High-Level
End-to-End Network

- Access
- Metro
- Core
- Internet

- National POPs
- Satellite POPs
- P&T (IXes etc.)

- Edge POPs
- Core POPs
Access

Fibre to the Home with
Gigabit Passive Optical Network (GPON)

- Optical Line Terminal (OLT) @ Edge POP
- Optical Network Terminal (ONT) @ Customer Premise
  - 2 box solution with Sky Hub 4 CPE router
Access

GPON

- Single-strand single-mode fibre
  - BiDi Optics
- Point-to-Multipoint using passive splitters
  - Commonly @ 32:1 or 64:1 but up to 128:1
- 2.5Gbps down
- 1.25Gbps up
Access

GPON - Downstream

- ~1490nm Downstream
- Broadcast to all ONTs
- AES Encrypted
Access

GPON - Upstream

- ~1310nm Upstream
- OLTs “range” to find each ONT’s distance from OLT
- Time Division Multiplexed (TDM)
- Cells then converge on the wire without collision
Access
Unbundled GPON

- Distinct Optical Distribution Network (ODN)
- Shared ducts / fibre bundles
Access
Unbundled GPON

- A Customer can “port” ISPs, by an engineer physically patching the ONT over to the other provider’s splitter.
Access
Future XGS-PON

- Shared optical network using distinct wavelengths
- XGS-PON
  - ~1270nm Up
  - ~1577nm Down
- Wavelength Mux at the OLT head end
- XGS-PON coloured ONT
Metro Network

- Access
- Edge POPs
- National POPs
- Metro
- Satellite POPs
- Core POPs
- P&T POPs (IXes etc.)
- Internet

skytel
Metro Network

Metro Aggregation

Aggregates multiple OLTs from multiple Edge POPs
Metro Network
Transport Aggregation

Aggregates multiple MA Pairs from multiple National POPs
Metro Network

EVPN E-LAN

• Active/Active Multi-homed
• No need for MC-LAG or STP
• BGP-Signaled
• Efficient MAC-learning
• Efficient handling of BUM packets
Metro Network
Future EVPN-VPWS

- Point-to-Point
  - No MAC-learning
- Single Active on BNG side
  - Using “Backup” control flag (RFC8214)
Core Network

- Access
- Metro
- Core
- Internet

- Edge POPs
- National POPs
- Satellite POPs
- Core POPs
- P&T POPs (IXes etc.)
Core Network

- Merchant Silicon
  - Cheaper per Gbps cost
  - High Port Density
  - Rigid Feature Set / Packet Processing
  - Bugs in Vendor Implementations
- Segment Routing (MPLS)
  - No LDP or RSVP-TE
- Virtualised Route Reflectors
- ECMP Everywhere
Segment Routing

- Extensions to IGP for label distribution
  - Not a distinct new “protocol”
- Uses existing MPLS label switching
- No Stateful Label Switch Paths (LSPs)
  - State is kept in the packet header
  - Scalable
- Supports Traffic Engineering
  - Label Stack or “Segment List”
  - Offline Path Computation
  - Dynamic TE using Link Delay metrics signaled in ISIS (RFC8570)
Segment Routing
Shortest Path

Ingress PE → P → P → PE (X) → PE (Y)

Payload → Label X → Payload → Label X → Payload → Explicit Null → Payload
Segment Routing
Traffic Engineering
Subscriber Termination

- Access
- Metro
- Core
- Satellite POPs
- National POPs
- P&T POPs (IXes etc.)
- Edge POPs
- Internet
Subscriber Termination
Broadband Network Gateways (BNG)

- **IPoE** – Less encapsulation overhead compared to PPPoE
- **Native IPv6** - /48 PDs
- **Port-based Authentication** – DHCPv6 Option 37 Remote-ID
  - Inserted by the OLT’s Lightweight DHCPv6 Relay Agent (LDRA)
- **Redundant BNG** – Proprietary vendor magic session state syncing, plus VRRP
- **IPv4aaS** – Mapping of Address and Port (MAP)

Future:
- **Subscriber Termination directly on EVPN-VPWS**
  - Vendor magic ties BNG backup state with EVPN state
IPv4 Addressing

• Starting with zero
• Join RIPE as a Local Internet Registry (LIR)
  – €2000 joining fee + €1400pa
  – Gets you a /22 or 1024 IPv4 addresses
  – Maybe this will be enough for infrastructure?
• Buy off the open market @ >$15USD / IP
• So you will probably also want to do some form of IPv4 address sharing
  – We’ve chosen to use MAP-T

---

**MOST RECENT TRANSFERS COMPLETED**

Individual buyers and sellers will agree to a specific nominated currency such as USD, EUR, GBP, etc. in the following table, prices per IP are illustrated in USD for comparative regional purposes.

<table>
<thead>
<tr>
<th>Block Size*</th>
<th>/24</th>
<th>/23</th>
<th>/22</th>
<th>/21</th>
<th>/20</th>
<th>/19</th>
<th>/18</th>
<th>/17</th>
<th>/16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price/IP (USD)</td>
<td>26.00</td>
<td>23.00</td>
<td>20.00</td>
<td>20.00</td>
<td>19.50</td>
<td>19.50</td>
<td>19.00</td>
<td>19.50</td>
<td>19.00+ depending on quality</td>
</tr>
</tbody>
</table>

Mapping of Address and Port (MAP)

- IPv4aaS.
  - IPv4 over the top of IPv6 transport.
  - IPv6-only Access Layer.
  - Reduces operational overhead.
- Allows IPv4 address sharing, or 1:1.
- No DNS synthesizing required.
- Doesn’t require an agent on end-hosts.
- Can operate in either encapsulation or translation modes.
- **Stateless.**
Mapping of Address and Port (MAP)

RFC7597: MAP-E
- Encapsulation
  - Larger per-packet overhead.
  - IPv4 header remains intact.

RFC7599: MAP-T
- Translation
  - Less per-packet overhead (not zero!)
  - Loses IPv4-only header attributes.
  - 5-tuple hashing. (E.g., with ECMP or over LAGs)
  - Border relay-bypass.
Packet Flow (MAP-T)

IPv4 LAN → NAT44 → IPv6 Access → IPv6 Core → MAP BR → IPv6 Core → IPv6 Access → NAT64 → IPv4 Internet

IPv6 LAN → NAT46 → Native IPv6 → NAT44 → IPv4 Internet

Native IPv6 → IPv6 Core → IPv6 LAN

RFC 6052 IPv4-embedded IPv6 addresses used for external host destination address.
Packet Flow Example (MAP-T)

1. Stateful NAPT on CPE translates both source IPv4 address and source port (when oversubscribing) based on the BMR

<table>
<thead>
<tr>
<th>TCP/UDP Header</th>
<th>IP Header</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPORT: 6789</td>
<td>SRC: 192.168.0.129</td>
</tr>
<tr>
<td>DPORT: 80</td>
<td>DEST: 8.8.8.8</td>
</tr>
</tbody>
</table>

2. MAP-T agent on CPE translates source IP address to v6 based on the BMR, and destination IP address to v6 based on the DMR (and RFC6052)

<table>
<thead>
<tr>
<th>TCP/UDP Header</th>
<th>IP Header</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPORT: 1230</td>
<td>SRC: 181.230.60.51</td>
</tr>
<tr>
<td>DPORT: 80</td>
<td>DEST: 8.8.8.8</td>
</tr>
</tbody>
</table>

3. MAP-T Border Router translates source IP address back to IPv4 based on the BMR, and destination IP based on DMR (and RFC6052)

<table>
<thead>
<tr>
<th>TCP/UDP Header</th>
<th>IP Header</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPORT: 1230</td>
<td>SRC: 181.230.60.51</td>
</tr>
<tr>
<td>DPORT: 80</td>
<td>DEST: 2001:db8:0:0:0:0:0:0/64</td>
</tr>
</tbody>
</table>

Note: Source address translation is BMR-dependent, it would not use RFC6052 as shown in this example.
Mapping of Address and Port (MAP)

Stateless

• No need to keep track of every flow
• Efficient packet processing
• Cheaper, more scalable hardware
• Already supported on existing linecards from some vendors.

• Some jurisdictions may require 5-tuple logging for compliance reasons, which stateless IPv4aaS methods don’t provide.
Appendix A – Additional MAP-T Awesomeness
MAP Border Relay Anycasting

- DMR IPv6 prefixes can be anycasted internally.
- Public IPv4 prefixes can be anycasted externally.
- Stateless translation/encapsulation allows for asymmetric packet flows.
MAP Border Relay Bypass for CDN

- On-net content servers can be numbered from within the IPv6 DMR prefix, allowing for Border Relay-bypass, using more specific destination-based routing.
- Allows for serving of IPv4-only clients from IPv6-capable CDNs.
MAP Forward Mapping Rules

- Allows direct CPE <-> CPE communication, bypassing Border Relays.
Appendix B – All the Load Balancing
Load-balancing with EVPN + ECMP

Ingress LAG Hashing

Step 1: Ingress LAG Hashing
Hashed across all LAG members
Load-balancing with EVPN + ECMP

EVPN “Aliasing”

Step 2: EVPN Active/Active Load Balancing
Both ingress PEs see 2 egress PEs advertising ESI A
Load-balancing with EVPN + ECMP

EVPN “Aliasing”

Step 2: EVPN Active/Active Load Balancing
Both ingress PEs see 2 egress PEs advertising ESI A
Load-balancing with EVPN + ECMP
ECMP of Egress PE Loopbacks in ISIS

Step 3: ECMP Next Hops
LSRs see multiple paths for both egress PE next hops.
Load-balancing with EVPN + ECMP

ECMP of Egress PE Loopbacks in ISIS

Step 3: ECMP Next Hops

LSRs see multiple paths for both egress PE next hops.
Load-balancing with EVPN + ECMP

Egress LAG Hashing

Step 4: Egress LAG Hashing
End result is the same in this instance. But interesting paths through the core