Przemysław PRZYBYLAK[1]

WIRTUALNA INFRASTRUKTURA – NOWE PODEJŚCIE DO SYSTEMÓW

Streszczenie

W artykule opisane zostały założenia stojące u podstaw technologii wirtualizacji w szeroko rozumianym obszarze usług ICT (Information and Communication Technologies). Opracowanie przedstawia przegląd technologii stosowanych dla celów wirtualizacji w głównych obszarach ICT, tj. sieciach, pamięciach masowych, serwerach oraz stacjach roboczych. Przedstawiono w nim także tezy dotyczące wpływu technologii wirtualizacji w obszarze ICT na biznes, który uwidacznia się przez obniżenie całkowitych kosztów posiadania (Total Cost of Ownership) i szybszego zwrotu z inwestycji.

Abstract

The article describes standing assumptions underlying virtualization technology in the wider field of ICT (Information and Communication Technologies). The study provides an overview of technologies used for the purpose of virtualization in key areas of ICT, such as networks, storage, servers and workstations. It presents the argument about the impact of virtualization technology in the field of ICT on business, which is reflected by the reduction in the TCO (Total Cost of Ownership) and a faster return on investment.

1. WSTĘP

Dzisiejszy sprzęt komputerowy wykorzystujący architekturę procesorów x86 został zaprojektowany do uruchamiania jednego systemu operacyjnego z pojedynczą aplikacją. Efektem takiej konstrukcji jest fakt, że moc obliczeniowa i wydajność większości obecnie produkowanych systemów komputerowych nie jest w pełni wykorzystana.

Wykorzystanie wirtualizacji pozwala uruchamiać wiele logicznych systemów (zwanych wirtualnymi maszynami) na jednym urządzeniu w ten sposób, że działając w różnych, często odseparowanych logicznie środowiskach, współdzielą jego zasoby sprzętowe. Maszyny wirtualne uruchomione na jednym fizycznym urządzeniu mogą

[1] Mgr inż. Przemysław Przybylak jest wykładowcą Warszawskiej Wyższej Szkoły Informatyki.

funkcjonować pod kontrolą różnych systemów operacyjnych i świadczyć wiele usług, dzięki aplikacjom, które mogą być uruchomione na każdej z nich.

Warunki makroekonomiczne oraz kryzys gospodarczy, który silnie uwidocznił się w ostatnich latach spowodowały, że funkcjonujące w takim otoczeniu organizacje zmuszone zostały do ograniczania wydatków na inwestycje. Trend ten nie ominął również wydatków związanych z inwestycjami w szerokorozumiane usługi ICT (ang. *Information and Communication Technologies*). Ograniczenia wydatków na nowe usługi, jak również redukcje i obniżanie kosztów już eksploatowanych dotknęły nie tylko działy IT funkcjonujące w ramach struktur organizacji, ale również firmy świadczące usługi w sektorze ICT.

W tej sytuacji wdrażanie nowych usług z wykorzystaniem środowisk wirtualnych oraz migracja już funkcjonujących do tego typu rozwiązań są i z pewnością będą w najbliższych latach poważnie rozważane przez profesjonalnie zarządzane organizacje. Potencjalne oszczędności, na jakie dzięki wirtualizacji może liczyć organizacja każdej wielkości, w zakresie użytkowania i wdrażania usług ICT wspierających biznes, są z pewnością jednym z głównych czynników przemawiających za gwałtownym wzrostem zainteresowania tą technologią. Organizacje podejmując decyzję o wdrożeniu usług z wykorzystaniem technologii wirtualizacji, argumentują to obniżeniem całkowitych kosztów posiadania (ang. *TCO – Total Cost of Ownership*), szybkim zwrotem z inwestycji, konsolidacją zasobów sprzętowych i programowych oraz elastycznością zarządzania usługami.

2. HISTORIA WIRTUALIZACJI

> *Dla każdego standardowego komputera trzeciej generacji może zostać skonstruowana wirtualna maszyna, jeśli zbiór instrukcji wrażliwych jest podzbiorem zbioru instrukcji uprzywilejowanych.*[2]

W 1974 roku na łamach flagowego miesięcznika organizacji ACM (*Association for Computing Machinery*), zrzeszającej profesjonalistów informatyki ukazał się artykuł pt.: *Formal Requirements for Virtualizable Third Generation Architectures.* Autorzy Gerald J. Popek (*University of California, Los Angeles*) i Robert P. Goldberg (*Harvard University, Cambridge*) określili w nim kryteria jakie muszą zostać spełnione aby możliwe było zbudowanie maszyny wirtualnej na zadanej architekturze komputera trzeciej generacji oraz kryteria właściwego funkcjonowania takiej maszyny wirtualnej.

[2] I Twierdzenie Wirtualizacji G. J. Popka i R. P. Goldberga sformułowane w artykule *Formal Requirements for Virtualizable Third Generation Architectures*; Communications of the ACM, July 1974, Volume 17, Number 7

Naukowcy zdefiniowali dla architektury komputera trzeciej generacji poniższe zbiory instrukcji, które powinny być dostępne w jej systemie operacyjnym:

- instrukcje uprzywilejowane, których efektem jest przerwanie lub wywołanie systemowe w trybie użytkownika lub ich brak w trybie jądra;
- instrukcje wrażliwe ze względu na kontrolę, które w trakcie wykonywania mogą zmienić konfigurację zasobów systemowych;
- instrukcje wrażliwe ze względu na wykonanie, których działanie jest zależne od konfiguracji systemu.

Zgodnie z ich przemyśleniami, aby maszyna wirtualna była skuteczna dla tak zdefiniowanej architektury musi spełniać 3 warunki:

- równoważności/wierności odwzorowania (ang. *Equivalence/Fidelity*) oznaczający, że program działający na wirtualnej maszynie musi zachowywać się w dokładnie taki sam sposób jak na rzeczywistym sprzęcie;
- kontroli zasobów/bezpieczeństwa (ang. *Resource controls/Safety*), oznaczający, że maszyna wirtualna musi w pełni kontrolować wszystkie zasoby, które są wirtualizowane;
- skuteczności/wydajności (ang. *Efficiency / Performance*), oznaczający, że większa część instrukcji musi być wykonywana bez udziału maszyny wirtualnej.

Uwzględniając zdefiniowane zbiory instrukcji i ograniczenia przedstawione w postaci warunków sformułowali zacytowane na wstępie twierdzenie, które stało się podstawą rozwoju technologii związanych z wirtualizacją.

W ciągu ostatnich 35 lat, opracowano wiele rodzajów wirtualizacji. Są wśród nich:

- emulacja, umożliwiająca wirtualizację innej architektury komputera na architekturze komputera hosta/gospodarza dzięki zastosowaniu tłumaczenia instrukcji;
- wirtualizacja natywna, bardziej wydajna niż tradycyjna emulacja, zapewniająca emulowanie tylko niektórych instrukcji, ograniczona do tej samej architektury co architektura hosta/gospodarza;
- parawirtualizacja, w której system hosta/gospodarza został tak zmodyfikowany aby działał na niższym poziomie uprzywilejowania, a odpowiedni, samodzielny program działający na poziomie sprzętu przydziela zasoby dla uruchomionych systemów wirtualnych.

Jednak dopiero odpowiednie zaprojektowanie architektury popularnych układów mikroprocesorowych serii x86 i opracowanie mechanizmów pełnej wirtualizacji zapoczątkowało lawinowy rozwój tej technologii w świecie ICT.

W chwili obecnej autorzy publikacji przedstawiających zagadnienia związane z wirtualizacją odnoszą się w znakomitej większości do technologii pełnej wirtualizacji.

Wirtualizację w środowisku ICT można ogólnie zdefiniować jako izolowanie jednego zasobu obliczeniowego od innych. Izolacja ta poprzez oddzielenie poszczególnych warstw logicznych pozwala na zwiększenie elastyczności oraz uproszczenie zarządzania zmianami.

Wirtualizację można zastosować w odniesieniu do serwerów, pamięci masowych, sieci, komputerów biurowych oraz aplikacji.

3. WIRTUALIZACJA SIECI

Wirtualizacja sieci przybiera dwie podstawowe formy. Są to rozwiązania typu „jeden do wielu", w ramach których na podkładzie jednej sieci fizycznej przenoszone są usługi wielu sieci wirtualnych, oraz rozwiązania typu „wiele do jednego", w ramach których jedno wirtualne urządzenie sieciowe konsoliduje wiele fizycznych urządzeń sieci.

Wirtualizacja sieci zwiększa elastyczność wykorzystania zasobów, poprawia wykorzystanie pojemności oraz umożliwia segmentację sieci, a co za tym idzie, podnosi poziom bezpieczeństwa sieci.

Wśród technologii związanych z wirtualizacją sieci w modelu „jeden do wielu" można wyróżnić:

- VLAN (ang. *Virtual Local Area Network*), czyli wirtualizowanie sieci lokalnych na poziomie warstwy II modelu ISO/OSI polegające na logicznym odseparowaniu wydzielonych lokalnych sieci komputerowych w ramach większych, fizycznych sieci lokalnych, budowane z wykorzystaniem przełączników sieciowych, których konfiguracja umożliwia podział jednego fizycznego urządzenia na większą liczbę urządzeń logicznych poprzez separację ruchu pomiędzy grupami portów;
- VPN (ang. *Virtual Private Networks*), czyli wirtualizowanie sieci rozległych, poprzez budowanie połączeń tunelowych, zapewniających poufność, integralność i dostępność przesyłanej informacji w niezaufanym otoczeniu sieci publicznej (np.: Internetu), będące wirtualnym odpowiednikiem sieci rozległych, w których jako połączenia odległych lokalizacji wykorzystuje się tzw. linie dzierżawione;
- VRF (ang. *Virtual Routing and Forwarding*), czyli wirtualizowanie routerów i ich tablic routingu w ramach jednego fizycznego urządzenia, w którym wirtualne instancje routerów wraz z tablicami routingu są logicznie odseparowane,

co pozwala na wykorzystanie tego samego zakresu adresacji IP w ramach wielu wirtualnych sieci nie powodując tym samym konfliktów; technologia ta często łączona jest z inną – MPLS (*ang. Multiprotocol Label Switching*) i funkcjonuje w świecie ICT pod nazwą MPLS VPN Technology.

Z kolei przykładem rozwiązania odwrotnego – „jeden do wielu" jest technologia tzw. wirtualnego przełącznika. W jej ramach przyjmuje się założenie, że grupując dwa lub więcej przełączników sieciowych w jeden wirtualny system można zapewnić większą niezawodność, wydajność oraz odporność na uszkodzenia w przypadku utraty jednego z nich. Rozwiązania tego typu stosowane są często w centrach przetwarzania danych, w sytuacji kiedy dostępność usług udostępnianych przez sieć komputerową jest krytyczna z punktu widzenia funkcjonowania biznesu i niedopuszczalna jest choćby najmniejsza przerwa w ich świadczeniu.

4. WIRTUALIZACJA PAMIĘCI MASOWYCH

Technologie pozwalające na agregowanie danych przechowywanych w dowolnych urządzeniach pełniących rolę pamięci masowych i umieszczanie ich w jednej wysoce dostępnej, łatwo zarządzanej, z jednego centralnego miejsca puli zasobów można określić mianem wirtualizacji pamięci masowych. Innymi słowy tego typu wirtualizacja, to nic innego jak odpowiednie ułożenie fizycznych pamięci masowych z wykorzystaniem nieszablonowych metod abstrakcji. Wirtualizacja pamięci masowych ukrywa przed użytkownikiem całą złożoność pamięci masowych, udostępniając jedynie logiczny obraz danych, który jest odseparowany od skomplikowanej struktury urządzeń fizycznych. Oprogramowanie dysponujące mechanizmami wirtualizacji pamięci masowych kolekcjonuje dane przechowywane na różnych typach pamięci masowych, m.in.: urządzeniach SAN (ang. *Storage Area Network*), NAS (ang. *Network Area Storage*) i systemach pamięci masowych dołączonych bezpośrednio do serwerów budując tym samym wspólną pulę danych, którą w łatwy sposób można monitorować i zarządzać.

Pamięć masową można wirtualizować na trzy sposoby:

- przy wykorzystaniu serwerów lub hostów;
- przy wykorzystaniu urządzeń zainstalowanych w sieci;
- przy wykorzystaniu macierzy.

W zgodzie z powyższą klasyfikacją można wyróżnić:

- wirtualizację symetryczną, zwaną też wirtualizacją „in-band", realizowaną w ramach pasma – czyli na ścieżce danych, w której urządzenie wirtualizujące, usytuowane na ścieżce danych pomiędzy serwerem a pamięcią masową, ma za zadanie przesłanie danych oraz całej inteligencji związanej z wirtualizacją do dołączonych pamięci masowych;

– wirtualizację asymetryczną, zwaną wirtualizacją „out-of-band", realizowaną poza pasmem – czyli na zewnątrz ścieżki danych, w której zarządzaniem danymi wymienianymi pomiędzy pamięciami masowymi a serwerami, przełącznikami i innymi urządzeniami transportowymi sieci zajmuje się serwer lub macierz.

5. WIRTUALIZACJA SERWERÓW

Najbardziej popularnym i rozpowszechnionym przykładem wykorzystania technologii wirtualizacji jest wirtualizacja oprogramowania serwerów świadczących ściśle określone i zdefiniowane usługi. Realizuje się ją w ten sposób, że system operacyjny oraz aplikacja są połączone w celu utworzenia tzw. maszyny wirtualnej, która jest umieszczana na fizycznym serwerze obsługiwanym przez warstwę pośrednią w postaci tzw. hipervizora (czyli minimalnej warstwy oprogramowania zapewniającej podstawowy interfejs umożliwiający obsługę sprzętu) lub tradycyjnego systemu operacyjnego.

Przyjmując przeciętnie, że funkcjonujące aplikacje obciążają zasoby fizycznego serwera na poziomie 15% widać wprost, że utylizowany jest jedynie ułamek możliwości sprzętowych urządzenia, zaś większość zasobów – 85% pozostaje niewykorzystana. Łącząc uzupełniające się wzajemnie obciążenie pracą z w zakresie przetwarzania danych i wykorzystania pamięci z zapewnieniem izolacji i bezpieczeństwa charakterystycznego dla osobnych komputerów można w prosty sposób zredukować liczbę fizycznych serwerów niezbędnych do świadczenia tej samej liczby usług. Zwiększając poziom obciążenia pojedynczego fizycznego serwera do 75% można nawet pięciokrotnie zmniejszyć zapotrzebowanie na przestrzeń, sprzęt oraz energię elektryczną niezbędną do zasilenia i chłodzenia zespołów serwerów.

Dobrze zarządzane, zwirtualizowane środowisko serwerowe zapewnia niższe koszty, wyższy poziom świadczonych usług oraz większą elastyczność dzięki:

– konsolidacji serwerów, czyli skupieniu wielu obciążeń pracą na jednej platformie sprzętowej przy zachowaniu rozwiązania „jedna aplikacja/jeden serwer" i jednoczesnej redukcji liczby fizycznych serwerów, dzięki czemu możliwe jest zachowanie tej samej funkcjonalności przy wykorzystaniu mniejszej ilości sprzętu;
– maksymalnej bezawaryjności, którą uzyskuje się poprzez podział obciążeń pracą, zapobiegając w ten sposób sytuacjom, w których niepoprawne funkcjonowanie jednej aplikacji ma negatywny wpływ na wydajność innych lub powoduje awarię systemu, co pozwala na używanie niestabilnych aplikacji w hermetycznym i bezpiecznym środowisku;

- skróceniu czasu niezbędnego do testowania kompatybilności oprogramowania, który dzięki wirtualizacji aplikacji i udostępnianiu ich na żądanie na komputerach biurowych zostaje skrócony do absolutnego minimum;
- zdolności obsługi starszych wersji programów oraz aplikacji biznesowych, często niewspółpracujących z aktualnie eksploatowanymi systemami operacyjnymi serwerów;
- zwiększeniu efektywności obsługi serwerów poprzez przenoszenie obciążeń pracą pomiędzy serwerami fizycznymi minimalizując zakłócenia w ich pracy, co powoduje, że obsługa serwerów może być przeprowadzana bez wpływu na wykonywane na nich operacje;
- usprawnionemu dostarczaniu zasobów dla określonych procesów biznesowych uniezależnionemu od procesów związanych z pozyskiwaniem sprzętu, co powoduje, że odpowiednie zmiany wprowadzane są szybko i efektywnie.

6. WIRTUALIZACJA STACJI ROBOCZYCH I APLIKACJI

W ostatnim czasie coraz większą popularność zyskuje wirtualizacja stacji roboczych/desktopów. W ramach tej gałęzi technologii wirtualizacyjnych można wyróżnić:

- tzw. wirtualizację „desktopa w desktopie", czyli nic innego jak emulację zasobów sprzętowych z wykorzystaniem oprogramowania uruchamianego w konkretnym systemie operacyjnym;
- wirtualizację aplikacji, czyli dostarczanie aplikacji na konkretną stację roboczą z serwera przy wykorzystaniu zainstalowanego na systemie operacyjnym desktopa oprogramowania, co pozwala w szybki sposób udostępnić dużej grupie użytkowników nową wersje aplikacji, lub zablokować dostęp do starej;
- wirtualizację klientów, która polega na oddzieleniu fizycznej lokalizacji, w której znajdują się zasoby sprzętowe z rezydującymi na nich obrazami klienckich maszyn wirtualnych, od miejsca, w którym znajduje się użytkownik i wykorzystaniu tzw. cienkich klientów (*ang.: thin client*) – bezdyskowych urządzeń terminalowych łączących się z wirtualną infrastrukturą desktopów przez lokalną lub rozległą sieć TCP/IP o dużej przepustowości.

Najbardziej rozbudowaną formą tego typu wirtualizacji jest wirtualizacja klientów, która od strony centrum danych, zawierającego repozytorium klienckich maszyn wirtualnych, udostępnianych na żądanie użytkownikom niewiele różni się od wirtualizacji środowisk serwerowych i może stanowić rozszerzenie usług wirtualizacyjnych w środowisku IT organizacji. Istotną różnicą jest funkcjonalność warstwy prezentacyjnej, która w tym wypadku musi zapewnić jednoczesny dostęp do pulpitów

wirtualnych maszyn wielu użytkownikom jednocześnie, w przeciwieństwie do warstwy prezentacyjnej wykorzystywanej przy wirtualizacji serwerów, gdzie dostęp do pulpitu systemu ma zwykle tylko administrator, zaś użytkownik korzysta z usług udostępnianych na serwerze.

7. PODSUMOWANIE

Każde z przedstawionych rozwiązań wirtualizacji zapewnia korzyści w zakresie organizacji IT. Wiele z wyżej wymienionych technologii związanych z wirtualizacją jest znanych i wykorzystywanych w procesach wspomagania biznesu od dawna, niektóre zaś, takie jak pełna wirtualizacja serwerów i desktopów, są stosunkowo nowe, uwzględniając skalę ich wykorzystania w środowiskach produkcyjnych.

Wirtualizację traktować należy nie jako sposób na zredukowanie liczby fizycznych urządzeń w infrastrukturze organizacji, ale jako pełny zestaw narzędzi wspomagających dostarczanie w sposób efektywny usług wspierających biznes.

Planując wirtualizaownie elementów infrastruktury IT organizacji należy jednak wybrać najodpowiedniejsze w bieżącej sytuacji rozwiązania i rozwijać je zgodnie z wcześniej opracowaną strategią wirtualizacji. Takie spojrzenie pomoże zmaksymalizować korzyści płynące z tej technologii.

Literatura

1. Formal Requirements for Virtualizable Third Generation Architectures. Gerard J. Popek, Robert P. Goldberg, Communications of the ACM, Lipiec 1974, Vol. 17, Nr 7
2. Materiały promocyjne Microsoft Corporation, VMware Inc., NetApp Corporation