

Computação em nuvem

Estudos de caso & exemplos de serviços



Prof. Dr. Marcos A. Simplicio Jr.
Laboratório de Arquitetura e Redes de Computadores
Departamento de Engenharia de Computação e
Sistemas Digitais
Escola Politécnica da Universidade de São Paulo



Objetivos – Aula 11

- Discutir alguns dos serviços de **IaaS** providos pela **nuvem pública** da Amazon: computação (EC2) e armazenamento (S3)



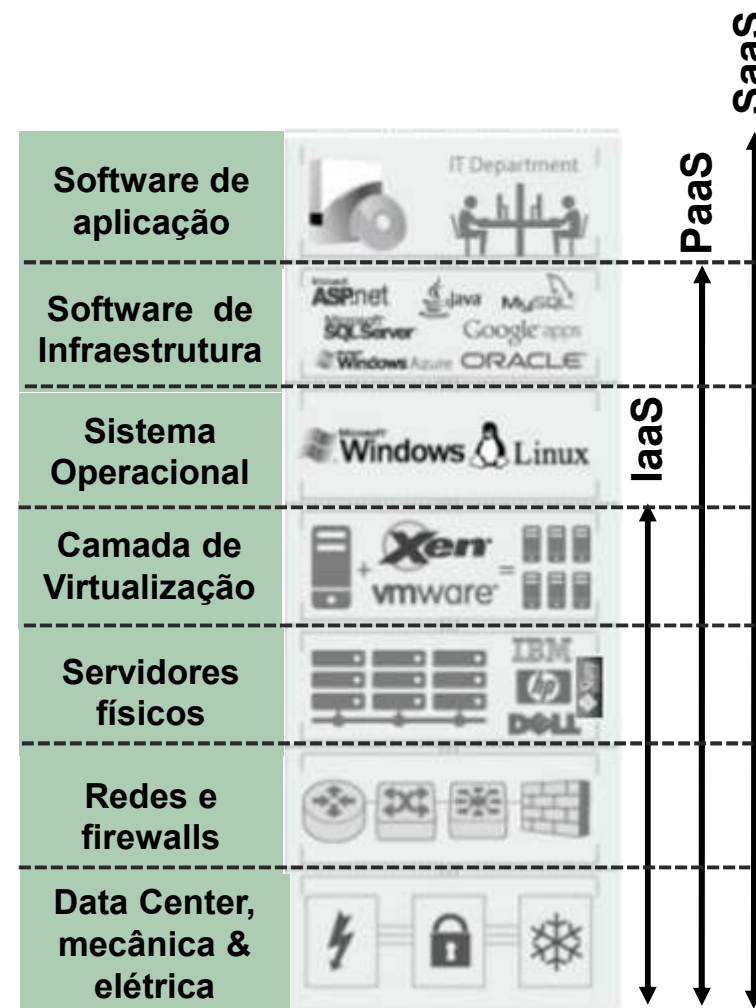
IaaS

□ Relembrando: **infraestrutura**

- Processamento
- Armazenamento
- Virtualização
- Outros recursos fundamentais

□ Consumidor pode executar software arbitrário

- Sistema operacional e aplicações





Amazon AWS


□ Diversos serviços

- ➤ **Computação:** máquinas virtuais (EC2), código reativo a eventos (Lambda), ...
- ➤ **Armazenamento:** diversos modelos (S3/Glacier)
- ➤ **Bancos de dados:** relacionais (RDS) ou não relacionais (DynamoDB: relacionamento chave-valor)
- ➤ **Administração e segurança:** gerenciamento de IDs, monitoramento (CloudWatch), ...
- ➤ **Análise de dados:** processamento de big data, ...
- ➤ **Serviços móveis:** envio de mensagens (SNS), identificação e sincronização de dispositivos (Cognito), dispositivos reais para testes (Device Farm), ...
- ➤ **Internet das Coisas:** envio e recepção de dados de sensores (IoT), botão no estilo "Amazon Dash" (IoT Button), ...
-








Amazon AWS: Serviços (obs.: lista não completa...)

 **AWS** ▾ **Services** ▾ **Edit** ▾






Marcos Simplicio ▾ Oregon ▾ Support ▾

Amazon Web Services





Compute

-  **EC2**
Virtual Servers in the Cloud
-  **Lambda**
Run Code in Response to Events
-  **EC2 Container Service**
Run and Manage Docker Containers




Storage & Content Delivery

-  **S3**
Scalable Storage in the Cloud
-  **Elastic File System** PREVIEW
Fully Managed File System for EC2
-  **Storage Gateway**
Integrates On-Premises IT Environments with Cloud Storage
-  **Glacier**
Archive Storage in the Cloud
-  **CloudFront**
Global Content Delivery Network








Database

-  **RDS**
MySQL, Postgres, Oracle, SQL Server, and Amazon Aurora
-  **DynamoDB**
Predictable and Scalable NoSQL Data Store
-  **ElastiCache**
In-Memory Cache
-  **Redshift**
Managed Petabyte-Scale Data Warehouse Service







Networking

-  **VPC**
Isolated Cloud Resources
-  **Direct Connect**
Dedicated Network Connection to AWS
-  **Route 53**
Scalable DNS and Domain Name Registration





Administration & Security

-  **Directory Service**
Managed Directories in the Cloud
-  **Identity & Access Management**
Access Control and Key Management
-  **Trusted Advisor**
AWS Cloud Optimization Expert
-  **CloudTrail**
User Activity and Change Tracking
-  **Config**
Resource Configurations and Inventory
-  **CloudWatch**
Resource and Application Monitoring
-  **Service Catalog**
Personalized Catalog of AWS Resources








Deployment & Management

-  **Elastic Beanstalk**
AWS Application Container
-  **OpsWorks**
DevOps Application Management Service
-  **CloudFormation**
Templated AWS Resource Creation
-  **CodeDeploy**
Automated Deployments
-  **CodeCommit**
Managed Git Repositories
-  **CodePipeline**
Continuous Delivery





Analytics

-  **EMR**
Managed Hadoop Framework
-  **Kinesis**
Real-time Processing of Streaming Big Data
-  **Data Pipeline**
Orchestration for Data-Driven Workflows
-  **Machine Learning**
Build Smart Applications Quickly and Easily




Application Services

-  **SQS**
Message Queue Service
-  **SWF**
Workflow Service for Coordinating Application Components
-  **AppStream**
Low Latency Application Streaming
-  **Elastic Transcoder**
Easy-to-use Scalable Media Transcoding
-  **SES**
Email Sending Service
-  **CloudSearch**
Managed Search Service
-  **API Gateway**
Build, Deploy and Manage APIs

Mobile Services

-  **Cognito**
User Identity and App Data Synchronization
-  **Device Farm**
Test Android, Fire OS, and iOS apps on real devices in the Cloud
-  **Mobile Analytics**
Collect, View and Export App Analytics
-  **SNS**
Push Notification Service

Enterprise Applications

-  **WorkSpaces**
Desktops in the Cloud
-  **WorkDocs**
Secure Enterprise Storage and Sharing Service
-  **WorkMail** PREVIEW
Secure Email and Calendaring Service

Resource Groups

A resource group is a collection of resources that share one or more tags. Create a group for each project, application, or environment in your account.

[Create a Group](#) [Tag Editor](#)

Additional Resources


[Getting Started](#)
See our documentation to get started and learn more about how to use our services.

[AWS Console Mobile App](#)
View your resources on the go with our AWS Console mobile app, available from [Amazon Appstore](#), [Google Play](#), or [iTunes](#).

[AWS Marketplace](#)
Find and buy software, launch with 1-Click and pay by the hour.

[AWS re:Invent - Register Now](#)
Join us for keynote announcements, technical sessions, bootcamps and more.

Service Health

 All services operating normally.

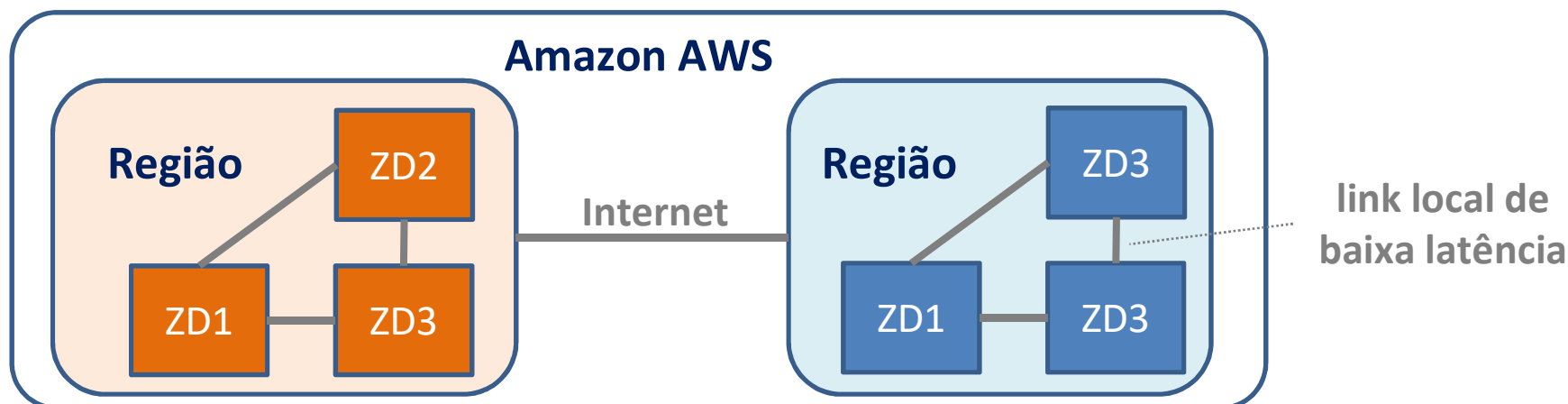
Updated: Aug 12 2015 17:38:00 GMT-0300

[Service Health Dashboard](#)

Amazon AWS: Localidade

□ Separação de **locais no AWS**

- **Região:** área geográfica (localização de data center)
 - Exemplos: Tokyo, Cingapura, Sydney (Asia); Frankfurt, Irlanda (Europa); São Paulo, Virginia, California, Oregon, Ohio (América)
- **Zonas de Disponibilidade:** locais independentes em região
 - Ex.: sa-east-1a, sa-east-1b, sa-east-1c
 - Para maior disponibilidade, pode-se fazer **replicação entre zonas**
 - Nota: "sa-east-1a" pode ser mapeado de forma diferente para contas distintas



IaaS: Amazon Elastic Compute Cloud (EC2)

❑ Serviço provido: **máquinas virtuais (VMs)** na nuvem

➤ Cliente configura **capacidade e características** desejadas



- Interface de monitoramento: cliente pode **ativar ou desativar** VMs
- VMs colocadas em **rede virtual privada** (isolamento lógico)
- Cliente escolhe **localização física** de cada VM (ex.: São Paulo, EUA,...)



➤ **CloudWatch**: visualização de **uso de recursos**

- Permite gerar estatísticas e configurar alarmes



➤ **Autoscaling**: métricas para aumentar ou reduzir número de instâncias de VMs **automaticamente**

- Métricas padrão (CPU, memória, banda, etc.) ou personalizadas



➤ **Balanceamento de carga elástico**: automaticamente distribui tráfego entre instâncias da aplicação, considerando carga atual

Amazon Elastic Compute Cloud (EC2)

❑ Grupos de **segurança**: “Firewall virtual”



- Controla **tráfego de entrada** de uma ou mais instâncias, conforme **regras** configuradas por usuário
 - **EC2-VPC** (Virtual Private Cloud): permite configurar **regras de saída**
- Cada máquina é associada a 1+ **grupos de segurança**
 - Nota: mais grupos = gerenciamento mais complexo...
- Política: “**tudo que não é expressamente permitido é proibido**”
 - Logo, regras adicionadas são para permitir algum tipo de tráfego

Editar regras de entrada

Tipo	Protocolo	Portas	Origem
HTTP	TCP	80	Qualquer 0.0.0.0/0
HTTPS	TCP	443	Qualquer 0.0.0.0/0

Adicionar regra
Cancel
Salvar



IaaS: Amazon Elastic Compute Cloud (EC2)

Modelos de **cobrança**



- “**Sob demanda**”: pagamento de acordo com horas de instâncias de VM ativas ou quantidade de dados transferidos



- “**Reservado**”: pagamento adiantado

- Descontos de até 60% em comparação com versão sob demanda

- “**Spot**”: cliente faz **oferta de preço** e, quando custo (conforme disponibilidade) fica abaixo do ofertado, cliente recebe serviço

- Interessante para aplicações que **podem ser interrompidas e reinicializadas** sem prejuízo para serviço: **serviço mensal** que pode esperar por preços mais baixos no mês; **simulações** podem ser **mais precisas** com processamento extra, mas isso é opcional.



- Pode também ser **combinado com outros tipos de instância** para acelerar processamento

- Descontos podem chegar a 85% comparados a versão sob demanda



IaaS: Amazon Elastic Compute Cloud (EC2)

- Memória: local ou EBS (Elastic Block Storage):
 - Memória **local**: persiste apenas enquanto máquina está ativa
 - Uso: armazenamento temporário
 - Memória **EBS**: persiste mesmo após VM ser desativada
 - **Propósito Geral** (SSD): desempenho normal (até 3000 IOPS/volume)
 - IOPS: Operações de Entrada/Saída por Segundo
 - **IOPS Provisionado** (SSD): alto desempenho (4000 IOPS/volume)
 - Aplicações com **uso intensivo** de memória (ex.: bases de dados)
 - **Disco magnético**: mais lentos e baratos (40-200 IOPS/volume)
 - Quando o **acesso aos dados é pouco frequente**
 - Dados espalhados por diferentes localidades físicas





IaaS: Amazon Elastic Compute Cloud (EC2)

- Diversos tipos de instâncias de **VMs**:
 - Uso **geral**: T2, M3 e M4
 - T2 (para tarefas em rajada, como servidores web): desempenho variável, com linha de base que pode ser ultrapassada eventualmente.
 - M3 (para tarefas com uso intenso de CPU, como codificação de vídeo): desempenho fixo
 - Otimizada p/ **computação** (ex.: processamento batch): C3, C4
 - Otmizada p/ **memória** (ex.: bancos de dados): R3, R4, X1
 - **Aceleração** : P2 (GPU de propósito geral), G2 (GPU para processamento gráfico), F1 (FPGA)
 - Otimizada para **armazenamento** (ex.: "big data"): D2 (armazenamento em disco rígido) e I2 (armazenamento SSD)



IaaS: Amazon Elastic Compute Cloud (EC2)

Exemplos de especificação e preço (Linux) – 2017

➤ Maiores informações: <http://aws.amazon.com/ec2>

Nome	vCPU	RAM (GiB)	Armazenamento (GB)	U\$/hora (sob demanda)
t2.nano	1	0.5	0 (EBS apenas)	0.0059 : 0.0101
t2.medium	2	4	0 (EBS apenas)	0.047 : 0.081
t2.2xlarge	8	32	0 (EBS apenas)	0.376 : 0.648
m4.large	2	8	0 (EBS apenas)	0.171 : 0.108
m4.16xlarge	64	256	0 (EBS apenas)	5.480 : 3.447
c3.large	2	3.75	2 x 16 SSD	0.163 : N/A
c3.8xlarge	32	60	2 x 320 SSD	2.600 : N/A
c4.large	2	3.75	0 (EBS apenas)	0.155 : 0.100
c4.8xlarge	36	60	0 (EBS apenas)	2.470 : 1.591
r3.8xlarge	32	244	2 x 320 SSD	5.597 : 2.660
i2.8xlarge	32	244	8 x 800 SSD	N/A : 6.820
d2.8xlarge	36	244	24 x 2000 SSD	N/A : 5.520

São Paulo : Ohio

IaaS: Amazon Simple Storage Service (S3)

- ❑ Serviço provido: **Armazenamento** na nuvem:



- Objetos de 1 byte a 5 terabytes
 - Qualquer número de objetos
- ❑ Níveis de disponibilidade
 - **Padrão**: replicação x3 e rapidamente acessível
 - **Redundância reduzida**: replicação x2 e rapidamente acessível
 - **"Glacier"**: replicação x3 e acessível após 3-5 horas
 - ❑ Maiores informações: <http://aws.amazon.com/s3/>

IaaS: Amazon Simple Storage Service (S3)

Exemplos de preço (São Paulo, 2017)

- Ex.: atualização de 10MB, com 1M downloads no mês → \$2500

Armazenamento: \$/mês por GB			
Terabytes	Padrão	Reduzida	Glacier
< 50	0.0405	0.0260	N/A
< 450	0.0390	0.0260	(Ohio: 0.004)
> 500	0.0370	0.0260	

Requisições: \$ / 10.000 requisições	
PUT, COPY, POST, LIST	0.005
Delete	Grátis
GET e outros	0.004

Transferência: \$ por GB	
Saída de dados	
< 1GB	Grátis
< 10 TB	0.250
< 40 TB	0.230
< 100 TB	0.210
< 350 TB	0.190
> 524 TB	Negociável
EC2 (mesma região)	Grátis
AWS (outra região)	0.160
Entrada de dados	Grátis

Prevê também serviço de aceleração, com tarifa adicional por GB

Amazon AWS: níveis gratuitos

- Por 12 meses (<https://aws.amazon.com/free/>):
 - **EC2**: t2.micro por 750h/mês (Win/Lin) com balanceamento de carga (Elastic Load Balancer)
 - **S3**: 5 GB de armazenamento padrão + (20k GET + 2k PUT + 15 GB entrada + 15 GB saída)/mês
 - **Elastic Transcoder** (transcodificação de vídeo): (20 min áudio + 20 min vídeo SD + 10 min vídeo HD)/mês
 - **Device Farm**: 1 teste de 250 dispositivos*minutos
 - **IoT** (recepção e envio de dados de sensores): 250K mensagens/mês
 - ...





Teste você mesm@!

- ❑ Crie um conta gratuitamente: <https://aws.amazon.com/free/>
 - Nota: necessário fornecer cartão de crédito
- ❑ Entre no serviço EC2 e crie uma instância
 - Escolha sistema operacional: você provavelmente vai querer começar com versões "free tier eligible"...
 - Escolha o tipo de instância: t2.micro são gratuitas
 - Configure instância, permitindo acesso externo (padrão)
 - Nomeie a instância
 - Configure acesso: tipos de acesso permitidos, controlados por firewall do AWS (por default, apenas SSH)
 - Crie chave de acesso SSH
 - Inicialize a instância
- ❑ Vários vídeos-tutorial na web
 - Ex. (explicativo e em português):
<https://www.youtube.com/watch?v=mCbiBcE1IA0>

Resumo

- ❑ Discutir alguns dos serviços de **IaaS** providos pela **nuvem pública** da Amazon: computação (EC2) e armazenamento (S3)
 - EC2: fornece diversos tipo de máquinas virtuais sob demanda, reservadas ou “quando disponíveis” (“spot”)
 - S3: fornece capacidade de armazenamento
 - Custos dependem do serviço contratado
 - Obs.: há também **cotas gratuitas**
 - **Calculadora:** <https://calculator.s3.amazonaws.com/index.html>
- ❑ Maiores informações:
 - <http://aws.amazon.com/>

