

Piotr Zadora

Problematyka zarządzania IT w warunkach przetwarzania rozproszonego

Ekonomiczne Problemy Usług nr 104, 461-470

2013

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

PIOTR ZADORA

Uniwersytet Ekonomiczny w Katowicach

PROBLEMATYKA ZARZĄDZANIA IT W WARUNKACH PRZETWARZANIA ROZPROSZONEGO

Wprowadzenie

Przetwarzanie rozproszone jest pojęciem funkcjonującym od dłuższego czasu w teorii systemów operacyjnych. W klasycznym rozumieniu dotyczy ono podziału problemu obliczeniowego na mniejsze części, rozdysponowania tych części w węzłach komputerowych połączonych siecią i rozwiązywania podproblemów we wspomnianych węzłach. W kontekście przetwarzania rozproszonego używane jest pojęcie przezroczystości rozumianej jako separacja użytkownika od implementacji. Przezroczystość dotyczy sprzętu, oprogramowania i komunikacji. W wyniku postępu technologicznego, a także zmniejszenia cen sprzętu sieciowego oraz przetwarzającego i magazynującego dane, wspomniane koncepcje przetwarzania rozproszonego doczekały się uwspółcześnienia w postaci określanej angielskim terminem *cloud computing* (chmura obliczeniowa, przetwarzanie w chmurze).

Cloud computing jest pojęciem nowym i dlatego nie istnieje jedna definicja tego pojęcia o charakterze normatywnym. Zamiast tego funkcjonuje wiele definicji opisowych. Wśród nich najważniejsze wydaje się jednak być podkreślanie rozproszonego charakteru tej technologii rozumianej jako model zarządzania zasobami IT dystrybuowanymi za pośrednictwem sieci. Ważna jest również mierzalność tej technologii oraz traktowanie jej jako podstawy do budowy szczegółowych usług i aplikacji. Przetwarzanie w chmurze powinno być analizowane wieloaspektowo, nie tylko od strony techniczno-funkcjonalnej, ale również organizacyjnej i społecznej. Istotne znaczenie dla powodzenia tego modelu w rzeczywistych zastosowaniach ma łączenie go z dotychczasowymi technologiami przetwarzania istniejącymi w firmach.

1. Technologia przetwarzania w chmurze

Cloud computing stanowi połączenie cech charakteryzujących następujące tradycyjne modele przetwarzania:

- model klient–serwer – najbardziej rozpowszechnioną architekturę przetwarzania, w której akcent położony jest na dostarczanie usług oraz występuje ścisły podział na dostawców i odbiorców (konsumentów) usług;
- *grid computing* – odmianę przetwarzania rozproszonego, w której komputer centralny składa się z klastra węzłów komputerowych pracujących równolegle;
- *supercomputing* – wykorzystanie komputerów centralnych o bardzo dużej mocy obliczeniowej do przetwarzania szczególnie istotnych danych;
- *peer-to-peer* – architekturę przetwarzania rozproszonego bez scentralizowanego zarządzania, w której uczestnicy są jednocześnie dostawcami jak i konsumentami informacji;
- SOA (*Service-Oriented Architecture*) – model przetwarzania, w którym nacisk położony jest na tworzenie komponentów programowych wielokrotnego użytku w środowisku rozproszonym.

Przetwarzanie chmurowe traktuje sprzęt i oprogramowanie jak zasoby, które podlegają udostępnianiu w różnych modelach biznesowych. W tym kontekście rozróżnia się następujące poziomy udostępniania:

- IaaS (*Infrastructure as a Service*) – usługa dostępu do infrastruktury. Dostawca udostępnia zdalnie klientowi posiadaną przez siebie infrastrukturę techniczną. Sprzedawane są: czas procesorów, miejsce na dyskach i przepustowość sieci. Klient uruchamia w tej infrastrukturze posiadane przez siebie oprogramowanie. Zwykle posługuje się w tym celu narzędziami do wirtualizacji środowisk operacyjnych lub aplikacji. Odpowiedzialność dostawcy ogranicza się do utrzymywania infrastruktury w gotowości do pracy. Infrastruktura wynajmowana jest odpłatnie klientowi na określony czas. Umowa gwarantuje utrzymanie uzgodnionych parametrów wydajnościowych infrastruktury. Ten poziom przetwarzania w chmurze w najmniejszym stopniu różni się od tradycyjnych modeli przetwarzania i ma najmniejszy wpływ na zmiany sposobu funkcjonowania organizacji (klienta), gdyż pozwala zaoszczędzić tylko na wykorzystywanym sprzęcie. Zasadniczo nie różni się on od okresowo następującej w firmach modernizacji posiadanego sprzętu. Wykorzystywane dotychczas oprogramowanie będzie funkcjonowało tak jak dotychczas bez widocznych zmian. Przykład: Amazon EC2.
- PaaS (*Platform as a Service*) – usługa dostępu do platformy programistycznej, usługodawca dostarcza infrastrukturę oraz narzuca konkretne środowisko programistyczne i operacyjne niezbędne do stworzenia aplika-

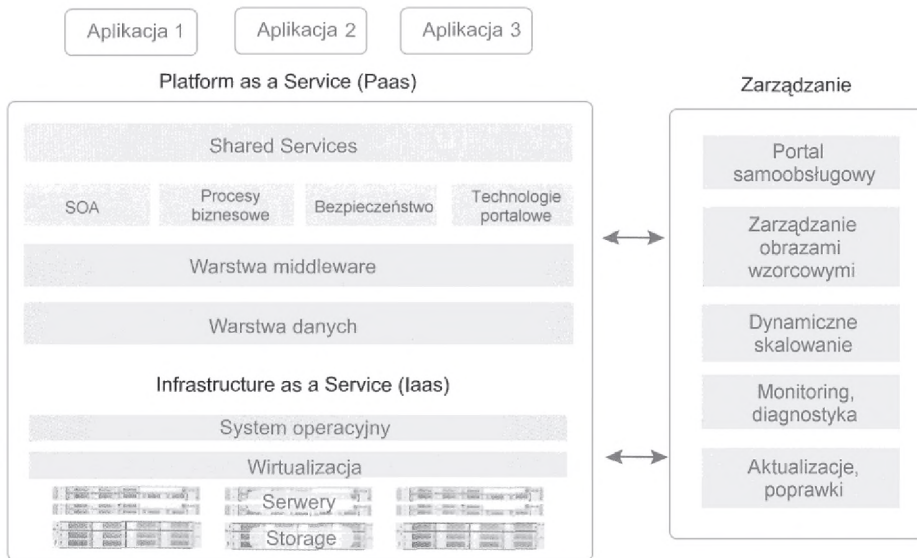
cji (np. Java, Python). Opłata dotyczy wykorzystania wirtualizowanych zasobów fizycznych (procesor/pamięć/sieć/dysk). Schemat architektury IaaS/PaaS przedstawiono na rysunku 1. Przykład: Google App Engine, Windows Azure.

- SaaS (*Software as a Service*) – usługa dostępu do aplikacji biznesowych, udostępnianie aplikacji biznesowych w formie usługi, a nie licencji. Dzięki temu klient nie ponosi kosztów początkowych, nie musi utrzymywać swojego środowiska, ani zatrudniać specjalistów do zarządzania nim. W tym wypadku podstawą rozliczenia jest zwykle ogólna liczba użytkowników lub liczba równoległe otwartych sesji. Przykład: Google Apps, Oracle CRM on Demand.

Innym rozróżnieniem jest podział na chmurę prywatną i publiczną. W chmurze publicznej (*public cloud*) usługodawcą jest firma świadcząca usługi dostępu do platformy programistycznej na zewnątrz. Dane i aplikacje firmy macierzystej przenoszone są do centrów danych należących do firm zewnętrznych. W chmurze prywatnej (*private cloud*) usługodawcą jest dział IT firmy macierzystej.

2. Udostępnianie środowisk programistycznych

Prywatna chmura obliczeniowa przedsiębiorstwa (*Enterprise Private Cloud*) realizowana jest zazwyczaj w wariantcie PaaS i udostępnia usługi IT działom oraz partnerom biznesowym (np. spółkom zależnym). Wdrożenie prywatnej chmury obliczeniowej wymaga zmiany sposobu myślenia o zasobach IT przedsiębiorstwa. Podobnie jak SOA nie jest tylko technologią opartą na usługach sieciowych (*web services*), lecz architekturą wymagającą szerszego spojrzenia na tworzenie i ponowne wykorzystanie aplikacji, podobnie *cloud computing* nie wymaga zmiany używanej technologii, lecz przede wszystkim zmiany w podejściu do zarządzania zasobami.



Rys. 1. Schemat koncepcyjny architektury IaaS/PaaS

Źródło: M. Kuratczyk: „Software Developers Journal” 2010, nr 2.

Technologia jest kwestią wtórną, która usprawnia ten proces i umożliwia zapewnienie dostępności i wydajności w warunkach skalowalnych. Korzystanie z usług publicznych dostawców *cloud computing* zwykle jest wybierane jako wariant początkowy umożliwiający testowanie nowych rozwiązań bez ponoszenia kosztów. Natomiast projekt budowy wewnętrznej platformy PaaS w firmie wymaga konkretnych nakładów początkowych, zmiany procesu zakupów sprzętu i licencji oraz prawdopodobnie zmiany sposobu rozliczania za korzystanie z nich. Ze względu na poufność przetwarzanych danych lub inne uwarunkowania wiele aplikacji nie może być uruchomionych w publicznych chmurach obliczeniowych. Budowa prywatnego środowiska programistycznego zgodnego z polityką bezpieczeństwa firmy może okazać się niezbędna. W okresie przejściowym lub dla celów testowych możliwe jest korzystanie z publicznego modelu PaaS, który gwarantuje natychmiastową dostępność wirtualnej platformy uruchomieniowej.

Ograniczeniem publicznego modelu PaaS jest narzucanie programistom konieczności korzystania z narzędzi i technologii udostępnianych przez platformę. Zwykle brak jest pełnej dowolności dodawania technologii programistycznych zgodnych z preferencjami programistów. Zaletą jest wyraźne obniżenie kosztów sprzętu, licencji, pomocy technicznej oraz personelu dzięki standaryzacji i konsolidacji środowisk. Istnieje możliwość prostego tworzenia dodatkowych środowisk na żądanie (szablony aplikacji, kopia środowiska, przenoszenie środowisk testowych do postaci produkcyjnej).

Prywatna chmura obliczeniowa organizacji umożliwia optymalizowanie wykorzystania wewnętrznych zasobów IT oraz poprawia skalowalność i odporność na awarie działających aplikacji. Prywatna chmura obliczeniowa powinna gwarantować łatwość synchronizacji z chmurami publicznymi. W konsekwencji pozwala to tworzyć rozwiązania międzycmurowe, które na co dzień wykorzystują zasoby prywatnej chmury obliczeniowej, ale w okresach wzmożonego zapotrzebowania wspomagane są zasobami zewnętrznymi.

W aspekcie technologicznym wyróżnia się szereg podejść i technologii tworzenia chmur obliczeniowych. Najważniejsze z nich to:

- Wirtualizacja – zwielokrotnianie kopii systemów operacyjnych lub środowisk uruchomieniowych działających na platformie sprzętowej realizowanej zwykle w architekturze wieloprocesorowej lub wielokomputerowej. Szacuje się, że w dużych firmach poziom utylizacji sprzętu komputerowego wynosi średnio rzecz biorąc 20%. Wirtualizacja pozwala zwiększyć stopień wykorzystania zasobów sprzętowych, zachować ciągłość działania w czasie prac konserwacyjnych (np. wymiana sprzętu) oraz pomaga zarządzać środowiskami. Pozwala tworzyć wzorcowe instalacje dostępnych na platformie komponentów, np. komponent bazy danych zawierający odpowiednio skonfigurowany system operacyjny z połączeniem do przestrzeni dyskowych, bazę danych w odpowiedniej wersji, narzędzia diagnostyczne itd.
- Warstwa pośrednia (*middleware*) – warstwa abstrakcji zwiększająca stopień przezroczystości rozproszenia komponentów sprzętowych i programowych oraz realizująca funkcje: klastrowania z synchronizacją sesji użytkowników, dynamicznego zarządzania wersjami bibliotek programowych, ułatwiania przeprowadzania testów integracyjnych i obciążeniowych.
- Serwery aplikacji działające bezpośrednio na nadzorcy (*hypervisor*) – bez pośrednictwa klasycznego systemu operacyjnego. Przykładem takiej architektury jest połączenie warstwy nadzorcy przechwytyjącej niskopoziomowe odwołania do sprzętu z wieloma instancjami wirtualnej maszyny Javy, w których uruchomiono kontenery serwletów zgodne ze standardem JEE (*Java Enterprise Edition*). Wykorzystywana jest specjalna wersja wirtualnej maszyny Javy obywająca się bez systemu operacyjnego, zwykle z dodatkowymi zabezpieczeniami (np. brak możliwości otwierania gniazd sieciowych czy plików). Źródłami danych dla aplikacji uruchamianych w kontenerach są bazy danych dostępne za pomocą standardowego mechanizmu JPA (*Java Persistence API*). Takie podejście pozwala nie tylko podnieść wydajność, dzięki poprawie zarządzania pamięcią, skróceniu ścieżki wykonania (*execution path*), ale również pozwala wyraźnie uprościć całe środowisko (brak systemu operacyjnego powoduje brak konieczności konfigurowania i aktualizacji).

- Technika *in-memory data grid* – pozwala przechowywać duże ilości danych obiektowych (np. obiekty sesji HTTP) w pamięci operacyjnej, co odciąża serwer aplikacji (np. kontener JEE). Widoczne jest wyraźne zwiększenie wydajności działania aplikacji. Takie rozwiązanie pozwala skalować aplikacje praktycznie w nieskończoność.
- Wirtualizacja pamięci masowej – bazy danych realizowane w technologii klastrów (węzłów komputerowych połączonych szybkim medium sieciowym). Takie rozwiązanie pozwala zachować ciągłość pracy baz danych, poprawia skalowalność (automatyczna migracja danych, dynamiczne odłączanie/przylączanie węzłów) i równomiernie rozkłada obciążenie pomiędzy węzłami klastra.

3. Udostępnianie aplikacji

Analizując sposoby udostępniania aplikacji w chmurze, można wyróżnić trzy podstawowe modele¹.

Pierwszy to oferowanie prostych aplikacji dużej grupie indywidualnych użytkowników. Ten model jest używany m.in. przez firmy Amazon, Google i Microsoft. Do wymienionych aplikacji należą: oprogramowanie do obsługi poczty elektronicznej, firmowej komunikacji, hostowania stron WWW, serwerów aplikacyjnych.

Drugi model dotyczy udostępniania w chmurze oprogramowania dużych producentów na potrzeby rynku małych i średnich przedsiębiorstw. Wymagana jest integracja udostępnianych rozwiązań z rozwiązaniami innych dostawców. Model ten obejmuje m.in. wdrażanie systemów klasy ERP. Rozpatrując systemy klasy ERP, można wyróżnić cztery sposoby udostępniania ich w chmurze:

- całkowite hostowanie i utrzymanie systemu klasy ERP przez firmę zewnętrzną;
- rozwiązanie pośrednie – hostowanie aplikacji i systemu operacyjnego, pozostałe czynności zapewniają administratorzy macierzystej firmy;
- rozwiązanie pośrednie – usługodawca zajmuje się również utrzymaniem aplikacji;
- *in-house* – aplikacja pracuje w obrębie infrastruktury IT macierzystego przedsiębiorstwa.

W modelu SaaS udostępniane są m.in. systemy Asseco SAFO ERP, Microsoft Dynamics, Oracle E-Business Suite, SAP ERP². Wymienia się cztery podstawowe

¹ A. Niziolek: *Podróż w chmury*, „Computerworld Polska” 2011, nr 8/924, s. 25.

² D. Piotrowski: *Cloud w małej i średniej firmie*, „Computerworld Guide” 2010, nr 42/910, s. 5.

korzyści, jakie gwarantuje wykorzystanie systemu klasy ERP w trybie SaaS. Są to korzyści: operacyjne, finansowe, techniczne i strategiczne³.

Trzeci model udostępniania aplikacji w chmurze to pełna wirtualizacja. Charakteryzuje ją dostęp do tzw. kompletnego centrum przetwarzania danych w chmurze. Integracja jest z góry zapewniona. Odbiorcą rozwiązań tego typu są duzi klienci biznesowi.

4. Zmiany w zarządzaniu IT

Obecnie problematyka zarządzania IT w organizacji dotyczy zarówno strony technicznej (sprzętu komputerowego i sieciowego, konfiguracji serwerów, sprzętowej strony magazynów danych itp.) jak i personalnej (specjaliści od utrzymywania sprzętu w gotowości i reagowania na awarie). Rozpowszechnienie się przetwarzania chmurowego spowoduje przeniesienie technicznej obsługi infrastruktury IT (*help-desk*) na zewnątrz do firm specjalizujących się w świadczeniu tego rodzaju usług. Firmy świadczące usługi typu *help-desk* będą dysponowały odpowiednimi narzędziami i doświadczonym personelem, a w konsekwencji ceny tych usług będą niższe od kosztów utrzymywania działów *help-desk* wewnątrz organizacji. Ponadto nastąpi przeniesienie tradycyjnie rozumianych usług serwerowych do amorficznej chmury obliczeniowej, w której nie będzie już można dokonać jednoznacznej relacji usługa–zasoby–sprzęt. W konsekwencji działy IT organizacji nie będą się już zajmowały instalowaniem, konfigurowaniem i utrzymaniem serwerów. Zakres ich obowiązków ulegnie zawężeniu do zarządzania projektami wykorzystującymi zasoby informatyki całkowicie lub prawie całkowicie wyprowadzone na zewnątrz macierzystej organizacji. W dalszym okresie czasu menedżerowie IT staną się w większym niż dotąd stopniu konsultantami biznesowymi.

Prawdopodobnie zwiększy się zakres odpowiedzialności menedżerów IT za zapewnianie działom organizacji ciągłego dostępu do Internetu o odpowiednich parametrach wydajnościowych. Nawet niewielki przestój lub ograniczenie przepustowości transferu może spowodować paraliż działania organizacji. Pracownik działu IT prawdopodobnie będzie współpracował z dostawcą łączności internetowej (ISP) w celu usuwania skutków awarii oraz będzie nadzorował utrzymywanie przez dostawcę wynegocjowanych parametrów transmisji. Dwa ostatnie wymienione zadania mogą jednak być również wyprowadzone (wynajęte) poza organizację macierzystą.

Model przetwarzania chmurowego okaże się zapewne bardzo przydatny dla organizacji wytwarzających i wdrazających oprogramowanie, gdyż pozwoli usunąć tradycyjnie istniejące w tej branży poważne napięcie między deweloperami

³ <http://www.isof.pl>.

a wdrożeniowcami. Polega ono na utrudnionej w tradycyjnych modelach przetwarzania synchronizacji potrzeb sygnalizowanych przez wdrożeniowców z kolejnymi wersjami oprogramowania wytwarzanymi przez programistów. Zwykle w tradycyjnych środowiskach programiści nie nadążają reagować na postulaty wdrożeniowców, co powoduje denerwujące opóźnienia. Środowisko chmurowe zdecydowanie powinno przyspieszyć wdrażanie kolejnych wersji oprogramowania dzięki praktycznemu pominięciu fazy reinstalowania (*deployment*).

Cloud computing z pewnością pomoże w rozwoju młodych firm (*start-up*), które zaoszczędzą na wydatkach na infrastrukturę. Muszą one jednak wziąć pod uwagę rachunek ekonomiczny i oszacować, czy opłaca im się płacić za dostęp do chmury, pasmo transferu i wolumen danych gromadzonych w tym rozproszonym środowisku. Może się okazać, że pewna część zasobów musi być dostępna ciągle i jest na tyle często modyfikowana, że warto ją utrzymywać wewnątrz organizacji.

Wiele wskazuje na to, że w efekcie rozpowszechnienia się przetwarzania chmurowego ulegnie zmianie sposób mierzenia niezawodności systemów informatycznych. Dotychczas najbardziej rozpowszechnioną miarą niezawodności systemów było wyrażenie:

$R = \text{MTBF} / (\text{MTBF} + \text{MTTR})$, gdzie

MTBF – średni czas pomiędzy awariami,

MTTR – średni czas do powtórnej uruchomienia.

Najbardziej widoczną słabością tej miary było ograniczenie jej tylko do kwestii sprzętowych. Wpływało to na decyzje inwestycyjne podejmowane przez działy IT organizacji, które najczęściej modernizowały i rozbudowywały sprzęt. Natomiast w dłuższej perspektywie okazywało się, że użytkowanie kolejnych wersji systemów skutkowało większym wpływem na całkowitą ich niezawodność również takich czynników jak: jakość oprogramowania oraz czas wdrażania. Przetwarzanie w chmurze, które zdejmuje odpowiedzialność działów IT ze sprzętu, a również może znacząco zmniejszyć czas wdrażania kolejnych wersji oprogramowania, na pewno przyczyni się do powstania innych sposobów mierzenia niezawodności systemów użytkowanych w organizacjach. Prawdopodobnymi kandydatami do zastąpienia tradycyjnie mierzonej niezawodności są: integralność, produktywność i wydajność oprogramowania.

Dla administratorów systemów znacznie lepszą miarą określania wydajności centrów danych w warunkach przetwarzania chmurowego jest stosunek przepustowości łącza (pasma) do poniesionych kosztów oraz stosunek ilości wykonanych transakcji do wydatkowanej energii⁴.

Rola sprzętu zmniejszy się znacząco w optyce menedżerów IT. Nawet w wewnętrznych chmurach prywatnych występuje przezroczystość infrastruktury sprzętowej i trudno jest ocenić bieżące obciążenie jej poszczególnych komponentów.

⁴ S. Greengard: *Cloud Computing Transforms IT*, <http://www.baselinemag.com>.

tów. W przypadku chmur zewnętrznych o dostępie publicznym decydenci po stronie klienta mogą się nigdy nie dowiedzieć, jaki konkretnie sprzęt uruchamia ich oprogramowanie. W warunkach przetwarzania rozproszonego w chmurze obliczeniowej zyskuje na znaczeniu stabilność i przepustowość łączy. Długi czas dostępu i wolny transfer może być traktowany tak samo jak awaria w tradycyjnych modelach przetwarzania.

Podsumowanie

Tradycyjne modele przetwarzania konstruowane w oparciu o własne, rozbudowane centra danych wymagające licznego personelu powoli odchodzą w przeszłość. Współczesne organizacje poszukujące oszczędności z dużym zainteresowaniem śledzą rozwój przetwarzania w chmurze, które ma istotny wpływ na sferę zarządzania zasobami IT. Przetwarzanie chmurowe nie jest tylko efektem integracji poprzednich modeli przetwarzania, lecz raczej ich rozwinięciem i przeniesieniem do nowego środowiska powstałego w wyniku postępu technologii oraz jej potaniecia. Wolne moce przerobowe nadmiarowo rozbudowanych, istniejących centrów danych spowodowały chęć ich lepszego zagospodarowania w okresach bezczynności. Tak postąpiła firma Amazon, która stworzyła w tym celu jedną z pierwszych publicznych chmur obliczeniowych. Pojawiły się firmy komercyjnie udostępniające wirtualizowaną infrastrukturę w jednym lub kilku z opisanych w artykule modeli.

Przetwarzanie w chmurze ma duży wpływ na zmiany zachodzące w sferze zarządzania infrastrukturą IT współczesnej organizacji niezależnie od jej potencjału i struktury wewnętrznej. Zmienia optykę postrzegania zasobów IT i stwarza szansę na korzystanie z nich w sposób bardziej efektywny i elastyczny. Otwiera nowe możliwości przed małymi firmami, firmami wirtualnymi albo prowadzącymi działalność o charakterze niszowym oraz takimi, które charakteryzują się okresowym zapotrzebowaniem na zasoby IT.

Literatura

1. Greengard S.: *Cloud Computing Transforms IT*, <http://www.baselinemag.com>.
2. Kuratczyk M.: *Enterprise Private Clouds*, „Software Developers Journal” 2010, nr 2.
3. Niziołek A.: *Podróż w chmury*, „Computerworld Polska” 2011, nr 8/924.
4. Piotrowski D.: *Cloud w małej i średniej firmie*, „Computerworld Guide” 2010, nr 42/910.
5. <http://www.isof.pl>.

IT MANAGEMENT IN DISTRIBUTED COMPUTING

Summary

As a consequence of progress in technology and decrease of the hardware prices the old concepts of distributed computing has become widely known as cloud computing. Cloud computing is the use of computing resources (hardware and software) that are delivered as a service over a network. This model of computing should be carefully analyzed either from the technical and functional point of view or from organizational and social point of view.

The ultimate goal of cloud computing is to allow IT to deliver more services more frequently to the business and customers at low cost. Hence it is driving a lot of changes to how applications are developed and delivered and how IT services are provided. Using cloud services, is not an all-or-nothing scenario, so to be successful, IT managers need to develop strategies for blending cloud technology with their existing internal infrastructure. In order to be effective IT management should to collaborate with developers. The IT manager will likely not be building or maintaining server equipment. The future will require managers to be more like business consultants.

Translated by Piotr Zadora