Bandwidth constrained bypass routing without reservation

draft-szarecki-teas-bw-aware-bypass-no-reservation

Rafal Szarecki
szarecki@google.com
Disclaimer

I’m speaking in behalf of myself.

Any content and information must not be associated with my employer.
Bypass without bandwidth reservation.

- Bypass w/o reservation takes shortest-path even if
  - Links are already 100% used/reserved
  - And link capacity is too small to carry protected traffic
- Congestion losses when bypass in use

[Diagram showing network flow and bandwidth usage]
Bypass with bandwidth reservation - Cost Prohibitive

- Reserved bandwidth is NOT USED most of time
  - ∼ 5 minutes after failure, and then nothing till next failure...
- Reservation for multiple bypasses are independent and additive.
- Highly overprovisioned network.

**Available bandwidth:** 0 bps

**Total reservable bandwidth:** 100Gbps

**Active reservations:**
- Blue bypass: 30 Gbps
- Violet bypass: 35 Gbps
- Revenue LSP: 35 Gbps

65 Gbps is wasted, not usable for revenue traffic.

NOTE: Under SPOF, only either blue or violet bypass is in use.
Decouple CSPF path computation form bypass signalling
  ○ Calculate path ERO with X bps bandwidth constraints against TED
  ○ Use above ERO to signal path with Y bps bandwidth.
  ○ The Y = f(X) such Y < X; Y could be constant, including 0

Derive value of X automatically - Protected Bandwidth (PBW)

Optimize Bypasses infrequently
Derive bypass Protected Bandwidth (PBw) automatically

- The PLR (bypass head) knows LSP->Bypass-tunnel (N:1) mapping and LSPs bandwidths.
- PLR periodically scan all MPLS-TE LSP and sum-up its bandwidth per Bypass-tunnel - Protected Bandwidth sample.
  - Interval < Bypass-tunnel reoptimization timer (⅓ or lower).
  - (Optional) smoothing function over most recent samples.
  - When Bypass-tunnel path need to be calculated, PBw value is used as CSPF constrain (only).
- Statically configured minimum bandwidth for CSPF (min_PBw)
  - New Bypass-tunnel
  - Fall-back if CSPF returns no path for dynamic PBw
- Other procedures possible
Bypass path calculation and signalling

- Calculate path ERO with PBW bps bandwidth constraints against TED.

- Use above ERO to signal path with Y bps bandwidth.
  - Y = f(PBW)
  - Y may be 0 (all further slides and consideration).

The r2>r4 and r4>r5 has available bandwidth (40) for:
  - More bypasses
  - New protected LSP

Time T0:
Blue bypass processing on r1
TED:
- r1>r2: available/total bw: 100/100
- r1>r4: available/total bw: 10/10
- r1>r5: available/total bw: 100/100
- r2>r4: available/total bw: 40/100
- r3>r2: available/total bw: 100/100
- r3>r4: available/total bw: 10/100
- r3>r5: available/total bw: 65/100
- r4>r5: available/total bw: 40/100
CSFP calc:
- Bw required "30"
- r1>r4 and r1>r5 pruned
ERO: r1, r2, r4, r5
RSVP signal
- Bw reservation "0"
ERO: r1, r2, r4, r5

Time T10:
Violet bypass processing on r3
TED:
- r1>r2: available/total bw: 100/100
- r1>r4: available/total bw: 10/10
- r1>r5: available/total bw: 70/100
- r2>r4: available/total bw: 40/100
- r3>r2: available/total bw: 100/100
- r3>r4: available/total bw: 10/100
- r3>r5: available/total bw: 65/100
- r4>r5: available/total bw: 40/100
CSFP calc:
- Bw required "35"
- r3>r4 and r3>r5 pruned
ERO: r3, r2, r4, r5
RSVP signal
- Bw reservation "0"
ERO: r3, r2, r4, r5

No changes in TED
Optimize Bypasses infrequently

- Calculating Bypass-tunnels BW and then path (ERO) can be computationally expensive.
- Triggers for Bypass-path computation
  - Periodic optimization (recommended to implement separate optimization timers for MPLS-TE LSP and Bypass-tunnels)
  - Bypass-Tunnel failure (get PathErr/PathTear) due to e.g. failure/drain of link traversed by bypass-tunnel.
  - Initialization of new bypass-tunnel.
  - Change in Bypass-tunnel bandwidth due to change of Sum of protected LSP Bandwidth
    - Too much churn
    - Not recommended
    - Area of future improvements
Applicability

- Not the perfect solution
  - Bypass-tunnel path is correct only at time of computation. Links available Bw may change right after it.
  - Multiple bypass-tunnels (e.g. from different PLR) protecting same SLRG may use same available bandwidth - risk of overloading.
    - Avoid “least-fill”, “most-fill”
- Local PLR behaviour - no changes to protocols
  - Gradual deployment, immediate benefits
  - No need for interoperability among implementations
- Dot capacity guarantee, but higher probability of no congestion.
Preview of -01

- Bandwidth-aware
  - -00 - PBW & Available-Bw aware, TE-metric optimized
  - -01 add alternative approach - Available-Bw only awate, unreserved BW optimized
  - More contributors

- Route bypass along max-available-capacity path
  - Simplified max-flow problem
  - Round max-available-capacity to “equalize” paths of similar max-available-capacity
  - Limits on: number of Hop, path TE-cost, etc