Storage

Problems

Storage devices develop faults

- It should be minimized the failures in storage devices and loss of data
- Failure is certain and cannot be ignored

Access to mechanical disks is slow (hard disks)

- Access Time = Translation time + Rotation Time
- More information -> higher impact of storage media

Problems

Solid State Devices (SSDs) have a limited number of write operations

- 2000-3000 writes per sector for MLC

Specific events may result in total data loss

- Fire, robbery, “energy peaks”, floods, user mistakes, attacks

May be required to distribute data in an intelligent manner

- To maximize performance
- To reduce costs

Solutions

Data backups
- Local
- Remote

Redundant Storage
- RAID
- Others: ZFS

Better storage devices, environments with higher control
- SLED (Single Large Expensive Disks)
- Enterprise Grade devices
- Temperature and Humidity Control

Infrastructures dedicated for storage
- Single policy control point

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 be encrypted

Full: Complete snapshot of the data volume

- Fast recovery
- Requires a large amount of space

Differential: Differences since the last full backup

- Slower recovery, but also lower storage requirements
- Daily differential backups will grow as changes increase

Incremental: Differences since the last backup

- Even slower recovery
  - Requires reconstruction of all intermediate backups since the last full
- Higher storage space efficiency

Backups

A backup is not an additional disk with data

- External or remote

It considers policies, mechanisms and processes to make, maintain and recover copies of the same data

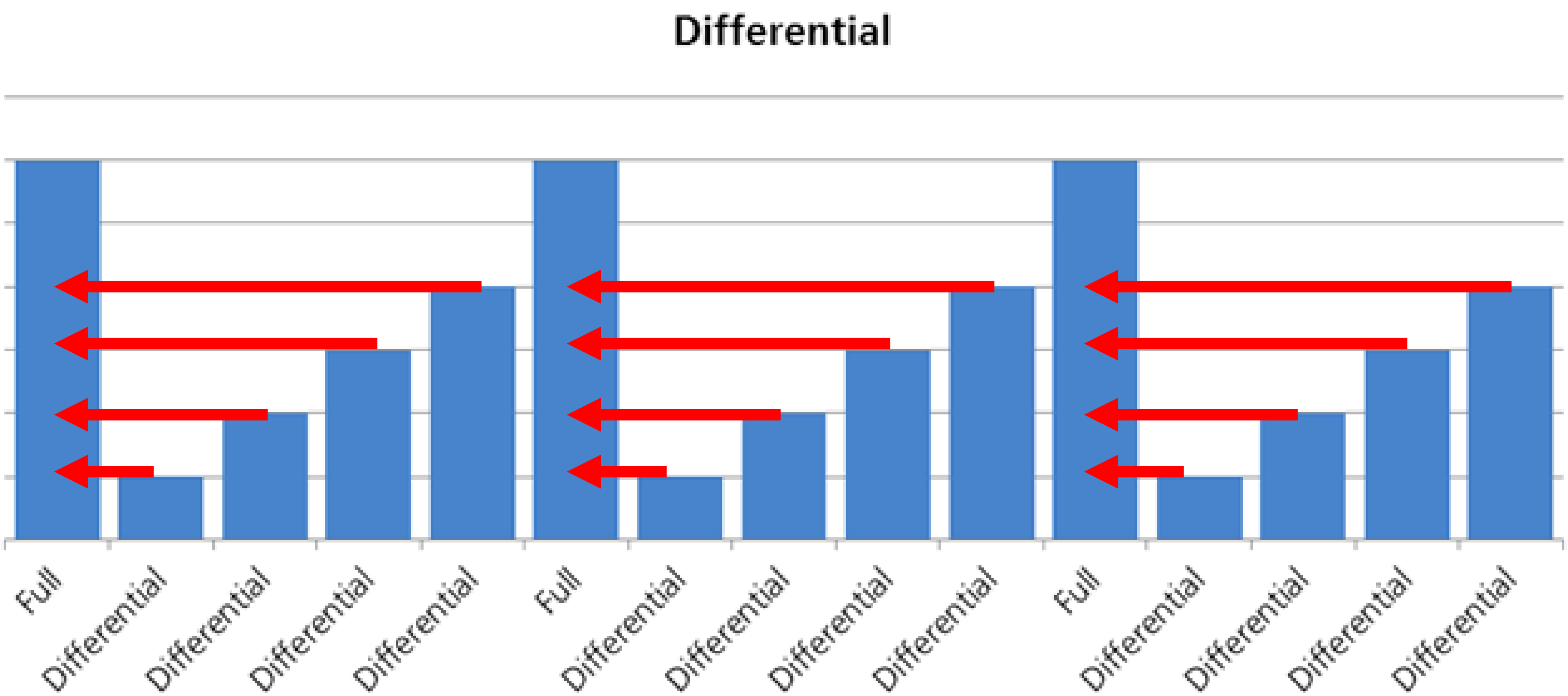
- Should resist specific situations
- Should be used only in emergency situations
- Important to consider both the copy, storage and recovery!

Legal framework implies a special care

- When dealing with personal data
- Frequently impose a retention policy
  - Backups should expire after some time

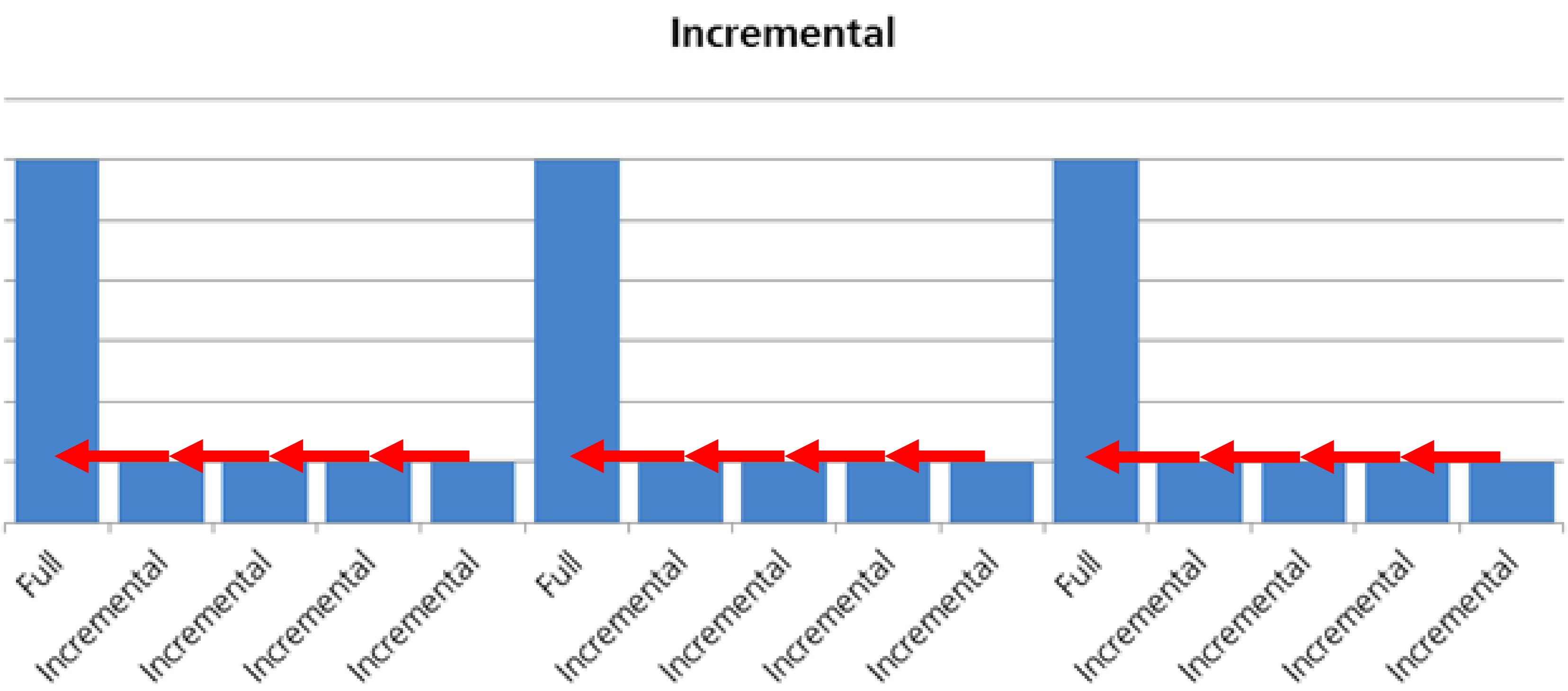
Backup Types: Differential

Backups types: Differential



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

Backups types: Incremental



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

Backups types: Incremental

		Totals			Existing Files		New Files	
Backup#	Type	#Files	Size/MB	MB/sec	#Files	Size/MB	#Files	Size/MB
657	full	143905	7407.3	2.07	143870	7360.4	59	46.9
658	incr	47	47.6	0.03	33	40.0	29	7.6
659	incr	153	39.5	0.02	132	32.1	36	7.4
660	incr	118	52.2	0.03	78	12.1	70	40.1
661	incr	47	47.4	0.02	32	40.0	32	7.4
662	incr	47	47.5	0.02	33	40.0	29	7.5
663	incr	47	47.5	0.01	33	40.2	29	7.3
664	incr	232	53.3	0.03	211	46.0	36	7.4
665	incr	91	51.4	0.05	35	1.2	85	50.2
666	incr	89	45.7	0.05	71	38.0	37	7.6
667	incr	47	47.7	0.02	18	9.2	44	38.5
668	incr	47	47.8	0.02	21	34.0	41	13.8
669	full	143937	7407.8	3.05	143824	7396.8	185	11.2
670	incr	95	35.0	0.04	68	27.0	54	8.0

Backups: Compression

Uses lossless compression algorithms and solutions

- Ex: ZIP

Copy only some parts of the information

- Only modified files

Deduplication

- Only store unique files/blocks
- Usually using full copy with offline deduplication
  - Of disk blocks using specific image formats
  - Of files using hard links

Backups: Compression and 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
```

du ignores duplicated hard links

Backups: Levels

Applications

- Extract data from applications (e.g. mysqldump)
- Represent a consistente view of the application
  - May be required to block the application state (e.g. Database changes)
- May be repeated for each individual application

Files

- Copy of individual files
- May backup any application in a filesystem
- State may be inconsistente
  - E.g. open files without data written, or applications change change many files at once

Backups: Levels

Filesystem

- Internal features provides by each individual filesystem
- Creation of periodic snapshots with record of all changes or current state
- May allow the recovery of individual files, or the entire filesystem

Device Blocks

- Copy of all blocks of a storage médium
- Independent of the filesystem or operation system in use
- May be implemented by the storage infrastructure
  - Transparent and without any impact to applications

Backups: Location of data

In the same volume or in the same server

- Allow users to rapidly recover information
- Protects against changes/deletions made by users
- May not protect against hardware malfunction
  - E.g. macOS Timemachine

In a system location in the same infrastructure

- Also with fast access time
- Protects against isolated storage failures
- Doesn’t protect data against events with broader reach
  - Floods, fire, robbery
- Examples: Most enterprise storage solutions, backuppc, TimeCapsule, Borg, Kopia

Backups: Location of data

Remote (off-site)

- Implemented to a system outside the local datacenter
  - Dedicated service or through the internet
    - E.g. Amazon S3, or to servers in a dedicated datacenter
    - Encryption if recommended (or mandatory) in the case of external services!
- Implemented with specialized secure transport
  - Armored car transporting backups to a secure place
- Allow recovery even if far reaching events occur
  - Terrorism, Earthquake
- Recovery will be slower
  - Limited by the speed of a network link or the physical transport

Selecting Storage Devices

There are diferent device grades: Enterprise vs Desktop

- Different construction quality and recovery features
- Different MTBF: Mean Time Between Failures
  - Enterprise HDD: 1.2M hours, at 45ºC, working 24/7, 100% use rate (1)
  - Desktop HDD: 700K hours, at 25º , working 8/5, 10-20% use rate(1)

Adjusted to each use case

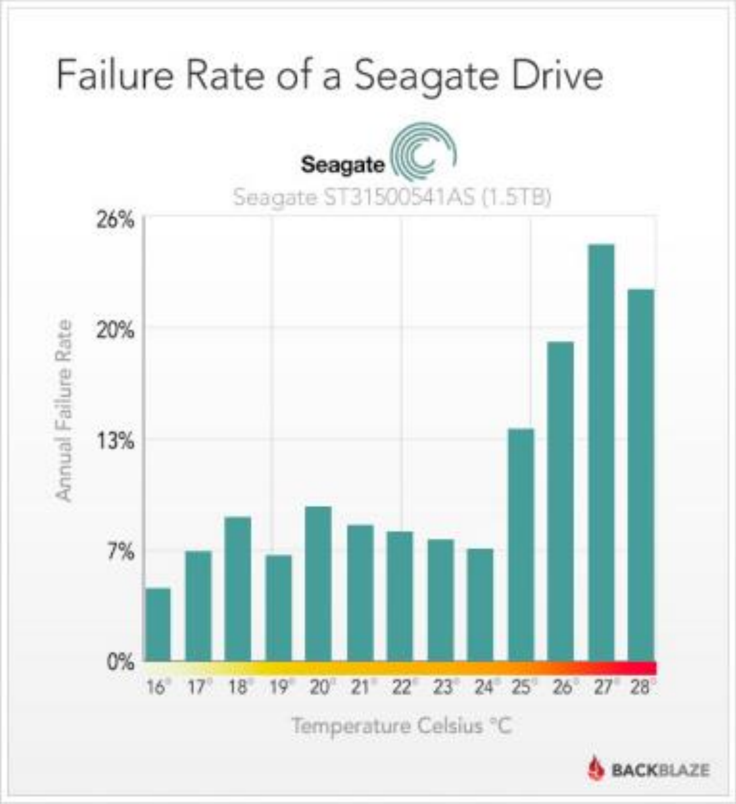
- Write intensive vs Read Intensive
- NAS vs Video vs Desktop vs Cold Storage vs Data Center
  - Differences in power consumption, reliability and performance

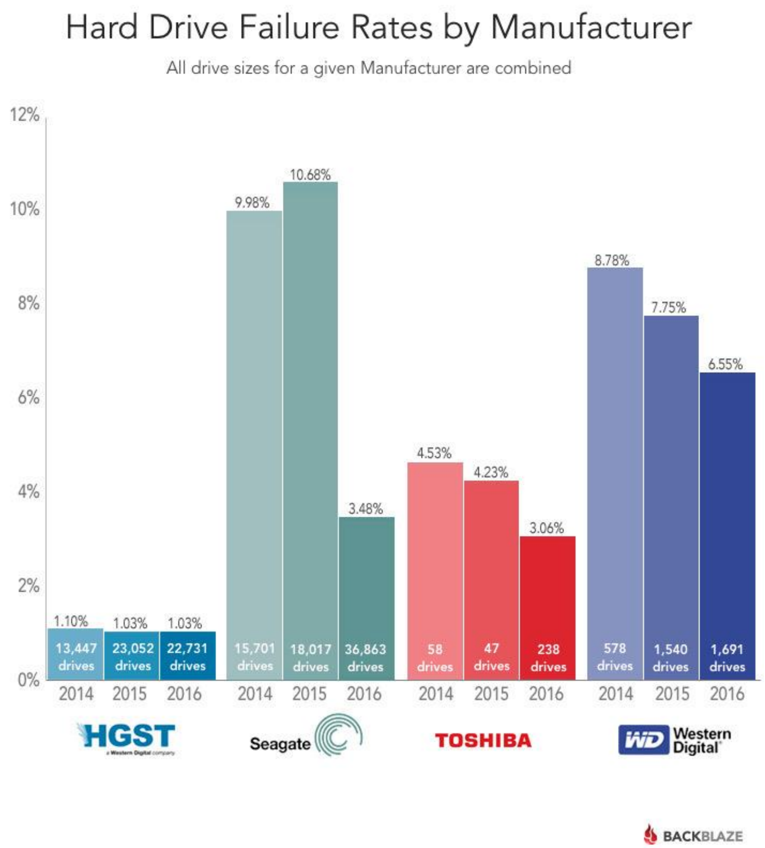
Adjusted to a specific performance level

- Tier 0: Highest performance, low capacity (PCIe NVME SLC SSD)
- Tier 1: Some performance, high capacity and availability (M2 SATA SSD)
- Tier 3: Low performance, high capacity, low price (SATA HDD)

1) Enterprise-class versus Desktop-class Hard Drives, rev 1.0, Intel, 2008

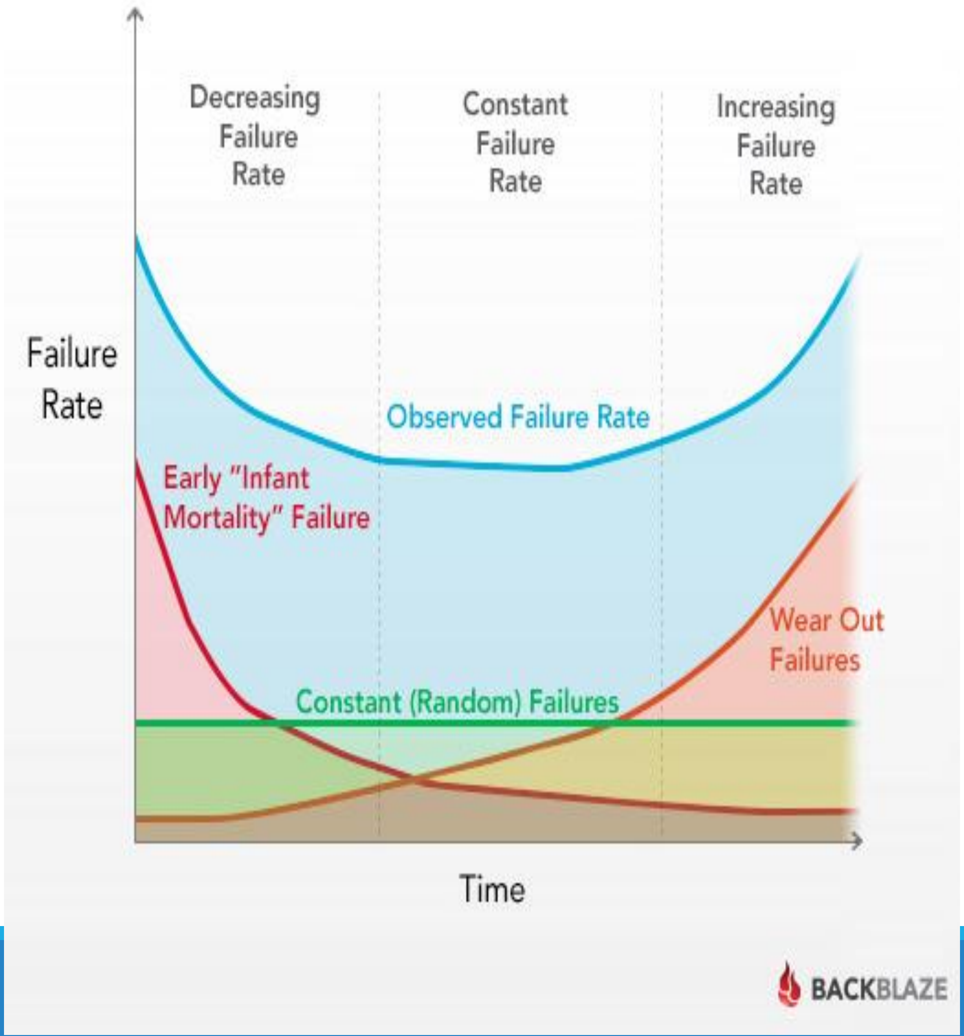
Controled Environment and Equipment



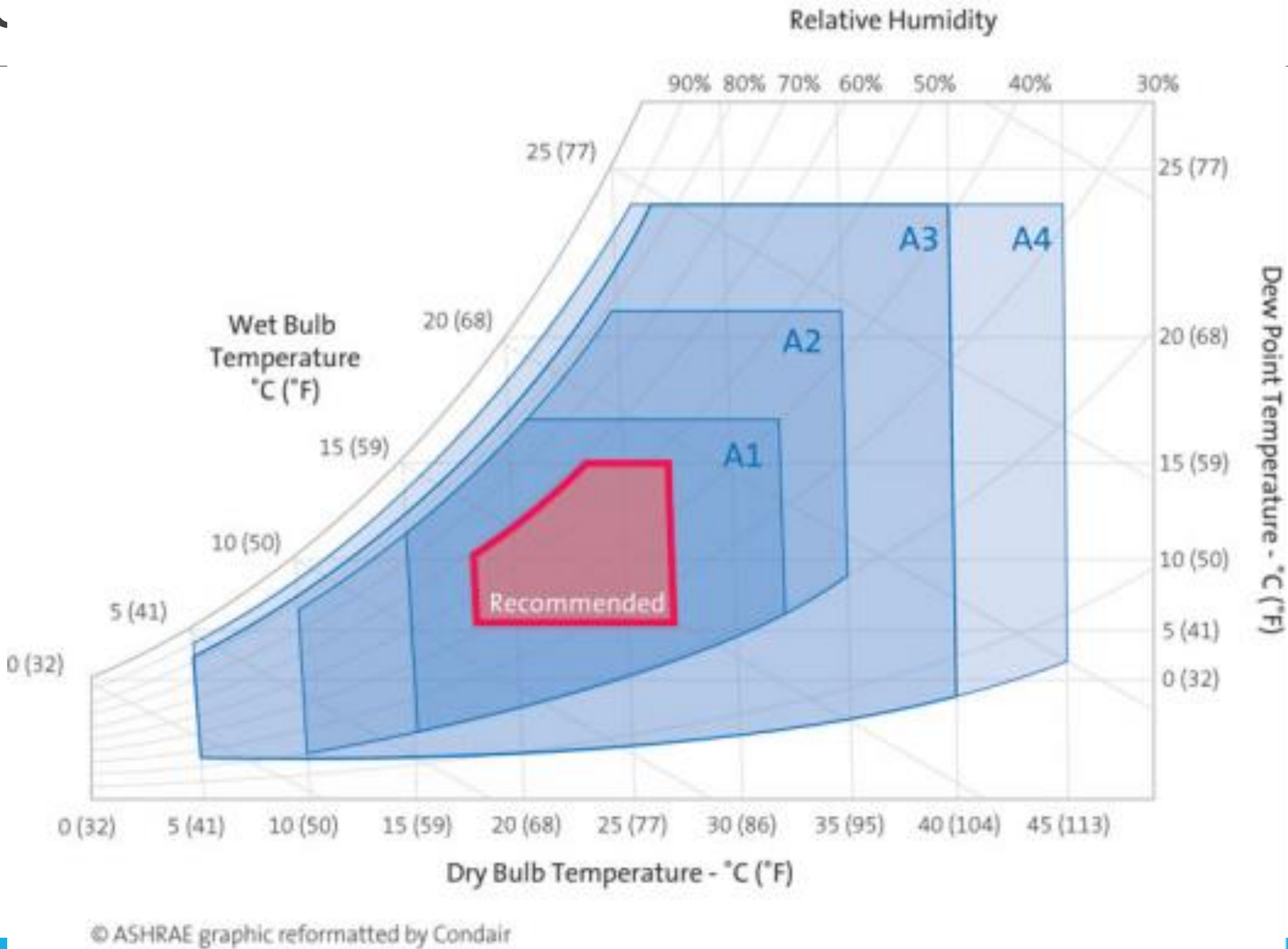


https://www.backblaze.com/b2/hard-drive-test-data.html

Controled Environment and Equipment



Controled Environment and Equipment



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

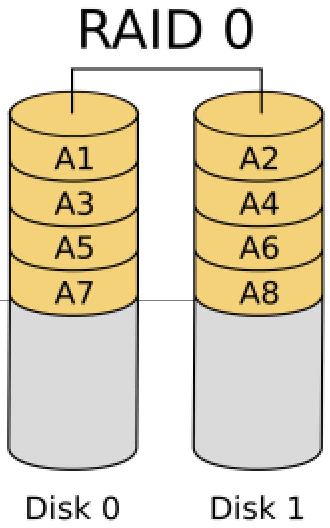
RAID doesn’t replace backups

- 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

RAID 0 (Striping)



Objectives

- Speedup data access

Approach

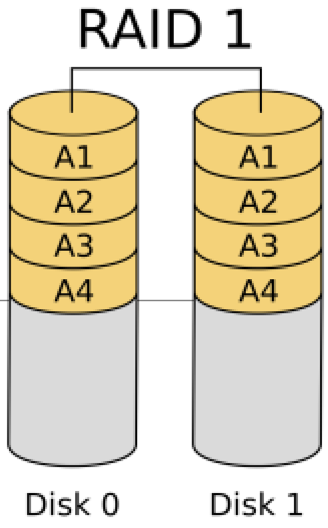
- Access disks in parallel
- Striping
  - Data is split in small chunks (stripes)
  - Stripes are stored among all disks in a distributed maner

Advantages

- May speedup performance as a factor of the number of disks

Disadvantages

- May increase the probability of loosing data
  - If Pf is the probability of failure of a single disk, a RAID 0 volume with N disks will have a failure probability of (1-(1-Pf))N
- Increases the number of devices
  - At least it will double the number



RAID 1 (Mirroring)

Objectives

- Tolerate the failure of disks

Approach

- Data duplication (mirroring)
  - Sincronized writing
  - Distributed read from any disk with or without comparison from another disk

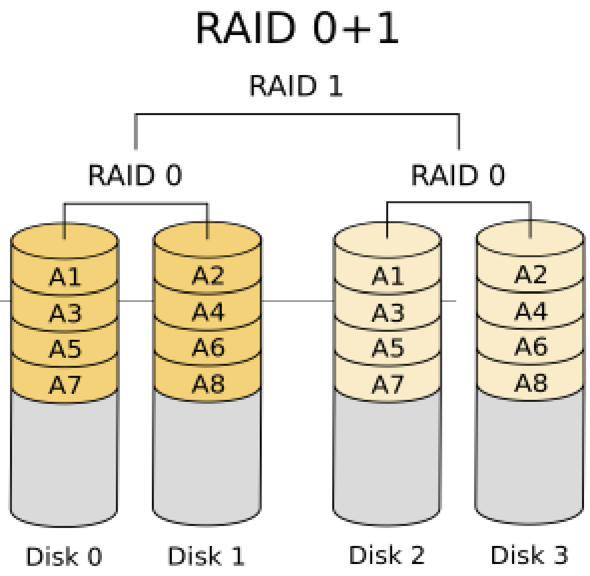
Advantages

- Decreases the probability of data loss
  - Considering the probability of failture of a single disk Pf, 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

RAID 0+1 (Nested)



Objectives

- Benefits of RAID 0 (performance)
- Benefits of RAID 1 (resilience)

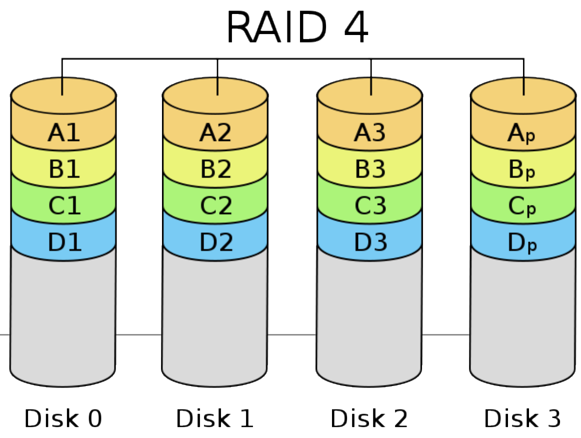
Approach

- A RAID 0 volume
  - Of RAID 1 volumes
- Result: Mirroring of striped volumes

Disadvantages

- Storage capacity waste
  - At least 50%
- Increase the number of devices

RAID 4



Objectives

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

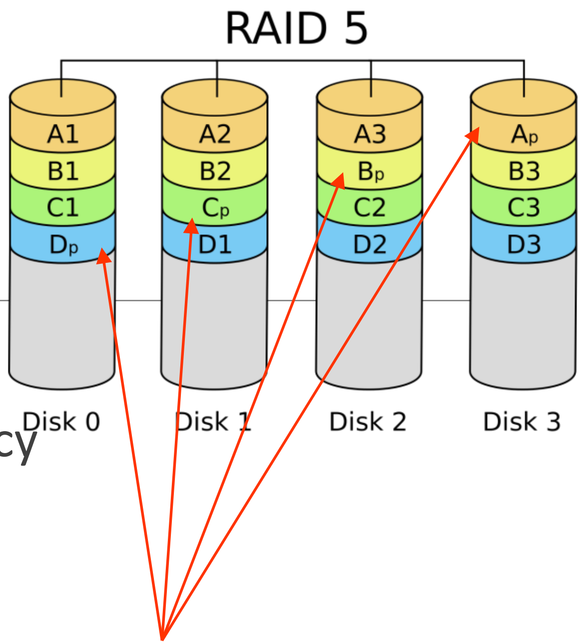
Approach

- Store data in N-1 disks
- Store parity data in a 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)
      - Read block from parity disk (Cp)
      - Compare old data block with new, and change the parity block (Cp')
      - Write the new data block (C1')
      - Write the new parity block (Cp')
  - Writes must be serialized due to the existance of a parity disk
- Recovery is way more complex than with RAID 1

RAID 5



Objectives

- Similar to RAID 4 but with higher write efficiency

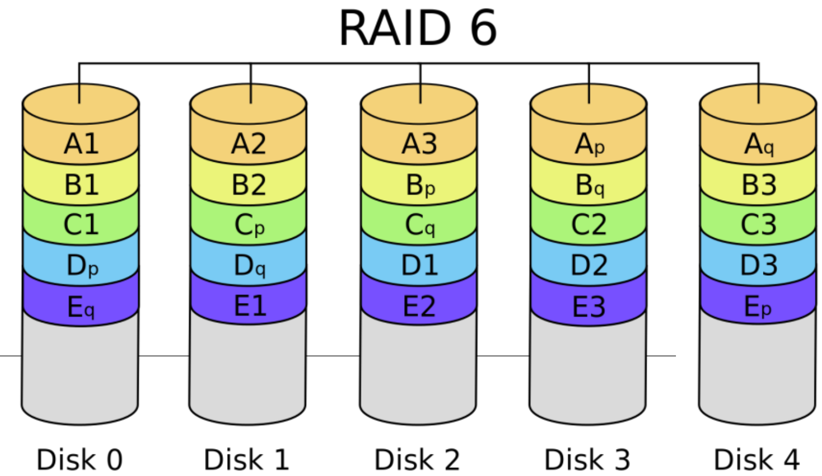
Approach

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

Disadvantages

- More complex to be implemented

RAID 6



Objectives

- Improve the realibility of RAID 5

Approach

- Use 2 parity blocks, distributed among all disks
- Capacity waste will be higher than in RAID 5 (eq to 2 disks)
- Concorrency is slightly worse than with RAID 5

Advantages

- Allows the failure of two disks without data loss

Disadvantages

- Even more than than RAID 5

NAS and SAN

NAS: Network Attached Storage
- Storage system available in the network
- Frequently created with RAID disks
- Cost: Hundreds to Thousands of Euro

SAN: Storage Area Network
- Set of systems available in a network
- Implemented distributed storage with redundancy
- Cost: Hundreds of Thousands to Millions of Euro

Advantages
- Allow centralizing the storage policies
- Provide a normalized interface, independente of the real storage
- May be used to distributed backups

Confidentiality of Data Storage

Problems

The protections provided by a traditional filesystem are limited

Physical Protections

- FS is limited to a physical device

Logical Protections

- Access Control to files, controled by the operating system
- Using ACLs na other confinement mechanisms

Problems

There is a relevant number of situations where standard protections are irrelevante

When there is direct and phyisical access to devices

- Access to host devices (laptops, smartphones, servers)
- Access to external storage devices
  - Tapes, CDs, DVDs, SSDs, NAS

Access through the system with the correct rights

- Non ethical access by system administrators
- With impersonation attacks

Problems

There is a prevalence of distributed storage

It imposes trusting multiple administrators, sometimes unknown

Authentication is made remotely

- Sometimes it is not clear what is the security level of said methods
- Storage Provider may have unkown integrations
- Interaction models are complex, through external networks
- Multiple entities involved

Information is transmited through communication channels

- May violate Confidentiality, Integrity and create Privacy issues

Solutions: Encrypt data

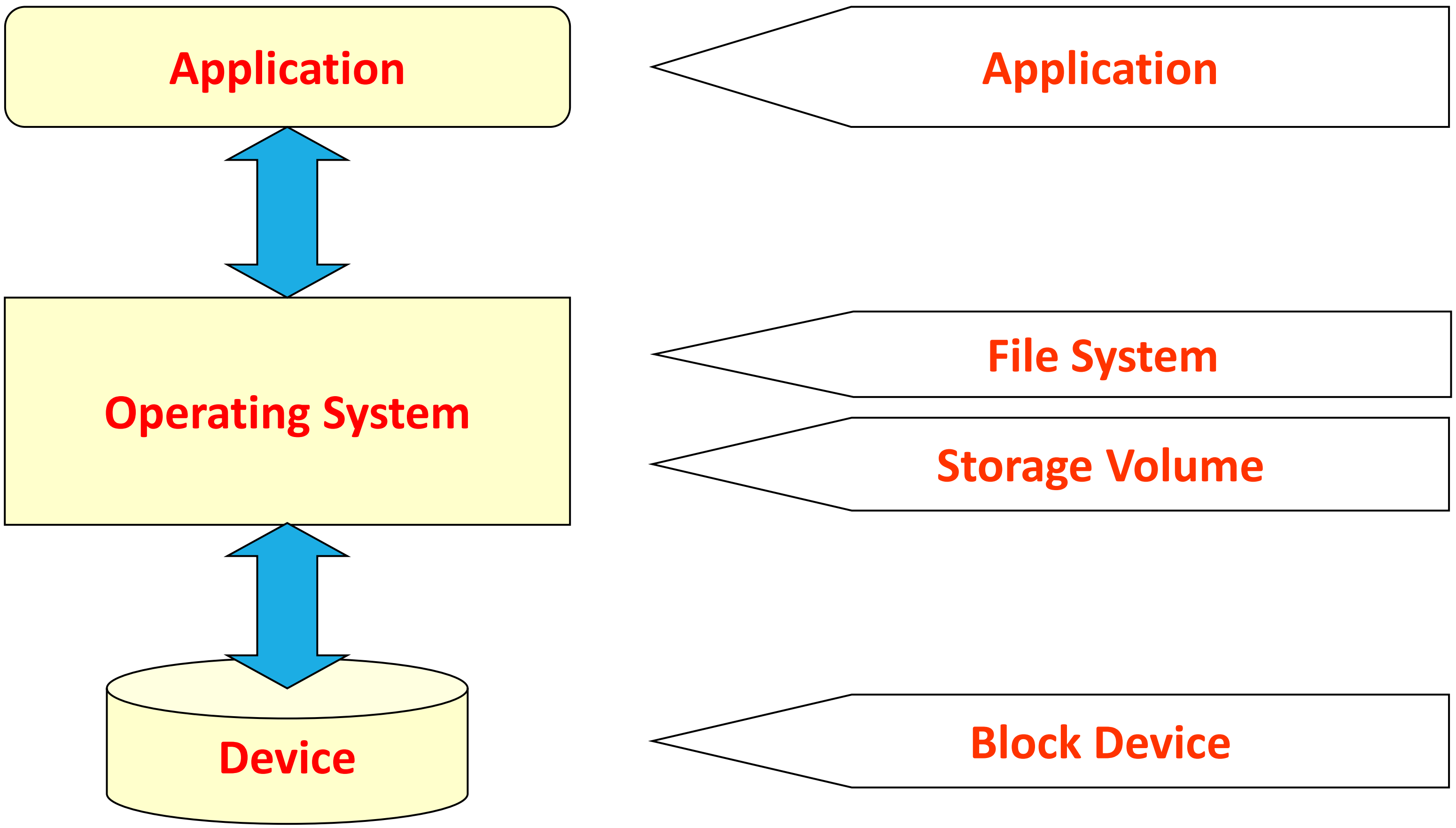
Encryption/Decryption of file contente

- Enable secure transfer over insecure networks
- Enable secure storage in insecure locations
  - Managed by external entities, or in shared storages

Problems of encryption

- Access to information
  - Users may lose the keys
    - Key loss = data loss
    - Key storage may reduce overall security
- File sharing
  - Sharing data implies sharing keys
- May interfere with standard management and recovert tasks
  - Content analysis, deduplication, indexing

Approaches


Application
Operating System
Device
Application
File System
Storage Volume
Block Device

Encryption in Applications

Information is transformed by each application

- Little or no integration with other applications
- Usually it is clear what is secure or not
  - Specific files with known file extensions

Present vulnerability windows

- Data must be encrypted to other files before it is access

Information may be processed by diferente algorithms/keys

- Adapted to a specific operating system or the security level
- May complicate the data recovery processes

May difficult sharing data inside the encrypted package

- May imply extract data which is stored in a clear format

Examples:

- PGP, AxCrypt, TrueCrypt, Veracrypt, etc..
- Also: RAR, ZIP, 7zip, LZMA…

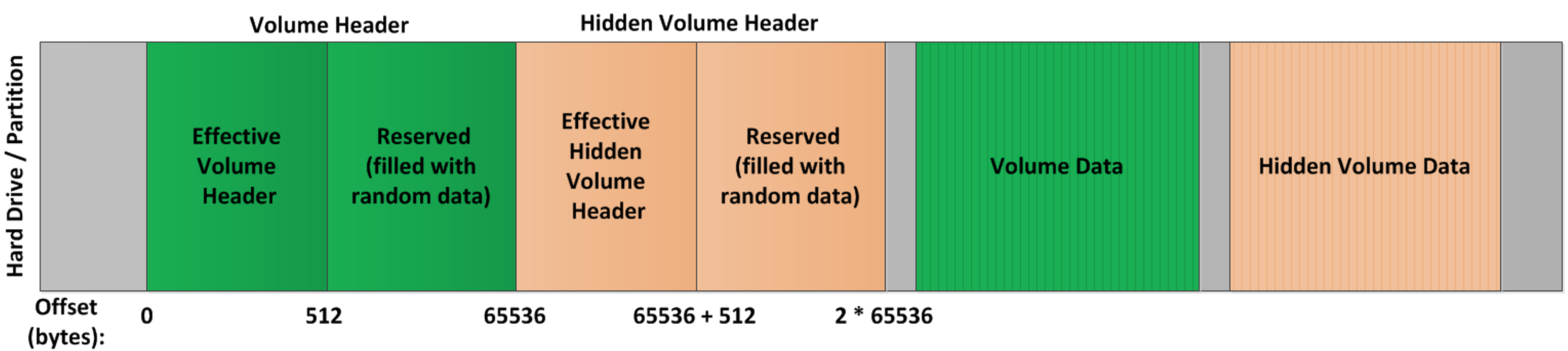
TrueCrypt

Was a popular application to manage FS over encrypted files

- Similar to a disk image used in virtual machines
- Used strong ciphers in cascade (e.g. AES+Twofish)
- Multiple modes: AES-CBC, AES-LRW, AES-XTS
- Key derivation processes: PKBDF2, SHA-512 and 2000 rounds

Interesting concept: Plausible Deniability

- Filesystems inside the file do not have obvious headers
- A file may have multiple volumes
  - It is not possible to prove how many volumes really exist
  - Different passwords unlock diferente volumes



Volume Encryption

Encryption in the File Systems

Information is transformed when is sent from memory to the filesystem

- May be broad, from the entire filesystem into the global memory cache
  - No protection in shared servers as data is available to all applications
  - Security mechanism is harder to implemente in distributed environments
    - Coordination of ACLs
- May be specific to the cache of a specific process
  - Protection in the case of shared servers as data access is contexto bound
  - Client API deficers data

Examples

- EncFS, EXT4, NTFS, CFS

Encryption at the volume level

Information is transformed by the volume driver

- Transparent to applications and almost transparente to OS
  - Requires support through a specific driver
- The entire volume will be made available (partition)

Policies defined through applications or the controller

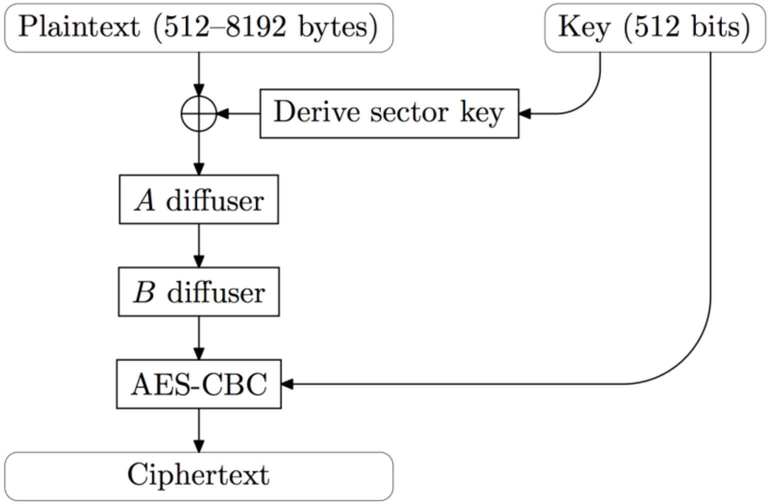
- Agnostic to the actual filesystem on top
  - Protects everything, including metadata
- But it doesn't differentiate between individual users

Unable to solve problems related with distributed systems, but solves those related with mobile devices

- Distributed systems expose the filesystem after decryption
- Mobile devices: lost of stolen devices will keep data secure

Examples:

- PGPDisk, LUKS, BitLocker, Filevault



Bitlocker (Windows)

Encrypts an entire volume

- Uses a small volume to bootstrap
- Key is composed (FVEK): K_{AES}, $K_{diffuser}$

Key Storage

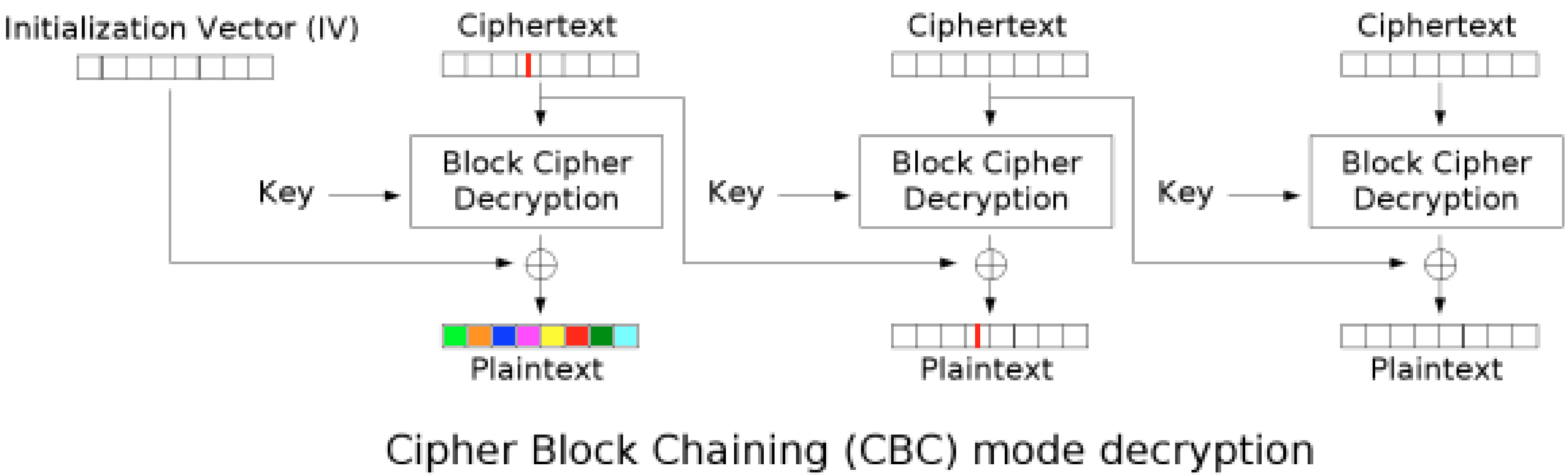
- FVEK encrypted with a Volume Master Key (VMK), Encrypted with a Key Protector Key
- Key Protector Key encrypted with secret provided by user on in a TPM

Encryption process

- AEC-CBC 128 or 256 applied to each sector, without MAC or intersector feedback
- IV = $E(K_{AES}, e(s))$, where e mapps the sector number to 16bits
- Sector Key = $E(K_{AES}, e(s))$ | $E(K_{AES}, e'(s))$
- Elephant diffuser: difusor for whitening, controlled by $K_{Diffuser}$

Bitlocker (Windows)

Malleability attack in AES-CBC



Encryption at the Device Level

Block Device applies security policy internally
- At boot. Device must be unlocked
  - After the correct credentials are provided
- Encryption is implemented at the hardware/firmware

Advantages
- No performance loss
- Data access is not trivial as keys are internal
- May be coordinated with applications (e.g USB devices)

Disadvantages
- After the device is unlocked, all data is made available
- Security is limited by the algorithms presente
- The possible existence of backdoors is difficult to find and correct

Encryption at the Device Level

Devices have two distinct áreas

- Shadow Disk: Read Only, ~100MB with software to unlock it
- Real Disk: Read Write. Contains user data

Two keys used

- KEK: Key Encryption Key (AuthenticationKey)
  - Provided by te 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 K hash(KEK)
- If it maches, MEK is decrypted and disk geometry is updated