Managing Modern IT

Measuring the Success of Cloud-Based Services

Mahesh Ramachandran
Senior Director, Product Management, Vistara

Eric Krock
Director of Product Marketing, Vistara
Agenda

- Cloud and metric driven management
- Relevant IT metrics and considerations
- Process and best practices
- How Vistara can help
Cloud and the imperative for IT metrics

Cloud...

- Represents a new model for IT service delivery
  - Resources consumed on demand, sourced from a mix of **private** and **public** clouds
- Brings transparency to **service levels** and **costs** of IT resources
- Requires IT to manage operations to meet specific service level and cost goals

Lines of Business

[Diagram of cloud services and IT operations]

**What metrics should Enterprise IT measure and track?**

<table>
<thead>
<tr>
<th>Resource</th>
<th>Service Level</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Instance</td>
<td>1CPU/2GB</td>
<td>$0.05/hour</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Cloud and the imperative for IT metrics

IT metrics should measure

- Service levels delivered to lines of business – **Service Level Metrics**
- Effectiveness and efficiency with which services are delivered – **Service Delivery Metrics**

### Service Level Metrics
- Availability
- Performance
- Service Fulfillment

### Service Delivery Metrics
- Problem
- Response
- Capacity & Cost

---

**Lines of Business**

**Enterprise IT**

**Public Cloud**

**Private Cloud**

<table>
<thead>
<tr>
<th>Resource</th>
<th>Service Level</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Instance</td>
<td>1CPU/2GB</td>
<td>$0.05/hour</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

### Lines of Business

- 
- 
- 

---

**Resource**

**Service Level**

**Cost**

| ...               | ...          | ...    |
| ...               | ...          | ...    |
| ...               | ...          | ...    |
Agenda

• Cloud and metric driven management
• Relevant IT metrics and considerations
• Process and best practices
• How Vistara can help
What Availability metrics measure

- Quantifies up or down status of services, as experienced by the end user
  - % of time service is available for use, over a time interval - e.g. last month

What to measure

<table>
<thead>
<tr>
<th></th>
<th>Typical Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Servers</td>
<td>Ping Response, OS service or agent status</td>
</tr>
<tr>
<td>Network Devices</td>
<td>Ping Response, Interface/Port Status</td>
</tr>
<tr>
<td>Applications, URLs, Web Services</td>
<td>Ping, HTTP GET</td>
</tr>
</tbody>
</table>
Availability

Considerations

• **Element level availability vs. availability of an entire service**
  – Do you track availability of individual elements or an entire service compose of multiple elements?

• **Is Availability a binary quantity (Up / Down) or more fine grained?**
  – How do you account services that are available, but running with degraded redundancy/performance?

• **Planned Downtime vs. Unplanned Downtime**
  – Do you count downtime due to planned activities (e.g. patching)?
Performance

What Performance metrics measure

- Quantifies responsiveness to user or application requests
  - Throughput and latency delivered by the service - OS, application ...

What to measure

<table>
<thead>
<tr>
<th></th>
<th>Typical Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Servers, Cloud Instances</td>
<td>CPU queue length, memory swap rate, I/O latency</td>
</tr>
<tr>
<td>Applications, URLs, Web Services</td>
<td>Response time to requests - transactions, http requests/response, API calls</td>
</tr>
</tbody>
</table>
Considerations

- **Focusing on the right metrics**
  - Apps, hypervisors expose tons of metrics – which ones are the best measures of performance?

- **Using the right statistical measures**
  - How do you define “bad” user response time – average, max, percentile?

- **Performance of cloud instances**
  - How can you tell if your cloud instances are getting promised resources?

---

**Cloud Instance Performance**

- CPU
  - Promised
  - Received
  - Time

- I/O Latency
  - Inconsistent I/O performance
  - Time
Service Fulfillment

What service fulfillment measures
• Quantifies timely fulfillment of lifecycle management tasks – e.g.
  – Service Requests, Patching, Anti-Virus, Network Configuration Backups ...

What to measure...

<table>
<thead>
<tr>
<th></th>
<th>Typical Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service Requests</td>
<td>Time to resolve service requests</td>
</tr>
<tr>
<td>Server Patching</td>
<td>Patching frequency, time lag between patch release and patch roll out</td>
</tr>
<tr>
<td>Anti-Virus</td>
<td>Servers up-to-date on anti-virus signature</td>
</tr>
</tbody>
</table>

Considerations
• Focusing on the right service attributes to measure
  – E.g. Patches – all types of patches vs. only critical security patches
What Problem metrics measure

- Provide actionable insight into causes of problems

What to measure...

<table>
<thead>
<tr>
<th>Type</th>
<th>Practical Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Common problem types</td>
<td>Problem and incident tickets by problem type</td>
</tr>
<tr>
<td>Failure modes of devices &amp; applications</td>
<td>OS / application error codes</td>
</tr>
</tbody>
</table>

Considerations

- Identifying the right data sources to mine
  - Do your tickets contain enough information to classify problems accurately?
- Interpreting collected data
  - How do you make sense of OS / application error codes?
Response

What Response metrics measure

- Quantifies how fast you are resolving incidents and problems

What to measure...

<table>
<thead>
<tr>
<th>Type</th>
<th>Practical Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incident and problem resolution times</td>
<td>Ticket resolution times</td>
</tr>
<tr>
<td>Incident and problem resolution rate</td>
<td>Ticket resolution rate</td>
</tr>
<tr>
<td>Correctness and completeness of resolution</td>
<td>Ticket re-opens, Ticket volume by problem type</td>
</tr>
</tbody>
</table>

Considerations

- Identifying the right data sources to mine
  - Do your tickets contain enough information to classify problems accurately?
Capacity and Cost

What Capacity and Cost metrics measure

• How efficiently you are using capacity
• How much capacity you will likely need in the future
• How much your services are costing you to deliver

What to measure...

<table>
<thead>
<tr>
<th>Type</th>
<th>Practical Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resource utilization</td>
<td>Utilization of server, storage, switch/router interfaces, unused virtual machines</td>
</tr>
<tr>
<td>Demand forecasts</td>
<td>Number of VMs over time, average VM sizes over time</td>
</tr>
<tr>
<td>Hardware costs</td>
<td>Server and software license costs</td>
</tr>
<tr>
<td>Resource usage accounting</td>
<td>Compute, storage, network used by different types of virtual machines</td>
</tr>
</tbody>
</table>

Considerations

• Resource usage accounting is challenging in a shared infrastructure
• Identifying the right data sources
Agenda

• Cloud and metric driven management
• Relevant IT metrics and considerations
• Process and best practices
• How Vistara can help
High Performing Organizations are Metrics-Driven

- Below-average performers show lower use of IT metrics across the board.
- Particularly true for external benchmarking.

<table>
<thead>
<tr>
<th>Metric used</th>
<th>Top</th>
<th>Average</th>
<th>Below-average</th>
</tr>
</thead>
<tbody>
<tr>
<td>“contribution to business value or business outcome”</td>
<td>56%</td>
<td>21%</td>
<td>9%</td>
</tr>
<tr>
<td>“comparison of benchmarking against others”</td>
<td>39%</td>
<td>15%</td>
<td>8%</td>
</tr>
</tbody>
</table>

IT Must Proactively Define & Drive Measurement

• Gathering data for benchmarking takes time.
• If IT must gather data reactively in response to management demands, delay harms credibility.
• Metrics without a historical baseline will be questioned.
• Agreeing on right metrics and achieving apples-to-apples comparisons takes time.

IT Metrics Definition and Measurement Process

Cloud & IT Self-Service Drive Cost Transparency

• Cloud gives organizations a choice between internal and cloud-based services
• IT self-service empowers line of business to make the choice
• Therefore embedding clear, understandable cost information in IT service catalog is critical to enable rational, optimal choices
Getting ready for service pricing

1. IT understands benefits of service pricing and why transition is happening
2. List of IT services validated with LOB
3. Prices can be compared with market prices
4. Business case approved for transition
5. IT has chart of accounts to enable pricing
6. Organization understands resources needed for initiative to succeed

Additional metrics to consider for cloud services

- **Agility:** Can IT service respond in real-time to changes in demand & business needs?
- **Elasticity:** How much can service’s capacity scale upward and downward to match demand?
- **Continuity:** Maximum time between interruptions
- **Consistency:** Can variations in service levels be reduced?

Agenda

• Cloud and metric driven management
• Relevant IT metrics and considerations
• Process and best practices
• How Vistara can help
How Vistara can help

- **Data**: collects relevant data
  - 100s of built-in monitoring templates to collect relevant metrics across
    - Servers, applications, storage, and network
- **Analytics**: analyzes data to extract relevant metrics
  - Availability, performance, problem, response, ...
- **Reports**: presents metrics to help you gain insight
  - Reports to spot key trends and patterns in key metrics
- **Action**: single tool for managing entire system in response to metrics
How Vistara can help: service catalog

- Service catalog
  - Internal and external services
  - Embedded cost model to drive rational choices by IT and LOB

Add Storage capacity to vSphere Environment
Add NFS Data store to VMware. It creates a NFS export for a given Storage Volume and then creates NAS Data store using the NFS mount.

Create Network

Managed Windows Instance in your secure Cloud
Running Microsoft Windows Server® 2012 is a fast and dependable environment for deploying applications using the Microsoft Web Platform, including ASP.NET, ASP.NET AJAX, Silverlight™, and Internet Information Server (IIS). You can run any compatible Windows-based solution on this high-performance, reliable, cost-effective, virtual

On Demand and Managed Cloud Instance - Pay as You Go
Get it going in minutes by requesting to provision a server with these specifications. Offered with decent customization on the configuration, the created server is allowed for root access. The server is automatically added as a managed instance for monitoring KPI (Key Performance Indicators) (or) health of the server and also enabled

Pay Per Use - LAMP Server - Gold Profile
Pay for your server usage on a rounded up hour-by-hour basis. The server comes with a private IP address and gets bundled with the Service Level as per the Profile associated. (View the Gold Profile specifications inside). vCPU 4 3Ghz, Memory 8 GB, NIC 1
Agenda

• Cloud and metric driven management
• Relevant IT metrics and considerations
• Process and best practices
• How Vistara can help
Let’s Get Started Today

Schedule a Demo!

Attend Our Next Webinars!

Email
sales@vistarait.com

Website
www.vistarait.com