

PetaSTAR

A Real World Data Storage and Management Solution

James M. Johnson

Ovation Data Services, Inc.

10650 Haddington Drive • Houston TX • 77043-3229

Phone: +1-713-464-1300 • FAX: +1-713-464-1615

e-mail: jimjohnson@ovationdata.com

Web: <http://www.ovationdata.com>,

Presented at the THIC meeting at the Nassau Bay Hilton Marina, Houston TX,
March 30, 1999

E & P Data Growth

Problems

E & P companies are acquiring & interpreting increasingly larger volumes of seismic data

Increasing complexity as old solutions are scaled-up

Challenge

Keep seismic assets available to explorationists on a timely, intuitive and seamless basis

Result...

- Disk is always full everywhere
- Constantly faced with buying more storage capacity
- Reduced productivity throughout the entire seismic life cycle
- Cost of storage doubles every year

Barriers to Data Access

- Technological Progress
- Device Obsolescence
- Media Obsolescence
- Media Degradation
- Data and Intellectual Growth

Barrier - Technological Progress

Description

- New Application Releases
- New file formats
- O/S Vs. Release Incompatibility

Mitigation Strategy

- Upgrade applications more than 1 release out
- Open/Save data files with new release

Barrier - Device Obsolescence

Description

- OS Driver Vs drive firmware mismatch
- OS Driver features dropped
- OS support dropped for interface

Mitigation Strategy

- Maintain current drive firmware
- Closely O/S releases for interface support

Barrier - Media Obsolescence

Description

- Short product life for drive technology

Mitigation Strategy

- Choose media with broad acceptance
- Copy to new media type as necessary
- Utilize translation specialists as needed

Barrier - Media Degradation

Description

- Tape stiction
- Loss of oxide coating
- Edge damage
- Contamination

Mitigation Strategy

- Test data and media integrity
- Copy to new media as necessary

Barrier - Data and Intellectual Growth

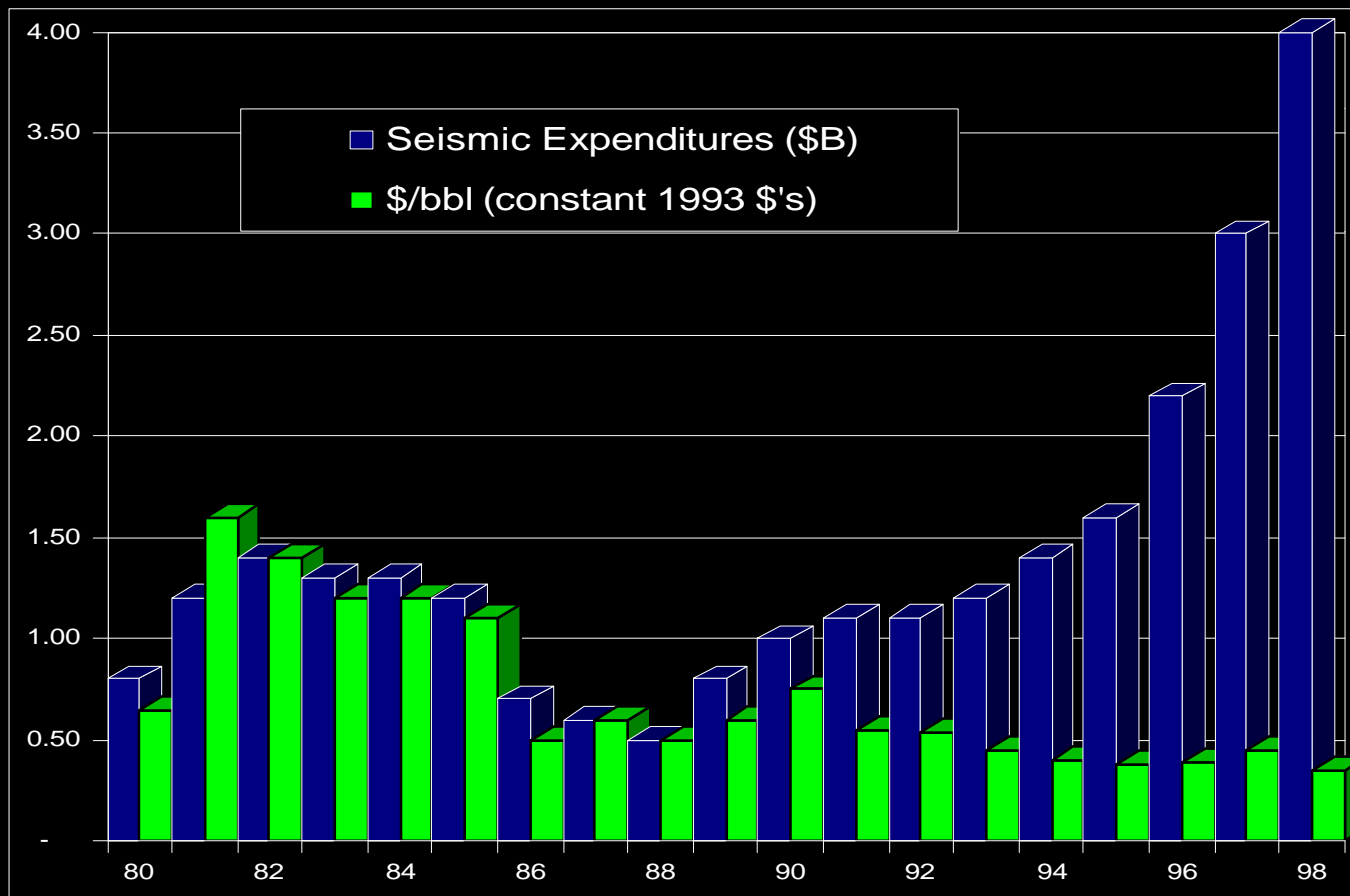
Description

- Exponential growth in data volume
- Costs double yearly despite price declines
- New data sources and applications
- Wider and wider geographic access

Mitigation Strategy

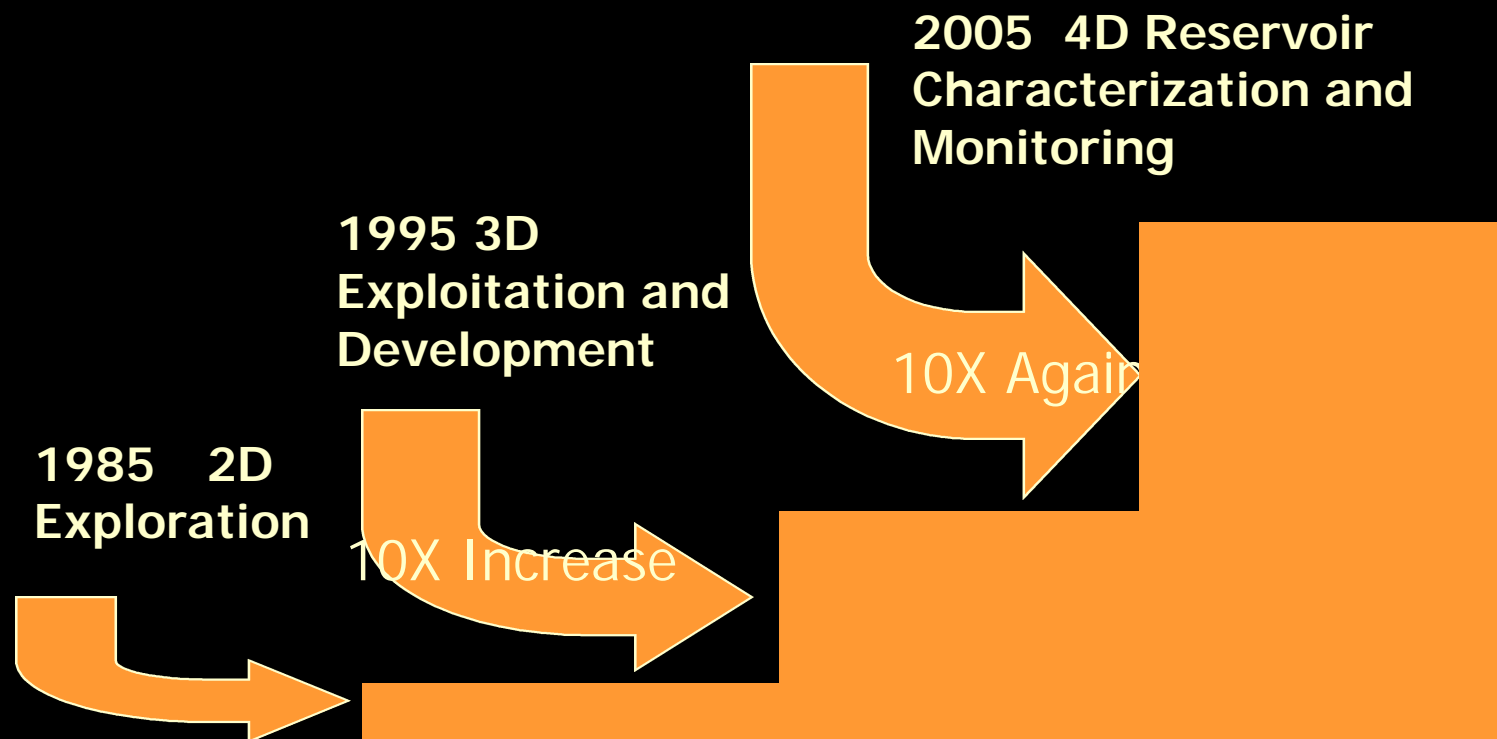
- Plan for increasing volumes of information
- Automate media handling
- Use mixed media solutions to minimize cost

Dramatic Increase in the volume of Data



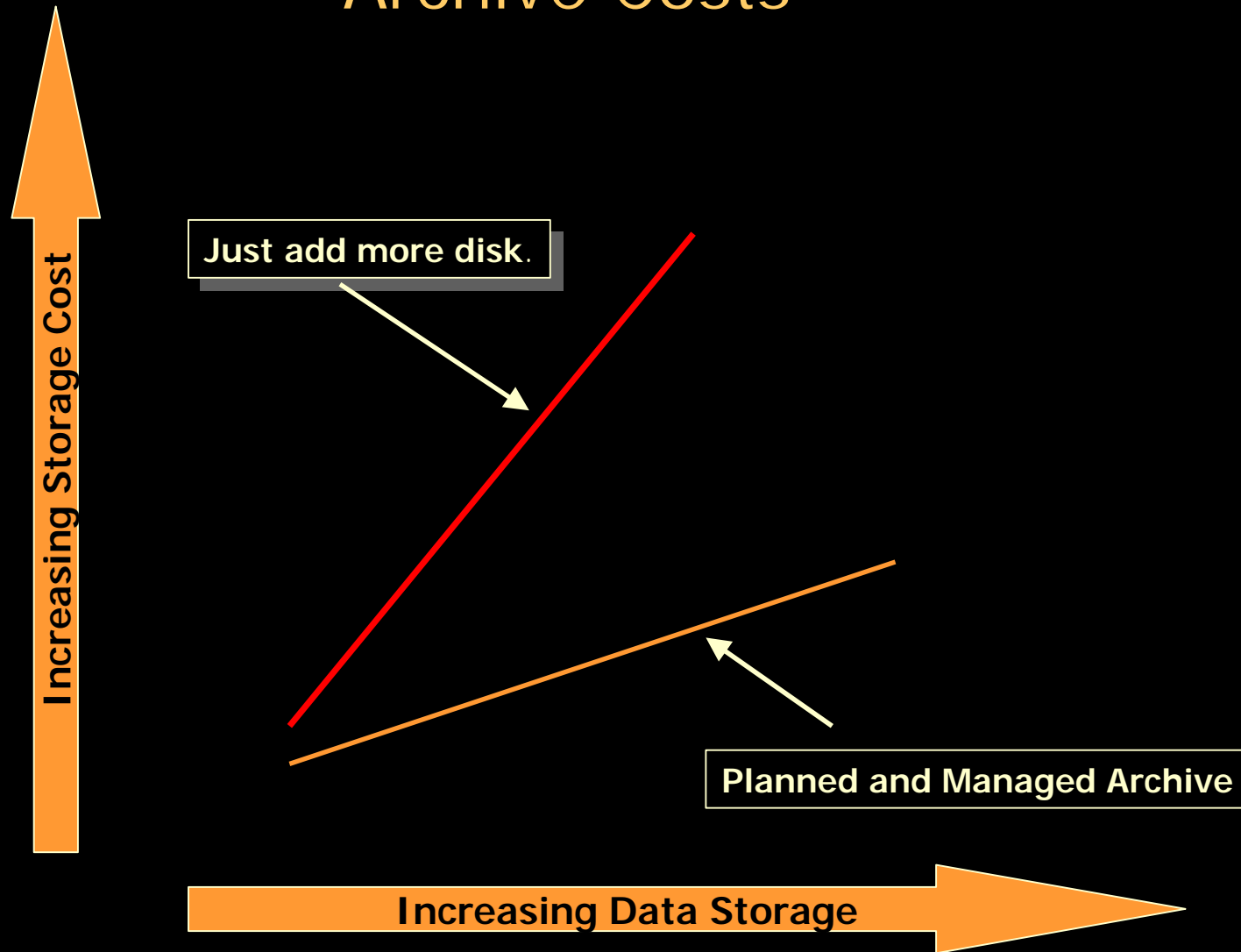
Data Volume Drivers

Seismic Expands into Life Cycle Oilfield Development



“Over the past 10 years the amount of seismic data acquired on just one kilometer has increased 40 fold”

Archive Costs



Data Access Barriers Timings

Barrier Driver

- Technological Progress
- Device Obsolescence
- Media Obsolescence
- Media Degradation
- Data and Intellectual Growth

Cycle Time

2-3 years

3-5 years

3-5 years

5+ years

Continuous

Solution Requirements...

- Easy integration with legacy systems
- Seamlessly cache seismic files into an application workflow automatically
- Integrate large volumes of tape and disk storage transparently
- Must simplify and streamline the centralized management and delivery of the right data to the right user in the right time
- Remove data quality and data loading concerns for end users
- Must be scaleable
- **Cost effective!**

The Solution...

PetaSTAR™

... A turnkey solution combining the latest in near-line data storage, seg-y data management, and file management.

Turn-key System Integration,
service and support



Data Management and Q/C
Panther SDMS and SDL



File Management - SamFS

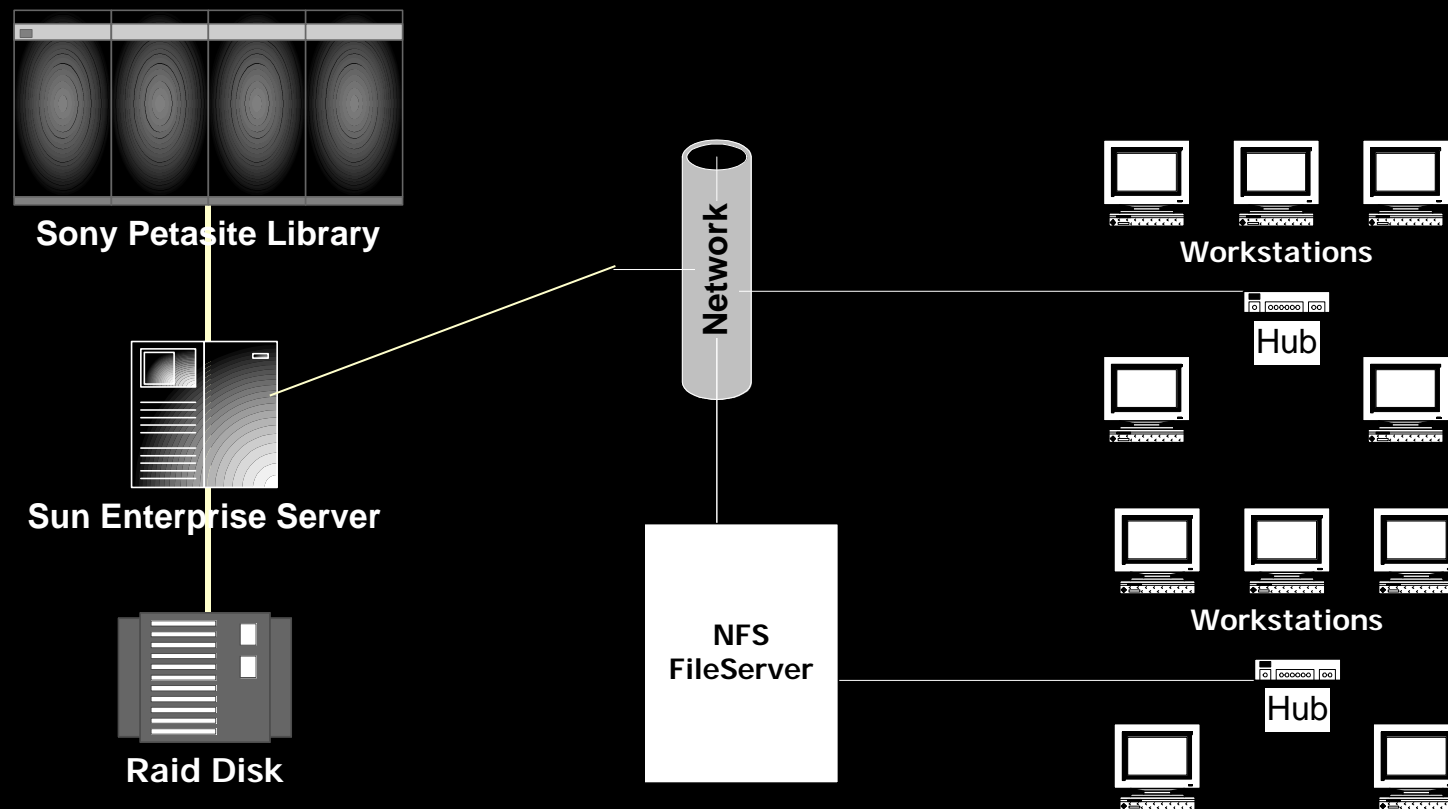


Hardware - Enterprise servers
and Sony Mass Storage

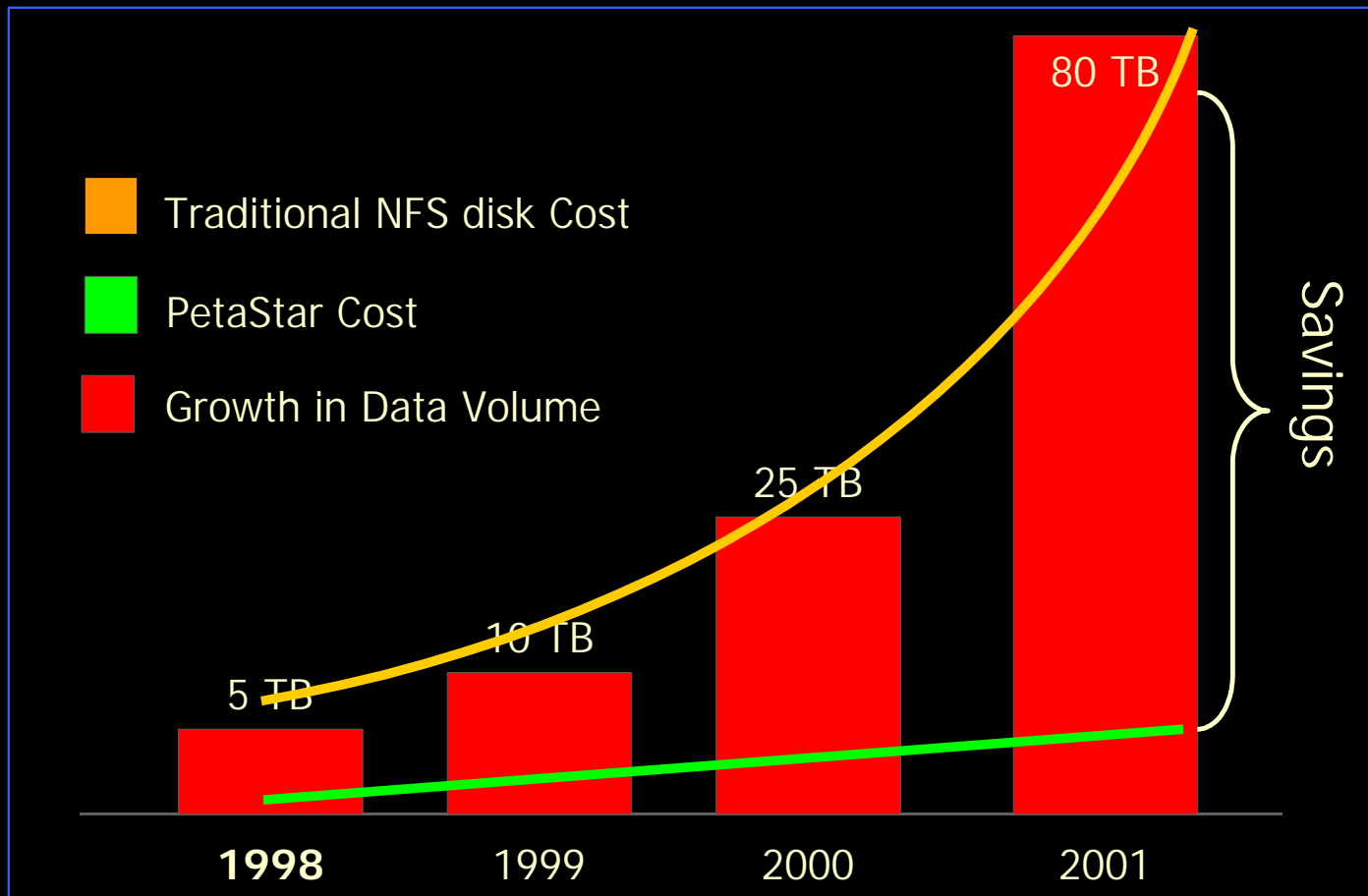
SONY



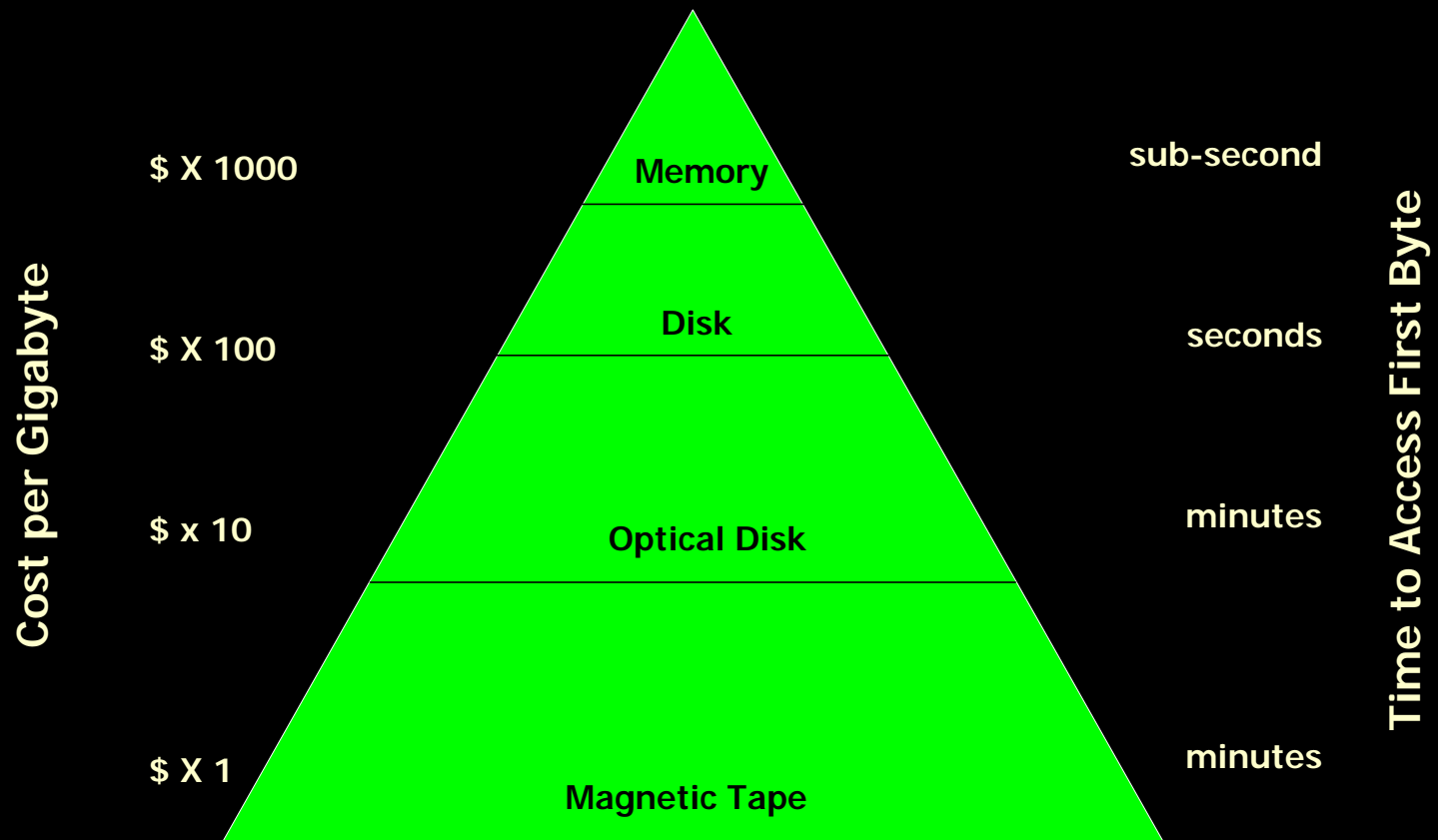
PetaSTAR configuration



Disk Vs.. Tape Cost Picture



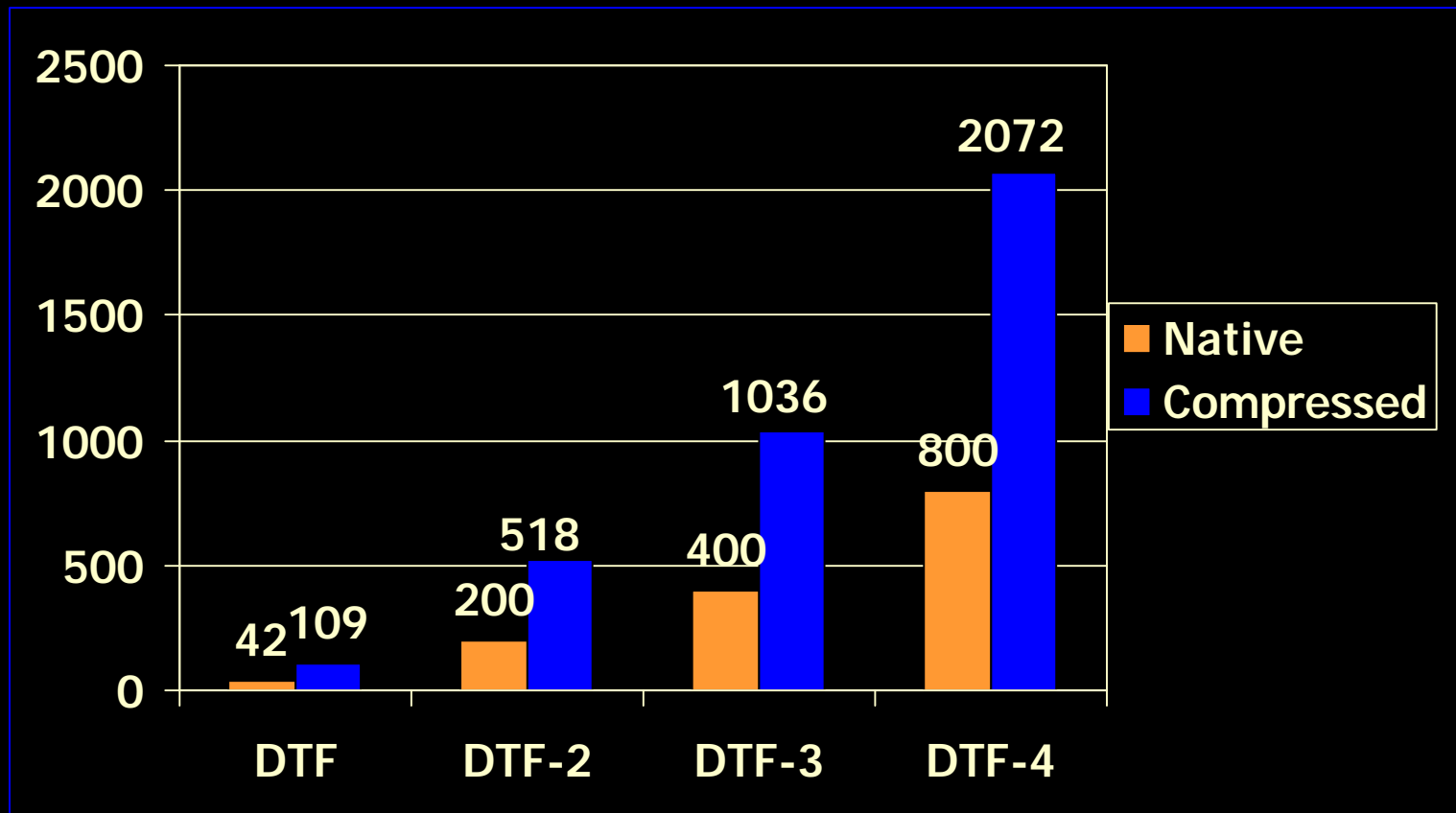
Storage Paradigm



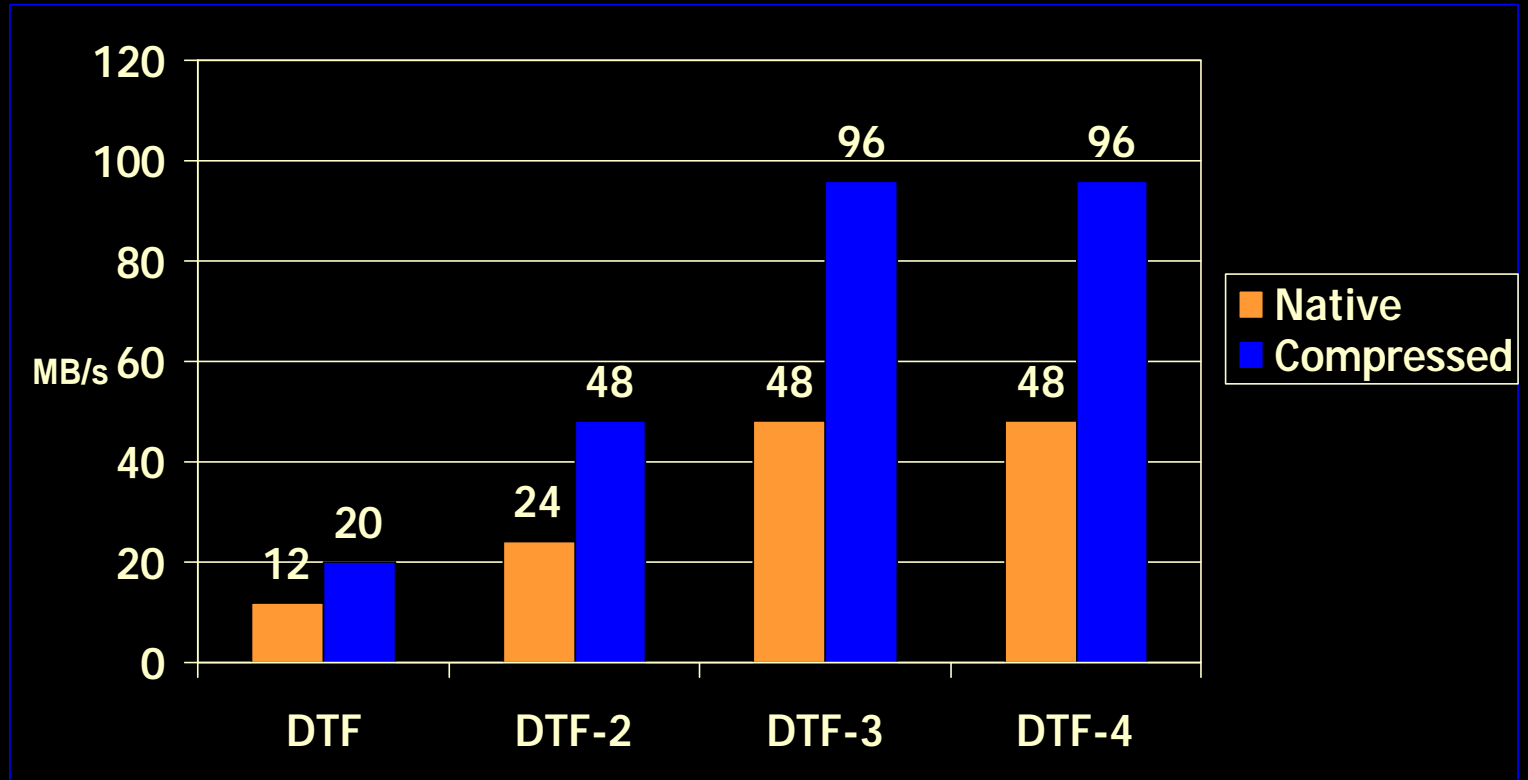
Storage Paradigm Definitions

Access to First Byte	Time to access first byte of the file, includes locate, mount, load, position and read.
Transfer Rate	Delivery speed of the device for read or write.
Multi-media Archive	Archive with mixed storage media capability
Data Verification	Assures data integrity throughout the archive and delivery process.

DTF Roadmap (storage GB/tape)



DTF Roadmap (Speed)



DTF-1 Available NOW

- 12 MB/sec sustained data rate
- 20 MB/sec burst rate
- 42 GB Native Capacity
- 109 GB Capacity w/compression
- Tape Drive Interface - Fast/Wide SCSI

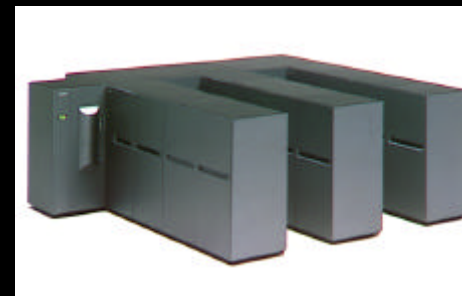
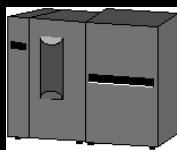


DTF-2 Available early 2000

- 24 MB/sec sustained data rate
- 40 MB/sec burst rate
- 200 GB Native Capacity
- 518 GB Capacity w/compression
- Tape Drive Interfaces
 - UltraSCSI (first development)
 - Fibre Channel (second development)
- Same tape media as DTF-1



Scalability



5.4 TB

- base console
- drive console
- 2 drives

~\$14/GB

→ 12.9 TB

- base console
- drive console
- tape console
- 2 drives

~\$10/GB

→ 88.0 TB

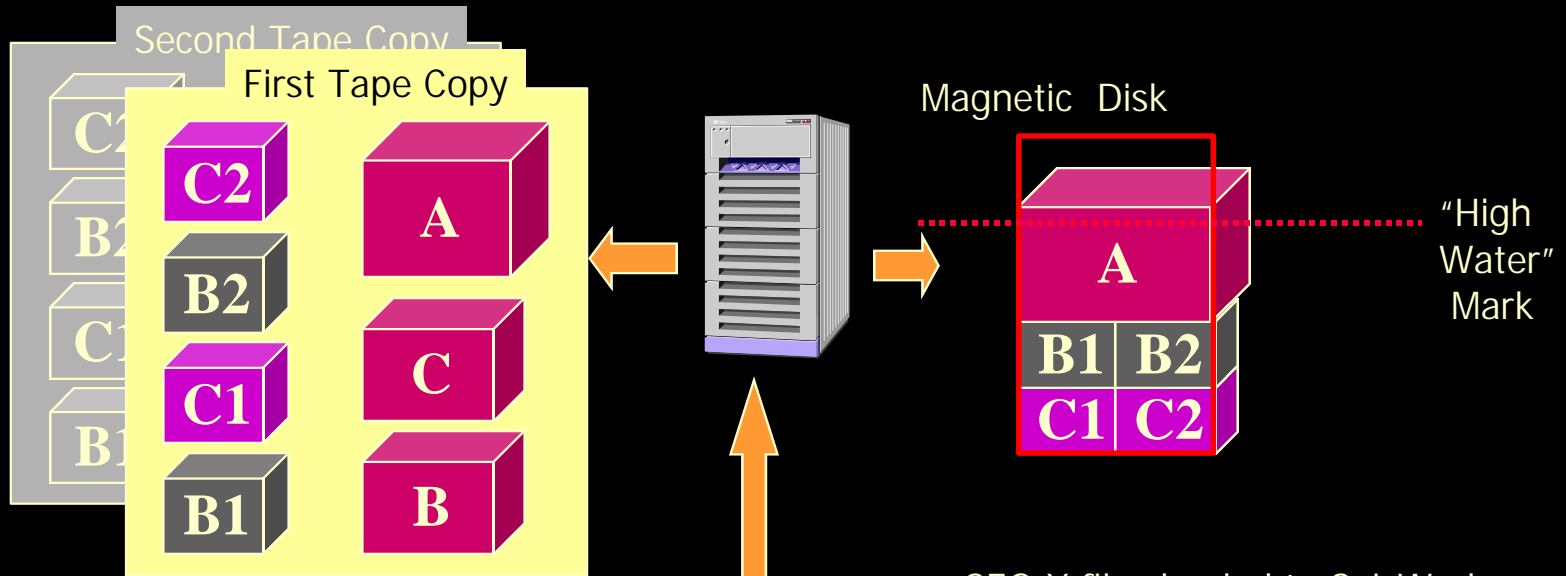
- upgrade to DTF2 drives
- cost ~\$100K

\$2/GB

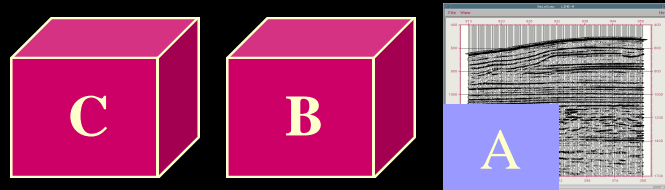
→ 2300 TB

Unlimited Virtual Disk

PetaSite Tape Library



New seismic data is loaded to disk via SDL



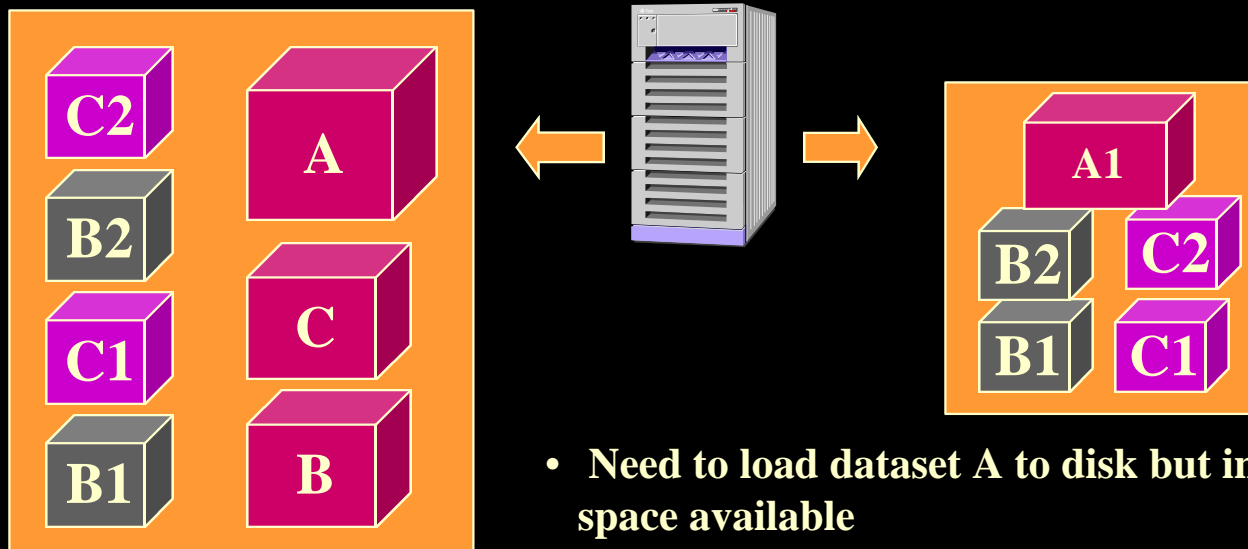
New seismic data arrives

- SEG-Y files loaded to SeisWorks or Charisma
- Copy of SeisWorks trace file made automatically to tape robot
- new version of project file built using SeisWorks tools
- Copy of new versions made automatically to tape robot

Virtual Disk Integration

PetaSTAR Tape Robot

Existing Disk



- Need to load dataset A to disk but insufficient space available
- SAM-FS “sweeps” sufficient files from disk to allow new data to be loaded
- new dataset is loaded to disk