AERO/OMNI and IP Parcels

IETF114 Intarea Working Group Session (July 28, 2022)
Fred L. Templin (The Boeing Company)
fred.l.templin@boeing.com
The Internet was Intended as a Network of Networks

• Beginning in the 1970’s, the ARPANET grew to become what we know today as the global public Internet - one of the first nodes was mobile (the SRI van)!

• Today’s public Internet is a single monolithic routing and addressing domain instead of a network of networks – incomplete architectural layering!

• Private Intranets connect to the Internet via security devices (firewalls, proxys, NATs, etc.) but use address translation – no true end-to-end global addressing!

• Internetworking between private Intranets problematic due to addressing and security incompatibilities – complicates global mobile Internetworking!

• But, the early pioneers envisioned true end-to-end communications over a global-scale network of networks. They called it:

“**The Catenet Model for Internetworking**”

“We are at a unique point in history where end-to-end can be restored”
The Catenet Model for Internetworking

- Documented in Internet Engineering Note 48 (IEN-48) written by Vint Cerf in 1978
- Incorporated still earlier concepts from Louis Pouzin beginning in 1974
- Envisioned a true “network of networks”
- They knew that Gateways were required to interconnect diverse Internetworks, but did not know how to traverse them
- They knew that end systems also required a “Local Gateway” to support end-to-end
- They did not know a new architectural layer was needed (the “Adaptation Layer”)
  - AERO/OMNI: an Adaptation Layer for the Internet

Original Catenet Figure from IEN-48 (1978)
AERO/OMNI and the Adaptation Layer

Net-A

Net-B

Net-C

Net-D

• Independent Internetworks currently in disjoint partitions
AERO/OMNI and the Adaptation Layer

- Independent Internetworks currently in disjoint partitions
- AERO/OMNI use Gateways to connect partitions and establish an (overlay) Adaptation Layer
Independent Internetworks currently in disjoint partitions
AERO/OMNI use Gateways to connect partitions and establish an (overlay) Adaptation Layer
End-to-end supported even across partition boundaries
AERO/OMNI and the Adaptation Layer

- Independent Internetworks currently in disjoint partitions
- AERO/OMNI use Gateways to connect partitions and establish an (overlay) Adaptation Layer
- End-to-end supported even across partition boundaries
- Adaptation based on Encapsulation and Fragmentation
OMNI Adaptation Layer (OAL)

- End (or near-end) systems configure OMNI interface (the Catenet “Local Gateway”)  
- OMNI Adaptation Layer (OAL) source uses IPv6 encaps/frag to produce “OAL packets/fragments”, then uses L2 encapsulation to produce “carrier packets”  
- Carrier packets traverse network to OAL destination which reassembles/decapsulates  
- Source can tune its packet sizes without loss to achieve best performance  
- Similar to ATM Adaptation Layer – 5 (AAL5)
AERO/OMNI Multinet Traversal

- Original source and final destination on different Internetwork segments
- OAL Source produces OAL packets/fragments, and OAL Destination reassembles
- OAL Gateways forward OAL packets/fragments below IP but above link layer
- Carrier packets transport OAL packets/fragments across first-hop segment, then undergo re-encapsulation and re-transmission at each next-hop segment
- True end-to-end in the spirit of Catenet
AERO/OMNI and the “6 M’s of Modern Mobile Internetworking”

• Adaptation Layer naturally eliminates many challenges that complicate diverse mobile Internetworking service models

• Incremental deployment on existing networks – no need for a “flag day”

• Security addressed at all layers of the architecture, including end-to-end

• Delay Tolerant Networking (DTN) naturally accommodated

• AERO/OMNI uniquely address the “6 M’s of Modern Mobile Internetworking”:

1. **Multilink** – the ability for a mobile node to utilize multiple diverse communications links simultaneously instead of just one at a time (improved performance and reliability)

2. **Multinet** – the Catenet network of networks model (coordinated in conjunction with the other M’s)

3. **Mobility** – the ability for a mobile node to move dynamically between communications link attachment points while maintaining uninterrupted end-to-end communications without readdressing

4. **Multicast** – the ability for a source to send a single packet stream that is received by multiple mobile node group members

5. **Multihop** – node-to-node relaying between mobiles out of range of fixed infrastructure

6. **MTU Assurance** – the ability for mobile nodes to send packets of diverse sizes without loss and to dynamically tune packet sizes for best performance – inspired new construct known as the “IP Parcel”
IP Parcels

• IP packets (both IPv4 and IPv6) contain data unit that becomes **retransmission unit** in case of loss

• Upper Layer Protocols (ULPs), e.g., TCP, QUIC/UDP, LTP/UDP etc., exchange **segments** with a **single segment** per IP packet

• **IP Parcels** permit **single packet** to carry **multiple ULP segments** ("packet-of-packets"), but segment still loss/retransmission unit

• Goal:
  • Support larger packets for better performance
  • Support flexible packaging/re-packaging for more efficient handling
  • Encourage larger and more diverse Maximum Transmission Units (MTUs)
IP Parcel Analogy

• “When a consumer orders 50 small items from a major online retailer, the retailer does not ship the order in 50 separate small boxes. Instead, the retailer puts as many of the small items as possible into one or a few larger boxes (or parcels) then places the parcels on a semi-truck or airplane. The parcels arrive at a regional distribution center where they may be further redistributed into different-sized parcels that are finally delivered to the consumer. But most often, the consumer will only find one or a few parcels at their doorstep and not 50 individual boxes. This greatly reduces handling overhead for both the retailer and consumer.”
IP Parcel Formation

• ULP identified by **5-tuple** (src-addr, dst-addr, src-port, dst-port, proto) produces buffer with **up to 64** segments

• All segments except final must be equal-length **up to 65535 octets (minus headers)**; final segment may be smaller

• ULP delivers buffer and non-final segment size to IP layer

• IP layer forms Parcel by appending **Jumbo Payload option**

```
+---------------+---------------+
| Opt Type | Opt Len | Jumbo Payload Length |
+---------------+---------------+
```

Jumbo Payload option
IP Parcels Based on IP Jumbograms

• IP Parcels use **Jumbo Payload option** with non-zero \(\{\text{Total, Payload}\}\) Length (true Jumbos use zero)
  
  • \(\{\text{Total, Payload}\}\) Length encodes length of first segment only
  
  • **Jumbo Payload Length** encodes length of entire Parcel

• IP Parcels defined for both IPv6 and IPv4
  
  • “IPv4 Jumbo Payload” reuses obsolete RFC1063 “IPv4 Probe MTU” option

• Maximum IP Parcel Size: \(~(64 \times 65535) = \sim 4\text{MB}\)
IP Parcel Structure

- Supports TCP; transports over UDP
- Includes up to 64 ULP segments, but only one {TCP,UDP}/IP header
- {TCP,UDP} checksum covers headers only with individual checksum trailer for each segment (all checksums calculated in single pass over data)
- For TCP only, each non-first segment is preceded by a 4-octet Sequence Number header (UDP transports encode their own start delimiter in each segment)
Transmission of IP Parcels

• IP Parcels traverse **Parcel-capable** links with sufficient MTU (same as packets)

• **Parcel-capable (physical) links** not yet available, but **OMNI (virtual) links** can forward IP Parcels using **Adaptation Layer**

• **OMNI Adaptation Layer (OAL)** uses **encapsulation/fragmentation** to break large Parcels into smaller (sub-)Parcels if necessary since largest that can undergo IP fragmentation is 65535 octets
  • 1\textsuperscript{st} pass: Parcel fragmentation (“loose” reassembly w/ opportunistic merging)
  • 2\textsuperscript{nd} pass: IP fragmentation (“strict” reassembly w/ fragment retransmission)

• Goal:
  • forward fewest and largest IP Parcels possible over network to final destination
  • minimize segment reordering due to re-Parceling if possible (not critical)
  • leverage IP fragmentation/reassembly if necessary
  • **loss unit single segment instead of entire Parcel**
Parcel Path Qualification

• Goal: qualify some or all of forward path as Parcel-capable (incremental deployment)

• **Parcel Probe** from source tests consecutive hops up to destination; router with non-Parcel-capable next hop
  • **Hop-By-Hop Option** (processed at each hop)

• **Parcel Reply** from destination/router informs source that some or all of forward path is Parcel-capable
  • **UDP/IP encapsulated ICMPv6** (processed only at source)

• After Parcel Path Qualification:
  • Parcels from source traverse Parcel-capable path same as ordinary IP packets up to destination/router
  • Destinations that receive Parcels can efficiently deliver them to upper layers
  • Routers that terminate Parcel-capable paths open Parcels and forward individual IP packets to destination
IP Parcel Integrity

• Link-layer integrity checks (e.g., CRC-32) can miss errors in packets larger than ~9KB – but, IP Parcels often much larger
• IP Parcels include separate integrity checks for each ULP segment
• Parcels improve integrity compared to same-sized packets/Jumbograms which only include single ULP segment and integrity check
• Only segments with correct integrity are accepted – individual segment (and not entire parcel) is the loss/retransmission unit

>IP Parcels encourage new link types with larger and more diverse MTUs plus improved integrity, resulting in major Internetworking performance improvements
Adoption Call

• IP Parcels ready for adoption as intarea wg document
• AERO/OMNI also ready for adoption as intarea wg documents

ADOPTION CALL:
• Adopt IP Parcels?
• Adopt AERO/OMNI?
• Automatic Extended Route Optimization (AERO)  
  • https://datatracker.ietf.org/doc/draft-templin-6man-aero/

• Transmission of IP Packets over Overlay Multilink Network (OMNI) Interfaces  
  • https://datatracker.ietf.org/doc/draft-templin-6man-aero/

• A Simple BGP-based Mobile Routing System for the Aeronautical Telecommunications Network  
  • https://datatracker.ietf.org/doc/draft-ietf-rtgwg-atn-bgp/

• IPv6 Fragment Retransmission and Path MTU Discovery Soft Errors  
  • https://datatracker.ietf.org/doc/draft-templin-6man-fragrep/

• IP Parcels  
  • https://datatracker.ietf.org/doc/draft-templin-intarea-parcels/
Additional Information – APNIC Blog

• APNIC Blog Fred Templin Publication Series
  • https://blog.apnic.net/author/fred-templin/

• OMNI: An Adaptation Layer for the Internet
  • https://blog.apnic.net/2022/02/18/omni-an-adaptation-layer-for-the-internet/

• OMNI: Integrity, Efficiency and Security
  • https://blog.apnic.net/2022/04/13/omni-integrity-efficiency-and-security/

• OMNI and the 6 M’s of Modern Internetworking
  • https://blog.apnic.net/2022/05/18/omni-and-the-6ms-of-modern-internetworking/

• AERO, OMNI and DTN: An internetworking architecture for mobility
  • https://blog.apnic.net/2022/06/22/aero-omni-and-dtn-an-internetworking-architecture-for-mobility/

• AERO/OMNI/DTN routing and route optimization
  • https://blog.apnic.net/2022/07/12/aero-omni-dtn-routing-and-route-optimization/
Backups
AERO/OMNI and the Adaptation Layer

- Independent Internetworks currently in disjoint partitions
- AERO/OMNI use Gateways to connect partitions and establish an (overlay) Adaptation Layer
- End-to-end supported even across partition boundaries
- Adaptation based on Encapsulation and Fragmentation
• Generic Segment/Receive Offload (GSO/GRO) implemented in some OS’s and NICs; ULP can supply multiple segments in single system call
• QUIC study showed significant performance increases using GSO/GRO
• Licklider Transmission Protocol (LTP) study showed moderate increases for small-to-medium segments using GSO/GRO, but significant increases for larger single segments even if IP fragmentation/reassembly needed
• BIG-TCP study considered end system-internal implications of Jumbograms for better performance
• IP Parcels combine GSO/GRO segmentation and IP fragmentation with IP Jumbograms for network transmissions