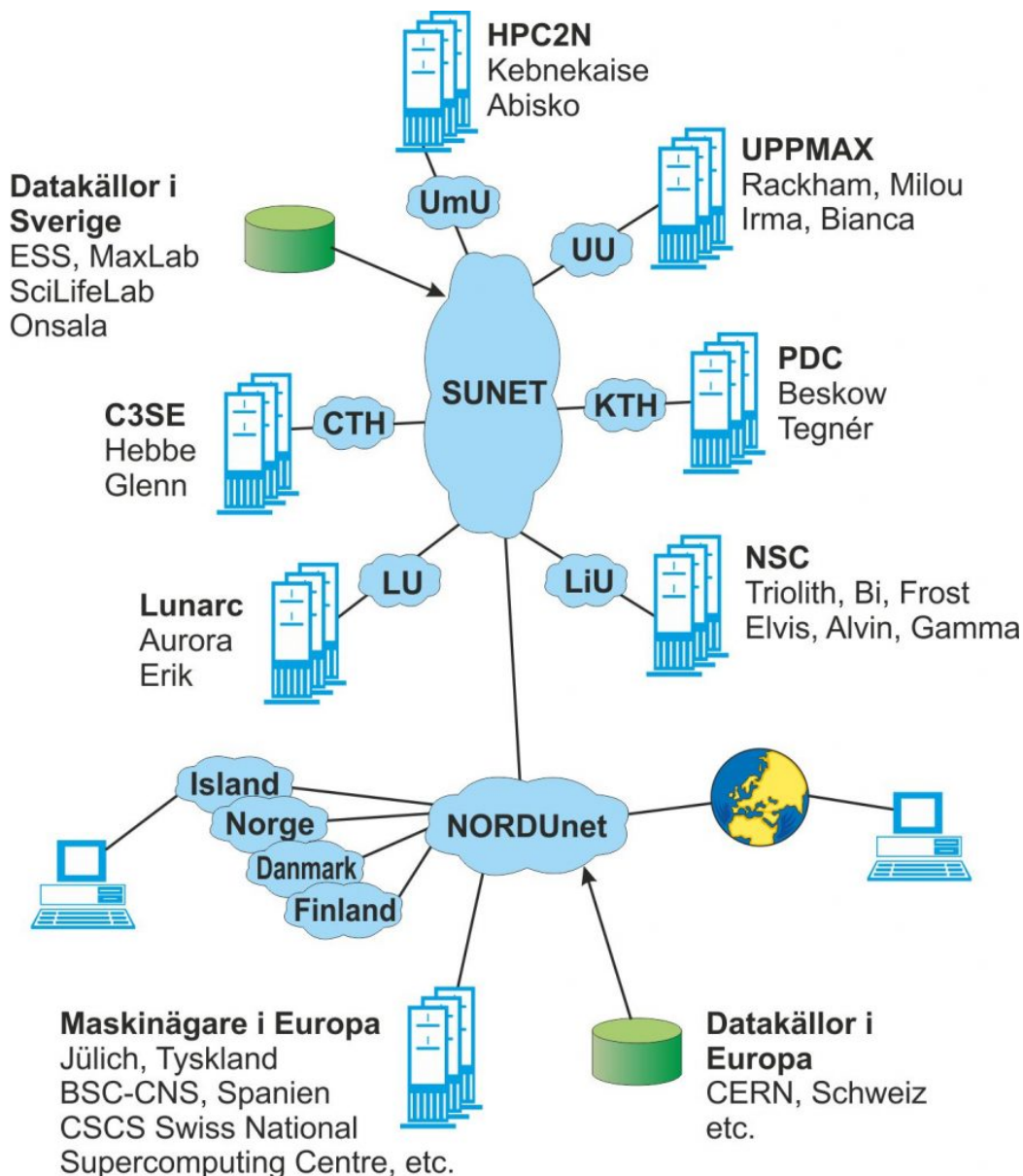


SNIC-SNACK

Sverige har ett antal datamaskintekniska stoltheter, bland dem superdatorerna som ingår i det svenska samarbetet SNIC. Datorerna står i sina datorhallar och användarna befinner sig på högskolor och universitet runt om i världen. Superdatorer kan producera oerhörda mängder utdata, så mycket att det inte skulle gå att transportera över ett datornätverk, inte ens över SunetC. Så, hur går det till när en superdator producerar data och hur överlämnas det till användaren?

Den svenska nationella infrastrukturen för databehandling (SNIC) är ett forskningssamarbete med avsikt att ge forskare från hela landet och inom alla ämnesråden balanserade och kostnadseffektiva beräkningsresurser, datalagring och användarstöd. SNIC är en del av Uppsala Universitet, lyder under Vetenskapsrådet och ska se till att det finns superdatorkraft till svensk akademisk forskning, både genom att hjälpa till ekonomiskt men också genom att skapa effektivt samarbete. Resurserna tillhandahålls efter en öppen ansökningsprocess, för att stödja den bästa svenska forskningen.



SNIC är sammansatt av sex samarbetsparter: Chalmers (med datorcentralen C3SE), Kungliga Tekniska Högskolan (PDC), Linköpings Universitet (NSC), Lunds Universitet (Lunarc), Umeå Universitet (HPC2N) och Uppsala Universitet (UPPMAX). En av

SNICs uppgifter är att se till att det alltid finns olika typer av datakraft tillgänglig och att de stora ägarna, som Stockholm och Linköping, inte byter maskinpark samtidigt, utan att det sker överlappande.

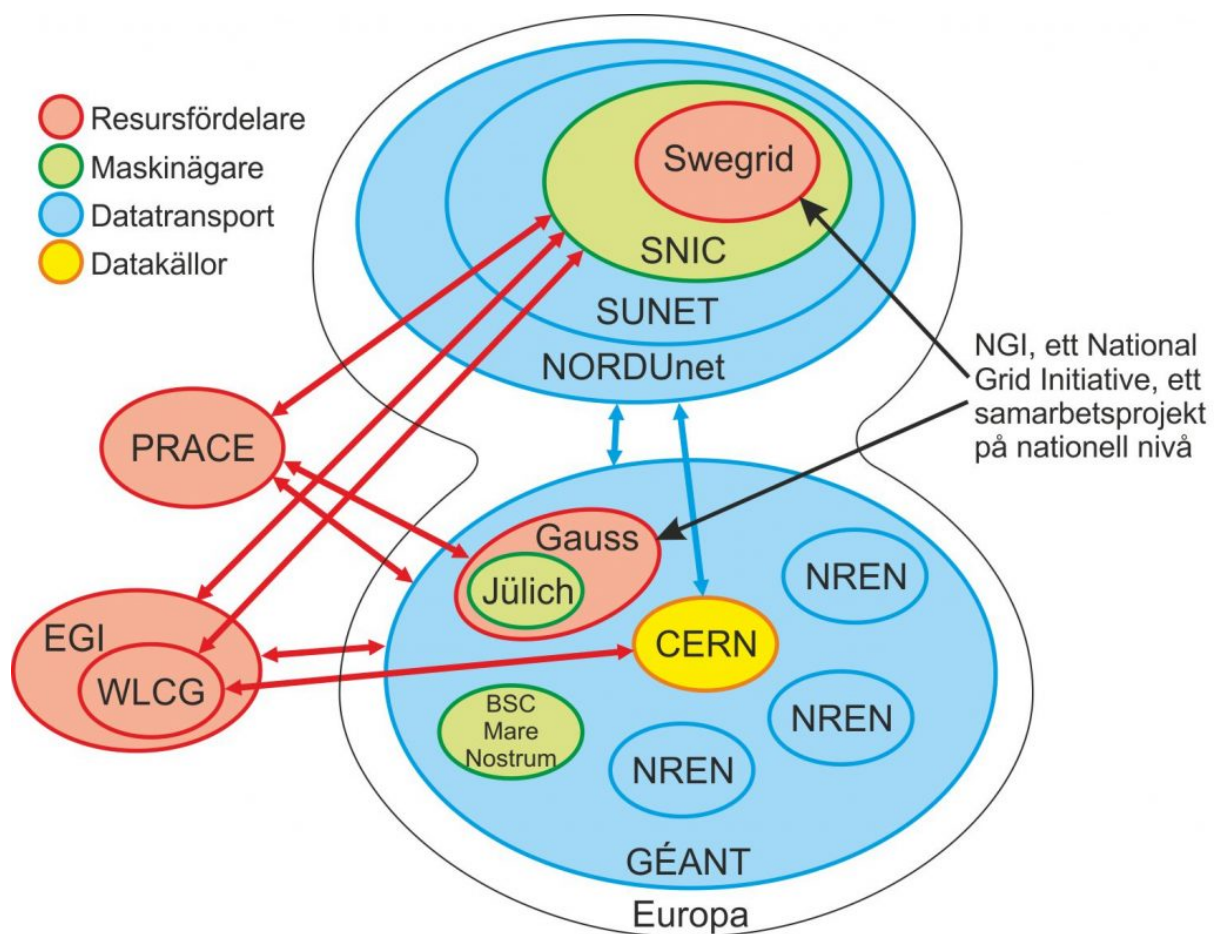
Superdatorerna är inte anslutna direkt till SunetC utan går genom lärosätenas intranät. Det kan de göra i och med att de enorma datamängderna som superdatorerna producerar, regelmässigt krymps genom efterbehandling i en lokal efterbehandlingsmaskin. Annars skulle dessa inte gå att transportera.

Användarna kan finnas var som helst i hela världen, inte alls bara i Norden. Detsamma gäller datakällorna. CERN är visserligen ansluten till Norden men besprutar i stort sett hela världen med sitt partikeldata. När ESS i Lund blir klar kan man räkna med att dess mätdata kommer att spridas över hela världen på samma sätt.

I Norge går det annorlunda till. Där är superdatorverksamheten organiserad som ett bolag kallat Uninett Sigma2, med tre centra. I Finland finns bara ett centrum, CSC i Helsingfors, som begär pengar direkt från staten utan någon mellanliggande samarbetsorganisation. Maskinen i Finland heter givetvis Sisu.

EUROPEISKT SAMARBETE

För att en superdator ska kunna användas till något praktiskt behövs nätverkskopplingar, dels inom det egna lärosätet, dels ut i världen. Dessutom behövs forskarsamarbeten med resten av världen.



Nu börjar organisationer och förkortningar hagla: Swegrid, National Grid Initiative (NGI), European Grid Infrastructure (EGI), Worldwide LHC Computing Grid (WLCG), LHCOPI (CERN), NeIC (Nordic e-Infrastructure Collaboration), EU-projektet PRACE (Partnership for Advanced Computing in Europe), med flera.

Det finns fyra olika klasser av aktörer i superdatorvärlden och många av dem har flera roller.

Centralt för superdatorintresserade är **maskinägare**. De enskilda lärosätena är de verkliga maskinägarna medan SNIC ordnar finansieringen av maskinvaran som andra får exekvera på.

För att nå maskinvaran behöver användarna olika **transportstrukturer**, som SunetC, som i sin tur arbetar mot NORDUnet, som i sin tur har anslutningar till transportnätet GÉANT och andra nationella forskningsnät (NREN) i Europa.

Eftersom en superdator är så dyr, kan den inte stå sysslös. Därför finns det en mängd organisationer som agerar **resursfördelare** (brokers) av datakraft, som meddelar när det finns ledig körtid på någon maskin som är med i just det speciella samarbetet, ett sk grid-samarbete. Olika superdatorer har ofta delar av maskinen eller delar av körtiden reserverad för olika fördelare. SNIC driver Swegrid som är en sådan fördelarorganisation som arbetar med att köa och fördela körtid inom delar av SNIC, medan exempelvis PRACE fördelar körtid mellan europeiska maskiner. Fördelningen kan gå till på olika sätt. I ett gridsamarbete sker fördelning av data och program automatiskt, medan man i ett samarbete som PRACE ansöker om tid och själv får se till att överföra data till anvisad superdator. SNIC motsvaras i Tyskland av Gauss, som också är medlem i PRACE och fördelar datorkraft från de tre tyska centren i Jülich, Stuttgart och Leibniz, alla tre med maskiner väl över en petaflop. Det finns ännu fler, som NeIC (Nordic e-Infrastructure Collaboration) som fördelar resurser och driver samarbeten på nordisk nivå.

Slutligen finns det rena **datakällor** som inte tar emot indata, utan bara åstadkommer mängder med utdata som måste behandlas i superdatorer. CERN i Schweiz är en sådan, som skapar många gigabyte per sekund som måste hanteras av datorer över hela världen och därför har sin egen resursfördelare, kallad WLCG (Worldwide LHC Computing Grid). Se vidare avsnittet **En stor dataspruta**, nedan. Den svenska neutronkanonen ESS kommer också att bli en sådan och acceleraterringarna MaxLab och Max IV är det redan.

Fördelarorganisationerna kan ha ytterligare uppgifter, såväl ekonomiska som administrativa, exempelvis att skaffa EU-pengar för utbyggnad av det europeiska superdatorsamarbetet, forskning kring nya maskintyper, upphandlingar osv. En annan viktig gren av samarbetet är erfarenhetsutbytet kring maskinvara, parallellprogrammering och effektiv utbyggnad av datorhallar. Det finns betydligt flera sådana organisationer än vad diagrammet visar och det kan därför bara bli översiktligt.

ANVÄNDARE NÄR OCH FJÄRRAN

Långt ifrån alla användare på en viss superdator befinner sig på universitetet där maskinen står, och kan använda den över intranätet. Det är här SunetC kommer in.

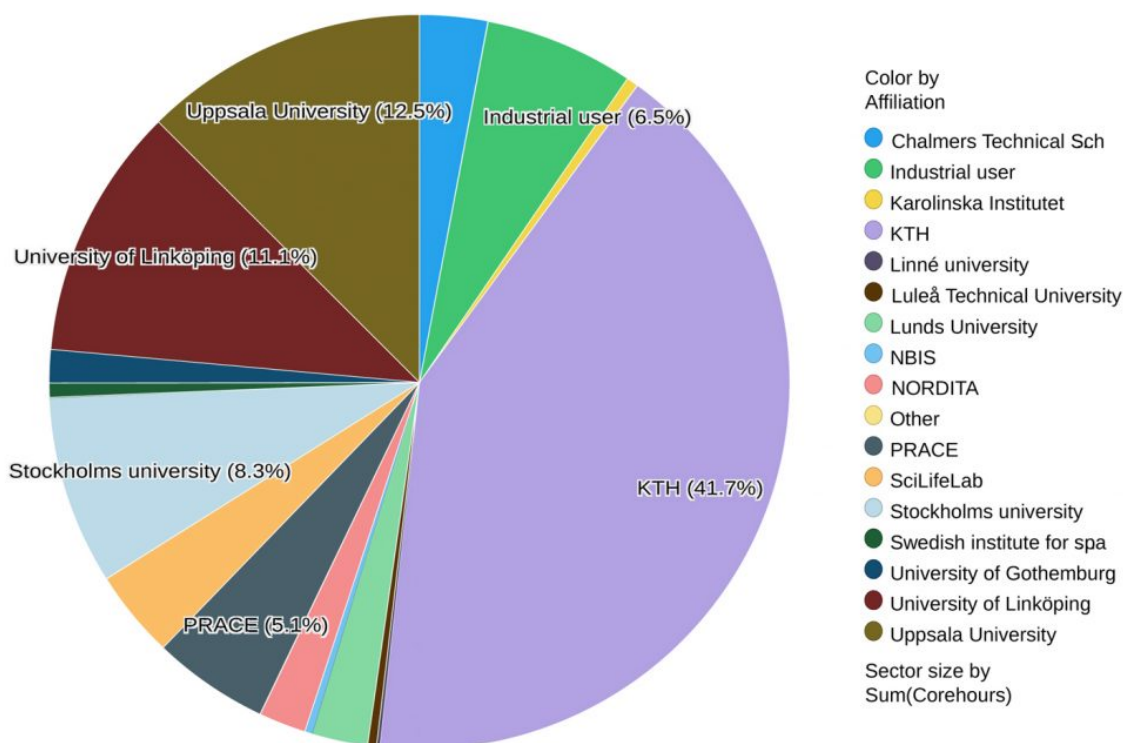


Bild: Henric Zazzi, KTH

Diagrammet visar utnyttjandet av maskinen Beskow på PDC på KTH. 41 procent av maskinkraften (core hours) används på KTH och resterande av externa användare. Som du ser flyter dataströmmarna ut från KTH till Stockholms Universitet, Linköpings Universitet och Uppsala Universitet. Den privata industrin (Industrial user), som delvis är med och finansierar maskinerna, får förbruka sin del av körtiden. Scania i Södertälje är exempel på en sådan användare. Du ser också att fördelarorganisationen PRACE har fått 5,1 % av körtiden.

MASKINERNA OCH TRAFIKEN

De sex svenska datorcentralerna inrymmer vardera ett antal superdatorer, eller High Performance Clusters (HPC). En modern superdator byggs inte upp som en enda jätte-processor utan är sammansatt av en stor mängd (kluster av) standardserverar (beräkningsnoder), som skalats av in på bara kretskortet på allt som inte behövs för att de ska kunna räkna, räkna, räkna!

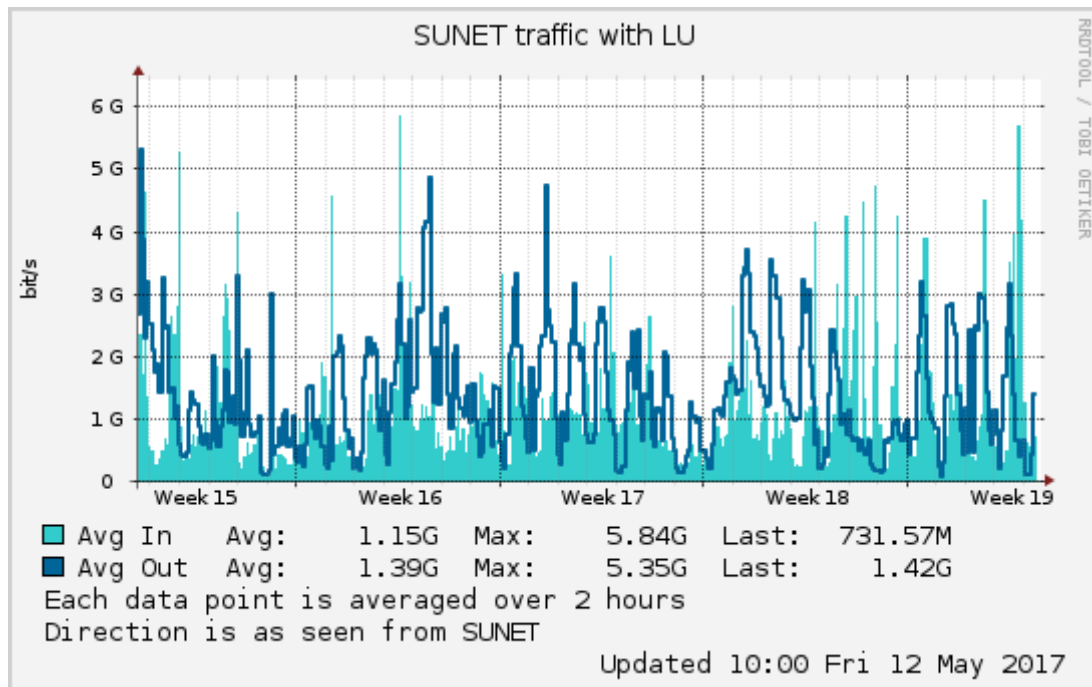
Hårddiskar finns oftast inte, ej heller nätaggregat, lokal redundans och en massa annat tingeltangel som återfinns i en bordsdator. Redundanta diskar och strömförsörjning finns istället centralt. Däremot har serverarna många kärnor och massor med primärminne. Superdatorn sätts samman av så många noder man har råd med och när kapaciteten behöver utökas, skaffar man ännu flera noder. Noderna kopplas samman med ett internt datornät som är betydligt snabbare och med kortare fördröjning än Ethernet. Det som gör beräkningarna snabba är dock inte mängden serverar som klarar massor av teraflops i sig själva, utan programvarans förmåga att arbeta parallellt. Det är där de stora forskningsinsatserna måste sättas in.

Maskinvaran finns definitivt och är till på köpet väldigt imponerande. Nog finns det beräkningsresurser i Sverige, minsann!

Aurora på Lunds universitet

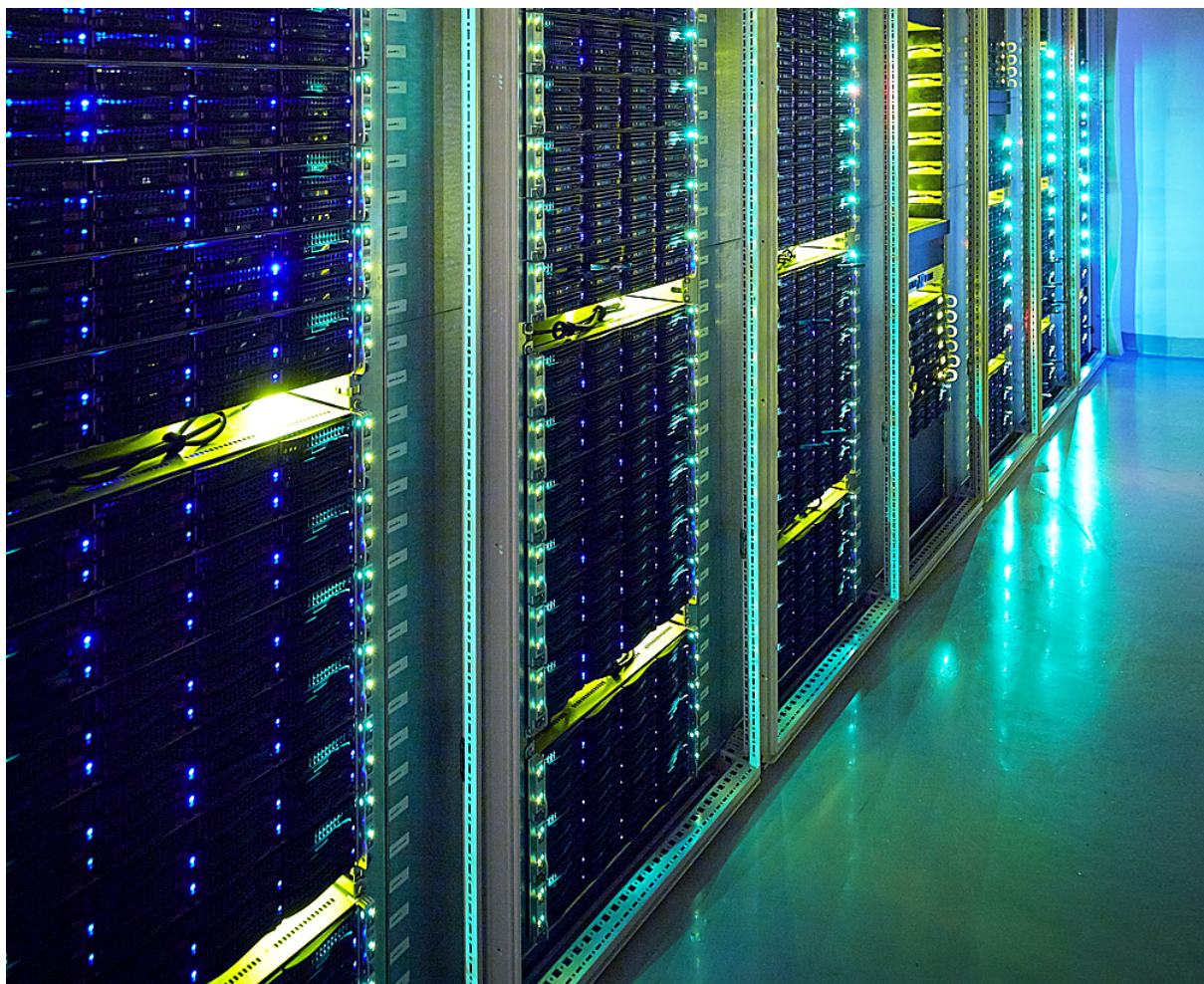


Aurora i Lunds universitets datorcentral Lunarc (Lund University NIC Application Research Center, där NIC i sin tur står för Numeric Intensive Computation) har 180 beräkningsnoder, med totalt 3600 kärnor och sammanlagt 11,5 TB primärminne, sammankopplade med Infiniband-nät.



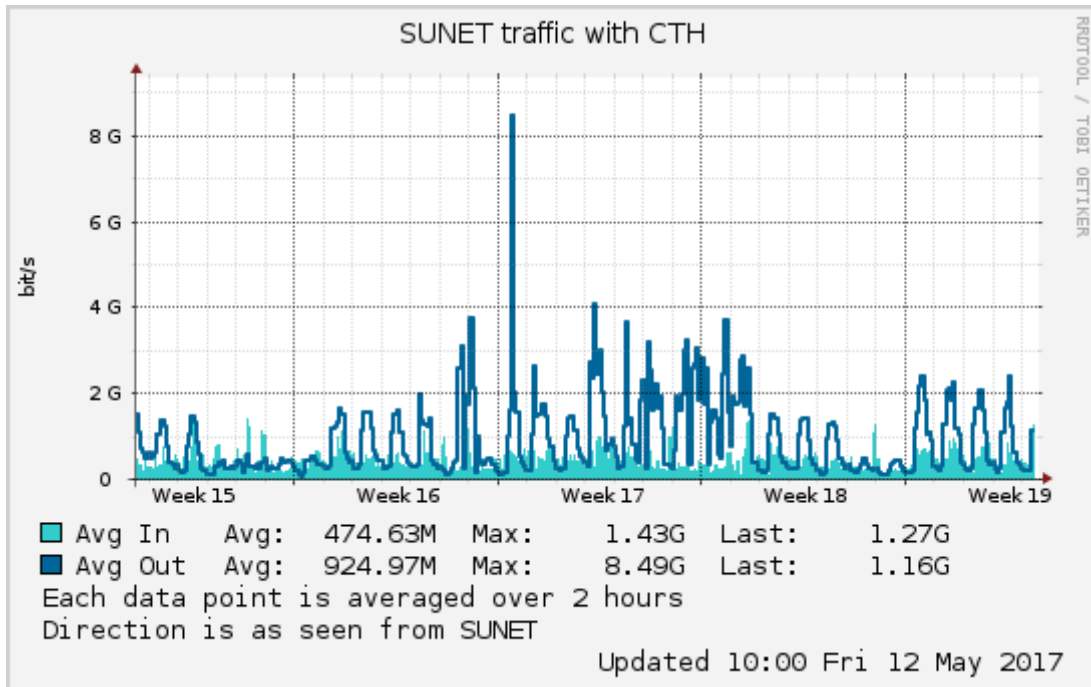
Trafiken från och till Lunds universitet och därmed Aurora den senaste månaden ligger kring 1,4 Gbps och toppar strax under 6 Gbps. Observera att man inte kan skilja datatrafiken från superdatorerna från all annan trafik från och till universitetet. Diagrammet är bara en översikt. Det är gott om rum i fibern på 100 Gbps.

Hebbe på Chalmers



Hebbe finns i Chalmers datorcentral C3SE (Chalmers Centre for Computational Science and Engineering). Hebbe har 315 beräkningsnoder av HPs fabrikat, med totalt 6300 kärnor och sammanlagt 26 TB primärminne. Noderna sammanbinds med 56 Gbps Infiniband-nät.

Namnet Hebbe syftar på den göteborgske skrotsamlaren Hebbe i TV-serien Albert och Herbert som var väldigt populär mellan 1974 och 1981.



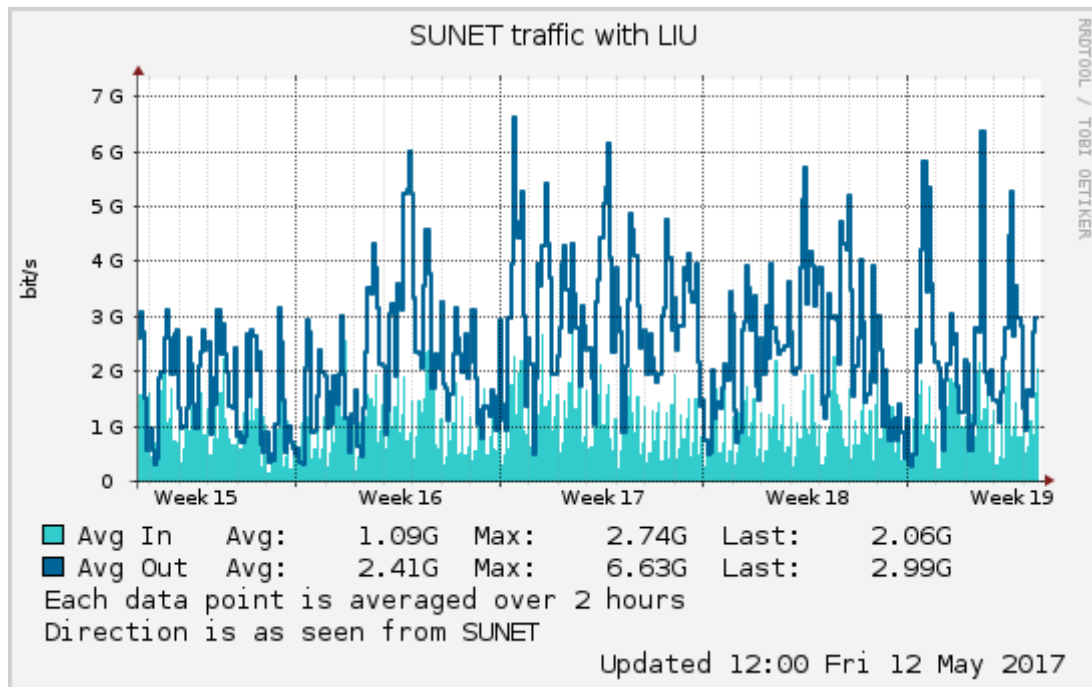
Trafiken från och till Chalmers och därmed C3SE de senaste 31 dagarna är i medeltal strax under 1 Gbps men hade en märklig topp på 8 Gbps vecka 17.

Triolith på Linköpings universitet



NSC (National Supercomputer Centre) i Linköping har landets största samling superdatorer, med namn som Triolith, Elvis, Bi, Frost, Alvin och Gamma. Flaggskeppet är Triolith med sina 16.368 kärnor i HP-servrar med totalt 35 TB primärminne, hopkopplade med Infiniband.

Fattar du? Bi+Frost, haha, bron till himmelriket! Namnkombinationen Bi+frost är särskilt rolig eftersom maskinerna huvudsakligen används av SMHI i Sverige och MET i Norge för meteorologiska beräkningar. Bifrost är också namnet på en Linux-distribution för routrar och brandväggar.



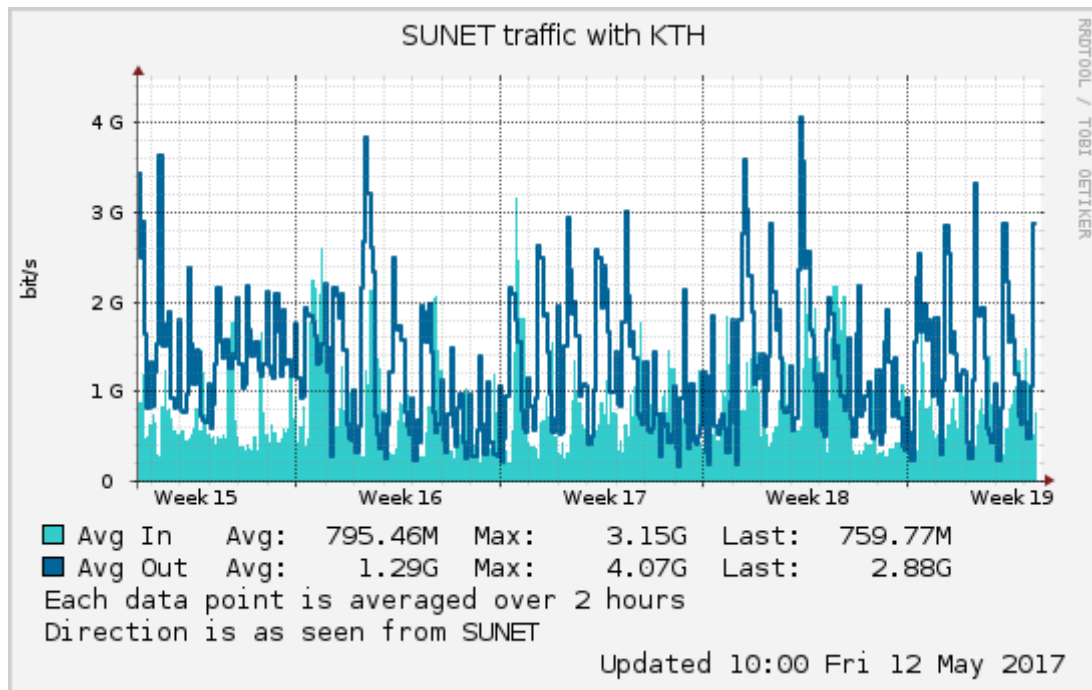
Linköpings universitet uppvisar mera trafik än de två tidigare. Medeltrafiken ligger kring 2,4 Gbps med toppar på 6,6 Gbps. Det är kanske naturligt att universitetet med landets största superdatorcentrum har den högsta medeltrafiken i SNIC?

Beskow på KTH



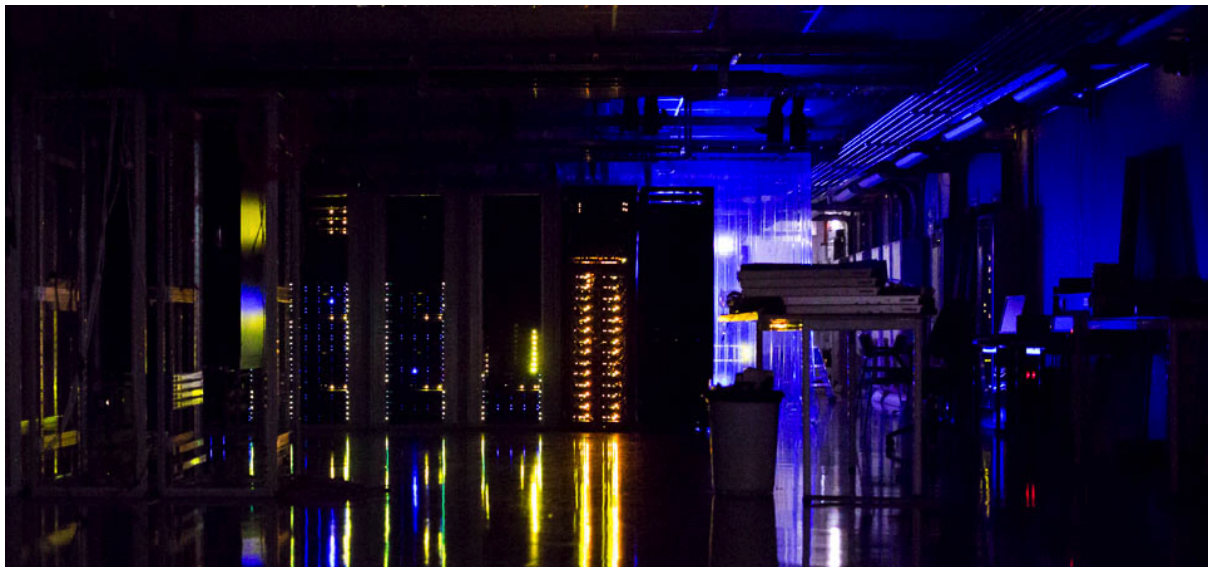
Maskinen Beskow på Paralleldatorcentrum på KTH i Stockholm är tillverkad av Cray. Den har 1676 noder med 53.632 kärnor, med totalt 104,7 TB primärminne som sammanbinds med Crays nätverk av typen Cray Aries.

PDC döper sina maskiner efter svenska författare och konstnärer, som Beskow (Elsa), Tegnér (Alice), Lindgren (Astrid), Zorn, Ferlin, Povel, Bellman osv.



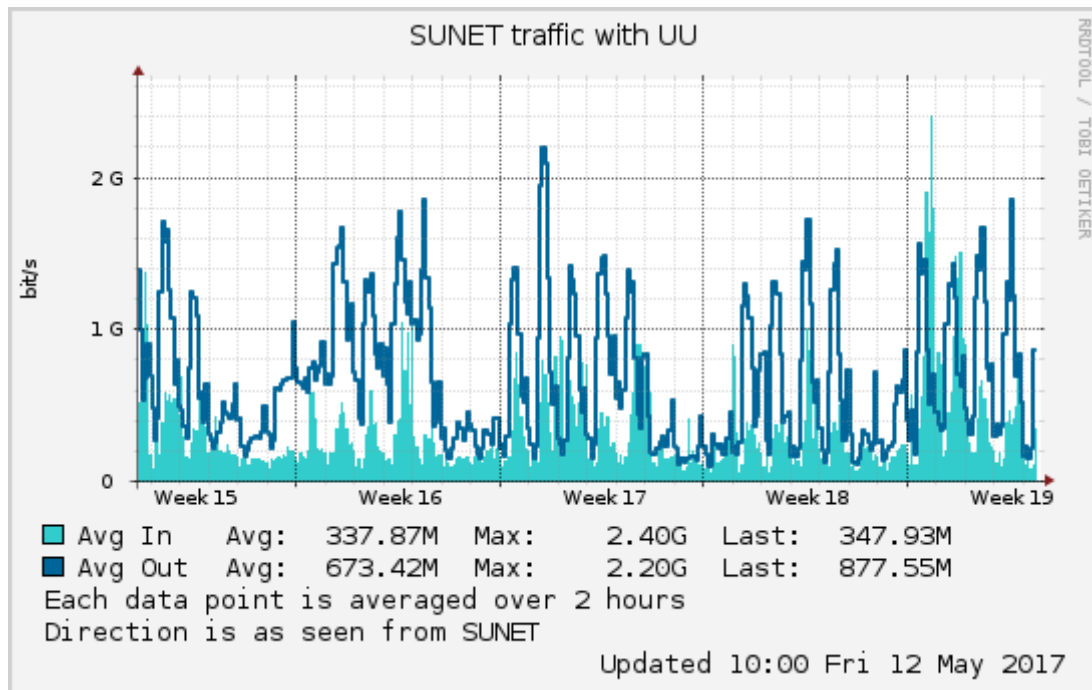
Tekniska högskolan presterar en medeltrafik på 1,3 Gbps, med toppar upp till 4 Gbps.

Rackham på Uppsala universitet



Maskinen Rackham vid UPPMAX (Uppsala Multidisciplinary Center for Advanced Computational Science) har 304 noder med totalt 6080 kärnor med 34,7 TB primärminne och noderna kopplas som brukligt samman med Infiniband.

Maskinerna i Uppsala har fått sina namn ur den tecknade serien Tintin, med namn som Rackam (den Röde, piratkapten), Milou (Tintins hund), Bianca (Castafiore, irriterande operasångerska) och Irma (Castafiores betjänt).



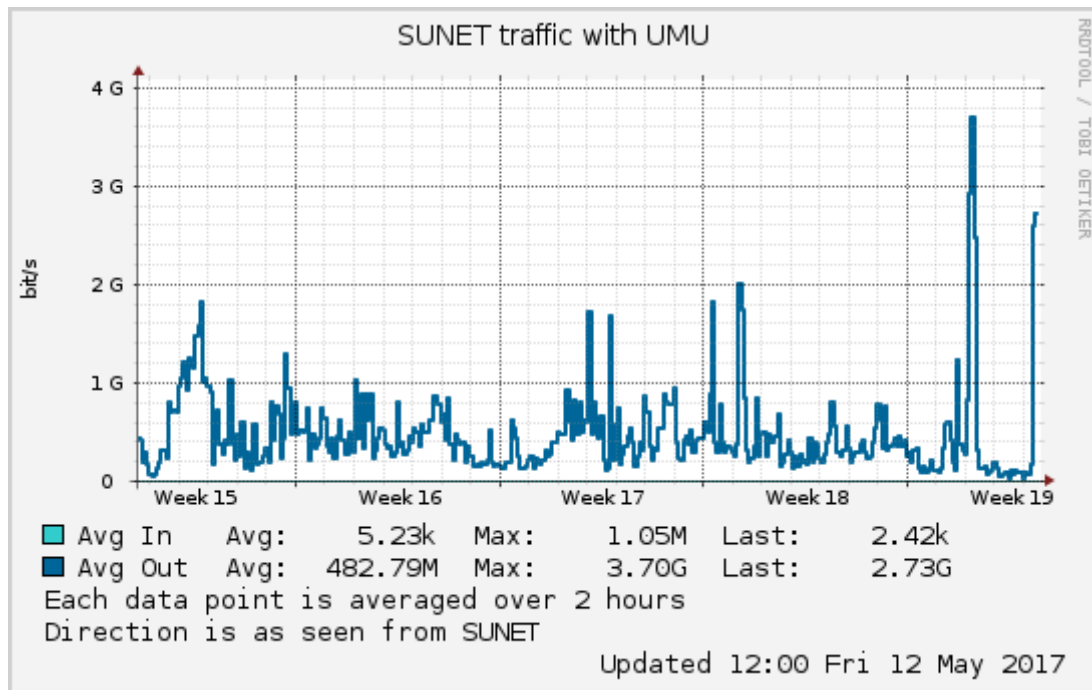
Uppsala universitet och UPPMAX har en medeltrafik på 0,7 Gbps och toppar uppåt 2,4 Gbps.

Kebnekaise vid Umeå universitet



Maskinen Kebnekaise vid HPC2N (High Performance Computing Center North) har 544 noder från Lenovo, med totalt 17.552 kärnor och 125 TB primärminne och kopplas givetvis ihop med Infiniband-nät.

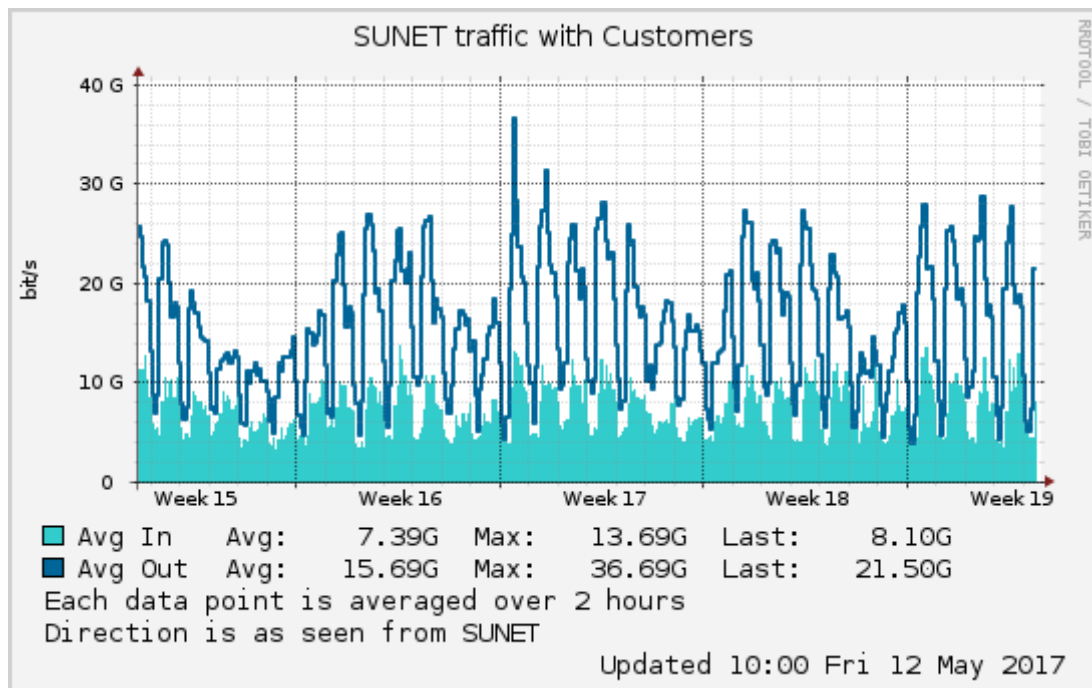
Maskinerna i Umeå har namn efter norrländska naturformationer som Kebnekaise (fjäll) och Abisko (nationalpark).



Trafiken till och från Umeå universitet ligger i medeltal på 0,5 Gbps och toppar på omkring 2 Gbps, med en enskild topp på 3,7 Gbps.

Totalt i landet

Det sista diagrammet visar en summering av all Sunets trafik på alla portar på alla routrar ute i landet.

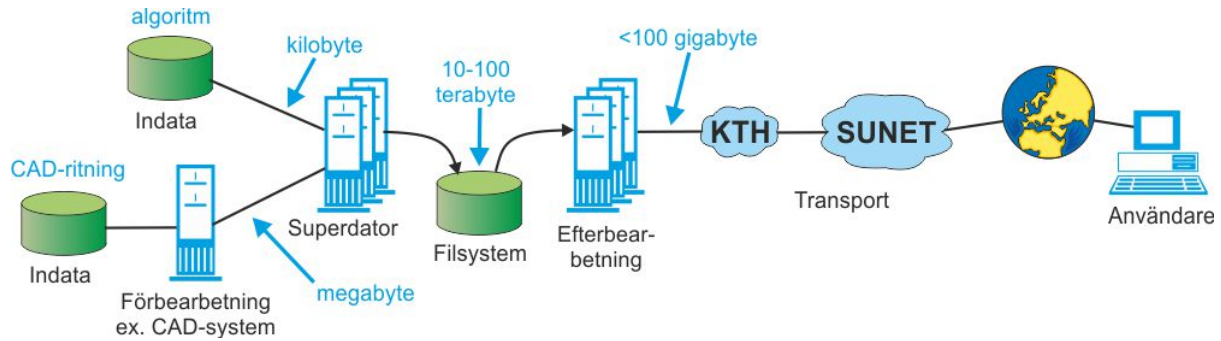


Diagrammet säger egentligen inget alls om något, utom att det ännu är långt kvar tills något universitet lyckts fylla en våglängd på SunetC till 100 Gbps. Vidare säger det att Sunet än så länge är väl tilltaget för att klara topptrafiken. När det börjar bli trångt finns det utmärkta utbyggnadsmöjligheter.

OM SUPERDATORBERÄKNINGAR

När man löser ett problem i en superdator, blir resultatet ofta ett sk data set, en väldig datamängd, som forskaren ska utvärdera för att få fram ett resultat som är begripligt för människan. Det handlar som vi sagt tidigare om enorma

datamängder, oftast allt för stora för att föra över på ett datornät i rimlig tid. Eftersom det blir så enorma datamängder försöker man göra visualiseringen eller vad det kan vara, inom datacentrets väggar, för att slippa transportera terabyte på SunetC.



Internt hos Paralleldatorcentrum på KTH

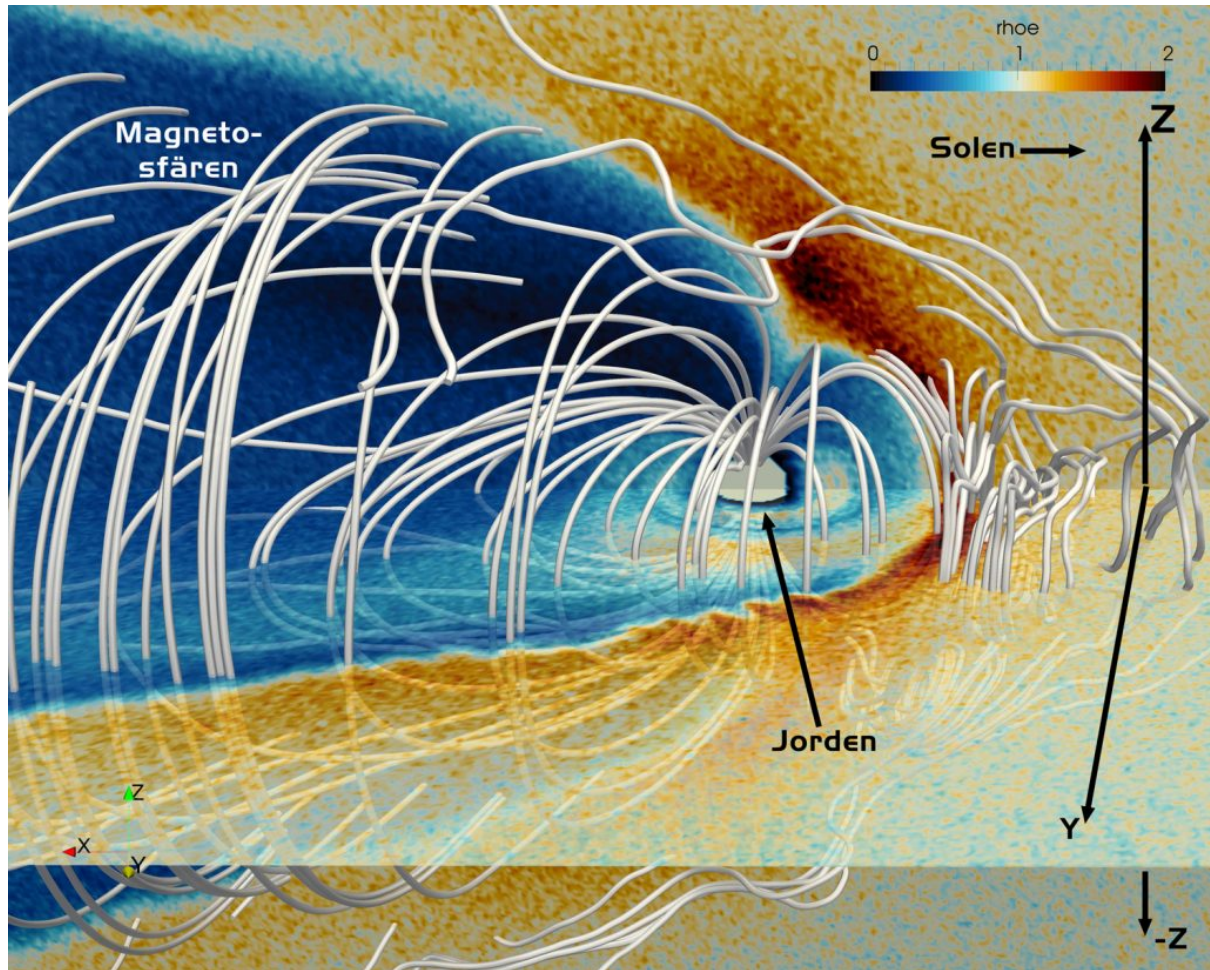
Från bara en liten fil med en algoritm på några kilobyte som indata kan en superdator mata ut data set på 10–100 terabyte. Data setet efterbehandlas i en lokal dator med något mindre kraft och blir till bilder eller data som kan förstås av en människa. Datareduktionen kan bli omkring 1000 gånger eller mer, och resulterar i högupplösta bilder eller en film. Det är särskilt användbart när man studerar flöden och virvlar, oavsett om det rör sig om friktion kring vingprofiler eller flöden av elektroner i solsystemet.

En annan typ av indata som är betydligt omfångsrikare är exempelvis CAD-modeller av vingprofiler, som måste snittas upp eller delas in "griddar", som superdatorns noder kan arbeta med.

Ett typiskt superdatorproblem kan köras på 16.000 kärnor i satser om 6 timmar. Ungefär en gång i timmen tar man ut en datadump på cirka 100 gigabyte som går vidare till efterbehandling och visualisering. När den tilldelade körtiden om 6 timmar är slut tar man en totaldump på ett par terabyte, kallad "restart file", som används senare när körningen ska återstartas och beräkningarna fortsätta. Slutresultatet av körningen är serier av bilder på cirka 100 megabyte stycket från visualiseringsdatorn.

SOLVINDEN BLÅSER PÅ JORDEN

Fysikern Stefano Markidis vid KTH är en av dem som använder superdatorer i sin forskning. Till exempel då man tar miljarder elektroner (alltså: ett plasma) och skickar dem i en våg från Solen mot Jorden, och beräknar hur var och en samverkar med övriga partiklar, stör Jordens magnetfält och antingen störtar ned mot planetytan eller virvlar bort i magnetfältets svans (magnetotail) bortom planeten.



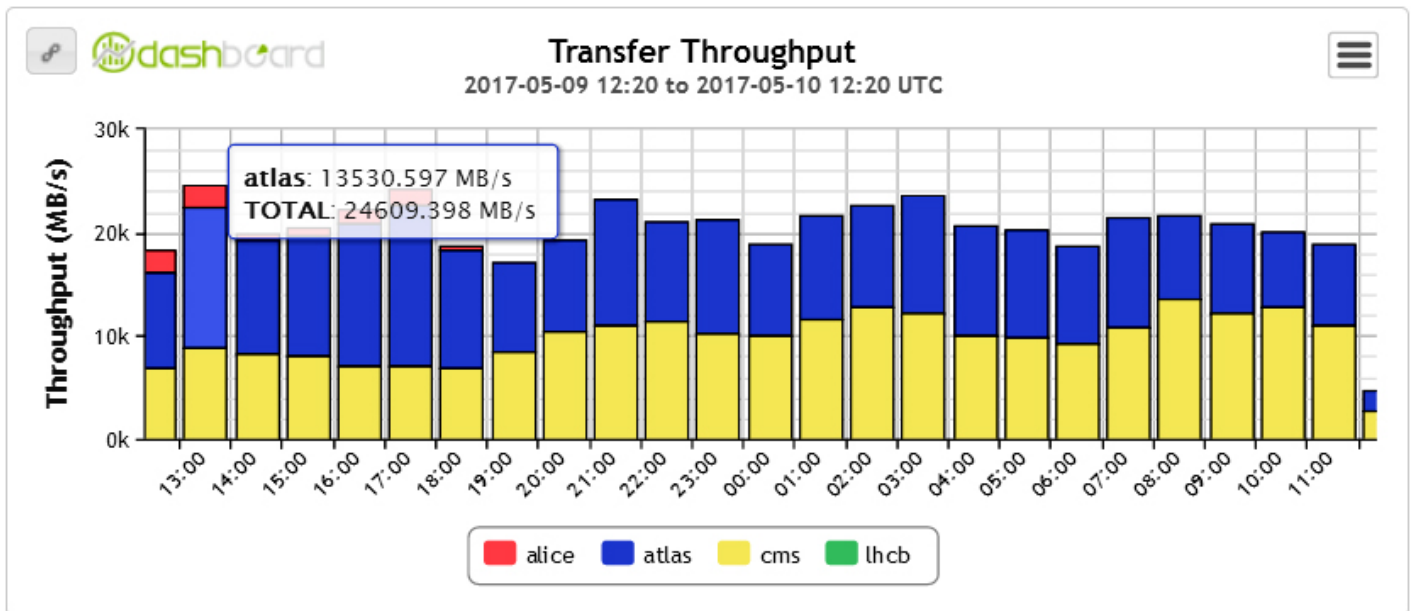
Låt oss titta på ett projekt med namnet iPIC3D i teoretisk rymdfysik som kräver superdatorkraft. Markidis har beräknat hur solvinden samverkar med Jordens magnetosfär, i avsikt att utvärdera en metod för att algoritmiskt bestämma sådan samverkan. Solvinden, i form av miljoner ampere kommer från höger och träffar det jordmagnetiska fältet och komprimerar och distorderar det på Jordens dagsida (höger) och drar ut det på nattsidan. De grå slangarna är de magnetiska fältlinjerna mellan Jordens nord- och sydpol och färgerna anger elektrontätheten, eller den elektriska strömmen om man så vill. Blått är låg ström och rött är hög, och man ser hur elektronerna trängs ihop framför Jorden. Men det jordmagnetiska fältet tränger undan elektronerna, bildar magnetosfären och hindrar därmed solvinden att riva med sig vår atmosfär. Sådan tur har inte marsianerna.

Körningen utfördes som en simulerad kub med totalt 3×10^9 partiklar och tog 24 timmar att utföra på KTHs superdator Beskow. Totalt användes 2048 kärnor och på 24 timmar ackumulerades 50.000 core-hours. Denna körning är att betrakta som ganska liten. Samma problem provades på superdatorn Mira vid Argonne National Laboratory i USA. Mira är kapabel till 10 petaflops på sina 786.432 kärnor och 768 terabyte primärminne.

Indata till Beskow var i storleksordningen ett par kilobyte i form av en beskrivning av geometrin i simulationen och hur problemet skulle spridas ut över de tillgängliga kärnorna. Körningen på KTH försiggick i fyra satsar om sex timmar och varje sats resulterade i 40 TB data som måste sparas och återanvändas i nästa sats. Data i primärminnet under körningen låg kring 50-100 TB. Allt data som skapades under processen, alltså 50-100 TB måste arkiveras, såväl som slutresultatet i form av bilder och filmsekvenser, som uppgick till 100-tals megabyte.

EN STOR DATASPRUTA

Som vi tidigare nämnt kan indata också komma från en dataleverantör. En mycket stor dataproducent är CERN och dess partikeldetektorer Atlas och CMS på acceleratoringen Large Hadron Collider i Meyrin i Schweiz.

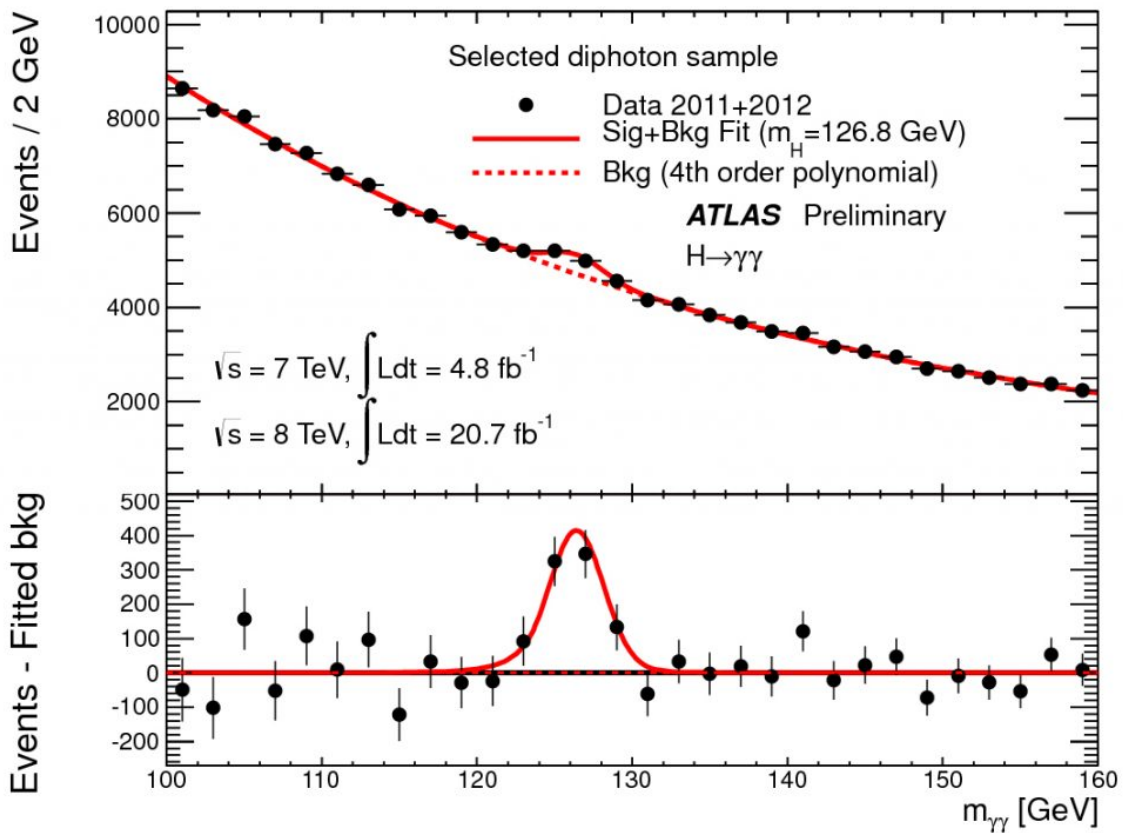


Här ser du dataflödet ut ur detektorerna när artikeln skrevs. Den infällda skylten visar att flödet ur Atlas-detektorn en dag var dryga 13 gigabyte per sekund och totalflödet från hela LHC-ringen var dryga 24 GBps. Gigabyte, inte gigabit. Så där håller det på, dygnet runt, året runt. Och det är efter bortfiltrering av ointressant känd fysik och efter datareduktion. Allt detta ska kontinuerligt transporteras ut i Europa och en del av det slutar i SunetC och på svenska lärosäten. Förhoppningsvis ska det resultera i mera information om mörk materia och higgspartiklar.

CERN är en del av ett globalt beräkningsprojekt kallat Worldwide LHC Computing Grid (WLCG) som fördelar körtid mellan fler än 170 beräkningscentraler i 42 länder som binds samman av ett antal internationella nätverk. Avsikten med WLCG är att kunna lagra, distribuera och analysera de omkring 50 petabyte data som förväntas komma ut ur LHC under 2017. CERN släppte nyligen ett enormt data set på 300 terabyte som innehåller data om partikelkollisioner och analyser av dessa, med ursprung i CMS-detektorn (gul stapel i bilden ovan). Med den nya linjäracceleratoren Linac 4 som blir klar år 2025 kommer dataflödet från LHC att öka tio gånger. Som CERN själva säger "The world is our calculator". Men SUNET kommer inte att ha några som helst problem att transportera de ökande datamängderna.

HIGGSPARTIKELN

Flera års körning på ett antal datorer över hela världen och petabyte med rådata som färdats från CERNs datorhall i Meyrin runt hela Jorden från Tokyo till New York, och så händer det!



Minsann! Higgspartiklar på 125 GeV, precis som Peter Higgs hade förutsagt. En topp av signaler höjer sig ur brusets och förändras från tanke till statistiskt signifikant med en säkerhet på 6 sigma eller en miljard gånger högre sannolikhet än slumpen. Higgspartikeln lever bara i 0,000000000000000000001 sekunder, men det räcker för att man ska kunna hitta den. Ett hål till i standardmodellen är fyllt och vi vet numera varför vi väger något och varför Universum faktiskt kunde skapas, få massa och inte bara flyga isär i ljusets hastighet. Nu är siktet inställt på den mörka materien.

AVSLUTNING

Forskare, ingenjörer, humanister och socialvetare i hela Sverige – många av dem vid universitet, högskolor och forskningsinstitut – använder avancerade digitala resurser och tjänster i sitt arbete varje dag. Saker som superdatorer, datasamlingar och nya programverktyg är avgörande för forskarnas framgång, som använder dem för att göra våra liv hälsosammare, säkrare och bättre. Därför finns SNIC – Swedish National Infrastructure for Computing.

LÄS MER

Om SNIC: <http://www.snic.vr.se/>

Datorcentralen vid Lunds universitet: Lunarc: <http://www.lunarc.lu.se>

Datorcentralen vid Chalmers: C3SE: <http://www.c3se.chalmers.se>

Datorcentralen vid Linköpings universitet: NSC: <http://www.nsc.liu.se>

Datorcentralen vid KTH: PDC: <http://www.pdc.kth.se>

Datorcentralen vid Uppsala universitet: UPPMAX: <http://www.uppmax.uu.se>

Datorcentralen vid Umeå universitet: HPC2N: <http://www.hpc2n.umu.se>

Så designar man en superdator: <http://techworld.idg.se/2.2524/1.486785/sa-bygger-kth-superdatorer>

Så designar Lenovo superdatorer: <http://www.sweclockers.com/artikel/20368-en-superdator-fran-grunden>

Så gör CERN i Meyrin: <https://www.sunet.se/blogg/cern-krossen-som-slar-sonder-materiens-minsta-byggstenar/>

SUNET och GÉANT: <https://www.sunet.se/blogg/geant-och-nordunet-basta-kompisar/>

Läs mer om den spännande rymdfysiken: http://www.qedata.se/idg-artiklar/voltmetare-i-verldrymden/med_voltmetare_i_verldrymden.htm

Skriven av



JÖRGEN STÄDJE

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