Windows Server 2003. Wysoko wydajne rozwiązania

Autorzy: Jeffrey R. Shapiro, Marcin Policht
Tłumaczenie: Paweł Gonera
ISBN: 83-246-0246-1
Tytuł oryginału: Building High Availability Windows Server (TM) 2003 Solutions (Microsoft Windows Server System)
Format: B5, stron: 472

Zbuduj wydajne środowisko Windows Server

- Dobierz sprzęt odpowiedniej jakości
- Wykorzystaj mechanizmy klastrowania do osiągnięcia maksymalnej dostępności
- Zapobiegaj awariom i przestojom w pracy serwerów

Platforma Windows Server 2003 zyskuje coraz większą popularność. Firmy odchodzą od rozwiązań opartych na innych technologiach, uruchamiając serwery wykorzystujące tę właśnie platformę. Jednak wymiana systemu operacyjnego na inny nie jest prostym zadaniem. Podczas wdrażania środowiska Windows Server 2003 należy uwzględnić wiele czynników, dzięki którym system pozostanie niezawodny przez 24 godziny na dobę, 7 dni w tygodniu, 365 dni w roku.

Książka „Windows Server 2003. Wysoko wydajne rozwiązania" przedstawia praktyczne zagadnienia związane z wdrażaniem i administrowaniem systemami operacyjnymi z rodziny Windows Server 2003. Opisuje proces planowania oraz implementacji rozwiązań opartych na klastrach, mechanizmach równoważenia obciążenia i technikach szybkiego przywracania serwerów do pracy po awariach i aktualizacjach. Ilustrowane przykładami zagadnienia oraz łatwe do wykorzystania instrukcje pomogą Ci podjąć szybkie i trafne decyzje.

- Wybór sprzętu
- Pamięci masowe przeznaczone dla serwerów
- Projektowanie sieci o maksymalnej dostępności
- Klasteryzacja Windows
- Wysoko wydajne serwery wydruków i plików
- Maksymalizacja wydajności i dostępności SQL Servera oraz Exchange Servera
- Równoważenie obciążenia
- Korzystanie z Microsoft Operation Manager

Chcesz zmaksymalizować dostępność, skalowalność i wydajność środowiska Windows Server? Koniecznie sięgnij po tę książkę.

Wydawnictwo Helion
ul. Kościuszki 1c
44-100 Gliwice
tel. 032 230 98 63
e-mail: helion@helion.pl

Spis treści

ROZDZIAŁ 1.

Świat przetwarzania wysoko wydajnego i wysokiej dostępności w Windows

Wstęp

W tym rozdziale przedstawimy przegląd pojęć związanych z *wysoką dostępnością* (HA), *przetwarzaniem wysoko wydajnym* (HPC) oraz z tym, w jaki sposób platforma Windows Server 2003 firmy Microsoft realizuje wymagane mechanizmy.

Rozpoczniemy od zapoznania się z podstawami poziomu obsługi, dostępności, pojęciami pracy pomimo usterek, nadmiarowości, skalowalności, wysokiej dostępności, jak również technologii zarządzania działaniem systemu. Zapoznamy się z zapobiegawczym zarządzaniem systemami komputerowymi na potrzeby wysokiej dostępności, a także z funkcjami reaktywnymi wymagającymi wcześniejszego zaplanowania, na przykład usuwaniem skutków awarii. Następnie zajmiemy się wysoko wydajnymi architekturami, sprzętem i oprogramowaniem.

Głównym punktem i przeznaczeniem tego rozdziału jest szerokie przedstawienie tematu, który staje się coraz bardziej skomplikowany. Następnie, w kolejnych rozdziałach, zostaną omówione szczegóły. Przedstawimy tu nowe terminy, wyjaśnimy stare, kilka z nich przedefiniujemy, a następnie przygotujemy grunt do przedstawienia nowości, jakie firma Microsoft dodała do swojego najnowszego systemu operacyjnego (mają one wspomóc zespół zarządzający oraz analityków systemu i architektów).

Zanim rozpoczniemy tę podróż, wspomnimy dwie publikacje, które w znacznym stopniu wpłynęły na nasz sposób postrzegania krytycznych systemów Microsoft — *Microsoft Operations Framework* (MOF) oraz *Microsoft Systems Architecture* (MSA).

MOF zawiera wiele informacji na temat tworzenia architektury działania, planowania operacji, zarządzania zmianami, funkcji zabezpieczeń, poziomu obsługi i tak dalej. MSA zawiera niezrównany opis architektury sieciowej, usług przechowywania i składowania, usług katalogowych, usług udostępniania plików i drukarek, klasteryzacji, usług rozproszonych i tak dalej. Obie publikacje są obowiązkowymi lekturami dla każdego architekta systemów wysoko dostępnych. MSA zawiera również tak zwane *Solution Accelerators* (przyspieszacze rozwiązań), które pomagają przy planowaniu, testowaniu i uruchamianiu różnych systemów.

Przy projektowaniu architektury systemów można zastosować kilka podejść. W książce tej autorzy używają podejścia bazującego na modelu Zachman. Z modelem Zachman można się zapoznać, przeglądając witrynę *www.zachmanframework.com*, która jest prowadzona przez Zachman Institute for Framework Advancement (ZIFA).

Poziom usługi

Poziom usługi to najczęściej nadużywany termin w zarządzaniu sieciami LAN i technologiami serwerowymi. To niewłaściwe używanie terminu wynika z istnienia wielu interpretacji. Na przykład, niektórzy menadżerowie i analitycy korzystają z tego terminu do określenia, jak „dostępny" jest lub powinien być pojedynczy system komputerowy. Może być to prawda w świecie schyłkowych systemów, które były w większym stopniu monolitami, a nie zbiorem rozproszonych i połączonych ze sobą komponentów, lub w świecie telekomunikacji, do opisania central i systemów PBX. Analitycy systemów i sieci są często proszeni o zapewnienie w systemie odpowiedniego poziomu obsługi, bez faktycznej wiedzy o tym, co ten poziom obsługi ma zapewniać.

SYSTEM

Systemem może być pojedynczy komputer, wiele serwerów (oraz klaster serwerów), a nawet wiele lokacji serwerów. System nie jest kompletny bez administratora systemu, będącego osobą lub komputerem, który wymaga również operatora. Do działania potrzebuje on zasilania, szaf, stojaków, budynków, ochrony przed ogniem i tak dalej.

Próby zachowania poziomu usługi i zapewnienia działania systemu komputerowego przez 24 godziny, 7 dni w tygodniu często prowadzą do złego ukierunkowania usług. Lepiej unikać poziomów usługi opisujących dostępność systemów komputerowych; zamiast tego powinno się podawać ilość czasu w okresie operacyjnym, przez który usługa lub aplikacja jest dostępna dla użytkowników lub subskrybentów.

Poziom usługi określa, na ile usługa jest dostępna w okienku serwisowym. Nie tylko sam system musi spełniać wymagania poziomu usługi, ale także usługa, być może wykorzystująca wiele systemów i komponentów, w tym zasoby ludzkie, musi być dostępna.

Gdy ocenimy lub określimy poziom usługi dla aplikacji lub usługi, na początek musimy zapytać, jaki jest oczekiwany okres działania dla tej usługi. *Gwarancja jakości świadczonych usług* (SLA), która zostanie opisana w jednym z kolejnych punktów, jest umową, najczęściej zatwierdzaną przez jej podpisanie przez zainteresowane strony, w której zakłada się, że usługa będzie dostępna przez większość czasu w okresie operacyjnym.

Na początek należy zdefiniować pojęcie okresu operacyjnego. Czy usługa jest wymagana przez 12 godzin dziennie, 18 godzin, czy też przez całą dobę? Po ustaleniu okresu operacyjnego musimy dowiedzieć się, przez jaki czas w okresie operacyjnym firma lub użytkownik usługi może tolerować przerwy w pracy usługi. Jeżeli okresem operacyjnym jest 12 godzin i w tym czasie wymagane jest nieprzerwane działanie usługi, to wymaganym poziomem usługi jest 100 procent. Jeżeli użytkownik lub firma może tolerować, załóżmy, 45-minutowy okres przestoju w oknie operacyjnym, to poziom usługi może być ustalony na 90 lub 95 procent.

Po określeniu i uzgodnieniu poziomu usługi, specjaliści IT oraz analitycy systemów mogą rozpocząć budowanie systemu składającego się z komponentów sprzętowych i programowych, które muszą być dostępne przez czas określony w SLA. Parametr ten jest nazywany *dostępnością* i jest mu poświęcony następny punkt. W tabeli 1.1 przedstawione są przykłady kategorii wymaganego poziomu usługi w typowej, obciążonej witrynie handlu elektronicznego.

Tabela 1.1. *Kategorie poziomu usługi dla witryny handlu elektronicznego*

Kategoria poziomu usługi	Wymagany poziom usługi
Godziny działania	♦ Usługa musi być dostępna 24×7×365.
	♦ Nie powinno być wyłączeń w czasie planowanej konserwacji komponentów.
	♦ Jedynym dozwolonym przypadkiem zawieszenia usługi jest przypadek poważnego zagrożenia bezpieczeństwa, na przykład atak na serwery, gdy konieczne jest zapobieżenie uszkodzeniu serwerów.
Wydajność	♦ Usługa musi być w stanie obsłużyć co najmniej 2500 jednoczesnych połączeń bez zauważalnego przez użytkowników obniżenia szybkości działania.
	♦ Opóźnienie pomiędzy przesłaniem i potwierdzeniem nie powinno być większe od dwóch sekund.
	♦ System musi obsługiwać około 25 000 transakcji na godzinę.

WSKAZÓWKA: W czasie negocjowania umowy SLA należy rozważyć wszystkie aspekty dostępności systemu, w szczególności zarządzanie działaniami, które zawsze obejmuje zarządzanie zasobami ludzkimi. Nawet w najbardziej dostępnych systemach wymagani są operatorzy, ponieważ nie istnieją systemy, które nigdy nie ulegają awariom.

Książka ta nie jest forum dyskusyjnym na temat SLA; dobrym źródłem informacji na ten temat jest konsorcjum International Engineering Consortium (IEC), którego witryna jest dostępna pod adresem *www.iec.org*.

Dostępność

Systemy, jakie opracowujemy w celu spełnienia określonych potrzeb biznesowych lub operacyjnych, powinny być dostarczane i budowane w taki sposób, aby spełnić uzgodniony poziom usługi. Można się z tym nie zgadzać i dowodzić, że podstawowym wymaganiem jest bezpieczeństwo systemu, integralność danych lub cena. Oczywiście zagadnienia te są ważne i są one częścią zadania.

Integralność systemu, bezpieczeństwo i zarządzalność są ważne, ale nie muszą być one rozważane *przed* określeniem poziomu usługi. W końcu warunki założenia systemu wpływają na jego dostępny potencjał. Jeżeli można włamać się do systemu, będzie on zbudowany z tanich komponentów lub źle zaprojektowany, to będzie on mniej dostępny i przez to nie będzie spełniał wymagań poziomu dostępności dla działających w nim aplikacji.

DOSTĘPNOŚĆ

Dostępność można zdefiniować jako ilość czasu w oknie obsługi, przez który aplikacja lub usługa jest dostępna dla użytkownika. Na przykład, macierze RAID to urządzenia pamięci masowej, które są dostępne również w przypadku awarii jednego, dwóch, a nawet trzech dysków. Nie jest to ten sam parametr co niezawodność, choć nie trzeba chyba tłumaczyć, że do naszych celów należy wykorzystywać niezawodne komponenty.

Mówiąc o dostępności systemu, na przykład Exchange, mamy na myśli procent czasu (w oknie operacyjnym), przez który usługa działa i możliwe jest wysyłanie i odbieranie poczty.

Istnieją różne poziomy dostępności. Możemy powiedzieć, że system jest średnio dostępny, gdy składa się z komponentów i technologii, które potencjalnie mogą powodować awarie systemu lub przerwy w działaniu, wpływające na poziom usługi i powodujące łamanie SLA.

Niektóre małe firmy mogą tolerować dłuższe okresy przerwy w działaniu niż większe firmy lub dostawcy usług, którzy muszą wypełniać swoje zobowiązania. Gdy z serwera wydruku korzysta tylko kilka osób w czasie normalnego dnia pracy trwającego od 9 do 17, to jego niedostępność przez godzinę lub dwie nie jest postrzegana jako krytyczne zagrożenie dla działania firmy. Jeżeli jednak w tym samym oknie obsługi z serwera wydruku korzysta kilkaset osób, to 15-minutowa przerwa w pracy może mieć katastrofalny wpływ na firmę.

W drugim scenariuszu wiadomo, że w celu spełnienia poziomu usługi i ciągłości działania, system musi być zbudowany z zastosowaniem architektury wysoko dostępnej i komponentów o dużej wydajności. Wcześniej stwierdziliśmy, że w przypadku mniej krytycznych potrzeb pojedynczy komputer może służyć jako serwer wydruku; jednak w drugim przypadku w celu zapewnienia poziomu usługi wymagany jest zaawansowany klaster serwerów, zapewniający natychmiastowe odtworzenie usługi w przypadku awarii jednego z węzłów.

Od dawna dostępność usługi systemów komputerowych i oprogramowania była mierzona procentem czasu działania. Model dziewiątkowy określa procent dostępności, gdzie 99,9999 (sześć dziewiątek) jest wartością największą, często wykorzystywaną do opisania systemu poczty elektronicznej lub serwera bazy danych, bez odpowiedniej wiedzy, do czego odwołują się te dziewiątki.

Dostępność jest typowo mierzona za pomocą *dziewiątek*. Na przykład, rozwiązanie o poziomie dostępności *trzech dziewiątek* jest w stanie udostępniać swoje funkcje przez 99,9 procent czasu, co jest odpowiednikiem rocznego czasu wyłączenia wynoszącego 8,76 godzin w przypadku działania w trybie 24×7×365 (24 godziny dziennie, siedem dni w tygodniu, 356 dni w roku). W tabeli 1.2 wymienione są standardowe poziomy dostępności, które próbuje osiągnąć wiele organizacji.

Tabela 1.2. *Dostępność opisana systemem dziewiątkowym*

Dostępność (%)	Roczny czas wyłączenia dla działania non stop (24×7×365)
99,9999	32 sekundy
99,999	5 minut i 15 sekund
99,99	52 minuty i 36 sekund
99,95	4 godziny i 23 minuty
99,9	8 godzin i 46 minut
99,5	1 dzień, 19 godzin i 48 minut
99	3 dni, 15 godzin i 40 minut
95	18,25 dnia
90	36,5 dnia

Sześć dziewiątek (99,9999 procent) oznacza, że rocznie system nie może być wyłączony dłużej niż przez 32 sekundy rocznie. Jasne jest, że w XXI wieku jest to niemożliwe do osiągnięcia. Zwykły system Windows 2003 Server musi być regularnie restartowany, aby mógł zacząć korzystać z poprawek zabezpieczeń oraz aktualizacji.

System zaprojektowany dla trzech dziewiątek jest bardziej realistyczny i jest w stanie spełnić uzgodniony poziom obsługi 99,9% czasu działania, co jest odpowiednikiem 8,76 godzin wyłączenia na rok dla okna działania 24×7×365. W tabeli 1.2 pokazane jest, jak procenty przekładają się na faktyczny czas wyłączenia.

Zanim przejdziemy do zastosowań praktycznych, zastosujemy bardziej naukowe podejście i wrócimy do współczynników równania dostępności. Dostępność jest faktycznie funkcją dwóch czynników: *średniego czasu między awariami* (MTBF) oraz *średniego czasu naprawienia usterki* (MTTR). Oba te czynniki są mierzone w godzinach, dlatego można zastosować następujące równanie:

```
Dostępność = MTBF/(MTBF+MTTR) = .9xxxxxx
```

Spróbujmy nieco bardziej uszczegółowić to równanie. Zajmujemy się tu równaniem dającym w wyniku prawdopodobieństwo awarii komponentu. MTBF określa średni odstęp czasu, mierzony w tysiącach lub dziesiątkach tysięcy godzin pracy (nazywanych również *godzinami czasu działania* lub POH), aż do wystąpienia awarii komponentu. Dlatego MTBF jest obliczany za pomocą następującego równania:

```
MTBF = (średni czas – całkowity czas przestoju) / liczba awarii
```

MTTR to średni czas (zwykle podawany w godzinach), jaki zajmuje naprawienie komponentu. Dlatego jeżeli system oferuje MTBF równe 60 (tysięcy godzin) oraz MTTR równe 4 godziny, to możemy wrócić do dziewiątek w następujący sposób:

```
60/(60+4) = .9375 lub 93.75%
```

Przy tych obliczeniach należy korzystać z tabeli 1.2. W taki sposób można zmniejszać niepożądany lub oczekiwany czas wyłączenia. Aby projektować i tworzyć bardziej niezawodne systemy, należy więc stosować konfiguracje nadmiarowe lub odporne na awarie. Inaczej mówiąc, jeżeli twardy dysk osiągnie milionową godzinę pracy i zawiedzie, ostatnią rzeczą, jaką będziemy się przejmować, jest MTTR. Jeżeli będziemy mieli inny dysk, który może zastąpić uszkodzony, to kto będzie zajmował się naprawą dysku? Obecnie czas MTTR oznacza czas potrzebny na zakupienie nowego dysku lub wyjęcie go z magazynu. Więcej informacji na temat nadmiarowych komponentów przedstawimy w następnym punkcie.

UWAGA: W świecie systemów komputerowych MTTR jest często rozwijany jako *Mean Time To Restore* (średni czas do odtworzenia).

Oczywiście, tak dużych współczynników dostępności nie da się osiągnąć, jeżeli weźmiemy pod uwagę potrzebę regularnej aktualizacji, zabezpieczania przeciwko robakom, wirusom i hakerom, a czasami wymiany uszkodzonego sprzętu. Choć formuła dziewiątkowa jest wygodnym wskaźnikiem referencyjnym, nie jest ona standardem do określania dostępności lub wymagań poziomu usługi. A poza tym, ilu inżynierów przy budowaniu systemu zapisuje MTBF i MTTR każdego komponentu, aby wstawić te wartości do magicznego wzoru dla całego systemu? W przypadku SLA, większość klientów nie rozumie formuły dziewiątkowej i wystarcza im obietnica oczekiwanego dziennego czasu wyłączenia. W dalszej części rozdziału i w całej tej książce będziemy omawiać warunki dostępności dla różnych tworzonych przez nas usług.

Podsumowując, na początek musimy określić poziom usługi, który jest wymagany do zaprojektowania systemu. Następnie należy określić jakiej dostępności oczekujemy od systemu, aby spełniał on poziom usług. Możemy określić trzy poziomy: niska dostępność, średnia dostępność i wysoka dostępność. Po określeniu poziomu usługi i dostępności możemy zacząć budować system spełniający oczekiwania zarówno użytkowników, jak i jego właściciela.

Wysoka dostępność, czas wyłączenia i awarie

System o wysokiej dostępności to taki, który spełnia wymagania wysokiej dostępności zapisane w SLA. Może to oznaczać dowolną technologię, konfigurację, projekt, technikę lub kombinację tych składników, które zapewniają spełnienie warunków SLA.

PROJEKTOWANIE SYSTEMÓW O WYSOKIEJ DOSTĘPNOŚCI

System o *wysokiej dostępności* to system zaprojektowany w celu spełnienia wymagań poziomu usług dla aplikacji lub usługi. Taki system korzysta z różnych komponentów, od nadmiarowych zasilaczy do zaawansowanych węzłów zapewniających pracę systemu pomimo awarii.

Niewiele mówimy na temat planowanego czasu wyłączenia. Inaczej mówiąc, wiemy, że o 2 w nocy serwery muszą być ponownie uruchomione w celu zakończenia zmian wprowadzonych do systemu operacyjnego i programów, które są wynikiem łatania, aktualizacji lub zainstalowania pakietów Service Pack. W świecie dużych komputerów jest to znane pod nazwą programu inicjującego ładowanie systemu (IPL) i jest to *planowany okres niedostępności*.

Nieplanowane wyłączenia to wyłączenia spowodowane awarią serwera lub występujące w przypadku braku odpowiedzi z powodu awarii. Awaria ta może być spowodowana przez błąd oprogramowania lub uszkodzenie jednego z komponentów serwera. Nieplanowane wyłączenia nie są oczywiście planowane. Nie wiemy, kiedy się zdarzą. Mogą wystąpić w godzinach porannych, gdy serwer jest mało obciążony, lub w momencie, gdy serwer obsługuje tysiące aktywnych użytkowników. (Według praw Murphy'ego nieplanowane wyłączenie zdarza się najczęściej w momencie, gdy system lub serwer jest najbardziej obciążony). Poszukując dużej dostępności i wysokiej wydajności, staramy się zminimalizować lub całkowicie wyeliminować nieplanowane wyłączenia. W kolejnych rozdziałach skorzystamy ze standardów i praktyk pozwalających wyeliminować wyłączenie systemu nawet w przypadku konieczności ponownego uruchomienia serwera. Można to zrealizować za pomocą klastrowania, na początek przenosząc wszystkie usługi do aktywnego węzła, a następnie aktualizując, łatając lub przebudowując pasywny serwer. Gdy *naprawiony* serwer zostanie uruchomiony, przenosi się do niego usługi, przerywając działanie usługi tylko na chwilę, w celu przeniesienia połączenia.

W kolejnych rozdziałach omówimy sposób pracy z komponentami i usługami, których awaria może spowodować utratę usługi i przestój. Przedstawiona poniżej lista zawiera przykłady tych usług i komponentów. Wiele z nich jest wymienionych w SLA wraz z dodatkowymi uwagami na temat sposobu ich obsługi.

- **Planowane wyłączenia administracyjne.** Obejmuje to wymianę sprzętu, instalowanie nowych sterowników, oprogramowania podstawowego, poprawek, pakietów Service Pack oraz nowych aplikacji wymagających ponownego uruchomienia komputera.
- **Awarie sprzętu serwera.** Obejmuje to awarie takich komponentów serwera jak kości pamięci, płyta główna, karty rozszerzeń, interfejsy (na przykład karty sieciowej), zasilacze, dyski, kontrolery dysków, procesory i wentylatory (szczególnie wentylatory procesorów).

- **Awarie komponentów sieciowych.** Obejmuje to awarie routerów, przełączników, koncentratorów, okablowania i kart interfejsów sieciowych.
- **Awarie oprogramowania.** Obejmuje to wycieki pamięci, ataki wirusów, uszkodzenie plików i danych, błędy oprogramowania i tak dalej.
- **Awarie lokacji.** Obejmuje to awarie zasilania, zalanie, pożar, huragany, trzęsienia ziemi i ataki terrorystyczne. Lokacja może ulec uszkodzeniu z powodu klęski lokalnej (na przykład lokalnej powodzi) lub regionalnej, takiej jak trzęsienie ziemi.

Pierwszą operacją służącą wykluczeniu wyłączeń będzie zbudowanie klastra działającego pomimo awarii, jednak, jak przedstawimy to w dalszej części książki, w skład projektu systemu o wysokiej dostępności wchodzi nawet sieć rozległa. Klasteryzacja jest faktycznie sposobem na zapewnienie nadmiarowości. Systemy komputerowe o wysokiej dostępności to systemy, w których można przenosić usługi z jednego serwera na inny z zachowaniem minimalnego czasu niedostępności, co pozwala zapewnić stałą dostępność usługi. Budowę klastrów działających pomimo awarii omówimy w dalszej części tego rozdziału, natomiast ich szczegółowy opis znajduje się w części II, „Tworzenie wysoko wydajnych systemów Windows Server 2003".

Na co jednak przyda się cała ta nadmiarowość, jeżeli zniszczeniu ulegnie lokacja? Ostatnio został zalany budynek, w którym znajdował się system komputerowy dużej firmy ubezpieczeniowej z Florydy, a dodatkowo uszkodzeniu uległo główne zasilanie. System został wyłączony (co miało wpływ na tysiące ubezpieczonych) na niemal trzy godziny. Nasz projekt polegał na przeniesieniu systemu do lepszej lokalizacji, w której można zapewnić dostępność systemu. Obecnie ten system działa poza firmą, ukryty w jednym z centrów operacji sieciowych, których właścicielem jest duża firma telekomunikacyjna z Miami. Budynek ten jest w stanie przetrwać zarówno lokalne, jak i regionalne klęski żywiołowe, najsilniejsze huragany, może też pracować na własnym zasilaniu nawet jeżeli cała Floryda zostanie pozbawiona prądu.

Takie udogodnienia stały się dostępne za rozsądną cenę, ponieważ wiele z centrów było budowanych w czasie lat boomu internetowego, a teraz są tylko częściowo wykorzystywane. Pełny stojak wraz z podłączeniem do internetu działającym z ogromną prędkością kosztuje nie więcej niż 3000 zł na miesiąc.

Wysoka dostępność obejmuje również *wyrównywanie obciążenia* (zarówno sprzętowe, jak i programowe), którego zadaniem jest wyrównywanie obciążenia zasobów w celu redukowania zatorów prowadzących w końcu do awarii usługi (system może pracować, choć nie będzie odpowiadał na żądania). Wyrównywanie obciążenia zakłada skalowalność systemu. Oczywiście, jeżeli system lub oprogramowanie nie daje się skalować, to bardzo trudno jest zapewnić wyrównywanie obciążenia

między wieloma hostami. Microsoft dostarcza rozwiązania do *klasteryzacji z wyrównywaniem obciążenia sieci* (NLB), *wyrównywaniem obciążenia komponentów* (CLB) oraz *klastry działające pomimo awarii*, jednak jeżeli samo oprogramowanie dostarczające usługę nie jest skalowalne ani zależne od zastosowania klastra NLB, to jego stosowanie wraz z pozostałymi usługami klastrowania dostępnymi na platformie Windows Server ma niewielki sens.

Wysoka dostępność korzysta również z nadmiarowości sprzętu. Nawet jeżeli system nie obsługuje pracy pomimo awarii, nadal można osiągnąć wysoką dostępność przez zastosowanie nadmiarowości sprzętowej. Najbardziej znanym zastosowaniem nadmiarowości jest nadmiarowość dysków; istnieje kilka technologii zapewniających mirroring, striping oraz kombinacje obu tych technik, dzięki czemu można zapobiec wyłączeniom spowodowanym awarią dysku lub wyeliminować je. Nadmiarowość sprzętową omawiamy bardziej szczegółowo w rozdziale 2., „Wybór sprzętu o wysokiej wydajności", oraz 3., „Pamięci masowe dla systemów wysoko dostępnych".

Jeżeli jesteśmy przy pamięciach masowych, trzeba pamiętać, że nie ma nic lepszego dla systemów o wysokiej dostępności jak technologie konsolidacji pamięci masowych. Zarówno systemy NAS, jak i SAN grają niezwykle ważną rolę w świecie wysokiej dostępności. Nie jest to tylko skonsolidowana centrala danych dostępna przy projektowaniu klastra działającego pomimo awarii. Cała technologia — przepustowość, łatwość serwisowania, zarządzania i tak dalej — pełni ważną rolę w spełnieniu podstawowego wymagania dostępności. Z tego powodu pamięciom masowym poświęciliśmy cały rozdział 3., „Pamięci masowe dla systemów o wysokiej dostępności".

Pamięci, procesory, wejście-wyjście, magistrale i tym podobne elementy również odgrywają krytyczną rolę, szczególnie w przypadku zapewnienia skalowania, wieloprocesowości, wielowątkowości i tak dalej. Komponenty te w systemach o wysokiej dostępności wymagają monitorowania dostępności, monitorowania wydajności oraz analiz, dzięki czemu można spełnić wymaganie wysokiej dostępności. Z tego powodu w rozdziale 13. przedstawiamy narzędzia do monitorowania działania, takie jak konsola *Performance* oraz *Microsoft Operations Manager* firmy Microsoft.

Na koniec przedstawimy czynniki pozasystemowe, które mogą wpłynąć na dostępność: ludzką zdolność do obsługi i utrzymania systemów, oraz ich właściwe projektowanie i implementowanie. W rozdziałach 5. i 6. wprowadzamy temat projektowania i implementacji. Kolejne rozdziały są ukierunkowane na projektowanie a zawarte w nich informacje są przedstawiane w postaci przykładów.

Książka na temat wysokiej dostępności i wydajności nie byłaby wyczerpująca bez omówienia zagadnień bezpieczeństwa. Aby zapewnić założony poziom obsługi, należy stale się upewniać, że system nie jest przedmiotem ataku. Ataki przyjmują

różne formy — wirusów, robaków, koni trojańskich, niskopoziomowych włamań interaktywnych i tak dalej. Choć tematem tej książki nie jest bezpieczeństwo komputerów, temat sam w sobie niezwykle obszerny, przedstawimy potrzeby zarządzania aspektami bezpieczeństwa z punktu widzenia poziomu usługi, dostępności oraz wydajności i czynnika ludzkiego.

Skalowanie dostępności w poziomie i Windows Server 2003

Mówiąc o skalowalności, mamy na myśli sposób rozbudowy usługi lub aplikacji w celu spełnienia rosnących wymagań co do wydajności zapisanych w SLA. Jeżeli w tej książce pojawia się pojęcie skalowalności systemów komputerowych, mamy na myśli możliwość dodania komputerów do istniejącego klastra, dzięki czemu obciążenie pozostałych komputerów może być skierowane do dołączonych, a w efekcie można spełnić SLA i zapewnić wymaganą wydajność. Przedstawimy tu dwie opcje skalowania: skalowanie pionowe i poziome.

Platforma Windows Server 2003 korzysta ze skalowania poziomego, ponieważ systemy korzystające z procesorów Intel najlepiej nadają się do wykorzystania w architekturze takiego właśnie skalowania. Skalowanie w poziomie wykorzystuje klasteryzację, czyli konfigurację, w której systemy mogą albo pracować równolegle jako systemy przetwarzania rozproszonego w celu obsłużenia dodatkowego obciążenia, albo jako klastry zapewniające pracę pomimo awarii oraz udostępniające usługi nadmiarowe.

Skalowanie poziome jako forma przetwarzania równoległego, wyrównywania obciążenia lub obu tych mechanizmów jednocześnie wymaga zastosowania systemów i oprogramowania, które pracuje zgodnie z zasadą „dziel i zwyciężaj", gdzie dane aplikacji i kod przetwarzający są rozproszone po wielu węzłach. Każdy węzeł może pracować na własnej części całego zbioru danych lub korzystać z jednego, wspólnego zbioru danych. W drugim przypadku główny proces integracji danych uruchamia system procedur transakcyjnych, przetwarzania rozproszonego i replikacji danych, dzięki czemu można zachować spójność danych.

Typowym przykładem jest popularna witryna typu B2C. Do rozproszenia połączeń i obciążenia uruchamianych jest wiele serwerów WWW. Połączenia są realizowane za pośrednictwem komponentów warstwy pośredniej, a przy przetwarzaniu wykorzystywane są bazy danych zainstalowane na jednym lub większej liczbie serwerów SQL. W celu zachowania spójności danych wykorzystywane są techniki replikacji i izolacji. Spójność danych oraz techniki takie jak replikacja i przesyłanie dzienników omówimy dokładniej w rozdziale 9., „Zastosowania SQL Server w rozwiązaniach o wysokiej dostępności i wydajności".

Jak już wcześniej wspomnieliśmy, aby aplikacje mogły być skalowane, muszą być do tego przygotowane. Można po prostu skorzystać z wyrównywania obciążenia oraz ruchu w sieci zapewnianego przez Windows Server 2003, o ile aplikacje obsługują NLB, albo podzielić działanie procesu na wiele serwerów, co jest znane jako federacja. Można skorzystać z wielu rozwiązań — a Microsoft dostarcza kilku interfejsów programowych korzystających ze wspólnych bibliotek, szczególnie tych znajdujących się w .NET Framework — które obsługują działanie programu rozproszonego pomiędzy kilka systemów o wysokiej wydajności (HPC).

Klasteryzacja

Jak się już niebawem okaże, termin *klasteryzacja* może odnosić się do więcej niż jednej techniki zapewnienia dostępności. Właśnie omówione klastrowanie zapewniające skalowanie w poziomie korzysta z usług dostarczanych przez Windows Server 2003 do zapewnienia wyrównywania obciążenia i rozpraszania przetwarzania pomiędzy poszczególne węzły. Z drugiej strony klastrowanie w celu zapewnienia pracy pomimo awarii to technika zapewniania dostępności. Jest to odmiana nadmiarowości.

Klaster *aktywno-pasyny* składa się z pary węzłów, w której węzeł pasywny nie wykonuje żadnej pracy, natomiast węzeł aktywny realizuje wszystkie zadania. Jeżeli awarii ulegnie węzeł podstawowy, działanie aplikacji jest przekazywane do węzła pasywnego, który jest w tym momencie aktywowany. Zauważalna jest niewielka przerwa w pracy usługi, jednak aplikacja jest odtwarzana z pomijalną przerwą w działaniu i przetwarzanie jest kontynuowane na nowym węźle podstawowym. Uszkodzony węzeł jest następnie odtwarzany i albo jest dołączany jako węzeł pasywny, albo aplikacja jest przenoszona na poprzedni system i odtwarzany jest stan aktywno-pasywny sprzed awarii.

W Windows Server 2003 Enterprise Edition oraz Datacenter Edition można mieć więcej niż jeden aktywny węzeł w klastrze. Jest również możliwe uruchomienie więcej niż dwóch aktywnych węzłów w klastrze i dodanie jednego pasywnego. Konfiguracja aktywno-aktywno-pasywna (n+1) jest omówiona dokładniej w rozdziale 5., „Przygotowywanie platformy dla sieci o wysokiej wydajności".

UWAGA: Nie można utworzyć klastra działającego pomimo awarii z wykorzystaniem Windows Server 2003 Standard Edition.

Pionowe skalowanie dostępności

Czasami programiści tworzą aplikacje, których nie da się łatwo skalować w poziomie, a wymagania tych aplikacji bardziej przystają do skalowania pionowego. Skalowanie w pionie polega na dodawaniu kolejnych komponentów sprzętowych,

najczęściej procesorów lub pamięci, w celu pełnego wykorzystania technik przetwarzania równoległego i wielozadaniowego.

Skalowanie pionowe wykorzystuje możliwości nowoczesnych procesorów, takich jak wielowątkowość, hyperthreading, blokady, semafory i inne funkcje atomowe. Systemy skalowane w pionie są zwykle łatwiejsze do zarządzania, ponieważ zwykle trzeba radzić sobie z jednym stanem systemu operacyjnego, jednym repozytorium danych oraz przestrzenią przetwarzania rozproszonego. Jednak niektóre aplikacje z powodzeniem korzystają z połączenia technik skalowania poziomego i pionowego.

Systemami tymi są wysoce dostępne systemy przetwarzania transakcji, w których złożone aplikacje wielowątkowe są skalowane poziomo na kilka klastrów obsługujących pracę pomimo awarii lub wyrównywanie obciążenia. Więcej na ten temat napiszemy w rozdziałach 5., 9., 11. i 12.

Skalowanie pionowe czy poziome?

Skalując system w pionie, dodajemy zasoby do istniejącego systemu komputerowego. Gdy czas odpowiedzi serwera zaczyna się powiększać z powodu zbyt dużego obciążenia, większej ilości żądań do bazy danych lub przepływu poczty elektronicznej, najczęściej stosowanym sposobem natychmiastowego rozwiązania problemów z wydajnością jest dodanie większego, szybszego (i droższego) sprzętu.

Obecnie producenci sprzętu podwajają wydajność urządzeń co 18 do 24 miesięcy. Jeśli można błyskawicznie i niemal w sposób nieograniczony powiększać moc systemu, podejście to wydaje się bardzo rozsądne, ale szybko natykamy się na pewien problem. Ze stałą aktualizacją sprzętu wiąże się wiele problemów.

Pierwszym i najważniejszym są ograniczenia urządzeń. Zakładając, że wydajność sprzętu podwaja się co dwa lata, i że mamy pieniądze na unowocześnianie sprzętu z taką częstotliwością, co możemy zrobić, gdy po dwunastu miesiącach moc nowego systemu stanie się niewystarczająca? Czy będziemy zmagać się z niewystarczającą wydajnością przez następny rok? Nie jest to dobre rozwiązanie, szczególnie po kosztownej rozbudowie.

Nawet gdy producenci będą produkowali komputery ośmioprocesorowe z układami pamięci 16 GB i macierzą SAN wykorzystującą połączenia światłowodowe, problem skalowalności nadal będzie istniał. Wcześniej czy później będziemy zmuszeni poczekać, aż nasz dostawca wypuści następną generację swojego supersprzętu, która zaspokoi nasze wymagania.

Sytuacja jednak może się jeszcze bardziej skomplikować. Gdy system osiągnie określony punkt, dalsze skalowanie w pionie staje się tak kosztowne, że wydane

pieniądze nie są warte osiąganych efektów. Nawet pomijając problemy ze zgodnością sprzętu, prawdopodobnie napotkamy problemy z oprogramowaniem po przekroczeniu pewnego punktu krytycznego.

Na przykład, zwróćmy uwagę na opcję `/3GB /PAE` z pliku *boot.ini* serwera Windows 2000. Mamy tu problem z prawidłowym wykorzystaniem przez OS dużych ilości pamięci (4GB i więcej). Niektóre systemy oprogramowania, na przykład serwery baz danych, mają wewnętrzne algorytmy obsługi transakcji, blokowania, wielodostępności i problemów z architekturą trójwarstwową. Architektura tych systemów ma ograniczenia efektywności. Ograniczenia ta mogą uniemożliwić dalsze skalowanie w pionie. Jest to podobne do krzywej dzwonowej: prędzej czy później, na szczycie krzywej będziemy potrzebowali bardzo drogich aktualizacji sprzętu, aby uzyskać niewielką poprawę wydajności. Skalowanie w poziomie oznacza zastosowanie większej ilości sprzętu.

Udostępnianie wszystkiego a nieudostępnianie niczego

Skupmy się teraz na zastosowaniu większej ilości sprzętu, a nie tylko sprzętu większego lub lepszego. Skalowanie poziome może być efektywnym rozwiązaniem problemów napotykanych w scenariuszu skalowania pionowego. Projektujemy system nie tak, aby *udostępniać wszystko*, ale raczej tak, by *nic nie udostępniać*.

W istocie, architektura współdzielenia niczego oznacza, że każdy system komputerowy w klastrze działa niezależnie. Każdy system w klastrze posiada osobne zasoby (CPU, pamięć, dyski). Aby rozwiązać problemy pojemności przez skalowanie poziome, dodajemy więcej sprzętu do puli — a nie do pojedynczej jednostki.

Skalowanie poziome pozwala rozwiązać problem czynnika kosztu związany ze skalowaniem pionowym, ponieważ dodanie kilku mniejszych systemów jest zwykle tańsze niż unowocześnianie dużego systemu klasy mainframe lub koszt oraz problemy związane z całkowitym przeniesieniem na nową platformę. W przypadku skalowania poziomego wielkość i szybkość działania pojedynczego systemu nie ogranicza całkowitej wydajności. Architektura współdzielenia niczego pozwala zlikwidować problem wąskich gardeł programowych przez dostarczenie architektury obsługującej wiele mechanizmów współbieżności. Ponieważ obciążenie jest dzielone na wiele serwerów, całkowita pojemność programowa i przepustowość zwiększa się.

Pomimo tego, że skalowanie poziome zapewnia rozwiązanie *integralnych* ograniczeń architektury skalowania pionowego, metoda jest związana z innymi problemami. Skalowanie poziome wymaga dodatkowych czynności administracyjnych, dogłębnej wiedzy i oczywiście pieniędzy. Pułapki mogą być potencjalnie

tak duże, jak uzyskiwany przyrost wydajności. Nawet pomimo tego skalowanie w poziomie może być doskonałym rozwiązaniem dla serwerów baz danych, które osiągnęły granice skalowalności sprzętu.

Mamy wiele do przemyślenia, szczególnie gdy klient ma tylko jedno wymaganie: „zapewnić, aby system był zawsze włączony".

Wysoko wydajne przetwarzanie danych

Wysoko wydajne przetwarzanie danych (HPC) nie może być mylone z wysoką dostępnością, ale jest to integralne pojęcie, bez którego nie będziemy w stanie osiągnąć odpowiedniego poziomu obsługi.

Można budować autonomiczne komputery o wysokiej wydajności albo komputery będące częścią farm serwerów czy federacji komputerów. Jeżeli jednak projekt nie będzie zapewniał wysokiej dostępności, możemy nie być w stanie zapewnić poziomu usługi wymaganego przez firmę lub właściciela procesu. I odwrotnie, projekt o wysokiej dostępności lub jego implementacja nie powiedzie się, jeżeli zastosowane komponenty nie będą pozwalały na osiągnięcie wysokiej wydajności środowiska komputerowego.

Mając na celu wysoką dostępność, można zbudować system komputerowy zapewniający poziom usługi przy wysokiej wydajności. Nie zawsze jednak zachodzi taki przypadek i nie zawsze jest on możliwy do zrealizowania.

Na przykład, sterta tanich dysków SCCI obracających się z prędkością najwyżej 7200 obrotów na minutę w konfiguracji RAID 1 lub RAID 5 oczywiście zapewni wysoki poziom dostępności, w przeciwieństwie do jednego dysku, którego awaria spowoduje niedotrzymanie warunków umowy. Czy można stwierdzić, że są to komponenty HPC? Na pewno nie. Implementacja SAN składająca się z kosza dysków o prędkości obrotowej 15 000 obrotów na minutę w konfiguracji RAID-5, przesyłających dane za pomocą włókna szklanego zamiast SCSI, zapewnia zarówno wysoką dostępność, jak i wysoką wydajność. Oczywiście to, co dla jednej firmy jest HPC, dla innej może być superkomputerem. Macierze SAN, podobnie jak inne technologie, mają również malejące punkty powrotu, które przedstawimy w rozdziale 3.

Potrzeba przetwarzania wysoko wydajnego

Skupmy się jeszcze chwilę na temacie HPC. Technologia napędza dzisiejsze firmy, a w większości napędza też społeczeństwo. Większość firm po prostu zawiesiłaby działanie, jeżeli na kilka dni odcięto by im dostęp do ich technologii. Aby pozostać

konkurencyjnym, nie tylko trzeba być dostępnym przez cały czas, ale przepustowość systemów musi być możliwie duża.

HPC jest krytyczne dla wszystkich firm. Jeżeli usługa działa powoli i mała grupa ludzi o minutę dłużej będzie wysyłać dane, to po podsumowaniu roku okaże się, że straciliśmy tydzień pracy. Ta strata na pewno wpłynie na wyniki finansowe firmy.

Serwery baz danych udostępniające witryny WWW lub wprowadzanie danych i obliczenia muszą działać z największą możliwą szybkością. Serwery WWW muszą być w stanie obsłużyć tysiące połączeń, a nie tylko kilka. Serwery plików i drukarek nie mogą zatrzymywać się na przeciążeniu procesora lub pamięci w przypadku drukowania ważnej notatki dla wszystkich pracowników.

Przetwarzanie wysoko wydajne dla każdego

Informatyka jest jedną z dziedzin nauki. Tak jak w każdej dziedzinie nauki, mamy tu prawa i wzory opisujące poszczególne hipotezy. W przetwarzaniu HPC możemy znaleźć twierdzenie, na którym mogą opierać się wszystkie firmy: koszt nowej technologii pozostaje znacznie wyższy, dopóki technologia ta jest postrzegana jako nowa. W momencie, gdy zostanie wymyślone coś lepszego, koszt starej już technologii znacznie spada. Bez wchodzenia w szczegóły *wskaźnika zwrotu z inwestycji* (ROI) i innych czynników, które usprawiedliwiają korzystanie z najnowszych technologii, dla większości potrzeb przedsiębiorstwa lepiej i taniej jest korzystać z technologii, gdy przestanie być traktowana jako nowość.

Przecież to, że rano coś lepszego pojawiło się w wiadomościach, nie oznacza, że stara technologia przestała być użyteczna dla firmy. Z drugiej strony, biorąc pod uwagę obecną szybkość rozwoju technicznego, urządzenia, które wyszły z mody, będą nadal użyteczne co najmniej przez kolejny rok. Większość firm może skorzystać z nowych technologii ze znacznym opóźnieniem.

Powodem tego są kłopoty ze zdobywaniem wiedzy. Ludzie, którzy najprawdopodobniej mogliby skorzystać z nowych technologii, nie zdobędą odpowiedniej wiedzy do momentu wykonania tzw. „transferu wiedzy". Dla wielu technologii proces ten może trwać kilka lat.

W jednym z ostatnich szokujących raportów firma Gartner poinformowała, że ponad połowa amerykańskich korporacji nadal korzysta z Windows 98. Firmy te mają takie opóźnienie w przyjęciu nowszych i bardziej skomplikowanych systemów operacyjnych, takich jak Windows XP i Windows Server 2003, że Microsoft musiał przedłużyć wsparcie dla swoich przestarzałych systemów o kolejne siedem lat. Podobne statystyki można znaleźć dla innych systemów operacyjnych dla serwerów. Tysiące firm korzystają nadal z Windows NT i pomimo tego, że Windows Server 2003 jest dostępny od początku roku 2003, większość z nich nie zmieni platformy co najmniej do roku 2005.

Przeanalizujmy więc następujące twierdzenie: nowy system lub technologia staje się dostępna dla firm w momencie, gdy dostępna jest wystarczająca ilość wiedzy i informacji, aby można było przeszkolić architektów, projektantów i operatorów. W momencie, gdy zostanie osiągnięta odpowiednia penetracja rynku, najprawdopodobniej powstanie nowa generacja tej technologii i w tym momencie cena tak zwanej przestarzałej technologii znacznie spadnie.

Na przykład, kilka lat temu cena włókna SCSI SAN była poza zasięgiem większości małych firm, które mogły sobie pozwolić tylko na pamięci masowe dołączane do sieci (NAS). Obecnie, przy zastosowaniu dysków SAN z włóknami światłowodowymi, stare miedziane włókna SCSI dla SAN znacznie potaniały. Jak się okaże w rozdziale 3., można zbudować podstawowy SAN dla małej firmy za mniej niż połowę kwoty, którą Twój zespół techniczny wydaje na obiad.

Komponenty HPC stały się dostępne cenowo w czasie potrzebnym architektowi do zaprojektowania systemu. Serwery stale tanieją, ponieważ coraz nowsze modele pojawiają się na witrynach producentów. Najtańsza linia serwerów była całkowitą nowością jeszcze sześć miesięcy temu i najprawdopodobniej nadaje się doskonale dla naszych celów, przy cenie równej jednej piątej kosztu najnowszego serwera.

Dzięki temu przetwarzanie HPC jest dostępne dla wszystkich. Po zaprojektowaniu systemu i określeniu potrzeb można łatwo zmieścić się w budżecie i kupić komponenty spełniające nasze wymagania.

Superkomputer w każdej szafie

Dzisiejsze systemy operacyjne nadal rozwijają zdolności jak najlepszego wykorzystania możliwości sprzętu, na którym zostały zainstalowane. W książce tej przedstawiamy jeden z tych systemów operacyjnych, który praktycznie każdej firmie daje możliwości obliczeniowe superkomputerów, przy zachowaniu ułamka wydatków, jakie byłyby przewidziane do tych celów jeszcze kilka lat temu.

Za nie więcej niż kilka milionów złotych możliwe jest zastąpienie antycznych systemów Novell NetWare, GroupWise następującą konfiguracją: kilkaset serwerów zainstalowanych w ponad 100 oddziałach z wysoko wydajną i wysoko dostępną implementacją Active Directory, która pozwala udostępniać pliki i drukarki przy wykorzystaniu kilku macierzy SAN, obsługujących 5 000 użytkowników Exchange i Outlook.

Przed rokiem 2000 taka wymiana systemu sieciowego byłaby uznana za zbyt drogą i kłopotliwą do przeprowadzenia. Tym, co spowodowało, że operacja jest obecnie łatwiejsza i tańsza, stało się użycie systemu operacyjnego Windows Server 2003 oraz Active Directory, omówionych w kolejnych rozdziałach.

Ten system operacyjny jest nie tylko tańszy, ale również działa taki sposób, że systemy z niego korzystające według większości naukowych definicji można traktować jak superkomputery.

Przetwarzanie i pamięć

Przetwarzanie wysoko wydajne zależy od kilku komponentów, ale zazwyczaj na początku zwraca się uwagę na procesor i pamięć. Procesory stają się tańsze, bardziej dostępne i wydajniejsze. Jednym z czynników, który przestał być problemem dla większości budżetów, jest rozmiar i ilość procesorów zamontowanych w serwerze. Większość firm obecnie kupuje serwery z obsadzonymi wszystkimi gniazdami serwerów dostępnymi na płycie głównej.

Prawo Moore'a zakłada, że „ilość tranzystorów w pojedynczym układzie podwaja się co 18 miesięcy". Do niedawna formuła ta sprawdzała się z niezwykłą dokładnością. Obecnie ilość tranzystorów podwaja się szybciej niż przewidział to Gordon Moore, były członek zarządu firmy Intel (największy producent mikroprocesorów i długoterminowy partner firmy Microsoft). Jak opiszemy w rozdziale 2., w przyszłości procesory staną się mniej zależne od tradycyjnych ograniczeń metalu nadprzewodzącego.

W roku 1998 komputer z procesorem 386 16 MHz z 1 MB pamięci RAM i 40 MB dyskiem twardym, kosztujący ponad 20 000 zł, był poza zasięgiem większości firm. Obecnie komputer z procesorem 1,5 GHz z ponad 256 MB pamięci RAM i dyskiem twardym 40 GB można bez problemu kupić za mniej niż 2 000 zł.

Rozwój pamięci również przebiega w niezwykłym tempie, dzięki czemu ilość pamięci zamontowanej w systemie przestaje być problemem. Jeżeli mamy potrzeby klasy HPC i SLA do spełnienia, zamawiając sprzęt dla nowego klastra SQL Server lub Exchange, najczęściej polecamy sprzedawcy, aby dodał pamięci „do pełna".

System operacyjny Windows Server 2003 zapowiada również erę systemów 64 bitowych, która spowoduje powstanie wielu firm, wielkich i małych, tworzących nowe aplikacje dla superkomputerów. Autostrada przetwarzania 64-bitowego jest niezwykle obiecująca. W rozdziale 2. przyjrzymy się, jak pojemność pamięci masowych wpływa na nasze potrzeby programowe.

Komponenty wysoko wydajne

Oprócz pamięci i procesorów, również wiele innych komponentów jest ważnych dla przetwarzania wysoko wydajnego. Żaden system nie obejdzie się bez pamięci masowej. Żaden system o wysokiej wydajności lub wysokiej dostępności nie

może też działać bez współdzielonego i nadmiarowego systemu pamięci masowej (RAID 1, 5, 10 i tak dalej). Pamięć masowa i kilka innych krytycznych komponentów jest opisanych w kilku kolejnych rozdziałach.

Pozostałe komponenty składające się na systemy HPC-HA to zasilacze (PSU), dyski i kontrolery, przełączniki i połączenia przełączników, okablowanie sieciowe (szczególnie włókna światłowodowe), karty sieciowe, adaptery magistrali i tak dalej.

Microsoft i Cornell Theory Center

Wzrost zainteresowania i potrzeby przetwarzania o wysokiej wydajności i wysokiej dostępności doprowadziły do powstania kilku standardów i organizacji zainteresowanych budową systemów o dużej wydajności. Jedną z takich instytucji promujących HPC na platformie Windows Server jest Cornell Theory Center (CTC).

Cornell Theory Center jest centrum badawczym zlokalizowanym w kampusie Cornell University w Ithaca. CTC posiada powiązania z ponad 500 badaczami w Cornell, zajmujących się różnymi dyscyplinami naukowymi i matematycznymi. Sieć ta rozciąga się na cały świat, poprzez badaczy, partnerów i zewnętrzne powiązania. W CTC opracowano doskonałe rozwiązania z zakresu przetwarzania wysoko wydajnego oraz informatyki z następujących dziedzin:

- finanse komputerowe,
- biologia i genomika komputerowa,
- komputerowa teoria materiałowa.

Według CTC, są oni „w pierwszym szeregu badań nad obliczeniami wysoko wydajnymi od wielu lat". Podobnie jak wiele innych tego typu organizacji, CTC we większości swoich prac wykorzystuje drogie, specjalizowane implementacje systemu UNIX. W związku z ograniczeniami budżetowymi, zwiększoną zależnością naukowców od infrastruktury wysoko wydajnej oraz zwiększającą się dostępnością komputerów i komponentów sieciowych, w roku 1999 w CTC opracowano unikatową strategię — wykorzystania do obliczeń wysoko wydajnych systemów Microsoft Windows.

„Od tego momentu CTC z sukcesem tworzy światowej klasy centrum superkomputerowe korzystające z Windows. Największym systemem CTC jest 256-procesorowy klaster Velocity II, który jest jednym z 10 najszybszych superkomputerów na świecie".

Choć w kolejnych rozdziałach skupiamy się w większości na dostępności, element HPC jest zawsze związany z każdą klasą sprzętu i oprogramowania, która spełnia nasze wymagania poziomu usługi.

Podsumowanie

W tym rozdziale przedstawiliśmy niektóre ważne pojęcia, przygotowując w ten sposób grunt pod kolejne rozdziały. Zdefiniowaliśmy kilka terminów: dostępność, wydajność, niezawodność, nadmiarowość, awaria, naprawa, czas działania i czas wyłączenia. Omówiliśmy również poziom obsługi oraz umowy SLA. Na koniec rzuciliśmy nieco światła na zagadnienia przetwarzania wysoko wydajnego i superkomputerów.

Oprócz przygotowania gruntu dla kolejnych rozdziałów, rozdział ten przedstawia kilka idei. Obecnie Windows Server 2003 nie jest zbyt trudny ani drogi, więc można go stosować przy budowaniu wysoko dostępnych lub wysoko wydajnych systemów komputerowych. Po umieszczeniu takiego systemu superkomputerowego w lokalizacji zabezpieczonej przed klęskami żywiołowymi, będziemy w stanie sprostać najbardziej wymagającym umowom SLA. Jeżeli w SLA prawidłowo zdefiniujemy czas wyłączenia i będziemy pilnować, aby wyłączenia nie zdarzały się w oknie działania, dla którego jest podpisana umowa, najprawdopodobniej będziemy w stanie osiągnąć cztery dziewiątki bez potrzeby obrabowania banku.