

TEWG  
Internet Draft

Francois Le Faucheur  
Editor  
Cisco Systems,  
Inc.  
December 2004

[draft-ietf-tewg-diff-te-russian-07.txt](#)

Expires: June 2005

Russian Dolls Bandwidth Constraints Model for  
Diff-Serv-aware MPLS Traffic Engineering

Status of this Memo

This document is an Internet-Draft and is subject to all provisions of [section 3 of RFC 3667](#). By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she become aware will be disclosed, in accordance with [RFC 3668](#).

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/lid-abstracts.txt>.

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

This Internet-Draft will expire on June 13, 2005.

Copyright Notice

Copyright (C) The Internet Society (2004). All Rights Reserved.

Abstract

This document provides specification for one Bandwidth Constraints Model for Diff-Serv-aware MPLS Traffic Engineering, which is referred to as the Russian Dolls Model.

## Specification of Requirements

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [[RFC2119](#)].

## Table of Contents

<a href="#">1.</a>	Introduction.....	<a href="#">2</a>
<a href="#">2.</a>	Contributing Authors.....	<a href="#">3</a>
<a href="#">3.</a>	Definitions.....	<a href="#">3</a>
<a href="#">4.</a>	Russian Dolls Model Definition.....	<a href="#">4</a>
<a href="#">5.</a>	Example Formulas for Computing "Unreserved TE-Class [i]" with Russian Dolls Model.....	<a href="#">6</a>
<a href="#">6.</a>	Receiving both Maximum Reservable Bandwidth and Bandwidth Constraints sub-TLVs.....	<a href="#">7</a>
<a href="#">7.</a>	Security Considerations.....	<a href="#">8</a>
<a href="#">8.</a>	Acknowledgments.....	<a href="#">8</a>
<a href="#">9.</a>	IANA Considerations.....	<a href="#">8</a>
<a href="#">10.</a>	Normative References.....	<a href="#">8</a>
<a href="#">11.</a>	Informative References.....	<a href="#">9</a>
<a href="#">12.</a>	Intellectual Property Considerations.....	<a href="#">10</a>
<a href="#">13.</a>	Editor's Address:.....	<a href="#">10</a>
<a href="#">14.</a>	Full Copyright Statement.....	<a href="#">10</a>
<a href="#">Appendix A</a>	- Addressing [ <a href="#">DSTE-REQ</a> ] Scenarios.....	<a href="#">11</a>
	Disclaimer of Validity.....	<a href="#">12</a>
	Copyright Statement.....	<a href="#">13</a>
	Acknowledgment.....	<a href="#">13</a>

## 1.Introduction

[DSTE-REQ] presents the Service Providers requirements for support of Diff-Serv-aware MPLS Traffic Engineering (DS-TE). This includes the fundamental requirement to be able to enforce different Bandwidth Constraints for different classes of traffic.

[DSTE-REQ] also defines the concept of Bandwidth Constraints Model

for DS-TE and states that "The DS-TE technical solution MUST specify at least one Bandwidth Constraints Model and MAY specify multiple Bandwidth Constraints Models."

This document provides a detailed description of one particular Bandwidth Constraints Model for DS-TE which is introduced in [DSTE-REQ] and called the Russian Dolls Model (RDM).

[DSTE-PROTO] specifies the IGP and RSVP-TE signaling extensions for support of DS-TE. These extensions support RDM.

## 2. Contributing Authors

This document was the collective work of several. The text and content of this document was contributed by the editor and the co-authors listed below. (The contact information for the editor appears in [Section 11](#), and is not repeated below.)

Jim Boyle  
Protocol Driven Networks, Inc.  
1381 Kildaire Farm Road #288  
Cary, NC 27511, USA  
Phone: (919) 852-5160  
Email: jboyle@pdnets.com

Kireeti Kompella  
Juniper Networks, Inc.  
1194 N. Mathilda Ave.  
Sunnyvale, CA 94099  
Email: kireeti@juniper.net

William Townsend  
Tenor Networks  
100 Nagog Park  
Acton, MA 01720  
Phone: +1-978-264-4900  
Email:  
btownsend@tenornetworks.com

Thomas D. Nadeau  
Cisco Systems, Inc.  
250 Apollo Drive  
Chelmsford, MA 01824  
Phone: +1-978-244-3051  
Email: tnadeau@cisco.com

Darek Skalecki  
Nortel Networks  
3500 Carling Ave,  
Nepean K2H 8E9  
Phone: +1-613-765-2252  
Email: dareks@nortelnetworks.com

### 3. Definitions

For readability a number of definitions from [[DSTE-REQ](#)] are repeated here:

**Class-Type (CT):** the set of Traffic Trunks crossing a link that is governed by a specific set of Bandwidth Constraints. CT is used for the purposes of link bandwidth allocation, constraint based routing and admission control. A given Traffic Trunk belongs to the same CT on all links.

**TE-Class:** A pair of:

- i. a Class-Type
- ii. a preemption priority allowed for that Class-Type. This means that an LSP transporting a Traffic Trunk from that Class-Type can use that preemption priority as the set-up priority, as the holding priority or both.

A number of recovery mechanisms under investigation or specification in the IETF take advantage of the concept of bandwidth sharing across particular sets of LSPs. "Shared Mesh Restoration" in [[GMPLS-RECOV](#)] and "Facility-based Computation Model" in [[MPLS-BACKUP](#)] are example mechanisms which increase bandwidth efficiency by sharing bandwidth across backup LSPs protecting against independent failures. To ensure that the notion of "Reserved (CTc)" introduced in [[DSTE-REQ](#)] is compatible with such a concept of bandwidth sharing across multiple LSPs, the wording of the "Reserved (CTc)" definition provided in [[DSTE-REQ](#)] is generalized into the following:

**Reserved (CTc):** For a given Class-Type CTc ( $0 \leq c \leq \text{MaxCT}$ ), let us define "Reserved(CTc)" as the total amount of the bandwidth reserved by all the established LSPs which belong to CTc.

With this generalization, the Russian Dolls Model definition provided in this document is compatible with Shared Mesh Restoration defined in [[GMPLS-RECOV](#)], so that DS-TE and Shared Mesh Protection can operate simultaneously, under the assumption that Shared Mesh Restoration operates independently within each DS-TE Class-Type and does not operate across Class-Types (for example back up

LSPs protecting Primary LSPs of CTx need to also belong to CTx; Excess Traffic LSPs sharing bandwidth with Backup LSPs of CTx need to also belong to CTx).

We also introduce the following definition:

Reserved(CTb,q) : let us define "Reserved(CTb,q)" as the total amount of the bandwidth reserved by all the established LSPs which belong to CTb and have a holding priority of q. Note that if q and CTb do not form one of the 8 possible configured TE-Classes, then there can not be any established LSP which belong to CTb and have a holding priority of q, so in that case, Reserved(CTb,q)=0.

#### 4. Russian Dolls Model Definition

RDM is defined in the following manner:

- o Maximum Number of Bandwidth Constraints (MaxBC)=  
Maximum Number of Class-Types (MaxCT) = 8
- o for each value of b in the range  $0 \leq b \leq (\text{MaxCT} - 1)$ :

Le Faucheur, et al.

[Page 4]

---

Russian Dolls Model for DS-TE

December 2005

- SUM (Reserved (CTc))  $\leq$  BCb,  
Where the SUM is across all values of c in the  
range  $b \leq c \leq (\text{MaxCT} - 1)$
- o BC0= Maximum Reservable Bandwidth, so that  
SUM (Reserved(CTc))  $\leq$  Max-Reservable-Bw,  
where the SUM is across all values of c in the  
range  $0 \leq c \leq (\text{MaxCT} - 1)$

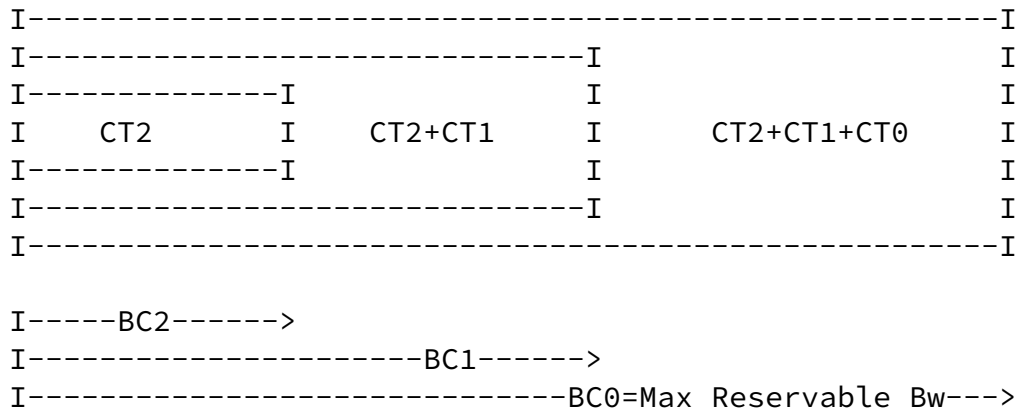
A DS-TE LSR implementing RDM MUST support enforcement of Bandwidth Constraints in compliance with this definition.

Both preemption within a Class-Type and across Class-Types is allowed.

Where 8 Class-Types are active, the RDM Bandwidth Constraints can also be expressed in the following way:

- All LSPs from CT7 use no more than BC7
- All LSPs from CT6 and CT7 use no more than BC6
- All LSPs from CT5, CT6 and CT7 use no more than BC5
- etc.
- All LSPs from CT0, CT1,... CT7 use no more than  
BC0 = "Maximum Reservable Bandwidth"

Purely for illustration purposes, the diagram below represents the Russian Dolls Bandwidth Constraints Model in a pictorial manner when 3 Class-Types are active:



While simpler Bandwidth Constraints models or, conversely, more flexible/sophisticated Bandwidth Constraints models can be defined, the Russian Dolls Model is attractive in some DS-TE environments for the following reasons:

- Although a little less intuitive than the Maximum Allocation Model (see[DSTE-MAM]), RDM is still a simple model to conceptualize.

- RDM can be used to simultaneously ensure bandwidth efficiency and protection against QoS degradation of all Class-Types, whether preemption is used or not.
- RDM can be used in conjunction with preemption to simultaneously achieve isolation across Class-Types (so that each Class-Type is guaranteed its share of bandwidth no matter the level of contention by other classes), bandwidth efficiency and protection against QoS degradation of all Class-Types.
- RDM only requires limited protocol extensions such as the ones defined in [\[DSTE-PROTO\]](#).

RDM may not be attractive in some DS-TE environments for the following reasons:

- if the usage of preemption is precluded for some

administrative reason, while RDM can still ensure bandwidth efficiency and protection against QoS degradation of all CTs, RDM cannot guarantee isolation across Class-Types.

Additional considerations on the properties of RDM can be found in [\[BC-CONS\]](#) and [\[BC-MODEL\]](#).

As a simple example usage of the "Russian Dolls" Bandwidth Constraints Model, a network administrator using one CT for Voice (CT1) and one CT for data (CT0) might configure on a given link:

- BC0 = Max-Reservable-Bw= 2.5 Gb/s (i.e. Voice + Data is limited to 2.5 Gb/s)
- BC1= 1.5 Gb/s (i.e. Voice is limited to 1.5 Gb/s).

#### 5.Example Formulas for Computing "Unreserved TE-Class [i]" with Russian Dolls Model

As specified in [\[DSTE-PROTO\]](#), formulas for computing "Unreserved TE-Class [i]" MUST reflect all of the Bandwidth Constraints relevant to the CT associated with TE-Class[i], and thus, depend on the Bandwidth Constraints Model. Thus, a DS-TE LSR implementing RDM MUST reflect the RDM Bandwidth Constraints defined in [section 4](#) above when computing "Unreserved TE-Class [i]".

Keeping in mind, as explained in [\[DSTE-PROTO\]](#), that details of admission control algorithms as well as formulas for computing "Unreserved TE-Class [i]" are outside the scope of the IETF work, we provide in this section, for illustration purposes, an example of how values for the unreserved bandwidth for TE-Class[i] might be computed with RDM, assuming the basic admission control algorithm which simply deducts the exact bandwidth of any established LSP from all of the Bandwidth Constraints relevant to the CT associated with that LSP.

We assume that:

TE-Class [i] <--> < CTc , preemption p >  
in the configured TE-Class mapping.

For readability, formulas are first shown assuming only 3 CTs are active. The formulas are then extended to cover the cases where more CTs are used.

If  $CT_c = CT_0$ , then "Unreserved TE-Class [i]" =  
 $[ BC_0 - \text{SUM} ( \text{Reserved}(CT_b, q) ) ]$  for  $q \leq p$  and  $0 \leq b \leq 2$

If  $CT_c = CT_1$ , then "Unreserved TE-Class [i]" =  
 $\text{MIN} [$   
 $[ BC_1 - \text{SUM} ( \text{Reserved}(CT_b, q) ) ]$  for  $q \leq p$  and  $1 \leq b \leq 2$ ,  
 $[ BC_0 - \text{SUM} ( \text{Reserved}(CT_b, q) ) ]$  for  $q \leq p$  and  $0 \leq b \leq 2$   
 $]$

If  $CT_c = CT_2$ , then "Unreserved TE-Class [i]" =  
 $\text{MIN} [$   
 $[ BC_2 - \text{SUM} ( \text{Reserved}(CT_b, q) ) ]$  for  $q \leq p$  and  $2 \leq b \leq 2$ ,  
 $[ BC_1 - \text{SUM} ( \text{Reserved}(CT_b, q) ) ]$  for  $q \leq p$  and  $1 \leq b \leq 2$ ,  
 $[ BC_0 - \text{SUM} ( \text{Reserved}(CT_b, q) ) ]$  for  $q \leq p$  and  $0 \leq b \leq 2$   
 $]$

The formula can be generalized to 8 active CTs and expressed in a more compact way in the following:

"Unreserved TE-Class [i]" =  
 $\text{MIN} [$   
 $[ BC_c - \text{SUM} ( \text{Reserved}(CT_b, q) ) ]$  for  $q \leq p$  and  $c \leq b \leq 7$ ,  
 $[ BC_{(c-1)} - \text{SUM} ( \text{Reserved}(CT_b, q) ) ]$  for  $q \leq p$  and  $(c-1) \leq b \leq 7$ ,  
 $\cdot \cdot \cdot$   
 $[ BC_0 - \text{SUM} ( \text{Reserved}(CT_b, q) ) ]$  for  $q \leq p$  and  $0 \leq b \leq 7$ ,  
 $]$

where:

TE-Class [i]  $\leftrightarrow$   $\langle CT_c, \text{preemption } p \rangle$   
in the configured TE-Class mapping.

## 6. Receiving both Maximum Reservable Bandwidth and Bandwidth Constraints sub-TLVs

[DSTE-PROTO] states that

" A DS-TE LSR which does advertise Bandwidth Constraints MUST use the new "Bandwidth Constraints" sub-TLV (in addition to the existing Maximum Reservable Bandwidth sub-TLV) to do so."



With RDM, BC0 is equal to the Maximum Reservable Bandwidth since they both represent the aggregate constraint across all Class-Types. Thus, a DS-TE LSR receiving both the "Maximum Reservable Bw" sub-TLV and the new "Bandwidth Constraints" sub-TLV (which contains BC0) for a given link where the RDM model is used, MAY ignore the "Maximum Reservable Bw" sub-TLV.

## 7. Security Considerations

Security considerations related to the use of DS-TE are discussed in [\[DSTE-PROTO\]](#). Those apply independently of the Bandwidth Constraints Model, including RDM specified in this document.

## 8. Acknowledgments

We thank Martin Tatham for his key contribution in this work. Tatiana Renko is also warmly thanked for her instantiation of the Russian Doll.

## 9. IANA Considerations

[\[DSTE-PROTO\]](#) defines a new name space for "Bandwidth Constraints Model Id". The guidelines for allocation of values in that name space are detailed in section 14.1 of [\[DSTE-PROTO\]](#). In accordance with these guidelines, IANA was requested to assign a Bandwidth Constraints Model Id for RDM from the range 0-127 (which is to be managed as per the "Specification Required" policy defined in [\[IANA-CONS\]](#)).

Bandwidth Constraints Model Id = TBD was allocated by IANA to RDM.

<IANA-note> To be removed by the RFC editor at the time of publication

We request IANA to assign value 0 for the RDM model.

Once the value has been assigned, please replace "TBD" above by the assigned value.

</IANA-note>

## 10. Normative References

[\[DSTE-REQ\]](#) Le Faucheur et al, Requirements for support of Diff-Serv-aware MPLS Traffic Engineering, [RFC3564](#).

[DSTE-PROTO] Le Faucheur et al, Protocol extensions for support of Diff-Serv-aware MPLS Traffic Engineering, [draft-ietf-tewg-diff-te-proto-08.txt](#), work in progress.

[RFC2119] S. Bradner, Key words for use in RFCs to Indicate Requirement Levels, [RFC2119](#), March 1997.

[IANA-CONS], T. Narten et al, "Guidelines for Writing an IANA Considerations Section in RFCs", [RFC2434](#).

## 11. Informative References

[BC-CONS] Le Faucheur, "Considerations on Bandwidth Constraints Model for DS-TE", [draft-lefaucheur-tewg-russian-dolls-00.txt](#), June 2002.

[BC-MODEL] Lai, "Bandwidth Constraints Models for DS-TE", [draft-wlai-tewg-bcmodel-03.txt](#), work in progress.

[DSTE-MAM] Le Faucheur, Lai, "Maximum Allocation Bandwidth Constraints Model for Diff-Serv-aware MPLS Traffic Engineering", [draft-ietf-tewg-diff-tet-mam-04.txt](#), work in progress.

[DSTE-MAR] Ash, "Max Allocation with Reservation Bandwidth Constraints Model for MPLS/DiffServ TE & Performance Comparisons", work in progress.

[OSPF-TE] Katz et al., "Traffic Engineering (TE) Extensions to OSPF Version 2", [RFC3630](#).

[ISIS-TE] Smit, Li, "Intermediate System to Intermediate System (ISIS) extensions for Traffic Engineering (TE)", [RFC 3784](#).

[RSVP-TE] Awduche et al, "RSVP-TE: Extensions to RSVP for LSP Tunnels", [RFC 3209](#).

[DIFF-MPLS] Le Faucheur et al, "MPLS Support of Diff-Serv", [RFC3270](#).

[GMPLS-RECOV] Lang et al, "Generalized MPLS Recovery Functional Specification", [draft-ietf-ccamp-gmpls-recovery-functional-02.txt](#), work in progress.

[MPLS-BACKUP] Vasseur et al, "MPLS Traffic Engineering Fast reroute: bypass tunnel path computation for bandwidth protection", [draft-vasseur-mpls-backup-computation-02.txt](#), work in progress.

## 12. Intellectual Property Considerations

The IETF 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 this 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. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR 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 this standard. Please address the information to the IETF at [ietf-ipr@ietf.org](mailto:ietf-ipr@ietf.org).

## 13. Editor's Address:

Francois Le Faucheur  
Cisco Systems, Inc.  
Village d'Entreprise Green Side - Batiment T3  
400, Avenue de Roumanille  
06410 Biot-Sophia Antipolis  
France  
Phone: +33 4 97 23 26 19  
Email: [flefauch@cisco.com](mailto:flefauch@cisco.com)

## 14. Full Copyright Statement

Copyright (C) The Internet Society (2004). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing

Le Faucheur, et al.

[Page 10]

---

Russian Dolls Model for DS-TE

December 2005

the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

## Appendix A - Addressing [\[DSTE-REQ\]](#) Scenarios

This Appendix provides examples of how the Russian Dolls Bandwidth Constraints Model can be used to support each of the scenarios described in [\[DSTE-REQ\]](#).

### 1. Scenario 1: Limiting Amount of Voice

By configuring on every link:

- Bandwidth Constraint 1 (for CT1=Voice) = "certain percentage" of link capacity
- BC0 (for CT1=Voice + CT0= Data) = link capacity

By configuring:

- every CT1/Voice TE-LSP with preemption =0
- every CT0/Data TE-LSP with preemption =1

DS-TE with the Russian Dolls Model will address all the requirements:

- amount of Voice traffic limited to desired percentage on every link
- data traffic capable of using all remaining link capacity
- voice traffic capable of preempting other traffic

## 2. Scenario 2: Maintain Relative Proportion of Traffic Classes

By configuring on every link:

- BC2 (for CT2) = e.g. 45%
- BC1 (for CT1+CT2) = e.g. 80%
- BC0 (for CT0+CT1+CT2) = e.g.100%

DS-TE with the Russian Dolls Model will ensure that the amount of traffic of each Class Type established on a link is within acceptable levels as compared to the resources allocated to the corresponding Diff-Serv PHBs regardless of which order the LSPs are routed in, regardless of which preemption priorities are used by which LSPs and regardless of failure situations.

By also configuring:

- every CT2/Voice TE-LSP with preemption =0
- every CT1/Premium Data TE-LSP with preemption =1
- every CT0/Best-Effort TE-LSP with preemption =2

DS-TE with the Russian Dolls Model will also ensure that:

- CT2 Voice LSPs always have first preemption priority in order to use the CT2 capacity
- CT1 Premium Data LSPs always have second preemption priority in order to use the CT1 capacity
- Best-Effort can use up to link capacity whatever is left by CT2 and CT1.

Optional automatic adjustment of Diff-Serv scheduling configuration could be used for maintaining very strict relationship between amount of established traffic of each Class Type and corresponding Diff-Serv resources.

### 3. Scenario 3: Guaranteed Bandwidth Services

By configuring on every link:

- BC1 (for CT1) = "given" percentage of link bandwidth (appropriate to achieve the Guaranteed Bandwidth service's QoS objectives)
- BC0 (for CT0+CT1) = 100% of link bandwidth

DS-TE with the Russian Dolls Model will ensure that the amount of Guaranteed Bandwidth Traffic established on every link remains below the given percentage so that it will always meet its QoS objectives. At the same time it will allow traffic engineering of the rest of the traffic such that links can be filled up.

#### Disclaimer of Validity

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY 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 HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Le Faucheur, et al.

[Page 12]

#### Copyright Statement

Copyright (C) The Internet Society (2004). This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

#### Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.

