draft-ginsberg-lsr-isis-flooding-scale

Les Ginsberg, Cisco
Peter Psenak , Cisco
Acee Lindem, Cisco
Tony Przygienda, Juniper

Changes Since V1

Revised example algorithm

Added discussion of Ack behavior (PSNP Rate)

Added Deployment Considerations

Example Flow Control Algo

```
MaxLSPTx:          Maximum # LSPs transmitted/second/interface
MinLSPTx:          Minimum # LSPs which may be transmitted/second/interface
UackSafe:          Safe level of unacknowledged LSP/Interface expressed as a percentage
                   of CurrentLSPTxMax(1-99)
UpdateBackoff:     Percent backoff when congestion occurs (1-99)
UpdateIncrement:   Percent increment when congestion has cleared (1-99)

CurrentLSPTxMax:   Current maximum number of LSPs which can be transmitted/second
CurrentUackLSP:    Current number of unacknowledged LSPs already transmitted

if (CurrentUackLSP > (CurrentLSPTxMax * UackSafe)) {
     CurrentLSPTxMax = max(MinLSPTx, (CurrentLSPMaxTx*UpdateBackoff))
} else { // CurrentUackLSP is at a safe level
     CurrentLSPTxMax = min(MaxLSPTx, CurrentLSPTxMax*((100 + UpdateIncrement)/100))
}
```

Configurable **vs** Calculated

Sending PSNPs

Tradeoff between acking immediately and delaying so as to minimize the number of PSNPs sent

ISO10589:

partialSNPInterval - This is the amount of time between periodic action for transmission of Partial Sequence Number PDUs. It shall be less than minimumLSPTransmission-Interval. The recommended value was 2 seconds.

Delay needs to be reduced.

Deployment Considerations

Inconsistent flooding rates have the potential to lengthen the period of LSPDB inconsistency in the network.

This increases duration of blackholes/loops.

Recommend not enabling faster flooding until all nodes in the network support it.

Can be enabled per area.

Discussion Points

“Optimize Goodput”

- Tony Li

Barriers to Implementation (hardware, dataplane QOS)

Issues with Static Controls

Comparisons to TCP

The Characteristics of IGP Flooding (Instability Bursts, Node Introduction/Removal/Maintenance)

Flooding Speed Goals (Target Speed, Consistency)

Barriers to Implementation

A solution which requires hardware/dataplane changes presents a higher bar

- QOS Specific to IS-IS PDUs is not widely available – particularly on receive
- Real time communication of dataplane state as regards IS-IS PDUs (queue state, drops, per interface statistics) is not commonly available
- Mapping hardware specific behaviors into a common notification to the protocol
- Rx based flow control depends on such data in order to provide optimal flow control

A solution where feedback is internal to the protocol avoids hardware/dataplane dependencies

- Tx based flow control uses data already available internal to the protocol
- Per interface statistics are inherent

Issues with Static Controls

What impacts the size of the LSPDB and number of PDUs which can be received?

- Number of nodes in the network
- Number of neighbors
- Flooding optimizations supported (mesh groups, parallel neighbor suppression, dynamic flooding) by each neighbor
- Other protocols (BGP, BFD, OAM, link PM)
- Link bandwidth
- Hardware speed/memory
- SRLG deployment
- …

Optimal Static control of flooding rate easily explodes into a very large number of cases

Comparisons to TCP

TCP	IS-IS
Byte Stream	Packet Based
Ordered delivery	Unordered delivery
Single independent data stream	Multiple interface streams
Resources managed by control plane	Resources dependent on dataplane

IS-IS Flooding Characteristics

Stable Topology

Refreshes. Distributed more sparsely at scale by using longer lifetimes

- default: 20 minutes
- Maximum: 18 hours

Link Topology Changes

Small number of LSPs updated (depends on optimal LSP Generation)

Multiplied (not-linearly) due to shared fate (SRLG)

Node State Changes

Node Up: Full LSPDB sync (Graceful Startup makes this less time critical)

Node Failure: Driven by number of neighbors

Maintenance: Similar to node failure – but can be mitigated by graceful shutdown techniques)

Flooding Speed Goals

Goal 1: Order of magnitude increase in flooding speed

- Currently in 10s of LSPs/second
- Goal in hundreds of LSPs/second
- Thousands of LSPs/second seems aggressive and likely not needed

This helps define the adjustment interval needed for flow control

Goal2: Keep flooding rate interface independent when possible