Agentúra Ministerstva školstva, vedy, výskumu a športu SR pre štrukturálne fondy EÚ

Operačný program VZDELÁVANIE

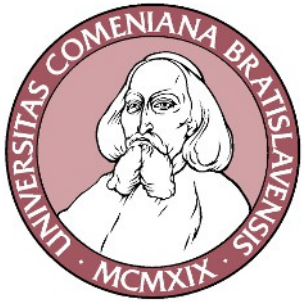
Európska únia
Európsky sociálny fond

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

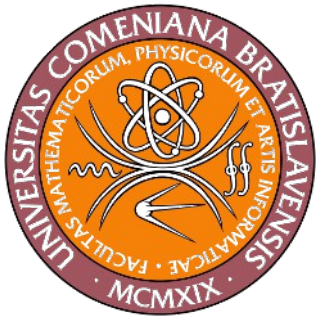
Príprava štúdia matematiky a informatiky na FMFI UK v anglickom jazyku

ITMS: 26140230008

dopytovo – orientovaný projekt



Moderné vzdelávanie pre vedomostnú spoločnosť/Projekt je spolufinancovaný zo zdrojov EÚ



Security of IT infrastructure

Data Availability Protection

RNDr. Jaroslav Janáček, PhD.

Causes of Availability Disruption

- technical failures in storage media
- destruction of entire systems or even rooms or buildings
  - fire, water, earthquake, ...
  - collapse of a building
  - theft
- intentional deletion of data (by an attacker)
- unintentional deletion of data (by mistake, due to a SW error)

Countermeasures for Availability Protection

- data backup
- RAID (Redundant Array of Independent Disks)
- backup data centre in another place
- access control

Data Backup

- backup media
  - magnetic tapes, disks, CD/DVD, ...
- media storage
  - local vs. remote place
- security of backup media
  - confidentiality, integrity, authenticity

Data Backup

- full backup
- incremental backup
  - only changes since the last higher level backup
  - reduces the amount of data
  - number of levels
    - small – newer backups are larger
    - large – more complicated recovery of a chosen state

Data Backup

- example 1 (3 levels)
  - full backup monthly
  - incremental backup weekly against the most recent monthly backup
  - incremental backup daily against the most recent weekly backup
- example 2 (2 levels)
  - full backup monthly
  - incremental daily against the most recent monthly backup
- example 3 (many (30) levels)
  - full backup monthly
  - incremental daily against the previous daily backup

Data Backup

- example 1 (3 levels)
  - recovery from 3 files
  - daily files contain only the changes since the beginning of the week
- example 2 (2 levels)
  - recovery from 2 files
  - daily files contain the changes since the beginning if the month
- example 3 (many levels)
  - recovery from many files
  - daily files contain only the changes since the previous day

RAID

- data blocks are spread across multiple disks
- RAID0 – striping
- RAID1 – mirror
- RAID5, RAID6
- combined RAIDs – RAID10, RAID0+1, ...

RAID0

- provides no redundancy
- increases performance
- min. number of disks: 2
- S=n*s, ef=1

disk 0	disk 1	disk 2	disk 3
0	1	2	3
4	5	6	7
8	9	10	11
12	13	14	15

RAID1

- n disks provide tolerance against failure of any n-1 disks
- min. number of disks: 2
- S=s, ef=1/n
- allows for faster reading

disk 0	disk 1	disk 2
0	0	0
1	1	1
2	2	2
3	3	3

RAID5

- provides tolerance against failure of one disk
- min. num. of disks: 3
- S=(n-1)*s, ef=1-1/n
- P=XOR of all others
- slower writes
- slow in degraded state

disk 0	disk 1	disk 2	disk 3
0	1	2	P
3	4	P	5
6	P	7	8
P	9	10	11

RAID6

- provides tolerance against failure of any 2 disks
- min. num. of disks: 4
- S=(n-2)*s, ef=1-2/n
- P a Q must be independent
  - XOR and Reed-Solomon
- slower writes
- very slow in degraded state

disk 0	disk 1	disk 2	disk 3
0	1	P	Q
2	P	Q	3
P	Q	4	5
Q	6	7	P

RAID0+1

- RAID1 of *m* RAID0 devices of *k* disks
- min. num. of disks: 4
- tolerance against failure of (m-1) RAID0 devices (from (m-1) to k(m-1) disks)
- S=k*s, ef=1/m
- higher performance

disk 0	disk 1	disk 2	disk 3
0	1	0	1
2	3	2	3
4	5	4	5
6	7	6	7

for m=2, k=2

- failure of 1 disk is OK
- failure of 2 disks is OK with the probability 1/3

After the failed disk is replaced, *k* disks have to be overwritten.

RAID10 = RAID1+0

- RAID0 of *k* RAID1 devices of *m* disks
- min. num. of disks: 4
- tolerance against failure of (m-1) to k(m-1) disks
- S=k*s, ef=1/m
- increases performance

disk 0	disk 1	disk 2	disk 3
0	0	1	1
2	2	3	3
4	4	5	5
6	6	7	7

pre m=2, k=2

- failure of 1 disk is OK
- failure of 2 disks is OK with the probability 2/3

After the failed disk is replaced, 1 disk has to be overwritten.

RAID Paradoxes

- as the disk capacity increases, it becomes more real that another disk in RAID5 fails during the reconstruction of the array after the failure of the first one
  - RAID6 is a partial solution, however, it only postpones the problem a bit to the future
  - that's why the popularity of RAID10 and RAID50 (RAID0 build of RAID5 devices) increases

Backup Data Centre

- offline replication
  - regular synchronization of data from the primary data centre to the backup data centre
- online replication
  - online synchronization of data from the primary data centre to the backup data centre (e.g. RAID1 across a network)

Summary

	backup	RAID	backup data centre	access control
technical failure	partially	yes (within limits)	partially – yes	no
system destruction	partially (if backup is in a remote place)	no	yes	no
unintentional deletion	yes (if detected early)	no	no – partially	partially
intentional deletion	yes (if detected early)	no	no – partially	yes