

# **SAS® IN DER AMAZON CLOUD (AWS) BEST PRACTISES**

**DIRK JOHANN, SYSTEMARCHITEKT**



- SAS® on AWS - Einführung
- SAS® on AWS - Grundsätze
- SAS® on AWS - Best Practices
- Demo
- Fragen!



## SAS® ON AWS - EINFÜHRUNG

- Mythen
- Quiz

Cloud / AWS ≠ Magic.

- Amazon weiß nicht, welche Software auf ihren Instanzen läuft - und es ist Amazon egal.
- Für AWS benötigt man erfahrene Administratoren.
- Bei AWS (und anderen Cloud-Deployments) muss man sich allen Security-Aspekten bewusst sein.
- Wenn eine Applikation nicht skalierbar ist, wird es durch AWS nicht besser.
- AWS / Cloud ist nicht immer billiger.

Welche Aussage stimmt?

- Alle SAS-Produkte und Lösungen sind bei allen Cloud-Providern supported
- Keine SAS-Produkte und Lösungen sind bei Cloud-Providern supported
- Alle SAS-Produkte und Lösungen sind nur auf AWS supported
- It depends



## SAS® ON AWS - GRUNDSÄTZE

- SAS®-Version für Amazon AWS.
- Support für SAS® on AWS.
- Allgemeine Grundsätze.

- Wir haben keine spezielle AWS-/Cloud-Version von SAS®.
- Aber: wir können SAS® in einer Cloud-Infrastruktur einsetzen:
  - Durch den Aufbau einer Umgebung, die einer on-premise / traditionellen Umgebung entspricht
  - und dann eine normale SAS®-Installation durchführen.
- SAS® 9.4 ist besser verträglich mit Cloud / AWS als SAS® 9.3 (siehe multicast).
- SAS® Viya hat „Cloud“ in den Genen und ist nochmals besser geeignet.

Aber natürlich: siehe [Virtualization Policy](#)

(Cloud/AWS ist nichts anderes als eine andere Form der Virtualisierung)

- Aber:
  - Supported ≠ Tested ↳ SAS®-Produkt / jede Version ist getestet
  - Supported heisst
    - Wenn es ein Problem gibt, kann man es nachstellen können
- Beispiel:
  - AWS erlaubt keine SAS®-Produkte
  - Viele Cloud Anbieter haben nur SAS®-Produkte bereit

Dies ist der Grund, warum man ein tiefes Verständnis von AWS benötigt!

There is no difference  
between theory and practice.  
In theory there is.

terra

ent

Deployments).

(Grid!), bei AWS

## Statement of support

- SAS® is supported on a Cloud IaaS provider (AWS, OpenStack, etc.) as long as that provider fulfills the requirements of the given SAS® solution.
- Not all SAS® solutions are supported in cloud infrastructures or have additional requirements that organizations need to adhere to. Example:

Solution	Notes/Limitation
SAS Grid Manager	 Whilst SAS Grid Manager is supported in the cloud, there is a dependency on a performant shared clustered file system being available

Der Prozess für ein Deployment von SAS® auf AWS ist der gleiche wie bei einem On-Premise-Deployment:

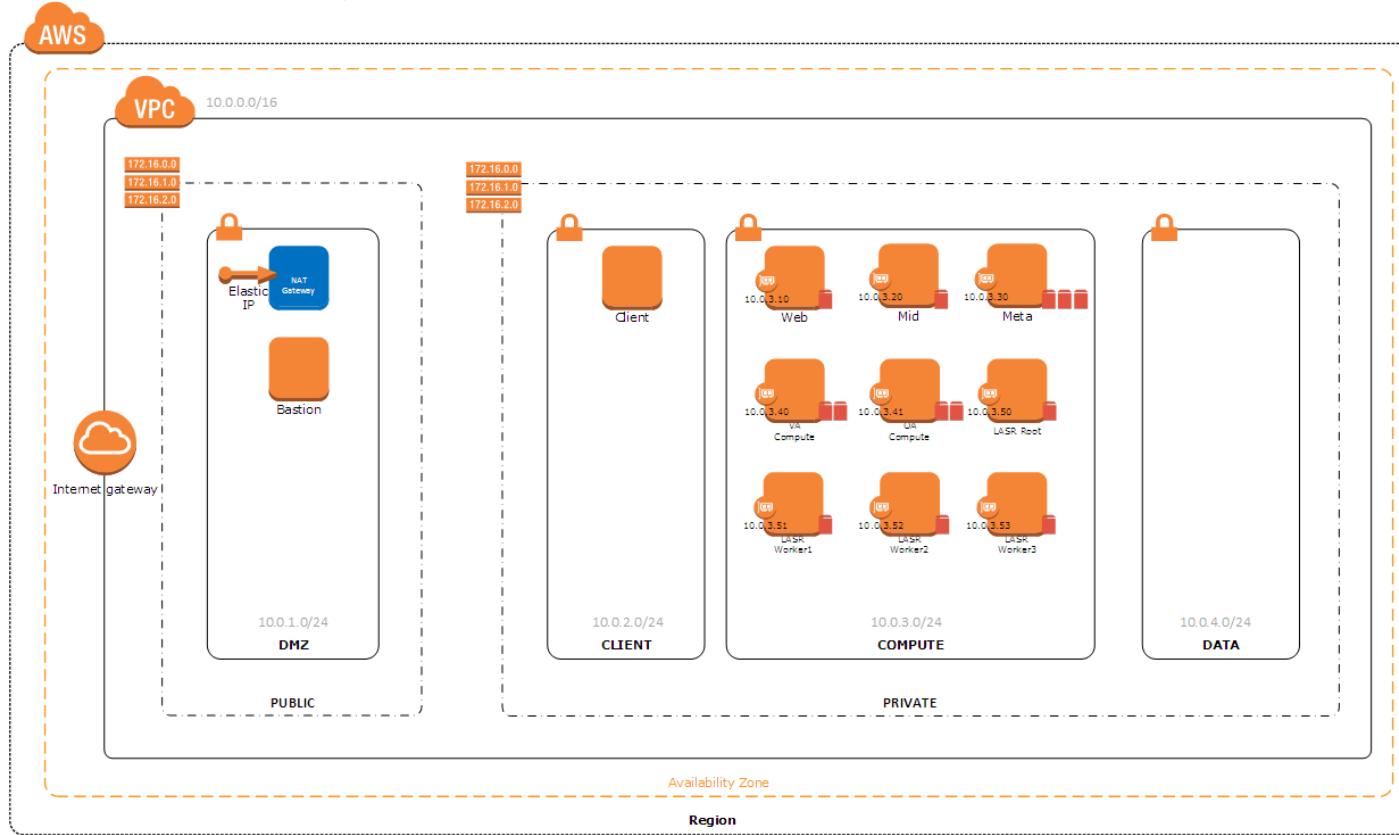
- Environment-Sizing und Architektur-Design ist notwendig.
- Die SAS® Software-Lizenierung ist in der Cloud die gleiche wie bei On-Premise (Besonderheit: AWS  $\Rightarrow$  halbe Cores).
- Cloud Administration, Security und Monitoring liegen im Verantwortungsbereich des Kunden.
- SAS® -Applikationen können Skalierungsmöglichkeiten von AWS nutzen. Wichtig ist aber, die Lizenzbedingungen zu beachten, was meist mit einer Outage verbunden ist.
- Themen wie Hochverfügbarkeit, Disaster Recover, Backup und Restore sind auf AWS genauso wichtig wie in klassischen SAS® -Deployments.



## SAS® ON AWS - BEST PRACTICES

- Referenzarchitektur:
  - SAS® 9.4 Office Analytics,
  - SAS® Visual Analytics,
  - SAS® Grid Manager mit LSF auf AWS.
- Kosten, Kosten und Kosten für SAS® on AWS.
- Security - Network - Data.
- Connectivity.
- Deployment und Verfügbarkeit.

# SAS® ON AWS REFERENZARCHITEKTUR



## Server-Kosten bei AWS

- Um die Kosten der AWS-Infrastruktur zu minimieren, kann man unterschiedliche Strategien nutzen. Diese Strategien werden hier diskutiert:  
Shut Down When Not In Use - <https://aws.amazon.com/premiumsupport/knowledge-center/stop-start-ec2-instances>
- Wenn die SAS®-Produktionsumgebung nur während der Geschäftszeiten genutzt wird, macht es aus Kostengründen Sinn, die Instanzen nachts herunterzufahren. Dies sollte automatisiert erfolgen, damit morgens die SAS®-Umgebung direkt bereitsteht, wenn sie benötigt wird.
- Ebenso sollte bei nicht-produktiven Umgebungen die Nutzung geprüft werden. Test-Umgebungen können provisioniert werden, wenn sie benötigt und ansonsten gestoppt werden. Wenn die Tests abgeschlossen sind und die Inhalte per Promotion nach Produktion übertragen wurden, können die Testumgebungen sogar terminiert und gelöscht werden.

## AWS Storage-Kosten

- AWS bietet unterschiedliche Storage-Formen (EBS, S3, ...). Diese unterscheiden sich in Bezug auf Kapazität, Limitierungen, Geschwindigkeit und Persistenz, aber auch in Bezug auf die anfallenden Kosten.
- Wenn man eine Instanz stoppt, wird der assoziierte Speicher nicht gelöscht. Das heißt, dass Storage-Kosten weiter anfallen, auch wenn die zugehörigen Instanzen gestoppt sind.
- Storage, der initialisiert wird, wenn eine Server-Instanz (durch-)gestartet wird (sog. Ephemeral Storage) ist keine gute Wahl für Anwendungs-Software oder -Daten, aber gut für temporären Speicher wie SAS® WORK und SAS UTIL.

## Data Transfer Kosten

- Bei AWS fallen keine Kosten für eingehende Daten an (Inbound).
- Sehr wohl fallen Kosten für ausgehende Daten oder den Datentransfer zwischen AWS-Zonen an (Outbound)! Dies kann dramatische Auswirkungen auf die Kosten haben, wenn Anwendungen große Datenmengen austauschen. Bei SAS®-Anwendungen ist es durchaus üblich, dass grosse Datenmengen aus Datenbanken oder SAS®-Datasets extrahiert werden.
- In der folgenden Tabelle sind Links zu Pricing-Informationen für AWS enthalten:

Cloud Provider	Link to Pricing Policy
Amazon Web Services Simple Monthly Calculator	<a href="http://calculator.s3.amazonaws.com/index.html">http://calculator.s3.amazonaws.com/index.html</a>
External Amazon AWS Calculator	<a href="http://www.ec2instances.info/">http://www.ec2instances.info/</a>

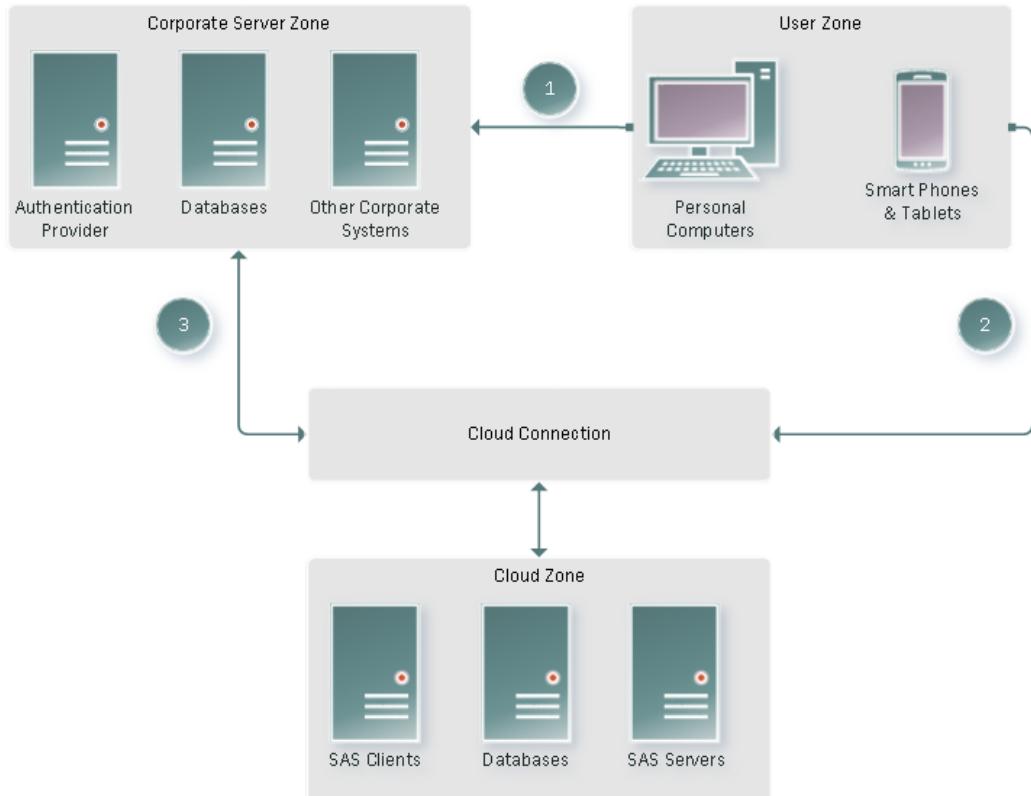
## Security, Networking und Data Location

- Server-Instanzen sollten nicht aus dem Public Internet erreichbar sein.  
3<sup>rd</sup>-Party-Zugriff sollte immer durch ein VPN erfolgen.
- Die Verbindung zwischen eigener On-Premise-Umgebung und AWS ist ein wichtiger Punkt in der Planung und sollte frühzeitig adressiert werden.
- Ausgehender Datentransfer kostet etwas, daher sollten Daten- und Server-Lokation frühzeitig geklärt werden.

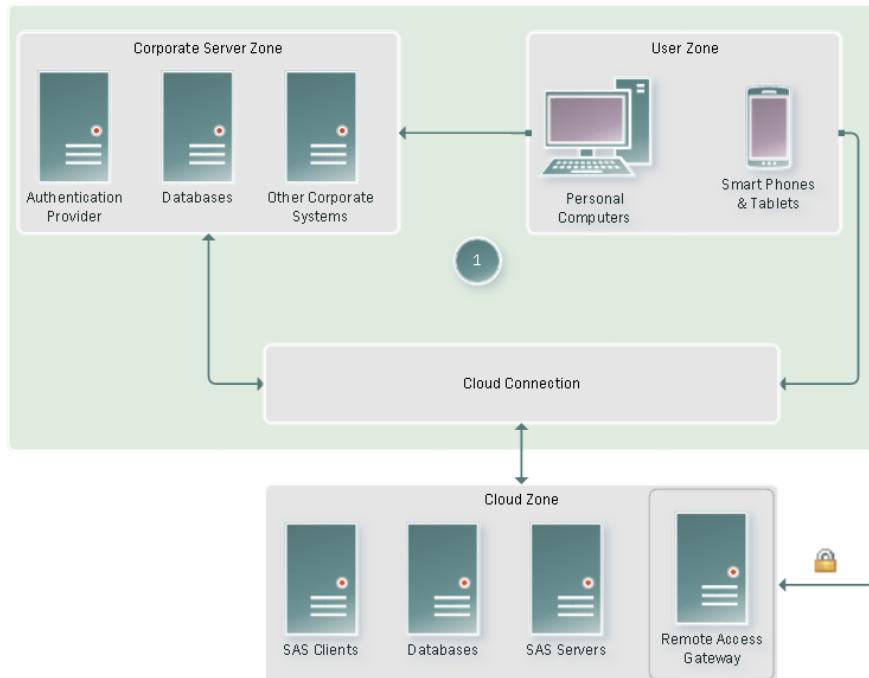
# SAS® ON AWS CONNECTIVITY

Drei wesentliche Zonen:

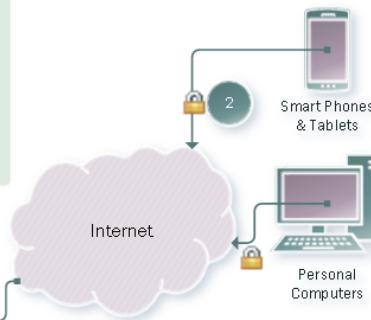
- User Zone
- Corporate Server Zone
- Cloud Zone



## Point to Site Connections

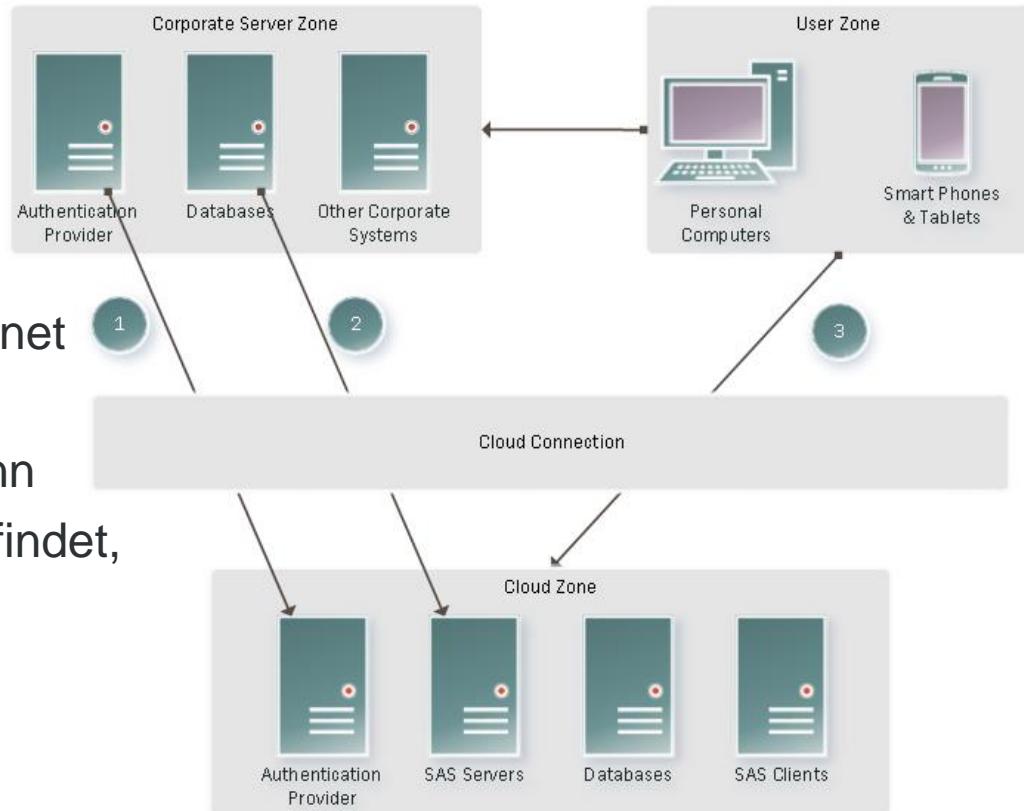


Verbindung zum Cloud-Environment wird direkt vom Client über eine gesicherte Verbindung aufgebaut



## Site to Site Connections

Verbindung zum Cloud-Environment wird vom Intranet über eine gesicherte VPN-Verbindung aufgebaut. Wenn man sich im Firmennetz befindet, ist ein Zugriff auf die Cloud-Ressourcen möglich.



## **AWS Direct Link - als besondere Site-to-Site-Connection**

- Reduziert Transferkosten zwischen Cloud und On-Premise-Environment.
- Gesicherte Verbindung zwischen beiden Environments mit garantierter Bandbreite.
- Zusätzliche permanente Kosten für die Verbindung.

- Deployment:
  - Automatisiertes Deployment möglich (CloudFormation, Ansible).
  - Automation / Scripting notwendig, um Umgebung reproduzierbar zu machen.
- Verfügbarkeit:
  - Verteilung auf mehrere Zonen in AWS notwendig, um Verfügbarkeit (keine Hochverfügbarkeit!) zu ermöglichen.
  - Bei der Auswahl der Zonen Unterschiede, Regulatorien und Performance beachten.



## SAS® ON AWS - DEMO

Können wir mal live sehen, wie so ein Deployment aussieht?

- CloudFormation,
- Ansible.

Topic	Link
Public Paper VA on Cloud	<a href="http://support.sas.com/resources/papers/proceedings_14/SAS298-2014.pdf">http://support.sas.com/resources/papers/proceedings_14/SAS298-2014.pdf</a>
Ways to Connect to Amazon Redshift with SAS/ACCESS	<a href="http://support.sas.com/resources/papers/proceedings_15/SAS1789-2015.pdf">http://support.sas.com/resources/papers/proceedings_15/SAS1789-2015.pdf</a>
SAS Grid Manager with Lustre on AWS	<a href="http://support.sas.com/rnd/scalability/grid/SGMonAWS.pdf">http://support.sas.com/rnd/scalability/grid/SGMonAWS.pdf</a>
AWS essentials	<a href="http://aws.amazon.com/training/course-descriptions/essentials/">http://aws.amazon.com/training/course-descriptions/essentials/</a>
Architecting on AWS	<a href="http://smportal.sas.com/training/Pages/ArchitectingonAWS.aspx">http://smportal.sas.com/training/Pages/ArchitectingonAWS.aspx</a>
Simple explanation of VA on Cloud	<a href="http://sasgf.v.sas.com/detail/videos/daily-reports/video/4184916822001/sas-tech-talk:-sas-in-amazon-cloud?autoStart=true">http://sasgf.v.sas.com/detail/videos/daily-reports/video/4184916822001/sas-tech-talk:-sas-in-amazon-cloud?autoStart=true</a>

- SAS ist cloud-ready, Besonderheiten bei Lösungen sind zu beachten.
- Das Kostenmodell von Amazon AWS ist beim Deployment und Betrieb von SAS zu beachten.
- Ein besonderes Augenmerk ist auf Security-Aspekte zu richten.
- Verfügbarkeitsanforderungen sind bei der Auswahl der Zone zu berücksichtigen.

- Wie sehen Ihre Cloud-Pläne aus?
- Inwieweit betrifft dies die SAS®-Installation?
- Welche Erwartungen haben Sie an SAS® bezüglich Cloud?

