Link Registry Protocol Options

Norman Finn, March 2017

HUAWEI

At least two obvious choices for P802.1CS Link Registration Protocol

- An “IS-IS-like” protocol.
- TCP (Transmission Control Protocol RFC793).

Common to any choice for LRP

- An “Applicant” at one end of a point-to-point link uses LRP to register information for an “application” in a “Registrar” at the other end of the link.
- Registered Information propagates hop-by-hop over a context.
- Information can be modified as it propagates.
- Registered information can be controlled from a central controller, as in P802.1Qcc.
- No support for shared media, at present.
- Loss of a link requires reinitialization and retransmission of the whole database.

LRP requirements

- Optimized for a database of registrations of ~ 1 Mbyte per port, according the P802.1CS PAR.
- Must have some kind of “keep alive” to time out information.
- Must operate very quickly if the database is small, and as quickly as possible for larger databases.
- Acknowledgement of successful reservations is essential.

An “IS-IS-like” protocol

- ISO 10589 defines the “Intermediate System to Intermediate System intra-domain routeing information exchange protocol…”
- This has similar requirements for conveying information hop-by-hop.
- ISIS defines a routing protocol that propagates address information through a network; we are interested only in the technique used to convey information hop-by-hop, hence, “IS-IS-*like*”, instead of “IS-IS”.
- Keep-alive is done with periodic transmissions as in IS-IS.

IS-IS-like hop-by-hop

- ISIS operates at Layer 2, using (for Ethernet) an LLC encoding (LLC = FE-FE-03), destination MAC addr 01-80-C2-00-00-14/5.
- The information to be registered is divided up, by the Applicant, into Blocks (IS-IS calls them ”LSPs”). A Block must fit into a frame. A frame can carry multiple Blocks. Each Block has an ID to distinguish it from other Blocks from that Applicant, and s sequence number to distinguish revised information.
- Each Block has a checksum.
- Acknowledgements are based on returning lists of Block ID, sequence number, and checksum.

IS-IS-like hop-by-hop

- The Applicant decides how to divide the database to be registered into Blocks.
- A Block must fit into a frame with ~ 60 bytes of header.
- If any data in a Block changes, the whole Block must be retransmitted. This argues for small Blocks.
- It takes ~16 bytes to acknowledge a Block. For 1 Mbyte of data, this argues for larger Blocks.

IS-IS-like issues

- Because ISIS operates at Layer 2, it requires some means to distinguish its PDUs on each medium. For Ethernet, LRP would use an EtherType, and presumably, one of the 16 reserved addresses, e.g. 01-80-C2-00-00-0E.
- ISIS has no means (other than configurable fixed rate parameters) to slow down its transmission rate to meet the capabilities of the Registrar. This is a problem for a 1 Mbyte data base (about 700 frames).

TCP

- TCP operates at Layer 4. It runs over IP.
- For LRP to use TCP, the Applicant would, presumably:
  - Divide up the database in some manner, so that a small change does not require retransmitting the whole database. For the moment, let's assume that means Blocks, as in IS-IS-like LRP, but that the Blocks are not limited in size.
  - Use top-level acknowledgements that would, again, be a summary of Blocks, sequence numbers, and checksums.
  - Transfer the blocks and acknowledgements over TCP, in order to get pacing and retransmissions.

TCP

- Because TCP operates over IP, the participants of LRP require IP addresses, and an IP stack.
- But, LRP defines its neighbors in L2 terms. A discovery protocol is therefore required. Presumably, this would be IEEE Std 802.1AB LLDP with a TLV for “LRP Registrar and/or Applicant address”.
- Keep-alive probably done at top-level as for IS-IS-like protocol, rather than with TCP keep-alive function.

Comparisons

- This author believes that the lack of transmission pacing in ISIS presents a real problem when transmitting hundreds of frames.
  - If transmitted at full rate, a large fraction may be dropped by the Registrar, and have to be transmitted, again. (And again, a large fraction may be lost and retransmitted.)
- TCP is certainly more complex, but it is very widely implemented, and provides well-known automatic methods for transmission pacing to reduce unnecessary packet retransmissions.
- TCP over IP requires an IP stack, and an L2 discovery protocol (LLDP).

Choices

- We could decide to do the IS-IS-like protocol more or less as presented here.
- We could decide to do the TCP/IP protocol more or less as presented here.
- We could add more TCP-like transmission pacing to IS-IS.
- We could invent “TCP-over-Ethernet” (there are precedents).
- Presentations on other choices would be in order.

Thank you