Systemy Operacyjne (6)

Marcin Gogolewski
marcing@wmi.amu.edu.pl

Uniwersytet im. Adama Mickiewicza w Poznaniu

Poznań, 16 grudnia 2018

Cechy hipernadzorcy

1. bezpieczeństwo (pełna kontrola nad zasobami)
2. wierność (programy powinny działać tak, jak na rzeczywistym sprzęcie)
3. wydajność (większość kodu powinna działać bez ingerencji hipernadzorcy)

Pierwsze dwa wymagania stosunkowo łatwo spełnić, np. za pomocą *Bochsa*, ale. . . pełna emulacja sprzętu w oprogramowaniu nie działa zbyt wydajnie.

Problem na architekturze x86

Niektóre instrukcje zachowują się inaczej w trybie użytkownika, a inaczej w trybie jądra (np. modyfikacja ustawień MMU) – *instrukcje wrażliwe*.

Instrukcje *uprzywilejowane* wywołane w trybie użytkownika powinny wykonywać rozkaz pułapki.

Problem

Nie można było bezpośrednio uruchomić programu wykonującego takie instrukcje w środowisku wirtualnym!

Problem na architekturze x86 (2)

Problem udało się rozwiązać dopiero w sprzęcie produkowanym po od 2005 (po ok. 25 latach od wprowadzenia niefortunnego rozwiązania). Technologia nazywa się VT (*Virtualization Technology* – SVM *Secure Virtual Machine* w przypadku AMD).

System można uruchomić w „kontenerze” tak, by wykonanie wrażliwej instrukcji powodowało przekazanie sterowania hipernadzorcy (technika *trap and emulate* – przechwyć i emuluj).

Uwaga

W rzeczywistości wirtualizacja działała wcześniej (VMware od 1999 roku), ale. . . w tym rozwiązaniu część kodu była „przepisywana w locie” (podobną technikę stosowały procesory *Transmety*, np. *Crusoe* wykorzystujący *Code Morphing*).

Tzw. *binary translation* podmieniało kod niebezpiecznych instrukcji na równoważny (czasami nie wymagało to nawet odwołania do hipernadzorcy).

Parawirtualizacja

Zamiast udawać rzeczywistą maszynę można **jawnie udostępnić zestaw „hiperwywołań”** do wykonywania instrukcji wrażliwych (coś takiego działało w systemie IBM'a o nazwie VM, z 1972 roku).

Wada

Nie da się tego zrobić bez przepisania choć fragmentu systemu operacyjnego (trzeba mieć dostęp do kodu).

Zaleta

System może działać bez zauważalnej straty wydajności (dodany jeden poziom wywołania funcji w przypadku operacji wrażliwych).

W przypadku dobrze zdefiniowanych hiperodwołań można by konstruować nowe systemy tak, jak w przypadku nowego sprzętu.

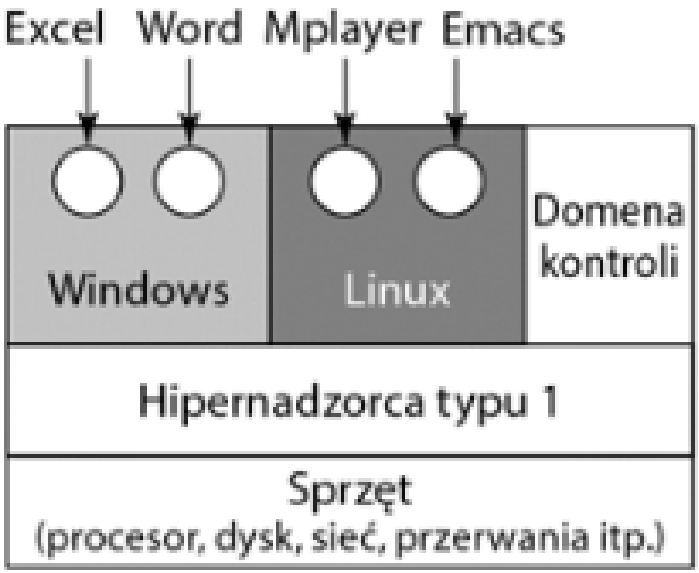
Inne podejście

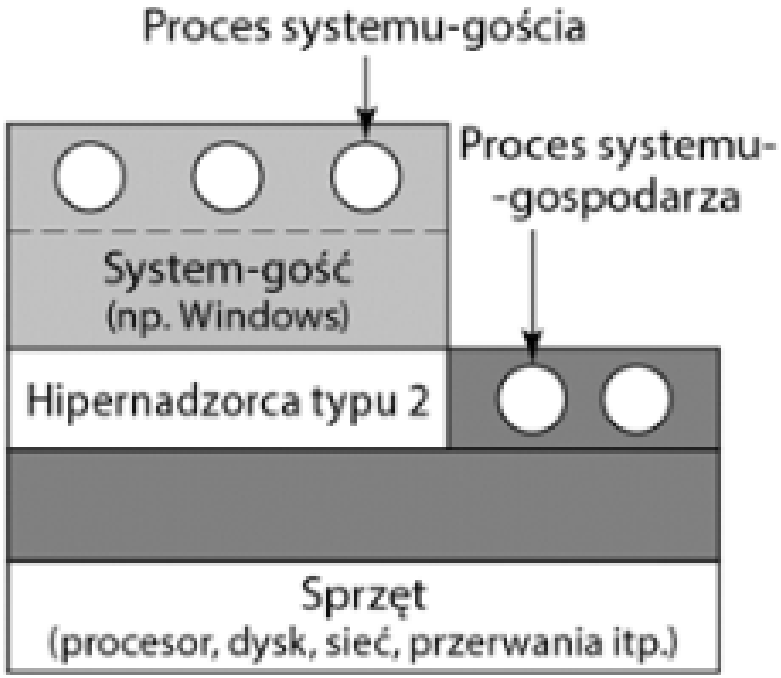
Wine (Wine Is Not an Emulator)

Programom napisanym dla Windows udostępniane są biblioteki umożliwiające uruchamianie ich pod kontrolą systemu Linux.

Uruchamiane są wyłącznie aplikacje trybu użytkownika (niepotrzebna obsługa instrukcji wrażliwych), na tej samej architekturze sprzętowej (brak konieczności emulacji).

Hipernadzorcy typu 1 i hipernadzorcy typu 2

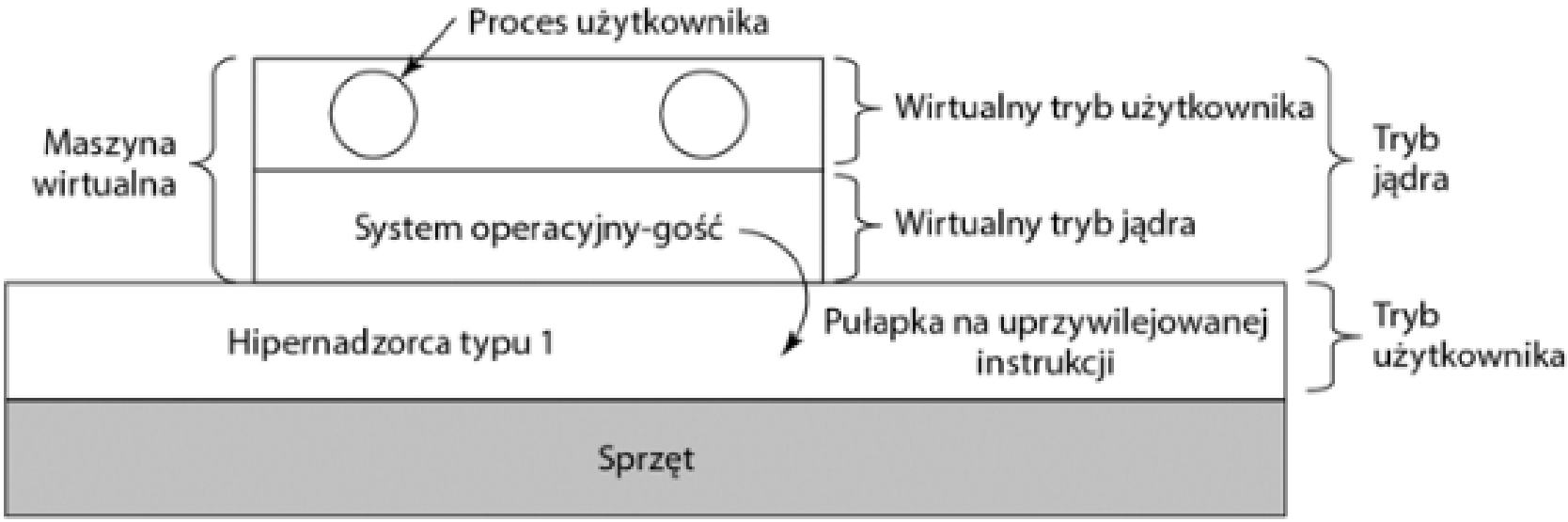




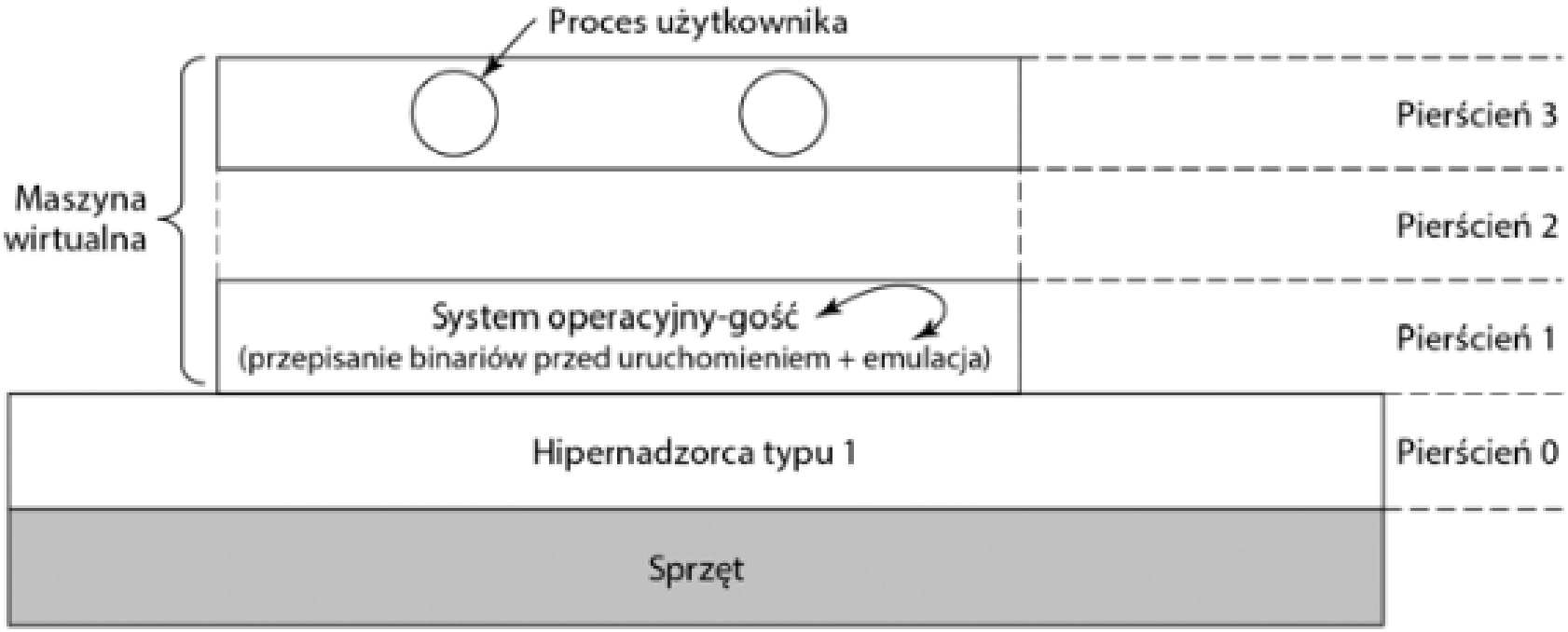
Hipernadzorcy typu 1 i hipernadzorcy typu 2

Metoda wirtualizacji	Hipernadzorcy typu 1	Hipernadzorcy typu 2
Wirtualizacja bez obsługi sprzętu	ESX Server 1.0	VMware Workstation 1
Parawirtualizacja	Xen 1.0	
Wirtualizacja z obsługą sprzętu	vSphere, Xen, Hyper-V	VMware Fusion, KVM, Parallels
Wirtualizacja procesów		Wine

Technologia VT



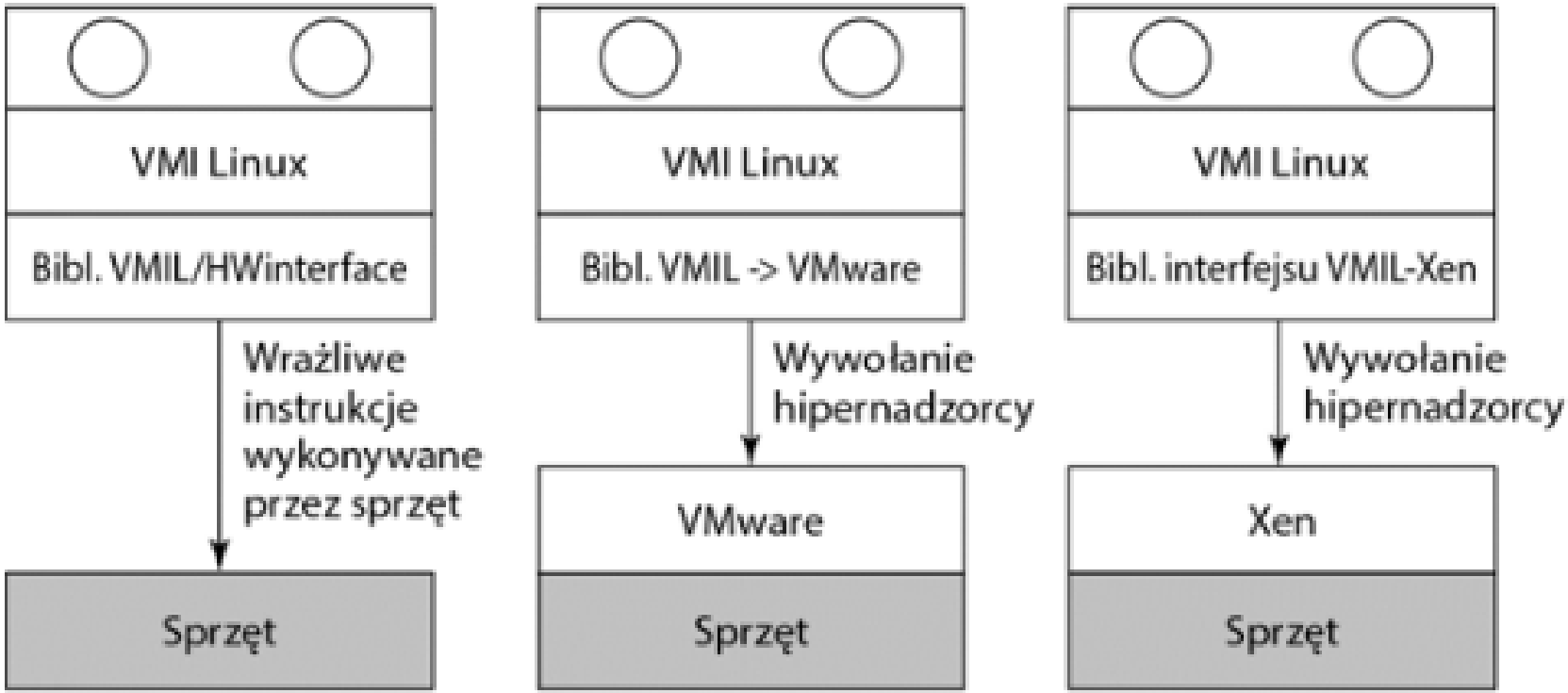
Bez technologii VT (m.in. wczesny VMware)



Przełączenie światów (*world switch*)

W przypadku, gry H. typu 2 chce dać dostęp do sprzętu musi „posprzątać" po operacji tak, by system hosta mógł działać poprawnie (sterownik w systemie hosta).

Virtual Machine Interface



Porównanie: na sprzęcia, na VMware, w Xen.

Wirtualizacja pamięci

Zwykle konieczne jest dodanie dodatkowej warstwy wirtualizacji. Mechanizm taki (*Extended Page Tablse* na Intelu) jest czasami sprzętowy.

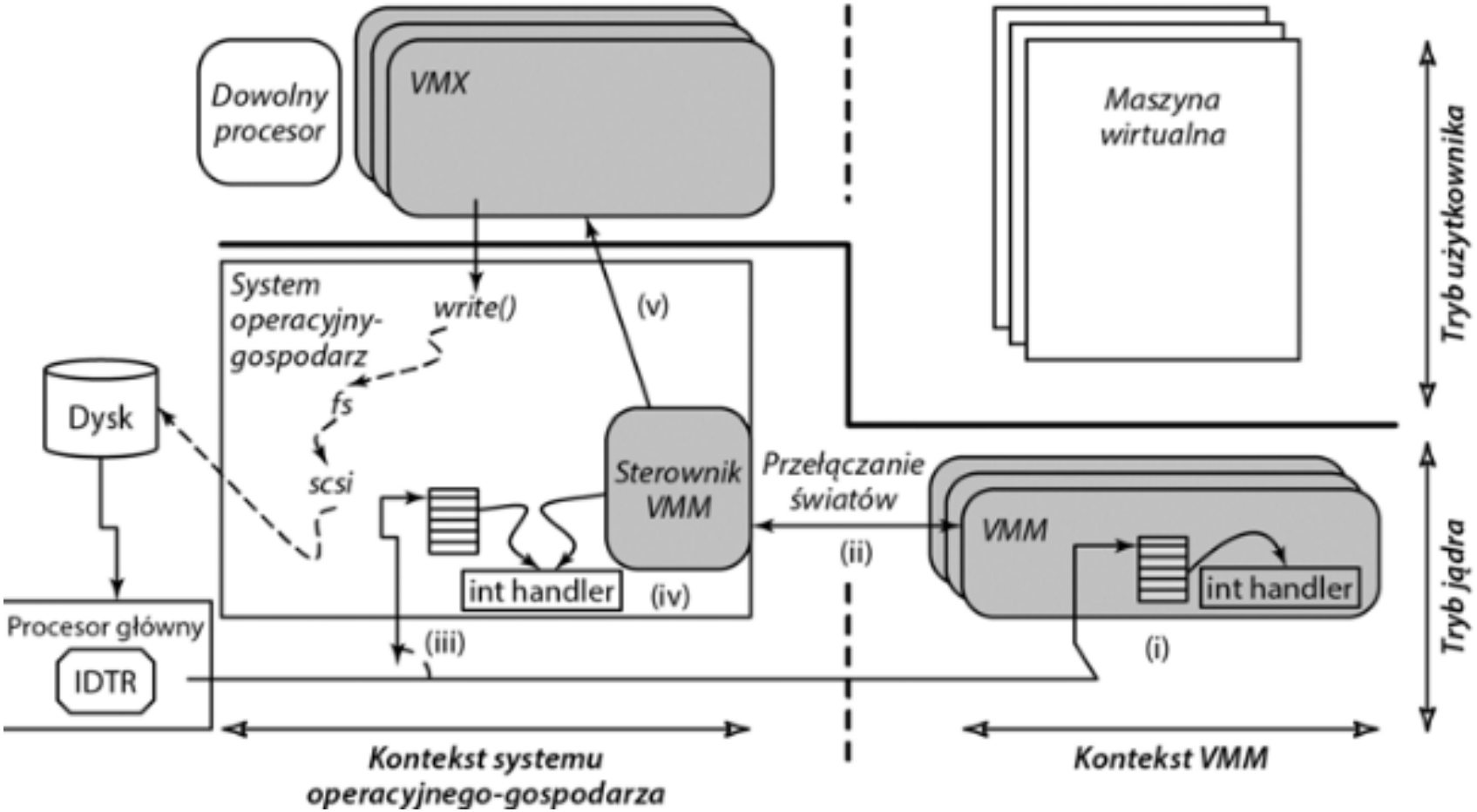
Rozszerzone tablice stron

balonikowanie (*balooning*), memory overcommitment, deduplikacja

Inne technologie wspomagające i pomysły

- wirtualizacja SR-IOV (PCIe) (*Single Root IO Virtualization*)
- urządzenia wirtualne (kompletne systemy do uruchomienia w VM)
- problemy licencyjne
- migawki
- chmury obliczeniowe (IaaS, PaaS, SaaS) – migracja maszyn
- VMware Hosted Architecture

VMware Hosted Architecture



Cele i zagrożenia

Cel	Zagrożenie
Poufność	Ujawnienie danych
Integralność	Uszkodzenie danych
Dostępność	Blokada usługi

Inne tematy związane z bezpieczeństwem

- słabe hasła
- bezpieczne przechowywanie haseł
- moduły TPM
- DRM
- hasła jednorazowe/CHAP
- kanarki
- podszywanie się pod ekran logowania

Wersje Windows

Rok	MS--DOS	Windows na bazie MS-DOS-a	Windows na bazie NT	Modern Windows	Uwagi
1981	1.0				Pierwsze wydanie dla komputera IBM PC
1983	2.0				Obsługa platformy PC/XT
1984	3.0				Obsługa platformy PC/AT
1990		3.0			Sprzedano dziesięć milionów kopii w dwa lata
1991	5.0				Dodano mechanizmy zarządzania pamięcią
1992		3.1			Stworzony tylko z myślą o platformie 286 i nowszych
1993		Windows NT 3.1			
1995	7.0	95			System MS-DOS wbudowano w system Win95
1996			Windows NT 4.0		
1998		98			
2000	8.0	Windows Me	2000		System Win Me był gorszy od systemu Win 98
2001			Windows XP		Zastąpił system Windows 98
2006			Windows Vista		System Windows Vista nie był w stanie zastąpić systemu Windows XP
2009			Windows 7		Znacznie poprawiony w porównaniu z systemem Vista
2012				Windows 8	Pierwsza nowoczesna wersja
2013				8.1	Firma Microsoft zaczęła stosować metodę „błyskawicznych wydań”

Geneza nazwy

Zwykle symbolem „chmurki” oznaczana jest „zewnętrzna sieć Internet”, w odróżnieniu od konkretnych lokalizacji ta jest „gdzieś tam”.

Gdzie użyta po raz pierwszy?

Prawdopodobnie w artykule: „The Self-governing Internet: Coordination by Design” napisanym przez: Sharon Eisner Gillett i Mitchell Kapor (na Workshop w 1996)
`http://ccs.mit.edu/papers/CCSWP197/CCSWP197.html`

Przetwarzanie w chmurze

1. pula zasobów dla każdego klienta
2. wirtualizacja w celu maksymalnego wykorzystania sprzętu
3. elastyczne skalowanie
4. automatyczne tworzenie/usuwanie instancji
5. opłata za wykorzystane zasoby

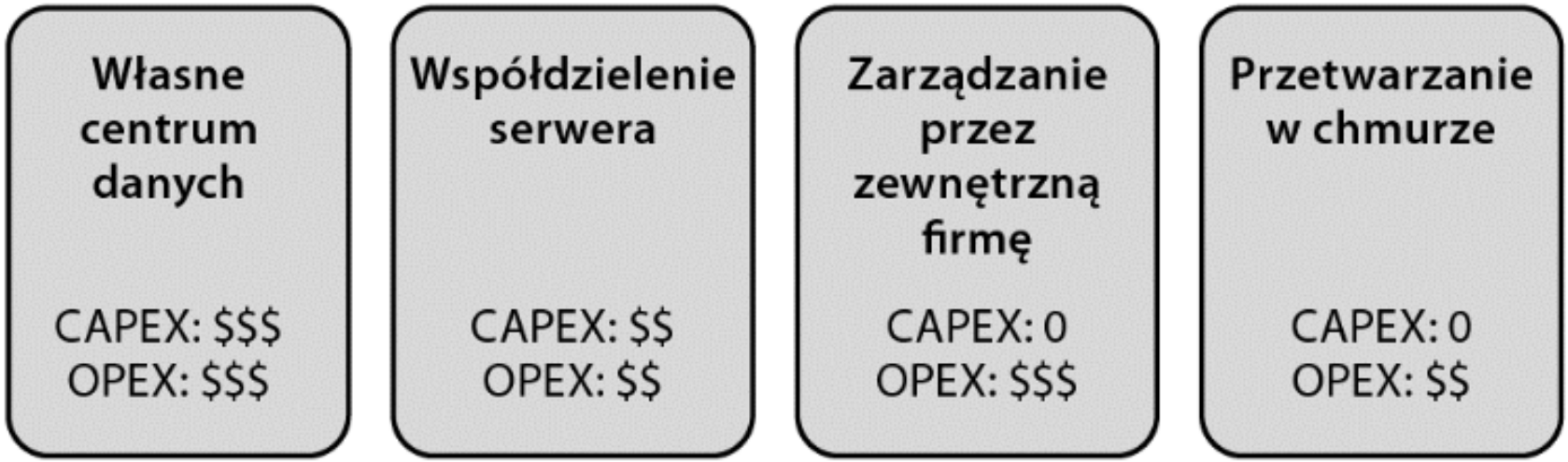
Ewolucja usług

- dedykowany serwer prywatny
- wynajmowany serwer (outsourcing)
- chmury obliczeniowe

Ewolucja kosztów – definicje

OPEX – operating expenditures

CAPEX – capital expenditures



Uwaga

W przypadku wynajmu serwera umowy sa zwykle na co najmniej rok, w chmurze obliczeniowej płacimy wyłącznie za używanie.

Dlaczego lepiej OPEX niż CAPEX?

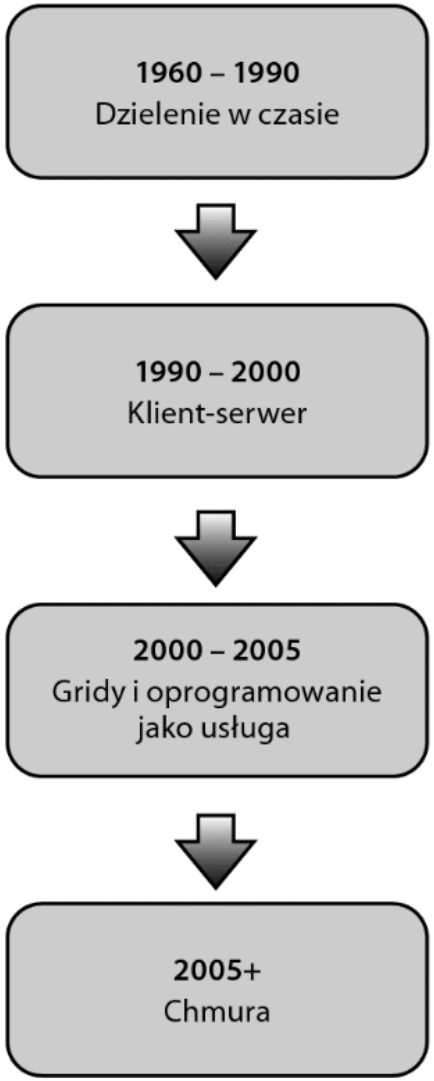
- zwykle w każdym przypadku niższe
- jak usługa nie wypali, mniejsza strata (dużo)
- właściciel chmury kupuje sprzęt i energię po cenach hurtowych!

Inne zalety chmur obliczeniowych

- gotowe narzędzia (zarządzanie, testy wydajności)
- niższy czas konfiguracji usług
- wyższe bezpieczeństwo

`http://coreygoldberg.blogspot.com/2009/02/pylot-web-load-testing-from-amazon.html`

Ewolucja przetwarzania



Wprowadzenie pojęcia „maszyny wirtualnej”

Wysoki koszt korzystania z maszyn *mainframe* powoduje, że czas niewykorzystany przez jednego użytkownika może zużyć inny

Informatyka zdefiniowana na nowo

Model klient-serwer dzieli zadania pomiędzy systemy klienckie, które wysyłają żądania, a serwery, które odpowiadają za pośrednictwem sieci

Wczesne koncepty „przetwarzania na żądanie”

Potężne zestawy ogólnodostępnego sprzętu wykonują duże zadania obliczeniowe. Przeglądarki internetowe umożliwiają dostęp do złożonych aplikacji internetowych

Przetwarzanie na żądanie zdefiniowane od nowa

Zasoby centrów danych dostępne na żądanie jako usługi dostępne za pośrednictwem przeglądarki internetowej

Nowy pomysł?

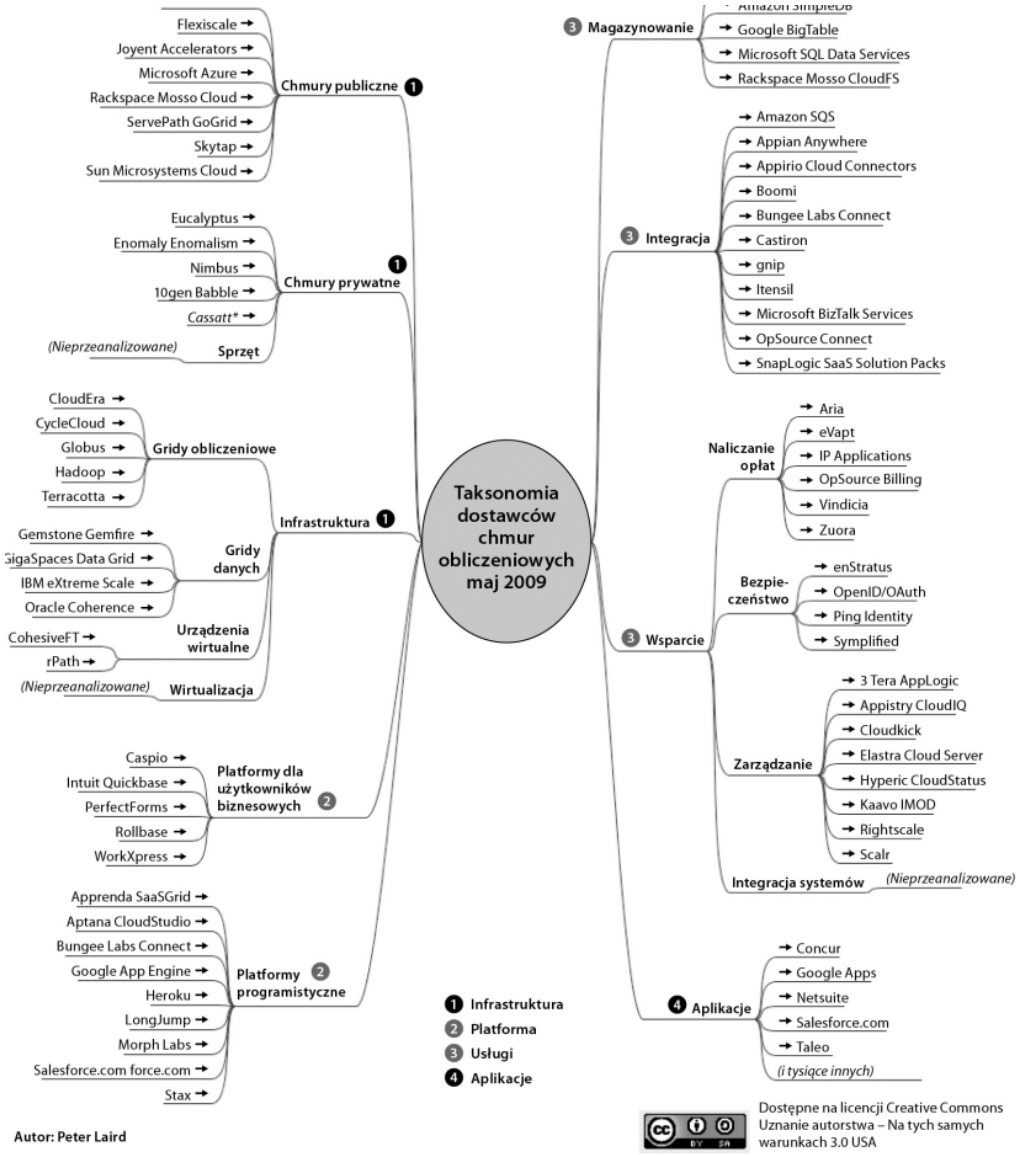
Niezupełnie nowy, wirtualizacja istniała na komputerach typu „mainframe”, jednak procesory na rynek konsumencki nie zapewniały wsparcia sprzętowego.

Architektura zorientowana na usługi

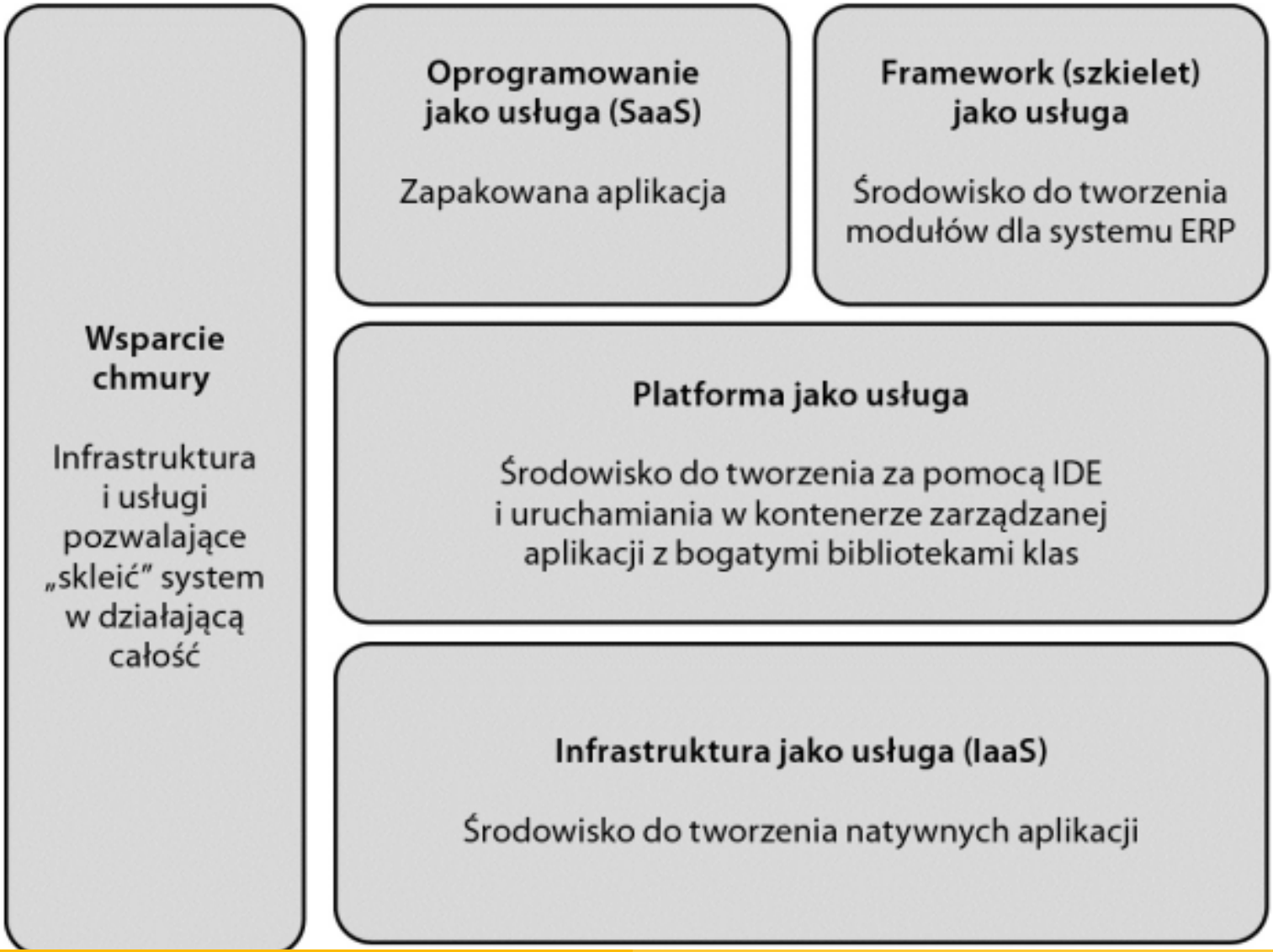
Zamiast dużych, monolitycznych systemów podział na tzw. usługi (*ang. service*) posiadające konkretne zadanie do wykonania i dobrze opisany interface.

SaaSi Software as a Service

Zmiana modelu biznesowego aplikacji. Zamiast zakupu (drogo) i opłaty za wsparcie techniczne (zwykle poniżej 20 procent), wynajem.


Taksonomia dostawców chmur obliczeniowych maj 2009
Chmury publiczne
Flexiscale
Joyent Accelerators
Microsoft Azure
Rackspace Mosso Cloud
ServePath GoGrid
Skytap
Sun Microsystems Cloud
Chmury prywatne
Eucalyptus
Enomaly Enomalism
Nimbus
10gen Babble
Cassatt*
Sprzęt
(Nieprzeanalizowane)
Infrastruktura
Gridy obliczeniowe
CloudEra
CycleCloud
Globus
Hadoop
Terracotta
Gridy danych
Gemstone Gemfire
GigaSpaces Data Grid
IBM eXtreme Scale
Oracle Coherence
Urządzenia wirtualne
CohesiveFT
rPath
Wirtualizacja
(Nieprzeanalizowane)
Platformy dla użytkowników biznesowych
Caspio
Intuit Quickbase
PerfectForms
Rollbase
WorkXpress
Platformy programistyczne
Apprenda SaaSGrid
Aptana CloudStudio
Bungee Labs Connect
Google App Engine
Heroku
LongJump
Morph Labs
Salesforce.com force.com
Stax
Autor: Peter Laird
Magazynowanie
Google BigTable
Microsoft SQL Data Services
Rackspace Mosso CloudFS
Integracja
Amazon SQS
Appian Anywhere
Appirio Cloud Connectors
Boomi
Bungee Labs Connect
Castiron
gnip
Itensil
Microsoft BizTalk Services
OpSource Connect
SnapLogic SaaS Solution Packs
Wsparcie
Naliczanie opłat
Aria
eVapt
IP Applications
OpSource Billing
Vindicia
Zuora
Bezpieczeństwo
enStratus
OpenID/OAuth
Ping Identity
Symplified
Zarządzanie
3 Tera AppLogic
Appistry CloudIQ
Cloudkick
Elastra Cloud Server
Hyperic CloudStatus
Kaavo IMOD
Rightscale
Scalr
Integracja systemów
(Nieprzeanalizowane)
Aplikacje
Concur
Google Apps
Netsuite
Salesforce.com
Taleo
(i tysiące innych)
1 Infrastruktura
2 Platforma
3 Usługi
4 Aplikacje
Dostępne na licencji Creative Commons Uznanie autorstwa – Na tych samych warunkach 3.0 USA

Podział


Wsparcie chmury
Infrastruktura i usługi pozwalające „skleić" system w działającą całość
Oprogramowanie jako usługa (SaaS)
Zapakowana aplikacja
Framework (szkielet) jako usługa
Środowisko do tworzenia modułów dla systemu ERP
Platforma jako usługa
Środowisko do tworzenia za pomocą IDE i uruchamiania w kontenerze zarządzanej aplikacji z bogatymi bibliotekami klas
Infrastruktura jako usługa (IaaS)
Środowisko do tworzenia natywnych aplikacji

Ewolucja chmur

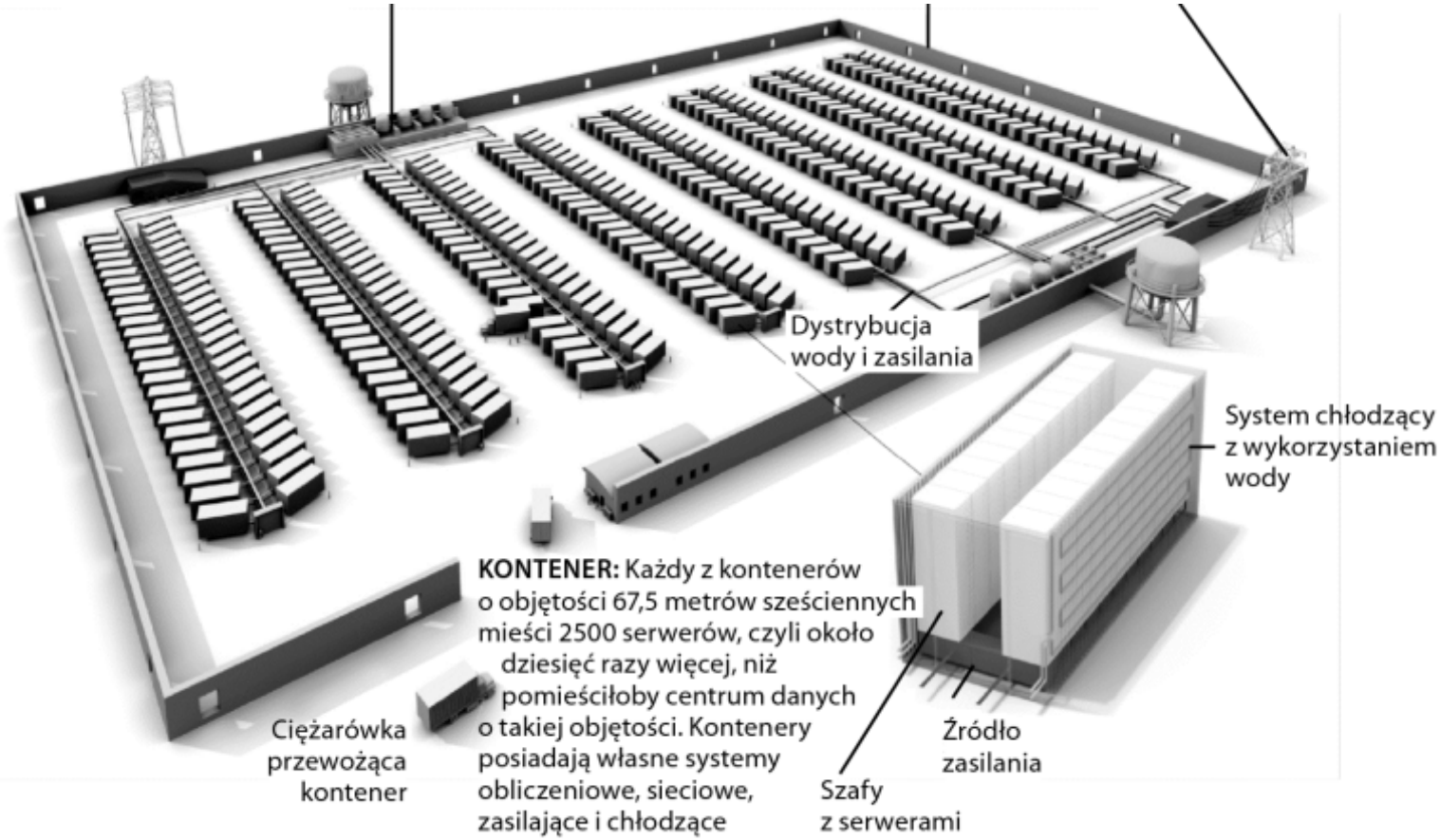
Chmury prywatne

Grupy serwerów (pracujących w *centrach danych* współpracujące w modelu „chmury”, ale dostępne wewnątrz organizacji (np. początki Amazon, Google – choćby wyszukiwarka, etc).

Chmury hybrydowe

W momencie, w którym nie wystarcza zasobów wewnętrznych, część usług przenosi się na zewnątrz.

Centrum danych (25tysm^2, 300MW, 400×2500 serwerów)



Współczynnik wydajności energetycznej (PUE)

Power Usage Effectiveness określa wydajność centrum danych. PUE wyznacza się, dzieląc ilość energii dostarczanej przez centrum danych przez ilość niezbędną do działania infrastruktury obliczeniowej wewnątrz. Wydajność poprawia się, gdy PUE zmniejsza się w kierunku wartości 1. Według organizacji Uptime Institute średnia wartość PUE typowego centrum danych to 2,5. Oznacza to, że z każdego 2,5W na liczniku tylko 1W jest wykorzystywany do obliczeń. Z szacunków Uptime wynika, iż większość centrów mogłaby osiągnąć wynik 1,6 PUE, stosując najlepszy sprzęt i dobre praktyki. Google i Microsoft zbliżają się do wartości 1,125 — to wynik nieosiągalny dla firmowych, a nawet współdzielonych centrów danych.