smekal.at :: IT Consulting

"IPv6 ist da. Was nun?"

Das neue Internet Protokoll - Grundlagen, Potenziale, Migration

smekal.at :: Goesta Smekal

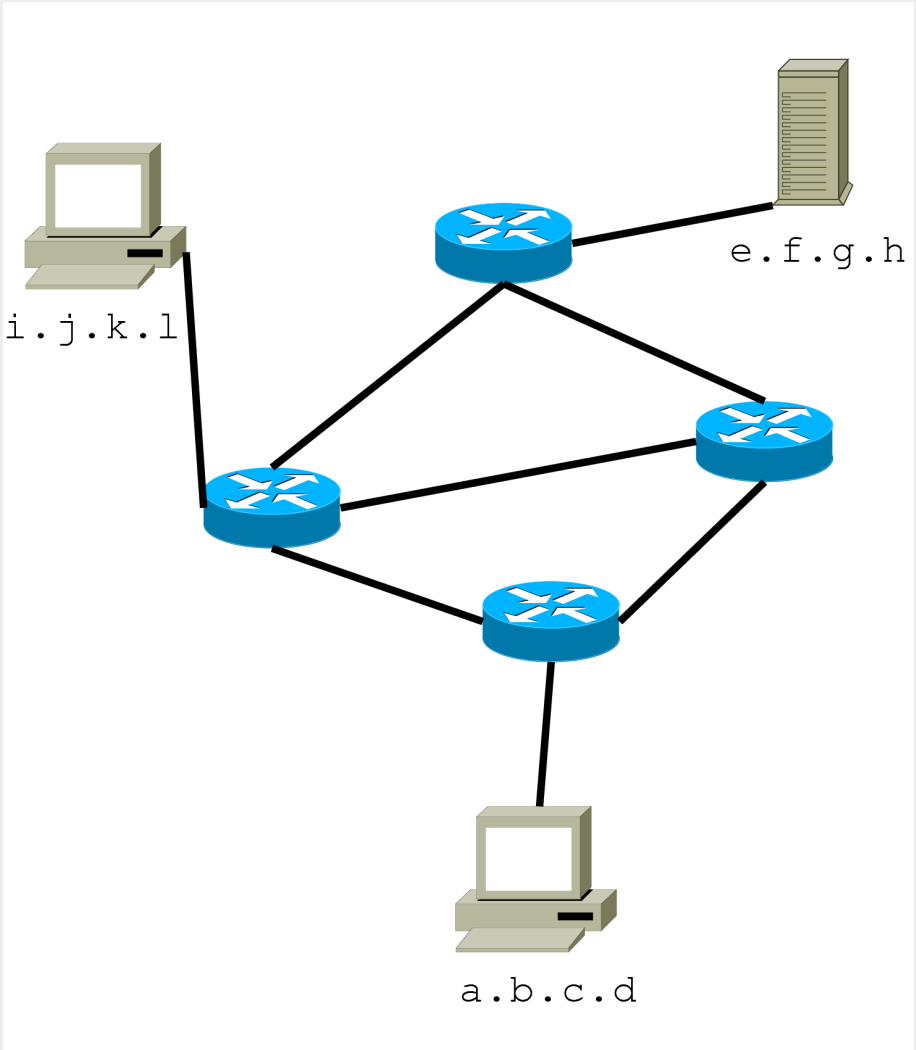
- 15 Jahre als IT Professional
- unterschiedliche Umgebungen:
  - 24x7 Operating im medizinischen Bereich
  - Support, Training, Implementierung
  - IT Leiter bei einer NPO
  - Senior Consultant Systemmanagement
- 15 Jahre Open Source Erfahrung
- enger Kontakt zur Community
  (Vorstandsmitglied der Linux User Group Austria)

Erfahrungen kann man nicht lernen, man muss sie machen

IP = Endpoint to Endpoint

1980:

- IPv4 verbindet als erstes Netzwerkprotokoll Hosts verschiedener Netze, unabhängig vom Übertragungsmedium
- einheitliches Adressierungsschema
- zwei Nodes können direkt miteinander kommunizieren



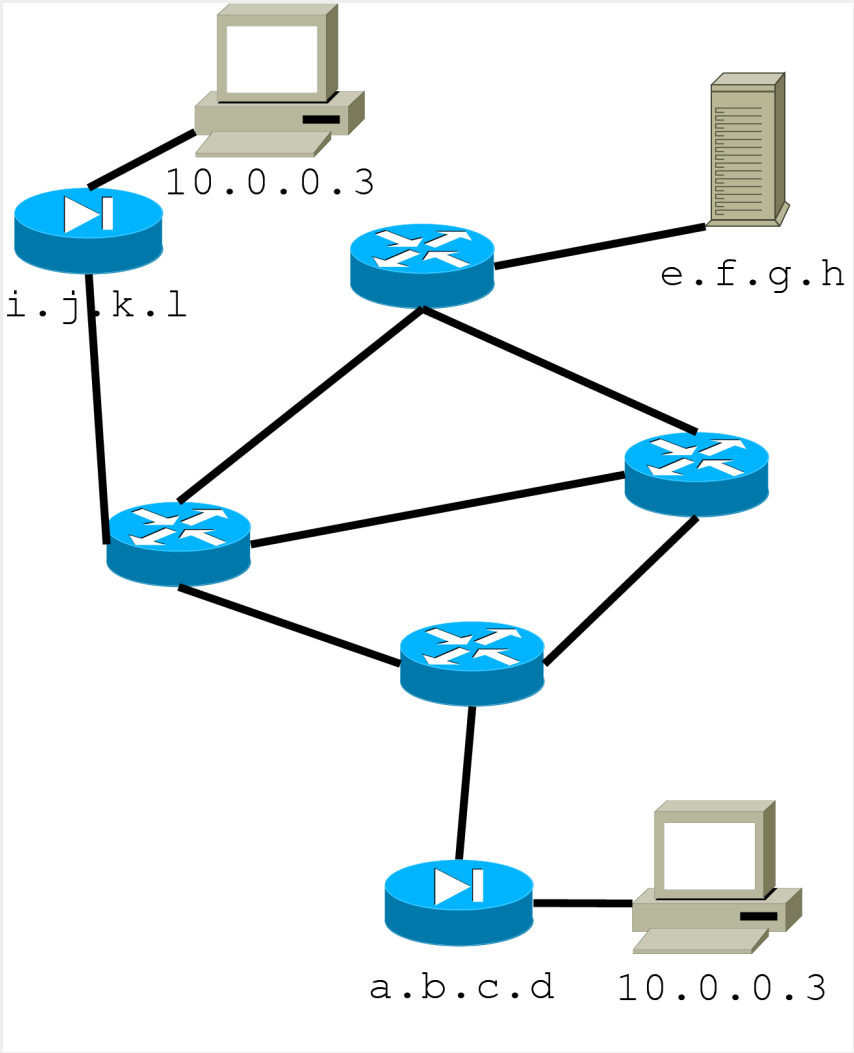
IP = Endpoint to Endpoint

derzeit:

direkter Verbindungsaufbau wird verhindert durch:

- Firewalls
- NAT

E2E nur über Vermittlung durch Dritte



IPv4 Adressen reichen bis ...

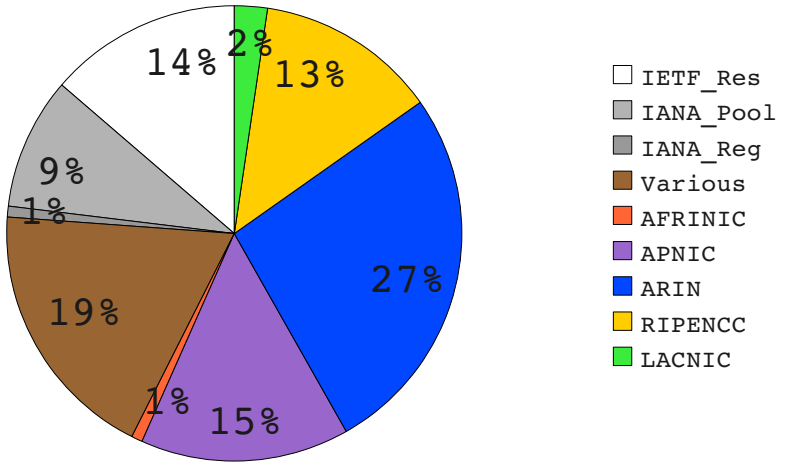
- "Nur" 4,29x10^9 IPv4 Adressen (32 Bit)
- anfangs nur "ganze Blöcke" (Class A,B,C) vergeben, teils leichtfertig
- rasantes Wachstum bei Mobilfunk, Homeautomation ...
- Verteilung sehr einseitig (Folie folgt)

Das Ende naht ... bei gleich bleibender Vergabe:

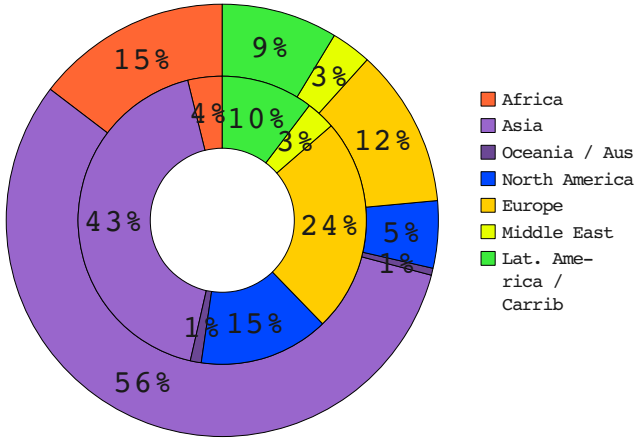
8.Sep.2011 (Stand: Jänner 2010)

3. Februar 2011

Internetnutzung vs. IPv4 Adressen



World Regions	Pop / Mio	Inet Users / Mio	Penetration (% Population)	Users % of Table
Africa	991,00	**67,37**	6.8 %	3.9 %
Asia	3808,07	**738,26**	19.4 %	42.6 %
Europe	803,85	**418,03**	52.0 %	24.1 %
Middle East	202,69	**57,43**	28.3 %	3.3 %
North America	340,83	**252,91**	74.2 %	14.6 %
Latin America/Caribbean	586,66	**179,03**	30.5 %	10.3 %
Oceania / Australia	34,70	**20,97**	60.4 %	1.2 %
WORLD TOTAL	6767,81	**1733,99**	**25.6 %**	100.0 %



Anatomie einer IPv6 Adresse 2001:0db8:1104:6100::1/64

- 128 Bit -> 16 Byte -> 8 Blöcke hexadezimal
- führende Nullen können entfallen
- ein zusammenhängender Block von Nullen kann durch "::" ersetzt werden
- Extremfall: 127.0.0.1(IPv4) -> ::1(IPv6)

2001:0db8:1104:6100:0000:0000:0000:0001

2001:db8:1104:6100::1 ✓

2001:0db8:1104:0000:0000:4bd3:0000:0ac3

~~2001:db8:1104::4bd3::ac3~~ ✘

2001:db8:1104::4bd3:0000:ac3 ✓

IPv6 Präfixe

analog zu CIDR in IPv4 werden IPv6 Netze durch Präfixe definiert:

- `fe80::230:5ff:febe:3084/64`
  `fe80:0000:0000:0000:0230:05ff:febe:3084`
- `2001:858:0002::/48`
  `2001:0858:0002:0000:0000:0000:0000:0000`

Je kleiner das Präfix, desto größer das Netz

Typische End User Präfixe:

- /64 -> ein Netz mit 2^64 Stationen
- /56 -> 256 Netze mit je 2^64 Stationen

IPv6 Adresstypen

- link local (`fe80::/10`)
  gilt innerhalb der Broadcast-Domain, wird nicht geroutet
- unique local IPv6 unicast (`fd00::/7`) (site local)
  wird nicht geroutet, trotzdem global eindeutig (40Bit random ID, 16Bit Subnet, 64Bit Interface) (vgl. RFC 1918, NAT+VPN)
- global unicast (`2000::/3`)
  `2001:0000::/32` Teredo (Tunnel Broker)
  `2001:db8::/32` Dokumentation
  `2002::/16` `6to4`
- multicast (`ff00::/8`)
  `ff02::1` all Nodes (link local)
  `ff02::2` all Routers
- anycast
  Paket wird dem nächsten Mitglied zugestellt. z.B: subnet-router anycast:
  `2001:5c0:1104:6100::` -> Subnet komplett, ID leer

Besondere IPv6 Adressen

- Multicastgruppen nach Scope:
  `ff01::x` am selben Interface
  `ff02::x` im selben Subnetz
  `ff05::x` site local
  `ff0e::x` global
- andere Multicastgruppen:
  all Nodes: `ff0x::1`
  all Routers: `ff0x::2`
  all NTP Servers: `ff0x::101`
- IPv4 in IPv6
  `::ffff:192.168.0.21`

Umgang mit IPv6 Adressen

Wie gibt man IPv6 Adressen ein?

- gar nicht - man nutzt DNS
- `[fe80::210:a4ff:fe88:ecd]:80`
  (z.B. im Browser, Achtung: Shell Escapes! `'[...]'`)

Suche mit Regular Expressions (z.B. Perl):

- `m/\d{1,3}\.\d{1,3}\.\d{1,3}\.\d{1,3}/`
  klappt nicht.
- `NetAddr::IP->version($IP)`
  liefert die Version der Adresse: 4,6 oder undef

Der IPv6 Header

- mehr Platz für Quell und Zieladresse
- feste Länge: 40B (statt 20-60 Byte)
- weniger Felder:
  - Version(4b)
  - Traffic Class(1B)
  - Flow Label (20b)
  - Payload Length (2B)
  - Next Header (1B)
  - Hop Limit (1B)
  - Source (16B)
  - Destination (16B)

- aus IPv4 nicht übernommen:
  - ~~Header Length~~
    (Länge ist fix)
  - ~~Identification~~
  - ~~Flags~~
  - ~~Fragment Offset~~
    IPv6 fragmentiert nicht
  - ~~Header Checksum~~
    findet am Media-Access-Layer und am Application-Layer statt

schnellere Bearbeitung durch Router

IPv6 Extension Header

Weitere Optionen werden bei Bedarf definiert

- Hop-by-Hop Options (QoS, Multicast, Jumbogram)
- Routing Header
- Fragment Header
- Destination Options Header
- Authentication Header
- Encryption Security Payload Header
- ...

ICMPv6

mehr als nur Ping ...

- informational Messages
  Echo Request, Echo Reply
- Error Messages
  Dest. unreachable, Packet too big, Time Exceeded, Parameter Problem
- Neighbor Discovery (ND)
- Autoconfiguration
- Path MTU Discovery
- Multicast Listener/Server Discovery (IPv4: IGMP)

ICMPv6: ND

- automatische Konfiguration von Adressen
  Neighbor Solicitation, Nieghbor Advertisment
  (Link local und global Unicast)
- Prefix und Router finden
  Router Solicitation, Router Advertisment
- Duplicate Address Detection (DAD)
- Link-Layer Address Resolution (IPv4: ARP)
- Secure Neighbor Discovery
  Router verwenden Zertifikate, Certification Path, Cryptographically generated Addresses

ICMPv6: Autoconfiguration, Path MTU Discovery

Ohne DHCP: "stateless Autoconfiguration"

- link local Addr. erzeugen (fe80::/10) "tentative"
- Multicast Gruppe ff02::1 (all Nodes) und "solicited-node" beitreten
- Nieghbor Solicitation mit Tentative-Addr als Dest. (DAD)
- Router Solicitaion (ff02::2)
- Präfixe und Routen eintragen

MTU Discovery:

- Framentierung findet am Quellhost statt
- MTU kleiner als Paket -> ICMP Packet too big + MTU
- Quelle verkleinert
- ab und zu größere Pakete -> bessere Route erkennen
- MTU mindestens 1280B: höhere Effizienz wegen geringerem Overhead

IPv6 und DNS

- neuer Record: AAAA
- kann auch via IPv4 abgerufen werden (bind seit v8.4)
- PTR:
  `3.e.d.3.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.b.0.0.0.0.0.4.1.0.c.5.0.1.0.0.2.ip6.arpa.`
- derzeit AAAA Records für acht Rootserver (+1 seit Jänner 2010)

Legacy Protokolle

mögliche Probleme:

- IP Adressen in der Payload
- IP Adressen Teil der Prüfsummen (IPSec)
- IP Adresse in ACLs

Beispiele:

- FTP
  - 32 Bit für IP in "Port" Kommando -> neues Kommando EPRT
  - EPSV für Passive Mode
- Telnet
  - funktioniert wie gehabt ... leider

Quality of Service

a) Integrated Services
   QoS Vereinbarung wird End to End getroffen, alle Router speichern QoS Anforderungen pro Flow (RSVP, RFC 2205) skaliert schlecht

b) Differentiated Services
   Flow wird innerhalb von DS-Domains (Menge aller Router innerhalb DS-Boundary Routers) einheitlich behandelt. PHB (Per Hop Behavior) Routing Decision

IPSec

- IPv4 IPSec ist ein Backport
- AH und ESP sind fester Bestandteil der Extension Header
- Tunnel Mode praktisch obsolet -> IPSec End to End
- in (fast) alle Protokolle höherer Ebenen integrierbar
- keine v4 <-> v6 Verbindung möglich

mobile IPv6 - Grundlagen

- Nodes bleiben in "fremden" Netzen unter bekannter IP erreichbar
- nutzt IPSec
- erspart VPN Tunnel
- lehrt Firewall Herstellern das Fürchten

Begriffe:

MN: Mobile Node
Rechner auf Wanderschaft

HA: Home Agent
führt Buch über MNs

CN: Correspondent Node
will MN kontaktieren

mobile IPv6 - Ablauf

1. MN sucht HA (Home Agent) unter link local
2. Meldung an HA mit neuem Präfix via global unicast
3. HA trägt neues Präfix in Liste ein
4. CN fragt nach MN
5. HA antwortet auf ND, sendet Anfrage an MN
6. beteiligte Router leiten Verkehr direkt zu MN (Route optimization), Präfix wird gekapselt -> transparent für CN und MN

IPv6 - Security

- durch global unicast neue Securitykonzepte:
  - Firewall ist keine einzelne Box
  - Security Policy zentral verwaltet, auf Clients ausgebracht
  - Isolation VLANs für unbekannte Rechner
- Endpoint to Endpoint IPSec - Firewall verliert Kontrolle über Inhalt
  - schlechte Nachrichten für Checkpoint, Phion & Co
  - gute Nachrichten für Dienstleister (s.o.)
- Intranet: Unique local unicast statt RFC-1918

IPv6 - Privacy

- Link local und global unicast enthalten MAC Adresse
  - kann zu Client-Tracking benutzt werden
    - Nutzer identifizierbar, auch in "fremden" Netzen (Notebook)
    - NIC identifizierbar -> Extremfall Gerät identifizierbar (Vorteil bei HW Support?)
  - kann abgestellt werden
    - DHCP mit kurzer Leasetime
    - pseudo zufälliger Interface-Identifier

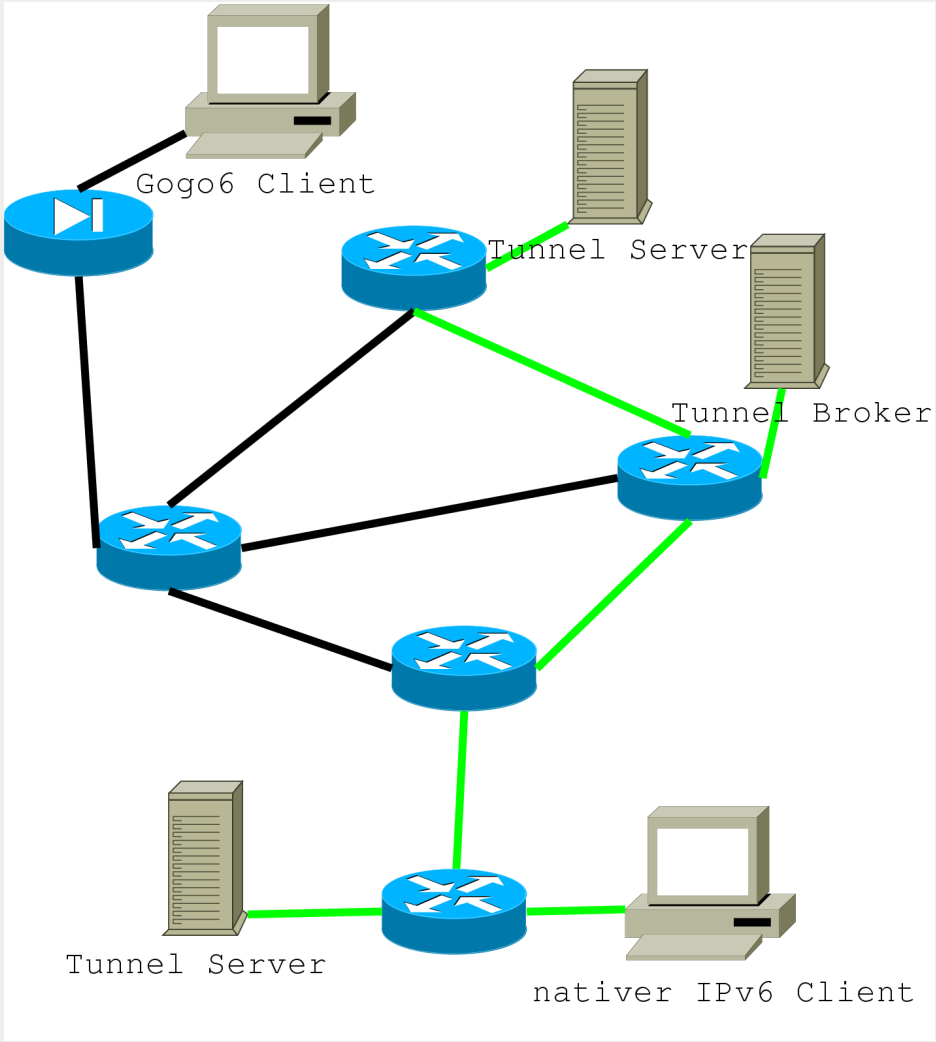
IPv4 - IPv6 Transition

- Dual Stack
  - Node spricht beide Protokolle (vor allem für öffentliche Server wichtig)
  - Proxy setzt Anfragen um
- Tunneling
  - 6to4 (erfordert offizielle IPv4 Addr)
  - Teredo (durch NAT möglich, zu Teredo Server)
  - Tunnel Broker (s.o.)
  - Silkroad
  - ISATAP
  - 6RD (RFC 5569, freies Präfix, von IANA zugeteilt)

und los!

Beispiel Tunnel Broker:

- Client -> T-Broker
- Broker verweist aus nächten T-Server
- Client tunnelt in UDPv4 zum T-Server
- ab jetzt IPv6



Quellen

- Liste der registrierten Netzblöcke:
  ftp://ftp.ripe.net/pub/stats/ripencc/membership/alloclist.txt
- allgemeine Statistiken zum Internet:
  http://www.internetworldstats.com/stats.htm
- Warum es bald keine IPv4 Adressen geben wird:
  http://www.potaroo.net/tools/ipv4/index.html
- diverse Hintergrundinfo
  http://www.tcpipguide.com