Introduction to
Storage Area Network (SAN)

Flavio Marinone
Dipartimento di Automatica e Informatica
Maggio 2004
Table of contents

- The storage evolution
  - Direct Attached Storage (DAS)
  - Small Computer System Interface (SCSI)
  - Network Attached Storage (NAS)
  - Storage Area Network (SAN)
Table of contents

- SAN architectures
  - Ethernet – TCP/IP
    - I-SCSI
    - FC-IP
  - Fibre Channel
- Complexities and performances
  - Fibre Channel vs. Ethernet
The storage evolution

- Traditionally each server has exclusive access to storage devices: Directly Attached Storage (DAS)
The storage evolution

- Small Computer System Interface (SCSI)

  - is a standard which define:
    - a command set
    - a protocol for transactions
    - A physical interface

  - is block-oriented, i.e. host’s o.s see the storage devices like a contiguos sets of fixed data blocks.
The storage evolution

- Low latency
  - through disk (~ms) and through cache (~μs)
- Very low error rate
  - Inefficient error recovery

Windows, Unix, Novell, MacOS, etc.
The storage evolution

Client

Local Network or Geographic

Server
The storage evolution

- Great difficulties on most important corporate richness’ management: the data
  - Resource administration must be done singularly
  - No optimization
  - Scalability
  - Performance
  - Limited maximum distance between devices
  - Inaccessibility to data during maintenance
  - Difficult backup management
The storage evolution

Solution:

- Separation between storage devices and computing resource
- Connection realized using network’s technologies
- Different implementations:
  - Costs
  - Performance
The storage evolution

Requirements:
- Consolidation of storage resources
- Centralized management
- Remote replication of data (disaster recovery)
- Centralized and transparent backup to LAN and computers
Network Attached Storage

- Object able to satisfy a *file* request received from the network
Network Attached Storage

- Usually a computer. In the future we will see more hardware based solutions
- Proprietary or heavily optimized operating system
- High storage capabilities
- Use of RAID and Hot-Swap to protect data and guarantee continuity of service
- Cost
- Use of the LAN
- File system invisible to client
Network Attached Storage

- Tipically TCP/IP over Ethernet
- File exported via NFS, CIFS (or both)
Storage Area Network

Client

LAN/WAN

Server

Storage Area Network

Disks
Storage Area Network

- Set of equipments and tecnologies to remote the storage in a network
- SAN features
  - Network resource used exclusively for the storage
  - SCSI protocol for end to end communication: minimum impact for DAS to SAN migration
  - Access to data through *logical blocks* and not to file
Storage Area Network

SAN features

- Requirement compatible to DAS technologies
  - High speed
  - Low latency
  - Very low error rate
- Compatibility with an high number of nodes
- Metropolitan distance coverage
- High reliability and ability to react to failures
Storage Area Network

- Used transport technologies:
  - Ethernet – TCP/IP
    - iSCSI
    - FC-IP
  - Fibre Channel

- Usable:
  - Infiniband (in the future?)
  - (Sonet/SDH)
  - (DWDM)
Storage Area Network
Benefits and of Ethernet – TCP/IP

- Network simplicity (! + !)
- Infrastructural and training costs are very low (!! + !!)
Ethernet – TCP/IP

- Benefits and lacks of Ethernet – TCP/IP
  - Evolution prospective uncomparable on respect to whatever rival technology (+)
  - No guarantees to receive transferred informations (! - !)
    - Frame loss is, today, a feature of Ethernet functioning
    - Error recovery relies on TCP
    - TCP hardware implementation is difficult
  - No guarantee on latency (! - !)
Fibre Channel

- Born from the needing of reliable support for serial Ultra3 SCSI
- Support high transfer rate
  - 1Gb/s, 2Gb/s, 10Gb/s
- Include a lossless mode
- Include a simple *data plane*
- The *control plane* is complex
Fibre Channel

- Direct connection
  - There is a full-duplex direct link between the two nodes
Fibre Channel

- Ring (Arbitrated Loop)
  - Up to 127 nodes connected in ring topology physically linked or through hub (better reliability)
Fibre Channel

- Meshed network (Switched Fabric)
  - Switches are linked to nodes and between themselves
  - Full duplex links

![Diagram of Fibre Channel network with switches and nodes connected in a meshed pattern]
Fibre Channel

- Protocol stack

```
+-----------------+   +-----------------+
|     SCSI        |  ->|        IP        |
+-----------------+   +-----------------+

| FC-4            |   | ULP Mapping |
| FC-3            |   | Common services |
| FC-2            |   | Signalling |
| FC-1            |   | Transmission |
| FC-0            |   | Physical layer |
```
Fibre Channel

- **FC-0**
  - Physical interface definition

- **FC-1**
  - Encoding and link’s low level control

- **FC-2**
  - End to end data transfer protocol
    - Frame format
    - Addressing
    - Segmentation
    - Flow control
    - Error detection/correction
Fibre Channel

- **FC-3**
  - Services common to every port
    - Cryptography
    - Compression
    - Channel bonding

- **FC-4**
  - Mapping between upper layer protocols and the transport layer for the delivery through the fabric
Fibre Channel

- Exist several kind of ports with specific functioning
  - N_port: HBA (Host Bus Adapter)
  - F_port: switches through HBA
  - E_port: connection between switches (ISL)
  - NL_port, FL_port: loop functioning
Fibre Channel

Diagram showing various port types:
- N_port
- F_port
- E_port
- NL_port
- FL_port
Fibre Channel

- **Addressing**
  - Nodes, ports and switches have a unique 64 bit ("World Wide Name") "Name identifier" which is assigned in factory
  - Dynamic assignment of 24 bit address for data exchanges

<table>
<thead>
<tr>
<th>Domain_ID</th>
<th>Area_ID</th>
<th>Port_ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>16</td>
<td>8</td>
</tr>
</tbody>
</table>
Fibre Channel

- **Addressing**
  - Domain_ID from 00h to EFh.
    - Usually one per switch. Assignment managed by a **main switch**
    - 239 switches supported
    - From F0h to FFh “Well Known Address”
      - Services offered by Fabric
      - Implemented with distributed protocols in the switch internals
  - Area_ID e Port_ID assigned to nodes. 65536 nodes per switch.
Fibre Channel

- Routing
  - FSPF: *Fabric Shortest Path First*. Link-state protocol similar to OSPF
  - FC doesn’t have TTL mechanism. Infinite loop of packets are possible
  - The network convergence must be as fastest as possible
Fibre Channel

- The communications between two nodes expects the an “exchange” opening
- Each exchange expects half duplex frame “sequences”
- Several kinds of communication
  - Flow control
  - Reservation of resources
  - Guarantee on ordered frame delivery
Fibre Channel

- Flow control
  - End to end
  - Buffer to buffer
  - Credits mechanism
Problematics

- Deadlock
- Traffic is blocked on the whole link due to lack of credits
Advanced aspects

- VSAN
  - Like VLAN, but on SAN
  - Interesting for Storage Providers
- Link Aggregation
- Load Balancing
Complexities and performances

- Frame Fibre Channel

<table>
<thead>
<tr>
<th>SOF</th>
<th>Header</th>
<th>Opt Hdr</th>
<th>Dati</th>
<th>CRC</th>
<th>EOF</th>
</tr>
</thead>
<tbody>
<tr>
<td>(4)</td>
<td>(24)</td>
<td></td>
<td>(da 4 a 2112)</td>
<td>(4)</td>
<td>(4)</td>
</tr>
</tbody>
</table>

- 36 byte di overhead
- Ethernet packet - TCP/IP
  - 18 byte Ethernet + 20 byte IP + 20 byte TCP
  - + iSCSI o FC-IP
Complexities and performances

- Fibre Channel vs. Ethernet
  - Performances
  - Guarantees
    - Investments protection
    - Developing
  - Prices
    - Target of consumers
Using scenarios

- **iSCSI**
  - Low cost connection of host (Ethernet) to SAN

- **FC**
  - Connection between servers and disk arrays

- **FCIP**
  - Connection of SANs through a geographic link
    - Es. Backup or redundancy
Main vendors

- SAN
  - Brocade
  - McData
  - Cisco
- NAS
  - NetApp
  - HP, Dell
- “turnkey” Solution
  - IBM, EMC
Bibliography

- Books

- Standard Fibre Channel
  - http://www.t11.org/index.htm
Bibliography

- Tutorials Fibre Channel
  - http://www.fibrechannel.org/OVERVIEW/
  - http://www.iol.unh.edu/knowledgeBase/training/fc

- Storage over IP

- Documentation from the top two vendors of Storage Switch equipments
  - http://www.brocade.com/san/resource.jsp