Electronic Records Storage Options and Overview

www.archives.nysed.gov

Objectives

- Understand the options for electronic records storage, including cloud-based storage
- Evaluate the options best suited for your records
- Understand electronic record backup methods
- Interpreting cloud computing contract provisions

Agenda

- Electronic Storage Overview
- Similarities of Paper and e-Record Storage
- Storage Landscape - Technology & Methods
- Cloud Computing Options
- Backup & Protection
- Strategies Moving Forward

Storage Challenges

- Growth and velocity - Annual growth of 40%
- Variety of data sources and uses
- Most stored data is non-critical or copies

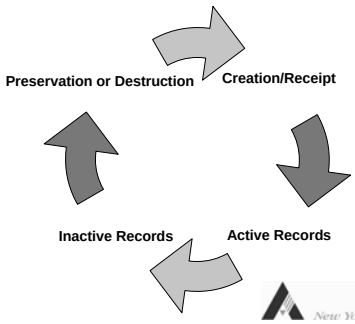
New York State Archives

Additional Storage Challenges

- Need to respond to eDiscovery and FOIL requests
- Migration and upgrades costs can be significant
- Disconnect with users, records managers and IT
- Storage options – which is best?

New York State Archives

Records Lifecycle



New York State Archives

Records Management is Key

- Storage planning begins with RM
  - Follow basic records management principles:
    - Conducting an inventory
    - Developing retention schedules
    - Managing the lifecycle of records

Additional Workshops

- Other workshops to assist you
  - Electronic Records Inventory
  - Managing Electronic Records
  - ECMS
  - Preserving Electronic Records webinar

What do you see are the challenges with electronic record storage?

How are you addressing it?

Similarities between hardcopy and electronic storage

NOT SO DIFFERENT

New York State Archives

Similar to Paper

- Centralized storage
- Managed
- Secure
- Protected
- Organized

New York State Archives

Centralized

Paper	Electronic
• Central file room – Controlled	• Central storage device – Server/network-based
• Few file rooms – More difficult	• Multiple servers – Multiple drives to manage
• No storage – Managing their own	• Desktop storage – C: Drive

Better

Avoid

New York State Archives

Managed

Paper	Electronic
• Records Manager/RMO	• ECMS Software • Administrator/RMO

Protected

Paper	Electronic
• Environmentally sound – Temperature and humidity • Fire protection	• Environmentally sound – Temperature and humidity • Off-site backup

Organized

Paper	Electronic
• Properly labeled boxes • Finding aids • Managed retention	• Properly indexed • Retention managed

Secure

Paper	Electronic
• Controlled access to storage area • Authority list	• Firewall • Passwords • Encryption

Can you think of other similarities between electronic and paper storage?

Technologies, methods and storage options

THE STORAGE LANDSCAPE

Storage Categories

- Local
- Network
- Cloud

Local Storage

- Workstation-based
  - PC, Laptop, tablet
    - C: drive
    - Desktop
- Avoid using for storage
  - Usually not backed up
  - Single drive failure = loss everything (apps and data)

Network Storage

- Single server storage
- Multiple individual servers
  - Application server(s)
    - Finance, GIS,
  - File, Email
  - Web, database
- Dedicated storage device(s)
  - Storage Area Network (SAN)
  - Network Attached Storage (NAS)

Server-class Storage

- Server hard drives built to run 24 x 7
  - Often hot swappable
- RAID (Redundant Array of Inexpensive Disks)
  - RAID 1 (2 discs mirrored)
  - RAID 5 (striping data across multiple disks)
    - e.g. 3 disks in array; any which can fail with no noticeable difference

Network Attached Storage (NAS)

- Single storage device connected to network
- Provides centralized storage
- Consolidates separate server storage
  - More efficient use of total storage space
  - Easier to backup than separate server-based storage
- Common for small to mid-sized organizations

Storage Area Network (SAN)

- Local network of multiple storage devices
  - Storage spread across multiple servers
- Higher performance than NAS
  - Often higher capacity
- Higher cost
- Ideal for larger operations

Cloud Storage

- Data stored remotely on 3rd party's servers
- Can include application and data, or just data
- Accessed through Internet or dedicated communication line
  - e.g. Google Docs, email, state agency server, BOCES (for schools), vendor's application, etc.

More discussion on cloud computing coming…

Which to Choose?

- Not an either/or decision
  - Often a hybrid approach is used
- Depends on overall IT strategy
  - But move towards consolidating repositories
- Consider
  - Current network architecture
  - Performance
  - Maintenance
  - Cost

Do you know the type of storage used in your government?

What are your top concerns when it comes to storage?

What is it?

CLOUD COMPUTING

What is Cloud Computing?

- General term for anything involving the delivery of hosted services over the Internet
- Scalable on-demand network access to shared pool of computing resources
  - e.g. servers, storage, applications and other services

Do you use cloud computing now?

Characteristics of the Cloud

- On-demand, self-service
  - Location independent – access form anywhere
- Massive scalability
  - Resource pooling & rapid elasticity
- Measured service
  - Use as much as needed
- Broad network access
  - Delivered via Internet

Additional Characteristics

- Cloud computing often leverages:
  - Resilient services
  - Distributed datacenters
  - Advanced security technologies
  - Low cost entry

Clouds vs. Traditional Hosting

- Sold on demand
  - Can be by the minute, the hour, or other time period
- Elastic
  - A user can utilize as much or as little of a service as needed at any given moment
- Service is fully managed by the provider
  - User only needs a PC with Internet access

Delivery Models

- Software as a Service (SaaS)
- Platform as a Service (PaaS)
- Infrastructure as a Service (IaaS)

Cloud Service Models

1. Software as a Service (SaaS)
   - Use provider's applications over a network
   - Log-on from anywhere for access to specific software
   - Typical applications can include e-mail, payroll, Human Resources, CRM, etc.
     - e.g. Google Apps, ADP, Microsoft Office 365, SugarCRM, Gmail

Cloud Service Models

2. Infrastructure as a Service (IaaS)
   - Provides *hardware* such as servers, network equipment, and data storage on a scalable basis.
     - e.g. Amazon's Cloud Formation EC2 Services; Rackspace Cloud; HP Cloud, etc.

Cloud Service Models

3. Platform as a Service (PaaS)
   - Provides an *application development environment* allowing users to collaborate, develop, test, deploy, host and maintain applications in the cloud.
     - e.g Google's AppEngine; Microsoft's Azure Compute, etc.

Cloud Deployment Models

- Private cloud
  - Operated solely for one organization
- Community cloud
  - Shared by several organizations in one community
- Public cloud
  - Available to anyone
- Hybrid cloud
  - Private cloud for mission-critical applications but hosting public website with cloud provider

Cloud Strengths

- Scalability
- Cost savings
  - Only paying for the storage needed
  - Less staff needed to maintain system
  - No need to own hardware
- Ease of implementation and expansion
- Elasticity of storage
- Disaster preparedness through redundancy

Potential Concerns

1. Security
2. Ownership & Control
3. Vendor Viability
4. Lack of Open Standards
5. Compliance
6. Access to Data (uptime)
7. Protection
8. Standardization Over Customization

Potential Concerns

- Trusting vendor's security model
  - Privacy of information sitting outside of own firewall
    - Where is the data stored and how is it secured?
      - Encryption in route?; Encryption while stored?
  - Authentication of data
- Ownership & control
    - Control of data/infrastructure
    - Depository agreement in place?
    - Indirect administrator accountability
    - Loss of physical control

Potential Concerns

- Vendor viability
  - What happens to the data if the provider goes under?
- Lack of open standards
  - Governing how data is stored & manipulated
  - How complicated is it to transition to another environment or system (portability)?
- Compliance
  - Inability to respond to audit requirements

Potential Concerns

- Performance & availability
  - Guarantee of availability and performance of services when 24/7 access is required
- Disaster protection
  - Backup procedures and standards
  - Redundancy in place?
  - Disaster recovery & business continuity
- Inability to customize
  - How tailored can the system be to meet your needs?

What concerns do you have?

Do you see benefits in your organization?

BACKUP AND PROTECTION

Why So Important?

- Disc drives are mechanical and will fail
  - Disc spins at 5,000 to 15,000 RPM
- Environmental conditions affect lifespan
  - Temperature and humidity
  - Dust
- Disaster protection
  - Fire and flooding
  - Static and electrical surges

Key Characteristics

- Performed regularly
- Reliability
- Off-site storage
- Tested
- Requires hardware and software

Methodologies

- Full
- Incremental
- Differential

Backup Storage Options

- Removable media
  - Tape
  - CD/R
  - DVD/R
  - USB device
- Disc to disc
  - Server to server
  - Server to external hard drive
- Off-site service

Removable Media

- Tape
  - Most often used
  - Large data sets making it
- External hard drive
  - Inexpensive, slower interface
  - Often left on-site
- CD/R & DVD/R, USB drive
  - Used for individual backups, not organization wide

Disc to Disc

- Becoming more popular
- Faster than tape
  - Backup & restoration purposes
- Caution with potential corruption
  - Must include multiple periodic snap shots of data

Off-site Service

- Data Vaulting
- Off-site data protection
- Part of a disaster recovery plan
- May be owned by organization or contracted via a service

Permanent Storage at Archives

- All media are temporary storage
- Media appropriate for transfer
  - CD-R, CD-ROM
  - DVD-R, DVD+R, DVD-ROM
- All media must be labeled
- Appropriate file formats
  - PDF/A
  - TIFF

Determine Acceptable Risk

- What is acceptable data loss?
  - A days worth of work? (can be up to 48 hours)

How much do you know about your data backup process?

How often is everything backed up?
Where is it stored?
What is your acceptable risk level?

Considerations moving forward

STRATEGIES

Get on the Same Page

- Talk to IT
- Talk to your users
- Talk to vendors

Questions to Ask

- What applications ‘create’ records
  - e.g. accounting, permit tracking; but not Acrobat, etc.
- Where are they stored?
  - Have the servers and storage devices pointed out
- How often is the data backed up?
  - Methodology
- How often is the backup taken off-site?
  - Where?

Regulations

- 188.17: Review of agency proposals to establish or lease facilities or to contract for services to store records.
  - The director, at the request of the Division of the Budget, shall review proposals by agencies to establish or lease facilities or to contract for services to store inactive records, vital records, master copies of microforms, or backup copies of electronic records.
  - The director shall recommend to the Division of the Budget approval of requests by agencies to establish or lease facilities or to contract for services to store inactive records when greater economy or efficiency can be achieved by such operation than by use of a records center facility and when such operation will adequately provide for the protection and servicing of records.

Cloud Contract Questions

- Costs
- Disaster Mitigation
- Ownership
- Records Transfer
- Accessibility - provider’s performance

Contract Questions

- Security
- Storage
- Retention and Disposition
- Termination of Contract

Develop Policies

- Create policies that address
  - Disaster response and recovery
  - Files management
  - Access

LAST WORDS OF ADVICE

In Conclusion

- Electronic data storage is an issue that must be addressed
- There is no one-size-fits-all solution
- Good records management is key to success

Summary

- Research what others have done
- Establish your business need
- Determine what you need to do
- Calculate costs and time saved
- Consider questions to ask vendors
- Understand and control your contract

Electronic Records Storage Options and Overview

Thank You!

For more information:
www.archives.nysed.gov