



# Data Stewardship

## Everyone's Responsibility

Rebecca Levesque  
President, CEO

21<sup>st</sup> Century Software - Booth 209

Leaders in Data Stewardship™

August 13, 2013 at 12:15 pm

Session #14226



# Overview

- Data Stewardship terminology
- Rationalization and fallacies
  - IT resiliency or disaster recovery
  - RTO/RPO
  - High availability solutions
  - Reality of testing
  - Change control
  - People factor - what you don't know
- Responsible data stewardship



## Data Stewardship

- “The conducting, supervising, or managing of something *especially*: the careful and responsible management of something entrusted to one’s care” (Merriam Webster, 2013)
- Related words - protection, safekeeping, trust
- Mundane issues cause downtime

*IT personnel are the stewards of our respective organizations most important asset – its data. This must be done in a responsible and careful manner because we are responsible for its protection and safekeeping and we are trusted with the responsibility whether the “whole” piece is our job or not.*

# Data Stewardship Trust

- *“IT...is like the heart pumping blood to the whole body, so any failure could threaten the whole organization’s survival” (IBM reputational risk and IT whitepaper, 2013)*
- IT potentially damages a company’s brand and/or reputation if the top three risks occur
  - Security breach
  - Data loss
  - System unavailability
- Management concern is at the reputation level
- Most enterprises take reactive approach

*“Underestimating the cost of reputational risk greatly exceeds the cost of protection. Being proactive is preferable to being reactive.”*

# News coverage to avoid



## 'IT Hall of Shame'

Presumably, a bank this size has worked out its business continuity and recovery plans with a level of efficiency only possible for a huge organization. Or maybe not.



## 'No Customer Transactions'

Wells Fargo didn't offer many details about the system failure, but it was serious enough that the company had to restore from backup. "Thanks to the hundreds of team members in our technology group for working so hard to resolve this problem."



## 'The Maintenance Fiasco'

The failure to apply the correct procedure to fix a simple problem in the data-storage system crashed most of the bank's systems. A procedural error triggered a malfunction in the multiple layers of systems redundancies, which led to the outage.

Suffice it to say that most of the bank's services were now down. Online banking was dead. ATMs were silent. Credit cards could not be used. Commercial banking was halted.

DBS' CEO issued a three-page detailed apology to the bank's customers.



## 'IT meltdown'

NatWest admitted that it could not say exactly how much money should be in individual accounts as the crisis caused by a failed software update last week spiraled out of control for days.



## 'IT failure train wreck'

Most IT failures go unnoticed by the public ... Maintenance on systems caused an error in our batch scheduler.

By the time the problem was isolated, the backlog was enormous, which created the long delay in getting operations back to normal.



## 'Outage with mainframe computer'

Customers were unable to use ATM cash machines, as well as online banking. "We cannot say, at this stage, what the problem was. HSBC is in the process of upgrading its systems but ruled this out as a reason for the failure."



# What do these brands have in common?



- System upgrades
- Very extended disruptions – missed SLAs
- Revenue & goodwill loss, possible fines
- Unable to find sync point to recover to SLA requirement
- Lost transactions
- All had a public display of unavailable systems causing customer dissatisfaction and brand damage



# Rationalization and Fallacies



# What is a disruptive event to business

"The **business risk** associated with non-recoverable systems, applications or networks are almost incalculable. If standard local backup recoveries fail at a rate of 20% plus, remote recoveries are probably two-to-three times that - and finding out during a real-time crisis is no way to live," (*Steve Duplessie, 2012 white paper*)

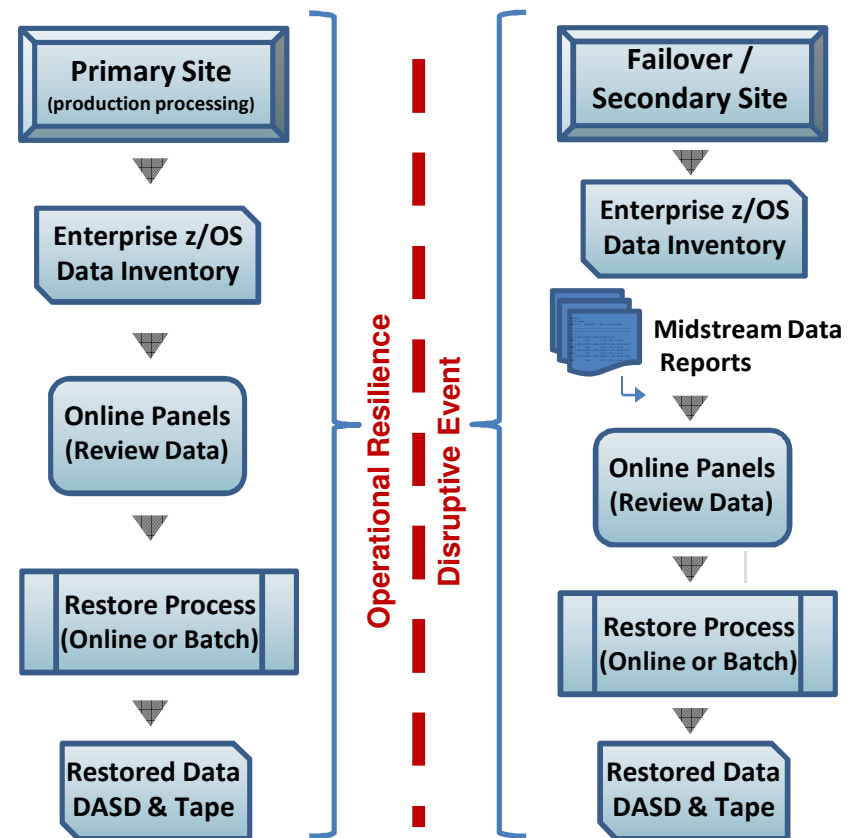
- Major issue in large data centers is coordination of application, storage, high availability and redundant team approaches to ever-changing business requirements and ability to recover
- How loyal are customers today in a global information world when any outage affecting THEIR business/finances depends on your system availability?
- The key is dwindling time frames of RTO and RPO



# IT Resilience or DR?

- Difference (from impact perspective) whether an outage is operational or considered disaster?
  - Vendors
  - Personnel
  - Hardware
  - Software
- Merely the side it's on

Operational and disaster recovery due to site issues, corruption, deletion or **any** disruptive event



# Clarify Terminology – Major Disconnect

## RTO Timeline

### Outage

#### Assessment

- Triage
- Decision on process (failure, restore, DR vendor)
- RTO Begins



## SLA Timeline

### Outage – seconds count



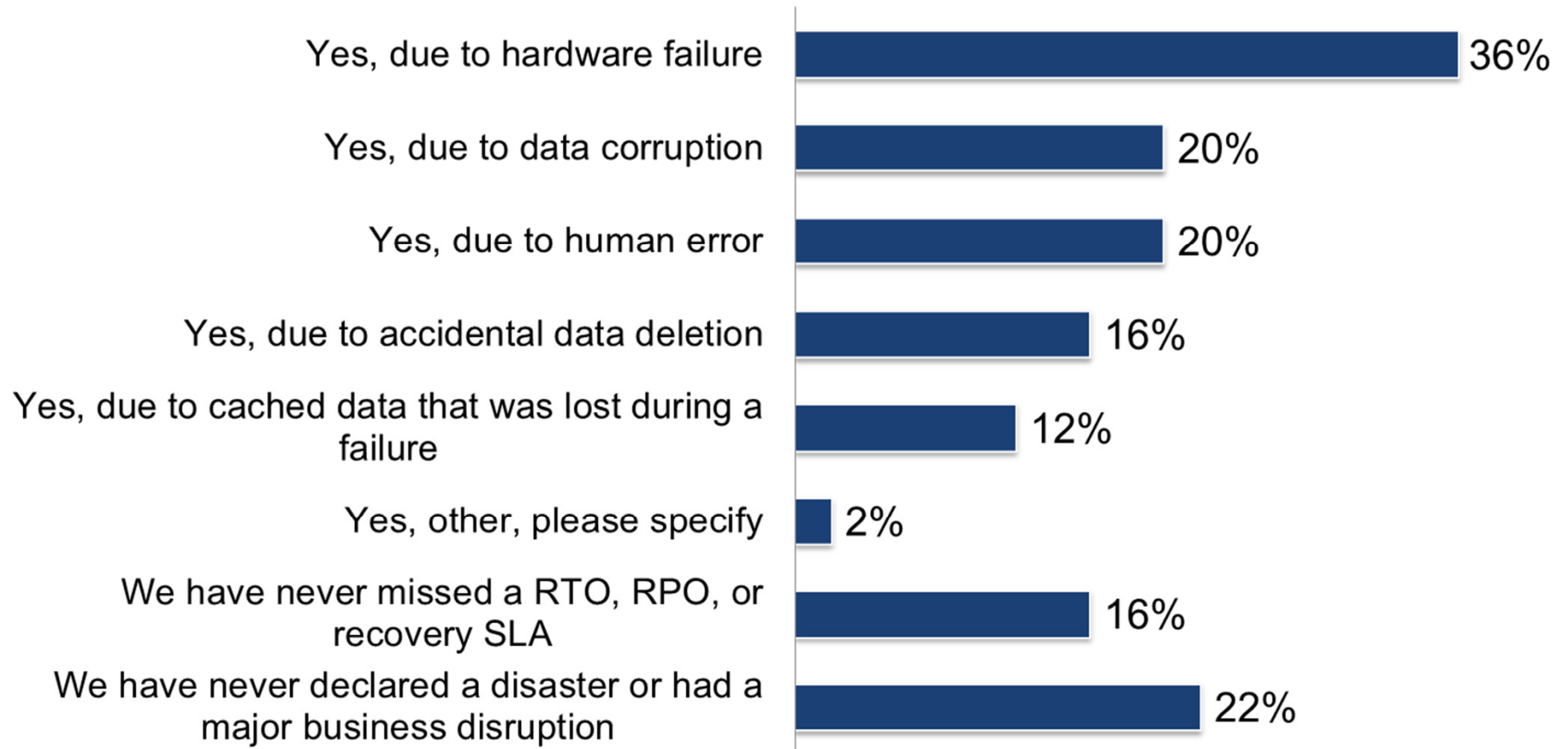
# What about RPO?

- Gartner reports top 5 complaints from business
  - Difficult meeting backup window
  - Need to troubleshoot and restart failed backup jobs
    - **Check points and automatic restart are not part of all applications especially VSAM and non-VSAM**
  - Time to restore data – missing the SLA
  - Lack of reporting to prove data is recoverable
  - Different point solutions (depending on platforms)
- Business cares about application availability
- IT focuses on infrastructure
- False sense of security around total system redundancy of applications and systems

# Recent Forrester Survey of large z/OS clients:



**“Have you ever had an operational disaster that made parts of your business unavailable for longer than your RTO, RPO, or recovery SLA?”**



Base: 50 disaster recovery decision-makers and influencers in North American enterprises

Source: “Final”, a commissioned study conducted by Forrester Consulting on behalf of 21<sup>st</sup> Century Software - February, 2011

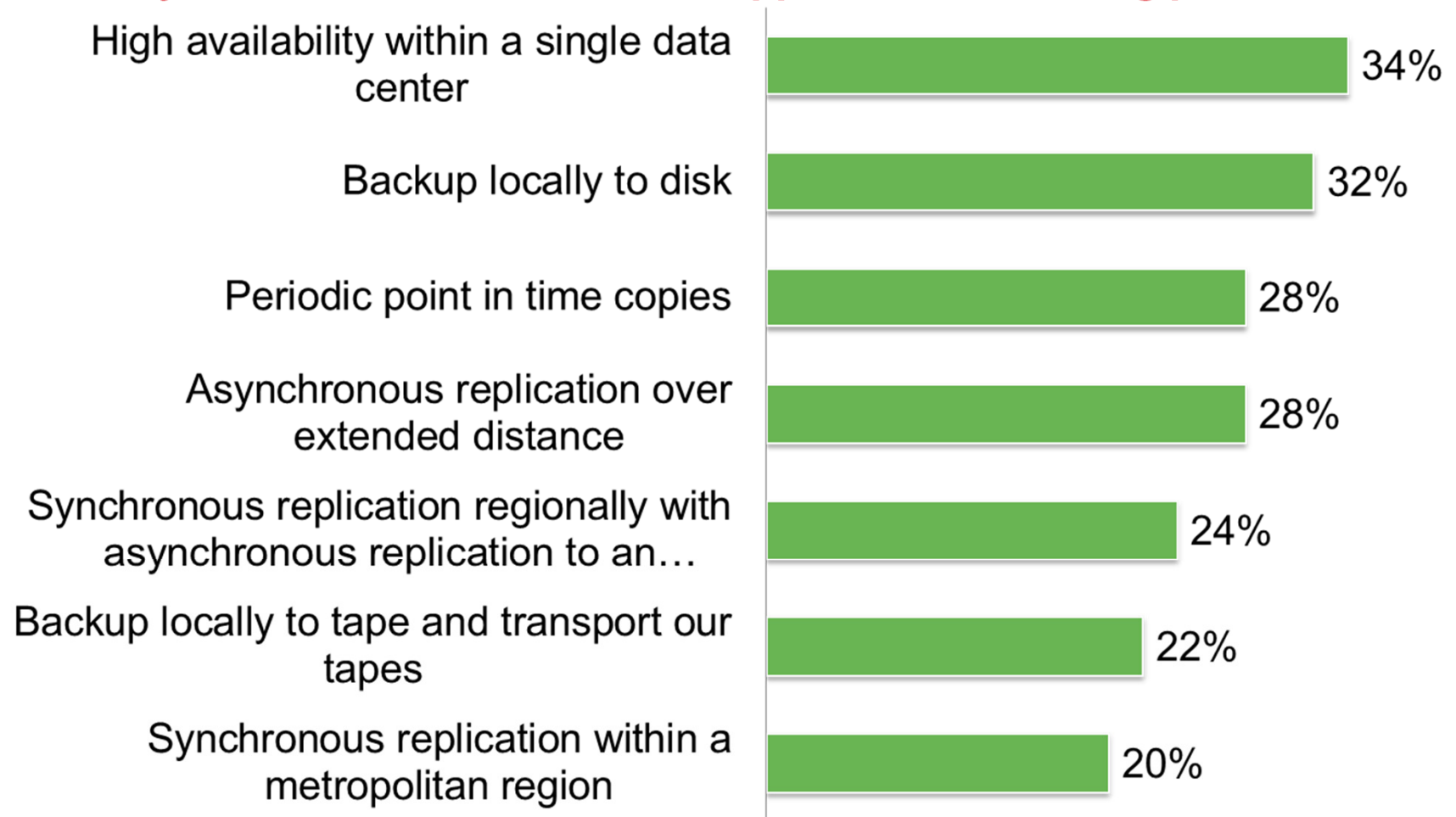
Complete your sessions evaluation online at [SHARE.org/BostonEval](http://SHARE.org/BostonEval)



# Recent Forrester Survey of large z/OS Clients:



“How is your most critical mainframe application data being protected?”



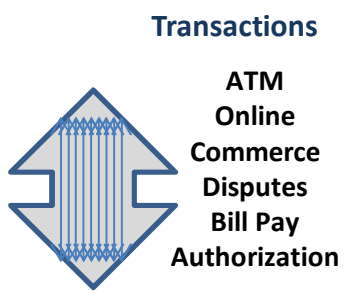
Base: 50 disaster recovery decision-makers and influencers in North American enterprises

Source: “Final”, a commissioned study conducted by Forrester Consulting on behalf of 21<sup>st</sup> Century Software - February, 2011

Complete your sessions evaluation online at [SHARE.org/BostonEval](http://SHARE.org/BostonEval)



# Perfect World



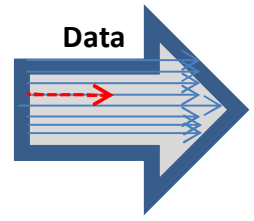
**DASD and Tape (including virtual), datasets are available at different times**

**Deferred copy consistency point: because of a delay performing the copy, access might or might not be available when a failure occurs.**

**Synchronous Replication (Under 100 Kilometers )**

**Asynchronous Replication**

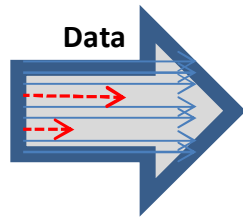
**Fourth Copy Snap/Flash/Full Volume**



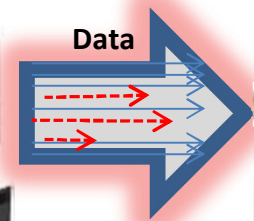
**Tape-on-DASD async at volume level**



**Tape-on-disk async at volume level, DASD mirrored**



**Tape-on-disk deferred mode at volume level, farther behind**



**Snap copies = fuzzy data**



**63% of respondents stated exercising tasks for all production applications was not automated.**

# Real World



## Transactions

ATM  
Online  
Commerce  
Disputes  
Bill Pay  
Authorization



## Production



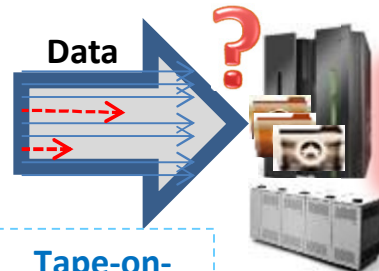
Synchronous  
Replication  
(Under 100  
Kilometers )



Tape-on-disk  
async at  
volume level

Application  
Run/Restart  
could fail

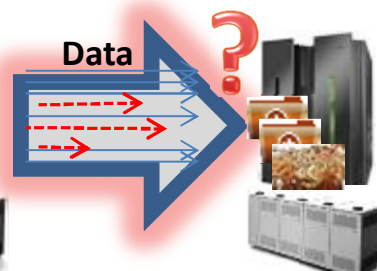
Asynchronous  
Replication  
Unknown data  
loss



Tape-on-  
disk  
deferred  
mode at  
vol level  
behind

App  
Run/  
Restart  
fails

Fourth Copy



Snap  
copies  
= fuzzy  
data

Application  
restarts fail  
what/when is  
fallback to  
clean data to  
sync with all  
apps?

### InFlight data

- VSAM data in buffer
- Tape not in sync with DASD
- VSAM re-org

Missed SLA – RTO - RPO



# How were your applications designed?

It is paramount to design availability and recovery into the very fabric of production operations as new systems, applications and data management techniques are deployed. Experience has proven that trying to retrofit these capabilities after new systems are brought online is often extremely difficult, if not possible.

*Integrating the resiliency needs of business and IT Functions, A holistic approach to recovery from an adverse event*

*July 2011*

*IBM Global Technology Services*



# Architecting for Availability

- Gartner states in many reports that companies have to architect for availability
- SLAs need to be tied to architectural standards for infrastructure, software and operations
- SLA planning and execution into the infrastructure design, application design and operational design
- Third party solutions (outsourcing, cloud, general service providers) have SLA penalties but don't pay for loss of customer perception of your system availability
- Network availability

# High Availability Solutions

## Good improvement – yet not complete



“With the emergence of new, combined business and IT functionality, it is mandatory to design infrastructures that can support availability and recovery from both a business function and an IT operational standpoint. **Often, availability and recovery are intertwined, with local availability being substituted for disaster recovery, or remote recovery being used as protection against wide-scale outages in production.**

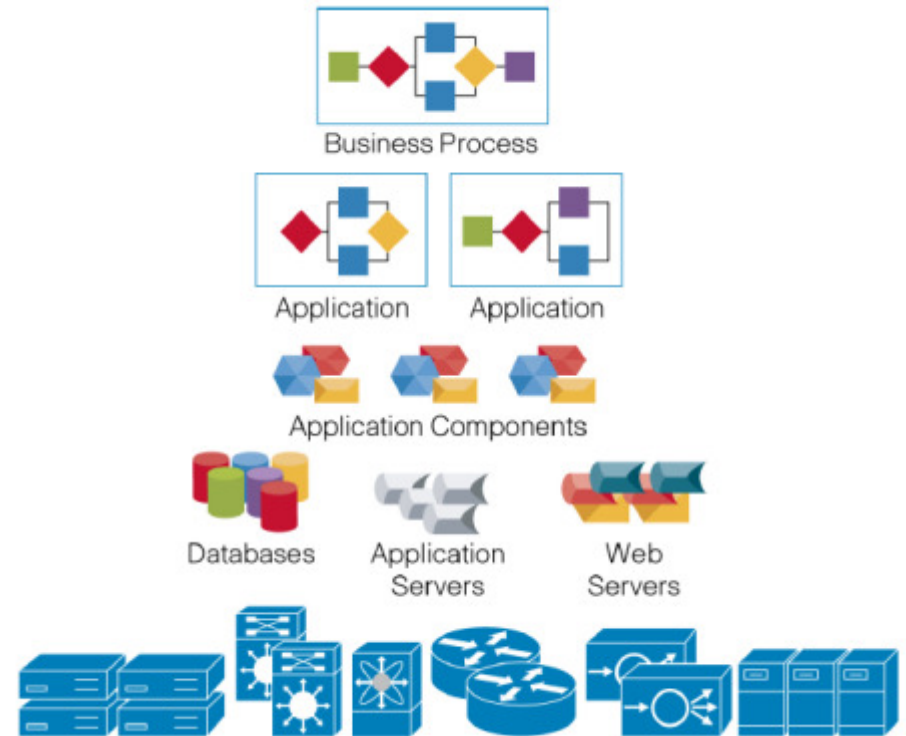
A critical factor to a ***successful recovery design is understanding the intricate data dependencies at the application level*** to enable the failing function integration back into the mainstream processing environment.”

*(IBM Global Technology Services, 2011)*

“**Remote mirroring is not well suited for granular recoveries, it is better qualified for recovery of the entire systems.**” *(SearchStorage.com, 2013)*

# Recognizing Data Dependencies

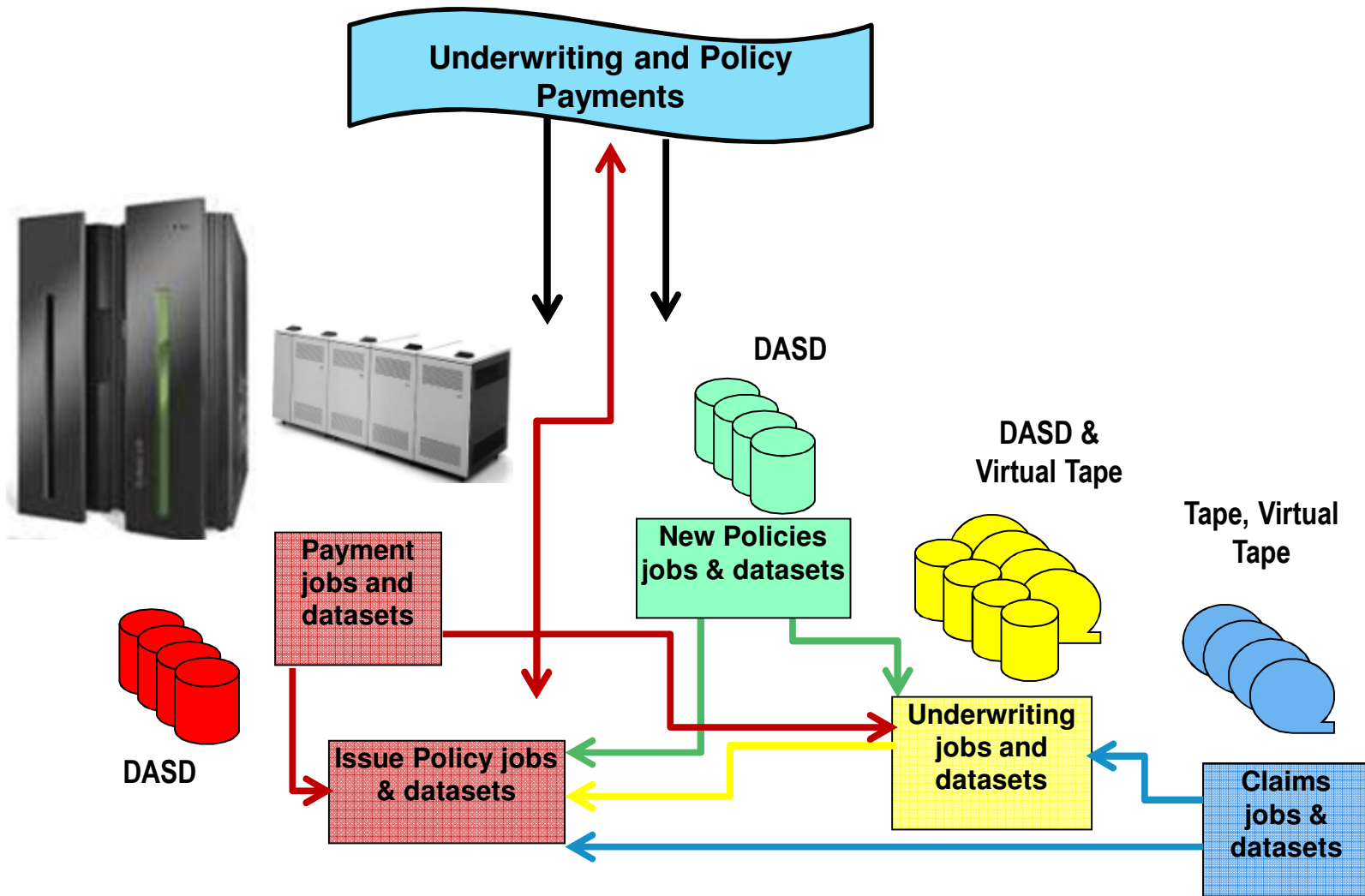
- Data in motion
  - Need dependency mapping
  - Improve data assurance
  - Detect data anomalies
- Preproduction pilots or production installation (technology assessments, modeling)
- Validation and gaps defined
- Result – improved predictability and efficiency



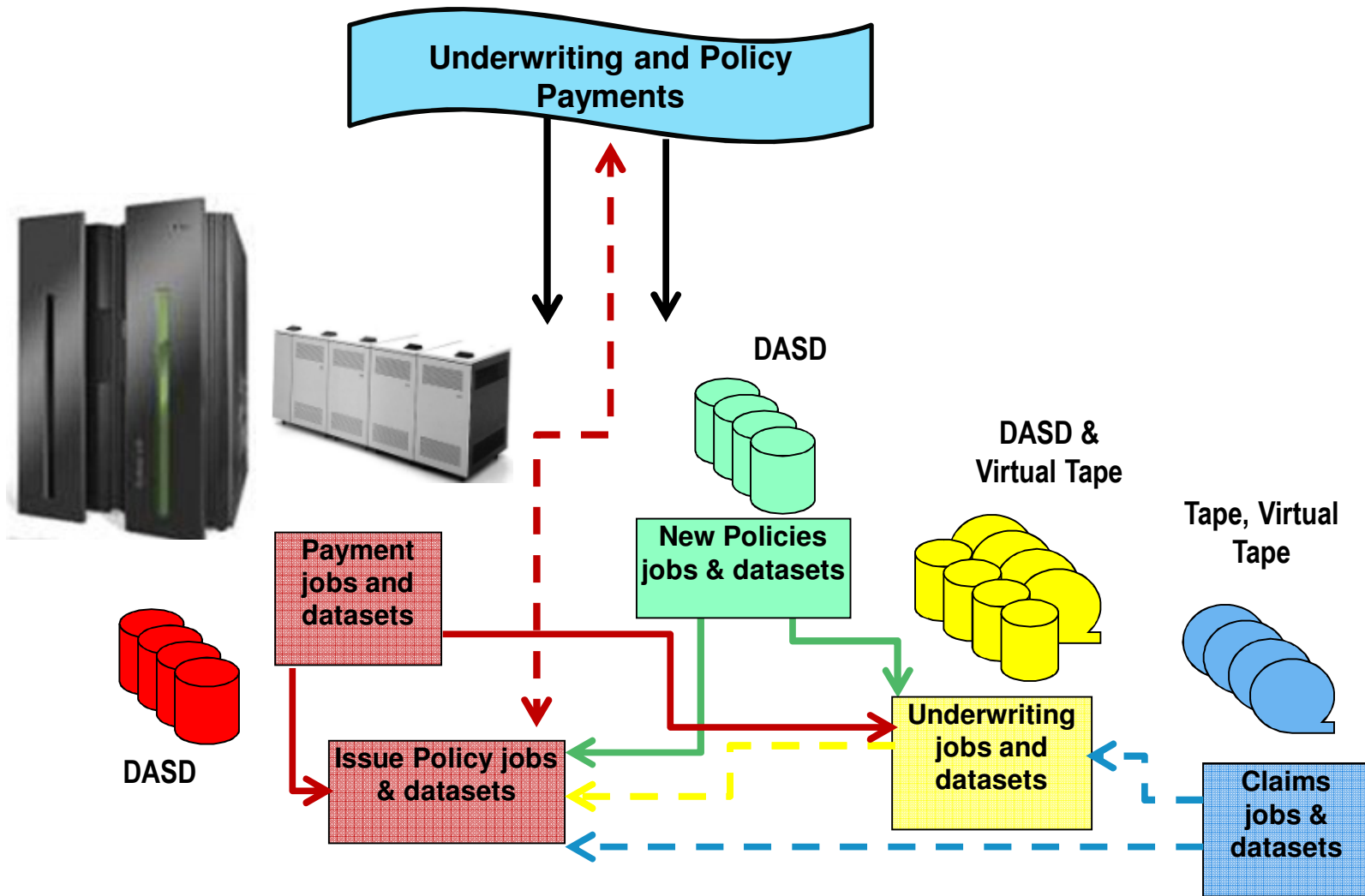
Gartner Hype Cycle for Business Continuity Management and IT Disaster Recovery Management, 2012  
 Published: 2012, Pages 14-15

Complete your sessions evaluation online at [SHARE.org/BostonEval](http://SHARE.org/BostonEval)

# Process/App Interdependencies



# Broken Interdependencies



# Resiliency Pitfalls

- Emphasis on system vs. application
- Plenty of data – where is generational – which data to use?
- How designed – for active/active?
- Backups for applications – where and when
- Intertwined dependencies are not known
- High availability solutions are not DR
- Batch jobs write at conclusion – interruption midstream create unknown status
- Corruption copied and not detected

# High Availability Panacea or More Pain?



"The current emphasis on disk-based mirroring and replication as a core technique for data protection is actually **exposing data protection strategies to epic failure**. Disks are failing at 50 to 1500 times the frequency stated by manufacturers, new information on **silent corruption** (the corruption of bits at the time of recording) is leading to increased questioning of data protection techniques such as RAID, and interconnect failures (array components other than disk drives) are starting to rival disk failures as causes of catastrophic data loss. **A witch's brew of issues with disk are under-reported by the industry but promise to make your next disaster a data disaster.**" *(Toigo, 2013)*

"The disadvantages [to replication] are that the **synchronization occurs at a very low level, below the ability to see data and file system errors or corruption**. This means that the **inevitable errors (some of which are potentially fatal to a restoration)** that occur naturally over time will be immediately replicated to the offsite location...and the hoped for protection is thereby immediately lost." *(Unitrends, 2012)*

# Reality of Testing

Why tests and exercises fall into rationalization:

- Artificial environment creates sense of competence and passes audit review
- Limited or staged integration of infrastructure & applications
- Rarely demonstrates entire infrastructure, applications, processes, rollback, generational data
- Serves as procedural checklist without all procedures (ex., interrupting batch cycle when data has not written)
- Would anyone dare pull the plug based on a “perfect” test?



**or**





# One story not making the headlines

- Outsourced data center, hard disk array failed, backup to HDA was RAID (also failed to restore)
- Jobsabend, mid batch dataset corruption replicated, and applications had to use run/restart procedures
- Had to roll back to generational data
- 24 hours to get systems/app programs restored
- Took 48 additional hours to restore and have backlogged transactions up to present
- Major logistics application unable to process any data until the app databases and dependencies sync'ed

# Change Control

If there is a configuration change in the production environment, it needs to be reflected in the recovery environment. And, **in most real world organizations, changes to the production environment are happening constantly** ... IT environments are getting more complex, so it is with their recoveries.”

*(SunGard Managing Recovery white paper, 2012)*

- How consistently are change control processes applied to all backup systems and **tested**?
- Best of intentions, policies are broken – just one pointer to a dataset that is not on production and not accounted for or backed up can crash the system
- Are effects of daily change management monitored for impact to recoverability
- What are true rollback procedures from change?

# People Factor

**“Therefore it helps to know where you stand in terms of your current security measures so that you’ll be better prepared to deal with the issues should they arise in the event of a disruption** – whether it is due to a data breach, natural disaster or other unplanned event. While many organizations have IT and security administrators, it’s a good idea to have an external consultant run the audit for you.” *(Philip Owens, CSO online, 2011)*

“Identifying critical resources and recovery methods is the most relevant aspect during this process since you need to ensure that all critical applications and data are included in the blueprint.” *(Tech Target, 2013)* ***How do you include in a blue print what you don’t know?***

# What you don't know... or realize

- Humans run technical recovery – not push button yet
- Who grants highest security access to get systems up if the primary go-to is not available
- Can the network (VPN) handle increased traffic?
- How is security maintained during a disruption when most vulnerable
- Tracking changes made on the fly to get things running
  - Change control
  - Documentation
- Return to normal – are tracked access changes undone?

# Gartner Best Practices

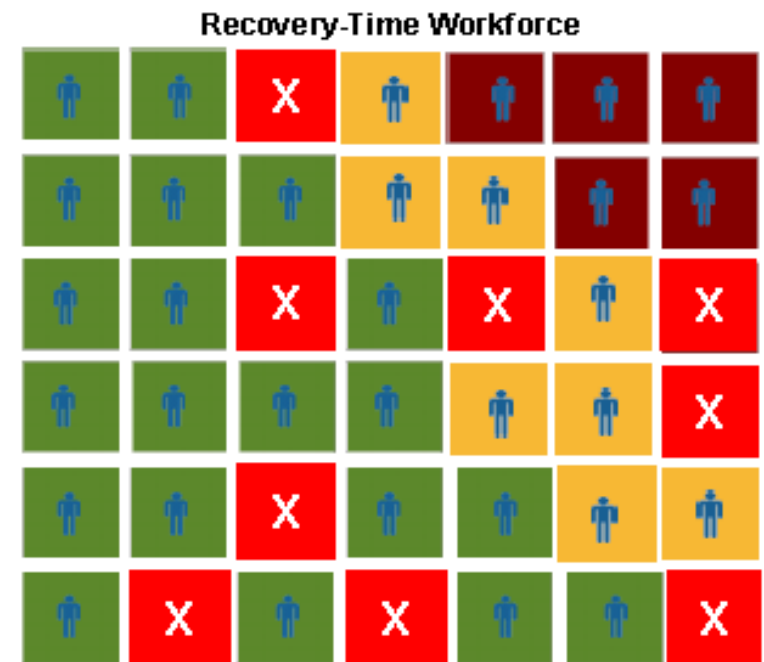
During a crisis, critical business and IT services need to be quickly restored or replaced for timely and effective recovery of the organization...Even with a plan, **the workforce will be operating under a heightened level of stress, such that business-as-usual and recovery operating procedures may not be diligently executed.**

Consider implementing a versatile authentication server or service to simplify the implementation of multiple authentication methods in multiplatform environments. It can be implemented and integrated only once — making support for alternative authentication methods during recovery much more viable.

*(Gartner, 2009)*

# Gartner Key Findings - Access

- Few programs include recovery access requirements in development processes and ongoing maintenance activities
- Not having needed access to mission-critical systems and recovery procedures delay the business coming back online
- Recovery access requirements must be managed for all workforce members
- Granting emergency recovery access in the midst of a disaster creates access policy violations that may not get closed after the situation



*Fewer people could mean increased access policy violations.*

**X is unavailable**

# Responsible Data Stewardship Keys

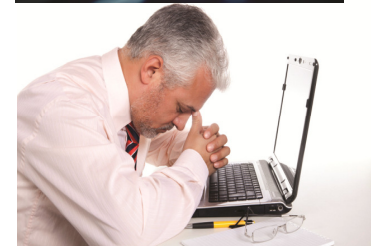
- Investment in need for speed and availability?
- Look for tools to validate, point to gaps, improve efficiency, reduce resources
- Plans not enough, tests not reality
- Find the gaps in silos – systems, data storage, infrastructure, apps
- Everyday changes amplify potential resiliency problems



# Evaluate your data stewardship

## Preventing damage to brand and reputation

- Understand and use recovered data reliably with respect to application sync points/dependencies
- Certify business SLAs through test simulation including month end, quarter end, year end
- Interdependencies readily identified to tier recovery data appropriately
- Improve CPU/storage efficiency through focused awareness
- Dynamic storage policies mapped to appropriate media based on actual size, use, and reference patterns of critical apps
- Modeling of policies to reduce hardware cost and boost resource efficiency
- Understand how change management affects applications, storage, and service providers for business goals assurance





# Final Thought

The lifeblood of almost every organization is data, and the requirements to **protect, maintain, and access the systems that store that data are absolutely essential to support business operations.** With the sheer amount and varying types of information that must be stored, the complexity of managing the data can grow as quickly as the amount of data itself.

**In this fast-moving day and age, businesses cannot afford to wait for anything** — if they do, they risk, at best, losing sales to the competition, or, at worst, falling behind the competition and **being pushed out of business.**

*(Storage Economics for Dummies - Hitachi Data Systems, 2012)*

 #SHAREorg

# Questions?



# Thank You!!



For additional information, please contact:

**Rebecca Levesque**  
**President, CEO**  
*Leaders in Data Stewardship™*

21<sup>st</sup> Century Software  
940 West Valley Road  
Suite 1604  
Wayne PA 19087  
U.S.A.

610 971 9946 x 200 (office)  
610 659 6521 (cell)

RebeccaL@21csw.com

Complete your sessions evaluation online at [SHARE.org/BostonEval](https://www.share.org/BostonEval)



# 21<sup>st</sup> Century Software recognition

**Gartner** **Research**  
 Publication Date: 24 April 2008 ID Number: G00156289  
**Cool Vendors in Compliance and Risk Management, 2008**  
 French Caldwell, Tom Eid, Jay Heiser, Jeffrey Wheatman, John E. Van Decker, Roberta J. Witty, Dave Russell

**Gartner** **Research**  
 Publication Date: 9 June 2008 ID Number: G00157852  
**Hype Cycle for Storage Software Technologies, 2008**  
 Carolyn DiCenzo, Dave Russell, Stanley Zaffos, Kenneth Chin, Pushan Rinnen, John P Morency, Robert E. Passmore, Donna Scott

**Gartner** **Research**  
 Publication Date: 23 July 2009 ID Number: G00169503  
**Hype Cycle for Business Continuity Management, 2009**  
 John P Morency, Dave Russell, Roberta J. Witty, Steve Bittinger, Monica Basso, Ted Chamberlin, Dan Miklovic, Richard J. De Lotto, Jeff Vining, Brian Gammage, Mark A. Marzevicius, Ronni J. Colville, Matthew W. Cain, Donna Scott, Les Stevens

**Gartner** **Research**  
 Publication Date: 20 July 2011 ID Number: G00214668  
**Hype Cycle for Business Continuity Management and IT Disaster Recovery Management, 2011**  
 John P Morency, Roberta J. Witty  
 Because of the major disaster events that have occurred over the past year, business continuity and IT disaster recovery management are getting increased attention. Use this report to make informed decisions on what can best support the recovery and availability

A Custom Technology Adoption Profile Commissioned by 21st Century Software  
**Enterprises Are Not Properly Protecting Mission-Critical Data**  
 March 2011