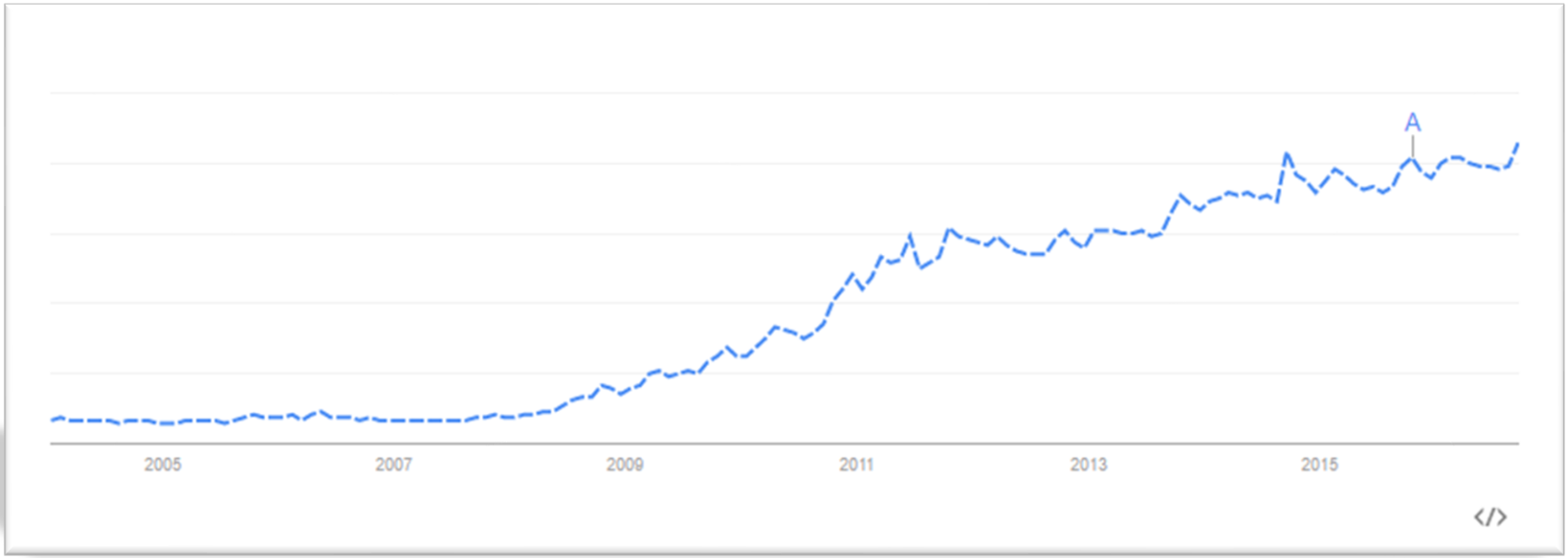
CLOUD COMPUTING

Goals

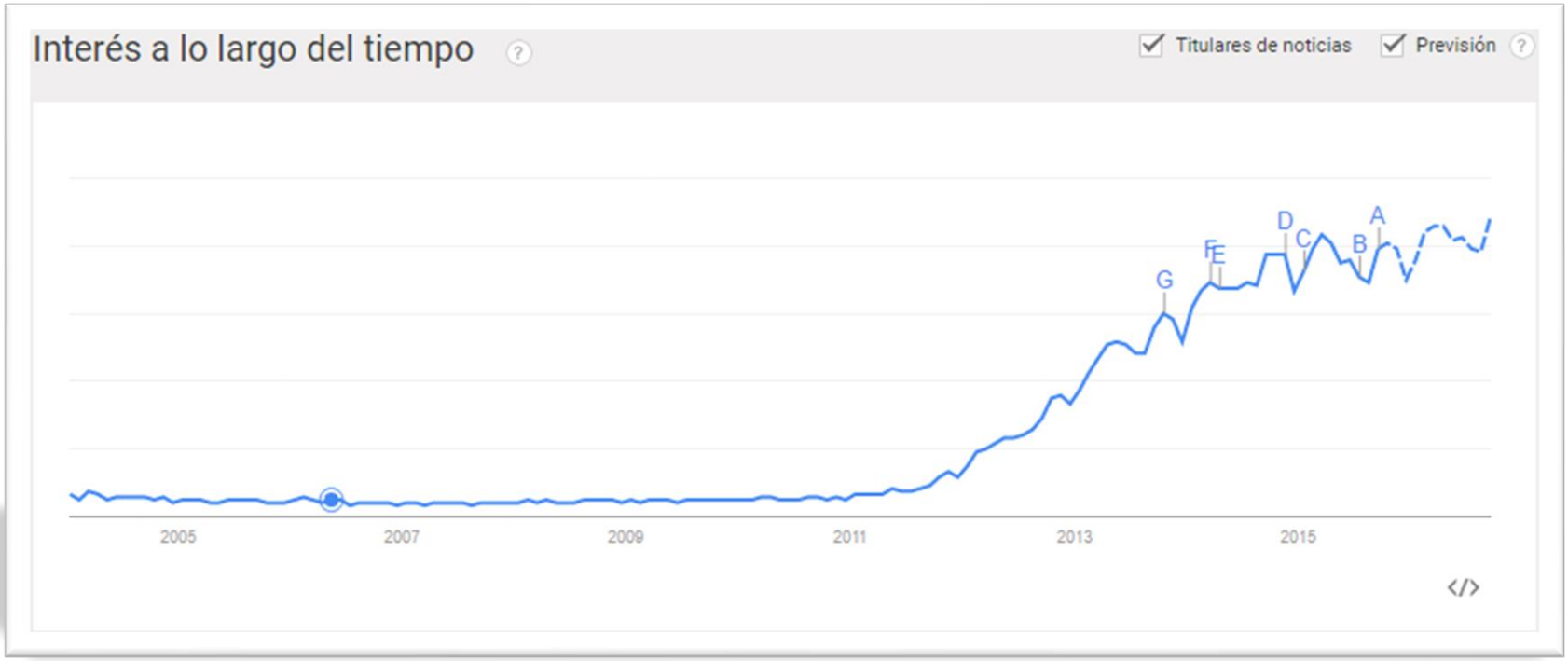
- Knowing a starting definition and characteristics of cloud computing.
- Knowing the evolution towards cloud computing, and its architecture.
- Knowledge of advanced technology for data processing.

Evolution of cloud computing



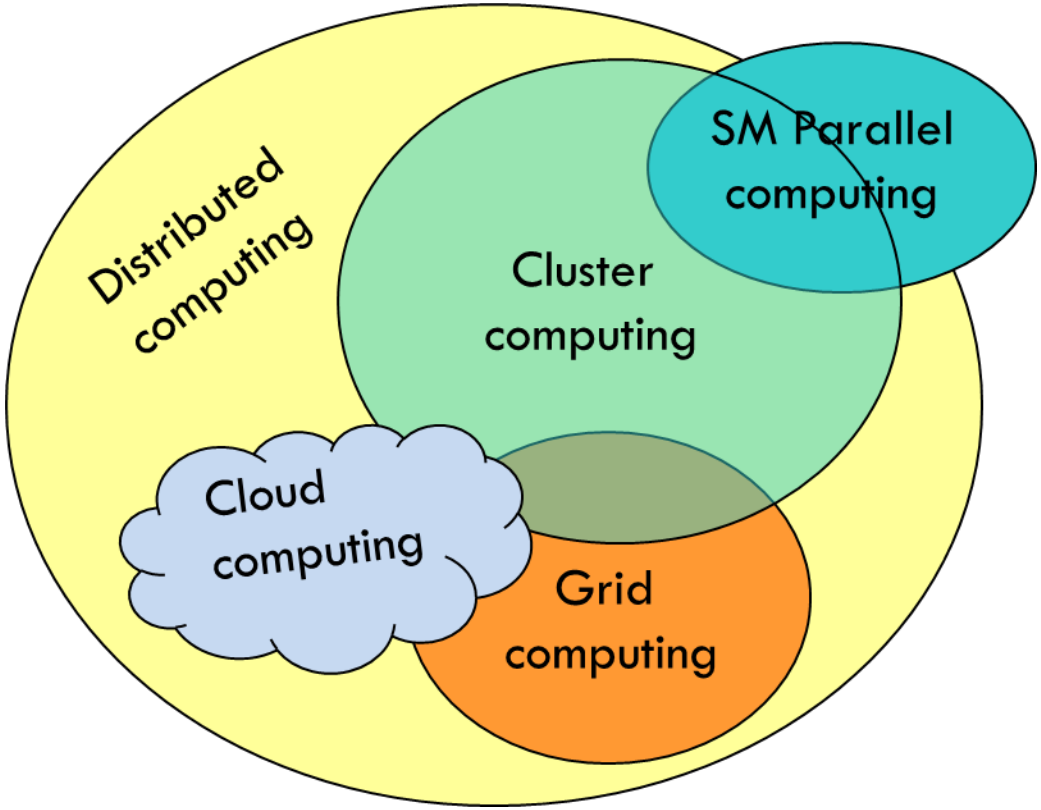
- Google trend can give us a clue about **cloud computing industry**: increasing interest

Just curiosity: big data on G.T.

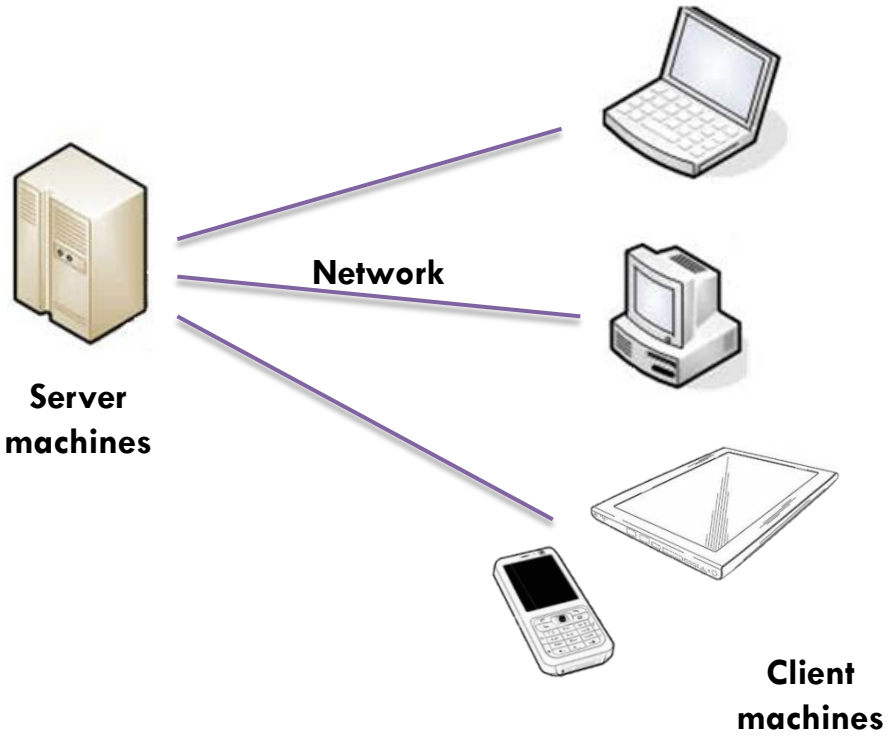


- Google trend on big data: increasing interest too

Scope

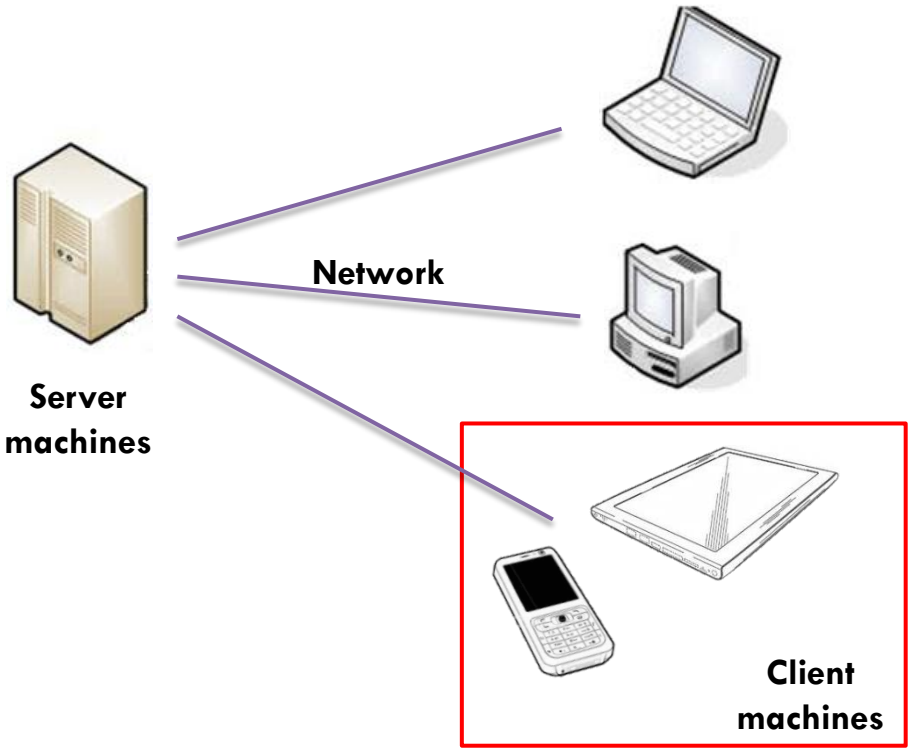

Distributed computing
Cluster computing
SM Parallel computing
Cloud computing
Grid computing

Scope



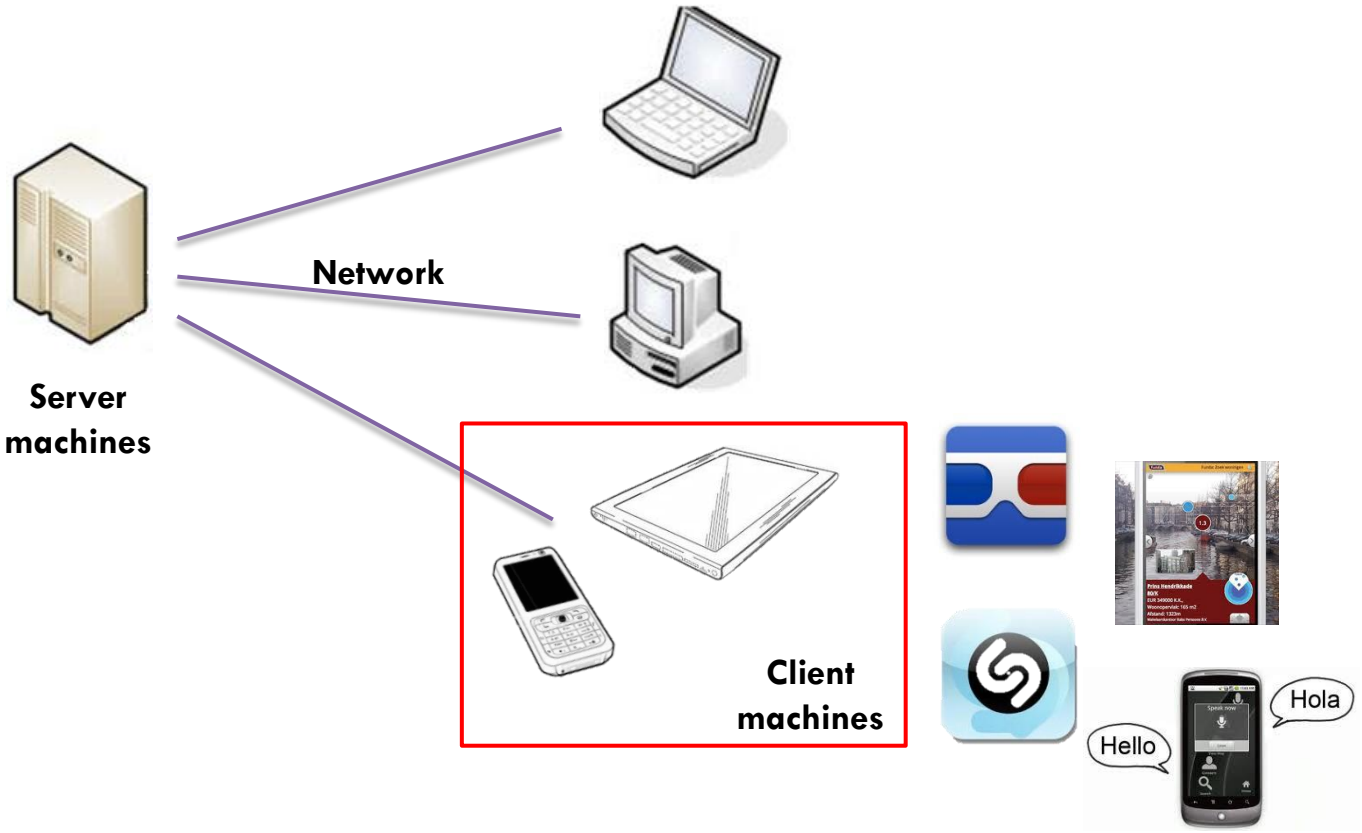
https://ardansirodjuddin.files.wordpress.com/2008/07/client_server_architecture.jpg

Scope



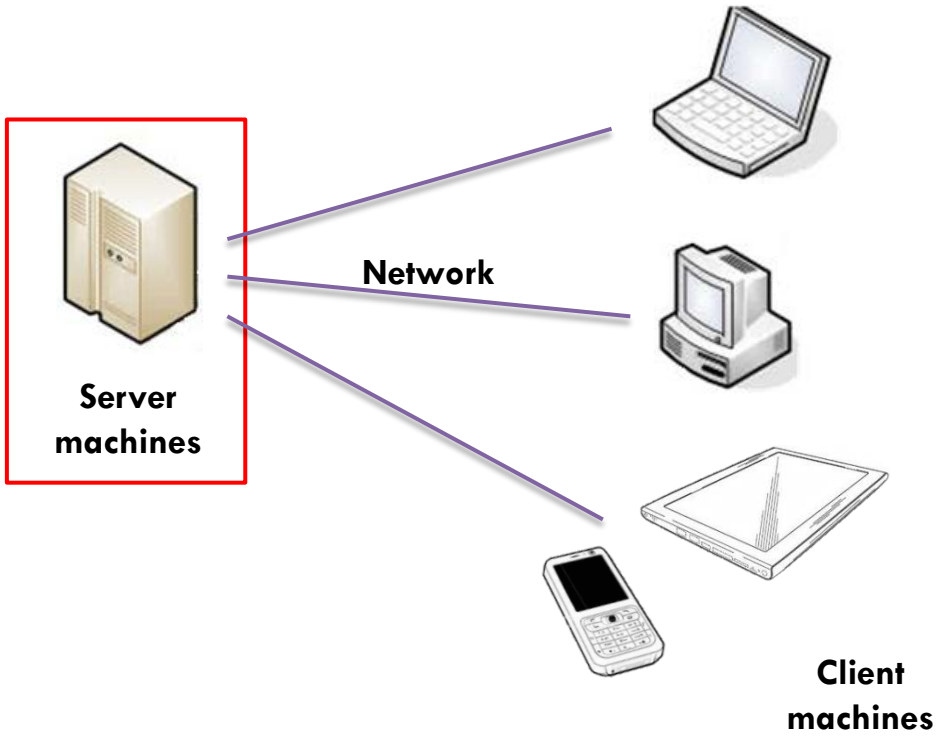
https://ardansirodjuddin.files.wordpress.com/2008/07/client_server_architecture.jpg

Scope



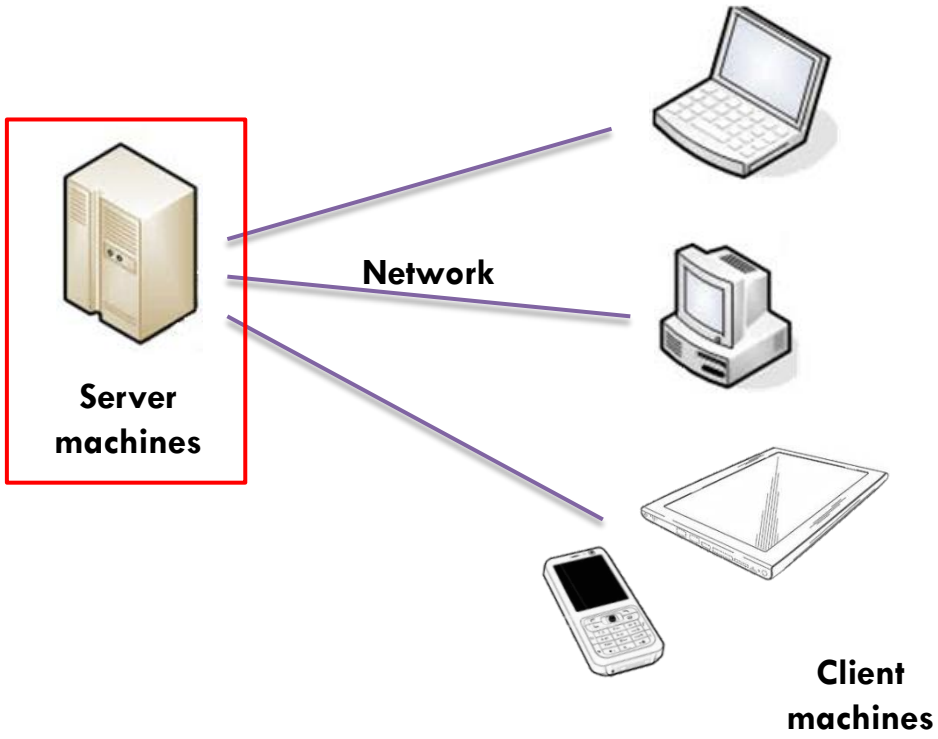
https://ardansirodjuddin.files.wordpress.com/2008/07/client_server_architecture.jpg

Scope



https://ardansirodjuddin.files.wordpress.com/2008/07/client_server_architecture.jpg

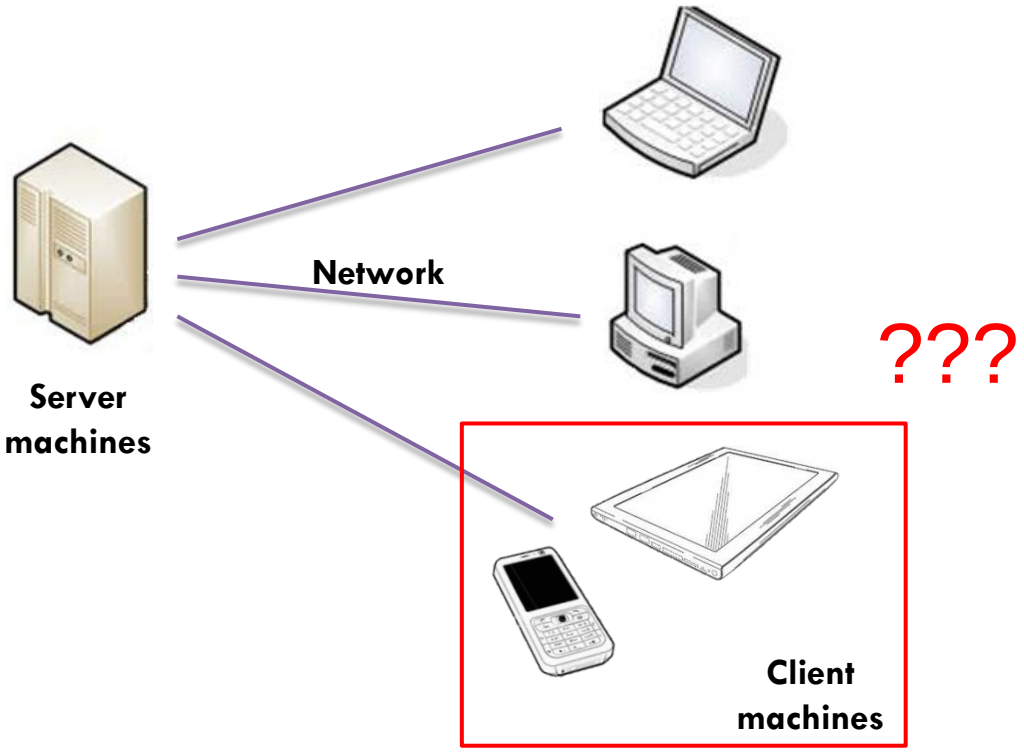
Scope



- Distributed supercomputing
- High throughput computing
- On-demand computing
- Data-intensive computing
- Collaborative computing
- Multimedia computing

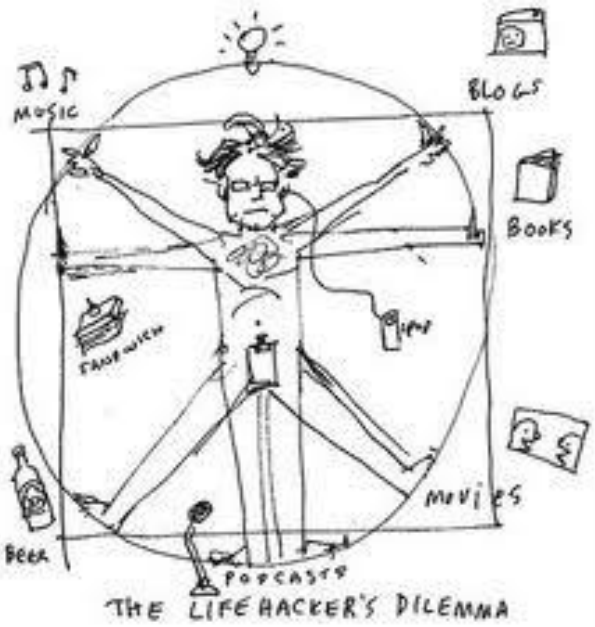
https://ardansirodjuddin.files.wordpress.com/2008/07/client_server_architecture.jpg

Client behaviour



https://ardansirodjuddin.files.wordpress.com/2008/07/client_server_architecture.jpg

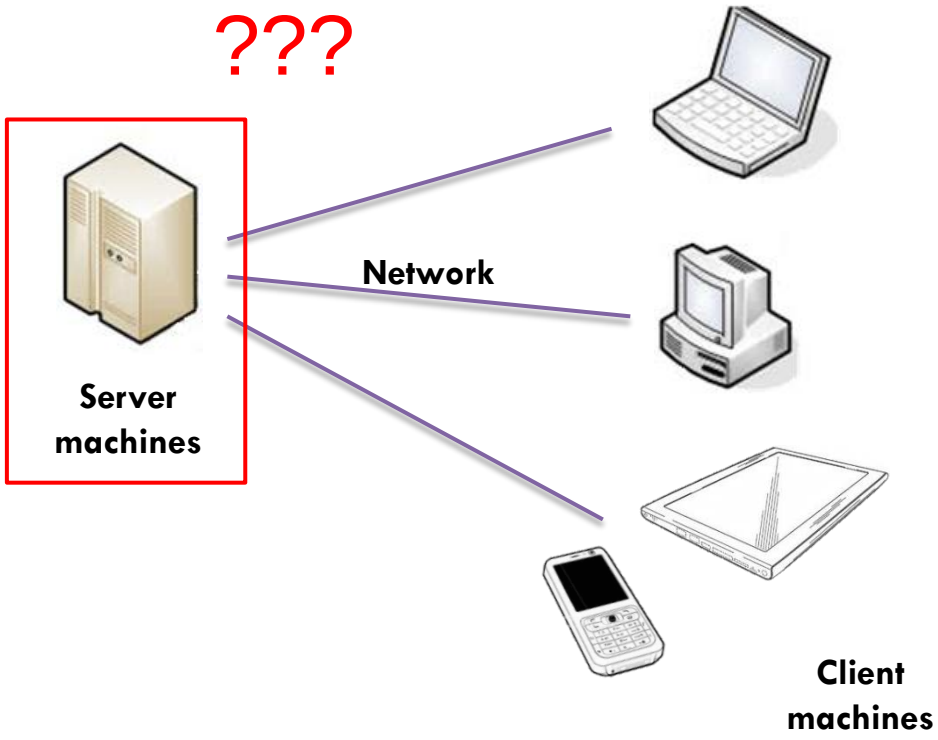
"vitruvio client"… (0 knowledge)



- Pareto principle or **80-20 rule**:
  - For many events, near 80% of the effects comes from 20% of the causes.
  - Example (business):
    - The 80% of incomes comes from 20% of clients
    - The 80% of sales comes from 20% of the products
- Zipf law, 1% rule, etc.

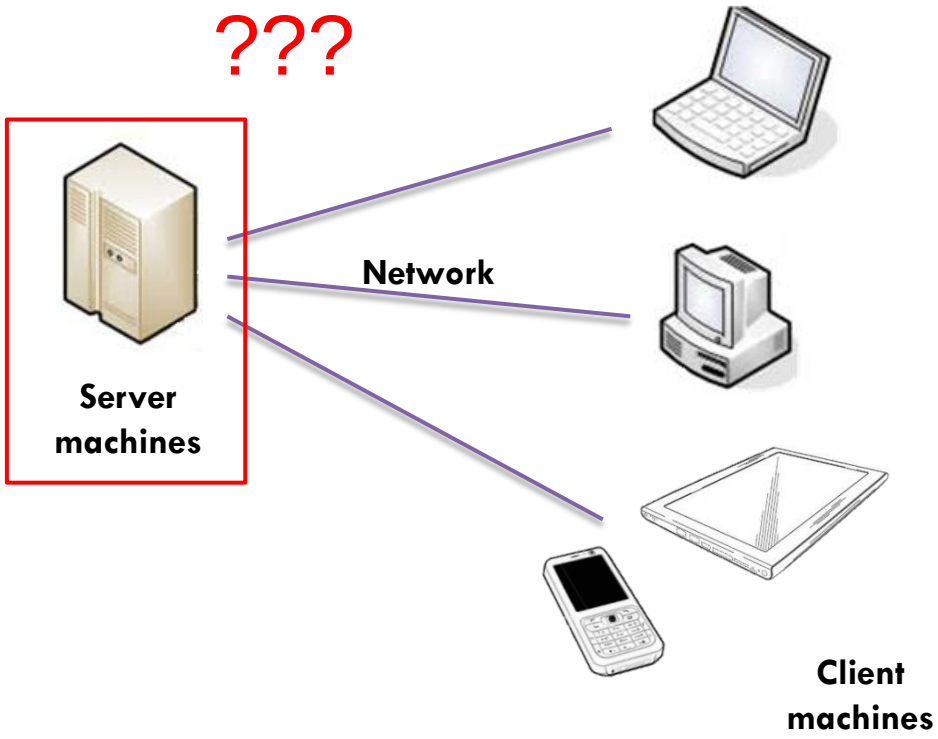
http://chistesgeeks.com/2008/10/25/hombre-de-vitruvio-geek/

Server platforms



https://ardansirodjuddin.files.wordpress.com/2008/07/client_server_architecture.jpg

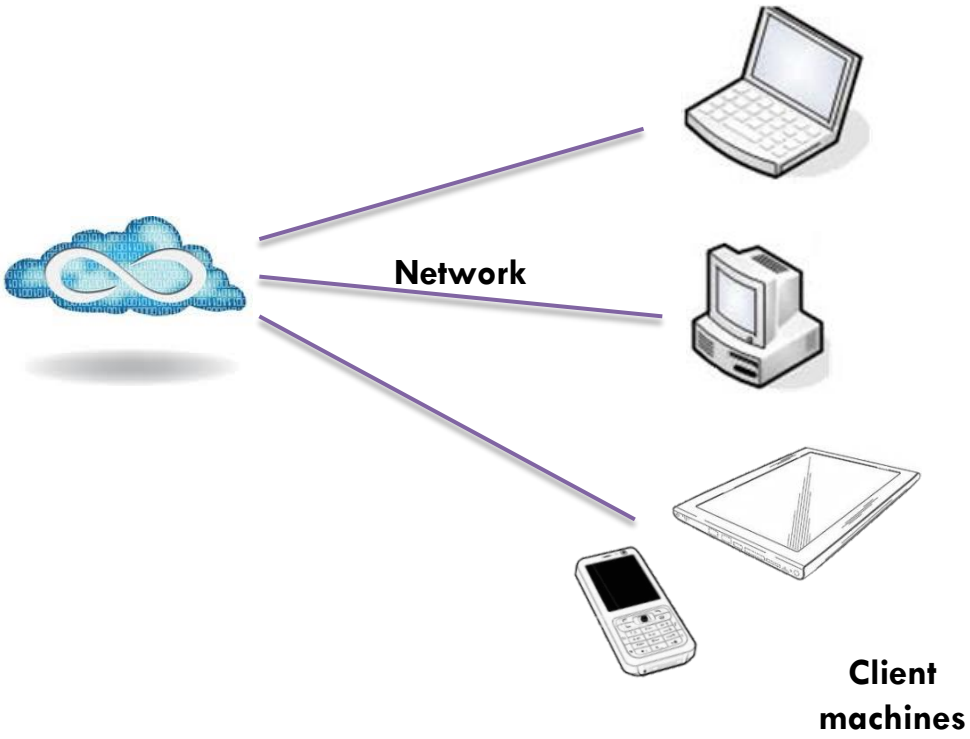
Server platforms



- Mainframes and supercomputers
- Clusters
- Grid computing
- Volunteer computing
- Cloud computing

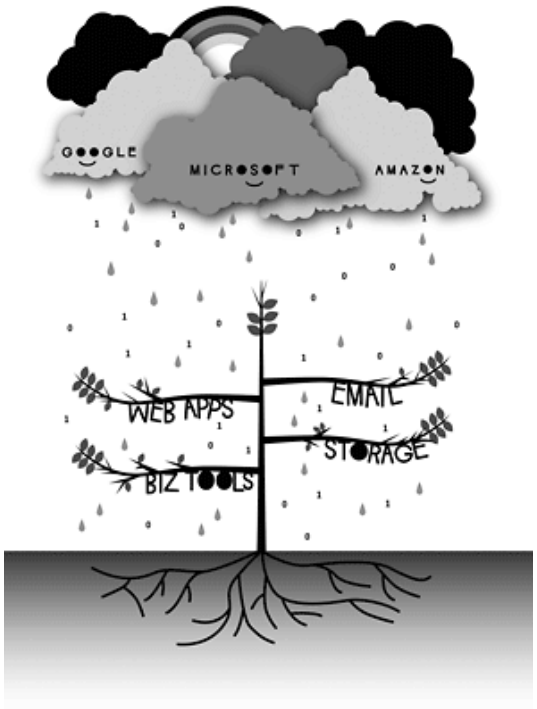
https://ardansirodjuddin.files.wordpress.com/2008/07/client_server_architecture.jpg

Scope



- Mainframes and supercomputers
- Clusters
- Grid computing
- Volunteer computing
- **Cloud computing**

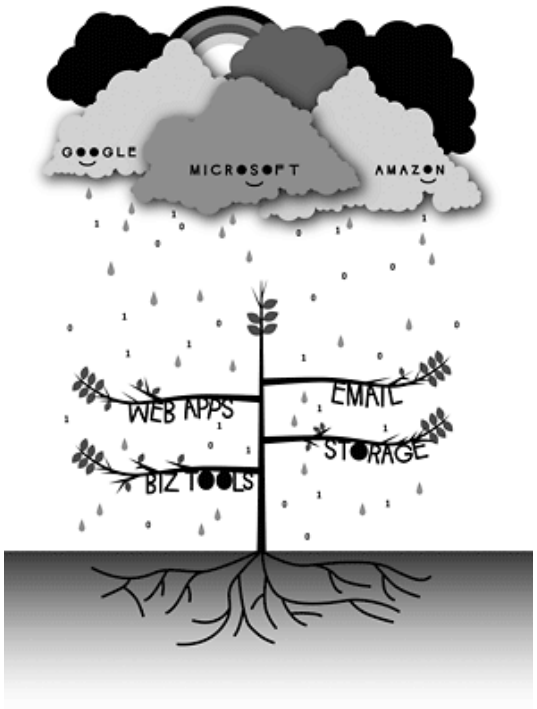
Case of use: Amazon



- Amazon design it data centers for being able to support Christmas sales.
- Problem: the rest of the time the system usage is around 10%

http://blogs.msdn.com/jim_glass/archive/2008/10/27/microsoft-s-cloud-computing-strategy.aspx

Case of use: Amazon



- Amazon design it data centers for being able to support Christmas sales.
- Problem: the rest of the time the system usage is around 10%
- Idea: to rent the computing oversizing to other companies to recover as much possible of the investment.
  - Utility computing

http://blogs.msdn.com/jim_glass/archive/2008/10/27/microsoft-s-cloud-computing-strategy.aspx

Utility computing…

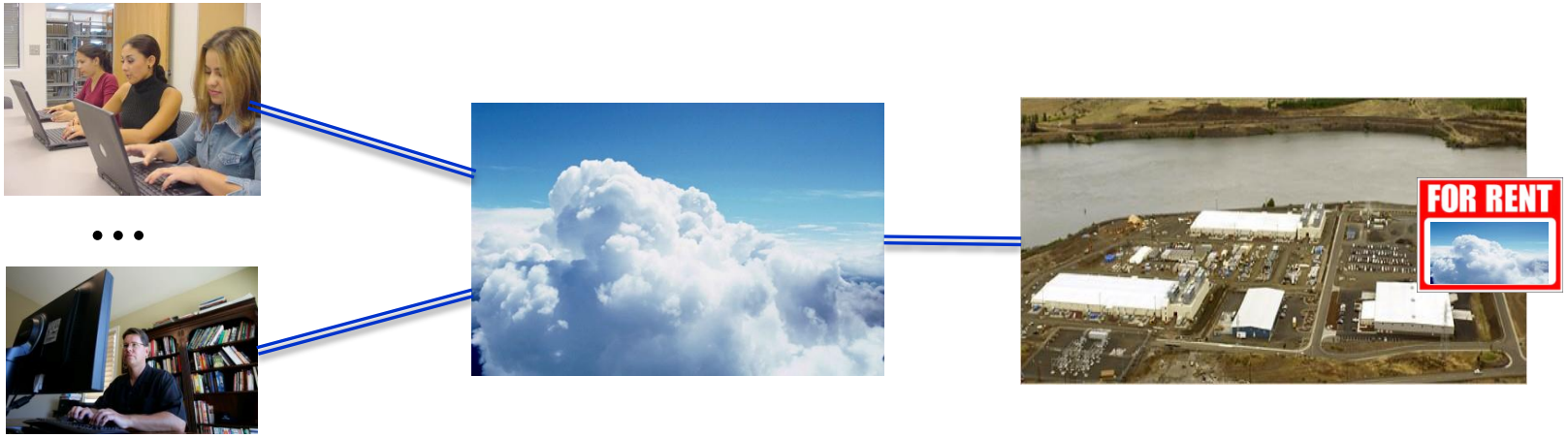
Windows Azure

amazon web services™

Google App Engine

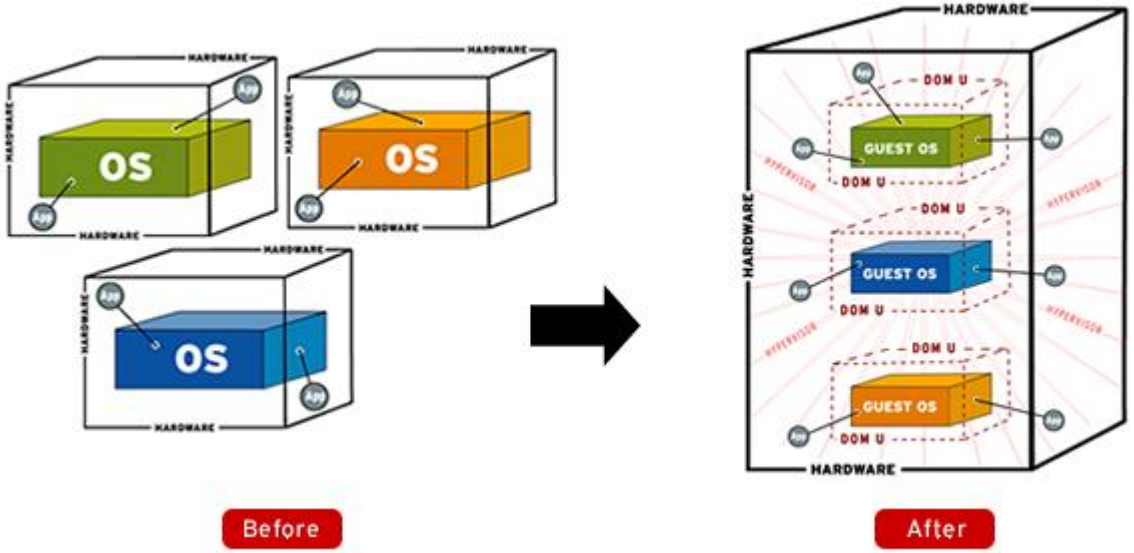
- Companies with extra computing capacity can rent its systems to clients.

Utility computing…



- Companies with extra computing capacity can rent its systems to clients.
- Companies with computing capacity demands can rent others companies resources with an interesting service/price relationship.
  - No pay for building big data-centers.
  - No pay for big data-center maintenance.
  - No pay for big data-center electrical bills.

Virtualization…



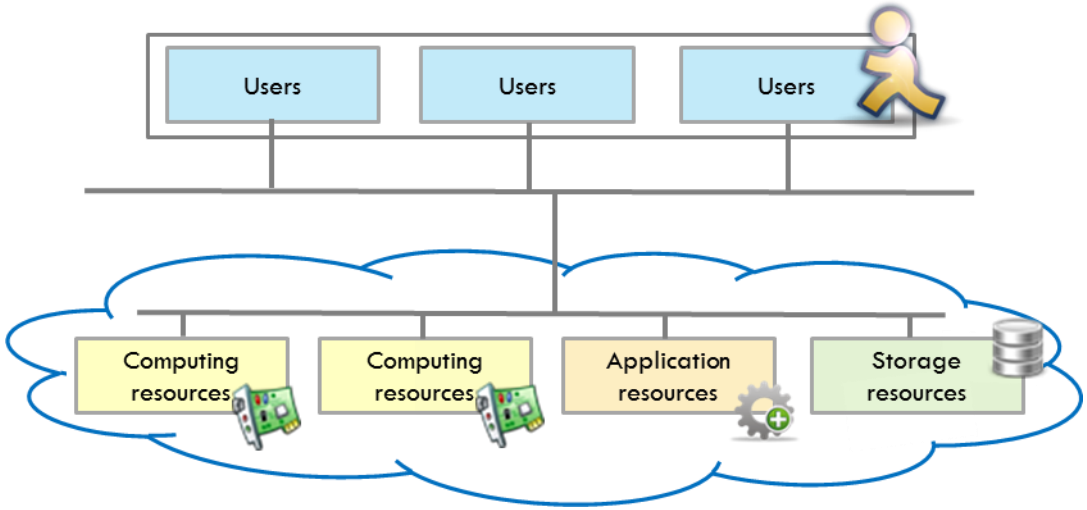
- Problem: different companies could have different environments
- Idea: to have a virtual computer (guest) on a computer (host) that could offer different environments with "few" penalties.
  - Virtualization

Virtualization…

- Problem: different companies could have different environments
- Idea: to have a virtual computer (guest) on a computer (host) that could offer different environments with "few" penalties.
  - Virtualization

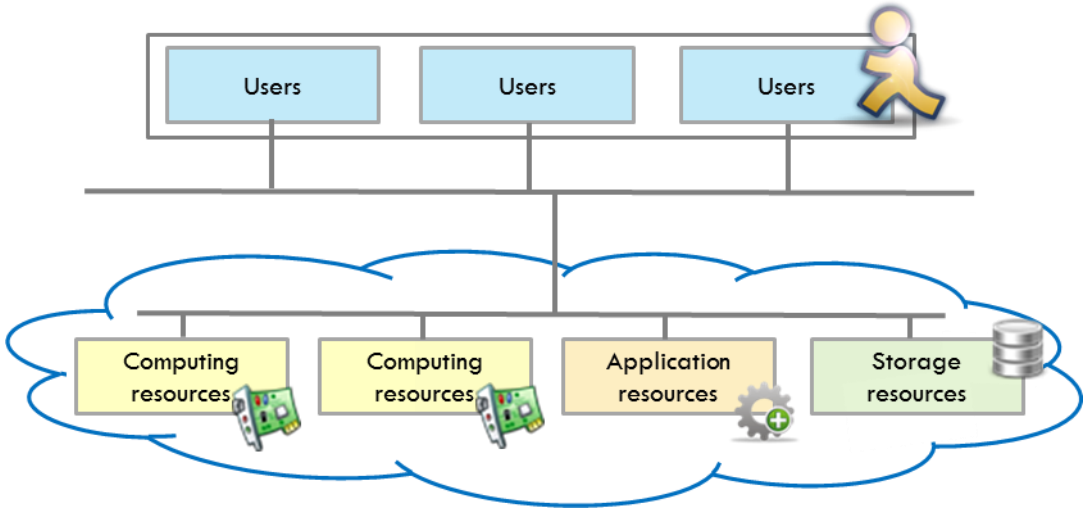
Cloud: Initial definition…

- Cloud computing can be defined in a rough way as an scalable usage of computational resources offered as a service outside the environment that use them, through a pay per use.



Cloud: Initial definition…

- Cloud computing can be defined in a rough way as an scalable usage of computational resources offered as a service outside the environment that use them, through a pay per use.



- Two main ingredients: utility computing and virtualization

Advantage and disadvantage…

- Main advantages:
  - Only used the resources needed, and the cost is tied to the use of resources (reduce extra cost)
  - Cloud resources can be accessed in any time, from any place of Internet.
- Main disadvantages:
  - Security and privacy warranties from hosts
    - E.g.: ad. based on the e-mail content.
    - E.g.: If sensible information is used…
  - High-speed networking is needed
    - E.g.: Networking-tied applications.
  - Among cloud systems: interoperability
    - E.g.: Divide on multiple cloud.

Advantage and disadvantage…

- Main advantages:
  - Only used the resources needed, and the cost is tied to the use of resources (reduce extra cost)
  - Cloud resources can be accessed in any time, from any place of Internet.
- Main disadvantages:
  - Security and privacy warranties from hosts
    - E.g.: ad. based on the e-mail content.
    - E.g.: If sensible information is used…
  - High-speed networking is needed
    - E.g.: Networking-tied applications.
  - Among cloud systems: interoperability
    - E.g.: Divide on multiple cloud.

Peter Deutsch

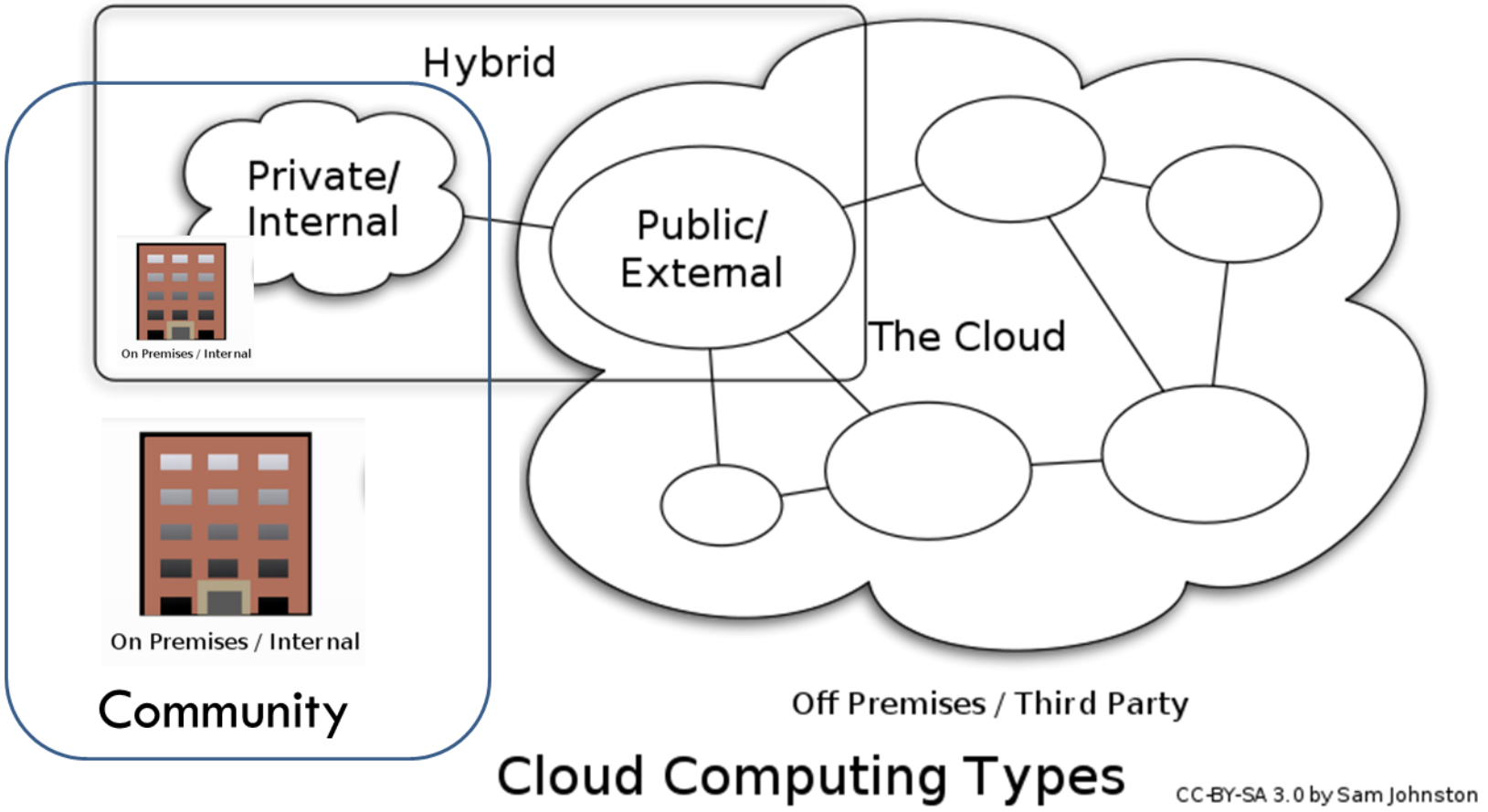
1. The **network** is reliable.
2. **Latency** is zero.
3. **Bandwidth** is infinite.
4. The **network** is secure.
5. **Topology** doesn't change.
6. There is one **administrator.**
7. **Transport cost** is zero.

James Gosling

8. The network is homogeneous.

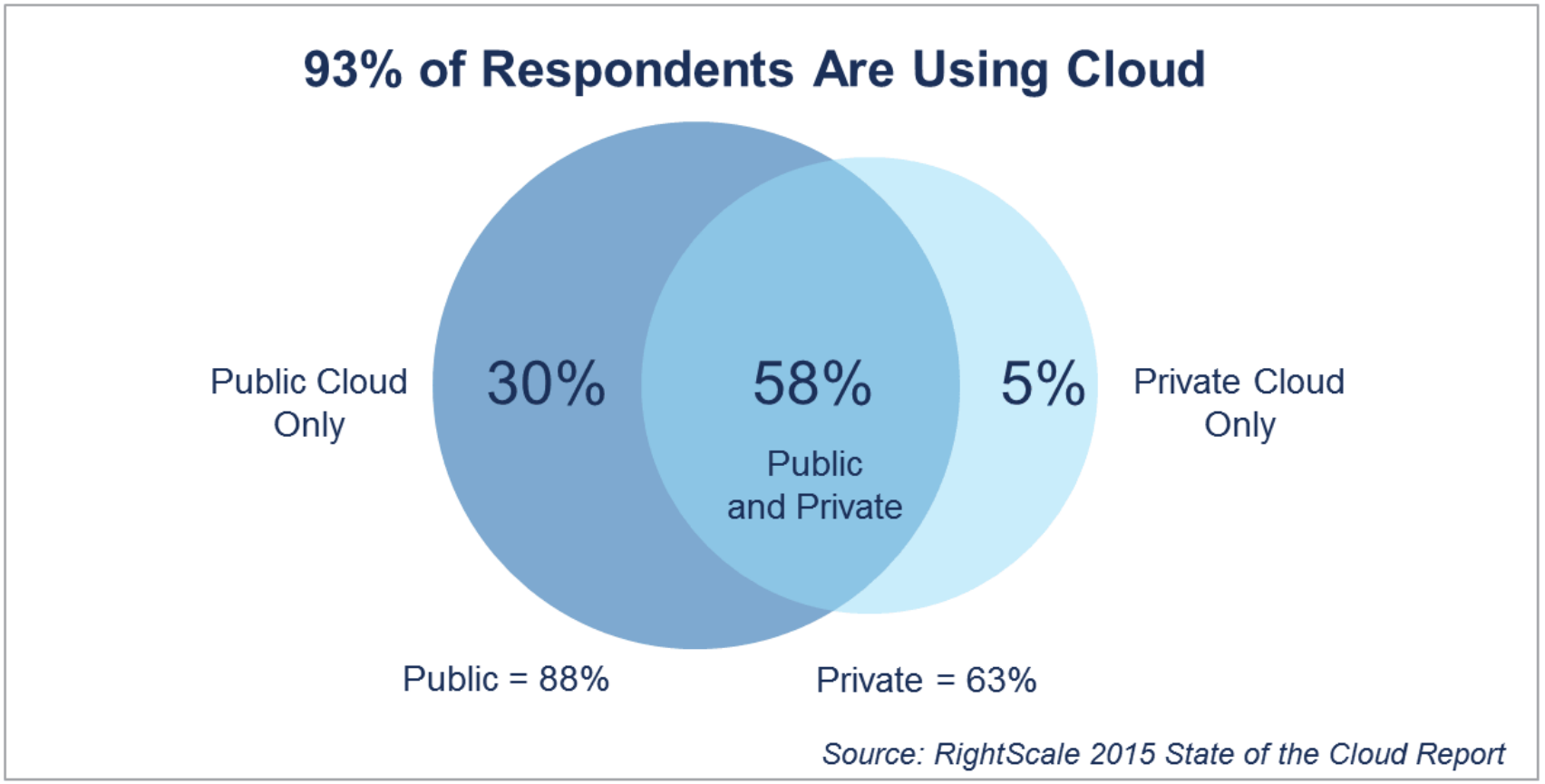
http://www.artifactconsulting.com/lapeira/2011/02/28/motivos-para-rechazar-el-cloud-computing/

Public, private, hybrid…



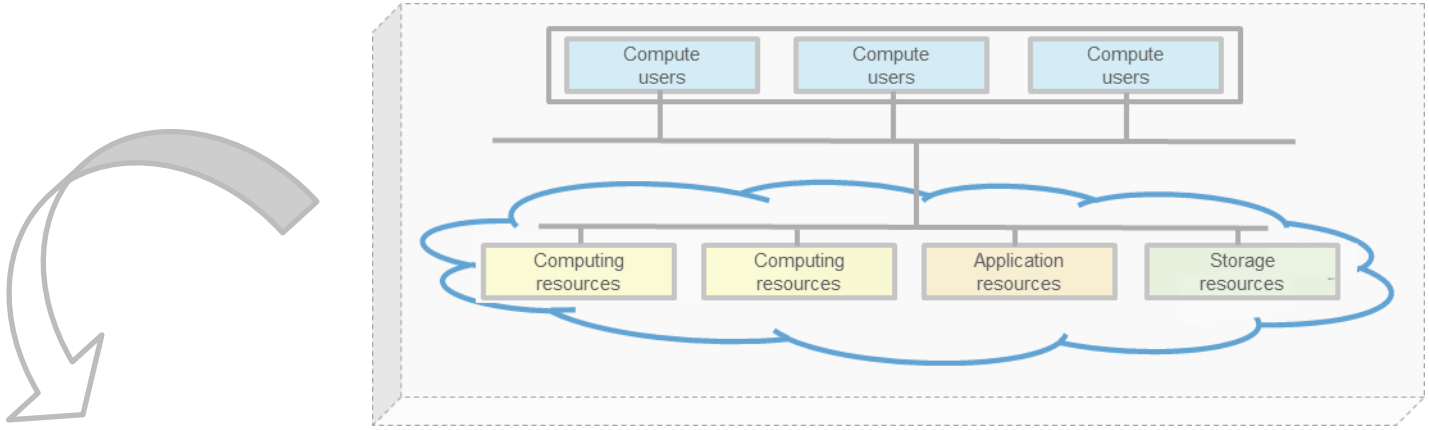
http://en.wikipedia.org/wiki/File:Cloud_computing_types.svg

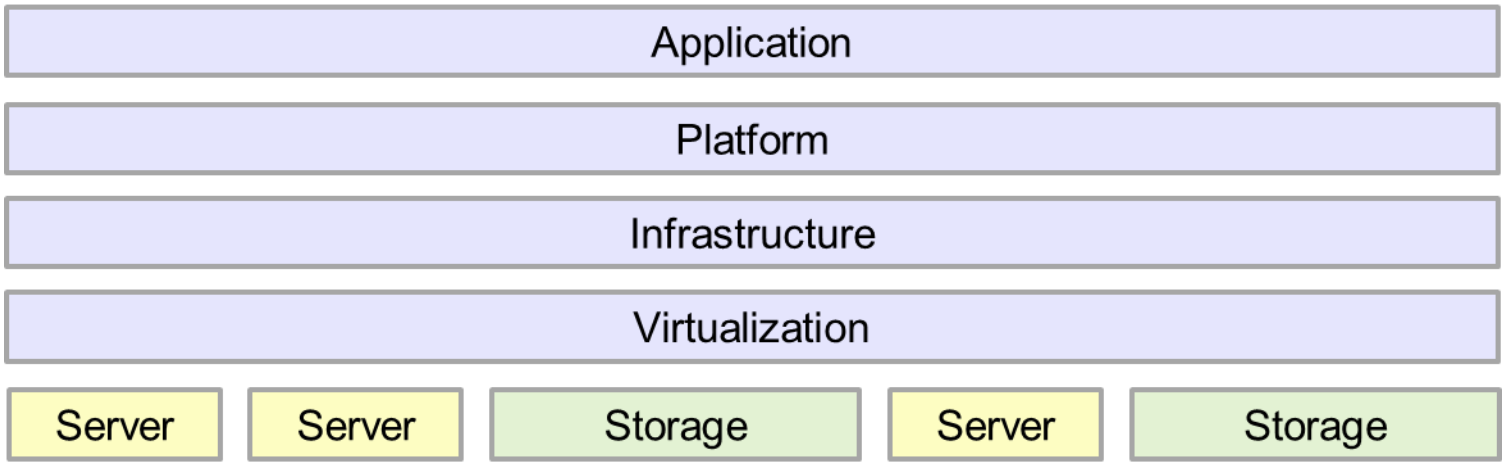
Public, private, hybrid…



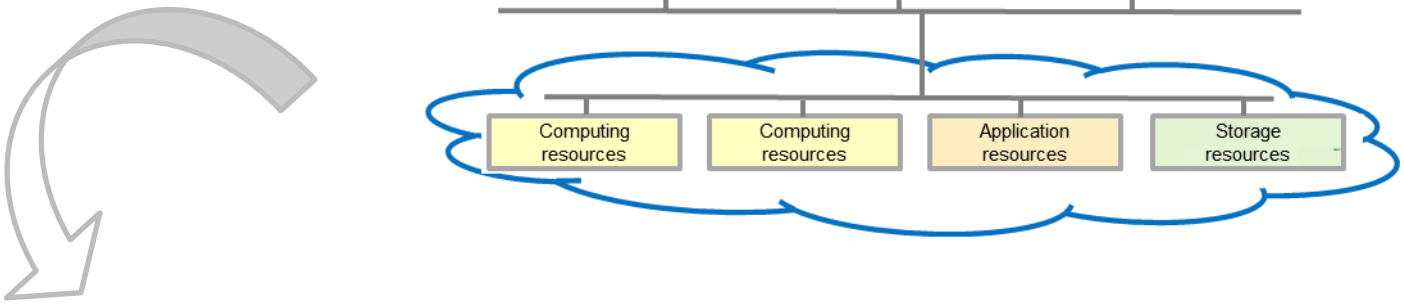
http://www.rightscale.com/blog/cloud-industry-insights/cloud-computing-trends-2015-state-cloud-survey

Anatomy of a Cloud



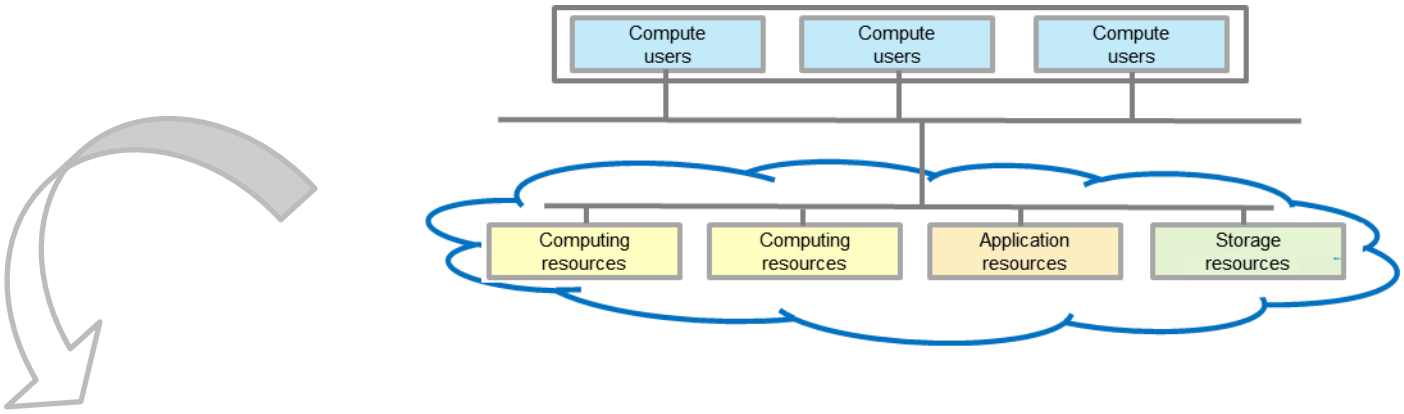


Anatomy of a Cloud



SaaS	Software as a service	Google Apps, Microsoft "Software+Services"
PaaS	Platform as a service	Google AppEngine, IBM IT Factory, Force.com
IaaS	Infrastructure as a service	Amazon EC2, Sun Grid, IBM Blue Cloud
dSaaS	Storage as a service	Amazon S3, Nirvanix SDN, Cleversafe dsNet

Anatomy of a Cloud



SaaS	Software as a service	Google Apps, Microsoft "Software+Services"
PaaS	Platform as a service	Google AppEngine, IBM IT Factory, Force.com
IaaS	Infrastructure as a service	Amazon EC2, Sun Grid, IBM Blue Cloud
dSaaS	Storage as a service	Amazon S3, Nirvanix SDN, Cleversafe dsNet

- High availability and simple storage services
- Costs by transfers and capacity needed

Cloud: dSaaS

	Amazon S3		
	Standard	Reduced Redundancy	Glacier
First 1 TB	0.095	0.076	0.01
Next 9 TB (1-10 TB)	0.08	0.064	0.01
Next 40 TB (10-50 TB)	0.08	0.064	0.01
Next 50 TB (50-100 TB)	0.07	0.056	0.01
Next 400 TB (100-500 TB)	0.07	0.056	0.01
Next 500 TB (500 TB-1000 TB)	0.065	0.052	0.01
Next 4000 TB (1000 TB - 5000 TB)	0.06	0.048	0.01
More than 5000 TB	0.055	0.037	0.01

	Google Cloud Storage	
	Standard	Durable Reduced Availability
First 1 TB	0.085	0.063
Next 9 TB (1-10 TB)	0.076	0.054
Next 40 TB (10-50 TB)	0.067	0.049
Next 50 TB (50-100 TB)	0.067	0.049
Next 400 TB (100-500 TB)	0.063	0.045
Next 500 TB (500 TB-1000 TB)	0.054	0.042
Next 4000 TB (1000 TB - 5000 TB)	0.054	0.042
More than 5000 TB	N/A [2]	N/A [2]

	Microsoft Windows Azure	
	Geographically Redundant	Locally Redundant
First 1 TB	0.095	0.07
Next 9 TB (1-10 TB)	0.08	0.065
Next 40 TB (10-50 TB)	0.08	0.065
Next 50 TB (50-100 TB)	0.07	0.06
Next 400 TB (100-500 TB)	0.07	0.06
Next 500 TB (500 TB-1000 TB)	0.065	0.055
Next 4000 TB (1000 TB - 5000 TB)	0.06	0.045
More than 5000 TB	0.055 [1]	0.037 [1]

Data as of January 22, 2013. Sources: Amazon, Microsoft, Google. (c) Tom's IT Pro

1 Pricing for 5000 TB - 9000 TB. Pricing beyond 9000 TB provided upon inquiry.

2 Pricing provided upon inquiry.

Amazon S3 Definitions:

Glacier: 99.999999999% durability. Designed for long term storage. Up to 5% of average monthly storage (pro-rated daily) can be retrieved free of charge. Retrieval fees start at $0.01 per gigabyte.

Reduced Redundancy: 99.99 percent durability.

Standard: Geographically redundant. 99.999999999% durability.

Windows Azure Definitions:

Geographically Redundant: Data is stored in two locations hundreds of miles apart within the same region.

Locally Redundant: Data is replicated three times within the same data center.

Google Cloud Storage Definitions:

Standard: Geographically redundant, 99.9 percent or better availability

Durable Reduced Availability: Currently in trial period, expected durability of 99 percent.

http://www.tomsitpro.com/articles/cloud_computing-it_storage-azure-amazon_s3-google_cloud_storage,2-528-2.html

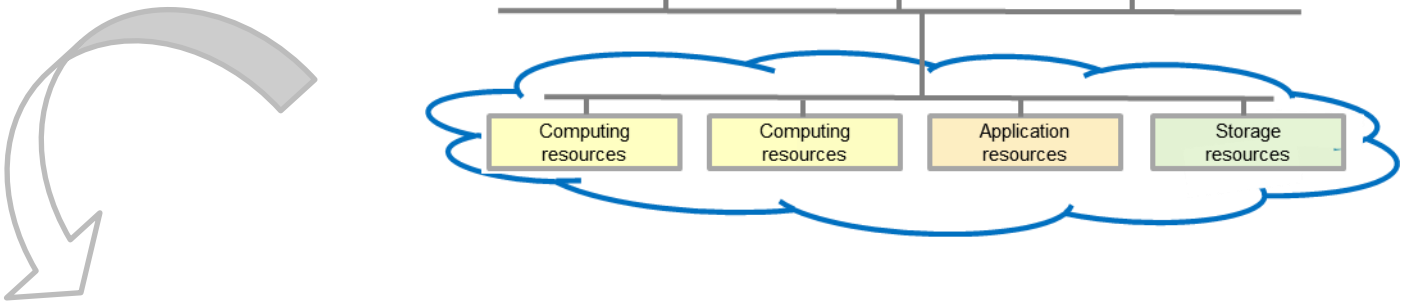
Cloud: dSaaS

- The S3 service bills by three concepts altogether:
  - Data stored:
    - Fare by GiB stored/month.
  - Data transfer:
    - Fare by GiB transferred.
  - Access requests:
    - Fare by object/file requests (GET, PUT, LIST, etc.).

http://aws.amazon.com/s3/

- Regardless the fares, Amazon services are competitive:
  - Around 2,5 GiB of data stored and 15GiB transferred in a month, will not reach 4 dollars (2,69€) in a month.
    - http://www.maestrosdelweb.com/editorial/por-que-utilizar-s3-el-sistema-de-almacenamiento-de-amazon/
  - Smugmug estimated that saved 1 million dollars in one year by using Amazon S3
    - http://www.error500.net/amazon-s3-ahorro-costes

Anatomy of a Cloud



SaaS	Software as a service	Google Apps, Microsoft "Software+Services"
PaaS	Platform as a service	Google AppEngine, IBM IT Factory, Force.com
IaaS	Infrastructure as a service	Amazon EC2, Sun Grid, IBM Blue Cloud
dSaaS	Storage as a service	Amazon S3, Nirvanix SDN, Cleversafe dsNet

- A full virtual platform that does NOT include all software stack (as PaaS does)
- It could be possible to manage (the system software) of the demanded platform

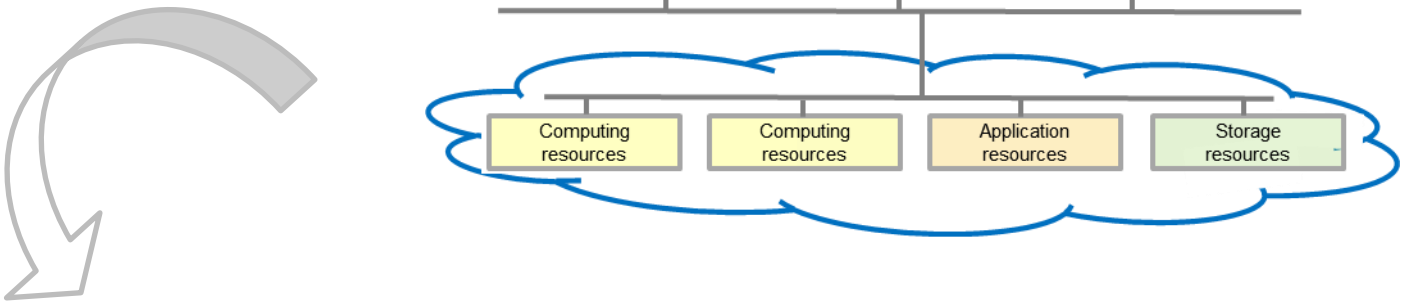
Cloud: IaaS

- The EC2 service bills the computing capacity requested & granted:
  - Instances of deployed virtual machine:
    - Several types of instances: small, large, etc.
  - Data transferred:
    - Fare by data transferred to/from EC2.
  - Storage:
    - Fare by the virtual machine data management (request and space).
  - Additional services:
    - Additional storage, monitoring, IP address, load balancing, auto-scale up, etc.

http://aws.amazon.com/ec2/

- Various uses:
  - Infrastructure for Applications, Content distribution, e-commerce, etc.
    - http://aws.amazon.com/solutions/case-studies/

Anatomy of a Cloud



SaaS	Software as a service	Google Apps, Microsoft "Software+Services"
PaaS	Platform as a service	Google AppEngine, IBM IT Factory, Force.com
IaaS	Infrastructure as a service	Amazon EC2, Sun Grid, IBM Blue Cloud
dSaaS	Storage as a service	Amazon S3, Nirvanix SDN, Cleversafe dsNet

- A full virtual platform (outsourcing of the company's infrastructure)
- It includes servers, operating systems, and specific applications

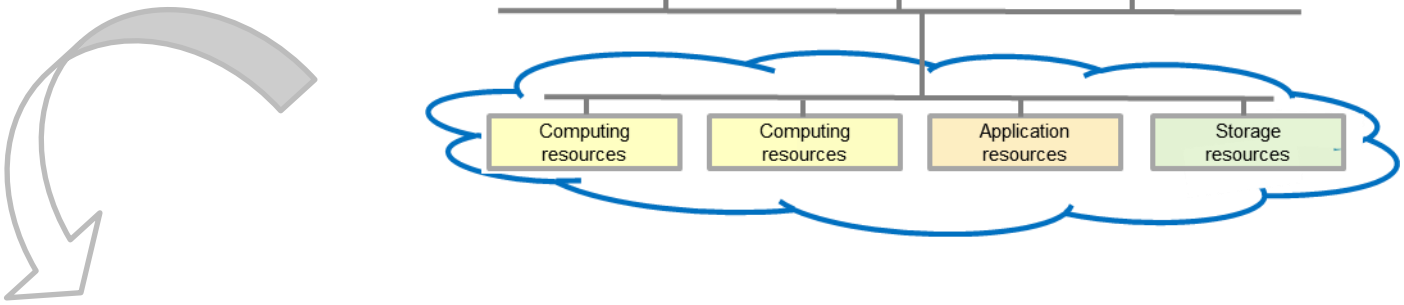
Cloud: PaaS

- Google App Engine let users execute Web applications in Google servers:
  - Google data centers offer scalability
  - Other google services can be integrated (e.g. : Google authentication can be used)

code.google.com/appengine/

- Google App Engine application are developed:
  - Python and Java are used for programming.
  - Data stored on GQL (similar to SQL)
- Google answer to:
  - Amazon Web Services, Microsoft Azure Services, Heroku, etc.

Anatomy of a Cloud



SaaS	Software as a service	Google Apps, Microsoft "Software+Services"
PaaS	Platform as a service	Google AppEngine, IBM IT Factory, Force.com
IaaS	Infrastructure as a service	Amazon EC2, Sun Grid, IBM Blue Cloud
dSaaS	Storage as a service	Amazon S3, Nirvanix SDN, Cleversafe dsNet

- Initial example can be found in old ASP (Application Service Provider)
- Another example: software offered as Web service that is used by a local application

Cloud: SaaS

- Applications for collaborative work:
  - Collaboration in the same document, rather than using attachment documents with the modifications.
  - Document sharing and calendar with work colleagues.
  - Access to all information from any computer.
  - Invitation to join a service team is not a complex task.

Google apps

www.google.com/apps/

- The main collaborative applications included are:

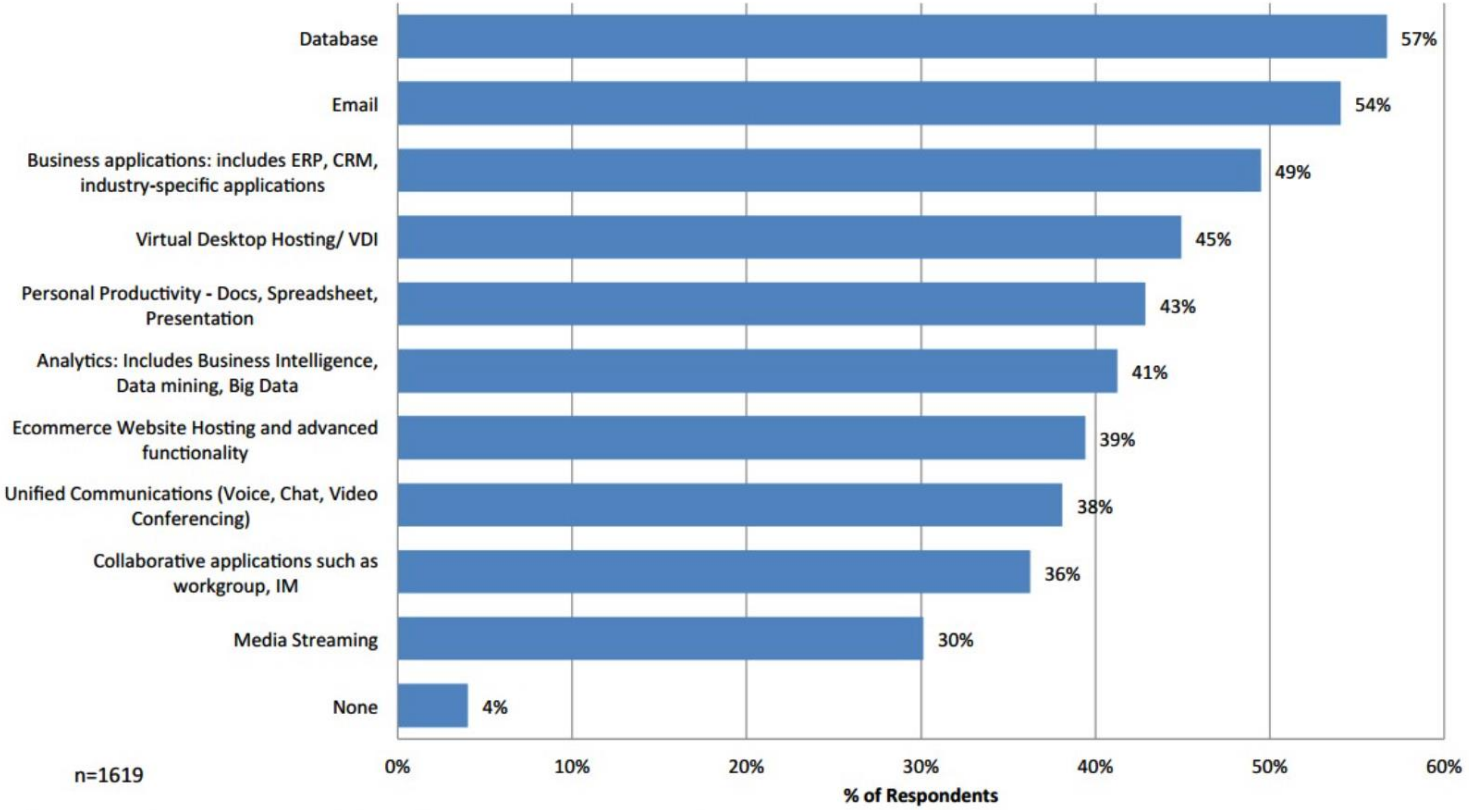
Google Docs: for editing documents, spreadsheets and presentations.

Google Calendar: for managing and introducing events (e.g.: meetings).

Google Hangout: for sending instant messages and video-conferencing.

Cloud: SaaS

Application Hosting: Spending Over Next Two Years



A1. Which of the following hosted applications – if any -- will comprise a significant portion of your Application Hosting spending with hosting & cloud providers over the next two years?

451 Research | Advisory Services

Page 18

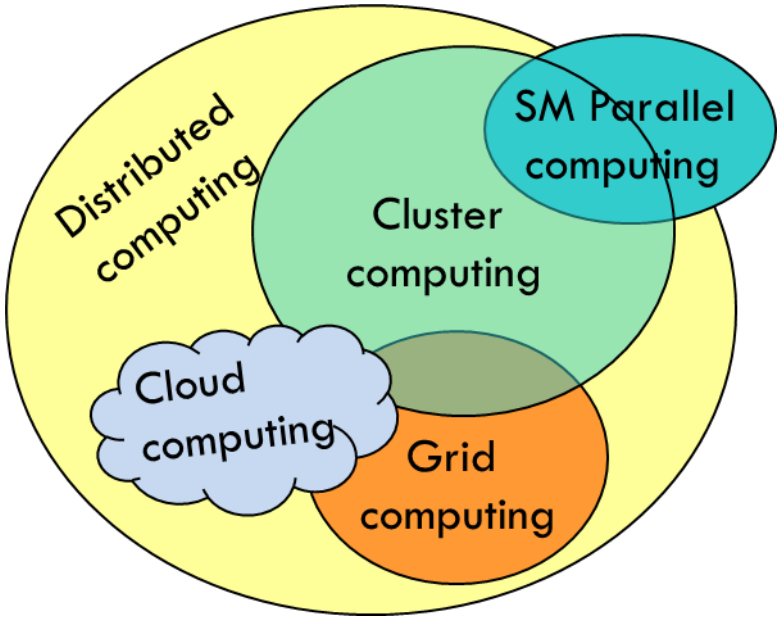
http://www.consultia.co/predicting-the-future-of-cloud-service-providers/

Summary

- Cloud computing is a model for enabling ubiquitous, 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.
  - 4 deployment models:
    - Private, Public, Hybrid, Community.
  - 3 service models:
    - SaaS, PaaS, IaaS.
  - 5 essential characteristics:
    - On-demand self-service, broad network access, resource pooling, rapid elasticity, measured service.

http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf

Scope



- **Mainframes and supercomputers**
- **Clusters**
- **Grid computing**
- **Volunteer computing**
- Cloud computing

Supercomputers and mainframes

- The supercomputers:
  - Mostly used for problems with great necessity of computing capacity (floating point ops.)
  - More frequently in research (science and army)
  - They use massive parallelism (they use many processors)

Supercomputers & mainframes
(SMP, MPP, Systolic, Array, …)

1950-1990

Supercomputers and mainframes

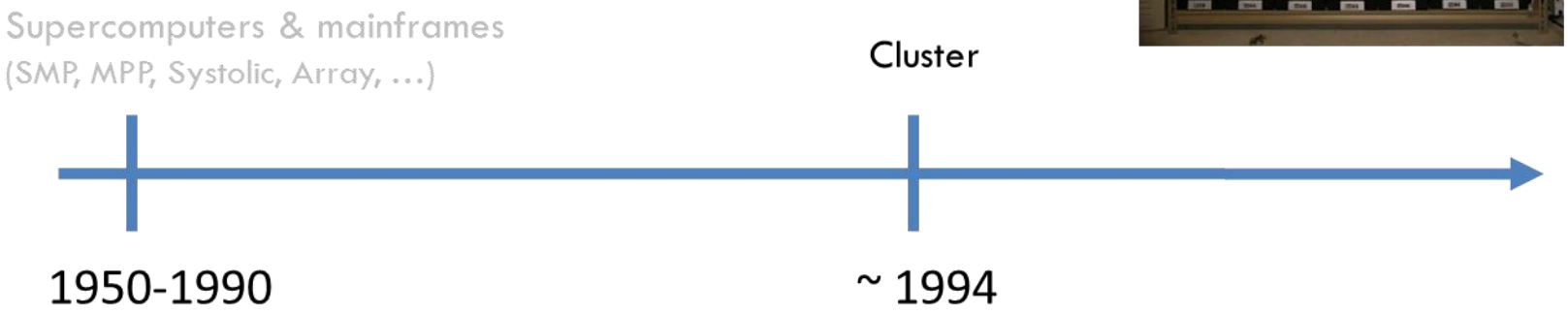
- The mainframes:
  - Mostly used for problems with great amount of data from I/O devices.
  - More frequently in companies and administrative environments.
  - Usually based on few number of processor (but with wide buses and higher frequency)
  - More sensible to reliability.

Supercomputers & mainframes
(SMP, MPP, Systolic, Array, …)

1950-1990

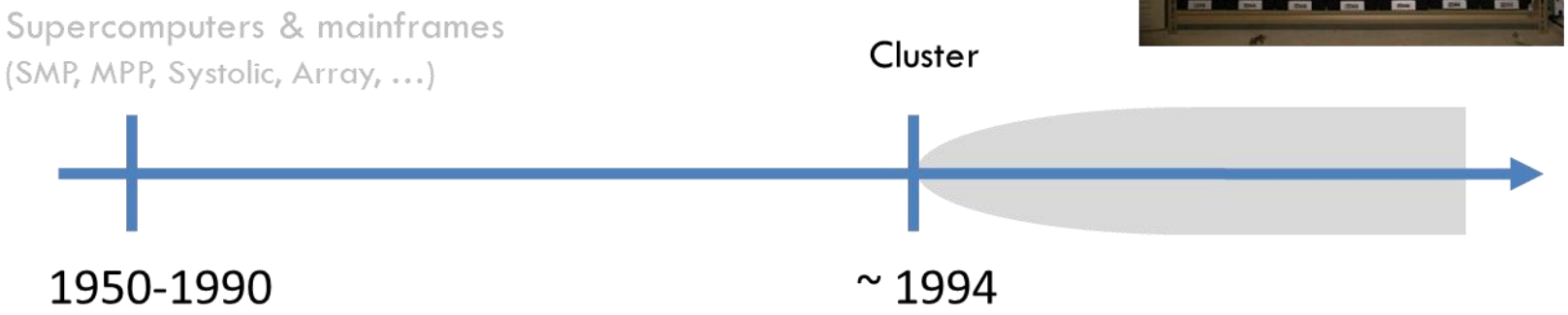
Clusters

- Built by Donald Becker and Thomas Sterling at 1994 (NASA)
- Consists of 16 personal computers with intel DX4 200 MHz processor interconnected by a Ethernet switch.
- Theoretical performance up to 3,2 Gflops
- Cluster let users to build cheaper supercomputers

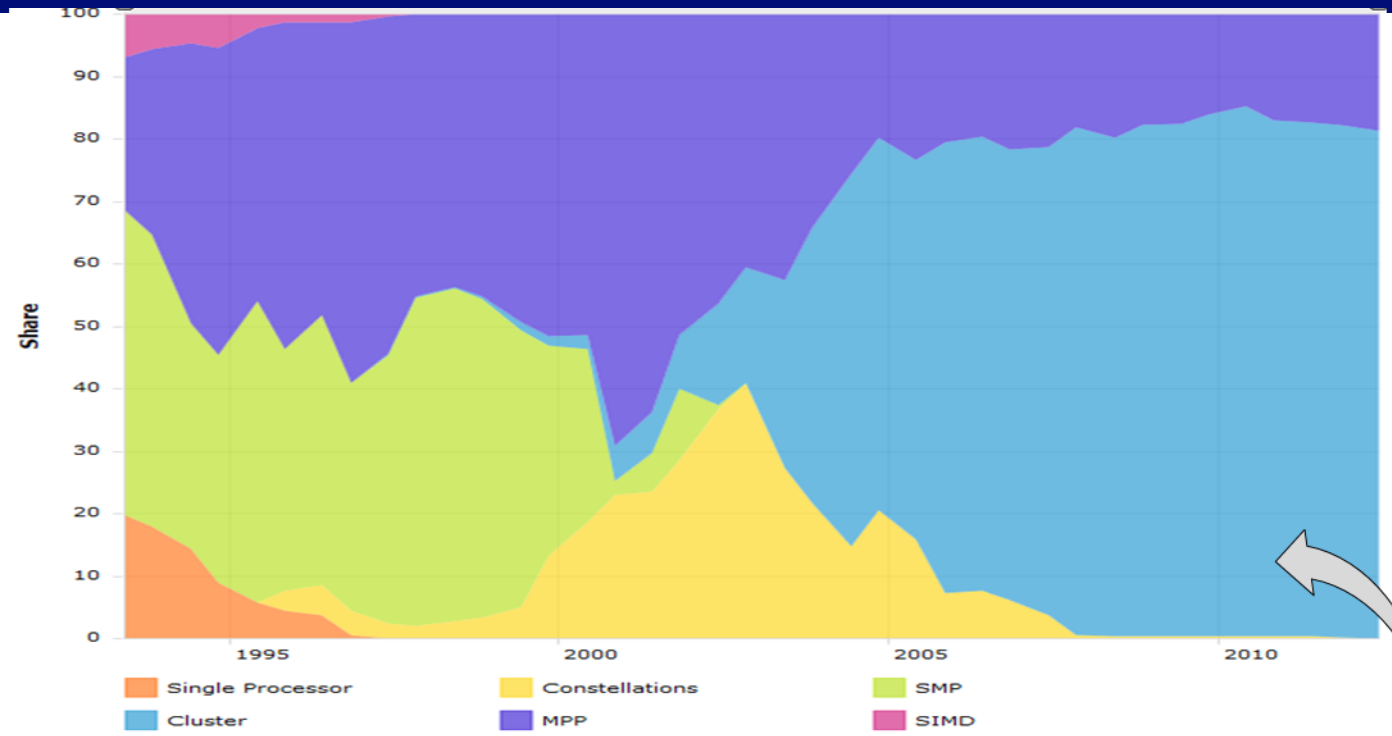


Clusters

- Built by Donald Becker and Thomas Sterling at 1994 (NASA)
- Consists of 16 personal computers with intel DX4 200 MHz processor interconnected by a Ethernet switch.
- Theoretical performance up to 3,2 Gflops
- Cluster let users to build cheaper supercomputers



Clusters



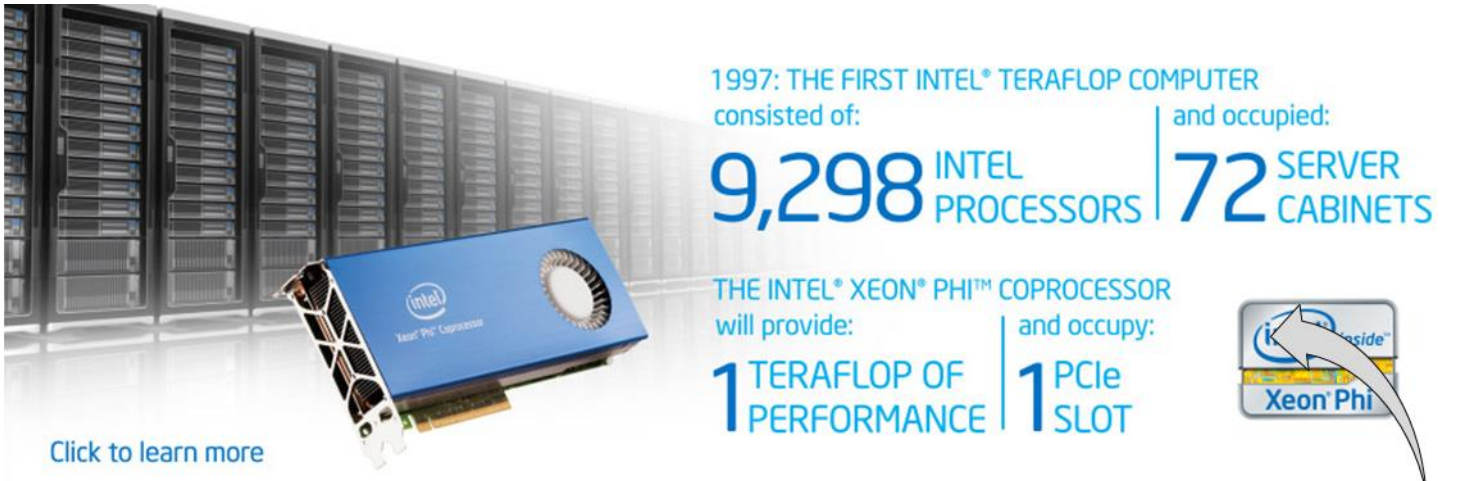
Supercomputers & mainframes
(SMP, MPP, Systolic, Array, …)

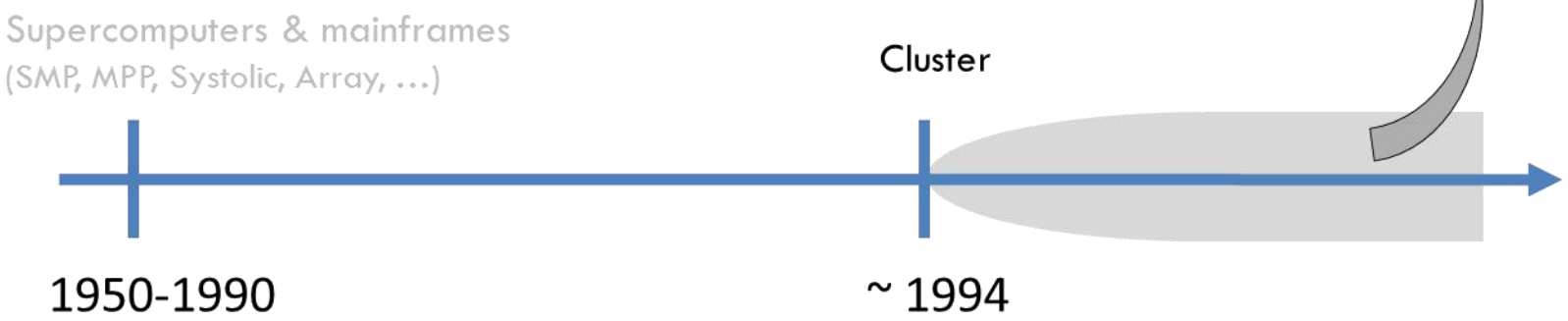
Cluster

1950-1990

~ 1994

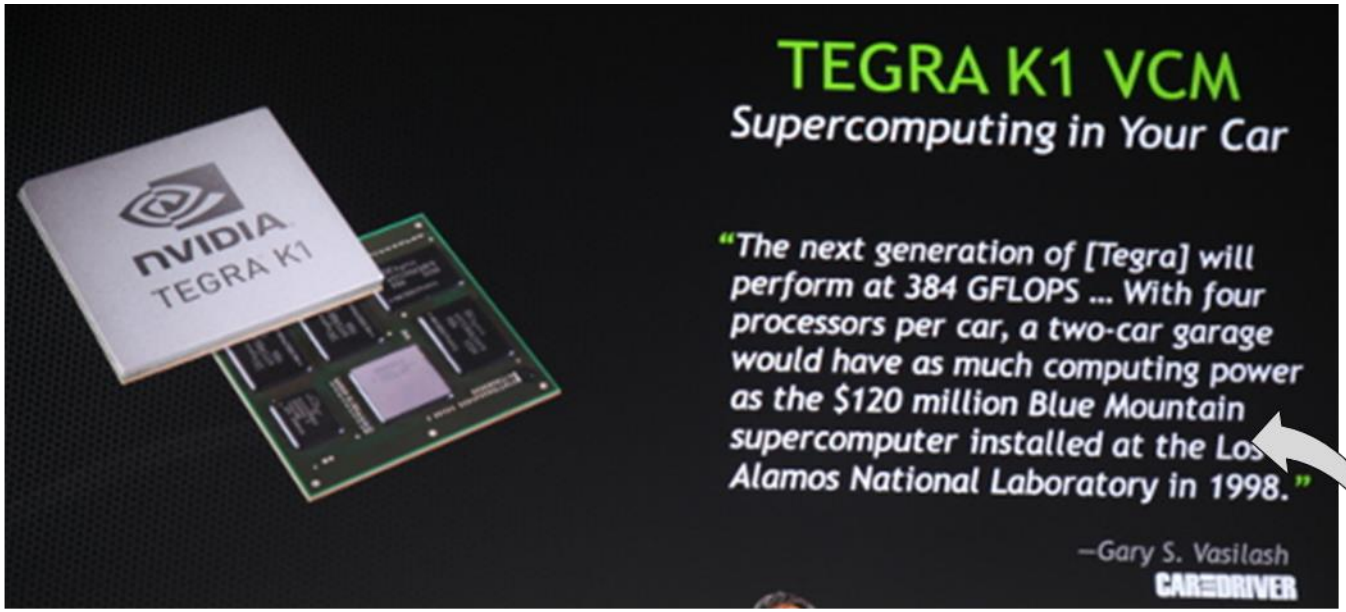
Clusters





http://es.wikipedia.org/wiki/Intel_MIC

Clusters



Supercomputers & mainframes
(SMP, MPP, Systolic, Array, ...)

Cluster

1950-1990

~ 1994

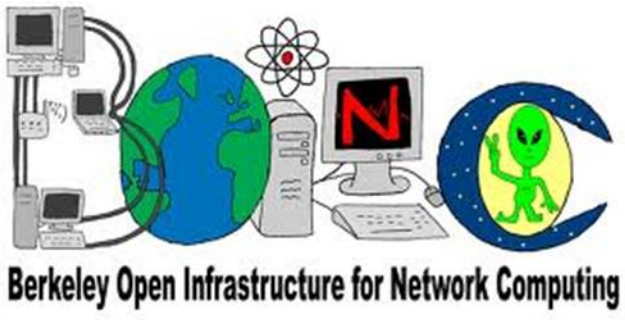
http://www.nvidia.com/object/visual-computing-module.html

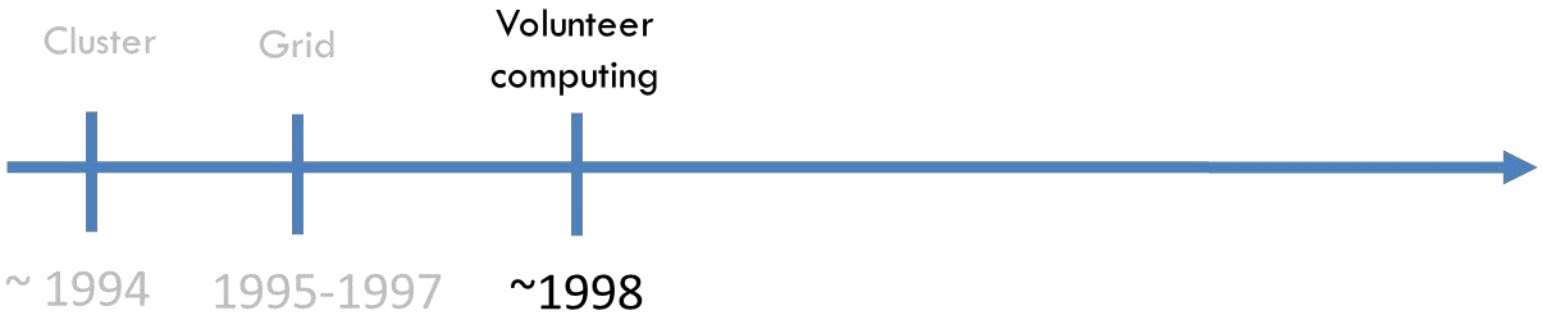
Grid

- Ancestor: metacomputing by Larry Smarr (NCSA) at beginning of 80's
  - Supercomputer centers interconnected: more resources available
  - I-WAY demonstrated on 1995
- Grid appears in a seminar given by Ian Foster y Carl Kesselman on 1997 at ANL

Volunteer computing

- Introduced by Luis F. G. Sarmenta (Bayanihan)
- On 1999 the proyects SETI@home and Folding@home start running
- On 16/oct/2010 all BOINC projects executes by using ~2,2 TeraFLOPS

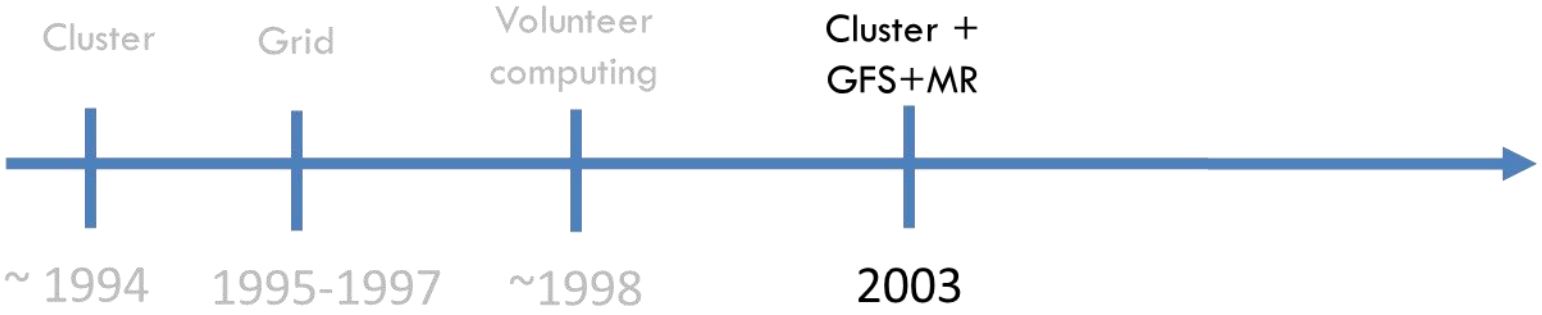




Google File System & MapReduce

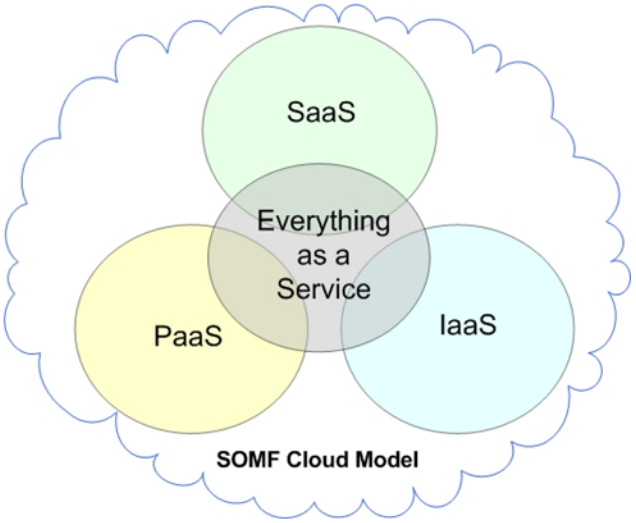
- Google introduces:
  - MapReduce as a framework for working with big dataset: the same function is applied to different data partition (map) and then results are combined (reduce)
  - GFS as a way to store petabytes of data (common computers, scalable distribution, and fault tolerance)
- GFS+MR let users to build cheaper "mainframes"

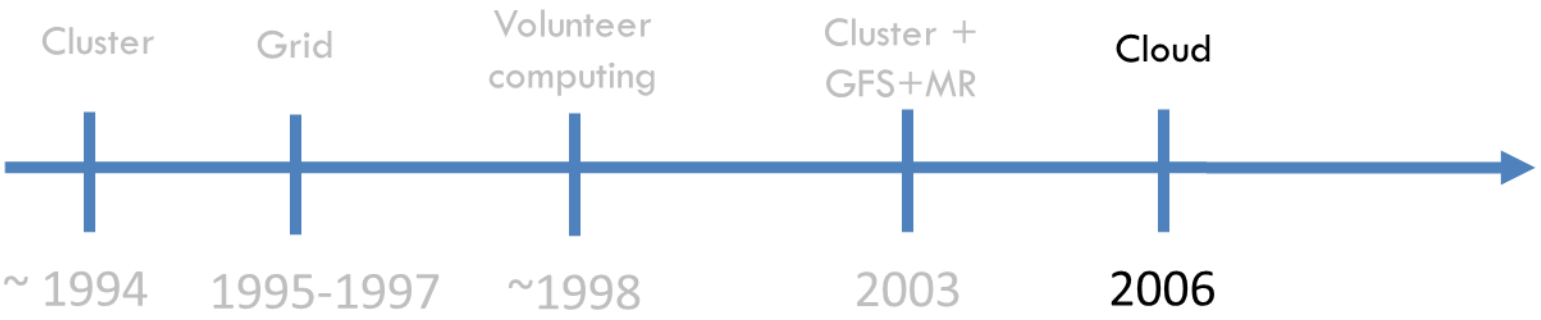
Doug Cutting and Hadoop



Cloud computing

- Amazon inspires:
  - It built its data center for the Christmas sales, but 10% of usage out there
  - Two main ingredients: utility computing and virtualization
- Advantages: agility, cost, scalability, maintenance, etc.
- OpenStack: let users to build a cloud from a cluster



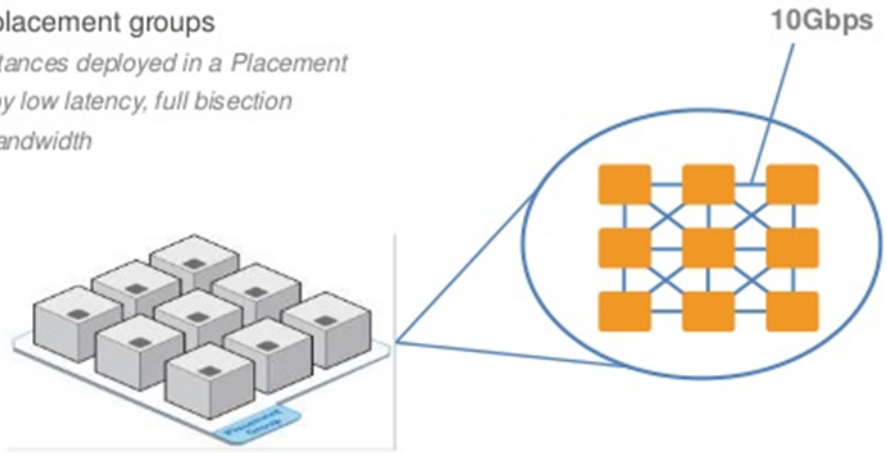


Amazon Cluster Compute Instance

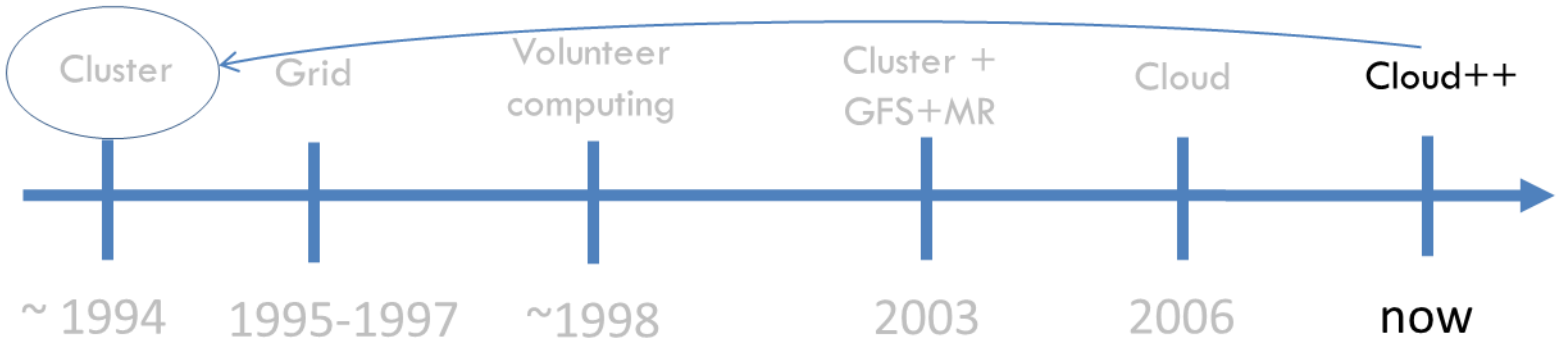
Tightly coupled

Network placement groups

Cluster instances deployed in a Placement Group enjoy low latency, full bisection 10 Gbps bandwidth

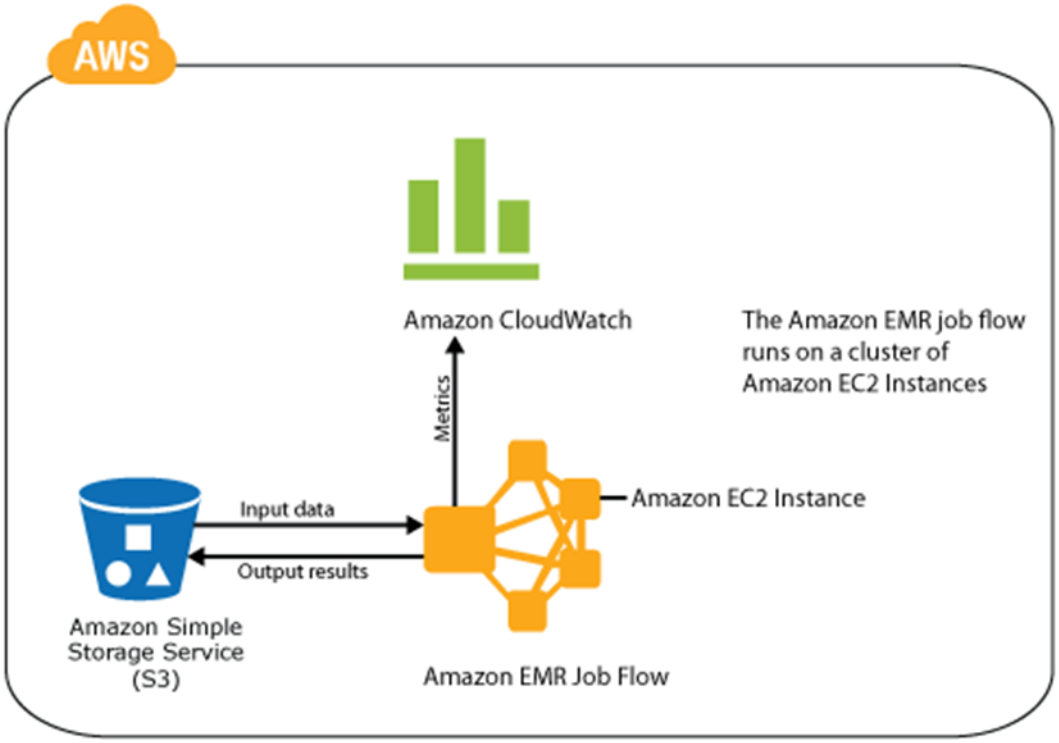


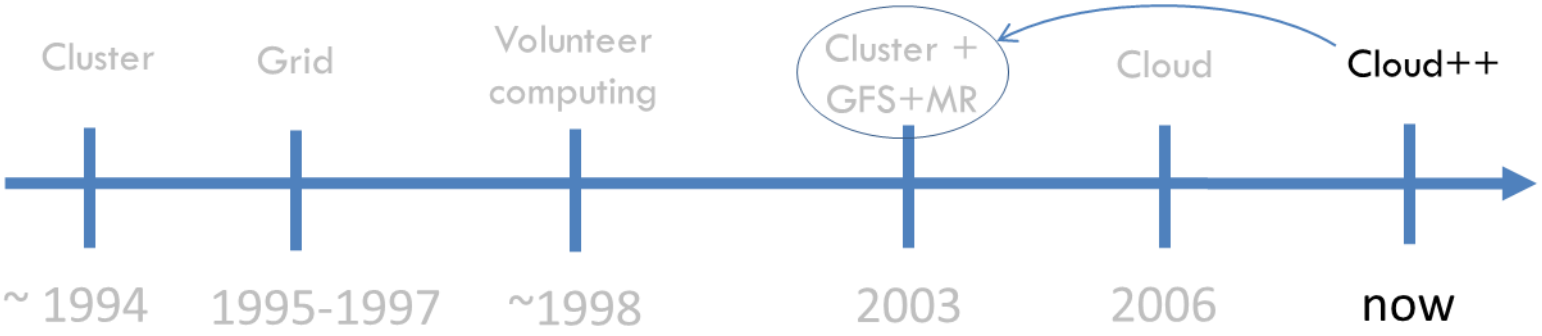
amazon web services



https://aws.amazon.com/es/blogs/aws/the-new-amazon-ec2-instance-type-the-cluster-compute-instance/

Amazon Elastic MapReduce





http://aws.amazon.com/es/elasticmapreduce/

Public cloud business model

- Large scale infrastructure available on a rental basis
  - Operating System virtualization provides CPU isolation
  - Network provisioning that provides network isolation
  - Locally specific storage abstractions
- Fully customer self-service
  - Service Level Agreements (SLAs) are advertised
  - Requests are accepted and resources granted via web
  - Customers access resources remotely via the Internet
- Accountability is e-commerce based
  - Web-based transaction
  - "Pay-as-you-go" and flat-rate subscription
  - Customer service, refunds, etc.

Amazon as testbed…

- http://aws.amazon.com/es/free/

Amazon EC2

Ofrece capacidad informática de tamaño variable en la nube.
Más información »

750 horas al mes de uso de las instancias t2.micro de Linux, RHEL o SLES

750 horas al mes de uso de la instancia t2.micro de Windows

Por ejemplo, puede ejecutar una instancia durante un mes o dos instancias durante medio mes.

Vence a los 12 meses a partir de la fecha de inscripción.

Amazon S3

Infraestructura de almacenamiento de alta escalabilidad, fiabilidad y baja latencia.
Más información »

5 GB de almacenamiento estándar

20 000 solicitudes Get

2 000 solicitudes Put

Vence a los 12 meses a partir de la fecha de inscripción.

Amazon DynamoDB

Base de datos NoSQL rápida y flexible con un escalabilidad perfecta.
Más información »

25 GB de almacenamiento

25 unidades de capacidad de escritura

25 unidades de capacidad de lectura

Suficiente para manejar hasta 200 millones de solicitudes al mes. No vence al finalizar el periodo de 12 meses de la capa gratuita de AWS.

CLOUD COMPUTING