

DEDUPLIZIERUNG IST PHYSISCH - STORAGE MANAGEMENT IST LOGISCH

WHITE PAPER

NORTHERN STORAGE SUITE
VERSION: V8
DATUM: 2009-01-05



SPEICHERRESSOURCENMANAGEMENT UND DEDUPLIZIERUNG

Selbst robustesten Speicherinfrastrukturen fällt es schwer, das starke Wachstum der täglich auf sie zuströmenden Datenmengen zu bewältigen. Die Frage, wie effizient ein Unternehmen seine Daten verwaltet, hat entscheidenden Einfluss auf Kosten und Produktivität.

Die wesentliche Herausforderung der IT-Branche ist es, zwei scheinbar unvereinbare Ziele in Einklang zu bringen: kritische Daten müssen jederzeit abrufbereit sein, während Speichergeräte, darunter Fileserver und NAS-Geräte, für ausreichend Raum sorgen. Oft ist das Hauptanliegen von Unternehmen der Einsatz einer Backup-Methode, die auf einer regelmäßigen Datensicherung beruht. Dabei wird jedoch in aller Regel ein Backup aller Daten erstellt, was zu Redundanz, schleppenden Netzwerkleistungen und vorschnell ausgelasteten Speicherkapazitäten führt.

Die Deduplizierung von Daten wird zunehmend zu einem entscheidenden Faktor im Backup-Prozess und bei der Speicherung der Daten auf Festplatten. Die Vorteile der Datendeduplizierung liegen auf der Hand: mehrfach vorhandene Datenobjekte werden unterdrückt, was den Backup-Prozess beschleunigt, vereinfacht und die Datenverfügbarkeit erhöht. Das Verfahren der Datendeduplizierung an sich ist jedoch häufig langwierig und erfordert umfassende Netzwerkressourcen und Rechenperformance, insbesondere weil die Technologie nicht die nötige „Intelligenz“ aufweist, um zwischen wesentlichen Geschäftsdaten und belanglosen bzw. unerwünschten Daten zu unterscheiden; dies führt zur Deduplizierung von Daten, die hier keine Daseinsberechtigung haben. Aus diesem Grunde erwägen zahlreiche Unternehmen den Einsatz von SRM-Tools in Kombination mit Dedup-Tools, um die Daten im Vorfeld der eigentlichen Deduplizierung zu strukturieren und zu rationalisieren. Dies führt zu einem bedeutenden Produktivitätsgewinn: die Arbeitslast der IT-Administratoren wird verringert, die Lebensdauer von Speichersystemen erhöht und die Verfügbarkeit von Daten optimiert, zudem haben Anwender sofortigen Zugriff auf die verlangten Daten.

Dem folgenden Text ist eine Aufstellung der Vorteile des Speicherressourcenmanagements im Vorfeld der Datendeduplizierung zu entnehmen. Wir wollen zeigen, wie ein SRM-Tool wie Northern Storage Suite in Kombination mit Datendeduplizierungs-Technologien eingesetzt werden kann, um Daten effizient zu strukturieren und eine optimale Leistung der unternehmenseigenen Storage-Infrastruktur zu gewährleisten.

WAS IST SRM?

Durchschnittlich etwa 70% der auf Speichersystemen gespeicherten Daten sind geschäftsirrelevant, überholt oder redundant. Datenmanagement mittels einer SRM-Lösung ermöglicht das Eliminieren irrelevanter oder veralteter Daten, die Freigabe verschwendeter Kapazitäten, die Steigerung der Geschwindigkeit und Effizienz des Dedup-Prozesses, was Systemressourcen schont.

IDENTIFIZIERUNG UNERWÜNSCHTER DATEN UND FREISETZUNG VON SPEICHERPLATZ

Durch die Integration umfassender Reportingtools in die Speicherressourcenmanagement-Lösung ist der Anwender genauestens über die reelle Auslastung seines Speicherplatzes, die auf dem System gespeicherten Datentypen, das Alter dieser Daten sowie über Speicherplatz-Benutzer und den Ort, an dem diese Daten abgelegt sind, informiert. Dies ermöglicht eine rasche Identifizierung unerwünschter Dateien (wie veralteter Dateien, Media-Dateien, Dateien zu längst abgewickelten Geschäftsabläufen usw.) sowie deren gezieltes

Unterdrücken bzw. Archivieren. Diese Prozesse können im Voraus geplant und nach Geschäftsschluss durchgeführt werden, um produktive Abläufe nicht zu beeinträchtigen.

KONTROLLIERTER EINSATZ UND WACHSTUM VON SPEICHERKAPAZITÄTEN

Die von einer Speicherressourcenmanagement-Software gebotene Quotenverwaltungsfunktion ermöglicht es dem Anwender, Einsatz und Wachstum seines Speicherplatzes gezielt zu kontrollieren, indem er akzeptable Grenzwerte für Benutzer bzw. Verzeichnisse festlegt. Durch die Konfiguration granularer Quoten je nach Anforderungsprofil von Unternehmen, Geschäftsbereichen oder Sachverhalten können zum Beispiel bestimmte Benutzer, Gruppen oder Verzeichnisse limitiert werden. Das Sperren von Dateien hält zudem Benutzer davon ab, unerwünschte Dateien bzw. Dateitypen wie zum Beispiel Media-Dateien zu speichern.

EINBEZIEHUNG VON BENUTZERN

Die Involvierung der User ist ein weiterer, nicht zu vernachlässigender Aspekt eines effizienten Storeressourcenmanagements, da nur der User selbst den genauen Wert seiner Dateien beurteilen kann. Der Systemadministrator kann zwar einen Zeitplan für das Löschen bzw. Archivieren von Dateien je nach deren Typ, Alter oder letztem Zugriffsdatum erstellen. Was aber passiert, wenn jemand feststellt, dass eine drei Monate alte Datei, die plötzlich benötigt wird, nicht mehr dort ist, weil das System aufgrund der voreingestellten, automatischen Quotenvorgaben ohne Vorwarnung die Datei gelöscht hat? Nur der Benutzer, der die Daten erstellt hat und dem diese gehören, kennt den realen Wert der Daten. Ein gutes Speichermanagementtool sollte in begrenzter Form Benutzern Handlungsspielraum und Entscheidungsfreiheit über die Daten einräumen, bei denen Bedarf an Datensicherung und ständigem Zugriff besteht. Durch ein solches Werkzeug wird Benutzern die Wichtigkeit der Speicherressourcenverwaltung bewusst, sie werden in die Verwaltung eigener Daten einbezogen, über zu alte, große oder duplizierte Dateien und Media-Dateien informiert und erhalten die Möglichkeit, diese entsprechend zu verwalten. Ausgehend von Einzelberichten und nach eigenem Ermessen können die Benutzer ihre Dateien löschen, verlagern bzw. archivieren.

WAS IST DEDUPLIZIERUNG?

Sinn der Dateneduplizierung ist das Eliminieren bzw. Entfernen von redundanten Dateien, Bytes oder Datenblöcken, um zu gewährleisten, dass nur „einmalige“ Daten auf einer Festplatte gespeichert werden. Während der Deduplizierung wird die Datei zerlegt, jedem Einzelteil wird eine separate Nummer zugeordnet, die wiederum in einem Index gespeichert wird; von jedem Teil wird nur eine Iteration gesichert.

Dateneduplizierung bietet zahlreiche Vorteile:

- Bedeutende Verringerung der erforderlichen Speicherbackupkapazität um durchschnittlich das 10- bis 20-fache.
- Weit größere Netzwerkbandbreite im Backup-Prozess
- Reduzierung von Übertragungszeiten
- Vereinfachte Backupverfahren, geringerer Einsatz von Sekundärspeicher, Tape etc.
- Enormes Einsparpotential bei den Kosten für Festplatten sowie unter anderem für Stromversorgung und Kühlung.

Die Dateneduplizierung erfordert jedoch einen hohen Bedarf an Rechenleistung und Speicherplatz. Die Aufrechterhaltung der Leistung setzt zudem eine Optimierung der

Größe des Indexes voraus; dieser befindet sich in aller Regel im Festplatten-Cache des Deduplizierungsservers, der einen begrenzten Speicherplatz aufweist.

VERSCHIEDENE ARTEN DER DEDUPLIZIERUNG

- **Inline-Deduplizierung:** die Deduplizierung findet Host-seitig in der Backup-Anwendung oder in einer Appliance auf dem Datenpfad statt. Diese Methode ist besonders effizient, wenn eine Verringerung des Speicherplatzes und eine starke Reduzierung der durch das Backup dargestellten Netzwerkbelastung angestrebt werden. An entfernt liegenden Niederlassungen mit schwachen Links und virtualisierten Servern ist die Inline-Deduplizierung besonders zu empfehlen, da sie die Ein- und Ausgänge auf den Netzwerkadaptern reduziert. Abstriche sind bei Performance und Kapazität hinzunehmen. Aus diesem Grunde ist sie besonders für Datenmengen unter 10 TByte geeignet.
- **Offline- bzw. Post-Process-Deduplizierung:** die Deduplizierung wird vom System oder einer Appliance vorgenommen, die sich außerhalb des Datenpfades befindet und tätig wird, wenn das Backup abgeschlossen ist. Dieses Verfahren ist weniger effizient, was die Freigabe von Speicherkapazitäten anbelangt, kann jedoch bei Volumen von mehr als 10 TByte eingesetzt werden, insofern das System für einen Datenempfang im Vorfeld der Deduplizierung ausgelegt ist.

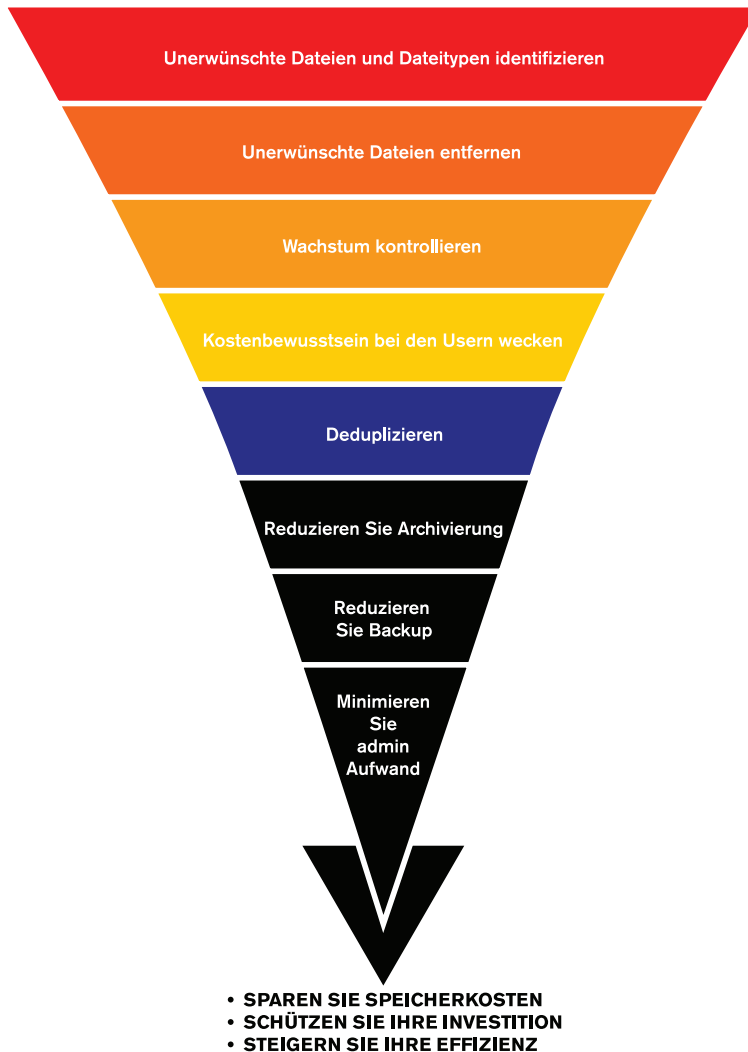
Beide Verfahren bieten Vorteile, je nach Storage-Infrastruktur und Volumen der zu bearbeitenden bzw. zu speichernden Daten stellt sich jedoch eine wesentliche Frage: warum sollten der Deduplizierung alter oder geschäftsirrelevanter Daten ohne Interesse für das Unternehmen wichtige Ressourcen zugeteilt werden?

STORAGEMANAGEMENT UND DEDUPLIZIERUNG: EINE KOMBINATION, DIE SICH LOHNT

Durch die Implementierung von SRM-Tools können in der Regel 30 bis 40% des Speicherplatzes auf dem Server freigesetzt werden, oft sogar mehr. Eine SRM-Lösung, die eine Art User-Selbstverwaltung beinhaltet, kann zum Deduplizierungsprozess beitragen, da die Benutzer ihre Daten nun selbst organisieren und strukturieren können. Hardware-Investitionen werden geschont und weisen dadurch eine längere Lebensdauer auf, der Verwaltungsaufwand, der rund 80% aller mit dem Storage verbundenen Kosten darstellt, wird bedeutend gesenkt, und IT-Administratoren bzw. -Mitarbeiter können sich vermehrt dem Tagesgeschäft widmen.

Die Verbindung von Speichermanagement und Deduplizierung führt zur Optimierung der Bandbreite und Verkürzung von Übertragungszeiten. Die Verarbeitung und Speicherbelegung in Verbindung mit Deduplizierungsprozessen stellen kein Hindernis mehr für Netzwerkleistung und Produktivität dar.

STORAGE-MANAGEMENT UND DEDUPLIKATION ARBEITEN HAND IN HAND



Eine SRM-Lösung, die Quoten- und Reportingfunktionen vereint und die Verrichtung täglicher Speicherarbeiten automatisiert, ist in heutigen High-Tech Unternehmen Gang und Gäbe, insbesondere da die Netzwerke der Unternehmen immer komplexere Speichertechnologien und -geräte aufweisen, von SAN bis NAS. Der Trend zur Einschränkung der freien Verfügung über Speicherplatz erfolgt nur schrittweise, obwohl die oben dargelegte Wertsteigerung für sich spricht. In Verbindung mit Deduplizierungs-Tools, die überflüssige Daten reduzieren und wichtige Benutzerdaten speichern, ist dieser Wertgewinn besonders offenkundig.

Selbstverständlich kann die Deduplizierungstechnologie nicht das von uns angesprochene Hauptproblem lösen: die Einwirkung auf das Benutzerverhalten durch eine Mischung aus Sensibilisierung, Leistungsansporn und Selbstverwaltungstools. Wie erwähnt verhindert die Deduplizierung nicht die Sicherung von Daten, sondern entfernt Dubletten auf physischer Ebene. Es ist auch nicht Aufgabe der Deduplizierung, Benutzer zu einer verantwortungsvolleren Speichernutzung zu veranlassen. Deduplizierung ist eine gute Lösung, deren Leistungsfähigkeit von der Reaktionsschnelligkeit des Prozesses abhängt. Es ist ebenfalls eine verborgene Lösung, denn dem Benutzer ist weder die Speicherpolitik seines Unternehmens noch die Tatsache, dass die Geschäftsleitung die Strukturierung von Daten als eine ihrer Hauptanliegen betrachtet, bewusst.

Aus Sicht der IT-Organisation verlagert ein solches SRM-Tool, das auf Deduplizierung von Daten basiert, einen guten Teil der Verantwortung auf Benutzerebene, was für jeden Vorteile bringt – die IT-Infrastruktur ist weniger überlastet, die Benutzer werden in die verantwortungsvolle Verwaltung ihrer Daten involviert, Kosten für Hardware und Verwaltungsaufwand werden drastisch reduziert, und Leistung und Produktivität werden gesteigert. Von diesem Standpunkt aus ist eine ausgereifte SRM-Lösung in Verbindung mit einem effizienten Dateneduplizierungs-Tool ein hervorragendes Mittel, Ihre Speicherinfrastrukturen optimal zu gestalten, was sich auch im Ergebnis widerspiegelt.

ÜBER NORTHERN

Northern ist ein internationaler Software-Anbieter, der auf die Entwicklung von zuverlässigen, flexiblen und einfach zu benutzenden Speichermanagementlösungen spezialisiert ist. Die vier erklärten Ziele des Speicherressourcenmanagement der Northern Storage Suite sind die Identifizierung und Freisetzung verschwendeter Speicherkapazitäten, die Kontrolle des Userverhaltens, die Planung zukünftigen Speicherbedarfs und die Integration der Enduser in das Speichermanagement. Mehr als 28.000 Unternehmen weltweit und mehr als die Hälfte der Unternehmen der Fortune Global 100 setzen die Software der Northern für die Kontrolle und Verwaltung ihrer Speicherumgebung ein. Die Northern ist über ihre sechs Niederlassungen auf dem globalen Markt vertreten und bedient von dort aus ihre Kunden: Vereinigte Staaten [Tampa, FL], Schweden [Stockholm], Frankreich [Biarritz], Vereinigtes Königreich [London] Deutschland [München] und BeNeLux [Amsterdam].

Die Northern Storage Suite sowie die vollständige, ergänzende Auswahl an Netzwerkadministrationshilfsprogrammen kann weltweit bei Wiederverkäufern (in 86 Ländern), autorisierten Verteilern und direkt bei Northern erworben werden. Kunden in Nordamerika sollten Northern (US-Geschäft) unter 1-800-881-4950 oder unter sales@northern.net kontaktieren. Kunden in anderen Teilen der Welt sollten die Webseite der Northern-Wiederverkäufer unter www.northern.net besuchen, um den nächsten Northern-Software-Lieferanten zu finden, oder Northern (Europa-Geschäft) unter +46 8 457 50 00 kontaktieren.