



Czym jest XEN?

- monitor maszyn wirtualnych
- powstał jako projekt badawczy na Uniwersytecie w Cambridge w 2001 r. - Ian Pratt
- projekt prowadzi firma XenSource, Inc.
- rozwijane są dwie wersje:
 - open source (GNU GPL)
 - komercyjna XenEnterprise

Obsługiwane platformy

- sprzętowe
 - x86 (od i686, PAE), x86/64, IA64, PowerPC, SMP
- programowe
 - gospodarz
 - Linux, NetBSD 3.0
 - gość
 - Linux, NetBSD 3.0, FreeBSD 7.0, OpenSolaris 10, Plan 9, Minix, Netware
 - Windows XP SP2, Windows 2003 Server (wymaga procesora ze wsparciem dla wirtualizacji – Intel Vanderpool (VT-x/VT-i), AMD Pacifica)

Model działania

- parawirtualizacja
 - Xen jako OS wyższego poziomu
 - prezentowany poziom abstrakcji jest podobny do sprzętowego (ale nie identyczny)
 - konieczność modyfikacji nadzorowanych systemów
 - nie trzeba modyfikować aplikacji użytkownika
- system – gospodarz (dom0)
- systemy – goście (domU)

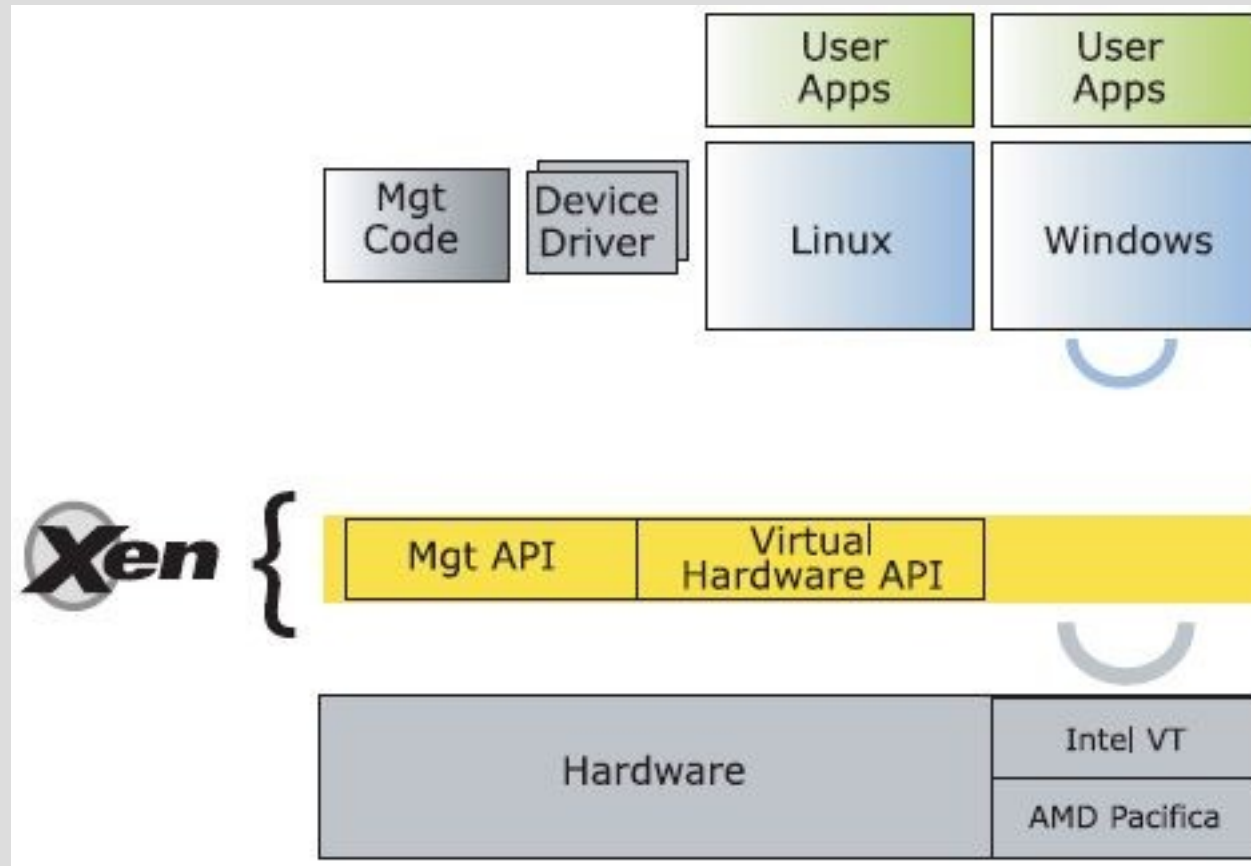
Co jest (para)wirtualizowane?

- CPU
- pamięć (MMU)
- operacje I/O
 - dyskowe
 - sieciowe

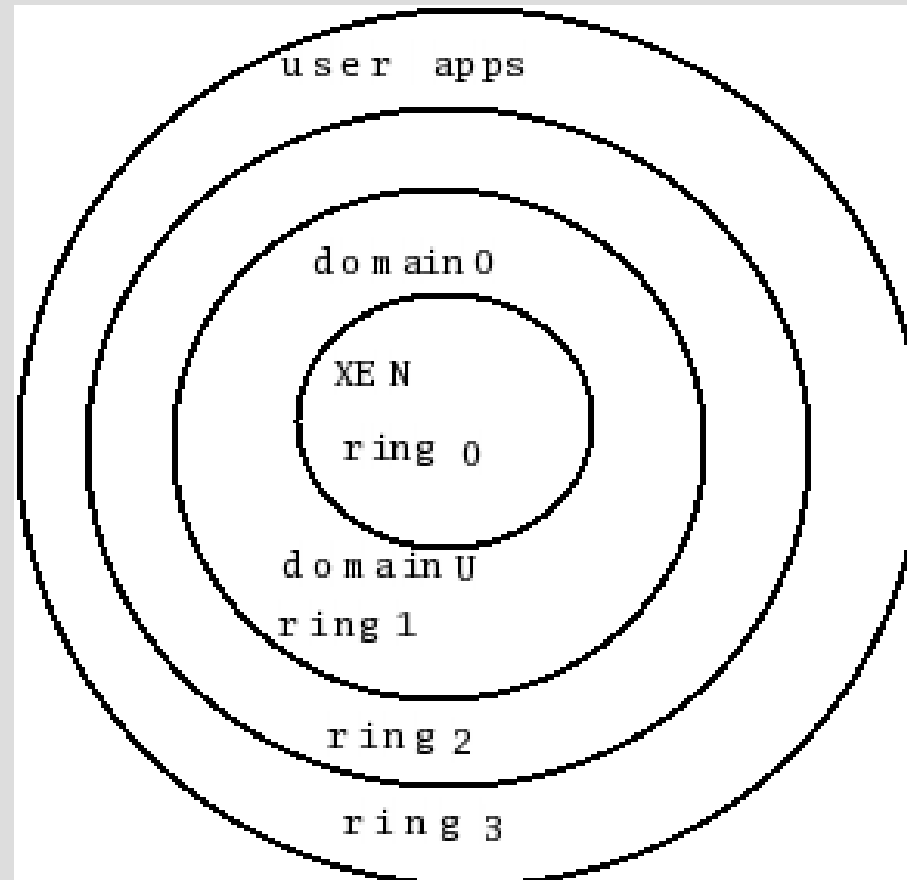
Parawirtualizacja

- hiperwywołania
 - zmodyfikowany system operacyjny – gość używa specjalnego API zamiast odwoływać się bezpośrednio do sprzętu
 - obsługa przerwań przez asynchroniczne kanały zdarzeń
 - współdzielone strony pamięci
 - mechanizm dosyć niskopoziomowy – wysoka wydajność

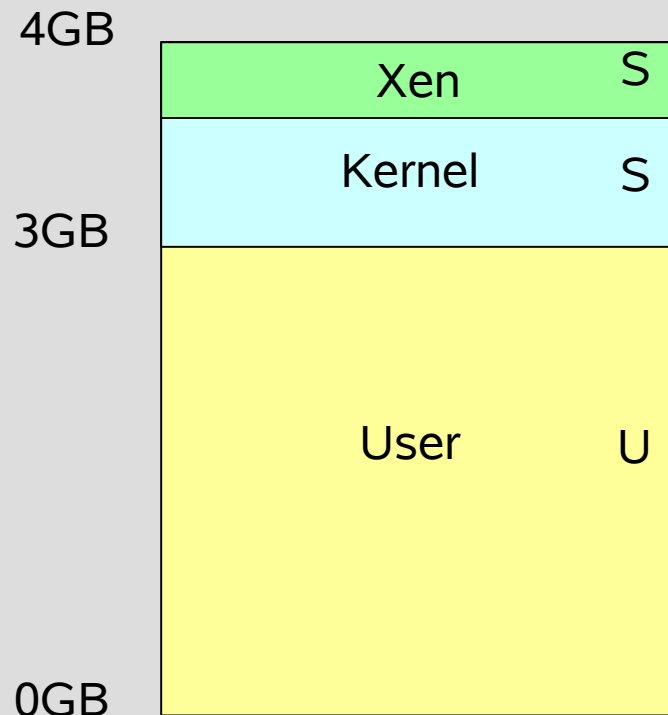
Architektura



Poziomy ochrony



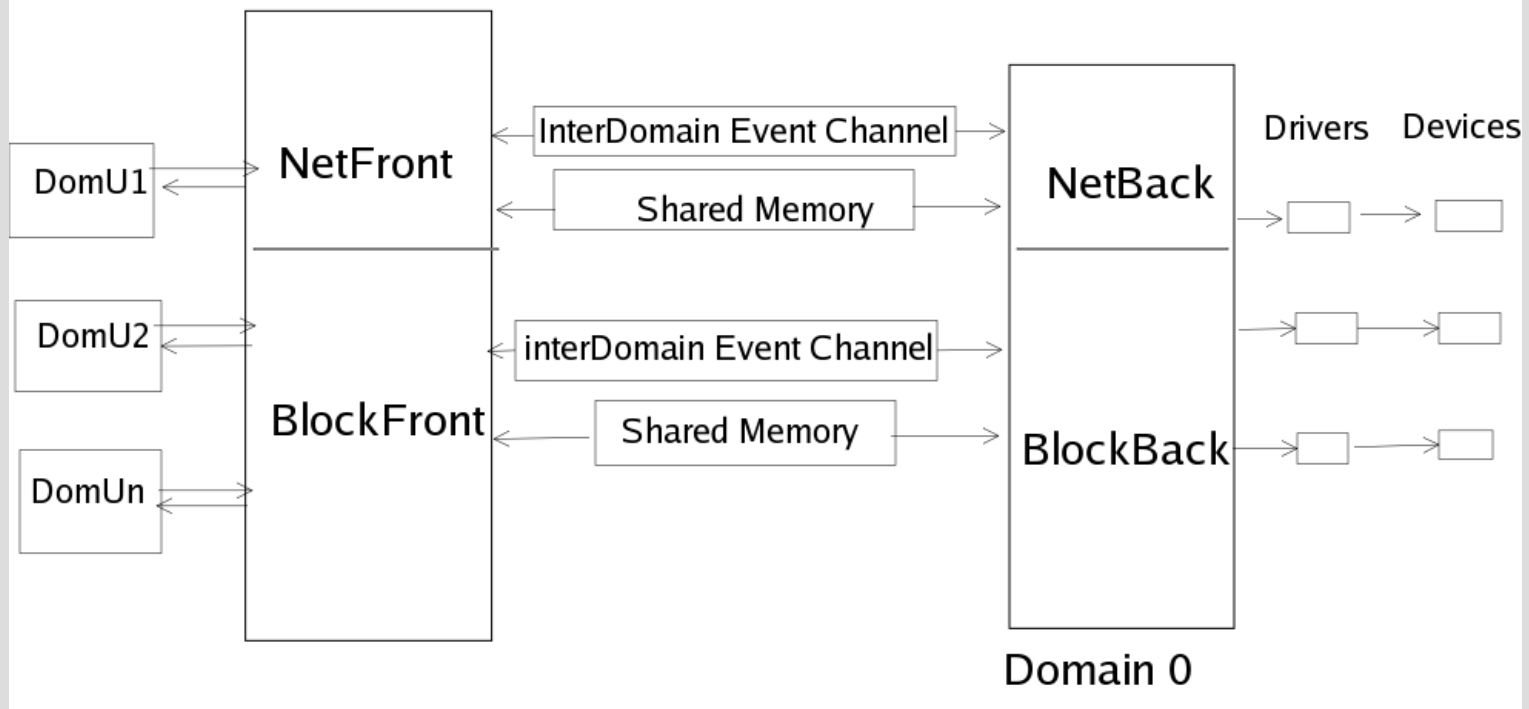
Zarządzanie pamięcią (x86)



- Xen rezerwuje górne 64MB przestrzeni adresowej, zapobiega czyszczeniu TLB przy wchodzeniu do nadzorcy
- segmentacja daje barierę bezpieczeństwa między Xenem a jądrem systemu operacyjnego
- stronicowanie – gość może otrzymywać (fizycznie) nieciągłe obszary pamięci, kolejkowanie i kontrola odwołań do tablicy stron

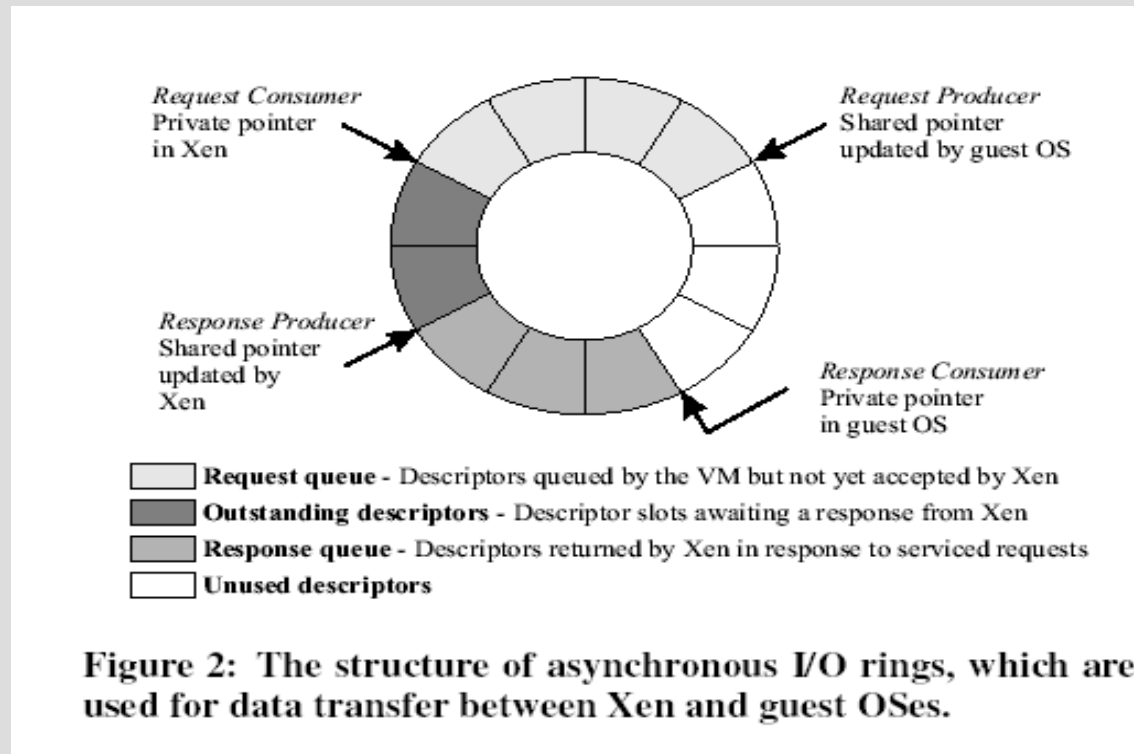
Sterowniki rozdzielone

Split Drivers Diagram



Operacje I/O

- transfer danych poprzez współdzielone, asynchroniczne bufory cykliczne deskryptorów
- mechanizm zdarzeń



Xen a czas

- liczniki czasu
 - cykli procesora
 - systemowy (liczony od uruchomienia)
 - absolutny
 - wirtualny czas domeny (liczony tak jak systemowy, ale aktualizowany tylko podczas faktycznego wykonywania danej domeny)
 - dostęp przez stronę współdzieloną
- aktualizacja czasu gościa co 10ms i przy szeregowaniu

Szeregowanie domen

- BVT
 - równomierny podział czasu oparty o wagi
 - wspiera zadania wrażliwe na opóźnienie
- SEDF (modyfikacja BVT) dostępny w wersjach unstable
- możliwość łatwej implementacji własnego algorytmu dzięki udostępnionemu API

Możliwości maszyn wirtualnych

- wstrzymywanie działania (z zapisem bieżącego stanu na dysk)
- przywracanie z zapisanego obrazu
- zmiana przydziału dostępnej pamięci w trakcie pracy
- migracja w trakcie pracy