



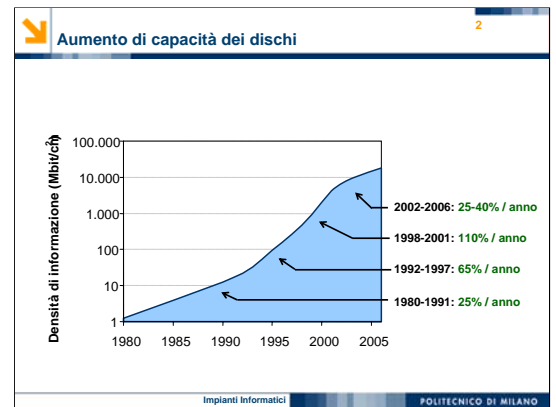
*Paolo Cremonesi*

# Impianti Informatici

 POLITECNICO DI MILANO



## Sistemi di storage



Se guardiamo più in dettaglio l'andamento della capacità dei dischi nel corso degli anni, vediamo che questo andamento non è stato regolare.

Osserviamo il grafico in figura, che mostra, sull'asse delle y, la densità dei dischi, espressa in Mbit per centimetro quadrato, in funzione dell'anno di produzione. Questa densità misura la capacità di un disco di memorizzare informazioni.

- Vediamo che, a partire dagli anni '80 la capacità media dei dischi è aumentata del 25% all'anno
- ma questo incremento è aumentato negli anni '90 passando prima al 65% all'anno
- per arrivare poi, verso il 2000, ad aumenti della capacità pari a più del 100% all'anno
- dopo il 2000 l'incremento della capacità dei dischi è ritornato ai valori iniziali del 25% all'anno

3

## Capacità vs. Prestazioni

### Crescita negli ultimi 30 anni

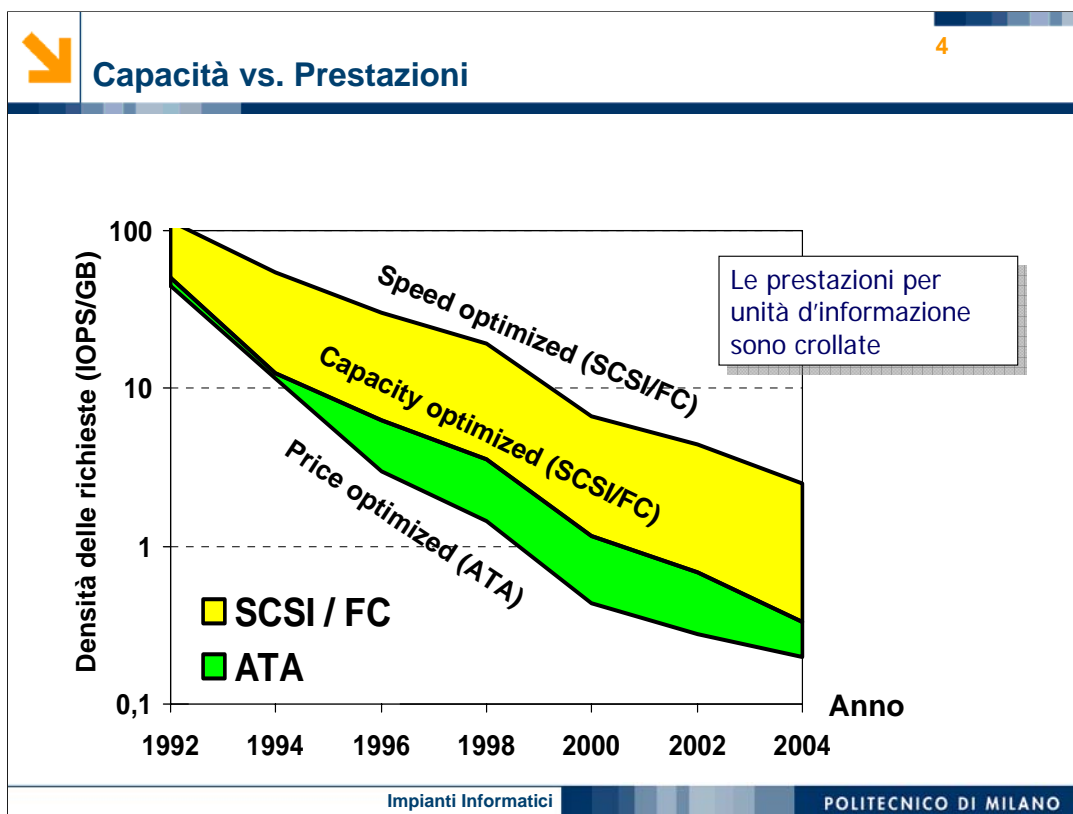
<i>Capacity:</i> Data per spindle	300 MB	300 GB	1000x
<i>Performance:</i> Random seek time	30 ms	3 ms	10x

Impianti Informatici

POLITECNICO DI MILANO

La capacità di memorizzare informazioni cresce molto più rapidamente rispetto alla velocità con cui si accede alle informazioni.

- In particolare, negli ultimi 30 anni, la capacità media dei dischi è aumentata di un fattore 1000, passando da 300 MB per disco a 300 GB per disco (si fa riferimento a dischi di fascia alta)
- Se osserviamo invece la velocità di accesso ai dischi, utilizzando come riferimento il seek time medio, notiamo che questa è aumentata solo di un fattore 10 nello stesso periodo di tempo



E' interessante a questo punto osservare come si è evoluto, nel tempo, il rapporto tra capacità e prestazioni di un disco. Il grafico mostra, sull'asse delle ordinate, il rapporto tra il numero medio di operazioni al secondo che un disco è in grado di eseguire (indicato con IOPS) e la capacità media di un disco espressa in gigabyte. L'evoluzione del tempo di questa grandezza mostra una costante riduzione delle prestazioni per unità di spazio.

Attenzione, questo non vuol dire che le prestazioni diminuiscono, ma che le prestazioni aumentano più lentamente della capacità.

- Per dischi di fascia bassa, che sono ottimizzati per avere un costo contenuto,
- e per dischi di fascia media, che sono ottimizzati per avere capacità elevata, questa tendenza è più marcata: il rapporto velocità su capacità è peggiorato di un fattore 100 in 10 anni
- Per dischi di fascia alta, che sono ottimizzati per avere prestazioni elevate, questa tendenza è sempre presente anche se meno marcata: il rapporto velocità su capacità è peggiorato di un fattore 20 in 10 anni

*Paolo Cremonesi*

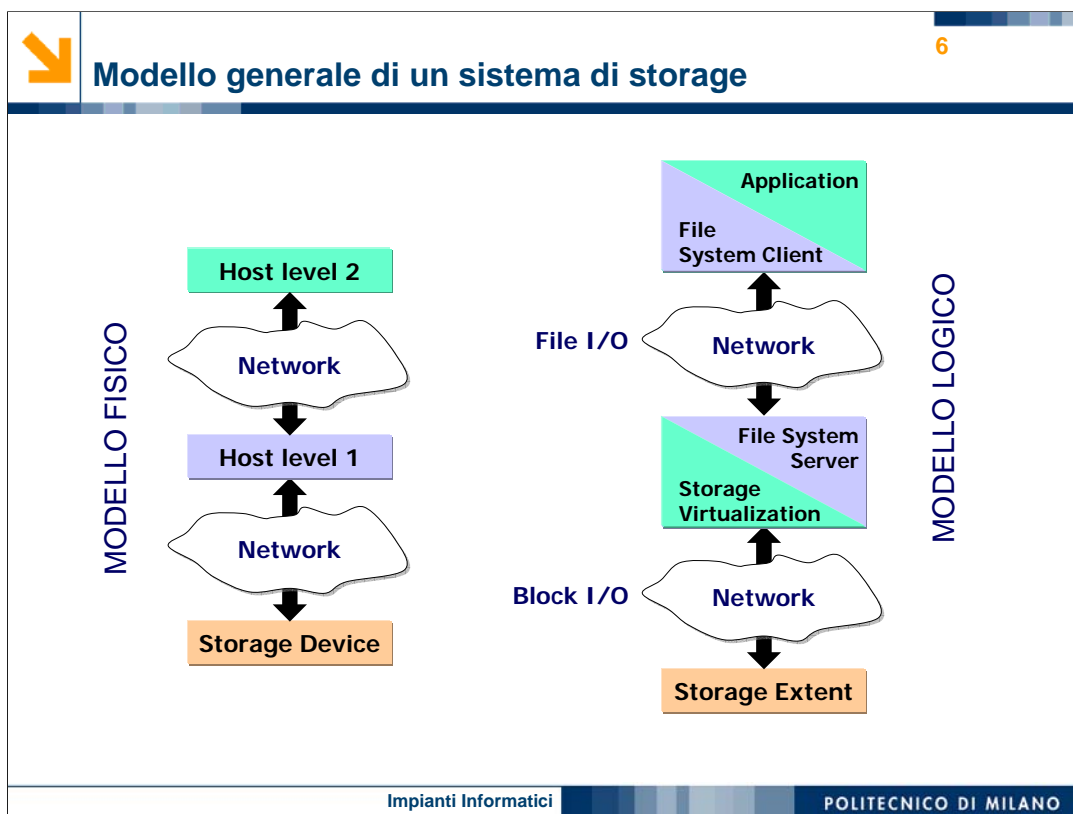
# Impianti Informatici

 POLITECNICO DI MILANO



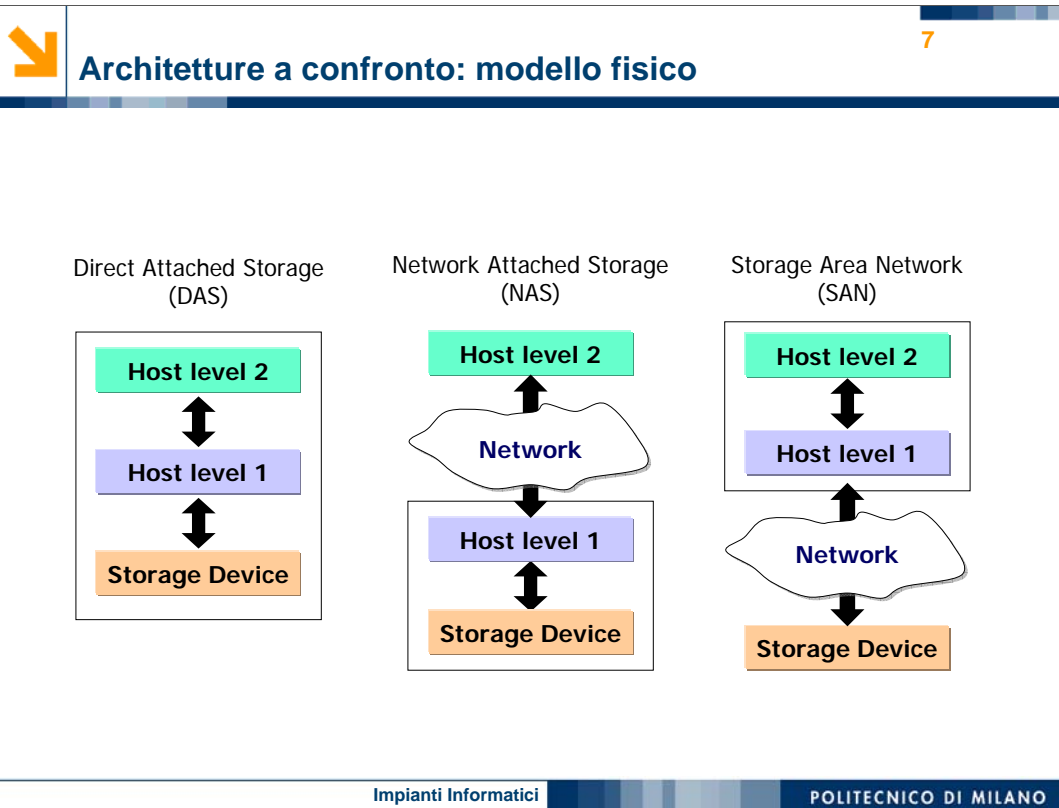
Rassegna delle architetture





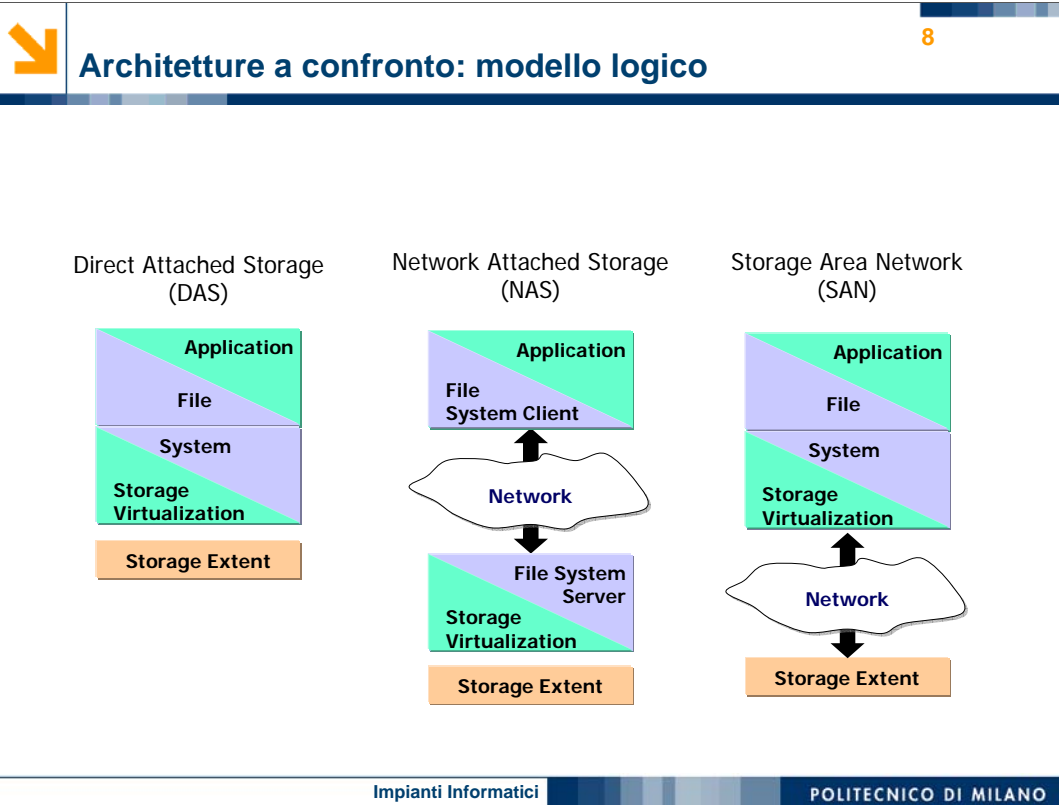
In generale, l'architettura di un sistema di storage può essere visto come composta da tre livelli. Vediamo prima di tutto l'architettura fisica di un sistema di storage.

- Al livello più alto abbiamo l'host, ossia il computer che deve accedere i dati.
- A sua volta l'host può essere scomposto in due livelli separati tra di loro da una rete
- Nell'ultimo livello abbiamo lo storage vero e proprio, composto da uno o più dischi configurati in modalità RAID. Host e storage dialogano tra loro tramite un'altra rete
- La stessa suddivisione in 3 livelli si ha analizzando l'architettura logica di un sistema di storage
- Al livello più alto abbiamo un'applicazione che deve accedere a dei file. Per fare questo l'applicazione si basa su dei servizi offerti dal file system
- Il file system si suddivide in due componenti, una componente client e una componente server. La componente server del file system si trova al livello intermedio dell'architettura e trasforma le richieste logiche in richieste fisiche
- mediante un driver che permette di dialogare con il device di storage. Da un punto di vista logico il device di storage viene chiamato storage extent. Uno storage extent è uno spazio continuo di memoria su un device di memoria di massa. Lo storage extent è la forma più "bassa" di storage logico, sotto i file, in cui lo storage è visto come un insieme di volumi (i dischi logici) e i volumi sono composti e indirizzati come **blocchi** di lunghezza definita
- E' interessante notare che, sia da un punto di vista fisico che logico, i tre livelli sono suddivisi da una rete. La rete che divide il livello superiore da quello intermedio trasporta richieste di I/O a livello di file, senza che vi sia conoscenza di come queste operazioni su file vengano spezzettate in operazioni su unità più semplici (i blocchi). La rete che divide il livello intermedio da quello sottostante trasporta richieste di I/O a livello di blocchi, senza che vi sia conoscenza dei file a cui fanno riferimento i dati trasportati.



Nel modulo precedente abbiamo visto che le architetture di storage si dividono in tre categorie: **DAS, NAS e SAN**. Vediamo adesso come le tre categorie si mappano sulla suddivisione in livelli fisici

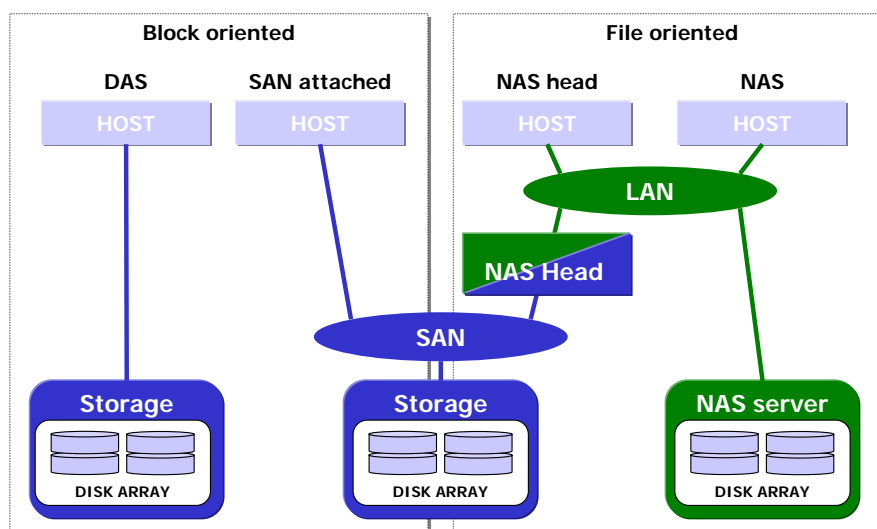
- Nei sistemi DAS i tre livelli sono collassati in un solo livello: host e storage si trovano sulla stessa macchina
- Nei sistemi NAS il livello più alto dell'host è separato dal livello più basso da una rete
- Nei sistemi SAN, i due livelli di host sono su una stessa macchina, il livello di storage è separato da una rete



Vediamo adesso come le tre categorie DAS, NAS e SAN si mappano sulla suddivisione in livelli logici

- Nei sistemi DAS i tre livelli sono collassati in un solo livello: application, file system e storage extent
- Nei sistemi NAS il file system client e il file system server sono separati da una rete
- Nei sistemi SAN, lo storage extent è separato dal file system da una rete





Le architetture di storage possono essere classificate in base al tipo di accesso ai dati eseguito dall'host (Blocchi o File).

1. Nel caso di accesso a blocchi, abbiamo i DAS e le SAN attached.
2. Nel caso di accesso a file, abbiamo i NAS e le SAN con NAS head.

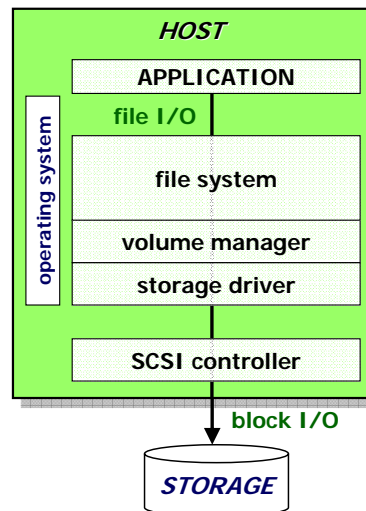
*Paolo Cremonesi*

# Impianti Informatici

 POLITECNICO DI MILANO

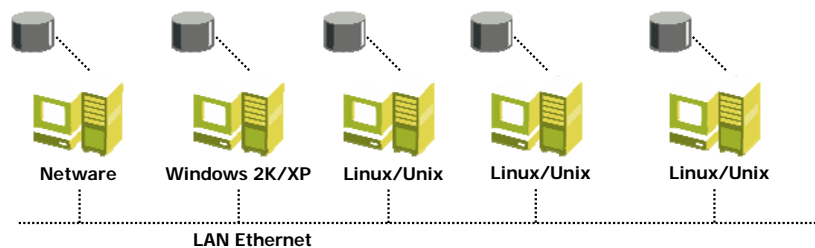


DAS: Direct Attached Storage



Studiamo ora in dettaglio l'architettura logica di un sistema DAS. Per capire il funzionamento di un DAS, studiamo cosa succede quando un'applicazione vuole accedere a delle informazioni nello storage.

- L'applicazione esegue una o più richieste di I/O (ad esempio, aprire e leggere un file di testo)
- La richiesta viene presa in carico dal sistema operativo, ed in particolare da quella parte del sistema operativo deputata alla gestione del file system.
- Il file system a sua volta inoltra la richiesta al volume manager che la inoltra al driver del disco
- Il driver invia la richiesta al controller del disco
- e il controller dialoga a sua volta con il disco



Nei sistemi DAS il dispositivo di storage è direttamente collegato direttamente ad un host. Questo fa sì che i dati siano “confinati” nel server che ospita il disco

- Espandere lo storage implica quindi acquistare e gestire altri server.
- La gestione, soprattutto in ambienti eterogenei, è complicata e la scalabilità e le prestazioni non sono buone

*Paolo Cremonesi*

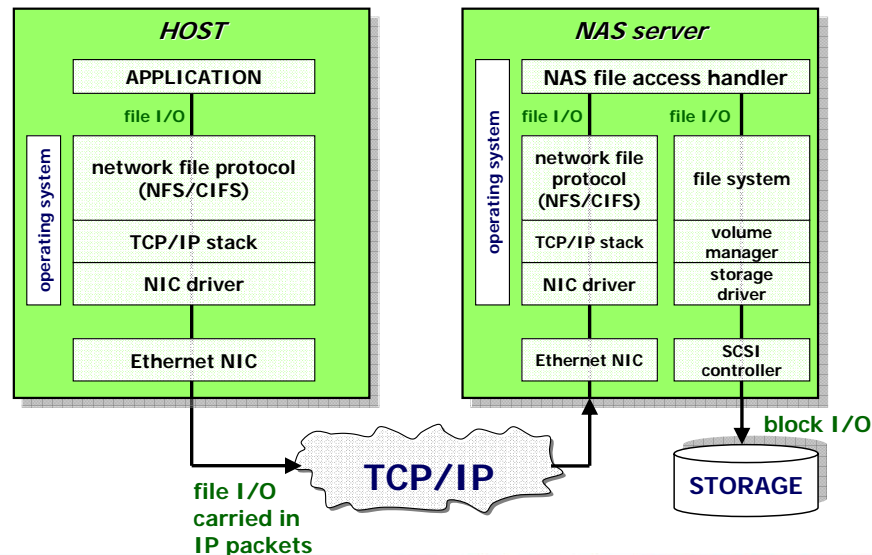
# Impianti Informatici

 POLITECNICO DI MILANO



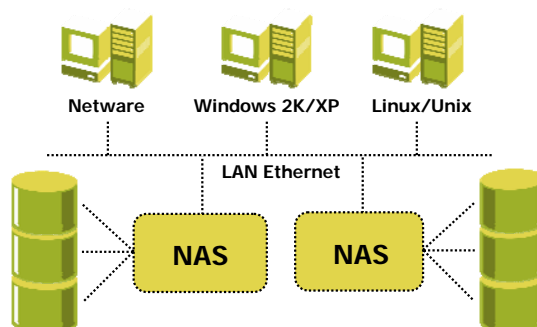
NAS: Network Attached Storage





Studiamo ora in dettaglio l'architettura logica di un sistema NAS. Per capire il funzionamento di un NAS, vediamo cosa succede quando un'applicazione vuole accedere a delle informazioni nello storage.

- L'applicazione su host esegue una o più richieste di I/O (ad esempio, aprire e leggere un file di testo)
- La richiesta viene presa in carico dal sistema operativo, ed in particolare da quella parte del sistema operativo deputata alla gestione del file system. In questo caso la richiesta viene presa in carico dalla parte client del network file system che utilizza un protocollo applicativo come NFS (Network File System) o CIFS (Common Internet File System)
- Il client del network file system a sua volta inoltra la richiesta allo stack protocollare TCP/IP che incapsula i dati in transito in pacchetti IP e li inoltra al driver della scheda di rete
- Il driver della scheda di rete invia la richiesta alla rete sotto forma di pacchetti IP incapsulati, ad esempio, in frame Ethernet
- I pacchetti vengono ricevuti dalla scheda di rete del NAS e risalgono la pila protocollare TCP/IP
- fino a ricomporre la richiesta di I/O originale tramite la parte server nel network file system.
- A questo punto la richiesta si comporta come in un DAS: la richiesta di I/O viene presa in carico dal file system.
- Il file system a sua volta inoltra la richiesta al volume manager che la inoltra al driver del disco
- Il driver invia la richiesta al controller del disco e il controller dialoga a sua volta con il disco



Con il termine di NAS ci si riferisce a elementi che possiedono un loro indirizzo IP di rete e provvedono servizi di accesso ai file per gli host connessi ad una rete TCP/IP. Con gli storage NAS i dati tra host e storage vengono quindi trasferiti attraverso una rete basata su TCP/IP. Eliminando gli accessi allo storage e la sua gestione, sia i programmi applicativi che i file possono essere serviti più rapidamente perché non competono più per le risorse dello stesso sistema.

- La scalabilità di un sistema basato su NAS può essere ottenuta incrementando il numero di storage device connessi ad un NAS
- o collegando nuovi NAS alla rete. Nei sistemi NAS lo storage è fisicamente separato dall'host applicativo permettendo una discreta **scalabilità e una notevole facilità di gestione.**

*Paolo Cremonesi*

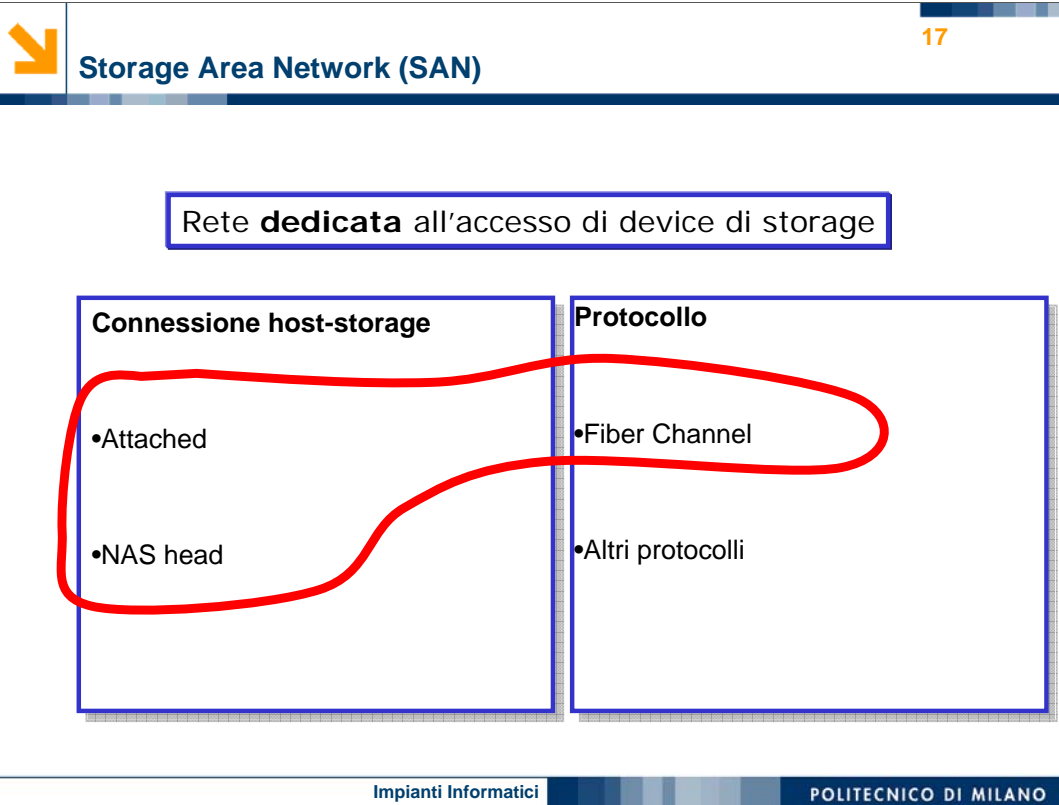
# Impianti Informatici

 POLITECNICO DI MILANO



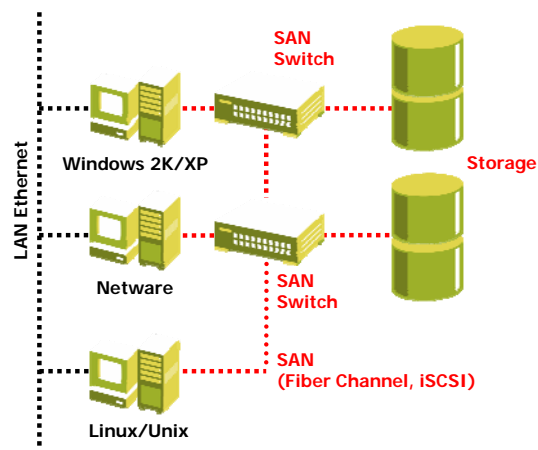
**SAN: Storage Area Network**





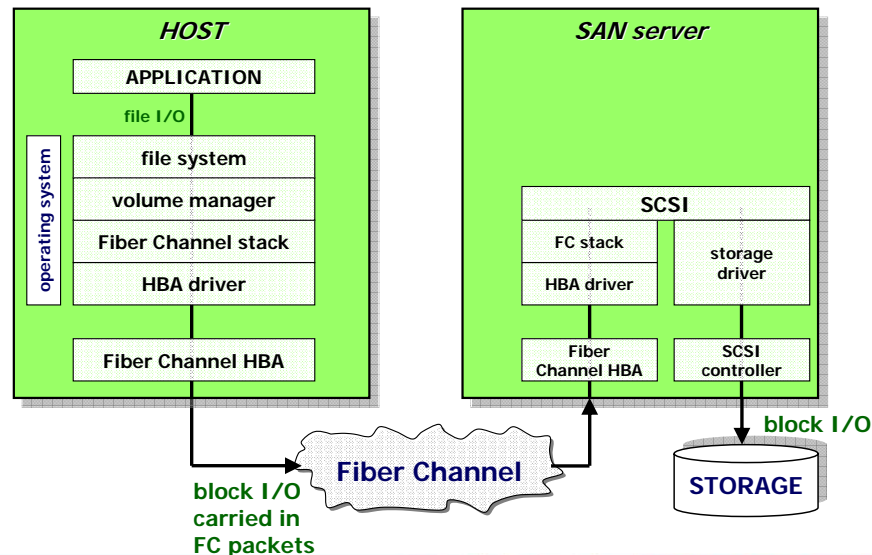
Con il termine di Storage Area Network (SAN) ci si riferisce ad una rete di elementi che forniscono accesso a dei storage extent per gli host connessi ad una rete.

- Le storage area network si dividono in due categorie a seconda di come gli host sono connessi allo storage
- le cosiddette attached SAN
- le SAN basate su NAS head
- Inoltre le SAN si possono classificare a seconda del protocollo di rete utilizzato
- abbiamo così le SAN che utilizzano il protocollo FC
- e le SAN che utilizzano altri protocolli
- In questo modulo vedremo le SAN (attached e con NAS head) basate su Fiber Channel. Nel modulo successivo vedremo le SAN basate su altri protocolli di rete



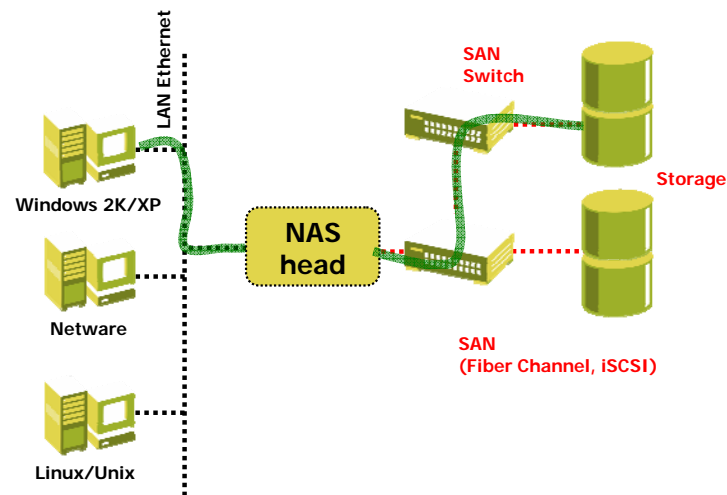
Nel caso di **SAN attached**, gli host sono connessi (attached) a due reti distinte:

- la rete locale TCP/IP
- una rete dedicata allo storage e tipicamente basata su protocolli ottimizzati per lo storage, quali ad esempio, il protocollo Fiber Channel
- La scalabilità di un sistema basato su SAN può essere ottenuta incrementando il numero di storage device connessi alla rete SAN. Nei sistemi SAN lo storage è fisicamente separato dall'host applicativo permettendo una notevole **scalabilità e una discreta facilità di gestione**.

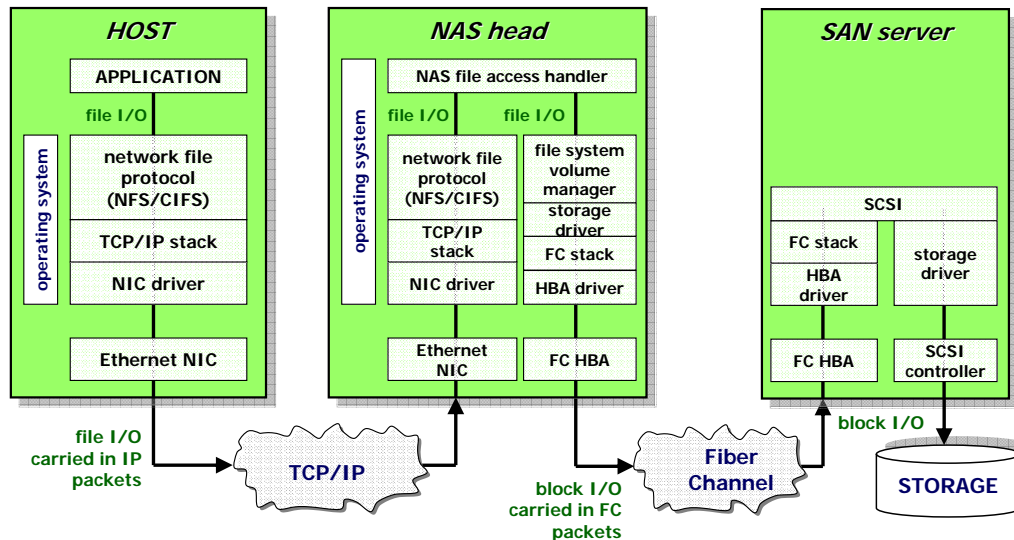


Studiamo ora in dettaglio l'architettura logica di un sistema di SAN attached basato su Fiber Channel. Per capire il funzionamento di un SAN, vediamo cosa succede quando un'applicazione vuole accedere a delle informazioni nello storage.

1. L'applicazione su host esegue una o più richieste di I/O (ad esempio, aprire e leggere un file di testo)
2. La richiesta viene presa in carico dal file system che inoltra le richieste al volume manager
3. Il volume manager incapsula le richieste SCSI nello stack protocollare Fiber Channel
4. Il protocollo FC invia le richieste sotto forma di pacchetti FC all'Host Bus Adapter (mediante il relativo driver)
5. I pacchetti vengono ricevuti dal Host Bus Adapter del sistema SAN e risalgono la pila protocollare FC
6. fino ad estrarre il comando SCSI originale
7. A questo punto la richiesta si comporta come in un DAS: il comando SCSI viene presa in carico dal driver del disco
8. Il driver invia la richiesta al controller del disco e il controller dialoga a sua volta con il disco



- Nel caso di SAN con NAS head, la rete locale TCP/IP e la rete dedicata allo storage sono separate da un NAS head che agisce da gateway tra le due reti.
- Gli host sulla rete locale TCP/IP vedono il NAS head come un normale NAS server
- Il NAS head, anziché avere uno storage direttamente connesso come accade con i NAS tradizionali, inoltra nella rete SAN le richieste di accesso allo storage



Studiamo ora in dettaglio cosa succede nel caso di un sistema SAN basato su NAS head e rete SAN con protocollo Fiber Channel

1. L'applicazione su host esegue una o più richieste di I/O (ad esempio, aprire e leggere un file di testo). La richiesta viene presa in carico dalla parte client del network file system che inoltra le richieste nello stack protocollare TCP/IP mediante il protocollo applicativo NFS o CIFS. Il protocollo TCP/IP invia le richieste alla scheda di rete Ethernet (mediante il relativo driver).
2. I pacchetti vengono ricevuti dalla scheda di rete del NAS e risalgono la pila protocollare TCP/IP fino a ricomporre la richiesta di I/O originale tramite la parte server nel network file system.
3. A questo punto la richiesta si comporta come in host direct attached

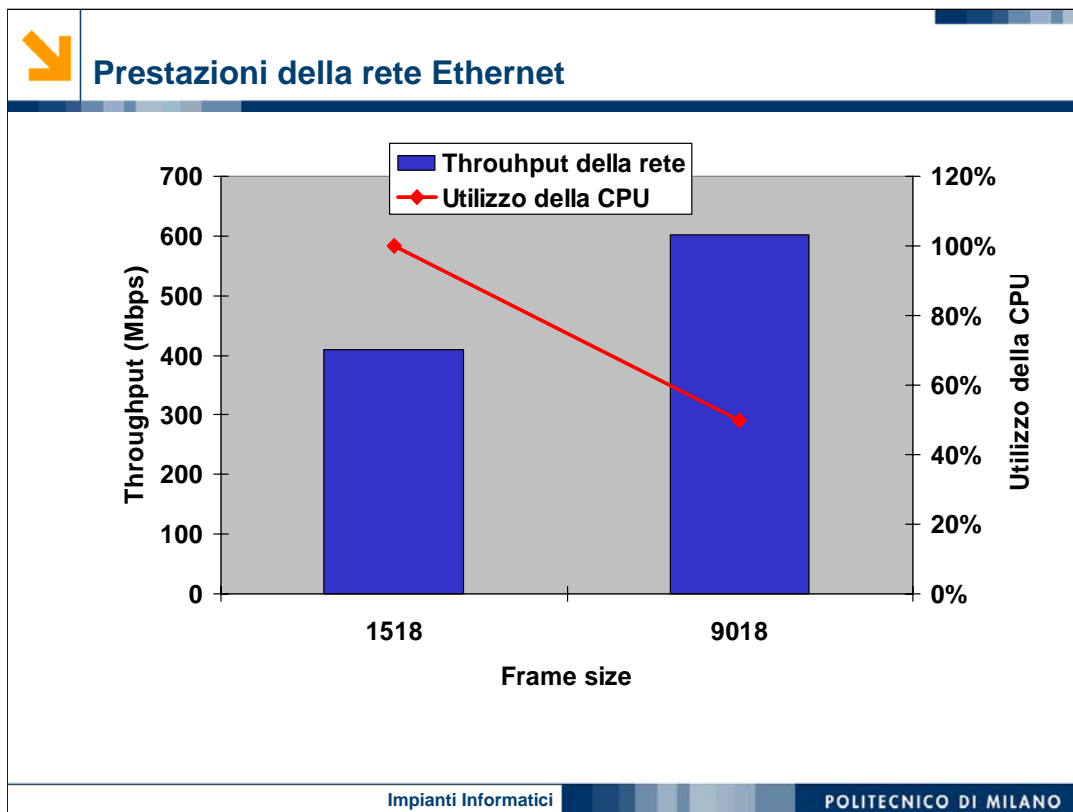
*Paolo Cremonesi*

# Impianti Informatici

 POLITECNICO DI MILANO



## Il protocollo Fibre Channel (FC)



Uno dei problemi principali legato alle architetture do storage è dato dalle prestazioni della rete che collega l'host con lo storage extent.

Il traffico tra host e storage extend è caratterizzato da pacchetti "piccoli", con una dimensione tipicamente più piccola rispetto ai 1518 byte che costituiscono la dimensione minima di un frame Ethernet.

Questo aspetto, legato agli elevati volumi di traffico tipici dei sistemi di storage, crea un notevole problema in termini di prestazioni.

•Guardiamo, ad esempio, le prestazioni di una Gigabit Ethernet con frame size di 1518 byte e confrontiamole con le prestazioni ottenibili nel caso di *jumbo frames* di 9018 bytes. Si vede che, nel primo caso, il throughput raggiunto effettivamente è pari a 400 Mbps, pari a meno della metà del throughput massimo teorico (1 Gbps). Nel caso di frame più grandi il throughput supera invece i 600 Mbps, con un incremento del 50% delle prestazioni.

•Questo divario nelle prestazioni ha, come causa principale, l'eccessivo carico di lavoro della CPU dell'host che, per sostenere un traffico di 1 Gbps, dovrebbe eseguire 80'000 interrupt al secondo, saturando la CPU. Questo emerge chiaramente dalla figura dove si osserva che, per frame di 1518 byte, la CPU dell'host è al 100% mentre per frame più grandi l'utilizzo della CPU scende al 50%

•tra cui l'overhead introdotto dal protocollo IP per il routing, l'overhead introdotto

No bus arbitration delays

Performance considerations:

e.g.,

1.5 KB Ethernet MTU

Requires processing 80,000 Host interrupts/sec @ 1 Gb/sec

or *Jumbo* frames, which also requires installing a new infrastructure

Which is why Fibre Channel was designed the way it is!



- La pila TCP/IP su Ethernet ha una serie di overhead
  - gestione dei frame Ethernet (overhead sulla rete)
  - dimensione dei frame Ethernet (overhead sulla CPU)
  - routing dei pacchetti IP (overhead sugli apparati di rete)
- Gli overhead aumentano con il volume dei dati trasferiti

Più in generale, le architetture di storage network basate su TCP/IP e su Ethernet hanno una serie di problemi di prestazioni

- alcuni problemi sono dovuti alla gestione dei frame Ethernet, che causano overhead sulla rete
- altri problemi sono dovuti alla dimensione dei frame Ethernet (la cui gestione causa overhead sulla CPU di host e storage server)
- altri problemi sono legati al meccanismo di routing dei pacchetti IP che è svolto a livello alto della pila protocollare (e questo introduce un overhead sugli apparati di rete)
- Inoltre, tutti questi overhead diventano pesanti nel momento in cui la rete deve gestire gli elevati volumi di traffico tipici del storage network. Questi problemi hanno spinto ad ideare una nuova pila protocollare ottimizzata per gli storage network





- Combina le caratteristiche migliori
  - del bus locali, per quanto riguarda la semplicità e la velocità
  - dei protocolli di rete, per quanto riguarda la flessibilità e la capacità di realizzare reti estese e con topologie complesse
- Caratteristiche
  - Collegamenti full-duplex
  - Throughput di 1600 Mbps
  - Supporto per connessioni fino a 10 km
  - Connettori piccoli
  - Utilizzo di componenti standard

Vediamo quindi come è strutturato il protocollo Fibre Channel

•Combina le caratteristiche migliori

del bus locali come il bus SCSI, per quanto riguarda la semplicità e la velocità  
dei protocolli di rete come il TCP/I, per quanto riguarda la flessibilità e la  
capacità di realizzare reti estese e con topologie complesse

•Caratteristiche

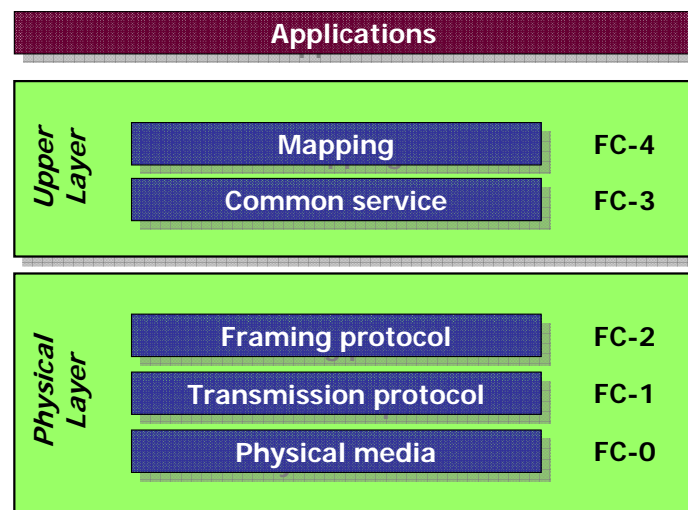
Collegamenti full-duplex

Throughput di 1600 Mbps

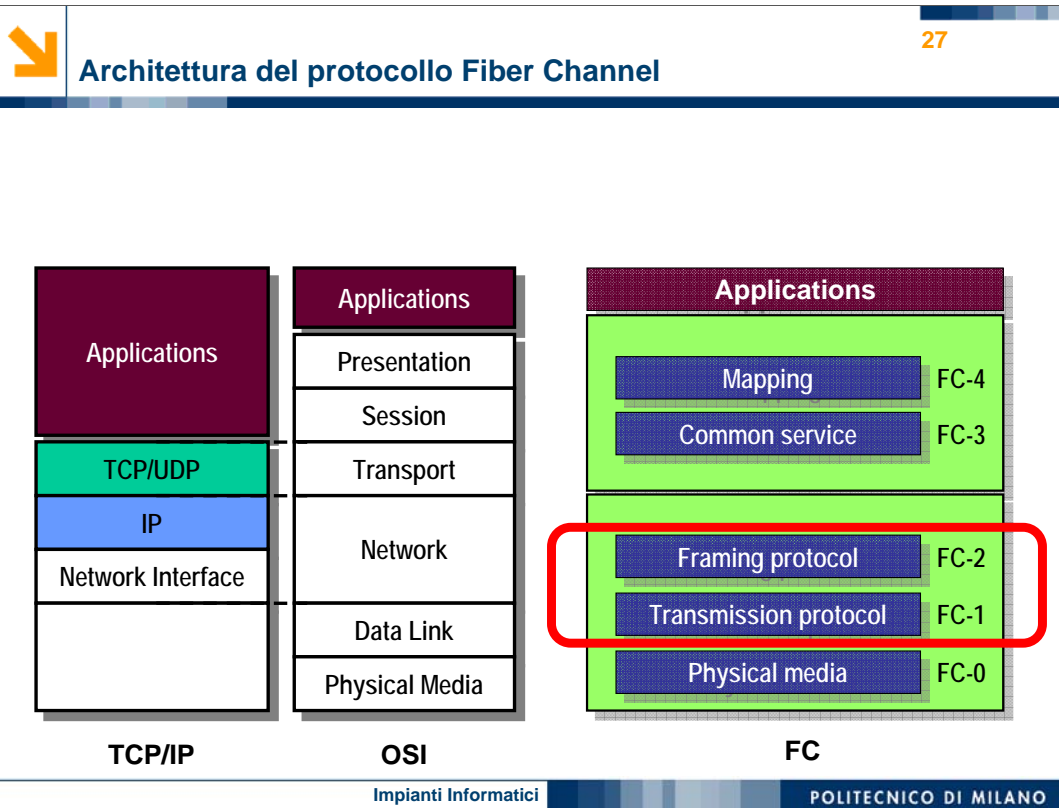
Supporto per connessioni fino a 10 km (MAN)

Connettori piccoli (come i SATA)

Utilizzo di componenti standard

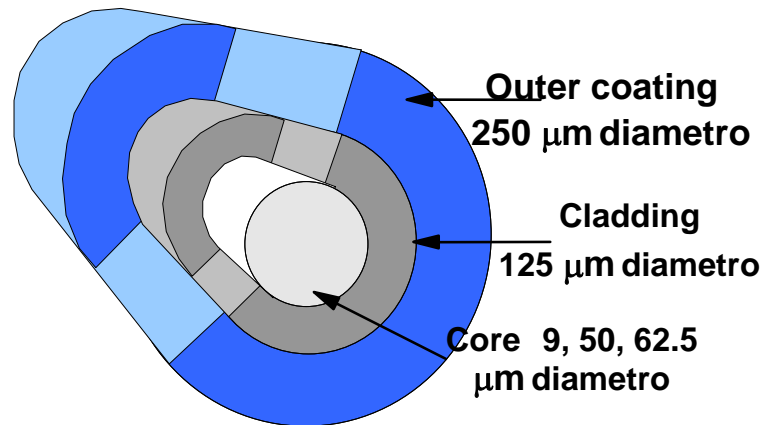


- Il protocollo Fiber Channel si articola su due macrolivelli, denominati Upper Layer e Physical Layer
  - Nel Physical layers abbiamo tre livelli, numerati da FC-0 a FC-2
  - Il livello **FC-0 Physical Media**, riguarda i mezzi fisici di trasmissione, che sono gli stessi di Ethernet, e comprendono fibre ottiche, cavi coassiali, doppini in rame
  - Il livello **FC-1 Transmission Protocol**, definisce gli schemi di codifica dei segnali e di sincronizzazione dei dati: questi schemi sono diversi da quelli di Ethernet
  - Il livello **FC-2 Framing Protocol**, definisce il formato dei frame, il flowcontrol e l'error control
  - Nell'Upper layers abbiamo due livelli, FC-3 e FC-4
  - Il livello **FC-3 Common Service**, gestisce tipicamente i multicast
  - Il livello **FC-4 Mapping**, definisce il mapping dei network protocol (ad esempio, IP) o dei channel protocol (ad esempio, SCSI)
- E' importante osservare che la pila del protocollo del Fiber Channel non rientra nella classica struttura protocollare ISO/OSI come accade, ad esempio, per il protocollo TCP/IP.



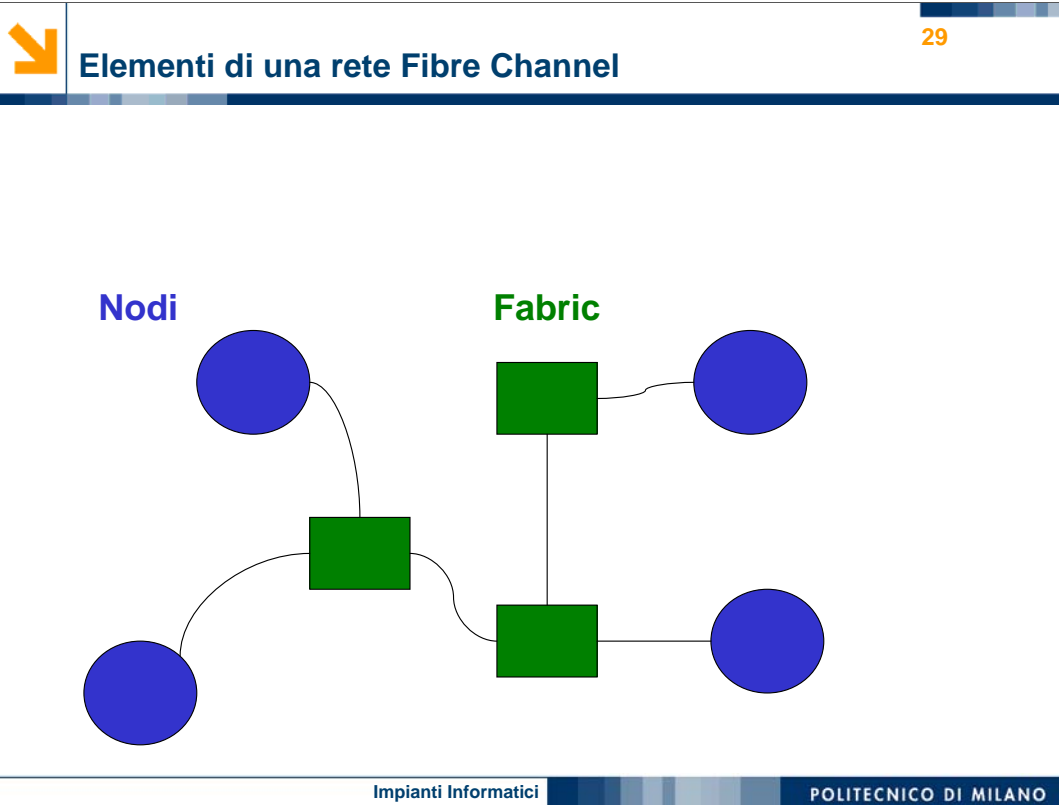
E' importante evidenziare una differenza importante tra il protocollo Fibre Channel e il protocollo TCP/IP su Ethernet: il protocollo Fibre Channel è **packet-oriented**, ossia è disegnato per il routing.

- Il protocollo Fibre Channel sposta molte funzioni di networking nei livelli più bassi, dove possono essere implementate in hardware dagli apparati di rete. Alcune di queste funzioni sono la gestione della frammentazione dei pacchetti e il routing dei pacchetti



E' importante sottolineare che, a dispetto del nome, il protocollo Fibre Channel non utilizza esclusivamente le fibre ottiche come mezzo trasmissivo, anche se le prestazioni migliori si ottengono con l'uso delle fibre ottiche. Una fibra ottica è composta da due strati:

- **Il Core** è il mezzo trasparente che trasporta il segnale laser (la velocità di propagazione è di 200.000 Km/sec e 1 Km è percorso in circa 5  $\mu\text{sec}$ )
- **Il Cladding** è lo strato ad esso concentrico che costringe l'onda a seguire il percorso della fibra ottica:



Vediamo adesso quali sono gli elementi che costituiscono una rete FC.

- Abbiamo innanzitutto i *nodi*, che sono le parti terminali della rete. I nodi sono dotati di connessioni chiamate *N\_ports*
- Abbiamo poi i *fabric*, che sono un insieme di elementi di switching tra nodi e altri fabric. Le porte dei fabric sono chiamate *F\_ports*. I fabric sono responsabili del buffering e del routing dei pacchetti



## Apparati di rete

- Il protocollo Fibre Channel è disegnato per operare con gli stessi tipi di cablaggi utilizzati per le reti Gigabit Ethernet
- Gli apparati di rete però sono diversi
- Esistono
  - Fibre Channel **hub**
  - Fibre Channel **switch**



Impianti Informatici

POLITECNICO DI MILANO

Il protocollo Fibre Channel è disegnato per operare con gli stessi tipi di cablaggi utilizzati per le reti Gigabit Ethernet

- Gli apparati di rete però sono diversi
- Esistono

Fibre Channel **hub**

Fibre Channel **switch**


Le prestazioni degli switch FC sono decisamente elevate, con una latenza molto bassa (di circa 1  $\mu$ sec)



## Vantaggi offerti dallo standard Fibre Channel

31

- Miglioramento delle prestazioni, capacità e affidabilità dei sistemi di memorizzazione e della rete di computer.
- Aggiunta o rimozione dei dispositivi (RAID, librerie, dischi, eccetera), senza alcuna interruzione del servizio; possibilità di collegare i diversi dispositivi direttamente in rete, raggiungendo maggiori distanze tra i server e i dispositivi di memorizzazione (fino a 10 km), e assicurando l'integrità dei dati lungo l'intera connessione.

	<div data-bbox="1177 226 1217 253">32</div> <div data-bbox="336 248 762 286">Fibre Channel vs. Ethernet</div>
<ul style="list-style-type: none"> <li>•A favore del Fibre Channel           <ul style="list-style-type: none"> <li>▪ Elevato overhead del TCP/IP su Ethernet</li> <li>▪ 10-gigabit Ethernet è costosa</li> <li>▪ Fibre Channel ha prestazioni migliori rispetto alla gigabit Ethernet</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>•A favore di Ethernet           <ul style="list-style-type: none"> <li>▪ 1-gigabit Ethernet è meno costosa rispetto al Fibre Channel</li> <li>▪ Le infrastrutture IP ed Ethernet sono presenti ovunque è ben conosciute</li> <li>▪ E' possibile riutilizzare apparati di rete e di storage già esistenti</li> </ul> </li> </ul>
Impianti Informatici	POLITECNICO DI MILANO

I punti a favore del Fibre Channel sono i seguenti:

Innanzitutto, il protocollo TCP/IP su Ethernet ha un overhead elevato del TCP/IP rispetto al FC

La Gigabit Ethernet da 10 Gigabit è costosa

mentre il Fibre Channel ha prestazioni migliori rispetto alla Gigabit Ethernet da un Gbit

•I punti a favore di Ethernet e del TCP/IP sono i seguenti:

La gigabit Ethernet è meno costosa rispetto al Fibre Channel

Le infrastrutture IP ed Ethernet sono presenti ovunque e le loro caratteristiche sono ben conosciute

E' possibile riutilizzare apparati di rete e di storage già esistenti





- Point to point
  - connessione singola fra due nodi
- Loop
  - la bandwidth è suddivisa fra tutti i nodi connessi al loop
  - ogni nodo è connesso ai suoi adiacenti
- Arbitrated loop
  - due o più porte sono fra loro connesse ma solo due alla volta possono comunicare nello stesso tempo
- Switched or fabric
  - uno switch permette connessioni multiple fra nodi
  - circuit switch stabilisce una connessione dedicata fra nodi
  - frame switch stabilisce la connessione solo quando è necessaria (switched fabric)

Vediamo adesso quali sono le topologie di rete possibili con il protocollo Fibre Channel

•Point to point:

una connessione singola fra due nodi - tutta la bandwidth è ad essi dedicata

•Loop:

la bandwidth è suddivisa fra tutti i nodi connessi al loop

ogni nodo è connesso ai suoi adiacenti

il dispositivo hub apre il loop quando il nodo è connesso e lo chiude quando si disconnette

•Arbitrated loop (127 addresses):

due o più porte sono fra loro connesse ma solo due alla volta possono comunicare nello stesso tempo

•Switched or fabric (16 MB addresses):

uno switch permette connessioni multiple fra nodi

circuit switch stabilisce una connessione dedicata fra nodi

frame switch stabilisce la connessione solo quando è necessaria (switched fabric)

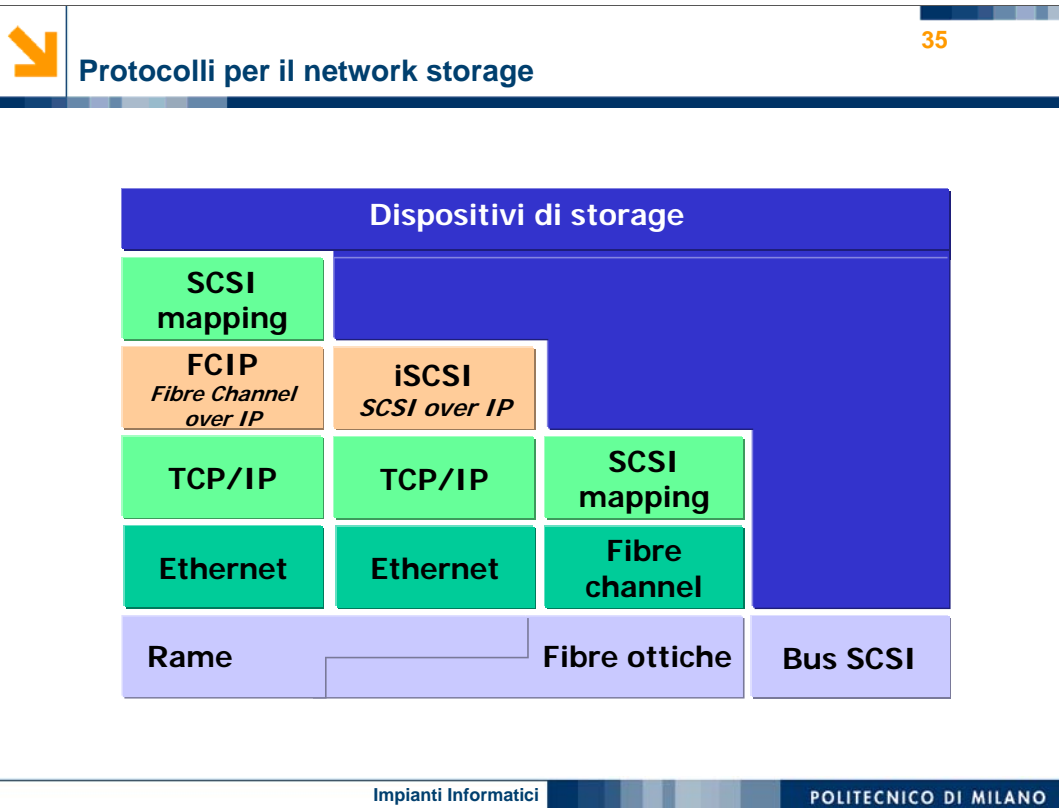
*Paolo Cremonesi*

# Impianti Informatici

 POLITECNICO DI MILANO



Altri protocolli per SAN



Vediamo ora quali sono alcuni dei protocolli disponibili per gestire un sistema di storage oltre al classico bus SCSI.

1. Tutti i protocolli che prendiamo in considerazione sono in grado di usare come canale fisico di trasmissione gli stessi doppini in rame o gli stessi cablaggi ottici disponibili per le reti Ethernet
2. Abbiamo già visto che il protocollo ideale, in termini di prestazioni, è il protocollo Fibre Channel
3. Una seconda possibilità consiste nell'incapsulare direttamente i comandi SCSI dentro i pacchetti IP mediante il protocollo iSCSI.
4. Un'ultima possibilità consiste nel portare il protocollo Fiber Channel su IP mediante il protocollo FCIP. Il protocollo FiberChannel è stato approfondito nei moduli precedenti. In questo modulo vedremo alcune caratteristiche dei protocolli iSCSI e FCIP



- Pregi
  - da 4 a 10 volte meno costosa
  - non ci sono costi di nuovo hardware
- Difetti
  - non ha le stesse prestazioni di FC

Il protocollo iSCSI offre una soluzione meno costosa per realizzare una SAN rispetto all'utilizzo di Fibre Channel.

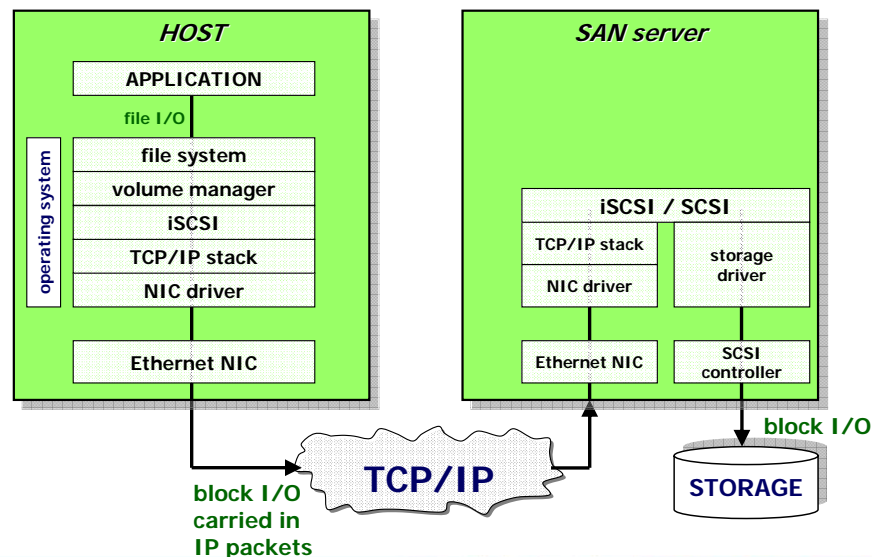
1. In particolare, la soluzione basata su iSCSI è da 4 a 10 volte meno costosa di soluzioni analoghe basate su FC
2. Il risparmio è dovuto alla possibilità di riutilizzo di storage e di apparati di rete esistenti, e quindi non ci sono costi di nuovo hardware
3. Per contro, una soluzione basata su iSCSI non ha la stessa scalabilità di una soluzione basata su FC



Session	FTP	SMTP	iSCSI	SNMP	DHCP
Transport	TCP			UDP	
Network	IP				
Data-link / physical	Ethernet	Token Ring	SONET	FDDI	

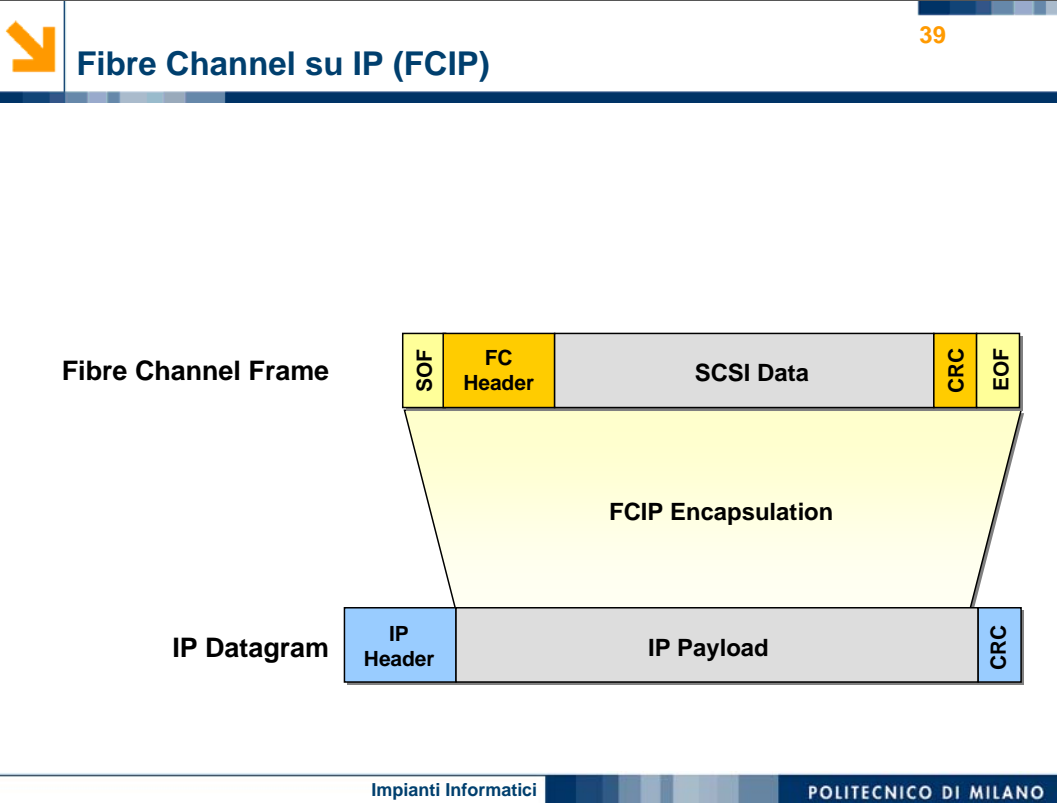
Il protocollo iSCSI è definito sopra il protocollo TCP

1. i blocchi di comando SCSI e i dati sono incapsulati in segmenti TCP e trasferiti su reti TCP/IP/Ethernet. Nelle SAN basate su iSCSI ogni target (storage) viene identificato da un indirizzo IP, un application port e uno SCSI identifier [SCSI device name]



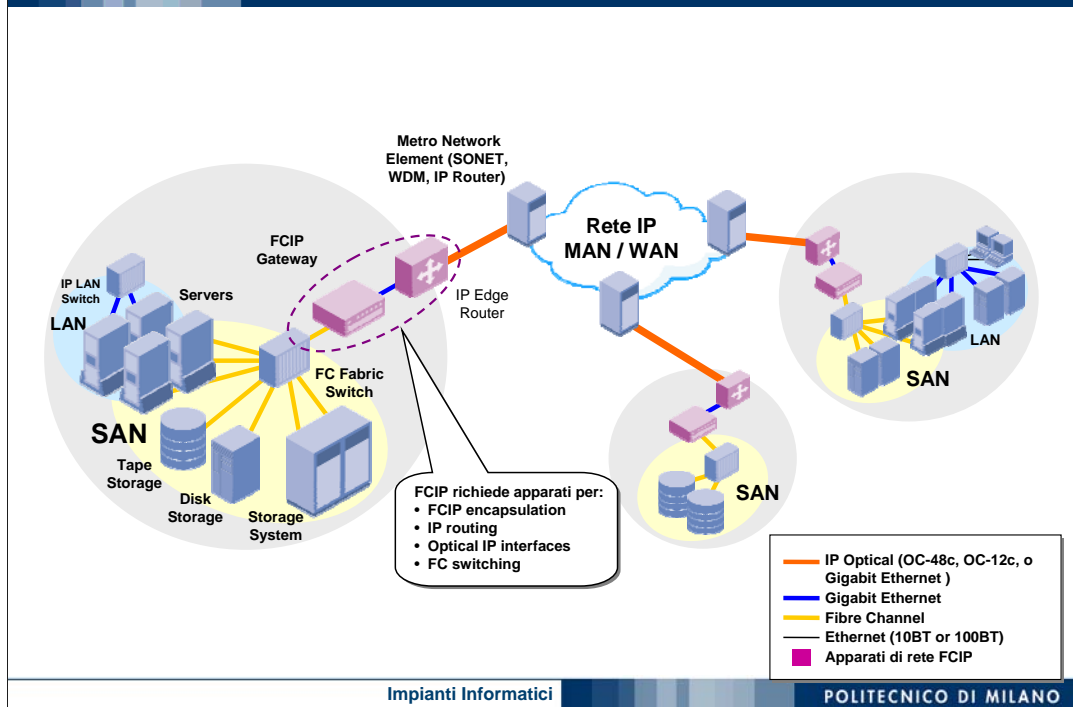
Studiamo ora in dettaglio cosa succede nel caso di un sistema SAN attached basato su protocollo iSCSI

1. L'applicazione su host esegue una o più richieste di I/O (ad esempio, aprire e leggere un file di testo)
2. La richiesta viene presa in carico dal file system che inoltra le richieste al volume manager
3. Il volume manager incapsula le richieste SCSI nello stack protocollare TCP/IP mediante il protocollo applicativo iSCSI
4. Il protocollo TCP/IP invia le richieste alla scheda di rete Ethernet (mediante il relativo driver)
5. I pacchetti vengono ricevuti dalla scheda di rete del sistema SAN e risalgono la pila protocollare TCP/IP
6. fino ad estrarre il comando SCSI originale
7. A questo punto la richiesta si comporta come in un DAS: il comando SCSI viene presa in carico dal driver del disco
8. Il driver invia la richiesta al controller del disco e il controller dialoga a sua volta con il disco



Vediamo adesso il secondo protocollo alternativo al FiberChannel, ossia il protocollo FCIP.

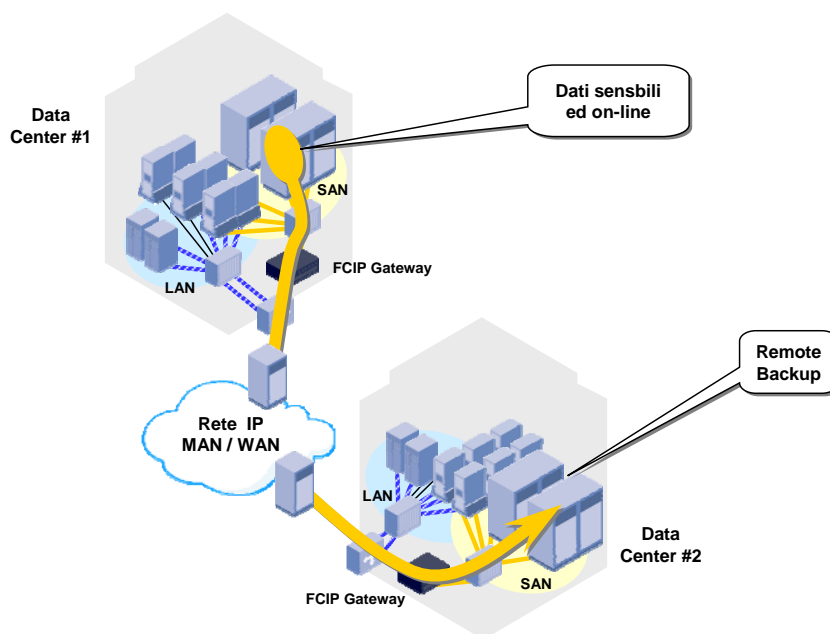
1. Si tratta di un sistema di tunneling del protocollo FC su IP. La stessa rete appare basata su FC per le SAN e basata su TCP/IP per le applicaioni c/s tradizionali



La creazione di reti SAN basate sul protocollo FCIP richiede l'uso di apparati di rete IP trasizionali

1. e apparati di rete dedicati al protocollo FCIP
2. Nell'esempio in figura vediamo un esempio di rete ibrida FC+FCIP. Abbiamo due reti locali SAN connesse tra loro da backbone IP in fibra ottica
3. e da reti gigabit ethernet
4. il FCIP gateway connette la rete IP alla rete con FiberChannel
5. Il traffico IP normale firetto alle lan non è influenzato dalla presnza del FCIP gateway





Vediamo adesso alcuni scenari di utilizzo di FCIP

- Data Backup & Restore
- Data Replication

# Impianti Informatici

 POLITECNICO DI MILANO



Esempi di apparati di storage



- Server Rack di 3-4 RackUnit con dischi rigidi ATA o SATA
- Costi bassi (2K€ al TB)
- Buona flessibilità nella gestione (Linux Box)
- Apparati basati su server biprocessore Intel con 16 dischi da 250 GB
  - 4TB Lordi, 3 TB Netti Raid5+2HotSpare
- Problemi rilevati più frequentemente:
  - Guasti frequenti di dischi e controller



Vediamo adesso alcuni esempi di DAS e NAS di fascia bassa

1. In generale si tratta di Server Rack\_mountable di 3-4 RackUnit “imbottiti” di dischi rigidi ATA o SATA gestiti da due RAID controller (3Ware, Adaptec ..)
2. Costi bassi (2K€ al TB)
3. Buona flessibilità nella gestione (Linux Box)
4. Si trovano comunemente sul mercato Apparati basati su server biprocessore Intel in grado di ospitare 16 dischi da 250 GB (4TB Lordi, 3 TB Netti Raid5+2HotSpare)
5. Problemi rilevati più frequentemente sono i guasti frequenti di dischi e controller



- Appareti complessi
- Produttori principali
  - EmC2 (DELL), HP, NetApp, Procom, IBM
- Costi alti (6-8K€ al TB)
- Scalano fino a dimensioni di decine di TB
- Sistemi chiusi
  - esportano i propri volumi via NFS, CIFS, HTTP e HTTPS
- Possono esportare porzioni di spazio disco via FC (Verso una SAN) o via iSCSI



Nel caso di NAS di fascia alta

1. Appareti complessi (spesso sono delle Storage Area Network chiuse con un sistema di gestione proprietario)
2. I principali produttori sono EmC2 (DELL), NetApp, Procom, IBM
3. I costi sono ancora alti (6-8K€ al TB) a fronte di una dichiarata maggiore continuità di servizio e una ridondanza a caldo di tutti i componenti essenziali
4. Scalano fino a dimensioni di decine di TB con un unico sistema di controllo. I sistemi di controllo e notifica sono in genere molto efficienti e permettono in genere espansioni dei volumi logici senza interruzioni di servizio
5. Sono sistemi chiusi e in genere esportano i propri volumi solo via NFS, CIFS, HTTP e HTTPS
6. Alcuni produttori possono esportare porzioni di spazio disco via FC (verso una SAN) o via iSCSI



- “Shelf” di dischi
  - dischi “semplici”
  - da 8 a 16 dischi
  - 1 o 2 controller RAID
  - numero di uscite Fibre Channel comprese fra 2 e 4
- Tecnologia dei dischi: FC, SCSI, ATA o SATA
- Configurazione raid set e creazione di partizioni logiche
- Più array con hot spare dedicati o con un unico hot spare globale
- Principali produttori
  - Axis (Brownie), Dell, Infortrend, RAIDTech,
- Prezzi indicativi per soluzioni ATA to FC:
  - 2-2,8 K€ al TB



Vediamo invece i tipici componenti delle architetture SAN

1. Si tratta di contenitori (“Shelf”) di dischi semplici (da 8 a 16 dischi ) dotati di 1 o 2 controller RAID con un numero di uscite Fibre Channel comprese fra 2 e 4
2. Tecnologia dei dischi: F.C., SCSI, ATA o SATA
3. I sistemi di gestione permettono la configurazione dei raid set e la creazione di partizioni logiche.
4. Tramite il sistema di gestione è spesso possibile definire più Array con Hot Spare dedicati o con un unico Hot Spare Globale
5. Principali produttori: Axis (Brownie), Dell, Infortrend, RAIDTech,
6. Prezzi indicativi per soluzioni ATA to F.C. (2-2,8 K€al TB)



## Esempio: BladeStore di Storagetek

46

- Blade:
  - è l'unità di base del sistema ed è composto da 5 dischi in tecnologia ATA
- Array:
  - contiene fino a 10 Balde
- Sistema Bladestore:
  - 1 Control Module + Max. 12 Array
- Massima Espandibilità
  - $12 \times 10 \times 5 \times 250\text{GB} = 150\text{TB}$

Blade (5 HD ADA) →

→ Array (10 Blade)



Impianti Informatici

POLITECNICO DI MILANO

Vediamo un esempio in particolare di sistema SAN: il BladeStore di Storagetek

1. Blade: è l'unità di base del sistema ed è composto da 5 dischi in tecnologia ATA
2. Array: contiene fino a 10 Balde
3. Sistema Bladestore: Control Module + Max. 12 Array
4. Massima Espandibilità  $12 \times 10 \times 5 \times 250\text{GB} = 150\text{TB}$