

**[BI BRIEF]****mayato Fact Sheets im Umfeld Business Intelligence und Customer Relationship Management****Titel: Nearline Storage in SAP BW 7.0****Autor(en): Dr. Marcus Dill****Themengebiet: SAP BW****Veröffentlicht: März 2008****Management Summary**

Der Begriff Nearline Storage (NLS) ist eine Verkürzung von "Near Online Storage" und repräsentiert eine Speichertechnologie, die zwischen den beiden klassischen Optionen einer Onlinespeicherung innerhalb der Anwendungsdatenbank (d.h. Daten sind direkt und schnell zugreifbar) und einer Offline Archivierung (d.h. Daten können nur nach erneuten Laden in den Onlinebereich genutzt werden) liegt. NLS erlaubt das Auslagern von Daten aus der Onlinedatenbank bei gleichzeitig fortgesetzter Verfügbarkeit der Daten für die jeweiligen Anwendungen. Während operative Anwendungen oftmals ohne historische Daten auskommen, benötigen analytische Anwendungen (z.B. OLAP-Reports) sehr oft gerade historische Daten. Aufgrund des Ad-hoc-Charakters von Datenanalysen ist auch oftmals schwer vorherzusagen, wie lange und wie oft bestimmte Daten benötigt werden. Aus diesem Grund ist in einem Data Warehouse wie im SAP BW die Entscheidung über die Freigabe von Daten zur Archivierung meist schwierig. Stattdessen bleiben die Daten i.d.R. relativ lange in der Onlinedatenbank des SAP BW liegen und tragen so zum gigantischen Datenwachstum (nicht selten im Terabytebereich) bei. Mit der Version 7.0 (NW 2004s) bietet SAP BW eine zertifizierbare Schnittstelle für NLS und ermöglicht so im Rahmen der Standardprozesse des Data Warehouse, aber mit Hilfe von zusätzlicher Software, was in früheren Releases nur als Projektlösung angeboten wurde: Daten können bereits frühzeitig aus der Onlinedatenbank des SAP BW entfernt und in günstigere Speichermedien umgelagert werden. SAP BW kann mit Hilfe dieser Schnittstelle signifikant entlastet werden, was für Anwender spürbar positive Effekte auf die Performance der Mehrheit von Datenanalysen hat. Benötigen Anwender dennoch auch Zugriff auf die historischen Daten, so wird dies durch die NLS-Schnittstelle ermöglicht. Die Laufzeiten für den Zugriff auf NLS sind gemeinhin nur wenig schlechter als im Onlinezugriff. Ein weiterer Vorteil für Anwender ist, dass sie Daten meist schneller erhalten, da die Ladeprozesse aufgrund des geringeren Datenvolumens verkürzt sind. Gleichzeitig sinken die Kosten für den Betrieb des SAP BW: Onlinedatenbanken erfordern meist teurere Hardware. Vor allem aber kostet die Aufrechterhaltung der Onlinedatenbank erheblichen Aufwand für den Systembetrieb (Backups, lange Ladezeiten etc.). Die Investitionen in NLS amortisieren sich typischerweise schon im zweiten Jahr.

Fakten und Empfehlungen

Die Chancen von NLS in BW 7.0 liegen in der Reduktion von Kosten im Systembetrieb für Speicherhardware und für Datenmanagementprozesse (z.B. Backups, ETL-Prozesse, Systemkopien) sowie in der verbesserten Systemperformance. Daten können mit geringerem Zeit- und Personalaufwand als bisher bereitgestellt werden, Querylaufzeiten sinken und die Dauer von Ladeprozessen (Datenladen und Aktivieren) nimmt deutlich ab. Die Auslagerung von Daten in NLS kann über Datenarchivierungsprozesse und Prozessketten automatisiert werden.

NLS birgt aber auch mehrere Einschränkungen: Daten in NLS können nicht mehr aktualisiert werden. Zu Zeitscheiben, die bereits in NLS geschrieben wurden, können keine neuen Daten in denselben InfoProvider geladen werden. Falls dies doch erforderlich ist, so müssen die Daten zuvor wieder in die Onlinedatenbank geladen werden, was die NLS-Schnittstelle zu diesem Zweck explizit vorsieht. NLS kann nur genutzt werden für Ladeprozesse auf der Basis von BW 7.0 Ladetechnologie mit Datentransferprozessen und Transformationen, da die herkömmliche Ladetechnologie mit InfoPackages und Fortschreibungsregeln keine Prüfung auf nearline gespeicherte Zeitscheiben durchführt. Eine weitere Einschränkung besteht darin, dass Zugriffe auf NLS-Daten nur für OLAP-Queries vorgesehen sind. Für Zugriffe innerhalb anderer BW-Anwendungen (z.B. InfoSets, Analyseprozessdesigner) oder durch externe Analysetools wie SAS, die direkt auf die BW-Datenbank zugreifen, müssen zusätzliche Lösungen entwickelt und vorhandene Analysen migriert werden. Auch für Ad-hoc-Queries über den BW OLAP-Server ist kein Zugriff auf NLS-Daten möglich. Alternative Frontends, die über OLE DB for OLAP auf BW-Daten zugreifen, haben hingegen vollen Zugriff auf NLS.

Die Investitionen in NLS umfassen Hardware, Software und Dienstleistungen. Die benötigte Hardware für NLS ist günstiger, da sie nicht dieselben Anforderungen hinsichtlich Datensicherheit und Verfügbarkeit hat. Die Lizenzkosten sind bei den meisten NLS-Anbietern an das gespeicherte Datenvolumen gekoppelt. Für eine Kostenabschätzung kann von 50.000 € pro einem Terabyte an Daten ausgegangen werden. Zusätzlich verlangen NLS-Anbieter eine Wartungsgebühr (ca. 20% p.a.). Der größte Kostenblock liegt im Bereich Personal und Dienstleistungen für das initiale Implementierungsprojekt, im Rahmen dessen geeignete InfoProvider identifiziert, Altersgrenzen für Onlinedaten bestimmt, Datenladeprozesse für NLS umgestellt, Standards definiert, die NLS-Software aufgesetzt und ggf. auch Nicht-OLAP-Anwendungen umgestellt werden. Der Aufwand für ein solches Projekt ist stark abhängig von Faktoren wie Datenvolumen, Anzahl auf NLS umzustellender InfoProvider und Ladeprozesse, Art der vorhandenen und zukünftigen Analysen. Für ein typisches Pilotprojekt (z.B. NLS von 1TB Onlinedaten aus 10-20 Data-Store-Objekten) ist durchaus mit einem Aufwand von über 100 Personentagen zu rechnen. Der Rollout auf weitere Datenbestände wird bei guter Standardisierung von Datenmodellen und Prozessen deutlich weniger Aufwand mit sich bringen. Rechnet man alleine die Kostenersparnis auf der Betriebsseite – Branchenexperten gehen von jährlichen Kosten von 200.000 € pro TB produktiver SAP-Daten aus - so amortisiert sich ein NLS-Projekt spätestens im zweiten Jahr.

Zurzeit sind drei NLS-Anbieter für die SAP-Schnittstelle zertifiziert: [FileTek](#), [SAND Technology](#) und [PBS](#). FileTek StorHouse/BI ist eine vollständige Lösung für das Speichermanagement, die u.a. auch Backups und einen späteren, altersabhängigen Wechsel des Archivmediums der Daten (billig statt schnell) ermöglicht. Kunden, die noch keine Archivlösung besitzen, könnten diese Lösung präferieren. SAND DNA hingegen speichert NLS-Daten hochkomprimiert und voll indiziert auf einer Festplatte. Dies ermöglicht eine Reduktion des Datenvolumens und sehr gute Performance für Querys gegen NLS-Datenbestände (~Faktor 2 relativ zu reinen Onlinequeries). Allerdings beinhaltet SAND DNA keine weiteren Datenmanagementfunktionen. Backups etc. müssen über zusätzliche Prozesse (ggf. auch Software) etabliert werden, sind angesichts der hohen Kompression allerdings deutlich weniger aufwändig als in der Onlinedatenbank. PBS schließlich bietet sein Addon CBW in verschiedenen Versionen: die Standardversion für 7.0 bietet zusätzlich zur Speicherung von Daten aus InfoCubes und Data-Store-Objekten auch die Speicherung von PSA-, Stammdaten- oder beliebiger SAP-Tabellen. Zusätzlich bietet PBS eine Version, die auf Sybase IQ basiert und funktional SAND DNA vergleichbar ist.

Vertiefung

- › Zwei Onlineartikel von Bill Inmon ([Teil 1](#) und [Teil 2](#)) bieten einen guten Überblick über NLS im Data Warehouse.
- › Die drei genannten NLS-Anbieter bieten auf ihren Homepages jeweils umfangreiche Informationen.
- › Die Onlinehilfe der SAP enthält alle wichtigen Details zur Umsetzung innerhalb SAP BW unter Business Intelligence → Data Warehousing → Data Warehouse Management → Information Lifecycle Management
- › Projektverantwortliche sollten in jedem Fall die relevanten OSS-Hinweise unter der Komponente BW-WHM-DST-ARC berücksichtigen. Die Hinweise enthalten wichtige Informationen über funktionale Einschränkungen und vorhandene Workarounds.