Why Virtualize My Automation Systems?

Henrik Mønster
VMware Danmark
Agenda

- Introduction
- The Challenges
  - And how we handle them…
- The Benefits
- The Possibilities
Introduction
More VMs are in motion than planes in flight

1 VM EVERY SIX SECONDS

That's faster than the rate of babies born in the U.S.

At any given time, more VMs are in motion than planes in flight.

20 MILLION VMs

If they were physical machines they would stretch 2x the length of Great Wall of China

>68,000

VMware-CERTIFIED PROFESSIONALS IN 146 COUNTRIES
\textbf{>95\% of apps can not make use of current server resources}

<table>
<thead>
<tr>
<th>Application Performance Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Cores</td>
</tr>
<tr>
<td>4 Cores</td>
</tr>
<tr>
<td>8 Cores</td>
</tr>
<tr>
<td>16 Cores</td>
</tr>
</tbody>
</table>

Server performance

The gap is widening

\( \€ \)
Operational Benefits of Virtualization

Based on Averages from Customer Operational Readiness Assessments, 2011

<table>
<thead>
<tr>
<th>Capital cost per workload</th>
<th>Admins per 100 workloads</th>
<th>Time to provision new workload</th>
<th>Datacenter outage cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before: $14,235</td>
<td>Before: 3 admins</td>
<td>Before: 3 days</td>
<td>Before: $30 MM</td>
</tr>
<tr>
<td>After: $5,694</td>
<td>After: 1 admin</td>
<td>After: 3 min</td>
<td>After: $4 MM</td>
</tr>
</tbody>
</table>

Based on Averages from Customer Operational Readiness Assessments, 2011
Better Platform for Business Critical Apps

- **Improve App Quality of Service**
  - Better performance with dynamic resources and scalability
  - Enhanced availability and automated DR for all apps

- **Improve App Efficiency**
  - Lower hardware and software costs with 5X-10X consolidation
  - Reduced Opex with intelligent policy management

- **Accelerate App Time-to-Market**
  - Provisioning times reduced from weeks to minutes
  - Optimized test/dev environments

---

- **SLAs**
- **Cost Reduction**
- **Agility**
The Trend Is Clear

% of Workload Instances Running on VMware in Customer Base

Source: VMware customer survey, Jan 2010 and April 2011 interim results,
Data: Total number of instances of that workload deployed in your organization and the percentage of those instances that are virtualized
The Challenges
Can we handle the resource consumption of critical apps?

Application’s Performance Requirements

<table>
<thead>
<tr>
<th>% of Applications</th>
<th>ESX 1</th>
<th>ESX 2</th>
<th>Virtual Inf. 3.0/3.5</th>
<th>VMware vSphere 4</th>
<th>VMware vSphere 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU</td>
<td>1 vCPU</td>
<td>2 vCPUs</td>
<td>4 vCPUs</td>
<td>8 vCPUs</td>
<td>32 vCPUs</td>
</tr>
<tr>
<td>Memory</td>
<td>2 GB per VM</td>
<td>3.6 GB per VM</td>
<td>16/64 GB per VM</td>
<td>255 GB per VM</td>
<td>1,000 GB per VM</td>
</tr>
<tr>
<td>Network</td>
<td>&lt; .5 Gb/s</td>
<td>.9 Gb/s</td>
<td>9 Gb/s</td>
<td>30 Gb/s</td>
<td>&gt; 36 Gb/s</td>
</tr>
<tr>
<td>IOPS</td>
<td>&lt; 5,000</td>
<td>7,000</td>
<td>100,000</td>
<td>365,000</td>
<td>1,000,000</td>
</tr>
</tbody>
</table>

% 95% of Apps Require

Yes we can!
History of Virtualizing Real time Applications

- **Customer demand is there.**
  - Common platform for “virtualization first” shops.
  - Why not virtualize real time applications also?

- **VMware cautious of risks in ESX 3.5 and before.**
  - Scheduler not designed for real time applications.

- **Goals of real time and virtualization very different.**
  - Real time:
    - Dedicate resources to ensure SLA.
  - Virtualization:
    - Spread resources across multiple VMs.
    - Over commitment of resources
What should we worry about when virtualizing real time apps?

- CPU scheduling latency
- Memory management and locality
- Network latency
- Storage performance
- “Noisy” Neighbor VM – resource isolation
And how we handle them…
CPU Recommendations

- Understand impact of larger SMP VMs
  - Size VMs Appropriately – Try to stay inside a NUMA node if possible
  - But 1 or 2 vCPU machines are totally fine
  - Host CPU Saturation occurs around 80%

- Don’t Set CPU Limits (unless you Really need to)

- Over provisioning is fine, Consolidation ratios very depending on application!

- HyperThreading is good, don’t disable it.

- %USED is more important than %Run

- Use the latest hardware, older hardware has less built-in virtualization support.

- ESXi Scheduler has evolved over the years
  - relaxed CO scheduler, Better HT usage, NUMA support, cores/socket
Advanced Tuning: CPU Affinity – Use as last resort

- **CPU Affinity (aka Pinning)**
  - Rumored to be critical.
  - Our data shows little gain with vSphere 4.x and before.

- **Affinity and vSphere 5.0**
  - Allows “Exclusive Affinity”.
  - Previously, cores still accessible to other VMs despite affinity.

![Graph showing Max DSP Execution Time in Milliseconds with SLA comparison between With and Without Exclusive Affinity.](image_url)
Initial Observations and Results: Intel 5450 Series, 4 vCPU

- **Initially**: Worst case latency – well over target
  - Memory management appeared to be the bottleneck.
The Effect of Extended Page Tables

- **Intel Nehalem and beyond**
  - Generally fast chips (without MMU assist)
  - EPT is critical for realtime performance
Memory - vNUMA

- Feature is ON by default for Large VMs
- Exposing the virtual NUMA topology can lead to significant performance gains for Large Memory intensive Applications
  - Close to native performance with HPC like applications/Benchmarks SPEC MPI V2.0 and SPEC OMP V3.2 when virtualized on vSphere 5.0 using vNUMA support
  - Virtual NUMA topology is chosen to closely match physical host topology, while preserving the virtualization benefits of portability and load balancing.
Reservations and Shares

- **Preallocate resources to VM.**
  - ESX by design is conservative in resource allocation.

- **Recommend only where necessary.**
  - Real time applications are good use case:
    - Both CPU and memory.
    - Quicker access to guaranteed resources.

- **Overuse impacts cluster usability.**
  - Customers may ignore recommendations.
  - Tradeoff between your application and others.

- **Largely idle systems.**
  - Some performance gains, even with little utilization.
Network performance

**Default configuration**

![Default configuration graph]

**SplitRxMode**

![SplitRxMode graph]
Advanced Tuning: Direct Path I/O

- **Direct Path I/O (aka NIC Passthrough)**
  - Disables vMotion.
  - Makes physical NIC available for only one VM.

- **Latency improvements**
  - But at substantial cost in vSphere functionality.

- **Cisco UCS gear provides alternative**
  - SRIOVC provides reusable NIC with vMotion and Passthrough.

---

![Graph showing worst case latency in milliseconds for SLA and Direct Path I/O](image-url)

- **SLA**
- **Without Direct Path I/O**
- **Direct Path I/O**
Storage Best Practices

- Size Storage Arrays appropriately for Total VM usage
- > 30ms Disk Latency could be a performance problem
- > 2ms Kernel Latency could be a performance problem or a undersized ESX device queue
- Use Storage vMotion to balance workloads across datastores, SDRS in vSphere 5.0
- Utilize SIOC for greater storage performance protection for important workloads
- Update Storage Firmware (if available) to take advantage of VAAI
- Follow Storage Vendor best practices for Block size and Alignment, Also investigate the use of Disk Defragmentation *cautiously* in the guest as needed
Protection from “Noisy” Neighbor with IO Controls

Business Critical

Storage I/O control

Network I/O control

Ensure critical apps get preferential access
Dynamic Load Balancing of Resources

**DRS**: Load balances VMs across Hosts

**Storage DRS**: Load Balances VMs across Storage Arrays

**Network & Storage IO controls**: Ensures critical workloads get preferential access to Network and Storage IO.
Tips for Latency-sensitive Applications

- **vSphere**
  - Use ESX 4.x or later
  - Prior versions were not engineered to support real-time workloads

- **Processor**
  - Must support HW assist (Intel VT, AMD V)
  - Use fewer vCPUs to reduce scheduling overhead
  - Monitor host utilization (<80%)

- **Memory**
  - Must support Hwmmu (Intel EPT, AMD RVI) – Critical!
  - Stay within NUMA boundaries
Tips for Latency-sensitive Applications

- **Power management**
  - Turn off power management in BIOS (set to **static high**)
  - Intel: Turn off C-States, but enable TurboBoost

- **Choice of Guest OS**
  - Use a modern OS, e.g. Windows 2008 or later, RHEL 6, SLES11SP1; these generally have lower virtualization overheads, e.g. tickless architecture

- **Physical NIC settings**
  - Use `ethtool` to disable interrupt coalescing on ESXi host

- **Virtual NIC settings**
  - Use VMXNET3 and disable interrupt coalescing

- **Virtual Machine settings**
  - Avoid over-committing resources; consider `monitor_control.halt_desched = “false”`
The Benefits
Improving Business Continuity At All Levels

Local Site
- vSphere High Availability
- vSphere Fault Tolerance
- vMotion and Storage vMotion

Data Protection
- vSphere Data Recovery
- Storage APIs for Data Protection

Failover Site
- vCenter Site Recovery Manager
- Includes vSphere Replication
Application-Aware High Availability

- Protect against h/w & n/w & storage connectivity failures
- Protect against app failures

Neverfail’s vAppHA and Symantec’s ApplicationHA currently use VMware App-aware API to create products to interfacing with vSphere HA to restart the VM
Recovery from Disasters

Site A (Primary)

VMware vCenter Server

Site Recovery Manager

VMware vSphere

Servers

Site B (Recovery)

VMware vCenter Server

Site Recovery Manager

VMware vSphere

Servers
Rapid provisioning or cloning of single or multiple VMs

**Full Clone (Default)**

- Entire disk copied
- New independent VM

**Fast Provisioning**

- No disk copy
- Delta disks spawned

VM A → VM B

Delta Disk
Base Disk (READ ONLY)
vShield: Better Security for Business Critical Apps

- Sensitive Data Detection provides visibility to meet standards & regulations
- vShield Endpoint protects the individual VM with offloaded anti-virus
- vShield App protects the applications with multi-VM trust zones
- vShield Edge protects the perimeter of your virtual data centers

More cost effective than physical firewalls

Simpler to manage than existing solutions – dynamic security that follows your application

Cloud Infrastructure (vSphere, vCenter, vShield, vCloud Director)
The Possibilities
vCenter Operations Management Suite

VMware’s Solution for Automated Operations Management

What does it deliver?

• Integrated and highly automated performance, capacity and configuration management
• Operations management console that integrates with 3rd party solutions
• Continuous compliance analysis and remediation

Why is it unique?

• Patented analytics self learn “normal” in a dynamic environment
• Integrated approach converges disciplines
• Designed for vSphere and built for Cloud but extensible to legacy solutions
Health Degradation – Smart Alerts (All Editions)

Overview

- Proactive alerts that provide early warning on building issues
- Identify upcoming health, performance and capacity issues
- Automatic root cause analysis of offending metrics across all layers

Benefits

- Advance notification of abnormal behavior help avoid incidents
- Immediately focus on the root cause rather than symptoms or false alerts
Everybody talks Cloud!

You will eat your cloud and you will like it!

Credits: VMworld 2011 presentation CIM2561 by Drue Reeves, Gartner.
VMware Offers Hybrid Cloud

Hybrid cloud offers optimal agility & flexibility:
- Control & SLAs of a private cloud
- Economics and speed of public cloud
Connect to external cloud resources, migrate workloads

- Connect Public and Private clouds
- Connect existing infrastructure to a Public cloud
Connect to a Public Cloud

- **Configure vCloud Connector**

  - vSphere Client with vCloud Connector Plug-in
  - vCloud Connector Virtual Appliance
  - vCloud Director
  - vApp (OVF)
  - vCenter Server
  - vSphere
  - Private vCloud
  - Public vCloud

Enterprise Datacenter
The vCloud Ecosystem on vcloud.vmware.com

Find a vCloud Services Provider

Choose a vCloud Service Provider and service type to match your public cloud needs from the world's largest enterprise service provider ecosystem. Connect with select partners providing "test drives" to first evaluate a public cloud. With vCloud services, manage and scale your private and public cloud initiatives using common technology, easily moving workloads, templates and virtual machines across your hybrid cloud.

Datacenters Located in EMEA

- All for One
- Micimarket AG
- Bitvimos IT Services
- Business Connexion
- Curranza
- Claranet
- Colt
- Coresio
- CSC

See List View
Questions?
External Resources

- VMware’s Performance Technology Pages
  - http://vmware.com/technical-resources/performance

- VMware’s Performance Blog – VROOM!

- Performance Community Forum

- RealTime Applications Running on vSphere
  - http://www.vmworld.com/docs/DOC-4727

- VMware Education
  - vSphere: Manage for Performance Class
External Resources

Case Study
http://www.vmware.com/a/customers/customer/819/Pinellas+County+Utilities

Case Study

Network I/O Latency in vSphere 5
http://www.vmware.com/resources/techresources/10256

Performance Tuning of Latency-Sensitive Workloads
http://www.vmware.com/resources/techresources/10220
Thank You!