Seminář IPv6@FEKT 28.1.2009

Petr Lampa: Zavádění IPv6 v praxi

1. Proč zavádět IPv6?
2. Jak lze zavádět?
3. Jak na VUT?
4. Přidělování IPv6 adres
5. Stav zavádění IPv6
6. Aktuální problémy

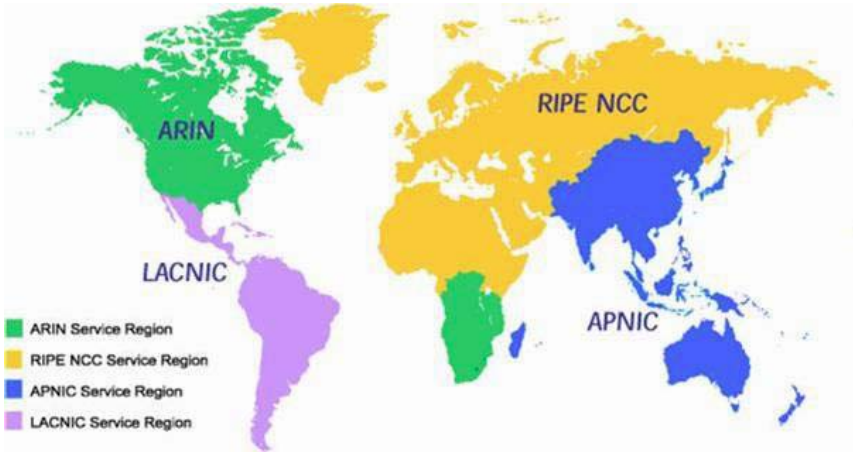
Přechod na IPv6 (2008/3)

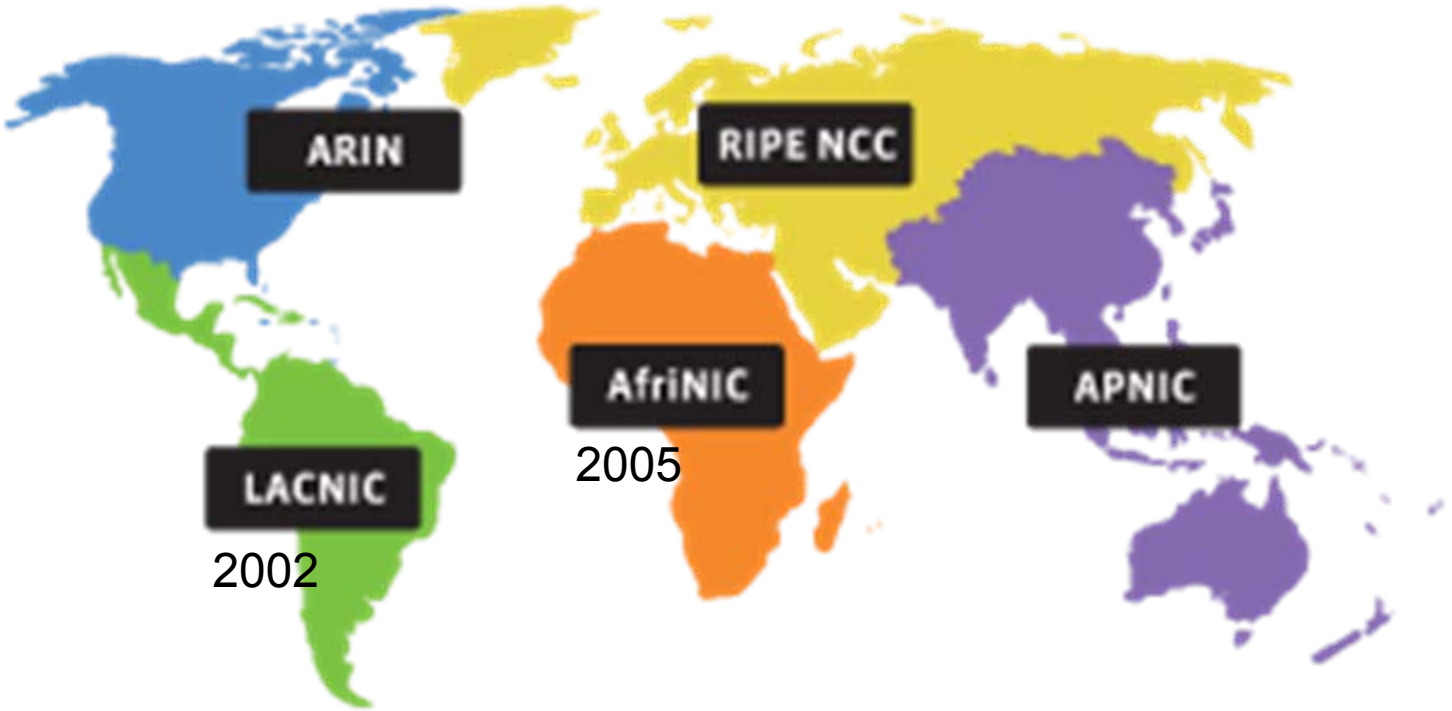
- **Říjen 2007** – 55. zasedání RIPE, usnesení vyzývá k přechodu na IPv6 vzhledem k vyčerpání IPv4 adres do 2-4 let
- **Únor 2008** - IANA po několikaletém úsilí uvolnila blok 14.0.0.0/8 (never ever)
- **Únor 2008** – vloženy IPv6 adresy 6 kořenových DNS serverů
- **Březen 2008** - projekce vyčerpání volných IPv4 adres v modelu: IANA 2/2011, RIR 5/2012
- **Březen 2008** – 71. zasedání IETF, draft „An Internet Transition Plan“, verze 0-7/2007 (2009/2010), verze 1/2 (2010/2011)

Přechod na IPv6 (2008/11)

- **Vint Cerf** (26.9.2008) – IPv4 adresy dojdou do 2010, dělá se málo pro přechod na IPv6.
- **CNNIC** – Číně dojdou IPv4 adresy do 830 dnů (2011), IPv6 je zatím pouze akademické, je třeba aby ISP začali s přechodem na IPv6.
- **EU** (27.5.2008) – ADVANCING THE INTERNET Action Plan for the deployment of Internet Protocol version 6 (IPv6) in Europe **-** veřejný sektor musí začít neprodleně s přechodem na IPv6 (své sítě, Web, e-gov), cíl 25% uživatelů IPv6 v roce 2010.
- **RFC5211** (7/2008) – Internet Transition Plan

RIR a IANA

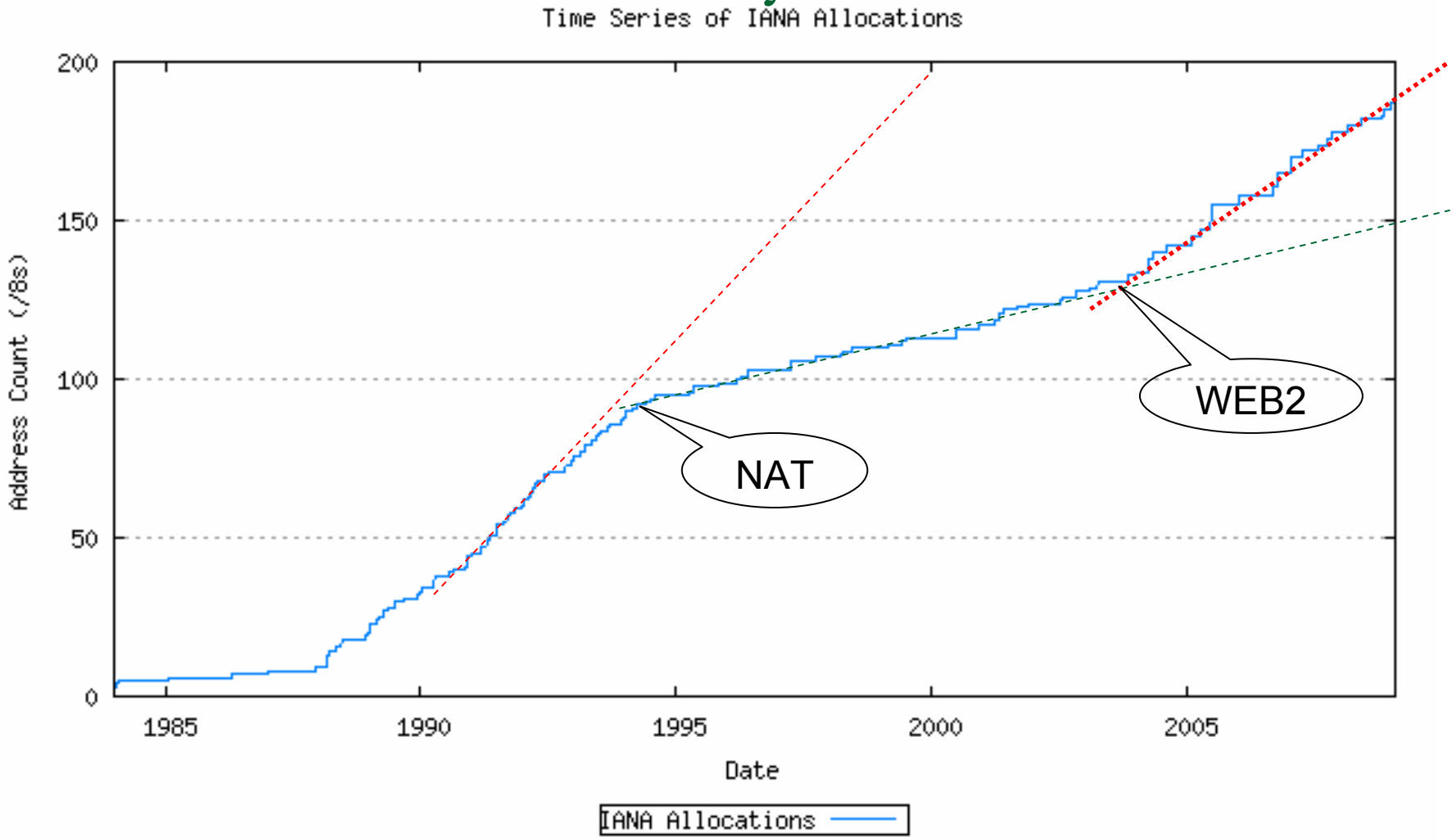

ARIN
RIPE NCC
LACNIC
APNIC
ARIN Service Region
RIPE NCC Service Region
APNIC Service Region
LACNIC Service Region


ARIN
RIPE NCC
AfriNIC
2005
APNIC
LACNIC
2002

Stav u RIR

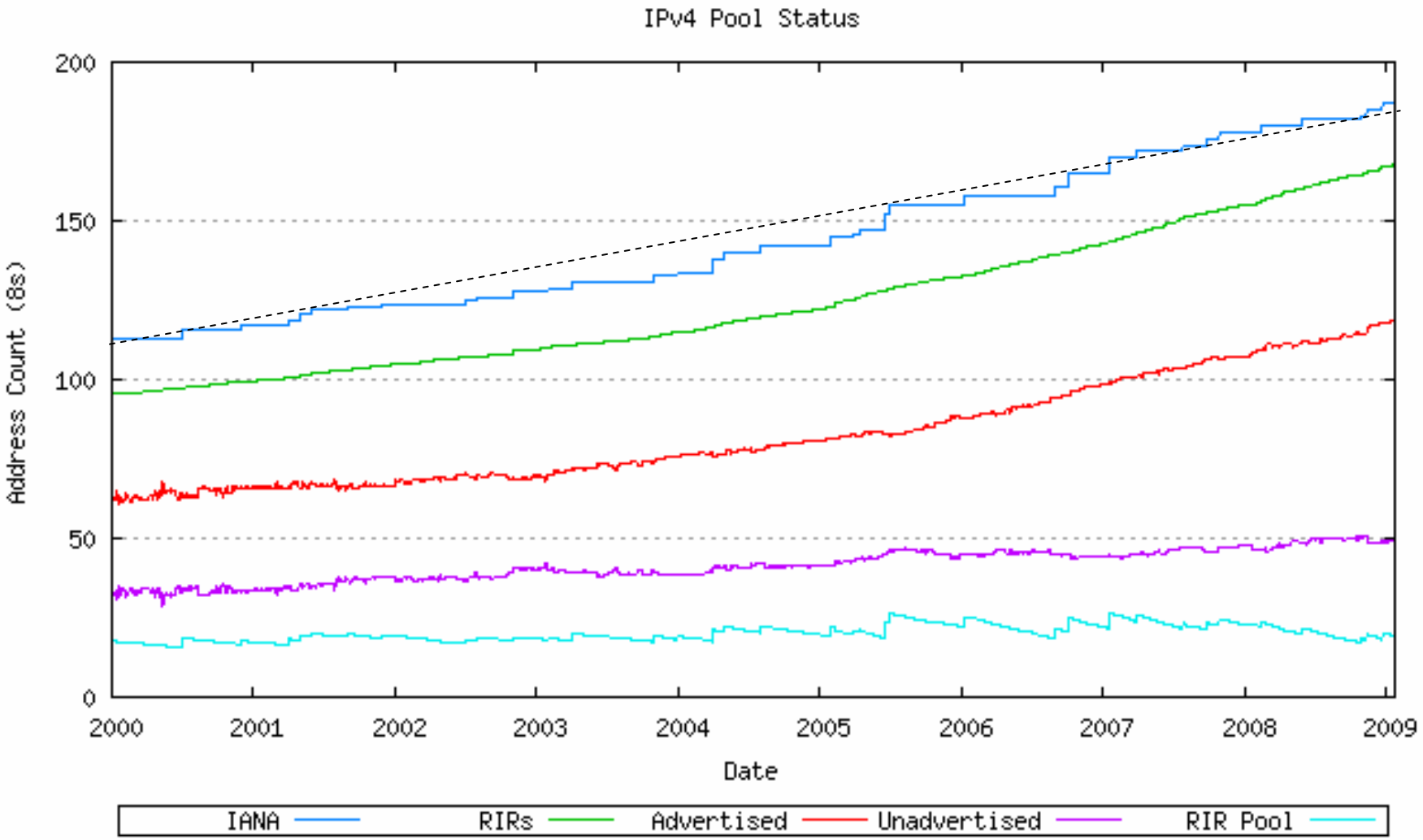
- IANA 2008/9 - Global Policy for the Allocation of the Remaining IPv4 Address Space (IPv4 Countdown)
- LACNIC - http://portalipv6.lacnic.net/en
- AFRINIC - http://www.afrinic.net/IPv6/
- ARIN - http://www.arin.net/v6/v6-info.html
- APNIC - http://www.apnic.net/services/ipv6_guide.html
- RIPE – www.ripe.net (nic?)

Přidělené IPv4 adresy

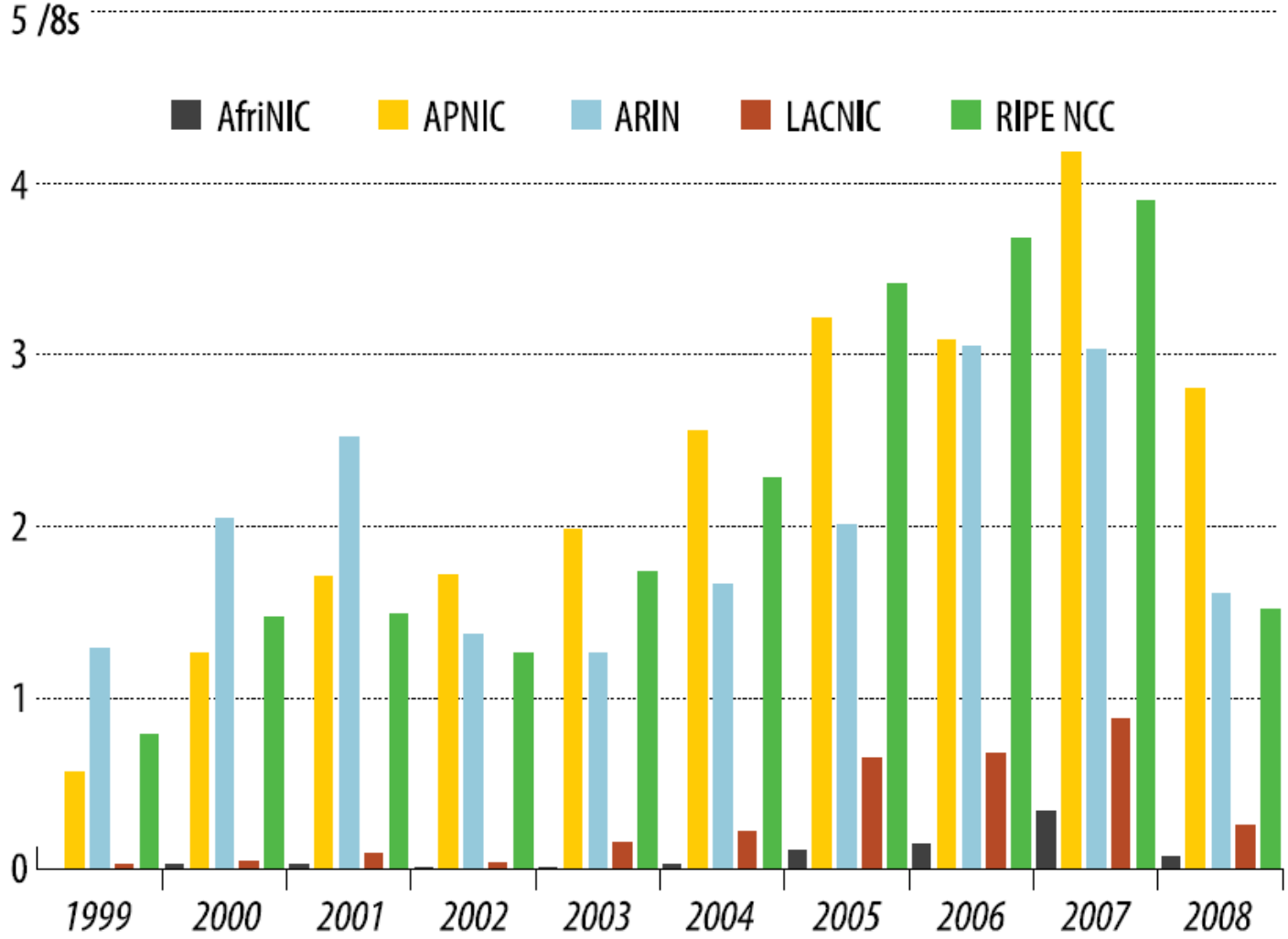


http://www.potaroo.net/tools/ipv4/index.html

Podrobněji od 2000 (2009/1)

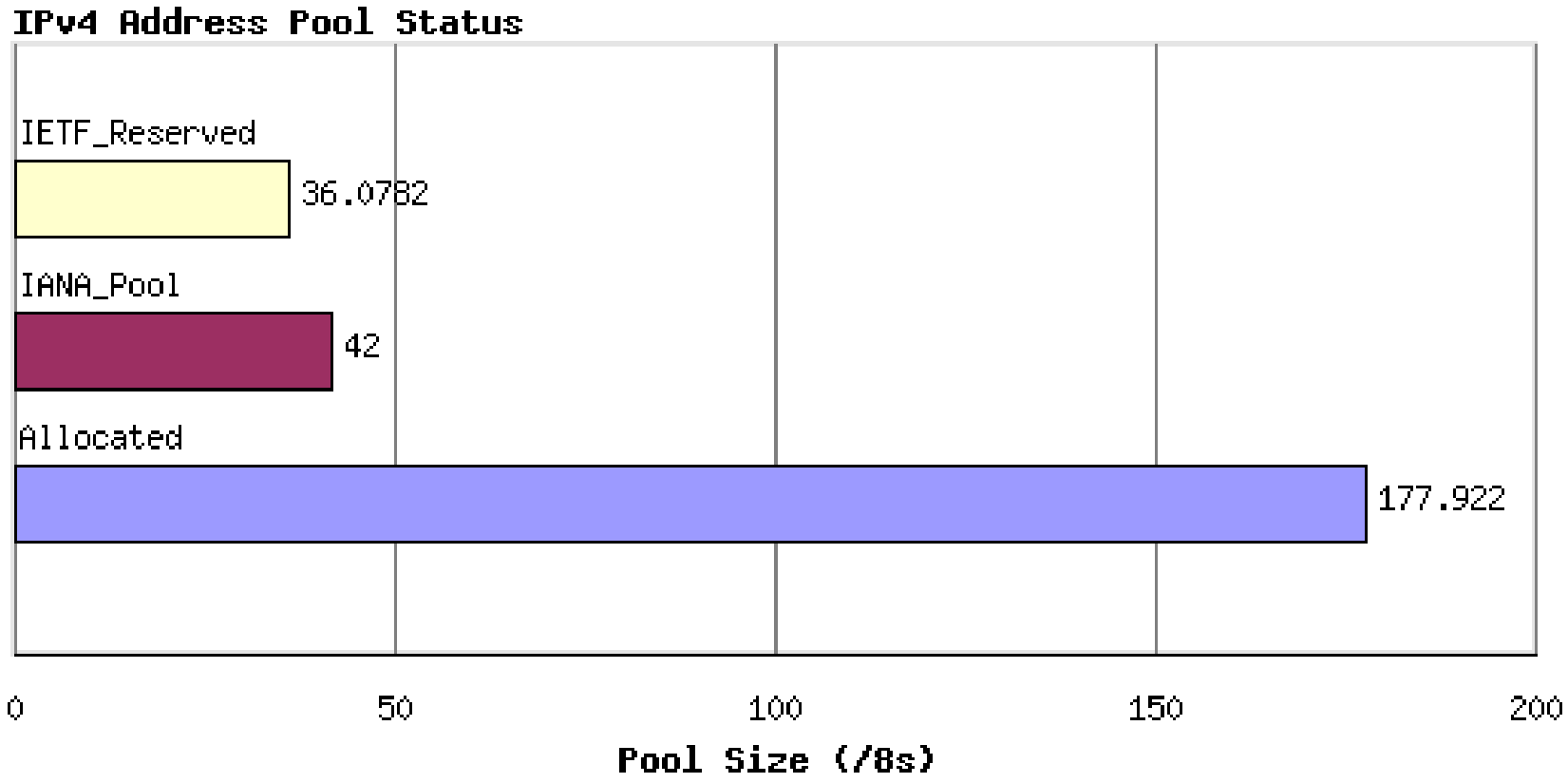


Alokace RIR v posledních letech

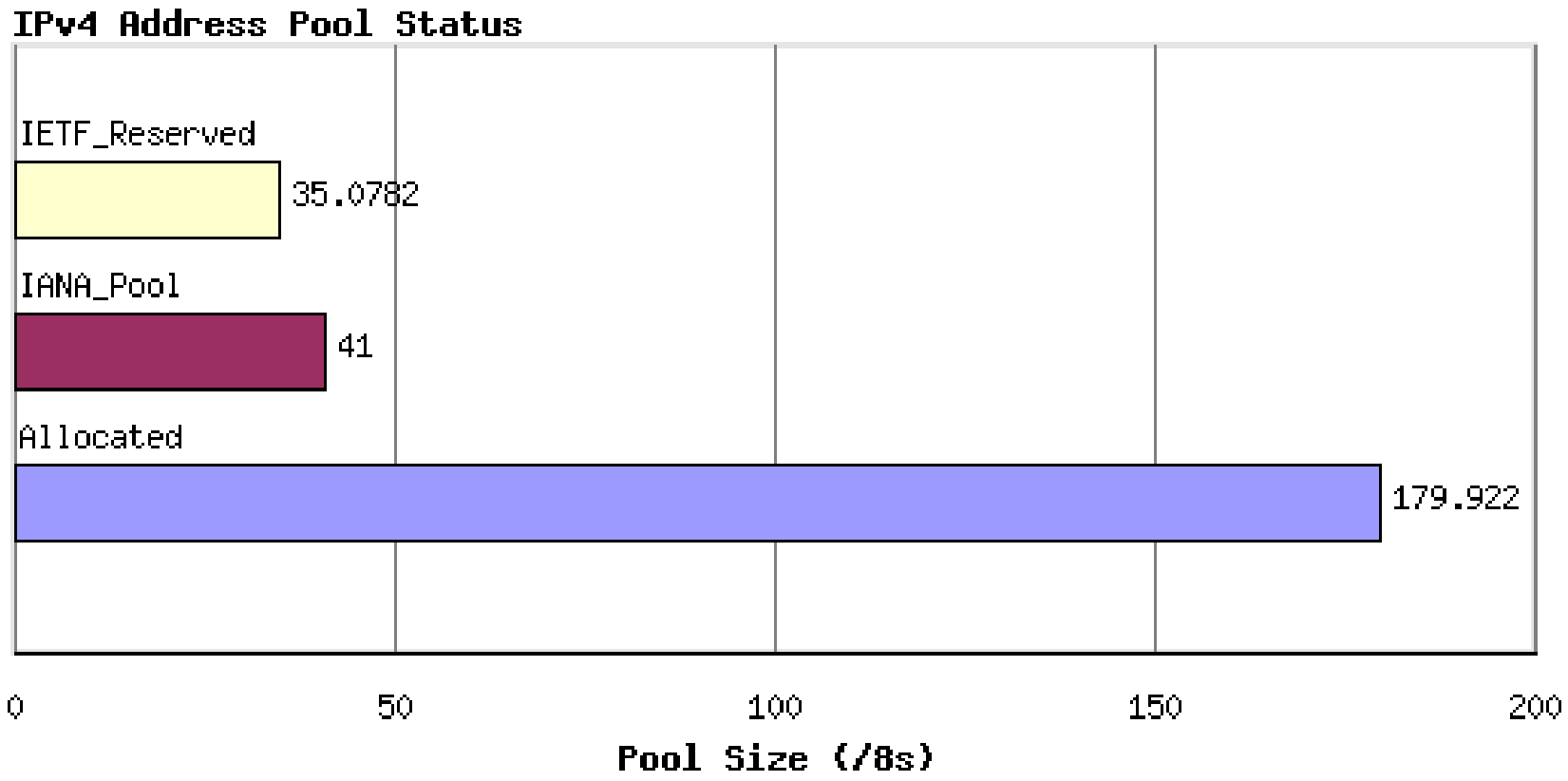


http://www.nro.net/documents/presentations/jointstats.v1.0608.pdf

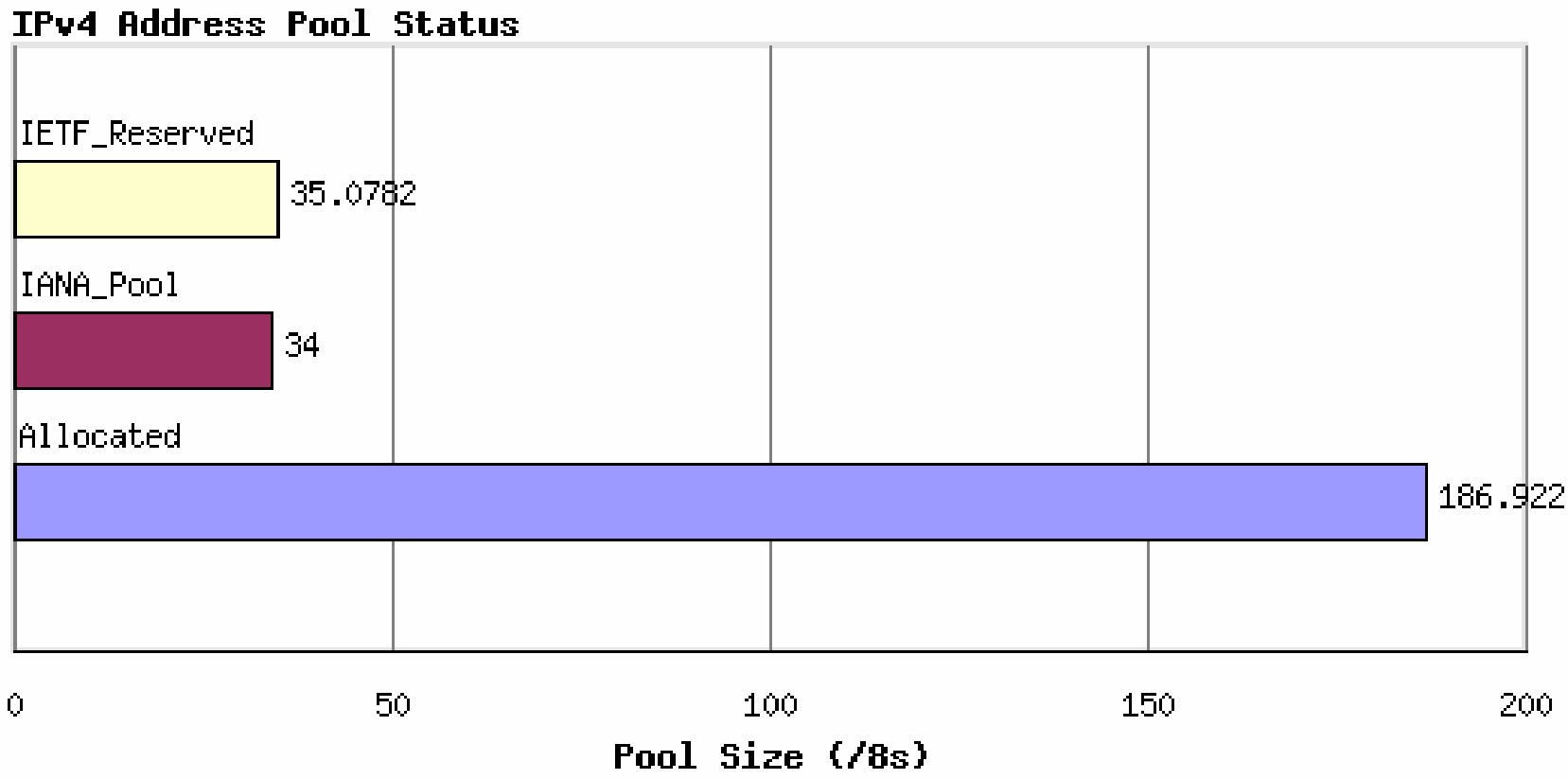
Alokace IPv4 adres (2007/11)

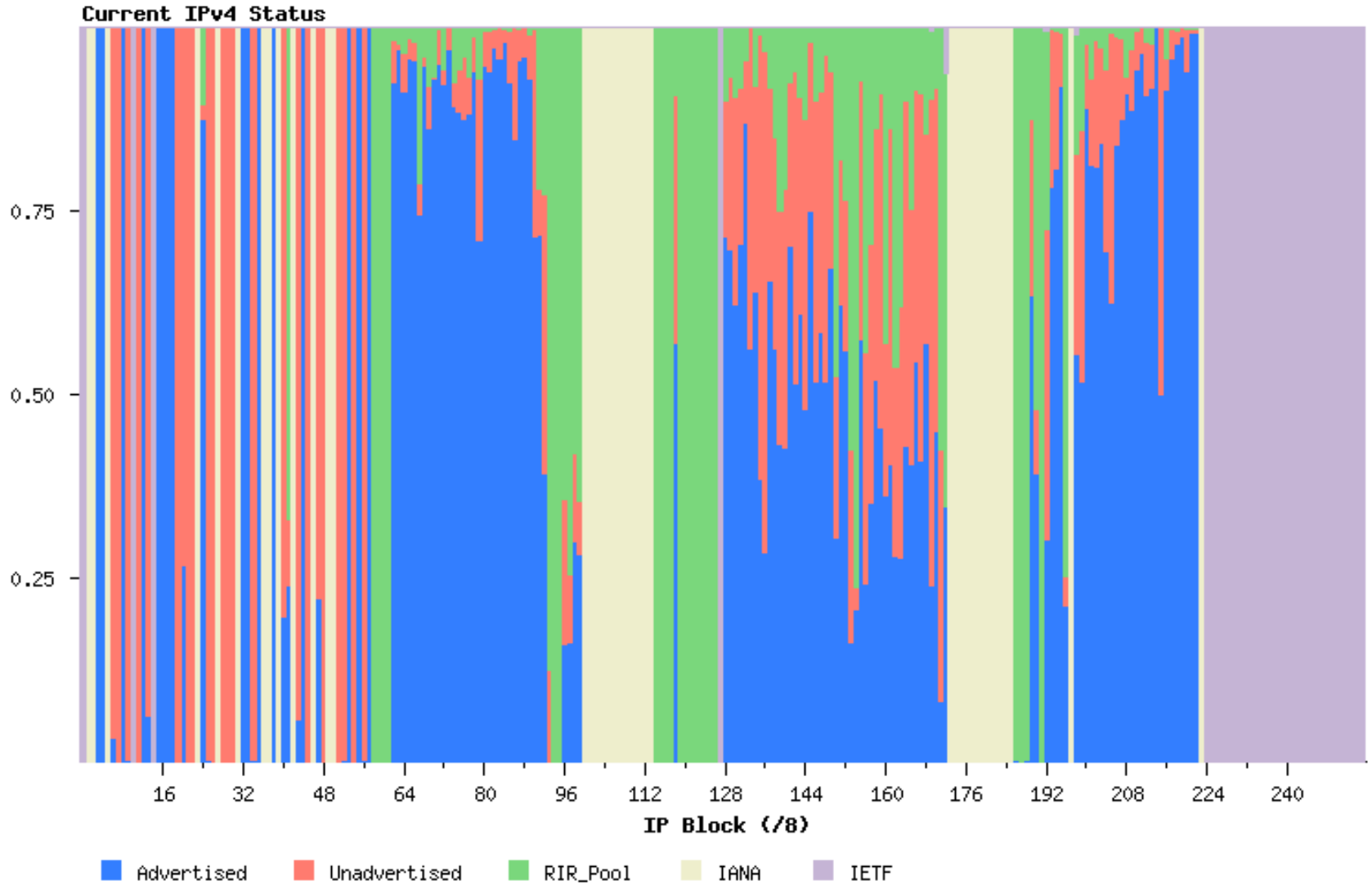


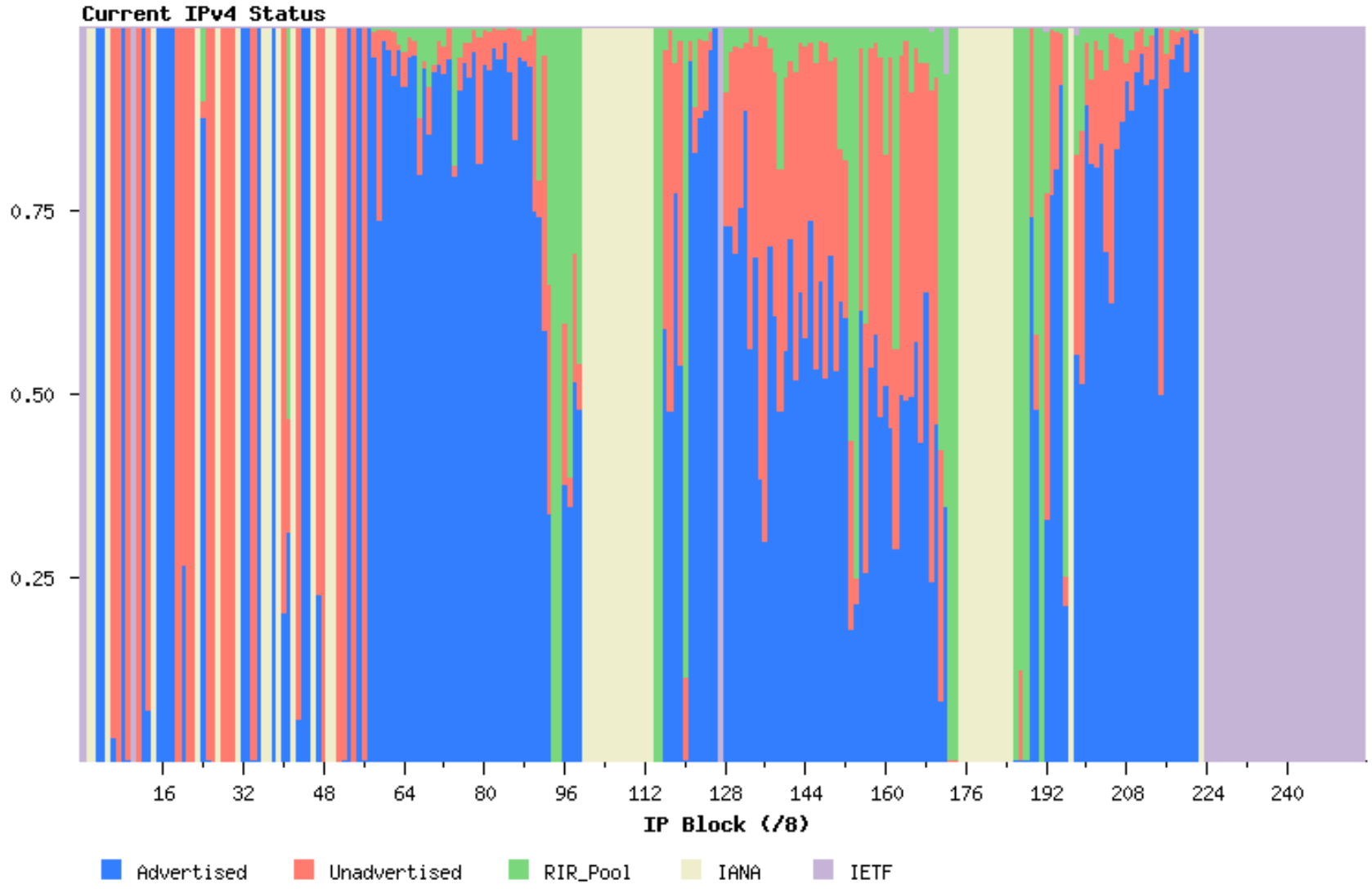
Alokace IPv4 adres (2008/3)

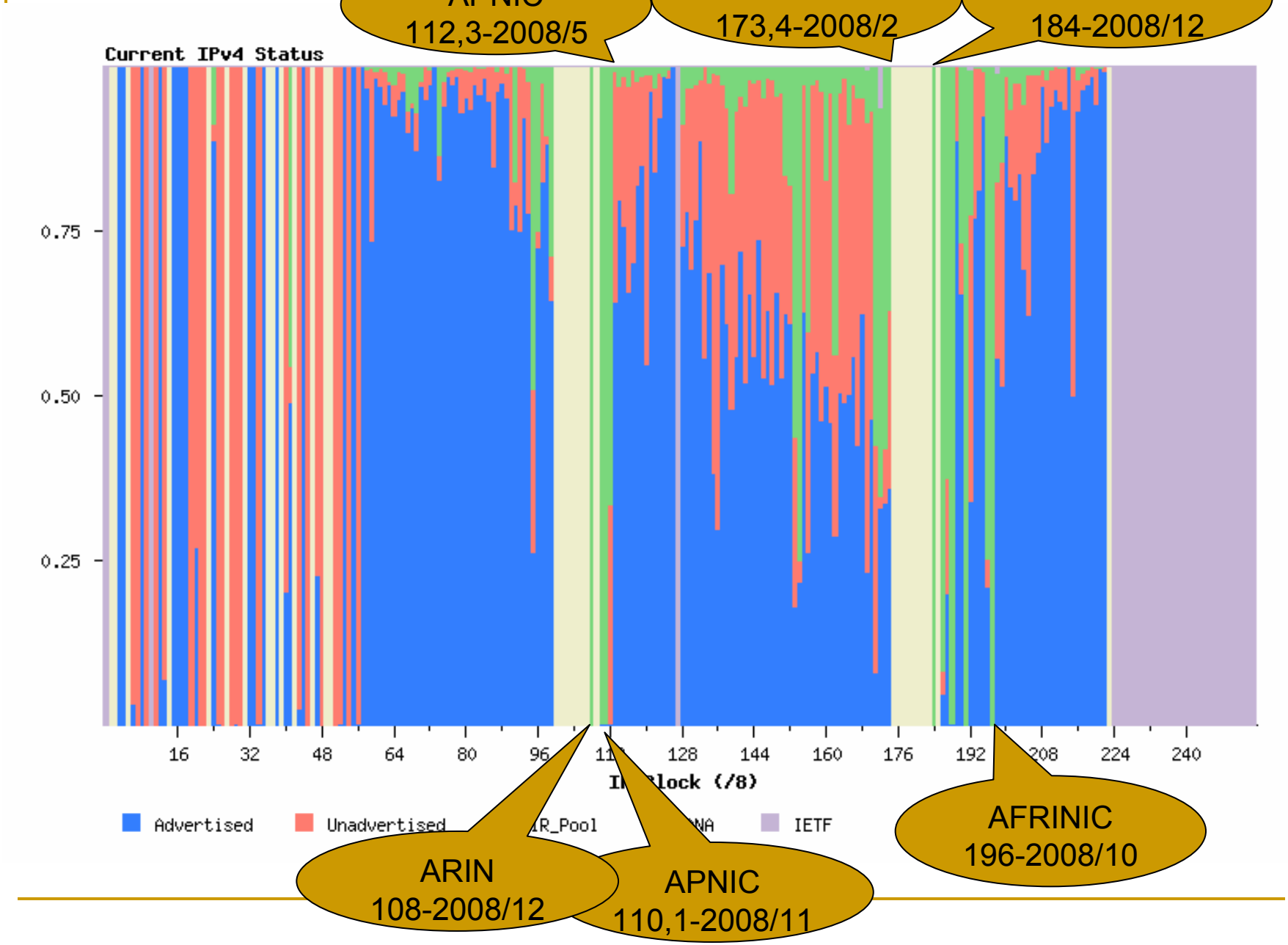


Alokace IPv4 adres (2008/12)

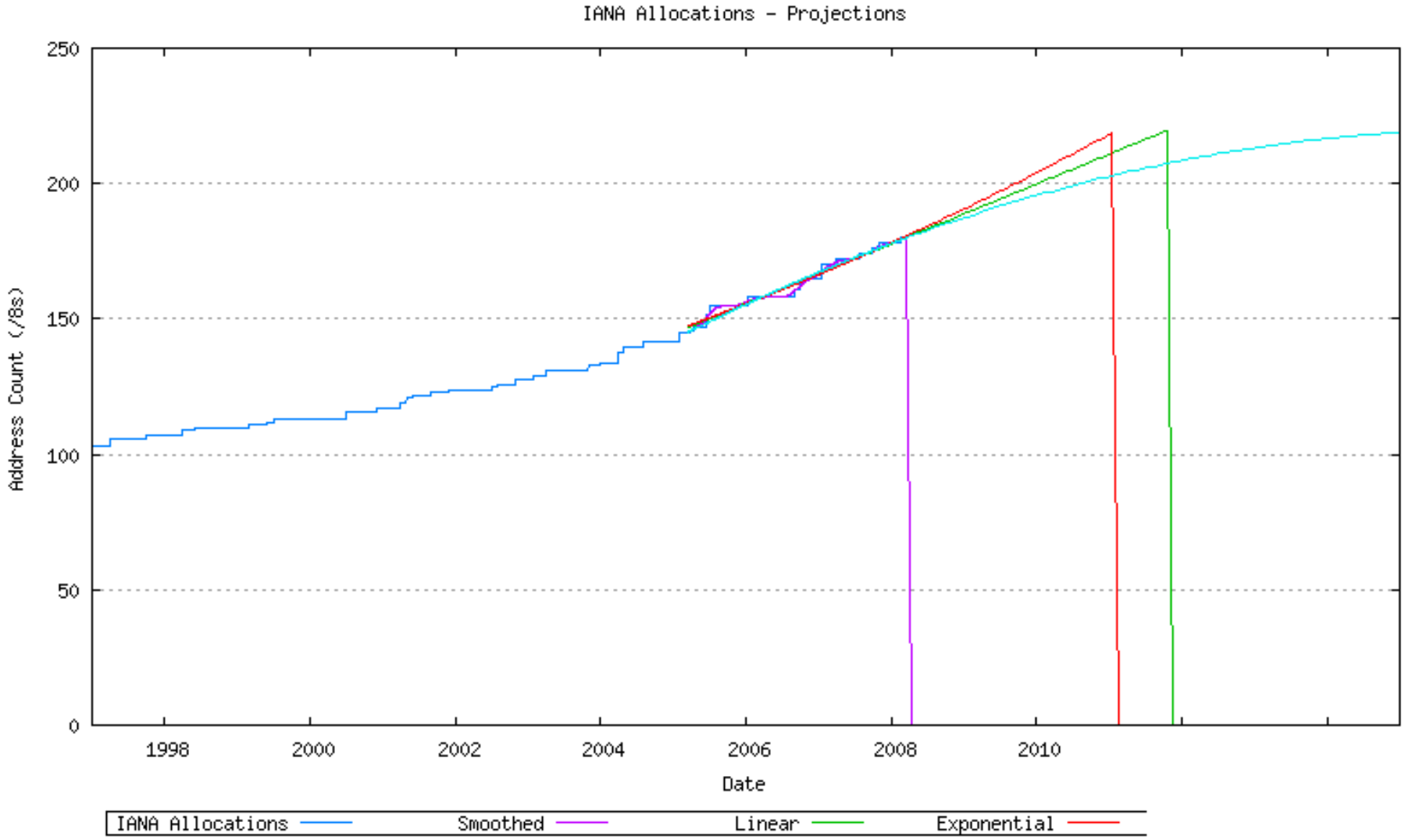



Current IPv4 Status
0.75
0.50
0.25
16
32
48
64
80
96
112
128
144
160
176
192
208
224
240
IP Block (/8)
Advertised
Unadvertised
RIR_Pool
IANA
IETF

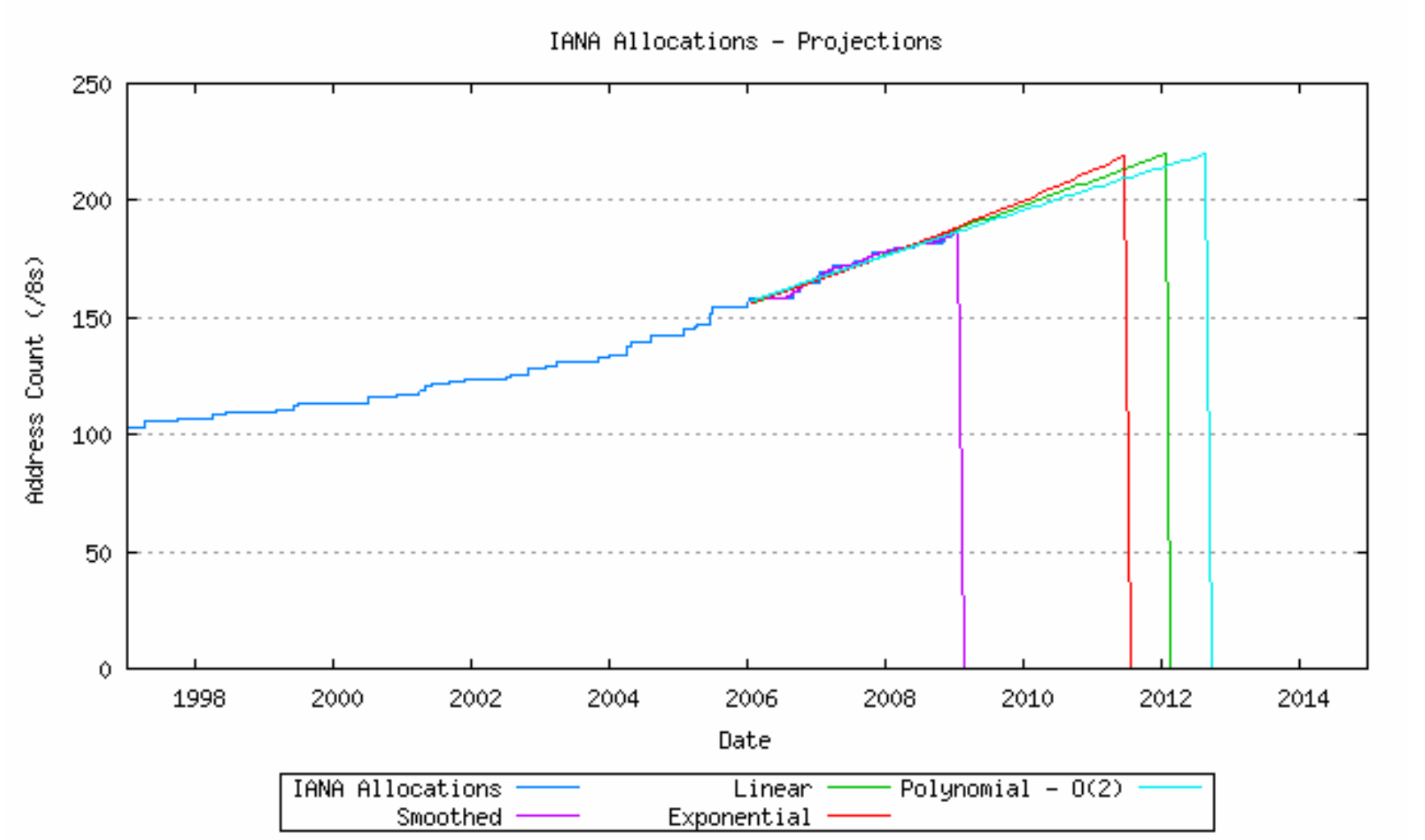

Current IPv4 Status
0.75
0.50
0.25
16
32
48
64
80
96
112
128
144
160
176
192
208
224
240
IP Block (/8)
Advertised
Unadvertised
RIR_Pool
IANA
IETF


APNIC
112,3-2008/5
ARIN
173,4-2008/2
ARIN
184-2008/12
Current IPv4 Status
0.75
0.50
0.25
16
32
48
64
80
96
128
144
160
176
192
208
224
240
Advertised
Unadvertised
IETF
ARIN
108-2008/12
APNIC
110,1-2008/11
AFRINIC
196-2008/10

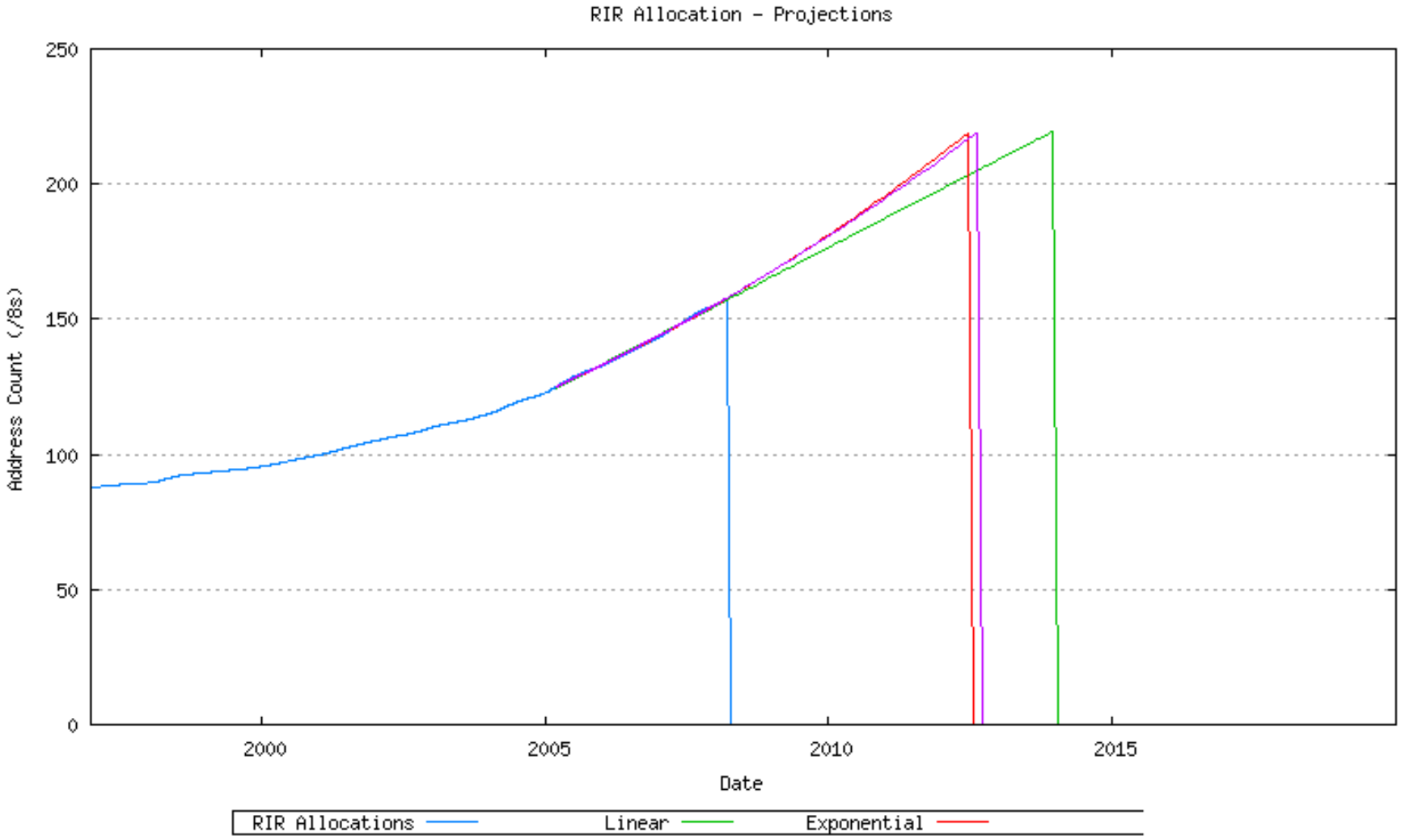
Projekce vyčerpání IANA (2007/11)



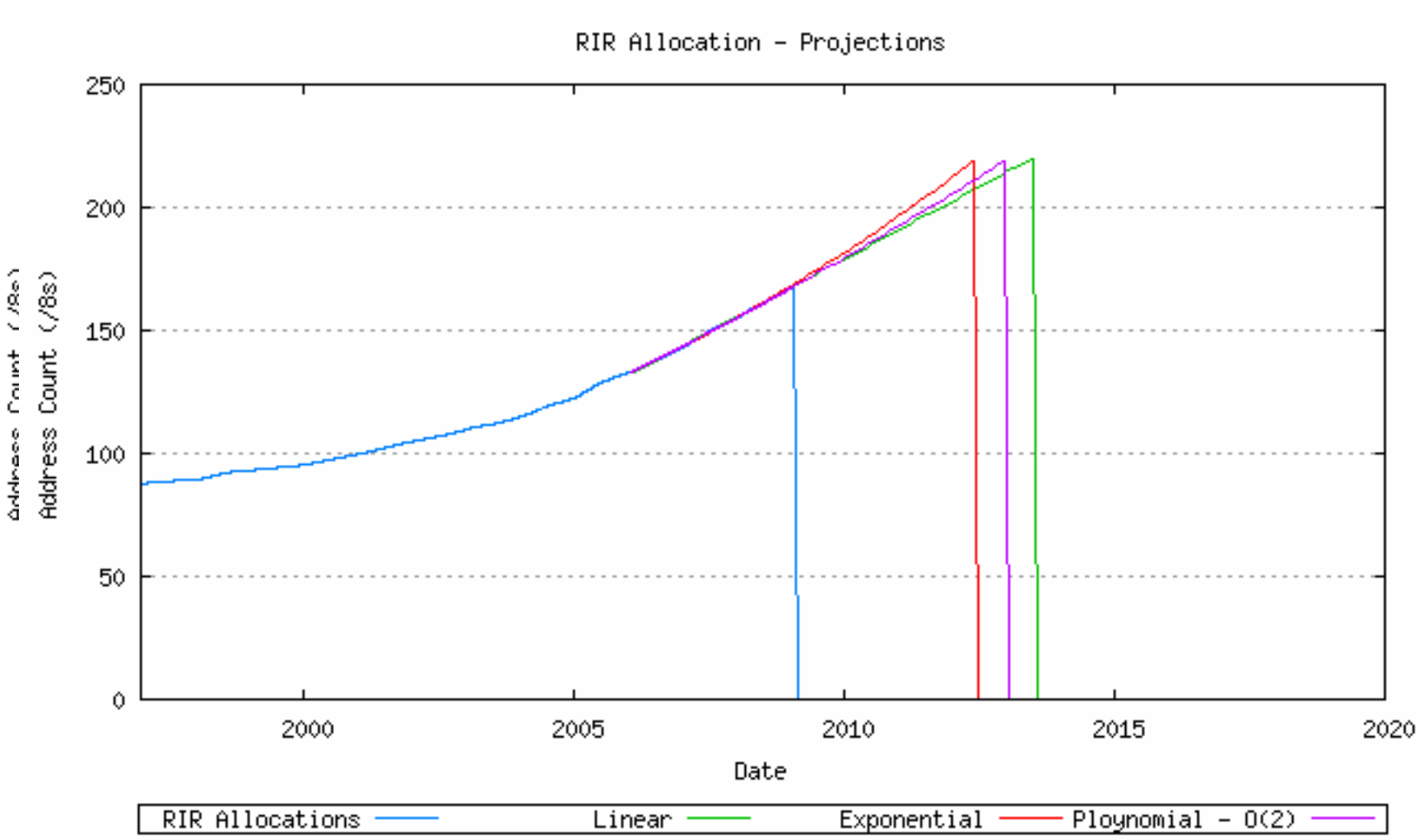
Projekce vyčerpání IANA (2009/1)



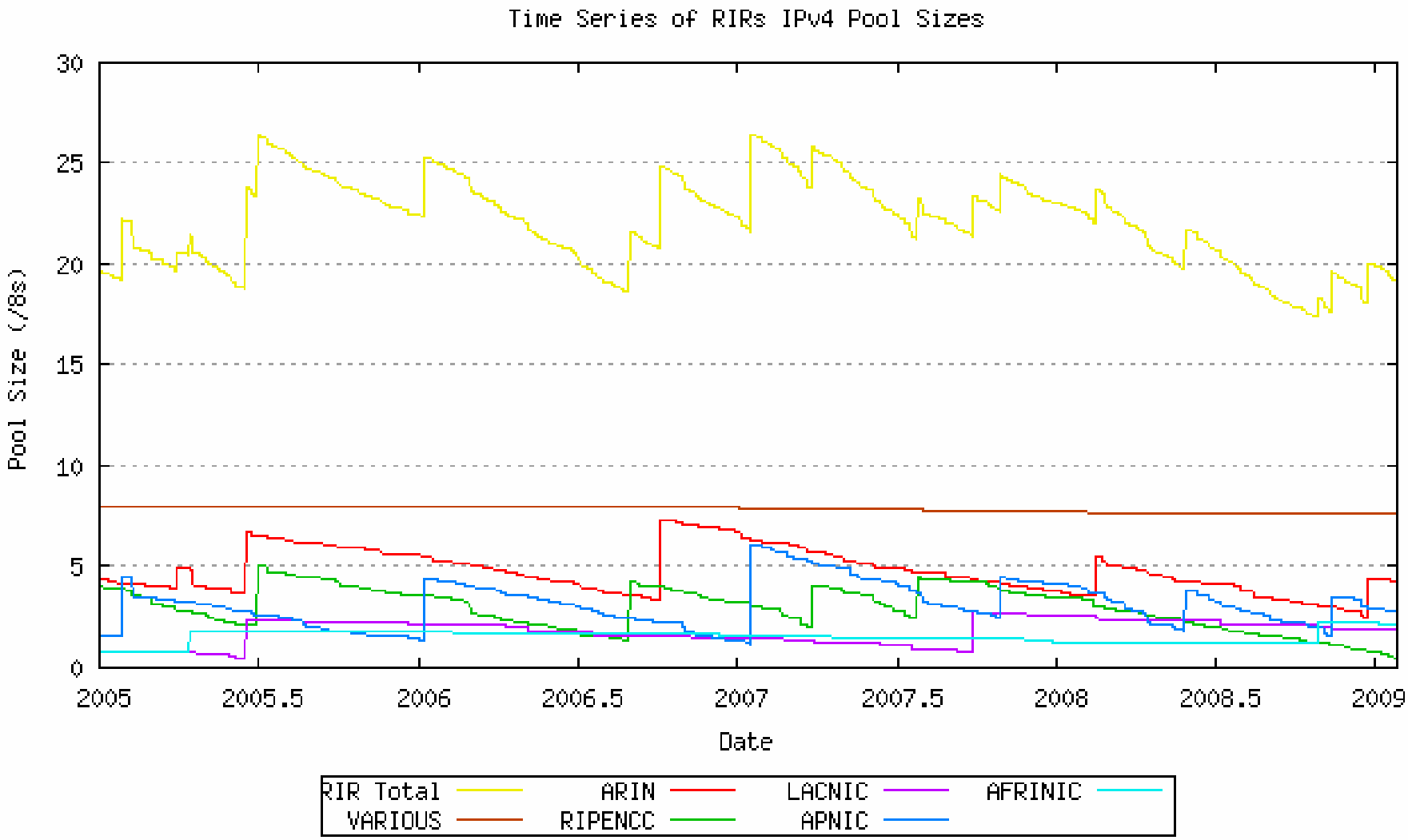
Projekce vyčerpání RIR (2007/11)



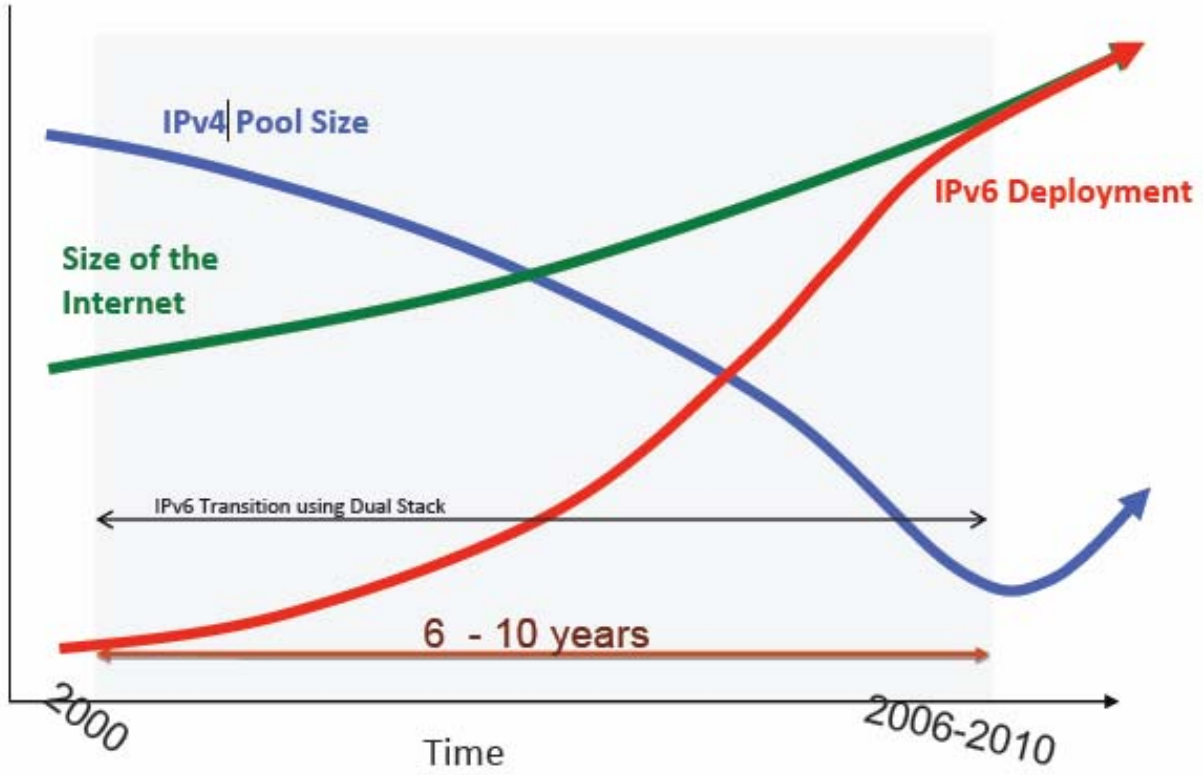
Projekce vyčerpání RIR (2009/1)



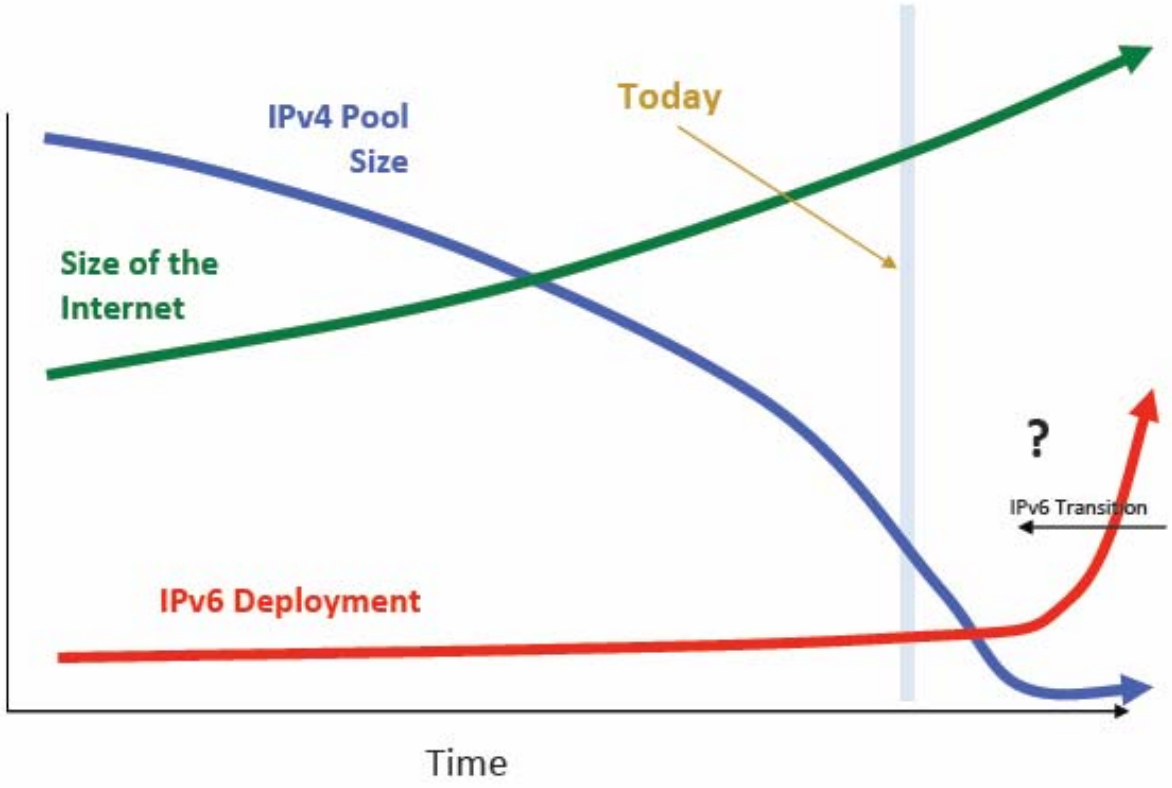
Stavy poolů RIR 2009/1



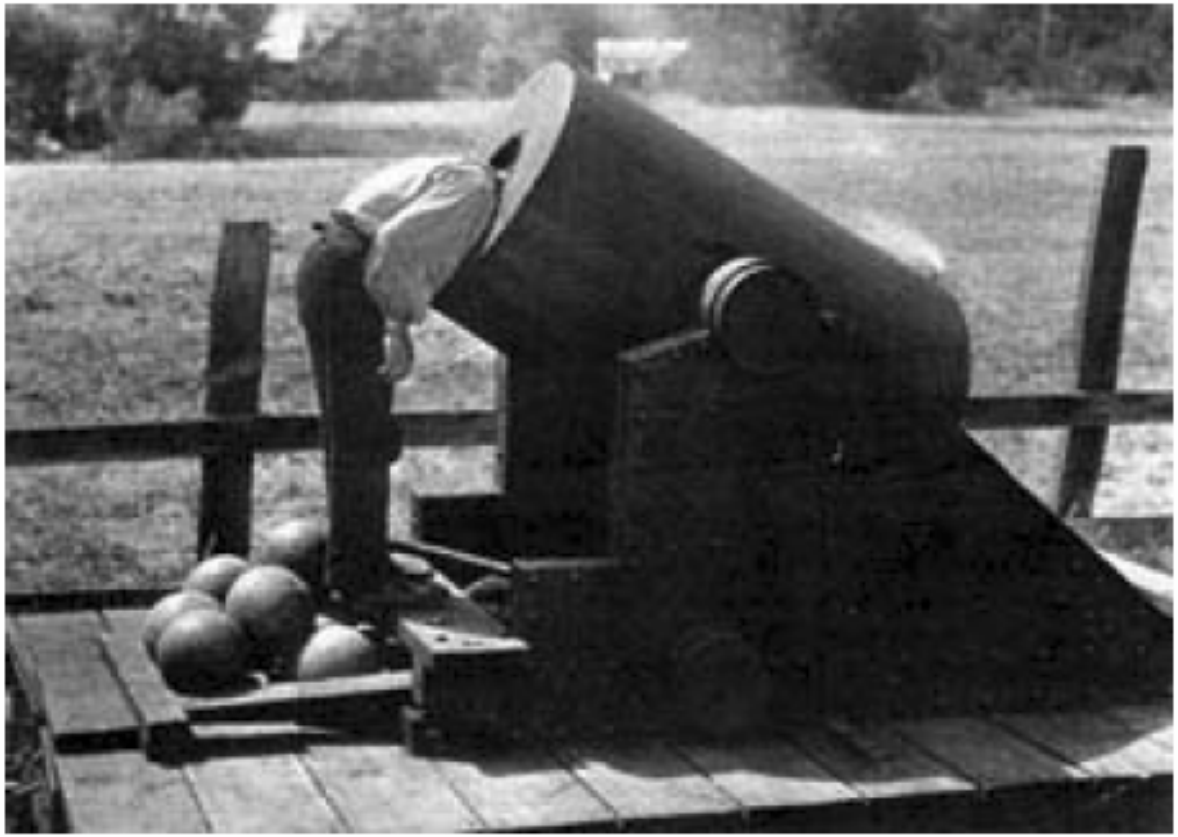
Ten years ago we had a plan …



What's the revised plan?



Its just not looking good is it?



http://www.ripe.net/ripe/meetings/ripe-56/presentations/Huston-Measuring_IPv6_Deployment.pdf

Viditelnost IPv6 prefixů (www.sixxs.net)

Pos	Country	Visible	Alloc	VP
1	United States	226	704	8.17%
2	Germany	109	210	3.94%
3	United Kingdom (Great Britain)	66	142	2.39%
4	Japan	72	137	2.60%
5	Netherlands, The	55	99	1.99%
6	France	32	74	1.16%
7	Switzerland	34	71	1.23%
8	Italy	34	68	1.23%
9	Australia	23	57	0.83%
10	Sweden	22	54	0.80%
11	Korea	19	54	0.69%
12	Canada	21	52	0.76%
13	Russia	21	52	0.76%
14	Czech Republic	24	42	0.87%

V ČR

Prefix	NetName	Owner	Allocated	First seen	Seenb
2001:678:1::/48	D-NS-NIC-CZ	CZ.NIC, z.s.p.o.	11.10.2006	11.12.2006 14:47	97%
2001:718::/32	CZ-TEN-34-20010521	CESNET Sub-TLA block	21.5.2001		100%
2001:7f8:14::/48	NIX-CZ-NET-IPv6-2003...	NIX.CZ, exchange point	3.2.2003		0%
2001:ae8::/32	CZ-IPEXNET-20030205	Ipexnet	5.2.2003	9.3.2003 5:30	100%
2001:af0::/32	CZ-GTS-20030211	GTS CZECH, a.s	6.2.2003	6.6.2003 11:21	40%
2001:b80::/32	CZ-PRAGONET-20030314	PragoNet	14.3.2003	28.7.2003 13:21	40%
2001:1478::/32	CZ-CECOM-20030618	CE Com Czech, s.r.o.	18.6.2003	15.9.2008 18:32	100%
2001:1488::/32	CZ-NIC-20030620	Top level domain .CZ Oper...	20.6.2003	8.9.2006 14:02	100%
2001:1508::/32	CZ-VOL-20030717	Czech On Line a.s.	17.7.2003	23.8.2003 15:41	100%
2001:1528::/32	CZ-CASABLANCA-200307...	Casablanca INT	24.7.2003	30.10.2003 12:50	100%
2001:1568::/32	CZ-BECOLINK-20030829	BECO Link spol. s r.o.	20.7.2007		0%
2001:1a48::/32	CZ-CDT-20040213	CD - Telekomunikace a.s.	13.2.2004		0%
2001:1ab0::/32	CZ-IGNUM-20040323	Ignum s.r.o.	23.3.2004	21.4.2004 11:00	100%
2001:1ae8::/32	CZ-RADIOMOBIL-200404...	Radiomobil a.s.	7.4.2004		0%
2001:41d8::/32	CZ-EUROTEL-20041118	Eurotel Praha, spol. s r....	18.11.2004	10.1.2005 12:17	100%
2001:4ba8::/32	CZ-SELFNET-20081001	CzechBone	1.10.2008		0%
2001:4cc8::/32	CZ-NETBOX-20050623	SMART Comp. s.r.o.	23.6.2005	25.6.2005 13:33	100%
2001:4de8::/32	CZ-INWAY-20051214	InWay. a.s.	14.12.2005	19.12.2005 2:17	100%
2a00:c40::/32	CZ-MAXPROGRES-200810...	MAXPROGRES, s.r.o.	6.10.2008		0%
2a00:ca8::/32	CZ-PODA-20081016	PODA a.s.	16.10.2008		0%

V ČR

Prefix	NetName	Owner	Allocated	First seen	Seen by
2a00:d38::/32	CZ-UPLTELECOM-200810...	UPL TELECOM s.r.o.	30.10.2008		0%
2a01:28::/32	CZ-SUPERNETWORK-2006...	SuperNetwork, s.r.o.	14.2.2006		0%
2a01:168::/32	CZ-HKFREE-20060714	Obcanske sdruzeni HKfree	14.7.2006	22.7.2006 23:32	100%
2a01:430::/32	CZ-MASTER-20070809	Master Internet s.r.o.	9.8.2007	17.8.2007 12:47	100%
2a01:490::/32	CZ-KLFREE-20070828	Sdruzeni Klfree.net	28.8.2007	16.9.2007 23:32	100%
2a01:510::/32	CZ-IJC-20071015	IJC, s.r.o.	15.10.2007		0%
2a01:538::/32	CZ-CZ-20071023	RFE/RL, Inc., organizacni...	23.10.2007		0%
2a01:5e0::/32	CZ-SLOANE-20071109	Sloane Park Property Trus...	9.11.2007	2.4.2008 16:17	100%
2a01:5f0::/32	CZ-COOLHOUSING-20071...	COOLHOUSING s.r.o.	13.11.2007	19.11.2007 10:02	0%
2a02:38::/32	CZ-NIX-20080214	NIX.CZ z.s.p.o.	14.2.2008	15.2.2008 15:32	100%
2a02:128::/32	CZ-ALFATELECOM-20080...	ALFA TELECOM s.r.o.	6.3.2008		0%
2a02:1d0::/32	CS-VISUAL-20080325	Visual Connection, spol. ...	25.3.2008	24.4.2008 22:47	98%
2a02:1f0::/32	CZ-MIRAMO-20080327	MIRAMO spol. s.r.o.	27.3.2008	16.5.2008 10:47	82%
2a02:350::/32	CZ-CL-NET-20080509	CL-NET s.r.o.	9.5.2008	14.5.2008 14:17	99%
2a02:3c0::/32	CZ-NETAIR-20080521	NetAir, s.r.o.	21.5.2008	6.8.2008 9:47	98%
2a02:4a8::/32	CZ-GLOBE-20080610	ACTIVE 24	10.6.2008		0%
2a02:4f0::/32	CZ-MAFRA-CZ-20080620	MAFRA, a.s.	20.6.2008		0%
2a02:570::/32	CZ-HA-VEL-20080701	ha-vel internet s.r.o.	1.7.2008		0%
2a02:598::/32	CZ-SEZNAM-20080704	Seznam.cz, a.s.	4.7.2008	17.8.2008 1:02	97%
2a02:618::/32	CZ-BROADNET-20080722	Broadnet Czech	22.7.2008		0%
2a02:768::/32	CZ-STARNET-20080908	STARNET s.r.o.	8.9.2008	30.9.2008 7:47	99%

Strategie nasazení IPv6

Jaké jsou možnosti:

1. Nic nedělat (nějak to dopadne)
2. IPv6 výhradně (hlavou do zdi)
3. Tunelování nebo překlad adres a protokolů
4. Dual stack IPv4 + IPv6

1. Nic nedělat

- Páteř nechat IPv4 only
- Klienti mohou využívat IPv6 přes automatické 6to4 tunely
  - Automatická konfigurace (Windows, Unix)
  - Je třeba akorát zprovoznit lokální 6to4 routery
- Řeší problém IPv6 uvnitř VUT, ale ne na venek - budou klienti požadující služby VUT pouze IPv6 protokolem
- Problém správy – klienti (Visty) budou používat tunelované IPv6 bez filtrování, správy, monitorování

2. IPv6 výhradně

- Vše pouze IPv6
  - Komplet nahradit síťovou infrastrukturu
  - Pro komunikaci s IPv4 nutný překlad (NAT-PT, proxy Web, apod.)
  - Nereálné, není k tomu na VUT důvod (máme dostatek IPv4 adres)
  - Zůstaly by nepropojené IPv4 ostrůvky nenahraditelných zařízení a aplikací

3. Tunelování a překlad

1. Pevně konfigurované tunely (propojení IPv6 ostrůvků tunely přes IPv4 síť)
2. RFC3053 Tunnel Broker (www.freenet6.net)
3. RFC3056 6to4 (doporučeno, lze použít pro host i subnet, vyžaduje veřejnou IPv4 adresu, tunel IPv4 protokolem 41, adresa 2002:IP:V4::/48)
4. RFC5214 ISATAP (Microsoft, obdobné, pouze pro host, adresa *prefix*::5EFE:IP:V4)
5. RFC4380 Teredo (tunelování přes NAT v UDP paketech, vyžaduje server)
6. Překlad IPv6/IPv4 (SIIT, NAT-PT, TRT+DNS-ALG)

4. Dual stack

a) Oddělené nezávislé sítě (broadcast domény)
   - samostatné port VLAN nebo protokol VLAN
   - nativní (netagovaná VLAN) IPv4, tagovaná vlan IPv6
   - topologie VLAN IPv4 a IPv6 se nemusí překrývat
   - problematická správa, údržba

b) Paralelní (překryvná) topologie
   - VLAN bude subnetem pro IPv4 i IPv6
   - VLAN bude mít dvě rozhraní routeru – IPv4 a IPv6 (fyzicky nemusí být na stejném HW)
   - Většina uzlů bude mít v dané VLAN obě adresy

Koexistence na Ethernetu

- Uzly napojeny netagovaně – je třeba rozlišit protokol:
  - Ethertype IPv4 = 0800
  - Ethertype IPv6 = 86DD
- Převod IPv6 na MAC = Neighbor Discovery (ND), součást ICMP6 (ne jako IPv4 ARP=0806)
- Zůstává problém multicastů, uzly budou zbytečně zatíženy Ethernet multicasty pro protokol, který nepoužívají (problém MLD):
  - IPv4 multicast = 01-00-5E-dolních 23 bitů IPv4 group
  - IPv6 multicast = 33-33-dolních 32 bitů IPv6 group

Cíle zavedení IPv6 na VUT

- Nedostatek IPv4 adres se projeví hlavně na straně klientů:
  - Pro servery (služby) je na VUT dostatek IPv4 adres
  - Služby budou v dohledné době poskytovány buď oběma protokoly nebo IPv4 only (lokální).
- Je třeba začít poskytovat všechny externě dostupné služby VUT oběma protokoly tak, aby se k nim dostali klienti IPv6 only (externí) → přechod serverů na dual stack.
- **Proto je nutné nejprve vybudovat páteř IPv6.**

Cíle zavedení IPv6 na VUT

- VUT má dostatek IPv4 adres pro klienty, po dohlednou dobu budou klienti provozováni s dual stackem.
- Vypnutí IPv4 neplánujeme, zůstane „trvale“ v provozu.
- Služby používané pouze v rámci VUT mohou zůstat IPv4 only (SAP, Oracle, BMS, dohledové systémy, VOIP, streaming, atd.).

An Internet Transition Plan - RFC5211

- **I. Preparation Phase** – do 2009/12
  - Poskytovatelé by měli začít poskytovat zákazníkům IPv6 konektivitu, nativní nebo tunelovanou.
  - Organizace by měly mít IPv6 konektivitu pro veřejné servery (Web, Email, DNS).
  - Organizace mohou začít poskytovat IPv6 konektivitu uživatelům.

An Internet Transition Plan

- **II. Transition Phase** – 2010/1 až 2011/12
  - Poskytovatelé musí poskytovat IPv6 konektivitu zákazníkům, měla by být nativní,ale může být tunelována.
  - Organizace musí mít IPv6 konektivitu pro veřejné servery. IPv6 služby by měly být poskytovány jako produkční.
  - Organizace by měly poskytovat IPv6 konektivitu uživatelům, včetně interních služeb (DNS, DHCP).

An Internet Transition Plan

- **III. Post-Transition Phase** – 2012/1 a dále
  - Poskytovatelé musí poskytovat IPv6 konektivitu zákazníkům, měla by být nativní.
  - Organizace musí mít IPv6 konektivitu pro veřejné servery. IPv6 služby musí být poskytovány jako produkční.
  - Organizace by měly poskytovat IPv6 konektivitu uživatelům, včetně interních služeb (DNS, DHCP).
  - Poskytovatelé mohou nabízet IPv4 konektivitu. Organizace mohou dále používat IPv4 konektivitu.

Zavedení IPv6 na VUT

- **Fáze I** – 2008 až 2009/12
  - výstavba páteřní sítě (metropolitní i areálové)
  - zprovoznění kritických IPv6 služeb
  - zprovoznění veřejných služeb přes IPv6
- **Fáze II** – 2010/1 až 2011/12
  - produkční stav páteřní sítě IPv6
  - produkční stav serverů a služeb
  - poskytování IPv6 uživatelům VUT
- **Fáze III** – 2012/1
  - produkční stav IPv6 pro uživatele VUT

IPv6 - fáze I (2008-2009)

- **Single arm IPv6 router** ve všech uzlech **(CVIS)**
  - Stačí PC, buď Gb nebo 10Gb (podle areálu, zátěže)
  - Definovat fakultní VLAN, ve kterých bude poskytováno IPv6 (výjimka např. kamery, VOIP, BMS, tiskárny, apod.), přidělit IPv6 prefixy
  - Všechny IPv6 VLAN protáhnout jedním portem do IPv6 routeru
  - Všechny páteřní propojovací IPv6 VLAN taky
  - Protokol OSPFv3 (Unix implementace XORP)
  - V 1. fázi pouze unicast, multicast pouze lokálně

Změny v páteři – fáze I

- Aktivovat podle dostupnosti MLDv2 (IGMP pro IPv6) na L2 prvcích – nekupovat nové prvky, které neumí! **(fakulty)**
- Aktivovat podle dostupnosti IPv6 management prvků
- Podle dostupnosti testovat a případně nasazovat HW L3 IPv6 switche:
  - V první fázi náhrada IPv6 PC routeru
  - V druhé fázi náhrada starého páteřního prvku
- Upravit nástroje pro monitorování stavu a zátěže sítě (ping, NetIS, …)

Změny v síti – fáze I

- Přechod hlavních serverů na dual stack = **upgrade systémů**:
  - FreeBSD – podpora od 4.x, doporučeno přejít na 7.x
  - Linux – záleží na distribuci, aktuální verze většinou bez problémů
  - Windows server – je třeba přejít na Windows Server 2008
  - Solaris – Solaris 8 a vyšší, doporučeno 10
  - **Realizace:** do konce 2009

Změny v síti – fáze I

- Přechod hlavních služeb na dual stack:
  - DNS – **CVIS + fakulty** do konce 2008:
    - zmapování situace a realizovatelnosti
    - upgrade OS a DNS serverů během prázdnin
    - AAAA záznamy pro NS *.vutbr.cz do konce 2008
    - delegace reverzních domén pod ip6.arpa
  - nepoužívat SLAAC autokonfiguraci pro servery (L2 adresa pro vytvoření části L3 adresy – změna síťové karty znamená změnu IPv6 adresy!)
  - DHCPv6 – produkční stav do konce 2009 **(fakulty)**
    - ponecháno na volbě fakulty, lze zůstat u SLAAC autokonfigurace (problém monitorování, správy sítě)
    - změna podle dosavadních zkušeností – viz dále

Upgrade aplikací – fáze I

- Oficiální Weby, E-mail a aplikace pro veřejnost (přihláška) – do konce 2009
- **Doporučení:**
  - **V 1. fázi nezavádět pro veřejné servery (Web, email) duální adresy, ale poskytovat IPv6 služby pod jiným jménem** (www.ipv6.vutbr.cz nebo www6.fit.vutbr.cz)
- E-mail MTA - do konce 2008
  - publikovat AAAA záznam pro MX servery do konce 2008
  - MX servery musí být všechny dual stack, musí mít oba záznamy A, AAAA (viz RFC3974)
  - problém RBL (zatím neexistuje pro src IPv6)

Upgrade aplikací – fáze I

- IMAP, POP do konce 2009
  - UW imapd, Cyrus, Courrier, dovecot – bez problémů
  - Qpopper – beta verze 4.1b12
  - MS Exchange 2007 SP1 + Windows Server 2008
  - klienti Outlook 2007, Windows Mail, Thunderbird (OE ne)
- Web servery a Web aplikace
  - Apache 2.x (Apache 1.3 neoficiální patch)
  - IIS 6.0 lze v nouzi (nedá se moc konfigurovat, neumí FTP, SMTP), doporučen přechod na WS2008/IIS 7
  - Oficiální Weby a aplikace pro veřejnost (přihláška) – do konce 2009

Upgrade aplikací – fáze I

- Zmapování dalších poskytovaných služeb a jejich možného přechodu na dual stack:
  - je třeba identifikovat, co vše vlastně běží a co bude nutno převádět na IPv6
  - zdokumentovat, co nebude převáděno
  - převést, co lze (do konce 2009)
  - Přehled BSD/Linux aplikací a jejich stavu:
    - http://www.deepspace6.net/docs/ipv6_status_page_apps.html
  - Přehled certifikovaných komerčních aplikací a produktů
    - http://www.ipv6ready.org/logo_db/approved_list_p2.php

IPv6 - fáze II (2010-2011)

- **Nasazování páteřních prvků s HW routingem IPv4 + IPv6**
  - OSPFv3 + PIM-SM (DM) + ACL
- Koncové prvky:
  - dual stack management
  - MLDv2
- Na konci fáze mají všechny dual stack VLAN ekvivalentní propustnost pro IPv4/IPv6
- Inzerování dual stack služeb v DNS (www.vutbr.cz)
- Zapnutí dual stack IPv6 na klientech

IPv6 – fáze III

- produkční stav IPv6 sítě a všech aplikací
- funkční IPv6 multicast
- nedostatek IPv4 adres může vést k IPv6 only subnetům
- IPv4 bude nadále provozováno

Fáze III - klienti

- Vista, WS 2008 – default, „bez problémů“
- Windows XP SP2 – `cmd ipv6 install`
- Windows XP nemá DNS resolver over IPv6, nemůže mít tedy nakonfigurovanou IPv6 adresu DNS serveru, buď musí být aktivní IPv4 nebo lze vyřešit instalací ISC BIND9
- Windows XP nemá DHCPv6, lze vyřešit free Dibbler http://klub.com.pl/dhcpv6/ (Made in Gdansk, PL)
- Windows XP nemá MLDv2 (IGMP pro IPv6)

Vista a IPv6

- Nakonfiguruje link-local IPv6 adresu a ověří DAD
- Zkusí Router Solicitation (viz dále)
- Pokud obdrží Router Advertisement
  - a je nastaveno „Managed Address“, zkusí DHCPv6 a nastaví DHCPv6 získanou adresu,
  - nastaví SLAAC adresu,
  - nastaví Temporary IPv6 adresu (celkem 4 IPv6 adresy!)
- Preferuje IPv6 adresy před IPv4 (pokud jsou obě pro daný DNS záznam)
- Aplikace používající P2P Framework vyžadují IPv6 (http://technet.microsoft.com/cs-cz/network/bb545868(en-us).aspx)

Windows XP/Vista/WS

- Pokud není IPv6 nativní spojení (detekován IPv6 router pomocí RS/RA), pak:
  - zkouší ISATAP (hledá DNS isatap.doména), pokud ano, preferuje
  - pokud je veřejná IPv4 adresa, nakonfiguruje 6to4 (**vždy!**)
  - obojí vyžaduje propouštění IP protokolu 41 (firewall!)
  - Teredo ručně (XP) (pouze v privátní síti za NAT!)
    - `netsh interface ipv6 set teredo client teredo.remlab.net`
    - `nesth interface ipv6 show teredo`
  - Vista/WS 2008 (implicitně zapnuto, ale nepoužije se pro DNS, pokud je pouze lokální adresa)
    - `netsh interface teredo set state disabled`

Problém Vista

- IPv6 implicitně aktivní, tunelované přes 6to4
- IPv6 je preferován (pokud je A a AAAA záznam)
- Nelze odinstalovat
- Lze zakázat IPv6 na interface, to ale neřeší tunel 6to4, ISATAP a Teredo
- Jedině přes registry - `HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\tcpip6\Parameters\DisabledComponents`
  - bit 0 = 1 vše, bit 1 = 1 6to4, bit 2 = 1 ISATAP, bit 3 = 1 Teredo, bit 4 = nativní IPv6, bit 5 = 1 preferovat IPv4
- Problém – uživatelům nefunguje spojení, protože nevědomky používají tunel 6to4

Pravidla přidělování IPv6 adres

- Úvod do problematiky
- Motivace
- Algoritmus

IPv6 adresa

- IPv4 adresa (typicky)

0..24	25..31
Prefix	customer
92.46.33	host

  - koncová síť prefix /24 a menší, 0 subnetů, max. 254 uzlů

- IPv6 adresa

0..15	16..31	32..47	48..63	64..127
prefix	RIR	ISP	customer	interface ID
2001	718	802	subnet	host

  - každý subnet 18,446,744,073,709,551,616 adres=uzlů!
  - koncový zákazník má obvykle více subnetů (běžně 65536)
  - 2^{32} sítí jen s prefixem 2001

Klasifikace IPv6 adres

- ::/0 (prefix délky 0) – default route
- ::/128 (samé nuly) – neexistující adresa (RFC4291)
- ::1/128 (1 na konci) – loopback (RFC4291)
- ::0:IP:V4/96 – IPv4 kompatibilní adresa (RFC4291 zruš.)
- ::ffff:IP:V4/96 – IPv4 mapovaná IPv6 adresa (RFC4291)
- 2000::/3 – globální unicast adresy (RFC3513)
- 2002::/16 – 6to4 globální unicast adresy (RFC3056)
- 3ffe::/16 – 6bone (RFC1897 historické)
- FC00::/7 – unique local unicast (RFC4193)
- FE80::/10 – link local unicast (RFC4291)
- FEC0::/10 – site local unicast (RFC3513 zrušeno)
- FFgs::/8 – multicast (g=flags, s=scope, RFC4291)

Struktura globální unicast adresy

48 bitů	16 bitů	64 bitů
globální prefix	subnet ID	interface ID
2001:718:802	*subnet*	*host*

- subnet prefix může být v rozmezí 32-64:
  - RIR dostávali prefix /23 (APNIC 2001:0200::/23, ARIN 2001:0400::/23, RIPE 2001:0600::/23), nyní /12 (2006 – APCNIC 2400::/12, ARIN 2600::/12, LACNIC 2800::/12, RIPE 2A00::/12, AFRINIC 2C00::/12)
  - prefix /32 (původně /35) je přidělován poskytovatelům
  - prefix /48 je přidělován koncovým zákazníkům (každý má 65536 subnetů), nyní omezováno na /56 a /64 (ARIN, APNIC 2008, RIPE?)
  - prefix /64 je pro jeden subnet (každý může mít 18 trilionů = 18 446 744 073 709 551 616 uzlů)

Algoritmus přidělování na VUT

48 bitů	16 bitů	64 bitů
network prefix	subnet ID	interface ID
2001:718:802	subnet	host

- Přidělovat IPv6 subnety primárně podle horní slabiky XX subnet ID (XXYY):
  - podle areálu (topologie)
  - podle fakulty (organizačně)
  - zachovat číselně IPv4 subnet (nejobvyklejší, nejhezčí)
- Obsah dolní slabiky YY subnet ID slouží uživateli pro vytvoření 256 subnetů v rámci přidělené lokality

Motivace

- Prostor 2^{16} je větší než 2^{8}
- IPv6 adresy jsou hodně dlouhé, špatně se pamatují
- Doporučení (RIPE):
  - Wherever possible, address space should be distributed in a hierarchical manner, according to the topology of network infrastructure. This is necessary to permit the aggregation of routing information by ISPs and to limit the expansion of Internet routing tables. Further, RIRs should apply practices that maximise the potential for subsequent allocations to be made contiguous with past allocations currently held.

Příklad FEKT Údolní

- horní část prefix fakulty/lokality = 40
- dolní část podle IPv4 subnetu:
  - 147.229.64-73.0/24 = 2001:718:802:**4040-4047**::/64
  - zbytek prefixu 2001:718:802:**4000**::/56 vyhrazen pro FEKT na Údolní
  - zbytek prefixu 2001:718:802:**4100-4f00**::/56 vyhrazen pro Údolní
  - je dostatek prostoru na růst všech sítí v areálu
  - podle horní slabiky se dá filtrovat (ACL)

Volba interface (host) ID

- Router pro subnet:
  - primární vždy ::1, záložní ::2, atd.
  - pozor ::1-15 jsou jediné adresy pro Embedded RP, nepoužívat na nic jiného než routery!
- Servery:
  - doporučeno plnou IPv4 adresu (::93e5:080C)
  - servery jsou v DNS, budou mít po dohlednou dobu IPv4 adresu, není důvod je utajovat
  - korespondence IPv6 a IPv4 adres je důležitá pro správu
- Klienti – složitější, původně mod. EUI-64

Modifikované EIU-64

- EIU-64 – FireWire, ZigBee MAC
- EIU-48 – Ethernet MAC-48
- jak z MAC-48 vyrobit mod. EUI-64
  - vloží se doprostřed FF:FE (správně mělo být FF:FF)
  - invertuje se bit 6 první slabiky (global/local)
- mod. EUI-64 se použije při autokonfiguraci rozhraní (router dodá prefix, horních 64 bitů)
- Bezpečnostní problém – ať jsem kdekoli, pořád mám dolních 64 bitů IPv6 adresy stejných (a navíc prozrazujících výrobce mého stroje/karty)!

Doporučení

- Nepoužívat automaticky přidělované mod. EUI-64
  - Vista implicitně náhodné – lze zakázat:
    - `netsh interface ipv6 set global randomizeidentifiers=disabled`
  - Obecně by bylo vhodnější použít DHCPv6 než stateless autoconfig
- Důležitý princip - nepřidělovat interface ID sekvenčně (RFC5157):
  - prostor 2^{64} je tak velký, že jej nelze v rozumném čase „propingat“, nelze tím pádem efektivně útočit a nakazit celou síť (ale také ne spravovat!)

Kolik musí uzel mít IPv6 adres?

1. Link-Local pro každý interface (FE80::) – auto EUI
2. unicast/anycast adresy (ruční nebo automaticky)
3. loopback adresu (::1) – jen loopback
4. *all-nodes* multicast (FF01::1, FF02::1) na všech
5. *solicited-nodes* multicast na všech pro každou unicast/anycast 1. a 2. adresu (FF02::1:FFxx:xxxx), kde x reprezentuje dolních 24 bitů interface ID
6. multicast adresy pro všechny skupiny, do kterých patří

(viz ifconfig, ifmcstat – BSD, ipconfig /all - Vista)

Kolik je typicky unicast adres?

- Vista:
  - Temporary IPv6 address (RFC 4941, random)
  - Stateful DHCPv6 address (pokud je povoleno)
  - Stateless autoconfig (random nebo EUI-64)
  - Link-local (random)
- Která se použije?
  - kdo ví?
  - `netsh interface ipv6 show prefixpolicy`
- Co je v reverzním DNS?
  - typicky žádná (DDNS může DHCP nebo SLAC)

Kolik musí mít adres router?

Stejně jako uzel a navíc:

1. *subnet-router* anycast (subnet::0)
2. *all-routers* multicast (FF01::2, FF02::2, FF05::2)
3. *ospf-routers* multicast (FF02::5)
4. *ospf-designated-routers* multicast (FF02::6)
5. *all-pim-routers* multicast (FF02::D)
6. *all-mldv2-capable-routers* multicast (FF02::16)
7. *all-dhcp-relay-agents-and-servers* multicast (FF02::1:2)

Aktuální stav IPv6 na VUT 2008/11

Stav přípravné fáze:

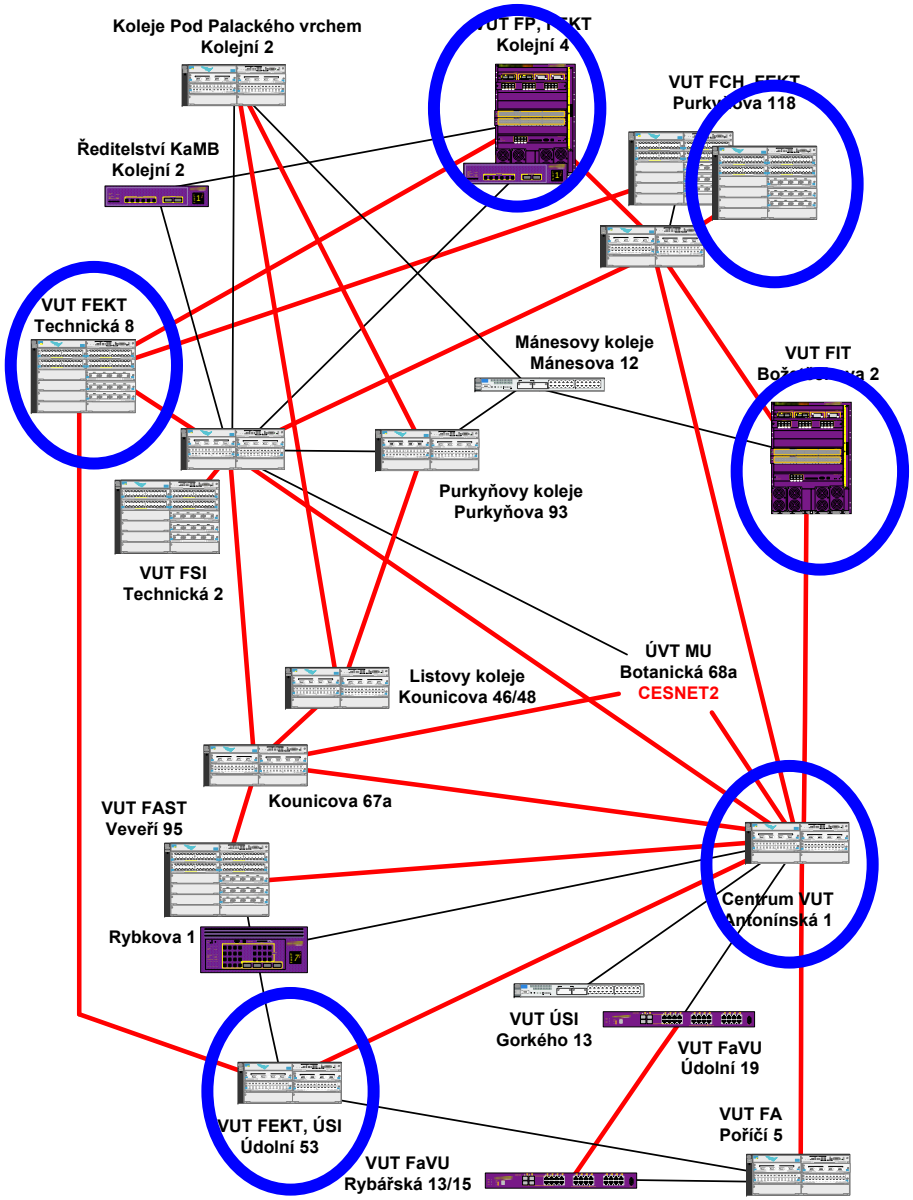
- Pravidla přidělování IPv6 subnetů na VUT
  - Hotovo, několik aktualizací podle nových RFC
  - Zveřejnit
- DNS
  - Spuštěn DNS po IPv6 pro vutbr.cz, fit.vutbr.cz, feec,.vutbr.cz, zaregistrovány a delegovány IPv6 reverzní zóny pro CVIS, FIT, FEKT
  - Registrovány routery a testovací servery, z praktických důvodů přípona 6 (guta6, radka6, apod.)
  - podíl IPv6 DNS dotazů/odpovědi 10% (FIT)

Stav fáze I – Aplikace

- Otestován Apache-1.3 (2.2), v ostrém provozu na www6.fit.vutbr.cz
- UW-IMAP, od verze imap-2004, bez problému
- Eudora Qpopper, dostupná beta 4.1b12, bez problémů
- OpenSSH, jede
- vsftpd-2.0, jede (ftp6.fit.vutbr.cz)
- X Window, jede
- DHCPv6 – ISC DHCP4.1beta, jede
- Squid, podpora IPv6 až ve verzi 3.1 (momentálně je 3.0STABLE10, 3.1RC)

Stav fáze I - Infrastruktura

- Páteřní síť VUT založena na:
  - HP 5406/5412, Extreme Networks BD 8810 a Summit X450
  - HP nemá firmware pro HW IPv6 routing (zatím jen dual stack pro management)
  - Instalované modely Extreme nemají HW IPv6 routing (až od verze a/c)
  - Testován H3C/3COM/Huawei switch, zdá se použitelný
  - Obecně používány Cisco a Juniper

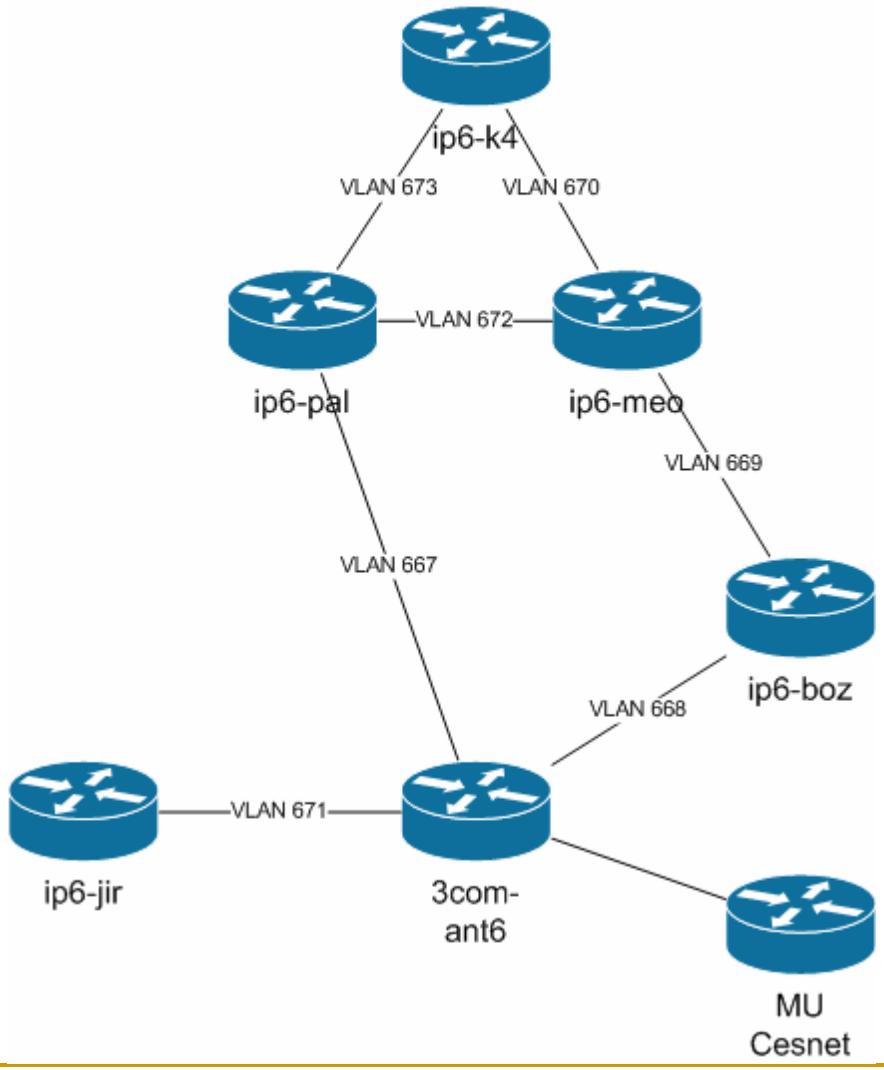

Koleje Pod Palackého vrchem
Kolejní 2
Kolejní 4
Ředitelství KaMB
Kolejní 2
VUT FEKT
Technická 8
Mánesovy koleje
Mánesova 12
Purkyňovy koleje
Purkyňova 93
VUT FSI
Technická 2
ÚVT MU
Botanická 68a
CESNET2
Listovy koleje
Kounicova 46/48
Kounicova 67a
VUT FAST
Veveří 95
Rybkova 1
Centrum VUT
Antonínská 1
VUT ÚSI
Gorkého 13
VUT FaVU
Údolní 19
VUT FEKT, ÚSI
Údolní 53
VUT FaVU
Rybářská 13/15
VUT FA
Poříčí 5

Stav fáze I - infrastruktura

- Testování infrastruktury OSPFv3:
  - Summit XOS se nedomluví s Ciscem (OSPFv3)
  - HP stále nemá firmware IPv6
  - Zebra nepoužitelná, Quagga padá
  - XORP 1.5– zásadní chyba při sestupném číslování Router ID (bug XORP#770, opraveno 13.10)
  - Chyba v IPv6 *recvmsg()* (bug FreeBSD#126349)
  - Ztrácení sítí v Ciscu po refresh LSA
  - Zatím PC routery a XORP-1.6 dedikovanými spoji po samostatných VLAN
  - H3C/3Com switch na CVIS (od 23.1.2009)

Topologie IPv6 na VUT

Tuesday, January 27, 2009



Stav leden 2009

- Dynamický routing se zálohou (OSPFv3)
- Překlopení zálohy pomalejší (protažení VLAN přes další switch neumožňuje detekovat link down na optice)
- Podle dohody ve vybraných VLAN puštěna inzerce routeru přes RA (nebezpečí Vista)
- Jména v DNS s příponou 6

```
traceroute6 to kos6.feec.vutbr.cz
   (2001:718:802:4848::93e5:480a) from
   2001:718:802:8b0::93e5:b00e, 64 hops max
1 ip6-boz  0.274 ms  0.241 ms  0.238 ms
2 routant61.net.vutbr.cz 0.611 ms 0.664 ms 0.492 ms
3 2001:718:802:ffff::2:fff2 0.875 ms 0.994 ms 0.994 ms
4 kos6.feec.vutbr.cz  1.135 ms  1.166 ms  1.240 ms
```

Co dále?

- Příprava doporučení pro filtrování IPv6 (první pokusy jsou)
- Další testování OSPFv3, zatím se zdá funkční
- FIT a FEKT hotovo, otázka co ostatní fakulty
- Zatím zůstaneme na PC platformě, provoz není natolik velký, aby to nestačilo (do cca. 500 Mb/s není problém)

Problémy nasazování IPv6

- Chyby a bezpečnostní problémy ND a RA (CVE-2008-2476) – akceptuje ND od nesousedních uzlů
- Není implementováno RFC5006 (DNS v RA)
- DHCPv6 téměř nepoužitelné (k ničemu)
- V4mapované adresy (BSD vypnuto, problém Linuxu) – obchází tcp wrappery, dns, atd.
- Reverzní DNS nepoužitelné – temporary IPv6 (RFC3041), náhodné SLAAC, dnes většina přístupů 6to4 - problém reverzního DNS pro 2002::/16 (a ISATAP)
- Různé firewally Windows často neumí (nebo špatně umít) IPv6

RD/ND, aneb co vše dělá ICMPv6?

- ICMPv6 =
  - ICMP destination unreachable, packet too big, time exceeded, parameter problem
  - Echo (ping)
  - Neighbor Solicitation/Advertisement (ND6=ARP)
  - Inverse Neighbor Discovery (RARP)
  - Router Solicitation/Advertisement (RD6=RDISC)
  - Router Renumbering
  - Mobile Prefix Solicitation/Advertisement, Home Agent Address Discovery Request/Reply (Mobile V6)
  - Multicast Listener Query/Report/Done (MLD=IGMP)
  - Multicast Router Advertisement/Solicitation/Termination (IGMP MRD)

Neighbor and Router Discovery

- **Jediný** způsob konfigurace gateway a subnet mask (IPv6 prefix) pro klienta (není v DHCPv6)
- Router Discovery – vyhledání všech routerů v dané broadcast doméně = síti (L2)
- Automatická konfigurace adresy (SLAAC):
  - přiřadí link local adresu, detekuje zda není duplicitní
  - vyšle **Router Solicitation** na all-routers group
  - všechny routery odpoví **Router Advertisement** (prefixy, router address, hop limit, MTU, ale ne DNS)
  - klient si vybere prefix a přidá interface-id
  - pomocí neighbor discovery ověří unikátnost adresy (DAD)
  - RFC 5006 (experimental 2007) doplňuje option 25 – adresy DNS serverů (není nikde implementováno)
  - Neřeší VRRP

DHCPv6 není příliš použitelné…

- Nelze zadat adresu routeru a subnet (jen ND6)
- Nelze přiřadit statickou IPv6 adresu podle MAC adresy:
  - pouze podle DUID (DHCP Unique Interface Id)
  - DUID se nedá odvodit z MAC adresy (obsahuje obvykle navíc timestamp nebo něco náhodného)
  - každý klient si generuje a ukládá DUID někde jinde
  - ISC dhcp vypisuje naprosto nepoužitelně: \000\001\000\001\020.\362\"\0000Hg\342x
- Nepočítá se s DHCP klient v4 a v6 současně (každý přepisuje /etc/resolv.conf)
- V existujících Unixech volá ISC *dhclient* špatný (existující) skript (takže nefunguje nastavení adresy)
- Nalezen drobný bug v *dhcrelay* (nefunguje pro více než 1 if, bug potvrzen a opraven v CVS 4.1.0rc1)

Vista DHCPv6, k čemu tam vlastně je?

- Vista/WS2008 SP1 chyba generování a parsování DNS domain search list (viz KB953268, hotfix na vyžádání)
- DHCPv6 klient se použije pouze za podmínky, že v RA flags je nastaven bit 7=M (Managed Address) nebo pro parametry pokud je bit 6=O (Other Stateful Config)

  Příklad: /etc/rtadvd.conf pro rtadvd

  ```
  vlan2:      :raflags#128:
  ```
- DHCP adresa se pak nastaví, ale Visty ji stejně nepoužijí - preferují temporary IPv6 adresu dle RFC4941 (platná je 7 dní, pak se generuje jiná), lze ji zakázat:

  ```
  netsh interface ipv6 set privacy state=disabled
  ```
- Ale ani pak nepoužijou při odchozím spojení DHCP adresu (použijí stateless adresu)!
- Při přechodu do jiného subnetu vrátí DHCP server status „not on-link“ a Visty nepřidělí nic (zatím nevyřešeno).

IETF August V6OPS:

- **Rogue router RA implications** - draft-chown-v6ops-rogue-ra-01
  - Manuálně konfigurovat default router
  - Použít SEND - SEcure Neighbor Discovery (RFC 3971), není implementováno, problémy RSA krypto, timestamp, atd.
  - RA (ND6) snooping v přepínačích (jako DHCP snooping) – není zatím implementováno
  - Použít RFC4191 router preference option – pouze volitelné, útočník ostatně může použít taky
  - Použít na koncových uzlech paketové filtry – nereálné
  - Monitorování a dynamické léčení – rafixd, ndpmon
  - Omezit dynamickou reakci, aplikovat RA po 2 hodinách
  - Doplnit do DHCPv6 default router – ale dhcp je taky nebezpečné
  - Konfigurovat ACL na přepínačích – filtrovat RA v opačném směru (nereálné)

Jak řešit

- Problém – Visty s nainstalovaným ICS po připojení vytvoří 6to4 tunel a začnou v síti inzerovat, že routují IPv6 (vysílají RA), navíc ten tunel většinou nejede
  - Okolní Visty si nastaví nativní IPv6 adresu a posílají vše přes tento nefunkční router…
- Pokud je v síti nativní IPv6, lze mu v RA nastavit vyšší prioritu – to samozřejmě nezabrání cílenému útoku
- Je nutné monitorovat RA (viz ndpmon), je na to ale nutný slušný management pro odhalování pachatelů (monitorovat všechny ND, MAC, IPv6, loginy, atd.)

IPv4-Mapped address vně

- Pro jednodušší převod aplikací na IPv6 bylo zavedeno, že IPv6 socket přijímá IPv4 spojení. Pro rozpoznání byla zavedena speciální adresa **::FFFF:IPv4-adresa**.
- To vedlo k několika bezpečnostním dírám:
  - IPv6 paket s cílovou adresou ::FFFF:127.0.0.1 může v rámci lokální sítě kdokoli podvrhnout a příjemce ho chápe jako lokální paket (localhost)
  - Paket obchází pravidla v paketovém filtru
- Doporučení RFC4942 – vně systému se nesmí vyskytnout (odesílat, ani přijímat)

IPv4-Mapped address uvnitř

- Problém implementační:
  - IPv6 socket posluchá na stejném portu pro AF_INET i AF_INET6, co když je jeden z nich obsazený?
  - co se stane po *setsockopt()* s IPv4 portem, když jsou nastavovány IPv6 options a naopak?
  - Co když jsou přijímány na portu multicasty?
  - Co s /etc/hosts a DNS PTR?
- Problém bezpečnostní:
  - IPv6 only služba (uzel) předpokládá, že se na něj nikdo přes IPv4 nedostane (a ejhle).
  - Pravidla v TCP wrapperu se přestanou pro dané IPv4 adresy uplatňovat.

IPv4-Mapped address

- Radikální řešení
  - OpenBSD – neimplementovat, nepřijímat (problém POSIX)
  - Windows XP/2003 – různé stacky, není implementováno
  - Dual stack aplikace musí vždy otevřít dva sockety (AF_INET a AF_INET6)
- Pragmatické řešení
  - RFC3493 socket option IPV6_V6ONLY (2003)
  - systémově implicitně nastavit na ON (FreeBSD, NetBSD sysctl net.inet6.ip6.v6only=1), Linux bohužel OFF
  - Problém – přestože to je známo už od roku 2002, pořád se vyskytují aplikace s implicitně zapnuto (zejména v Linuxu)

Nové nástroje

- ARP -> NDP
  - BSD: ndp –a
  - XP: ipv6 nc
  - Vista: netsh interface ipv6 show neighbor
- IPv6 adresa:
  - BSD: ifconfig *if* inet6 *addr* [prefixlen *len*]
    - /etc/rc.conf ipv6_enable="YES"
    - ipv6_ifconfig_*if*="*addr*"
  - XP: ipv6 adu *if I addr*
  - Vista: netsh interface ipv6 add address *if addr*

Nové nástroje

- default router (pokud je zakázáno RA):
  - BSD: route add –inet6 default *gwaddr*
    - /etc/rc.conf: ipv6_defaultrouter="*gwaddr*"
  - XP: ipv6 rtu ::/0 *if I gwaddr*
  - Vista: netsh interface ipv6 add route ::/0 *if gwaddr*
- Jak zakázat RA:
  - BSD: pokud je statická adresa, zakázáno
    - sysctl –w net.inet6.ip6.accept_rtadv=0
  - XP: ?
  - Vista: netsh interface ipv6 set interface *if* routerdiscovery=disabled [store]

Nové nástroje

- Jak přesvědčit Windows k DHCPv6, když jsme zakázali RA:
  - XP: nejde (nemá DHCPv6 klienta)
  - Vista: netsh interface ipv6 set interface *if* managedaddress=enabled
  - Kontrola: netsh interface ipv6 show interface *if*
- Funkce DHCPv6 klienta je jinak řízená flagem v RA, Visty sice DHCP adresu přidělí, ale současně taky temporary a SLAAC konfigurovanou (a používají temporary).

Nové nástroje

- Směrovací tabulka:
  - BSD: netstat -rn -f inet6
  - XP: ipv6 rt
  - Vista: netsh int ipv6 show route
- Ping:
  - BSD: ping6 ipv6.google.com
  - XP: ping6 ipv6.google.com
  - Vista: ping -6 ipv6.google.com

Nové nástroje

- Traceroute:
  - BSD: traceroute6 ipv6.google.com
  - XP: tracert6 ipv6.google.com
  - Vista: tracert -6 ipv6.google.com
- DNS:
  - BSD: host ipv6.google.com
  - XP: nslookup
set q=aaaa
ipv6.google.com
  - Vista: nslookup ipv6.google.com

Nové nástroje

- Telnet, FTP
  - BSD: telnet -6, ftp -6
  - XP: telnet, ftp
  - Vista: telnet, ftp
- Multicast skupiny:
  - BSD: ifmcstat
  - XP: neumí MLD
  - Vista: netsh interface ipv6 show neighbor

Závěr

- Doporučeno začít s IPv6, zkusit si chování klientů, serverů, začít se učit novým postupům
- Uživatelům IPv6 nenutit, ale nechat je to používat (nevědomky, viz Visty)
- Zprovoznit Web servery na IPv6, zkusit si chování, případně i zadat AAAA adresu pro normální jméno
- Používat lokálně FTP, SSH, IMAP, POP přes IPv6, pro seznámení