

Network Working Group  
Internet Draft  
Expiration Date: August 2001

K. Kompella (Juniper Networks)  
Y. Rekhter (Juniper Networks)  
A. Banerjee (Calient Networks)  
J. Drake (Calient Networks)  
G. Bernstein (Ciena)  
D. Fedyk (Nortel Networks)  
E. Mannie (GTS Network)  
D. Saha (Tellium)  
V. Sharma (Tellabs)

## OSPF Extensions in Support of Generalized MPLS

[draft-kompella-ospf-gmpls-extensions-01.txt](#)

### 1. Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of [Section 10 of RFC2026](#).

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>.

---

Internet Draft [draft-kompella-ospf-gmpls-extensions-01.txt](#) February 2001

## [2.](#) Abstract

This document specifies extensions to the OSPF routing protocol in support of Generalized Multi-Protocol Label Switching (previously known as Multi-Protocol Lambda Switching).

## [3.](#) Introduction

This document specifies extensions to the OSPF routing protocol in support of carrying link state information for Generalized Multi-Protocol Label Switching (GMPLS, previously known as Multi-Protocol Lambda Switching, MPL(ambda)S). For motivations and overall architecture of MPL(ambda)S see [\[1\]](#). The set of required enhancements to OSPF are outlined in [\[2\]](#). This document enhances the routing extensions [\[3\]](#) required to support MPLS Traffic Engineering. Some of these enhancements also need to be carried in the signaling protocols [\[6\]](#).

The organization of the remainder of the document is as follows. In [Section 4](#), we describe the types of links that may be advertised in OSPF TE LSAs. In [Section 5](#), we define a new set of Type/Length/Value (TLV) triplets, and describe their formats.

## [4.](#) GMPLS TE Links

Traditionally, a TE link is advertised as an adjunct to a "regular" OSPF link, i.e., an OSPF adjacency is brought up on the link, and when the link is up, both the regular SPF properties of the link (basically, the SPF metric) and the TE properties of the link are then advertised.

However, GMPLS challenges this notion in three ways. First, links that are not capable of sending and receiving on a packet-by-packet basis may yet have TE properties; however, an OSPF adjacency cannot be brought up on such links. Second, a Label Switched Path can be advertised as a point-to-point TE link (see [LSP-HIER]); thus, an advertised TE link need no longer be between two OSPF neighbors. Finally, a number of links may be advertised as a single TE link (perhaps for improved scalability), so again, there is no longer a one-to-one association of a regular adjacency and a TE link.

Thus we have a more general notion of a TE link. A TE link is a "logical" link that has TE properties, some of which may be configured on the advertising Label Switching Router (LSR), others which may be obtained from other LSRs by means of some protocol, and yet others which may be deduced from the component(s) of the TE link.

Internet Draft [draft-kompella-ospf-gmpls-extensions-01.txt](#) February 2001

A TE link between a pair of LSRs doesn't imply the existence of an OSPF adjacency between these LSRs. A TE link must also have some means by which the advertising LSR can know of its liveness (this means may be OSPF hellos, but is not limited to OSPF hellos). When an LSR knows that a TE link is up, and can determine the TE link's TE properties, the LSR may then advertise that link to its (regular) OSPF neighbors using the TE TLVs. In this document, we call the interfaces over which regular OSPF adjacencies are established "control channels".

[3] defines the canonical TE properties, and says how to associate TE properties to regular (packet-switched) links. This document extends the set of TE properties, and also says how to associate TE properties with non-packet-switched links such as links between Optical Cross-Connects (OXC's). [5] says how to associate TE properties with Label Switched Paths; [4] says how to associate TE properties with a "bundle" of links.

#### [4.1](#). Excluding data traffic from control channels

The control channels between nodes in a GMPLS network, such as OXC's (see [1], [2]), SONET cross-connects and/or routers, are generally meant for control and administrative traffic. These control channels are advertised into OSPF as normal IS links as mentioned in the previous section; this allows the routing of (for example) RSVP messages and telnet sessions. However, if routers on the edge of the optical domain attempt to forward data traffic over these channels, the channel capacity will quickly be exhausted.

If one assumes that data traffic is sent to BGP destinations, and control traffic to IGP destinations, then one can exclude data traffic from the control plane by restricting BGP nexthop resolution. (It is assumed that OXC's are not BGP speakers.) Suppose that a router R is attempting to install a route to a BGP destination D. R looks up the BGP nexthop for D in its IGP's routing table. Say R

finds that the path to the nexthop is over interface I. R then checks if it has an entry in its Link State database associated with the interface I. If it does, and the link is not packet-switch capable (see [5]), R installs a discard route for destination D. Otherwise, R installs (as usual) a route for destination D with nexthop I. Note that R need only do this check if it has packet-switch incapable links; if all of its links are packet-switch capable, then clearly this check is redundant.

Other techniques for excluding data traffic from control channels may also be needed.

## 5. OSPF Routing Enhancements

In this section we define the enhancements to the TE properties of GMPLS TE links that can be announced in OSPF TE LSAs. The Traffic Engineering (TE) LSA, which is an opaque LSA with area flooding scope [3], has only one top-level Type/Length/Value (TLV) triplet and has one or more nested TLVs for extensibility. The top-level TLV can take one of two values (1) Router Address or (2) Link. In this document, we enhance the sub-TLVs for the Link TLV in support of GMPLS. Specifically, we add the following sub-TLVs:

1. Outgoing Interface Identifier,
2. Incoming Interface Identifier,
3. Link Protection Type,
4. Link Descriptor, and
5. Shared Risk Link Group.

This brings the list of sub-TLVs of the TE Link TLV to:

| Sub-TLV | Type | Length | Name                         |
|---------|------|--------|------------------------------|
| 1       |      | 1      | Link type                    |
| 2       |      | 4      | Link ID                      |
| 3       |      | 4      | Local interface IP address   |
| 4       |      | 4      | Remote interface IP address  |
| 5       |      | 4      | Traffic engineering metric   |
| 6       |      | 4      | Maximum bandwidth            |
| 7       |      | 4      | Maximum reservable bandwidth |
| 8       |      | 32     | Unreserved bandwidth         |
| 9       |      | 4      | Resource class/color         |

|             |          |  |
|-------------|----------|--|
| 10          | 4        | Link Mux Capability                    |
| 11          | 4        | Outgoing Interface Identifier          |
| 12          | 4        | Incoming Interface Identifier          |
| 13          | 32       | Maximum LSP Bandwidth                  |
| 14          | 4        | Link Protection Type                   |
| 15          | variable | Link Descriptor                        |
| 16          | variable | Shared Risk Link Group                 |
| 32768-32772 | -        | Reserved for Cisco-specific extensions |

### [5.1. Outgoing Interface Identifier](#)

A link from LSR A to LSR B may be assigned an "outgoing interface identifier". This identifier is a non-zero 32-bit number that is assigned by LSR A. This identifier must be unique within the scope of A. If such an identifier has been assigned, A can advertise it as a sub-TLV of the Link TLV with type 11, length 4 and value equal to the assigned identifier.

### [5.2. Incoming Interface Identifier](#)

Suppose there is a link L from A to B. Suppose further that the link L' from B to A that corresponds to the same interface as L has been assigned an outgoing interface identifier by B. The "incoming interface identifier" for L (from A's point of view) is defined as the outgoing interface identifier for L' (from B's point of view). If A knows this value (by means beyond the scope of this document), A can advertise it as a sub-TLV of the Link TLV with type 12, length 4 and value equal to L's incoming interface identifier.

### [5.3. Maximum LSP Bandwidth sub-TLV](#)

The Maximum LSP Bandwidth takes the place of the Maximum Link Bandwidth. However, while Maximum Link Bandwidth is a single fixed value (usually simply the link capacity), Maximum LSP Bandwidth is carried per priority, and may vary as LSPs are set up and torn down.

The Maximum LSP Bandwidth of a bundled link at priority p is defined to be the maximum of the Maximum LSP Bandwidth at priority p of each component link.

If a component link is a simple (unbundled) link, define its Maximum LSP Bandwidth at priority  $p$  to be the smaller of its unreserved bandwidth at priority  $p$  and its maximum link bandwidth.

The Maximum LSP Bandwidth TLV has type 13 and length 32 octets. The value is a list of eight 4 octet fields in IEEE floating point format of the Maximum LSP Bandwidth of the link, with priority 0 first and priority 7 last.

Although Maximum Link Bandwidth is to be deprecated, for backward compatibility, one MAY set the Maximum Link Bandwidth to the Maximum LSP Bandwidth at priority 7 of the link.

#### [5.4. Link Protection Type sub-TLV](#)

The Link Protection Type sub-TLV represents the protection capability that exists on a link. It is desirable to carry this information so that it may be used by the path computation algorithm to set up LSPs with appropriate protection characteristics.

In the Traffic Engineering LSA, the Link Protection Type sub-TLV is a sub-TLV of the Link TLV, with type 14, and length of four octets, the first of which is a bit vector describing the protection capabilities of the link. They are:

##### 0x01 Extra Traffic

If the link is Extra Traffic, it indicates that these links are protecting other (primary) links. The LSPs on a link of this type will be lost if the primary links being protected fail.

##### 0x02 Unprotected

If the link is Unprotected, it means that there is no backup link for traffic being carried on the link.

##### 0x04 Shared

If the link has Shared protection, it means that for the  $N > 1$  primary data-bearing channels, there are  $M$  disjoint backup data-bearing channels reserved to carry the traffic. Additionally, the protection data-bearing channel MAY carry low-priority preemptable traffic. The bandwidth of backup data-bearing channels will be

announced in the unreserved bandwidth sub-TLV at the appropriate priority.

#### 0x08 Dedicated 1:1

If the link has Dedicated 1:1 protection, it means that for each primary data-bearing channel, there is one disjoint backup data-bearing channel reserved to carry the traffic. Additionally, the protection data-bearing channel MAY carry low-priority preemptable traffic. The bandwidth of backup data-bearing channels will be announced in the unreserved bandwidth sub-TLV at the appropriate priority.

#### 0x10 Dedicated 1+1

If the link has Dedicated 1+1 protection, it means that a disjoint backup data-bearing channel is reserved and dedicated for protecting the primary data-bearing channel. This backup data-bearing channel is not shared by any other connection, and traffic is duplicated and carried simultaneously over both channels.

#### 0x20 Enhanced

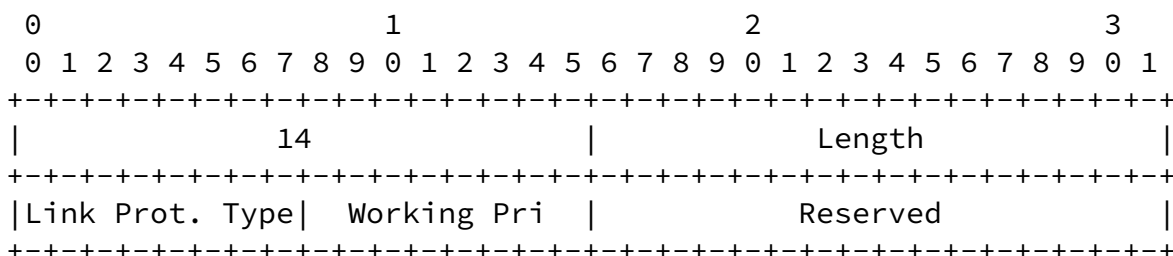
If the link has Enhanced protection, it indicates that a protection scheme that is more reliable than Dedicated 1+1 is being used, e.g., 4 fiber BLSR/MS-SPRING to protect this link.

#### 0x40 Reserved

#### 0x80 Reserved

The second octet gives a priority value such that a new connection with a higher priority (i.e., numerically lower than this value) is guaranteed to be setup on a primary (or working) channel, and not on a secondary (or protect) channel.

The format of the Link Protection Type sub-TLV is as shown below:



The Link Protection Type sub-TLV is optional and if an LSA does not carry the TLV, then the Link Protection Type is unknown. The protection capability of a link is typically pre-configured and does not change dynamically over time. The working priority value is pre-configured and does not depend on the traffic characteristics on the primary data-bearing channel.

#### [5.5. Link Descriptor sub-TLV](#)

The Link Descriptor TLV represents the characteristics of the link, in particular the link type and the bit transmission rate. These characteristics represent one of the constraints on the type of LSPs that could be carried over the link.

These characteristics should not be confused with the physical link encoding or multiplex structure (if any). For example there are systems where four OC-48s are switched and transported over a single fiber via wavelength division multiplexing (WDM) technology. There are systems where four OC-48s are transported in a transparent OC-192 time division multiplex (TDM) structure, i.e., all the overheads of the OC-48's are persevered. In both these cases the essential information from a routing perspective is that both of the links can transport media of type OC-48.

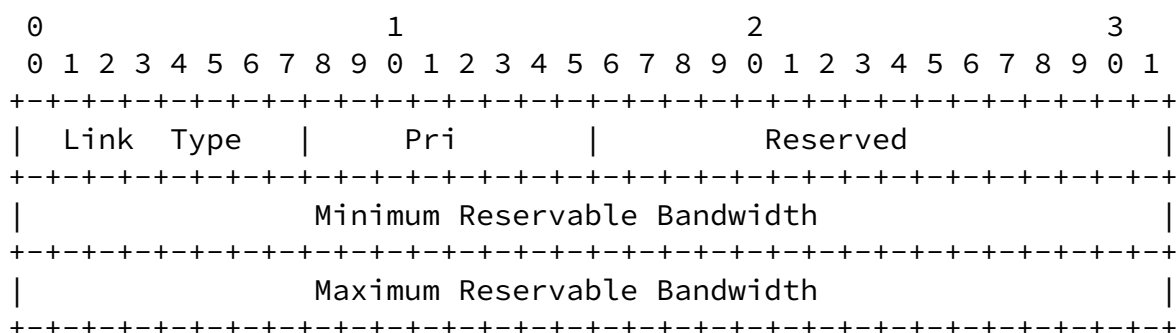
In the Traffic Engineering LSA, the Link Descriptor sub-TLV is a sub-TLV of the Link TLV, with type 15. The length is the length of the list of Link Descriptors in octets. Each Link descriptor element consists of the following fields: the first field is a one-octet value which defines the link encoding type, the second field is a one-octet value which defines the lowest priority at which that link encoding type is available, the next two-octets are reserved, the next field is four-octets and specifies the minimum reservable bandwidth (in IEEE floating point format, the unit being bytes per second) for this link encoding type, and the last four-octets specifies the maximum reservable bandwidth (in IEEE floating point format, the unit being bytes per second) for this link encoding type. Link encoding type values are taken from the following list:

| Value | Link Encoding Type |
|-------|--------------------|
|-------|--------------------|



- |   |                 |
|---|-----------------|
| 1 | Standard SONET  |
| 2 | Arbitrary SONET |
| 3 | Standard SDH    |
| 4 | Arbitrary SDH   |
| 5 | Clear           |
| 6 | GigE            |
| 7 | 10GigE          |

The format of the Link Descriptors is shown in the next figure.



A link having Standard SONET (or Standard SDH) link encoding type can switch data at a minimum rate, which is given by the Minimum Reservable Bandwidth on that link, and the maximum rate is given by the Maximum Reservable Bandwidth on that link. Typical values of these are enumerated in the GMPLS signaling draft [6]. In other words, the Minimum and Maximum Reservable Bandwidth represents the leaf and the root of one branch within the structure of the SONET (or SDH) hierarchy. An LSP on that link may reserve any bit transmission rate that is allowed by the SONET (or SDH) hierarchy between the minimum and maximum reservable values (the spectrum is discrete). For example, consider a branch of SONET multiplexing tree : VT-1.5, STS-1, STS-3c, STS-12c, STS-48c, STS-192c. If it is possible to establish all these connections on a OC-192 link, it can be advertised as follows: Minimum Reservable Bandwidth VT-1.5 and Maximum Reservable Bandwidth STS-192.

A link having Arbitrary SONET (or Arbitrary SDH) link encoding type can switch data at a minimum rate, which is given by the Minimum Reservable Bandwidth on that link, and the maximum rate is given by the Maximum Reservable Bandwidth on that link. Typical values of these are enumerated in the GMPLS signaling draft [6]. An LSP on that link may reserve any bit transmission rate that is a multiple of the Minimum Reservable Bandwidth between the minimum and maximum reservable values (the spectrum is discrete).

To handle the case where a link could support multiple branches of the SONET (or SDH) multiplexing hierarchy, one could use multiple LSP descriptors. For example, if a link supports VT-1.5 and VT-2 (which are not part of same branch of SONET multiplexing tree), then it could advertise two LSP descriptors, one for each one.

For a link with Clear encoding, the minimum and maximum reservable bandwidth will imply that the optical path is tuned to carry traffic within those range of values. (Note that it should be possible to carry OC-x as well as GigE or any other encoding format, as long as the bit transmission rate of the data to be carried is within this range.)

For other encoding types, the minimum and maximum reservable bandwidth should be set to have the same values.

Link Descriptors present a new constraint for LSP path computation.

On a bundled link we assume that either the whole link is configured with the Link Descriptor Types, or each of its component links are configured with the Link Descriptor Types. In the latter case, the Link Descriptor Type of the bundled link is set to the set union of the Link Descriptor Types for all the component links.

It is possible that Link Descriptor TLV will change over time, reflecting the allocation/deallocation of component links. In general, creation/deletion of an LSP on a link doesn't necessarily result in changing the Link Descriptor of that link. For example, assume that STS-1, STS-3c, STS-12c, STS-48c and STS-192c LSPs can be established on a OC-192 link whose Link Type is SONET. Thus, initially in the Link Descriptor the minimum reservable bandwidth is set to STS-1, and the maximum reservable bandwidth is set of STS-192. As soon as an LSP of STS-1 size is established on the link, it is no longer capable of STS-192c. Therefore, the node advertises a modified Link Descriptor indicating that the maximum reservable bandwidth is no longer STS-192, but STS-48. If subsequently there is another STS-1 LSP, there is no change in the Link Descriptor. The Link Descriptor remains the same until the node can no longer establish a STS-48c LSP over the link (which means that at this point more than 144 time slots are taken by LSPs on the link). Once this happened, the Link Descriptor is modified again, and the modified Link Descriptor is advertised to other nodes.

Note that changes to the Link Descriptor TLV will also affect the Unreserved Bandwidth sub-TLV with respect to bandwidth available on the link.

Internet Draft [draft-kompella-ospf-gmpls-extensions-01.txt](#) February 2001

## 5.6. Shared Risk Link Group TLV

A set of links may constitute a 'shared risk link group' (SRLG) if they share a resource whose failure may affect all links in the set. For example, two fibers in the same conduit would be in the same SRLG. A link may belong to multiple SRLGs. Thus the SRLG TLV describes a list of SRLGs that the link belongs to. An SRLG is identified by a 32 bit number that is unique within an IGP domain.

The SRLG of a LSP is the union of the SRLGs of the links in the LSP. The SRLG of a bundled link is the union of the SRLGs of all the component links. The SRLG values are an unordered list of 4 octet numbers that the link belongs to.

If an LSR is required to have multiple diversely routed LSPs to another LSR, the path computation should attempt to route the paths so that they do not have any links in common, and such that the path SRLGs are disjoint.

The SRLG sub-TLV is a sub-TLV of the Link TLV with type 16. The length is the length of the list in octets. The value is an unordered list of 32 bit numbers that are the SRLGs that the link belongs to. The format is as shown below:

```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               | 4 * No. of SRLGs in link          |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Shared Risk Link Group Value        |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               .....                               |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|                               Shared Risk Link Group Value        |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

The SRLG TLV starts with a configured value and does not change over

time, unless manually reconfigured. The SRLG TLV is optional and if an LSA doesn't carry the SRLG TLV, then it means that SRLG of that link is unknown.

Internet Draft [draft-kompella-ospf-gmpls-extensions-01.txt](#) February 2001

## 6. Security Considerations

The sub-TLVs proposed in this document does not raise any new security concerns.

## 7. Acknowledgements

The authors would like to thank Suresh Katukam, Jonathan Lang and Quaizar Vohra for their comments on the draft.

## 8. References

- [1] Awduche, D., Rekhter, Y., Drake, J., Coltun, R., "Multi-Protocol Lambda Switching: Combining MPLS Traffic Engineering Control With Optical Crossconnects", [draft-awduche-mpls-te-optical-02.txt](#) (work in progress)
- [2] Basak, D., Awduche, D., Drake, J., Rekhter, Y., "Multi-protocol Lambda Switching: Issues in Combining MPLS Traffic Engineering Control With Optical Crossconnects", [draft-basak-mpls-oxc-issues-01.txt](#) (work in progress)
- [3] Katz, D., Yeung, D., "Traffic Engineering Extensions to OSPF", [draft-katz-yeung-ospf-traffic-04.txt](#) (work in progress)
- [4] Kompella, K., Rekhter, Y., Berger, L., "Link Bundling in MPLS Traffic Engineering", [draft-kompella-mpls-bundle-05.txt](#) (work in progress)
- [5] Kompella, K., Rekhter, Y., "LSP Hierarchy with MPLS TE",

[draft-ietf-mpls-lsp-hierarchy-01.txt](#) (work in progress)

- [6] Generalized MPLS Group, "Generalized MPLS - Signaling Functional Description", [draft-ietf-mpls-generalized-signaling-02.txt](#) (work in progress)
- [7] Lang J., Mitra K., Drake J., Kompella K., Rekhter Y., Berger L., Saha, D., Sandick, H., and Basak D., "Link Management Protocol", [draft-ietf-mpls-lmp-02.txt](#) (work in progress)

[draft-kompella-ospf-gmpls-extensions-01.txt](#)

[Page 11]

---

Internet Draft [draft-kompella-ospf-gmpls-extensions-01.txt](#) February 2001

## 9. Authors' Information

Kireeti Kompella  
Juniper Networks, Inc.  
[1194](#) N. Mathilda Ave  
Sunnyvale, CA 94089  
Email: kireeti@juniper.net

Yakov Rekhter  
Juniper Networks, Inc.  
[1194](#) N. Mathilda Ave  
Sunnyvale, CA 94089  
Email: yakov@juniper.net

Ayan Banerjee  
Calient Networks  
[5853](#) Rue Ferrari  
San Jose, CA 95138  
Phone: +1.408.972.3645  
Email: abanerjee@calient.net

John Drake  
Calient Networks  
[5853](#) Rue Ferrari  
San Jose, CA 95138  
Phone: (408) 972-3720  
Email: [jdrake@calient.net](mailto:jdrake@calient.net)

Greg Bernstein  
Ciena Corporation  
[10480](#) Ridgeview Court  
Cupertino, CA 94014  
Phone: (408) 366-4713  
Email: [greg@ciena.com](mailto:greg@ciena.com)

[draft-kompella-ospf-gmpls-extensions-01.txt](#)

[Page 12]

---

Internet Draft [draft-kompella-ospf-gmpls-extensions-01.txt](#) February 2001

Don Fedyk  
Nortel Networks Corp.  
[600](#) Technology Park Drive  
Billerica, MA 01821  
Phone: +1-978-288-4506  
Email: [dwfedyk@nortelnetworks.com](mailto:dwfedyk@nortelnetworks.com)

Eric Mannie  
GTS Network Services  
RDI Department, Core Network Technology Group  
Terhulpsesteenweg, 6A  
[1560](#) Hoeilaart, Belgium  
Phone: +32-2-658.56.52  
E-mail: [eric.mannie@gtsgroup.com](mailto:eric.mannie@gtsgroup.com)

Debanjan Saha  
Tellium Optical Systems  
[2](#) Crescent Place  
P.O. Box 901  
Ocean Port, NJ 07757  
Phone: (732) 923-4264  
Email: dsaha@tellium.com

Vishal Sharma  
Jasmine Networks, Inc.  
[3061](#) Zanker Rd, Suite B  
San Jose, CA 95134  
Phone: (408) 895-5000