STORAGE DEVELOPER CONFERENCE

SDC 23

BY Developers FOR Developers

Data Immutability – Retention Locking / WORM

Sailu Yallapragada, Distinguished Engineer, **Dell Technologies**

Jagannathdas Rath, Software Senior Principal Engineer, **Dell Technologies**

Key Takeaways

- Data Immutability & Retention
  - What is Data Immutability and Data Retention?
  - System Implementations to enable data immutability
  - Retention Locks/WORM and its features
  - Enhanced Backup workflows with Data Immutability
  - Tackling Cyber and Ransomware attacks with Immutability
- Complete Data Immutability Techniques against attack vectors and best practices
- Data Immutability in action (example use cases)
  - Immutability in Replication
  - Air-gapped Cyber Secure Vault

Data Immutability & Retention

Data Retention, Backup workflows with and without immutability

Immutable Data

Immutable Data (unmodifiable, indelible)

Data that cannot be modified or deleted, once written. It can be read multiple times though.

Data Retention

Regulatory Requirements

- Securities and Exchange Commission (SEC), FINRA, SOX, GDPR etc.
- Organizations coming under these rules must comply with the policies
- Backup data (and its copies) must be protected in non-modifiable and non-erasable format for the required duration

Governance Requirements

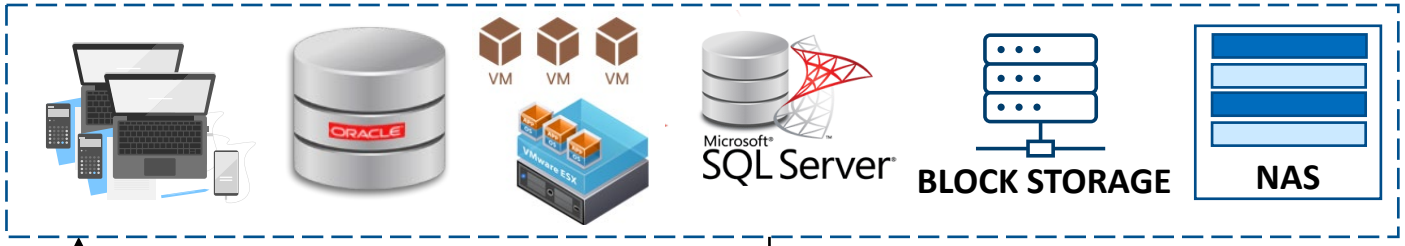
- Many companies have self-imposed retention policies
- Internal policies to preserve data, auditing purposes
- Meet privacy regulations where historical data might be requested by customers or government
- No mandatory duration to comply

Typical Data Backup Flow

Primary Data

Keeps on getting modified

Cannot be made Immutable



Assets

Primary data getting backed up as per policy schedule or manually

Backup Application

Backup Policies & Schedules

Backup Servers & Assets added

Backup Storage

Files, Archives

Virtual Disk images

Block Volume Snapshots

Filesystem Snapshots

……

Backup Data

Point in Time Copies

Never gets modified

Can be deleted at any time

Can be made Immutable

Cyber Attacks & Ransomware

Cyber Attacks

- Hackers/Attackers gain access of the data centers/storage servers
- Via Stolen credentials, Weak credentials, Phishing attacks, Insider attacks
- Objectives of such attacks – "Gain access to confidential data" OR Destructive breach - "Destroy all data, backups and copies – to bring down the organization"

2022-2023

25% of all data breaches[1]
Avg. $5.24 Million loss/attack[1]

Ransomware Attacks

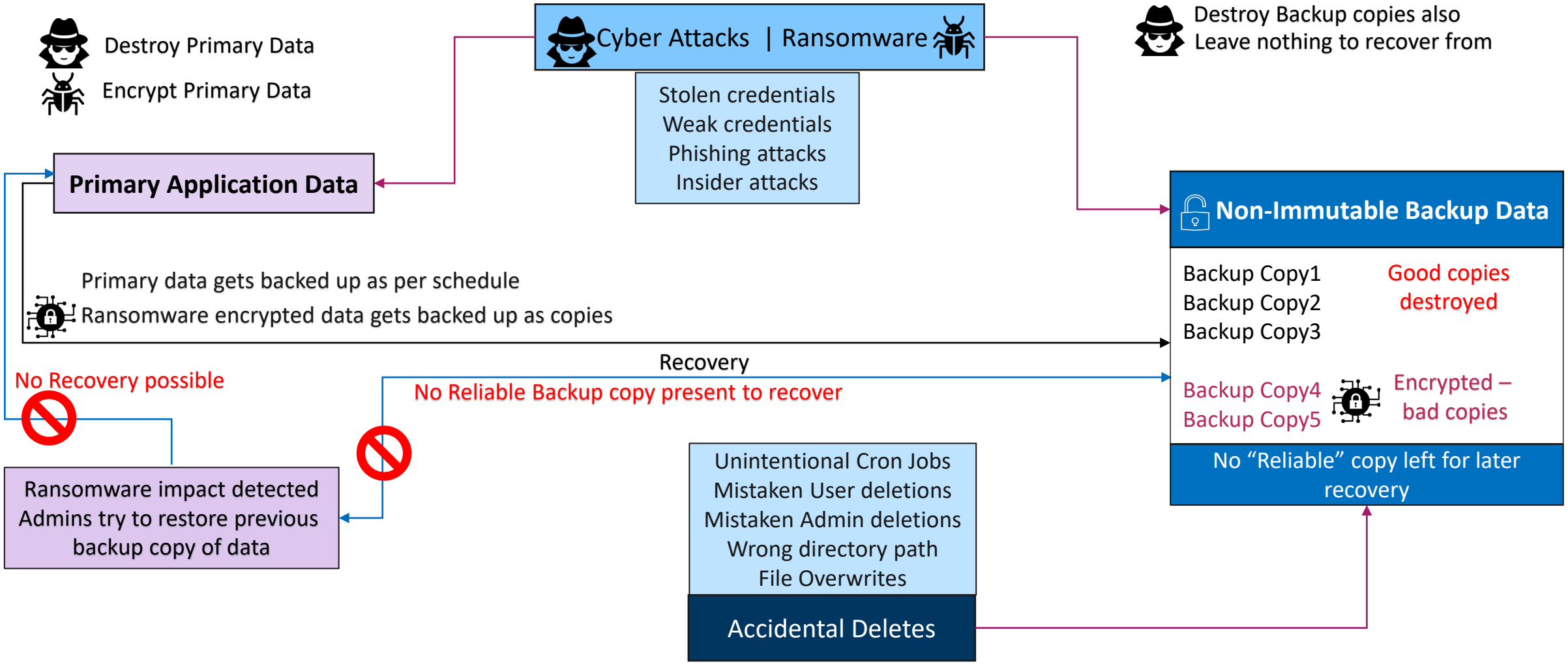
- Kind of a malware that creeps into the client systems
- Its attack model is to encrypt all the application/system data and ask for a significant fee to decrypt them

2022-2023

24% of all data breaches[1]
493 Million attacks[2]
(16 attacks/second)
Avg. $5.13 Million loss/attack[1]

1. Cost of Data Breach Report 2023 – Ponemone Institute and IBM Security
2. 2023 Sonicwall Cyber Threat Report

Cyber Attacks, Ransomware & Accidental Deletes



Ways to Make Data Immutable

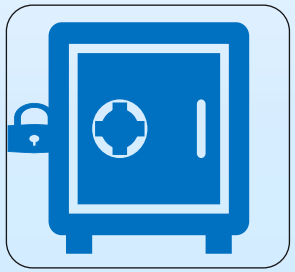
Make Read-Only (RO)

- RO data cannot be deleted or modified directly
- Still not enough protection from all threats
- Attackers can toggle RO mode on data and then destroy it
- No defined duration of protection

Retention Locking / WORM

- Data is allowed to be written only once
- No modifications or deletion until lock expires
- No way for attackers to toggle the lock mode
- They have to wait until the lock duration expires

Retention Locking Variants



Compliance Mode

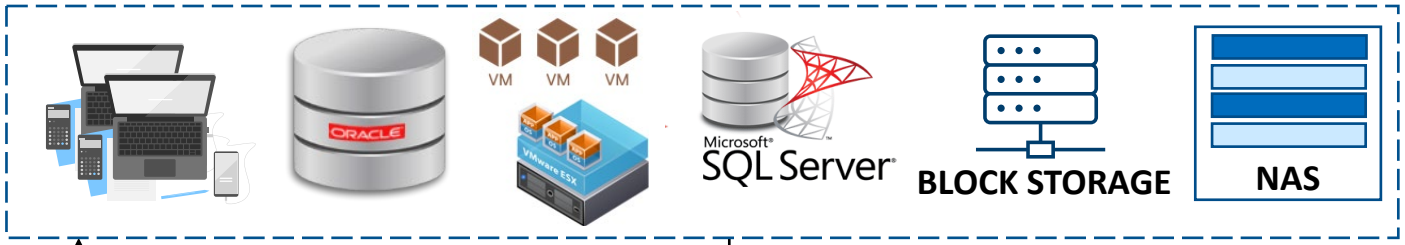
- Complaint with regulatory requirements like SEC 17f-4(a) and FINRA
- Stricter variant
- No lock reversal possible
- Enforces dual sign-on requirements
- Support for placing indefinite “legal hold” on the locked & expired data

Governance Mode

- Administration and Governance use cases within Organization
- Lenient Variant
- Admins can revert locks before expiry
- No dual authentication measures enforced
- Support for placing indefinite “legal hold” on the locked & expired data

Data Backup Flow – With Immutability

Assets



Primary data

Keeps on getting modified
Cannot be made Immutable

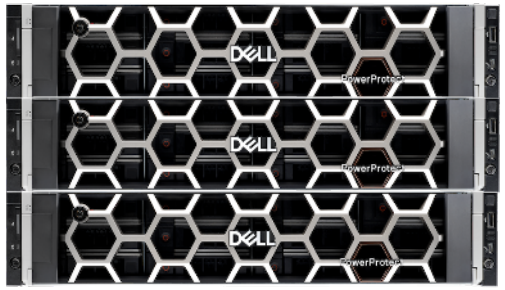
3. Add assets to be backed up to the backup application

5. Ingest data from client systems into the backup server

6. Utilize RL APIs/methods to lock newly ingested data in backup server

4. Create backup policies with RL enabled for the required duration

7. Cleanup old backup files after their locks expire



Backup data

Point in Time Copies
Never gets modified
Can be made Immutable

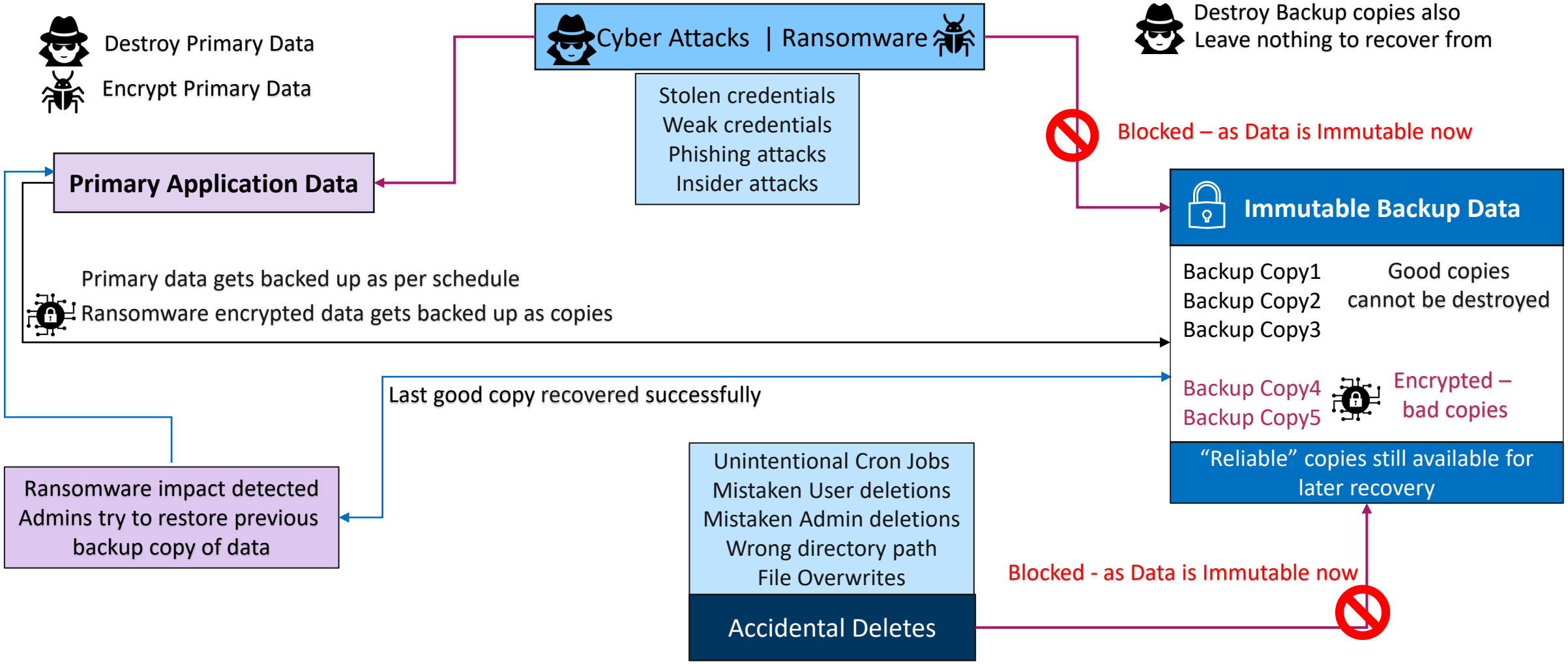
Integrated Backup Application

1. Integrate backup apps with the RL capability of backup servers

2. Add backup servers to the backup application

Backup Storage

Data Immutability - Protection Against Ransomware



Complete Data Immutability - Attack Vectors

Namespace and Beyond - Challenges & Best Practices

Physical Access to Data Center

Attackers gaining physical access to the datacenter that hosts the backup server is a major concern as well.

They can physically destroy the disks, shred them, or secure erase them

Such attackers are usually from within the organization and have seamless physical access

- Access is not refreshed periodically (revokes, grants)
- Absence of strict access guidelines in the organization
- Shared access between employees without any restrictions/roles

Shared Responsibility Mode

Customers need to ensure security & protection for the areas under their control

Grant physical access to datacenters on a need basis

Follow industry standard physical access guidelines

Dual Sign-on Model

- Requires two users System Admin & Security Officer (SO)
- SO credentials to be owned by a different individual in the organization
- To prevent data destruction by a single attacker
- Even stronger by - Multi-Factor-Authentication (MFA) for SO credentials
- Enforced by compliance variants

Namespace Level Protection – And Beyond

Filesystem Namespace Protection

- Namespace level protection via its Retention Lock capability
- Operations like file modify, delete, rename, resize, overwrite, truncate etc. are blocked
- Locked data cannot be tampered or deleted in any way via the namespace operations

Beyond Namespace?

- Attackers can exploit layers beyond the filesystem namespace to destroy retention locked data
- For example:
  - Clock & NTP
  - OS shell
  - Hypervisor
  - Boot loader
  - Platform management interface

Clock & NTP

Attack Vector

Can move the system clock forward and delete locked files prematurely before expiry

Can control the external NTP servers to manipulate date and time in the backup server

Mitigation

Restrict the frequency and amount of clock modifications and bring NTP configurations under Dual sign on model

Enable secure clock in the backup server software to detect clock skew

Restrict the amount of time skew that is allowed

Operating System

Attack Vector

Can enter #bash shell as root user and execute disk level destructive commands

No dual-sign on kind of protection available in operating systems

Mitigation

Use strong root user password or randomize it.

Prevent unlimited entry into the root shells

Enforce need for time bound unique token to access the OS root shell.

Hypervisors (Ex. ESXi, Hyper-V)

Attack Vector

Can enter the hypervisor console and perform destructive operations -

Delete virtual disks, delete virtual machines, corrupt physical disks holding the virtual disks etc.

Mitigation

Lockdown hypervisor console if supported

Block CLI, GUI, or REST API interfaces from outside access

Restrict Hypervisor console access

Bootloaders (Ex. GRUB, LILO)

Attack Vector

Can enter Single User Mode of OS and perform the destructive operations

Can exploit/misuse various disk management commands available in the bootloader console itself

Mitigation

Set randomized bootloader password

Prevent bootloader entry modifications,

Prevent bootloader console access

Provide access only via USB keys requiring when physical access is needed in the server

Platform Management Interfaces (Ex. iDRAC, ILO)

Attack Vector

Can enter the remote management interface (ex. IPMI, ILO, iDRAC etc.) and destroy disk volumes, disk groups, raid configs, initialize disks etc.

Mitigation

Disconnect management interfaces from the network so that physical presence is enforced

Randomize root user password

Disable platform management users by default (they can be enabled securely on need basis)

Advantage of Hyper-converged/Converged Appliances

All the components of a backup ecosystem are bundled into one single unit.

Hyper-converged/converged appliance vendors have additional control on more areas end-to-end and can hardened them effectively.

Ex. Dell Power Protect DM5500 Integrated Appliance

Backup Application
Power Protect Data Manager
RL Integrated Backup Application

Backup Server
Power Protect Data Domain Virtual Edition

Data Immutability via Retention Lock
Hardened Clock & NTP management
Secured OS Shell access
Special RLC Security Clock
Deduplication

Hardened GRUB Layer

(No GRUB console access)

Hardened Hypervisor layer

(Console and Interfaces Protected)

Hardened iDRAC Layer

(Remote Management interface restricted & users disabled by default)

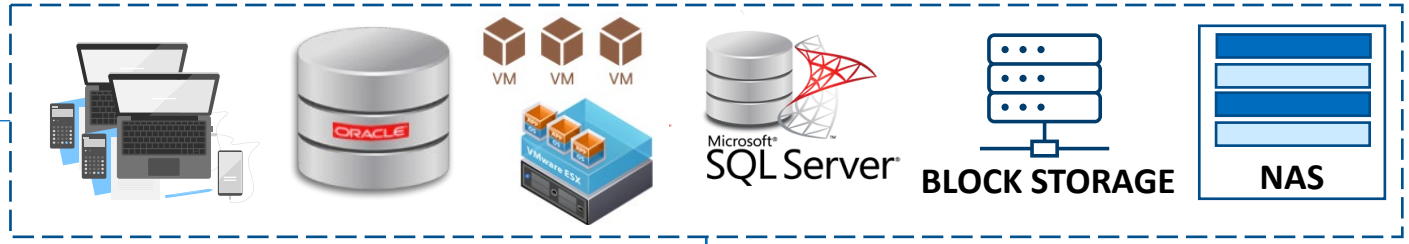
Dell Power Protect DM5500
Integrated Backup Appliance

Data Immutability in Action

Example use cases: Replication & Air-gapped Cyber Secure Vaults

Retention Locking in Replication Environment

Copy Data Management



Assets

Primary data get ingested as per schedule

Expiry : Dec 31, 2025
File1

1st Immutable Backup Copy

Retention Lock state of system and files also get replicated

Expiry : Dec 31, 2025
File1

2nd Immutable Backup Copy

New data get replicated regularly

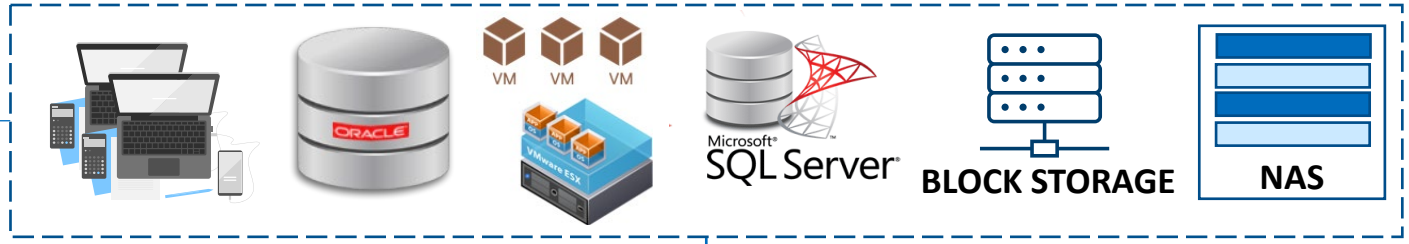
Replication policy between backup servers

Replication Source Backup Server

Replication Destination Backup Server

Data-protection via Air-gapped Cyber Secure Vaults

Copy Data Management



Assets

Primary data get ingested as per schedule

Automated Operational Air Gap
Gap is closed for a duration
to allow replication and then opened again

Replication Source Backup Server

AIR GAP

Replication policy between backup servers

Immutable Point in Time copies of backup data

CyberSense
AI/ML analysis to detect Ransomware infected data

Dell Power Protect Cyber Recovery

Dell Power Protect Data Domain Virtual Edition

Dell Power Protect Cyber Recovery Vault

Recovery path – From Restore points (PIT copies)

Please take a moment to rate this session.

Your feedback is important to us.

Auto Retention Lock (ARL)/Default Retention Lock

- After ingest, data gets auto-locked for a pre-configured duration
- Transforms from “Backup application controlled” to “Storage controlled” locking
- **Auto Retention Period**: Duration for which all new files would be auto-locked
- **Cooling Off Period (COP)**: No-modification duration after which files get auto-locked
- Non-integrated backup applications benefit the most

Legal Holds on Data

- Regulatory or Judicial asks to hold the compliance data until investigation is over.
- With a legal hold, retention lock expired data also cannot be deleted
- Legal holds stay until removed manually.
- Also called as Indefinite Retention Hold (IRH)