

Network Neutral Traffic Shaping using Linux

Ethan Sommer

Associate Director of Core Services

Gustavus Adolphus College



This is a LONG presentation!



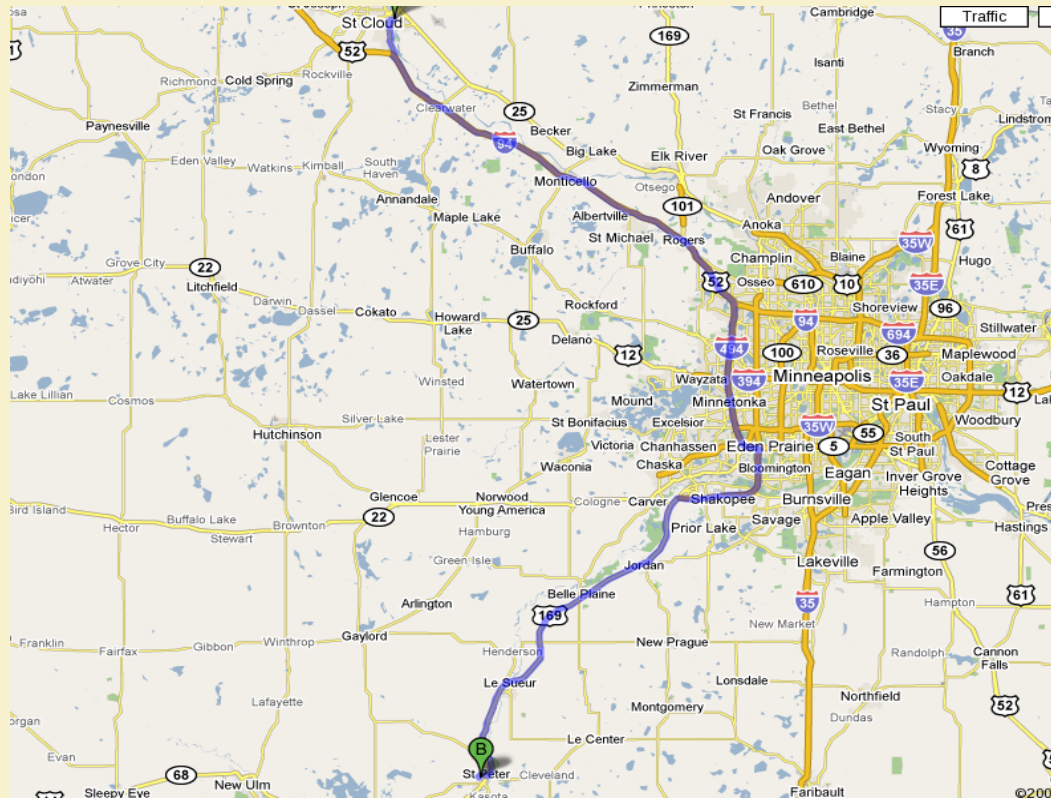
If you have a question
PLEASE ASK!



What is Gustavus Adolphus College

Gustavus is an Undergraduate Liberal Arts College

We are mostly residential with about 2000 of our 2600 students living on campus.





Gustauvs' Bandwidth History

Distant Past

56k leased line

1999 or so (onset of Napster)

we got a Packet Shaper 2500 which could handle 10Mbps

2004

we upgraded our connection to 10Mbps and got a Packet Shaper 6500 which could handle 45Mbps because the 2500 couldn't really handle 10Mbps

2005

we upgraded to 20Mbps

2006

we upgraded to 30Mbps I1 and 15ish I2 (whatever was left of the 45Mbps)



Gustauvs' Bandwidth History

- Fall 2007 (where things start to get difficult)
 - We get a 1Gig Fiber to UMN and 50M I1 from Onvoy
 - We rate limit down to 45M on I1 simplify our rules on the packet shaper 6500 and hope our plan works
- Fall 2007 (a bit later)
 - We start using our Linux Packet Shaper and everything goes fairly well
- Fall 2007 (even later)
 - We upgrade to 100Mbps of I1
- January 2009
 - We upgrade to 150Mbps of I1





Why packet shape at all?

- In 1999 Napster came out - first released in June.
- By December everyone's Internet connection was not nearly fast enough to keep up with demand.

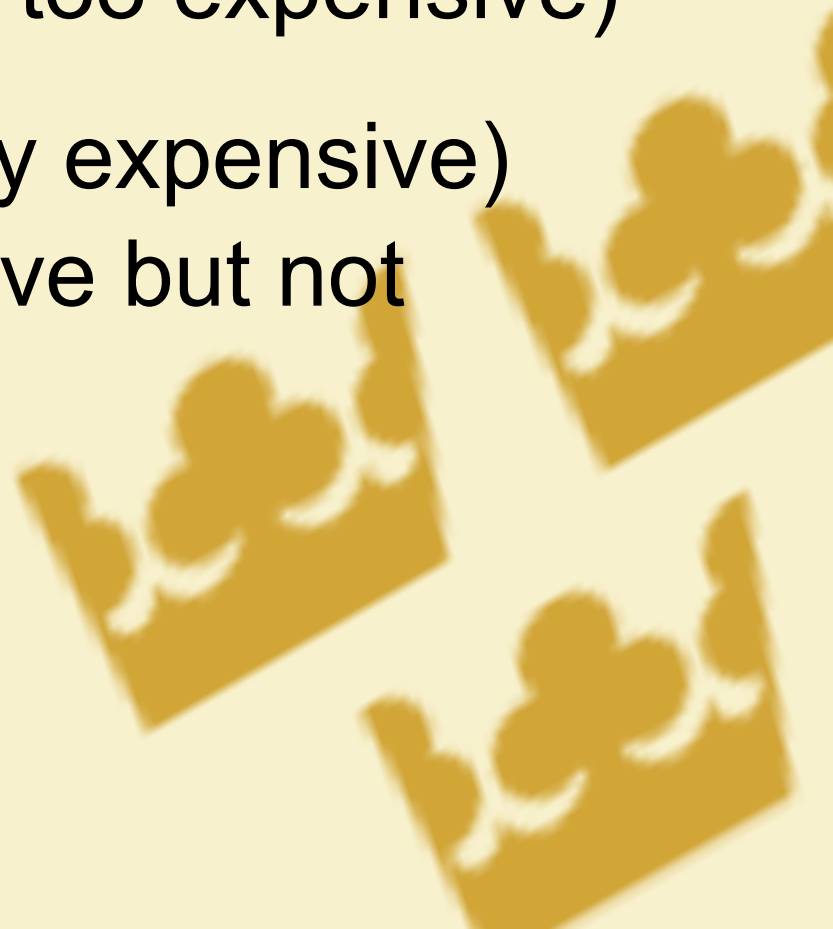




Why packet shape at all?

1999 Bandwidth cost formula

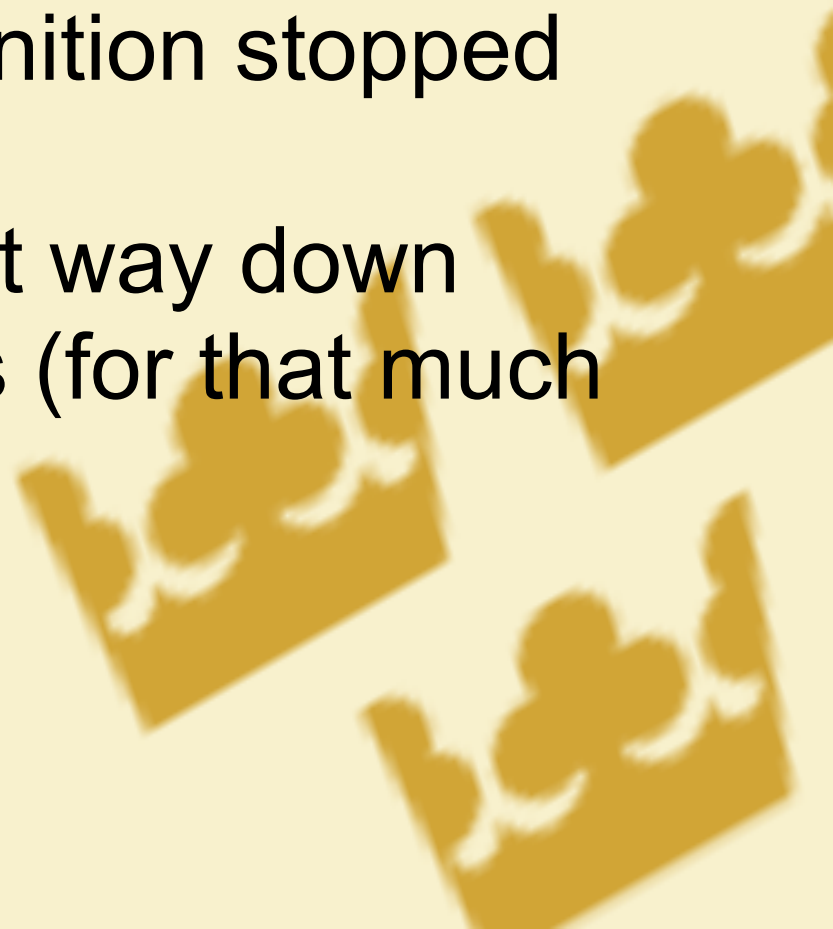
- More bandwidth = \$ (way too expensive)
- Packetshaper = \$ (painfully expensive)
- Block Napster = \$ (expensive but not entirely effective)





What Changed

1. Cat and Mouse game of P2P changing protocols and new P2P programs coming out
2. Layer Seven pattern recognition stopped working (encryption)
3. The cost of bandwidth went way down
4. The cost of packet shapers (for that much bandwidth) went way up.





What Changed

The New Reality

2009 Bandwidth Cost formula

- More bandwidth - \$ (\$7-\$20/Mbps)
- Packet Shaper - \$ (~\$40k for 100Mbps)
- Block P2P - \$ (but probably even less effective and less politically viable)



Why are we shaping again?

What is the problem we are trying to solve?

- Music Piracy?





Why are we shaping again?

What is the problem we are trying to solve?

- Music Piracy?
- Stop our students from getting sued?





Why are we shaping again?

What is the problem we are trying to solve?

- Music Piracy?
- Stop our students from getting sued?
- Administrative hassle of dealing with RIAA?
- Moral sense of "appropriateness of Internet Use"?





Why are we shaping again?

What is the problem we are trying to solve?

- Music Piracy?
- Stop our students from getting sue
- Administrative hassle of dealing with RIAA?
- Moral sense of "appropriateness of Internet Use"?
- Porn?
- Linux CDs?
- Sense of POWER?!





Why are we shaping again?

What is the problem we are trying to solve?

- Make the academic Internet use usable again!

How?

- By keeping users from using more than their share.





So... what do we do now?

Option 1: Do no shaping, buy lots more bandwidth

Cons:

- P2P insatiable
- Become a RIAA Target





So... what do we do now?

Option 2: Buy some more bandwidth, buy new packet shaper

Cons:

- Increasingly expensive
- Cat and Mouse game





So... what do we do now?

Option 3: Buy next gen shaper like NetEnforcer or NetEqualizer

Cons:

- Increasingly expensive (Allot NetEnforcer something like 60k)





So... what do we do now?

Option 4: Create our own Packet Shaper

Cons:

- Requires more staff time to develop/install (?)



How to Create a Packet Shaper



Linux has lots of QoS features built in.

- Classify traffic
 - Block traffic
 - Queue traffic
 - Dequeue traffic



How to Create a Packet Shaper



Classify Traffic

- Based on IP
- Based on Port
- Based on Layer-7 pattern



How to Create a Packet Shaper



Blocking vs Queuing Traffic

Some programs (kazaa was the first I think) will try to get increasingly sneaky if they detect that they are blocked.

If we aren't singling out programs, blocking would involve a hard quota per IP, which would stop academic work.

How to Create a Packet Shaper



What type of queue?

- FIFO
- Priority Queue
- SFQ (Stochastic Fair Queuing)
- CBQ (Class Based Queuing)
- HTB (Hierarchical Token Bucket)



How to Create a Packet Shaper



HTB

- Divides the available bandwidth among several subqueues according to rules



How to Create a Packet Shaper



HTB Rules

- Rate = The amount of bandwidth "guaranteed" for the child queue
- Ceil = The most bandwidth the child queue can use

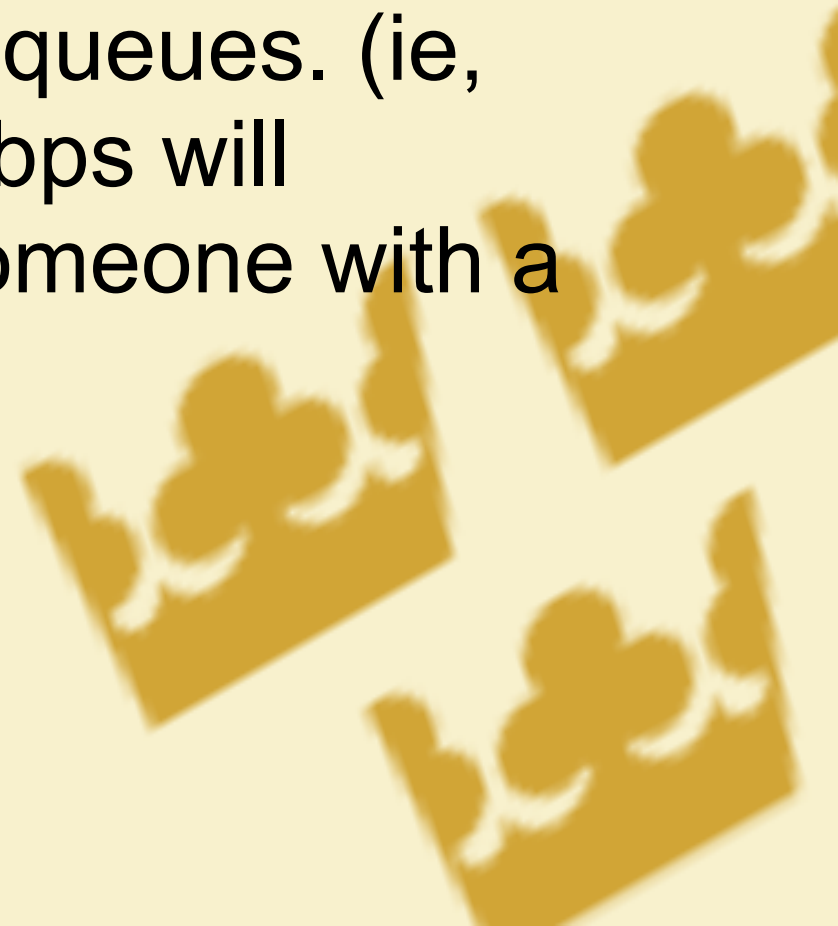


How to Create a Packet Shaper



HTB - What if the link is under or oversubscribed?

- The bandwidth will be divided *in proportion* to the rates of the different queues. (ie, someone with a rate of 1Mbps will download half as fast as someone with a rate of 2Mbps)





What is "Fair"?

1. Each IP gets an even share of the bandwidth at a given time?





What is "Fair"?

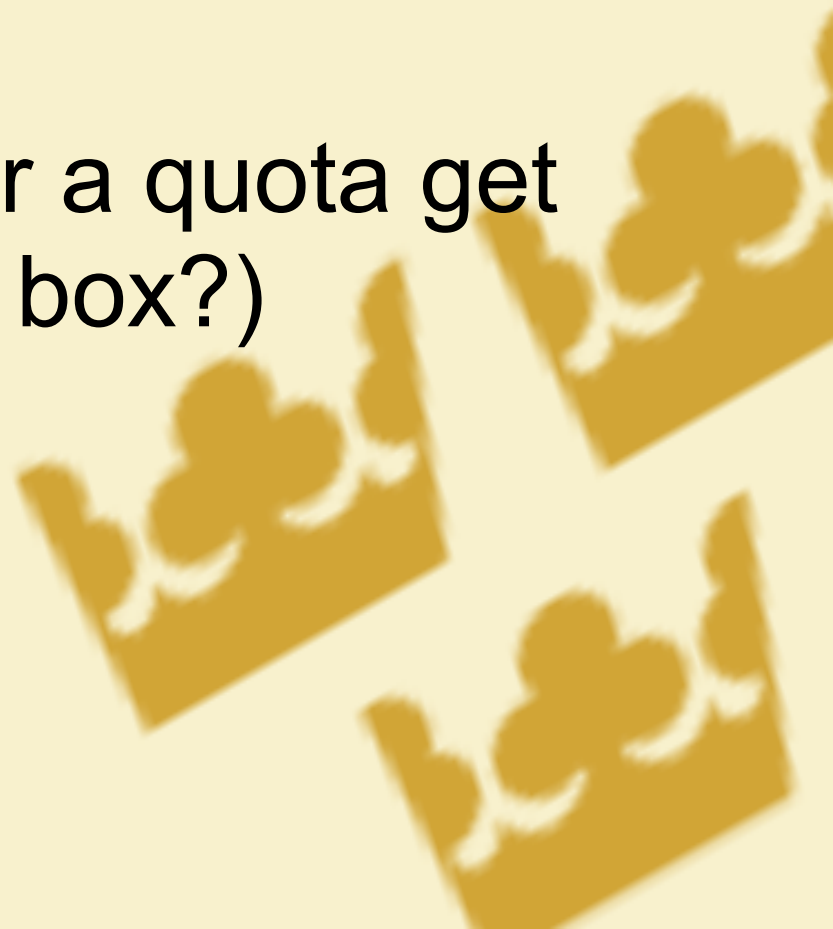
1. Each IP gets an even share of the bandwidth at a given time?
2. Each IP gets an even share over 5 minutes? 10 minutes?





What is "Fair"?

1. Each IP gets an even share of the bandwidth at a given time?
2. Each IP gets an even share over 5 minutes? 10 minutes?
3. IPs which have gone over a quota get less bandwidth? (penalty box?)





What is "Fair"?

1. Each IP gets an even share of the bandwidth at a given time?
2. Each IP gets an even share over 5 minutes? 10 minutes?
3. IPs which have gone over a quota get less bandwidth? (penalty box?)
4. As IPs pass successive thresholds they get successively less bandwidth.

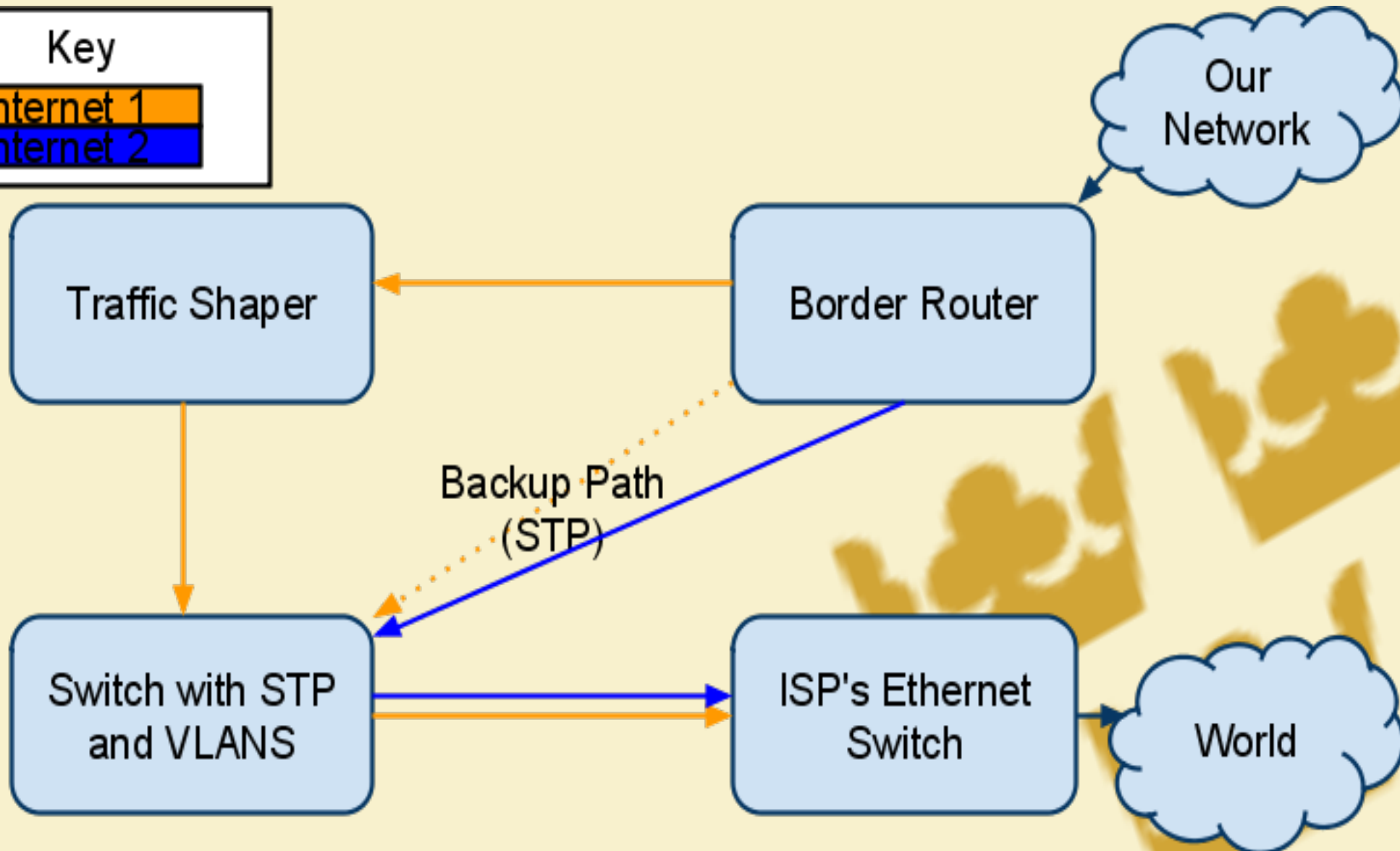
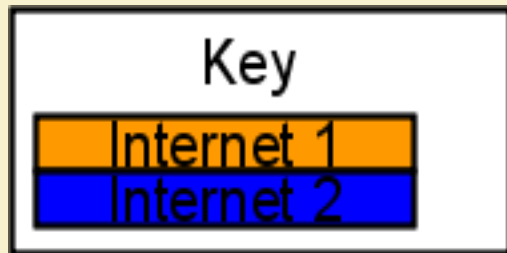


What is "Fair"?

Bandwidth Priority

| If a computer has downloaded | It will download this much for every 1MB which a computer which has downloaded less than 100MB can download |
|-------------------------------------|--|
| 1MB-99MB | 1MB |
| 100MB-499MB | 512KB |
| 500MB-999MB | 256KB |
| 1000MB-1499MB | 128KB |
| 1500MB-1999MB | 64KB |
| 2000MB-2499MB | 32KB |
| 2500MB-2999MB | 16KB |
| 3000MB-3499MB | 8KB |
| 3500MB-3999MB | 4KB |
| 4000MB+ | 2KB |

Network diagram!





Steps and Software

- Linux Bridging (brctl)
- iptables to count bytes for each IP
- perl/mysql to log the total bytes for each IP every 15 minutes
- Linux QoS (tc) to control the flow of traffic





Steps and Software

Linux Bridging (brctl)

- We set up the "traffic shaper" computer with 3 network cards.
 - In
 - Out
 - Management
- We then set up a bridge between In and Out. This way if there is a problem, we can unplug In and Out and put in a RJ-45 coupler.
 - We've since automated using STP



Steps and Software

IPTables

- We set up IPTables "chains" for each class C subnet in each direction (in and out)
- Within each chain there is a rule which "marks" the packet

```
"-A GAC-1-OUTBOUND -s 138.236.1.234 -i  
br0 -j MARK --set-mark 0x1ea1"
```

- IPTables then counts how many bytes have matched the rule, and we can use that data to...



Steps and Software

Perl/Mysql to log totals

- Every 15 min. we run a script which runs
`"iptables -t mangle -L -vnx"`
which outputs the total bytes which matched each rule.
- Then we enter that data into a mysql database to calculate how much data an ip has transferred in the last 24hours using this SQL statement:

```
SELECT ip_c,ip_d, max(time) maxtime, min(time) mintime, max(inbound)-min(inbound) as inbound, max  
(outbound)-min(outbound) as outbound, inspeed, outspeed  
FROM netusage nu LEFT JOIN current_prio cp on cp.c=nu.ip_c and cp.d=nu.ip_d WHERE  
(`time` > NOW()-INTERVAL 1 DAY) AND ip_c!=128  
GROUP BY ip_d,ip_c ORDER BY ip_c,ip_d
```



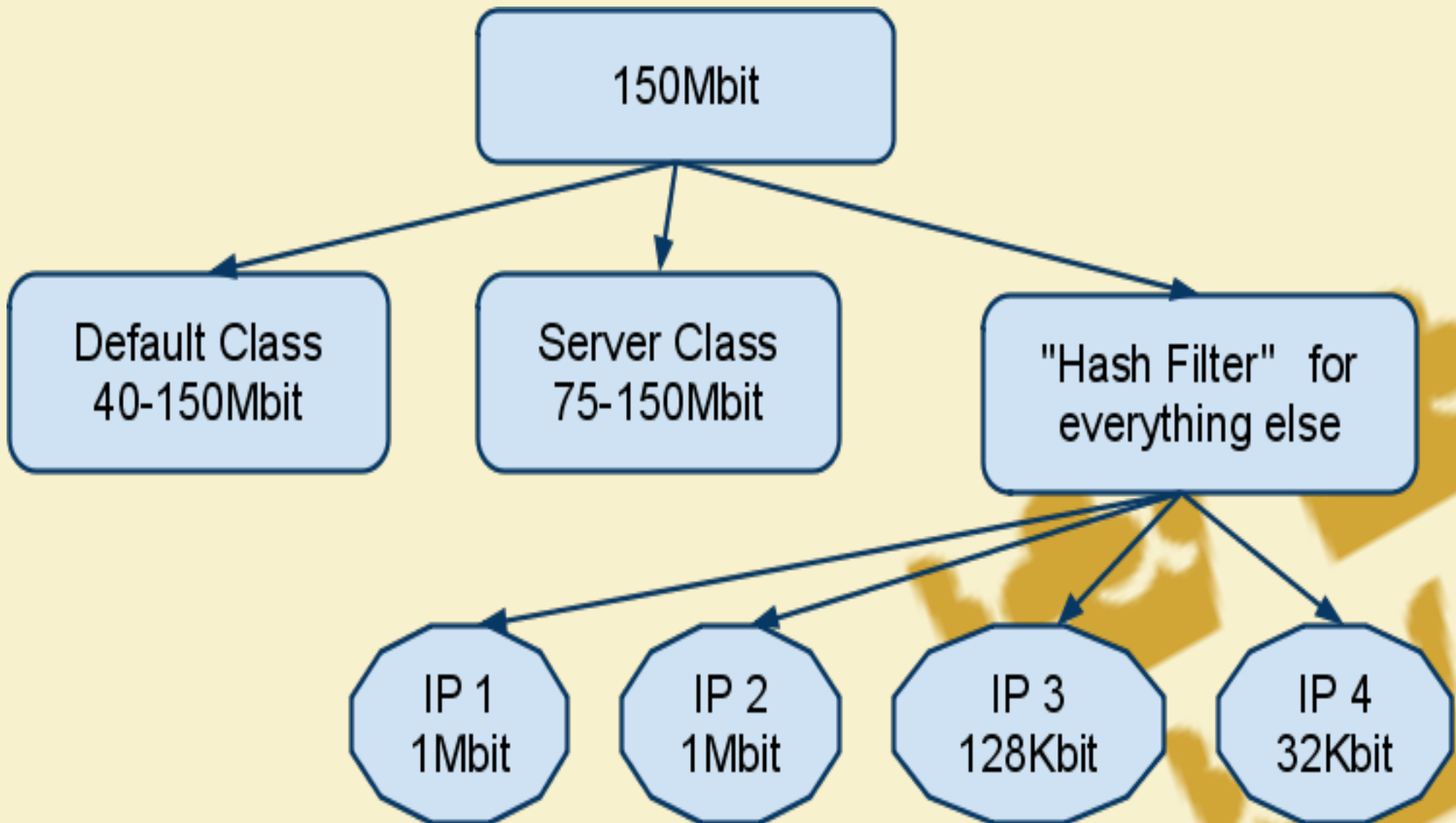
Steps and Software

Use "tc" to update the traffic shaping rules

A few tips

- We needed to use a "hash filter" so that lookup of which rule the packet went to used a hash table rather than a sequential search. (Otherwise it was too slow.)
- We wound up creating a "default class" for all the IPs that had sent less than 75M in the past 24 hours.

Tree diagram!





Tweaks

- Limit upload speeds to $20 * \text{rate}$ or 64kbps, whichever is higher.
- Allow for exceptions (eg streaming video from campus)
- A page to let you look up how much bandwidth you've used today and how that compares to others



Tweaks



Bandwidth Usage for 138.236.250.42

Logged in user: sommere with administrator privileges

Your [bandwidth](#) usage







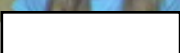



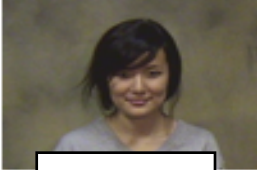





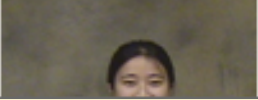

| | Inbound | Outbound |
|--|----------|----------|
| Your usage | 42 MB | 5 MB |
| Your priority in comparison to someone who does not use much bandwidth | 1024 KB | 1024 KB |
| % of total campus usage | 0.074 % | 0 % |
| Average usage | 107.1 MB | 40 MB |
| Total users our connection could support if everyone used the Internet like you | 19048 | 160000 |

Note: Totals exclude [Internet2](#) bandwidth usage.

Tweaks



Top bandwidth Users Inbound

| | | | |
|---|--|---|---|
| 1 | 2 | 3 | 4 |
|  |  |  |  |
|  |  | sommere |  |
| 138.236.62.62 | 138.236.224.87 | 138.236.71.64 | 138.236.253.112 |
| Prio: 2KB | Prio: 16KB | Prio: 20000KB | Prio: 128KB |
| 6401MB In 130MB Out | 2716MB In 185MB Out | 1259MB In 43MB Out | 1214MB In 19MB Out |
| 5 | 6 | 7 | 8 |
|  |  |  |  |
|  |  |  |  |
| 138.236.250.127 | 138.236.253.131 | 138.236.40.151 | 138.236.230.72 |
| Prio: 128KB | Prio: 128KB | Prio: 128KB | Prio: 128KB |
| 1150MB In 36MB Out | 1114MB In 25MB Out | 1082MB In 31MB Out | 1054MB In 21MB Out |
| 9 | 10 | 11 | 12 |
|  |  |  | |



Specs



Dual Intel Xeon 2.8Ghz (hyperthreading)

2G RAM

- Aprox cost \$2k



Not Vaporware!!!



Our source code is available **RIGHT NOW**

- <http://code.google.com/p/bandwidthfairness/>



Questions?



Contact info:

Ethan Sommer

Associate Director of Core Services

Gustavus Adolphus College

sommere@gustavus.edu

507-933-7042

<http://www.resnetsymposium.org/rspm/evaluation/>

