Transmission of IP Packets over Overlay Multilink Network (OMNI) Interfaces

IETF 111 intarea session (July 30, 2021)
Fred L. Templin (The Boeing Company)
fred.ltemplin@boeing.com
fltemplin@acm.org
Document Status

• “Transmission of IP Packets over Overlay Multilink Network (OMNI) Interfaces”
  • Single **Non-Broadcast, Multiple Access (NBMA)** network interface exposed to the IP layer with fixed 9KB MTU, but configured as an overlay over multiple underlying (physical or virtual) interfaces with heterogeneous MTUs
  • **OMNI Adaptation Layer (OAL)** – minimal mid-layer encapsulation that maps IP layer to underlying interfaces (similar to AAL5)
  • **Work closely related to ‘draft-templin-6man-aero’ (AERO)**
  • AERO/OMNI formerly discussed in 6man WG but not adopted there – now seeking adoption in rtgwg and/or intarea
  • Work also related to (but separate from) ‘draft-ietf-intarea-tunnels’
Clients connect to access Subnetworks via one or more data links

Proxy/Servers at Subnetwork edges act as IPv6 routers or IPv6 ND Proxies

Bridges establish “spanning tree” over one or more Internetwork w/IPv6 ULAs

OMNI Adaptation Layer (OAL) IPv6 encapsulation sees spanning tree as L2 service (same as bridged campus LAN)
AERO/OMNI Operational Big-Picture

• Arbitrarily-many **Internetwork segments** joined as **Non-Broadcast, Multiple Access (NBMA) link**
• **Segment Routing** over concatenated Internetworks using **OAL forwarding**
• Example Internetwork segments include IPv4 Internet, IPv6 Internet, corporate enterprise networks, cellular operator networks, etc.
• **NOTE:** Clients may connect to multiple distinct segments (not shown)
Overlay Multilink Network Interface (OMNI)

- AERO/OMNI Clients, Proxy/Servers and Bridges configure OMNI Interfaces (node’s connection to the OMNI link).
- Client OMNI Interfaces configured over multiple underlying interfaces; connect End-User Networks (EUN).
- OMNI interfaces use IPv6 encapsulation to span entire concatenated Internetworking path in a single NBMA L3 hop (all AERO/OMNI nodes are neighbors on the link).
OMNI Client Underlying Interfaces

- Clients connect to first-hop Proxy/Servers via underlying interfaces in one of the following ways:
  - **Direct-Connect (aka Point-to-Point)** – Client’s underlying interface connects directly to Proxy/Server at L2 over P2P media without traversing any L3 hops
  - **VPN** – Client’s underlying interface uses network-layer security to securely establish a virtual link to Proxy/Server over one or more L3 hops
  - **Access Network (ANET)** – Client’s underlying interface connects to a secured access network (e.g., enterprise/operator network) with L2/L1 security provisions on the path to the Proxy/Server, which then connects to an external Internetwork
  - **Internetwork (INET)** – Client’s underlying interface connects directly to an external Internetwork, where the Proxy/Server may be one or more L3 hops away. Client may be located behind one or several Network Address Translators (NATs)
  - **Clients may have diverse underlying interface types, which may connect to multiple distinct Internetworks (as opposed to all via the same Internetwork). For example, an INET interface connected to the public Internet, an ANET interface connected to a cellular operator network, a Direct interface connected to a dedicated enterprise network Proxy/Server, etc.**
OMNI Encapsulation and Segment Routing

- OMNI Interface inserts mid-layer encapsulations between IP layer (L3) and Underlying Interfaces (L2) – **minimal encapsulation with effective header compression**
- Mid-layer **segment routing** used as necessary to traverse spanning tree over disjoint Internetworking “segments”
- **Route optimization** avoids strict spanning tree paths whenever possible
AERO/OMNI Addressing

• Each Client advertises **IP Global Unicast Address (GUA)** prefix (aka **Mobile Network Prefix (MNP))** to downstream-attached End-User Networks (EUNs) (e.g., 2001:db8:1:2::/64)
  • EUN applications use MNP addresses as the source/destination addresses of original IP packets

• Each AERO/OMNI node configures an **IPv6 Link-Local Address (LLA)** for use in IPv6 Neighbor Discovery (IPv6 ND) messaging
  • Clients configure **MNP-LLAs** (e.g., for the delegated MNP 2001:db8:1:2::/64, the MNP-LLA is fe80::2001:db8:1:2/128)
  • Proxy/Servers and Bridges configure Administrative LLAs (**ADM-LLAs**) with the 32 least significant bits set according to SRT node numbering/subnetting (e.g., for the SRT subnet 2001:1000/24, ADM-LLAs within that subnet are assigned as fe80::2001:1001/112, fe80::2001:1002/112, fe80::2001:1003/112, etc.)

• Each AERO/OMNI node configures an **IPv6 Unique-Local Address (ULA)**
  • Clients, Proxy/Servers and Bridges configure ULAs by concatenating the lower 64 bits of their LLAs with the OMNI link 64 bit ULA prefix (e.g., for the ULA prefix fd00:0102:0304:0506::/64 the MNP-ULA taken from the above MNP-LLA is fd00:0102:0304:0506:2001:db8:1:2/128)
  • Clients, Proxy/Servers and Bridges use ULAs as OAL header source/destination addresses
AERO/OMNI Addressing (continued)

• AERO/OMNI Proxy/Servers configure IPv4 and/or IPv6 Anycast addresses to advertise OMNI link reachability into access networks

• Potential Clients that connect to the access network can send network join IPv6 ND messages with an Anycast encapsulation destination address to discover the nearest Proxy/Server for the desired OMNI link

• For IPv4, the Anycast address is 192.88.99.1 (formerly reserved as 6to4 relay anycast, but reclaimed for AERO/OMNI)

• For IPv6, the Anycast address is based on the “6to4-expansion” of 192.88.99.1 as 2002:c058:6301:MNP[64]:Link-Suffix[16]
AERO/OMNI Routing System

• Proxy/Servers are “stub” Autonomous System Border Routers (s-ASBRs)
• Peer with Bridges as “core” (c-ASBRs)
• Bridges use BGP to peer with the Bridges of different segments to create ULA-based spanning tree
• Secure spanning tree – L2 security with IPsec, Wireguard, etc.
• Unsecured spanning tree – no L2 security provisions
• Spanning tree hops visit all nodes, but often result in suboptimal paths
OMNI Interface and OMNI Adaptation Layer (OAL)

- **OAL source** OMNI interface admits original IP packets up to **9180 bytes (OMNI MTU)**
- OAL source encapsulates in IPv6 header with **ULA addresses**, includes trailing checksum, then fragments if necessary and includes *NET header(s) to create “Carrier Packets” guaranteed small enough to traverse all paths
- Carrier packets traverse one or more Internetworks to **OAL destination** which securely reassembles, verifies checksum, then forwards the original IP packet to final destination
- **Original source** can tune packet sizes to achieve optimal performance based on “hard/soft” MTU feedback

---

Image showing the flow of data from Original Source to Final Destination through OMNI Interface and OMNI Adaptation Layer (OAL). The Original IP Packet is processed by the OAL source, creating Carrier Packets that traverse the Internetworks before being reassembled and forwarded to the Final Destination.
Automatic Extended Route Optimization (AERO)

- Original source and final destination on different Internetwork segments
- OAL source produces carrier packets, and OAL destination reassembles
- OAL intermediate nodes (aka Bridges) join network segments in spanning tree
- Supports global mobility and route optimization, with dynamic MTU tuning
IPv6 ND Messaging (all ND message types)

• AERO/OMNI nodes use IPv6 ND messages for OMNI link neighbor coordination

• IPv6 ND messages include “OMNI option” to:
  1. Assert protocol conformance
  2. Synchronize windows
  3. Exchange configuration information
OMNI Option Sub-Options

- OMNI options carry “sub-options” that determine protocol operations
- Multilink Forwarding Parameters, Interface Attributes, Traffic Selectors determine addresses and other parameters for specific underlying interfaces
- DHCPv6, HIP, PIM-SM messages combine those functions with the IPv6 ND message as a “Wrapper”
- Reassembly Limit, Fragmentation Report, ICMPv6 Error, etc. provide OMNI destination feedback to source
IPv6 ND Messaging (RS/RA)

• Clients and Proxy/Servers exchange Router Solicitation/Advertisement (RS/RA) messages either directly or via a Proxy middlebox:
  • When a Client sends an RS message to Proxy/Server “A” with “All Routers” multicast or the ADM-LLA “A” as the destination, “A” assumes the ROUTER role and returns a unicast RA while injecting the Client’s MNP into the AERO/OMNI routing system
  • When a Client sends an RS message to Proxy/Server “B” with the ADM-LLA “A” as the destination, “B” assumes the PROXY role and forwards the RS to “A” which returns a unicast RA to the Client via “B”
  • Proxy/Server “A” then becomes the “hub” (aka default router / mobility anchor point) in a hub-and-spokes arrangement with all other Proxy/Servers “B*” as the “spokes”. Client can later switch to a different hub by sending new RS messages
Hub-and-Spokes Proxy/Servers

- First Proxy/Server contacted becomes the “Hub” and acts as Default Router
- Other Proxy/Servers become “Spokes” and act as IPv6 ND Proxys
- All Proxy/Servers equally capable of serving in the Hub role
- If the current Hub goes down, Client sends new RSes to select a new Hub.
MTU Hard and Soft Errors

- RFC1191 and RFC8201 provide ICMP Packet Too Big (PTB) error messages that ALWAYS report packet loss due to size restriction (i.e., a “hard error”)
- OMNI asks for a new PTB code to indicate a “soft error”, in which the packet is still delivered, but the error informs the source that it should reduce the size of future packets
- With PTB “soft errors”, sources can dynamically tune the sizes of packets they send to get the best performance w/o loss-related retransmissions
NAT Traversal Issues

• When source/target Client underlying interface connects via an open Internetwork, there is a very good chance it may be located behind one or more IP Network Address Translators (NATs)

• Each Client naturally establishes NAT mappings when it performs an RS/RA exchange with its FHS Proxy/Server, but these mappings generally not viable for Client exchanges with other AERO/OMNI nodes

• When AERO/OMNI route optimization is applied, route optimization peer nodes must perform NS/NA(NUD) and “bubble” exchanges to establish NAT mappings for itself in the NAT(s) on the path to the Client

• NAT traversal procedures are the same as in RFC4380 and RFC6081
IPv6 ND Messaging (unsolicited NAs)

- Hub Proxy/Servers announce Client state changes (e.g., mobility-related address changes, link quality changes, traffic selector changes) by sending unsolicited NAs (uNAs) to all nodes that it has recently sent an NA(AR)
  - When nodes receive uNAs, they update their state information for the Client which may require new NS/NA(NUD) path state exchanges
- New Hub Proxy/Servers send uNAs to inform old P/S that Client has “Departed”
  - When the old Hub Proxy/Server receives uNAs, it records the new Hub Proxy/Server ADM-LLA, withdraws MNP route and forwards uNAs to nodes that it has recently sent an NA(AR)
- Target nodes request selective link-layer retransmissions by sending uNAs to source with the Identifications (and ordinal numbers) of any missing fragments
  - When the source node receives the uNAs, it retransmits requested fragments if it still has them in its retransmission cache
  - Retransmission window should be brief, and determines link persistence
Document Adoption Call

• “Transmission of IP Packets over Overlay Multilink Network (OMNI) Interfaces”
  - Adopt as IETF intarea working group item?