CLOUD COMPUTING CONCEPTS

BY

Ravi Venkatasubbaiah

INTRODUCTION

This document will introduce various concepts within the Cloud Computing ecosystem. Cloud Computing is becoming ubiquitous, and soon most computing will become Cloud Computing and vice versa. There is much hype and jargon surrounding these topics; I have attempted to explain this in a simple manner. There are links to additional information throughout the document if you wish to learn more.

I wish to keep this document current, please send concepts to add & feedback to @ravi_venk

Contents

1. CONCEPTS AND TERMINOLOGY

Virtualization

Hypervisor

A Hypervisor is the component of the Virtualization software that runs directly on the hardware, and provides a platform to Virtual Machines to run. The Hypervisor functions like an Operating System in scheduling and managing the Virtual Machines, and connecting them to resources based on preset policies. Examples of Hypervisors are vSphere from VMware, Hyper-V from Microsoft, XenServer from Citrix, KVM, and others.

There are two types of Hypervisors:

Type 1 Hypervisors

Type 1 Hypervisors run directly on the hardware, without an Operating System. In fact, they function as the Operating System, and are also known as bare metal or native mode Hypervisors. All Virtual Machines run in the context of the Hypervisor. Type 1 Hypervisors offer better performance and better security. Examples of Type 1 Hypervisors are vSphere from VMware and Hyper-V from Microsoft. However, there are distinctions within this category as well. Type 1 Hypervisors can be either Monolithic or Microkernelized, each of which is quite distinct.

Microkernelized

A microkernelized Hypervisor is one that uses a full-fledged Operating System in place of a Hypervisor. This leads to larger memory utilization by the Hypervisor. A distributed driver model is used where each Virtual Machine has its own set of drivers, thus isolating Virtual Machines from each other. In addition third party code cannot execute within Virtual Machines, making them more secure. Microsoft Hyper-V is an example of a microkernelized Hypervisor.

Monolithic

A monolithic Hypervisor uses a proprietary driver model, and combines the virtualization stack into it as a single image. This helps in quicker boot times, and lower usage of memory by the Hypervisor. However, the drivers are shared between all Virtual Machines, and represents interaction between these Virtual Machines. VMware vSphere is an example of a monolithic Hypervisor.

This whitepaper details the differences between the architectures of these two Type 1 Hypervisors.

Type 2 Hypervisors

Type 2 Hypervisors run within the context of a regular Operating System, and is known as a hosted Hypervisor. Here, the Hypervisor itself functions similar to any other application running within the Operating System. Hosted Hypervisors were very popular when virtualization was still taking hold. The biggest advantage of a hosted Hypervisor is the ease of installation and use. However, Virtual Machines in this environment are limited in performance and functionality due to the additional layer between them and the hardware.

Additional information on these two types of Hypervisors is available here

Virtual Machine

A Virtual Machine is a set of computing resources that function together in the same manner as a physical computer, but exists only in the context of a Hypervisor, such as vSphere or Hyper-V. It is much easier to establish and manage a Virtual Machine than it is to manage the corresponding physical system. For example, we can quickly provide additional resources to a Virtual Machine compared with the time needed to upgrade the physical system. Virtual Machines use resources from their hosts, which are also shared with other Virtual Machines. This leads to much improved resource utilization. Each Virtual Machine runs an instance of an Operating System, much like its physical equivalent. Virtual machines can be easily moved to another host to facilitate hardware maintenance, or in response to a hardware failure.

Appliance/Virtual Appliance

An appliance or Virtual Appliance refers to a Virtual Machine that exclusively runs software provided by a particular vendor, rather than user software. This software customizes the Virtual Machine to provide a specific function. Some examples are network or security tools such as firewalls or virus scanners, monitoring tools, optimization tools, and more. The appliance runs within the context of the Hypervisor and can be viewed and managed like any other Virtual Machine. An appliance using commodity hardware thus replaces a custom designed hardware-based system. This lowers the cost and increases the speed of developing a new appliance.

Cloud

Cloud

The Cloud is a collection of large computing resources that are interconnected with high-speed networks. Clouds provide shared services to users, and enables sharing of resources at massive scales. This reduces the cost of computing for all users, since resource utilization is maximized.

Private Cloud

This is a set of computing resources that are entirely within a company's internal network. Access to this cloud is limited to users within the company, or that use VPN software provided by the company. These resources are otherwise not visible or accessible to other outside users at all. Examples of Private Cloud software are vCloud software from VMware, Hyper-V from Microsoft, and others, such as Cloud Stack, OpenStack, and Apache.

Public Cloud

This is a set of public computing resources that are accessible to all users via the Internet. Users establish an account with the Public Cloud provider of their choice, and proceed to select the resources they need to build their computing systems. They can upload their applications and data to these systems and start using them. Users can pay upfront for resources and services they intend to consume, or be billed by the Cloud Service Provider in a pay-as-you-go model. Amazon AWS and the OpenStack Consortium are the leading Public Cloud providers, while Google, Microsoft, and many others also compete in this market.

Hybrid Cloud

A Hybrid Cloud combines elements of the Private Cloud and the Public Cloud – it uses Public Cloud resources to augment, supplement, or replace the resources within Private Cloud as needed. This can occur due to a planned ramp-up of resources in anticipation of increased utilization, or a partial or total failure of resources within the private cloud. The addition of public cloud resources is seamless to the applications and mostly invisible to the users. Hybrid Cloud providers include VMware, Microsoft, and others.

Systems

Hosts

Hosts are the hardware systems that run software that we need. They provide resources for the most demanding of workloads, and can be configured with various amounts of resources based upon the workload it needs to support. Software that runs on hosts typically includes Hypervisors, Operating Systems, Databases, and Applications.

Platforms

Platforms are a combination of a host (or hardware) and an Operating System that runs in the context of the host. The Operating System could be a standard Operating system, such as Windows or Linux, or could be a Hypervisor, such as Hyper-V or vSphere. Examples of platforms are a host running vSphere 5.1, a host running RedHat Linux 6.1, or Windows 2008R2. Platforms constitute infrastructure upon which applications can be deployed.

Application platforms

Application platforms take the above concept a step further and include additional software, such as a Web Server, Java, or other application development environments. These application platforms also constitute infrastructure for applications, upon which platform-specific applications could be developed and deployed.

Servers

This was the previous name for large hardware systems that populated the datacenters of large enterprises. From the application end, it denoted one component of the client-server model of computing. These days, it represents any system that offers a system or application service to either users or other applications.

Clusters

Clusters are servers or hosts that are linked using software or hardware to offer more reliability to applications. Failure of a cluster member permits the application to continue on other surviving members within the cluster. Clustering technology improves the reliability of commodity hardware, lowering the costs of establishing highly reliable systems. In order to fully

utilize the benefits of a cluster, the Hypervisor, Operating System, and applications need to be cluster-aware.

Desktops/Laptops

These are the traditional client nodes for end users. Until recently, this was the only method we had to access our applications. Today, we have many other options, including Virtual desktops, mobile platforms, and other devices. Many of the recent advances have further lowered the costs of the client computer, while increasing available features and functionality.

Virtual Desktops

The benefits of virtualizing servers have also transformed the desktop environment, and consolidated them into corporate datacenters. This helps organizations easily deploy, update and manage desktops with ease. It also helps them reduce costs associated with periodic desktop replacements.

Services

Infrastructure as a Service (IAAS)

Companies that do not want to deal with establishing and maintaining a datacenter, managing equipment onsite, and building high-availability features can get all of these from a service provider. All they need is to request Virtual Machines based on their needs. All resources, such as CPU, Memory, Disk space, network bandwidth, etc. can be individually configured. In addition, they will need to setup their Virtual Servers, install, update, and manage them. However, this provides them with complete control over their environment in terms of configuration and security. AWS from Amazon is an example of an IAAS provider.

Platform as a Service (PAAS)

Companies often need more than just a set of Virtual Machines; they will need a development environment and tools to develop their applications. This is available as well, with tools to develop applications in categories such as web, mobile, social, gaming, etc. Microsoft's Windows Azure is an example of a PAAS provider.

Software as a Service (SAAS)

SaaS is the model of providing a specific software application, such as email, Accounting or Employee Benefits over the Internet. Users only need a browser, and there is no software to download, install or update. No user hardware is necessary for this, apart from a desktop, laptop, or mobile device to run a browser. The provider maintains the hardware, provides the application, and charges for usage, based upon a subscription model.

In each of the above service categories, the resources acquired or the service can be scaled up and down as needed. The benefits to companies result from the fact that they consume resources that have been implemented at large scales, leading to lower costs of usage. In addition, they have eliminated large capital expenditures, and replaced them with operating expenses.

The above three categories of services have been expanded by the industry to include desktops, storage, networks, and even IT. The official NIST definitions for the above services can be viewed here.

Software

Operating Systems

An Operating System is the program that runs on a computer and provides the environment for user programs and applications. While it was previously common for Operating Systems to require proprietary hardware manufactured by the seller, current Operating Systems run on industry standard hardware. Operating Systems provide a set of common services to user programs and applications, enabling application developers to quickly develop, migrate, and deploy their applications.

Commonly used computer Operating Systems are various versions of Windows for the server and end user environments, UNIX, Linux, MacOS, and other proprietary systems. A list of various computer Operating Systems can be found here, and a feature comparison of some of these can be found here.

Middleware

Middleware is the term for software or an application that provides services to programs enabling them to communicate with other applications. These services are distinct from the services provided by Operating Systems. In essence, they function as a translator between application and a database, or between two applications. Middleware often takes the form of messaging services, leading to integration of applications across a large enterprise. IBM, Oracle, and Microsoft are the leading vendors of middleware products.

Databases

Databases are repositories for large amounts of information. They are stored in specific formats that enable easy retrieval of information in response to queries for information. A Database Management Systems (DBMS) is software that enables easy organization, administration, development, storage and retrieval of information from the database. Collectively, the term database is loosely used to refer to both the DBMS and to the information repository.

Information within the repository is storage within tables, and further organized into rows and columns. They are likely to be indexed, which enables rapid access to information when needed. This access can be in response to users running queries for specific data; however it is much more likely that these queries originate from within user programs. As a result, it is highly important that the response time be low, otherwise the performance of the database will not scale.

Common databases in use include MySQL, PostgreSQL, SQLite, Microsoft SQL Server, Microsoft Access, Oracle, SAP, and IBM DB2.

Applications

Applications are software programs that provide specific functions to users. An application needs to be installed within a supported Operating System or platform in order to run. There are many categories of applications, depending upon the type of functions they provide, the industry they are used within, the technologies they use and support, and the types of users they cater to. A detailed listing of categories can be viewed in *Appendix A*.

Software Licensing Models

Applications can also be categorized by how they are provided and licensed by the vendor to the user. There are many types of licensing in use, but they fall into the two broad categories listed below.

Free Software Licenses

Free software permits users to modify the software for their needs and distribute it as well. However, any modified software must also provide the source code and offer freedom for further modification. There are multiple standards for free licenses, such as Open Source and Apache GNU General Public License. The Free Software Foundation and the Open Source Initiative are organizations that monitor license compliance with standards.

Non-Free Software Licenses

Non-free or proprietary software licenses restrict users in terms of their ability to view the source code, modify it for their use, or distribute it. The terms of the license are dictated by the provider of the software, and are governed by the manufacturer's End User License Agreement (EULA). The EULA is a legal contract between the manufacturer and the end user, defining the terms of usage, including what is permitted and what is denied.

Database

Traditional databases use a Structured Query Language (SQL) to provide easy access to information contained within the repository. Each database vendor provides SQL access to enable programs and applications to rapidly access data based on specific queries and update them as needed. Current databases can be broadly classified into two types - SQL based databases and NoSQL based databases.

SQL based

SQL based databases have been in use for many years for various purposes. However, they have fixed data formats; any modification would need the entire repository to be reorganized offline. In addition, distributed systems are needed to provide quick access to large amounts of information.

Examples of SQL based databases are Oracle, Microsoft SQL Server, IBM DB2 and many more.

NoSQL based

In the 2000s, databases grew rapidly, which limited the scaling in performance; in addition the format of data stored also changed frequently. This functionality was not possible using traditional databases, and gave rise to NoSQL databases. NoSQL databases provide quick response to requests, do not require fixed formats for data organization, and provide scaling in a distributed environment. This permits data to be organized in a flexible manner, and provide performance improvements.

Examples of NoSQL databases are MongoDB, CouchDB, Apache Cassandra and HBase. Each of these databases is also an Open Source database.

Networking

Network

A network is a mechanism of communication between a set of computers and devices that enables them to send and receive information from each other. A network permits sharing of information or resources between computers. Networks need a communication medium, common network communication protocols, and devices that work across different types of media and protocols.

LAN

A Local Area Network (LAN) interconnects computers that are in close proximity, such as a house, a building, or a corporate campus. By design LANs are limited to smaller distances, and support high-speed links. Computers, printers, storage controllers, and other devices are connected to LAN switches and can communicate with each other. While the connections and communication required network cables to be wired to all devices, wireless communication has established itself for LANs.

WAN

Wide Area Networks (WAN) span larger regions than LANs; they can data across cities, regions, countries, or around the world. They can interconnect LANs, as in connecting all campuses of a global corporation. They are not limited by distance, but their speeds are lesser than LAN speeds. In addition, they require telecommunication providers to establish connectivity, making this expensive. While WAN technologies have mostly required physical wiring, they are several wireless technologies as well.

Ethernet

Ethernet is the commonly used network medium that enables communication within a LAN. Starting out at 10 Megabit (Mb) speeds, most end connections today are at 1 Gigabit (Gb). Ethernet cables are limited to a maximum length of 100 meters each, which in turn places limits on the size of a LAN. Network devices, such as switches and routers can be used to extend Ethernet based networks.

10Gb Ethernet

During the past six years, 10Gb Ethernet has steadily become more popular, with more vendors releasing 10Gb network products. The cost per port has dropped during this period as well. This technology, which started as a high-end WAN technology is being used s a LAN technology within large datacenters. AT present 40Gb and 100Gb products are shipping from network vendors.

WiFi

Wireless communication has taken off over the past 10 years, with homes, businesses, campuses, and other public venues being equipped with WiFi connections. Since physical cables are not needed, it is less expensive; in addition, it can be used where it is not possible to lay physical cables. Usage of portable devices that are always connected to the Internet, such as laptops, MP3 players, game players, smartphones, and tablets has become widespread. This increased adoption has spurred growing demand and usage of WiFi connectivity.

WiFi connections have been considered less secure than wired connections, since an intruder does not need physical access to monitor traffic. As a result, this technology was not initially popular with corporations. However, advancements in encryption and security technologies for WiFi have led to widespread deployments. Additional information of WiFi technologies can be found here.

TCP/IP

Network communication protocols have been defined based upon conceptual models, such as the Open Systems Interconnection (OSI) model, which defines a seven layer hierarchy. A simplified form of this model, the TCP/IP model is the foundation of the Internet. The TCP/IP model simplifies the OSI Model into four layers. Additional information for the OSI Model and the TCP/IP Model can be viewed here and here.

IPv4

IP addresses are 32 bit addresses that uniquely identify a computer or other network connected device. This address length supports over 4 billion devices that can be connected to a public network. Addresses are assigned by the Internet Assigned Numbers Authority through Regional Internet Registries. These addresses were used much faster than anticipated, and eventually led to the development of IPv6.

IPv6

IPv6 uses 128 bit addresses, providing an extremely large number of addresses (3.4×10^{38}). IPv6 has been operating since 2008, but currently only about 1% of Internet traffic uses IPv6. Devices with IPv4 and IPv6 addresses cannot directly communicate with each other, and will exist as different networks. All devices will ultimately use IPv6 addresses; however this transition is expected to take a long time.

Additional information on IPv4 and IPv6 can be found here.

Switching

Switching moves network packets of information between two devices, a source and a destination, which are both part of the same LAN as the switch. These devices could be computers, printers, storage controllers, and many other types of devices. Network switches have powerful switching engines that permit communication at maximum line speeds. These switches vary greatly in terms of their capabilities – the number of ports, speeds supported on each port, and the types of media. All the ports within a switched network belong to a single broadcast domain, i.e., broadcast traffic reaches all of the connected devices.

Network switches can be broadly classified into three categories – access switches, distribution switches, and core switches.

Access switches

Access switches are used to establish network connectivity at homes and small businesses. These have smaller numbers of ports, and support lower speeds. They do not minimal management features, and are often used to establish Internet connectivity.

Distribution switches

Distribution switches support higher speeds and larger numbers of ports. They also aggregate traffic from smaller switches. The switches are more powerful, and have some management and redundancy capabilities. They are used by larger businesses that have centralized systems.

Core switches

Core switches are the most powerful switches used by large organizations, as well as major service providers. They have many high speed ports and provide reliable network communication with low latencies. Powerful computers that generate high volumes of traffic and aggregated traffic from smaller switches are typically connected to core switches. They possess advanced management features and fault-tolerance that increase the availability of the network.

Routing

Routing is a method of moving network packets between devices that are not necessarily in the same network as the router. Routers operate at Layer 3 of the OSI model, and use sophisticated routing protocols to determine the best route in order to direct the packets. Routers use a number of parameters to calculate the best route. These are prefix length, metric, and administrative distance, which are defined here. Routers maintain routing tables, which are updated based upon state information from neighboring routers. Routing can be static or dynamic, and respond to failures within the network by rerouting packets through an alternate path.

The aspect of the router that is responsible for moving packets is known as the data plane, while the aspect that maintains the routing tables and adds intelligence in deciding how to route packets is known as the control plane.

An alternate definition of routing can be viewed here, additional information can be accessed here, and a practical configuration example can be viewed here.

Layer 3 switching

Traditional switches operate on a single LAN, within Layer 2 of the OSI Model. They inspect packets to determine the source and destination addresses, and forward them appropriately. Layer 3 switches operate on multiple LANs, and support multiple routing protocols, similar to routers. Since they combine both switching and routing functions, Layer 3 switches are cost effective compared with routers. They offload many of the traditional router functions, while not possessing the WAN ports and features of traditional routers. Layer 3 switching enables many advanced features, such as VLANs, to be part of the switch configuration. A simple definition of a Layer 3 switch can be found here.

VLAN

Virtual LANs or VLANs are a method of partitioning networks to create smaller broadcast domains. Traffic between hosts in a VLAN is limited to the VLAN, which also reduces the amount of broadcast traffic within each domain. However, it also prevents network traffic from moving between domains, unless they are routed by a Layer 3 switch or router. See here for more information about VLANs.

Network Virtualization

Network virtualization is the process of abstraction to separate the logical behavior of network components from the underlying physical network resources. It can be classified into external and internal virtualization depending upon the resources being virtualized.

External Network Virtualization

Here, the external components of the network are logically partitioned to increase efficiencies. VLANs are a good example where a large physical network is divided into multiple logical networks. The benefits arise from limiting broadcast traffic, lower costs from Layer 3 switching instead of routing, and improved security and manageability.

Internal Network Virtualization

Virtual servers and containers can be provided with virtual interfaces, switches, and routers. This is referred to as Internal network Virtualization. Here multiple virtual network interfaces and virtual switches share the same physical network interfaces. This provides a number of benefits – lesser physical switch ports needed, elimination of physical cabling and configuration for virtual interfaces, and isolation of virtual switch traffic from the physical network.

Additional information on network virtualization can be viewed here.

Software Defined Networking (SDN)

Traditional network equipment operates in an autonomous manner, with the vendor's proprietary hardware integrating both the control lane and the data plane. However, this is unable to accommodate the needs of a virtualized datacenter, where administrators can set and dynamically manage network traffic. SDN abstracts the control plane into the Hypervisor itself, using external networks only to move data packets. There are several benefits that result from using SDN.

Agility

Using the higher processing power that a Hypervisor can access makes the control plane much more responsive than network equipment with lower processing power. Control plane traffic can be shared across all Hypervisors within the datacenter, which provides a comprehensive view of network traffic at the source.

Flexibility

Traditional networks are rigid, based on proprietary methods and a very large number of protocols, making them difficult to work with and optimize. It is difficult to modify control plane logic, since it is integrated with the hardware. The SDN Control plane can be modified easily, since it resides on standard hardware. In addition, the control plane logic can be modified by defined policies to handle various patterns of traffic.

Higher Efficiency

The control plane can easily balance traffic across multiple physical interfaces, dynamically shifting traffic away from a busy interface. Network interface utilization can be raised to high levels in a similar manner to increased resource utilization on servers from virtualization.

Lower Costs

Organizations do not have to invest in equipment that are complex and have a huge numbers of protocols and features that are rarely used. In addition, complex equipment requires extensive testing before deployment. The rigid nature of these networks also forces organizations to overbuild, at additional cost. With SDN, organizations can build and pay for only the capacity needed.

Resilient Infrastructure

Failed interfaces can be isolated sooner, and quicker responses lead to more robust infrastructure. Reduction in the numbers of physical network ports and cables lead to reduced failure due to human and non-human failures.

SDN Protocols

SDN is an evolving field, and the vendors involved are developing products and technologies, staking out positions, and influencing standards. Many existing protocols have been adapted for use within SDN environments, while new protocols have been defined as well. A couple of the key protocols are listed below.

OpenFlow

OpenFlow is a protocol that enables external control of routing decisions on switches and routers. While the data plane and associated logic still resides on the device, the Control plane logic is driven from the outside. This permits a centralized control plane for increased agility and flexibility. Most major vendors support the OpenFlow standard. See here for a white paper that provides more information on the OpenFlow architecture and use cases.

NVGRE

Network Virtualization using Generic Routing Encapsulation (NVGRE) is a protocol that enables VLANs to extend across data centers, across network boundaries, and span both Layer 2 and Layer 3 of the OSI Model (packets can be switched or routed to another part of the same VLAN). It uses encapsulation and tunneling, and enables load-balanced and multi-tenant networks to extend from a corporate network to the public cloud.

Microsoft is the main promoter of this standard, with several networking vendors supporting it.

VxLAN

Virtual Extensible LAN (VXLAN) is another standard similar to NVGRE that builds a virtual network across existing switched and routed networks. It enables communication between Virtual machines within the same VLAN across networks and datacenters, and enables the migration of VLANs to remote datacenters as well. VxLAN also extends the number of VLANS from 4096 to over 16 million. VMware and Cisco are the main promoters of VxLAN.

A larger list of SDN protocols can be viewed here.

Mobile

Mobile technologies have evolved over the past few years, with technology moving from the first generation known as 1G to the fourth generation, or 4G today. Mobile phones are smaller, more reliable, and much more affordable. They also have scores of new features that have increased their utility to users which in turn greatly increased their adoption. During the past seven years, phones have powerful computers, akin to desktops and laptops. They are capable of connecting to the Internet providing high speed access with 4G or WiFi networks.

Smartphones

Mobile technology has led to the evolution of powerful mobile phones that can do much more than make and receive calls. They can function as cameras, record audio and video, support email, send and receive text messages, browse the web, access internet sites for data, etc. They can be used to play games, access social media, and listen to music.

They also support applications (also known as apps), that can provide functions similar to a desktop or laptop computer. Popular smartphones include the Apple iPhone, Samsung Galaxy, Blackberry, as well as offerings from Nokia, HTC, and others.

Mobile Operating Systems

A smartphone is a powerful computer in a small form factor, and runs an Operating System, much like a desktop or laptop. The two main mobile Operating Systems available today are Apple's iOS and Google's Android. Together, they accounted for 88% of all smartphones sold. The other Operating Systems include Windows Phone and Blackberry. Each Operating System has access to an app store that provides access to updates and upgrades.

App Store

App stores provide users with more than Operating System updates. They have a large number of downloadable apps for various functions of interest. While many of these are free, there are apps that are sold as well. Availability of such apps improves the values of the smartphone platform to the users. Manufacturers offer application developers facilities to develop apps for their app store, and in turn take a share of the proceeds from the sale.

The main app stores are owned by Apple and Google; each of them has over 1 million apps in their app stores, and have recorded over 50 billion downloads.

Mobile computing

Email, web browsing, and apps to various functions have led to smartphones being used as end clients for corporate applications. The computing power has steadily increased as well, providing the ability to run multiple applications at the same time. Widespread availability of desktop applications on mobile platforms has led to smartphones becoming viable substitutes for desktop or laptop computers.

Tablets

Tablets are a new category of devices that have larger screen sizes than smartphones and usually connect over WiFi networks, instead of 4G connections. They function like a desktop or laptop computer, with access to similar applications. They have become popular due to the portability and convenience they offer. Apple's iPad is the leading player in this category, followed by Samsung's Galaxy Tab. There are many other products in this segment ranging from Amazon's Kindle Fire to Microsoft's Surface Pro.

A new category, known as the Phablet has been defined in between phones and tablets by phones with larger screen sizes, such as Samsung's Galaxy Note. These are attractive to users in developing economies that seek to purchase one device that functions as both a phone and as a tablet.

Bring Your Own Device (BYOD)

Large corporations have typically provided corporate cell phones to their employees. These are typically a set of phones qualified by the company and sourced through the mobile phone provider of the company's choice. With the advent of smartphones, companies initially restricted choices available to employees; they wished to protect company information and assets, since these phones could connect to and access the corporate network.

However, employees at all levels within the organization have adopted smartphones as their personal phones, and do not wish to carry multiple devices. Hence, corporations have begun to permit access to corporate networks using employee owned devices. This initiative, known as BYOD, has increased demands upon corporate IT and security departments to provide access while safeguarding company information.

Storage

Storage refers to repositories for databases, programs, and other types of information. Storage can be of many types, and some of these are below.

Disk Storage

We are all familiar with disk storage, which uses spinning magnetic disks to store information. Disk storage permits reliable and quick access to data, but magnetic disks are prone to mechanical failure. There are several categories of disk storage for various applications. Each of these provides a different set of features, including performance, reliability, cost, and risk of data loss.

Just a Bunch of Disks (JBOD)

This is the most common use of disks in desktops and laptops, where one or more disks together constitute the storage capacity. Failure of any disk means loss of information stored on that disk. This is relatively inexpensive, but provides lower performance in addition to the risk of losing data. As a result, this is not used when data availability is critical and data loss cannot be tolerated.

Redundant Array of Independent Disks (RAID)

This is a method of storing data spread across multiple disk devices. The entire set of disks is visible as a single logical disk to the Operating System. A failure of a single disk would not cause any loss of data, since it could be reconstructed from the remaining disks. However, multiple failures could result in loss of data. The additional tasks of spreading data across the set of disks and reassembling them can be performed by dedicated RAID controllers, or by the Operating System itself.

There are several methods of organizing disks in a RAID configuration, each of which provides some benefits at a certain cost. This method provides some protection against data loss; however it requires extra disks for this purpose, increasing the cost. In addition the speed of reading and writing to RAID disks is slower due to additional tasks involved. Additional information on various RAID configurations and their benefits can be found here.

Enterprise Storage Arrays

RAID arrays provide some amount of protection against data loss; however, they provide relatively poor performance. Large databases, real-time transaction systems, and other enterprise-class applications require higher performance from disk storage. Dedicated storage systems, known as enterprise storage arrays, provide performance benefits to such demanding applications.

Storage arrays employ dedicated hardware and software to organize and manage disks in RAID configurations to provide protection against data loss. They support various protocols for data transfer and provide multiple high-speed links to the system that hosts the application, and use advanced caching techniques to boost performance. An additional benefit of storage arrays is that they can provide storage to multiple systems. These arrays are highly reliable, with the ability to manage a variety of failures, including multiple disk failures without suffering data loss.

However, this performance and set of features comes at a very high cost. As a result, this solution is viable only for a small subset of enterprise applications, where availability and performance are both highly important.

Disk Protocols

Disk devices have moved from proprietary interfaces to open and standards interfaces. Some of the common interfaces are the Small Computer System Interface (SCSI), Serial AT Attachment (SATA), and Serial Attached SCSI (SAS), which are described below.

SCSI

The SCSI interface is a simple bus-based interface that connects a computer with a variety of devices, such as disk, tape, CDROM, scanners, printers, etc. SCSI cables, which could be copper or fiber cables interconnect SCSI devices with SCSI interfaces on host computers. SCSI interfaces can be found on high end servers, disk drives, and within enterprise storage arrays. Detailed information on multiple SCSI standards can be found here.

SATA

SATA is the primary interface for disk devices within desktop and laptop computers, and is an updated version of the previous parallel interface. SATA is popular due to its backward compatibility with its older standards. The current speeds support high performance making it cost effective as archival storage. The SATA standard defines cabling, protocols and device specifications, and can be found here.

SAS

SAS is a high-speed version of the SCSI interface, and used by systems, disks, tape devices, and storage arrays. Due to its performance and lower price, it is replacing fiber channel as the interface of choice within enterprise storage arrays. Detailed information on the latest version of SAS standards can be found here.

Tape Storage

Tape storage uses spools of magnetic tape on which data is stored. Magnetic tapes are inexpensive compared with disks and reliable with a long shelf life. However, they are quite slow in performance, and can only be used to save backup and archive copies of information on a periodic basis. Tapes also provide high density of storage, making it easy to store a copy of corporate information offsite, and can be used if total data loss occurs due to the occurrence of a disaster.

Higher disk capacities are available today, driving storage costs down close to the level of tape storage. In addition, technologies such as data compression, deduplication, and snapshots have enabled inexpensive disk storage to be effectively used to save multiple copies of data backups. This has resulted in a steady reduction of tape storage being used as a viable alternative for corporate data backups.

Flash Memory

Flash memory is a type of storage that does not need power to retain data. This type of memory is used in computers to store information, to make backup copies, and to transfer data between computers. While flash memory density has been low, it has increased over the recent years. In addition, the price has been decreasing as well, making it a contender to replace disk storage. Flash memory based disks are more reliable, since they do not have any moving parts.

Solid State Drive (SSD)

SSDs are increasingly used in laptops for a number of benefits, such as lower noise, smaller size, and greater reliability. In addition, SSD storage is not prone to fragmentation unlike disk storage. A new category of disks is the Hybrid drive. This is a magnetic disk with a small amount of flash storage. Frequently accessed data automatically moves to the flash storage, providing faster access.

SSDs are also used in Enterprise Storage Arrays, providing faster performance to critical workloads. These function in a manner similar to disk based storage arrays. Hybrid storage arrays incorporate a mix of disk storage and flash memory (similar to a hybrid drive, at a much larger scale), providing faster performance for frequently accessed information.

Storage Tiering

Storage tiering is a strategy to create multiple tiers of storage, with multiple performance tiers. The most frequently accessed information resides in the highest performance tier, moderately active data in the next tier, and so on. It is important to note that higher performance storage is much more expensive. Allocation of information to various storage tiers is automatic based upon the frequency of access. In case the frequency of access changes, the information is migrated to the appropriate tier over a period of time. This can provide a performance boost to infrequently used workloads that are periodic or seasonal.

Storage Protocols

Storage protocols enable systems to access storage in a simple manner, providing reliable high-performance links as part of a storage network. All components needed for communication to be accomplished are specified, from the physical wiring to the interconnect mechanism, and the various types of devices that can connect to a storage network. Commonly used storage protocols are Storage Area Network (SAN), Network Attached Storage (NAS), and the Internet Small Computer System Interface (iSCSI), which are described below.

SAN

A SAN is a dedicated network for storage devices to connect with systems to provide access to storage. Storage devices include disk arrays and tape libraries. These storage devices provide block level storage that the systems view and access as local disks. The storage device provides only raw storage capacity; file systems are created and managed by system that access storage volumes. All storage read and write operations take place on the SAN; these operations are not visible over the Local Area Network (LAN). Details about the definition of a SAN can be found here.

NAS

NAS is a file based network storage that connected to a standard LAN and provides file level access to clients. The NAS storage device manages the file system, and provides access to multiple clients. It manages the file system, while ensuring reliability of the

stored data, since it uses RAID. A NAS storage device uses one or more protocols to serve data to clients. Common protocols supported by NAS devices are (Network File System (NFS) and Common Internet File System (CIFS)

NFS

NFS is a distributed filesystem protocol for sharing files over a network. A client computer can access and update files on a remote computer in the same manner as accessing a local disk. Computers must support TCP/IP in order to use the NFS protocol. NFS protocol supports sharing of files and concurrent access from multiple systems. The current version of NFS is v4.1, released in 2010.

CIFS

CIFS is an enhanced version of Microsoft's Server Message Block (SMB), and is a protocol to share files and share access to printers over a network. CIFS protocol also supports sharing of files and concurrent access from multiple systems. The current version of CIFS/SMB is v3.0, released in 2012, and adds support for virtualized environments.

iSCSI

iSCSI is a storage protocol that permits systems to send SCSI commands using the Internet Protocol (IP) over networks to access storage. iSCSI works over a Local Area Network (LAN) or a Wide Area Network (WAN). The functionality is similar to that of a SAN; however it uses standard network cabling and interfaces instead of fiber cabling and dedicated Host Bus Adapters (HBA).

SAN Components

SANs require a dedicated network, with various components to be setup and managed. This is expensive to setup and operate. In contrast, NAS and iSCSI do not require any additional hardware or components, making them less expensive. However, all storage traffic in a SAN uses the dedicated fiber cables, providing higher available bandwidth. In addition, it is a block level protocol, which is more efficient than network protocols. With NAS and iSCSI, the network interface is shared between storage traffic and network traffic.

The components of a SAN are listed below.

SAN Switch

A SAN switch is a storage switch, very similar to a network switch. Storage Controllers and HBAs connect to ports on the SAN switch using fiber channel cables to create a SAN fabric. SAN switches can be interconnected to expand the fabric. Multiple SAN switches can be managed as a single logical switch, and allow the creation of zones. Each zone permits specific host systems to access storage volumes on specified storage controllers. Failed ports are automatically isolated upon failure, and alternate paths utilized for communication.

Storage Controller

Storage controllers are the brains of Enterprise Storage Arrays, and manage the storage volumes for presentation to host computers. Storage controllers offer multiple high speed ports, redundant controllers, RAID support, intelligent caching techniques, and many more performance and management features. Storage controllers can support multiple protocols including NAS and iSCSI, as well as storage virtualization features such as replication, snapshots, deduplication.

SAN HBA

SAN HBAs are interfaces on hosts systems that connect the computer to a SAN fabric to access SAN storage. The HBA performs all tasks associated with data storage and retrieval, offloading it from the host computer. Multiple HBAs can be configured to increase the storage throughput, as well as to provide the ability to withstand failure of a HBA.

Converged Network Adapters (CAN)

SAN HBAs have usually been more expensive compared with Ethernet adapters, but have provided more throughput and features. In addition, Ethernet adapters did not have sufficient bandwidth to support both network and storage traffic. With 10Gb Ethernet adapters becoming widely available, there is sufficient bandwidth that is available at lower costs. These adapters support multiple storage protocols, and consolidate all output from the host into one interface.

Storage Virtualization

Large organizations have a number of storage arrays, which are used to provide storage to various systems, workloads, and databases. Storage volumes have been sized based upon growth considerations for the next few years. As a result, corporations tend to but excess capacity for which they pay today, but do not use. In addition, they have a need for capacity on one array, while other arrays have excess capacity; in such cases, the "stranded" capacity may be unusable. A number of techniques have been developed to increase storage utilization. Some of these are listed below.

Consolidation

Consolidation of multiple storage arrays reduces the number of pools of "stranded" capacity. This has been helped by newer arrays that offer higher capacity and higher performance. In addition, virtualization software is available that provides a single logical view to multiple storage arrays. As a result, the entire storage is treated as a single pool, making utilization more effective.

Thin provisioning

The traditional concept is to allocate the storage required for a particular volume, regardless of whether it is immediately needed. In addition, allowances for expected growth and safety margins usually mean that storage allocated is many times the immediate need. Thin provisioning allocates the entire volume, but immediately provides only a small fraction as usable capacity. As this capacity is used and starts to fill up due to addition of data, the size is transparently extended. This process continues until the volume grows to the allocated size.

The main benefit of thin provisioning is that it reduces "stranded" storage that has been allocated to volumes, but is not in use. While it increases the utilization of the storage array, there is a small risk from oversubscription. In case multiple volumes quickly grow to their maximum allocated capacity, the storage array could run out of space to provide. This risk could be mitigated by monitoring the available space on the array. More information on thin provisioning can be viewed here.

Deduplication

Deduplication is a technique that eliminates duplicate copies of data. The intent is to eliminate storing multiple copies of the same block of data, and store it only once instead. This can be performed with primary disk storage, with backup to tape storage, and while transmitting information over networks. Deduplication thus increases

effectiveness of storage. There are different methods of deduplication, as listed below. Some interesting facts regarding data deduplication can be found here.

Post-process

The data is written to the disk immediately in the post-process method, without the delay involved with creating and checking deduplication checksums. Later, an automated process scans the disk to remove duplicate information. While post-process does not incur any delays during initial write, duplicate information is written to the disk and is not removed until the next scan.

In Line

With in line deduplication, checksums are calculated and checked for duplication in real time during the initial writing process to disk. In case the data exists, only a pointer to the existing data is written. This always uses the minimum amount of storage since no duplicates are ever present on the disk. However, the process of writing data to disk is slowed down due to calculation of checksums and checking for duplication.

Source based and target based Deduplication

Source deduplication performs the deduplication at the source, or at the primary storage. Periodic scanning locates and removes duplicate blocks of data, and references are placed pointing to the existing master block of data. In case one of a set of duplicate blocks is modified, the entire modified block is written to disk as a new unique block.

Target based deduplication locates duplicates and removes them from secondary storage, such as tape storage, or a virtual tape drive.

Replication

Storage replication is a method to copy all changes to storage volumes on a local storage array to a remote storage array, in order to maintain an exact replica of the local array. In case the local array fails, the remote array contains practically the same information. This replication technology can be enabled at the array level or at the individual volume level, and needs a dedicated link for replication to take place.

Snapshots

Snapshots are a method of establishing the state of a storage volume at a particular point in time. Snapshots can be created at the file, database, storage volume, or the array level. A set of pointers is recorded during the creation of the snapshot, and all subsequent changes are tracked separately. This technology has multiple uses – to perform a complete and consistent backup without pausing activity on the volume or to make a mirror copy of the entire volume. Snapshots can also track changes since the previous snapshot, enabling incremental backups, rollback, and roll forward operations.

Business Continuity

Business Continuity refers to a set of activities that are performed on a daily basis to ensure that the organization, its business operations, and its information are safe, consistent, and recoverable in the event of an accident, disaster, or emergency. Planning for these contingencies is known as Business Continuity Planning, and involves generating procedures, instructions, and other documents to address possible disasters and emergencies with a operational recovery plan. All designs, resources, and plans must be specified during this process to enable successful business recovery. Other aspects of Business Continuity Planning include Disaster Recovery Planning, and Risk Mitigation Strategies.

Disaster Recovery

Disaster Recovery is the process of recovering IT systems and information necessary for the business to operate in case of a major accident or disaster that causes a large-scale outage. Planning for Disaster recovery involves defining the policies, processes, and procedures involved in identifying the scope of the disaster, system and data recovery, and validation of business applications. The types of natural and manmade disasters should be specified, and the resources needed, including IT and business personnel, IT infrastructure components such as networks, systems, storage, and data backups should be delineated. In addition, two critical parameters need to be part of the planning process, with organizational buy-in. These are the Recovery Point Objective (RPO) and the Recovery Time Objective (RTO).

RPO

RPO represents a measure of how close a recovery to the time of the disaster is possible. It is also the maximum amount of time for which data loss could be tolerated. This parameter depends on the nature of the business, and the type of service or application, which need to be considered during the design of the

Disaster recovery Plan. For example, the corporate website is relatively static for many days, whereas the sales database records 20,000 transactions each hour. It clearly does not make sense to have a single value, say 2 hours for both of these services.

More information on RPO can be viewed here.

RTO

RTO represents the amount of time that the application or service could remain unavailable in the event of a disaster. This is the time during which the scope of the disaster is identified, and the Disaster Recovery Plan is initiated, until the application or service has been tested for use by the business users. Again, this needs to be individually specified for each application or service. Corporate email and the transaction database need to be brought back into service very quickly, while the corporate intranet site could wait for a day or two.

More information on RTO can be viewed here.

Security

Antivirus Software

Antivirus software is software that can prevent, detect, and remove various types of malware that can infect computers. Identification of viruses typically uses one of these methods – signature based detection, heuristic techniques to identify behaviors, or detecting if administrator privileges are being requested. Another method is to isolate the suspected program in a sandbox and examine its behavior.

While antivirus software protects computers, there are costs as well. Apart from financial costs, these programs use resources while scanning, and can render computers less usable during this period. Antivirus programs can malfunction, generate false positives, and interfere with the stability of the Operating System.

Firewall

A firewall is a security system that examines network packets that enter of leave a network and decides if they are permitted based upon a set of pre-defined rules. Firewalls may be hardware devices or implemented as software. They reside at the boundary between two networks of

different security or trust levels, i.e. between a secure network and an untrusted network. Practically, they are found between and organization's network and the Internet.

Firewalls also perform basic routing functions, and provide services such as Network Address Translation (NAT). Common firewall types are listed below.

Packet filters

Packet filters are firewalls that operate at lower levels of the OSI model, and perform filtering based upon a set of predefined rules. Stateful firewalls keep track of existing sessions while stateless firewalls do not need to know about sessions. However, stateful firewalls are slower and need more resources to complete it. Filtering can occur based on source IP address, destination IP address, source IP port, destination IP port, destination service, protocol, and many more parameters.

Application Firewalls

These firewalls operate at the application layer, and inspect packets directed to specific applications. In case packets are not correctly formatted for use by specific applications, or the packet contains the signature of a known virus, such packets are dropped without acknowledgment to the sender. Application firewalls understand the concept of sessions, i.e., they are able to understand the conversations that establish a session, and view packets sent to applications and filter based on malware signatures. Typically both types of firewalls are used in combination.

More information on firewall functionality and operation can be viewed here.

Virtual Private Network (VPN)

A Virtual Private Network (VPN) is the extension of a private network across a public network. A secure point-to-point connection is established between the private network and a client computer (VPN client) residing within a public network. However, the client receives access to network resources and is part of the security environment of the private network. VPNs are simple and easy to use, cost effective compared with dedicated lines, and improve productivity. More information on VPN technologies can be found here.

2. PRODUCTS

Virtualization

Hypervisors

vSphere

vSphere is VMware's virtualization platform, and the current release is v5.1 (August 2012). Supported Operating Systems on this platform include various versions of Microsoft Windows, and Microsoft Windows Server (32 bit and 64 bit versions), Red Hat Linux, SuSe Linux, Ubuntu, and Solaris.

vSphere is available in various licensing tiers, Essentials, Standard, Enterprise, and Enterprise Plus, with more advanced features limited to the higher tiers. Other products, such as VMware vSphere Storage Appliance, vSphere Data Protection, vSphere Replication and vShield Endpoint complement the vSphere platform.

Additional information on vSphere is available here.

Hyper-V

Hyper-V is the virtualization product of Microsoft, and the current release is v3.0 (June 2013). Supported Operating Systems are various versions of Microsoft Windows, and Microsoft Windows Server (32 bit and 64 bit versions), Red Hat Linux, and SuSe Linux.

Hyper-V is available in various licensing tiers, Foundation, Essentials, Standard, and Datacenter, with more advanced features limited to the higher tiers. Other products include Windows 2012 R2, Active Directory, and Internet Information Services 8.0

Additional information on Hyper-V is available here.

XenServer

XenServer is a Hypervisor that began as a research project at the University of Cambridge, and supported by XenSource until Citrix acquired it in 2007. The project has since moved to www.xen.org, which has an advisory board of industry majors.

The current version of XenServer is v4.3 (July 2013). Supported guests are various versions of Microsoft Windows, and Microsoft Windows Server (32 bit and 64 bit versions), a number of Linux distributions, and a number of UNIX distributions.

XenServer is available in Free, Advanced, Enterprise and Platinum Editions. A number of industry majors including Amazon Web Services (for EC2), Rackspace Cloud, Fujitsu Global Cloud. Many Internet hosting service companies also use XenServer to provide virtual private servers.

Additional information on XenServer is available here and here.

KVM

Kernel-based Virtual Machine (KVM) is a virtualization solution for Linux. While it began with support on Linux for x86 processors, it supports a number of hardware platforms. KVM is promoted by RedHat and is licensed under the GNU GPL license. The current release of KVM is 1.2.0 (September 2012)

See here for more information on KVM.

Cloud

Public Cloud

Amazon Web Services (AWS)

Since 2006, AWS has been providing a complete suite of Cloud Services accessible over the Internet. Popular for Elastic Cloud Compute (EC2), and Simple Storage Service (S3), AWS provides services for Compute, Networking, Content Delivery, Database, Storage, Application Services, and more.

AWS has a global presence in 8 regions, with multiple availability zones in each region. Customers can configure their virtual machines across regions and availability zones to assure availability and performance. Customers are billed for AWS resource utilization based upon actual usage, similar to a utility.

AWS offers IAAS on demand, enabling resources to rapidly scale up or down based on demand. In addition, AWS has been periodically reducing prices for its products; the most recent price reduction was in July 2013. The convenience at a low cost has made it attractive for companies to deploy applications using Cloud-based infrastructure – see here for a recent report about the growth of websites using the AWS platform.

More information on AWS services can be found here.

Rackspace Cloud

Rackspace offers Public Cloud services to users, known as RackSpace Cloud, and is an IAAS provider, similar to AWS. Billing for Rackspace users is based on actual usage, similar to a utility. Rackspace has three datacenters in the US and UK to provide Compute, Storage, and other services to users.

Rackspace also offers Cloud Sites, a lower functionality option at a lower cost. In 2010, Rackspace contributed source code from its "Cloud Files" product to the OpenStack project, for use as the OpenStack Object Storage Component. In 2012, Rackspace decided to use the OpenStack Compute as the technology for its Cloud Services.

More information on the Rackspace Cloud can be found here.

Windows Azure

Windows Azure the Cloud platform offering from Microsoft, that offers both IAAS and PAAS services. Many programming languages, tools and frameworks from Microsoft and third parties are part of the PAAS suite. Azure is also a Cloud application development platform; it can be used to develop web applications using available tools and deploy them within Microsoft's datacenter. These applications can communicate with a user's other applications within the Azure cloud.

In addition, Azure also has Compute services, Database, Business Analytics, Messaging and more. Windows Azure is hosted from a number of Microsoft's global datacenters.

Additional information on Windows Azure is available here and here.

Google App Engine

Google App Engine is a PAAS platform for developing and hosting web applications within Google's datacenters. Automatic scaling for applications is offered by Google App Engine; resources allocated to applications are automatically scaled up and down based upon the number of web requests.

Google App Engine is free for users until specified limits on resource usage are reached, and charges apply once these limits are exceeded.

See here for more information on Google App Engine.

Private Cloud

vCloud

VMware's vCloud is a suite of Cloud Service Components. Its vSphere virtualization platform is supplemented with a vCloud Director, vCloud Connector, Cloud service provisioning tools, Operations Management, and Business Management tools. It is available in Standard, Advanced, and Enterprise Editions.

The vCloud suite enables VMware's vSphere customers to move their virtualization platforms into deploying Private Cloud services. It enables them to implement a Software Defined Datacenter that enables a self-service model.

More information on the vCloud suite can be found here.

OpenStack

OpenStack is a free, open source, Cloud Computing project to provide IAAS software for use in Public and Private Clouds. It was formed in 2010 as an open source initiative between Rackspace and NASA in order to provide Cloud Computing services upon standard hardware. It is now managed by the OpenStack Foundation, and supported by over 150 companies, including all the Cloud Computing leaders.

The project is based upon collaboration within the OpenStack community, with periodic releases. OpenStack is available with various Linux distributions, and the current release is 2013.1.2 (June 2013). OpenStack is also compatible with Amazon AWS, and application written for AWS can be easily modified to work with OpenStack.

Additional information on OpenStack is available here.

CloudStack

CloudStack is open source software that helps to create and manage large numbers of virtual machines and services. It uses existing Hypervisors, such as vSphere, KVM, and XenServer. It is also compatible with AWS for Public Clouds, enabling customers to create Hybrid Clouds.

CloudStack was initially developed by cloud.com, which was acquired by Citrix in 2011. Citrix released the software to the Apache Foundation, and exited the OpenStack alliance. The current release of CloudStack is 4.1.0 (June 2013).

See here for more information on CloudStack.

Hybrid Cloud

vCloud

In May, 2013 VMware announced a vCloud Hybrid service that provides seamless integration between a customer's Private Cloud (based on vCloud) and the public cloud, offered by one of VMware's partners. Virtual Machines can be moved to the Public cloud based on increased usage, seasonal demand, planned maintenance, or unplanned outages.

More information on the vCloud Hybrid Service can be viewed here.

Hosts

HP Proliant Servers

HP is one of the leaders in providing Enterprise servers that are reliable and scalable. The familiar Proliant servers have been supplemented by the HP Bladesystem servers. In addition, HP recently released Moonshot servers based on Intel Atom processors to deliver high-density computing while reducing energy consumption. Supported Operating Systems include Windows, Linux, HP-UX, their proprietary OpenVMS and NonStop Operating Systems, in addition to most Hypervisors.

Additional information on HP servers can be viewed here.

IBM System x and System z

IBM is a leader in providing Enterprise computing solutions, with its Power Series, System x, System z mainframes, and the BladeCenter. Supported Operating Systems include Windows, Linux, and proprietary Operating Systems such as z/OS, AIX, and OS/400. Most Hypervisors are supported on open hardware.

See here for more information on IBM servers.

Dell PowerEdge

Dell provides PowerEdge servers for the Enterprise, with their R series servers and Blade Servers. Supported operating Systems on Dell Servers include Windows and Linux, in addition to most Hypervisors.

More information on Dell servers can be viewed here.

Oracle Sun Servers

Oracle acquired Sun Microsystems in January 2010, and has continued to develop the Sun server line. The current lineup includes SUN SPARC servers, Sun x86 servers, and Sun Blade Servers. The SPARC servers and the Blade servers are based on SUN's proprietary SPARC processors, and support the Solaris Operating System. The x86 servers contain Intel or AMD processors, and support Windows, Linux and most Hypervisors.

Detailed information on Oracle Sun Servers can be viewed here.

Cisco UCS Servers

Cisco, well known for its network equipment, entered the server hardware segment with its Unified Computing System (UCS) in 2009. UCS is a high-density server system that supports most Hypervisors. It offers redundancy and management features, in addition to large amounts of memory and Converged Network Adapters.

See here for more information on Cisco UCS servers.

Operating Systems

Windows

Microsoft Windows is the leading platform for desktop and server based computing.

Windows Server is qualified to run with x86 and IA64 based processors. The current version of Windows Server in use is 2008 R2 (July 2009) and Windows Server 2012 R2 was released in June 2013. Windows Server 2008R2 and Windows Server 2012 R2 also include Hyper-V, Microsoft's Hypervisor. There are four editions of Windows Server - Foundation, Essentials, Standard and Datacenter.

For the desktop, Windows 7 is the most popular release, with Windows 8 being released in May 2013. Windows 7 (released July 2009) runs on x86 hardware, whereas Windows 8 (released October 2012) supports x86 and ARM architectures. Windows 8 adds support for touchscreen devices, such as smartphones and tablets. A new feature of Windows 8 is the Metro style interface, which is a new user interface that has active tiles, and increased integration with online services offered by Microsoft.

See here and here for more information on Microsoft Windows.

Linux

Linux is a free open source Operating System first released in October 1991 by Linus Torvalds. The software may be modified and distributed under the GNU GPL license. There are many variants of Linux, with RedHat Enterprise Linux and SUSE Linux Enterprise Server being the chief Enterprise variants, and CentOS and Ubuntu the popular desktop variants.

Linux has also been modified to work on network devices, game consoles, and mobile devices (including the popular Android Operating System). Linux has been adapted to run on mainframes and proprietary architectures; over 90% of the Top 500 supercomputers run a variant of Linux.

More information on Linux can be viewed here.

End User Devices

HP

HP has a wide range of end user computing devices, for consumers as well as for business users. For consumers, HP has the Pavilion and Envy series of desktops and laptops, whereas the Pro and Elite series are targeted to Business users. In addition, HP has a wide range of workstations, Thin and Zero clients.

See here for more information on HP end user devices.

Lenovo

Lenovo is the second-ranked PC vendor by unit sales in 2012. It offers a full range of end user devices including desktops, laptops, workstations, and tablets. Lenovo acquired IBM's PC business in 2005, and ThinkPad laptops are one of its flagship products. In addition, it offers IdeaPad laptops, ThinkPad and IdeaTab tablets, ThinkCentre and IdeaCentre desktops, and ThinkStation workstations.

More information on Lenovo's end user products can be viewed here.

Acer

Acer has a wide range of end user computing devices, including laptops, desktops, tablets, and netbooks. Its consumer product offering include Aspire notebooks, Iconia

tablets, as well as Aspire and Predator desktops. In addition, Acer offers the TravelMate and Extensa laptops and the Veriton series of desktops targeted towards business users.

Additional information on Acer's end user products can be viewed here.

Dell

Dell offers a number of end user computing devices, for consumers as well as for business users. For consumers, Dell has the Inspiron and XPS series of desktops and laptops. The Latitude laptops and Optiplex desktops are targeted to Business users. In addition, Dell offers the Precision series, tablets, and Thin clients.

See here for more information on Dell's end user devices.

Database

Oracle

Oracle is an SQL based object-relational database management system. The current versions are Oracle 11g (September 2008) and Oracle 12c (June 2013). Oracle databases are available for various Operating Systems, including Microsoft Windows, Linux, Solaris, HP-UX, AIX, and HP OpenVMS.

Oracle databases come in three editions, Oracle Database Enterprise Edition, Oracle Database Standard Edition, and Oracle Database Standard Edition One. In addition, Oracle also offers several complementary products for Data Warehousing, Database Security, High Availability, Big Data, and Cloud-based databases.

More information on Oracle's database products can be viewed here

DB2

IBM DB2 is a relational model database server system. The current version of DB2 is v10.1 (April 2012). DB2 databases are available for various Operating Systems, including Microsoft Windows, Linux, Solaris, HP-UX, and AIX.

IBM DB2 is available in three editions, IBM DB2 Express-C, Workgroup Server Edition, and Enterprise Server Edition. IBM also offers the Informix database and the InfoSphere Data Warehouse.

Additional information on IBM's database products can be viewed here

SQL Server

Microsoft SQL Server is a relational database management system available on a number of Windows desktop and server Operating Systems. The current version is SQL Server 2012 (April 2012). There are several editions of SQL Server; Express, Workgroup, Business Intelligence, Web, Standard, and Enterprise. In addition specialized editions are available depending upon the type of environment.

See here for more information on Microsoft SQL Server.

MySQL

MySQL is an open source relational database management system, licensed under the GNU GPL license. The current release is v5.6.2 (June 2013). MySQL is available for various versions of Windows and Linux, as well as HP-UX, AIX, and Solaris.

MySQL is now owned by Oracle Corporationafter its acquisition of Sun Microsystems.

Additional information on MySQL products can be viewed here.

Enterprise Storage

Hitachi Unified Storage

Hitachi's Unified Storage platform is an offering to Enterprise customers that provides a high performance virtualized storage environment. It offers both SAN and NAS support, and provides many storage virtualization features, such as thin provisioning, Snapshots, Replication, Deduplication, and high levels of performance, availability, and data protection.

More information on Hitachi's storage products can be viewed here

IBM Storage Products

IBM has several families of storage products that are targeted to various market segments. The DS8000 series of storage is focused to Enterprise customers, offering features such as continuous access, redundancy, high scalability, and management features. The Storwize family is geared towards midrange customers and provides the ability to consolidate their storage into a virtualized platform. IBM also offers Flash based storage for demanding workloads.

Additional information on IBM's storage products can be viewed here

HP Storage

HP has storage solutions that range from entry level to midrange and the Enterprise. The 3PAR storage platform offers all flash, mixed, or disk based storage with large scale and high performance for demanding workloads. The StoreVirtual platform permits scale-out storage, increasing capacity and performance at the same time. The MSA family offers a robust entry level platform. In addition, HP has its traditional families of XP and EVA storage.

See here for more information on HP storage products.

NetApp FAS

NetApp offers the FAS family of storage to offer solutions to all categories of customers. The FAS6200 is targeted to Enterprise customers, and offers Flash, Disk, or mixed storage. The FAS3200 series is targeted to midrange customers, and FAS2200 is an entry level platform. NetApp also has the V series virtual storage, which manages third party storage. A wide variety of storage and management software is available from NetApp as part of the Data ONTAP system.

More information on NetApp's storage products can be viewed here

EMC Storage

EMC offers a number of storage products to customers familiar with its Symmetrix product line. In addition to the Symmetrix VMAX and the Symmetrix VMAXe, it also offers the ISILON line of scale-out NAS products, and the all Flash XtremIO product line.

EMC also offers a number of storage software and management solutions. Additional information on EMC's storage products can be viewed here.

Networking

Cisco Systems

Cisco is the largest networking company, popular for its routers and Catalyst switches. Current network product offerings include the Nexus series of switches, routers, Catalyst switches, wireless LAN products, carrier grade products, and remote connectivity solutions. Cisco also offers voice products, including IP phones, and Telepresence solutions for communications.

Additional information on Cisco's networking products can be viewed here.

Juniper Networks

Juniper Networks is best known for its high-end, carrier grade routers. It has several products including the T series core routers, the M and MX series of multiservice routers, and the E series edge routers. Juniper also has carrier grade MPLS switches and EX series of Enterprise switches.

See here for more information on Juniper networking products.

APPENDIX A

CATEGORIES OF APPLICATION SOFTWARE

A brief list of various categories of application software is below.

- Business function specific software
- Finance accounting and enterprise resource planning ERP software
- Computer game or entertainment software
- Content authoring and editing software
- Content management software
- Data management and query software
- Development software
- Educational or reference software
- Industry specific software
- Network applications software
- Network management software
- Networking software
- Operating environment software
- Security and protection software
- Utility and device driver software
- Information exchange software
- Electrical Equipment software
- System management software

More information on these categories, which is based upon the United Nations Standard Products and Services Code, can be viewed here.