Backups in the Cloud

Ron McCracken
IBM

August 8, 2011
Session 9844

Legal Information

The following are trademarks of the International Business Machines Corporation in the United States and/or other countries.

AIX*
DB2 Universal Database
e-business logo*
IBM*
IBM logo*

IBM eServer
OS/390*
OS/400*
S/390*
Tivoli*
Tivoli Storage Manager

TotalStorage
WebSphere*

* Registered trademarks of IBM Corporation

The following are trademarks or registered trademarks of other companies.

Intel is a trademark of the Intel Corporation in the United States and other countries.

Java and all Java-related trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc., in the United States and other countries.

Linux is a registered trademark of Linus Torvalds.

Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation.

SET and Secure Electronic Transaction are trademarks owned by SET Secure Electronic Transaction LLC.

UNIX is a registered trademark of The Open Group in the United States and other countries.

* All other products may be trademarks or registered trademarks of their respective companies.

Notes:

Performance is in Internal Throughput Rate (ITR) ratio based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput improvements equivalent to the performance ratios stated here.

IBM hardware products are manufactured from new parts, or new and serviceable used parts. Regardless, our warranty terms apply.

All customer examples cited or described in this presentation are presented as illustrations of the manner in which some customers have used IBM products and the results they may have achieved. Actual environmental costs and performance characteristics will vary depending on individual customer configurations and conditions.

This publication was produced in the United States. IBM may not offer the products, services or features discussed in this document in other countries, and the information may be subject to change without notice. Consult your local IBM business contact for information on the product or services available in your area.

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.

Information about non-IBM products is obtained from the manufacturers of those products or their published announcements. IBM has not tested those products and cannot confirm the performance, compatibility, or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

Prices subject to change without notice. Contact your IBM representative or Business Partner for the most current pricing in your geography.

This presentation and the claims outlined in it were reviewed for compliance with US law. Adaptations of these claims for use in other geographies must be reviewed by the local country counsel for compliance with local laws.

IBM's customer is responsible for ensuring its own compliance with legal requirements. It is the customer's sole responsibility to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer's business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer is in compliance with any law.

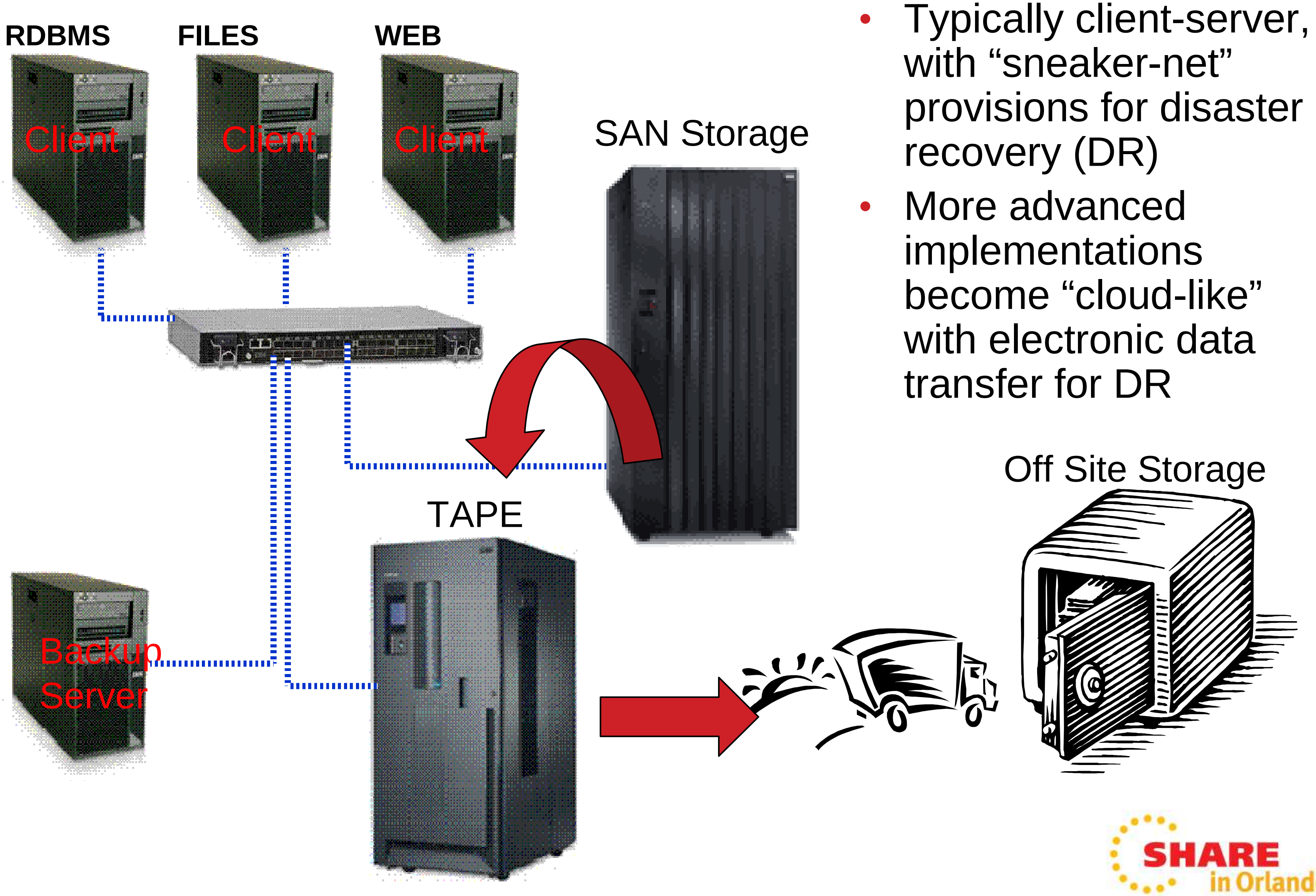
Agenda

This session is intended primarily to assist those considering the utilization of Managed Service Providers (MSPs) to perform some of the functions associated with enterprise data backup activities.

- What do we mean by cloud-based backups?
- Benefits
- Challenges and enabling technologies
- Architectures
- Summary and Recommendations
- Questions

Note: Mention of specific Companies/Products in this presentation are intended as examples and does not necessarily constitute IBM endorsement of said Companies/Products

Traditional Enterprise Backup Architecture



- Typically client-server, with "sneaker-net" provisions for disaster recovery (DR)
- More advanced implementations become "cloud-like" with electronic data transfer for DR

What are the characteristics of cloud-based backups?

- Still utilizes client-server architecture
- One or more elements of the traditional enterprise backup solution occur outside the company “firewall”
  - Even for a private cloud there is some exposure
  - Poses a security concern
- Generally involves web-based services, e.g:
  - Subscription to a backup service
  - Utilization of electronic data transfer to off-site DR storage
- Wide-Area networks are used

Simplified information management is the primary benefit of Cloud backups

- Offload one or more Continuity Management functions to an MSP, e.g.
  - Off-site DR data storage
  - Backup server management
- Simplify Capacity Planning (for backups, at least) to a simple fee-for-service

- Take better advantage of on-demand efficiencies

- Shift Service Level Management to a contractual, rather than technological, issue.

Data Security Poses a Concern

- Data security is about preventing unauthorized data access. The Cloud exposes organization data
  - If a public cloud is used, the backed-up data is outside the owning organization’s control
  - Once outside the firewall, data packets can be intercepted during transmission
    - The same concern exists for tape media traditionally used for off-site (DR) data copies, so this is not entirely a new risk
  - “Agentless” means “Well-Known” (to you AND attackers)
    - Keep this in mind when selecting backup technologies
  - The enabling technology is data encryption
    - Utilize encryption for all sensitive data sent to (or through) public facilities
    - Ideally, this should be a built-in capability of your backup software

Data Protection Concerns

- Data protection is about ensuring that data is not lost or corrupted. Use of an MSP does NOT eliminate the typical concerns:
  - Hardware failures
  - Media degradation/failure
  - Facility Disasters
  - Migration of data from obsolete technology
- Add the following concerns for an MSP:
  - Company failure
  - Legal disputes (including billing disputes)

Data Protection Strategies

- One possibility is to use an MSP only for your redundant or DR copies—continue to keep your own primary backups
  - High Availability Disaster Recovery (HADR) implementation
- Contract with more than one MSP
  - The best arrangement would be for MSP A to replicate your data to MSP B.
- Retain critical/sensitive data backups in house, use MSP for stuff you could survive losing, e.g. workstation backups.
- Or, very carefully evaluate your MSP for:
  - Their data protection strategy and capabilities
  - Their financial position (and this should be at minimum an annual review)

Network Bandwidth

- WAN networks re-introduce concerns for backup or DR windows that largely disappeared with the introduction of GB+ LAN/SAN capabilities.
- As with traditional LAN/SAN-based architectures, full DR restorations present the biggest challenge
  - The problem can be insidious with backup software using continuous incremental backup approach
  - Tends to drive current implementations to local primary backup, with DR copy in the cloud
- A combination of technology and management strategies are needed to deal with this issue

Network Bandwidth Acceleration

- The concept of network acceleration is that by a combination of buffering and compression technologies the true network load can be reduced, and spikes leveled, resulting in a higher apparent bandwidth
  - Data Compression
    - Object compression
    - Single Instance Store (SIS)
    - Deduplication
  - Network Accelerator Appliances
    - Either real (e.g. Riverbed Technologies) or virtual (e.g. Netex) implementations available
    - Typically a combination of local buffering and deduplication
    - Different “presentations”, e.g. network share, IP address, etc.
  - NOTE: These may, or may not, help much with restores—depends on how they cache chunks to reconstruct objects

Network Utilization Strategies

- Use Continuous Data Protection
  - Technology continuously backs up changes (file or block level) rather than doing periodic system-wide backup
- Use block-level (subfile) backup technology
- Employ disciplined Information Lifecycle Management (ILM) to reduce the DR problem
  - If you don't need it, delete it
  - If you might need it later, archive it (and delete from local storage)
- Have a prioritized DR plan
- Investigate whether your MSP can ALSO provide a DR site for business-critical servers.

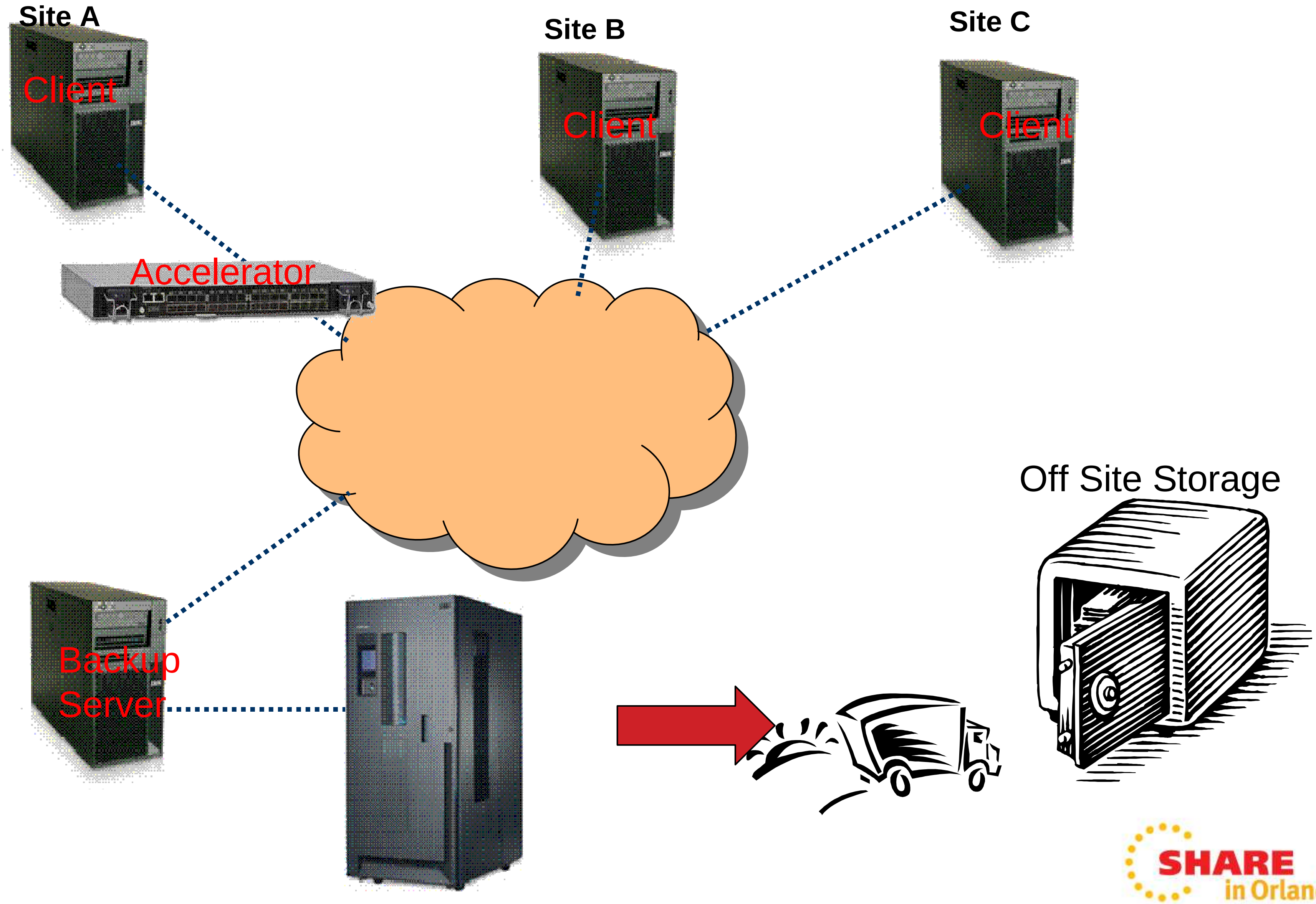
Factors in Selecting a Cloud Vendor

- What Service Level do they provide?
  - Backup and restore times
    - What about periodic DR tests? (The only way you'll know for sure)
  - Disaster Recovery
    - Are they prepared for a disaster at their own facility
    - Synchronous replication available?
    - Can they provide DR hot site facilities?
  - Can they work with a network accelerator, if so what kind?
- Do the actual facilities measure up?
  - What software, hardware, etc. do they use?
  - If possible, a site visit is recommended
- Data Security
  - Is data encrypted, and protected from unauthorized access?

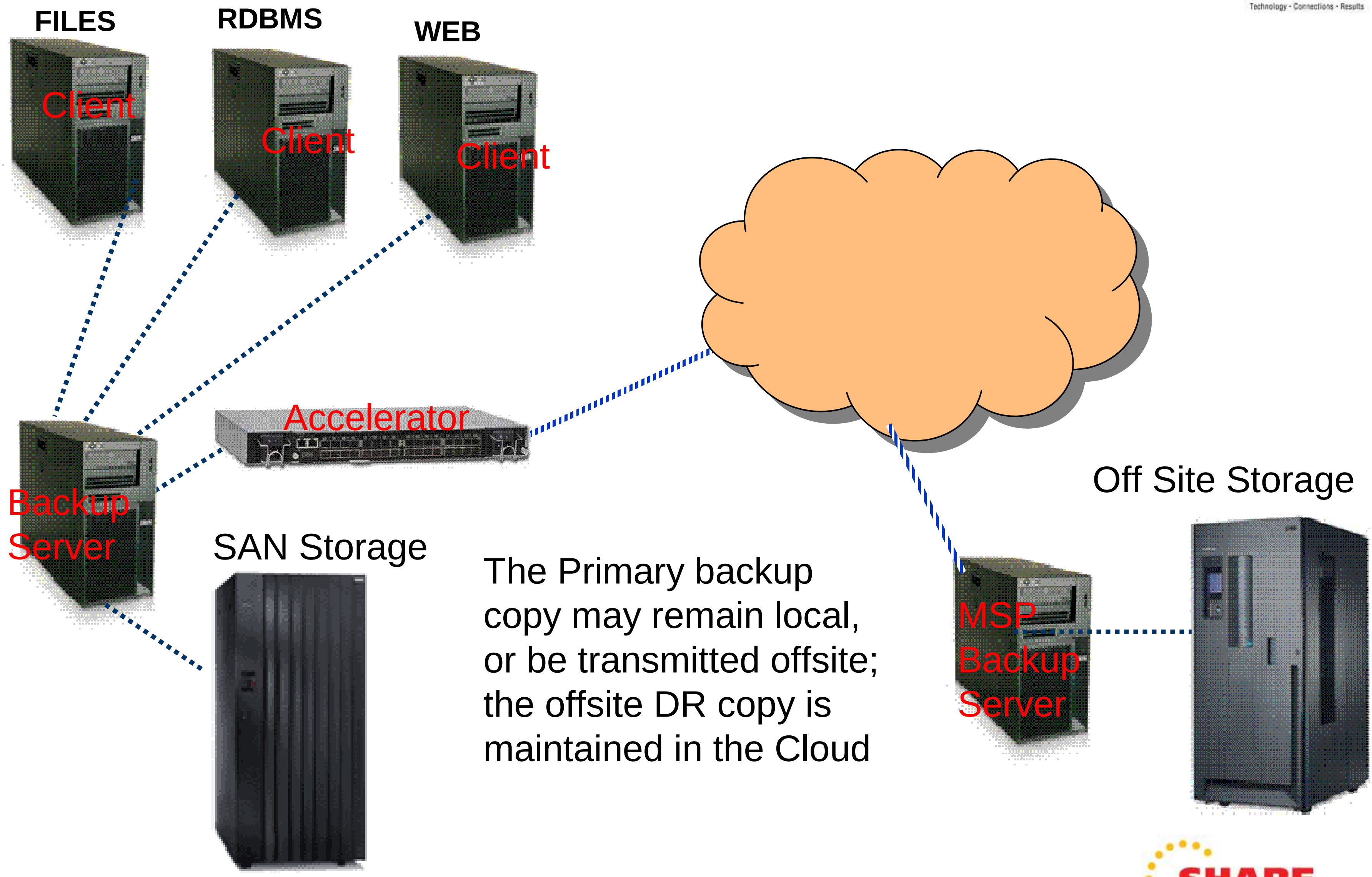
Architectural Options

- Private Cloud
  - Backup server(s) centrally located and administered
  - Remote sites back up to central servers
- Public Cloud
  - with local backup server
  - with cloud-hosted backup server
- Public-private Cloud
  - Private backup server
  - DR data copies sent to a public cloud storage service

Private Cloud (Remote Site) Backup Architecture

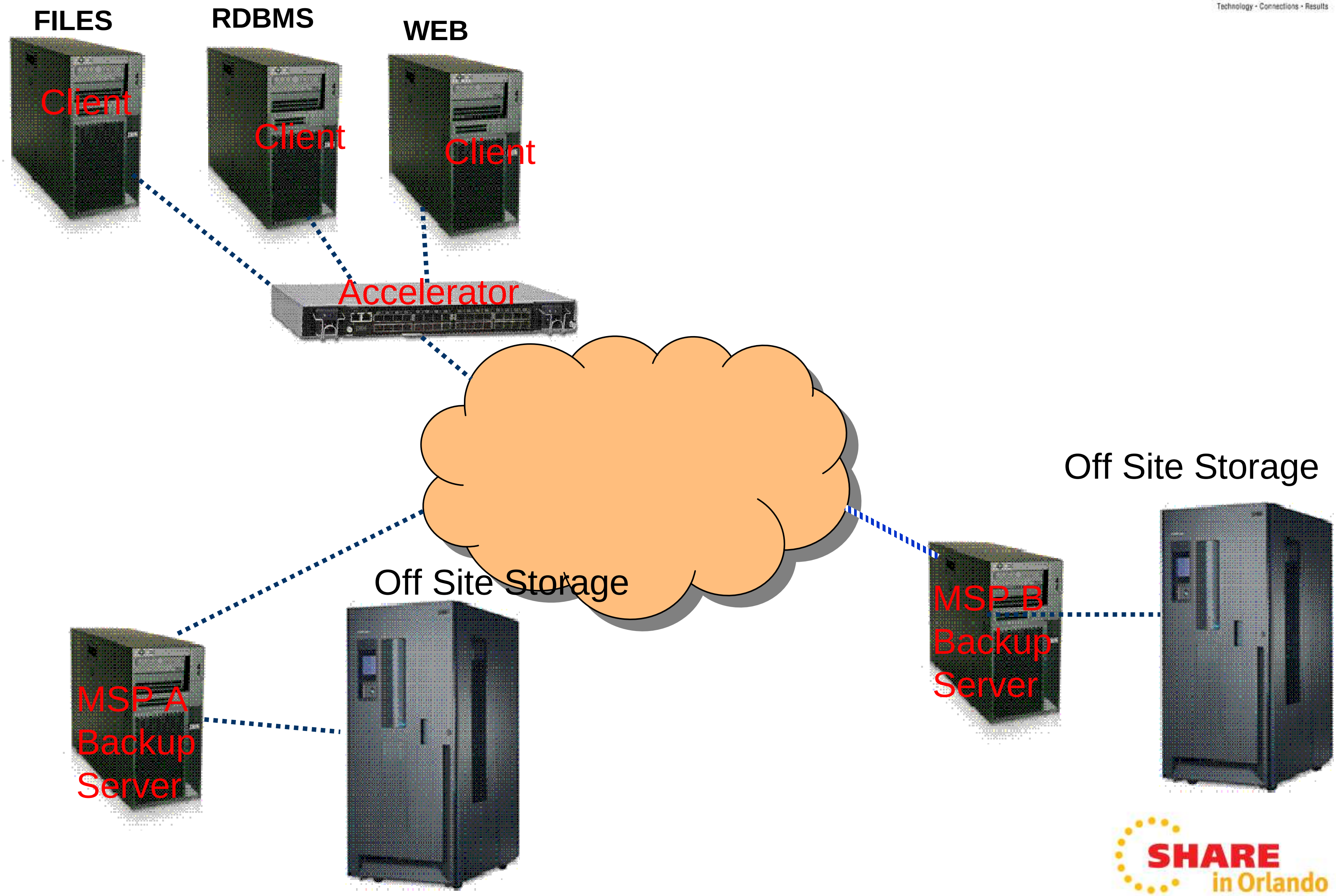


Public Cloud With Local Backup Server



The Primary backup copy may remain local, or be transmitted offsite; the offsite DR copy is maintained in the Cloud

Public Cloud With MSP Backup Server



When It’s Time to Railroad…

- Viable public cloud providers are plentiful for workstation backups
  - IBM (http://www-935.ibm.com/services/us/en/it-services/fastprotect-online.html)
  - Backblaze
  - Mozy
  - Carbonite
  - Etc.
- Less so for SMB enterprises, but coming on line
  - IBM
  - Starfire
- Just getting started for large enterprises
  - IBM

Summary and Recommendations

- If using a network accelerator, confirm compatibility with your (or your MSP's) backup software.
- If using a public cloud, evaluate carefully your candidate Managed Service Provider(s).
- Consider retaining conventional backups for large, business-critical servers.
- Disciplined ILM enabled with a good Content Management software package makes Cloud backups more viable (and for a variety of reasons, you really ought to be doing this anyway).

Questions and Discussion