Počítačové sítě

přednášky

Jan Outrata

říjen–prosinec 2010 (aktualizace září–prosinec 2013)

Tyto slajdy byly jako výukové a studijní materiály vytvořeny za podpory grantu FRVŠ 1358/2010/F1a.

Použitá a doporučená literatura

- Kabelová A., Dostálek L.: *Velký průvodce protokoly TCP/IP a systémem DNS (5. vydání)*. Computer Press, 2008. ISBN 978-80-251-2236-5
- Kállay F., Peniak P.: *Počítačové sítě LAN/MAN/WAN a jejich aplikace (2. vydání)*. Grada, 2003. ISBN 80-247-0545-1
- Pužmanová R.: *Moderní komunikační sítě od A do Z (2. aktualizované vydání)*. Computer Press, 2006. ISBN 8025112780
- Trulove J.: *Sítě LAN - hardware, instalace a zapojení*. Grada, 2009. ISBN 978-80-247-2098-2
- Zandl P.: *Bezdrátové sítě WiFi: Praktický průvodce*. Computer Press, 2003. ISBN 80-722-6632
- Tanenbaum A. S., Wetherall D. J: *Computer Networks (5th edition)*. Prentice Hall, 2010. ISBN 978-0132126953
- Forouzan B.: *TCP/IP Protocol Suite*. McGraw-Hill Science/Engineering/Math, 2009. ISBN 978-0073376042
- Archiv článků a přednášek Jiřího Peterky
- *Dokumenty RFC (Request For Comments)*, RFC Editor

Úvod

Úvod

- propojování počítačů nevyhnutelné – přístup (v reálném čase) k informacím na jednom místě z více míst, ideálně odkudkoliv
- **komunikační sítě**
  - dříve zvlášť **telekomunikační** (telefon, rádio), zábava (rádio, televize) a **datové** (počítačové sítě)
  - dnes hlas i obraz jako data (**digitalizace**) v telekomunikačních sítích a telekomunikační služby v datových sítích (Internet) $\rightarrow$ **konvergence**
- **počítačová síť** = skupina vzájemně propojených počítačů a dalších zařízení (**hostitelských/koncových uzlů**), komunikujících pomocí prvků **síťové infrastruktury**:
  - **přenosová/propojovací média**: metalické vodiče a optická vlákna = „drát“, elektromagnetické (rádiové) vlny = „bezdrát“
  - aktivní a pasivní **propojovací prvky**: opakovače, přepínače, směrovače, brány aj.
- **síťové prostředky (zdroje)**: SW a HW prostředky a služby poskytované hostitelskými uzly skrze síť

Historie počítačových sítí

- za posledních X desítek let neustálý nárůst objemu a komplexnosti informací ve formě dat
- od papíru k (přenosným) datovým médiím, s růstem počtu počítačů sílící požadavek na **výměnu dat**
- 50. léta – přenos dat mezi izolovanými počítači na samotných datových médiích („offline“), lokální využití počítačů
- od konce 50. let – **propojování** počítačů (drátovými a později bezdrátovými médii), data na jednom místě, přístup a výměna z jiných míst v reálném čase („online“), vzdálené využití počítačů

$\Rightarrow$ nutnost řešit **komunikaci mezi počítači (uzly)** $\rightarrow$ vývoj způsobů propojení a komunikace

Historie počítačových sítí: způsoby propojení

Na lokální úrovni:

- **dvoubodové spoje** – přímé propojení dvou počítačů (typicky přes HW porty), konec 50. let
- **terminálové sítě** – počítače jako vstupně/výstupní HW terminály připojené k hlavnímu počítači (mainframe), 60. léta, později SW emulátory terminálů na počítačích
- **lokální sítě** – propojení více (osobních) počítačů, od 70. let, různé **topologie**:
  - **polygonální** – dvoubodové spoje každý s každým $\Rightarrow$ velká spotřeba propojovacích médií
  - **sběrnicová** – minimum propojovacích médií, počítače napojené na **sběrnici** = sdílené přenosové médium $\Rightarrow$ vytížení sběrnice, varianta **kruhová** = uzavřená drátová sběrnice, prostorové a časové využití (přenosové kapacity) sběrnice
  - **hvězda**, **strom** – sběrnice $\rightsquigarrow$ propojovací prvek, dvoubodové spoje s počítači
- různá firemní (proprietární) řešení lokálních sítí – 80. léta, navzájem nekompatibilní $\rightarrow$ nutná **standardizace**

Historie počítačových sítí: způsoby propojení

Na globální úrovni:

- využití **telekomunikačních sítí** – oddělení přenosové (propojovací) části sítě od koncových zařízení (lokálních sítí), propojování lokálních sítí do **rozlehlých sítí**, od 60. let
- **globální sítě** - decentralizované a distribuované, od 70. let
- různé firemní (proprietární) sítě (ARPANET, CYBERNET, EIN) – vedle veřejných telekomunikačních (DATEX, EDS, TELENT), 70. léta
- ARPANET $\rightsquigarrow$ dominantní veřejná síť **Internet**, od 80. let

Historie počítačových sítí: způsoby komunikace

- přepojování fyzických **okruhů** (i komutovaných) – pronájem **komunikačního kanálu** = části přenosového média, podobně jako v telekomunikační síti, 50. léta
- přepojování (přenos) **zpráv** = celistvých dat – princip telegramu, ne v reálnem čase, 60. léta
- přepojování (přenos) **paketů** = „kousků zpráv“ – v reálném čase, řešení spolehlivosti přenosu, konec 60. let, 70. léta $\rightarrow$ **paketové sítě**
- „nespolehlivé“ (Internet) i „spolehlivé“ (X.25) paketové sítě

Konvergence sítí

= sbližování/využívání odlišných komunikačních technologií, telekomunikačních s hlasem a obrazem (přepojované sítě) a datových (paketové sítě)

- konvergentní telekomunikační sítě = integrace datových služeb (paketového přenosu dat) do telekomunikační sítě – přístup k Internetu, audio a video přenosy, „datová komunikace“, např. ISDN, GPRS
- konvergentní datové sítě = implementace telekomunikačních služeb v datové síti (Internetu), pomocné technologie pro garantovaný přenos (multimediálních) dat (hlasu a obrazu), např. streaming, virtuální telefonní ústředny

Klasifikace sítí

- podle různých kritérií: **rozlehlost**, rychlost přenosu (klasické a vysokorychlostní), forma aplikace aj.

Lokální (LAN, Local Area Network)

- propojení koncových uzlů s umožněním vzájemné komunikace a přenosu dat
- lokální = omezeny rozsahem (jednotky km, nejčastěji v budově nebo komplexu budov), v soukromé správě
- klasické přenosové rychlosti od 10 Mb/s do 1 Gb/s
- sdílené využití přenosového média
- př. Ethernet (10, 100 Mb, 1 Gb), Wi-Fi (jednotky až desítky Mb/s)
- dnes i virtuální

Klasifikace sítí

Metropolitní (MAN, Metropolitan Area Network)

- propojení a „prodloužení“ několika LAN, účelem přenosové sítě, charakterem lokální
- v rámci města (desítky km), soukromé i veřejné
- vyšší (několik Gb/s) i nižší ($<$ 1 Mb/s) rychlosti ve srovnání s LAN
- př. Ethernet (10 Gb), Wi-Fi (jednotky Mb/s)

Klasifikace sítí

Rozhlehlé (WAN, Wide Area Network)

- přenosové sítě propojující LAN/MAN (**páteřní sítě**, telekomunikační sítě – **broadband**)
- pro LAN má význam jen rozhraní přístupu k síti, zbytek „černá skříňka“
- velké vzdálenosti, pokrývají území států a kontinentů (neomezené), veřejné i soukromé (vlastní nebo pronájem kapacity)
- zpravidla vysoké přenosové rychlosti (desítky až stovky Gb/s), ale i nízké (desítky kb/s)
- (prostorově a časově) vyhrazené nesdílené využití přenosového média = pronájem kapacity sítě
- př. GPRS (desítky kb/s), xDSL (desítky Mb/s), Frame Relay, ATM (stovky Mb/s), DWDM (desítky až stovky Gb/s)

Klasifikace sítí

Personální (PAN, Personal Area Network)

- propojení zařízení, příp. k počítači, s umožněním vzájemné komunikace a přenosu dat, charakterem LAN
- omezeny dosahem, v okolí zařízení (jednotky až desítky m, nejčastěji „kolem osoby“), v soukromé správě
- nízké přenosové rychlosti (stovky kb/s)
- př. Bluetooth (stovky kb/s)

Klasifikace sítí

Z aplikačního hlediska:

- v **informačních systémech** jako komunikační subsystém s aplikačními službami pro poskytování a sdílení HW i SW prostředků a umožnění přenosu dat
- v **průmyslových aplikacích** jako komunikační systém pro řízení a automatizaci výroby (procesní úroveň), propojení a koordinace strojů (technologická úroveň) a napojení na informační systém (dispečerská úroveň)

Aplikace (v oblasti informačních systémů)

Počítačová síť (z pohledu informačního systému) = integrující prostředí pro vzájemné propojení komunikujících heterogenních prvků a systémů v rámci informačního systému

Vývoj informačních systémů kopíruje vývoj sítí:

- lokálně na 1 počítači (mainframe), s dávkovým zpracováním úloh, 50. léta
- **CIS** = centralizované informační systémy – v terminálových sítích, s interaktivním zpracováním dat, 60. léta
- **DIS** = distribuované informační systémy – lokální sítě s (osobními) počítači, se **souborovými servery** (downsizing), 70. až 80. léta
- architektury **klient-server**, distribuované zpracování s výkonem CIS, vznik dnešních informačních systémů (upsizing), od konce 80. let
- kombinace s počítači všech tříd (rightsizing)

Aplikace (v oblasti informačních systémů)

Služby poskytované (zejména rozlehlou) sítí, na aplikační úrovni:

- připojení k síti
- vzdálený přístup, sdílení výpočetních prostředků a přenos dat (sdílené soubory, databáze, peer-to-peer sítě)
- sdílení technických prostředků (tiskárny, disky, faxy, multimediální apod.)
- adresářové služby (jednotný přístup do informačního systému a k informacím z centrální databáze, např. LDAP, Active Directory)
- elektronická pošta a výměna dokumentů (služba EDI, objednávky, faktury)
- online komunikace/multimedia (např. ICQ apod., IRC, VoIP, VoD, video konference, streaming, hry) – vysoké nároky na síť
- informační služby, internetové aplikace (WWW, business a desktopové aplikace)
- monitorování a vzdálená administrace sítě (management, např. SNMP)
- ...

Aplikace (v oblasti informačních systémů)

Komunikace uzlů a propojovacích prvků sítě na různých úrovních:

- nižší – přenos bloků dat, (většinou) „nespolehlivý“ (bez potvrzení a opakování přenosu), založeno na cílové adrese (nespojová komunikace):
  - **unicast** = dvoubodová, základní
  - **multicast** = bod-skupina, např. streaming multimédií, virtuální sítě
  - **broadcast** = bod-všichni, např. konfigurace a zapojení do sítě
- vyšší – komunikace aplikací, (většinou) „spolehlivá“ (s potvrzením doručení a příp. opakováním), spojově orientovaná (vytvořeno „spojení“ mezi aplikacemi):
  - **peer-to-peer** = zpravidla rovnocenná výměna dat
  - **klient-server** = hierarchická, forma požadavek-odpověď, charakter nestavový i stavový (komunikace je v různých stavech)

Aplikace (v oblasti informačních systémů)

Typy koncových uzlů (počítačů) v síti:

- **pracovní stanice** (work station, klient)
  - převážně využívá služeb sítě
  - **tenký klient** = znakový/grafický HW terminál – pouze zprostředkování vstupu a výstupu pro vzdálený uzel (server), nemůže fungovat samostatně
  - **tlustý klient** = osobní počítač – i lokální úlohy, klientské části síťových služeb, může fungovat i samostatně (do určité míry)

Aplikace (v oblasti informačních systémů)

Typy koncových uzlů (počítačů) v síti:

- **server**, převážně poskytuje služby v síti, peer-to-peer nebo **dedicated**, nosné, pomocné apod.
  - souborový (FTP, NFS, SMB/CIFS) – operace se soubory, transparentní přístup k souborům po síti
  - databázový/adresářový (SŘBD/DBS, LDAP, AD) – strukturovaná data, prohledávání, adresáře uživatelských aj. účtů
  - poštovní (SMTP, POP3, IMAP) - přenos el. zpráv (emailů)
  - prezentační/terminálový (Telnet, SSH, VNC, Windows Terminal Server/RDC, Citrix Meta Frame/ICA)
  - informační/WWW (HTTP) – hypertextové stránky, dnes i aplikace
  - komunikační/multimediální – IM, VoIP, VoD, streaming
  - aplikační/výpočetní (RPC, DCOM/DDE, J2EE/SOAP) – spolupráce s databázovými a prezentačními servery

  - infrastrukturní – jmenné, přístupové, modemové, směrovače, brány aj.
  - tiskový – síťové tiskárny s tiskovou frontou
  - ...

Aplikace (v oblasti informačních systémů)

Více viz **informační systémy** (architektury host-terminal, file-server, client-server, intranet) a **multimediální systémy** (VoIP, VoD, konferenční služby, rezervace šířky pásma, prioritní řízení toku, časová synchronizace přenosu).

Síťové architektury

Síťová architektura

- snaha o vytvoření univerzálního konceptu sítě – topologie, formy a pravidla komunikace, poskytované služby atd.
- vytvářely (a vytvářejí) souběžně, ale nezávisle firmy (IBM), (telekomunikační) organizace, **normalizační instituce** (ITU-T, ISO, IEEE, IEC, ANSI, IETF a další (ČSNI)) a **průmyslová konsorcia** (GEA, WLANA aj.) $\rightarrow$ nekompatibilní řešení
- **požadavky**: decentralizace služeb, rozumná adresace uzlů, navazování spojení mezi uzly, data zasílána v nezávislých blocích, směrování bloků, zabezpečení, kontrola a řízení přenosu, aj.
- dříve proprietární uzavřená řešení, následně standardizace s koncepcí komunikace nezávisle na implementaci (výrobci zařízení)

$\rightarrow$ komunikace ve **vrstvách**:

  - definovaných službami poskytovanými (sousedním) vyšším vrstvám a využívajících služeb (sousedních) nižších vrstev, implementace skryté před okolními vrstvami
  - samostatné, s funkcemi podobnými v rámci vrstvy a odlišnými v různých vrstvách, nezávislé na implementaci

Síťová architektura

- komunikace mezi vrstvami (svislý směr) pomocí **mezivrstvových protokolů** – na každé komunikující straně zvlášť, skrze **programová rozhraní**, prostřednictvím přístupových bodů, využívajících tzv. služební primitiva, fyzická, př. komunikace člověka s překladatelem

Obrázek: Obrázek průvodce 2→16(5)

- obecná **služební primitiva** (druhé a poslední nepovinná):
  - žádost o službu (**request**)
  - oznámení poskytovatele o přijetí žádosti (**indication**)
  - odezva poskytovatele (**response**), příp. vytvoření spojení
  - potvrzení odezvy žadatelem (**confirmation**)
- komunikace mezi entitami (zařízeními) ve stejnolehlých vrstvách (vodorovný směr) pomocí **vrstvových protokolů** – entity z různých komunikujících stran, implementace služebních primitiv, fyzická na nejnižší vrstvě, jinak virtuální (zprostředkovaná nižšími vrstvami), př. komunikace cizinců

Síťová architektura

Protokol = souhrn pravidel (**norem** a **doporučení**) a procedur pro komunikaci (výměnu dat), synt. a sem. pravidla výměny protokolových datových jednotek

- **protokolové datové jednotky** = **režijní informace** a data, např. rámce, pakety, segmenty
- komunikace zprostředkovaná sousední nižší vrstvou
- na straně odesílatele od nejvyšší po nejnižší vrstvu **„zapouzdřování“** dat do protokolových jednotek, na straně příjemce v opačném směru **„rozbalování“** dat, př.
- pro komunikaci na jedné vrstvě je možné použít více různých protokolů na sousední nižší vrstvě
- protokol může garantovat příjem dat v pořadí odeslaní (typicky u spojovaných, spolehlivých služeb), ale také nemusí (typicky u nespojovaných, nespolehlivých služeb, přeskládání do správného pořadí řeší vyšší vrstva)
- vydávají normalizační instituce a průmyslová konsorcia, některé jsou zdarma (RFC, RIPE)

Síťová architektura

Síťová (protokolová) architektura = definice vrstev, služeb, funkcí, protokolů a forem komunikace

- **normalizované** de jure (normy OSI) i de facto (TCP/IP, doporučení a normy RFC)
- firemní proprietární (Novell NetWare, Apple Appletalk, Microsoft NetBEUI a SMB aj.)

Abstraktní referenční síťový model architektur od ISO

= **abstrakce** konkrétních síťových architektur, reference pro nové

- architektury nemusí podporovat všechny funkce modelu (např. průmyslové sítě nepodporují směrování, sítě jsou propojeny pomocí mostů a bran)

Referenční model ISO OSI (Open Systems Interconnection)

- **propojení otevřených systémů** = zařízení podporujících příslušné normy
- obecně platné principy implementace systémů (abstrakce síťové architektury), pozn. existuje i konkrétní architektura OSI s konktrétními protokoly!
- norma ISO IS 7498, 1979, referenční model ITU X.200, 1984
- definuje **koncové uzly** (**koncová datové zařízení, DTE**) a **mezilehlé uzly** zprostředkovávající komunikaci (**propojovací prvky, DCE**)
- vrstvy: fyzická, linková, síťová, transportní, relační, prezentační a aplikační

Obrázek: Obrázek sítě 32

RM OSI – Fyzická vrstva

- způsoby fyzické komunikace, **přenos sledu signálů** (bitů nebo skupin bitů) mezi přímo propojenými zařízeními, **bez ohledu na význam** bitů
- přenosové cesty elektrické, optické, drátové, bezdrátové
- komunikující zařízení na **fyzickém** nebo **virtuálním okruhu** (pevný nebo komutovaný)
- funkce a služby:
  - správa fyzických spojení a okruhů mezi DTE a DCE, identifikace okruhů
  - seřazování bitů (stejné na vstupu i výstupu)
  - udržování **parametrů** (přenosová rychlost, doba, ztráta) a oznamování poruch
- protokoly specifikující bity jako signály (kódování 0 a 1), tvary konektorů, typy médií (kroucená dvojlinka, optické vlákno, mikrovlny), přenosovou rychlost a jiné parametry apod.
- protokoly př. V.24/RS 232, **EIA/TIA 568A/B**, WiFi/Bluetooth, ISDN, DSL, vydávají organizace ITU-T, EIA/TIA aj.
- HW zařízení (nejsou součástí modelu) př. fyzické rozhraní síťové

RM OSI – Linková vrstva

- (dynamické) zajištění **výměny dat mezi sousedními zařízeními** (DTE) = **v dosahu protokolu** (v MAN/WAN nebo v rámci LAN), bity mají význam (data)
- zařízení má jednu **linkovou adresu**

Obrázek: Obrázek průvodce 4→21(5)

- jednotka přenosu = **datový rámec**: **záhlaví** s linkovou adresou příjemnce a odesílatele (př. MAC u Ethernetu) + data + **zápatí** s kontrolním součtem (CRC) **celého rámce**, přenášen fyzickou cestou
- funkce a služby:
  - správa linkových spojení, řízení fyzických okruhů, identifikace zařízení
  - formátování rámců
  - oznamování (neopravitelných) chyb, detekce a oprava chyb
- protokoly př. **Ethernet**, **WiFi**, Bluetooth, PPP/DSL, SLIP, ISDN, Frame Relay, FDDI aj.
- HW zařízení př. síťová karta/adaptér, přepínač, most, přístupový bod

RM OSI – Síťová vrstva

- zajišťuje **přenos dat mezi vzdálenými, nesousedními zařízeními** v různých sítích spojených do jedné rozsáhlé sítě (př. WAN, Internet)
- zařízení může mít více jednoznačných **síťových adres**

Obrázek: Obrázek průvodce 5→22(5)

- jednotka přenosu = **síťový paket**: **záhlaví** se síťovou adresou příjemce a odesílatele (např. IP u Internetu) + data + zápatí jen vyjímečně, přenášen v datovém rámci (datové části)
- funkce:
  - **abstrakce** různých linkových technologií
  - správa linkových spojení, **multiplexování** síťových spojení do linkových
  - formátování dat do paketů
  - **směrování** paketů
  - zjišťování a oprava chyb
  - vytváření **podsítí**

RM OSI – Síťová vrstva

- služby:
  - síťové adresování
  - správa síťových spojení
  - převod transportních paketů (datagramů) na síťové pakety
  - oznamování chyb, řízení toku dat
- přenos dat **se spojením (proudový = stream)** nebo **bez spojení (datagramový)**
- protokoly př. **IP** (bez spojení), CONP a CLNP, X.25 (WAN)
- HW zařízení př. síťová karta (vyšší funkce), směrovač, brána

RM OSI – Transportní vrstva

- zprostředkovává **transparentní spojení** s přenosem dat s požadovanou kvalitou **mezi klienty (aplikacemi)** v rámci jednoho síťového zařízení (počítače)
- aplikace může mít více **transportních adres**
- **propojení koncových zařízení**, nejnižší vrstva s entitami pouze v koncových systémech
- stojí mezi uživatelem a sítí

Obrázek: Obrázek průvodce 6→23(5)

- jednotka přenosu = **transportní paket (datagram)**: **záhlaví** s transportní adresou příjemnce a odesílatele (např. TCP/UDP port u Internetu) + data, přenášen v síťovém paketu

RM OSI – Transportní vrstva

- funkce:
  - adresování (transportní na síťové)
  - správa síťových spojení nebo přenosu datagramů
  - **multiplexování a větvení** transportních spojení do síťových
  - rozdělení dat na datagramy, formátování, **segmentace**
  - **řízení "proudu" dat** (správné pořadí datagramů), optimalizace služeb
  - koncová detekce a oprava chyb
- služby (parametrizované - propustnost/rychlost přenosu, doba):
  - transparentní přenos dat s **potvrzováním** (**"spolehlivý"**) nebo bez potvrzování (**"nespolehlivý"**)
  - správa transportních spojení
  - identifikace relační entity (transportní adresou)
  - **duplexní** přenos, **zacházení s daty jako s proudem**
- protokoly **TCP, UDP**, TP0-4, všechny koncové

RM OSI – Relační vrstva

- zabezpečuje **organizovanou výměnu dat mezi aplikacemi**, zprostředkovává **relaci/sezení** (např. sdílení síťového disku)
- jednotka přenosu = **relační paket**: pouze data, přenášen v datagramu
- funkce:
  - organizace a synchronizace dialogu výměny dat (pomocí **kontrolních bodů**)
  - zobrazení (několika) relačních spojení do (několika) transportních
  - správa transportních spojení
- služby:
  - správa a řízení relace (spojení)
  - různý přenos zpráv, řízení interakce
- protokol př. RPC, X.225, X.215 (OSI)

RM OSI – Prezentační vrstva

- poskytuje **jednotnou reprezentaci a zabezpečení informace** (dat, struktur), v jaké jsou dostupné aplikacím a v jaké se přenáší sítí
- funkce a služby:
  - transformace a výběr reprezentace dat (převod kódů, př. který je nejvyšší bit - **big**/little **endian**)
  - formátování, komprese, zapezpečení (šifrování), integrita dat
  - žádosti o správu relace, transparentní přenos zpráv (nezná jejich význam)
- "protokoly" př. **ASCII**, ASN.1 (kódování BER, DER), multimediální formáty, X.226, X.216 (OSI)

RM OSI – Aplikační vrstva

- poskytuje aplikacím **přístup ke komunikačnímu systému** a **aplikační funkce a služby**
- předepisuje **aplikační formát dat**, **záhlaví dat** + data
- funkce:
  - zprostředkování funkcionality sítě
  - řešení aplikační funkcionality – přenos zpráv, určení kvality, synchronizace
  - identifikace, stanovení pověření
  - dohoda o ochraně, dohody o opravách chyb a syntaxi (kódy, abecedy)
- protokoly př. SMTP, MHS (pošta), FTP, FTAM (přenos souborů), Telnet, VT (vzdálený přístup), SNMP, CMIP (management) a mnoho dalších

RM OSI – funkce společné více vrstvám

- výměna dat až po vytvoření spojení všemi nižšími vrstvami
- řízení toku, formátování a zabezpečení dat

Obrázek: Obrázek sítě 33

- **rozkládání a skládání datových jednotek** – fragmentace a segmentace: datagramy, pakety, rámce, sled bitů nebo oktety

RM OSI – funkce společné více vrstvám

- komunikace **se spojením** má 3 fáze: 1. navázání spojení, 2. přenos dat, 3. ukončení spojení
  - dohoda na parametrech, identifikace spojení
  - použití **potvrzování** přijetí či nepřijetí datových jednotek protokolu (“**spolehlivost**”)
  - **stejné pořadí dat** na vstupu i výstupu
- komunikace **bez spojení**
  - při každém přenosu vždy všechny parametry
  - **nezávislý přenos** datových jednotek
  - může být různé pořadí datových jednotek na vstupu a výstupu
  - **datagramová služba**, může být “spolehlivá” i “nespolehlivá”
- konverze mezi těmito typy služby (původně ale jen se spojením, transportní služby musí být se spojením)

TCP/IP (Transmission Control Protocol/Internet Protocol)

- použití v síti **Internet** (největší celosvětová síť propojených heterogenních sítí), **nejpoužívanější síťová architektura**
- všechny informace (konvence, protokoly, doporučení) v **RFC (Request For Comments)** od IAB (rada pro architekturu Internetu), de facto normy **IETF** (komise s pracovními skupinami Internetu)
- historie:
  - vyvinuta v 60.-70. letech na objednávku (D)ARPA USA: propojení počítačů vojenských, výzkumných a akademických pracovišť
  - **ARPANET** 1971 (23 uzlů, 1973 VB a Norsko, 1989 s více jak 1000 uzly zrušen, místo něj NSFNET)
  - původní protokol NCP (Network Control Protocol)
  - 70. léta univerzitní vývoj (Network Measurement Centre, UCLA, **Vinton G. Cerf**), vznikají RFC
  - 1982 **TCP/IP** = Internet, implementace v OS UNIX
  - od počátku 90. let i soukromé využití (výrobní společnosti, poskytovatelé služeb, soukromé osoby a další)
  - dnešní rozsah těžké odhadnout

TCP/IP (Transmission Control Protocol/Internet Protocol)

Obrázek: Obrázek průvodce 2→17(5)

- vrstvy: síťového rozhraní (odpovídá fyzické a linkové z RM OSI), mezisíťová (internet, síťová z RM OSI), transportní, aplikační (3 nejvyšší z RM OSI)
- **vlastní protokoly**, obecně nesrovnatelné s protokoly OSI (TCP/IP vznikla dřív), ale protokoly TCP/IP využívají protokolů OSI a naopak
- dominantní: rozšiřování Internetu, propojení (privátních) sítí, internetové aplikace
- síť tvořena: směrovači (modemy), specializovanými bránami (bezpečnostní, aplikační, telekomunikační), lokálními sítěmi a koncovými zařízeními

TCP/IP (Transmission Control Protocol/Internet Protocol)

Vrstva síťového rozhraní

- **přístup** k přenosovému médiu, **specifická** pro každé přenosové prostředí
- využívá všech typů přenosových prostředí a protokolů fyzické a linkové vrstvy z RM OSI, **využití** definováno v RFC

Vrstva internet

- řeší přenos a **směrování** datagramů na základě síťových (IP) adres
- protokoly **IP** (v4 a v6, síťový), (R)ARP (mapování adres), ICMP (řídící hlášení), OSPF, IGRP (směrování)

Transportní

- transportní služba se spojením ("spolehlivý" protokol **TCP**) nebo bez spojení ("nespolehlivý" protokol **UDP**)
- také směrovací protokoly RIP, BGP
- identifikace aplikačního protokolu **číslem portu** (seznam v RFC 1700)

TCP/IP (Transmission Control Protocol/Internet Protocol)

Aplikační

- **mnoho protokolů**, některé používají TCP, jiné UDP, některé oba, nelze o nich říct nic obecného, služby i protokoly se principiálně liší
- uživatelské protokoly:
  - TCP: **HTTP**, **SMTP**, Telnet, **SSH**, FTP, **IMAP**, POP3, Talk
  - UDP: NFS, BOOTP, TFTP, RPC
  - UDP, TCP: NTP
- služební protokoly (pro funkci sítě):
  - UDP, TCP: **DNS**
  - UDP: DHCP
  - TCP: směrovací, SNMP
- „prezentační-aplikační“ protokoly: **SSL**, S/MIME (zabezpečení dat), virtuální terminál (prezentace, **Telnet**, FTP, SMTP), ASN.1

TCP/IP (Transmission Control Protocol/Internet Protocol)

Obrázek: Obrázek průvodce 9→24(5)

Obrázek: Obrázek sítě 37

Ostatní síťové architektury

Firemní (proprietární) protokolové architektury ze 70.–90. let.

Novell NetWare

- vylepšení Xerox XNS, jednodušší než TCP/IP (spíše pro LAN), (v minulosti) nepoužívanější po TCP/IP
- distribuovaný systém klient-server skrze **volání vzdálených procedur (RPC)**
- nejnižší vrstva podporuje všechny typy přenosových prostředků
- síťová vrstva
  - protokol **IPX (Internet Packet eXchange)** - datagramový, nespojový, podobný IP
  - směrovací protokoly
- transportní vrstva: protokol **SPX (Sequenced Packet eXchange)** - spolehlivý, spojový
- vyšší vrstvy:
  - protokoly SAP (Service Advertising Protocol) a **NCP (NetWare Core Protocol)**
  - zprostředkování zpráv, doplňkové moduly (NLM)
  - emulátor NetBIOS (viz dále)

Apple AppleTalk

- distribuovaný systém klient-server
- spodní vrstvy podporují několik přenosových prostředků (př. EtherTalk) a **LocalTalk** (firemní protokol přístupu k médiu)
- síťová vrstva: dynamická adresace, vytváření sítí a zón, protokoly DDP a AARP
- transportní vrstva: několik transportních, směrovacích a specifických protokolů (ATP, RTMP)
- vyšší vrstvy: aplikační protokoly ADSP, PAP, **AFP** (přenos souborů)

IBM/Microsoft Network

- vlastní architektura založená na **IBM LAN Manager**
- původním základem protokol 3COM **NetBEUI** (NetBIOS Extended User Interface) implementující IBM **NetBIOS** (Network BIOS):
  - nejstarší **API pro LAN**
  - elementární I/O operace přenosu dat, 19 služeb (jmenné, relační, datagramové, všeobecné)
  - bez směrování, funkce linkové, transportní a částečně relační vrstvy, ne síťové $\Rightarrow$ použitelný jen v LAN
- nyní TCP/IP pro NetBIOS a aplikační protokol IBM/Microsoft **SMB** (Server Message Block) / **CIFS** (Common Internet File System):
  - nejpoužívanější pro souborové a tiskové servery v LAN
  - model klient-server se zabezpečným přístupem ke **sdíleným prostředkům** na různých úrovních (disky, adresáře, tiskové fronty)

Další (minulost): Xerox Networks Systems (XNS), Banyan Vines, Digital DECnet aj.

OSI

- řeší přenos dat mezi systémy nezávislými na fyzických prostředích, skrze spolupráci systémů na úkolech – **obecná řešení**
- definuje koncové a mezilehlé systémy, oblasti, správní domény aj.
- fyzická a linková vrstva: normalizovaná rozhraní a linkové protokoly (HDLC, LAPB)
- síťová vrstva: služby se spojením (CONS, protokol **CONP**) a bez spojení (CLNS, **CLNP**)
- transportní vrstva: spojové protokoly **TP0-4**
- vyšší vrstvy: relace pomocí tokenů, prezentační formát **ASN.1**, aplikační služby, systém zprostředkování zpráv, adresářový systém a další protokoly (FTAM, VTP)

Management sítě

= sledování zahajování, ukončování a monitorování činností síťových zařízení a optimalizace datových přenosů v síti, (automatická) rekonfigurace sítě

- součást aplikační vrstvy
- u OSI protokol **CMIP (Common Management Information Protocol)**:
  - centralizovaný
  - různé modely managementu, řešení poruch, konfigurace, účtování, výkonnosti, bezpečnosti zařízení a datových přenosů
- u TCP/IP protokol **SNMP (Simple Network Management Protocol)**:
  - distribuovaný, transakční, jednodušší, nejpoužívanější
  - agent (program řízeného systému, ukládá data) a manažer (aplikace řídící agenty, sbírá data)
- vzdálené monitorování (RMON) – vzdálené monitorovací sondy
- např. management založený na WWW (WBEM), Java JMAPI a další

Bezpečnost a ochrana sítě

- na odpovídajících vrstvách zajištění integrity rámce, paketu, datagramu atd.
- ochrana proti čemu?
  1. **obsah**: ideologie, ohrožující mravní výchovu, aj.
  2. **útoky** na činnost systému a neoprávněný přístup k datům
  3. organizační a fyzická - **sociální inženýrství** („ukecat“ pracovníka s právy, „servis“ si odnese disk s daty apod.)
- útoky zvenčí a zevnitř – řeší **podniková bezpečnostní politika**
- kritéria hodnocení bezpečnosti (ITSEC): důvěrnosti informací (dostupné jen oprávněným osobám), integrita (nenarušení neoprávněnou osobou), dostupnost (zaručení přístupu)
- obecné metody ochrany
  - omezování přenosu dat a přístupu k síti: blokování, filtrace
  - autorizace přístupu: obvykle jméno a (jednorázové) heslo, vícefaktorové, specializované protokoly
  - zabezpečení kanálu: šifrování, výměna klíčů
  - autenticita zpráv: digitální podpis (hashování), certifikáty a certifikační autority

Bezpečnost a ochrana sítě

OSI

- řešení rozpoznání neautorizovaného chování (autentizace, řízení přístupu, zajištění důvěrnosti a integrity dat)
- zabezpečovací protokoly
- snaha o minimalizaci zranitelných míst

TCP/IP

- **původně žádné zabezpečení** (**“Internet je nebezpečný!”**), ponecháno na aplikace
- typicky jednoduchá autorizace jménem a heslem (plain text)
- útoky:
  - falešná adresace (spoofing)
  - na hesla (analýza protokolů, „trojské koně“, apod.)
  - odposlech
  - odmítnutí služby (DoS = Denial of Service, zahlcení, vyčerpání zdrojů)
  - zneužití chyb aplikací (exploit, šíření přes služby WWW a email)
  - ...

Bezpečnost a ochrana sítě

TCP/IP

- ochrana
  - **firewall** (oddělení vnitřní sítě od vnější) s **demilitarizovanou zónou (DMZ)** – filtrace provozu a kontrola adres (prevence před DoS)
  - **překlad adres (NAT)** – vlastní „skrytá“ adresace
  - **aplikační brány (proxy)**, zástupné servery
  - autentizace komunikujících stran, autorizace přístupu k prostředkům (datům)
  - zabezpečení komunikace (šifrování)
  - opatření proti zahlcení aplikace
  - ...
- protokoly **bezpečnostní architektury pro IP**: **IPSec** (bezpečná komunikace na síťové vrstvě), **SSL/TLS** (na transportní vrstvě), **RADIUS** (autentizace a autorizace)

Technologie fyzické vrstvy

Přenos dat

- u protokolů nižších vrstev (fyzické, linkové, síťové) rozlišujeme typ přenosu, synchronizaci přenosu, použití virtuálních okruhů aj.

Sériový přenos

- dvojice vodičů, signálový a zem, bity dat přenášeny za sebou – sériově
- symetrický signál – zvlášť dvojice vodičů, např. pro příjem a vysílání dat, př. X.21
- asymetrický signál – více signálových vodičů oproti společné zemi, př. V.24

Paralelní přenos

- skupina, např. osmice, vodičů, signálové a zem, několik (8) bitů dat přenášeno zároveň – paralelně
- typické použití u vnitřních sběrnic v počítači nebo starší připojení periferních zařízení (tiskárna, modem)

Přenos dat

Synchronní přenos

- konstantní rychlostí, stejnoměrná garantovaná šířka pásma
- dříve blokový: bloky dat (**fyzické rámce**) konstantní délky rozložené do **slotů**, pro daný přenos vyhrazeny sloty se stejným pořadovým číslem, synchronizační bity pro synchronizaci přijímače s vysílačem na začátku bloku
- dnes kromě dat ještě **synchronizační signál** („hodiny“), zdrojem jedno zařízení, ostatní se přizpůsobí
- použití v telekomunikačních sítích (např. telefon 32 slotů po 64 kb/s), NE Internet

Paketový přenos

- proměnlivou rychlostí, bloky dat (**pakety**) obecně různé délky
- negarantovaná šířka pásma (maximální dosažení např. pomocí QoS), ale efektivnější využití pásma
- použití v datových sítích, např. Internet

Přenos dat

Asynchronní přenos

- kombinace předchozích, garance šířky pásma
- pakety stejné délky přenášeny proměnlivou rychlostí (start a stop bity), jednotlivé bity přenášeny synchronně (tzv. arytmický přenos)
- přenos bitů na vzorkovací frekvenci (řádově vyšší než bitová, kvůli rozpoznání bitů), vyšší režie
- např. síť ATM (pakety = **buňky**)

Přenos dat

Virtuální okruh

- vytvářený v síti některými protokoly (na nižších vrstvách, ale i síťové), např. Frame Relay, X.25
- nejprve sestaven (pomocí **signalizace**), pak přenos dat (s identifikací okruhu) po okruhu, v případě přerušení přenosu se vytvoří okruh nový
- spíše telekomunikační sítě, NE u Internetu – přerušení okruhu znamená přerušení spojení, IP pakety přenášeny samostatně
- typy:
  - **pevný (permanent)** – sestavené v síti napevno správcem
  - **komutovaný (switched)** – dynamicky vznikající dle potřeby přenosů

Strukturovaná kabeláž [LAN]

- síťové (a telefonní) rozvody: zásuvky, propojovací kabely, **propojovací (patch) panel**, optická vlákna, distribuční box optiky aj., ve skříni (rack)

Koaxiální kabel

- dnes se již nepoužívá
- **tlustý**: ∅ 1 cm (např. Belden 9880 PVC), max. 500 m, zakončený **terminátory** 50 Ω, připojení uzlu přes **transceiver** napíchnutý svorkou **vampír**, redukce i na tenký a dvojlinku
- **tenký**: RG 58, max. 185 m (u stejných síťových karet uzlů až 400 m), zakončený terminátory 50 Ω, připojení přes **BNC konektor** (existují i transceivery)

Strukturovaná kabeláž [LAN]

Kroucená dvojlinka (Twisted Pair)

- max. 100 m (závisí na kvalitě kabelu), přenos signálu kódováním Manchester II (log. $1 = -2\,\mathrm{V}$)
- 4 páry měděných vodičů, drát nebo lanko (licna, svazek drátků), po dvou kroucených
- nestíněná (**UTP**): kategorie EIA/TIA 3 (do 25 MHz), 5(E) (do 100 MHz), 6 (do 250 MHz), 7 (do 600 MHz)
- stíněná (**STP**)

Obrázek: Obrázek průvodce 56→61(5)

- **konektor RJ45**: nejčastěji zapojení podle EIA/TIA 568B s 1. párem (modrý) pro telefon a 2. a 3. párem (oranžový a zelený) pro datovou síť

Obrázek: Obrázek průvodce 56→61,62(5)

Strukturovaná kabeláž [LAN]

Optická vlákna (Fiber optic)

- dvě vrstvy skla: obal ($\varnothing$ 125 μm) a jádro – **vícevidové** ($\varnothing$ 50 a 62.5 μm, paprsky se odráží od rozhraní skel) a **jednovidové** (9 μm), buzení laserem (850, 1300, 1500 nm)
- primární ochrana – $\varnothing$ 250 μm, optický konektor SC s kouskem vlákna, tzv. **pigtail**, navařený na jiné vlákno
- sekundární, těsná sekundární ochrana – $\varnothing$ 0.9 mm, možné nasadit **různé optické konektory** (FC, LC, ST aj., dříve připojení přes **optické transceivery**)
- svazky mnoha vláken s (kevlarovou) ochranou v optických kabelech
- vlákno simplexní, pro duplex dvojice vláken – pro jednu frekvenci, dnes i „multifrekvenční“ duplexní vlákna
- dosah 2–3 km (vícevidové) nebo až 70 km (jednovidové), použití optických rozbočovačů pro páteřní sítě

Lokální sítě [LAN]

- v minulosti vyvinuta řada systémů LAN: Ethernet, FDDI, Token Ring a Token Bus, Arcnet aj., dnes jen Ethernet a FDDI
- IEEE: počátkem 80. let sjednocení a **normy IEEE 802.xx** pro systémy LAN, později převzaté ISO jako normy ISO 8802-xx

Obrázek: Obrázek průvodce 111→65(5)

- linková a částečně fyzická vrstva rozděleny do podvrstev:
  - **MAC (Medium Access Control)** – přístup na (sdílené) přenosové médium, zasahuje do fyzické i linkové vrstvy, řešená HW, závislost na topologii a HW, normy IEEE 802.3 – 802.15
  - **LLC (Logical Link Control)** – správa logických spojení, linková vrstva, řešená HW i SW, nezávislá na HW, IEEE 802.2
- připojení pomocí **síťové karty** – zčásti realizuje linkové protokoly

Ethernet [LAN]

- sdílené přenosové médium, v daném okamžiku využívá jeden uzel
- uzly samostatné, rovnocenné

Ethernet (II, IEEE 802.3)

- počátky koncem 70. let Xerox, 1982 DEC, Intel a Xerox jako DIX Ethernet (Ethernet II), 1985 IEEE 802.3
- 10 Mb/s, 8.5 MHz
- **segment** = počítače připojené na médium (kabel)
- **tlustý** (10BASE-5, DIX): tlustý koaxiální kabel, topologie sběrnice, konektor AUI (CANNON 15) na síťové kartě, max. 100 stanic
- **tenký** (10BASE-2, IEEE 802.3a): tenký koaxiální kabel, topologie sběrnice, připojení přes **konektor BNC-T** a konektor BNC na síťové kartě, max. 30 stanic

Ethernet [LAN]

Obrázek: Obrázek průvodce 61→69(5)

- **s kroucenou dvojlinkou** (10BASE-T, IEEE 802.3i):
  - konektor RJ45 na síťové kartě, kontrola integrity připojení pomocí signálu LinkBeat
  - připojení k **opakovači** (**linkový segment**), hvězdicová topologie, max. 100 m mezi počítačem a opakovačem
  - duplexní přenos (**Half Duplex**) – na uzlu 2. pár (oranžový) pro vysílání, 3. (zelený) pro příjem
  - při propojení dvou počítačů „překřížení“ – plně duplexní přenos (**Full Duplex**), teoreticky max. rychlost
- **s vícevidovými optickými vlákny** (10BASE-Fx, IEEE 802.3j): původně jen propojení optických opakovačů (FO-HUB), konektor AUI (CANNON 15) na síťové kartě, dnes mnoho různých konektorů (LC, SC, FC, aj.), max. 2 km

Ethernet [LAN]

Opakovač (Repeater)

- HW zařízení pro propojení segmentů, rozbočovač
- data jsou zopakována na všechna ostatní rozhraní (**porty**) opakovače, tj. do všech linkových segmentů
- **HUB** = opakovač pro kroucenou dvojlinku, propojení dvou HUBů „překříženým“ kabelem (nebo jeden port HUBu s přepínačem)
- možnost centralizované správy segmentu

Vícesegmentové sítě

- omezující metody Model I a II pro max. dosah a konfiguraci sítě
- omezení na počty opakovačů a vzdálenosti mezi nimi (Model I) nebo pomocí maximálního zpoždění přenosové cesty (Model II)

Ethernet [LAN]

Fast Ethernet (IEEE 802.3u)

- 1993 sítě 100BASE-T a 100VG-AnyLAN, z důvodu zpětné kompatibility u metody přístupu k médiu (viz linková vrstva) vybrána 100BASE-T
- 100 Mb/s, 125 MHz
- jen hvězdicová topologie s opakovači dvou tříd: **Class I** (retranslace signálu z linkového segmentu umožňující použití různých linkových segmentů, max. jeden na segmentu) a **Class II** (jen opakování signálu, jen stejné linkové segmenty, max. 2)
- fyziká vrstva (100BASE-X) podle FDDI: přenos čtveřic bitů (nibble) kódovaných do 5 bitů
- kroucená dvojlinka (100BASE-TX kategorie 5, 100BASE-T4 kategorie 3 25 MHz dva páry vodičů navíc) – max. 200 m
- optická vlákna (100BASE-FX) – max. 300 m (Full Duplex 2 km)
- volitelná duální rychlost 10/100 Mb/s a Half/Full Duplex: pomocný **Auto-Negotiation Protocol** využívající rozšířený signál integrity sítě

Ethernet [LAN]

Gigabitový Ethernet (IEEE 802.3z, 802.3ab)

- 1988 pro optické linky (IEEE 802.3z), pak pro kroucenou dvojlinku kategorie 5E (IEEE 802.3ab), vytlačil FDDI a ATM
- 1 Gb/s, 1062.5 MHz (optika)
- jen hvězdicová topologie s opakovači
- optická vlákna (jednovidová 1000BASE-LX, vícevidová 1000BASE-SX): fyzická vrstva podle Fibre Channel: přenos 8 bitů kódovaných do 10 bitů, max. 550 m (vícevidové, 850 nm) nebo 2 km (jednovidové, 1300 nm)
- kroucená dvojlinka (1000BASE-T): duplexní přenos na všech 4 párech u kategorie 5E, plně duplexní přenos u kategorie 6, max. 100 m

Ethernet [LAN]

10Gigabitový Ethernet (IEEE 802.3ae)

- 10 GB/s, velký dosah
- jen režim Full Duplex, ne sdílené médium
- fyzická rozhraní pro LAN a WAN (propojení s DWDM)
- 4 rozhraní odvozená od 1000BASE-X s rychlostí 2.5 GB/s
- optická vlákna (mnohovidová 10GBASE-S 400 m, jednovidová 10GBASE-L/E 10/40 km)
- kroucená dvojlinka (10GBASE-T 55 m kabel kategorie 5E nebo 6, 100 6A nebo 7)

FDDI [LAN]

- Fiber Distributed Data Interface – optická vlákna, 1989 ANSI X3T12, 1990 ISO 9314
- CDDI (Copper DDI) – kroucená dvojlinka
- vysokorychlostní páteřní sítě počátku 90. let, univerzitní sítě (campus)
- 100 Mb/s, max. 2 km (vícevidová vlákna), 60 km (jednovidová)
- zdvojená kruhová topologie: protisměrné páteřní kruhy, jeden primární, druhý záložní, v daném čase aktivní jen jeden
- zařízení: koncové stanice – porty pro oba kruhy (DAS) nebo jen jeden (SAS), **koncentrátory** – více portů pro připojení více konc. stanic, mosty

Bezdrátové lokální sítě (WLAN) – Wi-Fi [LAN]

- důvody pro WLAN (**Wireless LAN**): mobilita, snadná použitelnost, dostupnost, nižší náklady, rozšiřitelnost, roaming (vysílače si klienta předávají), atd., polovina 90. let
- použití pro vnitřní (původně, popř. v kombinaci s kabeláží) i vnější prostory (např. připojení k Internetu), propojení s drátovými LAN
- norma **IEEE 802.11** (1997), 2 Mb/s, mnoho rozšíření, např. 802.11b = **Wi-Fi (Wireless Fidelity)** – až 11 Mb/s v závislosti na poměru signálu k šumu, běžně 60 %, dosah až 11+ km (venku), 802.11a/g – až 54 Mb/s, 802.11n – až 500+ Mb/s

Wi-Fi [LAN]

Konfigurace (topologie)

- peer-to-peer/**ad-hoc**: přímá komunikace mezi stanicemi, do 10-ti stanic
- **infrastrukturní**/s **přístupovým bodem (access point, AP)**: propojuje WLAN a "drátovou" LAN (např. Ethernet), stanice komunikují jen prostřednictvím AP (nejdříve asociace a autorizace), bezpečnostní prvky (filtrace, šifrování, atd.), až 100 stanic
- s více přístupovými body (**roaming**): AP propojeny pevnou sítí, klient se přepojuje k AP s nejlepším poměrem signálu k šumu, když tento klesne pod nějakou mez
- point-to-point: propojení dvou sítí pomocí AP

Wi-Fi [LAN]

Přenosové médium

- rádiové vlny 2.4 (**802.11b/g/n**), 5 GHz (**802.11a/n**) – veřejné, není třeba licence, vzájemné rušení (také např. Bluetooth, RFID čipy, RC modely na dálkové ovládání a další)
- šíření signálu metodou rozptýleného spektra (v pásmu frekvencí):
  - přeskakování frekvencí (FHSS): 2.4 GHz pásmo dělené na 75 kanálů, při vysílání se periodicky přeskakuje mezi frekvencemi, př. starší Wi-Fi, Bluetooth
  - přímá sekvence (DSSS): 2.4 GHz pásmo dělené na 14 kanálů po 22 MHz, které se částečně překrývají, př. Wi-Fi 802.11b
  - ortogonální frekvenční multiplex (OFDM): 2.4 a 5 GHz, 802.11a/g, 802.11n technologie MIMO
- poloduplexní spoj, ale je možný i duplexní (dva páry antén)
- **antény**: horizontální, verikální a kruhové polarizace signálu, všesměrové, sektorové, směrové, provedením síťové, paraboly, šroubovice, Yagi, omezení na výkon vyzářený anténou normou ČTÚ (100 mW)

Bezdrátové personální sítě (WPAN) – Bluetooth [PAN]

- projekt „Blue Tooth“, Ericsson, 1994, bezdrátová komunikace mezi různorodými zařízeními (počítače, mobilní telefony, PDA, dig. fotoaparáty, kamery aj.)
- rádiové vlny 2.4 GHz, přenosová rychlost 1 nebo 2 Mb/s, max. 10 m (s opakovači do 100 m)
- norma **IEEE 802.15**
- komunikace po kanálech (tzv. piconetech) s pseudo-náhodnými skoky
- **Master** a **Slave** uzly (max. 7, další zaparkované)

Bluetooth [PAN]

- odlišná protokolová architektura: fyzická (Bluetooth radio, podvrstvy Radio a Baseband), linková, vyšší (identifikace a možnosti zařízení, podpora služeb, protokoly SDP, RFCOMM, TCS BIN, WAE/WAP)
- **profily zařízení** – definice parametrů protokolů služeb, GAP a SDAP pro vyhledávání (SDP), TCS-BIN pro telefonii, SPP pro emulace sériového propojení (RFCOMM, modem, PPP do LAN), GOEP pro souborové přenosy aj.
- podvrstva **Baseband**: adresace, tvorba sítí Piconet (uzly ve stavech a režimech, procedury Inquiry a Paging), zřizování linek (synchronní SCO, asynchronní ACL), řízení toku dat a zabezpečení přenosu

Rozlehlé sítě [WAN]

- velké vzdálenosti → odlišné technologie přenosu dat než v LAN
- dvojbodová propojení mezi prvky DCE nebo virtuální okruhy
- využití **telekomunikačních sítí**
- optické systémy:
  - **SONET/SDH**: synchronní vysokorychlostní přenosy, rychlosti 50 Mb/s až 10 Gb/s, aplikace v síti ATM
  - **DWDM**: multiplex na různých vlnových délkách, desítky virtuálních optických vláken v existujících fyzických, rychlosti řádově až Tb/s, full duplex po jednom vláknu
- rádiové – tzv. „last mile“:
  - dvojbodové: přímá viditelnost, až 20 km, 2.4, 3.5, 10 GHz – až 90 Mb/s, licencovaná pásma
  - **FWA**: pevné bezdrátové okruhy vzdálených uzlů se základovou stanicí, 26 GHz, buňková síť, dosah 5 km
  - **WiMAX**

Sériová linka [telekomunikační WAN]

Obrázek: Obrázek průvodce 40→49,52(5)

- propojení koncového zařízení (DTE), např. počítač, s propojovacím prvkem (DCE), např. modem, nebo dvou propojovacích prvků
- **ITU V.24 (ANSI RS232)**: sériový asynchronní arytmický přenos, rychlost desítky kb/s (64, 115.2 max), full duplex, konektory CANNON 9 a 25 (porty COM), propojení dvou počítačů pomocí „překřížení“ vodičů (**nulový modem**)
- dnes nahrazena **bezdrátovými PAN** (Bluetooth, infra)
- připojení modemu: signály DTR, DSR (signalizace), RTS, CTS (řízení toku) nebo znaky XON, XOFF, signály TD, RD (data, AT-příkazy)

Modem [telekomunikační WAN]

- pro připojení k datové síti pomocí analogové telefonní sítě – modulace a demodulace dat a zvuku
- **modulátor/demodulátor = modem** – připojen sériovou linkou/bezdrátovou sítí k počítači nebo vestavěný a telefonní linkou (kroucená dvojlinka/bezdrátová síť) k telefonní síti
- vytvoření okruhu v telefonní síti, dohoda stran na parametrech komunikace (nejvyšší rychlost, zabezpečení apod., protokol PPP) a přepnutí na data, poté uzly (DTE) propojeny transparentně

AT-příkazy (Hayes)

- znakové ovládání modemu počítačem a zprávy od modemu, např. ATDTčíslo, AT OK, CONNECT

Modem [telekomunikační WAN]

- přenosové rychlosti na telefonní drátové lince (doporučení ITU):
  - **přeložené pásmo** (Voice Band, překlad dat na zvuk v pásmu 0.3 až 3.4 kHz, komutovaná linka přes zesilovací stanice mezi ústřednami): nominální 9.6 (V.32), 14.4 (V.32bis), 28.8 (V.34), 33.6 (V.34+), 56/33.6 (download/upload, **V.90**, digitální ústředny a linky mezi nimi) kb/s
  - **základní pásmo** (Base Band, tzv. „širokopásmové modemy“, pevné linky): stovky kb/s až jednotky Mb/s (plný duplex), rozhraní V.35
- dnes bezdrátové sítě, např. GSM
- možná komprese dat (protokol MNP 5, ITU **V.42bis**) – rychlosti až stovky kb/s (v přeloženém pásmu), potřeba vyšší rychlosti na lince k počítači
- detekce chyb přenosu (V.42)

ISDN [telekomunikační WAN]

- připojení k datové síti pomocí digitální telefonní sítě s integrovanými službami, normy I.430 / I.431
- synchronní přenos dat, kroucená dvojlinka, konektor RJ45
- přenosové rychlosti (na telefonní drátové lince):
  - **Basic Rate** (**euroISDN2**, linka E0/T0): dva datové kanály B 64 kb/s, signalizační kanál D 16 kb/s, synchronizace
  - **Primary Rate** (**euroISDN30**, linka E1/T1): třicet datových kanálů B 64 kb/s, signalizační kanál D 64 kb/s

euroISDN2 (V.110)

- rozhraní U: dvojlinka mezi telefonní linkou a zařízaním **NT-1**
- rozhraní S/T: dvě dvojlinky z NT-1, sběrnice pro připojení digitálních zařízení (počítače pomocí „digitálního modemu“) nebo **terminálního adaptéru** pro připojení analogových zařízení, současně mohou komunikovat max. 2 (dva datové kanály B)

xDSL [telekomunikační WAN]

- dosažení maximální rychlosti na telefonní lince, různorodé technologie xDSL
- **ADSL** (Asymmetrical): rychlost 12/3.5 Mb/s (download/upload, ADSL2) nebo 24/1 Mb/s (ADSL2+), dosah do 7 km, využití dalších dvou kroucených párů vodičů pro přenos mimo telefonní pásmo (4 kHz) – potřeba **splitteru** u/v DSL modemu a zařízení DSLAM v ústředně, nátupce euroUSDN2
- HDSL (High data rate): rychlost 2 Mb/s, nástupce euroISDN30
- SDSL (Symmetrical), VDSL (Very-high-bit-rate, až 52 Mb/s) aj.

GSM [telekomunikační WWAN]

- bezdrátová původně analogová síť jen pro hlas, dnes digitální, normy ETSI
- pokryté území rozdělené do oblastí s (překrývajícími se) **buňkami** obsluhovanými jednou **BTS (Base Transceiver Station)** s max. 12 vysílači (běžně 4)

Obrázek: Obrázky průvodce 62→48,49(2)

- mobilní telefon komunikuje s BTS, roaming (síť si udržuje informaci, ve které oblasti buněk se telefon nachází a hledá jej ve všech buňkách oblasti)
- dvě frekvence: primární (900 MHz, rozsah 25 MHz po 200 kHz), sekundární (1800 MHz, rozsah 75 MHz), každá konkrétní frekvence rozdělěna do 8 **slotů**

GSM [telekomunikační WWAN]

- další zařízení: BSC (řídí BTS), NSS (přepíná okruhy, obsahuje databáze uživatelů), TRAU (převody rychlostí) aj.
- komunikace mezi telefonem a BTS (ve slotech): datový kanál TCH (9.6 kb/s, asynchronně), kombinované služební kanály synchronizace (GSM používá synchronní přenos), signalizace, „špehovací“ (telefon odesílá asi 80 bytů každé 2 minuty)
- počítač propojen s telefonem pomocí zařízení **RA-0** (= modem, součást telefonu), NSS připojeno na směrovač ve WAN, se kterým počítač vytvoří virtuální okruh
- **GPRS/EDGE**: místo virtuálního okruhu paketový přenos, teoreticky až ve všech 8 slotech (GPRS až 171.2 kb/s, EDGE až 500 kb/s), prakticky 4 sloty
- **UMTS/HSPA**: GSM sítě 3. generace, až 14 Mb/s, multimediální služby, 3.5 generace HSDPA, HSPA+ aj.
- **LTE**: GSM sítě 4. generace, až 300 Mb/s (?)

Bezpečnost na fyzické vrstvě

- útoky:
  - přerušení (drátové) linky $\rightarrow$ záložní linka, fyzická ochrana
  - rušení (bezdrátové) komunikace – cílené, ale i např. vadné konektory, vlivy okolního nebo i přenosového prostředí $\Rightarrow$ vadné linkové rámce
  - odposlech – fyzická ochrana linek a šifrování, omezení šíření bezdrátového signálu, užitečné pro správce
  - modifikace přenášených dat – neúměrně nákladná, spíše na vyšších vrstvách
- protokoly řeší ochranu a detekci chyb jen z technických příčin
- „inteligentní útočník“ $\rightarrow$ **fyzická ochrana** linek a omezení vysílačů + šifrování

Technologie linkové vrstvy

Propojování sítí IEEE 802 [LAN]

- původně LAN = uzly propojené stejnou síťovou technologií (např. segment Ethernetu), v rámci LAN stejný linkový protokol
- dnes LAN = propojení LAN s obecně **různými technologiemi** a linkovými protokoly pomocí **mostů** nebo **přepínačů**
- WAN = propojení (dnešních) LAN pomocí **směrovačů**
- **norma IEEE 802.1**: celková architektura sítí 802 (LAN/MAN), propojení sítí (na úrovni podvrstvy MAC), napojení na vyšší vrstvu, tvorba VLAN, bezpečnost, autorizace, atd.

Podvrstva LLC, norma IEEE 802.2 [LAN]

- řešena HW i SW, nezávislá na HW (fyzickém řešení sítě), rozhraní mezi podvrstvami MAC a LLC $\sim$ rozhraní mezi HW a SW
- navazování, správa a ukončování **linkových spojení**, řízení bezpečného (s rozpoznáváním chyb) přenosu dat mezi (dvěma) uzly sítě, identifikace vyšších protokolů
- pro protokoly bez vyšších funkcí, např. NetBEUI, poskytuje **datagramovou službu** a **virtuální linkové spoje** s potvrzováním příjmu (vychází z HDLC LAPB, viz HDLC)
- rámec: specifikace cílové (**DSAP**) a zdrojové služby (**SSAP**) (pro SNAP 0xAA, pro NetBIOS 0xF0, čísla viz RFC 1700), řídící pole HDLC (číslování, znovuzasílání atd., typ rámce I, U, S, viz HDLC, u IP rámce typu U, pole = 0x3)

Obrázek: Obrázek sítě 121→125(5)

Most (Bridge) [LAN]

- **norma IEEE 802.1d**, propojení (různých) LAN na úrovni MAC podvrstvy (**transparent MAC bridge**), např. Ethernet a WLAN, Ethernet a FDDI, možnost stanovení priorit přenosu s přiřazenou třídou (802.1p)
- transparentní vzhledem k vyšším protokolům a např. ve vícesegmentové homogenní síti Ethernet (síť se jeví jako jeden segment, např. Ethernetové segmenty s opakovači)

= multiportový opakovač, ale rámce jsou opakovány jen na to (jiné) rozhraní (port) mostu, ke kterému je připojen adresát rámce; všesměrové (broadcast) rámce jsou opakovány na všechny ostatní porty

- **filtrační tabulka**: linková (MAC) adresa vs. port – naplněná manuálně nebo automaticky samoučením (omezená doba platnosti položek, např. 300 s)
- **stavová tabulka** portů – seznam aktivních a blokovaných portů
- parametry: velikost filtrační tabulky, filtrační výkon (načtené rámce/s, přenosový výkon (zopakované rámce/s)

Most (Bridge) [LAN]

Algoritmus TRA (Transport Roading Algorithm)

- naplněn, aktualizace a použití filtrační tabulky

→ pokud adresa adresáta rámce není v tabulce, pracuje jako opakovač, ale pro nevšeobecné adresy navíc uloží do tabulky adresu odesílatele rámce vs. port, kterým rámec přišel = **learning**

→ pokud v tabulce je adresa adresáta rámce a pokud je asociovaný port jiný než port asociovaný s adresou odesílatele a není blokovaný, zopakuje rámec jen na port asociovaný s adresou adresáta = **forwarding**, jinak jej nezopakuje = **filtering**

- omezení na stromovou topologii sítě s více mosty (jinak cyklický oběh rámců!)

Most (Bridge) [LAN]

Protokol a Algoritmus výběru kostry (STA, Spanning Tree Algorithm)

- výpočet **stromové topologie sítě** s potlačením smyček v libovolné topologii

→ mosty identifikovány prioritou a MAC adresou, zvolen **kořenový most** (s nejnižším id), všechny ostatní mosty si označí jako **kořenový/root port** ten port, kterým vede nejlevnější (nejkratší) cesta ke kořenovému mostu (při stejných přes souseda s nejnižším id), z mostů na stejném segmentu se vybere ten s nejlevnější cestou (a nejnižším id) a jeho port do segmentu je označen (**designated port**), ostatní porty a neoznačené porty ostatních mostů jsou zablokovány (**blocked port**)

- periodicky (2 s) se opakuje, mosty si pomocí konfiguračních zpráv (BDPU rámce, SSAP a DSAP = 42 v LLC záhlaví) vyměňují info s id a cenou na speciální STP multicast MAC adrese

Most (Bridge) [LAN]

Protokol a Algoritmus výběru kostry (STA, Spanning Tree Algorithm)

Obrázek: Obrázek Wikipedie [Spanning Tree Protocol]

BDPU rámec: typ a příznak zprávy (např. konfigurace, změna topologie), id kořenového a aktuálního mostu, id portu, který odeslal rámec, cena cesty ke kořenovému mostu, čas odeslání rámce, aj.

Protokol GARP

- dynamická registrace atributů mostu a uzlů na speciálních MAC adresách (např. skupinová adresa, VLAN identifikátor aj.)
- **protokol GMRP** pro vytváření skupin se skupinovou adresou

Ethernet [LAN]

IEEE 802.3

- původně s opakovači propojujícími (linkové) segmenty
- rámce se šíří segmentem po sdíleném médiu nezávisle na sobě, stanice (síťové rozhraní) **“vidí” všechny, ale přijímá jen ty adresované jí** nebo všeobecně (“normální” režim/mód)
- v tzv. **promiskuitním režimu** přijímá (a předává OS) všechny rámce
- uzly rovnocenné, jen jeden v daném čase využívá sdílené přenosové médium pro vysílání rámců = režim (Half) Duplex
- 10Gigabitový Ethernet již nepoužívá sdílené médium (režim Full Duplex)

Ethernet [LAN]

Protokol CSMA/CD (Carrier Sense Multiple Access/Collision Detection)

→ **kolizní přístup** ke sdílenému médiu: stanice naslouchá (Carrier) a vysílá, až když nevysílá žádná jiná (tj. společné médium není používáno), když takto začne vysílat více stanic zároveň (zjistí porovnáním vysílaného a přijímaného signálu), dojde ke **kolizi**, první stanice, která ji detekuje, vyšle tzv. **signál JAM** (kvůli detekci kolize ostatními stanicemi) a všechny se na náhodný čas odmlčí (interval času odvozený od MAC adresy se v iteracích zdvojnásobuje, desítky μs, max. 16 pokusů)

= **stochastická, nedeterministická metoda** přístupu ke sdílenému médiu

- větší provoz = více kolizí, nejlepší využití a tedy propustnost sítě kolem 20 % (limit 40 %) (u FDDI 80-90 %), teoretické max. 30 stanic na segmentu
- propojení dvou počítačů (linkovým segmentem, např. kroucenou dvojlinkou) = **bezkolizní segment**

Přepínaný Ethernet [LAN]

- místo opakovače propojuje (linkové) segmenty přímo most
- **normy IEEE 802.1d** a **802.1q**

Přepínač (Switch)

= **multiportový most**, který zpracovává příchozí rámce na svých rozhraních "paralelně", vytváří "souběžné" **virtuální linkové segmenty** (dvobodové plně duplexní spoje) propojující odesílatele s adresátem . . . **přepínání** přenosů dat

- virtualní linkový segment je bezkolizní, **kolize** nastávají pouze pro segmenty s různými odesílateli, ale stejným adresátem
- pro přepínání používá **přepínací matici**, dokáže propojit sítě obecně s různými rychlostmi (má vyrovnávací paměť, store-and-forward) a může hned po načtení záhlaví rámce načítat další rámec (cut-through)

Ethernet [LAN]

Ethernet II

Obrázek: Obrázek průvodce 116→111(5)

- předepsaný pro lokální sítě přímo připojené do Internetu
- rámec: preambule pro synchronizaci hodin uzlů přijímajících rámec s vysílajícím uzlem (fyzická vrstva, (31× 10)11), adresy příjemce a odesílatele, specifikace protokolu vyšší (síťové) vrstvy, data 46–1500 B a kontrolní součet v zápatí
- **linkové (MAC) adresy**: globální (druhý bit = 0, první tři bajty identifikují výrobce, v trvalé paměti síťové karty), skupinová (nejnižší bit prvního bytu = 1), všesměrová (samé 1)

Ethernet [LAN]

Ethernet IEEE 802.3

Obrázek: Obrázek průvodce 118→125(5)

- rámec stejný jako u Ethernetu II, jen místo specifikace protokolu vyšší vrstvy je délka dat (max. 1500 B, čísla protokolů jsou vyšší)
- linkové (MAC) adresy mohou mít délku 2 až 6 B
- data nesou rámec podvrstvy **LLC** (802.2) se SNAP
- **SNAP (Sub-Network Access Protocol)** = specifikace protokolu vyšší vrstvy – kód organizace přidělující čísla a číslo protokolu (např. IP má 0x800, pro kód = 0 čísla stejná jako u Ethernet II, viz RFC 1700)

FDDI [LAN]

- podvrstvy LLC a MAC jako u IEEE sítí
- rámce **Token** a datový: typ, MAC adresa 2/6 B, kontrolní součet, stav (indikátor tokenu), max. 4500 B
- přístupová metoda **Token Passing**: deterministická, odevzdávání práva (tokenu) přístupu ke sdílenému médiu po kruhu (podobně jako u technologie Token Ring)
- mosty pro připojení jiných technologií LAN (Ethernet/FDDI, Token Ring/FDDI)

Wi-Fi [LAN]

DODELAT

Protokol CSMA/CA (CSMA/Collision Avoidance)

= přístupová metoda ke sdíleném bezdrátovému médiu

- nelze detekovat kolize jako u CSMA/CD, používá se **pozitivní potvrzování s vyhrazením pásma na určitý čas**
- asociace = „(znovu)přihlašování“ se k přístupovému bodu (AP)

Wi-Fi [LAN]

DODELAT

Bezpečnost

- obtížná ochrana proti odposlechu na fyzické vrstvě
- **SSID (Service Set ID)**: označení AP ("jméno sítě"), AP jej nemusí vysílat
- **WEP (Wired Equivalent Privacy)**: autentizace stanic vůči AP (40bitové sdílené tajemství = **heslo**, spolu s MAC adresou), symetrické šifrování přenosu (64bitový nebo 128bitový klíč, z toho 24bitů inicializační vektor měnící se s každým rámcem, proudová šifra RC4) – lze v krátkém čase zlomit = **nedostatečné**
- IEEE 802.1x: autentizace (EAP) oproti např. RADIUS serveru
- **WPA (Wi-Fi Protected Access)**, **WPA2**: autentizace (heslo, EAP), tvorba klíčů TKIP, silná šifra AES (WPA2)

Bluetooth [LAN]

DODELAT

VLAN sítě [LAN]

VLAN (Virtual Bridged LAN)

= virtuální síť vytvořená ve fyzické (přepínané) síti

- zpočátku jen proprietární, pak **norma IEEE 802.1q**
- přiřazení uzlů do VLAN pomocí **přístupových tabulek** na přepínačích, na základě portů, MAC adres nebo protokolu vyšší vrstvy – protokol GVRP
- identifikace VLAN pomocí čísla **VLAN ID** (1 až 2048), filtrační tabulka přepínače obsahuje pro každý port přístupovou tabulku s povolenými VLANy, rozšířená filtrační pravidla, priority
- **802.1q tagging**: rozšíření záhlaví linkového rámce (např. Ethernetu) o 4 byty pro prioritu a VLAN ID

(C)SLIP [WAN]

(Compressed) Serial Line IP (RFC 1055, RFC 1144)

- velice jednoduchý, vkládá síťové pakety přímo do asynchronní sériové linky

Obrázek: Obrázek průvodce 67→75(5)

- pro sychronizaci značka END (0xC0) na (začátku a) konci rámce, tento znak v datech nahrazen tzv. **Esc-sekvencí** (0xDB 0xDC, 0xDB nahrazen 0xDB 0xDD)
- nezabezpečuje detekci chyb, nenese info o přenášeném síťovém protokolu – může být jen jeden, nelze dohodnout konfigurační parametry, aj.
- varianta s kompresí (CSLIP):
  - **redukce záhlaví** protokolů IP a TCP (40 bytů) na 3 až 16 bytů, nově lze použít i pro UDP a IPv6
  - pouze vynechání neměnných položek záhlaví protokolu nebo uvádění malých změn (komprimovatelný paket) v sérii paketů

HDLC [WAN]

- více ISO norem, původně IBM SDLC, rozsáhlý protokol, využívají jej (jeho část) nebo jsou z něj odvozeny další protokoly (např. PPP, LAPB a LAPD u ISDN)
- synchronní i asynchronní přenos, detekce chyb (kontrolní součet, negativní potvrzování), řízení toku dat, možnost více síťových protokolů, stavy linky (odpojená, nastavování, přenos dat, odpojování)
- módy ABM (plně duplexní dvoubodový přenos), NRM (SDLC, polo-duplexní přenos), typy rámců **I** (přenos dat), **U** (i řídící funkce) a **S** (řízení toku), specifikované v řídícím poli

Obrázek: Obrázek průvodce 73→77(5)

- značka 0x7E, **bit stuffing** (v bitovém synchronním proudu za každých 5 jedniček nula), adresa 1 byte

PPP [WAN]

Point to Point Protocol (RFC 1661)

- využití pro připojení k datové síti (Internetu) pomocí telefonní sítě, virtuálních sítích aj.
- využívá rámce (je "zapouzřován" rámci) HDLC (u analogových telefonních linek), Ethernet (**PPPoE, PPP over Ethernet**, u ADSL), nebo i FrameRelay
- vyžaduje plně duplexní dvojbodový spoj

Obrázek: Obrázek průvodce 78→83(5)

- HDLC adresa 0xFF (všesměrová), značka pro asynchronní přenos 0x7E + Esc-sekvence (0x7D 0x5E, 0x7D 0x5D, i řídící znaky ASCII)
- **služební (pod)protokoly** pro navázání spojení, autentizaci, skupina protokolů NCP pro síťové protokoly aj. (šifrování, komprese)

PPP [WAN]

Protokol LCP

- protokol pro navázání a ukončení spojení, dohodě na autentizaci apod.
- linka ve **fázích** odpojena, navazování spojení, autentizace (nepovinná, i oboustranně), případné zpětné volání (s případnou kontrolou klientova tel. čísla), další protokoly (šifrování – ECP, MPPE, komprimace – CCP, MPPC, rozložení do více linek – MP, BAP, BACP aj.), síťový protokol (otevření linky pomocí odpovídajícího protokolu NCP), ukončování spojení (signalizace fyzické vrstvě)

Obrázek: Obrázek průvodce 82→85(5)

- rámec: kód příkazu/odpovědi (konfigurace, ukončení spojení, atd.), volby (jaká délka rámce, autentizační protokol, atd.)

PPP [WAN]

Autentizace

- terminálový dialog nebo autentizační protokoly
- **PAP (Password Authentication Protocol)** – příkaz se jménem a heslem, RFC 1334
- **CHAP (Challenge Handshake AP)** – RFC 1994
  - sdílené tajemství (heslo), autentizující strana zašle náhodný řetězec (příkaz **challenge**), autentizovaná strana spočte hash (např. MD5) z tajemství a řetězce a pošle zpět (**response**), první strana stejně spočte hash a porovná
  - varianty MS CHAP 1 a 2 – uložen hash (MD4) hesla (1), navíc šifrování dat (2), RFC 2433, 2759
- **EAP** – autentizace později (v rámci vlastního datového přenosu) libovolným autentizačním protokolem nebo mechanizmem (EAP-MD5 – obdoba CHAP, EAP-TLS), RFC 2284

PPP [WAN]

Protokol IPCP

- řídící protokol typu NCP pro otevření linky pro síťový protokol IP (v4), RFC 1332
- příkazy podobné LCP, volby pro IP adresu, adresy DNS serverů apod.

Frame Relay [WAN]

- datagramový, nespojovaný, "nespolehlivý" protokol

Obrázek: Obrázek průvodce 102→106(5)

- využívá (zejména) pevné **virtuální okruhy poskytovatele** (privátní síť) – parametry množství dat, které lze síti předat za sekundu, a povolené překročení
- připojení směrovače na Frame Relay přepínač, na fyzické vrstvě rozhraní V.35, X.21, rychlosti od 56 kb/s do 100 Mb/s
- rámec: záhlaví s identifikátorem **DLCI okruhu**, bity indikující možnost zahození, blížící se zahlcení okruhu (řeší se zvýšením doby odezvy, na vyšším protoloku snížením rychlosti) aj., data a kontrolní součet
- identifikace síťového protokolu (pole NLPID rámce): **Multiprotocol over FR** (RFC 2427) – např. IP (0xCC) nebo PPP
- **Protokol LMI (Local Management Interface)**: statistiky, účtování, informace o připojení rozhraní apod.

Bezpečnost protokolů linkové vrstvy

- zápatí rámce obsahuje **kontrolní součet**, který příjemce spočítá z přijatých dat a porovná – ochrana (jen) proti rušení
- na **LAN** nebo pevných linkách (např. telefonních) se útoky **neřeší**, uživatelé jsou v pracovně-právním vztahu
- na LAN promiskuitní režim síťové karty, útoky **podvrhnutím adresy** odesilatele (např. nastavením MAC adresy), podvrhnutím položky ARP cache (ARP spoofing a.k.a. **ARP cache poisoning**, viz protokol ARP)
- na **WAN**, komutovaných linkách, např. s protokolem PPP, nebo **WLAN autentizace**, zabezpečení přenosu apod.
- **Access Port Control (IEEE 802.1x)**: autentizace a autorizace přístupu prvku (uzel nebo i přepínač) k síti (přepínači, serveru) pomocí autorizační autority (např. RADIUS server), **protokol EAP** na speciální skupinové adrese, na základě portů, linkových adres nebo asociace (u WLAN)

Síťová vrstva

Síťové protokoly

- **linkové protokoly**
  - vyměňují data mezi **sousedními uzly** v rámci lokální nebo rozlehlé sítě, pomocí sdíleného komunikačního média (např. Ethernet) nebo dvoubodovými linkami (např. PPP)
  - standardizované normami IEEE 802.x
- **síťové protokoly**
  - vyměňují data mezi **libovolnými (nesousedními) uzly** v rozlehlé síti tvořené mnoha lokálními sítěmi

  → data jsou **směrována (routing)** rozhlehlou sítí pomocí směrovačů – nejdůležitější funkce síťové vrstvy (protokolu)
  - dříve různá řešení (TCP/IP, ISO OSI, firemní), dnes de facto standard TCP/IP

Síťové protokoly

- **směrovač (router)**:
  - propojuje lokální sítě (LAN) na úrovni síťové vrstvy, umožňuje libovolné topologie sítě (v praxi propojené hvězdicové)
  - řeší směrování z lokální sítě k následujícímu směrovači nebo koncovému uzlu (**next hop**), rozhoduje na základě svých **směrovacích tabulek**

  = běžný počítač nebo specializované zařízení (směrovač, router) s **více síťovými rozhraními**, předávající si data mezi rozhraními – **forwarding**
  - „vybaluje“ data (síťový paket) z linkového rámce a „zabaluje“ do jiného linkového rámce – i když jsou linkové protokoly sítí stejné!
  - nemění síťový paket (např. adresy)!, až na výjimky, např. položka TTL, fragmentace, volitelné položky aj.
- **koncové uzly** – vysílají a přijímají síťové pakety „zabalené“ do linkových rámců

Návaznost na linkovou vrstvu

- fyzická a linková vrstva implementována na síťové kartě (HW) a jejím ovladačem (driver, SW)

→ **rozhraní ovladače** – standardizovaný způsob přístupu ze síťové vrstvy k linkové, funkce:

  - výběr linkového protokolu (rámce)
  - identifikace a přepínání síťového protokolu (buffer, např. SSAP, DSAP)
  - „zabalování“ síťových paketů a „rozbalování“ linkových rámců
  - služby podvrstvy LLC (správa linkových spojů)

- standardizovaná rozhraní: PKDRV (Packet Driver, pro TCP/IP), **NDIS** (Network Driver Interface Specification, Microsoft/IBM, vyžaduje linkový protokolový ovladač, kterému NDIS ovladač předává rámce)

Internet Protocol (IP)

- 1980 RFC-760, 1981 RFC-791
- poskytuje **„nespolehlivou“ nespojovanou službu** – nevytváří spojení, nepotvrzuje příjem paketů
- spojuje lokální sítě do celosvětové sítě **Internet**
- tvořen několika dílčími protokoly: vlastní IP a služební **ICMP** (diagnostika a signalizace mimořádných stavů), **IGMP** (skupinové adresování), **ARP** a **RARP** (zjištění linkové adresy k IP adrese a opačně)
- síťové rozhraní uzlu má alespoň jednu **síťovou IP adresu**

IP paket (datagram)

– základní jednotka dat přenášených IP

= záhlaví 20 B povinných položek + volitelné položky, data, max. délka 64kB

Obrázek: Obrázek průvodce 132→131(5)

- délka záhlaví: v jednotkách 4 B, tzn. max. 60 B
- **typ služby (TOS)**: původně specifikace kvality přenosu (bity pro prioritu, min. zdržení a cena, max. výkon a dostupnost), dnes **DS (Differentiated Services)** – požadavky garance šířky pásma, protokol RSVP
- identifikace, příznaky a posunutí **fragmentu**: pro účely **fragmentace paketu**, bity příznaků pro zakázání fragmentace (DF) a indikaci dalších fragmentů (MF, tento není poslední)

IP paket (datagram)

- **doba života (TTL)**: zamezení nekonečného „toulání“ paketu, každý směrovač snižuje alespoň o 1 (a musí tedy změnit kontrolní součet záhlaví), při 0 se paket zahazuje a odesilateli je to signalizováno protokolem ICMP, nastavena v OS
- **protokol vyšší vrstvy**: čísla přiděluje IANA, např. ICMP 1, IGMP 2, IP 4, TCP 6, UDP 17, tunelování protokolů, např. IP over IP (privátní sítě, IPv6 over IPv4), IPX over IP

CVIČENÍ: zachytávání a inspekce IP paketů

IP adresa

- každé síťové rozhraní počítače (síťová karta) může mít jednu nebo více **jednoznačných** IP adres
- přidělení adresy síťovému rozhraní staticky pomocí programu `ipconfig` (MS Windows) nebo `ifconfig`/`ip` (UNIX, GNU/Linux)

  CVIČENÍ: zjištění IP adresy síťového rozhraní a jeho změna

= číslo délky **4 B** (pro protokol IPv4), notace zápisu s hodnotami bytů v desítkové soustavě oddělenými tečkou, např. **158.194.80.13** = 10011110.11000010.01010000.00001101

IP adresa

Historie

- od počátku Internetu až do roku 1993: RFC 796
- = dvě části adresy: **adresa sítě** a **adresa uzlu (rozhraní)** v síti
- jaká část pro síť určují počáteční bity prvního bytu, dělení sítí do 5 (základních) **tříd**:
  - třída **A**: adresa začíná (bitem) 0, 1 byte pro síť, 126 sítí (s hodnotami prvního bytu) 1 až 126 (0 a 127 mají zvláštní význam), $2^{24} - 2$ uzlů (0 a 255 mají zvláštní význam)
  - třída **B**: začíná 10, 2 byty pro síť, 2^{14} sítí 128 až 191, $2^{16} - 2$ uzlů
  - třída **C**: začíná 110, 3 byty pro síť, 2^{21} sítí 192 až 223, 254 uzlů
  - třída **D**: začíná 1110, nedělí se, 2^{28} skupinových adres 224.0.0.0 až 239.255.255.255 (**IP multicast**, RFC 1112)
  - třída **E** (a další): začíná 1111, 2^{28} adres 240.0.0.0 až 255.255.255.254 původně rezervovaných pro speciální a experimentální účely, dnes již také přidělené

IP adresa

Historie

- speciální adresy:
  - celá = 0: tento uzel (= loopback, bez přidělené adresy)
  - uzel = 0: **adresa sítě**
  - síť = 0: uzel na této síti (nepoužívá se)
  - uzel samé 1: **všesměrová adresa sítě (network broadcast)**
  - samé 1 (255.255.255.255): **všesměrová adresa lokální sítě (local broadcast)**, nesměruje se
  - 127.cokoliv: programová (lokální, SW) **smyčka (loopback)**, typicky **127.0.0.1**, odeslaný paket „ihned přijde“

CVIČENÍ: zjištění všech uzlů na lokální síti pomocí programu `ping`

IP adresa

Dnes – Subsítě

- od roku 1993: RFC 1517–1520, sítě se nerozlišují podle tříd, ale podle **síťové masky**:
  - $=$ 4B číslo (notace IP adres), bity $= 1$ určují v IP adrese adresu sítě
  - $\rightarrow$ určení adresy sítě: bitový součin IP adresy a síťové masky
  - počet uzlů v síti $= 2^{(\text{počet 0 v masce})} - 2$
  - masky odpovídající třídám adres $=$ **standardní** síťové masky, pro třídu A **255.0.0.0**, pro třídu B **255.255.0.0**, pro třídu C **255.255.255.0**
  - **notace sítě spolu s maskou**: adresa sítě/maska, např. 158.194.0.0/255.255.0.0
  - v binárním vyjádření ji tvoří (de facto) zleva souvislá řada 1 $\rightarrow$ notace adresa sítě/počet 1 v masce, tzv. **CIDR formát** (Classless Inter-Domain Routing), např. 158.194.0.0/16

IP adresa

Dnes – Subsítě

= **část sítě určená maskou**: část adresy pro uzel rozdělena na část pro subsíť a pro uzel, síťová maska pokrývá část adresy pro síť i subsíť

– výjimka: síť s maskou /32 je adresou samostatného uzlu

– např. síť 158.194.0.0/16 může být rozdělena např. do 256 subsítí s adresami 158.194.0.0/24 až 158.194.255.0/24

! **nejednoznačnosti**: subsíť samé 0 (adresa uzlu samé 0) – adresa subsítě nebo celé sítě?, subsíť samé 1 (adresa uzlu samé 1) – všesměrová adresa subsítě nebo celé sítě (tj. všech subsítí)? → nepoužívají se

– síť může být na subsítě rozdělena pomocí **konstantní síťové masky** (všechny subsítě mají stejnou, viz příklad výše) nebo **variabilní síťové masky** (subsítě mají různou masku, např. 158.194.1.0/30, 158.194.80.0/20, 158.194.92.0/22) – POZOR na omezení adres subsítí!

– subsíť je možné opět pomocí „prodloužení“ masky opakovaně rozdělit do (sub)subsítí

IP adresa

Supersítě a autonomní systémy

- **supersíť** – síťová maska nepokrývá celou adresu sítě, duální k subsíti
  - použití pro **agregaci adres sítí**, výhodné pro směrování, administrativu přidělování adres apod.
  - např. síť 158.194.92.0/24 je součástí supersítě 158.194.0.0/16
- z hlediska dopravy IP paketů (směrování) se Internet dělí na tzv. **autonomní systémy (AS)** = supersítě spravované největšími poskytovateli internetového připojení, bloky IP adres v rámci AS přidělují regionální a lokální **Internet Registry**
  - např. síť 158.194.0.0/16 (UPOL-TCZ) je součástí autonomního systému AS2852 (CESNET2), který je součástí bloku AS2830 – AS2879 patřícího **RIPE NCC** (regionální Internet Registry pro Evropu a přidružené země)
  - přidělené bloky adres pro (super)sítě a autonomní systémy a informace o nich lze zjišťovat programem `whois`, např. `whois 158.194.80.13`, `whois AS2852`

Lokální síť (Intranet)

- **Intranet** = lokální síť (pro informační systém), obvykle uzavřená nebo s omezením provozu z vnější sítě dovnitř, příp. i ven
  - v síti (Internetu) musí být IP adresy jednoznačné, v lokální síti:
    - libovolné adresy (jednoznačné v rámci lokální sítě) a **NAT (Network Address Translation)** = **překlad adres** lokální sítě na adresy ve vnější síti a naopak – typicky na rozhraní směrovače do vnější sítě, zvláštní případ tzv. **maškaráda** = překlad na 1 adresu (směrovače)
    - **vyhrazené rozsahy IP adres** pro uzavřené podnikové sítě (RFC1918): **10.0.0.0/8** (třída A), **172.16.0.0/12** (třída B), **192.168.0.0/16** (třída C) – použití dle libosti, **nesměrují se**
    - v praxi vyhrazený rozsah + NAT
  - propojení dvou a více lokálních sítí:
    - směrovačem – jednoduché, nevýhoda komunikace přes směrovač
    - adresami sousední tvořící supersíť s kratší maskou, např. /23 pro /24
- **nečíslovaná síť**: „síť“ propojující dva směrovače (např. pomocí sériových linek), tvořící jeden „virtuální“ směrovač

Lokální síť (Intranet)

Dynamické přidělování IP adres

= oproti pevnému (statickému) podle potřeby při připojení uzlu do sítě

- v lokální síti dnes aplikační protokol **DHCP** nahrazující dřívější protokoly RARP a BOOTP
- v rozlehlé síti (typicky komutované telefonní) linkový protokol **PPP**

Směrování (routing)

- **směrování (routing)** = odeslání paketů na další směrovač nebo cílový uzel (next hop), popř. do lokální sítě s cílovým uzlem
- **předávání (forwarding)** = předávání paketů v rámci směrovače mezi jeho síťovými rozhraními, základ procesu směrování

– děje se (zpravidla) bez vědomí vyšších vrstev, např. aplikační, konfiguruje se parametry (jádra) OS, výjimkou je filtrace paketů při předávání

Obrázek: Obrázek průvodce 184→186(5)

Směrování (routing)

Předávání paketů a filtrace

- předávání paketů umožňuje uzlu pracovat jako **směrovač** = pokud paket není adresován jemu, odešle (předá) ho dále (jiným rozhraním), stejně jako vlastní odchozí pakety
- lze v OS povolit/zakázat za běhu, u MS Windows hodnota 1/0 v klíči `IpEnableRouter` v registru, u GNU/Linuxu v souboru `/proc/sys/net/ipv4/ip_forward`
- pakety mohou být **filtrovány** – nastavením filtračních pravidel OS nebo pomocí aplikačního programu, na základě IP záhlaví (adres), TCP/UDP záhlaví (portů, příznaků) nebo aplikačního protokolu
- filtrace se často provádí (a doporučuje se) i u koncových uzlů na vstupech jejich síťových rozhraní – posílení ochrany a bezpečnosti systému
- filtrace bývá významnou funkcí tzv. **firewallů** – programů či stanic (směrovačů) chránících systém uzlu nebo (lokální) síť před útoky

Směrování (routing)

Směrovací tabulky

– pro paket, který není určený přímo směrovači, se musí **rozhodnout**, kterým síťovým rozhraním jej odeslat dále (next hop)

→ rozhoduje se pomocí **směrovací tabulky** se směry (cestami, **route**):

síť/uzel	maska	next hop (gateway)	rozhraní	metrika, vlajky aj.
158.194.92.0	255.255.255.0	0.0.0.0	Ethernet 1	...
158.194.80.0	255.255.255.0	158.194.80.1	Ethernet 2	...
(127.0.0.0	255.0.0.0	127.0.0.1	loopback	...)
10.0.0.0	255.255.0.0	0.0.0.0	Virtual Eth.	...
...	...	...	...	...
0.0.0.0	**0.0.0.0**	158.194.254.66	Ethernet 3	...

– **setříděna sestupně podle adresy sítě** (1. sloupec) – více specifická (s delší maskou) má přednost před obecnější v případě stejných směrů pro paket

Směrování (routing)

Směrovací tabulky

- **rozhodování**:
  1. průchod tabulkou odshora dolů, log. **vynásobení cílové adresy paketu s maskou** v tabulce (2. sloupec)
  2. pokud se výsledek **rovná adrese sítě**, popř. uzlu (maska samé 1) v tabulce (1. sloupec), paket se odešle skrze rozhraní (4. sloupec) na další směrovač nebo cílový uzel (**next hop**, 3. sloupec), popř. do lokální sítě s cílovým uzlem (next hop = 0.0.0.0, tzv. **přímé směrování**), jinak další řádek
  3. poslední řádek (adresa sítě i maska = 0.0.0.0) = **výchozí (implicitní) směr** pro paket nevyhovující žádnému předchozímu záznamu (žádné síti), typicky směr do Internetu
  - agregace záznamů tabulky u supersítí a autonomních systémů

Směrování (routing)

Směrovací tabulky

- **naplnění tabulky**:
  - staticky (**statické směrování**) ručně, automaticky při konfiguraci síťového rozhraní OS (nejčastější) nebo pomocí managementu sítě (např. aplikační protokol SNMP)
  - dynamicky (**dynamické směrování**) z ICMP zpráv (změny směrování) nebo **směrovacími aplikačními protokoly**

– výpis tabulky pomocí programu `netstat`, výpis a (statická) editace správcem OS pomocí programů `route`/`ip` (UNIX, GNU/Linux), ‘‘`Směrování a vzdálený přístup`’’ (MS Windows Server), `ip route` (CISCO) apod.

CVIČENÍ: výpis a editace směrovací tabulky (např. výmaz a vrácení směru default) programy `netstat`, `route`, `ip` apod.

Směrovací protokoly

- aplikační protokoly k vytvoření směrů, tj. k **dynamické aktualizaci směrovacích tabulek** směrovačů, NE k vlastnímu procesu směrování
- dělení: IGP (v rámci AS) a EGP (výměna směrovacích informací mezi AS, směrovací politiky), RVP a LSP (podle použitého směrovacího algoritmu)

RVP (Routing Vector Protocols)

= algoritmus **DVA (Distance Vector Algorithm), Bellman-Fordův**: směrovač opakovaně **odešle** svou **směrovací tabulku** sousedním a z přijatých tabulek si do své dočasně (2-5 minut) doplní záznamy (vektory) pro neznámé sítě nebo s **menší vzdáleností** (metrikou, počet směrovačů na cestě) s navýšenou metrikou (typicky o 1), konec při max. metrice (např. 16) = nedostupná síť

- jednoduché, ale při výpadku připojení směrovače do sítě nebo v rozlehlejších sítích (při vyšší max. metrice) mohou tabulky **oscilovat** → nedoplňovat záznamy, které směrovač sám dříve odeslal
- např. RIP (pouze pro standardní masky), **RIP 2** (multicast 224.0.0.9), **RIPng** (pro IPv6), **IGRP**, **BGP**, program `routed`

Směrovací protokoly

LSP (Link State Protocols)

- algoritmus **LSA (Link State Algorithm)**: směrovač opakovaně ohodnotí (metrika) cesty k sousedním (např. podle odezvy) a jejich seznam spolu se sítěmi rozešle do celé rozlehlé sítě, ze získané **topologie celé sítě** si pak (dočasně) doplní/upraví záznamy v tabulce pro sítě na základě nejkratších cest vypočtených **algoritmem nalezení nejkratších cest v grafu** (**SPF, Shortest Path First, Dijkstrův**)
- rozdělení rozlehlejších sítí na **oblasti (směrovací domény)**, z více směrovačů na jedné síti se vybere jeden
- oproti RVP méně dat, **stabilnější**, pružnější, ale složitější konfigurace
- např. **OSPF** (páteřní oblast, autentizace, IPv6 aj.), IS-IS, **EGP**, program `gated`

Protokol ICMP

- Internet Control Message Protocol, RFC 777

= služební protokol IP pro **diagnostiku a signalizaci mimořádných** (chybových) **stavů**

– OS většinou nepodporují všechny **zprávy**, směrovače mohou z bezpečnostních důvodů nějaké zahazovat

– **ICMP pakety** obsaženy v paketech IP, záhlaví (8B): typ (1B), kód (1B), kontrolní součet (2B) a proměnná část (4B), a data

Obrázek: Obrázek průvodce 135(5)

CVIČENÍ: zachytávání a inspekce ICMP paketů generovaných programem `ping` nebo `traceroute/tracert`

Protokol ICMP

Echo

- typ 8 (žádost, request) a 0 (odpověď, reply), kód 0
- použití pro **testování dosažitelnosti uzlu** pomocí programu `ping` – měří a vypisuje i čas mezi žádostí a odpovědí, tj. čas k uzlu a zpět (**Round Trip Time, RTT**), a použité TTL
- pole Identifikátor (v proměnné části záhlaví) pro spárování žádosti a odpovědi

CVIČENÍ: zjištění vzdálenosti (počtu směrovačů, hopů) uzlu pomocí programu `ping` se změnou TTL

Protokol ICMP

Čas vypršel (Time exceeded)

- typ 11, kód 0 (TTL = 0 a IP paket bude zahozen) a 1 (IP paket nelze v určeném čase sestavit z fragmentů)
- zahozený/částečný IP paket (prvních 64 B) v datové části ICMP paketu
- použití (kód 0) pro **zjištění cesty** (směrovačů) **k uzlu** pomocí programu `traceroute/tracert`:
  1. na cílový uzel odeslána ICMP žádost Echo nebo UDP datagram (`traceroute`, port lze nastavit) s TTL = 1
  2. první směrovač na cestě signalizuje zahození paketu (sníží TTL na 0)
  3. získání adresy směrovače a změření času od odeslání k přijetí signalizace (čas ke směrovači a zpět, RTT), výpis obojího
  4. toto třikrát, pak s TTL = 2 (zahodí druhý směrovač) atd. až do přijetí ICMP odpovědi Echo nebo signalizace nedoručitelného IP paketu (kód 3) od cílového uzlu

CVIČENÍ: zjištění cesty (směrovačů) k uzlu pomocí programu `traceroute/tracert`, zjištění autonomních systémů na cestě pomocí

Protokol ICMP

Nedoručitelný IP paket (Destination unreacheble)

- typ 3, signalizace odesilateli, pokud paket nemůže být předán dál nebo doručen a je zahozen
- zahozený IP paket (prvních 64 B) v datové části ICMP paketu
- **důvody** (kódy): nedosažitelná síť (0), uzel (1), protokol (2), UDP port (3), fragmentace zakázána, ale nutná pro další přenos (4), neznámá adresátova síť (6), uzel (7) atd.

Další

- **sniž rychlost odesílání** (typ 4, kód 0) – odesilateli signalizuje směrovač, který není schopen IP paket předat dál (je zahlcený)
- **změň směrování** (typ 5, kódy 0-3), **žádost+odpověď o směrování** (typy 9, 10, kód 0) – doporučení změny ve směrovací tabulce odesílatele (pro tento směr) nebo zjištění směrovačů (žádost na všeobecnou adresu, směrovače odpoví)
- . . . (mnoho)

Fragmentace

- linkové rámce mají omezenou velikost (jeden až dva, max. jednotky kB), maximální velikost dat v rámci = **MTU (Maximum Transfer Unit)**, např. u Ethernetu II 1500 B
- IP paket může být ale dlouhý až 64 kB $\rightarrow$ **fragmentace paketu**
- pokud je fragmentace zakázána (bitem DF v záhlaví IP paketu):
  - paket je zahozen (pokud nejde jinou linkou) a odesilateli je to signalizováno pomocí ICMP typu 3, kód 4 – využití v algoritmu zjištění nejmenší MTU na cestě k uzlu (**Path MTU Discovery, PMTUD**)
  - později byla tato signalizace doplněna o možnost informace o MTU linky (2 B proměnné části záhlaví ICMP paketu)
- zvyšuje režii přenosu dat $\rightarrow$ OS se snaží vytvářet pakety délky $\leq$ MTU, aby nebylo fragmentace potřeba

CVIČENÍ: zjištění nejmenší MTU k uzlu pomocí programu `ping` se zakázáním fragmentace a nastavením velikosti paketu (algoritmus PMTUD)

Fragmentace

= dělení IP paketu na fragmenty o celkové délce $\leq$ MTU linky, RFC 791

– **fragment** = samostatný **IP paket** se stejnou hlavičkou jako původní paket (s **identifikací fragmentu**), až na položky:

- celková délka = délka fragmentu ($\leq$ MTU)
- **posunutí fragmentu** – offset dat fragmentu v datové části původního paketu, tj. kolik dat původního paketu je v předchozích fragmentech, v jednotkách 8B
- **indikaci dalších fragmentů** (bit MF příznaků) – poslední fragment nemá nastavenu

Obrázek: Obrázek průvodce 144→145(5)

Fragmentace

- **skládání fragmentů** (se stejnou identifikací fragmentu a protokolem vyšší vrstvy) do původního paketu provádí **pouze příjemce paketu**! – nikdo jiný nemusí mít všechny fragmenty
- pokud příjemce nemůže paket sestavit, protože v určené době nemá všechny fragmenty (protože např. první byl na cestě odfiltrován, např. podle adresy vyššího protokolu), signalizuje to příjemci pomocí ICMP typu 11, kód 1
- mechanizmus umožňuje dále fragmentovat i fragmenty, směrovači na cestě

CVIČENÍ: zachytávání a inspekce IP fragmentů generovaných např. programem `ping` s nastavením velikosti paketu

Volitelné položky IP záhlaví

– max. 40 B za povinnými položkami IP paketu

Obrázek: Obrázek průvodce 145→146(5)

– bit kopírovat znamená kopírování položek do všech fragmentů, jinak jen prvního

– číslo volby specifikuje typ volitelné položky, 0 pro poslední položku, 1 pro výplň záhlaví na násobek 4 B

- **zaznamenávej směrovače** (číslo 7): každý směrovač na cestě k příjemci zapíše IP adresu svého výstupního rozhraní (max. 9), příjemce je může zopakovat v odpovědi s touto volbou
- **zaznamenávej čas** (68): každý směrovač na cestě k příjemci zapíše čas (v ms od poslední půlnoci UTC, 4B) nebo čas a IP adresu svého výstupního rozhraní (8B, max. 4)

Volitelné položky IP záhlaví

- **explicitní směrování** (131, 137): explicitní zadání směrovačů, přes které má paket jít, **striktní** = zadání všech, směrovače upravují adresu příjemce paketu na adresu následujícího směrovače, z bezpečnostních důvodů (průnik do privátní sítě) bývají pakety s touto položkou na směrovačích **filtrovány**
- **upozornění pro směrovač** (148): informace pro směrovače na cestě k cílovému směrovači, že v paketu mohou být informace (ohledně směrování) užitečné i ně

– některé volby jsou implementované v programu `ping`

CVIČENÍ: zachytávání a inspekce IP paketů s volitelnými položkami v záhlaví generovaných např. programem `ping`

Protokoly ARP a RARP

ARP (Address Resolution Protocol)

- odchozí IP paket se vkládá do linkového rámce (např. Ethernet), jak se zjistí linková adresa příjemce? → **protokol ARP** (RFC 826)

= **zjištění linkové adresy** příjemce ze znalosti jeho IP adresy

- uzel vyšle **ARP paket žádosti** obsahující IP adresu příjemce na všesměrovou linkovou adresu a příjemce odpoví **ARP paketem odpovědi** (přímo odesilateli)
- ARP paket se vkládá přímo do linkového rámce, NE do IP paketu – **ARP** je protokol **nezávislý na IP**

Protokoly ARP a RARP

ARP (Address Resolution Protocol)

Obrázek: Obrázek průvodce 154

- **ARP paket**:
  - typ linkového protokolu: číslo použitého linkového protokolu, např. 1 pro Ethernet II, 6 pro Ethernet podle IEEE 802.3 (viz IANA)
  - typ síťového protokolu: stejná čísla jako v poli Protokol u linkového rámce, např. 0x800 pro IP
  - HS a PS: délka linkové a síťové adresy
  - operace: 1 pro ARP žádost, 2 pro ARP odpověď
  - linková adresa příjemce je v ARP žádosti nulová
  - v ARP odpovědi jsou oproti žádosti adresy prohozeny

Protokoly ARP a RARP

ARP (Address Resolution Protocol)

- **ARP cache**
  - = **tabulka síťová adresa – linková adresa**, naplněná staticky (manuálně) nebo dynamicky z příchozích linkových rámců se síťovými pakety a ARP odpovědí
  - – použita při zjišťování linkové adresy k síťové adrese
  - – omezená doba uchování dynamických položek (nepoužitých např. 2 minuty, maximální např. 10 minut), parametr OS
  - – pro manipulaci slouží program `arp`
- **proxy ARP**
  - ARP pakety se nesměrují (přísně vzato ARP není síťový protokol), **ARP** funguje **v rámci lokální sítě** (v dosahu linkového protokolu)
  - = **konfigurace směrovače**, kdy v odpovědi na ARP dotaz se síťovou adresou za směrovačem uvede směrovač jako linkovou adresu příjemce svoji linkovou adresu
  - ⇒ automatické nastavení směrování pro uzly v lokální síti přes směrovač

CVIČENÍ: zobrazení a manipulace s ARP cache, zachytávání a inspekce ARP paketů (po vymazání ARP cache)

Protokoly ARP a RARP

RARP (Reverse ARP)

= **zjištění síťové adresy** odesílatele ze znalosti své linkové adresy

– dříve použití u bezdiskových stanic bootovaných po síti, které žádají o svoji síťovou adresu na základě linkové, tu přidělí a v odpovědi sdělí **RARP server**

– stejný paket jako u ARP, pole operace: 3 pro RARP žádost, 4 pro RARP odpověď

– dnes překonán aplikačním protokolem **DHCP**

Protokol IGMP (IP multicast)

= služební protokol IP k **šíření IP paketů na skupinové adresy** (**IP multicast**) s více příjemci **v rámci lokální sítě** (TTL=1)

– IP multicast výrazně snižuje síťový provoz a zátěž odesílatele

– několik verzí, zde verze 2 (RFC 2236)

– pro každou skupinovou adresu udržuje směrovač lokální sítě **skupinu členů** (uzlů) a pokud je nějaká skupina neprázdná, směrovač šíří multicast pakety s adresou skupiny zvenku dovnitř lokální sítě

– uzel (aplikace na něm) požadující příjem multicast paketů vyšle **IGMP paket** s požadavkem na členství ve skupině dané skupinovou adresou

Obrázek: Obrázek průvodce 158

Protokol IGMP (IP multicast)

- **IGMP paket** obsažen v IP paketu:
  - typ: dotaz směrovače na členství ve skupině (11), požadavek na členství ve skupině (16), opuštění skupiny (17)
  - **MRT (Maximum Response Time)**: pouze u typu 11, čas (v desetinách s), do kterého se musí uzly znovu přihlásit do skupiny, jinak jsou vyřazeni
  - **skupinová IP adresa**: nula u dotazu typu 11 (adresuje všechny skupiny), jinak z **třídy D**, rozsah **224.0.0.0/24** je pro vyhrazené účely (např. 224.0.0.1 je všeobecná pro všechny uzly, 224.0.0.2 pro všechny směrovače atd.)
- více směrovačů na lokální síti: dva režimy směrovače – **dotazovač** (posílá dotazy) a **posluchač** (dotazovač, který se přepnul, pokud detekoval v lokální síti dotazy směrovače s vyšší adresou, jen poslouchá)

Protokol IGMP (IP multicast)

Mapování síťových na skupinové linkové adresy

- „mapování“ jednoznačných IP adres (unicast) → ARP, mapování všesměrové → všesměrová linková adresa
- síťová karta zpracovává (v normálním, ne promiskuitním, režimu) pouze jí adresované a všesměrové rámce, navíc pak **skupinové rámce**, o které **zažádá síťová vrstva**
- Ethernet:
  - skupinová MAC adresa: nejnižší bit prvního bytu = 1
  - první tři byty MAC adresy pro výrobce – IANA má 00:00:5E, polovina jejího rozsahu pro **skupinové adresy**, prefix **01:00:5E**

  = **nejednoznačné mapování** 28 bitů skupinové IP adresy do 23 bitů skupinové MAC adresy: IP adresy lišící se pouze v nevyšších 5 bitech (po prefixu skupinových adres), např. 224.0.1.1 a 225.0.1.1, mapovány na stejné linkové adresy

Obrázek: Obrázek průvodce 162

⇒ pakety s nechtěnou IP adresou musí odfiltrovat síťová vrstva

IP multicast

IP multicast mimo lokální síť (v Internetu)

= šíření multicast paketů Internetem od odesílatele k příjemcům ve více lokálních sítích – poměrně **složitá záležitost**, cíl **zamezit nekontrolovanému lavinovitému duplikování paketů** v Internetu

– **úpravy směrovacích protokolů** pro výměnu směrovacích informací mezi směrovači – protokoly např. **DVMRP, MOSPF, MBGP**

– problémy se škálovatelností (počty příjemců v milionech), **aktivní výzkum**

- dříve experiment s **MBONE (Multicast Backbone)** = vybrané směrovače („jádro Internetu“) zabezpečující šíření multicast paketů pomocí tunelů
- dnes protokoly **PIM (Protocol Independent Multicast)** konstruující **distribuční strom multicastu** (pro každou skupinovou adresu), varianty Sparse Mode (SM), Source Specific Mode (SSM), Bidirectional Mode

– využití v **distribuci multimediálního obsahu (streaming)**, ne obecně jako způsob přenosu libovolných dat v Internetu

Protokol IP verze 6 (IPv6)

- $\sim$ **“IP nové generace”, IPng**, vyvíjen od roku 1991, 1995 RFC-1883, dnes RFC-2460 (základ + přidružená RFC)
- – odstraňuje **nedostatky IPv4**: řešení problému adresace, dynamické konfigurace, podpory bezpečnosti, mobility uzlů, multimédií aj.
- $=$ nejen **zvětšení IP adresy, nový pohled na IP paket** (revize):
  - zjednodušení záhlaví – přesun málo využívaných základních položek do (zřetězených) volitelných: pro směrování, fragmentaci, autentizaci aj.
  - (bezstavová) **automatická konfigurace uzlů**
  - bezpečnost – **autentizace a šifrování** na úrovni síťové vrstvy
  - podpora **mobility uzlů** – se snahou o zachování TCP spojení při přechodu uzlu ze sítě do sítě (!)
  - podpora multimédií – třídy dat (včetně real-time komunikace), směrování toku a ne jednotlivých paketů
  - ...

Protokol IP verze 6 (IPv6)

IPv6 paket

Obrázek: Obrázek průvodce 196→208(5)

= 40 B základní záhlaví + nepovinná rozšíření různé délky, data, max. 64 kB, ale možnost rozsáhlého paketu v rozšířeních

- **třída dat**: specifikace priority dat pro rozhodování o zahození paketu při zahlcení sítě, hodnoty 0 až 7 pro klasický provoz (datové přenosy, pošta, interaktivní atd.), 8 až 15 pro přenosy v reálném čase (multimédia)

Protokol IP verze 6 (IPv6)

IPv6 paket

- **identifikace toku dat**: spolu s adresou odesilatele jednoznačně identifikuje datový tok, pro potřeby **směrování** – řešení směrování jen u prvního paketu toku, ne u každého (na základě jen adresy příjemce u IPv4), nebo k **zajištění šířky pásma** – prioritní FIFO paketů na směrovači místo obyčejné (jako u IPv4), protokol RSVP
- **další záhlaví**: typ následujícího záhlaví **nepovinného rozšíření** IPv6 (včetně typu 59 pro žádné) nebo protokolu vyšší vrstvy, např. TCP (6), UDP (17), IP (v IP, 4)
- **počet hopů**: $\sim$ TTL u IPv4, k zahazovaní zatoulaných paketů nebo k nalezení nejkratší cesty (zvyšování TTL, obdoba traceroute)

Protokol IP verze 6 (IPv6)

IP adresa

- délka 16 B (128 b), tři typy:
  - jednoznačná síťového rozhraní (unicast)
  - skupinová (multicast) – zvláštní případ všeobecná (broadcast)
  - skupinová **anycast** = paket doručen jen nejbližšímu z adresátů skupiny, adresy z rozsahu unicast adres, např. subnet-router, DNS query anycast
- notace zápisu s až čtveřicemi šestnáctkových číslic oddělenými dvojtečkou, např. **2001:718:1401:50:0:0:0:0d**, nebo častěji zkrácená pomocí zdvojené dvojtečky (pouze jednou, nahrazuje sekvenci 0), např. **2001:718:1401:50::0d**, nebo i s posledními čtyřmi byty v notaci adresy IPv4 (tzv. kompatibilní adresy), např. **FE80::158.194.80.13**
- notace sítě spolu s maskou (**prefix**): prefix adresy pro síť/počet 1 v (binární) masce

Protokol IP verze 6 (IPv6)

IP adresa

- **rozdělení na poloviny** (RFC 2373, 2450): adresa sítě (64 b) a adresa uzlu (rozhraní, 64 b)
- **adresa sítě**: obdobně jako u IPv4, **globální prefix** (45 b za prvními třemi bity) pro Internet Registry a autonomní systémy, např. pro RIPE 2001:0600::/29 až 2001:07F8::/29, dále poskytovatele (supersítě) a organizace (sítě), pak pro subsítě (16 b)

Obrázek: Obrázek průvodce 214→227(5)

- globálně jednoznačné (unicast) adresy pro Internet: (zatím) **2000::/3**, bloky /23 až /12 pro Internet Registry

Protokol IP verze 6 (IPv6)

IP adresa

Obrázek: Obrázek průvodce 215→227(5)

- **adresa rozhraní**: vlastní, podle **IEEE EUI-64** = **MAC adresa** podle IEEE 802, kde doprostřed se vloží **0xFFFE** a nastavení druhého bitu prvního byte, např. pro 00:02:B3:BF:30:EA je 202:B3FF:FEBF:30EA, náhodně dynamicky generovaná (**Privacy Extensions**)
  - **bezstavová autokonfigurace, SLAAC**: adresa rozhraní podle IEEE EUI-64 nebo náhodná „samopřidělená“ na základě **oznámení směrovače (router advertisement, RA)** s adresou sítě (prefixem), obdoba 169.254.0.0/16 u IPv4, zabezpečení RA Guard, SEND (aymetrická kryptografie, šifrovaná adresa), access listy na přepínači
  - **DHCPv6**: bezstavové – SLAAC + další info (DNS servery na LAN, domény apod.) z DHCP serveru, stavové – jako DHCP pro IPv4, ale ne výchozí brána LAN (sic!), identifikace uzlu pomocí DUID místo MAC rozhraní – pro uzel, nezávislost na MAC, 3 typy, zabezpečení DHCP Snooping

Protokol IP verze 6 (IPv6)

IP adresa – speciální adresy:

- celá 0: nespecifikovaná, rozhraní ještě nebyla přidělena adresa
- ::1/128: **loopback**
- FE80::/10: automatické v rámci lokální sítě nebo linkově propojených sousedů (**link-local** unicast), nesměrují se, adresa rozhraní automaticky „samopřidělená“ podle IEEE EUI-64, pro objevování sousedů (viz dále), oznámení směrovače, směrovací protokoly aj.
- FC00::/7: unikátní (síťová adresa, prefix, z data a MAC rozhraní) v rámci organizace (**unique-local** unicast), použití u intranetu, nesměrují se, dříve FEC0::/10 – privátní v rámci organizace (**site-local** unicast), obdoba vyhrazených rozsahů u IPv4 (10.0.0.0/8 atd.)
- FF00::/8: skupinové adresy (**multicast**), první 4 bity z druhého byte specifikují rozsah skupiny, např. 1 v rámci uzlu, 2 lokální sítě, 5,8 organizace, E globální, vyhrazené adresy, např. FF02::1 pro všechny uzly = **broadcast**

Protokol IP verze 6 (IPv6)

IP adresa – speciální adresy:

- **přechodové z IPv4**: tunelovací (IPv6 v IPv4) 2002:AB:CD::/16 **6to4** – pro IPv4 $(A.B.C.D)_{16}$ adresu rozhraní, 6to4 relay směrovač (např. NIC.cz) na anycast adrese 192.88.99.1, **ISATAP** – relay ve firemní síti (v DNS), **Tunel Broker** – veřejný relay (HE, SixXS), 2001::/32 **Teredo** – pro uzly za NAT, UDP zapouzdření, veřejný Teredo server (např. Microsoft), mapování ::FFFF:a.b.c.d na IPv4 a.b.c.d. (**SIIT**, virtuální IPv4 rozhraní), překlady IPv6 (např. 64:FF9B::/96) na IPv4 **NAT64 & DNS64**, aj.
- 2001:db8::/32: pro dokumentace (obdobné i u IPv4)
- ...

Protokol IP verze 6 (IPv6)

Nepovinná rozšíření

Obrázek: Obrázek průvodce 199→211(5)

- **záhlaví rozšíření**: typ následujícího záhlaví (tvoří řetězec použitých položek na rozdíl od všech u IPv4), délka záhlaví, data
- **informace pro směrovače** (typ 0): informace = volby (pole typ, délka, hodnota, např. rozsáhlý paket délky až 4 GB, typ 194)
- **směrovací informace** (43): **explicitní směrování, hop-by-hop** – pole počet směrovačů, maska striktního směrování (bit = 1 = sousední směrovač), adresy směrovačů a příjemce, 2007 zrušeno

Protokol IP verze 6 (IPv6)

Nepovinná rozšíření

- **záhlaví fragmentu** (44): fragmentovat může pouze odesílatel (na rozdíl od IPv4, algoritmus PMTUD), pole posunutí fragmentu (hodnota v jednotkách 8B), indikace dalších fragmentů, identifikace fragmentu
- **autentizace** (51, protokol AH) a **bezpečnost/šifrování** (50, ESP): integrita a autentizace (místo kontrolního součtu, MD5 ze sdíleného tajemství a paketu), šifrování odesílatelem nebo směrovači, použití v IPSec, poslední záhlaví

– uspořádání od těch pro směrovače po ty pro koncový uzel

Protokol IP verze 6 (IPv6)

Protokol ICMP verze 6

= nepovinné rozšíření IP záhlaví, typ 58, RFC 2463
- stejně jako u IPv4 pro **signalizaci** chybových stavů a **diagnostiku**
- pole typ, kód, kontrolní součet a tělo

- např. echo (žádost, odpoveď), čas vypršel, nedoručitelný paket (není směr, adresa, administrativně), změň směrování, žádost+odpověď o směrování apod.
- **Neighbor discovery protokol, NDP**: objevování sousedů $\sim$ překlad IPv6 adresy na linkovou adresu (místo ARP a RARP u IPv4) = žádost a oznámení o linkové adrese (neighbor solicitation a advertisement), žádost o a oznámení směrovače (router solicitation a advertisement) – adresa sítě (prefix) a výchozí brány LAN (nově i DNS serverů), povolení SLAAC, aj., zasílaná na skupinovou adresu LAN (speciální FF02::1:FF00:0/104)

CVIČENÍ: zachytávání a inspekce IPv6 paketů, zjištění IPv6 adresy síťového rozhraní

Bezpečnost protokolu IP

IPv4

- **neřeší**, naopak např. některé volitelné položky (explicitní směrování) mohou být nebezpečné
- pouze kontrolní součet záhlaví – snadné přepočítat po modifikaci paketu
- útoky: podvržení IP adresy odesílatele a příjemce (**IP spoofing**), zahlcení sítě (např. flood ping) a odepření služby (**Denial of Service**, **DoS**)

→ řešení: **filtrace** (některých ICMP paketů, paketů s volitelnými položkami atd.), **šifrování** – privátní sítě (intranet, s překladem adres), DHCP Snooping aj.

IPv6 – útoky + řešení jako u IPv4

- autentizace (protokol AH) a šifrování (protokol ESP) v dalších záhlavích → **IPSec**
- zabezpečení autokonfigurace – SEND

Bezpečnost protokolu IP

Firewall

= oddělení vnitřní sítě (intranetu) od vnější (Internetu), ochrana systému uzlu před sítí

– služby: **filtrace provozu**, kontrola adres, překlad adres (NAT) – na základě IP záhlaví (a dále záhlaví vyšších protokolů), aplikační brána (proxy, protokol SOCKS), logování a detekce útoků (IDS, IPS)

– provozován na hraničních směrovačích (bráně) mezi sítěmi nebo na klientských počítačích

– nastavení pravidel (fitračních aj.) OS nebo pomocí aplikačního programu

- **demilitarizovaná zóna (DMZ)** – část sítě s počítači dostupnými z vnitřní (chráněné) i vnější sítě, např. aplikační (proxy) servery

Bezpečnost protokolu IP

Překlad adres (Network Address Translation, NAT) (RFC 1631)

= překlad IP adres paketů z vnitřní sítě (intranetu) na IP adresy vnější sítě (Internetu) a naopak

- **SNAT (Source NAT)** = překlad IP adresy odesílatele, **DNAT (Destination NAT)** = překlad IP adresy příjemce

– poskytuje skrytí vnitřní sítě, využití také při spojení více intranetů se stejným rozsahem adres

– provozován na hraničních směrovačích (bráně) mezi sítěmi, typicky v rámci firewallu

- **maškaráda** = SNAT na IP adresu hraničního směrovače ve vnější síti, překlad i (zdrojového) portu transportní vrstvy (**NAPT (Network Address and Port Translation)**)

– zasahuje i do vyšších vrstev, transportní (překlad portů) i aplikační (porozumění aplikačnímu protokolu)

Bezpečnost protokolu IP

IPSec (Internet Protocol Security) (RFC 2401 – 2412)

- – původně v rámci prací na IPv6 (jeho povinná součást), backportován i pro IPv4
- = zabezpečení komunikace mezi počítači (koncovými síťovými rozhraními) na úrovní síťové vrstvy ⇒ **bezpečná síť**
- = **autentizace** komunikujících rozhraní a **šifrování** IP paketů
- – poměrně komplikovaný protokol, závislý na architektuře TCP/IP
- – funkce: správa šifrovacích klíčů (certifikační autority, autentizace (digitální podpis, hashe), šifrování (DES, RSA)
- záhlaví IPSec mezi záhlavím IP a daty paketu, položky pro autentizaci (AH) a šifrování (ESP), viz IPv6, dále protokoly pro výměnu klíčů **ISAKMP** a **IKEY**
- – režimy:
  - transportní – šifrování datové části IP paketu, mezi koncovými uzly
  - tunelovací – tunelování IP sítě v IP síti, zapouzdření šifrovaných IP paketů do nových IP paketů (IPSec over IP), tunel mezi směrovači nebo vzdáleným uzlem a hraničním směrovačem sítě

Sítě WAN na bázi IP

– původní představa WAN jako propojení LAN pomocí směrovačů a pronajatých okruhů ATM nebo Frame Relay přestala stačit

→ páteřní sítě přímo na bázi IP, **homogenní IP síť**

- **IP over Fiber** = přenos IP prostřednictvím optických sítí, varianty
  - systém **SONET/SDH** – převod el. signálů na optické, IP over ATM (vysoká režie, 622 Mb/s), IP over SONET/SDH (IP pakety v PPP rámcích v kontejneru SONET/SDH, synchronní přenos, 155 Mb/s)
  - **IP over DWDM** (případně ještě se SONET/SDH) – transparentní přenos paketů bez převodu signálu a formátování do rámců, až 10 Gb/s, kombinace s MPLS (MPλS)
- **virtuální privátní sítě (VPN)**: virtuální IP síť v rozlehlé IP síti
- **MPLS**: přepínané IP sítě místo hop-by-hop sítí (se směrovači), na základě tzv. návěští po definované cestě (zaručení atributy spojení, QoS, VPN atd.)
- **QoS**: zabezpečení kvality přenosu pomocí rezervace zdrojů/upřednostnění paketů (InetServ/DiffServ), protokol RSVP

Virtuální privátní sítě (VPN)

= privátní sítě virtuálně v rozlehlé transportní síti (Internetu), často jako propojení (privátních) sítí nebo uzlu a (privátní) sítě, nahrazuje pronajaté telekomunikační okruhy

- privátní adresace – nutno řešit oddělení privátních sítí např. pomocí filtrace a NAT

→ **tunelování**

= zapouzdření paketů nebo celých rámců vnitřní sítě do paketů transportní sítě

– vytváření **tunelů** = (dvoubodových) logických spojení mezi uzly virtuální sítě, propojení do transportní sítě = **VPN gateway**

– zabezpečení tunelů a oddělení sítí: autentizace, šifrování

- tunelování linkové vrstvy (zapouzdřovány rámce): protokoly **PPTP**, **L2TP** (PPP rámce v IP, Frame Relay, ATM, autentizace, šifrování, komprese, vícebodové tunely)
- tunelování síťové vrstvy (zapouzdřování paketů): **IP over IP**, protokoly **GRE** (původní, dvoubodové tunely) a **IPSec**

– oddělení IP sítí – např. přepínaní, MPLS (MPλS)

Transportní vrstva

Transportní protokoly

"Proč dva protokoly?"

- síťové protokoly přepravují data mezi libovolnými **uzly** (počítači) v síti, adresují síťová rozhraní uzlu

→ přepravují data mezi dvěma (původně) **aplikacemi** běžícími na uzlech, adresují aplikaci na uzlu

→ zprostředkovávají **transparentní spojení** s požadovanou kvalitou mezi více aplikacemi v rámci síťových zařízení (uzlů)

Transportní protokoly

Služby

- **spojovaná (connection oriented)**:
  - mezi aplikacemi navázáno **spojení** (vytvořen virtuální okruh daných parametrů), s **plně duplexní** výměnou dat
  - (typicky) ztracená nebo poškozená data znovu vyžádána – **„spolehlivá“ služba**
  - integrita dat zabezpečena kontrolním součtem
  - zpracovává **souvislý proud/tok (uspořádaných) dat** od vyšší vrstvy (**stream**)
- **nespojovaná (connectionless)**:
  - nenavazuje spojení
  - data odeslána, (typicky) nezaručuje se doručení ani znovuzasílání ztracených nebo poškozených dat (ponecháno na vyšším protokolu) – **“nespolehlivá” (datagramová) služba**
  - integrita dat zabezpečena kontrolním součtem
  - zpracovává **(nesouvislé) části dat** od vyšší vrstvy (datagramy), rozdělení toku dat na datagramy řeší vyšší vrstva

Transportní protokoly

Port

= identifikátor aplikace (aplikace jich může používat víc), transportní adresa

= číslo délky 2 B, 0 až 65535

– porty 0 – 1023 jsou tzv. **privilegované** (může je použít pouze privilegovaná aplikace, např. systémová služba nebo privilegovaného uživatele), ostatní **neprivilegované** (může použít kdokoliv, pokud je volný)

– pro běžné služby (aplikační protokoly) Internetu všeobecně známá „standarní“ (**well-known**) čísla portů přidělovaná IANA, privilegovaných i neprivilegovaných

CVIČENÍ: zjištění čísel portů nejznámějších služeb Internetu, např. jmenné (aplikační protokol DNS), vzdáleného přihlášení (Telnet, SSH), přenosu dat (FTP(S), SMB), poštovní (SMTP, POP3(S), IMAP(S)), webové (HTTP(S)), a LAN (DHCP, SNMP)

Transportní protokoly

- aplikace jednoznačně určena: síťovou (IP) adresou, číslem portu a transportním protokolem (TCP/UDP), tzv. adresa **socketu** (síťového rozhraní **Socket API**)

Datagram/Segment

= základní jednotka přenosu, transportní paket/datagram/segment, vkládán do síťového paketu

- obsahuje část (toku) dat od odesílatele k příjemci od vyšší vrstvy

→ **segmentace** = rozdělení toku dat na části zapouzdřené do segmentů

Obrázek: Obrázek průvodce 219→231(5)

- max. délka = max. délka síťového paketu (64 kB u IP) - délka jeho záhlaví
- obsahuje záhlaví s porty příjemce a odesílatele + data

Transmission Control Protocol (TCP)

- RFC 962
- IP protokol poskytuje datagramovou (nespojovanou) „nespolehlivou“ službu, bez vyžadování opakování přenosu paketů, nanejvýš signalizace nemožnosti doručení (ICMP, nepovinná, potlačovaná)

→ poskytuje spojovanou „spolehlivou“ službu, řeší:

- navázání, udržování a ukončení **plně duplexního spojení**
- adaptivní přizpůsobení parametrů protokolu podle stavu spojení
- zaručení správného **pořadí dat**
- potvrzování přijetí dat (tzv. **pozitivní potvrzování**)
- vyžádání **opakování přenosu** ztracených nebo poškozených dat
- **řízení toku dat** a **předcházení zahlcení sítě** pomocí časových prodlev, opakovaného odeslání a potvrzení přijetí dat, bufferů a **posuvného okna** a **okna zahlcení**

- nezávislý rozsah portů pro TCP a UDP, TCP porty označeny **číslo/tcp**

TCP segment

Obrázek: Obrázek průvodce 219→232(5)

- záhlaví 20 B povinných položek + volitelné položky, data
- **identifikace spojení** (v Internetu) = zdrojový a cílový port, zdrojová a cílová IP adresa, transportní protokol (TCP)
- **pořadové číslo odesílaného bytu**: pořadové číslo 1. bytu dat z odesílaného toku dat (spojení) v segmentu, číslování při navázání spojení začíná od náhodného čísla (ISN, Initial Sequence Number), po dosažení $2^{32} - 1$ opět od 0 – pro zajištění správného **pořadí dat**
- **pořadové číslo přijatého bytu**: pořadové číslo následujícího bytu, který má být přijat – pro zajištění **pozitivního potvrzování** a opakování přenosu dat
- **délka záhlaví**: v jednotkách 4 B, max. 60 B

TCP segment

- příznaky:
  - **CWR, ECN** – pro (volitelné) oznámení zahlcení sítě, viz dále, bez zahazování dat, tzv. ECN (Explicit Congestion Notification), v kombinaci s IP (2 bity u položky TOS záhlaví IP paketu)
  - **URG** – segment nese naléhavá data, která má příjemce zpracovat přednostně (out of band data, použití vyjímečně, např. u Telnetu pro příkazy)
  - **ACK** – signalizace správného pořadového čísla přijatého bytu, tj. potvrzení správného přijetí bytů segmentu až do tohoto čísla - 1 = **pozitivní potvrzování**
  - **PSH** – segment obsahuje aplikační data, použití není ustáleno
  - **RST** – **odmítnutí** navazovaného TCP **spojení**
  - **SYN** – nová sekvence číslování odesílaných bytů, pořadové číslo odesílaného bytu je číslo 1. bytu toku dat (ISN), nastaven u 1. segmentu při **navazování spojení**
  - **FIN** – ukončení odesílání dat (dalších, tj. s výjimkou opakování přenosu dat), **ukončení spojení** **pro daný směr přenosu dat**
- **délka okna**: počet bytů, které je příjemce schopen přijmout – předcházení zahlcení přijímače v rámci **řízení toku dat**

TCP segment

- **kontrolní součet**: počítaný z některých položek IP záhlaví (IP adresy odesílatele a příjemce, 1 B bin. nul, protokol vyšší vrstvy, celková délka IP paketu), záhlaví TCP segmentu a dat (plus případně 1 B bin. nul výplně na sudý počet bytů), tzv. **pseudozáhlaví** – zajištění integrity dat

Obrázek: Obrázek průvodce 234(5)

- **ukazatel naléhavých dat**: počet bytů odesílaných naléhavých dat od začátku dat v segmentu nebo offset začátku naléhavých dat (závisí na aplikaci, např. příkazy Telnetu), pouze při příznaku URG

CVIČENÍ: zachytávání a inspekce TCP segmentů

Volitelné položky TCP záhlaví

– max. 40 B za povinnými položkami TCP segmentu

Obrázek: Obrázek průvodce 225→235(5)

- typ 0 pro poslední položku, 1 pro výplň záhlaví na násobek 4 B
- **max. délka segmentu (MSS)**, typ 2: max. délka dat přijímaných segmentů, dohodnutá stranami při navazování spojení, jen s příznakem SYN
- **zvětšení okna**, typ 3: délka bitového posunu doleva délky okna
- **povolení SACK** a **SACK (Selective ACK)**, typy 4 a 5: pro selektivní potvrzování segmentů mimo pořadí (vzhledem k pořadí dat), ukládání segmentů na straně příjemce do bufferu
- **časové razítko** a **echo časového razítka**, typ 8: echo je zopakování razítka z posledního přijatého segmentu, pro detekci starého zatoulaného segmentu při dlouhých oknech (stovky MB)
- a další, např. pro čítač spojení

Navazování spojení

- jedna strana spojení navazuje, druhá jej přijme nebo odmítne
- model **klient/server** (z hlediska aplikační vrstvy) – klient navazuje, server očekává a případně přijímá

- protokol TCP umožňuje navazovat spojení současně v obou směrech (v praxi ne příliš využívané) – POZOR!, neplést s **obousměrným přenosem dat** v rámci jednosměrně navázaného spojení!

→ obě strany **otevřou port** (pomocí **socketu**), klient v tzv. **aktivním režimu** (navázání spojení), server v tzv. **pasivní režimu** (očekávání spojení)

  - cílový port (na serveru) je daný aplikací
  - zdrojový port (na klientu) typicky náhodně vybrán OS z volných neprivilegovaných (≥ 1024)

Navazování spojení

Třífázový (Three-Way) handshake

Obrázek: Obrázek průvodce 226→238(5)

1. klient odešle segment (bez dat) s příznakem **SYN**, náhodně vygenerovaným pořadovým číslem odesílaného bytu jako startovacím číslem (**fiktivního**) 1. bytu dat od klienta (ISN) a navrhovanou max. délkou přijímaných segmentů (MSS)
2. server odešle segment (bez dat) s příznaky **SYN** a **ACK**, náhodně vygenerovaným ISN a navrhovanou MSS pro směr od serveru, pořadové číslo přijatého bytu je klientovo **ISN + 1** (potvrzuje přijetí předchozího segmentu, fiktivního 1 byte dat, od klienta)
3. klient odešle segment (bez dat) s příznakem **ACK**, pořadové číslo odesílaného bytu je klientovo ISN + 1 (další fiktivní byte dat, který server očekává), pořadové číslo přijatého bytu je serverovo ISN + 1 (potvrzuje přijetí předchozího segmentu, fiktivního 1 byte dat, od serveru)

Navazování spojení

- po navázání spojení, tj. příjmu segmentu s příznakem ACK oběma stranami, lze zasílat oběma směry data (datové segmenty s příznaky ACK a PSH) nebo jen **potvrzovací segmenty** (s příznakem ACK)
- první segment s příznakem SYN nepotvrzuje žádná přijatá data, tj. neobsahuje příznak ACK a pole pořadové číslo přijatého bytu není platné (bývá vyplněno bin. nulami)
- navrhované **MSS** je $\leq$ **MTU**, aby se zamezilo IP fragmentaci, pro Ethernet II 1460, Ethernet 802.3 1452

CVIČENÍ: zachytávání a inspekce TCP segmentů při navazování spojení, rozbor třífázového handshake

Navazování spojení

Stavy spojení při jeho navazování:

Obrázek: Obrázek průvodce 228→239(5)

- LISTEN – stav serveru, čekání na navázání spojení ze strany klienta
- SYN_SENT – na straně klienta, po odeslání prvního segmentu (s příznakem SYN), tj. navazování spojení
- SYN_RCVD – na straně serveru, po obdržení prvního segmentu (s příznakem SYN), tj. obdržena žádost o spojení
- ESTABLISHED – na obou stranách, po obdržení prvního segmentu s příznakem ACK, tj. spojení navázáno (pro přenos dat ve směru od strany, která segment **obdržela**)

Všechna spojení a jejich stavy lze zobrazit např. programem `netstat`.

CVIČENÍ: výpis všech spojení na z/do počítače, identifikace IP adres a portů (aplikací) stran a stavů spojení, např. pomocí programu `netstat`

Ukončování spojení

– ukončit/uzavřít spojení může libovolná strana, klient i server

Obrázek: Obrázek průvodce 229→240(5)

1. 1. strana odešle segment (možno i s daty) s příznakem **FIN** (vedle ACK), tzv. **aktivní uzavření spojení**, pak již nemůže odesílat datové segmenty (s příznakem PSH)
2. 2. strana odešle segment (potvrzovací, možno i s daty) bez příznaku FIN (jen s **ACK**), tzv. **pasivní uzavření spojení**, může dál odesílat datové segmenty 1. straně tzv. **polouzavřeným spojením**
3. 2. strana odešle segment (možno i s daty) s příznakem **FIN** (vedle ACK), tzv. **úplné uzavření spojení**
4. 1. strana odešle potvrzovací segment (bez dat, s příznakem **ACK**)

Ukončování spojení

- 2. krok je možné vynechat, při oboustranném uzavření spojení
- segment s příznakem FIN bez dat se potvrzuje (segmentem s ACK) jakoby měl 1 byte (fiktivních) dat

CVIČENÍ: zachytávání a inspekce TCP segmentů při ukončování spojení, rozbor sekvence segmentů ukončujících spojení

Stavy spojení při jeho ukončování:

Obrázek: Obrázek průvodce 230→241(5)

- FIN_WAIT1 – na 1. straně, po odeslání segmentu s příznakem FIN, tj. aktivní uzavření spojení
- CLOSE_WAIT – na 2. straně, po obdržení segmentu s příznakem FIN a odeslání segmentu jen s příznakem ACK (bez FIN), tj. pasivní uzavření spojení

Ukončování spojení

- FIN_WAIT2 – na 1. straně, po obdržení (potvrzovacího) segmentu bez příznaku FIN, po 11,25 min. nečinnosti polouzavřeného spojení (tj. bez přijetí segmentu) přechází do stavu CLOSED
- LAST_ACK – na 2. straně, po odeslání segmentu s příznakem FIN, tj. úplné uzavření spojení
- TIME_WAIT – na 1. straně, po obdržení segmentu s příznakem FIN a odeslání potvrzovacího segmentu, protože potvrzovací segment není potvrzován, po 30 s – 2 min. přechází do stavu CLOSED, kvůli možnosti opakování potvrzovacího segmentu po jeho vyžádání 2. stranou (při neobdržení)
- CLOSED – na obou stranách, na 2. straně po obdržení potvrzovacího segmentu

CVIČENÍ: výpis všech spojení na z/do počítače, identifikace IP adres a portů (aplikací) stran a stavů spojení, např. pomocí programu `netstat`

Odmítnutí spojení

- pokud cílový port na straně příjemce není otevřen (např. neběží aplikace serveru, nebo jsou segmenty zahazovány firewallem), klient, bez odpovědi serveru, po vypršení časového intervalu **opakuje** požadavek na **navázání spojení** (1. segment s příznakem SYN) do vypršení celkového času nebo počtu pokusů → časová prodleva

= kdykoliv zaslání segmentu s příznakem **RST** (bez dat) → **okamžité uzavření spojení** (v obou směrech) a přechod do stavu CLOSED na obou stranách

- použití např. u neúspěšného vytvoření šifrovaného kanálu u SSL/TLS
- použití také pro **rychlejší ukončení spojení**: nastavení příznaku RST místo FIN v 3. (nebo i 1.) segmentu při ukončování spojení, nebo po 4. segmentu ještě 2. strana odešle potvrzovací segment s příznakem RST, pro ušetření 1. straně čekání ve stavu TIME_WAIT

Ztráta segmentu (řízení toku dat)

Odesílatel:

- má definovaný časový interval pro příjem potvrzovacího segmentu od příjemce (retransmission timeout)
- při ztrátě nebo poškození segmentu (odeslaného nebo potvrzovacího) po vypršení intervalu nebo příjmu tří opakovaných stejných potvrzení od příjemce (viz dále) **opakuje odeslání segmentu**
- hodnota intervalu se dynamicky mění podle stavu sítě (linky) – na základě předpokládané doby odezvy (vypočítané z RTT), Karn-Jacobsonův algoritmus

Příjemce:

- má definovaný časový interval pro příjem následujícího segmentu s dalšími daty v toku dat (dle pořadových čísel)
- při neobdržení následujícího segmentu po vypršení intervalu nebo obdržení segmentu s dalšími daty mimo pořadí **opakuje potvrzení přijetí** předchozích dat
- ukládá si i data mimo pořadí do vstupního bufferu, po obdržení chybějícího segmentu **potvrdí příjem všech dat**

Ztráta segmentu (řízení toku dat)

CVIČENÍ: simulace ztráty segmentu (přerušením linky) a pozorování chování protokolu TCP při opakování odesílání a potvrzování dat

Zpoždění odpovědi

- výhodné u **interaktivních (konzolových) aplikací**, např. Telnet, FTP (příkazový kanál), SSH apod., vyměňujících **malé segmenty** (např. 1 B dat)

Obrázek: Obrázek průvodce 233→244(5)

- klasický průběh: uživatel stiskne klávesu, klient odešle znak serveru (v segmentu v IP paketu v linkovém rámci), server potvrdí příjem, zpracuje znak, odešle znak klientovi pro jeho zobrazení (interaktivita), klient potvrdí příjem a zobrazí, tj. min. 117 bytů (pro Ethernet) v každém směru – **velká režie**

→ snaha zmenšit objem přenášených dat a nebezpečí zahlcení sítě

CVIČENÍ: pozorování zpoždění odpovědi u aplikace Telnet (viz dále)

Zpoždění odpovědi

= **potvrzování příjmu dat** ne hned, ale **se zpožděním**, během kterého se mohou nahromadit data k odeslání:

Obrázek: Obrázky průvodce 234→244,245(5)

- **„delayed ACK“**: odesílání dat včetně potvrzení **v intervalech** např. 200 ms ($\leq$ 500 ms)
- **Nagleův algoritmus**: odesílání dat včetně potvrzení až **po obdržení dalších dat** od druhé strany nebo až je objem dat k odeslání $\geq$ MSS – vyrovnává dobu odezvy vůči kapacitě přenosové cesty v síti

– kombinace způsobuje konstantní zpoždění potvrzování (“ACK delay”) → zakázání Nagleova algoritmu pomocí volby **TCP_NODELAY** síťového API OS

Posuvné okno (sliding window)

Obrázek: Obrázek průvodce 235→246(5)

- využití při odesílání většího množství dat, **zamezení zahlcení příjemce**

= **segmenty se odesílají bez potvrzení každého zvlášť** až do počtu odeslaných bytů rovno délce **posuvného okna** (v položce délka okna v TCP segmentu, pak se ukládají do výstupního bufferu)

- délka okna vyjadřuje počet bytů, které je příjemce schopen přijmout (má plný vstupní buffer) či (v definovaném čase) zpracovat
- při navazování spojení **příjemce** navrhne počáteční délku (spolu s MSS, typicky 6–8 MSS) a pak ji může **v potvrzovacích segmentech měnit (inzerovat)** nebo i vynulovat (okno „uzavřít“), tj. zakázat odesílateli odesílat další data (když „nestíhá“)

Posuvné okno (sliding window)

- položka délka okna má 2 B, tzn. okno může být dlouhé max. 64 kB – malé u rychlých sítí $\rightarrow$ volitelná položka **zvětšení okna**, $n = 0$ až 14, délka okna je potom násobena 2^n (posun o n bitů doleva), tj. až téměř 1 GB, možno použít jen u segmentů s příznakem SYN při navazování spojení, nastavováno parametrem OS
- potvrzováním příjmu dat se okno po datech k odeslání „posouvá“ a mění velikost = řízení toku dat (**flow control**)

CVIČENÍ: identifikace a pozorování posuvného okna při přenosu dat

Zahlcení sítě (congestion control)

- posuvné okno udává množství dat akceptované příjemcem
- pokud je příliš velké a síť na straně příjemce plně využitá nebo pomalá, odesílatel může síť zahltit a ta (směrovače) začne data zahazovat

→ okno i na straně odesílatele: **okno zahlcení (congestion window)** = jaké množství **nepotvrzených dat je možné odeslat aniž by došlo k zahlcení sítě** – cíl: největší možné

→ odesílatel odesílá data do velikosti menšího z posuvného okna a okna zahlcení

- dvě fáze určování velikosti okna zahlcení: pomalý start a předcházení/vyhýbání se zahlcení

Zahlcení sítě (congestion control)

Pomalý start (slow start)

= od navázání spojení se **velikost okna zahlcení (CWND)** počínaje MSS s každým potvrzeným segmentem **zdvojnásobuje**, až do ztráty segmentu nebo pokud by se překročila velikost posuvného okna nebo parametru **SSTHRESH** – hranice pravděpodobnosti zahlcení, první hodnota je parametr OS, typicky 64 kB

– při ztrátě segmentu:

- po třech stejných potvrzeních předchozího segmentu se **CWND zmenší na polovinu** a na tuto hodnotu se také nastaví SSTHRESH (mininálně ale 2×MSS)
- po neobdržení potvrzení (v časovém intervalu) se **CWND nastaví na MSS** a SSTHRESH na 2×MSS a začne se **znovu**

Obrázek: Obrázek průvodce 238→248(5)

Zahlcení sítě (congestion control)

Předcházení/vyhýbání se zahlcení (congestion avoidance)

- následuje po pomalém startu, **pomalé zvětšování okna** s každým potvrzením, např. o MSS, $MSS^2/CWND + MSS/8$ apod.

→ algoritmy vyhýbání se zahlcení (**congestion avoidance algorithms**): Tahoe (první), **Reno**, **New Reno**, Hybla (pro rádiové spoje), **BIC** (rychlejší adaptace pro rozsáhlé rychlé sítě), **CUBIC** (CWND je kubická funkce času od posledního zahlcení) aj.

- **selektivní potvrzování (selective ACK, SACK)** = potvrzování i segmentů mimo pořadí, pomocí volitelných položek záhlaví (s dohodou při navazování spojení)

Obrázek: Obrázek průvodce 238→248(5)

Zahlcení sítě (congestion control)

- odesílatel udržuje pro každé spojení velikosti MSS, posuvného okna, okna zahlcení (CWND) a parametru SSTHRESH
- nalezená hodnota SSTHRESH pro daný směr se i po ukončení spojení použije jako výchozí u dalších spojení v tomto směru, uložená ve směrovací tabulce

Při ztrátě segmentu (během přenosu dat):

- po třetím stejném potvrzení se nastaví **SSTHRESH na polovinu aktuální CWND** (mininálně 2×MSS), zopakuje se segment, nastaví se **CWND na „o něco“ vyšší než SSTHRESH** a při opakovaných potvrzeních se zvyšuje o MSS
- po potvrzení ztraceného segmentu (celého okna zahlcení) se nastaví **CWND na původní SSTHRESH** (**rychlý start/zotavení**) a opět probíhá pomalé zvětšování okna (algoritmus vyhýbání se zahlcení)
- po neobdržení potvrzení (v časovém intervalu) **znovu pomalý start** (CWND = MSS, SSTHRESH = 2×MSS)

User Datagram Protocol (UDP)

- RFC 768
- poskytuje **nespojovanou (datagramovou) „nespolehlivou“ službu**: data odeslána, nezaručuje se doručení ani znovuzasílání ztracených nebo poškozených dat – ponecháno na vyšším (aplikačním) protokolu

→ **vyšší výkon** a rychlost přenosu dat než u TCP, za cenu „nespolehlivosti“ – využití u streamování multimediálního obsahu

- nezávislý rozsah portů pro TCP a UDP, UDP porty označeny **číslo/udp**
- snaha vyhnout se IP fragmentaci datagramů – **velikost datagramu** $\leq$ **MTU** linky (např. u DNS delší odpověď zkrácena na 512 B a na vyžádání poslána celá pomocí TCP)
- oproti TCP může být příjemcem skupina uzlů, tj. **IP adresa příjemce** může být **všesměrová** (např. u DHCP) nebo **skupinová** (multicast, typicky u streamování multimediálního obsahu) – jak dožádat nedoručená data (např. u přenosu souborů pomocí Multicast FTP)? → od nejbližšího směrovače (protokolem pro multicast)

UDP datagram

Obrázek: Obrázek průvodce 241→251(5)

- záhlaví 8 B, data
- **délka dat**: délka datagramu, tj. záhlaví a dat
- **kontrolní součet**: stejně jako u TCP počítán z tzv. pseudozáhlaví (některé položky IP záhlaví, UDP záhlaví a data), nemusí být povinně vyplněný (nulový), pro zrychlení (např. u NFS), ale může být nebezpečné (např. u DNS, pak počítán jen z linkového rámce, ale např. SLIP nepočítá)

CVIČENÍ: zachytávání a inspekce UDP datagramů

Bezpečnost protokolů TCP a UDP

TCP

- „spolehlivá služba“ – potvrzování příjmu dat a znovuzaslání ztracených a poškozených
- pouze kontrolní součet (i když i z části IP záhlaví a dat) – lze přepočítat
- náhodné pořadové číslo 1. bytu odesílaného toku dat (ISN) – pouze pro zaručení správného pořadí dat (a také zahození zatoulaných segmentů z předchozího přerušeného spojení ze stejného portu)
- **útoky**: **převzetí spojení** (**connection hijaking**, autentizovaného a dále nezabezpečeného!), **odepření služby** (**Denial of Service**, vyčerpání zdrojů systému pro spojení, maximum příznaků v záhlaví), zjišťování otevřených portů serveru (**port scanning**) a útok na aplikaci, aj.

→ **řešení**: **šifrování spojení** pomocí SSL, S/MIME apod. nebo vytvořením (šifrovaných) **tunelů** na jiných portech, omezování počtu spojení za daný čas, sledování (sekvenčního) skenování portů aj.

Bezpečnost protokolů TCP a UDP

UDP

- vyplnění kontrolního součtu je nepovinné, jinak lze přepočítat
- **musí** jej používat aplikace přenášející data na **skupinové nebo všesměrové adresy**, např. streamovaná multimedia nebo DHCP
- na směrovačích bývají povoleny porty pro **DNS** (53/udp, 53/tcp), dále např. UDP používá program traceroute na unixových systémech

Firewall

- **filtrace** paketů a segmentů/datagramů na základě TCP/UDP záhlaví

→ zejména **„bránění“ navázání TCP spojení nebo přenosu dat pomocí UDP na vybraných portech** ($\sim$ „blokování služeb“) = filtrování TCP segmentů s příznakem SYN (prvního při navazování spojení) a UDP datagramů na cílový port

- TCP záhlaví jen v prvním IP fragmentu – doporučené sledovat fragmenty a filtrovat i další

Bezpečnost protokolů TCP a UDP

Překlad adres (NAT)

- překlad IP adresy odesílatele paketů z vnitřní sítě na IP adresu hraničního směrovače ve vnější síti, tzv. **maškaráda** = **překlad** adresy a **zdrojového portu** spojení/přenosu (adresy socketu) na zdrojový port nového spojení/přenosu ze směrovače (**NAPT (Network Address and Port Translation)**)
- překlad portů u transparentních proxy (typicky v DMZ nebo přímo hraniční směrovač)
  - zasahuje i do aplikační vrstvy, v případě nutnosti porozumět aplikačnímu protokolu pro překlad IP adres/portů v datech, např. FTP

Aplikační vrstva

CVIČENÍ: aplikační programové rozhraní BSD Socket/Winsock

Jmenné služby

- aplikace používají pro identifikaci uzlů (síťových rozhraní) v síti **číselné** síťové (IP) adresy, např. 158.194.80.13
- pro člověka jsou číselné adresy těžko zapamatovatelné a sledují **fyzickou strukturu sítě** (na síťové vrstvě) – jedna organizace může mít podsítě po celém Internetu

→ **textové** označení uzlu přiřazené k adrese, **strukturované jméno** uzlu sledující **logickou strukturu sítě**

→ aplikace (používané člověkem) používají jména – **jméno se** nejdříve **přeloží na IP adresu** a ta se použije

- použití IP adres pouze nouzově při problémech s překladem
- historický vývoj:
  1. každý uzel udržuje vlastní databázi jmen – s počtem roste náročnost
  2. centrální databáze ve středisku InterNIC – uzké místo, proti duchu Internetu
  3. **decentralizovaná distribuovaná databáze** – systém DNS, 1985

Domain Name System (DNS)

- RFC 1035 a další
- strukturované jméno uzlu = symbolické, **doménové jméno**, např. **phoenix.inf.upol.cz**

= **decentralizovaná distribuovaná databáze** záznamů **doménových jmen vs. IP adres** (k jedné IP adrese může být přiřazeno více doménových jmen a obráceně)

= **systém překladu doménových jmen** na IP adresy a naopak

= **decentralizovaná distribuovaná (aplikační) služba** modelu klient/server

- záznamy rozmístěny na tzv. **jmenných (DNS) serverech**
- klient, tzv. **řešitel (resolver)**, žádá jmenný server o překlad doménového jména na IP adresu, popř. naopak

Domény

= pojmenované, **stromově hierarchické skupiny** logicky sdružených **uzlů** v síti (např. organizace, země, Internetu), podskupiny = **subdomény** (např. oddělení organizace), strukturní jednotky DNS

- **kořenová (root) doména** – nejvyšší doména stromu obsahující top-level domény, běžně se neuvažuje, existují i alternativní (OpenNIC, New.Net aj.)
- **top-level domény (TLD)**:
  - spravované IANA (ICANN), www.iana.org/domains/root/db/
  - infrastrukturní (historicky generické): jméno **arpa** (1985, Address and Routing Parameter Area), např. pro reverzní domény
  - **generické (gTLD)**: otevřené, jména com (1984, RFC 920), info (2000), net (1984), org (1984), i s omezeními na registraci, biz (2000), name (2000), pro (2000)
  - **sponzorované** (**sTLD**, uvažované jako generické): s omezeními na registraci, jména aero (2000), asia (2006), cat (2005), coop (2000), edu (1984), gov (1984), int (1988), jobs (2005), mil (1984), mobi (2005), museum (2000), post (2005), tel (2005), travel (2005), xxx (původně zamítnutá), od června 2008 jakákoliv (např. msn, google)

Domény

- **top-level domény (TLD)**:
  - **národní (country-code, cTLD)**: dvojznaková jména domén států a unií (ISO 3166), např. cz, sk, eu
  - **internacionalizované (IDN)**: pro testování národních abeced (arabské, cyrilice, čínské, řecké apod.)
  - rezervované: pro speciální účely v neprodukčních sítích

– top-level domény (domény 1. řádu) obsahují domény 2. **řádu** pro organizace (např. jména upol, google), ty zase domény 3. řádu (např. inf) atd. až po jména uzlů (např. phoenix, mail)

– domény spravovány jmennými servery – uloženy a poskytovány záznamy doménových jmen vs. IP adres

Obrázek: Obrázek průvodce 246→257(5)

Domény

Doménové jméno

= odráží příslušnost uzlu či subdomény k (sub)doméně, složeno ze jména uzlu v (sub)doméně a jmen nadřazených (sub)domén, např. uzel phoenix v subdoméně inf v subdoméně upol v doméně cz (v kořenové doméně)

– **tečková notace**: (zleva) jména uzlu a postupně nadřazených domén oddělená tečkou, max. 255 B

– jméno uzlu/domény: case-insensitive řetězec znaků, původně pouze **ASCII** znaky (a–z, 0–9, –, RFC 1034), od 1998 **IDN** (v některých TLD, v testovacím režimu, 2003 IDNA převod na ASCII, algoritmy ToASCII a ToUnicode), max. 63 B

– **kořenová doména** má **prázdné jméno**, poslední oddělující tečka se běžně nepíše (relativní jméno), i s tečnou (absolutní jméno) je tzv. **plně kvalifikované doménové jméno (FQDN)**

– např. **phoenix.inf.upol.cz.**

– uvnitř domény se obvykle vynechává část jména pro doménu, např. uvnitř inf.upol.cz jen phoenix

Domény

Reverzní domény

- pro **reverzní překlad IP adresy na doménové jméno**, např. z bezpečnostních důvodů (ověření IP adresy ke jménu)

= k IP adrese přiřazené doménové jméno v doméně **in-addr.arpa**: (standardně) zleva jména uzlu a reverzních subdomén jako čísla v IP adrese zprava

- např. pro IP adresu 158.194.80.13 jméno **13.80.194.158.in-addr.arpa**
- jména reverzních domén překládaná na doménová jména (stejným způsobem jako překlad na IP adresy)

Obrázek: Obrázek průvodce 248→258(5)

- reverzní doména **0.0.127.in-addr.arpa**: pro reverzní překlad zpětné smyčky uzlu (127.0.0.1) na jméno localhost, měla by být spravována každým jmenným serverem

Domény

Rezervované domény (RFC 2606)

- example (příklady do dokumentací, 192.0.2.0/24), invalid, localhost, test (také pro testování IDN), ...

Pseudodomény

- **local** – pro lokální sítě (intranety, 10.0.0.0/8), autokonfigurační protokol Zeroconf (multicast DNS), uzly bez přiděleného doménového jména (169.254.0.0/16, link-local) apod., záznamy pro překlad přímo na uzlu, ne na jmenném serveru
- pro jiné sítě: **onion** (pro anonymizační síť Tor), uucp (stará síť UUCP, bang notace jména), bitnet (síť BITNET) aj.

Zóna

Obrázek: Obrázek průvodce 249→259(5)

= **část** (prostoru jmen) **domény** spravovaná jedním jmenným serverem, kromě subdomén (podřízených zón) delegovaných jiným serverům

- **kořenové zóny** (= části kořenové domény), speciální zóny – pro implementaci jmenného serveru, např. stub (seznam jmenných serverů pro subdomény), cache/hint (seznam IP adres jmenných serverů pro kořenovou doménu/zóny)

Řešitel (resolver)

= **klient služby DNS** dotazující se jmenného serveru na překlad jména
- vyžaduje od serveru **konečnou odpověď**, kladnou (výsledek překladu) nebo zápornou (neexistující záznam)

= **komponenta OS**, knihovna nebo knihovní funkce standardní systémové knihovny používané aplikacemi pro jmennou službu
- má v konfiguraci **IP adresy** (!) **jmenných serverů místní domény**, kterých se dotazuje: v unixových OS soubor `/etc/resolv.conf`, v MS Windows záložka DNS v dialogu nastavení protokolu TCP/IP (plus záložka WINS pro systém LAN Manager, protokol NetBIOS a službu WINS poskytující jiný překlad jmen na IP adresy)
- může (dle konfigurace) k zadanému jménu bez koncové tečky (relativnímu jménu) přidávat **přednastavené domény** (v MS Windows i domény Windows), při negativních odpovědích dotazy postupně znovu bez nich
- konfigurace je možná ručně (staticky) nebo dynamicky pomocí protokolů DHCP nebo PPP

Řešitel (resolver)

- obsahuje cache se záznamy z výsledků předchozích dotazů (pozitivní i negativní), bez cache tzv. **pahýlový resolver**, např. v unixových OS (GNU/Linux), pro cache je pak caching-only jmenný server (viz dále, např. pdsnd, dnsmasq) nebo speciální daemon (např. nscd), v MS Windows 2000 a víš resolver s cache při volbě "Klient DNS" (výchozí)
- kromě DNS překladu (před ním) lze využít **lokální soubor** s (ručně zadanými) asociacemi jmen a IP adres

CVIČENÍ: konfigurace resolveru, IP adres jmenných serverů, přednastavené domény, lokální soubor

Jmenný server

- spravuje **záznamy pro svou zónu**, včetně seznamu jmenných serverů pro subdomény/podřízené zóny (stub) – tzv. **autoritativní záznamy**
- obsahuje **seznam IP adres serverů spravujících kořenovou zónu** (z konfigurace, cache/hint) a cache se záznamy z výsledků předchozích dotazů na jiné servery (pozitivní i negativní) – neautoritativní záznamy

= program poskytující klientům (resolver nebo jiný server v roli klienta) **odpověď na dotaz** = výsledek překladu jména, např. v unixových OS program BIND

- typy:
  - **primární** – jediný „hlavní“, autoritativní, server pro doménu/zónu (záznamy zóny v konfiguraci), poskytuje tzv. **autoritativní odpověď** pro autoritativní záznamy ze své zóny a neautoritativní odpověď pro záznamy z cache
  - **sekundární** – „vedlejší“, autoritativní, server pro doménu/zónu, pravidelně kopíruje záznamy zóny dotazem (**zone transfer**) z primárního serveru, poskytuje stejné odpovědi jako primární

Jmenný server

- typy:
  - **caching only** – neautoritativní server pro (žádnou) doménu nebo zónu, poskytuje pouze **neautoritativní odpovědi**
  - **kořenový** – primární server pro kořenovou doménu/zónu, je jich víc
  - **forwarder** – server provádějící překlad pro jiný server (v roli klienta)
- **pro každou doménu vždy minimálně dva** (nezávislé) jmenné servery, **primární a sekundární, v konfiguraci jmenného serveru nadřízené domény** – pravidlo Internetu
- jeden jmenný server může být primárním pro jednu doménu/zónu a zároveň sekundárním pro jiné domény/zóny
- **round robin**: při více IP adresách (různých strojů) k jednomu jménu cyklické vracení různých adres na dotazy na jméno, použití pro rovnoměrné vyrovnávání zátěže strojů (load balancing)

Překlad (vyřešení dotazu)

= překlad doménového jména z dotazu na IP adresu nebo IP adresy (reverzního doménového jména z dotazu) na doménové jméno

– požaduje resolver nebo jmenný server (v roli klienta), poskytuje jmenný server

– dotaz:

- **rekurzivní** – klient vyžaduje a server vrací **konečnou odpověď** (autoritativní nebo neautoritativní), typicky požaduje resolver
- **nerekurzivní** – server vrací **seznam IP adres jiných jmenných serverů**, typicky požaduje jmenný server v roli klienta – běžně označovaný jako **resolvující jmenný server**, **„resolver“**

Obrázek: Obrázek kombinace průvodce 253, 260 a 262→263 a 269(5)

Překlad (vyřešení dotazu)

1. aplikace žádá resolver o překlad
2. resolver prohledá cache (pokud ji má), v případě cache miss
3. resolver vznese (rekurzivní) **dotaz na jmenný server** (pro **místní doménu**, první z konfigurace) – pokud nedojde v časovém intervalu odpověď, opakuje dotaz na cyklicky další nebo stejný (pokud je v konfiguraci jen jeden) do vypršení celkového časového intervalu na překlad
4. server prohledá cache, v případě cache miss
5. server vznese **dotaz na jiný jmenný server** (DNS databáze je distribuovaná) – opakovaně v časových intervalech do vypršení celkového
   - defaultně **kořenový** (ze seznamu) – na nerekurzivní dotaz vrací **seznam IP adres jmenných serverů pro top-level doménu**, náš server vznese **dotaz na některý z nich**, ten v případě nerekurzivního dotazu vrátí seznam IP adres serverů pro subdoménu vyššího řádu atd. až do konečné odpovědi = **proces iterace**, **rekurzivní překlad**

Překlad (vyřešení dotazu)

5. • **výjimečně nadřazený** pro nadřazenou domény (zónu) nebo **forwarder** – vrací konečnou odpověď, náš server se chová jako resolver a vznáší rekurzivní dotaz, ale po vypršení časového intervalu provede překlad sám (pokud není tzv. **forwarder only**, např. v uzavřených sítích)

CVIČENÍ: vysvětlení úplného postupu rekurzivního překladu konkrétního jména (např. www.seznam.cz) z uzlu v konkrétní doméně (např. inf.upol.cz)

- **kořenové servery a servery pro TLD obsluhují pouze nerekurzivní dotazy** (kvůli zátěži, kritické místo systému DNS!), caching only server předává dotaz autoritativnímu serveru domény/zóny
- manuální překlad/diagnostika DNS: nástroje **nslookup, dig**

  CVIČENÍ: manuální překlad jména a IP adresy (reverzní), rekurzivní i nerekurzivní, programem nslookup (dig nebo host)
- veškerá komunikace (dotazy a odpovědi) pomocí **protokolu DNS**

Protokol DNS

= **aplikační protokol** pracující způsobem **dotaz-odpověď** poskytující službu typu **klient/server**: klient pošle dotaz, server odpověď

– základní **operace DNS Query** pro získání informací z DNS databáze na serveru, typicky překlad doménového jména na IP adresu

– další operace DNS, např. Update, Notify, Zone transfer aj.

– používá pro přenos dat transportní protokoly **UDP i TCP**, pro oba **port 53** (tj. 53/udp i 53/tcp)

  – stejný protokol pro dotaz i odpověď

  → pro běžné dotazy, např. překlad jména, **nejprve UDP** (kvůli režii TCP, časovým intervalům při nedostupnosti serveru), **odpověď** případně zkrácena na **512 B** (velikost UDP datagramu, kvůli IP fragmentaci)

  → pro kompletní odpověď nebo zone transfer dotaz přes TCP

  – protokol DNS (jmenná služba) **není zcela spolehlivý** – časový interval pro odpověď, datagramový protokol UDP

– pro různé operace různé **DNS pakety** – **neobsahují kontrolní součet!** → měl by obsahovat UDP datagram

DNS Query

= základní operace protokolu DNS: **dotaz** (klienta) a **odpověď** (serveru) s informacemi (**záznamy**) podle požadavků v dotazu (**pro doménové jméno, typ záznamu**) nebo negativní (záznam podle požadavků neexistuje)

– stejný formát DNS paketu pro dotaz i odpověď

Obrázek: Obrázek průvodce 266→294(5)

– 5 sekcí paketu: záhlaví (povinné), dotazy, odpovědi, autoritativní jmenné servery a doplňující informace (nepovinné)

- sekce **záhlaví (HEADER)**: v dotazu i odpovědi
  - **ID**: identifikátor, stejný v dotazu i odpovědi, pro spárování
  - **QR**: 0 pro dotaz, 1 pro odpověď
  - **Opcode**: typ dotazu (stejné v odpovědi), 0 pro standardní, 1 pro inverzní, 2 pro status, 4 pro operaci notify, 5 pro operaci update, aj.

DNS Query

- sekce **záhlaví (HEADER)**:
  - **AA**: 1 pro autoritativní odpověď
  - **TC**: 1 pro odpověď zkrácenou na 512 B
  - **RD, RA**: 1 pro požadavek (u dotazu) a možnosti (u odpovědi) rekurzivního překladu
  - **AD**: 1 pro požadavek (u dotazu) a vyznačení (u odpovědi) zabezpečené (podepsané) odpovědi (DNSSec)
  - **Rcode**: kód odpovědi, 0 (NoError) pro bez chyby, 1 (FormErr) pro chybu formátu dotazu, 2 (ServFail) pro neschopnost odpovědi, 3 (NXDomain) pro negativní odpověď (záznam pro jméno z dotazu neexistuje), 5 (Refused) pro odmítnutí odpovědi atd.
  - další: **počet záznamů** v dalších sekcích, při 1 formát odpovědi „one-answer“, při více „many-answer“, záleží na implementaci serveru
- sekce **dotazů (QUESTION)**: většinou jeden záznam (**doménové jméno a typ**), v dotazu i odpovědi (zopakovaný)
- ostatní sekce (**ANSWER, AUTHORITY, ADDITIONAL**): odpověď s požadovanými záznamy (ANSWER), jména autoritativních jmenných serverů pro doménu (příp. subdomény, AUTHORITY) a jejich IP adresy (příp. poštovní servery, ADDITIONAL)

DNS Query

- **komprese DNS paketu**: další výskyty (části) doménového jména v paketu jsou nahrazeny odkazem na první výskyt – oddělovací byte ve jméně (viz dále) je ≥ 192, tj. první dva bity 1, ostatní bity a další byte = pořadové číslo bytu prvního výskytu od začátku paketu (od 0)
- **inverzní dotaz** (Opcode = 1): jako reverzní, ale pro odpověď se místo záznamů typu PTR použijí záznamy typu A (viz DNS záznamy/věty RR dále), nemusí být servery podporován

CVIČENÍ: zachytávání a inspekce (záhlaví) DNS query paketů

DNS záznamy/RR věty

= **zdrojové věty (resource records, RR)** – forma dat **záznamů v** DNS **paketech** operací, např. u Query v dotazu a odpovědi

– forma **uložení záznamů** o doménových jménech vs. IP adresách a všech ostatních informací DNS v databázi na jmenném serveru, v textové podobě

Obrázek: Obrázek průvodce 264→272(5)

- **NAME**: **doménové jméno** uzlu nebo subdomény, řetězec proměnné délky – před řetězci mezi tečkami v tečkové notaci jmen je oddělovací byte s délkou řetězce (místo tečky) a nulový byte na konci, např. 7phoenix3inf4upol2cz0
- **TYPE**: **typ věty**, určuje význam pole RDATA (v odpovědi serveru):

DNS záznamy/RR věty

- **A** (1): **IPv4 adresa** (4B, v poli RDATA) ke jménu NAME
- **NS** (2): **jméno autoritativního jmenného serveru pro subdoménu** NAME (**na serveru nadřazené domény**) nebo pro doménu z věty SOA (na serveru domény), typicky primárního a sekundárního, pro jmenný server by měla být i věta A (tzv. **glue záznam**), domény z NAME postupně delegovány na servery od kořenových serverů stromem domén dolů
- CNAME (5): jméno jako alias k NAME
- SOA (6): informace o autoritativním (primárním) jmenném serveru pro doménu NAME (jeho jméno, email správce, časový interval pro zone transfer, výchozí hodnota TTL aj.)
- **PTR** (12): (FQDN) **jméno** k NAME **pro reverzní překlad**, reverzní domény z NAME postupně delegovány na servery od kořenových serverů stromem domén dolů
- **MX** (15): preference (2B číslo) a **jméno e-mailového serveru pro doménu** NAME

DNS záznamy/RR věty

- WKS (11), SRV (33): informace o počítači (jméno/IP adresa, port, priorita, váha) s aplikační službou (aplikační a transportní protokol, např. _http._tcp) pro doménu NAME
- HINFO (13), TXT (16): informativní, info o HW a SW uzlu NAME, lib. text
- AXFR (252), IXFR: požadavek transferu zóny (celé zóny nebo inkrementálního), v DNS paketech operace Zone transfer
- * (255): požadavek na všechny věty, v DNS paketech
- další: pro IPv6, DNSSec (zabezpečení DNS) aj., viz dále

DNS záznamy/RR věty

- CLASS: třída věty, IN (1) pro Internet, * (255) pro všechny
- **TTL**: time to live, doba **platnosti záznamu v cache** jiných serverů a resolveru (0 zabraňuje uchovávání v cache), v sekundách
- RDLENGTH: délka pole RDATA
- **RDATA**: **data (určená typem věty)**, jména jako řetězce proměnné délky
  - v dotazu operace Query jen položky NAME, TYPE a CLASS
  - **v konfiguraci serveru** (tzv. **zónových souborech**) zadané **v textové podobě**, jména v syntaxi doménových jmen, položky zadaná svým textovým označením, oddělené bílými znaky

CVIČENÍ: inspekce záznamů (RR vět) z jednotlivých sekcí DNS paketů z následujícího cvičení, rozpoznání komprese jména v paketu

CVIČENÍ: překlady programem nslookup (nebo dig): získání DNS záznamů (RR vět) pro dané jméno neexistujících, daných (A, NS, SOA, PTR, MX) a všech typů, ze serveru mimo místní doménu, inspekce TCP segmentů u delší odpovědi, s ladícím výstupem (úroveň debug)

DNS Update

- RFC 3007
- = operace DNS protokolu pro **dynamickou aktualizaci DNS záznamů** (RR vět) v konfiguraci primárního jmenného serveru (jiné na něj přepošlou)
- dotaz + odpověď, formát paketu podobný operaci Query: sekce zóny, předpokladů (na ne/existující věty), update (přidávané nebo rušené věty) a doplňkových informací
- změny jsou na serveru ukládány do **zónových žurnálových souborů** pravidelně ukládaných do zónových souborů konfigurace
- zabezpečení: Secure DNS Update, update dotazy povolené pouze z dané IP adresy aj.
- klient **nsupdate**

DNS Notify a Zone transfer

DNS Notify (RFC 1996)

= operace DNS protokolu pro **informování** sekundárních a podřízených jmenných serverů (tzv. notify set) **o změně záznamů** na primárním serveru (dříve než vyprší interval aktualizace)

– zprávu periodicky (různým serverům s různým zpožděním) zasílá primární server (formát paketu podobný operaci Query), sekundární nebo podřízený server potvrdí a požádá o transfer zóny

Zone transfer

= operace DNS protokolu pro **přenos záznamů zóny** z (typicky) primárního serveru

- **AXFR** = přenos všech záznamů
- **IXFR** = **inkrementální** – přenos pouze změněných záznamů (ručně v konfiguraci nebo operací Update), udržuje se historie stavů databáze, při příliš starém stavu nebo rozsáhlém IXFR se provede AXFR

Rozšíření DNS pro IPv6

- RFC 1886, 2874 aj.
- pro překlad doménového jména na IPv4 adresu se používá záznam (RR věta) typu A
- pro IPv6 adresu nejdříve a nyní záznam typu **AAAA** s 16B **IPv6 adresou**
- dříve dočasně záznam typu A6: počet bin. 1 v síťové masce, část IPv6 adresy pro uzel a doménové jméno domény uzlu, jedna IPv6 adresa volitelně uložena pomocí několika A6 záznamů, po doménách, na různých serverech – resolver musel sestavit tzv. A6 record chain
- **jméno pro reverzní překlad**: nejdříve a nyní jména uzlu a reverzních subdomén jako jednotlivé šestnáctkové cifry v IPv6 adrese (tzv. nibble formát), nejdříve v doméně ip6.int, nyní v doméně **ip6.arpa**, např. pro IPv6 adresu ??? jméno ???, dříve dočasně jména tvaru \[xcifry/bitů] (tzv. bitstring formát)
- záznam typu DNAME: analogie CNAME, jméno jako alias části doménového jména, např. pro postupnou delegaci reverzních subdomén místo záznamů typu NS

Zabezpečení DNS

DNSSec

- dřívější RFC 2535, 2538, dnes novější RFC 4033–5
- = **zabezpečení záznamů na jmenných serverech a v DNS paketech**, dříve od vybraných (top-level) domén/zón ve stromu domén níže, dnes od kořenové domény
- použití **el. podpisu**: soukromým klíčem subdomény/zóny podepsány všechny její záznamy (kromě RRSIG), podpisy v záznamech typu **RRSIG** (dříve SIG), pro ověření **integrity DNS paketu** (např. odpovědi DNS Query) záznamy pospojované do posloupnosti pomocí (podepsaných) záznamů typu **NSEC** (dříve NXT), plus poslední speciální záznam RRSIG podepisující celý paket
- ověření podpisů: veřejný klíč subdomény/zóny v (podepsaném) záznamu typu **DNSKEY** (dříve KEY), podepsaný (certifikovaný) soukromým klíčem **nadřízené domény**, veřejný klíč kořenové domény (self-signed, popř. vyšších zabezpečených domén) v konfiguraci resolveru

Zabezpečení DNS

DNSSec

- možnost uložení certifikátů (X.509 aj.) pro aplikace pomocí záznamů typu CERT
- nevýhody: soukromý klíč je potřeba pro podpis každého DNS paketu se záznamy (podpisy spojujících NSEC záznamů + celého paketu, podpisy jednotlivých záznamů již v podepsané konfiguraci zóny nebo cache)

TSIG (Transaction Signatures)

= **autorizace komunikace DNS serverů**, RFC 2845

- MD5 hash přenášených záznamů a sdíleného tajemství v záznamu typu **TSIG**
- sdílené tajemství vyměňováno Diffie-Hellmanovým algoritmem pomocí záznamů typu **TKEY**, nebo asymetrickou šifrou (tajemství zašifrováno zaslaným veřejným klíčem)
- použití u DNS Update – může jen autorizovaný server

Implementace jmenného serveru

Systém BIND (verze 4)

- DNS záznamy v textovém tvaru (**formát BIND**) udržovány v zónových souborech na primárním serveru
- udržovaná data: autoritativní záznamy zóny vč. záznamů delegujících správu části domény na jiné (podřízené) jmenné servery, záznamy zóny cache/hint (seznam IP adres kořenových jmenných serverů)

= program **named** na unixových systémech, služba **Server DNS** na MS Windows 2000 (může být součástí Active Directory)

BIND nové generace (verze 8 a 9)

- podpora dynamické aktualizace (**DNS Update ve spolupráci s DHCP serverem**), DNS Notify, IXFR, negativní caching, DNSSec, virtuální jmenné servery, propojení s MS Windows 2000, IPv6, ...
- oproti BIND 4: protokolování zpráv, ACL, master/slave místo primární/sekundární/atd., vícevláknový, implementace i pro MS Windows
- **lightweight resolver** = knihovna + (lokální) server jako caching-only jmenný server

Testování a ladění DNS

- chybně nastavené DNS $\Rightarrow$ **prodlevy** v aplikacích a OS kvůli časovému intervalu na překlad $\rightarrow$
  1. ověřit **fungování sítě**, např. pomocí ping
  2. ověřit **konfiguraci resolveru** – místní DNS servery, doména
  3. **testování** (místních) **jmenných serverů – dotazy** jako resolver i jako server v roli klienta
  4. kontrola a ladění konfigurace serveru – podle pravidel DNS (nástroje implementace serveru, např. rndc u BIND 9)
- nástroje (RFC 1713):
  - **nslookup** – rekurzivní i nerekurzivní dotazy, volba typů záznamů a jmenného serveru aj., interaktivní, ladicí výstup (úrovně debug a d2)
  - **dig** – rekurzivní i nerekurzivní dotazy, volba typu záznamů a jmenného serveru aj., formát BIND odpovědi
  - **dnswalk** – kontrola záznamů pro doménu (i reverzních) podle pravidel DNS, z transferu zóny

CVIČENÍ: testování DNS (dotazy) programy nslookup a dig (viz minulé cvičení), kontrola záznamů pro doménu programem dnswalk

Delegace a registrace domén

Delegace domény na vlastní jmenné servery

1. vytvoření **primárního jmenného serveru** pro doménu – připojení k Internetu by mělo být pevnou linkou (pravidlo Internetu)
2. vytvoření **sekundárního jmenného serveru** pro doménu – případně u poskytovatele Internetu
3. **delegace domény v nadřazené doméně** = záznamy typu NS v nadřazené doméně a typu PTR v nadřazené reverzní doméně pro jmenné servery delegované domény (plus glue záznamy typu A)

Delegace a registrace domén

Registrace domény 2. úrovně

1. **registrace domény** – v databázi (lokálního) **Internet Registry (IR)** pro TLD (= nadřazená doména, např. pro cTLD cz národní sdružení CZ.NIC), prostřednictvím **registrátora** (často poskytovatel připojení k Internetu), doména musí být **volná**
2. **registrace reverzní domény** – pro rozsah IP adres z bloku adres (= nadřazená reverzní doména), v databázi poskytovatele připojení k Internetu nebo regionálního IR (např. RIPE NCC), prostřednictvím registrátora

Příklad průvodce 370–372

Internet Registry (IR)

= organizace přidělující v Internetu IP adresy (RFC 1466), čísla autonomních systémů, jména domén (TLD a 2. řádu) aj.

- **IANA** (The Internet Assigned Numbers Authority) – nejvyšší, rozděluje mezi regionální IR
- **regionální** – spravují větší geografické oblasti Internetu rozdělené mezi lokální IR
  - **RIPE NCC** – Evropa, Blízký východ a Rusko (a bývalé sovětské republiky)
  - **ARIN** – Severní Amerika
  - **APNIC** – asijsko-pacifická oblast
  - **LACNIC** – Latinská Ameriku
  - **AfriNIC** – Afriku
- **lokální** – národní IR a poskytovatelé připojení k Internetu, sponzorují regionální IR, např. **CZ.NIC**, DE.NIC, **ICANN** (USA, gTLD, sTLD) atd.

Internet Registry (IR)

RIPE (www.ripe.net)

- **objekty** databáze = přidělená čísla a jména (inetnum, domain, aut-num), informace o zodpovědných osobách (správcích sítí = person, role, autorizovaných ke změnám = mntner), směrování = route aj.
- databáze **veřejně přístupná**, čtení pomocí programu **whois** nebo služby WWW, editace e-mailem

Protokol DHCP

model klient/server

Přidělení adresy

zprávy

Směrovací protokoly

RIP

OSPF

BGP

Elektronická pošta (e-mail)

Architektura

model klient/server, různé protokoly, záznam typu MX v DNS

Poštovní zpráva (e-mail), MIME

hlavičky

Protokoly SMTP a ESMTP

příkazy, rozšíření (např. 8BITMIME, potvrzení o doručení)

Protokoly POP3 a IMAP4

příkazy, stavy

Konference a diskuzní skupiny

Protokol NNTP

Informační služby – HTTP

Architektura

model klient/server, HTTP proxy a brána

URI

Dotaz a odpověď

metody dotazu GET a POST

Relace (session) a cookies

Přenos dat – FTP

Architektura

model klient/server, příkazový a datový kanál, módy přenosu dat

příkazy

FTP proxy a anonymní FTP

Aktivní a pasivní režim komunikace

Vzdálené přihlášení – Telnet, SSH

Virtuální terminál

Telnet

příkazy

SSH

port forwarding

Další aplikační protokoly

NTP

SMB

LDAP

Bezpečnost na aplikační vrstvě

Filtrace aplikačních protokolů

Aplikační proxy a brány, SOCKS

Autentizace uživatele a autorizace dat

Protokoly RADIUS a Kerberos

Prezentační protokol SSL/TLS a S/MIME

zabezpečení aplikačních protokolů (HTTP, FTP, IMAP aj.)