Quantum

CAPACITY FOR THE EXTRAORDINARY™

Network-Attached Storage Devices: Is It Time For A New Storage Paradigm ?

Paul Massiglia
Network Storage Strategist
Quantum Corporation
500 McCarthy Boulevard
Milpitas, CA 95035
paul.massiglia@quantum.com

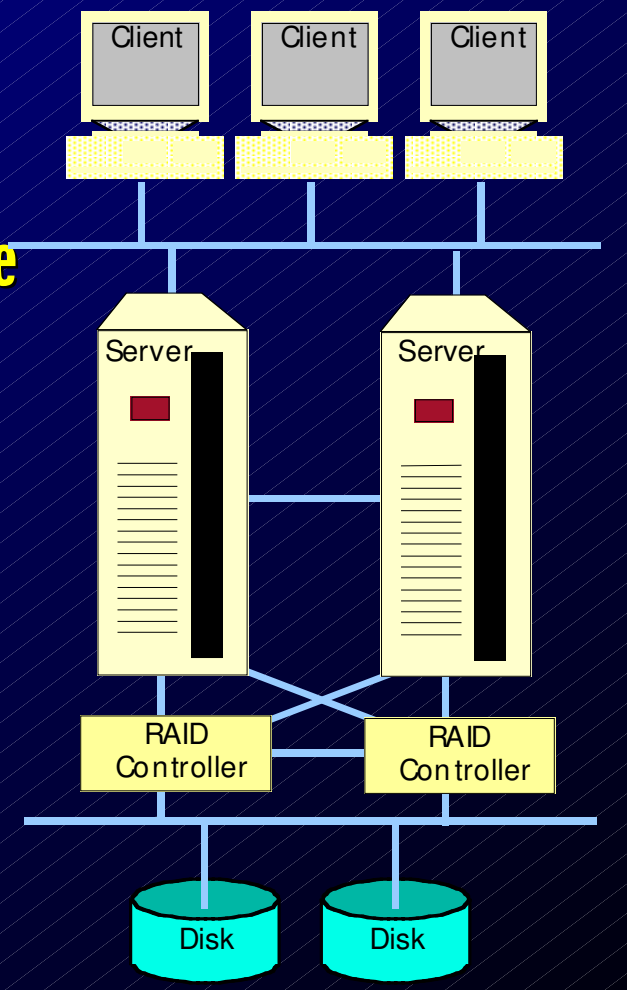
Quantum
CAPACITY FOR THE EXTRAORDINARY™

Today's I/O Systems

- A lot that's good:
  - cost-effective generic devices
  - mature, robust interconnects and software
  - well understood failure tolerance model
  - mature device and file models

but...

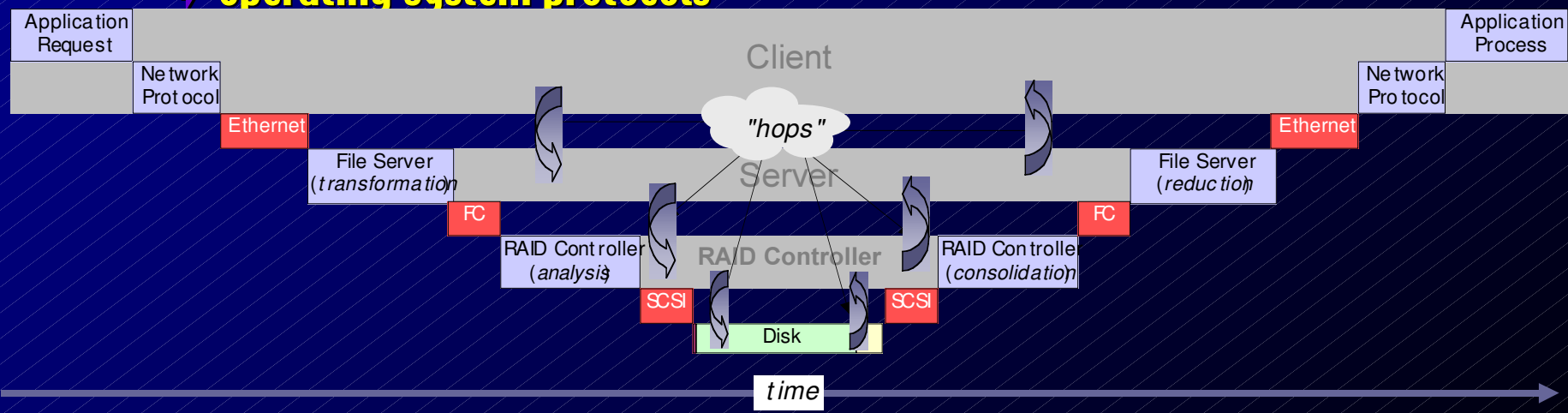
- The model is not tracking technology
  - system topologies
  - wire speed utilization
  - device capabilities



Quantum
CAPACITY FOR THE EXTRAORDINARY

Shortcoming 1:
Too Many "Hops" in the I/O Path

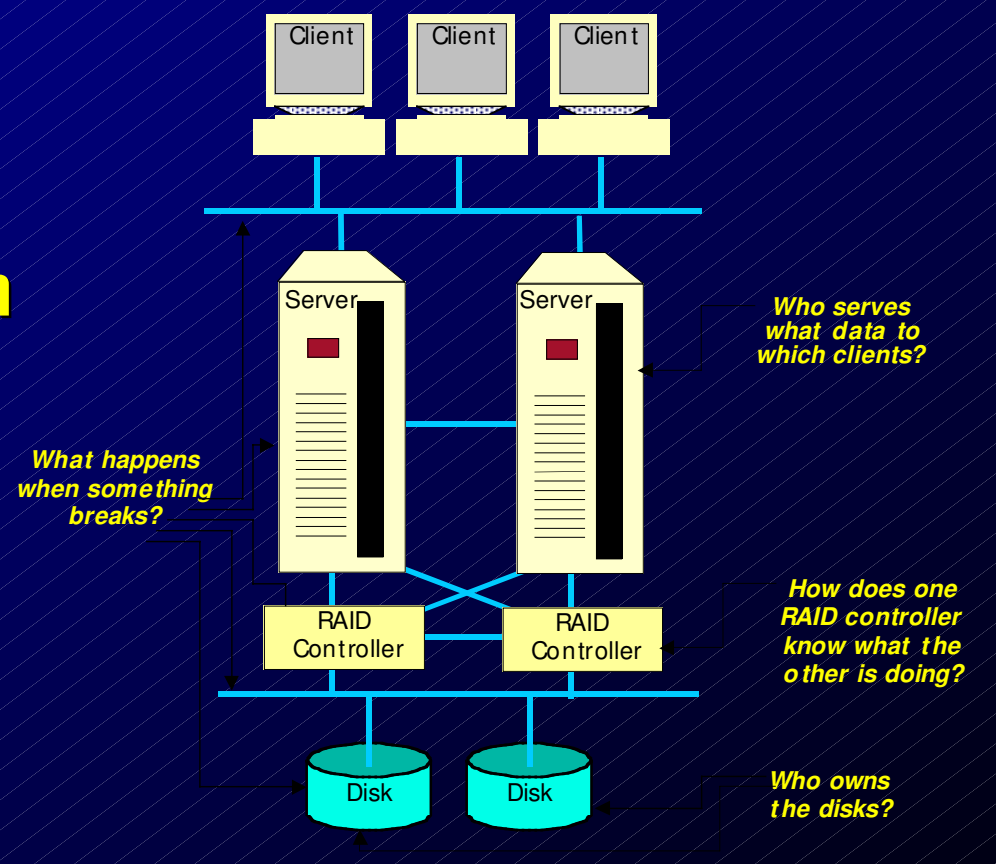
- Each hop represents
  - a trip across a wire.....fast and getting faster
  - a "store and forward" step..... minor slowdown
  - a "software stack"..... bottleneck
    - wire protocol conversion
    - operating system protocols



Quantum
CAPACITY FOR THE EXTRAORDINARY™

Shortcoming 2:
Not Well Suited To Clusters

- Servers must be involved in
  - data ownership
  - data access synchronization
  - I/O load balancing
  - failure recovery
- This limits
  - scaling
  - performance
  - availability



Enabler 1: Fibre Channel

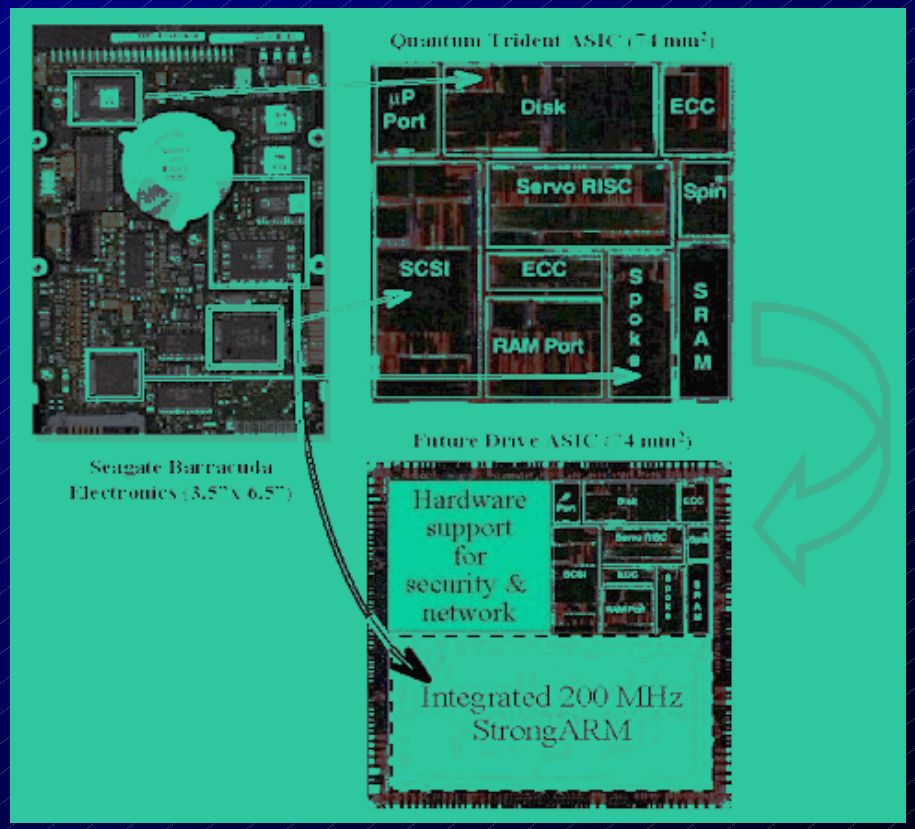
- Performance
  - bandwidth: 100 MBPS
  - latency: device-to-device data transfer
- Availability
  - dual loop, "hot" plugging, zoned switching
- Scaling
  - more devices, longer distances

- These have combined to create the "SAN" concept
- Question: does the conventional "block server" disk model utilize SANs effectively?

Quantum
CAPACITY FOR THE EXTRAORDINARY™

Enabler 2: "Smarter" Devices

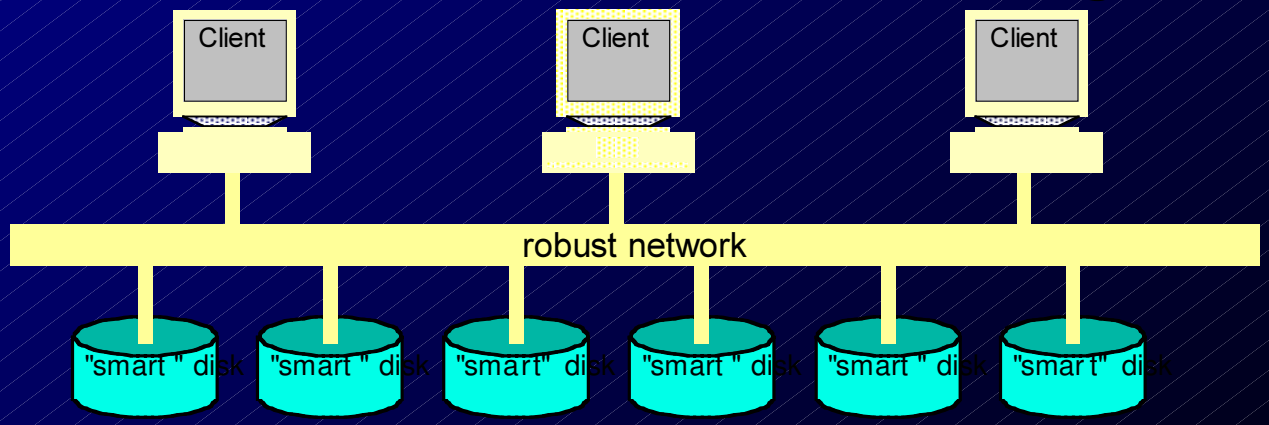
- Academic & industrial research
  - SWARM --U of AZ
  - GFS--U of Minn
  - Petal/Frangipani--DEC
  - NASD/Active Disks CMU
  - NASD/Active Disks CMU
  - etc.



Quantum
CAPACITY FOR THE EXTRAORDINARY™

Network Attached Secure Devices

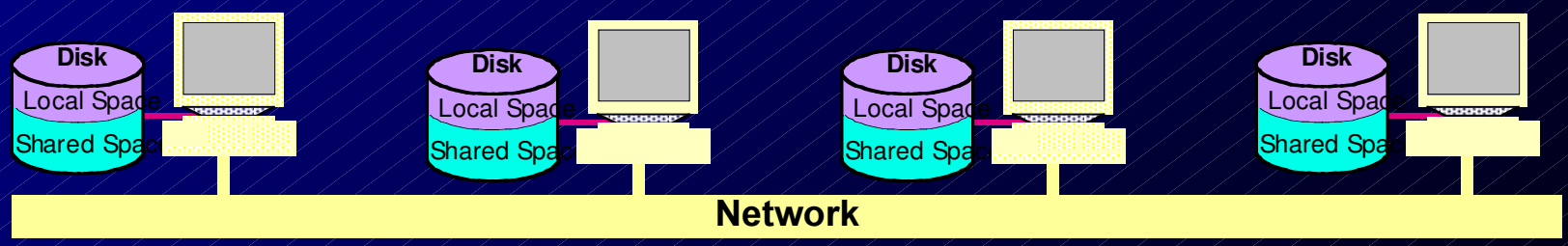
- Principles
  - make the disks "smarter"
  - wire clients directly to data
- Challenges
  - keeping data secure
  - using the "smart" in the smart disks beneficially



Quantum
CAPACITY FOR THE EXTRAORDINARY™

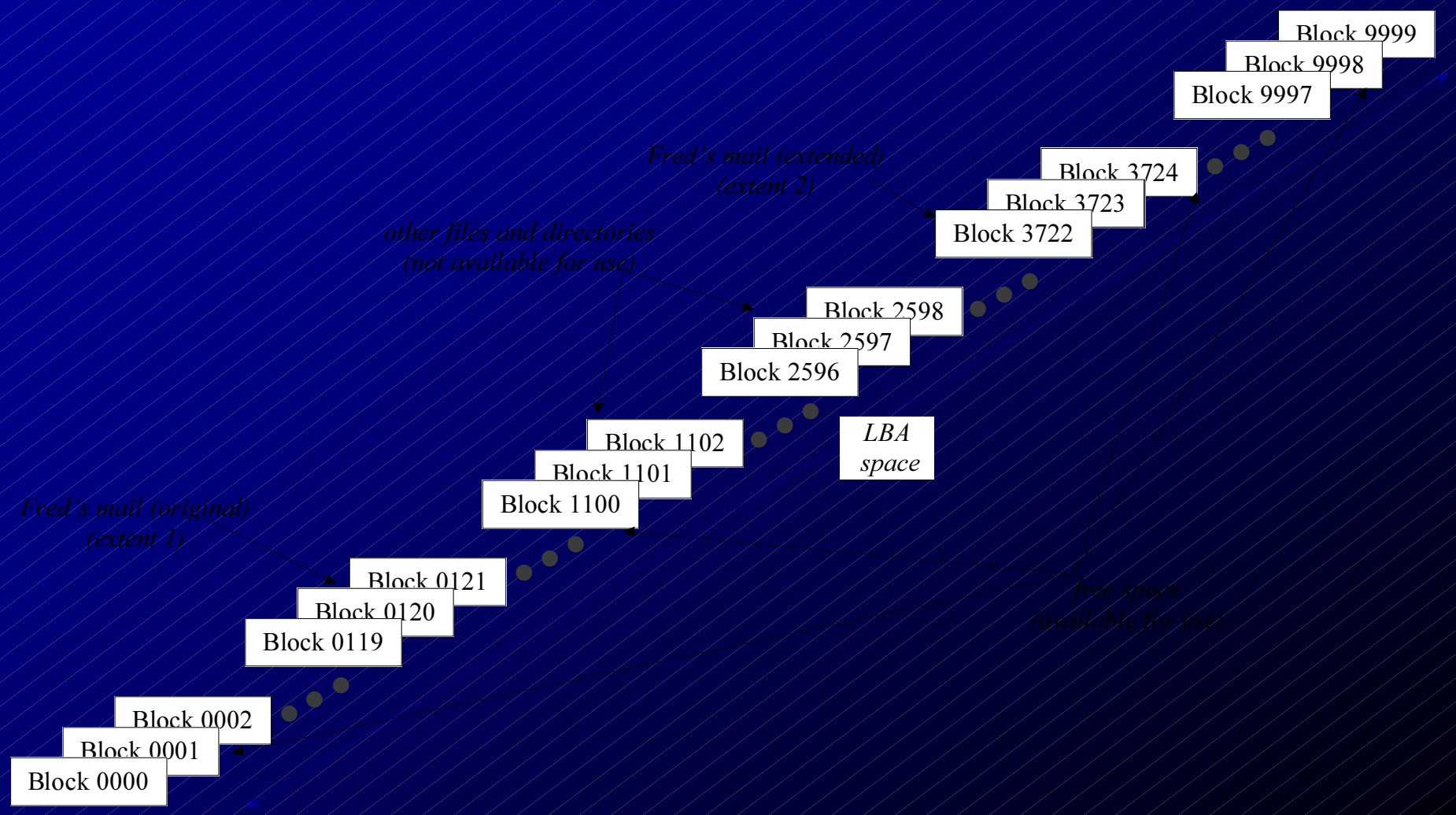
A Precursor: MangoSoft's *Medley*

- "Server-less" client-server computing
  - dynamic disk resource sharing
  - global file system
  - redundancy--mirroring of data objects
- Approximates the NASD model
  - disk intelligence is in the workgroup members



Quantum
Capacity for the Extraordinary™

File System Functions
Disk Space Management

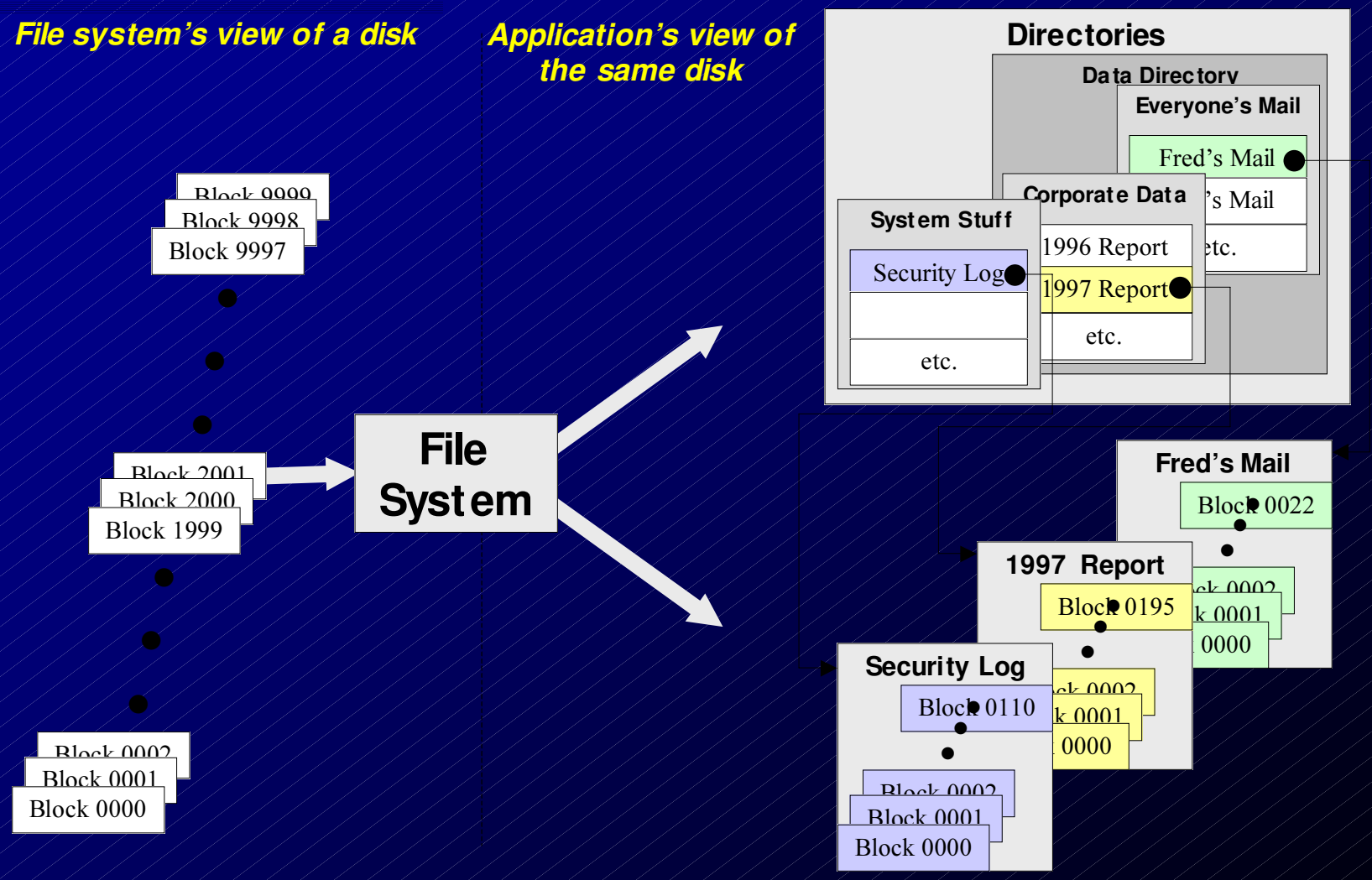


Quantum
Capacity for the Extraordinary™

File System Functions
Namespace Management

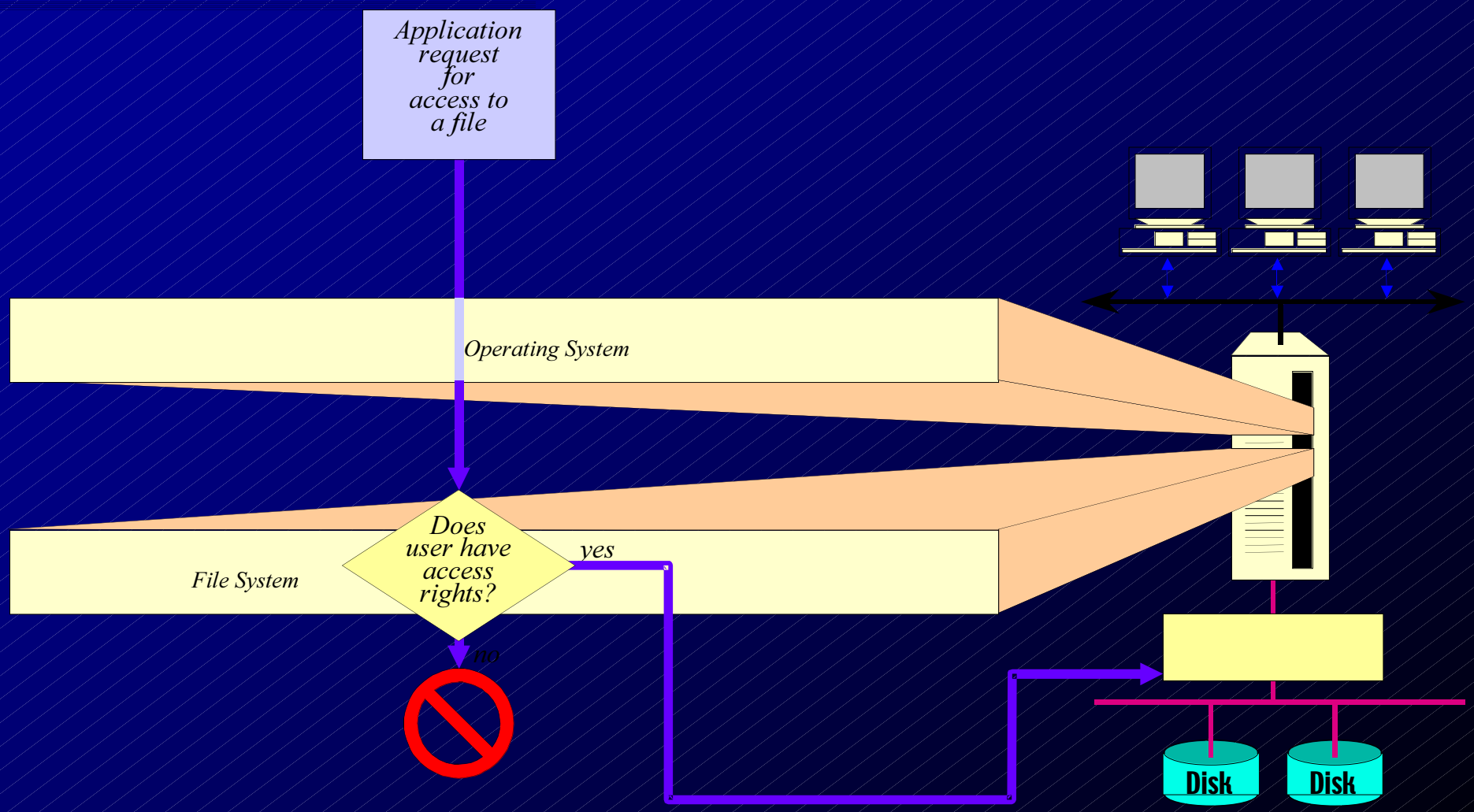
File system's view of a disk

Application's view of the same disk



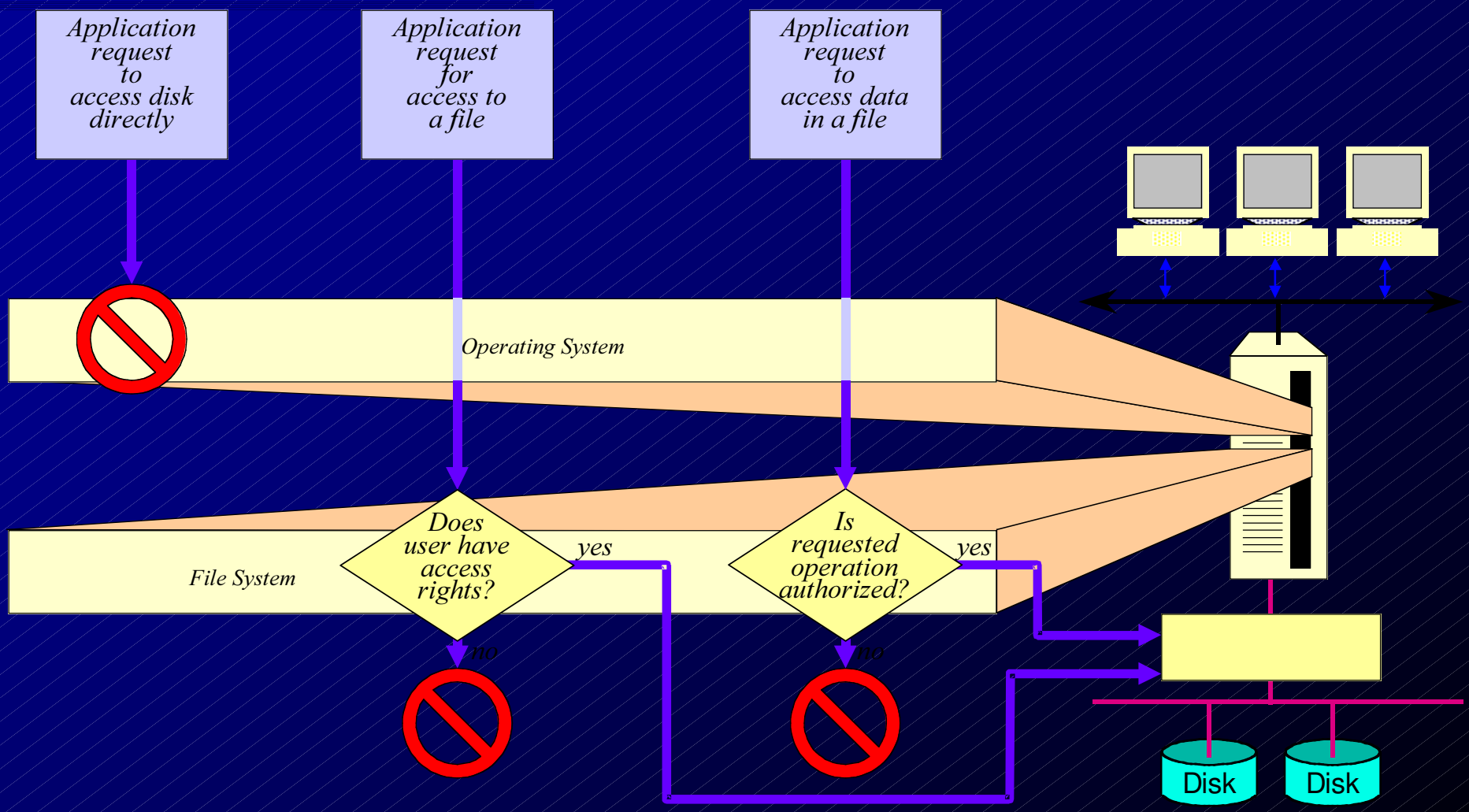
Quantum
CAPACITY FOR THE EXTRAORDINARY™

File System Functions
File Access Control



Quantum
CAPACITY FOR THE EXTRAORDINARY™

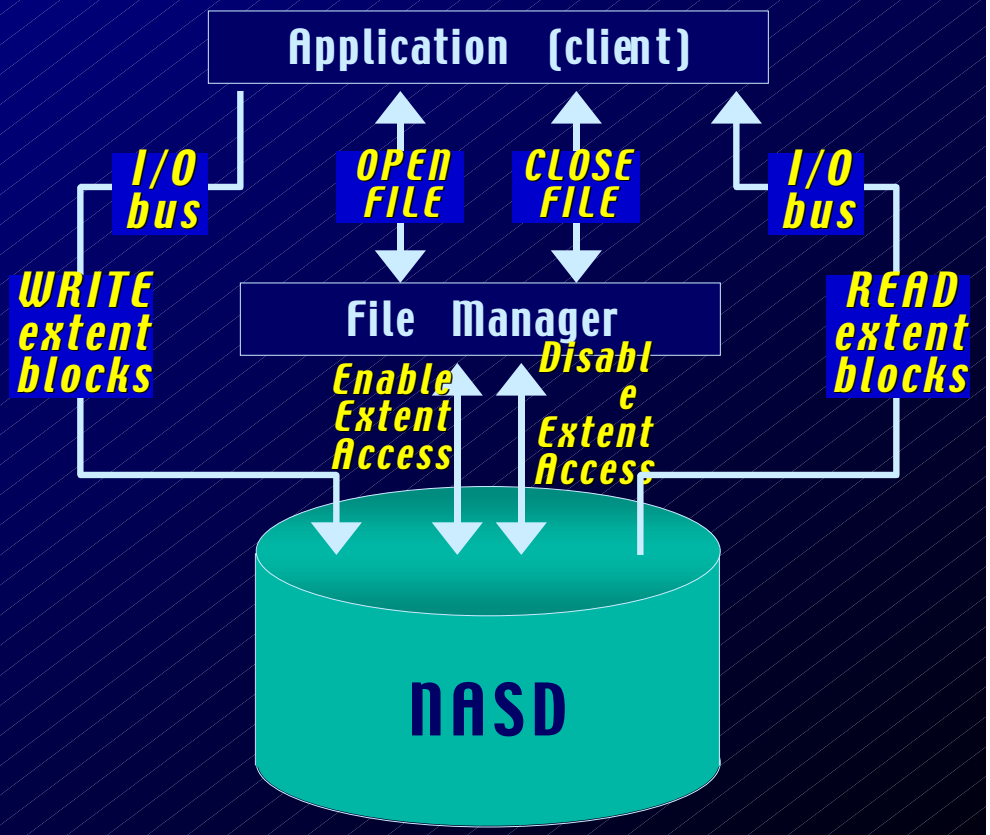
File System Functions
Data Access Control



Quantum
Capacity for the Extraordinary

NASD: The Basic Idea

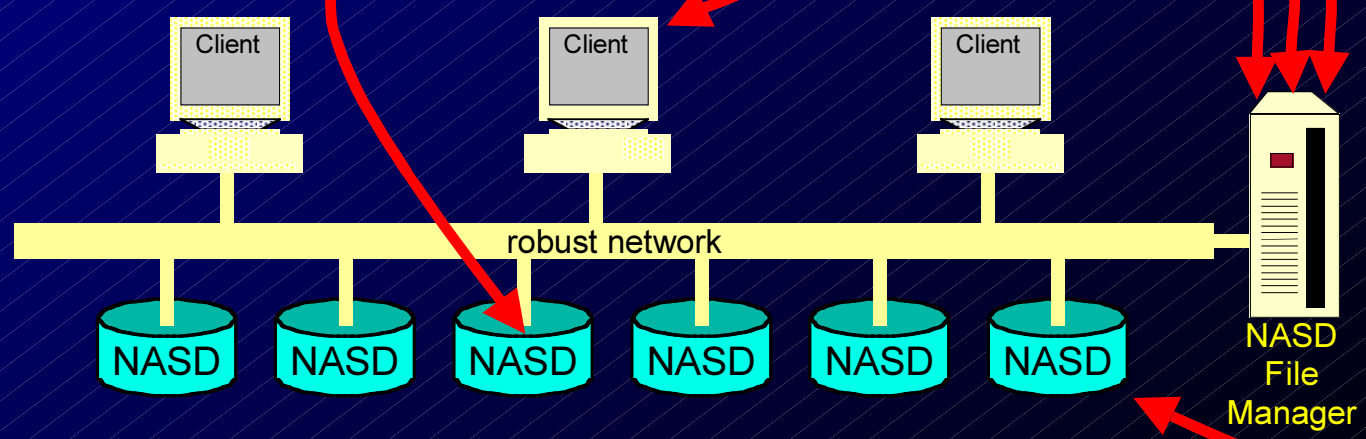
- A NASD "knows about" extents
- A "NASD File Manager" grants a client access to an extent
- Data transfers directly between NASD and client



Quantum
CAPACITY FOR THE EXTRAORDINARY

NASD Functional Partitioning

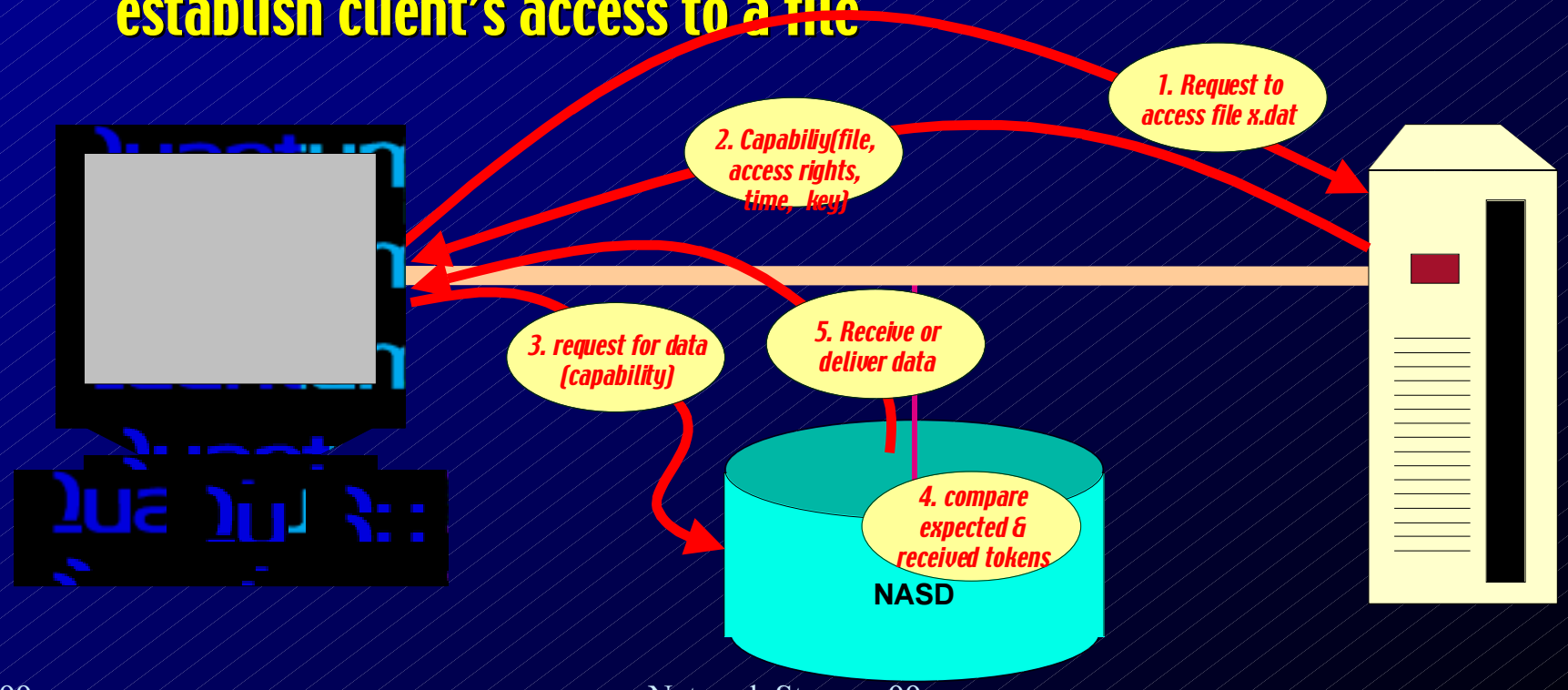
- Disk space management
- Namespace management
- File access control
- Data access control



Quantum
CAPACITY FOR THE EXTRAORDINARY

The "Secure" in NASD

- File Manager and NASD share a secret
- Secret is encrypted and handed to client
  - Key point: File manager and NASD need not communicate to establish client's access to a file



NASD Benefits

- Shorter I/O paths
  - data moves directly between storage and client
  - direct device-to-device transfers are possible
- "Smart" devices should be able to self-optimize
  - for performance
  - for failure tolerance
  - for security
- No loss of function
  - all conventional file system functions can be implemented without negating NASD advantages

Making NASD Happen

- Technical challenges
  - failure tolerance
    - server
    - network
    - devices
  - squeezing NASD functionality into disk-like cost
- Non-technical challenges
  - installed base transition from the block server model
  - motivation for the incumbents
    - RAID, OS, file system and database vendors

Making NASD Happen

- There is progress
- NSIC (National Storage Industry Consortium) working group
  - "NSIC/NASD"
  - members: CMU, Seagate, Quantum, IBM, StorageTek
  - meeting ~monthly to develop a NASD-like standards proposal
    - called "Object Based Disks"
    - roughly: SCSI commands for NASD function with security deferred
  - ~quarterly public meetings (next: Millbrae on 8/17/99)
  - target: ANSI X3T10 proposal in fall 1999

What's "Beyond" NASD

- NASD isn't even "here" yet—why ask what's beyond it?
- NASD does two important things
  - replaces the conventional block access disk paradigm
  - demonstrates the usefulness of intelligence in storage devices
- *but...* a NASD is still a fixed-function device

Beyond NASD

- What if storage device behavior could be adjusted in the field ?
- Would obviate questions of the "right" device capabilities
  - if a device's capabilities don't suit an application, download different capabilities
- Supporting research
  - CMU: Active Disk Project
  - UCB: "I-Disk" Project
  - Microsoft Research Group proposal: "Cyberbricks"

Sample Uses for "Active" Disks

- Data warehousing
  - 1000's of active disks search themselves simultaneously
- Database
  - 1000's of little transaction engines
- HSM/backup
  - each disk figures out what data needs to be moved or backed up
- Application customizations
  - "smart" caching, metered delivery, ...

Conclusion

- It's time for the block access disk to evolve
  - the technology is there
  - the need is there
- CMU NASD work shows the possibilities
  - potential benefits are great
    - performance, scaling, system cost, upgrade granularity
  - significant challenges remain
- The question:
  - will NASD become the next "disk", or is a more radical storage device paradigm change the right step?

Quantum
Capacity for the Extraordinary™

NSIC/NASD and SNIA

- NASD Goals
  - timeliness: get a proposal to X3T10 in 1999
  - quality
- Aligning with SNIA could be beneficial
  - lots more perspectives
  - potentially lots more "cycles"
- Aligning with SNIA could be disruptive
  - bringing the new people up to speed

- My "charter" from NSIC/NASD
  - get us aligned without delaying the submission