

SUNET mätserver för nätverksprestanda

Denna text är av rätt teknisk karaktär och riktar sig till användare med relativt god kännedom om dataöverföringar.

Det kan vara väldigt svårt att få bra överföringshastighet på dataöverföringar över långa distanser. Ganska ofta så är det inte nätet i sig som är begränsande, utan den sändande eller mottagande datorn. SUNET och NORDUnet har därför installerat servrar som kan användas för att testa sin nätverksprestanda över lite längre avstånd.

Bakgrund och räkneexempel

Överföring med TCP fungerar förenklat så att man skickar en mängd data och sedan förväntar man sig ett "mottaget OK" tillbaka innan man skickar mer data. Fram till att man fått detta OK så måste man spara en kopia på det data man sänt iväg, ifall att det man sände skulle försvinna på vägen och en omsändning skulle krävas. I väntan på "mottaget OK" så skickas inget data, och man använder därmed inte den kapacitet som finns tillgänglig. Om man har mer lagringsutrymme hos avsändaren, för att spara kopior av det man skickat, så kan man istället skicka flera datapaket på rad. För att använda all den tillgängliga kapaciteten gäller det att skicka så mycket data att det fyller hela förbindelsen innan man förväntar sig ett OK tillbaka.

Det tar t.ex. ca 5 ms för ett datapaket att ta sig 470 km från Stockholm till Göteborg. Om en speditorsfirma skulle köra lastbilar på samma sätt som TCP fungerar så skulle dom skicka iväg ett litet antal lastbilar med släp, låt säga 10st med en minuts mellanrum, och när lastbilarna kommer fram till målet efter 5 timmar så skickar man tillbaka en lastbil med en last som säger "OK - allt kom fram som det skulle - var snäll och skicka nästa omgång". Denna retur-bil kommer då inte fram förrän efter ytterligare 5 timmar, och först då kan nästa last skickas. Om man istället skickar en konstant ström med lastbilar från Stockholm till Göteborg så kommer man att ha skickat 600 lastbilar (60 lastbilar per timme och 10 timmar tur- och retur) innan den första retur-lastbilen kommer tillbaka och säger att det gick bra. Man måste alltså skicka 600 lastbilar innan man kan få en verifikation på att något har kommit fram och att det är OK att fortsätta skicka bilar. Dilemmat är att man hela tiden, för varje avsend lastbil, måste hålla en reserv-last kvar hos avsändaren, utifall att någon av bilarna skulle köra i diket – detta ställer alltså stora krav på lagringsutrymme hos avsändaren.

Om man ska kunna utnyttja 5 Gbps ledig kapacitet på denna sträcka så måste man skicka 6,25 Mbyte (50 Mbit = 5 Gbit per sekund med 5 ms dit och 5ms tillbaka) data innan man kan förvänta sig ett "mottaget OK". Dessa 6,25 Mbyte måste alltså lagras hos sändaren för att kunna återsändas, om så skulle behövas.

Väldigt få operativsystem är dock konfigurerade med såpass stora buffrar som standard, och kan därmed inte uppnå så höga dataöverföringshastigheter. Mer information om hur TCP fungerar och hur avståndet fungerar finns på:

<http://fasterdata.es.net/host-tuning/background/>

SUNETs mätserver

För att kunna utföra tester inom Sunet från anslutna nät så har vi installerat en mätserver i Luleå. Servern placerades i Luleå eftersom den då ligger ca 7 ms (enkel väg) från Sunets core-noder i Stockholm, vilket gör att alla högskolor och universitet (utom LTU) har minst 13 ms fördröjning dit (tur och retur). Ett längre avstånd gör det svårare att få bra överföringshastigheter mellan två maskiner. Om man kan få sin klient att få bra överföringshastighet till Luleå-servern så är det stor chans att överföringshastigheten till övriga värden ska vara skaplig. För att uppnå detta kan det dock krävas lite ändring av klientens TCP-inställningar (se nedan) och/eller översyn av det lokala nätverket.

Luleå-servern är utrustad med Internet2s "Perfsonar PS" distribution (se <http://psps.perfsonar.net/>). Denna programvara används bland annat i projektet LHCOPN (**Large Hadron Collider Optical Private Network**) för att kontinuerligt övervaka konnektiviteten mellan alla ingående sites. NORDUnet har för detta ändamål två liknande installationer i Köpenhamn. NORDUnet har också en mätserver i Stockholm med Geants distribution av perfsonar-mdm för BWCTL och OWAMP mätningar.

Servern i Luleå är direktansluten med 10GE till en av Sunets routrar, och klarar att generera data på den anslutningen med en paketstorlek på 9 kbyte. De flesta av de installerade testerna går även köra med IPv6.

Tester

SUNETs server är konfigurerad för engångsmätningar med hjälp av NDT, BWCTL, NDT, NPAD eller OWAMP (se förklaring nedan). Alla dessa tester har inbyggd schemaläggning, så att bara en test i taget kan genomföras, detta så att testerna inte påverkar varandra.

NDT

NDT är en webbaserad diagnos- och överföringshastighetstest mellan klient och NDT-sever. En mängd TCP överföringar görs under en tidsperiod på 10 sekunder i vardera riktningen. Under överföringen samlas detaljerad statistik in på servern, för att kunna ge information om vilka problem som eventuellt finns med det uppmätta överföringsresultatet.

Det finns även en kommandoradsversion för Linux (web100ctl)

Mer info:

<http://www.internet2.edu/performance/ndt/>

För att testa:

<http://lla-ps.sunet.se:7123/>

<http://lla-ps-v6.sunet.se:7123/>

BWCTL

BWCTL är en "kommandorads-wrapper" för Linux/MacOS runt programmet iperf för att hantera schemaläggning och accessrättigheter. Schemaläggningen tillåter bara en test åt gången, så att testerna inte påverkar varandra. Accessrättigheterna är uppsatta så att alla Sunet-anslutna nät kan köra TCP tester på upp till 9 Gbps under 120sekunder. För att göra en test måste man installera iperf och BWCTL på sin lokala dator.

Med en klient ansluten med 10GE och en MTU på 9k bör det gå att få hastigheter mellan 7-8,5 Gbps. Med en MTU på 1500 så bör det bli mellan 4-5,5 Gbps.

Mer info:

<http://www.internet2.edu/performance/bwctl/index.html>

<http://fasterdata.es.net/performance-testing/network-troubleshooting-tools/bwctl/>

För att testa:

tcp://lla-ps.sunet.se:4823

tcp://se-tug-ps-mdm.nordu.net:4823 (placerad i Stockholm)

NPAD

Är webbaserad test, utformad för att se vilken överföringshastighet du kan förvänta dig mot en server långt borta. Genom att testa överföringskapaciteten till en server relativt nära och mäta överföringsparametrarna dit, beräknar testet vilken överföring du borde få mot en server längre bort. För att göra testet måste du ha en server långt borta i åtanke och ange hur långt borta denna (RTT i ms) är och vilken hastighet du förväntar dig. Det finns även en kommandoradsversion för Linux.

Med en klient ansluten med 10G och en MTU på 9k kan du, mot SUNETs server, förvänta dig maximalt 3,5-4Gbps . Med en anslutning på 1G och MTU på 1500 så går det att få nästan 1Gbps.

Mer info:

<http://www.psc.edu/index.php/component/content/article/12-networking/644-npad-diagnostic-pathdiag>

För att testa:

<http://lla-ps.sunet.se:8000/> (finns tyvärr inte på ipv6)

OWAMP

På servern finns en Java-klient för "**One-Way Active Measurement Protocol**". Med ping kan man mäta tiden från klienten till server och tillbaka. Med OWAMP mäts avståndet i tid mellan server och klienten, för att detta ska fungera krävs det en exakt klocka på klienten eftersom mätningen är baserad på tidstämplrar. Denna är just nu lite ostabil, så tills vidare avråder vi från att förlita sig på de resultaten.

Tuning

TCP/IP-stacken på moderna klient-OS är ganska bra, men ofta är det nödvändigt att trimma inställningarna för TCP på sin klient för att få ut maximal överföringshastighet på långa avstånd. Våra kollegor på ESnet har sammanställt information för de vanligaste operativsystemen, samt också en bakgrund till dessa inställningar.

<http://fasterdata.es.net/host-tuning>