Deep packet inspection meets 'Net neutrality, CALEA

By Nate Anderson | Published 4 years ago

Throttle me this: An introduction to DPI

Imagine a device that sits inline in a major ISP's network and can throttle P2P traffic at differing levels depending on the time of day. Imagine a device that allows one user access only to e-mail and the Web while allowing a higher-paying user to use VoIP and BitTorrent. Imagine a device that protects against distributed denial of service (DDoS) attacks, scans for viruses passing across the network, and siphons off requested traffic for law enforcement analysis. Imagine all of this being done in real time, for 900,000 simultaneous users, and you get a sense of the power of deep packet inspection (DPI) network appliances.

Although the technology isn't yet common knowledge among consumers, DPI already gives network neutrality backers nightmares and enables American ISPs to comply with CALEA (government-ordered Internet wiretaps) reporting requirements. It also just might save the Internet (depending on who you believe).

Ars recently had the chance to talk with executives from DPI vendors Ellacoya and Procera Networks about their offerings and how they are already being deployed around the world, and we got a look at the newest boxes on offer from each company. Their top-of-the-line products can set you back several hundred thousand dollars, but some of them can inspect and shape every single packet—in real time—for nearly a million simultaneous connections while handling 10-gigabit Ethernet speeds and above.

That's some serious horsepower, and when major ISPs deploy these products in their networks, they suddenly know a whole lot more about their users and their traffic. They also gain the ability to block, shape, monitor, and prioritize that traffic—in any direction. That makes it suddenly simple to, say, prioritize all incoming traffic from any web site that has handed over a briefcase stuffed with unmarked bills while leaving every other site to fight its way through the tubes as best it can.

It also becomes trivial to start blocking or actively degrading services that a company dislikes—like VoIP, for example. Not that this would ever happen. But that's not how the technology is marketed, and there's little evidence that it's currently being used this way. DPI is generally sold on the premise that network operators can control entire classes of traffic (P2P, VoIP, e-mail, etc.) on a group or per-user basis. Let's take a look at how that happens and what it means for both network neutrality and legal interception (CALEA) compliance.

Inspecting packets, deeply

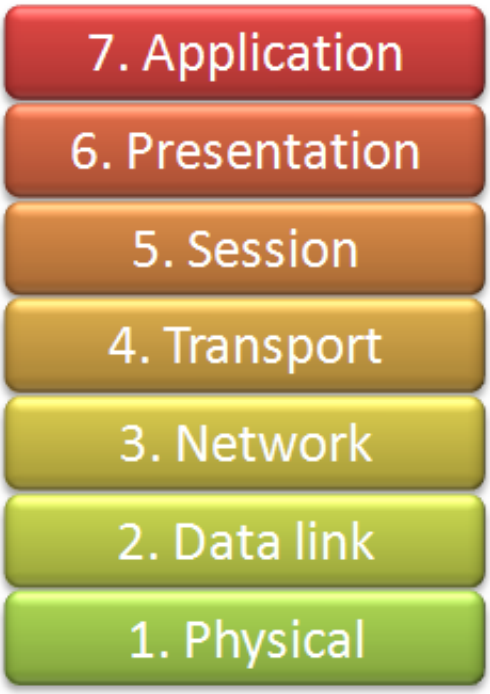
The "deep" in deep packet inspection refers to the fact that these boxes don't simply look at the header information as packets pass through them. Rather, they move beyond the IP and TCP header information to look at the payload of the packet. The goal is to identify the applications being used on the network, but some of these devices can go much further; those from a company like Narus, for instance, can look inside all traffic from a specific IP address, pick out the HTTP traffic, then drill even further down to capture only traffic headed to and from Gmail, and can even reassemble e-mails as they are typed out by the user.

But this sort of thing goes beyond the general uses of DPI, which is much more commonly used for monitoring and traffic shaping. Before an ISP can shape traffic, it must know what's passing through its system. Without DPI, that simple-sounding job can be all but impossible. "Shallow" packet inspection might provide information on the origination and destination IP addresses of a particular packet, and it can see what port the packet is directed towards, but this is of limited use.

Shallow inspection doesn't help much with modern applications, especially with those *designed* to get through home and corporate firewalls with a minimum of trouble. Such programs, including many P2P applications and less-controversial apps like Skype, can use many different ports; some can even tunnel their traffic through entirely different protocols.

So looking at the port doesn't give ISPs enough information anymore, and looking just at the IP address can't identify P2P traffic, for instance. Even for applications like web browsers that consistently use port 80, more information is needed. How much of that HTTP traffic is video? Ellacoya, which recently completed a study of broadband usage, says that 20 percent of all web traffic is really just YouTube video streams.

This is information an ISP wants to know; at peak hours, traffic shaping hardware might downgrade the priority of all streaming video content from YouTube, giving other web requests and e-mails a higher priority without making YouTube inaccessible.



OSI layer model

This only works if the packet inspection is "deep." In terms of the OSI layer model, this means looking at information from layers 4 through 7, drilling down as necessary until the nature of the packet can be determined. For many packets, this requires a full layer 7 analysis, opening up the payload and attempting to determine which application generated it (DPI gear is generally built as a layer 2 device that is transparent to the rest of the network).

Procera explains the need for this approach in marketing materials, saying that "layer 7 identification is a necessity today when most client software, like P2P file sharing, is customizable to communicate over any given port to avoid traditional port-based firewalls and traffic management systems."

But how does this work? Data packets don't often contain metadata saying that they were generated from eDonkey; the DPI appliances need to figure out this out. In real-time. For hundreds of thousands of simultaneous connections.