The CHART System: A High-Performance, Fair Transport Architecture Based on Explicit-Rate Signaling

Andy Bavier Princeton University
acb@cs.princeton.edu

Jack Brassil HP Laboratories
jack.brassil@hp.com

Brian Mark George Mason University
bmark@gmu.edu

Rick McGeer HP Laboratories
rick.mcgeer@hp.com

Raj Rajagopalan HP Laboratories
raj.rajagopalan@hp.com

Larry Roberts Anagran Inc.
lroberts@anagran.com

ABSTRACT

TCP/IP is known to have poor performance under conditions of moderate to high packet loss (> 1%) and end-to-end latency (> 100 ms). The HP CHART system uses a network of flow routers and a proxy-based explicit rate aware TCP/IP stack to improve end-to-end performance in these cases. In tests over such highly impaired links, CHART demonstrated a 20x performance improvement over legacy TCP/IP while preserving fairness to legacy flows.

Categories and Subject Descriptors: C.2.2: Network Protocols, C.2.5: Local and Wide-Area Networks

General Terms: Performance, Reliability

Keywords: TCP/IP, Quality of Service, Flow Management

1. THE CHART SYSTEM

TCP/IP was designed for transport over relatively low-loss, low-latency connections. The resulting protocol suite used in-band sensing of latency and packet loss as its primary tool for congestion avoidance. The result is a dramatic falloff in TCP/IP performance as end-to-end latency increases, and a catastrophic loss in throughput under conditions of corruption loss.

There have been various attempts to rectify this problem, from adjusting the congestion control algorithm in TCP/IP to special-purpose solutions for high-latency and loss links. Examples of the former include FAST-TCP[2], TCP Vegas[3] and Probe Control Protocol (PCP) [4]. Examples of the latter include adaptive forward error correction and performance enhancing proxies. These approaches have explored the space of solutions defined by in-network sensing and static information about specific links. The CHART system is an exploration of the efficacy of explicit communication between the network and the endpoints. Under the TIA-1039 protocol implemented in CHART[1], after every 128 packets each end-system receives a transmission rate from the network and then transmits at the granted rate. CHART consists of:

1. An Explicit-Rate option to IP[1], TIA-1039. The ER option permits a flow to request and receive a guaranteed or best-available transmission rate along the end-to-end path.
2. An explicit-rate transport protocol, TCP-Trinity, which exploits TIA-1039 to transmit at a fixed rate.
3. An in-network flow management service, which implements the TIA-1039 to signal the transmission rate. There are two alternate implementations, both of which allocate bandwidth to individual flows: a hardware flow-management infrastructure appliance for high-bandwidth links, and a software implementation based on the Click[5] platform.

The software and hardware flow-management components are alternate implementations of in-network active queue management with rate information communicated through the TIA-1039 protocol to the endpoints. Flow management devices, which are placed at network congestion points, differentiate between traffic classes, enhancing the end-to-end performance of high priority traffic.

A common decentralized network measurement and monitoring fabric communicates network state to CHART components. This sensing infrastructure aggregates and propagates measurements collected by monitors placed in the network Measurements include link transmission capacity, loss rate, bandwidth, latency, and hop count. We anticipate using the sensing infrastructure as an alternate oracle to in-network flow management.

In DARPA testing in the fall of 2006 and the fall of 2007 administered by SAIC, the CHART system demonstrated mean performance gains of 20x on a large test suite. SAIC testing in the fall of 2007 demonstrated that the presence of CHART traffic did not interfere with the flow of legacy UDP or TCP traffic.

2. ADDITIONAL AUTHORS

Stephen Schwab, Sparta Inc., `stephen.schwab@sparta.com`, Puneet Sharma, HP Labs, `puneet.sharma@hp.com`, Soony-ong Sohn, George Mason University, `ssohn@gmu.edu`, John Spies, HP, `john.spies@hp.com`, Nicki Watts, HP, `nicki.watts@hp.com` and Praveen Yalangadula, HP Labs, `praveen.yalangadula@hp.com`

3. REFERENCES

[1] L. Roberts, *TIA TR-34.1.02/12.04.05: QoS Signalling for IPv6 QoS Support*, TIA TR-34.1.7 Working Group (IP on Satellite), 2005.

[2] D. X. Wei, C. Jin, S. H. Low, and S. Hegde, *FAST TCP: motivation, architecture, algorithms, performance*, IEEE/ACM Trans. Networking 14.

[3] L. S. Brakmo, L. L. Peterson *TCP Vegas: End to End Congestion Avoidance on a Global Internet* IEEE Journal on Selected Areas in Communications.

[4] Thomas Anderson, Andy Collins, Arvind Krishnamurthy, and John Zahorjan, *PCP: Efficient Endpoint Congestion Control* NSDI, 2006.

[5] Eddie Kohler, Robert Morris, Benjie Chen, John Jannotti, and M. Frans Kaashoek, *The Click modular router*, ACM Transactions on Computer Systems 18(3), August 2000

SIGCOMM'08, August 17–22, 2008, Seattle, Washington, USA.
ACM 978-1-60558-175-0/08/08. This work was supported in part by DARPA Contract N66001-05-9-8904 (Internet Control Plane).