

算力感知网络(CAN)技术及标准

中国移动通信研究院 基础网络技术研究所 刘鹏

2022年10月

1 算力感知网络(CAN)的背景和发展现状

2 算力感知网络(CAN)的技术和IETF标准化

党中央、国务院

大力发展数字经济；技术和数据作为新的生产要素，作用日益凸显

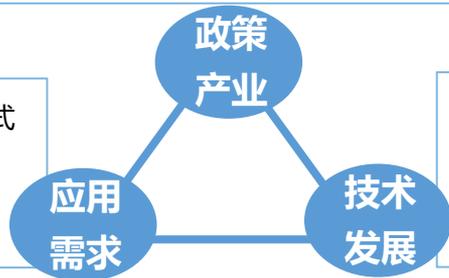
发改委等四部委

数据中心纳入“新基建”
启动“东数西算”工程、“双碳”目标

战略机遇

以网强算，弥补单点算力瓶颈，
提升我国算力整体能力

- **成本优化：**“算”和“网”各自领域资源独占模式成本高，需要组合优化
- **服务保障：**新型应用为保障用户体验对“网”和“算”提出新需求



- **技术突破：**传统网络面临单一学科理论难以突破的挑战
- **跨界融合：**DOICT技术持续融合，相互促进，共同发展

算力网络是**以算为中心、网为根基**，网、云、数、智、安、边、端、链 (ABCDNETS) 等深度融合、提供**一体化服务**④的新型信息基础设施。算力网络的目标是实现“**算力泛在、算网共生、智能编排、一体服务**”，逐步推动算力成为与水电一样，可“一点接入、即取即用”的社会级服务，达成“网络无所不达，算力无所不在，智能无所不及”的愿景。

——中国移动《算力网络白皮书》



起步：**泛在协同**

向算而生，算网协同
协同规划、协同运营



成熟：**融合统一**

算随网流，算智融合
统一编排、统一运营



跨越：**一体共生**

算网一体，算智共生
设施共生，一体供给

算网一体相关技术在学术、标准、产业等方面已经成为热点

学术研究

计算和网络的融合要追溯到分布式计算理论的提出

- 1965年，Dijkstra在其论文中首次提出并发计算的解决思路，开创了分布式计算的理论研究。
- 2003年以来，Google相继发表《MapReduce》、《Bigtable》、《Google File System》三篇论文，详细阐述Google如何打造基于分布式计算的全球一体化云服务技术底座。
- 基于FPGA的分布式机器学习参数聚合系统iSwitch发表于ISCA 2019。
- 基于可编程交换机的分布式机器学习参数聚合系统ATP (Aggregation Transmission Protocol) 获NSDI 2021最佳论文。

标准推进

国内外组织已经逐步开展在算网融合方面的相关工作，成为标准讨论热点

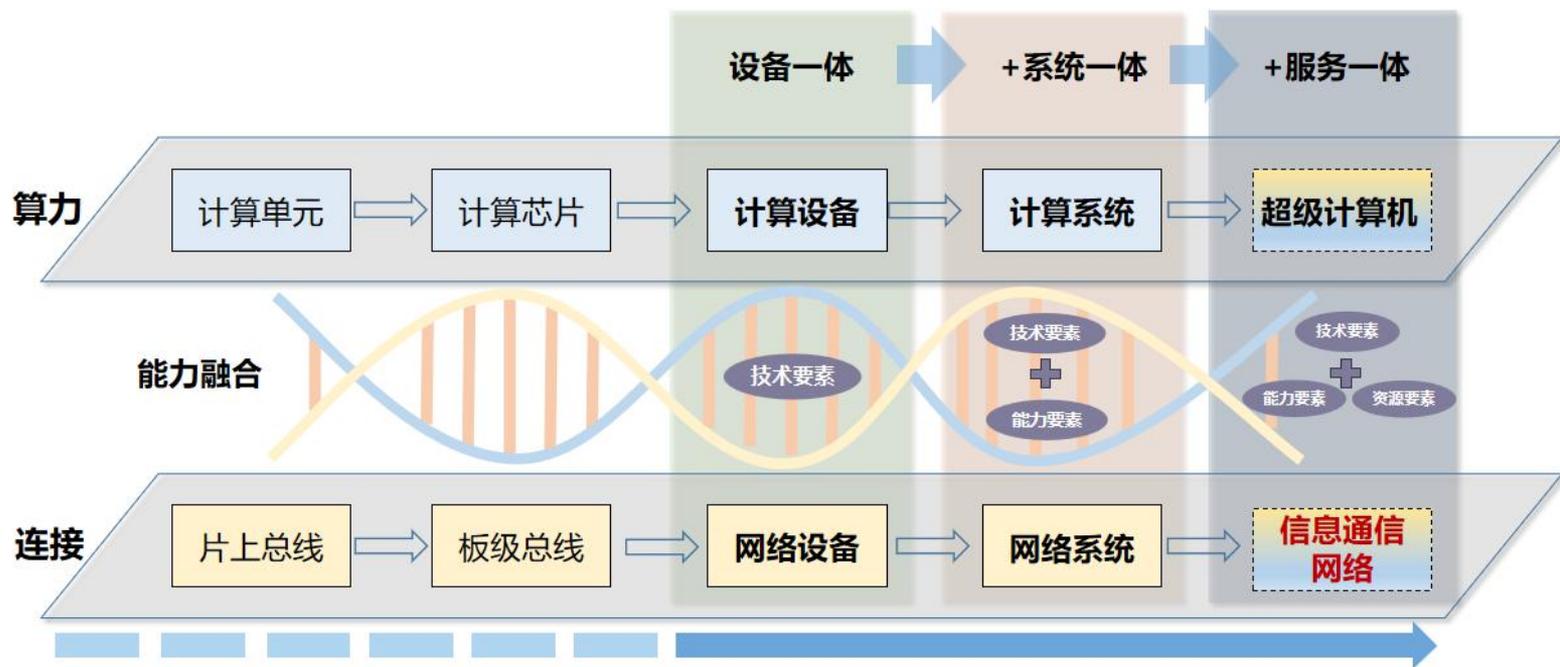
- ITU-T: 自2019年起开展相关标准化工作，2021年算力网络标准术语统一为CNC，初步构建算力网络国际标准体系。
- IETF: 2019年起，IETF开展算力路由协议工作，召开3次算力路由研讨会，2022年中国移动牵头召开CAN BOF，推动算力路由工作组。
- 3GPP: 中国移动在SA1发起移动算力网络项目讨论，引发大量关注。
- CCSA: 形成涵盖算力路由、度量、交易的行标体系，中国移动牵头完成首个行标《算力网络总体技术要求》。

产业生态

算网融合产业生态不断繁荣，相关技术逐步达成共识

- 2022年3月31日，算网融合推委会成立，以“计算网络化”和“网络计算化”两大方向为抓手，推动计算与网络技术的深度融合。
- 2021年9月，中国移动在世界计算大会上联合高校、企业发布“算力网络泛在调度”原型，推进算网融合技术落地。
- 2021年11月和2022年6月，中国移动分别发布《算力网络白皮书》《算力网络技术白皮书》
- 2022年7月，中国移动牵头在Openinfra启动算力网络CFN开源项目

算网一体最基本的组成单元是计算设备和网络设备，最初开始从设备层面呈现技术要素的融合，并且随着技术要素、能力要素、资源要素的不断驱动，由设备一体向系统一体发展，最终实现服务一体



算网一体将呈现出以计算为主和以网络为主的两种发展路径和目标

面向算网一体演进的未来网络设计原则需要顺应IP网络的发展历史和趋势，同时考虑自顶向下、演进与变革共存、跨域融合创新等原则



考虑自顶向下原则

以架构统一牵引算网的融合设计

算网一体的网络将不再是简单的连接媒体，计算元素的引入将带来设备、协议、调度、服务等不同层面的变化，需要自顶向下的从架构出发考虑整体设计。



演进与变革共存原则

平滑分阶段演进与有限域技术变革相结合

算网一体的创新很大程度上依赖于未来IP技术创新，IP创新难度大、周期长，是创新中的塔尖。新技术的创新需要考虑后向兼容，在继承传统IP优势的基础上，实现IP在有限域的更灵活应用。



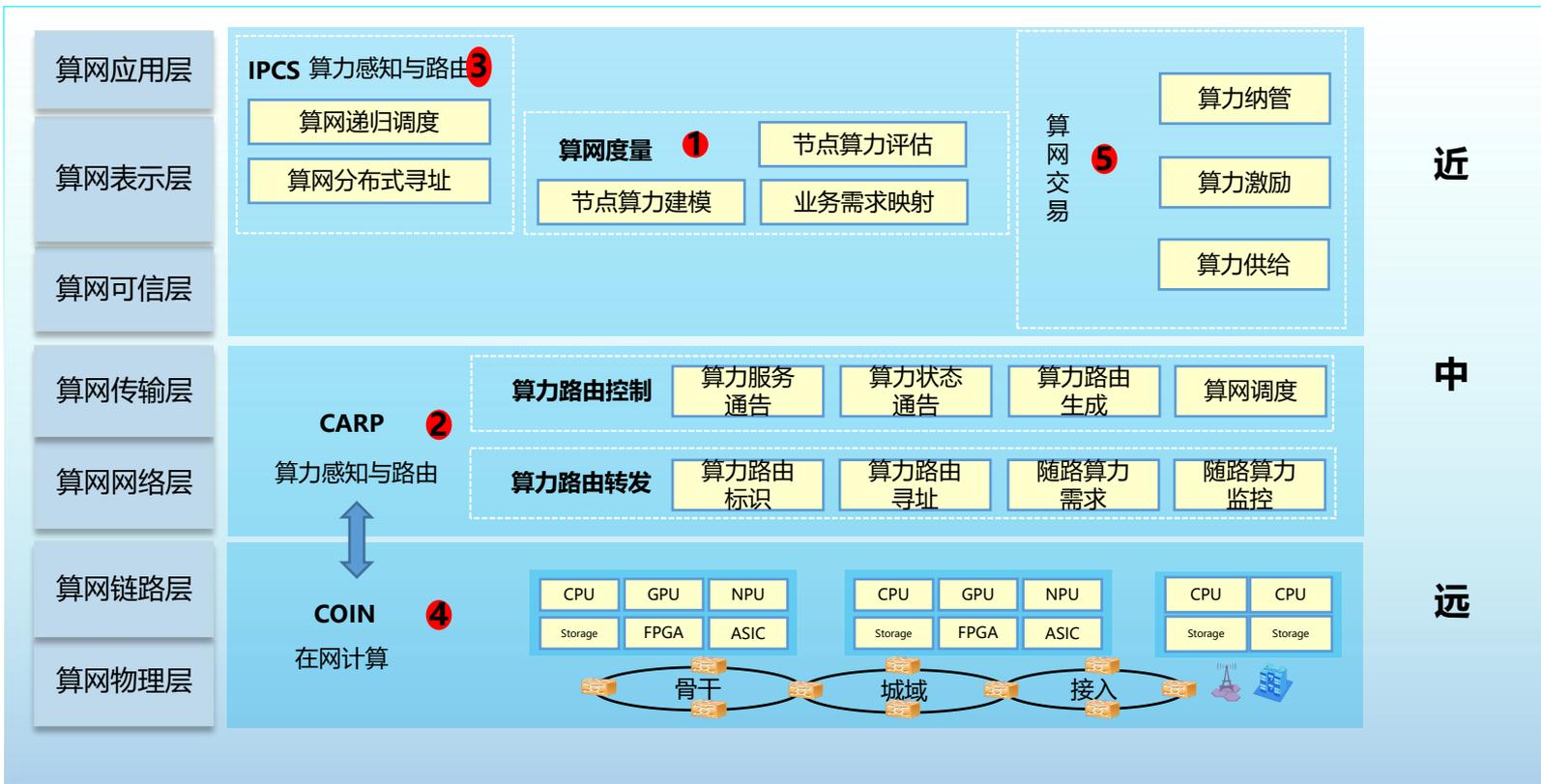
融合创新原则

融合应用、资源多要素创新

面向算网一体演进的网络需要充分发挥基础连接的优势，通过网络感知和融合应用需求、计算资源等多要素，实现算网共生，提供一体化服务。

面向算网一体的算力感知网络CAN技术体系

面向算网一体阶段，算网协同感知成为网络的核心诉求，提出算力感知网络CAN（Computing-aware Networking），基于5大核心技术实现全网算力资源和算力服务的可感、可控和可管



2019算力感知网络白皮书



2021算力感知网络白皮书



2022算网一体白皮书

近

中

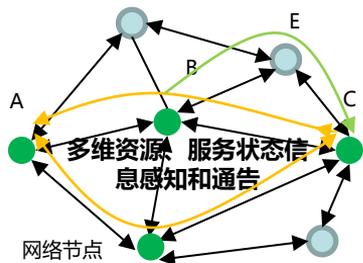
远

1 算力感知网络(CAN)的背景和发展现状

2 算力感知网络(CAN)的技术和IETF标准化

算力路由体系将算力信息引入路由域，通过统一控制和调度实现算力“自由流动”，优化算力和网络资源，提供用户极致体验

多维资源状态感知和通告



1.算力感知：路由层可感知带算力资源信息，生成并动态更新算力拓扑。

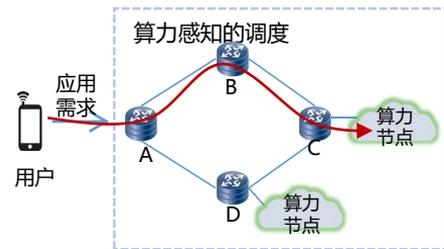
业务需求感知：通过IPv6/SRV6扩展头中加入应用的ID信息和SLA 需求信息。

算力路由生成与调优

路径	算力参数	网络参数
A->B->C	xx	xx
A->D->C	xx	xx
B->E->C	xx	xx

2.算力路由信息表：在网络信息表中新增算力信息，生成算力感知的路由信息表，并支持算力感知的路由优化

调度策略生成与更新



3.算力路由转发：结合算力路由信息表和业务需求，动态、按需生成算力调度和转发策略，可基于SRv6源路由实现网络和算力的可编程

算力信息通告需考虑的问题

1. 算力信息通告的信令开销与信息实时性的折中问题
2. 算力信息通告信令开销与路由调度成功率的均衡问题
3. 算力信息通告的范围问题

3 CFN/dyncast side meetings

2019 IETF106 2020 IETF109 2021 IETF110

Non-WG Forming CAN BoF

2022 Mar IETF113

CAN Progress Presentation

2022 July IETF114

WG Forming CAN BoF

2022 July IETF115



Dynamic-Anycast in Compute First Networking (CFN-Dyncast) Side Meeting
IETF 109, Online Meeting
18 November, 2020
Side meeting chairs:
Ping Lu (China Mobile)
Georgios Karagiannis (University)



Focus on the problem space and use cases

the solution is open for further work

Dynamic Anycast (Dyncast) Side Meeting
IETF 110, Online Meeting
10 March, 2021
Side meeting chairs:
Ping Lu (China Mobile)
Georgios Karagiannis (University)

Time	Topic	Chair
09:00	Introduction	Ping Lu
09:15	Network	Georgios Karagiannis
09:30	Open Discussion / Any other business	Ping Lu
09:45	Meeting	Ping Lu

- draft-liu-dyncast-reqs
- draft-liu-dyncast-ps-usecases
- draft-li-dyncast-architecture
- draft-gu-rtwgwg-cfn-field-trial
- draft-bormann-t2trg-affinity

- draft-liu-dyncast-reqs
- draft-liu-dyncast-ps-usecases
- draft-li-dyncast-architecture

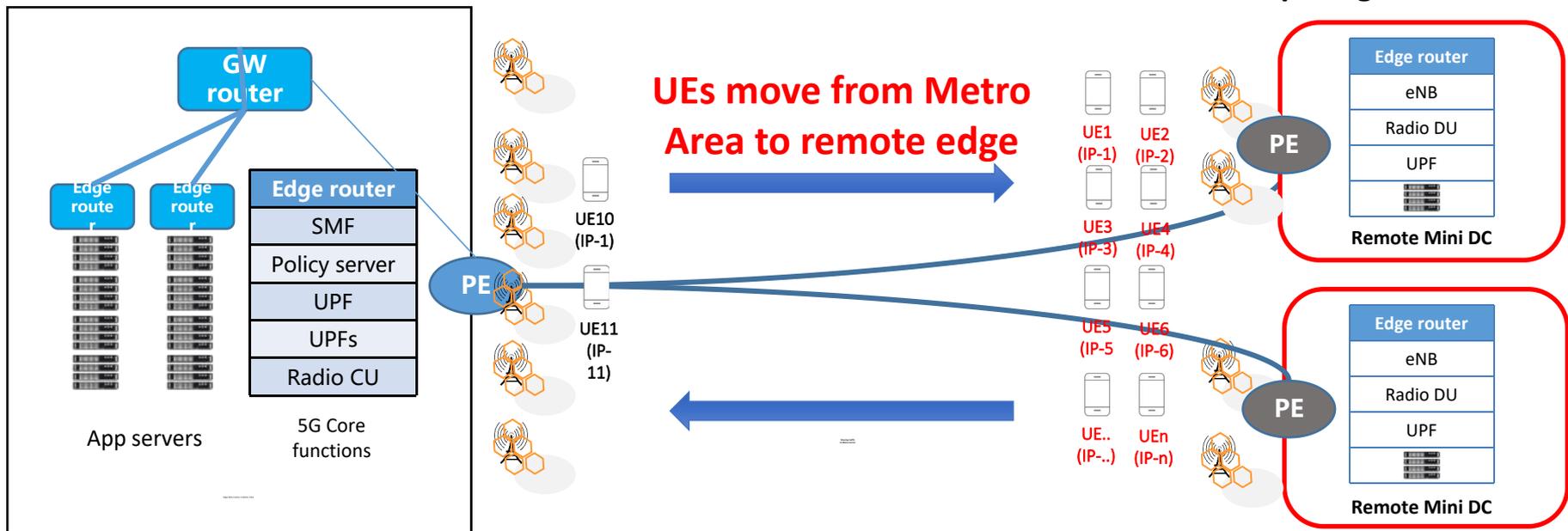
- draft-liu-can-ps-usecases
- draft-liu-can-reqs

经过几年的推进，问题和场景初步达成共识

用户在不同的区域往返引起计算和网络负载的变化，周期仅为每天/每周，难以频繁添加/卸载服务器

- Few UEs in Metro Area
- High computing resource

- Many UEs closes to remote edge
- Limited computing resource



Upper bound latency for motion-to-photon(MTP): less than **20ms** to avoid motion sickness, consisted of:

1. sensor sampling delay: <math><1.5\text{ms}</math> (client)
2. display refresh delay: $\approx 7.9\text{ms}$ (client)
3. frame rendering computing delay with GPU $\approx 5.5\text{ms}$ (server)
4. network delay (budget) = $20 - 1.5 - 7.9 - 5.5 = 5.1\text{ms}$ (network)

计算延迟和网络延迟的预算几乎相等

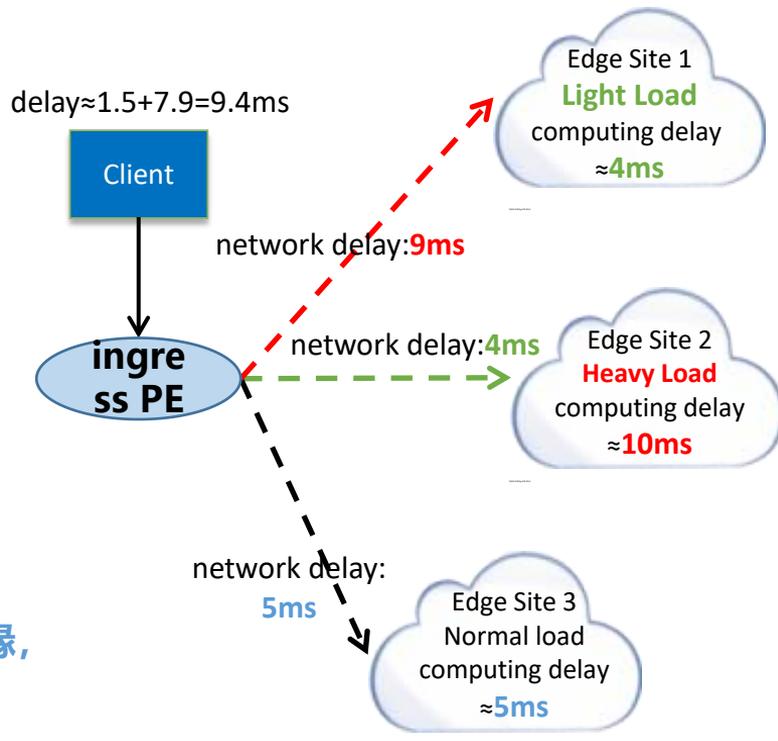


- choose edge site 1 according to load only, total delay $\approx 22.4\text{ms}$
- choose edge site 2 according to network only, total delay $\approx 23.4\text{ms}$
- choose edge site 3 according to both, total delay $\approx 19.4\text{ms}$

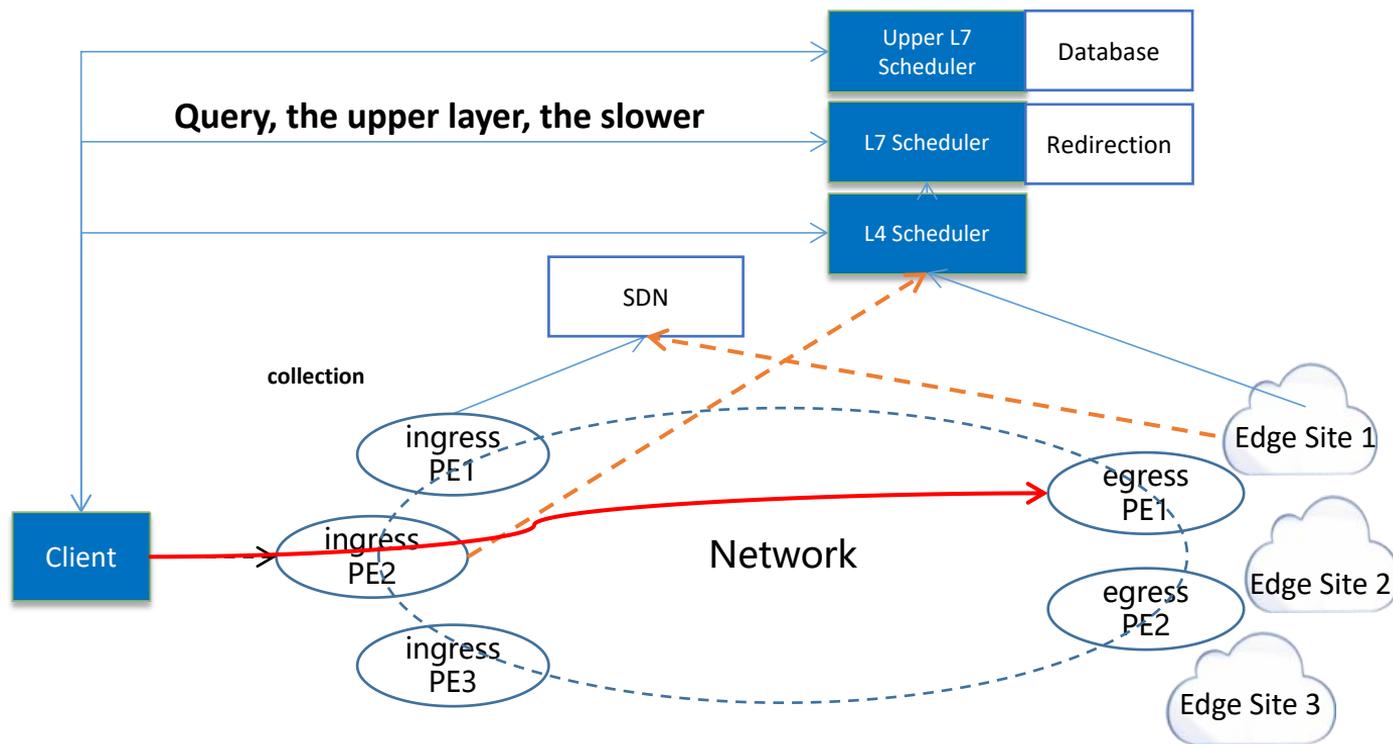
仅根据网络或计算资源状态, 找不到“最佳”服务器实例



需要通过考虑网络 and 计算资源状态, 将流量动态引导到适当的边缘, 以满足E2E延迟要求



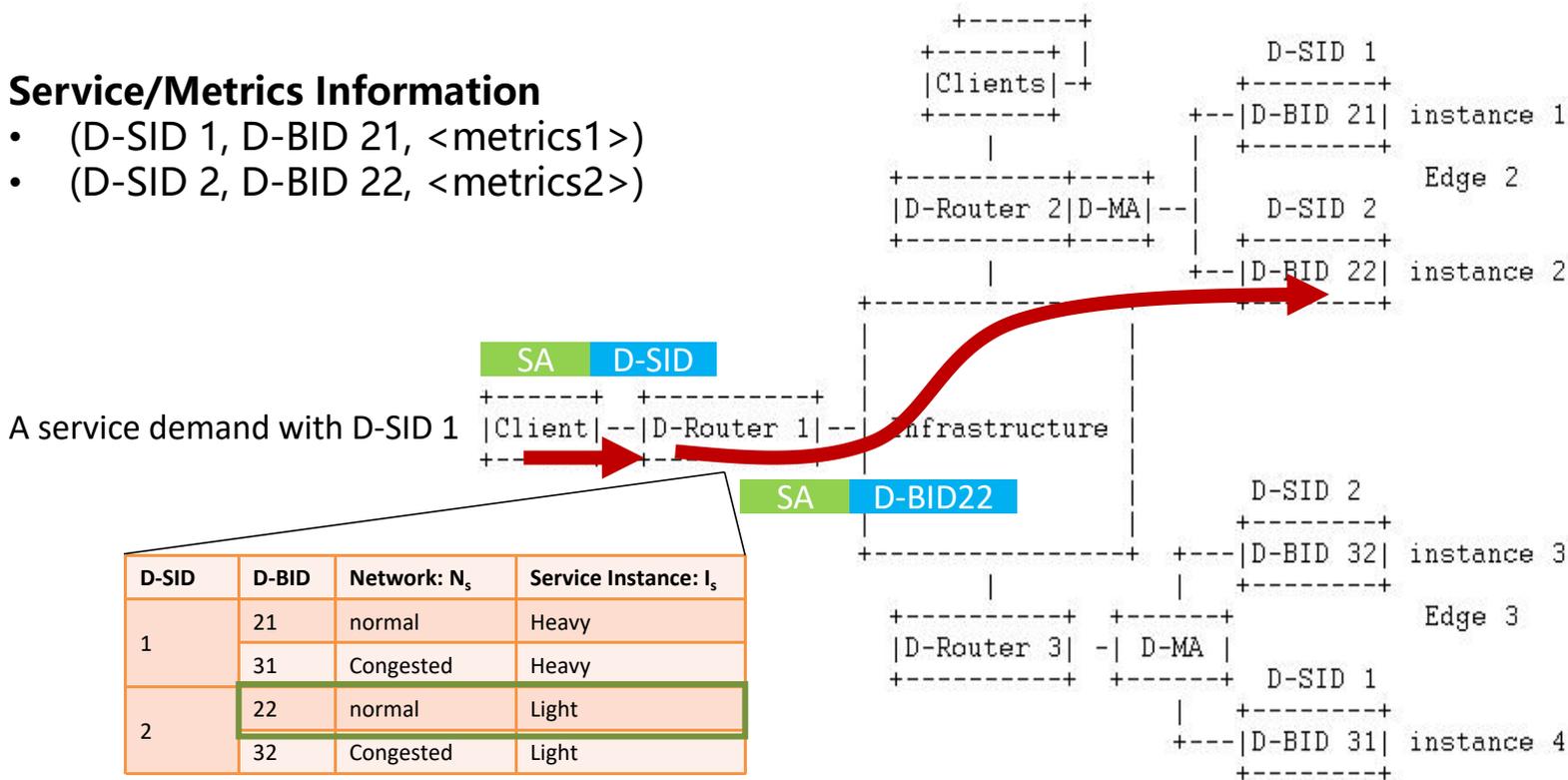
1. 当前缺乏将计算资源条件与网络条件相结合以决定最优路径和节点选择的方案。
2. 现有的解决方案通常为off-path，应用场景相对受限，如DNS、ALTO或L4/L7负载均衡



- Support **multi-access** to the available edge sites dynamically
 - The basic of being aware of the computing info of multiple edge sites
- Provide **both network and computing metrics** for further using
 - Computing resource related model for the metric
 - Rate control signaling of metrics
- Support **effective computing resource representation and encapsulation**
 - Single index or multi-dimensional information for specific purpose
- Support the **session continuity and service continuity**
 - Functional equivalency in different areas
- Preserve **Communication Confidentiality**
 - Weighted value of the metrics or computing information

Service/Metrics Information

- (D-SID 1, D-BID 21, <metrics1>)
- (D-SID 2, D-BID 22, <metrics2>)



- N_s : Network Metrics (congestion, latency....)
- I_s : Instance Service Metrics (load, resources available, ...)

Date: Monday, Nov 7, 2022

Room: Richmond 4

Area Director: John Scudder<jgs@juniper.net>

Time: 15:30-17:30 Session III

Area: Routing Area (rtg)

Mailing list: can@ietf.org

Agenda:

1. Administrivia and agenda---5min---Chairs
2. Problem statement and use cases---15min---Peng Liu/TBD
3. Gap analysis and Requirement---15min---Peng Liu/TBD
4. Quick presentation of proposed charter---10min---TBD
5. Charter discussion---45min---TBD
6. Wrap-up---10min---AD, Chairs

BoF documents and more details could be found here:



You are **very** welcomed to attend this BoF and join the discussions!

大规模确定性网络将包含更多的设备，承载更多种类确定性业务流量，需要扩展更多的队列机制

大规模确定性的增强需求

1. **容忍时间的异步**：时钟同步、频率同步、非同步
2. **支持较大的单跳传播延迟**：周期映射偏差
3. **适应更高的链路速度**：周期以及缓存的调整
4. **良好的设备和流量可扩展性**：流聚合
5. **容忍节点故障和拓扑改变**：双发选收的路径选择
6. **支持多样化队列机制的配置**：同时支持异构队列跨域配置

draft-huang-detnet-rdi-00	9 pages	2022-10-24	I-D Exists
BFD Extension for DetNet Remote Defect Indication (RDI)			New
draft-joung-detnet-async-detnet-framework-01	21 pages	2022-10-24	I-D Exists
Asynchronous Deterministic Networking Framework for Large-Scale Networks			New
draft-liu-detnet-large-scale-requirements-05	20 pages	2022-10-20	I-D Exists
Requirements for Large-Scale Deterministic Networks			New Candidate fo
draft-lyy-detnet-ref-delay-measurement-01	12 pages	2022-06-29	I-D Exists
One-way Delay Measurement Based on Reference Delay			
draft-peng-detnet-deadline-based-forwarding-03	23 pages	2022-10-22	I-D Exists
Deadline Based Deterministic Forwarding			New
draft-peng-detnet-traffic-shaping-solutions-01	16 pages	2022-09-28	I-D Exists
Traffic Shaping Solutions for Bounded Latency in Large-scale Networks			
draft-sw-detnet-network-slice-mapping-yang-00	17 pages	2022-05-06	I-D Exists
YANG Data Model for DetNet Mapping with Network Slice			Expires soon
draft-sx-detnet-mpls-queue-02	11 pages	2022-10-24	I-D Exists
MPLS Sub-Stack Encapsulation for Deterministic Latency Action			New
draft-tan-detnet-cap-discovery-00	15 pages	2022-10-24	I-D Exists
Echo Request/Reply for DetNet Capability Discovery			New

利用数字孪生和意图网络构建虚实交互的**算网数字孪生平台**，通过孪生验证闭环和意图管控闭环的“双闭环控制”，实现存、算、网资源的智能编排和调配以及创新优化策略的低风险、高效率部署

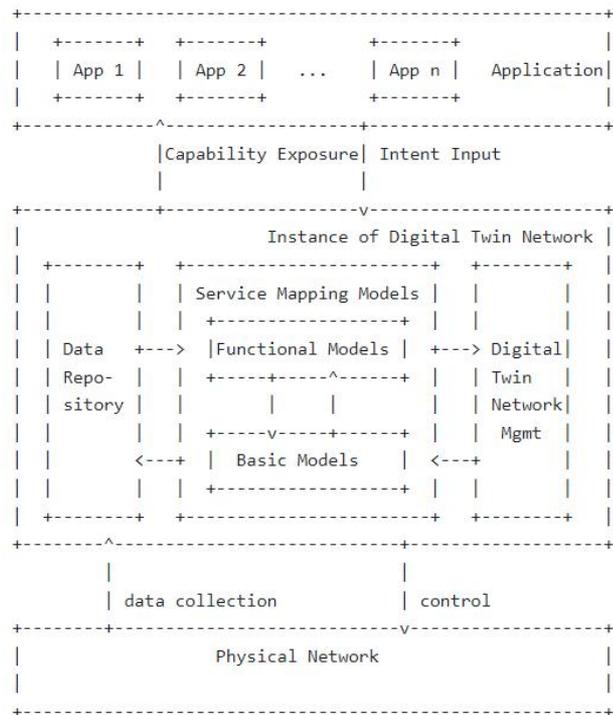


Figure 2: Reference Architecture of Digital Twin Network

工作组文稿:

- [draft-irtf-nmrg-network-digital-twin-arch-02](#)

待立项文稿:

- [draft-yang-nmrg-network-measurement-intent-06](#)

个人文稿:

- [draft-chen-nmrg-dtn-interface-00](#)
- [draft-hong-nmrg-ai-deploy-02](#)
- [draft-yc-nmrg-dtn-owd-measurement-01](#)
- [draft-yz-nmrg-dtn-flow-simulation-01](#)
- [draft-zcz-nmrg-digitaltwin-data-collection-00](#)

欢迎各位专家与中国移动在IETF建立更紧密的合作！！

CAN BoF IETF 115:

<https://datatracker.ietf.org/doc/bofreq-liu-computing-aware-networking/>

中国移动IETF文稿:

https://www.arkko.com/tools/allstats/c_chinamobile.html



CAN BoF

刘鹏 liupengyjy@chinamobile.com



中国移动研究院公众号



中移智库公众号