Data Center
Design & Virtualization

Md. Jahangir Hossain
Open Communication Limited

jahangir@open.com.bd

Objectives

Data Center Architecture

- Data Center Standard
- Data Center Design Model
- Application Design
- Server Virtualization
- Server Clustering
- Storage Technologies
- Case Study

Data Center 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.

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.

Green Data Center

Definition

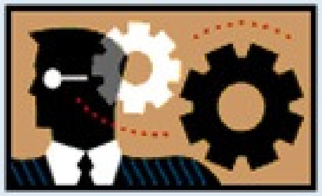
- **What is the Green Data Center?**

  A green data center is a repository for the storage, management, and distribution of data in which the lighting, electrical and computer systems are designed for maximum energy efficiency and minimum environmental impact.

Steps Involved in Creating Green Data Center

- Minimizing the footprints of the buildings
- The use of low-emission building materials, carpets and paints
- Sustainable landscaping
- Waste recycling
- Installation of catalytic converters on backup generators
- The use of alternative energy technologies such as Photovoltaic, heat pumps, and evaporative cooling
- The use of hybrid or electric company vehicles

Think Green Data Center



- **What is your data center carbon footprint?**
- **How much data center power can you reduce?**
- **How much harmful emissions can you help to reduce?**

By reducing 50 servers in you data center can reduce the following emissions over 3 years:

Carbon Dioxide (CO_2)	1,790,982 Lbs or 895 Tons
Methane (CH_4)	37 Lbs
Nitrous oxide (N_2O)	25 Lbs
Sulfur dioxide (SO_2)	7143 Lbs or 3.57 Tons
Nitrogen oxide (NO_x)	2239 Lbs or 1.11 Tons

Which is equivalent to: 177 *Passenger cars not driven for one year !!*

Data Center standard

Data Center Standard

- Tier Level 1
- Tier Level 2
- Tier Level 3
- Tier Level 4

Tier 1 to 4 data center is nothing but a standardized methodology used to define uptime of data center. This is useful for measuring:

a) Data center performance
b) Investment
c) ROI (return on investment)

Note: Access Control Should be considerable for Data Center standard

Tier Level 1

Single non-redundant distribution path serving the IT equipment
Non-redundant capacity components
Basic site infrastructure guaranteeing 99.671% availability

Tier Level 2

Fulfills all Tier 1 requirements
Redundant site infrastructure capacity components guaranteeing 99.741% availability

Tier Level 3

- Fulfills all Tier 1 and Tier 2 requirements
- Multiple independent distribution paths serving the IT equipment
- All IT equipment must be dual-powered and fully compatible with the topology of a site's architecture
- Concurrently maintainable site infrastructure guaranteeing 99.982% availability

Tier Level 4

- Fulfills all Tier 1, Tier 2 and Tier 3 requirements
- All cooling equipment is independently dual-powered, including chillers and heating, ventilating and air-conditioning (HVAC) systems
- Fault-tolerant site infrastructure with electrical power storage and distribution facilities guaranteeing 99.995% availability

Data Center Design Model

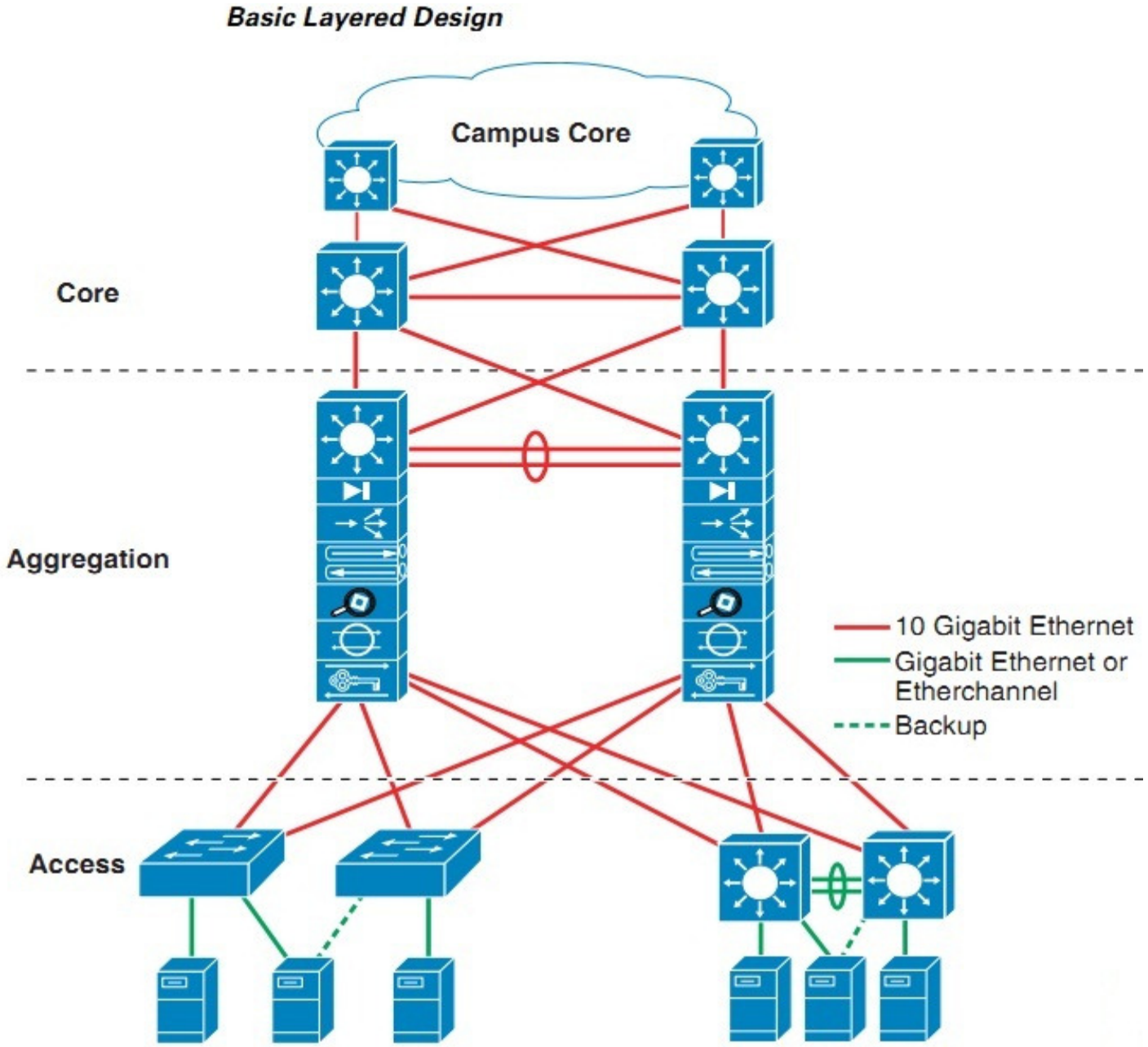
-Multi Tier Model

Multi -Tier Model

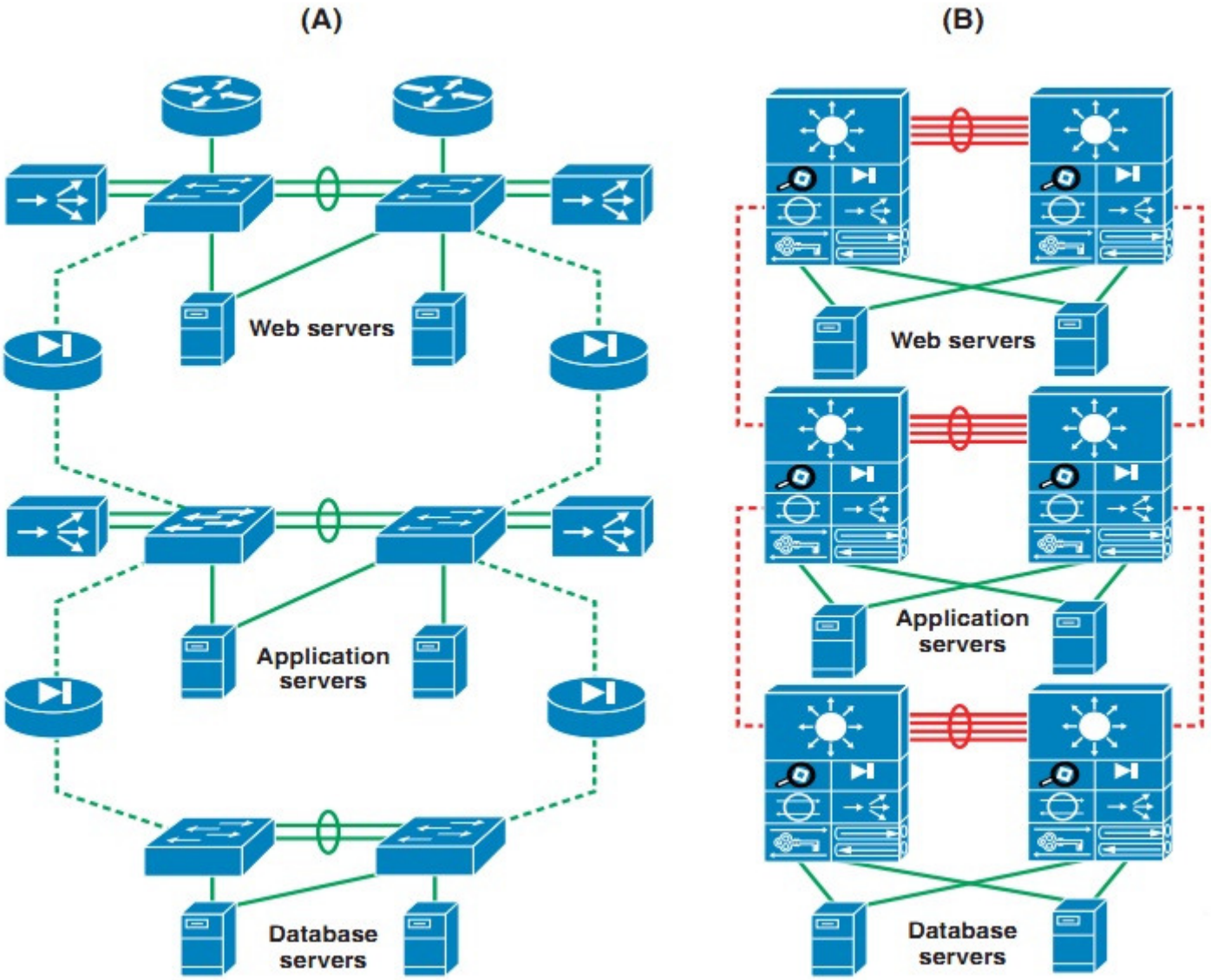
The multi-tier data center model is dominated by HTTP-based applications in a multi-tier approach. The multi-tier approach includes web, application, and database tiers of servers. Today, most web-based applications are built as multi-tier applications. The multi-tier model uses software that runs as separate processes on the same machine using interprocess communication (IPC), or on different machines with communications over the network. Typically, the following three tiers are used:

- Web-server
- Application
- Database

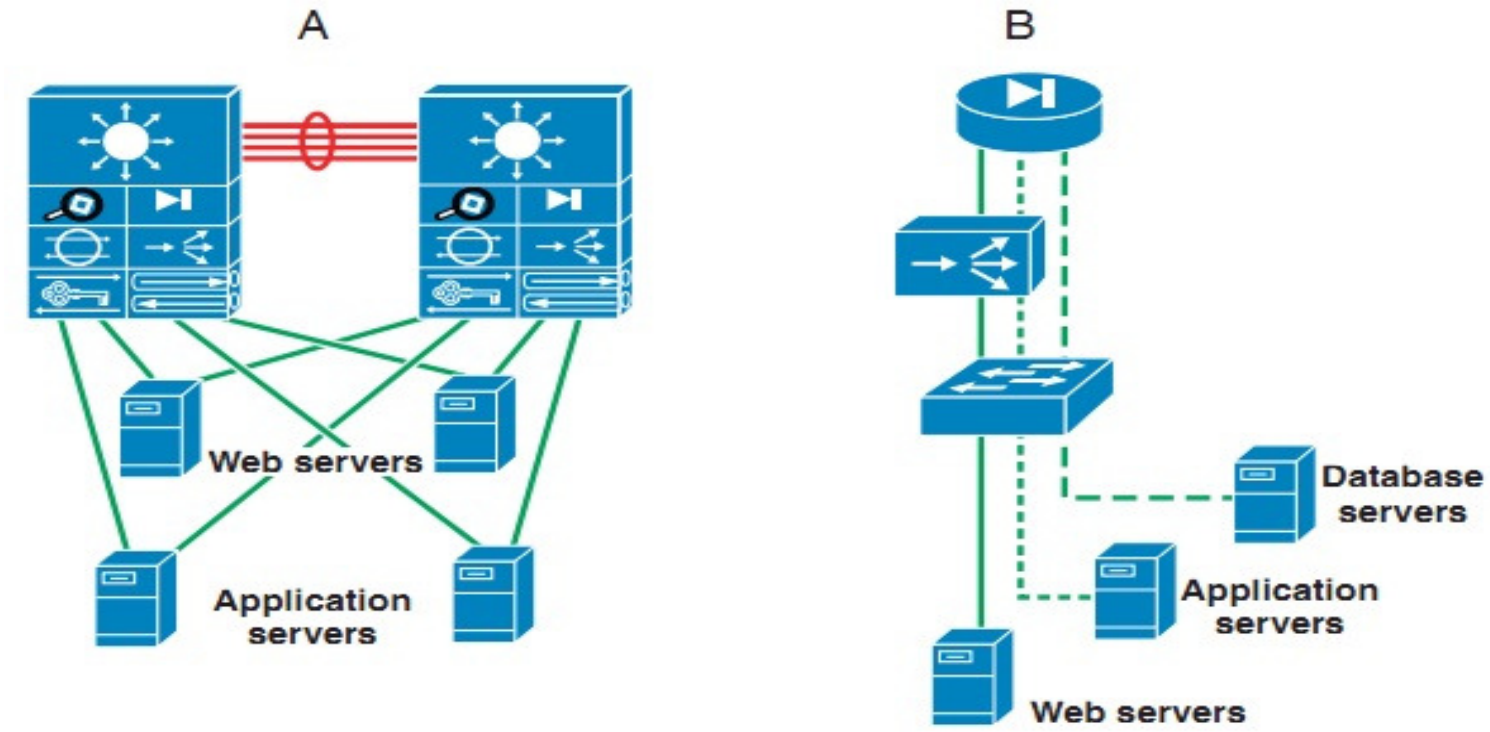
Multi-tier server farms built with processes running on separate machines can provide improved resiliency and security.



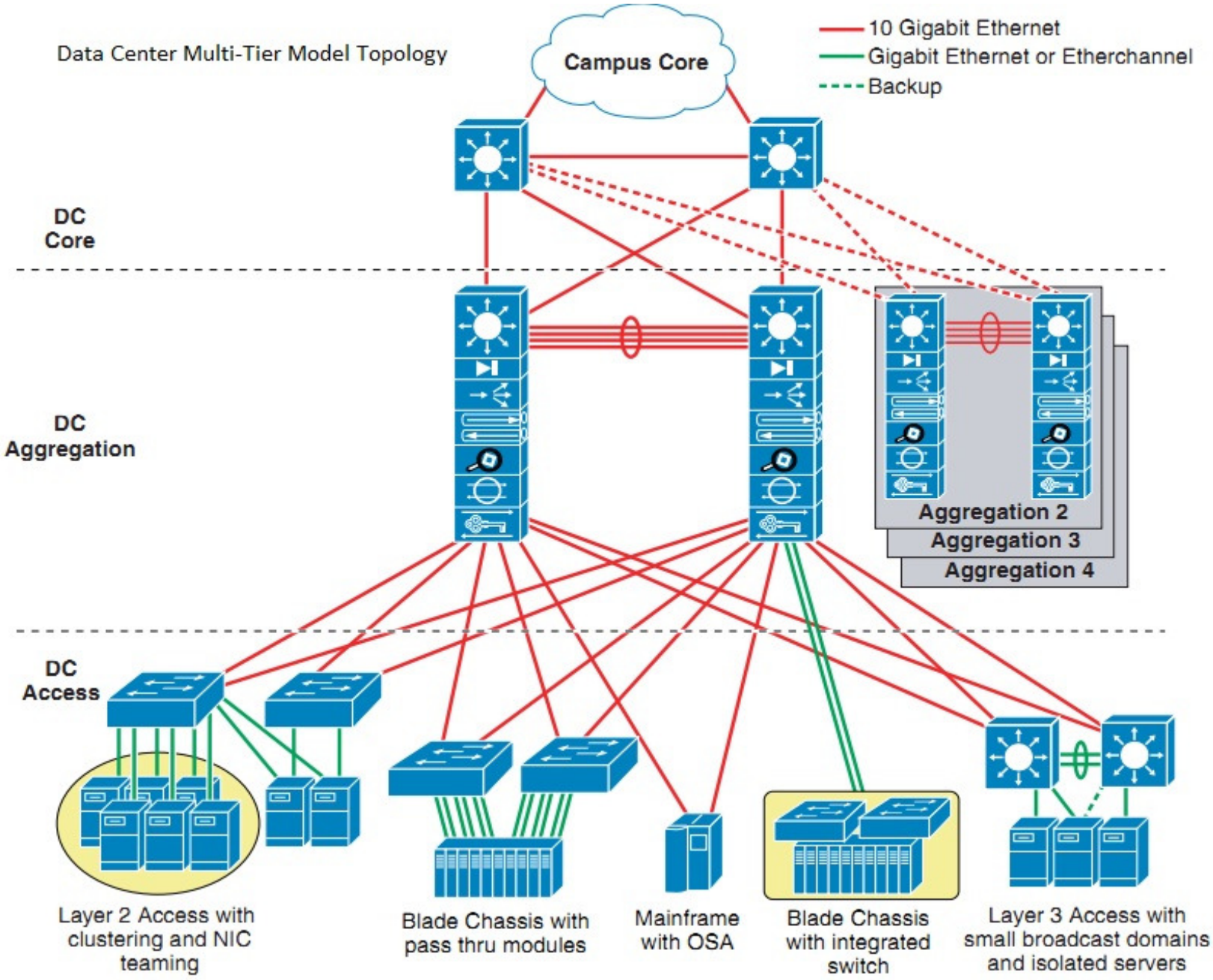
Physical Segregation in a Server Farm with Appliances (A) and Service Modules (B)



Logical Segregation in a Server Farm with VLANs



Physical segregation improves performance because each tier of servers is connected to dedicated Hardware. The advantage of using logical segregation with VLANs is the reduced complexity of the Server farm. The choice of physical segregation or logical segregation depends on your specific network performance requirements and traffic patterns.

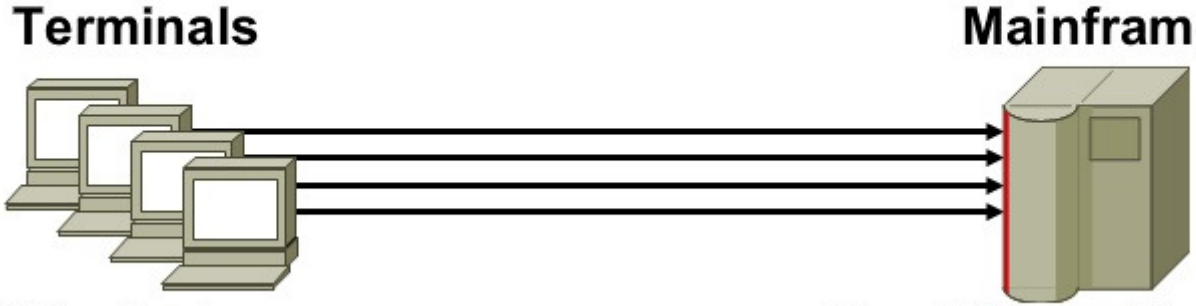


Application Design

Application Design Model

- One Tier Model
- Two Tier Model
- Three Tier Model
- N Tier Model

One-Tier Model



- Thin client
- Dumb terminal
- No local processing
- No storage

- Monolithic application
- Application intelligence
- Database system

- Limited scalability
- Lacks flexibility

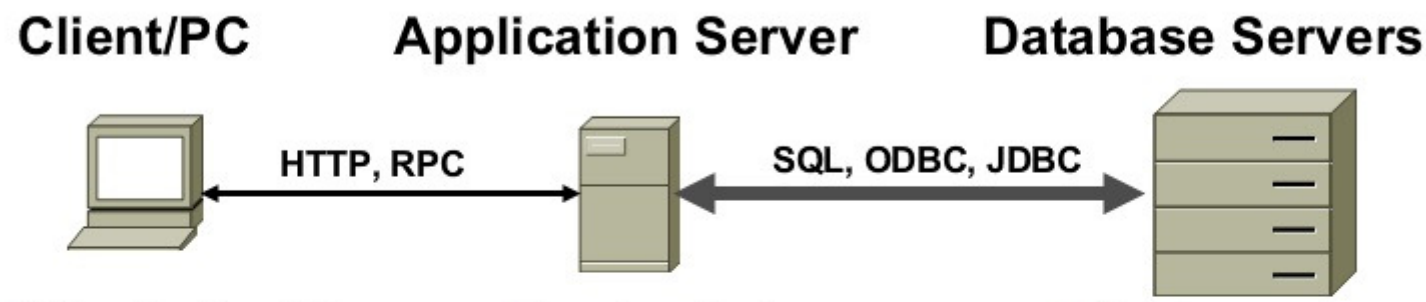
Two-Tier Model

- Direct interaction with DB server
- Local application processes

- Application intelligence
- Database system

- Limited scalability
- Generally not recommended for critical applications

Three-Tier Model



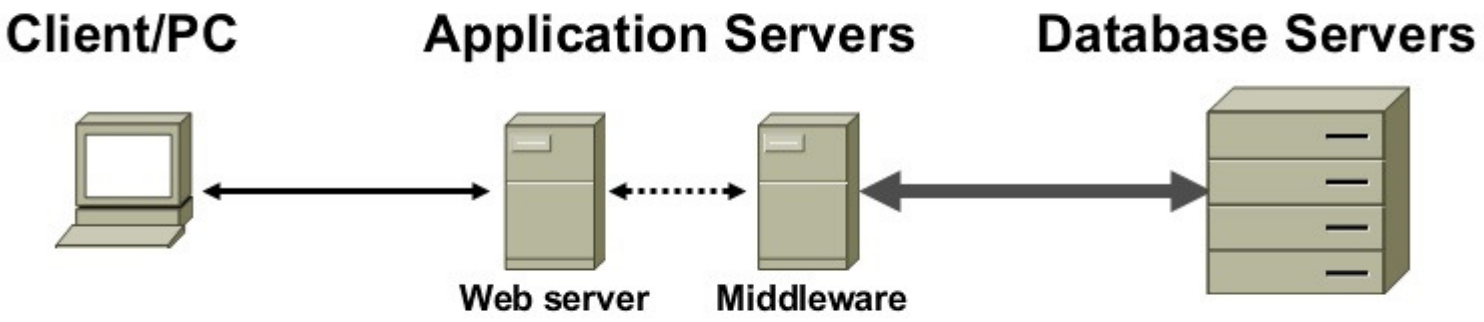
- Thin client (again)
- Direct interaction with application server only
- Presentation logic only

- Local application processes
- Application intelligence

- Database system

- **Scalability increase due to network insulation**
- **Lighter traffic to and from clients**
- **Heavier traffic to and from DB server**

N-Tier Model



- Thinner client
- Application logic spread out across multiple tiers
- Database system

- More scalability
- Robust, logical partitioning of application functionality
- N = (2 + (# of App Servers))

Web Services Model

Tasks performed by multiple hosts with specific roles

Web server

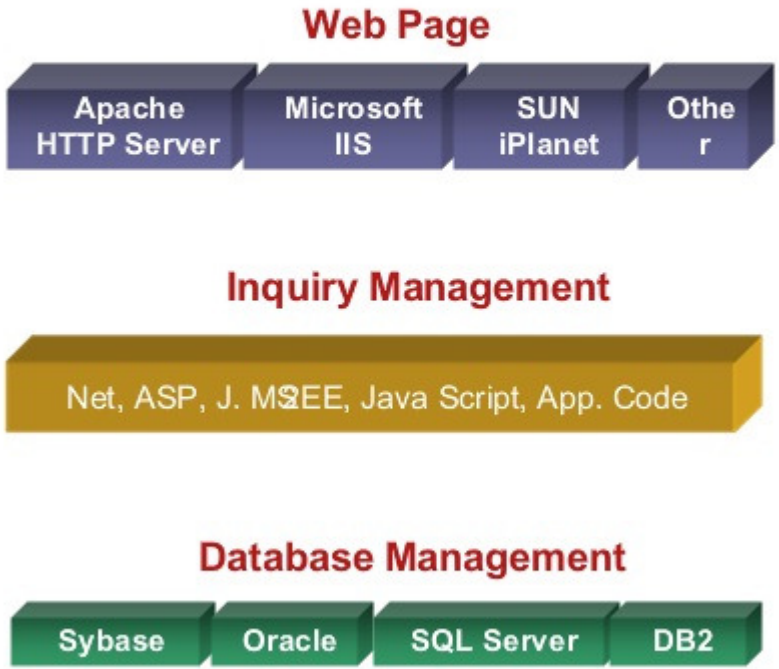
- Static versus dynamic content
- Serves web pages
- Web 2.0 and SaaS

Application server

- Implements business logic
- Manipulates data
- Data mining

Database server

- Accesses data store
- High transaction rate
- High Bandwidth
- Low Latency



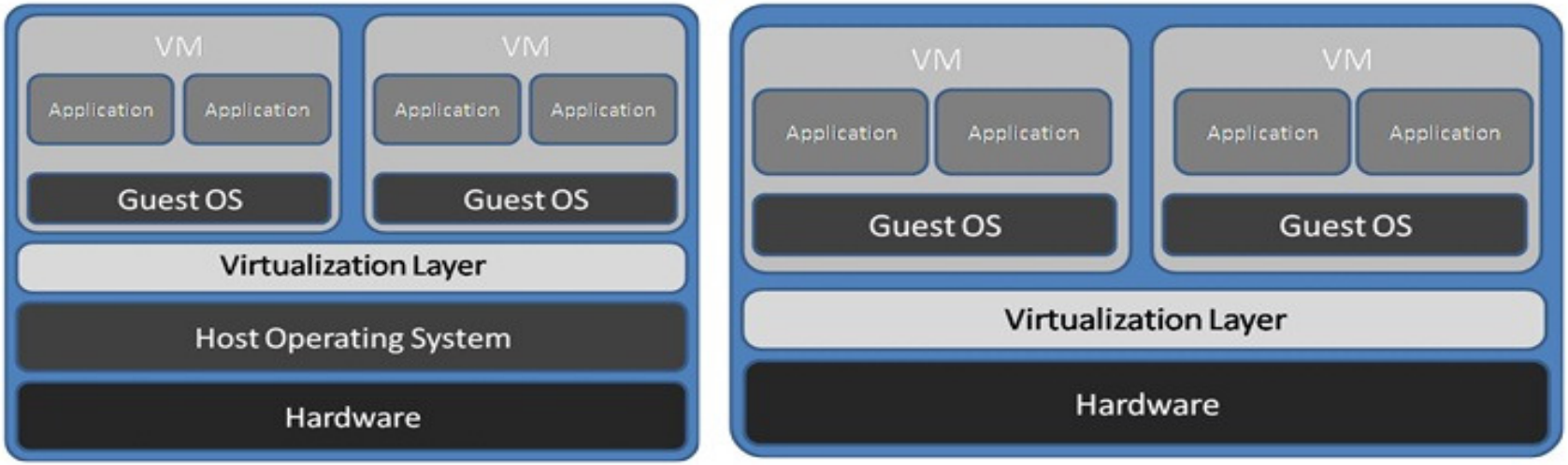
Virtualization

Objectives

- Virtualization
- Virtualization in Brief
- Server Virtualization
- The Technology of the future
- Server virtualization – best practices
- Solution for server virtualization

Virtualization is simple to understand

- **The term virtualization refers to the abstraction of system resources to allows multiple operating systems to run on one system at the same time.**
- **this is done by inserting a virtualization layer which uses either a hosted or hypervisor architecture.**



Why Virtualization?

1. Lower number of physical servers - you can reduce hardware maintenance costs because of a lower number of physical servers.

2. By implementing a server consolidation strategy, you can increase the space utilization efficiency in your data center.

3. By having each application within its own "virtual server" you can prevent one application from impacting another application when upgrades or changes are made.

4. You can develop a standard virtual server build that can be easily duplicated which will speed up server deployment.

5. You can deploy multiple operating system technologies on a single hardware platform (i.e. Windows Server 2003, Linux, Windows 2000, etc).

Still Need a Proof !!!

Healthcare Example

TCO Comparisons	Without Virualization	With Virtualization
Hardware and Software Costs		
Number of Physical Servers Required	62	6
Total Hardware Costs	$434,000	$38,757
Hardware Maintenance	$43,500	$16,757
VMware Software	$0	$21,000
VMware Software Support	$0	$5,250
VMware Training & Services	$0	$19,500
Total Hardware and Sofware Costs	**$477,500**	**$101,263**
Hardware and Software TCO Reduction		79%
IT Operations		
Affected Datacenter Costs: (SAN port and power costs)	$8,637	$31,526
Total Server Deployment Cost:	$59,520	$7,440
Server Development Time (hrs)	1488	186
Average Hourly Labor Cost	$40	$40
Server Support	N/A	N/A
Total Affected IT Operations Costs	**$68,157**	**$38,966**
IT Operations TCO Reduction		43%
Total Affected costs	$545,657	$140,230
Total TCO Reduction		**74%**
Six Month ROI		289%
Other Benefits		
Recovery Time (hrs)	12	1
Server Consolidation Ratio	10	1
Average CPU Utilization	5%	80%

Is that Right .. ???

- **We can say now that server virtualization has resulted in cost savings and efficiencies through server consolidation, so what's next in a virtualization strategy?**
  - Virtualizing Desktops
  - Virtualizing Networking
  - Using virtualization for DR and BC

Server Virtualization Best Practices

High-Level Network Design Guidelines

- The network design must be optimized to meet the diverse needs of applications, services, storage, administrators, and users.
- The goal is to design a network infrastructure that reduces costs, boosts performance, improves availability, provides security, and enhances functionality.

High-Level Network Design Requirements

- When planning the network design, consider how to determine and meet:
  - Connectivity requirements
  - Bandwidth requirements
  - Latency requirements
  - Availability requirements
  - Cost requirements

Number of Networks

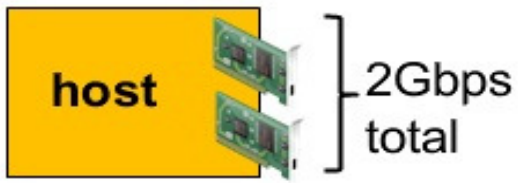
- How many networks or VLANs are required depends on the types of traffic required for:
  - Organization's services and applications
  - IP Storage
  - High Availability
  - Management

Why Separate Networks? (1)

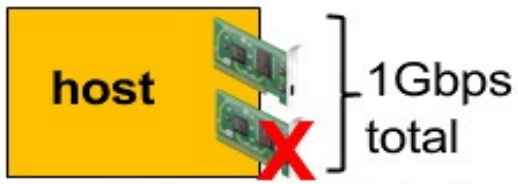
- There are two main reasons to separate different types of network traffic:
  - To reduce contention and latency, and improve performance:
    - High latency on any network has the potential to negatively affect performance.
    - This is especially important for IP storage, and depending on their workloads, some virtual machine networks.
  - To enhance security by limiting network access:
    - For example, IP storage traffic are not encrypted, so a separate network helps protect what could be sensitive data.

Why Separate Networks? (2)

- To avoid contention between different types of network traffic, configure enough physical NIC ports to satisfy bandwidth requirements.
  - Use a capacity-analysis report from the current-state analysis to determine bandwidth requirements.
  - Knowing the bandwidth per NIC port is more critical if VLAN trunk ports are configured.
  - Also consider network failures.
    - Will there be enough bandwidth remaining after a physical port or network failure?



sufficient bandwidth



insufficient bandwidth

Network Segmentation (1)

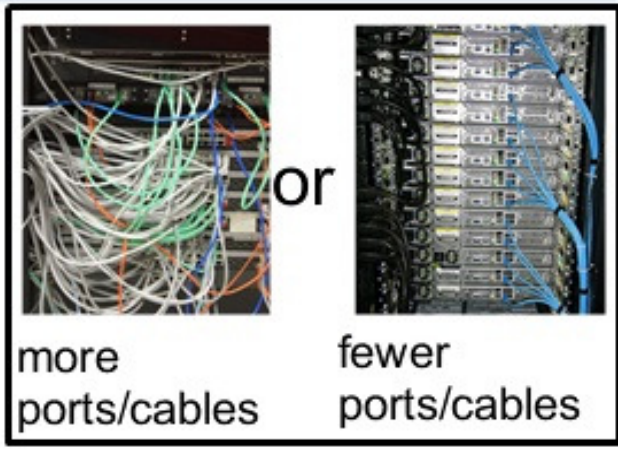
- The decision to use physical networks or VLANs depends on several factors, including:
  - The number of networks required and the number of available physical ports:
    - If the number of physical network ports is at least equal to the number of required networks, the use of VLANs is optional.
      - Consider redundancy, too.
    - If the number of networks required is greater than the number of available physical network ports, you have two choices:
      - You can configure VLANs to create separate virtual networks.
      - You can purchase hardware that supports a greater number of network ports.

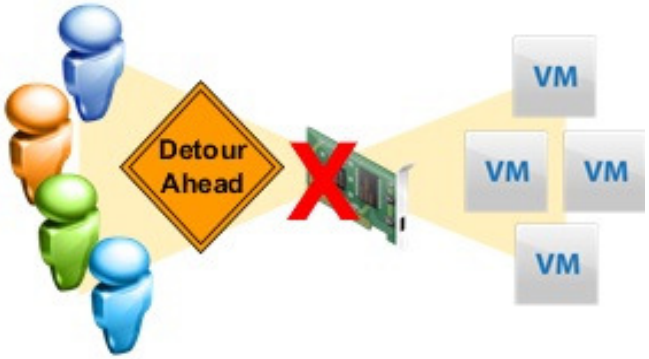
Network Segmentation (2)

- There are other factors that might determine the choice between physical or VLAN network segmentation:
  - Physical port speed:
    - A 10Gb Ethernet port often offers sufficient bandwidth for several VLANs.
  - Whether or not the physical network infrastructure supports VLANs
  - An organization's existing use of VLANs
  - Network security policies might require physical separation.

VLAN Benefits and Risks (1)

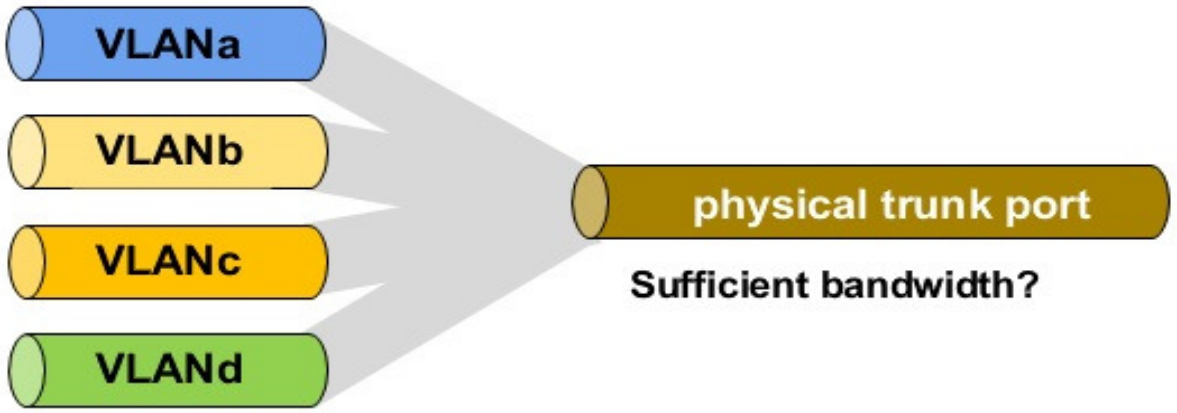
- VLANs provide many benefits, including:
  - Reducing the number of required physical ports and cabling
  - Allowing easier reconfiguration of networks and servers
  - Reducing server and administration costs
- Port failures in a VLAN environment are also potentially more serious.
  - Because many networks (and therefore many services) share a single port, a failure affects more services.
  - Network redundancy becomes more important in a VLAN environment.





VLAN Benefits and Risks (2)

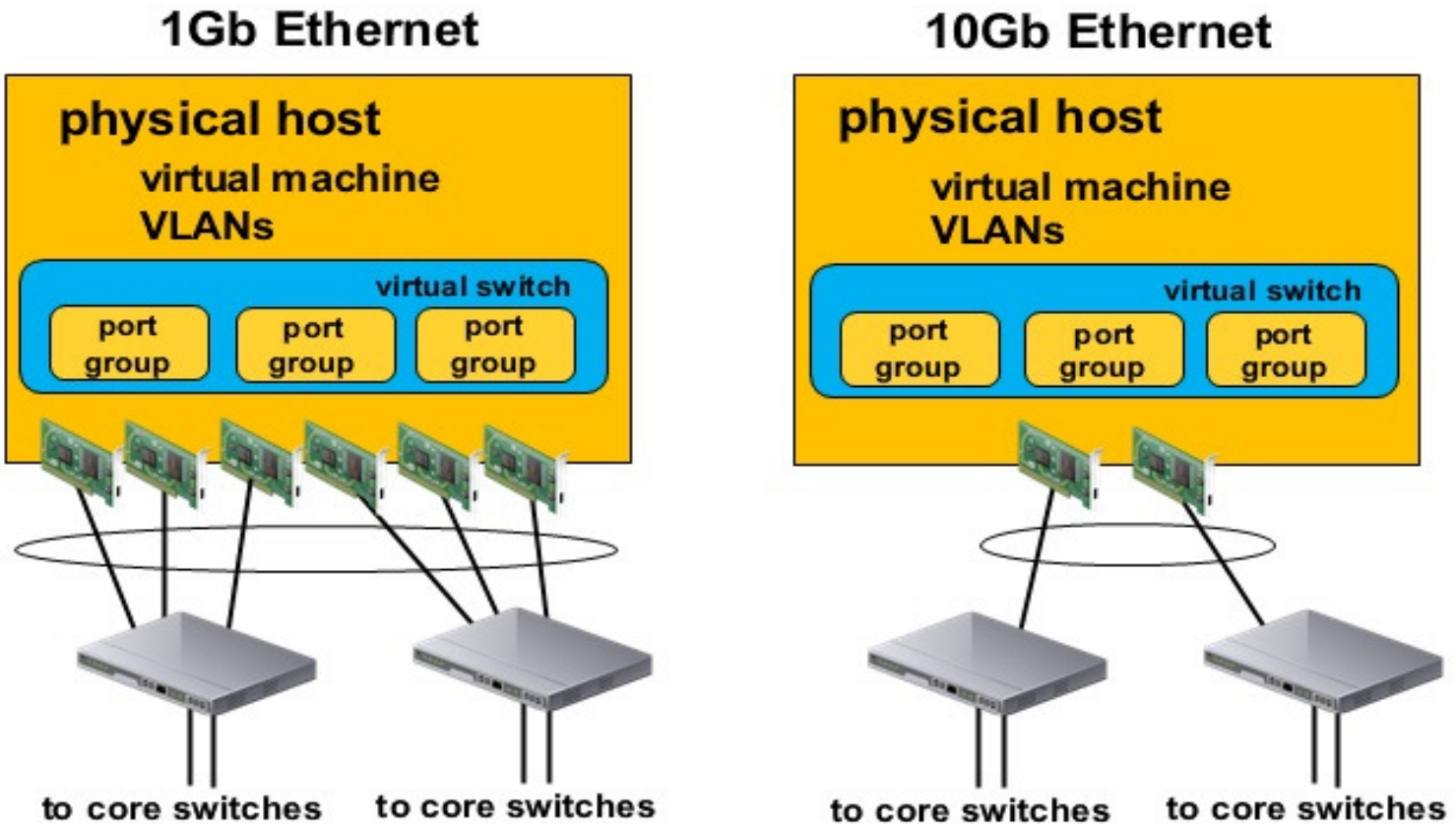
- It is easier to exceed available bandwidth on VLAN trunk ports.
  - NIC teaming can provide the additional bandwidth.
  - 10Gb Ethernet can provide additional bandwidth.



Implementing VLANs

- When implementing VLANs, consider:
  - Physical switches must support 802.1Q VLAN tagging.
  - Physical switch trunk ports must be manually configured.
    - Virtual switches are passive devices and do not participate in protocols like Dynamic Trunking Protocol (DTP) or Link Aggregation Control Protocol (LACP).
  - Configure PortFast mode on the physical switch ports.

Example VLAN Configurations



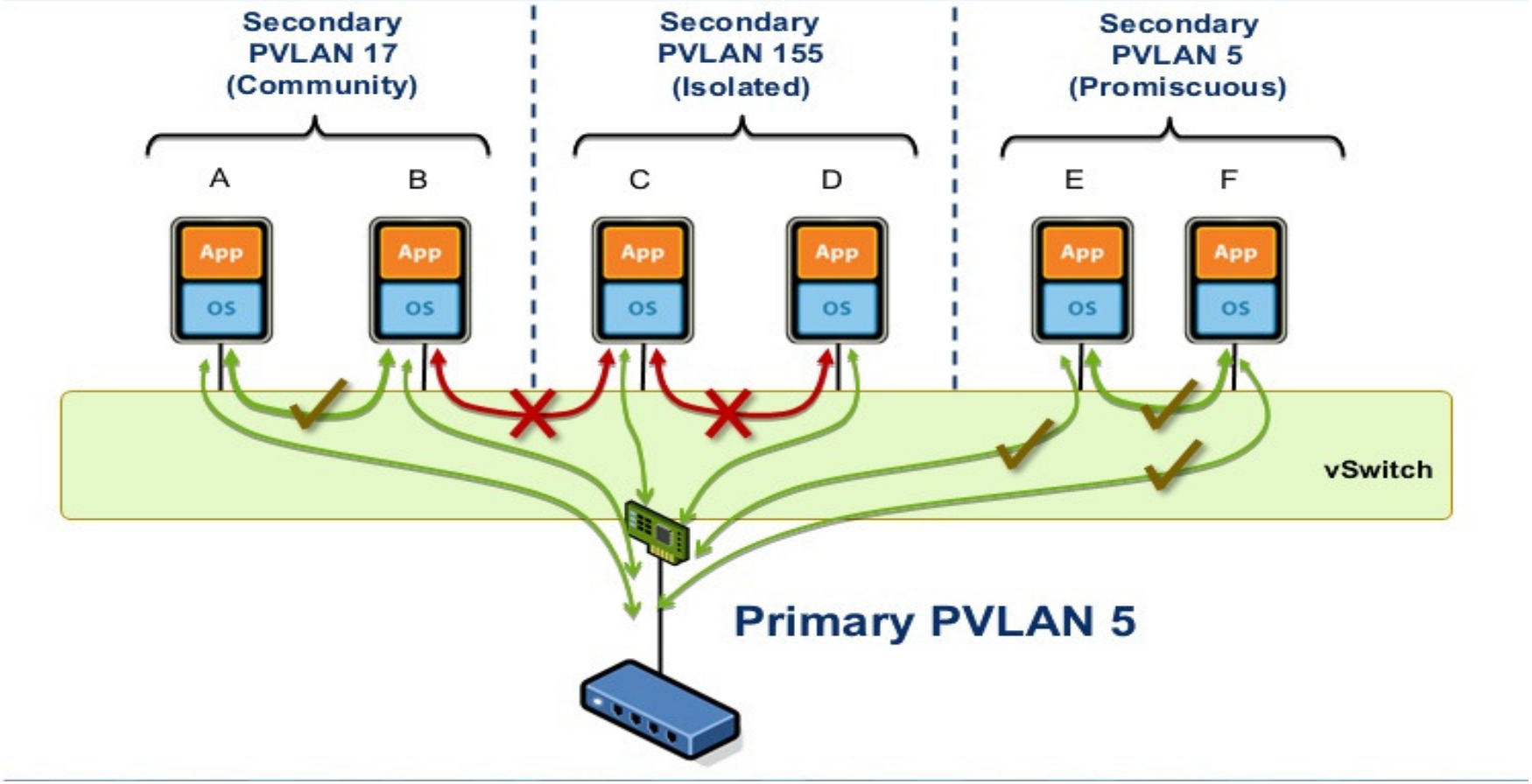
Private Virtual LANs

- PVLANs are a way of easily providing layer 2 network isolation between servers in the same subnet or network, without having to worry about such things as MAC access control lists.
- PVLANs significantly reduce the number of IP subnets needed for certain network configurations.

PVLAN Usage

- Configure PVLANs when you need to limit communication between servers in the same layer 2 network.
  - Configure PVLANs to protect the servers from each other if one becomes compromised.
- A common use for PVLANs in a corporate environment is in the DMZ network.
  - Configure a *community* PVLAN for each multitier application running across multiple servers.
    - Servers in the same community PVLAN can communicate with each other but are isolated from other servers in the same network.
    - This increases network security.
  - Configure an *isolated* PVLAN for all the other servers.
    - Each server is isolated from other servers on the same subnet.
    - This increases network security.

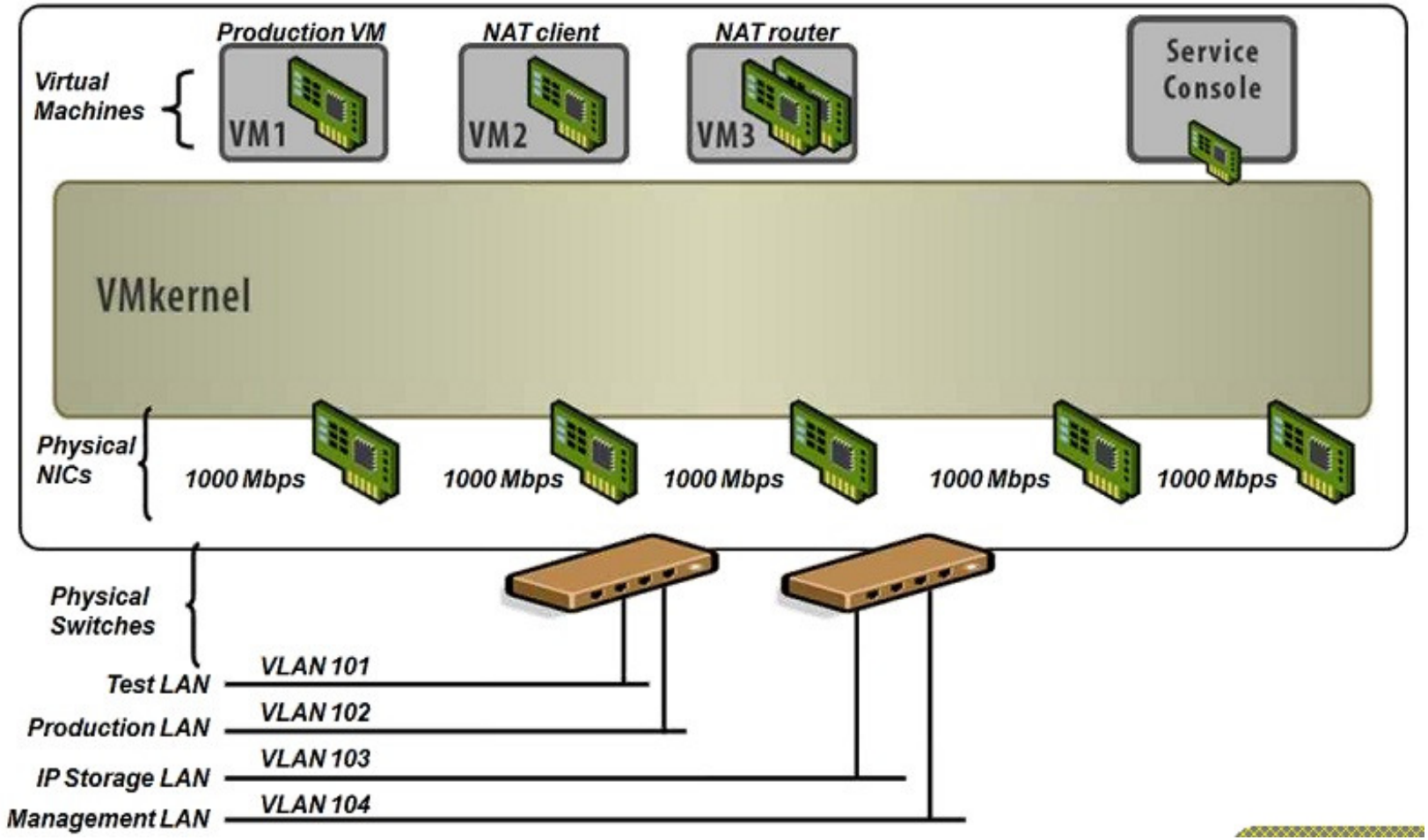
Private VLAN Architecture



Number of Virtual Switches

- Create fewer virtual switches, preferably one.
  - Configure a single virtual switch with a port group for each type of network traffic.
  - It simplifies configuration and monitoring.
  - One virtual switch with VLANs will work in environments with a limited number of physical network ports.
- If the organization has a policy that virtual machine–to–virtual machine traffic must pass through a physical firewall, the infrastructure will need multiple virtual switches.

A Networking Scenario

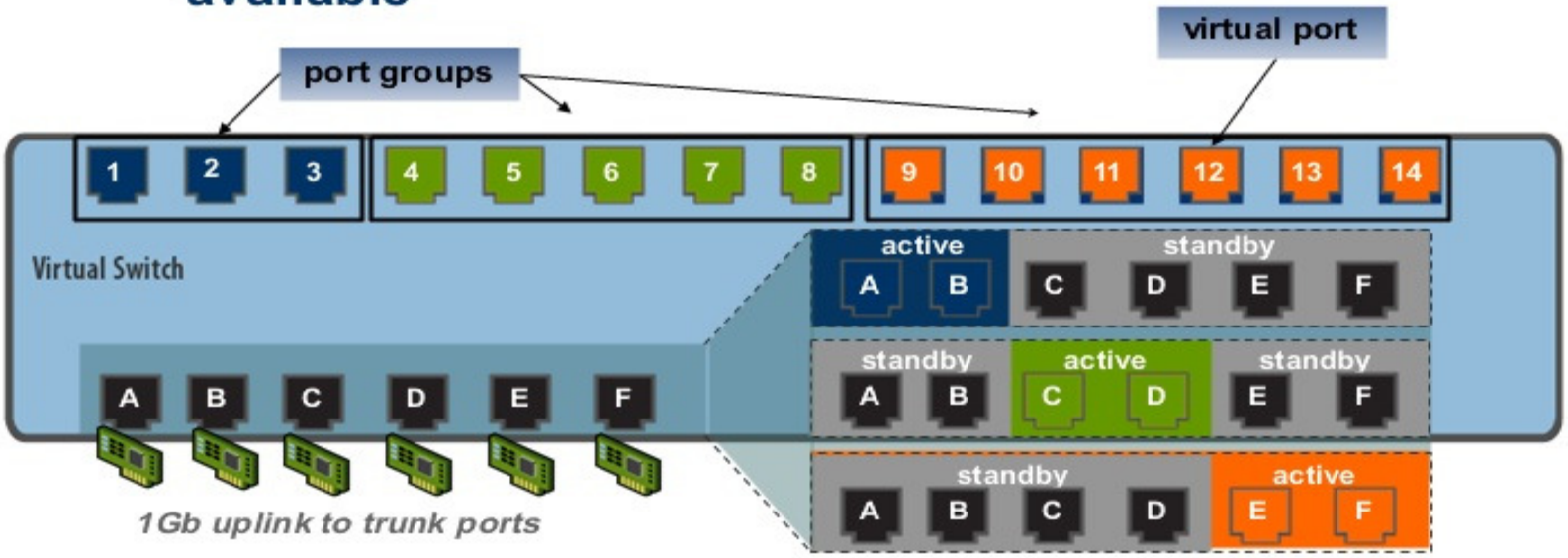


NIC Teams

- It increases the bandwidth available to a network path.
- It helps avoid any single point of failure.
  - Server consolidation compounds the effects of failure, which increases the need for redundancy.
  - Configure the NIC team by using ports from multiple NIC cards and motherboard interfaces to further reduce the number of single points of failure.
- NIC teaming requires:
  - Two or more NICs assigned to same virtual switch
  - That all NICs in the same port group are in the same layer 2 broadcast domain

NIC Teams and Availability

- Use NIC teams to reduce the required number of network ports while maintaining redundancy.
  - Use an active/standby port configuration.
  - Assumes that VLANs and sufficient bandwidth are available



Virtual Switch Security Design

- Change the default switch settings for Forged transmits and MAC address changes to Reject, unless there is an application that requires the default settings.
  - This increases security by preventing a compromised virtual machine from using MAC address spoofing to impersonate another server on the network.

Virtual Switch Security Design (2)

- If Forged transmits and MAC address changes are needed, enable them in a specific port group.
  - This reduces the security risk by reducing the possible number of servers affected.
- Leave Promiscuous mode disabled on all ports groups unless a specific application needs it.
  - If a specific application requires promiscuous mode operation, enable it only in the required port group.
  - This reduces the security risk by reducing the possible number of affected systems.
    - Applications that typically require promiscuous mode include intrusion- detection and intrusion-protection software, packet capture utilities, and performance monitoring tools.

Solution for Virtualization

- VMware
- KVM (Kernel-based Virtual Machine)
- Virtual Box
- Windows 2008 with Hyper-V
- XEN
- Citrix

Customer-Proven Solution

- Banks: commercial & saving
- Aerospace and Defense Companies
- Internet Service provider
- Airlines
- Chemical Companies
- Diversified Financial Companies
- Energy Companies
- Entertainment companies
- Pharmaceutical Companies
- Securities companies

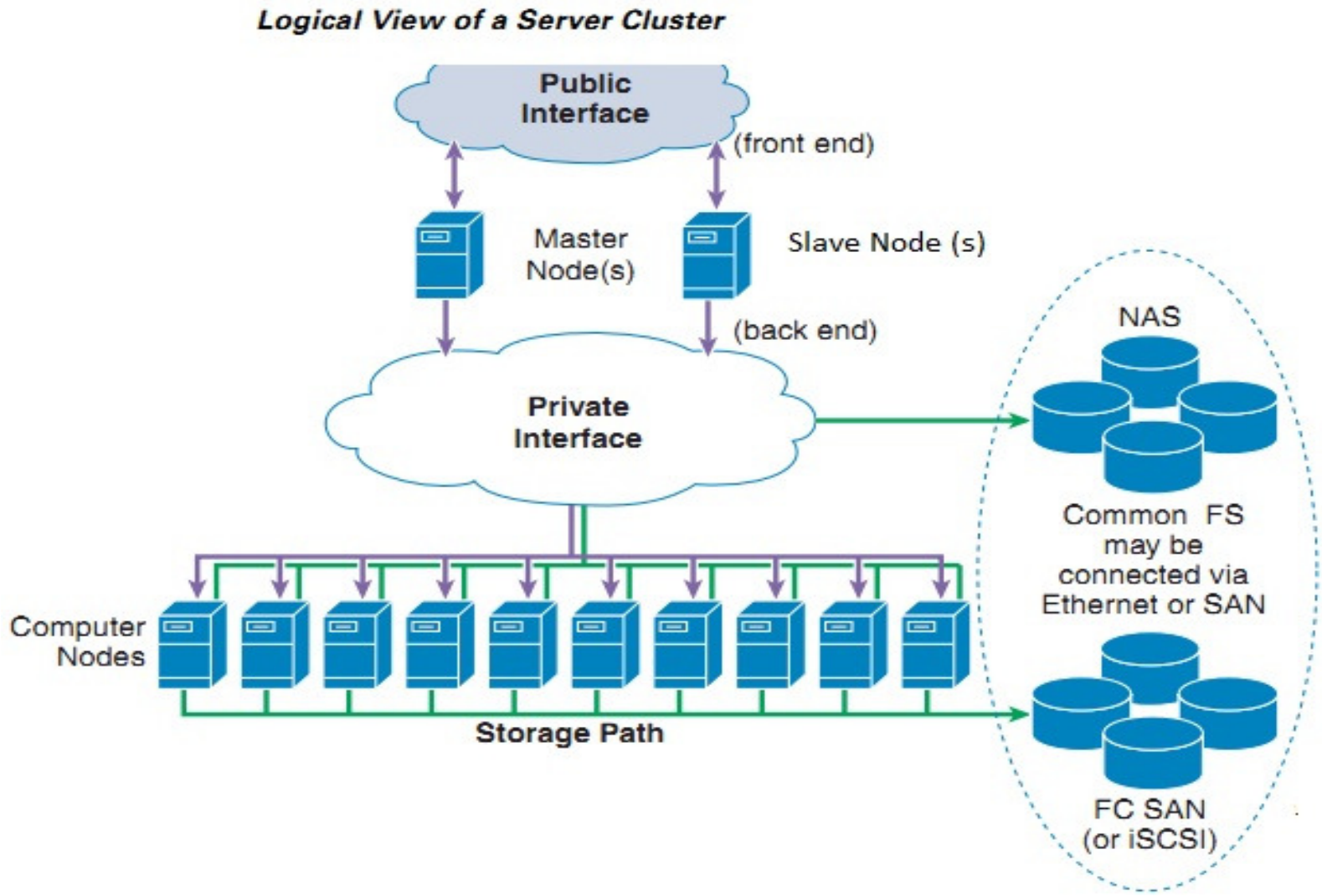
Server Clustering

Clustering

A **computer cluster** is a group of linked computers, working together closely thus in many respects forming a single computer. The components of a cluster are commonly, but not always, connected to each other through fast local area networks. Clusters are usually deployed to improve performance and availability over that of a single computer, while typically being much more cost-effective than single computers of comparable speed or availability.

Types of Cluster:

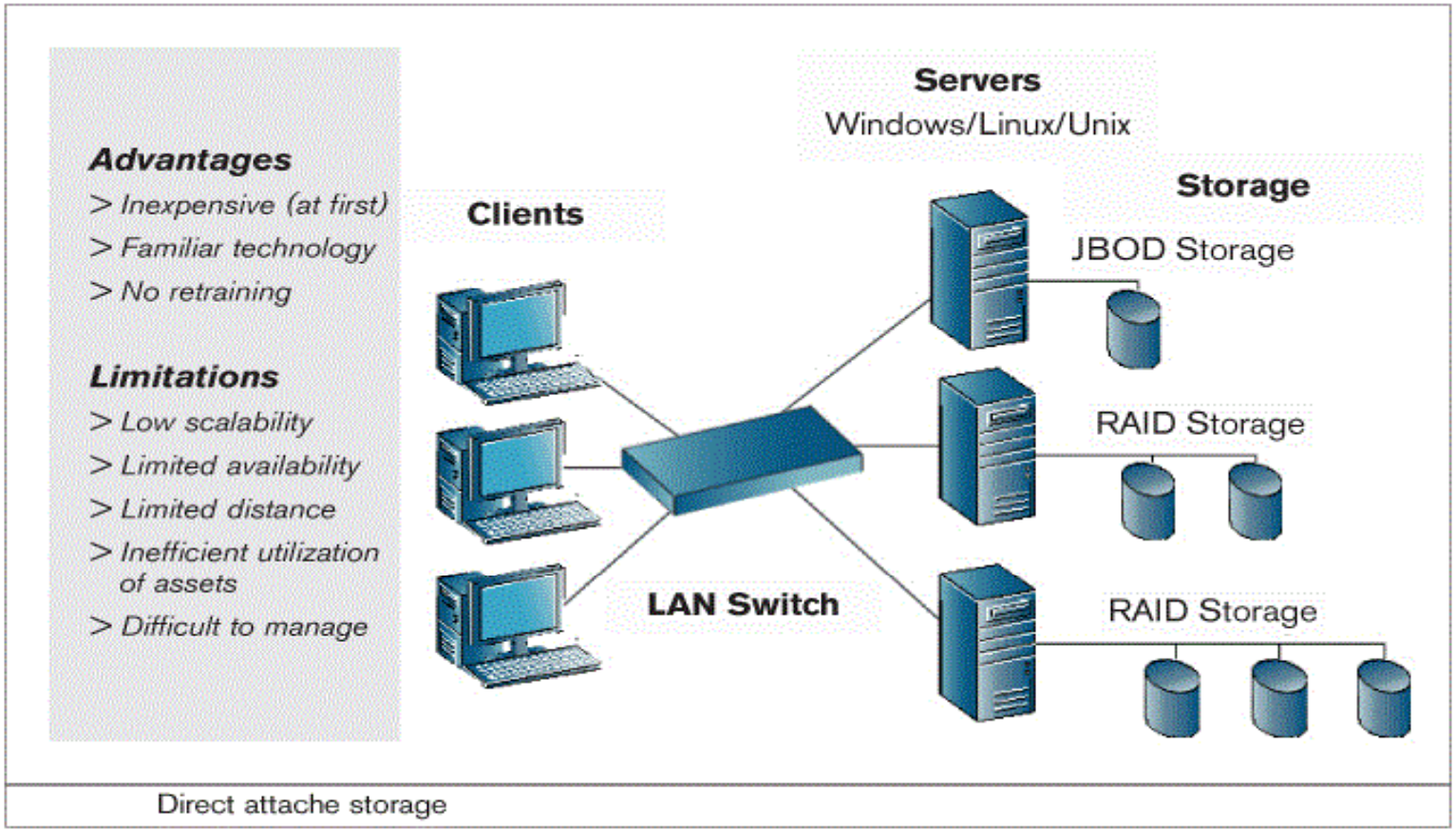
- High-availability (HA) clusters
- Load-balancing clusters
- Compute clusters



Storage Technologies

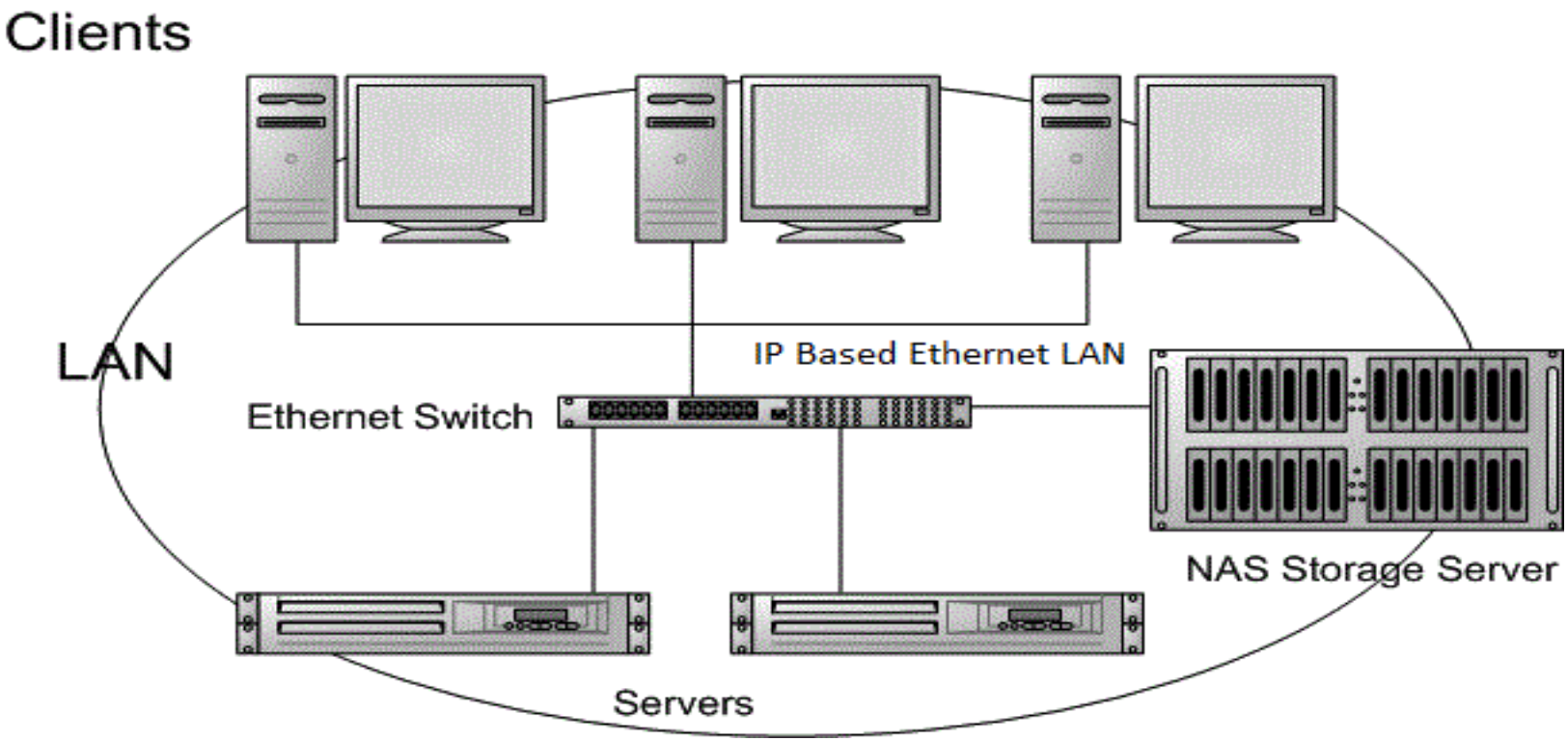
Storage Technologies

- DAS (Direct Attached Storage)
- NAS (Network Attached Storage)
- SAN (Storage Area Network)



DAS means each server has dedicated storage that is directly connected to that device. The storage device is seen and accessed by a single host system, and in the event that another host system will need additional storage, that host will add more physical storage and/or I/O interfaces or host bus adapters.

Network Attached Storage



NAS is storage that is connected directly to a network, such as a LAN, that provides file-level access to data using standard protocols such as NFS (Network File System) or CIFS (Common Internet File System).

NAS (1)

- Computer systems can access data from a NAS appliance over a network via a file "redirector" that changes the access to a file from the native file system (on the originating computer system) to a network operation using TCP (Transmission Control Protocol) to a remote server that is running software to provide the file system to support the individual client access.

- The file system on the NAS server determines the location of the data requested by the application client whether it is in its cache or on the storage. NAS mainly focuses on applications, users, and the files and data that they share.

SAN (Storage Area Network)

A **storage area network** (**SAN**) is a dedicated storage network that provides access to consolidated, block level storage. SANs primarily are used to make storage devices (such as disk arrays, tape libraries, and optical jukeboxes) accessible to servers so that the devices appear as locally attached to the operating system.

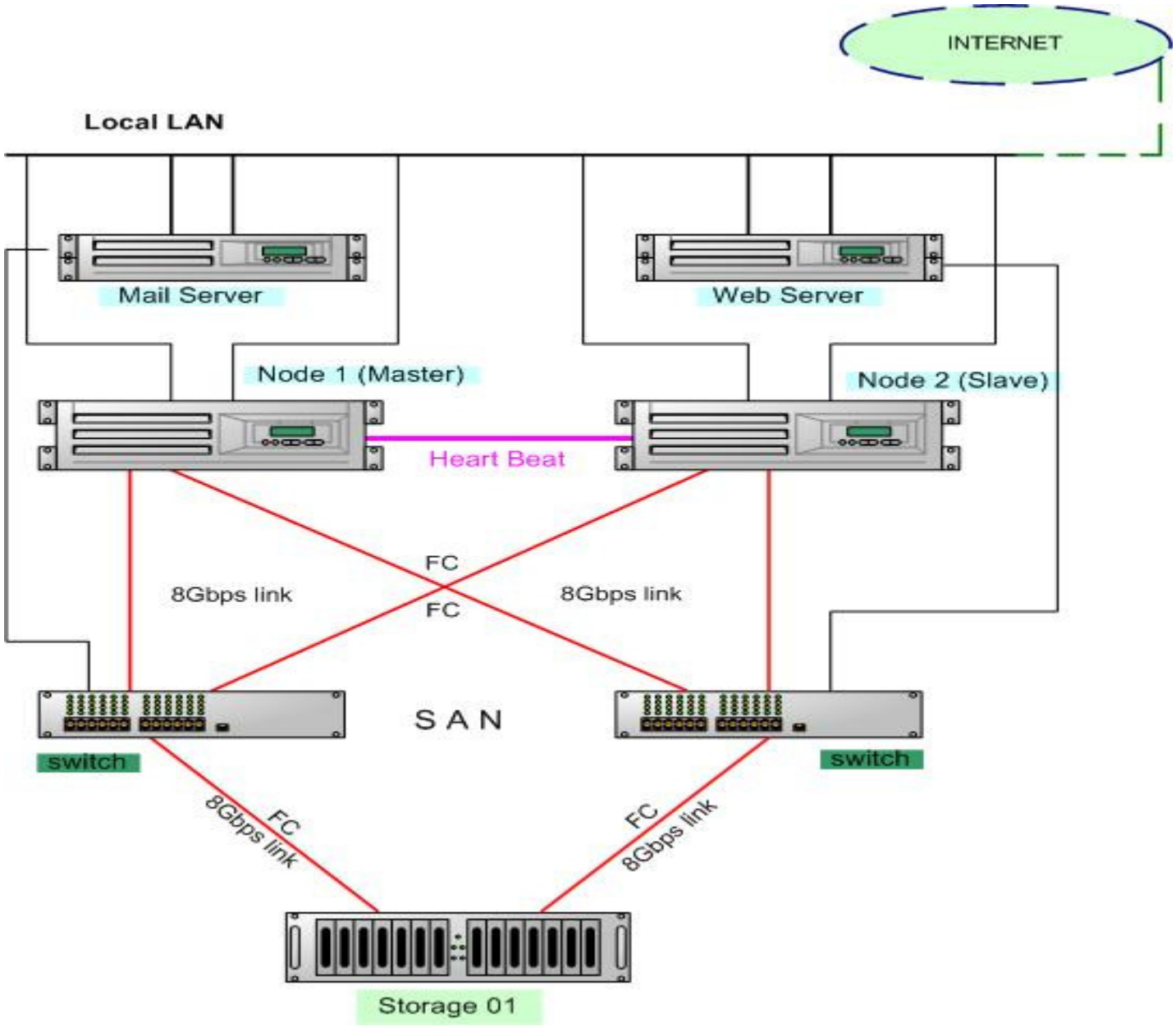
A SAN typically has its own network of storage devices that are generally not accessible through the regular network by regular devices.

--Focuses on disks, tapes, and a scalable, reliable infrastructure to connect them

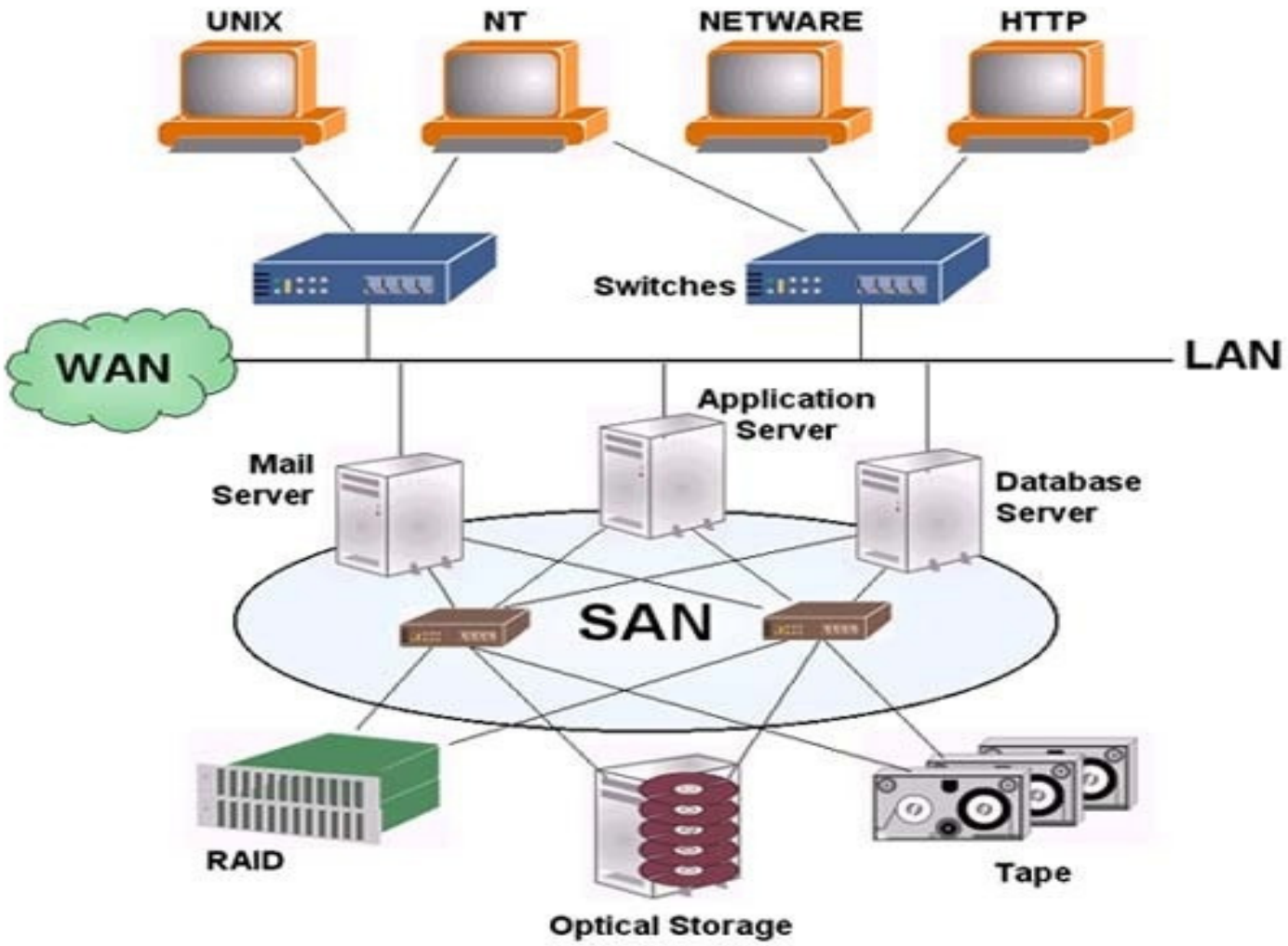
--Backup solutions (tape sharing)

--Disaster tolerance solutions (distance to remote location)

--Reliable, maintainable, scalable infrastructure



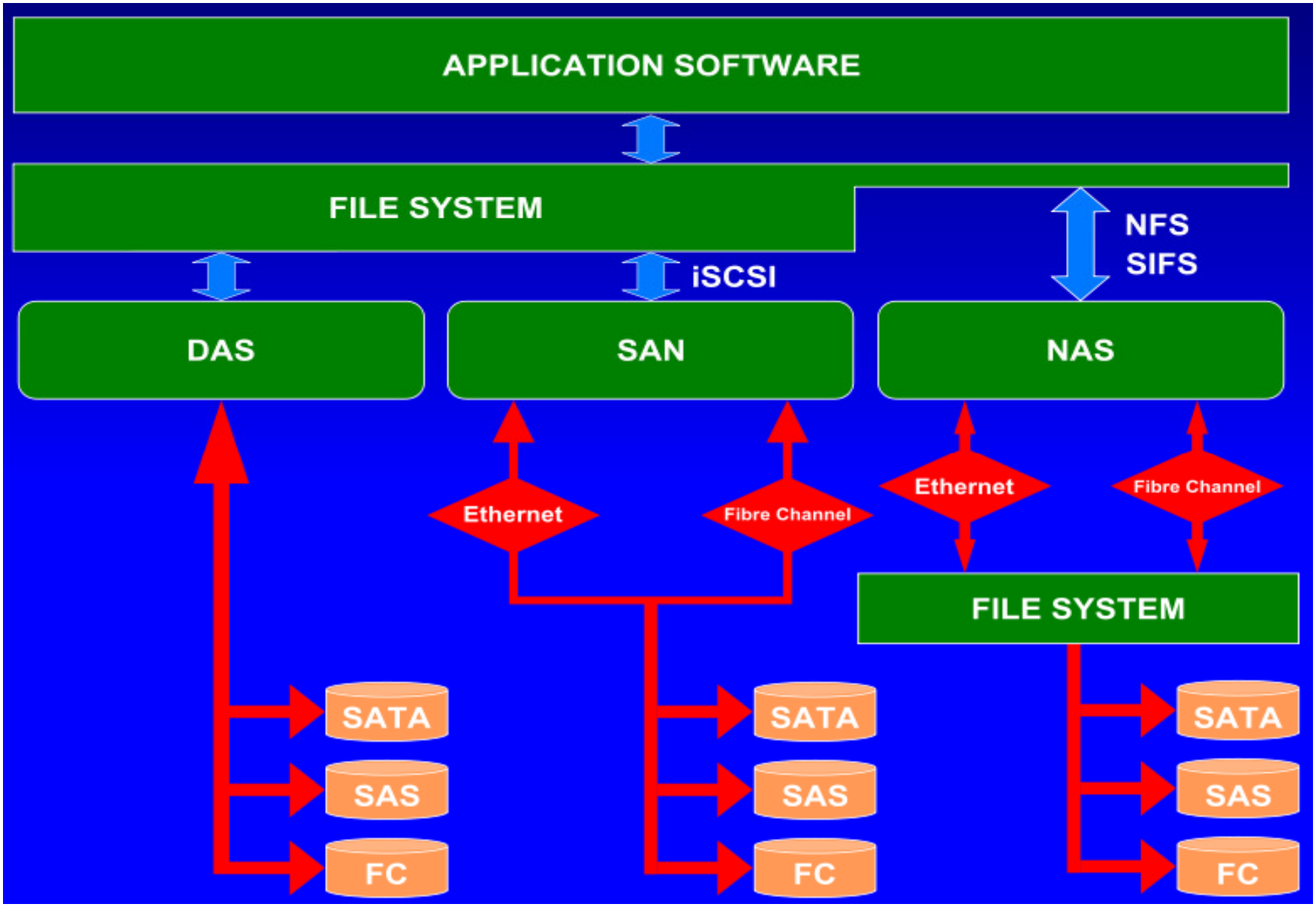
Storage Area Networks



Major Difference between NAS and SAN

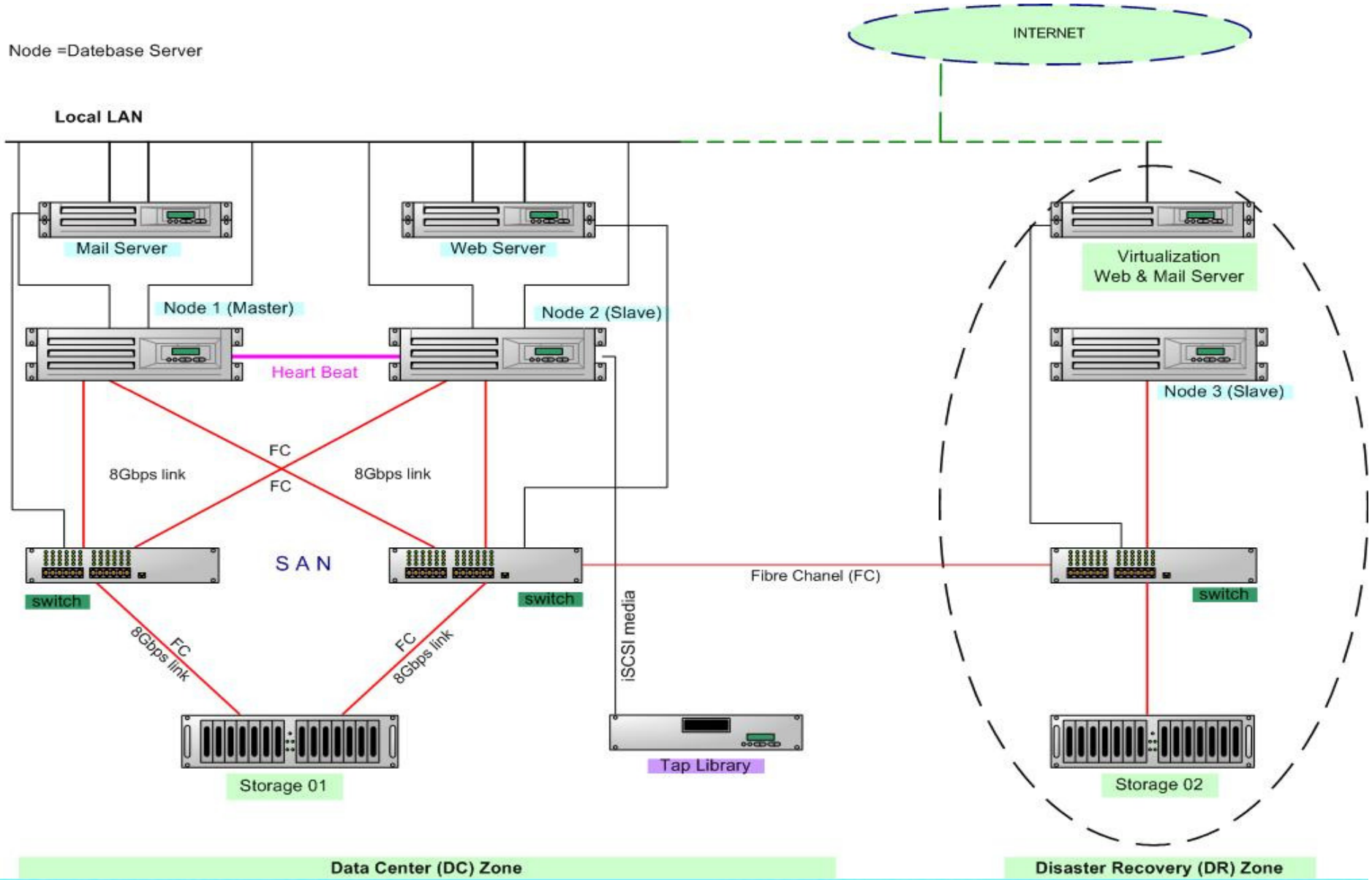
- **The Wires.**
  --NAS uses TCP/IP Networks: Ethernet, FDDI, ATM (perhaps TCP/IP over Fibre Channel someday)
  --SAN uses Fibre Channel, Ethernet.

- **The Protocols.**
  --NAS uses File Server Protocols: NFS, CIFS, and HTTP.
  --SAN uses Encapsulated SCSI.

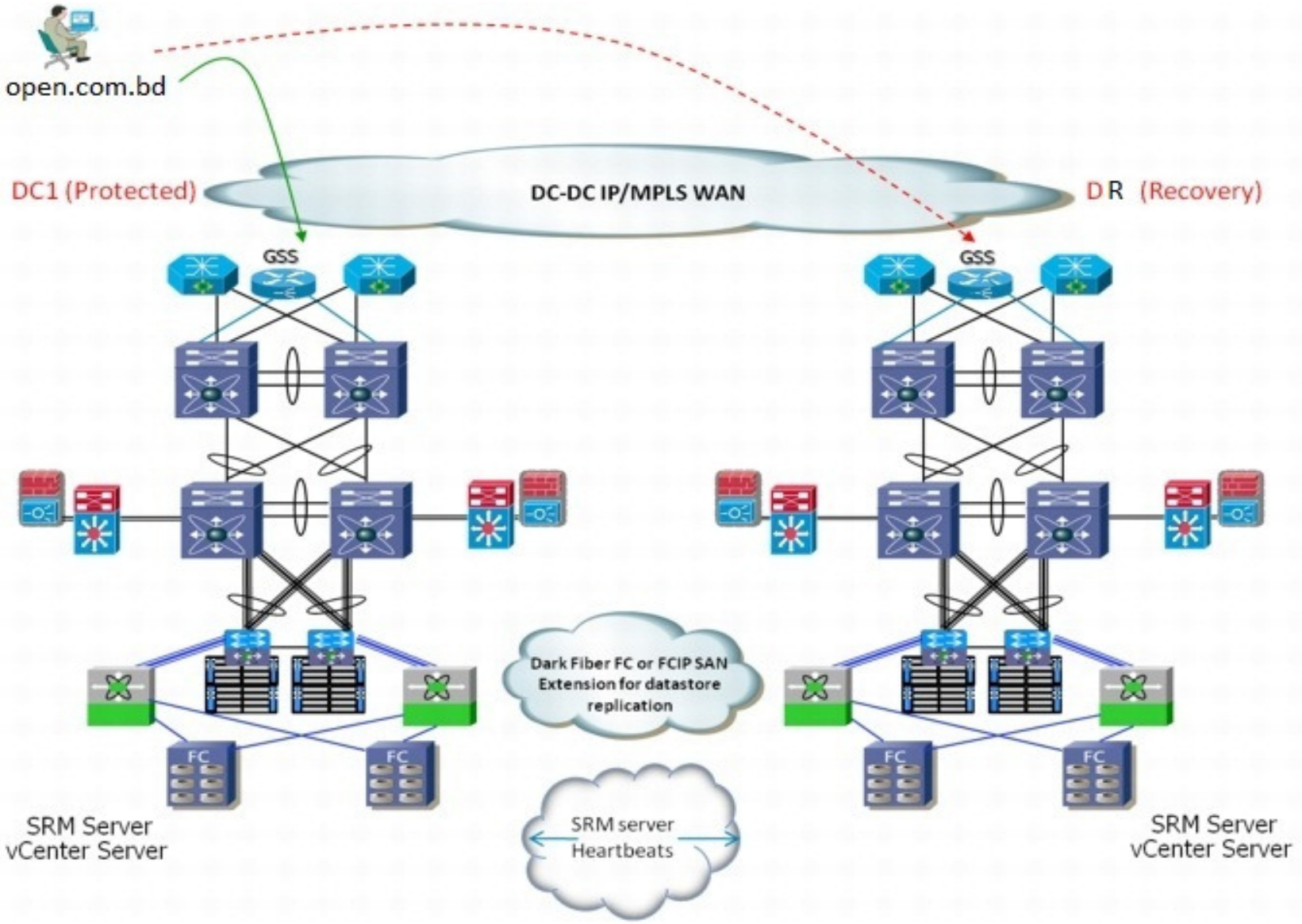


Case Study

Case Study (1), Data Center Design



Case Study (2),Data Center Design



Question ???