Shuang Wang
Software Developer
March 2013

Product Implementation Training (PIT)

IBM FileNet P8 Version 5.2
Thumbnail Generation Service

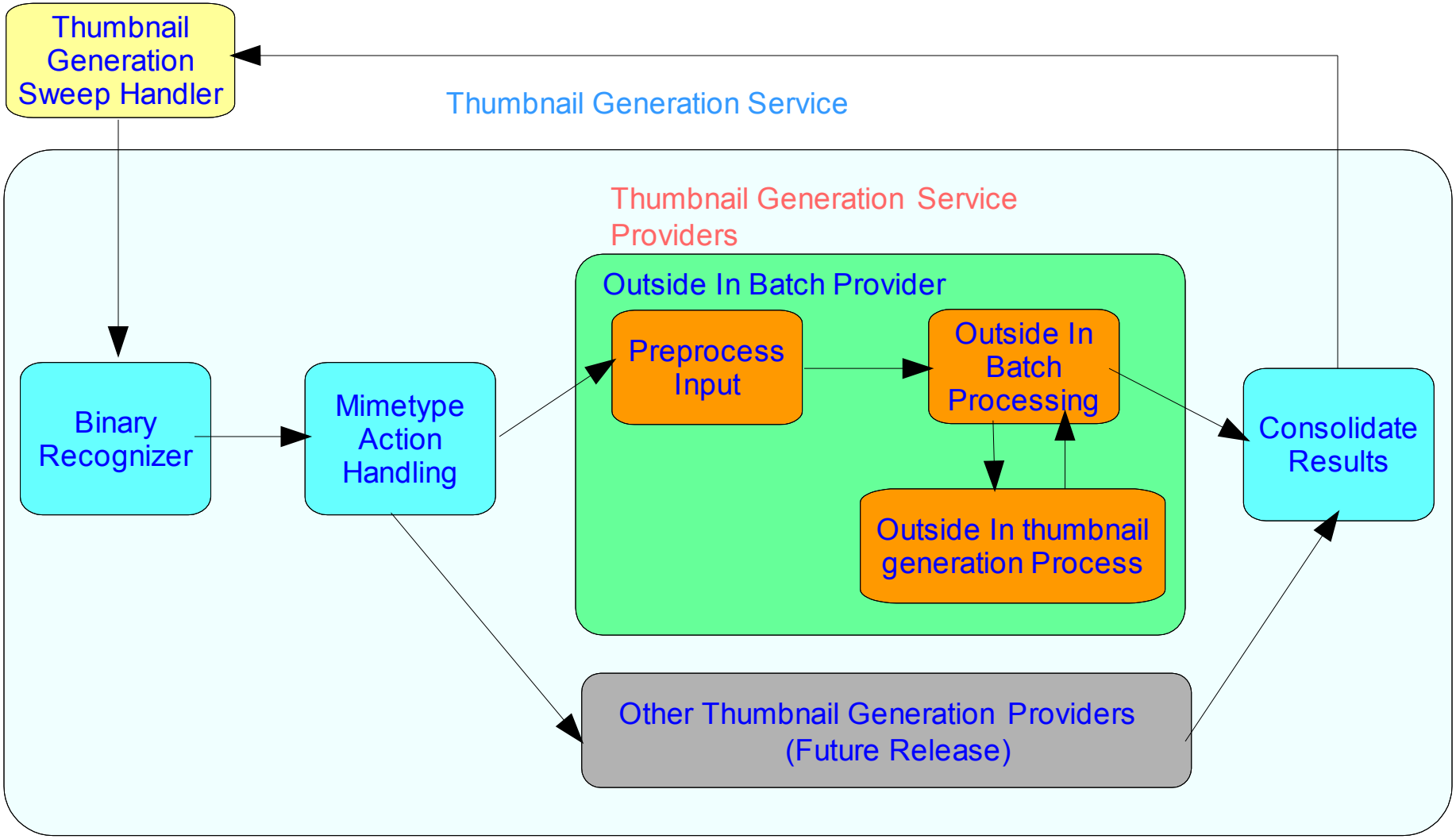
Introduction

- Course Overview
  - Introduces Thumbnail Generation Service Architecture
  - Provides details about Outside In Batch Thumbnail Generation
  - Installation, Configuration, Troubleshooting and Performance Considerations
- Target Audience:
  - Those who will be implementing and supporting Thumbnail Generation Service in FileNet P8
- Prerequisites:
  - Familiarity with IBM FileNet P8
  - Familiarity with Thumbnail Generation Sweep
- Version Release Date: March 2013

Course Roadmap

- ➔ Thumbnail Generation Service Overview
  - Architecture Diagram
  - Service Interface
  - Outside In Process Diagram
- – Thumbnail Generation Service Details
- – Installations and Configurations
- – Advanced Configurations, troubleshooting and Performance Considerations
- – Product Help/Documentation/Resources

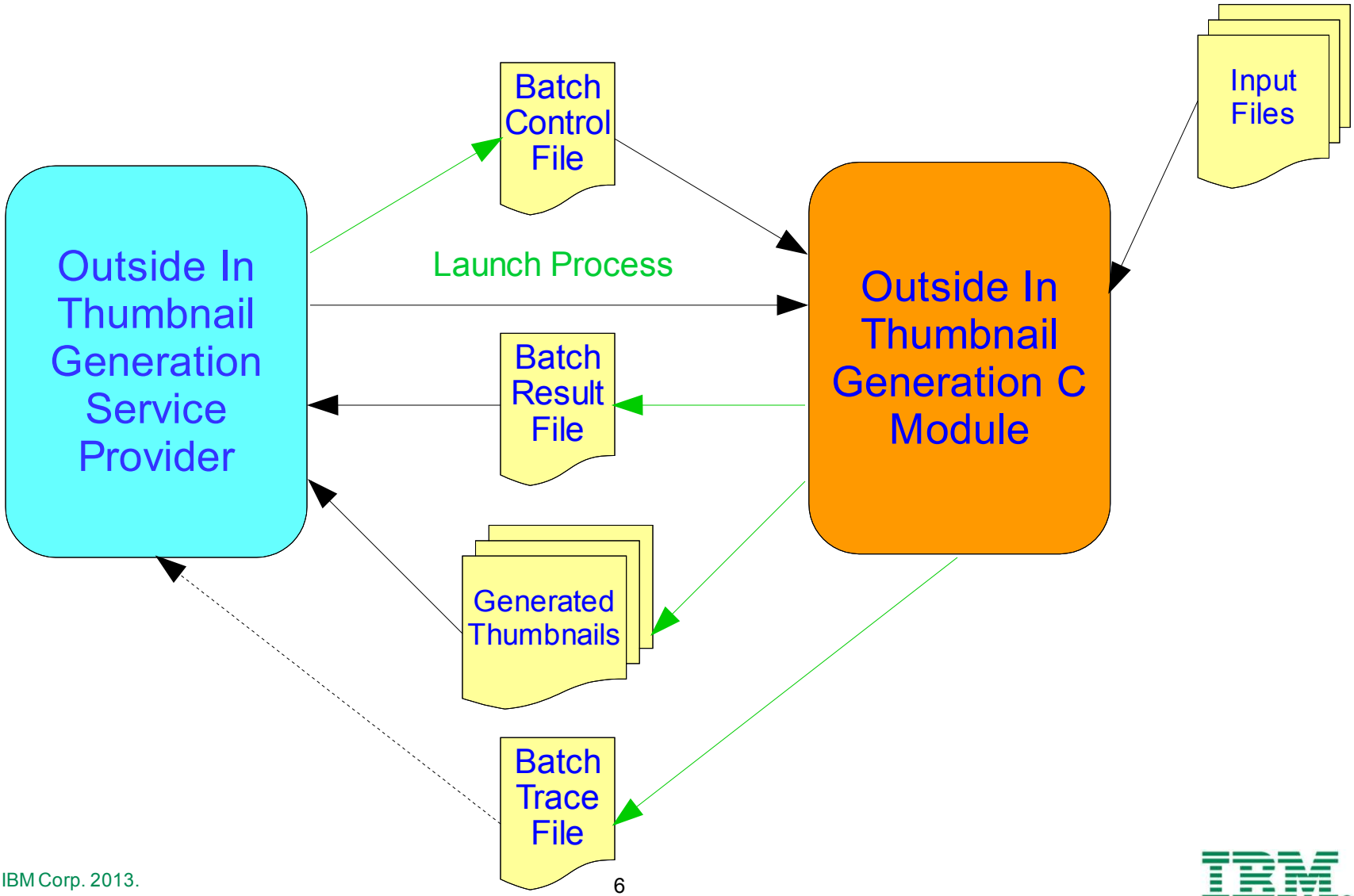
Thumbnail Generation Service Diagram



Thumbnail Generation Service Internal Interface

- ThumbnailGenerationResult[] generate(ThumbnailInputFile[] inputs, ThumbnailGenerationOptions options);
- ThumbnailInputFile specifies
  - FileName, Mimetype, File Extenstion
- ThumbnailGenerationOptions defines
  - Size (Small/Medium/Large)
  - Format (JPEG/GIF/PNG)
  - Color (not used now)
- ThumbnailGenerationResult
  - Can be an byte[] of generated thumbnail
  - or thumbnail generation exception with the information
    - If the exception is retriable
    - If the thumbnail input file has been processed

Outside In Thumbnail Generation Service Process



Course Roadmap

- Thumbnail Generation Service Overview
- ➔ Thumbnail Generation Service Details
  - Binary Recognizer
  - Mimetype Action Handling
  - Outside In Thumbnail Generation Provider
  - iWork File support
- Installations and Configurations
- Advanced Configurations, troubleshooting and Performance Considerations
- Product Help/Documentation/Resources

Binary Recognizer

- Detect the well-known mimetypes by checking the bytes in the file header
- Same algorithm used in other IBM product and CBR text extraction
- Default configuration file binary_recognizer_config.xml is packaged under /res folder of engine.jar file inside CE ear file, shared by thumbnail generation and CBR text extraction
- Configuration can be overridden by placing binary_recognizer_config.xml under the user.dir, user.home, java.home or OutsideIn SDK installation directory
- Binary recognition can be skipped by setting Thumbnail.OI.SkipBinaryRecognition = true in FileNet.properties

Mimetype Action Handling

- Decide how thumbnail generation will be handled based on mimetypes
- Default configuration is defined in thumbnail_mime_actions_config.xml
  - In the same location as binary recognizer configuration file and can be overridden in the same way
- Mimetype detection order
  - Binary Recognizer
  - File extension to mimetype mapping defined in thumbnail_mime_actions_config.xml
  - File extension to mimetype mapping defined in P8 default mimetypes.properties file
  - Declared mimetype passed into Thumbnail Generation Service
- Mimetype action handling can be skipped by setting Thumbnail.OI.SkipMimetypeAction = true in FileNet.properties
  - All input files will be passed to the default thumbnail generation service provider

Mimetype Action Handling (Cont'd)

- Mimetype actions:
  - supported, unsupported, supportedNoGeneration
  - With unsupported, supportedNoGeneration, files will not be sent to thumbnail generation provider.
- Extension based action override
  - Some file formats (like OpenDocument, iWork) are just in a common file format (like zip files) with their own extensions
  - It will be identified as the common mimetype (application/zip) by binary recognizer
  - Use the extensionMap section to override the default action handling for the common mimetype

```
<mimeTypeActionMap>
        <mimeType>application/zip</mimeType>
        <actionName>supportedNoGeneration</actionName>
        <!--  the following OpenDocument file types take the zip file fomat -->
        <extensionMap>
                <extension>odc, otc, odf, otf, odg, otg, odi, oti, odp, otp, ods, ots, odt, odm,ott, oth</extension>
            <extensionActionName>supported</extensionActionName>
        </extensionMap>
</mimeTypeActionMap>
```

Outside In Thumbnail Generation Provider

- Outside In Thumbnail Generation Provider calls the external C module ibmfnthb which integrates with Oracle Outside In Image Export version 8.3.7 patch 15954226
- Work with all formats Outside In Image Export supports
- Outside In Image Export supports converting many file formats to several image formats
  - Outside In 8.3.7 Supported Formats document is available at support.oracle.com (needs Oracle account) as an attachment to note 790088.1
- Default mimetype action handling configuration excludes certain file formats like audio, binary, zip files (supportedNoGeneration)

Apple iWork File Support

- Outside In 8.3.7 does not support iWork file thumbnail generation
- P8 Supports (on a best-efforts basis) Apple iWork '09 files in single file format with an embedded QuickLook thumbnail
- Single file format iWork files is in a zip file format with the iWork file extensions (.pages, .numbers, .key)
- QuickLook thumbnail is extracted and sent to Outside In for formatting and resizing
- For iWork files in the package bundle format with QuickLook thumbnails, can be processed if zipped into one file and set correct file extension
- For .key files only, if no QuickLook thumbnail available, thumbnail can still be generated using embedded thumbnails under /thumbs folder (image quality is low)
  - Need to set Thumbnail.OI.ProcessIWorkNativeThumb=true in FileNet.properties
- iWork file processing can be turned off by setting Thumbnail.OI.ProcessIWorkFiles=false in FileNet.properties file

Course Roadmap

- Thumbnail Generation Service Overview
- Thumbnail Generation Service Details
- ➔ Installations and Configurations
  - Outside In SDK Deployment
  - Prerequisites in Unix
  - Configurations in Content sub system
  - Configurations in FileNet.properties
  - Thumbnail Temp Directory Cleanup
- Advanced Configurations, troubleshooting and Performance Considerations
- Product Help/Documentation/Resources

Outside In SDK Deployment

- Outside In Image Export SDK is bundled and deployed together with Outside In Search Export used for text extraction
- For each supported Operating System, INSO.zip file includes the Outside In redistributables and P8 executable files
- CSSClient.jar includes the INSO.zip file for all Operating Systems
- In CMUI at the required bootstrap step, an appropriate INSO.zip file will be picked up based on the platform and added into CE .ear file.
- After CE ear file is deployed, during the CE server startup, the INSO.zip file will be extracted and installed under the FileNet folder
  - C:\ProgramFiles\IBM\WebSphere\AppServer\profiles\AppSrv01\FileNet\server1\INSO\bin\sx-8-3-7-win-x86-64
  - Under INSO folder, install.properties file shows installation status, version and library path
- A new version of the SDK can be delivered to the customer in CSSClient.jar, and re-bootstrap CE ear file. CE startup task will automatically check the version information and deploy the new SDK

Thumbnail Generation Outside In Prerequisites

- Unix
  - Access to a running X Windows System and the presence of Motif are required to convert from non-raster or vector formats files (such as Office documents, and pdf files)
    - VNC Server, Remote X Server, Windows X Servers, X Virtual Framebuffer (XVFB)
    - The X libraries Xm, Xt and X11 must be present
  - Bitmap graphic conversion does not require access to a running X Windows server (GIF, JPEG, TIFF, PNG, and BMP)

- Windows
  - Outside In Image Export is using the built in Windows Graphics Device Interface (GDI) to render documents to images

Thumbnail Configurations in Content Configuration

- Max Thumbnail Generation Processes (2)
  - The Max number of concurrent external thumbnail generation processes
  - Controlled by sweep framework using a semaphore
- Thumbnail Generation Batch Size (50)
  - The number of files that can be processed in one batch
  - Sweep framework has its own batch size (200)
- Thumbnail Generation Process Time Out (600 seconds)
  - The time out in seconds for processing a single batch
  - When a batch processing times out, thumbnails that have already been generated will be returned, others will be returned as time out exceptions. The files after the last processed file will also be marked as not processed.

Thumbnail Configurations in Content Configuration (Cont'd)

- Thumbnail Generation Temp Dir
  - The location of the temporary directory root for thumbnail generation
    - Example, C:\CE_Thum_temp\DaphneServer\{A77CE131-525D-4E86-BB17-B9D98A00DC3A}
  - If not set, fall back to Content Temporary Directory
  - The location for thumbnail generation input files, batch control, result and trace file, generated thumbnails and other temporary files

- Thumbnail Generation Delay (90 seconds)
  - The minimum number of seconds to delay generation of thumbnails for newly added content.
  - To avoid generating thumbnails before document content is finalized

Thumbnail Size Configuration

- FileNet.properties file location
  - C:\Program Files\IBM\WebSphere\AppServer\profiles\AppSrv01\FileNet.properties
- Sweep request defines thumbnail size as Small, Medium and Large
- By default, generated thumbnail will maintain the original aspect ratio and the actual pixel size is defined as the limit for both the width and height of the generated thumbnail images. (Small 100, Medium 250, Large 500)
  - Original size 2000X1600, the large thumbnail will be 500X400
  - Original size 1600X2000, the large thumbnail will be 400X500
- Default size can be changed by setting Thumbnail.SizeLimit.xxxxx (Small, Medium, Large) in FileNet.properties
- Other supported size configuration parameters for special requirements
  - Thumbnail.WidthLimit.xxxx, Thumbnail.HeightLimit.xxx
    - When any of these are specified, default size limit will not be used
    - If both SizeLimit and WidthLimit (or HeightLimit) are defined, then the minimum limit value will be used.
  - Thumbnail.Width.xxxx, Thumbnail.Height.xxx

X Windows Server Configuration

- For Unix system, define the X Windows sever in FileNet.properties
  - Thumbnail.OI.XWinDisplay=9.39.20.11:2.0
  - Thumbnail.OI.XWinDisplay=:0.0
- When Thumbnail.OI.XWinDisplay is not specified, environment variable $DISPLAY will be used if it exists
- Configure multiple X Windows Server
  - Thumbnail.OI.XWinDisplay=9.39.20.11:2.0, 9.39.20.12:2.0, 9.39.20.13:2.0
  - Thumbnail.OI.XServerSetAsideDuration (default 600 seconds)
    - Defines how long a X Windows server will be set aside (don't use unless no other Xservers available) when thumbnail generation time out is detected using a given X server

Thumbnail Generation Temp Directory Cleanup

- Background task to clean up any abandoned files in thumbnail temp directory
  - In a normal situation, all the temporary files will be deleted after thumbnail generation is completed
  - Will cleanup input files, batch control files, batch result files,batch trace files, generated thumbnail files, and dumped thumbnail files or error files in /verify folder
- Works similar to Inbound folder cleanup task
  - Content Queue dispatcher dispatches the work
- Enabled only when thumbnail extension add on is installed
- Configuration parameters (in Content subsystem configuration)
  - Abandoned Thumbnail Cleanup Interval
    - The frequency (in seconds) of the cleanup task
  - Thumbnail Temp File Life Time
    - The number of seconds temporary files can exist before being considered abandoned
- PCH counters available to show files examined, files accepted, files deleted and cleanup duration

Course Roadmap

- Thumbnail Generation Service Overview
- Thumbnail Generation Service Details
- Installations and Configurations
- ➔ Advanced Configurations, troubleshooting and Performance Considerations
  - Advanced Configurations
  - Logging and PCH Counters
  - Outside In Performance and Output Quality
  - Troubleshooting
- Product Help/Documentation/Resources

Advanced Configuration Parameters

- Defined in FileNet.properties for debugging purposes
- Thumbnail.OI.SDK (use a non-default SDK location)
- Thumbnail.OI.Executable (use a non-default executable)
- Thumbnail.OI.ExtraLaunchParams
  - comma separated extra parameters, for example “jpegquality=100”, that will be used to launch Outside In thumbnail generation process
- Thumbnail.OI.KeepBatchProcessFiles
  - Keep the batch control, result file and trace file after generation is done.
- Thumbnail.OI.KeepBatchProcessFilesForFailure
  - Keep the batch control, result and trace files only if the batch is not completed (like in a timeout case) or the process exits with a non-0 value
- Thumbnail.OI.DeleteFile
  - Set to false to keep generated thumbnails in temp dir after process is done

Advanced Configuration Parameters (Cont'd)

- Thumbnail.OI.DumpThumbnailResult
  - Set to true will dump the thumbnail files or error messages into \verify folder under thumbnail working directory
  - 20120216120801031_00002_Input_File_Name_thumb_restore.png or 20120216120801031_00003_Input_File_Name_thumb_error.txt
- Thumbnail.OI.DumpThumbnailErrorOnly
  - Set to true will dump error messages only into \verify folder under thumbnail working directory
- Thumbnail.OI.ConsumeStream
  - Set to true will start a thread to consume output stream from the generation process
  - When a batch can be processed successfully running command line but hangs in P8 processing, turn this on to verify the issue
- Thumbnail.OI.ShowStreamInLogs
  - Set to true will show Stream in CE logs at detailed trace level

Logging

- Use Thumbnail Generation subsystem for logging
- P8 Logging level also controls the C process logging level
  - Moderate trace
  - Detailed trace
- Thumbnail generation C process logs will be dumped into P8 logs at moderate trace level
- At the moderate trace level, logs will include errors for each input file that thumbnail generation service fails to generate thumbnails

PCH Counters

- Thumbnail Generation Invocations
  - The number of thumbnail generation calls by sweep handler
- Thumbnail Batches Processed
  - The number of batches thumbnail generation service has processed
- Batches Completed Processing
  - The number of batches thumbnail generation service completed processing
- Batches Failed to Complete
  - The number of batches that failed to complete processing (for example, failed to launch process, process timeout)
- Thumbnails Generated
  - The number of thumbnails generated successfully

PCH Counters (Cont'd)

- Thumbnail Generation Failed
  - The number of input files the thumbnail generation provider failed to generate thumbnails
- Thumbnail Inputs Ignored
  - The number of input files that have been filtered out (unsupported, supportedNoGeneration) and have not been processed by thumbnail generation providers
- Thumbnail Process Timeouts
  - The number of batches that have not been completed due to timeout
- Thumbnail Process Timeout Wait Seconds
  - The time spent on batch processing before timeout
- XServer usage: xserver_name
  - How many times a particular X server is used in thumbnail generation

Outside In Thumbnail Generation - Performance

- Unix
  - Using a local X server gives best performance
  - X server can be run using X Virtual Frame Buffer (XVFB)
    - An X server that can run on machines with no display hardware and no physical input devices, performs all graphical operations in memory
    - Using XVFB is covered in the documentation page "Configuring the display for thumbnail generation on X Windows"
  - Remote X servers (such as Exceed, Xmanger) can also be used, but performance may suffer
  - Depending on what X server is used the rendering quality will vary
- Windows
  - Use a fast processor and graphics card and adequate memory

Outside In Thumbnail Generation – Output Quality

- Font considerations with X server
  - Outside In Technology makes use of the fonts loaded on the system that are made available to the X Windows sever
    - Foreign language documents.
      - For example, when converting a Japanese Word file, Outside In makes use of the Japanese fonts loaded on the system performing the conversion. If an exact font match is not found, will try to find an appropriate replacement font.
    - Files using Windows fonts.
      - Since Unix systems will not have Windows fonts installed, Outside In must substitute for these fonts in its output, which can result in font quality differences in the output.

Thumbnail Generation Troubleshooting on Unix

- For Vector Graphic Conversion Problems on Unix
  - Check that the X libraries are located in your system-specific library path
  - Check that Thumbnail.OI.XWinDisplay or $DISPLAY environment variable is pointing to a running X server
  - Check the dependencies of one or more of the core Outside In libraries, especially libos_xwin.* or libsc_ut.*.
    - Tools to check dependencies will differ by platform, but ldd is a great tool to accomplish this. Please refer to the Unix man pages for more information.
    - If any dependencies are not resolved, locate these libraries on your system and make sure that they are included in your system specific library path. You may need to install additional packages that include libraries required by Outside In.

Course Roadmap

- Thumbnail Generation Service Overview
- Thumbnail Generation Service Details
- Installations and Configurations
- Advanced Configurations, troubleshooting and Performance Considerations
- ➔ Product Help/Documentation/Resources

Product Help/Documentation/Resources

- IBM FileNet P8 Platform Documentation Version 5.2
  http://www.ibm.com/support/docview.wss?uid=swg27036917
- IBM FileNet P8 Version 5.2 Information Center
  http://pic.dhe.ibm.com/infocenter/p8docs/v5r2m0/index.jsp
- Oracle support site for Outside In Documentation http://
  www.oracle.com/technetwork/middleware/webcenter/content/oit-all-085236.html
- Contact
  - General Sweep Framework
    - Bob Kreuch, Eric Edeen
  - Thumbnail Generation Sweep and thumbnail persistence
    - Kevin Bates
  - Thumbnail Generation
    - Shuang Wang, Kristoffer Gjevre
  - Development Manager: Grace Smith