StorageTek 5800 system (“Honeycomb”)  
Breakthrough Solution for Digital Repositories

Scott Parker  
Solutions Architect  
Storage Practice

Agenda

> Overview of ST5800 Design Goals & Objectives, Market Fit  
> A High-Level Look at the Architecture  
> Some Customer Examples  
> Deep Dive  
> Questions & Feedback Time
Storage Market Segmentation

Dynamic Data

Fixed Data
“Fixed Content market”

Structured

Unstructured

Media production, eCAD, mCAD, Office docs

Media-archive, DAM, Broadcast, medical imaging, Media-Internet

Transactional systems, ERP, CRM

BI, data warehousing, scientific, transaction archive

Object Storage

Storage Market Segmentation

Dynamic Data

Fixed Data
“Fixed Content market”

Structured

Unstructured

Media production, eCAD, mCAD, Office docs

Media-archive, DAM, Broadcast, medical imaging, Media-Internet

Transactional systems, ERP, CRM

BI, data warehousing, scientific, transaction archive

Object Storage
ST5800 is For Large Repositories & Archives - Fixed Content

New digital content is born fixed
Through the archive process even transactional data becomes fixed
STK 5800 is not designed for
> Transactional, throughput computing, ERP or other “live” database
> Strict compliance opportunities (today)
> Small-scale file sharing

Targeted at digital library/e-Science repositories

Fixed Content Archival Storage
Customer Requirements

“I need to capture, manage, and preserve digital assets over many years”

Assured data integrity to prevent loss or damage to archived digital assets

Organize, locate & retrieve hundreds of millions of “objects”
> User-definable metadata & search parameters

Minimize TCO
> Non-disruptive migration across generations of technology
> Simple, non-disruptive scaling as capacity grows
> Administrative efficiency: Large “TB managed” per administrator
> Low sustaining cost for service and support
> Low energy consumption
Sun Labs' Design Goal: Build a 3rd Generation Object Storage System

- Custom Data Services
- Metadata Mgmt and Query
- Object Storage

Changing the Game:
Converging Compute and Store

Extending the value

The Foundation

Sun StorageTek 5800 for Digital Repositories  The Only Open Source Petabyte- Scale Object Store

Organize, locate, retrieve & maintain integrity of hundreds of millions of digital library/e-Research data “objects”

- OpenSolaris project drives innovation & collaboration
  - Source code availability protects against lock-in/obsolescence
  - Storage Beans: community extends, customizes & shares
  - Run applications on the storage (closer to data)
  - User-definable metadata & search parameters
  - Integrated with open repository S/W platforms (Flexible Extensible Digital Object Repository Architecture, Dspace, Eprints)

Better data persistence: RAID6, Self healing, continuing data validation. MTTDL: Exceeds 1 million years

Performance scales with capacity to handle increased query and retrieval workloads

Radically simplified storage administration reduces cost of ownership
ST5800 Horizontal Scaling for Large Repositories

Full or half cell configurations
- Half cell 16TB raw
- Full cell 32TB raw
Add more Cells for horizontal scaling
Add or remove hardware online
- Data evacuation
- Data sloshing
Still only one IP address
RAID, healing and data integrity processing scale with capacity

ST5800 Extensible Metadata & Query

System decides unique handles (OID)
- Physical location transparency
- OID enables WORM
- Flat namespace – no directories
Data is tagged with attributes/metadata
The Application developer determines the metadata – and can search on it or define virtual file system views into the data
Metadata index is fault tolerant

RAIN Architecture (16 nodes per cell)
Digital Repository Solutions

<table>
<thead>
<tr>
<th>Solution</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>ST5800 with Fedora</td>
<td>Fedora is a key repository framework. This is being optimized by several institutions including Oxford U.</td>
</tr>
<tr>
<td>ST5800 with Dspace</td>
<td>The Dspace application was initially developed at MIT but is now a community developed application for repositories. Sun has several beta sites working on this application, such as Johns Hopkins U.</td>
</tr>
<tr>
<td>ST5800 with Eprints</td>
<td>EPrints is an electronic publication repository developed at U. Southampton. Southampton is now in early access with the ST5800. There are over 220 Eprints repository users worldwide.</td>
</tr>
<tr>
<td>VTLS VITAL</td>
<td>VTLS's VITAL digital library product interfaces with the Fedora open source repository software and is hosted on Solaris 10 and Linux.</td>
</tr>
</tbody>
</table>

Digital Repository 'Early Access' Program is Well Underway!

Stanford, Oxford, Southampton, Purdue, The Alberta Libraries, DRKZ/Max Planck, Johns Hopkins

These partners have
- An active digital repository effort
- A tight linkage to the community (thought leaders)
- An interest in StorageTek 5800 integration & use
- A willingness to contribute development resources

Sun provides
- Lots of StorageTek 5800 expertise & focused support
Stanford Digital Repository

Goal: 1TB/day ingest (1 x 5800/month)

ST5800 Also Fits Within eScience, HPC

Trends

- Traditional HPC customers need storage archives as well as raw compute power.
- “Cybrarians” have the mandate to preserve research data, applications, format and old equipment

Common requirements span high-end research and digital repositories. Examples: Early access customers at Max Planck Institute, Purdue.

Store and manage massive data sets and images

- Academic/scientific research
- Seismic (oil & gas), Mapping, Imaging
- Federal (satellite images, homeland security, high energy physics)

ST5800 can also be a back-end tier 2 target for Sun's HSM product (SAM)
Staging, Storing & Maintaining HPC Data

- Massive repository of on-line and near-line storage to support HPC retention needs
- High speed, scalable data movement to and from compute environment
- Archives data using standard formats (TAR), which allows technology refreshes and avoids vendor lock-in
- Policy driven, automatic migration of data to archives, user driven data recall
- Leverages Sun StorageTek SL8500/SL500 Libraries, ST6540/ST6140 Modular Arrays, SAM-QFS and SunFire Data Movers
- Sun 5800 – Fixed Content Archive - provides rich metadata attributes to organize, locate, and retrieve hundreds of millions of “objects”

Sun StorageTek 5800 Storage System

Technology Deep Dive
ST 5800 Availability

RAIN – Redundant Array of Independent/Inexpensive Nodes
Symmetrical Cluster
  > HA Database (Clustra)
Dual Level 2 switches, clustered
Service processor not in critical path
Call home with Service Connect, Service Delivery Platform or Control Tower Appliance (futures)

ST5800 High Availability Clustered Design

8 nodes shown (full cell = 16 nodes)
ST5800 Distributed Data placement & Self Healing

Mean Time To Data Loss > 1 Million Years

File 1
File 2
File 3

▲ = parity fragment  □ = data fragment

ST 5800 Data Integrity and Protection

• Unique Object IDs
  > Cryptographic Checksum/Hash (SHA-1)

• Data Placement Algorithm
  > 10,000 combinations

• Reed Solomon erasure coding
  > "error-correcting code that works by oversampling a polynomial constructed from the data"
  > Say again in English, please
  > 5+2 (RAID6) protection

• Ongoing healing services
  > Bit rot can run but can't hide

• Result is resilience to multiple hardware failures

• Backups (DR) with BakBone
ST5800 Serviceability

Deferred service model
- 8-node: lose 1 server
- 16 nodes: lose 2 servers

Service processor not in critical path

Add or remove hardware online
- Data evacuation
- Data sloshing

ST5800 Storage Services

Sun Systems Pack Will Give You MORE than Warranty
- Price break on hardware with purchase of support at sale
- Combines product/services for maximum customer value
- More tech coverage and support features than warranty
- 3 year, 24-hour tech assistance by phone available for greater support
- Extended coverage hours for on-site support
- Interoperability problem management and access to online learning resources

Available Services Include:
- Sun Spectrum Support: Less downtime and higher availability
- Installation: Confidence that it's installed right the first time
- Implementation: Speed the return on your storage investment
- Training: Acquire skills for your staff, limit interruptions
ST5800 in a Nutshell

- Symmetric, clustered design – compute, memory, performance grow with capacity
- Load-balanced horizontal scaling
- Dramatically reduced TCO
- Extreme reliability through self-healing
- For large-scale repositories (16TB – multi-PB)
- World’s first programmable storage solution based on Solaris & Java

ST5800 ("Honeycomb")
Breakthrough Solution for Digital Repositories

http://opensolaris.org/os/project/honeycomb/
http://www.sun.com/storagetek/disk_systems/enterprise/5800/