Virtualization, resource management and autonomous systems

Gastón Keller
PhD Candidate

The University *of* Western Ontario

Overview

- Virtualization
- Uses of virtualization
- Virtualization in data centers
- VM replication
- VM memory management

Overview

- ***Virtualization***
- Uses of virtualization
- Virtualization in data centers
- VM replication
- VM memory management

Old Concept

1960s: IBM designed the operating system CP-40/CMS.

Provided VMs that were indistinguishable from real machines by user programs.

Role in Industry

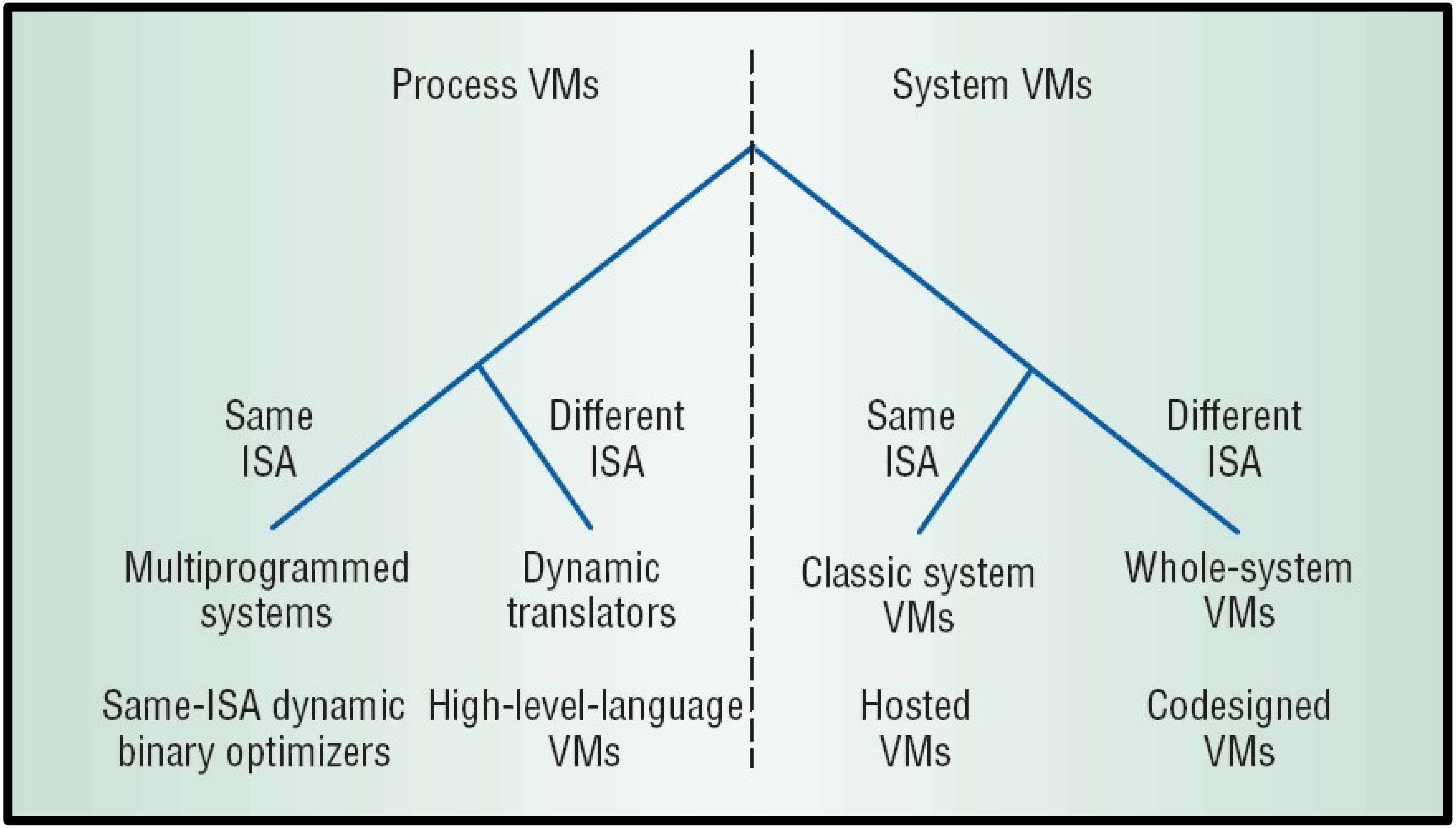
Level of adoption around 20 to 30% (expected to grow another 20%). *

Principal motivations:

- cost-cutting
- business continuity
- server manageability

* Frank Gillett, *Forrester Research*. Keynote talk at VMware Virtualization Forum 2008.

Taxonomy of VMs

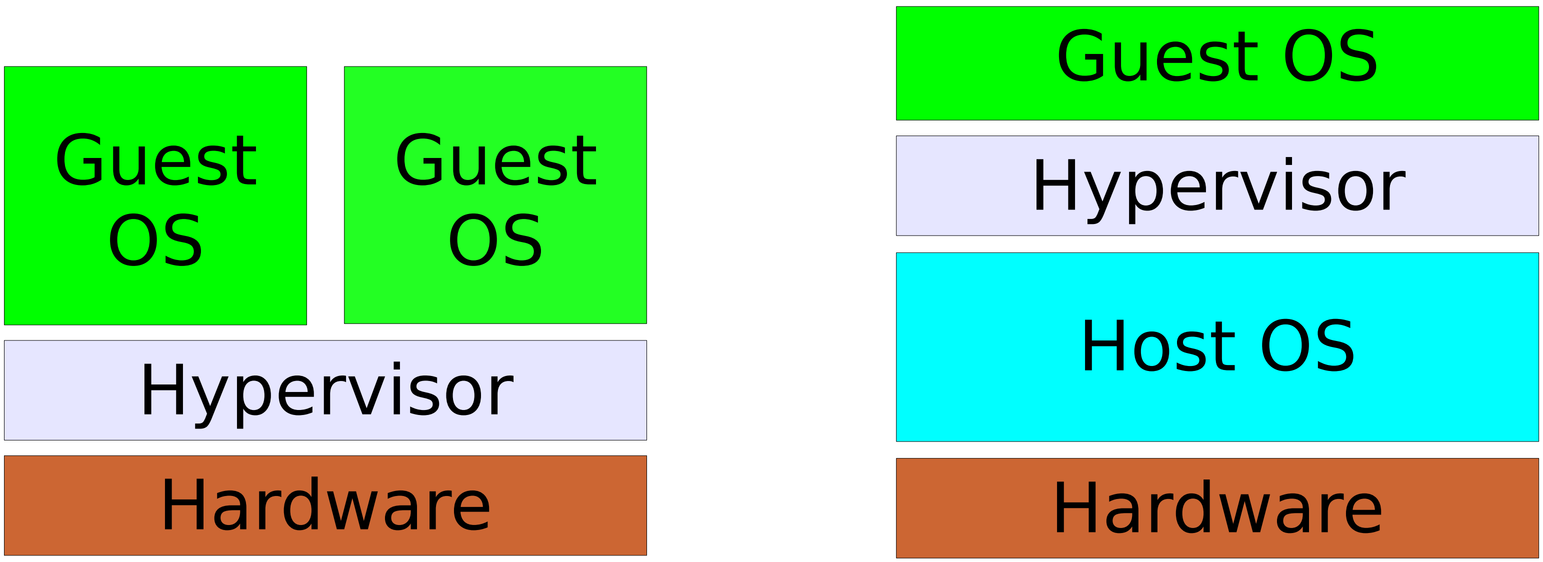


Source: Smith, J. E.; Nair, R.; *The Architectures of Virtual Machines.*

Definition

"*(System) Virtualization is a software technique that enables the simultaneous execution of multiple computer systems in one physical machine.*"

Hypervisor-based Virtualization

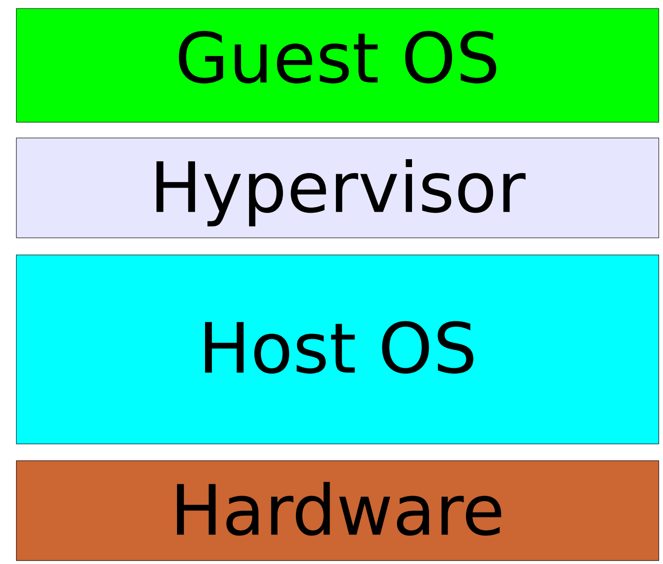


Type 1 Hypervisor

Type 2 Hypervisor

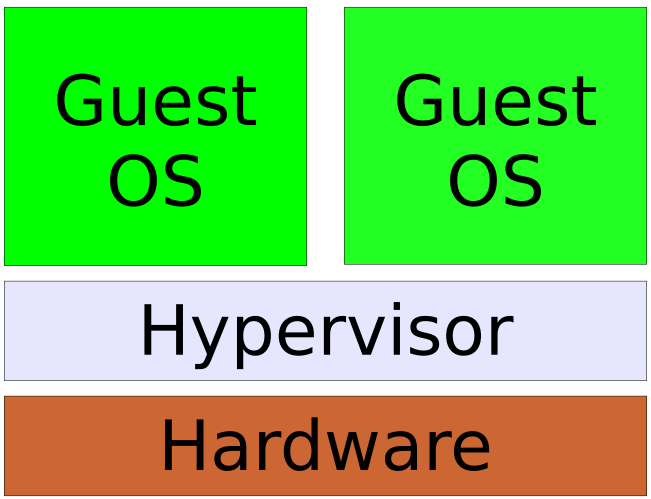
Type 2 Hypervisor

Interprets the code of the guest OS and its applications.



Type 2 Hypervisor

Type 1 Hypervisor

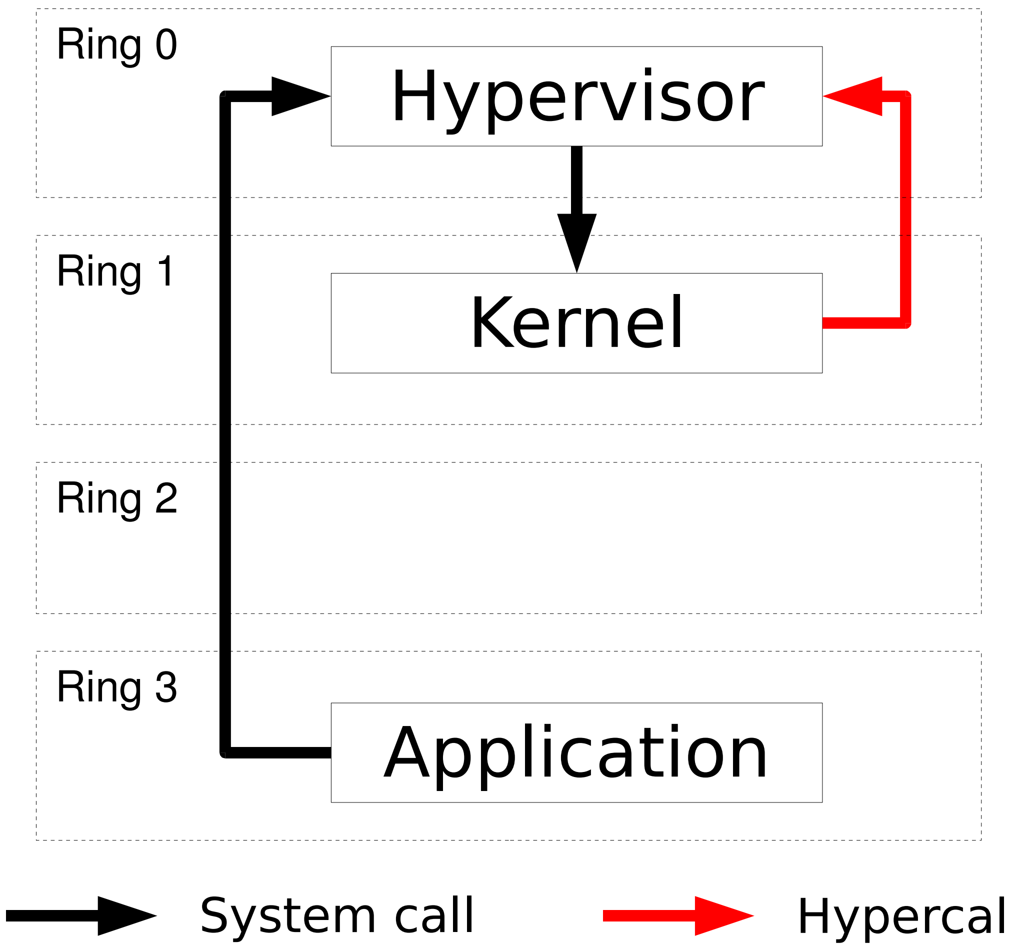


It can interpret the code of the guest OS as a Type 2 Hypervisor or use *paravirtualization*.

Type 1 Hypervisor

Paravirtualization

The guest OS source code is modified to enable communication with the hypervisor.



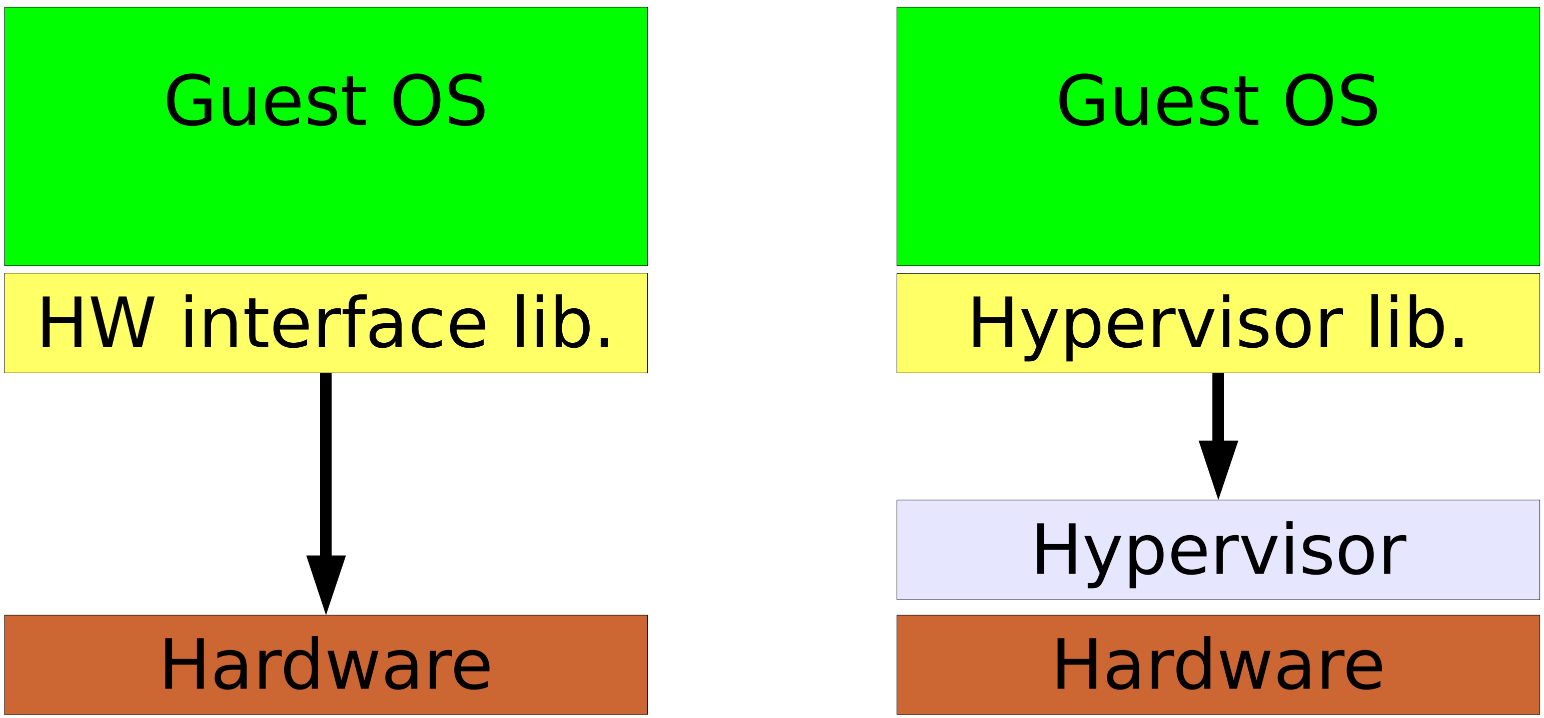
Full-virtualization

In 2006, Intel and AMD released CPUs with support for virtualization.

Full-virtualization enables unmodified guest OSes to run in VMs.

However, *paravirtualization* still offers better performance.

Evolution of paravirtualization



** Check out paravirt_ops in Linux.*

Xen

Started as a research project at the University of Cambridge (2003).

Is Open Source Software (which contributed to its success).

Supported by big companies such as IBM, Intel, and Oracle (among others).

Resource Management

(CPU) *Credit Scheduler*:

- weight value
- cap value

Memory:

- reservation value
- maximum value
- minimum value (Dom0)

Resource Management...

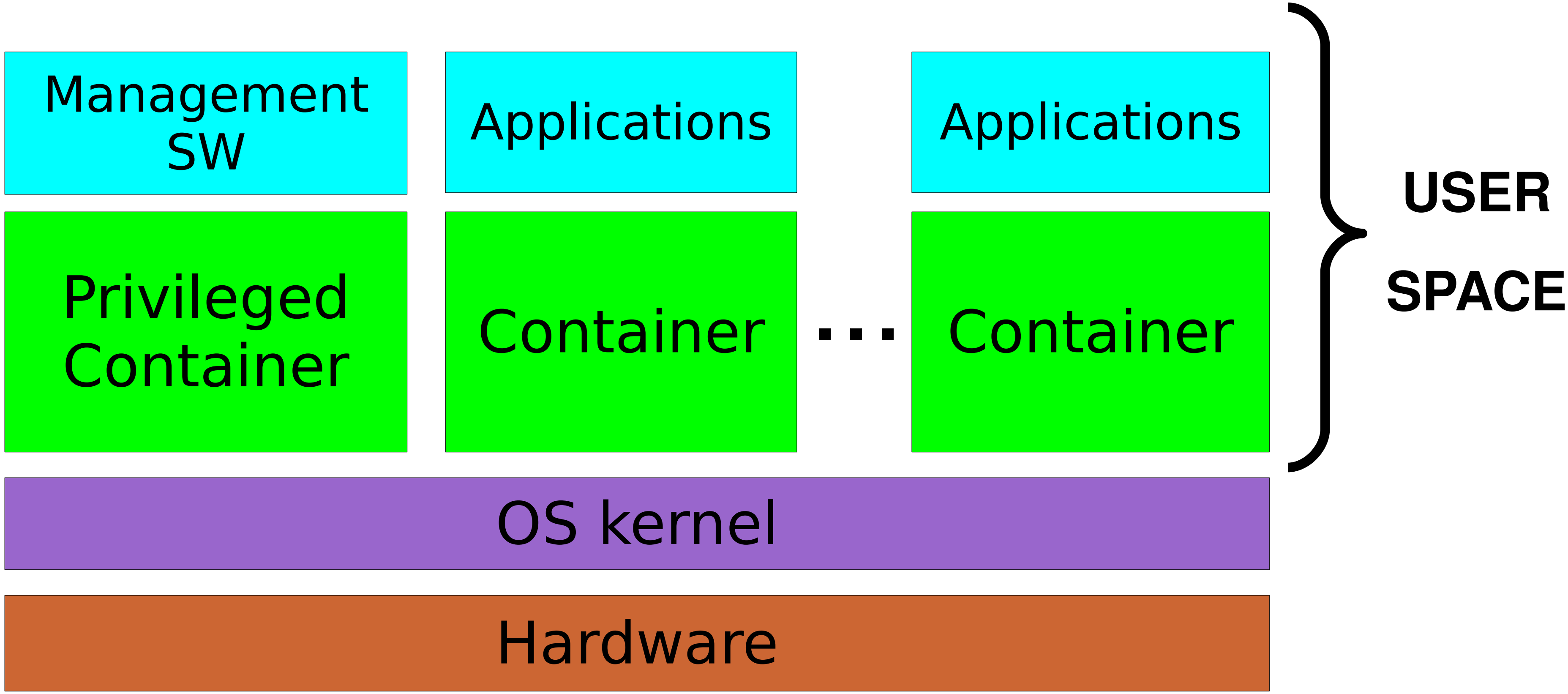
Networking:

- bridging
- routing

To reduce CPU overhead:

- dedicated NICs
- optimized inter-domain communication channel *(Dom0-DomU)*
- virtualization aware NICs

OS-level Virtualization



OpenVZ server virtualization

Linux VServer

solaris

OpenVZ

Provides *operating system-level virtualization*.

Modified Linux to run multiple, isolated *containers*.

Basis of Parallels Virtuozzo Containers.

OpenVZ

Advantages:

- simple deployment
- close to native performance
- great scalability

Disadvantage:

- only GNU/Linux-based virtual environments

Resource Management

Organized in two levels:

- storage subsystem
- CPU scheduler
- I/O scheduler

Memory:

- *User Beancounters* (per container)

Overview

- Virtualization
- ***Uses of virtualization***
- Virtualization in data centers
- VM replication
- VM memory management

Virtual Appliances

A software image containing a software stack (OS + app.) designed to run inside a virtual machine.

Makes software deployment easier and faster.

BitNami - http://bitnami.org/
VMware - http://www.vmware.com/appliances/

HPC

Potential benefits:

- resiliency
- scaling
- system-level portability
- observability

Snowflock - http://sysweb.cs.toronto.edu/snowflock

Palacios - http://www.v3vee.org/palacios/

Grid Computing

Virtualization provides Grid Computing with *isolated, customized environments*.

In addition:

- legacy systems
- security
- flexible resource allocation

Grid Computing…

Krsul et al. developed ***VMPlant*** Grid service: flexible and efficient resource sharing through virtualization.

Components:

- ***VMShop*** (front-end)
- ***VMPlants*** (hosts)

Grid Computing…

Emeneker and Stanzione developed ***Dynamic Virtual Clustering***: leverage an institution's clusters computing power through job forwarding and spanning.

VMs provided independence from the platform and encapsulation.

Overview

- Virtualization
- Uses of virtualization
- ***Virtualization in data centers***
- VM replication
- VM memory management

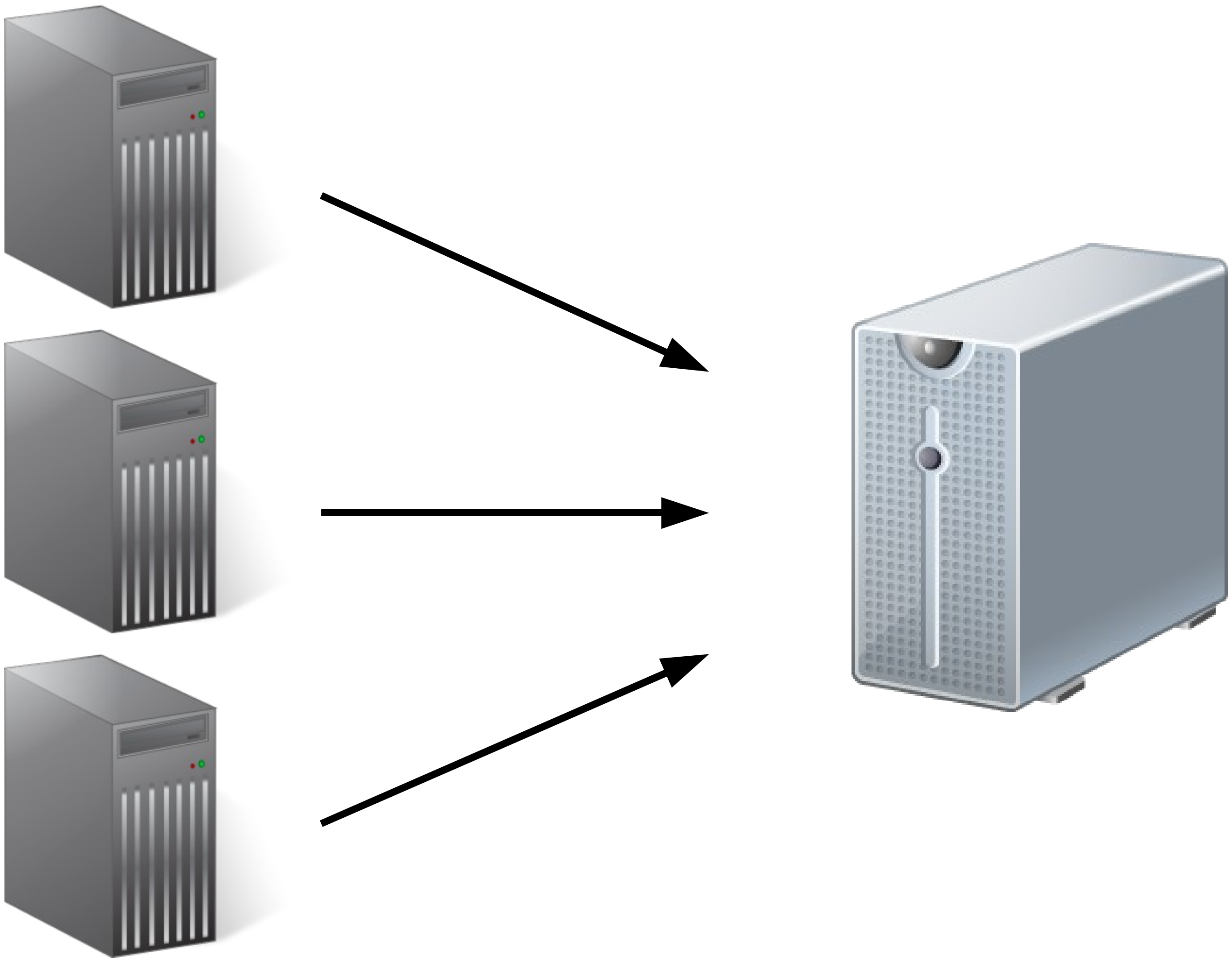
Definition

"*A **data center** is a collection of computing resources shared by multiple applications concurrently in return for payment by the application providers, on a per-usage basis, to the data center provider.*"

Data Centers

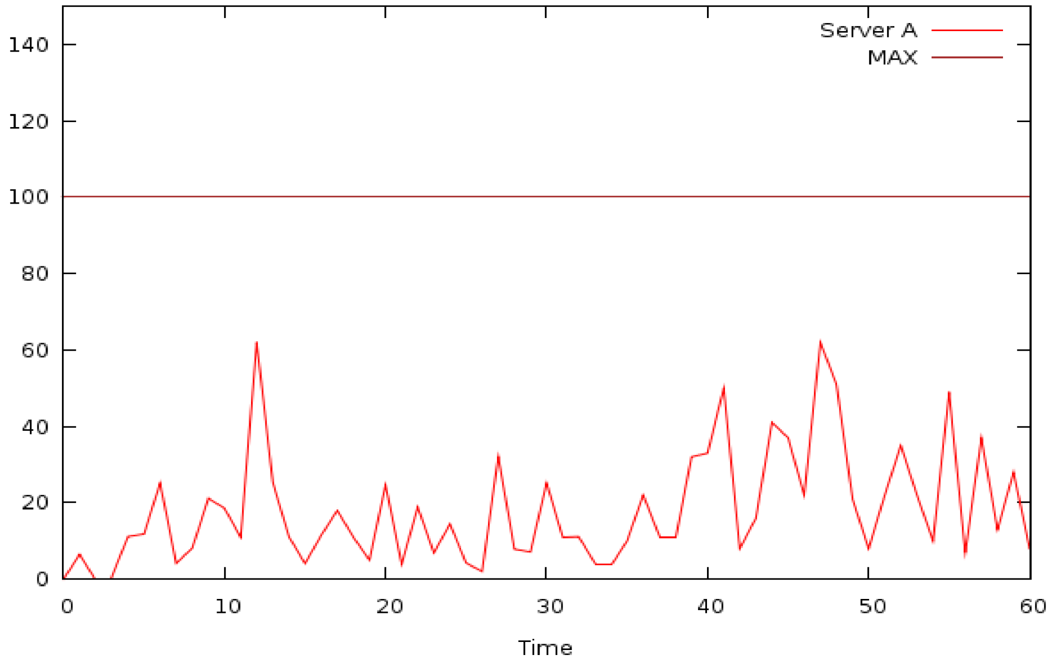
(*) http://scienceblogs.com/goodmath/2009/05/cloud_computing.php

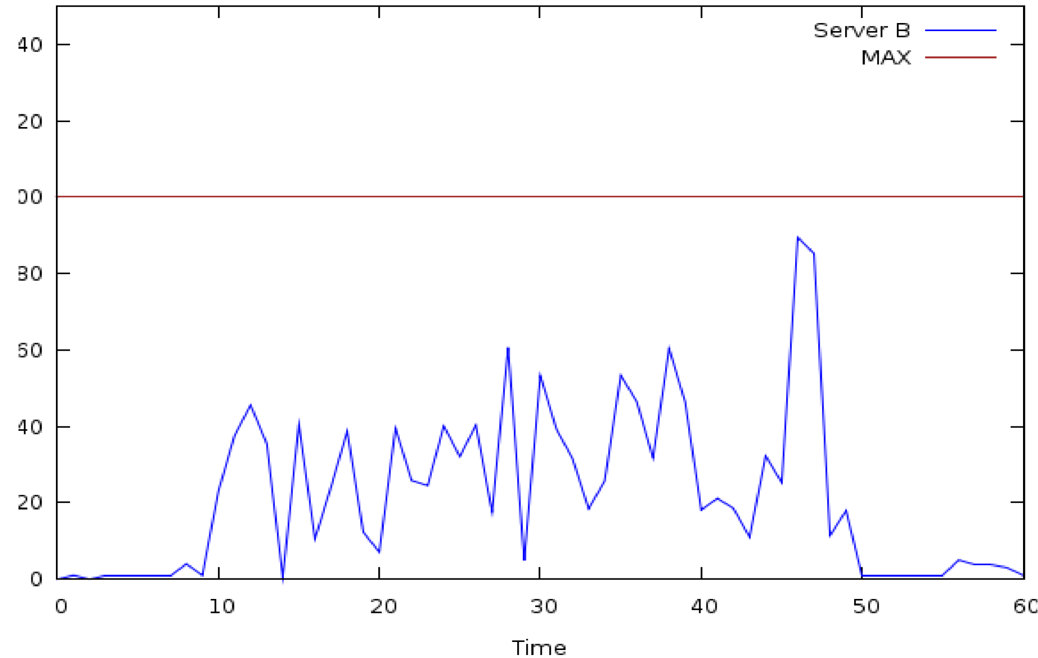
Server Consolidation

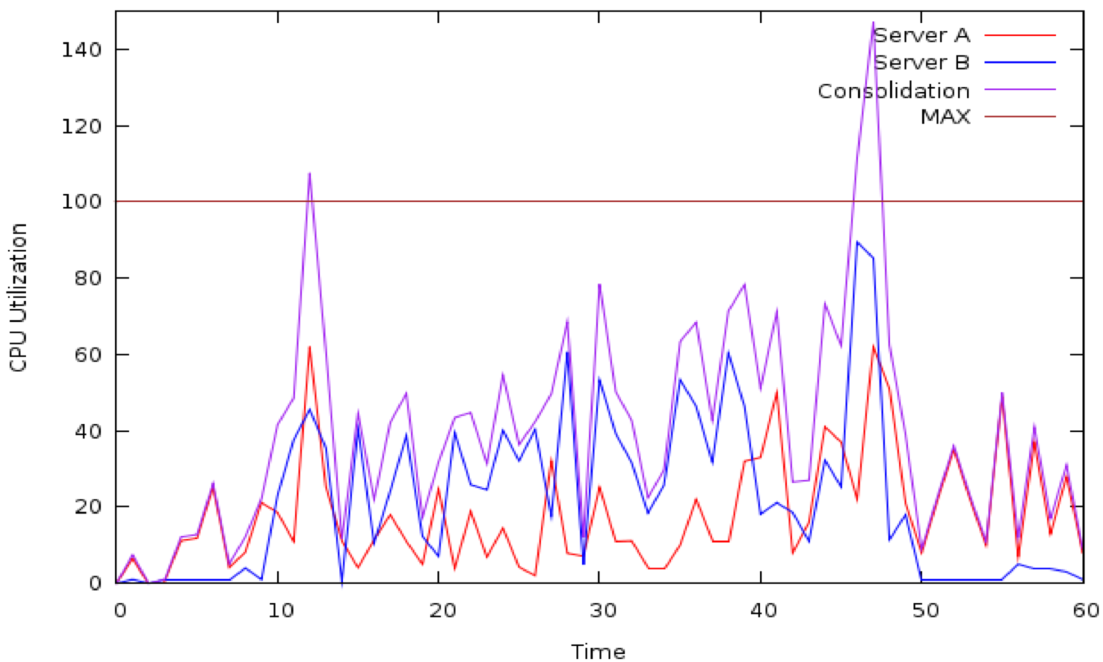


(*) http://www.visualpharm.com/

Resource Stress Situations







Challenges

Virtualization brings benefits to the data center, but also challenges.

The research literature shows that unsolved issues are abound...

Challenges...

- Resource monitoring
- Algorithms and policies
- Resource management systems
- VM migration process
- Management tools

Resource Monitoring

Wood et al. studied two approaches to monitoring:

- ***black-box***, and
- ***grey-box***.

Sandpiper used the data to detect ***hotspots*** and migrate VMs.

Algorithms and Policies

Gmach et al. studied workload consolidation through VM migration.

Developed:

- ***placement controller***, and
  - *(multiple policies)*
- ***migration controller***.
  - *(multiple thresholds)*

Resource Mgmt. Systems

Zhu et al. developed a hierarchy of controllers:

- ***node controller***,
- ***pod controller***, and
- ***pod set controller***.

Enable client and system admins to focus on policy setting.

VM migration process

Zhao and Figueiredo analyzed the VM migration process:

- in parallel,
- in sequence,
- cpu-intensive app.,
- mem-intensive app.

Predict time and performance of the VM migration process.

Management Tools

Vallée et al. extended OSCAR, a toolkit for cluster installation, configuration and management.

OSCAR-V enabled deployment and management of host OSes and VMs.

Challenges...

These were just a few research challenges that came with virtualization.

There are many more to be studied.

Overview

- Virtualization
- Uses of virtualization
- Virtualization in data centers
- ***VM replication***
- VM memory management

VM Replication

Migration: moving a VM from one host node to another.

Replication: instantiating a copy of a VM in a different host node:

- copy of **current** VM, or
- instance of a **stored** image.

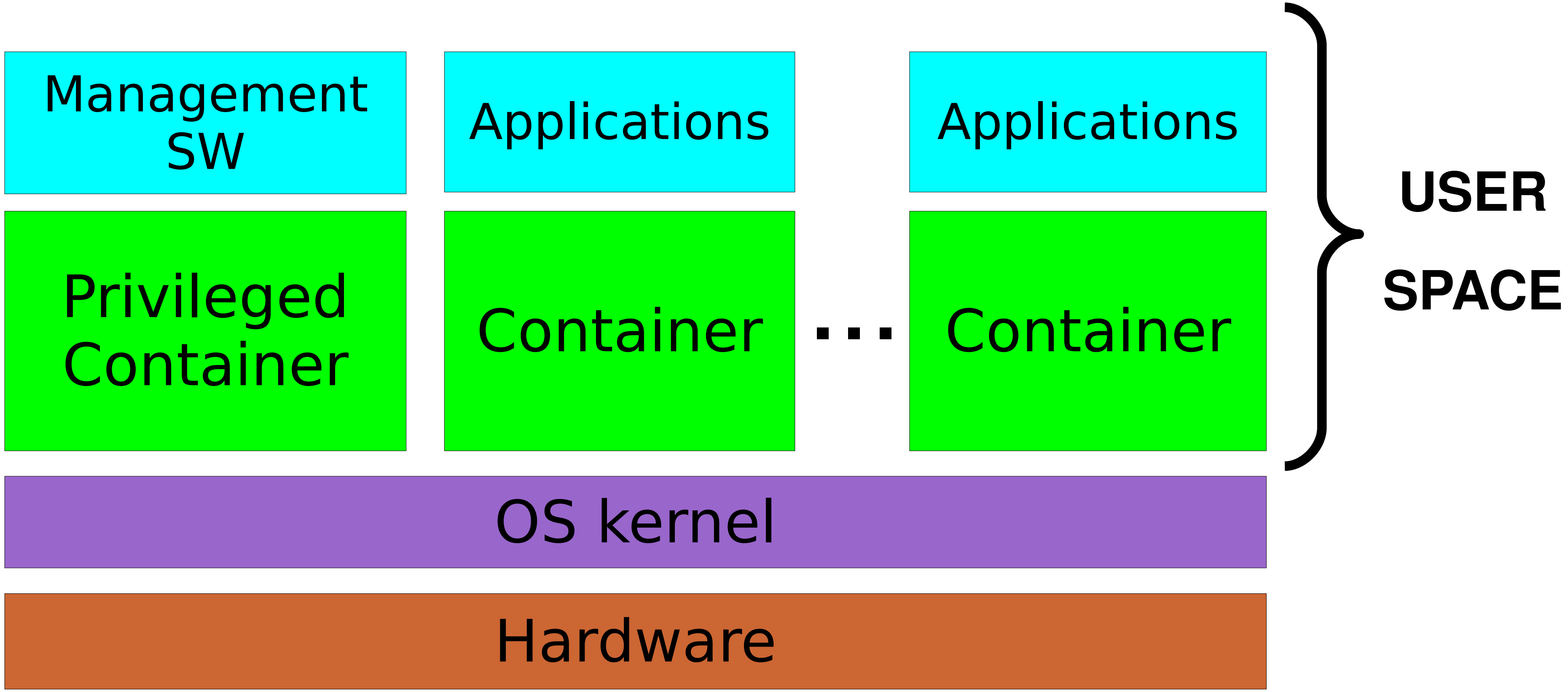
Our Work

- Built *Golondrina*, a resource management system for *OS-level virtualized* environments.

- Implemented ***replication*** mechanism to deal with resource stress situations

- Compared ***replication*** mechanism with ***migration*** mechanism

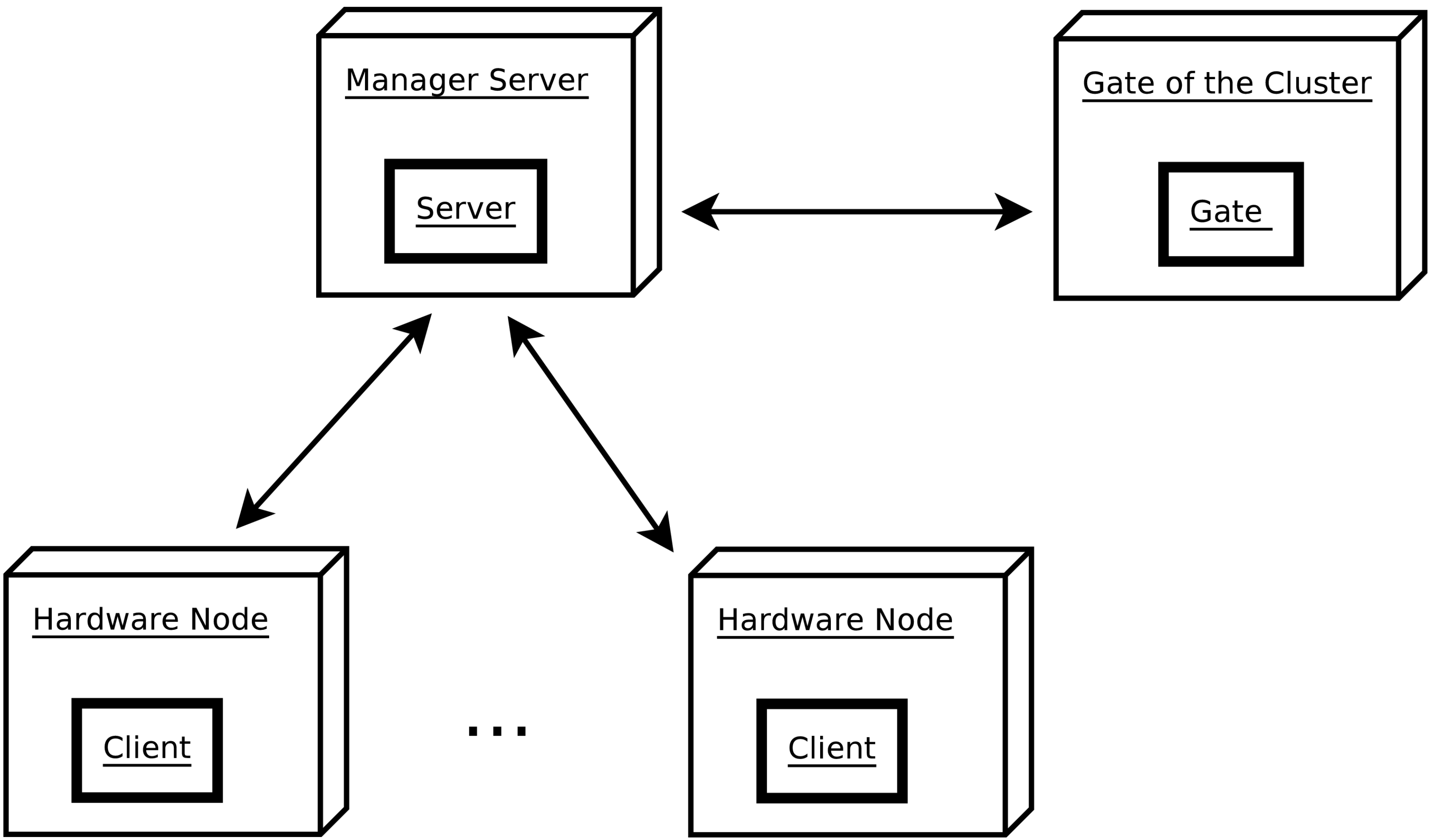
OS-level Virtualization



Our Work…

- Why use ***migration*** and ***replication*** to do resource management?
- Are both mechanisms needed? Why compared them?

Golondrina


Manager Server
Server
Gate of the Cluster
Gate
Hardware Node
Client
...
Hardware Node
Client

Implementation

OS: CentOS 5.2 / OpenVZ

Prog. Lang.: Python

Communications: Twisted (event-driven networking engine)

Load Balancer: Pound

Basic Responsibilities

(C) Gather CPU statistics

(S) Process Clients' statistics

(S) Search for resource stress situations

(S) Determine sequence of relocations

(C) Execute migration/replication

Basic Responsibilities

(C) Gather CPU statistics

(S) Process Clients' statistics

(S) *Search for resource stress situations*

(S) *Determine sequence of relocations*

(C) *Execute migration/replication*

Resource Stress Check

- periodic check on (almost) every hardware node

$$overloaded \leftarrow \underbrace{(\sum_{i=0}^{n}(\hat{u}_i > threshold) \geq k)}_{\textit{\textbf{K} out of the previous \textbf{N} checks were in excess}} \wedge \underbrace{(\hat{u}_{t+1} > threshold)}_{\textit{next predicted CPU utilization is in excess}}$$

Relocation Algorithm

```
decreasingLoadSort(stressed_HNs)
increasingLoadSort(non-stressed_HNs)
for each HN in stressed_HNs:
    decreasingLoadPolicy(containers)
    While HN is stressed:
        pick a CT and cycle through
        non-stressed_HNs until finding a HN_2
        that can host the CT
```

Replication Algorithm

1. generate CTID for the replica
2. bring CT image from central repository
3. process image
4. edit image configuration file
5. start replica

Experiments

- Cause resource stress situations (*httperf* – load generator)

- Configure *Golondrina* to react:
  - doing nothing
  - using replication
  - using migration

- Measure lost requests and throughput (*Apache web servers*)

Experiment 1

- 2 hardware nodes *(bravo02, bravo03)*
- 2 containers *(A, B)*
- *A* receives a load of around 70%
- *B* receives a load of around 105%
- *bravo02* experiences a load of 175%

Results Exp. 1

Web Server's Effectiveness			
Servers	Nothing	Replication	Migration
one.com	100.00%	99.11%	100.00%
two.com	100.00%	98.44%	100.00%

Throughput provided no conclusive results.

Experiment 2

- 2 hardware nodes *(bravo02, bravo03)*
- 2 containers *(A, B)*
- *A* and *B* receive a load of around 105%
- *bravo02* experiences a load of 200%

Results Exp. 2

Web Server's Effectiveness			
Servers	Nothing	Replication	Migration
one.com	77.55%	87.55%	78.00%
two.com	62.44%	88.00%	84.44%

Throughput provided no conclusive results.

Experiment 3

- 2 hardware nodes (*bravo02*,*bravo03*)
- 4 containers (*A*, *B*, *C*, *D*)
- e*ach container* receives a load of around 51%
- *bravo02* experiences a load of 200%

Results Exp. 3

Web Server's Effectiveness			
Servers	Nothing	Replication	Migration
one.com	88.00%	97.00%	94.00%
two.com	92.00%	96.33%	94.00%
three.com	87.00%	96.33%	95.33%
four.com	98.00%	96.33%	93.66%

Throughput provided no conclusive results.

Analysis of Results

- Both ***replication*** and ***migration*** offer an improvement over taking ***no action*** upon detection of a resource stress situation.

- ***Replication*** offers a better improvement over ***migration***.

Overview

- Virtualization
- Uses of virtualization
- Virtualization in data centers
- VM replication
- ***VM memory management***

VM Memory Management

VMs are allocated ***min*** and ***max*** amounts of memory.

We want more dynamism.

We are extending *Golondrina* to allocate memory as needed.
(Work in progress.)

Remark

Virtualization is finding its way into many environments:

industry, academia, government,

HPC, Grid, Data Center...

Research topics are abound...

THANKS