Data Storage for Research

Michael Pinch
Intro

• Data storage is a world full of tradeoffs
  – Read vs Write Speed
  – Availability vs Security
  – Cost vs Speed
• This presentation is to talk about the different types of data we deal with, options for storage, and risks of each
• We highlight University policy and cite how to get more information
Types of Data

• PHI
  – Identifiable health information that can be linked to a particular person

• PCI Account / Cardholder Data
  – Primary account number, cardholder name, service code, and expiration date

• Legally Restricted (UR Policy)
  – Data which must be protected by law such as PHI, Social Security numbers, and account data

• Confidential (UR Policy)
  – Information which may be sensitive or proprietary such as personnel records and unpublished
Types of Storage

- Hard Drive
- Server
- Database
- Share Drives
- Network Drives
- Removable Drive
- CD
- DVD
- Cloud
- USB
Share Drives / Network Drives

- Access is controlled via user account
- Data is backed up nightly
- Remotely accessible through VPN
Servers / Databases

- Servers must be in data center
- REDCap (Research Electronic Data Capture) may be a viable option – Contact Academic IT for more information
Cloud Storage

- Dropbox is extremely popular for convenience, capacity, and collaboration
- Dropbox is NOT an approved medium for data storage at URMC, especially legally restricted data
- URMC has an enterprise contract with Box (the major competitor to Dropbox). Legally, we will be able to store any data in Box we choose.
- Box allows for syncing to desktop, mobile, web, etc
Other Cloud Storage

• URMC is piloting Amazon Web Services
  – In addition to EC2, it offers S3, simple storage solution
  – Inquire with me if you’re interested in trying it

Storage Pricing

<table>
<thead>
<tr>
<th></th>
<th>Standard Storage</th>
<th>Reduced Redundancy Storage</th>
<th>Glacier Storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>First 1 TB / month</td>
<td>$0.095 per GB</td>
<td>$0.076 per GB</td>
<td>$0.010 per GB</td>
</tr>
<tr>
<td>Next 49 TB / month</td>
<td>$0.080 per GB</td>
<td>$0.064 per GB</td>
<td>$0.010 per GB</td>
</tr>
<tr>
<td>Next 450 TB / month</td>
<td>$0.070 per GB</td>
<td>$0.056 per GB</td>
<td>$0.010 per GB</td>
</tr>
<tr>
<td>Next 500 TB / month</td>
<td>$0.065 per GB</td>
<td>$0.052 per GB</td>
<td>$0.010 per GB</td>
</tr>
<tr>
<td>Next 4000 TB / month</td>
<td>$0.060 per GB</td>
<td>$0.048 per GB</td>
<td>$0.010 per GB</td>
</tr>
<tr>
<td>Over 5000 TB / month</td>
<td>$0.055 per GB</td>
<td>$0.037 per GB</td>
<td>$0.010 per GB</td>
</tr>
</tbody>
</table>
Protection

• Encryption in Rest
• Encryption in Motion
• Password Protected < Encrypted
Encryption Methods

- **VPN**
  - Provides encryption for data in motion

- **HTTPS**
  - Provides encryption for data in motion over the web

- **“Zip” Software**
  - Gives you the option to encrypt a zip file
  - Share the password with someone over a different channel than how you share the file!

- **Secure Email**
  - Put !secure in the subject of your email
Encryption Methods

• CD / DVD
  – Encrypted DVDs are available!

• USB / Removable /External Drive
  – Buy an encrypted USB Device
  – URMC Managed systems will have software on them in early 2014 that will turn all “dumb” USB drives into encrypted drives automatically
Encryption Methods

- Full disk for users
  - PointSec (URMC Enterprise Solution)
  - FileVault (Mac Native Solution)
  - BitLocker (Windows Native Solution)
  - TrueCrypt (Open Source Solution)

- Most of these tools can be used for full disk encryption or file/folder encryption
Encryption Performance

PassMark Composite Rating Comparison (Performance)

- SafeBoot
- GE
- PGP
- BitLocker
- Utimaco
- Check Point

Disk Performance in % of Maximum Performance (without Encryption)

Performance

- Check Point
- BitLocker
- Utimaco
- PGP
- GE
- SafeBoot
Encrypted Devices

• Camera
  – Samsung Galaxy Camera – Fully Encrypted!
    • Available in the bookstore - $399

• Voice Recorders
  – Olympus DS-3500 - $399
  – Philips LFH-9600 - $499
Backup / Redundancy

- URMC File server data backed up for 30 days
  - 14 Days on disk in the Primary Data Center
  - 30 Days on encrypted tape in the Med Ctr Annex
- Full backups performed weekly
- Incremental backup performed daily
Compliance

• HIPAA
  – Requires PHI to be encrypted at rest and in motion

• FISMA
  – New grants may include FISMA security requirements
  – Medium level security may require encryption of mobile devices and portable storage media

• PCI
  – All cardholder data must be encrypted in transit
  – Storing cardholder data significantly increases the cost to ensure compliance and is strongly discouraged
Recommended Practices

• All research data should have backup copies
• Minimize the use of unencrypted external hard drive, portable devices, and flash drives
• Store confidential and legally restricted data on network drives / shares

Your research data may not contain patient information, but it should remain your data!
Places not to store your data

• Unencrypted Devices
• Local Hard Drive
• Cloud Storage Providers
  – Dropbox
  – Google Drive
Policies

• Information Security Policies, URMC
  – https://intranet.urmc-sh.rochester.edu/policy/hipaa/PolicyManual/

• UR Data Classification
  – http://www.rochester.edu/it/policy/assets/pdf/INFORMATION_TECHNOLOGY_POLICY.pdf
What to do in case of loss?

• Contact the URMC Information Security Office
• If the lost device is encrypted, no further action should be necessary
Destruction of Data

• Legally restricted and confidential data must be securely removed when no longer needed
• Hard drives should be securely wiped before disposal, leaving the organization, or reuse
Questions

• Contact URMC Information Security Office
  – CISO: Michael_Pinch@urmc.rochester.edu
  – 585.305.5449