RESOURCE MONITORING AND MANAGEMENT IN CLOUDS

CSCI 5408: Data Management, Warehousing and Analytics

Prepared by: Suhaib Qaiser (suhaibqaiser@dal.ca)

Recap from last lecture…

Quick Review: Hypervisor

Quick Review: Load Balancers

Load Balancers Algorithms

Managed versus Unmanaged Services

SLA Management Systems

Billing Management Systems

Scenario: Speech Recognition LMS App

Resource management and monitoring

Resource Management Systems

Version Control System

Distributed Source Control

- *Version Control Systems*
- *Cloud Data Analytics*
- *Sales force Cloud analytic summary*
- *Right Scale Cloud analytic summary*
- *MS Azure Cloud analytic summary*
- *Application Metrics?*
- *Quiz*
- *Reading Material*
- *Any Question*

Recap from last lecture…

Q.1 What is the difference between Server Consolidation and Hardware Independence?

Q.2 What is the difference between Grid Computing and Parallel Computing?

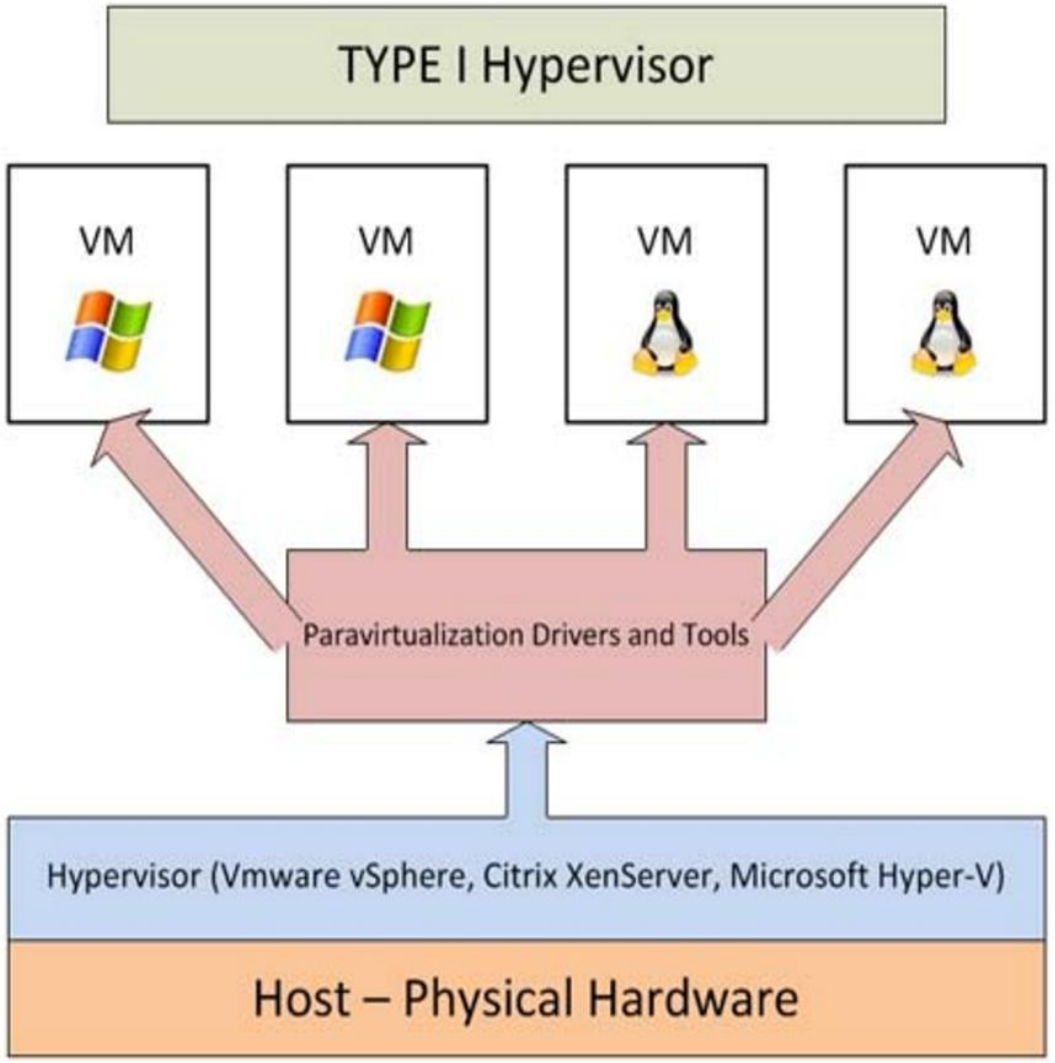
Q.3 What is the difference between resource pooling and resource replication?

Q.4 Why is JavaScript more important in Cloud computing than any other language?

Q.5 If a company server crashed or we get complains from service center, as a production support what should I do first?

Q.6 Are REST services more efficient then SOAP services? If yes than WHY

Quick Review: Hypervisor



The hypervisor architecture is the fundamental part of virtualization infrastructure

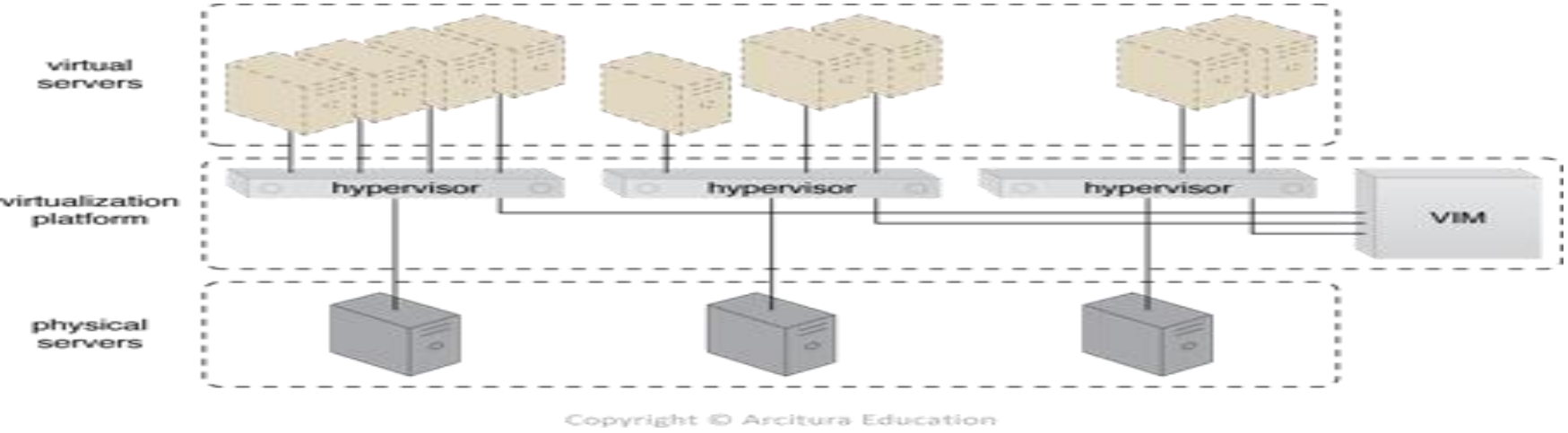
Hypervisor is primarily used to generate virtual server instances on physical servers

A hypervisor is generally limited to one physical server and can therefore only create virtual images of that server

Hypervisor can provide features of controlling, sharing, scheduling of hardware resources. It provides the on-demand provisioning and re-claiming of virtual server images

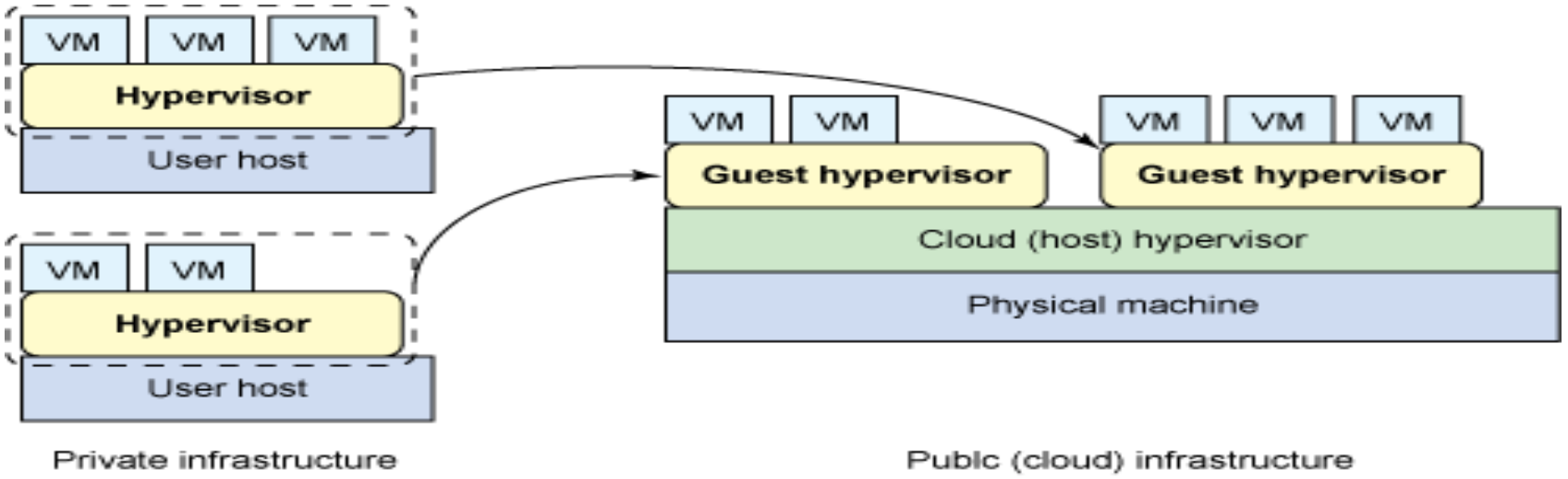
http://stackoverflow.com/questions/38130453/difference-between-bare-metal-hypervisor-based-and-host-virtualization-types

Quick Review: Hypervisor



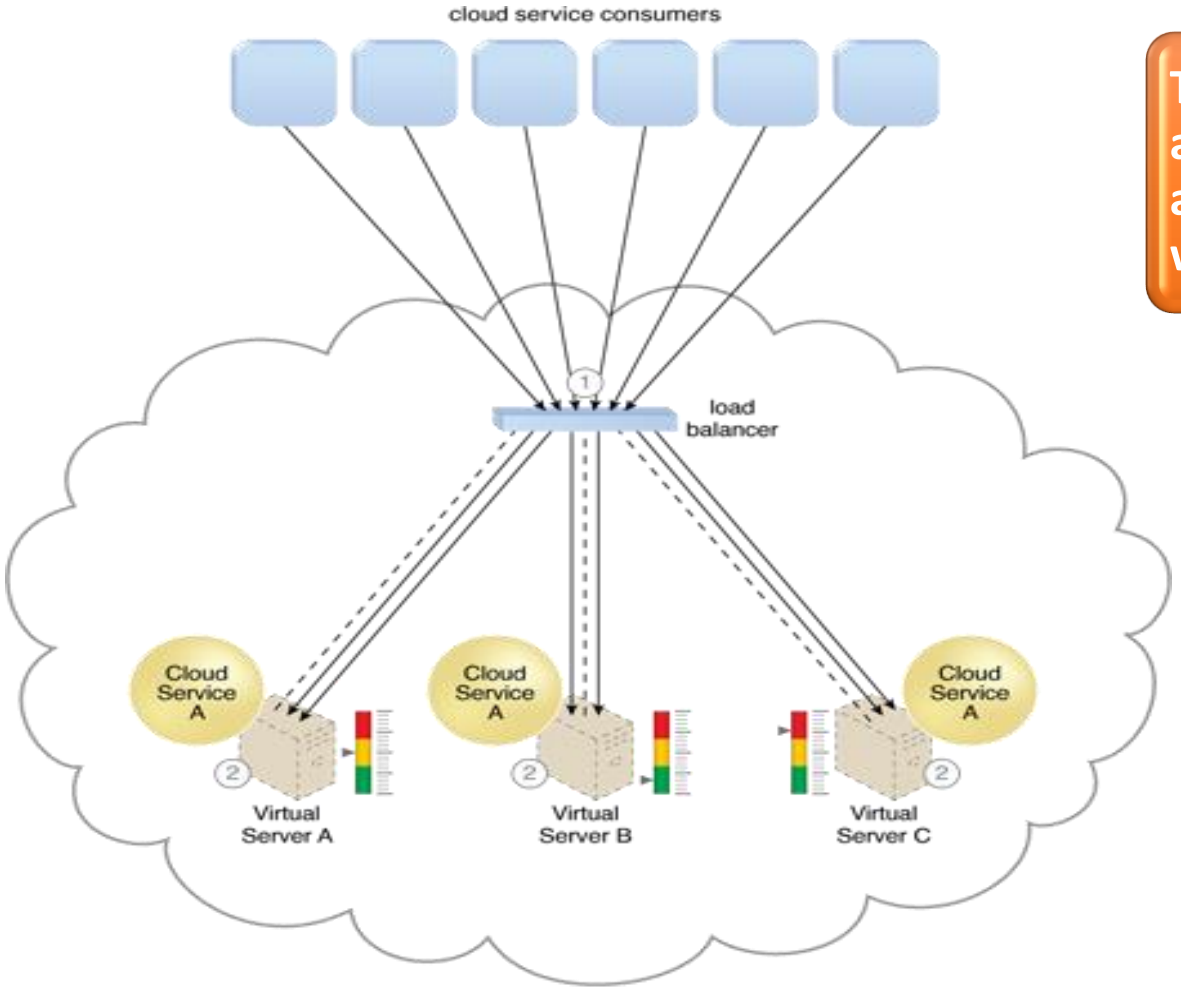
A **hypervisor** or virtual machine monitor (VMM) is computer software, firmware, or hardware, that creates and runs virtual machines.

A computer on which a **hypervisor** runs one or more virtual machines is called a host machine, and each virtual machine is called a guest machine.



http://is.muni.cz/th/396555/fi_b/BP_Vadym_Yanovskyy.pdf

Quick Review: Load Balancers



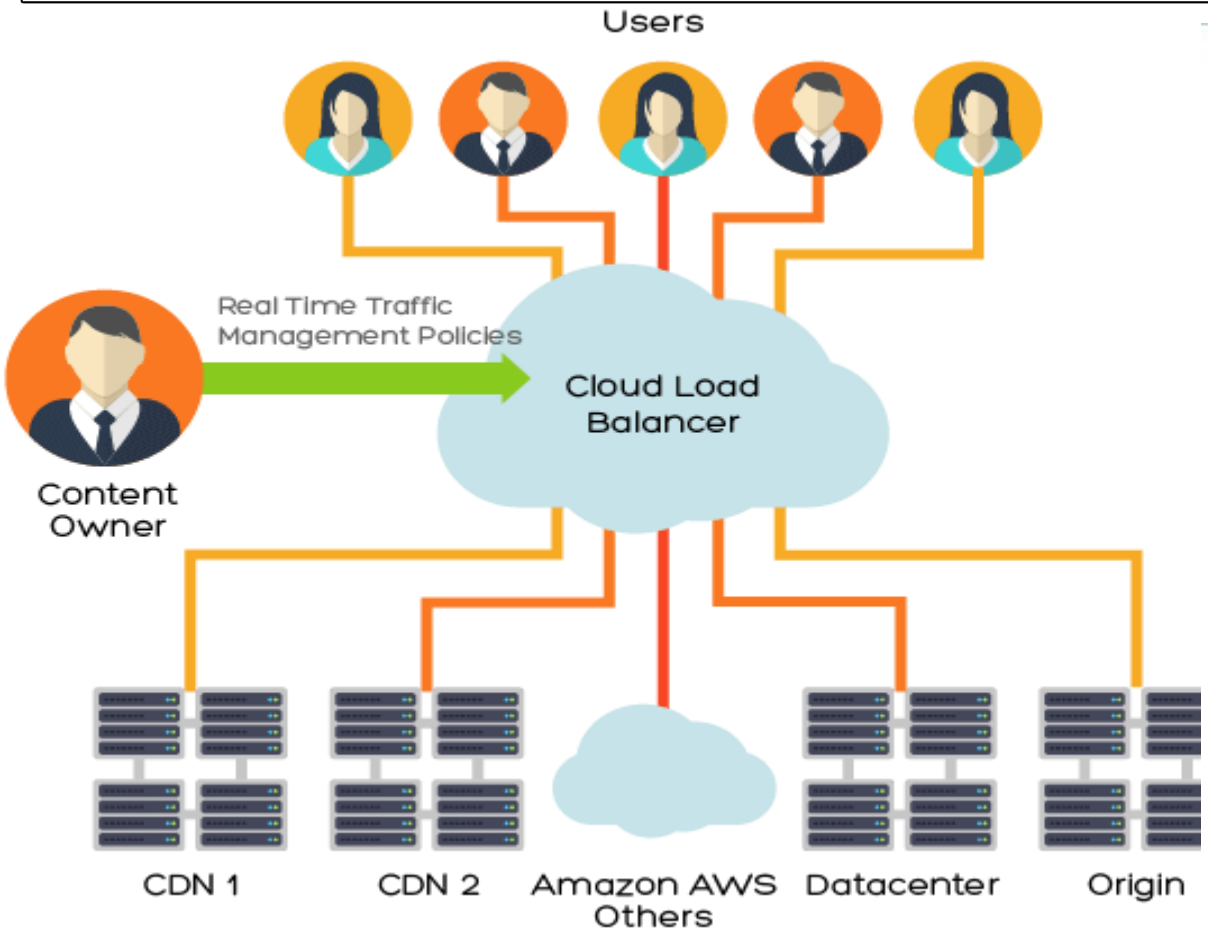
The service load balancer architecture can be considered a specialized variation of workload distribution

Load balancers can be used to distribute load on IT resources that are scaled horizontally

Duplicate cloud servers are placed in a resource pool and load balancer is positioned either external or built in component

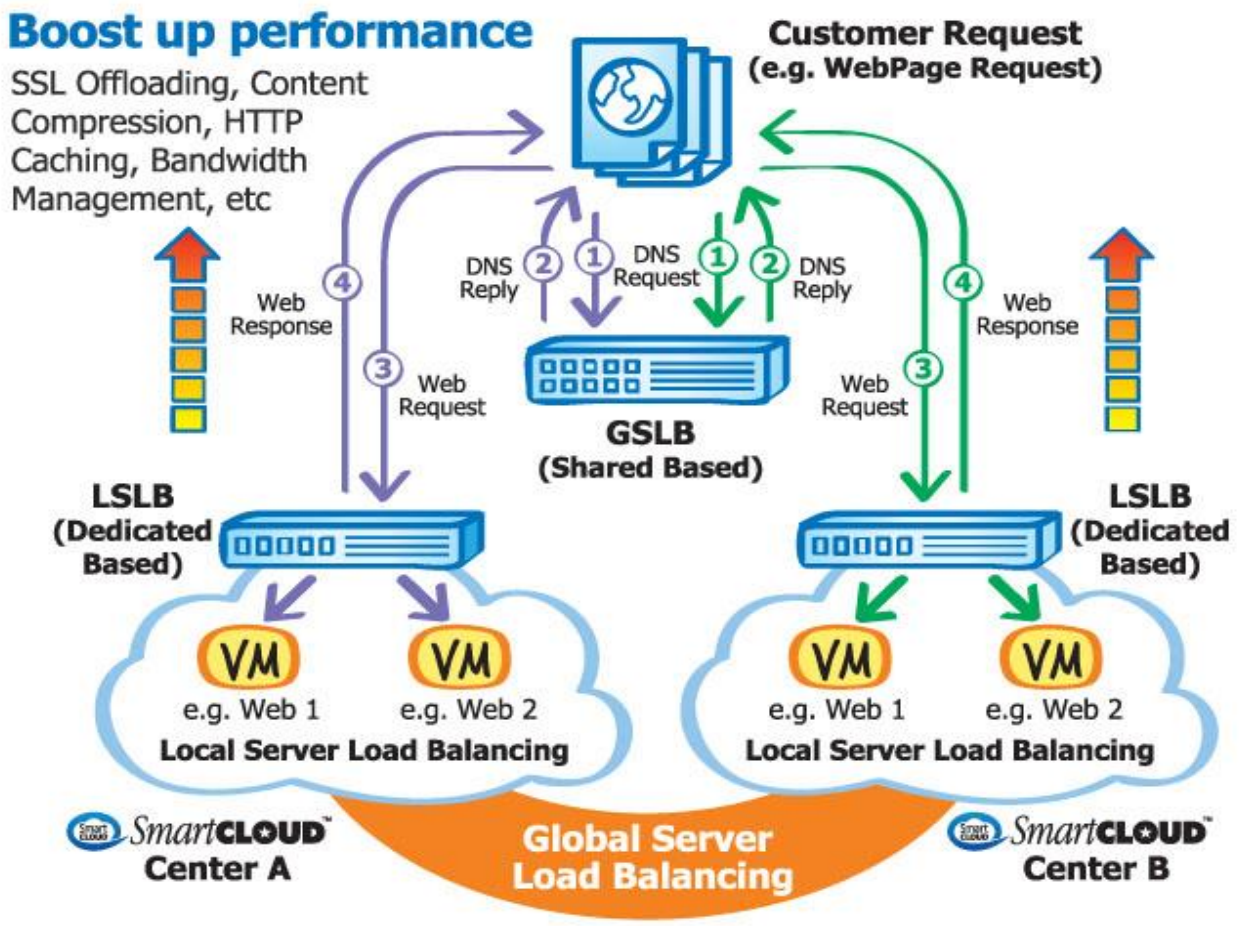
Source: http://ptgmedia.pearsoncmg.com/images/chap11_9780133387520/elementLinks/th11fig10.jpg

Load Balancers



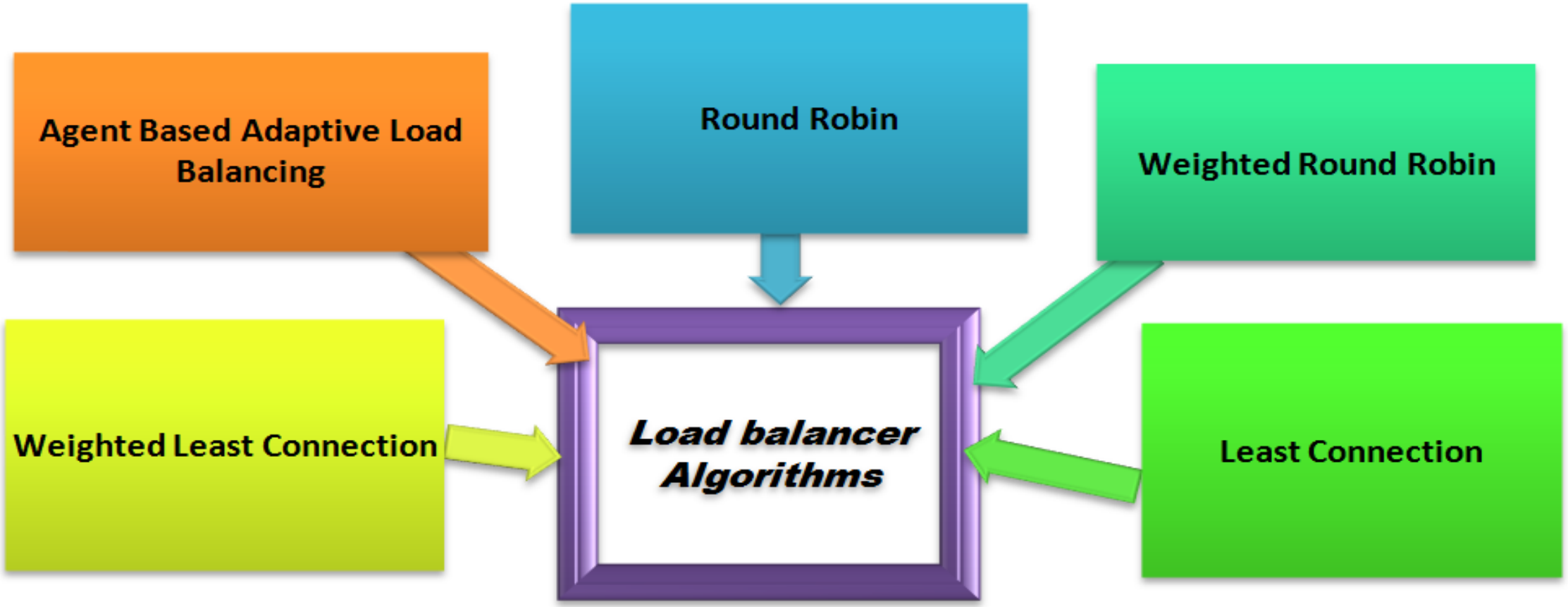
https://www.cdnetworks.com/products/cloud-load-balancer/

Load Balancers



https://www.citictel-cpc.com/EN/HK/Pages/product-services/smartcloud-tab

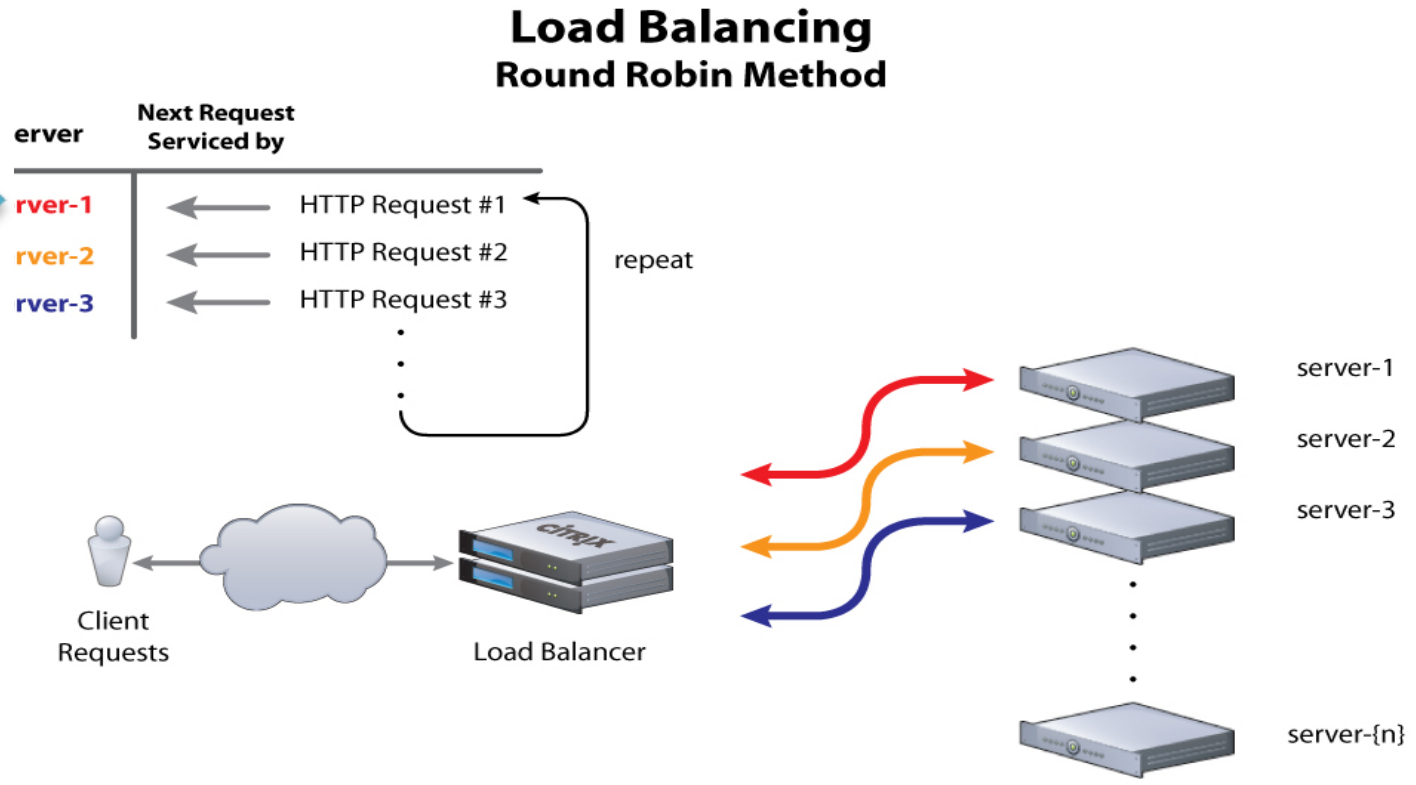
Load Balancers Algorithms



Source: https://kemptechnologies.com/load-balancer/load-balancing-algorithms-techniques/

Load Balancers Algorithms

Round Robin



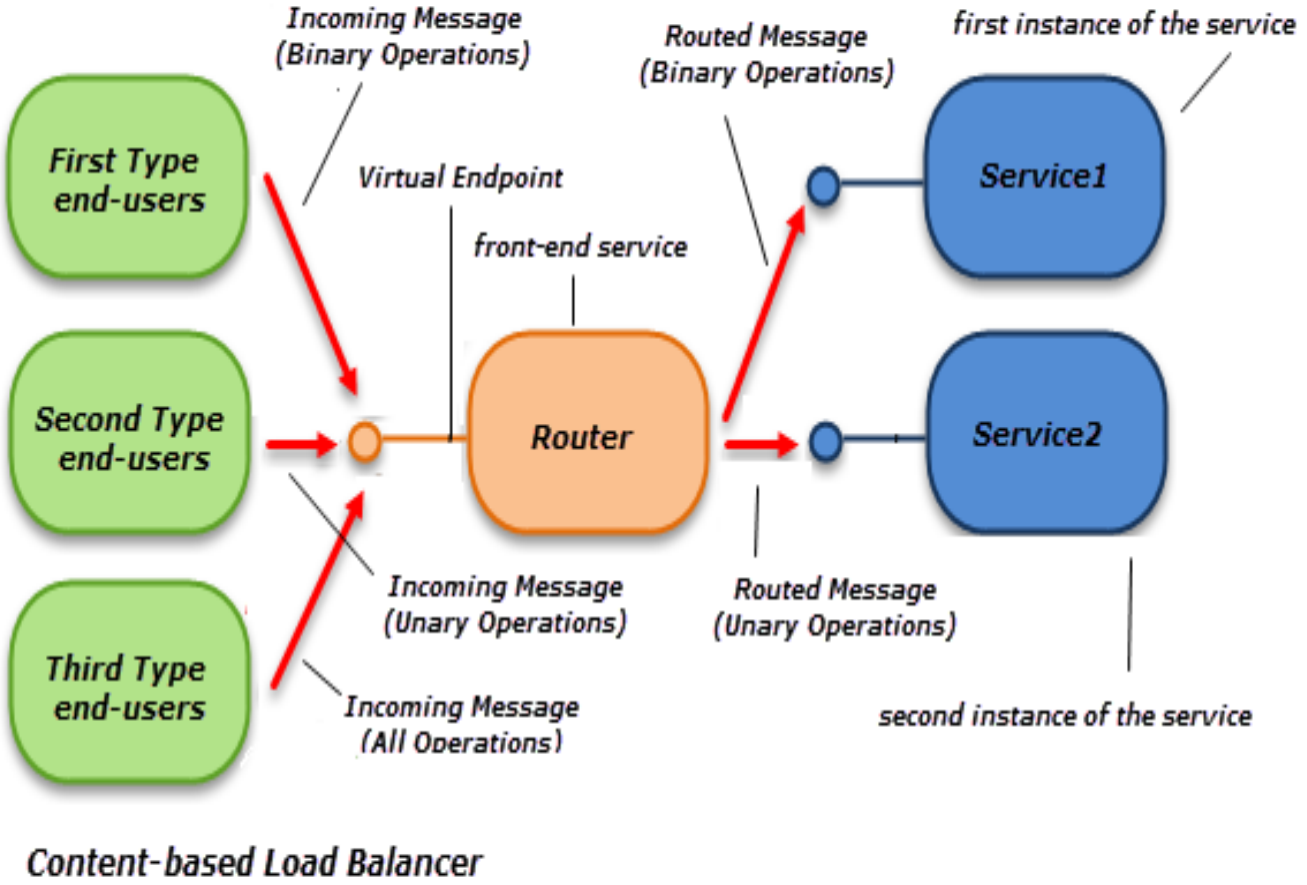
A simple method
of load balancing servers
or
for providing simple
fault tolerance

https://www.citrix.com/blogs/2010/09/03/load-balancing-round-robin/

Load Balancers Algorithms

Weighted Round Robin

Each server in the pool is given a static numerical weighting. Servers with higher ratings get more requests sent to them



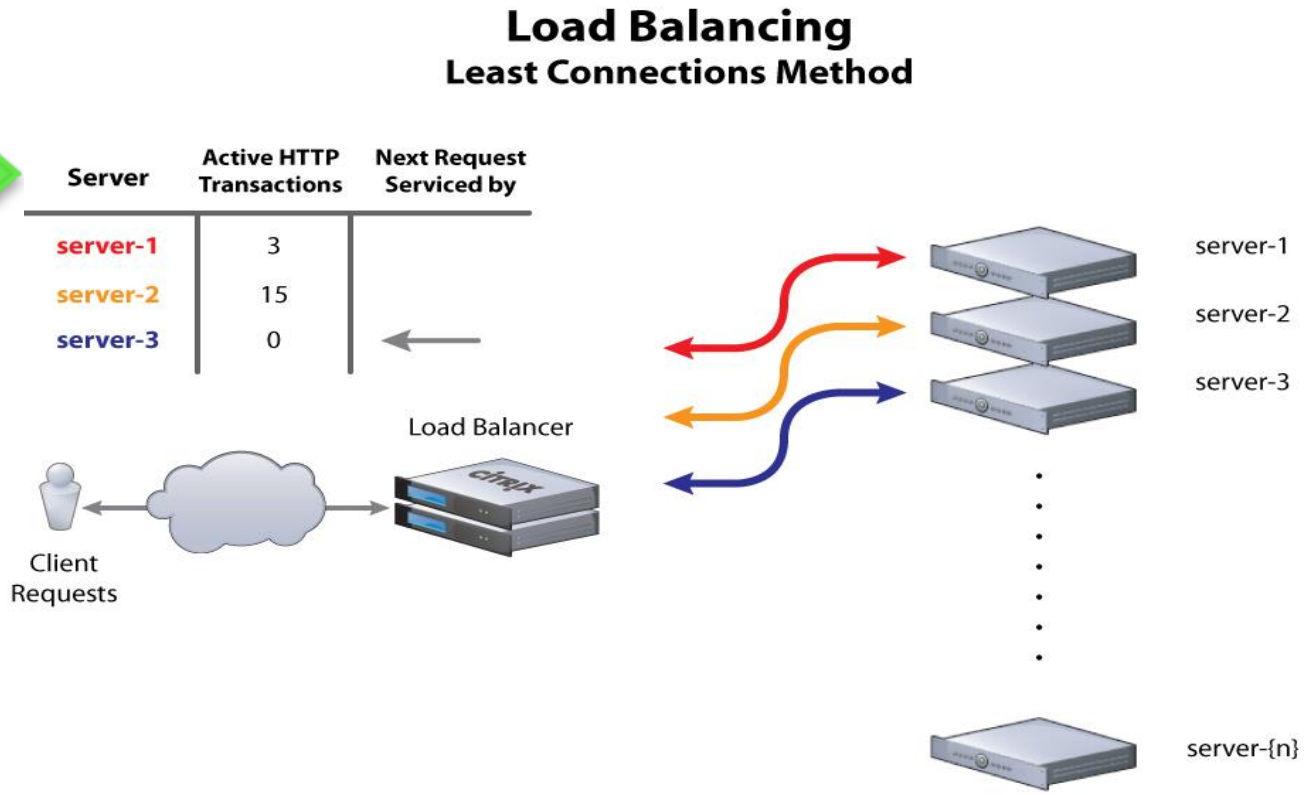
Content-based Load Balancer

https://www.codeproject.com/Articles/778575/WCF-Routing-Service-Part-III-Failover-Load-Balanci

Load Balancers Algorithms

Least Connection

The current request goes to the server that is servicing the least number of active sessions at the current time

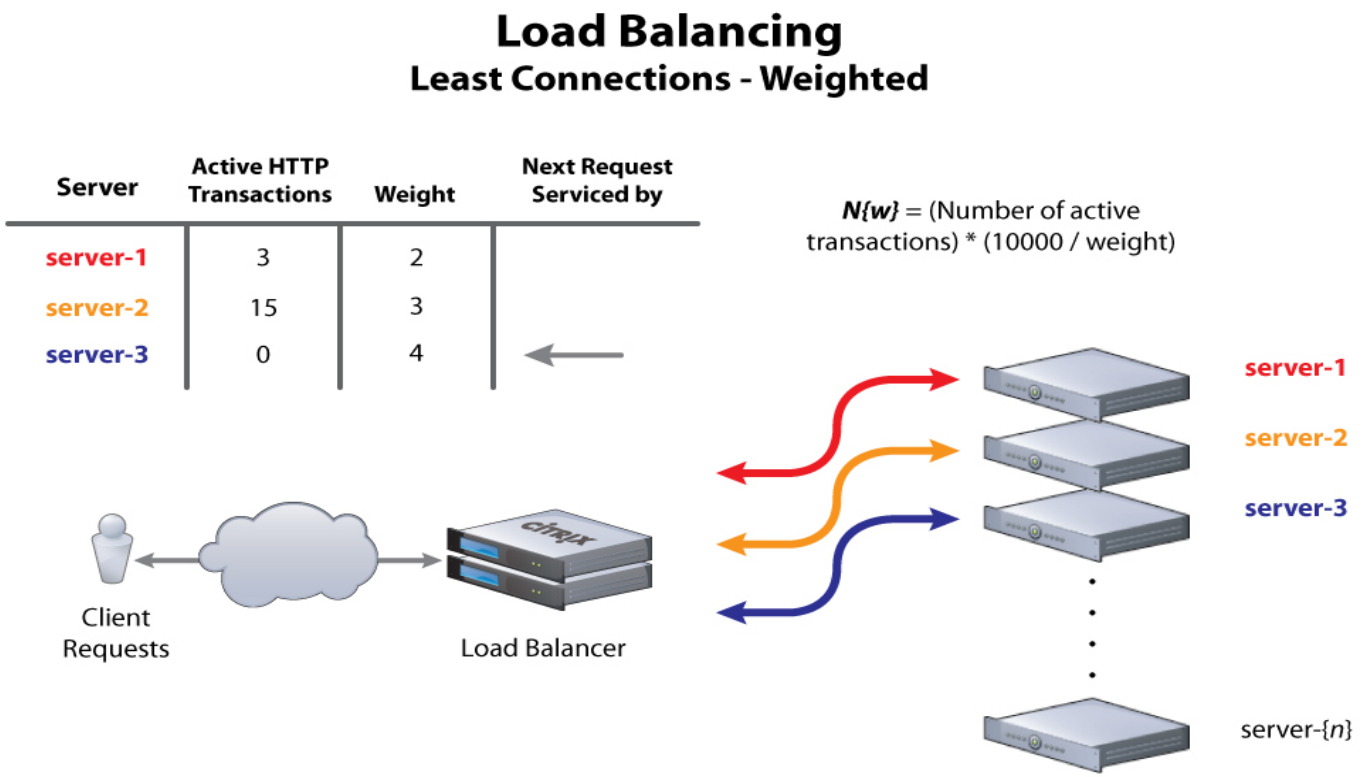


https://www.citrix.com/blogs/2010/09/02/load-balancing-least-connections/

Load Balancers Algorithms

Weighted Least Connection

Builds on the Least Connection method. If two servers have the same number of active connections then the server with the higher weighting will be allocated the new request

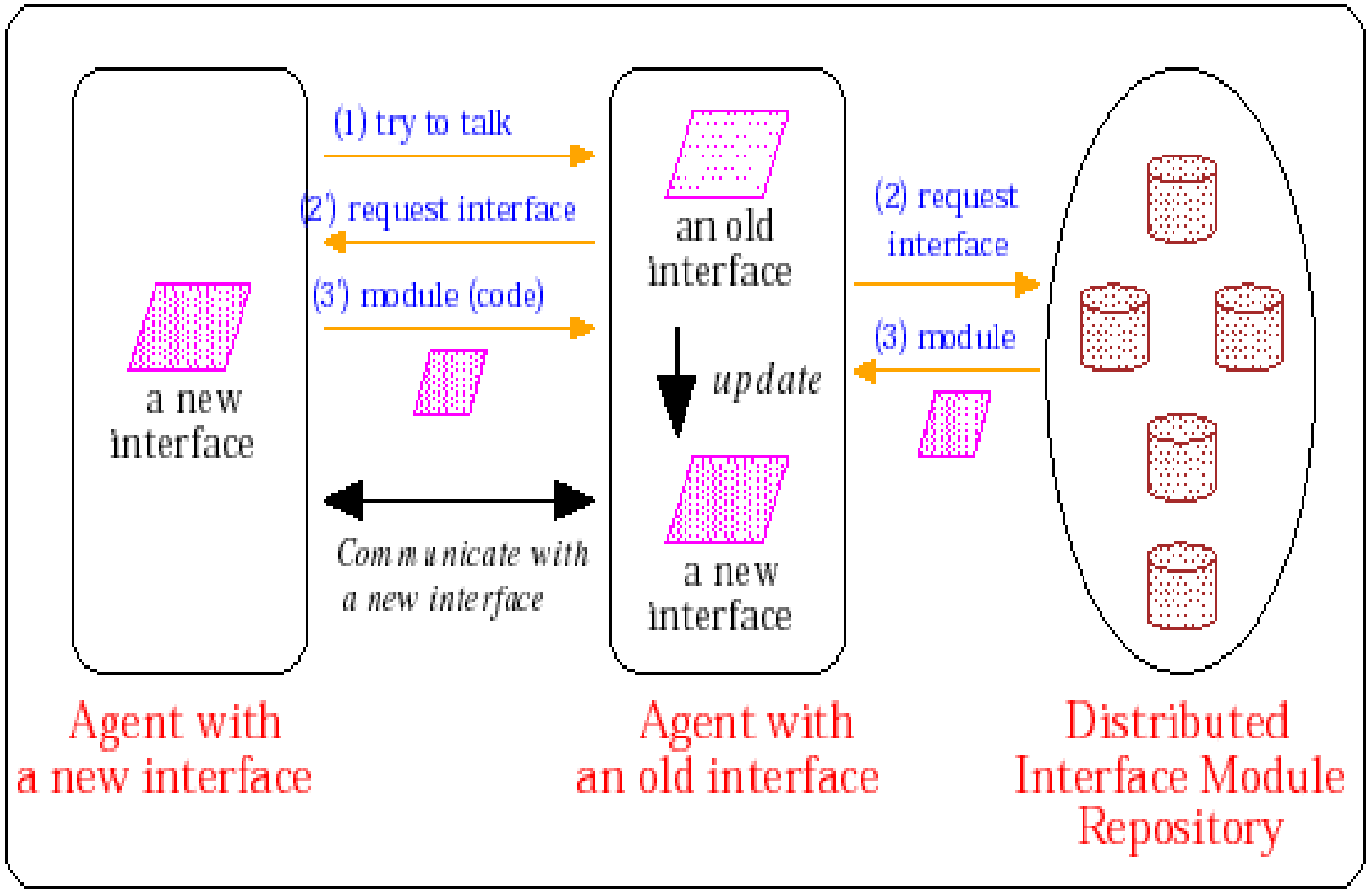


https://www.citrix.com/blogs/2010/10/01/load-balancing-weights/

Load Balancers Algorithms

Agent Based Adaptive Load Balancing

Each server in the pool has an agent that reports on its current load to the load balancer. This real time information is used when deciding which server is best placed to handle a request

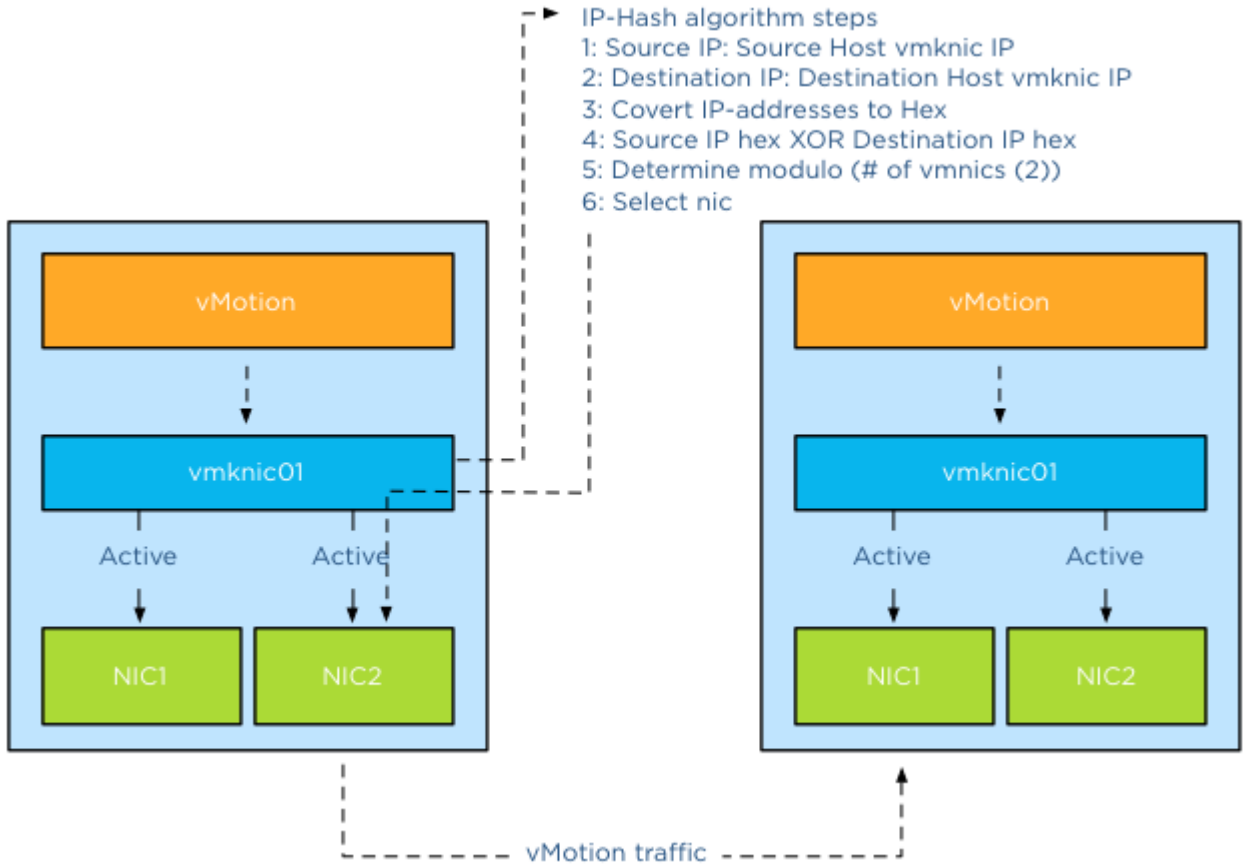


https://www.isoc.org/inet2000/cdproceedings/2f/2f_1.htm

Load Balancers Algorithms cont.

Source IP Hash

Source IP Hash load balancing uses an algorithm that takes the source and destination IP address of the client and server and combines them to generate a unique hash key. This key is used to allocate the client to a particular server



https://www.citrix.com/blogs/2010/09/04/load-balancing-hash-method/

Managed versus Unmanaged Services

Managed servers has an extra layer of administration pre built and provided by cloud service provider. That layer performs load balancing, resource replication, failover monitoring, logging and other operations as indicated in Service Level Agreements and service contracts

Unmanaged servers are virtual servers provisioned for cloud consumers for higher level of control into the system. They do not contain an extra layer of administration. Although, consumers are allowed to add their own configuration

SLA Management Systems

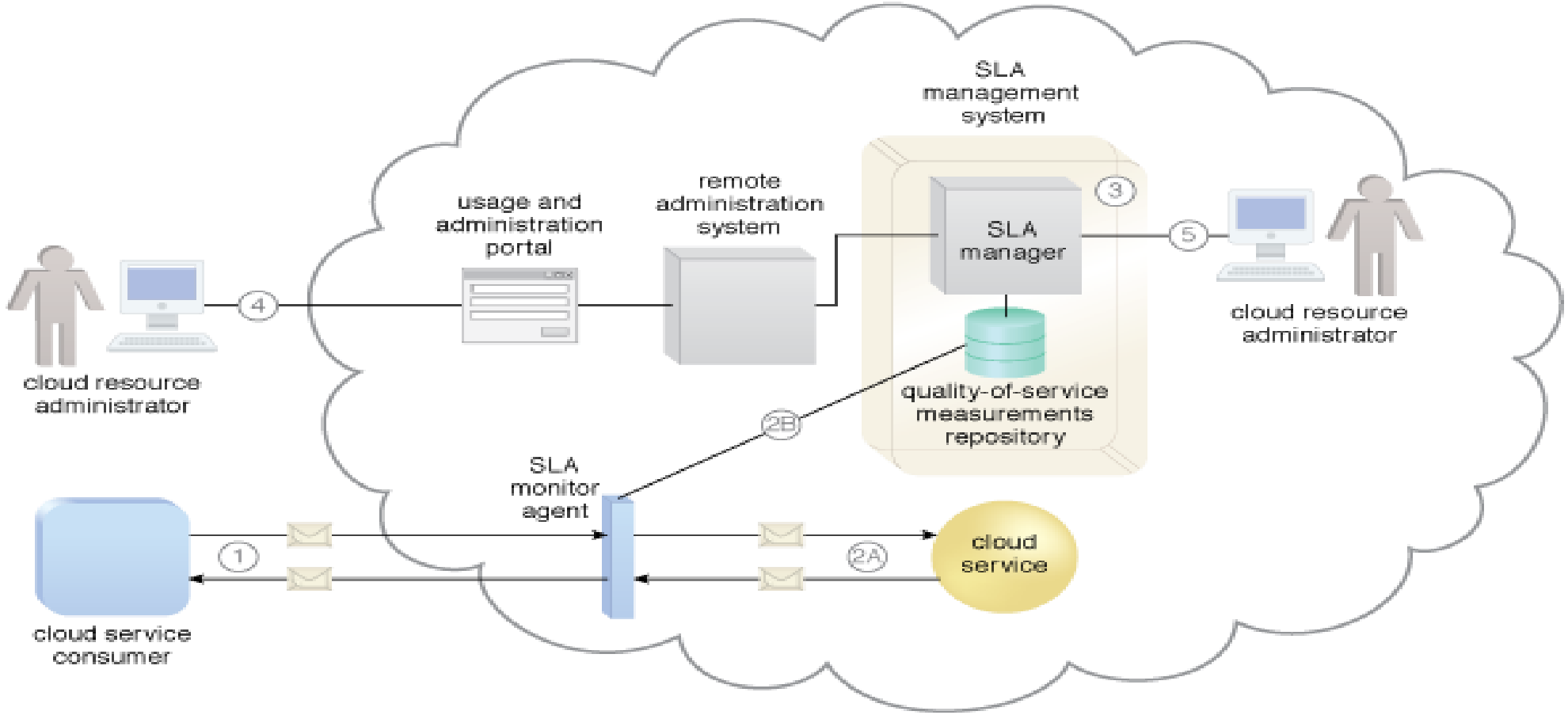
SLA Management systems use repositories to collect analytical information to report on pre-defined metrics and reporting parameters

SLA Management systems represents a range of commercially available cloud management products

SLA Management provides administration, collection, storage, reporting and runtime notification of SLA data

The SLA monitor mechanism is used to specifically observe the runtime performance of cloud services to ensure that they are fulfilling the contractual QoS requirements

SLA Management Systems



Source: http://cloudpatterns.org/mechanisms/sla_management_system

Billing Management Systems

The billing management system mechanism is dedicated to the collection and processing of usage data to calculate cloud consumer billing

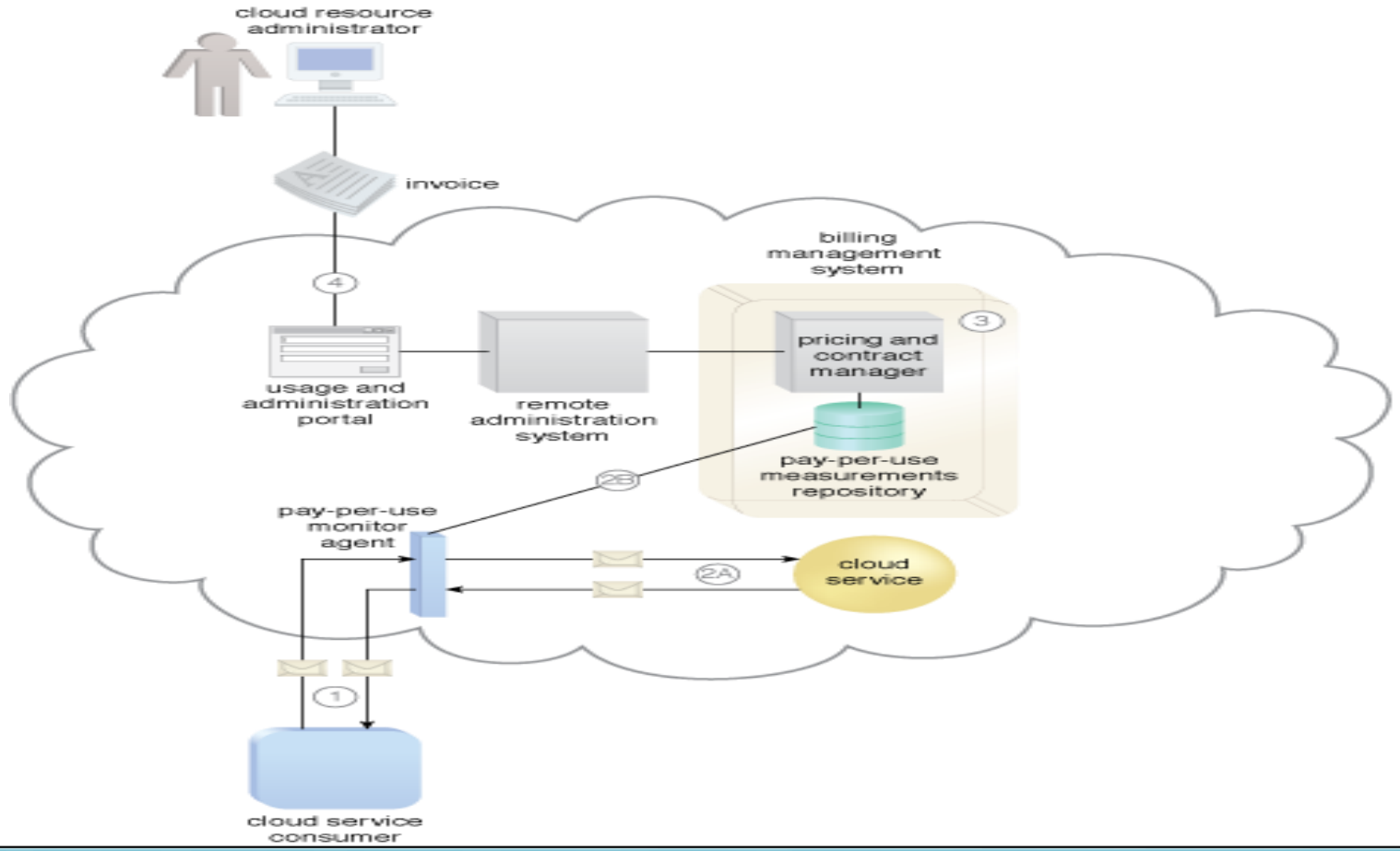
It relies on pay-per-use monitors

Billing system considers different pricing models and accommodate custom user pricing before reporting

When limits are defined they are in the form of usage quotas

The billing system can block further usage requests by cloud consumers

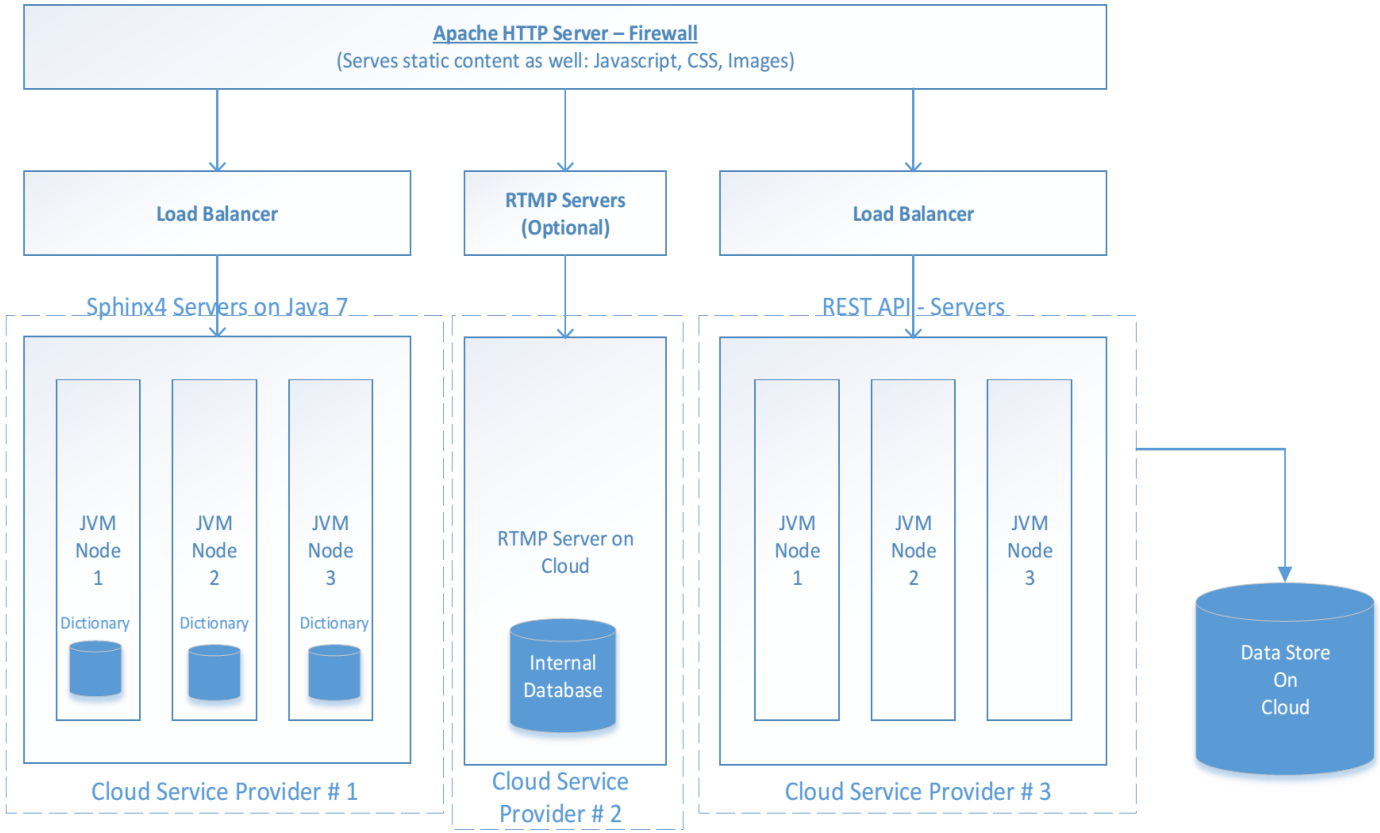
Billing Management Systems Overview



Source: http://cloudpatterns.org/mechanisms/billing_management_system

Scenario: Speech Recognition LMS App

ABC Company wants to shift all the below services to the cloud. Please provide them with an overview of how much cost would be spent on annual basis for the first year. Please provide cost saving alternatives.
How much revenue should company make per request to break even?



Pricing estimate per instance:

- Amazon Elastic Load Balancer: $0.03 per hour
- Amazon EC2: $0.094 per Hour
- Amazon RDS: $0.290 per Hour per 15 GB
- Amazon DynamoDB: $0.0128 per Hour per GB
- Amazon Elastic Beanstalk: Free
- Amazon S3: $0.0314 per GB per month
- Amazon F5 Routing: $0.400 per million / month
- Estimated development and managerial cost:
  - $ 500,000 per year

- **Please note:**
  - Expected traffic 5 million hits / month
  - Expected data size 100 KB / request-response
  - Every month traffic is expected to grow 5%

Resource management and monitoring

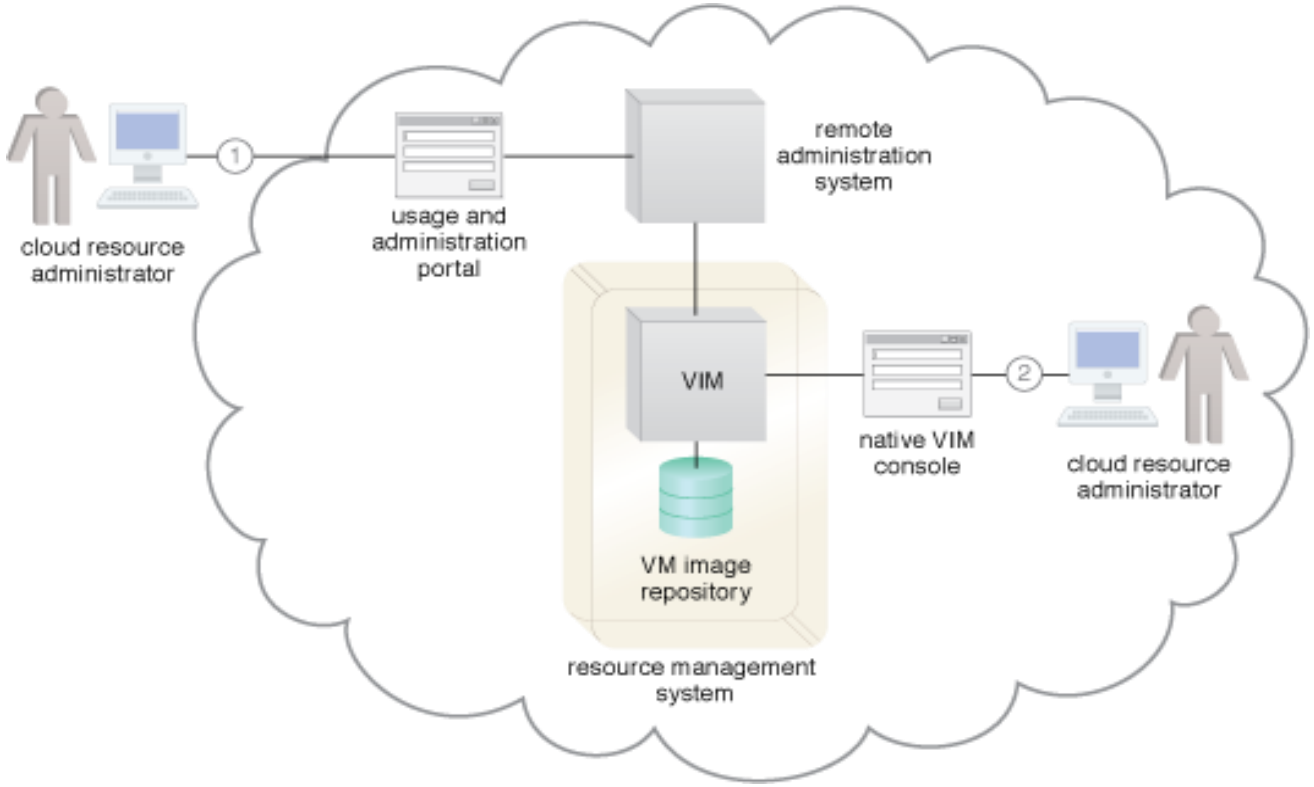
The resource management system mechanism helps coordinate IT resources in response to management actions performed by both cloud consumers and cloud providers

Tasks that are typically automated and implemented through the resource management system include:

managing virtual IT resource templates that are used to create pre-built instances, such as virtual server images
allocating and releasing virtual IT resources into the available physical infrastructure in response to the starting, pausing, resuming, and termination of virtual IT resource instances
coordinating IT resources in relation to the involvement of other mechanisms, such as resource replication, load balancer, and failover system
enforcing usage and security policies throughout the lifecycle of cloud service instances monitoring operational conditions of IT resources

Resource management systems typically expose APIs that allow cloud providers to build remote administration system portals

Resource Management Systems



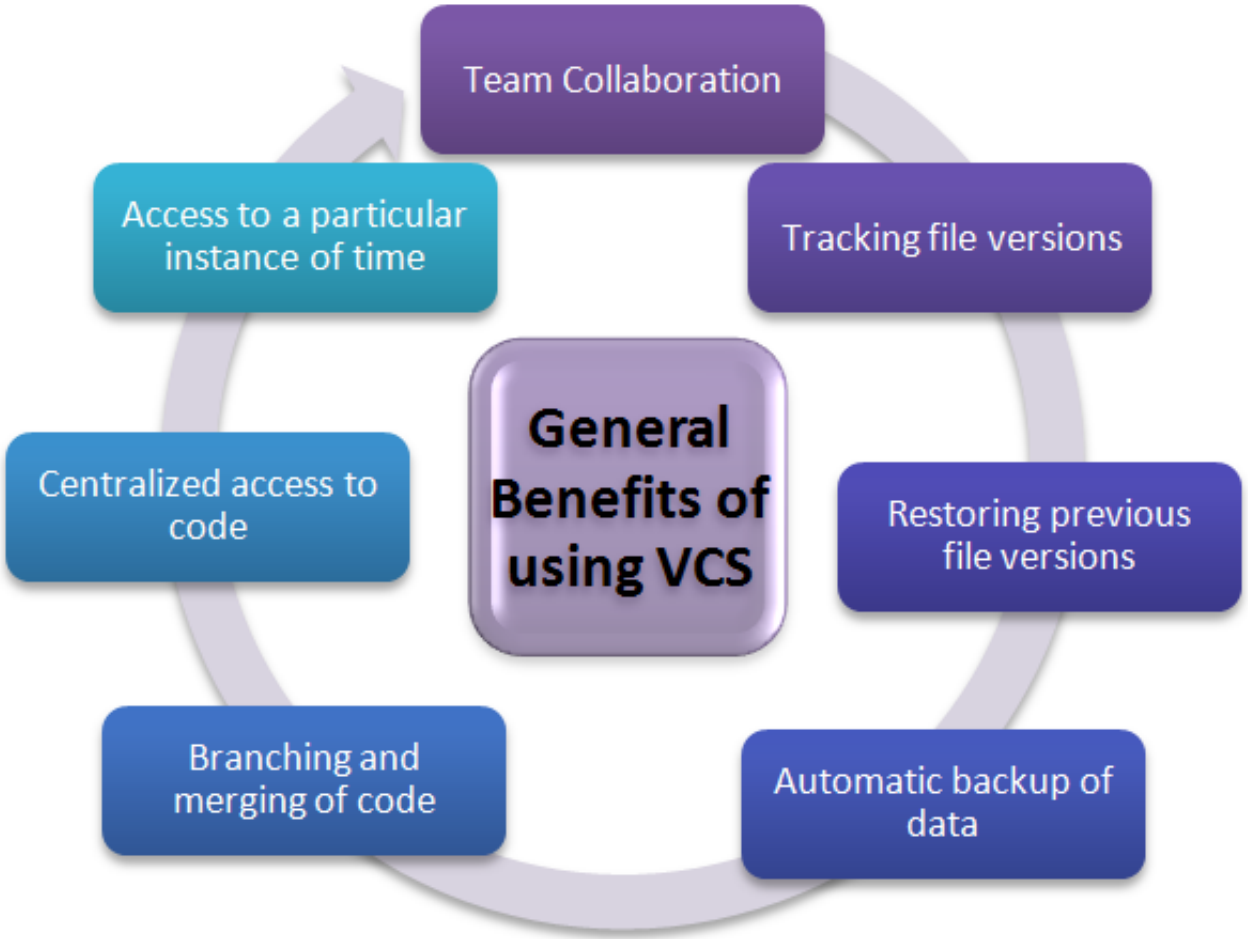
Source: http://cloudpatterns.org/mechanisms/resource_management_system

Version Control System

Version control system helps developers and system administrators to keep track of software changes done over a period of time. It also keep code consistent between multiple versions of a file

Source Control has many benefits. In Cloud computing it is important because it provides an online centralized access to a repository of code to be deployed

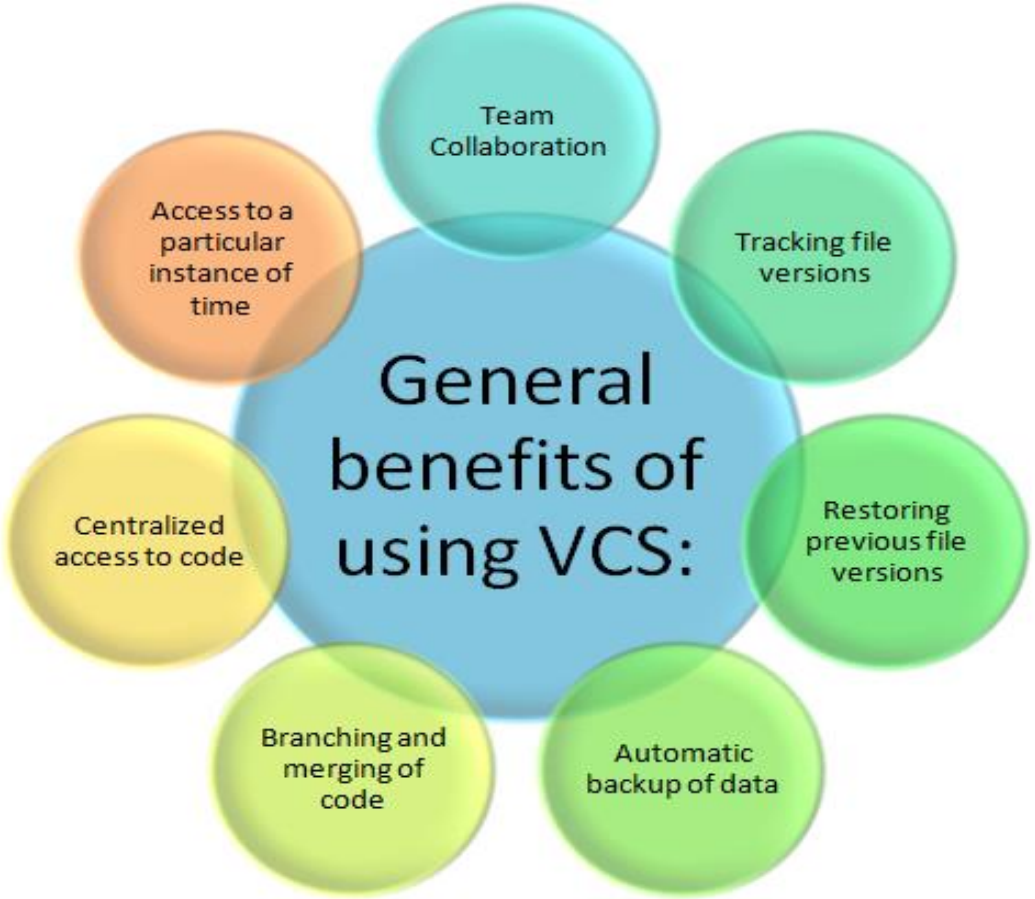
Some popular Source Control Systems are: CVS, Subversion, Git, Mercurial, IBM RTC, Bazaar, etc



Distributed Source Control

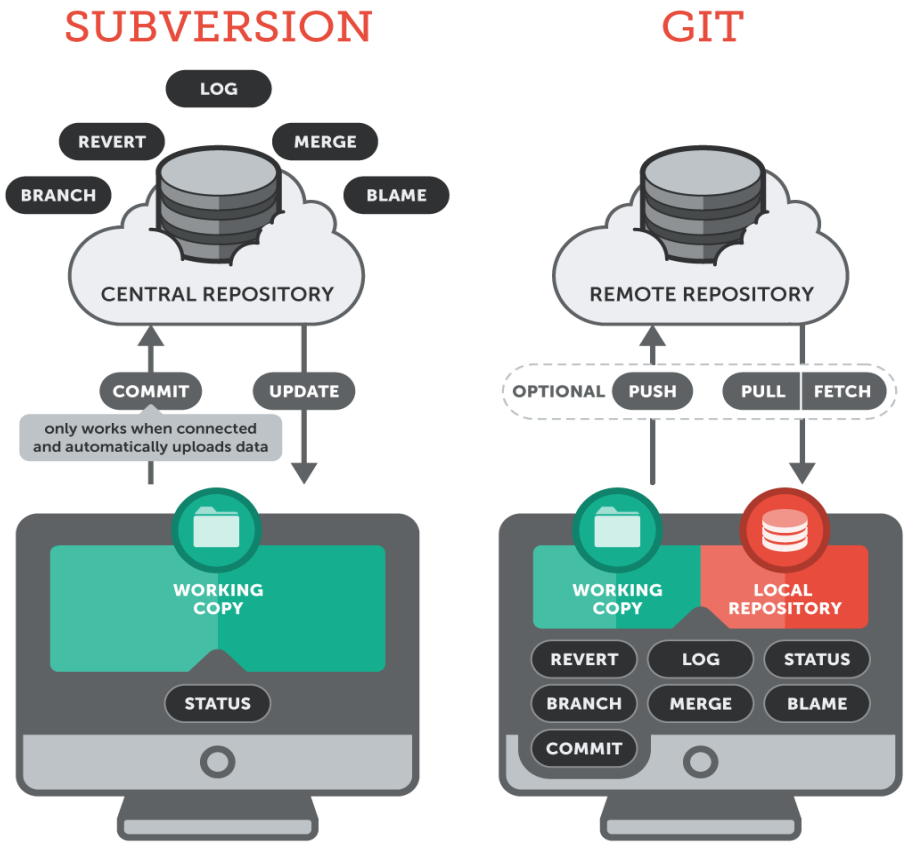
Distributed revision control takes a peer-to-peer approach to version control, as opposed to the client-server approach of centralized systems. Rather than a single, central repository on which clients synchronize, each peer's working copy of the codebase is a complete repository.

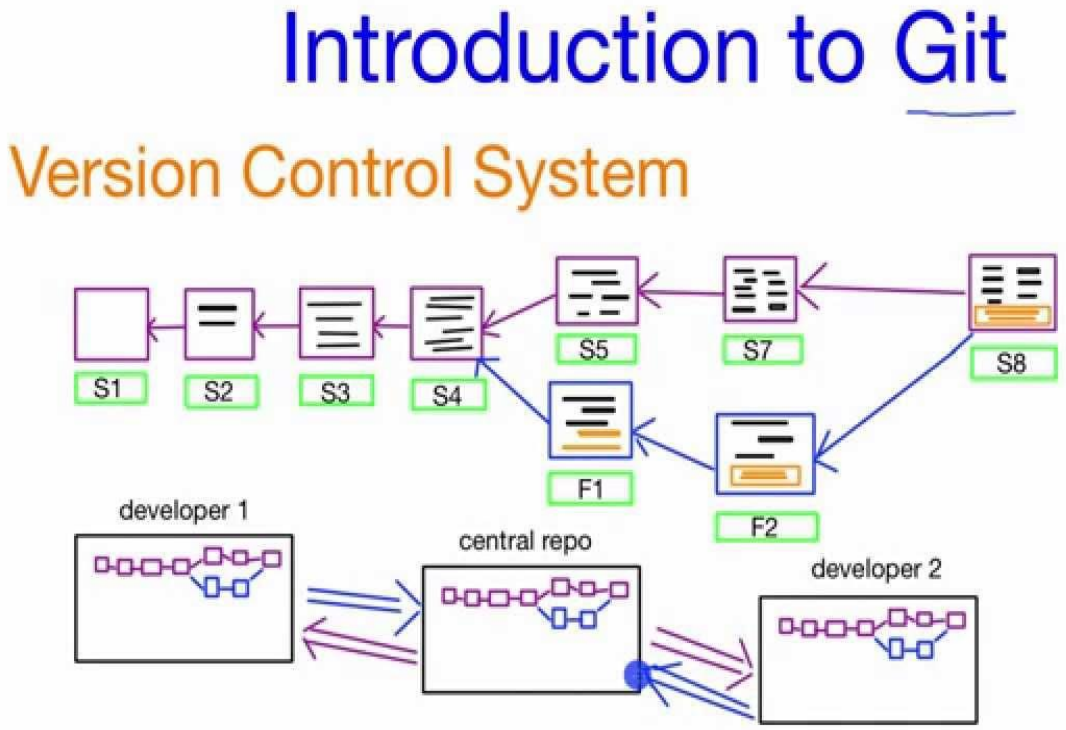
Distributed revision control synchronizes repositories by exchanging patches (sets of changes) from peer to peer. This results in some important differences from a centralized system.



Source: https://en.wikipedia.org/wiki/Distributed_version_control

Version Control Systems

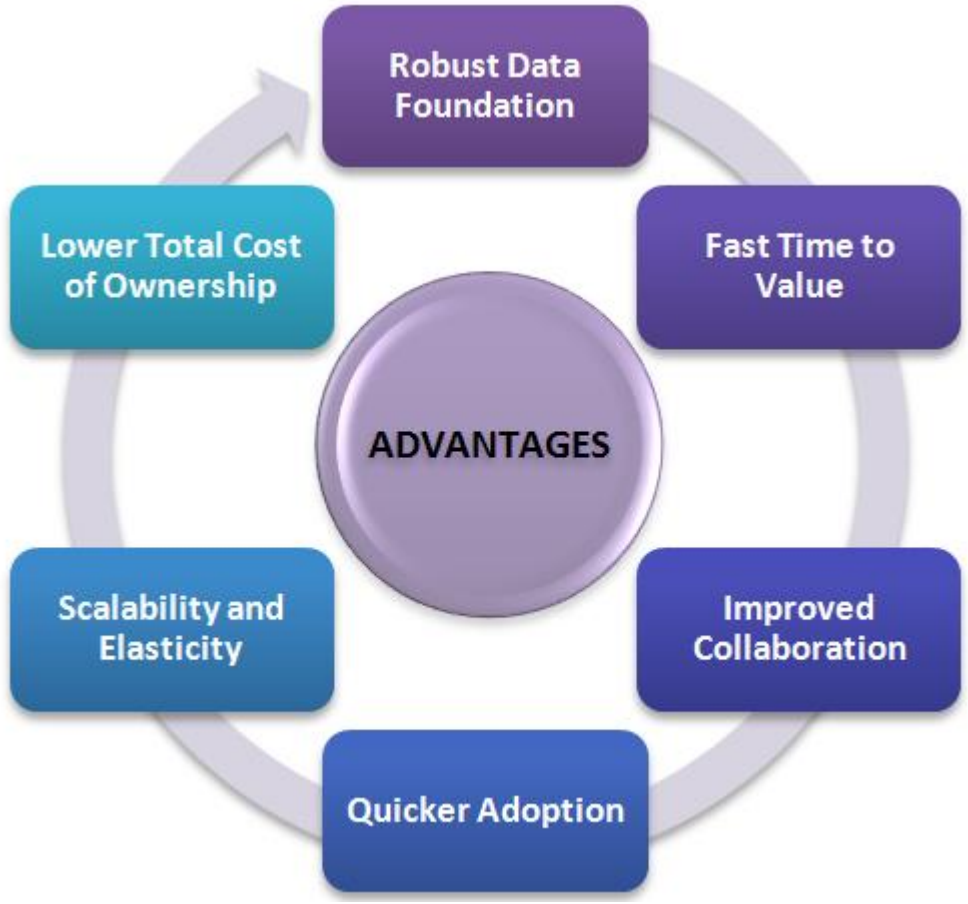




Source: https://www.git-tower.com/learn/content/01-git/01-ebook/en/02-desktop-gui/07-appendix/02-from-subversion-to-git/centralized-vs-distributed.png

Source: https://i.ytimg.com/vi/OqmSzXDrJBk/maxresdefault.jpg

Cloud Data Analytics



Cloud computing allows organizations to consolidate data from all sources, across all communication channels, and do it at a big data scale

Today, cloud and mobile technologies are providing enterprises of all sizes with opportunities to use big data and analytics to make better, data-driven decisions.

New-generation platforms (cloud, big data, analytics) bring analytics and operational applications together to deliver demonstrable ROI

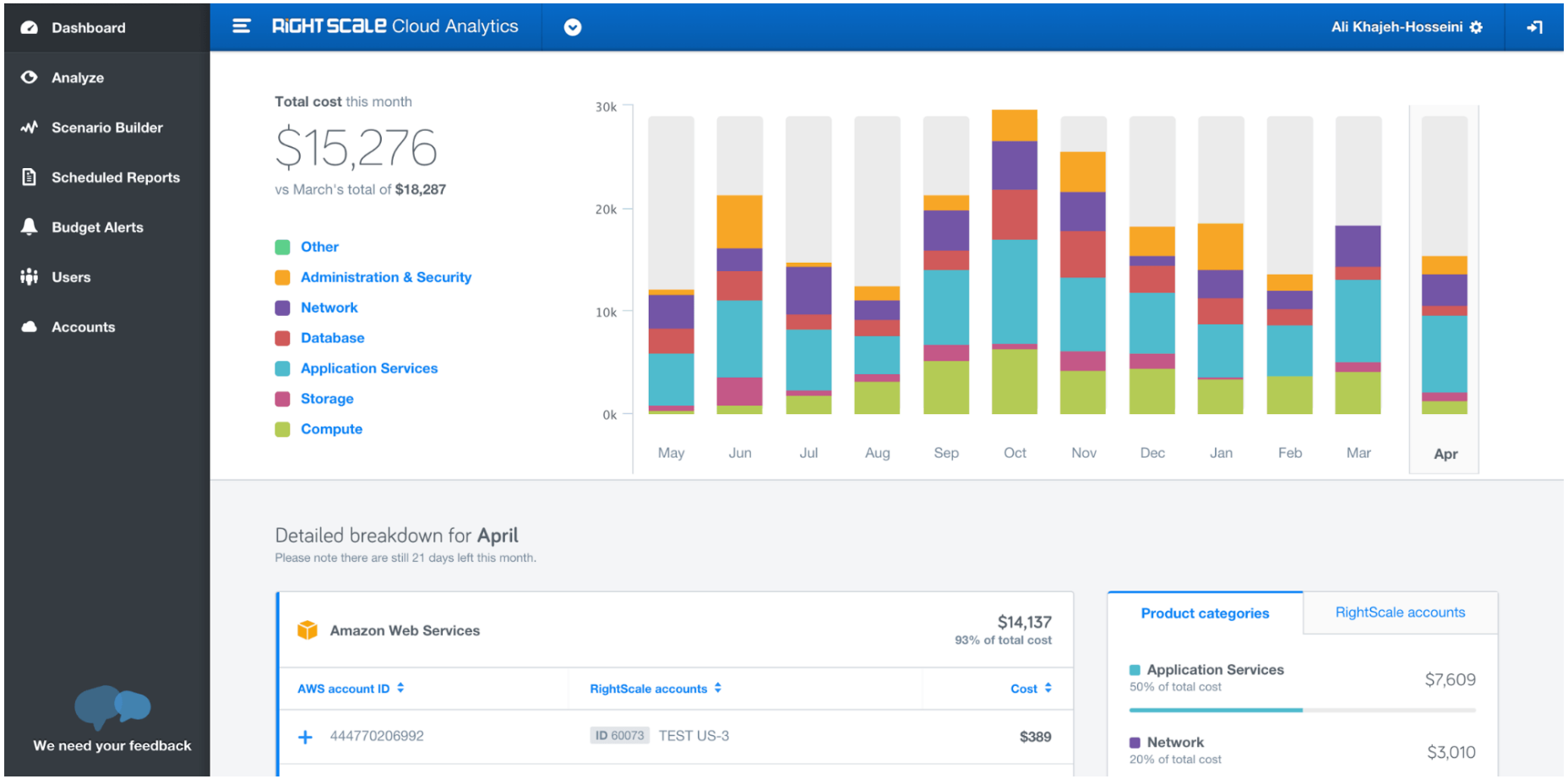
Source: http://data-informed.com/why-cloud-analytics-is-better-analytics/

Salesforce Cloud analytic summary

REGION	PRODUCT	SUM OF AMOUNT
Midwest	Bulldozer Track Roller	$25,600
Midwest	Concrete Mixer	$527,900
Midwest	Concrete Pump	$81,800
Midwest	Concrete Pump Motor	$274,900
Midwest	Crankshaft	$1,065,700
Midwest	Electric Chain Hoist	$596,100
Midwest	Heavy Duty Rock Bucket	$1,050,200

Source: http://data-informed.cm/why-cloud-analytics-is-better-analytics/

Right Scale Cloud analytic summary

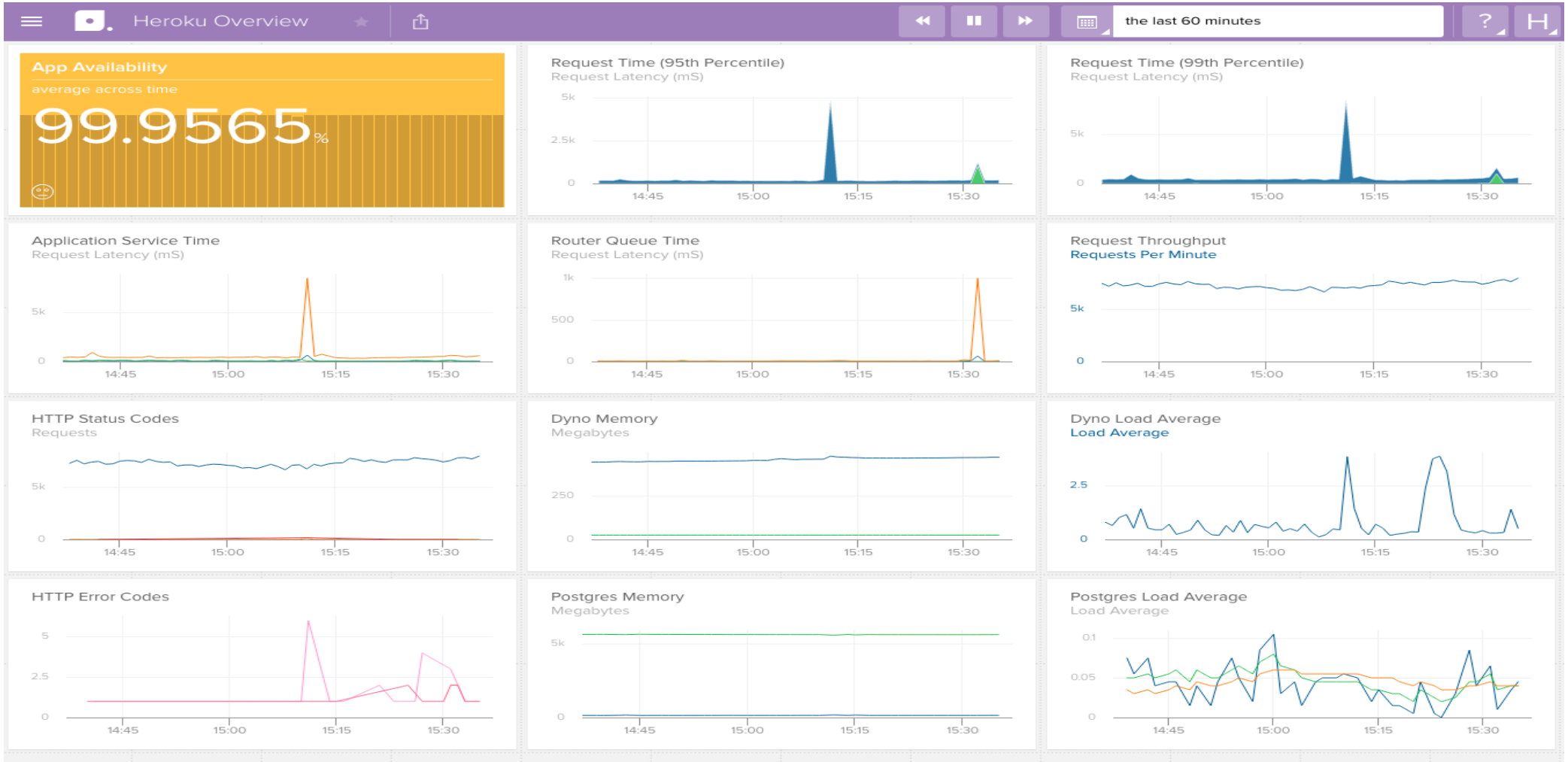


Source: http://www.rightscale.com/blog/cloud-cost-analysis/see-all-your-aws-costs

MS Azure Cloud analytic summary

Source: http://www.findbestwebhosting.com/web-hosting-blog/wp-content/uploads/2015/08/Microsoft-Azure.png

Application Metrics?



Source: https://s3.amazonaws.com/heroku-devcenter-files/article-images/1478908360-heroku-prefab.png

Quiz

Q.1 What should we use on x-axis and y-axis on request/response time graph?

Q.2 What is the difference between round robin and weighted round robin?

Q.3 What is the difference between centralized source control and distributed source control?

Q.4 Why analytical information is useful from cloud resource? Advantages / Disadvantages?

Q.5 Difference between guest machine and host machine?

Q.6 What are the roles of a hypervisor in private cloud versus public cloud?

Reading Material

Chapter 8: Cloud Computing: Concepts, Technology & Architecture
by Zaigham Mahmood, Thomas Erl, Ricardo Puttini

URL: https://www.safaribooksonline.com/library/view/cloud-computing-concepts/9780133387568/ch08.html

OPTIONAL READING:
Cloud Computing Tutorial

URL: https://www.tutorialspoint.com/cloud_computing/cloud_computing_tutorial.pdf


ANY
QUESTIONS
?