고려대학교
KOREA UNIVERSITY

KU-The Future

Cloud computing – Chapter 5

Heonchang Yu

Distributed and Cloud Computing Lab.

Broadband Networks and Internet Architecture

- Internetworks, or the Internet, allow for the remote provisioning of IT resources and are directly supportive of ubiquitous network access.
- Internet Service Providers (ISPs)
  - ✓ ISP network interconnects to other ISP networks and various organizations (**Figure 5.1**).
  - ✓ The concept of the Internet was based on a decentralized provisioning and management model.
  - ✓ Worldwide connectivity is enabled through a hierarchical topology composed of Tiers 1, 2, and 3 (**Figure 5.2**).
    - ❖ The core Tier 1 is made of large-scale, international cloud providers that oversee massive interconnected global networks, which are connected to Tier 2's large regional providers.
    - ❖ The interconnected ISPs of Tier 2 connect with Tier 1 providers, as well as the local ISPs of Tier 3.
  - ✓ Two fundamental components used to construct the internetworking architecture are *connectionless packet switching* (datagram networks) and *router-based interconnectivity*.

Broadband Networks and Internet Architecture

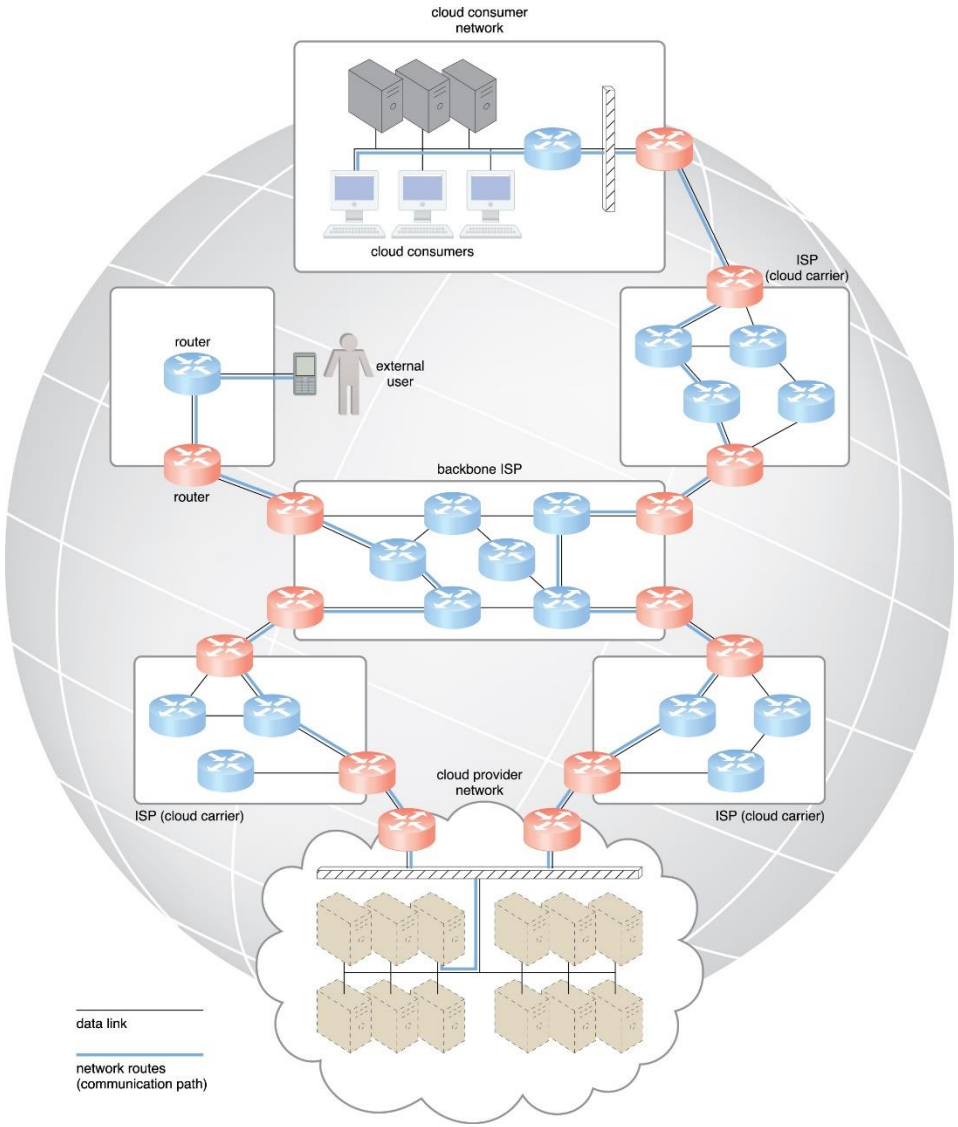

Figure 5.1 Messages travel over dynamic network routes in this ISP internetworking configuration

Broadband Networks and Internet Architecture

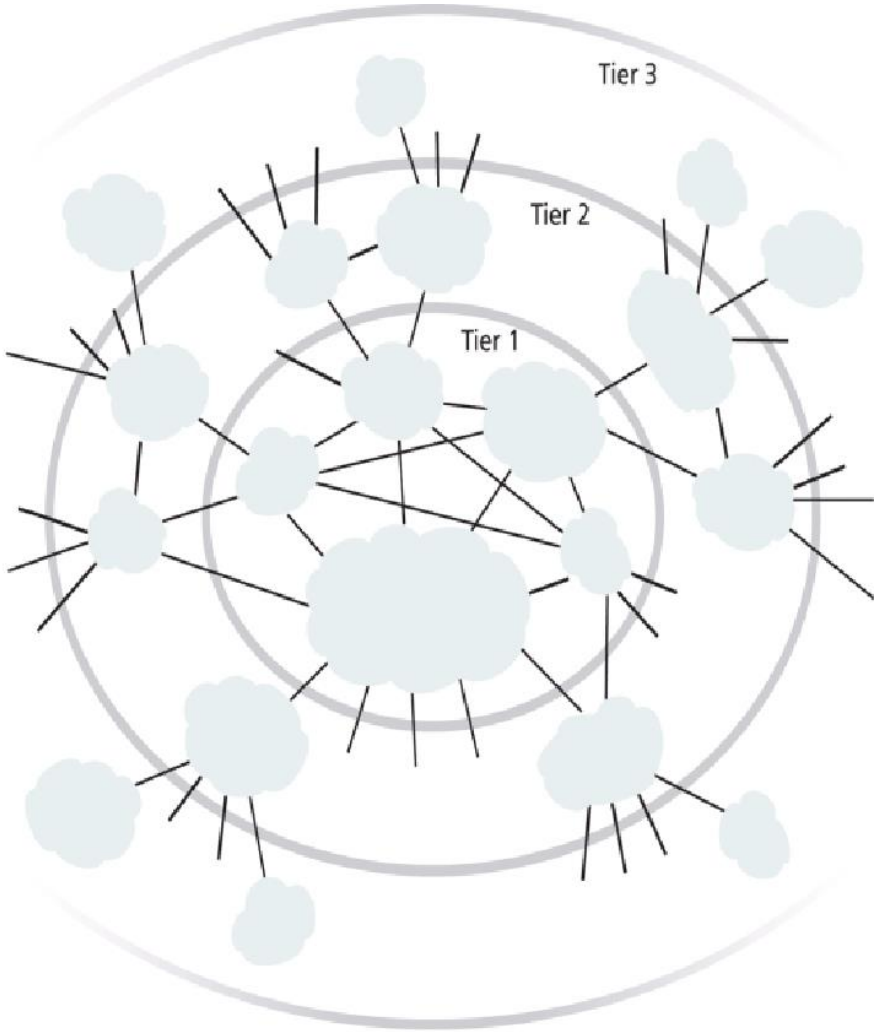

Figure 5.2 An abstraction of the internetworking structure of the Internet

Broadband Networks and Internet Architecture

- Connectionless Packet Switching (Datagram Networks)
  - ✓ End-to-end (sender-receiver pair) data flows are divided into packets of a limited size that are received and processed through network switches and routers, then queued and forwarded from one intermediary node to the next.
- Router-Based Interconnectivity
  - ✓ Even when successive packets are part of the same data flow, routers process and forward each packet individually while maintaining the network topology information that locates the next node on the communication path between the source and destination nodes.
- Communication can therefore be sustained even during simultaneous network failures, although using multiple network paths can cause routing fluctuations and latency.
- The basic mechanics of internetworking are illustrated in **Figure 5.3**, in which a message is coalesced from an incoming group of disordered packets.

Broadband Networks and Internet Architecture

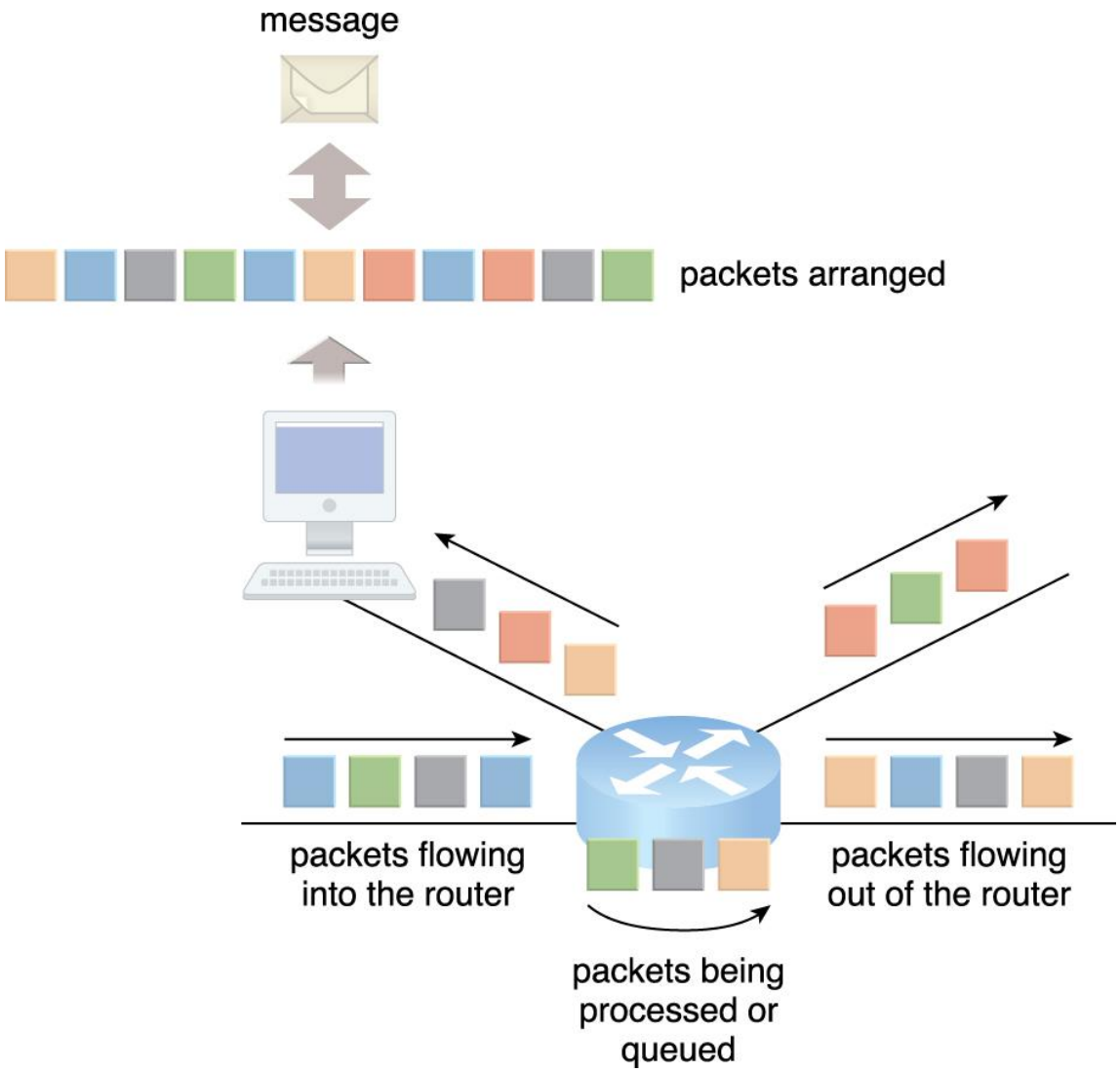

Figure 5.3 Packets traveling through the Internet are directed by a router that arranges them into a message

Broadband Networks and Internet Architecture

- Physical Network
  - ✓ IP packets are transmitted through underlying physical networks that connect adjacent nodes, such as Ethernet, ATM network, and the 3G mobile HSDPA.
    - ❖ A data link layer that controls data transfer between neighboring nodes
    - ❖ A physical layer that transmits data bits through both wired and wireless media.
- Transport Layer Protocol
  - ✓ Transmission Control Protocol (TCP) and User Datagram Protocol (UDP),
  - ✓ Use the IP to provide standardized, end-to-end communication support that facilitates the navigation of data packets across the Internet.
- Application Layer Protocol
  - ✓ HTTP, SMTP for e-mail, BitTorrent for P2P, and SIP for IP telephony

Broadband Networks and Internet Architecture

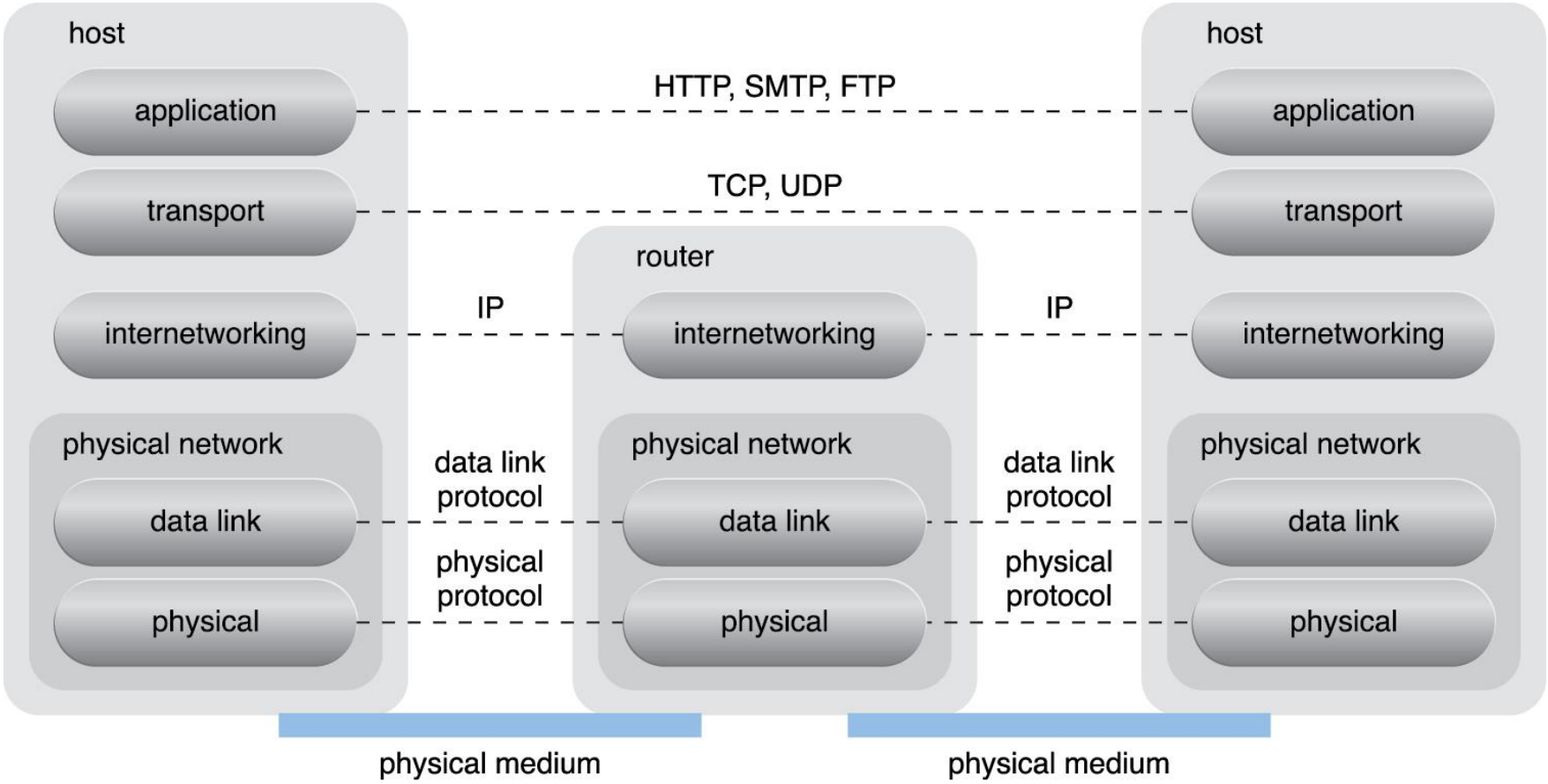

Figure 5.4 A generic view of the Internet reference model and protocol stack

Broadband Networks and Internet Architecture

- Technical and Business Considerations - **Connectivity Issues**
  - ✓ End-user devices, such as smartphones and laptops, access the data center through the corporate network, which provides uninterrupted Internet connectivity.
  - ✓ TCP/IP facilitates both Internet access and on-premise data exchange over LANs (**Figure 5.5**).
  - ✓ End-user devices that are connected to the network through the Internet can be granted continuous access to centralized servers and applications in the cloud (**Figure 5.6**).
  - ✓ Whether IT resources are on-premise or Internet-based dictates how internal versus external end-users access services, even if the end-users themselves are not concerned with the physical location of cloud-based IT resources (Table 5.1).

Broadband Networks and Internet Architecture

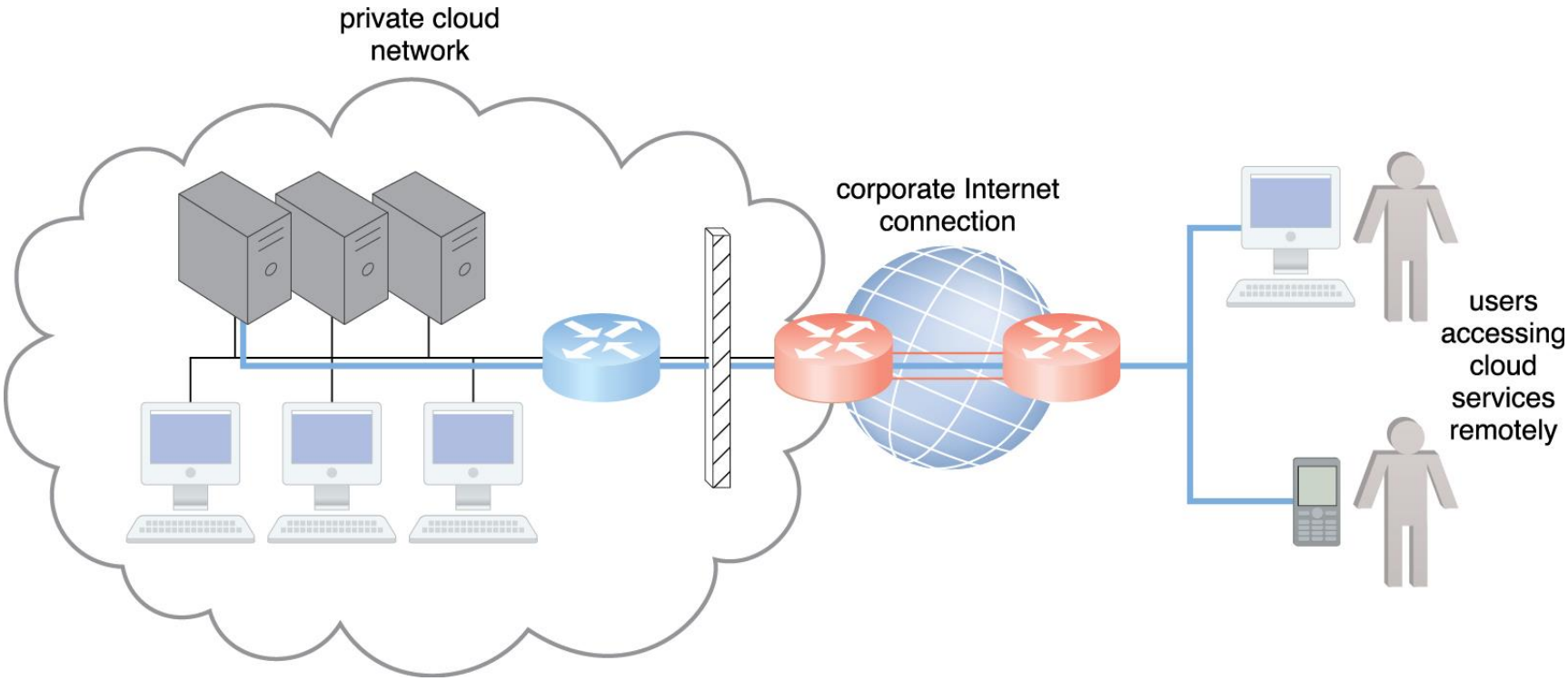

Figure 5.5 The internetworking architecture of a private cloud. The physical IT resources that constitute the cloud are located and managed within the organization

Broadband Networks and Internet Architecture

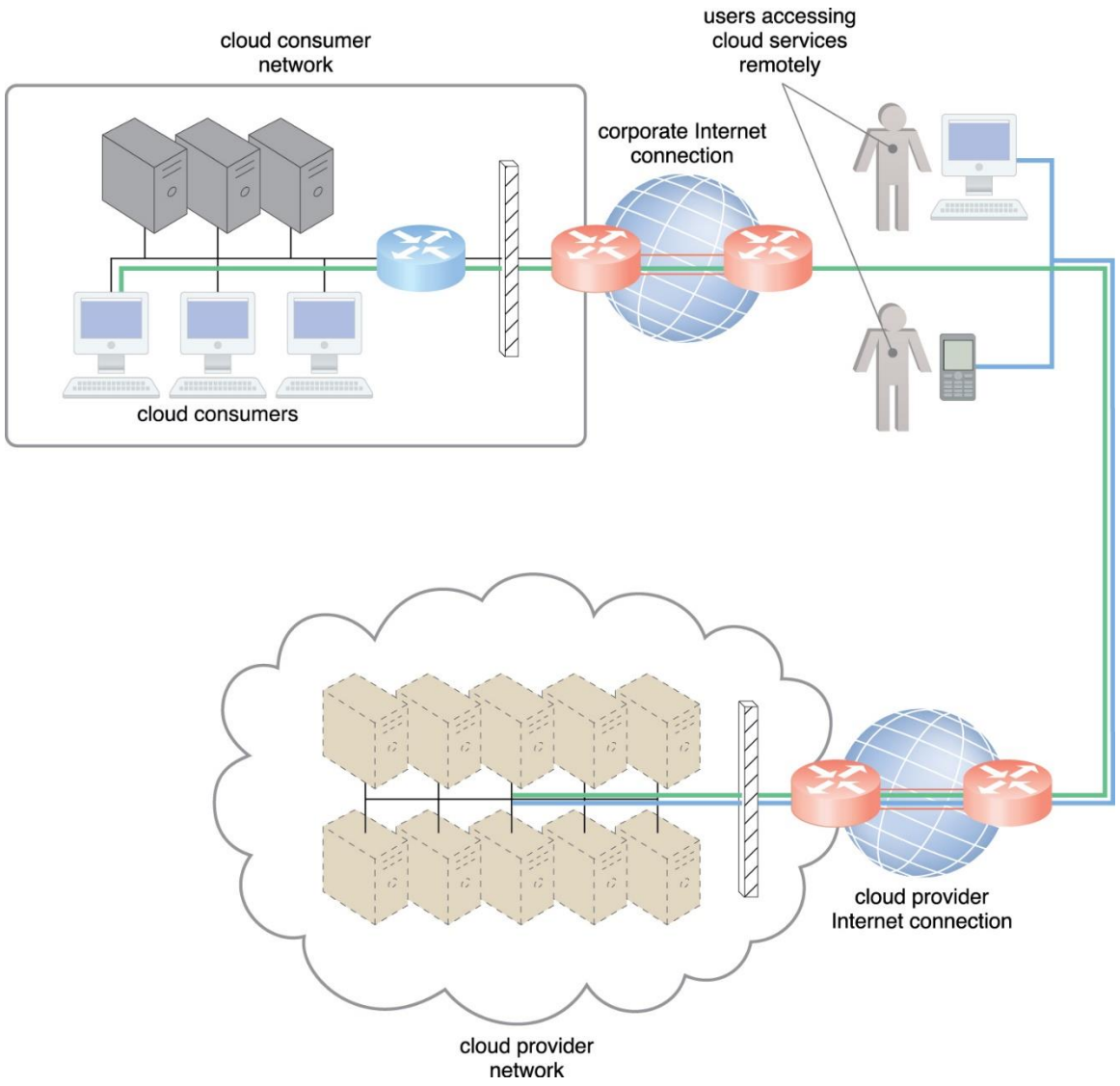

Figure 5.6 The internetworking architecture of an Internet-based cloud deployment model. The Internet is the connecting agent between non-proximate cloud consumers, roaming end-users, and the cloud provider's own network

Broadband Networks and Internet Architecture

Table 5.1 A comparison of on-premise and cloud-based internetworking

On-Premise IT Resources	Cloud-Based IT Resources
internal end-user devices access corporate IT services through the corporate network	internal end-user devices access corporate IT services through an Internet connection
internal users access corporate IT services through the corporate Internet connection while roaming in external networks	internal users access corporate IT services while roaming in external networks through the cloud provider's Internet connection
external users access corporate IT services through the corporate Internet connection	external users access corporate IT services through the cloud provider's Internet connection

Broadband Networks and Internet Architecture

- Technical and Business Considerations - **Network Bandwidth and Latency Issues**
  - ✓ In addition to being affected by the bandwidth of the data link that connects networks to ISPs, end-to-end **bandwidth** is determined by the transmission capacity of the shared data links that connect intermediary nodes.
  - ✓ Referred to as time delay, **latency** is the amount of time it takes a packet to travel from one data node to another.
    - ❖ Latency increases with every intermediary node on the data packet's path.
    - ❖ Networks are dependent on traffic conditions in shared nodes, making Internet latency highly variable and often unpredictable.

Broadband Networks and Internet Architecture

- ✓ IT solutions need to be assessed against business requirements that are affected by network bandwidth and latency, which are inherent to cloud interconnection.
  - ❖ Bandwidth is critical for applications that require substantial amounts of data to be transferred to and from the cloud
  - ❖ Latency is critical for applications with a business requirement of swift response times.

- Technical and Business Considerations - **Cloud Carrier and Cloud Provider Selection**
  - ✓ The service levels of Internet connections between cloud consumers and cloud providers are determined by their ISPs, which are usually different and therefore include multiple ISP networks in their paths.
  - ✓ Cloud consumers and cloud providers may need to use multiple cloud carriers in order to achieve the necessary level of connectivity and reliability for their cloud applications, resulting in additional costs.

Data Center Technology

- Modern data centers exist as specialized IT infrastructure used to house centralized IT resources, such as servers, databases, networking and telecommunication devices, and software systems.

- Virtualization
  - ✓ Data centers consist of both physical and virtualized IT resources. (**Figure 5.7**)
    - ❖ **The physical IT resource layer** refers to the facility infrastructure that houses computing/networking systems and equipment, together with hardware systems and their operating systems.
    - ❖ **The resource abstraction and control of the virtualization layer** is comprised of operational and management tools that are often based on virtualization platforms that abstract the physical computing and networking IT resources as virtualized components that are easier to allocate, operate, release, monitor, and control.

Data Center Technology

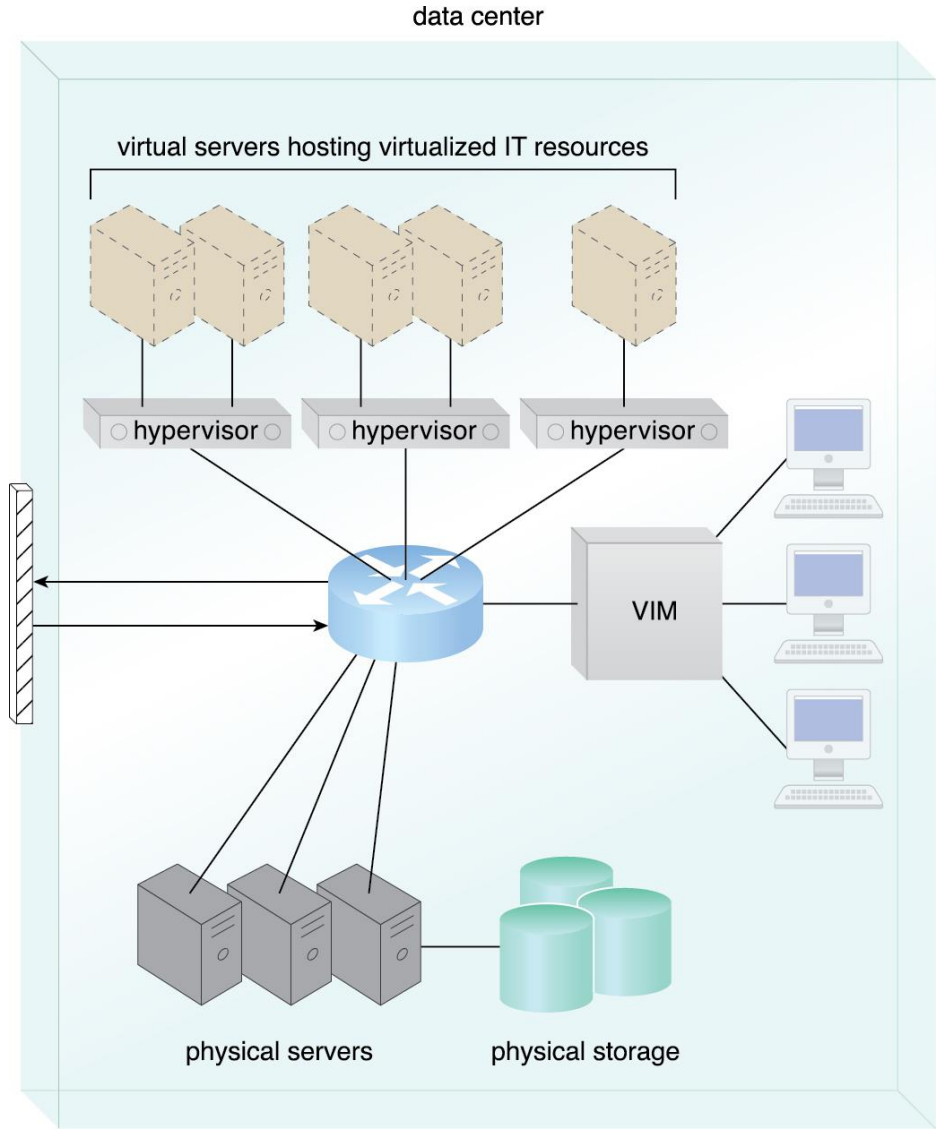

Figure 5.7 The common components of a data center working together to provide virtualized IT resources supported by physical IT resources

Data Center Technology

- Standardization and Modularity
  - ✓ Key requirements for reducing investment and operational costs as they enable economies of scale for the procurement, acquisition, deployment, operation, and maintenance processes
- Automation
  - ✓ Data centers have specialized platforms that automate tasks like provisioning, configuration, patching, and monitoring without supervision.
- Remote Operation and Management
  - ✓ Most of the operational and administrative tasks of IT resources in data centers are commanded through the network's remote consoles and management systems.

Data Center Technology

- High Availability
  - ✓ Since any form of data center outage significantly impacts business continuity for the organizations that use their services, data centers are designed to operate with increasingly higher levels of redundancy to sustain availability.
- Security-Aware Design, Operation, and Management
  - ✓ Requirements for security, such as physical and logical access controls and data recovery strategies, need to be thorough and comprehensive for data centers, since they are centralized structures that store and process business data.
- Facilities
  - ✓ Data center facilities are custom-designed locations that are outfitted with specialized computing, storage, and network equipment.

Data Center Technology

- Computing Hardware
  - ✓ Much of the heavy processing in data centers is often executed by standardized commodity servers that have substantial computing power and storage capacity.
  - ✓ Several computing hardware technologies are integrated into these modular servers, such as:
    - ❖ rackmount form factor server design composed of standardized racks
    - ❖ support for different hardware processing architectures
    - ❖ a power-efficient multi-core CPU architecture
    - ❖ redundant and hot-swappable components
  - ✓ Computing architectures such as blade server technologies use rack-embedded physical interconnections (blade enclosures), fabrics (switches), and shared power supply units and cooling fans.

Data Center Technology

- Storage Hardware
  - ✓ Data centers have specialized storage systems that maintain enormous amounts of digital information in order to fulfill considerable storage capacity needs.
  - ✓ Storage systems usually involve the following technologies:
    - ❖ Hard Disk Arrays - redundant arrays of independent disks (RAID) schemes
    - ❖ I/O Caching
    - ❖ Hot-Swappable Hard Disks
    - ❖ Storage Virtualization
    - ❖ Fast Data Replication Mechanisms
  - ✓ Networked storage devices
    - ❖ Storage Area Network (SAN)
    - ❖ Network-Attached Storage (NAS)

Data Center Technology

- Network Hardware
  - ✓ Data centers require extensive network hardware in order to enable multiple levels of connectivity.
    - ❖ Carrier and External Networks Interconnection
    - ❖ Web-Tier Load Balancing and Acceleration
    - ❖ LAN Fabric
    - ❖ SAN Fabric
    - ❖ NAS Gateways

Virtualization Technology

- Virtualization is the process of converting a physical IT resource into a virtual IT resource.
- Most types of IT resources can be virtualized
  - ✓Virtual Servers
  - ✓Virtual Storage Device or Virtual Disk
  - ✓VLAN
  - ✓Power - Virtual UPSs
- Virtualization software runs on a physical server called a host or physical host, whose underlying hardware is made accessible by the **virtualization software**.
  - ✓ Virtual machine manager or a virtual machine monitor (VMM)
  - ✓ hypervisor

Virtualization Technology

- Hardware Independence
  - ✓ The installation of an operating system's configuration and application software in a unique IT hardware platform results in many software-hardware dependencies.
  - ✓ Virtualization is a conversion process that translates unique IT hardware into emulated and standardized software-based copies.
  - ✓ Through hardware independence, virtual servers can easily be moved to another virtualization host.
- Server Consolidation
  - ✓ The coordination function that is provided by the virtualization software allows multiple virtual servers to be simultaneously created in the same virtualization host.
  - ✓ Virtualization technology enables different virtual servers to share one physical server.
  - ✓ Is used to increase hardware utilization, load balancing, and optimization of available IT resources.

Virtualization Technology

- Resource Replication
  - ✓ Virtual servers are created as virtual disk images that contain binary file copies of hard disk content. These virtual disk images are accessible to the host's operating system and paste, can be used to replicate, migrate, and back up the virtual server.
    - ❖ The creation of standardized virtual machine images commonly configured to include virtual hardware capabilities, guest operating systems, and additional application software
    - ❖ Increased agility in the migration and deployment of a virtual machine's new instances
    - ❖ The ability to roll back, which is the instantaneous creation of VM snapshots by saving the state of the virtual server's memory and hard disk image to a host-based file
    - ❖ The support of business continuity with efficient backup and restoration procedures, as well as the creation of multiple instances of critical IT resources and applications

Virtualization Technology

- Operating System-Based Virtualization
  - ✓ The installation of virtualization software in a pre-existing operating system, which is called the host operating system (**Figure 5.8**).
  - ✓ Hardware independence that is enabled by virtualization allows hardware IT resources to be more flexibly used.
  - ✓ Virtualization software translates hardware IT resources that require unique software for operation into virtualized IT resources that are compatible with a range of operating systems.
  - ✓ Operating system-based virtualization can introduce **demands and issues related to performance overhead** such as:
    - ❖ The host operating system consumes CPU, memory, and other hardware IT resources.
    - ❖ Hardware-related calls from guest operating systems need to traverse several layers to and from the hardware.
    - ❖ Licenses are usually required for host OSs, in addition to individual licenses for each of their guest OSs.

Virtualization Technology

- ✓ Implementing a virtualization layer will negatively affect overall system performance.

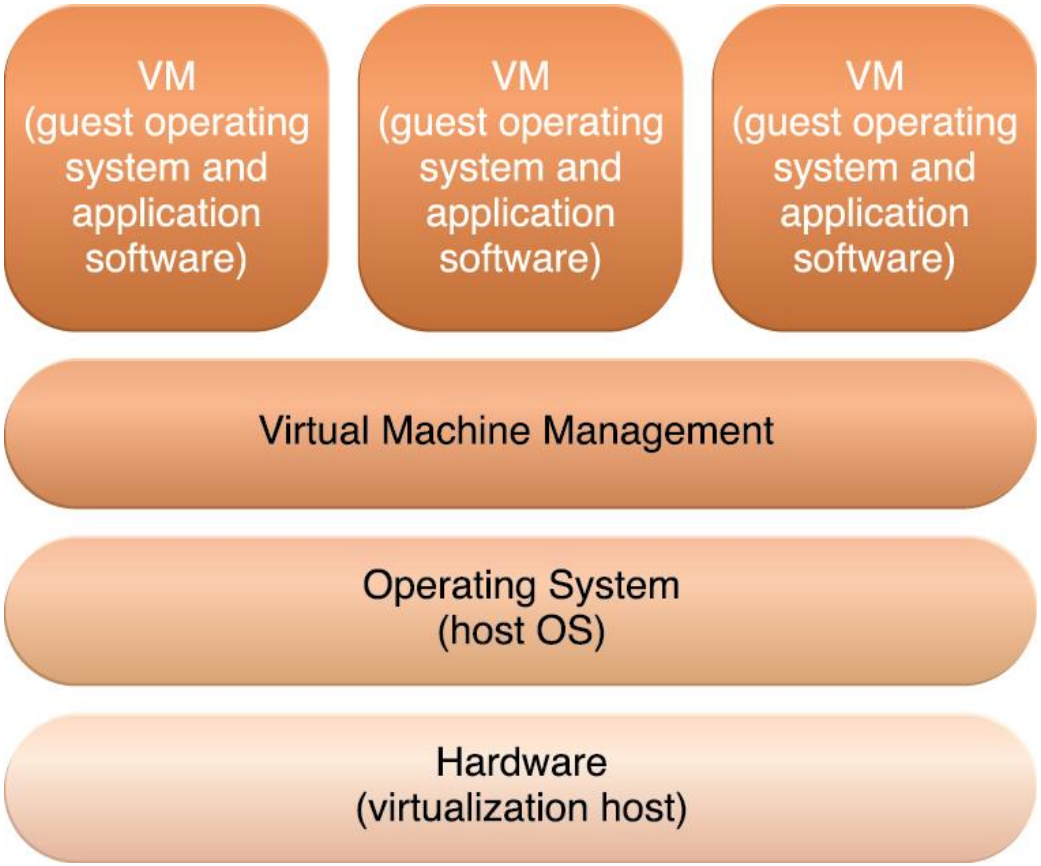

Figure 5.8 The different logical layers of operating system-based virtualization, in which the VM is first installed into a full host operating system and subsequently used to generate virtual machines

Virtualization Technology

- Hardware-Based Virtualization
  - ✓ This option represents the installation of virtualization software directly on the physical host hardware so as to bypass the host operating system (**Figure 5.9**).
  - ✓ Virtualization software is typically referred to as a **hypervisor**.
  - ✓ A hypervisor exists as a thin layer of software that handles hardware management functions to establish a virtualization management layer
  - ✓ One of the main issues of hardware-based virtualization concerns compatibility with hardware devices.
    - ❖ The virtualization layer is designed to communicate directly with the host hardware.
    - ❖ Hardware device drivers may not be as available to hypervisor platforms as they are to operating systems.

Virtualization Technology

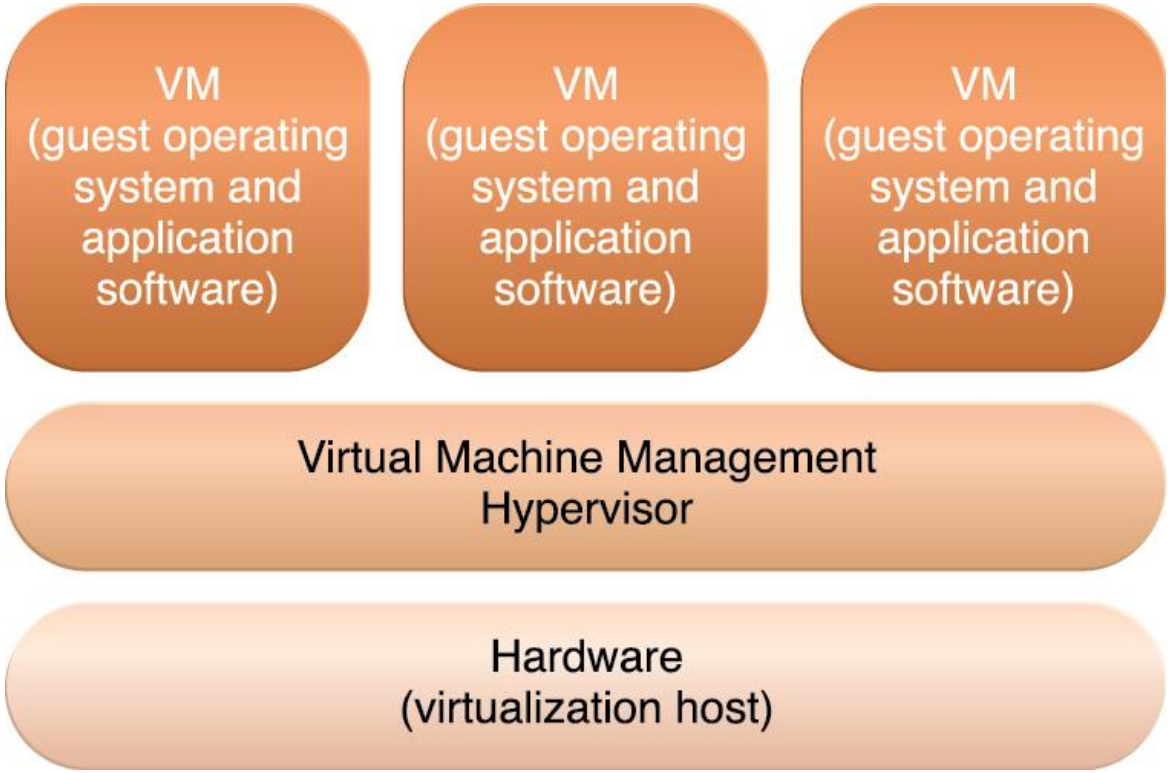

Figure 5.9 The different logical layers of hardware-based virtualization, which does not require another host operating system

Virtualization Technology

- Virtualization Management
  - ✓ Modern virtualization software provides several advanced management functions that can automate administration tasks and reduce the overall operational burden on virtualized IT resources.
- Other Considerations
  - ✓ *Performance Overhead* – Virtualization may not be ideal for complex systems that have high workloads with little use for resource sharing and replication.
  - ✓ *Special Hardware Compatibility* – Many hardware vendors that distribute specialized hardware may not have device driver versions that are compatible with virtualization software.
  - ✓ *Portability* – The programmatic and management interfaces that establish administration environments for a virtualization program to operate with various virtualization solutions can introduce portability gaps due to incompatibilities.

Web Technology

- Basic Web Technology
  - ✓ The two basic components of the Web are the Web browser client and the Web server.
  - ✓ Three fundamental elements comprise the technology architecture of the Web:
    - ❖ *Uniform Resource Locator (URL)* – A standard syntax used for creating identifiers that point to Web-based resources
    - ❖ *Hypertext Transfer Protocol (HTTP)* – This is the primary communications protocol exchange content and data throughout the World Wide Web.
    - ❖ *Markup Languages (HTML, XML)* – Markup languages provide a lightweight means of expressing Web-centric data and metadata.

Web Technology

- Web Applications
  - ✓ **Figure 5.10** presents a common architectural abstraction for Web applications that is based on the basic three-tier model.
    - ❖ The first tier is called the *presentation layer*, which represents the user interface.
    - ❖ The middle tier is the *application layer* that implements application logic.
    - ❖ The third tier is the *data layer* that is comprised of persistent data stores.

Web Technology

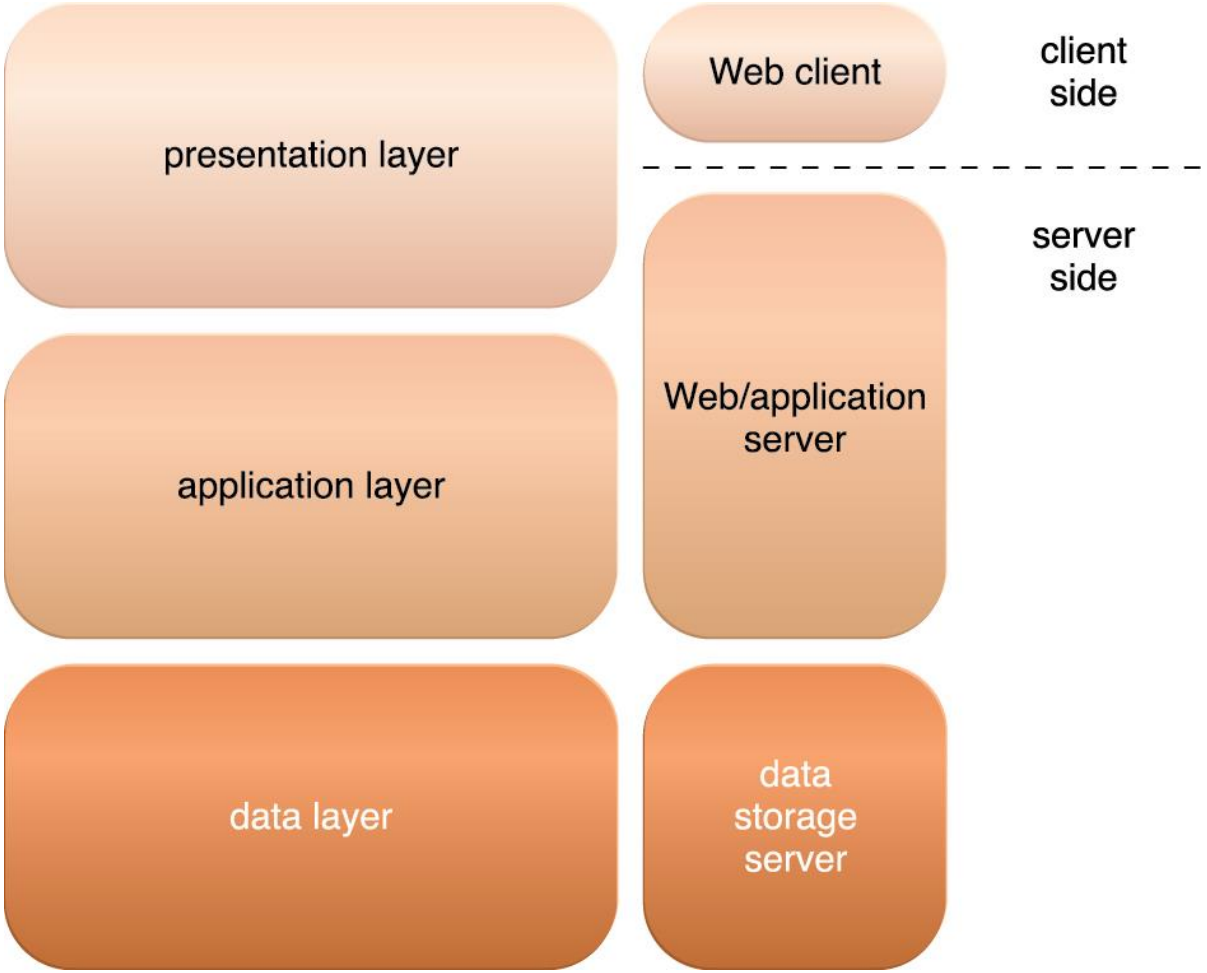

Figure 5.10 The three basic architectural tiers of Web applications

Multitenant Technology

- The multitenant application design was created to enable multiple users (tenants) to access the same application logic simultaneously.
- Tenants can individually customize features of the application, such as:
  - ✓ User Interface – Tenants can define a specialized "look and feel" for their application interface.
  - ✓ Business Process – Tenants can customize the rules, logic, and workflows of the business processes that are implemented in the application.
  - ✓ Data Model – Tenants can extend the data schema of the application to include, exclude, or rename fields in the application data structures.
  - ✓ Access Control – Tenants can independently control the access rights for users and groups.

Multitenant Technology

- Common characteristics of multitenant applications include:
  - ✓ Usage Isolation – The usage behavior of one tenant does not affect the application availability and performance of other tenants.
  - ✓ Data Security – Tenants cannot access data that belongs to other tenants.
  - ✓ Recovery – Backup and restore procedures are separately executed for the data of each tenant.
  - ✓ Application Upgrades – Tenants are not negatively affected by the synchronous upgrading of shared software artifacts.
  - ✓ Scalability – The application can scale to accommodate increases in usage by existing tenants and/or increases in the number of tenants.
  - ✓ Metered Usage – Tenants are charged only for the application processing and features that are actually consumed.
  - ✓ Data Tier Isolation – Tenants can have individual databases, tables, and/or schemas isolated from other tenants.

Multitenant Technology

- A multitenant application that is being concurrently used by two different tenants is illustrated in **Figure 5.11**.

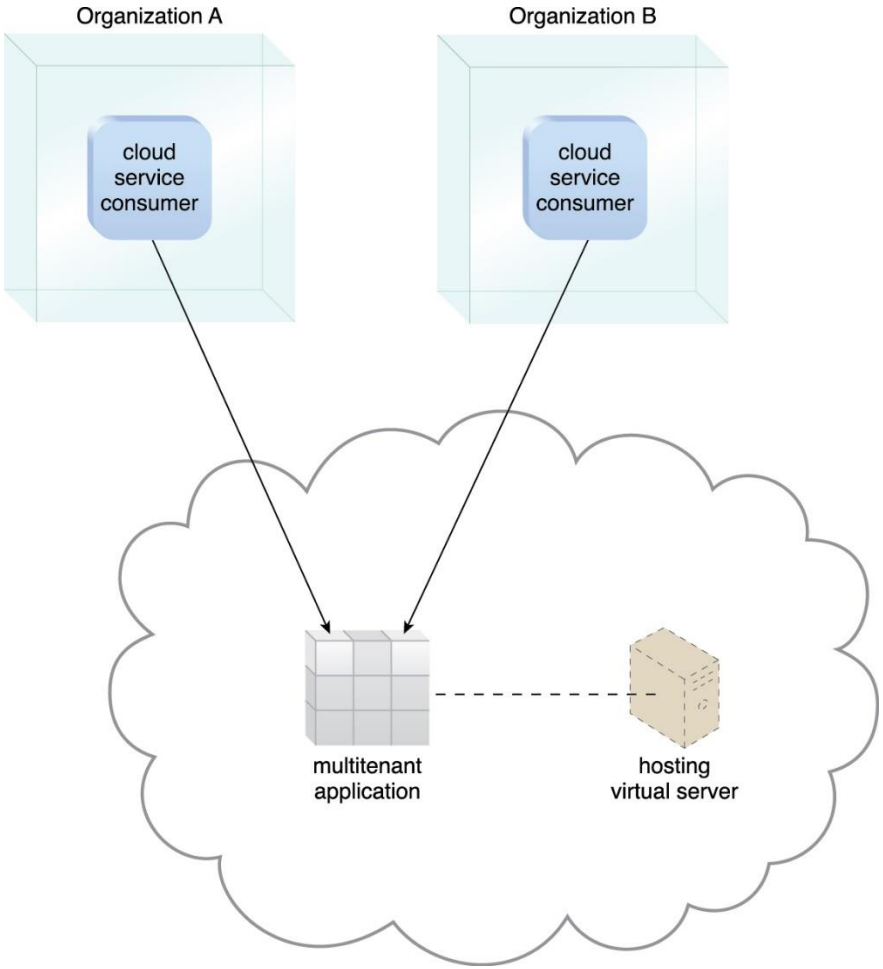

Figure 5.11 A multitenant application that is serving multiple cloud service consumers simultaneously

Multitenant Technology

- **Multitenancy** vs. **Virtualization**
  - ✓ With **virtualization**: Multiple virtual copies of the server environment can be hosted by a single physical server. Each copy can be provided to different users, can be configured independently, and can contain its own operating systems and applications.
  - ✓ With **multitenancy**: A physical or virtual server hosting an application is designed to allow usage by multiple different users. Each user feels as though they have exclusive usage of the application.

Service Technology

- Web Services
  - ✓ Commonly prefixed with "SOAP-based,"
  - ✓ The core technologies behind Web services are represented by the following industry standards:
    - ❖ Web Service Description Language (WSDL)
    - ❖ XML Schema Definition Language (XML Schema)
    - ❖ SOAP – Formerly known as the Simple Object Access Protocol
    - ❖ Universal Description, Discovery, and Integration (UDDI)
  - ✓ These four technologies collectively form the first generation of Web service technologies (**Figure 5.12**).

Service Technology

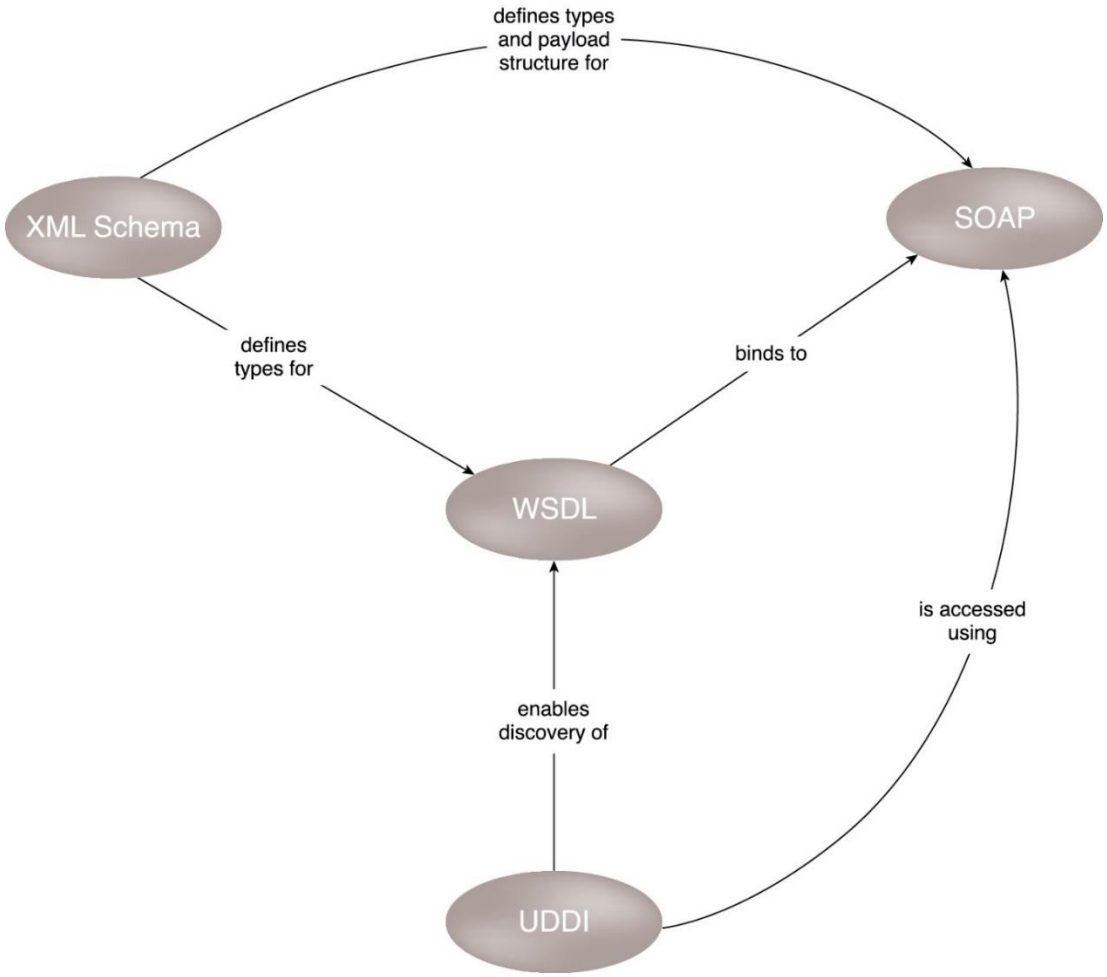

Figure 5.12 An overview of how first-generation Web service technologies commonly relate to each other

Service Technology

- Rest Services
  - ✓ REST services are designed according to a set of constraints that shape the service architecture to emulate the properties of the World Wide Web.
  - ✓ The six REST design constraints are:
    - ❖ Client-Server
    - ❖ Stateless
    - ❖ Cache
    - ❖ Interface/Uniform Contract
    - ❖ Layered System
    - ❖ Code-On-Demand
  - ✓ Each design constraint is described in detail at www.whatisrest.com.

Service Technology

- Service Agents
  - ✓ Event-driven programs designed to intercept messages at runtime.
  - ✓ **Active service agents** perform an action upon intercepting and reading the contents of a message.
    - ❖ Requires making changes to the message contents or changes to the message path itself
  - ✓ **Passive service agents** do not change message contents.
    - ❖ They read the message and may then capture certain parts of its contents, usually for monitoring, logging, or reporting purposes.
  - ✓ Cloud-based environments rely heavily on the use of system-level and custom service agents to perform much of the runtime monitoring and measuring required to ensure that features, such as elastic scaling and pay-for-use billing, can be carried out instantaneously.

Case Study Example

- DTGOV has assembled cloud-aware infrastructures in each of its data centers.
- **Figure 5.13** illustrates the physical layout of the server network connections inside of the data center.
- A separate network that connects the storage systems and servers is installed with clustered storage area network (SAN) switches and similar redundant connections to various devices (**Figure 5.14**).
- **Figure 5.15** illustrates an internetworking architecture that is established between every data center pair within the DTGOV corporate infrastructure.

Case Study Example

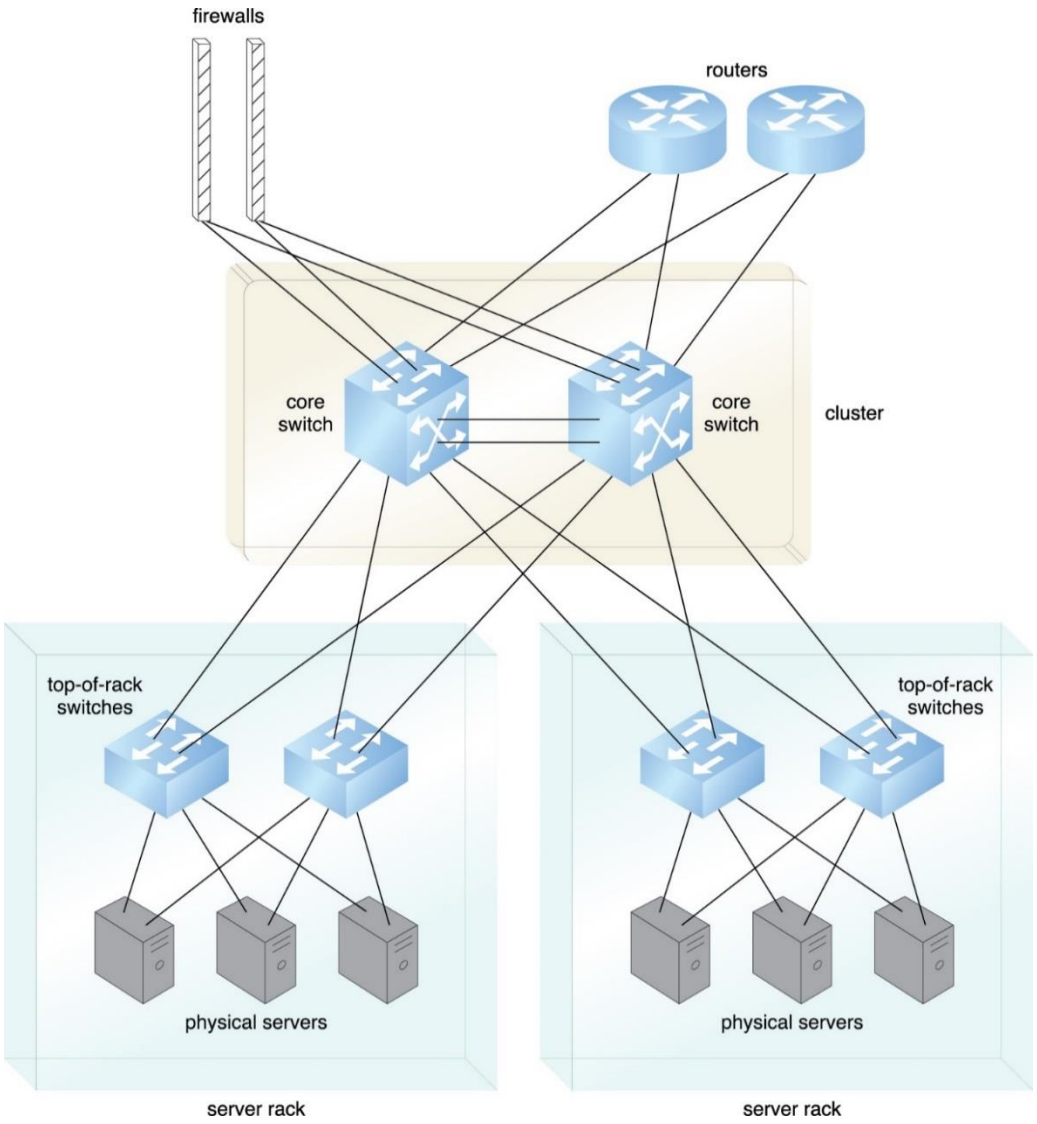

Figure 5.13 A view of the server network connections inside the DTGOV data center

Case Study Example

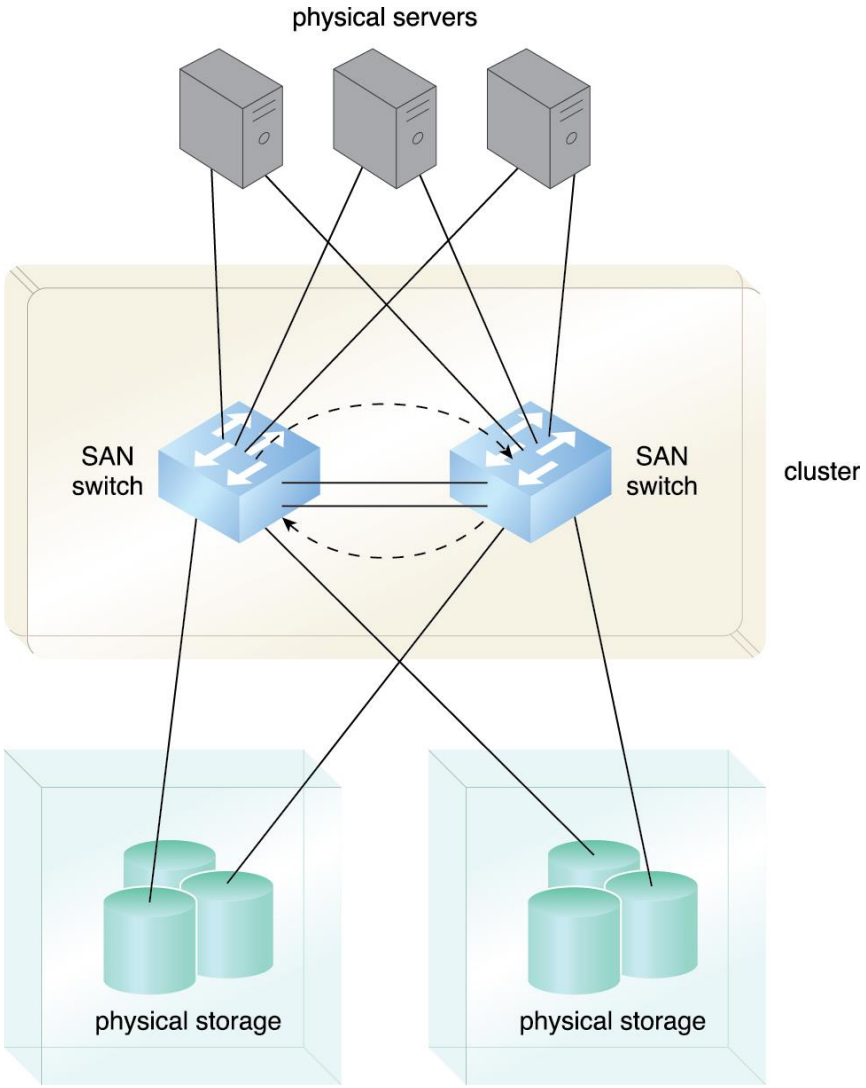

Figure 5.14 A view of the storage system network connections inside the DTGOV data center

Case Study Example

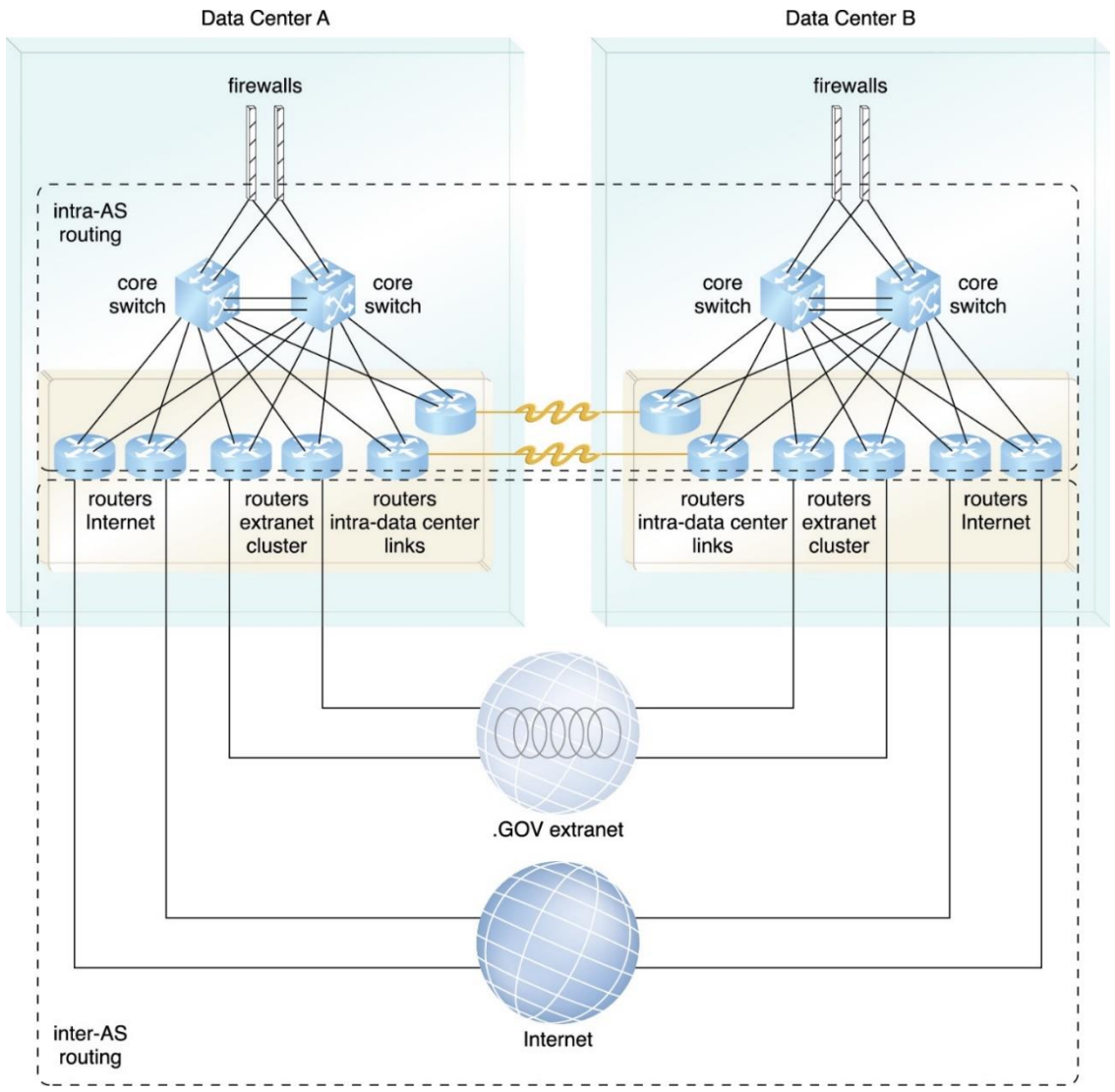

Figure 5.15 The internetworking setup between two data centers that is similarly implemented between every pair of DTGOV data centers. The DTGOV internetwork is designed to be an autonomous system (AS) on the Internet, meaning the links interconnecting the data centers with the LANs define the intra-AS routing domain. The interconnections to external ISPs are controlled through inter-AS routing technology, which shapes Internet traffic and enables flexible configurations for load-balancing and failover.