“VNA” Vendor Neutral Archive, The How and Why

Gary L Woodruff BASHA (RTR)
Enterprise Architect
Sutter Health
Data Growth overall in the Health Industry is matching most industry ranges of 20 to 50% per year.

Healthcare has been caught off guard in this explosion over the past 5-10 years.

We will explore the processes and some of the solutions that are growing within the Storage Industry to provide the Healthcare Industry with methods to solve these data growth problems.
Vendor Neutral Archive the How and Why

- Within the Health Industry, the “Medical Record Chart” has been the GOLD Standard which consisted of both paper and film type charts. This was the “Permanent” record.

- These solutions provided us with both a storage platform (a chart) and an organizational method for displaying that information (opening up tabs and sections to view).

- The problems in this format was only one interaction with the data from one person at a time. There also were problems moving this information around:
  - to multiple users
  - billing issues
The “solution” to address a more efficient management and sharing of information was to bring in computer systems:
- PACS systems for viewing -1990’s
- Electronic Billing (speed up revenue recovery) -1990’s
- Electronic Ordering: Laboratory, Radiology, Pharmacy. (printer based workflows) -1990’s/2000’s
- Thus the industry has been driven by “Application Solutions” that have been thought would solve these workflow problems. . .
What we have created has been, specialized silo’s of desperate information.

Supporting separate desperate Applications.
Siloed applications, support and storage

Radiology PACS
- UNIX
- RAID
- NAS Storage
- Vendor and App Support

Cardiology PACS
- LINX
- RAID
- File Storage
- Vendor and App Support

Document Scanning
- Windows
- RAID
- Block Storage
- Vendor and App Support
Vendor Neutral Archive the How and Why

Information Growth Rates:

- **Backup rates:**
  - 2010 ~ 0.8 PB
  - 2013 ~ 3.6 PB

- **Block Storage:**
  - 2010 ~ 0.4 PB
  - 2013 ~ 1.6 PB

- **File-NAS Storage:**
  - 2010 ~ 0.8 PB
  - 2013 ~ 2.6 PB
Vendor Neutral Archive the How and Why

- Lets look at the root causes here...

  - **Digital data Growth:**
    - 1975 CT scans 80 images per day 0.4 MB
    - 1984 CT scans 704 images per day 45.12 MB
    - 1998 CT scans 1,600 images per day 993.6 MB
    - 1998 MR scans 1,314 images per day 102.6 MB
    - 2009 CT scans 22,640 images per day 2,677.2 MB
    - 2009 MR scans 11,830 images per day 6,065.7 MB
    - 2009 PT scans 2,127 images per day 2,563.5 MB
Vendor Neutral Archive the How and Why

A Radiologist Reads...

- 1975  80 images per day          0.40  MB
- 1984  704 images per day         45.12  MB
- 1998  2,914 images per day       1,096.20  MB
- 2006  16,450 images per day      5,125.20  MB
- 2009  36,597 images per day      11,309.50  MB
Another look at Data Growth:

- 1998 Mammography (analog) 0.0 MB
- 2005 Mammography (digital) 137 GB
  > 240 / 500GB drives
- 2013 Mammography (new technology) 2138 GB
  > 626 / 3TB drives
### Grow Rates based on “owned data” from 2008 > 2026

<table>
<thead>
<tr>
<th>Year</th>
<th>Pedabytes</th>
<th>Year</th>
<th>Zedabytes</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>0.16</td>
<td>2007</td>
<td>0.0002</td>
</tr>
<tr>
<td>2008</td>
<td>0.33</td>
<td>2008</td>
<td>0.0003</td>
</tr>
<tr>
<td>2009</td>
<td>0.70</td>
<td>2009</td>
<td>0.0007</td>
</tr>
<tr>
<td>2010</td>
<td>1.46</td>
<td>2010</td>
<td>0.0015</td>
</tr>
<tr>
<td>2011</td>
<td>3.07</td>
<td>2011</td>
<td>0.0031</td>
</tr>
<tr>
<td>2012</td>
<td>6.44</td>
<td>2012</td>
<td>0.0064</td>
</tr>
<tr>
<td>2013</td>
<td>13.53</td>
<td>2013</td>
<td>0.0135</td>
</tr>
<tr>
<td>2014</td>
<td>28.42</td>
<td>2014</td>
<td>0.0284</td>
</tr>
<tr>
<td>2015</td>
<td>59.68</td>
<td>2015</td>
<td>0.0597</td>
</tr>
<tr>
<td>2016</td>
<td>125.33</td>
<td>2016</td>
<td>0.1253</td>
</tr>
<tr>
<td>2017</td>
<td>263.19</td>
<td>2017</td>
<td>0.2632</td>
</tr>
<tr>
<td>2018</td>
<td>552.70</td>
<td>2018</td>
<td>0.5527</td>
</tr>
<tr>
<td>2019</td>
<td>1160.68</td>
<td>2019</td>
<td>1.1607</td>
</tr>
<tr>
<td>2020</td>
<td>2437.42</td>
<td>2020</td>
<td>2.4374</td>
</tr>
<tr>
<td>2021</td>
<td>5118.58</td>
<td>2021</td>
<td>5.1186</td>
</tr>
<tr>
<td>2022</td>
<td>10749.02</td>
<td>2022</td>
<td>10.7490</td>
</tr>
<tr>
<td>2023</td>
<td>22572.94</td>
<td>2023</td>
<td>22.5729</td>
</tr>
<tr>
<td>2024</td>
<td>47403.18</td>
<td>2024</td>
<td>47.4032</td>
</tr>
<tr>
<td>2025</td>
<td>99546.68</td>
<td>2025</td>
<td>99.5467</td>
</tr>
<tr>
<td>2026</td>
<td>209048.03</td>
<td>2026</td>
<td>209.0480</td>
</tr>
</tbody>
</table>
Vendor Neutral Archive the How and Why

Now  Then
Vendor Neutral Archive the How and Why

Problems with Digital File Solutions:

- Physical Space Constraints (power, cooling)
- Need to consolidate data (compression, de-dup)
- Blend data sets for better access
- Create an unique identifiable data set
ILM elements (data retention guidelines)

- Title 22 (ca.gov), FDA (mqsa), HIPAA
- Sets an obligation criteria to retain data
  - Date of birth
  - Date of data creation
  - Data type, i.e. results, email, breast health, accounting
- 7 years from creation date (all medical records)
- Minor’s 18 years from birth +1
- 10 years from creation date (all mammograms)
- 3 years from creation date (all email transactions)
Analog File Management Solutions:

- Physical Space Constraints
- Need a method to consolidate data
- Evolved into a very sophisticated management system
- Creates an unique identifiable data set

Unique I.D.

Year to Destroy
ILM elements allow the institution to delete stored information within legal constraints. The average reduction of long term storage needed is reduced at a rate of 30%.

Tiering data by clinical need also helps reduce storage costs using SDD, SAS, SATA and Tape accordingly:

- Using comparative data for diagnosis. Aged data is not relevant to the current diagnosis if over 4-7 years old.
- This data becomes Legally constrained so the only reason it is retained is due to the possibility of legal intervention which does not need milli-second response times to retrieve.
Vendor Neutral Archive the How and Why

**DICOM METADATA**

<table>
<thead>
<tr>
<th>Tag</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0008,0008) [CS]</td>
<td>DERIVED\PRIMARY</td>
</tr>
<tr>
<td>(0008,0016) [UI]</td>
<td>1.2.840.10008.5.1.4.1.1.1</td>
</tr>
<tr>
<td>(0008,0018) [UI]</td>
<td>1.3.46.670589.30.1.2.1.72052720229.1177467833617.1</td>
</tr>
<tr>
<td>(0008,0020) [DA]</td>
<td>2007-04-24</td>
</tr>
<tr>
<td>(0008,0022) [DA]</td>
<td>2007-04-24</td>
</tr>
<tr>
<td>(0008,0023) [DA]</td>
<td>2007-04-24</td>
</tr>
<tr>
<td>(0008,0030) [TM]</td>
<td>19:22:04</td>
</tr>
<tr>
<td>(0008,0032) [TM]</td>
<td>19:22:49</td>
</tr>
<tr>
<td>(0008,0033) [TM]</td>
<td>19:22:49</td>
</tr>
<tr>
<td>(0008,0040) [US]</td>
<td>0</td>
</tr>
<tr>
<td>(0008,0041) [LT]</td>
<td>IMA NONE</td>
</tr>
<tr>
<td>(0008,0050) [SH]</td>
<td>MGR0711433</td>
</tr>
<tr>
<td>(0008,0060) [CS]</td>
<td>CR</td>
</tr>
<tr>
<td>(0008,0070) [LO]</td>
<td>Philips Medical Systems</td>
</tr>
<tr>
<td>(0008,0080) [LO]</td>
<td>Marin General Hospital</td>
</tr>
<tr>
<td>(0008,0081) [ST]</td>
<td>250 Bon Air Road Greenbrae CA 94904 USA</td>
</tr>
<tr>
<td>(0008,0090) [PN]</td>
<td>CLARK PA^NANCY^J.</td>
</tr>
<tr>
<td>(0008,1010) [SH]</td>
<td>mgEe3</td>
</tr>
<tr>
<td>(0008,1030) [LO]</td>
<td>Elbow</td>
</tr>
<tr>
<td>(0008,1032) [SQ]</td>
<td>Sequence with 1 Items</td>
</tr>
<tr>
<td>(0008,103e) [LO]</td>
<td>Lateral</td>
</tr>
<tr>
<td>(0008,1040) [LO]</td>
<td>ER</td>
</tr>
</tbody>
</table>
Vendor Neutral Archive the How and Why

ILM Policy Impact on Storage Needs

Managed vs Non Managed Archive

- Non Managed
- Managed

Years of Storage

Pedabytes

Managed vs Non Managed Archive

Years of Storage

Pedabytes
Vendor Neutral Archive the How and Why

Meta Data Management

Standard Vendor Archive Stacks Usually Application Controlled
**Problem:**

- Shared storage gives additional capacity, lowers cost:
  - Total: 375.0 TB Capacity
  - Total: 133.5 TB in reserve
  - 35.6% reserve capacity

- 100TB Capacity
  - 11TB in reserve
  - 89% FULL
  - Application A
    - Radiology PACS

- 200TB Capacity
  - 100TB in reserve
  - 50% FULL
  - Application B
    - Cardiology PACS

- 75TB Capacity
  - 22.5TB in reserve
  - 70% FULL
  - Application C
    - Document Scanning
Problem:

Over the last 18 years, typical Radiology PACS systems have “Upgraded” 3 times forcing migrations each time.
Vendor Neutral Archive the How and Why

Solution:

- VNA, Single archive which allows multiple Applications to sit on top. The inherent metadata is the common denominator to “glue” the desperate data into one system.
  - DICOM Modalities
  - Non DICOM elements:
    - PDF Scanning: document scanning
    - J-peg images: wound care, dermatology, endoscopy
    - Sleep Centers, Hearing Centers
    - XDS services (many archive functions from applications are now supporting this protocol)
Vendor Neutral Archive the How and Why

Migration costs over the life of the data:

- **Data Migration**
  - 1/1/2013
  - $15.3 Million 3K/TB
  - 1/1/2013

- **Data Migration**
  - 1/1/2014
  - $61.4 Million 1.5K/TB
  - 1/1/2015

- **Data Migration**
  - 1/1/2016
  - $155 Million .5K/TB
  - 1/1/2017

Vendor Neutral Archive
Vendor Neutral Archive the How and Why

The Digital Information Solution devised by the Government using funding as the carrot is the EHR, Electronic Health Record.

Over 640 Vendors

- EPIC
- NextGen
- Cerner
- SMS
- AthenaHealth
- Allscripts
- G.E. Healthcare
- McKesson
What has this caused?

- Large monolithic systems
  - One size fits all...
  - They need to “own” all applications
  - In order to provide the application depth, they glue on applications, additional silo’s of information in some cases.
- They do “archive” fixed data in the form of reports etc.
- Some of the “add-on” applications also archive fixed data.

This leads us to the final storage answer...
Vendor Neutral Archive the How and Why

 DataTypes Driven Archive’s

- Metadata rich archiving
- Able to store to multiple layers:
  - CIFS, NAS, Object, Tape, Support Multiple API’s
  - Migrate between platforms
  - Some offer Metadata scraping technology

- NetApp
  - Grid - (ByCast)

- Hitachi
  - HDS

- IBM
  - SONAS

- EMC (No tape)
  - ViPR
Vendor Neutral Archive the How and Why

VNA Ingestion Layer

Metadata Controlled

Archive Management Software Layer

Metadata Controlled

Object Storage
CIFS Storage
NAS Storage
Object Storage
Tape Storage

Radiology PACS
Cardiology PACS
Other "Ology"
EHR
Other App's

USERS
As the complexity of information increases along with the need to share that information with others, we have found that our standard methods of supporting applications and storage infrastructure are not meeting the needs of our customers. We have allowed vendors to build solutions that protect their interests and not provide us with tools that help us “own” the information so it can then be managed.

In the case of the Vendor Neutral Archive, we both “own” that information and manage the information, which allows us deploy more cost effective methods of ILM and storage infrastructure based on the intelligent use of embedded Metadata.
What we have discovered is the information that we are charged with in many cases has a life of 20 to 30 years. This far outweighs any technology that we have at our finger tips and causes us to constantly migrate data from one platform to the next again providing the vendors the opportunity to charge us for that privilege of owning and being responsible for that data. With the implementation of a VNA and Software driven storage solutions, the “owner” of the data now has the tools to migrate and prune data as needed so the most cost efficient methods can be deployed but not sacrifice performance since we now “understand the data” rather than just “storing the data”.