

Network Working Group  
Internet-Draft  
Intended Status: Informational  
Created: August 25, 2010  
Expires: February 25, 2011

O. Komolafe  
Cisco Systems  
A. Farrel  
Old Dog Consulting  
D. King  
Old Dog Consulting

An Analysis of Scaling Issues for Point-to-Multipoint  
Label Switched Paths in MPLS-TE Core Networks

[draft-komolafe-mpls-te-p2mp-scaling-analysis-04.txt](#)

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at <http://www.ietf.org/ietf/1id-abstracts.txt>.

The list of Internet-Draft Shadow Directories can be accessed at <http://www.ietf.org/shadow.html>.

Abstract

Traffic engineered Multiprotocol Label Switching (MPLS-TE) is deployed in providers' core networks, and the scaling properties have been analyzed to show how much control state must be maintained to support a full mesh of edge-to-edge point-to-point (P2P) Label Switched Paths (LSPs) in various network topologies and with several different scaling techniques.

Point-to-multipoint (P2MP) MPLS-TE LSPs are very interesting to service providers as a means to provide multicast services (such as TV distribution, or multicast VPN connectivity) across core MPLS networks. P2MP LSPs have different scaling properties than P2P LSPs, and service providers need to understand whether existing protocols and implementations can support the network sizes and service levels that they are planning in their P2MP MPLS-TE networks.

This document presents an analysis of the scaling properties MPLS-TE core networks that support P2MP LSPs.



Contents

- [1. Introduction](#) ..... [3](#)
- [1.1 Overview](#) ..... [3](#)
- [1.2 Definitions and Glossary of Notation](#) ..... [5](#)
- [1.3. Network Topologies](#) ..... [5](#)
- [1.4. Relationship Probabilities](#) ..... [6](#)
- [2. Issues of Concern for Scaling](#) ..... [7](#)
- [3. A New Method for Calculations for Point-to-Point LSPs](#) ..... [8](#)
- [3.1. Number of LSPs Traversing a P\(1\) Node](#) ..... [8](#)
- [3.2. Number of LSPs Traversing a P\(2\) Node](#) ..... [10](#)
- [4. Analysis of a P2MP LSP with Two Destinations](#) ..... [12](#)
- [4.1. Closest Destination is a Sibling of the Source](#) ..... [12](#)
- [4.2. Closest Destination is a First Cousin of the Source](#) ..... [14](#)
- [4.3. Closest Destination is a Second Cousin of the Source](#) ..... [15](#)
- [4.4. The Average Number of State Segments at P\(1\) and P\(2\) LSRs](#) .. [16](#)
- [4.5. Generic Formula for Number of State Segments at P\(1\) and P\(2\) LSRs](#) ..... [19](#)
- [5. Three PEs in Receiving Set of P2MP LSPs](#) ..... [22](#)
- [6. Any Number of PEs in the Receiving Set of a P2MP LSP](#) ..... [25](#)
- [7. Exemplar Numeric Results](#) ..... [27](#)
- [7.1. Impact of Varying Topological Properties of the Network](#) ..... [27](#)
- [7.2. Impact of Varying the Number of PEs in the Receiving Set](#) .... [28](#)
- [8. Conclusions and Open Issues](#) ..... [30](#)
- [9. Management Considerations](#) ..... [31](#)
- [10. Security Considerations](#) ..... [31](#)
- [11. Recommendations](#) ..... [31](#)
- [12. IANA Considerations](#) ..... [31](#)
- [13. Acknowledgements](#) ..... [31](#)
- [14. References](#) ..... [31](#)
- [14.1. Informative References](#) ..... [31](#)
- [15. Authors' Addresses](#) ..... [32](#)
- [16. Intellectual Property Consideration](#) ..... [32](#)
- [17. Disclaimer of Validity](#) ..... [33](#)
- [18. Full Copyright Statement](#) ..... [33](#)



## **1. Introduction**

Network operators and service providers are examining scaling issues as they look to deploy ever-larger traffic engineered Multiprotocol Label Switching (MPLS-TE) networks. Concerns have been raised about the number of Label Switched Paths (LSPs) that need to be supported at the edge and at the core of the network. The impact on control plane and management plane resources threatens to outweigh the benefits and popularity of MPLS-TE, while the physical limitations of the routers may constrain the deployment options. These issues and concerns are examined in [[RFC5439](#)] for networks carrying point-to-point (P2P) LSPs.

Point-to-multipoint (P2MP) MPLS-TE LSPs are very interesting to service providers as a means to provide multicast services (such as TV distribution, or multicast VPN connectivity) across core MPLS networks. The MPLS-TE and Generalized MPLS (GMPLS) signaling ([[RFC3209](#)] and [[RFC3473](#)]) has been extended to support the control plane establishment of P2MP LSPs [[RFC4875](#)]. P2MP LSPs have different scaling properties than P2P LSPs, and service providers need to understand whether existing protocols and implementations can support the network sizes and service levels that they are planning in their P2MP MPLS-TE networks.

This document presents an analysis of the scaling properties MPLS-TE core networks that support P2MP LSPs.

### **1.1 Overview**

Point-to-multipoint LSPs can be considered as a set of point-to-point fragments that are joined together to form a tree. The tree is rooted at the source, and has a number of leaves (destinations). Each P2P fragment runs between the source and a branch, between a pair of branches, or between a branch and a leaf (in the extreme case, a fragment may also run directly between the root and a branch).

The scaling benefit of a P2MP LSP comes in the fact that only one copy of the data for a set of leaves downstream of a particular branch is transmitted to that branch. This means that only one set of resources (for example, buffers) is consumed on the path upstream of the branch, leaving room for other LSPs to traverse the same path.

In the control plane there are savings, too. At each branch node, it is necessary to hold only one piece of upstream control plane state, and one piece of state for each downstream fragment. This compares with the P2P equivalent where it would be necessary to hold as much upstream state as downstream state.



Note that this is a slight simplification. P2MP control plane state must identify all downstream leaves and may contain certain leaf-specific information such as the paths to those leaves), and this information must be present in the upstream control plane state. Nevertheless, the reduction in quantity of state information is significant, and the effort required within the protocol implementation to maintain the state information is reduced exactly as described.

Thus, in order to quantify the scaling properties of P2MP LSPs in a network, we count control plane state per interface. That is, at any given node on the path of a P2MP tree we count the number of the node's interfaces that participate in the LSP. At the source, this is always one. At a branch node it is  $(1 + n)$  where there is just one upstream interface and  $n$  downstream interfaces. At a leaf we consider just one upstream interface. (Note that the special case of a bud node [[RFC4875](#)] that is a leaf but that also has downstream neighbors in the LSP is not considered in this document. This is a feature of the snowflake topology that is used as the model as described in [Section 1.3](#).)

The approach adopted in this document is to derive some formulae based on the probabilistic measure of where a node lies in the network, and what role it plays in the P2MP LSP. Furthermore, in order to simplify the equations that are derived, certain approximations are introduced - although these are believed to diminish in significance as the size of the network grows. This method is verified by applying it to P2P LSPs and comparing the results with those algorithms derived in [[RFC5439](#)]. As described in [Section 8](#), although this mechanism seems to deliver reasonable results and agrees with the P2P formulae of [[RFC5439](#)], the mathematical model needs further validation. Similarly, it would be valuable to attempt to quantify the approximations in the model.

This document is designed to help service providers discover whether existing protocols and implementations can support the network sizes that they are planning. To do this, it presents an analysis of some of the scaling concerns for MPLS-TE core networks, and examines the value of two techniques for improving scaling. This should motivate the development of appropriate deployment techniques and protocol extensions to enable the application of MPLS-TE in large networks.

This version of this document is limited to the analysis of P2MP LSPs in the artificially symmetric snowflake network, and the comparison of the results with those for P2P LSPs. The document does not address any scaling techniques such as hierarchical LSPs [[RFC4206](#)] (whether P2P or P2MP), nor does it look at protection mechanisms such as Fast Re-route (FRR) [[RFC4090](#)]. These features and their applicability to

other generic network topologies are for future study.

Komolafe, Farrel, and King

[Page 4]



## **1.2 Definitions and Glossary of Notation**

This document applies consistent notation to define various parameters of the networks that are analyzed. These terms are defined as they are introduced though-out the document, but are grouped together here for quick reference. Refer to the full definitions in the text for detailed explanations.

Where possible, these terms are kept consistent with those used in [[RFC5439](#)].

- n A network level.  $n = 1$  is the core of the network.
- $P(n)$  A node at level  $n$  in the network.
- $S(n)$  The number of nodes at level  $n$ , i.e., the number of  $P(n)$  nodes.
- $L(n)$  The number of LSPs traversing a  $P(n)$  node.
- $X(n)$  The number of LSP segment states maintained by a  $P(n)$  node, i.e.,  $X(n) = 2 * L(n)$  for P2P LSPs.
- $x(n)$  A single LSP segment state maintained by a  $P(n)$  node; i.e.,  $X(n) = \text{SUM}[x(n)]$
- $M(n)$  The number of  $P(n+1)$  nodes subtended to a  $P(n)$  node.
- $N$  The number of destination PEs for each P2MP LSP; i.e. the size of the receiving set is  $N$ .

Sibling nodes: PEs subtended to the same  $P(n)$ .

First cousin nodes: PEs subtended to the same  $P(n-1)$  but different  $P(n)$ .

Second cousin nodes: PEs attached to different  $P(n-1)$  nodes.

$C(\text{init})$ : The cost in terms of  $x(n)$  of setting up a P2MP LSP from a source to the closest PE in the receiving set.

$C(\text{inc})$ : The incremental cost in terms of  $x(n)$  of having a subsequent PE in the receiving set.

## **1.3 Network Topologies**

At this stage, analysis is limited to the snowflake topology defined in [[RFC5439](#)]. In this type of network, only the very core of the network is meshed. The edges of the network are formed as trees rooted in the core.

Thus, the snowflake topology is based on a hierarchy of connectivity within the core network. PEs are not directly interconnected, but have connectivity to exactly one  $P(n)$  (dual homing is not considered at this stage). Each  $P(n)$  is connected to precisely one  $P(n-1)$  and to no other  $P(n)$  node. There is one exception: the  $P(1)$  nodes are



assumed to be connected in a full mesh. Figure 1 shows an example snowflake topology.

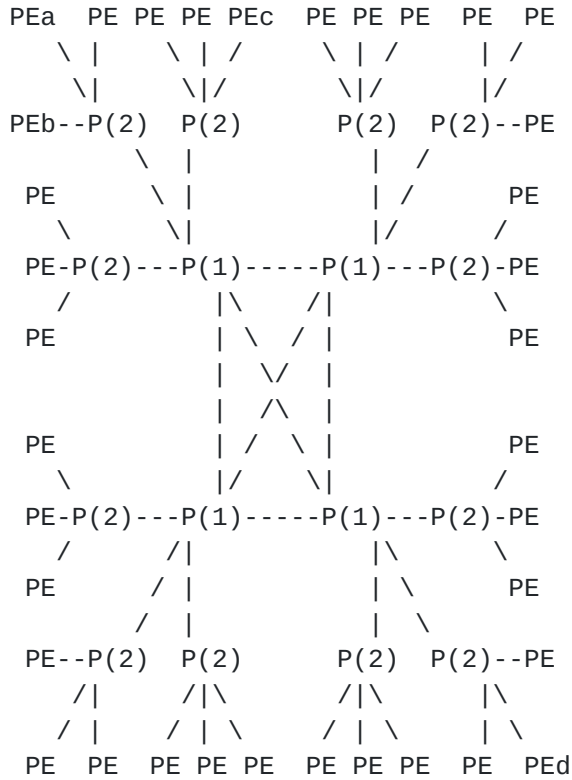


Figure 1 : A Snowflake Network

In Figure 1, the PEs marked PEa and PEb are siblings (subtended to the same P(n) - n=2 in this network). The PEs marked PEa and PEc are first cousins (subtended to the same P(n-1) but to different P(n)s). The PEs marked PEa and PEd are second cousins (subtended to different P(n-1) nodes).

Throughout this document, a snowflake network with 3 levels (i.e., there are P(1), P(2) and PE Label Switching Routers (LSRs)) is considered. However, the methodology employed may be easily applied to networks with a greater number of levels.

**1.4. Relationship Probabilities**

In a snowflake network with 3 levels, consider a given PE: there are a total of S(1)\*M(1)\*M(2) - 1 other PEs, of which M(2) - 1 are siblings. Hence, the probability of another PE chosen at random being a sibling is:

$$\text{Prob(sibling)} = (M(2) - 1) / (S(1)*M(1)*M(2) - 1)$$



Similarly:

$$\text{Prob(first cousin)} = (M(2)*(M(1) - 1)) / (S(1)*M(1)*M(2) - 1)$$

And:

$$\text{Prob(second cousin)} = M(1)*M(2)*(S(1) - 1) / (S(1)*M(1)*M(2) - 1)$$

## 2. Issues of Concern for Scaling

Issues of concern for scaling are discussed in detail in [[RFC5439](#)] and translate to P2MP LSPs. The issues directly affect the number of LSPs that an LSR can support, and therefore may limit the number of LSPs and hence customer services that can be supported by a network.

The issues may be summarized as follows.

### - LSP State

LSP state is the data (information) that must be stored at an LSR in order to maintain an LSP. This is mainly information used by the control plane protocols and grows in direct proportion to the number of LSPs. Since the memory capacity of an LSR is limited, there is a related limit placed on the number LSPs that can be supported.

### - Processing Overhead

The number of LSPs passing through an LSR may impact the processing speed for each LSP. For example, control block search times can increase with the number of control blocks to be searched. Thus, since CPU power is constrained in any LSR, there may be a practical limit to the number of LSPs that can be supported.

### - RSVP-TE Implications

RSVP-TE is a soft-state protocol which means that protocol messages (refresh messages) must be regularly exchanged between signaling neighbors in order to maintain the state for each LSP that runs between the neighbors. Observations of existing implementations indicate that there may be a threshold of around 50,000 P2P LSPs above which an LSR struggles to achieve sufficient processing to maintain LSP state. No observational figures are available for P2MP LSPs, but similar behavior may be expected. Various techniques including refresh reduction [[RFC2961](#)], increasing the refresh interval, use of the RSVP-TE Hello message [[RFC3209](#)], and the use of GMPLS extensions [[RFC3473](#)] (such as the use of [[RFC2961](#)] message acknowledgements on all messages) can improve the situation markedly.



- Management

Another practical concern for the scalability of large MPLS-TE networks is the ability to manage the network. This may be constrained by the available tools, the practicality of managing large numbers of LSPs, and the management protocols in use.

All of these consideration limit the number of LSPs supported within the network and at any particular LSR.

### **3. A New Method for Calculations for Point-to-Point LSPs**

#### **3.1. Number of LSPs Traversing a P(1) Node**

To calculate the number of LSPs seen by each P(1) LSR, consider the case when each PE establishes a single P2P LSP to a destination PE chosen at random. Zero, one, or two P(1) LSRs will be traversed depending on whether the destination is a sibling, first cousin, or second cousin respectively. Hence it may be said that each P(1) LSR traversed will either be:

- an ingress P(1) LSR: such a node is the only P(1) LSR traversed if the destination is a first cousin and the first P(1) LSR traversed if the destination is a second cousin. Hence, each P1 LSR will be an ingress P(1) LSR for each of the  $M(1)*M(2)$  subtended PEs that establishes an LSP destined for a PE that is not a sibling, the probability of which is  $1 - \text{Prob}(\text{sibling})$ .

or

- an egress P(1) LSR: such a node is the second P(1) LSR traversed as a result of the destination being a second cousin. Essentially, an egress P(1) LSR handles the LSPs destined for subtended PEs originating from second cousin PEs. There is a total of  $M(1)*M(2)*(S(1) - 1)$  source PEs which are second cousin nodes to the subtended PEs and the possibility of each of these establishing an LSP to one of the subtended PEs is  $\text{Prob}(\text{second cousin}) / (S(1) - 1)$ .

Figure 2 illustrates examples of ingress P(1) and egress P(1) nodes. The figure shows three exemplar LSPs from a common source to three distinct destinations (d, d1, and d2). d is a sibling of the source, d1 is a first cousin of the source, and d2 is a second cousin of the source. In these LSPs there is one ingress P(1) (marked as  $iP(1)$ ) and one egress P(1) (marked as  $eP(1)$ ).





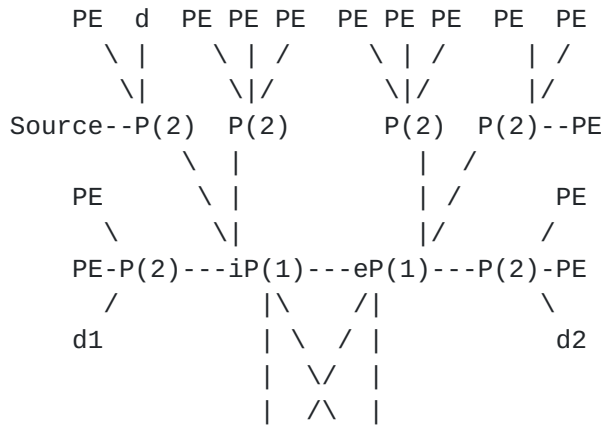


Figure 2 : Illustration of Ingress and Egress P(1) Nodes

Table 1 summarizes the probability of each P(1) LSR being an ingress P(1) or an egress P(1), and shows the frequency with which it may fulfill that role.

Role	Prob(role)	Frequency
Ingress P(1)	1 - Prob(sibling)	M(1)*M(2)
Egress P(1)	Prob(second cousin)/(S(1) - 1)	M(1)*M(2)*(S(1) - 1)

Table 1 : Probability and Frequency of Fulfilling a P(1) Role

Each P(1) LSR will be both an ingress and egress LSR depending on the relative location of the source and destination PEs. Therefore, l(1), the number of LSPs seen by a P(1) LSR if every PE were to establish an LSP to a randomly selected PE is:

$$\begin{aligned}
 l(1) &= M(1)*M(2)*(1 - \text{Prob(sibling)}) + \\
 &\quad M(1)*M(2)*(S(1) - 1) * \text{Prob(second cousin)/(S(1) - 1)} \\
 &= M(1)*M(2)*(1 - \text{Prob(sibling)}) + \\
 &\quad M(1)*M(2) * \text{Prob(second cousin)} \\
 &= M(1)*M(2) * (1 - ((M(2) - 1) / (S(1)*M(1)*M(2) - 1)) + \\
 &\quad M(1)*M(2) * (M(1)*M(2)*(S(1) - 1) / (S(1)*M(1)*M(2) - 1)) \\
 &= M(1)*M(2) * \\
 &\quad (2*S(1)*M(1)*M(2) - M(2) - M(1)*M(2)) / (S(1)*M(1)*M(2) - 1) \\
 &= M(1)*M(2) * (2*S(PE) - M(2) - M(1)*M(2)) / (S(PE) - 1)
 \end{aligned}$$



Lastly,  $l(1)$  is obtained when each PE establishes a single LSP to a randomly chosen destination.  $L(1)$  refers to when LSPs are established to  $S(\text{PE}) - 1$  destinations and so:

$$L(1) = M(1)*M(2)*(2*S(\text{PE}) - M(2) - M(1)*M(2))$$

This result is consistent with the result shown in [section 5.1 of \[RFC5439\]](#).

### **3.2. Number of LSPs Traversing a P(2) Node**

Again, assuming each source PE establishes a single P2P LSP to a randomly chosen destination PE, the manner in which the P(2) LSRs will be traversed by LSPs is described below, illustrated in Figure 3, and summarized in Table 2.

- An ingress P2 LSR is traversed when one of the subtended PEs establishes an LSP. There are  $M_2$  subtended PEs and the probability of each PE establishing an LSP that traverses the given P(2) node is 1.
- A first cousin egress P(2) LSR handles the LSP whenever it is established by one of the first cousin PEs and destined for one of the subtended PEs. There are  $M(2)*(M(1) - 1)$  first cousins and the probability of each of these establishing an LSP to a PE subtended to any given P(2) node is  $\text{Prob}(\text{first cousin}) / (M(1) - 1)$ .
- A second cousin egress P(2) LSR handles the LSP whenever it is established by one of the second cousin PEs and destined for one of the subtended PEs. The number of the former is  $M(1)*M(2)*(S(1) - 1)$  and the probability of the latter is  $\text{Prob}(\text{second cousin}) / M(1)*(S(1) - 1)$ .

Figure 3 shows examples of ingress and egress P(2) nodes. There are three exemplar LSPs from a single source to three destinations ( $d$ ,  $d_1$ , and  $d_2$ ).  $d$  is a sibling of the source,  $d_1$  is a first cousin of the source, and  $d_2$  is a second cousin of the source. The ingress P(2) node is marked  $iP(2)$ . There are two egress P(2) nodes; one is a first cousin egress P(2) (marked  $e1P(2)$  on the figure), and one is a second cousin egress P(2) (marked  $e2P(2)$ ).



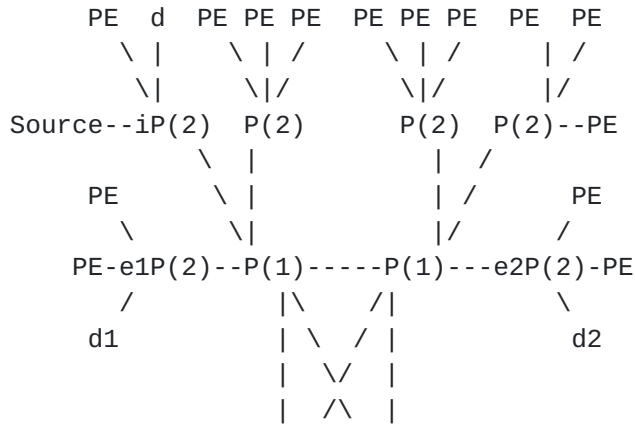


Figure 3 : Example of Ingress and Egress P(2) Nodes

Role	Prob(role)	Frequency
Ingress P(2)	1	M(2)
Egress first cousin P(2)	Prob(first cousin)/(M(1) - 1)	M(2)*(M(1) - 1)
Egress second cousin P(2)	Prob(second cousin) / (M(1)*(S(1) - 1))	M(1)*M(2)*(S(1) - 1)

Table 2 : Probability and Frequency of Fulfilling a P(2) Role

Therefore,  $l(2)$ , the number of LSPs seen by a P(2) LSR if every PE were to establish an LSP to a randomly selected PE is:

$$\begin{aligned}
 l(2) &= M(2) + M(2)(M(1) - 1) * (\text{Prob}(\text{first cousin}) / (M(1) - 1)) + \\
 &\quad M(1)*M(2)*(S(1) - 1) * (\text{Prob}(\text{second cousin}) / (M(1)*(S(1) - 1))) \\
 &= M(2) + M(2)*\text{Prob}(\text{first cousin}) + M(2)*\text{Prob}(\text{second cousin}) \\
 &= ( M(2) / (S(1)*M(1)*M(2) - 1) ) * \\
 &\quad ( S(1)*M(1)*M(2) - 1 + M(1)*M(2) - M(2) + \\
 &\quad\quad S(1)*M(1)*M(2) - M(1)*M(2) ) \\
 &= (M(2) / (S(1)*M(1)*M(2) - 1)) * (2*S(1)*M(1)*M(2) - M(2) - 1) \\
 &= M(2) * (2*S(PE) - M(2) - 1) / (S(PE) - 1)
 \end{aligned}$$



Therefore:

$$L(2) = M(2) * (2*S(PE) - M(2) - 1)$$

This result is also consistent with the conclusion reached in [section 5.1 of \[RFC5439\]](#). We can, therefore, assume that this approach is viable P2P MPLS-TE LSPs. The next sections extend this approach to P2MP LSPs.

#### 4. Analysis of a P2MP LSP with Two Destinations

Consider a P2MP LSP with a set of receiving nodes (destinations) called d1 and d2. If we count the number of hops from the source to each destination, we are able to designate d1 as the closer to the source, and d2 as being equidistant or further from the source.

According to our definitions, d1 must be a sibling, a first cousin, or a second cousin of the source. We consider these three possibilities in turn.

##### 4.1. Closest Destination is a Sibling of the Source

The probability that d1 is a sibling of the source is

$$\text{Prob(sibling)} = (M(2) - 1) / (S(1)*M(1)*M(2) - 1)$$

The cost associated with this possibility is that each P(2) LSR holds two state segments (one corresponding to the input interface and one for the output interface), giving  $C(\text{init}) = 2*x(2)$ .

The incremental cost of reaching d2,  $C(\text{inc})$ , is dependent on its relative location to d1. As shown in Figure 4, there are three distinct possibilities:

- d2 is a sibling of d1 (and so also a sibling of the source)

Excluding the source and d1, there is a total of  $S(1)*M(1)*M(2) - 2$  PEs of which  $M(2) - 2$  are siblings of d1 and the source. Thus, for d2 the probability of being a sibling is:

$$\text{Prob(sibling)} = (M(2) - 2) / (S(1)*M(1)*M(2) - 2)$$

Replication must occur at the P(2) LSR for the P2MP LSP to reach d2 and so the incremental cost is an extra state segment at the P(2) LSR, thus  $C(\text{inc}) = x(2)$ .





- d2 is a first cousin of d1 (and the source)

The probability that d2 is a first cousin of the source is

$$\text{Prob}(\text{first cousin}) = (M(2) * (M(1) - 1)) / (S(1) * M(1) * M(2) - 2)$$

Replication occurs at the first P(2) LSR, giving an extra state segment there, x(2). Since d2 is a first cousin, only a single P(1) is traversed, requiring that LSR to support two state segments, yielding 2\*x(1). Lastly, the P(2) LSR to which d2 is attached must be traversed by the LSP, leading to that LSR supporting two state segments, giving 2\*x(1). Aggregating, we arrive at:

$$C(\text{inc}) = x(2) + 2 * x(1) + 2 * x(2)$$

- d2 is a second cousin of d1 (and the source)

The probability that d2 is a second cousin of the source is

$$\text{Prob}(\text{second cousin}) = (M(1) * M(2) * (S(1) - 1)) / (S(1) * M(1) * M(2) - 2)$$

The only difference with the previous scenario is that d2 being a second cousin requires the LSP to traverse an extra P(1) LSR, giving  $C(\text{inc}) = x(2) + 2 * x(1) + 2 * x(1) + 2 * x(2)$

In Figure 4 we show a source and the nearest destination, d1. d1 is a sibling of the source. Three possible locations for d2 are considered: d2a is a sibling of d1 (and of the source), d2b is a first cousin of d1 (and of the source), d2c is a second cousin of d1 (and of the source).

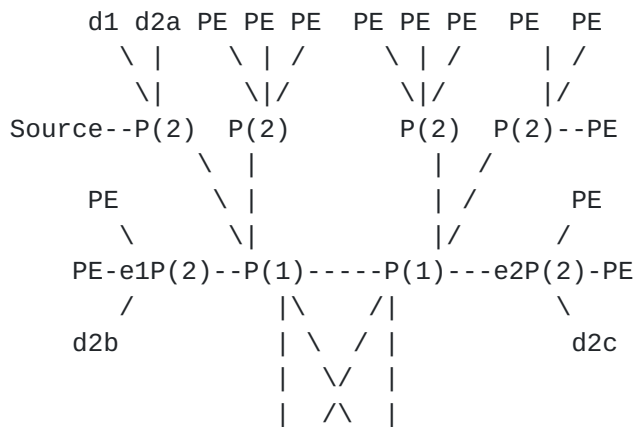


Figure 4 : d1 is a Sibling of the Source, Yielding Three Possibilities for the Location of d2



#### **4.2. Closest Destination is a First Cousin of the Source**

The probability that d1 is a first cousin of the source is:

$$\text{Prob}(\text{first cousin}) = M(2) * (M(1) - 1) / (S(1) * M(1) * M(2) - 1)$$

The associated initial cost is:

$$C(\text{init}) = 2 * x(2) + 2 * x(1) + 2 * x(2).$$

Given the requirement that d1 is equidistant or closer to the source than d2, d2 cannot be a sibling of the source. This leaves the following possibilities, also illustrated in Figure 5:

- d2 is a sibling of d1 (and a first cousin of the source)

$$\text{Prob}(\text{sibling}) = (M(2) - 1) / (S(1) * M(1) * M(2) - 2)$$

$$C(\text{inc}) = x(2)$$

- d2 is a first cousin of d1 (and of the source)

$$\text{Prob}(\text{first cousin}) = M(2) * (M(1) - 2) / (S(1) * M(1) * M(2) - 2)$$

$$C(\text{inc}) = x(1) + 2 * x(2)$$

- d2 is a second cousin of d1 (and of the source)

$$\text{Prob}(\text{second cousin}) = M(1) * M(2) * (S(1) - 1) / (S(1) * M(1) * M(2) - 2)$$

$$C(\text{inc}) = x(1) + 2 * x(1) + 2 * x(2)$$

In Figure 5 we show a source and the nearest destination, d1. d1 is a first cousin of the source. Three possible locations for d2 are considered: d2a is a sibling of d1, d2b is a first cousin of d1, d2c is a second cousin of d1.



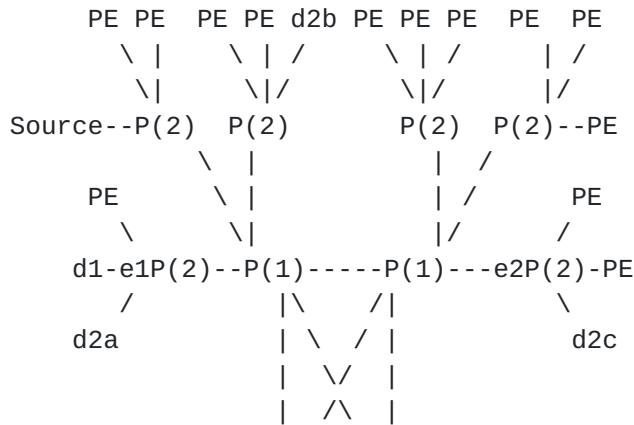


Figure 5 : d1 is a First Cousin of the Source, Yielding Three Possibilities for the Location of d2

**4.3. Closest Destination is a Second Cousin of the Source**

The probability that d1 is a second cousin of the source is:

$$\text{Prob(second cousin)} = M(1)*M(2)*(S(1) - 1) / (S(1)*M(1)*M(2) - 1)$$

The associated initial cost is:

$$C(\text{init}) = 2*x(2) + 2*x(1) + 2*x(1) + 2*x(2)$$

Given the requirement that d1 is equidistant or closer to the source than d2, d2 cannot be a sibling or a first cousin of the source. This leaves the following possibilities, also illustrated in Figure 6:

- d2 is a sibling of d1 (and a second cousin of the source)

$$\text{Prob(sibling)} = (M(2) - 1) / (S(1)*M(1)*M(2) - 2)$$

$$C(\text{inc}) = x(2)$$

- d2 is a first cousin of d1 (and a second cousin of the source)

$$\text{Prob(first cousin)} = M(2)*(M(1) - 1) / (S(1)*M(1)*M(2) - 2)$$

$$C(\text{inc}) = x(1) + 2*x(2)$$

- d2 is a second cousin of d1 (and of the source)

$$\text{Prob(second cousin)} = M(1)*M(2)*(S(1) - 2) / (S(1)*M(1)*M(2) - 2)$$

$$C(\text{inc}) = x(1) + 2*x(1) + 2*x(2)$$



In Figure 6 we show a source and the nearest destination, d1. d1 is a second cousin of the source. Three possible locations for d2 are considered: d2a is a sibling of d1, d2b is a first cousin of d1, d2c is a second cousin of d1.

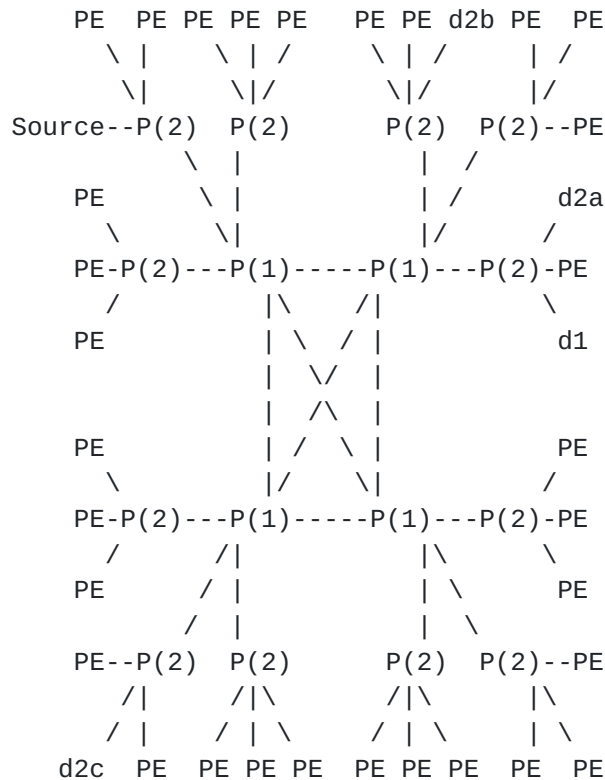


Figure 6 : d1 is a Second Cousin of the Source, Yielding Three Possibilities for the Location of d2

#### 4.4. The Average Number of State Segments at P(1) and P(2) LSRs

Having discussed the nine different permutations of establishing a P2MP LSP from a given PE to two destination PEs in a three-level snowflake topology, the next step is to explore the scalability of this configuration by computing the number of LSP state segments that must be handled by the P(1) and P(2) LSRs. To compute this value, the cost associated with each permutation, obtained by summing C(inc) and C(init) appropriately, has to be weighted by the corresponding probability and the overall cost aggregated. Table 3 gives the nine permutations, the associated costs in terms of state segments that must be managed at P(1) and P(2) nodes and the exact and approximate probabilities of the particular LSP being established.





In Table 3, the following three probabilities must be recalled:

$$\text{Prob}(\text{sibling}) = (M(2) - 1) / (S(1)*M(1)*M(2) - 1)$$

$$\text{Prob}(\text{first cousin}) = M(2)*(M(1) - 1) / (S(1)*M(1)*M(2) - 1)$$

$$\text{Prob}(\text{second cousin}) = M(1)*M(2)*(S(1) - 1) / (S(1)*M(1)*M(2) - 1)$$

For the sake of fitting the table on the page, the equations for the exact probabilities for d2 positions are numbered I through VI and are listed after the table.

d1	d2				LSR state	
-----	+-----				+ segments	
Position   Exact	Position   Exact	Approx	----- -----			
wrt src   probability	wrt d1   prob	probability	P(1)	P(2)		
-----+-----	-----+-----	-----+-----	-----	-----	-----	-----
sibling   Prob(sibling)	Sibling   I	Prob(sibling)	0	3		
sibling   Prob(sibling)	1st cousin   II	Prob(1stcousin)	2	5		
sibling   Prob(sibling)	2nd cousin   III	Prob(2ndcousin)	4	5		
1stcousin   Prob(1stcousin)	Sibling   IV	Prob(sibling)	2	5		
1stcousin   Prob(1stcousin)	1st cousin   V	Prob(1stcousin)	3	6		
1stcousin   Prob(1stcousin)	2nd cousin   III	Prob(2ndcousin)	5	6		
2ndcousin   Prob(2ndcousin)	Sibling   IV	Prob(sibling)	4	5		
2ndcousin   Prob(2ndcousin)	1st cousin   II	Prob(1stcousin)	5	6		
2ndcousin   Prob(2ndcousin)	2nd cousin   VI	Prob(2ndcousin)	7	6		

Equations:

- I (M(2) - 2) / (S(1)\*M(1)\*M(2) - 2)
- II M(2)\*(M(1) - 1) / (S(1)\*M(1)\*M(2) - 2)
- III M(1)\*M(2)\*(S(1) - 1) / (S(1)\*M(1)\*M(2) - 2)
- IV (M(2) - 1) / (S(1)\*M(1)\*M(2) - 2)
- V M(2)\*(M(1)-2) / (S(1)\*M(1)\*M(2) - 2)
- VI M(1)\*M(2)\*(S(1) - 2) / (S(1)\*M(1)\*M(2) - 2)

Table 3 : Summary of Probabilities and Cost of Establishing P2MP LSP To Two PEs

The approximate probabilities essentially assume the probabilities of d2 being in a certain location are independent of the location of d1, which is evidently not the case. However, such an approximation is attractive as it simplifies the equations significantly (a fact more important as the size of the receiving set grows). Furthermore, comparing the approximate and exact probabilities suggest that the approximate probabilities are not unreasonable, and actually would be expected to be close the exact probabilities for large networks (i.e., large values of M(1), M(2), and S(1) should reduce the discrepancy).



Therefore, approximate probabilities are used in the remainder of this document. In consequence, the total approximate number of state segments at P(1) and P(2) LSRs, X(1) and X(2) respectively, may be calculated as follows:

$$\begin{aligned}
 X1 &= 2*\text{Prob}(\text{sibling})*\text{Prob}(\text{first cousin}) + \\
 & 4*\text{Prob}(\text{sibling})*\text{Prob}(\text{second cousin}) + \\
 & 2*\text{Prob}(\text{first cousin})*\text{Prob}(\text{sibling}) + \\
 & 3*\text{Prob}(\text{first cousin})**2 + \\
 & 5*\text{Prob}(\text{first cousin})*\text{Prob}(\text{second cousin}) + \\
 & 4*\text{Prob}(\text{second cousin})*\text{Prob}(\text{sibling}) + \\
 & 5*\text{Prob}(\text{second cousin})*\text{Prob}(\text{first cousin}) + \\
 & 7*\text{Prob}(\text{second cousin})**2 \\
 \\
 &= 4*\text{Prob}(\text{sibling})*\text{Prob}(\text{first cousin}) + \\
 & 8*\text{Prob}(\text{sibling})*\text{Prob}(\text{second cousin}) + \\
 & 3*\text{Prob}(\text{first cousin})**2 + \\
 & 10*\text{Prob}(\text{first cousin})*\text{Prob}(\text{second cousin}) + \\
 & 7*\text{Prob}(\text{second cousin})**2
 \end{aligned}$$

$$\begin{aligned}
 X2 &= 3*\text{Prob}(\text{sibling})**2 + \\
 & 5*\text{Prob}(\text{sibling})*\text{Prob}(\text{first cousin}) + \\
 & 5*\text{Prob}(\text{sibling})*\text{Prob}(\text{second cousin}) + \\
 & 5*\text{Prob}(\text{first cousin})*\text{Prob}(\text{sibling}) + \\
 & 6*\text{Prob}(\text{first cousin})**2 + \\
 & 6*\text{Prob}(\text{first cousin})*\text{P}(\text{second cousin}) + \\
 & 5*\text{Prob}(\text{second cousin})*\text{Prob}(\text{sibling}) + \\
 & 6*\text{Prob}(\text{second cousin})*\text{P}(\text{first cousin}) + \\
 & 6*\text{Prob}(\text{second cousin})**2 \\
 \\
 &= 3*\text{Prob}(\text{sibling})**2 + \\
 & 10*\text{Prob}(\text{sibling})*\text{Prob}(\text{first cousin}) + \\
 & 10*\text{Prob}(\text{sibling})*\text{Prob}(\text{second cousin}) + \\
 & 6*\text{Prob}(\text{first cousin})**2 + \\
 & 12*\text{Prob}(\text{first cousin})*\text{Prob}(\text{second cousin}) + \\
 & 6*\text{Prob}(\text{second cousin})**2
 \end{aligned}$$

These two equations give the number of state segments that must be handled by the P(1) and P(2) LSRs respectively if a source sets up a P2MP LSP to any 2 PEs. The values for Prob(sibling), Prob(first cousin), and Prob(second cousin) from [Section 1.4](#) may be substituted into the equations allowing X(1) and X(2) to be expressed in terms of the topological properties of the network. Furthermore, the two equations for X(1) and X(2) correspond to the establishment of a single P2MP LSP, and so the equations must be appropriately scaled to handle more than one P2MP LSP. For example, the results should be multiplied by S(PE) to reflect every PE establishing an LSP.



Lastly, the equations for X(1) and X(2) give the total number of state segments handled by all P(1) and P(2) LSRs in the network and so must be divided by the number of each type of LSR to obtain the average values. Intuitively, the symmetry in the snowflake network suggests that there will be little variance around the average value provided each PE establishes an equal number of LSPs and the destinations are chosen from an identical statistical distribution.

#### 4.5. Generic Formula for Number of State Segments at P(1) and P(2) LSRs

Let D be the set of possible locations for a given destination in a P2MP LSP and so:

$$D = \{\text{sibling, first cousin, second cousin}\}$$

Let D1 and D2 be members of D. Using the patterns which may be observed in Table 3, generic formulae for X(1) and X(2) may be computed.

$$X1 = \text{SUM}_{D1} [ \text{SUM}_{D2} [ \text{Prob}(d1 = D1) * \text{Prob}(d2 = D2) * (fa(D1) + fb(D1, D2)) ] ]$$

where

$$fa(D1) = \begin{cases} 0 & \text{if } D1 \text{ is sibling} \\ 2 & \text{if } D1 \text{ is first cousin} \\ 4 & \text{if } D1 \text{ is second cousin} \end{cases}$$

$$fb(D1, D2) = \begin{cases} 0 & \text{if } D2 \text{ is sibling} \\ 1 & \text{if } D1 \text{ is not sibling AND } D2 \text{ is first cousin} \\ 2 & \text{if } D1 \text{ is sibling AND } D2 \text{ is first cousin} \\ 3 & \text{if } D1 \text{ is not sibling AND } D2 \text{ is second cousin} \\ 4 & \text{if } D1 \text{ is sibling AND } D2 \text{ is second cousin} \end{cases}$$

This equation for X(1) basically says that to compute the number of state segments at the P(1) LSR, iterate over all possible relative locations for the PEs in the receiving set and, for each permutation, calculate the product of the probability of that permutation occurring (i.e., Prob(d1 = D1)\*P(d2 = D2)) and the associated cost (i.e., (fa(D1)+fb(D1, D2))). Summing the products over all the permutations give the desired metric.

fa(D1) describes the number of state segments that must be handled by P(1) LSRs due to the location of the first PE in the receiving set, d1. No state segments are handled by any P(1) LSR if d1 is a sibling of the source. In this case, the LSP does not need to traverse any P(1) LSR to reach d1, but rather reaches it via the P(2) LSR to which d1 and the source are both attached. If d1 is a first cousin of the



source, the LSP reaches it via the P(1) LSRs to which they are both subtended and so two state segments must be handled by the P(1) LSR. Four state segments are handled by P(1) LSRs when d1 is a second cousin of the source due to the requirement for the LSP to traverse two P(1) LSRs.

fb(D1, D2) describes the incremental cost associated with any replication necessary for the P2MP LSP to reach the second PE in the receiving set, d2. If d2 is a sibling of d1, then no state segments need be handled by any of the P(1) LSRs. On the other hand, if d2 is a first cousin of d1, then one or two additional state segments must be handled by the P(1) LSRs, depending on whether or not d1 is a sibling of the source. If d1 is a sibling, then replication must occur at the first P(2) LSP and, since the LSP is destined for a first cousin, only a single P(1) LSR is traversed and so two state segments must be supported by the P(1) LSRs. When d1 is not a sibling, then it is either a first or second cousin of the source and so replication must occur at the first or second P(1) LSR respectively, with an additional state segment being necessary at that P(1) LSRs to reach d2. Lastly, if d2 is a second cousin of d1 and d1 is a sibling of the source, replication occurs at the first P(2) LSP and, since two P(1) LSRs must be traversed to reach a second cousin, four state segments must be handled by P(1) LSRs. If d2 is a second cousin of d1, and d1 is not a sibling of the source, replication occurs at the first P(1) LSR and, since the LSP must traverse a second P(1) LSR, gives a total of three state segments.

Similarly, for the number of state segments handled by the P(2) LSRs:

$$X(2) = \sum_{D1} [ \sum_{D2} [ \text{Prob}(d1=D1) * \text{Prob}(d2=D2) * (fc(D1) + fd(D1, D2))] ]$$

where

$$fc(D1) = \begin{cases} 2 & \text{if } D1 \text{ is sibling} \\ 4 & \text{otherwise} \end{cases}$$

$$fd(D1, D2) = \begin{cases} 1 & \text{if } D2 \text{ is sibling} \\ 2 & \text{if } D1 \text{ is not sibling AND } D2 \text{ is not sibling} \\ 3 & \text{if } D1 \text{ is sibling AND } D2 \text{ is not sibling} \end{cases}$$

fc(D1) simply states that the first PE in the receiving set, d1, requires two state segments to be supported at P(2) LSRs when it is a sibling of the source, and four states when it is not; two at each of the two P(2) LSR that must be traversed in this case.

According to fd(D1, D2), when the second PE in the receiving set, d2, is a sibling of d1 it results in a single additional state segment





being supported at the P(2) LSR, due to the replication that must occur there. If d1 is neither a sibling of the source nor d2, then replication must occur at a P(1) LSR and so reaching d2 adds two extra state segments at P(2) LSRs, corresponding to traversing the second P(2) LSR. If d1 is a sibling of the source but d2 is not, then replication occurs at the first P(2) LSR and, given the requirement to traverse a second P(2) LSR, produces a total of three state segments being supported by the P(2) LSRs.

These two equations for X(1) and X(2) may be readily expanded to obtain the results given in [Section 4.4](#) and have the advantage of being more easily adaptable as the size of the receiving set increases, an advantage exploited in the remainder of this document.

Furthermore, these two equations may be adapted to consider the cost of establishing a P2P LSP, and so used to validate the equations against the results set out in [\[RFC5439\]](#). For a P2P LSP, there is no d2, thus it may be said that:

$$X(1:p2p) = 2*\text{Prob}(\text{first cousin}) + 4*\text{Prob}(\text{second cousin})$$

This gives the total mean number of state segments at all the P(1) LSRs when a single P2P LSP is established. The values for Prob(first cousin) and Prob(second cousin) from [Section 1.4](#) may be substituted into the equation. To obtain the result matching [Section 3.1](#) and [\[RFC5439\]](#), we must multiply by S(PE) - 1 (since each source establishes an LSP to every other PE), multiply by S(PE) (since there are S(PE) sources), divide by 2 (since the number of state segments is twice the number of LSPs seen by the LSR for P2P LSPs), and divide by S(1) (since there are S(1) P(1) LSRs). This looks as follows:

$$\begin{aligned} L1 &= S(\text{PE}) * (S(\text{PE}) - 1) * \\ &\quad (2*\text{Prob}(\text{first cousin}) + 4*\text{Prob}(\text{second cousin})) / 2*S(1) \\ &= S(\text{PE}) * (S(\text{PE}) - 1) * \\ &\quad (2*M(2)*(M(1) - 1) / (S(1)*M(1)*M(2) - 1) + \\ &\quad 4*M(1)*M(2)*(S(1) - 1) / (S(1)*M(1)*M(2) - 1)) / 2*S(1) \\ &= 2*S(1)*M(1)*M(2)*(S(1)*M(1)*M(2) - 1) * \\ &\quad (M(2)*(M(1) - 1) + 2*M(1)*M(2)*(S(1) - 1)) / \\ &\quad 2*S(1)*(S(1)*M(1)*M(2) - 1) \\ &= M(1)*M(2)*(2*S(\text{PE}) - M(2) - M(1)*M(2)) \end{aligned}$$

Similarly:

$$\begin{aligned} X(2:p2p) &= 2*\text{Prob}(\text{sibling}) + 4*(1 - \text{Prob}(\text{sibling})) \\ &= 2*(2 - \text{Prob}(\text{sibling})) \end{aligned}$$



This gives the total mean number of state segments at all P(2) LSRs when a single P2P LSP is established. This result should be multiplied by  $S(PE) - 1$  (since each source establishes an LSP to every other PE), multiplied by  $S(PE)$  (since there are  $S(PE)$  sources), divided by 2 (since the number of state segments is twice the number of LSPs seen by the LSR for P2P LSPs), and divided by  $S(1)*M(1)$  (since there are  $S(1)*M(1)$  P2 LSRs) to give the same result as found in [Section 3.1](#) and [\[RFC5439\]](#). This is achieved as follows:

$$\begin{aligned}
 L2 &= S(PE)*(S(PE) - 1)* \\
 &\quad 2*(2 - \text{Prob}(\text{sibling})) / 2*S(1)*M(1) \\
 &= S(PE)*(S(PE) - 1)* \\
 &\quad 2*(2 - (M(2) - 1)/(S(1)*M(1)*M(2) - 1)) / 2*S(1)*M(1) \\
 &= 2*S(1)*M(1)*M(2)*(S(1)*M(1)*M(2) - 1) * \\
 &\quad (2*(S(1)*M(1)*M(2) - 1) - M(2) + 1) / \\
 &\quad \quad \quad 2*S(1)*M(1)*(S(1)*M(1)*M(2) - 1) \\
 &= M(2)*(2*S(PE) - M(2) - 1)
 \end{aligned}$$

## 5. Three PEs in Receiving Set of P2MP LSPs

The three nodes in the receiving set are d1, d2, and d3. Once again, d1 is the closest to the source, followed by d2, and lastly d3. Table 4 summarizes all the different ways the P2MP LSP may be established and the respective number of state segments that must be supported by P(1) and P(2) LSRs.

In reading Table 4, recall that:

$$\text{Prob}(\text{sibling}) = (M(2) - 1) / (S(1)*M(1)*M(2) - 1)$$

$$\text{Prob}(\text{first cousin}) = M(2)*(M(1) - 1) / (S(1)*M(1)*M(2) - 1)$$

$$\text{P}(\text{second cousin}) = M(1)*M(2)*(S(1) - 1) / (S(1)*M(1)*M(2) - 1)$$

In order to fit this table on the page, it is necessary to adopt some additional notation conventions as follows.

```

sib  sibling
1c   first cousin
2c   second cousin
p(x) Prob(x)

```

Additionally, the equations for exact probabilities are represented by the roman numerals I-XV and are listed separately.



d1		d2			d3			LSR state	
-----+		-----+			-----+			-----+-----	
Psn	Exact	Psn	Exact	Approx	Psn	Exact	Approx	P(1)	P(2)
wrt	prob	wrt	prob	prob	wrt	prob	prob		
src		d1			d2				
-----+		-----+			-----+			-----+-----	
sib	p(sib)	sib	I	p(sib)	sib	VII	p(sib)	0	4
sib	p(sib)	sib	I	p(sib)	1c	VIII	p(1c)	2	6
sib	p(sib)	sib	I	p(sib)	2c	IX	p(2c)	4	6
sib	p(sib)	1c	II	p(1c)	sib	XI	p(sib)	2	6
sib	p(sib)	1c	II	p(1c)	1c	X	p(1c)	3	7
sib	p(sib)	1c	II	p(1c)	2c	IX	p(2c)	5	7
sib	p(sib)	2c	III	p(2c)	sib	XI	p(sib)	4	6
sib	p(sib)	2c	III	p(2c)	1c	VIII	p(1c)	5	7
sib	p(sib)	2c	III	p(2c)	2c	XII	p(2c)	7	7
1c	p(1c)	sib	IV	p(sib)	sib	XIII	p(sib)	2	5
1c	p(1c)	sib	IV	p(sib)	1c	X	p(1c)	3	7
1c	p(1c)	sib	IV	p(sib)	2c	IX	p(2c)	5	7
1c	p(1c)	1c	V	p(1c)	sib	XI	p(sib)	3	7
1c	p(1c)	1c	V	p(1c)	1c	XIV	p(1c)	4	8
1c	p(1c)	1c	V	p(1c)	2c	IX	p(2c)	6	8
1c	p(1c)	2c	III	p(2c)	sib	XI	p(sib)	5	7
1c	p(1c)	2c	III	p(2c)	1c	VIII	p(1c)	6	8
1c	p(1c)	2c	III	p(2c)	2c	XII	p(2c)	8	8
2c	p(2c)	sib	IV	p(sib)	sib	XIII	p(sib)	4	6
2c	p(2c)	sib	IV	p(sib)	1c	X	p(1c)	5	7
2c	p(2c)	sib	IV	p(sib)	2c	XII	p(2c)	7	7
2c	p(2c)	1c	II	p(1c)	sib	XI	p(sib)	5	7
2c	p(2c)	1c	II	p(1c)	1c	X	p(1c)	6	8
2c	p(2c)	1c	II	p(1c)	2c	XII	p(2c)	8	8
2c	p(2c)	2c	VI	p(2c)	sib	XI	p(sib)	7	7
2c	p(2c)	2c	VI	p(2c)	1c	VIII	p(1c)	8	8
2c	p(2c)	2c	VI	p(2c)	2c	XV	p(2c)	10	8

Table 4 : Summary of Probabilities and Cost of Establishing a P2MP LSP to Three PEs

Equations:

- I  $(M(2) - 2) / (S(1)*M(1)*M(2) - 2)$
- II  $M(2)*(M(1) - 1) / (S(1)*M(1)*M(2) - 2)$
- III  $M(1)*M(2)*(S(1) - 1) / (S(1)*M(1)*M(2) - 2)$
- IV  $(M(2) - 1) / (S(1)*M(1)*M(2) - 2)$
- V  $M(2)*(M(1) - 2) / (S(1)*M(1)*M(2) - 2)$
- VI  $M(1)*M(2)*(S(1) - 2) / (S(1)*M(1)*M(2) - 2)$
- VII  $(M(2) - 3) / (S(1)*M(1)*M(2) - 3)$
- VIII  $M(2)*(M(1) - 1) / (S(1)*M(1)*M(2) - 3)$
- IX  $M(1)*M(2)*(S(1) - 1) / (S(1)*M(1)*M(2) - 3)$
- X  $M(2)*(M(1) - 2) / (S(1)*M(1)*M(2) - 3)$



- XI  $(M(2) - 1) / (S(1) * M(1) * M(2) - 3)$   
 XII  $M(1) * M(2) * (S(1) - 2) / (S(1) * M(1) * M(2) - 3)$   
 XIII  $(M(2) - 2) / (S(1) * M(1) * M(2) - 3)$   
 XIV  $M(2) * (M(1) - 3) / (S(1) * M(1) * M(2) - 2)$   
 XV  $M(1) * M(2) * (S(1) - 3) / (S(1) * M(1) * M(2) - 3)$

Using a similar approach to that used in [Section 4.5](#), the approximate number of state segments that must be handled by P(1) and P(2) LSRs may be readily computed.

Again, Let D be the set of possible locations for a given destination in a P2MP LSP and so

$$D = \{\text{sibling, first cousin, second cousin}\}$$

Using the patterns which may be observed in Table 4, generic formula for X(1) and X(2) may be computed as follows.

$$X(1) = \text{SUM}_{D1} [ \text{SUM}_{D2} [ \text{SUM}_{D3} [ \text{Prob}(d1=D1) * \text{Prob}(d2=D2) * \text{Prob}(d3=D3) * (fe(D1) + ff(D1, D2) + fg(D1, D2, D3))] ] ] ]$$

where

$$\begin{aligned} fe(D1) &= 0 \text{ if } D1 \text{ is sibling} \\ &2 \text{ if } D1 \text{ is a first cousin} \\ &4 \text{ if } D1 \text{ is a second cousin} \end{aligned}$$

$$\begin{aligned} ff(D1, D2) &= 0 \text{ if } D2 \text{ is sibling} \\ &1 \text{ if } D1 \text{ is not a sibling AND } D2 \text{ is a first cousin} \\ &2 \text{ if } D1 \text{ is a sibling AND } D2 \text{ is a first cousin} \\ &3 \text{ if } D1 \text{ is not a sibling AND } D2 \text{ is a second cousin} \\ &4 \text{ if } D1 \text{ is a sibling AND } D2 \text{ is a second cousin} \end{aligned}$$

$$\begin{aligned} fg(D1, D2, D3) &= 0 \text{ if } D3 \text{ is sibling} \\ &1 \text{ if } (D1 \text{ OR } D2 \text{ is not sibling}) \text{ AND } \\ &\quad D3 \text{ is first cousin} \\ &2 \text{ if } D1 \text{ is sibling AND } \\ &\quad D2 \text{ is sibling AND } \\ &\quad D3 \text{ is first cousin} \\ &3 \text{ if } (D1 \text{ is not sibling OR } D2 \text{ is not sibling}) \text{ AND } \\ &\quad D3 \text{ is second cousin} \\ &4 \text{ if } D1 \text{ is sibling AND } \\ &\quad D2 \text{ is sibling AND } \\ &\quad D3 \text{ is a second cousin} \end{aligned}$$





Similarly, for the number of state segments handled by the P(2) LSRs:

$$X(2) = \text{SUM}_{D1} [ \text{SUM}_{D2} [ \text{SUM}_{D3} [ \text{Prob}(d=D1) * \text{Prob}(d2=D2) * \text{Prob}(d3=D3) * (fh(D1) + fi(D1, D2) + fj(D1, D2, D3)) ] ] ]$$

where

$$fh(D1) = \begin{cases} 2 & \text{if } D1 \text{ is sibling} \\ 4 & \text{otherwise} \end{cases}$$

$$fi(D1, D2) = \begin{cases} 1 & \text{if } D2 \text{ is sibling} \\ 2 & \text{if } D1 \text{ is not sibling AND } D2 \text{ is not sibling} \\ 3 & \text{if } D1 \text{ is sibling AND } D2 \text{ is not sibling} \end{cases}$$

$$fj(D1, D2, D3) = \begin{cases} 1 & \text{if } D3 \text{ is sibling} \\ 2 & \text{if } (D1 \text{ is not sibling OR } D2 \text{ is not sibling) AND } \\ & \text{D3 is not sibling} \\ 3 & \text{if } D1 \text{ is sibling AND } \\ & \text{D2 is sibling AND } \\ & \text{D3 is not sibling} \end{cases}$$

It can be readily seen that there are similarities between the equations here and those for X(1) and X(2) in [Section 4.5](#). The main difference is that, in the former equations, when considering the incremental cost of reaching the third PE in the receiving set, d3, the relative position of all the previous PEs in the receiving set must be taken into consideration. These similarities will be exploited to allow generic equations for the mean number of state segments that must be supported at P(1) and P(2) LSRs as a function of the receiving set size in [Section 6](#).

## 6. Any Number of PEs in the Receiving Set of a P2MP LSP

Sections [4](#) and [5](#) have shown the scalability parameters for a single P2MP LSP from a source PE to two and three PEs, respectively. It did this by estimating the average number of state segments that must be maintained by P(1) and P(2) LSRs. This section extends these findings by considering the case where there are an arbitrary number of PEs, N, which are receivers for the LSP.

Comparing our two equations for X(1) (Sections [4.5](#) and [5](#)) we can deduce a generic equation about the mean number of state segments that must be supported at the P(1) LSR when there are N PEs in the receiving set for the P2MP LSP:



$$X(1) = \sum_{D1} [ \dots \sum_{DN} [ \text{Prob}(d1=D1) \dots \text{Prob}(dN=DN) * (fa(D1)+fb(D1,D2)+\dots+fb(D1, \dots, DN))] \dots ]$$

where

$$fa(D1) = \begin{cases} 0 & \text{if } D1 \text{ is sibling} \\ 2 & \text{if } D1 \text{ is first cousin} \\ 4 & \text{if } D1 \text{ is second cousin} \end{cases}$$

$$fb(D1,D2, \dots Dj) = \begin{cases} 0 & \text{if } Dj \text{ is sibling} \\ 1 & \text{if any of } D1, D2 \dots Dj-1 \text{ is not sibling AND} \\ & \text{Dj is first cousin} \\ 2 & \text{if all of } D1, D2 \dots Dj-1 \text{ are sibling AND} \\ & \text{Dj is first cousin} \\ 3 & \text{if any of } D1, D2 \dots Dj-1 \text{ is not sibling AND} \\ & \text{Dj is second cousin} \\ 4 & \text{if all of } D1, D2 \dots Dj-1 \text{ are sibling AND} \\ & \text{Dj is second cousin} \end{cases}$$

Similarly, for the number of state segments handled by the P2 LSRs,

$$X(2) = \sum_{D1} [ \dots \sum_{DN} [ \text{Prob}(d1=D1) \dots \text{Prob}(dN=DN) * (fc(D1)+fd(D1,D2)+\dots+fd(D1, \dots, DN))] \dots ]$$

where

$$fc(D1) = \begin{cases} 2 & \text{if } d1 \text{ is sibling} \\ 4 & \text{otherwise} \end{cases}$$

$$fd(D1,D2, \dots Dj) = \begin{cases} 1 & \text{if } Dj \text{ is sibling} \\ 2 & \text{if any of } D1, D2 \dots Dj-1 \text{ is not sibling AND} \\ & \text{Dj is not sibling} \\ 3 & \text{if all of } D1, D2 \dots Dj-1 \text{ are sibling AND} \\ & \text{Dj is not sibling} \end{cases}$$

These two equations for X(1) and X(2) give the values of the approximate mean number of state segments that must be supported by P(1) and P(2) LSRs when a P2MP LSP is set up from a single source to N destination PEa. These equations are used to obtain exemplar numerical results in [Section 7](#).



## 7. Exemplar Numeric Results

### 7.1. Impact of Varying Topological Properties of the Network

We can substitute the probabilities that a given PE is a sibling, first cousin, or second cousin (given by the equations in [Section 1.4](#)) into the equations in [Section 6](#) so that the number of state segments that must be supported by P(1) and P(2) LSRs, X(1) and X(2), can be computed as functions of the topological properties of the snowflake topology, M1, M2, and S1. This allows the impact of varying these topology parameters on the scalability to be investigated.

A simple snowflake topology with M1, M2, and S1 all initially equal to 4 was considered. A P2MP LSP from every PE to 3 other randomly chosen PEs was established. Each one of M1, M2, and S1 was individually varied from 4 to 20 (with the other two parameters kept equal to 4) to allow the impact of varying each of these parameters individually to be shown in isolation. Table 5 shows the results.

M(1)	M(2)	S(1)	X(1)	X(2)
4	4	4	134.855	31.428125
4	4	6	143.326	31.62091667
4	4	8	147.527	31.7165625
4	4	10	150.038	31.7735
4	4	12	151.707	31.81145833
4	4	14	152.897	31.83857143
4	4	16	153.788	31.85875
4	4	18	154.481	31.87458333
4	4	20	155.034	31.887125
4	4	4	134.855	31.428125
4	6	4	201.344	47.05175
4	8	4	267.837	62.675625
4	10	4	334.332	78.3
4	12	4	400.829	93.924375
4	14	4	467.325	109.54875
4	16	4	533.822	125.173125
4	18	4	600.32	140.7975
4	20	4	666.817	156.421875
4	4	4	134.855	31.428125
6	4	4	202.862	31.62091667
8	4	4	270.866	31.7165625
10	4	4	338.868	31.7735
12	4	4	406.869	31.81145833
14	4	4	474.87	31.83857143
16	4	4	542.87	31.85875
18	4	4	610.871	31.87458333
20	4	4	678.871	31.887125

Table 5 : Impact of Varying Topological Properties of a Network

Komolafe, Farrel, and King

[Page 27]

Table 5 shows that, typically, a greater number of average state segments must be supported by P(1) LSRs than P2 LSRs. This finding is unsurprising given that the topology of the snowflake network means it would be expected that a greater number of LSPs are concentrated in the core.

It may be seen from Table 5 that increasing M(1) or M(2) leads to a linear increase in the number of state segments that must be handled at P(1) LSRs. This observation may be explained by realizing that increasing M(1) or M(2) (equivalently) increases the number of PEs subtended to each P(1) LSR and so more LSPs will be expected to be handled by this LSR. In contrast, increasing S1 increases the number of P1 LSRs but since the number of PEs subtended to each P(1) LSR remains unchanged (since M(1) and M(2) are unvaried in this case), the mean number of state segments remain relatively constant.

Varying either M(1) or S(1) does not alter the number of PEs subtended to each P(2) LSR and so the mean number of state segments that must be handled by P(2) LSRs remains relatively unchanged. In contrast, increasing M(2) directly increases the number of PEs attached to each P(2) LSR and thus increases the number of state segments that such LSRs must handle accordingly.

## **7.2. Impact of Varying the Number of PEs in the Receiving Set**

The impact of varying the size of receiving set, N, when the topological properties of the network are kept constant is considered in this section. A simple snowflake topology in which M(1), M(2), and S(1) are all equal to 12, and in which N is varied from 1 to 10 is considered. Table 6 records the number of state segments that must be handled at P(1) and P(2) LSRs when the P2MP LSP is set up.

For comparison, Table 6 also presents the results of establishing a corresponding set of P2P LSPs, allowing any benefits obtained by using P2MP LSPs instead of P2P to be quantified. The cost of establishing N P2P LSPs to the PEs in the receiving set was calculated by multiplying the cost of establishing a P2P LSP given in the equations for P2P LSPs in [Section 4.5](#) by  $S(PE) * N$  since there are SPE sources of N P2P LSPs.

Table 6 shows that, again, there is are more LSPs handled by P(1) LSRs than P(2) LSRs. Also, as would be expected, increasing the size of the receiving set increases the amount of state that must be supported by LSRs.





Number of PEs	P2P X(1)	P2P X(2)	P2MP X(1)	P2MP X(2)
1	550.318	47.84715278	550.318	47.84715278
2	1100.636	95.69430556	958.465	71.84652778
3	1650.954	143.5414583	1365.71	95.77083333
4	2201.272	191.3886111	1772.94	119.6944444
5	2751.59	239.2357639	2180.18	143.6180556
6	3301.908	287.0829167	2587.41	167.5416667
7	3852.226	334.9300695	2994.65	191.4652778
8	4402.544	382.7772222	3401.89	215.3881944
9	4952.862	430.624375	3809.12	239.3118056
10	5503.18	478.4715278	4216.36	263.2354167

Table 6 : Impact of Varying the Number of PEs in the Receiving Set When P2P or P2MP LSPs are Used

Using P2MP LSPs leads to LSRs having to support less state than if P2P LSPs are used. The relative reduction in the amount of state that must be supported by LSRs when P2MP LSPs instead of P2P LSPs are used is illustrated in Table 7. It can be seen that the relative reduction obtained by using P2MP LSPs initially rises with the size of the receiving set, before levelling off at around 24% and 45% for P(1) and P(2) LSRs respectively. It is perhaps surprising that using multicast does not achieve greater performance improvement; this finding may be explained by realizing that the approximations made in assuming the position of a subsequent PE in a receiving set is independent of the position of the preceding PEs means that the P2MP results presented here are close to the worst-case results. Basically, in a topology with  $M(1) = M(2) = S(1) = 12$ ,  $\text{Prob}(\text{second cousin}) = 92\%$  and so the worst-case P2MP LSPs in which all the destinations are second cousins (and so there are N replications at the first P1 LSR) is statistically the most likely scenario. Consequently, using multicast reduces the state predominantly only on the path from the PE to the first P(1) LSR, via the first P(2) LSR. On the other hand, had positions of preceding PEs in the receiving set been taken into consideration, the possibility of a given PE being a second cousin will diminish with the number of PEs in the receiving set which are already second cousins and so the probability of establishing the worst-case P2MP LSP will be lower.



Number of PEs	Percentage Reduction Using P2MP LSPs Compared to P2P LSPs	
	X(1)	X(2)
1	0	4.64442E-09
2	12.91716789	24.92079089
3	17.2775256	33.28001928
4	19.45838588	37.45999632
5	20.76653862	39.96798254
6	21.6389433	41.63997336
7	22.26182991	42.83425251
8	22.72899487	43.7301433
9	23.0925473	44.42678598
10	23.38320753	44.98410012

Table 7 : Reduction in the Number of State Segments at LSRs Obtained by Establishing P2MP LSPs Instead of P2P LSPs

### 8. Conclusions and Open Issues

A framework for investigating the scalability of P2MP MPLS-TE deployments by estimating the amount of state that must be handled by different types of LSRs in a snowflake network has been developed. The approach works by exhaustively aggregating the sum of the product of the cost and probabilities of all the different P2MP LSPs configurations.

Open issues include:

- Validating the mathematical model further.

Successful validation against the results in [[RFC5439](#)] using two different approaches is encouraging, but these are for P2P variants of the model. It will be useful to attempt to sanity-check the results for the P2MP equations produce, perhaps using simulation.

- Quantifying/removing approximations in mathematical model

The approximation that the position of a subsequent PE in a receiving set is independent of the position of the preceding PEs should be investigated further and, ideally, removed. Getting rid of this approximation will complicate the mathematics more but is potentially worthwhile as it will make the results more accurate and remove the requirement that  $N < S(1)$  ,  $M(1)$  , and  $M(2)$ . (Essentially, the goal should be to use the exact probabilities, for example, in Table 4). Doing so may be possible by exploiting patterns in the probabilities.



## **9. Management Considerations**

Management of networks with large numbers of MPLS-TE LSPs pose a number of challenging network management problems. Some of these issues are indentified and discussed within the relevant sections of this document. Additional discussion of network management considerations will be included in future versions of the document.

## **10. Security Considerations**

Security considerations for MPLS can be found in [[MPLS-SEC](#)].

Increasing the number of LSPs in a network to the point that the network is unmanageable or fails to operate constitutes as serious security risk.

## **11. Recommendations**

The authors would like to expand the analysis and validate the mathematical models used in this study. Next steps are outlined in [section 8](#) (Conclusions and Open Issues) of this document. This document does not make any specific recommendation at this time.

## **12. IANA Considerations**

This document makes no requests for IANA action.

## **13. Acknowledgements**

The authors are grateful to Thomas Morin for his comments and Julie Morrison for useful discussions regarding the math.

## **14. References**

### **14.1. Informative References**

- [RFC2961] Berger, L., Gan, D., Swallow, G., Pan, P., Tommasi, F., and S. Molendini, "RSVP Refresh Overhead Reduction Extensions", [RFC 2961](#), April 2001.
- [RFC3209] Awduche, D., Berger, L., Gan, D., Li, T., Srinivasan, V., and G. Swallow, "RSVP-TE: Extensions to RSVP for LSP Tunnels", [RFC 3209](#), December 2001.
- [RFC3473] Berger, L., et al., Editor, "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions",

[RFC 3473](#), January 2003.

[RFC4090] P. Pan, G. Swallow, A. Atlas (Editors), "Fast Reroute Extensions to RSVP-TE for LSP Tunnels", [RFC 4090](#), May 2005.

- [RFC4206] Kompella, K. and Y. Rekhter, "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE)", [RFC 4206](#), October 2005.
- [RFC4875] Aggarwal, R., Papadimitriou, D., and Yasukawa, S., "Extensions to Resource Reservation Protocol - Traffic Engineering (RSVP-TE) for Point-to-Multipoint TE Label Switched Paths (LSPs)", [RFC 4875](#), May 2007.
- [RFC5439] Yasukawa, S., Farrel, A., and Komolafe, O., "An Analysis of Scaling Issues in MPLS-TE Core Networks", [RFC 5439](#), February 2009.
- [MPLS-SEC] Fang, L. (Ed.), "Security Framework for MPLS and GMPLS Networks", [draft-ietf-mpls-mpls-and-gmpls-security-framework](#), work in progress.

## **15. Authors' Addresses**

Olufemi Komolafe  
Cisco Systems  
96 Commercial Street  
Edinburgh  
EH6 6LX  
United Kingdom  
EMail: femi@cisco.com

Adrian Farrel  
Old Dog Consulting  
EMail: adrian@olddog.co.uk

Daniel King  
Old Dog Consulting  
EMail: daniel@olddog.co.uk

## **16. Intellectual Property Consideration**

The IETF Trust takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in any IETF Document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights.

Copies of Intellectual Property disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or





the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement any standard or specification contained in an IETF Document. Please address the information to the IETF at [ietf-ipr@ietf.org](mailto:ietf-ipr@ietf.org).

The definitive version of an IETF Document is that published by, or under the auspices of, the IETF. Versions of IETF Documents that are published by third parties, including those that are translated into other languages, should not be considered to be definitive versions of IETF Documents. The definitive version of these Legal Provisions is that published by, or under the auspices of, the IETF. Versions of these Legal Provisions that are published by third parties, including those that are translated into other languages, should not be considered to be definitive versions of these Legal Provisions.

For the avoidance of doubt, each Contributor to the IETF Standards Process licenses each Contribution that he or she makes as part of the IETF Standards Process to the IETF Trust pursuant to the provisions of [RFC 5378](#). No language to the contrary, or terms, conditions or rights that differ from or are inconsistent with the rights and licenses granted under [RFC 5378](#), shall have any effect and shall be null and void, whether published or posted by such Contributor, or included with or in such Contribution.

## **17. Disclaimer of Validity**

All IETF Documents and the information contained therein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY, THE IETF TRUST AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION THEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

## **18. Full Copyright Statement**

Copyright (c) 2009 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to [BCP 78](#) and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>)



in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document.