



确定性网络技术及应用场景研究

Deterministic Networking Technology and Scenarios

魏月华 /WEI Yuehua

喻敬海 /YU Jinghai

罗鉴 /LUO Jian

(中兴通讯股份有限公司, 中国 深圳 518057)
(ZTE Corporation, Shenzhen 518057, China)

DOI: 10.12142/ZTETJ.202004014

网络出版地址: <https://kns.cnki.net/KCMS/detail/34.1228.TN.20200224.1105.004.html>

网络出版日期: 2020-02-24

收稿日期: 2020-02-10

摘要: 结合多种网络场景, 探讨了确定性网络技术的实现原理。与现有技术进行对比研究, 提出了原型系统的设计方案。认为确定性网络技术是以太网和IP网络技术由“尽力而为”向“确定性”发展的新阶段, 可以提供超低的丢包率和可控的端到端时延, 以助力运营技术(OT)和IT的融合, 有效提升网络可用性, 并显著降低网络成本。

关键词: 确定性网络; 丢包率; 时延

Abstract: The implementation principles of deterministic network technology in combination with various network scenarios are discussed. Compared with the existing technologies, the design scheme of a prototype system is proposed. It's considered that deterministic network technology is a new stage in the development of Ethernet and IP network technology, which can provide ultra-low packet loss rate and bounded end-to-end latency, help the integration of operation technology (OT) and IT, effectively improve network availability, and significantly reduce network costs.

Keywords: deterministic network; data loss rate; latency

1 确定性网络的技术特性

1.1 确定性网络的技术概览

确定性网络^[1]是为确定性业务流提供服务的网络。它不限定特定拓扑且不限制连通性, 应用时可以通过网管系统、应用控制器等来进行网络资源预留。对确定性网络感兴趣的许多应用终端, 通常需要能够支持时钟同步至亚微秒级的精度。确定性网络中的一些队列控制技术也需要中继节点和传输节点之间的时间同步。

相对于非确定性网络, 确定性网络的最核心特征是更严格、更明确的服务质量(QoS), 包括: 从源到目标的最小和最大端到端时延和时延抖动, 在节点和链路的各种假设操作状态下

的丢包率, 数据包发生乱序(比率)的上限。

目前, 已经有方法实现可控的时延和丢包来满足很多应用, 例如基于优先级和冗余配置的技术, 然而, 这些技术通常只有在关键流在网络容量中占比很小、网络中的所有系统都运行正常、没有终端系统中断网络操作行为等情况下, 才能工作得很好。确定性网络关心的是端到端延迟在最糟糕情况下的值。平均值或典型值对确定性网络没有意义, 因为它们不代表一个实时系统执行任务的能力。一般来说, 一个普通的基于优先级的队列方案比确定性网络数据流有更好的平均延迟, 但在最坏情况下的延迟却可能是不受控的。确定性网络采用拥塞保护、显式路由、服务保护来提供QoS。

1.2 确定性网络堆栈模型

图1是一个概念性的确定性网络数据面层次模型。图中的“源”和“目的”为应用层。“报文定序”“副本消除”“流复制”“流合并”“报文编码”“报文解码”均是确定性网络服务层的一部分。其中, 报文定序为报文复制和副本消除提供顺序号。如果确定性网络流由更高层的传输协议来执行报文定序和副本消除, 那么就不需要这一层。副本消除基于报文定序功能提供的序列号, 丢弃确定性网络流复制所产生的任何报文副本。副本消除功能还可以对报文重新排序, 以便从因报文丢失而中断的流中恢复报文顺序。流复制将属于确定性网络复合流的报文复制到多条确定性网络成员流里。该功能与报文定序是分离的。流复制

可以是对数据包的显式复制和重标记，也可以通过例如类似于普通多播复制的技术来实现。流合并将属于特定确定性网络复合流的成员流合并在一起。确定性网络流合并与数据包定序、副本消除、确定性网络流复制一起执行报文复制和消除。报文编码可以替代报文定序和流复制的功能。报文编码将多个确定性网络报文中的信息进行组合（这些信息可能来自不同的确定性网络复合流），并将这些信息在不同的确定性网络成员流上用报文进行发送。报文解码可以替代报文合并和副本消除的功能，并从不同的确定性网络成员流中取得报文，然后从这些报文中计算出原始的确定性网络报文。

确定性网络在传输层提供“拥塞保护”。实际的队列和整形机制通常由下层的子网层提供。“显式路由”通过确定性网络传输层提供机制，确保为确定性网络流提供固定的路径。

操作、管理和维护（OAM）可以利用带内和带外信令验证服务在 QoS 约束下的有效性，并可以在数据包中添加特定的标记，以追踪网络运行、传递或发生错误。OAM 未在图 1 中画出，因为它可以存在于任意功能层中。

1.3 确定性网络数据平面

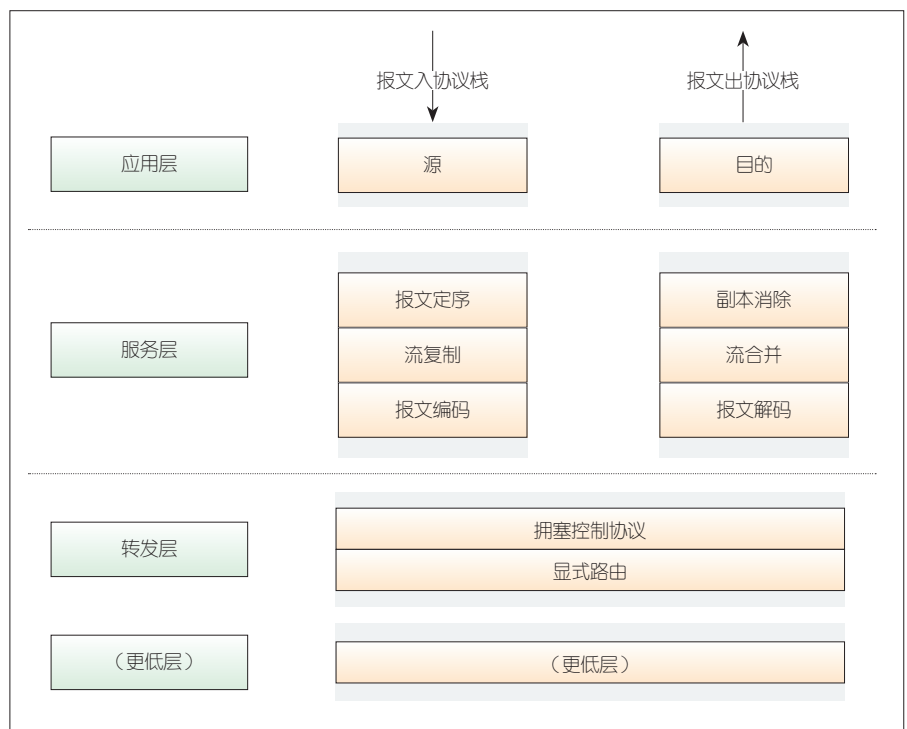
确定性网络由确定性网络使能的终端和节点组成。所有启用确定性网络的节点都连接到子网，其中点对点链接也被认为是简单的子网。这些子网提供确定性网络兼容服务，以支持确定性业务流。子网的例子包括 IEEE 802.1 时间敏感网络（TSN）和光传送网（OTN）。多层确定性网络系统也是可能的，例如将其中一个确定性网络作为子网为更高层的确定性网络系统提供服务。一个简单的确定性网络概念如图 2 所示。

分别为确定性网络服务层、传输

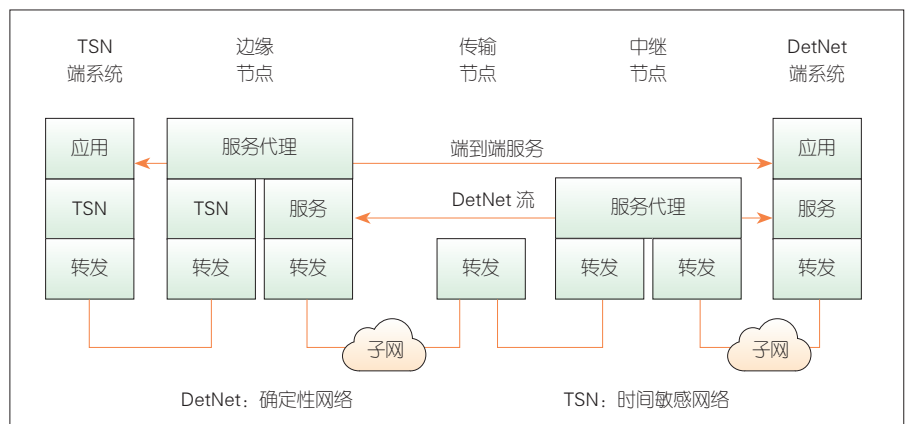
层选择一种技术方法，就可以为确定性网络流提供多种数据平面解决方案。不同的数据平面选项之间最根本的区别，是确定性网络端点系统使用的基本寻址方法和报文头各不相同。例如，可以基于多协议标签交换（MPLS）标签或 IP 报头来递送基本服务。传输层的技术选择会影响确定性网络服务层的基本转发逻辑。在这两种情况下，确定性网络节点都用 IP 地址来标示。所选的确定性网络传输层技术也需要

映射到用于互连确定性网络节点的子网技术，例如，确定性网络需要映射到 TSN 帧。

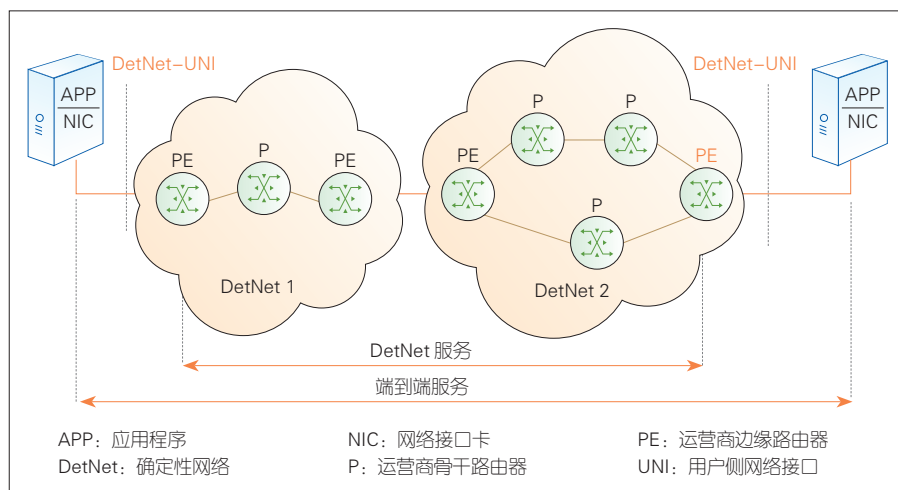
确定性网络用户侧网络接口（UNI）是基于分组的参考点，终端和 PE 通过分组网络提供连接，如图 3 所示。确定性网络 UNI 具有多种功能，例如，它可以将特定的联网技术专用封装添加到确定性网络流中，提供与预留相关的连接可用性状态，为终端系统提供同步服务。



▲图 1 数据面层次模型



▲图 2 确定性网络的使能网络组件



▲图3 确定性网络参考架构

1.4 确定性 QoS 的实现机制

1) 拥塞保护机制。该机制通过对确定性网络流所经过的路径进行预留资源来实现。预留的资源可能是缓存空间、链路带宽等。拥塞保护能极大地减少甚至完全消除网络中因输出报文拥塞造成的报文丢失，但它只作用于限定了最大报文大小和传输速率的确定性业务流。拥塞保护牵涉到确定性网络 QoS 的两个需求：延迟和丢包。鉴于确定性网络节点的缓冲区是有限的，拥塞保护必然导致最大的端到端延迟。缓冲区拥塞也是对丢包影响较大的因素。

2) 服务保护机制。除了拥塞，随机媒体错误和设备失效对丢包影响也很大。确定性网络使用报文复制和消除机制来解决这类丢包问题，从而实现服务保护。这种机制有时需要将确定性业务流重新编码并分发到多条路径上，使某条或某些路径的失效不会导致任何报文丢失。

3) 显式路由。显式路由通常通过特定的协议或者集中控制单元，根据确定性业务特性及网络约束条件计算出最佳确定性路径。这些确定性路径通常不会因路由或桥接协议的收敛而发生改变。

这 3 种机制可以独立或组合应用（有 8 种可能的组合），例如，在 IEEE 802.1CB 中采用显式路由和服务保护来实现无缝冗余机制。显式路由通过限制网络的物理拓扑为一个环来实现。顺序化、复制和副本消除是通过在以太网帧的头部或尾部加上报文标签来实现的。IEEE 802.1Qat、IEEE 802.1Qca 可以提供拥塞保护。只要网络不失效，就可使用流预留协议（SPR）或者路径控制与预留协议（PCR）在每个交换节点上做整形。如果将 3 种机制结合起来使用则低时延业务的可靠性可以获得最大程度的保障。

确定性网络通过在确定性网络流路径的每一跳预留带宽和缓冲区资源，实现拥塞保护和有限的传送延迟，然而在穿越多供应商网络时，预留本身是不够的。这是因为一个系统中的时延变化，会导致下一跳系统需要额外的缓冲空间，从而增加最坏情况下每跳的延迟。

标准排队和传输选择算法允许中央控制器计算每个传输节点对端到端延迟的贡献，和每个传输节点中每个增量确定性网络流所需的缓冲区空间量。IEEE 802 已经规定（并且正在制订）一组排队、整形和调度算法，使每个

传输节点（网桥或路由器）和 / 或中央控制器能够计算这些值。这些算法包括基于信用的整形器^[2]、基于时间同步的时间门控队列调度^[2]、基于时间同步的双重（或三重）循环队列转发调度^[2]，以及基于优先级抢占的传输机制^[2-3]。除了分组抢占技术外，它们都是可以应用到其他非以太网的媒介上。

2 应用场景举例分析

2.1 交通领域

2.1.1 汽车车内控制及娱乐通信网络

汽车领域的应用包括汽车内部的音视频和总线，也包括自动驾驶网络。前者是应用以太网局域网技术，后者是应用 IP 互联网技术。

汽车控制网络的主要特点有：高度工程化、固定拓扑、物理规模很小、通常在 30 m 五跳之内，但是会有很多端口（可能达到 100 个设备）。网络需要连接控制器、传感器、驾驶辅助视频、雷达以及娱乐用的影音，还需要通过网关连接控制器的局域网技术（CAN）、FlexRay、面向媒体的系统传输（MOST）等。

传统车载娱乐网络的主要技术是 MOST。MOST 技术的专有性质妨碍了自身应用。如果没有更多的开放性和更高的承受能力，MOST 技术可能会让位给以太网。

汽车控制网络应用需要确定性的极小时延，例如备用或驾驶辅助相机。如果使用 100 MB 的物理层（PHY），那么每跳的时延需要小于 20 μ s。

车上控制回路系统中的传感器和控制消息都是预先安排调度的。调度循环周期一般在 30 μ s~10 ms 之间，大部分情况是 125 μ s。控制消息一般在 128~256 字节之间。传感器需要的带

宽变量范围比较大。已有的应用所需带宽比较低,但是新型机器视觉应用在本地区域中需要更高的带宽。

IEEE 802.1Qav 排队和转发协议通过网络调度高优先级流量,确保低优先级数据不会干扰时间敏感内容。带宽保留在流启动之前保留整个网络的端到端带宽可用性,保证带宽直到明确释放。带宽预留可以预先配置为最小启动。为预期流量模式配置静态预留,从而保留所需网络资源的系统功能是汽车用例中的默认功能。

确定性网络协议作为更开放的标准,可用于下一代信息娱乐和驾驶员辅助系统的解决方案。

2.1.2 轨道交通控制及娱乐通信网络^[4]

随着轨道交通的发展,自动化、安全、舒适、娱乐需求的提升对轨道交通的通信系统提出新的要求。轨道交通中通信系统提升来自两个驱动力:一个是旅客信息系统及旅客外部网络接入的需求,二是机车自动化的需求。机车自动化又分为机车控制和机车操作维护。表1为轨道交通各种信号的网络需求。

这些参数中,机车控制信号优先级最高,机车操作维护次之,旅客信息服务优先级最低。实际上,高优先级的控制信号会被低优先级的娱乐、上网信息淹没。当前的解决方案是各种信号分开采用专用系统,但这将会

带来系统割裂、价格高昂、维护困难等问题。

能够在一个统一、标准的通信系统内共存,又能够满足各种信号的需求,是最优的解决方案。确定性网络技术正是这样的解决方案。

2.2 工业领域

2.2.1 电力传输与保护系统

电力设施部署依赖于下层网络的高可用性和行为确定性。在电力传输中,传输保护是一个非常重要的需求。电力保护包括操作者、电力设备的保护,以及电网的稳定性和频率的保持。如果出现错误,将会对操作者、电力设备和电网本身造成损害并导致断电。通信链路结合保护中继,可在最短时间内发出命令信号,以切断高压线路上的错误部分。

当前,电力传输与保护还依赖于在复杂环境下采用的技术,这包括时分复用(TDM)网络技术和一些应用特定网络技术。这种网络环境无法将OT和IT集成到同一个网络中,反而会产生信息孤岛。未来电力设施会向集成的基于开放和标准化的IP基础设施发展。

2.2.2 楼宇自动化系统

在典型的楼宇自动化系统(BAS)架构中,管理网络采用基于IP的通信

协议。在中大型楼宇中,管理系统部署在楼宇中;对于小型办公室或住宅,管理系统布放在远程,以节约成本。现场网络主要采用非IP的通信协议。本地控制器(LC)连接几十或上百个使用“现场协议”的设备,这些设备包括环境监控器、火源探测器、反馈控制等。LC一般是一个可编程逻辑控制器,负责测量设备状态,以提供信息给楼宇管理服务器或人机接口,还负责发送控制指令给设备(单方面的,或作为控制环路的反馈)。BAS中的“现场协议”五花八门,有多种介质接入控制(MAC)/PHY模块和接口,这导致BAS比较昂贵,存在很多厂商锁定的管理应用。

管理网络通常是尽力而为的,但是现场网络采用的是非IP的设计,导致无法互通。现场网络有特定的时间同步、定时等要求,未来BAS将可以提供更复杂和精确的管理控制,而且楼宇网络能够连接到企业网、家庭网和互联网。融合的网络可以代替现场网络中的现有技术。新的BAS网络架构和技术应该要保证低通信时延、低抖动、“6个9”的可靠性、容灾备份以及设备和网管之间的鉴权和认证。

2.2.3 专业音频和视频传输网络

专业音频和视频行业包括音乐和电影内容创作、广播、电影制作、现场声音、大型场馆的公共广播。这些

▼ 表1 轨道交通的网络通信需求

服务	上下行	每节点带宽需求 / (Mbit/s)	优先级	端到端时延	可靠性	安全	数据完整性	每车厢节点数	总带宽需求 / (Mbit/s)
控制信号	下行,上行	<1	最高 (=1)	<100 ms	>99.999%	最高	强制	2	2
实时视频监控	上行	>4	高 (=2)	<500 ms	>99.99%	高	推荐	20	80
乘客信息系统	下行	<1	低 (=4)	<1 s	<99.99%	中	-	6	3
紧急语音通知	下行,上行	<1	高 (=2)	<200 ms	>99.99%	高	-	8	6
乘客上网	下行	≥ 0.2	低 (=4)	<10 s	<99.9%	中	-	500	100
机车检测信号	上行	>0.1	中 (=3)	<1 s	<99.99%	高	推荐	50	5

行业已经将音频和视频信号从模拟转换为数字，但是数字互联系统主要是点对点的，每条链路上只有一个（或少量）信号与专用硬件相互连接。专业音频和视频的确定性网络的典型用例包括不中断的流播放、同步的流播放和声音增强。

如今专业音频和视频行业正急需基于数据包的基础设施，其应用程序可以基于 IEEE 802.1 TSN 标准，创建和传输确定性的流，但是不能通过 IP 路由，因此不能有效地分布到更广阔的区域（例如跨越广阔地理区域的广播事件）。如果这样的局域网流可以跨越 IP 路由网进行连接，则可以为专业音频和视频应用提供更为灵活的网络解决方案。

2.3 电信领域

3GPP 定义的典型蜂窝网络架构包括前传网、中传网和回传网 3 个网段。前传网连接基站处理单元（BBU）到远端射频头（RRH），中传网将基站互连起来，回传网将无线基站连接到网络控制器或网关。为了适应移动互联网和物联网的业务场景，低时延、高可靠是 5G 的 4 大关键技术特征之一。

留给前传网络的总体传输时间受限于基带处理无线帧后的可用时间。对基于分组的传输，分配的传输时间需要供天线与基带处理单元之间所有节点和缓存加上线路延迟使用。在当今的前传网网络技术中，队列、调度和发送组件成为主要因素，链路时延反而不是主要因素了，因为前传网链路相对来说很短。

一般来说，前传距离是一个给定参数，因为 RRH 和 BBU 通常放置在预定地点。然而，一个 RRH 可能会增加更多的天线，来提升多输入多输出（MIMO）容量或支持大规模 MIMO。这意味着增加共享相同的前传链路的

前传流量。确定性网络可以控制前传链路的带宽分配和流量调度，并提供足够的缓冲区来减少丢包率。

对于中传网，时延约束主要受站点之间的无线功能驱动，例如协同多点处理（CoMP）。CoMP 的设计原则是将当前一个蜂窝向多个用户终端（UE）发射的模式，扩展为通过基站站间的协作把多个蜂窝向多个 UE 发射的模式。CoMP 的“中传时延”和“信道状态信息（CSI）报告和精度”是两个时延敏感的性能参数。CoMP 的基本特征是在演进型基站（eNB）之间互访信令，所以中传网的时延是 CoMP 性能最主要的限制。站点间的 CoMP 是 5G 的关键需求。

确定性网络技术通过控制和减少队列、调度和传输操作所需要的时间，从而为链路传输留出更多时间以支持更长的传输距离。通过提供与尽力而为流量的隔离，确定性网络技术还可以满足 5G 传输网的不同网络切片差异化的性能需求。

中传和回传网络已经向着支持精

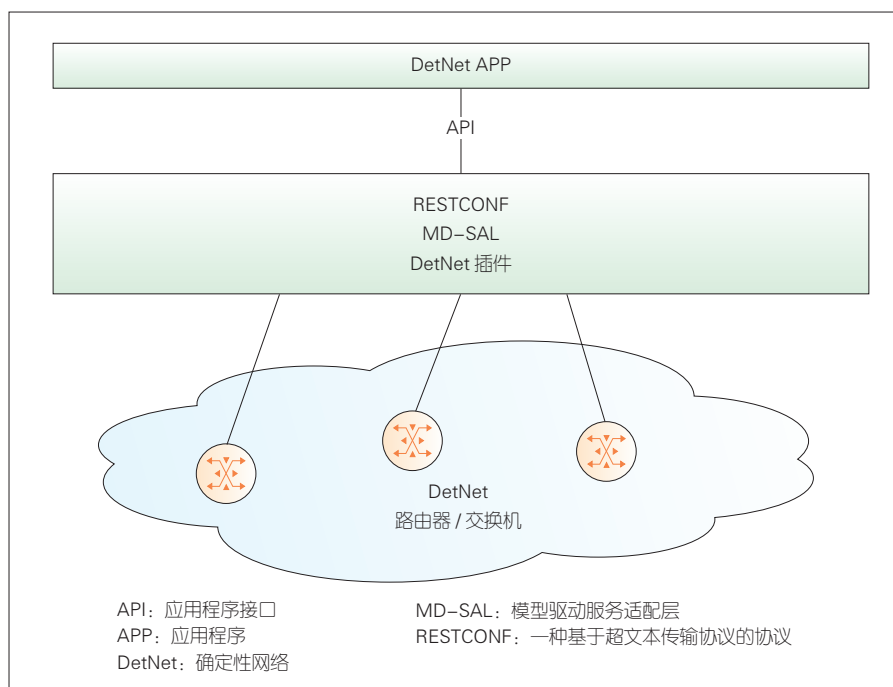
确时间同步的传输网络发展。传输网络本身为了满足带宽和成本的需求，事实上已经过渡到全 IP 的基于分组的网络，因此如此实现高精度的时钟分发已经成为一个挑战。

3 集中式确定性网络原型系统

如图 4 所示，典型的集中式确定性网络原型系统包括 3 个部分：确定性网络（DetNet）应用程序（APP）、DetNet 控制器、DetNet 路由器/交换机。

DetNet APP 向用户提供可视化界面，以方便用户对网络进行配置；DetNet 控制器基于 OpenDayLight，提供 DetNet 管理插件，主要负责拓扑收集、业务部署、路径计算、资源预留、操作维护管理（OAM）等工作，并通过网络配置（NETCONF）/边界网关协议-链路状态（BGP-LS）等南向接口与设备侧进行交互；DetNet 路由器/交换机执行转发面的工作，根据控制器的指令进行转发。

控制器的主要功能模块如图 5 所示。第 1 层静态处理包括 4 大功能模



▲图 4 确定性网络原型系统架构

块：“拓扑管理”模块负责拓扑收集，并获取相关链路属性（链路度量值、带宽、时延等）以及各网络设备类型和能力，提供北向接口，支持手动配置拓扑和链路属性，存储拓扑信息；“流特征管理”为动态处理单元提供确定性业务数据支撑；“域划分”模块划分 TSN 域或 Detnet 域等，并进行各域的特性数据管理；“时间同步”模块提供北向接口用于时间同步参数配置，调用南向接口插件下发配置至网络设备。

第2层次动态处理包括两大功能：“路径计算”模块根据拓扑以及业务约束条件（带宽和时延），计算满足需求的路径，供端到端业务整合模块使用；“端到端业务整合”模块调用路径计算单元（PCE）模块获得业务路径，然后进行端到端业务的部署——根据路径形成转发表来处理 TSN 域内业务，利用传输标签和业务标签的形成，以及在边缘节点和中继节点上的流映射配置来处理 DetNet 域内业务。

第3层执行层包括4大功能：“门控操作”模块根据带宽和业务路径，规划对应队列的门操作序列，并调用 SB 插件模块向网络设备下发门操作列表，确保业务的带宽和有边界时延；“带宽调整”模块用于确定性业务的可用带宽管理和实际带宽预留工作；“域间流映射”根据端到端业务整合模块指令，对 Detnet/TSN 域间业务进行封装、流映射规则生成；“队列调整”根据端到端业务整合模块指令，生成各节点对应出端口队列过滤、入队模板生成。

确定性网络系统要想实现“确定性”的 QoS 目标，最终还须依赖转发面设备的芯片能力。

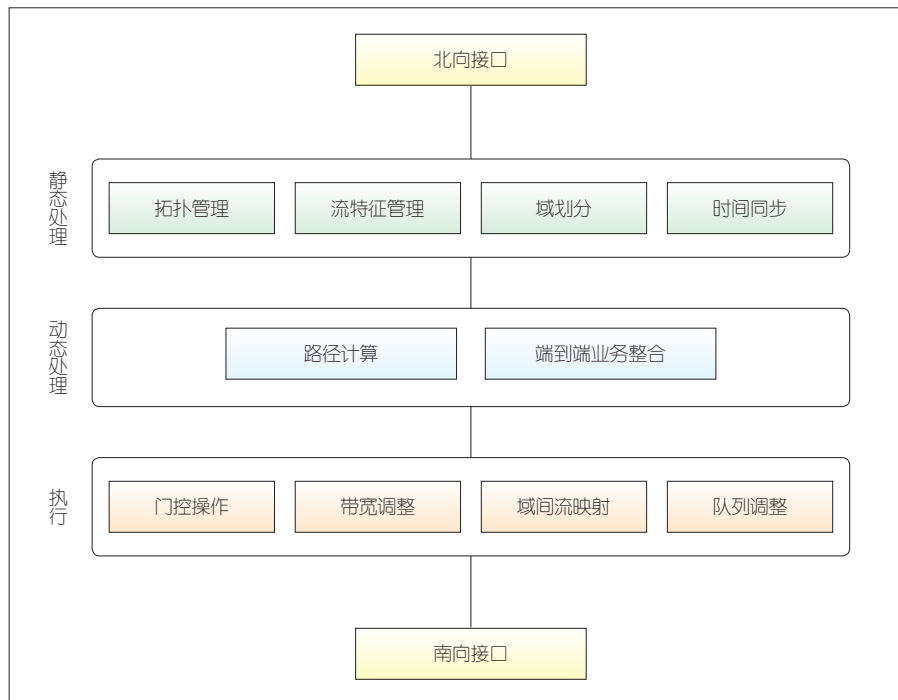
4 结束语

确定性网络应用场景的共性需求，是低且可控的端到端时延保证、可控的时延抖动、极低的丢包率，以及网络规模的动态伸缩。在此基础上，端到端分发协议、标准化的转发行为、

软件定义网络（SDN）集中管理技术、标准化的流信息模型、多时钟域的不同步等还要继续发展，才可能被应用在大规模网络中。多种多样的 OT 网络，会逐步采用 IT 网络所用的通用及泛在网络技术，从而实现 OT 和 IT 基础架构的互通和集成，以提升网络管理效率和路由灵活性，降低技术壁垒和网络成本。

参考文献

- [1] IETF. Deterministic networking architecture: RFC 8655 [S]. 2017
- [2] IEEE. IEEE standard for local and metropolitan area networks, bridges and bridged networks: IEEE 802.1Q [S]. 2018
- [3] IEEE. Specification and management parameters for interspersing express traffic: IEEE 802.3Br [S]. 2016
- [4] 中国信息通信研究院. 车联网白皮书 [R]. 2017



▲图5 确定性网络控制器软件模块

作者简介



魏月华，中兴通讯股份有限公司承载网标准预研总工；从事以太网、IP 路由、云计算数据中心网络、SDN 等技术和标准研究；拥有 15 年以上数据网络产品研发、设计及新技术预研经验；发表论文 3 篇，获授权专利 40 余项。



喻敬海，中兴通讯股份有限公司算法标准部技术预研项目经理；从事确定性业务承载技术、以太网、IP 路由、云计算数据中心网络、SDN 等技术和标准研究；拥有 20 年以上数据网络产品研发、设计及新技术预研经验；发表论文 6 篇，申请发明专利 100 余项。



罗鉴，中兴通讯股份有限公司有线研究院总工；从事数据产品研发、系统设计、网络架构和标准研究以及技术规划等工作。