Virtualization Overview

Yao-Min Chen

The new look of computing

Outline

- Intro to Virtualization (V14n)
- V14n and Cloud Computing
- V14n Technologies

What is Virtualization?

- "Pretending"
  - One thing pretends to be 2 or more things
  - E.g., Linux and Windows on the same system
- "Sharing"
  - You don't really own it but use it as if you own
  - E.g., virtual memory
- "Multiple Personalities"
  - A different OS, or application, or infrastructure, depending on when and where you need it

Recent Gartner Report

- STAMFORD, Conn., September 27, 2010 — More than 80 percent of enterprises now have a virtualization program or project, but only 25 percent of all server workloads will be in a virtual machine (VM) by year-end 2010
- ***"Virtualization will continue as the highest-impact issue challenging infrastructure and operations through 2015***
- "Virtualization now drives efficient IT from all angles, including ***data center design***, ***platform updates***, and application and infrastructure modernization, as well as traditional and ***new delivery models***, such as ***infrastructure utility*** and ***cloud computing***"

Why Virtualize?

- Efficient sharing of resources
  - Old problem: multi-user, multi-tasking
  - New problem: multi-core CPUs

Processor	Number of cores	Threads per core	Total threads
SPARC T3	16	8	128
Power7	8	4	32
Nehalem EX	8	2	16

More than just Cost Saving

- Faster deployments
- Reduced downtime
- Disaster recovery
- Usage accounting and chargeback
- Capacity planning
- Hosted virtual (thin client) desktops (HVDs)

Technology Enablers

- Hypervisors
  - VMware, Xen, KVM, Solaris Hypervisor, Hyper-V
- CPU
  - MMU enhancement - Intel VT, AMD-V
  - IOMMU – Intel VT-d, AMD-Vi
- Networking
  - Single Root I/O Virtualization (SR-IOV)
  - Virtual switching (vswitch)
- Storage
  - Network-attached storage (NAS)
  - Storage Area Networks (SAN)

Hypervisors

- Also known as VM Manager (VMM)
- Acquired the name from "supervisor"
  - Operating system is also known as "supervisor"
  - OS for v14n is thus called "hyper-supervisor," or just "hypervisor"
- Thin layer, providing arbitration between VMs and hardware, and task scheduling
- Repying on Dom0 VM and other guest VMs for hardware device drivers

CPU Virtualization Technology

- AMD Virtualization (AMD-V)
  - Originally known as Secure Virtual Machine (SVM)
  - 2nd generation technology is called Rapid Virtualization Indexing (RVI)
- Intel Virtualization Technology (Intel VT)
  - Extended Page Table (EPT) is Intel's adoption of RVI

CPU Virtualization - IOMMU

- Input/output memory management unit
- Enables guest virtual machines to directly use peripheral devices
  - Ethernet, accelerated graphics cards, hard-drive controllers, etc
- Through DMA and interrupt remapping

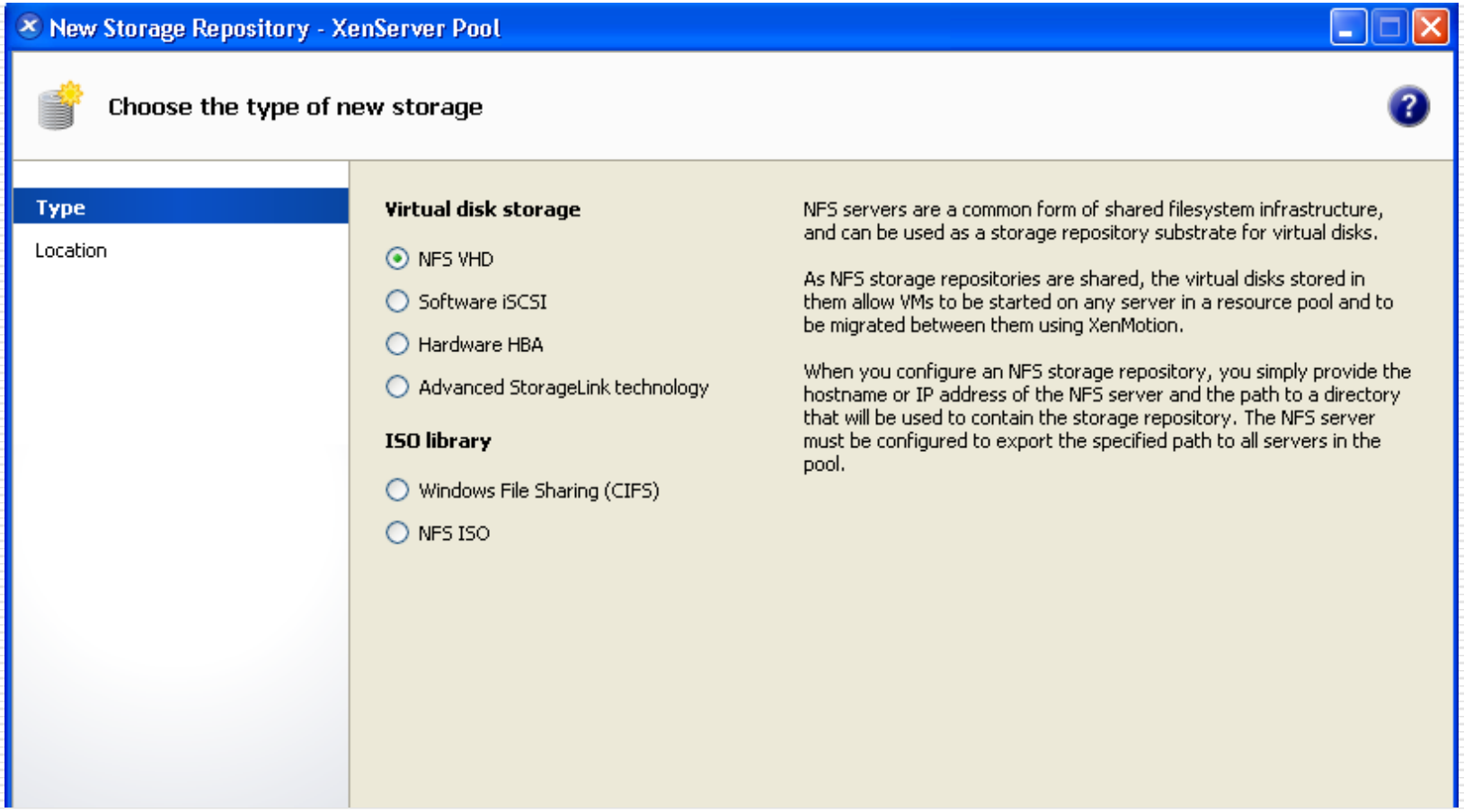
Network Virtualization

- SR-IOV
  - PCI-SIG standard, which allows virtual functions (VF), such as queue pairs, to be created that share the resources of a physical function (PF)
  - Primarily useful in allowing a single PCI device to be shared amongst multiple VMs
- Virtual switch
  - Virtual network switch connecting and exchanging traffic amongst VMs in a physical host
  - Cisco Nexus 1000V, VMware vswitch, Open vSwitch

Storage Virtualization

- Storage is for persistent (non-volatile) data
- Difficult to move or migrate if bundled with compute node
- Network-based storage is more suitable for cloud computing
  - Network-attached storage (NAS)
  - Storage Area Network (SAN)

Storage Types



Storage V14n – Current State

- Rethinking volume and LUN-based rigid file systems
  - Once configured, it is hard to increase or decrease the "volume"
  - Consider your experience in resizing a Linux or Windows partition
- Scalable file systems
  - ZFS pools
  - Other "bytes" based file systems

V14n and Cloud Computing

Cloud Computing (C^2)

- The term started when Google and IBM launched a university initiative to address Internet-scale computing back in 2007
- Services have been evolving since the 90s
- Previous incarnations include ***Grid*** and ***Utility computing*** and the ***Software as a Service*** a decade ago
- ***On-demand***, ***self-service***, ***pay-as-you-go*** utility
- Evolved from a combination of grid computing, ***virtualization***, and automation

Money Talks

- According to Cloud Computing Journal article:
  - "Experts estimate that this industry will grow to a ***42 billion*** dollar industry by 2012"
- What business owners and CIOs "think":
  - Shifting from ***Capex*** (capital expenditure) to ***Opex*** (operational expenditure)

Definition of C^2

- NIST defines cloud computing as
  - ***"a pay-per-use model for enabling available, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, services) that can be rapidly provisioned and released with minimal management effort or service provider interaction"***

C^2 Models

- Infrastructure as a Service (IaaS)
- Platform as a Service (PaaS)
- Software as a Service (SaaS)
- Hybrid variations

Infrastructure as a Service (IaaS)

- Provisioning of servers, storage, networking, OS, middleware, etc
  - E.g., EC2 (Amazon), Dynamic Infrastructures (Fujitsu), services from ISPs
- Subcategories
  - IaaS for workplace
  - IaaS for server
  - IaaS for storage
- Subscribers have control over selected portions of network
  - E.g. firewalls, OS, applications and storage

Platform as a Service (PaaS)

- Derivative of 70's Framework as a Service
- Provide services to support an application development life-cycle
  - e.g. AppEngine (Google) and Azure (Microsoft)
- Subscriber will use tools supplied by provider with no responsibility for cloud infrastructure (network, severs, OS and storage etc.)
- Locked into vendor's languages (Python, Java, Ruby, etc), tools and API

Software as a Service (SaaS)

- Applications running on a cloud infrastructure that can be accessed via a web browser interface.
- Started out as a service to companies who don't want to own SW (and the HW that runs it) in house
  - Salesforce.com as most famous example
- When ecosystem is formed, may expose internal platform and becomes PaaS
  - Force.com, Google

Software as a Service (Cont'd)

- Subcategories by application areas
  - CRM and business applications
  - Collaboration as a service
    - WebEx, GotoMeetings
  - Testing as a Service
  - Monitoring and management as a service
  - Development tools as a service
    - Force.com, Intuit
  - Security as a service
    - Qualys
  - Compliance and goverance as a service

Hybrid

- IT as a Service
  - Customized, packaged solutions provisioned to subscribers
  - Cut cross multiple layers, from infrastructure to applications
- XYZ as a Service
  - Expense Report as a Service ☺
  - You name it

Cloud in a Box

http://www.oracle.com/us/products/middleware/exalogic/index.html

- 30 servers with 360 cores
- Networking and storage
- Virtual machine technologies
- Solaris and Linux guests
- JRockit and HotSpot JVMs
- WebLogic Java EE
- Coherence caching

C^2 Delivery Methods

- ***Private Cloud***
  - A.k.a. Corporate Cloud
  - Behind a corporate firewall
  - Single tenant
- ***Public Cloud***
  - Provided for lease external to an entity's physical location
  - Multiple tenants
- ***Hybrid Cloud***
  - Mixture of public and private
  - E.g., non-confidential data externally whilst keeping private data in-house
- ***Managed Cloud***
  - Cloud physically located in house
  - Managed by provider

C^2 Issues and Concerns

- Software licenses
- Single point of failure
- Portability
- Security
- Auditing
- Compliance
- Other data access
- ***They mean research opportunities***

C^2 Security

- Intrusion/hacking
- Data piracy
- Privacy
- Multi-tenant issue
- NSA Trusted Computing Initiative

C^2 and Virtualization

- Many aspects of C^2 rely on server virtualization to realize
  - On-demand rapid provisioning
  - Elasticity (bidirectional scalability)
  - Portability and mobility
  - Security and Isolation
- C^2 has virtualization as the core, along with vendor differentiators:
  - Automation
  - Management

Virtualization Technologies

And the players who own them

V14n Types

- **According to where it happens**
  - **Desktop**
    - **VMWare**
  - **Server**
    - **ESX/ESXi/vSphere**
    - **Xen**
    - **LDom**
  - **Network**
  - **Storage**
- **According to technology**
  - **OS Virtualization (a.k.a., containers)**
    - **Solaris/OpenSolaris**
    - **SWSoft**
  - **Hardware Emulation (a.k.a., binary translation)**
    - **VMWare, QEMU/KVM**
  - **Para-virtualization**
    - **Xen, ESXi, Microsoft**

V14n Benefits

- Improve CPU utilization
- Improve security and isolation
- Save power, cooling, and space
- Facilitate migration and performance scaling
- Support software patching and upgrading
- Support legacy application and multiple OS versions

Drivers of Server Virtualization

- V14n 1.0
  - CPU Utilization
  - Workload Consolidation
  - Energy efficiency, building cooling and footprint
- V14n 2.0
  - Reliability
  - Security
  - Elasticity

Types of Server Virtualization

- Full virtualization
  - No modification to guest OS
  - Privileged instructions intercepted and translated by Hypervisor
- Para-virtualization
  - Modifications to guest OS
  - Reduced translations
  - Reuse of guest OS drivers

CPU Virtualization

- Multi-core technologies
- X86 hyper-threading
- Sparc chip multi-threading
- SMP architecture
- NUMA
- P = N*S*C*T

  P = No. of virtual processors (CPUs)

  N = No. of nodes

  S = No. of sockets

  C = No. of cores

  T = No. of threads

CPU Resource Allocation

- Virtual processors form a processor pool
- Knapsack problem
  - Partition P virtual processors into K workloads (or M virtual machines)
  - Subject to memory, network and storage constraints
- Power efficiency
  - Only activate a subset of P
- High Availability
  - Multiple VMs for a workload (thus M > K)

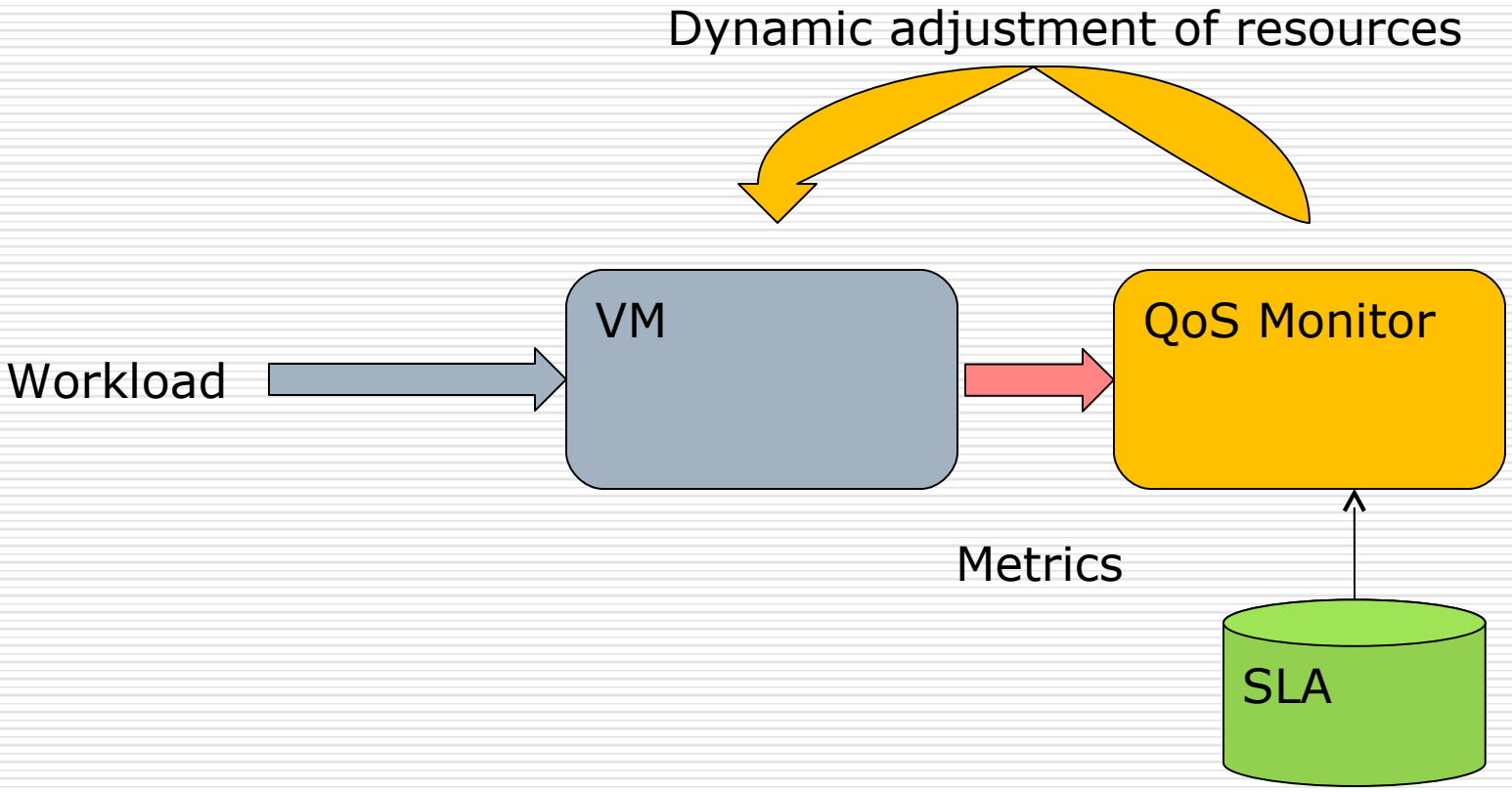
Elasticity

- Horizontal Scaling
  - Scale up and down the virtual processors allocated for a workload
- Mobility
  - Live migration from one set of virtual processors to another set
  - Main challenge in how to copy memory content while the memory pages are being dirtied
  - Xen VM migration, VMware vMotion

Quality of Service (QoS)

- No real-world application runs at 100% CPU utilization
  - Delay becomes unbounded
- Challenge is to determine the "optimal" ***operational point***
  - Synthetic workload (benchmarks) used for sizing does not capture the real-time dynamics
- QoS (e.g., transaction delay) of a server is monitored in real time
  - If QoS satisfied, then scale down resources
  - Else, scale up

QoS Monitoring



Welcome Further Discussions

- Email: yaominchen@gmail.com
- Skype: yaominchen
- LinkedIn: Yao-Min Chen