

Kommunikationsarkitektur för samhället

Del 3: Teknisk specifikation

Version 0.14

Fem små hus presenteras i tre dokument:

- Del 1. Sammanfattning
- Del 2. Översikt, grundläggande krav och principer, organisation
- Del 3. Teknisk specifikation

Fem små hus – Del 3: Teknisk specifikation Version 0.14	2 (16)	
Innehåll		
Innehåll	2	
Teknisk målsättning	3	
Husregioner	3	
Underhåll och vidmakthållande	6	
Fiber mellan husregioner	6	
Infrastrukturtjänster i ett hus	7	
Tillgänglighetsövervakning	7	
Routing	8	
ICMP	10	
MTU	10	
Tid och frekvens i en husregion	11	
Vilken funktion erbjuder infrastrukturen?	11	
Kryptosystem	12	
Routing registry-funktion	13	
Fibervägskoordinering		

14

Förkortningslista

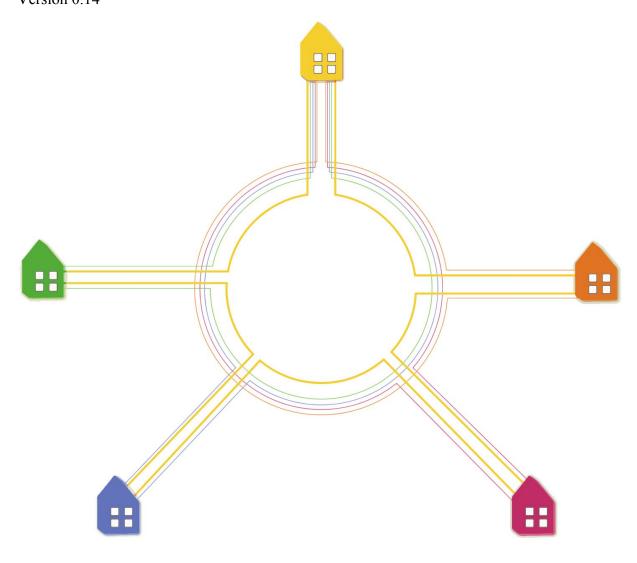
Teknisk målsättning

Den övergripande tekniska målsättningen med Fem små hus är att en användare eller producent av internetbaserade tjänster inom Sverige aldrig ska drabbas av ett avbrott som varar längre än tre minuter. Kravet gäller även när infrastrukturen har skadats och en eller flera operatörer saknar egen anslutning till en eller flera regioner. Under 99,8% av tiden ska inga paket förloras i Fem små hus-infrastrukturen och under de 17 återstående timmarna per år ska maximal paketförlust vara 2% av paketen i en riktning. Envägsfördröjningen av paket inom Sverige ska under 99,8% av tiden vara kortare än 30 ms. Under normala driftsituationer ska en användare ha maximalt 10 avbrott per dag om maximalt 200 ms inom en operatörs nät.

Användare med högre krav rekommenderas att anordna sin trafik så att den i normalfallet har flera alternativa vägar genom nätet. Det kan exempelvis åstadkommas genom att ansluta sig till flera operatörer som ingår i den nationella infrastrukturen och ha en IPv6-tunnel genom varje operatörs nät. Användaren kan kombinera ett internt routingprotokoll (IGP), t.ex. IS-IS, med BFD för att snabbare detektera avbrott eller vägar som inte uppfyller SLA. Efter upptäckt kan användaren flytta sin trafik till en bättre framföringsväg.

Husregioner

En husregion täcker en yta med ungefär 60 000 – 120 000 invånare. Inom en husregion placeras fem hus med ett minsta inbördes avstånd på 7 km, med undantag av att två hus per husregion tillåts stå inom 1 km från varandra. Fiberavståndet mellan två hus bör hållas under 80 km och ska i normalfallet vara under 50 km. I extremfall kan fiberavstånd upp till 110 km tillåtas. Samma gränsvärden för fiberavstånd gäller mellan hus i olika husregioner.



Figur 1 Fiberförbindelser mellan hus i en husregion.

Figur

Alla hus i en region ska vara förbundna med varandra via diversifierad fiber.

Fiberanslutningarna kan antingen anläggas som en ring som ligger minst 500 meter från husen eller som punkt-till-punkt-anslutningar mellan husen. I alla fall ska förbindelserna mellan husen i en husregion vara separerade minst 100 meter på alla punkter under dess väg till andra hus.

Kanalisationen av fiber till hus i en annan husregion bör vara separerad från kanalisationen mellan hus i samma region. Om möjligt bör även anslutningsfiber till operatörer som ansluter stadsnät och användare inom regionen ha separat kanalisation fram till husen.

Fyra fibervägar till ett hus skulle kunna fördelas på följande vis.

- Väg 1 Fjärrnät till annan husregion.
- Väg 2 Stadsnät, kommunikationsoperatörer och lokala operatörer samt en eventuell andra förbindelse till annan region.
- Väg 3 Första vägen till alla andra hus i regionen samt en eventuell tredje förbindelse till annan region.
- Väg 4 Andra vägen till alla andra hus i regionen samt en eventuell fjärde förbindelse till annan region.

Ett referenshus kan dimensioneras enligt följande.

- EMP-skydd. Kraftförsörjningen har skyddad galvanisk genomföring.
- Vatten, avlopp, grovstädskrubb och toalett.
- Reservdelsförråd med hyllor (20m²).
- Batterirum.
- Datorskärm (VGA, HDMI, USB-C), tangentbord och mus.
- 15 serverrack med vardera 11 kW elektrisk effekt (3×16A trefas¹).
- Elmatning och elverk för 240 kW elektrisk effekt (400A/fas).
- Bränsle för 7 dagars drift på reservkraft.
- Kyla 222 kW kyleffekt, exempelvis fördelad på 6 kretsar om 37 kW med en 37 kW-krets i reserv.
- UPS 170 kW elektrisk effekt i 5 minuter².
- Skyddsnivå 3 enligt bilaga 1 i MSB:s Vägledning för fysisk informationssäkerhet i it-utrymmen.
- Trådlöst nätverk med anslutning och autentisering liknande Eduroam.

Hus med halv och en fjärdedels elektrisk effekt jämfört med referenshuset specificeras också. De har samma fysiska dimensionering som referenshuset med nedanstående skillnader.

1/2 effekt referenshus:

- 15 serverrack med vardera 7 kW elektrisk effekt (2×10A trefas).
- Elmatning och elverk för 160 kW elektrisk effekt (250A/fas).
- Kyla 148 kW kyleffekt, exempelvis fördelad på 4 kretsar om 37 kW med en 37 kW-krets i reserv.
- UPS 110 kW elektrisk effekt i 5 minuter.

¹ Gemensam försörjning av 48V likström har medvetet valts bort. Operatörer med behov av 48V likström tillhandahåller egen utrustning för detta.

² Dimensioneras så att generatorn hinner starta vid strömavbrott. Generatorn konfigureras att ladda UPS-batterierna vid generatordrift. Återväxling från generatordrift skall ej ske innan UPS-batterierna har laddats så att de har tillräckligt med energi för att överbrygga återväxlingen till elnät.

1/4 effekt referenshus:

- 15 serverrack med vardera 2,3 kW elektrisk effekt (1×10A enfas).
- Elmatning och elverk för 120 kW elektrisk effekt (125A/fas).
- Kyla 74 kW kyleffekt, exempelvis fördelad på 2 kretsar om 37 kW med en 37 kW-krets i reserv..
- UPS 40 kW elektrisk effekt i 5 minuter.

För hus som inrättas i redan befintliga byggnader gäller samma regler för tillträde, kyla, kraft etc. som för nya hus.

Underhåll och vidmakthållande

En underhålls- och reservdelsutredning måste genomföras. Målet bör vara att vid en mycket svår kris ska minst två av vägarna vara fungerande med full kapacitet och stabilitet efter 90 dagar. Minst en väg ska fortsatt fungera 270 dagar efter krisens inledning. Detta förutsätter att bränsle tillförs vid reservkraftdrift längre än sju dagar.

En materielomsättningsplan behöver också tas fram. Delar såsom batterier, kraftenheter, fläktar m.m. behöver omsättas med jämna mellanrum. För all utrustning behövs utbytesplaner med omkring 5-7 års intervall beroende på utrustning. Installation av utrustning bör spridas ut över tiden mellan olika operatörer så att materielomsättningar kan ske stegvis.

Fiber mellan husregioner

För att planera husregionernas och husens placering behöver ett simuleringsverktyg utvecklas. För att simuleringsverktyget ska kunna planera topologi och förläggning effektivt behövs indata i form av:

- Geografiskt fördelad befolkningsmängd.
- Existerande fiberinfrastruktur och geodata som beskriver vilka delar av kanalisationen där vägar är närmare varandra än 500 meter.
- Kostnader för fiberförläggning i olika typer av terräng.
- Terrängdata.
- Existerande anläggningar som kan användas som hus.
- Andra krav på önskade anslutningspunkter.

Parametrar och regler som planeringsverktyget ska kunna ta hänsyn till är:

- Högst en operatör per hus får använda huset som "transitnod", d.v.s. mer än en förbindelse till annan husregion.
- Fördelning av transitnoderna så att det finns 5 oberoende av hus och kanalisation "vertikala" (från norr till söder) ("stamnät").

- Varje operatör ska ha 2 till 3 separata vertikala stamnät som är horisontellt förbundna med fem horisontella fiberförbindelser, där de olika operatörernas horisontala ihopkopplingar ej har något gemensamt hus.
- Kluster av husnoder. Intilliggande husregioner kan ses som en "superregion" med färre än fem vägar ut, trots att de individuella regionerna har fem anslutningar vardera.

Efter noggranna simuleringar specificeras och anläggs de saknade fiberstråken. Lämpligen förstör man insamlade och genererade data när man är klar.

Infrastrukturtjänster i ett hus

- Husrouter f\u00f6r anslutning av resurser och tj\u00e4nster som finns i huset. Den ansluts till
 alla operat\u00f6rer i regionen.
- Redundant DNS-resolver att användas av användares klienter i husregionen.
 Resolvrarna har två unicastadresser som är gemensamma för alla hus i landet.
 Resolvern kan förses med specialfunktioner så det är möjligt att exempelvis ladda ner hela zoner i förväg.
- DNSsec-möjlighet med publikt tillgängliga nycklar.
- DNS-server f\u00f6r root, se, in-addr.arpa och gemensam infrastruktur. Fram- och bakl\u00e4ngesuppslagning.
- Enumhantering i både DNS-server och DNS-resolver. Innehåller funktioner för omdirigering av vissa nummer, exempelvis 112, i händelse av att husregionen blir isolerad.
- Routingannonserare och analysator av tillgänglig konnektivitet.
- Eventuellt **tid**, **frekvens och fas** (stratum-1).
- **NTP-tid** (stratum-2).
- Webbserver med statusinformation och distribuerad krisinformation, nåbar via anycast som innebär att servrar på olika platser i infrastrukturen delar på samma IP-adress.
- Pingserver som svarar på en nationell anycastadress och en anycastadress som är unik för husregionen.

Den slutliga förteckningen över infrastrukturtjänster i varje hus tas fram av en arbetsgrupp.

Tillgänglighetsövervakning

Funktionen för den nationella infrastrukturen upphandlas enligt beskrivning i del 2. Den måste kontinuerligt testas och övervakas under normal drift. För det krävs ett system som testar trafikkapacitet och routingfunktionalitet. Systemet ska ha automatiserad felrapportering och funktioner för statistisk uppföljning i efterhand.

Övervakningssystemet bör slumpmässigt testa den upphandlade aktiva infrastrukturen så att det under en period av 14 dagar har verifierat mer än 85% av infrastrukturens geografiska täckningsområde. Under varje 60-dagarsperiod bör alla delar i systemet ha testats. Förslagsvis sker testerna mellan hus i olika husregioner. Om man har 100 GE-interface på datorer i vart och ett av husen kan man testa upp till 500 Gbit/s.

Tester ska utformas med utgångspunkt från de dokument och metoder som tagits fram av IETF:s IPPM-arbetsgrupp. Framtagna tester ska publiceras. Kapacitet ska testas till 95% av upphandlad kapacitet vad gäller genomströmning. Funktioner som MTU, pakettransparens samt routingsystem ska testas till 100%.

Routing

För all routing mellan operatörer, utrustning och abonnenter med speciella behov används routingprotokollet BGP4. Nedan följer några riktlinjer och tankar om hur routing kan konfigureras i en framtida infrastruktur. En arbetsgrupp administrerad av sekretariatet får till uppgift att ta fram den slutgiltiga konfigurationen vid införandet av infrastrukturen.

Routingannonseraren annonserar alla prefix som används i husregionen, se adressplan. Oberoende av vilken operatör som prefixen kommer från annonseras de kortare prefixen med AS som originator. Routingannonserarna i alla hus delar på samma AS-nummer.

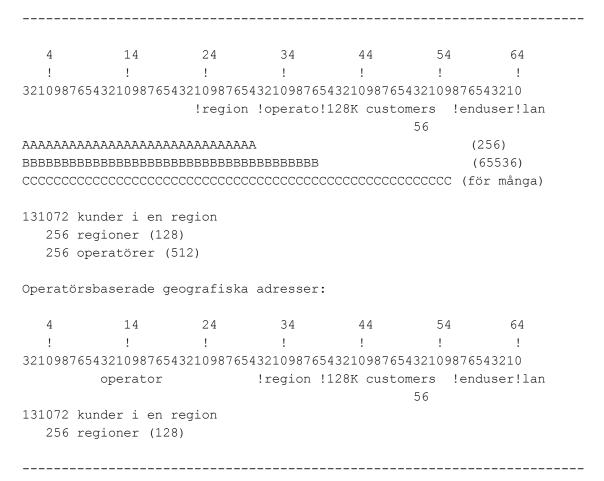
En viktig funktion är att se till att en operatör som förlorat sin egen konnektivitet till en husregion kan nå sina egna ändpunkter då annonseringen kan komma via routingannonserarens AS.

Operatörerna sänder alla routes från samtliga prefix som ingår i den svenska infrastrukturen till en routingannonserare. Det bör gå att sammanställa dessa routes lokalt för att få överblick över aktuell nåbarhetsinformation.

Normalt filtreras BGP-prefixannonseringar erhållna via EBGP som innehåller eget AS eftersom de sannolikt är orsakade av routingloopar. Exempel: Om operatör 1 har AS 10, operatör 2 AS 20 och routingannonseraren AS 30 kommer operatör 1 att se sina egna prefix i en husregion den förlorat egen förbindelse till som "20 30 i", vilket är giltigt. Annonseringen i regionen från operatör 1 till 2 kommer däremot att ha "20 10 i", vilket inte fungerar i normalfallet. Begränsningen med den föreslagna modellen är att om nätet fragmenteras till den grad att operatör 1 ser sig själv via operatör 2 och en ö av operatör 1:s nät som operatör 1 inte har direktkontakt med kommer annonseringen att bli exempelvis "20 10 30 i" vilket kommer att droppas av operatör 1:s router. (Se RFC 4271.)

Genom samarbete med alla operatörer ordnas signering av routeannonseringarna.

Figur 2 visar en tänkt adresseringsplan för Fem små hus. Ett sätt att realisera "husgruppsprefix" annonserade av både routingannonseraren och operatörens egen router är att operatören skickar två uppdateringar med en bit längre prefix än routingannonseraren. Med adresseringsexemplet i figur 2 skulle operatörens router skicka två stycken /32 och routingannonseraren ett /31. Lämpligen läggs adressplanen ut så att sista biten i regiondelen är ledig. Det möjliggör en framtida fördubbling av regionernas storlek.



Figur 2 Modell för geografisk IPv6-adressering i Fem små hus ("Sverigeprefixet").

Genom arbete i IETF eller kontakter med leverantörer adderas funktionalitet i routrarna så att det går att definiera en lista av AS-nummer vilka man accepterar sina egna AS-nummer bakom, en eller flera gånger.

IETF har tagit fram en metod för verifiering av routingannonseringar i BGP – RPKI. Tyvärr innebär metoden ett mer eller mindre centralt system som kan användas som en avstängningsbrytare på AS- eller prefixnivå. Om detta ska användas inom Sverige måste en

egen infrastruktur med distribuerade RPKI-valideringsservrar upprättas. I så fall krävs att uppdatering och kontroll av dessa inte hamnar i fel händer och skyddas från konfigurationsfel med mera.

Förslagsvis uppdras åt varje operatör i den gemensamma infrastrukturen att driva egna servrar för detta ändamål i varje region. En gemensam databas för konfigurationsinformation skapas. Därigenom skulle en eventuell "utslagning" begränsas och en attackerare måste hantera minst fem olika organisationers uppdateringsprocesser. Med andra ord: när man hämtar från den gemensamma databasen kan man validera att data i den inte stör de upphandlade reservfunktionerna hos respektive operatör.

ICMP

Till skillnad från IPv4 stödjer inte IPv6 fragmentering av paket. Istället förutsätts ändsystem använda minsta gemensamma MTU på en given sträcka som maximal paketstorlek. Om en nod tar emot ett paket större än vad som är möjligt att vidarebefodra skickar den ett ICMP-paket till paket avsändaren med information om detta. Om en operatör av någon anledning filtrerar bort ICMP-paket uppfylls inte kravet och stödet för IPv6 blir bristfälligt. Det blir då möjligt att hamna i situationer där förbindelsen bryts på grund av paketförlust utan ändsystemen blir informerade om detta.

MTU

Infrastrukturmodellen utgår från tre olika IPv6 MTU. Detta används för att beskriva paketstorklekar eftersom olika lager 2-media har olika ramstrukturer. Ett Ethernet-paket är vanligen 16 eller 18 byte större beroende på vilken Ethernet-standard som används.

Operatörsstamnät – IPv6 MTU 9180 byte

Med operatörsstamnäten avses den del av infrastrukturen som förmedlar trafik mellan operatörer i sammankopplingspunkter för trafik. Det kan exempelvis vara punkt-till-punkt-förbindelser mellan hus i en region. En operatör kan internt i sin infrastruktur ha stöd för större paket, men de kommer inte garanterat att kunna hanteras när en annan operatör används som reservväg. Valet av MTU-storlek har gjorts så att en operatör kan tunnla en användares IPv6-trafik (med MTU 9000 byte) inom eget nät.

Användarnät – IPv6 MTU 9000 byte

Alla nätanvändare och hela den svenska infrastrukturen ska klara av att förmedla IPv6-paket med MTU på 9000 byte. Se IETF BCP 39.

Krypto- och VPN-användare – IPv6 MTU 8192 byte

Användare av en tjänst som produceras genom tunnling av trafik över en användaranslutning måste begränsa sin MTU så att det finns utrymme för den information som tillförs av tunnlingsprotokollet.

Tid och frekvens i en husregion

Varje hus har en stratum-2 NTP-server som synkroniserar mot de tillgängliga nationella tidsnoderna och distribuerar vidare till klienter i husregionen. Den lokala NTP-serverns tidbas kan vara en dators interna klocka eller en lokal lågpresterande tidsnormal. Därigenom kan alla klienter i en husregion följa den lokala stratum-2-servern och driva med denna när regionen blir isolerad.

Husen är försedda med GNSS-mottagare och skillnaden mellan GNSS-tid och NTP-serverns tid loggas, men används ej till distributionen. Om GNSS-systemet ger tillförlitlig information kan man i efterhand skapa spårbarhet vid behov.

Finns det andra användare med högre krav och möjlighet till finansiering kan tid- och frekvensfunktionen i en husregion uppgraderas till samma standard eller bättre än den som erbjuds av de nationella tidsnoderna.

Vilken funktion erbjuder infrastrukturen?

Vidarebefordran av IPv6-paket med giltiga adresser från godtycklig avsändare till godtycklig mottagare vid husregion inom 30 ms efter att routing upprättats enligt specifikationen av den dynamiska routingens prestanda.

Accepterar paket med en IP MTU på lägst 9000 byte mot användare och lägst 9180 byte mot infrastrukturen.

Operatörens teknik för lager 0 – lager 2 inom eget nät är ett lokalt val, helt upp till respektive operatör. Uppsatt SLA ska levereras.

Operatörerna väljer internt routingprotokoll (IGP) och struktur själva.

Routing och paketförmedling skall vara distribuerad och automatisk så att godtyckliga öar av operatörens nät kan fungera autonomt. Eventuella topologier av routeservrar ska vara distribuerade och placerade på platser som minst motsvarar kraven för ett hus.

Avlämning vid en husregion sker med lämplig Ethernetteknologi som är felrättande och klarar distanserna mellan husen utan aktiv utrustning utanför huset. 100G staircase DPSK är ett förslag som stöds av alla routertillverkare och de flesta långdistansoptikmoduler. Även

400G och högre hastigheter bör beaktas när interoperabla optiska format har standardiserats. Utvärdering sker löpande av kansliet och anpassas till vid varje tidpunkt aktuella och tillgängliga standarder.

I en husregion kopplas operatörerna mot varandra med en full mesh utan någon gemensam aktiv utrustning förutom respektive operatörs avlämningsrouter. En operatör kan ha flera avlämningsroutrar i en region.

Husets regionsoperatör tillhandahåller två Ethernetswitchar som används som redundant anslutning till regionresurser. Dessa switchar är inte lämpliga för trafikutbyte mellan operatörer eller mellan operatörer och regionala distributörer.

Utbyte av routinginformation inom en husregion mellan operatörer och med routingannonseraren sker med BGP. Det dynamiska kravet är att från det att man har skickat en komplett BGP-uppdatering från en routingannonserare att ett prefix är nåbart ska det annonseras till samtliga routingannonserare i alla nåbara husregioner inom 200 ms. Det ska vara möjligt att skicka trafik annonserad väg när annonseringen tas emot. Samma krav gäller när man annonserar att ett prefix inte längre är nåbart. Även då ska uppdateringen nå alla husregioner inom 200 ms.

En operatör ska kunna acceptera minst 128 000 BGP-prefix, genererade inom den svenska infrastrukturen.

Det är inte tillåtet att annonsera aggregerade husregionsprefix till andra operatörer vid husregionerna. Ut mot resten av världen är externa annonseringar upp till respektive operatör.

Kryptosystem

För de användare som tidigare använt fasta förbindelser eller någon form av VPN-tjänst tillhandahåller Fem små hus-infrastrukturen ett kryptosystem. Kryptosystemet ger skydd mot obehörig trafik, avlyssning och överbelastningsattacker.

Kryptosystemets princip är att det överför Ethernet-paket mellan två punkter genom att man kapslar in det krypterade Ethernet-paketet i ett IPv6-paket som skickas över infrastrukturen. Olika metoder används för att skydda kryptots ändpunktsadresser mot överbelastningsattacker. Har man flera korresponderande motparter via samma krypto identifieras de med ett VLAN per motpart.

Kryptot ansluts till infrastrukuren med 100/400Gbit och tjänsten mot användaren är förmedling av Ethernet-paket med 100/10Gbit-anslutning. Maximal överförd lager 2-Ethernet-MTU är 8210 byte.

13 (16)

För detaljspecifikation av kryptot, se del 4 som tas fram av en separat arbetsgrupp.

En enklare form av tunnling, motsvarande MPLS, är att använda L2TPv3 över IPv6. Funktionaliteten finns i de flesta kommersiella routrar och kan i vissa fall kombineras med routerns kryptofunktion.

Routing registry-funktion

"Offlinefunktion" som genererar konfigurationer till routeannonserare och annan dokumentation.

Fibervägskoordinering

Procedurer och protokoll för erforderlig koordinering måste tas fram av en arbetsgrupp administrerad av sekretariatet eller PTS. Det bör lämpligen göras utan man skapar ett centralt register.

Förkortningslista

Förkortning	Förklaring	Standard
A	ampere	
AS	Autonomous System	RFC 1930
BFD	Bidirectional Forwarding Detection	RFC 5880
BGP	Border Gateway Protocol	RFC 4271
DNS	Domain Name System	RFC 1034
DNSsec	Domain Name System Security Extensions	RFC 4033
DPSK	Differential Phase-shift Keying	
EBGP	Exterior Border Gateway Protocol	RFC 4271
EMP	Elektromagnetisk puls	
GE	Gigabit Ethernet	
GNSS	Global Navigation Satellite System	
HDMI	High-Definition Multimedia Interface	
ICMP	Internet Control Message Protocol	RFC 792
ICMPv6	Internet Control Message Protocol version 6	RFC 4443
IETF	Internet Engineering Task Force	
IGP	Interior Gateway Protocol	
IP	Internet Protocol	
IPv4	Internet Protocol version 4	RFC 791
IPv6	Internet Protocol version 6	RFC 8200
IPPM	IETF arbetsgrupp IP performance measurement	
IS-IS	Intermediate System to Intermediate System	ISO/IEC 10589:2002
IT	Informationsteknik	
km	kilometer	

Förkortning	Förklaring	Standard
kW	kilowatt	
L2TP	Layer 2 Tunneling Protocol	RFC 3931
LAN	Local Area Network	
m	meter	
MPLS	Multiprotocol Label Switching	RFC 3031
ms	millisekund	
MSB	Myndigheten för samhällsskydd och beredskap	
MTU	Maximum Transmission Unit	
NTP	Network Time Protocol	RFC 5905
PTS	Post- och telestyrelsen	
RFC	Request for Comments	
RPF	Reverse Path Forwarding	
RPKI	Resource Public Key Infrastructure	RFC 6480
SLA	Service-Level Agreement	
UPS	Uninterruptible Power Supply	
USB-C	Universal Serial Bus Type-C	
VGA	Video Graphics Array	
VLAN	Virtual LAN	

Om TU-stiftelsen

Stiftelsen för telematikens utveckling, förkortat TU-stiftelsen, bildades 1996 och har som mål att arbeta för ett gemensamt internet i Sverige som klarar alla påfrestningar och tjänar som samhällets enda kommunikationsinfrastruktur.

TU-stiftelsen äger Netnod som driver flera kritiska infrastrukturtjänster i den nordiska grenen av internet och delar ut stipendier till relevant forskning och utveckling inom dessa områden.

Författare

Dokumentet är framtaget av TU-stiftelsens arbetsgrupp för infrastruktur med deltagare från nätoperatörer, leverantörer och myndigheter. Deltagarna har inte representerat sina respektive arbetsgivare utan deltagit som experter. TU-stiftelsen tackar alla deltagare för sin medverkan och för de resurser som ställts till gruppens förfogande i form av möteslokaler, databaser och andra typer av verktyg.