BITS Pilani presentation

BITS Pilani
Hyderabad Campus

D. Powar
Lecturer,
BITS-Pilani, Hyderabad Campus

innovate achieve lead

BITS Pilani
Hyderabad Campus

SSZG527

Cloud Computing

Agenda:

- **Service License Agreements:** Lifecycle and Management
  - INSPIRATION
  - TRADITIONAL APPROACHES TO SLO MANAGEMENT
  - TYPES OF SLA's
  - LIFE CYCLE OF SLA
  - SLA MANAGEMENT IN CLOUD
  - AUTOMATED POLICY-BASED MANAGEMENT
- Managing Clouds: Services and Infrastructure
  - Managing IaaS
  - Managing PaaS
  - Managing SaaS

Amazon EC2 SLA ?

innovate achieve lead

aws.amazon.com/ec2/sla/

Employees' Prov... Laird OnDemand: Vi... Imported From Firef... Cloud Login to BITS Pilani ... Call for papers Admin Login Watch Live Indian T... Cloud Computing B...

amazon web services

English Sign In Sig

PRODUCTS & SERVICES

- Amazon EC2
- Product Details
- Instances
- Pricing
- Previous Generation Instances
- Purchasing Options
  - Amazon EC2 Spot Instances
  - Amazon EC2 Reserved Instances
  - Amazon EC2 Dedicated Instances
- Developer Resources
- FAQs
- Amazon EC2 SLA
- AWS Management Portal for vCenter

RELATED LINKS

- Windows Instances
- VM Import/Export
- Management Console

Amazon EC2 Service Level Agreement

Effective Date: June 1, 2013

This Amazon EC2 Service Level Agreement ("SLA") is a policy governing the use of Amazon Elastic Compute Cloud ("Amazon EC2") and Amazon Elastic Block Store ("Amazon EBS") under the terms of the Amazon Web Services Customer Agreement (the "AWS Agreement") between Amazon Web Services, Inc. ("AWS", "us" or "we") and users of AWS' services ("you"). This SLA applies separately to each account using Amazon EC2 or Amazon EBS. Unless otherwise provided herein, this SLA is subject to the terms of the AWS Agreement and capitalized terms will have the meaning specified in the AWS Agreement. We reserve the right to change the terms of this SLA in accordance with the AWS Agreement.

Service Commitment

AWS will use commercially reasonable efforts to make Amazon EC2 and Amazon EBS each available with a Monthly Uptime Percentage (defined below) of at least 99.95%, in each case during any monthly billing cycle (the "Service Commitment"). In the event Amazon EC2 or Amazon EBS does not meet the Service Commitment, you will be eligible to receive a Service Credit as described below.

Definitions

- "Monthly Uptime Percentage" is calculated by subtracting from 100% the percentage of minutes during the month in which Amazon EC2 or Amazon EBS, as applicable, was in the state of "Region Unavailable." Monthly Uptime Percentage measurements exclude downtime resulting directly or indirectly from any Amazon EC2 SLA Exclusion (defined below).
- "Region Unavailable" and "Region Unavailability" mean that more than one Availability Zone in which you are running an instance, within the same Region, is "Unavailable" to you.
- "Unavailable" and "Unavailability" mean:
  - For Amazon EC2, when all of your running instances have no external connectivity.
  - For Amazon EBS, when all of your attached volumes perform zero read write IO, with pending IO in the queue.

Amazon EC2 SLA (contd..) ?

https://aws.amazon.com/ec2/sla/

Amazon Web Services
PRODUCTS & SERVICES

- Amazon EC2
- Product Details
- Instances
- Pricing
- Purchasing Options
- Developer Resources
- FAQs
- Getting Started
- Amazon EC2 Dedicated Hosts

RELATED LINKS

- Amazon EC2 Spot Instances
- Amazon EC2 Reserved Instances
- Amazon EC2 Dedicated Instances
- Windows Instances

Service Commitments and Service Credits

Service Credits are calculated as a percentage of the total charges paid by you (excluding one-time payments such as upfront payments made for Reserved Instances) for either Amazon EC2 or Amazon EBS (whichever was Unavailable, or both if both were Unavailable) in the Region affected for the monthly billing cycle in which the Region Unavailability occurred in accordance with the schedule below.

Monthly Uptime Percentage	Service Credit Percentage
Less than 99.95% but equal to or greater than 99.0%	10%
Less than 99.0%	30%

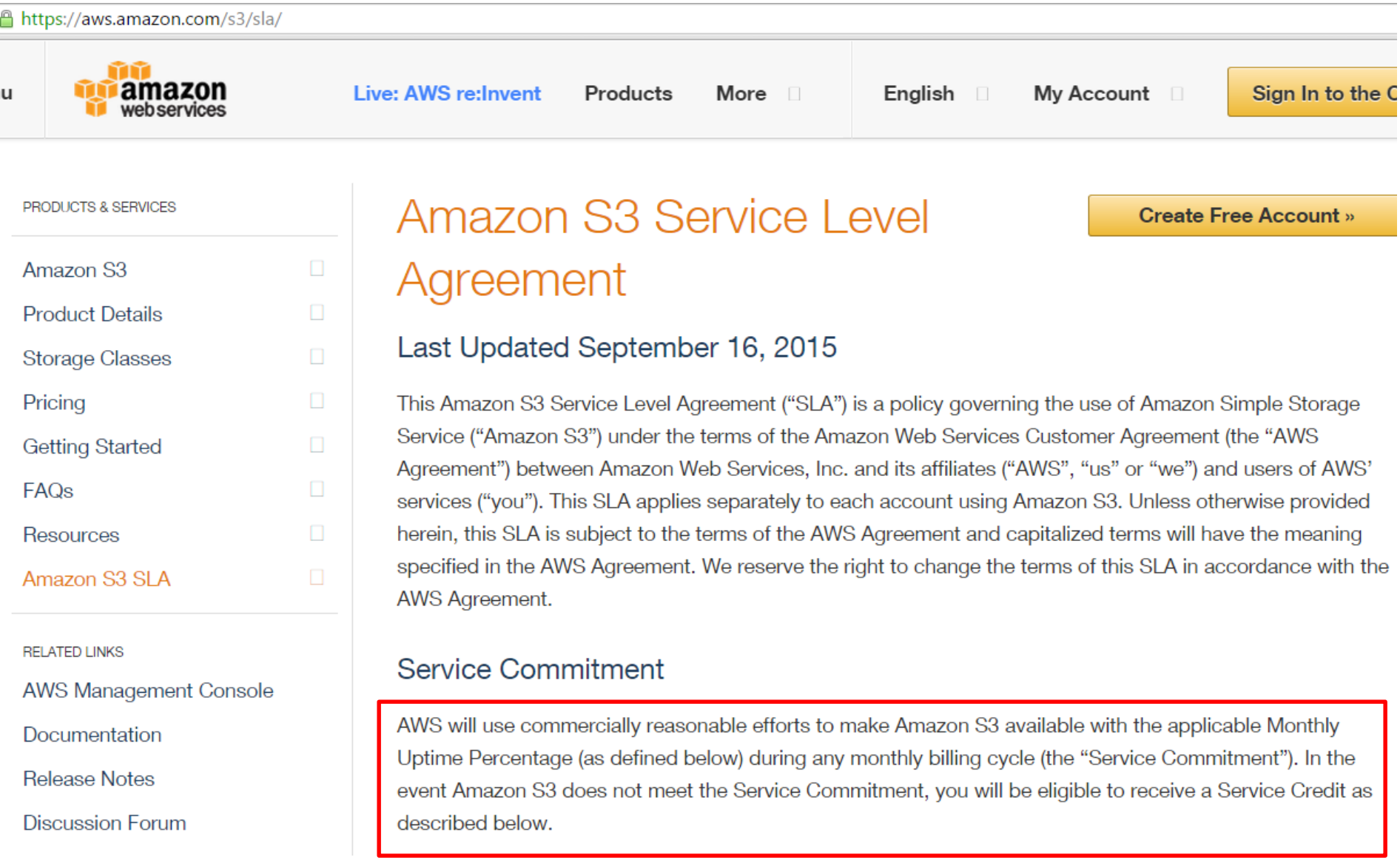
We will apply any Service Credits only against future Amazon EC2 or Amazon EBS payments otherwise due from you. At our discretion, we may issue the Service Credit to the credit card you used to pay for the billing cycle in which the Unavailability occurred. Service Credits will not entitle you to any refund or other payment from AWS. A Service Credit will be applicable and issued only if the credit amount for the applicable monthly billing cycle is greater than one dollar ($1 USD). Service Credits may not be transferred or applied to any other account. Unless otherwise provided in the AWS Agreement, your sole and exclusive remedy for any unavailability, non-performance, or other failure by us to provide Amazon EC2 or Amazon EBS is the receipt of a Service Credit (if eligible) in accordance with the terms of this SLA.

Credit Request and Payment Procedures

To receive a Service Credit, you must submit a claim by opening a case in the AWS Support Center. To be eligible, the credit request must be received by us by the end of the second billing cycle after which the incident occurred and must include:

1. the words "SLA Credit Request" in the subject line;

Amazon S3 SLA ?



Amazon S3 SLA (contd..) ?

Service Commitment

AWS will use commercially reasonable efforts to make Amazon S3 available with the applicable Monthly Uptime Percentage (as defined below) during any monthly billing cycle (the "Service Commitment"). In the event Amazon S3 does not meet the Service Commitment, you will be eligible to receive a Service Credit as described below.

Definitions

- "Error Rate" means: (i) the total number of internal server errors returned by Amazon S3 as error status "InternalError" or "ServiceUnavailable" divided by (ii) the total number of requests for the applicable request type during that five minute period. We will calculate the Error Rate for each Amazon S3 account as a percentage for each five minute period in the monthly billing cycle. The calculation of the number of internal server errors will not include errors that arise directly or indirectly as a result of any of the Amazon S3 SLA Exclusions (as defined below).
- "Monthly Uptime Percentage" is calculated by subtracting from 100% the average of the Error Rates from each five minute period in the monthly billing cycle.
- A "Service Credit" is a dollar credit, calculated as set forth below, that we may credit back to an eligible Amazon S3 account.

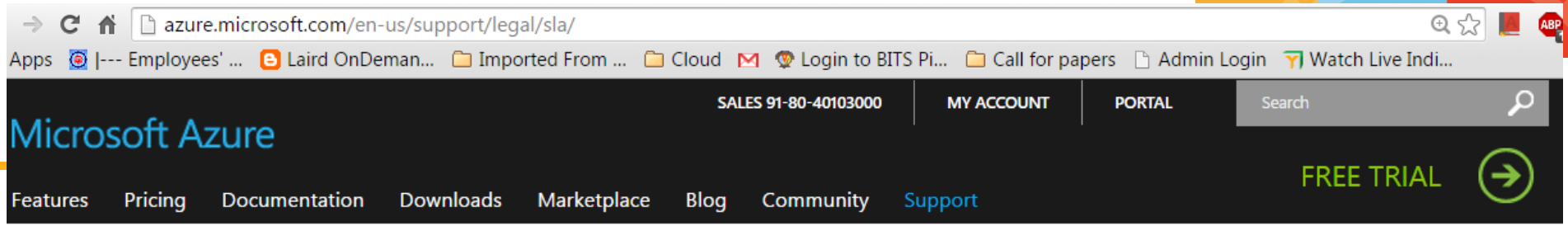
Service Credits

Service Credits are calculated as a percentage of the total charges paid by you for Amazon S3 for the billing cycle in which the error occurred in accordance with the schedule below.

For all requests not otherwise specified below:

Monthly Uptime Percentage	Service Credit Percentage
Equal to or greater than 99.0% but less than 99.9%	10%
Less than 99.0%	25%

Microsoft Azure SLA ?



Cache

We guarantee at least 99.9% of the time that customers will have connectivity between the Cache endpoints and our Internet gateway.

CDN

We guarantee that at least 99.9% of the time CDN will respond to client requests and deliver the requested content without error. We will review and accept data from any commercially reasonable independent measurement system that you choose to monitor your content. You must select a set of agents from the measurement system's list of standard agents that are generally available and represent at least five geographically diverse locations in major worldwide metropolitan areas (excluding PR of China).

Cloud Services, Virtual Machines and Virtual Network

- For Cloud Services, we guarantee that when you deploy two or more role instances in different fault and upgrade domains, your Internet facing roles will have external connectivity at least 99.95% of the time.
- For all Internet facing Virtual Machines that have two or more instances deployed in the same Availability Set, we guarantee you will have external connectivity at least 99.95% of the time.
- For Virtual Network, we guarantee a 99.9% Virtual Network Gateway availability.

ExpressRoute

We guarantee a minimum of 99.9% ExpressRoute dedicated circuit availability.

Microsoft Azure SLA (contd..) ?

innovate achieve lead

Microsoft Azure

Why Azure Products Documentation Pricing Partners Blog Resources Support

stopping or deleting the Virtual Machines.

"**Virtual Machine**" refers to persistent instance types that can be deployed individually or as part of an Availability Set.

Downtime

The total accumulated minutes that are part of Maximum Available Minutes that have no External Connectivity.

Monthly Uptime Percentage

The Monthly Uptime Percentage is calculated using the following formula:

(Maximum Available Minutes - Downtime) / Maximum Available Minutes * 100

Service Credit

MONTHLY UPTIME PERCENTAGE	SERVICE CREDIT
< 99.95%	10%
< 99%	25%

INSPIRATION

Provisioning server resources:

- Deciding hardware configuration, determining the number of physical machines, and acquiring them upfront so that the overall business objectives could be achieved

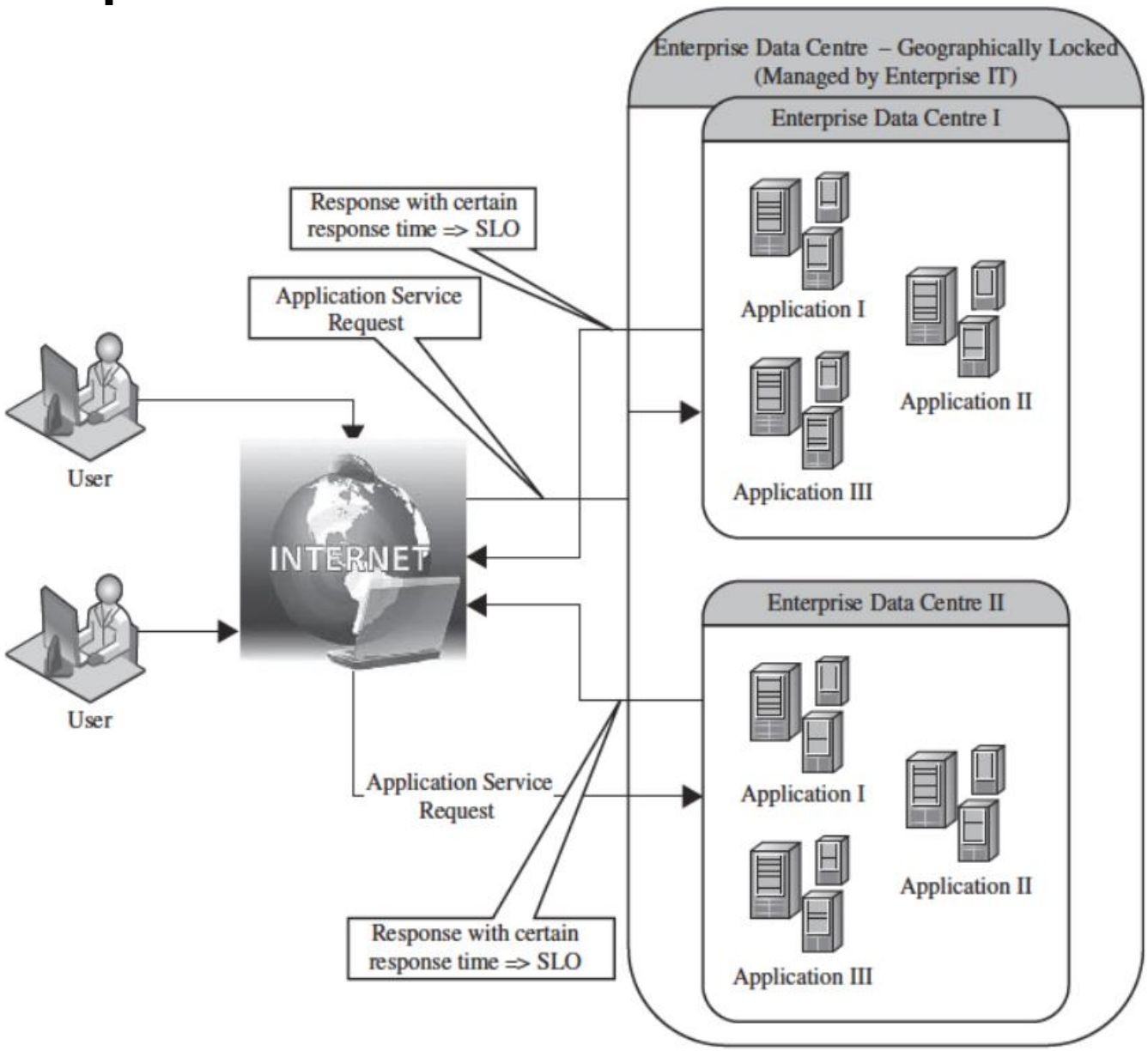
Service level Objectives (SLO):

- ❖ Response time
- ❖ Throughput of the application to end user requests

Capacity Planning:

- ✓ The activity of determining the number of servers and their capacity that could satisfactorily serve the application end-user requests at peak loads

Example



Hosting of applications on servers within Enterprise's data centers

Can I deploy my app some where else??

- Outsourcing of applications hosting to 3rd party infrastructure provider is economical for the following reasons:
  - The enterprise need not invest in procuring expensive hardware upfront without knowing the viability of the business
  - Maintenance of the applications and hardware were non core activities of the enterprise
  - Cost to maintain the data centers increased as the applications grew because sophisticated tools were required
- Therefore enterprises developed the app's and deployed on the infrastructure of the 3rd party service providers

SLA – Service level Agreement

- Hence enterprises enter into a legal agreement – SLA (Service Level Agreement)with the infrastructure service providers to guarantee a minimum quality of service (QoS)
- General QoS Parameters:
  - System CPU
  - Data storage
  - Network bandwidth
- SLA rules:
  - ✓ The application's server machine will be available for **99.9%** of the key business hours of the application's end users, also called core time, and 85% of the non-core time
  - ✓ The service provider would respond to a reported issue in less than 10 minutes during the core time, but would respond in one hour during non-core time
- These SLA are termed as Infrastructure SLA and Providers are called ASP (Applications Service Providers)

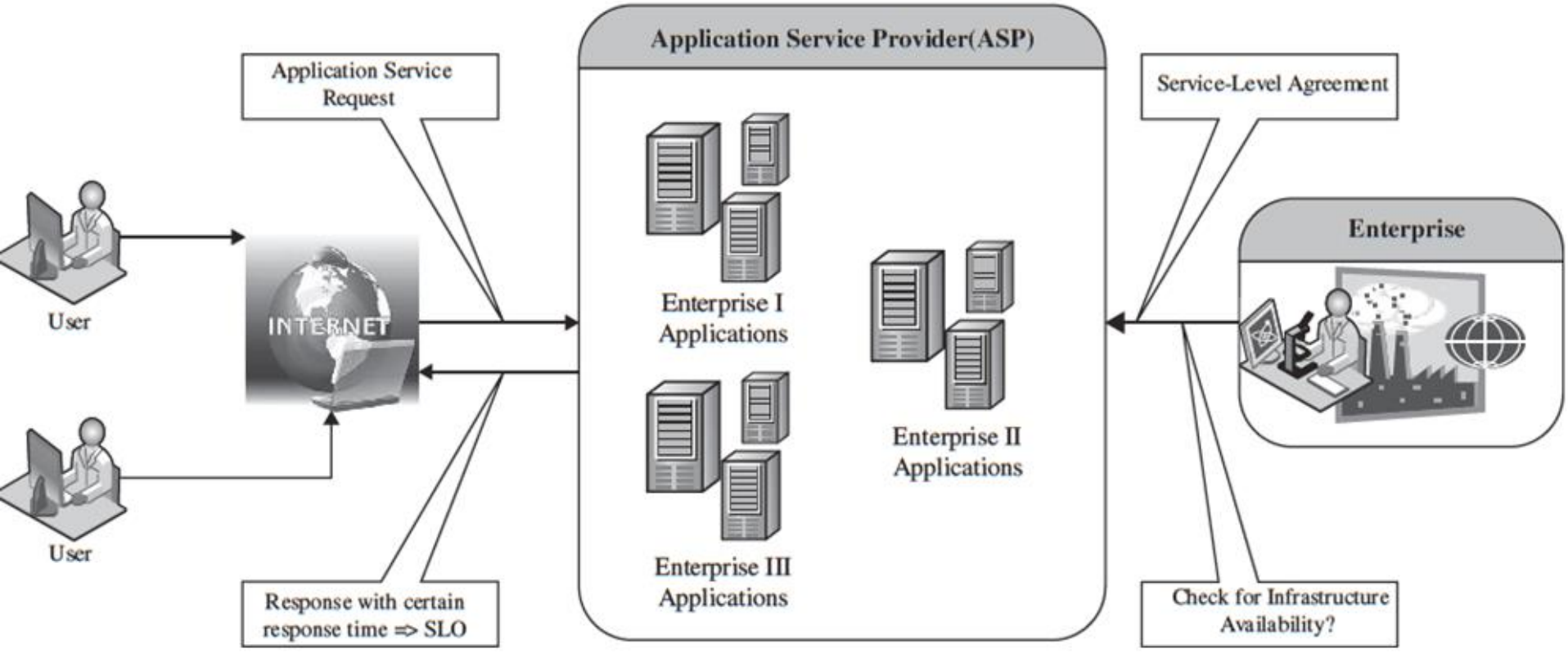


FIGURE 16.2. Dedicated hosting of applications in third party data centers.

Advantages of App's deployed on ASP

- ASP's data centers servers underutilized because the applications were not fully utilizing their servers' capacity at nonpeak loads.
- To reduce the redundancies and increase the server utilization in data centers, ASPs started co-hosting applications with complementary workload patterns.
- Co-hosting of applications means deploying more than one application on a single server.

Consumer and Provider Perspective

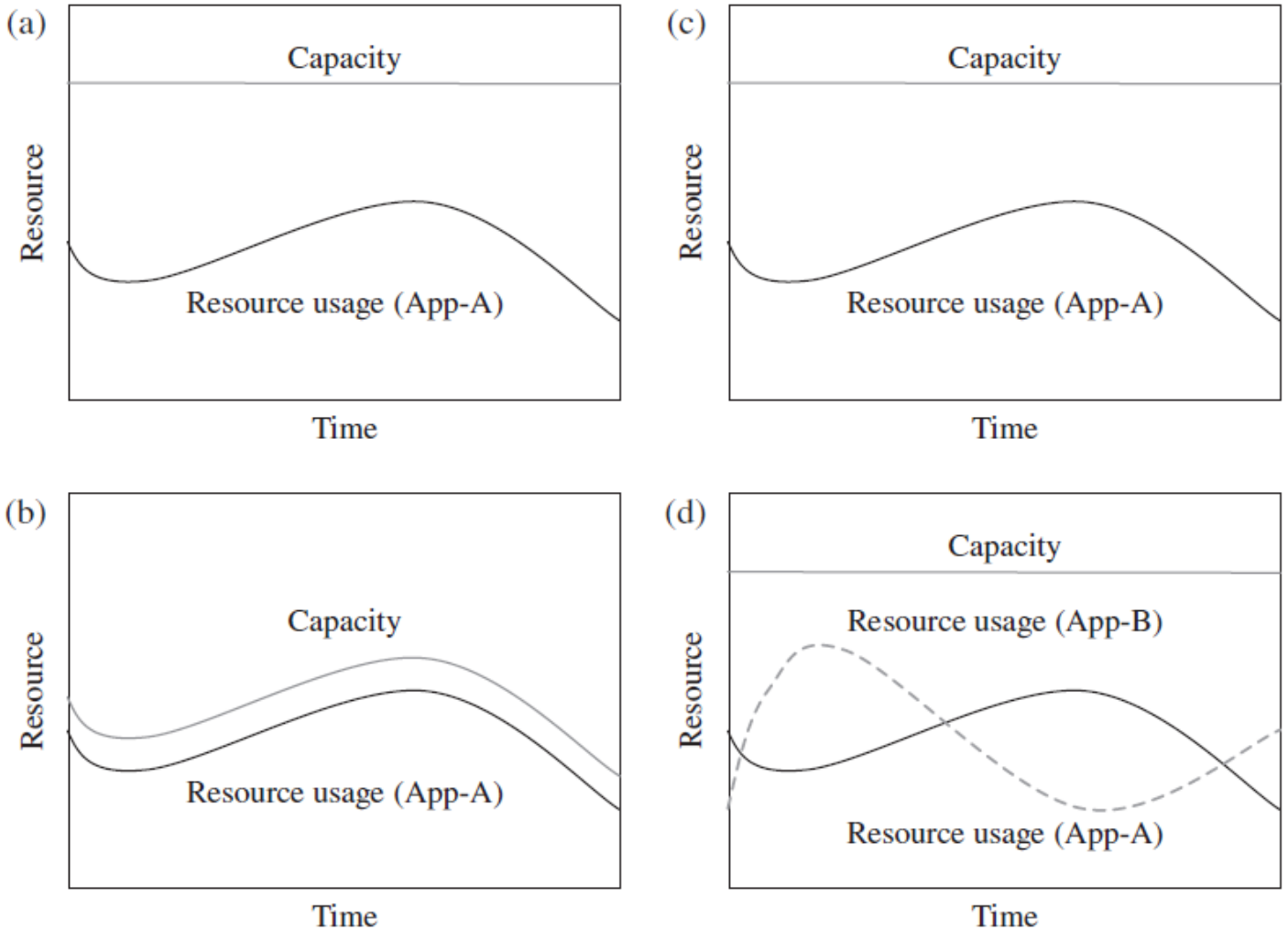


FIGURE 16.3. Service consumer and service provider perspective before and after the MSP's hosting platforms are virtualized and cloud enabled. (a) Service consumer perspective earlier. (b) Service consumer perspective now. (c) Service provider perspective earlier. (d) Service provider perspective now.

Issues with Co-Hosting

- Application performance isolation
  - Performance isolation implies that one application should not steal the resources being utilized by other co-located applications
  - For example, application A is required to use more quantity of a resource than originally allocated to it for duration of time t (Peak time traffic). For that duration the amount of the same resource available to application B is decreased
  - This could adversely affect the performance of application B
- Security guarantee
  - One application should not access and destroy the data and other information of co-located applications
- These challenges prevented ASPs from fully realizing the benefits of co-hosting

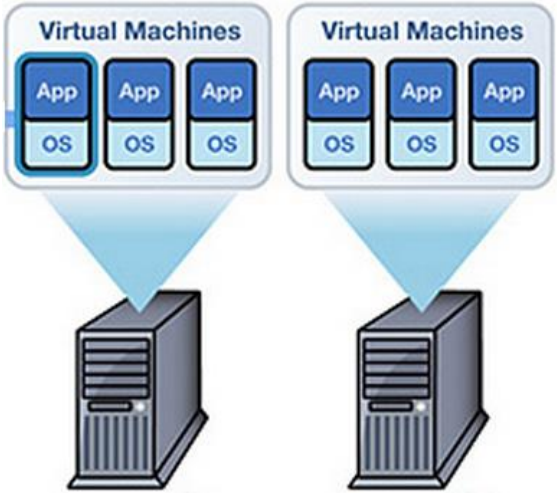
- **Solution**: Virtualize the data center

Virtualization

- The applications, instead of being hosted on the physical machines, can be hosted using virtual machines
- Resource allocation to the virtual machines can be made in two modes:

Conserving: A virtual machine demanding more system resources (CPU and memory) than the specified quota cannot be allocated the spare resources that are remain un-utilized by the other co-hosted virtual machines

Non-Conserving: The spare resources that are not utilized by the co-hosted virtual machines can be used by the virtual machine needing the extra amount of resource

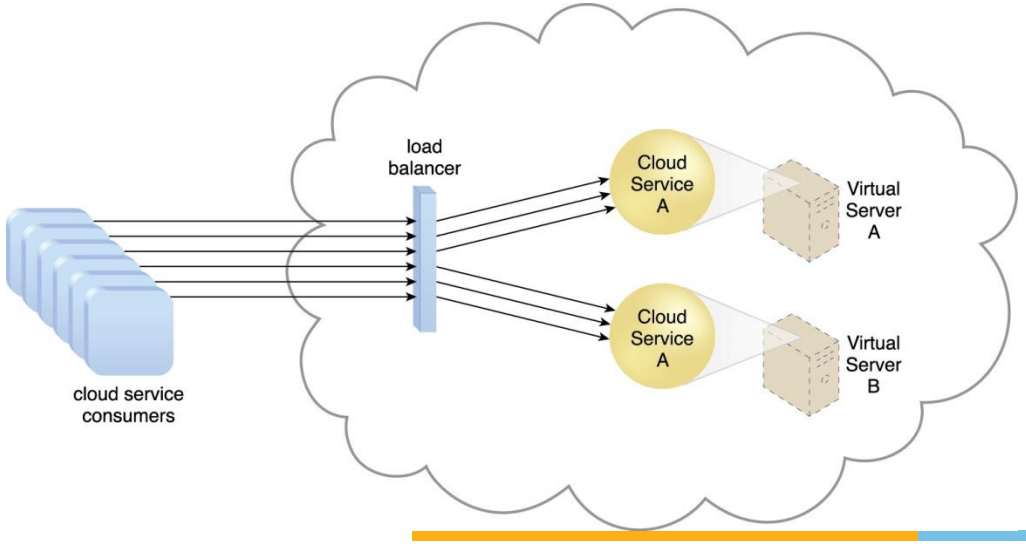


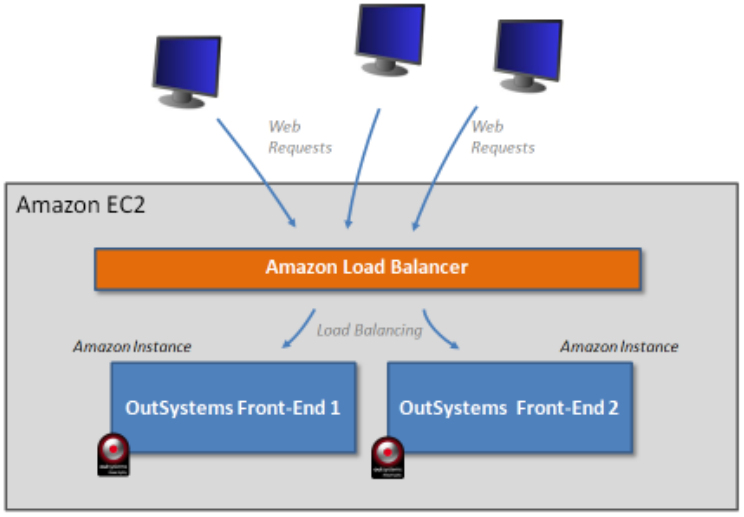
Virtualization (contd..)

- ASPs can allocate system resources more efficiently to these applications on-demand, so that application-level request rates and response times can be monitored and met effectively
- Different SLAs than Infrastructure SLAs are required. These SLAs are called application SLAs
- These service providers are known as Managed Service Providers (MSP) because the service providers are responsible for managing the application availability too
- To make IT infrastructure elastic, MSPs require in-depth understanding of the application's behavior
- Elasticity implies progressively scaling up the IT infrastructure to take the increasing load of an application

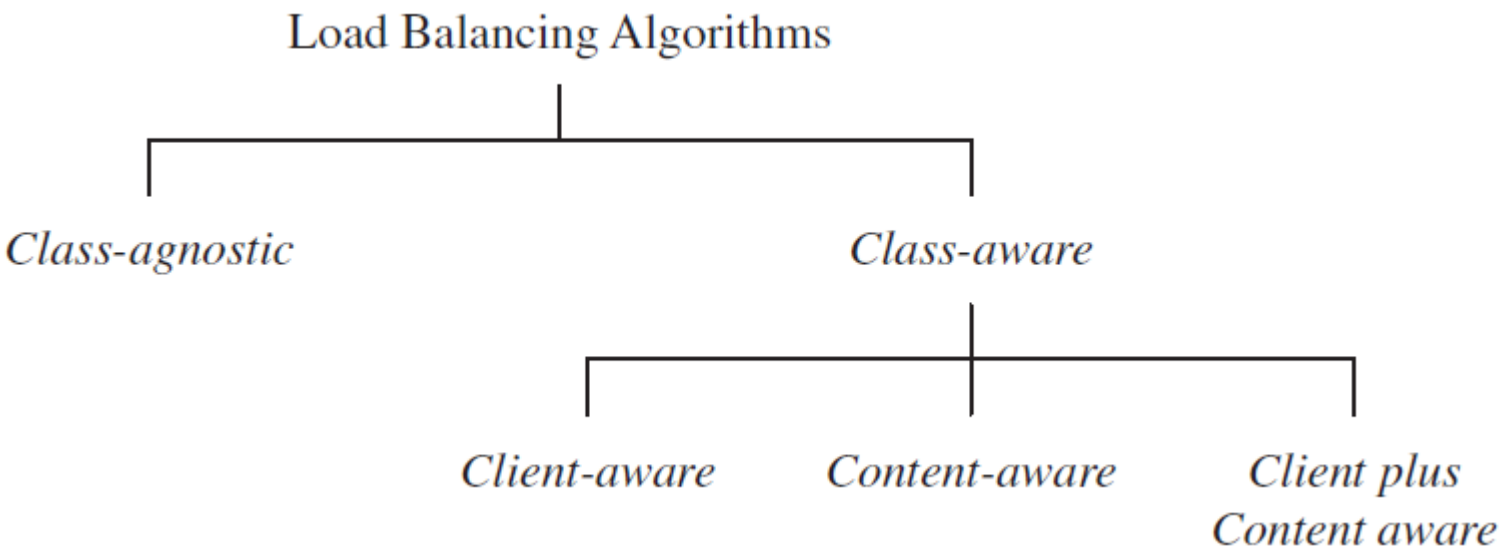
TRADITIONAL APPROACHES TO SLO MANAGEMENT

- To provide guaranteed quality of service (QoS) for hosted web applications following mechanisms are used:
  - Load balancing techniques
  - Admission control mechanism
- Load balancing is to distribute the incoming requests onto a set of physical machines, each hosting a replica of an application



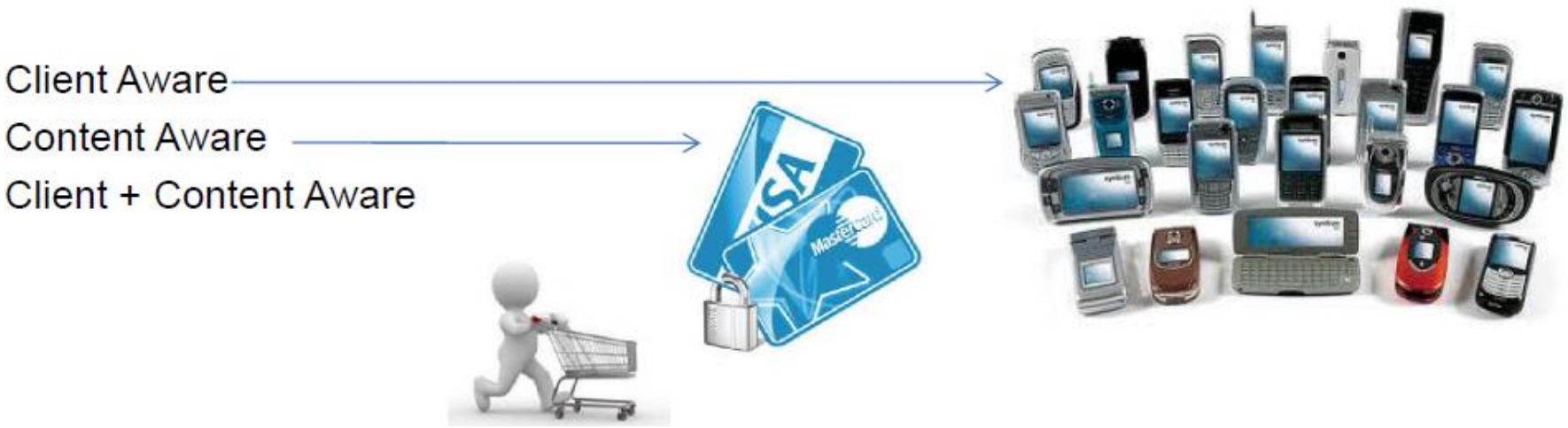


Load Balancing Algorithms



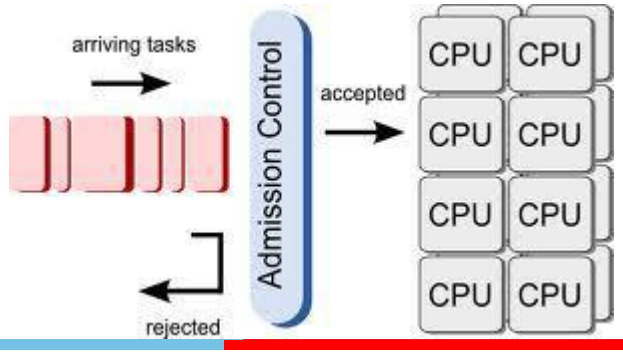
Load Balancing Algorithms (contd..)

- Class – agnostic
  - ✓ This means that the front-end node is neither aware of the type of client from which the request originates nor aware of the category (e.g., browsing, selling, payment, etc.) to which the request belongs to
- Class – aware
  - ✓ With class-aware load balancing and requests distribution, the front-end node must additionally inspect the type of client making the request and/or the type of service requested before deciding which back-end node should service the request

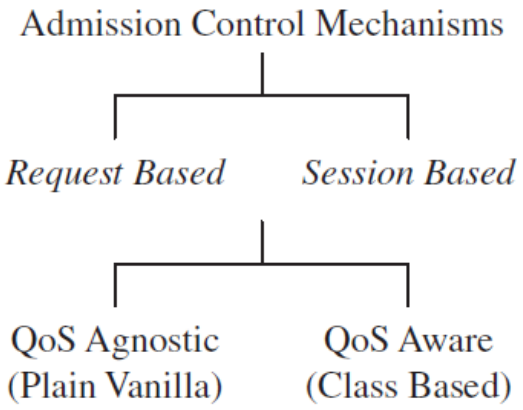


Admission Control

- Admission control algorithms play an important role in deciding the set of requests that should be admitted into the application server when the server experiences “very” heavy loads
- During overload situations, since the response time for all the requests would invariably degrade if all the arriving requests are admitted into the server, it would be preferable to be selective in identifying a subset of requests that should be admitted into the system so that the overall payoff is high
- The objective of admission control mechanisms, therefore, is to police the incoming requests and identify a subset of incoming requests that can be admitted into the system when the system faces overload situations



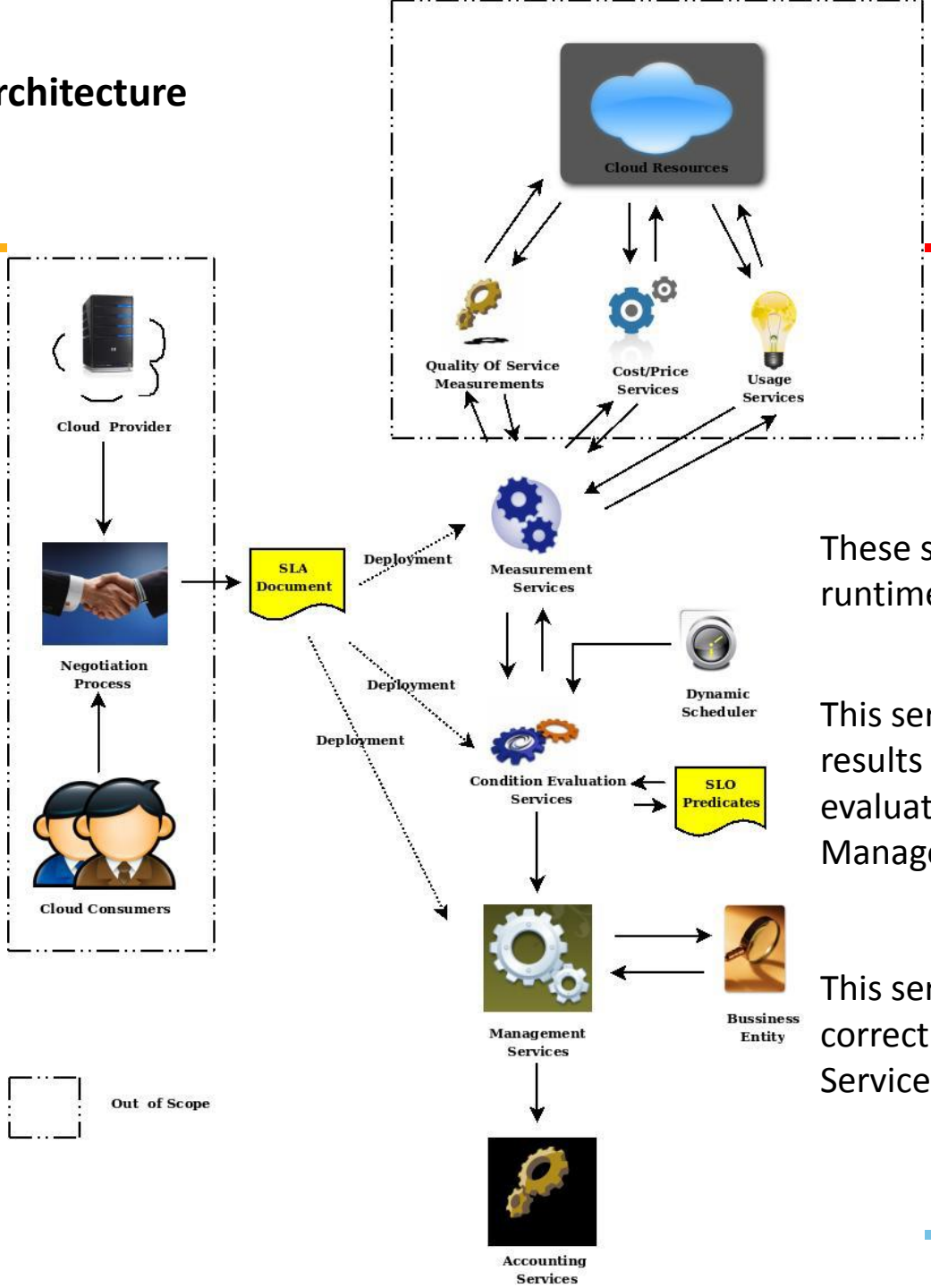
Admission Control Mechanisms



- **Request-based admission control algorithms** reject new requests if the servers are running to their capacity
  - ✓ Disadv: a client's session may consist of multiple requests that are not necessarily unrelated. Consequently, some requests are rejected even if there are others that are honored.
- **Session-based admission control mechanisms** try to ensure that longer sessions are completed and any new sessions are rejected

- **QoS Aware:** Requests from low priority users & requests that are likely to consume more system resources can be rejected

SLA architecture



These services are responsible for measuring the runtime parameters of cloud providers resources.

This service is responsible of getting the results from measurement services and evaluating the SLO's. If there are violations the Management service will be contacted.

This service is responsible for taking corrective actions on violation of the Service Level Objectives.

TYPES OF SLA

Key components of SLA:

- **Service Level Parameter**: Describes an observable property of a service whose value is measurable (reasonable, attainable, enforceable, measurable,)
- **Metrics:** These are definitions of values of service properties that are measured from a service providing system (server uptime is 98% for a period of 10 weeks)
- **Function:** A function specifies how to compute a metric's value from the values of other metrics and constants
- **Measurement directives:** These specify how to measure a metric

Infrastructure SLA: The infrastructure provider manages and offers guarantees on availability of the infrastructure, namely, server machine, power, network connectivity, etc.

Application SLA: In the application co-location hosting model, the server capacity is available to the applications based solely on their resource demands

Infrastructure SLA

innovate achieve lead

TABLE 16.2. Key Contractual Elements of an Infrastructural SLA

Hardware availability	• 99% uptime in a calendar month
Power availability	• 99.99% of the time in a calendar month
Data center network availability	• 99.99% of the time in a calendar month
Backbone network availability	• 99.999% of the time in a calendar month
Service credit for unavailability	• Refund of service credit prorated on downtime period
Outage notification guarantee	• Notification of customer within 1 hr of complete downtime
Internet latency guarantee	• When latency is measured at 5 min intervals to an upstream provider, the average doesn't exceed 60 msec
Packet loss guarantee	• Shall not exceed 1% in a calendar month

Application SLA

innovate achieve lead

TABLE 16.3. Key contractual components of an application SLA

Service level parameter metric	• Web site response time (e.g., max of 3.5 sec per user request)
	• Latency of web server (WS) (e.g., max of 0.2 sec per request)
	• Latency of DB (e.g., max of 0.5 sec per query)
Function	• Average latency of WS (latency of web server 1 + latency of web server 2) /2
	• Web site response time Average latency of web server + latency of database
Measurement directive	• DB latency available via http://mgmtserver/em/latency
	• WS latency available via http://mgmtserver/ws/instanceno/latency
Service level objective	• Service assurance
	• web site latency < 1 sec when concurrent connection < 1000
Penalty	• 1000 USD for every minute while the SLO was breached

LIFE CYCLE OF SLA

Five phases in SLA life cycle:

1. Contract definition
2. Publishing and discovery
3. Negotiation
4. Operationalization: SLA operation consists of
   - SLA monitoring
   - SLA accounting
   - SLA enforcement
5. De-commissioning

SLA MANAGEMENT IN CLOUD

SLA management of applications hosted on cloud platforms involves five phases:

1. Feasibility
2. On-boarding
3. Pre-production
4. Production
5. Termination

SLA MANAGEMENT IN CLOUD

Feasibility Analysis

- Technical feasibility
- Infrastructure feasibility
- Financial feasibility

On-Boarding of Application

- Moving an application to the MSP's hosting platform is called on-boarding

Preproduction

- The application is hosted in a simulated production environment

Production

- The application is made accessible to its end users under the agreed SLA

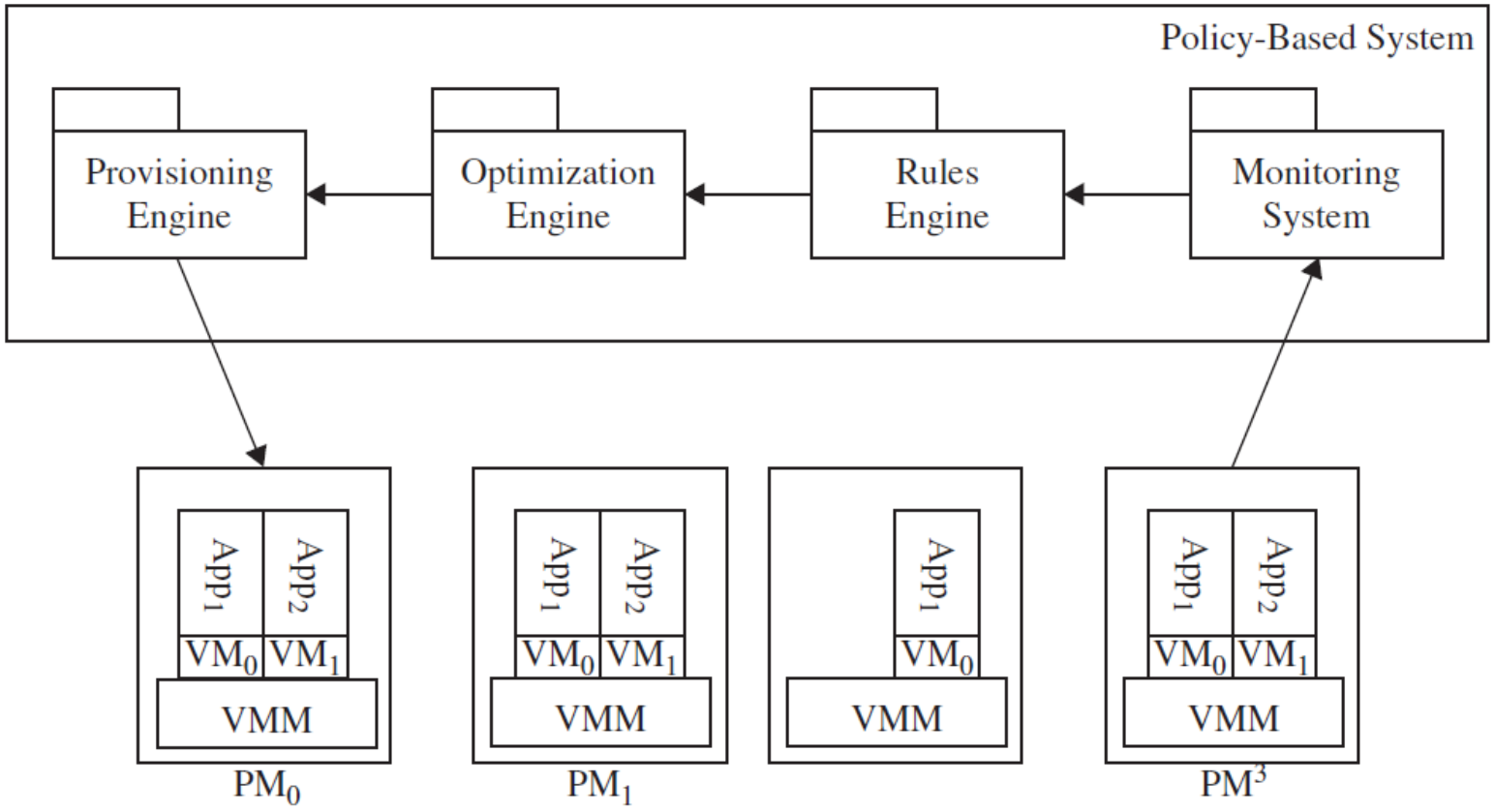
Termination

- When the customer wishes to withdraw the hosted application and does not wish to continue to avail the services of the MSP for managing the hosting of its application, the termination activity is initiated. On initiation of termination, all data related to the application are transferred to the customer and only the essential information is retained for legal compliance

AUTOMATED POLICY-BASED MANAGEMENT

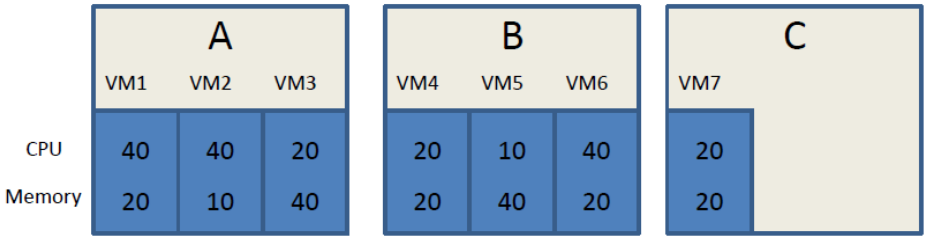
- The parameters often used to prioritize action and perform resource contention resolution are:
  - ✓ The SLA class (Platinum, Gold, Silver, etc.) to which the application belongs to.
  - ✓ The amount of penalty associated with SLA breach.
  - ✓ Whether the application is at the threshold of breaching the SLA.
  - ✓ Whether the application has already breached the SLA.
  - ✓ The number of applications belonging to the same customer that has breached SLA.
  - ✓ The number of applications belonging to the same customer about to breach SLA.
  - ✓ The type of action to be performed to rectify the situation

Policy based Management System



Policy based Management System (example)

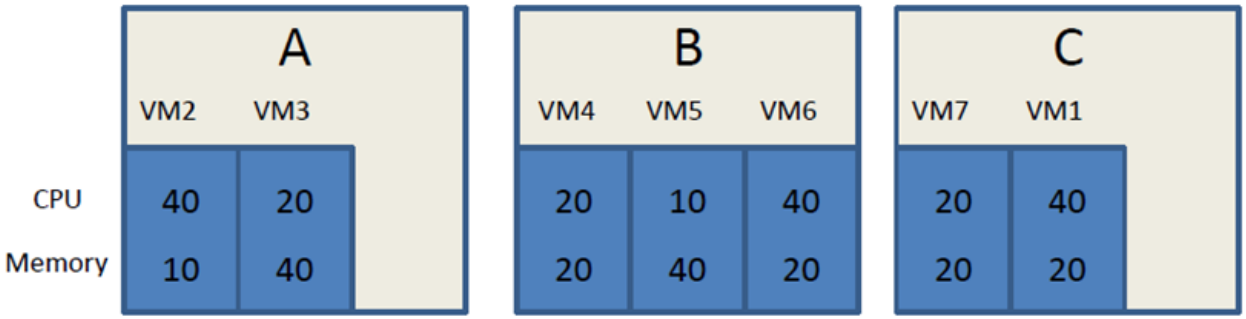
Consider a Policy Based Management system used by some cloud computing environment where 80% of the load is optimal for physical servers. There are 3 physical machines (or servers)–A, B and C with CPU and memory capacity of 100 units each. The data center in which A, B and C are hosted follows Green Computing methodology for conserving power resources, meaning until absolutely required physical machines are not turned on. Following figure shows the resource allocation to different virtual machines in the data center. Answer the following



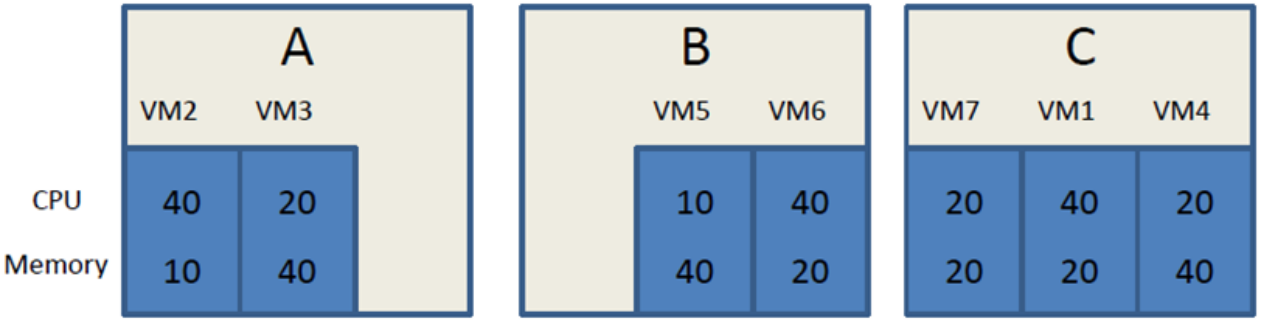
Resource allocation to VMs in a data center

1. Is the current load optimal for all the physical machines? If not, what can be done to enforce the optimal load policy?
2. How can you allocate 20 more units of memory to VM4 dynamically? Show the resource allocation diagram after allocation.
3. If the data center gets a request of provisioning a new virtual machine (VM8), with CPU and memory requirements of 70 units each, which physical machine will be used?

a. No



b.



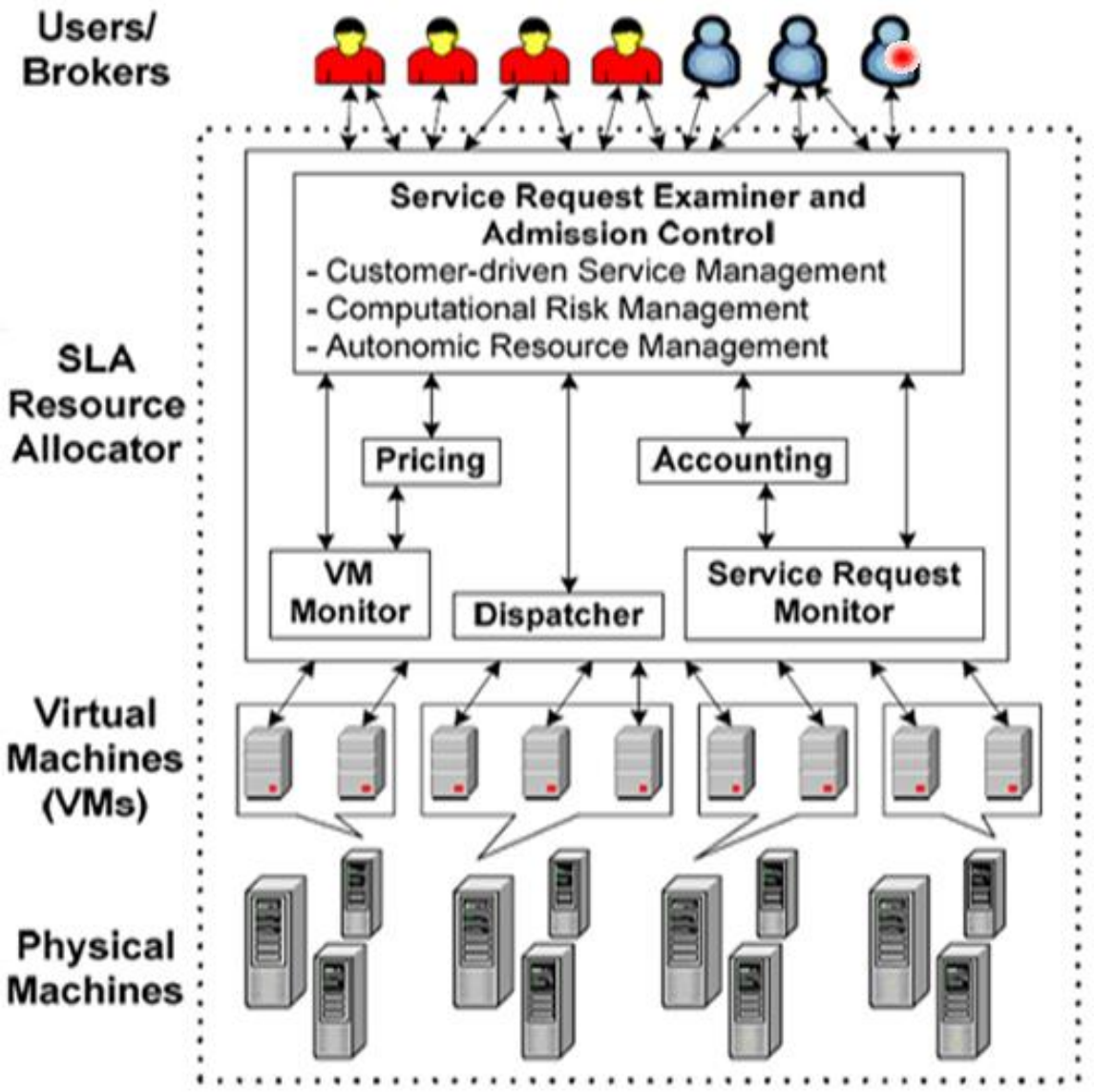
c.

None of the physical machines A, B, and C has the resources to provision new VM (VM8). So, a new physical machine, D, needs to be switched on for hosting it.

Managing Clouds: Services and Infrastructure

Introduction:

- The essential key requirement for cloud architecture is “efficient management” of resources at all the three layers of cloud stack



Introduction (contd..)

- The key questions to be addressed are:
  - How to monitor performance?
  - How to monitor health of resources?
  - How to perform fault diagnostics and recovery?
  - How to enforce SLAs during the operation of the resources?
- As cloud is a **distributed environment** with large scale of systems being managed, support of multi-tenancy, there is need for better management to maintain SLAs
- And there is need for automation to replace manual operations and to reduce overall cost

Managing IaaS

Two IaaS Systems, namely

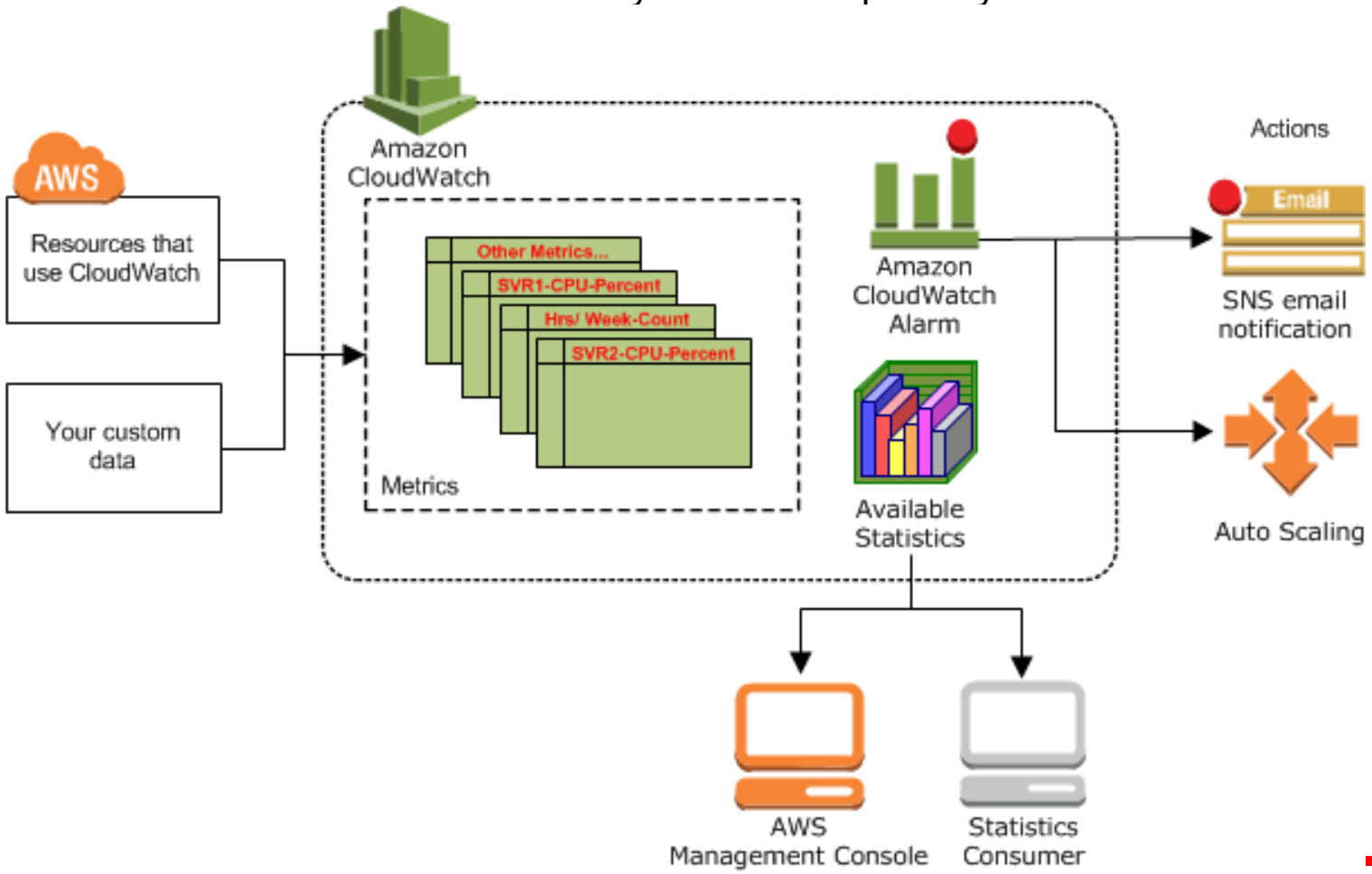
- HP CloudSystem Matrix
- Amazon EC2
- OpenStack

Managing IaaS - HP Cloud System Matrix

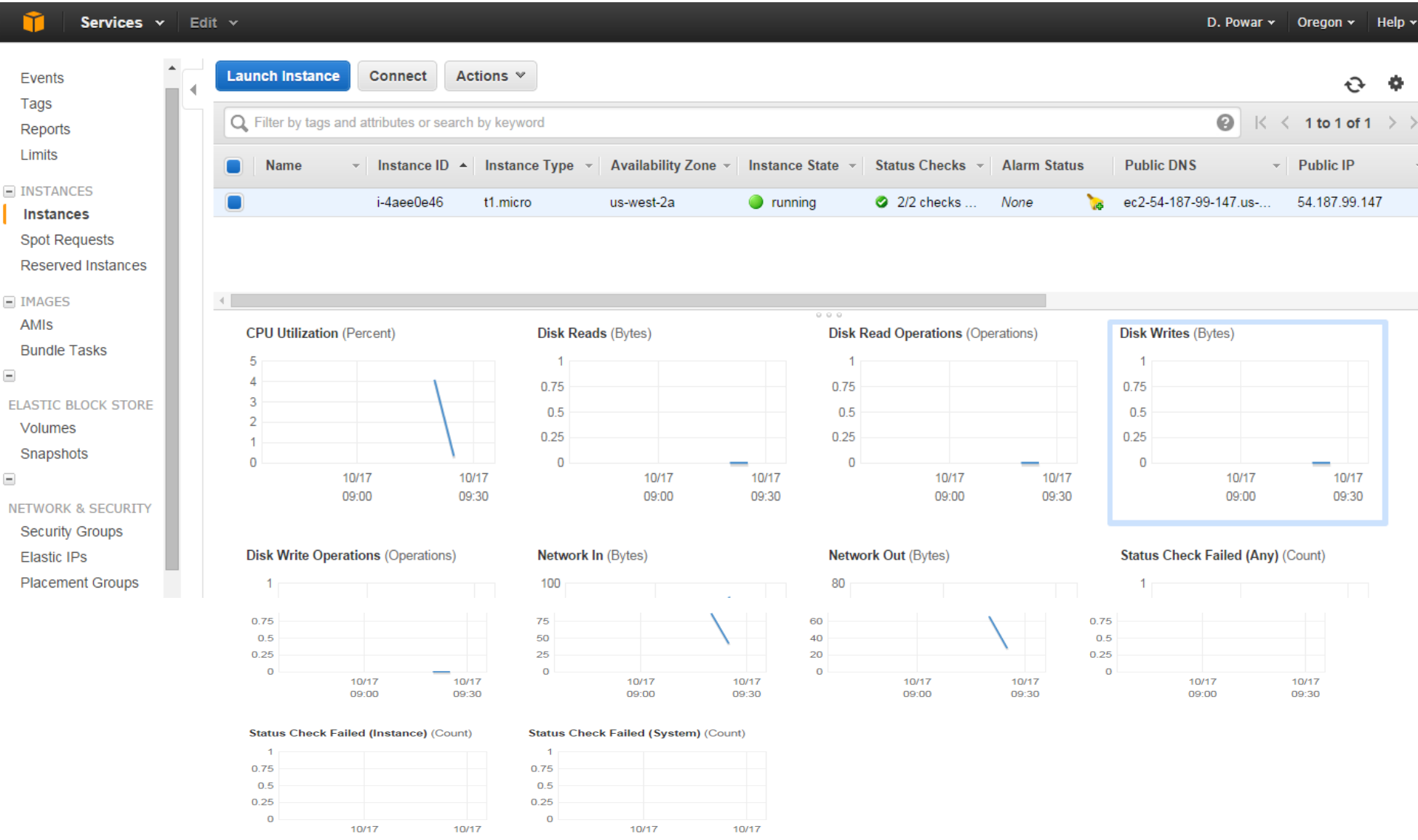
- It provides a self-service interface that consumers and administrators can use to perform operations.
- These could be simple activities such as
  - re-boot instances / virtual machines
  - getting console access to the environment
  - adjusting the resources assigned to the service
  - expanding resources to meeting demand growth
  - reducing resources during low utilization periods
- These and other such operations are available from the CloudSystem Matrix administrator portal as well as via WebService APIs

Managing IaaS - Amazon EC2

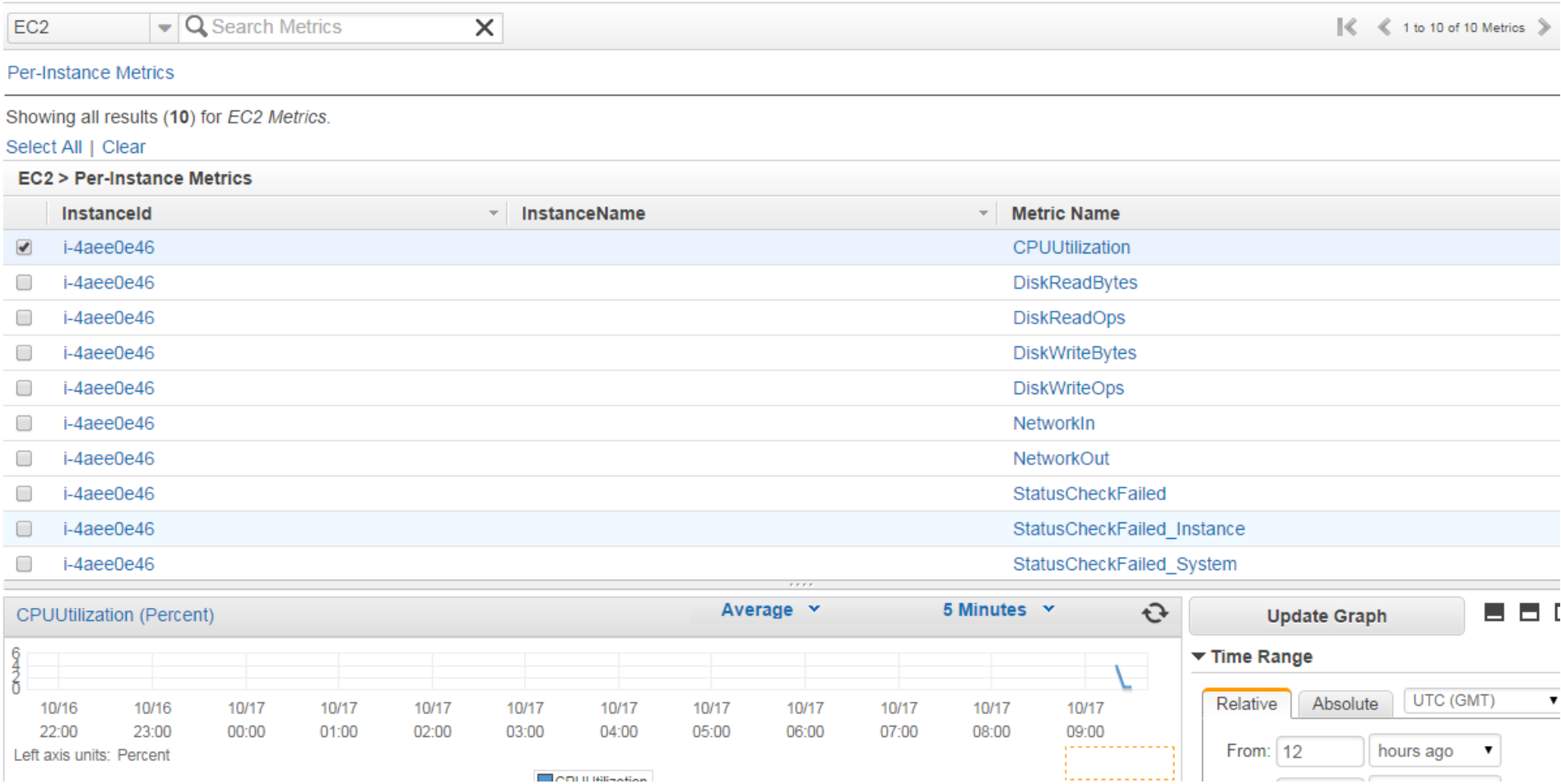
- Amazon CloudWatch Architecture
  - ✓ Amazon CloudWatch is basically a metrics repository



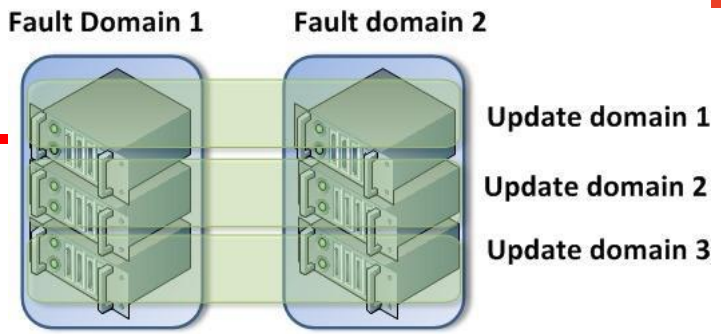
Monitoring


Services
Edit
D. Powar
Oregon
Help
Events
Tags
Reports
Limits
INSTANCES
Instances
Spot Requests
Reserved Instances
IMAGES
AMIs
Bundle Tasks
ELASTIC BLOCK STORE
Volumes
Snapshots
NETWORK & SECURITY
Security Groups
Elastic IPs
Placement Groups
Launch Instance
Connect
Actions
Filter by tags and attributes or search by keyword
1 to 1 of 1
Name
Instance ID
Instance Type
Availability Zone
Instance State
Status Checks
Alarm Status
Public DNS
Public IP
i-4aee0e46
t1.micro
us-west-2a
running
2/2 checks ...
None
ec2-54-187-99-147.us-...
54.187.99.147
CPU Utilization (Percent)
Disk Reads (Bytes)
Disk Read Operations (Operations)
Disk Writes (Bytes)
Disk Write Operations (Operations)
Network In (Bytes)
Network Out (Bytes)
Status Check Failed (Any) (Count)
Status Check Failed (Instance) (Count)
Status Check Failed (System) (Count)
10/17 09:00
10/17 09:30

CloudWatch


EC2
Search Metrics
1 to 10 of 10 Metrics
Per-Instance Metrics
Showing all results (10) for EC2 Metrics.
Select All | Clear
EC2 > Per-Instance Metrics
InstanceId
InstanceName
Metric Name
i-4aee0e46 CPUUtilization
i-4aee0e46 DiskReadBytes
i-4aee0e46 DiskReadOps
i-4aee0e46 DiskWriteBytes
i-4aee0e46 DiskWriteOps
i-4aee0e46 NetworkIn
i-4aee0e46 NetworkOut
i-4aee0e46 StatusCheckFailed
i-4aee0e46 StatusCheckFailed_Instance
i-4aee0e46 StatusCheckFailed_System
CPUUtilization (Percent)
Average
5 Minutes
Update Graph
Time Range
Relative
Absolute
UTC (GMT)
From: 12
hours ago
10/16 22:00
10/16 23:00
10/17 00:00
10/17 01:00
10/17 02:00
10/17 03:00
10/17 04:00
10/17 05:00
10/17 06:00
10/17 07:00
10/17 08:00
10/17 09:00
Left axis units: Percent

Managing PaaS

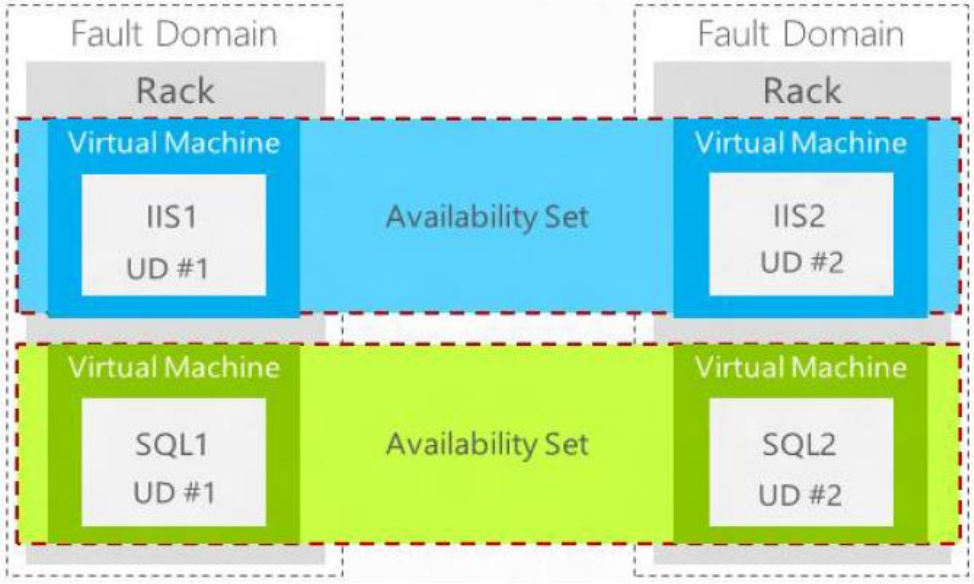


Managing Applications in Azure

- Availability and Fault tolerance
  - Upgrade Domain: an upgrade domain is a strategy to ensure an application stays up and running, while undergoing an update of the application.
  - Upgrading a deployment, is carried out one upgrade domain at a time.
- The steps are:
  1. stopping the instances running in the first upgrade domain
  2. upgrading the application,
  3. bringing the instances back online followed by repeating the steps in the next upgrade domain
- An upgrade is completed when all upgrade domains are processed
- By stopping only the instances running within one upgrade domain, Windows Azure ensures that an upgrade takes place with the least possible impact to the running service.

Fault Domain:

- The purpose of identifying/organizing fault domains is to prevent a single point of failure.
- To provide redundancy, configure more than one virtual machine to perform the same function or role
- Ex: a computer by itself connected to a power outlet is a fault domain



Managing SaaS

Monitoring Force.com using Netcharts:

- Provides dashboard to view key performance indicators
- Dashboard provides data analytics, help users to make optimal decisions and increase operational efficiency

Monitoring Force.com using Nimsoft:

- Monitoring solution that apply to IaaS, PaaS, SaaS
- Metrics monitored include
  - Average Transaction Time
  - No. of Transactions
  - Instance Status
  - File and data storage
  - Login/Logout time logs
  - API calls
  - Query execution time


NetCharts Performance Dashboards
Cost Center IS
New
Edit
Center for Information & Standards
NETCHARTS
PERFORMANCE
DASHBOARDS
Top 5 Expenses by Project Groups (as of April 30, 2009)
3%
11%
7%
16%
6%
56%
300-Inform America Magazine
305-Publications
320-Publications - Order Processing
330-Marketing & Communications
360-Sales & Marketing - Indirect Costs
Other
Variance to Budget by Project Groups (as of April 30, 2009)
360-Sales & Marketing - Indirect Costs
335-Trade Show/Exhibits
332-Partner Program
330-Marketing & Communications
320-Publications - Order Processing
315-Publications - Promotion
310-Third Party Resales
305-Publications
300-Inform America Magazine
NetCharts Performance Dashboards
Membership Statistics
Membership Rollforward - YTD
Membership Rollforward - Monthly
Three Year Trend (as of April 30, 2009)
Monthly Activity (as of April 30, 2009)
Year to Date Net Income
TOTAL ASSETS
LIABILITIES
TOTAL LIABILITIES
NET ASSETS
TOTAL NET ASSETS
TOTAL LIABILITIES & NET ASSETS
Comments
Enter comments as necessary

Other Cloud-Scale Management Systems

HP Cloud Assure

HP SaaS solution for assessing

- Security
- Performance
- Availability of cloud services

RightScale:

- Cloud Management Environment – Providers management dashboard
- Cloud-Ready Server Templates – Provides pre-packaged solutions based on best practices for common application scenarios
- Adaptable Automation Engine – Resource allocation as required by system demand
- Multi-Cloud Engine – Cloud infrastructure APIs for multiple cloud provider

Compuware:

- Cloud monitoring suite that directly measure performance by end users and customers

Summary

- **Service License Agreements:** Lifecycle and Management
  - INSPIRATION
  - TRADITIONAL APPROACHES TO SLO MANAGEMENT
  - TYPES OF SLA's
  - LIFE CYCLE OF SLA
  - SLA MANAGEMENT IN CLOUD
  - AUTOMATED POLICY-BASED MANAGEMENT
- Managing Clouds: Services and Infrastructure
  - Managing IaaS
  - Managing PaaS
  - Managing SaaS