

Microsoft's Datacenter TCP (DCTCP): TCP Congestion Control for Datacenters

draft-bensley-tcpm-dctcp-05

Stephen Bensley, Lars Eggert, Dave Thaler,
Praveen Balasubramanian and Glenn Judd
TCPM, IETF-93, Prague, Czech Republic
July 22, 2015

RAND-Z IPR disclosure! <https://datatracker.ietf.org/ipr/2319/>

DCTCP recap

- **TCP variant for datacenters**
- **Goals**
 - Low latency for short flows
 - High throughput for long flows
- **Approach**
 - Use ECN to quantify the extent of congestion
 - Scale `cwnd` proportionally

Known DCTCP implementations

- **Microsoft Windows Server 2012**
- **FreeBSD-CURRENT** (by Midori Kato, Keio)
 - <http://lists.freebsd.org/pipermail/freebsd-net/2014-February/037915.html>
 - Conforms to draft (plus some variants; ODCTCP)
- **Linux 2.6.38.3** and **ns2** (by Stanford & NEC)
 - <http://simula.stanford.edu/~alizade/Site/DCTCP.html>
 - Unclear conformance to draft

DCTCP momentum

- More implementations
 - **Linux kernel patch** (By Borkmann, D. and F. Westphal)
<https://git.kernel.org/cgit/linux/kernel/git/davem/net-next.git/commit/?id=e3118e8359bb7c59555aca60c725106e6d78c5ce>
 - **Windows Server 2012 R2**
 - **Windows Server 2016 (Technical Preview 2 / May 4, 2015)**
- More deployments
 - **Morgan Stanley**
<https://www.usenix.org/conference/nsdi15/technical-sessions/presentation/judd>
 - **Under evaluation at Microsoft and NetApp**
- Academia research interest
- Other TCP proposals based on ideas from DCTCP
 - Accurate ECN
 - ODCTCP
 - DualQ Coupled AQM

Major changes since Revision 0

- Improved draft for completeness & clarity
 - Handling of SYN/RST/“pure ACK” packets
 - Co-existence with conventional TCP
 - Other TCP changes for datacenter deployment
 - Starting to address loss handling
- Addressed feedback from tcpm mailing list
 - Ongoing effort

Next Steps

- We think the draft is near completion
 - Open to suggestions
 - Open to working with tcpm to address feedback
- Request tcpm to consider publishing this as Informational RFC
 - Accelerate deployments
 - Improve interop