#### `IEEE 5G World Forum 2020 ver. D

## Header Compression for ***5G*** Transport, Without Internet Protocol, Saves Bandwidth and Latency

Header compression to save bandwidth has been used for decades, but only over a single link at a time. To preserve end-to-end sessions, compressed headers of arriving packets are decompressed before routing. Processing at each hop burdens the router and adds latency, so header compression has been applied largely to choke points such as access links and satellite hops.

A class of real-time connectivity, "conversational," places high value on low latency. Telephone calls, live video conferencing, telemedicine, and gaming fall in this class. Time limits bar retransmitting errored or missing packets because replacements arrive too late to be of use. Traffic trends are pushing new volumes of conversations onto packet transport as the legacy circuit switched public phone network winds down and earlier generations of cellular technology reach sunset.

For this class, a new form of header compression, Session Bridging, saves bandwidth and reduces end-to-end latency without imposing processing loads on intermediate nodes.

Originally there was a reason to carry all headers everywhere: early data transmissions faced noisy lines and frequent errors. VoIP and conversational sessions arrived more recently, when transmission quality had improved. For an easy transition from circuits, VoIP kept existing packet headers and added RTP to help emulate digital circuits. Today, live media sessions can't replace lost or damaged packets quickly enough to be of use, so nobody tries, eliminating the reason for most of the header functions for "conversational" streams.

So why do packets still carry all those other headers--IP, UDP/TCP, RTP, etc.--despite a high cost in bandwidth? Tradition plays a part.

* Some applications and LAN connectivity depend on multiple headers. Carrying all the headers preserves an end-to-end session at each protocol layer. For example, one TCP session or chain of sessions between end points transport information reliably by resending errored packets.
* MultiProtocol Label Switching (MPLS) developed as a tweak to IP routing to speed transport and provide some security features at a time when enterprises migrated off of leased lines. Maintaining all headers and simply adding a Label ahead of them was easy to implement.

Even before MPLS, packet payloads crossed carrier networks with no regard for two pairs of globally unique addresses in the IP and Ethernet headers. Bridged Ethernet and virtual circuits also treat IP addresses as overhead. A Label Switched Router (LSR) examines only the Label value to decide how to forward a packet on a Label Switched Path (LSP). RFC--5795 discusses why legacy RObust Header Compression doesn't work over MPLS.

But Session Bridges do. For conversational connections, Session Bridge instances at the end points achieve the savings of header compression across an entire network without updating existing switches and routers.

The innovative process separates a connection into three parts, two sessions and a bridge to join them (Fig. 1). Session Bridges communicate with user equipment via packets with full header sets. Between them the payloads bear a much smaller compressed header, the bridge tag.

Fig. 1: The Sessions Bridge identifies each stream as it enters the network and replaces the full headers with a tiny tag. As the packets exit the network at the far end the full headers are rebuilt in a separate session so downstream systems are unaware.

The Session Bridge ORIGin module terminates protocol sessions of conversational flows at the front of a network (LAN Session 1), extracts the payloads, and reframes them with new bridge tags to route and manage the packet stream:

* A standard MPLS Label
* A unique field to identify how the headers are to be recreated at the TERMination module when the packet exits

A separate session (LAN Session 2) extends from the far-end appliance to the user equipment. TERM places a complete set of new headers on each exiting payload. Appliances hide the bridging operation from the end-customer equipment.

The terms ORIG and TERM refer to the direction of the flow of packets. Each appliance acts in both capacities to maintain full-duplex connections. While on the central network segment, the payload carries only the tag.

Custom code in the working proof of concept shows Session Bridging has minimal impact on perceived quality of voice. It works well because ensured packet order and low error rate remove the need for packet sequence numbers. With no reordering process, buffers in TERM can shrink which reduces latency.

The TERM function inserts new sequence numbers in delivered packets--not necessarily identical to the source. When needed, the ORIG process can add a separate short sequence number to detect lost packets or to create skips in the output sequence.

As an alternative to LSPs, virtual Ethernet LAN connections (for example, vLINE and other Carrier Ethernet services) also support Session Bridging. In Ethernet cases the Label may be omitted and only a control word accompanies the payload. For most carriers the LSP promises more bandwidth savings than Carrier Ethernet.

### A Practical Service

Quickly available LSPs open the new avenue of Session Bridge bandwidth savings for real-time "conversational" connections. In the past, MPLS services took weeks or months to engineer and turn up. Rrecently, software defined network (SDN) orchestration applications create LSPs on demand. Better yet, the LSP can extend beyond a carrier's previous boundary, into customer premises equipment (CPE).

5G wireless service makes low latency a key issue, but also strains bandwidth capacity as faster radio access networks funnel more traffic into core networks. Shorter packets benefit both bandwidth and latency. That is, Session Bridging:

* Applies header compression savings over Radio Access Networks, aggregation segments, and core transport with processing required only at the edge.
* Eliminates compression processing in intermediate nodes. Routers forward payloads based on an MPLS Label or bridged Ethernet path.
* Reduces CPU loads needed for routing, which increases the packet-per-second capacity of existing routers and extends their useful life.
* Adds no latency for compression/decompression for any number of links.

Saving from Session Bridges is in addition to payload compression.

* Voice has reached a point of diminishing returns as the bandwidth per conversation was dropped 8:1 or more. Headers can be many times the size of the voice payload.
* Video codecs continue to improve. The search for lower video latency points to smaller video payloads which enhance the relative importance of header overhead and raise the potential savings from header compression.

### Use Cases

Breaking connections into two sessions joined by an efficient bridge has three initial uses cases.

#### Session Bridge at IP and Above

A proof of concept (custom software on commercial hardware platforms) improves efficiency for VoIP by replacing RTP/UDP/IP headers (about 100 bytes) with a tag header of 8 bytes. The demonstration operates over an Ethernet switch with the appropriate Labels or transport via existing MPLS networks.

A Session Bridge instance acts as a Label Edge Router: the bridge tag starts with an MPLS Label. For bridged Ethernet connections the tag may omit the Label, which makes the tag only 4 bytes.

Table 1: Demonstrated VoIP Bandwidth Savings of Session Bridge Appliances

| Voice Codec  | RTP*a* | Session Bridge **Instance*b*** | **Saved** |
| --- | --- | --- | --- |
| G.729 (8 kbit/s) IPv4 | 34,400 bit/s |  14,800 bit/s | **57.0%** |
| IPv6s | 42.400 bit/s | 14,800 bit/s | **65.1%** |
| G.711 (64 kbit/s) IPv4 |  90,400 bit/s |  70,800 bit/s | **21.7%** |
| IPv6 | 98,400 bit/s | 70,800 bit/s |  **28.0%** |

*a Most conservative assumptions: minimum size VoIP headers, no extensions, includes Ethernet header.
b Session Bridge appliance places two payload samples in each Session Bridge packet, halving packets per second.*

#### Layer 2 Links Between Routers

Today's routers and switches use Ethernet for transport because of the silicon built into the hardware. That is, an Ethernet header is required on any packet to cross a link. Recent announcements of programmable chips to interface with links means that the Ethernet requirement no longer applies. A Session Bridge process in the Network Interface Card (NIC) replaces the Ethernet header with a smaller Layer 2 header. At the link level, Transparent Link software reduces the header size for every packet, not limited to conversational connections. Separation between protocol layers allows both forms of Session Bridge compression to operate independently or together.

Transparent Link connections apply to point-to-point connections between routers or switches. Call capacity for VoIP with compressed payloads doubles again compared to an Ethernet link. Gains apply to all packets, for example within a data center, and not only to media streams.

Additional saving may accrue from HDLC because it does not require inter-packet intervals needed by Ethernet at some speeds.

#### User Equipment on Radio Access Network

For carriers that adapt their networks to extend the LSP into MPLS-capable user equipment, there is no need to rebuild headers at the delivery end. The savings in reduced bandwidth then apply to the backhaul, fronthaul, and the Radio Access Network as well as the core.

### Conclusion

Existing features of carrier networks allow more efficient transport of conversational or real-time media streams when headers are compressed in a way that applies across entire networks. Session Bridging can save bandwidth over constrained resources such as radio links and access to remote sites.