



ARQUITETURA DE SOFTWARE EM NUVEM

Tiago Coelho Ferreto - Aula 03

Professores

MANOEL VERAS

Professor Convidado

Graduado em Engenharia Elétrica pela Universidade Federal do Rio Grande do Norte (UFRN), é mestre em Engenharia Elétrica pela Universidade Estadual de Campinas (UNICAMP) e doutor em Administração na Universidade de São Paulo (USP). Atualmente, é professor e consultor vinculado ao Departamento de Administração da Universidade Federal do Rio Grande do Norte (UFRN). Atua também no Programa de Pós-Graduação em Administração como coordenador, com vínculo com a linha de pesquisa Estratégia, Gestão e Tecnologia da Informação. É certificado como PMP pelo PMI desde 2001, ITIL pelo itSMF e COBIT pelo ISACA.

TIAGO COELHO FERRETO

Professor PUCRS

Professor adjunto da Pontifícia Universidade Católica do Rio Grande do Sul, possui doutorado em Ciência da Computação pela PUCRS (2010), com doutorado-sanduíche na Technische Universität Berlin, na Alemanha (2007-2008). Tem experiência na área de Ciência da Computação, com ênfase em Redes de Computadores, atuando principalmente nos seguintes temas: computação em nuvem, grades computacionais, virtualização, processamento de alto desempenho e gerência de infraestrutura de TI.

Ementa da disciplina

Características do paradigma de computação em nuvem (self-service, elasticidade, pay-as-you-go). Estudo dos modelos de computação em nuvem: SaaS, PaaS e IaaS. Análise dos principais provedores de nuvem pública (AWS, Google e Azure) e seus serviços. Estudo sobre arquiteturas de software centradas na nuvem (Cloud-Native Architecture). Estudo de casos de aplicações que utilizam computação em nuvem. Desafios e tendências em computação em nuvem.

PUCRS online

Desenvolvimento Full Stack

Arquitetura De Software Em Nuvem

Tiago Ferreto

tiago.ferreto@pucrs.br

Quem sou eu?

Tiago Ferreto



<https://www.linkedin.com/in/tiagoferreto>



tiago.ferreto@pucrs.br



PUCRS online

PÓS-GRADUAÇÃO

Ciência de Dados e
Inteligência Artificial

PPGCC/PUCRS

Programa de Pós-Graduação
em Ciência da Computação

https://bit.ly/ppgcc_pucrs

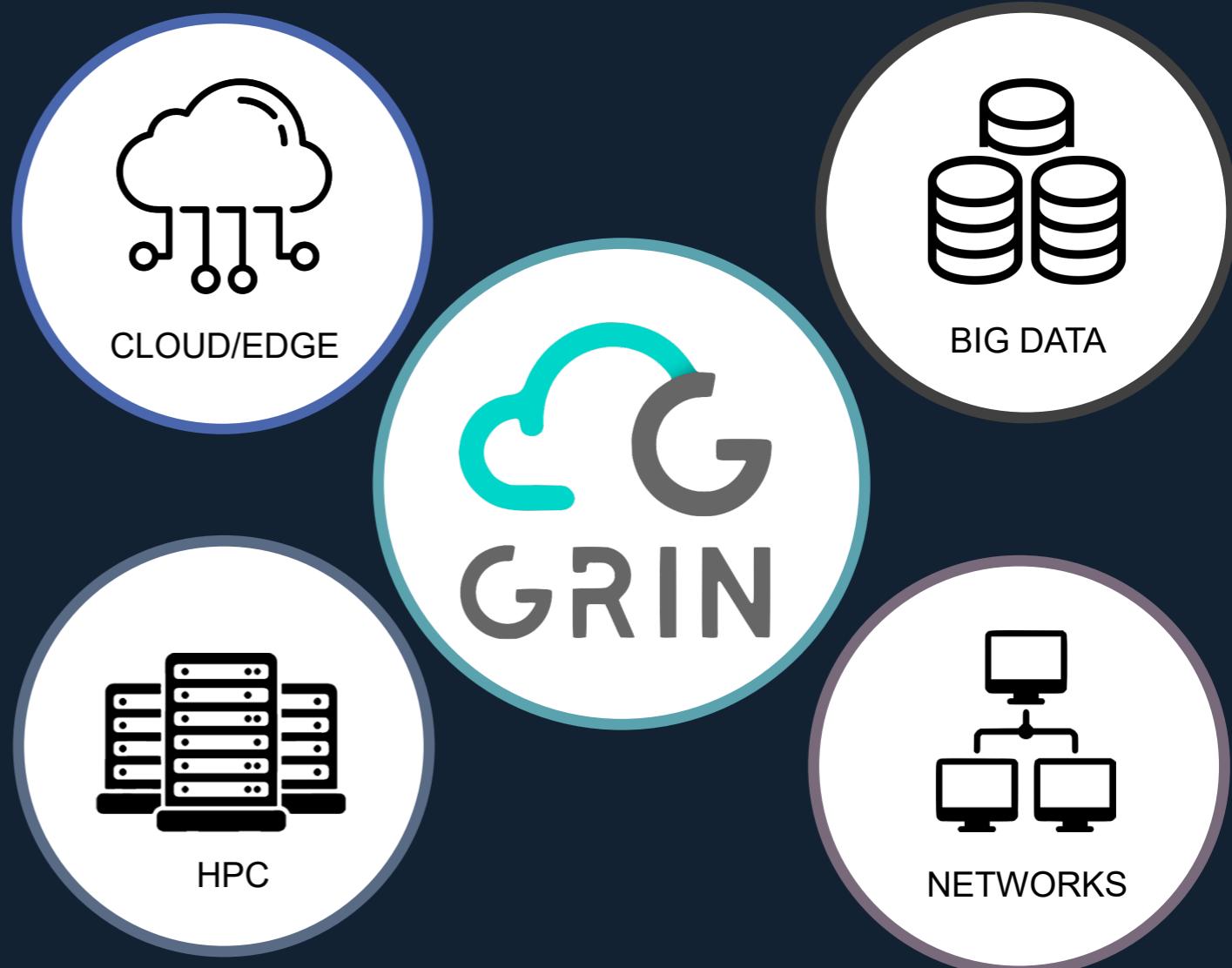


Grupo de Pesquisa
em Redes, Infraestrutura e
Computação em Nuvem

https://bit.ly/grin_pucrs

GRIN – Grupo de Pesquisa em Redes, Infraestrutura e Computação em Nuvem

- Projetos de **pesquisa, desenvolvimento e inovação**
- **Formação e capacitação**
 - Mestrado e Doutorado (PPGCC/PUCRS)
- Atuação em **projetos governamentais** e em **colaboração com empresas**

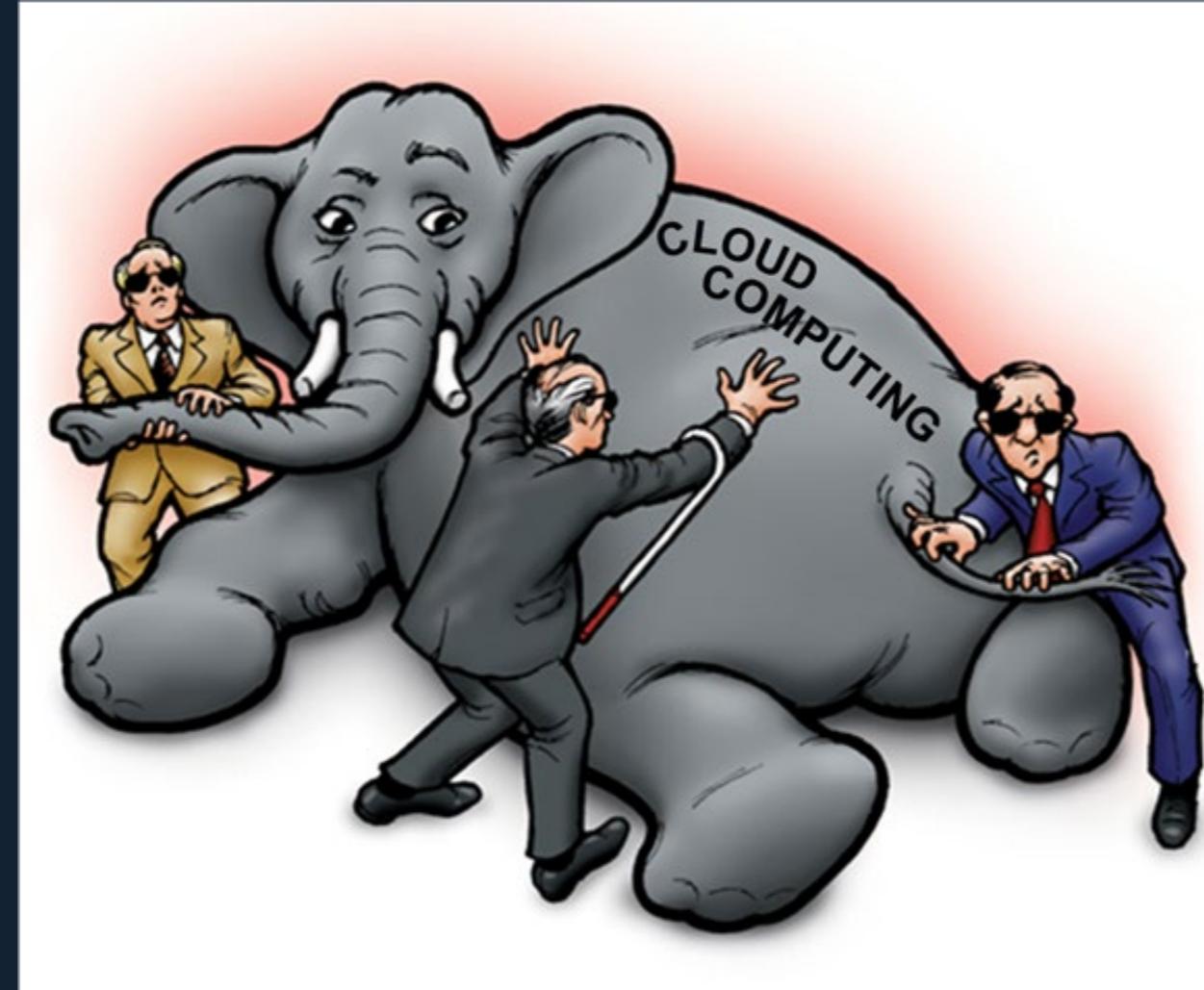


Vamos falar de nuvem...



<https://dilbert.com/strip/2011-01-07>

O que é Computação em Nuvem?



<https://www.cloudpulsestrat.com/posts/the-blind-men-and-cloud>

HISTÓRICO

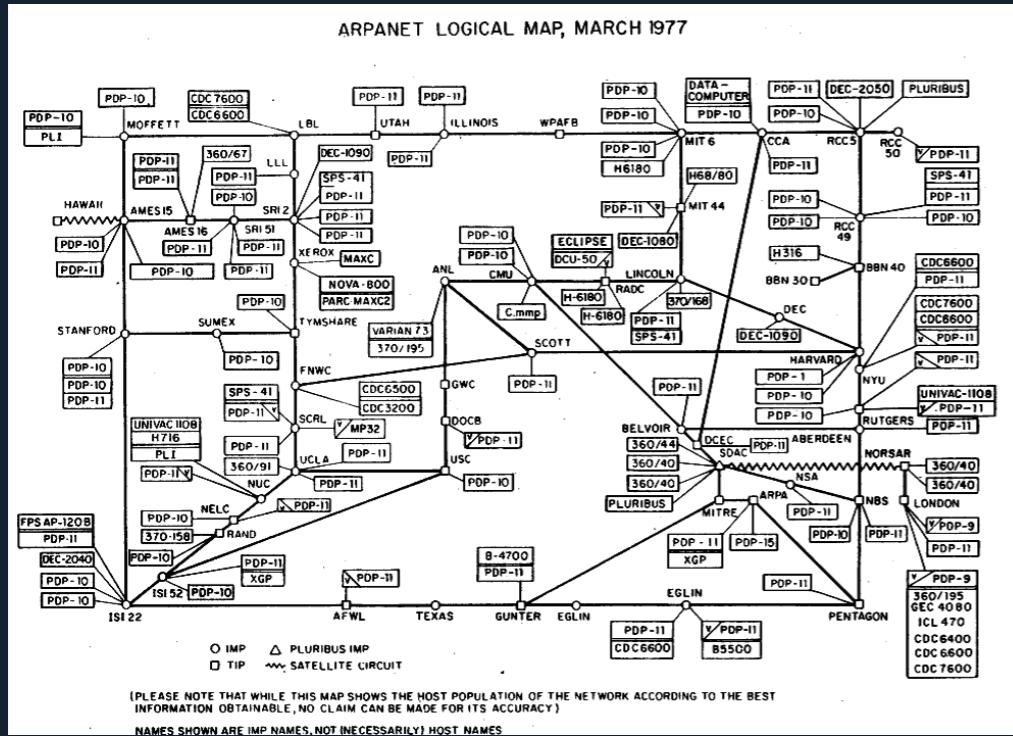
Histórico – Mainframes (Anos 60 e 70)

- IBM System/360
- Sistema operacional CP/CMS
- Utilização de máquinas virtuais
 - Time sharing – compartilhamento da máquina entre diversos usuários através de “terminais burros”

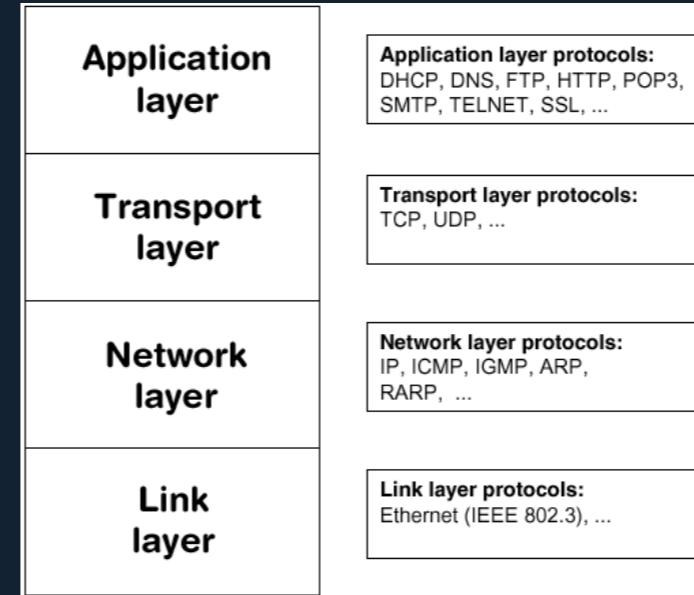


Histórico – Internet (Anos 70 e 80)

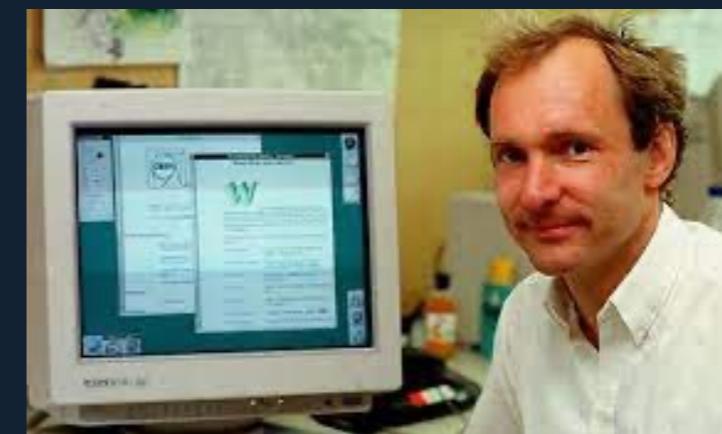
- ARPANET
- TCP/IP
- WWW – World Wide Web



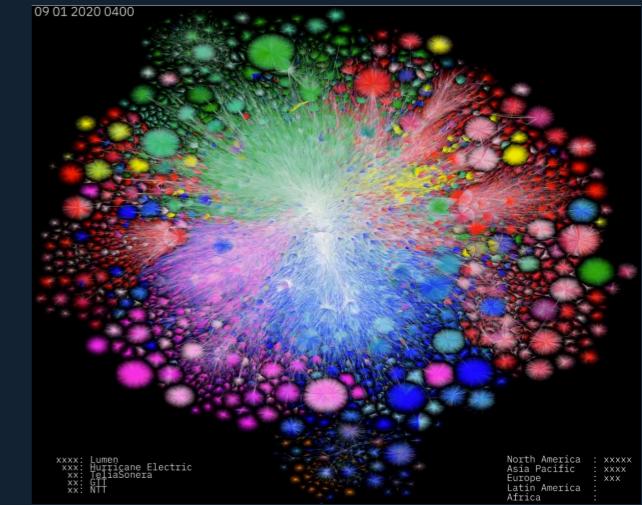
https://commons.wikimedia.org/wiki/File:Arpanet_logical_map,_march_1977.png



<https://www.phaedsys.com/principals/segger/seggeremtcip.html>



<https://home.cern/science/computing/birth-web>



<https://www.opte.org/the-internet>

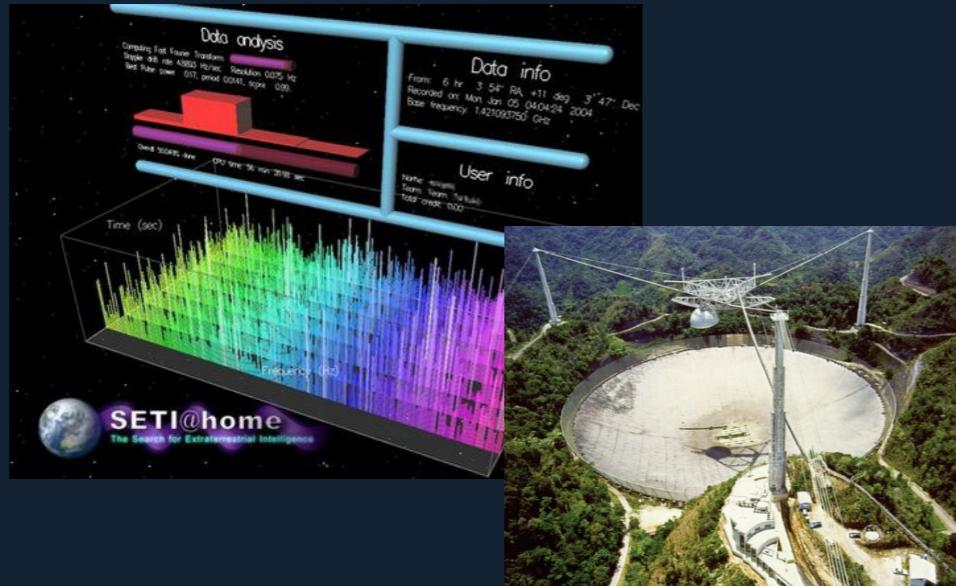
Histórico – PCs e Supercomputadores (Anos 80 e 90)

- Máquinas com menor custo
- Supercomputadores
- Cluster computing



Histórico – Grid computing (Anos 90 e 2000)

- Demanda por maior capacidade computacional
- Aplicações com demandas por recursos remotos (ex. Banco de dados, Telescópio)
- Utility computing
 - Modelos de economia para “leasing” de recursos
 - Acesso sob demanda
 - “Recursos infinitos”

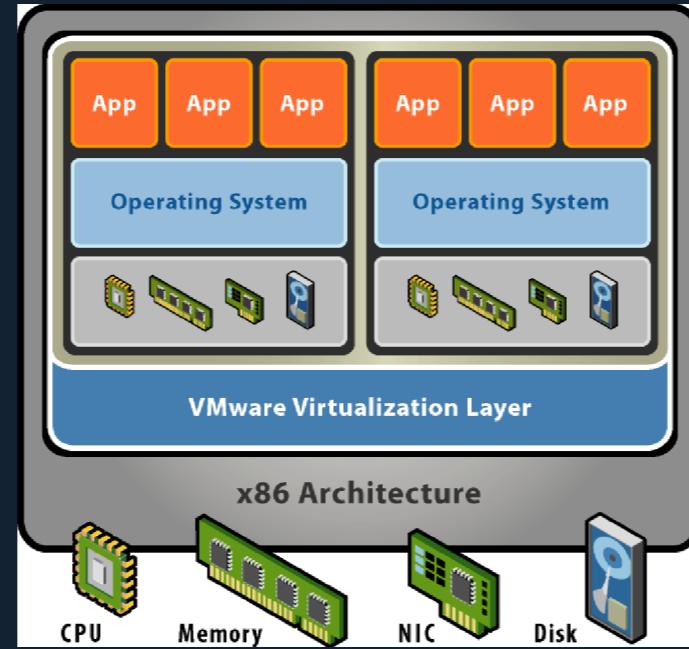
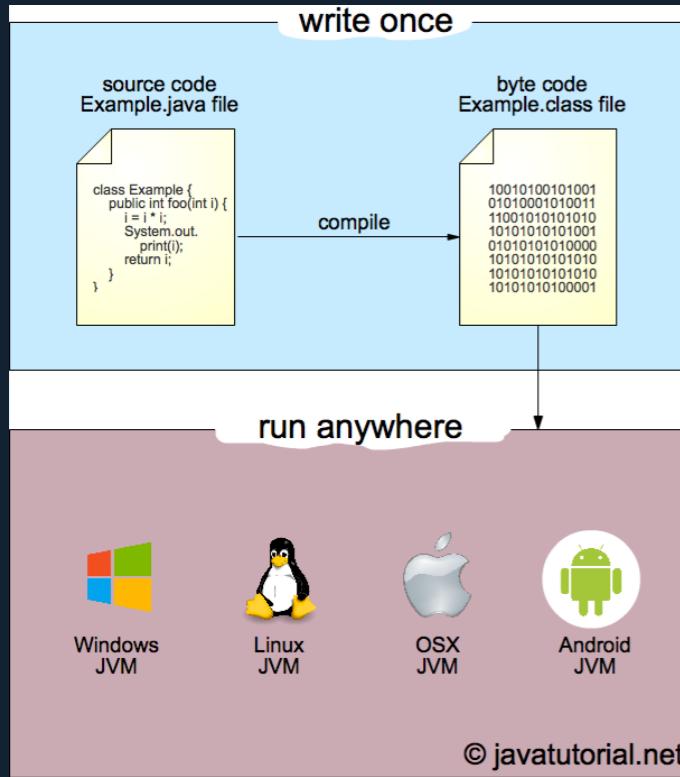


Histórico – Virtualização (Anos 90 e 2000)

- Aplicações multi-plataforma (Java)
- Compartilhamento de recursos (VMWare, Xen)
 - Uso intensivo em data centers



<https://www.vmware.com/br/timeline.html>



Principais motivadores para Computação em Nuvem

- Reduzir os investimentos com infraestrutura
 - Custo inicial fixo e alto
 - Necessidade de investimentos contínuos
 - Prejudicam o foco no core business
- Evitar a subutilização de recursos
 - Data centers com aproximadamente 20% de utilização
- Necessidade de uma plataforma única
 - Garantias de fácil acesso, alta disponibilidade, escalabilidade, segurança

CONCEITO E CARACTERÍSTICAS

O que é Computação em Nuvem?

“é um **modelo pay-per-use** para permitir o acesso pela rede **sob demanda**, com disponibilidade e conveniência a um **conjunto compartilhado de recursos computacionais** configuráveis (e.g., redes, servidores, armazenamento, aplicações, serviços) que podem ser **rapidamente provisionados e liberados** com custo de gerência e interação mínimos com o provedor de serviços.”
(NIST 2009)



Special Publication 800-145

The NIST Definition of Cloud Computing

Recommendations of the National Institute of Standards and Technology

Peter Mell
Timothy Grance

O que é Computação em Nuvem?



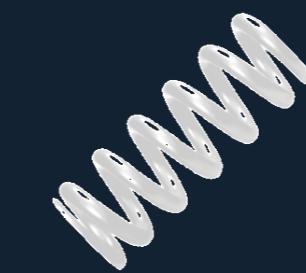
Pay-per-use



Multitenant

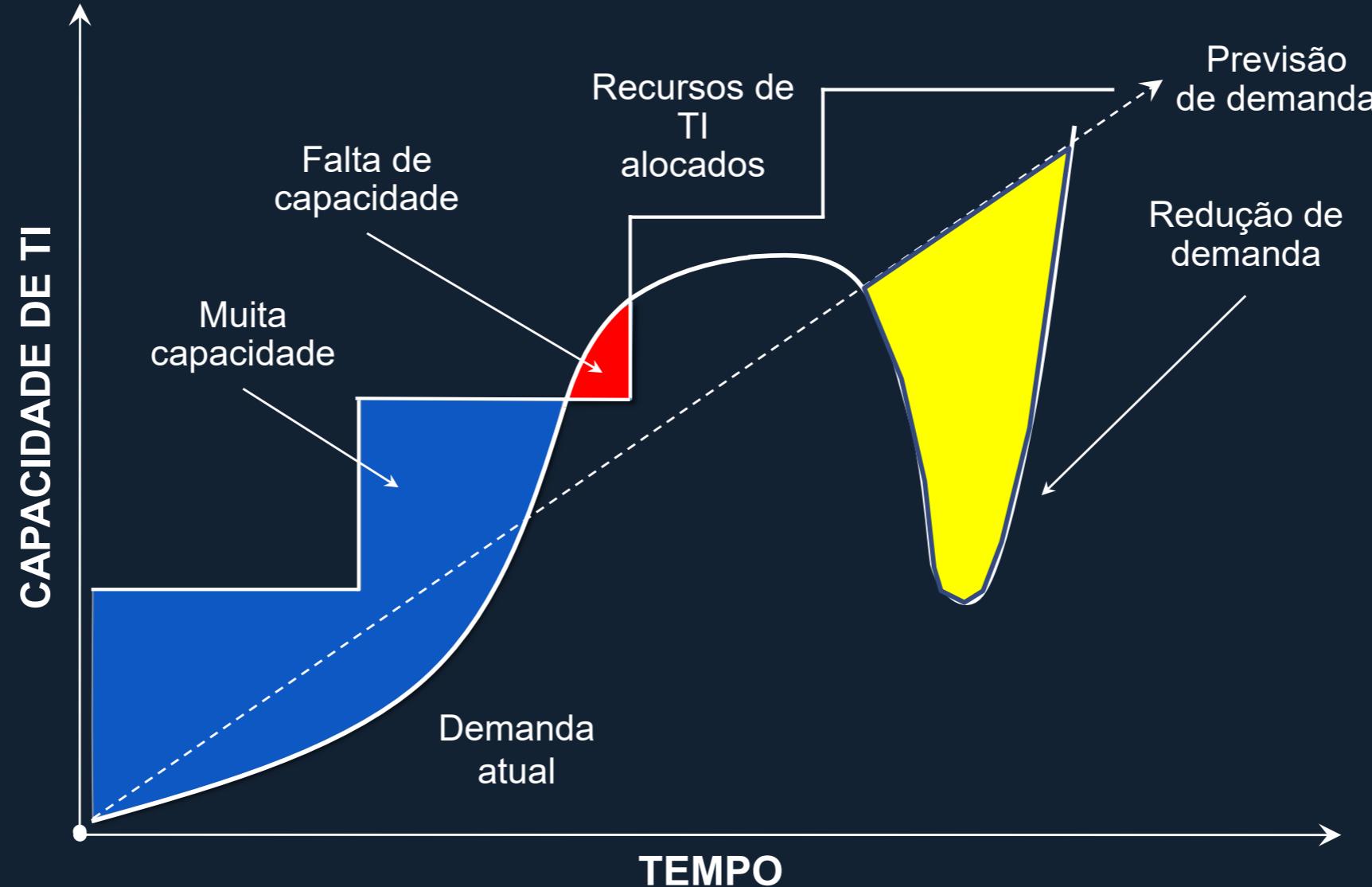


Self-service



Elasticidade

Gerência de infraestrutura tradicional



Gerência de infraestrutura na Nuvem



Infraestrutura local x Cloud Computing



Modelos de serviço

CONSUME



SaaS – Software como Serviço

DEVELOP



PaaS – Plataforma como Serviço

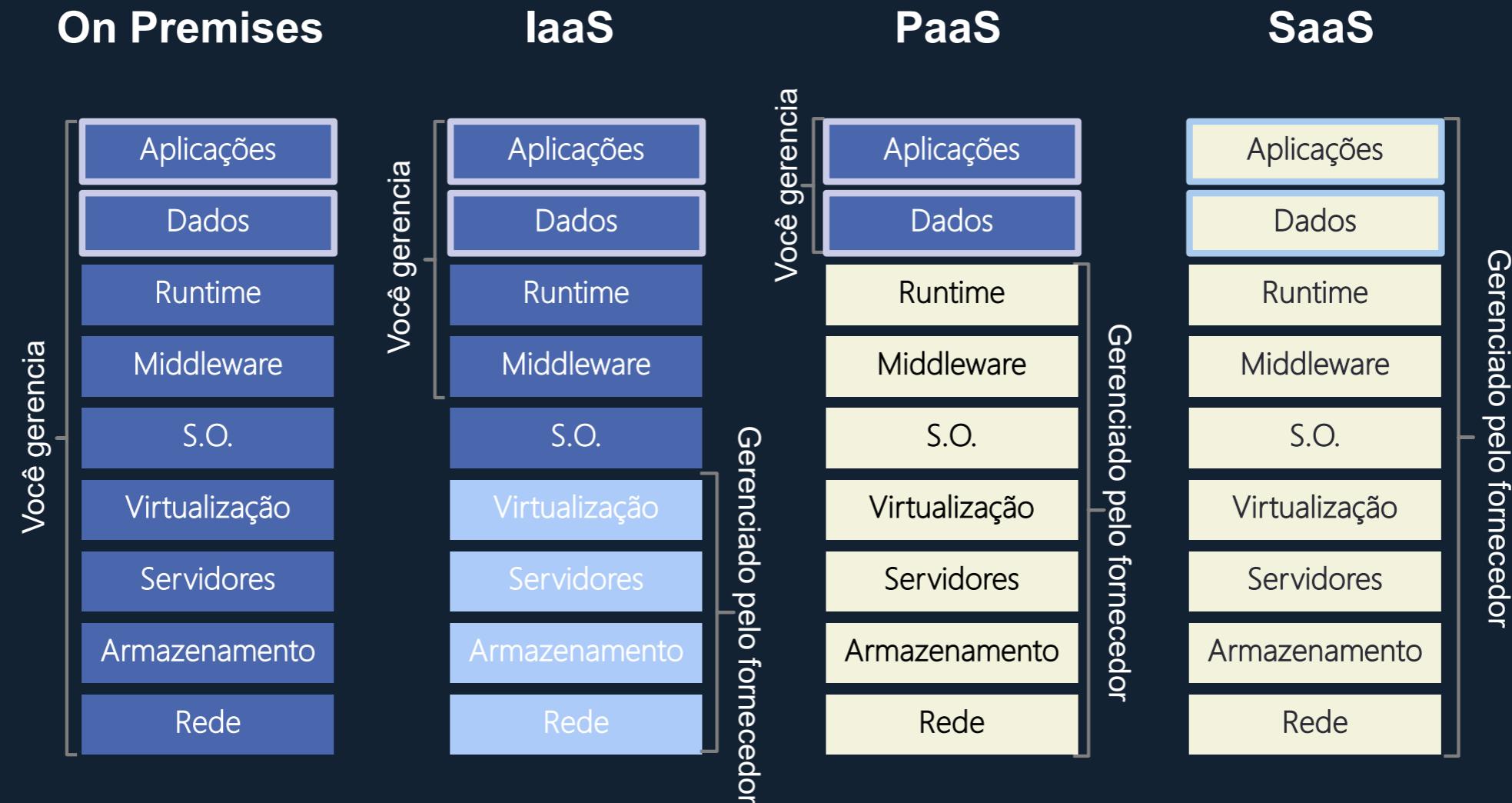
BUILD



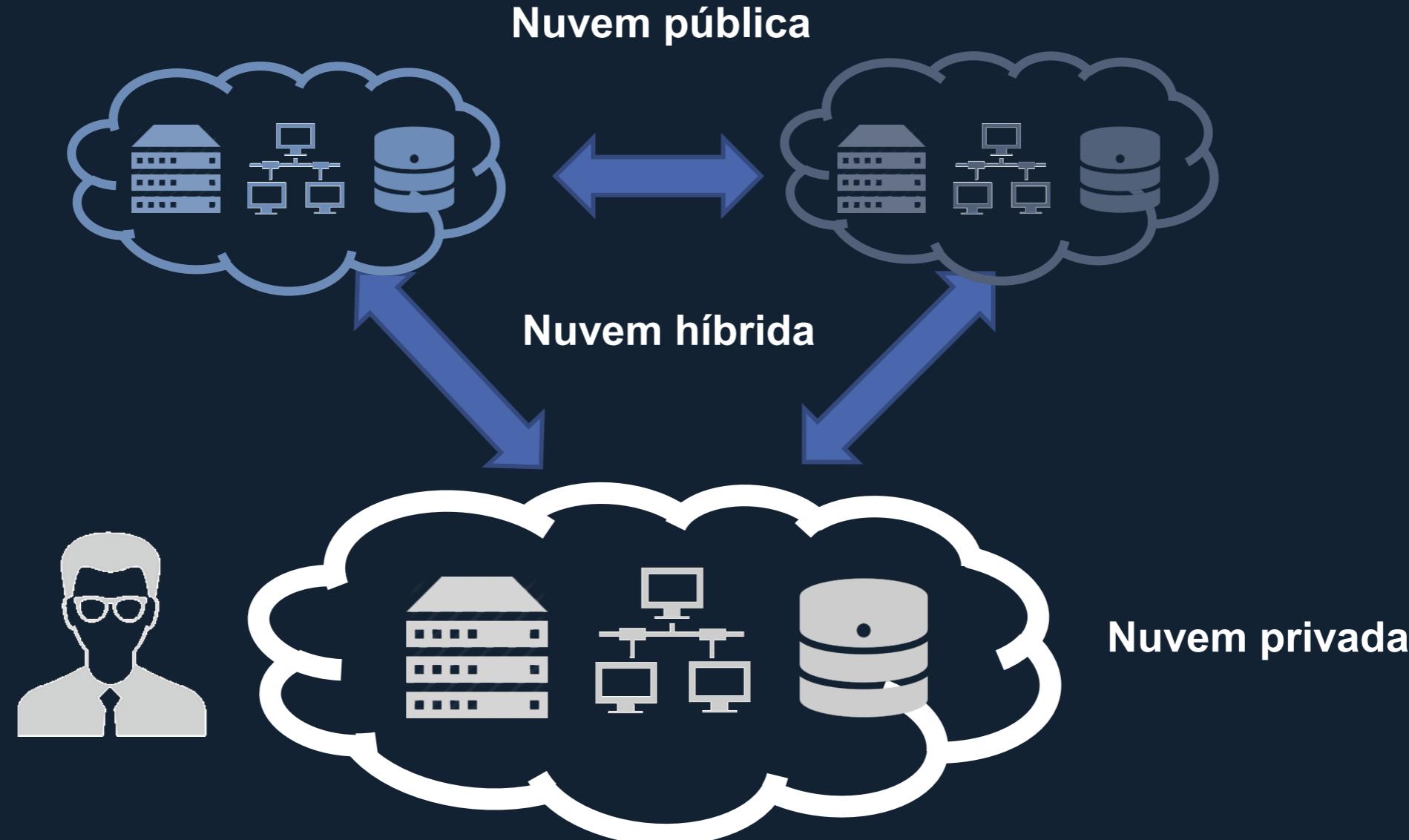
IaaS – Infraestrutura como Serviço

- Aplicações e serviços disponibilizados aos usuários pela Internet usando uma infraestrutura escalável
- Ambiente com suporte para desenvolvimento de aplicações na nuvem
- Gerência baseada nas aplicações e não nos recursos computacionais
- Provisionamento de recursos computacionais básicos
- Usuário é responsável por instalar e gerenciar seus sistemas nos recursos alocados

Modelos de serviço - responsabilidades



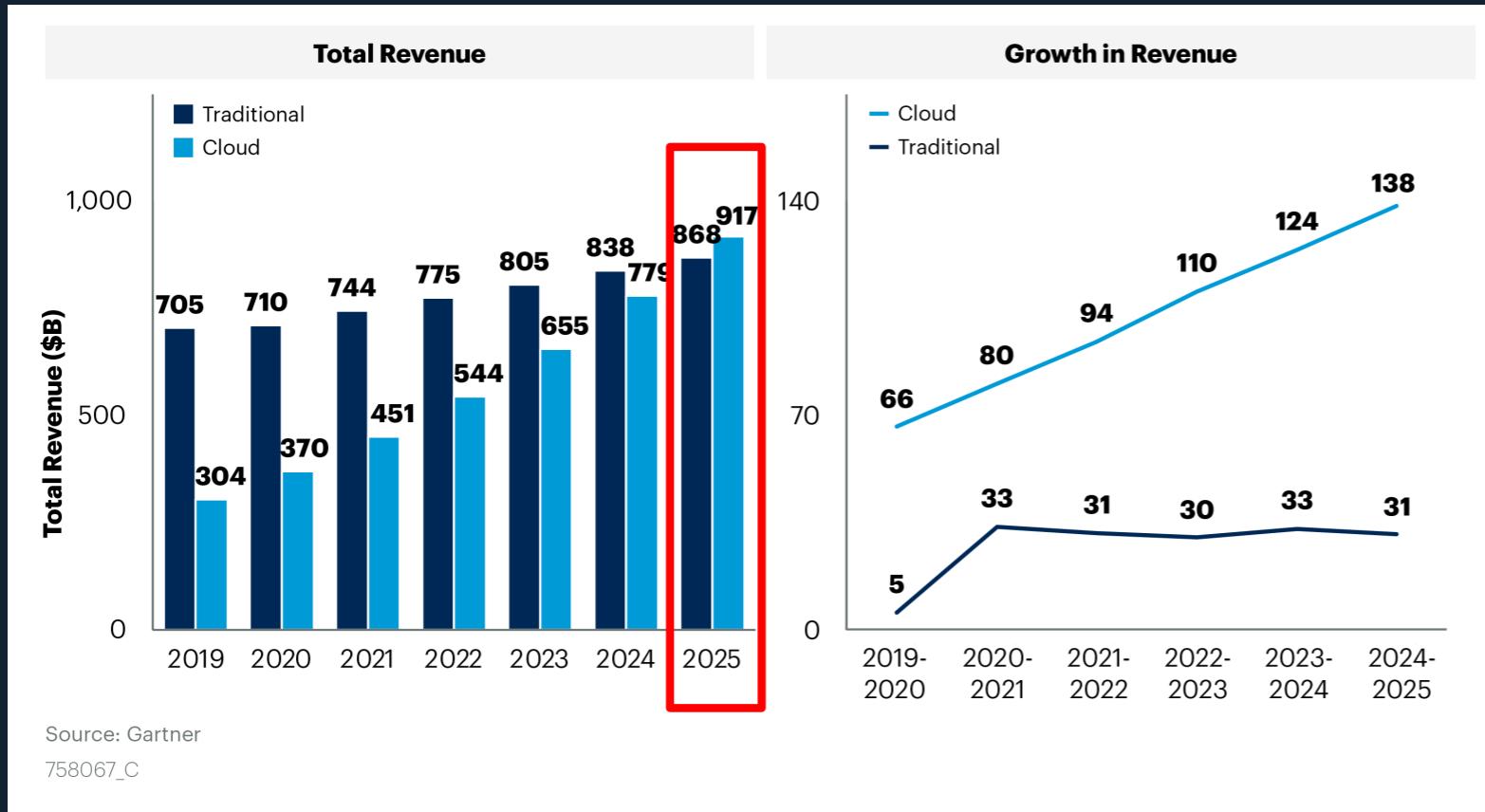
Modelos de implantação



NUVENS PÚBLICAS

Adoção de nuvens públicas

Despesas em TI pelas organizações – tradicional x nuvem pública



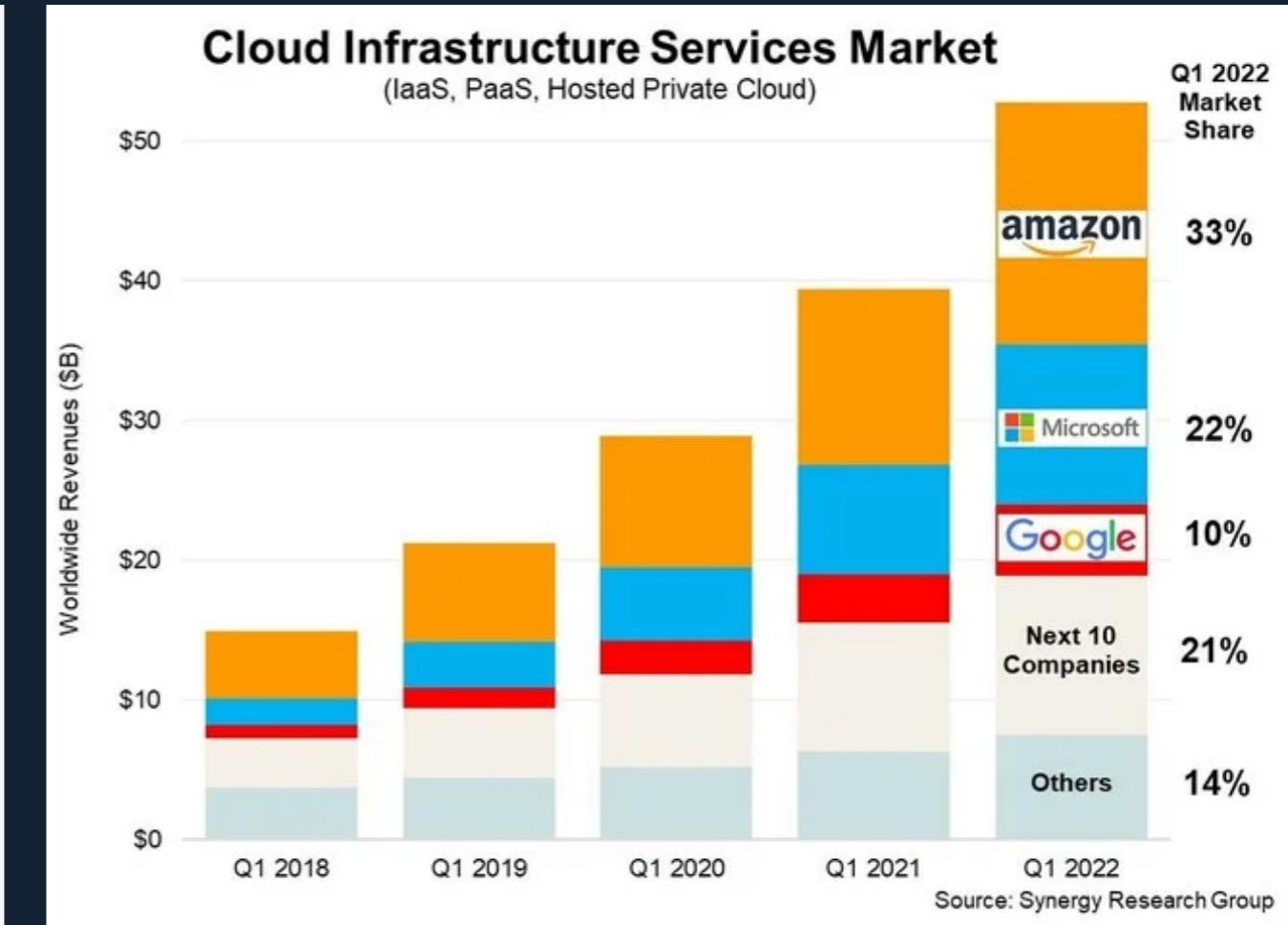
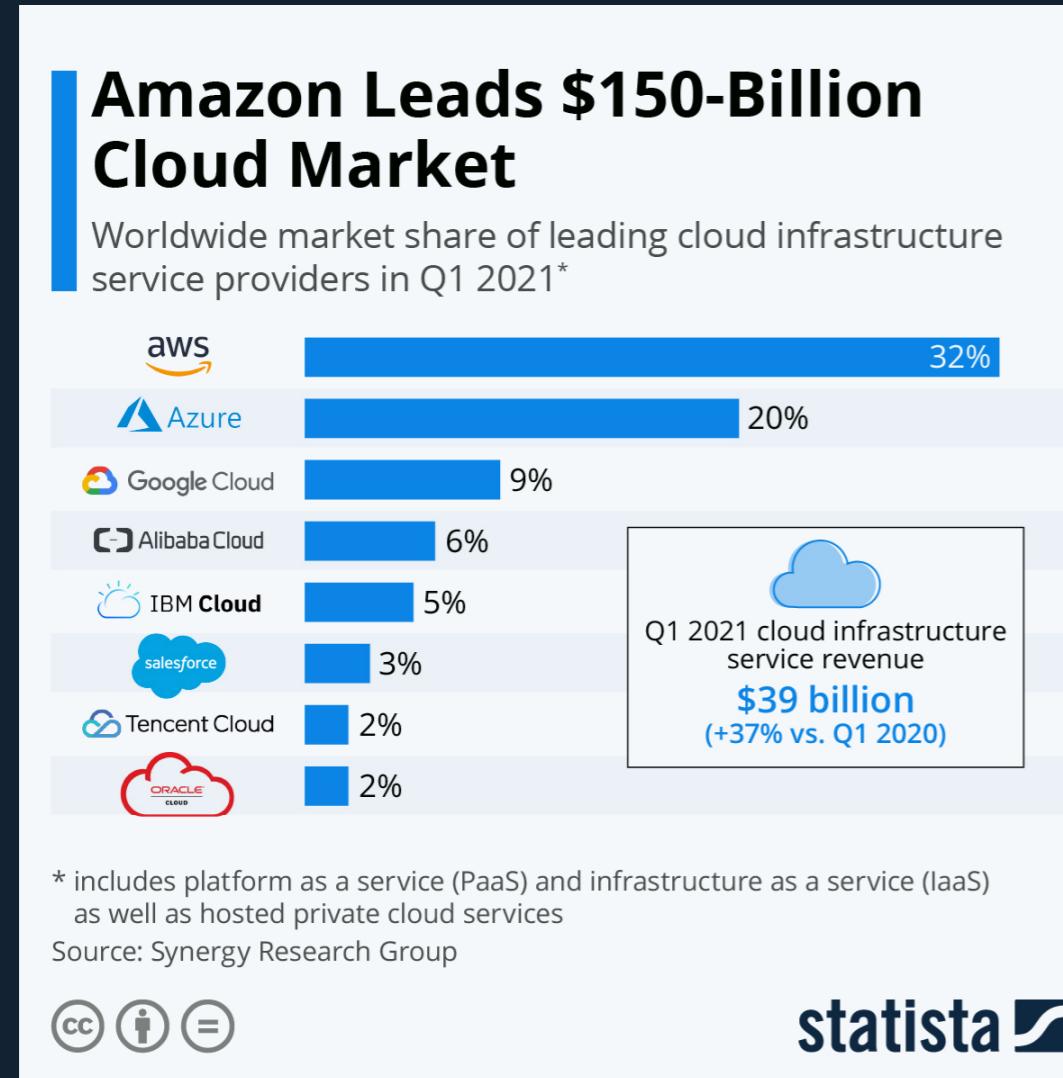
<https://www.gartner.com/en/newsroom/press-releases/2022-02-09-gartner-says-more-than-half-of-enterprise-it-spending>

Quadrante mágico do Gartner



<https://aws.amazon.com/pt/resources/analyst-reports/gartner-mq-cips-2021/>

Adoção de nuvens públicas



https://www.theregister.com/2022/05/02/cloud_market_share_q1_2022/

Amazon Web Services (AWS)



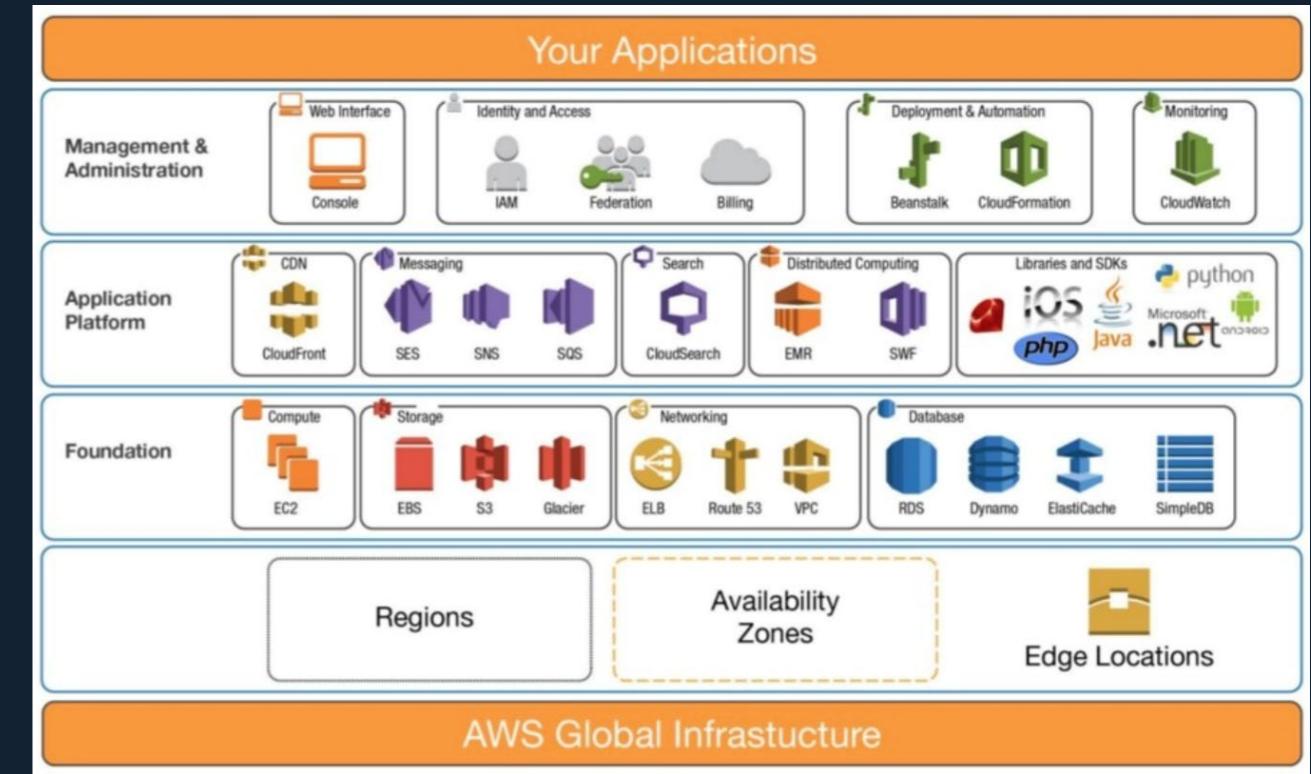
- Plataforma de computação na nuvem oferecida pela Amazon.com

- <https://aws.amazon.com>

- Primeiros serviços

- SQS (comunicação) – 2004
- EC2 (computação) – 2006
- S3 (armazenamento) – 2006

- Usuários

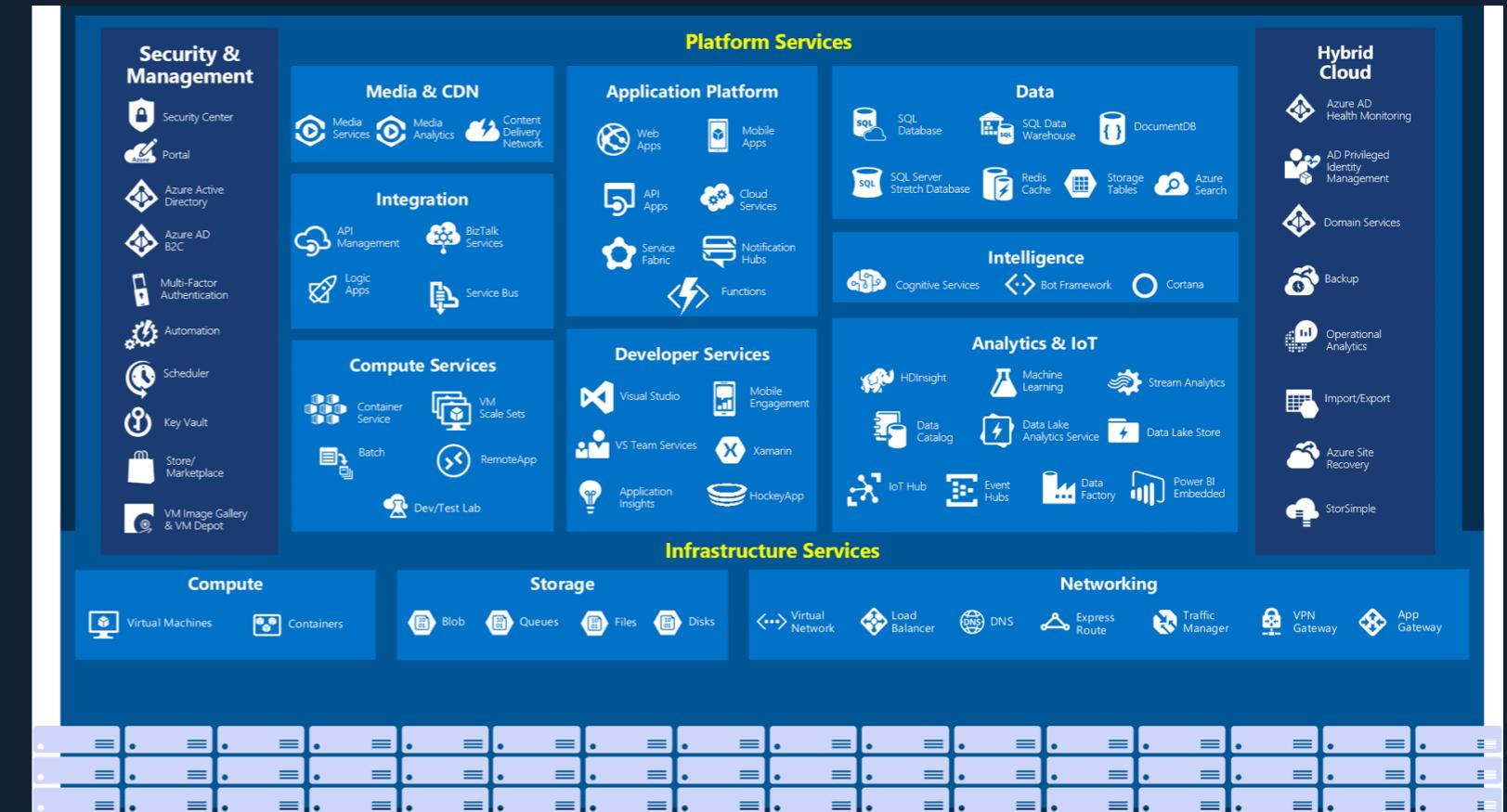


<https://www.clickittech.com/aws/why-aws-cloud/>

Microsoft Azure



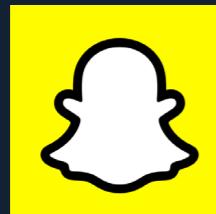
- Plataforma de computação na nuvem oferecida pela Microsoft
 - <https://azure.microsoft.com>
- Windows Azure, 2010
 - Foco em PaaS
- Microsoft Azure, 2014
 - Inclusão de IaaS e SaaS
- Usuários



Google Cloud Platform



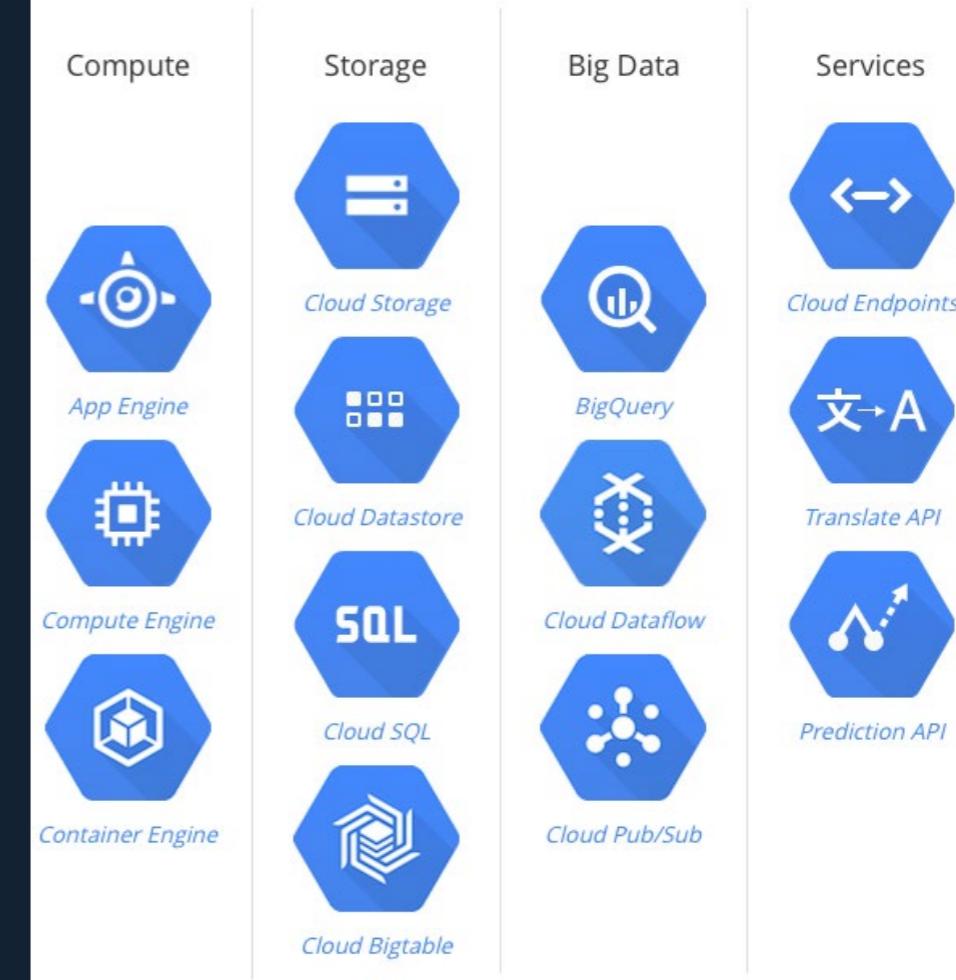
- Plataforma de computação na nuvem oferecida pela Google
 - <https://cloud.google.com>
- Google App Engine, 2008
 - Foco em PaaS
- Google Compute Engine, 2012
 - IaaS
- Usuários



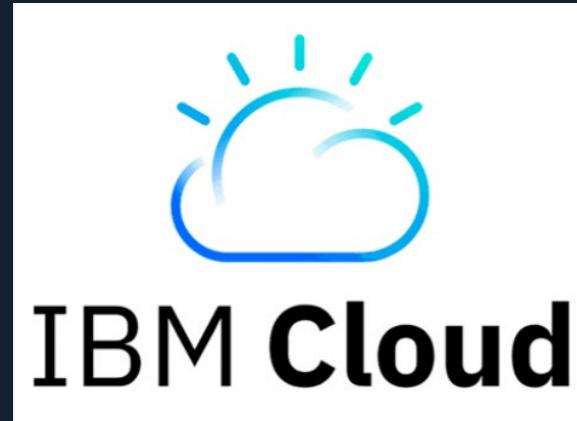
snapchat



Airbnb



Outros provedores



<https://cloud.ibm.com/>



<https://www.oracle.com/br/cloud/>



<https://www.rackspace.com/pt-br/cloud>



<https://www.alibabacloud.com/>



HUAWEI CLOUD

<https://www.huaweicloud.com/>

INFRAESTRUTURA DE NUVEM

Infraestrutura de nuvem - Data Centers



Data center em Douglas County, Georgia - Google



Saint-Ghislain, na Bélgica - Google



Data center em Clonee, Ireland - Facebook



Projeto Natick - Microsoft

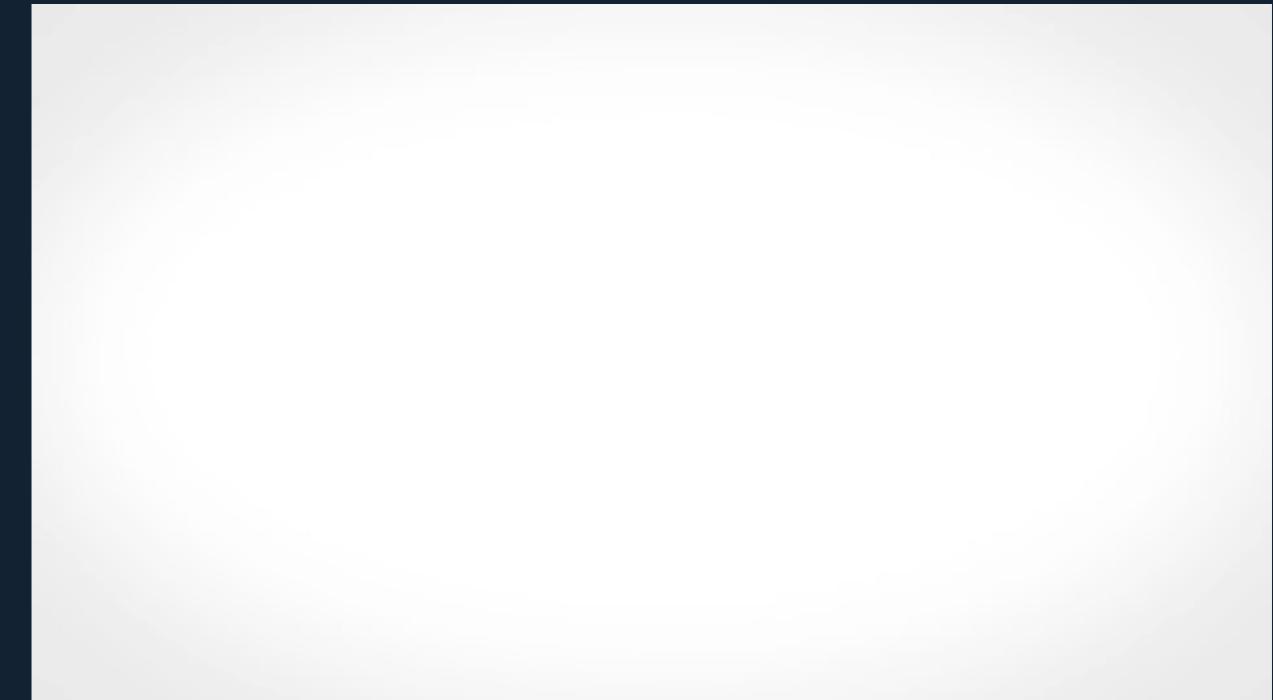


<https://www.youtube.com/watch?v=XZmGGAbHqa0>

Visitando um Data Center da Google



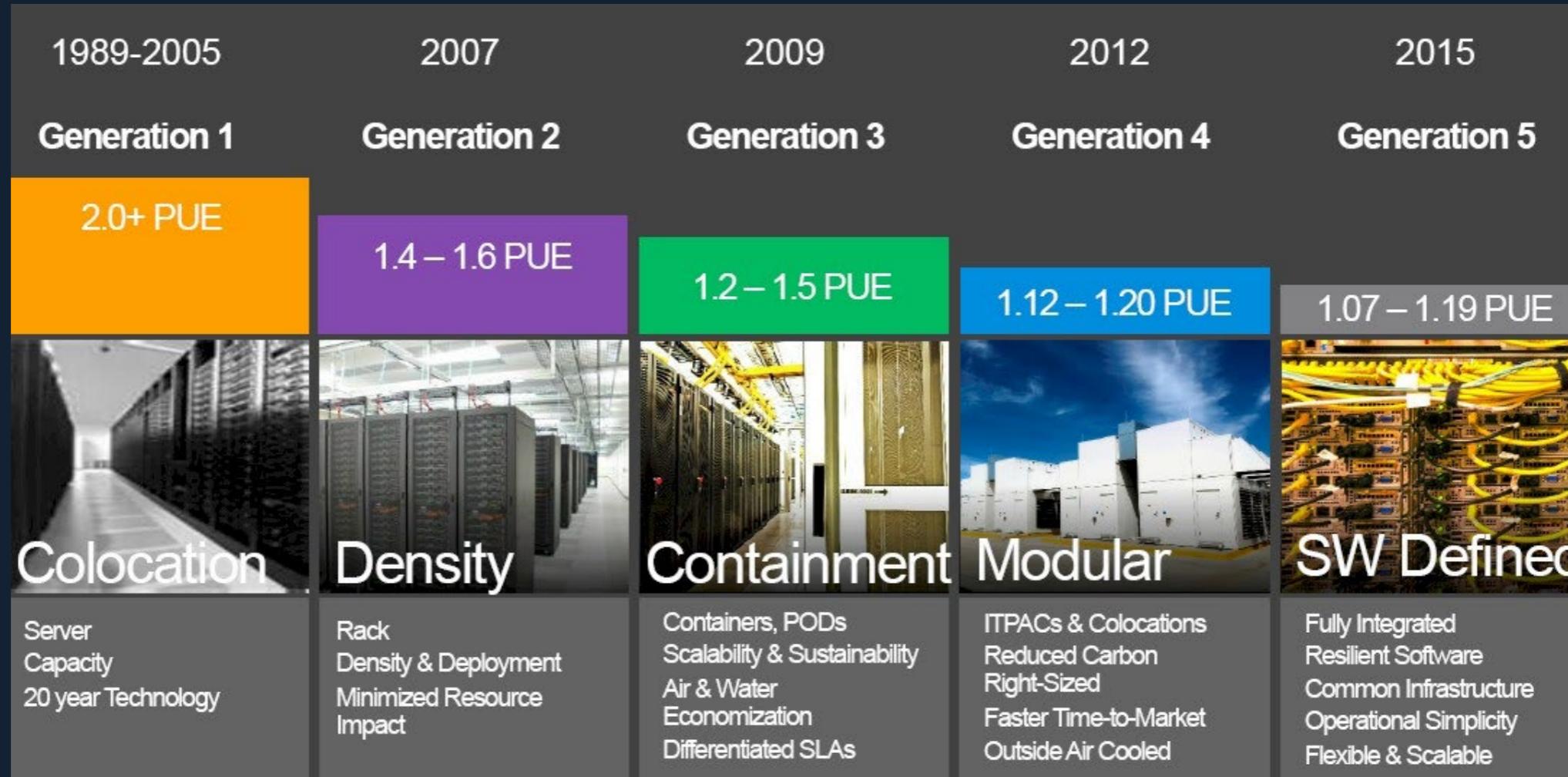
<https://goo.gl/maps/6NRN5fLVCHSQZrne9>



<https://www.youtube.com/watch?v=avP5d16wEp0>

Data Centers – Evolução

Microsoft Data Centers



<https://www.nextplatform.com/2016/09/26/rare-tour-microsofts-hyperscale-datacenters/>

AWS Global Infrastructure

- **Região**
 - Localização geográfica que agrupa um conjunto de Data Centers da Amazon
 - Data Centers são isolados fisicamente
 - Energia, refrigeração e segurança independentes
- **Zona de disponibilidade (AZ - Availability Zone)**
 - Um ou mais Data Centers com energia, rede e conectividade redundantes localizados em uma Região AWS
 - Comunicação entre AZs é com baixa latência e segura
- **Local Zones**
 - Hospelam recursos de computação, armazenamento, banco de dados e outros serviços mais próximo dos usuários finais
 - Permitem a execução de aplicações sensíveis à latência



(2022)

- 26 regiões geográficas
- 84 zonas de disponibilidade
- 17 local zones

AWS Global Infrastructure Map



SERVIÇOS DE NUVEM

Serviços de computação em nuvem

<https://console.aws.amazon.com/>

<https://azurecharts.com/overview>

AWS Management Console

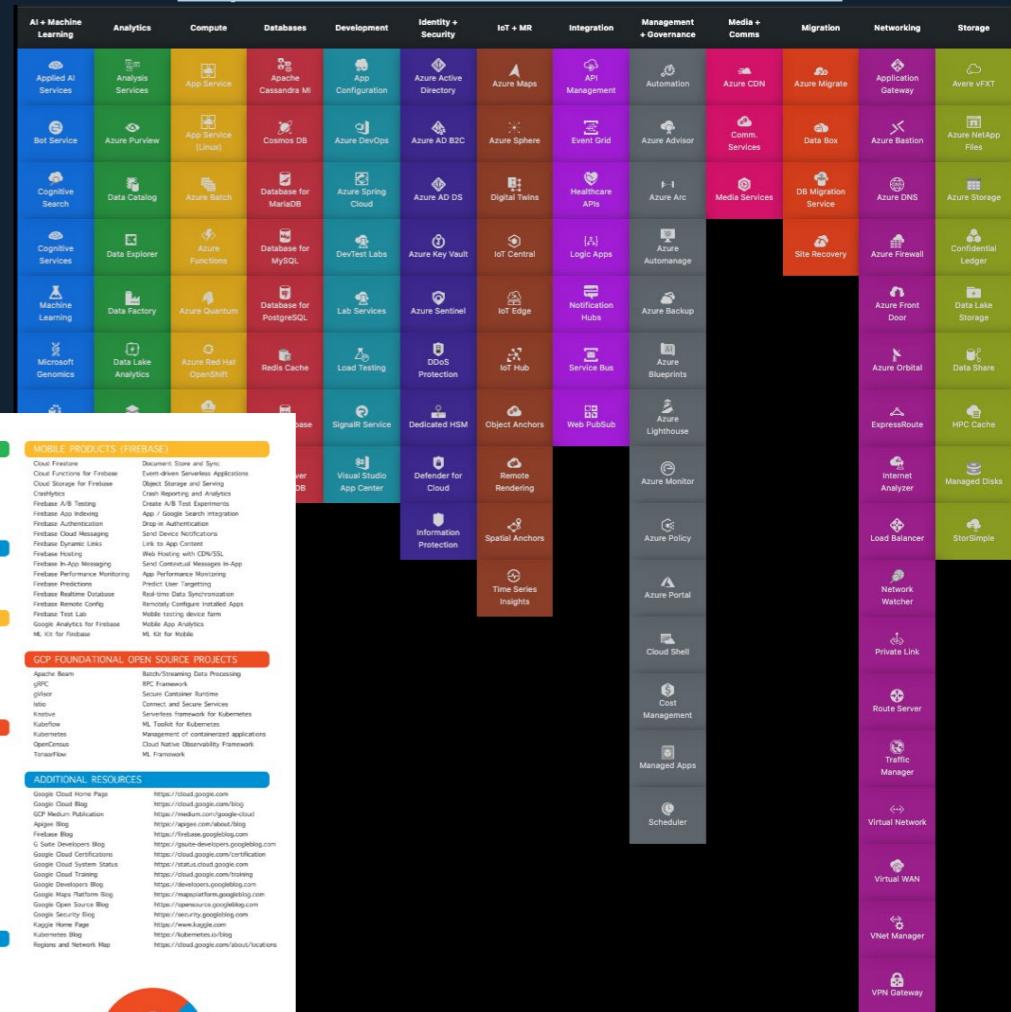
it for Vision
it for Equipment
it for Metrics

Google Cloud
DEVELOPING CHEAT SHEET

DEVELOPER'S CLOUD SHEET	
v2018.11.30	
Maintained at https://github.com/gregstevens/google-cloud-dotnet	Feedback Twitter GitHub Blog
COMPUTE PRODUCTS	
Compute Engine	Virtual Machines, Brios, Networks
Cloud Functions	Managed App Services
Kubernetes Engine	Managed Kubernetes/Containers
Cloud Functions	Event-driven serverless functions
STORAGE PRODUCTS	

Cloud Services	
Cloud Compute	Amazon Lambda
Cloud Data	Amazon Redshift
Cloud Database	Amazon Aurora
Cloud DevOps	Amazon CloudWatch Metrics
Cloud Machine Learning	Amazon SageMaker
Cloud Migration	Amazon CloudWatch Metrics
Cloud Storage	Amazon S3
Cloud Compute	Amazon Lambda
Cloud Data	Amazon Redshift
Cloud Database	Amazon Aurora
Cloud DevOps	Amazon CloudWatch Metrics
Cloud Machine Learning	Amazon SageMaker
Cloud Migration	Amazon CloudWatch Metrics
Cloud Storage	Amazon S3
DATABASE PRODUCTS	
Cloud Relational	Petabytes-scale, low-latency nonrelational
Cloud Retention	Horizontally Scalable Document
Cloud Stream	Document Management Services (Document DB)
Cloud Memcached	Managed Redis
Cloud NoSQL	Horizontally Scalable Relational DB
Cloud Spanner	Managed MySQL and PostgreSQL
DATA & ANALYTICS PRODUCTS	
Cloud Composer	Managed Workflow Orchestration Service
Cloud Dataproc	Stream/Batch data processing
Cloud Data Pipeline	Managed Apache Flink
Cloud DataSync	Virtual Data Pipeline
Cloud Dataproc	Managed Spark and Hadoop
Cloud Data Pipeline	Global Real-time Messaging
Cloud Pub/Sub	Data Ingestion Service
Cloud Data Pipeline	Collaborative Data Exploration/Combining
Google Data Studio	Google Analytics
Google Genomics	Managed Genomics Platform
AI AND MACHINE LEARNING PRODUCTS	
Cloud AutoML, Natural Language	Custom text classification models
Cloud AutoML, Computer Vision	Custom object detection, image recognition
Cloud AutoML, Vision	Custom image classification models
Cloud AutoML, Job Discovery	Job Search with ML
Cloud AutoML, Video ML Engine	Pre-trained models for deep learning
Cloud Machine Learning Engine	Managed Platform for ML
Cloud Natural Language	Text, Parsing and Analysis
Cloud Speech-to-Text	Convert Text to Speech
Cloud Text-to-Speech	Convert Audio to Text
MANAGEMENT	
Cloud API	Cloud API
Cloud Audit	Cloud Audit
Cloud Identity	Cloud Identity
Cloud Key Management	Cloud Key Management Service
Cloud Security	Cloud Security
Security Keys	Security Keys
Shadow VM	Shadow VM
Time Sync	Time Sync
MANAGER	
Cloud APM	Cloud APM
Cloud Asset	Cloud Asset
Cloud Consensus	Cloud Consensus
Cloud Dev	Cloud Dev
Cloud Shell	Cloud Shell
Stackdriver	Stackdriver
DEVELOPER	
Cloud SDK	Cloud SDK
Cloud IDE	Cloud IDE
Cloud Schema	Cloud Schema

<https://github.com/gregsramblings/google-cloud-4-words>



Principais serviços de computação em nuvem



Computação

Máquina virtual

Container

Função como serviço



Armazenamento

Baseado em objetos

Baseado em blocos

Sistema de arquivos distribuído

Banco de dados



Comunicação

Rede virtual privada

Load Balancer

CDN - Content Delivery Network

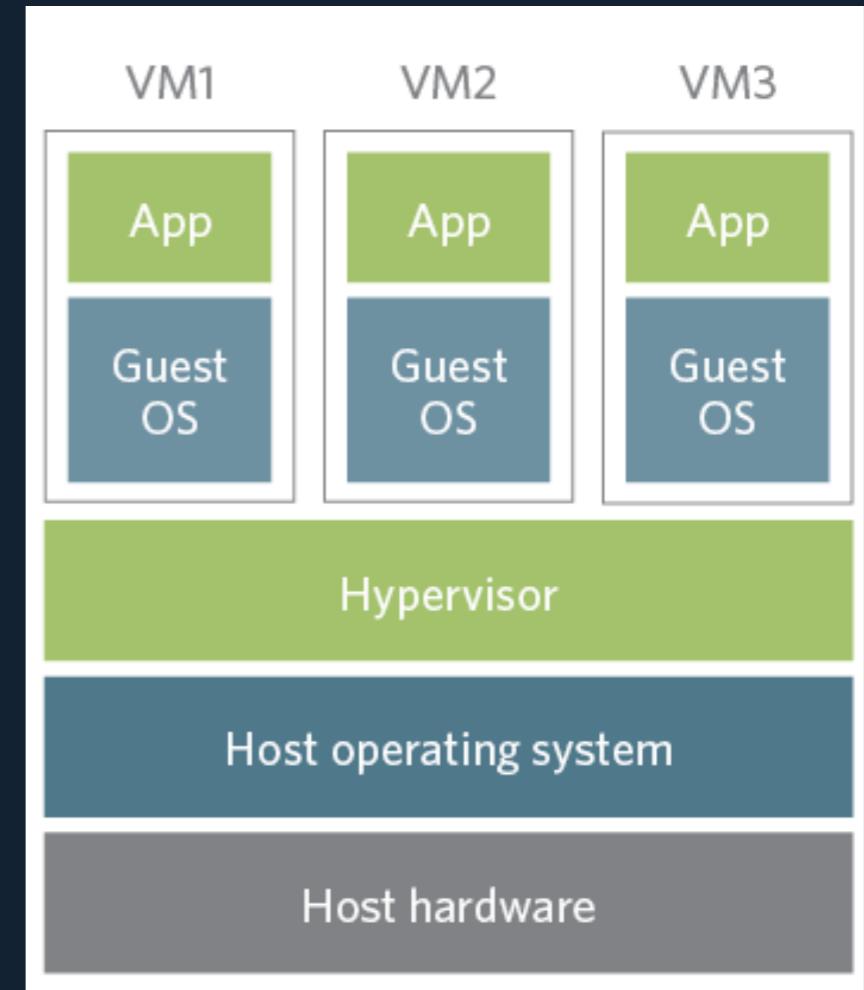


Segurança

Controle de acesso

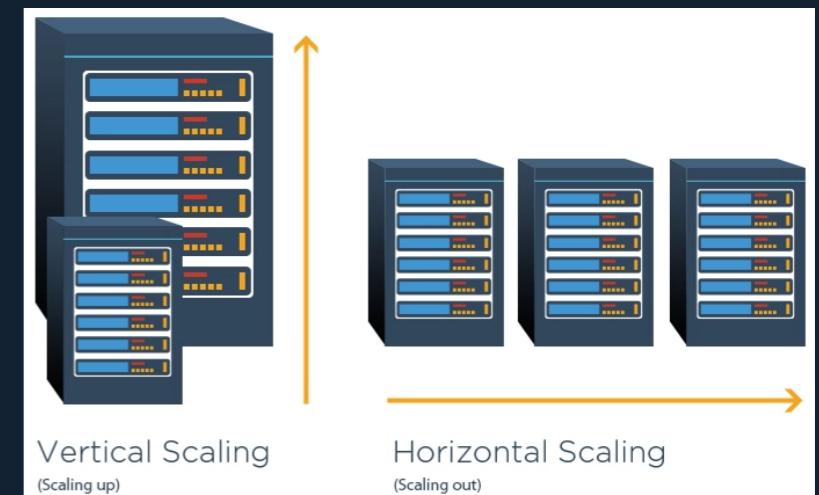
Computação / Máquina Virtual

- Uma máquina virtual (VM) é um recurso de computação similar a uma máquina física implementada em software
- Permite a execução de um sistema operacional e aplicações tradicionais
- Permite que um recurso físico (servidor) seja compartilhado por diversos usuários, cada um com sua máquina virtual, com garantias de isolamento (desempenho, falhas e segurança)
- Exemplos
 - EC2 (Elastic Compute Cloud)
 - Azure Virtual Machine
 - Google Compute Engine



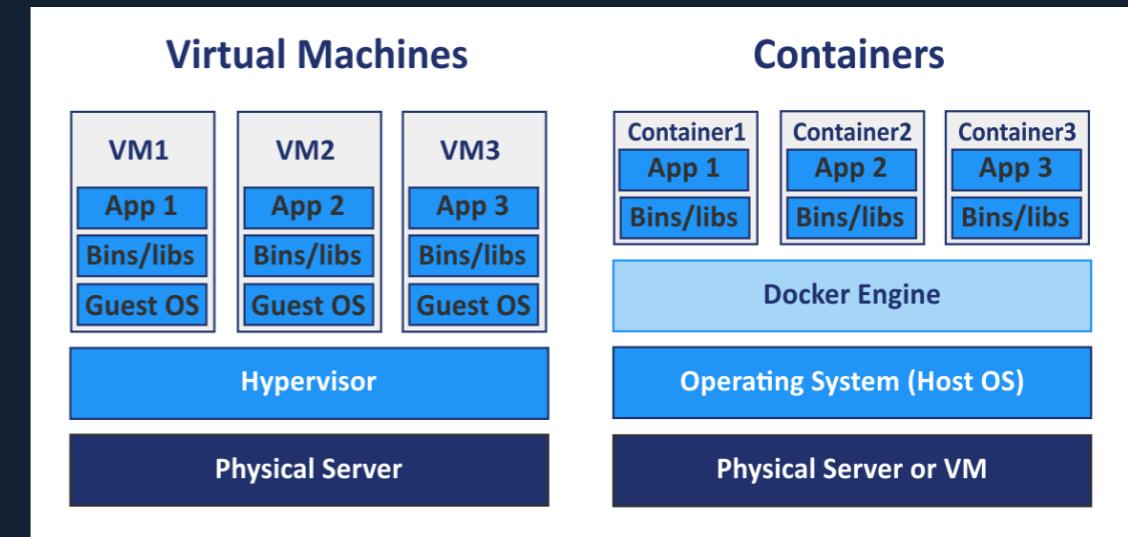
Computação / Máquina Virtual

- Criação de máquina virtual na nuvem
 - **Imagen base**: sistema operacional e aplicações
 - **Tipo de instância**: capacidade da máquina virtual (vCPUs, GPUs, memória, rede, disco)
 - Normalmente existem categorias otimizadas para diferentes cenários (web, DB, HPC, etc)
 - Tipos de instância EC2 - <https://aws.amazon.com/pt/ec2/instance-types/>
 - **Rede**: interfaces de rede e endereçamento
 - Utilização de rede privada ou alocação de endereço público (DNS)
 - **Disco**: capacidade, tipo (magnético ou SSD) e velocidade
- Escalabilidade vertical x horizontal
 - Vertical: altera o tipo de uma instância para alterar sua capacidade (requer reinicialização)
 - Horizontal: adiciona mais instâncias ao conjunto de computação (normalmente acessados via load balancer)



Computação / Container

- Provisionamento de aplicações encapsuladas em containers
 - Container inclui o executável da aplicação e dependências
- Vantagens
 - Imagens com menor tamanho (em relação a VMs), menor tempo de provisionamento, menor demanda por recursos
- Utilização do padrão Docker
 - Dockerfile – manifesto descrevendo a imagem base do container
- As imagens dos containers são armazenadas em um registro
 - Docker Hub, Amazon ECR, Azure Container Registry



<https://www.nakivo.com/blog/docker-vs-kubernetes/>

Computação / Container

Dockerfile

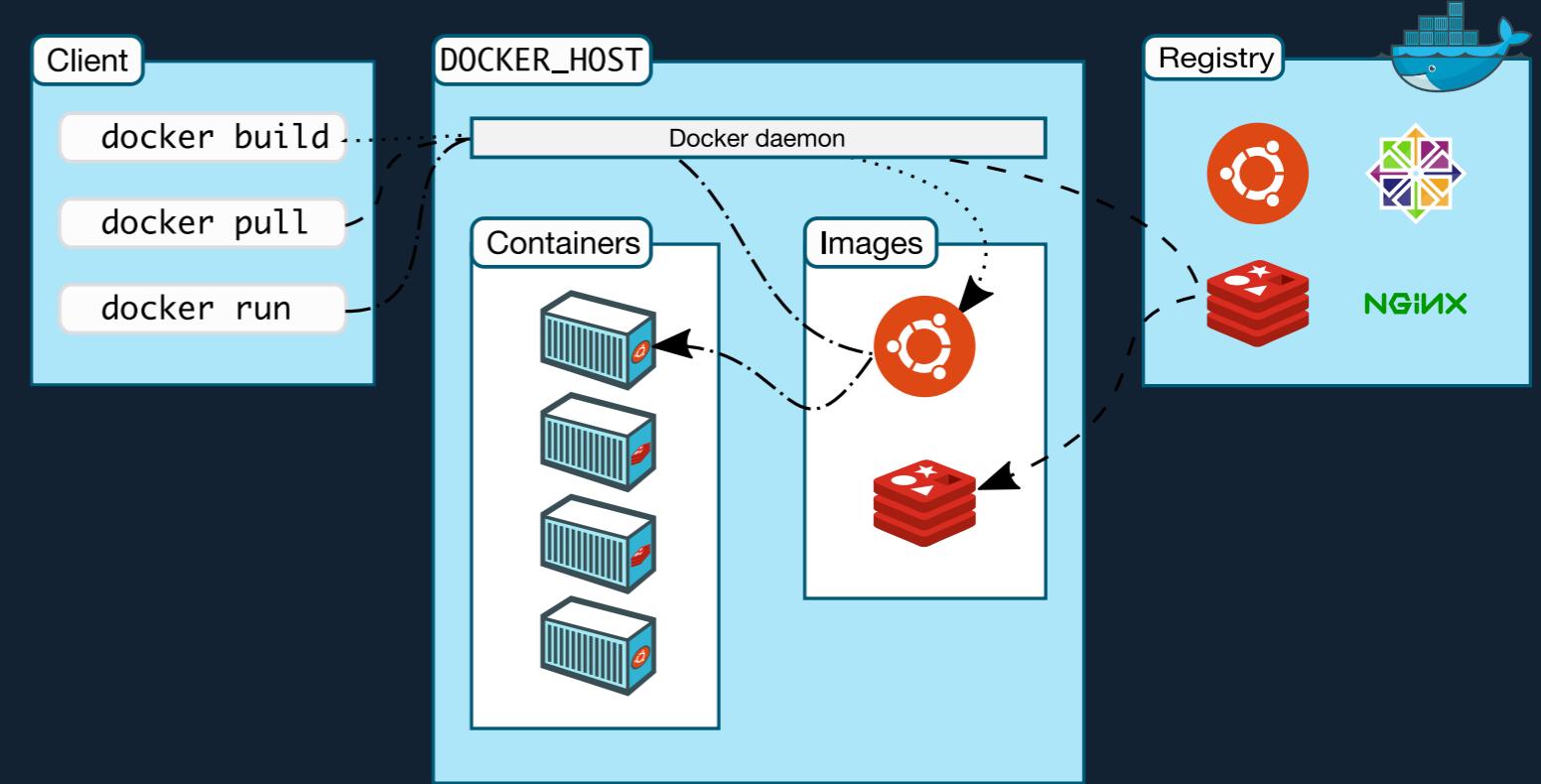
```
FROM ubuntu:16.04

RUN apt-get update
RUN apt-get install -y python3
COPY . /app

WORKDIR /app

EXPOSE 5000

CMD python3 /app/app.py
```



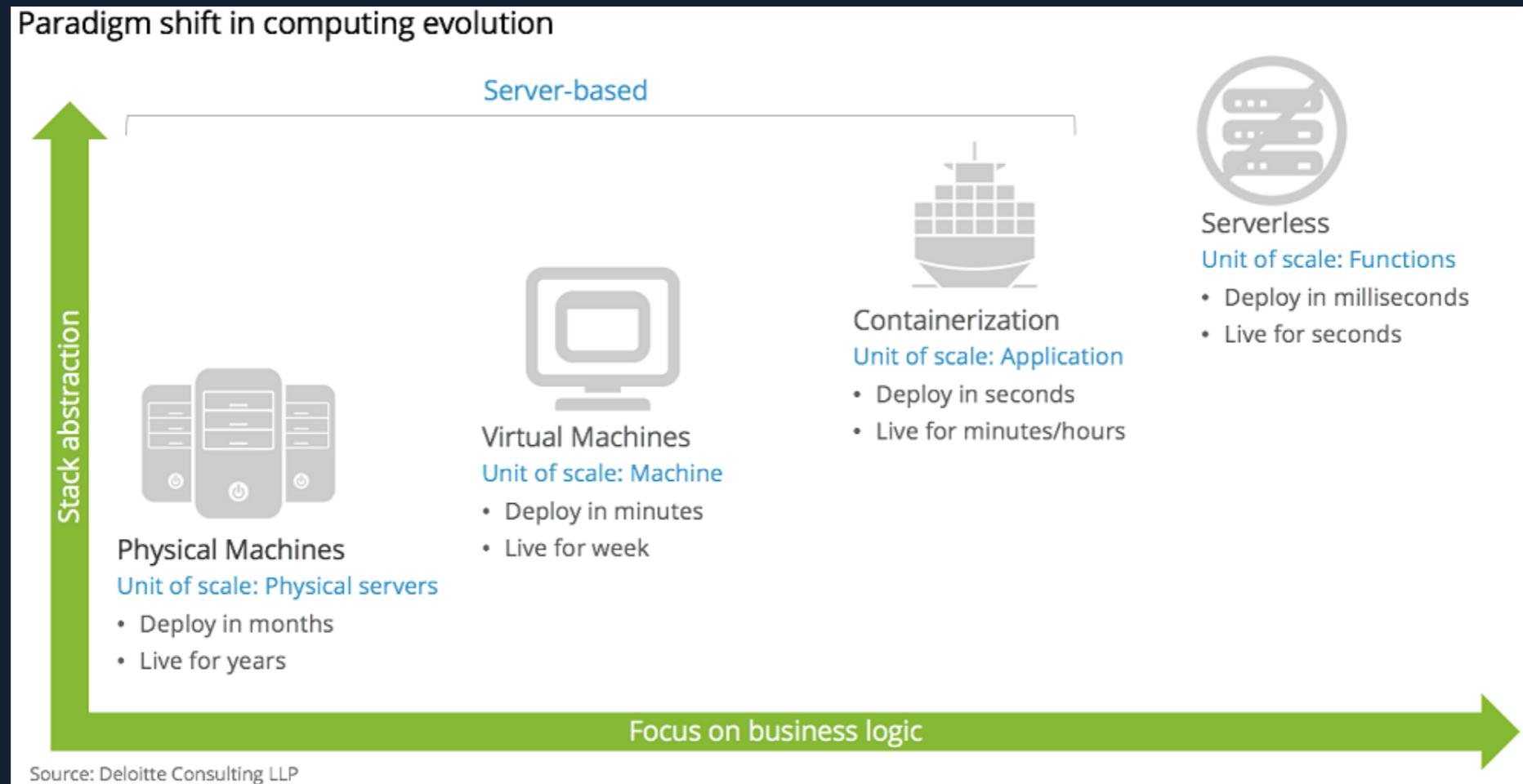
<https://docs.docker.com/get-started/overview/>

Computação / Container

- Utilizando containers em nuvem
 - Criação manual de containers dentro de VMs
 - Utilização de um orquestrador (ex. Kubernetes ou outro) em um cluster de VMs
 - Gerencia o ciclo de vida dos containers
 - Provê resiliência e escalabilidade
- Exemplos
 - ECS (Elastic Container Service), EKS (Elastic Kubernetes Service), ECR (Elastic Container Registry)
 - Azure Container Instances, AKS (Azure Kubernetes Service), Azure Container Registry
 - Kubernetes Engine, Artifact Registry

Computação / Função como serviço

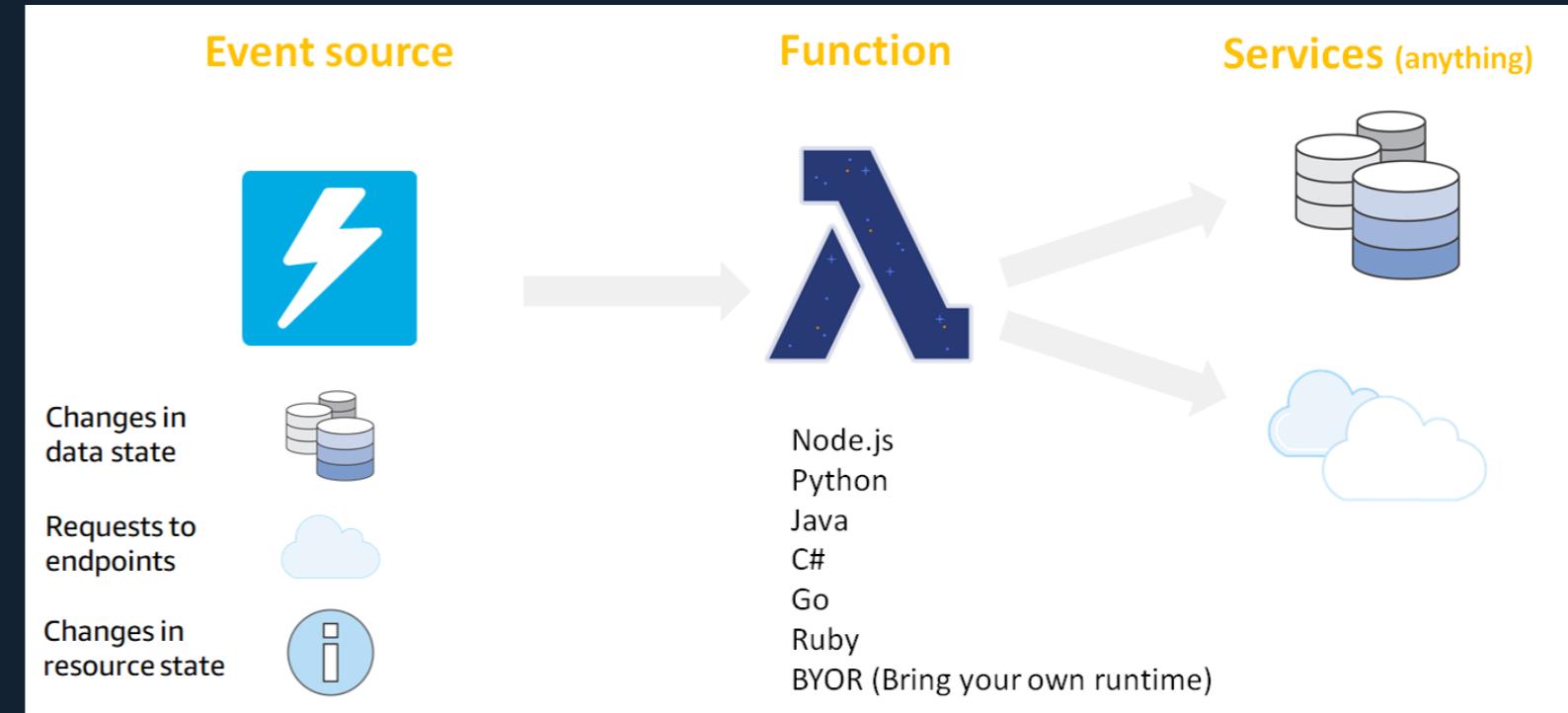
- Também chamado de serverless computing (computação sem servidor)



Computação / Função como serviço

- Evita a necessidade do gerenciamento dos servidores
- Foco é no provisionamento de código através de funções
- Características
 - Execução por tempo limitado Baseado em eventos (triggers)
 - Stateless
 - Elasticidade automática
 - Suporte a diversas linguagens

- Exemplos
 - AWS Lambda
 - Azure Functions
 - Google Cloud Functions

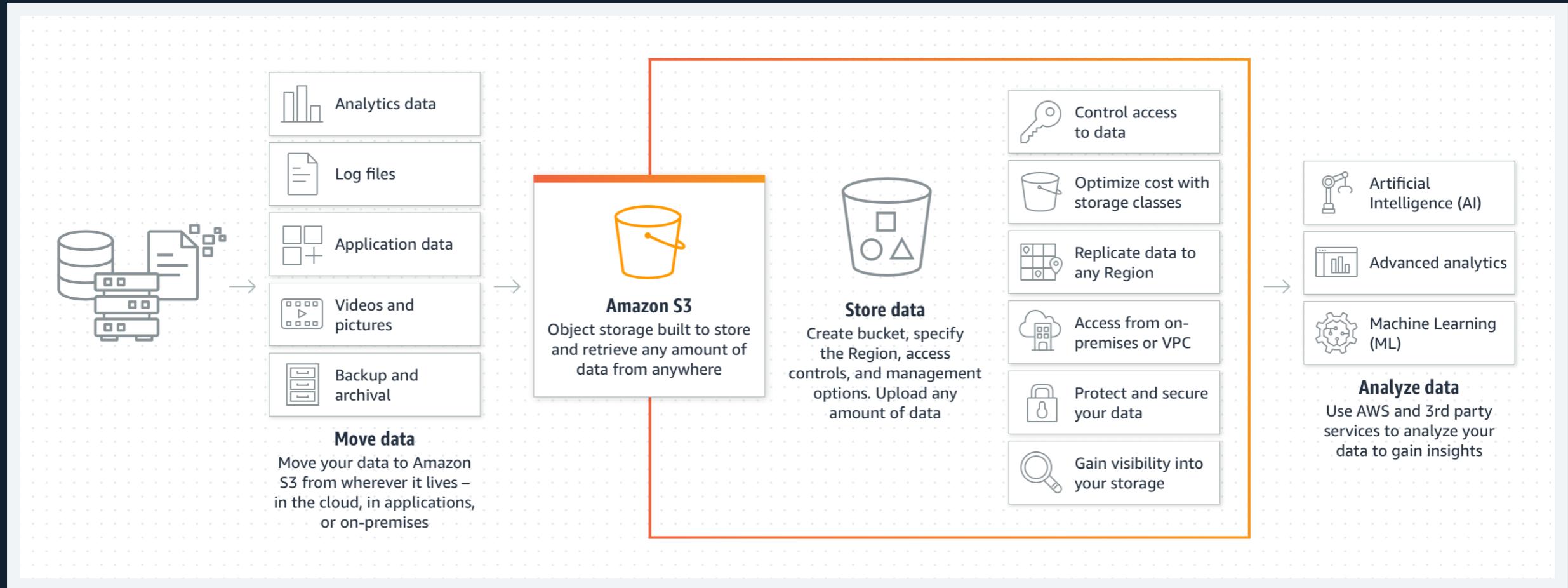


<https://aws.amazon.com/getting-started/deep-dive-serverless/>

Armazenamento / Baseado em objetos

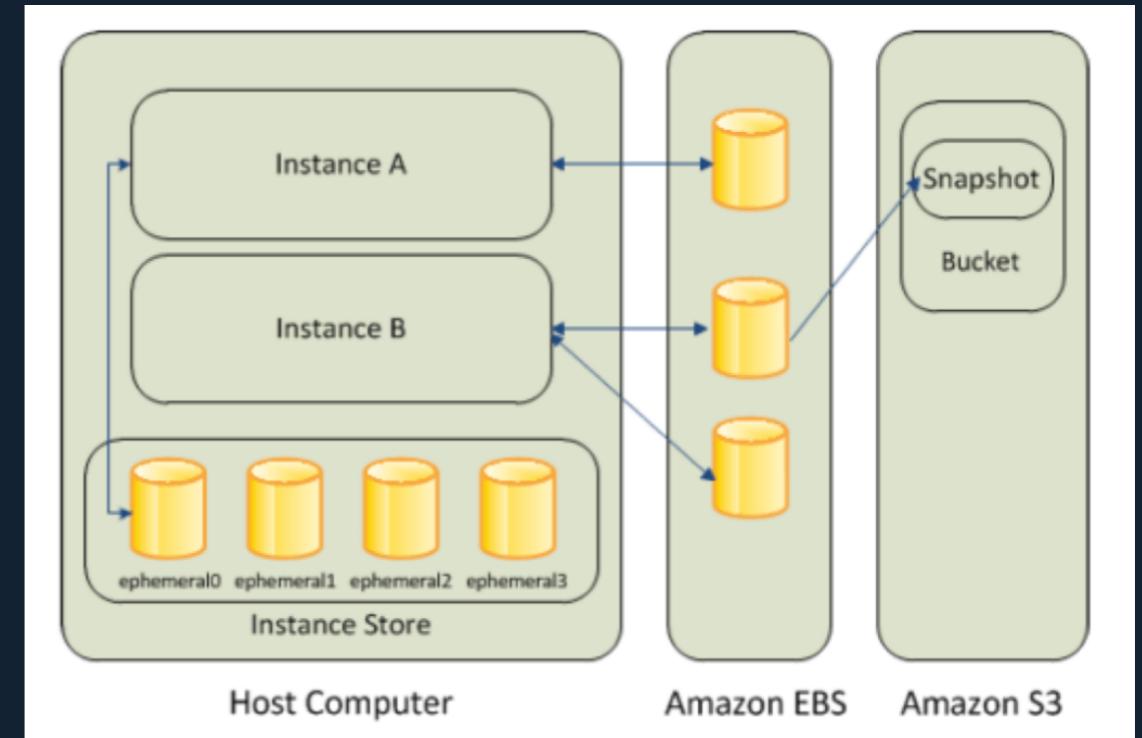
- Armazenamento de arquivos (objetos) com acesso pela Internet
 - Objetos nomeados através de URLs
- Garantem alta disponibilidade
 - Objetos são replicados em diversas zonas de disponibilidade
 - Caso alguma réplica fique indisponível, uma nova cópia é realizada automaticamente
- Diferentes níveis de permissão de acesso
 - Público, privado, uso de chaves, temporário
- Permite armazenamento de arquivos com grande tamanho
 - Por exemplo, Amazon S3 suporta arquivos com até 5TB
- Suporte a encriptação dos dados para garantir segurança
- Exemplos
 - Amazon S3 (Simple Storage Service)
 - Azure Blob Storage
 - Cloud Storage

Armazenamento / Baseado em objetos



Armazenamento / Baseado em bloco

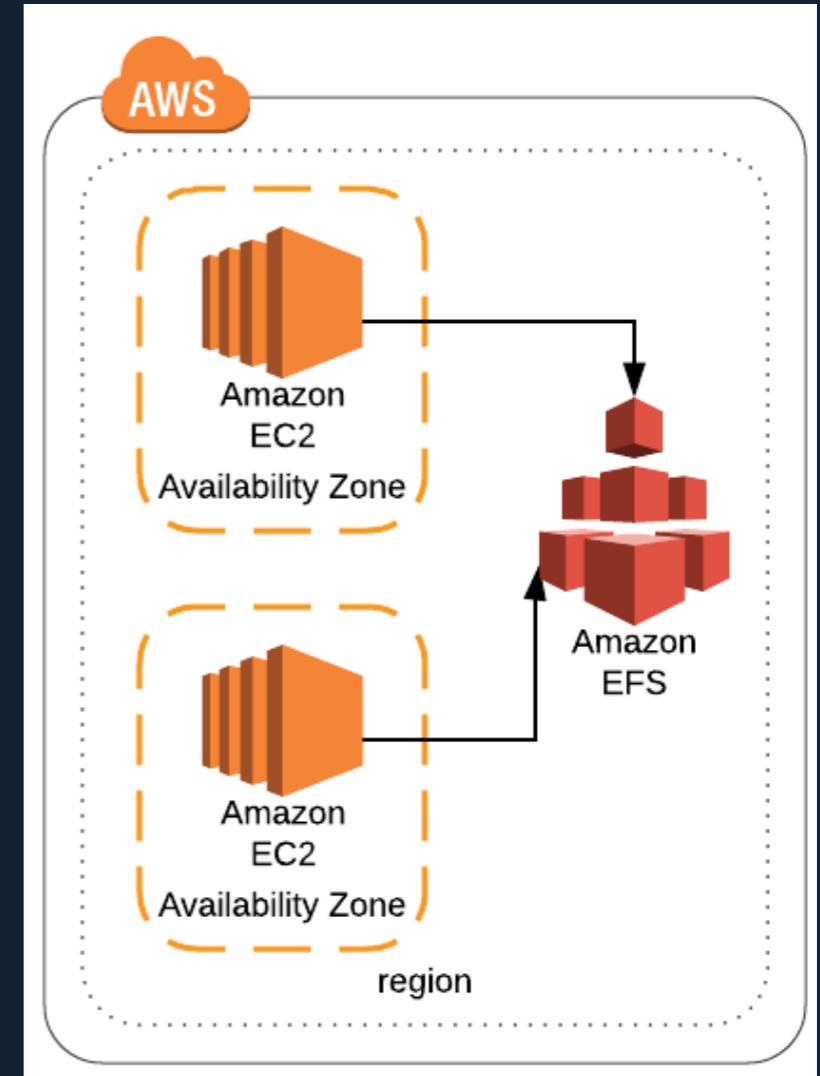
- Serviço de discos virtuais para ser utilizado pelas VMs
- Garantia de alta disponibilidade
 - Dados são replicados na zona de disponibilidade
- Suporte a snapshots
 - Cópia para somente leitura de um disco virtual
 - Permite o retorno para um estado específico do disco
- Encriptação dos dados para garantia de segurança
- Diferentes níveis de desempenho (IOPS - *Input/Output Operations Per Second*)
 - Permite a execução de aplicações com demanda de alta velocidade (bancos de dados)
- Exemplos
 - Amazon EBS (Elastic Block Storage)
 - Azure Page Blobs, Managed Disks
 - Persistent Disk



<https://blog.estabil.is/aws-na-pratica-armazenamento-ebs-e-s3/>

Armazenamento / Sistema de arquivos distribuído

- Sistema de arquivos compartilhado entre diversas instâncias (VMs, containers, funções)
- Alta escalabilidade e disponibilidade
- Pay-per-use: pagamento decorrente do uso
 - Diferente do armazenamento baseado em blocos (pagamento pela capacidade alocada)
- Uso de protocolos padrão
 - SMB e NFS
- Exemplos
 - Amazon EFS (Elastic File System)
 - Azure Files
 - File Store

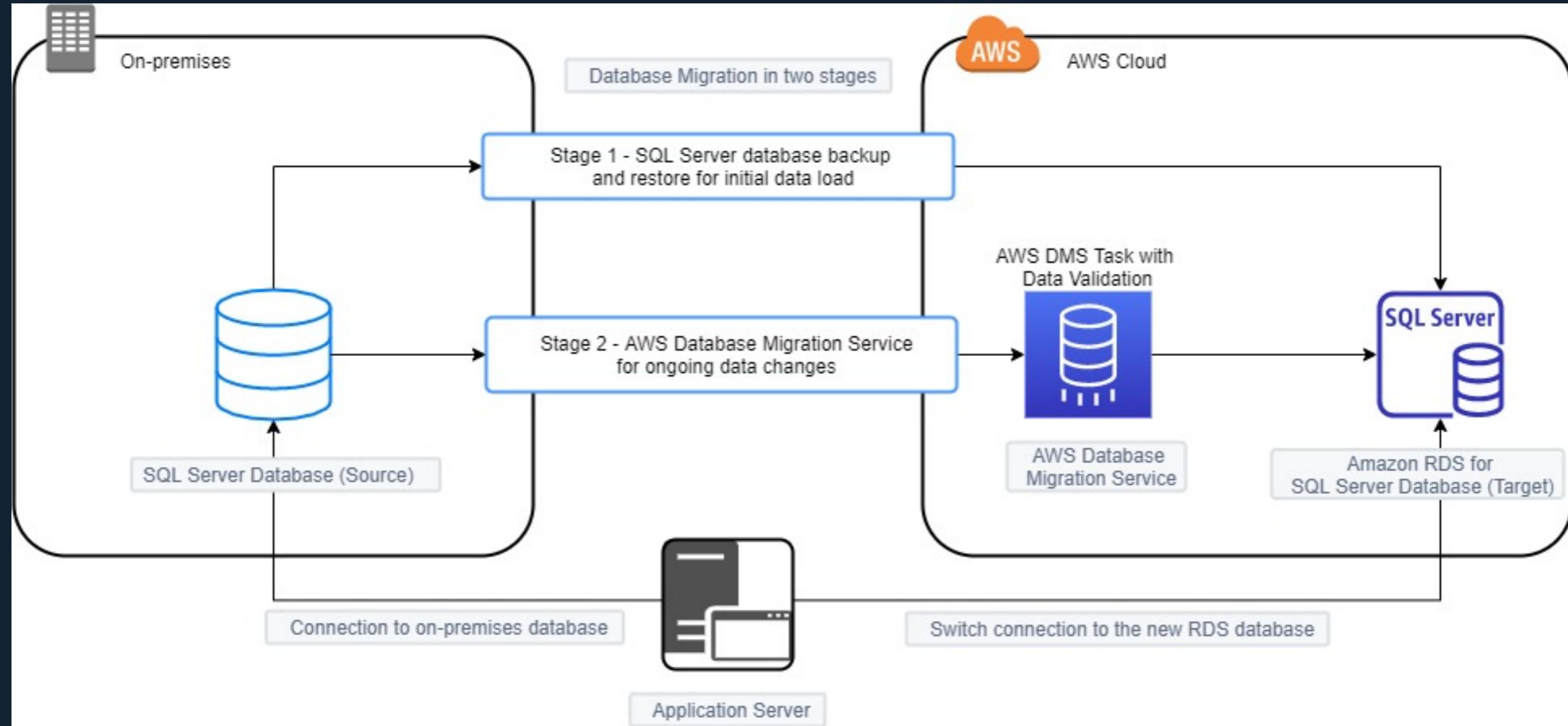


<https://medium.com/@surajjadhav525k/deployment-of-application-on-aws-using-terraform-efs-service-d3cf522242fa>

Armazenamento / Banco de dados

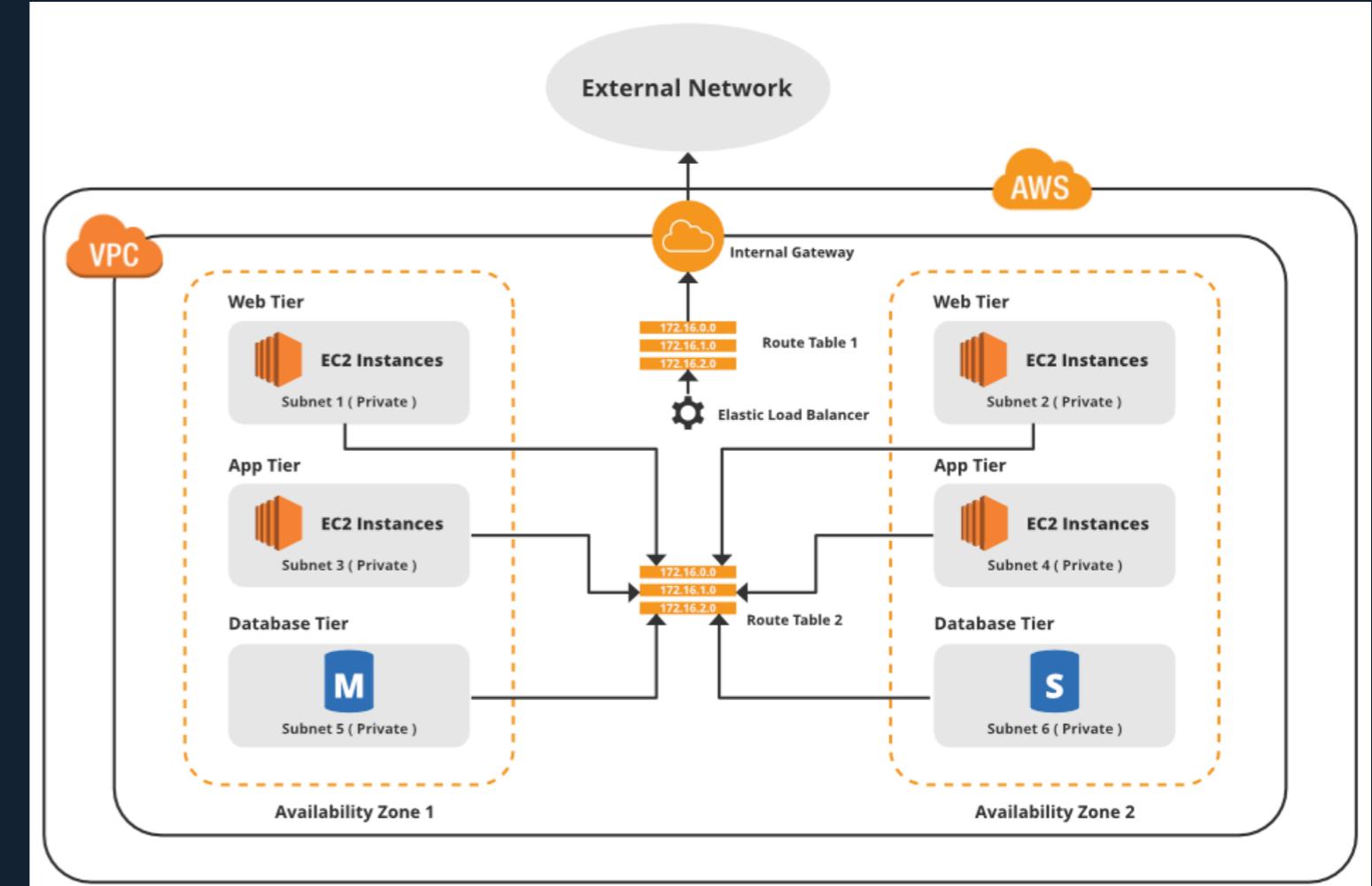
- Bancos de dados relacionais e não relacionais (NoSql) gerenciados pelo provedor de nuvem
- Provisionamento, escalabilidade, patching e backups automáticos
- Alta disponibilidade (*Disaster Recovery*)
- Desenvolvedor é responsável pela estrutura dos dados, ingestão, consultas e permissões de acesso
- Compatibilidade com engines de bancos de dados tradicionais
 - MySQL, PostgreSQL, SQL Server, Oracle
- Exemplos
 - Amazon RDS, Amazon DynamoDB
 - Azure SQL Database, Azure CosmosDB
 - Cloud SQL, Cloud Datastore

Armazenamento / Banco de dados



Comunicação / Rede virtual privada

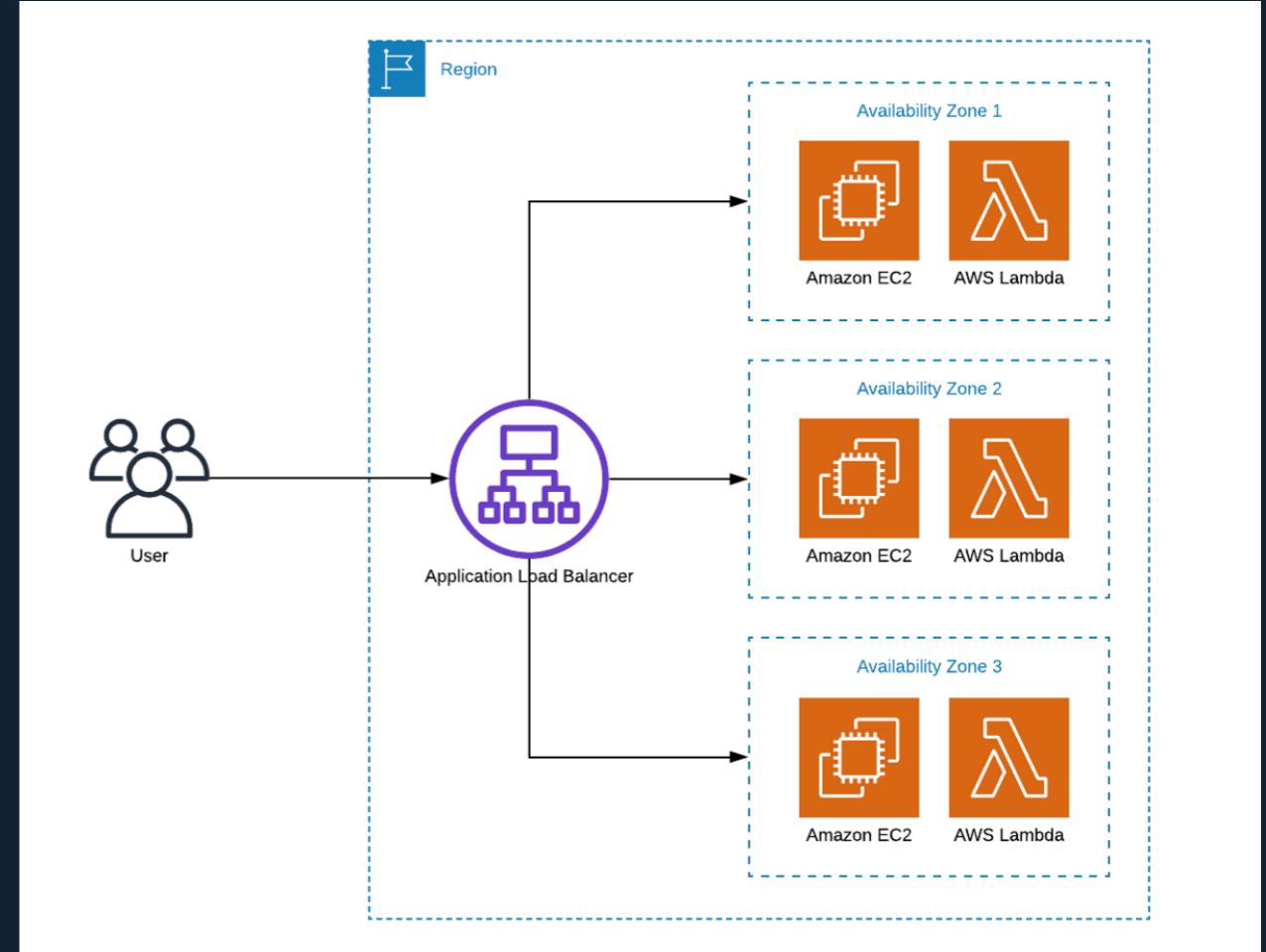
- Rede virtual isolada logicamente de outros usuários
- Provê controle e segurança
- Alta configurabilidade
 - Endereçamento, subredes, tabelas de roteamento
- Integração com appliances
 - Router, Internet Gateway, Load Balancer, VPN, NAT, Firewall
- Viabiliza a conexão com redes virtuais de outros usuários ou redes internas da organização
- Exemplos
 - Amazon VPC
 - Azure VNet
 - (GCP) Virtual Private Cloud



<https://acervolima.com/amazon-vpc-introducao-ao-amazon-virtual-cloud/>

Comunicação / Load balancer

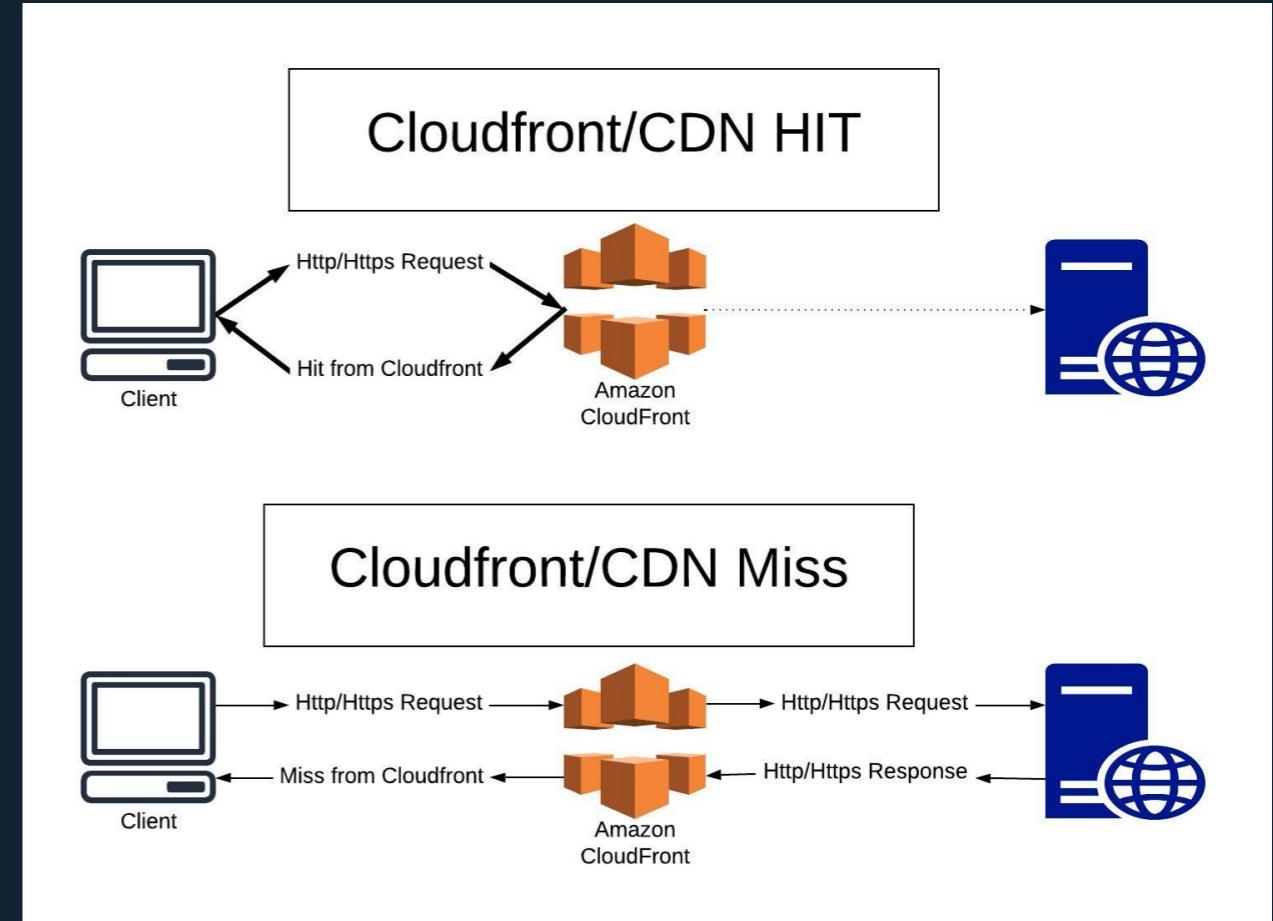
- Distribuição do tráfego de rede entre instâncias de computação
- Permite a elasticidade da arquitetura de software
 - Auto-scaling
- Suporte a sticky sessions, health checks
- Operação em diferentes camadas
 - Aplicação: HTTP, HTTPS, WebSocket
 - Transporte: TCP, TLS, UDP
 - Rede: IP
- Exemplos
 - (AWS) Elastic Load Balancing
 - Azure Load Balancer
 - Cloud Load Balancing



<https://medium.com/ankercloud-engineering/10-reasons-why-you-should-think-about-using-an-aws-application-loadbalancer-945f57816c34>

Comunicação / CDN

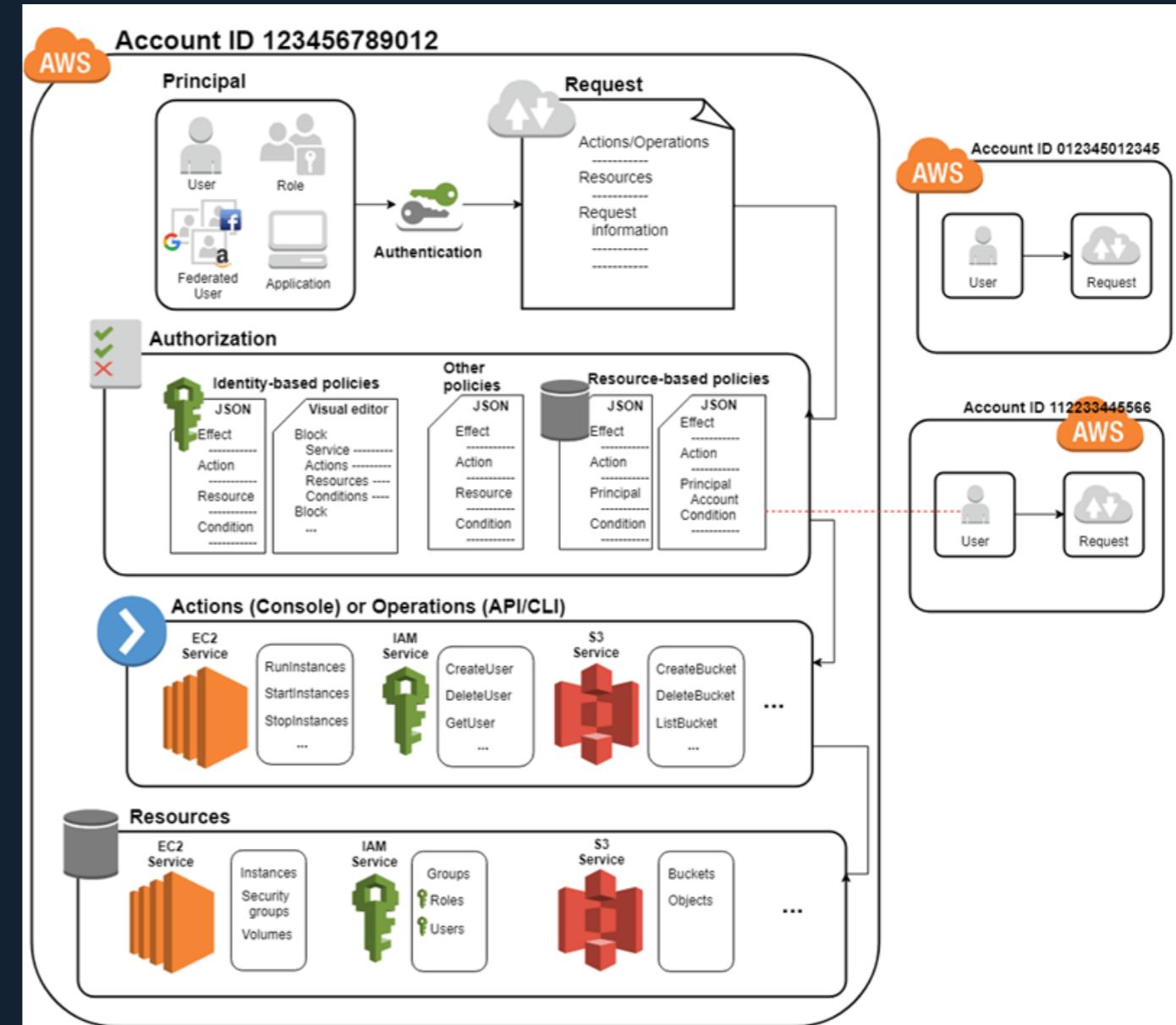
- Content Delivery Network
- Rede rápida para distribuição de conteúdo
- Utilização de edge locations
 - Cache de conteúdo
- Clientes acessam conteúdo mais próximo
 - Menor latência
- Maior segurança à ataques
 - Servidor edge recebe tráfego malicioso e protege servidor de origem
- Exemplos
 - Amazon CloudFront
 - Azure CDN
 - Cloud CDN



https://www.drupal.org/project/cloudfront_path_invalidate

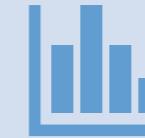
Segurança / Controle de acesso

- Controle centralizado sobre o acesso aos serviços de nuvem
- Gerência de usuários, permissões, chaves, recursos, etc
- “Least privilege principle”
- Exemplos
 - AWS Identity and Access Management (IAM)
 - Azure Active Directory
 - Cloud IAM



https://docs.aws.amazon.com/pt_br/IAM/latest/UserGuide/intro-structure.html

Serviços avançados de nuvem



Big Data Analytics



Edge Computing

Big Data Analytics

- Permite a criação de um cluster para processamento de dados de Big Data
- Utilização de plataformas tradicionais – Hadoop e Spark

Welcome to Amazon Elastic MapReduce

Amazon Elastic MapReduce (Amazon EMR) is a web service that enables businesses, researchers, data analysts, and developers to easily and cost-effectively process vast amounts of data.

You do not appear to have any clusters. Create one now:

Create cluster

How Elastic MapReduce Works

- Upload
- Create
- Monitor

Upload your data and processing application to S3.

Configure and create your cluster by specifying data inputs, outputs, cluster size, security settings, etc.

Monitor the health and progress of your cluster. Retrieve the output in S3.

Learn more Learn more Learn more

AWS – EMR

RUNNING

myhdinsightcluster
HDInsight Cluster

Settings Dashboard Secure Shell Scale Cluster Delete

Essentials

Resource group myresourcegroupsouth

Status Running

Location South Central US

Subscription name

Subscription id

URL myhdinsightcluster.azurehdinsight.net

Cluster Type hadoop

Operating system Linux

Worker Nodes Standard_D12, Small

Head Node Standard_D12

All settings →

Microsoft Azure HDInsight

Cloud Dataproc

Clusters Jobs

Create a cluster

Name example-cluster

Region global Zone us-central1-a

Master node

Contains the YARN Resource Manager, HDFS NameNode, and all job drivers

Machine type n1-standard-4 (4 vCPU, 15.0 GB ...)

Cluster mode Standard (1 master, N workers)

Primary disk size (minimum 10 GB) 500 GB

Worker nodes

Each contains a YARN NodeManager and a HDFS DataNode. The HDFS replication factor is 2.

Machine type n1-standard-4 (4 vCPU, 15.0 GB ...) 2

Primary disk size (minimum 10 GB) 500 GB Local SSDs (0-8) 0 x 375 GB

YARN cores 8 YARN memory 24.0 GB

Preemptible workers, bucket, network, version, initialization, & access options

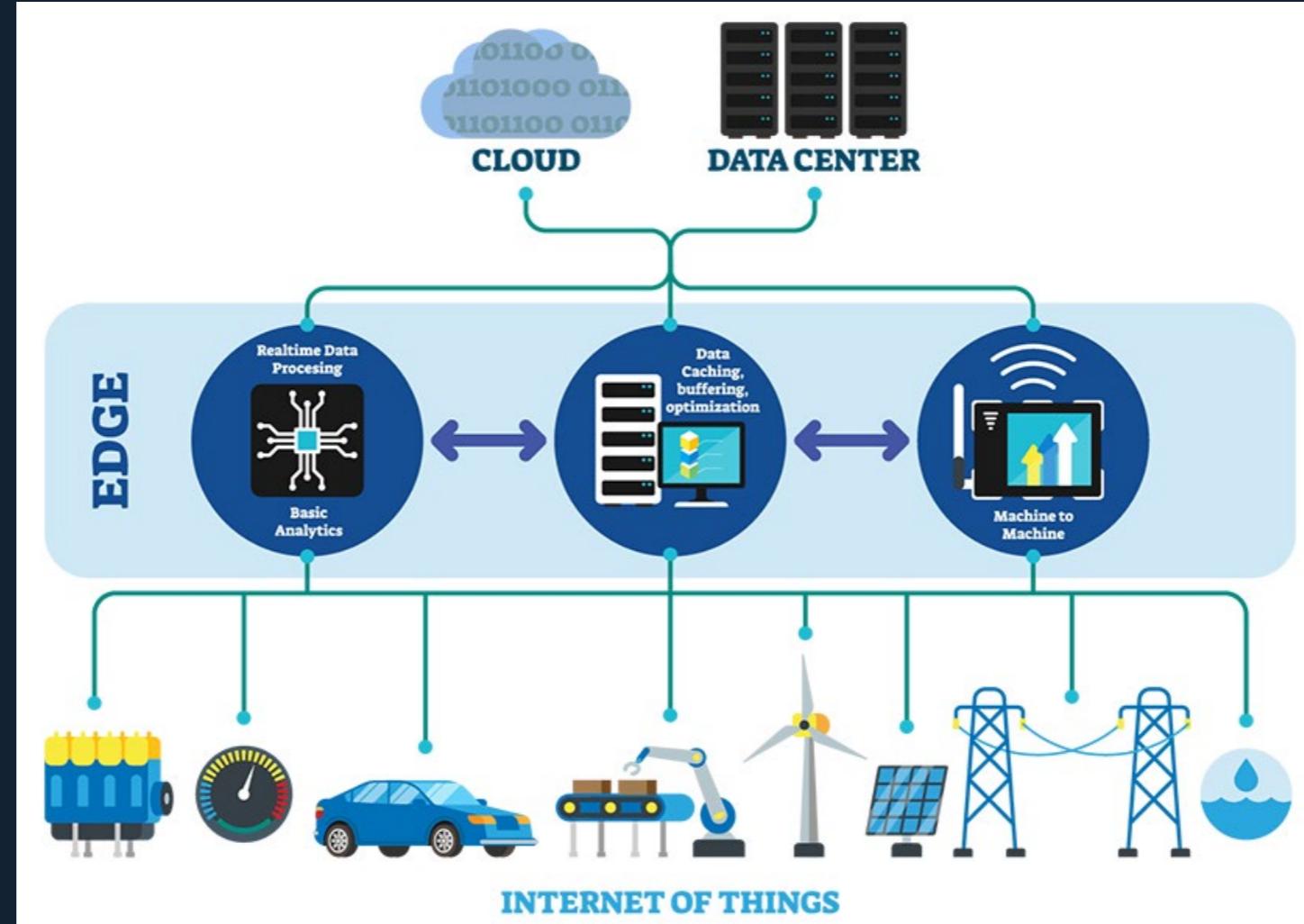
Create Cancel

Equivalent REST or command line

Google Cloud Dataproc

Edge Computing

- Método de otimização de sistemas de computação em nuvem, executando o processamento de dados na borda da rede, próximo à fonte dos dados.



Edge Computing

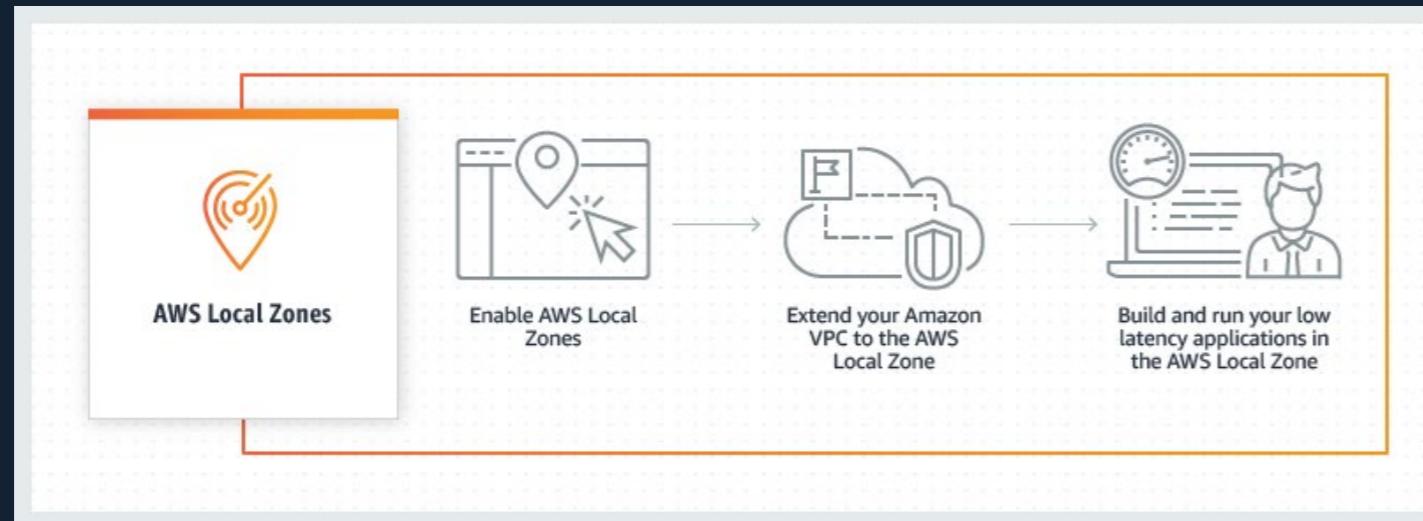
<https://aws.amazon.com/about-aws/global-infrastructure/localzones/locations/>

- AWS Local Zones

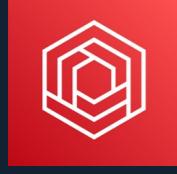
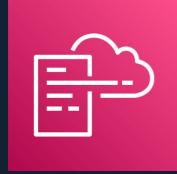
- Infraestrutura (computação, armazenamento, banco de dados, entre outros) disponível próximo aos usuários
- Menor latência para execução de aplicações com requisitos de atraso máximo (jogos em tempo real, streaming, realidade virtual/aumentada, etc)
- Permite a migração de aplicações que executam na nuvem para a edge

AWS Local Zones locations

• Generally Available
• Announced



Outros serviços – Exemplos AWS



- AWS Cloud Formation
 - Serviço de IaC (Infraestrutura como Código) para provisionamento e gerência da infraestrutura
- Amazon Managed Blockchain
 - Serviço para gerência de redes de blockchain escaláveis
- Amazon HoneyCode
 - Criação de aplicação NoCode/LowCode
- AWS CodeBuild e AWS CodePipeline
 - CI/CD (Continuous Integration/Continuous Delivery)
- Amazon WorkSpaces
 - VDI (Virtual Desktop Infrastructure)
- AWS IoT Core
 - Conexão de dispositivos IoT com a nuvem

Outros serviços – Exemplos AWS



- Amazon GameLift
 - Hospedagem de servidores de jogos com baixa latência



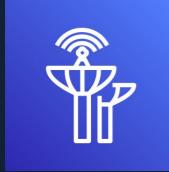
- Amazon SageMaker
 - Serviço de Machine Learning



- Amazon Elastic Transcoder
 - Transcodificação de mídia na nuvem



- Amazon Braket
 - Serviço de computação quântica



- AWS Ground Station
 - Gerência automatizada de estações terrestres para comunicação por satélites



- AWS Sumerian
 - Criação de aplicações 3D de realidade aumentada e realidade virtual

Comparação entre serviços nas nuvens públicas

Public Cloud Services Comparison

[Edit this page on Github](#) [Star 756](#) [Follow @iilyas-it83](#) [136](#) [Fork 688](#) [Issue](#)

Category	Service	AWS	Azure	Google Cloud	IBM Cloud	ORACLE CLOUD	Alibaba Cloud
Compute	Shared Web hosting		Azure shared App Services		Web hosting services		Web Hosting Simple Application Server
Compute	Virtual Server	Amazon EC2	Azure Virtual Machine	Compute Engine	Classic Virtual Server Virtual Server for VPC Power Systems Virtual Servers VMware Shared Server Instance VMware Dedicated vCenter Service Hyper Protect Virtual Server	Oracle Cloud Infrastructure Compute	Alibaba ECS
Compute	Bare Metal Server	Amazon EC2 Bare Metal Instance	Azure Bare Metal Servers (Large Instance Only for SAP Hana)		Bare Metal Servers	Oracle Bare Metal Servers	ECS Bare Metal Instance
Compute	Virtual Dedicated Host	Amazon EC2 Dedicated Hosts AWS Nitro Enclaves	Azure Dedicated Host	Sole Tenant Node (Beta)	Dedicated Virtual Servers Infrastructure (VSi)	Dedicated Virtual Machine Hosts	Dedicated Host
Compute	High Performance Computing	High Performance Computing AWS ParallelCluster Elastic Fabric Adapter NICE DCV					
Compute	Container Registration Service	Amazon Elastic Container Registry	Azure Container Registry	Container Registry	IBM Cloud Container Registry	Oracle Cloud Infrastructure Registry	Container Registry

<https://comparecloud.in/>

SOLUÇÕES, ARQUITETURAS E CUSTO NA NUVEM

Exemplos de soluções/arquiteturas

- AWS Reference Architecture Diagrams
 - <https://aws.amazon.com/pt/architecture/reference-architecture-diagrams>
 - Exemplo: [Asynchronous Online Gaming](#)
- Centro de Arquitetura do Azure
 - <https://docs.microsoft.com/pt-br/azure/architecture/>
 - Exemplo: [Renderização de vídeo 3D](#)
- Centro de arquitetura do Cloud
 - <https://cloud.google.com/architecture>
 - Exemplo: [Implantar uma rede hub and spoke usando o peering de rede VPC](#)

Quanto custa a nuvem?

<https://calculator.aws>

The screenshot shows the AWS Pricing Calculator interface. At the top, it displays a summary of the estimate: "First 12 months total" is 7,956.48 USD, with "Up front" being 0.00 USD and "Monthly" being 663.04 USD. Below this, there's a section for "West Coast Servers" (Region: US West (Oregon)) with options to "Add service", "Edit Region", and "Action". The main table lists resources and their costs:

	Up front	0.00 USD	Monthly
Amazon EC2			
1 t3.xlarge Linux instance with a consistent workload	Up front	0.00 USD	Monthly
Amazon EBS			
30 GB General Purpose SSD (gp2)	Up front	0.00 USD	Monthly
Amazon EC2			
5-10 t3.xlarge Linux instances with a daily workload	Up front	0.00 USD	Monthly
Amazon EBS			
30 GB General Purpose SSD (gp2) with 2x daily snapshots	Up front	0.00 USD	Monthly
Data Transfer			
Outbound: 100 GB	Up front	0.00 USD	Monthly
Group total	Up front	0.00 USD	Monthly

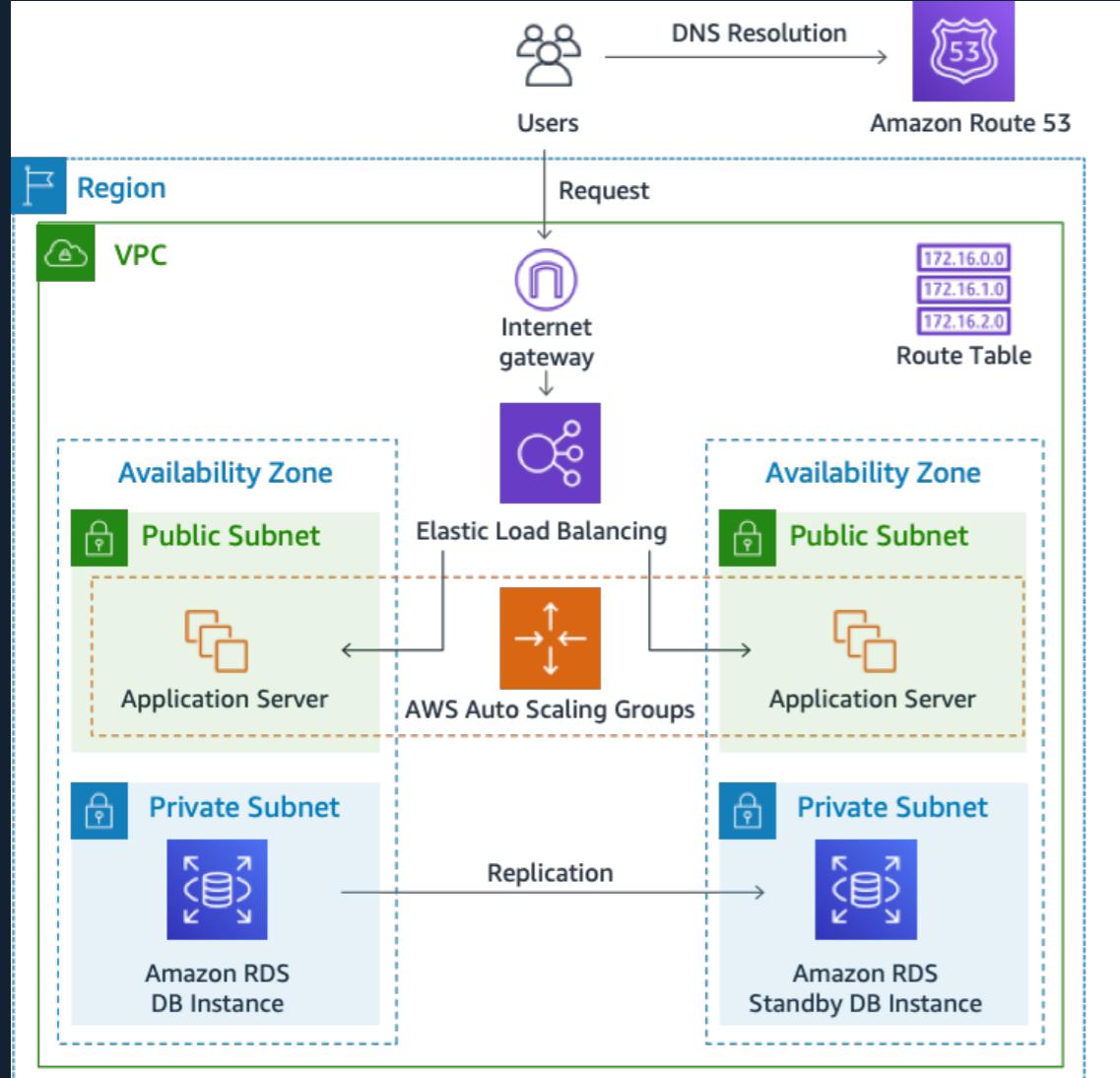
<https://cloud.google.com/products/calculator>

The screenshot shows the Google Cloud Pricing Calculator interface. At the top, it has a navigation bar with icons for COMPUTE ENGINE, GKE STANDARD, GKE AUTOPILOT, BACKUP FOR GKE, CLOUD TPU, CLOUD RUN, ANTHOS, VERTEX AI TRAINING, and VPC PRE. Below this is a search bar with the placeholder "Search for a product you are interested in.". The main area is titled "Pricing calculator" and contains a digital clock-like display showing "01134". It says "Calculate your estimated hourly or monthly costs for using Azure." Below this is a section for "Azure customers" with a "Sign in" link. The interface includes dropdown menus for "Number of Instances", "These instances for?", "System / Software", "Model", "Billing purpose", "Processor", "Core", "Per core", and "Persistent disk". A "Products" tab is selected, showing a grid of services:

- Virtual Machines
- Storage Accounts
- Azure SQL Database
- App Service
- Azure Cosmos DB
- Azure Kubernetes Service (AKS)
- Azure Functions
- Azure Cognitive Services
- Microsoft Cost Management

<https://azure.microsoft.com/en-us/pricing/calculator>

Exemplo – arquitetura para website dinâmico na AWS



- **VPC (Virtual Private Cloud)**
 - Cria uma rede lógica isolada para comunicação entre instâncias
 - O Internet Gateway habilita a comunicação entre o VPC e a Internet
- **AWS Auto Scaling Group**
 - Aumenta ou diminui instâncias EC2 para atender variações de demanda
- **Elastic Load Balancing**
 - Distribui o tráfego entre as instâncias EC2 localizadas nas AZs
- **Amazon RDS**
 - Banco de dados relacional provisionado em AZs distintas com replicação de dados em modo primário/secundário com failover automático
- **Amazon Route 53**
 - Serviço de DNS (Domain Name Service) para registro do domínio do website

Estimativa de recursos necessários

Característica	Uso Estimado	Descrição
Utilização	100%	Componentes da infraestrutura executam 24 horas por dia, 7 dias por semana
Instância	t3a.xlarge	16 GB memória, 4 vCPU
Número de instâncias	4	Na média por dia, existem 4 instâncias em execução.
Armazenamento	Amazon EBS SSD gp2	1 volume EBS por instância com 30 GB de armazenamento por volume
Backup da Dados	Snapshots EBS Diários	1 volume EBS por instância com 30 GB de armazenamento por volume
Transferência de Dados	Data in: 1 TB/mês Data out: 1 TB/mês	10% de incremento por dia
Balanceamento de Carga	20 GB/Hora	Elastic Load Balancing é usado 24 horas por dia, 7 dias por semana. Processa um total de 20 GB/Hora (data in + data out)
Banco de Dados	MySQL, instância db.m5.large com 8 GB memória, 2 vCPUs, 100 GB armazenamento	Provisionamento Multi-AZ com réplica standby sincronizada em uma AZ separada

Estimativa de custo (US East Region)

Serviço	Mensal	Anual	Configuração
Elastic Load Balancing	\$87.60	\$1051.20	Número de Network Load Balancers (1), Bytes processados por NLB para TCP (20 GB por hora)
Amazon EC2	\$439.16	\$5269.92	Sistema Operacional (Linux), Quantidade (4), Armazenamento para cada instância EC2 (General Purpose SSD (gp2)), Quantidade de armazenamento (30 GB), Tipo de instância (t3a.xlarge)
Amazon Elastic IP address	\$0	\$0	Número instâncias EC2 (1), Número EIPs por instância (1)
Amazon RDS for MySQL	\$272.66	\$ 3271.92	Quantidade (1) db.m5.large, Armazenamento para cada instância RDS (General Purpose SSD [gp2]), Quantidade de armazenamento (100 GB)
Amazon Route 53	\$183.00	\$2,196.00	Hosted Zones (1), Número de Elastic Network Interfaces (2)
Amazon Virtual Private Cloud (Amazon VPC)	\$92.07	\$1,104.84	Custo de transferência de dados, Inbound (da Internet) 1 TB por mês Outbound (para Internet) 1 TB por mês

Análise de custos

- Os preços com nuvem dependem de vários fatores
 - Especificação e número de servidores a serem provisionados
 - Volume de dados para armazenamento e transferência
 - Sistema operacional utilizado pelas instâncias provisionadas
 - Pagamento por minute, hora ou mensal
 - Opções de pay-as-you-go, instâncias reservadas ou contrato de longo prazo
 - Localização do data center de nuvem escolhido

<https://kinsta.com/blog/cloud-market-share/>

- Cloudorado - <https://www.cloudorado.com/>
 - Ferramenta de comparação de provedores de nuvem

Nível gratuito da AWS (Free Tier)

Detalhes do nível gratuito

Filtrar por:

[Limpar todos os filtros](#)

▼ Tipo de nível

- Em destaque
- 12 meses gratuitos
- Sempre gratuito
- Testes

▼ Categorias de produtos

- Análise
- Integração de aplicações
- Produtividade empresarial
- Computação
- Contêineres
- Envolvimento de clientes
- Banco de dados
- Ferramentas do desenvolvedor
- Computação de usuário final
- Web e plataforma móvel front-end
- Game Tech
- Internet das coisas
- Machine learning
- Gerenciamento e governança
- Serviços de mídia
- Migração e transferência
- Redes e entrega de conteúdo
- Robótica
- Segurança, identidade e conformidade
- Sem servidor
- Armazenamento

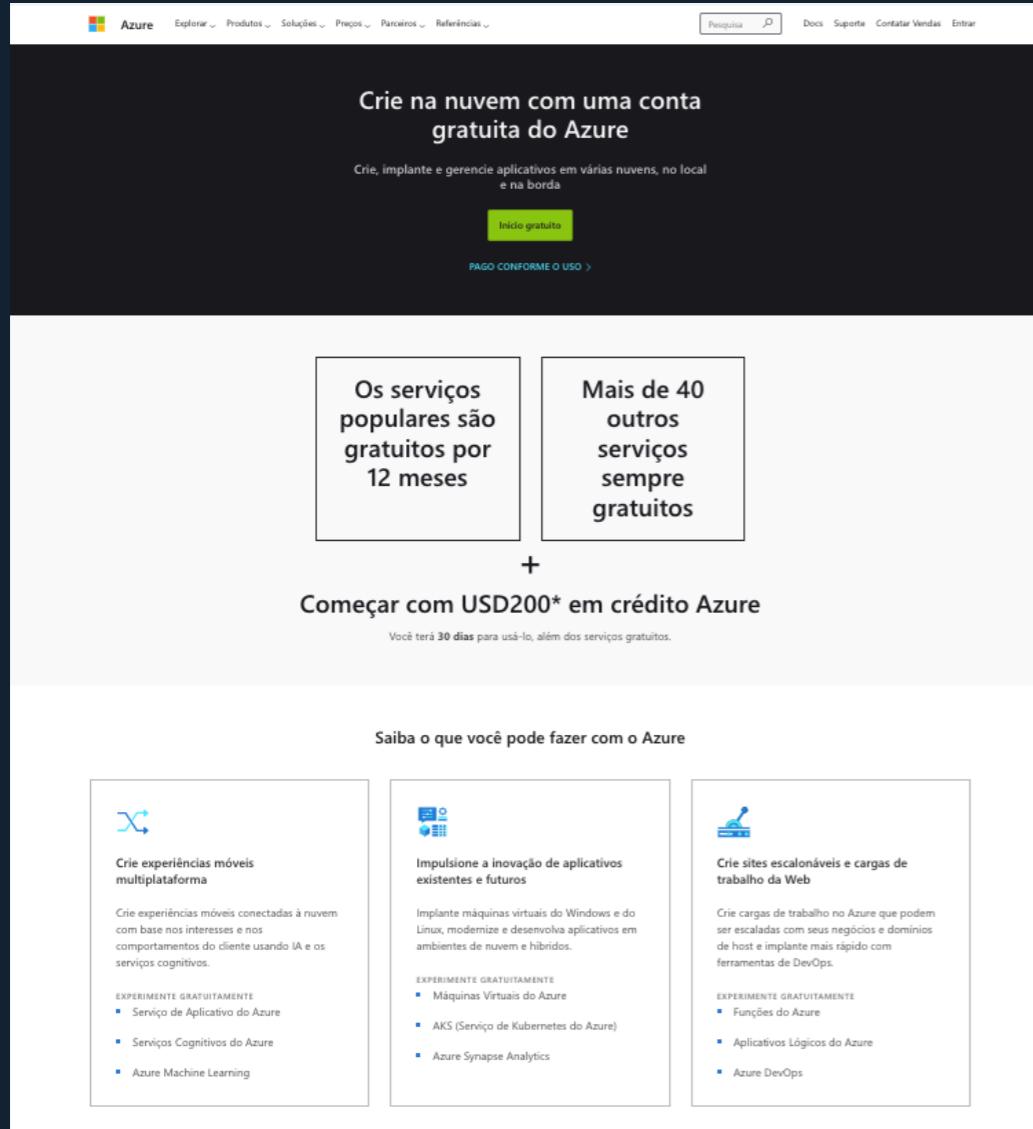
The screenshot shows the AWS Free Tier landing page with a search bar at the top. Below it, there are six service cards arranged in two rows of three. Each card includes the service name, its free tier, and a summary.

CATEGORIA	SERVICE	TIPO DE NÍVEL	DETALHES
COMPUTAÇÃO	Amazon EC2	Nível gratuito	12 MESES GRATUITOS 750 horas por mês de uso da instância t2.micro ou similar.
ARMAZENAMENTO	Amazon S3	Nível gratuito	12 MESES GRATUITOS 5 GB de armazenamento padrão
BANCO DE DADOS	Amazon RDS	Nível gratuito	12 MESES GRATUITOS 750 horas por mês de uso do banco de dados db.t2.micro (mecanismos de banco de dados). Serviço de banco de dados relacional gerenciado para MySQL, PostgreSQL, MariaDB, Oracle BYOL ou SQL Server.
BANCO DE DADOS	Amazon DynamoDB	Nível gratuito	SEMPRE GRATUITO 25 GB de armazenamento
MACHINE LEARNING	Amazon SageMaker	Nível gratuito	NOVO AVALIAÇÃO GRATUITA 2 meses de avaliação gratuita
COMPUTAÇÃO	AWS Lambda	Nível gratuito	SEMPRE GRATUITO 1 milhão de solicitações gratuitas por mês

<https://aws.amazon.com/pt/free>

- Permite o uso de alguns recursos da AWS de forma gratuita ao criar uma nova conta
- Tipos de ofertas
 - 12-month Free Tier
 - Always Free
 - Short term trials
- Ao terminar o uso gratuito, é realizada a cobrança no modelo pay-as-you-go

Azure – conta gratuita



- US\$ 200 em créditos ao criar uma conta
 - Válido por 30 dias
- Serviços populares gratuitos por 12 meses
 - VMs Linux e Windows – B1s (1 core e 1 GB de memória)
 - Banco de dados SQL – 250GB
 - ...
- Outros 40 serviços sempre gratuitos

<https://azure.microsoft.com/pt-br/free/>

GCP – Teste gratuito

The screenshot shows the Google Cloud Free Trial landing page. At the top, there's a navigation bar with links for 'Por Que O Google?', 'Soluções', 'Produtos', 'Preços', 'Prir >', a search bar, 'Language' dropdown, and 'Fazer login'. Below the navigation is a button 'Entre em contato com nossa equipe' and a blue button 'Comece a usar gratuitamente'. The main section is titled 'Produtos de Nível gratuito' with a sub-section note: 'Esses produtos podem ser usados até o [limite de uso gratuito](#) especificado. O limite de uso gratuito não expira, mas está sujeito a alterações. Disponíveis para clientes qualificados.' Below this, there are six product cards:

- Compute Engine**: Máquinas virtuais escalonáveis de alto desempenho. 1 instância e2-micro por mês. [→](#)
- Cloud Storage**: Melhor desempenho, confiabilidade e preço para todas as necessidades de armazenamento. 5 GB por mês na classe Standard Storage. [→](#)
- BigQuery**: Armazenamento de dados de análise totalmente gerenciado, capaz de processar petabytes de dados. 1 TB de consultas por mês. [→](#)
- Google Kubernetes Engine**: Orquestração de contêineres com um clique usando clusters do Kubernetes, gerenciados pelo Google. Um cluster zonal ou de [região](#). [→](#)
- App Engine**: Plataforma para criação de aplicativos da Web e back-ends de dispositivos móveis escalonáveis. 28 instâncias/hora por dia. [→](#)
- Cloud Run**: Um ambiente totalmente gerenciado para executar contêineres sem estado. 2 milhões de solicitações por mês. [→](#)

- US\$ 300 em créditos ao criar uma conta
 - Válido durante 90 dias
- Mais de 20 produtos gratuitos
 - 1 instância e2-micro por mês
 - 5GB de Cloud Storage por mês
 - ...

<https://cloud.google.com/free>

CERTIFICAÇÕES DE COMPUTAÇÃO EM NUVEM

Certificações AWS



- Conhecimentos básicos da plataforma AWS
- Desenvolvimento de aplicações no AWS
- Construção de soluções no AWS
- Gerência e operação de workloads no AWS

Outras certificações



<https://aws.amazon.com/pt/certification/exams/>

Certificações Azure e GCP

- Azure

- <https://docs.microsoft.com/pt-BR/learn/certifications/>



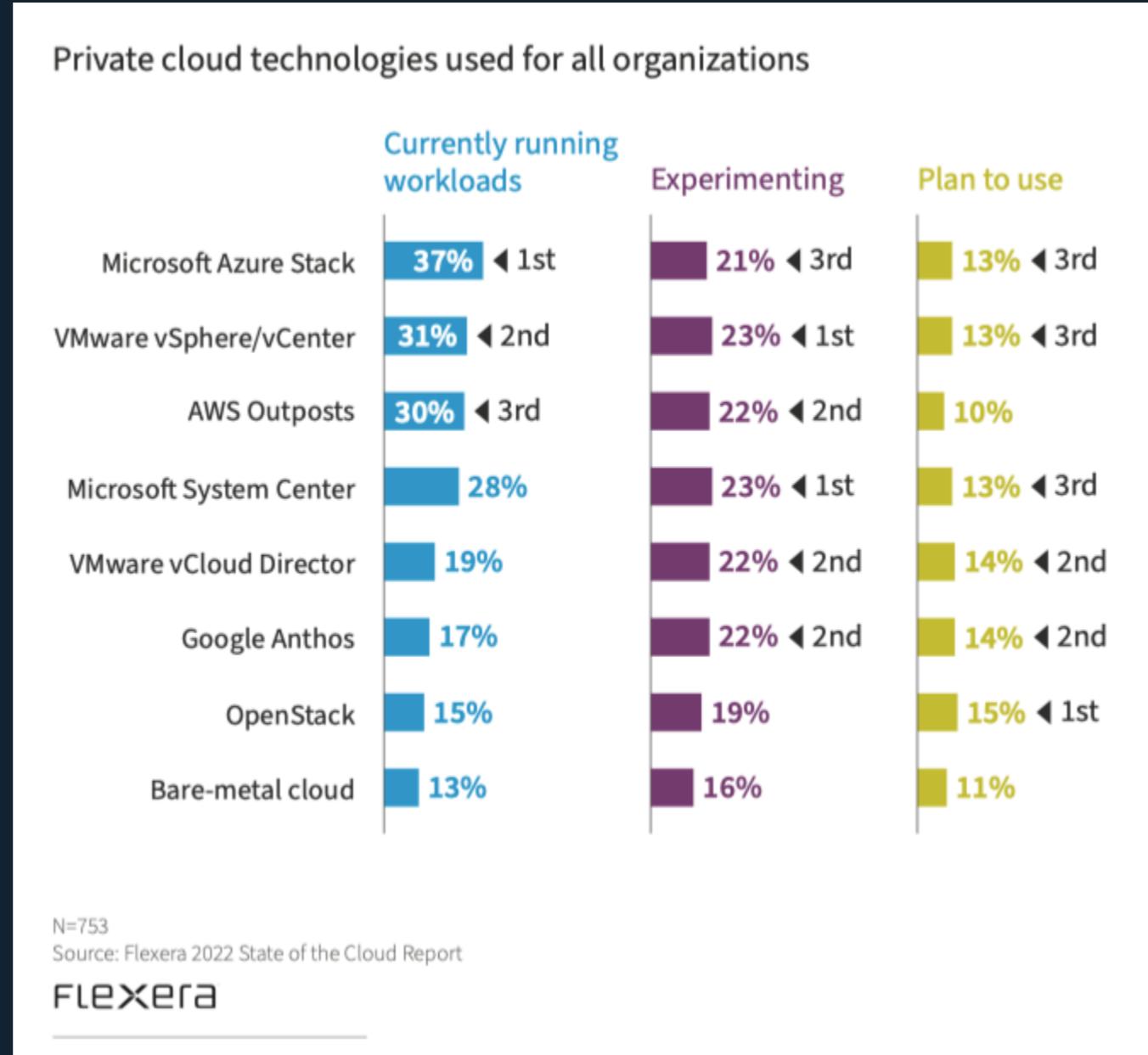
- Google Cloud Platform

- <https://cloud.google.com/certification>



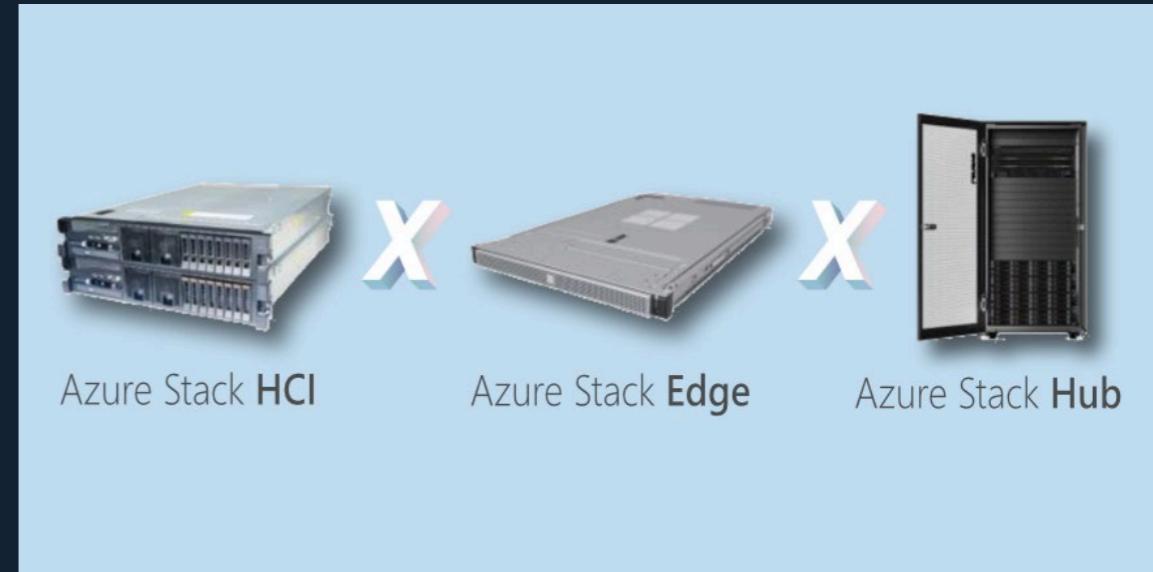
NUVENS PRIVADAS

Adoção de nuvens privadas



Microsoft Azure Stack

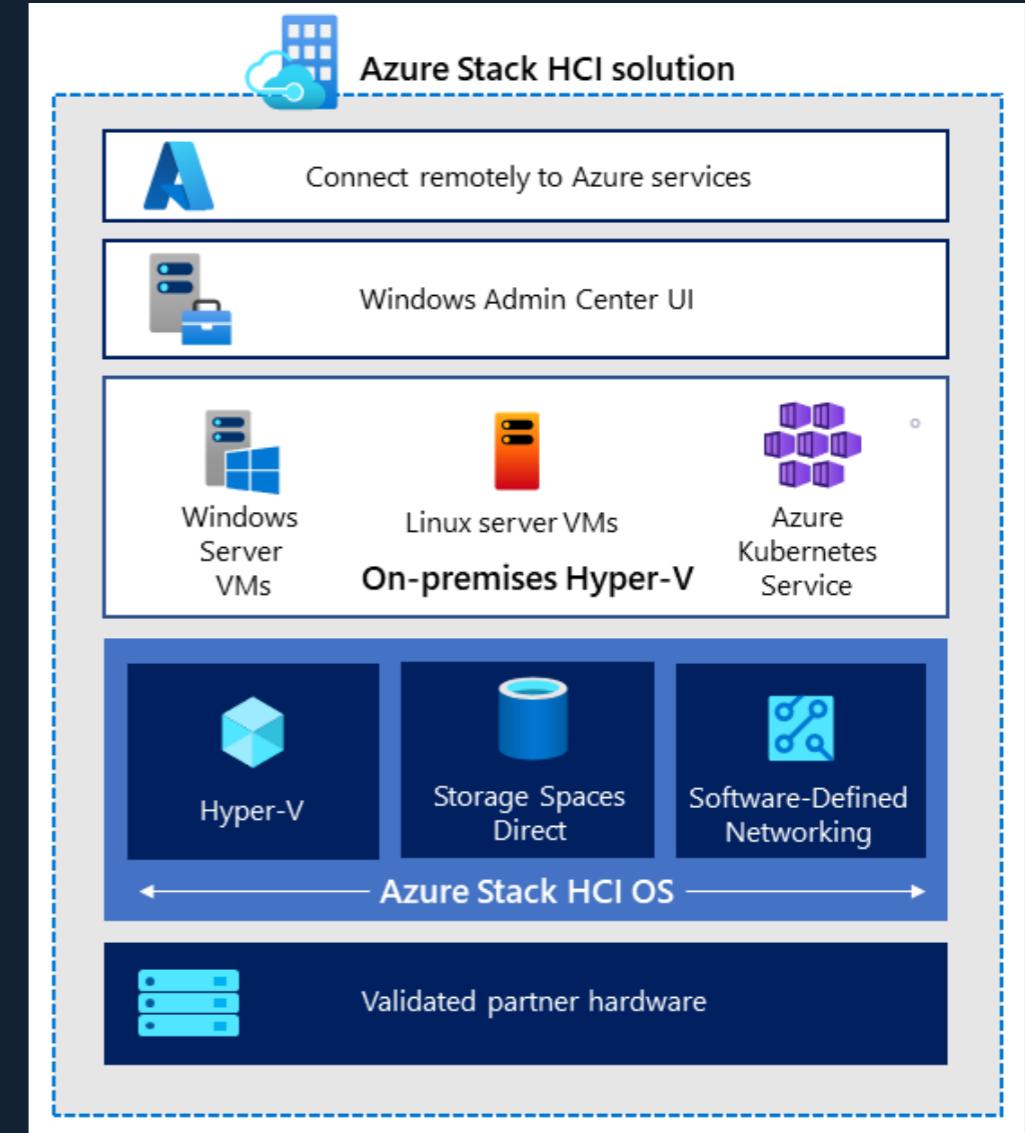
- Extensão do Microsoft Azure que permite que as organizações forneçam aplicativos, funções e máquinas virtuais da nuvem por meio de seus próprios data centers locais
- Versões
 - **Azure Stack HCI**: solução hiperconvergente para execução de aplicações virtualizadas localmente
 - **Azure Stack Edge**: dispositivo para acelerar cargas de trabalho com IA (antes de ir para a nuvem) e prover serviços Azure na borda
 - **Azure Stack Hub**: extensão para execução de aplicativos da nuvem em ambiente local, mesmo sem conexão ou para atender regulamentos



<https://lattinegroup.com/microsoft-azure/microsoft-azure-stack-o-que-e/>

Azure Stack HCI

- Solução de cluster HCI (infraestrutura hiperconvergente) que hospeda cargas de trabalho virtualizadas Windows e Linux e seu armazenamento em um ambiente híbrido que combina a infraestrutura local com os serviços de nuvem do Azure
- Aquisição através de parcerias de HW da Microsoft com o sistema operacional Azure Stack HCI pré-instalado, ou instalando o sistema operacional em um hardware compatível



<https://docs.microsoft.com/pt-br/azure-stack/hci/overview>

Azure Stack Edge

- Dispositivos de computação de borda entregues como uma solução de hardware como serviço.
- O dispositivo é enviado pela Microsoft para o cliente
- Opções para rack ou áreas externas



Azure Stack Edge Pro 2



Azure Stack Edge Pro R

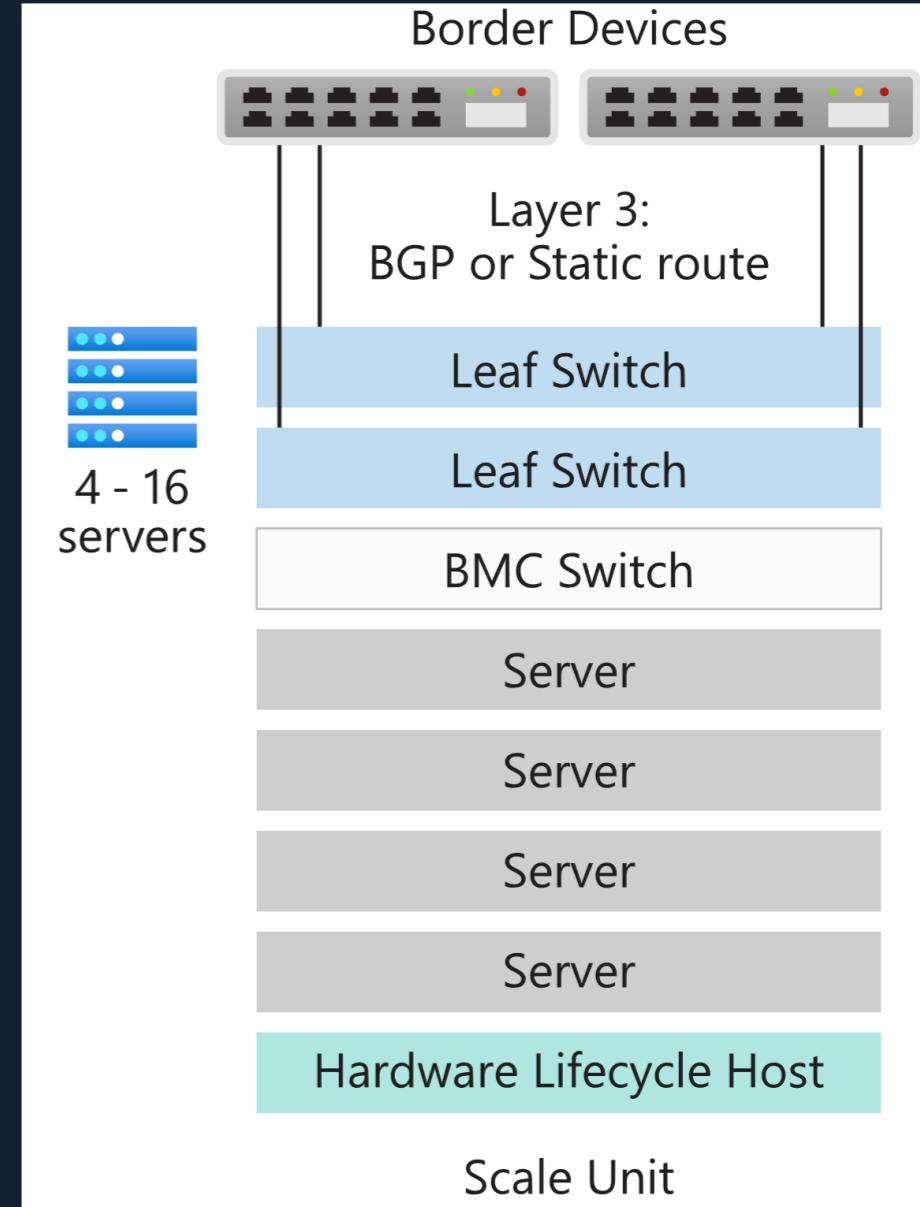


Azure Stack Edge Mini R

<https://docs.microsoft.com/pt-br/azure/databox-online/>

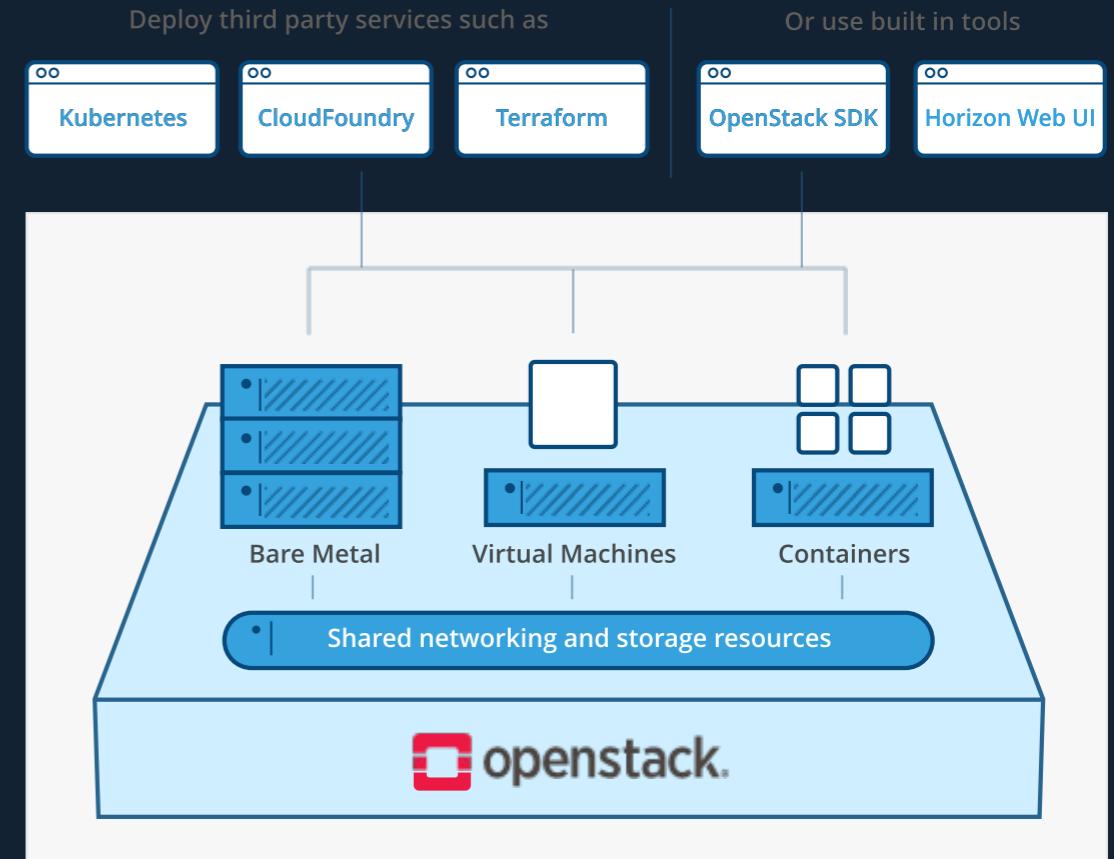
Azure Stack Hub

- Azure Stack Hub é uma extensão do Azure que permite executar aplicativos em um ambiente local e fornecer serviços do Azure em seu datacenter
- Os sistemas integrados do Azure Stack Hub são oferecidos por meio de uma parceria da Microsoft e de parceiros de hardware
- Pode ser utilizado de forma isolada ou conectado ao Microsoft Azure

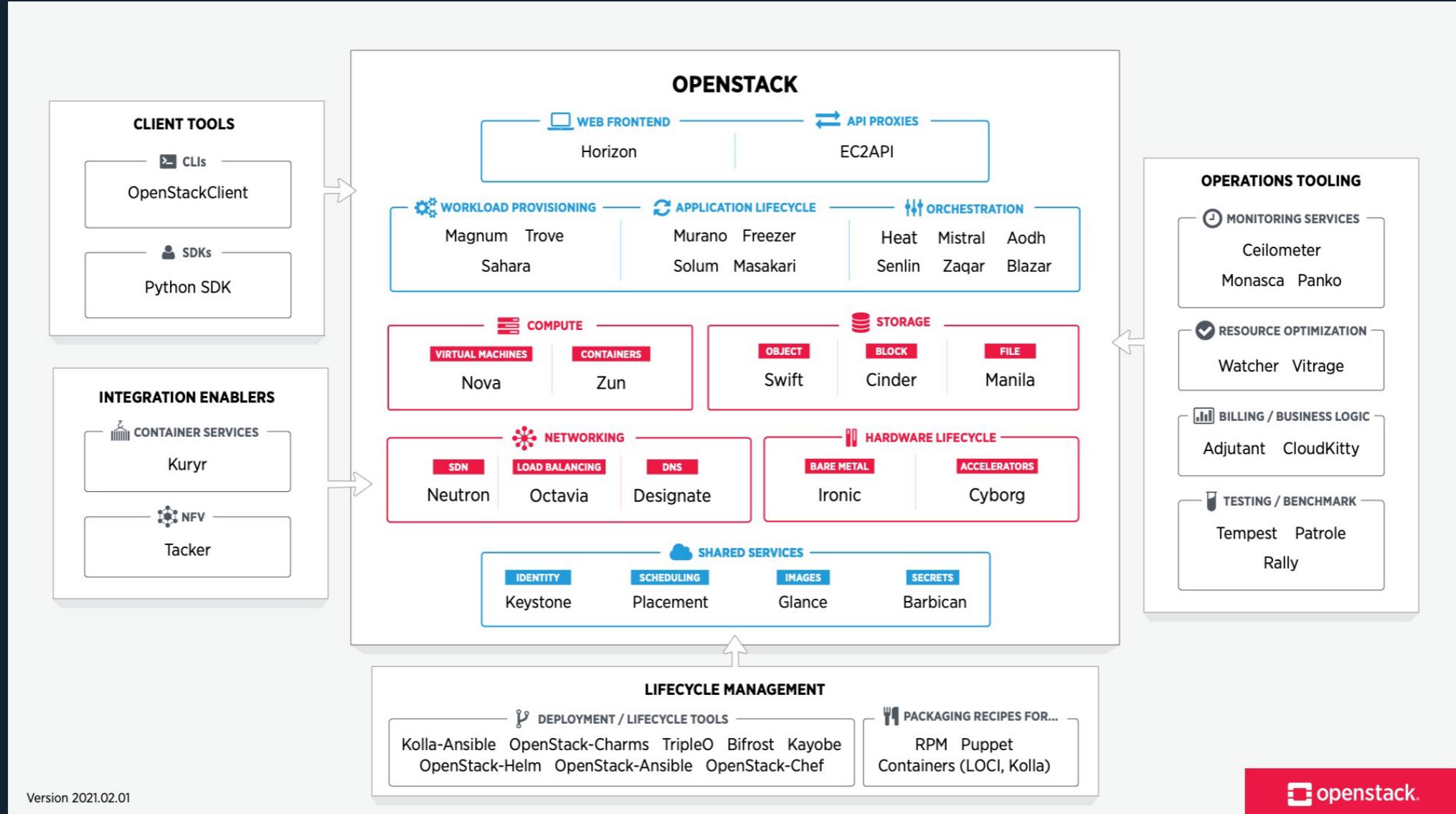


OpenStack

- Início: 2010 (parceria entre Rackspace e NASA)
- Gerenciado pela Open Infrastructure Foundation
 - <https://openinfra.dev/>
- Provê um conjunto de componentes para montar uma nuvem
 - Privada ou pública (RackSpace)



OpenStack Landscape



openstack Default admin

Project API Access Compute

Overview

Limit Summary

Category	Current Usage	Total Limit
Instances	Used 0 of 10	Used 0 of 20
VCPUs	Used 0 of 50	Allocated 0 of 50
RAM	Used 0Bytes of 50GB	Allocated 0 of 50
Floating IPs	Used 2 of 10	Used 0 of 10
Security Groups	Used 0 of 10	Used 0 of 10
Volumes	Used 0 of 10	Used 0 of 10

Usage Summary

Select a period of time to query its usage:
The date should be in YYYY-MM-DD format.

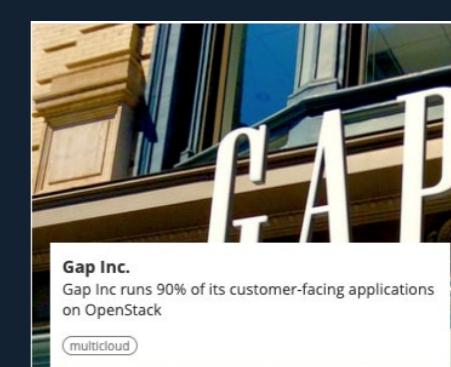
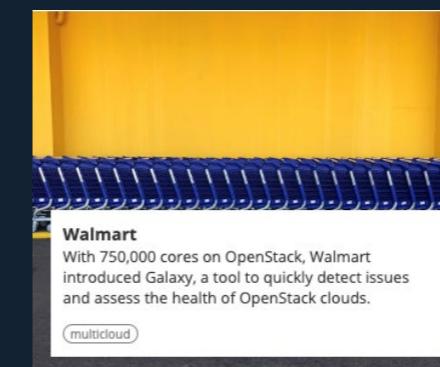
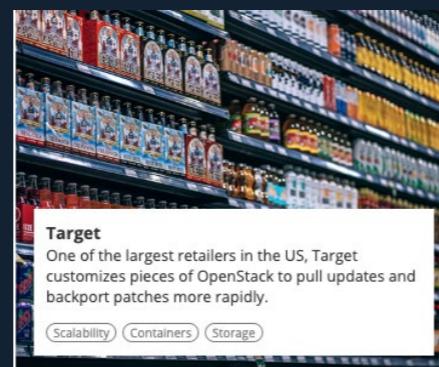
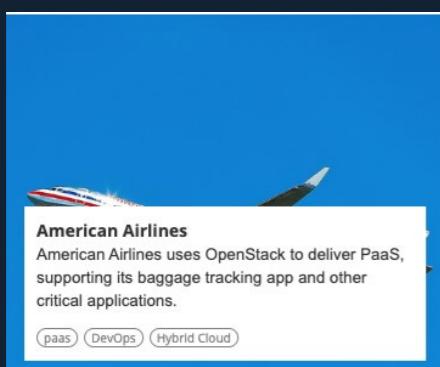
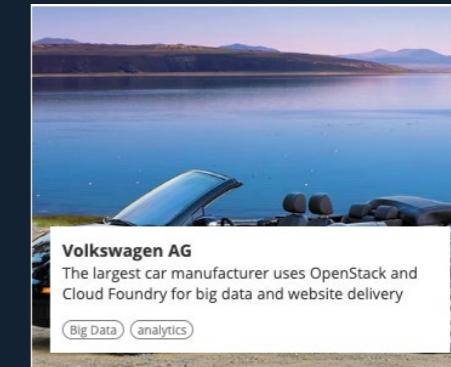
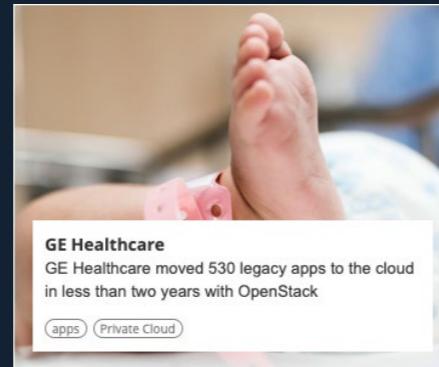
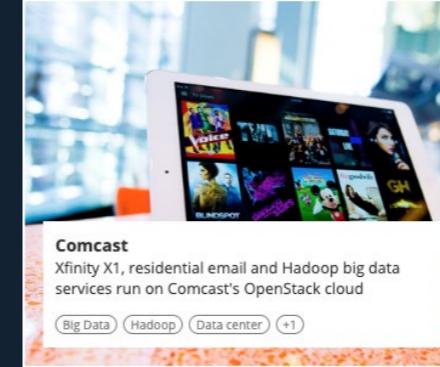
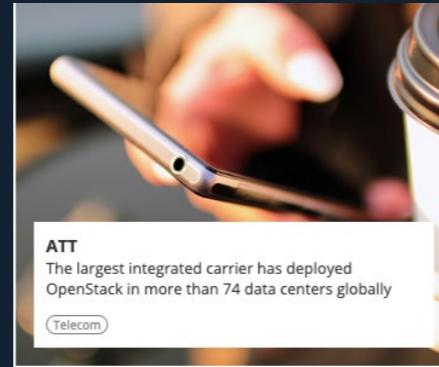
2018-01-05 to 2018-01-06

Active Instances: 0
 Active RAM: 0Bytes
 This Period's VCPU-Hours: 0.00
 This Period's GB-Hours: 0.00
 This Period's RAM-Hours: 0.00

Usage

Instance Name	VCPUs	Disk	RAM	Time since created
No items to display.				

OpenStack - Usuários

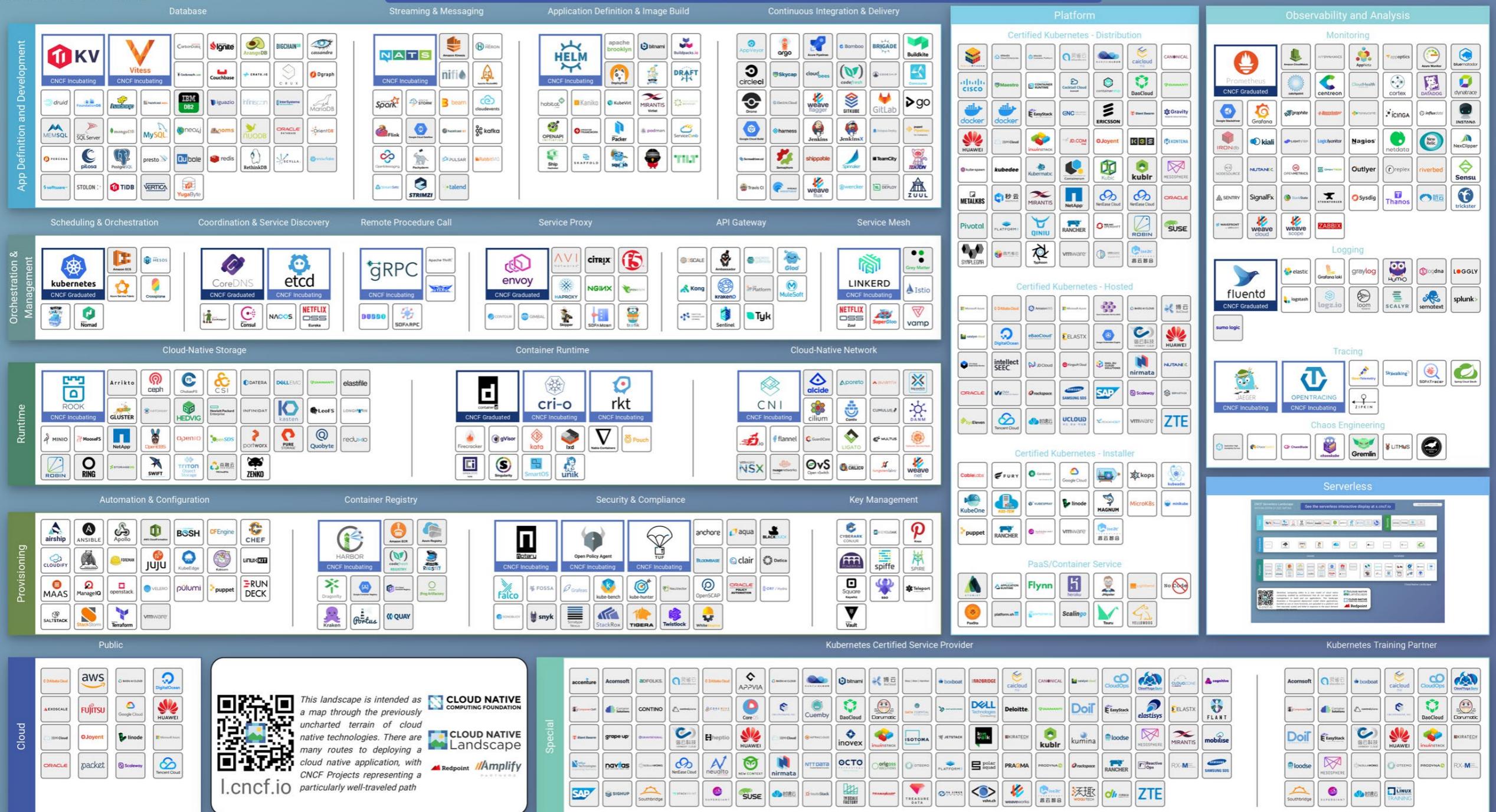


CLOUD NATIVE COMPUTING

Cloud Native Computing

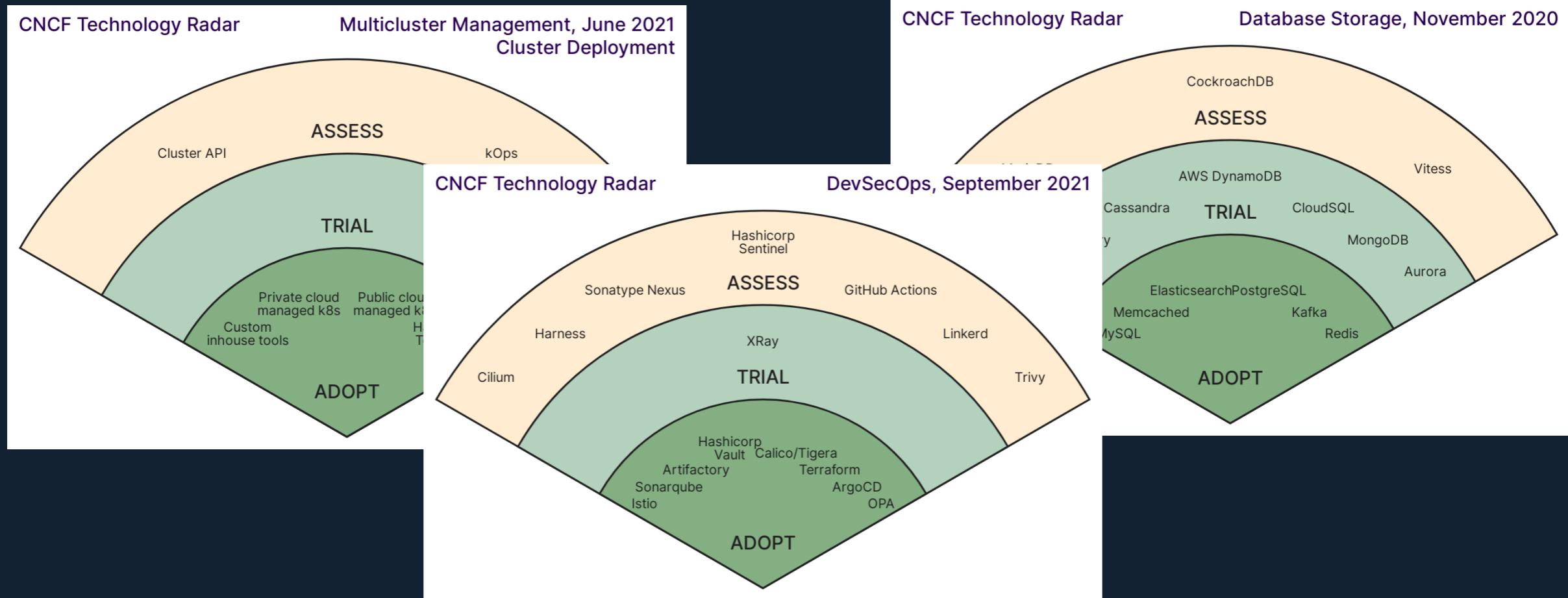
- **Cloud native** é uma abordagem para criar e executar aplicativos que exploram totalmente as vantagens do modelo de computação em nuvem
- Principais Drivers: Microserviços, Containers, DevSecOps, CICD
- 12 Factors app methodology (<https://12factor.net/>)
 - Metodologia para criar aplicativos de software como serviço
- **Cloud Native Computing Foundation** (<https://www.cncf.io>)
 - Conduz a adoção do paradigma de Cloud Native Computing auxiliando e sustentando um ecossistema de projetos de código aberto e não atrelados a nenhum fornecedor
 - Projetos: Kubernetes, Prometheus, envoy, rkt, etcd, etc

Arquitetura de Software em Nuvem – Tiago Ferreto



CNCF Technology Radar

- Avalia tecnologias de cloud native computing

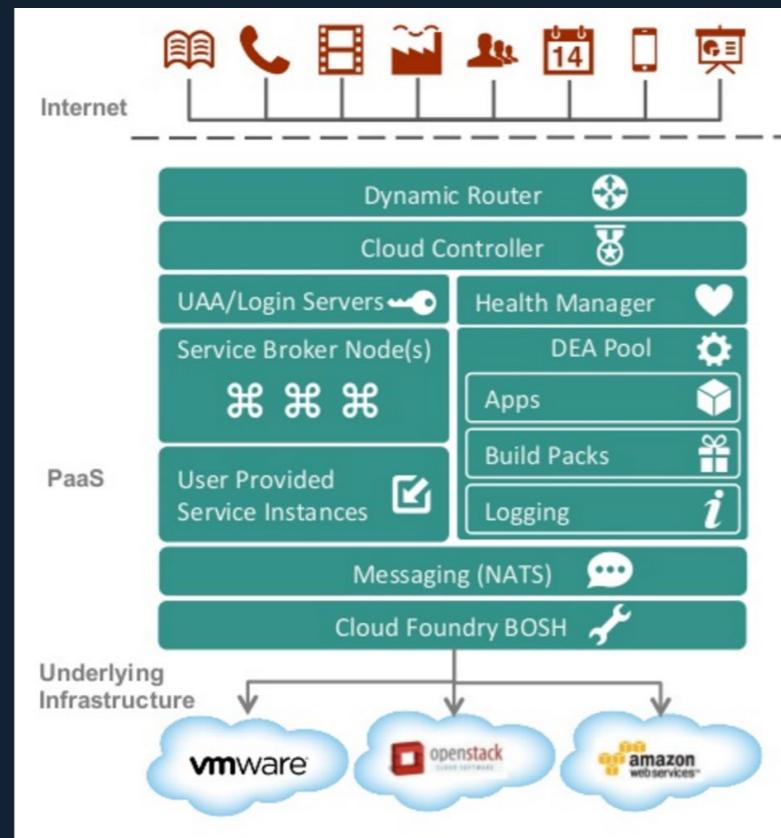


<https://radar.cncf.io/>

Cloud Foundry

<https://www.cloudfoundry.org/>

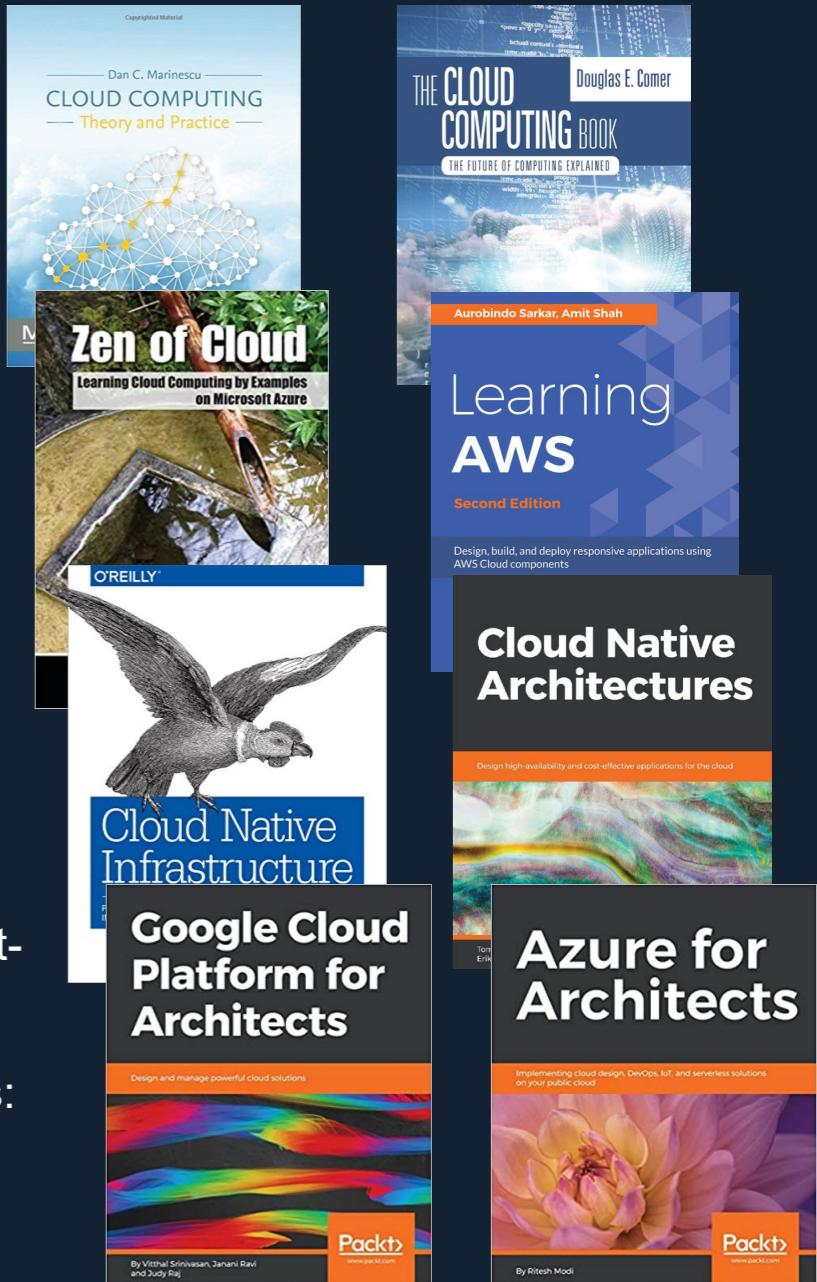
- Cloud Foundry é uma plataforma-como-serviço (PaaS) de código aberto, permitindo que o usuário escolha a nuvem, framework de desenvolvimento e serviços que desejar



<https://www.cloudfoundry.org/the-foundry/#cert-distros>

Referências – Cloud Computing

1. MARINESCU, Dan C. *Cloud computing: theory and practice*. Morgan Kaufmann, 2017.
2. COMER, Douglas E. *The Cloud Computing Book: The Future of Computing Explained*. Chapman and Hall/CRC, 2021.
3. BAI, Haishi. *Zen of Cloud: Learning Cloud Computing by Examples*. CRC Press, 2019.
4. SARKAR, Aurobindo; SHAH, Amit. *Learning AWS: Design, build, and deploy responsive applications using AWS Cloud components*. Packt Publishing Ltd, 2018.
5. GARRISON, Justin; NOVA, Kris. *Cloud native infrastructure: Patterns for scalable infrastructure and applications in a dynamic environment*. O'Reilly, 2017.
6. LASZEWSKI, Tom et al. *Cloud Native Architectures: Design high-availability and cost-effective applications for the cloud*. Packt Publishing Ltd, 2018.
7. SRINIVASAN, Vitthal; RAVI, Janani; RAJ, Judy. *Google Cloud Platform for Architects: Design and manage powerful cloud solutions*. Packt Publishing Ltd, 2018.
8. MODI, Ritesh. *Azure for architects: Implementing cloud design, DevOps, IoT, and serverless solutions on your public cloud*. Packt Publishing Ltd, 2017.



Referências – Cloud Computing

9. BENATALLAH, B. Cloud Computing – Methodology, Systems, and Applications. CRC Press. 2011.
10. ERL, T.; PUTTINI, R.; MAHMOOD, Z. Cloud Computing: Concepts, Technology & Architecture. Prentice Hall. 2013.
11. FOX, G; DONGARRA, J.; HWANG, K. Distributed and Cloud Computing. Morgan Kaufmann. 2013.
12. JACKSON, K. OpenStack Cloud Computing Cookbook. Packt Publishing Ltd. 2012.
13. SABHARWAL, N; SHANKAR R. S.. Apache CloudStack Computing. Packt Publishing Ltd. 2013.
14. LUCIFREDI, F; RYAN, M. AWS System Administration. O'Reilly Media, Inc. 2018



PUCRS online

Desenvolvimento Full Stack

Arquitetura De Software Em Nuvem

Tiago Ferreto

tiago.ferreto@pucrs.br

PUCRS online  uol edtech