International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

FG Cloud TR

Version 1.0
(02/2012)

Focus Group on Cloud Computing
Technical Report

Part 3: Requirements and framework architecture of cloud infrastructure

FOREWORD

The procedures for establishment of focus groups are defined in Recommendation ITU-T A.7. The ITU-T Focus Group on Cloud Computing (FG Cloud) was established further to ITU-T TSAG agreement at its meeting in Geneva, 8-11 February 2010, followed by ITU-T study group and membership consultation.

Even though focus groups have a parent organization, they are organized independently from the usual operating procedures of the ITU, and are financially independent. Texts approved by focus groups (including Technical Reports) do not have the same status as ITU-T Recommendations.

INTELLECTUAL PROPERTY RIGHTS

The ITU draws attention to the possibility that the practice or implementation of this Technical Report may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU-T Focus Group participants or others outside of the Technical Report development process.

Table of Contents

1. Scope

This document identifies capabilities to support the cloud infrastructure.

The document is focused on:

- the capability to efficiently link existing network services, Internet connectivity, and L2/L3 VPN, to public or private cloud services.
- the capability to link a flexible L2 & L3 network management and cloud technology to form an integrated cloud infrastructure enabling cloud services.
- the capability of resource pooling and automation, which is formed by cloud infrastructure using server virtualization and storage-virtualization technologies, to provide computing and storage capabilities for services and applications deployed in the cloud.
- the other supporting functions, such as resource management in the resource pool of cloud computing environments, power management, and so on.

2. References

None.

3. Definitions

3.1 Terms defined elsewhere:

None

3.2 Terms defined in this Technical Report

This Technical Report defines the following terms:

Cloud infrastructure: The basis of a cloud, which provides capabilities for computing, storage, and network resources, including resource orchestration, virtualization, and sharing. It also provides relevant cross-layer supporting functions to support the upper-layer cloud services as well.

VDC: The virtual data centre (VDC) is an evolutionary computing model that presents the data centre as a service view to a single computer, which virtualizes all hardware and software resources behind it.

NOTE - Variations of the VDC model include grid, fabric, and utility computing, each of which has a goal of satisfying quality-of-service (QoS) requirements at guaranteed resource costs. At the highest level, the VDC divides resources into two distinct runtime environments - one for virtual services and one for virtual resources.

4. Abbreviations and acronyms

This Technical Report uses the following abbreviations and acronyms:

ARP	Address Resolution Protocol
BGP	Border Gateway Protocol

BOD	Bandwidth on Demand
BSS	Business Supporting System
CaaS	Computing as a Service
CapEx	Capital Expenditure
CDN	Content Delivery Network
CIFS	Common Internet File System
CRUD	Create, Read, Update and Delete
DCB	Data Centre Bridging
DDOS	Distributed Denial of Service
DHT	Distributed Hash Table
FCoE	Fibre Channel Over Ethernet
GLBA	Gramm-Leach-Bliley Act
IaaS	Infrastructure as a Service
I/O	Input/Output
ICSC	Integrated Cloud Service Control
IS-IS	Intermediate System-to-Intermediate System
ITIL	Information Technology Infrastructure Library
MAC	Media Access Control
MAN	Metropolitan Area Network
MPLS	Multi-Protocol Label Switching
NAS	Network Attached Storage
NFS	Network File System
NIC	Network Interface Card
OS	
OSS	Operational Supporting System
OTN	Optical Transport Network
OSPF	Open Shortest Path First
PaaS	Platform as a Service
PCE	Path Computing Element
QoS	Quality of Service
RSVP-TE	Resource ReSerVation Protocol-Traffic Engineering
SaaS	Software as a Service
SAN	Storage Area Network
SLA	Service Level Agreement
SME	Small and Medium Enterprises

SNIA	Storage Networking Industry Association
SOX	Sarbanes-Oxley
STP	Spanning Tree Protocol
TCO	Total Cost of Ownership
TRILL	Transparent Interconnection of Lots of Links
VDC	Virtual Data Centre
VDI	Virtual Desktop Infrastructure
VEPA	Virtual Ethernet Port Aggregator
VI	Virtual Infrastructure
VM	Virtual Machine
VPN	Virtual Private Network
WAN	Wide Area Network

5. Conventions

In this Technical Report:

The keywords "is required to" indicate a requirement which must be strictly followed and from which no deviation is permitted if conformance to this document is to be claimed.

The keywords "is prohibited from" indicate a requirement which must be strictly followed and from which no deviation is permitted if conformance to this document is to be claimed.

The keywords "is recommended" indicate a requirement which is recommended but which is not absolutely required. Thus this requirement need not be present to claim conformance.

The keywords "is not recommended" indicate a requirement which is not recommended but which is not specifically prohibited. Thus, conformance with this specification can still be claimed even if this requirement is present.

The keywords "can optionally" indicate an optional requirement which is permissible, without implying any sense of being recommended. This term is not intended to imply that the vendor's implementation must provide the option or that the feature can be optionally enabled by the network operator/service provider. Rather, it means that the vendor may optionally provide the feature and still claim conformance with the specification.

6. General requirements and framework of a cloud infrastructure

6.1 General requirements of a cloud infrastructure

A cloud infrastructure includes servers, storages, networks, and other hardware appliances. It can deliver infrastructure resources as a service. Virtualization allows the splitting of a single physical piece of hardware into independent, self-governed environments, which can be extended in terms of CPU, RAM, Disk, I/O and other elements.

The common characteristics of a cloud infrastructure include:

- Network centric: The framework of a cloud infrastructure consists of a large number of computing resources, storage resources, and other hardware devices that connect with each other through the network.
- Service provisioning: Cloud infrastructure provides a multi-level on-demand service mode according to the individualized demand of different customers.
- High scalability/reliability
- Resource pooling/ transparency: The underlying resources (computing, storage, network, etc.) of the cloud infrastructure are transparent to the customer, who does not need to know how and where resources are deployed.

6.1.1 Reduction of capital expenditure

Customers can avoid spending large amounts of capital on the purchase and installation of their IT infrastructure or applications, by moving them to the cloud infrastructure. Capital expenditure on IT reduces the amount of working capital available for other critical operations and business investments. Cloud infrastructure offers a simple operational expenditure that is easier to budget for month-by-month, and prevents wasting money on depreciating assets. Additionally, customers do not need to pay for excess in-house resource capacity to meet fluctuating demand.

6.1.2 Reduced administration costs

It is desirable that IT solutions are deployed extremely quickly and managed, maintained, patched and upgraded remotely by service providers. This means that customers are free to focus on business-critical tasks, and businesses can avoid incurring additional labour and training costs. Cloud infrastructure must allow organizations to streamline procurement processes, and eliminate the need to duplicate certain computer administrative skills related to setup, configuration, and support.

6.1.3 Anywhere access

Cloud infrastructure allows customers to access their application and data securely from any location via an Internet connection. As both the application and the data are stored in the cloud, multiple users can work together on the same project. The network is therefore very important since, if the users' Internet connections fail, they will not be able to access the cloud. Consequently, service providers must construct robust networks and offer several types of cloud-accessing connection (such as fixed and mobile networks), so that customers can access their cloud service from the fixed connection in their office or the nearby Wi-Fi enabled point.

6.1.4 Improved resource utilization

Combining of resources into a cloud infrastructure not only reduces costs and environmental impact but also maximises utilization by delivering resources only when they are needed. There is no need to worry about over-provisioning for a service whose use rate is not meeting predictions, or under-provisioning for one that becomes unexpectedly popular. The moving of an increasing number of applications, infrastructure, and even customer support, into the cloud can save precious time, effort, and budgets, which can be concentrated on the real job of exploiting technology to improve the mission. Sharing IT resource power among multiple tenants can improve utilization rates, as servers are not left idle, and thereby reduce costs significantly while increasing the speed of application development.

6.1.5 Scalability on demand

Cloud infrastructure will have to adapt quickly and flexibly to changing customer requirements, and realize high scalability and high reliability through various mechanisms.

Cloud infrastructure customers can benefit from the economies made through large-scale deployment by service providers who, typically, have very large-scale data centres operating at much higher efficiency levels, and multi-tenant architecture to share resources between many different customers. This model of IT provision allows service providers to pass on savings to their customers. Scalability and flexibility are highly-valuable advantages offered by cloud infrastructure, allowing customers to react quickly to changing IT needs, adding or subtracting capacity and users as required, and responding to real rather than projected requirements. Even better, because cloud follows a utility model in which service costs are based on actual consumption, customers pay only for what they have really used.

Scalability features are not restricted to physical systems but also to the scalability of host. Cloud deployments are required to scale to meet business requirements, and eliminate the need to switch platforms when an existing host reaches its maximum capacity. Customers with smaller configurations can be confident that the platform has been engineered to grow as their requirements increase.

6.1.6 Quality of service guarantee

Without the need to purchase hardware, software licences, or implementation services, cloud infrastructure customers can get their resources arrangement off the ground within minutes. Service providers should provide 24/7 customer support, an immediate response to emergencies, reliability, and guaranteed service levels, e.g., ensure that a customer's applications and/or services are always online and accessible. On the other hand, the disaster recovery and backup mechanism is necessary.

QoS mechanisms ensure that business applications continue to receive the necessary performance guarantee even though they no longer run on dedicated hardware. The QoS provided by an infrastructure is partially determined by its overall performance and efficiency. However, QoS is also the ability to prioritize specific workloads and allocate the needed resources to meet required service levels. It should offer a powerful way to allocate processor, memory, I/O and network traffic resources among applications and virtual guests.

6.2 Functional description of cloud infrastructure

6.2.1 On-demand resource provisioning

On-demand resource provisioning is a fundamental function of cloud infrastructure. With the cloud infrastructure, consumers and SME-users no longer need to spend money to set up their IT or portal system in data centres, but instead obtain or release computing, storage, or network resources as required. This flexibility means that cloud customers can increase IT resources when the volume of tasks or data increases, and reduce IT resources when the need decreases. Ultimately, costs can be reduced by avoiding that servers remain idle.

Network-resource provisioning is very important for cloud infrastructure, and the resources include: the intra-cloud network, the core transport network, and the inter-cloud network. The network influences the user experience directly, so the network resource provisioning must guarantee the QoS and meet the requirements of customer services.

The IT resource provisioning in cloud infrastructure includes the computing and storage resource provisioning. The unique management of IT resource (especially the virtualized resources) is a key issue.

6.2.2 Resource pooling

As infrastructures have become more complex and more brittle, cloud infrastructure needs to automate routine tasks and provisioning, and to flexibly deliver computing/storage/and network resources when and where needed, without over-provisioning. With cloud infrastructure there is potentially no limit to customers' future choices of operating systems, hardware, or applications, when their systems need to expand.

Applications and resources are provisioned automatically while resources are pooled and delivered on demand according to business-driven policies. Such resources let customers build a self-managing virtual infrastructure for increased efficiency and agility while cutting costs at the same time through resource pooling, and resource management from discrete hardware to virtualized pools of shared resources, including servers, storage, and network. By virtualization, resources will be dynamically distributed towards high-priority applications.

6.2.3 Customer self-service model

The dynamic of IT provisioning is one of the most significant forces affecting IT enterprises today, so cloud infrastructure needs a customer self-service model. Enterprise customers demand a modern self-service experience for IT in the workplace, and the same shift is taking place for the infrastructure services provided by corporate data centres.

6.2.4 Convergence

Convergence means that the cloud infrastructure is required to get management traffic, backup traffic, and storage traffic, from centralized storage arrays to the servers on the same network that carries IP data. In general terms, the concept of convergence is 'one wire', which allows a flexible I/O infrastructure with greatly reduced cable requirements. fibre channel over Ethernet (FCoE) is an example of this technology.

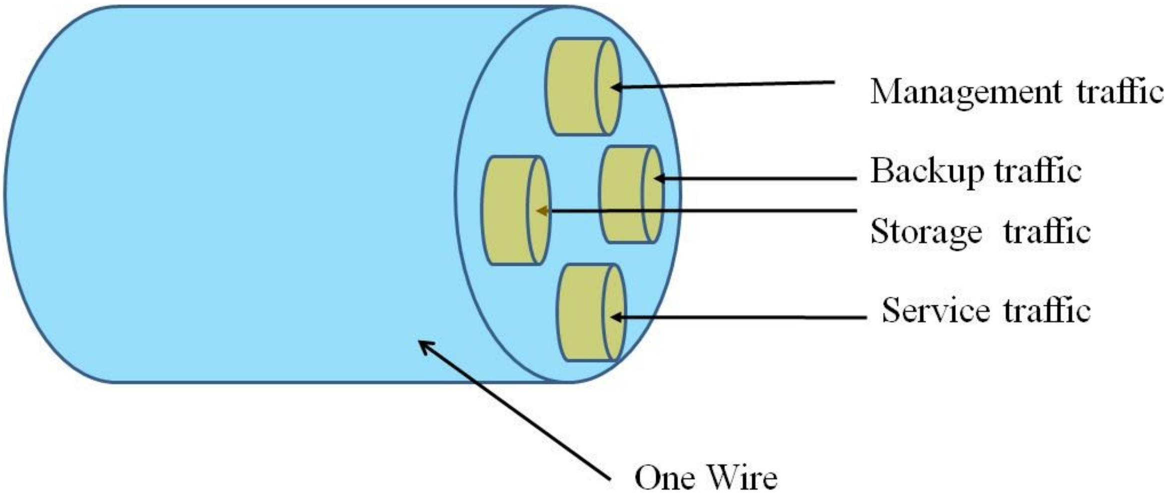


Figure 1 - 'One wire' concept for convergence

6.2.5 Resource orchestration

The network resources are required to be orchestrated with IT resources, which means that the network equipment should be kept aware of the status and configuration of IT resources.

The orchestrator receives requests for cloud services from customers (through a portal or an API) and dispatches these requests to the different components involved in the cloud service composition and delivery. The different components include: the data centre cloud operational supporting system (OSS), the transport network OSS, and possibly the inter-cloud network OSS. This vision represents a major difference from the existing situation, where provisioning and management tools of these worlds are completely separated.

6.3 Framework of cloud infrastructure

Cloud infrastructure is an essential component of cloud architecture. It provides computing, storage and network capabilities, as well as relevant cross-layer supporting functions to support the upper-layer cloud services.

Figure 2 shows the framework of cloud infrastructure. Logically, the cloud infrastructure includes: the resource-layer and cross-layer functions. The resource layer is divided into three sub-layers: resource orchestration, pooling and virtualization, and physical resources. The cross-layer functions include: security and privacy, cloud operational management, and cloud performance functions. However, in this Technical Report, the cross-layer functions largely cover resource-layer relevant functions, such as resource management and power management.

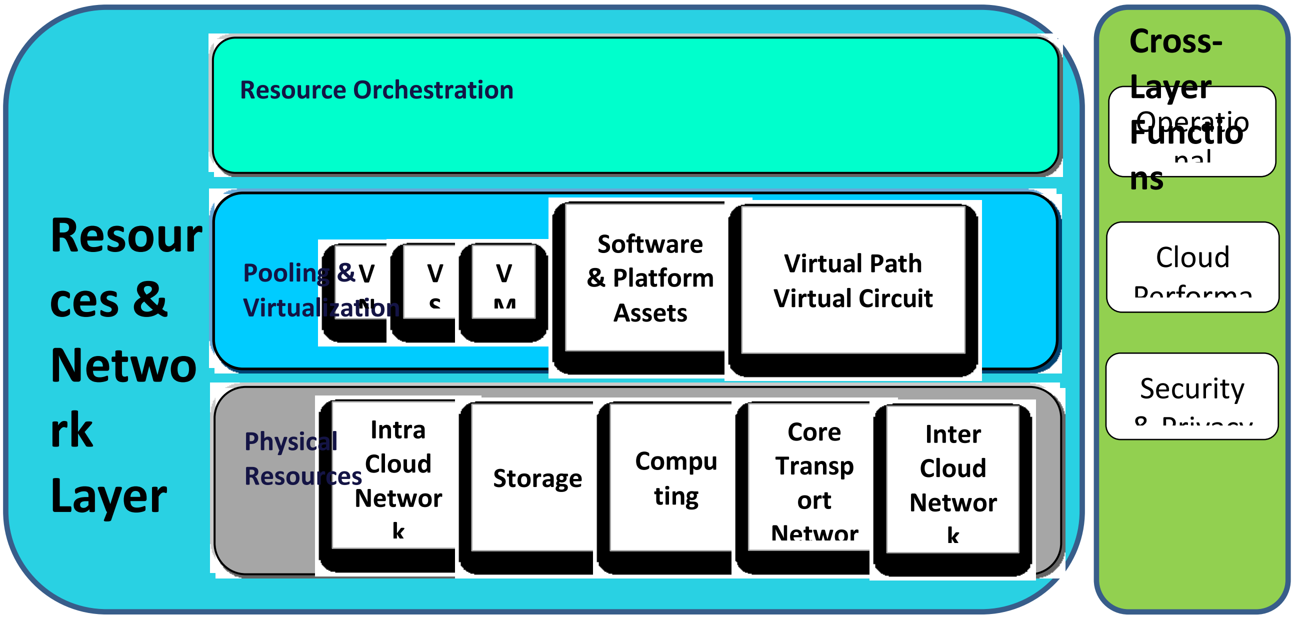


Figure 2 - Framework of cloud infrastructure

Resource orchestration is defined as the channelling of management, monitoring, scheduling of computing, storage, and network resources, into consumable services by the upper layers and users.

Resource orchestration controls the creation, modification, customization and tearing down of virtualized resources. It holds capability directories about what is possible within a cloud segment, based upon the total resource capacity and the incremental allocations that have been implemented to respond to current requests.

Pooling and virtualization of physical resources are essential means to achieve the on-demand and elastic characteristics of cloud computing. Through these processes, physical resources are turned into virtual machines, virtual storage, and virtual networks. These virtual resources are in turn managed and controlled by the resource orchestration, based on user demand. Software and platform assets in the pooling and virtualization layer are the runtime environment, applications, and other software assets used to orchestrate and implement cloud services.

Physical resources refer to the computing, storage, and network resources that are fundamental to providing cloud services. These resources may include those that reside inside cloud data centres (e.g., computing servers, storage servers, and intra-cloud networks) and those that reside outside of a data centre, typically, network resources such as inter-cloud networks and core transport networks.

The cross-layer functions include security and privacy, cloud operational management, and cloud performance functions. This layer caters to a service provider's need for monitoring of resources, and generates a consolidated view of current resource allocations and how efficiently the resources are being utilized. Resource monitoring allows the service provider to exercise load balance to ensure the performance of cloud services. It also flags network and service-related problems, such as hardware or software resource failures, missing SLA targets, or if a provider's network is experiencing security violations or other forms of compromised situations. As the point that collects availability, performance and security information, it is the central source of information on a cloud service provider's service excellence.

For the mapping between this document and RA, please refer to Annex A.

6.4 Network model for cloud infrastructure

There are several network components involved in cloud computing services delivery and composition. They are presented in Figure 3.

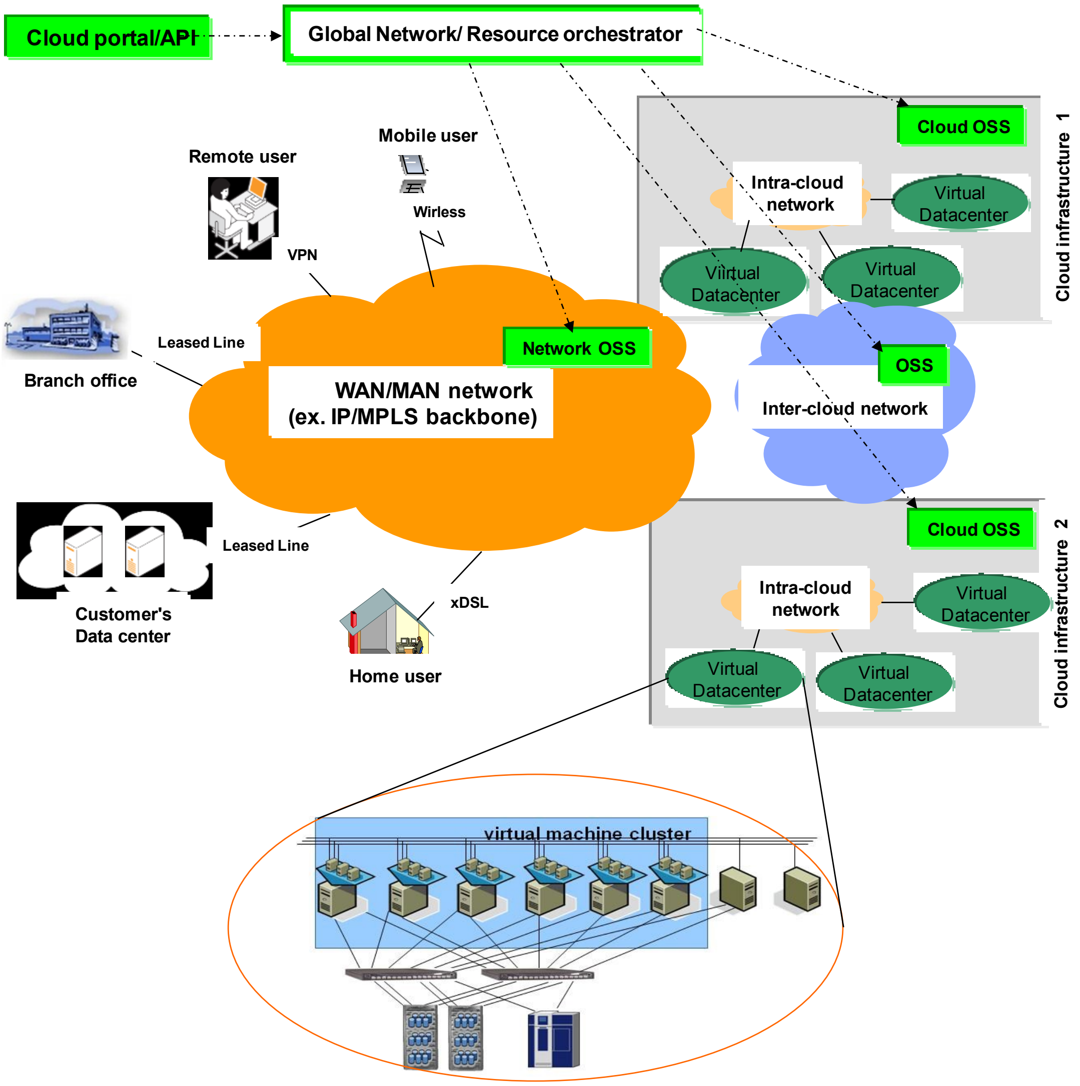


Figure 3 - Cloud network model

A brief description of the role of these network components is given as follows.

- Intra-cloud network: this network connects local cloud infrastructures, such as data centre LAN used to connect servers, storage arrays and L4-L7 services (firewalls, load balancers, application acceleration devices, IDS/IPS, etc..). This network comprises three kinds of traffic:
  - Network between VMs at the same server

    Layer 2 interconnection is needed here. There are two kinds of solutions according to different perspectives: local VMs interconnection based on physical server (e.g., IEEE 802.1Qbg) and local VMs interconnection based on access switch (e.g., IEEE 802.1Qbh).
  - Layer 2 network between servers and storage systems

    e.g., FCoE can be regarded as a typical implementation.
  - Network between VMs at different servers
- Core transport network (WAN/MAN): this is the network used by customers to access and consume cloud services deployed within the cloud provider's data centre.
- Inter-cloud network: this network role is to interconnect cloud infrastructures together. These cloud infrastructures may be owned by the same cloud provider or by different ones.

These three network components are in the heart of cloud services composition and delivery. In order to provide a real added value for cloud services, they must answer cloud services requirements in term of flexibility, scalability, and on-demand resources provisioning, and offer the necessary advanced network functions to ensure performance, security and availability of cloud services.

Operations Support System (OSS): the OSS in the cloud is derived from the OSS for the telecommunication network. There are two categories of OSS in the cloud:

- Network OSS: the traditional OSS is a system dedicated to providers of telecommunication services. The processes supported by network OSS systems include service management and maintenance of the network inventory, configuration of particular network components, as well as fault management.
- Cloud OSS: OSS of cloud infrastructure is the system dedicated to providers of cloud computing services. Cloud OSS supports processes for the maintenance, monitoring and configuration of cloud resources.

7. Functional requirements for computing capability

7.1 Functional requirements for the virtual machine

The virtual machine can provide a virtualized and isolated computing environment for each guest operating system. For a guest OS designed for a physical machine environment, the virtual machine environment should be compatible.

The running state of a virtual machine instance, e.g., the guest OS becoming suspended, cannot influence other running instances.

7.1.1 CPU virtualization and scheduling

CPU virtualization means to virtualize one physical CPU into multiple virtualized CPU (vCPU) for multiple virtual machine instances using time sharing technologies, so that one instance may obtain at least one vCPU. The system administrator can assign vCPUs for virtual machines, and when virtual machines run into instances, the performance of virtualized CPU can be guaranteed or limited.

As an advanced requirement, it is not necessary for the vCPU's capability to be the integral multiples of physical cores, e.g., the vCPU is as powerful as 1/2 the physical core. In this way, system resources can be assigned much more accurately.

The virtual machine (hypervisor) also implements the CPU scheduling function which determines the mapping between the vCPUs and physical CPUs managed by the hypervisor. At any time, a running vCPU of running instance is either associated with a physical CPU or is in an idle state.

7.1.2 Memory virtualization

The memory visualization function divides the physical memory, allocates memory for virtual machine instances when starting up, and collects memory from virtual machines when shutting down. With memory virtualization, every running instance OS may see a continuous memory space. However, every running instance should not see the memory space of other instances. Therefore, the hypervisor is responsible for the memory address conversion from the guest instance physical memory address to the machine physical address. The operating system of a running instance maps the application virtual memory to guest instance physical memory.

As an advanced requirement, the memory manager of the hypervisor can detect whether the virtual memory is actually used by the guest OS. If not, the hypervisor can assign this part of the memory to another guest OS, so that the memory can be shared among the guest OS because, in many cases, the guest OS occupies 2GB memory but does not actually use all of them.

7.1.3 I/O device virtualization

A virtual machine needs to implement the input/output virtualization function so that each virtual machine OS can equip its own virtual I/O devices abstracted from the I/O devices of the physical machine. I/O virtualization of the virtual machine then implements the mapping of virtual and physical devices. Though the maximum number of physical cards depends on the number of system board slots, the number of virtual I/O devices should not be constrained. Also, the physical I/O devices can be shared by multiple virtual machines as the information transfer channel, but the data transferred or stored by the physical I/O devices are never shared among the virtual guests' OS.

7.1.4 Duplication of virtual machines

An existing virtual machine with a deployed operating system and installed applications can be duplicated to a new virtual machine with the same software stack (OS and application). This function facilitates fast production of new virtual machines and virtual machine backup in the execution environment.

7.1.5 Static migration of the virtual machine

Static migration of the virtual machine means the moving of an operating system and its applications between a virtual machine and a physical machine, or between virtual machines on different physical machines, with the operating system temporarily stopped.

7.1.6 Multi-tenancy/User self-service

The operational mode of cloud computing is to rent resources (network, storage, computing, etc) to a person or an enterprise. From the network side, different enterprises are required to be identified and isolated. From the tenancy view, a logical network which can be managed should be provided, including server, IP, bandwidth, enhancement service, policy, and topology.

The tenant/user is required to request, manage, and access cloud services by himself. By using role-based access control and authorization, the system is required to provide the capability to assign certain aspects of administration (such as starting/stopping/removing VMs) to designated "tenant administrators".

7.1.7 Automation

It is critical to have the capability to automate all expected operations over the lifetime of a hardware or software component. In the absence of a deeply-embedded capability across all layers of the infrastructure, dynamic processes will be halted unless user intervention or other manual processing occurs. The system could perform operations such as starting or stopping a VM, rebooting a server, and applying software updates automatically. The atomic units of automation are required to be associated and implemented by higher-level management systems.

8. Functional requirements for the cloud network

8.1 General requirements

8.1.1 Scalability

Cloud service providers, including telecommunication players, will have to adopt new network and server technologies that will enhance the scale of their data centres, allowing an evolution from hundreds (or a few thousands) of servers to tens or even hundreds of thousands of servers. There are many limitations for current technologies to provide large scale Ethernet domains. Some of these limitations are: address resolution protocol (ARP) broadcast limitations, MAC size table constraints, and spanning tree protocol limitations. A lot of initiatives within the industry and the academia are proposed to solve these issues. An example is the transport interconnection of a lot of links (TRILL) protocol, under standardization at the IETF [1]. A more long-term initiative is to work on solutions for separating names from locations (such as locator/ID separation protocol (LISP) under standardization at the IETF). The goal here is to be able to assign any IP address to any VM, thereby breaking all the scalability constraints of current IP sub-networking.

8.1.2 Performance

The nature of cloud-generated traffic is very random [2]. Cloud providers will have to transform their data centres LAN interconnect into a local area distributed system, with reliable high-speed direct (point-to-point) communications between servers with congestion-free links, and uniform bandwidth between any two arbitrary servers within the data centre. From an architecture perspective, the current three-tier topology (access, aggregation and core) used in data centres is not well-adapted to provide these requirements. A second item to consider is how to maximize the use of available links. Here, also, there are some research initiatives that studied traffic randomization in order to maximize the utilization of available links with 802.1ah or IP in IP encapsulation [4].

In addition to this, a cloud-ready data centre network is required to ensure application performance, and means to provide and control applications level SLAs. Application acceleration and services optimization is required when designing a network infrastructure to deliver cloud services.

8.1.3 Agility and flexibility

A high requirement of a cloud-ready data centre is to be able to follow the high-dynamicity nature of cloud resources. The network is required to be "VM-aware", and to be able to adapt itself dynamically to events such as VM mobility. Several propositions are under standardization to enable this VM-aware networking (e.g., VEPA and VN-tag ongoing standards in the IEEE).

Another aspect that should be studied is the finding of solutions to have a fine-grained control of flows routing within the data centre. This is mandatory to be able to define policy-based routing (in opposition to shortest path routing) to force a given flow passing through an ordered sequence of L4-L7 network services devices (e.g., firewall, load balancer and application acceleration). Current solutions to do this are very costly (based on dedicated VLANs and Access Lists ACLs).

8.1.4 Convergence of data and storage networks

Reducing the total cost of ownership (TCO) of the overall cloud infrastructure is mandatory for a cloud provider (including telecommunication players) to be able to propose cloud services at competitive prices. In this area, one of the items to consider is the deployment of new hardware technologies that reduce the cloud infrastructure capital expenditure (CapEx). A major hardware evolution seen today is unified Ethernet and Fibre channel fabric. A unified protocol framework for communications within data centres for Ethernet and storage should be taken seriously into consideration when building a cloud infrastructure.

For more information, please refer to Annex B.

8.1.5 Network interface card virtualization

With the network interface card (NIC) virtualization, a system administrator can create and delete a virtual network interface card for a guest virtual machine OS, regardless of the number of physical NICs. One physical NIC may be shared by multiple NICs belonging to multiple running guest instances. To implement virtual NICs, a virtual bridge (Ethernet bridge) could be set up on the virtual machine hypervisor or the host OS. The virtual network bridge links the virtual NICs with the physical NIC. The virtual NICs will need to have different MAC addresses since the local area network identifies network nodes with MAC addresses.

As an advanced requirement, the vNICs of one hypervisor can be grouped into a same VLAN with the vNICs of another hypervisor. A standard needs to be developed between the hypervisor and the network switch, otherwise they cannot cooperate with each other.

8.1.6 Dynamic migration of a virtual machine

The dynamic migration of a virtual machine means the moving of an operating system and its applications from a virtual machine to another virtual machine, without the operating system being stopped. Dynamic migration is also called live migration. Dynamic migration enables the operation of virtual machine high availability, dynamic load distribution, and service continuation during hardware maintenance.

With dynamic migration, the IP address of VM should be kept unchanged where possible (i.e. when migrating within the same Layer 3 subnet). The QoS policy, security policy, and traffic policy, for the VM on network equipment should be migrated at the same time, so that the services' continuity, reliability, and security can be guaranteed.

The virtual machine is required to be migrated among the physical servers. During the process, the guest OS might not be aware of the migration. The virtual environment, including memory, storage, and network configuration, are kept the same as before.

NOTE: there will inevitably be a brief interval while the final dynamic state information (e.g., virtual CPU register values) is transferred to the new VM host.

Once the cloud system has finished this process, the VM on the target host will be starting to run. Meanwhile, the VM on the source host will be terminated. During this process, Layer-2 network devices (e.g., switch) must be made aware of this change.

In addition, if the migration was to a host in a different IP subnet, the layer 3 network will also need to be updated, and the IP address of the VM will change.

8.1.7 Seamless migration of virtue machine

Cloud network infrastructure is recommended to support seamless migration of VM. Seamless migration follows the same process of dynamic migration, with the additional restriction that services will not be interrupted during the migration. It means that this will only be possible for the migration within the same subnet, so that the IP address and any open connections, sockets, handles, etc., will remain unchanged.

8.1.8 IPv4/IPv6 support

For the cloud infrastructure's systems, services and applications' data switching, transmission and routing over the IPv6 network, the following conditions should be met:

- Network layer of clouds is recommended to support IPv6 packets switching and routing for data transportation and route selection.
- OS of cloud service platforms is recommended to support IPv6 address allocation, configuration and Ipv6 protocol resolution.
- In the virtualization environment, a virtualization layer of cloud infrastructure is recommended to support IPv6 address allocation, configuration, addressing, data packet parity, and other IPv6 protocol stack functions.

Cloud computing services provide users with services with the following types of IP protocol networks:

- IPv4 only
- IPv6 only
- Dual stack, both IPv4 and IPv6

8.2 Functional requirements for the core transport network

8.2.1 General requirements for the core transport network

8.2.1.1 Reachability

The core transport network is recommended to be compatible with multiple terminals and multiple access modes, making a wide range of customers use cloud computing services and maximizing the expansion of the scope of cloud computing services. Such compatibility reduces the terminal physical device requirements of customers to the lowest, such as thin terminal, dumb terminal, intelligent soft-terminal and browser. Customers could enjoy cloud services at anytime, anywhere, in any terminal and in any way.

8.2.1.2 Reliability

The access to core transport network shall be reliable, and shall ensure that the network connection is sufficiently available to meet the SLA.

8.2.1.3 Flexibility

The core transport network should enable convenient access to new data centres, new customers' terminals and deliver new cloud services as quickly as possible.

8.2.1.4 Awareness of the user/terminal attributes

The core transport network provides connections between the user/terminal and the service, and it plays an essential role in the experience of the cloud computing service. Because of the CapEx, the network cannot be constructed as large as possible, and the network resources are utilized statistically. In most cases, the resource of network, computing, and storage, needs to be used and managed dynamically. Thus, the resource planning is very difficult. In some extreme cases, the network resources are not enough for the cloud-computing service. In these cases, they should be used according to the user/terminal attributes.

In order to transform the core transport network into "smart-pipes", the core transport network should allocate the adequate resource to every user/terminal. Thus, the network is required to know the attributes of the users and terminals. The attributes of the terminal include the terminal type, access media, and mobility, etc. The attributes of the user include the subscriber profile, location, roaming, the network prioritization of the using application, etc.

If the user/terminal changes its attributes via the portal, or in another way, the core transport network is required to know the change of the attributes.

8.2.1.5 Awareness of the application attributes

In many cases, the cloud computing service needs interconnecting cloud data centres to recover, maintain, and optimize resources. If so, more WAN resources are needed. Because of the CapEx, the network cannot be constructed as large as possible, and the network resources are utilized statistically. In most cases, the resource of network, computing and storage needs to be used and managed dynamically. Thus, resource planning is very difficult. In some extreme cases, the WAN network resources are insufficient for the cloud computing service. Since the cloud computing services can include all varieties of applications and they do not have the same SLA, in such cases, the network should be smart enough to identify the specific application and allocate the necessary resources to the important one.

Virtual data centres can offer several or many kinds of applications and services. The cloud service provider could identify the applications' attributes according to the IP addresses of applications and the service deployed in the virtual data centre. The cloud service provider could also be aware of the application attributes through many ways, such as deploying a deep packet inspection (DPI) system in the core transport network. The attributes of the application include the SLA profile, the application type (such as video, voice and data), and the cloud computing service to which it belongs. In some cases, if the application's attributes are changed, the core transport network should know the change and adjust network properties for better delivery of the cloud services or applications.

8.2.2 Network topology requirement for core transport network

Customers of cloud services access and consume cloud services through this network component. From a telecommunication provider perspective, the efforts should be concentrated on preventing this component from being "dump pipes" for tier cloud providers. To achieve this, the core transport network will have to provide advanced network features for cloud services delivery. These features could be offered to cloud services customers either by the telecom provider itself or by third party cloud services providers.

As traffic becomes multidirectional across the core transport network, the network layers need to provide enhanced visibility into the cloud data centre services flowing across them. The general network requirements for core transport network are described as follows:

- It is recommended to provide network intelligence and proximity services.

There are several existing techniques that determine and implement the proximity services in the network, such as the domain name system (DNS), round-trip time (RTT) measurements, and manual configuration. However, these approaches do not provide interworking between the network and the application layer, bringing inflexibility and unreliability. Therefore, the core transport network is recommended to provide capabilities/interfaces to optimize application traffic, which are based on accurate information like management statistics, and policy database.

8.3 Functional requirements for intra-cloud network

8.3.1 Network topology requirements for intra-cloud network

Figure 4 shows a traditional "hierarchy of switches" tree topology inside a data centre, which contains three layers: access, aggregation and core. Basically, the access layer and the aggregation layer are linked by the data link layer which the spanning tree protocol (STP) runs at. This scheme guarantees that only one path exists between any two points (servers) and that there are no loops with other links being blocked. For traditional data centres, this STP-based network tree topology works.

Inside a virtualized cloud infrastructure, each server can be virtualized as many VMs, and each VM behaves like one single server and has the same connection requirement, which generates more and more horizontal traffic between VMs inside a data centre. However, network topology based on STP keeps only one path between these servers and cannot satisfy these requirements. Moreover, traditional MAC addresses cannot be layered, leading to exhausted address space, especially in a virtualized data centre.

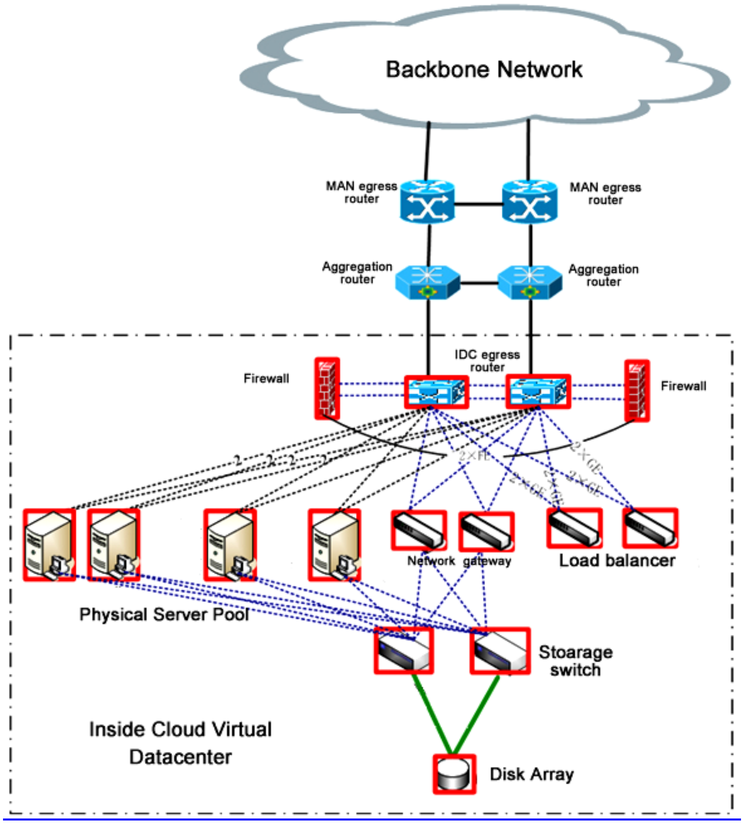


Figure 4 - Traditional tree topology inside a data centre

The general network requirements for the intra-cloud network are described below.

It can optionally provide an efficient layer 2 network, because more and more horizontal traffic among servers will predominate.

It can optionally support a large number of addressable virtual machines.

It is recommended to support live migration of virtual machines.

It is recommended to support traffic monitoring among virtual machines

It is recommended to provide multi-paths on layer 2

For more information, please refer to Annex C.

8.4 Functional requirements for the inter-cloud network

8.4.1 Network topology requirements for the inter-cloud network

With the development of cloud-oriented services, there will be much more traffic between data centres but not limited within them. In this scheme, all the traffic between data centres must run through the backbone, which brings a great burden to it. The traditional data centre is attached to the backbone network as shown in Figure 5.

The traffic models will be varied with applications, such as data synchronous, huge data replication, transmission relay of high-resolution video, and data floating of frequently-used services, etc. In addition, different data centres may have a respective workload. For instance, there could be transaction-related, data-processing related or storage-related workloads. They require different traffic models to guarantee various requirements of latency, packet loss sensitivity and bandwidth.

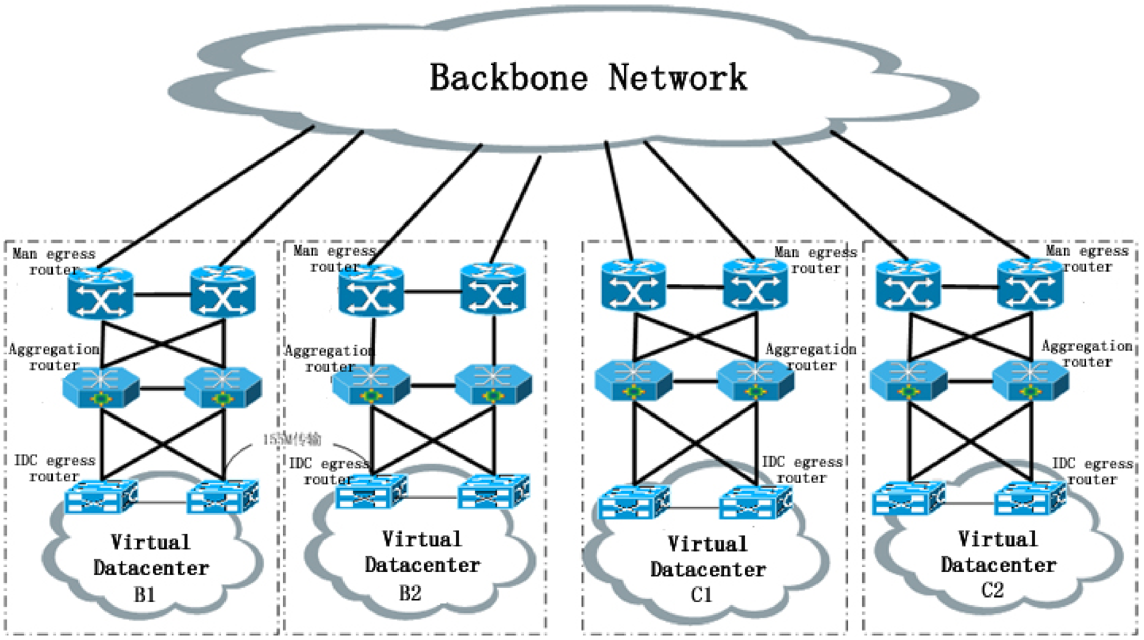


Figure 5 - Traditional network topology between cloud data centres

The inter-network topology of future data centres needs to satisfy the following requirements:

It may provide high degrees of agility. In order to lower the burden of backbone, topology policies should be different from the services to distinguish traffic which can run through the backbone, like normal long-distance data and traffic which can be floated in the dedicated network among data centres. This has implications where, in addition to being able to route packets effectively and efficiently, one can interact with network in a way that topologies can be adjusted to the specific requirements. This ability to change topologies can be driven through an API so that agents can deal with the required changes without human intervention.

There are many use cases where an interconnecting cloud data centre is required (recovery, maintenance, resources optimization, hybrid clouds, etc.). Layer 3 data centre interconnection does not raise important challenges, and there are well-proven technologies that have been widely used and are in continual enhancement to answer any new requirements. On the other hand, there is an emerging technical challenge with cloud computing that telecommunication players should resolve when choosing the technology to use for providing layer 2 extensions across geographically-distributed data centres to interconnect cloud infrastructures at layer 2. This is, for example, required to perform live virtual machines migration between data centres. Existing layer 2 solutions have several limitations that make them inadequate to fulfil advanced requirements in terms of flexibility, scalability and availability of cloud computing services.

A typical network topology between cloud data centres and network topology inside cloud data centres is shown in the following figures, which can be taken as a reference for the network connections between the IDCs and inside the IDCs.

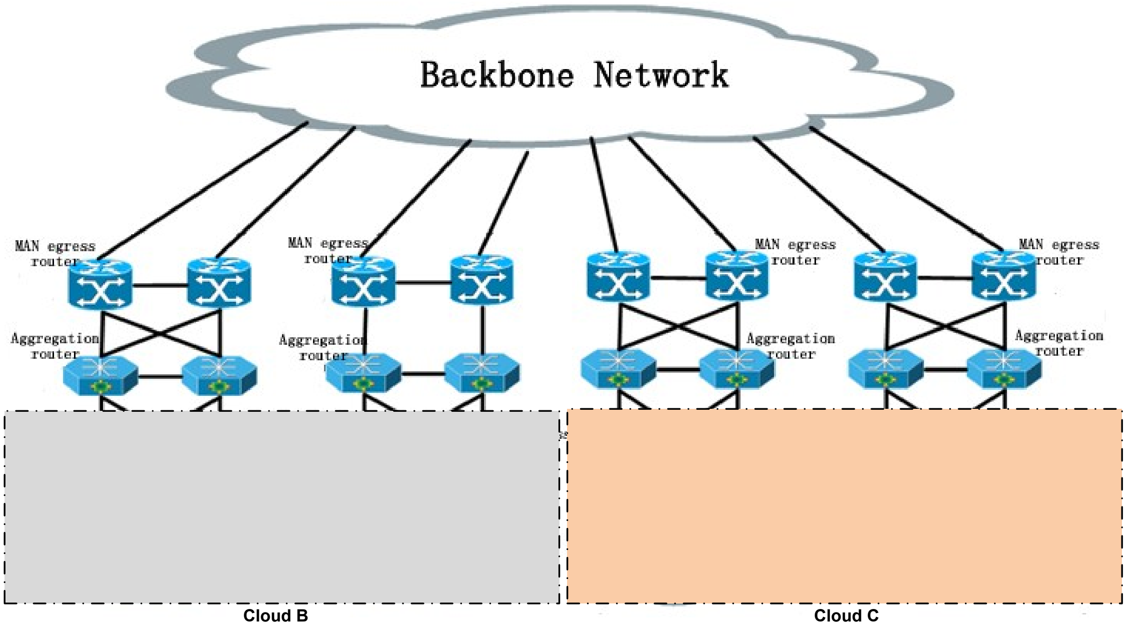


Figure 6 - Network topology between cloud data centres

Figure 6 shows the network topology between cloud data centres. As for a traditional data centre, the egress router of the cloud data centre will be connected to the aggregation router of a metropolitan area network. Since different cloud data centres may contain the infrastructure and devices that belong to the same cloud service provider, these cloud data centres may be connected via a special line or directly connected via a fibre channel which could be considered as part of the intra-network in the provider's cloud infrastructure.

8.5 Requirement of network services

8.5.1 Bandwidth on-demand services

In the traditional network-planning process, the fluctuation of the traffic can be estimated and predicted with the required confidence level. In the case of traffic generated in the cloud computing environment, the bursts in the network produced, for example, by VM migration, make traffic flows strongly non-linear and unpredicted. Therefore, the network structure should be dynamically adapted and, in consequence, the bandwidth modification service should be considered in order to decrease this impact.

The goal of the bandwidth on demand (BoD) service is to propose the ability to realize on demand the bandwidth modifications on the particular links. This is useful in cases such as:

- massive data recovery transfer for cloud services
- intensive bandwidth consuming applications hosted in the cloud data centre
- partial link utilization of applications hosted in the cloud data centre

The main advantage for the BoD service is to increase, as well as to decrease, the bandwidth allocated on the particular links. In consequence, the business use case assumes the “pay-as-you-go” model and the client is charged for real-link utilization.

The BoD service is a dedicated service implemented at the layer 3 or below in the OSI network model. The BoD service could be provided in the network by different technical solutions depending on domain network technology. The granularity of the BoD service may be defined according to the network transport technologies used for cloud computing services. The BoD service can be used as a supporting service for other business services in the cloud.

In order to realize the BoD service provisioning, standardized interfaces should be provided to users and applications for resource modifications. The interfaces can support dynamic signalling for instantaneous provisioning. Requests concerning bandwidth modifications may be sent through a user interface, which can be a web interface or an application-programming interface (API). Such requests may be accepted by utilization of available resources in the cloud environment. In case of insufficient resources, such requests may be denied or modified during the negotiation process to deliver the highest possible parameters.

Bandwidth on demand is both economical and practical. It makes sense to use a switched line and pay only for services as they are needed, rather than lease an expensive dedicated line that may go underused part of the time. Networks based on frame relay, ATM or Ethernet can automatically provide more capacity without the need to add additional physical lines, but the capacity is limited by the size of the trunk that connects a customer to the network.

The BoD service should guarantee capacity, providing connection-oriented and point-to-point service between every BoD service stakeholder. The BoD service should have the following characteristics:

- BoD service can optionally provide a seamless solution of bandwidth allocation, independent of network technology and architecture used by the cloud service provider
- BoD service can optionally provide symmetric or asymmetric capacity, according to the provider policies and restrictions imposed by the technology used
- BoD service can optionally be bi-directional and provide identical forward and return paths in the case of some QoS sensitive services, i.e. time synchronization, and videoconference streaming
- BoD service can be fully protected by creating two independent paths from source to destination, including the physical layer, in order to secure some emergency services i.e. e-health remote operations
- BoD service is required to be realized as fast as possible.
- Resources for BoD service can optionally be reserved in advance.

Note: Stakeholders are all the subjects who have interests in a BoD service community, in other words, all the persons (i.ex. end-user), groups and other subjects (e.g., operator, CSP, and so on) who can affect or can be affected by a BoD service. The different stakeholders have different roles in a BoD service enterprise.

For more information on cloud-specific use case descriptions related to the bandwidth on-demand services, refer to Appendix IV.

8.5.2 L2-L3 networking services

The following L2-L3 service types are identified:

- L2 service type
  - The mostly relevant network connection between clouds. Limited applicability is needed between cloud provider and cloud service requester.
- L3 service type
  - Global IPv4 or IPv6 touting tables: Main interconnection service type between clouds and between cloud service requester and cloud service provider. There is no traffic separation between Internet traffic in the network and traffic that accesses cloud resources. This service type is applicable if the application level security is sufficient.
  - Private IPv4 or IPv6 routing tables: This service type will assure enterprises and commercial customers to access cloud resources over dedicated network access. This service type is a prerequisite for hosting critical applications in the cloud.

The following service attributes are identified:

- Latency
- Jitter
- Packet loss
- Prioritization of traffic types
- Or, expressed differently, available bandwidth between clouds and available bandwidth between cloud service requester and cloud service provider

In order for such users and applications to consider cloud to host their critical mission applications, the following network level attributes are very important:

- Disaster recovery options: There are arrangements for guaranteed access through the network for events such as earthquakes, terrorist attacks, Tsunami, etc.
- Network reliability and availability: During network events (such as fibre cuts, router crashes, etc.), the fastest level of detection and service restoration is applied
- Secured access to cloud resources: Many layers of security are applied to traffic/queries destined to critical mission applications
- Distributed servicing: Applications content/data is cloned across multiple DCs, flows/request are hashed to multiple DCs, and there is no single point of vulnerability/failure
- Network QoS: Adequate service differentiation to protect against any traffic starvation DDOS
- Multicast: Resilient and redundant multicast architecture in the network to support the highest availability for multicast flows of critical mission applications

8.5.3 L4-L7 networking services

In order to transform the core transport network into "smart-pipes" with real added value for cloud services delivery, this network component is required to provide on-demand L4-L7 network services and to ensure the performance, security and availability. Some examples of this requirement are:

- Security functions: provide on-demand security functions within the core transport network to protect and control customer's traffic to cloud services. An example is firewalling functions and intrusion detection and prevention.
- Performance: provide on-demand application acceleration and optimization services for cloud services. This is essential to ensure application performance because these applications will be accessed remotely (from customers sites to the cloud providers sites), and it is well-known that most business applications were initially designed for the LAN environment and are negatively impacted by the delays and packet loss of the core transport network.

8.5.4 L2-3 VPN service requirements

Virtual machines migration between data centres will maintain its existing IP and MAC addresses, which requires L2 interconnection across WAN. The Layer 2 VPN is required to provide secured Layer 2 connections across WAN.

If VPN is deployed, the general requirements are identified as follows:

- The VPN is recommended to provide a transparent L2 network through WAN to connect data centres;
- The VPN is recommended to support SLA;
- The VPN is recommended to be able to provide traffic load balance over the link;
- The VPN is recommended to provide multi-homing for data centre access;
- The VPN is recommended to be able to provide traffic protection if the links or nodes in WAN are failed;
- The VPN edge device is recommended to be able to provide node redundancy or load balancing;
- The VPN is recommended to have good scalability to accommodate large number of VPN edge devices and customer devices;

When the cloud interconnected is a private cloud for one customer, the private cloud may be an L2 or an L3 network. The VPN used to connect the private cloud should be able to provide both L2 and L3 connections between data centres. All applications running on a data centre's servers, and data stored in a data centre's storage devices, should allow traffic segregation per VPN.

If the service provider has already provided L3VPN to the customer, the L2VPN interconnecting cloud is recommended to be integrated into existing L3VPN, so as to provide one integrated VPN to the customer.

8.6 Network automation

Network automation is required for telecommunication players if they want to provide self-managed and on-demand network services for cloud customers. Telecommunication operators have already deployed backbone automation tools within their operational support system (OSS) that interfaces directly with the business support system (BSS) to directly push network-level configurations to the network infrastructure. In the context of on-demand cloud and network services, the new challenge that should be addressed here is to allow customers' action on a self-management portal, to directly perform management actions (modification or creation) on the core transport network infrastructure.

Network automation enables the administrator to implement a policy-based automation for managing the cloud data centre, while ensuring highest uptime for business-critical services. It provides a unified platform for managing physical and virtual network devices. The high-level requirements of network automation provisioning could be summarized as follows:

Network automation could shield users from underlying complexity and enable the administrator to change network device configuration easily and deploy the access control list (ACL), and other policy changes that support adds, moves, and changes of other service resources, such as servers, without service disruptions.

Major requirements are described as follows:

- Deployment change without rebooting the network devices.
- ACL deployment to network devices without exposing the network to potential security vulnerability.
- Virtual network provisioning (i.e. VRF, switches, firewalls, load balancers and WAN accelerators).
- Provisioning actions for cloud data centre switch/router and hypervisor virtual switches.
- Real-time and historical visibility into configuration and change attributes.
- Pre-built templates which can simplify configuration process.
- Automatic scripts creation.

The following sub-clause describes a few examples of network automation.

8.6.1 Cloud switch/router

In the core transport network, the cloud switch and router are needed to ensure better services, just like the content delivery network (CDN) which ensures a better streaming service.

First, the cloud switch and router should identify different types of services. For example, the cloud switch and router can obtain information of the service type from the control stratum. The cloud switch and router should then have the ability to reserve appropriate different kinds of network resources and provide a logically isolated network partition for every user, according to the type of service (e.g., multimedia, voice).The cloud switch and router should also take full advantage of the limited network resources to achieve a better user experience.

8.6.1.1 Cloud switch

With the development of cloud computing, it is more and more difficult to keep up with the traditional switch and to meet the increasing requirements of network connection. The traditional switch cannot recognize different virtual machines in the same physical machine. Therefore, different control strategies cannot be set to different virtual machines. The traditional switch cannot monitor horizontal flow between virtual machines in the same physical machine, which can introduce security risks. The traditional switch cannot keep the control strategy unchanged when a virtual machine is migrated.

The cloud switch is a kind of switch that can meet the above requirements. In the cloud switch architecture, switch management and control strategy management have been abstracted to the cloud switch server so as to realize unified management and simplify the cloud switch structure (as shown in Figure 7). The main function of the simplified cloud switch is data switching. In addition, each virtual machine has a corresponding virtual port in the cloud switch which can achieve virtual machine recognition. It also has a virtual control strategy module to receive and save control strategies from the cloud switch server for each virtual machine. The cloud switch server has the control strategy of each virtual machine. When a virtual machine migrates to a new cloud switch, the server just has to distribute the control strategy to the new cloud switch to achieve control strategy migration with the virtual machine.

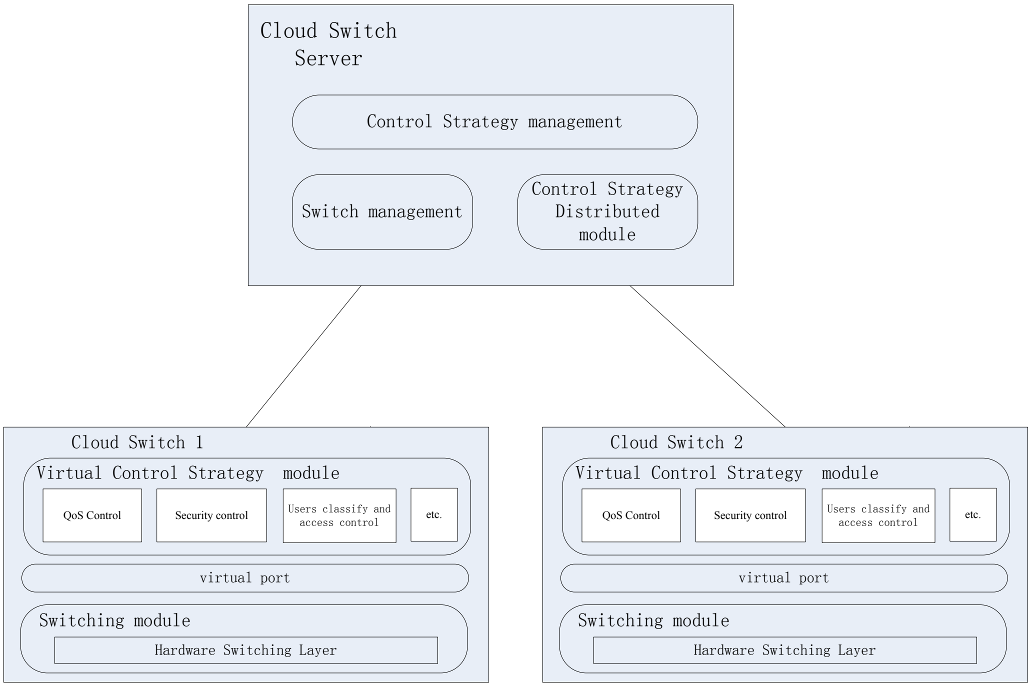


Figure 7 - Cloud switch system architecture

8.6.1.2 Cloud router

With the development of Internet applications, the traditional router is encountering more and more difficulty in meeting the increasing computing requirements of new services. The traditional router operates as a stand-alone system in the existing network, uses only its own computing resource to finish routing and forwarding, and cannot share the computing capability of other routers inside the network. In this regard, at the very beginning of network planning, it is necessary to upgrade the stand-alone router processing capacity to guarantee QoS, especially in the scenario of occasional burst traffic. However, in the actual operation, especially in the carrier-class, this planning will bring much redundancy and low efficiency.

Cloud router architecture is put forward in order to solve the problem described above.

In the cloud router architecture, routing management, resource management, and routing protocol computing have been abstracted to the cloud router so as to realize unified management and scheduling of computing resource (as shown in Figure 8). The reduced routers only retain basic functions, such as transport, switching, etc. In addition, the middleware layer has been offered to communicate between the cloud router and ordinary routers. This design can be used in the access layer and the aggregation layer in the existing network, where there is much redundant computing ability. The major components of the cloud router can be described as follows.

Routing protocol computing layer: to implement the routing protocols such as BGP, ISIS, OSPF, etc.

- Resource management layer: to manage all the computing resources in the whole network.
- Routing management layer: to maintain the routing table for the whole network.
- Control layer: to schedule the computing resources of the whole network and realize a dynamic load balance.

The cloud router design has the following advantages:

- Simplify management: The cloud router is managed as a container. It can reduce management points and improve management efficiency.
- Simplify routing: The cloud router runs as a virtual network element. It can simplify network topology, reduce network scale, reduce IGP flooding, improve routing computing efficiency, reduce routing convergence time, and improve routing stability.
- Balance traffic: The cloud router manages the links and nodes. It can realize unified traffic management and balance, avoid congestion, and improve resource utilization.

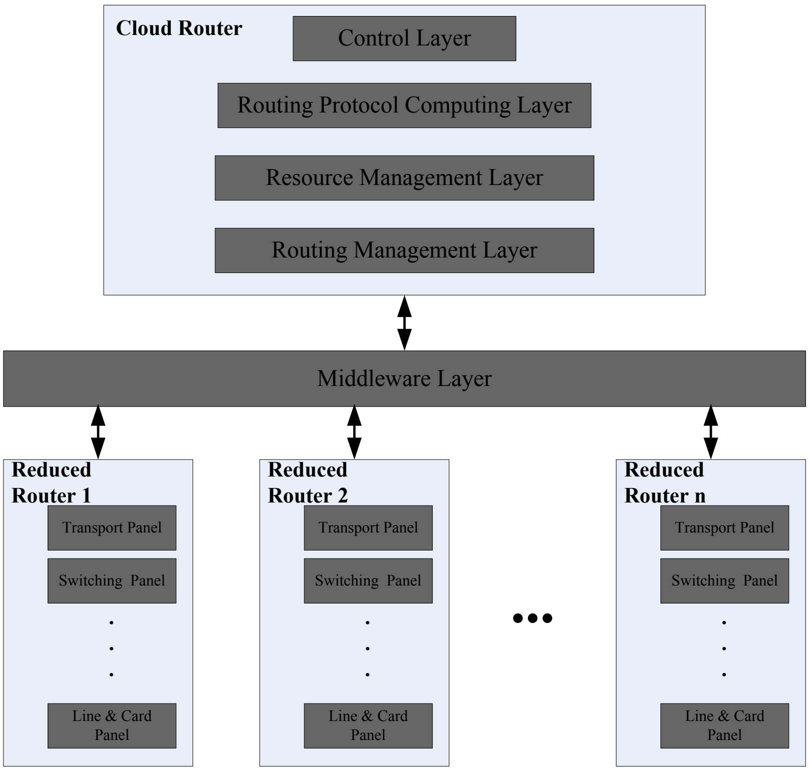


Figure 8 - Cloud router architecture

9. Functional requirements and architecture for storage capability

9.1 Functional requirement of cloud storage

Storage plays an important role in cloud infrastructure. All the applications and data are required to be delivered reliably. Storage array based on fibre channel or iSCSI must be connected to the servers. So FCoE/DCB becomes necessary in the cloud computing environment with which the physical path could be well converged.

9.1.1 Storage space requirements

Cloud storage requires no space limit, and needs to support dynamic expansion. Cloud storage needs to support at least the PB level of storage capacity. The file system namespace of cloud storage needs to support the number of files on the billion-level.

9.1.2 Storage interface requirements

Cloud storage needs to provide standard storage interfaces which can be categorized into four types:

- Block storage interface, for example SCSI, FC, FCoE, LUNs
- POSIX file system interface, for example NFS, CIFS, WebDAV
- Object storage accessed via a RESTful HTTP data path interface
- Sharing of structured data access interface, for example database access interface JDBC, ODBC, XML, and so on.

9.1.3 Management requirements

The major managed objectives are described as follows:

- **user management**

Cloud storage needs to manage users according to the specific needs of the application and the actual business, including: add users, delete users, change user description, modify user password, initialize user password, modify user levels, retrieve user information, user login, user log off, and user authentication.

- **authentication management**

Cloud storage needs to have a user authentication and identification system that allows legitimate users to use storage systems, while preventing unauthorized user access to the system, and preventing illegal operation by legitimate users.

- **device management**

Cloud storage needs to provide a device and resource management function which can manage all the storage nodes, hard disk, CPU, memory, power supplies, fans, network cards, and even switches and other network equipment, while monitoring the status of the system, including storage devices, network equipment, etc.

- **configuration management**

Cloud storage provides basic configuration functions, including storage domain configuration, file system namespace configuration, the storage node configuration, the local file system configuration, and even the network device configuration.

- **status monitoring and statistic**

Cloud storage needs to provide performance monitoring and statistics. The system needs to monitor in real-time the performance of read and write, disk space usage, CPU utilization, memory utilization, job completion, etc. The system can send out a warning when software or hardware has faults, and collect statistics to present to the user.

- **data services and storage applications**

Cloud storage needs to provide a means to apply data services, such as backup, replication, archiving, and retention, to the data that is stored in the cloud. The best means to accomplish this is to use metadata to express these requirements to the cloud for each item of data.

9.1.4 Availability requirements

Cloud storage is essentially to provide data storage capabilities. Authorized users can get storage space (disk volume, formatted file system namespace, or database storage space) through legitimate channels, to access stored user data without any time and place limits.

Cloud storage needs to provide data backup and recovery abilities. The data backup function can be synchronous or asynchronous to data generation and deletion. When one of the storage nodes or hard disks fails, or there is data loss or destruction, cloud storage can automatically detect failure, and recover data using backup in other storage nodes.

Cloud storage is essentially to provide cloud data verification capabilities. Generally, cloud storage stores data into data blocks which need backup to different location. Cloud storage needs to support data block verification and synchronization to keep all these data blocks consistently with the same data.

9.1.5 Scale-out storage requirements

Traditional block storage architectures are built on a single client to server pair interaction which does not meet the needs for scale and elasticity in the cloud. The scale-out storage needs to happen inside data centres, or across two geographically separate data centres, for both capacity and elasticity.

In addition to the standard functionality expected from a cloud storage solution, such as local and remote storage access interfaces, the dynamic provisioning of storage and storage users, and low power/high capacity storage scaling to petabytes (PBs) and beyond, the scale-out storage is recommended to provide a complete set of tools to automate management of the cloud storage. These tools include a high-availability metadata service that supports user-defined policies and self-management of storage placement, reliability, compression, and de-duplication.

9.2 Functional architecture for cloud storage

The traditional storage system utilizes tightly-coupled symmetry architecture which aims to work out high-performance computing problems and could scale out to cloud storage. The next generation system adopts loosely-coupled asymmetry architecture which centralizes metadata and control manipulation. This architecture is not suitable for high-performance computing, but is designed to solve large-capacity storage requirements, based on cloud deployment.

- Tightly-coupled symmetry architecture

In tightly-coupled symmetry architecture, many nodes have the functions of distributed lock service (lock the write operation of one file in different parts) and catch consistency. This scheme could solve the throughput issue of the single file.

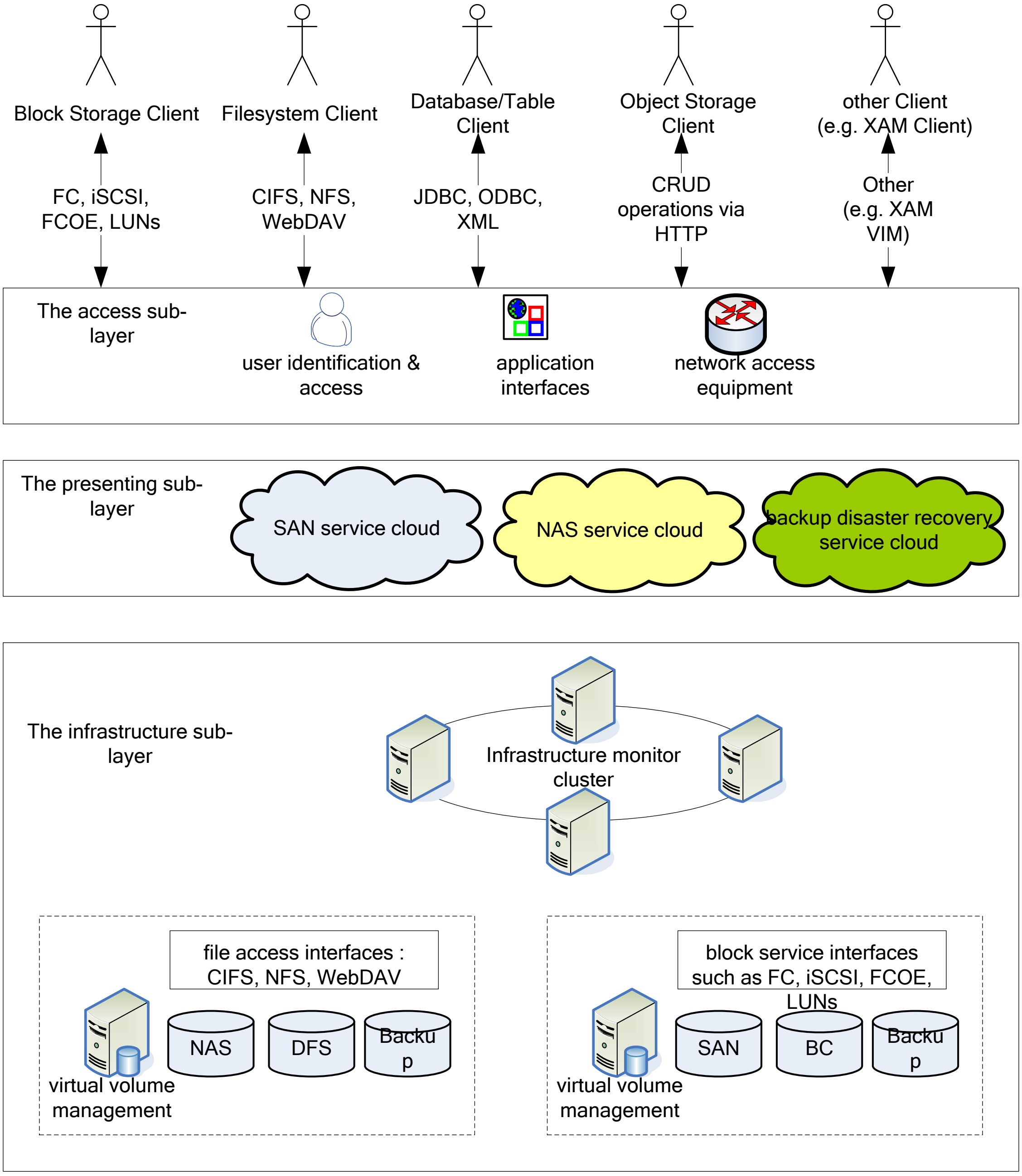
- Loosely-coupled asymmetry architecture

Loosely-couple asymmetry architecture introduces a different scale out method which does not make each node know each action via the specific policy, but makes use of the central metadata control server outside the data path. This scheme allows scaling out in a new layer.

9.2.1 Layered reference model for cloud storage

Cloud storage delivers virtualized storage on demand over the network, based on a cluster, grid and distributed file system. When storage and management of large-scale data is the key issue of operation and processing in cloud computing, a large number of storage equipment needs to be deployed. Hence, cloud storage is a kind of cloud computing system based on data storage and management.

Cloud storage is the cooperating operation of multiple-storage devices, multi-applications and multi-services. A single storage system cannot be called cloud storage. A cloud storage system can provide services such as SAN, NAS, data backup, and disaster recovery as well.



NOTE: The interfaces and protocols shown inside this figure are used as an example.

Figure 9 - Cloud storage reference model

Figure 9 plots the reference model for cloud storage which is divided into three sub-layers: infrastructure, presenting, and access. The infrastructure sub-layer is the core of cloud storage. The presenting sub-layer can provide SAN, NAS, data backup, and the disaster recovery service, based on the stratum infrastructure. The access sub-layer realizes the connecting service with application software and related hardware.

9.2.2 The infrastructure sub-layer

It is the part of data storage in the system which consists of three parts, as follows:

- **Storage infrastructure** is composed of fibre channel storage devices, NAS and iSCSI IP storage devices. It is suggested to deploy it in several working nodes within a distributed network to support high availability and reliability. The working node can include a virtual volume management element, NAS, and a distributed file system (DFS) device to provide file access interfaces such as CIFS, NFS, and WebDAV. Another type of working node can include a virtual volume management element, SAN, and a block control (BC) device to provide block service interfaces, such as FC, iSCSI, FCOE, LUNs.
- **Backup infrastructure** is composed of type library, virtual type library, CD database and related software.
- **Infrastructure monitor cluster** is composed of many servers which manage and monitor all kinds of storage and backup devices, repair related links, check the redundancy, and carry out centralized management. It can support the global schedule function of cloud storage to locate the resource location in the storage infrastructure, according to the accessing requests and requested resource. The servers shall support the distributed hash table (DHT) (networking to provide a general accessing interface of name space management, load balance, metadata management, routing management, and duplication management. It can access the virtual volume management element of storage infrastructure to realize unified volume management and policy management.

There are usually many devices in the storage and backup infrastructure and, because of the wide distribution, they are connected with the IP network (such as WAN and LAN) and FC SAN.

9.2.3 The presenting sub-layer

The presenting sub-layer is the core of service logic and consists of several service clouds which are related to storage, such as the SAN service cloud, NAS service cloud, and the backup disaster recovery service cloud.

SAN and NAS service clouds provide key storage services. The two clouds manage cloud storage logically, detect and repair the faulty links, and offer status monitoring and QoS.

The backup disaster recovery service cloud supplies high level data protection so that it is unnecessary to establish a special disaster recovery network.

9.2.4 The access sub-layer

The access sub-layer consists of an application interface, network access equipment, user identification and relevant access functions.

The privileged user could be authorized to log in and make use of the cloud storage service via a standard public application interface and storage protocols, such as CIFS, NFS, FC, iSCSI.

It can connect with the infrastructure sub-layer through the private or public network. It supports the sending of user requests to the infrastructure monitor cluster, and has direct access to resources in the storage infrastructure according to location information, which is located by the infrastructure monitor cluster.

The user can visit the access sub-layer through four client types:

- **Block Storage Clients** support interfaces like FC, iSCSI, FCoE, LUNs
- **Filesystem Clients** support interfaces like CIFS, NFS, WebDAV
- **Database/Table Clients** support interfaces like JDBC, ODBC, XML
- **Object Storage Clients** support interfaces like CRUD operations via HTTP, such as SNIA CDMI

- **Other Clients: e.g., SNIA XAM Clients,** support interfaces like XAM VIM for SNIA CDMI

9.2.5 Data verification functional architecture

The data verification functional architecture is shown in Figure 10. Cloud storage delivers virtualized storage over the distributed file system. The distributed file system data verification device is composed of many servers, which include: the metadata server, data block servers and storage media.

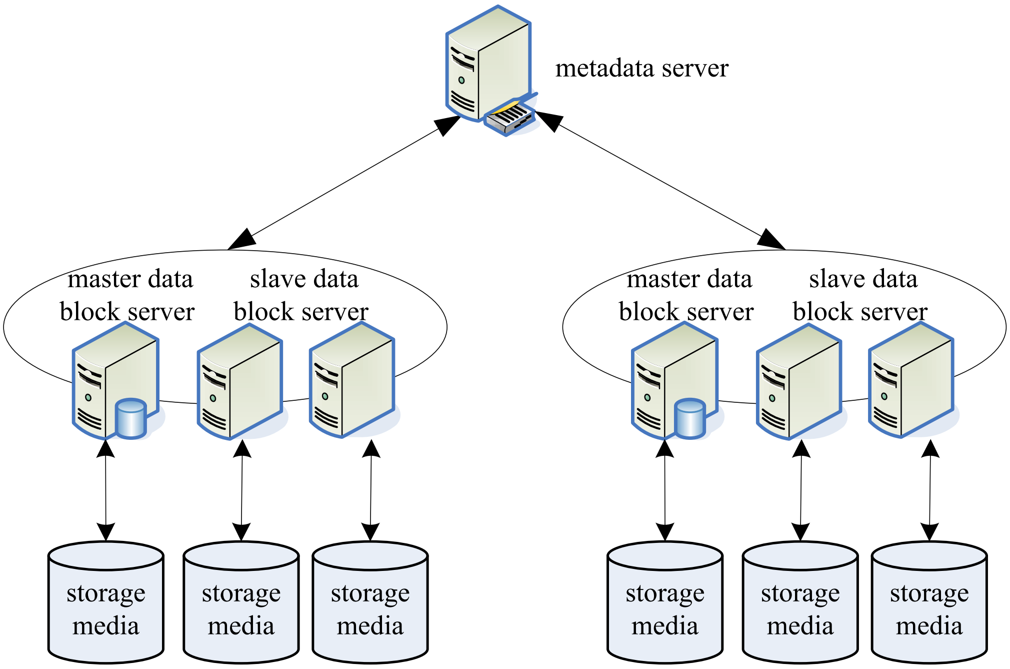


Figure 10 - Data verification functional architecture

9.2.5.1 The metadata server

The metadata server is the core of data verification It is in charge of administration of all file metadata, including file name, data block, and mapping information in the distributed file system. It can assign the master data block server and slave data block servers in a group.

It supports sending of a data block verification request to the master data block server by timer-triggering, and updates metadata information according to the return result of verification and synchronization.

9.2.5.2 The data block server

The data block server is in charge of interacting with storage media and performs data block write/read operations. There are master data block and slave data block servers. The master data block server will verify all data block information managed by slave data block servers in its group, synchronize with the verified result, and then return the result of verification and synchronization to the metadata server.

The verification and synchronization include several steps:

- The master data block server obtains all data block servers' information from the data block verification request of the metadata server.

- The master data block server sends a data block collection request to the slave data block servers in its group.
- The slave data block servers report managed data block information to the master data block server. The master data block server will then store this information into the buffer.
- The master data block server finishes verification after obtaining all data block information from the slave data block servers. The verification will verify the consistency of the master data block and the slave data blocks.
- When the verification result is inconsistent, the master data block server will synchronize the inconsistent data from the master data block server to the slave data block servers. The synchronization includes full duplication or partial duplication.

9.2.5.3 The storage media

The storage media consists of storage devices, such as disk, tape, and so on. The storage media is used to physically store data.

10. Functional requirements for resource management

10.1 General requirements for resource management

10.1.1 Physical and virtual resource management

Resource management is required to support computing, storage, and network resources management. For example, for resource template management, it is required to have the capability to describe all kinds of resources, including computing, storage, and network resources.

Resource management is required to support both virtual and physical resources. For example, resource packaging management is required to provide a unified interface of heterogeneous resources, whether it is virtual or physical. Resources can be managed hierarchically to satisfy traditional applications/services or an enterprise hosting environment.

10.1.2 Heterogeneity shielding

The cloud resource management is required to shield the heterogeneity of cloud resources for the user. It is required to provide the user with a unified interface so that he can use all the heterogeneous resources without caring about their real type.

To simplify resource management, it is recommended to support a unified resource management interface between different types of hypervisors and the cloud resource management so as to integrate different types of heterogeneous resources.

10.1.3 Dynamicity shielding

The cloud resource management is required to shield the dynamicity of cloud resources for the user. The dynamicity of cloud resources means that their performances change at any moment. The cloud resource management is required to be able to evaluate the performance of each resource to fulfil the QoS of each user request.

10.2 Functional requirements for resource management

Cloud infrastructure should provide a unified resource management function for the upper-layers (including virtualized and physical computing resource, storage resource, and network resource). The resource management function should provide resource packaging, resource deployment, and resource scheduling, whilst managing templates and assets. Resource packaging provides a unified interface of heterogeneous resource, whether virtualized or physical, to upper-layers for management and utilization. Resource deployment and scheduling provide elastic, dynamic, on-demand and automation management for the down-layers, based on user-defined policies. They also provide resource access control interfaces to the upper-layers, and can dynamically allocate the resource by the real-time monitoring of applications and SLAs. Template management provides the capability to describe groups of computing, storage and network resources within their life cycles. With the template management, resources can be easily allocated and deployed to satisfy application/service resource demand of the upper-layers. Asset management provides unified management of the physical devices, including configuration information and topology of assets.

10.2.1 Resource encapsulation

It is recommended that all the physical and virtual resources be managed in a unified manner through the resource encapsulation. The heterogeneous resources are integrated and provide a unified interface for the upper layer to create, locate, deploy, provision, recover and delete the resources. The attribute of each resource, including resource deployment, status, capacity, execution, exception, error and interrupt, can be measured and searched.

10.2.2 Resource orchestration and provisioning

All resources are recommended to be flexible, on-demand and automation orchestrated, deployed, and provisioned, based on the pre-defined policies that include high availability, load balance, resource migration, energy efficiency and storage deployment. It should be possible for the services to be analyzed and to be translated into resource requirements and to trigger appropriate actions.

Resource deployment and scheduling management is recommended to provide on-demand and automation management for the lower-layers based on pre-defined policies, and access control interfaces to the upper-layers. It can dynamically allocate the resources by real-time monitoring of applications and SLAs. It is composed of resource deployment and strategy management.

Resource packaging management is required to provide a unified interface to upper-layers for management and execution. Resource packaging is composed of asset auto discovery, resource-acceptance management, resource topology management, resource pool management, and API management.

10.2.3 Assets management

It is recommended that assessment attributes and topology of physical devices be managed in a unified manner. The assessment attribute can be divided into hardware and software assessment. Hardware assessment means racks, servers, storage devices, network equipment, and VMs, while software assessment means hypervisors, operating systems, middleware, databases, applications, licenses, and so on. It is recommended to automatically update the assessment attribute when the physical devices are changed. Multi-layer topology can be searched automatically or manually.

Asset management is required to provide unified management of the physical devices, including asset information management and asset topology management.

10.2.4 Template management

Template management is one of the important functions of resource management. The resource template provides the life cycle management of each template, including creation, publication, activation, revocation, deletion, template provision, etc.

Resource template management is required to provide management of life cycles, including creation, publication, activation, revocation, deletion, template provision, etc. Resource template management is composed of template management and image management.

10.2.5 Cloud service monitoring

Monitoring is essential to ensure the availability, security and usability of cloud services. The cloud computing system introduces many new challenges with new technologies being deployed. First, server virtualization introduces a new layer that needs to be monitored - the hypervisor. At the early beginning, the primary objects being monitored were applications, services, operating systems, and physical infrastructure, including storage and networking. When virtualization is introduced, the operating systems are working with virtual resources and the hypervisor adds a new dimension to monitoring. Second, the distributed computing and distributed file system（storage）system consist of multiple independent working nodes, and clustering software or programs made these nodes work together to finish the jobs or tasks, thus monitoring for distributed system concerns not only the individual nodes but also the status and performance of the whole cluster.

Cloud infrastructure environments consist of a number of physical and virtual entities such as physical servers, virtual machine monitors/hypervisors, virtual machines (VMs), physical and virtual disks, physical and virtual network, and applications. All of these elements are associated through complex relationships. Data is captured for analysis and measurement by monitoring various attributes of these elements and their relationships. The following aspects are the main functional requirements of cloud service monitoring and usage metering.

The real-time monitoring is recommended to be executed for all physical and virtual resources. The architecture of resource monitor is multi-layered, including service instance monitoring, physical resources monitoring, resource pool monitoring, user connection monitoring, software monitoring, etc. It is recommended that the system detect the exception or error of computing, storage, network equipment, and the resources pool while the currently connected number and the users' IP address, login time, idle time, etc., are monitored.

10.2.5.1 Health monitoring

Monitoring the health/status of the complete infrastructure requires monitoring of the physical server hardware status, hypervisor status, virtual machine status, physical and virtual network switches and routers, and storage systems.

Cloud service needs an integrated approach that combines a top-down and bottom-up approach to monitoring and spans both the physical and virtual infrastructure. This requires a consolidated operations bridge that can help IT track and respond to infrastructure events occurring on the ground floor of IT (bottom-up), and maintain service health in real time by monitoring the end-user experience (top-down). Requirements include an accurate run-time service model and tools that enable collaboration, support automation, and facilitate cloud computing.

The service model can be regarded as a map which displays all of the technology components, including transactions, applications, web servers, network switches, virtualized components, and third-party cloud services. Having such a model can play a critically important role in effective business service management because when there is an application or transaction problem, it can help pinpoint the infrastructure components that may be playing a role in service disruptions.

In addition, the service model is recommended to provide run-time monitoring, because the service infrastructure is constantly changing. It is necessary to ensure the "currency" of this service model on a continuous 24/7 basis so that an accurate service definition can be used to troubleshoot service problems and manage service level agreements. Nonintrusive probes can be used to automatically detect infrastructure, application, and transaction changes in near real time.

Supporting heterogeneous and comprehensive management of virtualized services, for example, end-user monitoring, can help IT monitor target service levels. When performance drops below predefined thresholds, IT can trace the problem back to its source for rapid resolution.

An integrated operations bridge consolidates event and performance data from both physical and virtual sources to reduce duplicate monitoring and boost productivity. Automated remediation capabilities can help reduce mean time to repair (MTTR).

10.2.5.2 Performance monitoring

Basic performance monitoring looks at the CPU, memory, storage and network performance metrics from the VM guest OS as well as from the hypervisor. These metrics typically get monitored even in non-virtualized environments. The virtualization-specific metrics could be for specific entities that are introduced by various virtualization technologies, e.g., the cluster and data enter concepts in VMware. The behaviour of other virtualization features can also be measured as metrics, such as how frequently VM migrations are occurring or when other availability features are engaged. Then there are specialized applications built by virtualization, for example, desktop virtualization (VDI). Monitoring for such solutions requires more parameters to be collected from the virtual machine as well as the hypervisor layer, for example, how quickly VMs are provisioned to a requesting end user.

10.2.5.3 Capacity monitoring

Modern organizations are truly dynamic and their IT resource utilization/requirements are continuously evolving. So, continuous planning of various resources such as servers, desktops, network, storage, and also many kinds of software is required. This requirement demands periodic audits of physical as well as virtual resources. The capacity monitoring requires end-to-end continuous capacity monitoring of the following key metrics:

Server utilization: Peak/average server resource utilization - memory /CPU/resource, server bottlenecks and correlation with a number of users/VMs.

Memory usage: Memory utilization on each server, capacity bottlenecks and relationship with number of users/VMs and with different cloud services.

Network usage: Peak/average network utilization, capacity/bandwidth bottlenecks and relationship with a number of users/VMs and with different cloud services.

Storage utilization: Overall storage capacity metrics, VM/virtual disk utilization, I/O performance metrics, snapshot monitoring and correlation with a number of users/VMs and with different cloud services.

10.2.5.4 Security and compliance monitoring

Virtualization introduces a new set of security risks due to VM sprawl, and the introduction of new threat targets - the hypervisor layer, VI configurations, and potential conflicts in the way access control is managed and policies are applied. IT security and compliance monitoring becomes critical for securing the virtualized environment. Security and compliance monitoring requires end-to-end VI activity monitoring for:

VM sprawl: Metrics to monitor the VM activities as they get cloned, copied, V-Motion-ed within VI, move of network, move to different storage media.

Configuration metrics: Virtual server configuration monitoring to ensure that they are compliant with standards and hardening guidelines, VM configuration monitoring for software licensing policy enforcement. VI Events that help enforce/detect violations of IT policy. This includes individual security and organization security policy monitoring.

Access control: Access control monitoring and reports for role-based access control enforcement.

Compliance monitoring: Metrics to validate/audit IT for standards such as HIPAA, SOX, ITIL and GLBA.

10.2.5.5 Monitoring and metering for charging and billing

In a virtualized environment, where the infrastructure is centralized, it is important to measure resource usage by different business units, groups, and users. This information can be used to distribute/amortize and, in some cases, recover the cost correctly across the organization through a proper chargeback mechanism. Chargeback could be based on dynamic parameters such as resource usage and/or fixed parameters. To compute the correct chargeback information in a dynamic virtualized environment, it is important to monitor virtual/physical resource usage and allocations and be able to normalize the measurement across the cloud infrastructure. The monitoring and metering data for service charging should be collected and well kept for a period of time as needed. For more information about cloud service metering please refer to Appendix III.

Chargeback monitoring requires end-to-end VI (virtual infrastructure) activity monitoring and service usage metering for:

Standard metrics: All chargeable resource metrics like CPU usage, memory usage, storage usage (volume and time), and network usage (bandwidth and network traffic).

Key VI events: VI events for virtual resource life cycle events like start date and end date of VM creation and allocation.

Configuration monitoring: VM configuration in terms of assigned resources and reservations and also applications installed to an account for software licensing costs.

VM usage metrics: VM uptime, number of VMs can vary, depending on how the charging model is employed in the organization.

10.2.5.6 Application and service monitoring

The need for application and service monitoring is important in the cloud computing environment, especially for the SLA/QoS evaluation aspect because the application or service may have problems even if the VM or the physical server on which it is running looks perfectly normal. Application and service monitoring is required to monitor the basic health of application servers, with the help of application-specific response time and throughput metrics. The analytics on this data is recommended to be able to correlate the application-observed and service-observed metrics to all layers of the infrastructure to perform a root-cause analysis, in the event that something is going wrong. Application and service performance monitoring using the capture of network traffic is used more and more commonly in this area.

There are a few other aspects to virtual infrastructure monitoring that add to the complexity of building a comprehensive monitoring solution. All kinds of virtualization software (server virtualization, storage virtualization, network virtualization, etc.,) provide API to be able to collect metrics. However, each kind of virtualization software has its own object models and APIs. There are wide differences in features and even in the behaviour of the common features. Therefore, the analytics that are to be built on the collected metrics must be developed for each kind of virtualization software.

10.2.6 User resource environment management

User resource environment includes the resources allocated to a user, the state of the resources (such as running, stopped for a virtual machine), and the topologies among the resources.

Cloud infrastructure is required to ensure the secure isolation between different user resource environments; it should not happen that the management activities in one user resource environment impact other user resource environments.

A user is recommended to have appropriate control of his user resource environment, which is managed by cloud provider, or IT administer for enterprise users. For instance, a user may shut down a virtual machine, but cannot release it.

If cloud support is provided, a user can release his user resource environment as a whole, and recover it when needed. This is useful for users who need to apply and release the cloud resources periodically. For example, if a user wants to utilize cheap resources in the night and release them in the daytime; the easy recovery of his user resource environment will avoid him the trouble of applying and organizing the same resource environment every day.

10.3 Cloud resource management model

The cloud resource management is responsible for the entire process of resource allocation, authentication, accounting, and related security, in the cloud. Such a process covers tasks that range from the automatic negotiation of developers requirements to the execution of their applications. It has four main modules: negotiation, resource allocation, access and control, and resource security. Figure 11 shows the framework of resource management.

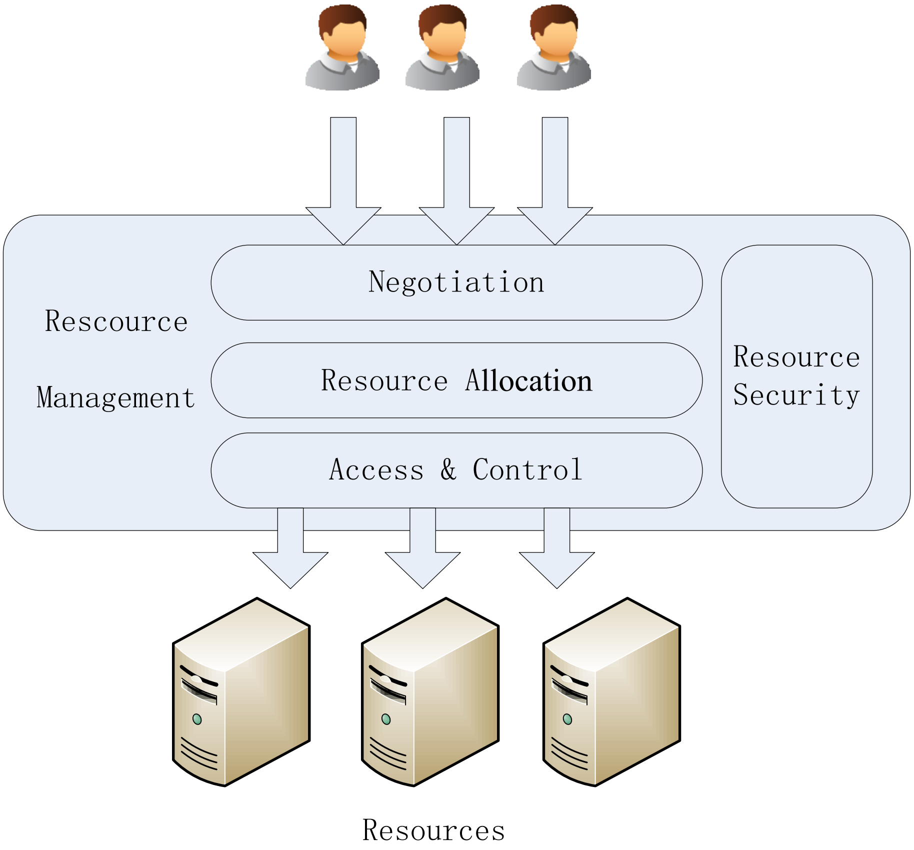


Fig 11 – Resource management framework

The negotiation module deals with the interface between the resources and network layer and the service layer. Different cloud service model implementations could use the resources layer differently, but the interfaces exposed can be defined in a common way to accommodate all service models implemented at the higher levels of the cloud reference architecture..

The resource allocation module is responsible for the optimal allocation of applications for obtaining the maximum usage of resources. This function requires advanced strategies and heuristics to allocate resources that meet the contractual requirements, as established with the services layer. These may include service quality restrictions, jurisdiction restrictions, elastic adaptation, and so on.

The access and control module encompasses all of the functions needed to enforce decisions generated by the other modules. Beyond the tools used to configure the cloud resources effectively, all communication protocols used by the cloud are included in this module.

The resource security module is responsible for the user authentication, identity and access management, resource monitoring and auditing for compliance, the patch and update status of software, and so on. For detailed information, refer to the Security Technical Report.

10.3.1 Resource allocation

Resource allocation mechanisms provide an automated provisioning of resources while aiming for the best utilization of available resources, as shown in Figure 12. Along with this guarantee to the services layer, resource allocation mechanisms are recommended to consider the current status of each resource in the cloud environment, in order to apply algorithms to better allocate physical and/or virtual resources to the upper layers, thus minimizing the operational cost of the cloud environment.

Resource allocation is divided into four categories:

- resource modelling and description,
- resource offering and treatment,
- resource discovery and monitoring, and
- resource selection.

The development of a suitable resource model and description is the first step to describing the resources present in the cloud. Resource offering and treatment is to represent the applications requirements. The automatic and dynamic resource allocation must be aware of the current status of the cloud resources in real time. Thus, mechanisms for resource discovery and monitoring are an essential part of resource allocation. These two mechanisms are also the inputs for optimization algorithms, since it is necessary to know the resources and their status in order to elect those that fulfil all the requirements.

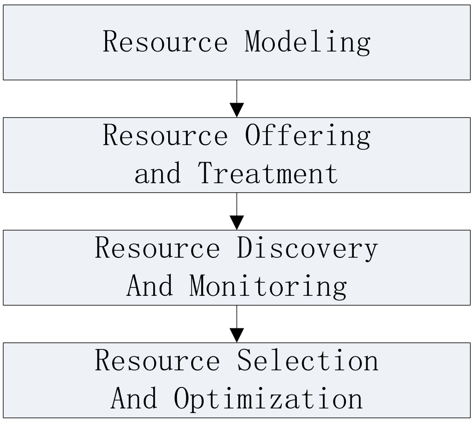

Resource Modeling
Resource Offering and Treatment
Resource Discovery And Monitoring
Resource Selection And Optimization

Figure 12 – Resource allocation mechanism

10.3.1.1 Access and control

Resource management provides an efficient and effective way to access, control, manage, deploy, schedule, and bind resources when they are provided by resource providers and requested by resource clients. The architecture of the resource management system includes the functional entities as listed in Figure 13.

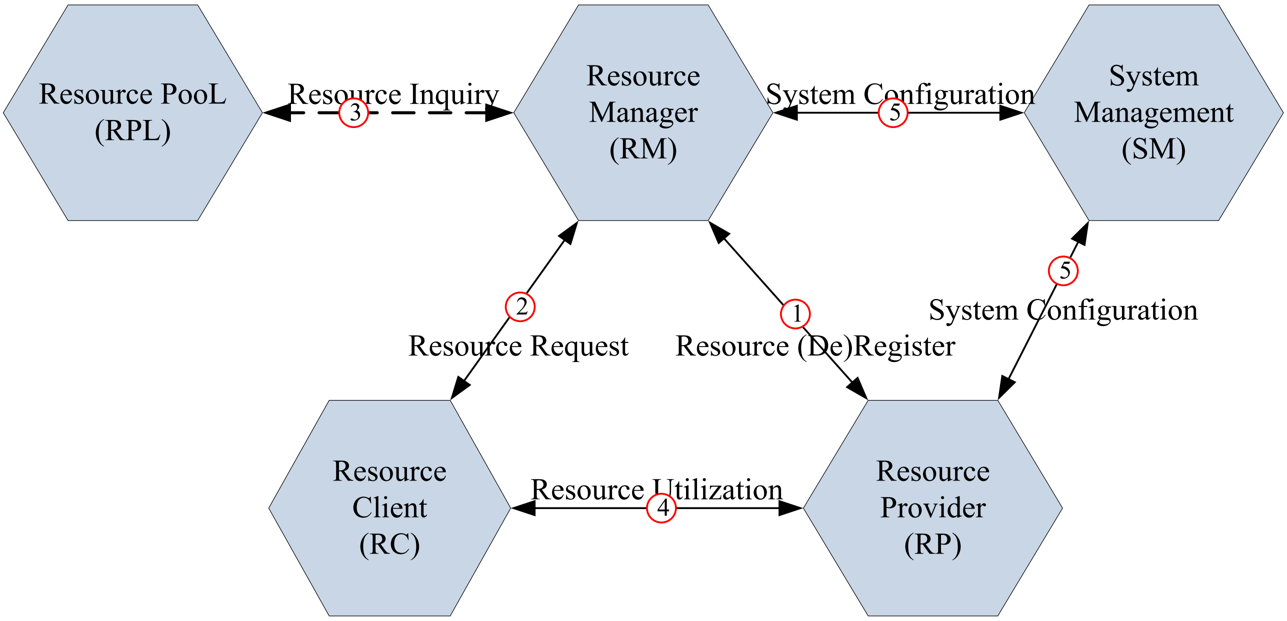


Figure 13 – Overview of resource management model

The following five general function entities exist in the architecture.

Resource manager (RM).

RM provides registry of virtual resources and allocation of virtual resources to resource clients (consumers).

RM has its focus on providing flexibility, transparency and reliability.

RM interacts with resource provider (RP) to implement the resource registry and resource allocation function. When powered on, the hypervisor implementing resource virtualization or the management agent running on the device, are responsible for the resource registry. RM reports registry information, such as resource type, available capacity, host address, etc. RM records the registry information and classifies the resource into a resource pool. Based on service/application requirements to resource, RM performs the resource allocation procedure and allocates a suitable amount of resources to satisfy service/application deployment and running requirements.

RM interacts with the resource client (RC) to provide a resource allocation, resource update and resource query function, as well as resource template upload, update and query functions.

RM needs to provide a unified interface to the resource client, and also needs to provide a unified interface to manage heterogeneous resources provided by resource providers.

Resource provider (RP)

RP is the entity which provides virtualized resources to the resource management platform. In the architecture above, it mainly refers to the hypervisors (virtualized case) and hosts (traditional non-virtualized case).

RP represents a collection of devices providing virtual or physical resources, such as, hosts, virtual machines, storages, memories, and network related resources.

RP needs to support a unified interface to register available resource and provide resources according to a resource request from the resource manager.

Resource client (RC)

RC refers to the entity which consumes resources to perform any services or applications.

RC, who is the service provider or consumer consuming resources to deploy his services/applications, is the recipient of the allocated resources. RC uses the unified interface supported by RM to request required resources.

System management (SM)

SM provides monitoring and administration of the RMP with the objective of keeping the system operating normally.

SM provides monitoring, logging, auditing and administration of the resource management platform, with the objective of keeping the system operating normally.

SM needs to support both traditional network management and virtual network management.

Resource pool (RPL)

RPL is a collection of virtualized resources abstracted from physical resources, and it can be organized hierarchically to partition the available virtualized resources.

Resources in the RPL can be used exclusively by one RC or can be shared among several RCs.

A set of allocated resources from (child) RPL with its setting data can be used to constitute virtual DC, virtual LAN, virtual SAN, or virtual network, etc.

RPL is the inventory of the registered resources, which records resource status, such as configurable capacity, base capacity, available capacity, allocated capacity, consumed capacity and reserved capacity, sub-resource pools information, corresponding host, sharing or exclusion, etc.

RPL organizes resource hierarchically and partitions the available virtualized resources into groups. The grouped resources can be used to constitute a system, such as virtual DC, virtual LAN, or virtual SAN, etc.

11. Power management

Data centres are amongst the world's highest consumers of electricity . One distinct advantage of cloud computing is its ability to power-manage hardware and devices, i.e., devices are recommended to be dynamically powered on and off. Data centre resources are often arranged in trees. As resources become idle, it can opt to power off complete twigs or branches on trees. As data centre resource usage trends are measured by orchestration, it would be possible for these networks to put energy back into the grid by giving it accurate time-based predictions of energy use. The grid can use this information to redirect energy to other destinations, or take other intelligent decisions.

Power management represents a collection of IT processes and supporting technologies geared toward optimizing data centre performance against cost and structural constraints. This includes increasing the deployable number of servers per rack, when racks are subject to power or thermal limitations, and making power consumption more predictable and easier to plan for.

Power management comes in two categories: static and dynamic. Static power management deals with fixed power caps to manage aggregate power, while policies under dynamic power management take advantage of additional degrees of freedom inherent in virtualized cloud data centres, as well as dynamic behaviours supported by advanced platform power management technologies.

For further information, refer to Appendix IV.

12. Security considerations

Refer to Part 5 of this Technical Report.

Annex A
Mapping between this Technical Report and Technical Report for RA of cloud

Technical Report *Requirements and framework architecture of cloud infrastructure* , based on cloud service data connection and transmission point of view (telecom and network operator), focuses on L2/L3 network transportation ability, connectivity, network infrastructure requirements and reference architecture to connect different cloud services, including intra-cloud networks and inter-cloud networks, to effectively and efficiently support cloud providers.

Technical Report *Functional requirements & reference architecture*, based on cloud service consumption (service user) and cloud service operation (service provider) points of view , introduces L4 and layers above data transmission, including functional architecture, functional entities and reference points. The mapping between the two Technical Reports could be plotted as follows:

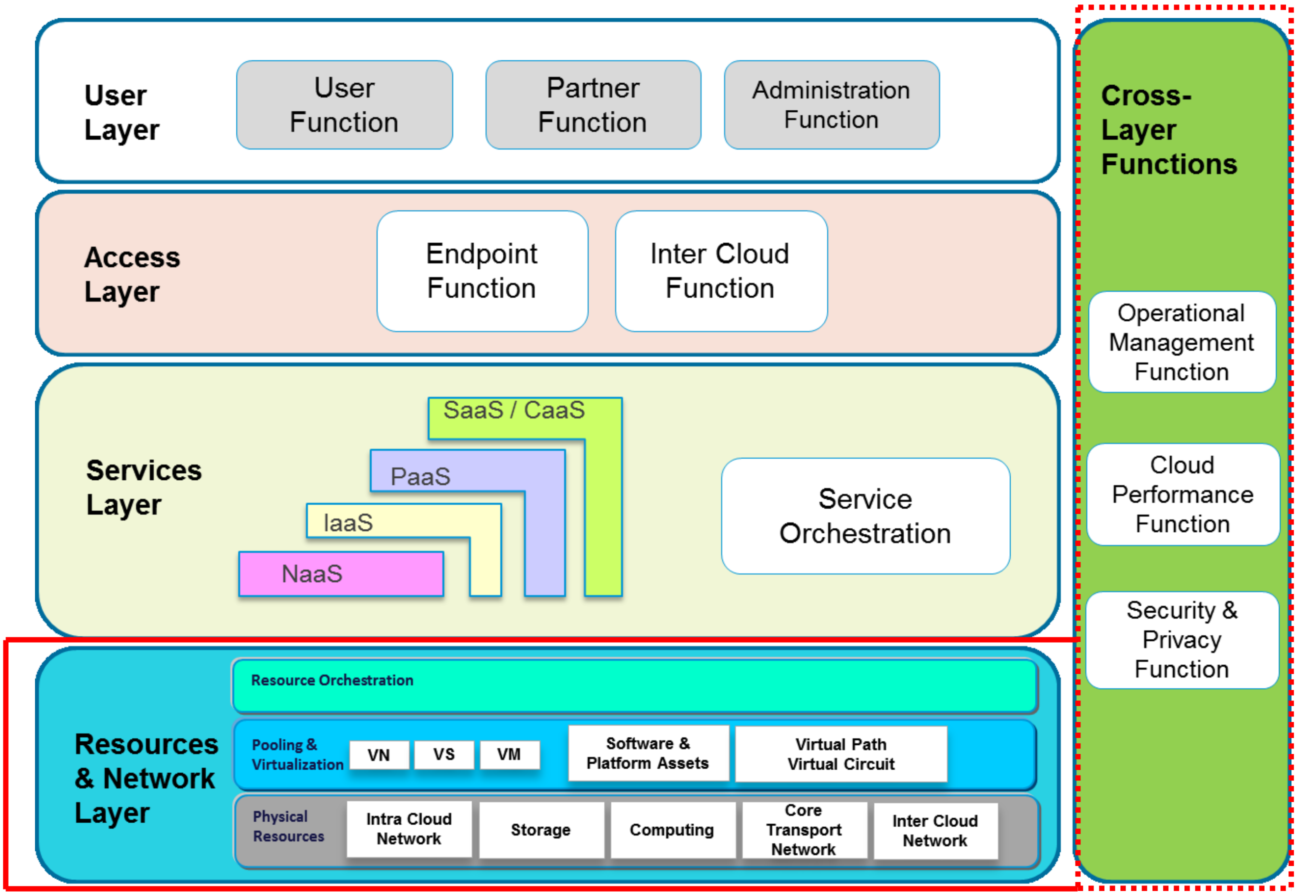


Figure A.1 – Mapping between infrastructure and RA

The cloud computing functional reference architecture proposed by RA Technical Report defines the layering framework, while the infrastructure Technical Report focuses on the requirements and interaction of specific functional blocks in the resources and network layer and cross-layer functions.

Annex B
Convergence of data and storage networks

Many connectivity protocols and separate networks for handling network traffic and storage protocols run within data centres. This creates redundancy and potentially underutilized connectivity while, at the same time, increasing cost in additional ports/links and management complexity at scale. Unified fabric seeks to minimize the number of ports required to the fewest links for efficient performance and minimum cost, while consolidating the dominant network (ex.IP) and block and file storage protocols (ex. FCoE, iSCSI, NFS, CIFS) onto L2 backplane (ex.10 Gb Ethernet).

Figure B.1 describes the convergence of data and storage networks using FCoE, which reduces switches and lowers the complexity of management.

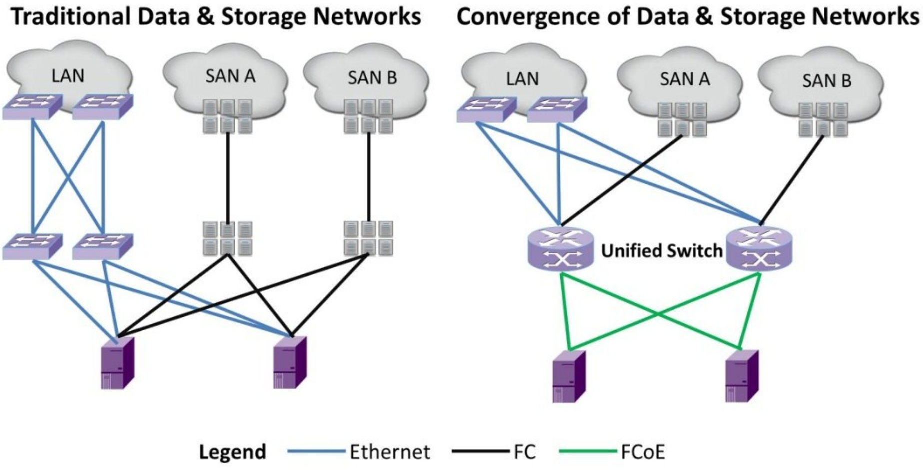

Traditional Data & Storage Networks
Convergence of Data & Storage Networks
LAN
SAN A
SAN B
LAN
SAN A
SAN B
Unified Switch
Legend
Ethernet
FC
FCoE

Figure B.1 – Convergence of data and storage networks

Annex C
Network topology inside the virtualized data centre

The network topology inside the virtualized data centre can be divided into two kinds.

Option 1: based on multi-paths on data link layer

Current data link layer is too easy to plan a suit of data-forwarding schemes based on the learned address, because there is only one data plane but not a control plane. This causes switches on data link layer that cannot execute "routing" and brings a series of problems such as STP.

This option rebuilds a control plane in order to support data centre extension based on multi-paths on data link layer, such as TRILL (IETF) and 802.1aq (IEEE), as shown in Figure C.1. This control plane comprises the following basic functions:

(1) Maintains a routing database based on link state;

(2) Supports agile addressing;

(3) Supports equivalent routes;

(4) Keeps simplicity of original data link layer.

Figure C.1 describes the typical TRILL or 802.1aq networking which is deployed between access and aggregation devices with no STP, and multi-links can forward data.

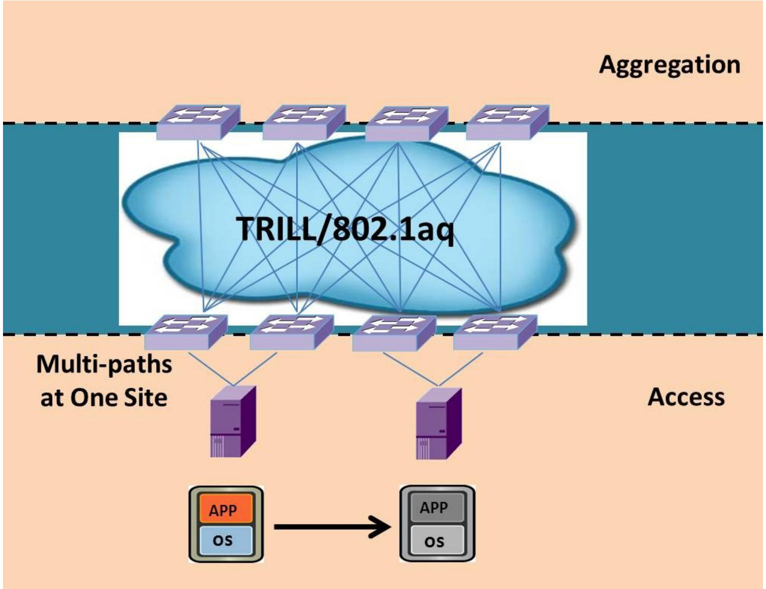


Figure C.1 – Future network topology inside a data centre based on multi-paths on the data link layer

Option 2: based on logical switch for an entire data centre

This option proposes one logical switch for an entire data centre as shown in Figure C.2. It connects thousands of devices but behaves and is managed as a single switch. It is totally different from the legacy network method and presents a revolutionary thinking. All ports are directly connected to all other ports, meaning that no matter where an application is located, it is only one hop away from its resources. This scheme shows there is one device to manage and no interactions within the fabric itself. TRILL, 802.1aq and STP will not run on the fabric, so there are no loops.

This scheme brings unprecedented simplicity and agility in managing the cloud to improve the end user experience and data centre economics. It delivers benefits that simplify server and storage administration and make business processes more agile.

Because it is a radical departure from the traditional data centre and needs to replace the original network architecture, it is suitable for the data centre to be built in the future but not for transformation on the basis of existing ones.

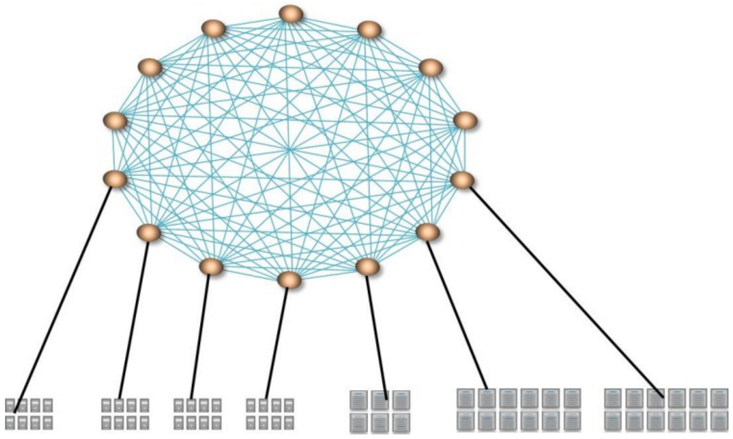


Figure C.2 – Logical switch for an entire data centre

Appendix I
Implementation examples of cloud infrastructure

I.1 Example of service providing with an integrated service control model

This clause shows examples of service providing with an integrated service control model.

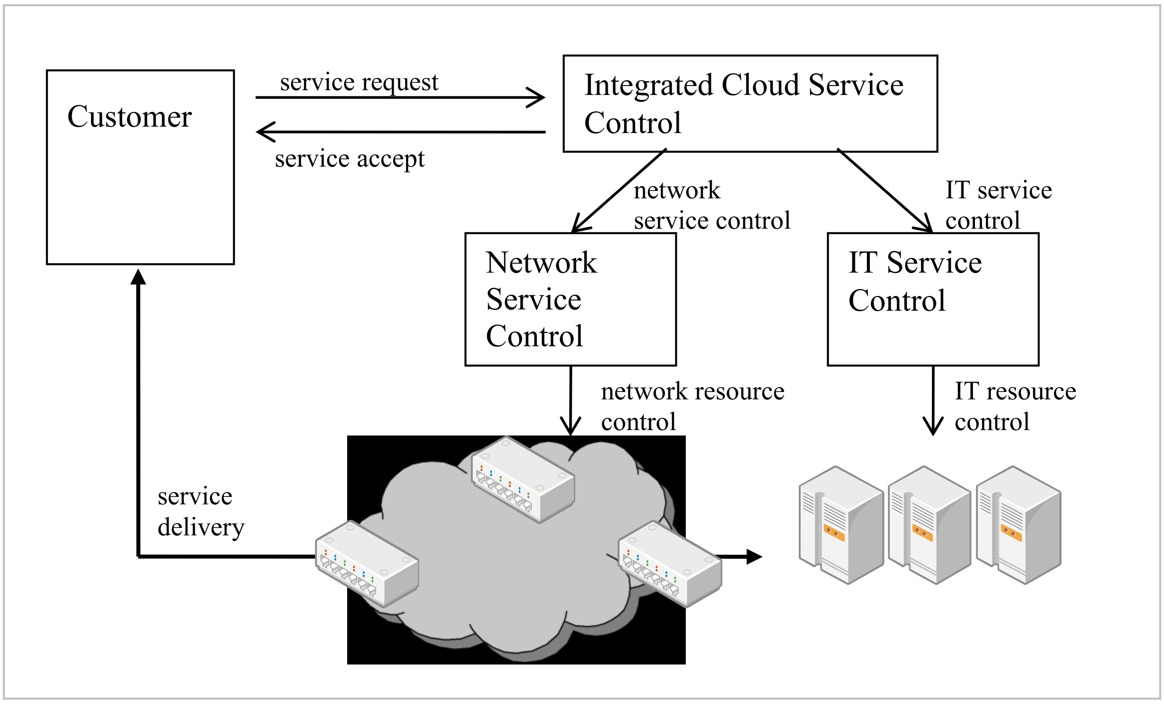


Figure I.1 – Integrated control model of network and IT resources for cloud computing services

Example depiction:

1) Service initiation

The integrated cloud service control (ICSC in this document) gathers the current status of available resources in network and IT infrastructures. When a service request arrives, the ICSC analyzes the request and determines if it can accept the request. If current available resources are enough to support the request, then the request is accepted. If not, the request fails.

2) Resource allocation

The ICSC tries to allocate the resource and activate the network and IT services. Using network service functions, the requested network sources are allocated. With IT service functions, required IT resources are also allocated. If one of the required resource allocation is failed and retry is failed (network resources or IT resources), then other resources that have already been allocated are released. This will be notified to the user with a fail message and cause of that failure.

3) Dynamic service control

The service control function can monitor and control the network and IT resources via their own control functions while the service is activated. For example, if some network resources within the path are congested, then the ICSC can choose an alternative route or allocate some of the IT resources to the other geo-location to minimize the congestion.

4) Fault handling

If some errors occur during service time, the ICSC can gather the fault log for the network and IT service. If the error occurs in the network service, then the IT resources can be moved to another location to solve the problem. Otherwise, if the problem is caused by IT resources, the IT resources can be re-allocated with unused IT resources.

Appendix II
Reference model for cloud service monitoring and usage measurement

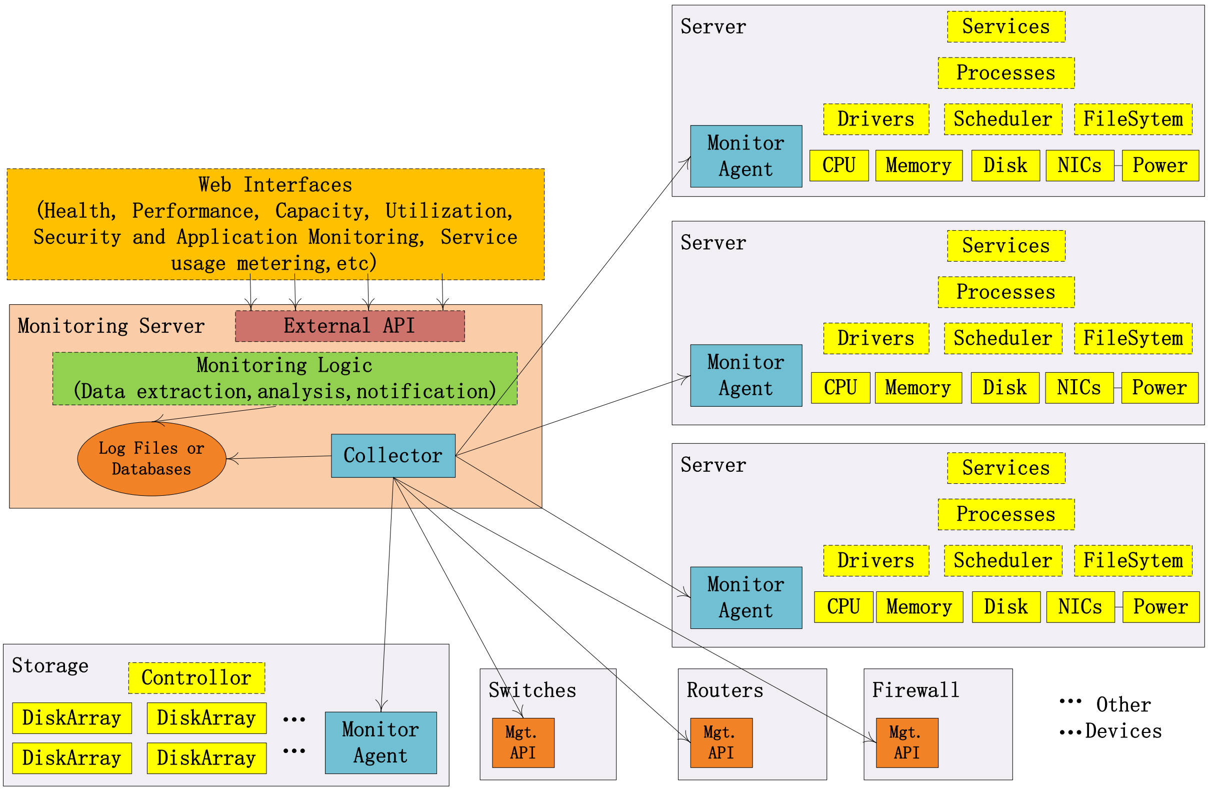


Figure II.1 – Reference monitoring model

Figure II.1 depicts reference cloud service monitoring and usage measurement model. It is the same information flow as traditional IT devices monitoring. There are three main steps:

1) System usage and performance data acquisition:

   First, the metric data of cloud infrastructure, including the metric data of the server（e.g., power, CPU, memory, disk, network cards, storage and network devices, is acquired and stored. Second, the status and performance data of system and platform software information is collected. Third, the status and performance data of services and applications is collected.

2) Monitoring data collection and reservation:

   The monitoring server is responsible for collecting the metric data of the monitored servers, storage, network devices via various monitoring protocols, such as SNMP, ICMP, etc., and for preserving the collected data in log files or monitoring databases.

3) Monitoring and measurement information presentation.

   Different monitoring services and applications, including system and application health status, performance, capacity planning, utilization, security status, and usage measurement, are presented to the end user by web interfaces.

Table II.1 shows the reference cloud service usage measurement unit.

Table II.1 – Reference cloud service usage measurement unit

Service type	Measurement dimension	Unit
Compute(CPU & memory)	Time	Hour
Disk (space or volume)	Volume*Time	GB/month
Network	Outgoing data traffic (date flow)	GB/in/month, GB/out/month
	Regional data transfer	GB/month
	IP addresses hold time	Hour
File storage	Volume*Time	GB/month
	Outgoing data traffic(flow)	GB/in/month, GB/out/month
	Operation requests	Number countered
Dedicated network (VPN)	Duration of connection	Hours/month
	Outgoing data traffic (data flow)	GB/in/month, GB/out/month
Content delivery	Volume per unit time	GB/month
	Operation requests	Number countered
Database service	Provisioned storage volume	GB-month
	Outgoing data traffic (data flow)	GB/in/month, GB/out/month
	Regional data transfer	GB/month
	I/O requests	Number countered

Appendix III
Cloud power management function

Power management comes in two flavours: static and dynamic. While static power management deals with fixed power caps to manage aggregate power, policies under dynamic power management take advantage of additional degrees of freedom inherent in virtualized cloud data centres, as well as the dynamic behaviours supported by advanced platform power management technologies.

Other SDOs are working on green technologies for data centres and cloud computing, like DMTF. For further details, please refer to relevant SDO websites.

Power management requires two important logical components in the cloud network.

1. Cloud inventory manager (CIM) – this component discovers and collects information about devices and creates an inventory. The discovery can include physical topologies which can be discovered by having networking device expose connectivity information through protocols or APIs. Host connectivity is for instance known in ARP tables. Network neighbours are discovered via protocols like CDP and LLDP. Such information is also available through SNMP based MIB operations on prior provisioned IP addresses.

2. Cloud power manager (CPM) – this component interfaces with circuit breakers, or other power distribution mechanisms. In addition, the CPM may use native power control methods on devices, for example through power MIBs, to control the level of power used by them. The CPM must support interfaces or APIs to request power management operations, including power on/off devices or putting them into lower power modes (if supported).

The CIM and CPM have to work with the cloud service broker (CSB). The broker can request inventory information from the CIM to check if there is powered off inventory available. The CSB can then request the CPM to power on the inventory. When the inventory is powered on, the CSB may itself learn of the existence of new resources or can request CIM to re-discover it.

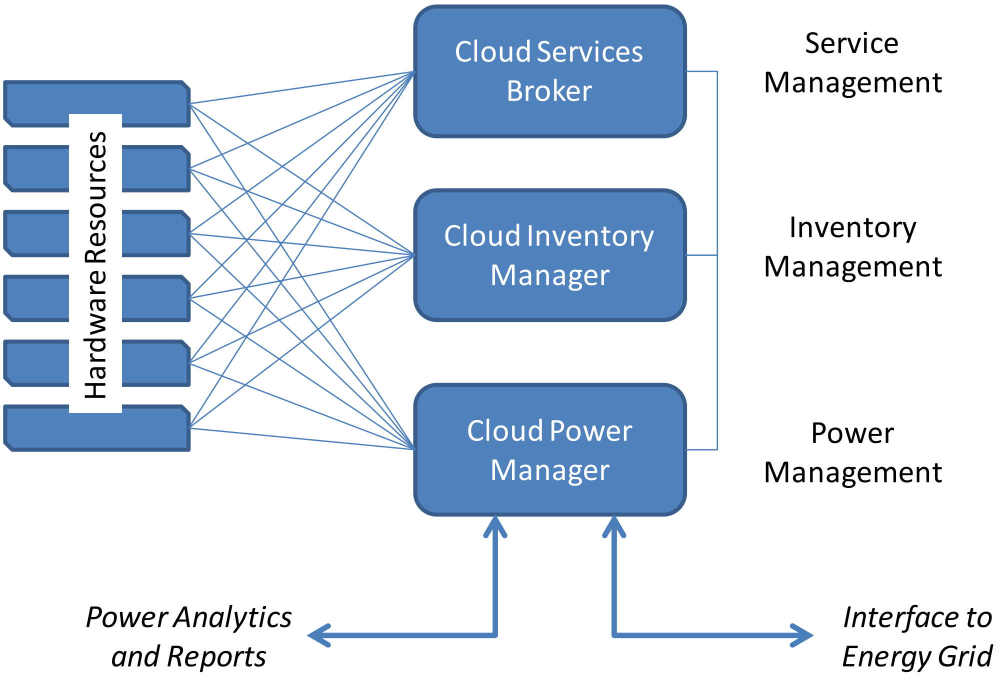


Figure III.1 – Service, inventory and power management

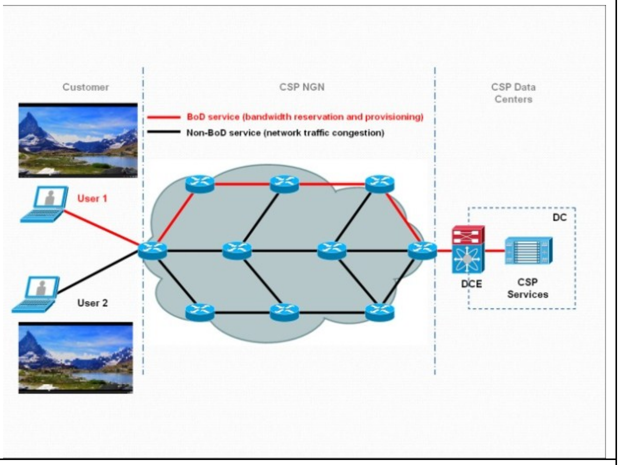
Appendix IV
Use case description related to the bandwidth on-demand services

In traditional cloud computing approach the applications or services hosted by CSP in data centre have no visibility on the end-to-end network connection. This could have a negative influence on the performance of these services/applications expressed by QoE end-users (stockholders). In consequence, the CSP offer may become less attractive and less competitive.

According to this, the concept of bandwidth on-demand services is targeted on the NaaS as a support of existing IT services offered in the cloud computing model. The BoD service is a complementary way for dynamic network resource allocation in order to leverage high performance of IT services located in data centres.

Table IV.1 – Cloud specific use-case description related to the bandwidth on-demand services

	Legend	**Use case**
Row 1	Use case title	Bandwidth on-demand service between an end user and CSP.
Row 2	Relevant business roles (Cloud service users, cloud service providers, cloud service partners, inter-cloud service broker – from clause 6.3)	Cloud service users, cloud service providers, cloud service partners, inter-cloud service broker
Row 3	Types of actors (Telecommunication service providers, Internet service providers, third-party provider, user, cloud broker – from Appendix IV)	Internet service providers, CSP, third-party provider, users
Row 4	Relevant cloud services categories (IaaS,PaaS,SaaS,NaaS,CaaS – from clause 7)	NaaS
Row 5	Relevant ecosystem deployment types (Private cloud, public cloud, community cloud, hybrid cloud, Internet or telecom cloud, hosted enterprise cloud, virtual network cloud – from clause 6.1)	Private cloud, public cloud, hybrid cloud, Internet or telecom cloud, hosted enterprise cloud, virtual network cloud
Row 6	Use case description *Note: this description should possibly use the above identified roles, service categories and deployment types (unless this is not necessary to describe clearly the use case)*	In this scenario we consider end-user access to cloud computing service (i.ex. VDI, video streaming) offered by CSP. The CSP serves the services on the basis of its own data centres and has no information on performance of particular connectivity between end users to the data centre where the service is hosted. According to this, from an end-user point of consideration, the QoE of the service is dependent on a combination of

		data centre and network performance. The CSP is able to guarantee a certain service quality limited to its own data centre. This quality could be easily downgraded by network performance on the way between the end user and the particular data centre of CSP. CSP acting alone is not able to impact network performance without interaction with the network operator. In order to prevent such a situation, the bandwidth reservation in the network between end-user and data centre is recommended. It allows guaranteeing certain network performance and allows the setup of a contract for end-to-end SLA for CC services between an end user and CSP. To fulfil these needs, CSP interacts with the network operator. The network operator can be any actor that has the ability to offer connectivity between the CSP data centre and the end user.
Row 7	Information flow (between the business roles) *Note: information flow should possibly use the above identified roles, service categories and deployment types (unless this is not necessary to describe clearly the use case)*	
Row 8	High level figure describing the use case	 Customer CSP NGN CSP Data Centers BoD service (bandwidth reservation and provisioning) Non-BoD service (network traffic congestion) User 1 User 2 DC DCE CSP Services
Row 9	Derived requirements (and required capabilities) for the cloud ecosystem *Note: requirements should be written as much as possible in a clear way such that we can then move (re-use) all of them to the requirements clause (clause 9) and possibly identifying common requirements derived for different use*	

	cases beyond those specific to each case.	
Row 10	Other information specific to the use case	

Bibliography

[1] IETF RFC5556: *Transparent Interconnection of Lots of Links (TRILL): Problem and Applicability Statement*

[2] Kandula, A., Sengupta, S., Patel, P. *The nature of data center traffic: Measurements and analysis*. In: ACM SIGCOMM IMC (November 2009)

[3] Farrington, N., Rubow, E., Vahdat, A.: *Data Center Switch Architecture in the Age of Merchant Silicon*. In: IEEE Hot Interconnects, New York (August 2009)

[4] Niranjan, N., Mysore, A., Pamboris, N., Farrington, N., Huang, P., Miri, S., Radhakrishnan, V.: Subramanya and Vahdat A Portland: *A scalable fault-tolerant layer 2 data center network fabric*. In: SIGCOMM 2009 (2009)

Printed in Switzerland
Geneva, 2012