

IPv6

Vladimír Kotal
vlada@openbsd.cz

March 15, 2003

Problémy IPv4

- malý adresní prostor
- škálovatelnost
- bezpečnost (máme IPsec, ale...)
- Quality of Service

Nedostatek IPv4 adres

- mobilní komunikace typu 3G
- státy, které zmeškaly Internetovou horečku
- domácí sítě
- kolem roku 2008 nebudou internet registries (RIRs, např. RIPE) schopny přidělit adresy
- nárůst velikosti globálních routovacích tabulek (cca 60k záznamů)
- problém: stávající algoritmy nejsou lineárně škálovatelné

Nedostatek IPv4 adres (pokrač.)

krátkodobá řešení: :

- CIDR (agregace routovacích informací)
- NAT (Network Address Translation)

Proč je NAT špatný ?

- globální dostupnost : znemožňuje používat jednoznačnou adresu k identifikaci (rlogin/rsh, Kerberos,IPsec,NFS/RPC) a komunikaci (clustery), problémy s laděním problémů (DNS out of sync)
- globální jednoznačnost : aplikace nemohou asociovat stavy per-host s jeho adresou (jednoznačné identifikátory)
- persistence vazby host-adresa : NAT heuristiky mohou mít jiné pojetí session času než aplikace

Proč je NAT špatný ? (pokrač.)

- struktura adres : aplikace předpokládající, že dva různé porty jsou na tomtéž stroji (fail-over pro aplikace)
- zavádění nových aplikací (proxy moduly na NAT routeru, jejich upgrade, multicast aplikace)
- spolehlivost (single point of failure)
- škálovatelnost (M:N problém, stavové tabulky NAT routeru)
- privátní adr. prostor a VPN : adresní prostory se mohou překrývat

Three bears problem

2002 - cca 350 mil. zařízení na Internetu
IPv4 - teoreticky 4 bil. IP adres

V čem je problém ?

- minimum organizací potřebuje adr. prostor třídy A
- B je tak akorát pro většinu

Three bears problem (pokrač.)

- C je příliš malý
 - použijeme několik tříd C pro jednu organizaci
 - * vyčerpání adresního prostoru
 - * exploze velikosti routovacích tabulek

CIDR

- řešení: CIDR (Classless Inter-Domain Routing)
 - standardní IP adresa
 - informace o délce (počtu bitů) síťového prefixu

Důsledky CIDR

- vznikají supernety: sítě, jejichž prefix je kratší než přirozený prefix např. supernet 192.32.0.0/16 má přirozený prefix /24
- je možné provádět route aggregation např. ISP1 má přidělen prefix /16 a z něj přiděluje zákazníkům např. /24, /20 atd. výměna routovacích informací s ostatními : oznamuje pouze /16 prefix
 - hierarchická agregace routovacích informací
 - minimalizace velikosti routovacích tabulek

IPv6/IPng - základní vlastnosti

- 128-bitový adresní prostor (adresy pro všechna zařízení)
- bezstavová autokonfigurace (jednodušší správa IP adres)
- hierarchická struktura (agregovaná alokace adres už od začátku)
- "lepší" QoS - 20bit Flow label a 8bit Traffic Class indikátor přímo v IP hlavičce
- povinný IPsec (IPv4 má IPsec taky)

Hlavičky IPv4, IPv6

- IPv4
 - 12 polí, celková délka 160 bitů
- IPv6
 - IP hlavička má pevnou délku (8 polí, 40B)
 - checksumy nejsou v síťové vrstvě (jsou až v transportní)
 - routery neprovádějí zpracování fragmentů (PMTU detekce) to je starost zdroje
 - MTU $>$ 1280 povinně (doporučeno 1500+)
 - Hop Limit \sim TTL
 - přidána pole 'Next Header' a 'Flow Label'

IPv6 header - detaily (pokrač.)

- **Flow Label** odlišuje tok na síťové vrstvě
- **Traffic Class** podobně jako Type Of Service v IPv4 pro differentiated services
- **Next Header** podobně jako Protocol v IPv4 charakterizuje hlavičku následující po IPv6 hlavičce
- **Hop Limit**
 - maximální počet hopů
 - každý router po cestě sníží o jedničku
 - nejsou checksumy na IP vrstvě - nižší overhead

IPv6 adresy - formát zápisu

- 16 bytů oddělených dvojtečkami
- :: nahrazuje souvislou posloupnost nul může se vyskytovat pouze jednou v celém zápisu adresy
- souvislé bloky adres vyjadřujeme pomocí CIDR-like zápisu IPv6/"prefix"

IPv6 adresy - příklad

0:0:0:0:0:0:0:1 loopback, totéž co ::1

0:0:0:0:0:0:0:0 není specifikováno, totéž co ::

3ffe:80ee:038f:0000:0000:0000:0000:0002
unicast (viz. dále), totéž co 3ffe:80ee:38f::2

Adresace (RFC 2373)

- Unicast
- Anycast
- Multicast

(žádný broadcast)

IPv6 unicast adresa je maskovatelná podobně jako CIDR v IPv4. bity nejvíce vlevo pro síť, bity nejvíce vpravo pro hosty

Alokace v IPv6 (RFC 1884)

routable Aggregatable Global Unicast Address Public Topology prefixes:

- formát: TLA(16b)-NLA(s)(32b)-SLA(16b)-Interface id(64b)
- 3FFE::/16
TLA (6bone testbed)
- 2001::/16
TLA (produkční alokace)

Alokace v IPv6 - SLA

- SLA jsou přidělovány organizacím
 - obdoba subnetu v IPv4
 - mnohem větší prostor (65536 subnetů)

Unicast

- Link-local unicast
FE80::/10
použitelné pouze lokálně (ND, RD)
- Site-local unicast
FEC0::/10
formát FEC0::<subnet id (16bit)>:<interface id>

Unicast (pokrač.)

- global unicast
 - `::1` (loopback)
 - `::<IPv4 adresa>` (auto tunneling)
- IPv4 kompatibilní adresy
 - `0:0:0:0:0:0:192.168.30.1 =`
 - `::192.168.30.1 = ::COA8:1E01`

Multicast (RFC 2375)

- FF00::/8
- FF02::1
(všechny uzly na lokální síti)
- FF02::2
(všechny routery na lokální síti), užitečné pro testování tunelů

Anycast (RFC 2526)

- adresa přiřazená několika interfacům/uzlům
- paket směřovaný na anycast adr. je doručen vždy nejbližšímu interface (DNS discovery)

Autokonfigurace (RFC2462)

- není potřeba žádný centrální server pro přidělování adres
- vše je Plug-n-Play
- adresa typu Link-local (prefix FE80::0/10) na interface (automaticky) je nutnou prerekvizitou

ND (Neighbor Discovery) (RFC 2461)

- pomocí ICMPv6 paketu
 - neighbor solicitation packet (request)
 - neighbor advertisement message (reply)
 - odesílané na adresy typu solicited-node multicast
- obdoba ARP protokolu v IPv4
- routery se konfigurují ručně, ne automaticky !

ND (Neighbor Discovery) (pokrač.)

- předpokládá jednoznačný ident u interface
- neobsahuje info o DNS serveru (nevýhoda)
- detekce duplikátních adres (pomocí Neighbor Solicitation, multicast)

Autoconfig v IPv6 (pokračování)

Router solicitation

- posílá klient, router okamžitě odpovídá

Router advertisement

- routery periodicky posílají ICMPv6 pakety router solicitation paket
- ty mohou obsahovat i více prefixů
- stateless konfigurace
stateless - neexistuje tabulka přiřazení adres

Bezpečnost v IPv6 (IPsec)

- povinně v IPv6
- oproti IPv4 IPsec může chránit data end-to-end (absence NAT)

Prostředky

- authentication (AH) extension header
- ESP extension header

Bezpečnost v IPv6 (pokrač.)

Zaručuje

- confidentiality (šifrování)
- ochrana integrity
- autentizace
- ochrana proti replay útokům

Mobilita

- zařízení si udrží adresu při cestě mezi sítěmi
- návrh není ještě ustálen
- pomocí Extension hlavičky
- zastupuje ho jeho *domácí agent*
- zařízení průběžně informuje agenta o jeho aktuální poloze (Binding Update)

Mobilita (pokrač.) - algoritmus

- externí zařízení pošle žádost o navázání spojení na adresu z DNS
- žádost dorazí k domácím agentovi
- agent odpoví externímu stroji pomocí Ext. hlavičky daty o aktuální pozici MN
- další komunikace pak probíhá přímo

Mobilita (pokrač.) - algoritmus

- mobilní zařízení změní adresu
 - pošle aktualizaci (Binding Update)
 - * domácímu agentovi
 - * všem strojům s kterými komunikuje
- každá aktualizace musí být autentizována

DNS v IPv6 (RFC 1886)

Teorie

- AAAA zaznam - ekvivalent A zaznamu
 - kiberdigi IN AAAA 3ffe:80ee:38f::2
- ip6.int - reverzní záznamy (obdoba PTR)
 - 2.<20 nul>.f.8.3.0.e.e.0.8.e.f.f.3.ip6.int.

DNS v IPv6 (pokrač.)

Praxe

- existují patche pro Bind 8.1.2 pro podporu IPv6 dotazů
- Bind 9.1.x umí IPv6 dotazy

Přechod na IPv6

- nelze překlopit celý Internet v jediném dni
- proces přechodu nebude kompletní před 2010
- NAT zvolnil tlak na zavedení elegantnějšího řešení - dává více adres pro koncová zařízení

Strategie přechodu (RFC 1933)

- dual-stack zařízení
- tunnelling
 - automatické tunelování
 - * 6to4
 - * IPv4 kompatibilní IPv6 adresy
 - konfigurované tunelování

Strategie přechodu (pokrač.)

Translation gateways

- NAT-PT
 - Network Addr. Translation - Protocol transl.
 - umožňuje komunikaci IPv6-only hostům s IPv4 only hosty
 - mění hlavičky IPv6 na IPv4
 - vyžaduje ALG (Application Level Gateway) překlad mezi IPv4 a IPv6
DNS dotazy a odpověďmi
- faithd - IPv6-to-IPv4 relay
 - pouze jeden port

IPv6 tunely přes IPv4 sítě

- Automatické
 - routery na koncích tunelu musí umět IPv4 i IPv6
- Ručně vytvářené
 - tunnel source a tunnel destination automaticky
 - IPv4 kompatibilní IPv6 adresy

6to4 tunely

- dovolí IPv6-only hostům komunikovat s IPv6-only hosty přes IPv4 síť
- zapouzdřuje IPv6 do IPv4
- border router musí být dual-stack
- routery na hranicích IPv6-nativní sítě se propojí tunely
 - (NLA = IPv4 adresa)
 - 2002:IPv4 adresa border routeru::/48

6to4 tunely (pokrač.)

Terminologie

- 6to4 border router - 6to4 + sit interface
- 6to4 relay router - 6to4 <-> IPv6 native

6to4 tunely (pokrač.)

RFC 3068

- 2002:C058:6301:: anycast adresa všech 6to4 relay routerů
- default route na 6to4 border routerech (::/0) -> 6to4 IPv6 anycast address
- 6to4 relay routery ohlašují anycast adresu via IGP
- správci mohou nechat relay routery propagovat do sousedních AS

Obecné strategie přechodu

IPv6 prosakuje postupně od okrajů sítě k jádru dovoluje kontrolovat cenu přechodu (oproti celk. upgradu)

Stupně integrace pro ISP

- poskytnout IPv6 zákazníkům
 - nenáročné na cenu
 - neovlivní IPv4-based služby

Obecné strategie přechodu (pokrač.)

- IPv6 v jádře sítě (backbone)
 - použití dual-stack routeru
 - IPv6-only routery jakmile převáží IPv6 traffic
- propojení IPv6 s ostatními ISPs

Aplikace pro IPv6

KAME (www.kame.net)

- japonsky projekt, core developereři z telco firem
- vývoj IPv6 kódu pro jádra BSD* systémů
- IPv6 patche pro populární sw
 - SSH
 - Apache (2.x)
 - Sendmail (MX zaznamy jako AAAA)
 - Mozilla
 - IKE, IPsec

IPv6 a OpenSource

- FreeBSD, NetBSD, OpenBSD - KAME stack současný stack vznikl sloučením stacku od KAME, Inria, NRL v stabilní fázi:
 - autoconfig (rtsold, rtadvd)
 - paket filtery (ipf, pf, ipfw)
- Linux - USAGI projekt

IPv6 a OpenSource (pokrač.)

MIP6 (Mobile IPv6)

- Net,FreeBSD
- had (home agent discovery daemon)
- rtadvd
- IPsec pro ochranu binding, tunelovaných zpráv

IPv6 a OpenSource (pokrač.)

USAGI - Linux IPv6 Devel. Project

- WIDE, KAME, ...
- TAHI project - conformance testing
- další stable verze: léto 2003

USAGI - status

- status:
 - 2.4.x
 - anycast
 - IPsec
 - ICMPv6
 - ISATAP
 - lepší selekce zdrojových adres
 - IPv6-over-IPv4 tunely
 - opraveny chyby v podpoře autokonfigurace, NDP

IPv6 a ClosedSource

- Cisco vývojové verze IOSu
 - oficiální verze pro IPv6 12.2(2)T release train
 - 12.2(13)T/B
- Nortel
- Juniper - ready
- Nokia (R. Hinden)

IPv6 a ClosedSource (pokrač.)

- Windows - XP, 2003 server
 - 6to4 tunely
 - ISATAP

Proč nabízet IPv6 ? (očima ISP)

- protože konkurence bude IPv6 nabízet také
- dát domácím uživatelům více než jednu IP adresu
- protože routery už umí IPv6
- protože domácí uživatelé začnou používat IPv6
 - dokonce i tehdy když o tom nebudou vědet
 - Windows budou používat 6to4 pokud ISP nebude nabízet IPv6
 - zákazníci budou na helpdesk volat s divnými otázkami ohledně tunelování

Proč nabízet IPv6 ? (pokrač.)

- všechny operační systémy jsou IPv6 ready
- Microsoft tlačí na zavedení IPv6
 - peer-to-peer networking (online gaming)

IPv6 u nás

situace březen 2003

- CESNET má IPv6 síť delší dobu (CESNET2)
- tři komerční ISPs mají od RIPE produkční prefix
- NIX je IPv6 ready
 - dva komerční ISP peerují přes IPv6
- nikdo nenabízí nativní IPv6 konektivitu
- XS26 je občas nestabilní

Stav IPv6

Extrémní přístupy

- o IPv6 se hodně mluví, ale nic se neděje
- "IPv6 vyřeší všechny naše problémy"

Konec