Hálózatba kapcsolt erőforrás platformok és alkalmazásaik

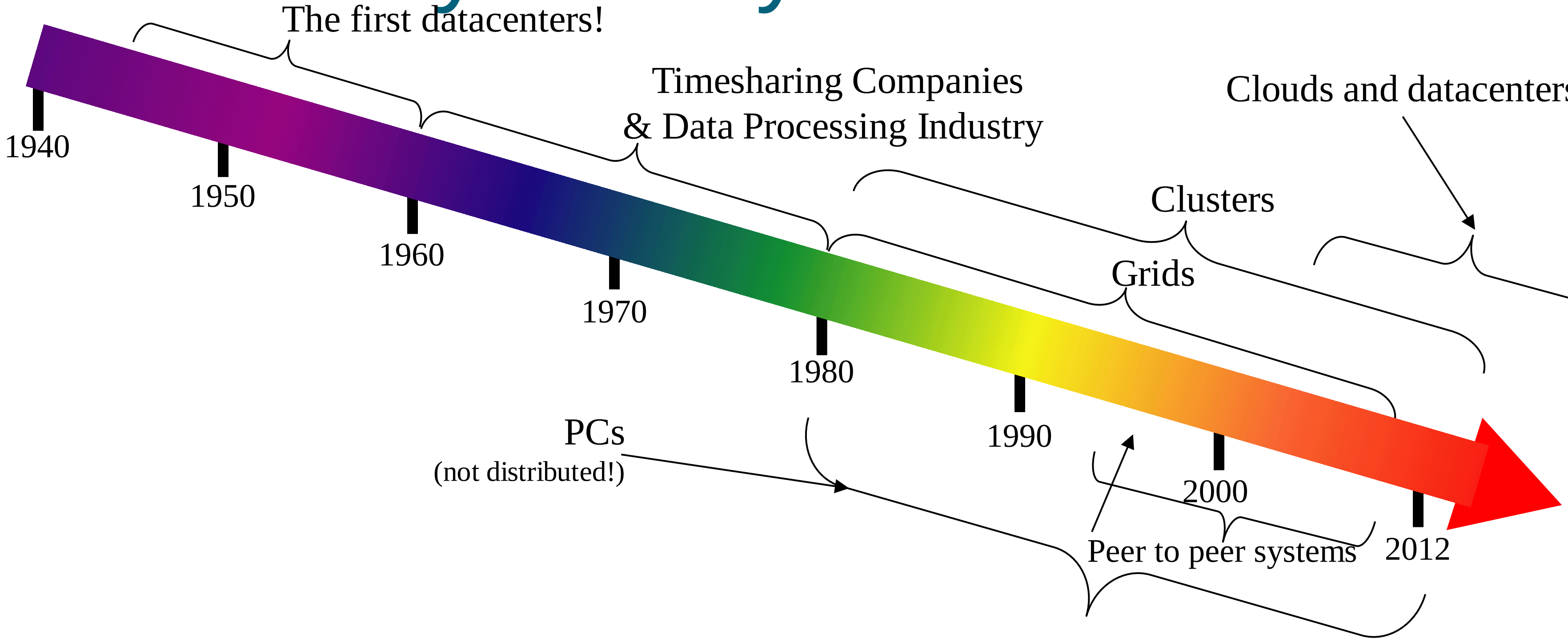
Simon Csaba
TMIT

2017

P2P Computing vs Grid Computing

- Differ in Target Communities
- Grid system deals with more complex, more powerful, more diverse and highly interconnected set of resources than P2P.
- VO

"A Cloudy History of Time"

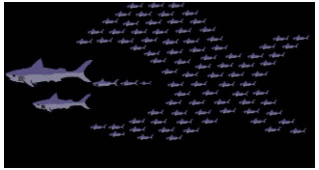


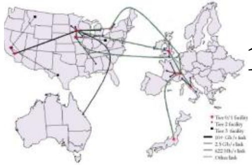
"A Cloudy History of Time"

First large datacenters: ENIAC, ORDVAC, ILLIAC
Many used vacuum tubes and mechanical relays

1940
1950
1960
1970
1980
1990
2000
2012

Clouds





Berkeley NOW Project
Supercomputers
Server Farms (e.g., Oceano)

P2P Systems (90s-00s)
- Many Millions of users
- Many GB per day

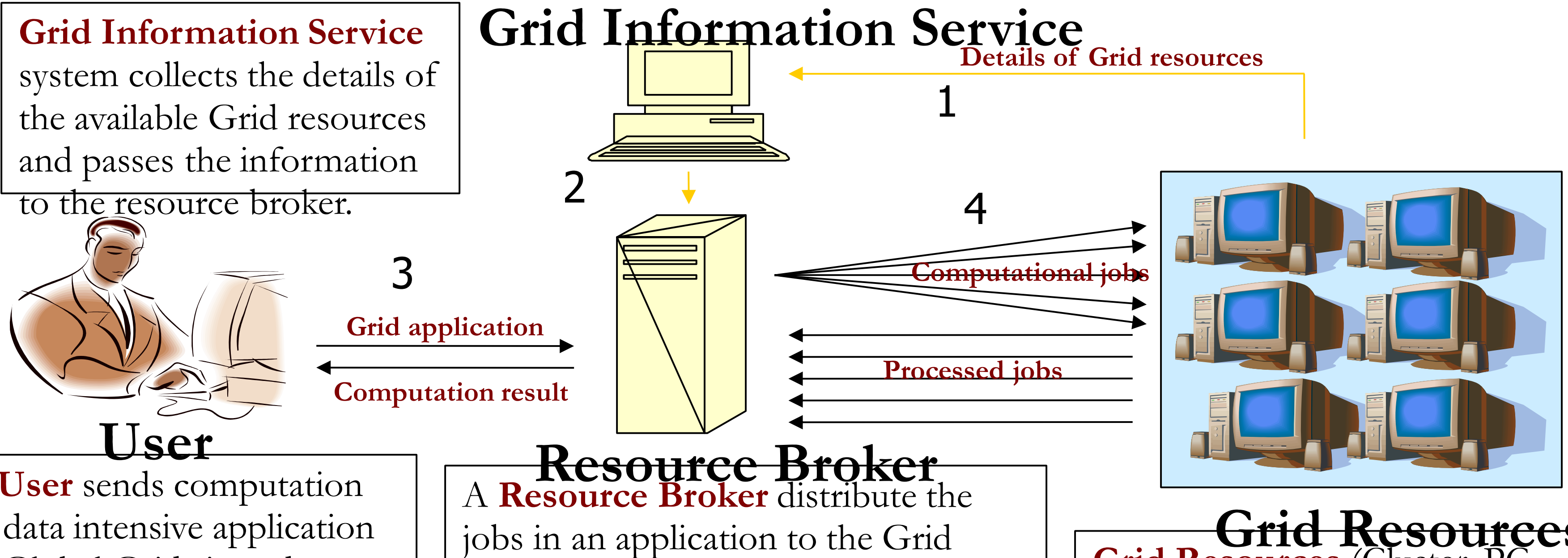
Data Processing Industry
- 1968: $70 M. 1978: $3.15 Billion

Timesharing Industry (1975):
- Market Share: Honeywell 34%, IBM 15%,
- Xerox 10%, CDC 10%, DEC 10%, UNIVAC 10%
- Honeywell 6000 & 635, IBM 370/168, Xerox 940 & Sigma 9, DEC PDP-10, UNIVAC 1108

Grids (1980s-2000s):
- GriPhyN (1970s-80s)
- Open Science Grid and Lambda Rail (2000s)
- Globus & other standards (1990s-2000s)

Grid - intro

A typical view of Grid environment



Grid Information Service system collects the details of the available Grid resources and passes the information to the resource broker.

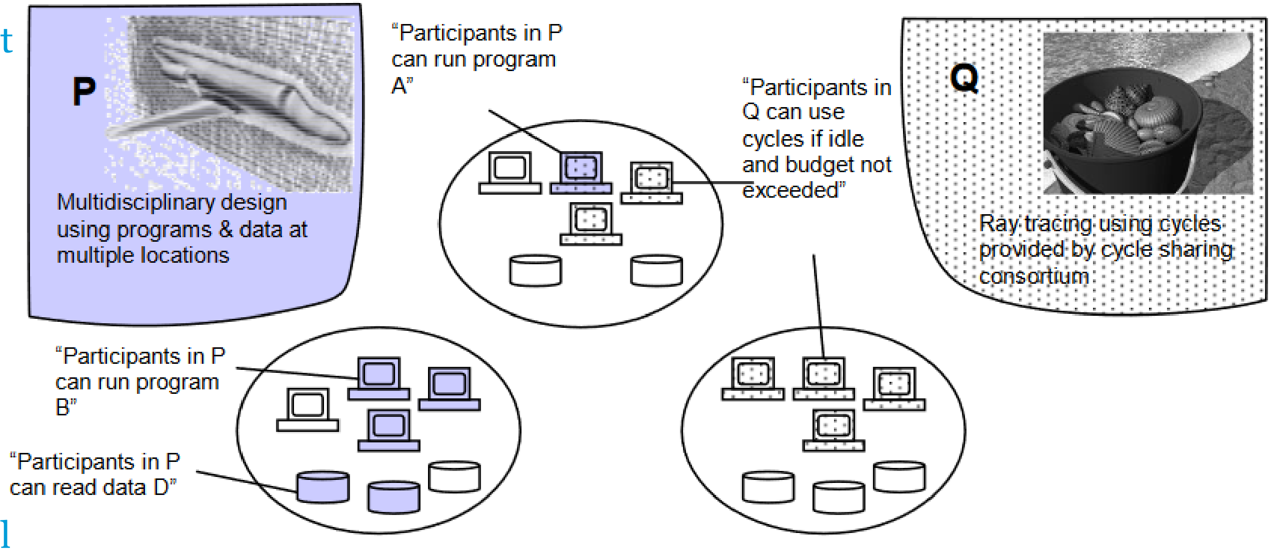
A **User** sends computation or data intensive application to Global Grids in order to speed up the execution of the application.

A **Resource Broker** distribute the jobs in an application to the Grid resources based on user's QoS requirements and details of available Grid resources for further executions.

Grid Resources (Cluster, PC, Supercomputer, database, instruments, etc.) in the Global Grid execute the user jobs.

History and Evolution of Grid
The emergence of virtual organisations (VO)

- VO = virtuális szervezet
  - Azonos jellegű erőforrás-mixet használ
  - Hasonló feladat-típus elvégzésére
  - Erőforrás felhasználás és biztosítás összefügg (freeride csökkentése/elkerülése)
- Virtuális, mert
  - Elosztott, több helyszínről csatlakozik a gridbe
  - egy erfőforrás overlayt használ
- Szervezet
  - Felhasználókat csoportosít
  - „alkalmazottai" = grid felhasználók



Foster I. et al (2003)

History and Evolution of Grid
The Emergence of Virtual Organisations

- Sharing resources:
  - The degree of service availability – which resources will be shared
  - The authorization of the shared resource – who will be permitted
  - The type of the relationship - Peer to peer
  - A mechanism to understand the nature of the relationship
  - The possible ways the resource will be used (memory, computing power, etc.)

Grid Middleware

- Grids are typically managed by grid ware - a special type of middleware that enable sharing and manage grid components based on user requirements and resource attributes (e.g., capacity, performance)
- Software that connects other software components or applications to provide the following functions:
  - Run applications on suitable available resources
    - Brokering, Scheduling
  - Provide uniform, high-level access to resources
    - Semantic interfaces
    - Web Services, Service Oriented Architectures
  - Address inter-domain issues of security, policy, etc.
    - Federated Identities
  - Provide application-level status
  - monitoring and control

Middlewares

- Globus – Chicago Uni
- Condor – Wisconsin Univ– High throughput computing
- Legion – Virginia Uni – virtual workspaces- collaborative computing
- IBP – Internet Back Pane – Tennesse Uni – logistical networking
- NetSolve – solving scientific problems in heterogeneous environments – high throughput & data intensive

Two Key Grid Computing Groups

The Globus Alliance (www.globus.org)

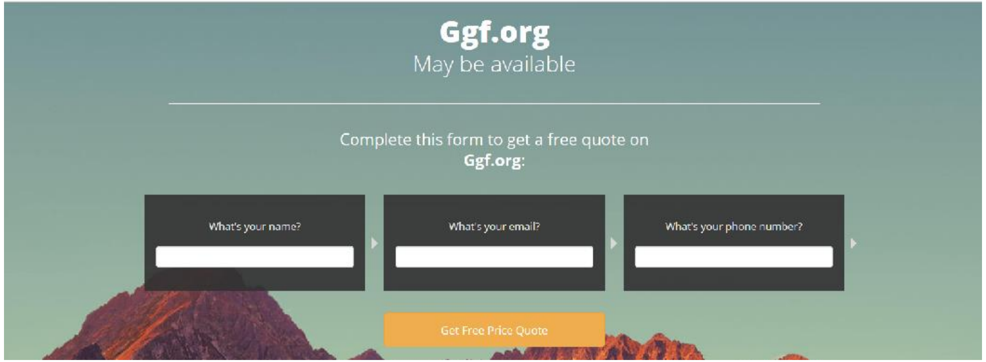
- Composed of people from:
  Argonne National Labs, University of Chicago, University of Southern California Information Sciences Institute, University of Edinburgh and others.
- OGSA/I standards initially proposed by the Globus Group

The Global Grid Forum (www.ggf.org)

- Heavy involvement of Academic Groups and Industry
  - (e.g. IBM Grid Computing, HP, United Devices, Oracle, UK e-Science Programme, US DOE, US NSF, Indiana University, and many others)
- Process
  - Meets three times annually
  - Solicits involvement from industry, research groups, and academics
- Open Grid Forum

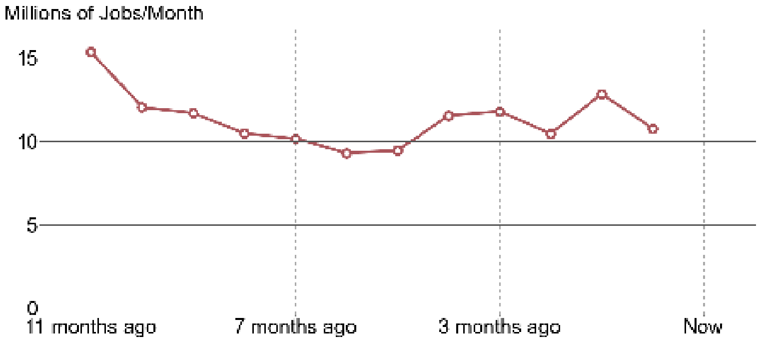
Grid-mánia lecsengőben van

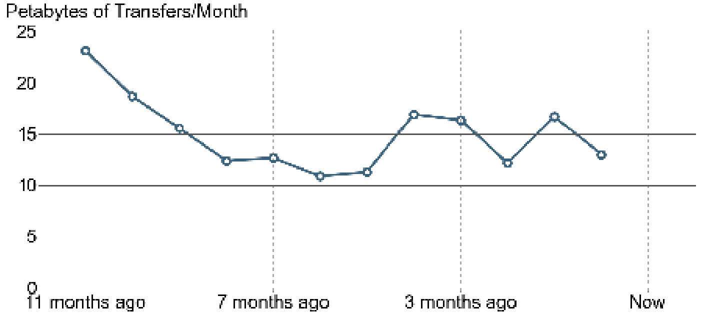
- Pl. ggf.org honlap: (ggf.org beolvadt az Open Grid Forum-ba)
  - Enterprise Grid Alliance (EGA)-val közösen



Open Science Grid (OSG)

Open Science Grid (OSG) – működő rendszer





„The Compute Element of OSG needs to run the globus-gatekeeper, tomcat5, globus-gridftp-server „

„The global pool of resources is a HTCondor pool; the batch system worker node startup script launches a HTCondor worker node"

geni
Exploring Networks of the Future

Mikor jó a grid? (OSG best practices)

- Linux application for the x86 or x86_64 architecture.
- single- or multi-threaded but does not require message passing.
- small runtime between 1 and 24 hours.
- can handle being unexpectedly killed and restarted.
- is built from software that does not require contact to licensing servers.
- The scientific problem can be described as a workflow consisting of jobs of such kind.
- The scientific problem requires running a very large number of small jobs rather than a few large jobs.

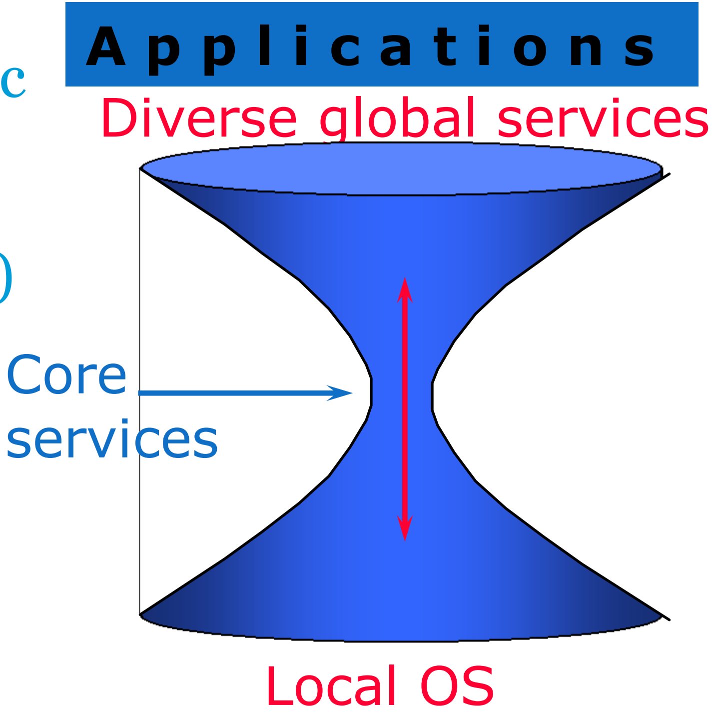
Some of the Major Grid Projects

Name	URL/Sponsor	Focus
EuroGrid, Grid Interoperability (GRIP)	eurogrid.org European Union	Create tech for remote access to super comp resources & simulation codes; in GRIP, integrate with Globus Toolkit™
Fusion Collaboratory	fusiongrid.org DOE Off. Science	Create a national computational collaboratory for fusion research
Globus Project™	globus.org DARPA, DOE, NSF, NASA, Msoft	Research on Grid technologies; development and support of Globus Toolkit™; application and deployment
GridLab	gridlab.org European Union	Grid technologies and applications
GridPP	gridpp.ac.uk U.K. eScience	Create & apply an operational grid within the U.K. for particle physics research
Grid Research Integration Dev. & Support Center	grids-center.org NSF	Integration, deployment, support of the NSF Middleware Infrastructure for research & education

Grid Architecture

The Hourglass Model

- Focus on architecture issues
  - Propose set of core services as basic infrastructure
  - Used to construct high-level, domain-specific solutions (diverse)
- Design principles
  - Keep participation cost low
  - Enable local control
  - Support for adaptation
  - “IP hourglass” model



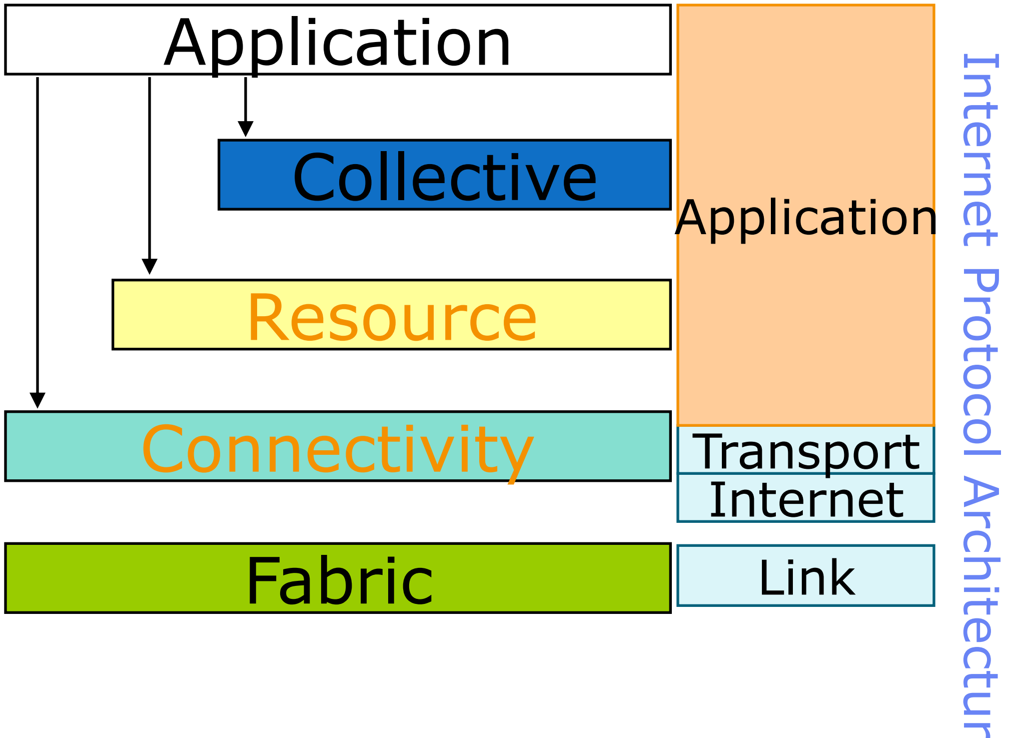
Layered Grid Architecture (By Analogy to Internet Architecture)

"Coordinating multiple resources": ubiquitous infrastructure services, app-specific distributed services

"Sharing single resources": negotiating access, controlling use

"Talking to things": communication (Internet protocols) & security

"Controlling things locally": Access to, & control of, resources



Example: Data Grid Architecture

App	Discipline-Specific Data Grid Application
Collective (App)	Coherency control, replica selection, task management, virtual data catalog, virtual data code catalog, …
Collective (Generic)	Replica catalog, replica management, co-allocation, certificate authorities, metadata catalogs,
Resource	Access to data, access to computers, access to network performance data, …
Connect	Communication, service discovery (DNS), authentication, authorization, delegation
Fabric	Storage systems, clusters, networks, network caches, …

Globus Toolkit™

- A software toolkit addressing key technical problems in the development of Grid-enabled tools, services, and applications
  - Offer a modular set of orthogonal services
  - Enable *incremental* development of grid-enabled tools and applications
  - Implement standard Grid protocols and APIs
  - Available under liberal open source license
  - Large community of developers & users
  - Commercial support

General Approach

- Define Grid protocols & APIs
  - Protocol-mediated access to remote resources
  - Integrate and extend existing standards
  - "On the Grid" = speak "Intergrid" protocols
- Develop a reference implementation
  - Open source Globus Toolkit
  - Client and server SDKs, services, tools, etc.
- Grid-enable wide variety of tools
  - Globus Toolkit, FTP, SSH, Condor, SRB, MPI, …
- Learn through deployment and applications

Key Protocols

- The Globus Toolkit™ centers around four key protocols
  - Connectivity layer:
    - *Security*: Grid Security Infrastructure (GSI)
  - Resource layer:
    - *Resource Management*: Grid Resource Allocation Management (GRAM)
    - *Information Services*: Grid Resource Information Protocol (GRIP) and Index Information Protocol (GIIP)
    - *Data Transfer*: Grid File Transfer Protocol (GridFTP)
- Also key collective layer protocols
  - Info Services, Replica Management, etc.

Technologies

Condor

- It is a specialized job and resource management system. It provides:
  - Job management mechanism
  - Scheduling
  - Priority scheme
  - Resource monitoring
  - Resource management

Technologies
Condor Terminology

- The user submits a job to an agent.
- The agent is responsible for remembering jobs in persistent storage while finding resources willing to run them.
- Agents and resources advertise themselves to a matchmaker, which is responsible for introducing potentially compatible agents and resources.
- At the agent, a shadow is responsible for providing all the details necessary to execute a job.
- At the resource, a sandbox is responsible for creating a safe execution environment for the job and protecting the resource from any mischief.

Cloud - intro


Google App Engine
amazon web services
RIGHT SCALE
ORACLE
GridGain 2.0
Elastic Server
VirtualBox
virtualLogix
vmware
force.com platform as a service
CITRIX
OpenVZ
10gen
KVM
Microsoft
parascale Powering Cloud Storage
hp
Xen
GIGASPACES
Parallels Optimized Computing
RIGHT SCALE
enomaly
elastra
ORACLE
BUNGEEconnect
PowerVM
Virtualiron

Cloud és a többiek…

Cluster

- szorosan csatolt, azonos rendszerek (HW és OS)
- központosított feladatkezelés és ütemezés
- dedikált kis késleltetésű és nagysebességű hálózat
- egy adott feladat megoldására

Grid

- lazán csatolt, különböző rendszerek (HW és OS)
- autonóm elemek, saját erőforráskezelővel, elosztott vezérlés
- egy nagyobb feladat osztódik szét
- nagy teljesítményű számítógépek
- Interneten keresztül összekötve

Cloud

- állhat azonos és különböző rendszerekből is
- független elemek, ahol az erőforrásokat a virtuális gép menedzser kezeli
- több különböző alkalmazás futtatása egyidőben
- dedikált kis késleltetésű és nagysebességű hálózat

The Hype!

- Forrester in 2010 – Cloud computing will go from $40.7 billion in 2010 to $241 billion in 2020.
- Goldman Sachs says cloud computing will grow at annual rate of 30% from 2013-2018
- Hadoop market to reach $20.8 B by by 2018: Transparency Market Research
- Companies and even Federal/state governments using cloud computing now: fbo.gov

Many Cloud Providers

- AWS: Amazon Web Services
  - EC2: Elastic Compute Cloud
  - S3: Simple Storage Service
  - EBS: Elastic Block Storage
- Microsoft Azure
- Google Cloud/Compute Engine/AppEngine
- Rightscale, Salesforce, EMC, Gigaspaces, 10gen, Datastax, Oracle, VMWare, Yahoo, Cloudera
- And many many more!

Two Categories of Clouds

- Can be either a (i) public cloud, or (ii) private cloud
- Private clouds are accessible only to company employees
- Public clouds provide service to any paying customer:
  - Amazon S3 (Simple Storage Service): store arbitrary datasets, pay per GB-month stored
    - As of 2015: 1-3 c per GB month
  - Amazon EC2 (Elastic Compute Cloud): upload and run arbitrary OS images, pay per CPU hour used
    - As of 2015: 1.3 c per CPU hr to $5.52 per CPU hr (depending on strength)
  - Google cloud: similar pricing as above
  - Google AppEngine/Compute Engine: develop applications within their appengine framework, upload data that will be imported into their format, and run

Customers Save Time and $$$

- Dave Power, Associate Information Consultant at Eli Lilly and Company: “With AWS, Powers said, a new server can be up and running in **three minutes** (it used to take Eli Lilly **seven and a half weeks** to deploy a server internally) and a **64-node Linux cluster** can be online in five minutes (compared with three months internally). … It's just shy of instantaneous.”

- Ingo Elfering, Vice President of Information Technology Strategy, GlaxoSmithKline: “With Online Services, we are able to reduce our IT **operational costs** by roughly **30%** of what we’re spending”

- Jim Swartz, CIO, Sybase: “At Sybase, a private cloud of virtual servers inside its datacenter has saved nearly **$US2 million annually** since 2006, Swartz says, because the company can share computing power and storage resources across servers.”

- 100s of startups in Silicon Valley can harness large computing resources without buying their own machines.

But what exactly IS a cloud?

What is a Cloud?

- It’s a cluster!
- It’s a supercomputer!
- It’s a datastore!
- It’s superman!

- None of the above
- All of the above

- Cloud = Lots of storage + compute cycles nearby

What is a Cloud?

- A single-site cloud (aka “Datacenter”) consists of
  - Compute nodes (grouped into racks)
  - Switches, connecting the racks
  - A network topology, e.g., hierarchical
  - Storage (backend) nodes connected to the network
  - Front-end for submitting jobs and receiving client requests
  - (Often called “three-tier architecture”)
  - Software Services
- A geographically distributed cloud consists of
  - Multiple such sites
  - Each site perhaps with a different structure and services

Trends: Technology

- Doubling Periods – storage: 12 mos, bandwidth: 9 mos, and (what law is this?) cpu compute capacity: 18 mos
- Then and Now
  - Bandwidth
    - 1985: mostly 56Kbps links nationwide
    - 2015: Tbps links widespread
  - Disk capacity
    - Today's PCs have TBs, far more than a 1990 supercomputer

Trends: Users

- Then and Now

  Biologists:

  - 1990: were running small single-molecule simulations
  - Today: CERN’s Large Hadron Collider producing many PB/year

Prophecies

- In 1965, MIT's Fernando Corbató and the other designers of the Multics operating system envisioned a computer facility operating “like a power company or water company”.
- Plug your thin client into the computing Utility and Play your favorite Intensive Compute & Communicate Application
  - Have today’s clouds brought us closer to this reality? Think about it.

Four Features New in Today's Clouds

I. Massive scale.

II. On-demand access: Pay-as-you-go, no upfront commitment.
   - And anyone can access it

III. Data-intensive Nature: What was MBs has now become TBs, PBs and XBs.
   - Daily logs, forensics, Web data, etc.
   - Humans have data numbness: Wikipedia (large) compressed is only about 10 GB!

IV. New Cloud Programming Paradigms: MapReduce/Hadoop, NoSQL/Cassandra/MongoDB and many others.
   - High in accessibility and ease of programmability
   - Lots of open-source

Combination of one or more of these gives rise to novel and unsolved distributed computing problems in cloud computing.

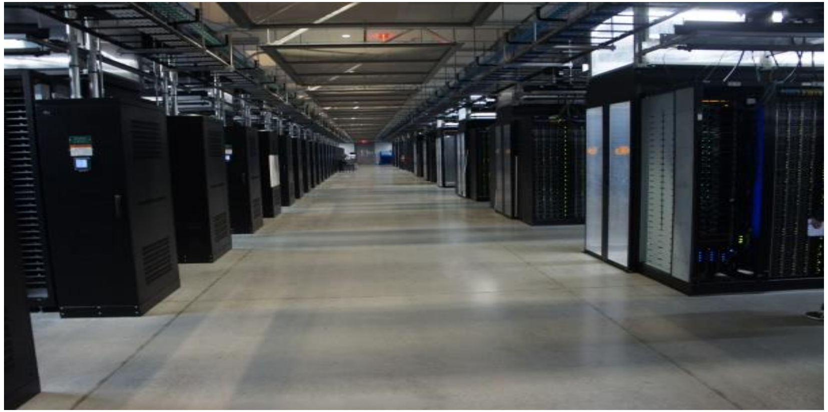
I. Massive Scale

- Facebook [GigaOm, 2012]
  - 30K in 2009 -> 60K in 2010 -> 180K in 2012
- Microsoft [NYTimes, 2008]
  - 150K machines
  - Growth rate of 10K per month
  - 80K total running Bing
  - In 2013, Microsoft Cosmos had 110K machines (4 sites)
- Yahoo! [2009]:
  - 100K
  - Split into clusters of 4000
- AWS EC2 [Randy Bias, 2009]
  - 40K machines
  - 8 cores/machine
- eBay [2012]: 50K machines
- HP [2012]: 380K in 180 DCs
- Google [2011, Data Center Knowledge] : 900K

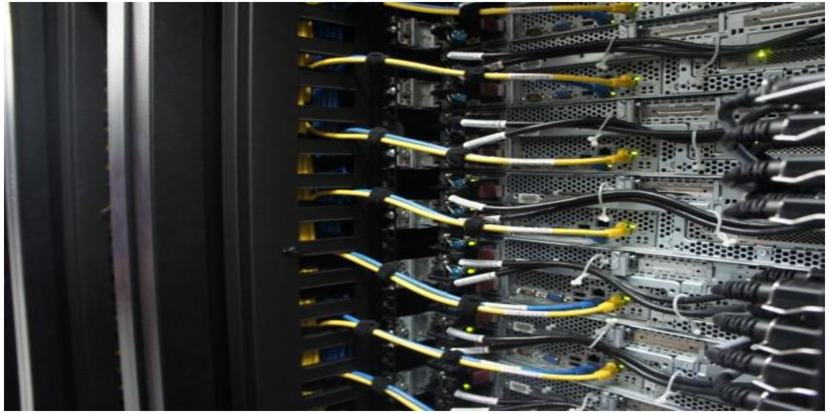
What does a datacenter look like from inside?

- A virtual walk through a datacenter
- Reference: http://gigaom.com/cleantech/a-rare-look-inside-facebooks-oregon-data-center-photos-video/

Servers



Front

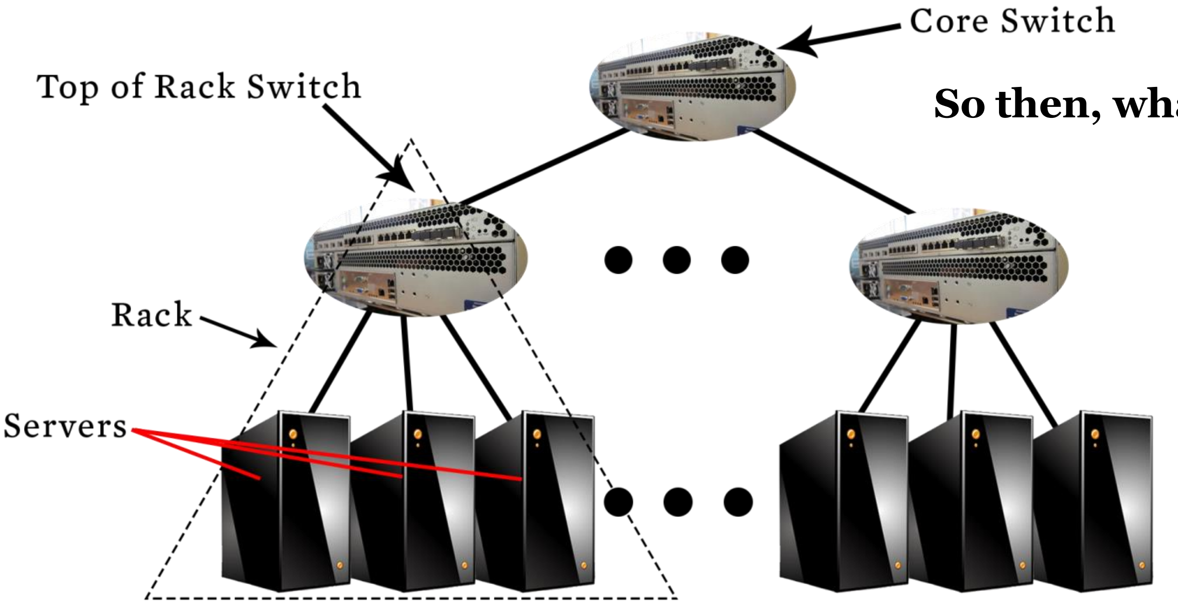


Back

In

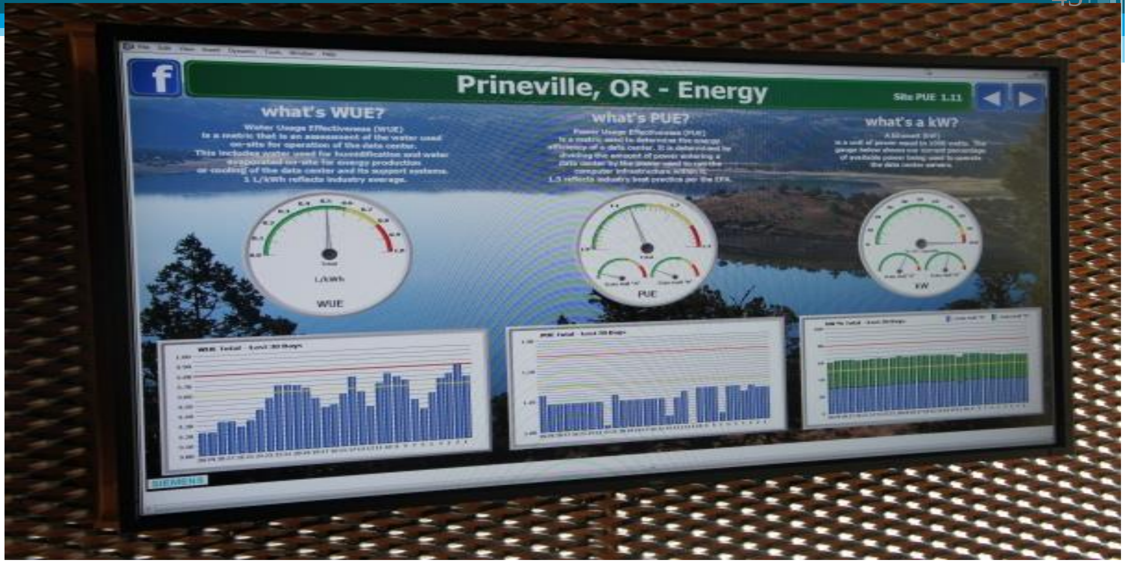
Some highly secure (e.g., financial info)

A Sample Cloud Topology



So then, what is a cluster?

Power



Off-site

On-site

- WUE = Annual Water Usage / IT Equipment Energy (L/kWh) – low is good
- PUE = Total facility Power / IT Equipment Power – low is good (e.g., Google~1.1)

Cooling

Air sucked in from top (also, Bugzappers)

Water purified

Water sprayed into air

15 motors per server bank

II. On-demand access: *aaS Classification

On-demand: renting a cab vs. (previously) renting a car, or buying one. E.g.:

- AWS Elastic Compute Cloud (EC2): a few cents to a few $ per CPU hour
- AWS Simple Storage Service (S3): a few cents per GB-month

- HaaS: Hardware as a Service
  - You get access to barebones hardware machines, do whatever you want with them, Ex: Your own cluster
  - Not always a good idea because of security risks
- IaaS: Infrastructure as a Service
  - You get access to flexible computing and storage infrastructure. Virtualization is one way of achieving this (cgroups, Kubernetes, Dockers, VMs,…). Often said to subsume HaaS.
  - Ex: Amazon Web Services (AWS: EC2 and S3), OpenStack, Eucalyptus, Rightscale, Microsoft Azure, Google Cloud.

II. On-demand access: *aaS Classification

- PaaS: Platform as a Service
  - You get access to flexible computing and storage infrastructure, coupled with a software platform (often tightly coupled)
  - Ex: Google’s AppEngine (Python, Java, Go)
- SaaS: Software as a Service
  - You get access to software services, when you need them. Often said to subsume SOA (Service Oriented Architectures).
  - Ex: Google docs, MS Office on demand

III. Data-intensive Computing

- Computation-Intensive Computing
  - Example areas: MPI-based, High-performance computing, Grids
  - Typically run on supercomputers (e.g., NCSA Blue Waters)
- Data-Intensive
  - Typically store data at datacenters
  - Use compute nodes nearby
  - Compute nodes run computation services
- In data-intensive computing, the focus shifts from computation to the data: CPU utilization no longer the most important resource metric, instead I/O is (disk and/or network)

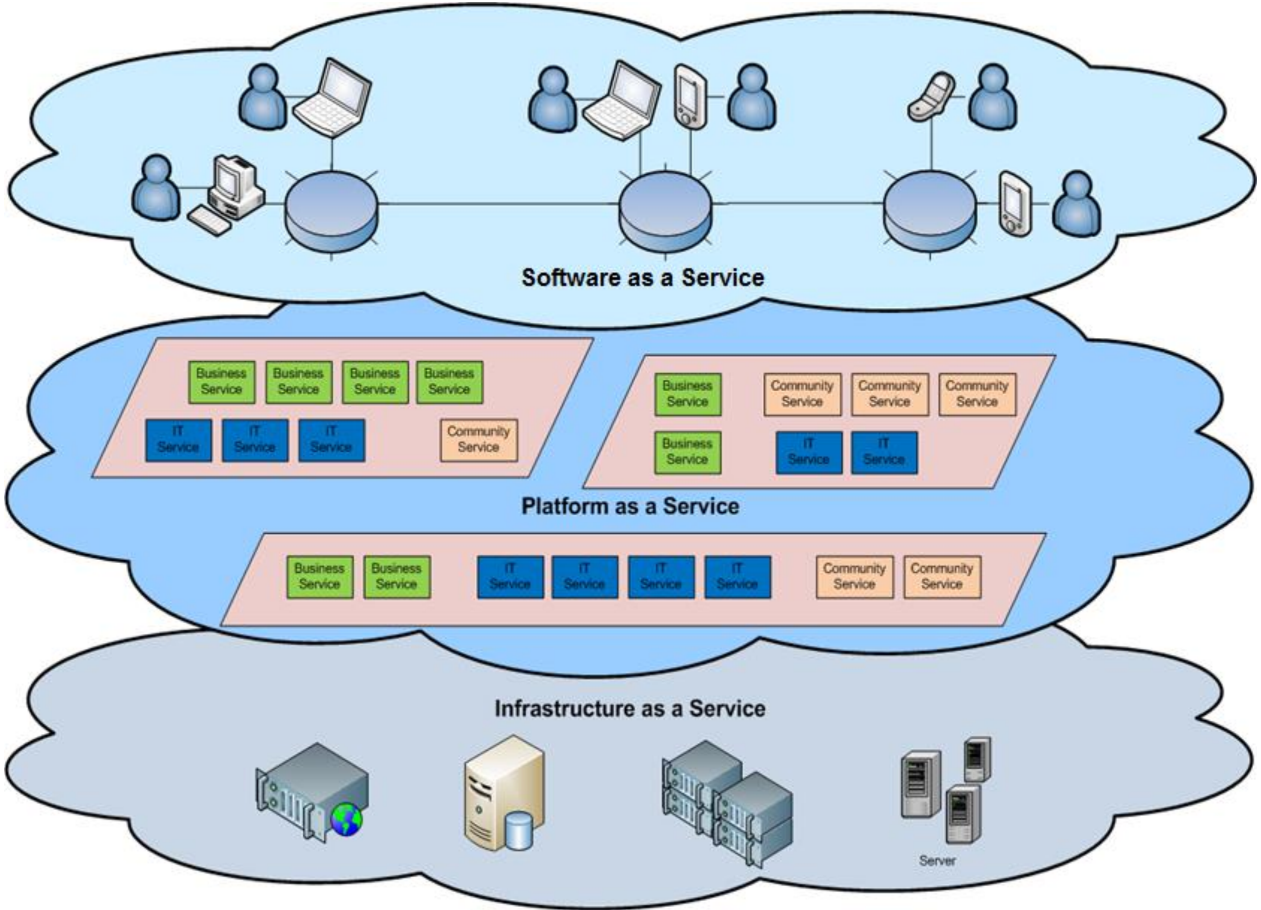
IV. New Cloud Programming Paradigms

- Easy to write and run highly parallel programs in new cloud programming paradigms:
  - Google: MapReduce and Sawzall
  - Amazon: Elastic MapReduce service (pay-as-you-go)
  - Google (MapReduce)
    - Indexing: a chain of 24 MapReduce jobs
    - ~200K jobs processing 50PB/month (in 2006)
  - Yahoo! (Hadoop + Pig)
    - WebMap: a chain of several MapReduce jobs
    - 300 TB of data, 10K cores, many tens of hours (~2008)
  - Facebook (Hadoop + Hive)
    - ~300TB total, adding 2TB/day (in 2008)
    - 3K jobs processing 55TB/day
  - Similar numbers from other companies, e.g., Yieldex, eharmony.com, etc.
  - NoSQL: MySQL is an industry standard, but Cassandra is 2400 times faster!

Introduction (revisites)

- **Scalable resource allocation**
- **Tailored services**
  - Software as a Service (SaaS)
  - Platform as a Service (PaaS)
  - Infrastructure as a Service (IaaS)

Cloud Arch.


Software as a Service
Business Service
Business Service
Business Service
Business Service
IT Service
IT Service
IT Service
Community Service
Business Service
Community Service
Community Service
Community Service
Business Service
IT Service
IT Service
Platform as a Service
Business Service
Business Service
IT Service
IT Service
IT Service
IT Service
Community Service
Community Service
Infrastructure as a Service
Server

Exapmles: Cloud Computing Layers

- **Application Service (SaaS)**
  - MS Live/Exchange, Google Docs, Salesforce.com, Quicken Online, Jupyter
- **Application Platform (PaaS)**
  - Google App Engine, Heroku, AWS
- **Server Platform (IaaS)**
  - Google Compute Engine, Amazon EC2, OpenStack, Eucalpytus

Cloud Computing Layers

	Services	Description
Application Focused	Services	Services – Complete business services such as PayPal, OpenID, OAuth, Google Maps, Alexa
	Application	Application – Cloud based software that eliminates the need for local installation such as Google Apps, Microsoft Online
	Development	Development – Software development platforms used to build custom cloud based applications (PAAS & SAAS) such as SalesForce
Infrastructure Focused	Platform	Platform – Cloud based platforms, typically provided using virtualization, such as Amazon ECC, Sun Grid
	Storage	Storage – Data storage or cloud based NAS such as iCoud, Dropbox, CloudNAS
	Hosting	Hosting – Physical data centers such as those run by IBM, HP, Amazon, etc.

Two Categories of Clouds

- Can be either a (i) public cloud, or (ii) private cloud
- Private clouds are accessible only to company employees
- Public clouds provide service to any paying customer

- You're starting a new service/company: should you use a public cloud or purchase your own private cloud?

- +

Cloud categories - refined

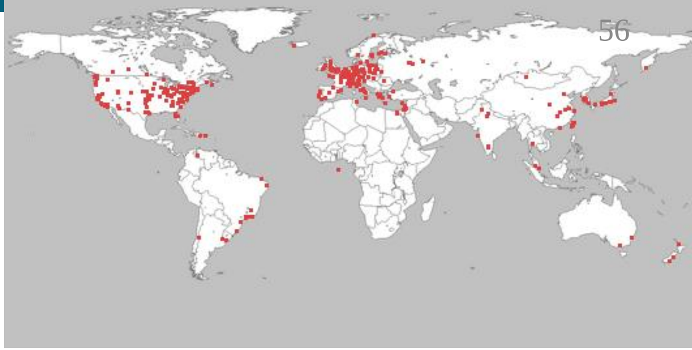
- **Community Cloud**
  - Governmental organizations
- **Public Cloud**
  - Marketed based on
    - Resources offered, availability, security, price
- **Local/Private Cloud**
  - Cloud architectures tailored to an organization's needs.
- **Hybrid Cloud**
  - Combination of public and local cloud resources.

nic Clouds: Emulab

- A community resource open to researchers in academia and industry. Very widely used by researchers everywhere today.
- https://www.emulab.net/
- A cluster, with currently ~500 servers
- Founded and owned by University of Utah (led by Late Prof. Jay Lepreau)

- As a user, you can:
  - Grab a set of machines for your experiment
  - You get root-level (sudo) access to these machines
  - You can specify a network topology for your cluster
  - You can emulate any topology

PLANETLAB
An open platform for developing, deploying, and accessing planetary-scale services



- A community resource open to researchers in academia and industry
- http://www.planet-lab.org/
- Currently, ~ 1077 nodes at ~500 sites across the world
- Founded at Princeton University (led by Prof. Larry Peterson), but owned in a federated manner by the sites

All images © PlanetLab

- Node: Dedicated server that runs components of PlanetLab services.
- Site: A location, e.g., UIUC, that hosts a number of nodes.
- Sliver: Virtual division of each node. Currently, uses VMs, but it could also other technology. Needed for timesharing across users.
- Slice: A spatial cut-up of the PL nodes. Per user. A slice is a way of giving each user (Unix-shell like) access to a subset of PL machines, selected by the user. A slice consists of multiple slivers, one at each component node.
- Thus, PlanetLab allows you to run real world-wide experiments.
- Many services have been deployed atop it, used by millions (not just researchers): Application-level DNS services, Monitoring services, CoralCDN, etc.
- PlanetLab is basis for NSF GENI https://www.geni.net/

Public Research Clouds

- Accessible to researchers with a qualifying grant
- Chameleon Cloud: https://www.chameleoncloud.org/
  - HaaS
  - OpenStack (~AWS)
- CloudLab: https://www.cloudlab.us/
  - Build your own cloud on their hardware

- http://www.cloud-lounge.org/EN/clouds-and-grids-compared.html