Storage System
High-Availability & Disaster Recovery Overview
[638]

John Wolfgang
Enterprise Storage Architecture & Services

jwolfgang@vicominfinity.com
26 June 2014

Agenda

- **Background**
- Application-Based Replication
- Storage-Based Replication
  - Tape Replication
  - Point-in-Time Replication
  - Synchronous Replication
  - Asynchronous Replication
- Automation
- Replication Examples
- Key Questions for Any Solution

My Background

- **West Virginia University**
  - BS Electrical Engineering
  - BS Computer Engineering
- **Carnegie Mellon University**
  - MS Electrical and Computer Engineering
  - Data Storage Systems Center
- **Lockheed Martin & Raytheon**
- **IBM**
  - Development (Tucson) - Software Engineer – 10 years
  - Data replication, disaster recovery, GMU, eRCMF, TPC for Replication
  - Global Support Manager (New York City) – Morgan Stanley – 2 years
  - IBM Master Inventor
- **Vicom Infinity**
  - Enterprise Storage Architecture and Services – 1+ years

Why are High Availability & Disaster Recovery Important?

- Information is your most important commodity – need to protect it
- What happens to your company if you don't have access to your production data for a minute? An hour? A day? A month?
- How much money does your company lose every minute?
  - Amazon.com loses $66,240 per minute (Forbes.com – 8/19/2013)
  - Ebay.com loses $120,000 per minute (ebay.com)

Lessons Learned from Previous Disasters

- Rolling disasters happen
- Distance is more important
- Redundancy may be smoke and mirrors
- If you have not successfully tested your exact DR plan, you do not have a DR plan
- Automate as much as possible
  - Increase dependency on automation and decrease dependency on people
  - Automation provides the ability to test over and over until perfect
  - Automation will not deviate from procedures
  - Automation will NOT make mistakes (even under pressure!)
  - Automation will not have trouble getting to the DR site
- Recovery site Considerations
  - Site capacity (MIPs and TBs) needs to be sized to handle the production environment
  - What is the DR Plan after successful recovery from disaster
  - Disasters may cause multiple companies to recover and that puts stress on the commercial business recovery services

Replication Beyond Disaster Recovery

Disaster Recovery/ Business Continuity

Minimize data loss
Minimize restart time
Increase distance
Enable automation

Availability Improvements

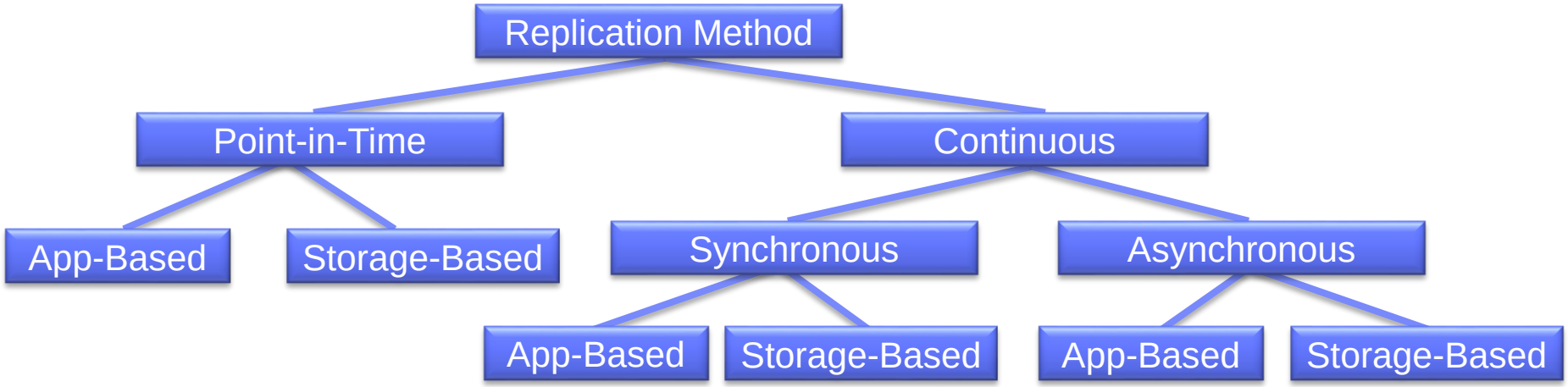
Backup Window
Tape Backup
Data Migration
Archival

Operational Efficiency

Data Mining
Content Distribution
Software Testing

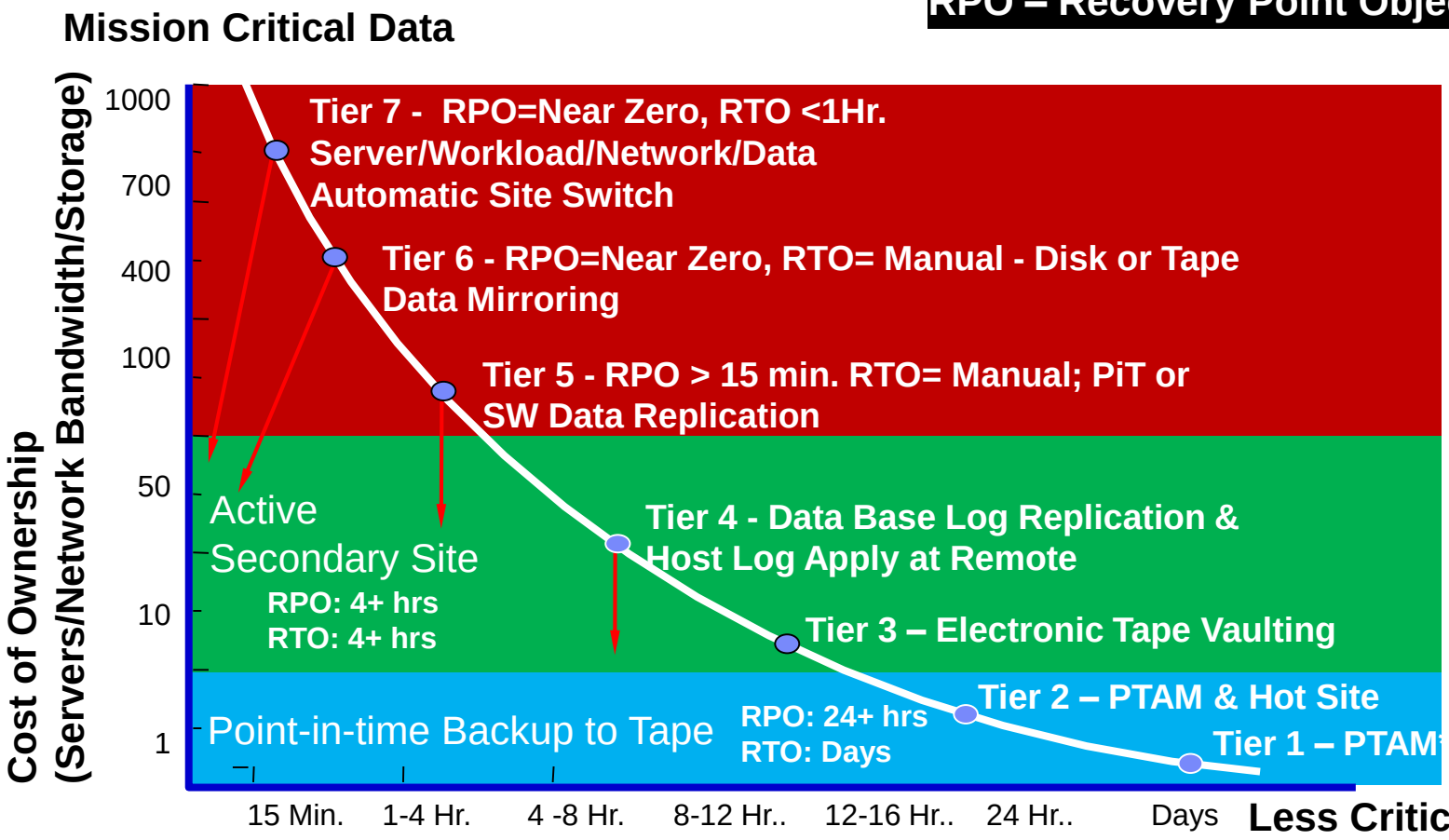
Some Definitions

- Recovery Point Objective (RPO)
  - How much data can you tolerate losing during a disaster
- Recovery Time Objective (RTO)
  - How much time will it take to get your systems up and running again after a disaster



7 Tiers of Business Recovery Options

Key Customer Objectives:
RTO – Recovery Time Objective
RPO – Recovery Point Objective



Time to Recover – How quickly is an application recovered after a disaster?

*PTAM – Pickup Truck Access Method

Agenda

- Background
- **Application-Based Replication**
- Storage-Based Replication
  - Tape Replication
  - Point-in-Time Replication
  - Synchronous Replication
  - Asynchronous Replication
- Automation
- Replication Examples
- Key Questions for Any Solution

Application-Based vs. Storage-Based Replication (1)

Application/File/Transaction Based

- Specific to application/file system/database
- Generally less data is transferred
  - Lower telecommunication costs
- No coordination across applications, FSs, DBs, etc.
- Applications change - replication may need to change
- May forget “other" related data necessary for recovery
- With many transfers occurring in a corporation, it may be difficult to determine what is where in a disaster. RTO/RPO may not be repeatable, auditing may be difficult
- Many targets possible (ex. millions of cell phones)

Application-Based Replication Examples

DB2 HADR

- High availability solution for both partial and complete site failures
- Log data is shipped and applied to standby database
  - One or more standby databases
- If primary database fails, applications are redirected to the standby database
- Standby database takes over in seconds
  - Avoids database restart upon a partial error

LVM Mirroring

- Create more than one copy of a physical partition to increase data availability
- Handled at the logical volume level
- If a disk fails, can still have access the data on an alternate disk
- Remote LVM mirroring enables use of disks located at multiple locations
  - Replication between multiple storage systems via a Storage Area Network (SAN)

Agenda

- Background
- Application-Based Replication
- **Storage-Based Replication**
  - **Tape Replication**
  - **Point-in-Time Replication**
  - **Synchronous Replication**
  - **Asynchronous Replication**
- Automation
- Replication Examples
- Key Questions for Any Solution

Application-Based vs. Storage-Based Replication (2)

Storage-Based – Block Level Replication

- Independent of application, file systems, databases, etc.
- Common technique for corporation
  - Managed by operations
- Generally more data transferred
  - Higher telecommunication costs
- Consistency groups yield cross volume/storage subsystem data integrity/consistency
- Independent of application changes.
  - Mirror all pools of storage
- Consistent repeatable RPO.
- RTO depends on server/data/workload/network
- Generally a handful of targets
- Specific to data replication technique (tied to specific architecture & devices that support it)

What Does Data Consistency Really Mean?

- For storage-based replication, we are talking about “power fail” consistency
- Typical Database transaction:
  1. Update log – database update is about to occur
  2. Update database
  3. Update log - database update complete

- Host is very careful to do each of the transactions in order
  - This provides power fail data consistency
- BUT, these transactions are likely done to different volumes possibly on different control units
- Failure to be careful about transaction order results in loss of data consistency and data may become unusable
- In order to ensure data consistency at secondary site, dependent writes must be done in order
- How does a storage system know which writes are dependent?
  - It doesn’t
  - What it does know is that writes that are done in parallel are not dependent
  - Any writes NOT done in parallel are assumed to be dependent
- This is exacerbated for asynchronous replication

Storage-Based Replication Techniques

Tape

- Pickup Truck Access Method (PTAM)
- Virtual Tape Replication

Disk

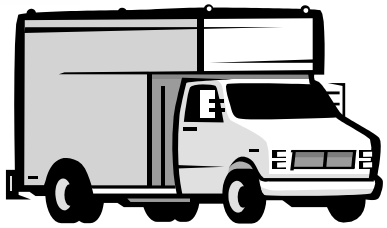
- Point-in-Time Copy
- Synchronous Replication
- Asynchronous Replication
- Three-site Replication (Synchronous & Asynchronous)

Automation

- Hyperswap
- Tivoli Storage Productivity Center for Replication
- Globally Dispersed Parallel Sysplex (GDPS)

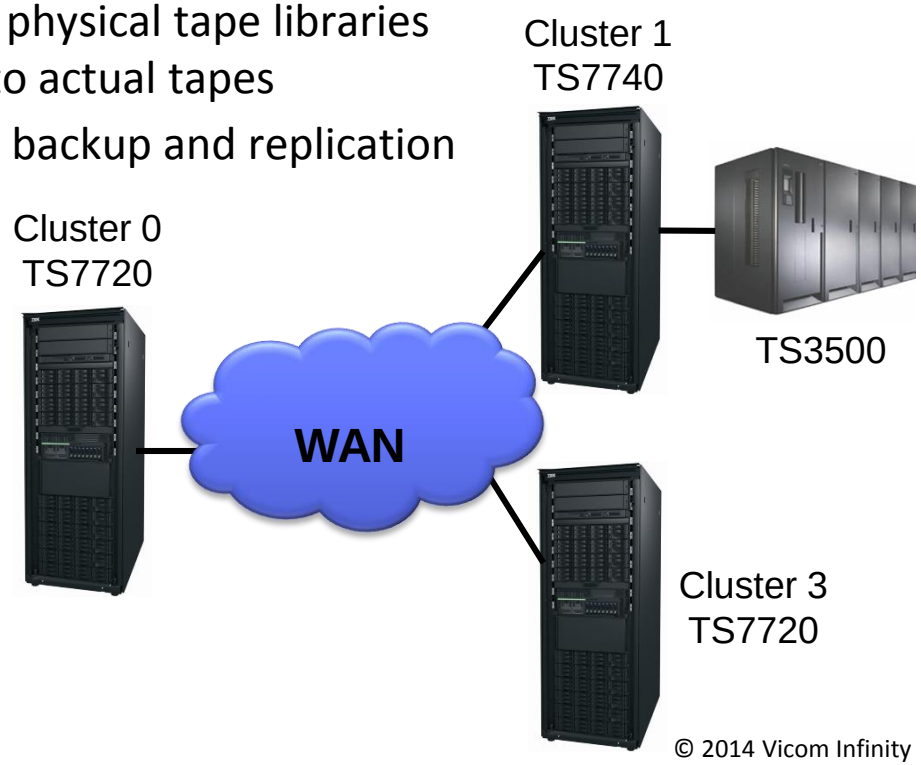
Tape Replication - PTAM

- Backups created and dumped to physical tapes
  - Recovery Point Objectives quite high – 24 hours at best?
- Tapes are literally picked up by a truck and taken to another location
  - Hot site
  - Storage only
  - Recovery Time Objective fairly high in both cases
- Lower cost and simpler option than disk replication



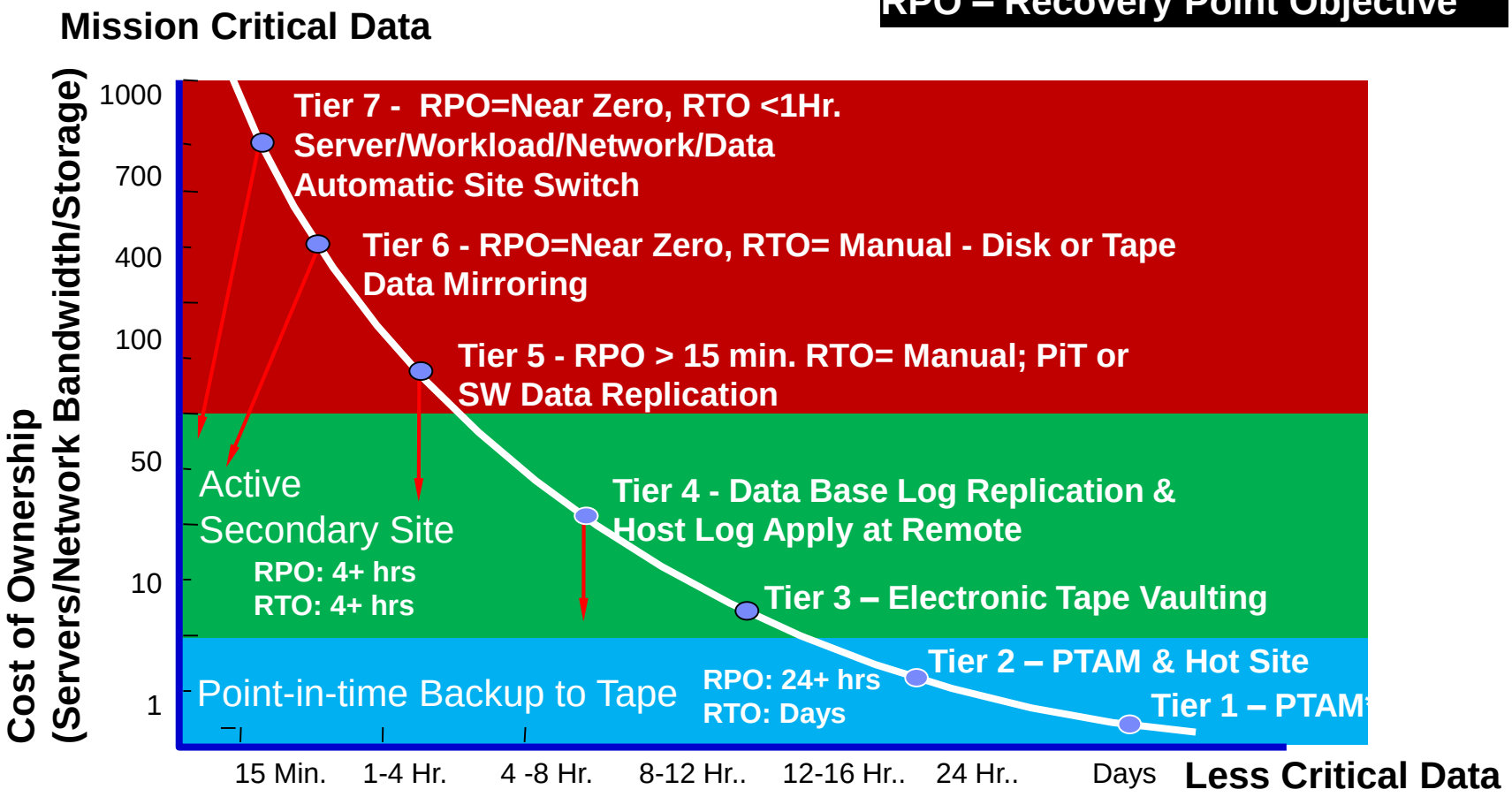
Tape Replication – Virtual Tape Grid

- Virtual Tape Servers appear to hosts as standard tape volumes
  - May or may not actually contain tape drives and tapes
- Multiple clusters can be put together into a tape grid
- Tape volumes can be selectively replicated to one or more other clusters
- Tape volumes can be accessed through any cluster in the grid
  - Whether or not the tape volume physically resides on that cluster
- Certain virtual tape server models have physical tape libraries behind them that can offload volumes to actual tapes
- Hybrid with characteristics of both tape backup and replication
  - Recovery Point Objective much better than PTAM



7 Tiers of Business Recovery Options

Key Customer Objectives:
RTO – Recovery Time Objective
RPO – Recovery Point Objective



Time to Recover – How quickly is an application recovered after a disaster?

*PTAM – Pickup Truck Access Method

Point-in-Time vs. Continuous Replication

Point-in-Time

Local copy of data
Data "Frozen"

- Provides protection against logical corruption, user error
- Data is not the most current

Continuous Replication

Remote copy of the data

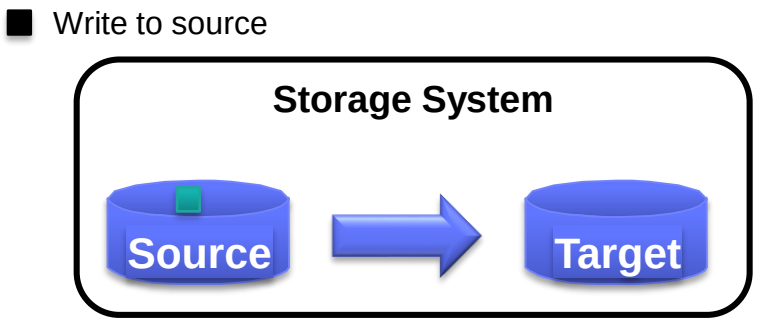
- Provides protection against primary storage system or data center issue

Continuously updated

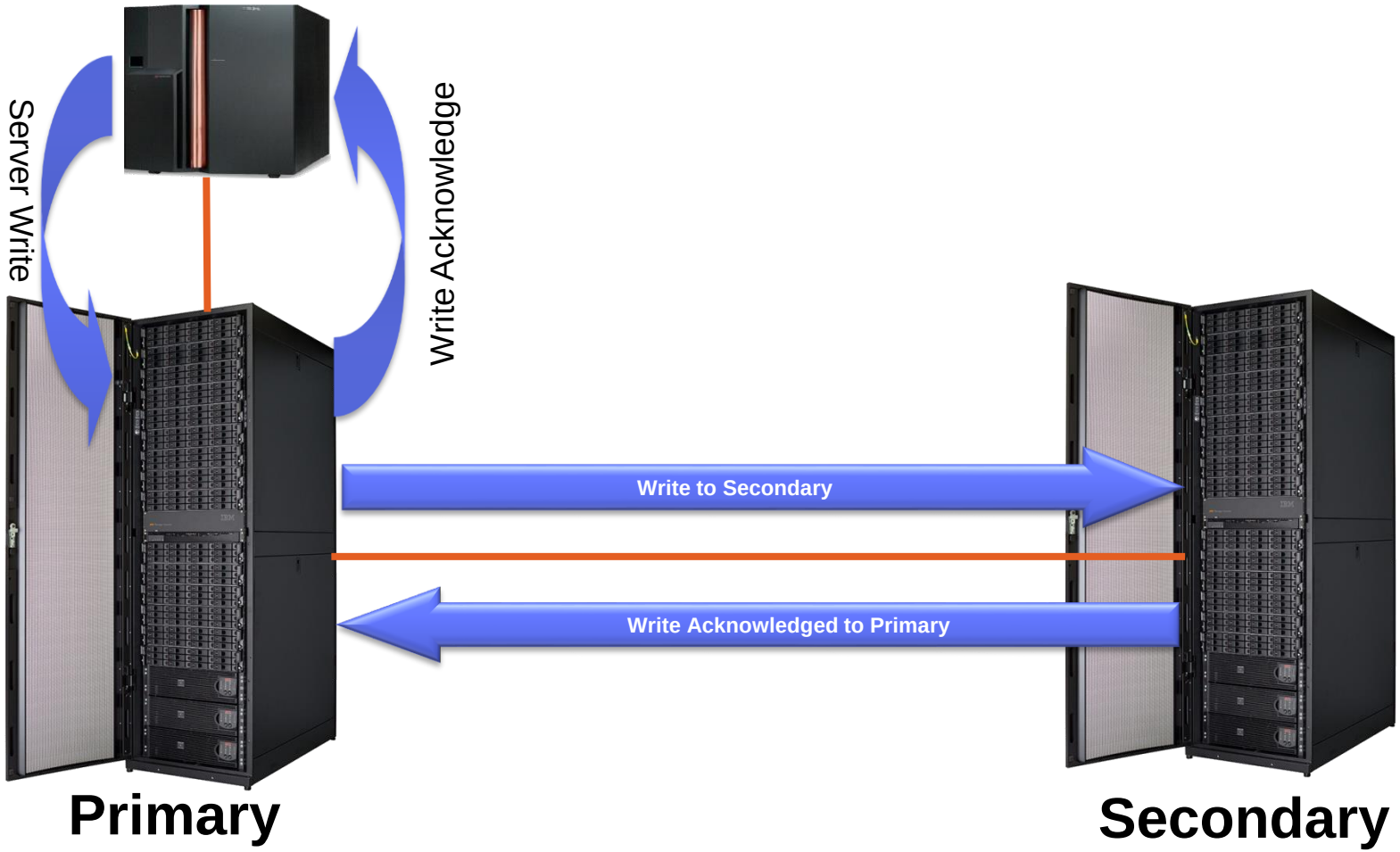
- Data is always current (or close to it)
- Corruption/Errors on the primary site will be transferred to the secondary

Point-in-Time Copy

- Internal to Storage System
- New copy created and available immediately
- Possible to read & write to both volumes
- No-Copy
  - No data is copied to Target unless updated on the Source
- Copy on Write
  - Data must be copied to Target before being updated on Source
- Background Copy
  - All data from Source copied to Target
  - Relationship typically ends when copy is complete
- Incremental Copy
  - Full background copy is done the first time
  - Only changes copied subsequently
- Space Efficient/Thin Provisioned
  - Only allocate space as it is used

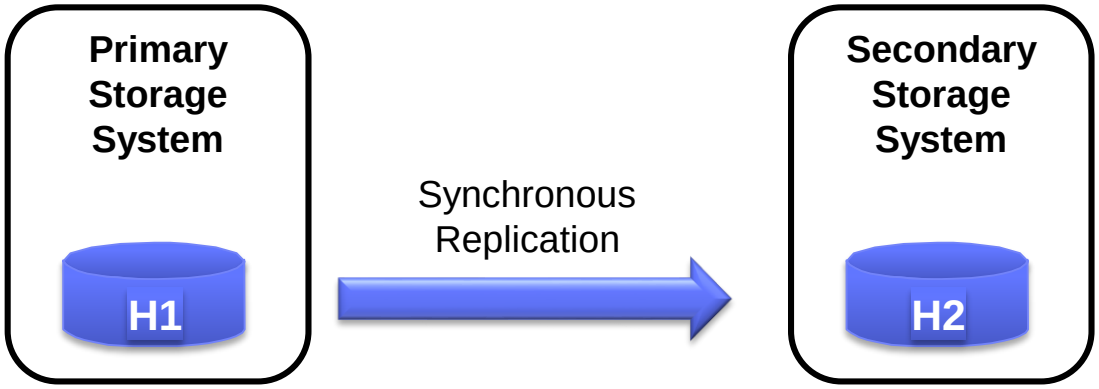


Synchronous Replication Overview



Synchronous Replication

- Data on secondary storage system is always identical to primary
  - Recovery Point Objective of 0
- Standard implementation for many storage vendors
- There is an impact on application I/Os
  - Dependent on distance between primary and secondary
  - Distance to 300 km
  - Bandwidth must be sufficient for peak
- Data Freeze technology keeps all pairs in consistency group consistent
  - Requires automation to guarantee consistency across multiple storage systems

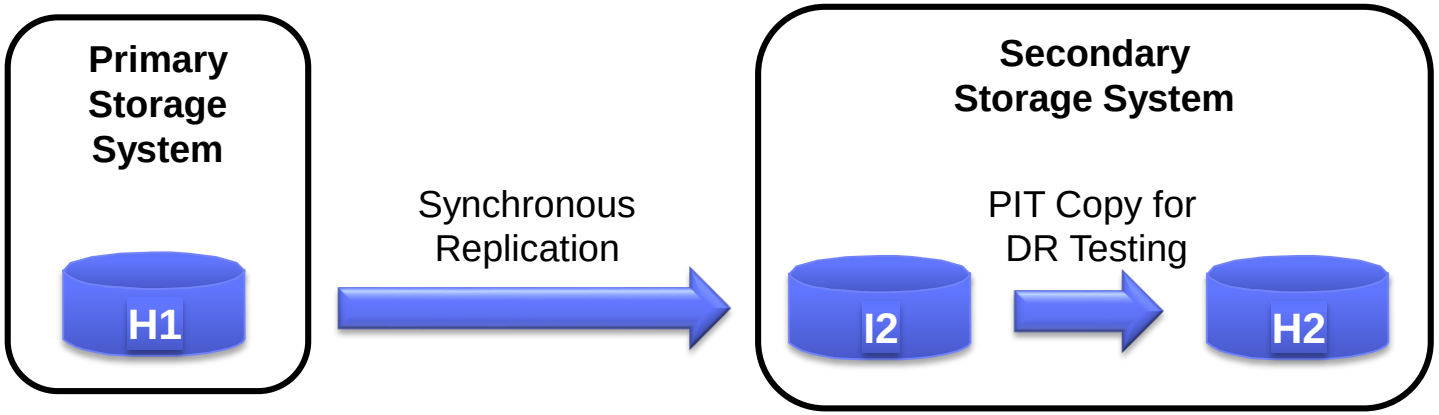


Practice How you Recover and Recover How you Practice

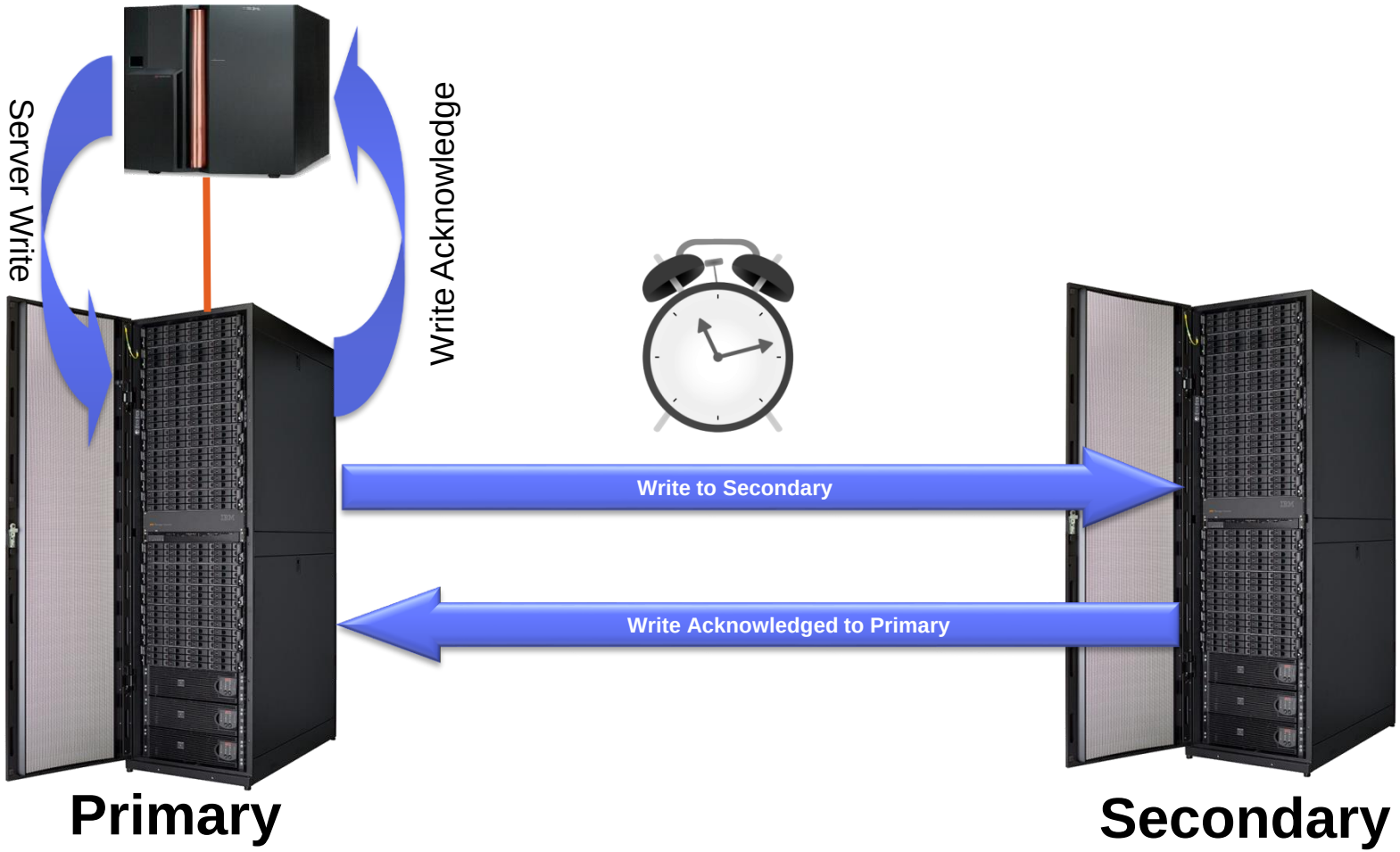
- Proper Disaster Recovery Tests require time & effort & commitment
- If you haven't successfully tested your exact DR plan, you don't have a DR plan
- A DR test may require you to stop data replication temporarily
- Use Practice Volumes to test properly while continuing replication
- Practice Volumes can also be used for other activities
  - Development, testing, data analytics
- Make sure you always recover to the Practice Volumes – even in a real disaster

Synchronous Replication with Practice Volumes

- Standard synchronous replication as the basis
- Typical synchronous replication requires replication outage for DR testing
- Practice volumes provide capability to continue replication during DR testing
- Data is recovered to secondary storage system
- Point-in-Time copy created on secondary storage system
- Replication is restarted while access to H2 volume still available
- Should recover in actual disaster using the same method

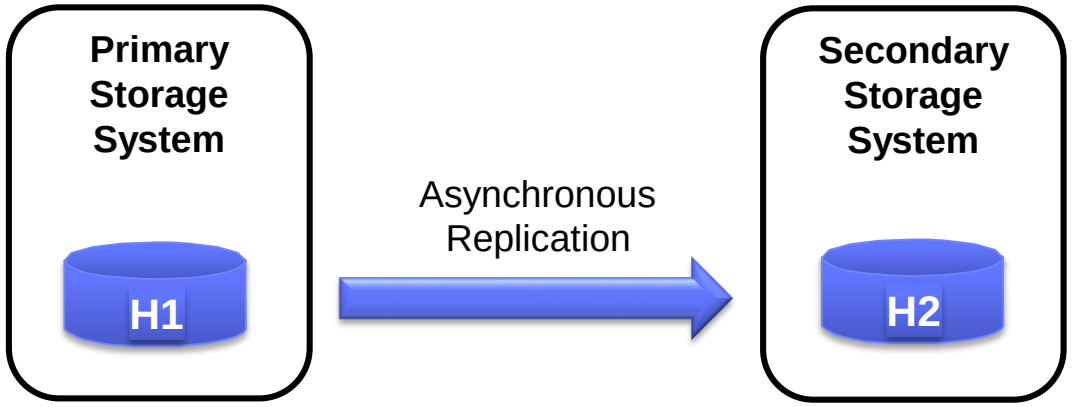


Asynchronous Replication Overview



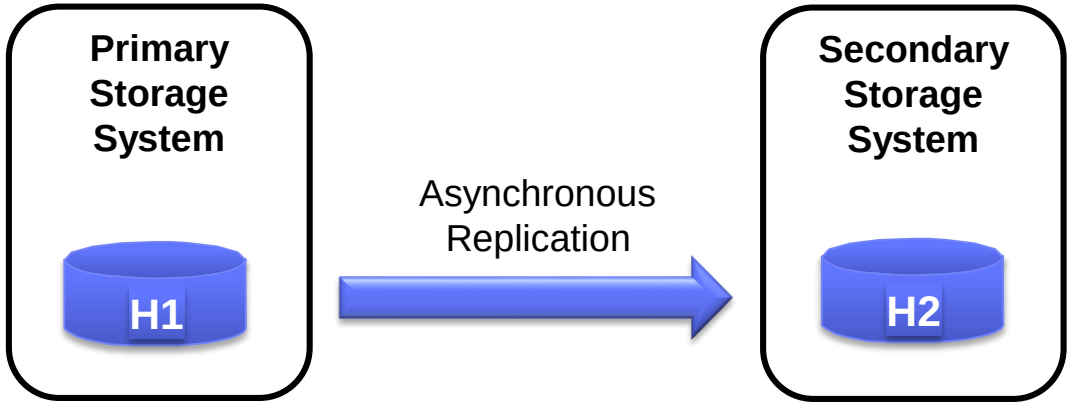
Asynchronous Replication – No Consistency

- Asynchronous transfer of data updates
- No distance limitation
- Little impact on application I/Os
- Secondary not guaranteed consistent
  - No write ordering
  - No consistent data sets
  - Hosts/Applications must be shut down to provide consistency
- Most useful for migration
- Can transition to/from Synchronous replication



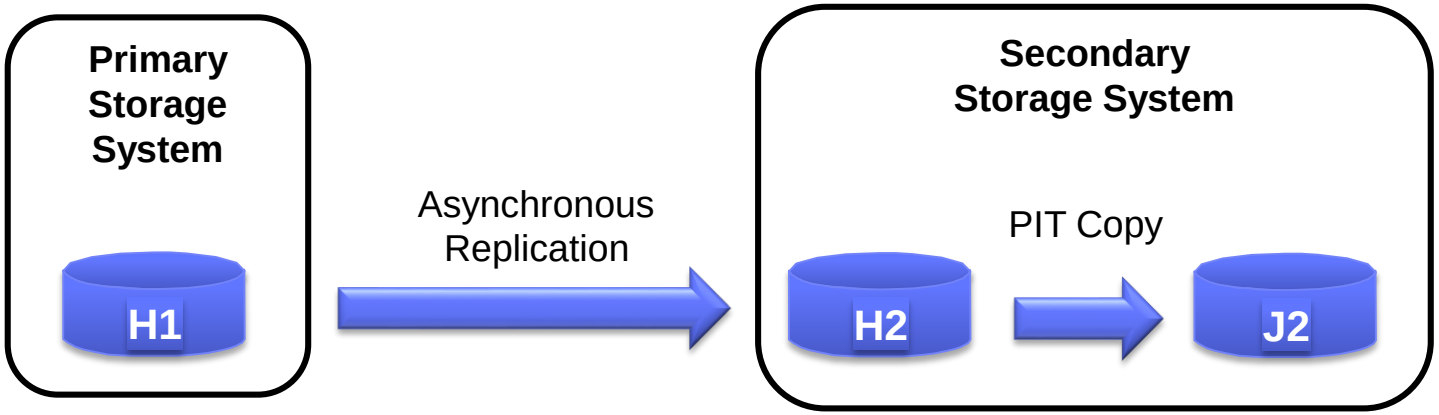
Asynchronous Replication – Two Volumes

- Asynchronous transfer of data updates
  - Recovery Point Objective > 0
- No distance limitation
- Little impact on application I/Os
- Data consistency maintained via:
  - Write ordering
  - Consistent data sets
- If bandwidth is not sufficient for peak, data will back up on the primary
  - Some vendors require extra cache



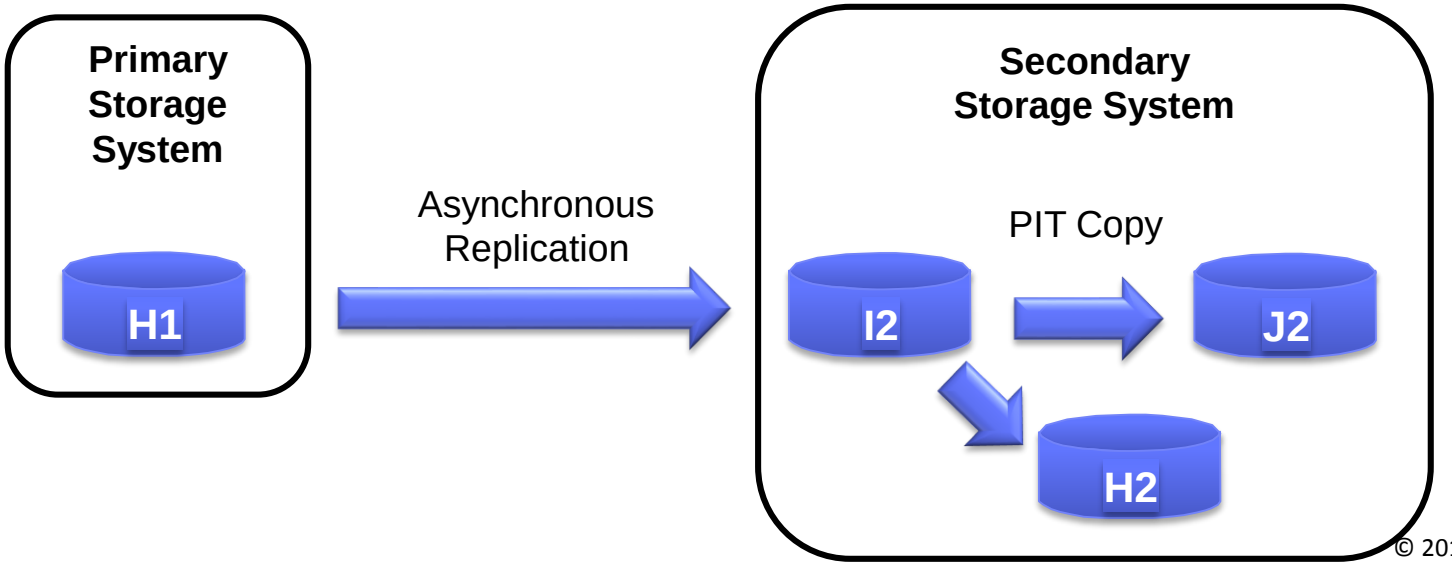
Asynchronous Replication – Three Volumes

- Asynchronous transfer of data updates
  - Recovery Point Objective > 0
- No distance limitation
- Little impact on application I/Os
- Data consistency created using 3rd volume
- Consistency coordinated by primary storage system
- If bandwidth is not sufficient for peak, RPO will grow and “catch up” later



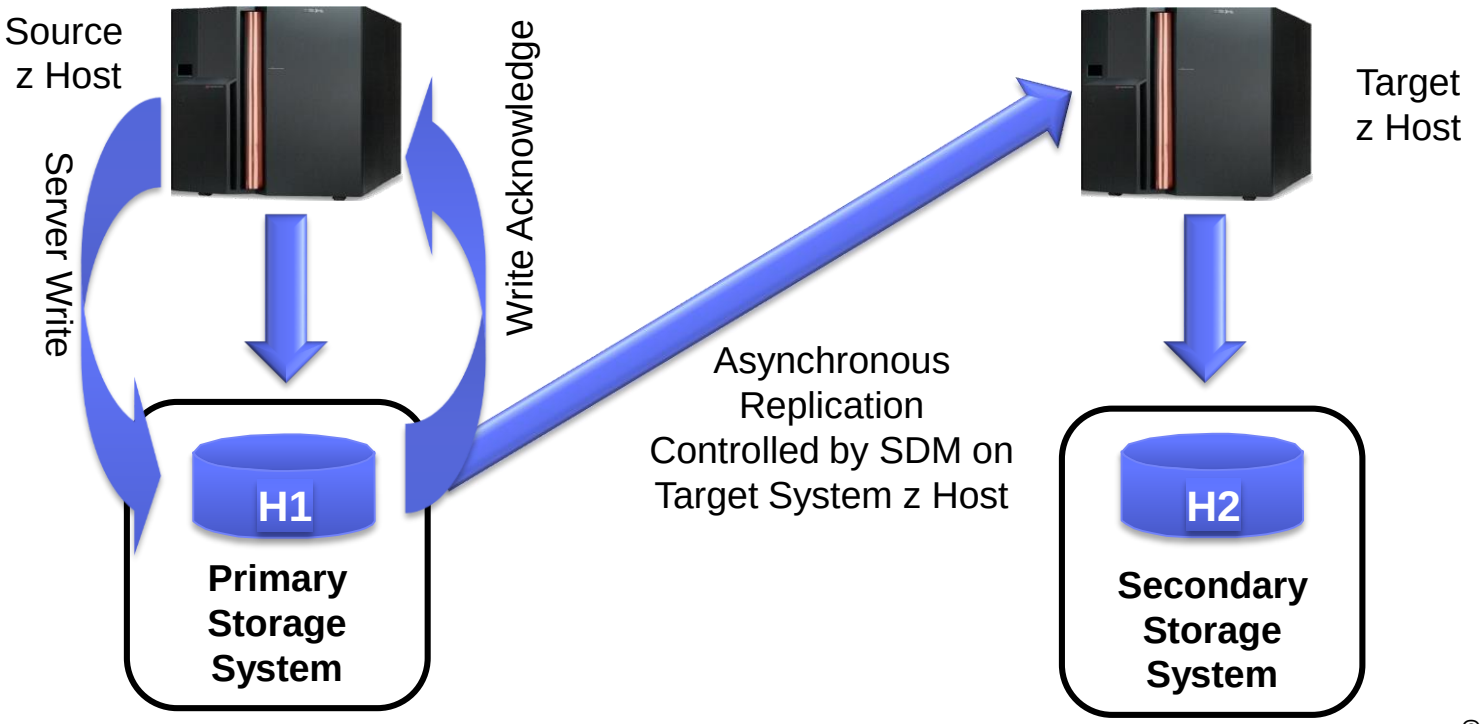
Asynchronous Replication With Practice Volumes

- Standard asynchronous replication as the basis
  - Could be any of the consistent variants
- Typical asynchronous replication requires replication outage for DR testing
- Practice volumes provide capability to continue replication during DR testing
- Data is recovered to secondary storage system in typical manner
- Point-in-Time copy created on secondary storage system
- Replication is restarted while access to H2 volume still available
- Should recover in actual disaster using the same method



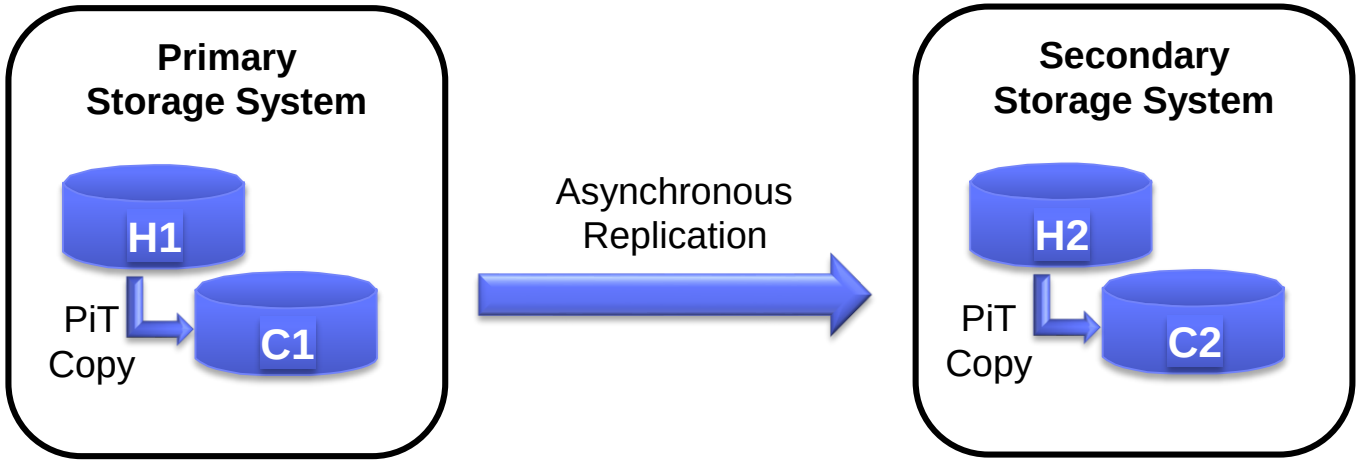
Asynchronous Replication – z/OS Interaction

- Asynchronous transfer of data updates
  - Recovery Point Objective > 0 but very low (~seconds)
- No distance limitation
- Little impact on application I/Os
- Managed by System z
- Multiple Storage vendors



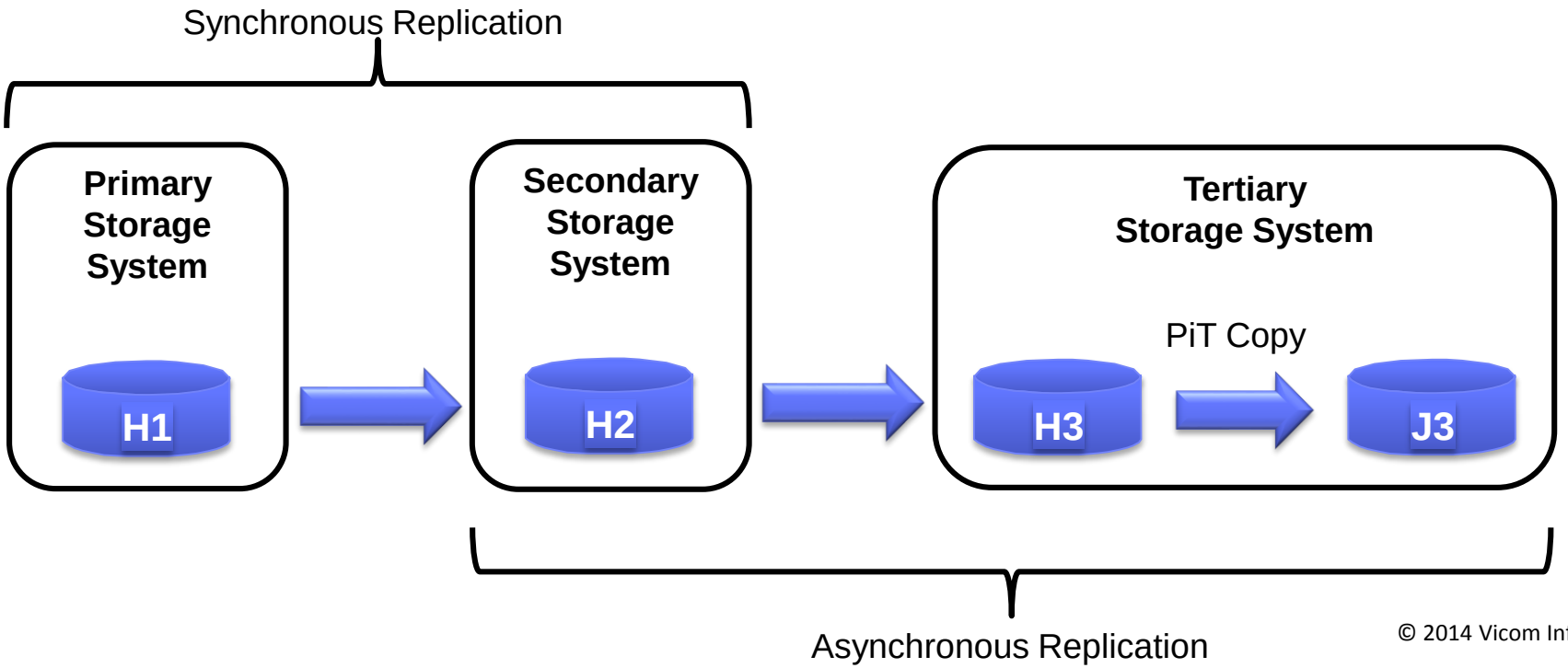
Asynchronous Replication – Four Volumes/PiT Copies

- Asynchronous transfer of data updates
  - Recovery Point Objective typically higher than previously discussed implementations
  - RPO is 2x the “cycling period”
  - Can tolerate lower network bandwidth
- Little impact on application I/Os
- Periodic consistent PiT copies are created from primary volumes
- PiT copies are replicated to secondary volumes
  - Does not require a consistent replication mechanism
- After copy is complete, PiT copies are created from secondary volumes for protection



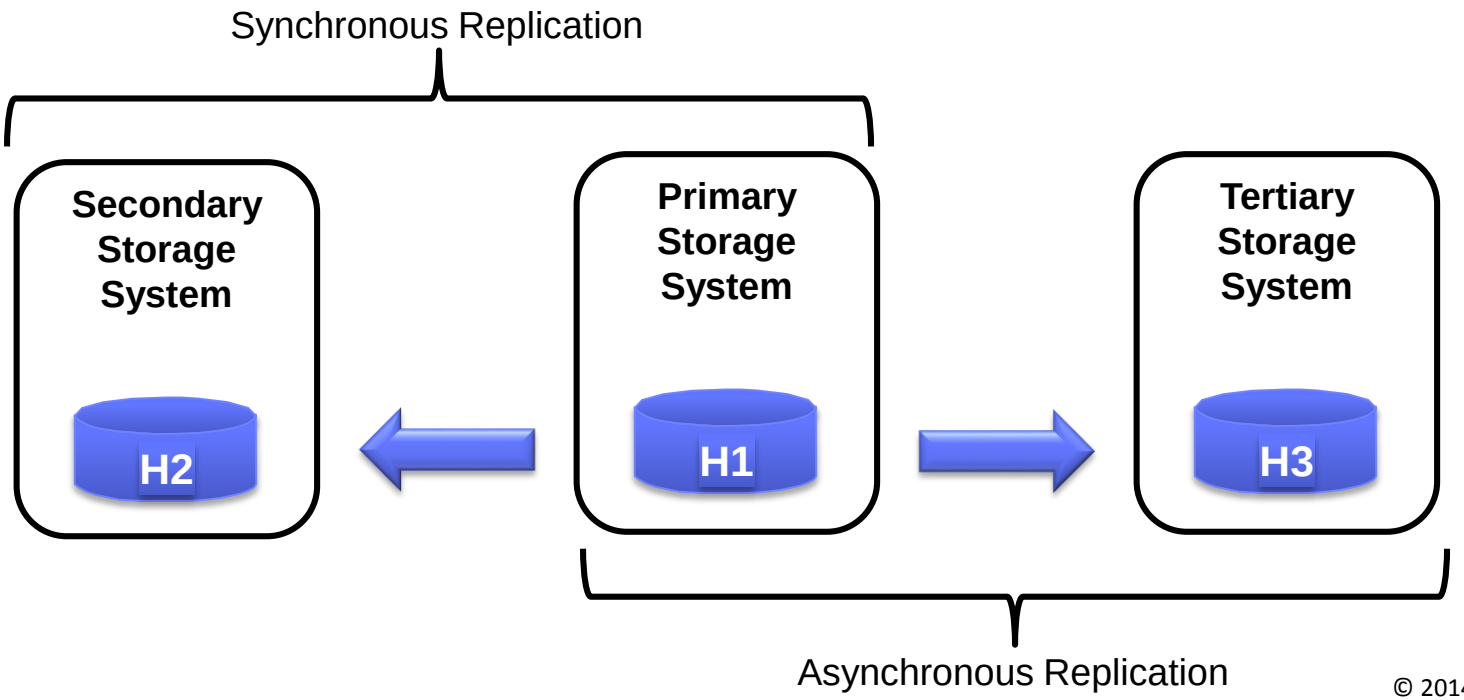
Three-Site Replication - Cascading

- Combination of Synchronous & Asynchronous replication techniques
- Synchronous replication to provide High Availability at metro distances
  - Protect against storage system & data center disasters
- Asynchronous replication to provide disaster recovery capability at global distances
  - Protect against regional disasters
- Ability to switch production between primary and secondary systems
- Incremental resynchronization between primary and tertiary if secondary lost
- Requires automation to handle the various transitions



Three-Site Replication – Multi-Target

- Combination of Synchronous & Asynchronous replication techniques
- Synchronous replication to provide High Availability at metro distances
  - Protect against storage system & data center disasters
- Asynchronous replication to provide disaster recovery capability at global distances
  - Protect against regional disasters

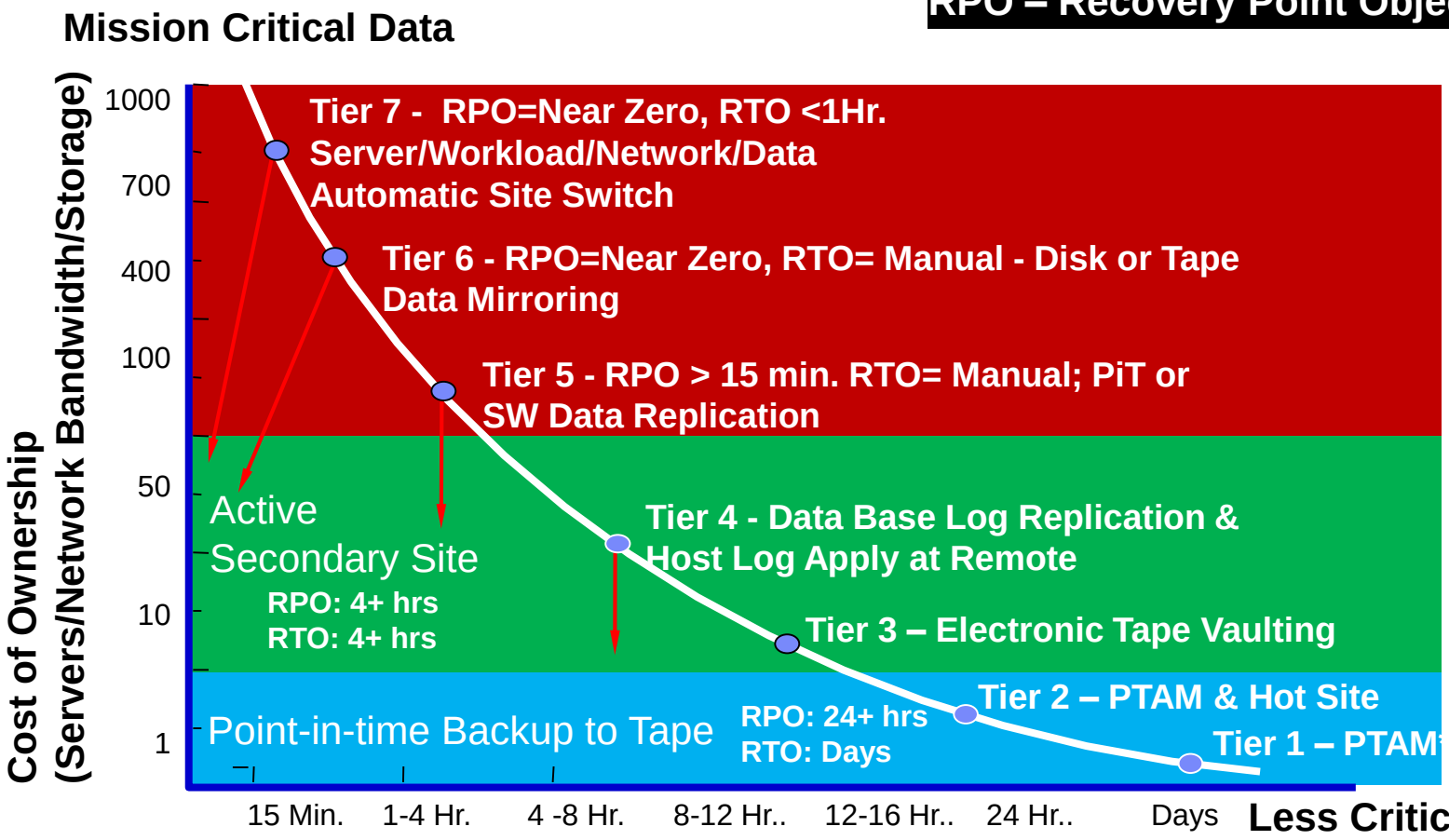


Agenda

- Background
- Application-Based Replication
- Storage-Based Replication
  - Tape Replication
  - Point-in-Time Replication
  - Synchronous Replication
  - Asynchronous Replication
- **Automation**
- Replication Examples
- Key Questions for Any Solution

7 Tiers of Business Recovery Options

Key Customer Objectives:
RTO – Recovery Time Objective
RPO – Recovery Point Objective



Time to Recover – How quickly is an application recovered after a disaster?

*PTAM – Pickup Truck Access Method

High Availability

Hyperswap for Synchronous Replication Configurations

- Triggered when there is a problem writing or accessing the primary storage devices
- Swap from using primary storage devices to secondary storage devices
- Transparent to applications (brief pause on the order of seconds)
- Steps
  - Physically switch the secondary storage devices to be primary and allow access
  - Logically switch the OS internal pointers in the UCBs
  - Applications are not aware that they are now using the secondary devices
- No Shutdown, No Configuration Changes, No IPL
- Managed by Automation software (GDPS, TPC for Replication)
- Planned Hyperswap
  - Use for maintenance, production site move, migration
- Unplanned Hyperswap
  - Automated to protect against storage system failure

Tivoli Storage Productivity Center for Replication

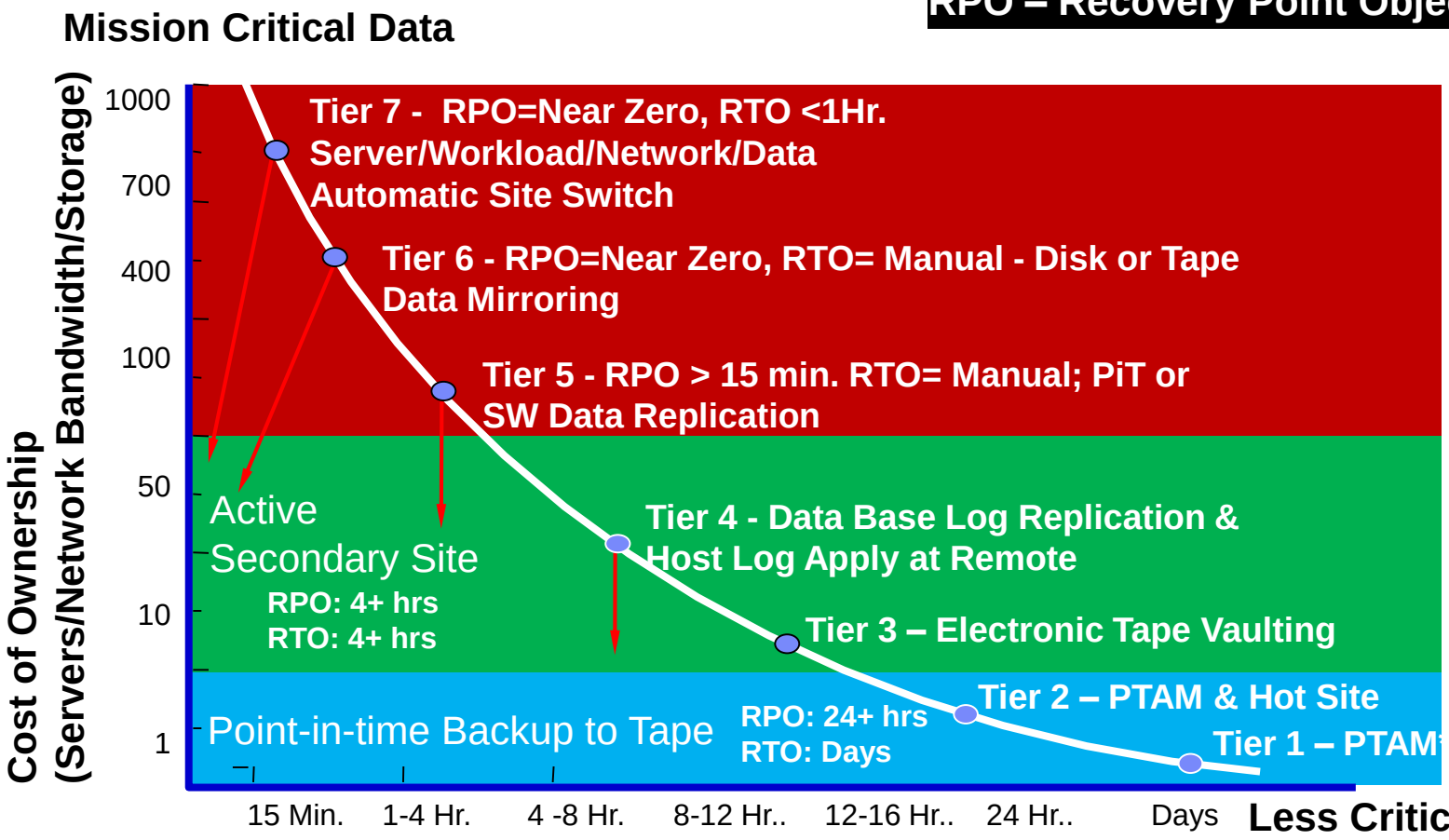
- Automate and simplify complex data replication tasks
- Control multiple replication types and storage systems from a single pane
  - Including CKD and FB volumes
- Added Error Protection
- Added Ease of Use
- Facilitates DR Testing and DR Recovery
- Enables Basic Hyperswap, Hyperswap, Open Hyperswap
- GUI-based
  - Operational control of replication environment via a GUI rather than DSCLI scripts or TSO commands
    - Also provides a CLI
- Linux, Windows, AIX, z/OS

Globally Dispersed Parallel Sysplex (GDPS)

- Enables Business Recovery Tier 7 capability
- Manage all forms of replication
- Manage Hyperswap
- Drive down RTO through automation
- Scripting Capability provides ability to automate the recovery process at the DR site
  - Enable CBU
  - Automate Recovery of Disk systems
  - Automate IPL of LPARs
  - Automate application startup

7 Tiers of Business Recovery Options

Key Customer Objectives:
RTO – Recovery Time Objective
RPO – Recovery Point Objective



Time to Recover – How quickly is an application recovered after a disaster?

*PTAM – Pickup Truck Access Method

Data Replication Considerations

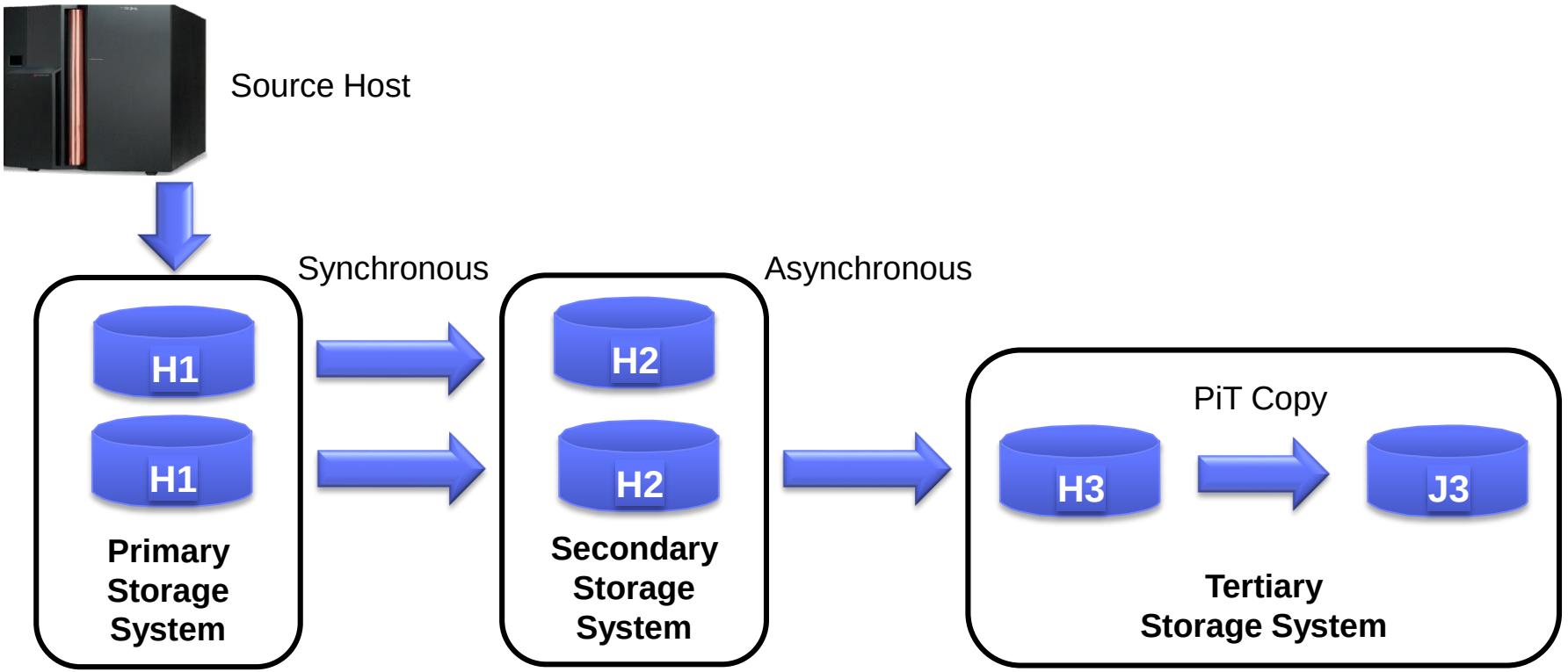
- Synchronous solutions do not work at distance
- Asynchronous solutions have data loss and potential problems managing consistency, particularly across different storage platforms
- Maximizing use of long distance link is critical for many customers
  - Smaller customers may want to purchase extended links which meet maximum transfer requirements for a shift, not their 15 second peak
- Being able to test, recover data at the recovery site, and replicate back to the production site after resolution is critical
  - If you have not successfully tested your DR procedures, you do NOT have DR procedures
  - Practice how you recover, and recover how you practice

Agenda

- Background
- Application-Based Replication
- Storage-Based Replication
  - Tape Replication
  - Point-in-Time Replication
  - Synchronous Replication
  - Asynchronous Replication
- Automation
- **Replication Examples**
- Key Questions for Any Solution

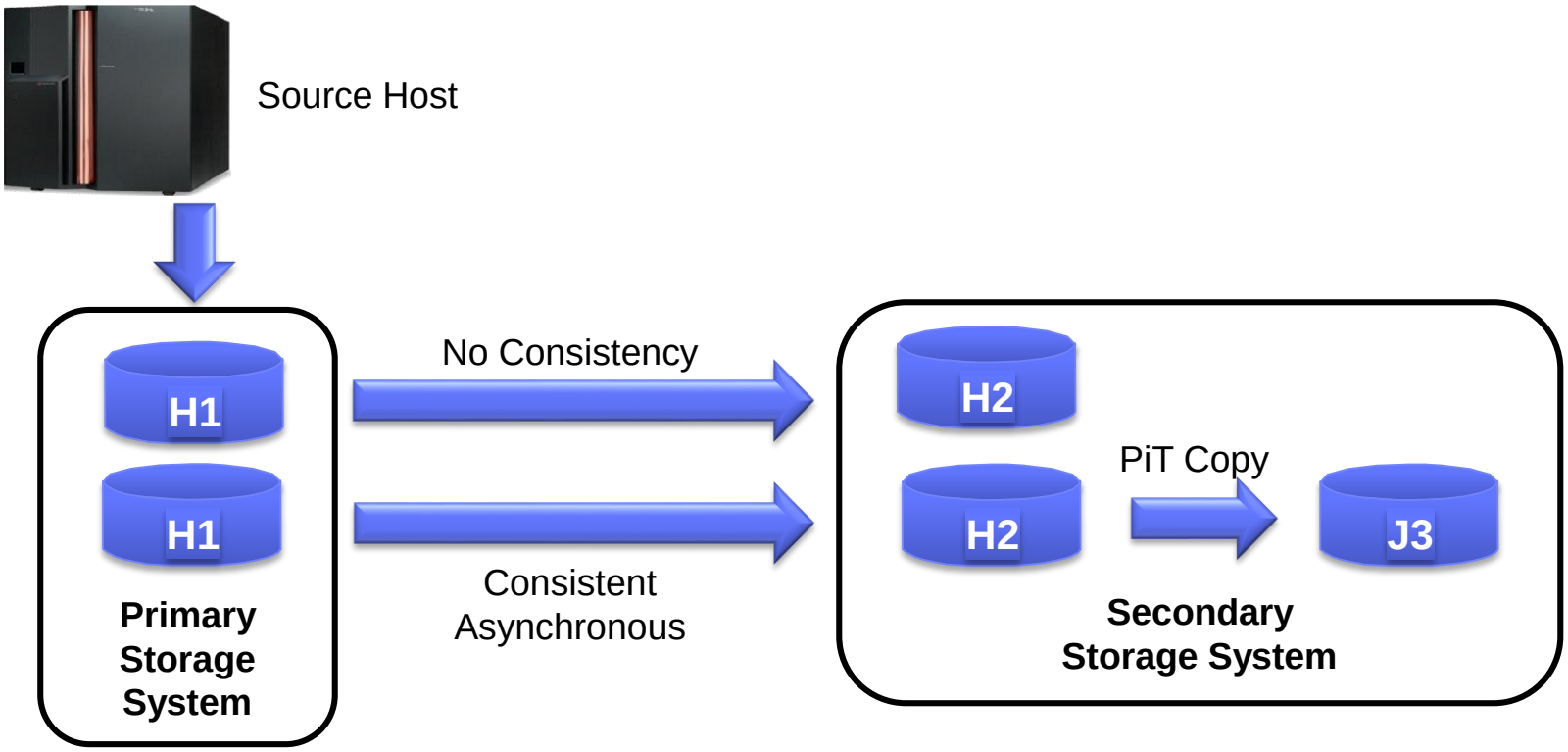
Fit For Purpose – Two & Three Site Replication

- Tailor your solution to your needs (and budget)
- Synchronous replication for everything
- Three-site Synchronous/Asynchronous only for your most important data



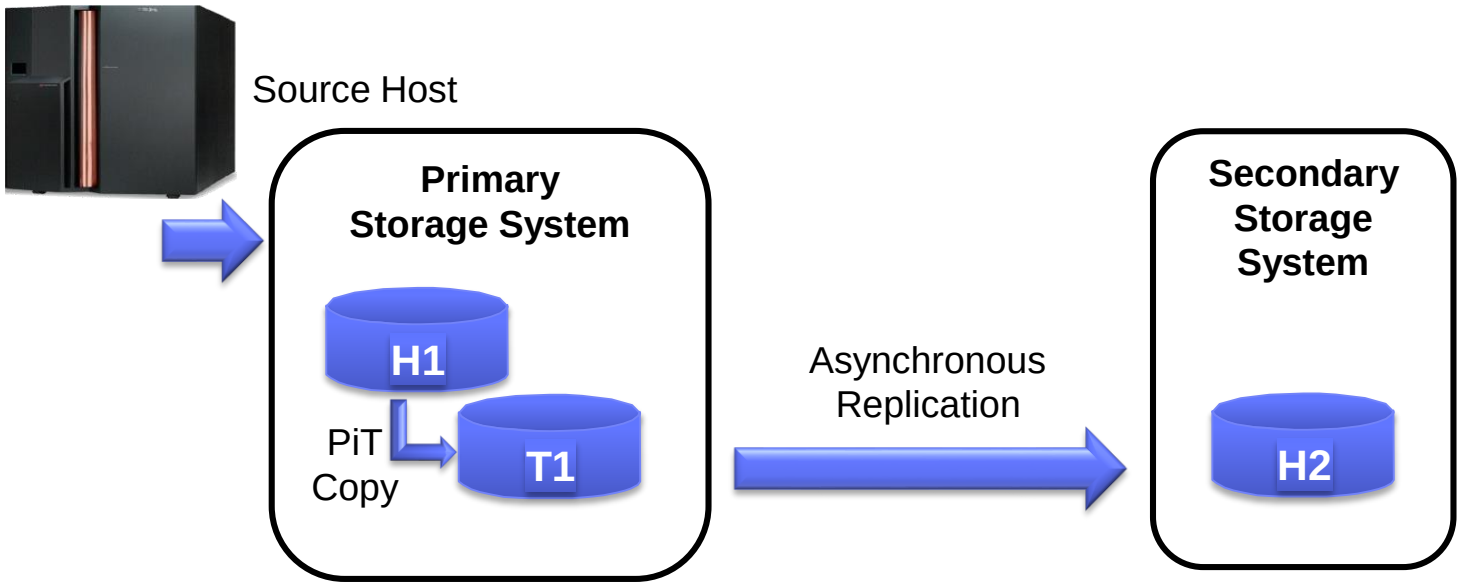
Fit For Purpose – Asynchronous Replication

- Save money and reduce complexity by replicating some data consistently and other data with no consistency
- Make sure you understand the ramifications of these decisions



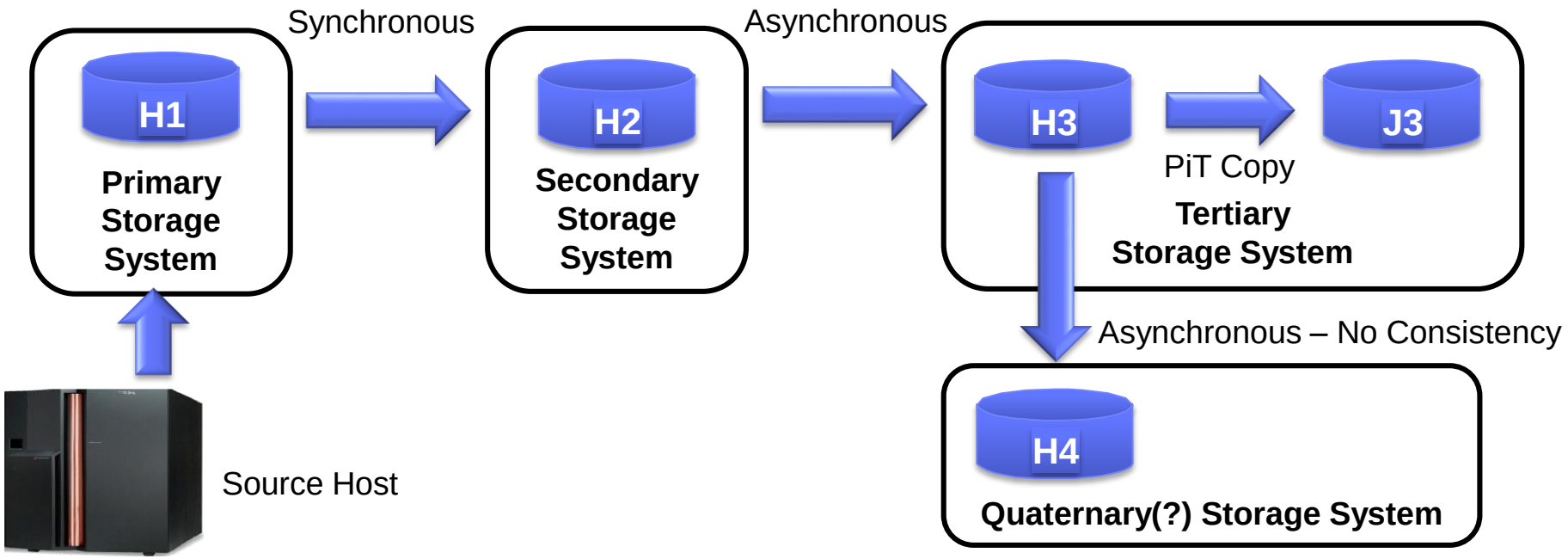
Low Cost Asynchronous with Consistency

- Create periodic consistent PiT copies of primary production volumes
- Use asynchronous replication with no consistency to copy all data to secondary
- When all the data is copied, the secondary volumes are consistent
- Lower network bandwidth requirements
- Avoids extra volume at secondary site
- Use Space Efficient PiT copy to conserve even more space
- Useful for testing, data analytics, or high RPO requirements



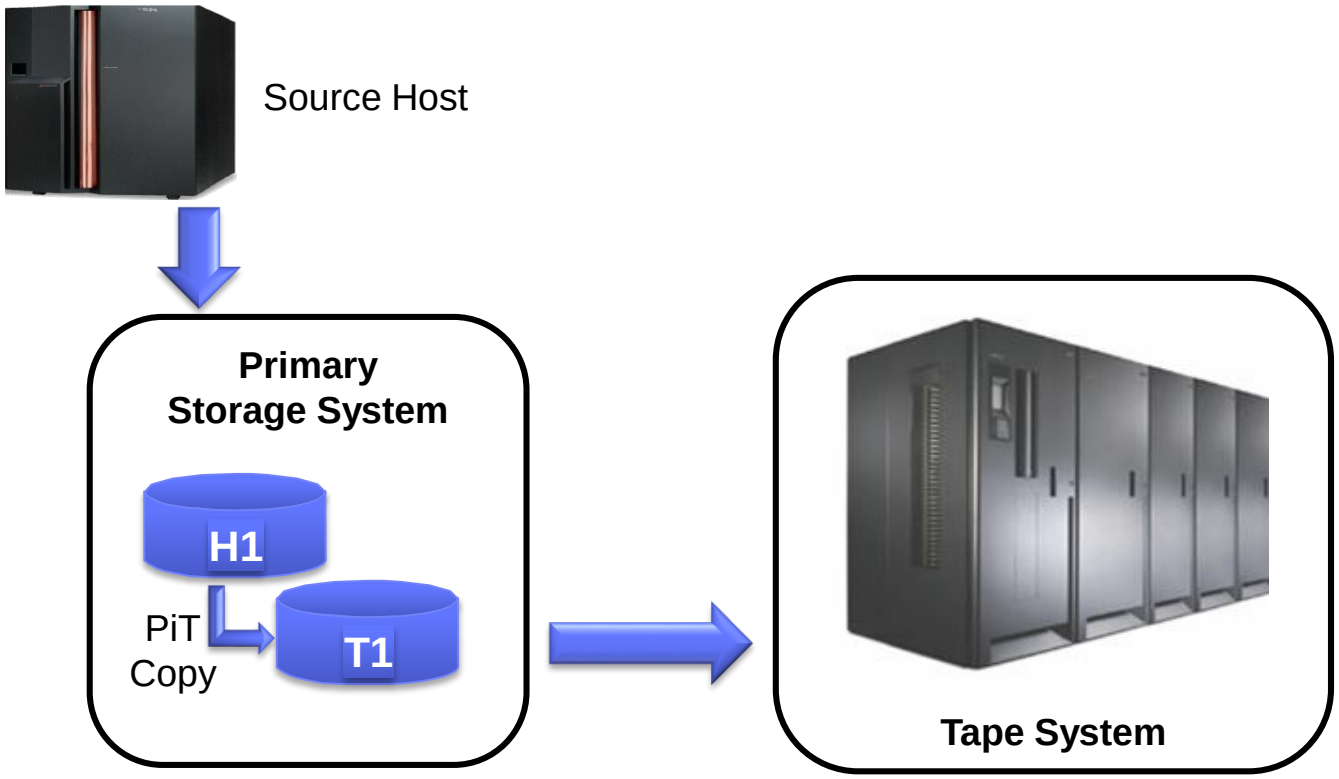
Four Site Replication

- Synchronous replication to provide high-availability
- Asynchronous cascaded replication to provide global distance DR capability
- Cascaded asynchronous leg to provide another copy of data
  - Used for development, testing, data analytics
  - Only consistent periodically



PIT Copies for Tape Backup

- Create periodic PiT copies of primary production volumes
- Dump these PiT copies to tape for backups
- Avoids the tape backup software accessing production volumes
- Use minimum space by employing Space Efficient PiT copies



Key Questions for Any Potential Solution

- How does the solution provide cross volume/cross subsystem data integrity/data consistency ?
- What is the impact to the primary application I/O ?
- What happens if data replication fails or slows down ?
- Interoperability with other data replication solutions ?
- Cost of installing & maintaining solution?
- Do solutions provide “concurrent maintenance” ?
- What flexibility does the solution provide ?
- If I recover to the secondary site, how do I replicate back to the primary?
- If I use different “types" of disk subsystems, after recovery can I maintain my QoS to my users?

Thank you!

John Wolfgang
Enterprise Storage Architecture & Services

jwolfgang@vicominfinity.com
26 June 2014

About Vicom Infinity

For More Information please contact…

Len Santalucia, CTO & Business Development Manager
Vicom Infinity, Inc.
One Penn Plaza – Suite 2010
New York, NY 10119
212-799-9375 office
917-856-4493 mobile
lsantalucia@vicominfinity.com

About Vicom Infinity

- Account Presence Since Late 1990's
- IBM Premier Business Partner
- Reseller of IBM Hardware, Software, and Maintenance
- Vendor Source for the Last 8 Generations of Mainframes/IBM Storage
- Professional and IT Architectural Services
- Vicom Family of Companies Also Offer Leasing & Financing, Computer Services, and IT Staffing & IT Project Management