

Die 10 häufigsten Probleme bei Datenmigrationen

Von Ralf Draeger, Technischer Leiter, dynaMigs

Datenmigrationen sind komplex und erfordern ein hohes Maß an Expertenwissen, damit sie ohne Probleme über die Bühne gehen. Organisationen, die potenzielle Problemfelder schon vor der Migration erkennen und entsprechend vorausplanen, sparen bei ihren Migrationsprojekten Zeit und Geld. Hier eine Liste aus der Praxis mit den zehn häufigsten Problemen bei Datenmigrationen – und wie man sie lösen kann.

1. Mangelhaftes Datenmanagement

Eines vorweggenommen: Datenmanagement ist keine Voraussetzung für die Portierung von Daten. Wer aber keinen Überblick über seine Daten hat, hat es schwer, diese planvoll, korrekt und schnell zu migrieren. Betreiben Unternehmen kein konsequentes Datenmanagement, sind die Bestände schnell unstrukturiert auf verschiedenen Systemen verteilt – und damit an sich Grund für eine Datenmigration. Denn oft liegen dann kalte Daten auf teuren Medien, unternehmenskritische Informationen dagegen auf nicht ausreichend gesichertem Speicher. Das kann im schlimmsten Fall zum Produktionsausfall führen, weil wichtige Informationen nicht mehr verfügbar sind. Mangelhaft verwaltete Daten sind zudem teuer, weil die Verantwortlichen laufend unkoordiniert Speicher erweitern und teure Medien auf Vorrat kaufen, die nicht unbedingt nötig sind. Außerdem kann eine Migration gewachsener Datenbestände viel Zeit und Geld verbrauchen. Ein Aufräumen der Datenstruktur im Vorfeld empfiehlt sich – nicht zuletzt, um sich vielleicht von Altlasten zu befreien und Speicherkosten zu optimieren.

Lösung: Eine Migration ohne professionelles Datenmanagement geht natürlich auch – aber mit ist es oft schneller, sicherer und besser.

2. Tasks manuell abarbeiten

Die Datenberge einer Organisation von Hand zu migrieren, kostet viel Zeit und Ressourcen. Insbesondere wenn im Lauf der Zeit innerhalb von Jahrzehnten angehäufte Daten zu verarbeiten sind, wird die manuelle Migration im Prinzip bereits unmöglich. Zu viele Mitarbeiter wären dafür nötig und zeitaufwändig einzuarbeiten. Ganz entscheidend: Die manuelle Datenmigration ist bei solchen Mengen sehr fehleranfällig.

Lösung: Sich wiederholende Teile der Migration mit fortschrittlichen Migrationstools automatisieren. Prozessautomatisierung spart einerseits Zeit und Geld. Andererseits erhöht es die Qualität und damit die Datensicherheit.

3. Umlautproblematik

Im Mainframe war die Menge an Zeichen mit ASCII noch auf 127 begrenzt und kannte keine Umlaute. UTF-8 / Unicode (Windows NT 4.0 und höher/Unix/NAS) stellt Zeichen mit einem, zwei oder drei Bytes dar, was mehr als eine Million Zeichen ermöglicht. Bei unterschiedlichen Netzwerkprotokollen, insbesondere beim Wechsel auf NFSv4, können Sonderzeichen oder Umlaute in Dateinamen bei Datenmigrationen Probleme bereiten. Denn jedes dieser Zeichen kann in unterschiedlichen Zeichenfolgen auf dem NAS abgelegt sein, je nach Konfiguration des Clients. Ein „ä“ findet sich unter Umständen im Filesystem der NAS als „e4“, „c3 a4“ oder „c3 83 c2 a4“, wird aber je nach Konfiguration des Clients womöglich auch als Element aus der Liste „δφηΨ⋯“ interpretiert. Dieses Problem betrifft rund 400 Sonderzeichen aus unterschiedlichen Sprachen und Sprachvarianten. Mit einem neuen falschen Namen lassen sich die Dateien nach der Migration nicht mehr ohne Weiteres finden.

Lösung: Das Wissen, dass Umlaute oder andere Zeichen Probleme bereiten können, ist der erste Schritt. Nur jahrelang erfahrene Experten können solche Probleme und deren Effekte antizipieren und schon vor der Migration entschärfen.

4. Kein einheitliches Usermapping bei Multiprotokoll-Umgebungen

Unternehmen, die aus historischen Gründen unterschiedliche Protokolle auf ein und demselben Speicher verwenden, etwa NFS unter Linux oder CIFS unter Windows, haben oft kein einheitliches Usermapping. Das kann bei der Migration Probleme bereiten, wenn beispielsweise Windows-Nutzer im Active Directory stehen, die Linux-Anwender jedoch nur lokal angelegt sind. So weiß das Storage-System nach einer Datenmigration nicht mehr, welchem Nutzer es welche Zugriffsrechte geben soll.

Lösung: Vor der Migration tragen die Verantwortlichen alle User z.B. via LDAP unternehmensweit mappen. Dies sorgt dafür, dass die Storagesysteme die User in beiden Welten korrekt erkennen und die ansprechenden Rechte ziehen.

5. Kopierfehler durch Verifikation erkennen und vermeiden

Wer Daten migriert, will die Sicherheit haben, dass diese korrekt am Ziel ankommen. Das Caching-Verhalten und Fehler im IP-Stack des Copyhost-Betriebssystems, falsches Codieren von Umlauten oder nicht korrekt dargestellte Verzeichnisse können ein falsches Überspielen der Daten in den beteiligten Speichersysteme verursachen. Die Logfiles der Migration, auf die viele Verantwortliche für die Verifikation zurückgreifen, können die korrekte Migration nicht hinreichend bestätigen, da sie etwaige Schwächen von Kopiertools außer Acht lassen.

Lösung: Moderne Verifikations-Tools, wie etwa das externe Prüfsystem von dynaMigs, scannen Daten, Dateien und Ordner samt Metadaten an Quelle und Ziel und gleichen diese ohne weitere Downtime ab. So können sie Fehler finden und zeigen auf, im welchen Teilschritt der Migrationskette sie passierten.

6. Distributed File Systems (DFS) und nicht erkannter Reparse-Point

Speicherarrays unterschiedlicher Anbieter legen DFS-Links oder Widelinks unterschiedlich an. Dies kann dazu führen, dass das Kopiertool den DFS-Link nicht mehr als Reparse-Point erkennt. Mit unter Umständen überraschenden Folgen: Denn das Kopiertool merkt dann nicht, dass der Link den zu kopierenden Speicherbereich verlässt. Die Folge kann sein, dass ungewollt Daten kopiert oder sogar irrtümlich gelöscht werden.

Lösung: Um dieses Problem zu umschiffen, muss man wissen, dass es das Problem potenziell gibt – und man benötigt die nötige Erfahrung des Verhaltens verschiedener Speicherarrays und Kopiertools.

7. Wechsel des Speichersystems

Die Probleme bei Storage-Migrationen wachsen, wenn eine IT-Administration nicht nur auf ein neues Array, sondern auf ein komplett neues System migrieren will oder muss - beispielsweise anlässlich eines Anbieterwechsels oder bei Vertragsende. Das neue System ist dann so einzustellen, dass es sich genauso verhält, wie das alte. Ansonsten bindet es sich nicht in die IT-Gesamtinfrastruktur ein und bleibt ein Fremdkörper. Abläufe zum Verarbeiten der Daten – und damit Geschäftsabläufe – werden folglich unterbrochen.

Lösung: Die Migrationsverantwortlichen konfigurieren die Zielsysteme vor der Migration grundlegend und sorgfältig. Sie berücksichtigen dabei vor allem das Zusammenspiel mit der gegebenen IT.

8. Kein Blick für Bandbreite

Die immer größeren Datenmengen bringen Netzwerke unter Umständen an ihre Leistungsgrenze. Ein Datentransport über WAN kann lange dauern und zu Abbrüchen der Migration und Datenverlust führen. So manche Speichersysteme sind bereits heute zu groß geworden, um sie in einem sinnvollen Zeitrahmen über weite Strecken zu transportieren. Stichwort: Data-Gravity.

Lösung: Das lokale Kopieren der Dateien vermeidet diese Probleme. Es macht vieles einfacher und schneller.

9. Offline-Cache

Offline-Caches sind unter CIFS ein Problem. In diesen Zwischenspeichern halten Clients Dateien vor, die offline geändert werden können und daher später mit dem Server zu synchronisieren sind. Nach der zwischenzeitlichen Migration kann der Client diese nicht mehr synchronisieren, falls das Ziel einen neuen Namen trägt. Dann verwirft der Client diese nach einiger Zeit gänzlich. Die Daten gehen also verloren.

Lösung: Hier hilft nur die Analyse, ob ein Client den Offline-Cache serverseitig unterstützt sowie ein entsprechender Recall vor der Migration.

10. Übernahme von IP-Adresse und Servernamen

Will man Servernamen und IP-Adresse bei einer Datenmigration übernehmen, kann das die Address-Resolution-Protocol (ARP) - Caches der Clients verwirren. Dann sind falsche Ziel- oder Ursprungsadressen hinterlegt. Im schlimmsten Fall ist ein Reboot des Clients nötig: Bei Windows wegen korumpierter ARP-Zwischenspeicher, bei Unix etwa wegen eines veralteten Mounting-Punkts (Stale Mount).

Lösung: Stellt man auf neue Namen / IPs um, dann ergibt sich das Problem nicht, allerdings müssen dann auf allen Clients Änderungen vorgenommen werden.

Bei Übernahme sollte man alle beteiligten Windows Client Versionen im Vorfeld in einer Testmigration auf ihr Verhalten prüfen. Anfällige Versionen werden dann am besten während der Migration heruntergefahren. Bei Unixderivaten ist ein umount der betroffenen Exports zwingend, hierfür sollte man neben den fstabs ggf auch die Automounterkonfigurationen prüfen.

Migrationsprobleme idealerweise lösen, bevor sie aufkommen

Die Liste möglicher Probleme bei Migrationen ist lang. Zahlreiche Fallstricke sorgen dafür, dass die meisten Migrationsprojekte früher oder später ins Stocken geraten. Dabei sind die Unternehmen im Vorteil, die bereits ein professionelles Datenmanagement aufgebaut haben. Wer noch kein Datenmanagement umgesetzt hat, oder wem Manpower, Tools, Knowhow und Expertise fehlen, holt sich besser frühzeitig Rat bei externen Spezialisten für Datenmanagement und Datenmigrationen. Sowohl bei der Planung als auch bei der Umsetzung. Insbesondere durch die Prozessautomatisierung spart dies am Ende Nerven, Zeit und Geld.