Welcome to:

Service Management & System Administration in Cloud

Unit objectives

After completing this unit, you should be able to:

- Understand the concept of IT service management
- Learn the characteristics of Cloud Service Management
- Learn cloud architecture & workflow in the Cloud
- Learn cloud provisioning
- Learn monitoring the usage of the Cloud & Metering and Billing the cloud Usage
- Learn system administration and tasks included in the system administration
- Learn System Health check, Patching and Updates
- Learn maintenance outages
- Learn sending notifications
- Learn and maintain the service catalog & system troubleshooting

Concept of service management

- Management and implementation of quality information technology services

- Process-focused

- Prototype shift from managing IT

- Associated with the field of management information systems

- The use of the term "service management" is interpreted by many in the world as ITSM

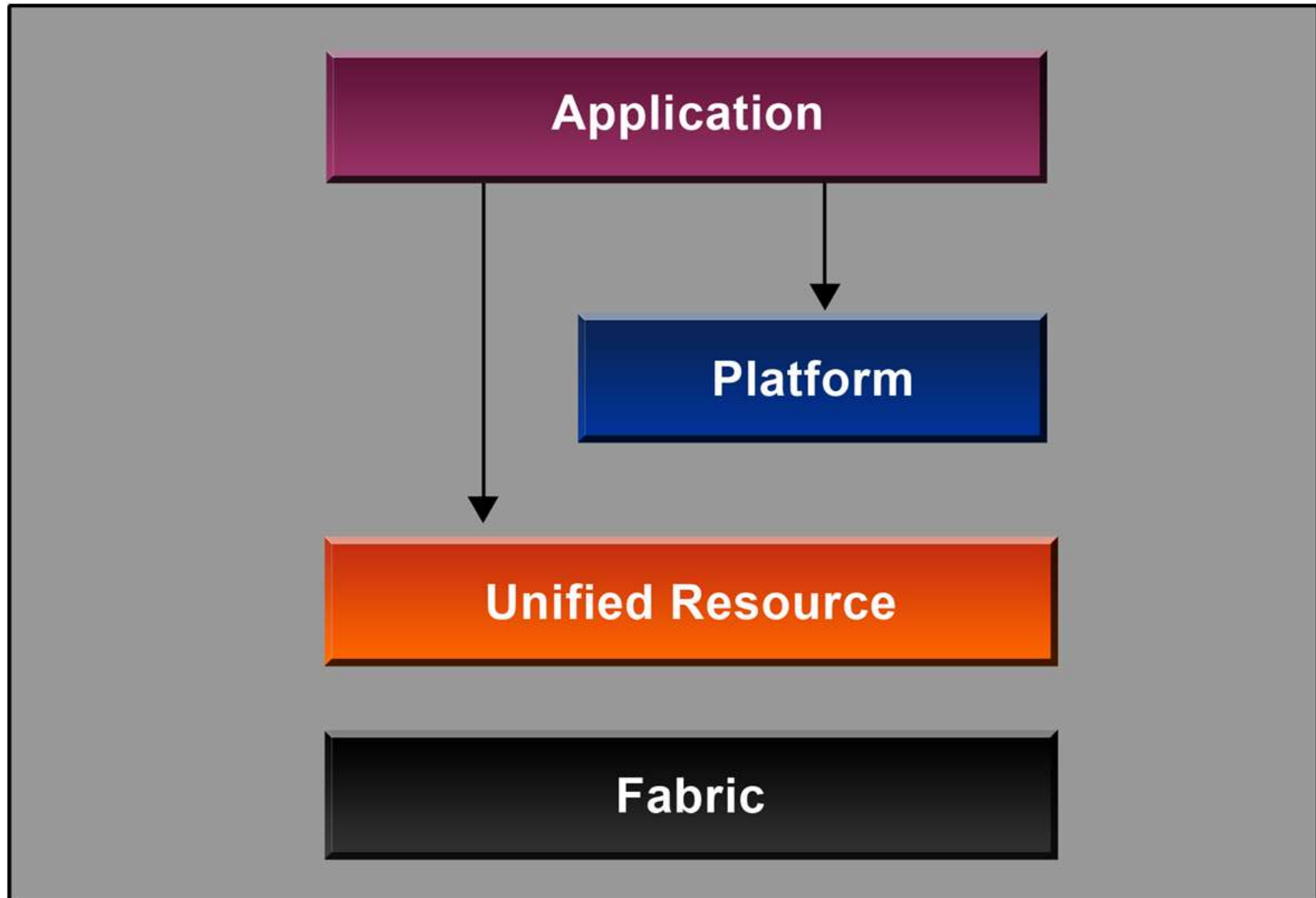# Characteristics of Cloud Service Management

Cloud Service Management

- Cloud service delivery and cloud monitoring
  - To confirm optimal performance, stability and productivity in virtualized, on-demand environments

- What are the key processes associated with cloud service management?
  - Cloud service-level agreement (SLA) management
  - Organize resources for fast provisioning

- Cloud service management platforms

- Cloud service management is key to survival, though it has challenges
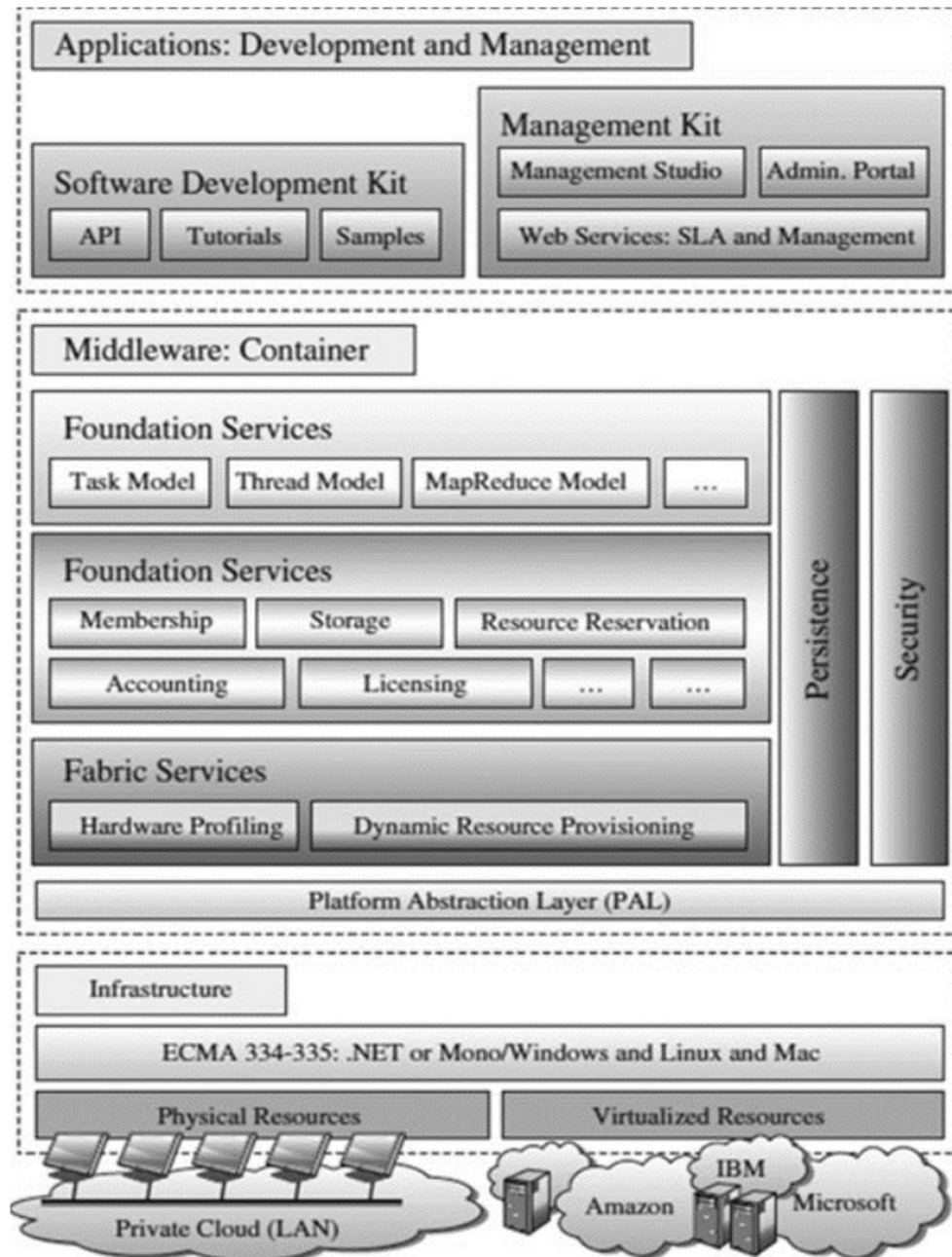
Cloud Architecture

Aneka Cloud Architecture
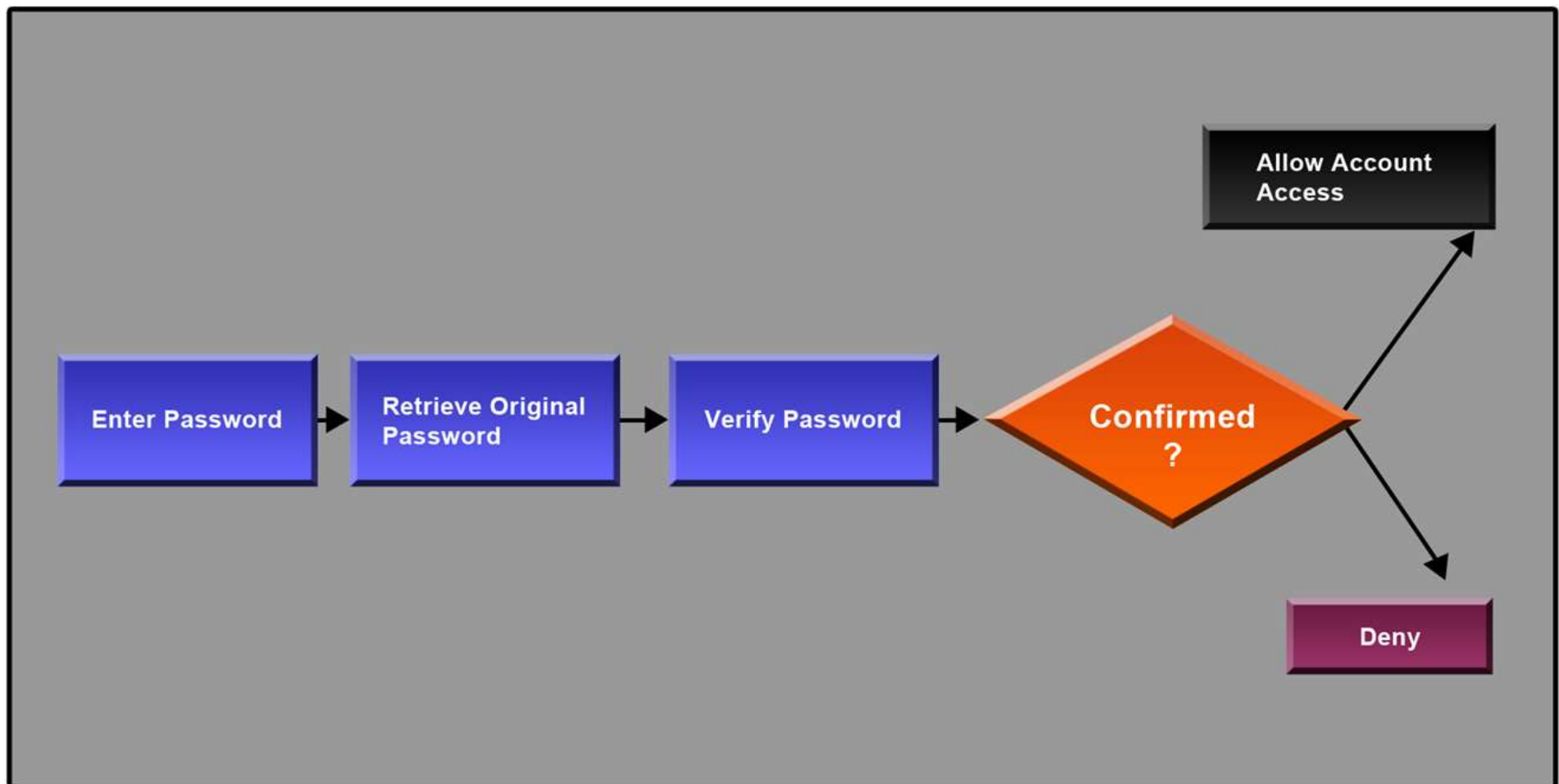
Characteristics of Cloud

Workflows in Cloud

- Workflows can be implemented with less complexity in cloud

- Scalability permits real-time provisioning in comparison with grid

- Workflows are denoted by a Directed Acyclic Graph (DAG)

Microsoft is using the WINSAT (Window System Assessment Tool)
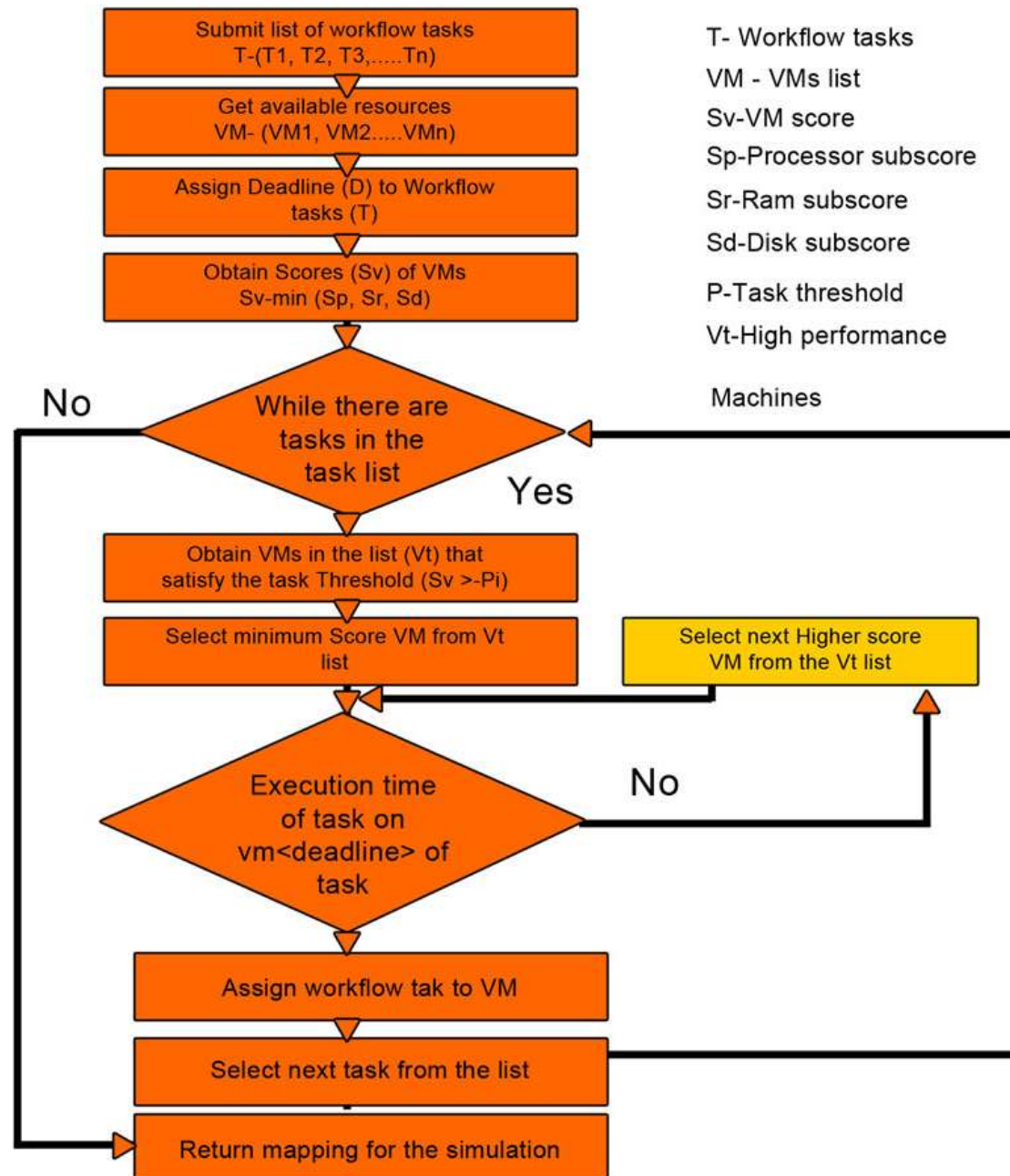
Workflows in Cloud

- Simplified account verification workflow process

Algorithm Description

Submit list of workflow tasks
T-(T1, T2, T3,.....Tn)

Get available resources
VM- (VM1, VM2.....VMn)

Assign Deadline (D) to Workflow tasks (T)

Obtain Scores (Sv) of VMs
Sv-min (Sp, Sr, Sd)

While there are tasks in the task list

Obtain VMs in the list (Vt) that satisfy the task Threshold (Sv >-Pi)

Select minimum Score VM from Vt list

Select next Higher score VM from the Vt list

Execution time of task on vm<deadline> of task

Assign workflow tak to VM

Select next task from the list

Return mapping for the simulation

No

Yes

No

T- Workflow tasks

VM - VMs list

Sv-VM score

Sp-Processor subscore

Sr-Ram subscore

Sd-Disk subscore

P-Task threshold

Vt-High performance

Machines

Calculating Score of Machine

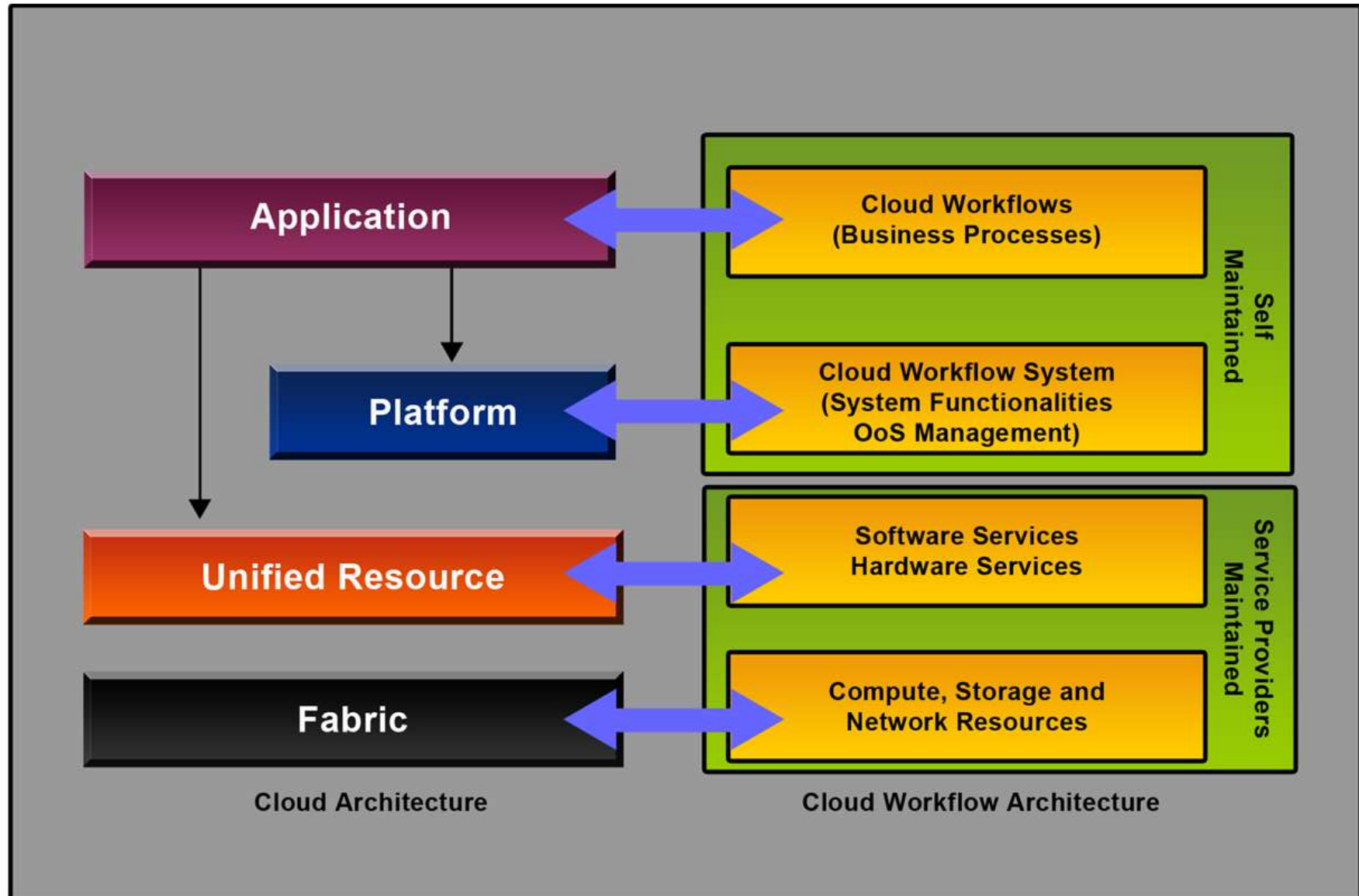| Component	What is rated	Sub score	Final score
Processor	Calculation per second(Mips)	6	4 (Determined by lowest sub score)
Ram (memory)	Memory operation per second	5	
Storage space	Disk data transfer rate	4	

Deadline Distribution
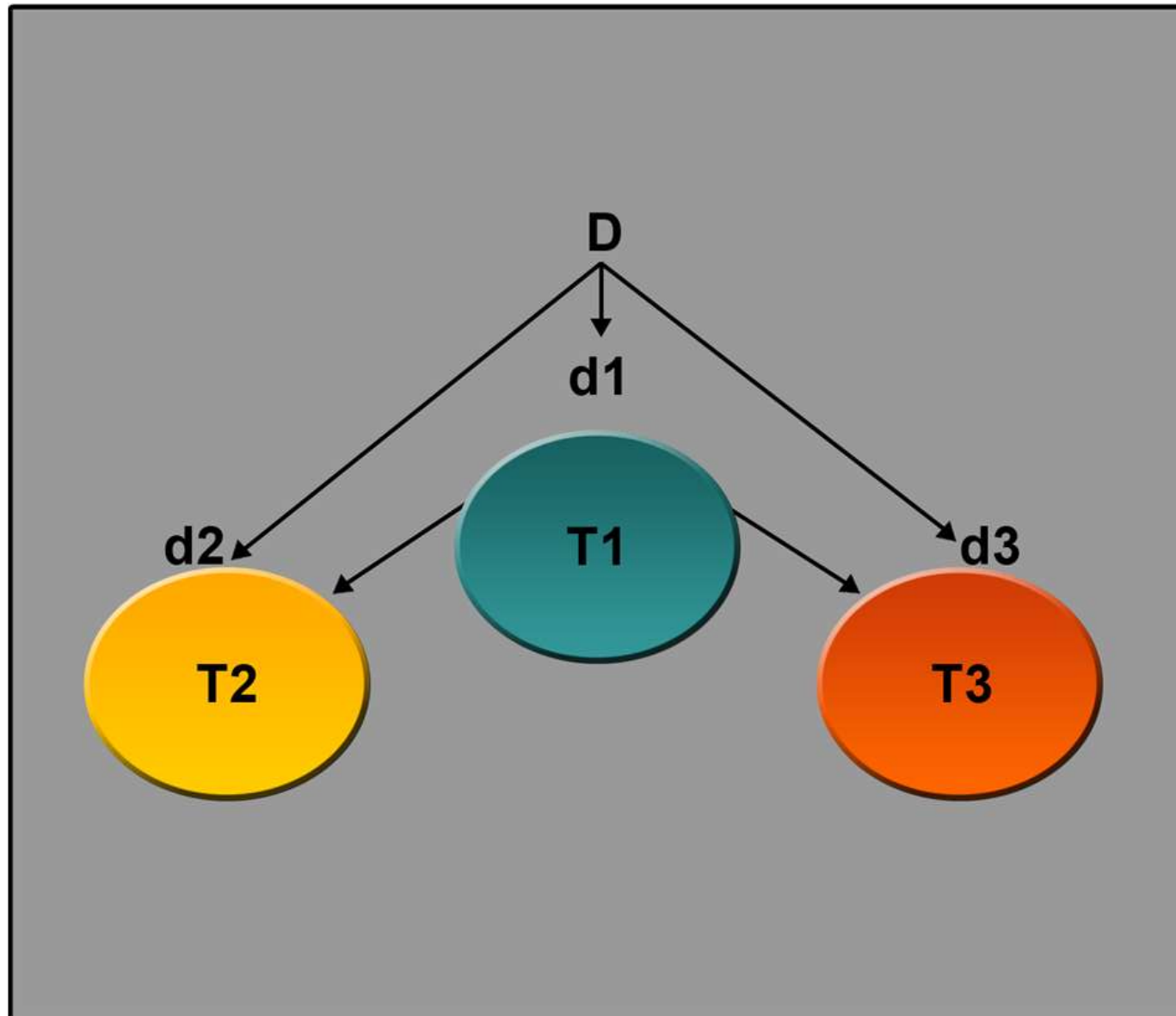
General Architecture of Cloud Workflow Systems

Cloud Provisioning

- Cloud provisioning is user account provisioning in a cloud environment

- Self-provisioning - (also referred as cloud self-service), the customer buys resources from the cloud provider over a web form, generating a customer account and paying for resources with a credit card

- Dynamic provisioning - the provider assigns more resources as they are needed and eliminates them when they are not

- Advance provisioning - the client deals with the provider for services and the provider organizes the appropriate resources in advance, before starting service Term provisioning just means "to provide"

Integrate and Automate

- Deploying applications requires that you configure and deploy a variety of application delivery services before going live

- Streamline utilization through automation with common frameworks such as Puppet, Chef, and Open Stack, and with enterprise-class solutions from Microsoft, VMware, and IBM

- Permit self-service deployment with repeatable application delivery facilities

Enable elastic applications

- Distribute demand and capacity monitoring data between solutions and provisioning management systems from HP, Microsoft, IBM, and VMware

- Respond to scaling and security events dynamically

Maximize Efficiency

- Designing application delivery services manually can take days

- Confirm fast and repeatable application deployments

- Confirm that processes can effortlessly span environments and systems

- Streamline application-specific service configuration

- Cloud provisioning also involves developing the methods for interfacing with the cloud's applications and services

Four dimensions of cloud provisioning (1 of 4)

- Payment is not only for the tool but the space it uses

- Time: The memory essential for each image to be accommodated within the cloud data center, regularly billed by how much time the memory is dynamically accessible (what GoGrid calls the "gigabyte-hour")

- Space: The bandwidth used by the image's Internet connection during its normal progress of operation, payable by the megabyte

- Breadth: The storage essential to host the database used by the software and middleware these images will run

- Depth: The software that makes the image useful, usually comprising the operating system for which Windows Server is seldom, if ever, the least expensive choice

Four dimensions of cloud provisioning (2 of 4)

- "Federated" provisioning

- Advance provisioning -- like classic "on-boarding" provisioning where the provisioning is done before the user knows it

- Just-in-time provisioning -- unlike "on-boarding," this is known as do-it-yourself provisioning, accomplished when the user first contacts the application or service. It can be attribute-based, role-based or hinge on a number of different triggers, which govern if that specific user can gain contact to that service

- "Cloud Storage Parameters

- Cloud storage frequently uses centralized services, so some interpret storage as a potential goal for offenders or hackers. This has constantly been the instance for any valuable resource. The user can alleviate this by applying the suitable security controls

- Multitenancy offers concerns, with the potential for data-isolation tools that may either fail in process or in a rollback action from a backup system

- Storage systems comprise of complex hardware and software applications. There's always the potential for disastrous failure modes that might destroy the data or expose the data of one client to another
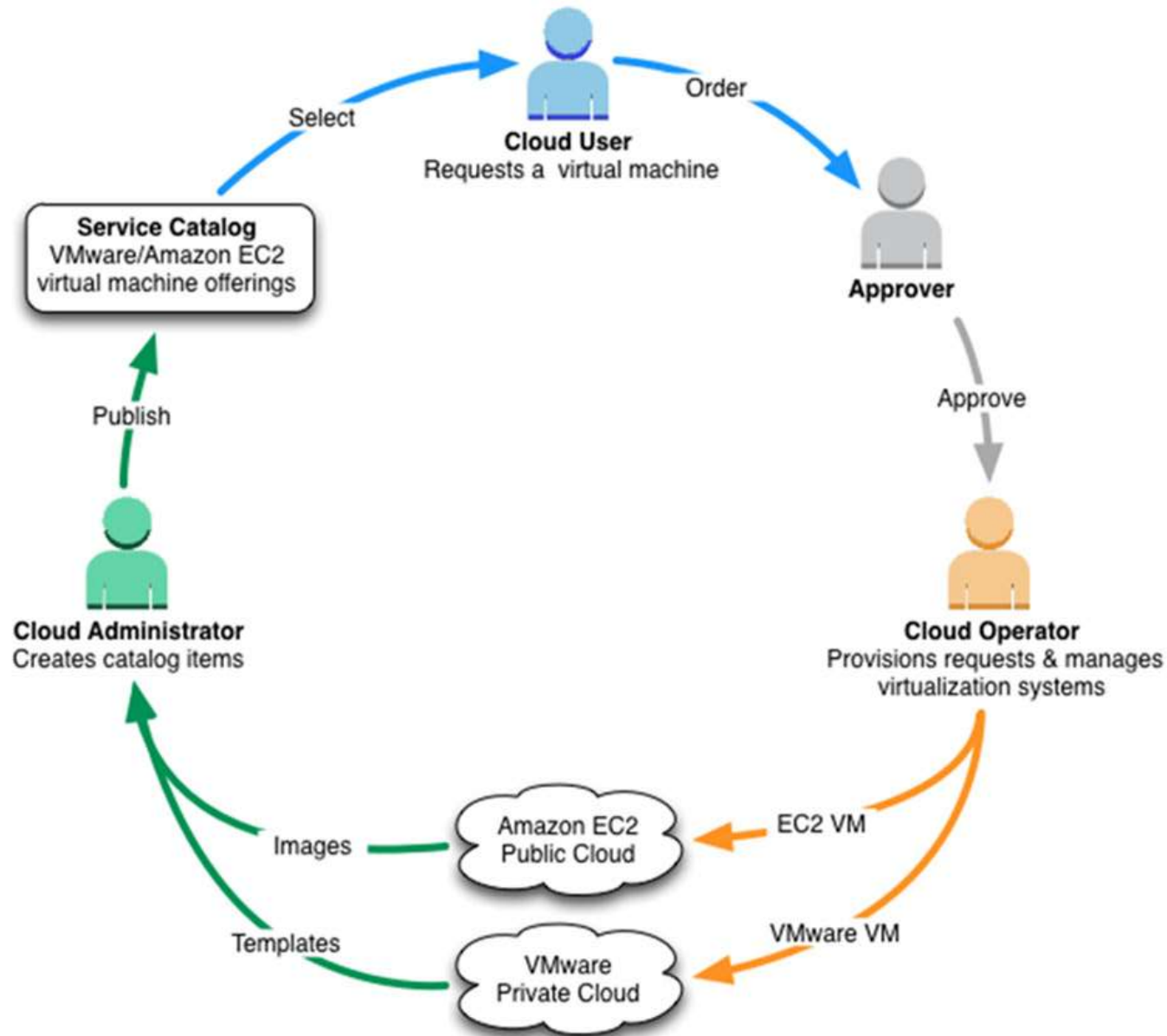
- Encryption has several other uses in the cloud environment as well, comprising:

    - Controlling contact to the control interfaces for resources

    - Directing access for administrators to VMs and OS images

    - Governing access to applications

How cloud provisioning works (1 of 2)

How cloud provisioning works (2 of 2)

- Virtual Provisioning Cloud Administrator
- ✓ Members of this group possess the cloud provisioning environment and are in authority for configuring the different virtualization providers used by cloud provisioning

- Virtual Provisioning Cloud Operator
- ✓ Members of this group satisfy provisioning requests from customers

- Virtual Provisioning Cloud User
- ✓ The group members can request virtual machines from the service and use the My Virtual Assets portal to manage any virtual machines that are allocated to them

Table . User groups, roles and rights

Group	User Roles	Privileges
Virtual Provisioning Cloud Users	cloud_user	Demand virtual machines from the service catalog and use the My Virtual Assets portal to manage any virtual machines that are assigned to them.
Virtual Provisioning Cloud Operators	cloud_operator	Satisfy provisioning requests from users by completing tasks that appear on the Cloud Operations Portal. Cloud operators are allocated to specific virtualization providers and must be technically skillful with the products they upkeep. This group also includes all members of the child groups EC2 Operators and VMWare Operators.
Virtual Provisioning Cloud Administrators	cloud_admin	Own the cloud provisioning environment and are responsible for configuring the different virtualization products used by cloud provisioning. The cloud provisioning environment can be monitored by cloud administrators using the Service Monitoring Portal.

Cloud provisioning security (2 of 2)

Contd- Table . User groups, roles and rights

Group	User Roles	Privileges
EC2 Approvers	itil	Accept or reject requests for Amazon EC2 virtual machine resources, which includes requests for new virtual machines, lease extensions, and state changes to existing virtual machines. Approvers have no technical responsibilities.
EC2 Operators	ec2_operator	Satisfy Amazon EC2 provisioning requests from users by completing jobs that appear on the Cloud Operations Portal. Customers in the EC2 Operators group are members of Virtual Provisioning Cloud Operators parent group.
VMware Approvers	itil	Accept or reject requests for VMware virtual machine resources, which includes requests for new virtual machines, lease extensions, and modifications to existing virtual machines. Approvers have no technical responsibilities.
VMware Operators	vmware_operator	Satisfy VMware provisioning requests from users by completing tasks that appear on the Cloud Operations Portal. VMware Operators users are members of the Virtual Provisioning Cloud Operators parent group.

Specific provisioning actions

- Modify VM (VMware only)

- Update Lease End

- Pause VM

- Stop VM

- Delete Snapshot

- Start VM

- Cancel

- Terminate

- Take Snapshot

- Restore from Snapshot

Self service provisioning for cloud computing services

- On-demand self-service is defined by the NIST as a vital characteristic of Cloud computing.

- The self-service feature of cloud computing permits end users gain and eliminate cloud services including applications, configuration, and the infrastructure supporting the applications, themselves without needing the assistance of an IT associate.

- The automatic self-servicing may aim different application goals and constraints (e.g. deadlines and cost), as well as managing different application designs (e.g., workflows and bags-of-tasks).

- Cloud users can gain cloud services through a self-service portal or a cloud service catalog.

- Since business users can achieve and design cloud services themselves, this means IT staff can be more productive and gives them more time to accomplish cloud infrastructures.

- However, one problem of cloud service provisioning is that it is not instantaneous.

# Cloud-based Infrastructure Services Provisioning

- Cloud Computing technologies are developing as a common and constant way of provisioning infrastructure services on-demand that may include resources from multiple providers and multiple domains, comprising integration with the legacy services and setups. Like this, clouds represent a new step in evolutional computing and communication technologies development chain by introducing a new type of services and a new abstraction layer for the general infrastructure services mobility and virtualization (similar to utilities).

- The projected InterCloud Architecture describes four complimentary components addressing inter-cloud integration and interoperability:

- ✓ Multi-layer Cloud Services Model (CSM) that associates commonly adopted cloud service models, such as SaaS, PaaS, IaaS, in one multilayer model with corresponding inter-layer edges;
- ✓ InterCloud Control and Management Plane (ICCMP) that supports cloud based applications interaction;
- ✓ InterCloud Federation Framework (ICFF);
- ✓ InterCloud Operation Framework (ICOF).

Cloud Usage Monitor

- The cloud usage monitor tool is a lightweight and independent processing module responsible for gathering and processing IT resource usage data.

- Based on the type of usage metrics they are intended to collect and the manner in which usage data needs to be collected, cloud usage monitors can be present in different setups.

- The forthcoming sections describe three common agent-based implementation setups for the cloud usage monitor tool.

- All can be designed to forward collected usage data to a log database for reporting and post-processing purposes.

Monitoring Agent

- Figure - A cloud service consumer sends a request message to a cloud service (1). The monitoring agent intercepts the message to collect relevant usage data (2) before allowing it to continue to the cloud service (3a). The monitoring agent stores the collected usage data in the log database that is stored outside of the cloud (3b). The cloud service replies with response messages (4) that is sent back to the cloud service consumer without being intercepted by the monitoring agent (5).
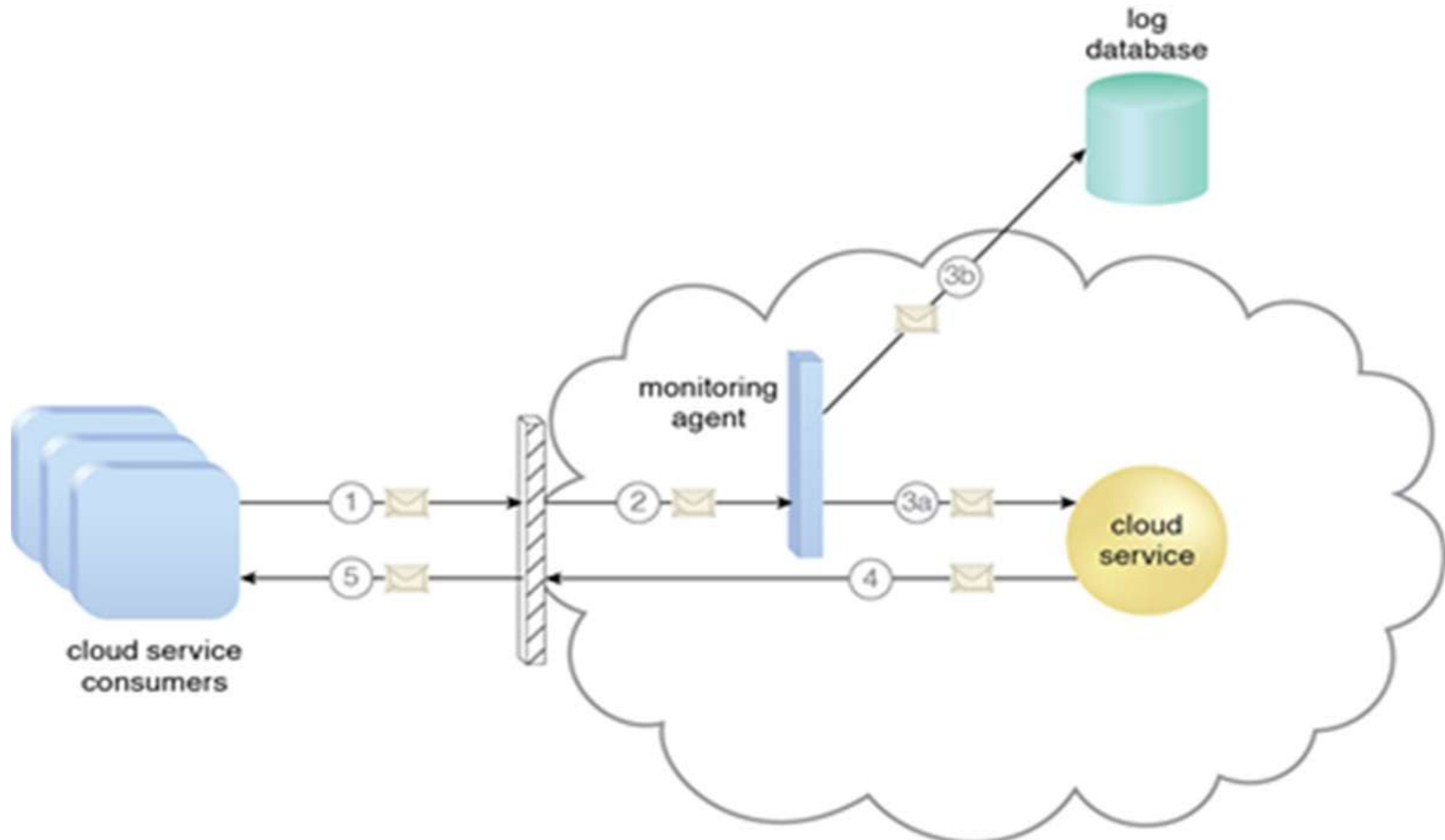
Resource Agent

- Figure - The resource agent residing within the cloud detects an increase in usage between a cloud service consumer and a virtual server (1).

- The resource agent receives a notification from the underlying resource management program that the virtual server is scaled up (2) and stores the collected usage data in the log database, as per its monitoring metrics (3).
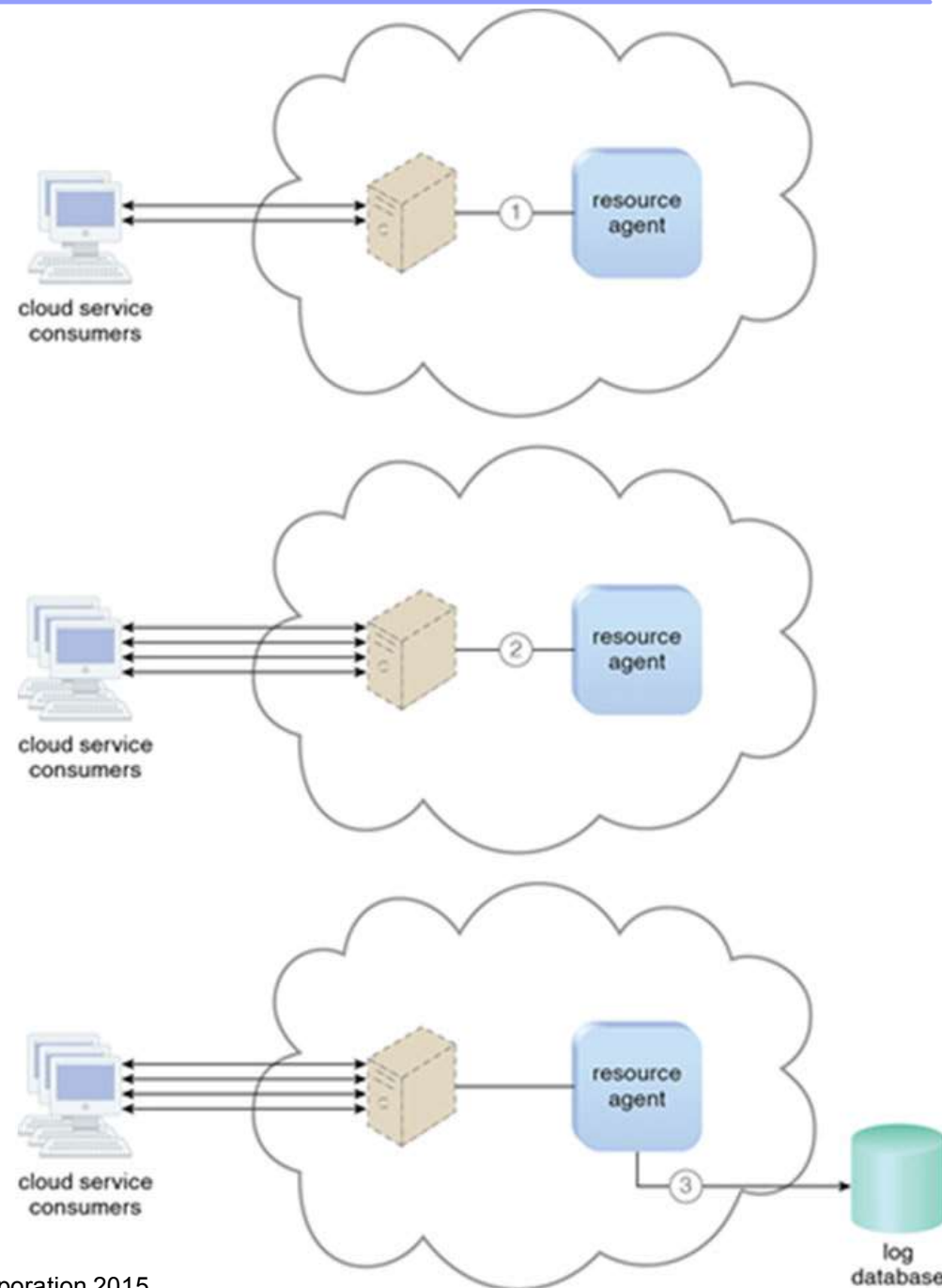
Polling Agent

- Figure - The polling agent monitors the status of a cloud service hosted by the virtual server by sending periodic polling request messages (MREQ1 to MREQN).

- It receives polling response messages (MREP1 to MREPN) that report usage status "A" after every polling cycle (1a).

- The polling agent stores the collected usage data in the log database (1b). Subsequently, the polling agent sends polling request messages (MREQN+1 to MREQN+M) and receives response messages (MREPN+1 to MREPN+M) that report usage status "B" after every polling cycle (2a). The polling agent stores the new usage data in the log database (2b).
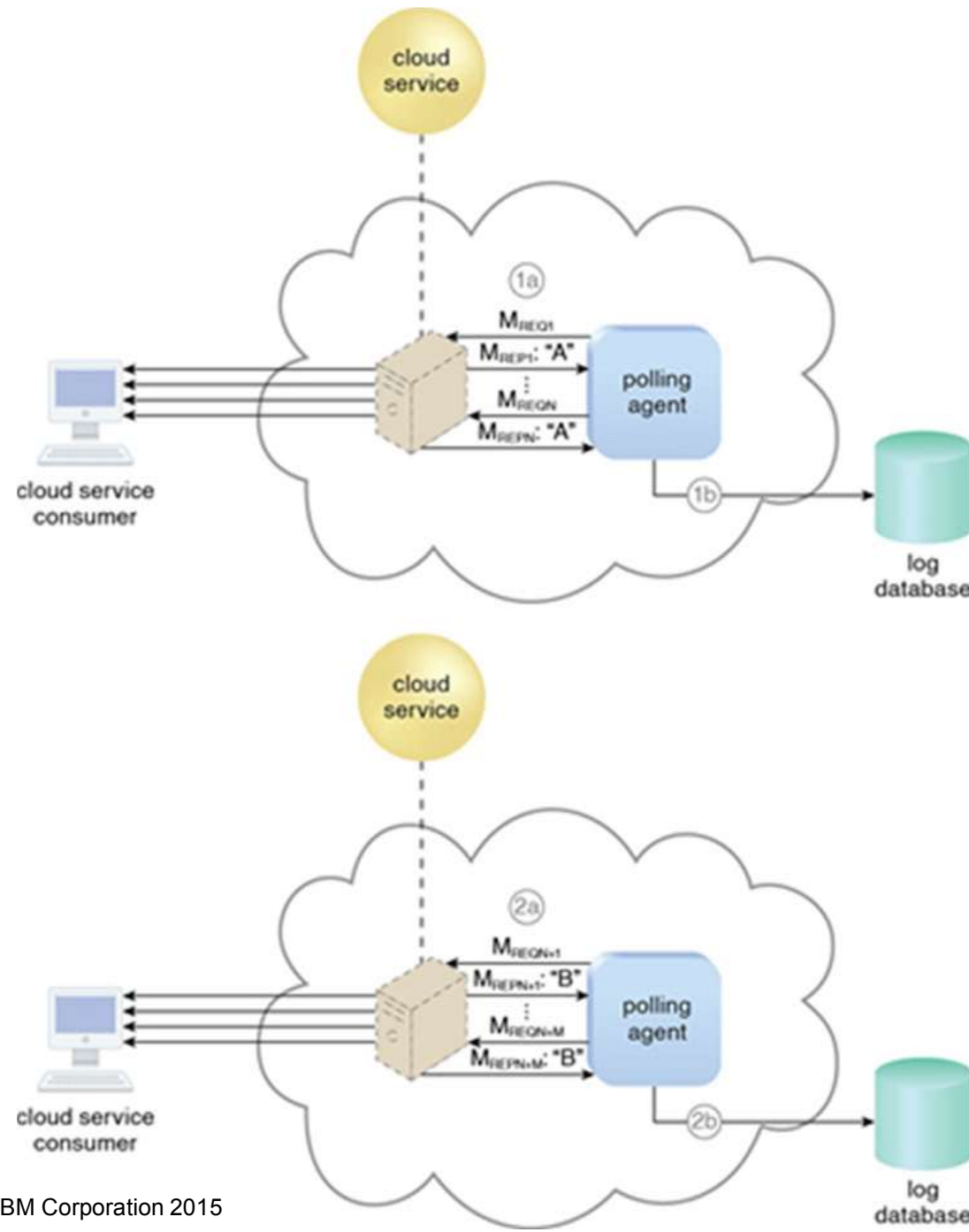
Key Benefits and Features

- Manage "compute anywhere"

- Monitor resources in public, private, or hybrid cloud atmospheres

- Incorporate management and visibility of cloud and data center resources

- Report on availability and performance of business-critical applications in any cloud

- Report on cloud-computing rate estimates at any given time

- Utilize public cloud providers' native APIs for real-time views of performance

- Monitor health and performance of cloud computing resources

Cloud Monitoring Features

Benefits of using Applications Manager's Cloud Monitoring Feature:

- The virtual, physical and cloud setup are monitored with a single console

- Applications presented in the cloud as well as inside the corporate data center are monitored

- Confirm high uptime and peak performance of your Amazon cloud resources

- Obtain insight into the performance of your cloud computing environment; troubleshoot and resolve problems before end users are affected

- Propose capacity and make educated decisions about allocating cloud resources

- To set up and manage easily, agentless monitoring solution is used

Metering and Billing
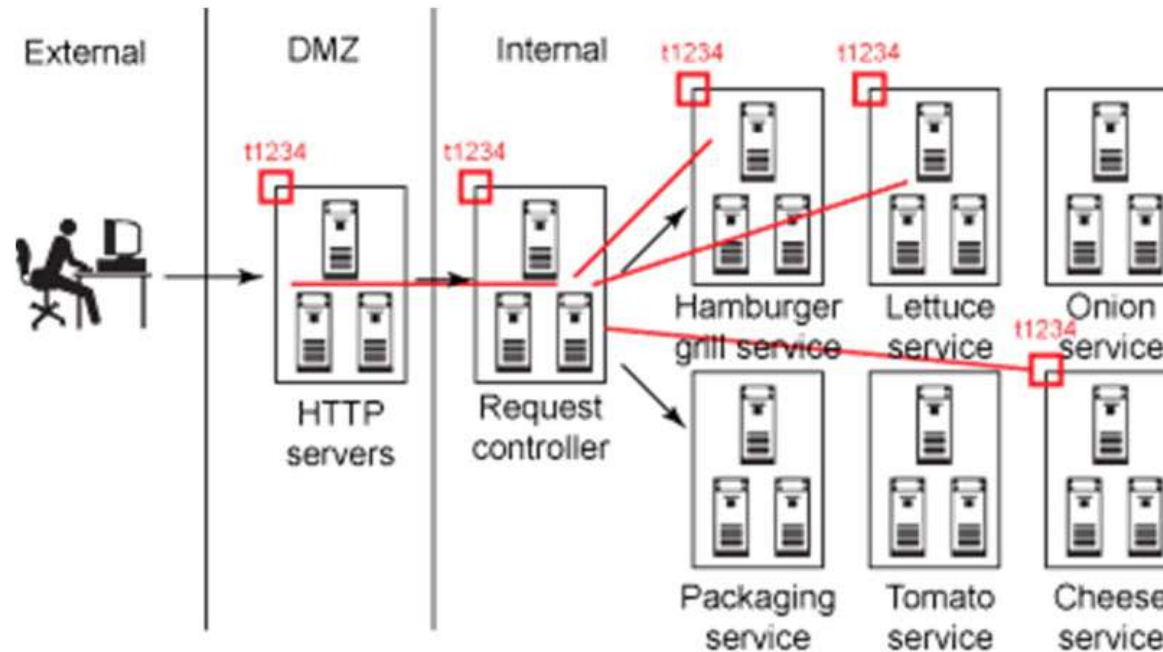
Transaction example using SOA services and transaction IDs

- Cloud computing billing impacts
- Accounting for cloud cost allocation in the code
- Operations
- Established service models

IaaS Billing and Metering Services (1 of 2)

- Servers serving per hour an on-demand model

- Reserved servers for better planning

- Greater and lesser compute resource units based on application performance

- The number of instances consumed is based on volume-based metering

- Reserved and prepaid infrastructure resources

- Clustered server resources

Figure -. Monthly account activity for an IaaS platform

PaaS Billing and Metering Service

- Outgoing and incoming network bandwidth

- CPU time per hour

- Stored data

- High availability

- Monthly service charge

SaaS Billing and Metering Service

- Monthly subscription fees

- Per-user monthly fees

Up and Coming Service Models

- Secondary service models are in development, and many have regular billing and metering models that have gained acceptance in all levels of commerce.

- Since SaaS is gaining approval, it is possible that these up-and-coming models will improve in adoption, too.

- For instance, Database as a Service (DaaS) and Monitoring as a Service (MaaS) are being leveraged from SaaS providers and gaining traction for cloud computing and SaaS IT-focused enterprises.

Definitions

- Pricing rules: Pricing information specified by the service provider corresponding to the load on the system

- Load Range: The difference between minimum and maximum load values between which a price value holds

- Threshold load: The load values on the cloud, which identify the load condition

- Instance: A virtual machine installed on the Eucalyptus cloud infrastructure

- Current price: The cost of utilizing the cloud service in a specified load condition. This is obtained from the pricing rules specified by the service provider and the current load estimated

Billing Model and Pricing

- Pricing and billing are two important aspects addressed in the Smart Metering implementation. The mechanism of pricing and billing are described in this section.

- **Pricing**

- ✓ Pricing for a cloud service can be applied based on multiple considerations.

- **Billing**

- ✓ Dynamic billing on cloud is a function of the instantaneous load on the cloud and the pricing information obtained as per the configuration specified by the service provider.
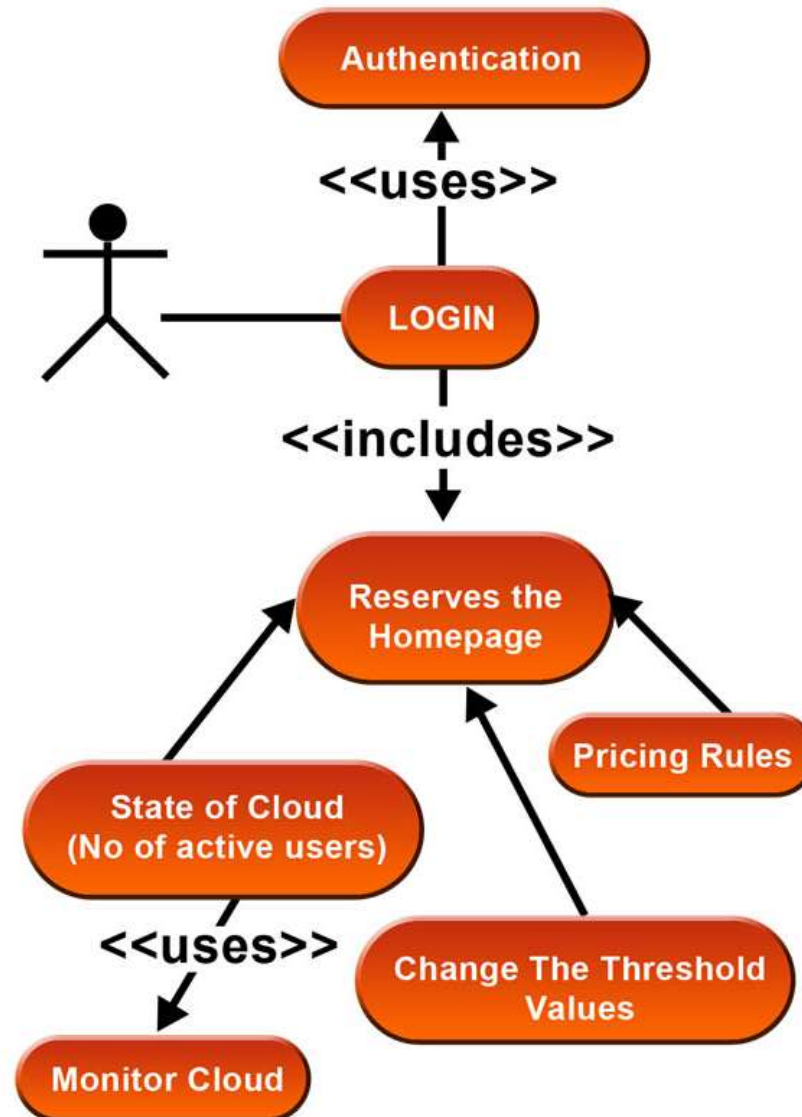
Smart Metering - Architecture
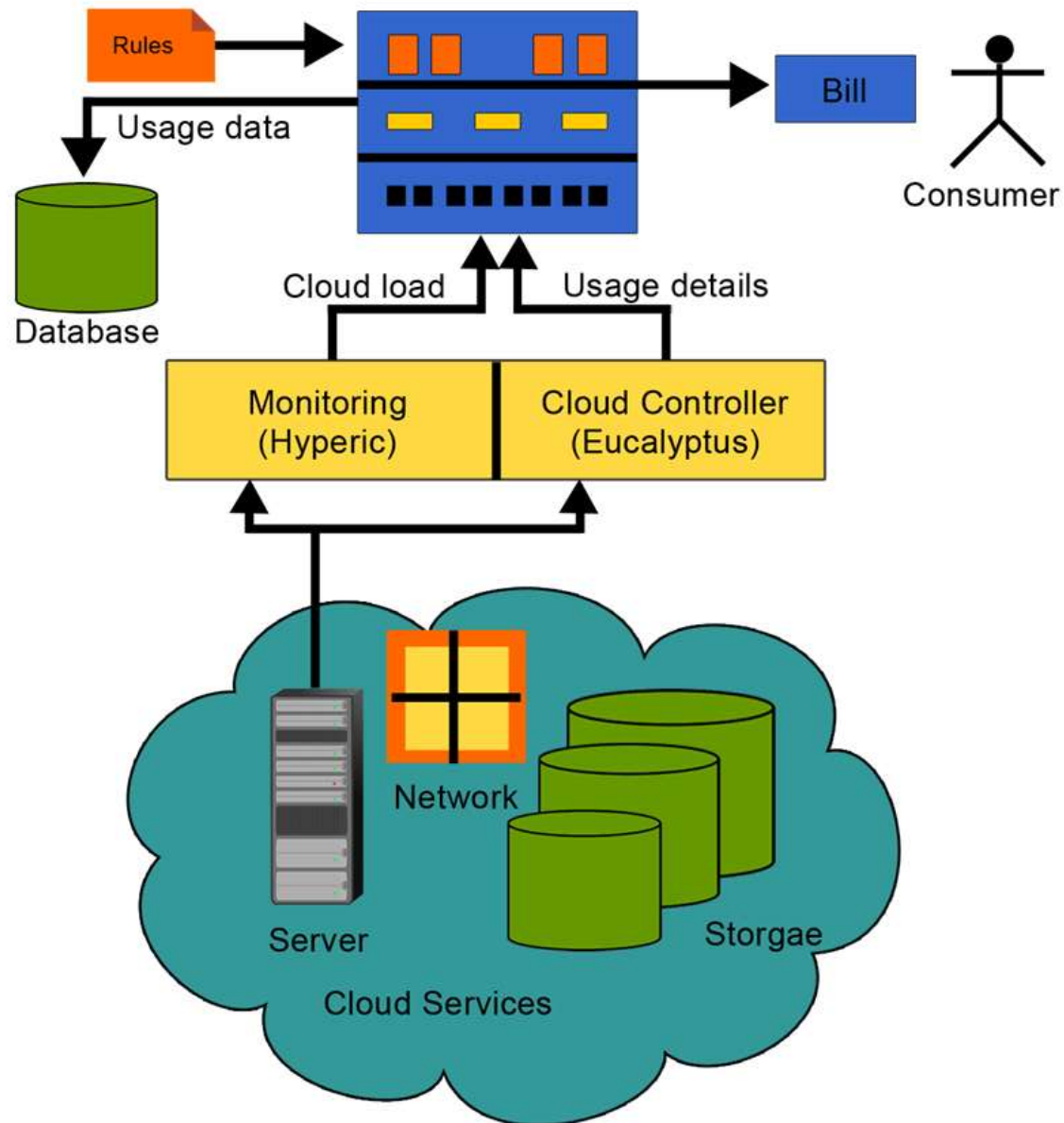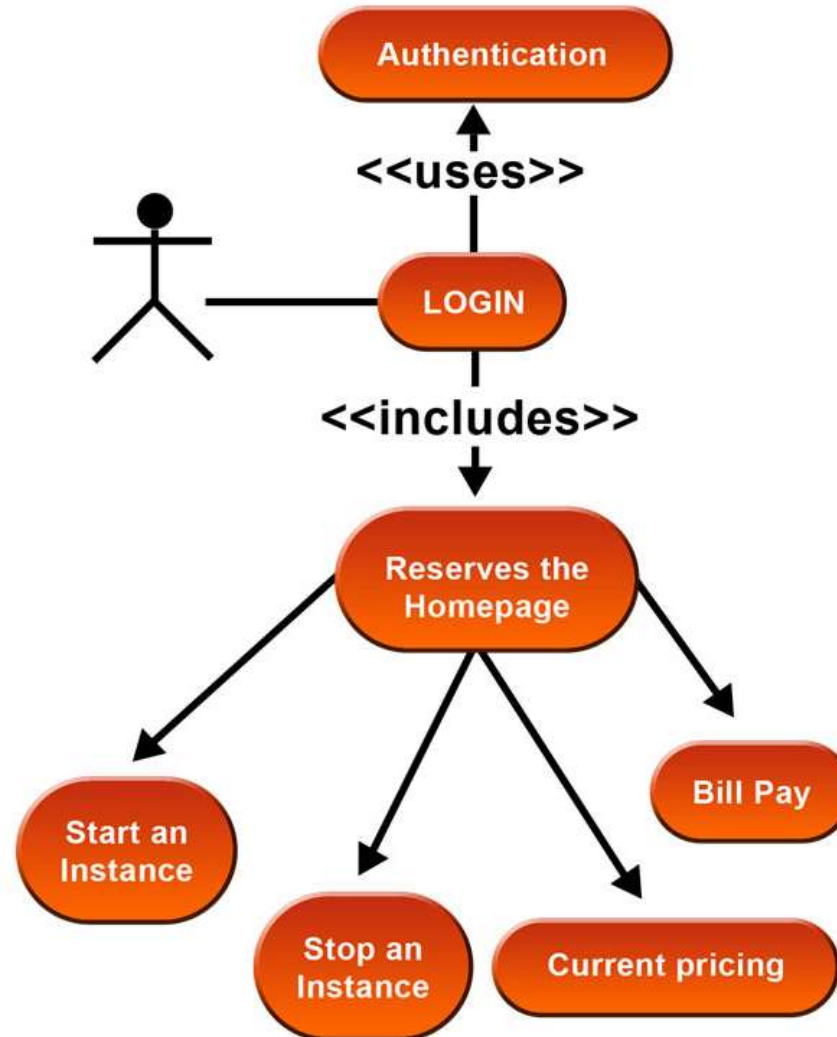
Smart Metering - Workflow
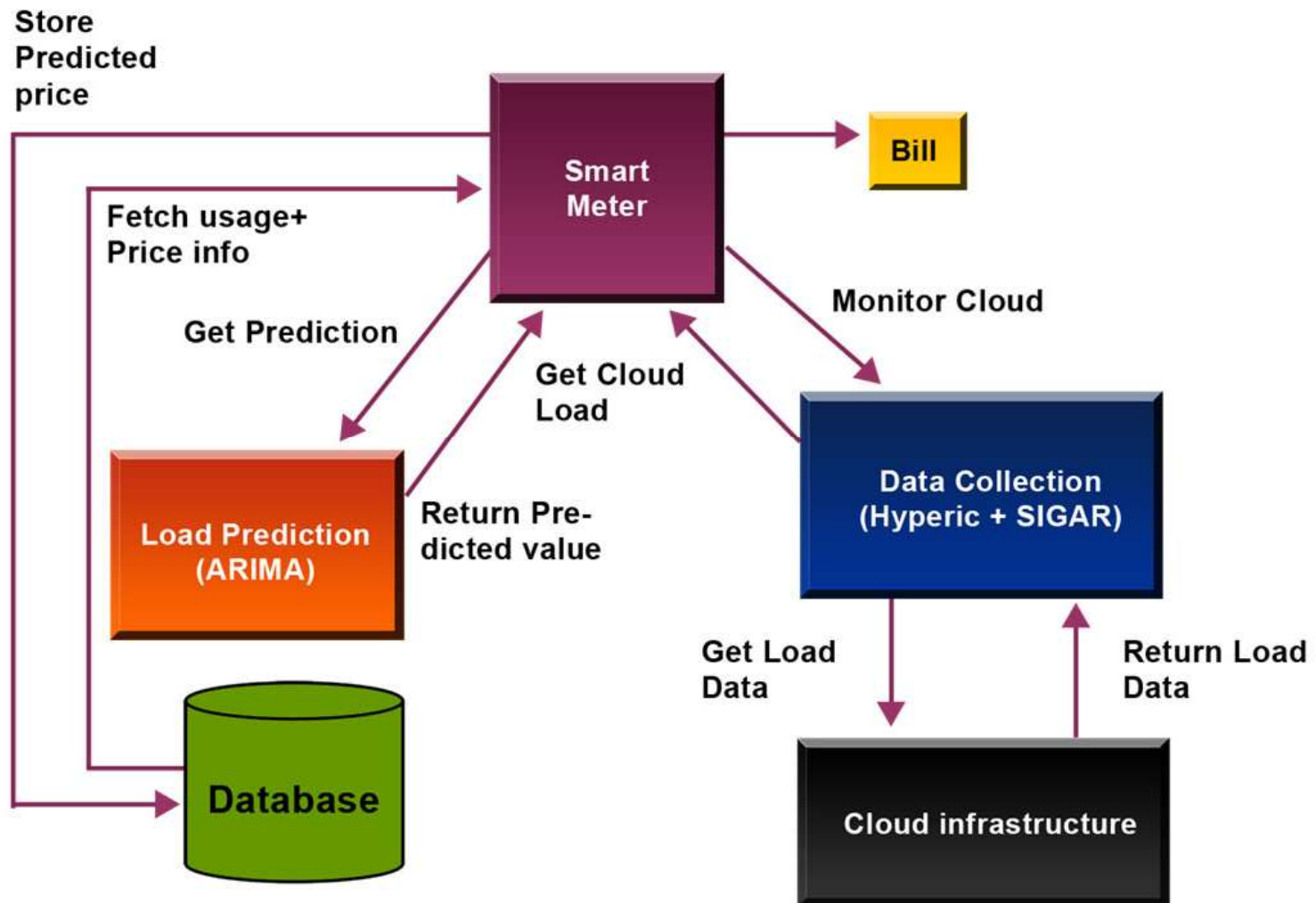
Smart Metering – Cloud Set-Up
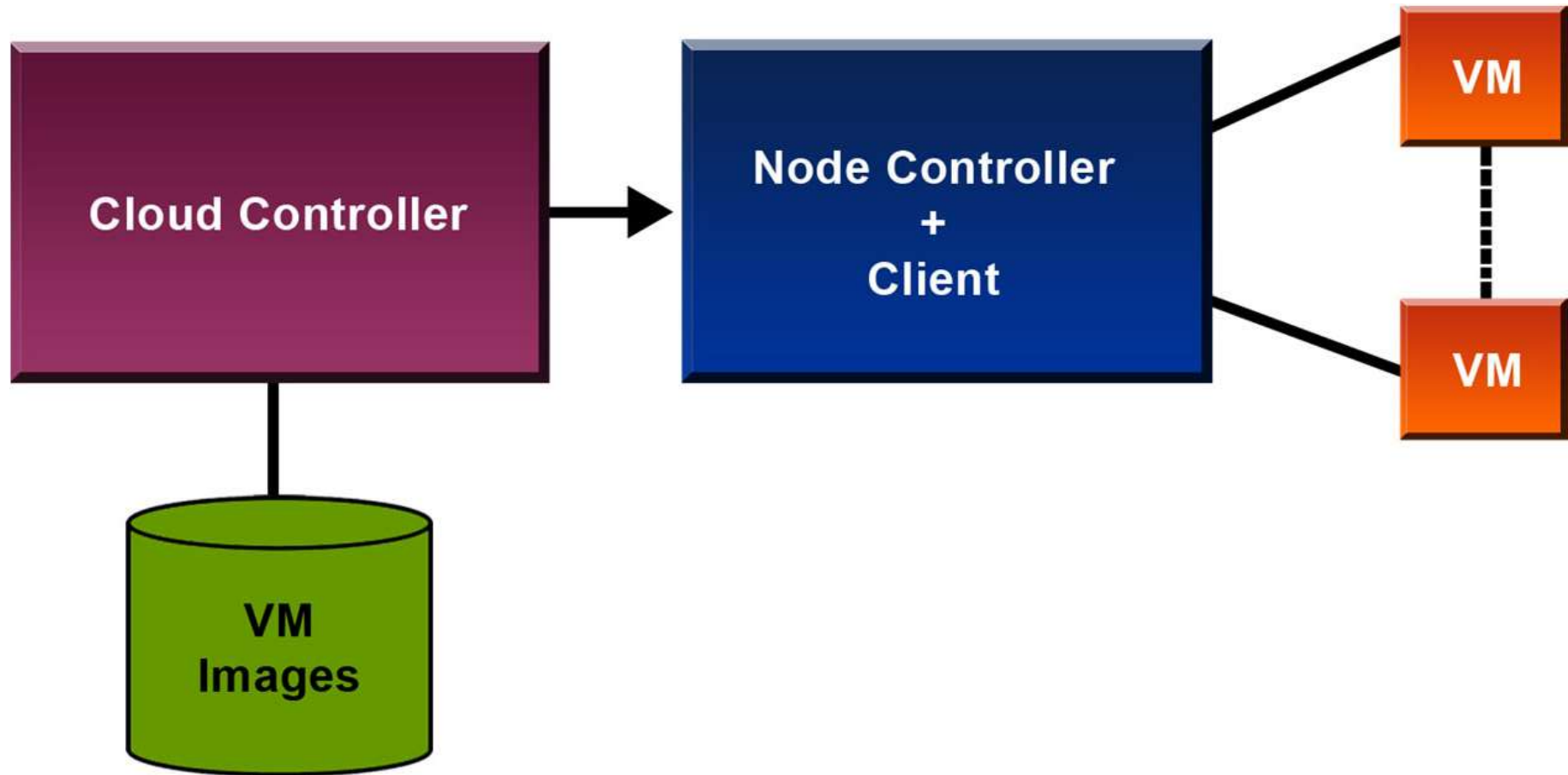
ARIMA

- ARIMA is used to predict the load based on the cloud based on the metric data collected from the Hyperic monitoring tool.

- The Auto Regressive model assumes that the current value of a process xt, can be defined by a finite linear aggregate of the previous values of the process along with the current value of a white noise signal at.

IOPs and Capacity Based Metering Metric

Example: How IOPs + Capacity Based metering works

- 24 powered on VMs with a total of 2400 IOPs
- 8 Powered off VMs do not produce IOPs but consume capacity

vDC A

vDC B
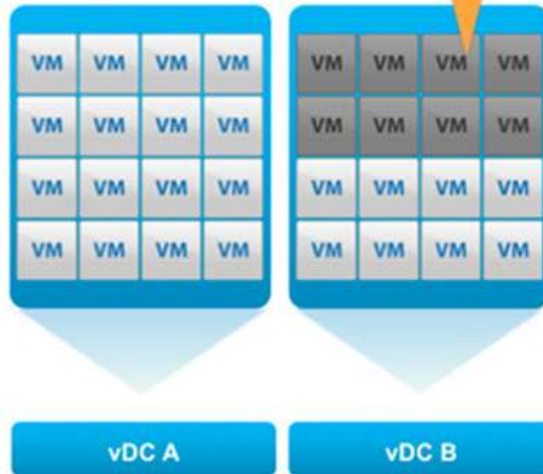
Virtual Datacenter A
16 VMs Powered On

Virtual Datacenter B
8 VMs Powered On

Powered-on VM   Powered-off VM

Example

- User creates 32 VMs and their total IOPS consumption is 2400

Virtual Datacenter A	
Total Count of Powered On VMs	16
Total IOPs	1600
Total Capacity	100GB X 16=1600GB

Virtual Datacenter B	
Total Count of Powered On VMs	8
Total IOPs	800
Total Capacity	100GB X 16=1600GB

Total Billed to SP for 30-Day Month	
Total IOPs	2400 (1600 + 800)
Total Capacity	3200 GB
Total Points (Some Bundle = x Pts & y Pts) (Total IOPs * X Points) + (Total Capacity * Y Points)	2400 X + 3200 Y

Slab Based Billing

Vcenter Server
VM1 VM2 VM3

Get IOPs and Capacity every month

Save the Data to DB

DB

Usage meter

Monthly Reporting / Report on demand

Get Data for report

((capacity*Xpt) + (IOPs* Ypt))

Calculate IOPs and capacity to points

Total points (p) for current month

Slab based calculator

IF P<S1 — No

IF S1<P< S2 — No

P>S2

Yes — Slab ID

Yes — Slab ID

Yes

Slab ID

P & Slab ID

Slab ID

Total charges= Px Slab charges

Cloud Computing Services

Cloud Service Provider
(Google, Amazon, Microsoft, IBM, Oracle, Sales force)

Cloud Computing Services

IaaS (windows/ Linux VM)

PaaS (Middleware/ Database stack)

SaaS (ERP/email CRM)

Cloud Computing Clients
(Enterprises/Individuals)

Cloud System Administration

Cloud Computing Benefits to a Client/Consumer

- Reduced Capital Cost to setup IT Infrastructure

- Utility Pricing Model

- Rapid scalability with the help of dynamic infrastructure

- Self Service by using Automated Provisioning

- Resource availability from anywhere of the world

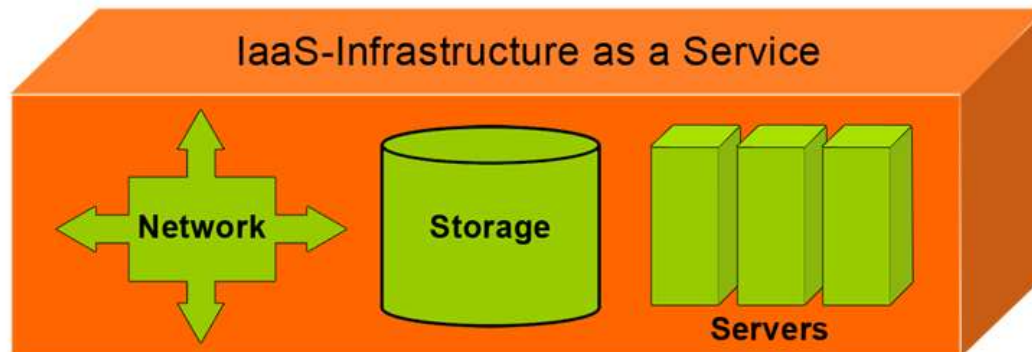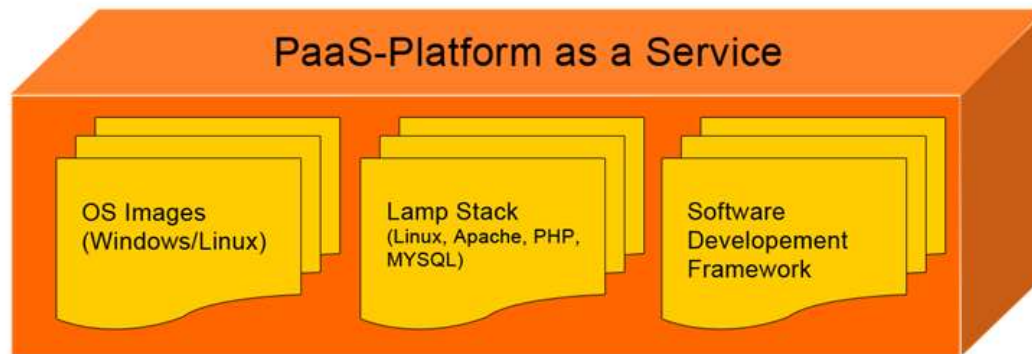Cloud Computing Solutions

Figure Possible Cloud Computing Solutions for Enterprise

Public and Private Cloud for Enterprise

Public Cloud Solution for Enterprise

- Computing service offered on the internet, and real computing resources are available

Private Cloud Solution for Enterprise

- Who within the organization can utilize and pay for cloud computing resources as if they are using public cloud resources

Hybrid Cloud Solution for Enterprise

- Hybrid cloud solution allows enterprise use both public cloud and private cloud resources at the same time depending on the importance and criticality of the business function

Virtual Private Cloud Solution

- Create their own private cloud environment within the public cloud by using different network/firewall guidelines to avoid external access to the enterprise resources

Infrastructure Support Team

- Hardware administration

- Operating System Builds

- Operating System Administration

- Network Services Administration

How Cloud Computing affects the Job roles in the Infrastructure Support Team
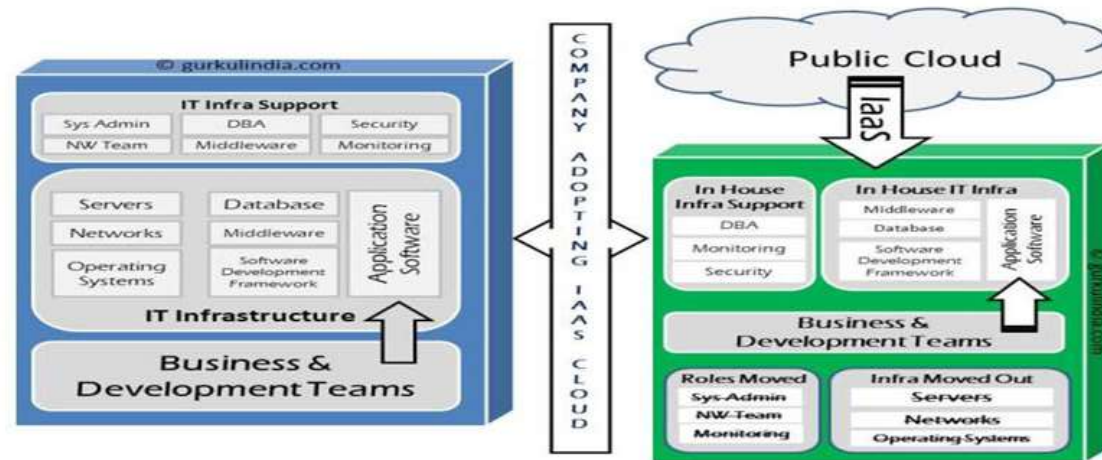
Figure Job Role movement with IAAS Cloud Computing Solution
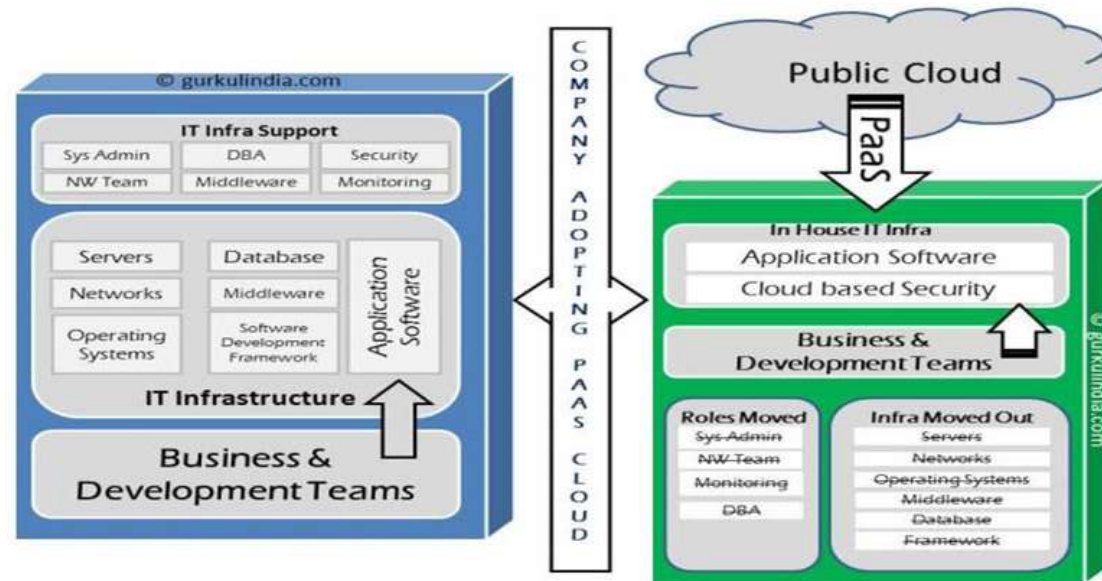
Figure. Job Role movement with PAAS Cloud Computing Solution

System Administration

- Accomplish routine audits of systems and software

- Execute backups

- Install operating system patches, configuration changes, and updates

- Troubleshoot any reported problems

- Adjust system performance

- Assure that the network infrastructure is up and running

- Handle basic programming

- Conduct basic end user training

- Accomplish projects

- Contribute in systems selection

- Offer periodic reports to management

The System Administrator Role

- The systems administrator role is possibly one of the earliest in technology.

- In the initial days of computing, technical professionals of electronics built prototype computers, and newly minted "programmers" wrote logical instructions for these computers.

- Meantime, the systems administration role owned the installation, operation, tuning, and configuration of these systems when they went into production and use on a larger measure.

Tasks involved in System Maintenance

- Corrective maintenance refers to any remedial steps needed to correct problems with the software, hardware, network, or utilization of information systems.

- Customized maintenance refers to modifying features already in existence in information systems that need to be updated or modified for user needs.

- Enhancement maintenance involves improving the performance of the application, hardware, network, or even the role of people in their development and use of these mechanisms.

- Preventive maintenance includes taking steps in advance to reduce the risk of a serious problem.

How the Cloud Changes Things

- The systems administrator has the same concerns and impacts of "the cloud" as the Systems Architect and the DBA.

- They need to educate themselves on the options within this knowledge, try a few test solutions out (experience) and of course work with others on various parts of the implementation.

- The three big buckets of cloud computing are dealing with Virtual Machines (IaaS), writing code (PaaS) and using software that is already written and being delivered via an Application Programming Interface (API).

- The systems administrator role normally tackles the first "bucket" most often - IaaS, which has at its base the technology of virtualization.

Virtualization

- One of the first areas the systems administrator is involved with "the cloud" is in the area of virtualization. This technology is not new.

- It is the process of using software to emulate hardware - which has implications far beyond that simple sentence.

Cloud Computing Architecture - Private, Public and Hybrid

- It is significant to note that IaaS can be a facility, or on-premises, or both. The first is called "private cloud", the second "public cloud", and the third "hybrid cloud". These are marketing languages, but they are useful in describing where the decisions are for deploying a structure.

- If data security is vital, then private cloud may be the right choice for a given assignment. If agility or cost is a concern, public cloud may be the right answer for another assignment. In many cases - perhaps most - using both architectures is the right way to split the workload.

- The key is to understand the workload well. Earlier the system administrator needed to know the component necessities, such as how much memory, CPU, network and storage a workload required. In cloud computing also these are concerns, however the user needs to add in the questions of location of users, cost, business use, security and other vectors. These apprehensions bring the systems administrator closer to the business and its goals.

DevOps

- One new term introduced into cloud computing is "DevOps" - short for Developer Operations. Everyone does not agree that this is even a real "thing" - that it is a made-up term by cloud dealers. Nevertheless, there is a new set of tasks that the cloud brings that may sit within the purview of the system administrator.

- Basically it involves the administration needed at the level of PaaS or SaaS. The IaaS function of cloud computing holds most of the same characteristics as an on-premises structure, defined the in the first list. But when the organization uses Platform as a Service, the operating system, much of the scale, security and other components of infrastructure are abstracted into the platform, and are often even controlled by the designer.

- However once the application "goes live", there is a host of controlling, billing, scaling and additional security questions that developers aren't equipped to handle. As companies are discovering, they need to appoint someone to cover these overlapped areas between developers and administrators.

PC Health Check

- Do a virus scan

- Use the disk defragmenter

- Diagnostic tools

- Empty the recycle bin

What Slows Down a Computer?

- Data clutter

- Insufficient memory

- System Resources

- Failing Hardware

Running a PC Health Check

- Do a virus scan to make sure the machine isn't infected with viruses or malware.

- Use the disk defragmenter tool to defrag your hard disk. This is probably an overnight process.

- Verify any diagnostic tools that came pre-loaded with the machine (like Dell Diagnostics).

- Verify your devices in the Control Panel to make sure nothing is malfunctioning.

- Check your hard drive for errors (usually found in the Properties dialog of your C: drive)

- Update your software / hardware drivers.

- Empty the recycle bin and the Internet browsing history.

Patching and updates

- Fix a software bug

- Install new drivers

- Address new security vulnerabilities

- Address software stability issues

- Upgrade the software

Why Manual Patching is mandatory in Enterprise Framework?

- System and server administrators need complete analysis regarding the impact of newly released patch before replacing it with previous one.

- Apparently, pre-assessment process is not viable with automated patching
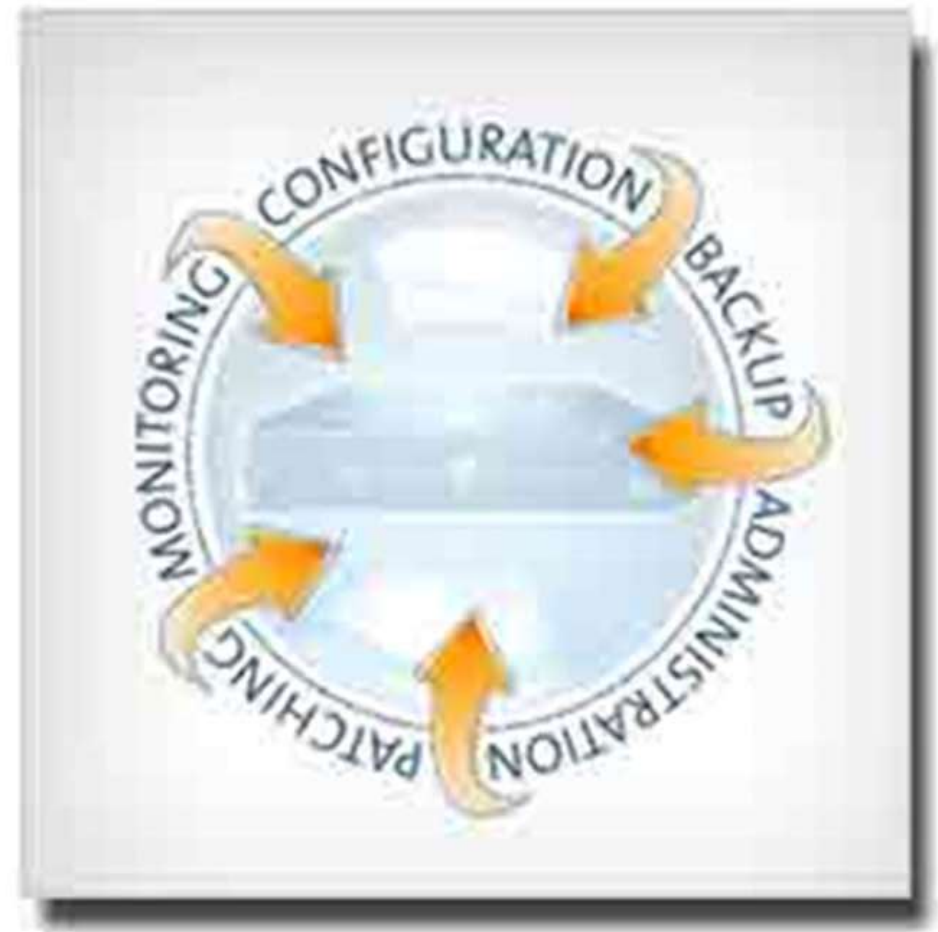
Tools for Patch Management

- Windows Server Update Service

- Red Hat Network

- VMware vCenter Update Manager

- Tivoli Endpoint Manager

Patching demands Proper due Diligence
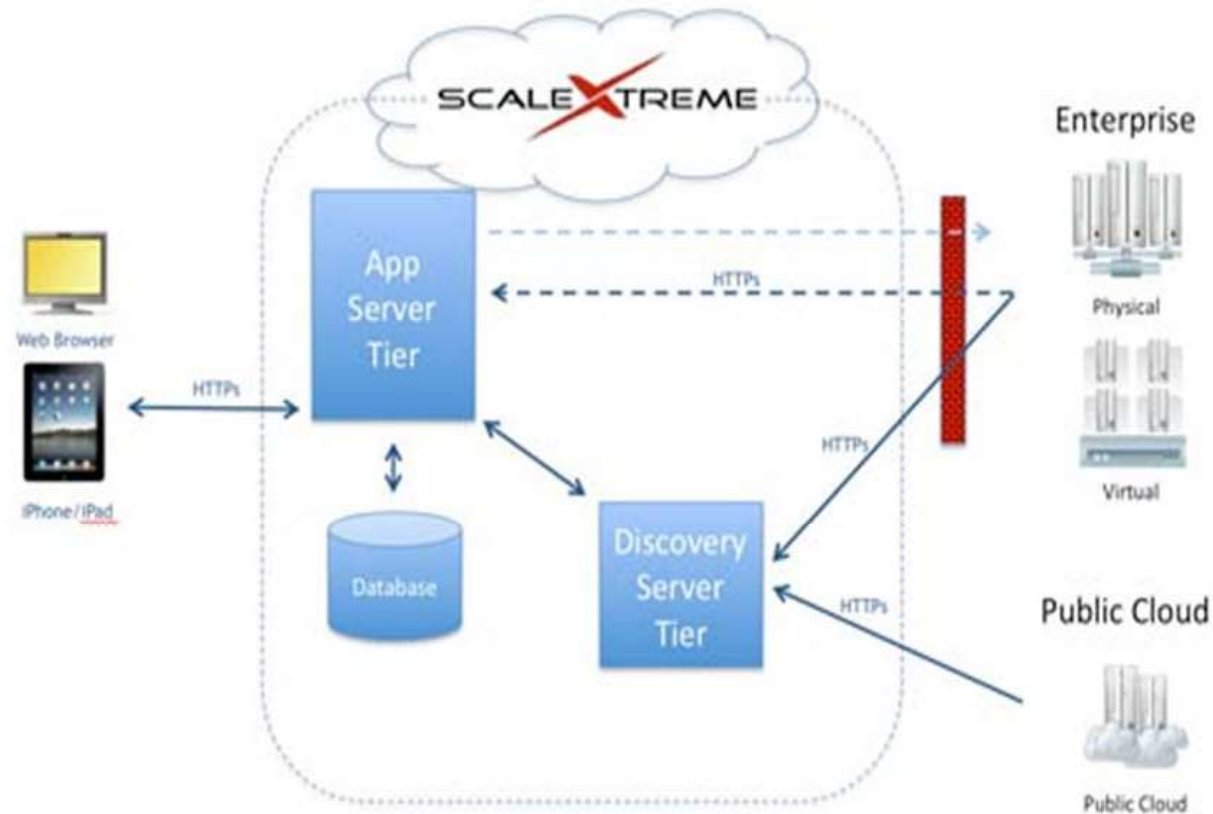
Figure Scalextreme

Patch Management in the Cloud – It is About Consistency and Automation

- There are 2 separate thoughts on how patches can be pushed to keep an environment reliable.

- Patch management can either be done via replacement or update of active workloads or virtual machines.

- In-Place Update

- In-Place Swap

Maintenance outages

- Servers and network systems require occasional planned outages for repairs and improvements

- Whenever possible, UIS will schedule these planned outages during our defined outage window.

- An outage window is defined in advance so that event schedulers know to avoid scheduling events during these times.

- All planned outages will be announced, with as much advance notice as possible, on the UIS main page, via emails to appropriate user lists, and (when outages affect wide populations) via mass broadcast emails.

- **Maintenance & Outages** - necessary to keep our IT systems well maintained so that they function dependably and securely. Each maintenance situation is different depending on what needs to be done and what systems are being affected

- Outages come in two varieties - planned and unplanned

System Maintenance

- Database application maintenance

- Operating system maintenance

- Hardware maintenance

- Patches and minor updates

Service levels (1 of 2)

- Service level agreements, it is a common feature to mention a percentage value (per year or per month) that is calculated by dividing the sum of all downtimes timespans by the total time of a reference time span.

- If the server was available every time, it means 0% downtime.

- For Internet servers downtimes above 1% per year or worse can be regarded as unacceptable as this means a downtime of more than 3 days per year.

- In case of e-commerce and other industrial use any value above 0.1% is usually considered unacceptable.

Planning

- Deferred maintenance

- Diagnostics to isolate a detected fault

- Hardware fault repair

- Fixing an error or omission in a configuration database or omission in a recent configuration database change

- Fixing an error in application database or an error in a recent application database change

- Software patching/software updates to fix a software fault

Avoidance - For most websites, website monitoring is available. Website monitoring (synthetic or passive) is a service that "monitors" downtime and users on the site.

Downtime

- Downtime can also refer to time when human capital or other assets go down. For example, if employees are in meetings or unable to perform their work due to another limitation, they are down. This can be equally expensive, and can be the result of another asset (i.e. computer/systems) being down. This is also commonly known as "dead time". This term is also used in factories or industrial use.

- Measuring downtime - There are a many external services, which can be used to monitor the uptime and downtime as well as availability of a service or a host. A notable example is Pingdom.

- Sending notifications - "Service X will be unavailable between 00:00 CST January 4th and 12:00 CST January 10th. We apologize for the inconvenience." Sending notifications is to tell people about the problem the administrator is having or fixing, and what he is doing to fix it.

A well-defined and managed service catalog offers:

- A self-service opening for customers

- A solitary portal to present all service and product contributions

- A consistent approach to request satisfaction

- Administration of customer prospects

Setting up a Service Catalog (1 of 2)

Role Title [Name]	Description
Administrator [admin]	Can manage all aspects of the Service Catalog application, with scripting functions such as creating UI macros or business rules.
Catalog administrator [catalog admin]	Can manage the Service Catalog application, including catalogs, categories, and items, but not including scripting functions available to administrators.

- Administrators and catalog administrators, catalog_admin role users, can use the Service Catalog application to define service catalog content and layout

- Administrators and catalog administrators can further extend the service catalog to provide more powerful structures, using specialized catalog stuffs, configuration options, and scripting functions

Service Catalog

Menus and Modules

- Catalogs: Define and enable a homepage for multiple catalogs

- Catalog: Contact the service catalog portal page to view and request items

- Open Records: View open Requests, Tasks, or Items

Menus and Modules

- Catalog Definition

- Maintain Catalogs

- Maintain Categories

- Renderers

- Maintain Dynamic Categories

- Maintain Items

- Content Items

- Ordered Item Links

- Order Guides

- Record Producers

Catalog Policy

- Scriptable Order Guide Failures

- Properties

- Execution Plans

- Fulfillment Groups

- Catalog Client Scripts

- Catalog UI Policies

Catalog Variables

- All Variables: View and edit all service catalog variables

- Item Variables: View and edit only the variables that are associated to a catalog item

- Plan Variables: View and edit variables used with execution plans

- Variable Sets: View and edit groups of variables that can be shared as sets between catalog items

- **Mobile Admin**

  Mobile Layout: Configure the service catalog smartphone layout

Service Catalog Management (1 of 2)

- Objective of ITIL Service catalog Management aims to ensure that a Service catalog is produced and maintained, covering accurate information on all operational services and those being prepared to be run operationally.

- Service Level Management process mentioned the concept of a Service catalog.

- ITIL V3 takes this concept further, presenting a dedicated process to ensure that the Service catalog is up-to-date and contains reliable information

Roles | Responsibilities

- Remarks

- Phases of maintaining the Service catalog

Accountability Matrix: ITIL Service catalog Management	
ITIL Role / Sub-Process	Service catalog Manager
Service catalog Management (no sub-processes specified)	A R

Service catalog Contents (1 of 3)

Service catalog Contents (2 of 3)

- **Business or Technical**

  enables to specify which of the CIs in the system are Business Services (those services used by the end-users)

- **Request Mode**

  enables to specify which of the Services in the system can be requested only once by each user through the Service catalog and which can be requested many times

- **Bundles**

  The Bundles tab allows to group services together into a Bundle that can then be requested by end-users as a single item from the Service catalog

- **Publishing**

  The Publishing tab allows to define which groups and roles have access to which services in the Service catalog. This permits to control access to the different services, where services are available only to those users who are entitled to them

Service Design Process

- This process defines and explains the supporting Service Design processes. They are responsible for providing key information to the design of new or hanged service solutions.

 5 aspects of design, that need to be considered :

- Design of the services, comprising all functional requirements, capabilities and resources

- Design of Service Management systems and tools, specifically the Service Portfolio, for the control and management of services through their lifecycle

- Design of technology planning and management systems required to provide the services

- Design of the processes needed to design, operate, transition and improve the architectures, the services and the processes themselves

- Design of the measurement metrics and methods of the services, the architectures and their constituent components and the processes.

Service Catalog Management

The Service Catalog Management events should include:

- Definition of the service

- Construction and maintenance of an accurate Service catalog

- Dependencies, interfaces and consistency between Service catalog and Service Portfolio

- Dependencies and interfaces between all services and supporting services within the Service catalog and the CMS

- Dependencies and interfaces between all services, and supporting components and Configuration Items (CIs) within the Service catalog and the CMS.

Service Catalog - Value to the business

- The Service catalog provides a central source of information on the IT services delivered by the service provider organization.

- It ensures that all areas of the business can view an accurate, reliable picture of the IT services, its status and the details.

- It contains a customer-facing view of the IT services in practice, how they are intended to be utilized, the business processes they facilitate, and the levels and quality of service the customer can expect for each service

Policies, principles and basic concepts

- Every organization should develop and sustain a policy with regard to both the Catalog and the Portfolio, relating to the services within them, the details to be recorded and the statuses to be recorded for every service. The policy must also cover the details of responsibilities for each section of the overall Service Portfolio and the scope of each of the constituent sections.

The Service Catalog has two aspects:

- The Business Service Catalog: containing details of all the IT services delivered to the client, together with relationships to the business units and the business process that rely on the services of IT. This is the customer view of the Service Catalog.

- The Technical Service Catalog: containing details of all the IT services delivered to the client, together with relationships to the components, shared services, supporting services and CIs necessary to support the provision of the service to the industry. This should underpin the Business Service Catalog and not form part of the customer view.

The Service Catalog

The Service Catalogue

Business Process 1

Business Process 2

Business Process 3

Business Service Catalogue

Service A

Service B

Service C

Service D

Service E

Technical Service Catalogue

Support Services

Hardware

Software

Applications

Data
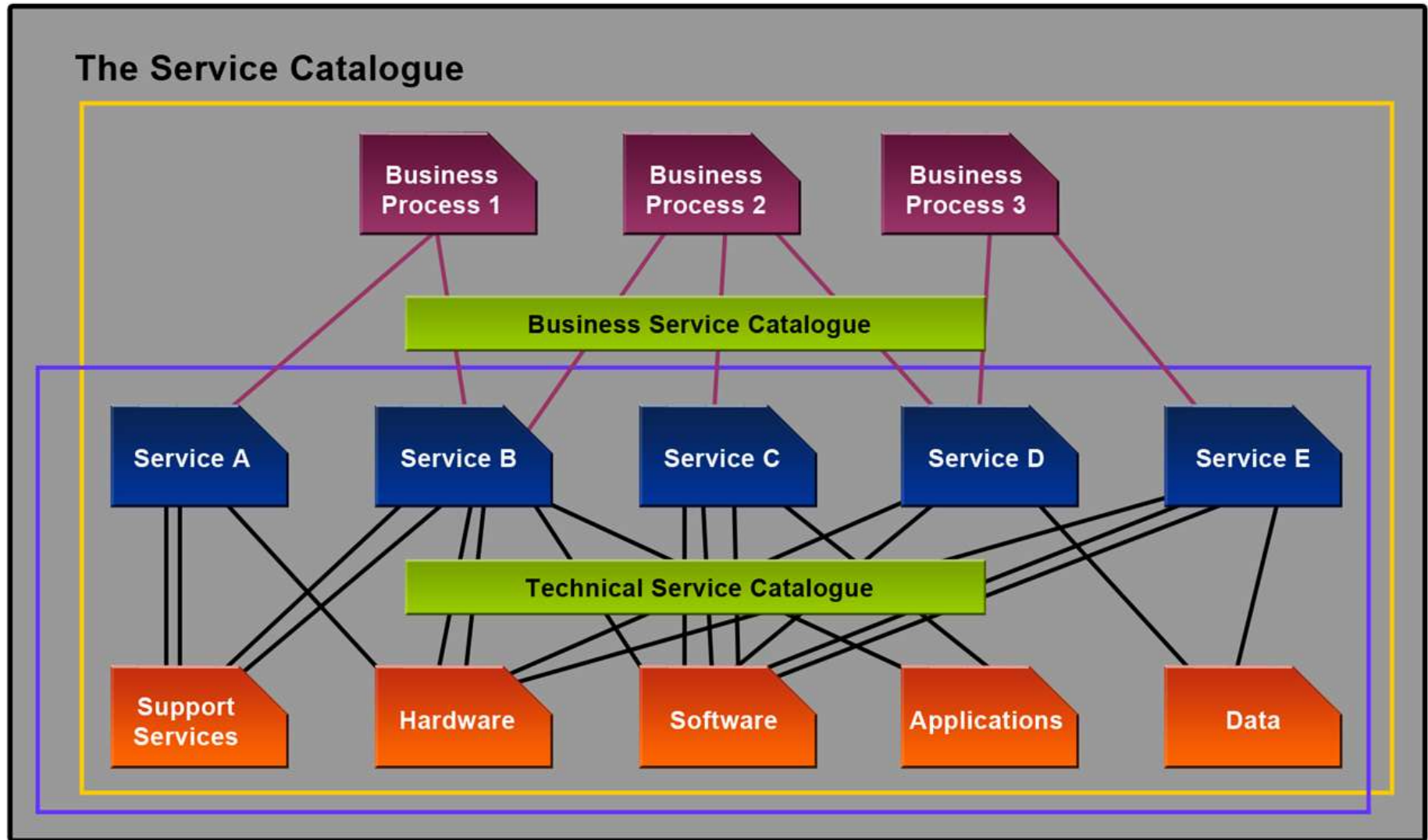
Figure The Service Catalog

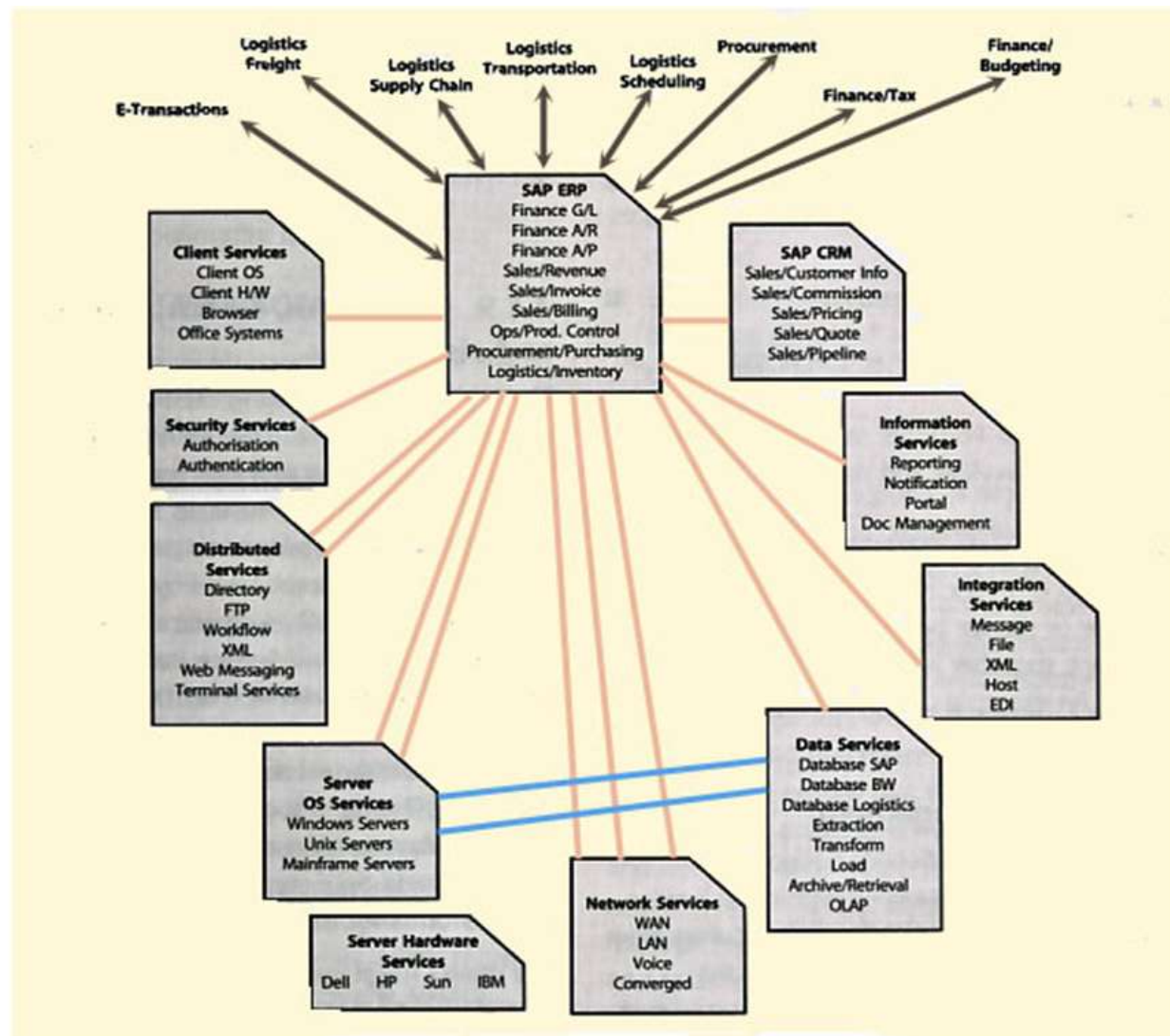# Process activities, methods and techniques

Figure The Services

Information Management

- The key information within the Service catalog Management process is that contained within the Service catalog.

- The core input for this information comes from the Service Portfolio and the business via either the Business Relationship Management (BRM) or Service Level Management (SLM) procedures.

Key Performance Indicators

- The quantity of services recorded and managed within the Service catalog as a percentage of those being delivered and transitioned in the live environment

- The quantity of variances detected between the information contained within the Service catalog and the 'real-world' situation

Challenges, Critical Success Factors and Risks

- The chief Critical Success Factors for the Service catalog Management process are:
- ✓ An accurate Service catalog
- ✓ Corporate users' awareness of the services being provided

- IT staff awareness of the technology supporting the services. The risks associated with the provision of an accurate Service catalog are:
- ✓ Inaccuracy of the information in the catalog and it not being under rigorous Change control
- ✓ Poor acceptance of the Service catalog and its usage in all operational methods. The more actively the catalog exists, the more likely it is to be accurate in its content
- ✓ Inaccuracy of information received from the IT, the Service Portfolio and business, with regard to service information
- ✓ The resources and tools required to maintain the information
- ✓ Poor contact to accurate Change Management information and processes
- ✓ Avoidance of the use of the Service Portfolio and Service catalog
- ✓ The information is either too detailed to maintain accurately or at too high a level to be of any importance. It should be consistent with the level of detail within the CMS and the SKMS.

Service Level Management

- Service Level Management (SLM) negotiates, approves and documents appropriate IT service targets with representatives of the industry, and then monitors and produces reports on the service provider's ability to deliver the agreed level of service

- SLA is effectively a level of assurance or warranty with regard to the level of service quality delivered by the service provider for each of the services delivered to the business

Purpose/goal/objective

The objectives of SLM are to:

- Define, document, monitor, agree, measure, report and review the level of IT services provided

- Deliver and improve the relationship and communication with the business and customers

- Confirm that specific and measurable targets are developed for all IT services

- Observe and improve customer satisfaction with the quality of service delivered

- Make sure that IT and the customers have a clear and unambiguous expectation of the level of service to be delivered

- Make certain that proactive measures to improve the levels of service delivered are implemented wherever it is cost-justifiable to do so.

Scope

- SLM needs to manage the expectation and perception of the business, customers and users and ensure that the quality of service delivered by   the service provider is matched to those expectations and needs.

- In order to do this successfully, SLM should establish and maintain SLAs for all current live services and manage the level of service provided to meet the targets and quality measurements contained within the SLAs. SLM should also produce and agree to SLRs for all planned, new or changed services.

Troubleshooting (1 of 2)

- Troubleshooting is a method of problem solving, regularly applied to repair failed products or procedures

- It is a logical, organized search for the source of a problem so that it can be solved, and the product or process can be made operational for another time

- Troubleshooting is needed to develop and maintain complex systems where the symptoms of a problem can have many possible reasons

- Troubleshooting is the identification of, or diagnosis of "trouble" in the management flow of a corporation or a system caused by a failure of some type

- Half-splitting - Efficient methodical troubleshooting starts with a clear understanding of the expected behavior of the system and the symptoms being perceived

- Reproducing symptoms - One of the core principles of troubleshooting is that reproducible problems can be reliably isolated and resolved

- Intermittent symptoms - Difficult troubleshooting issues relate to symptoms which occur intermittently

- Multiple problems - Segregating single component failures which cause reproducible symptoms is relatively straightforward

- If a user follows these steps, it helps to identify a problem and direct in finding the resolution to a problem.

- Any Error messages?

  If the user is getting any error messages, he has to write down the error and do a search for that error note. Computer Hope and millions of other Internet sites have documents relating to error messages.

- Reboot the computer

  If the computer is acting abnormal, frozen, or encountering errors but can reboot, then boot the computer. Mostly rebooting the computer can solve many computer problems. We highly recommend not doing any of the steps below until the computer has been rebooted.

- Any new hardware or software added?

- Has the computer moved?

- Is there any power outages or electrical storms?

- Hardware or software issue

  If the problem is a hardware issue and if it is possible to determine what hardware device is causing the problem, try the basic troubleshooting steps for that hardware device. If the problem is hardware related, but unable to determine which hardware device is failing, for instance, when a computer does turn on, the problem can be resolved using the process of elimination.

- Install latest patches or update drivers

- Scan for malware and viruses

- The last run of computer with no problems

  If the computer has worked in the past and the user is running Microsoft Windows consider running a Windows system restore to restore the computer back to an earlier date. The system restore will not erase any data but will get settings back to how they were on an earlier date.

Basic Troubleshooting Techniques

The computer goes blank before the Word document was saved. The explorer window freezes for no reason. The user can't hear anything from the speakers

- Always check the cables

- Isolate the problem

- Take notes about error messages

- Remember the steps taken, otherwise write them down

Simple solutions to common problems

- Most of the time, problems can be fixed by using simple troubleshooting methods, such as closing and reopening the program.

- It is essential to try these simple solutions before resorting to more extreme processes.

- If the problem still isn't fixed, the user can try other troubleshooting techniques, similar to reinstalling the software

Program isn't working properly or runs slowly

- If a program isn't working properly or otherwise running slowly, the first thing the user should try is closing the program and reopening it.

- The user can also shut down the system, wait a few seconds, and boot it up again. Certain minor problems will work themselves out when the user does this.

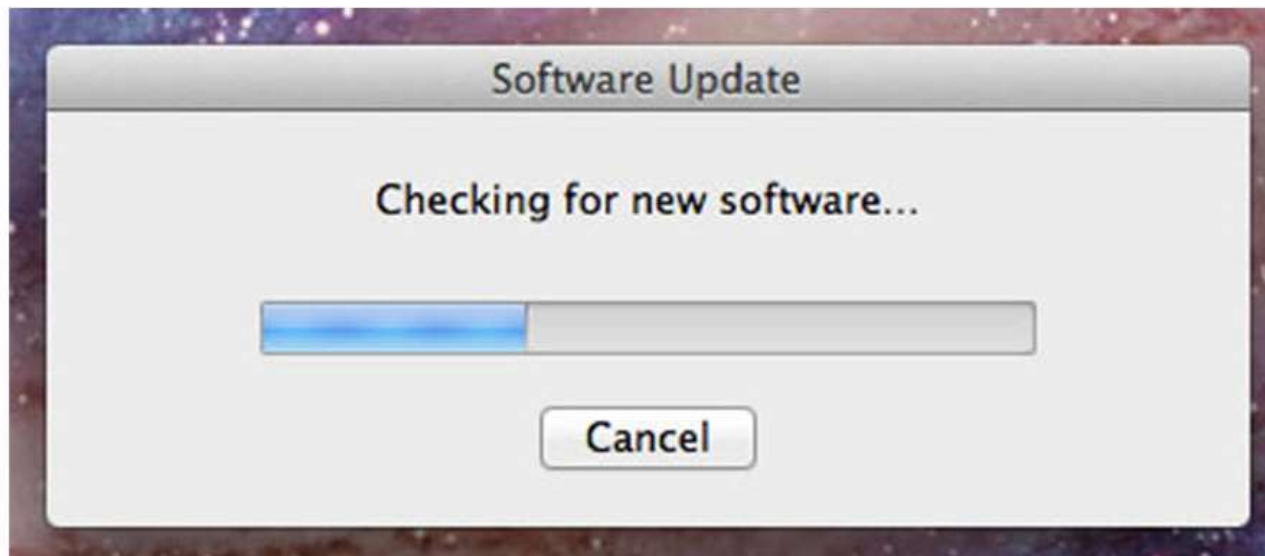- Check with the company for any known problems or updates to the software.

Figure Software Update Status

Program is completely unresponsive

- If a program has become completely unresponsive, the user can press (and hold) Control+Alt+Delete on the keyboard to open the Task Manager.

- The user can then select the program that isn't working and click End Task. If the user is using a Mac, he can press Option+Command+Esc to open a similar dialog box.
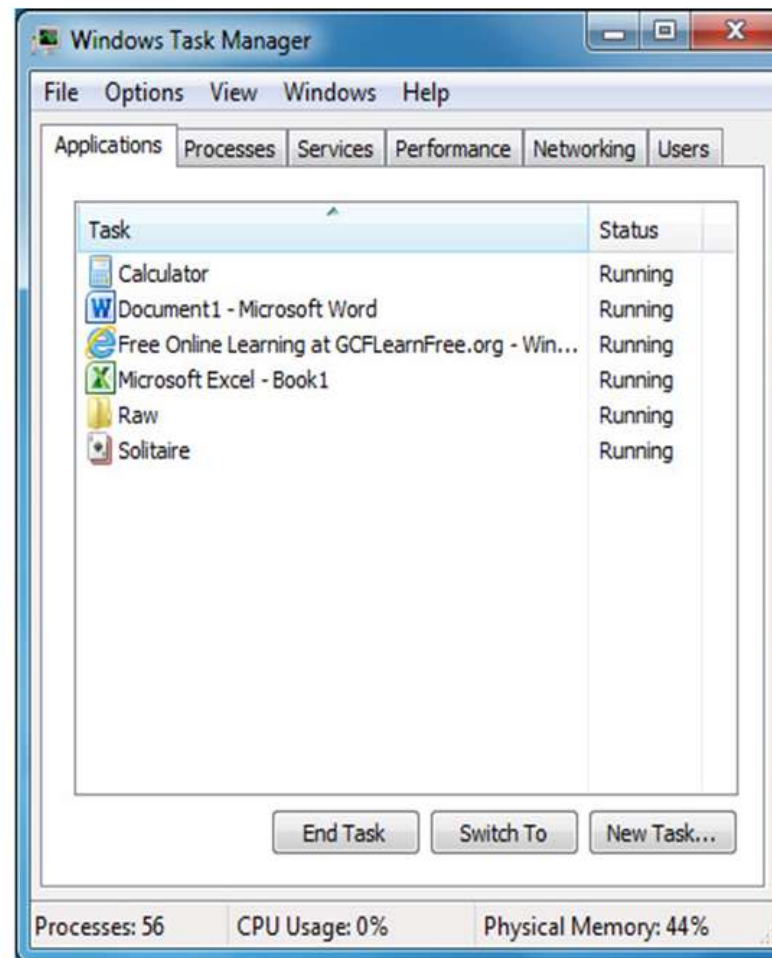
Figure Windows Task Manager

Problems starting or shutting down the computer (1 of 2)

- If the computer does not start, start by checking the power cord to confirm that it is plugged into power outlet.

- If it is plugged into an outlet, confirm it is a working outlet.

- If the computer is plugged in to a surge protector, confirm that it is turned on.

- If the user is using a laptop, the battery may not be charged. Try to turn on the laptop by plugging the AC adapter into the wall.

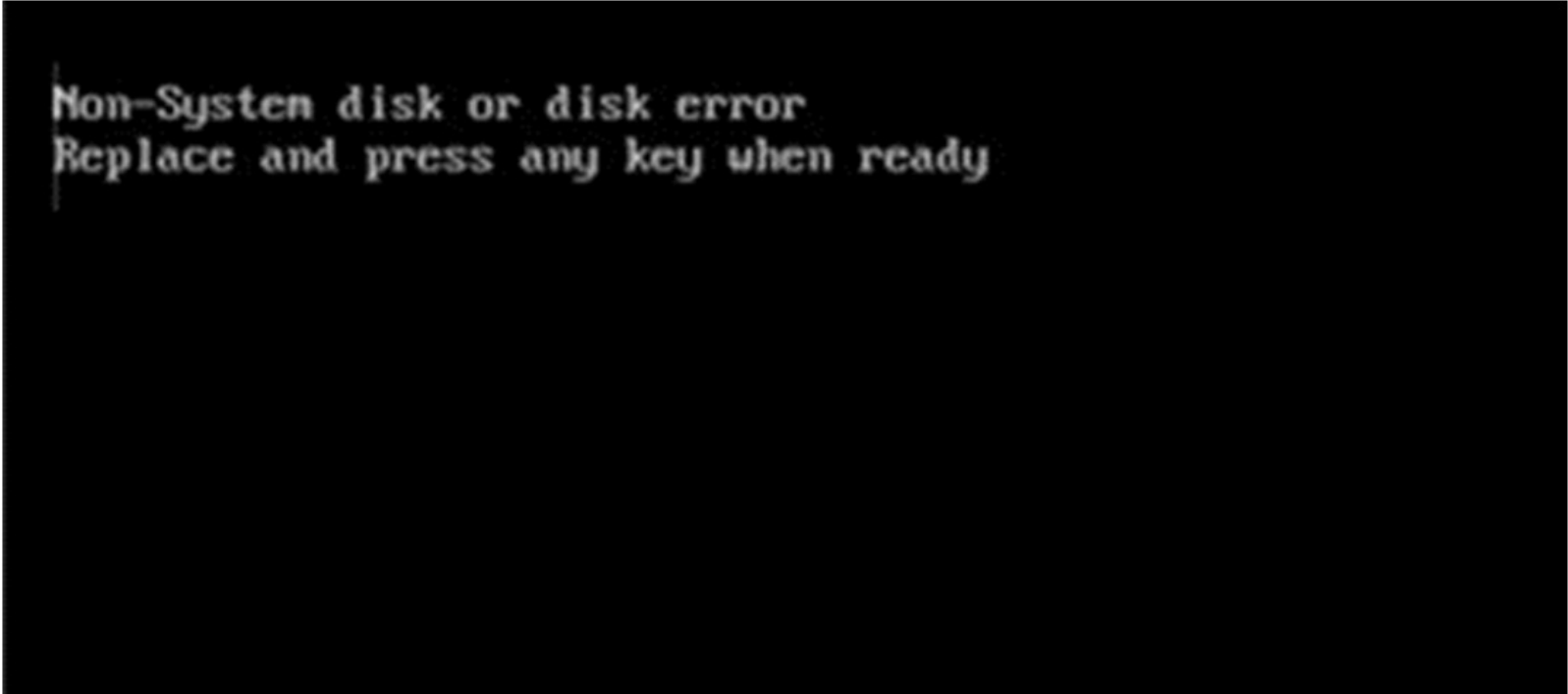# Problems starting or shutting down the computer (2 of 2)

"Non-system disk or disk error" Message

- If the user gets this message when he boots up the system, it usually means there is a DVD, CD, floppy disk, or USB flash drive in the system, which is interfering with the system's booting process. Take out the disk from the drive and then restart the computer.

Solving more difficult problems

- If the user still has not found a solution to the problem, he may need to ask someone else for help.

Configuration Management - Principles

- Configuration management supports the management of services by providing information about how the services are being provided.

- This information is crucial to the other service management processes, particularly such processes as incident management, problem management or change management.

- This is also crucial to ensure meeting all agreed-to service levels.

Configuration Management and the Cloud

- The need for good configuration management practices does not end when services (or parts of services) are moved to the cloud.

- Still, it is the service provider's responsibility to ensure that services are being delivered as agreed to with the clients.

- Different cloud-based services, like IaaS, PaaS, or SaaS, would require different levels of configuration management.
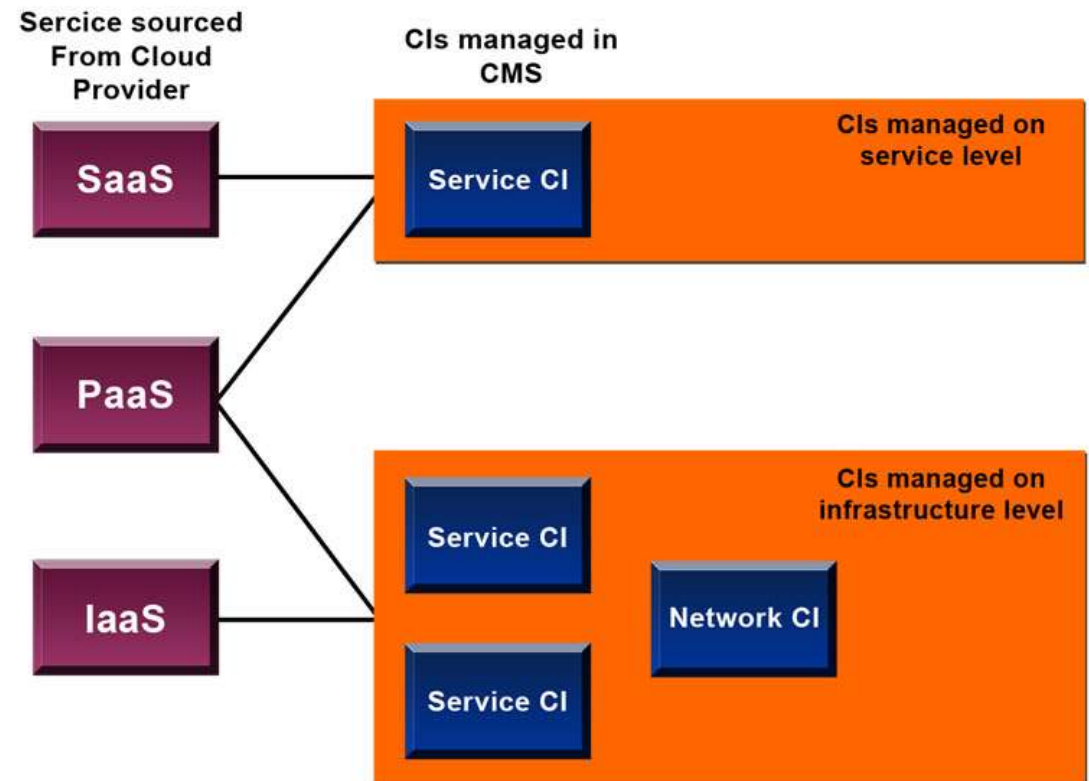

Figure Configuration Management and the Cloud

- Chef is an infrastructure automation structure that makes it easy to configure, setup, manage, and deploy servers and applications to any environment physical / virtual / cloud).

- With Chef, the user can code the infrastructure (called 'recipes') and use the recipes to setup the infrastructure.

- Once automated, the user holds a blueprint for the setup, enabling to build (or rebuild) automatically in minutes or hours – not weeks or months. Still, in case of disasters (hardware, network, or geographical) Chef makes disaster recovery job easier.

- Chef has become one of the widely used tools for configuration management. Besides Chef, tools supporting cloud environments are Salt, Puppet, Ansible.

- AWS OpsWorks is an application management service that makes it easy for DevOps to model and manage the entire application from load balancers to databases. Chef is supported by Amazon OpsWorks.

With Chef, the user will be able to:

- Manage servers by writing recipes.

- Incorporate tightly with applications, databases and more.

- Design applications that require knowledge about your entire infrastructure

- Construct perfect clones of QA environments, partner preview environments, pre-production environments, and more.

Recipe	Configuration element within an enterprise. Recipes are used to install, configure software and deploy applications
Cookbook	Fundamental unit of configuration and policy distribution. Every cookbook defines a setup, such as everything needed to install and configure MySQL.
Knife	Knife is a command-line tool that provides an interface between the Chef server and a local chef-repo. Knife helps provisioning resources, handle nodes, recipes/cookbooks & more.
Chef-repo	Chef-repo is located on the workstation and contains roles, recipes, cookbooks. Knife is used to upload data to the chef server from the chef-repo.
Workstation	A workstation is a computer that is configured to run Knife, in order to synchronize with the chef-repo, to interact with a single server. It is the location from which most users will do most of their work.
Node	A node is any virtual, physical, or cloud machine that is configured to be maintained by a chef-client
Run_list	A run_list is an ordered list of roles and/or recipes that are run in an exact order.
Chef-client	An agent that runs locally on every node is called chef-client.

The Chef Server

3 types of Chef Servers

- Hosted Chef

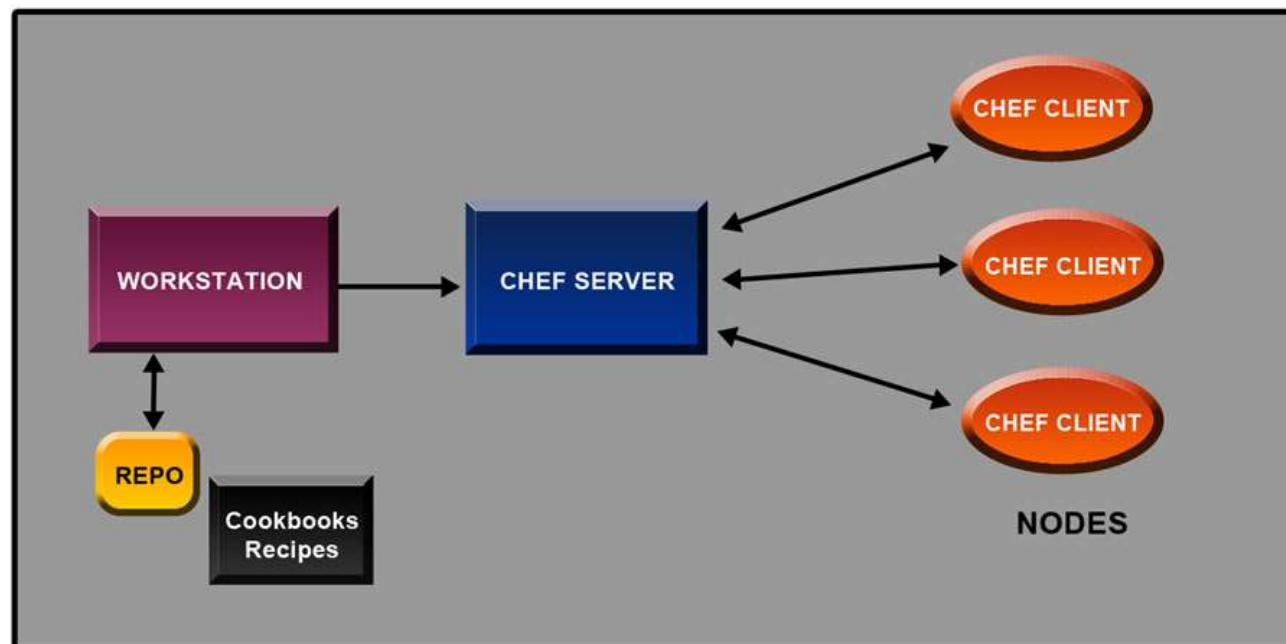- Enterprise Chef

- Open Source Chef

Figure The Chef Server