

# Virtualizace



**Břetislav Bakala**

# Okolnosti využívání virtualizace

- Rostoucí význam **izolace a bezpečnosti** v moderních systémech
- Závady v zabezpečení a spolehlivosti standardních **operačních systémů**
- Sdílení jednoho počítače mezi mnoha nezávislými uživateli, **datová centra** nebo **cloudová řešení**
- Dramatické **zvýšení rychlosti** procesorů, které zpřístupňují **režii VM**
- Správa softwaru - VM poskytuje abstrakci může spustit **kompletní konkrétní SW vybavení**
- Správa hardwaru - **každá aplikace** vlastní **kompatibilní verzi** operačního systému v samostatných VM – **zvýšení spolehlivosti**
- **Podpora migrace** běžícího VM do jiného počítače

# Vlastnosti virtualizace

- **Úspora** energií, financí, síťové infrastruktury
- Možnost rychlé **migrace** - VM definován konfiguračním souborem a souborem virtuálního disku
- Vysoká dostupnost - Hyper-V failover cluster - replikace virtuálního stroje v režimu aktiv-pasiv ( VMware v Sphere **High Availability** )
- **Snadný upgrade** - rychlá změna výkonnostních parametrů VM
- Rychlé nasazení serveru - instalace probíhá **z předpřipraveného obrazu**

# Možnosti a principy virtualizace

- Nejširší definice VM zahrnuje **např. i platformu Java VM**
- VM **nejčastěji** podporuje **stejnou ISA jako fyzický hardware** (VMware ESX Server), může být však i odlišná.
- Vyžaduje **Monitor virtuálního počítače (VMM) – Hypervisor**
- VMM určuje, jak **mapovat** virtuální zdroje na fyzické
- Fyzický zdroj může být **sdílený časem, rozdělen** nebo dokonce **SW emulován**

# VM Monitor - Hypervisor

- **VMM** je mnohem **menší než tradiční operační systém** (např. pouze 10 000 řádků kódu)
- **Náklady na virtualizaci** procesoru závisí na **pracovní zátěži**
- **Nároky na I / O** spouští mnoho systémových volání a privilegované pokyny, **vedou k vysoké režii virtualizace.**

# Softwarová virtualizace

- Vyžaduje nainstalovaný operační systém hostitele
- OS hostitele snižuje výkon – vyžaduje systémové zdroje
- Např. Virtualbox – využití pro testování

# Hardwarová virtualizace

- **nevyžaduje operační systém hostitele**
- **Využívá VM monitor** – hypervizor jako vrstvu mezi HW a VM
- Technologie pro hardwarovou virtualizaci **Intel VT-x, AMD -V.**
- Nutno povolit v **BIOSu** (EFI).

# Požadavky na VMM

- Hostovaný SW by se měl chovat na VM přesně tak, **jako kdyby běžel na nativním hardware** s omezeními souvisejících se sdílením prostředků více VM.
- **Hostovaný SW by neměl** být schopen **změnit alokaci** skutečných **systémových zdrojů přímo**.
- **VMM řídí přístup k privilegovanému stavu, překlad adres, I / O, výjimky a přerušení** – během momentálního využívání hostujícími VM
- Např. v případě přerušení časovače VMM pozastaví aktuálně spuštěný VM, uloží jeho stav, obslouží přerušení, určí, který hostitelský VM bude pokračovat, a poté načte jeho stav.

# Požadavky na VMM II.

- **VMM musí být na úrovni vyšších oprávnění než hostovaný VM** (obvykle běží v uživatelském režimu)
- **VMM zpracovává jakoukoli privilegovanou instrukci**
- Požadavky jsou téměř shodné s požadavky na virtuální paměť:
  - Minimálně dva režimy procesoru, **systém a uživatel**.
  - **Privilegovaná podmnožina instrukcí, která je k dispozici pouze v režimu systému**, provedení v uživatelském režimu má za následek chybu. Všechny **systémové zdroje** musí být **kontrolovatelné** pouze prostřednictvím těchto instrukcí.

# Vliv VM na virtuální paměť

- **Hostující systém** využívá pouze virtuální prostředky, které běží v **uživatelském módu nad VMM**
- Zápis nebo čtení **do tabulky stránek** musí provádět pomocí **privilegované instrukce přes VMM**
- Každý hostující OS v každém **VM** spravuje **vlastní sadu tabulek stránek**.
- **VMM odděluje** virtuální paměť **jednotlivých VM** od fyzické pomocí **stínové tabulky stránek VMM**
- Položky tabulky stínových stránek VMM používají **hardware pro překlady**, odpovídající pozicím hostitelského operačního prostředí s nahrazením pro skutečné stránky v tabulkách hostů.

# Vliv VM na vstup / výstup

- Problémem je sdílení skutečného zařízení mezi více VM.
- **Každá VM** verze ovladače I / O přistupuje **k reálným V/V přes VMM**
- Mapování závisí na typu zařízení
- AMD-Vi a VT-d - **přemapování DMA a IRQ** někdy nazýváno PCI passthrough

Např. u virtuálních disků pro hostitelské VM jsou pomocí VMM mapovány virtuální stopy a sektory k fyzickým.

Síťové rozhraní jsou sdíleny mezi VM ve velmi krátkých časových intervalech, VMM spravuje adresy virtuální sítě.

# Hypervizor a rozhraní API

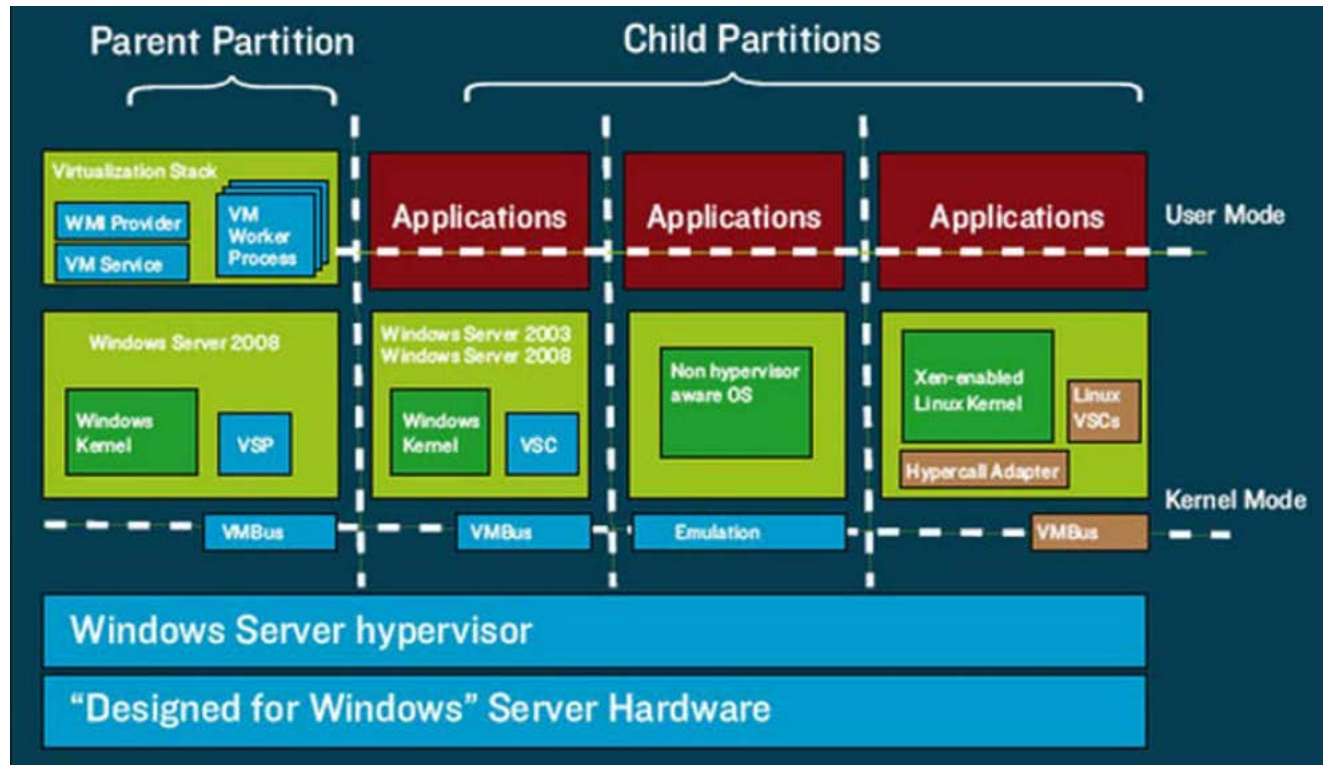
- Vrstva vložená mezi HW a OS, **vytváří jednotné rozhraní API**
- **ASPI** - Advanced SCSI Programming Interface
- **HAL** - Hardware abstraction layer
- **SR-IOV** (single root input/output virtualization) - umožňuje rozdílným virtuálním strojům sdílet stejný HW rozšiřující sběrnice PCIe
- **Execute Disable Bit** (Intel) NX bit (AMD) - oddělení části paměti pro procesorové instrukce a pro data – **ochrana přetečení**

# Hypervizor - SLAT

- Technologie SLAT - **Second Level Address Translation**:  
Intel - **Extended Page Table** - EPT,  
AMD - **Rapid Virtualization Indexing** - RVI - vnořené stránkování -  
virtualizační technologie s hardwarovým řízením, zabraňuje režii spojené  
se stínováním stránkovacích tabulek

# Hyper-V

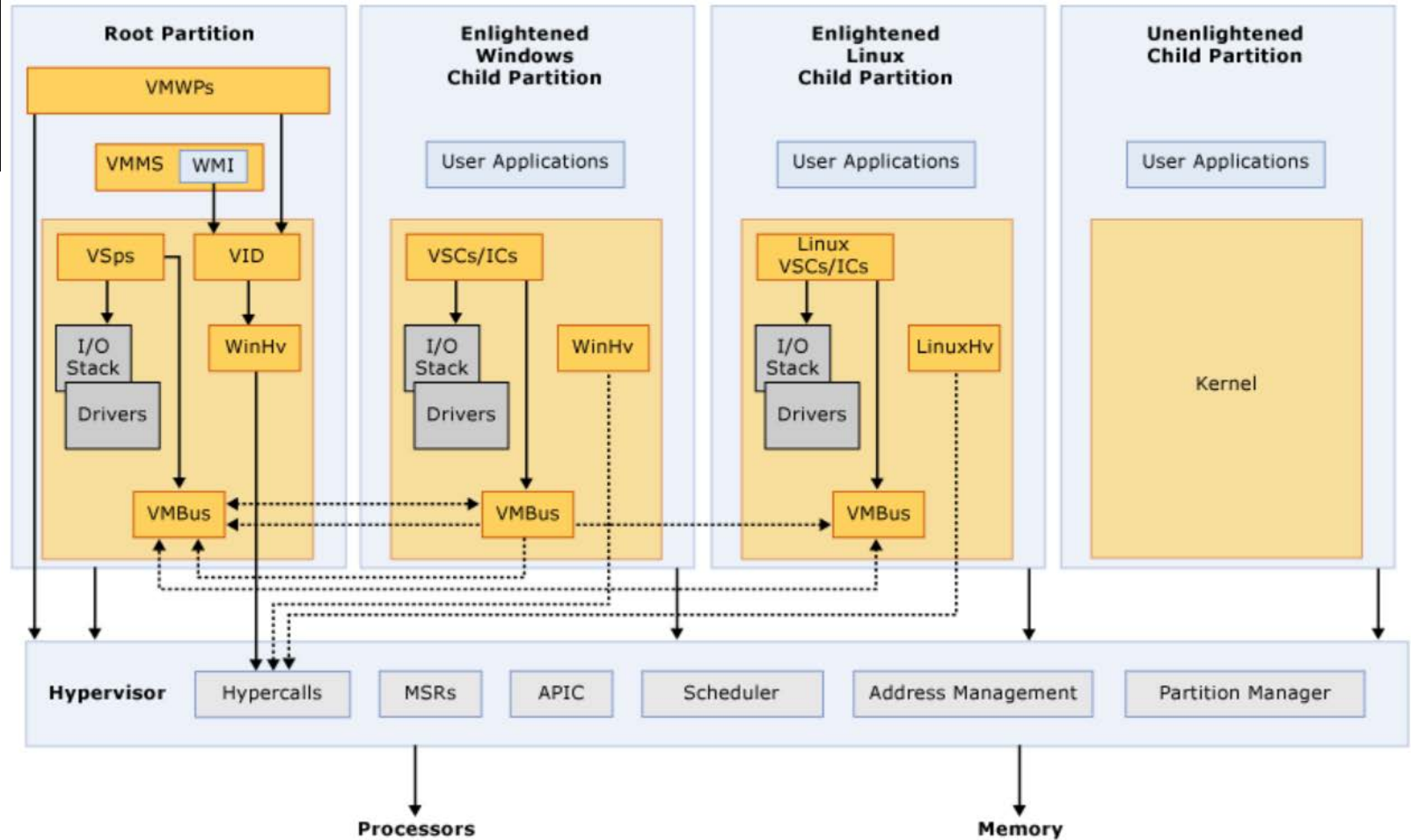
- původní OS se stává virtuálním OS - Parent partition, další VM jsou nazývány Child partition



# Hyper-V

- **Mikrokernel** - virtualizační platforma zabírá pouze několik MB – zajišťuje běh virtualizace (CPU, diskový prostor, RAM)
- má **výhradní právo přístupu k HW**
- parent partition a HW komunikují přes VM BUS
- Virtualization Service Provider – VSP
- Virtualization Service Clients – VSC
- IOMMU (Input Output Management Unit) pracuje nezávisle na hardwaru - přemapování adres fyzické paměti adresy, které používají podřízené oddíly.

## Hyper-V High Level Architecture



# Použité zkratky

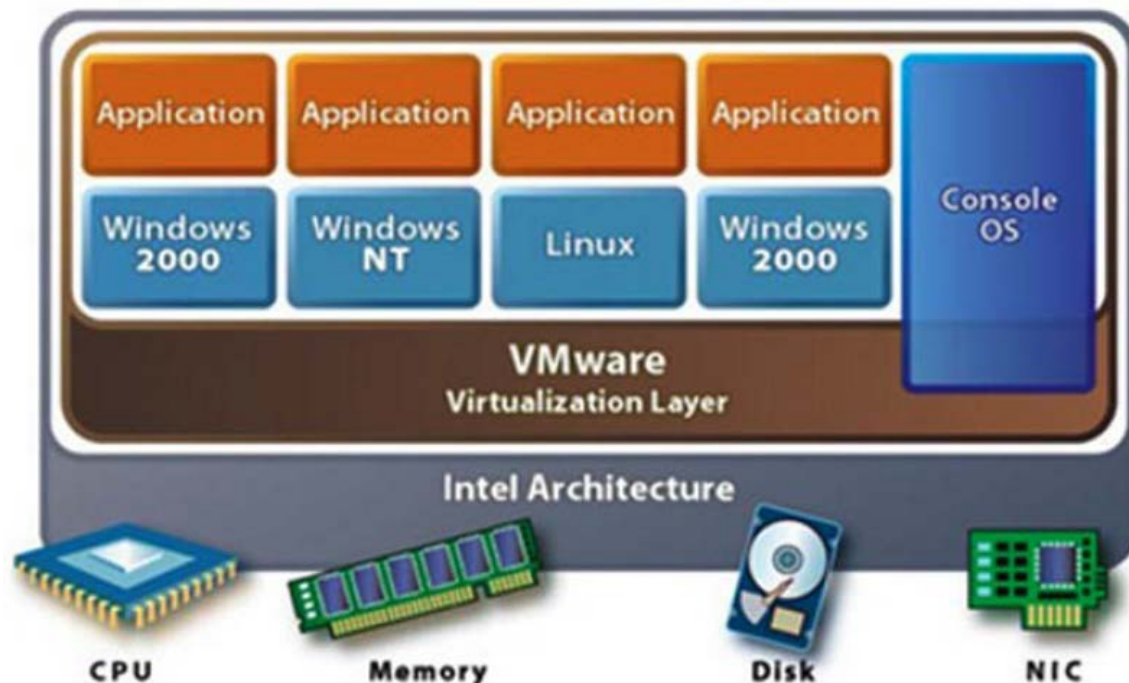
- **APIC** – Advanced Programmable Interrupt Controller
- **MSR** – Memory Service Routine
- **VID** – Virtualization Infrastructure Driver - služby správy diskových oddílů, služby správy virtuálních procesorů a služby správy paměti pro oddíly
- **VMMS** – Virtual Machine Management Service – řízení stavů VM
- **VMWP** – Virtual Machine Worker Process - poskytuje služby správy virtuálního stroje z instance systému Windows Server 2008 v nadřazené oblasti hostitelským operačním systémům v podřízených oddílech - pro každý spuštěný virtuální stroj samostatný pracovní proces
- **WinHv** – Windows Hypervisor Interface Library - můstek mezi ovladači operačního systému a hypervisorem - ovladačům volat hypervisor pomocí standardních konvencí volání systému Windows

# Licence Hyper-V

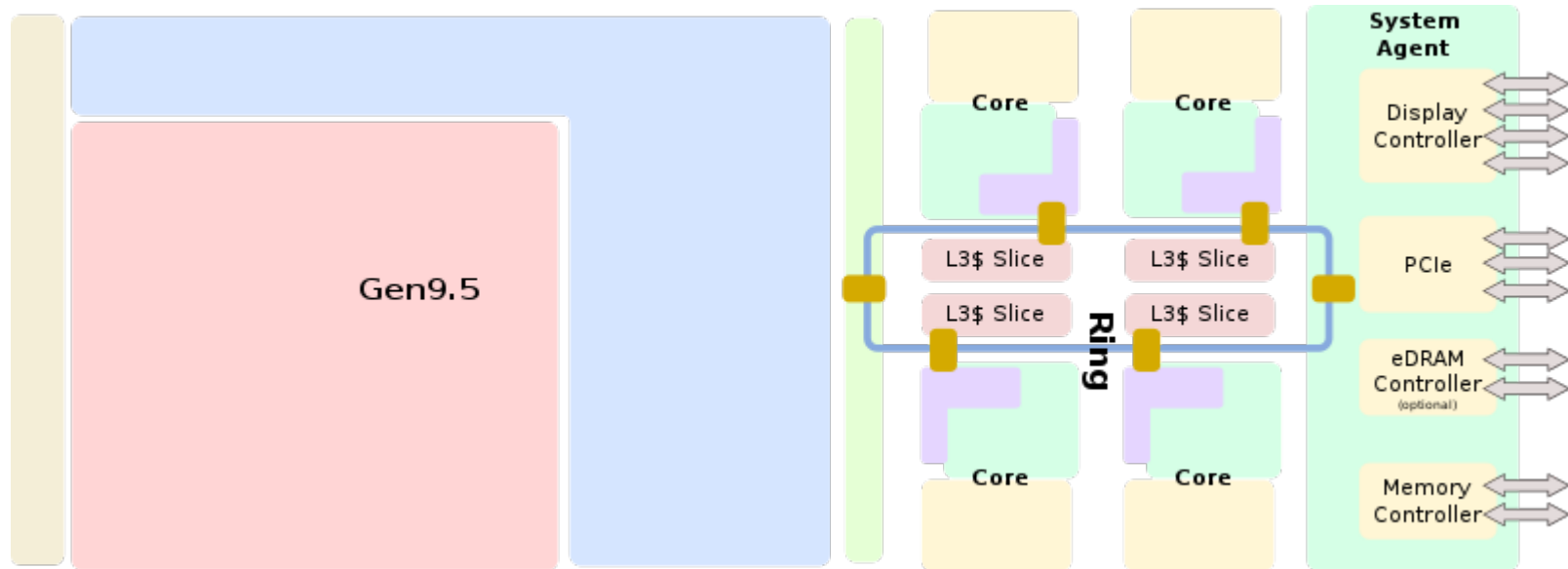
- licence standard Microsoft obsahuje jednu instalaci parent partition a 2 VM
- Pokud má parent partition další role stává se jednou licencí pro VM
- Instalace SW pro správu a zálohování nenárokují licenci VM
- Verze Datacenter má počet VM neomezen

# VMware

Každý virtuální stroj představuje kompletní systém s procesory, pamětí, připojením k síti, úložištěm a systémem BIOS včetně všech samostatných ovladačů.



# Vnitřní sběrnice procesoru



kaby lake soc block diagram

## skylake block diagram

