

TEAS WG  
Internet-Draft  
Intended status: Informational  
Expires: 25 August 2026

C. Barth  
T. Li  
V. P. Beeram  
R. Bonica  
HPE  
21 February 2026

A Power Conserving Path Placement Strategy (PCPPS)  
draft-many-teas-power-steering-00

## Abstract

A robust network has enough capacity to satisfy demand during peak hours. It has extra capacity to ensure fault-tolerance.

Many networks have a daily utilization pattern. For example, a network might be busy during the day and less busy at night. These networks have sufficient capacity during peak hours, and excess capacity during non-peak hours. Excess capacity increases energy costs and environmental impact.

This document introduces a Power Conserving Path Placement Strategy (PCPPS). When possible, PCPPS concentrates traffic onto a small set of network resources. When traffic is concentrated onto a small set of network resources, other network resources become idle and can be powered down until they are needed again. This solves the problem of excess capacity during non-peak hours.

## Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

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."

This Internet-Draft will expire on 25 August 2026.

## Copyright Notice

Copyright (c) 2026 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 (<https://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. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

## Table of Contents

1. Introduction . . . . .	2
2. Conventions and Definitions . . . . .	3
3. Terminology . . . . .	3
4. Constraint-based Shortest Path Forwarding (CSPF) . . . . .	4
5. PCPPS and CSPF . . . . .	4
6. The PCPPS Metric . . . . .	5
6.1. TE Metric . . . . .	5
6.2. Power Save Capability . . . . .	5
6.3. Power Groups . . . . .	5
6.4. Interface Power . . . . .	6
6.5. Unidirectional Sleeping Bandwidth . . . . .	6
7. Recovering Sleeping Bandwidth . . . . .	6
7.1. Sleeping Links . . . . .	6
8. Power Groups . . . . .	6
8.1. Example Architecture . . . . .	6
8.2. Definition . . . . .	8
8.3. Interfaces and Power Groups . . . . .	10
8.4. Power-Save Capability and Power Group Hierarchies . . . . .	11
9. Security Considerations . . . . .	11
10. IANA Considerations . . . . .	11
11. Acknowledgements . . . . .	11
12. References . . . . .	11
12.1. Normative References . . . . .	11
12.2. Informative References . . . . .	12
Authors' Addresses . . . . .	12

## 1. Introduction

A robust network has enough capacity to satisfy demand during peak hours. It has extra capacity to ensure fault-tolerance.

Many networks have a daily utilization pattern. For example, a network might be busy during the day and less busy at night. These networks have sufficient capacity during peak hours, and excess capacity during non-peak hours. Excess capacity increases energy costs and environmental impact.

This document introduces a Power Conserving Path Placement Strategy (PCPPS). When possible, PCPPS concentrates traffic onto a small set of network resources. When traffic is concentrated onto a small set of network resources, other network resources become idle and can be powered down until they are needed again. This solves the problem of excess capacity during non-peak hours.

Network operators can control the degree to which traffic is concentrated onto a small set of network resources. They can configure constraints that prevent traffic flows from being assigned to a path that does not satisfy their requirements. They can also configure the degree to which power conservation is prioritized in path placement.

## 2. Conventions and Definitions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

## 3. Terminology

This document uses the following terms:

- \* Path - An ordered set of links that connect a source node to a destination node. In a robust network, there are many paths from a particular source to a particular destination.
- \* Traffic flow - A set of packets that have the same source and destination, and traverse the same path. Packets on an MPLS [RFC3031] Label Switched Path LSP are an example of a traffic flow
- \* Constraint - A rule that prevents a traffic flow from traversing some set of paths. For example, a constraint might prevent a particular traffic flow from traversing a path that contains low-speed links.

- \* Traffic Engineering (TE) metric - An administratively assigned attribute of each link in a network. This attribute represents the cost of traversing the link. The cost need not be monetary. It may represent latency (i.e., circuit miles), or any other link attribute.
- \* Sleeping bandwidth - Bandwidth between two nodes that is not currently available because the resources that support it have been powered down.

#### 4. Constraint-based Shortest Path Forwarding (CSPF)

PCPPS leverages Constraint-based Shortest Path Forwarding (CSPF). CSPF can be centralized or distributed onto each node in the network. When it is centralized, it calculates a path for every traffic flow in the network. When it is distributed, each node calculates a path for every traffic flow that originates on it.

As stated in Section 3, many paths can connect a source node to a destination node. CSPF computes a path:

- \* that does not violate any of the traffic flow's constraints
- \* whose links have sufficient bandwidth to support the traffic flow
- \* whose links have the lowest cumulative TE metric

CSPF requires the following inputs:

- \* Information regarding traffic flows (e.g., source, destination, required bandwidth, constraints)
- \* The network topology (i.e., nodes, node attributes, links, and link attributes including the TE metric)

CSPF acquires this information from a Traffic Engineering Data Base (TED). Typically, an Intradomain Gateway Protocol (IGP) populates the TED.

#### 5. PCPPS and CSPF

As stated in Section 4, PCPPS leverages CSPF. However, when PCPPS leverages CSPF, CSPF does not compute a path whose links have the lowest cumulative TE metric. Instead, it computes a path whose links have the lowest cumulative PCPPS metric. Section 6 describes the PCPPS metric.

Furthermore, when PCPPS leverages CSPF and CSPF cannot compute paths due to bandwidth scarcity, it can recover sleeping bandwidth by powering up network resources that were previously powered down. Section 7 describes inputs to the sleeping bandwidth recovery process.

## 6. The PCPPS Metric

The PCPPS Metric is greater than or equal to the TE Metric. The difference between them reflects the cost of a link's power utilization.

The algorithm used to compute the PCPPS is beyond the scope of this document. However, the balance of this section describes the inputs to that algorithm.

### 6.1. TE Metric

The TE Metric is described in [RFC5305].

### 6.2. Power Save Capability

Each TED interface entry includes a Power Save Capability Bit. This bit determines whether the interface can be powered down when idle or nearly idle.

For interfaces that originate on the local node, this bit is administratively assigned and advertised by an IGP. For interfaces that originate on a remote node, this bit is learned by an IGP. See [I-D.many-lsr-power-group].

If the interface is not power save capable, the TE metric and PCPPS metric are equal.

### 6.3. Power Groups

Each TED interface entry includes zero or more references to a Power Group. A Power Group is a hierarchical abstraction of power consumed by hardware components that support the interface. See Section 8.

For interfaces that originate on the local node, this data is administratively assigned or learned from hardware. It is advertised by an IGP. For interfaces that originate on a remote node, this data is learned by an IGP. See [I-D.many-lsr-power-group].

#### 6.4. Interface Power

Each TED interface entry includes a power value, measured in milliwatts. This value represents the amount of power that the interface uses. It does not include the power used by Power Groups to which it is a member.

For interfaces that originate on the local node, this value is administratively assigned or learned from hardware. It is advertised by an IGP. For interfaces that originate on a remote node, this value is learned by an IGP. See [I-D.many-lsr-power-group].

#### 6.5. Unidirectional Sleeping Bandwidth

Each TED interface entry includes a unidirectional sleeping bandwidth value, measured in bits per second. This value represents the sleeping bandwidth on a link. This is useful for LAG adjacencies that have some sleeping members.

For interfaces that originate on the local node, this value is administratively assigned or learned from hardware. It is advertised by an IGP. For interfaces that originate on a remote node, this value is learned by an IGP. See [I-D.many-lsr-power-group].

### 7. Recovering Sleeping Bandwidth

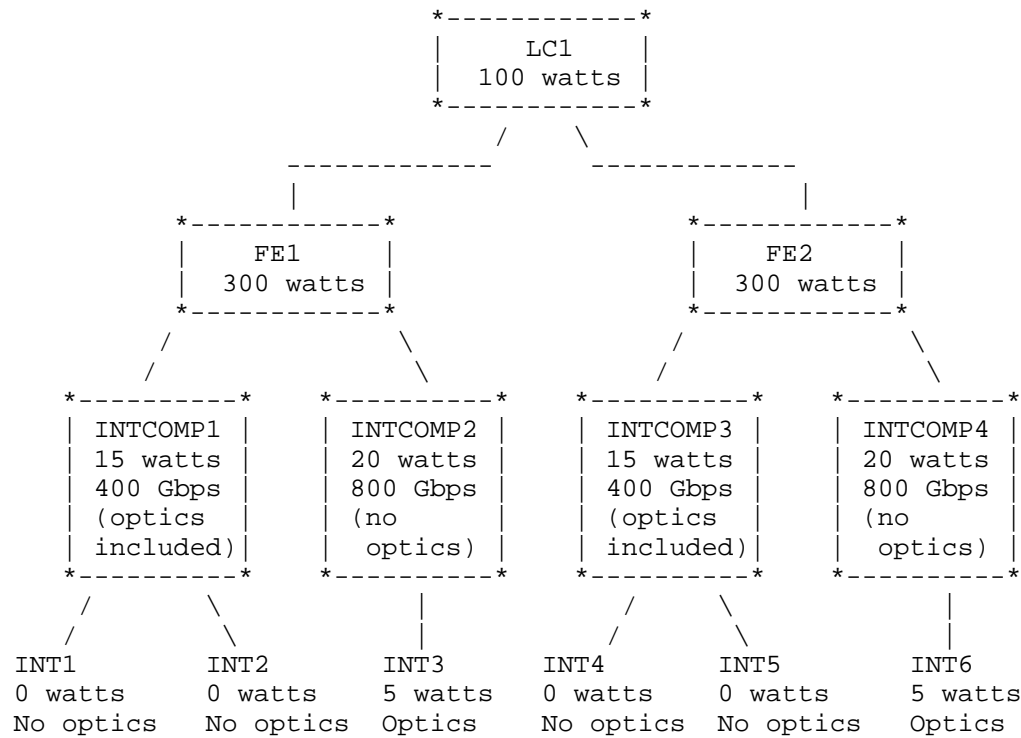
The algorithm used to recover sleeping bandwidth is beyond the scope of this document. However, the balance of this section describes the inputs to that algorithm.

#### 7.1. Sleeping Links

When PCPPS cannot calculate a path due to bandwidth scarcity, it must wake up a sleeping link that might allow the path to be calculated. Therefore, the TED must include information regarding sleeping links. In the TED, sleeping links must be distinguishable from active links. See [I-D.many-lsr-power-group].

### 8. Power Groups

#### 8.1. Example Architecture



Line Card 1 (LC1) consumes 100 watts

Figure 1: Line Card 1

Figure 1 depicts a line card (LC1). LC1 contains two forwarding engines (FE1 and FE2) and four interface complexes (INTCOMP1 through INTCOMP4). INTCOMP1 supports in two interfaces (INT1 and INT2). Likewise, INTCOMP3 supports in two interfaces (INT4 and INT5). INTCOMP2 and INTCOMP4 support one interface each (INT3 and INT6).

An interface complex includes PHY, MAC, encryption, gearbox, and other related circuitry. INTCOMP1 and INTCOMP3 also contain optics. INTCOMP2 and INTCOMP4 do not contain optics. Therefore, the interfaces that they support have their own optics.

INTCOMP1 and INTCOMP3 provide 400 Gbps of forwarding capacity each, while INCOMP2 and INTCOMP4 provide 800 Gbps of forwarding capacity each.

Each hardware component consumes power. LC1 consumes 100 watts while FE1 and FE2 consume 300 watts each. INTCOMP1 and INTCOMP3 consume 15 watts each, while INTCOMP2 and INTCOMP4 consume 20 watts each. INT3

and INT6 contain optics that consume 5 watts each. INT1, INT2, INT4 and INT5 do not have separate optics. Therefore, they do not consume power beyond what is consumed by the complex.

INT1 and INT2 depend upon INTCOMP1. If INTCOMP1 fails, so do INT1 and INT2. Likewise, INT3 depends upon INTCOMP2. If INTCOMP2 fails, so does INT3.

INTCOMP1 and INTCOMP2 depend on FE1. If FE1 fails, so do INTCOMP1, INTCOMP2, INT1, INT2, and INT3. Likewise, INTCOMP3 and INTCOMP4 depend on FE2. If FE2 fails, so do INTCOMP3, INTCOMP4, INT4, INT5, and INT6.

FE1 and FE2 depend on LC1. If LC1 fails, so do all of the forwarding engines, interface complexes, and interfaces in the diagram.

## 8.2. Definition

A Power Group is a hierarchical abstraction of power consumed by hardware components. Each Power Group, except for the one at the top of the hierarchy, has exactly one parent. The Power Group at the top of the hierarchy does not have a parent. Many Power Groups can have the same parent.

Each Power Group has one or more components and each component consumes power. The power consumed by a Power Group is equal to the sum of the power consumed by each of its components. The power consumed by a Power Group does not include the power consumed by its ancestors or by its children.

The parent-child relationship reflects dependency. One Power Group is the child of another if any one of the child components depends upon any one of the parent components.

A network device's power consumption characteristics can be described by any number of equivalent Power Group hierarchies. The paragraphs below demonstrate how two equivalent Power Group hierarchies can describe the power consumption characteristics of the line card in Figure 1.



Identifier	Parent	Power Consumption	Hardware Components
1	None	100 watts	LC1
2	1	300 watts	FE1
3	1	300 watts	FE2
4	2	15 watts	INTCOMP1
5	2	20 watts	INTCOMP2
6	3	15 watts	INTCOMP3
7	3	20 watts	INTCOMP4
8	5	5 watts	INT3
9	7	5 watts	INT6

Table 1: A Granular Power Group Hierarchy

Table 1 describes the power consumption characteristics of the line card in Figure 1 using a granular Power Group hierarchy. We call it granular because each Power Group contains only one component. The power consumed by each Power Group is equal to the power consumed by its component.

In Table 1, Power Group 7 is the child of Power Group 3 because INTCOMP4 depends upon FE2. Likewise, Power Group 3 is the child of Power Group 1 because FE2 depends on LC1. Furthermore, Power Group 8 is the child of Power Group 5 because INT3 depends upon INCOMP2. Likewise, Power Group 9 is the child of Power Group 7 because INT6 depends on INTCOMP4.

Identifier	Parent	Power Consumption	Hardware Components
1	None	700 watts	LC1, FE1, FE2
2	1	15 watts	INTCOMP1
3	1	20 watts	INTCOMP2
4	1	15 watts	INTCOMP3
5	1	20 watts	INTCOMP4
6	1	5 watts	INT3
7	1	5 watts	INT6

Table 2: A Less Granular Power Group Hierarchy

Table 2 describes the power consumption characteristics of the line card in Figure 1 using a less granular Power Group hierarchy. We call it less granular because Power Group 1 contains three components (LC1, FE1 and FE2). Its power consumption is equal to the sum of the power consumed by LC1, FE1 and FE2 (i.e., 700 watts).

Power Group 2 and Power Group 3 are children of Power Group 1 because INTCOMP1 and INTCOMP2 depend on FE1. Likewise, Power Group 4 and Power Group 5 are children of Power Group 1 because INTCOMP3 and INTCOMP4 depend on FE2. Finally, Power Group 6 and Power Group 7 are children of Power Group 1 because INT3 and INT6 depend on INCOMP2 and INCOMP4..

Section 8.4 describes how a network device's power-save capability determines which of the equivalent Power Group hierarchies it should advertise.

### 8.3. Interfaces and Power Groups

An interface is not part of a Power Group, even if it contains optics and consumes power. However, an interface can reference a Power Group. When it references a Power Group, it MUST reference the Power Group that contains the interface complex that supports it.

Therefore, Power Groups can be used to associate interfaces that depend on a common set of hardware components and have common power consumption characteristics.

A Link Aggregation Group (LAG) interface requires support from multiple interface complexes. Therefore a LAG interface references every Power Group that contains an interface complex that supports it.

#### 8.4. Power-Save Capability and Power Group Hierarchies

A network device SHOULD advertise the least granular Power Group hierarchy that can exercise its complete power-savings capability.

Assume that a network contains line cards that are power-save capable. Those line cards contain forwarding engines and interface complexes that are also power-save capable. This means that the line cards, forwarding engines and interface complexes can be powered on and off independently of the chassis.

In order to exercise its complete power savings capability, information regarding line card, forwarding engine and interface complex dependencies is required. Therefore, the line card must advertise the granular Power Group hierarchy in Table 1.

Now assume that another network contains line cards that are power-save capable. Those line cards contain interface complexes that are also power-save capable. However, the forwarding engines are not power-save capable.

In order to exercise its complete power savings capability, information regarding line card, and interface complex dependencies is required. However, information regarding forwarding engine dependencies is not required. Therefore, the line card could advertise either the granular Power Group hierarchy in Table 1 or the less granular Power Group hierarchy in Table 2.

#### 9. Security Considerations

TBD

#### 10. IANA Considerations

This document makes no IANA requests.

#### 11. Acknowledgements

TBD

#### 12. References

##### 12.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/rfc/rfc2119>>.
- [RFC5305] Li, T. and H. Smit, "IS-IS Extensions for Traffic Engineering", RFC 5305, DOI 10.17487/RFC5305, October 2008, <<https://www.rfc-editor.org/rfc/rfc5305>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/rfc/rfc8174>>.

## 12.2. Informative References

- [I-D.many-lsr-power-group] Barth, C., Li, T., Beeram, V. P., and R. Bonica, "Using IS-IS To Advertise Power Group Membership", Work in Progress, Internet-Draft, draft-many-lsr-power-group-02, 25 January 2026, <<https://datatracker.ietf.org/doc/html/draft-many-lsr-power-group-02>>.
- [RFC3031] Rosen, E., Viswanathan, A., and R. Callon, "Multiprotocol Label Switching Architecture", RFC 3031, DOI 10.17487/RFC3031, January 2001, <<https://www.rfc-editor.org/rfc/rfc3031>>.

## Authors' Addresses

Colby Barth  
HPE  
United States of America  
Email: [Jonathan.barth@hpe.com](mailto:Jonathan.barth@hpe.com)

Tony Li  
HPE  
United States of America  
Email: [tony.li@tony.li](mailto:tony.li@tony.li)

Vishnu Pavan Beeram  
HPE  
United States of America  
Email: [vbeeram@hpe.com](mailto:vbeeram@hpe.com)

Ron Bonica  
HPE  
United States of America  
Email: ronald.bonica@hpe.com