

Backup über das Netzwerk – neue Systeme

Dr. Jens Döbler | jd@cms.hu-berlin.de

Im CMS wird ein zentraler Backupservice betrieben, durch den Daten von Servern gesichert werden. Die Datensicherung hat auch heute noch eine hohe Bedeutung, obwohl vielfach redundante und zuverlässige Systeme verwendet werden. Datensicherung bleibt trotzdem unverzichtbar, einerseits da kein System eine Zuverlässigkeit von 100% erreichen kann. Andererseits ist eine der häufigsten Ursachen für Datenverlust ein Nutzerfehler – z. B. durch versehentliches Löschen oder Überschreiben von Dateien – vor dem keine Systemzuverlässigkeit schützen kann.

Zur Zeit werden ca. 350 Serversysteme gesichert. Das Gesamtvolumen der gesicherten Daten beträgt 120 Terabyte in 300 Millionen Dateien, bei einem täglichen Datenvolumen von durchschnittlich 3,5 Terabyte. Um eine effiziente Verarbeitung dieser Datenmenge sicherzustellen, wurde die Infrastruktur des Backupservice seit Mitte 2008 schrittweise erneuert.

Speichersysteme

Für die Speicherung der Backupdaten wurden zwei Anlagen vom Typ Quantum Scalar i2000 angeschafft, die im Erwin Schrödinger-Zentrum und im Jacob-und-Wilhelm-Grimm-Zentrum aufgestellt wurden. Die Systeme sind automatisierte Bandspeichersysteme, in denen sich Magnetbänder und Bandlaufwerke befinden. Die Bandkassetten werden durch einen Greifarm innerhalb der Anlage bewegt.

Als Speichertechnik kommt LTO4 (Linear Tape Option – 4. Generation) zum Einsatz. Hierbei handelt es sich um eine herstellerübergreifende Magnetband-Technologie, die äußerst zuverlässig ist. Auf einer Bandkassette lassen sich 800 GB Daten speichern. Die Daten werden hierbei in mehreren parallelen Spuren auf das Magnetband geschrieben. Bei der Herstellung der Magnetbänder werden

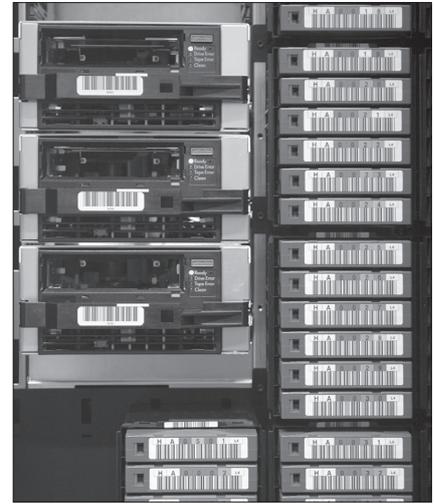


Abb.1 : Innenansicht einer Scalar i2000

sogenannte Servospuren aufgebracht, die von den Laufwerken für die Positionierung der Magnetköpfe verwendet werden. Hierdurch ist stets die korrekte Spurlage sichergestellt und Spurfehler durch falsche Positionierung der Köpfe sind ausgeschlossen. Die LTO-Technologie verwendet Fehlerkorrekturalgorithmen, ähnlich wie bei CD oder DVD, so dass die Daten auch bei einer begrenzten Anzahl an Bitfehlern lesbar sind. Zusätzlich werden geschriebene Daten immer von einem zweiten Kopf zurückgelesen. Wenn die Daten bei dieser Überprüfung nach der Fehlerkorrektur nicht identisch sind, wird der betroffene Block automatisch erneut geschrieben. Auf diese Weise ist immer sichergestellt, dass die geschriebenen Daten lesbar sind. Das Auftreten dieser Art von Fehlern („soft error“) wird in einer Statistik festgehalten, so dass Laufwerksprobleme durch eine erhöhte Anzahl von soft errors frühzeitig erkennbar sind.

Die Speichersysteme sind mit 9 Laufwerken und 1300 Bändern (Adlershof) bzw. 12 Laufwerken und 1600 Bändern (Mitte) ausgestattet. Sie sind modular aufgebaut, so dass sich bei Bedarf sowohl weitere Laufwerke als auch Stellplätze für Bandkassetten nachrüsten lassen.

Der CMS betreibt einen Backupservice auf Basis des Tivoli Storage Managers. Die Infrastruktur wurde im Jahr 2009 erneuert und wird im Folgenden vorgestellt.

Backupsoftware

Bei einer klassischen Backupstrategie wird zwischen einem vollständigen Backup und inkrementellen Backups unterschieden. Bei letzteren werden ausgehend von einem vorherigen Backup nur geänderte Daten gesichert, um das Backupvolumen zu reduzieren. Da die inkrementellen Backups aufeinander aufbauen, nimmt der Aufwand einer Rücksicherung stetig zu, da nacheinander zunächst das vollständige und anschließend die inkrementellen Backups eingespielt werden müssen. Daher werden normalerweise regelmäßige (z. B. wöchentlich) vollständige Sicherungen durchgeführt.

Ein solches Konzept ist bei dem vorliegenden Umfang von ca. 120 TB nicht praktikabel. Eine Alternative stellt der Tivoli Storage Manager (TSM) des Herstellers IBM dar. TSM verwendet ein vollinkrementelles Verfahren, das mit einer SQL Datenbank realisiert ist. Die Datenbank speichert die Metadaten der gesicherten Dateien (Dateiname, Zugriffsrechte, Erstellungsdatum, Dateigröße, usw.), der eigentliche Dateiinhalt wird auf einem Speichermedium abgelegt und mit einem Index verknüpft. Dieser Index dient als Verweis auf den Dateiinhalt und ist gemeinsam mit den Metadaten in der Datenbank gespeichert. Bei einer Erstsicherung erfolgt zunächst eine Sicherung aller Dateien. Bei darauf folgenden Sicherungen werden dann die Metadaten des bisherigen Datenbestandes aus der Datenbank an den Klienten übertragen. Dieser prüft nun anhand der Informationen, ob sich Dateien geändert haben, überträgt diese sowie neue Dateien zum Server und meldet gelöschte Dateien an den Server. Bei geänderten oder gelöschten Dateien wird dann auf dem Server der Datenbankeintrag der Datei als inaktiv markiert. Hierdurch ist zu jeder Zeit eine Rücksicherung des aktuellen Datenbestandes in einem Schritt möglich, indem aus der Datenbank die aktiven Dateien abgefragt werden. Ältere Dateiversionen bleiben ebenfalls eine Zeitlang gespeichert; in der jetzigen Konfiguration sind maximal 4 Versionen einer Datei vorhanden (zusätzlich zur aktuellen Version verbleiben

maximal drei vorherige Versionen für 30 Tage). Gelöschte Dateien sind 60 Tage verfügbar. Durch dieses Sicherungsschema beträgt das tägliche Datenvolumen nur etwa 3,5 TB.

Die Datenbank ist für die Funktion des Servers erforderlich und die Leistungsfähigkeit der Datenbank ist von entscheidender Bedeutung für den ordnungsgemäßen Betrieb des Servers. Dies stellte für die bis Anfang 2009 verfügbaren TSM-Versionen ein Problem dar, da die integrierte Datenbank bei mehr als 150 Mio. gesicherten Dateien Performanceschwächen zeigte. Dieses Problem wird durch die im März 2009 veröffentlichte Version TSM 6.1 behoben, in der die Datenbankfunktionalität nicht mehr im TSM integriert ist, sondern auf eine externe DB2 Datenbank umgestellt wurde. Die Version 6.1 ist im CMS seit Oktober 2009 im Einsatz und hat sich als zuverlässig und performant herausgestellt.

Backupkonzept

Seit Oktober 2009 werden insgesamt vier TSM-Server betrieben. Es befinden sich jeweils zwei Server im Erwin Schrödinger-Zentrum und im Jacob-und-Wilhelm-Grimm-Zentrum. An beiden Standorten wird ein Server mit der Version 5.5 und der zweite mit Version 6.1 betrieben. Es handelt sich um Server vom Typ IBM p520, die mit 10 Gigabit-Ethernet an das HU-Netz angebunden sind. Die Backups erfolgen normalerweise zwischen 18:00 und 4:00 Uhr. Die zu sichernden Systeme übertragen die Daten per Netzwerk an den jeweiligen Server, dort werden die Backups zunächst auf Festplattensysteme zwischengespeichert. Die Daten werden hierbei auf zwei getrennte Festplattensysteme gespiegelt, um die Sicherheit der Daten auch bei Ausfall eines Plattensystems garantieren zu können. Ab 5:00 Uhr werden die Daten dann an beiden Standorten auf Band kopiert. Von allen Daten sind immer zwei Kopien vorhanden, um sicherzustellen, dass die Daten verfügbar sind, auch wenn eine Kopie der Daten nicht lesbar oder zerstört ist. Das zweistufige Konzept der Datenspeicherung (erst auf Disk, dann auf Band) erlaubt es, dass die Bandlauf-

werke kontinuierlich Daten schreiben. Die verwendete LTO4-Technik schreibt mit maximal 120 MB/s auf Band und verringert die Geschwindigkeit dynamisch, um einen Start-Stopp-Betrieb zu verhindern. Hierzu muss jedoch eine Datenrate von 30 MB/s erreicht werden. Diese Schwelle wird bei der Übertragung der Daten über das Netzwerk oftmals nicht erreicht; die Zwischenspeicherung auf Disk erlaubt es, die Daten effizienter auf die Bandmedien zu transferieren.

Da TSM alle Metadaten in der Datenbank speichert und diese auch die Zuordnung von Dateien zu Backupmedien enthält, ist ohne sie eine Wiederherstellung von gesicherten Daten unmöglich. Daher wird auch die Datenbank täglich – an beiden Standorten – auf Band gesichert. Neben dieser Sicherung werden im TSM alle Änderungen nach dem letzten Datenbankbackup in einem sogenannten Log gespeichert. Bei einer Beschädigung der Datenbank ist es normalerweise möglich, den aktuellen Stand aus einem Datenbankbackup und den Logs wiederherzustellen. Im Notfall ist auch eine Wiederherstellung ohne Logs möglich, in diesem Fall gehen jedoch die Änderungen in der Datenbank seit dem Datenbankbackup verloren. Da die Datenbankbackups nach dem Sicherungsfenster, in dem die Klienten sichern, angefertigt werden, ist der Datenverlust auch in diesem ungünstigsten Fall minimal. Durch die Verteilung der Daten und der Server auf die zwei Standorte ist eine weitreichende Datensicherheit gewährleistet, selbst wenn die Daten an einem Standort (z. B. durch Feuer) vernichtet werden.

Die Details für eine Teilnahme am Backupservice sind unter http://www.cms.hu-berlin.de/dl/systemservice/fileservice/tsm_t_t.html aufgeführt. Das TSM-Team ist unter der E-Mail tsm@cms.hu-berlin.de zu erreichen.