Dr. D. Y. Patil Pratishthan's Institute for Advanced Computing &Software Development IACSD

Data Center Management

INDEX

What is a Data Center?

A data center is a building that contains a large amount of computer hardware. This hardware consists of the following:

- Central Processing Units (CPU's) - this is the brains of the data center, and is made up of literally thousands of processors that perform the work needed by the business.
- Storage - this is static storage (storage that maintains its contents even after the power is removed) the system has at its disposal. This is typically a combination of hard disk drives (regular storage), solid state drives (high-speed storage), and tape drives (backup).
- Communications - Depending on the age of the data center, this can consist of modems (telephone line communications), datasets (dedicated telephone line communications), traditional networking (what you're likely familiar with), and high-speed networking (fiber optics or similar).
- Software - this is the programming that the business needs to operate. Generally, this falls into two categories. First, there is infrastructure software. Things like the tools used to manage the hardware, database management systems, email systems, and such. Second is application software. These are the programs that employees typically use on a day-to-day basis. Microsoft Office immediately comes to mind.

Why are data centers important to business?

In the world of enterprise IT, data centers are designed to support business applications and activities that include:

- Email and file sharing
- Productivity applications
- Customer relationship management (CRM)
- Enterprise resource planning (ERP) and databases
- Big data, artificial intelligence, and machine learning
- Virtual desktops, communications and collaboration services

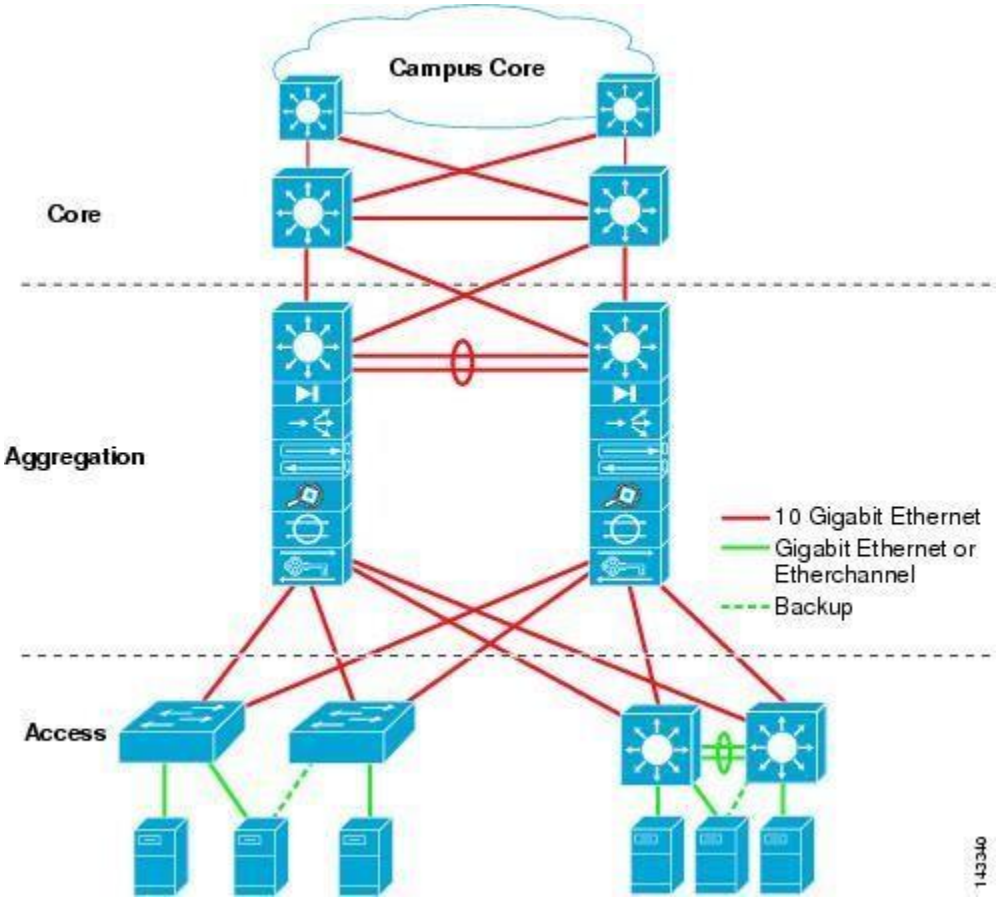
Data Center Architecture Overview

The data center is home to the computational power, storage, and applications necessary to support an enterprise business. The data center infrastructure is central to the IT architecture, from which all content is sourced or passes through. Proper planning of

the data center infrastructure design is critical, and performance, resiliency, and scalability need to be carefully considered.

Another important aspect of the data center design is flexibility in quickly deploying and supporting new services. Designing a flexible architecture that has the ability to support new applications in a short time frame can result in a significant competitive advantage. Such a design requires solid initial planning and thoughtful consideration in the areas of port density, access layer uplink bandwidth, true server capacity, and oversubscription, to name just a few.

The data center network design is based on a proven *layered* approach, which has been tested and improved over the past several years in some of the largest data center implementations in the world. The layered approach is the basic foundation of the data center design that seeks to improve scalability, performance, flexibility, resiliency, and maintenance. Figure 1-1 shows the basic layered design.



described as follows:

- Core layer—Provides the high-speed packet switching backplane for all flows going in and out of the data center. The core layer provides connectivity to multiple aggregation modules and provides a resilient Layer 3 routed fabric with no single point of failure. The core layer runs an interior routing protocol, such as OSPF or EIGRP, and load balances traffic between the campus core and aggregation layers using Cisco Express Forwarding-based hashing algorithms.

- Aggregation layer modules—Provide important functions, such as service module integration, Layer 2 domain definitions, spanning tree processing, and default gateway redundancy. Server-to-server multi-tier traffic flows through the aggregation layer and can use services, such as firewall and server load balancing, to optimize and secure applications. The smaller icons within the aggregation layer switch in Figure 1-1 represent the integrated service modules. These modules provide services, such as content switching, firewall, SSL offload, intrusion detection, network analysis, and more.
- Access layer—Where the servers physically attach to the network. The server components consist of 1RU servers, blade servers with integral switches, blade servers with pass-through cabling, clustered servers, and mainframes with OSA adapters. The access layer network infrastructure consists of modular switches, fixed configuration 1 or 2RU switches, and integral blade server switches. Switches provide both Layer 2 and Layer 3 topologies, fulfilling the various server broadcast domain or administrative requirements.

Data Centre Requirements

The internet is a universal network – and its 'universality' is only expanding further! As web technologies are maturing and the number of internet-enabled devices in the average business or household increases, there is a corresponding increase in demand for data centers.

A data center is a facility that houses hardware pertaining to data storage, processing and transfer. A regular data center contains hardware units catering to the computing, data storage and networking needs of the system. They are designed to centralize data processing and keep processes running with as little downtime as possible.

While many people have a broad understanding of what goes into building a data center, here's a list of the hardware components required to build a data center to help provide a better understanding.

Servers

Servers are the heart of a data center. These units of hardware are responsible for storage, analysis and transmission of data. The three major types of servers are blade servers, rack servers and tower servers. Modern data centers use rack servers, wherein multiple servers are stacked within a rack.

Racks

Selecting the right racks is an important consideration in the setup or management of a data center, and is also a factor that's often overlooked. Server racks are of two major

kinds – Open Racks and Cabinets – and support the physical weight of the servers, storage devices and switches. While Open Racks allow for better visibility and easier handling, Cabinets are the more secure and stable units. It is important to select the right racks or cabinets for your requirements, in order to balance accessibility and functionality.

Cables

Coaxial, twister and fiber optic cables are used in data centers to connect switches, storage devices and servers. Fiber optic cables allow for faster transmission, and they are therefore used heavily in modern data centers. As cables are the units ultimately responsible for the smooth and timely transmission of data across the network, they play a central role in the functioning of a data center. The usage of fibre cable raceways helps to protect and segregate fiber cables, for increased efficiency.

Switches

The connector units of data center hardware, switches are responsible for linking multiple devices together. Switch's function based on the hardware addresses of devices.

Storage

NAS (Network Attached Storage) and SAN (Storage Area Network) are the two most commonly used types of storage systems in modern data centers, with SAN providing higher speeds and more efficient access to the stored files and block level data.

Apart from these basic units, without which the Data Center cannot function, many other factors also come into play which are sometimes not given the same importance by those of us setting up or engaging the services of a data center.

Power Equipment

Data centers require 24x7 power availability and must therefore employ backup generators and multiple data routes to ensure Uninterrupted Power Supply (UPS). This is true no matter the physical location of the center, but is especially the case when the data center is located in a country or area where power supply is known to be unreliable.

Security Systems

The physical space where the data center is set up must be resistant to natural disasters like flooding, earthquakes, fires, etc. This is of course in addition to physical security systems and access control systems to keep the data private.

Cooling Systems

Data centers require that multiple servers be placed in close proximity to one another, in a limited space. That being the case, there are high chances of equipment overheating rapidly, in spite of the units' internal fans. Server rooms must therefore be maintained at a specific temperature as prescribed and planned during setup, through the use of external HVAC equipment, for smooth functioning.

Policies and Procedures

Maintenance and management of the physical hardware requires the strict adherence to certain policies and procedures that need to be followed. This could include scheduled servicing, downtime protocols, or even housekeeping.

Data Center Prerequisite

Structure
Cabling performance
Redundancy
Grounding/potential equalization
Tier classification
Cable routing
Ceilings and double floors
Floor load
Space requirements (ceiling height, door width)
Power supply/UPS
Fire protection/safety
Cooling
Lighting
Administration/labeling
Temperature/humidity

DataCenter

Required Physical Area for Equipment and Unoccupied Space

Designing a data center is a huge task that requires a lot of time, effort, and expense. When done properly, a data center facility can house servers and other IT equipment for decades into the future. Whether planning out a modest facility for a specific company, or a massive, million-plus square foot facility for cloud technologies, doing

everything properly is critical. The tips listed here will give you a great place to start in your data center design.

Leave Room for Growth

Investing in a data center is going to be extremely costly. Whether retrofitting an existing facility or creating a new one, the expense is going to be significant. This is why many businesses create a data center design that meets their current needs, but don't invest in the future. Consider the following points when looking at the needs your company will have down the road:

- **Floorspace** – How many square feet of floorspace do you need today? Do you expect this to grow over time? It is much less expensive to build what you need now than to try to perform a renovation in a few years.
- **Power Requirements** – The electrical needs of a data center can be quite massive. Take time to plan out the needs you have today, and the potential requirements you will have in the future.
- **Cooling Requirements** – As you add more and more hardware into a data center, the heat produced will need to be eliminated. New cooling units are extremely costly, so investing in the right ones up front is essential.
- **Server Space** – Choosing the right server racks now will allow you to house your equipment properly while leaving space for growth as well. Many new data centers have rows of empty racks that help to facilitate proper airflow until they are filled.

Plan for the Support Team

When planning a data center design, most of the effort is going to be focused on the actual areas where the servers and other equipment are kept. Another important aspect, however, is going to be where the support team must work from. In most cases, this is going to be an office area just outside the data center itself.

This team must use advanced monitoring equipment to see what is going on in the data center at all times. In most cases, the staff will include both IT professionals who will support the hardware and software within the data center, and other personnel who will support the cooling systems, humidity levels, wiring, physical server racks, and more. Having a place for them to work will allow for rapid responses to outages and easy access for ongoing maintenance and upgrades.

Optimize Data Center Cooling

Keeping your equipment operating at appropriate temperatures is one of the most important things one must consider when designing a data center. If your facility gets too hot, it will cause potentially catastrophic failures in hardware, which could cost millions to replace. For this reason, you should spend a significant portion of your time on planning the cooling and airflow systems for your data center.

The first thing to look at will be what type of air-cooling system should be used. There are quite a few options available, and the one you choose will depend on things like budget, region of the world, electricity cost, and more. These are some of the most popular types of cooling systems available today.

- **Traditional Air Conditioning** – Industrial air conditioners create reliably chill air and bring it where it needs to go. While energy-intensive, air conditioning units will be able to keep a data center at the precise temperature you need.
- **Water Cooling Units** – Water cooling is a lot more efficient than many other methods. People often build data centers near the ocean or other large bodies of water specifically so this type of cooling is an option.
- **Outdoor Air Cooling** – In regions where the temperature outdoors is cold for much of the year it is possible to use outdoor air for cooling a data center.
- **Localized Cooling** – This is an option where a cooling unit (or multiple units) is placed in each 'warm row' of the data center so the air doesn't have to be transported through ducting, which can make it more efficient. It also allows for precision cooling based on the needs of your equipment.

Smart Data Center Airflow Management

The cooling is just the first step in keeping the temperature where it needs to be in a data center. The next step is airflow. Bringing the cool air where it needs to be, and getting rid of the heated air, is extremely important. Not only is this critical for temperature control, but also for keeping the cooling costs as low as possible.

A smart airflow plan can reduce your cooling expenses by as much as 40% in many cases. There are multiple steps to planning out your airflow strategy, including:

- **Main Intake & Exhaust** – Planning where the cool air will come in and where the warm air will exit is critical. Learn about a concept called hot aisle/cold aisle to get the best results in this area.
- **Server Rack Airflow** – Having proper airflow within each server rack is essential. Using filler panels (also called blanking panels), for example, will make planning your strategy more effective.
- **Segmented Aisles** – If using the hot aisle/cold aisle strategy, it is important to use physical barriers above and around each aisle to direct the air where it belongs.

Don't Neglect Physical Security

Data centers contain a lot of very expensive equipment. On top of that, most facilities will have important, or even sensitive, data flowing in and out all the time. Keeping all of this safe is one of the most important reasons why data centers are built. With this in mind, it is critical to consider physical security when designing a new facility.

Even if designing a modest data center with no sensitive information, it is still important to ensure only authorized people are coming and going. This is because it is necessary

to know who is working on machines, and why. On top of that, every time someone comes into the data center, they are bringing with them dust and other contaminants that should be kept to a minimum, which is why only authorized individuals should enter.

For large facilities, or those that contain highly sensitive data, physical security becomes even more important. Many large-scale data centers will have physical barriers preventing unauthorized access to the property. Once on the property, there are locked doors with security guards required to get into the building itself. Finally, to get into the actual data center, there are more secured doors. In most cases, these doors come with biometric scanners to ensure only approved individuals can gain access.

Finally, investing in enclosed server racks also helps in long term data center physical security. These server racks help to prevent both theft and potential damage, as they are typically lockable. Data center managers can control who truly has access to the servers.

Focus on Proper Wiring from the Beginning

A smart wiring strategy will help to reduce outages, increase the speed troubleshooting can occur, make adding new equipment easier, and much more. Without a good wiring plan in place, a data center can become a huge mess very quickly.

There are two main areas where a good cable management plan will be necessary. First, within the server racks. Dozens, or even hundreds, of cables have to come in and out of server racks, so it is important to ensure they are run properly to avoid tangles and other issues. Using horizontal or vertical cable managers can help tremendously with this. Next, the cables must be run neatly to where they are going. This typically means running them either under the flooring or in the ceiling, depending on the data center design. In all of these areas, make sure to properly label the cables on both ends for easier troubleshooting.

Plan Properly for Your Data Center Design

Taking the planning of your data center design seriously is very important. These tips are a great place to start and will help ensure your facility operates well from the very beginning.

Selecting a Geographic Location

IT Data Centers are a huge investment and most midsize companies need to setup their own data centers to run their business operations. Building new data centers is a time-consuming effort needing millions of dollars and months of planning. Companies cannot afford to make a wrong decision when designing and setting up a new data center as these need to build for a life of around 20 years.

This article describes the challenges faced by enterprise data centers and explores various criteria which can be used for selecting a location for a new data center. Most of the businesses run their operations based on information technology processes and data centers form the heart of IT infrastructure of any organization. For all medium and multinational organizations, data centers form the backbone of critical business processes.

Variety of services like web hosting, telecom services, banking services depend on the data center infrastructure to offer services to clients. With the advent of ecommerce, smart phones, social networking and mobile communications, requirements of enterprise-wide data centers keep on getting complex.

Companies like Facebook, Netflix and Google cannot afford a downtime of even a single minute. They need a robust facility which can be operated 24×7 without disruptions and provide scalable computing power to customers around the world.

Companies want to build data centers in locations which are not prone to natural disasters, terrorism and also provide cheap and reliable resources like electricity, network and transport facilities.

Although Data Centers are such a critical part of any business, decisions regarding locations are sometimes arbitrary. In the past, companies have mostly setup their Data Centers near to their offices. This is rapidly changing now. As most of the monitoring and equipment administration work can be done remotely. Companies are looking for different sites for their data centers and proximity to the corporate office is not required any more. Main economic, political, geographical & social factors which influence the decision regarding the site are discussed below:

(1) Disaster Avoidance

Reducing the physical risk of catastrophic facility failure starts with site selection. Beyond the direct physical threat posed by natural disasters (earthquakes, hurricanes, lightning storms, Tornadoes, floods, etc.) data centers can be crippled by damage to supporting infrastructure or rendered inoperable because key employees and vendors are unable to reach the facility. When deciding where to locate a data center, CIOs must choose an area with a low risk of natural disasters. Areas which fall in earthquake zones should be strictly avoided.

(2) Network Carrier Availability

Communications is at the core of a data center functioning. All equipment in a data center needs to communicate with devices spread around the world for different business processes. Companies need good fiber connectivity. Location should have more than one network carrier for redundancy. Google set up one of its largest data center in Dalles (Oregon, USA) due to availability of cheap power and easy access to fiber optic network. A Data center with poor connectivity is a useless setup.

(3) Availability of Power

In US, Data centers consume power equal to 2500 typical homes. Power requirements have been raising from the data centers and this is the reason that data center energy efficiency is one of the hot topics in IT industry today. Data center electricity costs have become the second highest expense in the data center operations at 13 %. Therefore, cheap and abundant power supply is one of the main criteria for site selection. Water availability and cost also needs to be considered if Data Center design uses water-based cooling.

(4) Transport / Accessibility

Although data centers can be setup in small cities, these should be still easily accessible by any means of transport. Suppliers of IT equipment (like HP, Cisco, Dell, IBM, Intel etc.) and other construction crew need to be able to travel to the site easily. Air connectivity is also preferred so that support personnel can be flown to the site ASAP if needed for outages.

(5) Land and Building Cost

Since Data Centers are massive facilities, cost to acquire the land is a crucial factor in the site selection. Usually, land is cheaper in small remote cities. Land in big cities is 5 times the cost of same land in a small town and data centers need lots of lands. Therefore, building data centers outside the main cities is cost effective. It is also not easy to upgrade the building of a 24×7 data center, therefore Building Construction cost also needs to be looked into account

(6) Tax Structure, Incentives and Subsidies

Political Factors impacts all investment decision even if the operation of a data center is a purely technical issue. Corporate Tax structure varies in different states. As data centers involve massive capital investments, companies look for ways to reduce their taxes. Locations offering lower taxes appeal to companies interested in making huge capital investments in data centers. For example, Washington State has one of the lowest corporate taxes in the US. Oregon State also offers lots of tax incentives to companies setting up Data Centers. Taxes is one of the main reasons that Amazon, Microsoft, Yahoo, Providence have setup their data centers in different cities of Washington & Oregon.

(7) Availability of Skilled Manpower

Data Centers not only create lots of job during construction, but they also need on site engineers – electrical/telecom/IT engineers to operate the site. Therefore, IT managers need to consider the local talent pool when looking at a potential data center site.

(8) Safety and Security

Crime rate of the city needs to be checked. Also, terrorism threat needs to be evaluated for the city being considered for building the new data center. Physical safety of the data centers is of extreme importance as any disturbance to the facility

can impact business processes. Data Centers host critical business data and data integrity cannot be compromised. Therefore, Data Center access should be controlled and regularly monitored.

(9) Urban Planning and Environment
What are the environmental and business laws of the various cities? Data Center is not a very environmentally friendly business due to its huge energy consumption. Some communities do not want any industries. Permits to build data centers are required from local municipalities – which can be time consuming. Therefore DC planners need to be very conscious of the communities when looking for cities for building new data centers.

(10) Climate Conditions
Environment and climate are very important when considering a location for data center. Since data centers need colder temperatures and humidity control, cities with moderate climate are preferred. This is the reason many companies are setting up their new data centers in extremely cold places like Scandinavian countries in Europe and Pacific Northwest in US.

Selecting and Existing building

The first best practices are in the selection process itself.

1. **Define a location selection process to list critical and desirable selection criteria.** Assign a weight or score to each one, compared to the others. Then you can objectively compare different possibilities for sites.
2. **Remember that the selection process should not be limited to the situation today.** Simulate future trends and needs to make sure a choice made now will also be the right one into the future.

Security and Safety Considerations

Security and safety are typically the highest priority in data center site selection:

1. **Avoid high-risk areas,** such as aircraft glide paths, but also proximity to major highways (risk of fuel truck accidents.)
2. **Choose a site that has easy access for emergency responders.**
3. **Choose a site with good air quality** to protect the health of your employees, as well as to lower costs and avoid equipment malfunction.
4. **In a building with multiple tenants, prefer the end of the building** rather than the middle, to minimize disruption caused by other tenants.

Building Considerations

For the data center building:

1. **For existing construction, single story buildings with large floor areas are often best.** Lower rental and operating costs, better security, and higher flexibility are the main advantages.
2. **Check there is sufficient area around the building for parking, water, and fuel storage, as well as for access for delivery trucks.** Building and parking lot expansion needs should be evaluated and checked too.

Real Estate and Negotiations Concerns

The cost of real estate and negotiations can be showstoppers:

1. **Besides aiming for a cost per square foot that falls within your budget, consider possible add-on costs** for upgrading power, networking, and any other necessary facilities.
2. **For speedier negotiations, prefer sites with single owners,** rather than multiple owners.
3. **Check on the availability of public incentives,** offered for example by municipalities to attract high-technology businesses.
4. **Make sure you know about any site or zoning restrictions** that could affect the type of building or operations you plan (for example, the operation of diesel generators.)

Climate Considerations

Climate and natural conditions can affect the quality of a site for a data center in many ways.

1. **Get weather data for a sufficiently long period (10 years)** to understand if there is a history of natural disasters. Avoid such sites.
2. **Consider sites that make outside air cooling viable,** thus lowering a major cost in data center operation.
3. **Check that the site humidity ranges are compatible with the IT equipment** you plan to use in the data center (or face extra costs to make this so inside the data center).

Network Connectivity and Power Considerations

Network connectivity and power availability must be properly assessed too.

1. **Ensure you have adequate networking and power** for your needs today, and that these facilities can keep pace with your needs into the future.

2. **Prefer sites that offer redundant, separate network links and power lines.** The best is to have the entry points on opposite sides of the building (north and south for power, east and west for networking, for example.)

Staffing Considerations

Even the most automated data center needs staff to run it:

1. **Moderate economic conditions are often a good trade-off,** with sufficient access to skilled people, but without other expenses driving up overall costs.
2. **Locating a data center near a university or IT training school** can ensure the availability of staff with appropriate skills.
3. **A site offering good commute times and quality of life,** in general, will help avoid staff turnover.

Modular Cabling Design

Data centers possess an immensely intricate cabling structure. A proper cabling design is integral to the functioning of the data center. A minute error in the data center cabling can lead to structural challenges such as spaghetti cabinets, ill-connection between network switches, and tedious product installation. The complexity of data center cable architecture eventually leads to problems in error-identification, maintenance, and troubleshooting. Adding to it, today's modern technologies like the Internet of Things (IoT), cloud computing, big data, etc. demand optimum access to the highest storage capacity of the data center. Again, any errors in data canter cable infrastructure can lead to operational challenges like data loss, insufficient storage, and abrupt breakdown of the data storage system. To avoid these risks, it is essential to integrate data center cabling in specific types of cabling that are called topologies while paying attention to a few significant considerations. This white paper guides readers through ways to perform effective and efficient data center cabling, its considerations, and performance-enhancing strategies.

General Considerations of Data Center Cabling

To perform efficient data center cabling in specific topologies, it is essential to consider a few significant factors. These factors add to the quality and efficiency of the data center cabling.

Data Center Cabling Standards: Increasing data demands and shrinking loss budgets offer several challenges to data center engineers. However, gaining a fair understanding of infrastructure cabling standards may help them to minimize risks.

ANSI/TIA-942: This standard offers specific recommendations for efficient cabling in the data center architecture. These recommendations include a minimum number of pathways, space between horizontal cabling and backbone, environmental considerations, and guidelines for cable management, utility, and redundancy, etc.

ISO/IEC 24764: This standard is a combination of TIA-942 and EN- 50173-5. It describes the types of cabling suitable for specific data center requirements.

ANSI/BICSI 002-2014: This standard offers guidelines for specifics of the design and operation of data center cabling. It includes aspects like infrastructure design and planning, commissioning, maintenance, troubleshooting, etc. It covers minute aspects like pathways, spaces, infrastructure detailing, cabling classes, etc for different types of data centers.

Up-Time and Security: It is essential to determine the data center uptime and security requirements of the overall data center. To increase the performance and life expectancy of the data center efficient establishment of cabling uplinks, power sources and cablings, etc are essential.

Scalability: The network must be scalable. Therefore, it is important to calculate the expectancy of network performance, determine power and resource requirements according to calculated performance needs. The consideration of space between the servers and switches of different functional areas defines the scalability of the network.

Manageability and Performance Flexibility: It is essential to have a manageable and flexible cabling network. Therefore, it is important to consider factors like configuration and reconfiguration frequency, cabling limitations, pathways and routing requirements, etc. The network must be functionally flexible to admit such changes after the first layout of the cabling infrastructure.

Cost of Operation: The data center cabling includes several components like switches, connectors, media converters, etc along with the cables. Thus, capital investment, cost of operation, maintenance and damage coverage, etc. becomes essential considerations.

Functional Areas of Data Center Cabling: The data center is distributed in different functional areas through which the cabling runs and adds value to the data center operation. To establish an efficiently designed data center infrastructure, the consideration of the following functional areas is suggested under the standard TIA-942.

Entrance Room: The entrance room is the main interaction area of the data center from where the data center information can be accessed via an interface. This entrance room provides an access to operational equipment, demarcation points, and switch access to different functional areas.

Main Distribution Area (MDA): The main distribution is area is where functional components are housed. It houses switches, routers, cross-connect, and inter-connect equipment, etc. The LAN essentials are housed in the MDA.

Horizontal Distribution Area (HDA): HDA houses the connecting equipment and switches to interconnect equipment distribution area (EDA) to local area network (LAN), storage network area (SAN), and/or KVM switches.

Zone Distribution Area (ZDA): ZDA is basically a coherent point between the HAD and EDAs.

Equipment Distribution Area (EDA): EDA is the functional zone of the data center that houses end equipment like servers, racks, cabinets, etc. Also, the HAD is terminated in the EDA region by using patch panels.

Backbone Cabling: Backbone cabling is the cable line that interconnects all other functional areas of the data center cabling. Error in backbone cabling can terminate the function of all other functional zones of data center cabling. That is why critical care of backbone cabling is essential to perform.

Taking all the above-mentioned factors into consideration, commonly three topologies of data center cabling are adopted across industries. These strategically planned topologies ensure lesser possibilities of errors, easy maintenance, and highly-efficient performance of the data center infrastructure.

Data Center Cabling Designs/Topologies

In order to successfully connect the data center cables, one of the following topologies can be utilized as per the performance requirements.

Centralized Cross-Connect Topology: Centralized cross-connect data center cabling can be adopted in two ways:

Centralized Cross-Connect via MDA: In this type of topology, the backbone cable directly connects the switches from the main distribution area (MDA) to the horizontal distribution area (HDA), which is further centralized by cross-connecting the distribution rows or lines from HDA to the equipment distribution areas (EDAs). The distribution lines from HDA to the EDAs are connected by using three-connector channels, which allows the HDA lines to terminate at patch panels and replicate or mirror it to the switch ports of EDA. The termination point of the first patch panel is further cross-connected with the second patch panel so that each switch port from individual EDAs are cross-connected in the network.

Centralized Cross-Connect via ZDA: To implement this type of centralized cross-connect data center cabling, a four-connector channel is utilized. At first, the backbone cable connects the zone distribution area (ZDA) to the HDA. Further, by using a four-connector channel, the ZDA distribution lines are cross-connected to the patch panels of EDAs. However, the patch panels of the four-connector channels are interconnected with a cross-link and one or/and both links are cross-connected to the switch ports of different EDAs. Although centralized cross-connect via ZDA type of data center cabling is not widely adopted, it still can be used for data centers with critical security needs. It secures the cabling by efficient cooling and increases efficiency due to the easy reconfiguration feature.

Benefits of Adopting Centralized Cross-Connect Topology
The following are a few benefits of choosing a centralized cross-connect topology.

- This infrastructure enables easy connections from any switch to any device in the data center cabling network.
- Permanent connections to the switches can be established in order to avoid over-interaction with sensitive devices.
- The cabling infrastructure enables easy configuration and reconfiguration, device addition, line integration, and so on.
- Network flexibility and manageability are higher in these topologies.
- The centralized cross-connect topology assures optimized space utilization which reserves server cabinet spaces.

Partially Centralized Inter-Connect Topology: The partially centralized inter-connect topology of data center cabling features a backbone cabling from the main distribution area that runs through multiple HDAs in a row of specific EDA. While running through HDAs the backbone cable inter-connects the switches from each HDA with the help of two-connector channels. The second distribution lines from the two-connector channels are inter-connected to the patch panel of the EDA.
This topology is can be implemented in either of the following ways.

- **Point-to-Point:** To implement the point-to-point approach to partially centralizing inter-connect topology, the switches from HDAs are allocated at a distance from the EDA. However, no cross-connect or inter-connect approach is deployed in this type. Long patch chords are used to connect the switches from HDAs with the patch panels in the EDA.
- **One-Connector Channel:** To implement a one-connector channel approach, one-connector channels are used instead of two-connector channels. The patch panels are only deployed in the HDAs and then they are directly connected to different equipment housed in the EDA. This approach is suitable for smaller data center applications. It also allows complete utilization of mounting rail in the server cabinets.

Benefits of Adopting Partially Centralized Inter-Connect Topology
The benefits of partially centralized inter-connect topology are listed below.

- It reduces the requirement of copper or optical cabling run.
- It optimizes the pathway spaces and increases the safety of the connections.
- The number of terminations is comparatively lesser than other types of cabling due to a lesser number of switches and patch panels being interconnected.
- It allows the switches and patch panels to locate in the adjacent cabinets without jeopardizing the cabling efficiency.

Distributed In-cabinet Switching: In the distributed in-cabinet switching topology of data center cabling, the connections are simplified by eliminating the use of HDAs totally. This topology features a backbone cable that connects the MDA directly to different cabinets or switches of EDAs. This way, the backbone is directly connected to performing devices in the data center connections.
This type of cabling also features a direct connection to the storage area network (SAN) which further simplifies the data storage operations in the data center.

Benefits of Adopting Distributed In-cabinet Switching Topology
The following are the benefits of implementing distributed in-cabinet switching topology in the data center cabling.

- This type of cabling drastically reduces cabling complexity and expenses.
- For applications with high-bandwidth like 40gbps to 100gbps, this topology gives maximum returns on investment (ROI).
- In special applications where a specific functional area of the data center is required to be separated or secured.
- It is an effective solution for limited cabinet space availability.

Important Devices for Data Center Cabling

In addition to the aforementioned data center cabling topologies, it is also important to choose the right devices for building a robust data center cabling infrastructure. The following are a few important devices for consideration.

- **Media Converters**: Media conversion has emerged as a critical function in enterprise data centers. This is because most data centers utilize hybrid cabling infrastructure where copper cables and fiber optic cables are used together. These cable types have different media, which possess challenges when used together. This is where media converters help. The media converters enable data center managers to optimize the value of existing legacy infrastructure while taking advantage of fiber optic cables. Copper to fiber media converters and fiber to fiber media converters are two popular types of media converters used in data centers.
- **Network Switches**: The enterprise network infrastructure underwent a great change in the decade 2010-2020. With increasing emphasis on network-function virtualization (NFV), software-defined networking (SDN), and software-defined WAN (SD-WAN), choosing the right network switches have become important for data center engineers. You can see different types of data center-class network switches in the market. They are available with PoE/PoE+ options with flexible ports in managed and unmanaged versions.
- **Optical Transceivers**: Until a decade ago, fiber optic technology was only used to mitigate critical interconnection challenges. However, today, it has evolved as a critical element of data centers. The fiber optic cables and devices assure

better scalability, security, and performance, which makes them an integral part of data center cabling. Optical transceivers with SFP modules are used to extend distances, reduce power consumption, optimize cabling density, and leverage existing multimode fiber. Owing to all these reasons choosing the right optical transceivers is important. Today, these transceivers are available in various specifications such as 10/100/1000Mbps, 1000 Mbps, 100 Mbps, and 10 Gbps speeds.

Guidelines to Improve the Performance of Data Center Cabling

In order to enhance the efficiency of a data center, several tactics are being utilized. Implementing fiber optic cables instead of copper cables for long run cross or interconnections. Even if the data center features legacy devices that are not compatible with direct fiber optics connections, the copper cables from the device switches can be connected to fiber optics cables to get higher data transmission speed, bandwidth, and data security. It is achieved by using devices like media converters, fiber optics switches, fiber optics connectors, etc. as copper-to-fiber moderator equipment. Along with replacing copper cables with fiber optics, there a few tactics that can help in enhancing the overall efficiency of the data center cabling. The guidelines to adopt the tactics are listed below.

- Utilize pre-terminated cabling to perform the connections. It enables plug-and-play operation which saves efforts and time on cabling integrations.
- Use high-quality interconnecting products like media converters, optical transceivers, networking switches, etc.
- Adopt flexible yet structured types of cabling topologies.
- Utilize back-to-back vertical cable managers to secure cable bulking at high-density locations.
- Form patching zones between the equipment and cable distribution lines in order to identify and categorize the elements.
- Prevent cable slack by using Velcro ties and guide the patch chords thoroughly.

Data centre physical security

Physical security of a data center comprises various kinds of built-in safety and security features to protect the premises and thereby the equipment that stores critical data for multi-tenant applications. For the safety and security of the premises, factors ranging from location selection to authenticated access of the personnel into the data center should be considered, monitored, and audited vigorously. To prevent any physical attacks, the following need to be considered:

- proximity to high-risk areas, such as switch yards and chemical facilities

- availability of network carrier, power, water, and transport systems
- likelihood of natural disasters, such as earthquakes and hurricanes
- an access control system with an anti-tailgating/anti-pass-back facility to permit only one person to enter at a time
- single entry point into the facility.

Organizations should monitor the safety and security of the data center rack room with authenticated access through the following systems:

- closed-circuit television (CCTV) camera surveillance with video retention as per the organization policy
- vigilance by means of 24×7 on-site security guards and manned operations of the network system with a technical team
- periodic hardware maintenance
- checking and monitoring the access control rights regularly and augmenting if necessary
- controlling and monitoring temperature and humidity through proper control of air conditioning and indirect cooling
- uninterruptible power supply (UPS)
- provision of both a fire alarm system and an aspirating smoke detection system (e.g., VESDA) in a data center. A VESDA, or aspiration, system detects and alerts personnel before a fire breaks out and should be considered for sensitive areas.
- water leakage detector panel to monitor for any water leakage in the server room
- rodent repellent system in the data center. It works as an electronic pest control to prevent rats from destroying servers and wires.
- fire protection systems with double interlock. On actuation of both the detector and sprinkler, water is released into the pipe. To protect the data and information technology (IT) equipment, fire suppression shall be with a zoned dry-pipe sprinkler.
- cable network through a raised floor, which avoids overhead cabling, reduces the heat load in the room, and is aesthetically appealing.

Data centre Logical security

Logical security partially boils down to cybersecurity. The white paper notes that cybersecurity threats are often just as serious as physical threats to data centers.

The white paper urges data center operators to consider that “despite wildly different approaches and measures, physical and logical/cybersecurity are inextricably linked.” The white paper says that the best data center providers find ways to merge the two disciplines to “create the tightest web of protection between bad actors and the data center.”

Building off the concern over bad actors, Data Centre Dynamics says that data center specialists need to have access to a regularly updated and published threat list containing known dangerous IP ranges.

In addition to cybersecurity, logical security also includes meeting government regulations and compliance mandates. Governments – both national and local – are rapidly passing new data privacy laws. These laws, Data Centre Dynamics explains, add even more complexity to the security landscape.

Data center users must look for a data center with "highly mature operations and processes for managing government regulations, risk, and compliance." The white paper says that an organization having those bases covered provide a "signal that the data center meets or exceeds their contractual duty to maintain compliance and the ability to efficiently perform thorough regulatory audits if necessary."

Reasons for data centre Consolidation

In simple terms, data center consolidation is a way in which businesses streamline their IT systems. Many companies will have grown rapidly over the years, and it's often the case that IT systems have struggled to keep up with increasing demand. And that's why a huge number of companies stand to benefit from data center consolidation. By consolidating, companies can make IT architecture far more efficient. In the long run, the strategy saves time and money, while also paving the way for future growth by maximizing capabilities.

How to consolidate: a step-by-step guide

There are a number of different methods commonly used by IT professionals working on data consolidation projects. In our experience, the most effective methods tend to incorporate a hybrid cloud work environment, which enables more effective running on fewer resources. Take a look at our step-by-step guide to data center consolidation for a brief guide to how it works.

Step 1: Capacity planning

First things first, think about space. Consider whether or not you're using all the space you currently have, and whether the unused space in your data center is costing you money. Far too many companies end up paying for space that they've never used and have no plans to use, simply because their existing contracts are inflexible. This is something you'll definitely want to avoid, so make it a priority in new data center consolidation plans.

Look into data centers offering cage free infrastructure, which will allow you to scale up or down as required. Other things to look out for in new data centers include free

cross connects, set up, rack and stack, and free remote hands-on that will consolidate unused space or cabinets for you, as time goes on and requirements change.

Step 2: Remove or replace inefficient hardware and software
We bet you've already heard all about the Marie Kondo method, the trailblazing organization strategy that's taken the world by storm, sparking so much joy in the process! But did you know you can apply the same approach in data center consolidation?

Take a look at your existing hardware - if you immediately start to feel frustrated then you can safely say that it doesn't spark joy, and it needs to be replaced or upgraded. Remove inefficient hardware or software and look out for better alternatives that'll help you stay one step ahead.

Step 3: Enjoy the fruits of your labor
Once you've finished consolidating you'll be able to sit back, relax and enjoy your hard work. Consolidation can be enormously helpful for all kinds of businesses, helping to improve their total cost of ownership and allowing teams to regain control over their data center space. If this sounds like something your company might find useful, now's the time to look into data center consolidation. Start thinking about consolidation and considering how you might approach the task, and before long you'll have a thorough plan of action ready to go.

Introduction of virtualization

Virtualization is a technique to divide the computer resources logically. It's achieved by abstracting away the underlying complexity of resource segregation. Although an old technology, it's still a popular technique and highly relevant in this era of cloud computing.

The hypervisor manages the virtualization technique and creates, runs, and monitors multiple virtual machines (guest) simultaneously, on single computer hardware (host).

So, hypervisors regulate the virtualization process, creates multiple virtual machines that allow you to work on several computing instances at once. This is the key difference between Virtualization and Hypervisors.

The Virtual Machine Monitor or VMM or a Hypervisor act as a supervisor. It's implemented on computer hardware as code embedded in a system's firmware or as a software layer.

Hypervisors create, start, stop, and reset multiple VMs while virtually sharing its resources like RAM and Network interface controller.

VMM governs the guest operating systems and manages execution on a virtual operating platform. It furthermore separates Virtual Machines (VMs) from each other logically, so even if one OS crashes for some reason, the other VMs can function unhindered.

Two Types of Hypervisors: Type 1 and Type 2

Based on their working system Hypervisors are divided into two categories-

- **Type 1 – Bare Metal hypervisor**
- **Type 2 – Hosted hypervisor**

The primary contributor to why hypervisors are segregated into two types is because of the presence or absence of the underlying operating system.

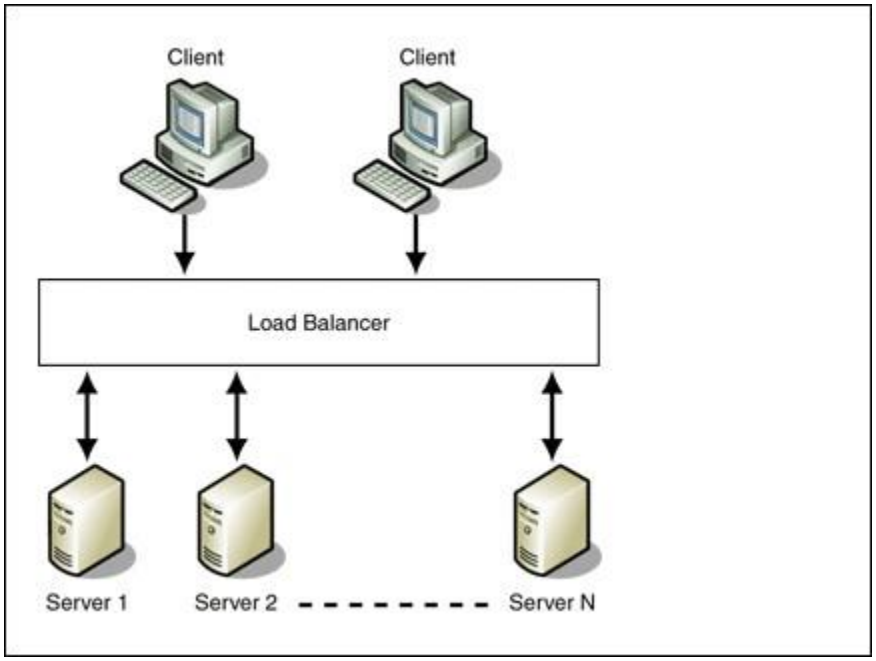
Type 1 runs directly on the hardware with Virtual Machine resources provided. Type 2 runs on the host OS to provide virtualization management and other services.

Hypervisor Type 1 vs. Type 2 in Tabular Form

Criteria	Type 1 hypervisor	Type 2 hypervisor
AKA	Bare-metal or Native	Hosted
Definition	Runs directly on the system with VMs running on them	Runs on a conventional Operating System
Virtualization	Hardware Virtualization	OS Virtualization
Operation	Guest OS and applications run on the hypervisor	Runs as an application on the host OS
Scalability	Better Scalability	Not so much, because of its reliance on the underlying OS.
Setup/Installation	Simple, as long as you have the necessary hardware support	Lot simpler setup, as you already have an Operating System.
System Independence	Has direct access to hardware along with virtual machines it hosts	Are not allowed to directly access the host hardware and its resources
Speed	Faster	Slower because of the system's dependency
Performance	Higher-performance as there's no middle layer	Comparatively has reduced performance rate as it runs with extra overhead
Security	More Secure	Less Secure, as any problem in the base operating system affects the entire system including the protected Hypervisor
Examples	• VMware ESXi • Microsoft Hyper-V • Citrix XenServer	• VMware Workstation Player • Microsoft Virtual PC • Sun's VirtualBox

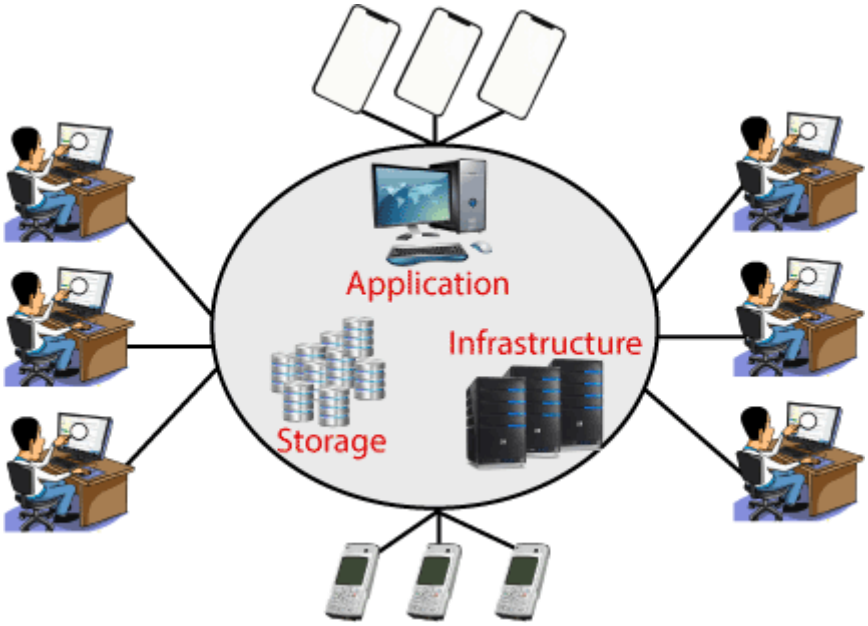
Cluster Architecture

In cluster architecture, requests or parts of the user requests are divided among two or more computer systems, such that a single user request is handled and delivered by two or more than two nodes (computer systems). The benefit is unquestionably the ability of load balancing and high-availability. How? If one node fails, the request is handled by another node. Hence, there are less or negligible chances of complete system failures.



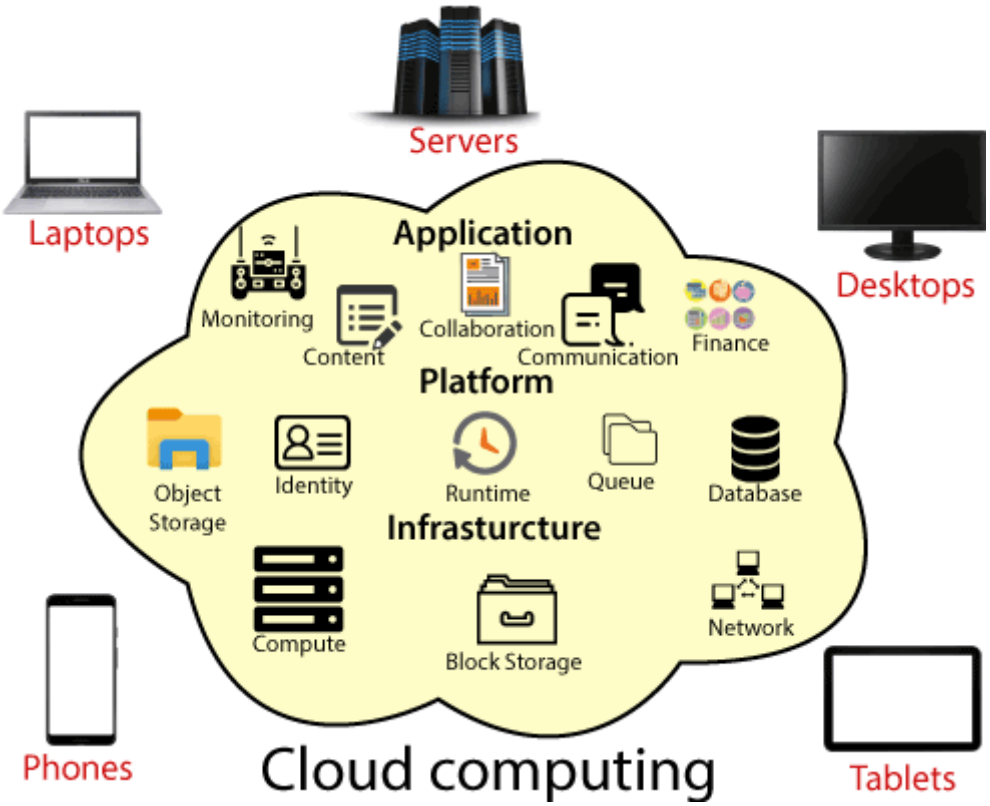
Introduction to Cloud

Cloud Computing is the delivery of computing services such as servers, storage, databases, networking, software, analytics, intelligence, and more, over the Cloud (Internet).



Cloud Computing provides an alternative to the on-premises datacentre. With an on-premises datacentre, we have to manage everything, such as purchasing and installing hardware, virtualization, installing the operating system, and any other required applications, setting up the network, configuring the firewall, and setting up storage for data. After doing all the set-up, we become responsible for maintaining it through its entire lifecycle.

But if we choose Cloud Computing, a cloud vendor is responsible for the hardware purchase and maintenance. They also provide a wide variety of software and platform as a service. We can take any required services on rent. The cloud computing services will be charged based on usage.



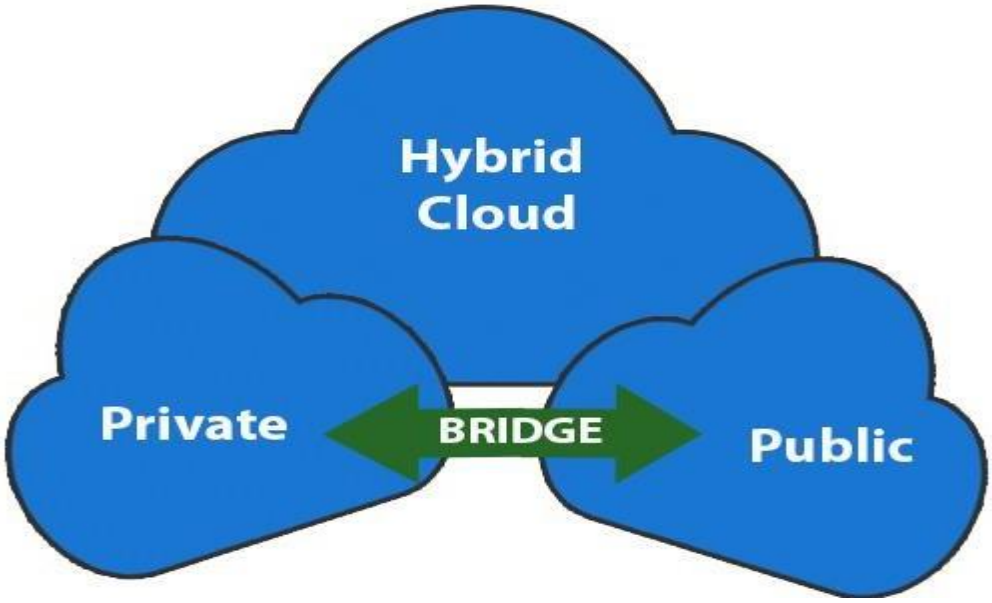
The cloud environment provides an easily accessible online portal that makes handy for the user to manage the compute, storage, network, and application resources. Some cloud service providers are in the following figure.

Advantages of cloud computing

- **Cost:** It reduces the huge capital costs of buying hardware and software.
- **Speed:** Resources can be accessed in minutes, typically within a few clicks.
- **Scalability:** We can increase or decrease the requirement of resources according to the business requirements.

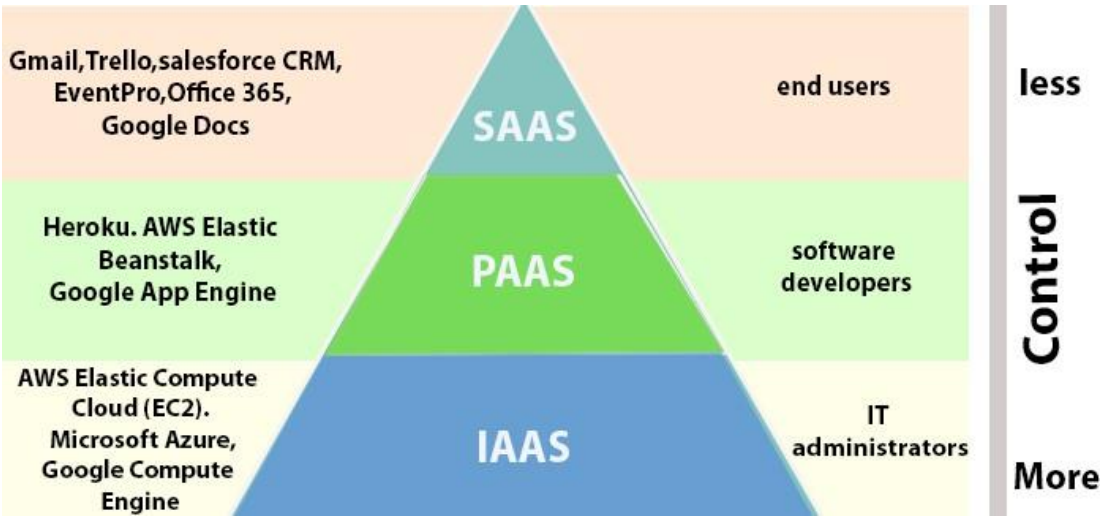
- **Productivity:** While using cloud computing, we put less operational effort. We do not need to apply patching, as well as no need to maintain hardware and software. So, in this way, the IT team can be more productive and focus on achieving business goals.
- **Reliability:** Backup and recovery of data are less expensive and very fast for business continuity.
- **Security:** Many cloud vendors offer a broad set of policies, technologies, and controls that strengthen our data security.

Types of Cloud Computing

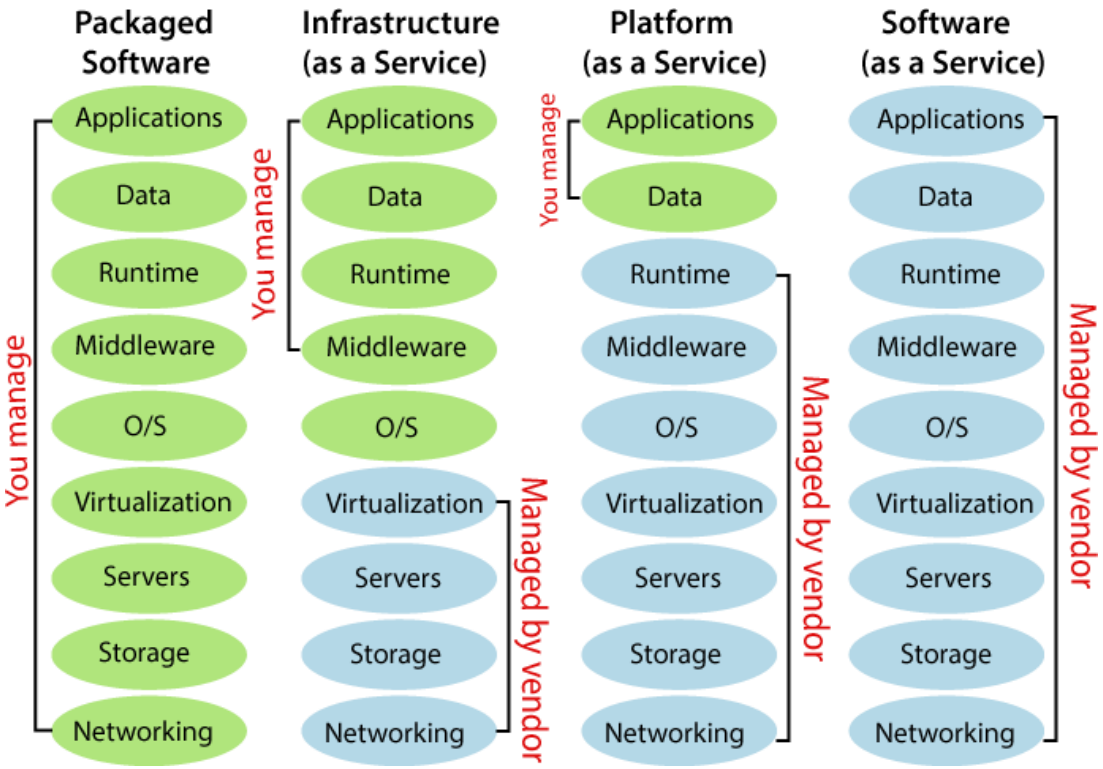


- **Public Cloud:** The cloud resources that are owned and operated by a third-party cloud service provider are termed as public clouds. It delivers computing resources such as servers, software, and storage over the internet
- **Private Cloud:** The cloud computing resources that are exclusively used inside a single business or organization are termed as a private cloud. A private cloud may physically be located on the company's on-site datacentre or hosted by a third-party service provider.
- **Hybrid Cloud:** It is the combination of public and private clouds, which is bounded together by technology that allows data applications to be shared between them. Hybrid cloud provides flexibility and more deployment options to the business.

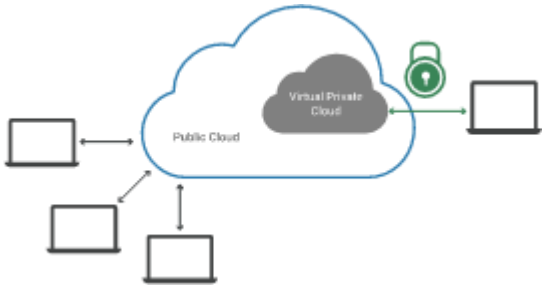
Types of Cloud Services



1. **Infrastructure as a Service (IaaS):** In IaaS, we can rent IT infrastructures like servers and virtual machines (VMs), storage, networks, operating systems from a cloud service vendor. We can create VM running Windows or Linux and install anything we want on it. Using IaaS, we don't need to care about the hardware or virtualization software, but other than that, we do have to manage everything else. Using IaaS, we get maximum flexibility, but still, we need to put more effort into maintenance.
2. **Platform as a Service (PaaS):** This service provides an on-demand environment for developing, testing, delivering, and managing software applications. The developer is responsible for the application, and the PaaS vendor provides the ability to deploy and run it. Using PaaS, the flexibility gets reduce, but the management of the environment is taken care of by the cloud vendors.
3. **Software as a Service (SaaS):** It provides a centrally hosted and managed software services to the end-users. It delivers software over the internet, on-demand, and typically on a subscription basis. E.g., Microsoft One Drive, Dropbox, WordPress, Office 365, and Amazon Kindle. SaaS is used to minimize the operational cost to the maximum extent.



Introduction Virtual Private Cloud (VPC)



A virtual private cloud (VPC) is a secure, isolated private cloud hosted within a public cloud. VPC customers can run code, store data, host websites, and do anything else they could do in an ordinary private cloud, but the private cloud is hosted remotely by a public cloud provider. (Not all private clouds are hosted in this fashion.) VPCs combine the scalability and convenience of public cloud computing with the data isolation of private cloud computing.

Imagine a public cloud as a crowded restaurant, and a virtual private cloud as a reserved table in that crowded restaurant. Even though the restaurant is full of people, a table with a "Reserved" sign on it can only be accessed by the party who made the reservation. Similarly, a public cloud is crowded with various cloud customers

accessing computing resources – but a VPC reserves some of those resources for use by only one customer.

What is a public cloud? What is a private cloud?

A public cloud is shared cloud infrastructure. Multiple customers of the cloud vendor access that same infrastructure, although their data is not shared – just like every person in a restaurant order from the same kitchen, but they get different dishes. Public cloud service providers include AWS, Google Cloud Platform, and Microsoft Azure, among others.

The technical term for multiple separate customers accessing the same cloud infrastructure is "multitenancy" (see What Is Multitenancy? to learn more).

A private cloud, however, is single-tenant. A private cloud is a cloud service that is exclusively offered to one organization. A virtual private cloud (VPC) is a private cloud within a public cloud; no one else shares the VPC with the VPC customer.

How is a VPC isolated within a public cloud?

A VPC isolates computing resources from the other computing resources available in the public cloud. The key technologies for isolating a VPC from the rest of the public cloud are:

Subnets: A subnet is a range of IP addresses within a network that are reserved so that they're not available to everyone within the network, essentially dividing part of the network for private use. In a VPC these are private IP addresses that are not accessible via the public Internet, unlike typical IP addresses, which are publicly visible.

VLAN: A LAN is a local area network, or a group of computing devices that are all connected to each other without the use of the Internet. A VLAN is a virtual LAN. Like a subnet, a VLAN is a way of partitioning a network, but the partitioning takes place at a different layer within the OSI model (layer 2 instead of layer 3).

VPN: A virtual private network (VPN) uses encryption to create a private network over the top of a public network. VPN traffic passes through publicly shared Internet infrastructure – routers, switches, etc. – but the traffic is scrambled and not visible to anyone.

A VPC will have a dedicated subnet and VLAN that are only accessible by the VPC customer. This prevents anyone else within the public cloud from accessing computing resources within the VPC – effectively placing the "Reserved" sign on the table. The VPC customer connects via VPN to their VPC, so that data passing into and out of the VPC is not visible to other public cloud users.

Some VPC providers offer additional customization with:

- **Network Address Translation (NAT):** This feature matches private IP addresses to a public IP address for connections with the public Internet. With NAT, a public-facing website or application could run in a VPC.
- **BGP route configuration:** Some providers allow customers to customize BGP routing tables for connecting their VPC with their other infrastructure. (Learn how BGP works.)

What are the advantages of using a VPC instead of a private cloud?

Scalability: Because a VPC is hosted by a public cloud provider, customers can add more computing resources on demand.

Easy hybrid cloud deployment: It's relatively simple to connect a VPC to a public cloud or to on-premises infrastructure via the VPN.

Better performance: Cloud-hosted websites and applications typically perform better than those hosted on local on-premises servers.

Better security: The public cloud providers that offer VPCs often have more resources for updating and maintaining the infrastructure, especially for small and mid-market businesses. For large enterprises or any companies that face extremely tight data security regulations, this is less of an advantage.

Introduction to AWS

Amazon Web Services (AWS) is a cloud service from Amazon, which provides services in the form of building blocks, these building blocks can be used to create and deploy any type of application in the cloud.

These services or building blocks are designed to work with each other, and result in applications that are sophisticated and highly scalable.

What are the services provided by AWS?

Each type of service in this "What is AWS" blog, is categorized under a domain, the few domains which are widely used are:

- Computer
- Storage
- Database
- Migration
- Network and Content Delivery
- Management Tools
- Security & Identity Compliance
- Messaging

Compute Services
The computer domain includes services related to compute workloads; it includes the following services:

- EC2 (Elastic Compute Cloud)
- Lambda
- Elastic Beanstalk
- Amazon LightSail

Storage Services
The **Storage** domain includes services related data storage; it includes the following services:

- S3 (Simple Storage Service)
- Elastic Block Store
- Amazon Glacier
- AWS Snowball

Database Services
The Database domain is used for database related workloads, it includes the following services:

- Amazon Aurora
- Amazon RDS
- Amazon DynamoDB
- Amazon RedShift

Migration Services
The Migration domain is used for transferring data to or from the AWS Infrastructure, it includes the following services:

- AWS Database Migration Service
- AWS SnowBall

Networking and Content Delivery Services
The Networking and Content Delivery domain is used for isolating your network infrastructure, and content delivery is used for faster delivery of content. It includes the following services:

- Amazon Route 53
- AWS CloudFront

Management Tools
The Management Tools domain consists of services which are used to manage other services in AWS, it includes the following services:

- AWS CloudWatch
- AWS CloudFomation
- AWS CloudTrail

Security & Identity, Compliance Services
The Security & Identity, Compliance domain consist of services which are used to manage to authenticate and provide security to your AWS resources. It consists of the following services:

- AWS IAM
- AWS KMS
- AWS Shield

Messaging Services
The Messaging domain consists of services which are used for queuing, notifying or emailing messages. It consists of the following domains:

- Amazon SQS
- Amazon SNS
- Amazon SES
- Amazon Pinpoint

Cloud API integration

API integration refers to the system-to-system connection, via APIs, allowing those two systems to exchange data. APIs are designed so you can use a system remotely and connect systems, people, IoT devices, and more.

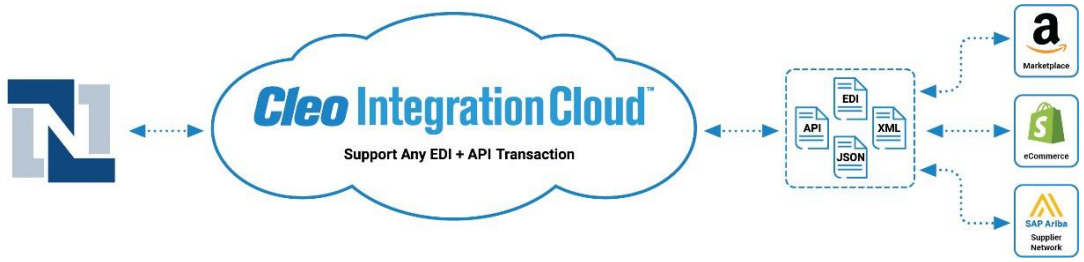
Two or more systems that have APIs can interoperate in real-time using those APIs, which saves time, money, and is far more reliable in terms of information currency and data accuracy.

For example, let us say your company has a TMS (transportation management system) and my company has an ERP (enterprise resource planning) system, and these two systems need to exchange data.

In the old days, we might have faxed or emailed this information or discussed it on the phone.

With API integration, it happens digitally, without human interaction. API integration is what opens a channel that enables our companies to, quite literally, conduct business faster and more accurately.

In this diagram, you can see a visual representation of API integration with a Netsuite ERP instance connecting to Amazon Marketplace, Shopify, and SAP Ariba:



By keeping data in sync in all connected systems, productivity is enhanced, so you can leverage that data to improve efficiency and drive more revenue.