Software-Defined Infrastructure Control

Capacity Management for Hybrid IT

Dan Adirim
SVP, Customer Management
Factors like cost and risk are now part of every placement decision:
- Initially for new workloads
- Ongoing re-evaluation of existing workloads

Cost paradigm has shifted from capital to expense:
- Not cost avoidance
- Provider benefits from overcommit
- Rightsizing leads to real savings
How Does it Work?

- Cirba predictively analyzes workload patterns to optimize workload density
- While at the same time reducing workload contention and operational risk

VM1 – Busy in the morning

VM2 – Busy in the evening

Predictively identifying complementary workload patterns drives higher density, dramatically reducing infrastructure costs

Predictively identifying potential conflict prevents workload contention before it occurs, making environments run better
Reactive approaches don’t move VMs until the damage is already done. Resource contention will have already occurred for at least 5-10 minutes, significantly impacting end users.

(Note: In this example DRS will never act, as it doesn’t look at Disk I/O)

In the month after Cirba was enabled the number of reactive VM moves during business hours dropped over 80%

This is a clear indication of reduced resource contention.
Policy Enablement

- Cirba uses a concept we call “policy” to encode how you want to manage your environment.
- These rules make the same infrastructure behave differently depending on the workload you are running.

<table>
<thead>
<tr>
<th></th>
<th>Production</th>
<th>Critical Production</th>
<th>Non Critical Production</th>
<th>Dev/Test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Density</strong></td>
<td>Low</td>
<td>Med</td>
<td>Med</td>
<td>High</td>
</tr>
<tr>
<td><strong>Performance</strong></td>
<td>High</td>
<td>Med</td>
<td>Med</td>
<td>Low</td>
</tr>
<tr>
<td><strong>Availability</strong></td>
<td>N+2</td>
<td>N+1</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Compliance</strong></td>
<td>Rigorous</td>
<td>Medium</td>
<td>Multi-Tenant</td>
<td>None</td>
</tr>
<tr>
<td><strong>Volatility</strong></td>
<td>Low</td>
<td>Med</td>
<td>High</td>
<td>Med</td>
</tr>
<tr>
<td><strong>Operational Cycles</strong></td>
<td>Business Defined</td>
<td>IT Defined</td>
<td>Unbound</td>
<td>Simulated</td>
</tr>
<tr>
<td><strong>Licensing</strong></td>
<td>Host-Based</td>
<td>Catalog-Based</td>
<td>Transient</td>
<td>Developer Edition</td>
</tr>
</tbody>
</table>

Cirba uses a concept we call “policy” to encode how you want to manage your environment. These rules make the same infrastructure behave differently depending on the workload you are running.
Cirba’s sophisticated analytics and policy can be used to safely rightsize VMs

Business cycle has peaks of high utilization throughout the month.

Cirba workload analysis sees 85% utilization and recommends a bump-up.

Other solutions using 90th percentile data see 3.25% and recommend a **bump-down**, which would be catastrophic to the app.
The Importance of Analyzing Workload Patterns

- CPU Intensive
- Start of Day

- Memory Intensive
- End of Day
The Importance of Analyzing Workload Patterns

Average Increase in VM Density: 48%
Average Hardware Savings: 33%
Average Software License Savings: 55%

Operational Risk
Stranded Capacity
Using Analytics to Automatically Route New Applications

New Application or Cloud Request

Windows OS
Requires SQL server
Has customer data
Needs Gold tier storage
Must run on west coast

New Application or Cloud Request

Public Cloud

New York
San Francisco
Toronto
London
Singapore

SOFTLAYER®
Bare Metal
What’s Different for Public Cloud?
Challenge with Public Cloud Adoption

- Customers need agility and real-time provisioning *and* enterprise-class governance and control over hosting decisions
- Security, Compliance, Performance, Proximity, Cost

Codify
Govern
Optimize
Automate

- Amazon Web Services
- Microsoft Azure
- SoftLayer
Cirba and Hybrid Cloud Optimization

Off-Prem IaaS
- Placement Optimization
- Bump-Up
- Reclaim

Off-Prem Bare Metal
- Placement Optimization
- Bump-Up
- Reclaim

On-Prem Virtual & Cloud
- Placement Optimization
- Bump-Up
- Reclaim

Intelligent Hybrid Routing
- Cross-Provider Routing
- Cross-Cluster Workload Routing
- Capacity Reservation
Cirba’s public cloud data collection discovers cloud instances running in Amazon, Azure and SoftLayer, providing seamless views and analysis across both on-prem and public cloud workloads.

Detailed workload patterns are tracked across CPU, memory, disk and network I/O. Full historical data retention enables analysis of operational patterns and business cycles, and Cirba’s benchmark system enables normalization between platforms.
Public Cloud Optimization

Data is analyzed against cloud catalogs to uncover allocation risks, which put hybrid cloud initiatives in danger, or inefficiencies, which directly lead to over-spend.

Detailed recommendations are generated to remediate issues and optimize public cloud spend.

Allocation Risk and Efficiency - System Details

<table>
<thead>
<tr>
<th>System</th>
<th>Overall Status</th>
<th>OS</th>
<th>Recommended Allocations</th>
</tr>
</thead>
<tbody>
<tr>
<td>172.30.0.197</td>
<td>At Risk</td>
<td>Linux</td>
<td>platform: m3.medium-1X3.75-Linux, CPU: 1, Memory: 4, Disk: 4</td>
</tr>
<tr>
<td>172.30.0.198</td>
<td>Excess Capacity</td>
<td>Windows</td>
<td>m4.large-2X8-Windows, CPU: 1, Memory: 30, Disk: 30</td>
</tr>
<tr>
<td>172.30.0.203</td>
<td>Excess Capacity</td>
<td>Windows</td>
<td>m4.large-2X8-Windows, CPU: 1, Memory: 40, Disk: 40</td>
</tr>
<tr>
<td>172.30.0.212</td>
<td>At Risk</td>
<td>Linux</td>
<td>m3.medium-1X3.75-Linux, CPU: 1, Memory: 6, Disk: 6</td>
</tr>
</tbody>
</table>
The Impact of Properly Right-Sizing Cloud Workloads

- T-Shirt instance sizing model
- Cost based on catalog size
- Typically sized to peak utilization
- User pays for capacity whether it is used or not (no overcommit)

Cirba analyzed **983 real workloads** to determine the impact of optimization analysis on the 1-year hosting cost in Amazon AWS:

**Sized As-Is (No Optimization):**

$2,368,899

**Sizing Optimized with Cirba:**

$1,892,733

**Net Impact:** **20%**

Note that more aggressive reclaim policies were found to yield up to 46% savings.
Cirba Version 9: Hybrid Routing Analysis
Modeling Detailed Application Requirements
Cirba Provides Seamless Capacity Planning for Hybrid IT

- Enables truly automated, real-time provisioning
  - Factor security and reputational risk into placements
  - Use policy to place workload in the right cloud environment

- Enables immediate and ongoing cost savings
  - Rightsizing
  - Predict the financial impact of placements
  - Be able to identify and correct poor placements quickly

- Cirba can help you be successful in a hybrid world
Thank You!