Virtual Environments Optimization

A simple and powerful recipe to reduce costs

Prepared for: St Louis Computer Measurement Group
October 2016
Renato Bonomini - renato.bonomini@moviri.com
Ettore Trevisiol – ettore.trevisiol@moviri.com
Agenda

• Background of the case Study: what are we trying to solve?
• Part 1 of the solution: decommissioning process
• Part 2 of the solution: rightsizing process
Study Case Background

• Modern datacenters increasingly, if not exclusively, rely on virtual technologies
• Such environments are highly dynamic and flexible in their nature, making IT resource provisioning quick and easy
• As a side effect though, virtual machines may be kept even after they are no longer needed by client

Virtual Estates size

+10 new VMs/day

Increase avg. VM size

CMG StLouis 2016 - Virtual Environments Optimization - © Moviri 2016
Some thoughts from previous CMG Meetings

Moore's law

Now achieved by increasing the number of cores

The number of transistors on integrated circuits doubles approximately every two years
Decommissioning – What’s wrong with this picture?

CPU Utilization % distribution

- Average
- Maximum of hourly samples

Number of VMs vs. CPU Utilization [%]
Decommissioning – Process

1. Collect Performance Data
   - Extraction from hypervisors (VMware, Hyper-V, AIX, KVM, Xen)

2. Retrieve VMs’ owners
   - Extraction from asset mgmt. and inventory platforms

3. Spot unused VMs
   - Define utilization thresholds and decommissioning rules
   - Evaluate rules on all virtual machines and list unused
   - Contact virtual machines’ owners and share thresholds

4. Shut down VMs
   - Obtain VM’s owners approval
   - Shut down virtual machines
   - Repeat periodically
Decommissioning – Performance Data Collection

Capacity Manager
Virtual/Cloud Admin
Storage Admin
Head of Infrastructure
Service Manager
Business Owner

Demand Mgmt
Discovery/CMDB

Capacity Management Information System

Performance, Capacity and Configuration Metrics
Business KPIs

Physical/Virtual/Clouds
Databases
Storage
Networks
Big Data
Facilities
Business Drivers

CMG StLouis 2016 - Virtual Environments Optimization - © Moviri 2016
Decommissioning – Criteria

Performance conditions that can be applied:
- N-th percentile of $\text{Normalized\_CPU\_Utilization} < \text{CPU\_THRESHOLD}$
- N-th percentile of $\text{Network\_bit\_rate} < \text{NETWORK\_THRESHOLD}$
- N-th percentile of $\text{Disk\_transfer\_rate} < \text{IO\_THRESHOLD}$
- Availability of data for CPU/NETWORK/DISK_IO > AVAILABILITY_THRESHOLD

Enforced conditions
- VM is a Server (not a VDI VM)
- VM is currently present in Virtual Infrastructure
- VM is powered on
Decommissioning – Rules

Quick-wins

CONDITION 1 (Server not using network)
  – Network traffic $X^{th}$ perc = THRESHOLD_A

CONDITION 2 (Server very scarcely using CPU)
  – Normalized CPU utilization $Y^{th}$ perc < THRESHOLD_B

Process-like conditions

CONDITION 3 (Idle servers, condition A)
  – Normalized CPU utilization $X^{th}$ perc < THRESHOLD_X AND Network traffic $Y^{th}$ perc <= THRESHOLD_Y AND Disk transfer Rate $Z^{th}$ <= THRESHOLD_C

CONDITION 4 (Idle servers, condition B)
  – Normalized CPU utilization $W^{th}$ perc < THRESHOLD_X AND Network traffic $K^{th}$ perc <= THRESHOLD_Y AND Disk transfer Rate $Z^{th}$ <= THRESHOLD_C

All conditions impose data availability > 80% over last X months
Decommissioning – Results Breakdown

- **All VMs**
  - 62% Server
  - 38% VDI

- **Utilized**
  - 90% Server
  - 90% VDI

- **Un-utilized**
  - 10% Server
  - 10% VDI

- **VDI**
  - 38%

- **Utilized Request sent**
  - 81%

- **Unutilized**
  - 10%

- **Utilized**
  - 90%

- **Request sent**
  - 81%

- **Responded**
  - 51%

- **No owner**
  - 19%

- **No answers**
  - 49%

- **OK to shutdown**
  - 51%

- **Reallocated**
  - 10%

- **Spare VMs**
  - 34%

- **Wrong owner**
  - 5%

- **Reallocated VMs**
  - Owners confirmed VMs are not longer used

- **Answer solicited each months, then shut down after 6 months if still under threshold**

- **Decommissioning process started after 6 months**

CMG StLouis 2016 - Virtual Environments Optimization - © Moviri 2016
Rightsizing – Situation after Decommissioning

We have identified machines that we can turned off. How do further improve?

Rightsizing machines reduces risk of over-commitment and decreases licensing costs or potentially hosting costs
Rightsizing – CPU 95th Percentile Distribution

- VMs with only 1 vCPU have been removed from the analysis: they are not down-sizeable.
- The most part of the selected VMs have a 95th percentile wide under the warning suggested thresholds.
- About 66% of the VMs are under the 10% of CPU utilization.
Rightsizing – CPU 99th Percentile Distribution

- Considering the 99th percentile the distribution moves **slightly to the right**
- However the most part of the selected VMs 99th percentile keeps wide under the warning suggested thresholds
- About 37% of the VMs are under the 10% of CPU utilization
- About the 70% is under the 20% of CPU utilization
Rightsizing – Process

1. Collect Performance Data
   Extraction from hypervisor (VMware, Hyper-V, AIX, KVM, Xen)

2. Spot oversized VMs
   - Define target utilization thresholds and other algorithm parameters
   - Evaluate utilization forecasts
   - Produce recommendations on all VMs and list oversized

3. Right size the VMs
   - Downsize virtual machines
   - Repeat periodically

CMG StLouis 2016 - Virtual Environments Optimization - © Moviri 2016
# Rightsizing – Scope Breakdown

<table>
<thead>
<tr>
<th>Category</th>
<th>In scope</th>
<th>Not in scope for decommissioning</th>
<th>Data availability &gt;= 80% (180 days)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All VMs</td>
<td>92%</td>
<td>9%</td>
<td>73%</td>
</tr>
<tr>
<td>Server</td>
<td>77%</td>
<td>23%</td>
<td>27%</td>
</tr>
<tr>
<td>VDI</td>
<td>1 vCPU</td>
<td>23%</td>
<td>27%</td>
</tr>
</tbody>
</table>

**CMG StLouis 2016 - Virtual Environments Optimization - © Moviri 2016**
What more than already available?

- Historical data **timeframe** has been increased to 6 months (from 30 days)
- Historical **trends** has been considered to ensure consumption growths
- Specific **metrics** for specific **technologies** (e.g. entitlement considerations on AIX LPARs, …)
- Everything is **parametric**: different recommendations for different environments (e.g. very conservative on PROD, aggressive on TEST, …)
- Ad-hoc rules for the specific environment
Rightsizing – Example

Configuration = 4 vCPUs
99th percentile = 8.4%
Target Threshold = 40%

CMG StLouis 2016 - Virtual Environments Optimization - © Moviri 2016
Rightsizing – Example

Configuration = 4 vCPUs
99\textsuperscript{th} percentile = 8.4%
Target Threshold = 40%
Growth = 2.97%
Rightsizing – Example

Configuration = 4 vCPUs
99th percentile = 8.4%
Target Threshold = 40%
Growth = 2.97%
Suggestion = 2 vCPU

Configuration = 4 vCPUs
99th percentile = 8.4%
Target Threshold = 40%
Growth = 2.97%
Suggestion = 2 vCPU
Rightsizing – Example

Configuration = 4 vCPUs
99th percentile = 8.4%
Target Threshold = 40%
Growth = 2.97%
Suggestion = 2 vCPU

Chosen target threshold

With 2 vCPU («suggested» configuration)

With 4 vCPU (current configuration)

CMG StLouis 2016 - Virtual Environments Optimization - © Moviri 2016
Rightsizing – Example

Configuration = 4 vCPUs
99th percentile = 8.4%
Target Threshold = 40%
Growth = 2.97%
Suggestion = 2 vCPU

With 2 vCPU («suggested» configuration)

With 4 vCPU (current configuration)

CMG StLouis 2016 - Virtual Environments Optimization - © Moviri 2016
Rightsizing – Target CPU utilization choice

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
<th>Scenario 4</th>
<th>Scenario 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>99th</td>
<td>99th</td>
<td>99th</td>
<td>99th</td>
<td>99th</td>
<td>99th</td>
</tr>
<tr>
<td>Data Availability</td>
<td>&gt; 80%</td>
<td>&gt; 80%</td>
<td>&gt; 80%</td>
<td>&gt; 80%</td>
<td>&gt; 80%</td>
</tr>
<tr>
<td>Target CPU Utilization</td>
<td>20%</td>
<td>30%</td>
<td>40%</td>
<td>50%</td>
<td>60%</td>
</tr>
<tr>
<td>Savings [% of all vCPUs]</td>
<td>24%</td>
<td>32%</td>
<td>36%</td>
<td>39%</td>
<td>41%</td>
</tr>
</tbody>
</table>

Best trade-off: Performance-Savings

Savings increase:
- +30% for Scenario 1
- +14% for Scenario 2
- +8% for Scenario 4
- +6% for Scenario 5
Rightsizing – Results

Virtual Machines Distribution by CPU Utilization

Downsizing Implications

- **36% of saved vCPU → $ savings**
- No impact on VMs over «target threshold»
- No impact on performance
- Less software licenses
- Increased host efficiency and decreased density
Conclusion

The solution is a **non-invasive improvement** of de-provisioning processes; usage of both performance data from any Virtual Estate and owners’ data from several heterogeneous repositories is achievable in a fully automated fashion.

Several M$/year saved achieved by identifying machines to be optimized and reaching out to hundreds of virtual machines’ owners and getting their approval: **shut down 9% of Virtual Machines**

Most important benefits:

• Less storage allocated
• Less servers in scope of Business Continuity and Disaster Recovery
• Less 3rd party software licenses

CMG StLouis 2016 - Virtual Environments Optimization - © Moviri 2016
## Contacts

<table>
<thead>
<tr>
<th>Headquarters</th>
<th>USA East</th>
<th>USA West</th>
</tr>
</thead>
<tbody>
<tr>
<td>Via Schiaffino 11</td>
<td>283 Franklin Street</td>
<td>425 Broadway Street</td>
</tr>
<tr>
<td>20158 Milan Italy</td>
<td>Boston, MA 02110</td>
<td>Redwood City, CA 94063</td>
</tr>
</tbody>
</table>