Emerging edge computing has seen latency-sensitive services moving rapidly from cloud to the edge to take advantage of its close vicinity to end users, while cloud is retained for carrying out latency-insensitive and computation-intensive tasks. When more edge computing service providers come to the market, the network will become increasingly more fragmented because of their proprietary services and policies deployed in the network. This means that the Internet can become more cumbersome and riskier as there will be more tiers and potential vulnerabilities that could be exploited. To tackle this issue, we envisage a federated service chaining paradigm in which operators can share and put service functions in other participants’ networks so as to improve resource utilization, collaboratively mitigate cyber threats, and enable service innovations. In this position article, building on our past experience in enabling federated cloud infrastructure and heterogeneous service chaining, we present a Federated Service Chaining architecture followed by discussions of its key components. Several key research challenges are described for the successful realization of such architecture. We hope this article can open a discussion in the research community and generate enough research interest to significantly advance this field.

**Introduction**

In recent years, due to the advancement of computing technologies (e.g., cloud computing and edge computing), most application service providers (ASPs), which provide application functionality and associated services to users, tend to deploy and deliver their services across distributed network domains, such as data centers (DCs), Internet service provider (ISP) networks, and edge clouds, to improve service responsiveness. In fact, end-user experience depends on a collection of those networks working together through interoperation and resource sharing, for example, networking access bandwidth (e.g., ISP peering) and storage and computing service (e.g., content delivery networks, CDNs). This is becoming increasingly more important as we have witnessed a remarkable shift of computation from centralized clouds to distributed network edges. However, the emphasis is in fact on the cloud-edge collaboration in which less resource-demanding tasks are performed on the edge while computation-intensive tasks are carried out in the cloud due to its vast amount of computation power and network bandwidth.

In the meantime, network stakeholders, such as ISPs, carriers, edge cloud providers (ECPs), and cloud DCs, often need to implement complex network policies that match their business objectives, such as traffic engineering, quality of service (QoS), and security. Those policies are usually composed of a sequence of various service functions (SFs) called a service chain (SC), such as firewall, video transcoding, caching, and monitoring [1]. Distinct business objectives of different networks mean that they have various policy requirements and configurations, which have crucial impact on the performance and security of each network [2]. To deliver efficient and reliable services across multiple network domains to users, these stakeholders should cooperate to provide consistent and improved policies and services. Figure 1 illustrates an example scenario for such purpose.

However, existing works mainly focus on service chaining within individual networks (e.g., SCs for cloud DCs [3] or edge networks) [4]. In such scenarios, all SFs and SCs are only the concern of a single stakeholder, and the deployment and management of SCs can be achieved easily. Different stakeholders across multiple network domains make it much more challenging. Furthermore, current underlying network models are usually static and network-dependent, and lack auto-configuration. This imposes significant limitations and challenges on the handling of the rapid change of user traffic patterns as a result of fast service innovations. Although software defined networking (SDN) and network function virtualization (NFV) enable the flexibility to deploy internal policies efficiently, there is still no suitable mechanism at the moment for providing consistent policies across different networks.

As a result, it is important to study and develop techniques to federated service chains with an integrated interoperation layer to deploy applications securely and efficiently across federated network domains so as to improve QoS, resource utilization, and network security. Our previous investigations into service chaining in DCs [2] and mobile edge environments [5] have demonstrated that both resource utilization and latency can be greatly improved when heterogeneity of servers and network devices is...
considered. Such heterogeneity will multiply significantly across multiple network domains. From an economical point of view, we have already seen the success of cloud federation whose global cloud service brokerage market size was valued at US$4.96 billion in 2018 and is predicted to expand at a compound annual growth rate (CAGR) of 17.3 percent from 2019 to 2025 [6]. Thus, our position is that effective federation of networking policy enforcement across multiple network domains is fundamental for the deployment of future Internet applications.

In light of this position, in this article, we conceptualize Federated Service Chaining (FSC), which extends the traditional service chain to cross multiple network domains. A use case is described and analyzed to demonstrate the importance of being able to federate service chains across networks to meet specific security and performance requirements of applications. Then we propose a reference architecture for FSC and investigate essential components and issues for federation based on our past experience in building federated micro cloud infrastructure [7, 8]. Finally, we identify several research directions in FSC. To the best of our knowledge, this is the first study on federation for service chaining.

Federated Service Chaining

FSC Definition

Many applications and services need to be delivered across multiple networks. Each network, which is also referred to as a domain, belongs to one stakeholder. We refer to deploying policies with SFs across a federation formed by two or more domains as FSC. These SFs can be abstracted functions provided by each domain through global services based on virtual machines (VMs) or containers. We also note that network functions and service chains are usually referred to as service functions and service function chains (SFCs), respectively, in Internet Engineering Task Force (IETF) documents [11, 10]. In this article, these terms are used interchangeably.

Since multiple stakeholders are involved, we assume that each domain has a controller, which is responsible to coordinate with other networks in FSC. The whole federated service chain includes two levels of chaining: Intra-domain service chaining refers to chaining of SFs that are available within the same domain, whereas Inter-domain service chaining is the chaining of SFs that reside in two or more distinct domains.

Motivations and Benefits

The main benefits of FSC include:

- **Service level agreement (SLA) assurance:** FSC will let all domains provide consistent and efficient service to users (e.g., consistent priority, latency, and bandwidth) throughout the whole FSC path.

- **Cost effectiveness:** FSC allows each network to share/rent unused resource to reduce cost. Also, if SCs are deployed in each domain independently, duplicated operations and functions are inevitable. With coordination of FSC, redundant functions can be eliminated within FSC.

- **Security:** FSC enables coordination among domains, ensuring deployment of consistent policies. Also, by deploying SFs in the edge, unwanted traffic can be blocked at an early stage of the chain in upstream providers’ networks.

- **Flexibility and Innovation:** Through inter-domain SCs across multiple domains, new applications and services can be designed and deployed easily. For example, functions for prepossessing traffic can be deployed on edge clouds dynamically to utilize distributed edge resource to improve performance.

Use Case: Deploying Micro Services in the Edge

We demonstrate benefits of FSC through a use case in Fig. 1. The introduction of ubiquitous Internet of Things (IoT) and edge computing is fundamentally changing the computational landscape while raising significant cyber security concerns. There have been incidents where IoT devices have been compromised to launch large-scale attacks that have brought down well-known Internet services [11].

Traditional means for mitigating these cyber threats are usually costly. In blackholing, upstream providers must be explicitly informed (e.g., by phone calls) to create policies that drop malicious packets. Local mitigation often employs SFs to, for example, shape incoming traffic, rebalance traffic by manipulation of anycast policies. And applying internal filtering, but they are very costly. With FSC, ASPs can deploy policies at network edges to detect and push back attacks.

Service innovation on collaborative defense: Ideally, cyber threats should be detected and stopped at the network edge. This is achievable because edge devices have abundant computation capability. Should malicious activities take place, lightweight edge-based detection and protection SFs running on these devices can promptly, in collaboration with their adjacent and remote peers, detect Botnet infection, alert about attacks, and subsequently terminate malicious traffic at an early stage before it can cause harm to backbone services.

In FSC, functions located at the network edge can be dynamically instructed to run intrusion detection systems (IDSS) to find attack patterns.
and signatures. In the event of attack, a dynamic SC can be set up in specific providers’ edge networks to force suspected packets to traverse through firewall filtering and IDS. Such collaboration can effectively terminate malicious traffic from the edge of networks. This idea can be extended easily to accommodate any types of innovations on the cross-domain computation such as edge caching and edge data analytics.

**SLA assurance:** More specifically, in this use case, the target end systems and upstream network providers’ links would otherwise be too saturated to serve legitimate users if the attacks are not stopped at an early stage. Thus, it is clear that the FSC can provide better SLA assurance through resource sharing across providers.

**Discussions**

FSC extends traditional SC, but they are significantly different as shown in Table 1. Traditional SC normally operates in silos, deployed within a single network. FSC is deployed across multiple domains from cloud DC to edge networks. ASPs can utilize the heterogeneity of different networks to carry out varied services to safeguard and improve application performance. Since multiple stakeholders are involved, the management of FSC is much more challenging.

Similar to FSC, federated cloud, which has strong industry demand recently, enables efficient and secure deployment of resources and services across distributed cloud infrastructures [12]. For example, BEACON [13] enables provision of federated cloud infrastructures over different cloud management platforms. Those works offer great insights for designing a federation architecture of FSC. Compared to BEACON, FSC further considers more heterogeneous networks than cloud DCs (e.g., edge networks) with special emphasis on deployment of SFs and supporting mobility. Moreover, FSC allows a federated network to share virtualized services in addition to sharing of computation, storage, and networking resources as commonly seen in federated cloud.

**FSC Architecture**

**FSC Three-Plane View**

Figure 2 provides a three-plane illustration of FSC, demonstrating a federated service chain deployed across an ECP and a cloud DC, which includes a federation data plane, a control plane, and an application plane.

The federation data plane consists of all resources that are eligible to be shared through federation. Those resources can be at any abstraction level in the system stack, ranging from instances of VMs/containers (for user deployed functions) to arbitrary application-level functions (e.g., IDS). Each domain has at least one federation controller (FC), which forms the federation control plane. The federation controller is responsible for providing necessary federation functions and interfaces. It can be organized in various manners, such as peer-to-peer or hierarchical deployment (Fig. 4). Different implementation approaches may have different issues concerning consistency, communication latency, and so on. Trust should be established among federated service chain deployment (Fig. 4). Different implementation architec-

|                      | Traditional SC | Federated SC |
|----------------------|----------------|--------------|
| **Stakeholders**     | Single         | Multiple     |
| **Complexity**       | Controller has full control | Complex, limited access to other domains |
| **Consistency**      | Easy to achieve | Difficult, needs interoperation |
| **Flexibility**      | Limited to each domain | Can offload some SFs to other domains |
| **Services diversity** | Limited | More diversity |
| **Troubleshooting**  | Easy | Difficult |

Table 1. Comparison of traditional and federated SC.

Figure 2. Three-plane illustration of FSC.

Federation Application Plane

Federation Control Plane

Federation Data Plane

Figure 3 depicts a high-level view of the components necessary to carry out FSC. A federation is composed of network entities that can be widely geographically dispersed. They can be connected through the federation communication network with both authentication and encryption.

The network provider maintains all resources and includes two main components: service management and the resource abstraction and control layer.

Service management provides interfaces with the federation controller and is broken down into provision and configuration and portability and interoperability functions. Provision and configuration allow automatically deploying SFs based on the requested service and resources, adjusting configurations, discovering and monitoring resources, and SLA enforcement according to defined rules. Portability and interoperability allow customers to move/use their data and SFs across multiple domains with a unified management interface.

IEEE Communications Magazine • March 2020
The resource abstraction and control layer manages a catalog for a set of resources and their attributes that describes how they are exposed and accessed in a federation. These resources can not only be infrastructure resources (e.g., VMs/containers), but also arbitrary application-level services (e.g., instances of SFs). Furthermore, resources to be accessed can be either physical resources or data resources.

Each federation controller comprises three components, service chain management, resource management, and federation management, which can support multiple federation instances, or simply federations. Details of those components are explained in the following.

Federation Instance

Service Chain Management

A federation may observe a number of SCs that would be deployed and scheduled across multiple domains.

Service Composition: A network service is delivered using one or more SFs. The composition of inter-domain SCs is usually high-level and abstract, determined by each ASP according to their requirements. For example, an ASP may decide to deploy firewalls in the edge to filter malicious traffic close to the source, and deploy a cache in regional ISPs to reduce latency of users. On the other hand, intra-domain SC composition is handled by the controller of each domain (e.g., edge cloud). To implement the service required by an ASP, the controller should select appropriate existing instances of functions or create new instances at appropriate locations. Federation controllers can employ their own policies during composition (e.g., multiple instances for load balancing or adding a rate limiter along with the cache).

Service Networking and Routing: Service chaining is concerned with stitching SFs in correct logical order to implement desired network services. This involves placing SFs in the most suitable place in the federated environment. Hence, we advocate the idea of inter-domain chaining and intra-domain chaining, respectively. Inter-domain chaining refers to stitching SFs across different domains, whereas intra-domain chaining means composing the chain within a specific domain. The general principle for FSC is that nodes with better computational resource should be shared as much as possible, and the functions that are available in the edge and IoT environments should be as distributed as possible. When SCs are successfully composed, it is important to steer policy-bounded traffic flows through them. The federation communication network should be responsible for allocating network resources to manage the federated network and overlay networks across dispersed domains. Existing mechanisms such as Generic Routing Encapsulation (GRE), the virtual extensible local area network (VXLAN), and segment routing may be used for steering traffic over service chains in intra-domain chaining. However, inter-domain will need more careful consideration and more complex techniques, depending on types of federated networks, such as GRE, Geneve, or other tunneling protocols. The federation communication network enables all domains to be able to communicate with each other when they belong to the same federation.

Service Chain Deployment: Deploying specific functions onto the network is not only subject to the constraint of underlying physical resources, but also other operational objectives such as service latency. For example, if both operators A and B have sufficient resource to accommodate operator X’s request for a firewall, since the latency to operator B’s network is smaller, it will favor operator B’s network for placing this firewall. Hence, the actual SC deployment is a combinatorial optimization in relation to physical resources. More specifically, two types of chaining are involved:

- Intra-domain service chaining: Intra-domain service chains are private/local service chains that are directly deployed by the controller of the current domain according to its own policy requirement.
- Inter-domain service chaining: Inter-domain service chains maintain the global chaining between two endpoints. Since multiple domains are involved, a common negotiation scheme is required to exchange policy configurations/requirements, routing, and reachability information across domains.

Resource Management

Service Availability and Discovery: We envision a highly dynamic service demand in the FSC as users and things that demand ser-
FSC CHALLENGES

HIGHLY HETEROGENEOUS ENVIRONMENT

With dramatically increased heterogeneity in terms of computing and networking resources across different domains, especially in edge and IoT environments, achieving effective and efficient SCs over heterogeneous networks in a federation environment is very challenging. One fundamental challenge is how resource can be represented and understood, such that a proper decision can be made. Even if a member makes resources available within a federation, portability and interoperability would be needed for members to access and deploy services with reasonable cost. Some challenges concerning heterogeneous environments include:

• Effective sharing of resources and services, including representation of different types of resource/service, on-demand sharing and retrieval, real-time monitoring, enforcement of user-defined policy, and so on.
• Efficient federated SC deployment and scheduling, including inter-domain SC and intra-domain SC deployment.
• Network construction establishment across domains (e.g., overlay network for inter-domain SC).
• Interaction with other services and applications; for example, how to leverage federated clouds, which mainly focus on integration of different cloud platforms, to enfore FSC across cloud service providers.

SC VALIDATION AND VERIFICATION

When SCs are deployed in a federation, it is important to verify whether a particular SC is correctly deployed and enforced as desired to avoid any policy violations. Verifying an SC is a complicated process, especially in a federated environment with multiple stakeholders. A federation controller normally has limited visibility on other members’ networks, while existing solutions for cloud federation assume complete network visibility. Hence, mechanisms are needed for effectively probing and validating federated SC. Such operations must be able to be implemented easily by all federation members. We have previously studied verification and validation of service chaining in the SDN environment [14]. Our experience suggests that for FSC, a more feasible approach is to create a passive probing (black box) scheme that is less intrusive.

Challenges in verification and performance measurement of FSC mainly concern:

• SF availability check (e.g., how to characterize availability of SFs under a federated environment and how to design fine-grained mechanisms for functionality check of SF instances).
• SF performance measurement (e.g., passive measurement using live traffic or active measurement with synthetic probe packets).
• Federated service chain availability check and performance measurement, for example, FSC forwarding path validation on both inter-domain SC and intra-domain SC, policy violation detection, and performance measurements over the entire FSC or a specific segment within the FSC.

FEDERATION MANAGEMENT

Membership management keeps track of active participants of the federation with both federated identity credentials and attributes.

Monitoring and reporting are basic functions that support other components. These include resource usage, performance, health status, and so on.

Accounting and billing are used to track resource usage by federation members, which may need to be associated with pricing or a cost schedule.

Data portability is needed to enable members to access and retrieve data with reasonable cost and format. Similarly, system portability allows images (VMs or containers) and SFs to be able to be moved among federation partners. In the meantime, different federation domains should have a unified management interface or middleware that presents a more unified application programming interface (API) to achieve better service interoperability.

Demand on different services can change over time as a result of the ebb and flow of user traffic on the Internet. For example, large events such as music festivals often create Internet service hotspots in specific areas. One fundamental requirement, therefore, is to adapt the federated SFs with the dynamism of user traffic.
Security and Trust

Different from cloud federation, the biggest challenge of FSC is to enable trust. The cloud federation brings together established businesses that can be trusted partly by their business reputation. However, FSC contains a lot of small operators at the edge of the networks that demand and provide services. They may not have the business reputation needed to be considered trustworthy.

While FSC can be used to improve security for the whole Internet, necessary mechanisms must also be established to defend the federation against attack. One of the most prominent is to prevent the federation from insider attack. This includes a member intentionally removing the firewall or IDS that is required for the downstream members, letting through unfiltered traffic onto the downstream members’ unprotected network.

Concerns regarding trust and secure FSC includes:

• How to go beyond the minimum common denominator of services across domains to enable richer FSC applications and services (e.g., encouraging more ESPs to join FSC)
• How to establish and coordinate trust relationships, authorization, access control, and billing across different network stakeholders, including transitivity and delegation of trust through federation
• How to establish a secure communication channel for both inter-domain and intra-domain

Traffic Flow Classification

An incoming traffic flow should be classified according to classification rules to determine which federation it belongs to and which service chain of the federation will handle it. The classification is based on the content of one or more packet header fields; so the classification rule may vary in granularity. This may cause a problematic case in which an incoming packet matches two or more classification rules of different SCs or federations, which can result in incorrect operations on flows.

Another is the billing mechanism. In cloud federation, it is easy to track the source and bill the upstream providers for the services. However, for FSC, the immediate upstream provider may not be the source of the request, hence making both metering and billing difficult.

Moreover, federation brings additional challenges to the classification in which two similar service chains may exist. For example, if two service chains that belong to two different federations have installed rules to filter traffic destined to the same destination but with contradicting actions (e.g., Drop vs. Allow), the classification system will be unable to correctly classify incoming traffic flows on to the correct service chaining that governs the traffic.

Major challenges in this area include:

• Fine-grained classification schemes for inter-domain SC and intra-domain SC
• Effective violation detection mechanisms for both inter-domain SC and intra-domain SC

Migration and State Consistency

The dynamic nature of the Internet means the services running atop change frequently. On the other hand, the underlying physical resources, edge and IoT in particular, can join and/or leave federations unexpectedly. This dictates the needs for dynamically migrating individual SFs to better utilize physical resources (for improving the return on investment). For this to take place efficiently, the first challenge is that the controller will need to have a very timely and accurate view of the resource availability across the federation, and the service chaining algorithms need to be adaptive. The second challenge is to ensure consistent state after migrating any stateful SFs as incorrect states can easily lead to policy violations.

Major challenges in this area include:

• SFs’ migration within and across network domains, such as algorithms to determine when and where to migrate an SF, and migration schemes to ensure non-interruption of services
• Network update for SF migration, for example, network re-configuration for migration within a domain and overlay network coordination for inter-domain SC
• Efficient state migration to ensure consistency (e.g., consistency levels characterisation, strong or weak), and inconsistency detection and recovery
• Fault tolerance mechanisms for FSC reliability, such as passive or active replication of SF instances and federation controllers, and a topology-independent loop-free alternate path for both inter-domain and intra-domain SC.

Conclusions

In this article, we propose a new networking paradigm called Federated Service Chaining, which focuses on cross-network collaboration harnessing computation resources for greater network-wide resource utilization, enforcing functions, and promoting service innovation. We present the design of the FSC architecture, as well as its major components. Since realization of FSC architecture is not a straightforward task, we also lay out the most prominent research challenges we can foresee. We hope that this article will inspire more further research works in the field of federated service chaining.

Acknowledgment

This work is partially supported by the National Science Foundation of China (NSFC) Nos. 61772235, 61532013, and 61872239; 6007/2018/A1, 0060/2019/A1, the DCT-MoST Joint-project No. 025/2015/AMJ of Science and Technology Development Fund, Macao SAR (FDCT); University of Macau Grants MYRG201600327-FST, CPG2020-0005-SAR, and SRG2018-00111-FST; and the U.K. Engineering and Physical Sciences Research Council (EPSRC) grants EP/P004407/2 and EP/P004024/1.
REFERENCES

[1] P. Quinn and T. Nadeau, “Problem Statement for Service Function Chaining,” IETF RFC 7498, Apr. 2015; https://rfc-editor.org/rfc/rfc7498.txt, accessed Jan. 19, 2020.

[2] L. Cui et al., “Enabling Heterogeneous Network Function Chaining,” IEEE Trans. Parallel and Distributed Systems, vol. 30, no. 4, 2019, pp. 842–54.

[3] M. T. Surendra et al., “Service Function Chaining Use Cases in Data Centers,” Working Draft, IETF Secretariat, Internet-Draft draft-ietf-sfcdc-use-cases-06, Jan. 2017.

[4] H. Huang and S. Guo, “Proactive Failure Recovery for NFV in Distributed Edge Computing,” IEEE Commun. Mag., vol. 57, no. 5, May 2019, pp. 131–37.

[5] T. A. Genez, F. P. Tso, and L. Cui, “Latency-Aware Joint Virtual Machine and Policy Consolidation for Mobile Edge Computing,” 2018 15th IEEE Annual Consumer Commun. & Net. Conf., 2018, pp. 1–6.

[6] G. V. Research, “Cloud Service Brokerage (CSB) Market Size, Share & Trends Analysis Report by Service, by Deployment Model, by Organization Size, by Platform, by End-Use Industry, by Region, and Segment Forecasts, 2019–2025”; https://www.grandviewresearch.com/industry-analysis/cloud-service-brokerage-csb-market, accessed Jan. 19, 2020.

[7] S. J. Johnston et al., “Commodity Single Board Computer Clusters and Their Applications,” Future Generation Computer Systems, vol. 89, 2018, pp. 201–12.

[8] J. Singer et al., “Next Generation Single Board Clusters,” IEEE/IFIP Network Operations and Management Symp., 2018, pp. 1–3.

[9] E. Loukis, M. Janssen, and I. Mintchev, “Determinants of Software-as-a-Service Benefits and Impact on Firm Performance,” Decision Support Systems, vol. 117, 2019, pp. 38–47.

[10] J. M. Halpern and C. Pignataro, “Service Function Chaining (SFC) Architecture,” IETF RFC 7663, Oct. 2015; https://rfc-editor.org/rfc/rfc7663.txt, accessed Jan. 19, 2020.

[11] C. Kolias et al., “DDoS in the IoT: Mira and Other Botnets,” Computer, vol. 50, no. 7, pp. 80–84, 2017.

[12] R. Buyya et al., “A Manifesto for Future Generation Cloud Computing: Research Directions for the Next Decade,” ACM Computing Surveys, vol. 51, no. 5, 2019, p. 105.

[13] E. Huedo et al., “Interoperable Federated Cloud Networking,” IEEE Internet Computing, vol. 21, no. 5, 2017, pp. 54–59.

[14] Y. Zhang et al., “Track: Tracerouting in SDN Networks with Arbitrary Network Functions,” 6th Int’l Conf. Cloud Networking, 2017, pp. 1–6.

BIOGRAPHIES

Lin Cui is currently with the Department of Computer Science at Jinan University, Guangzhou, China. He received his Ph.D. degree from City University of Hong Kong. He has broad interests in networking systems, with focus on cloud data center resource management, data center networking, software defined networking, network function virtualization, congestion control, and so on.

Fung Po Tso is currently a senior lecturer in the Department of Computer Science at Loughborough University. He received his B.Eng., M.Phil. and Ph.D. degrees from City University of Hong Kong in 2006, 2007, and 2011, respectively. His research interests include network policy management, network measurement and optimization, cloud data center resource management, data center networking, SDN, distributed systems, as well as mobile computing and systems.

Weijia Jia is currently a Chair Professor and deputy director of the State Kay Laboratory of Internet of Things for Smart City at the University of Macau. His research interests include smart city, IoT, knowledge graph constructions, multicast and anycast QoS routing protocols, wireless sensor networks, and distributed systems. He has over 500 publications in prestigious international journals/conferences and research books and book chapters.