Datensicherung und -wiederherstellung

Jens Piegsa, Martin Bialke, Stefan Ostrzinski, Christoph Havemann

04.08.2016

Motivation

Backups können helfen, allerdings nur wenn man sie hat

- häufig macht erst ein schmerzlicher Datenverlust klug

Manuelle und spontane Sicherungen sind oftmals nicht effektiv genug:

- sie erzeugen große Mengen „Datenmüll“
- der "wichtige" Ordner oder die "wichtige" letzte Änderung war noch nicht in der Sicherung enthalten
- das Sicherungsformat ist im Schadensfall nicht lesbar; die Wiederherstellbarkeit wurde nie getestet
- der Verlust wird zu spät erkannt und komplexe Rekonstruktionsszenarien entstehen

Datensicherungs- und wiederherstellungskonzept

Der notwendigen Schutz kann durch ein strukturiertes Vorgehen und ein effektives Datensicherungs- und Wiederherstellungskonzept erzielt werden.

Dies beinhaltet:

- Die Bestimmung von Anforderungen und akzeptiertem Restrisiko
- dokumentierte Entscheidungen über einzusetzende Sicherungs-maßnahmen
- getestete Wiederherstellungsprozeduren
- klar definierte Verantwortlichkeiten

Welche Gefahren sind realistisch?

Kategorie	Beispiel
höhere Gewalt	Feuer, Wasser, Blitz, ...
organisatorische Mängel	unzureichende Regelungen
menschliche Fehlhandlungen	versehentliches Löschen oder Überschreiben, Laptopverlust
technisches Versagen	Festplattenausfall/ -fehler, Ausfall anderer Hardwarekomponenten, Stromausfall, Softwarefehler
vorsätzliche Handlungen	Verfälschung, Löschung, Diebstahl, Softwareviren

Welche Gefahren birgt die Praxis?

- Sicherung wird zu spät eingerichtet
- Schadensfall wird zu spät erkannt und Sicherungen reichen nicht lange genug zurück
- Zugriffsrechte gehen über Betriebssystemgrenzen hinweg verloren
- Sicherungen sind unzureichend vor unberechtigtem Zugriff geschützt

Welche Arten der Datensicherung existieren?

Art	Funktionsweise	Eigenschaft
unstrukturiert	alle oder die wichtigsten Daten werden manuell auf spontan verfügbares Medium gesichert	Oftmals schlecht dokumentiert, zu unregelmäßig
vollständig	Daten zu einem Zeitpunkt unabhängig vom Änderungszeitpunkt	Speicherplatzbedarf hoch, Wiederherstellung einfach und schnell
inkrementell	Zyklen aus einer vollständigen und mehreren inkrementellen Sicherungen, wobei letztere nur die Änderungen seit der letzten Sicherung beinhaltet	spart Speicherplatz und Zeit bei der Sicherung, erfordert jedoch höherer Aufwand bei der Wiederherstellung
differenziell	alle Daten, die seit der letzten *vollständigen* Sicherung hinzugefügt oder geändert wurden	Kompromiss aus *inkrementell* und *vollständig*
invers-inkrementell	inkrementell, jedoch so, dass der aktuellste Stand immer vollständig erscheint	Kompromiss aus *Spiegelung* und *inkrementeller* Sicherung

Was ist per Definition KEINE Datensicherung?

- Datei-Synchronisation
- Datenbanksynchronisation
- Messaging-Konzepte auf Anwendungsebene
- Dokumentenmanagementsysteme
- Versionsmanagementsysteme
- Langzeitarchivierung

Was sind wichtige Einflussfaktoren?

Einflussfaktor	Beschreibung
Spezifikation der Daten	Betriebssystem, Anwendungssoftware, Konfiguration, Anwendungsdaten, Protokolldaten
Verfügbarkeitsanforderungen	maximal tolerierbare Ausfallzeit (ggf. unter zeitweiliger Weiterarbeit auf Papier)
Wiederherstellungsaufwand	Betriebssystem, Anwendungssoftware, Daten
Rekonstruktionsaufwand für zum Ausfallzeitpunkt ungesicherte Daten	können und müssen diese wiederhergestellt werden und wenn ja: wie hoch ist der Aufwand dafür?
Datenvolumen, Änderungsvolumen und -zeitpunkt	Sicherung abends / nachts und unmittelbar vor / nach wesentlichen Änderungen
Aufbewahrungs- und Löschfristen	Aufbewahrungsfristen (Nachvollziehbarkeit von Studiendaten), Löschfristen (Datenschutz)
Vertraulichkeitsbedarf des Sicherungsmediums	steigt bei Kombination verschiedener Daten
Integritätsbedarf	bei Rekonstruktion aus verschiedenen Quellen

Was sind individuelle Parameter?

Parameter	Argument	Beispiel
Sicherungs-intervall P	$A_P \leq P \leq B_P$ A_P: *Wie häufig werden Daten voraussichtlich verändert und wie detailliert sollen Zwischenstände rekonstruierbar sein?* B_P: *Welcher Rückfallverlust bezüglich neuer Daten wird toleriert?*	$P = 1\ Tag$
Aufbewahrungs-dauer D	$A_D \leq D \leq B_D$ A_D: *Wie lange in der Zeit zurück sollen Veränderungen erkenn- und rückgängig machbar sein?* B_D: Richtet sich nach dem Sicherungsmodus und danach wie viele Sicherungsstände der verfügbare Backup-Speicher voraussichtlich fassen kann: - fortlaufend: alle Sicherungsstände werden aufbewahrt oder - umlaufend: ältere Stände werden überschrieben oder gelöscht	$D = 6\ Wochen$; umlaufend
Sicherungs-modus M	$M = F$ *oder* $I(n)$ F: Es werden immer vollständige Sicherungsstände erstellt; empfohlen bei geringer Datenmenge bzw. größeren Datenänderungen $I(n)$: eine vollständige Sicherung erfolgt nur alle n-mal, sonst inkrementell (Änderungen gegenüber der letzten vollständige Sicherung); empfohlen bei großen Datenmengen mit wenigen Änderungen; mit höherem Wiederherstellungsaufwand verbunden	$M = I(7)$: einmal pro Woche vollständig, sonst täglich inkrementell

Was bleibt zu klären?

- Detaillierte Vorgehensweise und eingesetzte Softwarewerkzeuge
- Art und Menge der Speichermedien
- Verantwortlichkeiten
- Aufbewahrungsort
- Transportmodalitäten

Fazit

In Forschungsprojekten fallen oftmals vielseitig, umfänglich und verteilt Dokumente und Daten an, die für einen spezifischen Zeitraum effektiv vor Verlust und ungewollter Änderung geschützt werden müssen.

- Durch Datenredundanz werden Verlustrisiken eingeschränkt.
- Die redundante Datenhaltung verursacht Kosten, die gegenüber Restrisiken abzuwägen sind.
- Die Einrichtung von Sicherungsautomatismen und Wiederherstellungsmaßnahmen muss rechtzeitig geplant und umgesetzt werden.
- Sprechen Sie die Datensicherung mit den zuständigen Administratoren ab. Nutzen Sie die **Datensicherungs- und wiederherstellungsvorlage** zur Vorbereitung.

Weiterführende Literatur

- Bundesamt für Sicherheit in der Informationstechnik:
  www.bsi-fuer-buerger.de/BSIFB/DE/MeinPC/Datensicherung/datensicherung_node.html
- Heise-Artikel „Besser sichern! Backup-Blues und -Strategien“:
  www.heise.de/ct/artikel/Besser-sichern-290430.html
- Galileo openbook Shell-Programmierung. Backup-Strategien (Linux):
  openbook.galileo-press.de/shell_programmierung/shell_017_003.html
- BMI Österreich: IT-Sicherheitshandbuch. Datensicherung und Notfallvorsorge:
  www.bmi.gv.at/cms/BK/praevention_neu/info_material/files/IT_Sicherheitshandbuch.pdf
- Rat für Sozial- und WirtschaftsDaten: Langzeitarchivierung von Forschungsdaten:
  ratswd.de/dl/downloads/langzeitarchivierung_von_forschungsdaten.pdf
- TMF Musterdokumente für ein IT-Sicherheitskonzept
  https://www.tmf-ev.de/Produkte/Mustertexte_DSKonzepte2.aspx