Unshackle the Cloud: Commoditization of the Cloud

Hakim Weatherspoon
Assistant Professor, Dept of Computer Science

CS 5412, Guest Lecture, Cornell University
January 24, 2012
• The promise of the Cloud
  – A computer utility; a commodity
  – Catalyst for technology economy
  – Revolutionizing for health care, financial systems, scientific research, and society
The promise of the Cloud

- ubiquitous, 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.

NIST Cloud Definition
• The promise of the Cloud
  - ubiquitous, 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.
• The promise of the Cloud
  – ubiquitous, 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.

• However, cloud platforms entail significant risk
  – Vendor Lock-in
  – Storage Lock-in
  – Computation Lock-in
Challenge

• How to use the cloud?
  – Storage
  – Computation
  – Network

• Without being locked into a single cloud provider?
Outline

• Breaking Cloud Storage Lock-in
• Breaking Cloud Computation Lock-in
  – (Nested) Virtualization
Vendor Lock-in: Cloud Storage

- Large organizations considering using the cloud
  - New York Times
  - Netflix
  - Nintendo
  - Cornell
  - Library of Congress

- The more data you have, the harder it is to move
  - Switching providers entails paying for bandwidth *twice*
  - Inhibits opportunistic migration
Vendor Lock-in: Cloud Storage

- How hard is it to move a PetaByte?

Titan tech boom, randy katz, 2008
Vendor Lock-in: Cloud Storage

• All my valuable data/computation is in the cloud
  Am I locked in to one provider forever?
  – The more data you have, the harder it is to move

• RACS: Redundant Array of Cloud Storage
  – Collaboration with the Internet Archive and IBM
Vendor Lock-in: Cloud Storage

• All my valuable data/computation is in the cloud. Am I locked in to one provider forever?
  – The more data you have, the harder it is to move

• RACS: Redundant Array of Cloud Storage
  – Collaboration with the Internet Archive and IBM

- RACS(3,4)
  - Object 100 KB
  - Relative Storage $n/k$
  - Relative Upload Bandwidth $n/k$
  - Relative Download Bandwidth 1
Vendor Lock-in: Cloud Storage

Estimated Cost of Switching Cloud Providers

<table>
<thead>
<tr>
<th>Provider</th>
<th>Cost in $K</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Provider</td>
<td>25</td>
</tr>
<tr>
<td>RACS (4,5)</td>
<td>15</td>
</tr>
<tr>
<td>RACS (6,7)</td>
<td>10</td>
</tr>
<tr>
<td>RACS (8,9)</td>
<td>5</td>
</tr>
</tbody>
</table>
Vendor Lock-in: Cloud Storage

• Graduate Students
  – Hussam Abu-Libdeh
  – Lonnie Princehouse
  – Ji Yong Shin

• Collaborators
  – Sandra Payette (Fedora Commons)

• Website:
  – http://racs.cs.cornell.edu
Vendor Lock-in: Cloud Computation

• Cloud storage is only a half third of the story
  – What about computation?

• How can I move my computation between clouds?
Vendor Lock-in: Cloud Computation

• Move computation via Virtualization
  – Virtualize processor Instruction Set Architecture
  – Full Virtualization vs Paravirtualization (of hardware)
  – VMWare vs (Original) Xen

• Xen
  – Separation of policy and mechanism
  – DomU hosts guest operating system in virtual machine
  – Dom0 manages devices and guests
  – Control Transfer: Hypercalls and Events
    (like syscalls and device interrupts)
Vendor Lock-in: Cloud Computation

- Xen
- Linux Dom0
  - Backend
  - PCI-Driver
- Shared Memory Device: Ring
- Linux DomU: VM
  - Frontend
- Baremetal
- Cloud Computation

- Vendor Lock-in
Can I compute in the cloud if some of my data is in a vault at home or on another provider?
Vendor Lock-in: Cloud Computation

- Popular IaaS clouds are becoming **feature-rich**
  - Integrated monitoring
  - VM migration
  - CPU bursting

- **Hypervisor-level** innovations are emerging
  - Availability (e.g. Remus [Cully et al., NSDI 2008])
  - Security (e.g. Revirt [Dunlap et al., OSDI 2002])
  - Efficiency (e.g. Overdriver [Williams et al., VEE 2011])
Vendor Lock-in: Cloud Computation

- Cloud users can be large enterprises with 100's or 1000's of VMs
- Provider must expose hypervisor-level features
- Tools and features lead to lock-in

- **Users can't implement hypervisor-level features themselves**
Unshackle the Cloud: xClouds

- Bring **extensibility** into IaaS clouds
- Allow users to run or implement **their own** hypervisor-level services
- Avoid lock-in with **user-centric homogenization**
How to Build xClouds

- Users are isolated
- VMM composed of **modules**
How to Build xClouds

- Users are isolated
- VMM composed of modules
  - User / Provider (U / P)
How to Build xClouds

- Users are isolated
- VMM composed of modules
  - User / Provider
    - (U / P)
  - Mutable / Immutable
    - (U / P)
How to Build xClouds: Alternatives

Download VMM Extensions
- e.g. SPIN, VINO

Providers must adopt new VMM

![Diagram showing VMs and users]

Hardware
## How to Build xClouds: Alternatives

<table>
<thead>
<tr>
<th>Download VMM Extensions</th>
<th>Expose Hardware Through VMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>e.g. SPIN, VINO</td>
<td>e.g. Exokernel</td>
</tr>
</tbody>
</table>

| Providers must adopt new VMM | Providers must adopt new VMM |

### Diagram:

```
Hardware

User 2

User 3

<table>
<thead>
<tr>
<th>VM</th>
<th>VM</th>
</tr>
</thead>
<tbody>
<tr>
<td>U</td>
<td>P</td>
</tr>
<tr>
<td>P</td>
<td>U</td>
</tr>
<tr>
<td>P</td>
<td>U</td>
</tr>
<tr>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>P</td>
<td>P</td>
</tr>
</tbody>
</table>

LibVMM

User 2

User 3

<table>
<thead>
<tr>
<th>VM</th>
<th>LibVMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>U</td>
</tr>
<tr>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>P</td>
<td>P</td>
</tr>
<tr>
<td>P</td>
<td>P</td>
</tr>
</tbody>
</table>
```
# How to Build xClouds: Alternatives

<table>
<thead>
<tr>
<th>Download VMM Extensions</th>
<th>Expose Hardware Through VMM</th>
<th>Add Another VMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>e.g. SPIN, VINO</td>
<td>e.g. Exokernel</td>
<td>e.g. Turtles Project</td>
</tr>
<tr>
<td>Providers must adopt new VMM</td>
<td>Providers must adopt new VMM</td>
<td>Turtles needs VMM support, but...</td>
</tr>
</tbody>
</table>

![Diagram showing VMs and users interacting with hardware.](image-url)
How to Build xClouds: Another Layer

- Xen
- Linux Dom0
  - Backend
  - PCI-Driver
- Shared Memory Device: Ring
- Linux DomU: VM
  - Frontend

- Baremetal
- Xen

Diagram showing the integration of Linux Dom0 and DomU with Shared Memory Device: Ring and PCI-Driver.
How to Build xClouds: Another Layer

- **Linux Dom0**
  - Backend
  - PCI-Driver

- **Dom0 for Xenblanket**
  - Frontend
  - Backend

- **PV-DomU**
  - Frontend

- **PV-DomU**
  - Frontend

- **Xen - 4.1.1 Blanket**

- **1st-Layer Xen**

- **Device**

- **Baremetal**

- **Hypercalls**
Hypercall Passthrough

• Need Hypercall Passthrough
  – Nested Dom0 must be able to get information about shared memory devices from 1st Layer-Xen

  – Nested Dom0 can only issue hypercall to Nested Xen
    • So, nested Xen should help passthrough related hypercalls
Will xClouds Perform?

- Compared single and nested setups with Xen (PV) as the second-layer hypervisor
- Microbenchmarks
  - Nested perf. comparable to single-layer PV
- Device I/O benchmarks
  - Xen is not designed to run on PV hardware
  - Nested PV is essential for device I/O
Configuration for Comparison

- Compared single and nested virtualization
- Xen (PV) top layer hypervisor
- Xen (HVM) and KVM bottom layer hypervisor
<table>
<thead>
<tr>
<th></th>
<th>Baseline</th>
<th>PV</th>
<th>KVM</th>
<th>PV/KVM</th>
<th>HVM</th>
<th>PV/HVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double div (ns)</td>
<td>7.19</td>
<td>7.55</td>
<td>7.41</td>
<td>7.35</td>
<td>7.61</td>
<td>7.57</td>
</tr>
<tr>
<td>Null call (µs)</td>
<td>.19</td>
<td>.37</td>
<td>.20</td>
<td>.38</td>
<td>.21</td>
<td>.37</td>
</tr>
<tr>
<td>Fork proc (µs)</td>
<td>65</td>
<td>250</td>
<td>87</td>
<td>337</td>
<td>79</td>
<td>280</td>
</tr>
</tbody>
</table>
Disk Write Throughput

![Chart showing Disk Write Throughput for different configurations: Baseline, PV, KVM, PV/KVM, HVM, and PV/HVM. The y-axis represents Disk Write Throughput (MB/s), and the x-axis represents different configurations. The chart indicates the performance comparison among these configurations.]
Nested PV is essential
xClouds works Today!

- Nested paravirtual device drivers
- Xen on EC2
xClouds works Today!

- Nested paravirtual device drivers
- Xen on EC2

- Can create your own
  Cloud-within-a-Cloud
xClouds works Today!

• Graduate Students
  – Dan Williams
  – Zhefu Jiang
  – Ji Yong Shin

• External Collaborators
  – Hani Jamjoom (IBM)
“With great power comes great responsibility”

– Cloud technology can be used to address economic concerns

Treating the cloud as a commodity

– Users need to be able to trade-off overhead and vendor mobility
  – Providers need to be accountable to users and environment

Lots more research to do to achieve the promise of the Cloud
Paper Trail Theme: Cloud & Vendor Lock-in
  – xCloud/Xen-Blanket in EuroSys-2012
  – xCloud in HotCloud-2011
  – Overdriver in VEE-2011
  – RACS in SOCC-2010

More at http://fireless.cs.cornell.edu
and also http://xcloud.cs.cornell.edu

Email: hweather@cs.cornell.edu
Backup