



CENTRES DE PROCESSAMENT DE DADES (CPD)



David López

2022-2023/Q1



UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH



LET'S BEGIN

David López

V.1.1.7

UPDATED SPRING 2022



UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Outline

- 1.1 – Summary
- 1.2 – Why Data Centers?
- 1.3 – Big Data / Analytics
- 1.4 – Mobility and Social
- 1.5 – Smart & Green Computing
- 1.6 – Is this *Cloud Computing*?
- 1.7 – Offering Services
- 1.8 - Some Key Concepts

Acknowledgements:

If I have seen further it is by standing on ye sholders of giants.

Sir Isaac Newton. Letter to Robert Hooke, circa 1676

Part of these slides have been borrowed from several talks by Jordi Torres

1- INTRODUCTION

3

Key technical concepts

- IT evolution
- Big data centers
- Opex & Capex costs
- High Performance Computing
- Cloud Computing
- Software as a Service (SaaS)
- Platform as a Service (PaaS)
- Infrastructure as a Service (IaaS)
- Faults & Errors
- FT, HA, DR, SLA

References

- Empreses en el núvol
 - Jordi Torres 2011

Additional references

- NIST definition of SaaS, IaaS and PaaS:
csrc.nist.gov/publications/drafts/800-146/Draft-NIST-SP800-146.pdf

1.1. SUMMARY

4

Every 15 years

A technological leap
in IT

redefining IT but
also business

1950 1965 1980 1995 2010 2025

1980: The PC is born



1950 1965 1980 1995 2010 2025



1950 1965 1980 1995 2010 2025

1995: The Internet



1950 1965 1980 1995 2010 2025

2010: Big Data Centers



1950 1965 1980 1995 2010 2025

1.2. WHY DATA CENTERS?

2025: ?

The best way to predict the future is to invent it.



Alan Kay

1950 1965 1980 1995 2010 2025

1.2. WHY DATA CENTERS?

Why big data centers?

Mobility & Social

Big Data / Analytics

Smart & Green Computing

1.2. WHY DATA CENTERS?

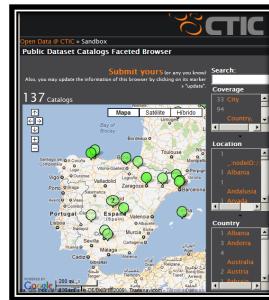
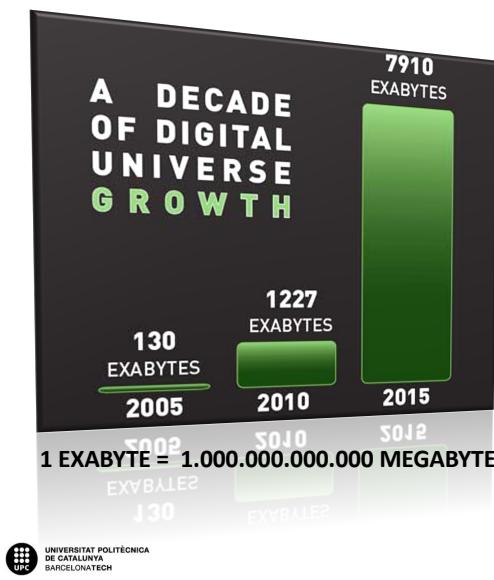
Why big data centers?

Mobility & Social

Big Data / Analytics

Smart & Green Computing

1.3. BIG DATA/ ANALYTICS



2013 International Data Corporation (IDC) Prediction for 2020:
40 zettabytes



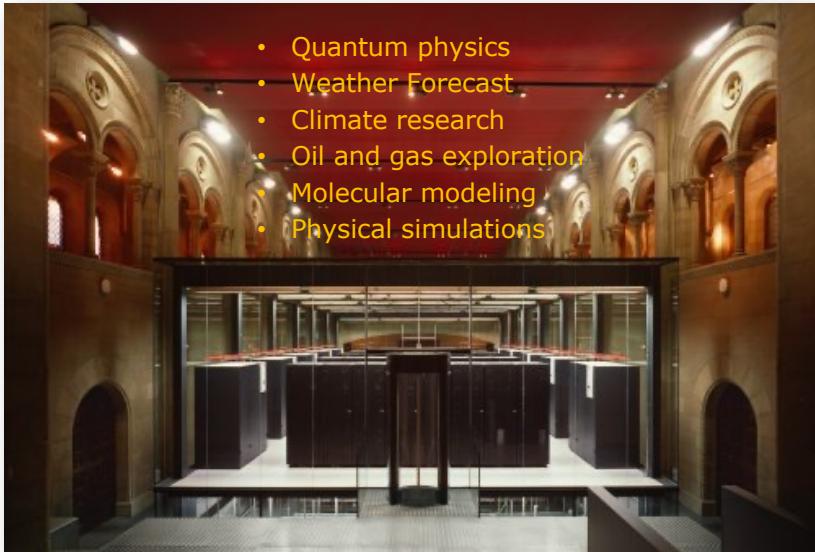
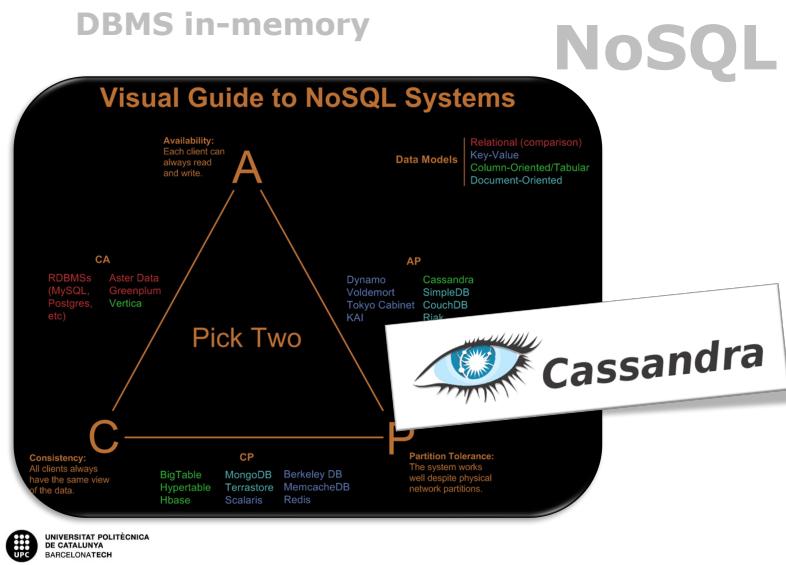
2013 International Data Corporation (IDC) Prediction for 2020:
40 zettabytes

2018 International Data Corporation (IDC) Prediction for 2025:
175 zettabytes



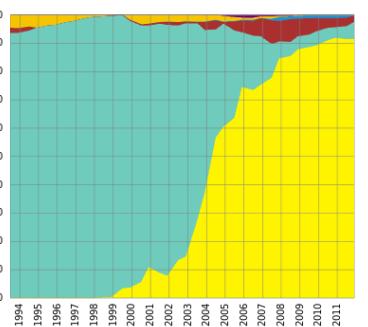
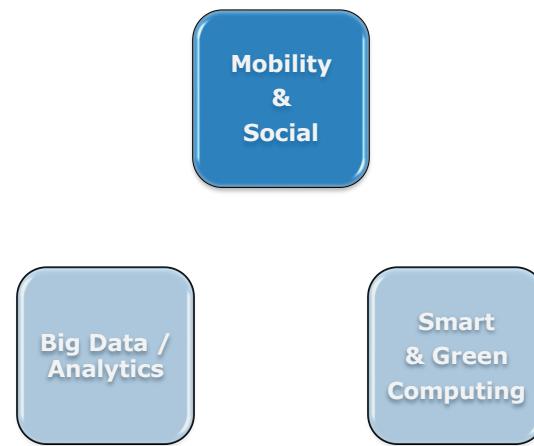
175 ZB in 2025

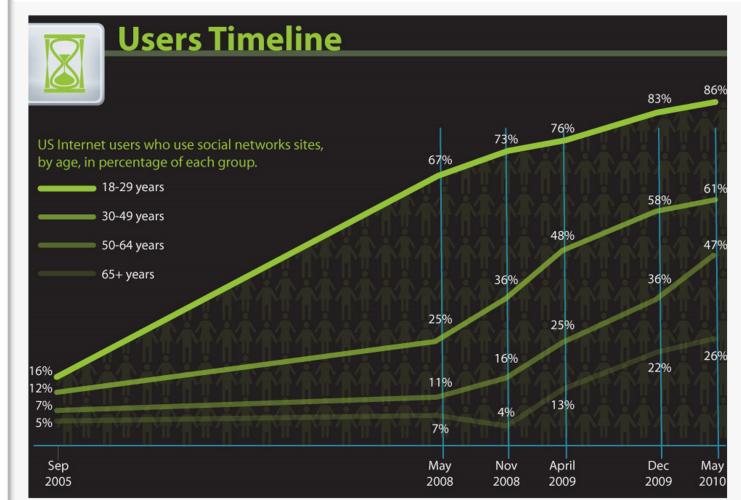
- 175 ZB onto BluRay discs: a stack 23 times the distance to the moon
- 90 ZB expected on IoT devices
- 49% stored in public cloud environments
- 30% consumed in real time

**High Performance Computing for Users**

- Parallel computing?
- Distributed computing?
- Grid computing?

Not for the final user!

**Supercomputer as a facility****Why big data centers?**

Social Networks

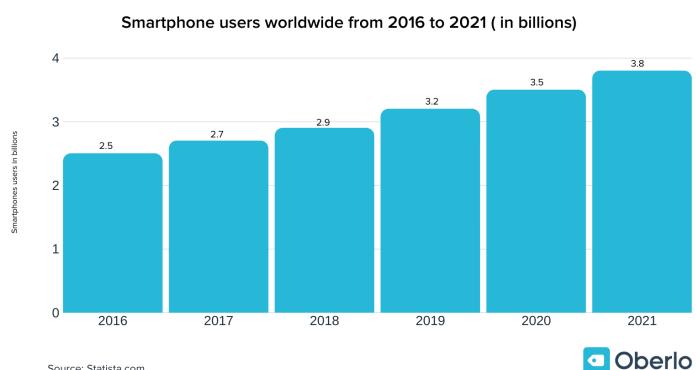
UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

1.4. MOBILITY AND SOCIAL

21

Some Internet facts

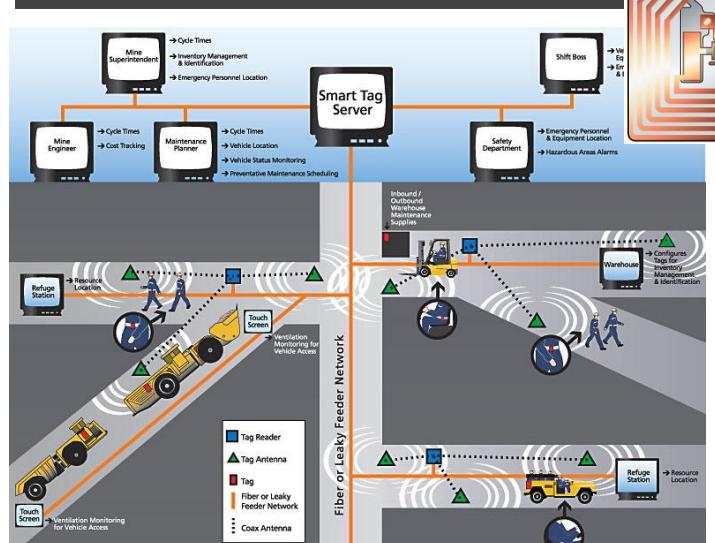
- Internet users: From 1.9 billion in 2015 to 4.57 billion (march 2020, source www.internetworldstats.com)
- Activity?

2020 This Is What Happens In An Internet Minute**Mobile devices**

UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

1.4. MOBILITY AND SOCIAL

23

RF tags (and so)

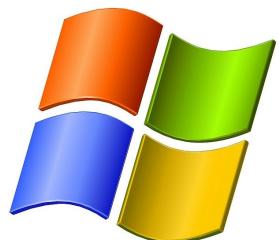
UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

So...



Servers in the world? June 2013. Source Steve Ballmer (Microsoft CEO) speech

- Facebook operates over 250,000 servers
- Amazon operates over 900,000 servers
- Microsoft operates about 1,000,000 servers
- Google operates over 1,000,000 servers
- Probably, there are about 400,000,000 physical servers in the world today



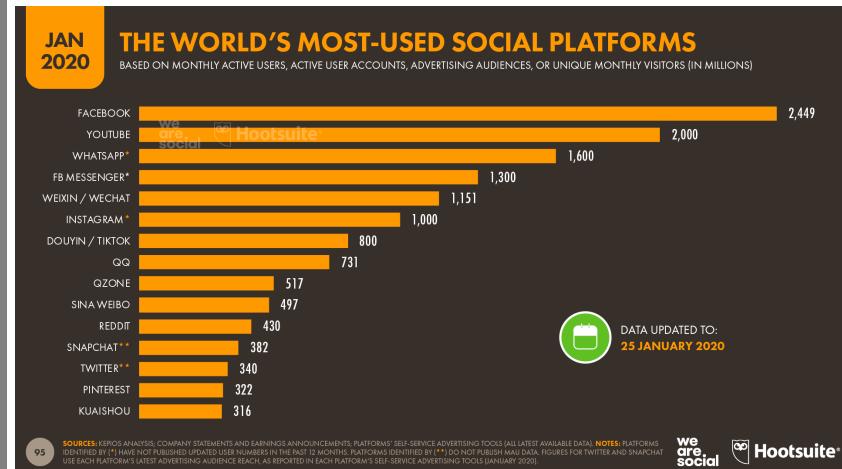
So...



Servers



Growing ratio?

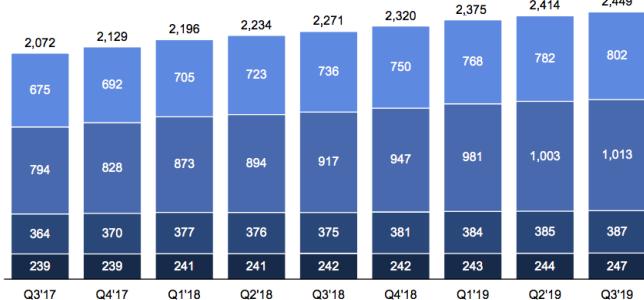


Growing ratio?

Monthly Active Users (MAUs)

In Millions

- Rest of World
- Asia-Pacific
- Europe
- US & Canada



Please see Facebook's most recent quarterly or annual report filed with the SEC for definitions of user activity used to determine the number of our DAUs and MAUs. The numbers for DAUs and MAUs do not include Instagram, WhatsApp, or Oculus users unless they would otherwise qualify as such users, respectively, based on their other activity. Beginning in Q3 2016, our MAU metrics reflect an update to our calculation methodology to exclude certain data signals that were previously misclassified as user account activity. This update resulted in the removal of a small percentage of accounts for Q3 2016. Excluding this update, MAUs in Q3 2016 would have been: Worldwide: 2,280 million; Rest of World: 1,740 million; Asia-Pacific: 921 million; Europe: 377 million; and US & Canada: 242 million. Periods prior to Q3 2016 have not been adjusted to reflect this updated methodology because the change was imminent.



1.4. MOBILITY AND SOCIAL

29

**Mobility
&
Social**

**Big Data /
Analytics**

**Smart
& Green
Computing**

Moving services to servers

Some services rely on servers

- E-mail
- Corporate web
- ...

Can be external centers with my machines?

- Opex (OPeration EXPenditures) reduction

Can be (directly) external servers?

- Opex & Capex (CAPital EXPenditures) reduction

1.5. SMART & GREEN COMPUTING

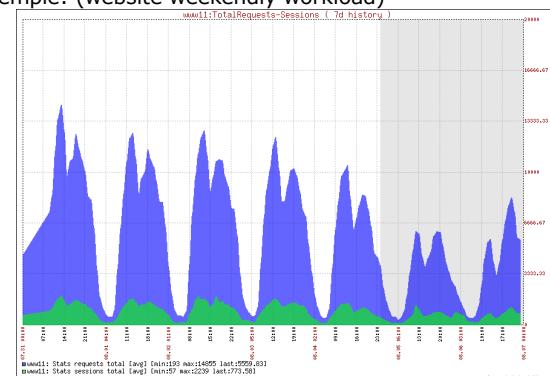
31

Why big data centers?

UPC
UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Moving services to servers

- The business community has begun to embrace Data Centers as a viable option
 - To reduce costs (opex & capex)
 - Scalability
 - Business agility
- Target Exemple: (website weekly load workload)

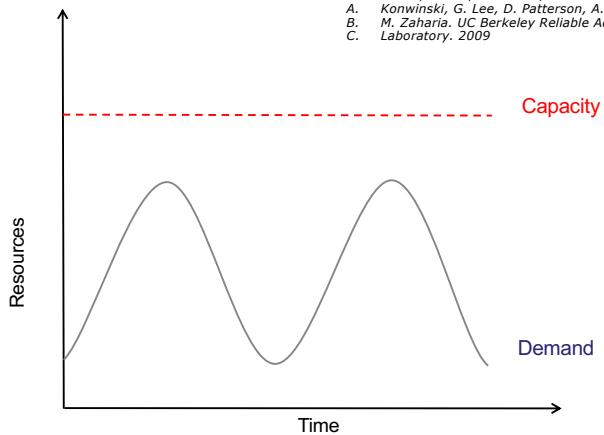


UPC
UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

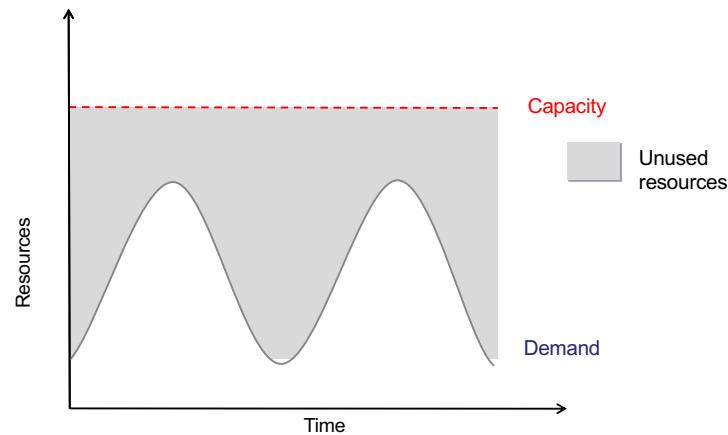
32

Capacity vs. demand

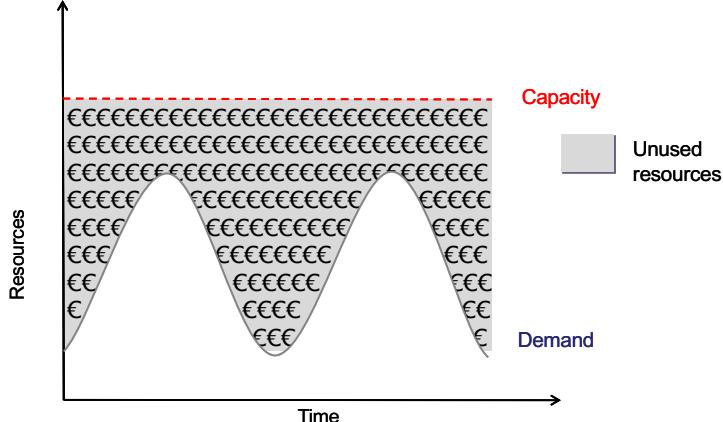
Source: *Above the Clouds: A Berkeley View of Cloud Computing*.
M. Ambrust, A. Fox, R. Griffith, A.D. Joseph, R. Katz,
A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica,
B. M. Zaharia. UC Berkeley Reliable Adaptive Distributed Systems
C. Laboratory. 2009



Lost of Resources

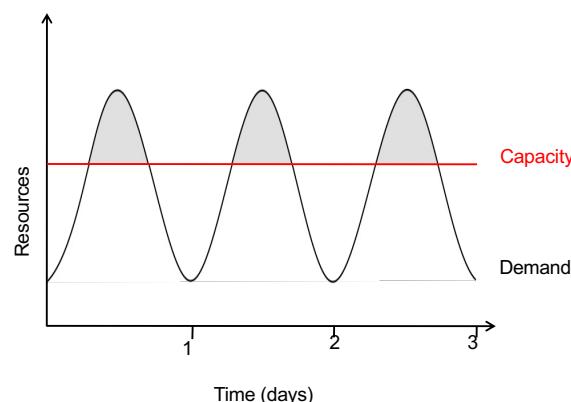


Overestimation means lost of money



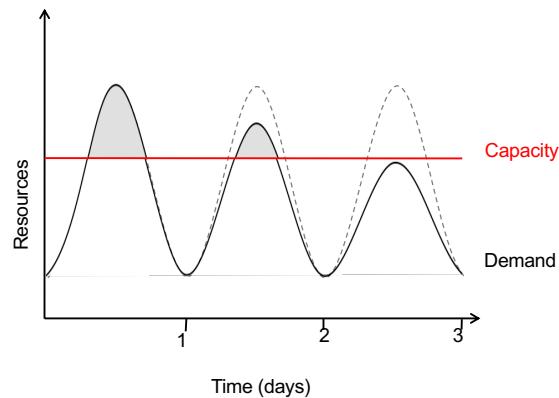
Underestimation is even worse!

- Can you measure users' dissatisfaction?

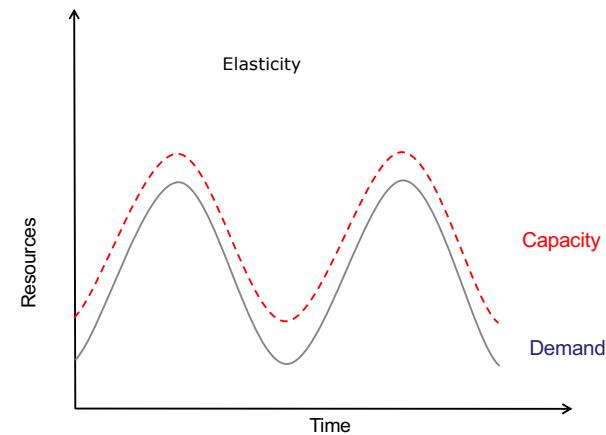


Clients are lost!

- Yes, you can!

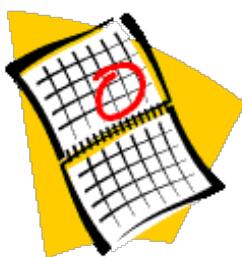


A perfect world...



Reasons for adoption

- Easy of management (no configuration or backup needed)
- Ubiquity of access (just a browser!)
- Easy to maintain (not upgrading lots of clients)
- Low cost (due to resources optimization)
 - And greener ☺
- Computing capacity (search services)
- Agility (hours for provision vs. weeks)



Is this *Cloud Computing*?

Yet another step

- Our data?
 - In the cloud
- Our applications?
 - In the cloud!
- High availability!
- From where you want!
- With whatever you want!
- Almost unlimited capacity!



Picture by Jordi Torres

1.6. IS THIS CLOUD COMPUTING?

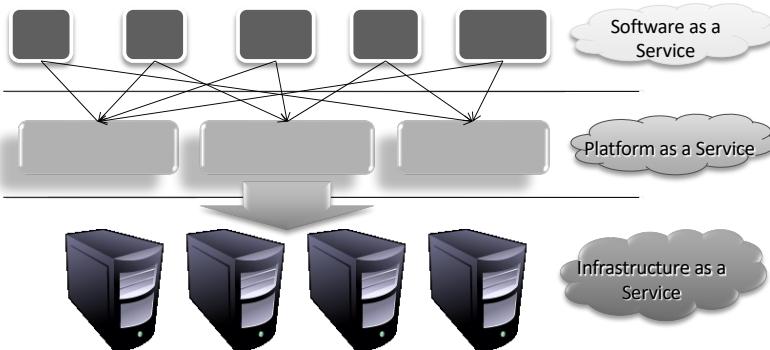
41

The name is not important. The concept is important

- Cloud computing offers the economies of scale of the utility model
 - Most organizations don't build their own power generators and water treatment plants
 - So why should every IT organization need to acquire and maintain all of the hardware, software, and development resources it uses?



The Cloud is offering SERVICES



Source: <http://www.slideshare.net/Cumulux/what-is-cloud-computing-presentation?nocache=4789>

1.7. OFFERING SERVICES

43

The Cloud is offering SERVICES

- According to NIST (National Institute of Standards and Technology) in document 800-146¹, cloud computing services can be classified as:
 - **Infrastructure as a Service (IaaS)**
 - **Platform as a Service (PaaS)**
 - **Software as a Service (SaaS)**

¹ csrc.nist.gov/publications/drafts/800-146/Draft-NIST-SP800-146.pdf

1.6. IS THIS CLOUD COMPUTING?

42

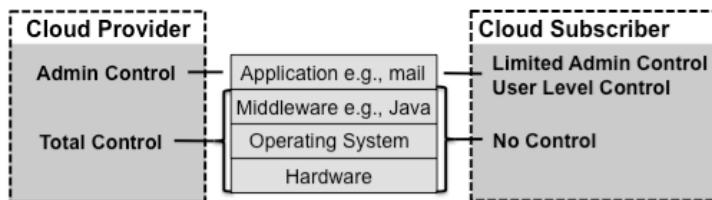
1.7. OFFERING SERVICES

44

Software as a Service (SaaS)

- NIST definition:

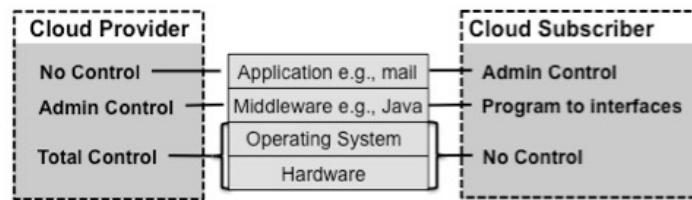
The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through a thin client interface such as a Web browser (e.g., Web-based email). The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, storage, or even individual application capabilities, with the possible exception of limited user-specific application configuration settings.



Platform as a Service (PaaS)

- NIST definition:

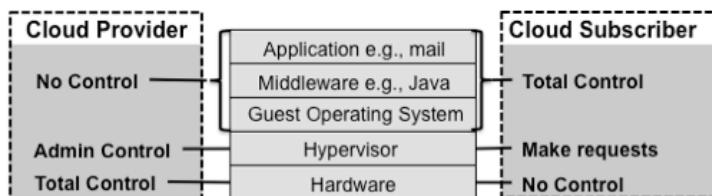
The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or -acquired applications created using programming languages and tools supported by the provider. The consumer does not manage or control the underlying cloud infrastructure including network, servers, operating systems, or storage, but has control over the deployed applications and possibly application hosting environment configurations.



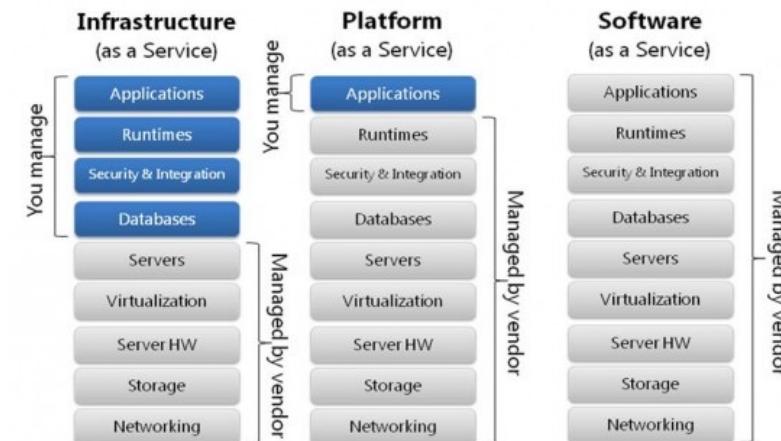
Infrastructure as a Service (IaaS)

- NIST definition:

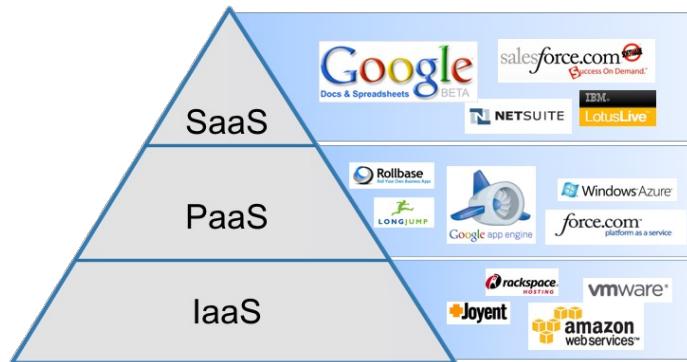
The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications. The consumer does not manage or control the underlying cloud infrastructure but has control over operating systems, storage, deployed applications, and possibly limited control of select networking components (e.g., host firewalls).



Differences



Some Examples



Risks of Cloud Computing

Availability

Privacy

Legislation

Data theft and Loss

...

→ Vendors are addressing these by offering:

- In continent / in country hosting
- **Private clouds**
- Secured IaaS (e.g. using VPN)
- ...

→ Customers are addressing these risks by:

- including cloud computing in backup, recovery and data protection (encryption).

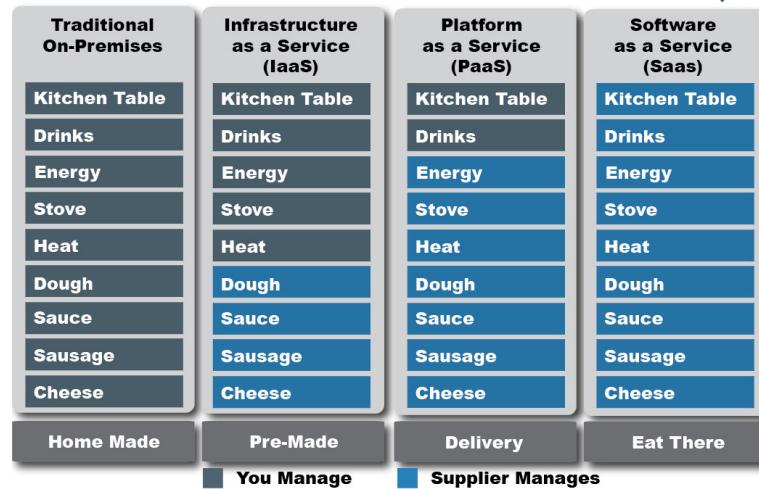


1.7. OFFERING SERVICES

49

Understanding IaaS, PaaS and SaaS ©

Pizza as a Service



1.7. OFFERING SERVICES

51

On-premises, colocation, virtual, hybrid DC

→ On premises:

- You own the hardware
- Maintain the software (updates, backup)
- Take care of faults, errors, software security
- Control building (access security, power, refrigeration, connection, fire protection, ...)

→ Colocation:

- The same, but you rent a space (including access security, power, connection, fire protection)

→ Virtual (Cloud):

- SaaS, PaaS, IaaS

→ Hybrid:

- Some on premises (or colocation) some virtual
- Advantages?

1.7. OFFERING SERVICES

50

1.8. SOME KEY CONCEPTS

52

Fault & Error (I)

Fault: some problem (hardware, software, bugs, cyberattack)

Error: unnoticed problem or an impossible to recover error

- Server with ECC RAM detects erroneous bit and corrects it before sending it to CPU
- Server with ECC RAM detects several erroneous bits and cannot correct them before serving the CPU
- Server without ECC RAM has an erroneous bit and the CPU reads it
- An Ethernet data packet has been received and an erroneous bit has been detected
- A magnetic disk cannot read a data block
- A ransomware attack took place and data has been encrypted

SLA, strategy, technology

SLA (Service Level Agreement): A type of agreement between companies on the level of service agreed between them (and the horrible consequences of not maintaining it).

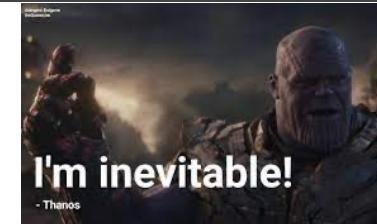
- Defines policy in FT, HA and DR

Strategies. Schemes to achieve FT, HA and DR characteristics in a design (e.g. geographic dispersion, redundancy). They are abstractions

Technology. Concrete solutions to implement strategies (RAID, Distributed File Systems, Copy-on-Write) are three technologies that offer disk fault-tolerance

Fault and Error (II)

Faults are inevitable



A company application cannot accept errors, so a company requires:

- **FT (Fault Tolerance):** Ability to continue to function error-free despite existing faults
- **HA (High Availability):** Ability to not stop services even with a large number of faults
- **DR (Disaster Recovery):** Ability to maintain service and not lose data even in large-scale disasters



LET'S BEGIN

David López