Whole Slide Imaging in DICOM

Ole Eichhorn
Chief Technical Officer, Aperio
August 20, 2008

ole@aperio.com

Whole Slide Imaging in DICOM

- Introduction
- Characteristics of Whole Slide Images (WSI)
  - Image size
  - Image access patterns
- Storing WSI in DICOM
  - NEMA Working Group 26
  - Proposed DICOM supplement for WSI
- IT Considerations for WSI
  - Storage
  - Bandwidth
- Solutions for Managing WSI
  - Access Models and Security
  - Compliance Characteristics
  - Performance

Introduction

- Pathology is undergoing a transformation
  - Digital imaging is increasingly important
  - Driven by availability of instruments for digitizing microscope slides, resulting in whole-slide images (WSI)
- WSI are different to other images
  - Size - larger
  - Access patterns – pan, zoom, focus

Image Size

Image dimensions, data size

- Typical: 20mm x 15mm @ .5mpp (“20X”) = 40,000 x 30,000 pixels = 1.2Gp = 3.6GB
- Typical: 20mm x 15mm @ .25mpp (“40X”) = 80,000 x 60,000 pixels = 4.8Gp = 15GB
- Extreme: 50mm x 25mm @ .1mpp (“100X”) = 500,000 x 250,000 pixels = 125Gp = 375GB
  - x 10 Z-planes => 3.75TB

Image Size

File size

- Use lossy compression
  - JPEG – fast, standard, yields 15:1
  - JPEG2000 – better, yields 25:1

	Image Size	File Size JPEG	File Size JPEG2000
20mm x 15mm @ 20X	3.6GB	240MB	140MB
20mm x 15mm @ 40X	15GB	960MB	580MB
50mm x 25mm @ 100X	375GB	25GB	15GB

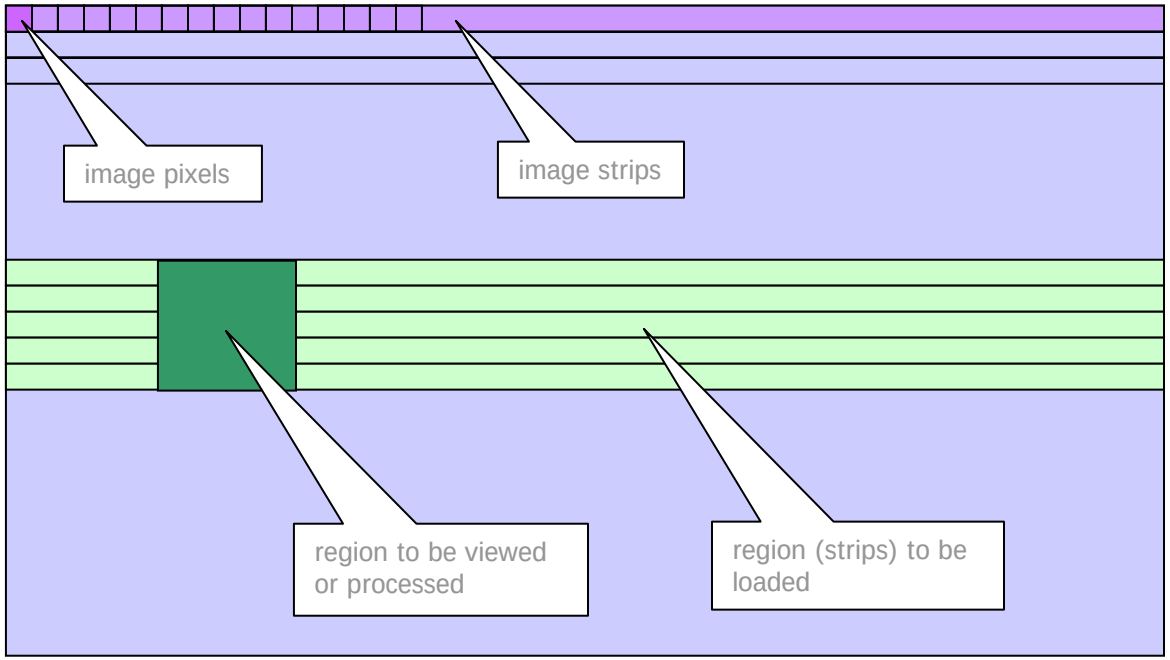
Access Patterns

Access patterns

- Pathologists cannot view entire microscope slide at diagnostic resolution
- Pan around at low res (2.5mpp = “4X”), then zoom in at high res (.25mpp = “40X”)
  - Requires organization to support rapid panning & zooming
- Multiple Z-planes may be captured
  - Requires organization to support rapid focusing

Access Patterns

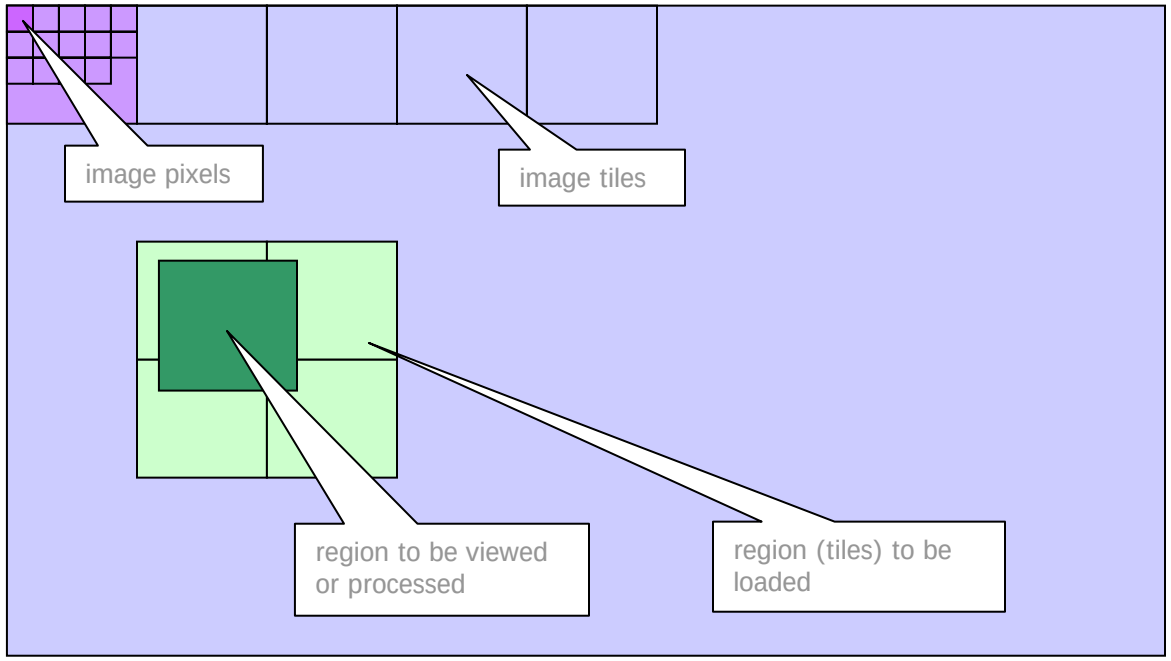
Simple image organization: Stripped



Does *not* support efficient panning

Access Patterns

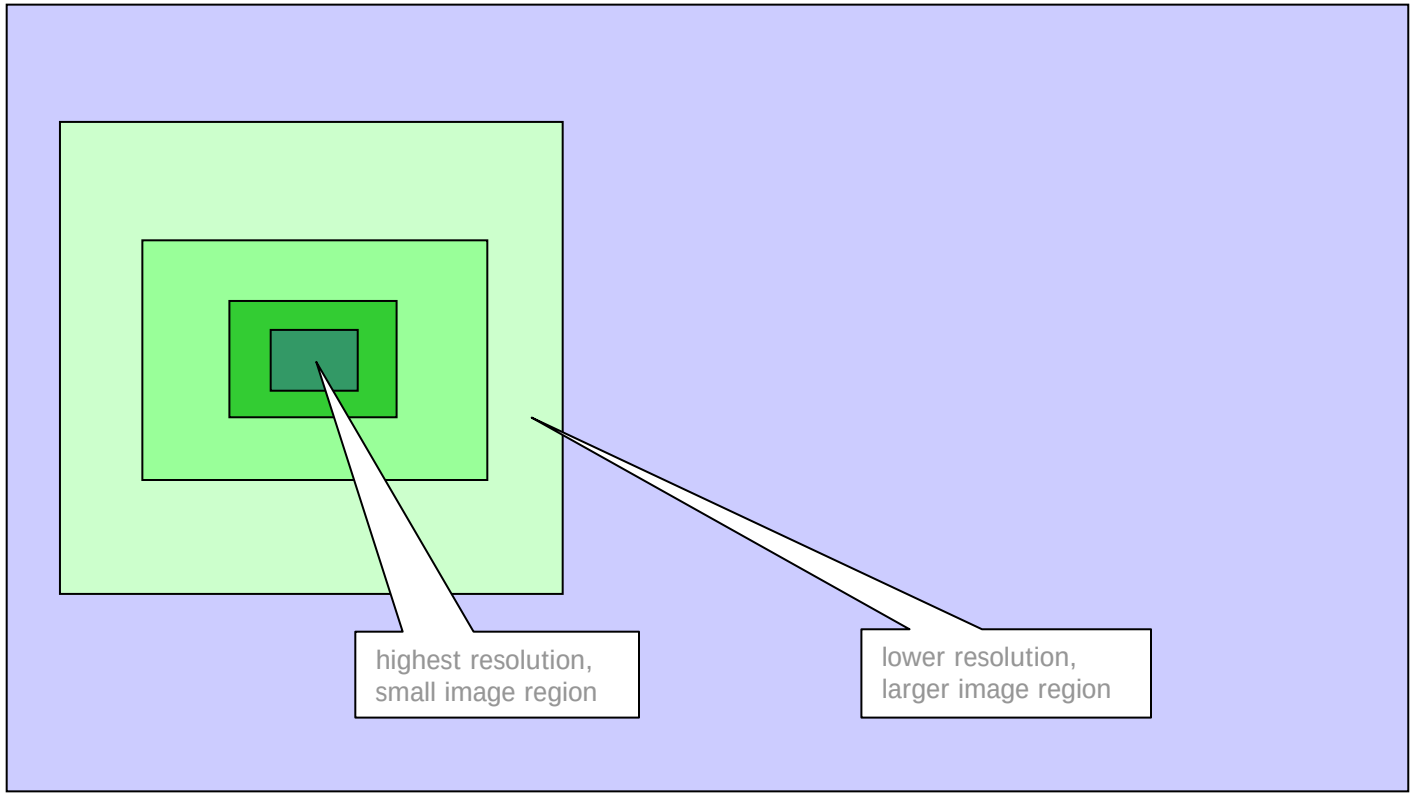
More complex image organization: Tiled



Does support efficient panning

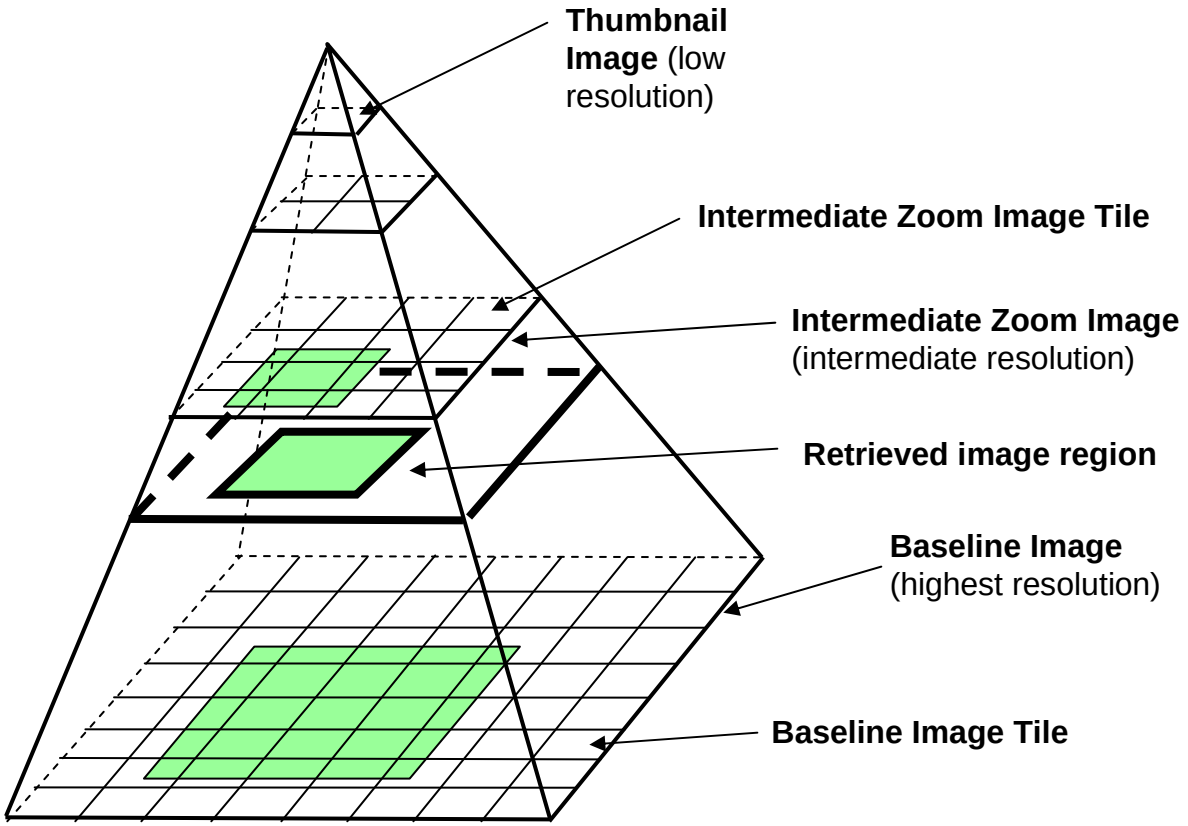
Access Patterns

Still have an issue with zooming



Access Patterns

Pyramid addresses zooming



Access Patterns

Incremental Transmission

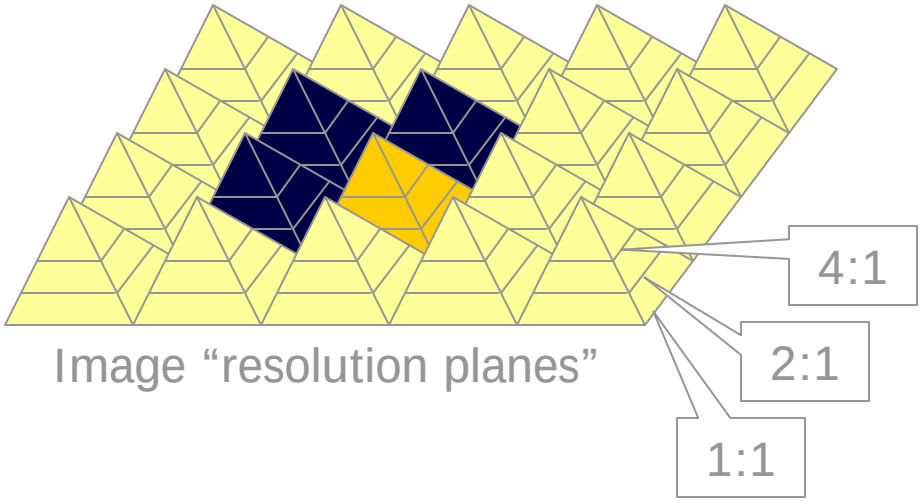
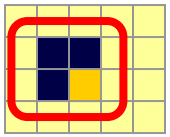
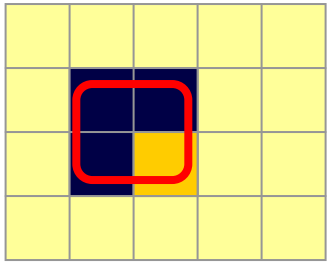


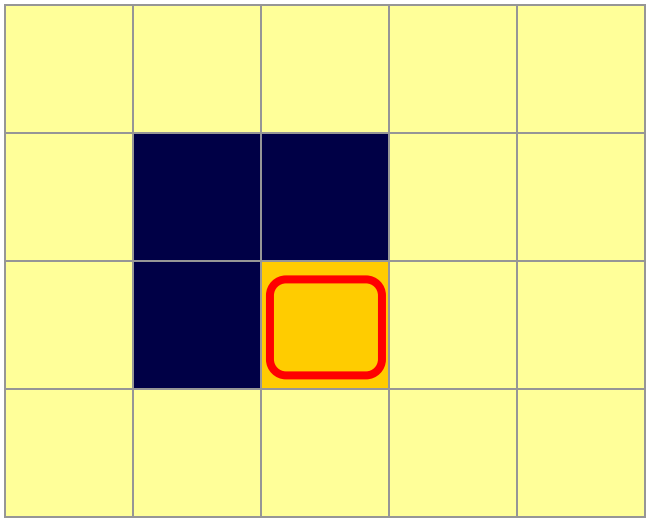
Image “resolution planes”



4:1 - 16 x 1/16



2:1 – 4 x 1/4



1:1 – 1 x 1

Whole Slide Imaging in DICOM

- Introduction
- Characteristics of Whole Slide Images (WSI)
  - Image size
  - Image access patterns
- **Storing WSI in DICOM**
  - NEMA Working Group 26
  - Proposed DICOM supplement for WSI
- IT Considerations for WSI
  - Storage
  - Bandwidth
- Solutions for Managing WSI
  - Access Models and Security
  - Compliance Characteristics
  - Performance

Storing WSI in DICOM

- DICOM = Digital Imaging and COmmunications in Medicine.
  - Standard for **file formats** and **inter-system communication**
  - Used in hospitals and labs for storing and managing radiology images (X-rays, MRIs, CTs, etc)
  - Systems which support DICOM are called PACS = Picture Archive and Communication Systems
  - Increasingly other medical disciplines also use DICOM to store images in PACS, such as cardiology; in this context such disciplines are called modalities
  - Pathology is the newest modality which is beginning to store images using DICOM in PACS

NEMA Working Group 26

- NEMA = National Electronics Manufacturer's Association
  - Administers DICOM standard
  - Working Groups propose Supplements to standard

- Working Group 26 formed two years ago to address storage of Pathology images in DICOM
  - Comprises pathologists, digital pathology vendors, and PACS vendors
  - Have created supplement with specimen model
  - Working on supplement for WSI storage and access

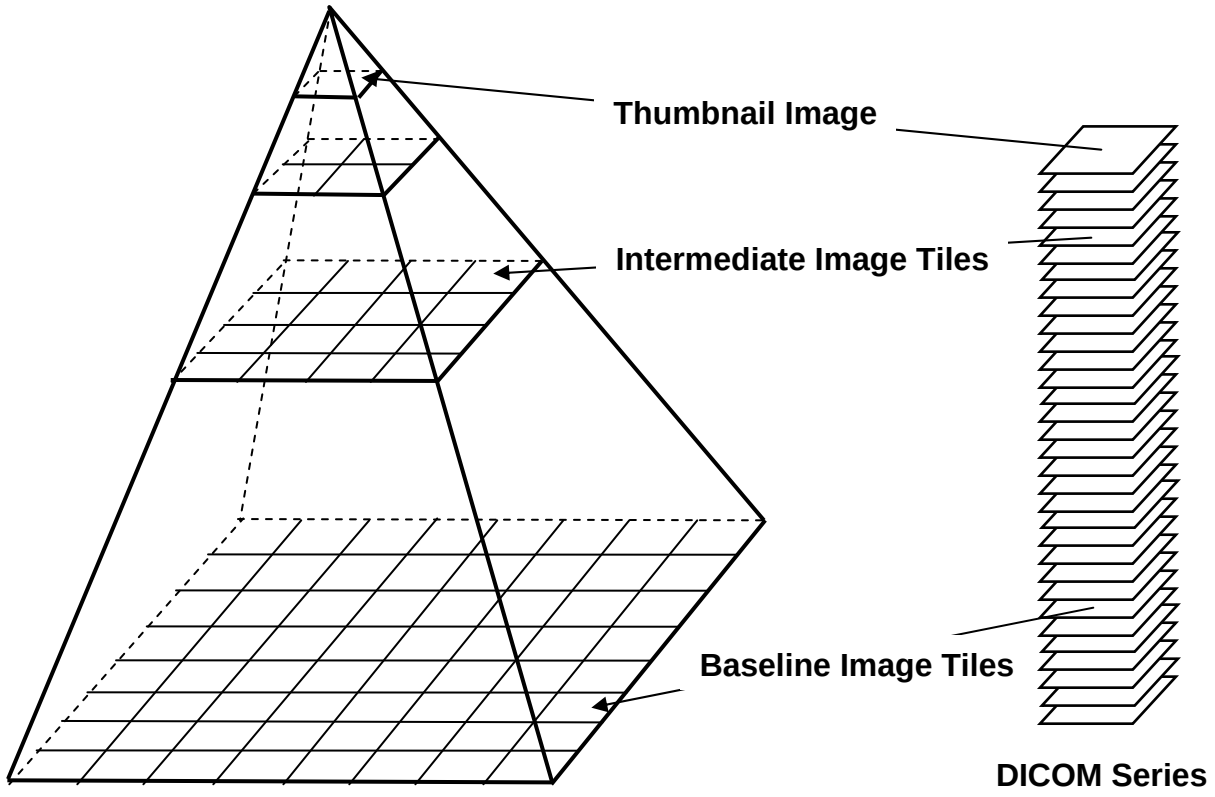
DICOM Limitations

DICOM Limitations

- 16-bit integer for pixel dimensions: 64K x 64K
- 32-bit integer for image object size: 2GB
- No subregion access to images
- Designed for store-and-forward access

Proposed Approach

Each “tile” of each level of each plane of WSI corresponds to image in series



Proposed Approach

Store WSI Pyramid as DICOM Series

- Tile size can vary (256 … 1K … 4K … 32K)
- Each resolution level stored separately in the series (typically thumbnail is first)
- Each Z-plane stored separately in the series
- Each image in series identified by coordinates
  - X, Y offsets
  - Zoom (resolution)
  - Z offset (plane)

Proposed Approach

Characteristics of Proposed Solution

- Works around DICOM limitations
  - 16-bit pixel dimension
    - tiles will be less than 64K x 64K
  - 32-bit image size
    - tiles will be smaller than 2GB
  - Subregion access to images
    - Each tile is individually addressable as an image
  - Designed for store-and-forward access
    - Supports incremental transmission via tile access
- Encompasses storage of single image as subset
  - Supports direct access to smaller images

Whole Slide Imaging in DICOM

- Introduction
- Characteristics of Whole Slide Images (WSI)
  - Image size
  - Image access patterns
- Storing WSI in DICOM
  - NEMA Working Group 26
  - Proposed DICOM supplement for WSI
- **IT Considerations for WSI**
  - Storage
  - Bandwidth
- Solutions for Managing WSI
  - Access Models and Security
  - Compliance Characteristics
  - Performance

Storage

Concerns: Capacity, Cost, Backup

Storage: Capacity, Cost, Backup

- **Typical slide = 250MB**
  - Digital slide = **20mm x 15mm** @ **20X** (.5 micron/pixel)
  - = 40,000 x 30,000 pixels = 1.2 Gpixel
  - x 3 bytes/pixel = 3.6GB
  - / 20:1 compression (JPEG2000) = 180MB
- **Typical lab: One year = 5TB = $5,000**
  - (**100 slides / day**) x (200 days / year) = 20,000 slides
  - = 5TB (1TB can hold 4,000 slides)
  - x $1,000 / TB ($500 disk + $500 server, etc) = $5,000
- **Typical server: storage up to 80TB = 16 years**
  - 2 controllers, each with 4 racks, each with up to 12 drives
  - With RAID 5, capacity = 2 x 4 x 11 (12-1) drives = 88TB
  - With RAID 6, capacity = 2 x 4 x 10 (12-2) drives = 80TB

IT Considerations for WSI
Bandwidth

High bandwidth
(high capacity, lower speed)

Low bandwidth
(low capacity, higher speed)

Concerns: Capacity, Cost

Bandwidth: Capacity, Cost

- **Typical user = 20K/s**
  - Computer screen = 1280 x 1024 pixels = 1.3Mpixel
  - x 3 bytes / pixel = 4MB
  - / 20:1 compression (JPEG2000) = 80KB / screen
  - / **screen every 4 seconds** = 20KB/s
- **Typical lab: 3 T1 lines = $18,000 / year**
  - **20 concurrent users** x 20KB/s = 400KB/s
  - / 200KB/s (T1 line = 1.5Mb/s = 200KB/s) = 2 T1 lines
  - + 50% (safety utilization) = 3 T1 lines
  - x $500 / month = $1,500 / month

- DS3 (aka OC3 aka T3) = 24 x T1 = 5MB/s
  - = 125 concurrent users = $6,000 / month

Whole Slide Imaging in DICOM

- Introduction
- Characteristics of Whole Slide Images (WSI)
  - Image size
  - Image access patterns
- Storing WSI in DICOM
  - NEMA Working Group 26
  - Proposed DICOM supplement for WSI
- IT Considerations for WSI
  - Storage
  - Bandwidth
- **Solutions for Managing WSI**
  - Access Models and Security
  - Compliance Characteristics
  - Performance

Access Models and Security

- Encrypt network data transmission
  - Use SSL (https://...)
- Authenticate all access to system
  - Positive identification of users
  - Timeout sessions, etc.
  - Log all sessions
- Establish Roles to manage user permissions
- Establish Data Groups to manage access permissions

Compliance Characteristics

- Encrypt network data transmission
  - Use SSL (https://...)
- Positively identify users
  - Support electronic signatures for “signing” reports
  - Support electronic signatures for data changes
- Use WSI image signatures
- Audit everything
  - All updates, all data access
  - Support sorting, filtering on audit reports

Performance

- Planning: Create usage model
  - Project users, sessions
  - Project number of cases, slides, images
  - Diagram workflows
- Measuring: Record activity
  - Sessions
  - Data access and updates
  - Image access and updates

Whole Slide Imaging in DICOM

- Thank you for your attention!
- **Questions** – now, via chat, or to ole@aperio.com