Distributed Systems

31. The Cloud: Infrastructure as a Service

Paul Krzyzanowski
Rutgers University
Fall 2013
Motivation for the Cloud

• Self-service configuration
  – Provision the servers, storage, and network connectivity they need

• Hosting
  – Don’t worry about hosting and maintenance

• Elastic growth & metered service
  – Users can add more to their environment as needed
  – Pay for what you use

• CAPEX → OPEX
  – No capital purchases, depreciation
XaaS

Software as a Service (SaaS)
Ability to use software in the cloud

Platform as a Service (PaaS)
Ability to deploy apps in the cloud

Infrastructure as a Service (IaaS)
Create a computing environment in the cloud

Users
Web-based email, Google apps, Microsoft Office Online, Apple iWork for iCloud

Programmers
Java/Eclipse programming platform as a network service; Google App Engine: Python, Java, PGP, Go

IT
Create a web service with several load-balanced web servers, app servers, and a back-end database
Public vs. Private

• Private cloud
  – Dedicated to one organization; hosted on premises

• Public cloud
  – Hosted Enterprise Cloud
    • Set of infrastructure dedicated to one organization
    • Hosted and operated by a third party (e.g., Rackspace)
  – Multi-tenant Public Cloud
    • Shared and dedicated resources
    • Elastic scaling
    • Services accessible by multiple customers over the Internet

• Hybrid Cloud
  – Services provided with a combination of private and public cloud
Vendors

• Many vendors: frameworks & providers
• Popular open efforts: CloudStack, Openstack, & Eucalyptus
  – Share some of the same components
  – Enable the use of a variety of VMs, network controllers, and storage systems

• Also:
  – Google Compute Engine (GCE)
  – Amazon EC2 (Elastic Compute Cloud)
    • + Amazon S3 (storage) + Amazon RDS (relational DB) + …
  – Microsoft Azure
<table>
<thead>
<tr>
<th>CloudStack</th>
<th>OpenStack</th>
<th>Eucalyptus</th>
</tr>
</thead>
<tbody>
<tr>
<td>[initially Citrix]</td>
<td>[initially Rackspace]</td>
<td>Mostly written in Java</td>
</tr>
<tr>
<td>Mostly written in Java</td>
<td>Mostly written in Python</td>
<td>Focus on private cloud</td>
</tr>
<tr>
<td>&gt;100 production clouds</td>
<td>Not widely used commercially</td>
<td>KVM &amp; Xen hypervisor support</td>
</tr>
<tr>
<td>Up to 30K physical hosts</td>
<td>Supports lots of hypervisors</td>
<td>Amazon EC2 and S3 compatible</td>
</tr>
<tr>
<td>Supports lots of hypervisors</td>
<td>Network is a point of failure</td>
<td></td>
</tr>
<tr>
<td>Multiple datacenters</td>
<td>Partial support for Amazon EC2</td>
<td></td>
</tr>
<tr>
<td>Amazon EC2 compatible</td>
<td>Difficult to install</td>
<td></td>
</tr>
</tbody>
</table>
The following examples will focus on CloudStack
Cloud infrastructure hierarchy

Managed collection of hosts, networks, storage: may span data centers

Collection of hosts, networks, storage: one or more zones per data center

A rack

Group of hosts & associated primary storage

Server Farm

Zone

Pod

Secondary storage

Cluster

Hosts

Primary storage

VMs

Templates, snapshots, & ISO images

VM storage
Core Components

Dashboard
Management
Identity
Compute
Network
Storage
Compute
Virtual Machines

• Each physical machines runs a hypervisor
  – Allows running of multiple operating systems
  – Provides virtual interfaces to hardware (storage, NIC, MMU, …)
  – Management server interacts with hypervisor

• Setup: user actions
  – Select OS
  – Number of CPUs and RAM
  – Disk size, number of volumes
  – Network

• Provisioning
  – Primary storage volume(s) allocated on a cluster
  – VM configured from template & configuration data
  – Instance of VM copied from secondary storage to primary storage
The Network

• Administrator actions
  – Define networks (local addresses)
  – Attach VMs to networks
  – Acquire public address(es) for NAT
  – Configure firewall rules for access and load balancing
The Network: keeping things separated

- Physical hosts share a physical network
- Virtual network (VLAN)
  - Tag on ethernet frame identifies virtual network
  - Traffic from multiple hosts multiplexed on the same physical network
  - Management traffic is untagged
- Hypervisor involvement
  - Traffic labels defined by management interface and configured on hypervisor
  - Hypervisor provides virtual LAN interface
  - Filters traffic to VM
The Network: connecting networks

• Virtual router
  – DHCP: assign internal addresses
  – DNS: domain name lookup
  – Load balancing: define a pool of internal servers
  – VPN: secure channel between internal & remote networks/machines

• Virtual routers manage connectivity between
  – Internal subnets
  – Internal & external networks

• May deploy a combination of virtual and physical network hardware
  – E.g.,
    dedicated firewalls and load balancers,
    virtual router for DHCP & DNS
Storage

• Primary Storage
  – The VM’s “disks”
  – Presented as a block device (looks like a raw disk)
    • The VM writes a file system on it
  – Accessible from on VM host or a cluster within a pod
  – Uses local disk, iSCSI, Fibre Channel, or NFS

• Secondary Storage
  – Zone-level configuration
  – Stores templates, snapshots, ISO images
  – Uses NFS or OpenStack Swift object store

• Other (beyond IaaS)
  – Access other storage servers
    • Distributed file systems, Bigtable, Amazon S3, Databases
Snapshots

• Backup images of a VM
  – Allows restoring state of a machine to the latest snapshot
  – Administrator configures frequency of snapshots
  – Hypervisor writes them

• Initially written to primary storage (fast)

• Then moved to secondary storage

• Backup network traffic can be segregated to lower impact
High Availability
High Availability

- Poll VMs for liveness (heartbeats)
- Fencing
  - Disallow dead/unknown VM from accessing storage
- Restart VM
- Use snapshots for restoring VM state
- Other mechanisms for HA (beyond IaaS)
  - Active/active, active/passive redundancy
  - Application-level heartbeats
  - Paxos/Virtual synchrony
The End