

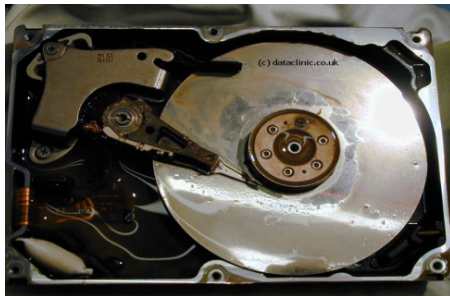
Backups

- Backups – noções gerais
- Soluções integradas

Motivação

“Existem dois tipos de administradores de sistemas: os que já perderam pelo menos um disco, e os que hão de vir a perder...”

-- W. Curtis Preston, *UNIX Backup & Recovery*



Motivação

- A solução de cópias de segurança deve inserir-se no **plano de sobrevivência a catástrofes** da organização.
- A incapacidade de recuperar um sistema quando necessário é **causa frequente de despedimento** de administradores de sistemas!



"You should check your e-mails more often. I fired you over three weeks ago."

Motivação

1. **WHY ?** → Impacto para o negócio. Estimativa de perdas.
2. **WHAT ?** → Volume de dados que devem ser guardados.
3. **WHEN ?** → De dia? À noite? Qual a periodicidade.
4. **WHERE ?** → Para disco? Para Tape?
5. **WHO ?** → Responsabilidades de execução e análise.
6. **HOW ?** → Comandos e procedimentos

Motivação

- Quando é que é necessário recuperar os dados?
 - Perda accidental de dados (p.e. $rm -Rf *$)
 - Falhas de discos
 - Medida MTBF dos discos
 - Erros humanos, por engano ou propositadamente
 - Questões legais (p.e. prova em tribunal)
 - Dados corrompidos por engano, propositadamente ou outros

Os backups têm de ser fiáveis

Plano de contingência (1)

- 1) Quanto tempo pode demorar a recuperar de...
 - uma falha total do sistema?
 - uma falha parcial do sistema?

- 2) Enquanto o sistema está activo será possível...
 - fazer recuperações?
 - fazer cópias de segurança?

- 3) Quais as falhas que provocarão mais perdas?

Plano de contingência (2)

- 4) Quem perde mais por erro humano ou falha hw?
- 5) Quais as prioridades na recuperação?
- 6) Quantas cópias de segurança são necessárias?
- 7) Quanto tempo será necessário manter as cópias?
- 8) Qual o orçamento disponível?

Plano de contingência (3)

Cinco passos que devem constar do plano de contingência:

1) Definir “*Perdas Aceitáveis e Inaceitáveis*”

- Informação que pode ou não ser perdida
- Tempo de recuperação

2) Salvaguardar todos os dados

- Além dos dados, armazenar meta-dados e instruções de recuperação dos dados

3) Organizar e documentar todas as ações

Plano de contingência (4)

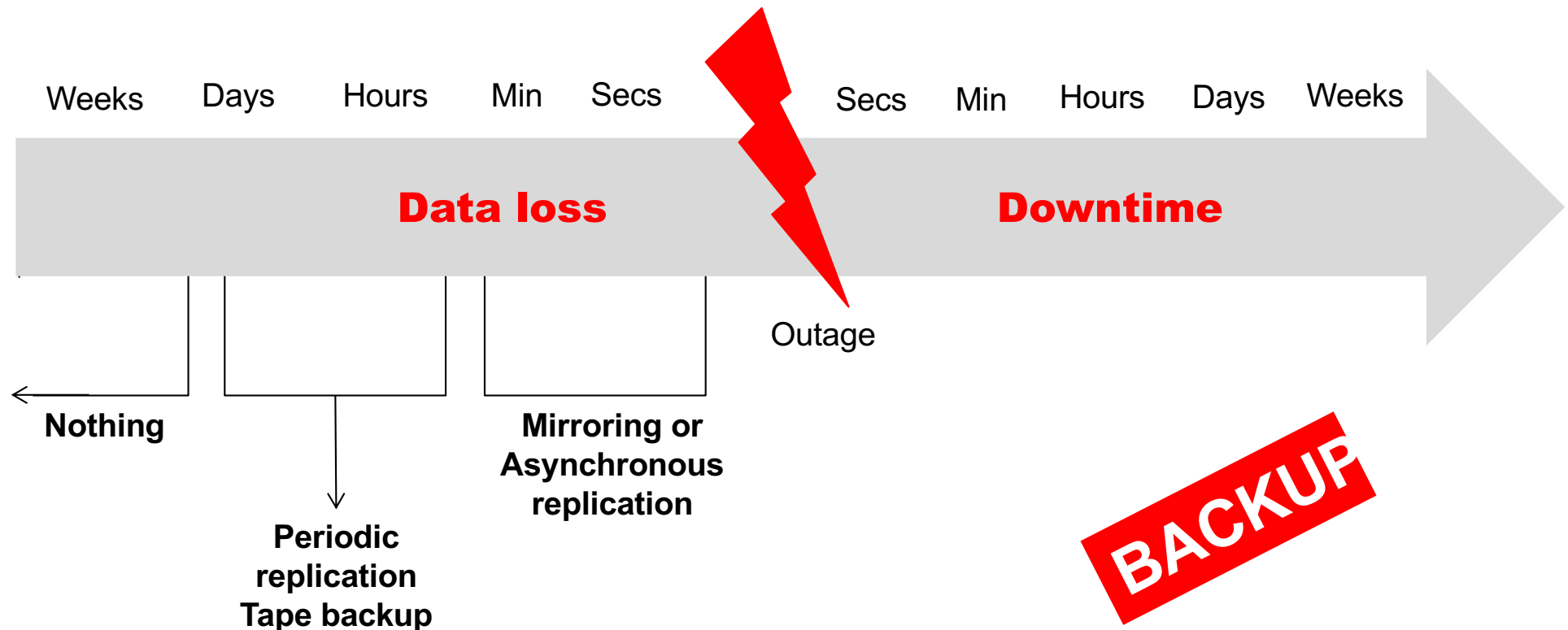
4) Protecção contra todo tipo de desastres

- Os desastres naturais não são os únicos
 - Ex.: ataque ao WTC
 - Protecção anti-fogo e ... anti-água!
 - Roubos...
- O suporte das cópias de segurança também pode falhar!
- Pode haver uma combinação de várias causas!

5) Testar, testar e ... testar!

- Plano de contingência não testado? É apenas uma “proposta”.
- Para prevenir surpresas, testar, testar e testar ainda mais!

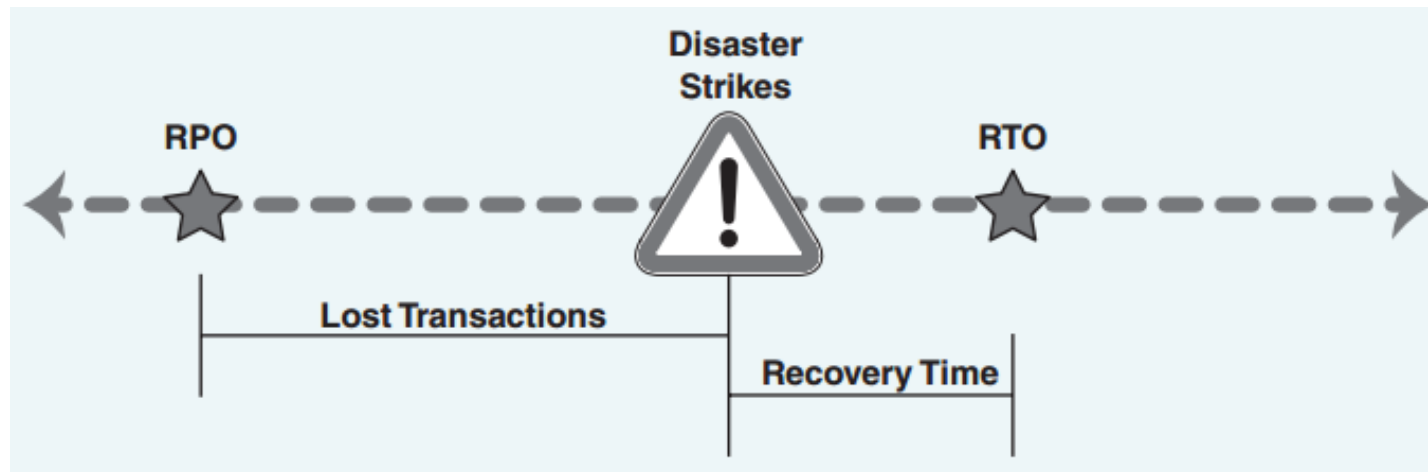
Estratégias de mitigação - dados



Objetivo: minimizar perda de dados e garantir a consistência

Estratégias de mitigação - dados

- Quantos dados está a empresa disposta a perder?
- Ao fim de quanto tempo terá o sistema de estar disponível após o desastre?



RPO = Recovery Point Objective

RTO = Recovery Time Objective

Políticas de backup e SLA

Definir convenientemente:

- Qual a necessidade de fazer backups
- Em que consistem as operações de backup
- Que dados devem ser salvaguardados
- Quais os requisitos legais
- A que horas deverão ser realizados
- Período de tempo usado para a recuperação de dados
- Período de tempo de retenção dos dados
- Método de escalonamento das operações de backup
- Planeamento de capacidade e utilização das tapes

Exemplo de um SLA

“Customers should be able to get back any file with a granularity of one business day for the last six months and with a granularity of one month for the last three years.

Disk failures should be restored in four hours, with no more than two business days of lost data.

Archives should be full backups on separate tapes generated quarterly and kept forever.

Critical data will be stored on a system that retains user accessible snapshots made every hour.”

Tipos de backup

1) Completo

- Cópia completa e integral de todos os ficheiros existentes no disco
- Para restaurar a informação apenas é necessário restaurar a cópia completa
- O tempo de salvaguarda é longo
- Reposição do backup é simples. Implica apenas uma cópia.

Tipos de backup

2) Incremental

- Copia apenas ficheiros modificados/criados desde a última salvaguarda (incremental ou completa)
- Para restaurar → necessário restaurar primeiro a ÚLTIMA cópia completa e TODAS as salvaguardas incrementais
- Salvaguarda é rápida: apenas as modificações são gravadas
- Recuperação dos dados pode ser longa
 - Último backup completo + backups incrementais

Tipos de backup

3) Diferencial (ou integral cumulativo)

- Copia todos os ficheiros modificados/criados desde a última **salvaguarda completa**
- Para restaurar → repôr a última cópia completa e depois a salvaguarda diferencial
- Salvaguarda pode não ser rápida
- A restauração dos dados pode ser medianamente longa
 - Último backup completo + último backup diferencial

Níveis de salvaguarda

Níveis - Definição de níveis 0 até 9

Nível 0

- Backup integral

Nível 1

- Backup **diferencial**

Nível i ($i=2$ até 9)

- Backup **incremental** em relação ao último backup nível $i-1$
- *Backup nível 3* → backup incremental em relação ao último nível 2

Table 5-1: Backup Level

<i>Backup Level</i>	<i>Description</i>
Level 0	A full backup; backs up all files.
Level 1	The first incremental level; gets all files that have changed since the last level 0 backup. It acts like a differential by getting everything that has changed since the level 0/full backup.
Levels 2-9	Backs up whatever files have changed since the next <i>lower</i> level backup. Can work as an incremental if used sequentially, or as a differential if all the same number is used. Can also be "mixed" to create hybrid backup strategies.

Linux Troubleshooting Bible

Christopher Negus and
Thomas Weeks

Níveis de salvaguarda

Exemplos:

Dom.	2ª	3ª	4ª	5ª	6ª	Sab.
Full/0	Full/0	Full/0	Full/0	Full/0	Full/0	Full/0
Full/0	Diff/1	Diff/1	Diff/1	Diff/1	Diff/1	Diff/1
Full/0	Incr/1	Incr/2	Incr/3	Incr/4	Incr/5	Incr/6

Quando? Geralmente, durante a noite

Porquê?

- **Integridade:** Pouca atividade e maior estabilidade do sistema (ficheiros)
- **Velocidade:** Sistema e rede mais libertos para operações de backup.

Proteção contra falha de suporte

1) Possuir dois conjuntos de suporte



2) Ciclo rotativo A + B

- Dois conjuntos completos: A e B
- **A** é empregue nos dias pares, **B** nos ímpares
- Cada arquivo incremental tem os ficheiros dos **dois** últimos dias
- Há redundância no arquivo (ficheiros duplicados nos dois conjuntos) excepto para o último dia

Características do suporte de backup

- 1) **Custo:** precisamos de muito mais espaço de cópias de segurança do que espaço ocupado pelos dados
- 2) **Fiabilidade:** de nada serve uma cópia de segurança se o meio de suporte falhar
- 3) **Velocidade:** suficientemente rápido para permitir concluir a cópia durante o tempo disponível para o efeito
- 4) **Disponibilidade:** não é possível fazer cópias de segurança para CD-R sem um gravador de CDs!
- 5) **Usabilidade:** quanto mais fácil de fazer melhor

Que servidores proteger?

- 1) Cópias de segurança devem ficar num servidor diferente daquele que se pretende proteger
- 2) Fazer cópias de segurança de todas as máquinas numa mesma LAN ou hub para um servidor local
- 3) Fazer cópias de segurança de todos os servidores locais de um edifício para um servidor noutro edifício.

ONSITE *versus* OFFSITE !!!!

Que suporte utilizar?

- **Disquete:** uma cópia integral de /etc **** OBSOLETO****
 - **Drive ZIP:** ~200 MB **** OBSOLETO****
 - **CD-R:** 600 MB **** OBSOLETO****
 - **DVD:** 4.7 GB (ou 9 GB se for double-sided)
 - **HD-DVD** 17 GB p/camada), **Blu-Ray** (25 GB p/camada)
 - **Banda magnética:** 40GB até vários TB (LTO-4/120MB/s)
 - **Discos rígidos:** 500 GB a muitos TB...
 - **Bibliotecas robotizadas:** muitos TB
- } **amovíveis**

Tapes

Drive de tape - *Sun StorageTek T10000C*

- Capacidade até 5 TB
- 252 MB/segundos



Unidades robotizadas (*libraries*)

(+) Grande capacidade de armazenamento

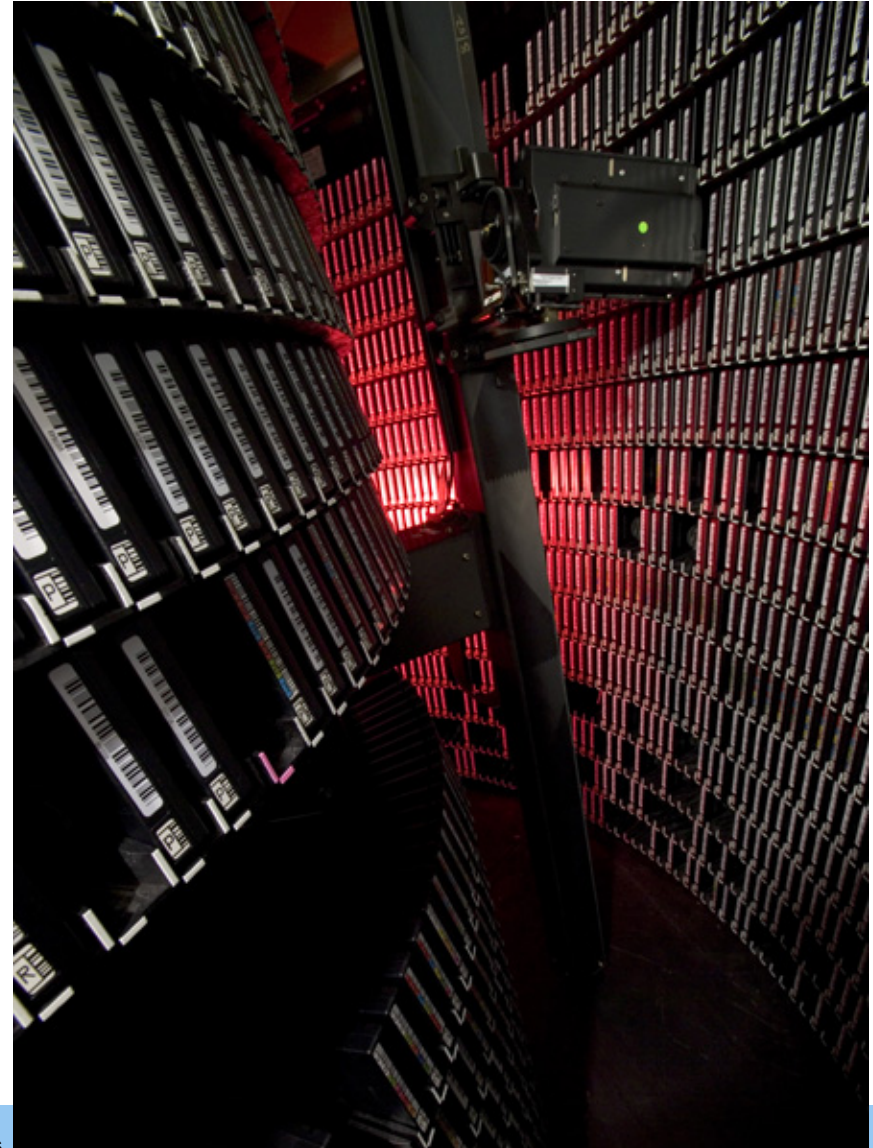
Facilmente atingem vários terabytes

(-) Acesso muito lento

Robô tem que ir buscar a tape

A tape tem que ser lida

(-) Caros!



Unidades robotizadas (*libraries*)

Sun StorageTek SL8500 Modular Library system

- até 100.000 slots
- até 1.000 PB (2:1)
- modular
- 64 drives *T10000C* = 55.3 TB/hora



Soluções integradas opensource

The logo for Amanda, featuring the word "Amanda" in a stylized, cursive font with a rainbow color gradient.The logo for BackupPC, featuring a blue folder icon with a white arrow pointing right, followed by the text "BackupPC" in a bold, sans-serif font.The logo for Bacula, featuring a red dragon head icon on the left, followed by the text "BACULA.ORG" in a bold, sans-serif font.

Soluções integradas

Open source

Linux e/ou Windows

Escaláveis

**Integração com unidades
robotizadas de backup**

Soluções integradas comerciais



<http://www.backupcentral.com/>

Leituras...

- “*Backup and Recovery*”; Curtis, 2006