

# Archiv & Backup



Andreas Gerdes

[gerdes@rrzn.uni-hannover.de](mailto:gerdes@rrzn.uni-hannover.de)

# Inhalt

- 1 Definitionen: Archivierung, Backup und Restore
- 2 Festplatten - RAID (Cache) und und Magnetbandroboter
- 3 Zugang zum Archivsystem
- 4 Backup als zentraler Service
- 5 Von Netbackup zu TiNa

## Einordnung: Backup & Restore, Archiv

Umstrukturierung des Rechenzentrums nach ITILv3  
(siehe Vortrag *IT-Reorg-Umsetzungsprojekt* von gestern)

Beide bereiche gehören zum AP-Bereich 3 (Services)

- Start eigentlich 2013-01
- AP *Backup & Restore* vorgezogen auf 2012-01
- Zur Zeit wird ein Zeitplan zur Migration (der Institute) erstellt
- Vorbereitungen des SLA in Zusammenarbeit mit dem zuständigen FIO

# Lebensdauer von Festplatten

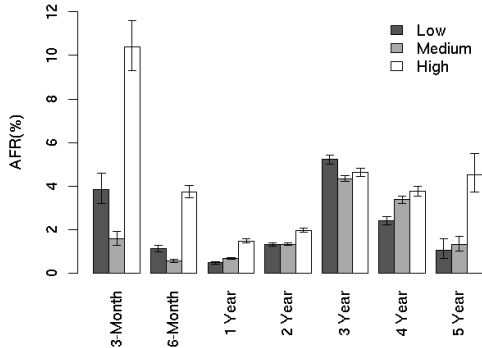


Figure 3: Utilization AFR

Quelle:

Eduardo Pinheiro, Wolf-Dietrich Weber, Luiz André Barroso (Google Research):  
*Failure Trends in a Large Disk Drive Population*

5th USENIX Conference on File and Storage Technologies (FAST 2007), pp. 17-29.

# Archivierung

## Definition

Daten sammeln, klassifizieren (sortieren, aufräumen) und (langfristig + sicher) abspeichern.

Eigenschaften / Ziele:

- nicht automatisiert
- einfache Wiederherstellung konsistenter Daten
- (komplett oder teilweise)
- gelegentlicher Vorgang
- einfaches Auswahl-/ Verwaltungswerkzeug

# Backup

## Definition

(Kurzfristige) Speicherung von Daten zur schnellen Wiederherstellung im Falle des Datenverlustes.

Eigenschaften / Ziele:

- möglichst automatisiert mit Möglichkeit des manuellen Eingreifens
- einfache schnelle und vollständige Wiederherstellung aller (verlorengegangenen) Daten
- (komplett oder teilweise)
- regelmäßiger Vorgang
- simples Scheduling
- einfaches Auswahl-/ Verwaltungswerkzeug

# Restore

## Definition

Wiederherstellung von Daten (zu einem bestimmten Zeitpunkt in der Vergangenheit)

Eigenschaften / Ziele:

- nicht automatisiert
- einfach und nutzerfreundlich
- (komplett oder teilweise)
- gelegentlicher Vorgang
- wird gewöhnlich vom Benutzer angestoßen

# System am Rechenzentrum: Bilder





## System am Rechenzentrum: Zahlen

- Bandroboter : Quantum Scalar 10K, LT04 Bänder und Laufwerke
- (zur Zeit Migration auf LT05)
- Band-Füllstand (zur Zeit): ca. 2.5 PB
- Cache: Nexsan RAID 20 TB, kommende Erweiterung um 36 TB
- 8880 Stellplätze für HLRN und Uni
- Monatliche Fullbackups werden 3 Monate aufgehoben
- Inkrementelle Backups jeden Tag werden 6 Wochen aufgehoben
- Brand-, wasser- und explosionsgeschützte Zelle

## Archiv: Mechanismus, Grenzen

- Zugangsaccount (6er Netz) nötig
- Über Zugangsrechner zen via LFTP
- Oder aus dem Uni-Netz mit Filezilla
- Zur Zeit 500 (+200) GB Quota pro Nutzer
- Möglichst wenige (große) Dateien (auch hier Quota)
- Details: <http://www.rrzn.uni-hannover.de/archiva3.html>
- Daten werden 8 Jahren aufgehoben und migriert auf neue Medien

## Archiv: Mechanismus, Grenzen

- Zugangsaccount (6er Netz) nötig
- Über Zugangsrechner zen via LFTP
- Oder aus dem Uni-Netz mit Filezilla
- Zur Zeit 500 (+200) GB Quota pro Nutzer
- Möglichst wenige (große) Dateien (auch hier Quota)
- Details: <http://www.rrzn.uni-hannover.de/archiva3.html>
- Daten werden 8 Jahren aufgehoben und migriert auf neue Medien

### AP 3.12:

- Start: 1. Quartal 2013
- Verdoppelung der Quota

## Backup: (noch) Symantec Netbackup

- Historisch gewachsen
- Teilweise Server, teilweise Einzelplatz-PCs
- 1 TB pro Client
- 1x pro Monat: Fullbackup
- 1x pro Woche: Cumulative Differential Backup
- 1x pro Tag: Inkremental Differential Backup
- Details: <http://www.rrzn.uni-hannover.de/archivb1.html>

## Backup: (noch) Symantec Netbackup

- Historisch gewachsen
- Teilweise Server, teilweise Einzelplatz-PCs
- 1 TB pro Client
- 1x pro Monat: Fullbackup
- 1x pro Woche: Cumulative Differential Backup
- 1x pro Tag: Inkremental Differential Backup
- Details: <http://www.rrzn.uni-hannover.de/archivb1.html>

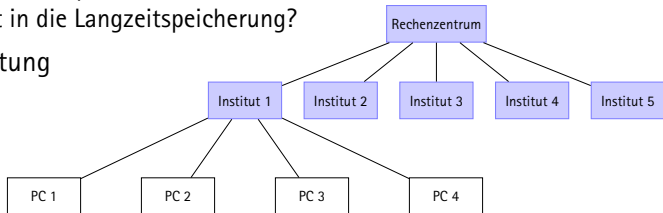
### Probleme:

- Speichermengen entwickeln sich schneller als Netzwerk-Zugänge
- ⇒ Der Service wird sich ändern!

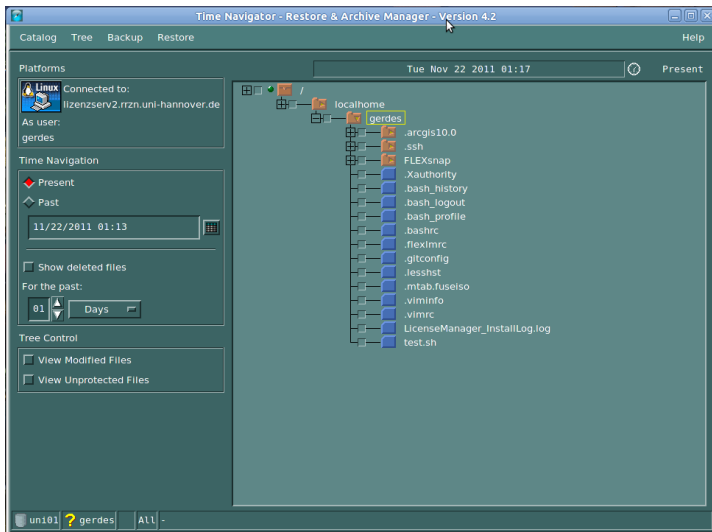
# Backup: Service mit TiNa

Software: Atempo Time Navigator

- Pro Institut wird nur noch der(die) zentrale(n) Server gesichert.
- Pro Maschine max 2 TB
- Verschlüsselte Datensicherung wird eingeführt (auf Wunsch)
- Sensibilisierung der User: Man muss sich Gedanken machen
  - Was soll ins Backup (auf den Server des Instituts)?
  - Was gehört in die Langzeitspeicherung?
- SLA in Vorbereitung



# Backup: TiNa Restore Manager



## Verschlüsselung und Zertifikate

- Time Navigator verwendet unter Linux/Unix einen Zufallszahlengenerator (/dev/random or /dev/urandom), unter Windows eine Kryptographie-API
- Integrierte Verschlüsselung (NED und MED)
- Zertifikate nach X.509
- Integriertes Schlüssel-Management
- Zertifikate hierarchisch organisiert: ausgestellt vom Elternzertifikat oder selbst unterschrieben
- Private Schlüssel passwortgeschützt
- Passwort für Backup auf den Server nicht nötig, nur für Restore, Signierung von Kind-Zertifikaten oder andere sicherheitsrelevanten Operationen



## Fazit

Backup wird als Service des zentralen IT Dienstleisters neu aufgestellt.

- Es gibt Grenzen (Netz, Speichermenge)
- Es gibt auch noch das Archiv!
- Ein Umdenken über Datenspeicherung wird nötig
- Für Institute mit großen Wünschen gibt es Alternativen:
  - RAID - System (lokal oder bei uns)
  - ⇒ Serverhousing (Service 3.7)
- Wir finden eine Lösung!

Vielen Dank für die Aufmerksamkeit!