

Backup mit Linux und Amanda

Johannes Nieß

7. November 2000

Warum Backups?

- Daten kosten Arbeitszeit zur Erstellung
- Erstellte Daten müssen immer verfügbar sein
- Ausfall des Servers ist teuer

50 Mitarbeiter * 50 DM/h * 8 h = **20 000 DM**

Backup-Hardware

	Kapazität	Geschw.	Medien	Laufwerk
	GB (unk.)	MB/sec	DM/GB	DM
ZIP-Disk	0,25	0,5	60	200
CD-RW	0,65	0,6	25	400
Festplatte	20	10	15	300
Floppy-Streamer	0,12	0,03	125	†
Travan 5	20	1	10	1000
DAT DDS2	4	0,5	7	1000
DAT DDS3	12	0,8	4	1700
DAT DDS4	20	3	3,5	2500
DLT7000	35	5	4	9000

Besonderheiten von Tapes und Backups

- Bänder haben kein Dateisystem. Der Zugriff erfolgt zeichenweise (Char-Device, wie z. B. Modems). Informationseinheiten („Dateien“) werden durch feste Bytefolgen („Filemarks“) getrennt.
- Bänder können nur vorwärts und rückwärts gespult werden. Es existieren für jedes Bandlaufwerk ein „rewinding device“ (Band wird nach dem Schliessen jeder „Datei“ zurückgespult) und ein „non-rewinding device“ ohne diese Eigenschaft (/dev/nst0).
- Um weniger Daten speichern zu müssen, werden oft nur die Unterschiede zur letzten Sicherung auf Band geschrieben. „Level 0“ ist ein volles Backup; ein „Level 1-Backup“ sichert die Änderungen gegenüber Level 0 usw. . Ein „Level 2“ benötigt beim Zurückspielen drei Bänder: Level 0, 1, 2 .

Einfache Backup-Software

dump Dateisystemspezifisches Sichern von Partitionen

tar Portables Sichern von Verzeichnissen

cpio Grundlage von RPM-Paketeten

aio Advanced File Input Output

taper Textbasiertes Menüsystem für dump und tar

kdat KDE-basiertes Interface für Backups

Netzwerkfähige Backup-Software

Arkeia Kommerzielles Backup unter Linux, netzwerkfähig?

Legato Paralleles Backup von Windows, Netware, Unix, Linux

Tivoli (ADSM) In der Uni eingesetztes System von IBM

Amanda GPL-Backup-System für Unix (und Windows)

Advanced MArylaNd Disk Archiver



Amanda

Autoren u. a:

James da Silva: Erstentwickler, Univ. of Maryland, MD, USA

Alexandre Oliva: u. a. Samba-Entwickler, angestellt bei Redhat

John R. Jackson: Entwickler, sehr aktiv in der Mailingliste

Einsatz:

John R. Jackson, Purdue-Univ., Fachbereich Informatik:

“Ich bin froh, wenn meine Backups in 24h fertig sind”

Institut für Landtechnik: 20 GB, DAT DDS3

Johannes Nieß zu Hause: Desktop und Laptop, Floppy-Tape

Eigenschaften von Amanda

- Betriebssysteme: Unix, Linux, Windows (smbtar über Unix)
- Unterstützung von Wechsler-Systemen
- Daten-Kompression: Software (Client), Software (Server), Hardware (Tape), ohne
- Selbständige, bedarfsangepasste Planung der Backup-Level
- Statusmeldungen per E-Mail
- Software-Labels auf den Bändern
- Zwischenspeicherung der Daten auf Server-Festplatten
- Primitive Beschränkung der Netzwerkbelastung
- Restore mit primitivsten Mitteln möglich
- Interaktives Restore mit FTP-ähnlicher Syntax

Nachteile von Amanda

- keine Block-Devices (z. B. CD-RW) als Backup-Medium
- nur ein Tag pro Band (Grund: Datensicherheit)
- Eine Partition kann nicht auf mehrere Bänder aufgeteilt werden
- kein grafisches Bedien-Interface
(Kommandozeilen-Tool vorhanden)
- Firewalls zwischen Server und Client sind problematisch
- Selbstkompilation empfohlen

Ablauf eines Backups (vereinfacht)

1. Server startet `amdump`-Prozess
2. Clients schätzen Platzbedarf für letzten Backup-Level
3. Server errechnet optimale Backup-Level für Bandausnutzung und Priorität mittels Client-Schätzung und Datenbank alter Kompressionsfaktoren und Backup-Geschwindigkeiten
4. Client schickt Dateien des befohlenen Levels zum Server
5. Server speichert Dateien auf Holdingdisk
6. Vollständige Backups werden auf Band geschrieben

Allgemeine Installation

1. Quellcode holen (<http://www.amanda.org>)
2. entpacken (z. B. über Midnight Commander)
3. auf die Rechner-Plattform abstimmen lassen
(`./configure` ; Optionen mit `--help` anzeigen)
4. Übersetzungsdatei (Makefile) editieren
5. Kompilieren (`make`)
6. erzeugte Dateien kopieren (`make install`)

Amanda-spezifische Installation

1. Server konfigurieren
2. Client-Partitionen zum Backup eintragen
3. Services eintragen in `/etc/inetd.conf` (Server und Clients)
4. nach Bedarf andere Programme patchen / aktualisieren
(tar, samba, dump (Linux))
5. `amcheck config`
6. Verzeichnisse erstellen (Fehlermeldungen abwarten)
7. Restore üben
8. ganz am Schluß: regelmässige Ausführung starten
(`crontab -e`)

Konfiguration des Servers: amanda.conf

Allgemeines

org Name Name der Konfiguration: taeglich, archiv etc.

mailto email Adresse(n) für Berichte: root@localhost

dumpcycle tage Volles Backup mindestens alle tage: 7

runspercycle tage Backup-Durchläufe pro Zyklus: 5

tapecycle zahl Anzahl der Bänder in der Konfiguration:
6 oder mehr

tapedev Band-Device Non-Rewinding!: /dev/nst0

labelstr regexp Namen der Software-Labels auf den Bändern:
taeglich*

tapetype Bandtyp Benutzte Bandtyp-Definition

Sektionen in amanda.conf

Dumptype Programm (tar/dump), Komprimierung (Client/Server; fast/best/none), Priorität (low/medium/high), Ausschliessung von Dateien eines Backup-Typs

Holdindisk Verzeichnis und Größe der Holdingdisk(s)

Tapetype Beschreibung von Band-Länge, -Geschwindigkeit und Länge der Filemarks. Bei unbekanntem Laufwerk: tapetype.c kompilieren, ausführen und warten.

Interface Beschreibung der Netzwerk-Interfaces und nutzbarer Bandbreiten

Konfiguration der Clients: disklist

Eine Zeile beschreibt eine Backup-Einheit, Trennung der Einträge mit Tabulator

- Rechnername: r50.landtechnik.uni-bonn.de
Unix-Samba-Rechner bei Windows-Backups!
- Verzeichnis bzw. Partition: /home
//Windows-Rechner/Freigabe; Passwort in /etc/amandapass
- Dumptype: client-best-tar
aus amanda.conf
- Spindle
Kontrolliert gleichzeitige Backups von diesem Client
- interface
aus amanda.conf

Ausgabe von amstatus

Using /var/lib/amanda/taeglich/amdump

```

server:/                1      96k finished
server:/usr             0 485487k wait for dumping
server:/mnt/windoofs-data 1    1696k finished
server:/var            1  23618k wait for dumping
server:/home          3  66421k wait for dumping
server:/mnt/root-save  0  14079k dumping    4192k ( 29.77%)
server:/var/log       1   1216k finished
r50:/                 1   5413k dumping     32k (  0.59%)

```

```

SUMMARY      part      real estimated
              size      size

partition    :    8
estimated    :    8          599235k
failed       :    0              0k
wait for dumping:    3          575526k
dumping to tape :    0              0k
dumping      :    2    4224k    19492k
dumped       :    3    3008k    4217k
wait for writing:    0         0k         0k
writing to tape :    0         0k         0k
failed to tape :    0         0k         0k
taped        :    3    3008k    4217k
all dumpers active
taper idle

```

Mail-Meldung nach erfolgreichem Backup

Subject: taeglich AMANDA MAIL REPORT FOR September 14, 2000

These dumps were to tape taeglich12.

Tonight's dumps should go onto 1 tape: taeglich7.

STATISTICS:

	Total	Full	Daily	
	-----	-----	-----	
Dump Time (hrs:min)	1:04	0:45	0:03	(0:08 start, 0:08 idle)
Output Size (meg)	963.9	828.7	135.2	
Original Size (meg)	2182.0	1778.7	403.2	
Avg Compressed Size (%)	44.2	46.6	33.5	
Tape Used (%)	8.0	6.9	1.1	(level:#disks ...)
Filesystems Dumped	8	3	5	(1:4 2:1)
Avg Dump Rate (k/s)	308.5	313.6	280.6	
Avg Tp Write Rate (k/s)	341.5	314.5	719.1	

NOTES:

planner: Forcing full dump of r50:/ as directed.

taper: tape taeglich12 kb 987296 fm 8 [OK]

DUMP SUMMARY:

HOSTNAME	DISK	L	ORIG-KB	DUMPER STATS			TAPER STATS		
				OUT-KB	COMP%	MMM:SS	KB/s	MMM:SS	KB/s
r50	/	0	1771290	832736	47.0	44:37	311.1	44:37	311.1
server	/	0	37590	14624	38.9	0:25	582.0	0:18	800.1
server	/home	2	267560	120832	45.2	4:08	487.3	2:27	820.7
server	/mnt/root-save	1	190	32	16.8	0:01	45.1	0:20	3.2
server	-windoofs-data	1	10900	960	8.8	0:31	30.7	0:02	591.1
server	/usr	1	6910	928	13.4	2:38	5.9	0:03	352.7
server	/var	1	127360	15712	12.3	0:56	282.7	0:21	760.6
server	/var/log	0	12530	1216	9.7	0:04	303.8	0:02	518.8

(brought to you by Amanda version 2.4.1p1)

interaktives Restore einzelner Dateien

```
cd Verzeichnis
```

```
amrecover -d Band-Device
```

```
sethost Client
```

```
setdisk Disklist-Eintrag
```

```
setdate YYYY-MM-DD
```

```
cd unter/ver/verzeichnis
```

```
ls
```

```
add Datei oder Verzeichnis
```

```
list
```

```
extract
```

Bänder nach Aufforderung einlegen

Restore ganzer Disklist-Einträge mit Amanda

```
amrestore Band-Device hostname diskname
```

von einer Notfall-Diskette

Band positionieren (mit mt)

```
dd if=Band-Device bs=32k skip=1 |zcat>Verzeichnis
```

wichtige Kommandos

- `amlabel config band-name`
Band mit Software-Label versehen
- `amcheck config`
Test von Clients (Erreichbarkeit)
und Server (richtiges Band eingelegt und beschreibbar)
- `amdump config`
Backup per Hand
- `amflush config`
Manueller Transfer von Holdingdisk auf Band
- `amadmin config befehl`
Administrations-Tool
 - `force disklist-Eintrag`: Volles Backup erzwingen
 - `tape`: Nächstes Band herausfinden

Versionen von Amanda

2.4.1p1 Stabile Version vom 21. Nov. 1998

2.4.2 Beta-Version, lt. Mailingliste sehr stabil

2.5 Entwickler-Version

Implementierung von

- Dumper-API: Nutzung anderer Programme (afio, bzip-Kompression, Anhalten/Kopieren von Datenbanken)
- Änderungen im Protokoll: ausschließlich TCP
- Unterstützung verschlüsselter Datenübertragung (z.B. SSH)

Hilfe und Informationen

- Readmes, Textdateien und “manpages” aus dem Amanda-Quellcode-Paket
- <http://www.amanda.org/> Homepage des Projektes
- sourceforge CVS-Repository für neueste Entwickler-Versionen und Patches
- [mail:amanda-users@amanda.org](mailto:amanda-users@amanda.org) Mailingliste für Probleme (empfehlenswert)
- [mail:amanda-hackers@amanda.org](mailto:amanda-hackers@amanda.org) Mailingliste der Entwickler
- [mail:bolug@uni-bonn.de](mailto:bolug@uni-bonn.de) “unsere” Mailingliste
- [mail:j.niess@uni-bonn.de](mailto:j.niess@uni-bonn.de) Meine persönliche Adresse