

Workgroup: CATS
Internet-Draft:
draft-shi-cats-analysis-of-metric-
distribution-00
Published: 13 March 2023
Intended Status: Informational
Expires: 14 September 2023
Authors: H. Shi
Huawei Technologies

Design analysis of methods for distributing the computing metric

Abstract

This document analyses different methods for distributing the computing metric from the service instance to the ingress router.

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 14 September 2023.

Copyright Notice

Copyright (c) 2023 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. Conventions and Definitions](#)
- [3. Requirement of distributing computing metric](#)
- [4. Option 1: Centralized C-SMA + Centralized C-PS](#)
- [5. Option 2: Centralized C-SMA + Distributed C-PS](#)
- [6. Option 3: Distributed C-SMA + Centralized C-PS](#)
- [7. Option 4: Distributed C-SMA + Distributed C-PS](#)
- [8. Comparison](#)
- [9. Security Considerations](#)
- [10. IANA Considerations](#)
- [11. References](#)
 - [11.1. Normative References](#)
 - [11.2. Informative References](#)
- [Acknowledgments](#)
- [Author's Address](#)

1. Introduction

Many modern computing services are deployed in a distributed way. Multiple service instances deployed in multiple sites provide equivalent function to the end user. As described in [\[I-D.yao-cats-ps-usecases\]](#), traffic steering that takes computing resource metrics into account would improve the quality of service. Such computing metrics are defined in [\[I-D.du-cats-computing-modeling-description\]](#). This document analysis different methods for distributing these metrics.

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 BCP 14 [\[RFC2119\]](#) [\[RFC8174\]](#) when, and only when, they appear in all capitals, as shown here.

This document uses terms defined in [\[I-D.ldbc-cats-framework\]](#). We list them below for clarification.

*Computing-Aware Traffic Steering (CATS): An architecture that takes into account the dynamic nature of computing resources and network state to steer service traffic to a service instance. This dynamicity is expressed by means of relevant metrics.

*CATS Service Metric Agent (C-SMA):Responsible for collecting service capabilities and status, and reporting them to a CATS Path Selector (C-PS).

*CATS Path Selector (C-PS): An entity that determines the path toward the appropriate service location and service instances to meet a service demand given the service status and network status information.

3. Requirement of distributing computing metric

The function components of the CATS is defined in [I-D.ldb-cats-framework](see [Figure 1](#), the figure is replicated here for better understanding). C-SMA is responsible for collecting the computing metric of the service instance and distributing the metrics to the C-PS. C-PS then select the path based on the computing metric and network metric.

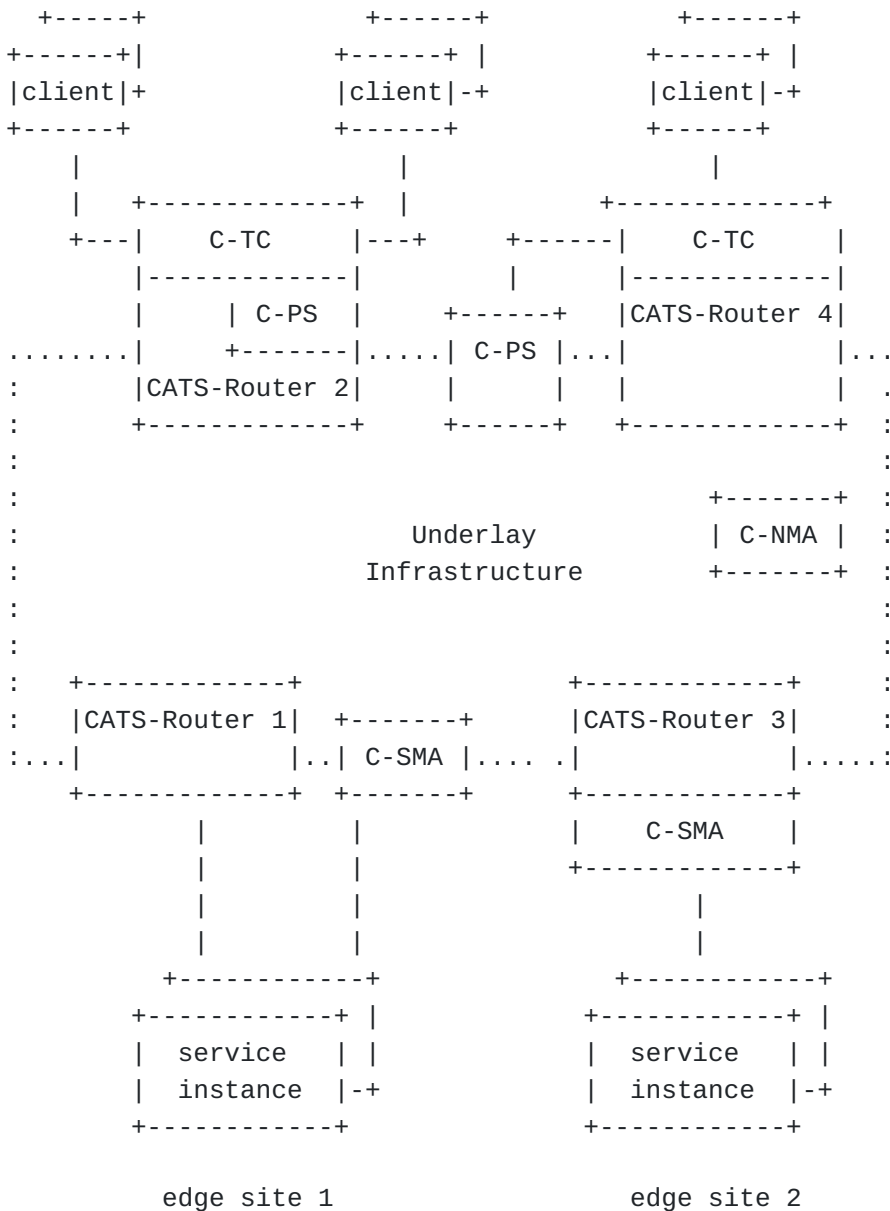


Figure 1: CATS Functional Components

4. Option 1: Centralized C-SMA + Centralized C-PS

The computing metric can be collected internally by a centralized cloud monitor of the cloud platform. Various tools such as Prometheus for this purpose. The cloud monitor can pass the metric to the network controller. Then the network controller calculate the optimal path and distribute the path to the CATS ingress router. The ingress router just need to steer the flow to the path. The network controller distributed the metric update to the C-PS using south-bound protocol.

5. Option 2: Centralized C-SMA + Distributed C-PS

Similar to option 1, the network controller does not calculate the path. It just passes the computing metric received from the cloud monitor to the C-PS in the CATS ingress router. The C-PS at each CATS ingress router will calculate the best path independently.

6. Option 3: Distributed C-SMA + Centralized C-PS

The C-SMA can be deployed in a distributed way. For example, C-SMA running at each site collects the computing metric of the service instance running in the site. Then it reports the metric to the network controller. The network controller calculates the best path and distribute the path to the CATS ingress router.

7. Option 4: Distributed C-SMA + Distributed C-PS

Similar to option 3, each C-SMA collect the computing metric of each site. Then it needs to distribute the metric to C-PS at each ingress router. It can do so directly or through a network controller.

8. Comparison

	Option 1	Option 2	Option 3	Option 4
Protocol	None	Southbound	Southbound	Southbound or Eastbound
CATS router requirement	Low	High	Low	High
Network controller requirement	High	Low	High	Low

Table 1: Comparison between different option

9. Security Considerations

TBD

10. IANA Considerations

This document has no IANA actions.

11. References

11.1. 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>>.
- [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>>.

11.2. Informative References

- [I-D.du-cats-computing-modeling-description] Du, Z., Fu, Y., Li, C., and D. Huang, "Computing Information Description in Computing-Aware Traffic Steering", Work in Progress, Internet-Draft, draft-du-cats-computing-modeling-description-00, 5 March 2023, <<https://datatracker.ietf.org/doc/html/draft-du-cats-computing-modeling-description-00>>.
- [I-D.ldbc-cats-framework] Li, C., Du, Z., Boucadair, M., Contreras, L. M., Drake, J., Huang, D., and G. S. Mishra, "A Framework for Computing-Aware Traffic Steering (CATS)", Work in Progress, Internet-Draft, draft-ldbc-cats-framework-01, 10 March 2023, <<https://datatracker.ietf.org/doc/html/draft-ldbc-cats-framework-01>>.
- [I-D.yao-cats-ps-usecases] Yao, K., Eardley, P., Trossen, D., Boucadair, M., Contreras, L. M., Li, C., Li, Y., and P. Liu, "Computing-Aware Traffic Steering (CATS) Problem Statement and Use Cases", Work in Progress, Internet-Draft, draft-yao-cats-ps-usecases-00, 3 March 2023, <<https://datatracker.ietf.org/doc/html/draft-yao-cats-ps-usecases-00>>.

Acknowledgments

The author would like to thank Xia Chen, Guofeng Qian, Haibo Wang for their help.

Author's Address

Hang Shi
Huawei Technologies
China

Email: shihang9@huawei.com