\\_high-availability\\_linux\\_cluster\\_with\\_sap\\_hana\\_installed](https://cloud.google.com/solutions/sap/docs/sap-hana-ha-dm-deployment#creating_a_high-availability_linux_cluster_with_sap_hana_installed) (Zugriff am 20.12.2021)
10. <https://cloud.google.com/solutions/sap/docs/sap-hana-ha-config-sles> (Zugriff am 20.12.2021)
11. [https://cloud.google.com/solutions/sap/docs/sap-hana-ha-dm-deployment#checking\\_the\\_sap\\_hana\\_system\\_using\\_sap\\_hana\\_studio](https://cloud.google.com/solutions/sap/docs/sap-hana-ha-dm-deployment#checking_the_sap_hana_system_using_sap_hana_studio) (Zugriff am 20.12.2021)
12. [https://cloud.google.com/solutions/sap/docs/sap-hana-ha-dm-deployment#performing\\_post-deployment\\_tasks](https://cloud.google.com/solutions/sap/docs/sap-hana-ha-dm-deployment#performing_post-deployment_tasks) (Zugriff am 20.12.2021)
13. [https://cloud.google.com/solutions/sap/docs/sap-hana-operations-guide#setting\\_up\\_your\\_sap\\_support\\_channel\\_with\\_saprouter](https://cloud.google.com/solutions/sap/docs/sap-hana-operations-guide#setting_up_your_sap_support_channel_with_saprouter) (Zugriff am 20.12.2021)
14. <https://cloud.google.com/solutions/sap/docs/netweaver-deployment-guide-linux> (Zugriff am 20.12.2021)
15. <https://cloud.google.com/solutions/sap/docs/netweaver-deployment-guide-windows> (Zugriff am 20.12.2021)
16. <https://cloud.google.com/solutions/sap/docs/netweaver-ha-config-sles#nw-ha-example-config-file> (Zugriff am 20.12.2021)
17. [https://cloud.google.com/solutions/sap/docs/netweaver-deployment-linux-manual#formatting\\_and\\_mounting\\_disk\\_drives](https://cloud.google.com/solutions/sap/docs/netweaver-deployment-linux-manual#formatting_and_mounting_disk_drives) (Zugriff am 20.12.2021)
18. [https://cloud.google.com/solutions/sap/docs/netweaver-deployment-linux-manual#preparing\\_the\\_operating\\_system](https://cloud.google.com/solutions/sap/docs/netweaver-deployment-linux-manual#preparing_the_operating_system) (Zugriff am 20.12.2021)
19. [https://help.sap.com/viewer/p/SAP\\_NETWEAVER\\_750](https://help.sap.com/viewer/p/SAP_NETWEAVER_750) (Zugriff am 20.12.2021)
20. [https://help.sap.com/doc/18cb1a50b9924bc3b94c2988cc8c51d9/7.5/en-US/mg\\_nw\\_75.pdf](https://help.sap.com/doc/18cb1a50b9924bc3b94c2988cc8c51d9/7.5/en-US/mg_nw_75.pdf) (Zugriff am 20.12.2021)
21. <https://cloud.google.com/solutions/sap/docs/netweaver-ha-config-sles> (Zugriff am 20.12.2021)
22. <https://cloud.google.com/solutions/sap/docs/netweaver-ha-config-sles#enable-back-end-comms> (Zugriff am 20.12.2021)
23. <https://cloud.google.com/compute/docs/instances/signing-up-committed-use-discounts> (Zugriff am 20.12.2021)
24. <https://cloud.google.com/billing/docs/reports> (Zugriff am 20.12.2021)
25. <https://cloud.google.com/compute/docs/instances/reservations-overview> (Zugriff am 20.12.2021)
26. <https://cloud.google.com/docs/quota> (Zugriff am 20.12.2021)
27. <https://cloud.google.com/billing/docs/how-to/budgets> (Zugriff am 20.12.2021)
28. <https://cloud.google.com/products/calculator> (Zugriff am 20.12.2021)
29. *The estimated fees provided by Google Cloud Pricing Calculator are for discussion purposes only and are not binding on either you or Google. Your actual fees may be higher or lower than the estimate. A more detailed and specific list of fees will be provided at time of sign up. To sign up for Google Cloud and purchase services, please click on one of the product links above.*
30. <https://cloud.google.com/products/calculator/#id=44aa2107-51c8-46ee-95a7-07c9bbeadc6d> (Zugriff am 20.12.2021)
31. [https://cloud.google.com/compute/docs/disks/create-snapshots#create\\_zonal\\_snapshot](https://cloud.google.com/compute/docs/disks/create-snapshots#create_zonal_snapshot) (Zugriff am 20.12.2021)
32. <https://cloud.google.com/compute/docs/disks/scheduled-snapshots> (Zugriff am 20.12.2021)
33. <https://cloud.google.com/compute/docs/disks/snapshot-best-practices> (Zugriff am 20.12.2021)
34. [https://cloud.google.com/solutions/sap/docs/sap-hana-ha-config-sles#back\\_up\\_the\\_databases](https://cloud.google.com/solutions/sap/docs/sap-hana-ha-config-sles#back_up_the_databases) (Zugriff am 20.12.2021)

35. [https://cloud.google.com/solutions/sap/docs/sap-hana-operations-guide#creating\\_a\\_snapshot\\_of\\_sap\\_hana](https://cloud.google.com/solutions/sap/docs/sap-hana-operations-guide#creating_a_snapshot_of_sap_hana) (Zugriff am 20.12.2021)
36. [https://cloud.google.com/solutions/sap/docs/sap-hana-backint-guide#configuring\\_the\\_backint\\_agent\\_and\\_hana](https://cloud.google.com/solutions/sap/docs/sap-hana-backint-guide#configuring_the_backint_agent_and_hana) (Zugriff am 20.12.2021)
37. <https://cloud.google.com/solutions/sap/docs/sap-hana-backint-guide> (Zugriff am 20.12.2021)
38. <https://cloud.google.com/solutions/sap/docs/dm-templates-for-sap> (Zugriff am 20.12.2021)
39. <https://github.com/terraform-google-modules/terraform-google-sap> (Zugriff am 20.12.2021)
40. <https://cloud.google.com/scheduler/docs/start-and-stop-compute-engine-instances-on-a-schedule> (Zugriff am 20.12.2021)
41. <https://cloud.google.com/blog/products/sap-google-cloud/best-practices-for-sap-app-server-autoscaling-on-google-cloud> (Zugriff am 20.12.2021)
42. [https://cloud.google.com/solutions/sap/docs/sap-hana-operations-guide#cloning\\_your\\_sap\\_hana\\_system](https://cloud.google.com/solutions/sap/docs/sap-hana-operations-guide#cloning_your_sap_hana_system) (Zugriff am 20.12.2021)
43. <https://cloud.google.com/solutions/sap/docs/sap-lama-connector-planning> (Zugriff am 20.12.2021)
44. <https://cloud.google.com/solutions/sap/docs/sap-lama-connector-deployment> (Zugriff am 20.12.2021)
45. [https://cloud.google.com/solutions/sap/docs/sap-hana-dr-planning-guide#triggering\\_a\\_take-over](https://cloud.google.com/solutions/sap/docs/sap-hana-dr-planning-guide#triggering_a_take-over) (Zugriff am 20.12.2021)



# Zusammenfassung und Ausblick

10

## Zusammenfassung

Dieses Kapitel fasst die wichtigsten Punkte aus allen vorherigen Kapiteln zusammen und zeigt einen Ausblick auf die zukünftige Entwicklung von SAP S/4HANA-Systemen in der Public Cloud.

### 10.1 Das Momentum von SAP S/4HANA

Das Ende der traditionellen SAP-Systeme wurde von der SAP auf 2025 festgelegt. Dann soll der Support für alle nicht-HANA-basierten SAP-Systeme enden. Kunden, welche bis dahin noch nicht auf HANA oder S/4HANA umgestellt haben, laufen Gefahr, die Geschäftsprozesse mit einem ERP-System abzuwickeln, welches keinen Support mehr hat. Dies kann sich kein Unternehmen leisten.

Aus dieser Frist von SAP entsteht nun das Momentum von SAP S/4HANA und der Transformation in die Cloud. Kunden haben wenig Möglichkeiten, die Frist von 2025 zu ignorieren. Es gibt sicherlich Unternehmen, welche aktuell mit dem verwendeten ERP-System mehr als zufrieden sind. Hier ist es wenig sinnvoll, allein aus dem Druck heraus auf S/4HANA umzusteigen.

So schauen sich die Unternehmen aktuell auch nach Alternativen zu dem Szenario der Transformation um. Es gibt Kunden der SAP, welche ganz bewusst, keine Transformation beginnen werden, sondern weiterhin auf das SAP ERP-System setzen werden. Die Entscheidung ist oftmals getrieben aus dem wenigen Mehrwert von SAP S/4HANA und HANA. Die Unternehmen sind mit ihren SAP-Systemen, z. B. ECC6.0 EhP 7 auf Oracle, sehr zufrieden und die Transformation zu S/4 kann nicht den zusätzlichen Nutzen generieren, wie SAP das propagiert.

Nach dem Jahre 2025 wird die SAP keine Updates und keine Patches mehr für die Systeme liefern. Der Support im Fehlerfall wird ebenfalls nicht mehr geleistet und Kunden, welche die Systeme nicht umgestellt haben, nehmen ein signifikantes Risiko in Kauf. Dennoch wäre es auch denkbar, diese Frist zu ignorieren.

Auch wenn es einige Kunden geben wird, welche nicht zu SAP S/4HANA transformieren werden, so ist davon auszugehen, dass der Großteil diesen Schritt bereits gemacht hat, derzeit macht oder in Zukunft machen wird. Bei der Vielzahl der Kunden geht mit jeder Transformation zu S/4HANA auch der Schritt in die Cloud einher. Dieses Momentum wird also noch länger als 2025 bestehen bleiben, da noch einige Nachzügler die Systeme nach 2025 umstellen werden.

---

## 10.2 Public Cloud als etablierter Trend

Die Public Cloud ist kein zeitlich begrenzter Trend, sondern die Public Cloud hat sich bereits so stark etabliert, dass neue Geschäftsmodelle nur auf Basis von Public Cloud aufgebaut werden. Seien es die neuen Internetbanken, welche ausschließlich auf Basis von Cloud Computing arbeiten oder aber Start-Ups, welche erst durch die Features der Public Clouds ein Geschäft generieren können. Allen gemein ist, dass Public Cloud das Rückgrat bildet.

Dennoch zeigen manche größeren Unternehmen immer noch Berührungsängste, wenn es um die Frage nach „SAP in der Cloud?“ geht. Hierbei verfolgen die Unternehmen die Strategie des geringsten Risikos, was für kritische SAP-Systeme absolut richtig ist. So zeigen sich Unternehmen sehr zögerlich und starten mit der Adaption der Public Cloud mit unkritischen Applikationen, welche jedoch auch keinen signifikanten Einfluss auf das Geschäft haben. Ein Ausfall von diesen unwichtigeren Applikationen bleibt meist ohne Folgen. Würde ein SAP-System wegen der Public Cloud komplett ausfallen, wäre diese tragisch und würde das Geschäft und alle Prozesse des Unternehmens signifikant beeinflussen.

In dieser Diskussion und bei dieser Überlegung sollten aber immer die Fakten herangezogen werden. Diese zeigen, dass die Public Clouds, genauso wie auch die Private Clouds, immer mal wieder einen Ausfall bei der Hardware haben oder aber Systeme davon beeinträchtigt sind. Wie auch bei den jüngsten Ausfällen von Facebook und WhatsApp sind nicht selten die Mitarbeiterinnen und Mitarbeiter verantwortlich, während die Hardware tadellos funktionierte.

Aktuell verfolgen alle größeren Unternehmen eine Strategie, die Public Cloud als Zielplattform zu nutzen. Manche Unternehmen haben bereits frühzeitig die Public Cloud genutzt, wie Carlsberg, welche strategisch frühzeitig alle Systeme in die Cloud migriert haben. Andere Unternehmen nutzen die Cloud als neue Platform für die Greenfield-Implementierung von SAP S/4HANA-Systemen. Für die Unternehmen, welche bislang noch keinen Fußabdruck in der Public Cloud besitzen, gibt es zwei Möglichkeiten:

1. Die Unternehmen nutzen bereits Public Cloud, jedoch ist dies nie wirklich publik gemacht worden und die zuständige IT-Abteilung hat keine Kenntnisse darüber. Dies geschieht und geschah bei sehr großen Konzernen immer wieder und muss durch fehlende Innovationskraft in der IT-Abteilung erklärt werden.
2. Die Unternehmen nutzen keinerlei Public Cloud aus Sicherheitsbedenken. Dieses Szenario ist nicht bei großen Konzernen, sondern eher im Mittelstand zu sehen. Hier existieren Unternehmen, welche hochsensible Daten verarbeiten, sodass eine Auslagerung der IT nicht infrage kommt.

Alle anderen Unternehmen und Konzerne befinden sich schon in der Public Cloud.

Die drei wichtigsten Anbieter von Public Clouds, Amazon Web Services, Microsoft Azure und Google Cloud, haben bereits eine lange Tradition und die Maturität der angebotenen Services ist sehr hoch. Dies ist auch der Grund, warum SAP ebenfalls das Cloud-Geschäft strategisch forciert. Neben der SAP HANA Enterprise Cloud erweitert die SAP AG stetig das Portfolio durch neue Cloud-Services und neue Cloud-Angebote. Hierzu gehört auch RISE, welches strategisch auf die Hyperscaler als Plattform für den Betrieb der SAP-Systeme setzt.

Alle diese Punkte zeigen, dass die Konzerne an der Public Cloud nicht vorbeikommen und diese alsbald adaptieren müssen. Die SAP-Systeme werden ebenfalls in die Richtung gehen, da die Vorteile die (wahr genommenen) Nachteile überwiegen.

---

## 10.3 Public Cloud als Innovationstreiber

Die Hyperscaler arbeiten mit Hochdruck daran, die Public Clouds und deren Ressourcen so einfach, wie möglich, nutzbar zu machen. Damit soll die Hürde zum Eintritt in die Welt der Public Clouds so gering, wie möglich, gehalten werden. Dies allein ist jedoch nicht der wichtigste Grund für die strategische Nutzung von Public Clouds. Vielmehr konzentrieren sich die Unternehmen und Kunden darauf, neue Features aus den Clouds einzusetzen.

In den vorherigen Kapiteln haben Sie gesehen, wie einfach es ist, neue SAP S/4HANA-Systeme in den Hyperscaler Clouds bereitzustellen. Hierbei wurden nicht nur einfache Umgebungen, sondern auch hochkomplexe Umgebungen gezeigt. Ein einfaches SAP S/4HANA-System mit einer Hochverfügbarkeit zu erstellen, benötigt viele Komponenten, wie den Cluster oder aber auch die Fileshares. Dies kann in den Hyperscaler Clouds sehr einfach bereitgestellt werden. Selbstverständlich sind dazu einige Konfigurationsschritte notwendig, aber Kunden können alle notwendigen Komponenten einfach aus der Cloud beziehen und müssen sich keine weiteren Hardware- oder Softwarekomponenten aneignen.

Insbesondere bei dem Thema des Katastrophenfalls (Disaster Recovery) zeigt sich der Vorteil der Hyperscaler Clouds. Alle Anbieter bieten den Kunden eine Vielzahl von Optionen für den Schutz vor einem Katastrophenfall. Dies kann durch ein-

fache Verfügbarkeitszonen innerhalb einer Region sein; dies kann aber auch durch die Absicherung vor dem Ausfall einer Region sein (Dublin – Frankfurt). Somit können alle Kunden die üblichen Ausfallszenarien auch in den Clouds nachbilden. Dies gilt ebenso für die Datensicherung und Datenwiederherstellung. Hierbei bieten die Hyperscaler zwar nur die Grundfunktionalitäten an, jedoch sind diese für einige kleinere Systemumgebung durchaus ausreichend. Bei hochkomplexen und produktiven, hochkritischen SAP S/4HANA-Systemen sollte jedoch eine etablierte Backup/Restore-Lösung eines Drittanbieters eingesetzt werden.

Einen besonders hohen Stellenwert in der Cloud erhält die Automatisierung. Terraform und Ansible haben wir in den Kapiteln teilweise schon betrachtet, jedoch sind deren Potenziale noch lange nicht ausgeschöpft. Viele Unternehmen folgen bereits dem Grundsatz der „Immutable Infrastructure“, welcher darauf basiert, dass neue Infrastrukturkomponenten nur noch vollständig automatisiert und getestet provisioniert werden können. Da SAP S/4HANA-Systeme jedoch sehr komplexe Gebilde sind und eine hohe Komplexität in der Architektur besitzen, wird es noch etwas dauern, bevor neue SAP S/4HANA-Systeme ohne Zutun des Administrators komplett bereitstellen lassen. Es existieren schon Möglichkeiten, wie die SAP Cloud Appliance Library, welche genau hier ansetzt, aber es darf hier noch mehr erwartet werden.

Beim Digital Decoupling werden die zusätzlichen Funktionen aus dem Kern von SAP S/4HANA-Systemen entkoppelt. Auf diesem Wege können die SAP S/4HANA-Systeme genau die Funktionen erfüllen, welche wichtig sind. Alle zusätzlichen Funktionalitäten und Integrationsszenarien werden durch weitere unterstützende Systeme bereitgestellt. Auf diesem Wege sind die SAP S/4HANA-Systeme zukunftsfähig geworden. Somit lassen sich neue Funktionen aus den Clouds einfach an die bestehenden Systeme konfigurieren. Unternehmen gehen hier schon sehr weit und implementieren die Steuerung von SAP über Amazon Alexa oder nutzen den großen Datenbestand aus den SAP-Systemen in Data Lakes, um danach pr Analytics die Daten zu analysieren.

Die Innovationskraft der Public Clouds ist enorm und es ist davon auszugehen, dass auch die SAP diese Kraft nutzen wird, um neue Features und neue Arbeitsweisen für die SAP S/4HANA-Systeme zu implementieren und anzubieten.

---

## 10.4 Ausblick

Die vorausgegangenen Kapitel haben Ihnen die Idee, die Funktion, die Konzeptionierung, die Implementierung und die Nutzung von Hyperscaler Clouds für SAP S/4HANA-Systeme im Detail näher gebracht. Es ist ersichtlich, wie einfach sich solche Systeme auf den Hyperscaler Clouds erstellen lassen und wie einfach ein Betrieb der Systeme auf den Clouds wird, wenn die immanenten Vorteile der Clouds genutzt werden.

Wir nehmen an, dass es auf absehbare Zeit nur noch sehr wenige Kunden geben wird, welche keine SAP-Systeme auf den Hyperscaler Clouds betreiben/betreiben lassen. Dieser Trend wird durch die Wirtschaftlichkeit der Umgebungen, der hohen Sicherheit,

der starken Homogenisierung und der einfachen Steuerung der SAP-Landschaften verstärkt. Auf diesem Wege können sich Unternehmen nicht mehr lange davor verschränken und werden sich früher oder später öffnen.

Die SAP setzt strategisch auf die Cloud und wird hier auch noch weitere Investitionen tätigen, um SAP S/4HANA-Systeme vollständig Cloud-fähig zu machen und insbesondere die Systemarchitektur anpassen. Auf diesem Wege ließe sich die Ausfallsicherheit, die Verfügbarkeit, die Datensicherheit und die Resilienz von den SAP S/4HANA-Systemen signifikant steigern. Alle Kunden werden hiervon profitieren.

Der Cloudmarkt wird weiterhin hochdynamisch bleiben. Dennoch kann davon ausgesehen werden, dass sich die Kräfteverhältnisse auf absehbare Zeit (3 Jahre) nicht mehr zu stark ändern werden. Neue Marktteilnehmer werden es schwer haben, einen signifikanten Anteil an dem Markt zu erlangen, auch wenn der Markt noch sehr viel Potenzial hat. Es ist sicherlich nicht verkehrt, wenn Kunden auf einen der drei großen Anbieter im Markt setzen, welche hier in diesem Buch beschrieben worden sind: Microsoft Azure, Amazon Web Services und Google Cloud. Alle drei Hyperscaler investieren weiterhin viel in die Zukunft von Public Cloud und alle Kunden (egal, ob klein oder groß) werden hiervon profitieren können.