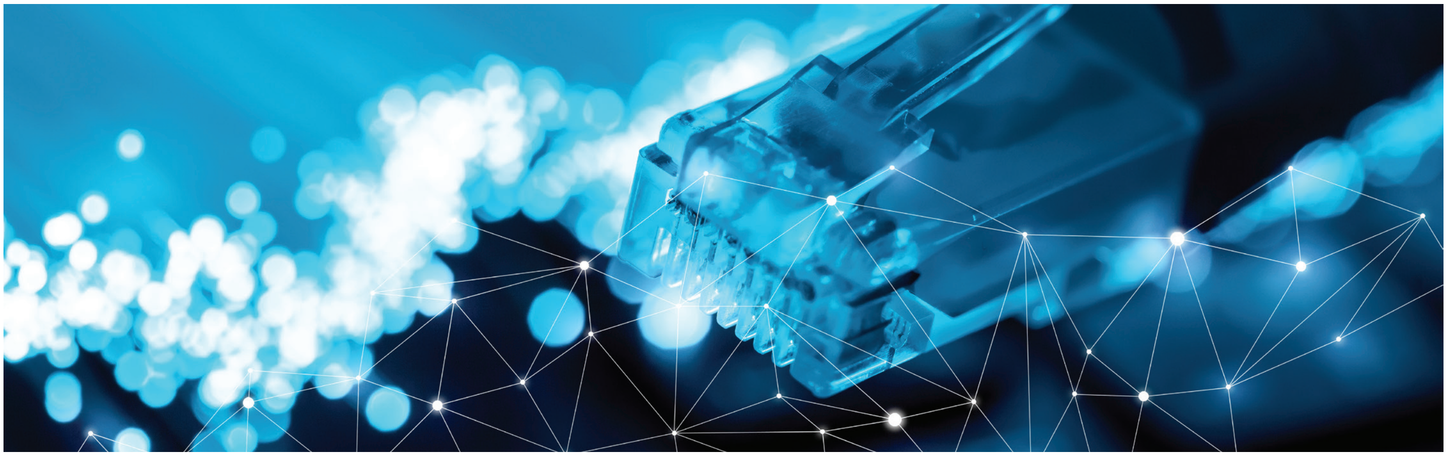
WHITE PAPER

The Case for Segment Routing in the Utility Industry

By Andrew Silvius, PE, and Dustin Williams, PE

Utility networks that use label distribution protocols have proved reliable over time and capable of transporting mission-critical traffic, but those protocols also introduce operational complexity and excess network overheads. Software to better manage traffic priority and flow can alleviate some of these challenges. Segment routing on existing multiprotocol label switching (MPLS) networks can deliver those improvements.



Utility communications networks have evolved to support more sophisticated applications and now require technologies that scale to thousands of sites. The modern utility's wide area network (WAN) resembles a regional carrier network in scale but needs to carry mission-critical operations traffic. Carrier networks are leading the way to simpler operations and addressing the need for better controls faster than standard control protocols can provide. Better control of traffic flow and priority through software is an answer to the problem, and it can be implemented on existing multiprotocol label switching (MPLS) networks using segment routing (SR).

Implementing SR can simplify network deployment and eliminate manual traffic engineering that can be complex and time-consuming. SR streamlines configuration, reduces the challenges of breaking traffic engineering paths, and mitigates network outages. This paper compares label distribution protocol (LDP), resource reservation protocol (RSVP) and SR. Furthermore, it illustrates the main reasons for the utility adoption of SR. Benefits include less manual configuration and fewer protocols to configure and troubleshoot.

Internet protocol (IP)/MPLS networks use protocols like LDP and RSVP, and now SR, as technologies to create and distribute labels that forward customer/application packets through the network without per-hop routing lookups. Instead, the router uses a smaller table called the label forwarding information base (LFIB). These tables use less memory and forwarding time than routing lookups and tables. This makes switching packets through networks more efficient.

Label Distribution Protocol (LDP)

One of the most fundamental components of MPLS networks is the distribution of labels throughout the network, and the simplest protocol to achieve this is LDP. Once labels are distributed, label switched routers (LSRs) switch traffic from one destination to another in the MPLS network. LSRs will perform actions based on the label-determined path as the packet traverses the network. Traffic ingress to an MPLS network without a label will have a label pushed onto the packet. This label is assigned to a class, referred to as the forwarding equivalence class (FEC), and all traffic matching the FEC follows the same forwarding path through the network.

Because LDP is a simple protocol, LSRs have no end-to-end view of the tunnels traversing the node or the associated overhead of managing a stateful tunnel. An LSR only has label actions from ingress to egress across itself and associated FECs. The introduction and use of labels require each node to maintain a database to manage the label/FEC pairings on the LSR.

An LSR has different methods to create labels for local addresses and any other FEC for receiving a label. These methods are called downstream on-demand and downstream unsolicited. LSR behavior generates labels for local residing addresses, such as the loopback or management IP of the LSR, and distributes to other LSRs without an explicit request. This is downstream unsolicited label distribution. Network operators can alter this LSR behavior and create additional labels for additional local addresses to be populated to the rest of the MPLS network, if desired.

For any traffic flow, traffic is forwarded to the next-hop neighbor using the label in the LFIB matching the destination prefix. The receiving router then installs the prefix and label in its route table and LFIB, respectively. A packet traversing an MPLS network with labels set up and distributed utilizing LDP will always follow the interior gateway protocol (IGP) best path through the network. The network's convergence time and failure detection time using LDP rely on the IGP convergence and failure detection times. Implementing bidirectional forwarding detection (BFD) on the IGP links reduces IGP failure detection from seconds to milliseconds. BFD is one method for a more robust IGP deployment, and thus more robust LDP deployment.

While LDP is a solution that can provide adequate transport label distribution, the drawback is that it follows the IGP with no control other than to turn off LDP on links in the network intended for exclusion. This exclusion is manual and limits the ability to provide the actual shortest paths. The modern practice uses traffic engineering (TE) to control services like protective relaying. While LDP is simple to deploy, it lacks active-standby tunnels and traffic engineering support. It therefore extends network convergence times. However it is simplified, the inability to steer traffic based on link metrics and bandwidth requirements makes LDP a less desired transport label distribution protocol in utility networks.

Resource Reservation Protocol (RSVP)

RSVP-TE is another transport label distribution protocol implemented on IP/MPLS networks to signal label switched paths (LSPs). One of the main features that makes RSVP-TE desirable for a utility is providing active protection tunnels. These tunnels provide alternate paths through the network if the primary route fails because of fiber or system failures. RSVP-TE configures path constraints such as bandwidth reservation, hop counts and admin groups. These different types of controls allow bandwidth assurance and the ability to define and restrict paths to predefined routes in the network, referred to as RSVP-TE for traffic engineering. LSPs allow for a few different types of path selection. The simplest form is using fast reroute (FRR), which comes with two methods: facility FRR and one-to-one FRR.

Both methods of FRR use a constrained shortest path first (CSPF) algorithm to calculate the shortest path through the network to reach the configured destination. CSPF uses the IGP and the traffic engineering database (TED) to find the shortest path using RSVP constraints that the network administrator configures. Facility FRR is unique in protecting path failures from a node failure and the final link connecting to the destination or egress node. If a node does fail, the path bypasses the failed node and quickly, through local repair, finds a way back on the original path.

One-to-one FRR also uses node and link protection but differs by creating LSPs that detour around failed nodes or the final link from each node in the path of the service. In other words, every node in the traversal path creates a detour LSP around the failure node or final link. The detour LSP defines a path that is the shortest path from the node located before the failure point to the egress node of the service, even if it's not on the original path. What is important to note is that both methods of FRR use local repair and reduce the need to signal the headend router to complete the restoration of the path, therefore accomplishing service restoration in under 50 ms.

This feature of FRR is the primary reason RSVP is so desirable to utilities. RSVP also provides the capability to define paths through the network and steer the traffic as the network administrator requires by defining each path as a primary, with or without a standby path. These defined network paths also achieve sub-50 ms restoration as long as the two paths through the network are separated and the secondary path is active as standby. Utilities use these traffic engineering methods to transport protective relaying as the defined paths are known, and failover is predictable.

Segment Routing (SR)

Traditional IP/MPLS networks that utilize LDP and RSVP have been well suited for the utility industry for the past decade. However, the use of label distribution protocols such as LDP and RSVP often have added complexities, making it difficult for network operators to troubleshoot and scale. The introduction of SR through the Internet Engineering Task Force (IETF) and Source Packet Routing in Networking (SPRING) working groups aimed to address these concerns.

These working groups charted a course for SR to provide quicker transit with less configuration and more automation on the network. Using a controller to make network-based decisions and utilize automation is often grouped into a broader term: software-defined networking (SDN). Segment routing is used on a per-node basis, but you can also use a path computational element (PCE) to create an SDN. A PCE is a device — such as a computer or a network node — that can calculate sophisticated routing decisions for a network based on constraints such as metrics, latency and jitter. The PCE is centralized and possesses more processing power and memory than a typical network element. The advantage of a PCE is that it has a holistic network view. Therefore, the PCE communicates routing changes to the whole network and changes a distributed system algorithm to a centralized algorithm. This centralized processing enables a network to make calculated decisions faster to orchestrate network routing decisions, rather than waiting for multiple nodes to converge. SDNs become possible by combining PCE with SR.

SR emerged in 2013, while MPLS has been used since 1999. SR is used primarily by hyperscaled web providers and large companies. These companies prefer SR for its ability to precalculate and create paths for point-to-point or layer 3 services, referred to as SR paths. These SR paths optimize network process and memory use with less overhead than IP/MPLS. The main attraction of implementing SR in an MPLS network is eliminating additional label distribution protocols, like LDP and RSVP, and working toward a central control using a PCE. SR's flexible definition of end-to-end paths within IGP topologies is accomplished by encoding these paths as sequences called segments. These segments are made possible through IGP extensions to link-state protocols. Extensions add optional information elements called type-length-values (TLVs) to the link state information sent to other routers. These TLVs allow routers within the SR domain to communicate additional information and capabilities, including segment identifiers (SIDs).

SR uses different types of SIDs across the network. The first is the prefix SID. The prefix SID sub-TLV is associated with a prefix advertised by the router and must be unique in an IGP domain. A node SID is a particular type of prefix SID often associated with the router loopback. Another SID type is the adjacency SID. A router may assign the adjacency SID for each adjacency between two nodes within a given IGP topology. Unlike prefix SIDs, adjacency SIDs are locally significant in an SR domain. The adjacency SID value may be utilized on another router in the SR domain. Global and local segment labels are assigned by the segment routing global block (SRGB) and the segment routing local block (SRLB). These ranges are defined in an MPLS-SR domain throughout the network, and it is common practice to set the same SRGB and SRLB ranges across all routers within an SR domain. These SIDs compute the source routing to an endpoint in the network after the SR network has converged and acts like the labels traditionally distributed in an MPLS network deployment.

Traffic Differences

For IP/MPLS, a packet traversing the network stops at each LSR along the network path for processing to determine the next hop. The process first pops the ingress label from the label stack and queries it against the LFIB. The next-hop label found from the query pushes onto the top of the label stack, and the LSR sends the packet out to the associated next-hop interface. Each router along the service path performs the same function until the packet arrives at the destination. An analogy can explain the difference between a packet traversing an MPLS network versus an SR network.

Imagine driving to work in the morning. Every time you reach an intersection, you must provide a note to a traffic director to determine the next route. At the first intersection, the traffic director asks to see the note. With the appropriate information to look up the next route, the director gives you the next direction and another note to take to the next intersection. The journey continues; this action is repeated at every intersection until the destination. This process requires as many stops as there are intersections along the route. These multiple stops add time to the overall trip, and the longer the trip, the more time is spent exchanging information. This trip would be analogous to a journey through an IP/MPLS network.

Applying the same analogy to SR, consider the same trip to work through an SR network. Before leaving the point of origin, the initial note contains instructions on what direction to take at every intersection along the route. The result is no interactions with traffic directors, as the path is predetermined. This approach saves time and reduces overhead. The behavior is analogous to the implementation of the SR label stack.

As mentioned previously, the SR protocol has determined the path through the network upon entry to an SR network. An ordered list of instructions, in the form of SIDs, is calculated by the IGP, encoded as labels, and pushed onto the packet, guiding the network traffic. The network administrator can also choose to engineer the path using SR policies. These policies might diverge from the shortest path calculation determined by the IGP to recover from link failures or bypass a specific link. Additionally, SR policies may trigger changes in network paths by dynamic or static conditions in the network.

In a standard IGP network, reconvergence of a network after a link failure could take from hundreds of milliseconds to tens of seconds. This time frame for recovery is not acceptable to mission-critical utility application traffic. SR networks might utilize topology-independent loop-free alternate (TI-LFA) to improve the resiliency and reliability of the operations network. TI-LFA reduces the reconvergence time to tens of milliseconds by precalculating alternate next-hops if the active next-hop path fails. This recovery time closely mirrors capabilities provided by RSVP FRR in traditional MPLS networks.

As utility operations and network requirements continue to evolve, new protocols like SR improve networks by addressing growing challenges of more traffic with better convergence and sustained resiliency. SR can implement networks in place of LDP and RSVP for label distribution to provide ease of use and feature-rich policies requiring less manual configuration and overhead to create those configurations.

LIST OF ACRONYMS

BFD: bidirectional forwarding detection
CSPF: constrained shortest path first
FEC: forwarding equivalence class
FRR: fast reroute
IETF: Internet Engineering Task Force
IGP: interior gateway protocol
IP: internet protocol
LDP: label distribution protocol
LFIB: label forwarding information base
LSP: label switched paths
LSR: label switched routers
MPLS: multiprotocol label switching
PCE: path computational element
RSVP: resource reservation protocol
SDN: software-defined networking
SID: segment identifier
SPRING: Source Packet Routing in Networking
SR: segment routing
SRGB: segment routing global block
SRLB: segment routing local block
TE: traffic engineering
TED: traffic engineering database
TI-LFA: topology-independent loop-free alternate
TLV: type-length-value
WAN: wide area network

Conclusion

IP/MPLS networks that use label distribution protocols such as LDP and RSVP have proved reliable and capable of transporting mission-critical traffic. However, these protocols often introduce operational complexities and excess network overheads. As a result of these complexities, operational networks may become challenging to manage and maintain.

Migrating existing MPLS networks to an SR MPLS network removes the network overheads of label distribution protocols such as LDP and RSVP. Utility operations networks are often running in a steady state, making them an excellent candidate for SR MPLS networks with TI-LFA. SR MPLS offers the reliability and resiliency that utility operations depend on while removing some overhead complexities. SR MPLS networks also position utilities to quickly adapt and expand network capabilities with further automation in the future.

Adopting SR will simplify day-to-day operations as changes continue to improve SR. Introducing PCE will help automate the network further and move it more toward becoming an SDN.

About Burns & McDonnell

Burns & McDonnell is a family of companies bringing together an unmatched team of engineers, construction and craft professionals, architects, and more to design and build our critical infrastructure. With an integrated construction and design mindset, we offer full-service capabilities. Founded in 1898 and working from dozens of offices globally, Burns & McDonnell is 100% employee-owned. For more information, visit **burnsmcd.com**.