Solution Direction for Long-Term Archive

Donna Harland
Oracle Optimized Solutions: Solutions Architect
Program Agenda

- Archive Layers
- SAM QFS connectivity for complete solution
- Oracle solutions and 3rd party applications
- Features in SAM QFS to enable long term archive
- Data distribution option
- What Next?
Data Archive Layers

1. Create Information

2. Manage
   Manage content

3. Preserve Storage
   Preservation

4. Store Storage Devices

Data Preservation and Content Management Applications

- Flash
- Disk
- Tape
What Do the Parts Need to Do?

- Search and Access Dynamically
- Meet preservation requirements including
  - Audit
  - Data Integrity
- Meet ingest requirements
  - Performance
  - Metadata Creation
- Seamlessly migrate to the next generation technology, both HW and SW
- Data is ALWAYS available
Program Agenda

• Archive Layers
• SAM QFS connectivity for complete solution
• Oracle solutions and 3rd party applications
• Features in SAM QFS to enable long term archive
• Data distribution option
• What Next?
Basic Architecture of an Unstructured Data Archive Solution

- **Application**
  - Captures Data
  - Creates Content Metadata; Optionally stored in DB
  - Stores Content in a File Store
  - Provides Search Engine
  - Provides data preservation features

- **Database Server**
  - Content Metadata
  - Security
  - Improved search performance
SAM QFS As The File Store

• SAM-QFS
  - Dynamically maintains data on defined tiers of storage
  - Dynamically stages data for access when requested by application
  - Standard file access via FC, NFS, CIFS

![Diagram of SAM QFS as the file store with applications, database server, and file store connected to SAM-QFS managed tiered storage.](image-url)
How the Application Interfaces with SAM QFS

SAM QFS and Application on the same server

• Requirements
  – Application runs on Solaris
  – Application writes to a POSIX compliant file system
  – Up to 4 archive copies

• Good Points
  – Good Performance
  – Single server purchase
How the Application Interfaces with SAM QFS
SAM QFS and Application on the Network

• Requirements
  – Application runs on any operating system that has NFS and/or CIFS
  – Application writes to a Network File System
  – Up to 4 archive copies

• Good Points
  – 10GigE gets needed performance
  – Use of Link Aggregation on GigE gets improved performance
  – More Apps available
  – Easy access to storage
How the Application Interfaces with SAM QFS

QFS Client

• Requirements
  – Applications writes to POSIX compliant file system
  – Up to 4 archive copies
  – QFS Client runs on Solaris or Linux
  – Up to 65,280 clients

• Good Points
  – Scales application horizontally
  – I/O performance
Program Agenda

- Archive Layers
- SAM QFS connectivity for complete solution
- Oracle solutions and 3rd party applications
- Features in SAM QFS to enable long term archive
- Data distribution option
- What Next?
Oracle Optimized Solutions
Lifecycle Content Management

• Complete technology stack. Integration includes:
  – Engineering
  – Tuning
  – Collaboration
  – Software is optimized to take advantage of hardware features
  – Hardware is engineered to take advantage of software features

• Tuned for performance, availability, security, and manageability

• Full-stack testing
  – Use real-world workloads
  – Validated to work
  – Quickly deploy their enterprise infrastructure and achieve a faster return on their IT investments.

Oracle’s Optimized Solution for Lifecycle Content Management

- Web and content management
- Image management; forms recognition
- Integration with ERP apps
- Manage content in all forms, even paper
- Secure search, share and access data
What Solutions Include SAM QFS?

• Third Party Applications and SAM QFS
  – Scalable On-line Archive Repository (S.O.A.R.) from Moca/Arrow and their Channel Partners (see Mark Legott preso)
    – Sun tested and partner marketed
    – Uses Open Source Software Drupal and Fedora
    – Fully supported by Oracle partner for implementation and 1st call
  – Ex Libris
    – New Zealand National Library implementation and validated solution
  – Storage Resource Broker (SRB)
    – Customer implementation at DOD
    – Tight integration with SAM
  – PACS Applications
    – Been in production in many sites since STK was STK
  – Home-Grown-Application
    – Norwegian National Library “it just works”
    – 6PB under SAM management (1 on disk archive 2 on tape archive)
Partner Integration and Solution
Scalable Online Archive and Repository (SOAR)
Scalable Online Archive and Repository (SOAR)

- http://islandora.ca/
- http://www.youtube.com/watch?v=6IzUH44ytGw&feature=youtube_gdata_player
- Scalable Online Archive and Repository
  - Open source digital asset management framework
  - Based on Drupal and DuraSpace Fedora Repository
  - Delivered fully integrated and installed and supported through Arrow ECS and Discovery Gardens (DGI)
  - Industry focus is Higher Education and Teaching Hospitals, BioScience and Research, State and Local Government
  - Customization and support through Discovery Gardens, Inc (DGI)
  - Solution packs meet industry requirements
Ex Libris Rosetta

- http://www.exlibrisgroup.com/category/
- Designed in collaboration with the National Library of New Zealand
- Enterprise-class digital preservation system
- An Archive Strategy, not just an Archive solution
Nirvana SRB (Storage Resource Broker)

Files | Apps | Images | DB

Logical Data Collectors

Nirvana SRB

Metadata Catalog

SAM QFS

SAM QFS Disk Cache

Archive Devices
Storage Resource Broker (SRB)

- http://www.nirvanastorage.com/
- Builds an enterprise-wide view of high-value unstructured data with a one-time implementation of Nirvana SRB
  - Authorized users can collaborate around virtual collections, no matter where it resides, through familiar interfaces
  - A true Global Namespace for strategic collaboration
  - Captures, leverages metadata in a central Metadata Catalog (MCAT)
- Federates multiple static data stores into a single view
  - A unified view of all data, all storage, at all locations
  - ILM, cross-location tiering, technology migrations, data synchronization
  - Optimize distributed storage resources
  - Implement unstructured data management across locations
- **SAM QFS Connection**
  - Tight integration with SAM QFS to properly handle latency of tape access
Picture Archiving and Communication Systems (PACS) and SAM QFS
PACS and SAM QFS

• Improve the reliability and performance of PACS systems used for local storage of diagnostic images and other DICOM content.
• Increase the security, integrity and reliability of medical image archives retained for legal and compliance reasons.
• Build reliable and efficient repositories of images used in clinical trials, research and education.
Program Agenda

- Archive Layers
- SAM QFS connectivity for complete solution
- Oracle solutions and 3rd party applications
- Features in SAM QFS to enable long term archive
- Data distribution option
- What Next?
The Painful Truth about Archives

Economic conditions will change

People will come and go

Hardware
3-10 Years

Hardware and media will become obsolete

Software
5-20 Years

Software will be upgraded

Data will live forever

Years
Data Migrated to New Storage Technology

People Retire
Technology Changes
Uninterrupted Access

- Recycle: migrate to new media
  - Dynamic
  - Non-Disruptive
  - Never off-line
- Information always available

SAM QFS Server
Application

Tier 1 Storage

Tier 2 Storage

Tier 3 Storage
On-Line Grow and Shrink
Introduce New Disk Technology

- Existing Filesystem
- GROW the Filesystem
- SHRINK the Filesystem

- Two Options
  - SAM: Release all; stage on access
  - SAM: Copy data from old LUN to new LUNs
Content Format

• Oracle Content Manager Writes multiple formats based on policy
  – Original format
  – Transform to open formats
    • PDF
    • JPEG
    • JPEG 2000
    • XML
    • Others
Program Agenda

- Archive Layers
- SAM QFS connectivity for complete solution
- Oracle solutions and 3rd party applications
- Features in SAM QFS to enable long term archive
- Data distribution option
- What Next?
• Run SAM archiver on data to be moved to central location
• Run `samfsdump`; send to central site
• Remove media from library
• Load into truck
• Drive truck to central site
• Load media into library
• Run `samfsrestore`
What Next?
What Next for SAM QFS Archive Solutions

• Expand into other industries
  – Flexible application frontends to accept industry data
    • Healthcare
    • Geophysical
    • Health Sciences
    • Research
    • Other big data applications
  – Metadata creation that meets industry requirements

• Application Integration
  – Identify and document customers using SAM QFS and third party applications
  – Partners to sell solutions using third party applications
Questions..
We encourage you to use the newly minted corporate tagline “Hardware and Software, Engineered to Work Together.” at the end of all your presentations. This message should replace any reference to our previous corporate tagline “Hardware. Software. Complete.”
Presentation Instructions

The rest of the slides in this presentation contain additional graphic elements, templates, and instructions for creating your own presentations. It’s recommended that you save all the following slides—including this slide—as a new PowerPoint file and name it “Oracle PPT Resources,” and then use it as a graphic resource and instructional deck for creating future presentations.