

Traffic Engineering op een Force10



Door: Marco van Doorn en Roeland Kluit
Datum: 01 juli 2004
Vak: Analytisch Netwerk Project
Begeleiding: Freek Dijkstra

FORCE ™



Management summary

Voor het Analytisch server project is er onderzoek gedaan naar verkeersbeperking en reservering op een Force10 switch geplaatst bij de UvA. Wat zijn de mogelijkheden om op een Force10 switch verkeerscontrole te activeren en op welke wijze kan dit het beste worden gerealiseerd.

Er zijn verschillende manieren van verkeersbeperkingen en reserveringen. De overkoepelende term welke wordt gehanteerd is traffic engineering.

Het onderzoek bestaat uit een aantal onderdelen. Er zal worden gekeken naar verkeersbeperking, reservering en verschillende varianten hierop.

Om de testen te kunnen uitvoeren is er gebruik gemaakt van een Force10 switch van het type E600 welke is voorzien van een loopbackverbinding van 1 gigabit/s. Verkeer over deze verbinding wordt gegenereerd tussen 4 systemen welke zijn aangesloten op de switch.

Binnen de testomgeving zullen de volgende testen worden uitgevoerd. Zonder traffic engineering zal verkeer zonder beperking worden bekeken om zo een beeld te kunnen vormen van de normale verkeersstromen. Bij beperking wordt gekeken in hoeverre het mogelijk is verkeer te limiteren tot maximale waarden op de switch. Tevens zal er worden gekeken naar reservering, in dit geval zal een bepaald deel van de bandbreedte wordt toegewezen. Een zelfde soort meting zal worden uitgevoerd voor reservering met piek. In dit geval is het toegestaan, wanneer er meer bandbreedte beschikbaar is, deze te gebruiken. De totaal gereserveerde bandbreedte kan echter meer zijn dan de beschikbare bandbreedte. In dit geval wordt er een garantie gegeven voor een bepaalde bandbreedte. Echter als deze niet gebruikt wordt mag deze bandbreedte worden gebruikt voor andere verbindingen.

De conclusie die uiteindelijk uit dit onderzoek kan worden getrokken is dat het beïnvloeden van datastromen soms erg af kan wijken van ingestelde waarden. Voor elke beperking van verkeer, zowel UDP als TCP, geldt dat er fluctuaties in verkeersstromen optraden.

Uiteindelijk bleek dat voor TCP de functie rate shaping een goede manier is om verkeer te beperken.

Voor UDP kan de rate-police functie het beste worden gebruikt. Tevens blijkt uit dit onderzoek dat verkeersbeperking pas moet worden toegepast als men zeker weet hoeveel de beperking moet zijn, welk protocol wordt gebruikt en of er een peak-waarde moet worden ingesteld. Het blijkt dat als er een peak waarde wordt ingesteld, deze ook direct volop wordt benut. Soms kan dit vervelend zijn.

Tevens is gebleken dat verkeersbeperking op alle poorten wel een goed resultaat geeft, in tegenstelling tot wanneer de beperking op 1 poort wordt ingesteld.

Inhoud

1	Inleiding	5
2	Doel van het onderzoek	6
2.1	Vraagstelling	6
2.2	Traffic engineering	6
2.3	Onderzoek	7
2.4	Soorten Traffic engineering	7
3	Uitvoering	9
3.1	Opzet	9
3.1.1	Netwerk	9
3.1.2	Uit te voeren testen	10
3.1.3	Performance metingen	12
3.2	Overige testen	13
4	Uitvoering van het onderzoek	14
4.1	Geen traffic engineering	14
4.2	De metingen	15
5	Uitkomst van het onderzoek	22
6	Conclusie	24
7	Bronnen en bijlagen	25

1 Inleiding

De hedendaagse netwerkkapparatuur wordt steeds sneller. Netwerkverbindingen van enkele gigabits zijn tegenwoordig eerder regel dan uitzondering. Ondanks de grote snelheden is het in sommige gevallen belangrijk om bandbreedte te reserveren of juist te beperken, zodat bijvoorbeeld het real-time video verkeer voorrang krijgt op andere verkeersstromen.

De netwerkkapparatuur wordt geleverd met veel verschillende opties om de deze te configureren. Met dit onderzoek wordt alle aandacht gericht op een bepaalde functionaliteit van een switch, van het merk Force10 (type E600). Een van de vele functionaliteiten die deze switch biedt is *rate limiting* en *traffic shaping*. In dit document wordt uitvoerig beschreven wat dit inhoudt en op welke manier deze functionaliteit zal worden getest.

In de periode van begin tot eind juni is er een project gaande genaamd Analytisch Netwerk Project (ANP). Bij dit project moet men een onderwerp kiezen dat te maken heeft met de studie Systeem- en Netwerkbeheer en dit voornamelijk binnen het vakgebied dat het netwerkbeheer omvat. In dit geval zal er worden gekeken naar *rate limiting* en *traffic shaping* op een Force10 switch.

2 Doel van het onderzoek

2.1 Vraagstelling

Bij de Universiteit van Amsterdam (UvA) is een Force10 switch geplaatst, deze switch wordt gebruikt voor diverse testdoeleinden. Zo wordt deze bijvoorbeeld gebruikt om (high speed) internationale lijnen op af te monteren. Momenteel worden de functionaliteiten van de switch niet volledig benut en graag wil men zien in hoeverre deze, volgens de fabrikant gegeven specificaties, kunnen worden toegepast.

Volgens de specificaties heeft het model van de switch dat wordt gebruikt voor dit onderzoek (E600) verschillende features, waaronder *rate limiting* en *traffic shaping*. Met dit onderzoek zal worden gekeken naar verschillende mogelijkheden die verkeersbeperking biedt en of dit al dan niet in overeenstemming is met de specificaties van deze switch.

Op welke manieren is het mogelijk op een Force10 switch het gebruik van bandbreedte te beheersen? Op welke wijze kan dit het beste worden gerealiseerd?

2.2 Traffic engineering

Rate limiting en traffic shaping zijn 2 verschillende methoden om het verkeer van een bepaalde verbinding te manipuleren volgens bepaalde ingestelde condities. Beide methoden hebben gemeen dat de beschikbare bandbreedte kan worden beïnvloed en daarom wordt er in dit document gebruik gemaakt van de overkoepelende term traffic engineering. Dit omvat alle mogelijke manieren om bandbreedte te limiteren. Als er in dit document over traffic engineering wordt gesproken dan wordt hiermee de algemene terminologie voor het beïnvloeden van verkeersstromen bedoelt.

Een switch heeft een of meerdere netwerkaansluitingen naar andere apparaten. Deze koppelingen hebben vaak een vastgestelde maximale data doorvoer. In veel voorkomende gevallen is het praktisch bepaalde verbindingen een minimale ook wel gegarandeerde bandbreedte toe te kennen. Een voorbeeld hiervoor is een videostream, men wil tenslotte niet dat het beeld gaat haperen als andere mensen de lijn overbelasten.

Om een beperking op te leggen aan het gebruik van de bandbreedte wordt er gebruik gemaakt van rate-limiting of traffic-shaping. Een verschil tussen de 2 varianten zit hem voornamelijk in de manier waarop het verkeer wordt gelimiteerd.

In het geval van rate-limiting worden de pakketten die vanwege de bandbreedte niet verzonden kunnen worden weggegooid (Drop). Dit heeft gevolgen voor de data die verstuurd wordt - de data gaat ten slotte verloren. Het protocol zal deze data nogmaals moeten verzenden om er voor te zorgen dat de data correct wordt afgeleverd.

Bij traffic-shaping wordt het verkeer niet weggegooid maar wordt deze tijdelijk in een input buffer geplaatst. Het protocol zal dit sneller doorhebben en de snelheid waarmee data wordt aangeboden verlagen.

2.3 Onderzoek

Om een antwoord op eerder gestelde vragen te kunnen geven, zal er het een an ander op het gebied van traffic engineering moeten worden onderzocht. Om dit te kunnen realiseren zal er binnen een testopstelling gebruik worden gemaakt van de beschikbaar gestelde apparatuur, waaronder de Force10 switch.

De beschikbare apparatuur en software:

- Force10 Switch
- 4 systemen (computers) met in elk systeem een 1 Gbs netwerkkaart
- Koppeling tussen de systemen en de switch m.b.v. fiber
- Iperf applicatie om netwerkverkeer te genereren (<http://dast.nlanr.net/Projects/lperf/>)

De 2 systemen zullen op de switch worden aangesloten om een verbinding met elkaar te kunnen realiseren. Daarna zal er gebruik worden gemaakt van verschillende opties die de switch biedt t.b.v. Traffic Engineering.

Door veel (gecontroleerde) data te genereren tussen de systemen is het mogelijk om na te gaan hoeveel bandbreedte er gebruikt zal worden en hoeveel data de switch door zou mogen laten. Aangezien er verschillende mogelijkheden zijn voor Traffic Engineering zal elke mogelijkheid worden bekeken en worden getest met bovenstaande testopstelling

2.4 Soorten Traffic engineering

Binnen traffic engineering valt onderscheid te maken tussen de verschillende soorten van limitatie of reservering van bandbreedte. Zo is het mogelijk om een bepaalde poort een minimale bandbreedte toe te wijzen maar wel meer verkeer toe te staan. In het volgende stuk zullen de verschillende type beperkingen en reserveringen worden besproken. Ook zal een aantal mogelijke toepassingen worden gegeven.

◆ Beperking (Limitation)

In het geval van een beperking wordt de bandbreedte welke beschikbaar is beperkt tot een vastgestelde grens. Verkeer kan niet sneller worden verstuurd dan de waarde die is ingesteld.

- Hierbij valt voornamelijk te denken aan de beperkingen die ADSL en kabel leveranciers toepassen op hun netwerk.

◆ Reservering (Reservation)

In het geval van reservering wordt de bandbreedte welke beschikbaar is vastgelegd. Te allen tijde is de bandbreedte die is vastgesteld beschikbaar voor de verbinding.

- Hierbij valt voornamelijk te denken aan een leased-line (huurlijn). Deze wordt veelal met een vaste bandbreedte afgeleverd.

◆ Reservering met peak (Reservation with peak)

In het geval van reservering met peak wordt de bandbreedte welke beschikbaar is vastgelegd, het is echter toegestaan, wanneer er meer bandbreedte beschikbaar is, deze te gebruiken. Te allen tijde is de bandbreedte die is vastgesteld beschikbaar voor de verbinding maar hogere doorvoersnelheden zijn mogelijk, maar deze worden niet gegarandeerd.

- Hierbij valt voornamelijk te denken aan toepassingen voor Voice over IP en Video Streams.

◆ **Reservering met overboeking (Reservation with overbooking)**

In het geval van reservering met overboeking wordt de bandbreedte welke beschikbaar is vastgelegd. Te allen tijde is de bandbreedte die is vastgesteld beschikbaar voor de verbinding. Echter, wanneer niet de volledige bandbreedte wordt gebruikt mag deze worden gebruikt voor andere verbindingen. Het gereserveerde verkeer heeft altijd voorrang op het niet gereserveerde verkeer.

- Hierbij valt voornamelijk te denken aan toepassingen voor Voice over IP en Video Streams of Backup systemen.

Bij de eerder genoemde mogelijke traffic engineering is het altijd mogelijk bepaalde wenselijke combinaties toe te passen.

3 Uitvoering

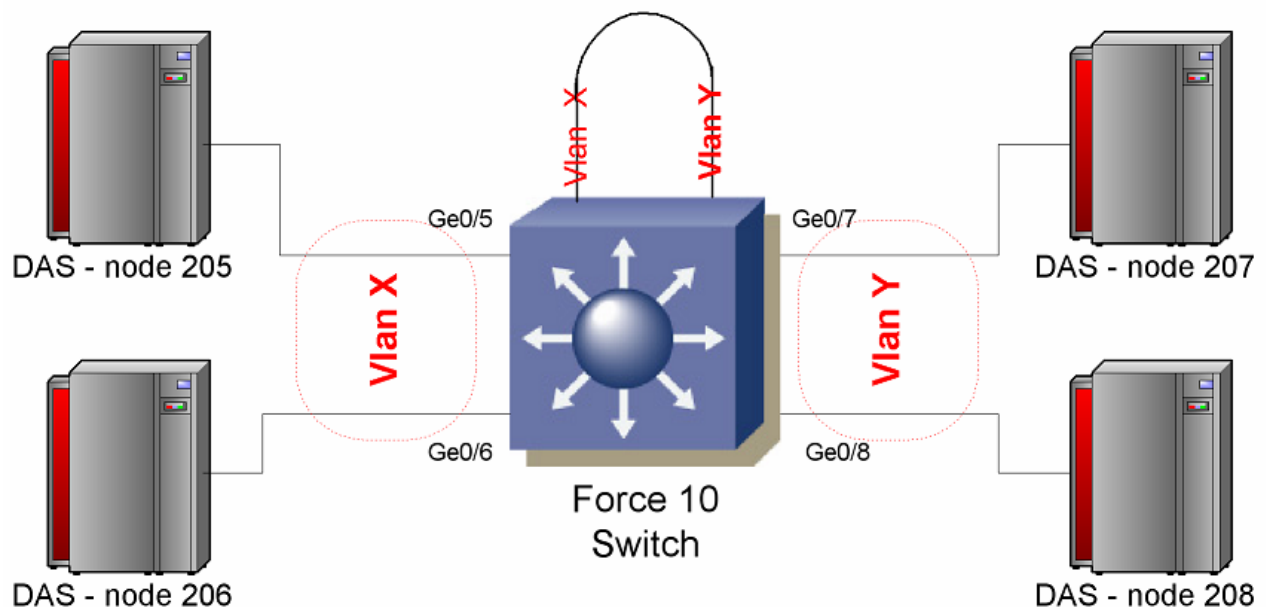
In dit hoofdstuk wordt de uitvoering van het project beschreven. Paragraaf 3.1 geeft een indicatie van de werkzaamheden die vooraf zijn bedacht om te uit te voeren. Hoofdstuk 4 geeft een beschrijving van de uitvoering, zoals het in de praktijk is gebeurd. Ook zijn hier de configuratieoverzichten te vinden die zijn gebruikt om de uitvoering te realiseren.

3.1 Opzet

Om de functionaliteiten te kunnen vergelijken zal een aantal testen worden gedaan. De uit deze testen gehaalde resultaten zullen worden gebruikt om een uitkomst te kunnen vaststellen.

3.1.1 Network

Om de traffic engineering correct te kunnen testen is er voor gekozen de, al, aanwezige koppelingen met een supercomputer (de DAS-2) te benutten. De metingen worden in de volgende opstelling uitgevoerd:



Afbeelding 3.1-1: Network overzicht

Aan de linkerkant van de afbeelding staan de twee systemen die verkeer genereren via de force10 switch naar de systemen aan de rechterkant van de afbeelding. Op de switch wordt door middel van een loopback koppeling een stukje netwerk tussen de verschillende systemen gesimuleerd.

Wegens de beschikking over slechts 1 switch was het niet mogelijk om gebruik te maken van een *trunk*. Een trunk zou kunnen voorzien in het limiteren van data per VLAN. Er is uiteindelijk gekozen voor bovenstaande loopback koppeling op de switch (zie bovenstaande afbeelding) om alsnog de standaardmetingen uit te kunnen voeren. Limitatie op een trunk wordt in dit verslag dus niet behandeld.

De IP adressen die gebruikt worden binnen het netwerk zijn als volgt;

DAS 205: 145.146.100.15 / 26 (Ge0/5)
DAS 206: 145.146.100.16 / 26 (Ge0/6)

DAS 207: 145.146.100.17 / 26 (Ge0/7)
DAS 208: 145.146.100.18 / 26 (Ge0/8)

De aanduiding *Ge0/7* geeft aan dat dit om een interface op de Force10 switch gaat en wel een Gigabit interface op kaart 0, positie 7.

3.1.2 Uit te voeren testen

De volgende testen dienen uitgevoerd te worden om te kijken welke soorten traffic engineering de Force10 switch ondersteund.

◆ **Meting 1: Geen Traffic Engineering.**

Als eerste moet er worden gekeken naar de standaard verkeersstromen, wanneer er op de interfaces geen verkeersbeperkingen zijn toegepast. De resultaten hiervan kunnen worden gebruikt voor de vergelijking met de andere metingen.

Op geen van de interfaces zal een limitatie worden ingesteld.

Er zullen binnen de testopstelling 2 verschillende soorten verkeer worden gegenereerd. De eerste meting zal worden gedaan met behulp van TCP verkeer. De tweede meting zal worden gedaan met UDP verkeer.

◆ **Meting 2: Beperking**

De Force10 switch ondersteunt beperking van de bandbreedte van een interface. Met deze beperking van de bandbreedte zou het mogelijk moeten zijn om te testen of deze ingestelde waarde niet wordt overschreden. (Bij beperking heeft men een bepaalde bandbreedte, maar deze bandbreedte kan wel verschillen. Het kan op sommige momenten, bij veel gebruik van de gehele lijn, minder zijn dan de totale opgestelde beperking.)

Volgens het document met alle command-line opties kan dit worden bereikt met het 'rate-limit' commando. Dit commando zou het mogelijk moeten maken om een bepaalde interface een maximale bandbreedte te geven.

Om bovenstaande te realiseren zal het commando met de volgende waarde moeten worden aangeroepen:

- *committed rate* (Hiermee wordt de te gebruiken bandbreedte vastgelegd. Dit wordt uitgedrukt in Mbps)

Het commando voor het limiteren van de bandbreedte op b.v. 100 Mbps kan er als volgt uitzien:

```
Force10# config terminal
Force10(conf)# qos-policy-input QoSPolicy25
Force10(conf-qos-policy-in)# rate-limit 100
Force10(conf-qos-policy-in)# end
Force10 #
```

Hierboven wordt een QOS policy aangemaakt met een bepaalde naam. Deze aangemaakte policy is te koppelen aan een bepaald interface.

◆ **Meting 3: Reservering**

Reservering van een bepaalde bandbreedte zou kunnen worden gerealiseerd met het commando *bandwidth-percentage*. Hiermee kan worden aangegeven wat voor percentage van de beschikbare lijn mag worden gebruikt. In theorie is de toegekende bandbreedte altijd beschikbaar en nooit meer of minder.

Om bovenstaande te realiseren zal het commando met de volgende waarde moeten worden aangeropen:

- *percentage* (Het percentage dat van de totale lijn mag worden gereserveerd)

Het commando in zijn geheel (voor het b.v. limiteren op 10%) zou kunnen zijn:

```
Force10# config terminal
Force10(conf)# qos-policy-output PolicyName25
Force10(conf-qos-policy-out)# bandwidth-percentage 10
Force10(conf-qos-policy-out)# end
Force10 #
```

◆ **Meting 4: Reservering met peak**

Het zou mogelijk moeten zijn om door middel van peak-waarden aan te geven hoeveel een bepaalde lijn boven zijn eigen grens mag komen als er (ongebruikte) bandbreedte over zou zijn op een bepaalde lijn. Het commando *rate-limit* zou hierin voorzien.

Om bovenstaande te realiseren zal het commando *rate-police* met de volgende opties moeten worden uitgevoerd:

- *committed rate* (Instellen van de bandbreedte (Mbps) dat maximaal mag worden gebruikt)
- *peak (peak-rate)* (Instellen van de hoeveelheid bandbreedte (Mbps) die boven de committed rate mag worden gebruikt. Dit getal is de committed rate + peak waarde)

Het commando dat hiervoor zou moeten zorgen is:

```
Force10# config terminal
Force10(conf)# qos-policy-output QoSPolicy25
Force10(conf-qos-policy-in)# rate-police 100 peak 1000
Force10(conf-qos-policy-in)# end
Force10 #
```

Met de hierboven getoonde opties zou er een rate-limit van 100 Mbps moeten zijn met een peak-level van 1000 Mbps.

Meting 5: Reservering met overboeking

Er is geen direct commando om dit te realiseren. De praktijk moet uitwijzen hoe dit in zijn werk gaat en hoe dit is te limiteren.

3.1.3 Performance metingen

Om testverkeer te kunnen genereren zal er gebruik worden gemaakt van de tool *NLANR/DAST Iperf 1.7.0*. Deze tool beschikt over functionaliteit voor het genereren van verkeer en heeft daarnaast de mogelijkheid om een duidelijk overzicht van de gehaalde snelheden weer te geven. Hieronder staat een voorbeeld van een UDP meting.

```
-----
Client connecting to 192.168.2.3, UDP port 5001
Sending 1470 byte datagrams
UDP buffer size: 63.0 KByte (default)
-----
[1856] local 192.168.2.2 port 2230 connected with 192.168.2.3 port 5001
[ ID] Interval      Transfer      Bandwidth      Jitter    Lost/Total Datagr
[1856] 0.0- 1.0 sec    129 KBytes    1.06 Mbits/sec
[1968] 0.0- 1.0 sec    122 KBytes    1000 Kbits/sec  8.699 ms    0/ 85 (0%)
- cut -
[1856] 0.0-10.0 sec  1.25 MBytes   1.05 Mbits/sec
[1968] 0.0-10.0 sec  1.25 MBytes   1.05 Mbits/sec  8.189 ms    0/ 893 (0%)
[1856] Server Report:
[1856] 0.0-10.0 sec  1.25 MBytes   1.05 Mbits/sec  5.439 ms    0/ 893 (0%)
[1856] Sent 893 datagrams
```

Om deze metingen toepasbaar te maken voor het project zullen er 2 toepassingen voor dit programma zijn.

Ten eerste zal er een verbinding tussen tussen node 206 en node 208 (afbeelding 3.1-1) worden gecreëerd welke zoveel als mogelijk data genereert. De meetresultaten zijn waarschijnlijk niet uiterst interessant maar dit verkeer zorgt ervoor dat we de beperkingen en reserveringen zoals ingesteld op de verbinding tussen node 205 en node 207 kunnen testen. In de praktijk zal moeten worden gekeken of het effectiever is met UDP dan wel TCP te werken. Om deze verkeersstroom te kunnen genereren worden de volgende commando's gebruikt:

Voor UDP verkeer:

Node 206; Server

```
./iperf --server --udp --interval 2 --window 512 [--bandwidth 1000M]
```

Node 208; Client

```
./iperf --client <ip_node_206> --udp --interval 2 [--bandwidth 1000M] --
window 512 [--time <seconden>]
```

Voor TCP verkeer:

Node 206; Server

```
./iperf --server --interval 2
```

Node 208; Client

```
./iperf --client <ip_node_206> --interval 2 [--time <seconden>]
```

Voor een gedetailleerde omschrijving van de commando's en hun werking wordt verwezen naar de website van iperf.



Ten tweede is er de verbinding waarop de traffic engineering zal worden toegepast. Voor de metingen tussen node 205 en node 207 zullen commando's worden gebruikt die gerelateerd zijn aan de commando's uit het bovenstaande overzicht. De nadruk zal hier meer liggen op de resultaten van de metingen. In het geval van reservering, zal ondanks de verkeersstroom welke wordt gegenereerd tussen node 206 en node 208, toch minimaal de gereserveerde bandbreedte beschikbaar moeten zijn. In de uitwerking van de metingen zijn de meer gedetailleerde commando's terug te vinden.

3.2 Overige testen

Wanneer na het uitvoeren van de eerder genoemde testen nog mogelijkheden zijn voor het uitvoeren van andere testen zullen de volgende testen worden uitgevoerd:

Transatlantisch; In dit geval wordt de loopbackverbinding, zoals midden bovenaan in afbeelding 3.2-1 te zien is, worden vervangen door een verbinding welke loopt via Chicago, America. Het doel hiervan is te kijken welke invloeden de langere round-trip-time heeft op bijvoorbeeld TCP sessies en traffic engineering.

Stress-testing; Wat gebeurt er met TCP sessies wanneer er een enorme hoeveelheid sessie tegelijk een grote hoeveelheid data wordt verzonden. Wat valt hier uit op te maken en is hier een patroon in te ontdekken.

Traffic engineering over trunks; Wat zijn de mogelijkheden die de Force10 switch biedt aan traffic engineering over trunks (Tagged interfaces).

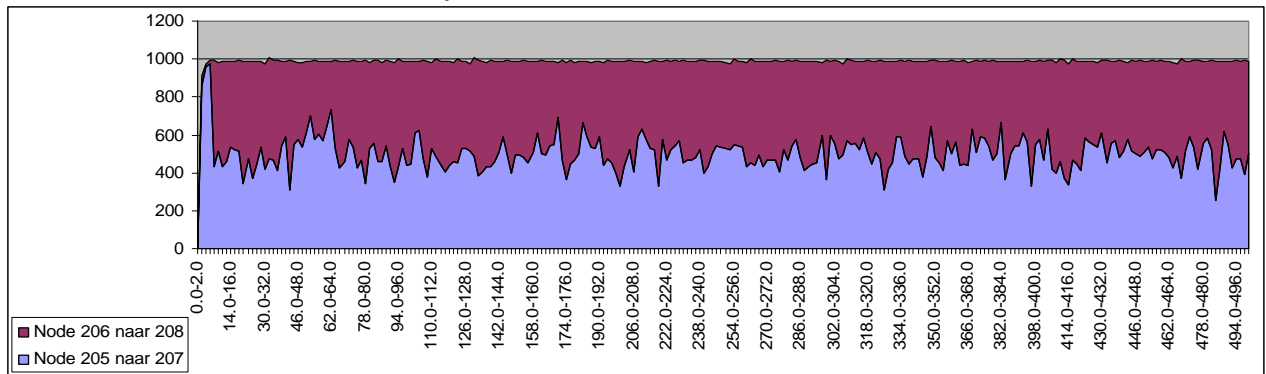
4 Uitvoering van het onderzoek

4.1 Geen traffic engineering

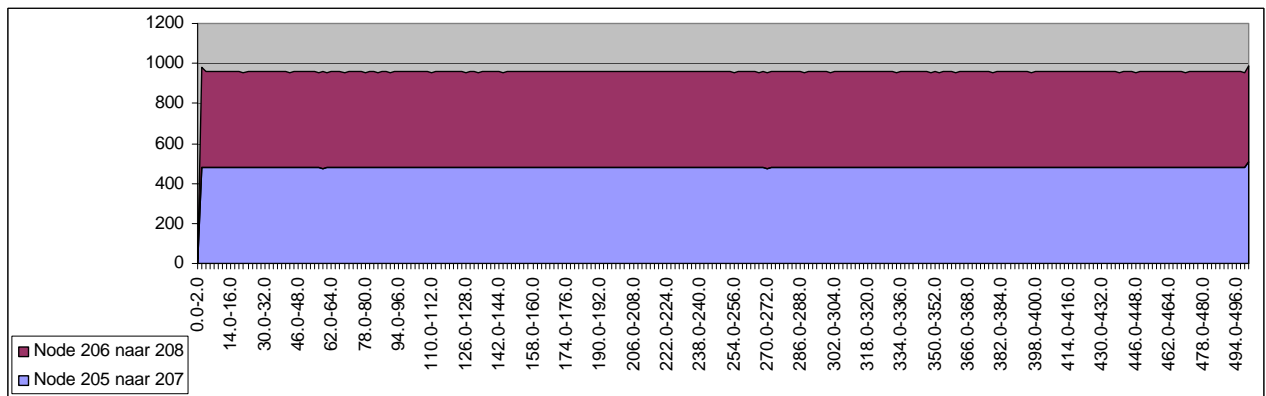
Door verkeer te genereren tussen verschillende nodes (onderling) kan er worden gekeken hoeveel bandbreedte er maximaal (bij elkaar opgeteld) wordt gebruikt. In onderstaande grafiek is het resultaat van deze eerste meting te zien.

TCP:

Verticaal: bandbreedte, Horizontaal: tijd



UDP:



De paarse kleur geeft het verkeer aan van node 206 *naar* node 208. De blauwe kleur geeft verkeer aan van node 205 *naar* node 207.

In bovenstaande grafieken is het duidelijk te zien dat er nagenoeg 1 gigabit/s over de loopback-lijn gaat. Waar het verkeer van node 206 *naar* node 208 inzakt, wordt dit gat opgevangen door het verkeer dat loopt van node 205 *naar* node 207.

Het verschil tussen beide grafieken is dat TCP erg veel pieken en dalen heeft ten opzichte van UDP. De reden hiervoor is dat TCP aan flow-control doet en bij een te hoge bandbreedte neemt de snelheid van de ene datastroom af, waardoor de bandbreedte van de andere datastroom toeneemt.

4.2 De metingen

De metingen die zijn uitgevoerd op de Force10 switch worden in dit hoofdstuk behandeld. Om de configuratie sneller te kunnen aanpassen is er bij de uitvoering gekozen om de configuratie direct per interface te configureren. Dit in tegenstelling tot de voorbeelden waar deze in de QoS policies werden gedefinieerd.

Binnen de testomgeving was er helaas geen mogelijkheid Traffic Engineering toe te passen op *trunks*, omdat er geen 2^e switch beschikbaar was.

rate police (incoming traffic)

Bij deze eerste meting is er een test gedaan met het commando *rate police* zonder piek-waarden (peak). De metingen met een ingestelde piekwaarde zullen later worden vermeld.

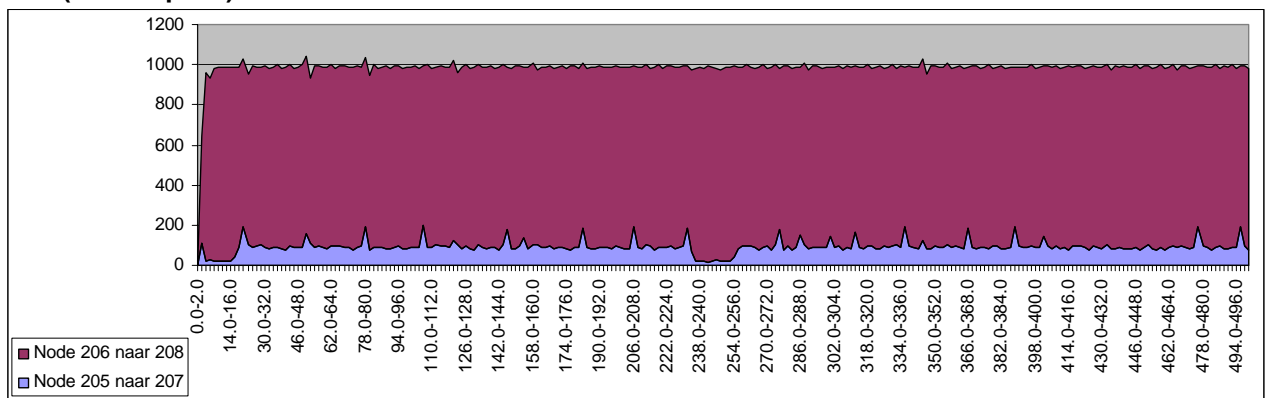
Er moet tevens een **burst-waarde** worden ingesteld. Een burst waarde is een buffer die wordt gereserveerd op de line-card van een bepaalde interface. De standaard waarde is 60 kilobyte, maar omdat de switch met deze waarde niet volledige pakketten kwijt kan in het geheugen komen er vreemde waarden tijdens het testen. Door de waarde hoger in te stellen is het vrijwel zeker dat de pakketten in het geheugen van de linecard passen en daardoor wordt een goede bandbreedte gehaald.

```
Force10(conf-if-gi-0/5)# rate police 100 26000
```

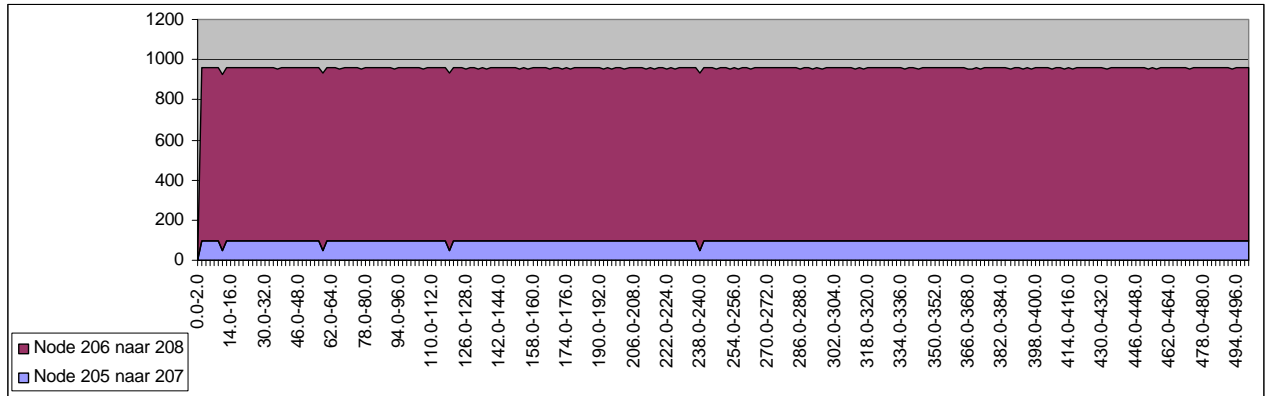
De meting betreft het verkeer van node 205 naar node 207 waar een beperking is geplaatst op de **inkomende** interface (gigabit ethernet 0/5) op de switch. Hierop is node 205 aangesloten. Ter verduidelijking: het inkomende verkeer wordt beperkt

Dit houdt in dat hij niet boven de ingestelde waarde mag komen. Zelfs niet als er meer bandbreedte beschikbaar zou zijn.

TCP (zonder peak):



UDP (zonder peak):

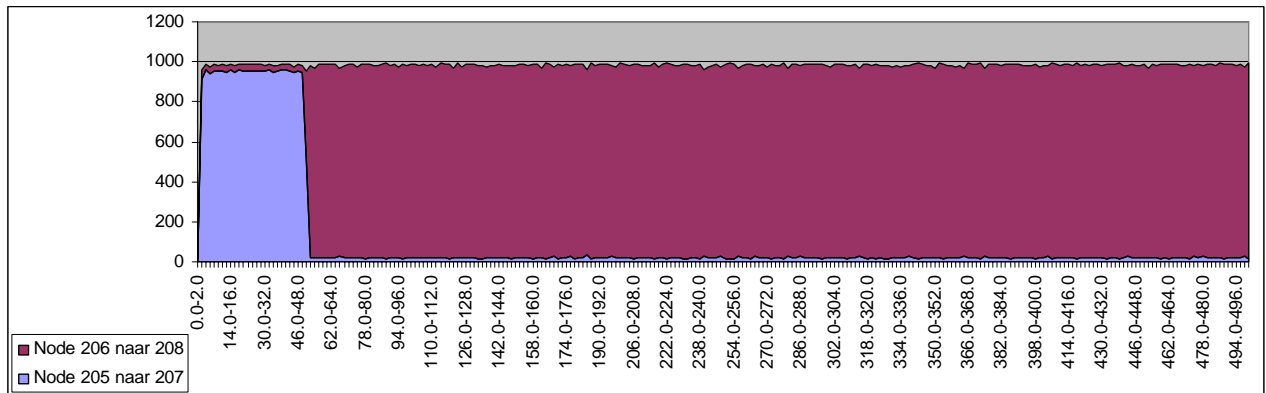


Bovenstaande grafieken laten zien dat er gemiddeld 92.2 megabit/s wordt gebruikt (blauw) van de totale bandbreedte. De overige bandbreedte wordt gebruikt door de gelimiteerde datastroom tussen node 206 en node 208 (paars).

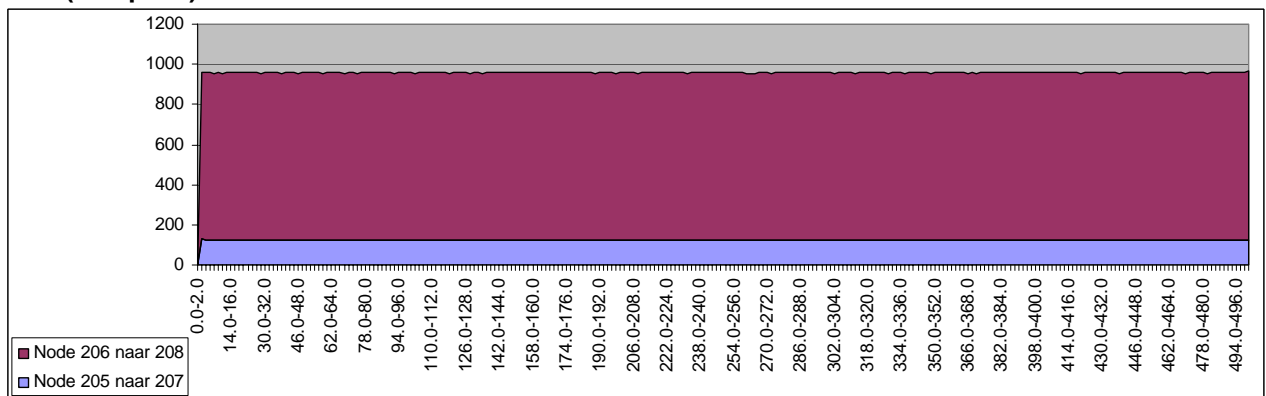
Bij beperking van het verkeer is een aantal opvallendheden geconstateerd. Zoals eerder beschreven zijn er verschillende mogelijkheden om verkeer te beperken over een lijn of een VLAN. Vervolgens is er een test gedaan met het commando *rate police met peak*.

De uitkomst van deze verkeersbeperking was erg verrassend, omdat de beperking niet direct effect had op de lijn. Onderstaande grafiek laat de uitkomst van deze meting zien.

TCP (met peak):



UDP (met peak):



Bij deze test zijn er 4 nodes gebruikt. De paarse kleur geeft het verkeer aan van node 206 naar node 208. De blauwe kleur geeft het verkeer aan van 205 naar 207. De traffic limiting wordt hier gerealiseerd op de verbinding van node 205 naar node 207 (blauw).

Het verschil tussen TCP en UDP is hier dat UDP een strakke bandbreedte heeft en dat TCP sterker pieken en dalen vertoont.

De onderstaande configuratieregel zorgt ervoor dat er een limiet van 10 megabit/s wordt ingesteld met een maximale piek (peak) van 1 gigabit/s. Dit commando is gerealiseerd op het interface waar het systeem aan zit en dus **niet** op de trunk of de loopback interface.

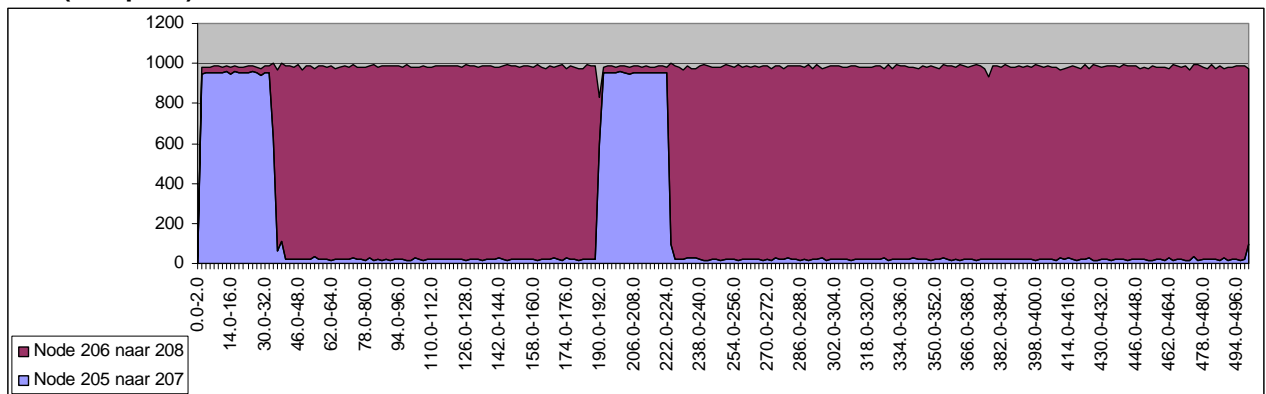
```
Force10(conf-if-gi-0/5)# rate police 10 peak 1000 26000
```

In bovenstaande grafiek is duidelijk te zien dat de datastroom afgebeeld met de blauwe kleur na +/- 30 seconden pas reageert op het ingevoerde commando. Hierna zakt de bandbreedte in naar +/- 20 megabit/s (terwijl de ingestelde waarde 10 megabit/s is!) en de datastroom afgebeeld met de paarse kleur (data van node 206 naar node 208) gebruikt de overige bandbreedte. Hiervoor is geen extra commando ingevoerd. De paars gekleurde verbinding gebruikt daarom de bandbreedte die op dat moment beschikbaar is. Zo is de totaal gebruikte bandbreedte nagenoeg 1 gigabit/s.

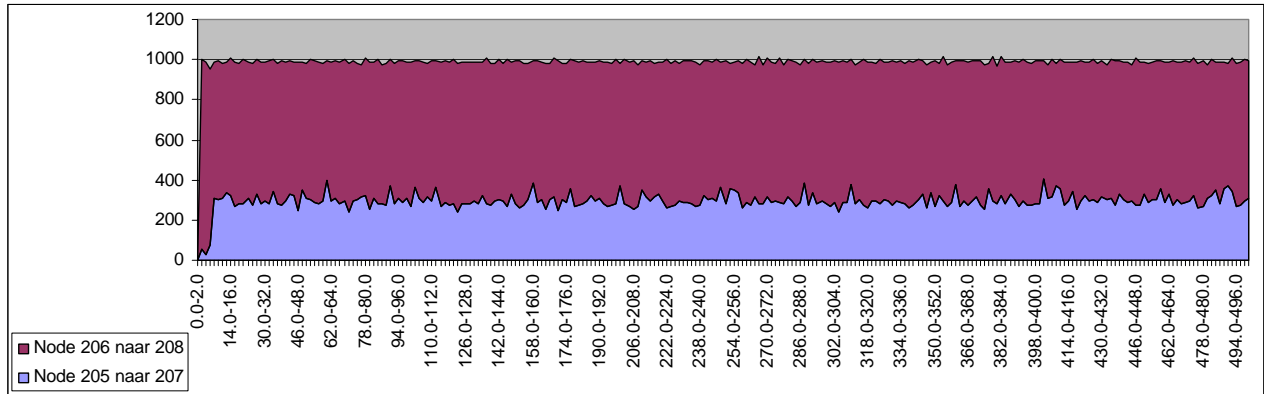
Het is met deze meting dus zichtbaar dat het verkeer (paars) waar geen *rate police* is ingesteld voorrang heeft op het verkeer waar dit wel (blauw) is ingesteld. Dit is af te leiden uit het feit dat het blauw gekleurde verkeer een piek-niveau van 1 gigabit/s mag hebben (zie bovenstaande instelling).

Het vreemde van bovenstaande meting is dat de beperking pas na +/- 30 seconden actief wordt en dan zelfs maar +/- 20 megabit/s, terwijl er 10 megabit/s is ingesteld. Om er zeker van te zijn dat de meting niet verkeerd is gegaan is de meting nogmaals uitgevoerd. In onderstaande grafiek is te zien dat de verbinding (na 190 seconden) opeens niet meer wordt gelimiteerd en op de +/- 222 seconden wel weer inzakkt naar de **niet**-ingestelde 20 megabit/s.

TCP (met peak):



Rate police (met peak en op beide interfaces)



Bij de eerste meting (TCP met peak) kwam er een vreemd resultaat tevoorschijn bij het limiteren op 1 interface. Bovenstaande grafiek laat zien wat er gebeurt als er op beide interfaces wordt gelimiteerd (een interface op 300 megabit/s en een interface op 700 megabit/s) met beide 1000 megabit/s als peak. Het gemiddelde komt hiermee op de ingestelde waarde.

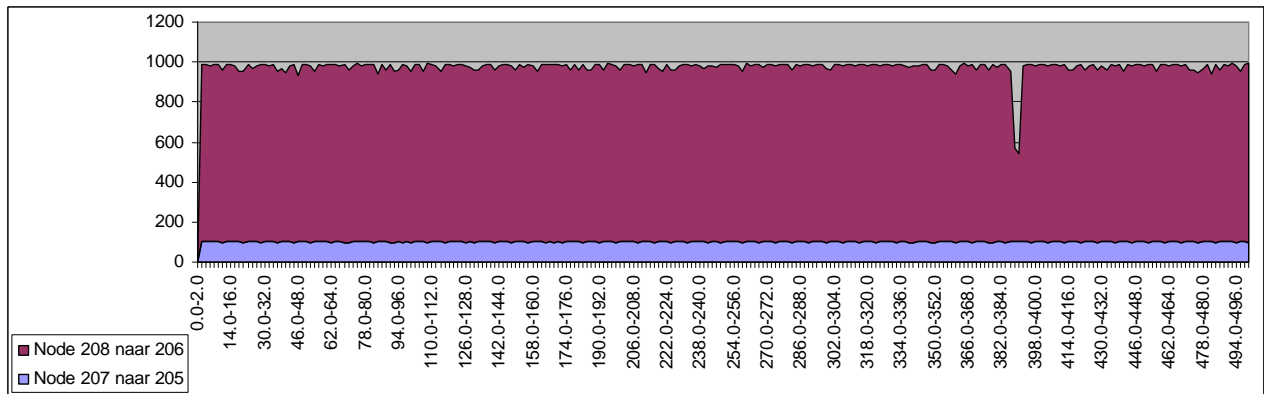
Onderstaande regels moeten op elk interface worden geconfigureerd om bovenstaande te realiseren.

```
Interface 0/5: Force10(conf-if-gi-0/5)# rate police 300 peak 1000 26000
Interface 0/6: Force10(conf-if-gi-0/6)# rate police 700 peak 1000 26000
```

rate shape (outgoing traffic)

De meting betreft het verkeer van node 207 naar node 205 waar een beperking is geplaatst op de **uitgaande** interface (gigabit ethernet 0/5) op de switch. Hierop is node 205 aangesloten. Ter verduidelijking: het uitgaande verkeer wordt beperkt

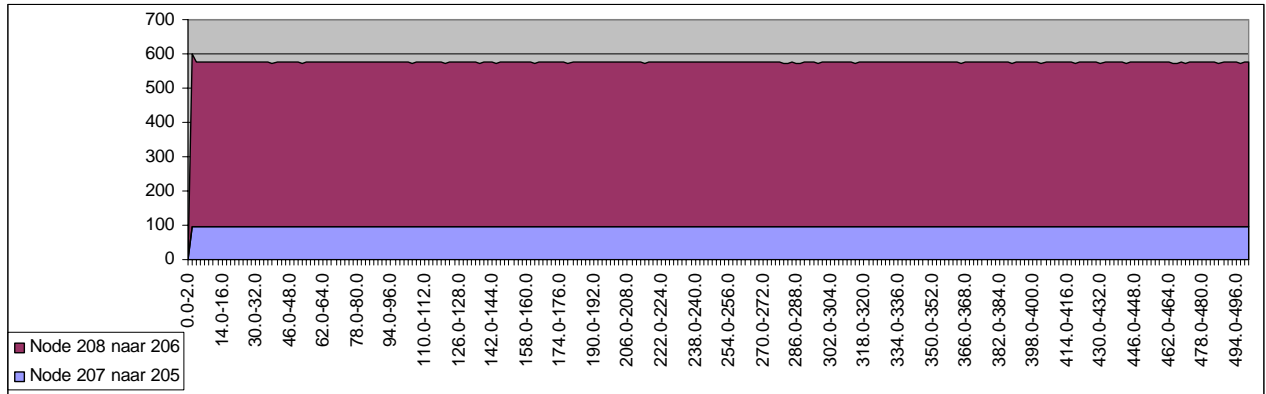
TCP (zonder peak):



Bij traffic-shaping op TCP verkeer wordt er gelimiteerd op de uitgaande interface (dus nadat de data door de switch is gegaan). Dit houdt in dat de node die de data verzendt zijn snelheid moet verlagen om een stream van 100 megabit/s aan te houden. Uiteindelijk gaat er dus 100 megabit/s door de switch en zo blijft er voor de andere stream ongeveer 900 megabit/s over.

Opvallend is dat deze meting een aantal pieken en dalen vertoont. De oorzaak hiervan is dat TCP soms sneller probeert te versturen en daarna (omdat het wordt gelimiteerd) wordt afgeremd.

UDP (zonder peak):



Bij traffic-shaping op UDP verkeer is een leuk resultaat zichtbaar. Het is namelijk zo dat het verkeer pas wordt beperkt op de uitgaande interface van de switch naar de node. Dit houdt in dat er door de switch gewoon 2 streams van +/- 500 megabit/s gaat en dat er uiteindelijk 1 van de 2 streams wordt gelimiteerd naar 100 megabit/s. Zo is de totale gebruikte bandbreedte (na shaping) maar 600 megabit/s.

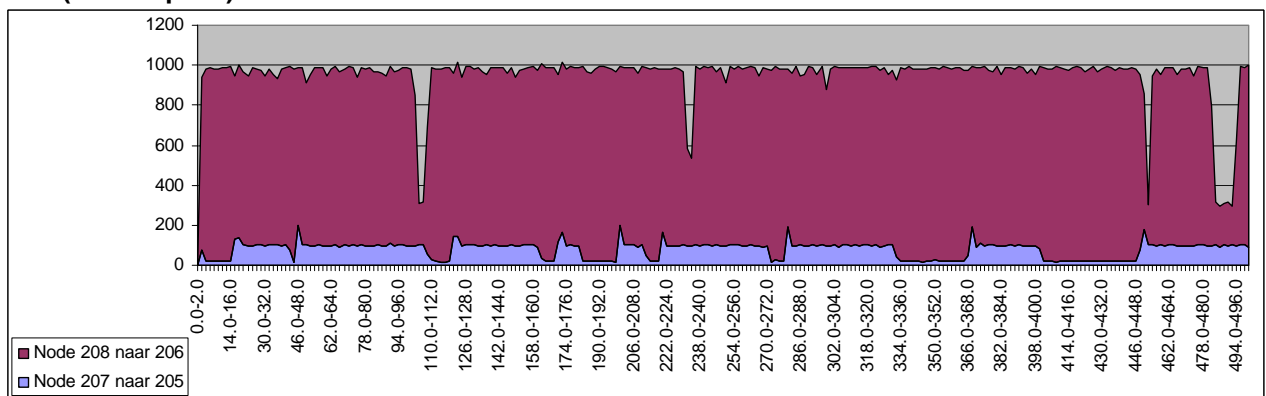
In tegenstelling tot de vorige meting zijn hier nauwelijks pieken en dalen zichtbaar, omdat UDP de pakketten dropt om zo tot de bandbreedte te komen waarop is gelimiteerd. De configuratie luidt:

```
Force10(conf-if-gi-0/7)# rate shape 100
```

rate limit (outgoing traffic)

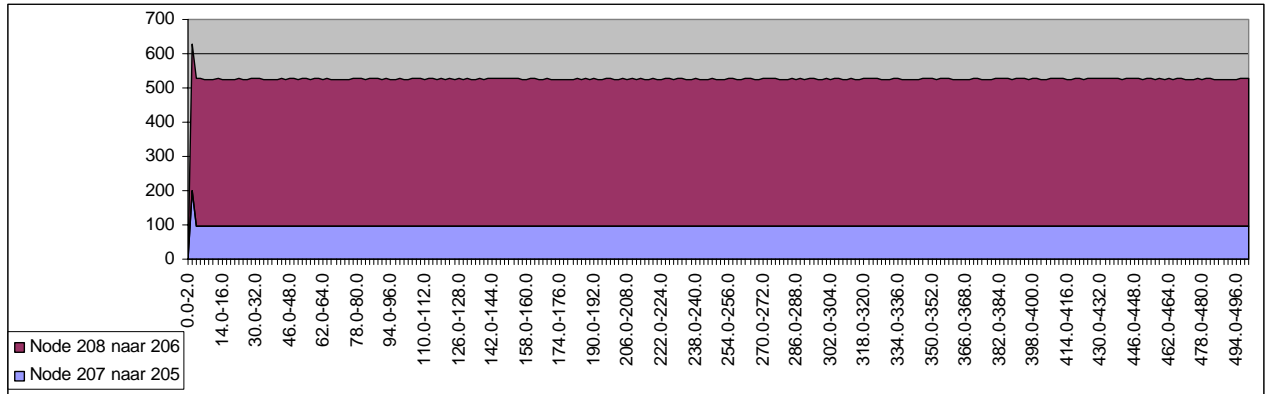
De meting betreft het verkeer van node 205 naar node 207 waar een beperking is geplaatst op de **inkomende** interface (gigabit ethernet 0/5) op de switch. Hierop is node 205 aangesloten. Ter verduidelijking: het inkomende verkeer wordt beperkt

TCP (zonder peak):



Bovenstaande meting geeft de datastroom weer tussen de verschillende nodes met het TCP protocol. Hier is te zien dat de limitatie van 100 megabit/s niet constant blijft en zelfs sterke pieken en dalen vertoond.

UDP (zonder peak):

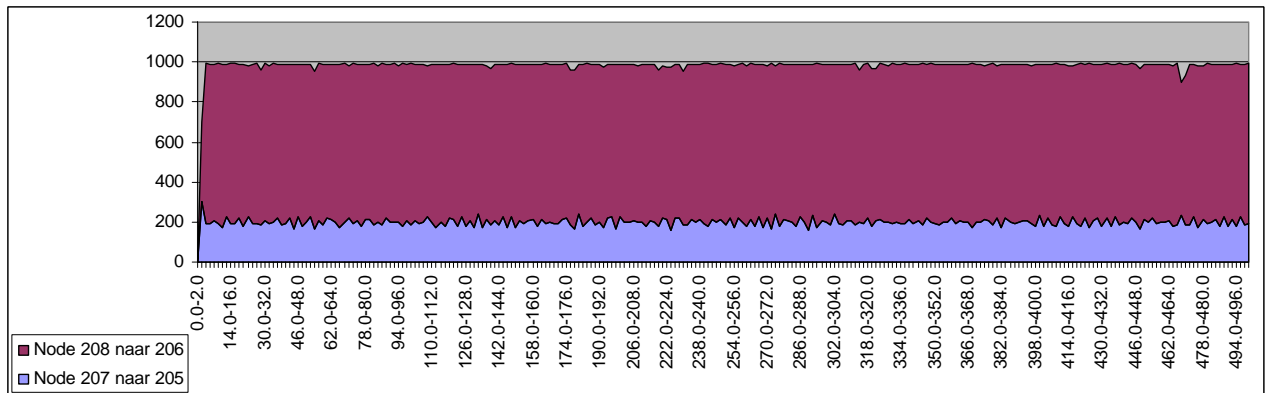


De UDP meting ziet er erg strak uit. De limitatie is 100 megabit/s en er zijn in bovenstaande grafiek geen pieken en dalen zichtbaar. Limitatie gebeurt op de uitgaande interface, waardoor er bij UDP totaal maar 600 megabit/s kan worden gehaald (zelfde als bij rate-shaping met UDP).

De configuratieregel voor bovenstaande test is als volgt:

```
Force10(conf-if-gi-0/5)# rate limit 100 26000
```

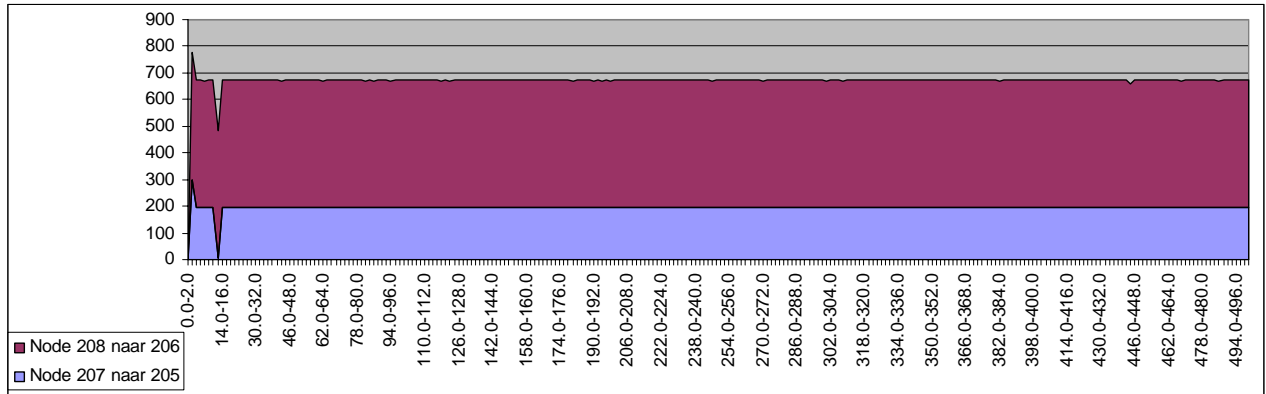
TCP (met peak):



Deze grafiek geeft het resultaat van een rate-limit voor TCP met een piek-waarde van 200 megabit/s voor de datastroom van node 207 naar node 205. De rate-limit is hier redelijk nauwkeurig en het is dan ook goed te zien dat de bovengenoemde datastroom zich rond de 200 megabit/s bevindt. Omdat het TCP protocol aan flow-control doet komen er (zoals in eerdere metingen) tevens pieken en dalen. Over het algemeen geeft deze meting een net resultaat.

Omdat TCP signalen geeft om de snelheid af te remmen (er worden geen pakketten weggegooid!) gaat er uiteindelijk 1 gigabit/s aan totale bandbreedte door de switch. De andere datastroom kan namelijk gebruik maken van de overgebleven bandbreedte.

UDP (met peak):



Er is ook een rate-limit ingesteld voor het UDP protocol met als piek-waarde 200 megabit/s. Voor UDP geldt dat de pakketten worden weggegooid als de bandbreedte over de ingestelde waarde gaat. Dit houdt (net zoals eerdere metingen) dat er bij UDP een strakke rechte lijn komt op de 200 megabit/s. Aangezien er bij rate police wordt gefilterd op de uitgaande interface is de totale bandbreedte dus ongeveer 700 megabit/s.

```
Force10(conf-if-gi-0/5)# rate limit 100 peak 200 26000
```

Beide configuraties zijn gestart met bovenstaande opdrachtregel. Er is uiteindelijk waarneembaar dat de piek waarde wordt benut, ook als de andere datastroom geen limitaties heeft.

5 Uitkomst van het onderzoek

De uitkomst van dit onderzoek is gebaseerd op de testen die zijn verwerkt in dit document. Zoals bij de inleiding is vermeld is er gebruik gemaakt van het programma *iperf* om data te genereren en zo een beeld te kunnen krijgen hoe de switch omgaat met bepaalde hoeveelheden data. Vanwege tijdgebrek is er niet gekeken naar alternatieven voor *iperf* en wordt er vanuit gegaan dat de gegenereerde data juist is en voor goede resultaten zorgt.

De metingen die zijn verricht worden hieronder in een tabel afgebeeld met daarachter het uiteindelijke resultaat van de desbetreffende meting. Ook wordt in de tabel een indicatie gegeven waarvoor het betreffende commando kan worden toegepast.

	TCP	UDP	Toepassing	Opmerking
rate police	Redelijk	Goed	Reservering	Tijdens de TCP meting zijn er opvallende pieken en dalen.
Rate police op beide interfaces met peak	Goed	Goed	Reservering met peak / Reservering met overboeking	Als er op beide interfaces wordt gelimiteerd met een peak waarde dan is de limitatie vrij constant
rate police met peak	Slecht	Uitstekend	Reservering met peak / Reservering met overboeking	Tijdens de meting kwamen er bij TCP onverwachte hoge pieken uit de metingen.
rate shape	Uitstekend	Matig	Beperking / reservering	Bij UDP wordt niet volledige bandbreedte benut, zie hoofdstuk 4.2.
rate limit	Slecht	Matig	Beperking	Zeer slechte resultaten met TCP en bij UDP wordt niet de volledige bandbreedte benut.
rate limit met peak	Voldoende	Matig	Beperking / Reservering met overboeking	Bij UDP wordt niet volledige bandbreedte benut, zie hoofdstuk 4.2.

- ◆ **Rate police**
Tijdens de metingen met rate-police viel voornamelijk de slechte performance van het TCP verkeer op. Grote dalen en verschillende pieken, welke volgens de instelling niet zouden mogen voorkomen, traden op. In sommige gevallen werd de datastroom tot bijna 0 gereduceerd.
De UDP meting was over het algemeen goed. De bandbreedte was nagenoeg constant. Het is wel belangrijk om een grote burst-size op te geven, omdat de performance anders erg laag is.
- ◆ **Rate police met peak**
Tijdens de TCP meting bleek dat de instelling voor rate police met peak niet geheel correct functioneert. De verkeersbeperking wordt pas na tientallen seconden actief, terwijl de andere datastroom al de volledige bandbreedte zou willen benutten. In de tweede meting bleek het effect zelfs meerdere keren op te treden.
De UDP metingen daarentegen waren uitstekend en zeer constant.
Het is wel belangrijk om een grote burst-size op te geven, omdat de performance anders erg laag is.
- ◆ **Rate police op beide interfaces met peak**
Bij bovengenoemde rate police waar de limitatie op een interface is gedaan kwam er geen constante lijn uit als uitkomst. Als er wordt gelimiteerd op 2 interfaces (300 megabit/s om 700 megabit/s) met bijbehorende piek-waarden van 1 gigabit/s is de grafiek gemiddeld genomen veel meer constant.

- ◆ **Rate shape**
De verkeersbeperking met rate shape geeft uitstekende resultaten met TCP. Het verkeer wordt vrijwel constant gelimiteerd.
De UDP meting is op zich een strakke lijn. Echter waar het verkeer pas bij het verlaten van de switch wordt gelimiteerd is de datastroom door de loopback verbinding nog maximaal. Wanneer de rate shape kon worden toegepast op inkomend verkeer zou dit tot veel betere resultaten lijden.
- ◆ **Rate limit**
De TCP metingen gedaan met rate limit waren uiterst slecht. Het verkeerspatroon had veel pieken en grote dalen. Daarnaast werd ook het niet gelimiteerde verkeer ernstig belemmerd.
De UDP meting is wederom een strakke lijn. Echter waar het verkeer pas bij het verlaten van de switch wordt gelimiteerd is de datastroom door de loopback verbinding nog maximaal. Wanneer de rate limit kon worden toegepast op inkomend verkeer zou dit tot veel betere resultaten lijden.
- ◆ **Rate limit met peak**
De TCP meting was redelijk, maar de switch gebruikte wel continu de bandbreedte gelijk aan de ingestelde piek waarde. Desondanks is de verkeersstroom goed te beperken men moet alleen een correcte bovengrens kiezen.
De UDP meting is een redelijk strakke lijn. Echter waar het verkeer pas bij het verlaten van de switch wordt gelimiteerd is de datastroom door de loopback verbinding nog maximaal. Wanneer de rate limit met piek kon worden toegepast op inkomend verkeer zou dit tot veel betere resultaten lijden.
- ◆ **Bandwidth percentage**
Binnen de testomgeving lukte het niet om de het correct te laten functioneren. Waarschijnlijk is dit alleen te testen met een trunk, maar helaas kon dit niet worden gerealiseerd in deze testopstelling. Dit is dan ook de reden waarvoor deze niet in bovenstaande tabel is opgenomen

6 Conclusie

De conclusie die er uit dit onderzoek kan worden getrokken is dat het, al dan niet selectief, beïnvloeden van datastromen soms erg af kunnen wijken van ingestelde waarden.

In het onderzoek is bijvoorbeeld naar voren gekomen dat rate shaping en rate limiting voor UDP niet echt handig zijn, omdat er erg veel bandbreedte verloren gaat aan UDP verkeer dat wordt gedropt bij de uitgaande interface. Dit is niet zozeer een beperking van de switch, maar dit komt meer door het UDP protocol. Als er rate police wordt gebruikt (voor UDP, inkomend verkeer) blijkt dat er hele mooie strakke resultaten uitkomen.

Voor TCP is rate-shaping een goede oplossing voor het beperking van datastromen en geeft rate police weer een slecht resultaat. Wanneer er op beide interfaces een limitatie werd gezet dan waren de resultaten wel als verwacht. Het gemiddelde resultaat was nagenoeg precies hetzelfde als de instellingen.

Voordat er een beperking wordt ingesteld op een bepaalde lijn zal er eerst moeten worden gekeken naar de totale beschikbare bandbreedte en in hoeverre iets moet worden beperkt (met of zonder piek-waarde). Uit het onderzoek blijkt dat een ingestelde piek-waarde direct volledig wordt benut. Dit kan in sommige gevallen vervelend zijn.

Als men dus verkeer wil beperken zal er eerst moeten worden gekeken naar het protocol dat over de desbetreffende lijn gaat. Aan de hand hiervan kan dan een bepaalde functionaliteit worden gekozen die de Force10 biedt om verkeer te beperken.

De Force10 switch biedt al met al diverse mogelijkheden voor het beheren van verkeersstromen. De verschillende methodes hebben elk hun eigen resultaten en daarmee een verschillende toepassing. Wanneer men een beperking wil plaatsen op de verkeersstroom kan dat worden gedaan met een rate limit met peak. Deze peak wordt dan net iets boven de limit waarde ingesteld. Deze geeft duidelijke en voorspelbare verkeersstromen en beperkt de bandbreedte.

Wanneer een bepaalde hoeveelheid bandbreedte voor een poort moet worden gereserveerd is het verstandig gebruik te maken van rate shape. Deze reserveert de benodigde bandbreedte en geeft goede resultaten. Als er gebruik wordt gemaakt van reservering met peak is het verstandig gebruikt te maken van rate police op meerdere interfaces. Duidelijke en goede verkeerstromen bij zowel UDP als TCP verkeer zijn het gevolg.

Binnen de testomgeving is er slechts gebruik gemaakt van een beperkt aantal systemen. Hierdoor was het niet mogelijk om te kijken wat de gevolgen zijn van overboeking. Echter, wanneer de resultaten worden bekeken is het aannemelijk dat rate police op de verschillende interfaces het beste resultaat oplevert. Dit omdat de metingen met rate police op verschillende interfaces overal de beste resultaten geven.

Uiteindelijk is te stellen dat uit de testen naar voren komt dat rate police de voorkeur heeft voor traffic engineering op de Force10 switch. Uit de test komt naar voren dat voor zowel UDP als TCP hiermee goede resultaten kunnen worden behaald.

7 Bronnen en bijlagen

Websites:

UVA Force10 Website: <http://air:air@www.science.uva.nl/research/air/network/force10/>
Force10 Documentation Website: <http://www.force10networks.com/support/customer/documentation.asp>
Force10 Software Website: <http://www.force10networks.com/support/customer/software.asp>
UvA Force10 Cricket: <http://195.169.124.34/cricket/grapher.cgi?target=/f10>
lperf: <http://dast.nlanr.net/Projects/lperf/>

Bijlagen:

Force 10 QoS configuratie guide: FTOS Configuration Guide, Version 5.3.1.0 April 2004