Reliable and Secure Storage

SIO

deti universidade de aveiro
departamento de eletrónica, telecomunicações e informática

João Paulo Barraca

Problems to solve

- Storage devices develop faults
  - It should be minimized the failures in storage devices and loss of data
  - Failure is certain and cannot be ignored
- Solid State Devices (SSDs) have a limited number of write operations
  - 2000-3000 writes per sector for MLC (2 bits per cell)
- Specific events may result in total data loss
  - Fire, robbery, “energy peaks”, floods, user mistakes, attacks

Problems to solve

- Access to storage disks is slower than Memory
  - Hard Disks: Access Time = Translation time + Rotation Time
  - Device Interface also limits bandwidth
- Data may be exposed to malicious actors
  - Which can access and or change data
- May be required to distribute data in an intelligent manner
  - To maximize performance
  - To reduce costs
  - To increase availability, confidentiality or integrity

Generic Solutions (some)

- Hardware Selection
- Infrastructure Selection
- Data Encryption
- Storage Redundancy
- Data Backups

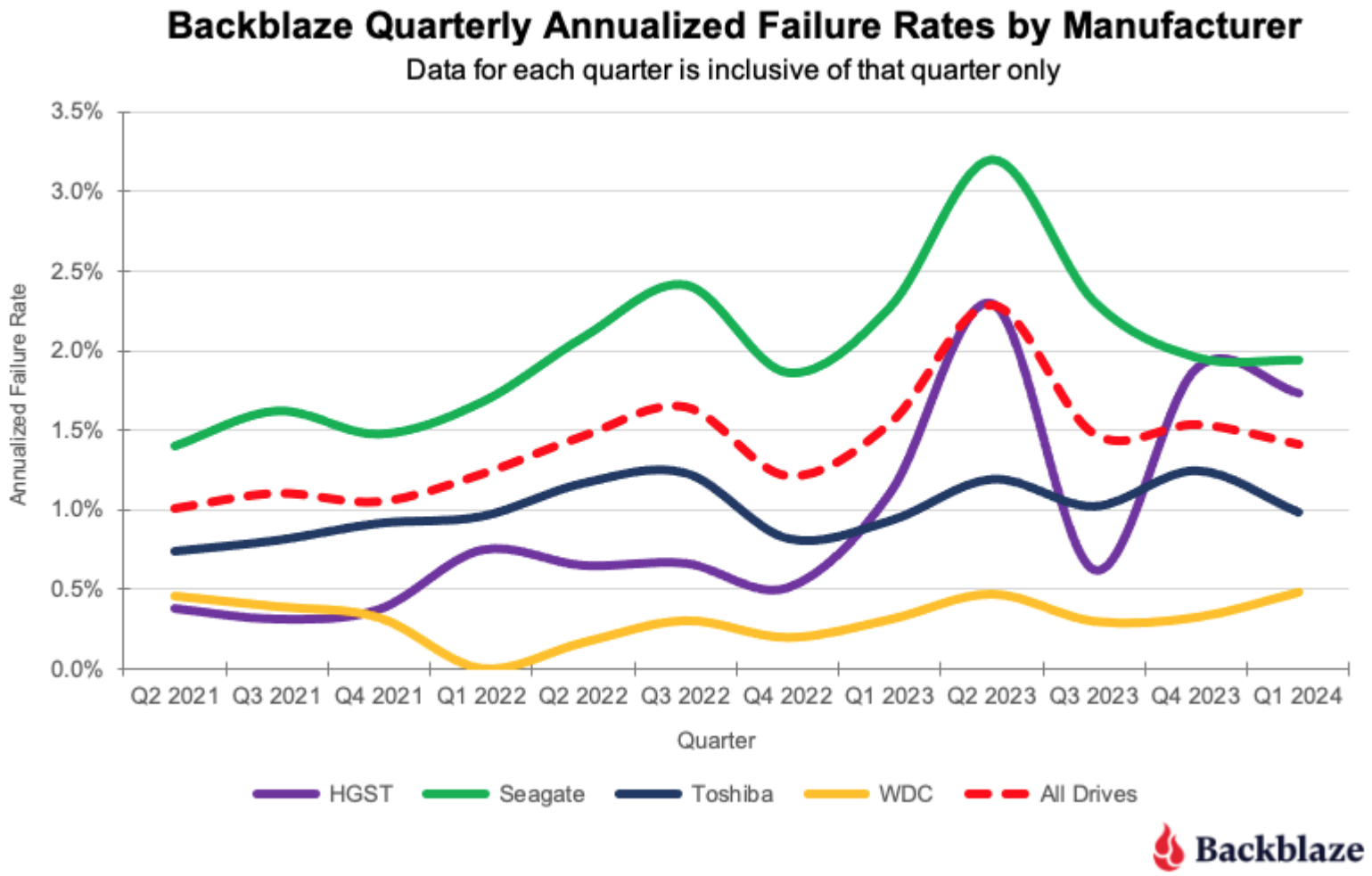
Hardware Selection

- Different device grades: Enterprise vs Desktop
  - Different MTBF: Mean Time Between Failures
    - Enterprise HDD: 1.2M hours, at 45°C, working 24/7, 100% use rate (1)
    - Desktop HDD: 700K hours, at 25°C, working 8/5, 10-20% use rate(1)

- Adjusted to each use case
  - Write intensive vs Read Intensive
  - NAS vs Video vs Desktop vs Cold Storage vs Data Center

- Differences in power consumption, reliability and performance
  - **Tier 0**: Highest performance, low capacity (PCIe NVME SLC SSD)
  - **Tier 1**: Some performance, high capacity and availability (M2 SATA SSD, SAS)
  - **Tier 3**: Low performance, high capacity, low price (SATA HDD)

Hardware Selection



Infrastructure Selection

- Storage Infrastructures require specific controls
  - Physical Access Measures: Walls, Server Cages
  - Access Control
  - Redundancy

- Certifications and Standards
  - Reliability: TIA-492
    - Grades Datacenters in 4 reliability tiers
  - Normative
    - Quality: ISO 9001
    - Security ISO 27001
    - Environmental: ISO 14001
    - Privacy: GDPR

Infrastructure Selection

- Information resilience can consider multiple strategies
  - Multi Site: locate systems in two locations
  - Multi Region: locate systems on multiple regions of the same provider
  - Multi Cloud: locate systems on two or more cloud providers

- Information types can limit provider selection
  - Personal Data: GDPR implies limitations on data movement over Europe
  - Health Data: Must be encrypted, access must be controlled, and risk assessed (HIPAA)
  - Financial Data: Strictly restricted based on multiple contexts: DORA, PCI DSS, Bank Secrecy

Infrastructure Selection

Fields	Tier I Basic	Tier II Redundant Components	Tier III Concurrently Maintainable	Tier IV Fault Tolerant
Number of Delivery Paths	Only 1	Only 1	1 Active 1 Passive	2 Active
Redundant Components	N	N + 1	N + 1	2(N + 1)S + S
Support Space to Raised Floor Ratio	20%	30%	80–90%	100%
Initial Watts/Ft	20–30	40–50	40–60	50–80
Ultimate Watts/Ft	29–30	40–50	100–150	150+
Raised Floor Height (Inch)	12″	18″	30–36″	30–36″
Floor Loading Pounds/Ft	85	100	150	150+
Utility Voltage	208,480 V	208,480 V	12–15 KV	12–15 KV
Months to Implement	3	3 to 6	15 to 20	15 to 20
Year First Deployed	1965	1970	1985	1995
Construction Rupees/Ft Raised Floor	₹40,000	₹40,000	₹40,000	₹40,000
Annual IT Downtime due to failure	28.8 h	22.0 h	1.6 h	0.4 h
Site Availability	99.671%	99.749%	99.982%	99.995%

Pomnar, Ashok & Rajawat, Anand & Tatkar, Nisha & Bhaladhare, Pawan. (2023). Sustainable Power Prediction and Demand for Hyperscale Datacenters in India. 124. 10.3390/engproc2023059124.

Data Encryption

- Data Encryption is vital for current storage strategies
  - **Encryption at Rest:** Data is never stored in clear text
    - Personal Information, Financial data, Health data
    - Mandatory in multiple regulatory contexts
  - Reduces Risk of malicious actors and direct data or hardware access
    - E.g. Actor tries to access data stolen on a stolen database
    - E.g. Actor tries to access data from a lost laptop

- How:
  - Data encrypted by databases engines, such as MySQL and Postgres
  - Encrypted Storage Volumes, such as Bitlocker and Luks
  - Filesystem features, such as in NTFS and EXT4
  - Application Encryption, such as 7ZIP and Joplin
  - Devices Encrypt data before writing to media

Data Encryption

Encrypted Volumes and Filesystems

- Popular solution supported in current systems with wide support
  - Mandatory by current security practices
  - Protects data from direct access to device
    - Lost/stolen storage device or laptop
  - File Based Encryption: encrypts files or folders
  - Full Disk Encryption: encrypts the entire storage volume

- Kernel component encrypts/decrypts data as required
  - **Linux**: EXT4 encrypted folders and the Linux Unified Key Setup (LUKS)
  - **Windows**: Bitlocker
  - **MacOS**: FileVault
  - **Android**: File Based Encryption (and FDE on Android
  - **iPhone**: Data Protection

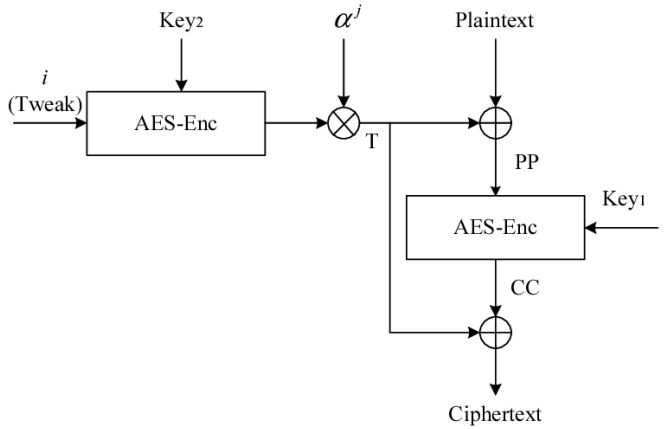
Data Encryption

Bitlocker

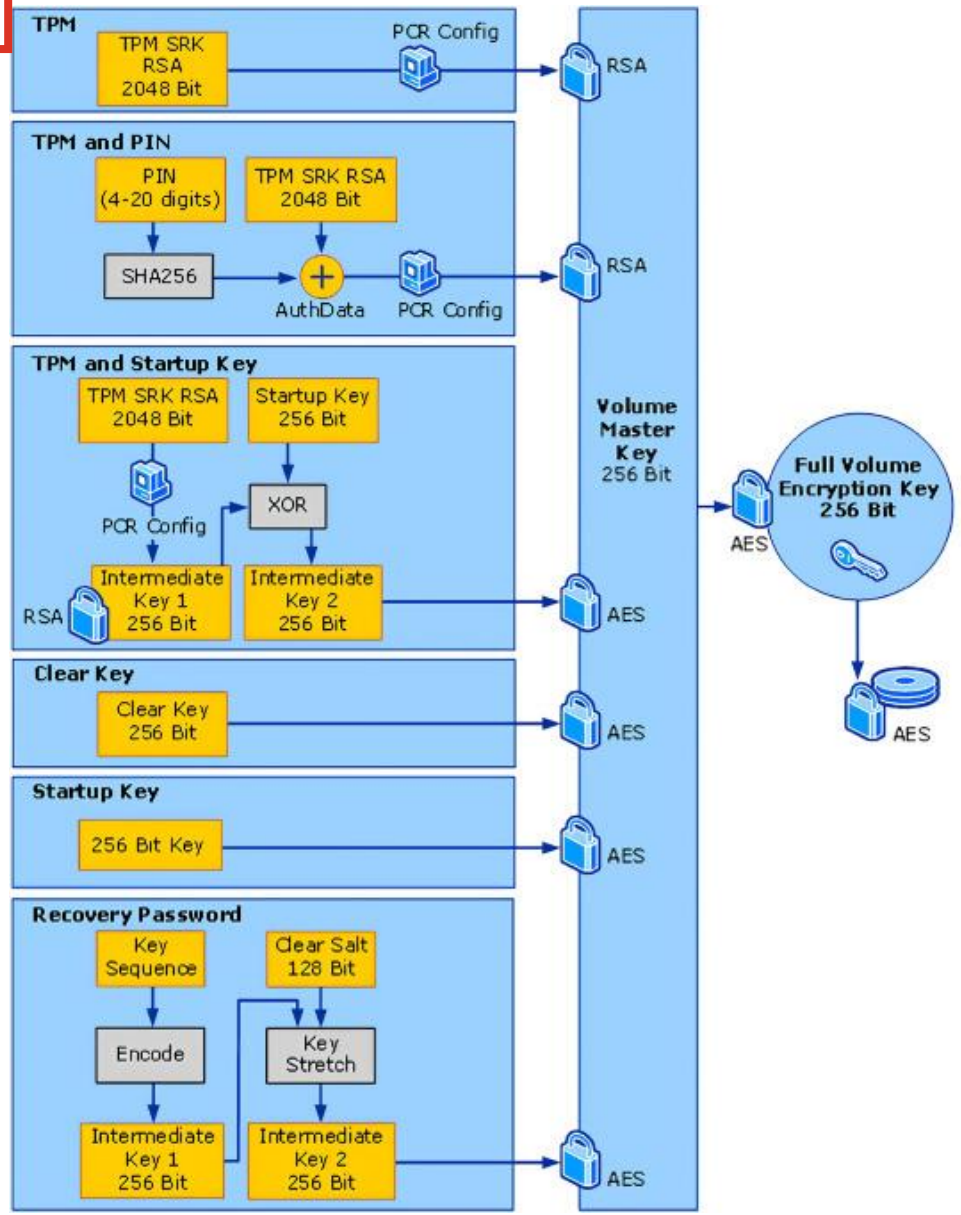
- Full Disk Encryption solution native in MS Windows Systems
  - Key can be on **Trusted Platform Module** (W11)
  - Key can be on Active Directory for management systems
  - Several other key options, some with MFA

- Encrypts Storage Blocks with 256 bit key
  - XTS-AES 256 for Internal block devices
  - CBC-AES 256 for USB devices

XTS Mode:



Default on W11

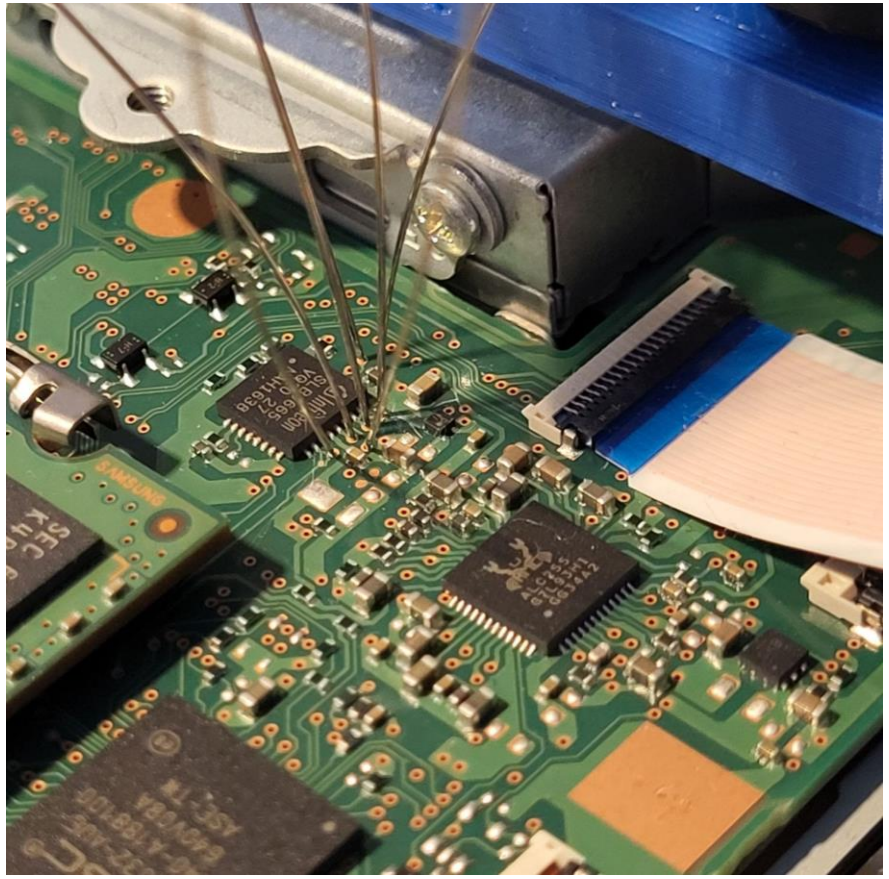


Data Encryption

Bitlocker

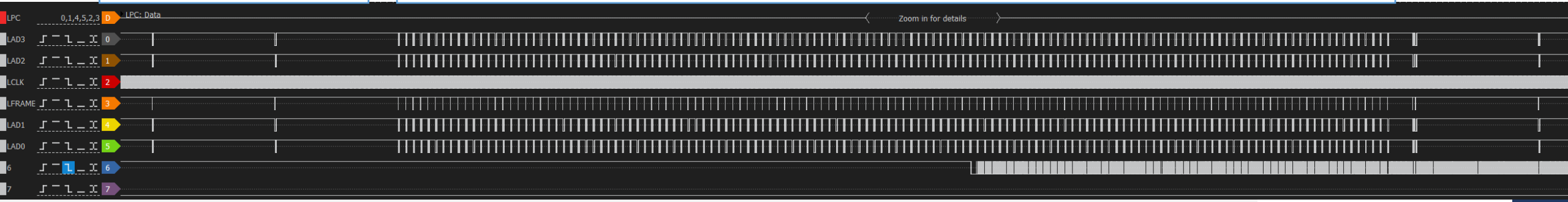
- Standard configuration is vulnerable to MITM
  - Key is stored in TPM and is provided to system on boot
  - Attacker can sniff TPM bus and get the key
  - Renders Bitlocker ineffective
  - Attack costs <10 € (RPI Pico + wires)

- Solutions
  - Firmware TPMs
  - Integrated TPMs
  - Use Boot PIN or Password: TPM will only provide key when unlocked



Other LPC Messages

TPM Transaction



Data Encryption

Self Encrypting Devices (SED)

- Devices have two distinct areas
  - Shadow Disk: Read-Only, ~100MB with software to unlock it
  - Real Disk: Read/Write. Contains user data

- Keys used
  - **KEK**: Key Encryption Key (Authentication Key)
    - Provided by the user. Digest stored in the Shadow Disk
  - **MEK** (or DEK): Media (Data) Encryption Key
    - Encrypted with the KEK

- Boot process
  - BIOS will access Shadow Disk and boots
  - Application in Shadow Disk requests password, decrypts KEK and verifies hash(KEK)
  - If it matches, MEK is decrypted, and disk geometry is updated

- Opal Storage Specification
  - TCG Storage Security Subsystem Class: Opal Specification

Storage Resilience

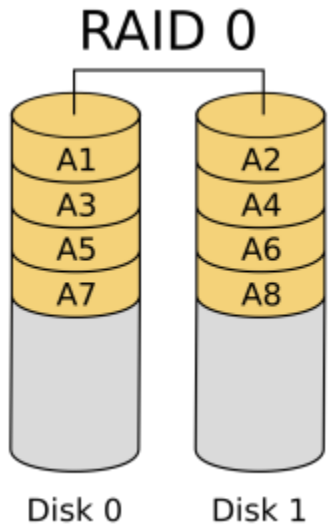
RAID: Redundant Array of Inexpensive Drives

- Improves the survivability of information
  - Data is only lost after several devices are lost
  - The number of lost devices is configurable
- Low cost and efficient solution
  - Can use cheap, lower quality hardware
  - Can improve read and write performance
- RAID doesn’t replace backups
  - Only tolerates the failure of a limited number of devices
  - Cannot cope with user mistakes (file modification/deletion)
- RAID can even increase the failure probability
  - As it can be tweaked towards performance

Storage Resilience

RAID 0 - Stripping

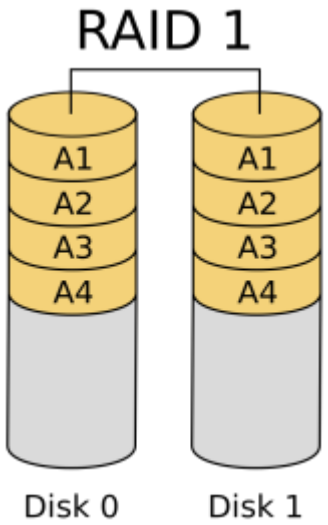
- Objectives
  - Speedup data access
- Approach
  - Access disks in parallel
  - Striping: Data is split in small chunks (stripes)
    - Stripes are stored among all disks in a distributed manner
- Advantages
  - May speedup performance as a factor of the number of disks
- Disadvantages
  - Increases the probability of loosing data
  - If Pf is the probability of failure of a single disk, an N-disk RAID 0 volume will have a 1-(1-Pf)N failure probability
  - Increases the number of devices
    - At least it will double the number of devices required



Storage Resilience

RAID 1 - Mirroring

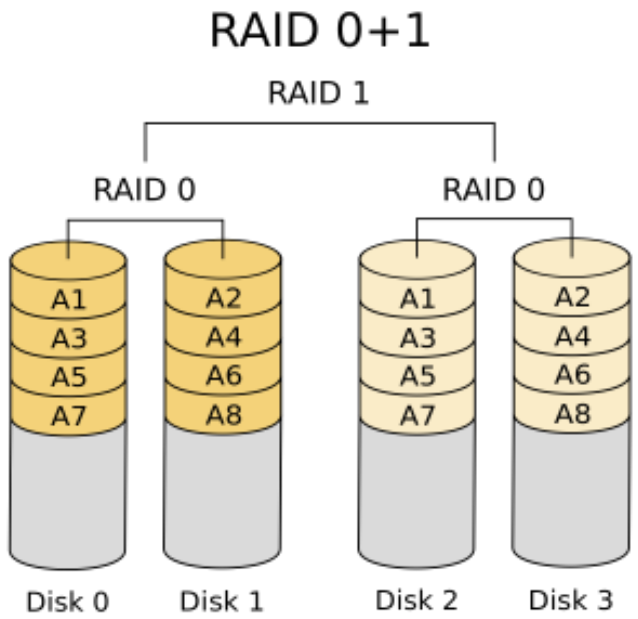
- Objectives
  - Tolerate disk failures
- Approach
  - Data duplication (mirroring)
    - Synchronized writing
    - Distributed read from any disk with or without comparison from another disk
- Advantages
  - Decreases the probability of data loss
    - If Pf is the probability of failure of a single disk, the probability of failure with N disks is Pf^N
- Disadvantages
  - Storage inefficiency: Will lose at lease 50% of the total capacity
    - For 3 disks it will lose 66%... Loss is (N-1)/N
  - Increase the number of devices
    - At least to the double
- Extremely common for the boot disks

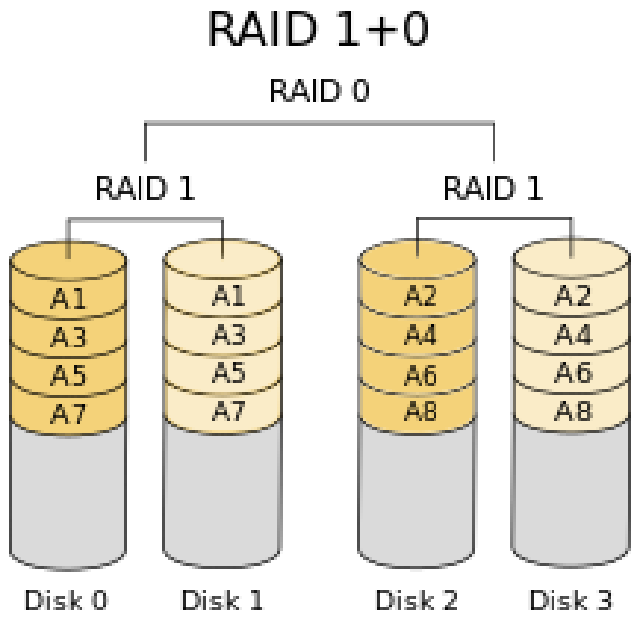


Storage Resilience

RAID 0+1 and RAID 1+0 – Nested RAID

- Objectives
  - Benefits of RAID 0 (performance)
  - Benefits of RAID 1 (resilience)
- Approach
  - 0+1: A RAID 1 volume using RAID 0 volumes
    - Mirroring of striped volumes
  - 1+0: RAID 0 over RAID 1 volumes
    - Striping over mirrored volumes
- Disadvantages
  - Storage capacity waste
    - At least 50%
  - Increase the number of devices





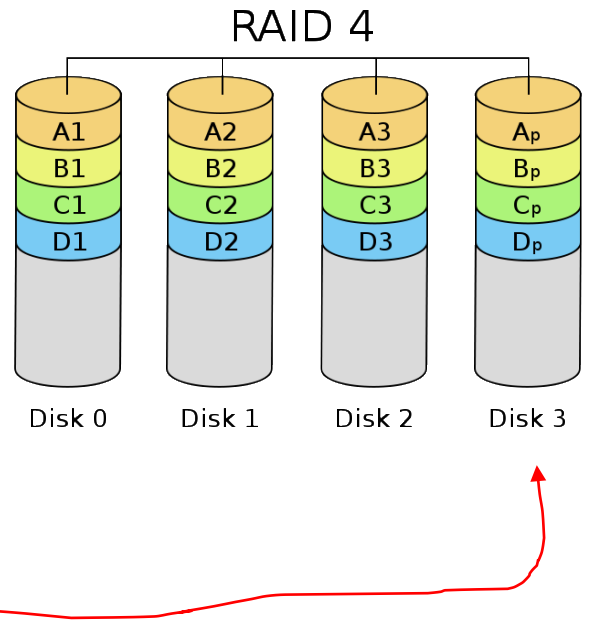
Storage Resilience

RAID 4

- Objectives
  - Have some resilience as RAID 1
  - With a performance close to RAID 0

- Approach
  - Store data in N-1 disks
  - Store parity data in an additional disk
    - Total waste is dependent on the capacity and number of disks
    - Data from any N-1 disk can be used to recreate another one

- Disadvantages
  - Requires at least 3 disks
    - Updating parity data is complex and will require specific hardware
    - Imposes the need to read before any write
      - Read data from existing block (e.g., C1) and from the corresponding parity disk (Cp)
      - Compare old data block with new, and change the parity block (Cp')
      - Write the new data block (C1') and the new parity block (Cp')
    - Writes must be serialized due to the existence of a parity disk
  - Recovery is way more complex and slow than with RAID 1

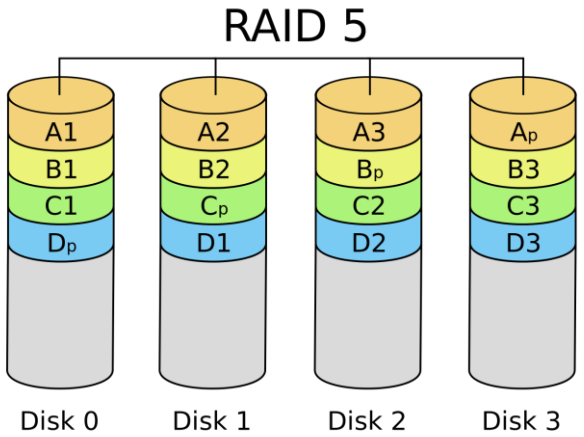


Storage Resilience

Ranking menos desperdicio:
1. RAID 0 (0%)
2. RAID 5 (1 disco)
3. RAID 4 (1 disco)
4. RAID 6 (2 discos)
5. RAID 1 (50%)
6. RAID 0+1 e 1+0 (50%)

RAID 5

- Objectives
  - Similar to RAID 4
  - But with higher write efficiency



- Approach
  - Distribute the parity blocks among all disks
  - Waste is similar to RAID 4
  - Write concurrency is improved

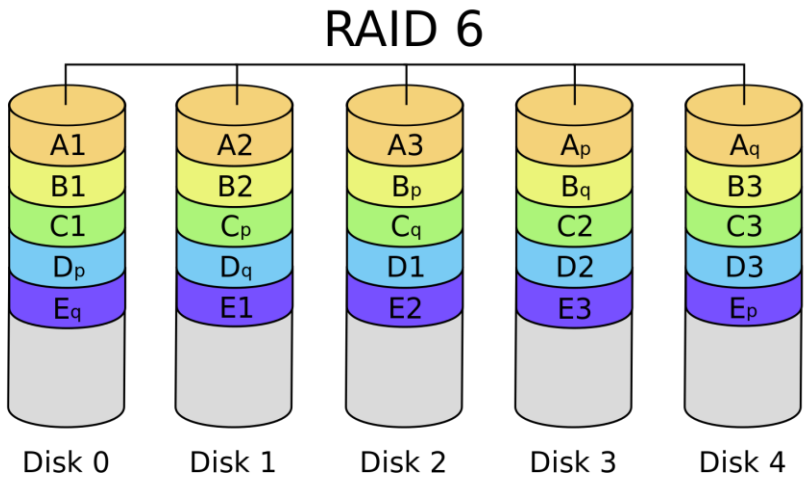
- Disadvantages
  - More complex to be implemented, typically requiring dedicated hardware

- Extremely common in servers

Storage Resilience

RAID 6

- Objectives
  - Improve the reliability of RAID 5
- Approach
  - Use 2 parity blocks, distributed among all disks
  - Capacity waste will be higher than in RAID 5 (equal to 2 disks)
  - Concurrency is slightly worse than with RAID 5
- Advantages
  - Allows the failure of two disks without data loss
- Disadvantages
  - Even more complex than RAID 5



Backups

- Periodic copy of data
  - Snapshot of the storage state in a specific moment
  - Copies will allow to set files to a previous version
  - May have additional integrity and confidentiality controls

- Data is only a backup if it constitutes an archival copy of other data
  - There are at least 2 locations with the same data

- Basic and Essential mechanism for Disaster Recovery
  - Copy data from backups into systems

Backups

Desired Properties

- Availability
- Redundancy
- Deduplication
- Data protection
- Immutability

Backups

Properties: Availability

- The ability to quickly restore and access that when needed.
  - Not a simple measure of the backup system availability (e.g. uptime)

- **Recovery Time Objective (RTO):** Maximum amount of time to restore a system or data after a failure.
  - A lower RTO indicates faster recovery and higher availability.

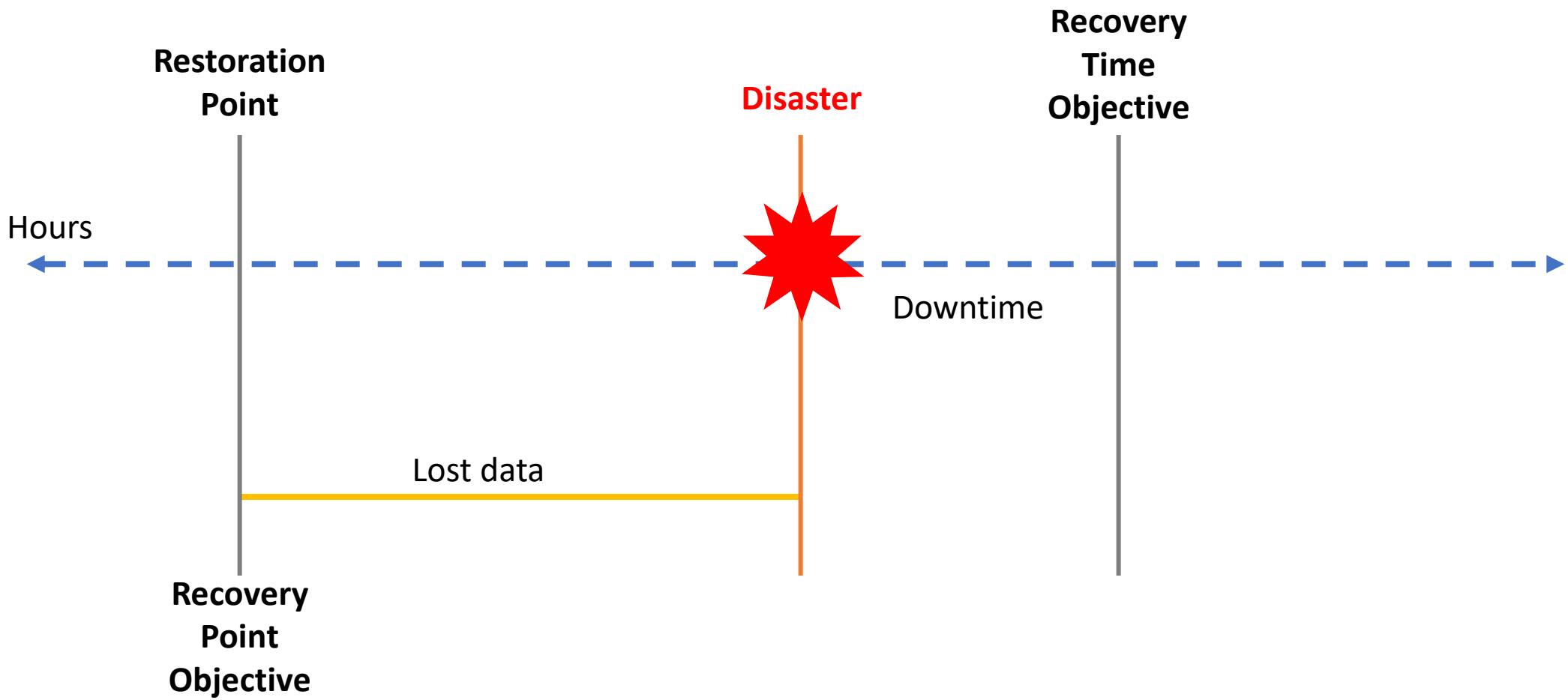
- **Recovery Point Objective (RPO):** Maximum amount of data that can be lost during a recovery process.
  - A lower RPO means less data is at risk of being lost.

- **Backup Frequency**: The time interval between two copies
  - More frequent backups generally lead to higher availability, as there are more recent versions of data to restore from.

- **Backup Storage and Retrieval**: The speed and ease of retrieving backups from storage solution

Backups

Properties: Availability



Backups

Redundancy

- How and how much are data objects duplicated in a backup strategy

- **Number**: number of copies of the same object
  - Must be at least 1 (by definition)

- **Diversity**: Which media and locations are used for backups
  - Media: tape, hard disk, SSD optical media
  - Location: same system, same rack, same infrastructure, remote

- **Retention**: How far in time can we go
  - How many copies of the same object over time
  - Allows capture changes to an object

Backups

Deduplication

- How systems optimize storage to minimize data storage requirements for similar data

- Backup systems scrub data looking for duplicated items
  - Files, Block Sectors, chunks inside files
  - As long as redundancy criteria are met, additional data is deleted and linked
  - Many files are duplicated
    - Same file over time without large changes
    - Multiple VMs with the same OS

- **Deduplication ratio**: Total backup size / deduplicated size
  - Media and compressed files had low ratio
  - Storage management can improve ratio
    - Keep similar files on the same backup system

Backups

Deduplication

			Existing Files			New Files		
Backup#	Type	Comp Level	Size/MB	Comp/MB	Comp	Size/MB	Comp/MB	Comp
657	full	3	7360.4	6244.5	15.2%	46.9	9.4	80.0%
658	incr	3	40.0	9.0	77.6%	7.6	1.7	76.9%
659	incr	3	32.1	8.6	73.1%	7.4	1.7	77.3%
660	incr	3	12.1	3.2	74.0%	40.1	9.0	77.6%
661	incr	3	40.0	8.3	79.4%	7.4	1.7	76.7%
662	incr	3	40.0	8.8	77.9%	7.5	1.7	76.8%
663	incr	3	40.2	8.3	79.3%	7.3	1.7	77.2%
664	incr	3	46.0	12.3	73.2%	7.4	1.7	77.1%
665	incr	3	1.2	0.4	68.2%	50.2	10.5	79.2%
666	incr	3	38.0	9.1	76.0%	7.6	1.9	74.8%
667	incr	3	9.2	1.2	86.5%	38.5	8.4	78.2%
668	incr	3	34.0	7.2	78.9%	13.8	3.4	75.4%
669	full	3	7396.8	6251.1	15.5%	11.2	2.9	74.5%
670	incr	3	27.0	6.5	76.0%	8.0	2.0	75.7%

```
$ du -hs 669
6.2G    669
$ du -hs 657
6.2G    657
```

```
$ du -hs 669 657
6.2G    669
106M    657
6.3G    total
```

Deduplication in makes backup 657 to mostly contain hard links.
Filesystem presents two files, but storage space is shared

Backups

Data Protection

- Backups contain sensitive data
  - Personal data, system logs, financial data, health data
  - Access to backups must be restricted according to data sensitivity
  - Atackers may go after backups as they may be easier targets

- Best practices
  - Data is compressed, and encrypted at the source (Encrypt Data at Rest)
    - Backup keys are never stored in the backup data
  - Backups are made through secure interfaces: VPNs (Encrypt Data in Transit) and dedicated links
  - Backup location is carefully selected
    - external provider location and jurisdiction
    - Physical access controls
  - Air Gaps
  - Audit logs to help validate improper access to backups
  - Backups are verified

- Data retention also applies to backups
  - GDPR Right to be Forgotten

Backups

Immutability

- Backup data cannot be modified
  - Or delete altogether
  - This is a typical action for malicious actors (e.g. Ransomware)

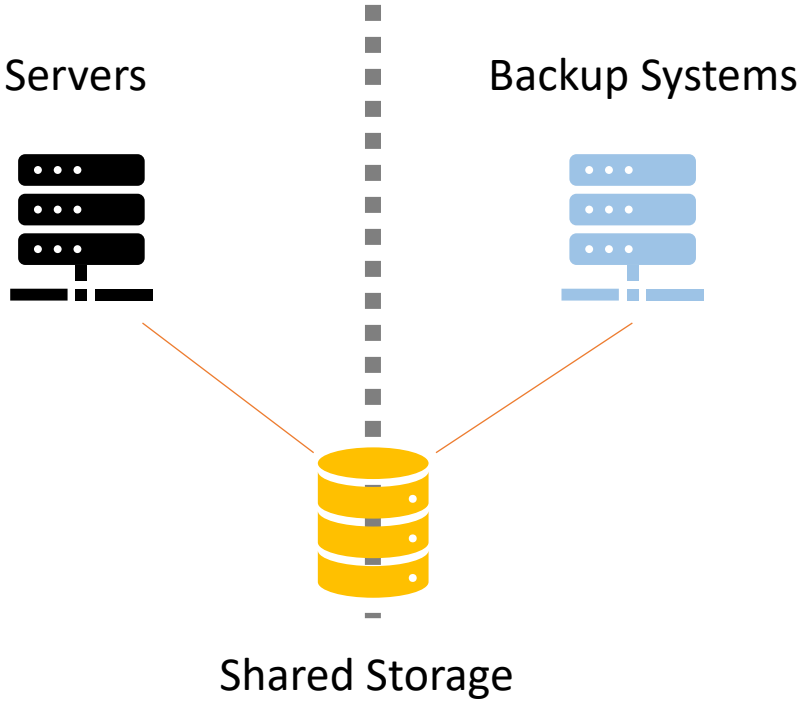
- Write Once Read Many Storage (WORM): destination storage only allows appending data
  - Writes or deletions are denied
  - Frequently: access to data requires special set of Administration Keys
  - Storage device acts as a secure storage
    - and not a generic storage allowing full control

- Air Gaps: Storage is isolated from client network
  - Logical Airgap: Backup system is isolated using a highly restrictive firewall
  - Physical Airgap: Backup media is offline and not available
    - May imply transporting media to different locations

- Blockchain: Use distributed ledger technology to lock data
  - Data is grouped in blocks
  - Blocks are locked using hash chains

Backups

Immutability – Air Gap using Shared Storage

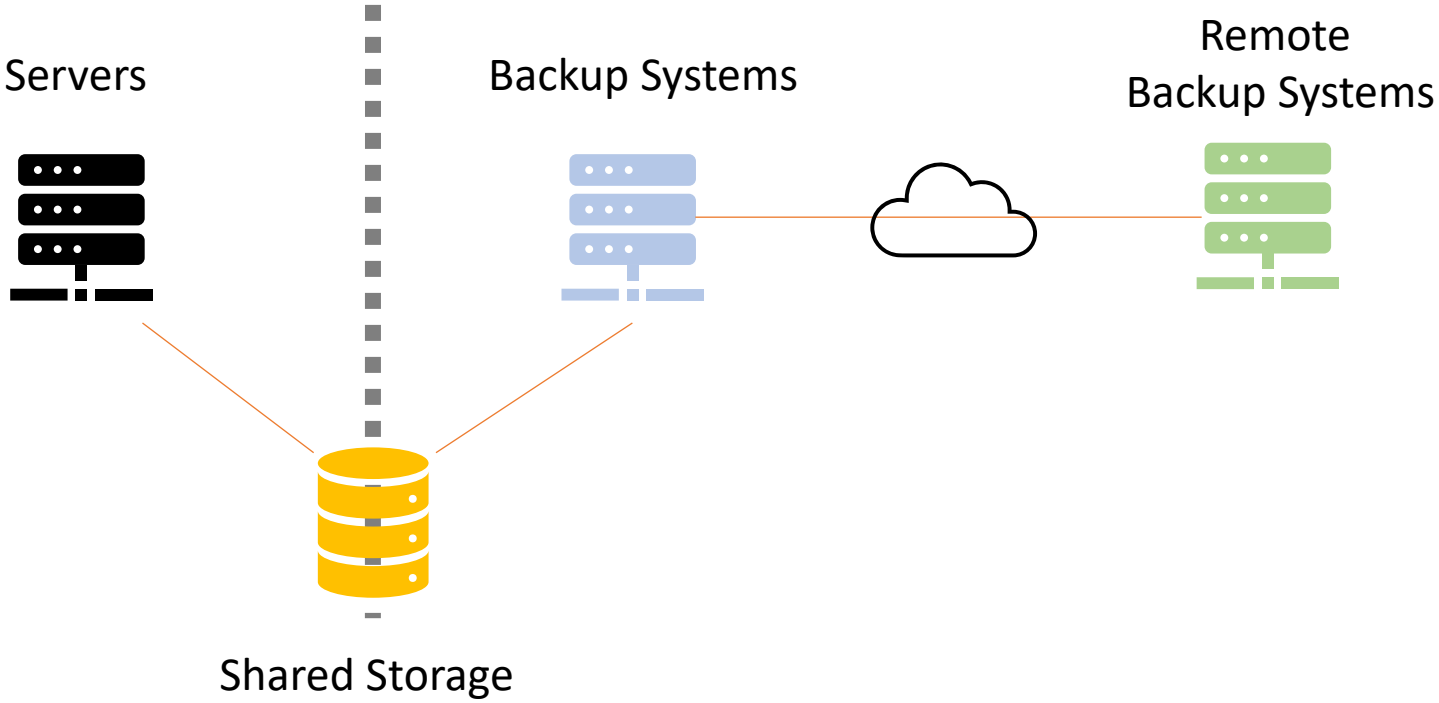


Both the Servers and Backup Systems can access data, but cannot communicate

Backup Systems are protected from compromised server

Backups

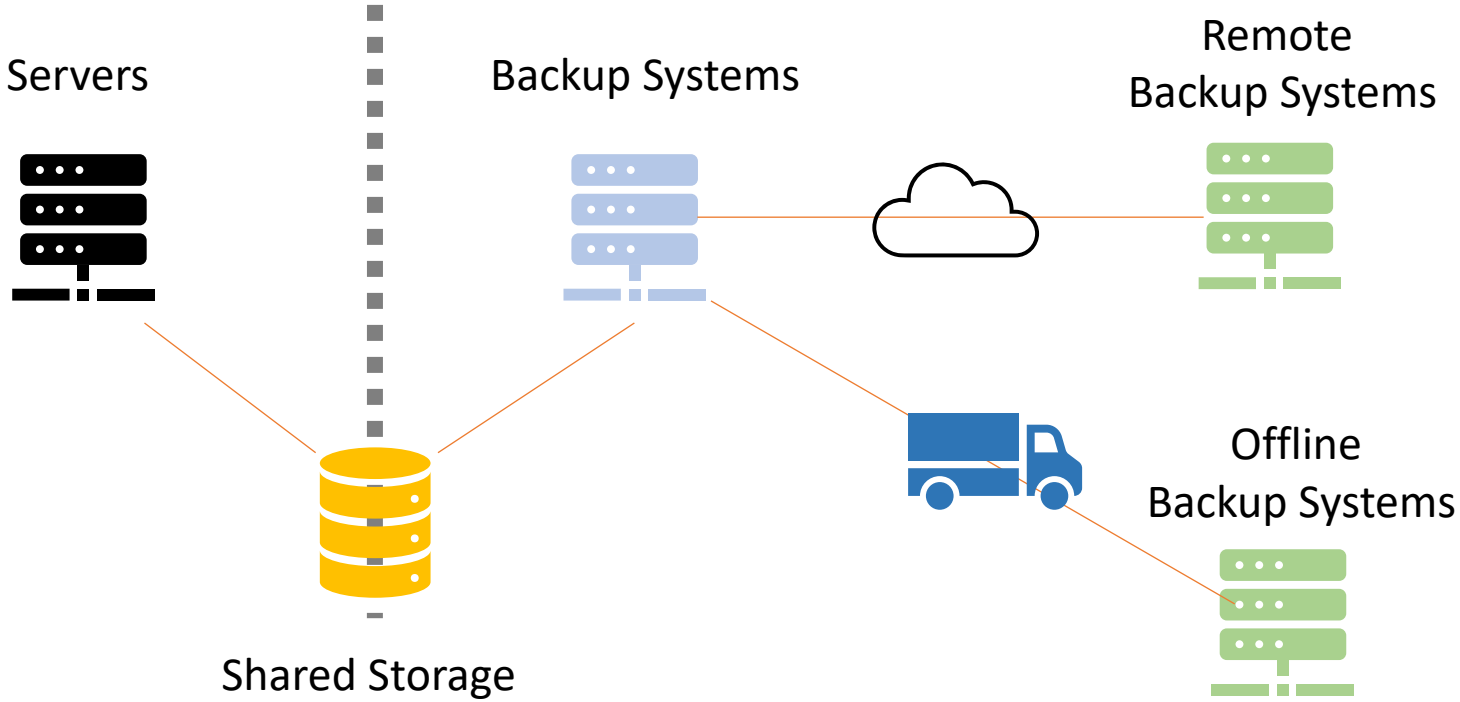
Immutability – Air Gap using Secondary Replication



Both the Servers and Backup Systems can access data, but cannot communicate
Backup Systems are protected from compromised server
Remote Backup System is further isolated

Backups

Immutability – Air Gap using Secondary Replication and Manual Transport



Both the Servers and Backup Systems can access data, but cannot communicate
Backup Systems are protected from compromised server
Remote Backup System is further isolated

Backups

The 3-2-1 Rule

- Keep 3 copies of any important file: 1 primary and 2 backups.
- Keep the files on 2 different media types
  - to protect against different types of hazards.
- Store 1 copy offsite
  - e.g., outside your home or business facility)
- Limitations
  - The notion of offsite was considered as a cloud, but systems already are in the cloud
  - The notion of "different media" considered HDDs and tapes, but there are other options
  - What about backup verification? How many errors can we have?

Backups

The 3-2-1-1-0 Rule

- Keep 3 copies of an important file
  - Observe the consideration of “important”
- Keep the files on 2 different media types
  - Any media
- Keep 1 copy at an alternative location (alternative cloud)
  - Alternative infrastructure
- Keep 1 copy offline
  - system is disconnected when not doing backups
- Verify backups to keep then with 0 errors

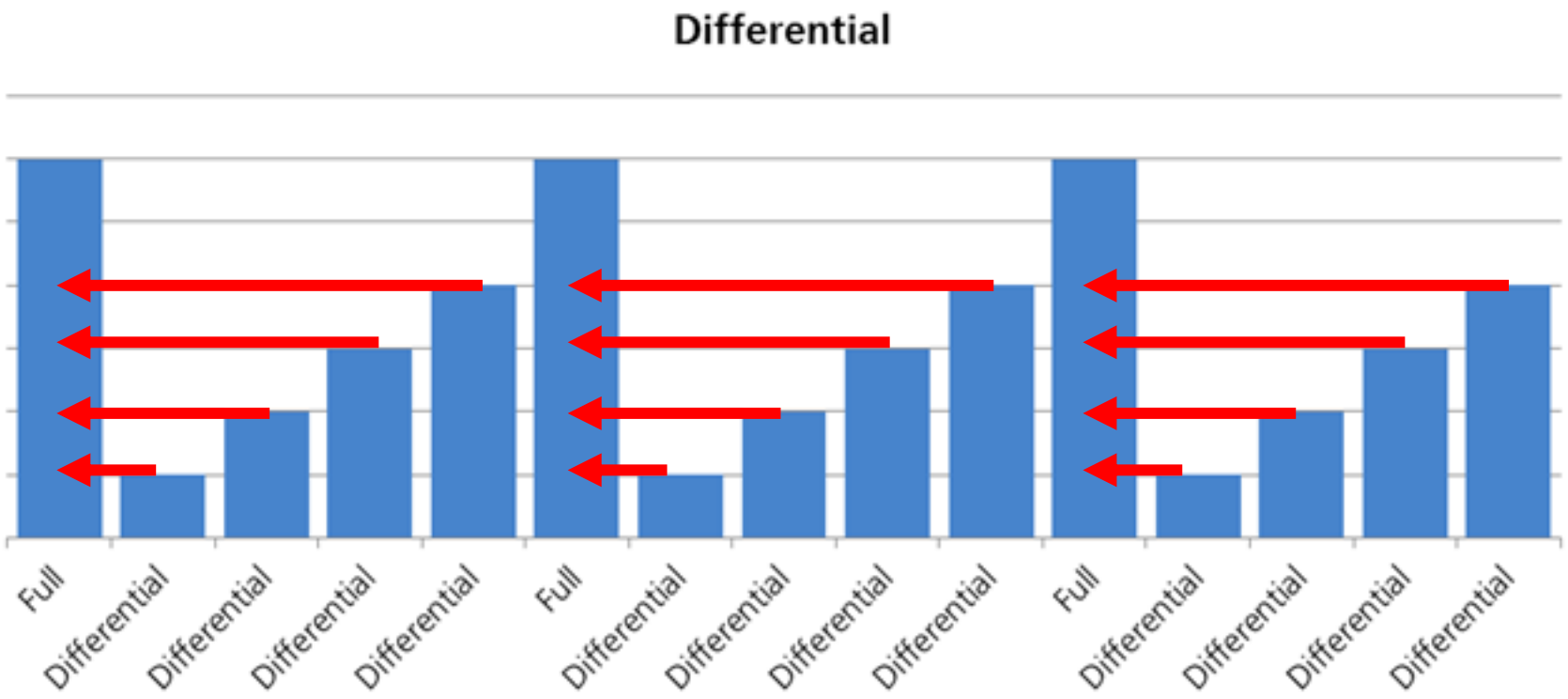
Backups

- Periodic copy of data
  - Snapshot of the storage state in a specific moment
  - Copies will allow to set files to a previous version
  - May have additional integrity and confidentiality controls

- Strategies in relation to data selection
  - Full: Complete snapshot of the data volume
    - Fast recovery
    - Requires a large amount of space
  - Differential: Differences since the last full backup
    - Daily differential backups will grow as changes increase
  - Incremental: Differences since the last backup
    - Higher storage space efficiency

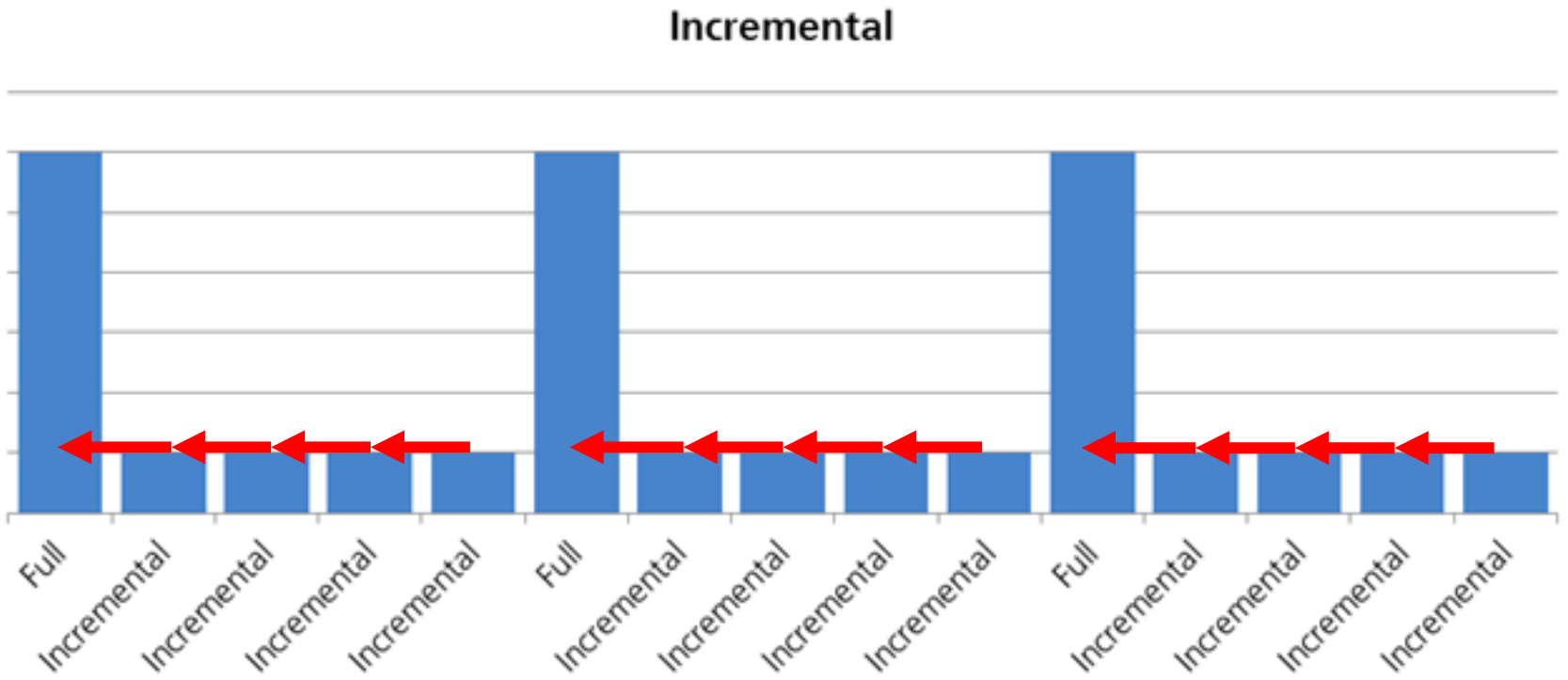
Mais complexa para restaurar

Backups types: Differential



http://www.teammead.co.uk/

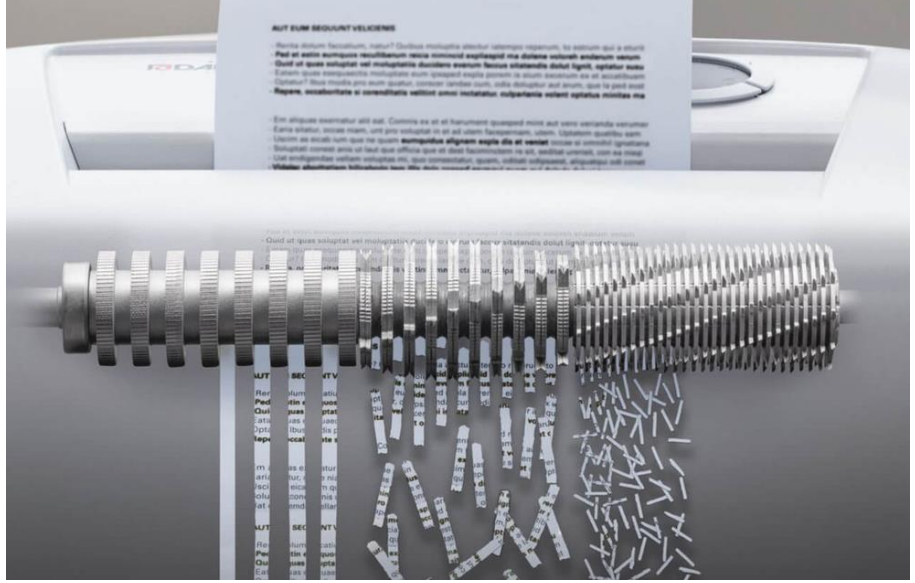
Backups types: Incremental



http://www.teammead.co.uk/

Data Destruction

- Secure storage requires a data destruction program
  - Prevents leaking critical information to actors
    - Dumpster Diving Attack: https://threatpost.com/hackers-dumpster-dive-covid-19-relief-scams/155537/
  - Standard practice in classified and even business information
    - Paper shredders
  - May be part of standard operations in IT



- Aspects to consider
  - Effectiveness of the method according to the media
  - Audit Trail created
  - Standard compliance
    - NIS 800-88 r1
    - IEEE 2883-2022
  - Documentation: Destruction certificate

	REUSE	RECYCLE		
	ERASE	DEGAUSS	CRUSH	SHRED
MAGNETIC DRIVES	✓	✓	✓	✓
SOLID STATE DRIVES	✓	✗	✓	✓
TAPES	✗	✓	✗	✓
BENEFIT	Financial Return	Quick and Cost Effective	Quick Visual Confirmation	High Volume Visual Confirmation