IP Parcels

draft-templin-intarea-parcels-09
IETF113 intarea WG – March 22, 2022
Fred L. Templin (fltemplin@acm.org)
The Boeing Company
Draft Status

  • draft-templin-intarea-parcels-00 published 12/17/2021
  • Closely related to AERO/OMNI/DTN:
    • [https://datatracker.ietf.org/doc/rfc9171/](https://datatracker.ietf.org/doc/rfc9171/)

• draft-templin-intarea-parcels-09 published 2/10/2022
• This presentation is a narrative of -09
IP Parcels

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

• Upper Layer Protocols (ULPs), e.g., QUIC, LTP, TCP, etc., exchange **segments** and include 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 Maximum Transmission Units (MTUs) in the Internet
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 boxes 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 slightly smaller parcels that get delivered to the consumer. But most often, the consumer will only find one or a few parcels at his 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-IP, src-port, dst-IP, dst-port, protocol) 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

<table>
<thead>
<tr>
<th>Opt Type</th>
<th>Opt Len</th>
<th>Jumbo Payload Length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Jumbo Payload option
IP Parcels Based on IP Jumbograms

- IP Parcels include Jumbo Payload option but with **non-zero Payload Length** (true Jumbos use zero)
- Payload Length gives length of first segment only; Jumbo Payload Length gives length of entire Parcel
- IP Parcel support for both IPv6 and IPv4 (Jumbo Payload option for IPv4 defined)
- Maximum IP Parcel Size: \(~(64 \times 65535) = \approx 4\text{MB}\)
Related Work

- 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
### IP Parcel Types

<table>
<thead>
<tr>
<th></th>
<th>Segment-1 len = L</th>
<th>Segment-2 len = L</th>
<th>Segment-3 len = L</th>
<th>Segment-J len = K</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jumbo Payload</td>
<td>Jumbo Payload</td>
<td>Jumbo Payload</td>
<td>Jumbo Payload</td>
<td></td>
</tr>
<tr>
<td>len = IHL</td>
<td>len = IHL + L</td>
<td>len = IHL + L</td>
<td>len = IHL + (J-1)L + K</td>
<td></td>
</tr>
<tr>
<td>IP Header</td>
<td>IP Header</td>
<td>IP Header</td>
<td>IP Header</td>
<td></td>
</tr>
<tr>
<td>len = IHL</td>
<td>len = IHL + L</td>
<td>len = IHL + L</td>
<td>len = IHL + L</td>
<td></td>
</tr>
</tbody>
</table>

- **Null Parcel** - IP header with Jumbo Payload option; “No Next Header” (singleton w/segment length L=0)
- **Singleton Parcel** - IP header with Jumbo Payload option; single ULP segment of length L
- **Multi-Segment Parcel** - IP header with Jumbo Payload option; (J-1) ULP segments of length L; final ULP segment of length (K <= L)
Transmission of IP Parcels

• IP Parcels traverse **Parcel-capable** links with sufficient MTU (same as packets)
• Parcel-capable links not yet widely deployed in heterogeneous Internetworks, but **Adaptation Layer** can forward Parcels in **overlay**
• **OMNI Adaptation Layer (OAL)** uses **encapsulation** and **fragmentation**
• OAL breaks large Parcels into smaller (sub-)Parcels if necessary since largest that can undergo IP fragmentation is 65535 octets
  • 1\(^{\text{st}}\) pass: Parcel fragmentation (“loose” reassembly w/ opportunistic merging)
  • 2\(^{\text{nd}}\) 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 for greater performance
  • loss unit single segment instead of entire IP Parcel
Parcel Path Qualification

- **Goal**: qualify head of (src->dst) forward path as Parcel-capable (support incremental deployment)

- **Parcel Probe** from src tests consecutive hops up to dst or router with non-Parcel-capable next hop
  - **Hop-By-Hop Option** (processed at each hop)

- **Parcel Reply** from dst/router informs src head of forward path Parcel-capable
  - **Destination Option** (only processed at src)

- **After Parcel Path Qualification**:
  - Parcels from src traverse Parcel-capable path same as ordinary IP packets up to the end dst/router
  - Routers that terminate Parcel-capable paths open Parcels and forward individual IP packets to dst

```
+----------------+-----------------+
| Type           | Length          |
+----------------+-----------------+
| Nonce-1        | Nonce-2 (0-1)   |
+----------------+-----------------+
| Nonce-2 (2-3)  | Reserved| Check |
+----------------+-----------------+
|               | PMTU             |
+----------------+-----------------+
```

Parcel Probe/Reply option
IP Parcel Integrity

• Link-layer checks (e.g., CRC-32) can miss errors in packets larger than ~9KB – but, IP Parcels often much larger

• IP Parcels include separate integrity check for each ULP segment

• Parcels improve integrity compared to same-sized Jumbograms (Jumbos only include single ULP segment and integrity check)
Next Steps

• IP Parcels increase efficiency and performance for end systems
• IP Parcels provide path forward for larger MTUs in the Internet
• Adopt IP Parcels as WG Item?
Next Steps (2)

• AERO/OMNI (discussed here many times) provide Parcel-capable Adaptation Layer service

• AERO/OMNI specs also in advanced stages:
  https://datatracker.ietf.org/doc/draft-templin-6man-aero/
  https://datatracker.ietf.org/doc/draft-templin-6man-omni/

• Also adopt AERO/OMNI as WG items?