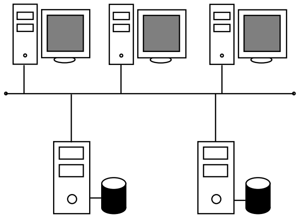
Storage Area Networks: Performance and Security

Presented by **Matthew Packard**

July 27, 2003

SAN Architecture - Definition & DAS Limitations

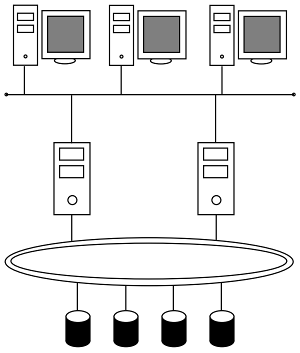
- Storage Area Network (SAN)
  - universal storage connectivity
    - free from interconnection implementation
  - dedicated storage network
    - reduces overhead on data networks
- Directly Attached Storage (DAS)
  - widely used - host centric storage
  - high overhead on data networks
  - failover/clustering more difficult and expensive



DAS Storage Model

SAN Architecture (Cont) - Storage Design & Applications

- Storage and data traffic isolation
  - out of band signaling
- Based on high capacity, redundant links
- On-the-fly storage allocation
  - plug, configure, mount
- Centralized backups
  - fast, one stop repository
- Easy clustering
  - all hosts see same data, same view
- Easy failover
  - with volume managers, swap mounts and run



SAN Storage Model

SAN Architecture (Cont) - SAN vs NAS

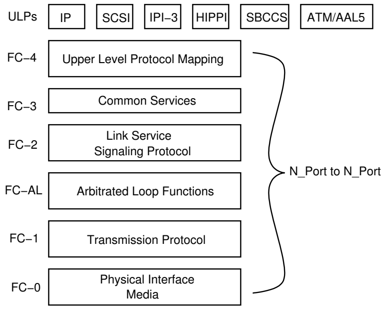
- Network Attached Storage (NAS)
  - similar to SAN
    - direct network connection
  - uses TCP/IP protocol
  - internal filesystem
    - shared to remote hosts (NFS/CIFS)

- SAN
  - direct network connection
  - uses FC with encapsulated SCSI commands
  - no internal filesystem
  - relies on controlling host for representation

SAN Architecture (Cont) - Interconnection: Fibre Channel

- Fast: 100Mb/s - 3.2Gb/s up to 10km

- FC de facto standard
  - direct connect (N_Port)
  - arbitrated loop (FL_Port)
    - FC-AL giant bus
  - switched fabric (F_Port)
    - fast
  - bridging for SCSI devices
  - FC-0 (physical)
  - FC-1 (error-free conditioning)
  - FC-2 (most important)
    - framing, flow control, segmentation, errors
  - FC-3 (striping)
  - FC-4 (ULP)



ULPs: IP, SCSI, IPI–3, HIPPI, SBCCS, ATM/AAL5

FC–4 Upper Level Protocol Mapping

FC–3 Common Services

FC–2 Link Service
Signaling Protocol

FC–AL Arbitrated Loop Functions

FC–1 Transmission Protocol

FC–0 Physical Interface
Media

N_Port to N_Port

Fibre Channel Hierarchy

SAN Architecture (Cont) - Interconnection: iSCSI

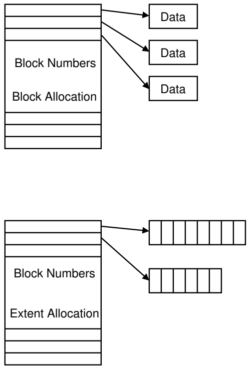
- SCSI over IP
- Slow, since it uses software stack conversions
- Best for sites using existing wiring plants and long distance storage

SAN Architecture (Cont) - Interconnection: Infiniband

- Intel led
  - adds ASIC support for SAN technology in processors
- x86 OSes have SAN support built-in for little cost
  - Linux, Solaris, and Windows
- Replacement for PCI
- SAN and NAS integration with VIA

SAN Performance - Filesystem Performance: UFS

- UNIX Filesystem (UFS)
  - support under Solaris, Linux, BSD, AIX, HP-UX
- Metadata logging
  - transaction rollback on mid-write failure
  - good for large volumes - no fscking
- Block allocation
  - disk block allocated per requested data block



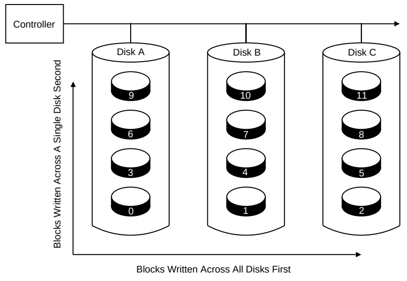
Block vs Extent Allocation

SAN Performance (Cont) - Filesystem Performance: VxFS

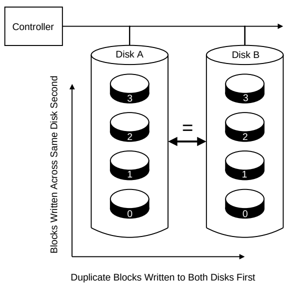
- Another UNIX filesystem
- Part of Veritas suite - Volume Manager add on
- Full data and metadata logging
  - data can be rolled back or forward with logs
- Extent allocation
  - series of blocks allocated per requested write
  - blocks accessed as offset from master block
  - slower than UFS for heavy random I/O

SAN Performance (Cont) - Hardware RAID

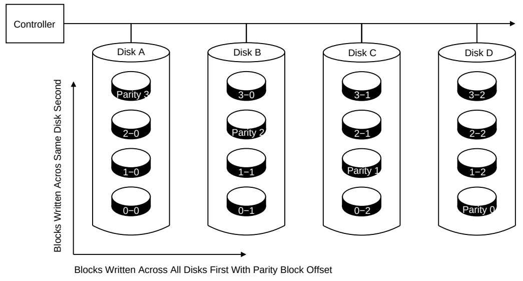
- Redundant Array of Inexpensive Drives (RAID)
- Performed on storage array controller
  - very fast, depending on RAID type
- Various RAID levels (0, 1, 3, 4, 5, 6, 0+1, 1+0)
  - no one better than another
  - based on performance and failure resilience tolerances
  - RAID-0 (striping, no failure tolerance, fast)
  - RAID-1 (mirroring, can lose one drive, fast reads)
  - RAID-5 (distributed parity, one drive, fast reads)
  - RAID-0+1 (mirrored stripes, one drive, fast r/w)
  - RAID-1+0 (striped mirrors, one drive, fastest r/w)


Controller
Disk A
Disk B
Disk C
9
6
3
0
10
7
4
1
11
8
5
2
Blocks Written Across A Single Disk Second
Blocks Written Across All Disks First

RAID-0: Disk Striping


Controller
Disk A
Disk B
3
2
1
0
=
3
2
1
0
Blocks Written Across Same Disk Second
Duplicate Blocks Written to Both Disks First

RAID-1: Disk Mirroring



RAID-5: Disk Striping w/ Dist. Parity

SAN Performance (Cont) - Volume Management

- RAID configuration through software
- Works on host, rather than on storage
- Slower than hardware RAID, but more options
  - tighter volume creation parameters
  - cluster support
  - failover support

SAN Performance (Cont) - Backups: Online

- Performed live on SAN storage array
- Incurs heavy I/O penalties due to at least two read requests
- Does not require separate storage mechanisms or hardware
- Cannot deal with open files (databases)
  - open file agents can, but not well

SAN Performance (Cont) - Backups: Third Mirror

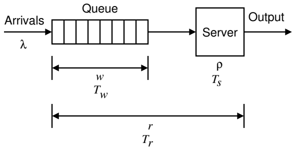
- Volume management intervention
  - regular RAID-1, with additional mirror set
- Data synched, then split for backup
- Greatly reduced I/O for backup
  - not performed on actual production storage array
- Still lacking in open file backups

SAN Performance (Cont) - Backups: Frozen Image

- Succeeds in backing up open files
- Applications must be backup-aware
- Apps go in hot backup mode during backup
  - data files in consistent, quiet state
  - must cache client data requests during backup
- Oracle 2 minute default, then clients time out

SAN Performance (Cont) - QT: One Server & Queue

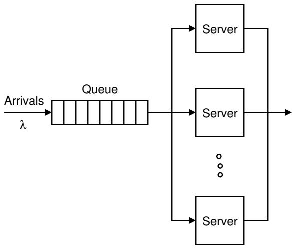
- Queueing theory - study queues, determine performance
- Arrival rate λ
- Queue items w
- Server utilization ρ
- Total items r
- Avg. time in queue, server, overall T_w, T_s, T_r



Single Queue Single Server Model

SAN Performance (Cont) - QT: Multiserver Single Queue

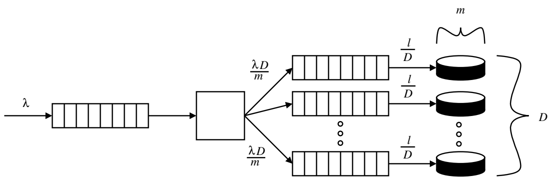
- Performance increased over single server model
- Each server receives percentage of λ
- Bank line with multiple tellers example



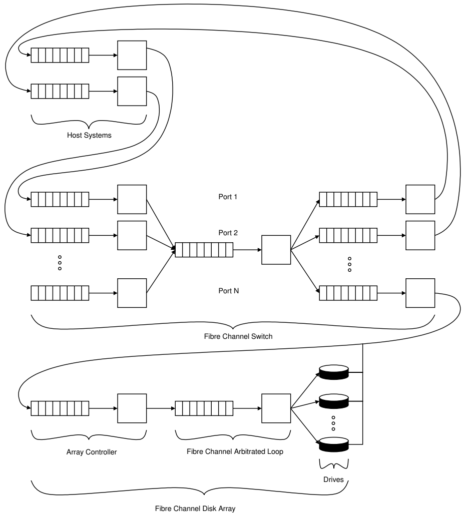
Single Queue Multiple Server Model

SAN Performance (Cont) - QT: Modeling Disks and Nets

- ❖ Extending QT to disk subsystems and networks
- ❖ SCSI array controller and disks have queues
- ❖ FC switch, has queues per interface
- ❖ FC array has controller, FC-AL, and disk models



Disk Array Model



Storage Network Model

SAN Security - Zoning, LUN Masking & Mapping

- Limit the access to SAN storage
  - goes against complete storage visibility
  - necessary for security, software access mechanisms

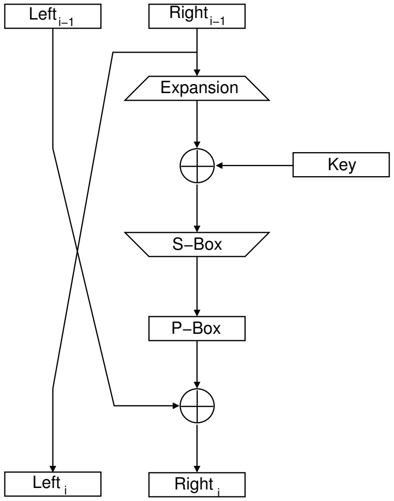
- Zoning lets switches determine which ports can talk to other ports

- LUN masking lets array controllers determine which LUNs are visible to a port
  - single RAID device can contain multiple volumes (LUNs)

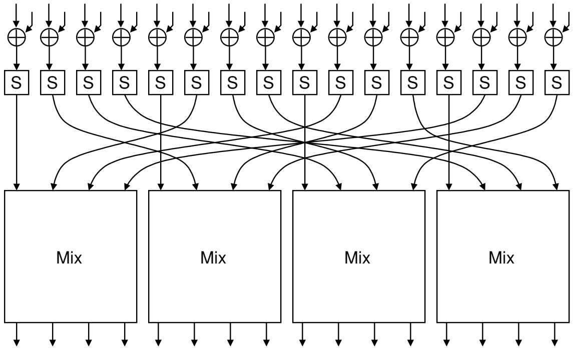
- LUN mapping lets host SAN drivers limit OS disk driver's access to storage

SAN Security - Cryptography: Private Key

- Cryptography - obscure data through math functions
- Private key crypto - same en/decryption key
  - Fast, but hard to distribute and manage key securely
- Data Encryption Standard (DES)
  - government standard since 1977
  - block cipher, 64 bits, 16 rounds, symmetric
  - aging, crackable, 56 bit key, slow in software
- Advanced Encryption Standard (AES)
  - opened to public for submission
  - Rijndael accepted as standard (Twofish, Lucifer)
  - fast with small memory footprint, 16 byte block size
  - 10-14 round, 128 - 384 bit key



Single DES Round



Single AES Round

SAN Security (Cont) - Crypto: PubLic Key & Key Exchange

- Separation of encryption and decryption keys
  - public key published for all to use (encryption/signature verification)
  - private key held by user (decryption/signature hashing)

- RSA (Rivest, Shamir, Adleman) most widely known
  - security lies with factorization of huge integers with only two, non trivial factors
  - patent expired recently - freely available now

- Diffie-Hellman key exchange allows for key swapping over insecure channel
  - solution for private key sharing

SAN Security (Cont) - Cryptography: Digital Signatures

❖ Alice writes data, Bob wants to verify it's from her, and was not tampered with

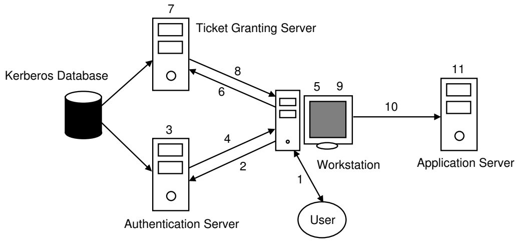
❖ Alice hashes data, encrypts with private key (signature), attaches to data

❖ Bob retrieves Alice's public key, decrypts hash, computes hash, compares both

❖ If they match, data is valid and belongs to Alice

SAN Security (Cont) - Kerberos

- Kerberos developed for use in MIT's Athena project
- Allows users to authenticate to a realm
  - without revealing passwords
- Authentication based on tickets
  - granted by TGS and AS for application server
- Beats MS domain authentication schemes
  - L0phtcrack?



Kerberos Authentication Procedure

Summary

- SANs require careful planning with focus on performance and security
- Very high speeds over redundant links
- Dynamic storage allocation
- Separation of storage/control traffic

Discussion

❖ Questions?