

## **Agenda**

- Definição de "Computação na nuvem"
  - Enquadramento histórico e tecnológico, a 5<sup>a</sup> utilidade, centros de dados à escala global
- Cenários de uso, Cloudnomics: CAPEX vs. OPEX
- Potencialidades e desafios
- Nuvens públicas e oferta de serviços
  - Amazon EC2, Microsoft Azure, Google Cloud Platform
  - Demonstração de uso de serviços
- As "novas" buzzwords, depois da Cloud



# Qual a definição de "Computação na Nuvem"?

- Definição académica
  - Vaquero et al. [VMC+08] estudou mais de 20 definições para extrair uma definição consensual
    - Definições comuns incluem termos: «escalabilidade, modelos pay-per-use, e virtualização»
  - Investigadores da University of California at Berkeley também apresentaram a sua visão [AFG+09]
    - «Cloud Computing refers to both the applications delivered as services over the Internet and the hardware and systems software in the datacenters that provide those services.»

[VMC+08] A Break in the Clouds: Towards a Cloud Definition, <a href="http://ccr.sigcomm.org/online/files/p50-v39n1l-vaqueroA.pdf">http://ccr.sigcomm.org/online/files/p50-v39n1l-vaqueroA.pdf</a>, visited 06-03-2022

[AFG+09] Above the Clouds: A Berkeley View of Cloud Computing, <a href="https://www2.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.pdf">https://www2.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.pdf</a>, visited 06-03-2022



# Qual a definição de "Computação na Nuvem"?

- Definição Institucional
  - A Comissão Europeia [EU12] define Cloud Computing como «the storing, processing and use of data on remotely located computers accessed over the internet.»

### Computação e armazenamento distribuídos

- Incluí várias camadas: hardware, middleware, plataformas;
- Uso de hardware de forma otimizada usando redes de computadores;

### Virtualização

Facilidade de mudança de requisitos de hardware;

#### **Economia**

Possibilidade de pagar conforme a utilização (pay-per-use)

[EU12] Unleashing the Potential of Cloud Computing in Europe, http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0529:FIN:EN:PDF visited 06-03-2022



# **USA: Cloud Computing (2011)**

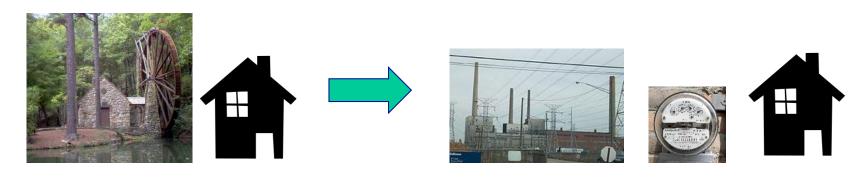
# National Institute of Standards and Technology Information Technology Laboratory

"A model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."



# Cloud Computing: Uma nova utility

- Comum a todas as definições está a noção subjacente de Utility Computing
  - Cloud Computing a 5<sup>a</sup> utilidade (depois da eletricidade, água, gás e telefone) [BYV+09]



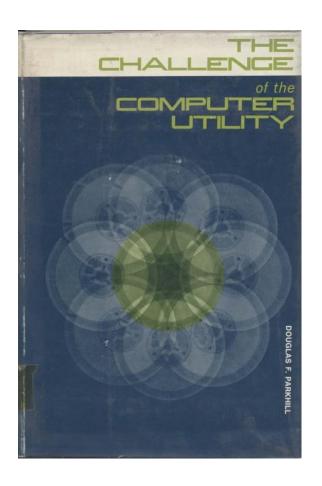
- Tal como outras utilidades, a computação está agora amplamente disponível, medida e paga por utilização
  - Existem vários "ambientes" de execução (máquinas virtuais, contentores, web app, serverless) e bases de dados

[BYV+09] Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility, Rajkumar Buyya, Chee Shin Yeo, Srikumar Venugopal, James Broberg, and Ivona Brandic, Future Generation Computer Systems, Volume 25, Issue 6, June 2009, Pages 599-616



## Utility Computing... de 1966

«A computer utility is basically a mechanism for permitting a number of remotely located **customers** to utilize the facilities of large central computer complex, with much the same ease and flexibility as **if the complex** was located in their own premises»

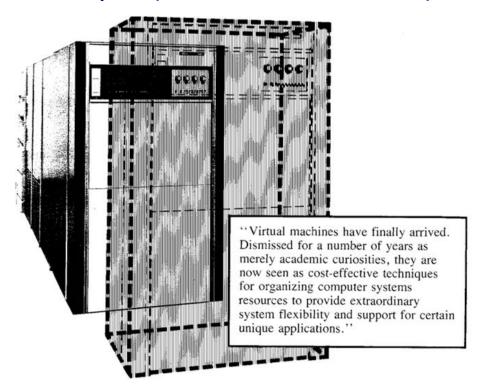


Douglas F. Parkhill, Challenge of the computer utility. Addison-Wesley Publishing Company (1966)



## Máquinas virtuais... em 1974

#### **IEEE Computer (Volume: 7, Issue: 6, June 1974)**



### Survey of Virtual Machine Research

Robert P. Goldberg

Honeywell Information Systems and Harvard University



«IBM System/370 Models 155 and 165, announced today, are designed for the emerging data processing needs of the Seventies -- large data bases, remote computing and high-throughput multiprogramming.»

in IBM Data Processing Division press technical fact sheet distributed June 30, 1970



# Então, porquê só agora?

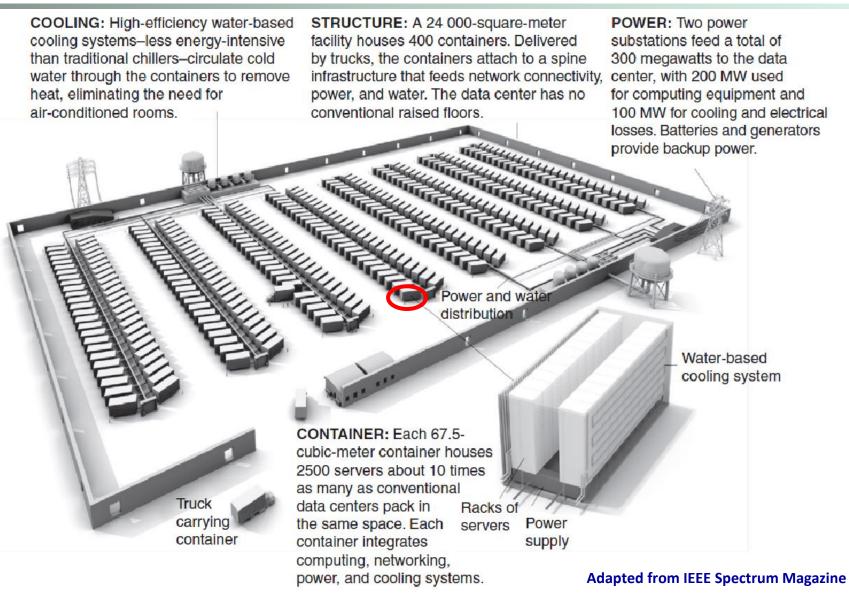
Nos anos 60 estávamos a tentar chegar à Lua...



- (Procura) Onde estavam os milhões de dispositivos móveis, o comércio eletrónico?
- (Oferta) O hardware e as redes de comunicação não eram suficientes para uma grande escala
- Hoje em dia temos:
  - Economia de escala
    - Centros de dados muito grandes com ganhos de eficiência energética de 5x a 7x, quando comparados com *clusters* privados
  - Amplo suporte de hardware para virtualização
    - Suporte ao nível de estruturas de CPU e memória para uma virtualização mais eficiente; diversas ferramentas
  - Investigação de técnicas para gestão de recursos partilhados



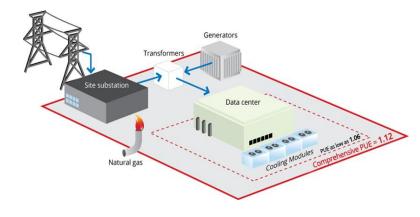
# Visão geral de um centro de dados moderno





# Energia – variável com impacto

- Cada fornecedor de computação na nuvem define o seu modelo económico
  - Existe diversidade nos elementos de computação: VMs, contentores, funções serverless, armazenamento, rede
  - Diferentes métricas: horas/minutos/segundos de utilização, número de nós, pedidos por intervalo, memória por pedido, Gbytes in e out
- A energia é um custo relevante para o fornecedor (e para o meio ambiente...)
  - Power Usage Efficiency =  $\frac{Total\ Facility\ Energy}{IT\ Equipment\ Energy}$ ;
    - Em 2020, a Google afirmava que em média conseguia um PUE de 1.12 (por cada 1 watt gasto em IT gasta +0.12 watts em arrefecimento)
  - O consumo de energia dos centros de dados nos Estados Unidos, representa 1.2% do consumo total, e está a subir [RM11]



[RM11] "The Cloud at Your Service: The when, how, and why of enterprise Cloud Computing", J. Rosenberg, A. Mateos, 2011

https://www.google.com/about/datacenters/efficiency/internal/



## Projeto "Natick"

"The vision of operating containerized datacenters offshore near major population centers anticipates a highly interactive future requiring data resources **located close to users**.

Deepwater deployment offers ready access to **cooling** and a controlled environment, and has the potential to be powered by **co-located renewable power sources.**"







# Regiões e número de zonas

- Os fornecedores de Cloud têm datacenters distribuídos por várias partes do mundo
  - Cada região tem um número de zonas associadas
  - Existem recursos que são específicos de uma zona, de uma região ou multi-região





Amazon AWS

Google GCP



# **Modelos/Tipos de Cloud**

**Private cloud** – Exclusiva a uma organização, eventualmente com múltiplas unidades de negócio;

**Community cloud** – Exclusiva a uma comunidade de consumidores que partilham os mesmos objetivos;

**Public cloud** - Pertence a um fornecedor (envolvendo múltiplos centros de processamento, em vários países). Os consumidores usam a Cloud segundo modelos de pay-per-use, com diferentes níveis de serviço (SLA - Service Level Agreement);

*Hybrid cloud* - Composição de infraestruturas de Cloud distintas (*private*, *community*, ou *public*) que interoperam através de modelos e tecnologias, que facilitam a portabilidade de aplicações ou dados.



# **Cloud Computing Stack**

### **Service Class**

### **Access & Tools**

### **Service contents**

SaaS Software as a Service

**Web Browser** 

Cloud Applications: Social Networks, Email, Office suites (Google docs), ERP, CRM, IAM (Identity and Access Management), Video processing, ...

PaaS
Platform as a
Service

**Development Environments** 

**Cloud Platform:** Programming languages, frameworks, *Mashups* editors, Web APIs, Data Storage models (Relational, NoSQL), *Data Analytics, ...* 

IaaS
Infrastructure
as a Service

Virtualization Manager **Cloud Infrastructure:** Computer servers, Data Storage, Firewall, Load Balancer, IP Addressing, VPN, ....



### **Cloudnomics - Conhecer bem os custos**

- Custos totais (TCO Total Cost of Ownership):
  - Custos de capital/investimento (CAPEX Capital Expenditure);
    - São despesas que criam benefícios futuros. Quando uma empresa gasta dinheiro para comprar ativos fixos, incorre em uma despesa de capital. Normalmente, o CAPEX requer um grande investimento inicial a ser amortizado à medida que o seu valor diminui ao longo do tempo. [RM11]
  - Custos operacionais (OPEX Operational Expenditure):
    - O OPEX é um custo contínuo para a execução de um produto, empresa ou sistema. É uma despesa diário que pode muito mais facilmente ser aumentada ou diminuída conforme as necessidades do negócio. [RM11]
- Como antecipar ou prever os custos mensais/anuais de soluções de IT
  - Licenciamento de software;
  - Contratos manutenção de hardware e software;
  - Custos investimento em novo hardware (ciclos 2/3/4 anos);
  - Custo de recursos humanos (serviços) em ações de manutenção;
  - Custos energéticos (eletricidade, ar condicionado);



### **Cloudnomics**

- De custos de instalação para custos de operação (CAPEX vs. OPEX)
  - O investimento de capital inicial (CAPEX) pode ser significativo
  - Custo de operação (OPEX) são dinâmicos, e variam consoante a evolução do negócio: e.g. mais tráfego dita maior alocação de recursos
  - Não é necessário comprar e instalar servidores, sistemas operativos, licenças,
  - Estes ganhos de eficiência podem ser uma vantagem competitiva
- As tabelas de preços dos fornecedores nem sempre são claras e podem variar ao longo do tempo.



# Caso de uso (i)

### **In-premisses**

2x Xeon

16 cores

**128 GB RAM** 

300 Gb SSD

Network

7900€

36 months

= 220€/month

### Não inclui:

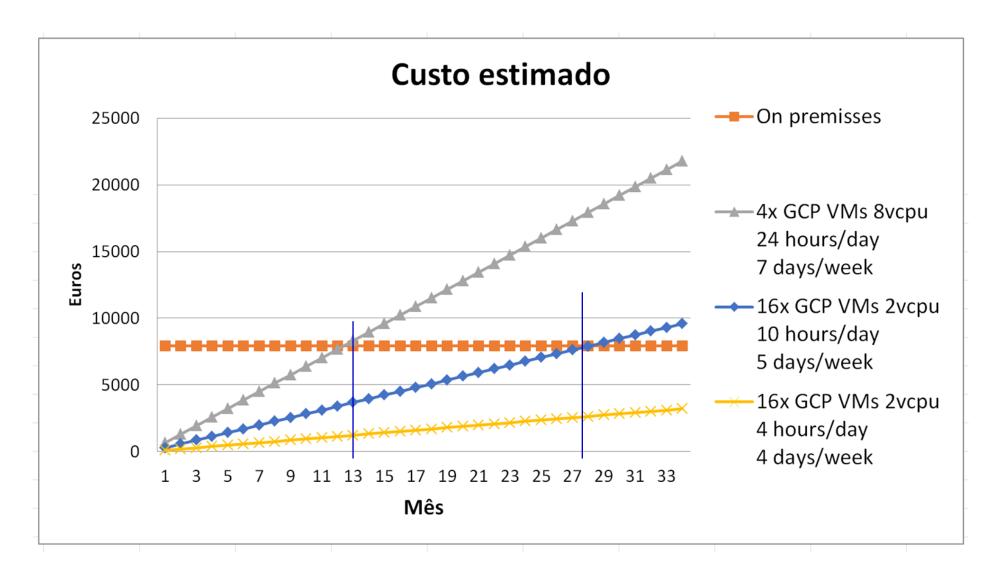
- Energia
- Arrefecimento
- Pesssoal de manutenção
- Edifício

Virtual Machine		
8 vCPU	2vCPU	
30 Gb RAM	8 Gb RAM	
300 Gb SSD	50 Gb SSD	
24 hours day	10 hours day	4 hours day
7 days	5 days	4 days
4 VMs	16 VMs	16 VMs
640€ Per month	282€ Per month	94€ Per month

Valores com base na calculadora de preços do GCP <a href="https://cloud.google.com/products/calculator/">https://cloud.google.com/products/calculator/</a>



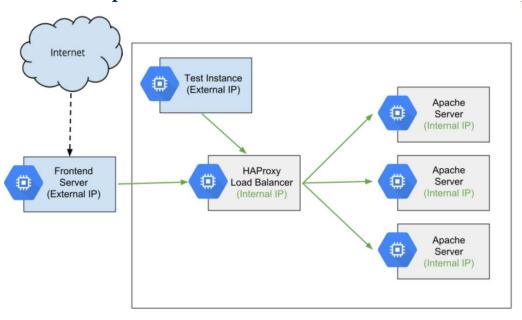
## Caso de uso (ii)

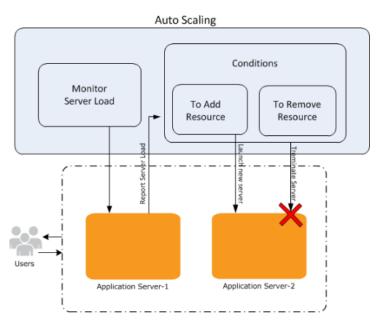




# Elasticidade (Auto Scaling)

- Capacidade de efetuar escalabilidade (scale up and scale down) automaticamente;
- Algoritmos para controlar e alocar recursos com as cargas associadas (workloads);
- Requer a capacidade de distribuir computações e dados por múltiplos nós;
- A coordenação dos nós é um ponto crítico, onde se pretende orquestrar múltiplas atividades executadas em múltiplos nós;







## Potencialidades e Oportunidades

- Cloud Computing é a maior mudança de paradigma desde o aparecimento do computador pessoal e da Internet:
  - Pay-per-use (utility prices)
  - Recursos computacionais virtuais infinitos Elasticidade
  - Interfaces self-service;
  - Evita a aquisição/instalação/manutenção de centros de processamento de dados (empresas/organizações start-up)



## O caso *Instagram*: de zero a mil milhões

- Outubro 2010: Kevin Systrom e Mike Krieger, criam uma aplicação de partilha de fotos:
  - 1º dia: 25.000 utilizadores;
  - 3 meses depois: 1 milhão de utilizadores (só para iphone);
  - Abril de 2012: 100 milhões de utilizadores (com Android), adquirida pelo Facebook por mil milhões de dólares;
  - Dezembro de 2014: 300 milhões de utilizadores.

Seria possível dois jovens construírem um centro de processamento com esta capacidade e num tempo tão curto?

Alojamento na Cloud da Amazon tirando partido da elasticidade



### **Desafios na Cloud**

- É um serviço de computação acessível por rede mas não resolve todos os problemas
  - A transferência de grandes quantidades de dados implica ter em conta a diferença entre:
    - Bandwidth Quantidade de dados que podemos transferir por unidade de tempo;
    - Latency Tempo para transferir dados de um ponto para outro.

Transferir 1TB a 1Mbps demora ~102 dias Transferir 1TB a 1Gbps demora ~2,3 horas

### Vendor lock-in

- É a "5ª utilidade" mas a interoperabilidade entre fornecedores é difícil ou inexistente (apesar de algumas ferramentas poderem minimizar este ponto)
- Sistemas legados que requerem configurações específicas de hardware
- Aplicações críticas para a organização, com requisitos de "tempo real"
- Problemas de segurança/privacidade dos dados
  - e.g. Registos de saúde?
  - Garantia de conformidade com o papel de processador de dados do RGPD?



## Desafios na Cloud: posição americana

### Clarifying Lawful Overseas Use of Data (CLOUD) Act

- Emenda ao *Stored Communications Act* (SCA), que já permitia a autoridades federais solicitar dados armazenados em solo dos EUA
- Pelo CLOUD Act as autoridades dos EUA têm a possibilidade de solicitar também dados armazenados fora dos EUA por empresas americanas (ex, Google, Amazon, Microsoft ...)

https://legalict.com/2018/11/08/the-cloud-act-and-its-consequences/ visited 06-03-2022



# Partilha de responsabilidades

Software-as-a-Service Platform-as-a-Service Infrastructure-as-a-Service (SaaS) (PaaS) (IaaS) ntorface Interface Interface Utilizador **Application Application Application Operating System Operating System Operating System** Limite Responsabilidade **Hypervisor Hypervisor Hypervisor** Computing hardware Computing hardware Computing hardware **Fornecedor** do serviço Storage hardware Storage hardware Storage hardware Network Network Network Local Infrastructure Local Infrastructure **Local Infrastructure** 



<sup>\*</sup> Figura adaptada de «Cloud Computing – Theory and Practice, 2nd Edition, Dan Marinescu, Morgan Kaufmann, 2017»

# BigData na Cloud

- As organizações geram enormes volumes de dados não estruturados oriundos de múltiplas fontes (email, redes sociais, etc.);
- Esses dados podem conter sinais importantes para o negócio, por exemplo sobre os clientes da organização;
- A análise destes dados é feita recorrendo a técnicas de Data Mining para extrair informação ou padrões implícitos, não triviais e desconhecidos;
- Tais técnicas requerem grande capacidade de recursos computacionais, de armazenamento, de processamento e de aplicações de aprendizagem automática, só ao alcance de algumas organizações;

Hoje os fornecedores de Cloud oferecem múltiplas e potentes soluções (*SaaS*) que permitem a qualquer organização realizar análise de dados quase em tempo real num contexto *BigData* 

### Ameaças?

- Privacidade dos cidadãos;
- Cyber Crime;
- Cyber Terrorismo;
- . . .



### A Cloud também falha

- Em 29 de junho de 2012 uma tempestade na região Este dos Estados Unidos, afetou vários *datacenters* da *Virginia* e empresas como o *Instagram* 
  - https://mashable.com/2012/06/30/aws-instagram/
  - Falha demorou a ser resolvida; Geradores de *backup* não arrancaram
  - O processo de balanceamento entre datacenters também teve problemas e as
     VMs demoraram a voltar a funcionar
- Google Cloud Status Dashboard
  - <a href="https://status.cloud.google.com">https://status.cloud.google.com</a>
  - <a href="https://status.cloud.google.com/incident/compute">https://status.cloud.google.com/incident/compute</a>



### Resumo dos desafios e dificuldades

- Evitar situações de "Vendor lock-in";
- Incerteza sobre o controlo físico de Hardware, Software e Dados;
- Incerteza sobre o suporte para auditar ou investigar falhas;
- Incerteza sobre leis aplicáveis;
- Dificuldade de definir SLAs adequados;
- Ausência de standards e de interoperabilidade entre fornecedores;
- Adaptar/migrar aplicações para Clouds públicas pode ser complexo;
- Multi-tenant versus multi-instance, isto é, uma única instância do serviço, serve múltiplas organizações versus instâncias do serviço separadas para cada organização;
- ✓ Para chegar à Cloud é necessário ter comunicações adequadas
  - "A dead Internet connection means no work"



### **Análise SWOT**

	INTERNO				
	STRENGTHS (forças)	WEAKNESS (fraquezas)			
POSITIVO	<ol> <li>Custo eficaz</li> <li>Flexibilidade e inovação</li> <li>Modelos de consumo e custos simples</li> <li>Rapidez de fornecimento de sistemas e aplicações</li> <li>Infraestrutura segura</li> <li>Instalações de acordo com normas</li> <li>Resiliência e recuperação de falhas</li> <li>Redução de custos de manutenção</li> <li>Níveis de acessibilidade convenientes</li> <li>Controlo de recursos adequado</li> <li>Independência de tempo e localização</li> <li>Poupança de energia</li> <li>Proteção do ambiente</li> <li>Utilização amigável</li> <li>Expansibilidade (Elasticidade)</li> </ol>	<ol> <li>Requistos de formação/aprendizagem</li> <li>Novo desenvolvimento das aplicações</li> <li>Aumento da dependência de terceiros</li> <li>Ligação à internet com alta largura de banda</li> <li>Dificuldade de integrar software local</li> <li>Botlenecks na transferência de dados</li> <li>Perda de controlo fisico dos dados</li> <li>Perda de controlo e garantia sobre QoS</li> <li>Dificuldade na garantia de localização dos fornecedores</li> </ol>	EGATIVO		
╚	OPPORTUNITIES (potencialidades)	THREATS (ameaças)	Z		
	<ol> <li>Pagar o que se usa (pay-per-use)</li> <li>Acessível a pequenas empresas sem grandes investimentos</li> <li>Facilidade de expansão do negócio</li> <li>Melhoria no posicionamento no mercado</li> <li>Adapatação a necessidades futuras</li> <li>Procesos de negócio normalizados</li> <li>Rapidez na solução de problemas</li> <li>Ambientes de trabalho modernos</li> <li>Soluções de acordo com as últimas tecnologias</li> </ol>	<ol> <li>Segurança dos dados</li> <li>Falta de regulação (nacional e internacional)</li> <li>Dificuldade de migração para outro fornecedor (vendor lock-in)</li> <li>Custos escondidos (backup, recuperação e resolução de problemas)</li> <li>Compatibilidade reduzida</li> <li>Resistência dos envolvidos à mudança</li> </ol>			
	EXTERNO				



### Oferta de Serviços na Cloud

- Pela sua importância existem serviços cuja oferta é comum a todos os fornecedores, independentemente dos nomes comerciais que os designam:
  - Armazenamento massivo de dados: Amazon S3; Azure storage; Google Cloud Storage; ...
  - Criação de VMs com múltiplas capacidades de processamento, de memória
     RAM e de storage: Amazon EC2; Azure VM; Google Compute Engine; ...
  - Armazenamento de dados NoSQL (Key, Value): Amazon DynamoDB; Azure
     CosmosDB e DocumentDB; Google Bigtable; Firestore; ...
  - Mensagens: Amazon Simple Notification Service; Azure Queues; Google Pub/Sub; ...
  - Data analytics (BigData; Data Mining; Business Intelligence, MapReduce, ...):
     Amazon Elastic MapReduce (EMR); Azure HDInsight; Google MapReduce;
     DataFlow; TensorFlow; ...



### Exemplos de ofertas públicas de Cloud









**vm**ware

VMware vCloud® Air™

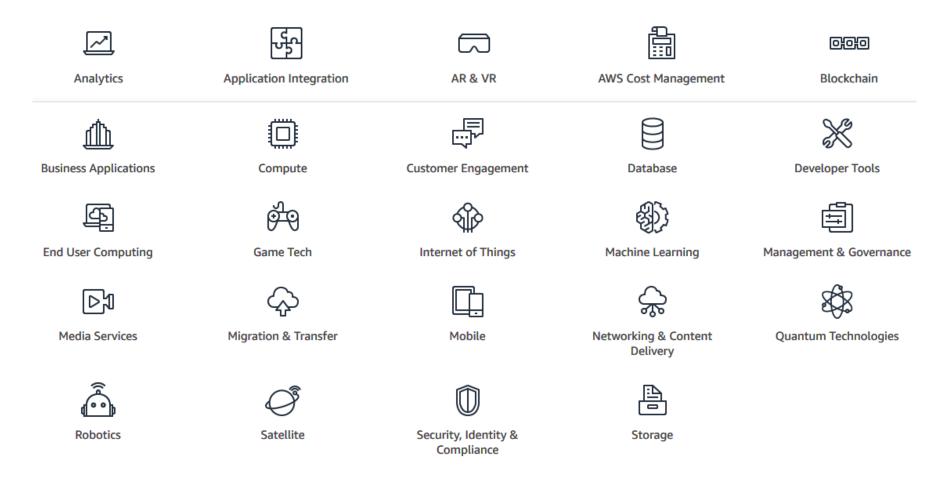






## **Amazon Web Services (AWS)**

### https://aws.amazon.com/products/





### **Microsoft Azure**

#### Virtual Machines



Provision Windows and Linux virtual machines in minutes

#### Storage



Durable, highly available and massively scalable cloud storage

#### **App Service**



Create web and mobile apps for any platform and any device

#### **Cloud Services**



Create highly available, infinitely scalable cloud applications and APIs

#### **SQL** Database



Managed relational SQL Database-as a-service

#### DocumentDB



Managed NoSQL document database-as-a-service

#### **Azure Active Directory**



Synchronise on-premises directories and enable single sign-on

#### Apache Spark for Azure HDInsight



Apache Spark in the cloud for mission-critical deployments

#### Backup

https://azure.microsoft.com/en-gb/services/



Simple and reliable server backup to the cloud

#### Apache Storm for HDInsight



Real-time stream processing made easy for big data

#### HDInsight



Provision cloud Hadoop, Spark, R Server, HBase and Storm clusters

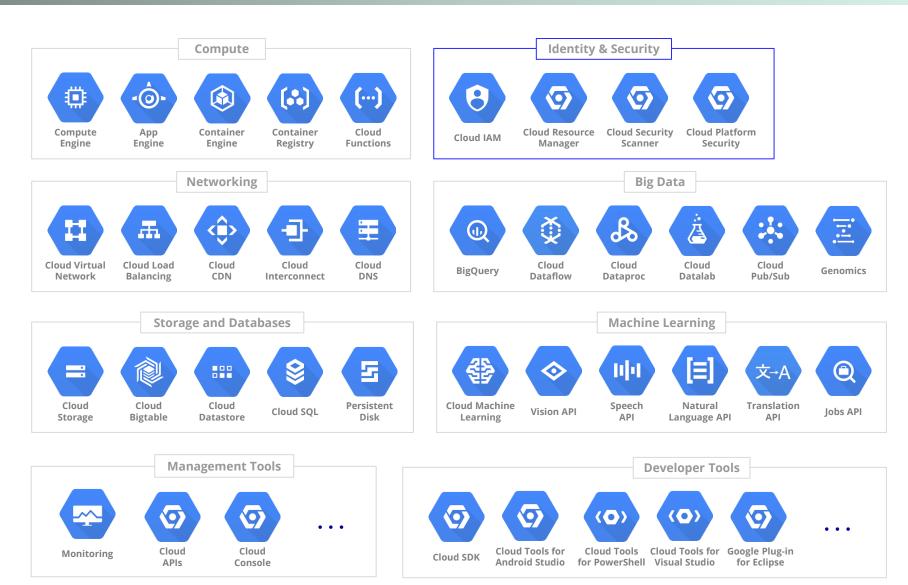
#### R Server for HDInsight



Predictive modelling, machine learning and statistical analysis for b...



# **Google Cloud Platform (GCP)**





## Google Cloud Platform

### Computação

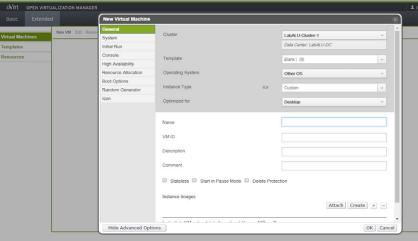
- Virtual Machines (incluindo máquinas preemptíveis)
- Managed Containers (Kubernetes, a cluster of containers)
- Cloud functions
- Armazenamento
  - Discos persistentes
  - Cloud Storage, 4 tipos de armazenamento (multi-regional, regional, nearline, coldline)
  - Bases de Dados relacionais e NoSQL (Firestore)
- Message Queueing Publish/Subscribe (Pub/Sub)
- Machine Learning and Big Data
  - Speech API, Vision API, ...
  - Big Query, ...



## **Clouds privadas**

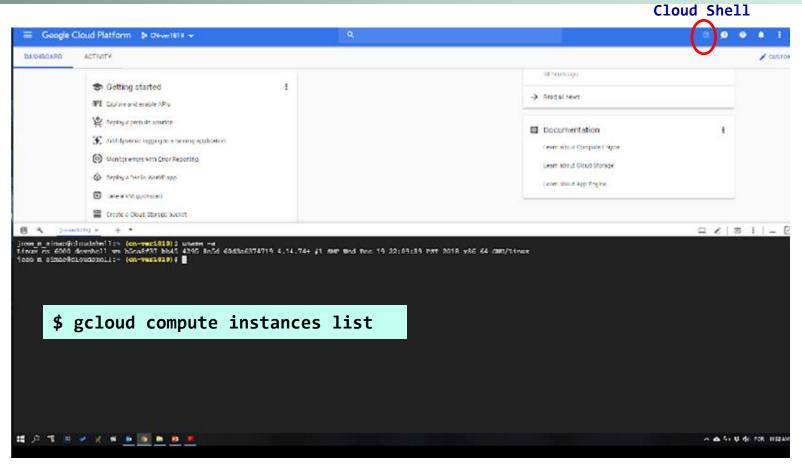
- A segurança, disponibilidade, comunidade de utilizadores, economia de escala, podem tornar a opção on-premise a melhor solução
- Exemplos: Eucalyptus, OpenNebula, OVirt
- ISEL / ADEET usa OVirt no seu centro de dados privado
  - OVirt usa o hypervisor Kernel-based Virtual Machine (KVM)







### **Demonstração 1**



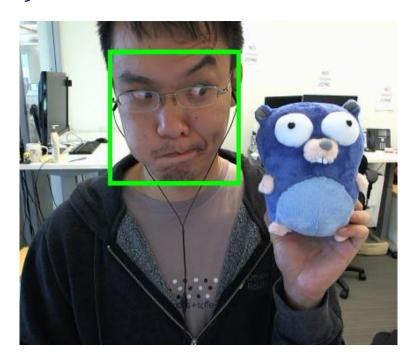
- A console web do GCP reúne o acesso a todos os serviços
  - Cloud shell: máquina virtual que corre num datacenter da Google, associada à conta do utilizador e com acesso aos diversos serviços



### Demonstração 2

 Exemplo de acesso programático à API de reconhecimentos de imagem (faces de pessoas e labels)





```
Labels for image data\cat.jpg:
Cat (score: 0.996)
Small to medium-sized cats (score: 0.976)
Felidae (score: 0.963)
```



# As "novas" buzzwords, depois da Cloud

- A computação distribuída de maior escala começou com terminais a aceder a centros de dados remotos
  - Quando os computadores pessoais se tornaram mais capazes, a computação ficou concentrada no computador local
- As aplicações web moveram a computação de novo para um ponto central
  - Mas de novo para o browser, (e.g. single-page application) e para os dispositivos móveis
- A Grid reinventou o centro de dados remoto e a Cloud tornou-o num serviço acessível a todos, com um modelo de custos
- Vemos mais uma vez um movimento de aproximar a computação dos utilizadores e das fontes de informação
  - IoT, Fog, Edge -- the cloud continuum



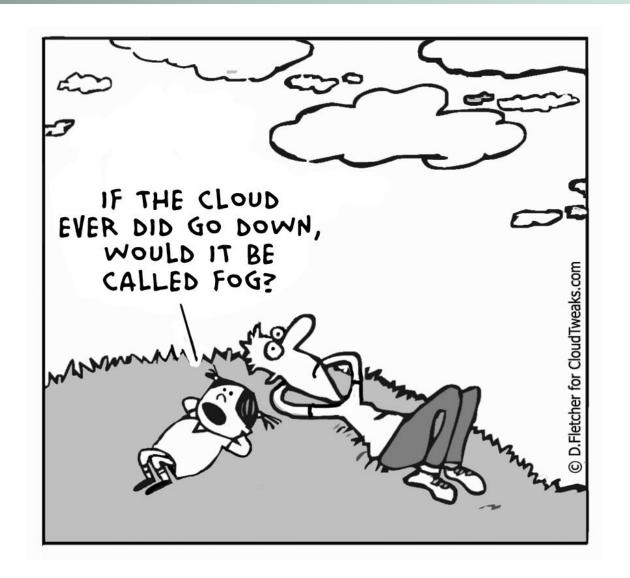




Figure 1. Connecting More and Different Kinds of Things Directly to the Cloud Is Impractical



### **Cisco Whitepaper**,

"Fog Computing and the Internet of Things: Extend the Cloud to Where the Things Are"



# Interação entre Fog e Cloud

### Fog nodes

- Recebem dados de dispositivos IoT usando múltiplos protocolos de tempo real (por exemplo *Sigfox*, NFC, *Zigbee*, *Bluetooth*, etc.) e protocolos M2M (*Machine-to-Machine*) por exemplo MQTT\*;
- Armazenam dados transitórios (1 a 2 horas);
- Executam aplicações de controlo ou análise de dados de pequena dimensão com tempos de resposta da ordem dos milissegundos;
- Enviam resumos de dados para infraestruturas Cloud.

### Plataforma Cloud

- Recebe e agrega os sumários de dados recebidos dos Fog nodes;
- Realiza análise dos dados oriundos dos dispositivos IoT conjuntamente com outros dados relativos ao negócio;
- Envia dados e diretivas para os Fog nodes baseados no negócio

