Link Discovery and Liveness

What do we really need?

Randy Bush <randy@psg.com>
We Are Here
Trying to Discover
IIJ is Building a Second Medium Scale Data Center (MSDC) in Shiroi/Chiba Capacity of 6k Racks
How Can We Route In Something of This Scale?
OSPF OK to 500 Nodes
IS-IS good to 1,000

Limited Because They Repeatedly Flood Everything
BGP Is Great as Updates are Infrequent
BGP Scales Because It Signals Only Changes

So BGP has become common in MSDCs
ECMP can be Very Wide
32, 64, even 128
The Problem is Topology Discovery
Two Kinds of Standards

**Union** – the accumulation of all the features anybody wanted

**Intersection** – only those things everybody absolutely had to have

Either Tony Hoare or Klaus Wirth – I can not find the quote <blush>
IETF asks the ITU

Q: So you add features until the “NO”s stop

A: We don’t like to think of it that way
Must Haves

• Discover Nodes and Links

• Discover Link Encapsulations:
  - IPv4, IPv6, MPLS4/6, ...

• Maintain Layer-2 Liveness

• Northbound API to BGP-SPF
Security?

- Datacenter Ops seem not to think of security at this layer (or any!)
- We need Authentication. Maybe Integrity?
- One of the things which are likely to drive PDU size over 1,500
Non-Features

• Routing Data, BGP-SPF does that

• Access to IGP Databases, This is discovery and liveness, not routing

• Just want the Link

• Transport, not our job
Desiderata

- Discovery & Liveness for BGP-SPF
- Simple but usable in Massively Scalable networks of >10,000 nodes
- May be useful for other applications
- Simple
- Extensible (e.g. authentication, cost)
- Simple
- No IPR
Why Simple?

We are here to produce easily understood, implementable, and securable standards, not build résumés.
Why Simple?

A high goal of software engineering is to remove the need for features. It's a vital part of designing for simplicity, even invisibility. -- Rob Pike
Candidates?

- LLDP and its children
- IS-IS link discovery
- Edge Control Protocol (Alvaro)
- BGP Neighbor Autodiscovery
- Link State Over Ether
LLDP

• IEEE Protocol
• IPR over 1,500 bytes
• A bit complex
• Won’t go through a switch (feature or bug?)
• Beacons, not KeepAlives
• Viable but
IS-IS Discovery

• IETF now has control
• Complex enough that BGP-LS was invented so normals could get the link state database
• IS-IS not commonly implemented on MSDC devices, so would need to profile and develop
Edge Control Protocol

- It is a transport controlled by IEEE
- A Reliable layer two transport, on top of LLC
- Has flow control, reliable, non-reorder, ... transport
- used for EVP and PD/CSP
- Reinventing TCP over 802.1
BGP Neighbor Autodiscovery

- IETF protocol
- Very new
- Needs the peering address to get the peering address
- AS Based, can not use other idents
- Not really discovery at all, configuration
- No liveness
Link State Over Ether

• Custom made for the job
• Very bare bones, brutally simple
• Only does discovery and liveness
• New, therefore risky
• But so is BGP-SPF
• No measurement or monitoring tools
<table>
<thead>
<tr>
<th></th>
<th>LLDP</th>
<th>IS-IS</th>
<th>ECP</th>
<th>BNA</th>
<th>LSOE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Who Owns</td>
<td>IEEE</td>
<td>IETF</td>
<td>IEEE</td>
<td>IETF</td>
<td>IETF</td>
</tr>
<tr>
<td>Maturity</td>
<td>Mature</td>
<td>Mature</td>
<td>Recent</td>
<td>New</td>
<td>New</td>
</tr>
<tr>
<td>Complexity</td>
<td>Somewhat</td>
<td>Very</td>
<td>Rather</td>
<td>Somewhat</td>
<td>Almost too Simple</td>
</tr>
<tr>
<td>Discovery</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Configure</td>
<td>Yes</td>
</tr>
<tr>
<td>Liveness</td>
<td>Beacons</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>IPR</td>
<td>IPR</td>
<td>No</td>
<td>?</td>
<td>?</td>
<td>No</td>
</tr>
</tbody>
</table>
Discussion