

Workgroup: ALTO
Internet-Draft:
draft-zhou-alto-dbors-framework-00
Published: 22 October 2022
Intended Status: Standards Track
Expires: 25 April 2023
Authors: F. Zhou X. Qian D. Yuan
 ZTE Corporation ZTE Corporation ZTE Corporation
 Database-based Open Resource Service Framework

Abstract

This document proposes the framework of Database-based Open Resource Service. It contributes to the overall integration of network and cloud by providing fine-granularity differentiated services and increasing resource utilization rate over the cloud and network.

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 April 2023.

Copyright Notice

Copyright (c) 2022 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. Requirements Language](#)
- [3. Terminology](#)
- [4. DB-ORS Framework and Key Components](#)
 - [4.1. DB-ORS Framework Description](#)
 - [4.2. DB-ORS Implementation Procedures](#)
 - [4.3. DB-ORS Handling Requirements](#)
- [5. Illustration and Designs](#)
 - [5.1. Services Abstraction](#)
 - [5.1.1. VDlink](#)
 - [5.1.2. VTlink](#)
 - [5.1.3. Link Identification](#)
 - [5.1.4. Link Employment](#)
 - [5.2. Services Publishing](#)
 - [5.3. Services Orchestration](#)
- [6. Use Cases](#)
 - [6.1. Cloud Access Scenario](#)
 - [6.2. DCI Scenario](#)
- [7. Security Considerations](#)
- [8. Considerations in the Future](#)
- [9. Acknowledgements](#)
- [10. IANA Considerations](#)
- [11. Normative References](#)
- [Authors' Addresses](#)

1. Introduction

With the rapid development of cloud computing and profound utilization of mobile Internet, the cloud has become an increasingly popular platform for hosting software applications and daily information in a variety of domains such as e-retail, finance, news, and social networking. The demand to connect the clouds accelerates urgently for not only enterprises and companies but also government departments which leads to rapid growth of traffic between terminals and clouds or different data centers.

However, gaps exist among current solutions including complexity in configuration, time-consuming provisioning cycle, constrained access and coarse-granularity services provisioning as clarified in [[I-D.zhou-alto-dbors-requirement-usecase](#)] .

Therefore, a systematic architecture to perform integrated operation of clouds and the network for future scenarios which is defined as Database-based Open Resource Service is proposed in this draft, aiming to provide fine-granularity differentiated services and increase resource utilization rate over the cloud and network.

2. Requirements Language

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.

3. Terminology

*DB-ORS: Database-based Open Resource Service

*SRv6: Segment Routing over IPv6

*VDLink: Virtual Direct Link

*VTLINK: Virtual Tunnel Link

*SID: Segment Identifier

*CPE: Customer Premise Equipment

*DCI: Data Center Interconnection

4. DB-ORS Framework and Key Components

4.1. DB-ORS Framework Description

Network functions are possible to be shared as services by abstracting and encapsulating its resources. Applications subscribe services on their interests and further binds them, so as to satisfy the fine-grained SLA requirements in the context of multiple clouds connected by a unique network area.

DB-ORS abstracts the atomic service capabilities of the network and by introducing common database techniques, realizes service delivery which merely requires single point access. DB-ORS expands and enhances perception abilities of the network by successive procedures including the abstraction of services and capabilities, service publishment and re-orchestration of network services which is shown in Figure 1.

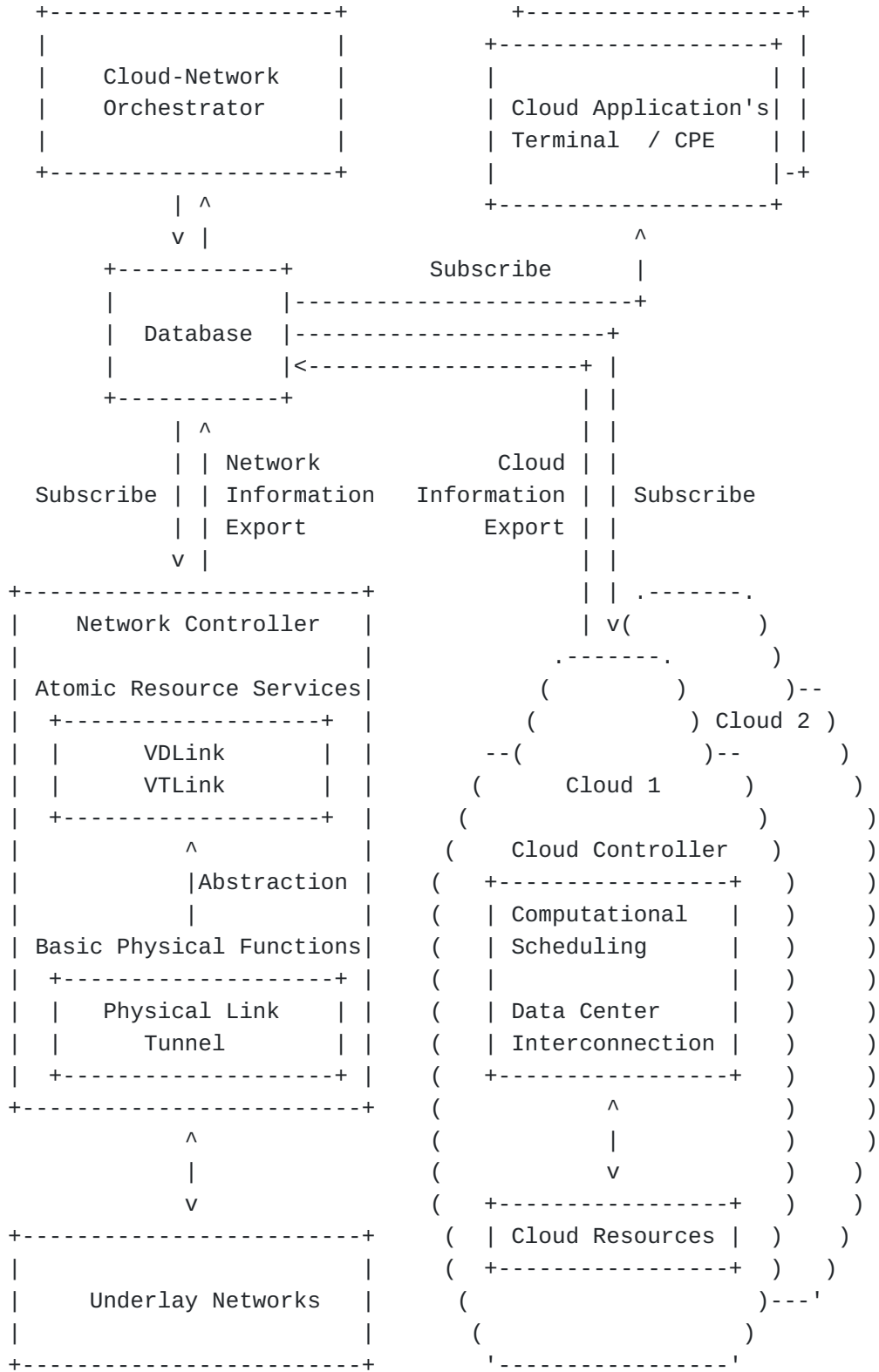


Figure 1: DB-ORS Framework

The key components are introduced as follows.

- *Network controller: the controller for the network domain, whose specially assigned duty in this framework is to draw an abstraction towards network basic physical functions like link and tunnel by extracting key attributes while neglecting unnecessary details, and further encapsulate them into a series of virtual services. Each service can be treated as an unique atomic element without interfering any other services.
- *Cloud-Network Orchestrator: an orchestrator for both the network domain and the cloud domain.
- *Database: a distributed Key-Value database owned by the Cloud-Network orchestrator and shared by both the cloud and Network, with the publish/subscribe mechanism.
- *Application Terminal and CPE: terminals and customer premise equipment which turn out to be service customers, which subscribe their required services from the database.

The key interfaces are introduced as follows.

- *The Northbound Interface (NBI) of the Network Controller: the network capabilities and physical functions is abstracted and exported via this interface to the distributed database, which will be translated into key-value schemes.
- *The Southbound Interface (SBI) of the Network Controller: the network physical topology and fine-granular network information are enforced via this interface from the underlay network. The candidate protocols for this interface are PCEP, BGP, YANG-based protocols, etc.

4.2. DB-ORS Implementation Procedures

The network controller in the bearer network processes the service abstraction of network capabilities which translates basic networking functions into a series of open atomic services. The outgoing atomic services of bearer network are modeled by applying a key-value scheme and further stored into a designated database. Here, distributed databases, ETCD for instance, are recommended for which sustains strong consistency among its operators or visitors. Since the bearer network and cloud applications commonly communicates with the database, a typical subscribe/publish mechanism is applied. To be noted, the unification of a specific key-value scheme is indispensable since multiple manufacturers of network devices and cloud providers are possible to be involved and released information needs to be identified by every participants.

So a schema description file is proposed and utilized to unify various descriptions as a template.

The cloud controller subscribes the updated information of network's atomic services published by the network which stored in the database, and further parses them out referring to the schema description file. Then the cloud controller rearrange and bind the services according to designated principles. Mapping relationships and updated results can also be informed back to the database in need.

4.3. DB-ORS Handling Requirements

Currently, clouds communicate with the network through specific interfaces, Restful for instance. Since the cloud and the network locates in separate domains, third-party infrastructures addressed as super controllers is highly demanded to orchestrate traffic bi-directionally. When a newly registered service capability is provided by the network to deliver to the cloud, a respective application program interface (API) is required which has deficiencies in scalability and simplicity. Also, as analyzed in [[I-D.zhou-alto-dbors-requirement-usecase](#)] , fine-granular service provisioning and enhanced resources utilization is urgently required. Thus, DB-ORS handles the mentioned requirements:

- *With network capabilities abstracted as atomic services, fine-granularity service identification, provisioning can be achieved.

- *With services subscribed by orchestrators, explicit information of the network domain reveals which enables intelligent traffic orchestration and scheduling and increases network resources utilization.

- *Since the database communicates the network domain and the cloud and the inherent publish/subscribe mechanism, the communication framework is flattened and thus the interactions is relatively simplified which brings a profound integration and convergence of cloud and network.

5. Illustration and Designs

5.1. Services Abstraction

Cloud applications are regarded to be important customers of bearer network. In order to meet the customized requirements from different cloud applications at the same time, the bearer network needs to reserve link resources (layer 2) or topology resources (layer 3) in advance respectively, and enable the corresponding cloud applications to invoke the resources allocated to them exclusively.

The resources reserved by the bearer network can be abstracted into the following atomic services:

*virtual direct link (VDLink)

*virtual tunnel link (VTLink)

5.1.1. VDlink

VDLink is a virtual direct link abstracted from a physical direct link. With the definition physical direct link of BGP-LS in [\[RFC7752\]](#) , VDLink is identified by local node descriptor, remote node descriptor, local interface address, remote interface address, and other fields about its resource capabilities.

The attributes of VDLink include: Logic ID, Local Node Descriptor, Local Node Interface Address, Remote Node Descriptor, Remote Node Interface Address, Minimal Unidirectional Link Delay, Maximal Unidirectional Link Delay, Maximal Reservable Link Bandwidth, Unidirectional Link Loss, IGP Metric, TE Metric, END.X SID. The definitions of Logic ID, Local Node Descriptor, Local Node Interface Address, Remote Node Descriptor, Remote Node Interface Address, Maximal Reservable Link Bandwidth, IGP Metric and TE Metric are described in [\[RFC7752\]](#) . The definitions of Minimal Unidirectional Link Delay, Maximal Unidirectional Link Delay and Unidirectional Link Loss are illustrated in [\[RFC8571\]](#) . The definitions of End.X SID are stated in [\[I-D.ietf-idr-bgpls-srv6-ext\]](#) and [\[RFC9086\]](#) .

5.1.2. VTlink

VTLink is a virtual tunnel. There are multiple tunnel types over different dataplanes and SRv6 dataplane is analyzed as an instance in this draft. A VTLink can be abstracted from a SRv6 policy tunnel and its attributes include: Logic ID, Local Node Descriptor, Remote Node Descriptor, Minimal Unidirectional Link Delay, Maximal Unidirectional Link Delay, Maximal Reservable Link Bandwidth, Unidirectional Link Loss, IGP Metric, TE Metric, Binding SID. The definitions of mentioned attributes are described in [\[RFC7752\]](#) , [\[RFC8571\]](#) , [\[I-D.ietf-idr-bgpls-srv6-ext\]](#) and [\[RFC9086\]](#) .

Among them, Maximal Reservable Link Bandwidth is the maximum constrained bandwidth in SRv6 Policy reserved for the VTLink. It should be subtracted from physical bandwidth when any VTLink is allocated, which is similar to the process in VDLink.

Minimal/Maximal Unidirectional Link Delay is the accumulation of the minimal/maximal delay of each node and each segment in segment-list. If multiple sigment-lists are applied, the largest accumulation among all segment-lists is adopted.

IGP Metric is the accumulation of IGP metric of each segment in segment-list. If there are multiple segment-lists in a path, the largest accumulation among all segment-lists is adopted.

TE Metric is the accumulation of TE metric of each segment in segment-list. If there are multiple segment-lists in a path, the largest accumulation among all segment-lists is adopted.

Unidirectional Link Loss is the quotient of the total packets lost in each segment in segment-lists divided by the total number of sent packets. If there are multiple segment-lists in a path, the largest quotient among all segment-lists is adopted.

Regarded as a virtual SRv6 policy tunnel, VLink can be abstracted as a link whose granularity is optional according to service scenario. For instance, for the scenarios of interconnection between data centers, the number of required nodes orchestrated in the segment-list is 2 or 3. However, fulfillment of communications between a terminal and applications in clouds, a VLink may be a complete SRv6 policy path, and the number of required nodes orchestrated in the segment-list may reach 10.

5.1.3. Link Identification

To facilitate path calculation and routing in cloud, a new logic-id is defined to identify a VLink or VLink. The logic-id should be globally unique in the network domain. Its data type is unsigned integer.

The format of virtual link identifier is shown below.

Bit 0 to 4: Cloud-ID (CI), which is used to differentiate applications (such as different cloud service providers).

Bit 5 and 6 are reserved bits for future usage.

Bit 7: Link Type(LT), value 0 stands for VLink while value 1 stands for VLink.

Bit 8 to 31: the value of Logic-id, which ranges from 1 to 16777215 and value 0 is reserved.

SID, local interface address, remote interface address, delay, bandwidth, packet loss rate, TE metric.

*Node: The information of head node and tail node in a real layer 3 link, including, IGP ROUTE ID, AS index, etc.

5.3. Services Orchestration

Based on the published network atomic services, unified service orchestration, path calculation and routing can be fulfilled at an integrated layer of the cloud and the network by an overall cloud-network orchestrator as shown in Figure 1.

In a typical scenario of data center interconnection, the cloud have relatively powerful processing capabilities compared to network terminals. When the cloud controller is informed of its latest assigned virtual links from the bearer network through a subscription scheme, it combines the network information with the collected cloud information and further performs an integrated orchestration. Under current conditions, various services running in the cloud raise respective differentiated requirements for the network. As mentioned above, the cloud is configured as a starting point of the service which also obtains the network open atomic services. After re-orchestration, various paths that satisfy the requirements of different services respectively are calculated. The binding relationship between paths and specific service traffic which help to achieve fine-grained perception for network services.

In another scenario when terminals access cloud services, some POP points or terminals do not have sufficient capabilities to process complex calculation. Information about virtual links like VLink or VDLINK will not published to the network domain. As a substitute, the Cloud-Network orchestrator maps virtual links to an specific identifier, namely a SAN-ID described in [\[I-D.service-identification-header-of-san\]](#) . The SAN-ID represents the SLA information of the network, which is published to a terminal or a CPE. The terminal or CPE carries the SAN-ID in the IPv6 packet as described in [\[I-D.encapsulation-of-san-header\]](#) , and the network edge router forwards it according to the SAN-ID matching the required network path in reference to [\[I-D.computing-segment-for-service-routing\]](#) and [\[I-D.compute-aware-advertise-route-san-database\]](#) .

6. Use Cases

6.1. Cloud Access Scenario

As shown in Figure 3, when terminals access the cloud to acquire specific applications, the orchestration process within the architecture of DB-ORS mainly includes:

- *A Cloud-Network orchestrator is deployed over the network and cloud domain, and a key-value scheme database is configured. The cloud controller informs the running status of computing resources and service SLA requirements to the database.
- *The bearer network allocates exclusive resource for the cloud application which is abstracted as open atomic services like VTLink and is further written into the database.
- *Based on the information collected from the cloud and the network, the Cloud-Network orchestrator implements path calculation, generates appropriate SRv6 policies and assigns a unique SAN-ID to bind respective policies. Then, the SRv6 policy and SAN-ID are written into the database in a key-value scheme.
- *Terminals to access the cloud obtains a specific SAN-ID from the database by subscription while the bearer network controller obtains the SRv6 policy and SAN-ID and advertises the information to all edge routers which serve as traffic entrance by BGP, NETCONF or other feasible means.
- *Service traffic to the cloud sent from the terminal is encapsulated with SAN-ID at CPE. When receiving a packet, edge routers of the bearer network match a specific policy by the identified SAN-ID, and forward the packet to the gateway of the cloud.

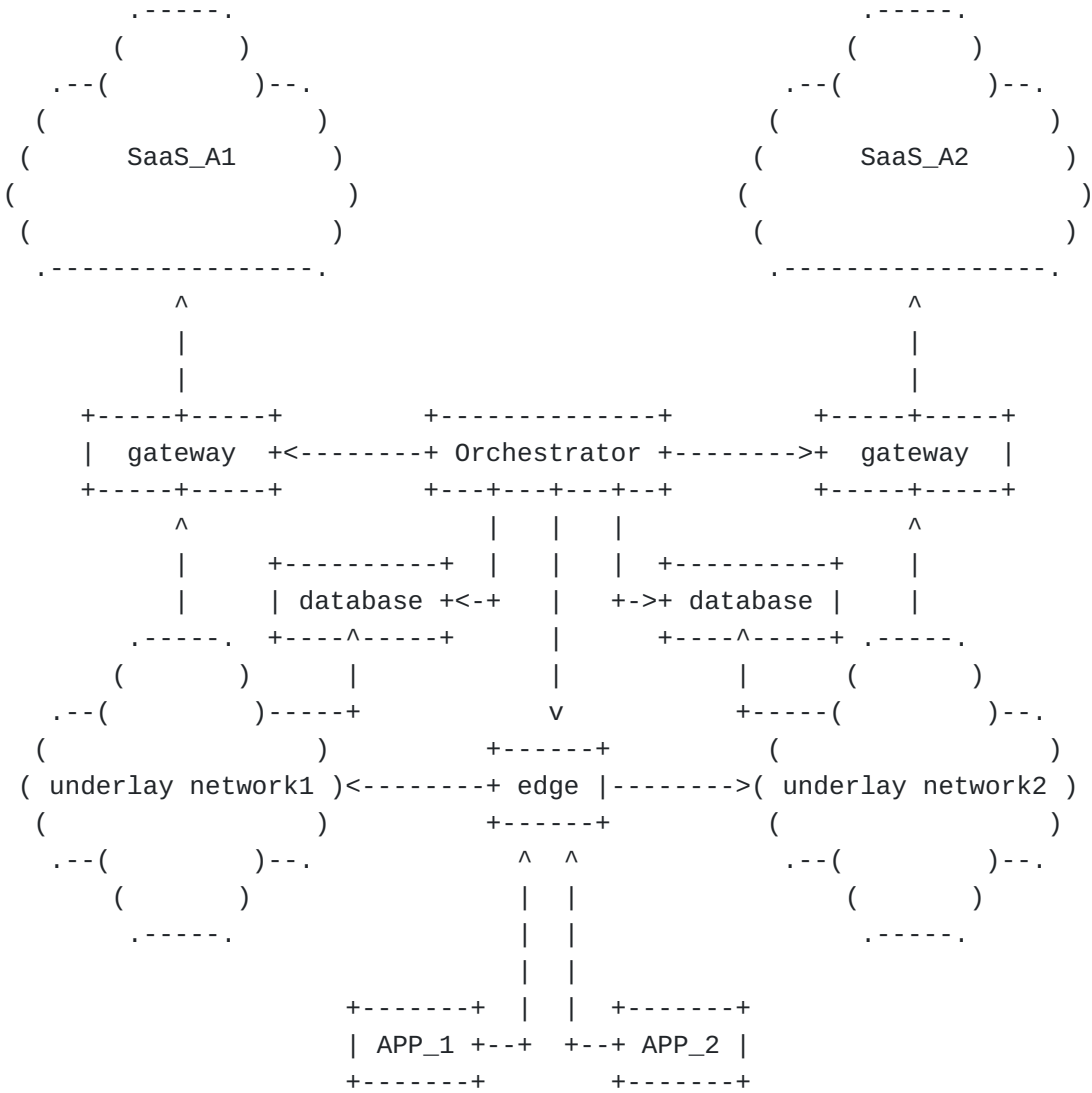


Figure 3: Cloud Access Scenario

6.2. DCI Scenario

As shown in Figure 4, When data centers interconnects with each other, the orchestration process within the architecture of DB-ORS mainly includes:

*Based on the demand for the interconnection between data centers, the bearer network allocates exclusive resources respectively which are abstracted as atomic services and stored in a key-value scheme database.

*Respective controllers subscribe concerning information, and observes variations promptly including additions, deletions and modifications.

*Controllers perform analysis through the unified schema template to obtain the latest network atomic services like VTLink and VDLink. Based on the latest visible topology within both the inner and outer domain of the cloud, the re-orchestration and path calculation can be achieved for the interconnection between data centers.

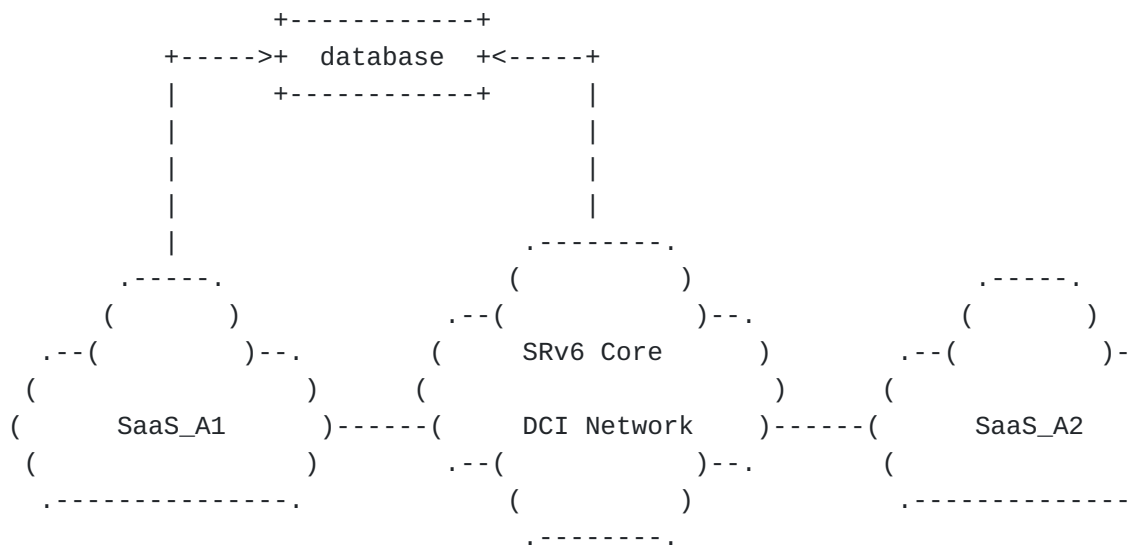


Figure 4: DCI Scenario

7. Security Considerations

Considering the network domain, revealing internal resources obviously brings security problems that must be considered. The exposure of internal topologies, metrics and other privacies in the network is possible to encounter more malicious attacks. The unawakened security drawbacks within database or cloud will also increase the risk. So security protection measures such as anti DDOS attack methods, network security audit, database anti attack schemes, database intrusion detections, access authorization and verification of the database, and other necessary techniques SHOULD be applied. Specific methods will be discussed in detail in the future.

8. Considerations in the Future

The framework of DB-ORS described in this document takes the cloud and the network as a whole for resource allocation and service orchestration, thus improving the service delivery efficiency and enhancing user experiences. Furthermore, It is easy to expand more opened services beyond the link resource services described in this document, such as topology services, security resource services, and deterministic QoS services, which make the framework of DB-ORS be capable of satisfying future requirements appearing with the trend of the convergence of the cloud and the network.

9. Acknowledgements

TBA.

10. IANA Considerations

TBA.

11. Normative References

[I-D.compute-aware-advertise-route-san-database] Zhou, F. and D. Yuan, "Computing Status Awareness, Advertisement and Service Routing methods of SAN Based on Databases", Work in Progress, Internet-Draft, draft-compute-aware-advertise-route-san-database-00, 10 October 2022, <<https://www.ietf.org/archive/id/draft-compute-aware-advertise-route-san-database-00.txt>>.

[I-D.computing-segment-for-service-routing] Zhou, F. and D. Yuan, "Computing Segment for Service Routing in SAN", Work in Progress, Internet-Draft, draft-computing-segment-for-service-routing-00, 9 October 2022, <<https://www.ietf.org/archive/id/draft-computing-segment-for-service-routing-00.txt>>.

[I-D.encapsulation-of-san-header] Ma, L., Zhao, D., and F. Zhou, "Encapsulation of SAN Header", Work in Progress, Internet-Draft, draft-encapsulation-of-san-header-00, 18 August 2022, <<https://www.ietf.org/archive/id/draft-encapsulation-of-san-header-00.txt>>.

[I-D.ietf-idr-bgppls-srv6-ext] Dawra, G., Filsfils, C., Talaulikar, K., Chen, M., Bernier, D., and B. Decraene, "BGP Link State Extensions for SRv6", Work in Progress, Internet-Draft, draft-ietf-idr-bgppls-srv6-ext-11, 14 October 2022, <<https://www.ietf.org/archive/id/draft-ietf-idr-bgppls-srv6-ext-11.txt>>.

[I-D.service-identification-header-of-san]

Ma, L., Zhou, F., and H. Li, "Service Identification Header of Service Aware Network", Work in Progress, Internet-Draft, draft-service-identification-header-of-san-00, 18 August 2022, <<https://www.ietf.org/archive/id/draft-service-identification-header-of-san-00.txt>>.

[I-D.zhou-alto-dbors-requirement-usecase] Zhou, F., Wang, S., and D.

Yuan, "Requirements and Use Cases of DB-ORS (Database-based Open Resource Service)", Work in Progress, Internet-Draft, draft-zhou-alto-dbors-requirement-usecase-00, 22 October 2022, <<https://datatracker.ietf.org/api/v1/doc/document/draft-zhou-alto-dbors-requirement-usecase/>>.

[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/info/rfc2119>>.

[RFC7752] Gredler, H., Ed., Medved, J., Previdi, S., Farrel, A., and S. Ray, "North-Bound Distribution of Link-State and Traffic Engineering (TE) Information Using BGP", RFC 7752, DOI 10.17487/RFC7752, March 2016, <<https://www.rfc-editor.org/info/rfc7752>>.

[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/info/rfc8174>>.

[RFC8571] Ginsberg, L., Ed., Previdi, S., Wu, Q., Tantsura, J., and C. Filsfils, "BGP - Link State (BGP-LS) Advertisement of IGP Traffic Engineering Performance Metric Extensions", RFC 8571, DOI 10.17487/RFC8571, March 2019, <<https://www.rfc-editor.org/info/rfc8571>>.

[RFC9086] Previdi, S., Talaulikar, K., Ed., Filsfils, C., Patel, K., Ray, S., and J. Dong, "Border Gateway Protocol - Link State (BGP-LS) Extensions for Segment Routing BGP Egress Peer Engineering", RFC 9086, DOI 10.17487/RFC9086, August 2021, <<https://www.rfc-editor.org/info/rfc9086>>.

Authors' Addresses

Fenlin Zhou
ZTE Corporation
No.50 Software Avenue
Nanjing
Jiangsu, 210012

China

Email: zhou.fenlin@zte.com.cn

Xiaocong Qian
ZTE Corporation
No.50 Software Avenue
Nanjing
Jiangsu, 210012
China

Email: qian.xiaocong@zte.com.cn

Dongyu Yuan
ZTE Corporation
No.50 Software Avenue
Nanjing
Jiangsu, 210012
China

Email: yuan.dongyu@zte.com.cn