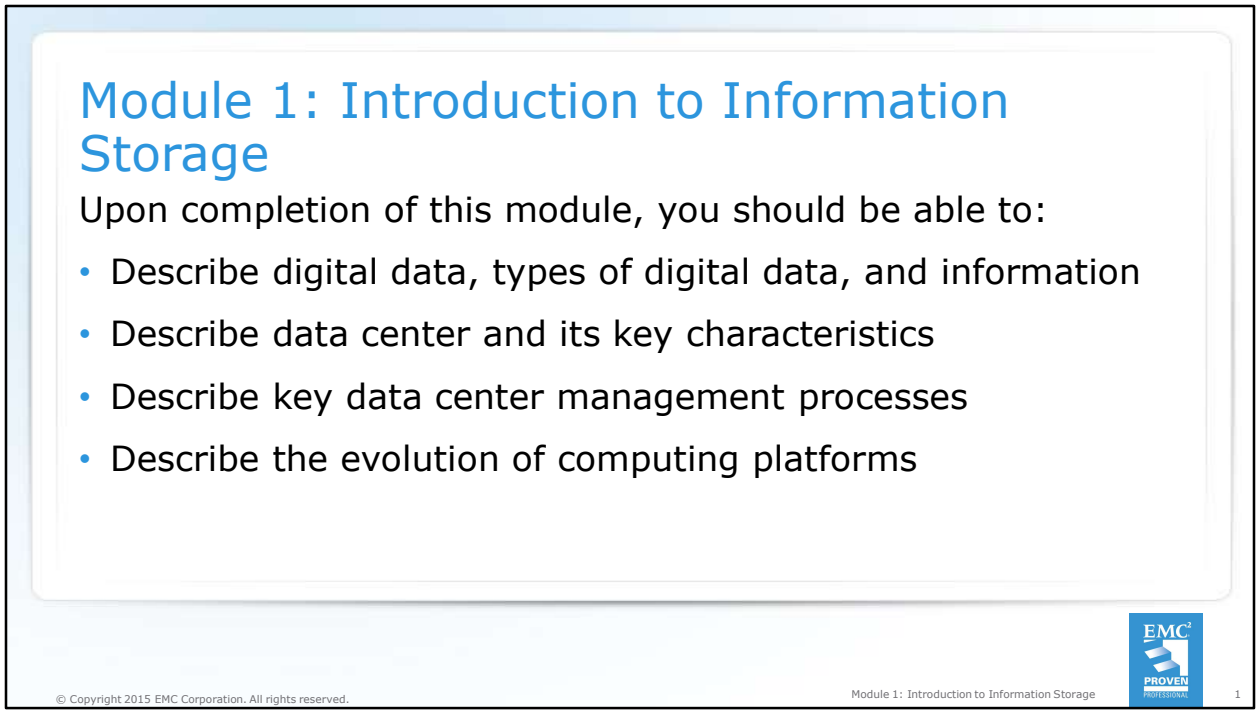


This module focuses on digital data, the types of digital data, and information. This module also focuses on data center and its key characteristics. Further this module focuses on the key data center management processes. Finally, this module focuses on the evolution of computing platforms.

The Growth of the Digital Universe

- The digital universe is created and defined by software
  - Digital data is continuously generated, collected, stored, and analyzed through software
- The digital universe generates approximately 4.4 trillion GB of data annually(44 in 2020) by International Data Corporation (IDC)
  - Proliferation of IT, Internet usage, social media, and smart devices adds to data growth
- The Internet of Things (IoT) is also adding to data growth
  - IoT is made up of Internet-connected equipment and sensors

We live in a *digital universe* – a world that is created and defined by software. A massive amount of digital data is continuously generated, collected, stored, and analyzed through software in the digital universe. **According to the 2014** Digital Universe Study conducted by **International Data Corporation (IDC**), it is estimated that the digital universe produces approximately 4.4 trillion gigabytes (GB) of data annually, which is doubling every two years. By these estimates, it is projected that by the year 2020, the digital universe will expand to 44 trillion GB of data. The data in the digital universe comes from diverse sources, including individuals living and working online, organizations employing information technology (IT) to run their businesses, and from a variety of "smart" electronic devices connected to the Internet.

In organizations, the volume and importance of information for business operations continue to grow at astounding rates. Individuals constantly generate and consume information through numerous activities, such as web searches, e-mails, uploading and downloading content and sharing media files. The rapid proliferation of online social networking and Internet-enabled smartphones and tablets has also contributed significantly to the growth of the digital universe.

The advent of the *Internet of Things* (IoT) is also gradually adding to the growth of the digital universe. The IoT is a technology trend wherein "smart" devices with embedded electronics, software, and sensors exchange data with other devices over the Internet. Examples of such devices are wearable gadgets – smartwatches and fitness activity trackers; electronic sensors – temperature sensors and heart monitoring implants; and household appliances – televisions, thermostats, and lighting. The IoT has vast applications and is driving the development of several innovative technology solutions. Some application areas include weather monitoring – remote monitoring and analysis of temperature and atmospheric conditions; healthcare – health monitoring devices can enable doctors to remotely monitor patients and be notified in case of emergencies; and infrastructure management – technicians can remotely monitor equipment and proactively schedule repair activities for maintenance crews.

Why Information Storage and Management?

- Organizations are dependent on continuous and reliable access to information
- Organizations seek to effectively store, protect, process, manage, and leverage information
- Organizations are increasingly implementing intelligent storage solutions
  - To efficiently store and manage information
  - To gain competitive advantage
  - To derive new business opportunities

Organizations have become increasingly information-dependent in the twenty-first century, and information must be available whenever and wherever it is required. It is critical for users and applications to have continuous, fast, reliable, and secure access to information for business operations to run as required. Some examples of such organizations and processes include banking and financial institutions, government departments, online retailers, airline reservations, billing and transaction processing, social networks, stock trading, scientific research, and healthcare.

It is essential for organizations to store, protect, process, and manage information in an efficient and cost-effective manner. Legal, regulatory, and contractual obligations regarding the availability, retention, and protection of data further add to the challenges of storing and managing information.

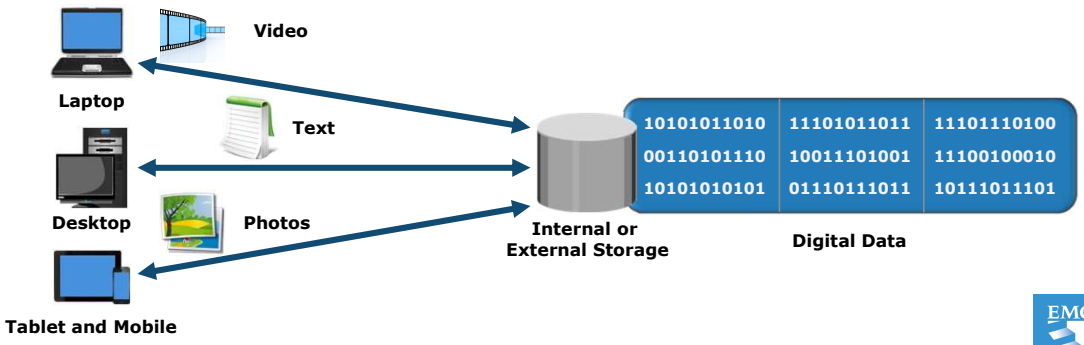
Organizations also face newer challenges in the form of requirement to extract value from the information generated in the digital universe. Information can be leveraged to identify opportunities to transform and enhance businesses and gain a competitive edge. For example, an online retailer may need to identify the preferred product types and brands of customers by analyzing their search, browsing, and purchase patterns. The retailer can then maintain a sufficient inventory of popular products, and also advertise relevant products to the existing and potential customers. Furthermore, the IoT is expected to lead to new consumer and business behavior in the coming years creating new business opportunities.

To meet all these requirements and more, organizations are increasingly undertaking digital transformation initiatives to implement intelligent storage solutions. These solutions not only enable efficient and optimized storage and management of information, but also enable extraction of value from information to derive new business opportunities, gain a competitive advantage, and create new sources of revenue.

What is Digital Data?

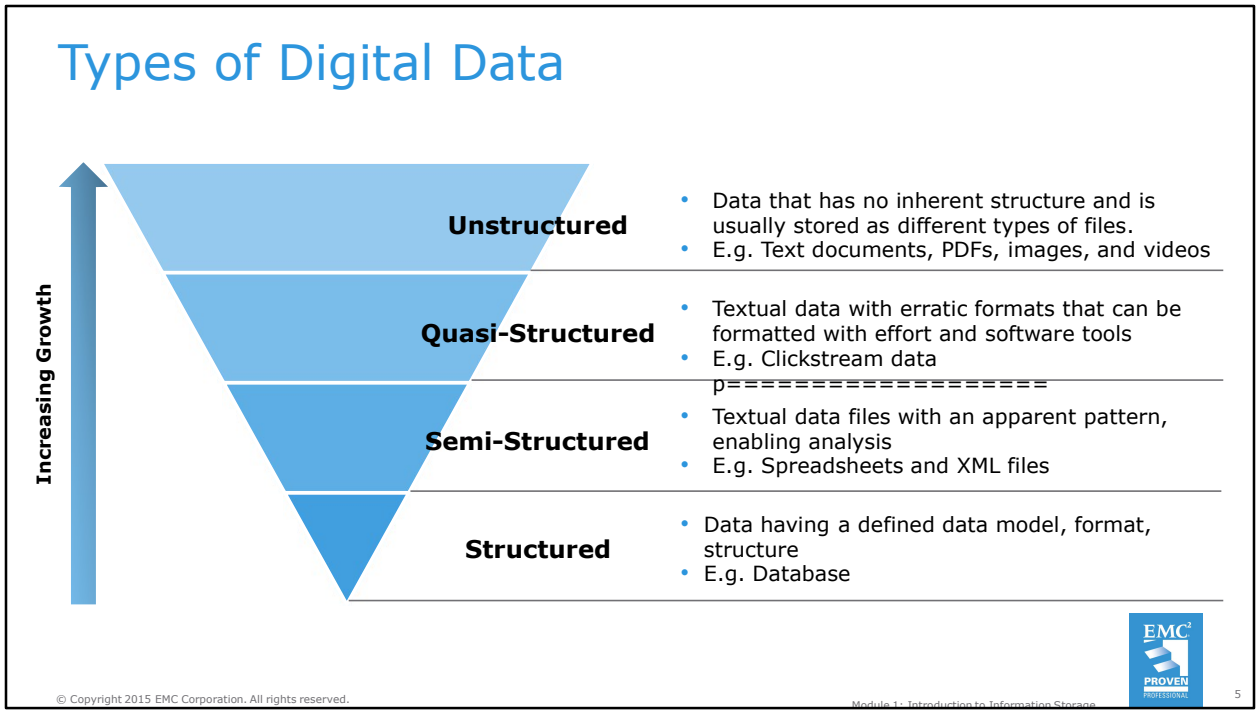
Digital Data

A collection of facts that is transmitted and stored in electronic form, and processed through software.



A generic definition of *data* is that it is a collection of facts, typically collected for the purpose of analysis or reference. Data can exist in a variety of forms such as facts stored in a person's mind, photographs and drawings, alphanumeric text and images in a book, a bank ledger, and tabled results of a scientific survey. Originally, data is the plural form of "datum". However, data is now generally treated as a singular or mass noun representing a collection of facts and figures. This is especially true when referring to digital data.

In computing, *digital data* is a collection of facts that is transmitted and stored in electronic form, and processed through software. Digital data is generated by various devices, such as desktops, laptops, tablets, mobile phones, and electronic sensors. It is stored as strings of binary values (0s and 1s) on a storage medium that is either internal or external to the devices generating or accessing the data. The storage devices may be of different types, such as magnetic, optical, or solid state storage devices. Examples of digital data are electronic documents, text files, e-mails, e-books, digital images, digital audio, and digital video.



Based on how it is stored and managed, digital data can be broadly classified as either structured data or unstructured data. *Structured data* is organized in fixed fields within a record or file. For data to be structured, a data model is required. A *data model* specifies the format for organizing data, and also specifies how different data elements are related to each other. For example, in a relational database, data is organized in rows and columns within named tables. *Semi-structured data* does not have a formal data model but has an apparent, self-describing pattern and structure that enable its analysis. Examples of semi-structured data include spreadsheets that have a row and column structure, and XML files that are defined by an XML schema. *Quasi-structured data* consists of textual data with erratic data formats, and can be formatted with effort, software tools, and time. An example of quasi-structured data is a "clickstream" or "clickpath" that includes data about which webpages a user visited and in what order – which is the result of the successive mouse clicks the user made. A clickstream shows when a user entered a website, the pages viewed, the time spent on each page, and when the user exited. *Unstructured data* does not have a data model and is not organized in any particular format. Some examples of unstructured data include text documents, PDF files, e-mails, presentations, images, and videos.

As indicated by the figure on the slide, the majority, which is more than 90 percent, of the data generated in the digital universe today is *non-structured data* (semi-, quasi-, and unstructured). Although the figure shows four different and separate types of data, in reality a mixture of these is typically generated. For instance, in a call center for customer support of a software product, a classic relational database management system (RDBMS) may store call logs with structured data such as date/time stamps, machine types, and problem type entered by the support desk person. In addition, there may be unstructured or semi-structured data, such as an e-mail ticket of the problem, call log information, or the actual call recording.

What is Information?

Information

Processed data that is presented in a specific context to enable useful interpretation and decision-making.

- Example: Annual sales data processed into a sales report
  - Enables calculation of the average sales for a product and the comparison of actual sales to projected sales
- New architectures and technologies have emerged for extracting information from non-structured data

The terms "data" and "information" are closely related and it is common for the two to be used interchangeably. However, it is important to understand the difference between the two. Data, by itself, is simply a collection of facts that needs to be processed for it to be useful. For example a set of annual sales figures of an organization is data. When data is processed and presented in a specific context it can be interpreted in a useful manner. This processed and organized data is called *information*. For example, when the annual sales data is processed into a sales report, it provides useful information, such as the average sales for a product (indicating product demand and popularity), and a comparison of the actual sales to the projected sales. Information thus creates knowledge and enables decision-making.

As discussed previously, processing and analyzing data is vital to any organization. It enables organizations to derive value from data, and create intelligence to enable decision-making and organizational effectiveness. It is easier to process structured data due to its organized form. On the other hand, processing non-structured data and extracting information from it using traditional applications is difficult, time-consuming, and requires considerable resources. New architectures, technologies, and techniques (described in Module 2, 'Third Platform Technologies') have emerged that enable storing, managing, analyzing, and deriving value from unstructured data coming from numerous sources.

Information Storage

- Information is stored on storage devices on non-volatile media
- Types of storage devices:
  - **Magnetic storage devices**: Hard disk drive and magnetic tape
  - **Optical storage devices**: Blu-ray disc, DVD, and CD
  - **Flash-based storage devices**: Solid state drive, memory card, and USB thumb drive
- Storage devices are assembled within a storage system or "array"
  - Provides high capacity, scalability, performance, reliability, and security
- Storage systems along with other IT infrastructure are housed in a data center

In a computing environment, storage devices (or simply "storage") are devices consisting of non-volatile recording media on which information can be persistently stored. Storage may be internal (for example, internal hard drive), removable (for example, memory cards), or external (for example, magnetic tape drive) to a compute system. Based on the nature of the storage media used, storage devices can be broadly classified as given below:

- **Magnetic storage devices**: For example, hard disk drive and magnetic tape drive.
- **Optical storage devices**: For example, Blu-ray, DVD, and CD.
- **Flash-based storage devices**: For example, solid state drive (SSD), memory card, and USB thumb drive (or pen drive).

Storage is a core component in an organization's IT infrastructure. Various factors such as the media, architecture, capacity, addressing, reliability, and performance influence the choice and use of storage devices in an enterprise environment. For example, disk drives and SSDs are used for storing business-critical information that needs to be continuously accessible to applications; whereas, magnetic tapes and optical storage are typically used for backing up and archiving data. The different types of storage devices are covered in Module 3, 'Data Center Environment'.

In enterprise environments, information is typically stored on storage systems (or storage "arrays"). **A storage system** is a hardware component that contains a group of homogeneous/heterogeneous storage devices assembled within a cabinet. These enterprise-class storage systems are designed for high capacity, scalability, performance, reliability, and security to meet business requirements. The compute systems that run business applications are provided storage capacity from storage systems. Storage systems are covered in Module 4, 'Intelligent Storage Systems (ISS)'. Organizations typically house their IT infrastructure, including compute systems, storage systems, and network equipment within a data center.

What is a Data Center?

Data Center

A facility that houses IT equipment including compute, storage, and network components, and other supporting infrastructure for providing centralized data-processing capabilities.

- A data center comprises:
  - **Facility**: The building and floor space where the data center is constructed
  - **IT equipment**: Compute, storage, and network equipment
  - **Support infrastructure**: Power supply, fire detection, HVAC, and security systems

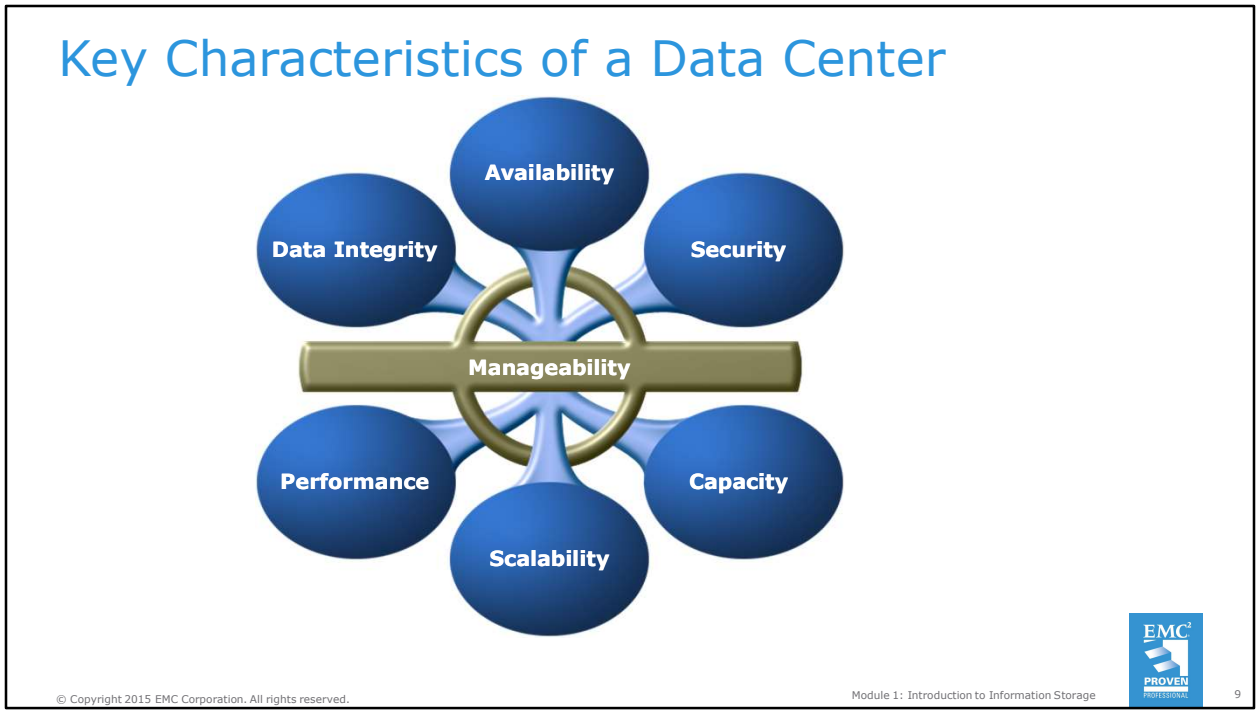
A data center is a dedicated facility where an organization houses, operates, and maintains back-end IT infrastructure including compute systems, storage systems, and network equipment along with other supporting infrastructure. A data center centralizes an organization's IT equipment and data-processing operations, and is vital for carrying out business operations.

A data center typically comprises the following:

- **Facility:** It is the building and floor space where the data center is constructed. It typically has a raised floor with ducts underneath holding power and network cables.
- **IT equipment:** It includes equipment such as compute systems, storage systems, network equipment and cables, and cabinets for housing the IT equipment.
- **Support infrastructure:** It includes all the equipment necessary to securely sustain the functioning of the data center. Some key support equipment are power equipment including uninterruptible power sources, and power generators; environmental control equipment including fire and water detection systems, heating, ventilation, and air conditioning (HVAC) systems; and security systems including biometrics, keycard, and video surveillance systems.

An organization may build a data center to provide open access to applications over the Internet, or for privately executing business applications within its operational environment. A data center may be constructed in-house and located in an organization's own facility, or it may be outsourced, with equipment being located at a third-party site. Large organizations often maintain multiple data centers to distribute data-processing workloads and for disaster recovery.

Organizations are increasingly focusing on energy-efficient technologies and efficient management practices to reduce the energy consumption of data centers and lessen the impact on the environment. Such data centers are called as "green data centers".



Data centers are designed and built to fulfill the key characteristics shown in the figure on the slide. Although the characteristics are applicable to almost all data center components, the discussion here primarily focuses on storage systems.

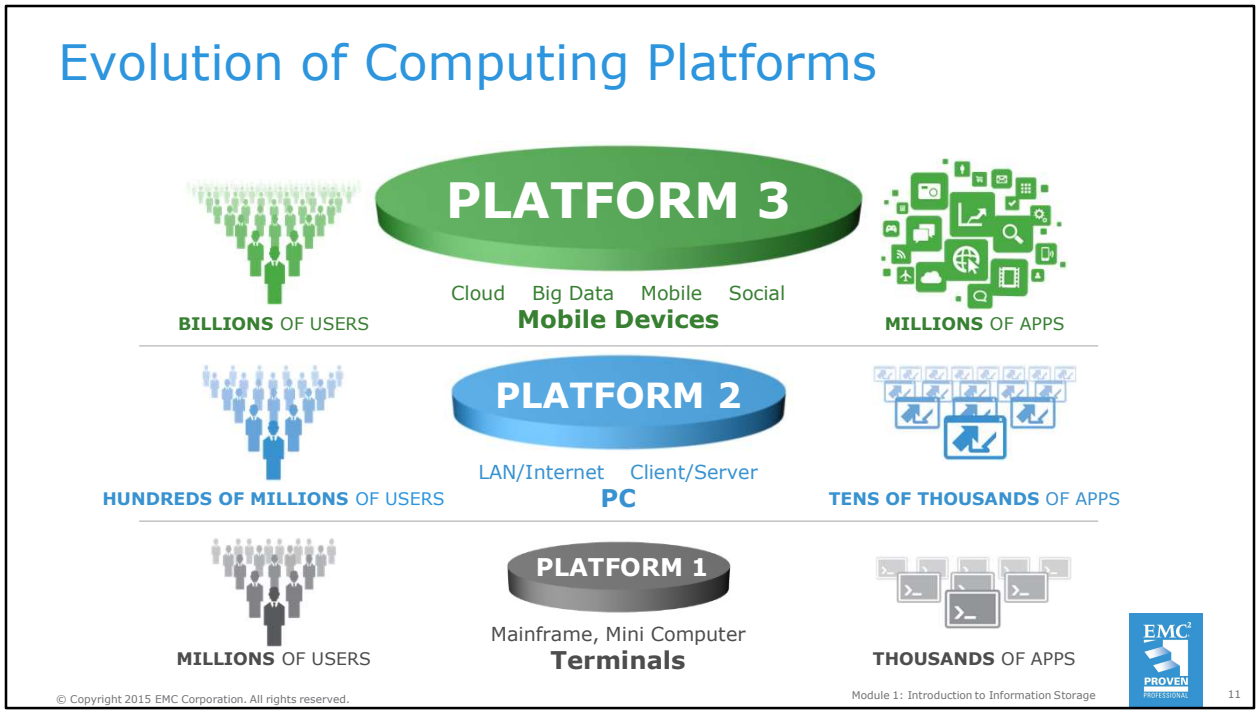
- **Availability:** Availability of information as and when required should be ensured. Unavailability of information can severely affect business operations, lead to substantial financial losses, and damage the reputation of an organization.
- **Security:** Policies and procedures should be established, and control measures should be implemented to **prevent unauthorized access** to and alteration of information.
- **Capacity:** Data center operations require adequate resources to efficiently store and process large and increasing amounts of data. When capacity requirements increase, additional capacity should be provided either without interrupting the availability or with minimal disruption. Capacity may be managed by adding new resources or by reallocating existing resources.
- **Scalability: Organizations** may need to deploy additional resources such as compute systems, new applications, and databases to meet the growing requirements. Data center resources should scale to meet the changing requirements, without interrupting business operations.
- **Performance:** Data center components should provide optimal performance based on the required service levels.
- **Data integrity:** Data integrity refers to mechanisms, such as error correction codes or parity bits, which ensure that data is stored and retrieved exactly as it was received.
- **Manageability:** A data center should provide easy, flexible, and integrated management of all its components. **Efficient manageability** can be achieved through automation for reducing manual intervention in common, repeatable tasks.

Key Data Center Management Processes

Management Process	Description
Monitoring	Continuously gathering information on data center resources
Reporting	Presenting the details on resource performance, capacity, and utilization
Provisioning	Configuring and allocating resources to meet the capacity, availability, performance, and security requirements
Planning	Estimating the amount of resources required to support business operations
Maintenance	Ensuring the proper functioning of resources and resolving incidents

The activities carried out to ensure the efficient functioning of a data center can be broadly categorized under the following key management processes:

- **Monitoring:** It is a continuous process of gathering information on various resources in the data center. The process involves monitoring parameters such as configuration, availability, capacity, performance, and security of resources.
- **Reporting:** It is a process of collating and presenting the monitored parameters such as resource performance, capacity, and utilization of resources. Reporting enables data center managers to analyze and improve the utilization of data center resources and identify problems. It also helps in establishing business justifications and chargeback of costs associated with data center operations.
- **Provisioning:** It is the process of configuring and allocating the resources that are required to carry out business operations. For example, compute systems are provisioned to run applications and storage capacity is provisioned to a compute system. Provisioning primarily includes resource management activities to meet capacity, availability, performance, and security requirements.
- **Planning:** It is a process of estimating the amount of IT resources required to support business operations and meet the changing resource requirements. Planning leverages the data collected during monitoring and enables improving the overall utilization and performance of resources. It also enables estimation of future resource requirements. Data center managers also determine the impact of incidents and devise contingency plans to resolve them.
- **Maintenance**: It is a set of standard repeatable activities for operating the data center. It involves ensuring the proper functioning of resources and resolving incidents such as malfunctions, outages, and equipment loss. It also involves handling identified problems or issues within the data center and incorporating changes to prevent future problem occurrence.

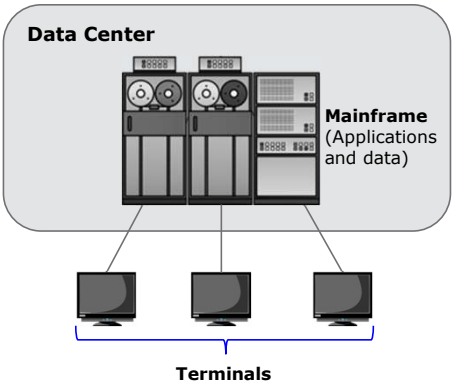


In general, the term "platform" refers to hardware and software that are associated with a particular computing architecture deployed in a data center. Computing platforms evolve and grow with advances and changes in technology. The figure on the slide displays the three computing platforms of IT growth as specified by IDC. The *first platform* (or Platform 1) dates back to the dawn of computing and was primarily based on mainframes and terminals. The *second platform* (or Platform 2) emerged with the birth of the personal computer (PC) in the 1980s and was defined by the client-server model, Ethernet, RDBMSs, and web applications. The *third platform* (or Platform 3) of today comprises cloud, Big Data, mobile, and social technologies.

Each computing platform is defined not so much by the comprising technologies but by the scale of users and the scope of applications the technologies enable. The first platform supported millions of users, with applications and solutions in the low thousands. The second platform supported hundreds of millions of users and tens of thousands of applications. The third platform is already supporting a user base of billions and has millions of applications and solutions. This is evident from the fact that over 2.4 billion people (~36 percent of the world's population) are currently connected to the Internet (more than half of them through mobile devices), and that there are over one million applications available for iOS and Android devices alone.

First Platform

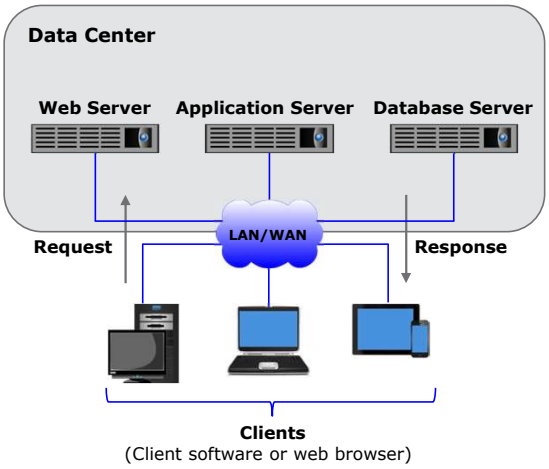
- Based on mainframes
  - Applications and databases hosted centrally
  - Users connect to mainframes through terminals
- Challenges with mainframes
  - Substantial CAPEX and OPEX
    - High acquisition costs
    - Considerable floor space and energy requirements



Mainframes are compute systems with very large processing power, memory, and storage capacity and are primarily used for centrally hosting mission-critical applications and databases in an organization's data center. Multiple users simultaneously connect to mainframes through less-powerful devices, such as workstations or terminals. **All processing is performed on the mainframe, while the terminals only provide an interface to use the applications and view results**. Although mainframes offer **high reliability** and **security**, there are several cost concerns associated with them. Mainframes have high acquisition costs, and considerable floor space and energy requirements. Deploying mainframes in a data center may involve substantial capital expense (CAPEX) and operating expense (OPEX). Historically, large organizations such as banks, insurance agencies, and government departments have used mainframes to run their business operations.

Second Platform

- Based on client-server model
  - Distributed application architecture
  - Servers receive and process requests for resources from clients
  - Users connect through a client program or a web interface
- Challenges with client-server model
  - Creation of IT silos
  - Hardware and software maintenance overhead
  - Scalability to meet the growth of users and workloads



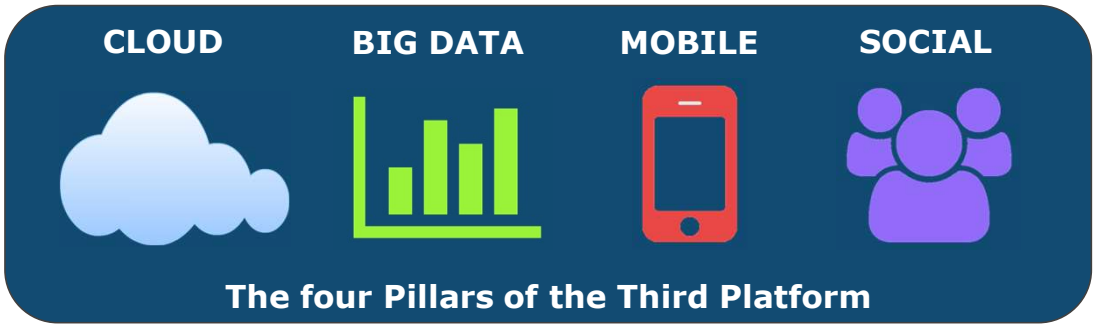
The *client-server model* uses a distributed application architecture, in which a compute system called "server" runs a program that provides services over a network to other programs running on various end-point devices called "clients". Server programs receive requests for resources from client programs and in response to the requests, the clients receive access to resources, such as e-mail applications, business applications, web applications, databases, files, and printers. Client devices can be desktops, laptops, and mobile devices. Clients typically communicate with servers over a LAN or WAN, with users making use of either a client application or a web interface on a browser.

In the client-server model, both the clients and the servers may have distinct processing tasks that they routinely perform. For example, a **client** may **run** the business application while the server may run the database management system (DBMS) to manage storage and retrieval of information to and from a database. This is called a *two-tier architecture*. **Alternatively**, a client may use an application or web interface to accept information while the server runs another application that processes the information and sends the data to a second server that runs the DBMS. This is called the ***three-tier architecture***. This distributed application architecture can be extended to any number of tiers (*n-tier architecture*). Because both client and server systems are intelligent devices, the client-server model is completely different from the mainframe model.

The figure on the slide shows an example of the client-server model. In the example, clients interact with the web server using a web browser. The web server processes client requests through HTTP and delivers HTML pages. The application server hosts a business application and the database server hosts a DBMS. The clients interact with the application server through client software. The application server communicates with the database server to retrieve information and provide results to the clients. **In some implementations, applications and databases may even be hosted on the same server**.

(Cont'd)

Third Platform



- The four pillars are transforming the way organizations are using technology for business operations

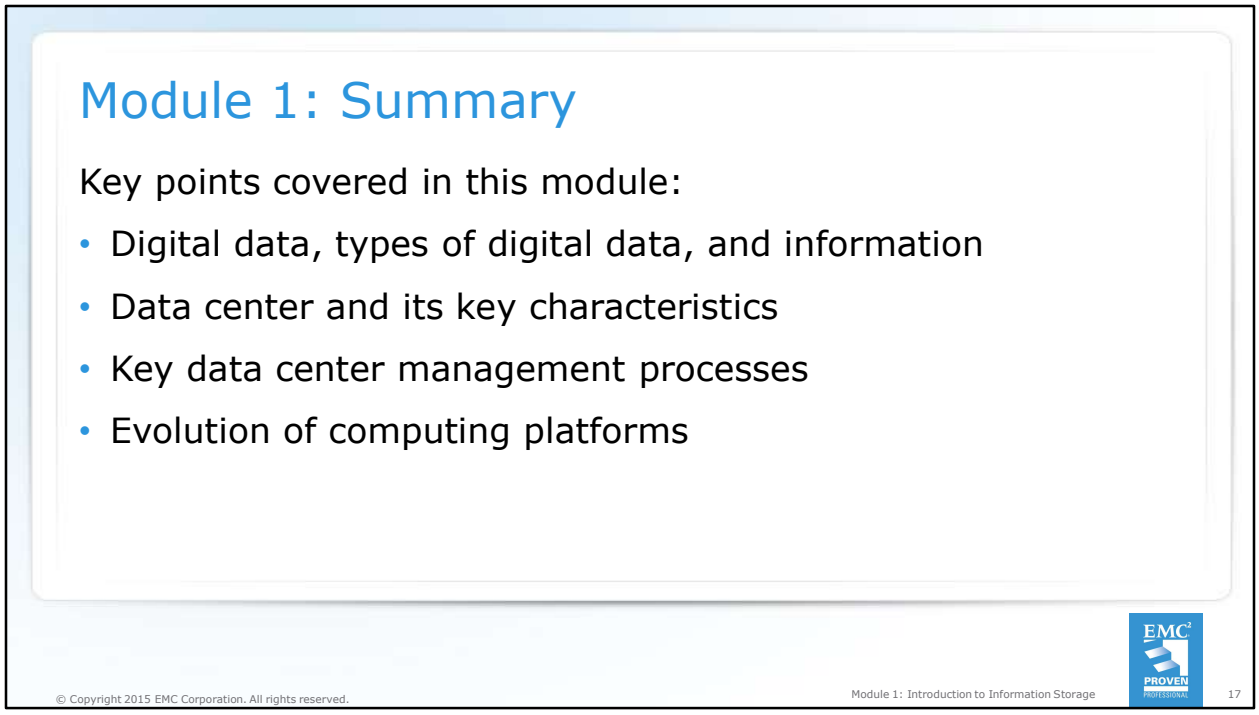
The term "third platform" was coined by IDC, and Gartner refers to the same as a "nexus of forces". The third platform is built on a foundation of cloud, Big Data, mobile, and social technologies. These are the four major "disruptive" technologies that are significantly transforming businesses, economies, and lives globally.

At its core, the third platform has the cloud that enables a consumer to provision IT resources as a service from a cloud provider. Big Data enables analytics that create deeper insights from data for improved decision-making. Mobile devices enable pervasive access to applications and information. Social technologies connect individuals, and enable collaboration and information exchange.

Over the past three decades, it was essential for organizations to intelligently leverage the second platform for their businesses. According to IDC, over the next three decades, the third platform will represent the basis for solution development and business innovation. The third platform is being used for the digital transformation, evolution, and expansion of all industries and for developing major new sources of competitive advantage. Business strategists, IT leaders, and solution developers are already building disruptive new business models and consumer services around third platform technologies.

Third platform technologies are an enhancement of second platform technologies rather than a substitution. A key aspect of third platform is that it is a convergence of cloud, Big Data, mobile, and social technologies and not just each technology taken in isolation. The real key is combining two or more of the technologies to create high-value industry solutions known as "*mashups*". For example, some of the top drivers of cloud include social and mobile solutions. This means that organizations already see the greatest value in solutions that are mashups across all four technologies. The combinations of third platform technologies are already transforming organizations such as retail, financial services, government departments, telecommunications, and healthcare.

(Cont'd)



This module covered digital data, the types of digital data, and information. This module also covered data center and its key characteristics. Further, this module covered the key data center management processes. Finally, this module covered the evolution of computing platforms.