Armazenamento

Problemas

- **Os dispositivos de armazenamento avariam**
  - É preciso minimizar a falha de discos ou a perda de informação
  - É uma certeza para qualquer dispositivo! Resta saber quando.
- **O acesso mecânico à informação é lento (Discos)**
  - Tempo = tempo de translação + tempo de rotação
  - Mais informação -> maior estrangulamento
- **Dispositivos sólidos (SSD) possuem número de escritas reduzidas**
  - 2000—3000 escritas para tecnologia MLC
- **Existem eventos que levam à perda total de dados**
  - Incêndios, roubos, “picos de energia”, inundações, erros do utilizador, ataques informáticos....
- **Pode ser necessário distribuir dados de forma inteligente**
  - Para maximizar desempenho
  - Para reduzir custos

Soluções

- **Cópias de segurança (backups)**
  - No local
  - Remotos
- **Armazenamento Redundante**
  - RAID
  - Outros: ZFS
- **Discos mais caros, ambientes mais controlados**
  - SLED (Single Large Expensive Disks)
  - Discos “Enterprise grade”
  - Controlo de Temperatura e humidade
- **Infraestruturas dedicadas de armazenamento**
  - Ponto único de aplicação de políticas

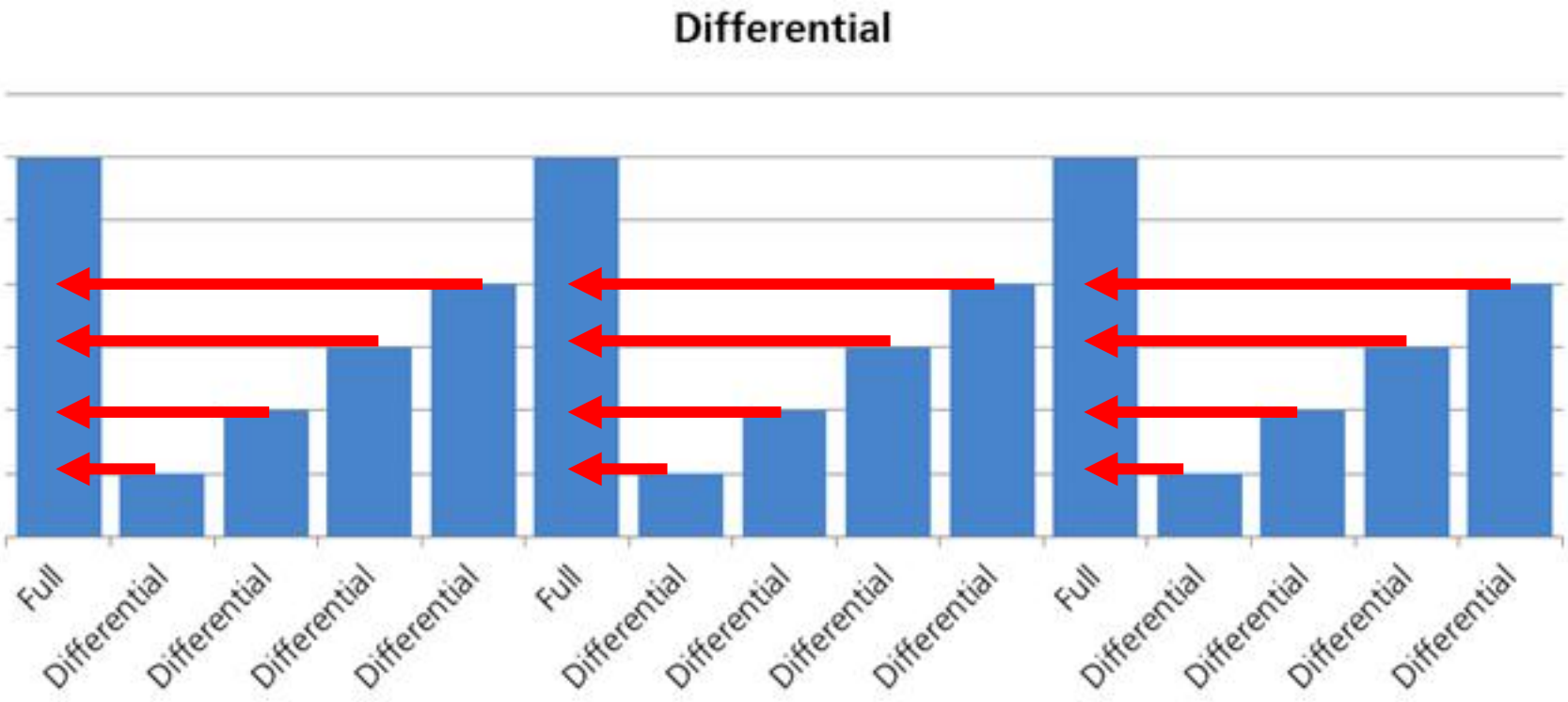
Backups

- **Cópias periódicas dos dados**
  - Imagem do estado do armazenamento naquele momento
  - Cópias permitem repor ficheiros para versões anteriores
  - Por vezes cifradas
- **Completos: Imagem completa da informação**
  - Recuperação rápida
  - Necessário muito espaço
- **Diferenciais: Diferenças desde o último backup completo**
  - Recuperação mais lenta com redução de espaço
  - Diferenciais diários vão aumentando progressivamente de tamanho
- **Incrementais: Diferenças desde o último backup**
  - Recuperação muito mais lenta
    - Reconstrução incremental desde o último backup completo
  - Grande eficiência de espaço

Backups

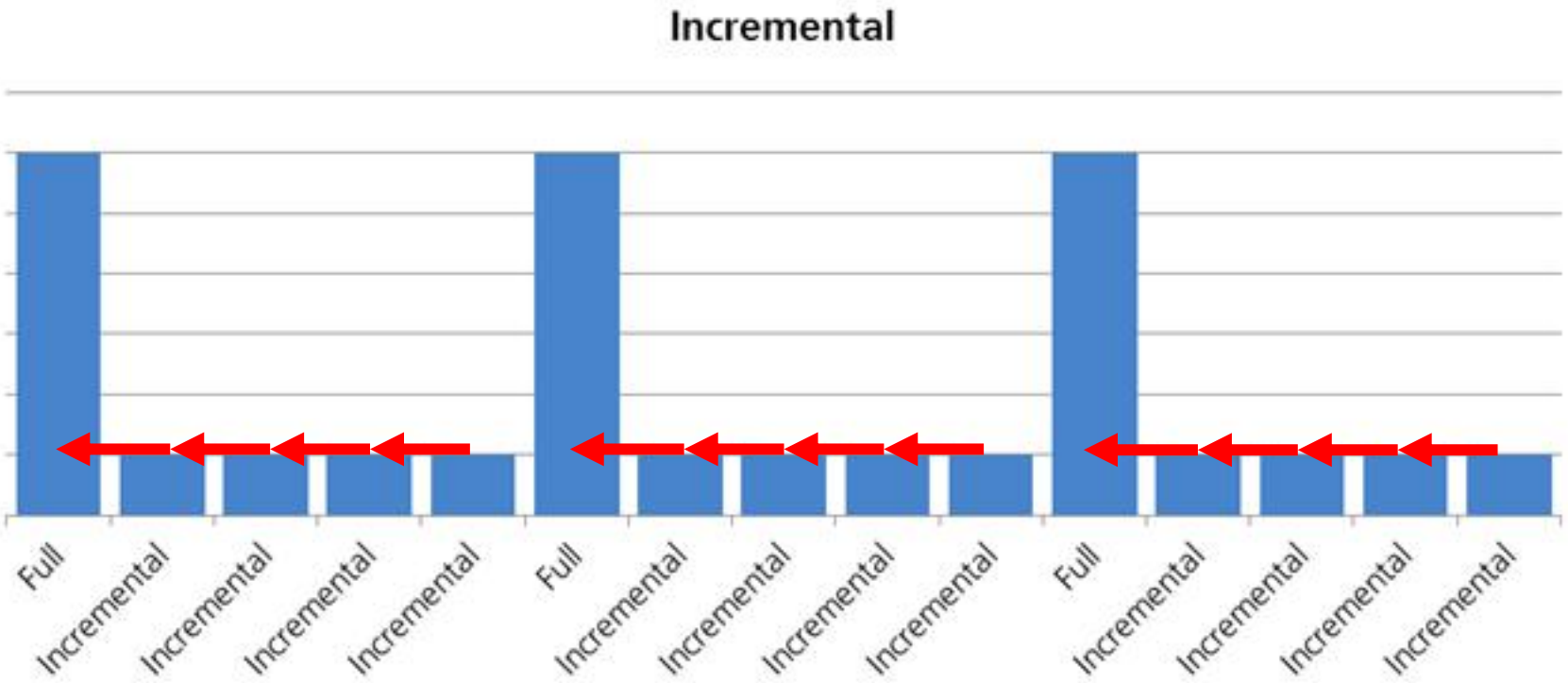
- **Não é armazenar informação num disco adicional**
  - externo, remoto
- **Considera políticas, mecanismos e processos para realizar, manter e recuperar cópias de informação**
  - Que resista a várias situações
  - Apenas usado em situações de catástrofe
  - Que considere a realização da cópia, armazenamento e restauro
- **Enquadramento legal obriga a cuidado especial**
  - Podem existir dados pessoais
  - Necessitam de ter uma política de retenção
    - Backups têm de expirar

Backups: Tipo Diferencial



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

Backups: Tipo Incremental



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

Backups: Tipo 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: Compressão

- **Compressão por algoritmos sem perdas**
  - Ex: zip
- **Cópias seletivas da informação**
  - Apenas os ficheiros que foram alterados (inc, ou diff)
- **Deduplicação**
  - Armazenar apenas ficheiros/blocos únicos
  - Cópias totais com processo de redução posterior
    - De blocos usando formatos de imagens adequados
    - De ficheiros através de ligações (ex, hardlinks)

Backups: Compressão e Deduplicação

			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 ignora hardlinks repetidos

Backups: Níveis

- **Aplicacional**
  - Extração dos dados da aplicação (ex mysqldump).
  - Representa uma vista consistente para a aplicação
    - Pode ser necessário bloquear o estado da aplicação (ex. escritas na DB)
  - Necessário repetir para todas as aplicações existentes
- **Ficheiros**
  - Cópia dos ficheiros individuais
  - Permite copiar qualquer aplicação
  - Estado guardado pode ser inconsistente
    - Ex. Ficheiros abertos com dados não escritos para o disco

Backups: Níveis

- **Sistema de Ficheiros**
  - Mecanismos próprios do sistema de ficheiros
  - Criação de registos de alterações periódicos
    - Snapshots temporais
  - Pode permitir recuperar ficheiros individuais ou não
- **Blocos**
  - Cópia dos blocos do suporte de armazenamento
  - Agnóstico do sistema de ficheiros e sistema operativo
  - Pode ser realizado pela infraestrutura de armazenamento
    - Transparente e sem impacto

Backups: Local da Cópia

- **No mesmo volume ou sistema**
  - Permitem aos utilizadores rapidamente recuperarem informação
  - Protege contra alterações/remoções indevidas de ficheiros
  - Não protege contra avarias do armazenamento
  - Ex: OS X TimeMachine
- **Num sistema localizado na mesma infraestrutura**
  - Também de acesso rápido
  - Protege contra falhas isoladas do armazenamento
  - Não protege contra eventos com maior âmbito
    - Inundações
    - Incêndios
    - Roubos
  - Ex: Maioria dos sistemas de armazenamento, Backuppc, Apple TimeCapsule

Backups: Local da Cópia

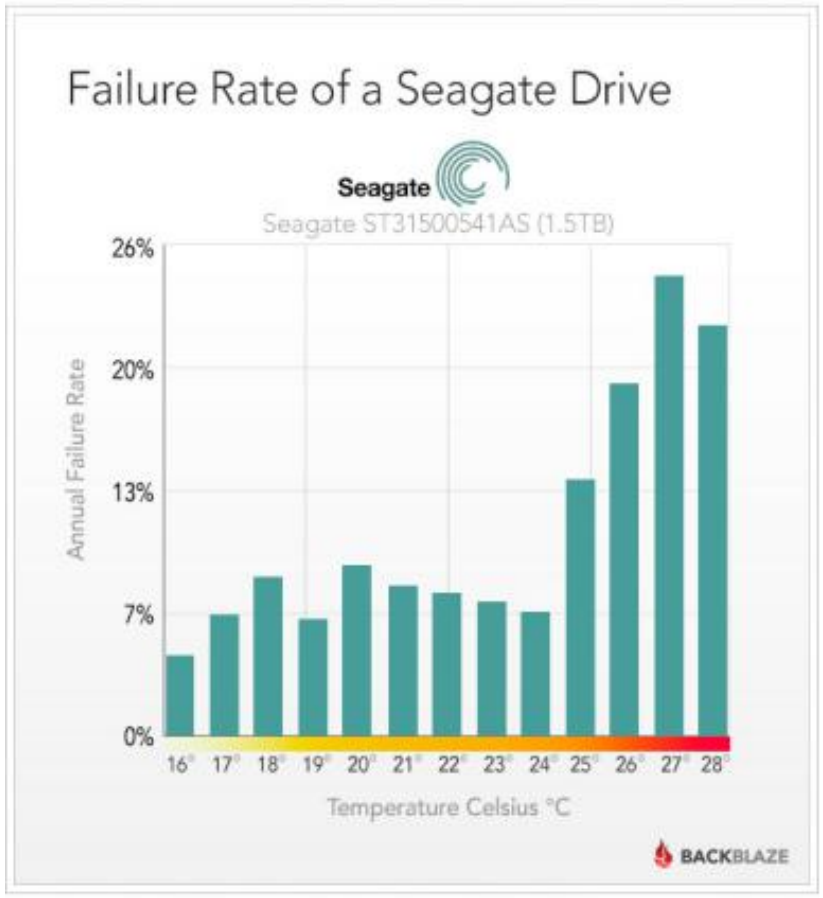
- **Remotos (Off-site)**
  - Realizados para um sistema a uma grande distância
    - Serviço disponível via rede dedicada ou Internet
      - ex, para Amazon S3, ou para servidores num DC alternativo ou alugado
      - Cifras são recomendadas (obrigatórias) no caso de serviços externos!
    - Transporte especializado para local seguro
      - ex, via um veículo seguro que transporte os suportes de armazenamento
  - Permitem recuperar informação em caso de evento com grandes danos
    - Incêndio, roubo, inundação, terrorismo, terramoto...
  - Recuperação de informação muito mais lenta
    - Necessário ir buscar fisicamente a informação, ou transferir a informação via a Internet

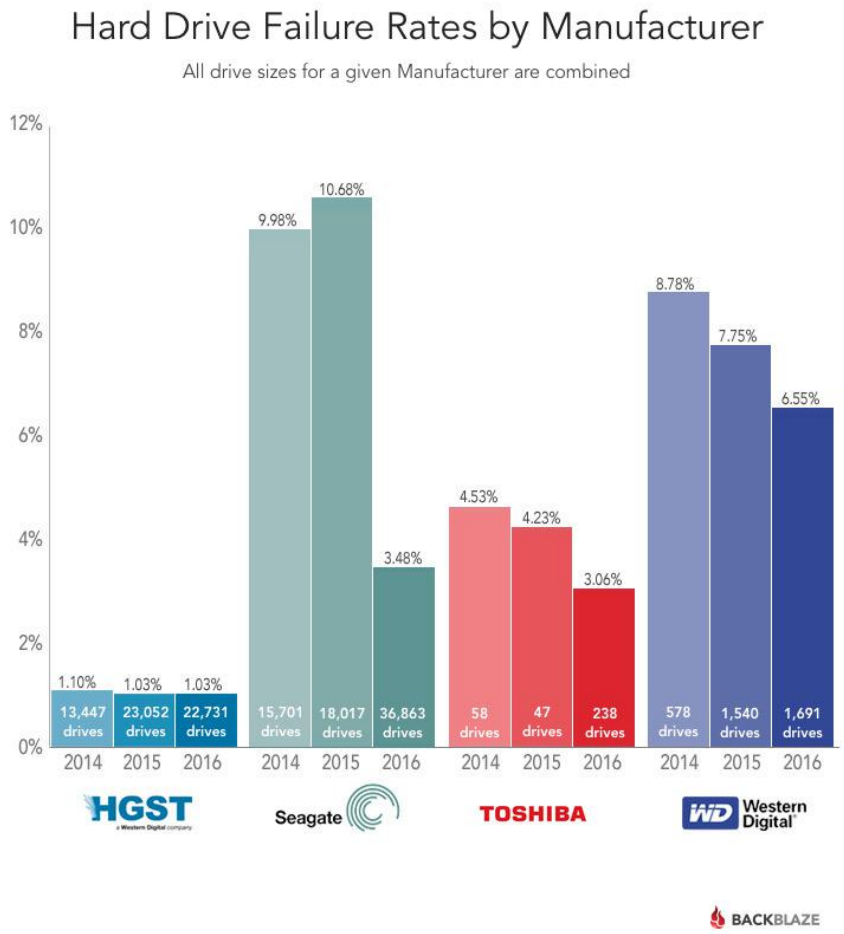
Seleção do Equipamento

- **Gamas Diferentes: Enterprise vs Desktop**
  - Qualidade de construção e mecanismos de recuperação
    - Qualidade... alegadamente
  - MTBF: Mean Time Between Failures
    - Enterprise HDD:: 1.2M hours, at 45ºC, 24/7, 100% use rate(1)
    - Desktop HDD: 700K hours, at 25º, 8/5, 10-20% use rate (1)
- **Ajustado ao caso de Uso**
  - Write Intensive vs Read Intensive
  - NAS vs Video vs Desktop vs Cold Storage vs Data Center
    - diferenças a nível do consumo, fiabilidade, desempenho
- **Ajustado ao nível de desempenho**
  - Tier 0: Desempenho muito alto e baixa capacidade (PCIe NVMe SSD)
  - Tier 1: Desempenho, capacidade e disponibilidade altos (M2 SATA SSD)
  - Tier 2: Desempenho baixo, alta capacidade (SATA HDD)

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

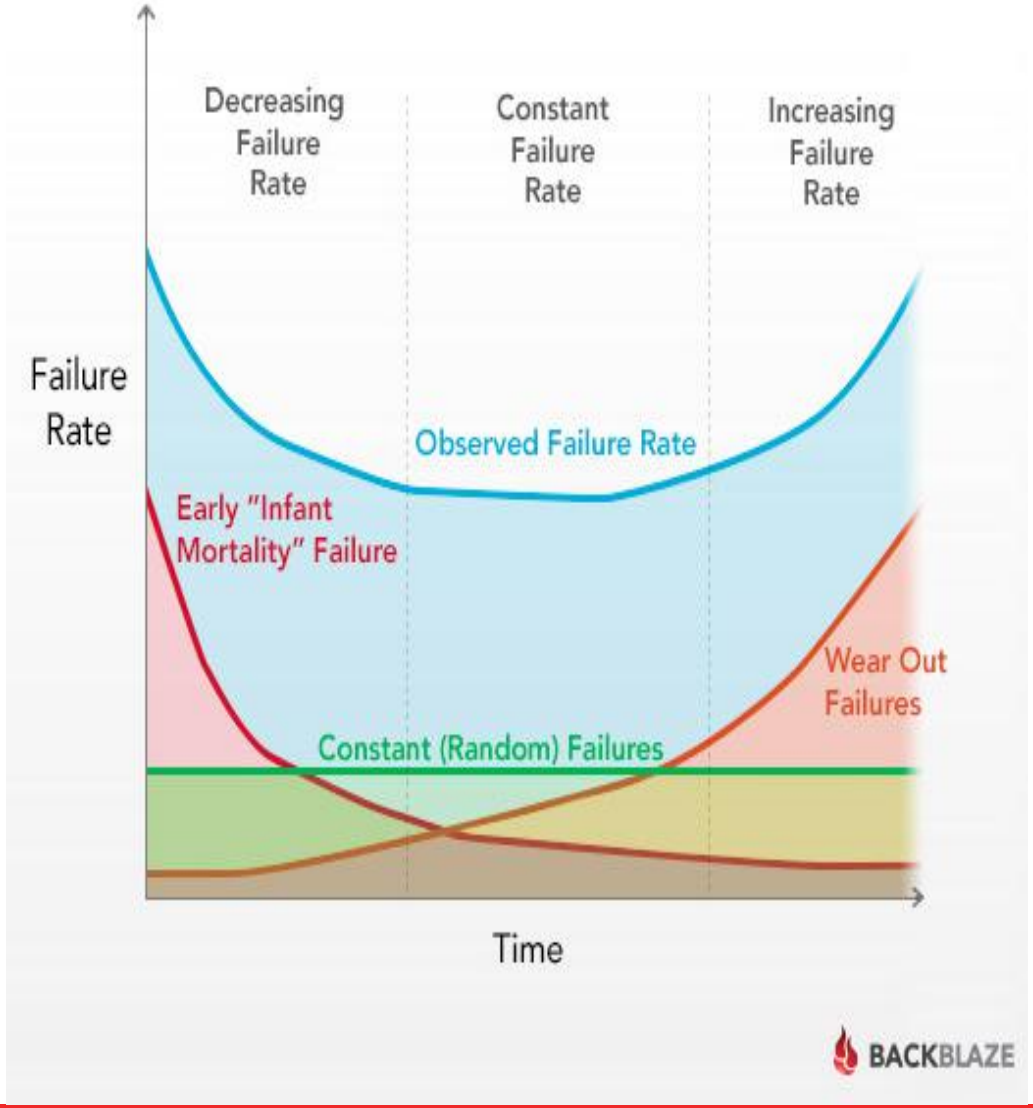
Ambientes e Equipamentos Controlados



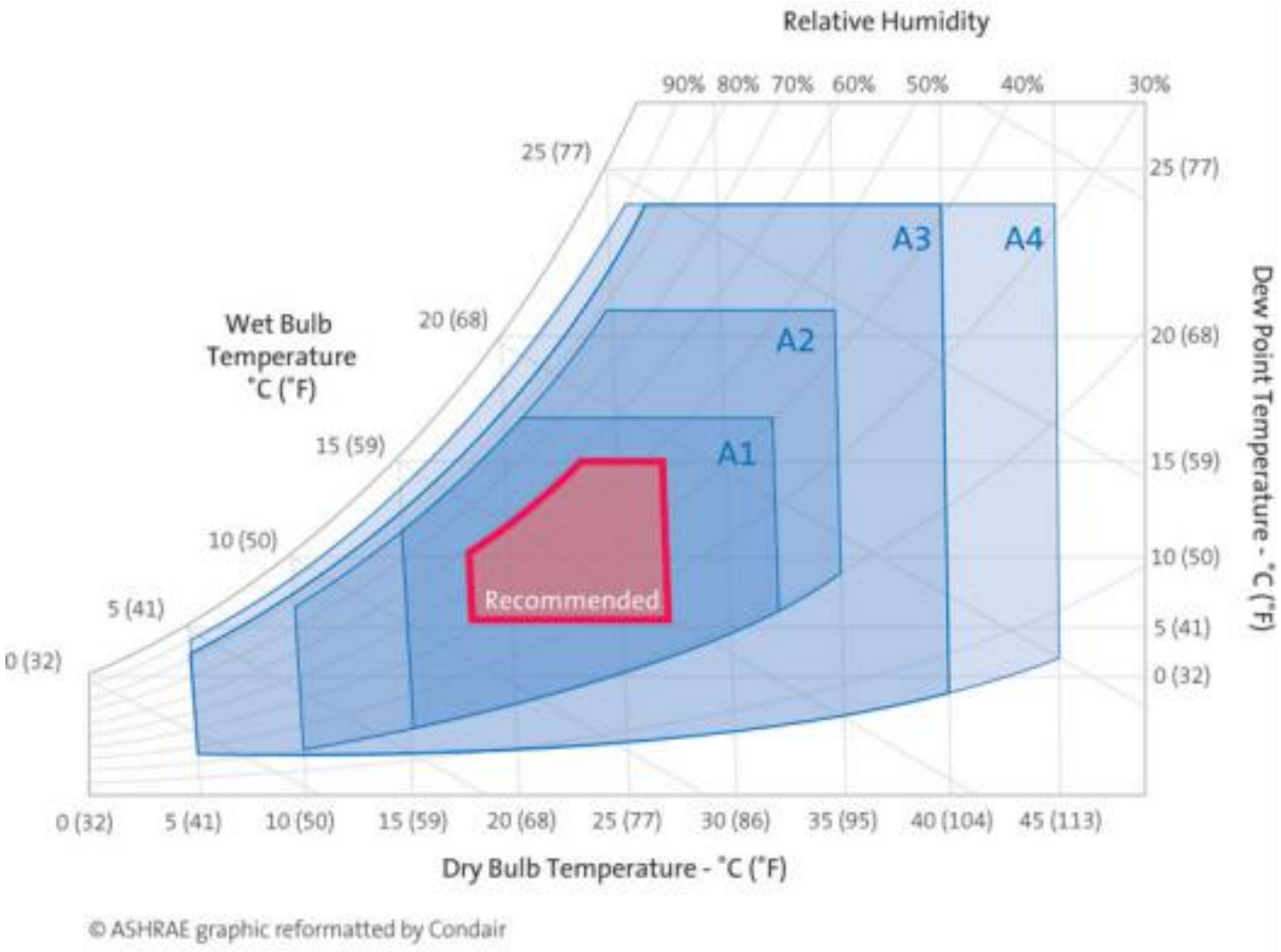


https://www.backblaze.com/b2/hard-drive-test-data.html

Ambientes e Equipamentos Controlados



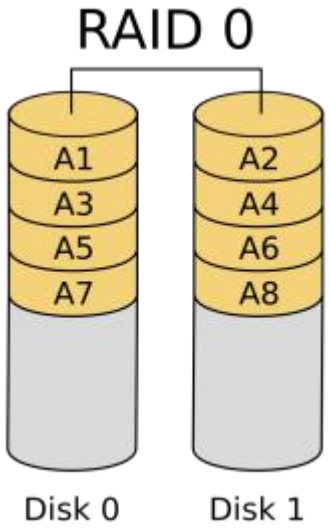
Ambientes e Equipamentos Controlados



RAID
Redundant Array of Inexpensive Drives

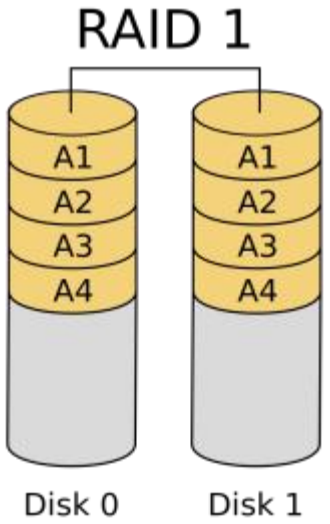
- **Garantir a sobrevivência da informação**
  - Os dados só se perdem se falharem mais do que X discos do RAID
  - O valor de X depende do tipo de RAID
- **Solução de baixo custo e eficiente**
  - Permite usar hardware barato, falível
  - Acelerar o desempenho nas leituras e escritas em discos
- **Mas o RAID não substitui o backup!**
  - Não tolera falhas catastróficas em mais do que X discos dos N do RAID
  - Não tolera erros dos utentes ou do sistema
- **E o RAID pode aumentar a probabilidade de falha do sistema!**
  - Se o objetivo for apenas acelerar o mesmo

RAID 0 (striping)



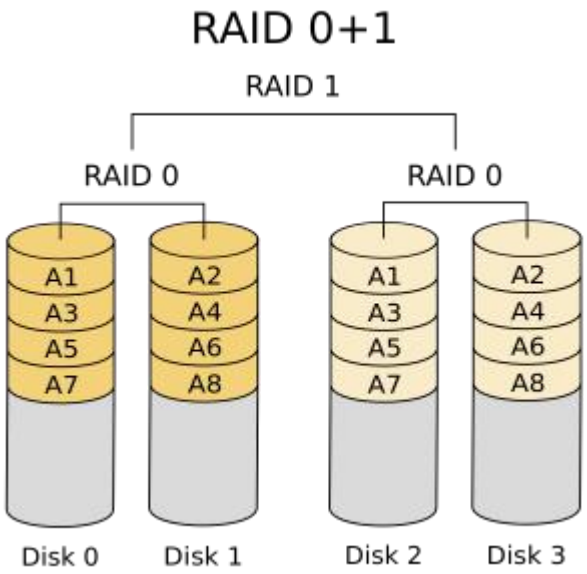
- **Objetivos**
  - Acelerar o acesso à informação em disco
- **Aproximação**
  - Acesso a discos em paralelo
  - Striping
    - A informação lógica de um volume é subdividida em fatias (stripes)
    - As fatias são intercaladas nos discos
- **Prós**
  - Aceleração dos acessos aos discos até N vezes
- **Contras**
  - Aumento da probabilidade de perda de informação
    - Se PF for a probabilidade de falha de um disco, a probabilidade de perder informação com um RAID 0 com N discos é 1–(1-PF)N
  - Aumento do número de dispositivos
    - Pelo menos para o dobro

RAID 1 (mirroring)



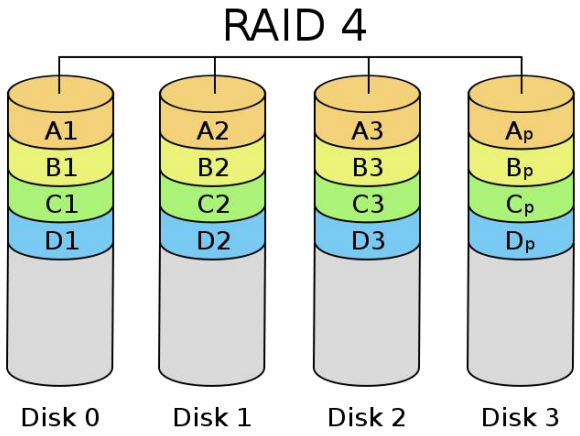
- **Objetivo**
  - Tolerar falha de discos
- **Aproximação**
  - Duplicação da informação (mirroring)
    - Escrita sincronizada
    - Leitura com comparação ou de apenas um disco (mais rápido)
- **Vantagens**
  - Diminuição da probabilidade de perda de informação
    - Considerando a prob. de falha de um disco PFD , a prob. de perda de dados com N discos é (PFD)-N
      - Ignorando falhas não isoladas (ex, pico de energia, temperatura excessiva)
- **Desvantagens**
  - Desperdício da capacidade de armazenamento
    - Perdido pelo menos 50% da capacidade (2 discos, 66% em 3 discos, .. (N-1)/N)
  - Aumento do número de dispositivos
    - Pelo menos para o dobro

RAID 0+1



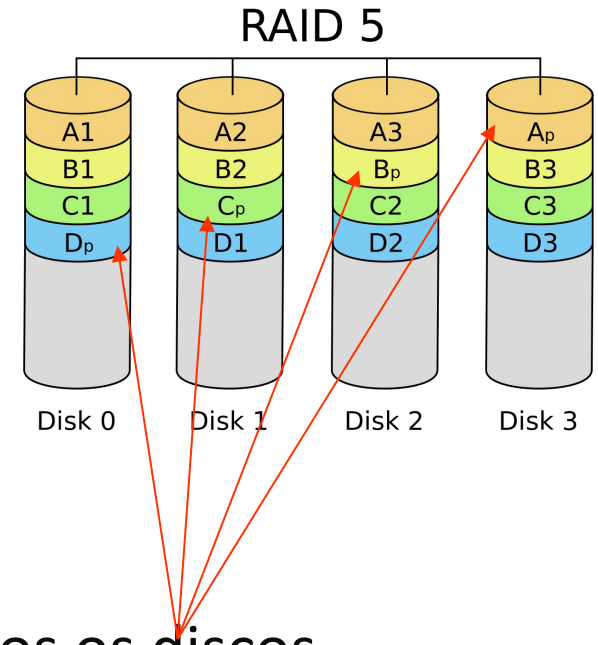
- **Objetivos**
  - Benefícios do RAID 0 (desempenho)
  - Benefícios do RAID 1 (resistência a falhas)
- **Aproximação**
  - Um nível RAID 0
    - ... de volumes em RAID 1
  - Ou seja: mirroring de volumes striped
- **Contras**
  - Desperdício de capacidade de armazenamento
    - Pelo menos 50% da capacidade é perdida
  - Aumento do número de dispositivos necessários

RAID 4



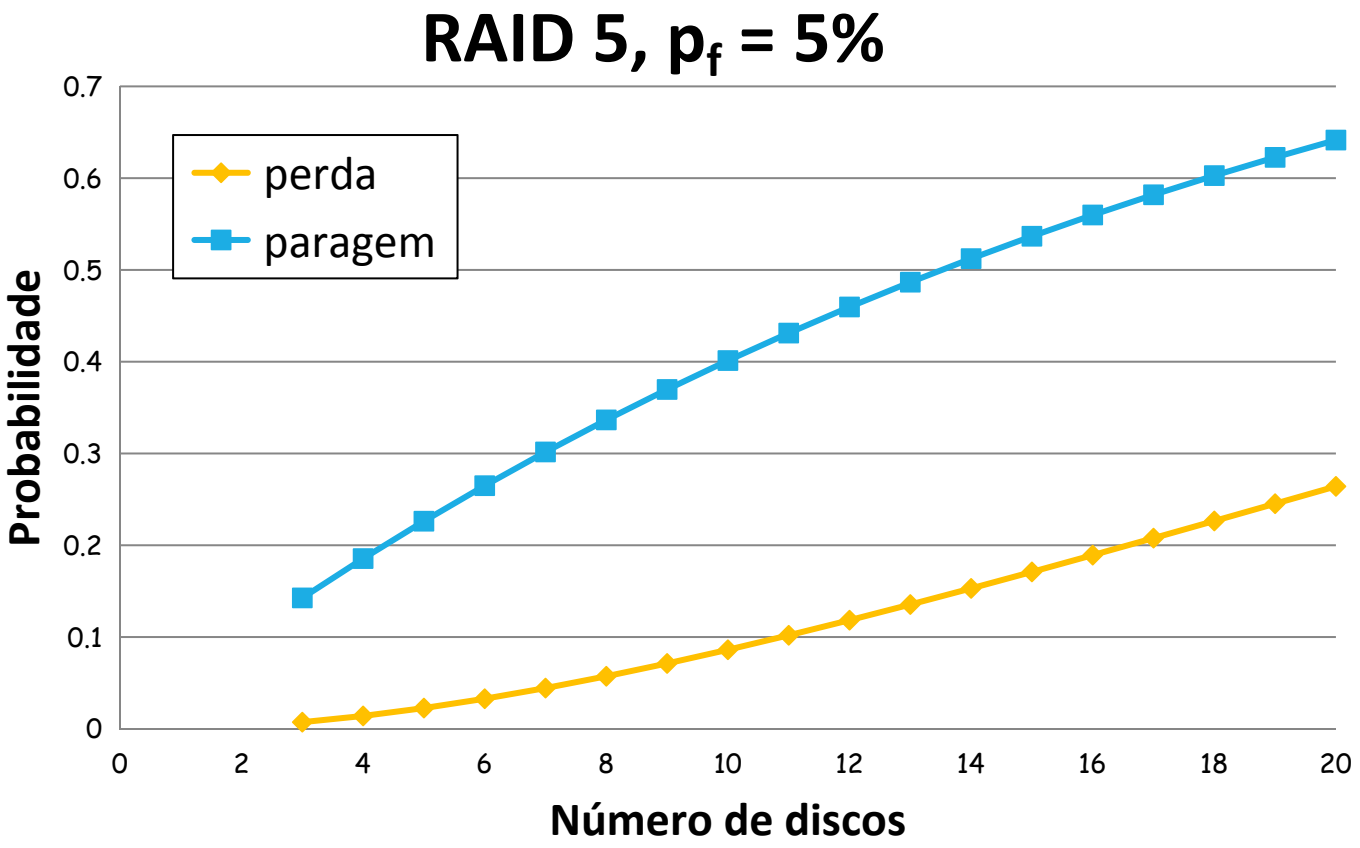
- **Objetivos**
  - Ter a proteção do RAID 1
  - Ter um desempenho e um eficiência de espaço próximos do RAID 0
- **Aproximação**
  - Armazenamento de dados em N-1 discos
  - Armazenamento de paridade num disco
    - O desperdício de espaço é igual a à capacidade de cada disco
    - Os dados de quaisquer N-1 discos podem ser gerar um outro
- **Problemas**
  - Necessita de 3 ou mais discos
  - A atualização da paridade é complexa e demorada
    - Obriga a leituras antes das escritas
      - Ler bloco de dados antigo (e.g. C1)
      - Ler bloco de paridade antigo (Cp)
      - Comparar bloco de dados antigo com novo, alterar o bloco de paridade (Cp')
      - Escrever bloco de dados novo (C1')
      - Escrever bloco de paridade novo (Cp')
    - As escritas têm de ser seriadas por causa do acesso ao disco de paridade
  - A recuperação é mais demorada do que com RAID 1

RAID 5

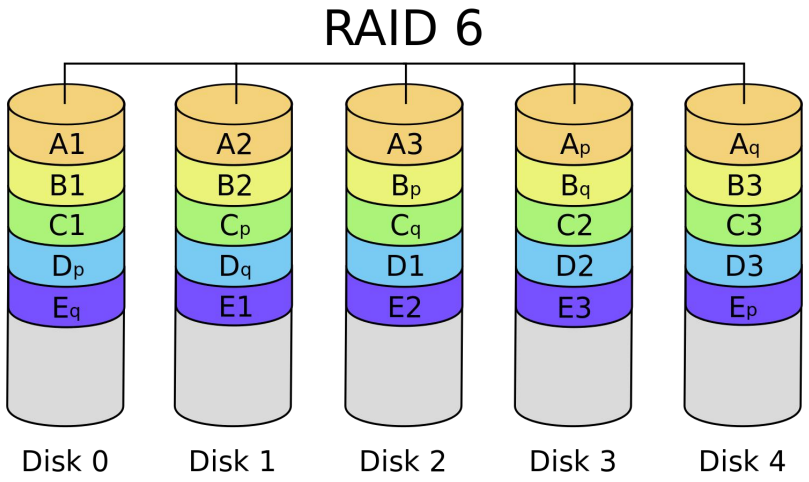


- **Objetivos**
  - Similar ao RAID 4 mas mais eficiente nas escritas
- **Aproximação**
  - Blocos de paridade espalhados por todos os discos
  - O desperdício de espaço é igual ao do RAID 4
  - A concorrência nas escritas é melhorada
- **Problemas**
  - Mais complexo do que RAID 4

RAID 5: Probabilidade de falha



RAID 6



- **Objetivos**
  - Melhorar fiabilidade do RAID 5
- **Aproximação**
  - 2 Blocos de paridade espalhados por todos os discos
  - O desperdício de espaço é maior do que o RAID 5
  - A concorrência nas escritas é ligeiramente pior que o RAID 5
- **Problemas**
  - Mais complexo do que RAID 5
- **Vantagens**
  - Permite falha simultânea de 2 discos

NAS e SAN

- **Network Attached Storage**
  - Sistema disponível por rede
  - Frequentemente com vários discos em RAID
  - Custo: centenas a milhares de euros

- **Storage Area Network**
  - Conjunto de sistemas disponíveis por rede
  - Pode implementar qualquer esquema de redundância
  - Custo: centenas de milhares a milhões de euros

- **Vantagens**
  - Permitem centralizar políticas de armazenamento
  - Fornecem interface normalizado independente do armazenamento real
  - Utilizados para armazenamento e cópias

Confidencialidade do Armazenamento

Problema

O sistema de ficheiros tradicional possui proteções que são limitadas

- **Proteções Físicas**
  - Sistema de ficheiros é confinado a um dispositivo
- **Proteções Lógicas**
  - O controlo de acesso é aplicado pelo sistema operativo
  - Faz-se uso de ACLs e outros mecanismos de confinamento

Problema

Existe um número de situações onde esta proteção é irrelevante

- **No caso de acesso direto e físico aos dispositivos**
  - Acessos aos dispositivos anfitriões (portáteis, smartphones)
  - Dispositivos de armazenamento discretos, por vezes externos
    - Tapes, CDs, DVDs, SSD, ...
- **Acesso através dos mecanismos de controlo de acesso**
  - Acesso não ético pelos administradores
  - Personificação de utentes

Problema

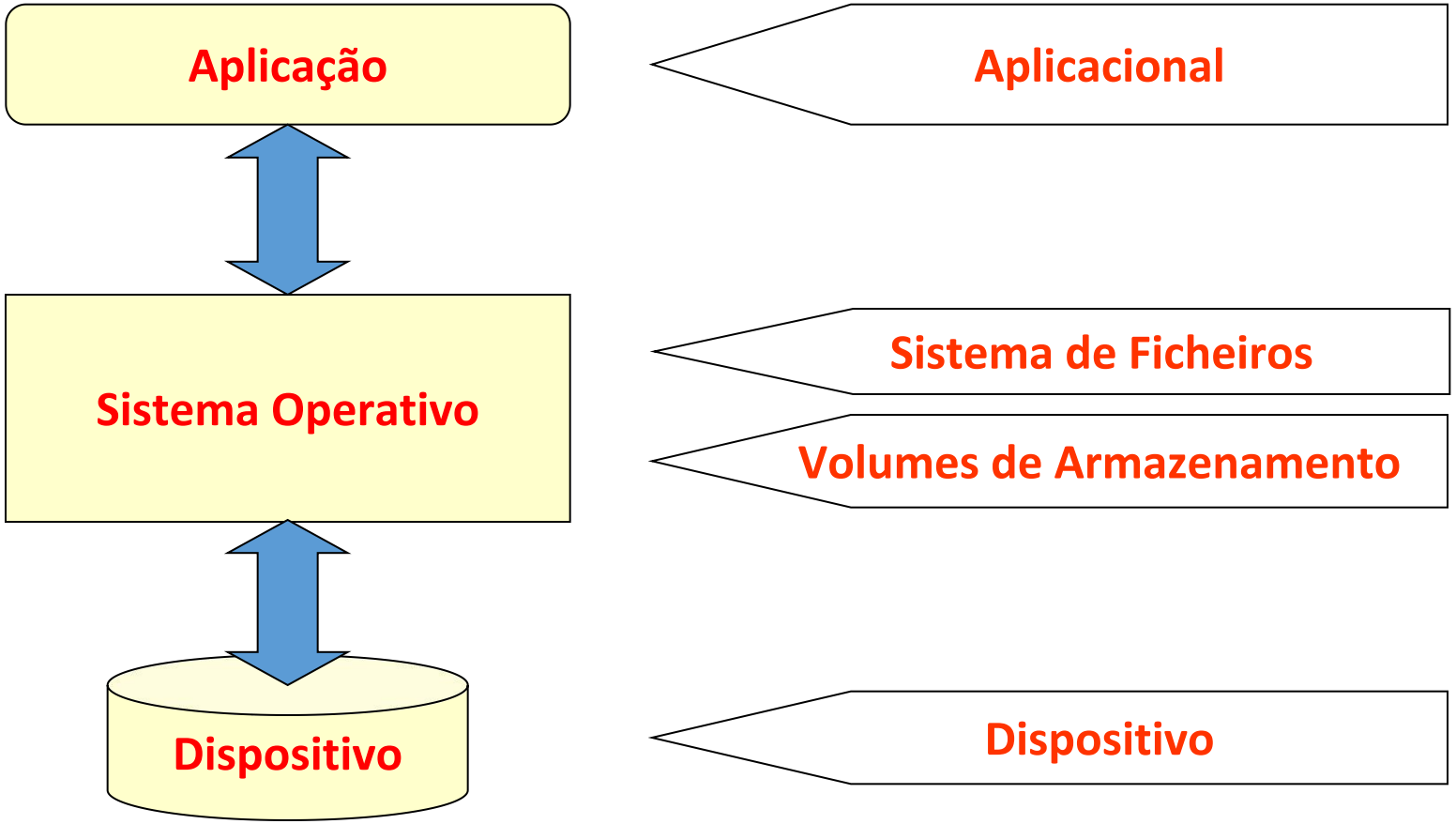
Prevalência de armazenamento distribuído

- **Necessária confiança em vários administradores (por vezes anónimos)**
- **Autenticação é efetuada remotamente**
  - Por vezes não é claro qual o nível de segurança
  - Existem integrações múltiplas e por vezes desconhecidas
  - Modelos de interação complexos
  - Diversos sujeitos
- **Informação é transmitida na rede**
  - Confidencialidade, Integridade, Privacidade

Soluções: Cifra de Informação

- **Cifra/Decifra do conteúdo dos ficheiros**
  - Permite a disponibilização segura sobre uma rede insegura
  - Permite o armazenamento em meios inseguros
    - Geridos por externos, ou em meios de armazenamento partilhados
- **Problemas**
  - Acesso à informação
    - Utentes não podem perder as chaves
      - perda das chaves = perda dos dados
      - cópias da chave diminuem a segurança
    - Cifra ilegítima ou abusiva da informação
      - Dados do empregador
  - Partilha de ficheiros
    - Implica libertação dos ficheiros ou das chaves
  - Possível interferência com tarefas comuns de administração
    - análise de conteúdos, deduplicação, indexação...

Aproximações

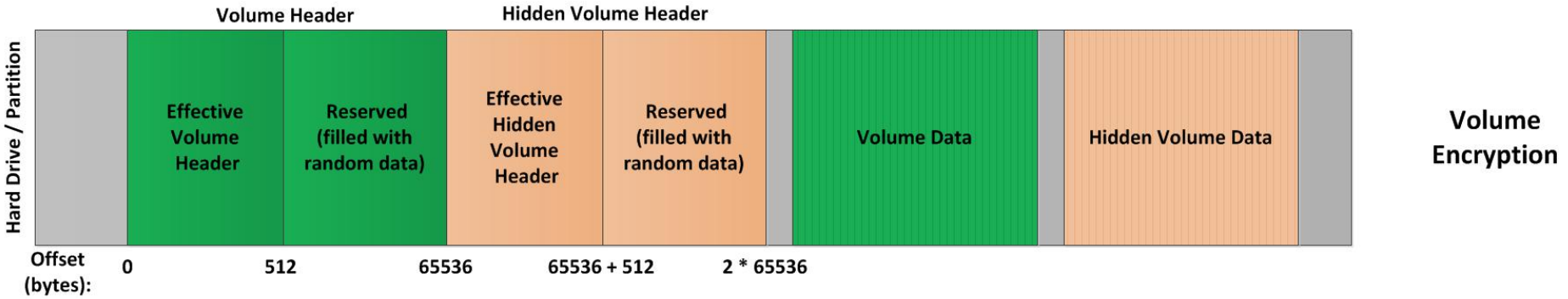


Nível Aplicacional

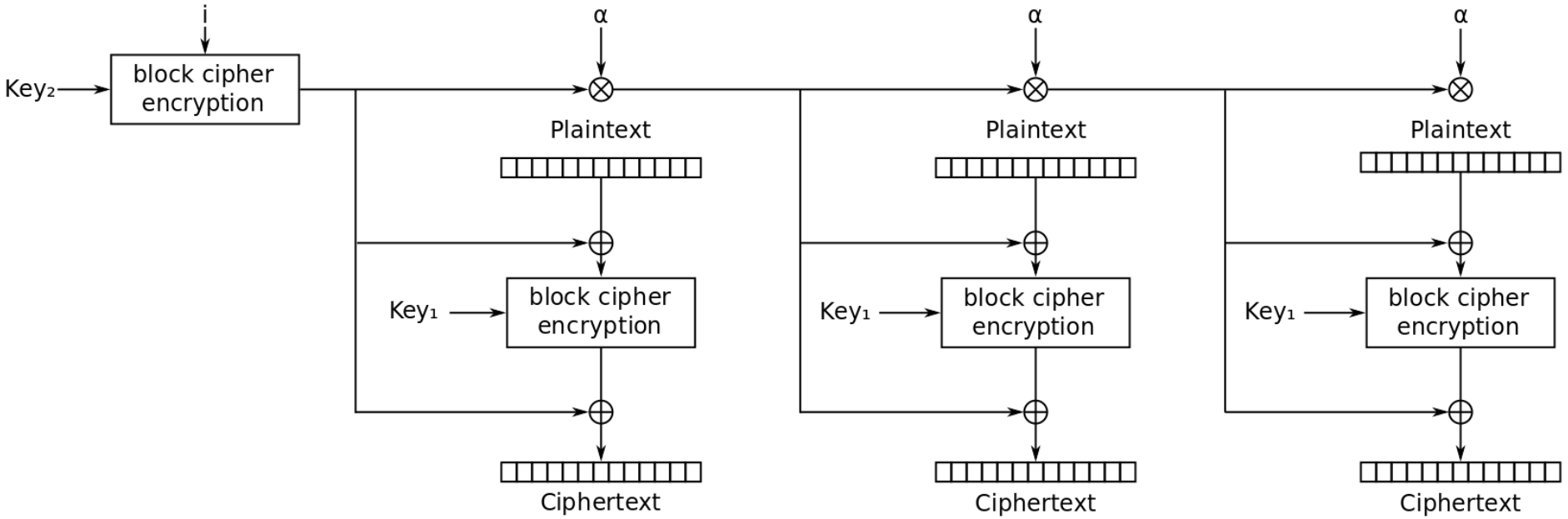
- **Informação é transformada por aplicações autónomas**
  - Pouca ou nenhuma integração com outras aplicações
  - Usualmente é claro o que é seguro ou não
    - Ficheiros específicos com extensões específicas
- **Apresenta janelas de vulnerabilidade**
  - Dados são extraídos para serem acedidos por outras aplicações
- **Informação pode ser transformada por algoritmos/aplicações diferentes**
  - Adaptados ao sistema operativo ou à segurança pretendida
  - Complica os processos de recuperação de informação
- **Difícil partilhar ficheiros internos ao pacote cifrado**
  - Pode implicar extrair e tornar a cifrar
- **Exemplos:**
  - PGP, AxCrypt, TrueCrypt, etc.
  - Também… RAR, ZIP, 7zip, LZMA, ...

Nível Aplicacional: TrueCrypt

- **Cria um ficheiro no FS que contém vários volumes**
  - Semelhante a uma imagem de um virtualizador
  - Cifras fortes, em cascata (e.g. AES+Twofish)
  - AES-CBC, depois AES-LRW, depois AES-XTS
  - Chaves criadas com PBKDFS2, SHA-512 e 2000 rounds
- **Suporta Negação Plausível**
  - FSs internos não possuem cabeçalhos óbvios
  - Um ficheiro pode ter um ou mais volumes
    - Não é óbvio determinar quantos volumes existem



Nível Aplicacional: TrueCrypt



XEX with tweak and ciphertext stealing (XTS) mode encryption

Nível dos Sistemas de Ficheiros

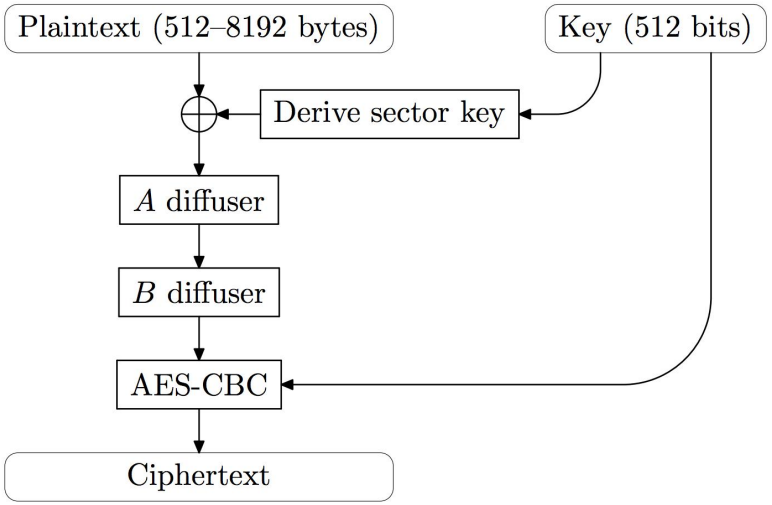
- **Informação é transformada entre a memória e a escrita no volume**
  - Dispositivo físico -> Cache em Memória
    - Sem proteção no caso de servidores (servidor decifrou informação quando lhe acedeu)
    - Mecanismo é mais complexo de implementar em ambientes distribuídos
      - Coordenação com ACLs
      - Partilha das chaves pelo SO
  - Cache -> memória das aplicações
    - Proteção no caso de servidores (é o cliente que decifra)
    - Pode ter lugar fora do ambiente de armazenamento (aplicação, cliente)

- **Exemplos:**
  - CFS (Cryptographic File System)
  - EFS (Encrypted File System)
  - NTFS (NT Filesystem)

Nível dos Volumes

- **Transforma informação a nível do controlador**
  - Transparente para aplicações e quase transparente para o SO
    - requer a existência de um controlador
  - Granularidade do acesso ao nível de um volume inteiro
- **Políticas de cifra definidas ao nível da aplicação ou controlador**
  - Agnóstico do sistema de ficheiros
    - Proteção integral de dados, metadados, ACLs, ...
  - Não permite diferenciação entre diferentes utilizadores
    - Uma das chaves desbloqueia volume
- **Não resolve questões com sistemas distribuídos mas sim de dispositivos móveis**
  - Distribuídos: Volume está acessível ou não, para o mundo
  - Móveis: Protege contra roubo ou perda de equipamento
- **Exemplos:**
  - PGPdisk, LUKS, BitLocker, FileVault

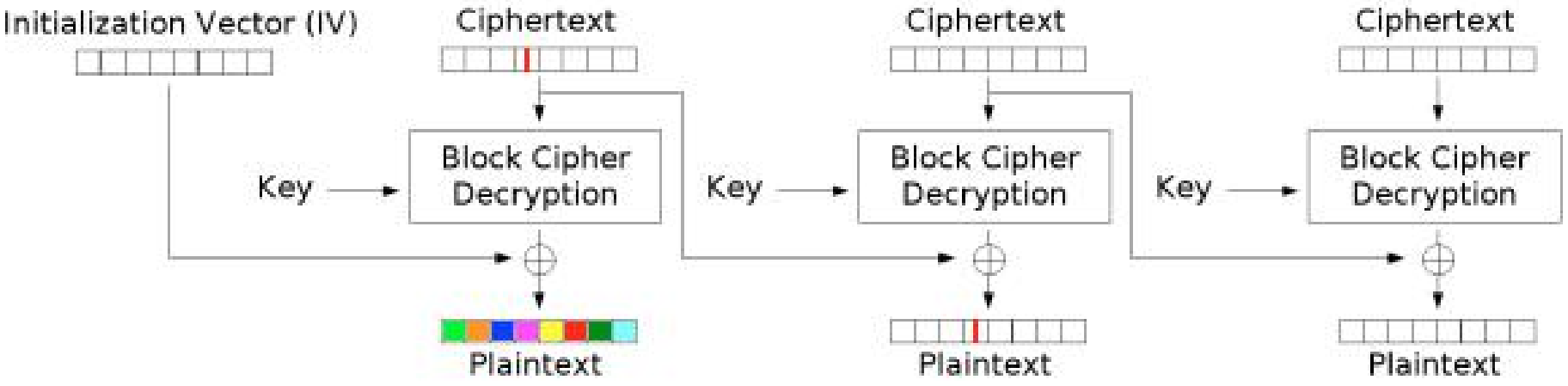
BitLocker (Windows)



- **Cifra um volume inteiro**
  - Utiliza um pequeno volume para iniciar processo de decifra
  - Chave de cifra composta (FVEK): K_{AES} e $K_{Diffuser}$
- **Armazenamento da Chave**
  - FVEK cifrada com Volume Master Key (VMK), cifrada com Key Protector Key
  - Key Protector Key cifrada com senha ou segredo no TPM (recentemente retirado)
- **Processo de Cifra**
  - AES-CBC 128 ou 256, aplicado a cada sector, sem MAC e sem feedback
  - IV = $E(K_{AES}, e(s))$, onde e mapeia o número do sector para um valor de 16bits
  - Sector Key = $E(K_{AES}, e(s)) \mid E(K_{AES}, e'(s))$
    - e' = igual a e mas terminado em 128
  - Elephant Diffuser: Difusor de bits controlado por $K_{Diffuser}$ (entretanto removido)

Bitlocker (Windows)

Malleability attack no CBC



Cipher Block Chaining (CBC) mode decryption

Nível do dispositivo

- **Dispositivo aplica política de segurança internamente**
  - No boot, dispositivo tem de ser desbloqueado
    - Fornecendo as credenciais corretas
  - Cifras implementadas em hardware/firmware

- **Vantagens**
  - Sem perda de performance (grátis)
  - Pode não trivial a extração de informação ou chaves
  - Possível de coordenar o processo com aplicações

- **Desvantagens**
  - Quando o dispositivo é desbloqueado, dados ficam acessíveis
  - Segurança é limitada aos algoritmos presentes
  - Possível presença de erros ou backdoors é difícil de detetar e corrigir

Nível do dispositivo

- **Dispositivos possuem 2 áreas**
  - Shadow Disk: Read Only, ~100MB; Possui software para desbloqueio; disponível
  - Real Disk: Read Write, contém dados; protegido
- **Duas chaves**
  - KEK: Key Encryption Key (Authentication Key)
    - Fornecida pelo utente. Síntese armazenada no Shadow Disk
  - MEK (ou DEK): Media (Data) Encryption Key
    - Cifrada com o KEK
- **Boot Process**
  - Bios vê o Shadow Disk e utiliza-o para iniciar o sistema
  - Aplicação pede senha ao utilizador, decifra KEK e verifica o valor de Hash(KEK)
  - Sucesso: decifra-se MEK para a memória e geometria é atualizada