Migrating a million documents from Stellent to FileNet: A Customer Perspective

Session Number 2224

John Herzer,
Sandia National Laboratories

IBM Software

Information On Demand **2011**
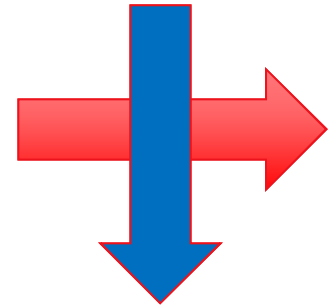
About Sandia

- A government-owned/contractor operated facility
    - Managed by Lockheed Martin for the U.S. Department of Energy's National Nuclear Security Administration
    - Vision: Helping our nation secure a peaceful and free world through technology
- 5 Key Mission Areas
    - Nuclear Weapons
    - Energy, Climate & Infrastructure Security
    - Nonproliferation
    - Defense Systems & Assessments
    - Homeland Security & Defense
- Two primary sites: Albuquerque NM and Livermore CA
- A workforce of 10,795
- More information available at www.sandia.gov
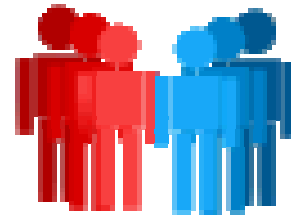
Our information management environment

- A multi-program research and development facility

- Both project and process focused organizations

- Multiple content management systems

  - Stellent

  - Documentum

- Decision to replace these systems with FileNet

- Horizontal vs Vertical deployment

  - Vertical deployments can focus on a few key processes

  - Horizontal deployments bring more diverse support issues

  - Less customization may be achievable with horizontal deployments
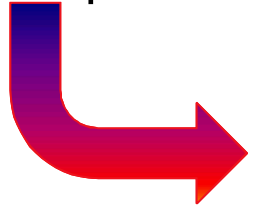
Project Organization

- Project team matrixed across several organizations

  - Content Team

    - Defines use cases, gathers customer requirements

    - Helps customers define information mgmt. practices

    - Handles customer calls

  - Technical team

    - Installs software and applies updates/patches

    - Provides customizations to FileNet where needed

    - Resolves technical problems with IBM support

  - Middleware team
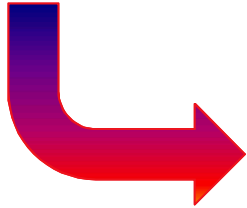
  - DBAs

  - Systems support
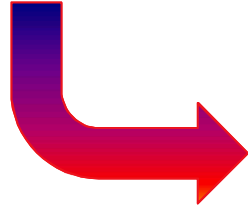
Creating a layered support structure

Corporate Computing Help desk

Extended Enterprise

Content Team

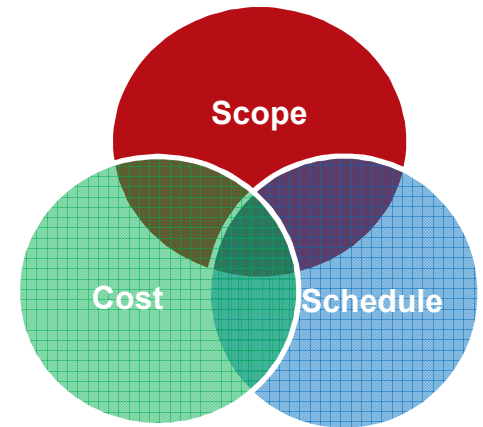Technical Team

Defining the objectives of the migration

- Possible objectives
  - Minimize the duration of the migration
  - Minimize the cost of the migration
  - Re-architect the information environment
  - Provide a smooth transition

- Evaluate the tradeoffs in your decision
  - These objectives may be mutually exclusive
  - Your top objectives may change over time

Scope

Cost

Schedule

Adapting the migration process

- Initial Goals
  - Identify owners of each document collection
  - Work with them to re-organize their content
    - Create document classes with appropriate metadata
    - Define a folder structure for their content
    - Create entry, search and publishing templates
- Adjusted Goals
  - Complete a mass copy of content to FileNet
  - Use one document class with superset of metadata
  - Mirror folder structure of the old system in FileNet
  - Hide content in FileNet until migration complete
- Work with customers after the migration to re-organize their content

The Customization Decision

- Whether to customize is a cost/benefit decision

- Possible reasons for customizing FileNet

  - Supporting your companies unique processes

  - Adding new functionality

  - Fixing a limitation or problem with the product

- Costs associated with customization

  - Development cost of the customization

  - Cost of integrating changes with vendor updates

  - Increased complexity of the environment

FileNet Customizations

- Created default document class setter

  - for Add document and entry template wizards

  - set 'Sandia Document' as the default

- Expanded publication title to 1000 char

- Created dynamic author lookup control

  - Implemented as an AJAX control

  - Accesses the HR database to retrieve names

- Developed custom property value picker

  - Combines multiple dialog screens

  - Allows selection of multiple values

- Defined our own web services layer for CE
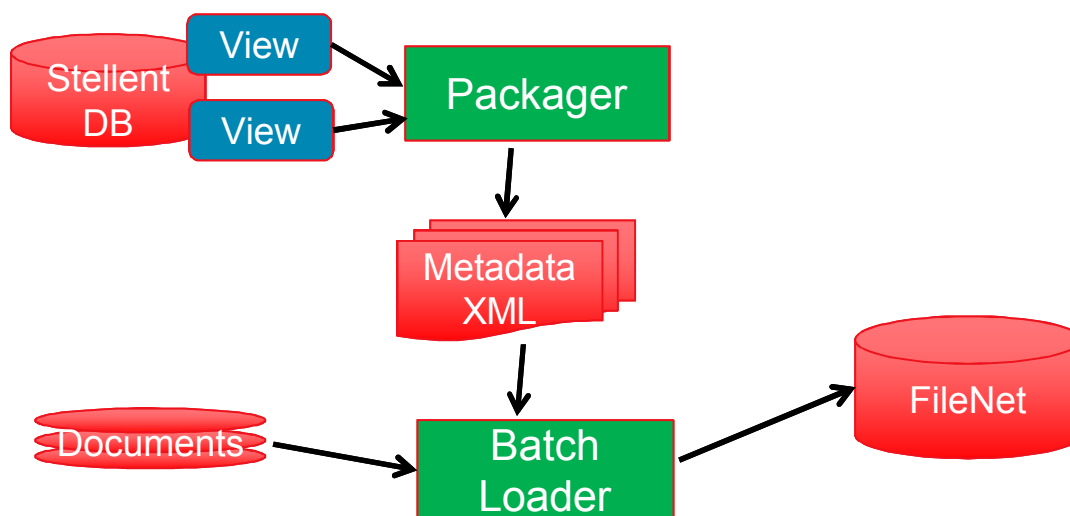
Custom Web Services layer

- addDocument

- checkOutDocument

- checkInDocument

- fetchDocumentMetadata

- fetchFolderMetadata

- findDocuments

- updateDocumentMetadata

- publishDocument

- deleteDocument

- fileDocument

- unfileDocument

- cancelCheckOut

- getServiceVersion
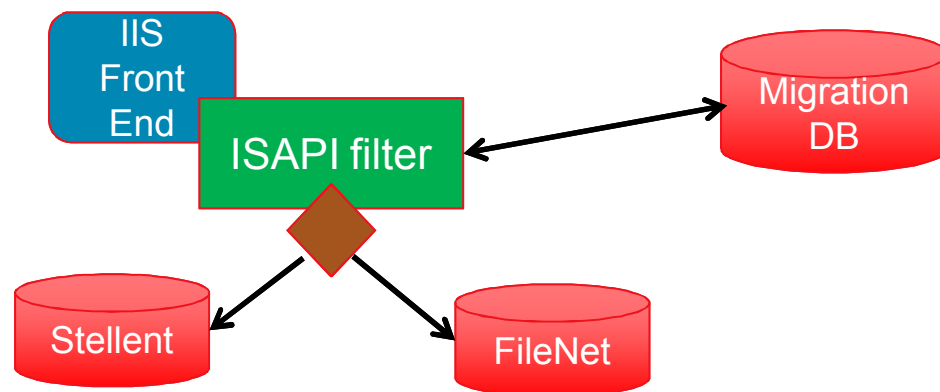
Batch Loader Design

- Pulled metadata directly out of Stellent DB

- Retrieved Stellent documents off of the SAN

- Converted metadata to XML files for transfer

- Used the FileNet Java API to load content

Redirector Design

- Implemented as an ISAPI filter

- Written in C++ for performance

- References a migration database to map URLs

Batch Loader Statistics

- Load rates varied significantly by machine

  – Desktop PC loaded files at .8 MB / sec

  – CE Server loaded files at 3.9 MB / sec

- Ran up to 3 batch loader jobs concurrently

- Migrated a total of 1.2 M files (3.5 TB of data)

- Actual migration took 6 months

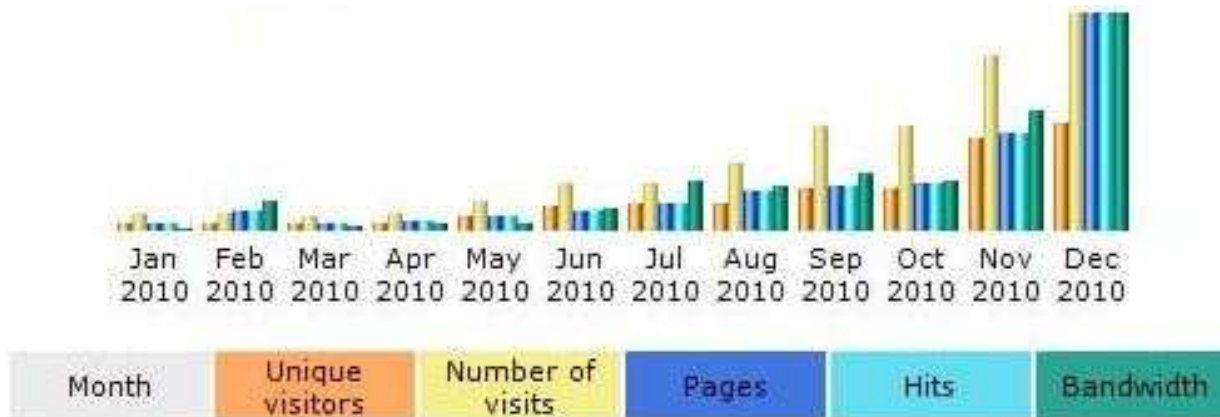- Elapsed time from FileNet install was 2.5 years

Batch Loader Issues

- Kickouts from each batch had to be resolved

    – Non-ASCII characters in metadata

    – UserIds not found on the LDAP server

    – Filenames without extensions

    – Missing metadata

- Several programming bugs popped up as well

    – Default attributes were updated incorrectly

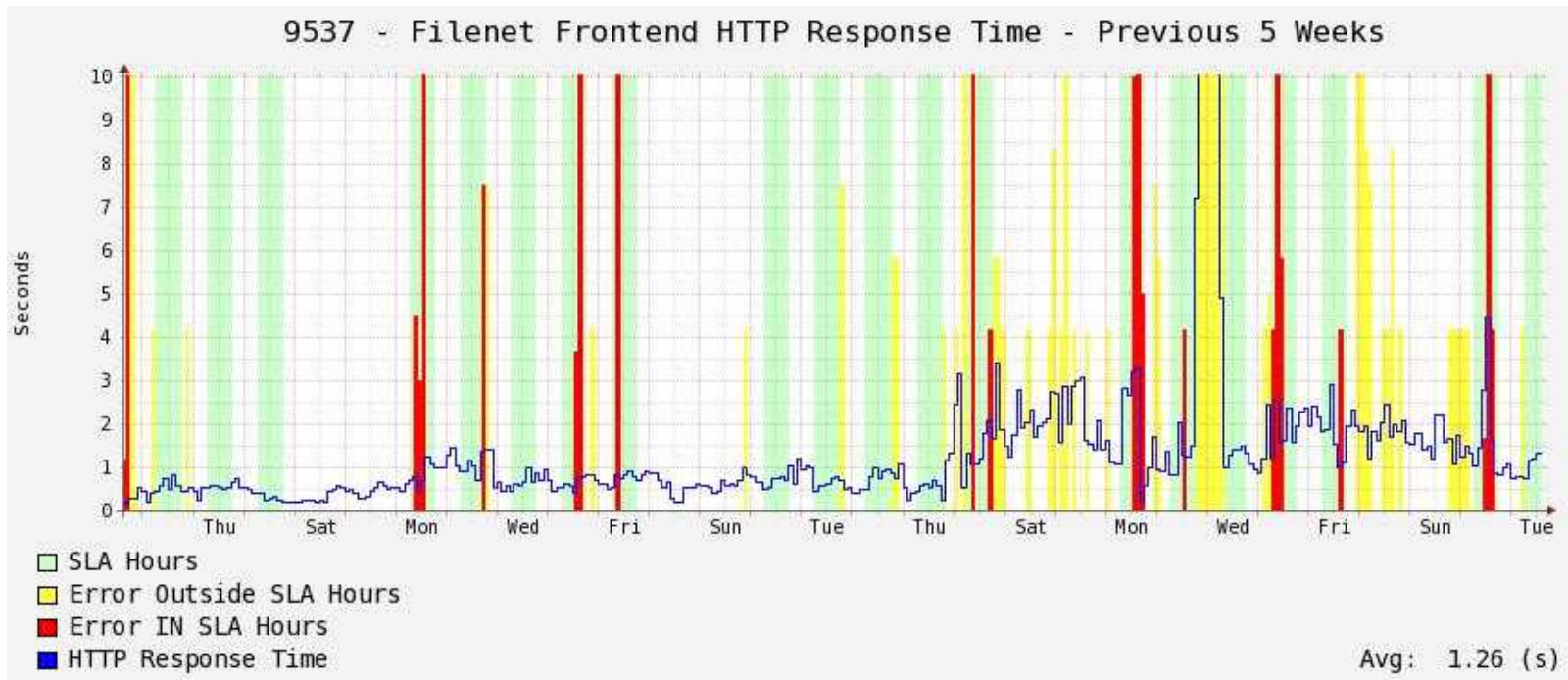   – Revisions were ordered incorrectly

FileNet Usage during 2010

- Metrics from the AWStats log file analyzer

- Web service load testing completed in February

- Migration completed in October
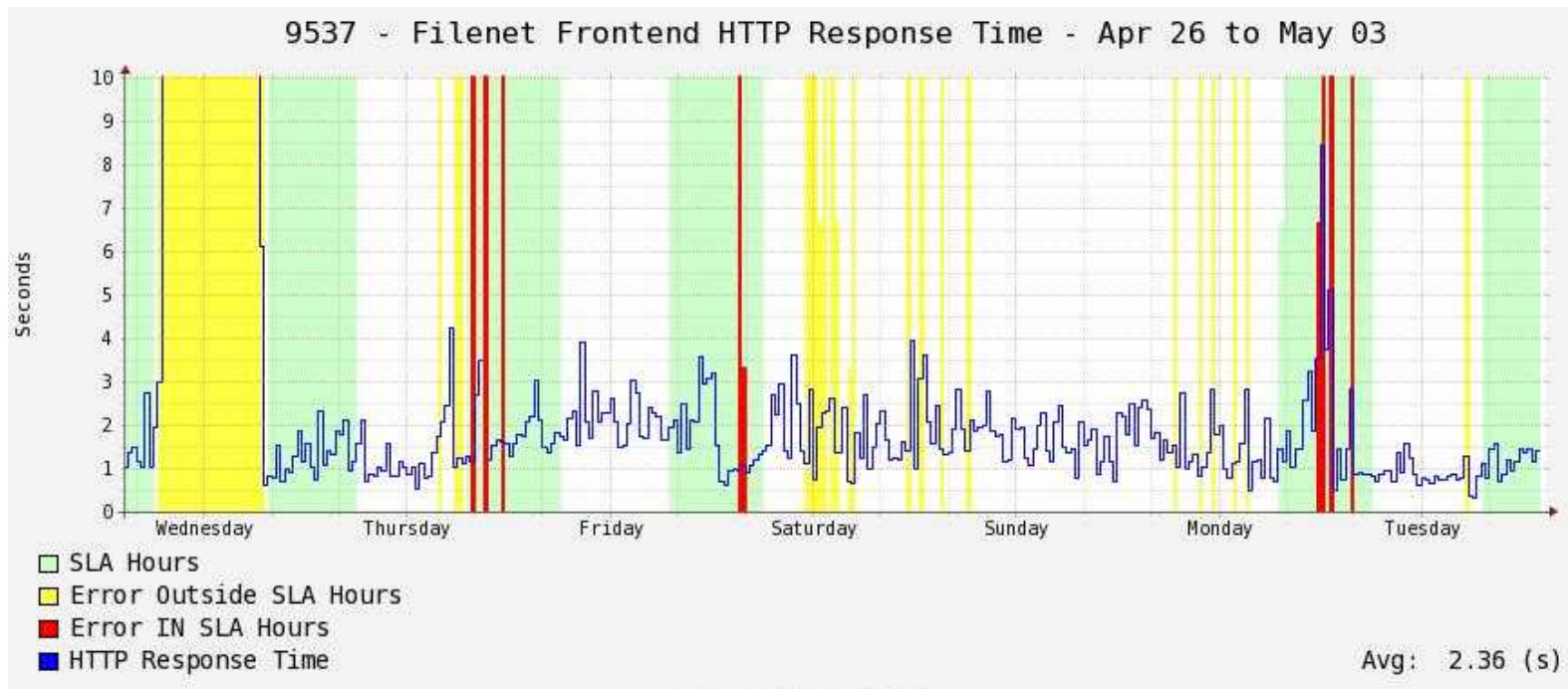
Performance Impact

- Ran into memory constraints on Windows AE machine

- Switched to AE on Linux after thorough load testing

- To our surprise, performance got worse

Performance Improvement

- We found up to 900 active sessions at one time

- Created 4 WebLogic nodes – each with their own Workplace

- Reduced session timeout from 60 minutes to 15 minutes

9537 - Filenet Frontend HTTP Response Time - Apr 26 to May 03

Legend:
- □ SLA Hours
- □ Error Outside SLA Hours
- ■ Error IN SLA Hours
- ■ HTTP Response Time

Avg: 2.36 (s)

Lessons learned from the migration

- Maximize your throughput for content
  - Pull content directly off the file store or database
  - Use the network to move files, not the API
  - Look for choke points
- Bring over all of your metadata
  - Storage is cheap but reruns are costly
  - Consider eDiscovery issues
- Clean up the content as you migrate
  - Encourage customers to re-organize before the migration
  - Have to provide cutoff dates
- Perform 100% validation of the data

Thank You!
Your Feedback is Important to Us

- Access your personal session survey list and complete via SmartSite
  - Your smart phone or web browser at: iodsmartsite.com
  - Any SmartSite kiosk onsite
  - Each completed session survey increases your chance to win an Apple iPod Touch with daily drawing sponsored by Alliance Tech