Agenda

- Transport Layer
- Network Layer
- Link Layer
CUBIC Fixes

- BIC → CUBIC (unfairness where $RTT$ is small)

- CUBIC Fixes and Tuning
  - Fix time resolution bugs where $HZ < 1000$ (HR Timers)
  - ACK train delta now a parameter
  - Commit 6b3d626321c
IW10

- `#define TCP_DEFAULT_INIT_RCVWND 10`
- Commit 442b9635c569 (`#define TCP_INIT_CWND 10`)
- Via dst metrics cache modifiable
MD5 for Sequence Numbers

- ISNs not guessable
- Computers have become a lot faster
- MD5 is a safer hash function as MD4
IPsec Extended Sequence Numbers

- IPsec extended (64-bit) sequence numbers for ESP
- RFC 4303 (December 2005)
- Userspace tools need modifications too (see iproute2 package)
Team Network Device

- Bonding “replacement”
  - Fast, simple, userspace-driven
- Netlink socket for communication (not sysfs)
- Planned support for 802.3ad (IEEE 802.3ad Link Aggregation Control Protocol)
PPTP Support

- Point-to-Point Tunneling Protocol
- Dramatically speeds up PPTP VPN connections (compared to userspace poptop/pptpclient)
- Example: High-Performance PPTP NAS
- 00959ade36acadc0
Random Early Drop

- Drop packets before queue is full: pro-actively avoid queue overruns

- RED maintains an exponentially-weighted moving average of the queue length which it uses to detect congestion

- To be effective the router requires buffer space that amounts to twice (see buffer bloat debate) the bandwidth-delay product (adds considerable end-to-end delay and delay jitter)

- Configuration not simple and error prone
SFB

- Perform queue management based directly on packet loss and link utilization (rather average queue lengths)
- If the queue is continually dropping packets due to overflow: increase packet drop/mark probability
- If the queue becomes empty: decrease packet drop/mark probability
- `tc qdisc add dev $dev root sfb`
Shaping, Scheduling and Policing

- Random Early Detection (RED and GRED)
- Stochastic Fair Blue (SFB)
- Stochastic Fairness Queueing (SFQ)
- Generic Random Early Detection (GRED)
- CHOose and Keep responsive flow scheduler (CHOKE)
- Class Based Queueing (CBQ)
- Hierarchical Token Bucket (HTB)
- Token Bucket Filter (TBF)
- Hierarchical Fair Service Curve (HFSC)
- Quick Fair Queue scheduler (QFQ)
- Netem
Berkeley Packet Filter

- Kernel side packet filter functionality (e.g. tcpdump, wireshark)
- Provides filter functionality (e.g. host 192.168.20.0 and TCP)
- Since April 2011: JIT Compiler (for x86_64)
- Default disabled (enable via echo 1 >/proc/sys/net/core/bpf_jit_enable)
Thank You!

► Any questions?

► hagen@jauu.net

► GnuPG Key-ID: 0x98350C22