

INNOWACYJNE ROZWIĄZANIA PAMIĘCI MASOWYCH NETAPP DLA NOWOCZESNYCH TECHNOLOGII INFORMACYJNYCH

Bogusz Błaszkiwicz

NetApp

Zbigniew Błaszkiwicz

Institut Inżynierii Rolniczej, Uniwersytet Przyrodniczy w Poznaniu

Streszczenie. W pracy przedstawiono nowoczesne pamięci masowe NetApp poprzez krótką prezentację ich cech sprzętowych oraz oprogramowania zarządzającego, niezbędnego dla efektywnego zawiadywania przechowywanymi danymi. Zaprezentowano zagadnienia najnowocześniejszych rozwiązań sprzętowych, konsolidacji i administracji danych oraz ochrony przed ich utratą. Poruszono również kwestię oprogramowania systemowego i jego możliwości współpracy z nowoczesnymi technologiami informacyjnymi, efektem czego jest umiejętność sprostania zróżnicowanym wymaganiom odbiorców i użytkowników.

Słowa kluczowe: pamięci masowe, ochrona danych, deduplikacja, konsolidacja, oprogramowanie

Wstęp

Szybki, wydajny dostęp do stale rosnącego zasobu informacji jest niezbędny dla gromadzenia coraz większej liczby danych w rozwijanej innowacyjnej gospodarce, w tym w szeroko pojętym rolnictwie i w jego otoczeniu oraz szkolnictwie wyższym dla zwiększania efektywności i rozwoju dydaktyki oraz badań naukowych. Dla spełniania postępujących oczekiwań postępuje modernizacja rozwiązań technologii informacyjnych w zakresie konsolidacji danych (wykorzystywanych i często zmienianych przez wielu użytkowników naraz), administracji danymi rozproszonymi czy ochrony przed awariami sprzętowymi, błędami ludzkimi, kataklizmami – ogólnie utratą danych. Jednym z głównych zadań pamięci masowych oraz oprogramowania zarządzającego jest konsolidacja wielorakich środowisk informatycznych – aplikacji biznesowych, serwerów pocztowych, a przede wszystkim baz danych i systemów wirtualnych, niezależnie od metody dostępu do informacji, przy zapewnieniu redundancji środowiska informatycznego w szerokim stopniu niezawodności i przy najkrótszym czasie odtwarzania danych. Szczególne możliwości w tym zakresie zapewniają należące do najnowocześniejszych na świecie rozwiązania firmy NetApp.

W niniejszej pracy przedstawiono produkty w dziedzinie pamięci masowych NetApp oraz dokonano analizy innowacyjnych rozwiązań zwiększających możliwości ich zastosowania w technologiach informacyjnych.

Pamięci masowe NetApp

NetApp jest firmą o zasięgu globalnym zatrudniającą ponad 8 tysięcy pracowników. Jej bogate portfolio obejmuje wdrożenia w największych organizacjach, koncernach i firmach na świecie, w tym od ponad pięciu lat w Polsce. Dotyczą one nowoczesnych środowisk technologii informacyjnych wielu działów gospodarki, kultury, szkolnictwa i nauki, itp. obejmując także coraz liczniejsze grono uczelni wyższych, w tym przyrodniczych i rolniczych. Podstawowy podział produktów jest następujący:

- małe (low-end): FAS2xxx (65 - 99TB, 68 – 104 dysków, SAS, FC & SATA),
- średnie (midrange): FAS3xxx (84 - 504TB, 240 - 960 dysków, SAS, FC & SATA),
- bardzo duże (high-end): FAS62xx (840 - 1176TB, 1200 - 1440 dysków, SAS, FC & SATA),
- integrujące środowiska heterogeniczne: V3xxx (84TB - 504TB, 168 - 504 LUNs) i V6xxx (840 - 1176TB, 840 - 1176 LUNs).

Firma NetApp dostarcza rozwiązania umożliwiające nie tylko efektywne zarządzanie pamięcią masową, ale również zmniejszające jej kompleksowość i koszty utrzymania oraz skuteczną kontrolę zmian. Za pomocą macierzy dyskowych oraz oprogramowania zarządzającego infrastrukturą firma NetApp konsoliduje wielorakie środowiska informatyczne – aplikacje biznesowe, serwery pocztowe, a przede wszystkim bazy danych i systemy wirtualne – niezależnie od metody dostępu do informacji. Technologia Netapp integruje wszystkie swoje macierze dyskowe, niezależnie od platformy, za pomocą jednolitego systemu operacyjnego o nazwie Data Ontap. Systemy FAS używają redundantnych, wymiennych „online” komponentów oraz stosują technologię RAID-DP gwarantującą ochronę bitów bez strat wydajności poprzez zastosowanie dwóch dysków parzystości w jednej grupie RAID gwarantującej dostęp do danych nawet przy awarii dwóch jednostek dyskowych.

Ochrona danych

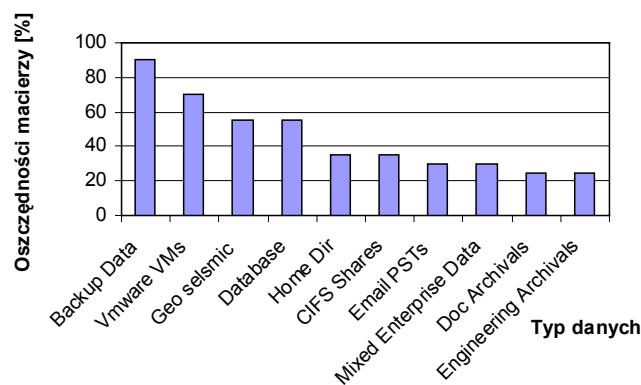
Do najważniejszych funkcji macierzy pamięci masowych w zakresie konsolidacji danych wykorzystywanych i często zmienianych przez wielu użytkowników naraz, należy ochrona przed utratą danych spowodowaną awariami sprzętowymi, błędami ludzkimi czy kataklizmami.

Unikalnym, opatentowanym przez Netapp rozwiązaniem ochrony danych jest tzw. kopia migawkowa, z angielskiego Snapshot (nazwa wymyślona i opatentowana przez Netapp, a używana przez inne firmy na podstawie licencji). W teorii jest to tablica wskaźników do bloków danych, która zapamiętuje ich treść zachowując bardzo małe rozmiary (kilka kilobajtów). Snapshot utrzymuje trwałość tych danych wymuszając na nich zapis zmiany do nowych bloków na dysku, dzięki czemu nawet, jeśli zostaną usunięte pliki po stronie systemu operacyjnego, tak długo jak mamy Snapshot, możemy je odtworzyć. Unikalność rozwiązania wynika z faktu, iż Snapshoty Netapp, jako jedyne na rynku nie opierają się na technice COFW. Migawki są statycznym zbiorem wskaźników i we współpracy z systemem plików WAFL nie mają najmniejszego wpływu na wydajność macierzy, dzięki czemu można tworzyć ich bardzo dużą ilość (do 255 per wolumen logiczny) dla zabezpieczenia danych w różnych punktach w czasie.

Efektywność macierzy

Wspomniane bloki danych to najmniejsza jednostka logiczna, w jakiej można definiować pojemność macierzy. Wszelkie informacje trafiające na dyski są zapisywane właśnie w tych blokach o z góry zdefiniowanej, stałej wielkości czterech kilobajtów. Stworzona przez Netapp od podstaw technologia o nazwie Deduplikacja dba o to, aby na macierzy nie były przechowywane dokładnie takie same bloki. Uruchamiany przez administratora algorytm skanuje wskazany wolumen danych porównując sumy kontrolne (identyfikatory) bloków i w przypadku natrafienia na dwa identyczne, weryfikuje bity w nich zapisane; jeżeli są identyczne, drugi blok z tą samą zawartością jest zastępowany wskaźnikiem do pierwszego. Co istotne, jest to proces całkowicie odwracalny, co ułatwia decyzję o jego zastosowaniu. Pojemność macierzy, jaką można w ten sposób zaoszczędzić, waha się pomiędzy piętnaście, a nawet dziewięćdziesiąt procent (dla systemów wirtualnych) powierzchni deduplikowanej. W konsekwencji uzyskuje się wiele oczywistych korzyści, ostatecznie też finansowych.

Macierze NetApp dają bardzo wymierne korzyści w postaci oszczędności powierzchni dyskowej (rys. 1). Jako jedyne na rynku IT umożliwiają bezstratne dla pojemności tworzenie klonów systemów 1:1 na potrzeby testów rozwojowych czy produkcyjnych. Jedna grupa dysków okupowana przez dynamiczną bazę danych obciążaną przez szereg testów dokonywanych przez programistów szybko osiągnie swój limit wydajności, co więcej, zmiany przez nich dokonywane będą trudne do zsynchronizowania. Stworzenie klonu takiego systemu na bazie szerzej opisywanego wcześniej Snapshotu nie tylko odciąża bazę danych – wszelkie zmiany są zapisywane do nowych bloków, co zaangażuje więcej dysków, ale również zredukuje do minimum ryzyko dla produkcyjnych danych; testy będą odbywać się poza obszarem krytycznym. Co więcej, takich klonów, z zerowym wstępnym kosztem pojemnościowym, można stworzyć dziesiątki. Tego rodzaju rozwiązanie jest wyprzedzającym inne technologie macierzowe w środowiskach testowych i produkcyjnych.



Źródło: zbiór referencji klienckich firmy NetApp

Rys. 1. Typowe wartości oszczędności pojemności macierzy dzięki deduplikacji wg typu danych.
 Fig. 1. Typical capacity savings by dataset that can be achieved with deduplication (Source: NetApp proven customer stories)

Do nowatorskich rozwiązań należy efektywne zarządzanie pojemnością macierzy. Dotyczy to usługi wbudowanej w oprogramowanie Data Ontap - *Thin Provisioning* (Cienkie Rezerwowanie) - pozwalającej na „pozorną” alokację przestrzeni przeznaczonej na dysk fizyczny po stronie systemu-gościa. W praktyce polega to na zdefiniowaniu logicznego dysku *LUN* o konkretnej pojemności alokowanej w całości logicznie, natomiast fizycznie – w miarę przyrostu danych na tym dysku, (tj. dysk zajmuje na macierzy dokładnie tyle bajtów, ile w rzeczywistości się na nim znajduje, niezależnie od zdefiniowanej logicznej wielkości widocznej po stronie systemu operacyjnego). Pozwala to na bardziej efektywne zarządzanie pojemnością macierzy, gdyż nigdy nie marnujemy przestrzeni dysków fizycznych realnie zużywając bloki niewykorzystane przez systemy.

Konsolidacja pamięci masowych

Systemy Netapp zapewniają wszechstronne konsolidacje pamięci masowych. Umożliwiają symultaniczny dostęp do danych typu blokowego i plikowego za pomocą połączeń światłowodowych oraz sieci IP. Jedna macierz Netapp może przechowywać dane środowisk różnorodnego typu, poprzez biznesowe aplikacje, serwery pocztowe i sieciowe, bazy danych, itp., oraz złożonych z serwerów obsługiwanych przez systemy operacyjne MS Windows® czy UNIX® i opartych na protokołach FC, iSCSI i IP (TCP/UDP). Macierze NetApp posługują się tym samym systemem operacyjnym oraz korzystają z tej samej grupy dysków o wysokiej gęstości, rozbudowa danego modelu do odpowiednika wyższej generacji może odbyć się w trybie „online”.

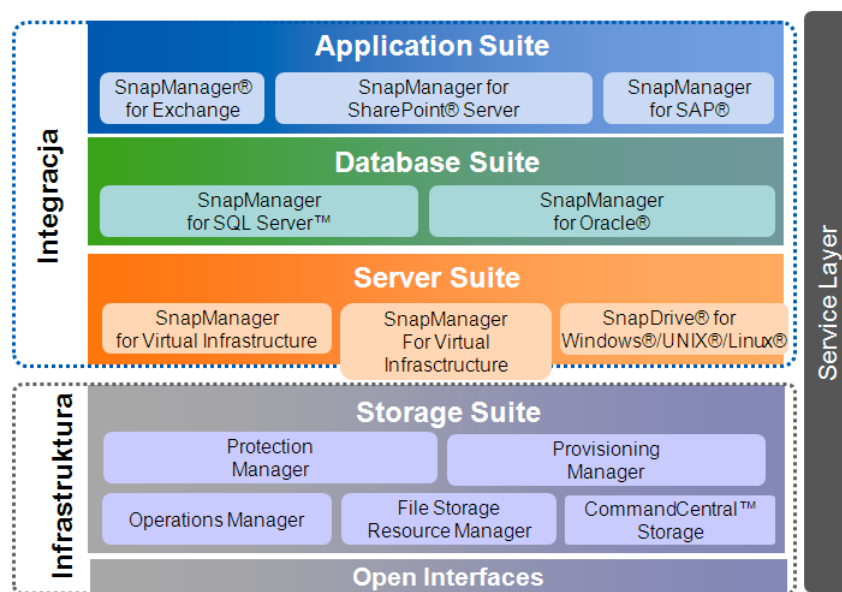
Oprogramowanie

Pamięć masowa to również szerokie spektrum oprogramowania (rys. 2), za pomocą którego macierze realizują zadania administracji systemem kopii zapasowych, zapewniają funkcjonalność DR (*Disaster Recovery*), utrzymują synchroniczną kopię danych produkcyjnych. Firma Netapp udostępnia łącznie ok. 100 produktów oprogramowania dla stabilnego i wydajnego rozwiązania IT. Aby ułatwić użytkownikowi kompleksowo poruszanie się po oprogramowaniu zależnie od pożądanej funkcji, Netapp rozwija system graficznego zarządzania macierzami o nazwie Netapp System Manager, która integruje w sobie dostęp do wszystkich modułów programowych niezbędnych do całościowej administracji środowiskiem pamięci masowych. To nowatorskie podejście do zarządzania pozwala na zwiększenie przystępności technologii dla użytkownika.

Podsumowanie

Nowoczesne pamięci masowe NetApp, oferowane w wielu elastycznych konfiguracjach, są narzędziem o szerokim zastosowaniu w wielu dziedzinach gospodarki, szkolnictwa i nauki. Mogą stanowić jeden z wydajnych i efektywnych elementów infrastruktury informatycznych uczelni wyższych, opierających na nowoczesnej informatyce swoje systemy nauczania, badań naukowych i zarządzania jednostkami. Prawidłowe przechowywanie informacji, szybki i wydajny do nich dostęp oraz zarządzanie nimi zapewniają unikalne i postępowe rozwiązania w zakresie konsolidacji danych, administracja danymi rozproszo-

nymi czy ochrona przed ogólnie pojmowaną utratą informacji. Wysoką efektywność, uniwersalność i funkcjonalność zapewnia bogate portfolio oprogramowania integrującego zróżnicowane środowiska informatyczne zgodnie z wymaganiami użytkowników i odbiorców.



Źródło: materiał firmy NetApp

Rys. 2. Rodzina oprogramowania firmy NetApp z podziałem wg przeznaczenia i zakresu działania

Fig. 2. Netapp Management Software Family as per design and scope of functions

Bibliografia

NetApp Promotion Materials. Materiały reklamowe. Dostępne w Internecie: <http://www.Netapp.com>

INNOVATIVE NETAPP STORAGE SOLUTIONS FOR MODERN INFORMATION TECHNOLOGY

Abstract. The paper presents modern NetApp storage systems through a brief description of their features hardware and software management, necessary for effective governing of stored data. This document presents the issues of cutting-edge hardware, consolidation and administration of data and protection from loss. It also issues system software and its capabilities in advanced information technologies, the effect of which is the ability to meet the diversified demands of customers and users

Key words: management software, storage, disaster recovery, data protection, backup, data consolidation, virtualization

Adres do korespondencji:

Zbigniew Błaszkiwicz; e-mail: zbigniewblaszkiewicz@up.poznan.pl
Instytut Inżynierii Rolniczej
Uniwersytet Przyrodniczy w Poznaniu
Wojska Polskiego 50
60-637 Poznań