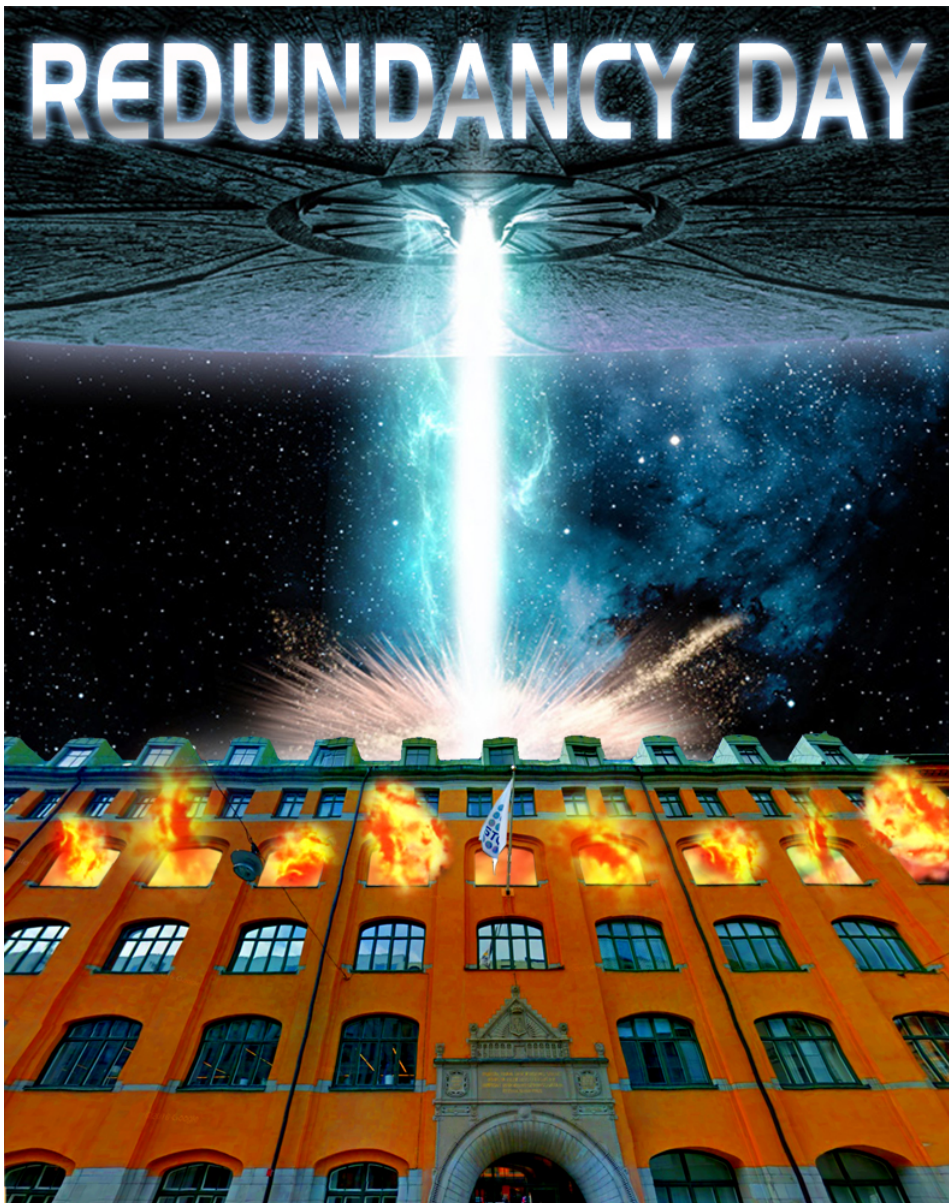


## REDUNDANS ÄR ALLT

Den 27 maj 2017 stannade hela världen, eller åtminstone den del av världen som betjänades av British Airways. Alla flygrutter stod still på grund av bristen på redundans. BA vill i skrivande stund inte avslöja grunden till problemet, men skyller på ett kraftavbrott och på att reservkraften också strejkade. Det fråntar dem inte från skulden att inte ha haft tillräcklig redundans i sin världsomspännande lösning för datadrift.

Allt stannade när serverna för BAs "Fly"-system stannade, e-post, webb och bokningssystem, bagagehantering och passkontroll. Flygplanen blev kvar på marken eftersom bolaget inte kunde kommunicera internt. Rent juridiskt kan BA dock inte skylla ifrån sig på tekniska problem. Slutnotan visade sig hamna kring 890 miljoner kronor.

Samtidigt kan man konstatera att redundans är enda utvägen. Alla andra flygbolag som flög från Heathrow fungerade normalt, för de hade andra lösningar för sin databehandling. Flygledningen vid Heathrow fortsatte som vanligt, eftersom de sannolikt inte har sina system utkontrakterade till någon med bristfällig reservkraft.



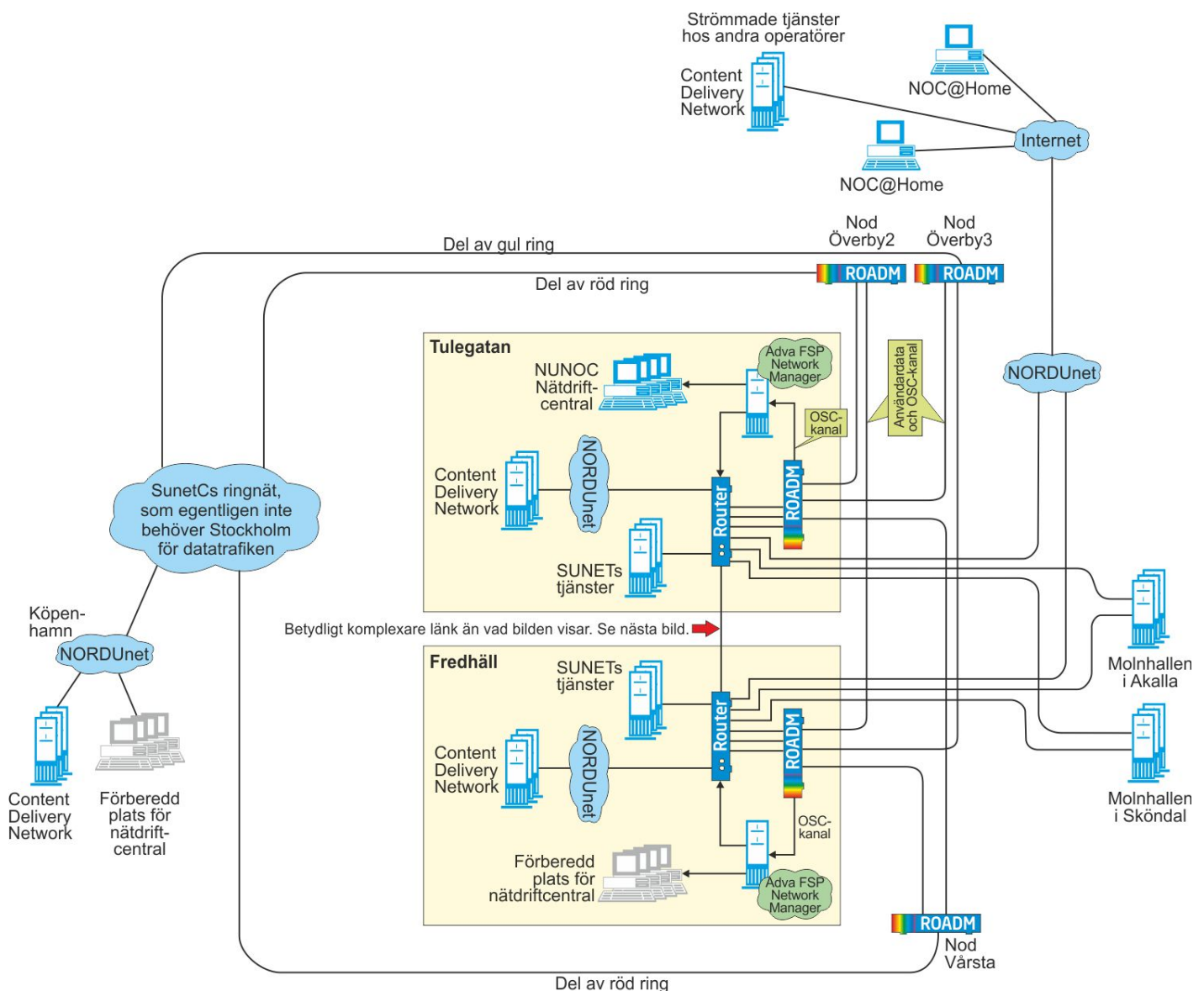
*SUNET Films visar REDUNDANCY DAY, ett drama om vad som händer om det värsta händer. Medverkande: SUNET*

*Regissör: SUNET Manus: SUNET Distribution: SUNET  
Friskrivning: Filmen skildrar en verklig situation och alla  
schemor och personer i filmen är verkliga.*

Det väcker frågan om hur redundansen i SunetC är konfigurerad. SUNETs hallar i Stockholm är inte längre nätets centrum. Sanningen är att SunetC är konstruerat för att fortsätta fungera även om Stockholm "försvinner". SunetCs väg genom Stockholm är inte nödvändig för att datatransporten i landet ska fungera och kommunikation som normalt går genom Stockholm kan gå andra vägar, exempelvis runt Mälaren på röd <--> grön <--> gul ring. Landet skulle uppleva något försämrade prestanda, men det vore också allt som kunde hända.

SUNET har som bekant två centrala hallar i Stockholm som inte alls befinner sig i närheten av varandra. Om ett UFO störtar på Tulegatan 11 påverkar det inte alls hallen i Fredhäll, 3,6 kilometer därifrån.

De båda hallarna körs helt synkroniserat tack vare en väldigt bred länk dem emellan. Genom länken hålls alla databaser ständigt uppdaterade och SUNETs tjänster fungerar lika bra från vardera hallen.



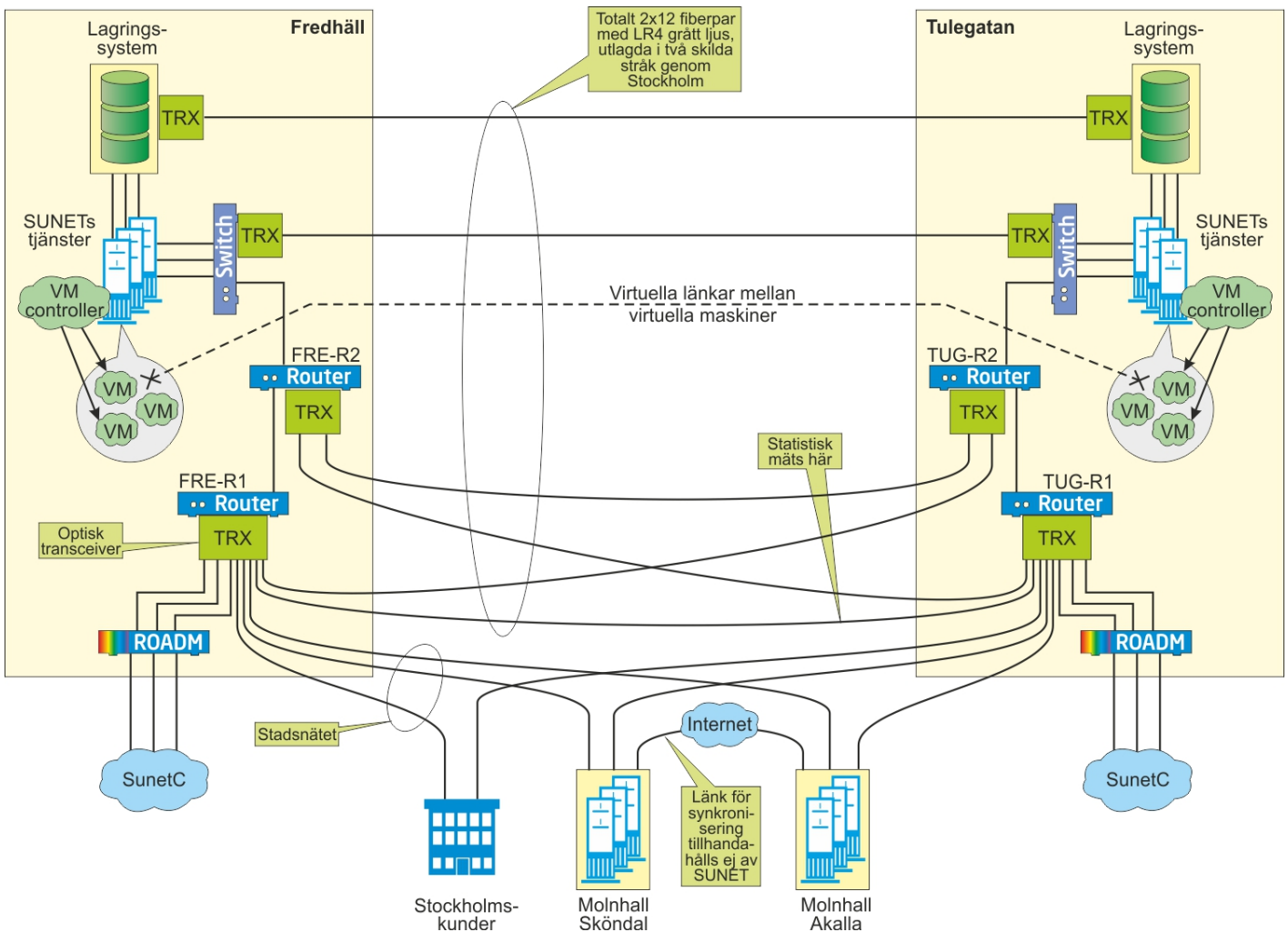
Nätdriftcentralen NUNOC befinner sig emellertid på tredje våningen på Tulegatan och bara där. Om Tulegatan slås ut, slås även driftcentralen ut. Därför finns det en förberedd plats i berggrummet i Fredhäll där man kan ställa ett par datorer och köra nätdriften vidare. Driftdata från ROADM ute i landet kommer in till driftcentralen via den sk OSC-kanalen, en separat optisk kanal som skiljs ut i den mottagande ROADM och tas om hand av en särskild server som exekverar Advas hanteringsmiljö FSP Network Manager. (Se vidare artikeln <https://www.sunet.se/blogg/we-have-liftoff-del-2-av-2/>).

I nödfall kan denna server även nås via Internet och nätdriften skötas av de anställda hemifrån. Utöver detta är alla stamnätssnoder nåbara genom ett antal 4G-modem i form av dataöverföring via mobiltelefonnätet, vilket normalt bara används vid total kommunikationsförlust till en stamnätssnod, eller när den nyinstallerats. Detta nät kan också nås hemifrån via Internet.

Skulle även denna möjlighet slås ut finns en driftplats förberedd i NORDUnets datorhall i Köpenhamn. Används denna får man emellertid försöka koppla sig till varenda stamnätssnod manuellt via SunetC. Man kan dock föreställa sig att i en situation så extrem som denna, kommer det akademiska Sverige att ha andra problem än att hålla SunetC igång.

Återstår så de strömmande tjänsterna. Olika leverantörer av kommersiella tjänster, som Akamai, Netflix och Google har mellanlager (Content Delivey Network, CDN) i SUNETs hallar, i olika utsträckning. I TUG finns Akamai, Netflix och Google, i FRE finns Netflix och Google och i Köpenhamn finns Google. Dessa CDN är inte nödvändiga för att SUNETs åtaganden till sina kunder ska kunna uppfyllas. Skulle en hall slås ut kan de kommersiella aktörerna gå runt problemet och ta sig in på SunetC via Internet i och med att de har CDN hos andra kommersiella nätoperatörer också.

## SYNKRONISERINGSLÄNKARNA I DETALJ



Detta är den fullständiga beskrivningen av den redundanta länken för synkronisering med mera mellan hallarna på Tulegatan och Fredhäll. Det är inte en, utan 2 gånger 12 fiberpar som ligger dragna i två separerade stråk genom Stockholm. Var och en av fibrerna är 10-100 Gbps i form av fyra våglängder "grått" ljus enligt metoden LR4 (Long Range 4). Man har kunnat undvika en dyrare, färgad DWDM-lösning i och med att avståndet är mindre än en mil.

Utöver det som visas i den något grövre ritningen överst, visas här de grundläggande åtgärderna för synkronisering mellan hallarna och tvärtransporten av data mellan de gula och röda ringarna i stockholmsområdet. Se vidare avsnittet "Ett praktiskt exempel – Tulegatan" i artikeln <https://www.sunet.se/blogg/we-have-liftoff-del-4-av-2/>

## DATATRANSPORT GENOM STOCKHOLM

De båda molnen "SunetC" är i själva verket samma nät som försetts redundant med data från vardera hallens ROADM. Routrarna ute i landet skulle i varje stund exempelvis kunna välja att datatransporten från **Uppsala** i gul ring överförs till **Linköping** i röd ring som en transport mellan nod Överby in till TUG-R1 över mellanlänken och ut genom FRE-R1 och vidare mot nod Vårsta för att detta råkar vara den minst belastade vägen. Skulle exempelvis den optiska transceivern som hanterar detta gå sönder i TUG kan data istället flyta genom den korsande länken TUG-R1 – FRE-R2 – FRE-R1 och vidare till Vårsta.

Nästa dataström från exempelvis **Jönköping** på röd ring till **Stockholms universitet** (stockholmskund i schemat) skulle kunna flyta in från nod Vårsta och ut från TUG-R1, men gick TUG-R1 sönder skulle data routas om till FRE-R1 och därifrån ut på stadsnätet.

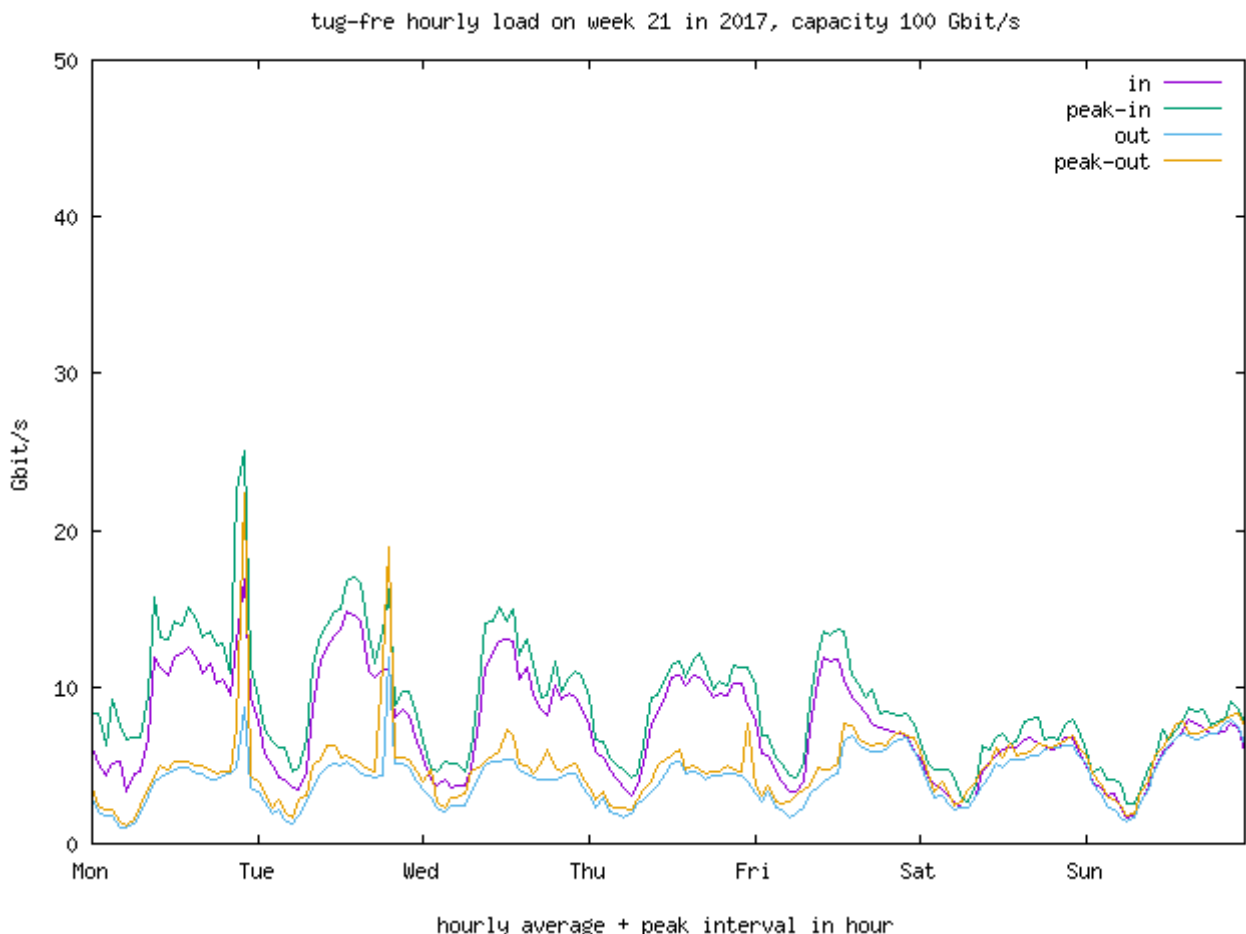
## TJÄNSTERNAS VIRTUELLA MASKINER OCH DATADELNING

Det kritiska för stockholmshallarna är att kunna leverera SUNETs tjänster till användarna. Tjänsterna exekverar som virtuella maskiner (VM) i servrar i de båda hallarna. I den ena hallen körs den skarpa tjänsten, medan en redundant spökkopia exekverar i den andra hallen. Exakt vilken hall som har den skarpa tjänsten är inte viktigt. Det viktiga är att de båda controller-applikationerna är överens om det. Data som skapas i den skarpa tjänsten överförs via en virtuell länk till spökapplikationen i den andra hallen. Skulle kontakten förloras på grund av maskinfel i endera hallen, blir den andra hallen "skarp" och fortsätter att leverera tjänsten som om inget hade hänt. Själva programkoden för tjänsten eller dess lagrade data överförs aldrig via länken, utan finns alltid på båda ställena.

Den virtuella länken dem emellan är bara symbolisk. Det finns givetvis inga virtuella länkar utan att det samtidigt finns fysiska länkar att överföra dem på.

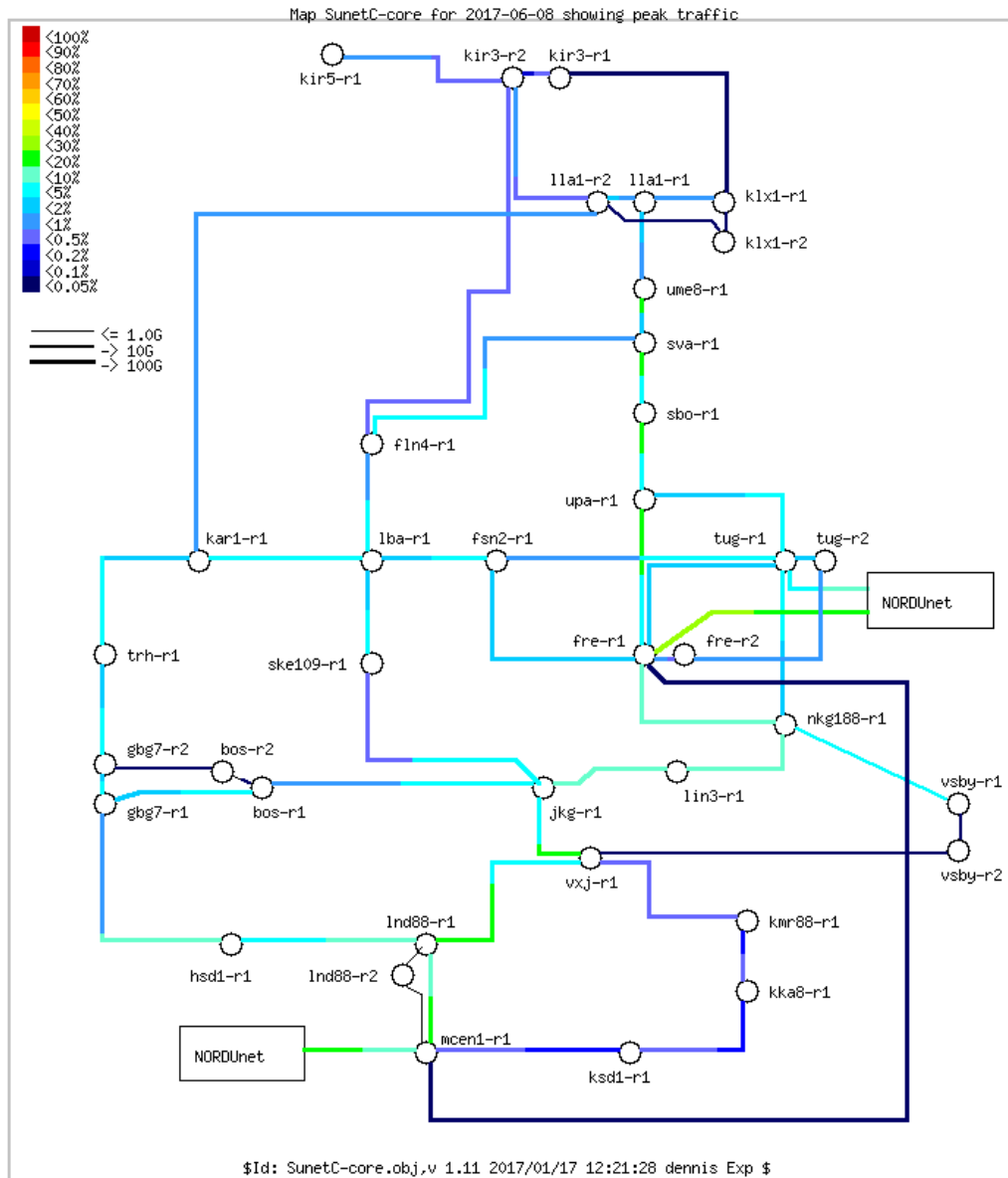
Data som ska delas direkt mellan serverna flyter företrädesvis som Ethernet-paket på de sammankopplade switcharna. Det är mest rakt på sak. Ethernet kan inte routas, men det spelar ingen roll i detta fall. Oavsett vad som händer med servernas datatransport, så synkroniseras lagringssystemen genom sin egen Ethernet-länk. Det tar hand om själva datadriften.

Data från tjänsterna som ska ut till användarna flyter från switchen ut på router R2, därifrån till R1 och ut på SunetC. Går R1 sönder i någon hall kan data utväxlas mellan routrarna R2 och flyta vidare ut på den andra hallens R1. Går någon optisk transceiver sönder i endera routern kan data alltid flyta på en andra transceiver. Skulle ROADM-konstellationen i endera hallen gå sönder, kommer alla tjänster att tas över av den andra. Mer om ROADM och deras konfiguration finns i artikeln <https://www.sunet.se/blogg/we-have-liftoff-del-4-av-2/>



Här ser du nivån på NORDUnets tvärtrafik mellan TUG och FRE som utgörs av synkroniseringen och lastdelningen mellan hallarna. Statistiken är uppmätt där så angivits i schemat ovan. Medeltrafiken ligger kring 10 Gbps. Belastningen är en blandning av tvärtrafiken mellan de olika ringarna i SunetC, stockholmkunderna, synkroniseringstrafiken mellan hallarnas servrar, Facebook osv. Skiljer man ut trafiken för SunetC mellan samma ställen rör det sig omkring 150 Mbps och har inget med synkroniseringen att göra utan utgör enbart lastdelningen, bestående av trafik mellan lärosäten. Ska trafik lämna SunetC och gå vidare ut på Internet så görs det via NORDUnet i Stockholm eller Köpenhamn.

## SNABBASTE OPTISKA VÄGEN



Vilken väg tar ett datapaket egentligen i SunetC? Alla länkar i SunetC har åsatts en "kostnad" baserad på dess längd och antalet hopp genom routrar och systemet ska normalt välja den "billigaste" vägen. Antag att data ska färdas mellan **Luleå** och **Karlstad**. Då kan man se att direktvägen mellan LLA1-R2 och KAR1-R1 är direktast och egentligen alltid borde vinna, medan vägen via UME-SVA-SBO etc är åtta hopp och alltid borde väljas i andra hand. Så blir det inte på grund av kostnadsresonemanget.

Vill lule-studenterna ut på Internet och få tag i Google måste de köra via NORDUnet i Stockholm. Vägen via KAR är bara tre hopp till TUG men är väldigt lång geografiskt eftersom den går längs en kraftledning mitt i landet. Går man kustvägen längs motorvägen E4:an så är det kortare väg. Det tar sex hopp att komma till Stockholm. Systemet väljer först och främst den kortaste vägen.

Vägen via KAR är belastad till <2% medan den andra vägen längs kusten ligger kring 20%, men den långa vägen via KAR är att betrakta som redundant. Räkna med att den kommer att behövas en dag när en grävsropa tagit ett stort bett i kabeln norr

om Uppsala.

Om data å andra sidan ska från **Luleå** till **Karlstad** är vägen LLA1-R2 till KAR1-R1 den kortaste och kommer att väljas i första hand, såvida inte en grävskopa varit elak norr om Karlstad till exempel. Då är vägen UME-SVA-SBO etc att betrakta som en sämre, redundant väg.

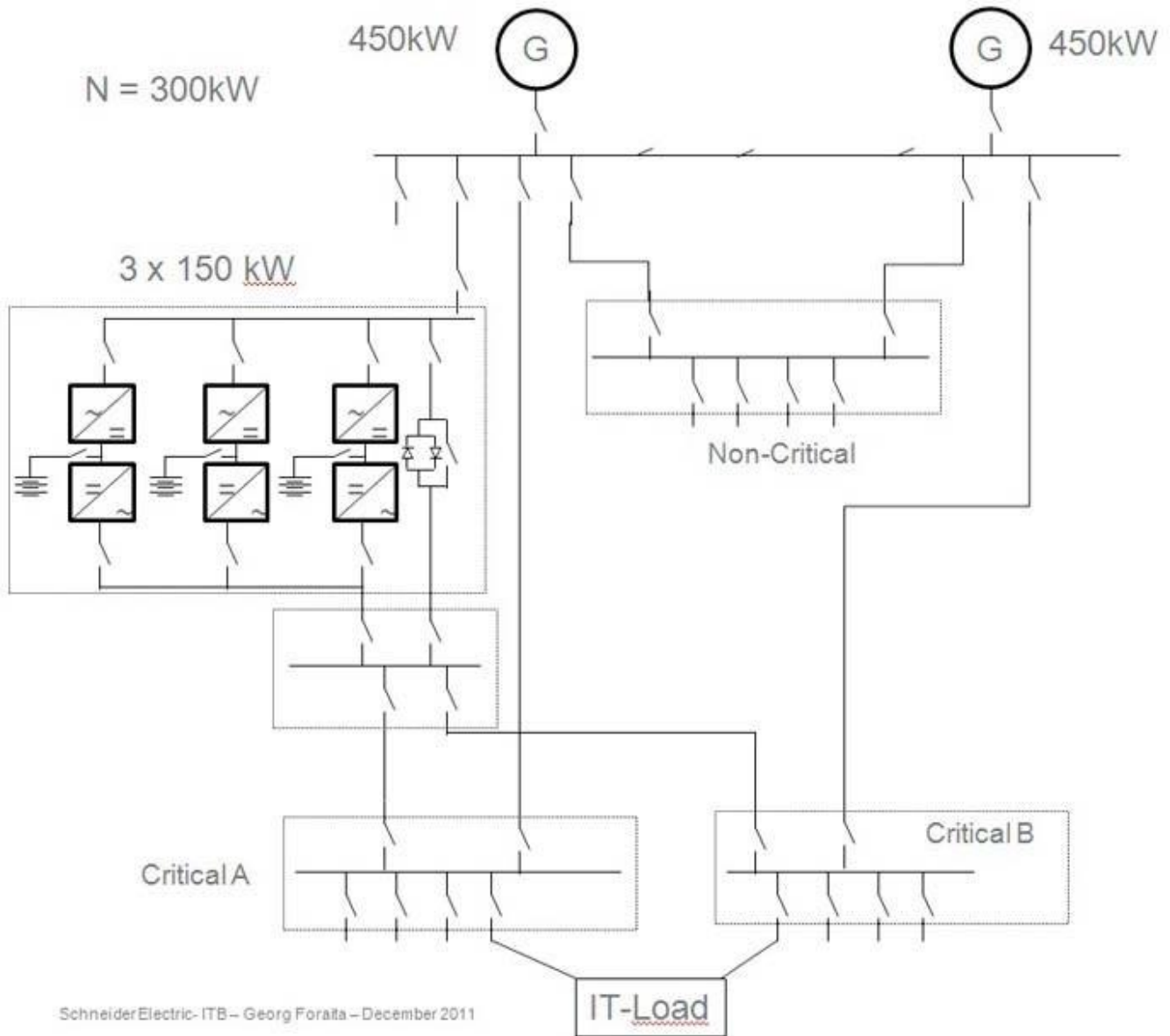
En kommersiell operatör skulle kanske inte resonera på samma sätt. De måste väga in kostnaden för att köra trafik en viss väg och kan komma att välja en trafikmässigt sämre väg för att den bättre vägen kostar flera kronor att hyra. Det är inget som SUNET behöver bry sig om.

## DATORHALLAR MED TIER-3-REDUNDANS

Notera slutligen de båda molnhallarna i detaljbilden ovan. I skrivande stund finns bara hallen i Akalla, varför molntjänsterna för närvarande saknar redundans. En redundant hall håller dock på att inrättas hos en kommersiell aktör i Sköndal så att tjänsterna kan dubbleras. Notera dock att SUNET inte håller med en eventuell synkroniseringslänk mellan hallarna utan att det är upp till den enskilde att ordna en sådan via Internet. Samma regel som vanligt gäller: kör alltid med livrem och hängslen.

Låt oss emellertid inte förringa säkerheten i en Tier-3-hall. Sådana hallar är ganska vanliga idag. Tier-begreppet har tagits fram av amerikanska Uptime Institute och nivåerna 1–4 betecknar olika tillgänglighetsnivåer. Ju högre man stiger i graderna, desto dyrare blir det, för desto mera redundant utrustning får man köpa till och desto säkrare måste byggnaden och kanaliseringen utföras. Avsikten med Tier 3 är att ge 99,98% upptid eller strax under 2 timmars oplanerad nedtid per år och att det ska gå att utföra underhåll och utbyten av infrastrukturen utan att anläggningen behöver tas ur drift. Uptime Institute kallar det för "concurrently maintainable". Det ska finnas redundant leverans av kraft och kyla.

Vi har snuddat vid elkraftproblemet tidigare i artikeln. Redundant kraft är alltid en god idé. Det kommer att bli strömavbrott. Dubbel spänningsmatning till serverna är också en god idé, för nätaggregat i serverar kommer att gå sönder.



Så här bygger man i princip upp ett elsystem för 400 volt trefas för en anläggning på Tier 3-nivå. Du ser att den är uppbyggd med en UPS-skyddad vänstersida (Critical-A) och en direktmatad högersida (Critical-B), samt direktmatning till icke-kritiska laster som belysning (Non-critical). Dessutom ser du att man kan utföra diverse korskopplingar för att dirigera om kraft om något eller några ställverk skulle gå sönder eller behöva underhåll.

Normalförbrukningen i denna hypotetiska anläggning är 300 kW. Det hela matas av två inkommande kraftledningar, respektive två reservgeneratorer på 450 kW stycket. UPS:erna är konfigurerade som N+1. Man har alltså två UPS:er på 150 kW och en tredje för redundans. Varje rackskåp (IT Load) och därmed varje server, får kraft från två håll och är redundant anslutna till varsitt ben av kraftmatningen. Det sitter två elfördelare (PDU) i varje rackskåp och varje server förutsätts ha dubbla, redundanta nätaggregat som kan bytas under drift.

Du kanske funderar på varför man inte har UPS på det högra kraftbenet också, som nu bara matas direkt från elnätet? Skulle man göra det (och en del annat), skulle man hamna på Tier 4-nivå, med 99,995% tillgänglighet och det var inte avsikten.

## AVSLUTNING



*This is our Redundancy Day!*

SUNET har satsat så hårt på redundans att SunetC i skrivande stund är ett av de bästa och modernaste näten i världen. Sedan SunetC kom igång 2016 har man stängt av den ena av routrarna R1 på Tulegatan då och då för att exempelvis byta ut delar, utan att någon någonsin har märkt att det hänt!

Satsa på redundans du också. Ingenting är någonsin helt säkert. Att stora molnlösningar ute i det kommersiella samhället stannar titt som tätt är något vi rycker på axlarna åt några veckor senare när systemen är igång igen.

Oavsett hur mycket redundans du har, fritar det dig inte från bruket av sunt förnuft. Att ha sin säkerhetskopia i andra änden av en hundra kilometer lång ledning är inte ägnat att förbättra nattsömnen. Den säkerhetskopia du inte kan ta på, den har du inte.

## LÄS MER

Om ROADM och Adva: <https://www.sunet.se/blogg/we-have-liftoff-del-2-av-2/>

Se driftläget SUNET: <http://stats.sunet.se/stat-q/load-map/SunetC-core,,traffic,peak>

Se driftläget NORDUnet: <http://stats.nordu.net/stat-q/load-map/ndn-map,,traffic,peak>

Molntjänsterna: <https://www.sunet.se/blogg/sunet-in-i-molnet-det-har-far-du/>

Driftläget för alla tjänster får du av tjänstens respektive ansvarige.

Så bygger man en redundant datorhall: <https://techworld.idg.se/2.2524/1.649909/stadje-datacenter>

Allt om Tier-klasserna: <https://journal.uptimeinstitute.com/explaining-uptime-institutes-tier-classification-system/>



Skriven av



**JÖRGEN STÄDJE**

Jag heter Jörgen Städje och har skrivit om teknik  
och vetenskap sedan 1984. Friskt kopplat, hälften  
brunnet!