

Tricks und Fallstricke bei Prefix-Delegation in DHCPv6

Dr. Dietrich Schroff (GIP AG)



GIP AG :: Thinking Beyond Computing

Real-time OSS Lösungen basierend auf der Service-Automatisierungsplattform **Xyna Factory**

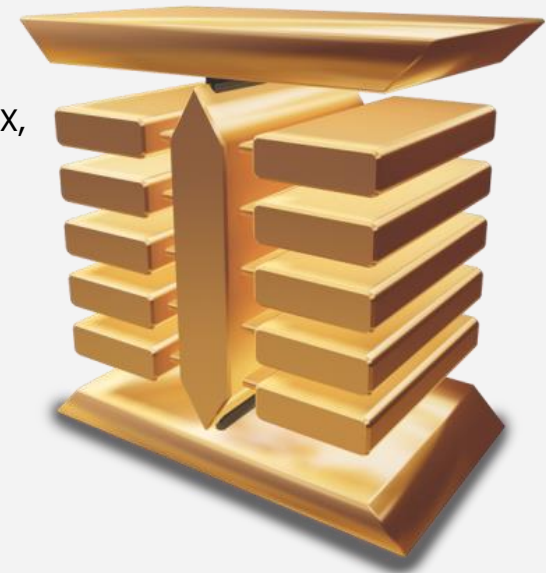
GIP

Gesellschaft für Industriephysik

- Gründung 1992, heute > 50 Mitarbeiter
- Enge Verbindungen zu regionalen Hochschulen, u.a. Universität Mainz, FH Bingen, TU Kaiserslautern
- Über 15 Jahre Erfahrung mit komplexen Kommunikationsnetzen auf internationaler, nationaler und regionaler Basis

Xyna Factory Lösungen

- VPN Services (IP VPN, L2-VPN)
- Managed Services (IP-Centrex, IP-PBX, UC / UCC, etc.)
- Triple-play / 4-play Services für Privatkunden
- Service Provisioning
- Subscriber Self Service
- Visual Service Engineering
- Auto-Configuration Server (ACS)
- Order management
- Assurance & Monitoring
- Device Management
- Network Activation



Erfolgreiche Xyna-Lösungen bei führenden Tier-1 und Tier-2 Providern

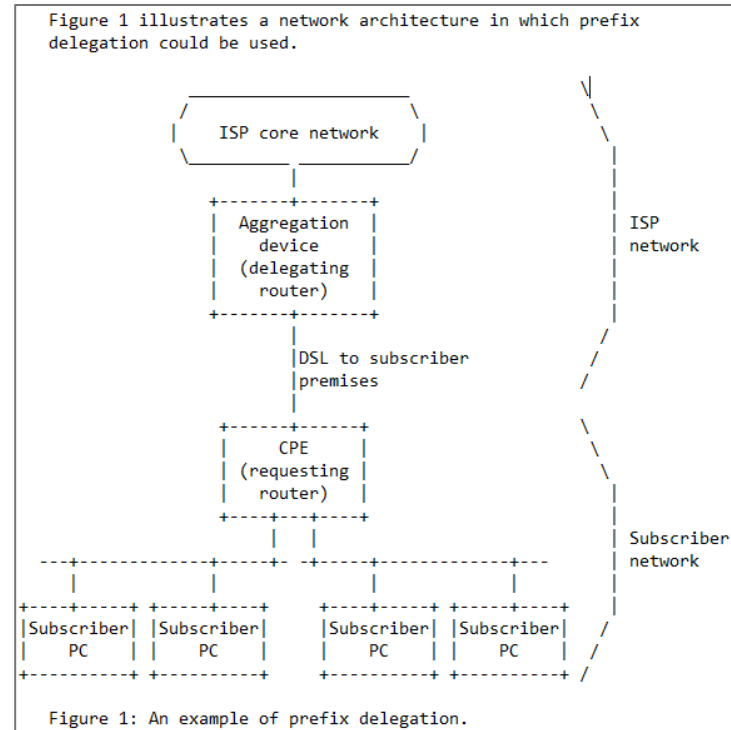
- Einführung in Prefix-Delegation
 - RFCs & Parameter
- Prefix-Delegation bei Privatkunden
 - Typische Einsatzszenarien
- Renumbering
 - Vorgaben & RFCs
 - Vorgehensweise bei Privatkunden
 - Zusammenfassung: Privatkunden
 - Unterschiede bei Geschäftskunden
 - Mehrstufige Hierarchien
 - Herausforderungen
 - Infrastruktur
 - Offene Punkte
 - Ist-Zustand
- Zusammenfassung

RFC 3633 „IPv6 Prefix Options for DHCP version 6”

- It is appropriate for situations in which the delegating router does not have knowledge about the topology of the networks to which the requesting router is attached, and the delegating router does not require other information aside from the identity of the requesting router to choose a prefix for delegation. For example, these options **would be used by a service provider to assign a prefix to a Customer Premise Equipment (CPE) device acting as a router between the subscriber's internal network and the service provider's core network.**

RFC 3769 „Requirements for IPv6 Prefix Delegation“

- The prefix delegation mechanism should allow for delegation of prefixes of lengths between /48 and /64, inclusively. Other lengths should also be supported. The mechanism should allow for delegation of more than one prefix to the customer.

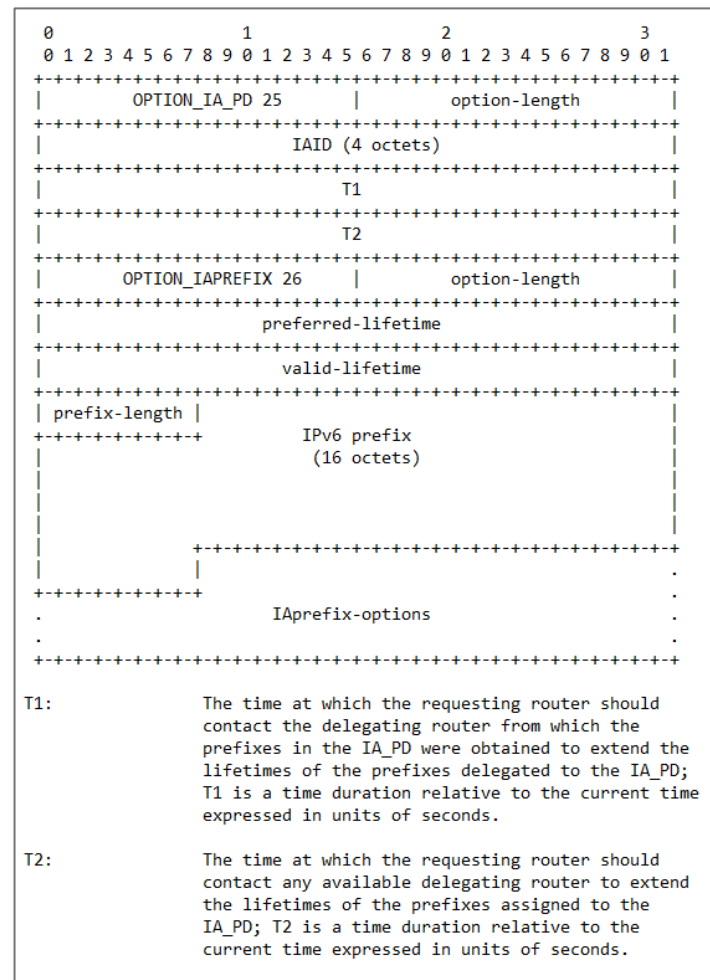


Allgemeine Werte

- IPv6 und Prefix-Länge
- Preferred-lifetime (in Sekunden)
 - the length of time that a valid address is preferred (i.e., the time until deprecation). When the preferred lifetime expires, the address becomes deprecated. A deprecated address should no longer be used as a source address in new communications, but packets sent from or to deprecated addresses are delivered as expected.
- Valid-lifetime (in Sekunden)
 - When the valid lifetime expires, the address becomes invalid.

DHCPv6 spezifische Werte

- T1 (Renew-Intervall)
 - Vorschlag in RFC 3633: 50% von Preferred-lifetime
- T2 (Rebind-Intervall)
 - Vorschlag in RFC 3633: 80% von Preferred-lifetime



Unitymedia KabelBW

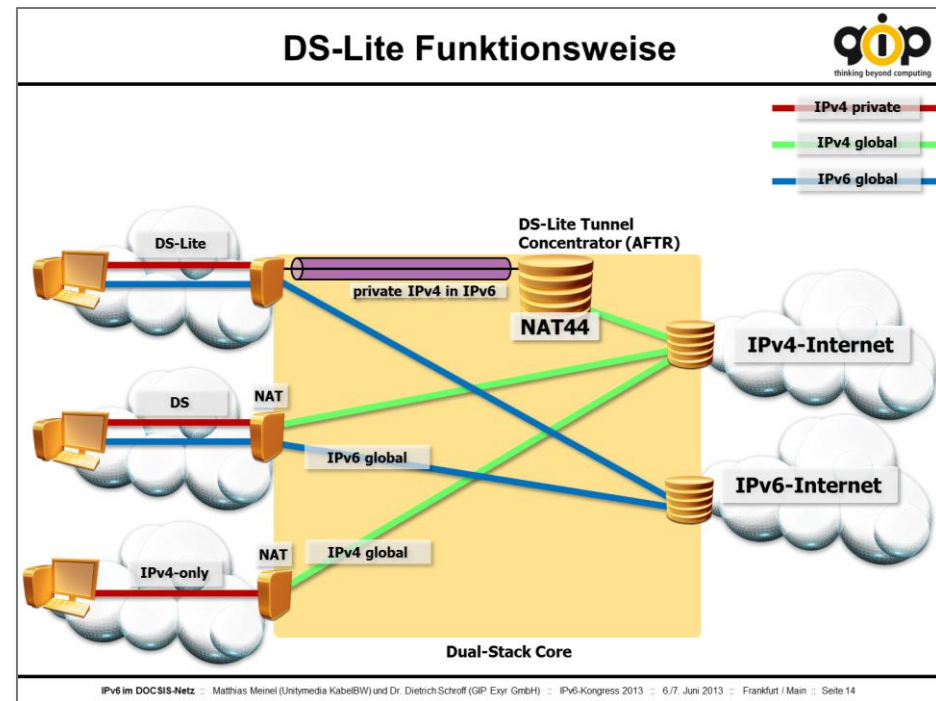
- DHCPv6-PD: /56 von Provider
 - /64 per SLAAC
- CPE meist AVM FritzBox 6360
- Anbindungsvarianten
 - Dual Stack
 - DS-Lite

Deutsche Telekom

- IPv6CP in PPP mit /64
- DHCPv6-PD mit /56
- Anbindungsvarianten
 - Dual Stack

Heimnetz

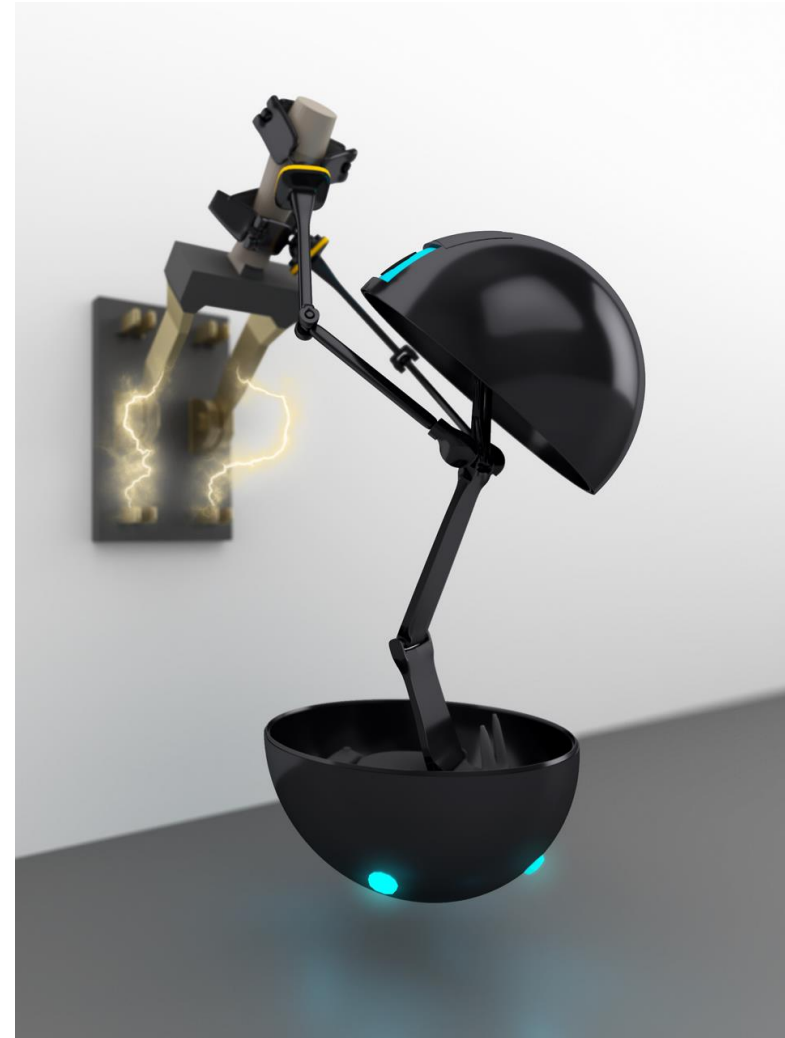
- Vergabe der IPv6-Adressen mit SLAAC (Stateless address autoconfiguration / RFC 4862)



RFC 4192 „Procedures for Renumbering an IPv6 Network without a Flag Day“

- 8 Phasen:
 1. Stable using the old prefix
 2. Preparation for the renumbering process
 3. Configuring network elements for the new prefix
 4. Adding new host addresses
 5. Stable use of either prefix
 6. Transition from use of the old prefix to the new prefix
 7. Removing the old prefix
 8. Stable using the new prefix
- inklusive Betrachtung von DNS
- keine Betrachtung von Zeiten!

→ Eher Grobspezifikation



Normalbetrieb

- Prefix wird im Zustand *preferred* gehalten

Wechsel

- Altes Prefix wird in den Zustand *deprecated* gebracht
- Neues Prefix wird im Zustand *preferred* gehalten

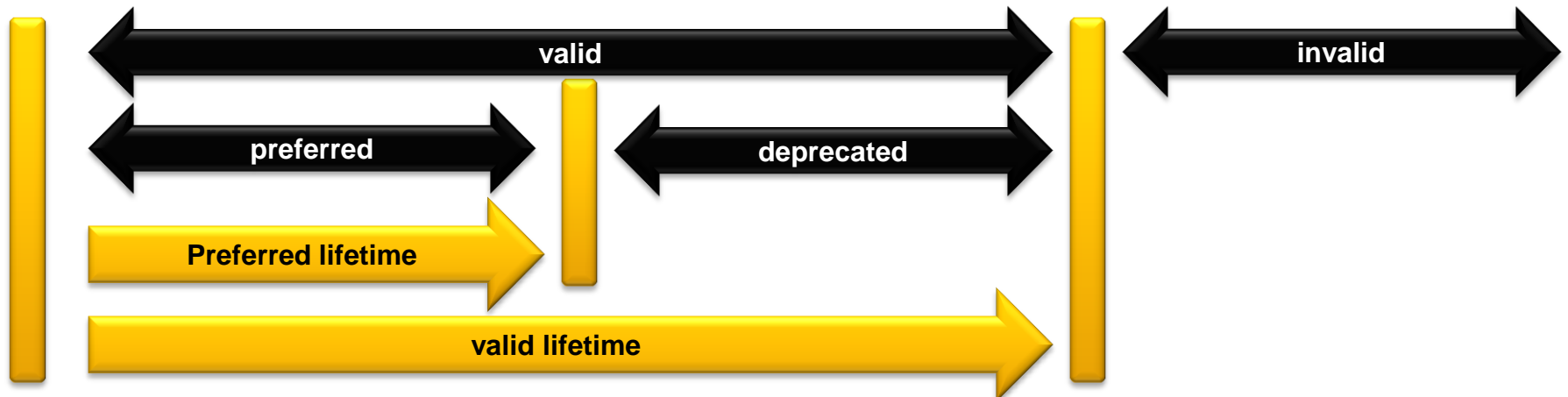


CPE

- 2 Prefixes im Zustand *valid*
- Zu Beginn *preferred/preferred* → *preferred/deprecated*

Devices

- 2 IPv6-Adressen im Zustand *valid*
- Zu Beginn *preferred/preferred* → *preferred/deprecated*
- Damit Kommunikation zwischen Devices jederzeit möglich
- *Deprecated*-Zeitraum >> typischer Kommunikationszeitraum

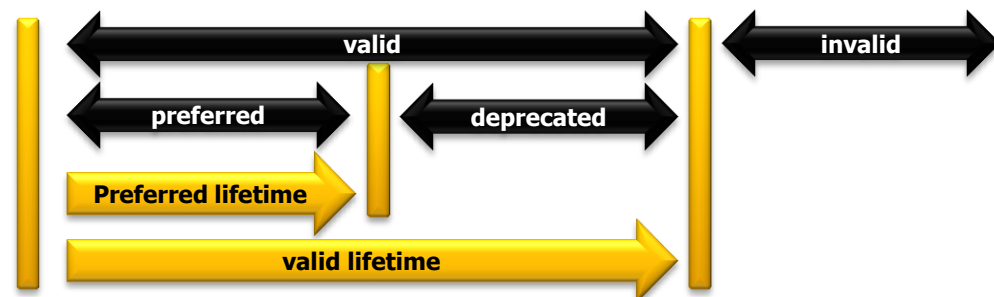


Betrachtung der Zeiten

- Vor dem Wechsel (Prefix valid und preferred)
 - `2001:db8:cafe:cafe::/64` [Valid 86400, Preferred 82800]
 - Renew-Intervall/RA-Intervall: 3600s
- Start des Wechsels (alle Prefixes valid, gemischter Modus preferred/deprecated)
 - `2001:db8:cafe:cafe::/64` [Valid 86400, Preferred 0]
 - `2001:db8:cafe:affe::/64` [Valid 86400, Preferred 82800]
 - Renew-Intervall/RA-Intervall: 3600s
 - Nach einer Stunde haben alle CPEs das neue Prefix und Preferred=0 für das alte Prefix gelernt
Nach einer weiteren Stunde haben alle Devices vom CPE die Änderungen gelernt
- Nach 24h (Prefix valid und preferred)
 - `2001:db8:cafe:affe::/64` [Valid 86400, Preferred 82800]
 - Renew-Intervall/RA-Intervall: 3600s
 - Nach einer Stunde kennen alle CPEs nur noch das neue Prefix
Nach einer weiteren Stunde haben alle Devices vom CPE die Änderungen gelernt

Prefix-Wechsel nur für eine „Kaskade“

- Richtige Konfiguration von 4 Parametern:
 - Renew-/RA-Intervalle
 - Valid-/Preferred-Lifetimes



Privatkunden

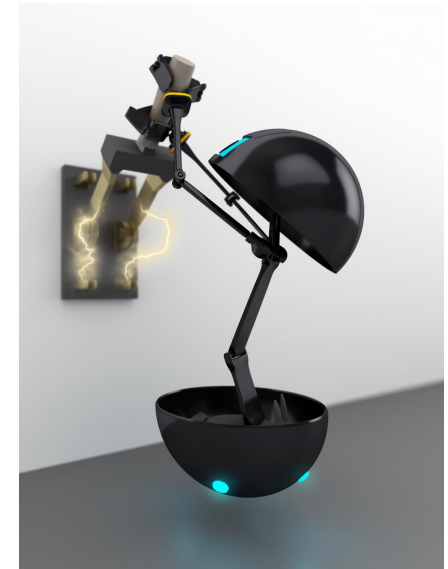
- Unitymedia KabelBW: für Prefix-Wechsel ist Reboot notwendig aber nicht hinreichend
- DTAG: Prefix-Wechsel durch Reboot



Prefix-Wechsel nur mit Service-Unterbrechung

Geschäftskunden ?

- Verwendung von IPv6 analog zu IPv4
 - DHCPv6 ebenfalls analog zu DHCPv4
 - Typische Szenarien
 - zentraler DHCP-Server mit unternehmensweiter Konfiguration
 - dezentrale DHCP-Server (pro Standort, pro Unternehmenseinheit, ...)
- Automatischer Prefix-Wechsel nur mit folgendem Aufbau möglich
 - zentraler DHCPv6-Server verteilt Prefixes an untergeordnete DHCPv6-Server
- Offener Punkt: Koppelnetze
- Szenarien
 - Netzbau (Provider / intern)
 - Provider-Wechsel



In welchen Größenordnungen bewegen wir uns?

Zeiten

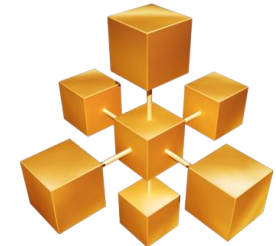
- Option 1: Preferred Lifetime direkt auf 0 setzen
 - $N * \text{Renew-Intervall}$, mit N Anzahl der Stufen
- Option 2: Preferred Lifetime um vergangene Zeit vermindern
 - $\text{Preferred Lifetime} + N * \text{Renew-Intervall}$



Netze

- Koppelnetze werden mit öffentlichen /64 Netzen betrieben (symmetrischer Baum mit n Stufen und je k Verzweigungen)

$$\text{Anzahl Netze} = \frac{k^{(n+1)} - 1}{k - 1}$$



→ Verhältnis von Zugangsnetzen zu Gesamtanzahl bei 80-90%,
d.h. 10-20% der verfügbaren /64 Netze werden für Koppelnetze verwendet

Mehrstufige Hierarchien realisieren

Koppelnetze

- öffentliche Netze
 - /64 → 10%-20% aller verfügbaren Netze werden „verschwendet“
 - /126 → zwischen Routern inzwischen häufig eingesetzt, geringe Probleme
- link-lokale Adressen
 - Erreichbarkeit des Routers?
 - Management über IPv4 ;-)
 - Verwendung von öffentlichen Loopback-Adressen

Öffentliche Netze

- Probleme
 - Renumbering in jedem Abschnitt → Jedes Interface muss IPv6-Adresse wechseln

Link-lokale Adressen

- Probleme
 - Änderung von Loopback-Adressen nicht out of the box automatisiert änderbar

Herausforderungen

Szenarien

- Netzbau (Provider / intern)
- Provider-Wechsel

Das Relay-Problem

- DHCPv6-Server sind fast nie direkt von den Clients erreichbar
- DHCPv6-Relay arbeitet als Proxy und leitet die DHCP-Requests an die IPv6-Adresse des DHCPv6-Servers weiter
- DHCPv6-Server müssen öffentliche IPv6-Adressen verwenden

- Sobald die Adresse bekannt wäre
 - DHCPv6-Relay muss IPv6-Adresse des DHCPv6-Server verwenden
 - DHCPv6-Server muss den Dienst auf neue IPv6-Adresse wechseln
 - Zeitliche Koordinierung wichtig!

- Es fehlt ein Protokoll mit dem ein übergeordneter DHCPv6-Server allen Relays seine neue IPv6-Adresse mitteilt
Alternative: Eine Prozedur durch die ein untergeordneter Relay nur den übergeordneten DHCPv6-Server nach seiner Adresse fragt
 - Broadcast auf FF05::1:3 dafür nicht verwendbar

Anforderungen an die Infrastruktur

DHCPv6-Server

- Konfiguration wird durch eigenen DHCPv6-Client durchgeführt
- Wechsel der IPv6-Adresse im laufenden Betrieb
- Eigentlicher Prefix-Wechsel
 - Hinzufügen des neuen Prefix inkl. Leasevergabe aus altem und neuem Prefix
 - Löschung des alten Prefix

DHCPv6-Relays

- Bestimmung der DHCPv6-Server-Adresse bei einem Prefix-Wechsel
- Wechsel des Relay-Ziels im laufenden Betrieb

SLAAC-Router

- Konfiguration wird durch eigenen DHCPv6-Client durchgeführt
- Eigentlicher Prefix-Wechsel
 - Hinzufügen des neuen Prefix inkl. Leasevergabe aus altem und neuem Prefix
 - Löschung des alten Prefix



Anforderungen an die Infrastruktur

„Nur Prefix-Wechsel“

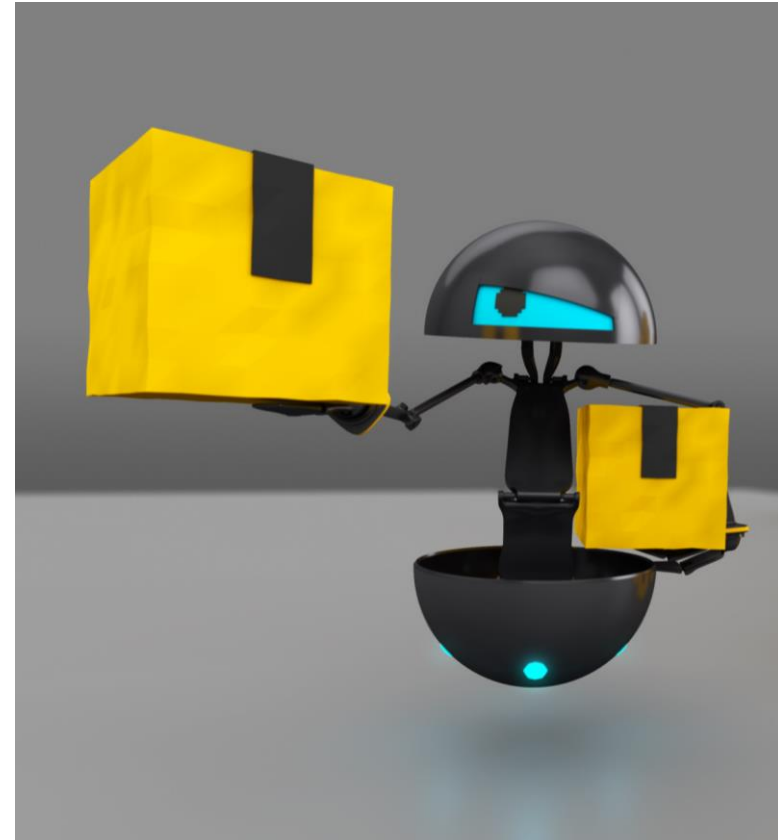
- Konfigurationsparameter
 - Zeitpunkt für Bekanntgabe des neuen Prefixes
 - Zeitpunkt für Status=deprecated für altes Prefix
 - Zeitpunkt für Status=invalid für altes Prefix /Entfernung des alten Prefixes

ISC

- Konfigurationsfile muss für jeden Zeitpunkt erstellt werden
- Kein Tool für die Erstellung der Konfigurationsfiles vorhanden

RADVD

- Konfigurationsfile muss für jeden Zeitpunkt erstellt werden
- Kein Tool für die Erstellung der Konfigurationsfiles vorhanden



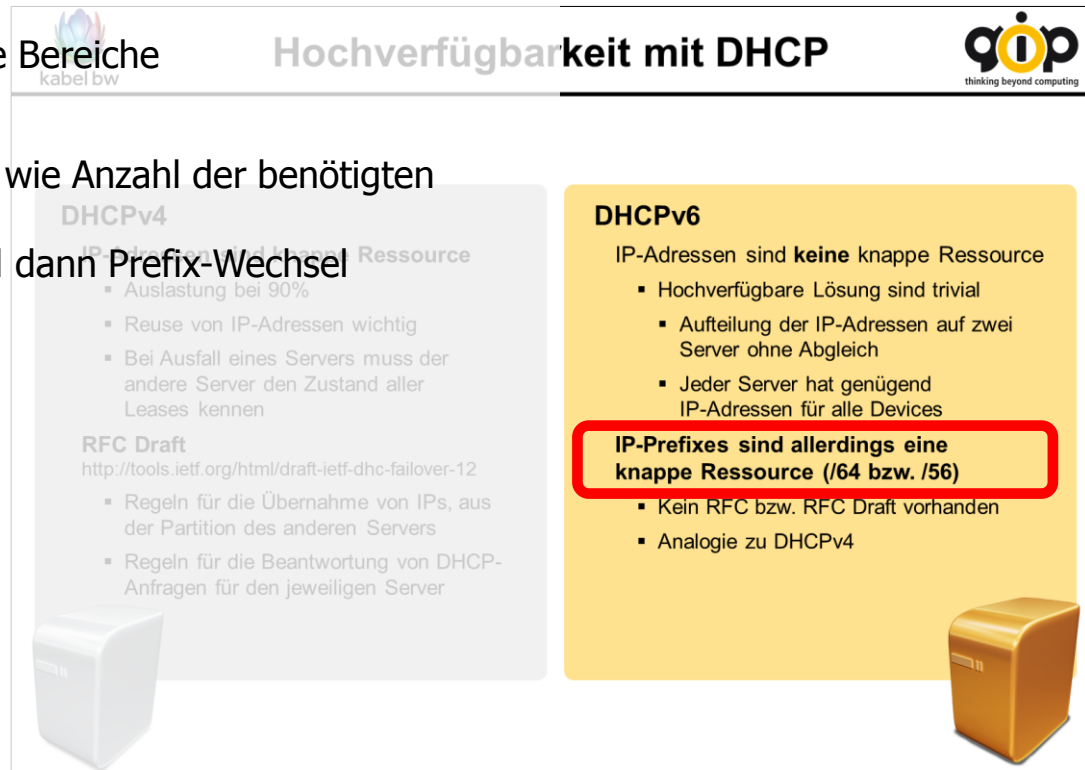
Zwei Optionen

DHCPv6-Option 7: Preference Option

- Sub-DHCPv6-Server (1)
 - Nur Einzel-Adressen aus einem /64 Subnetz: Kollisionen selten
 - Zuständigkeit per Preference Option (Hot-Standby)
- Sub-DHCPv6-Server (2)
 - Aufteilung der Prefixe auf zwei disjunkte Bereiche
- DHCPv6-Server
 - Falls Anzahl der Prefixe doppelt so groß wie Anzahl der benötigten Prefixe
→ Einsatz von 2 Servern und mit Notfall dann Prefix-Wechsel
 - Verwendung von langen Laufzeiten?
 - Kurzes MTTR?

DHCPv6-Cluster

- Konfiguration analog zu einem DHCPv6-Server



Hochverfügbarkeit mit DHCP

DHCPv4

- IP-Adressen sind eine knappe Ressource
 - Auslastung bei 90%
 - Reuse von IP-Adressen wichtig
 - Bei Ausfall eines Servers muss der andere Server den Zustand aller Leases kennen

RFC Draft
<http://tools.ietf.org/html/draft-ietf-dhc-failover-12>

- Regeln für die Übernahme von IPs, aus der Partition des anderen Servers
- Regeln für die Beantwortung von DHCP-Anfragen für den jeweiligen Server

DHCPv6

IP-Adressen sind **keine** knappe Ressource

- Hochverfügbare Lösung sind trivial
 - Aufteilung der IP-Adressen auf zwei Server ohne Abgleich
 - Jeder Server hat genügend IP-Adressen für alle Devices

IP-Prefixes sind allerdings eine knappe Ressource (/64 bzw. /56)

- Kein RFC bzw. RFC Draft vorhanden
- Analogie zu DHCPv4

Offene Punkte

Prefix-Länge

- Root-DHCPv6-Server
 - Vergabe von /64 Prefixes für jede Anfrage
 - Anfragen unterscheiden sich in Relay-Agent-IP
 - Potentiell auch verschiedene DUID verwendbar
 - Zerschneidung des vorliegenden Prefix in größere Teile?
 - Hierarchien > 2 Stufen (unsinnig?)
 - Hoher Planungsaufwand (Auf welche Anfragen müssen welche Prefixlängen vergeben werden)
 - Signalisierung von Sub-DHCPv6-Servern zu DHCPv6-Server unklar

Netzdokumentation

- DHCPv6-Server
 - DHCPv6-Relays: IPv6-Adressen \leftrightarrow MAC-Adressen
 - DHCPv6-Server: IPv6-Adressen, IPv6-Prefixe \leftrightarrow MAC-Adressen
- Sub-DHCPv6-Server
 - Devices: IPv6-Adressen \leftrightarrow MAC-Adressen

Offene Punkte

Firewalls

- Duplizierung aller Regeln auf die neuen IPv6-Adressen nötig
- Signalisierung über Zuteilung von Prefixen zu besonders geschützten Netzbereichen fehlt
- Löschung der Regeln zum alten Prefix

DNS

- Dynamisches DNS (häufig auf IP-Adressen eingeschränkt)
 - Wechsel auf Authentifizierung nötig, da sich IP-Adressen der DHCPv6-Server ändert
- Zeitweise zwei IP-Adressen zu einem Host
 - DNS-Eintrag verschwindet erst bei invalid, sollte aber schon bei deprecated gelöscht werden (50% aller neuen eingehenden Verbindung auf deprecated!)

Router

- Wechsel der Management-Adresse
- Konfigurationsänderungen?

Derzeitige Situation

Privatkunden

- Bei diversen Providern im Einsatz
- Prefix-Wechsel ist mit Service-Unterbrechung verbunden
- CPEs müssen noch dazulernen

Geschäftskunden

- Verwendung von Provider Independent (PI) Address Space
 - >22.000 objects in ripe database
- Verwendung von privaten IPv6-Adressen und NAT
 - Zu finden in Universitätsnetzwerken
- Administration wie bei IPv4
 - Service-Unterbrechung
 - Einige Server/Applications/Datenbanken funktionieren nur mit statischen IPs



Renumbering

Offene Quellen / Internet

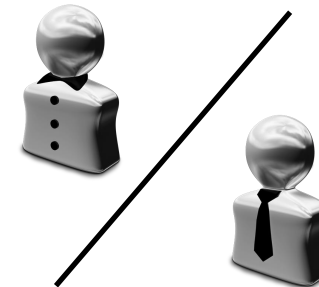
- Wenig Literatur zu finden
- Vortrag von Fred Bovy 2011
- Häufig nur auf SLAAC beschränkt

Infrastruktur

- Noch nicht für Prefix-Wechsel geeignet
- Keine Planungstools / Migrationstools vorhanden
- Protokolle z.T. nicht ausreichend

Einsätze/Realität

- Privatkunden
 - Noch kein Thema, da viele Geräte im Heimnetz nur IPv4 unterstützen
 - Keine öffentlichen IPv6-Adressen im Einsatz → keine Serviceunterbrechung
- Geschäftskunden
 - PI-Adressen oder IPv4 oder Site-Local + NAT



"On the other side of the screen,
it all looks so easy."
Kevin Flynn



Vielen Dank.

Dr. Dietrich Schroff
GIP AG

Research Fellow