

Operační systémy

Jiří Zacpal



DEPARTMENT OF COMPUTER SCIENCE
PALACKÝ UNIVERSITY, OLOMOUC

KMI/YUDIT Úvod do informačních technologií

Osnova



- Co je to operační systém
- Proces
- Správa operační paměti
- Souborový systém
- Správa zařízení

Literatura

- Aleš Kepřt: Operační systémy
- Andrew Tanenbaum: Modern Operation Systems
- <http://phoenix.inf.upol.cz/~outrata/courses/udit/syllabus.html>

Operační systém

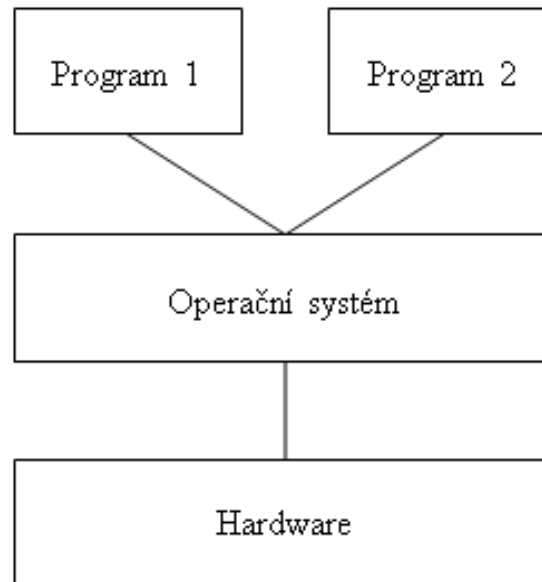
Co to je operační systém



- základní softwarové vybavení počítače - rozhraní mezi uživatelem a hardware počítače
- umožňuje programům (aplikacím) běh na/v počítači pomocí programového rozhraní (API) a uživatelům práci s počítačem pomocí svého uživatelského rozhraní (UI) a programů
- cíl: snadné a efektivní využití počítače (pro uživatele i aplikace)
 - víceméně protichůdné požadavky - dříve důraz na efektivitu (a vůbec možnost), nyní spíše snadnost
 - kompromis, závisí na způsobu využití a typu počítače -> různé OS
- poskytuje abstrakci (funkcí) hardware počítače, odstiňuje uživatele a aplikace od hardware

Operační systém

- **Virtuální počítač** – OS prostředník mezi hardwarem a aplikačními programy
- **Správa prostředků** – OS přiděluje a odebírá systémové prostředky procesům



- **jádro (kernel)** - vlastní OS,
- **základní obslužné programy** - pro práci s OS a zdroji počítače, např. administrátorské a diagnostické nástroje, diskové utility, programy pro práci se zařízeními, sítí apod.
- **uživatelské rozhraní (UI)** - součást OS (jádra?) nebo programy? -> záleží na typu OS, neinteraktivní (dávkové), interaktivní - textová konzole s interpretem příkazů (shell) nebo grafické s okenním systémem

- Jádro je základním kamenem operačního systému. Zavádí se do operační paměti počítače při startu a zůstává v činnosti po celou dobu běhu operačního systému.
- Druhy:
 - **Monolitické jádro** – jádro je jedním funkčním celkem. Tento typ jádra je použit v operačním systému UNIX.
 - **Mikrojádro** – jádro je velmi malé a všechny oddělitelné části pracují samostatně jako běžné procesy.
 - **Hybridní jádro** – kombinuje vlastnosti monolitického jádra i mikrojádra. Tento typ jádra je použit u operačních systémů Windows.

- **univerzální** - pro desktopové a přenosné počítače typu PC, servery, mainframe aj.
- **embedded** - specializované pro embedded zařízení, dnes i upravené univerzální (např. Linux, MS Windows)
- **reálného času** - zaručení vyřízení požadavku/odpovědi v pevně daném čase, např. VxWorks, QNX, upravené univerzální (např. RTLinux, MS WindowsCE) i HW řešení, např. pro řízení strojů
- **distribuované** - pro běh současné na více počítačích, simulace např. jedné společné paměti, pro počítačové klastry (cluster)
- dnes nejvíce používané: na desktopových PC MS Windows (majoritní), Mac OS X, GNU/Linux, Unix, na (síťových) serverech unixové, na embedded zařízeních různé (MS Windows, Linux)

- Linux je označení pro jádro operačního systému. Na počítači se setkáte s tzv. distribucí.
- Nejznámější distribuce:
 - **Debian** – Jedna z nejstarších distribucí
 - **Fedora** – Sponzoruje firma Red Hat, jejíž distribuce je na tomto systému založena.
 - **Slackware** – Jedna z prvních distribucí, vhodná spíše pro pokročilejší uživatele.
 - **Mandriva** (dříve Mandrakelinux) – Tato distribuce používá balíčkovací systém RPM.
 - **SUSE** – Původně samostatná distribuce, později koupena firmou Novell.
 - **Ubuntu** – Tato distribuce vhodná pro začátečníky vychází z Debianu.

- Operační systém (či správně operační systémy) Windows jsou asi nejznámější operační systémy.
- Druhy:
 - Windows pro DOS
 - 3.1x (1992)
 - Win 95, 98, Me
 - Windows NT (1993)
 - Win 2000, XP, 7,8,8.1
 - Windows CE
 - Windows Mobile, Windows Phone

- **program** = sekvence (binárních) kódů instrukcí, registrů procesoru a dat (čísla, texty, hodnoty adres do operační paměti a vstupně/výstupních zařízení)
- stejná (RISC) nebo proměnná (CISC) délka kódů instrukcí - 1 až 2 byty
- **operandy** = parametry instrukcí, registry a data, specický počet (obvykle 0 až 2), přípustné kombinace pro každou instrukci
- výsledek instrukce často ukládán do prvního operandu
- vykonávání instrukce
 - trvá určitý počet taktů (na vnitřní frekvenci procesoru), jednotky až desítky
 - až 7 fází: např. načtení, dekódování, načtení operandů, provedení, uložení výsledku
 - **pipelining** – částečně současné provádění instrukcí, fáze za sebou, nelze vždy, např. podmíněné skoky

- **Jazyk symbolických adres** („assembler“)

- jazyk (textově) pojmenovaných instrukcí, např. MOV, ADD, MUL, AND, CMP, JE, JMP, a registrů procesoru, (zápisů) čísel a textu, hodnot adres, proměnných atd.

MOV eax, promenna1; CMP ebx, promenna2; JE adresa

- překládán do kódu instrukcí
- přímá a nepřímá adresa do paměti - adresa vypočítána z hodnot v registrech a zadaných přímo, např. posunutí + báze + index x faktor, použití pro přístup do pole, k lokální proměnné apod.

- **Vyšší programovací jazyky**

- vyšší úroveň abstrakce, např. iterace přes prvky seznamu -> cyklus průchodu seznamem -> jména instrukcí procesoru (přesuny z/do paměti, log. operace, skoky aj.) -> kódy instrukcí
- **překladač** - přeloží (přepíše) program z jednoho (vyššího) prog. jazyka do jiného (nižšího) jazyka, typicky až do kódů instrukcí
- **interpret** - přeloží program z prog. jazyka do programových bloků interní formy a tyto vykoná

Přerušení (Interrupt)



- původně pro řešení komunikace (rychlého) procesoru s (pomalými) zařízeními:
 - dříve: vyslání požadavku, **aktivní čekání** na vyřízení (= smyčka testující stav oznamující vyřízení), pokračování ve výpočtu
 - dnes: vyslání požadavku, pokračování ve výpočtu zatímco zařízení zpracovává požadavek, oznámení vyřízení požadavku = **přerušení procesoru**
 - př. procesor vyšle požadavek čtení sektoru z disku (dá požadavek s číslem sektoru na sběrnici) a pokračuje ve výpočtu, disk najde sektor, načte do své cache a vyvolá přerušení, procesor vyšle požadavek zaslání bytu dat, disk pošle, procesor uloží do operační paměti, další byte atd.
- pozastavení vykonávání programu, vykonání programu (**rutiny**) **obsluhy přerušení** implementované OS (např. ovladači zařízení), pokračování vykonávání programu (od následující instrukce)
- během vykonávání obsluhy přerušení další přerušení zakázána nebo systém **priorit přerušení**

Přerušení (Interrupt)

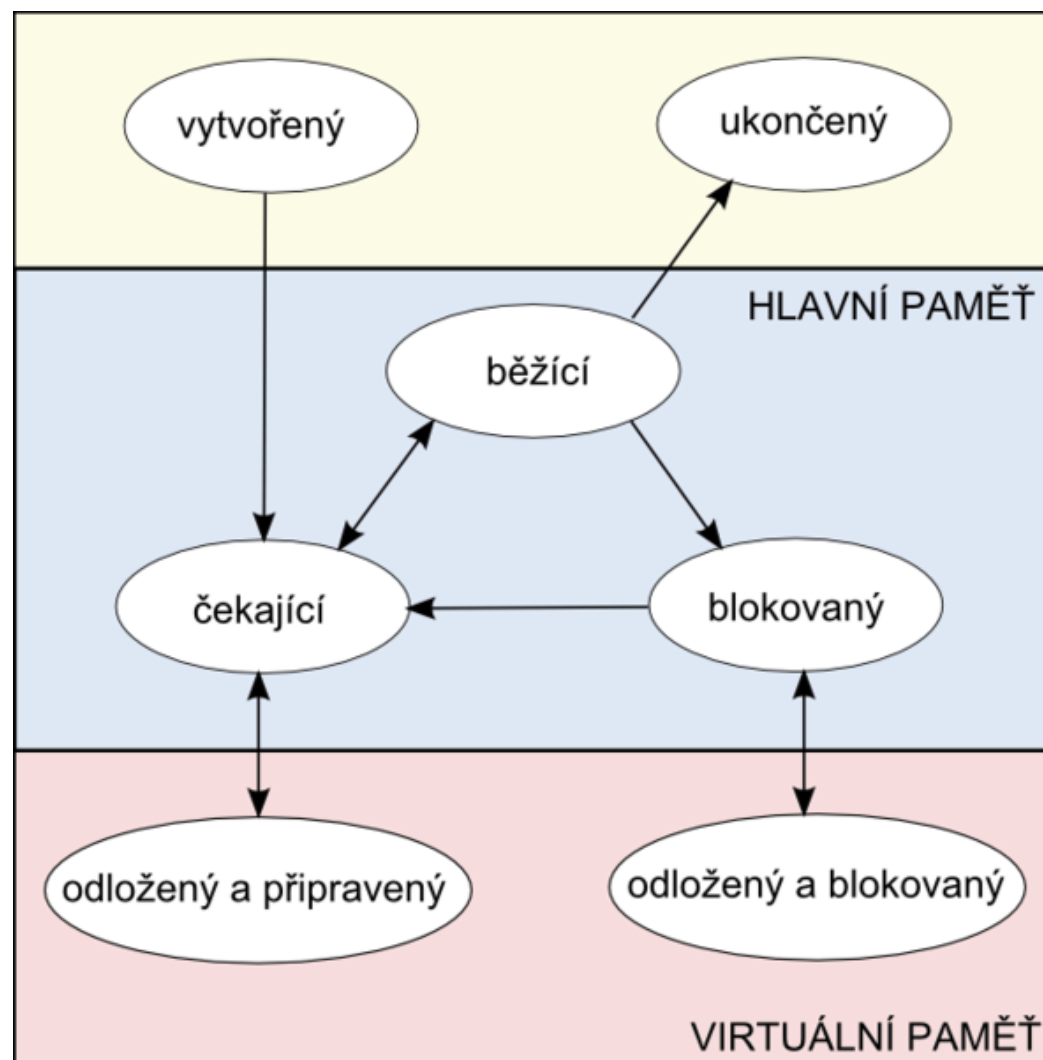


- **hardwarová** - přídavné karty (dříve), disková zařízení (dříve), vstupně/výstupní zařízení aj., 256 přerušení u Intel 80x86
- **softwarová** - vyvolána OS pro vlastní potřeby fungování, programy pro služby OS (tzv. systémová volání)
- DMA (Direct Memory Access) - způsob přenosu dat mezi zařízením a pamětí přímo, pro větší množství dat, např. disková zařízení, procesor pouze naprogramuje řadič DMA a vyšle prvotní požadavek, zbytek řeší řadič

Procesy

- **Proces** je spuštěný program
- **Vlákno** je prvek reprezentující vykonávání kódu procesu. Neformálně řečeno, vlákno je „vykonávání kódu“, zatímco proces je „paměť a další prostředky“.
- Stavy procesů:
 - **Vytvořený** (created)
 - **Běžící** (running)
 - **Připravený** (ready)
 - **Blokovaný** (blocked)
 - **Ukončený** (terminated)

Stavy procesů



- Plánování procesů řeší výběr, kterému následujícímu procesu bude přidělen procesor a proces tak poběží.
- Tuto úlohu má na starosti **plánovač (scheduler)**, který je součástí operačního systému.
- Plánovače používají dvě základní strategie:
 - **cyklickou obsluhu** - každému procesu je přiděleno stejné časové kvantum
 - **systém priorit** - každý proces má určitou prioritu s tím, že procesům s vyšší prioritou je přidělován procesor přednostně
- V praxi se používá kombinace obou strategií

Přepnutí kontextu



- Při přepínání procesů je nutné, aby proces po opětovném spuštění pokračoval od stejného místa, ve kterém byl přerušen a aby v procesu až na časové zpoždění nebylo poznat, že k přerušení došlo => **přepínání kontextu**
- **uložení kontextu** (anglicky context save) - kompletní uložení stavu procesu
- **obnovení kontextu** (anglicky context restore) – zpětné načtení stavu procesu do procesoru
- **změna kontextu** (anglicky context switch) - uložení kontextu jednoho a obnovení kontextu druhého

- pro procesy a pro vlákna
- procesy jsou paměťově oddělené (každý má svoji přidělenou paměť) - komunikace pomocí speciální **sdílené paměti** a **posíláním zpráv**
- vlákna sdílejí paměť procesu - komunikace pomocí sdílení paměti -> soupeření (race) o sdílenou paměť a jiné zdroje počítače, zvláště při více procesorech
- **chyba souběhu (race condition)** = chybné pořadí běhu vedoucí k nekonzistentním stavům při konfliktních operacích, např. čtení-zápis - fáze operací (i instrukcí procesoru) např. čtení z paměti, operace, zápis do paměti

- určení specifického pořadí běhu
- **atomické operace** = nedělitelné, nepřerušitelné, sekvenčně prováděné (ty konfliktní), např. nastavení nebo inkrementace atomické proměnné, použití pro implementaci tzv. synchronizačních primitiv
- **synchronizační primitiva**: zámek (mutex), semafor (počítadlo), kritická sekce, událost, monitor a další
- **hardwarová podpora**: atomické instrukce procesoru (test-and-set, fetch-and-add, compare-and-swap aj.), zakázání přerušování (při jednom procesoru), preempce (při více procesorech)
- **softwarové implementace**: Dekkerův (1965), Petersonův (1981) algoritmus, vyžadují atomické uložení hodnoty do proměnné
- implementovaná a poskytovaná OS, ale i využívaná v rámci samotného OS!

Uváznutí (deadlock)

- vzájemné čekání na výlučně vlastněné zdroje (např. chráněné zámky) při modelu využívání „požadavek na přivlastnění-používání-uvolnění“
- podmínky vzniku:
 1. výlučné vlastnictví,
 2. čekání při vlastnictví jiného,
 3. vzájemné (cyklické) čekání,
 4. nemožnost preempce (násilné odebrání prostředku)
- řešení deadlocku:
 - neřešení (ignorování),
 - detekce a zotavení,
 - prevence (zamezení vzniku, tj. nenaplnění podmínek),
 - vyhýbání se (přidělování prostředků tak, aby nenastaly podmínky)
- dnešní OS neřeší (ignorují)

Správa operační paměti

- S operační pamětí souvisí především tyto funkce, na kterých se podílí hardware a operační systém:
 - přidělování (alokace) a uvolňování (dealokace) paměti procesům na požádání
 - udržování informací o obsazení paměti
 - zabezpečení ochrany paměti – zabránění přístupu procesu k paměti mimo jeho přidělený region
 - realizace virtuální paměti

Přidělování operační paměti

- Operační systém může přidělovat procesům buď:
 - **souvislé úseky paměti**
 - přidělování pevných bloků paměti
 - přidělování bloků paměti proměnné velikosti
 - nebo malé bloky paměti – **stránky**
 - fyzický adresový prostor - **rámcce**
 - logický adresový prostor – **stránky**
 - Kdykoliv proces přistupuje do paměti, je logická adresa hardwarově přeložena na fyzickou adresu

- Operační paměť (zde se jí říká primární) rozšířena o místo na pevném disku (sekundární paměť), které je sice výrazně pomalejší, ale také výrazně lacinější
- Systém virtuální paměti funguje stejně jako stránkování, jen fyzický prostor je rozšířen o místo na disku
- Při přístupu procesu ke stránce, která není v primární paměti zajistit její nahrání do primární paměti.
- Pokud není v primární paměti volné místo, je nutno ještě nějakou stránku z primární paměti přemístit do sekundární. Tomuto procesu se říká **swapování**.

Souborový systém

- Úkolem operačního systému je především:
 - Umožnit procesům ukládat velké množství informací obvykle ve formě souborů.
 - Umožnit přístup k těmto informacím všem procesům, které o to požádají.
- Aby bylo možné tyto úkoly správně plnit, musí být data ve vnější paměti nějak organizována. Organizaci těchto dat se říká **souborový systém**.
- Nejčastější způsob organizace dat je pomocí hierarchického uspořádání **souborů** a **adresářů**.

Souborový systém



- Pevné disky jsou obvykle na fyzické úrovni rozděleny na **oddíly (partitions)**.
- Souborový systém se rozkládá jen na konkrétním oddílu a ne na celém disku.
- Operační systém vnímá jednotlivé oddíly jako tzv. **svazky (volumes)**, které se do systému připojují pomocí tzv. **mountování (mounting)**.

- **Soubor** je pojmenovaná posloupnost bytů, která je uložena na nějakém datovém médiu (pevný disk, CD disk, flash disk, ...)
- Každý soubor je charakterizován svým **jménem**.
- Část jména, která je za znakem „.“, se nazývá **přípona** názvu souboru.
- Většina operačních systémů podporuje různé druhy souborů. Nejzákladnější rozdělení je na:
 - **Normální soubory** – soubory, které obsahují uživatelské informace. Tyto se pak rozdělují na textové a binární soubory.
 - **Adresáře** – systémové soubory, které udržují strukturu souborového systému.

- Atributy - další informace o souboru:
 - délka – velikost souboru v počtech bajtů,
 - přístupová oprávnění – kdo smí se souborem pracovat (čtení, zápis, ...) pro uživatele a skupiny,
 - vlastník – uživatel vlastnící soubor (též skupina),
 - časové informace (čas vytvoření, čas posledního přístupu k souboru, čas poslední změny v obsahu souboru, ...).
- Tyto atributy jsou většinou uloženy v adresáři, ve kterém je soubor uložen

- **Adresář** je zvláštní druh souboru, který obsahuje množství záznamů
- Každý záznam se týká jednoho souboru (normálního nebo adresáře) a obsahuje jeho jméno, atributy a adresu, kde je na disku uložen.
- Některé operační systémy umožňují vytvořit jen jeden adresář (tento se nazývá **kořen** – root), který obsahuje všechny soubory
- Většina moderních operačních systémů však umožňuje vytvářet libovolné množství adresářů, které vytvářejí hierarchickou strukturu souborového systému.
- **Cesta** - určuje posloupnost adresářů od kořenového adresáře až k samotnému souboru

- V souborovém systému rozlišujeme dva druhy linků:
 - **Hard link** což je ukazatel na fyzické tělo souboru. Každý pojmenovaný soubor na disku je tedy vlastně hard link. Moderní operační systémy umožňují vytvářet více hard linků k jednomu souboru.
 - **Soft link** je odkaz na jiný soubor, který je specifikován jeho cestou.

- **FAT32** je jednoduchý souborový systém, proto je podporován prakticky všemi operačními systémy
 - FAT tabulka
 - kořenový adresář
- **NTFS** byl navržen jako souborový systém pro Windows NT
 - podpora pro přidělování práv k souborům
 - kompresi na úrovni souborového systému
 - šifrování
 - diskové kvóty umožňující nastavit maximálně využitelné místo na diskovém oddíle pro konkrétního uživatele
 - struktura
 - bootovací sektor
 - master file table (MFT), což je seznam všech souborů na disku
 - systémové soubory
 - ostatní soubory.
- **UFS (Unix File Systém)** je souborový systém používaný v Unixu

Správa zařízení

- Jednou ze základních funkcí operačního systému je správa vstupně výstupních zařízení
- Vstupní zařízení se dělí na:
 - **Bloková** (například disky). V těchto zařízeních se data ukládají v blocích (nejčastější velikost bloku je od 128 B do 1 024 B), které mají svou adresu. Základní vlastností těchto zařízení je to, že umožňují čtení či zápis celého bloku.
 - **Znaková** (například klávesnice, myši, tiskárny). Tato zařízení pracují s proudy znaků, které nemají žádnou adresu. Data se tedy čtou a jsou i ukládána po znacích.

- Úkolem operačního systému při správě zařízení je především:
 - **Zajistit nezávislost na zařízení.** To znamená, že například můžeme napsat program, který vytváří a ukládá soubory na libovolná zařízení (disk, flash disk, ...) a pro každé z těchto zařízení není potřeba tento program měnit.
 - **Zajistit ošetření chyb zařízení.** Při čtení nebo zápisu na zařízení se může vyskytnout chyba. Operační systém by měl tuto chybu odstranit, pokud to jde, a pokud ne, měl by tuto chybu ošetřit (pokusit se o danou akci znovu, ohlásit chybu, ...).
 - **Zajistit sdílení zařízení.** Každé zařízení může využívat více uživatelů v tu samou chvíli.
- Tyto cíle zajišťuje operační systém pomocí **ovladačů zařízení**

- Disky jsou vstupně-výstupní zařízení, které slouží pro zápis a čtení dat, obvykle ve formě souborů
- Nevýhodou oproti operační paměti je větší časová náročnost při čtení i ukládání dat. Čas pro tyto operace se skládá z:
 - hledání – přesunu ramena nad odpovídající cylindr,
 - rotace – čas pro najetí správného sektoru pod čtecí hlavu,
 - přenos dat – čtení sektoru.

- RAID je metoda zabezpečení dat proti selhání pevného disku
- Zabezpečení je realizováno specifickým ukládáním dat na více nezávislých disků, kdy jsou uložená data zachována i při selhání některého z nich
- Metody:
 - Pruhování
 - Zrcadlení
- Druhy: RAID 0, RAID 1, RAID 5, RAID 6

Podrobnější informace

- J. Hronek: Struktura počítačů
- P. Tišnovský: Seriál Co se děje v počítači (<http://www.root.cz/serialy/co-se-deje-v-pocitaci/>)
- <http://phoenix.inf.upol.cz/~outrata/courses/udit/syllabus.html>

- Počítačové sítě, technologie a principy fungování. Celosvětová síť Internet a její služby.
- **Studijní texty:**
 - P. Příhoda: Počítačové sítě
 - uvt_3_pocitacove_site.pdf (Vyuka\KMI_UVT\vyukovy_text)
 - Andrew S. Tanenbauma Computer Networks.
 - Jiří Peterka: Báječný svět počítačových sítí (http://www.earchiv.cz/i_serial.php3)