Retrospective on "Measured Capacity of an Ethernet: Myths and Reality"

Jeffrey C. Mogul
jeffmogul@acm.org

Christopher A Kantarjiev
cak@acm.org

This article is an editorial note submitted to CCR. It has NOT been peer reviewed.
The authors take full responsibility for this article's technical content. Comments can be posted through CCR Online.

ABSTRACT

The original Ethernet design used CSMA/CD on a broadcast cable. Even after it became commercially popular, many people expressed concerns that Ethernet could not efficiently use the full channel bandwidth. In our 1988 paper, "Measured Capacity of an Ethernet: Myths and Reality," we reported on experiments we ran showing that, even under relatively heavy loads, Ethernet typically still performed well. We describe the context in which we ran those experiments, and some subsequent research conducted by others.

1 ETHERNET IN PRACTICE AND THEORY

Ethernet was the first widely-successful LAN technology, and in some ways, it is also the last wired LAN standing. Between its invention in the 1970s at Xerox PARC [4] and today, however, Ethernet has changed in many ways, informed by both practical experience and theoretical analysis. Sometimes theory and experience do not agree.

When we published "Measured Capacity of an Ethernet: Myths and Reality" [3] in 1988[1], Ethernet had already achieved considerable commercial success, probably because it was designed to be "relatively inexpensive" yet scalable to "several buildings full of personal computers." [4]. The design relied on a passive, branching broadcast medium (coaxial cable) and stations that implemented a distributed packet-switching protocol using Carrier-Sense Multiple Access with Collision Detection (CSMA/CD).

The Aloha wireless network [1], starting in 1968, had introduced the use of a multi-access broadcast channel for packet-based networking, but did not have either Carrier Sense or Collision Detection. With Carrier Sense, an Ethernet station could avoid transmitting until it sensed that the channel was idle. Two or more stations could sometimes, with this mechanism, start transmitting at about the same time; this "collision" corrupts the packets as they interfere on the broadcast medium. Collision Detection allows the stations to stop transmitting as soon as possible, to avoid wasting an entire packet's duration. When the channel is contended, as discovered via collision detection, a CSMA/CD Ethernet station uses a "binary exponential random backoff" algorithm to delay its subsequent attempt to acquire the channel; therefore, it transmits with probability $1/Q$ when there are Q stations contending simultaneously. In effect, the stations cooperate to fairly service a queue of depth Q pending packets.

Metcalfe and Boggs [4] provided a simple analysis explaining how this queue-estimation works. They calculated that the theoretical efficiency of a highly-loaded Ethernet ($Q = 256$) varies from 95%, for large packets, down to Aloha's asymptotic efficiency of $1/e$ for minimum-size packets. (Efficiency, in this case, is defined as the ratio of goodput to the maximum possible bit-rate of the network.) They observed that such high loads were unlikely (in fact, packet headers in their original design carried only 8-bit addresses, so $Q = 256$ would be unlikely in practice). A workload of only minimum-size packets was also unlikely.

In subsequent years, various papers studied the performance of Ethernet, using both theory and simulation. Some of these papers suggested that CSMA/CD Ethernet would perform poorly under various scenarios; we refer readers to our 1988 paper for a survey of these papers.

While our paper implicitly criticized some prior publications for using unrealistic assumptions, we recognized these were all serious, peer-reviewed papers by skilled researchers. However, for many years, Dave Boggs had on his office door a copy of a letter from a physicist asserting that he had proved that Ethernet could not possibly work. Dave was quite proud of that letter.

2 OUR EXPERIMENTS

We perceived that, perhaps as a result of some of these studies (and *perhaps* due to the marketing efforts of companies with competing LAN technologies) many believed that "Ethernets saturate at an offered load of 37%."[2] Since we had some experience to the contrary, we decided to perform our own experiments.

We were lucky to be working at the Digital Equipment Corporation Western Research Lab (DECWRL), whose first major project was the design of the Titan RISC processor [5], which ran at about 15 MIPS – quite fast for the time, especially since we treated these mostly as single-user personal computers. (Titans ran a weird version of UNIX, but for these tests we used a lightweight operating system designed for diagnostics, and so our experiments avoided most of the overheads of the typical network stacks of the day, which were not nearly as clever as they are now.)

We were able to run experiments on a set of 24 Titans connected to a 10Mbit/sec Ethernet, the original standard. We ran a load generator that varied the length of its packets between 64 bytes

[1]David Boggs was the lead author of that paper, but we have been unable to reach him to obtain his contribution to this article.

Jeff Mogul is currently affiliated with Google LLC. This article represents his own opinion and does not represent the views of Google.

[2]This belief appears to stem from a mis-reading of the simplified model analyzed in the original paper [8].

(the minimum) and 4000 bytes (well over the standard maximum, but it was our private network, with hardware we had designed). We tried using both short and long cables for the shared Ethernet medium (because propagation delay affects how quickly a station can detect a collision), and varied the number of hosts between 2 and 24.

Our paper concluded that "our measurements, and a careful reading of the theoretical analyses, show that Ethernet is capable of good performance for typical applications, even at high offered load." That is, prior suggestions that CSMA/CD Ethernet was inefficient or unstable were myths. We also showed that some of the prior reports of poor performance were likely due to implementation problems, including bugs that caused "broadcast storms." One downside of Ethernet's "distributed packet switching" design was, we said, that Ethernets are also "distributed single points of failure," because a hardware or software fault could make the entire network useless.

3 SOMETHING WE MISSED

One metric we reported in our paper was the unfairness between senders, measured as the standard deviation of the senders' achieved bit rates; we noted that "When $N = 3$, for example, there is a high probability that one host will continually defer to the other two for several collision resolution cycles, and the measured standard deviation becomes a sizeable fraction of the mean bit rate." In retrospect, we might have been seeing evidence of the "capture effect" identified by Ramakrishnan and Yang in 1994 [7], in which "a station transmits consecutive packets exclusively for a prolonged period despite other stations contending for access." However, we did not further investigate this aspect of our results, so we completely failed to understand the capture-effect mechanism, or how to avoid it; if we had, it would have saved our employer the cost of redesigning an Ethernet NIC chip.

4 ETHERNET TODAY: GRANDFATHER'S AXE

Today, most wired LANs, including most warehouse-scale datacenter networks, use something we call "Ethernet," even though today's switched Ethernets share only a few things with the CSMA/CD Ethernet we used in 1988. Like the proverbial "grandfather's axe," almost every mechanism in the original Ethernet design has been replaced with a new part, but it's still Ethernet.

What has been preserved are the aspects most valuable to software stability and network interoperability: the basic format of packet headers and the specific format of station addresses. Because these formats have stayed largely the same, you can connect systems that were designed for various generations of Ethernet and they can easily communicate without much concern for compatibility. (We're ignoring innovations such as VLAN headers, but these are intentionally invisible to most of the network software stack, and jumbo frames, which increase efficiency but can be tricky to deploy incrementally.)

What has changed is that the world has moved to physical, rather than distributed packet switches – CSMA/CD is no more.[3] Centralized packet switches, and complex networks with many switches, required numerous innovations (including the Spanning Tree Protocol [6] or Clos networks [2], and high-speed multi-port switch chips) but they largely freed us from the limited scaling potential of a shared broadcast channel, the $1/e$ efficiency limit for small packets, and perhaps most important, the "distributed single point of failure" that could cause an entire network to fail when someone improperly drilled into the coaxial cable while installing a "vampire tap."[4]

5 SUMMARY

The measurements in our 1988 paper are today mostly of historical interest for wired LANs (and the techniques used by modern wireless LANs to achieve high performance are quite different from CSMA/CD). Perhaps, though, our work helped Ethernet make it through its early years to its current domination.

[1] Norman Abramson. 1970. THE ALOHA SYSTEM: Another Alternative for Computer Communications. In *Proceedings of the November 17-19, 1970, Fall Joint Computer Conference (AFIPS '70 (Fall))*. 281–285.

[2] Mohammad Al-Fares, Alexander Loukissas, and Amin Vahdat. 2008. A Scalable, Commodity Data Center Network Architecture. In *Proceedings of the ACM SIGCOMM 2008 Conference on Data Communication (SIGCOMM '08)*. 63–74.

[3] D. R. Boggs, J. C. Mogul, and C. A. Kent. 1988. Measured Capacity of an Ethernet: Myths and Reality. In *Symposium Proceedings on Communications Architectures and Protocols (SIGCOMM '88)*. 222–234.

[4] Robert M. Metcalfe and David R. Boggs. 1976. Ethernet: Distributed Packet Switching for Local Computer Networks. *Commun. ACM* 19, 7 (July 1976), 395–404.

[5] Michael J. K. Nielsen. 1986. *Titan System Manual*. Research Report 86/1. Digital Equipment Corporation Western Research Laboratory.

[6] Radia Perlman. 1985. An Algorithm for Distributed Computation of a Spanningtree in an Extended LAN. In *Proceedings of the Ninth Symposium on Data Communications (SIGCOMM '85)*. 44–53.

[7] K. K. Ramakrishnan and H. Yang. 1994. The Ethernet capture effect: analysis and solution. In *Proceedings of 19th Conference on Local Computer Networks*. 228–240.

[8] Charles E. Spurgeon and Joann Zimmerman. 2014. *Ethernet: The Definitive Guide: Designing and Managing Local Area Networks*. O'Reilly Media, Inc.

[3] 802.11 WiFi uses CSMA/CA – for "Collision *Avoidance*" – with various recent versions using techniques such as frame aggregation and MIMO to reduce the drawbacks of CSMA/CA.

[4] A 1988 Ethernet was a typically a thick yellow coaxial cable. Stations were attached via "vampire taps," which, after a careful installation procedure that involved a special tool, had an insulated central spike that pierced the cable's core conductor, and shorter spikes that pierced just the shield conductor. A single-drop cable connected each tap to a NIC. Vampire taps allowed adding stations without having to cut the cable.