A Computing Power Method Based on EVPN-VXLAN Extension in Data Center Networks

Nan Wu*
China Telecom Corporation Limited
Research Institute
Beijing, China
wun3@chinatelecom.cn

Guozhen Dong
China Telecom Corporation Limited
Research Institute
Beijing, China
donggz@chinatelecom.cn

Wenjuan Xing
China Telecom Corporation Limited
Research Institute
Beijing, China
xingwj@chinatelecom.cn

Siwei Ji
China Telecom Corporation Limited
Research Institute
Beijing, China
jisw@chinatelecom.cn

Yue Zhang
China Telecom Corporation Limited
Research Institute
Beijing, China
zhangy390@chinatelecom.cn

Abstract

Data center networks need to deliver various cloud services to large number of users, necessitating efficient and flexible management of network resources. However, the widely utilized EVPN-VXLAN (Ethernet Virtual Private Network - Virtual Extensible Local Area Network) technology for network virtualization currently lacks support for computing power resource scheduling within the overlay network. While computing power services can conserve network bandwidth and reduce latency for high-demand applications, they encounter challenges related to end-to-end delivery, path selection, and resource allocation in the EVPN-VXLAN data center network.This paper proposes a novel computing power scheduling method based on EVPN-VXLAN that defines a new EVPN routing type to facilitate the delivery of multi-resource information, which is then utilized for computing power routing selection and forwarding. The proposed method enhances the performance and efficiency of computing power services within the data center EVPN-VXLAN network at the overlay layer while maintaining compatibility with existing EVPN-VXLAN technology.

CCS Concepts

• **Networks** → Network protocols; Network layer protocols.

Keywords

Datacenter networks, EVPN-VXLAN, Computing power scheduling, EVPN routing type

ACM Reference Format:
Nan Wu, Guozhen Dong, Wenjuan Xing, Siwei Ji, and Yue Zhang. 2024. A Computing Power Method Based on EVPN-VXLAN Extension in Data Center Networks. In *2024 5th International Conference on Computer Science and Management Technology (ICCSMT 2024), October 18–20, 2024, Xiamen, China.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3708036.3708277

*Corresponding author.

ICCSMT 2024, October 18–20, 2024, Xiamen, China

ACM ISBN 979-8-4007-0999-9/24/10
https://doi.org/10.1145/3708036.3708277

1 INTRODUCTION

The evolution of network organizational forms has been a gradual transition from traditional networks and physical connections to cloud-based networks. This shift is further accelerated by the explosive growth of application data and network traffic, which has positioned new businesses centered on cloud services and cloud computing at the forefront of the industry. As a result, both business operations and resources increasingly need to be centralized in the cloud, rendering traditional network models inadequate to meet these emerging demands. In the foreseeable future, the resources required for applications and businesses will continue to exist in the form of cloud infrastructure. As cloud network integration progresses, traditional networks are expected to evolve into cloud networks. The cloud network represents not only a new type of network architecture but also an innovative model for delivering cloud network services. Data center networks dissolve the conventional boundaries between networks and services, serving as a crucial component of the structural reform necessary for enhancing cloud network integration capabilities. A data center network is an architectural framework that interconnects hardware resources—including servers, networking devices, and storage—through specific topological structures, routing protocols, load balancing techniques, and other technologies. This architecture provides a high-performance, highly available, and scalable network that supports a variety of application scenarios, such as cloud computing, big data, and artificial intelligence.

Current research on computing power network technology is still in a phase of diversified development and has yet to establish a unified and systematic standard. As Moore's Law approaches its limits and we face exponential growth in both computing power and network demand, leveraging network advantages to optimize computing power utilization and enhance overall computational capacity has become a critical focus in the industry. The evolution of the internet has facilitated the expansion of computing power, streamlined data flow, and improved user accessibility. Consequently, computing resources have increasingly moved to the edge of networks, resulting in a deployment trend characterized by diverse and heterogeneous distributed computing power. As cloud networks transition from centralized, single deployments to dispersed and diverse edge clouds, the allocation of computing power resources is shifting toward a distributed cloud-edge-end model. Thus, a pressing requirement in contemporary computing power networks is the effective integration of dispersed computing resources across the network, enabling efficient collaboration between cloud, edge, and end-user systems. To address varying computing power business requirements, networks must be capable of dynamic scheduling based on real-time conditions and the utilization of computing resources. This approach facilitates the optimal matching of the most suitable business applications to computing nodes, ultimately leading to enhanced utilization of computing resources and improved network efficiency, as well as fostering a deeper integration of cloud resources.

The efficient transmission of computing power through networks has emerged as a significant technical focus in network research. Existing computing power networks primarily utilize two technical solutions. The first approach is based on traditional three-layer network architectures for computing power resource scheduling, including extensions of routing protocols such as OSPF and BGP. However, this solution can be cumbersome in data center environments that rely on virtualization technology. The reliance on underlying routing increases resource demands, as each network device must maintain multiple computing power routing tables, leading to additional complexity and pressure on network devices. The second approach employs a centralized management and scheduling scheme utilizing a controller, which issues instructions for the scheduling and management of computing power resources. Nonetheless, this method is constrained by the controller's performance, the size of the applicable network, and its inability to synchronize and schedule information automatically based on real-time network data.

In this paper, we propose a management scheme for scheduling computing power resources in data center networks based on VXLAN EVPN. This approach implements end-to-end overlay routing of computing power, which is highly compatible with existing data center networks and eliminates the need for hop-by-hop transmission over the underlay network. As a result, it reduces resource waste, alleviates the burden on intermediate network devices, and allows tunnel endpoint devices to provide effective support.

2 RELATED WORK

The foundation of cloud computing lies in virtualization technology, which allows for the division of network resources into multiple virtual networks. Each virtual network can be allocated to various virtual machines for sharing, effectively separating the logical network from the underlying physical infrastructure. Network virtualization is primarily achieved through overlay technology, which employs tunneling encapsulation to construct a logical Layer 2 network over traditional networks, facilitating end-to-end transmission. Currently, data center networks predominantly utilize EVPN-VXLAN technology. VXLAN (Virtual eXtensible Local Area

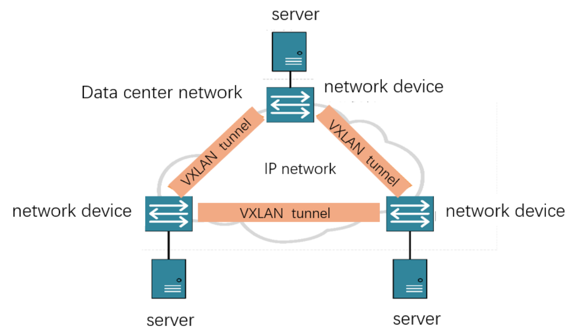

Figure 1: Illustration of EVPN-VXLAN network.

Network) is an NVO3 standard defined by the IETF and is a network virtualization solution derived from VLAN extension. It addresses scalability challenges encountered in large-scale cloud computing deployments. Essentially, VXLAN creates a virtualization tunnel on the IP network, enabling communication between source and destination network devices. From the user's perspective, servers connected to the network function as if they are linked to different ports of a virtual Layer 2 switch, facilitating more convenient communication. As illustrated in the following Figure 1, VXLAN has emerged as the predominant technology for new data center network architectures. Its capabilities to support the dynamic migration of virtual machines, meet multi-tenant requirements, reduce IT operational costs, and enhance business deployment flexibility contribute to its widespread adoption.

Numerous network application cases based on EVPN-VXLAN. However, the transmission of computing power within this framework has yet to be extensively explored. Reference [1] addresses computing power routing technology based on microservices, which establishes end-to-end routing between the network and cloud using SRv6. In this approach, computing power services are treated as segment routes within SRv6 end-to-end routing, orchestrating them into a unified computing network routing strategy. In scenarios involving cross-cluster microservice collaboration and routing addressing, VXLAN Network Identifiers (VNIs) facilitate the virtual networking of application requirements and microservice associations, thereby enabling cross-cluster microservices. The core of the solution proposed in this article hinges on the extension of SRv6 technology, emphasizing the implementation of unified routing orchestration through the Service Function Chaining (SFC) mechanism. Currently, research has not addressed the extension of VXLAN beyond the use of VNI identification, making this approach a novel theoretical exploration. Reference [2] employs an SDN architecture in conjunction with VXLAN and EVPN technology to tackle the challenge of high traffic, commonly referred to as "east-west traffic," between virtual machines. This application aligns with conventional research on VXLAN in existing data center networks but does not address the transfer of computing resources. In Reference [3], a strategic analysis is conducted on key technologies and network deployment strategies for IP bearer networks in computing power networks. It suggests that SDN and IPv6 segmented routing can serve as critical IP bearer technologies, with L3VPN

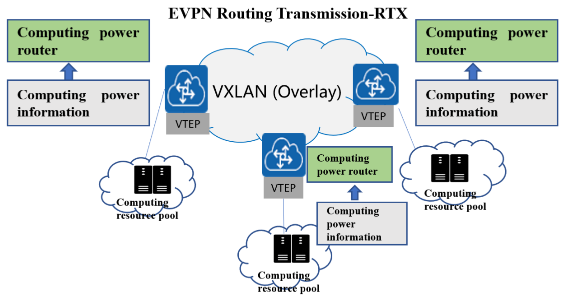

Figure 2: Illustration of computing power information announcement.

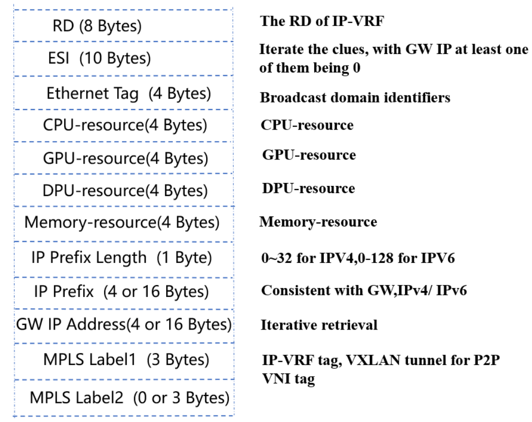

Figure 3: The MGP-BGP computing power route transports the TLV-RT-X(VXLAN) packet format.

over SRv6 proposed as an IP bearer strategy for computing power networks. By incorporating IPv6+, SRv6 technology, network orchestration technology, and network awareness technology, the network structure and performance of computing power networks are optimized. These advancements enhance the flexibility and scalability of the network, enabling dynamic adjustments to network structure and configuration according to varying needs and business scenarios. Both reference [4] and reference [5] focus on the basic technical implementation of EVPN-VXLAN, as well as its usage within and between data centers, without mentioning the transfer of computing power in the data center.

Research on the deep expansion of VXLAN technology and its application in computing power transmission has not been extensively conducted. Currently, studies focused on computing power transmission within computing power networks predominantly emphasize the orchestration and scheduling capabilities of Software-Defined Networking (SDN) and SRv6. Similarly, research on EVPN-VXLAN remains largely centered on traditional cloud computing paradigms.

This article proposes a novel approach based on the EVPN-VXLAN network, wherein computing power information is encapsulated within protocol packets for transmission. This method does not alter the existing architecture of data centers and is independent of controller size, making it suitable for large-scale data centers. The approach also facilitates dynamic protocol periodic announcements and enables real-time awareness of computing power routing.

3 METHOD

In this paper, we propose a new route type (RT) in Ethernet VPN protocol to achieve the computing power routing announcement method, which synchronizes computing power routing information of virtual tunnel end point nodes in the domain and enables computing power scheduling. The new RT in EVPN delivers computing power routing resources, while the original RT remains unchanged, and the computing power routing resources can be announced among VTEP nodes in the local domain. This way, the overlay-based computing power routing forwarding is achieved. Figure 2 shows the illustration of computing power routing information announcements.

There are five commonly used types of EVPN routing, and we name the new EVPN RT routing type RTX, for example RT7, which announces the multicast resource information in the EVPN VXLAN network, realizes the end-to-end computing power information delivery, and does not need to maintain the computing power route forwarding representation on each node, saving equipment resources. The description of the message advertised by the route is as follows:

1. Ethernet A-D Route (Type1): Ethernet automatic discovery.
2. MAC/IP Advertisement Route (Type2): MAC address advertisement, used to announce the host MAC address, host ARP and host routing information.
3. Inclusive Multicast Route (Type3): Integrated multicast for automatic VTEP discovery and dynamic VXLAN tunnel establishment.
4. Ethernet Segment Route (Type4): ES members are automatically discovered.
5. IP Prefix Route (Type5): IP Prefix route (type5) is used to advertise incoming external routes as well as route information to the host.
6. Computing Power Route (TypeX): Computing power route (TypeX) is used to advertise computing power route information.

As shown in Figure 3, the self-defined MGP-BGP computing power route transmits TLV-RT-6(VXLAN) packets in the following format:

This scheme only extends the reserved fields in the standard protocol fields of EVPN, without affecting the integrity of the EVPN standard protocol. It is fully compatible with the standard protocol, and its security and data consistency are consistent with the EVPN standard protocol.

Business process flow (as shown in Figure 4):

1. When the service node is online, VTEP1 node converts the multicast route according to the computing power route information and announces it through RTX routing.
2. VTEP2 receives the RTX route to establish the EVPN VXLAN tunnel.

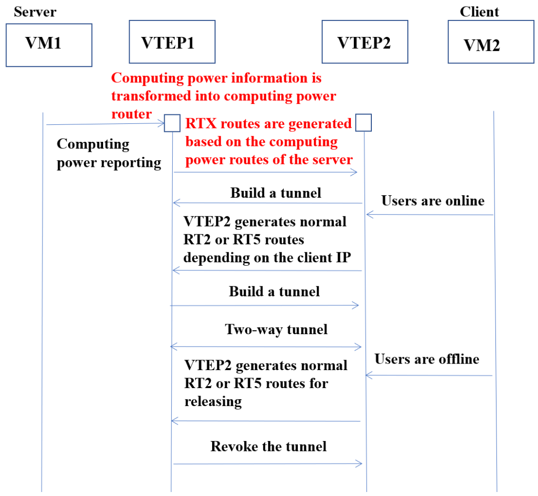

Figure 4: The business process of announcing computing power routing RTX.

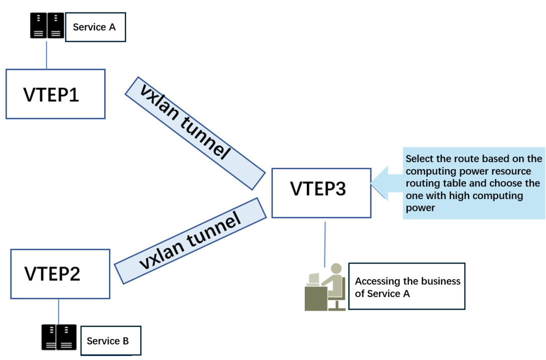

Figure 5: The rule of Route selection.

3. When client VM2 goes online, VTEP2 generates normal RT2/RT5 routes according to client IP for delivery.

4. VTEP1 establishes the EVPN VXLAN tunnel after receiving the RT2/RT5 route.

5. Bidirectional tunnel is established to realize bidirectional communication.

When selecting routes on VTEP nodes, the computing power information transmitted through RT is used as one of the reference criteria for route selection, and the routing calculation of computing power resources is added to the original RT route. The added routing allows users to choose the size of their computing power. The routing rules are as follows (as shown in 5): when selecting routes, VTEP nodes add a metric for computing power routing, When accessing the same service, the path with higher computing power is preferred for transmission. For example, routing can be selected based on the computing power of CPU, GPU, and DPU in RTX routing information, and a VXLAN tunnel can be established with the corresponding VTEP to achieve VXLAN forwarding based on computing power resources.

In certain embodiments, the computing power information encompasses the size of computing power resources. In this context, the first Virtual Tunnel Endpoint (VTEP) identifies a group of servers with computing power resources exceeding a predetermined threshold based on this information. It then designates the VTEP connected to one server within this group as the second VTEP, among multiple available VTEPs. For instance, if the computing power information includes CPU metrics, the first VTEP determines a set of servers with CPU capacities greater than the specified threshold and selects one of these servers to provide services to the client. The VTEP connected to this selected server is then identified as the second VTEP. Furthermore, when the computing power information comprises three types of computing power—CPU, GPU, and DPU—and the client's service request necessitates multiple types of computing power (specifically CPU and GPU), the first VTEP can perform a weighted analysis of these resources. By applying weights to the CPU and GPU resources, it derives a weighted calculation result, enabling the identification of a group of servers whose weighted results exceed the preset value. One server from this group is chosen to service the client, with the VTEP connected to this server designated as the second VTEP. By establishing preset values to filter servers and determine the second VTEP, this approach ensures that the computing power of the servers providing client services meets the requisite standards, thereby enhancing service quality. Moreover, selecting the server with the highest computing power for service delivery further contributes to improving overall service quality.

4 ADVANTAGES

The proposed scheme advertises the computing power route through a self-defined route type X(RT7), facilitating end-to-end delivery of computing power routes in overlay scenarios. Within the data center context, this scheme introduces computing power routing delivery rules that do not deplete underlay computing power routing resources or protocols, thereby enhancing forwarding efficiency. Furthermore, in the EVPN-VXLAN data center network, the scheme enables end-to-end delivery of computing power routing while keeping intermediate links unaware, effectively conserving network resources. This approach establishes an overlay computing power transmission tunnel between tunnel endpoints to connect servers and users, eliminating the need for traditional hop-by-hop transmission and the maintenance of extensive computing power routing entries. The incorporation of routing rules based on computing power allows users to flexibly select routes according to their specific business requirements, thereby providing diversified options for optimal path selection. This flexibility enhances service quality and achieves efficient utilization of network resources.

In the actual application process, we find that this mechanism can be widely used in data center EVPN-VXLAN scenarios. VTEP nodes use computing power routing as one of the forwarding rules to realize the rational application of resources. It can solve the end-to-end delivery problem of computing power routing and the path selection problem of multicast computing power routing in the current data center EVPN-VXLAN network. On the basis of not

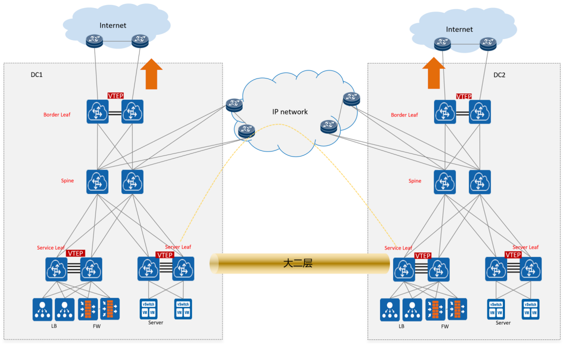

Figure 6: The network diagram of multicast forwarding process based on EVPN-VXLAN.

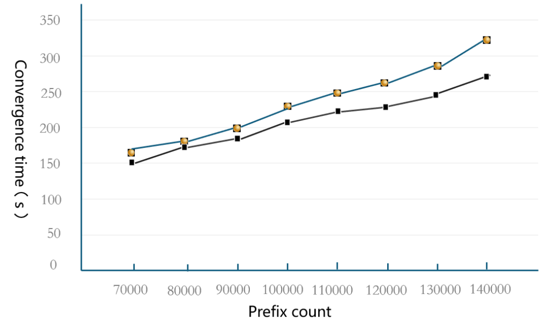

Figure 7: The diagram of experimental results.

changing the original RT routing type, the new selection of computing power routing can meet the requirements of computing power routing scheduling in the data center EVPN-VXLAN network.

5 EXPERIMENTAL SIMULATION

In this paper, we validate the computing power forwarding process based on EVPN-VXLAN through network simulation. As illustrated in Figure 6, we employ a traditional three-layer CLOS architecture for the data center, with servers and users connected to the Virtual Tunnel Endpoint (VTEP). A comparative simulation experiment was conducted to evaluate computing power routing transmission. Specifically, we implemented two transmission methods: one utilizing the underlying BGP protocol and the other employing the overlay EVPN-VXLAN protocol extension proposed in this study. As illustrated in Figure 7, both methods were tested under the same conditions, with an equal number of computing power routing entries, allowing us to monitor and compare the convergence times of the respective routing protocols. The experimental results indicate a significant improvement in routing convergence time with the proposed EVPN-VXLAN method, particularly as the number of routing entries increases. This performance advantage becomes increasingly pronounced, demonstrating that our approach outperforms traditional routing protocols in terms of convergence time. These findings align with theoretical expectations, confirming the feasibility of the computing power routing transfer method based on EVPN-VXLAN. This method not only optimizes the generation of computing power routing tables but also fulfills the requirement for rapid routing convergence, indicating its practical applicability in real-world scenarios.

Additionally, the reserved fields in EVPN are sufficiently long and offer good extensibility, allowing users to define their own uses. If more specific requirements arise in the future, further definitions can be expanded upon. Therefore, the approach presented in this paper demonstrates a high degree of flexibility and does not interfere with standard protocols. For network devices that do not support extended fields, these will be unrecognized and treated according to the standard protocol processing procedures. The security and confidentiality remain consistent with the standard EVPN protocol. The implementation of this method is based on dynamic protocols for the transmission of computational power, which differs from the approach of using controllers for this purpose. However, the method described in this paper can be compatible with controller scheduling. It employs an overlay approach to report computational power, which is beneficial for global computational resource scheduling management, effectively balancing the utilization efficiency of various resource pools, and holds significant application value in next-generation data center networks.

6 CONCLUSION AND FUTURE DIRECTION

This paper proposes a computing power scheduling method based on EVPN-VXLAN for data center networks, aimed at enhancing the efficiency and performance of computing power services within the overlay network. The method introduces a new EVPN routing type for delivering computing power resource information, which serves as the foundation for computing power routing selection and forwarding. This approach effectively addresses challenges related to end-to-end delivery, path selection, and resource allocation of computing power routing in the EVPN-VXLAN data center network.

In the future, we plan to conduct additional experiments to evaluate the performance of the proposed method under various network conditions and compare it with other computing power scheduling techniques. We also intend to explore other potential applications of the new EVPN routing type in data center networks, including load balancing, security, and fault tolerance.

References

[1] Huang, Shi, Tan. 2021. Computing Power Network Resources Based on SRv6 and Its Service Arrangement and Scheduling, ZTE Technologies. 2021 Vol.27 No.3 23-28

[2] Yu, Li, Guo, Luo., Meng. 2023. Novel SDN Architecture and Application Based on VxLAN Protocol, Command Information Systems and Technology. 2023,02,95-100, DOI:10.15908/j.cnki.cist.2023.02.016

[3] Ma, Yang, Mao., 2022. Implementation of Dual-fordwarding-plane Intelligent Path Selection based on SRV6+MPLS (2022). Vol.229 No.1 71-74 DOI:10.13756/j.gtxyj.2022.01.013

[4] Goralski, W. 2017. Chapter 21 - EVPN and VXLAN. In: The Illustrated Network (Second Edition), W. Goralski Ed. Boston: Morgan Kaufmann, San Francisco. pp. 535-560. doi:10.1016/B978-0-12-811027-0.00021-7.

[5] Singh, T., Jain, V. and Babu, G. S. 2017. VXLAN and EVPN for data center network transformation. In: 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT). Delhi. pp. 1-6. doi: 10.1109/ICCCNT.2017.8203947.