AGGREGATION AT THE PHYSICAL LAYER

HOWARD M. FRAZIER

Aggregation at the physical layer (APL) allows a single IEEE 802.3 media access control (MAC) sublayer to treat a collection of underlying physical layer (PHY) devices as a single logical link. APL creates Ethernet links with bandwidths greater than can be achieved with a single physical link. The primary purpose of APL is to support the next generation Ethernet link speeds of 40 and 100 Gb/s.

At the most basic level, APL decimates Ethernet packets into fragments, and distributes the fragments across a collection of physical interfaces. A sequence number identifies each fragment, allowing a receiving APL sublayer to reconstruct and deliver packets to the receiving MAC without altering the temporal order of a sequence of packets.

APL borrows somewhat from the physical medium entity aggregation function (PAF) defined in Clause 61 of IEEE Std 802.3ah. The APL sublayer is located below the MAC in the International Organization for Standardization (ISO) seven-layer reference model. APL resides just below the reconciliation sublayer, as shown in Fig. 1.

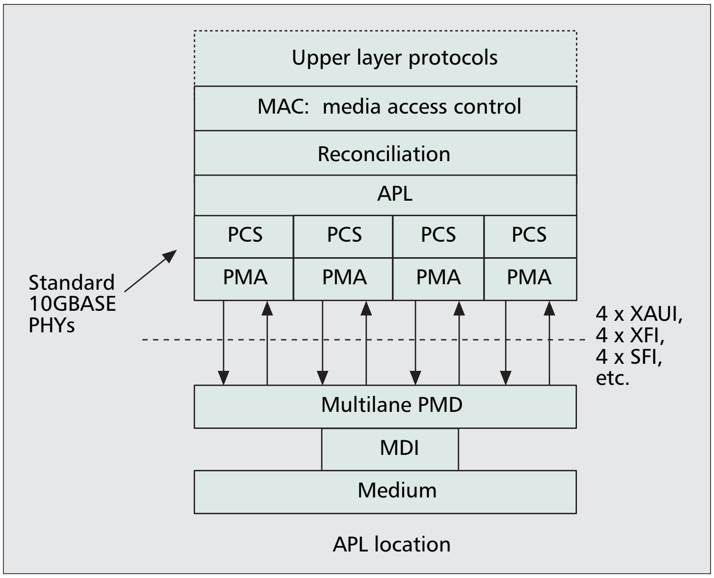

Figure 1.

APL is designed to be compatible with all extent 10GBASE PHYs, as well as all extent 10G compatibility interfaces, such as XGMII, XAUI, XFI, and SFI. APL will also be compatible with any newly defined PHYs, as may be specified for 100 Gb/s operation. In addition, APL preserves and transports the Ethernet preamble sequence unchanged, thus supporting various proprietary and standards-based uses of the preamble pattern.

Fragments can be of variable size, up to a maximum of 256 bytes in length. Fragment sizes are always a multiple of 8 bytes, which provides maximum compatibility with the 64b/66b encoding method used for 10GBASE-R.

Each fragment includes a header field and a trailer field. The length of the header field is 4 bytes, and the length of the trailer field is 1 byte. The header field includes a sequence number and a CRC-8 field to protect the fragment header. After decimating a packet into 256-byte fragments, the APL sublayer encapsulates any remaining portion of the packet in a single fragment that is a multiple of 8 bytes in length.

An important concept of APL is that fragments are delimited just like packets as they are transmitted across any given compatibility interface, through any given physical layer, and across any given medium. This makes it possible to aggregate any of the already standardized and implemented 10GBASE physical layer devices. APL enables aggregation of PHYs in groups of four or 10 to support 40 or 100 Gb/s operation on a wide variety of physical media, including multimode fiber, single-mode fiber, copper cable assemblies, and backplane assemblies. By employing APL, the industry will be able to bring 40 and 100 Gb/s products to market very rapidly, addressing customer needs while preserving the knowledge base that has been developed for 10 Gb/s Ethernet.

APL can also be used to aggregate higher-speed physical links, such as a group of four optical links running at 25 Gb/s, as another means of achieving 100 Gb/s operation. Taken a step farther, APL can aggregate 100 Gb/s links to provide bandwidths of 400 Gb/s or 1 Tb/s. Thus, APL is a forward-looking protocol that will serve the industry and the vast community of users for many years to come.

BIOGRAPHY

HOWARD M. FRAZIER (hfrazier@broadcom.com) is a technical director at Broadcom Corporation, responsible for technical strategy development within the Enterprise Networking Group. Previously, he was a distinguished engineer at Cisco Systems. He is one of the co-inventors of 10 Gigabit Ethernet, Gigabit Ethernet, and Fast Ethernet. He also served as chairman of the IEEE 802.3 Task Forces that wrote the standards for Ethernet in the First Mile (IEEE Std 802.3ah), Gigabit Ethernet (IEEE Std 802.3z), and Fast Ethernet (IEEE Std 802.3u). He has served as chairman of the IEEE Standards Association Review Committee (RevCom) and vice chairman of the IEEE-SA Standards Board. Among his accomplishments in high-speed networking, he led the development of the world's first 10/100BASE-T network interface controller in 1993 while working at Sun Microsystems. He has also co-authored a book titled *Ethernet in the First Mile: Access for Everyone*, published by the IEEE Standards Information Network/IEEE Press. He is a graduate of Carnegie-Mellon University.