Cloud middleware

Part1: What a cloud middleware is good for ?

Sergio Maffioletti

S3IT: Service and Support for Science IT,
University of Zurich
http://www.s3it.uzh.ch/

Part1: content

1. What is large scale computing ?
2. What is virtualization ?
3. What is cloud computing ?

1. What is large scale computing ?

We intend the processing of large amount of information in a controlled manner to produce knowledge.
Let's start with few examples . . .

Large Hadron Collider (LHC) @ CERN

Purpose

Demonstrate theories of high-energy physic (Higgs boson).

What it is

The world's largest particle collider

What it does

Allows collision of proton beams at 14 TeV

Volume of data produces

On average 600 million proton collisions/s.
Data flow: 700MB/s (post-filter),
20TB/day, 25PB/year

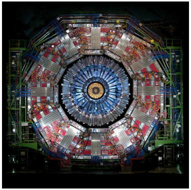


Figure : CMS detector © CERN

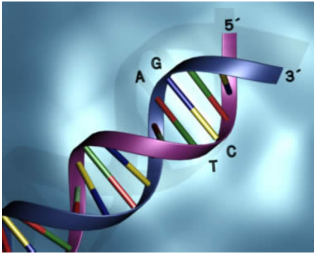
Genomics

Purpose

Study of the genomes of organisms.

What it is

DNA Sequencer



What it does

Determines the order of the four bases composing pieces of the sample.

Volume of data produces

60Gbp/day (high frequency)
Data flow: 50GB/Week

Trends in data grow: BigData

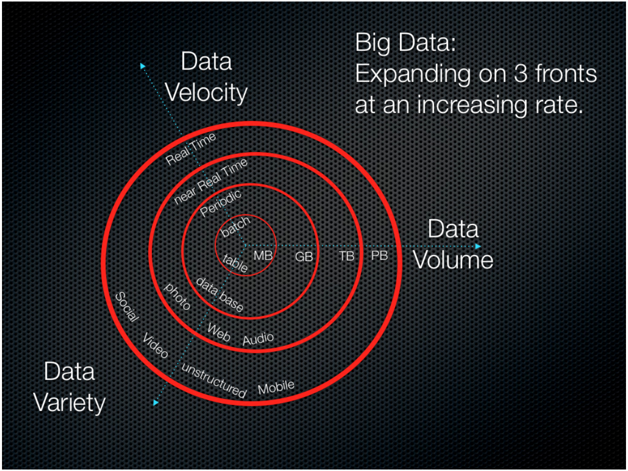


Figure : source: http://www.datasciencecentral.com

Data Volume

- Facebook has **2.5PB** of user data **+ 15 TB/day** (40 billion photos).
- Google processes **20PB** a day
- eBay has **65PB** of user data **+ 50TB/day**
- Human Brain project: simulate electrical signal propagation through the neurons.
  - Scale up number of neuron to **86 billions**.
  - Model data coarse grained up to **27PB**.
  - Model fine grained: up to **0.5EB**.

Data Velocity

- Twitter
  http://www.statisticbrain.com/twitter-statistics/
  - 500M tweets per day.
  - 9’100 tweets per second.
- YouTube statistics.
  http://www.youtube.com/t/press_statistics/
  - 100 hours of video uploaded to YouTube every minute
  - 6 billion hours of videos watched every month.

Data Variety

– Large variety of data representation formats due to new emerging data sources
– Pure text, photo, audio, video, web, GPS data, sensor data, relational data bases, documents, SMS, pdf, flash, etc etc etc.

What infrastructure are available ?

Supercomputing Centers

Figure : (Swiss National Supercomputing Center. http://cscs.ch)

- Monte Rosa: 402 TFlops system with 47872 compute cores and 46TB memory available.
- Tuned for High Performance Computing (HPC)
- Tightly coupled (either MPP or large SMP)

Supercomputing Centers

- **Large-scale problems**?
  - scale up hardware
  - Provide bleeding-edge solutions (top performances)
  - Very expensive investments (normally at 3year cycle)
- **Trend**: centralization of computing resources in large data centers
- **Important Issues**:
  - Redundancy, Efficiency, Utilization, Management

Large Scale Computing on Supercomputing Centers

- The size of your solution depends on the size of the single system you can use
- Applications and data need to be adapted to the system (OS, compilers, programming environment, tools)
- Very specialized system (portability is not taken into account)
- Simple primitives for communication and synchronization (MPI, OpenMP)

Grid

Figure : European Grid Initiative http://egi.eu

– Resources owned by independent institutions, federated together forming a Europe-wide distributed facility
– Tuned for High Throughput Computing (HTC)
– Middleware is required to aggregate geographically distributed resources
– Present a uniform data-processing infrastructure
– Nominally among the largest computing system available (e.g. WLCG is about 300k cores)

Grid

- **Large-scale problems**? aggregate resources geographically distributed
- **Trend**: de-centralization of computing resources, centralization of management and coordination
- **Important Issues**:
  - Need agreed standards to guarantee uniform access to resources
  - Grid middleware are not trivial (Authorization and Authentication, resource description and access, ...)
  - Resource allocation and accounting that span across institutions and user communities

Large scale computing on Grid

- Fault-tolerance: very large variety of possible failures that needs to be addressed
- Shared resources
- Users and Application developers need an understanding of how the underlying infrastructure works
- Applications and data need to be adapted to the system (OS, compilers, programming environment, tools)
- In this case is even worst because one has to consider an heterogeneous system

Cloud

- Commercial provider
- Utility Computing
- Virtualization and pay as you go as core concepts

Cloud

- **Large-scale problems**?
  - Veeeery large data centers
  - keeping infinite resource availability perception
- **Trend**:
  - User is provided with full control of his/her environment
  - No restriction on when, how, and how much (provided one can pay)
- **Important Issues**:
  - Virtualization as a mechanism to compartmentalize applications
  - Virtualise computing (e.g. EC2) and storage (e.g. S3)
  - Provide integration tools (APIs, Platform services)
  - Give me your credit card. . .

Large Scale Computing on Cloud

- The size of your solution depends on the size of your bank credit
- Provided one could negotiate with cloud provider, it is possible to scale up indefinitely
  - The business model behind commercial clouds makes them very flexible
  - dynamically scaling (elasticity)

Large Scale Computing on Cloud

- The size of your solution depends on the size of your bank credit
- Provided one could negotiate with cloud provider, it is possible to scale up indefinitely
  - The business model behind commercial clouds makes them very flexible
  - dynamically scaling (elasticity)

Large Scale Computing on Cloud

- Virtualization allows to customize the host environment to fit application and data
- Portability is granted by VM portability
- Users and Application developers do not need an understanding of how the underlying infrastructure works
- Performance penalty is, most of the time, outweighed by higher flexibility and manageability

2. What is virtualization?

"A virtual machine is taken to be an *efficient, isolated duplicate* of a real machine."

Reference: G. J. Popek and R. P. Goldberg (1974): "Formal Requirements for Virtualizable Third Generation Architectures" *Communications of the ACM* 17 (7): 412–421.

Virtualization characteristics

Fidelity Software on the VM exhibits “essentially identical” effects to its execution on hardware.

Efficiency “A statistically dominant subset of the virtual process instructions are executed directly by the real processor, with no software intervention.”

Resource control The VM control software manages hardware allocation, “it is not possible for a program running in the VM to access any resource not explicitly assigned to it.”

Why virtualization?

Execution environments with resource limits and/or resource guarantees.

– Provide secure, isolated sandboxes for running untrusted applications.

Make systems independent of the hardware.

– Run legacy applications.
– Provide binary compatibility.

Co-locate and consolidate independent workloads.

– Run multiple operating systems.

Treat application suites as appliances by “packaging” and running each in a virtual machine.

Why virtualization in IaaS clouds

Execution environments with resource limits and/or resource guarantees.

- Provide secure, isolated sandboxes for running untrusted applications.

Make systems independent of the hardware.

- Run legacy applications.
- Provide binary compatibility.

Co-locate and consolidate independent workloads.

- Run multiple operating systems.

Treat application suites as appliances by “packaging” and running each in a virtual machine.

Pros of using Virtualization

Server Consolidation

Testing and development

Dynamic Load Balancing

Disaster Recovery

Reduction in cost of infrastructure

Cons of using Virtualization

Magnified physical failures

Degraded performance

Complex root cause analysis

New management tools

Classical virtualization, I

(Popek and Goldberg's model of a "third-generation" machine.)

The processor has two modes of operation: **supervisor mode** and **user mode**. Only a *subset* of all machine instructions is available in user mode.

Memory addressing is:

– *relative* to a relocation register

– *limited* by a bounds register

Classical virtualization, II

An instruction is **privileged** iff it can only be executed in supervisor mode.

An instruction is **control sensitive** iff it attempts to alter the configuration of machine resources (e.g., change the memory bounds register).

An instruction is **behavior sensitive** iff its effect depends on the configuration of resources (e.g., the memory bounds or relocation register).

Classical virtualization, III

An instruction is **trapping** if its execution causes a jump to supervisor code at a fixed location.

Theorem (Popek + Goldberg): **If all sensitive instructions are trapping, then a machine can be virtualized.**

Basic idea: execute all instructions in user mode; the sensitive ones will trap and special supervisor code (the *VMM* or *control program*) can emulate the requested behavior.

Note: Popek and Goldberg's Theorem is a *sufficient* condition, but not a *necessary* one.

Solution: Hardware-assisted virtualization

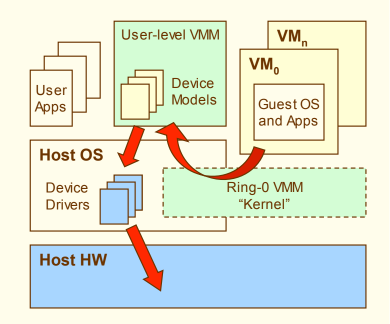
About 2005, Intel and AMD introduced extensions to the x86 instruction set to enable classical virtualization.

The CPU has an *additional* protection bit, selecting **host** or **guest** mode.

Code in *host supervisor* mode can set which instructions in *guest supervisor* mode should trap and on which occasion.

The virtualized OS runs in *guest supervisor* mode, and will not notice any difference with a non-virtualized CPU.

Solution: Paravirtualization



Paravirtualization relies on a *modified* guest OS: forward hardware access to host OS.

Notable examples in Linux are Xen and User-Mode Linux.

(Diagram stolen from Smith and Uhlig's presentation http://www.hotchips.org/archives/hc17/1_Sun/HC17.T1P2.pdf)

KVM: HW-assisted virtualization in Linux

Linux kernel infrastructure to leverage Intel or AMD virtualization extensions to run virtual machines.

Included into mainstream Linux kernel since version 2.6.20.

Reference: A. Kivity, Y. Kamay, D. Laor, U. Lublin, A. Liguori: "**`kvm`**: the Linux Virtual Machine Monitor", Ottawa Linux Symposium (July 2007), pp. 225–230

What do you need to run Virtual Manchines ?

1. Hypervisor
2. Image file
3. Network configuration
4. Virtual Block Bevice

What do you need to run Virtual Manchines ?

1. Hypervisor
   - KVM, VirtualBox, Xen, HypeV, VMWare, . . .
   - A program that allows multiple operating systems to share a single hardware host.
2. Image file
3. Network configuration
4. Virtual Block Bevice

What do you need to run Virtual Manchines ?

1. Hypervisor
2. Image file
   - An image file stores the contents of the virtual machine's hard disk drive (base OS).
3. Network configuration
4. Virtual Block Bevice

What do you need to run Virtual Manchines ?

1. Hypervisor
2. Image file
3. Network configuration
   – Allows to connect virtual machines to the external network.
4. Virtual Block Bevice

What do you need to run Virtual Manchines ?

1. Hypervisor
2. Image file
3. Network configuration
4. Virtual Block Device
   – A remote storage device accessed using, for example, (iSCSI) protocol.
   – From within the VM, the remote storage device will be seen as a regular block device (can create filesystem and mount it).

3. What is cloud computing ? (Wikipedia)

It is a paradigm shift whereby details are abstracted from the users who no longer need knowledge of, expertise in, or control over the technology infrastructure "in the cloud" that supports them. [1]

---

[1] Wikipedia

What is cloud computing ? (NIST)

Cloud computing is 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. [2]

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

Fundamental components for a data processing infrastructure

Computing: the Virtual Appliances where application will run

Image Repository: Where Virtual Appliances wil be stored

Storage: object storage used for persistent data (e.g. computation resutls)

Identity management: authentication system to access Virtual Appliances and deployed services

Fundamental components for a data processing infrastructure

Computing: the Virtual Appliances where application will run

Image Repository: Where Virtual Appliances wil be stored

Storage: object storage used for persistent data (e.g. computation resutls)

Identity management: authentication system to access Virtual Appliances and deployed services

Fundamental components for a data processing infrastructure

Computing: the Virtual Appliances where application will run

Image Repository: Where Virtual Appliances wil be stored

Storage: object storage used for persistent data (e.g. computation resutls)

Identity management: authentication system to access Virtual Appliances and deployed services

Fundamental components for a data processing infrastructure

Computing: the Virtual Appliances where application will run

Image Repository: Where Virtual Appliances wil be stored

Storage: object storage used for persistent data (e.g. computation resutls)

Identity management: authentication system to access Virtual Appliances and deployed services

Fundamental components for a data processing infrastructure

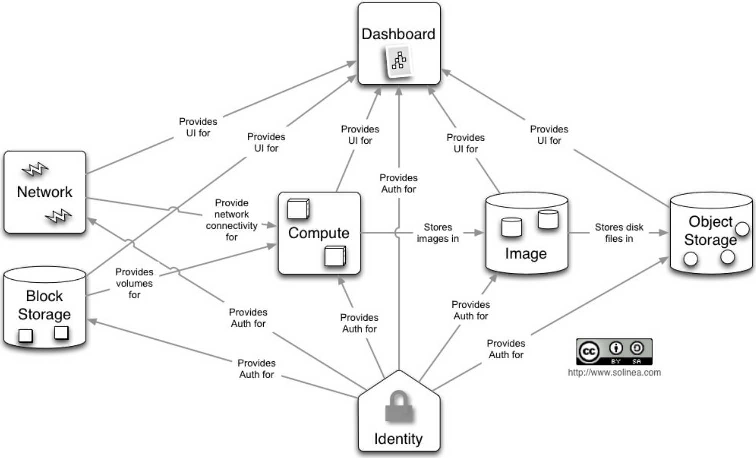
Computing: the Virtual Appliances where application will run

Image Repository: Where Virtual Appliances wil be stored

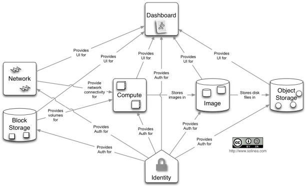
Storage: object storage used for persistent data (e.g. computation resutls)

Identity management: authentication system to access Virtual Appliances and deployed services

Prototype of a cloud: the OpenStack example



Prototype of a cloud: the OpenStack example



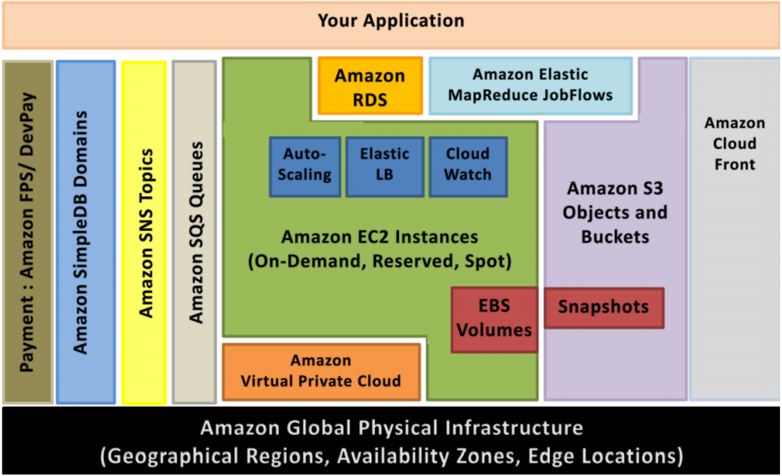
Computing provisioning: provision and manage large networks of virtual machines

Storage blocks: create redundant, scalable object storage

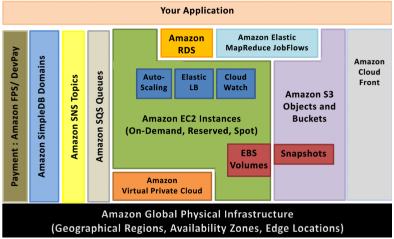
Image repository: discovery, registration, and delivery services for virtual disk images

Identity management: authentication system across the cloud operating system

Prototype of a cloud: the Amazon example

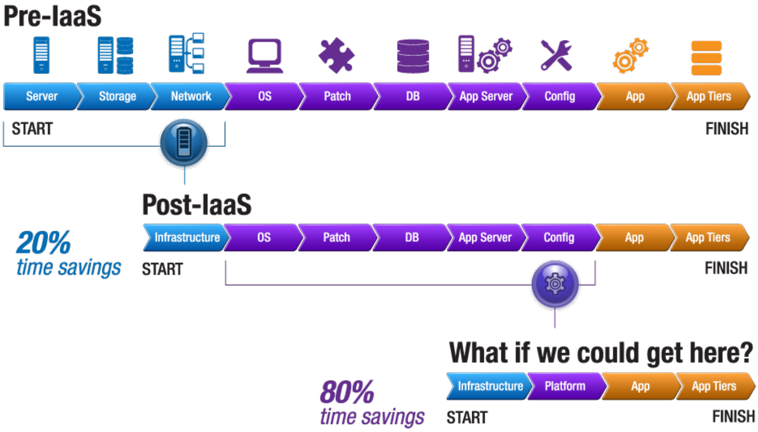


Prototype of a cloud: the Amazon example



Computing provisioning: Elastic Compute Cloud (EC2)
Storage blocks: Elastic Block Storage (EBS) and Simple Storage Service (S3)
Image repository: Amazon Machine Image (AMI) are stored on S3 buckets
Identity management: Identity and Access Management (IAM)

Classification of cloud provisioning



Infrastructure as a Service (IaaS)

Provisions and manages the physical processing, storage, networking and the hosting environment and cloud infrastructure.

- Fully outsourced service so businesses do not have to purchase servers, software or equipment.

Infrastructure providers can dynamically allocate resources for service providers.

Users have to create/install, manage and monitors services for IT infrastructure operations.

Examples: Amazon's EC2, RightScale, OpenStack, OpenNebula.

Classification of cloud provisioning

Platform as a Service (PaaS)

Provisions and manages cloud infrastructure and middleware; provides development, deployment and administration tools.

Infrastructure providers can transparently alter the platforms for their customers' unique needs.

Users have to develop, test, deploy and manage applications hosted in a cloud environment.

Example: Google App Engine, SalesForce.

Classification of cloud provisioning

Software as a Service (SaaS)

Defined as service-on-demand, where a provider will license software tailored.

Installs, manages, maintains and supports the software application on a cloud infrastructure.

Infrastructure providers can allow customers' to run applications off their infrastructure, but transparent to the end user.

Users interact with application/service for process operations.

Example: Gmail, Facebook, Flickr, . . .

Types of Cloud infrastructure

Public Cloud
Private Clouds
Hybrid Cloud

Types of Cloud infrastructure

Public Cloud

- The cloud infrastructure is made available to the general public or a large industry group and is owned by an organization selling cloud services.
- Only operational expenses
- No control on cloud stack, dependency on external partner

Private Clouds
Hybrid Cloud

Types of Cloud infrastructure

Public Cloud

Private Clouds

- The cloud infrastructure is operated solely for an organization. It may be managed by the organization or a third party and may exist on premise or off premise.
- Owner organization provides cloud services to his own customers
- Full control on cloud stack, accounting, allocation

Hybrid Cloud

Types of Cloud infrastructure

Public Cloud
Private Clouds

Hybrid Cloud

- The cloud infrastructure is a composition of two or more clouds (private and public) that remain unique entities but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load balancing between clouds).
- Constraints on own cloud stack: needs to interoperate with public cloud