Hier finden sich meine Notizen zur Vorbereitung eines Workshops zum Thema Datensicherung/Datensicherheit. Ergänzungen und Hinweise bitte hier im Wiki oder direkt an mich RalphGL

Die S5-Präsentation für den Vortrag ist noch nicht fertig aber hier.

Die Gefahr des Datenverlusts und Vorkehrungen zum Schutz

Ein paar Überlegungen zur Gefahr des Datenverlusts.

Welche Gefahren und Risiken drohen?

- Hardwarefehler z.B. Ausfall einer Festplatte oder eines Controllers
- Benutzerfehler (versehentliches Löschen von Daten und Sabotage)
- Softwarefehler (bzw. Schadsoftware, Cracker etc.)
- höhere Gewalt (Brand, Wasser, Blitzschlag)
- Diebstahl (Einbruch)

Welche Maßnahmen taugen gegen welche Gefahren, gegen welche nicht?

Mit unendlichem Aufwand lässt sich JEDES Risiko minimieren. Kein Aufwand bedeutet MAXIMALES Risiko. In jedem Einzelfall gilt es eine Abwägung zu treffen und das **Optimum zwischen Risiko und Sicherheit** unter **Berücksichtigung der entstehenden Kosten (des Aufwandes)** zu beurteilen.

manuelle Sicherung durch Benutzer

Ergebnis: leider völlig unbrauchbar, wird erfahrungsgemäß nicht dauerhaft regelmäßig durchgeführt, oft unvollständig und zu zeitaufwändig;

automatische Sicherung durch Bandlaufwerk:

Bedingt geeignet, allerdings müssen Prüfverfahren für die Sicherungsmedien implementiert werden. Es besteht die Gefahr dass die Datensicherung im Bedarfsfall nicht brauchbar ist. Schützt nicht vor höherer Gewalt, wenn die Bänder im Gerät verbleiben. Mehrere Versionsstände nur umständlich verwaltbar.

Absicherung gegen höhere Gewalt:

Ist eingeschränkt möglich durch Wahl eines guten Standortes - z.B. reduziert es die Chance eines Wasserschadens, wenn der Server nicht gerade auf dem Boden neben der Waschmaschine steht; z.B. Schutz des Servers durch Überspannungsschutz der Stromzufuhr etc.

Welche Kriterien sollten vor der Wahl der optimalen Maßnahmen geprüft werden?

- Welche Kosten würden sich aus einem Totalverlust aller Daten ergeben?
- Welche Daten müssen gesichert werden, welche sollten gesichert werden, auf welche kann verzichtet werden?
- Wo befinden sich die Daten, auf welchen Computern und Festplatten?
- Welchen Schaden würde ein Verlust der geänderten Daten der letzten 3 Tage, letzten 8 Tage, letzten 14 Tage, letzten 4 Wochen ergeben? Außer der Angabe in € ist auch die Einteilung in folgende Kategorien möglich: (A) existentiell bedrohend, (B) unvereinbar mit einem positiven Betriebsergebnis,(C) Gewinnmindernd, (D) vertretbare Kosten, (E) unerhebliche Kosten;
- Wie wichtig ist es, dass ein unterbrechungsfreier Betrieb gewährleistet ist?
- Sind mehrere Versionsstände wünschenswert?

Mögliches Anforderungen an die Datensicherung:

Ein mögliches Anforderungsprofil könnte sein:

- Bei Totalverlust entstehen Kosten die mit einem einem positiven Betriebsergebnis unvereinbar sind, evtl. sogar ein existentiell bedrohliches Ausmaß annehmen.
- Zu sichernde Daten:

MUSS Inhouse-Server „A“ /home und mysql-Datenbanken

MUSS Windows-Arbeitsplatz „A“ Benutzerverzeichnis

MUSS Internet-Server „B“ /var/www /etc und mysql-Datenbanken

SOLL Windows-Arbeitsplatz „C“ komplett,

SOLL Internet-Server „B“ Komplettsystem

NICHT: Windows-Arbeitsplatz „C“

NICHT: Inhouse-Server „A“ /tmp /var/log .Papierkorb /unsicher

NICHT: Internet-Server „B“ /tmp /var/log

- Dimension der Schäden:
  - Verlust der letzen 3 Tage: „unerhebliche Kosten“
  - Verlust der letzen 8 Tage: „vertretbare Kosten“
  - Verlust der letzen 14 Tage: „vertretbare Kosten“
  - Verlust der letzen 4 Wochen: „Gewinnmindernd“

- Auf einen unterbrechungsfreien Betrieb im Störungsfall kann verzichtet werden.
- Mehrere Versionsstände sind gewünscht.
- Möglichst vollautomatische Sicherung ohne Benutzerinteraktion
- Einfache Verifizierung der Sicherung
- Einfache komplette oder partielle Wiederherstellung
- Nur bedingte Sabotagesicherheit
- Totalverlust der letzten 14 Tage stellt „nur“ eine nichtexistenzielle Bedrohung dar
- kompletter Totalverlust ist existentiell bedrohend
- Mehrere ältere Versionen sollen wiederherstellbar sein.

Falls ein **unterbrechungsfreier Weiterbetrieb** des Systems auch nach einem Hardware-Totalausfall nötig ist, MÜSSEN ein gespiegelter Server betrieben und Vorkehrungen zum automatisierten Umschalten getroffen werden. Hiermit beschäftige ich mich im folgenden NICHT!

Sollen **versehentlich durch Benutzer gelöschte Daten vom Benutzer selbst wiederhergestellt werden**, so empfiehlt sich einen „Papierkorb“ ähnlich wie unter Windows einzurichten. Dies lässt sich mit Samba leicht bewerkstelligen. (Stichwort vfs object = recycle, siehe Samba-Manpage; automatisiertes leeren der alten Dateien im Papierkorb mittels Cronjob **find /home -name .Papierkorb -ctime +14 -exec rm {} \;**)

Geht es darum **im Schadensfall das System neu einrichten zu können und die möglichst aktuellen Daten wieder einzuspielen**, bietet **rsnapshot** eine gute Möglichkeit.

Das Konzept von rsnapshot

rsnapshot ist eine Sammlung von perl-Skripten, welche Unix-Standardmechanismen (wie z.B. rsync, hardlinks, ssh) nutzen um eine frei konfigurierbare automatisierte Sicherung von Daten vorzunehmen.

- 1:1 Kopie aller Dateien genau einmal
- Automatisches setzen von hardlinks
- keine Unterscheidung zwischen inkrementellen und Full-Backup sondern transparente Darstellung des kompletten Backups jedes gesicherten Zeitpunkts (hardlinks).

Fix Me!

Die Installation

Versionen: Debian 1.2.9 - 1.3 von den Quellen installieren. Fix Me!

Die Konfiguration

Fix Me!

Ist ein Backup der Datensicherung sinnvoll

Durch eine Sicherung der Daten mittels rsnapshot auf eine weitere Festplatte (nicht auf die gleiche Platte sichern, wie die zu sichernden Daten!) im Server können die meisten Anforderungen gut erfüllt werden.

- Nicht erfüllt sind die Anforderungen die sich stellen, wenn ein Schutz vor höherer Gewalt erfolgen soll.
- Nur bedingt erfüllbar sind die Anforderungen vor einem Schutz vor Sabotage und Softwarefehlern. (Zugriffsrechte auf Sicherung minimieren!)

Um diese beiden Kriterien besser zu erfüllen wird eine zweite Form der Datensicherung eingeführt, das Backup des Backups auf eine externe Festplatte, welche im Normalbetrieb NICHT am System angeschlossen ist. Das Backup auf die externe Festplatte soll so konfiguriert sein, dass das System zu definierten Zeit (z.B. nachts) per cron überprüft ob eine externe Backup-Festplatte angeschlossen ist, und wenn ja dorthin sichert. Als Erfolgsmeldung soll eine E-Mail versendet werden.

Fix Me!

Ausblicke

- Datensicherung von Windows-Arbeitsplätzen mit rsnapshot Fix Me!
- Datensicherung von Apple OS X-Arbeitsplätzen mit rsnapshot Fix Me!

weiterführende Links

- www.rsnapshot.org
- Artikel in Linux-User 2006-08
- Gezielter Fernschuss: Rsnapshot, von Charly Kühnast
- Jürgen Schmidt: „Beruhigungsmittel - Backups für kleine Linux-Server“ in: c't. 7/06, S. 212
- Mike Rubel über automatisierte Backups mit rsync, Hinweise zur Sicherheit des Backups(eng., 2004)