WIRTUALIZACJA
teoria i praktyka

Oskar Skibski, Piotr Sikora, Mateusz Kruszyński

Plan prezentacji

- Wstęp
  - co to jest?
  - po co to jest?
- Rodzaje wirtualizacji
  - emulacja pełna
  - emulacja API
  - wirtualizacja
- Problemy w architekturze x86
- Sprzętowe wspomaganie wirtualizacji w architekturze x86
  - Intel VT-x i VT-i
  - AMD-V

Plan prezentacji cd.

- Realizacja wirtualizacji
  - Wine
  - User Mode Linux
  - Xen
  - QEMU
  - VMware
  - VirtualBox
  - Virtual PC
- Przedstawienie wybranych programów w akcji

Wirtualizacja

Co to jest?

"a technique for hiding the physical characteristics of computing resources from the way in which other systems, applications, or end users interact with those resources"

różne postacie:
wirtualna pamięć, wątki, VM

Wirtualizacja

Po co to jest?

- Używanie programów z innego systemu operacyjnego (linux na windowsie, gry commodore 64)
- Symulowanie wielu komputerów przez jeden komputer fizyczny
- Symulowanie jednego komputera przez wiele słabszych komputerów
- Odpalanie programów niebezpiecznych dla komputera
- Symulowanie innego niż rzeczywisty stanu zasobów (np. mała pamięć)
- Testowanie kompatybilności

Rodzaje wirtualizacji

- Emulacja pełna
- Emulacja API
- Wirtualizacja

Rodzaje wirtualizacji

Emulacja pełna

- Emulacja całego komputera - wirtualny system operacyjny, wirtualna pamięć
- Komputer wykonuje w pętli wszystko co robiłby system operacyjny emulowany
- Każda operacja jest emulowana, bo przez warstwę wirtualną przechodzi do jądra systemu

Rodzaje wirtualizacji

Emulacja pełna

- Zalety:
  - Duża przenośność
  - Pełna kontrola systemu głównego nad systemem emulowanym (zrzuty)
  - Brak zależności od tego co jest uruchamiane
- Wady:
  - Bardzo wolne
  - Słaby kontakt z emulowaną maszyną (wielka niewiadoma)

Rodzaje wirtualizacji

Emulacja API

- Podmiana instrukcji sięgających poza własne środowisko
- Część programów nie działa, ciężko o kompatybilność
- Doczepienie bibliotek danego systemu operacyjnego
- Przykład: Wine

Rodzaje wirtualizacji

Emulacja API

- Zalety:
  - Dużo szybszy niż emulacja pełna
  - Brak potrzeby instalowania systemu operacyjnego
- Wady:
  - Zupełnie nieprzenośne - bezużyteczne dla innych instalowanych systemów

Rodzaje wirtualizacji

Wirtualizacja

- Drugi system działający równolegle z pierwszym
- Tłumaczenie niewielkiej liczby instrukcji
- Takie same systemy oparte na tym samym standardzie

Rodzaje wirtualizacji

Wirtualizacja

- Zalety:
  - Szybsza niż emulacja pełna
  - Bardziej wszechstronne od emulacji API
- Wady:
  - Brak przenośności na inne systemy
  - Brak możliwości instalowania innych systemów innej architektury niż architektura hosta

Problemy w architekturze x86

Trochę formalizmu

Kryterium Popka - Goldberga

- Odpowiedniość – program działający na maszynie wirtualnej musi zachowywać się w dokładnie taki sam sposób, jak na rzeczywistym sprzęcie
- Kontrola zasobów – wirtualna maszyna musi w pełni kontrolować wszystkie zasoby, które są wirtualizowane
- Wydajność – większa część instrukcji musi być wykonywana bez udziału maszyny wirtualnej

źródło: Wikipedia

Problemy w architekturze x86

W systemie operacyjnym wyróżniamy trzy zbiory instrukcji dostępnych na danej architekturze:

- Instrukcje uprzywilejowane
- Instrukcja wrażliwe ze względu na kontrolę
- Instrukcje wrażliwe ze względu na wykonanie

Twierdzenie:

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

W x86 tak nie jest !

http://www.floobydust.com/virtualization/lawton_1999.txt

Problemy w architekturze x86

Inne problemy

- Aliasowanie poziomów przywilejów
- Kompresja przestrzeni adresowej
- Zmiana trybu pracy procesu (użytkownik - kernel)
- Wirtualizacja przerwań
- Dostęp do ukrytego stanu
- Kompresja poziomów przywilejów
- Częste dostępy do uprzywilejowanych zasobów

Wspomaganie sprzętowe

Intel VT-x oraz VT-i - architektura

- VT-x
  - VMX root operation i VMX non-root operation
  - Virtual Machine Control Structure
    - guest-state oraz host-state
- VT-i
  - dodatkowy bit - PSR.vm
  - Rozszerzenia Processor Abstraction Layer
  - PAL service

Wspomaganie sprzętowe

Intel VT-x oraz VT-i – rozwiązania problemów

- Rozróżnienie poziomu przywilejów 0 dla guesta i hosta
- Przeładowanie (VT-x) lub rozszerzenie (VT-i) przestrzeni adresowej
- Kontrola przerwań
- Ukryty stan procesora w VCMS
- Ulepszona obsługa wyjątków

Wspomaganie sprzętowe

AMD-V

- host mode i guest mode
- VMCB
- vmrun
- VMCALL
- Migracja VM (64Bit)

Wine

- Wine = Wine is not emulator
- Implementacje bibliotek .dll pod Linuxa
- Dwa komponenty: do .exe i Winelib
- Zalety:
  - Nie wymaga posiadania Windowsa
- Wady:
  - Niepełny zbiór .dlli
  - Tylko Linux

User Mode Linux

- Zwykły proces
- Wirtualny system
- Zintegrowany z Linuxem (2.6.0)
- 2 tryby działania:
  - Tracing Thread
  - Separate Kernel Adress Space

User Mode Linux

- Zalety:
  - Debugowanie jąder
  - Honeypot
  - Wirtualne serwery
  - Prosty w obsłudze
  - Szybko działa...
- Wady
  - ...ale wolniej od np. Xena

Xen

- Parawirtualizacja
- Obsługiwane architektury : x86, AMD64
- Systemy hosta: Linux, NetBSD, Solaris
- Systemy guesta: j.w. + inne BSD, Windows XP i Server, Plan 9
- Zalety:
  - Wydajny
- Wady:
  - Wymaga modyfikacji w systemach guestów

QEMU

- Pełny emulator
- 2 tryby emulacji: User Mode i cały system
- Emuluje: IA-32 (x86) PCs, AMD64 PCs, MIPS R4000, Sun's SPARC sun4m, Sun's SPARC sun4u, ARM (Integrator/CP i Versatile/PB), SH4 SHIX, PowerPC (PReP i Power Macintosh), ETRAX CRIS.

QEMU

- Zalety:
  - Copy-On-Write
  - Obsługa wielu architektur
  - Snapshoty
- Wady:
  - Słabo wspiera Windows
  - Nie kompiluje się z GCC 4.x

VMware

- **VMWorkstation**
  - Wielokrotne snapshoty
  - Teamy wirtualnych maszyn
  - Wirtualne dyski twarde .vmdk, wirtualne CD .iso
  - Sieć przez NAT
  - Klonowanie systemów
  - Komercyjny

Virtual Box

- Open source
- Snapshoty
- Zapisywanie stanu
- Sieciowość: NAT, HIF, wewnętrzna
- Guest Additions (np. seamless mouse)
- Możliwość współdzielenia obszarów dysku pomiędzy hostem a guestem

Virtual Box

- Obrazy dysków twardych
  - O stałym rozmiarze, rozszerzalne, różnicowe, fizyczne
  - W trybie normalnym, niezmiennym i write-through
  - Obsługuje też *.vmdk (na razie tylko write-through)
  - iSCSI
- Interfejsy: VirtualBox, VBoxManage, VBoxSDL, zdalny VRDP

Microsoft Virtual PC

- Większość funkcjonalności jak VirtualBox
- Darmowy
- Drag'n'drop
- Dyski Undo
- Virtual PC dla PowerPC
- Utrudnienia dla Linuxa
  - http://vpc.visualwin.com/

Podsumowanie

- (dyskusja)