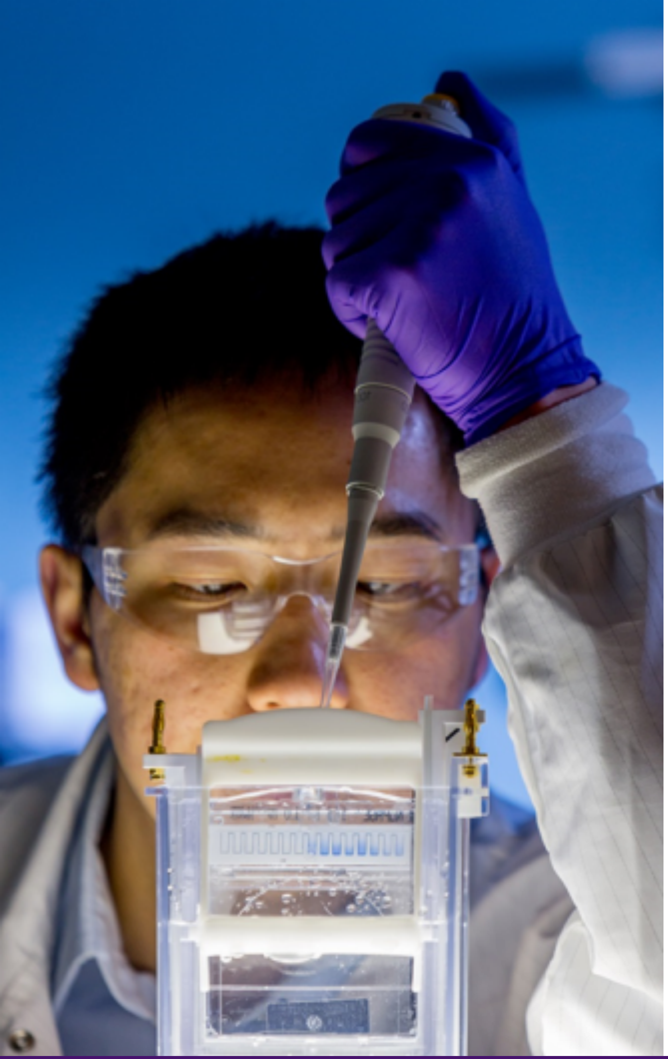


Data Integrity on Large Networks

Scott Davis
Associate Director
PPD Information Technology

scott.davis@ppdi.com

HELPING DELIVER LIFE
CHANGING THERAPIES

PPD®

Agenda

1 What is data?

2 What is a large network?

3 Protecting your data

4 The cloud

5 Cloud compliance

6 Cloud sharing

What is Data?

Data (plural of *Datum*) – Information in digital form that can be transferred or processed.[1]

Types of Data

- Static
  - A static record format, such as a paper or electronic record, is one that is fixed and allows little or no interaction between the user and the record content. 2
  - Examples: PDF Report
- Dynamic
  - Records in dynamic format, such as electronic records, allow an interactive relationship between the user and the record content. 2
  - Examples: Database, Proprietary File Format

What Data to Keep for Regulated Studies?

Static Data

+

Dynamic Data

=

ALL THE DATA!

All The Data?

Includes

- Reports
- Tables
- Export Files
- Proprietary Instrument Data Files
- Databases

YUMMY DATA!

EVERYTHING!

How Much Data?

Data acquisition instruments

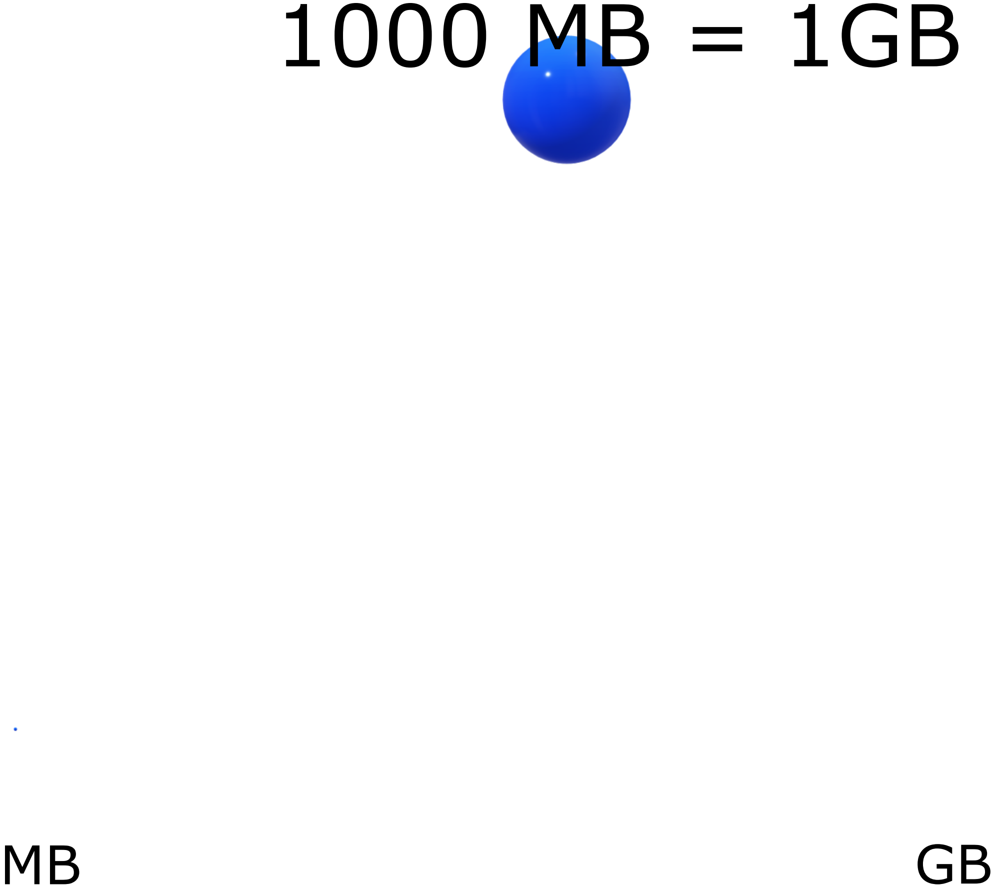
- In the past, less than 100MB per batch
- Now, upwards of 50+GB per batch = over 1TB per month!

Enterprise systems

- Multiple TB of data every year!

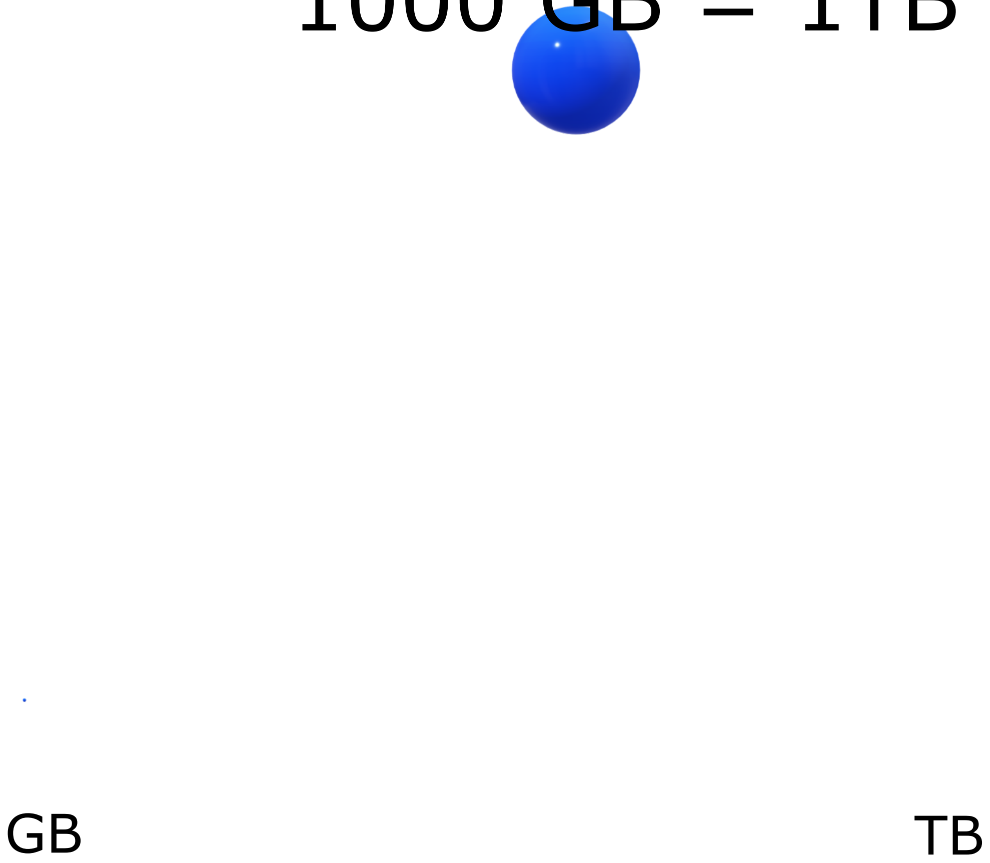
Scale MB to GB

1000 MB = 1GB



Scale GB to TB

1000 GB = 1TB



Seriously?!
It's the same
slide! 😒

Overall Scale

1,000,000 MB = 1000 GB = 1TB

Example:

1 MP3 Song = 5MB

1GB = 200 Songs

1TB = 200,000 Songs

What is a Network?

Network – a system of computers and peripherals that are able to communicate with each other.[1]

What is a Large Network?

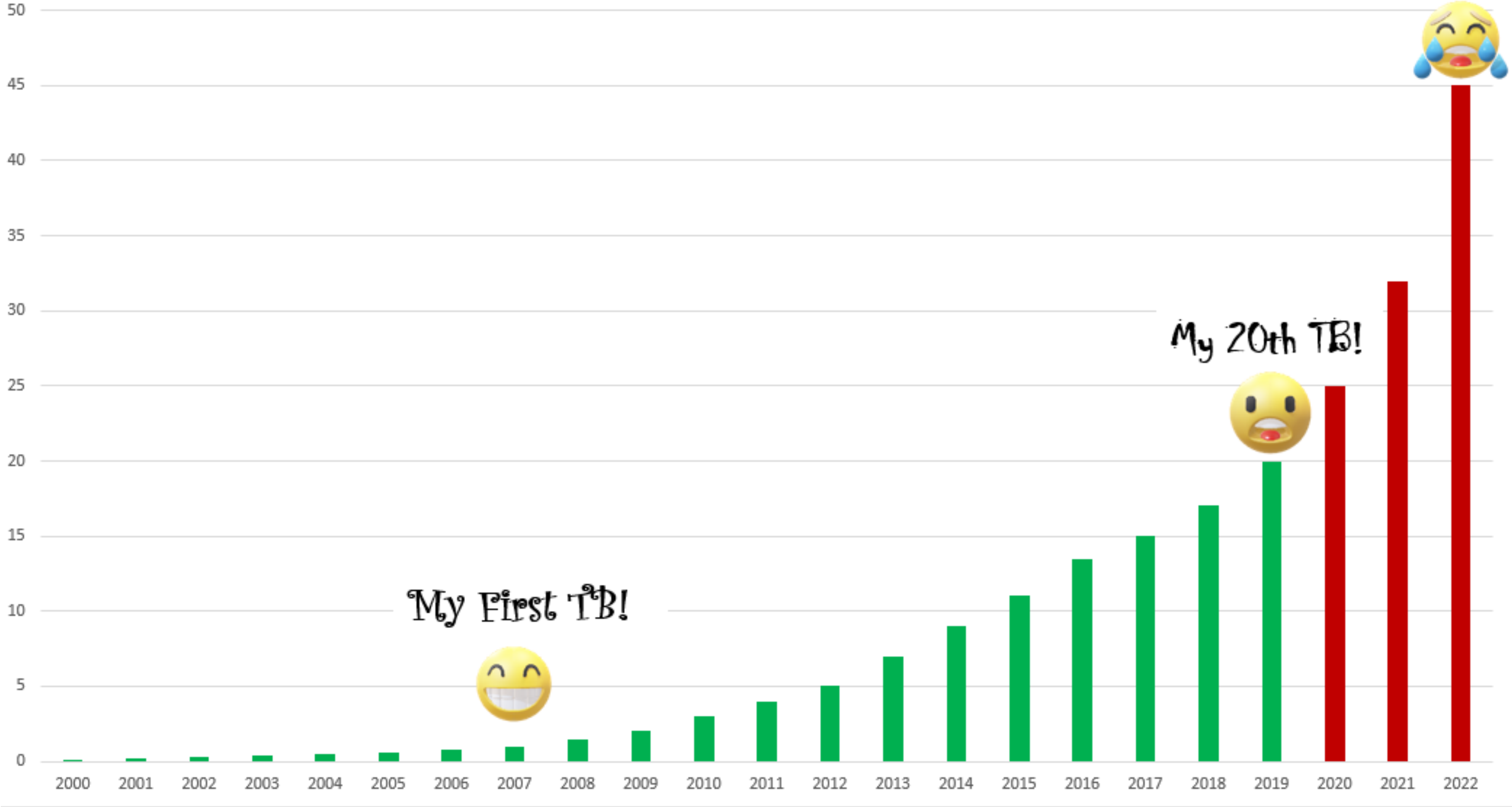
“Large” is a relative term

-One site with many users

-One organization across several sites

For the purposes of this presentation, a ***large network*** *is an organization with multiple locations that share a common network structure.*

PPD Bioanalytical Lab Data Over Time



Protecting Your Data

- Creation
- In-transit
- Storage
- Backup
- Disaster recovery

Data Creation

- Utilize an instrument subnet
    - Segregated from the rest of your network
    - Instruments and their computers only
    - No access to the internet
- Save to a network location if possible
- If system requires you to save to the local computer, move the data as soon as possible to a network location
- Utilize compliance functions if available

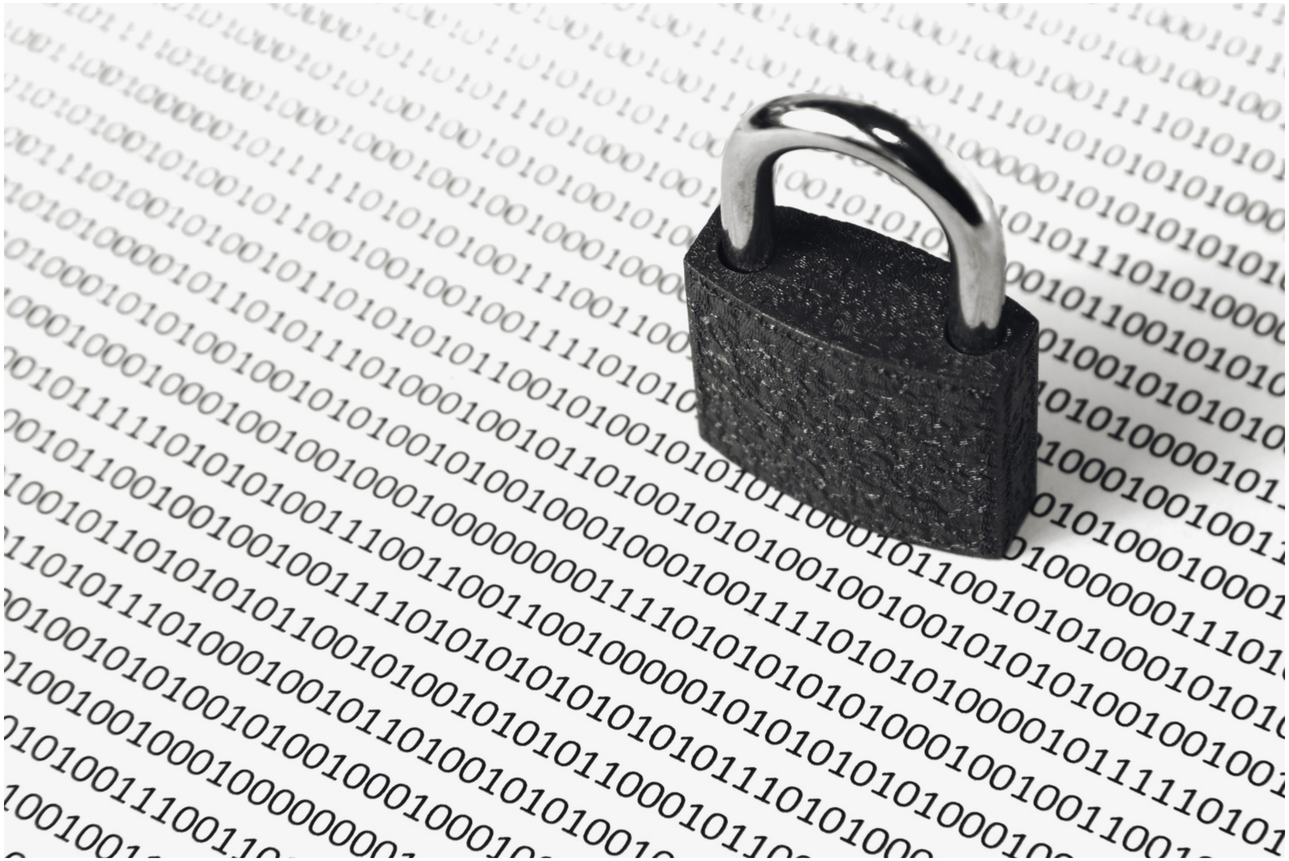
Data In Transit

- Save to a read-only location or
- Move to a read-only location
- File Mover
  - An application used to move files from one data directory to another[3]
- File Monitor
  - An application used to monitor and provide audit trails on network data directories[3]
  - Indicates who saved, changed, moved or deleted a file
  - Audit trail information should be archived

- **Combining a file mover and a file monitor can close the gap on compliance of editable data files to a large degree!**

Data Storage

- Data files should be protected for the life of the file
- 10-20 YEARS



Data Backup Onsite

For onsite data backup, data should be backed up to a separate system and kept offsite, if possible.

Knowledge System

An in-house developed or third-party system used to monitor, version and store files on the network. These also can function as an electronic data repository.[3]

Data Backup Offsite

- Utilize a **sister site** for electronic data backup
- Sister site
  - A separate site within an organization that is not in the same location as the original site
- Utilizing a sister site allows for data backup to another region without utilizing third-party resources
- Connections between sites must be dedicated and encrypted at a minimum

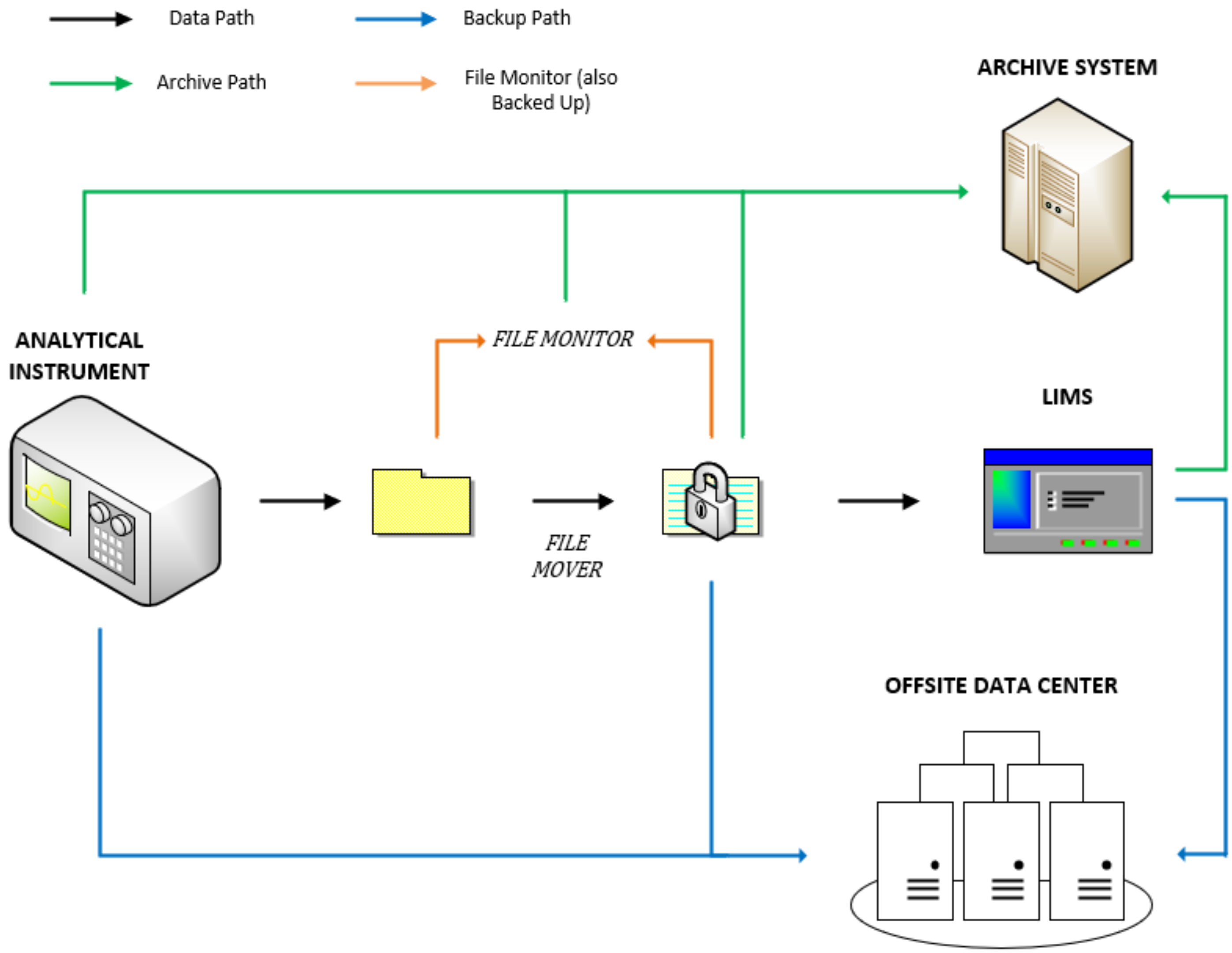
Disaster Recovery

- Perform disaster recovery (DR) testing at least every two years on **all critical applications**
- Prioritize restoration of critical applications
  - In what order will applications be restored?
  - All at once?
- Core business applications whose data is stored in your **knowledge system** can be covered by the DR testing for that one system, which should be defined as a critical application
- If possible, utilize a **sister site** within the organization to restore applications to

Example: Within the Organization

- Instrument
  - No compliance functions
  - Saves to any network location
  - Users log on using Windows account
  - Exports data to editable text (.txt) file
  - Text file is imported into an onsite laboratory information management system (LIMS)
- Data is archived to an onsite knowledge system
- Data is backed up to a sister site within the organization

Example: Within The Organization



Using Third-Party Resources

- Storage is cheap
- Upkeep is not
- IT employees spend a lot of time maintaining storage and backup processes

- But if you use third party resources, that means you will be working in…

The Cloud?

The Cloud!

The Cloud

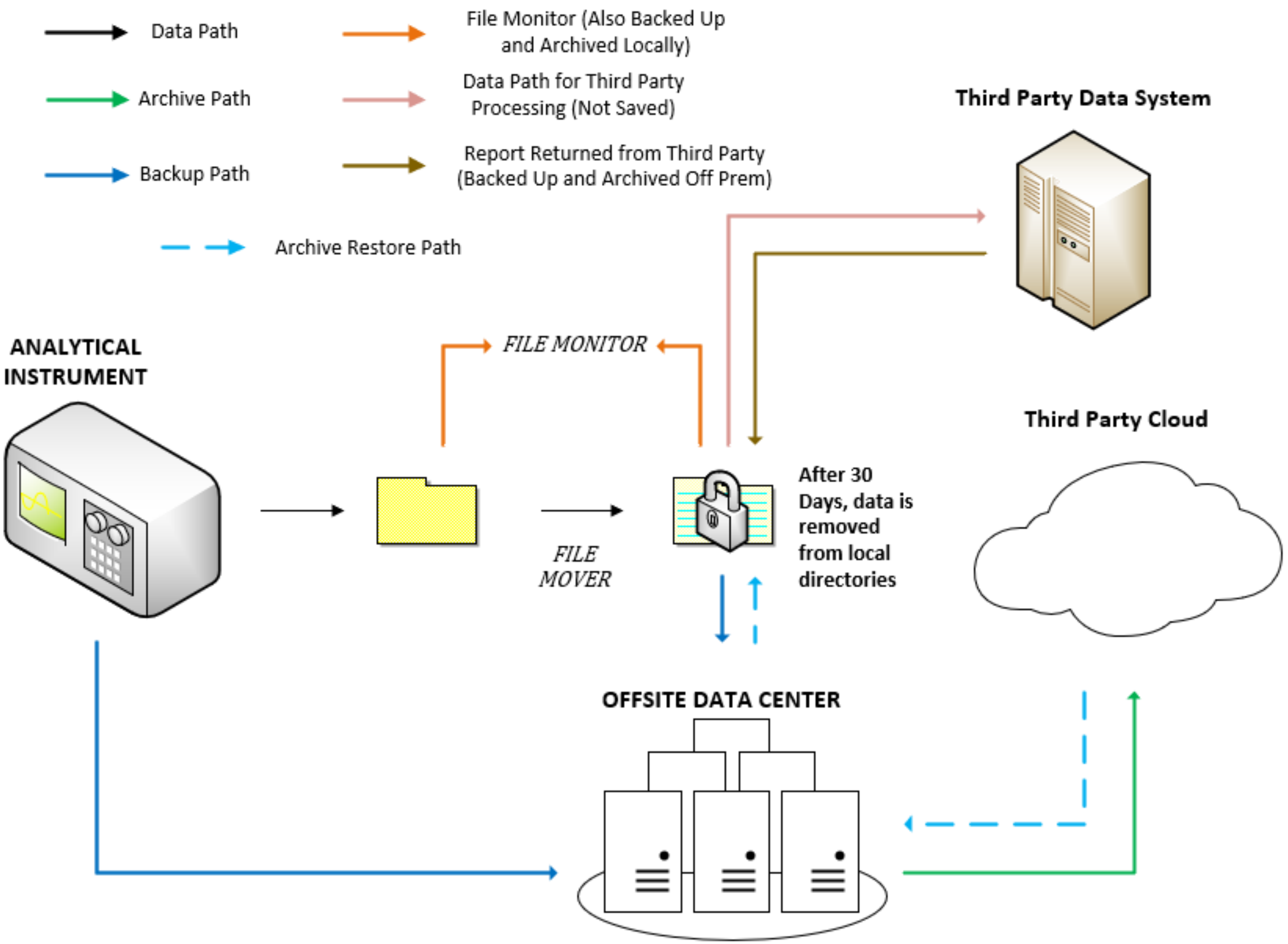
- **Cloud computing** is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction[4]

- A **cloud network** is a system where an organization keeps its network on third-party resources[3]

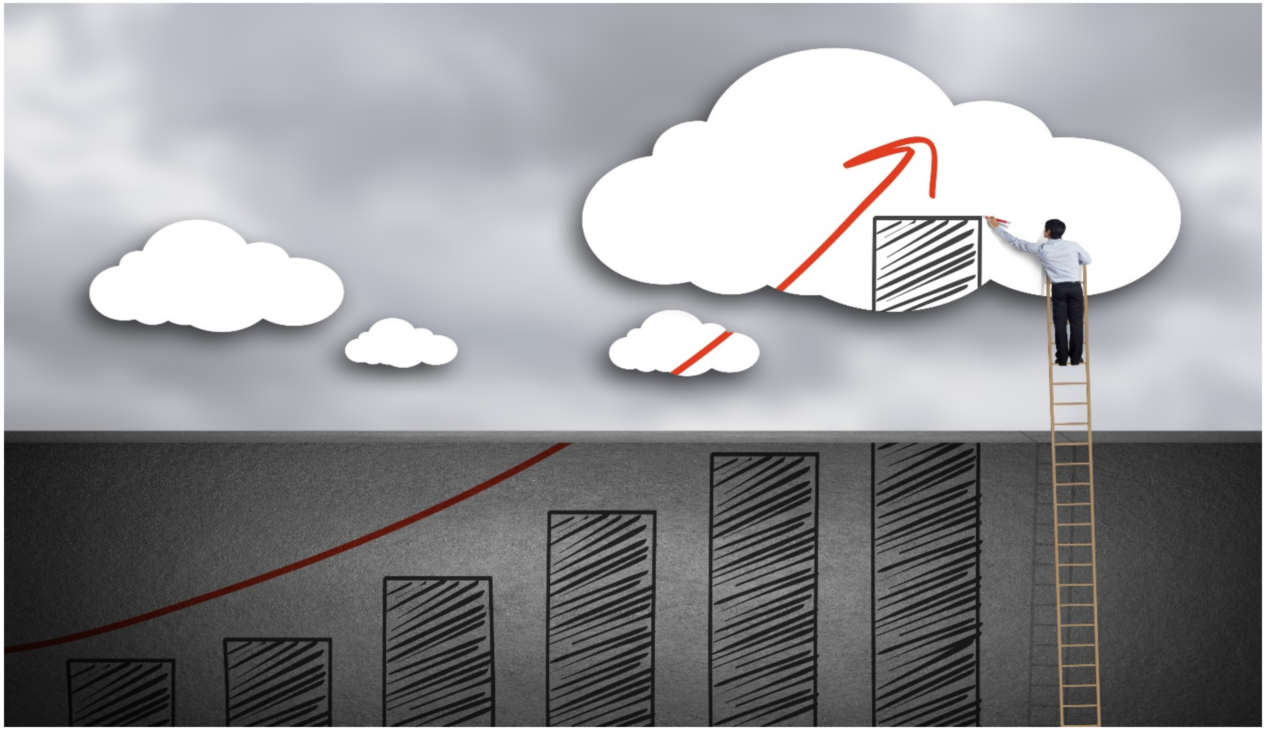
Example: Outside the Organization

- Instrument
  - No compliance functions
  - Saves to any network location
  - Users log on using Windows account
  - Exports data to editable text (.txt) file
  - Text file is imported into an offsite, **third-party** application for data processing
  - Reports are downloaded back to organization
- Data is backed up to a sister site within the organization
- Data is archived to a **third-party** data storage system

Example: Outside the Organization


Data Path
Archive Path
Backup Path
File Monitor (Also Backed Up and Archived Locally)
Data Path for Third Party Processing (Not Saved)
Report Returned from Third Party (Backed Up and Archived Off Prem)
Archive Restore Path
Third Party Data System
ANALYTICAL INSTRUMENT
FILE MONITOR
FILE MOVER
Third Party Cloud
After 30 Days, data is removed from local directories
OFFSITE DATA CENTER

Cloud Benefits



- Large Amounts of Storage
- Data backup and disaster recovery are the responsibility of the cloud provider
- Your IT employees can concentrate more on serving your users

Working With Cloud Providers

- Vendor audit
  - At least every other year
  - Compliance
  - Penetration testing history
  - Backup/disaster recovery
  - Escrow
- Confidentiality agreement
- Data segregation
  - Your data should be separated from other client data
- Secure connection
  - **Any information** between you and a third-party provider should be encrypted including email and any data transfer
- Co-location
  - An agreement where an organization has dedicated third-party resources

Cloud Compliance

- There are many different sets of guidance and regulations
  - Examples: FDA 21 CFR Part 11, CLIA, GLP, GCP, GMP, OECD, MHRA
- Many cloud providers offer compliance functions

- *Investigate compliance functions with any potential cloud provider to make sure they meet your needs*
- ***YOU** will most likely be responsible for setting these functions up and using them!*

Cloud Data Sharing

- The cloud makes sharing *easy*
- A central laboratory can share data with sister sites more fluidly
- An organization can share data with other organizations, such as with their clients or with auditors

Summary

- Large networks can be intimidating in terms of data protection, but they offer more efficient, more secure ways of handling data
- Data can be protected across multiple sites within a large organization, allowing for a multi-site data protection protocol to be used
- Using third-party resources (i.e. **the cloud**) is not as intimidating as it once was and is now a viable way to process, store, share and protect data without your organization absorbing the burden of these processes

AAPS Data Storage Working Group

Saad Abed
Jeb Adams
Bargav Bhesaniya
Phyllis Conliffe
Sean Crawford
Michelle Dawes
John Evens
Boris Gorovits
Geoffrey Grove
Hannes Hochreiner
Kimberly Honrine
John Kellie
Stephen MacMannis
Heather Myler
Samuel Pine
Nanda Subbarao
Phillip Sundman
Elizabeth Tran
Teruyo Uenoyama
Joel Usansky
Dominic Warrino
Joleen White
Eric Woolf

Join Us!

Questions?

THANK YOU

References

- 1. Merriam-Webster Online, https://www.merriam-webster.com/dictionary/data, accessed 03 October 2019

- 2. *'GXP' Data Integrity Guidance and Definitions*, Medicines & Healthcare products Regulatory Agency (MHRA), March 2018

- 3. Davis, *Closing the Gap on Data Integrity,* The Journal of GXP Compliance, September 2018, Volume 22, Issue 5

- 4. Mell, Peter; Grance, Timothy; NIST Special Publication 800-145, The NIST Definition of Cloud Computing, September 2011