Data Protection Modernization: Meeting the Challenges of a Changing IT Landscape

Tom Clark
IBM Distinguished Engineer, Chief Architect Storage Software
Data growth is continuing to explode

IoT will connect 30 billion devices by 2020

6.75 TBs of data will be created per person per day by 2018

2.5 Billion gigabytes of data per day

90% of data created in last two years

Growth of data from different sources

SOURCE: *2014 IBM Institute for Business Value Study on Infrastructure Matters; Gartner IT Metrics*
Storage trends

Continued dramatic growth in data driven by mobile, analytics, IoT, hybrid cloud, cognitive business and big data

Transition of data types with new workloads, new applications

- 60% of clients committed to SDS
- 23% interested

New storage deployment models powered by software

- 70% of clients deploying object storage or plan to within 24 months
- By 2018, storage-rich servers are expected to account for 50% of new capacity purchases

Transition to cloud models

- "IDC predicts that over 80% of enterprise IT organizations will commit to hybrid cloud architectures by 2017."

Sources: ESG, IDC
Traditional and New Generation Applications

Traditional application workloads

New generation application workloads

Traditional workloads require further optimization, leveraging features in applications, hypervisors and storage

Scale out infrastructure to support new workloads requires new approaches to management

Optimize here

Extend here
Traditional Data Protection Approach

**Purpose**
- Recover from storage failure
- Disaster recovery
- Recover from application or operator error
- Recover from malware
- Archive or regulatory copy

**Client or Proxy**
- Data Mover
  - Integration with APIs and CLIs
  - Incremental forever or Periodic full + incremental copy

**Backup Server**
- Integration with backup server
  - Disk/VTL
  - Tape
  - Object Storage

**Backup Server Integration**
- Hypervisors
- Storage Systems
- Traditional Applications
Traditional Data Protection Challenges

Backup/restore are times too long for critical applications
Traditional Data Protection Challenges

Scale out applications don’t map well to traditional backup infrastructures
Traditional Data Protection Challenges

Data protection for apps moving to cloud and cloud apps

On Premises

Cloud

IaaS, PaaS, SaaS
Traditional Data Protection Challenges

It’s hard to reuse backup data for other purposes

- Dev-Ops
- Analytics
- Backup Server
- Object Storage
- Tape
- Disk/VTL
Traditional Data Protection Challenges

Compliance with increasing regulations

82% have more than 10 copies of each DB
This amounts to more than 2000 database copies to manage

What are the typical number of physical copies of a given production instance that are made for the purpose of test/development, QA, DR testing, data warehouse loading, reporting, etc?

- 5 or fewer
- 60 to 10 copies
- 11 to 15 copies
- 16 to 20 copies
- More than 20 copies

82% have more than 10 copies
Consistent with other IDC research
Traditional Data Protection Challenges

Avoiding and recovering from ransomware

Air Gap?: Periodic Connectivity Replication

Backup Server

Disk/VTL

Object Storage

Tape
Evolution of Data Protection: Storage Snapshot Backup

Advantages:
• Fast backup, fast recovery

Challenges:
• End to end orchestration
• Complexity
• The old way works (or used to)
Evolution of Data Protection: Self-Protecting Storage

Advantages:
- Fast backup, fast recovery

Challenges:
- End to end orchestration, reporting
- Data movement for large backups

Snapshot or Differential Backup

Direct Copy to Storage

Object Storage

Tape

Disk/VTL
Evolution of Data Protection:
Scale-out Application Backup
Self-protecting Applications, Self-service

Advantages:
- Operates at scale

Challenges:
- End to end management, reporting
- Lack of capability in current applications
- App developer expertise
Evolution of Data Protection: Data Protection in Clouds

Advantages:
- Similar operations to on premises

Challenges:
- Centralized hybrid management
- Lack of capability or APIs in cloud *aaS

<table>
<thead>
<tr>
<th>On Premises</th>
<th>Cloud</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disk</td>
<td>Object Storage</td>
</tr>
<tr>
<td>Object Storage</td>
<td>Disk</td>
</tr>
</tbody>
</table>

- Traditional Backup
- Snapshot or Application-driven

IaaS, PaaS, SaaS
Evolution of Data Protection: Data Reuse

Advantages:
- Leverage collected data
- Immediate data availability

Challenges:
- End to end orchestration
- Performance of mounted data

Snapshot or Differential Replication

DevOps and Analytics

Direct Mount

Backup Server

Object Storage

Tape

Disk/VTL

Backup

DevOps and Analytics

Client

Snapshot or Differential Replication

Backup

Disk/VTL

Object Storage

Tape
What We Need in Data Protection

**Data Sources**
- Traditional Apps
- Hypervisors
- Scale-out Apps
- Storage Systems
- IaaS, PaaS, SaaS

**Key Capabilities**
- Software-defined solutions
- Simplicity
- Orchestration
- Central catalog
- Central reporting
- Self-service
- Complete data functions

**Data Functions**
- Data Mover
- Integration
- Mount/Re-use
- Application or Storage Replication

**Data Sinks**
- Backup Server
- Object Storage
- Tape
- Disk/VTL
- App Instance
- Storage Systems
Thank You!