

Inhalt

Vorwort	11
---------------	----

1 Risky Business – ausfallsichere Server 17

1.1 Grundsätzliche Überlegungen zur Redundanz	17
1.1.1 Parallelität, MTBF, MTTR und einiges mehr	19
1.2 Tool-Time	24
1.2.1 Päckchen... ..	25
1.2.2 ... und Betriebsanleitungen	26

2 Lokale Hochverfügbarkeit 27

2.1 Netzteile, CPUs und mehr	29
2.1.1 Redundante Netzteile	30
2.1.2 Kernfrage, Teil 1: CPU	31
2.1.3 Kernfrage, Teil 2: Control Groups und Cpusets	33
2.1.4 monit	40
2.1.5 Fessel-Spielchen – Netzwerkkarten-Bondage	44
2.2 Lokaler Storage	52
2.2.1 The road ahead – (un-)solide SSD's?	57
2.2.2 Disk-Überwachung: Clever mit SMART	58
2.3 RAID	64
2.3.1 Raid-Level	65
2.3.2 Exkurs: Raid-Kombilevel	71
2.3.3 Hard- oder Software-Raids	71
2.3.4 Softraids unter Linux	75
2.3.5 Sofraid-Performance	77
2.3.6 Einsatzzweck und Bootloader	78
2.3.7 Weitere Vorbetrachtungen zum Raid-Setup	79
2.3.8 Exkurs: Journalisten – Journaling-Dateisysteme unter Linux	81
2.3.9 Mirror, Mirror – Teil 1: Sofraid Level 1	98
2.3.10 Dynamisches Duo – Sofraid Level 1 + 0 und 10	108
2.3.11 Sofraid Level 5	110
2.3.12 Sofraid Level 6	115
2.3.13 Partitionable Raids	116
2.3.14 Bootredundante Arrays für das OS	118

2.3.15	Raid-Upgrade: Single-Disk mit installiertem Linux-OS zum bootredundanten Softraid Level 1 aufbohren	119
2.3.16	Weitere wichtige Raid-Parameter und -Szenarien	132
2.4	LVM – Storage-Virtualisierung mit dem Logical Volume Manager	138
2.4.1	How it works	139
2.4.2	Logische Schaltzentrale – lvm.conf	141
2.4.3	Let's get physical – Erstellen und Administrieren der PVs	142
2.4.4	Gruppentherapie – Administration der Volume Groups	145
2.4.5	Administration der Logical Volumes	146
2.4.6	Bitte lächeln – Snapshots	151
2.4.7	Spieglein, Spieglein ... LVM-Mirror statt unterliegendem RAID	160
2.4.8	Just me, myself and I – lineares LV zu einem Mirror erweitern	162
2.4.9	Auto-Rebuild von defekten PVs im LVM-Mirrorset	163
2.4.10	LVM für das root-Dateisystem	165

3 HA auf Netzwerkebene 169

3.1	How it works – Clustertypen	169
3.1.1	Hochverfügbarkeits-Cluster	170
3.1.2	Load-Balancing-Cluster	170
3.1.3	High-Performance-Computing-Cluster	170
3.1.4	Active/Passive – (Hot-)Failover-Cluster	172
3.1.5	Active/Active-Cluster	173
3.1.6	Cluster-Reigen	174
3.1.7	Cluster-Lobotomie: Split-Brain, Fencing und Stonith	174
3.1.8	Volksentscheid: Cluster-Quorum	177
3.1.9	Exkurs: Alles oder nichts – Shared-all-/Shared-nothing- Cluster	178
3.2	HA-Basics	178
3.2.1	Alles eine Frage der Kommunikation: konzeptionelle Cluster-Übersicht	181
3.2.2	Das technische (Schichten-)Modell von Heartbeat OpenAIS/Corosync und Pacemaker	182
3.2.3	Cluster-Kommunikation am konkreten Beispiel	184
3.3	Clustersoftware	185
3.3.1	Heartbeat	185
3.3.2	OpenAIS/Corosync	186
3.3.3	Pacemaker	188

3.3.4	Upgrade der Cluster-Software und einzelner Applikationen	190
3.4	NTPD – The Time Machine	192
3.4.1	Exkurs: Virtual Time Machines	196
3.5	Setup der Cluster-Kommunikation	200
3.5.1	OpenAIS/Corosync-Setup	201
3.5.2	Heartbeat-Setup	211
3.6	First Contact	216
3.6.1	Zählerstände, Klebrigkeiten, Regeln, Platzanweiser und Rangordnungen	217
3.6.2	Konfigurations-Layout	221
3.6.3	Die CIB-Files	222
3.6.4	Die crm-Shell	223
3.6.5	cibadmin und crm_*-Tools	231
3.6.6	Die Pacemaker-GUI	235
3.6.7	HAWK	238
3.6.8	DRBD MC	239
3.7	Management von Cluster-Ressourcen	242
3.7.1	At your Service – gestatten: Ihre Service-IP	242
3.7.2	Integration von Ressourcen	243
3.7.3	Örtlichkeiten – Setzen von Location-Constraints	245
3.7.4	Failover-Simulation und Klebrigkeiten	246
3.7.5	Schattenspielchen im Sandkasten – Testsimulation mit crm_shadow	247
3.7.6	CIB-Templating	251
3.7.7	Ressourcen verwalten	253
3.7.8	Ressourcen und Constraints löschen, Integritätsprüfung der CIB	254
3.7.9	Einfache Ressourcen-Integration am Beispiel von Apache	256
3.7.10	The Clone Wars – Hot-Standby mit Clonesets	261
3.7.11	Migration-Threshold und Failure-Timeout	264
3.7.12	Voll vernetzt – die ping-/pingd-Ressource	266
3.7.13	Failover einer Samba-3-Ressource	269
3.8	Ausfallsichere Shared-Nothing-Cluster mit DRBD	275
3.8.1	Wozu – und wozu nicht: Einsatzmöglichkeiten von DRBD	277
3.8.2	Die DRBD-Funktionalität im Detail	278
3.8.3	Rollenspielchen – DRBD-Modi	279
3.8.4	Replikations-Varianten	279
3.8.5	DRBD-Standalone-Setup	281

3.8.6	Exkurs: drbd-overview und /proc/drbd	289
3.8.7	Manueller DRBD-Funktionstest (Master/Slave)	293
3.8.8	Konfiguration der DRBD-Master/Slave-Ressource im Cluster	294
3.8.9	Exkurs: DRBD on top of LVM – Backup des Secondary per Snapshot	299
3.8.10	Exkurs: LVM on Top of DRBD (Master/Slave)	302
3.8.11	DRBD-Dual-Primary-Mode	304
3.8.12	Exkurs: DRBD-Split-Brain-Recovery	305
3.8.13	Exkurs: Manuelles DRBD-Split-Brain-Recover	308
3.8.14	Cluster-Dateisysteme	310
3.8.15	DRBD Dual Primary mit OCFS2	313
3.8.16	DRBD Dual Primary mit GFS2	328
3.8.17	DRBD Dual Primary mit OCFS2 und CLVM	330
3.8.18	Three-Node-DRBD für Disaster Recovery	337
3.8.19	DRBD-Online-Device-Verification	348
3.8.20	DRBD Replication Traffic Integrity Checking	349
3.9	iSCSI im Cluster	352
3.9.1	iSCSI-Basics	352
3.9.2	Setup der iSCSI-Ressourcen (DRBD im Primary/Secondary)	353
3.9.3	Einrichtung des iSCSI-Initiators	356
3.9.4	Setup der iSCSI-Ressourcen (Hot Failover mit DRBD im Dual-Primary)	359
3.10	Exkurs: Node-Fencing mit STONITH und/oder Watchdog	361
3.10.1	Stonith-Setup	362
3.10.2	Exkurs: Passwortlose ssh-Key-Autorisierung	363
3.10.3	Integration der external/ssh-Stonith-Ressourcen in unseren Cluster	365
3.10.4	Watchdog	366
3.10.5	SBD – Stonith per Split-Brain-Detector	366
3.11	Debugging im Cluster	369

4 Virtualisierung im Cluster 373

4.1	Virtualisierungskonzepte – oder: Die wundersame Welt der Zwiebel	374
4.2	XEN	376
4.2.1	Xen-Terminologie	379
4.2.2	Xen-Setup	379
4.2.3	Installieren einer Xen-DomU	383

4.2.4	Manuelle Live-Migration von Xen-DomUs	388
4.2.5	Exkurs: Snapshots/Backups für Xen-DomUs und Blockdevice-Formate	395
4.2.6	Monitoring von Xen-DomUs im Cluster	397
4.2.7	Remus	398
4.3	KVM/qemu	399
4.3.1	KVM-Setup	400
4.3.2	KVM-Netzwerksetup	400
4.3.3	KVM-Gast manuell erzeugen	400
4.3.4	KVM-Live-Migration	401
4.3.5	Backup/Snapshots von KVM-Gästen	404
5 Backup und Disaster Recovery		407
5.1	Analyse	408
5.2	Umsetzung	409
5.2.1	Kategorie 1: Backup und Recover des reinen Datenspeichers	409
5.2.2	Kategorie 2: Backup und Recover des Systems	415
5.2.3	Kategorie 3: Backup und Recover von virtuellen Maschinen	419
Anhang		421
A.1	Beispieldateien	421
A.2	Paketlisten	421
A.3	Manpages	431
Index		443