Gestión de Centro de Cómputos

Prof. Cristian Cabanellas

Prof. Gerardo Riveros

Facultad Politécnica, Universidad Nacional de

AGENDA

- ***Clusters***
- **Tecnologías de *Storage***
- **Proceso de Diseño y Proyecto de un DC**

AGENDA

- ***Clusters***
- **Tecnologías de *Storage***
- **Proceso de Diseño y Proyecto de un DC**

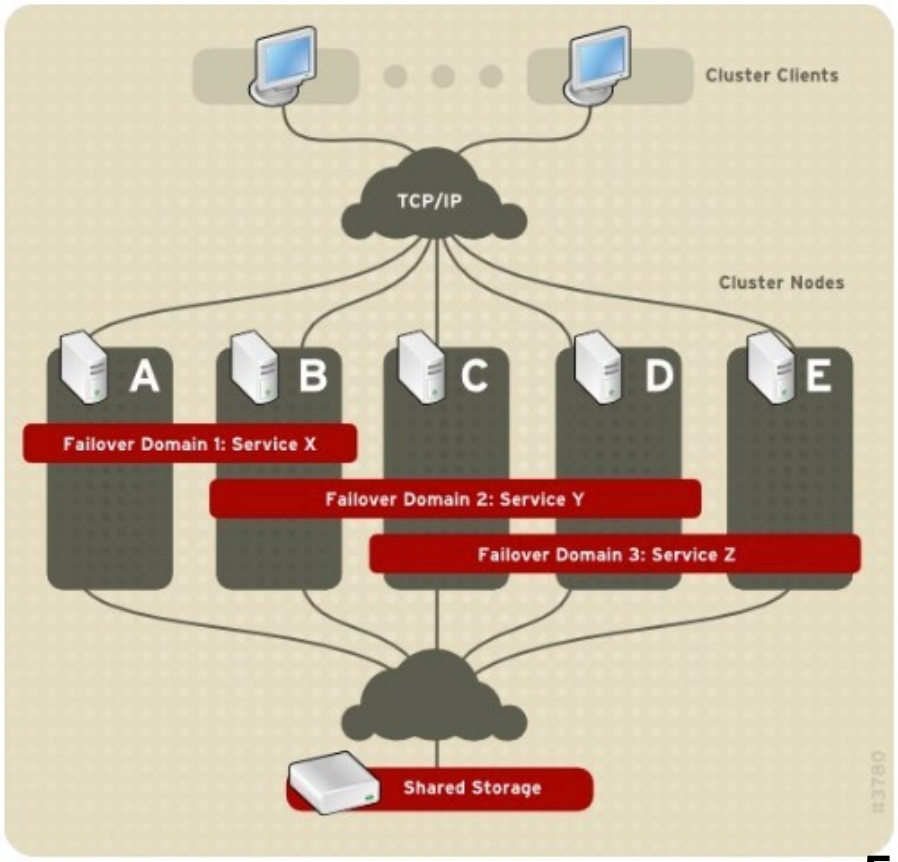
CLUSTERS

WHAT IS A CLUSTER?

Is a set of computers working together on a single task. What task is performed, how is that task performed, can differ greatly from cluster to cluster.

Type of cluster:

- **Load Balancing Cluster**
- **Compute Cluster**
- **High-Availability Cluster**



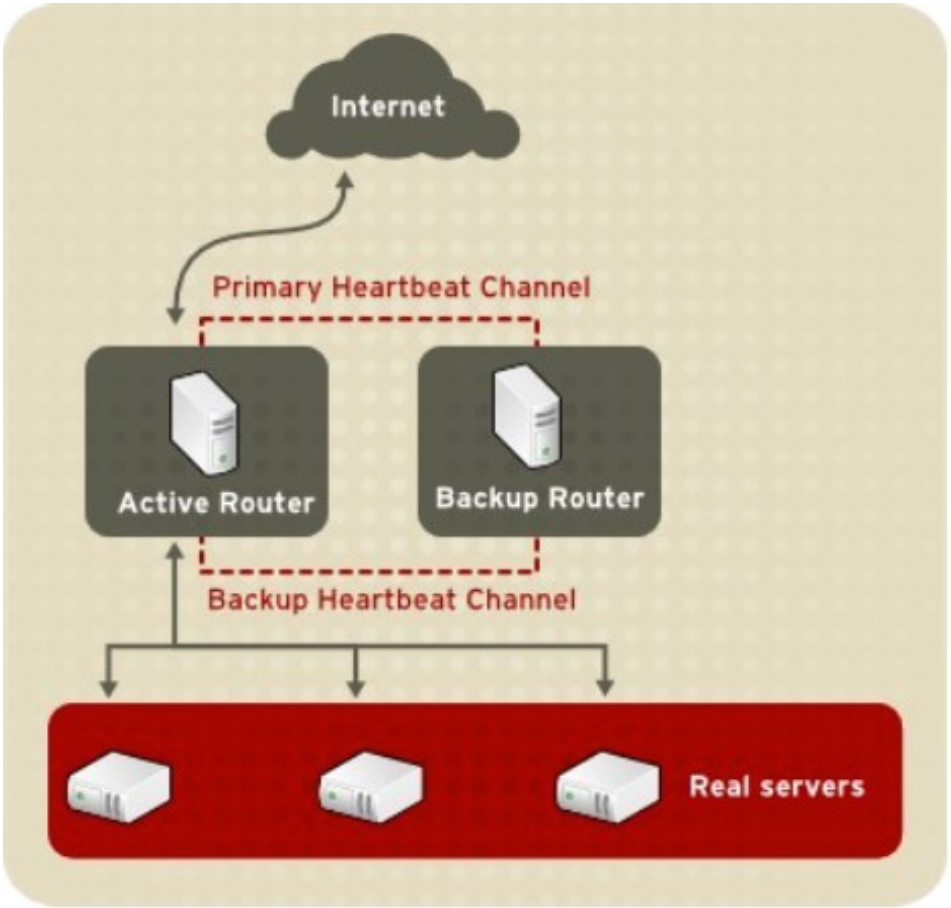
LOAD BALANCING CLUSTER

In a load-balancing cluster all members (nodes) run the same software and performed the same task at the same time. Requests from clients are distributed over the nodes, either a dedicated load-balancer, or basic by employing techniques such a round-robin DNS.

Examples:

SMTP setups with multiple incoming SMTP servers and multiples MX records, and web hosting farms.

LOAD BALANCING: LINUX VIRTUAL SERVER (LVS)



Fuente: Red Hat Cluster

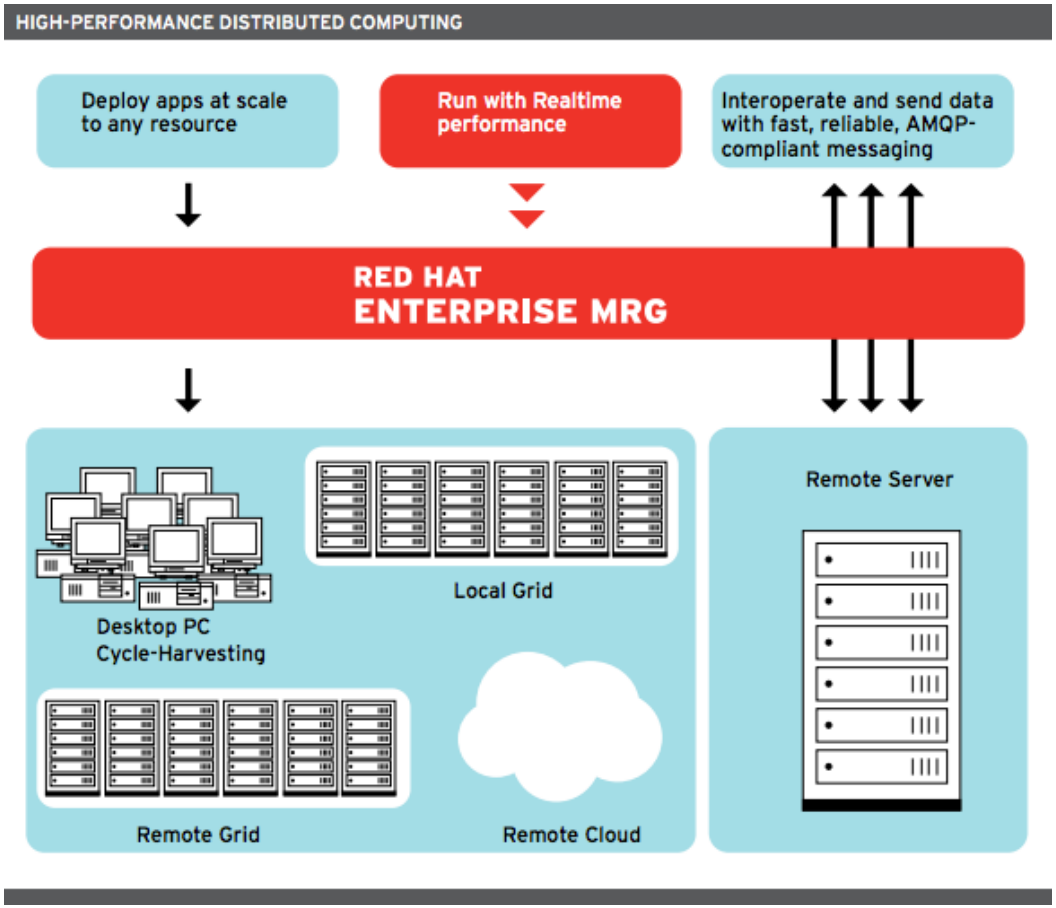
COMPUTE CLUSTER

In a compute cluster, also knows HPC (High Performance Computing) cluster, tasks as divided into smaller chunks with the get computed on different machines. This form of computing is sometimes also referred to as grid computing.

Examples:

- **Red Hat Enterprise MRG**
- **Condor Project**

RED HAT ENTERPRISE MRG



Fuente: Red Hat Cluster

HIGH-AVAILABILITY CLUSTER

The goal of a high-availability cluster (also knows HA cluster, or failover cluster) is to keep running the services as available as they can be. This is primarily achieved by having the nodes of the high availability cluster monitor each other for failures, and migrating services to a node that is still considered “healthy” when service or node fails.

HA cluster can be grouped in to tow subsets:

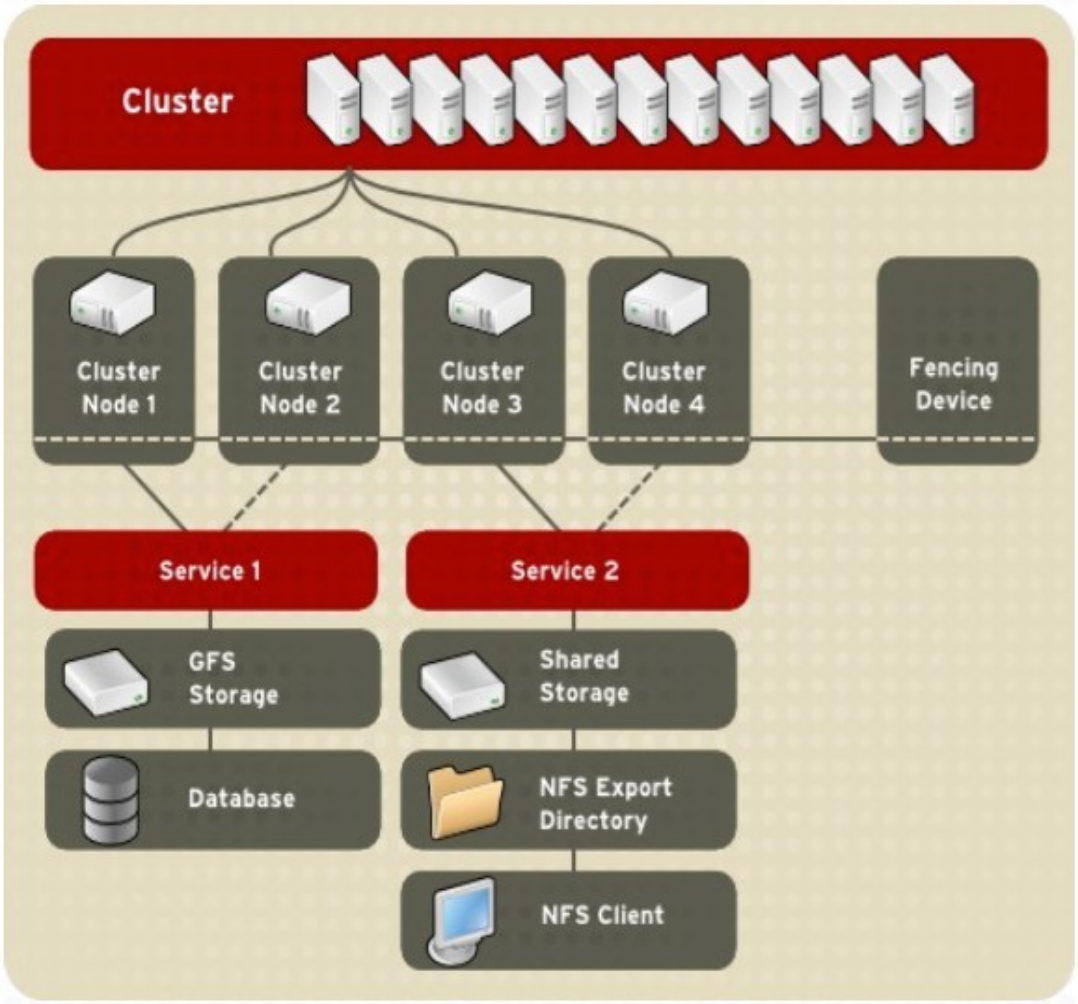
- **Active-Active**
- **Active-**

Passive

Examples:

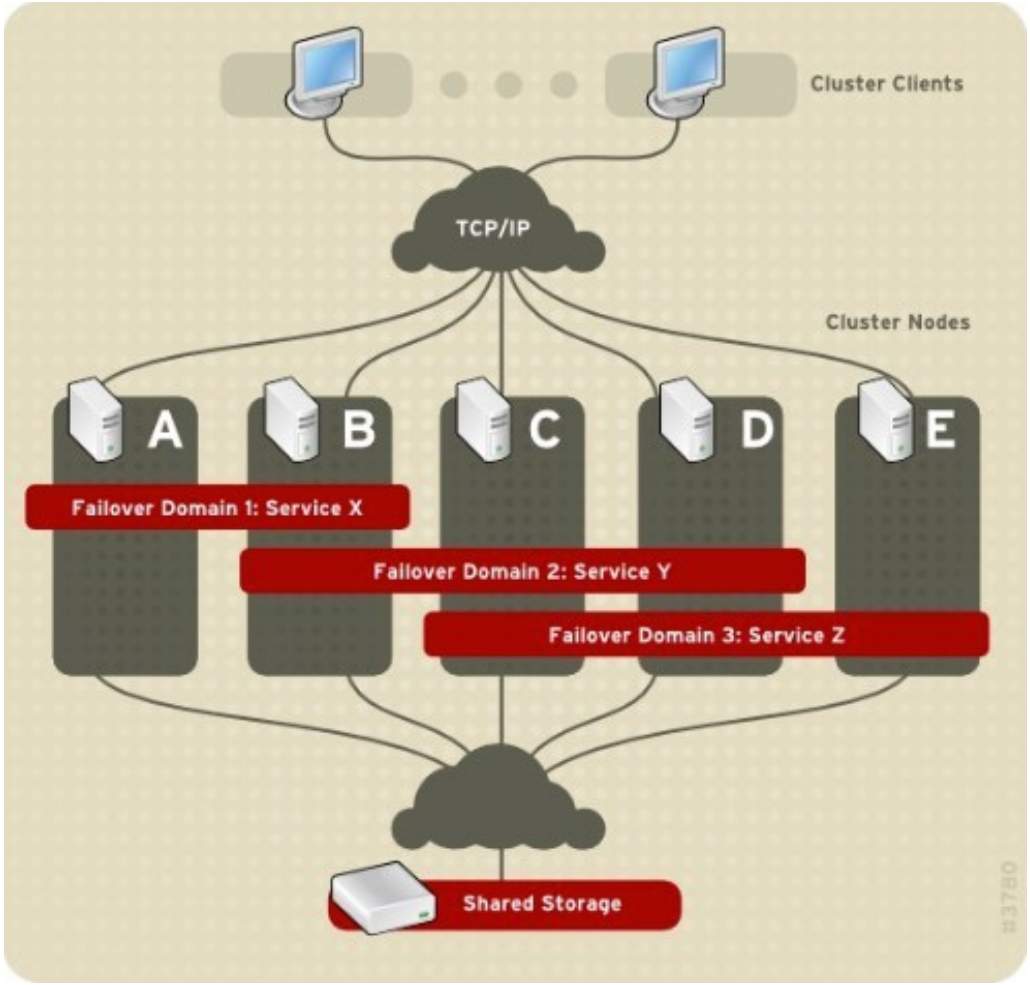
- Pacemaker
- Red Hat HA Add-on.

HA SERVICE



Fuente: Red Hat Cluster

FAILOVER DOMAINS



Fuente: Red Hat Cluster

WHEN TO USE HA-CLUSTERING?

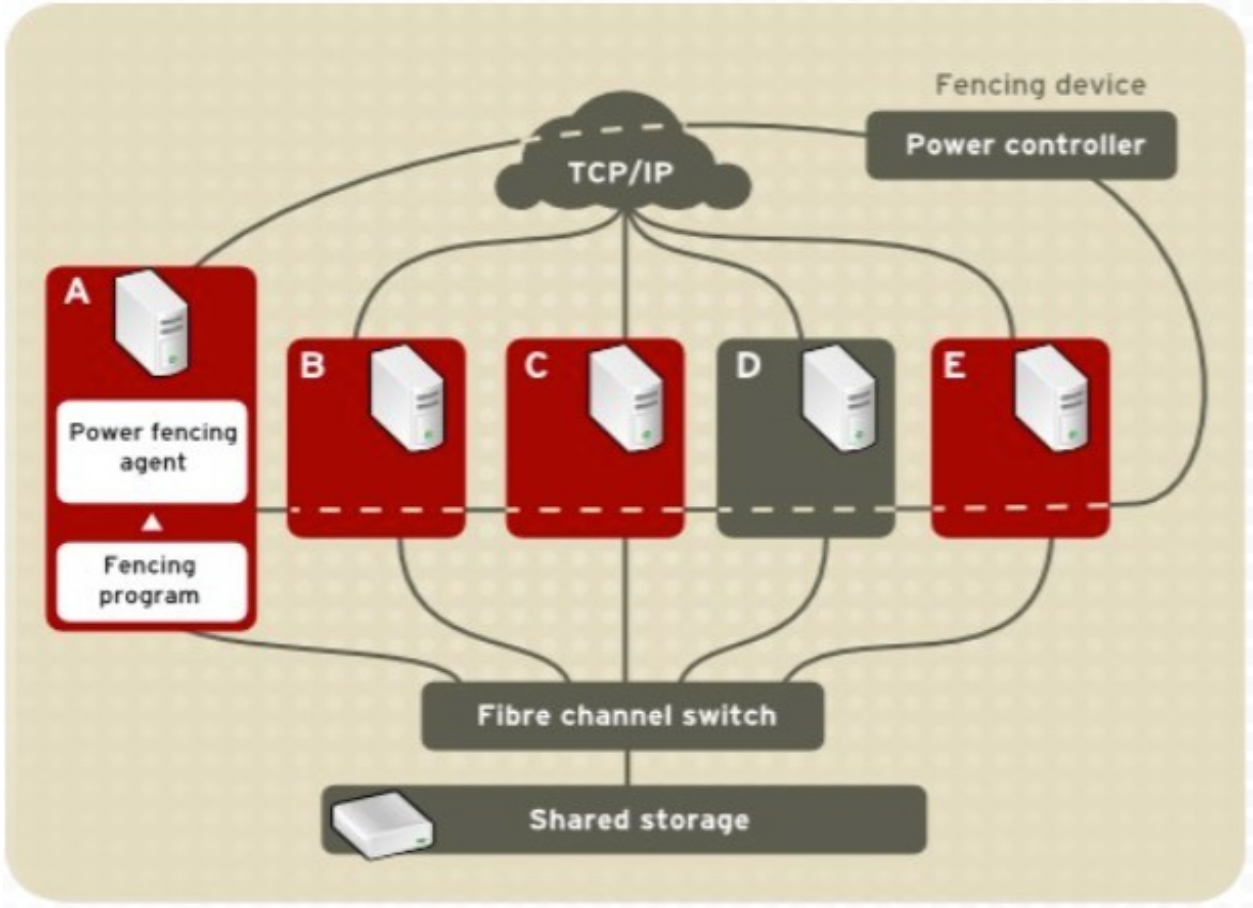
Before you start planning a HA cluster you must always ask yourself one important question: “Will availability of my service increase by putting in on a HA cluster?”

Examples:

- DNS services
- LDAP services

FENCING

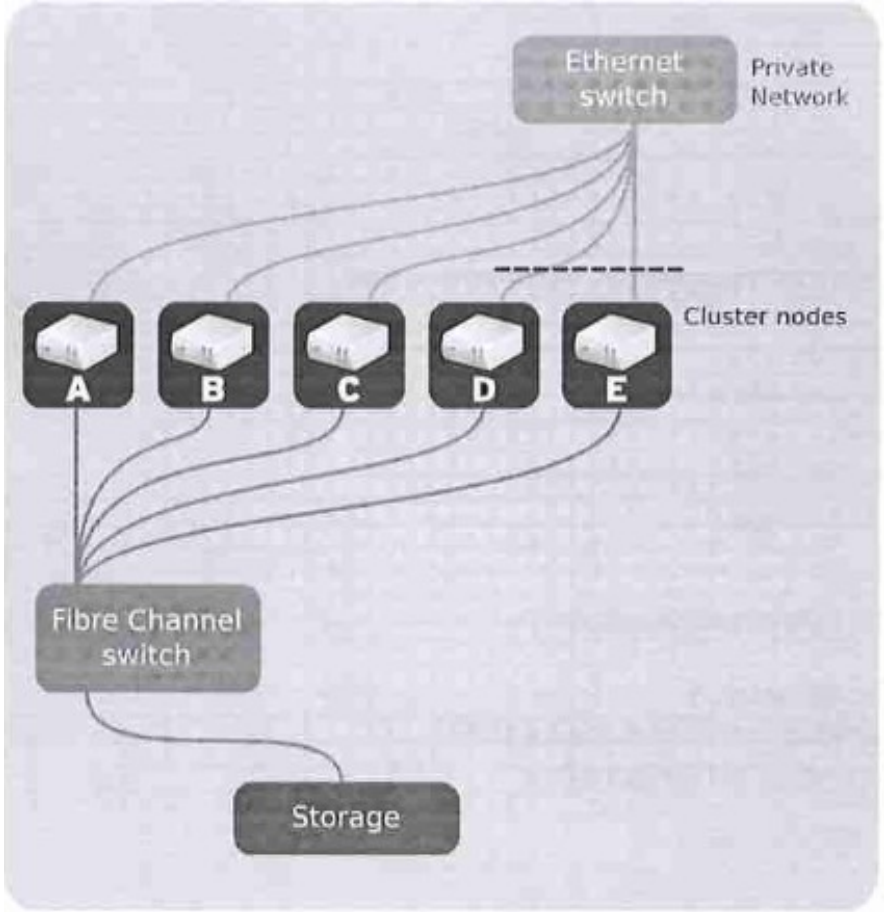
- Disconnection of a node from the cluster's shared storage
- Necessary to prevent file system corruption
- Methods:
  - Power fencing
  - Fabric fencing
  - Combining both



Fuente: Red Hat Cluster

QUORUM

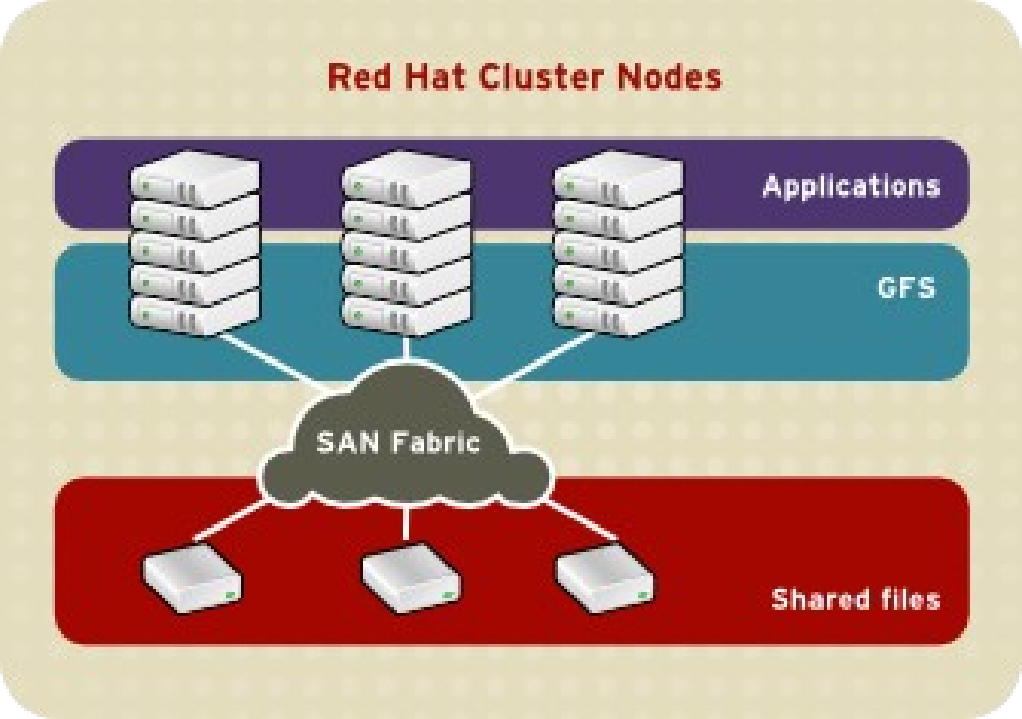
- majority voting scheme
- Each node gets to cast one or more votes
- Only is more than half of all possible
  votes are successfully cast will the cluster be operational.
- The minimum number of votes needed to achieve more than half of the votes is called the quorum



Fuente: Red Hat Cluster

RED HAT GLOBAL FILE SYSTEM (GFS)

- GFS/GFS2 is a native file system
- GFS/GFS2 provides data sharing among nodes with
  single, consistent view of the file system name space
- GFS/GFS2 file system must be created on an LVM logical volume (linear or mirror)



HOMEWORK, READ ABOUT

- **FENCING**
- **QUORUM**
- **GFS**

AGENDA

- *Clusters*
- **Tecnologías de *Storage***
- **Proceso de Diseño y Proyecto de un DC**

DIFFERENT TYPE OF DATA

When you take a closer look at your data you can probably distinguish multiple type of data. Among the attributes you can look at are:

- **System data vs. User data**
- **Static data vs. Dynamic data**
- **Current data vs. Achieved data**
- **Easy-to-replace data vs. Hard-to-replace data**
- **Often access data vs. Rarely access data**
- **System specific data vs. Common data**
- **Etc.**

Looking at your data this way allows you analyze what type of storage you might wan to use.

STORAGE TECHNOLOGIES

To actually store data there are many technologies available, some faster than others, while others can be more reliable or cheaper. Among the different technologies, current and obsolescent, are the following:

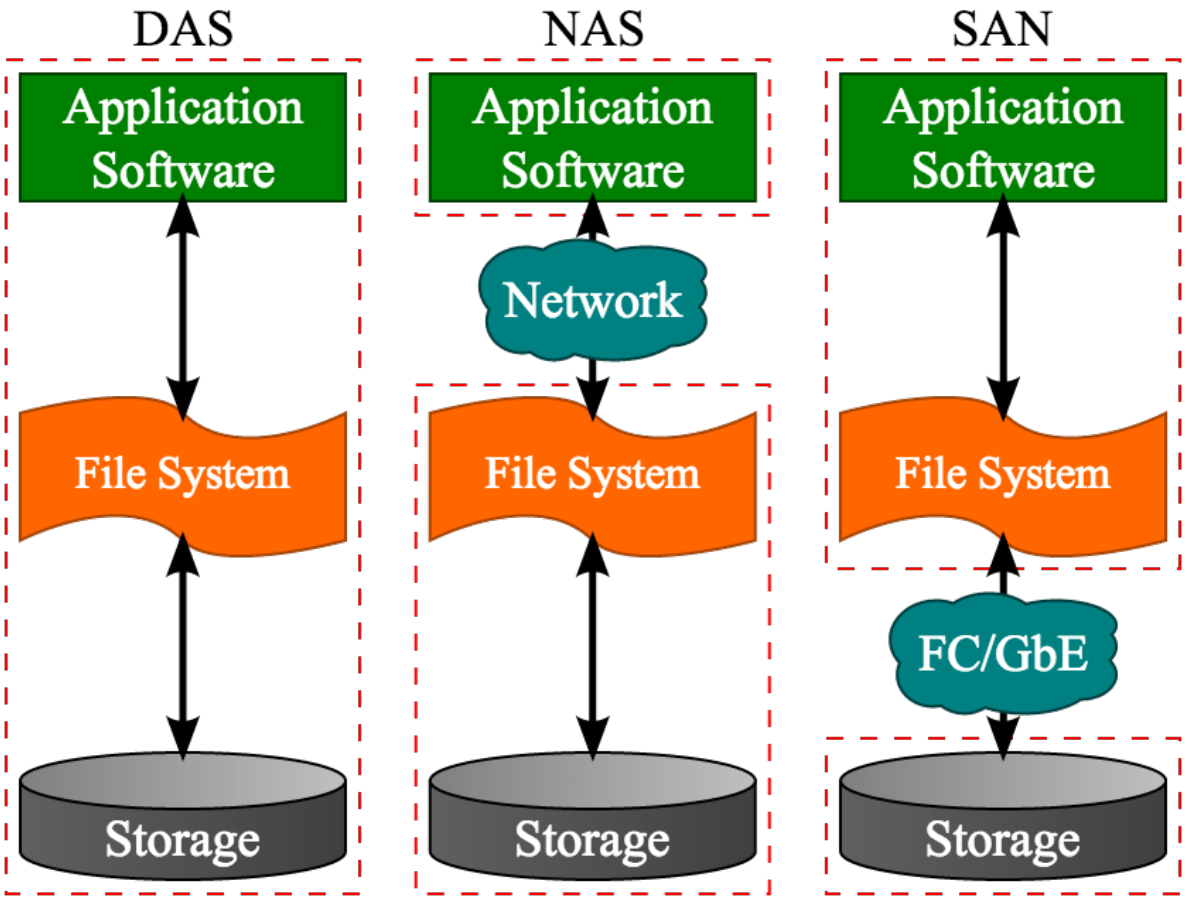
- **Paper**
- **Flash-base storage**
  - USB sticks
  - Flash card
  - Solid State Drives (SSDs)
- **Magnetic tape**
  - 9-track reel tapes
  - DDS (Digital Data Storage on DAT)
  - DLT (Digital Linear Tape)
  - LTO (Linear Tape-Open)

STORAGE TECHNOLOGIES

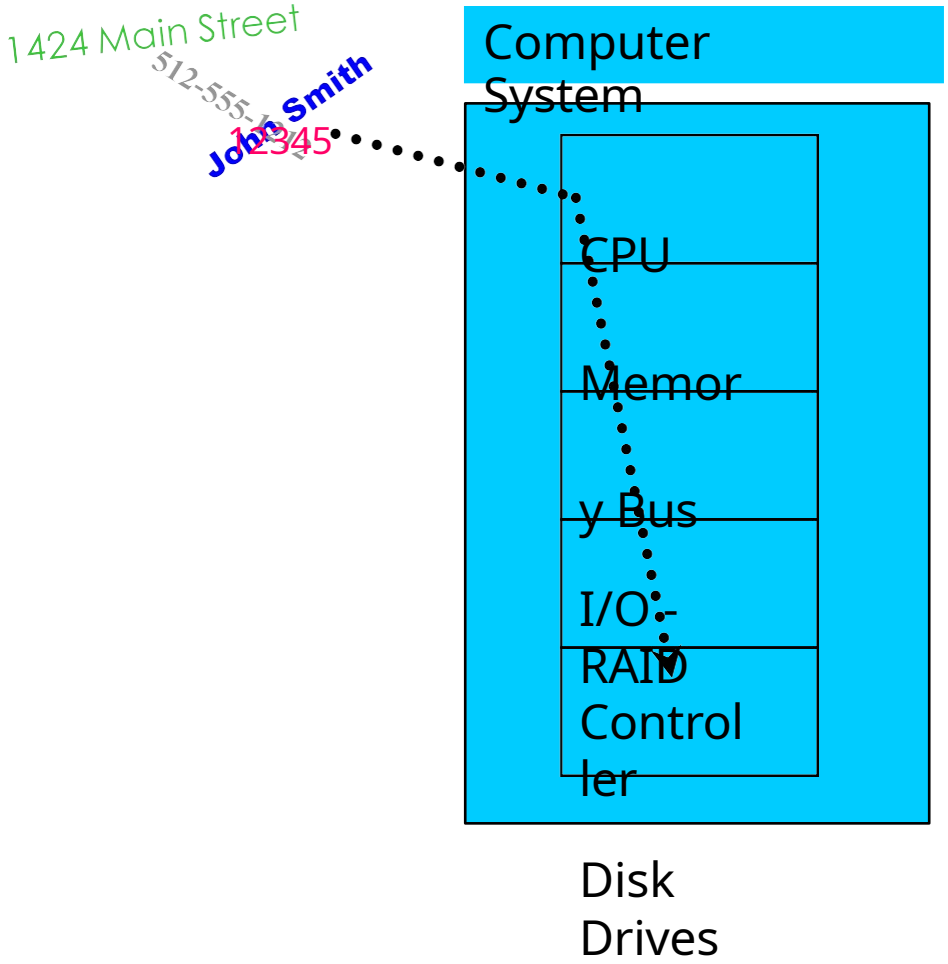
- **Rotating Magnetic Platters**
  - Hard Disk (using various interconnect)
- **Optical Storage**
  - CD-ROM, CD-R, CD-RW
  - DVD-ROM, DVD-R(W), DVD+R(W)
  - Blu-Ray discs: BD-ROM, BD-R, BD-RE

When you need to connect a data carrying medium to a computer there are again multiple options, both for the physical connection and protocol used to talk to the storage hardware, although these are mostly tied together.

STORAGE TYPES

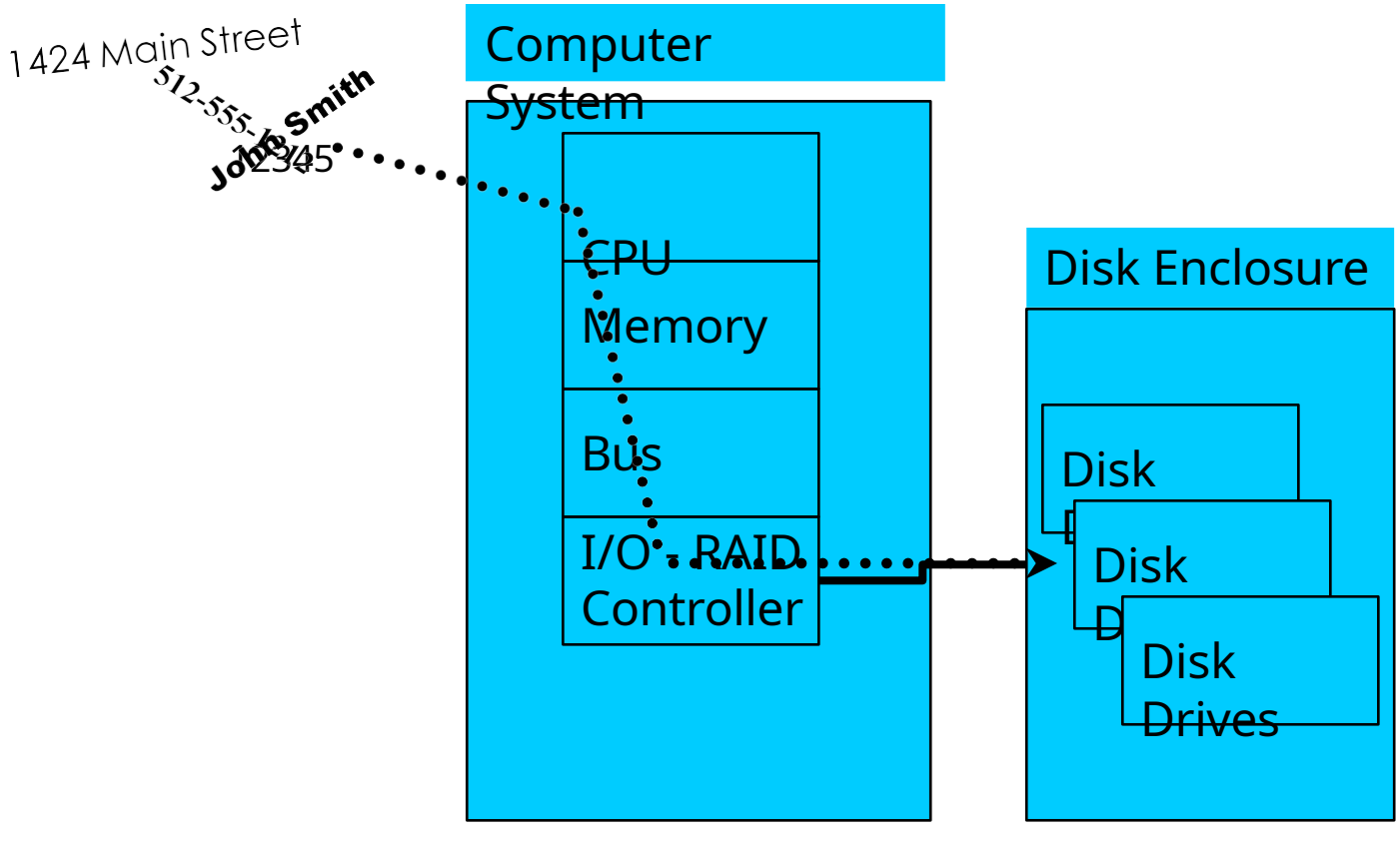


DIRECT ATTACHED STORAGE (INTERNAL)



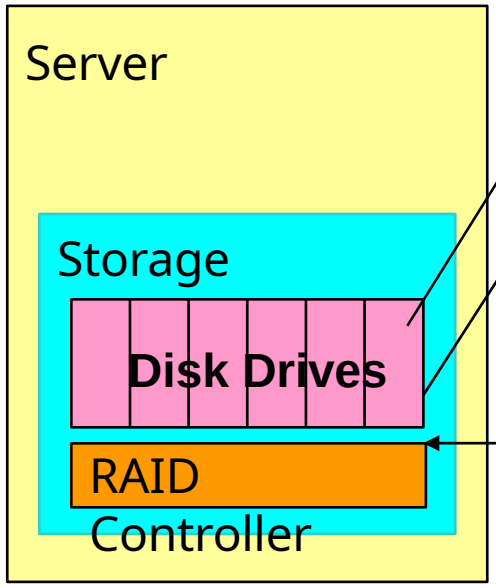
Fuente: Intro to SAN,

DAS W/ INTERNAL CONTROLLER AND EXTERNAL STORAGE



Fuente: Intro to SAN,

COMPARING INTERNAL AND EXTERNAL STORAGE



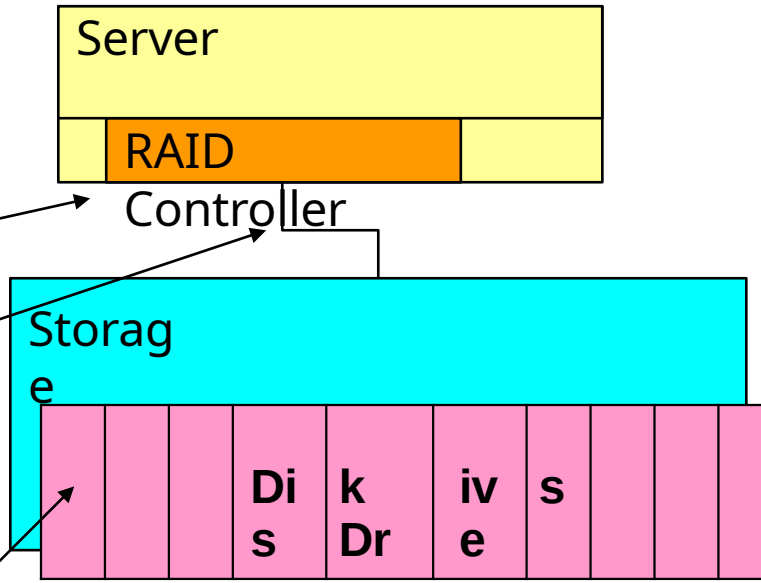
Internal Storage

RAID controllers and disk drives are internal to the server

SCSI, ATA, or SATA protocol between controller and disks

RAID controller is internal

SCSI or SATA protocol between controller and disks

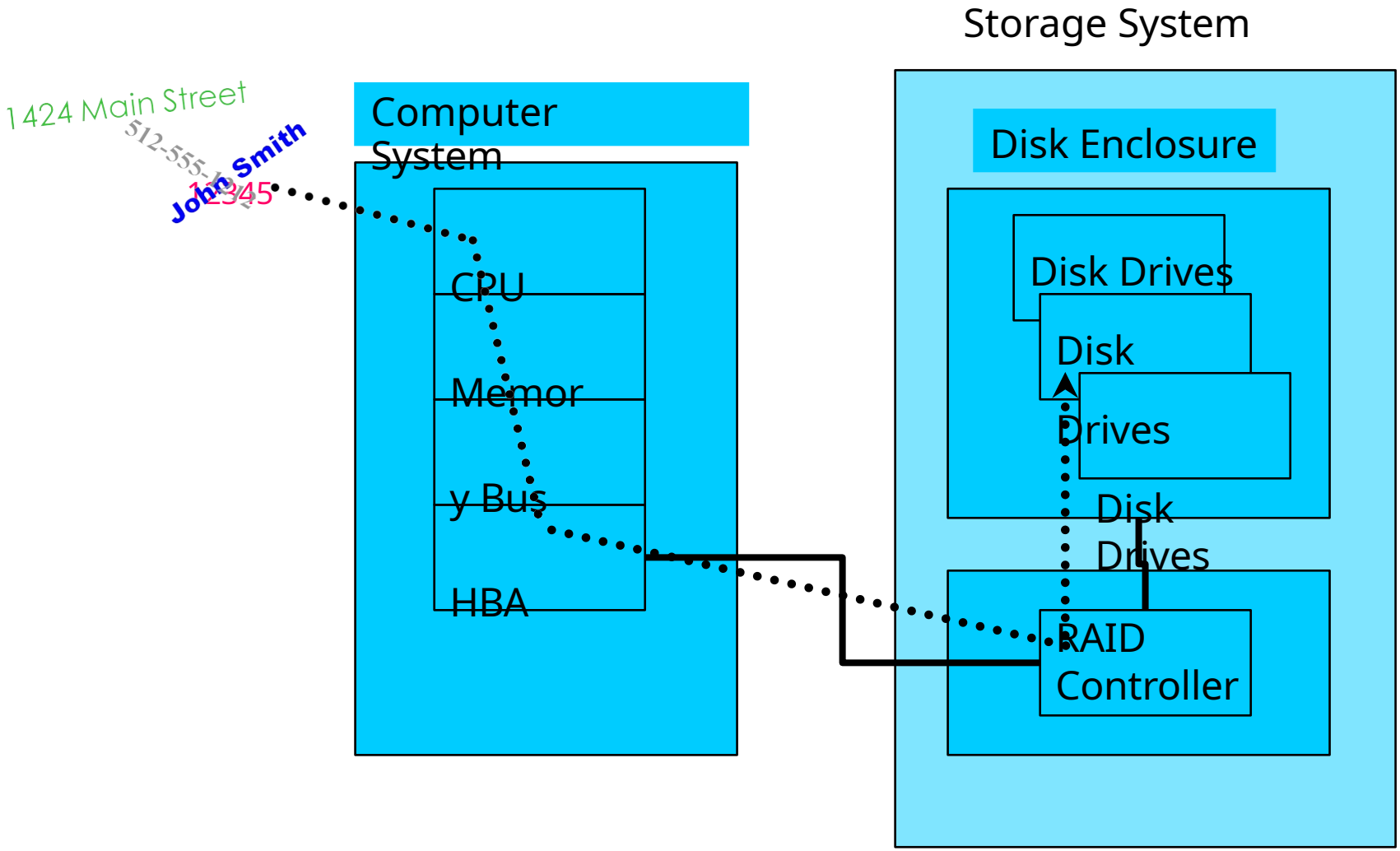


SCSI Bus w/ external storage

externa

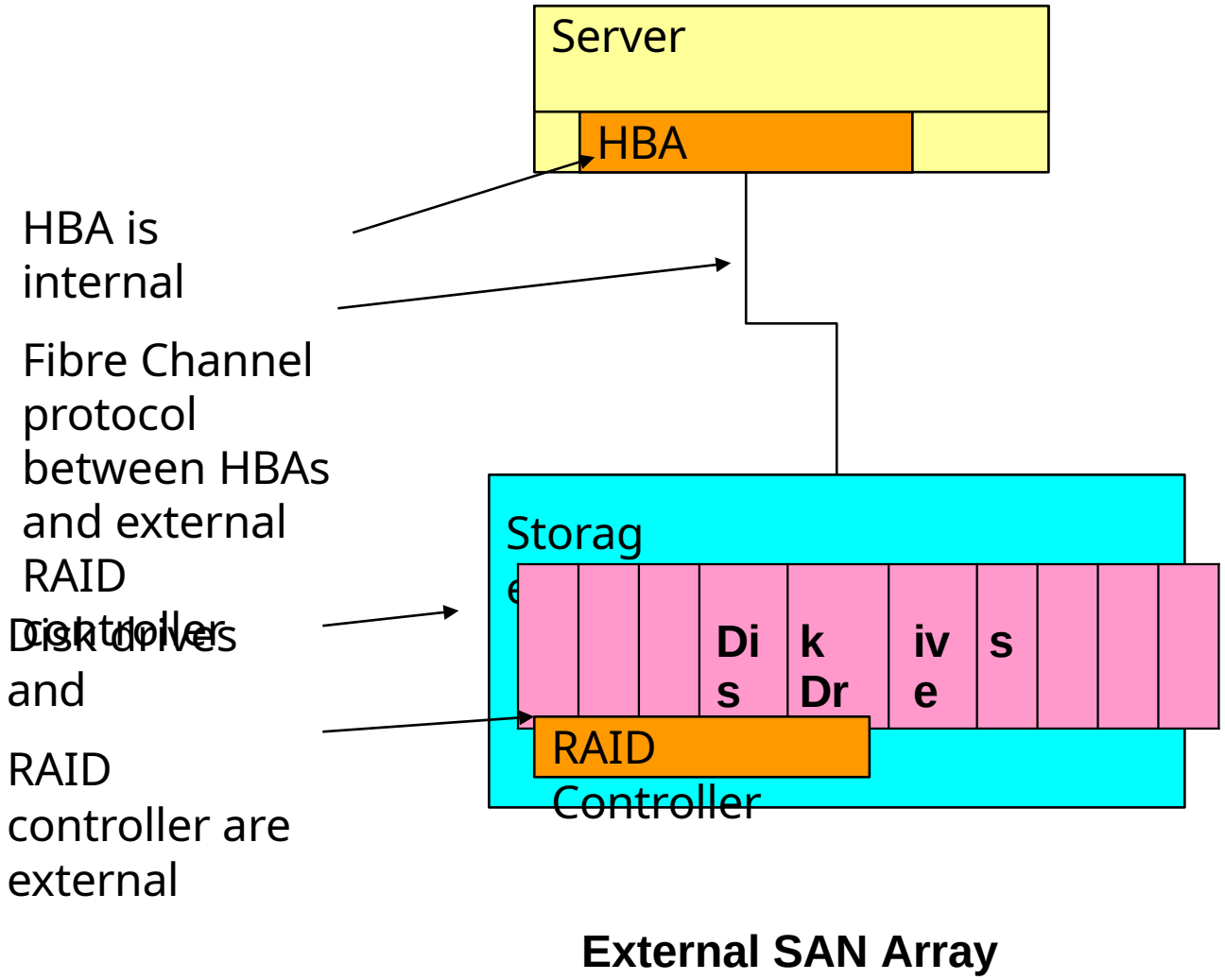
Fuente: Intro to SAN,

DAS W/ EXTERNAL CONTROLLER AND EXTERNAL STORAGE



Fuente: Intro to SAN,

DAS OVER FIBER CHANNEL



Fuente: Intro to SAN,

I/O TRANSFER

RAID Controller

- Contains the “smarts”
- Determines how the data will be written (striping, mirroring, RAID 10, RAID 5, etc.)

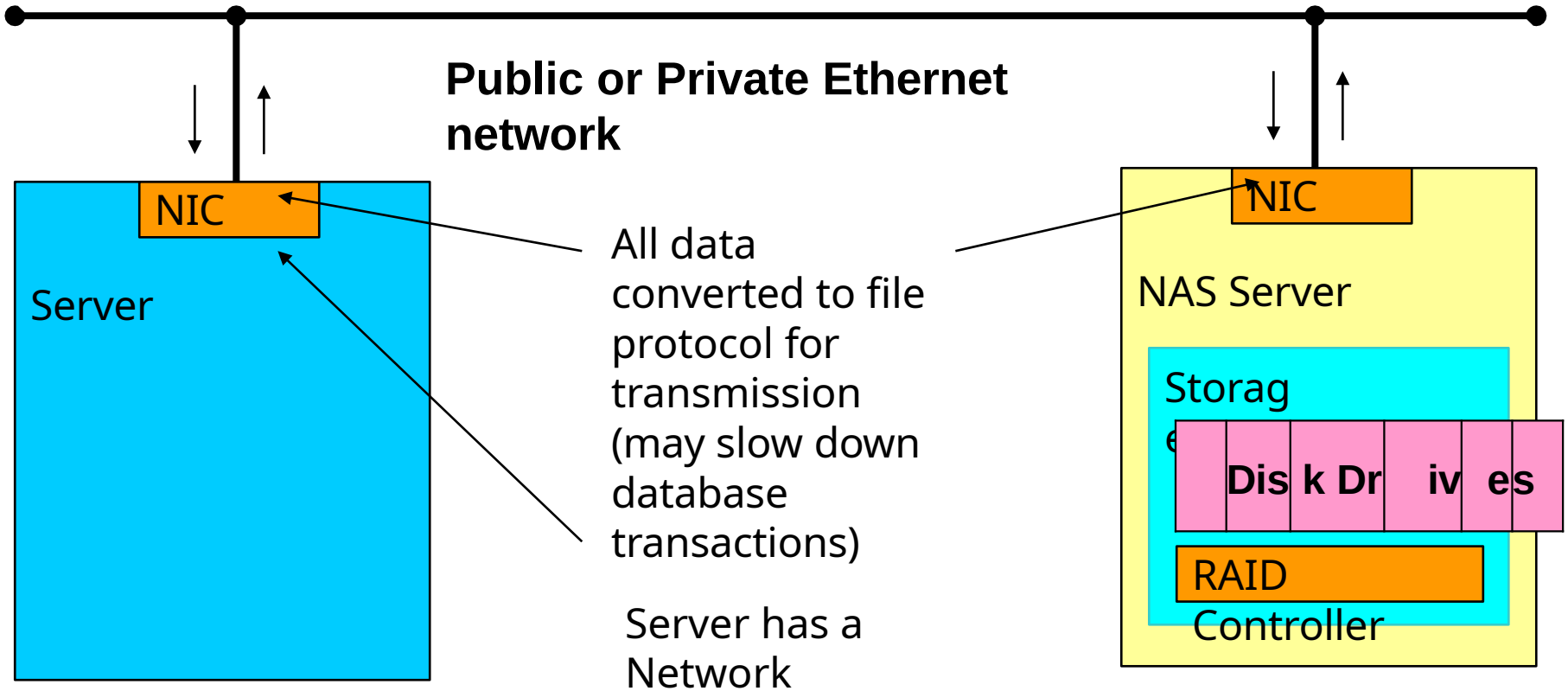
Host Bus Adapter (HBA)

- Simply transfers the data to the RAID controller.
- Doesn't do any RAID or striping calculations.
- “Dumb” for speed.
- Required for external storage.

NAS: WHAT IS IT?

- **Network Attached Storage**
- **Utilizes a TCP/IP network to “share” data**
- **Uses file sharing protocols like Unix NFS and Windows CIFS**
- **Storage “Appliances” utilize a stripped-down OS that optimizes file protocol performance**

NETWORKED ATTACHED STORAGE



All data converted to file protocol for transmission (may slow down database transactions)

Server has a Network Interface Card

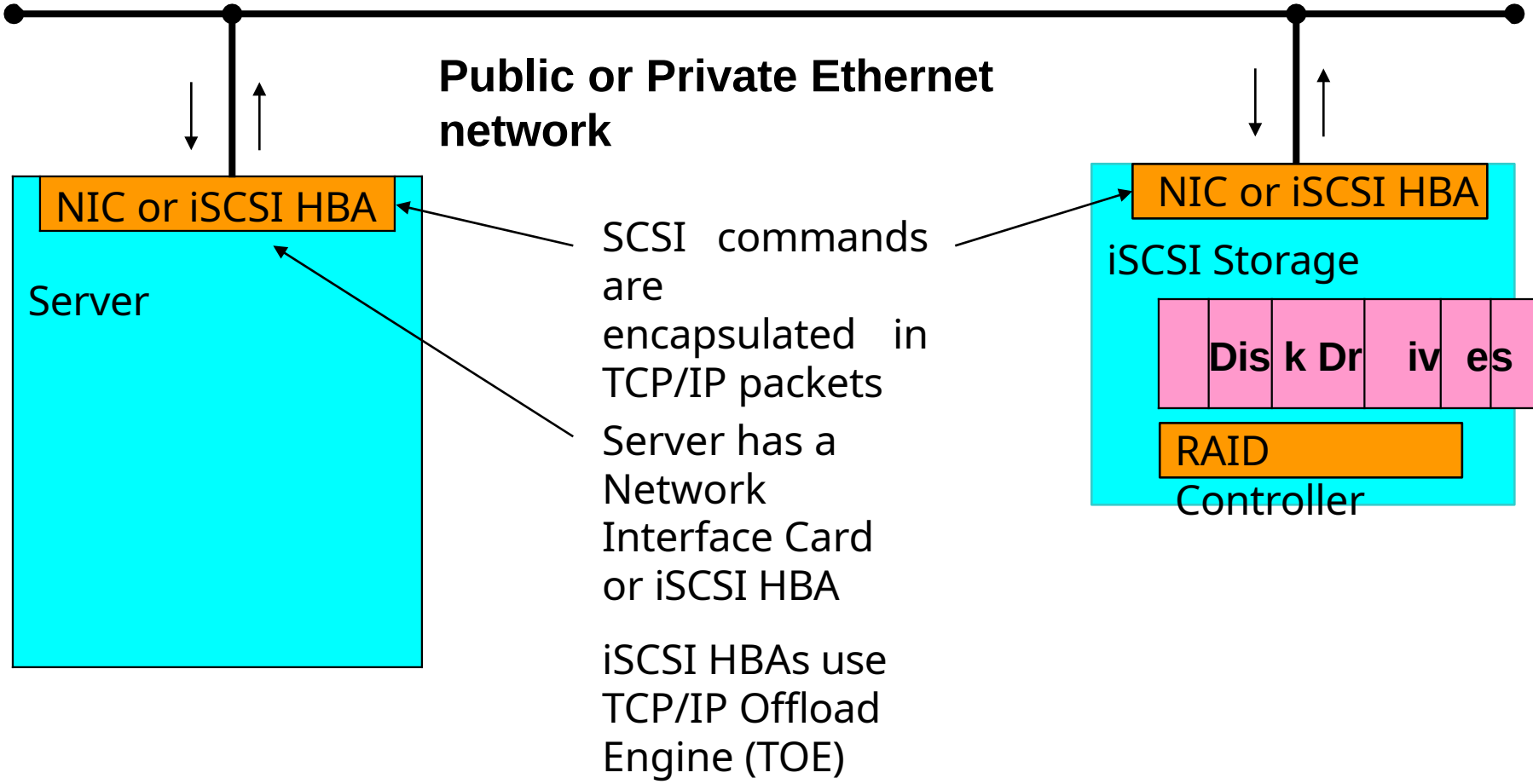
No RAID Controller or HBA in the server

Fuente: Intro to SAN,

ISCSI: WHAT IS IT?

- **An alternate form of networked storage**
- **Like NAS, also utilizes a TCP/IP network**
- **Encapsulates native SCSI commands in TCP/IP packets**
- **Supported in Windows 2003 Server and Linux**
- **TCP/IP Offload Engines (TOEs) on NICs speed up packet encapsulation**

ISCSI STORAGE

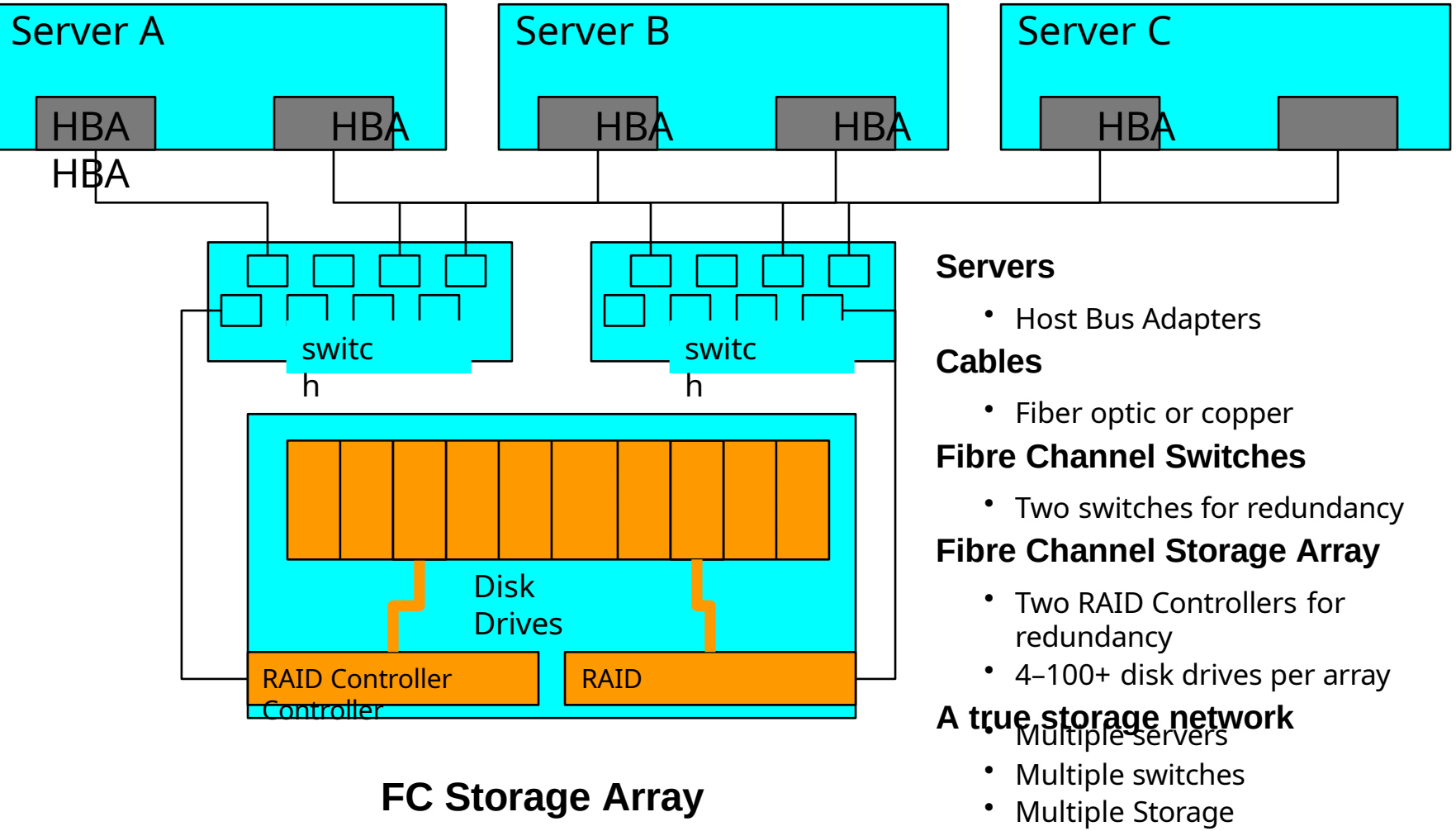


Fuente: Intro to SAN,

FIBER CHANNEL: WHAT IS IT?

- **Fiber Channel is a network protocol implemented specifically for dedicated storage networks**
- **Fiber Channel utilizes specialized**
  - Switches
  - Host Bus Adapters
  - RAID controllers
  - Cables

FIBER CHANNEL COMPONENTS



FC Storage Array

Servers

- Host Bus Adapters

Cables

- Fiber optic or copper

Fibre Channel Switches

- Two switches for redundancy

Fibre Channel Storage Array

- Two RAID Controllers for redundancy
- 4–100+ disk drives per array

A true storage network

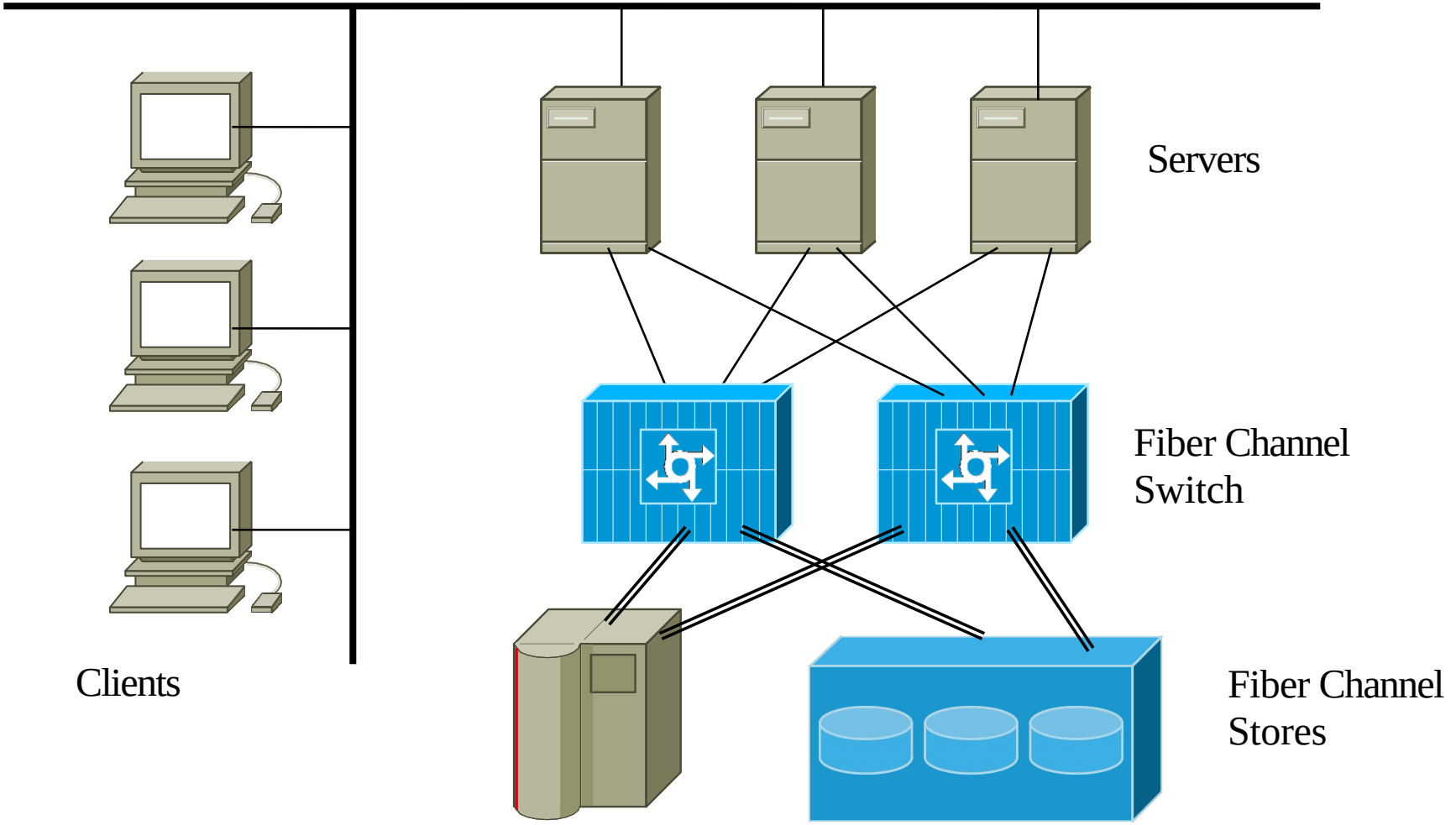
- Multiple servers
- Multiple switches
- Multiple Storage Arrays

Fuente: Intro to SAN,

SAN: WHAT IS IT?

- **Storage Area Network**
- **A network whose primary purpose is the transfer of data between storage systems and computer systems**
- **Fiber Channel is the primary technology utilized for SANs**
- **Recently, SANs have been implemented with dedicated iSCSI networks**

FC - STORAGE AREA NETWORK (REDUNDANT ARCHITECTURE)

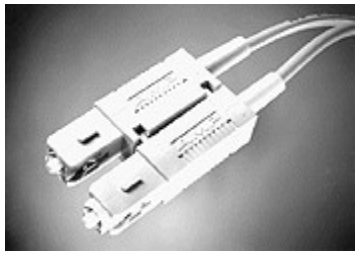


BENEFITS OF SAN/CONSOLIDATED STORAGE

- **Reduce cost of external storage**
- **Increase performance**
- **Centralized and improved tape backup**
- **LAN-less backup**
- **High-speed, no single-point-of-failure clustering solutions**
- **Consolidation with > 70TB of storage**

FIBER CHANNEL TECHNOLOGY

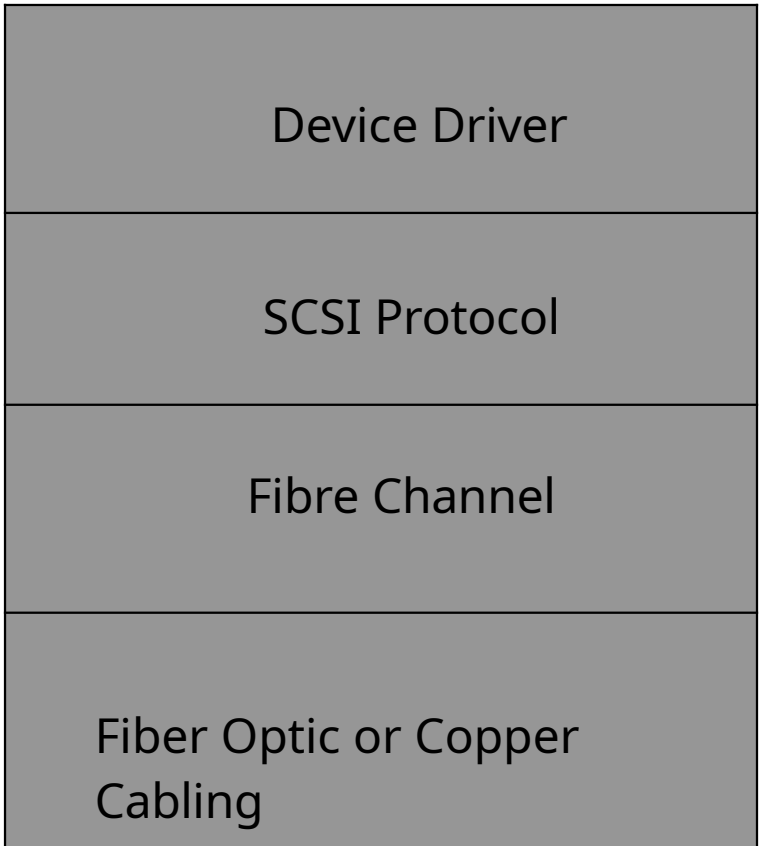
- **Provides concurrent communications between servers, storage devices, and other peripherals**
- **A gigabit interconnect technology**
- **FC1: Over 1,000,000,000 bits per second**
- **FC2: Over 2,000,000,000 bits per second**
- **A highly reliable interconnect**
- **Up to 127 devices (SCSI: 15)**
- **Up to 10 km of cabling (3-15 ft. for SCSI)**
- **Physical interconnect can be copper or fiber optic**



FIBER CHANNEL TECHNOLOGY

- **Hot-pluggable - Devices can be removed or added at will with no ill effects to data communications**
- **Provides a data link layer above the physical interconnect, analogous to Ethernet**
- **Sophisticated error detection at the frame level**
- **Data is checked and resent if necessary**

FIBER CHANNEL INTERFACE LAYERS



SCSI VS. FIBER CHANNEL PROTOCOL

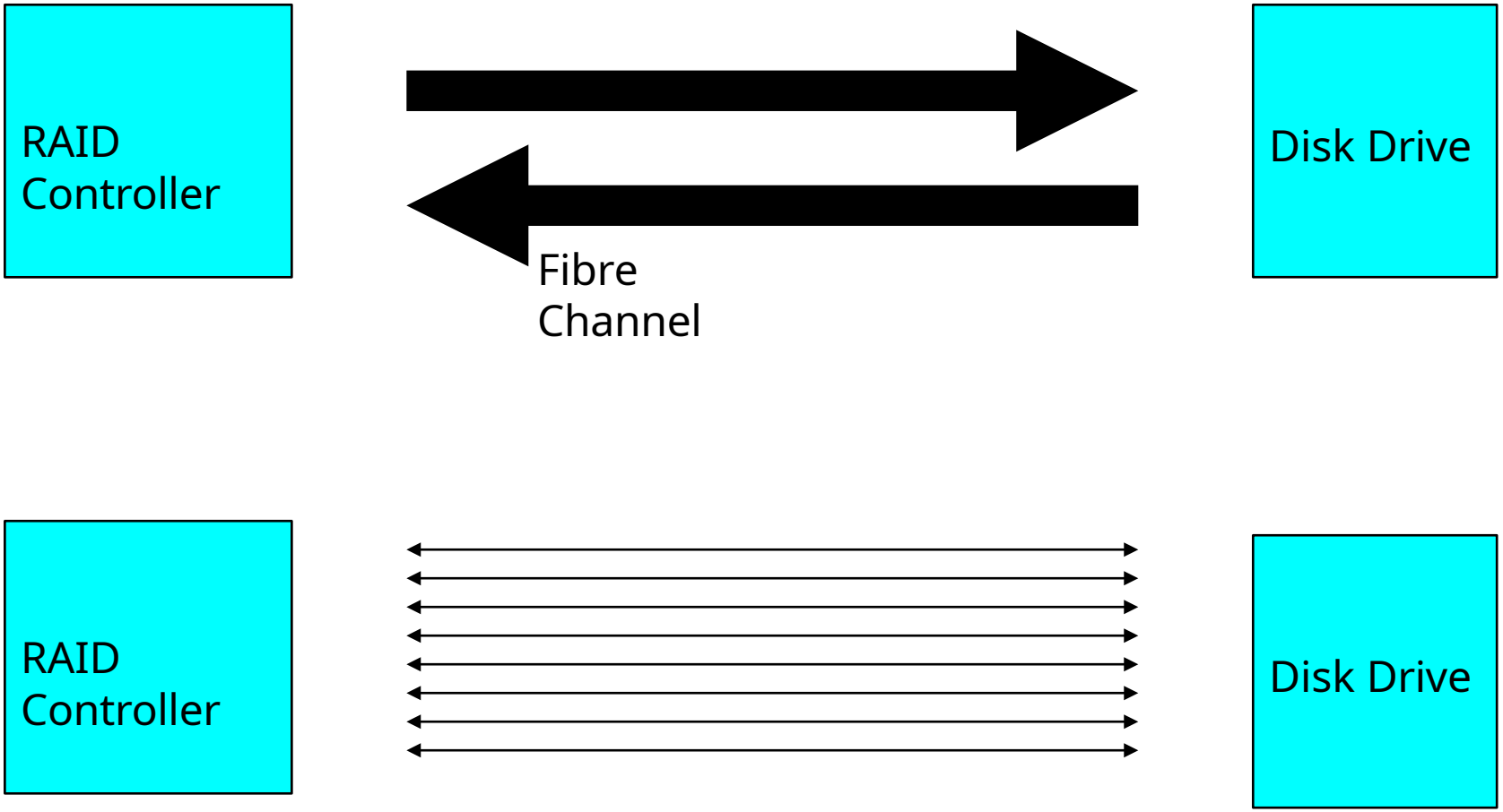
SCSI

- SCSI protocol vs. SCSI device
- SCSI is an established, tried and true protocol
- Provides services analogous to TCP/IP
- Supported in every major OS on market

Fiber Channel

- Fiber Channel runs on top of SCSI
- No re-inventing the wheel
- Immediate OS support

SCSI VS. FC TRANSMISSION

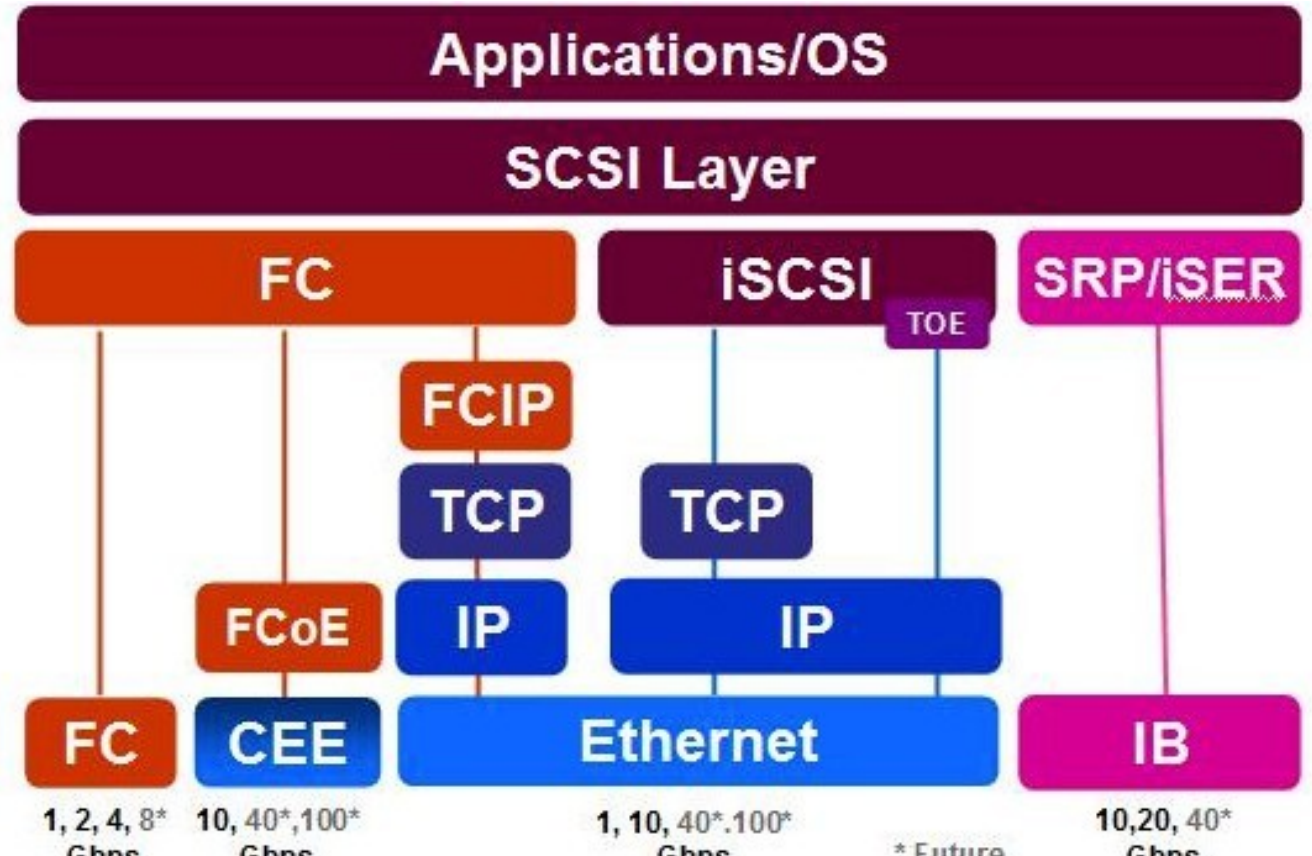


CHOOSING

SCSI VS. FIBER CHANNEL

Fibre Channel	iSCSI
Designed for enterprise markets	Designed for SMB markets
High bandwidth of 1 Gbps, 2 Gbps, 4 Gbps	Low bandwidth of 1 GigabitEthernet, 10 Gigabit-Ethernet
Low latency of 2 msec per port	High latency; IP has msec latency
Large payload of up to 2112 bytes	Smaller payload of up to 1500 bytes; up to 9000 bytes using jumbo frames
Low overheads of 5.5% for 1KB payload, 3% for 2KB payload	Higher overheads of 8% for 1KB payload
Short distance of 10 km per link for single mode fiber	Long distance; no theoretical limit over IP networks, but high latency
High cost HBAs and switch ports	Low cost: use existing NiC and LAN; iSCSI HBAs are expensive

FCOE VS. FC VS. ISCSI VS. IB



Fuente: http://www.imexresearch.com/newsletters/fcoe.html

FIBER CHANNEL VS. ISCSI

Fiber Channel

- The current market leader for shared storage technologies
- Provides the highest performance levels
- Designed for mission-critical applications
- Cost of components is relatively high, particularly per server HBA costs
- Relatively difficult to implement and manage

iSCSI

- Relatively new, but usage is increasing rapidly
- Performance can approach Fiber Channel speeds
- A better fit for databases than NAS
- A good fit for Small to Medium Size Businesses
- Relatively inexpensive, compared to Fiber Channel
- Relatively easy to implement and manage

QUICK OVERVIEW

Factor	iSCSI	SAS	FC	FCoE	IB-SRP	FCIP	iFCP
Best Suited Environment	SMB, ROBO	SMB, Enterprise	Enterprise	Enterprise	HPC	Enterprise	
Cost	**LOW**	**LOW**	HIGH	V.HIGH	V.HIGH	V.HIGH	HIGH
Performance	MEDIUM	HIGH	HIGH	HIGH	HIGH	LOW	LOW
TCP Required	YES	**NO**	**NO**	**NO**	**NO**	YES	YES
CPU Overhead	HIGH	LOW	LOW	LOW	LOW	HIGH	HIGH
Gateway Overhead	HIGH	LOW	LOW	LOW	HIGH	HIGH	HIGH
Security Robustness	MEDIUM	**LOW**	HIGH	HIGH	**LOW**	MEDIUM	LOW
FC SW on Host Needed	**NO**	**NO**	YES	YES	**NO**	**NO**	YES
Management Features	Low	HIGH	HIGH	HIGH	Medium	**NO**	YES
IP Routable	YES	**NO**	**NO**	**NO**	**NO**	YES	YES

Fuente: http://www.imexresearch.com/newsletters/fcoe.html

15 MINUTOS

AGENDA

- *Clusters*
- Tecnologías de *Storage*
- **Proceso de Diseño y Proyecto de un DC**

IMPACTO EN EL DISEÑO

- **Mayores niveles de seguridad**
- **Mayores niveles de resilencia/redundancia**
- **Mayor responsabilidad antes las autoridades normativas**
- **Requisitos de auditoría más exigentes**
- **Requisitos añadidos de almacenamiento**
- **Especificaciones en las coberturas de los seguros más exigentes**
- **Necesidad de una operación más eficiente**
- **Tiempos de procesamiento de las transacciones más rápidas**

TOMA DE DECISIONES

- **Antes de que se tome una decisión de diseño relativa a un centro de cómputos, debe tenerse en cuenta a todas las partes interesadas, entre las que se incluyen:**
  - La clase “C”
    - CEO, CFO, CIO, COO, CTO, etc.
  - Accionistas
  - Clientes
  - Personal
  - Proveedores

PROCESO DE DISEÑO Y PROYECTO

Fases del proceso de diseño y proyecto:

- **Instrucciones del cliente-establecimiento a alto nivel de las exigencias del cliente.**
  - Aumento de la capacidad TI a lo largo de la vida del activo
  - ¿Ajuste incremental de los espacios desocupados?
  - Requisitos de resilencia de componentes electromecánicos. Ej.: n+1, 2n, etc.
  - Requisitos de resilencia de telecomunicaciones
  - Objetivos en PUE

FASES PARA EL PROYECTO DE DC

- **Planificación**
- **Diseño**
- **Construcción**
- **Entrega**
- **Post-Construcción – Entrega al cliente**

FASE DE PLANIFICACIÓN

- **Se designan consultores**
- **Se toman decisiones cruciales**
- **Documentos de objetivos**
- **Plan de inversión**
- **Cronogramas**
- **Interacción con el cliente para determinar necesidades**
- **Financiación (aprobar)**
- **Autorizaciones**
- **Cronograma de implantación (entrega, recepción/instalación de equipos, etc.)**

FASE DE DISEÑO

- **Designar contratista principal**
- **Definir equipo de proyecto**
- **Sistemas a poner en marcha**
- **Revisión edilicia**
- **Requerimientos de IT**
- **Entrega de documento con objetivos del diseño**
- **Revisar alcance del proyecto**
- **Proveedores de servicios básicos**
- **Plan de puesta en marcha**
- **Aprobar infra IT y plan lógico de red**

FASE DE CONSTRUCCIÓN

- **Contratistas (electromecánicos, civiles, etc.)**
- **Consultores definen criterios de aceptación**
- **Informes de progreso**
- **Pruebas pre-funcionales**
- **Documentación**
- **Aprobación de modificaciones al diseño original**
- **Inicio de fase de puesta en funcionamiento**

FASE DE ENTREGA

- **Recepción de documentación de equipos (y equipos)**
- **Pruebas de aceptación de obras no IT (civiles, electro..)**
- **Pruebas de rendimiento funcional e integración**
- **Auditoría de equipamientos adquiridos**
- **Auditoría de garantías (del fabricante)**
- **Recepción de documentación final y resultados de las pruebas**

FASE DE POST-CONSTRUCCIÓN. ENTREGA AL CLIENTE

- **Definición de procedimientos**
  - Operativos Estándar (SOP)
  - Operativos de emergencia (EOP)
  - Procedimientos de Mantenimiento Planificados (PMP)
- **Instalación de equipamientos de IT (2 parte)**
- **Definición de políticas y procedimientos para:**
  - Traslado
  - Incorporación
  - Cambios
  - Control de Cambios

FUENTES Y LECTURAS ADICIONALES

- **Red Hat Enterprise Cluster and Storage Management**
- **Red Hat Cluster Suite for Red Hat Enterprise Linux 5**
- **Intro to SAN, Microsoft**
- **http://www.imexresearch.com/newsletters/fcoe.html**
- **Data Center Design Awareness [L1], DC Professional Development.**

THE END