面向低轨卫星与公有云融合网络的遥测系统研究

张 妍[1]，潘 恬[1]，郑 奘[1]，吴慧文[2]，赵晓红[3]，陈平平[1]

（1. 紫金山实验室，江苏 南京 211111；
2. 国网山东省电力公司青岛供电公司，山东 青岛 266071；
3. 国网山东省电力公司电力科学研究院，山东 济南 250004）

摘要：随着低轨卫星网络与公有云平台的深度融合，跨域、跨层的网络运维面临严峻挑战，租户业务网络难以进行快速、精准测量。为此，提出一种面向低轨卫星与公有云融合网络的遥测系统。该系统基于带内网络遥测（INT），提出端到端跨域遥测模式，实现了多VXLAN隧道场景下遥测数据的连续采集和跨动态星地链路的报告回传；通过Overlay层逐跳回传和Underlay层逐跳回传的测量模式，支持从虚拟网络到物理网络的逐层故障定位；利用实时INT信息驱动动态重路由，实现链路故障的快速自愈。实验结果表明，该系统能以低于5%的时间开销实现全链路时延精准监测，并有效定位跨域链路故障。

关键词：低轨卫星网络；公有云；带内网络遥测；故障定位

中图分类号：TP393

文献标志码：A

doi：10.11959/j.issn.2096−8930.tdyt2025086

Research on the telemetry system for LEO satellite and public cloud converged networks

Zhang Yan[1], Pan Tian[1], Zheng Yan[1], Wu Huiwen[2], Zhao Xiaohong[3], Chen Pingping[1]

1. Purple Mountain Laboratories, Nanjing 211111, China
2. State Grid Shandong Electric Power Company Qingdao Power Supply Company, Qingdao 266071, China
3. State Grid Shandong Electric Power Company Electric Power Research Institute, Jinan 250004, China

Abstract: With the deep integration of low earth orbit (LEO) satellite networks and public cloud platforms, cross-domain and cross-layer network operation and maintenance face severe challenges, which make it difficult to achieve rapid and accurate measurement of tenant service networks. To address this issue, a telemetry system for LEO satellite and public cloud converged networks was proposed. Based on in-band network telemetry (INT), the system introduced an end-to-end cross-domain telemetry mode, which enabled continuous collection of telemetry data in multi-VXLAN tunnel scenarios and reliable report feedback across dynamic satellite-to-ground links. Through the hop-by-hop feedback measurement modes at both the overlay and underlay layers, it supported layered fault localization from virtual to physical networks. Dynamic rerouting driven by real-time INT information was utilized to achieve rapid recovery from link failures. Experimental results show that the system can achieve precise monitoring of full-link latency with less than 5% time cost and effectively localize cross-domain link faults.

Key words: LEO satellite network, public cloud, in-band network telemetry, fault localization

0 引言

近年来，低轨卫星网络的快速发展与云计算技术的广泛应用，正驱动着网络基础设施向空天地一体化深度演进。低轨卫星网络凭借其广域覆盖与较低传输时延的优势，成为弥补地面网络覆盖盲区、实现全球连接的关

收稿日期：2025−12−10；**修回日期：**2026−08−13

通信作者：潘恬，pantian@pmlabs.com.cn

基金项目：国家电网有限公司科技攻关登高行动计划项目（No.5700-202455369A-3-1-DG）

Foundation Item：State Grid Corporation of China Science and Technology Breakthrough and Climbing Action Plan Project (No.5700-202455369A-3-1-DG)

键组成部分。与此同时，主流公有云服务商（如AWS、Azure）[1]积极推出卫星云服务，推动地面云基础设施与太空业务深度融合，将地面虚拟化计算与存储能力通过卫星链路延伸至全球任意角落，形成了“卫星—公有云”一体化网络架构。

然而，这种异构、跨域、高动态的融合网络给网络运维与故障诊断带来了严峻挑战。在公有云侧，多租户的虚拟化覆盖网络通过虚拟扩展局域网（virtual eXtensible LAN, VXLAN）[2]等多层隧道封装，在实现灵活性与隔离性的同时，也遮蔽了底层物理网络的拓扑与状态，使得数据包的真实转发路径成为“黑盒”；在卫星网络侧，卫星节点的高速运动导致网络拓扑持续变化，星间/星地链路质量动态波动、链路切换频繁，使得遥测信息回传困难，难以维持全链路测量的连续性。同时卫星节点星上计算、存储资源与星地链路带宽均为稀缺资源，无法承载大规模、高开销的测量任务。带内网络遥测（in-band network telemetry，INT）[3]可以实现细粒度遥测，收集设备内部状态信息（如排队时延、拥塞程度、丢包计数等），具有多种测量模式以应对不同场景，但其标准工作模式均针对静态地面数据中心设计，若直接部署于融合网络场景，将会引入显著的带宽与设备处理开销，无法适配高动态拓扑、多层隧道封装的特殊环境。

具体而言，面向低轨卫星与公有云融合场景的测量系统需攻克以下核心难题：其一是虚拟与物理网络的关联感知与故障定位，快速协同虚拟网络与物理卫星网络的测量数据，准确判定故障位置，实现跨层根因分析；其二是多次隧道封装下的遥测数据继承，业务流历经多次VXLAN隧道的封装与解封装，需确保测量上下文（INT元数据）在隧道端点不被丢弃，维持全链路测量的连续性；其三是跨动态星地链路的可靠遥测回传，在卫星与地面站间歇连通的动态链路上，确保INT报告能够高效、可靠地回传至地面分析器；其四是多租户环境下的开销控制，在保证细粒度测量能力的同时，将遥测开销（带宽、CPU）控制在可接受范围内。

为解决上述问题，本文设计了一种面向低轨卫星与公有云融合网络的遥测系统，基于INT提出了一种支持动态星地链路的端到端遥测与回传机制。针对跨层感知与定位难题，提出了端到端、Overlay层逐跳、Underlay层逐跳3种测量模式，通过“端到端监测→Overlay段定位→Underlay段精确定位”的协同流程，实现了从租户业务视角到物理链路的逐层精准故障定位，破解了网络跨层状态不可见的困境；针对多隧道场景下INT测量数据的继承难题，提出一种基于VXLAN报头保留位的扩展方案，通过定义VXLAN报头中的继承标志位，在不改变标准报文格式的前提下，确保该扩展兼容现有商用网络设备，且不会导致非支持设备丢弃报文，确保INT测量数据在“上星”与“下地”的多次隧道操作中无缝继承，实现端到端连续采集；针对动态链路回传难题，结合卫星星历的可预测性，设计适配断续连通特性的可靠报告回传机制，保障星上INT数据高效、可靠地回传至地面分析器；针对多租户开销管控难题，采用“默认低开销端到端测量+故障时按需触发逐跳报告”的混合机制，在保障细粒度测量能力的同时，将遥测开销控制在可接受范围内。在此基础上，系统构建了基于实时INT遥测信息的智能重路由策略，实现链路故障的快速自愈与服务质量保障。

本文的主要贡献概括如下。

1) 提出面向低轨卫星与公有云融合网络的遥测系统，设计了端到端、Overlay层逐跳、Underlay层逐跳的分层协同多模式测量机制，实现了物理网络与虚拟网络的关联感知和交互式故障定位，解决了融合网络跨层状态不可见、故障无法精准定位的问题。

2) 设计轻量级跨域遥测数据继承机制，提出了基于VXLAN报头继承标志位的INT元数据跨隧道保留方法，结合星历感知的动态星地链路可靠回传方法，解决了多次隧道封装导致的测量信息丢失问题，实现了端到端连续遥测和跨动态星地链路的可靠回传。

3) 构建基于实时INT遥测数据的重路由机制，利用INT提供的细粒度网络状态，结合动态等价多路径路由算法，通过权重自适应调整实现智能路径选择与流量调度。

4) 构建卫星与公有云融合的仿真实验场景，实验结果表明，本文所提系统能以低于5%的额外时间开销实现全链路精准测量，可快速完成网络异常感知与故障定位。

1 背景与挑战

1.1 研究背景

（1）低轨卫星网络

低轨卫星网络可以实现全球范围内的低时延和高吞吐量通信，正成为全球通信基础设施的重要组成部分。与传统地面网络相比，低轨卫星网络呈现出其独有特征：其一是广域覆盖与灵活接入，能够为海洋、沙漠等偏远地区提供无缝的网络服务；其二是拓扑高动态性，卫星相对于地面的高速运动导致网络链路与端到端路径持续、可预测地变化，链路通断、拓扑变化的时间窗口均可通过卫星

星历实现精准预测；其三是异构多层结构，网络包含星间链路与星地链路，共同构成一个复杂的空间信息承载网。这些特征在拓展全球服务能力的同时，也给网络稳定运行、性能保障与运维管控带来了严峻挑战：星间链路频繁切换导致拓扑持续动态变化，星上资源受限，复杂的星间/星地链路环境，大幅提升了网络测量与运维的难度。此外，现有的一些卫星网络测量方案[4-5]仅关注星上网络测量，对星地网络传输少有涉猎。

（2）地面公有云网络

以虚拟化为核心的公有云技术已成为现代计算的基础。公有云通过覆盖网络虚拟化技术，在物理网络之上为多租户构建了逻辑隔离的虚拟网络，可以同时为数以百万计的租户提供服务。VXLAN 等技术通过隧道封装，实现了跨物理子网的虚拟机逻辑连通，构成了云数据中心的覆盖网络。覆盖网络比底层物理网络更复杂，其服务质量直接影响云厂商与租户达成的服务等级协议。在此架构中，公有云内部构建了虚拟私有云（virtual private cloud, VPC）[6]，虚拟交换机是虚拟机流量的入口，云网关则是实现跨子网、跨 VPC 乃至跨地域流量转发的核心枢纽。数据包在虚拟网络中被多次封装与解封装，使得租户乃至云平台运维人员都无法获知其真实的物理路径，极大地增加了故障诊断与性能优化的复杂度。一些传统的基于简单网络管理协议（simple network management protocol, SNMP）[7]的周期轮询式方法存在数据滞后、粒度粗糙的固有缺陷；一些方案集中于物理网络的测量[8-10]，仅通过信令探针监测底层链路质量，无法有效感知上层虚拟网络的业务状态，存在测量视野盲区。

（3）卫星与公有云融合网络

卫星与公有云融合的一体化网络架构，实现了虚拟机到卫星乃至跨卫星的虚拟机到虚拟机的直接通信，用户感知到的将不再是分离的“卫星通信服务”和“云计算服务”，而是一个统一的、全球覆盖的云平台。这种架构极大地减少了数据在地面的迂回，降低了传输时延，提升了整体性能，同时，可以分配卫星网络资源（如带宽、星上计算资源），实现真正的全球部署和流量调度。企业租户通过卫星网络接入公有云，要求业务具备高可靠性、低时延。然而，卫星网络与公有云的融合，也使得网络路径包含了从地面虚拟化设备到空间网络设备的复杂场景，这对融合网络的测量提出了新的需求：运维需具备租户业务视角的端到端细粒度测量能力，故障发生时可快速完成地面/卫星段、Overlay/Underlay 层的精准定位，而现有方案均无法适配融合场景的特殊要求。目前一些数据中心测量方法（如 Pingmesh[11]、VNET Pingmesh[12] 等）缺少细粒度精准观测；一些研究关注虚拟与物理网络感知[13-15]，如阿里云的 Zoonet[13]，专为大规模多租户虚拟网络探测设计，实现了租户业务视角的测量，但未涉及卫星网络的高动态特性测量困境，无法满足卫星与公有云融合网络的测量需求。

（4）新型网络测量技术

INT 是一种新兴的数据面测量技术，根据 P4 官方发布的规范[16]，INT 主要存在 3 种模式：INT-MD（eMbed data）模式是最经典的工作模式，也是 INT 的默认模式，它将指令和 INT 元数据（INT metadata）写入数据包中，INT 源端（INT source）插入指令，每个 INT 传输节点（INT transit）将自己的元数据压入包内的元数据栈，进行逐跳收集，INT 接收端（INT sink）移除整个 INT 头部和元数据栈，并将数据发送给监控系统；INT-MX（eMbed instructions）模式仅嵌入指令，INT 源节点将 INT 指令嵌入数据包头部，所有 INT 节点（源端、传输节点、接收端）根据包中的指令，直接向监控系统发送元数据；INT-XD（eXport data）模式也被称为“明信片（Postcard）模式”，每个节点独立发送报告，它无须向业务数据包中插入 INT 头部和逐跳元数据栈，而是根据其观察列表（watchlist）匹配目标流量，从数据平面收集元数据，独立生成并向监控系统导出遥测报告。INT 技术虽突破了传统测量方法的粒度瓶颈，但其标准模式无法适配低轨卫星与公有云融合网络的高动态拓扑、多层隧道封装、超大规模多租户特性，无法解决前文所述的四大核心难题。

1.2　挑战与需求

卫星网络与公有云虚拟化网络的深度集成，有助于构建天地一体化云网，带来服务能力的飞跃，但也因其架构的复杂性，网络测量面临前所未有的复杂挑战。

（1）虚拟与物理网络的关联感知与故障定位问题

云网络由底层物理网络和上层虚拟网络组成，二者共同构建了可靠的网络环境，保障每个租户应用程序的正常运行。物理网络提供基本的网络连接能力，虚拟网络为租户提供高级网络服务，例如虚拟路由、网络地址空间隔离、租户跨地域通信等。由于物理网络和虚拟网络都会影响租户的网络性能，因此，理想情况下，在租户感知到网络性能下降之前，任何由物理网络或虚拟网络引起的网络故障都应当被检测、定位和解决。然而，虚拟网络和物理网络之间没有精确的拓扑映射，一个虚拟网络功能网元（如云网关、虚拟路由器）可能分布式

部署在多个物理设备上，一条端到端虚拟转发路径可映射为多条异构的物理转发路径。Overlay 网络与 Underlay 网络无法相互感知，底层的物理链路故障难以定位。因此，如何实现虚拟网络和物理网络的交互，快速协同覆盖网络与底层网络，利用遥测数据形成统一的故障诊断视图，实现逐层、逐跳的故障定位，是一个亟须解决的技术难题。

（2）多隧道场景下的遥测数据继承问题

在卫星云网的端到端路径中，数据包会经过虚拟交换机、网关、卫星等多类设备，跨越不同 VPC 或地域网络，并经历多次 VXLAN 隧道的封装与解封装。传统的网络遥测方案通常无法按租户粒度覆盖跨地域网络并提供端到端故障排除，因为每一次解封装都可能导致携带的 INT 元数据丢失，造成测量链路的断裂，破坏遥测的连续性。因此，低轨卫星与公有云融合网络的测量需要一种轻量且高效的机制，保存跨隧道的遥测上下文，在隧道端点处能够识别、保存并继承前一阶段的测量数据，避免端到端路径状态感知的中断。

（3）跨动态星地链路的可靠遥测回传问题

低轨卫星网络具有高动态拓扑、链路频繁切换和时变信道等特性。卫星与地面站之间存在持续的相对运动，单个地面站的卫星可视时间通常仅为数分钟。在低轨卫星与公有云融合网络的测量中，遥测数据包需要通过卫星网络段传输，卫星等网络节点也需要将测量报告回传至地面的控制与分析器，而卫星与地面站之间的可视窗口可能受限，造成链路断续连通，星地链路动态性制约了数据报文在星地之间的传输效率。因此，测量系统需要设计高效、可靠的回传与聚合机制，以保障报文的端到端转发与测量报告的有效回传，从而实现关键测量信息的无丢失传输与低时延回传。

（4）多租户环境下的遥测开销管控问题

在公有云环境中，物理网络可能包含数十万台服务器，而虚拟节点的数量比物理节点至少高出两个数量级。例如，一台物理服务器可虚拟化出数十至上百台虚拟机，这使得公有云供应商能够服务海量租户，因此，简单的遥测方案在这种超大规模云网络中可能产生巨大的遥测开销。此外，收集和处理探测包也需要庞大的带宽和计算成本，例如，直接进行 INT 操作（元数据插入、报告生成）会消耗宝贵的 CPU 与内存资源。所以，如何在不影响宿主服务器上其他虚拟机网络性能的前提下为特定租户提供精细化的测量服务，并能精确控制测量行为带来的性能开销，是测量方案的关键之一。

2　低轨卫星与公有云融合网络遥测系统

2.1　整体架构与测量模式

针对前文所述的四大挑战，本文设计了一套面向低轨卫星与公有云融合网络的遥测系统，其整体架构如图 1 所示。

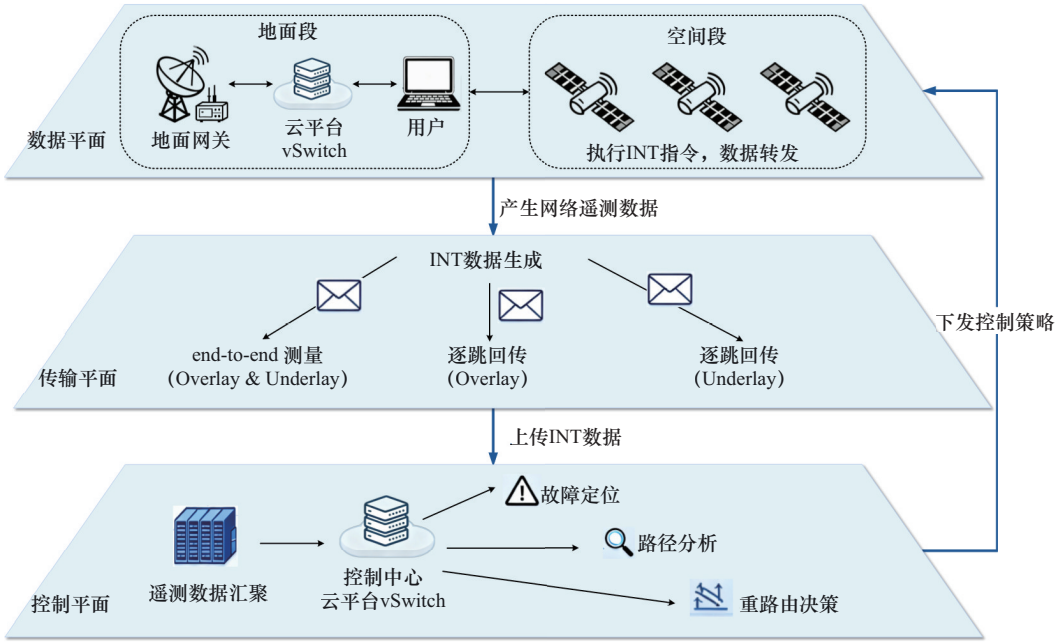


图 1　面向低轨卫星与公有云融合网络的遥测系统整体架构

数据平面由物理和虚拟设备构成，如虚拟交换机、地面网关、卫星节点等，负责网络流量转发与测量信息采集。其中，地面网关作为天地通信的枢纽，负责将数据上行接入卫星以及接收卫星下行数据。同时，它不仅作为VXLAN隧道的端点，进行封装与解封装，还作为INT的源端和终端，继承、插入并转发遥测数据；卫星节点分布在空间轨道，负责数据包的转发和INT指令的执行；用户侧的虚拟交换机是虚拟机流量的第一跳，负责处理地面云网络的数据流转发，并作为INT测量的起始点和最终报告收集点。同时，它还可以作为用户流量的控制器，根据遥测信息实时调控用户流量。

控制平面负责遥测数据的汇聚、分析与决策。在面向租户的遥测系统中，控制器可以是租户端的虚拟交换机，它接收来自所有数据平面节点的INT报告报文，并进行解析、关联与存储。控制平面基于实时和历史遥测数据，实时检测网络异常（如时延突增、丢包），并通过触发多层报告模式，快速定位故障点；结合卫星星历与网络状态，分析当前和预测未来的端到端路径性能，识别潜在瓶颈。根据路径分析与故障定位的结果，动态生成路由优化策略，并下发至数据平面执行。

传输平面并非由独立设备构成，而是作为一组核心功能内嵌于数据平面节点中，主要实现遥测数据的生成与传输。数据平面节点在转发数据包时，根据预设指令采集本地的设备内部状态信息（如设备ID、入口/出口时间戳、时延、队列占用率等）。传输平面负责执行3种核心测量模式，以适应不同的运维场景。测量模式如图2所示。

1) 端到端（end-to-end）模式：默认测量模式，用于日常性能监控。数据包在转发路径上逐跳采集INT元数据，实现端到端网络遥测。由图2可知，虚拟源端向目的端发送遥测报文，虚拟目的端向源端发送报告报文，汇报遥测数据，而遥测路径上的中间节点不需要向源端发送报告报文。

2) Overlay层逐跳（hop-by-hop Overlay）模式：当检测到性能异常时，首先触发此模式。采用Postcard方式，不修改原始数据包，而是将采集到的遥测数据封装成独立的报告报文单元。遥测路径上的所有Overlay节点（如虚拟节点a、b）向指定的INT Sink（如虚拟源端）回传遥测报告，用于将故障快速定位到某个Overlay网络段。

3) Underlay层逐跳（hop-by-hop Underlay）模式：在Overlay定位故障后，用于进一步精确定位。指定Underlay节点（如物理网络节点x、y）向指定的INT Sink（如虚拟源端）回传报告，以定位具体的物理链路或设备故障。

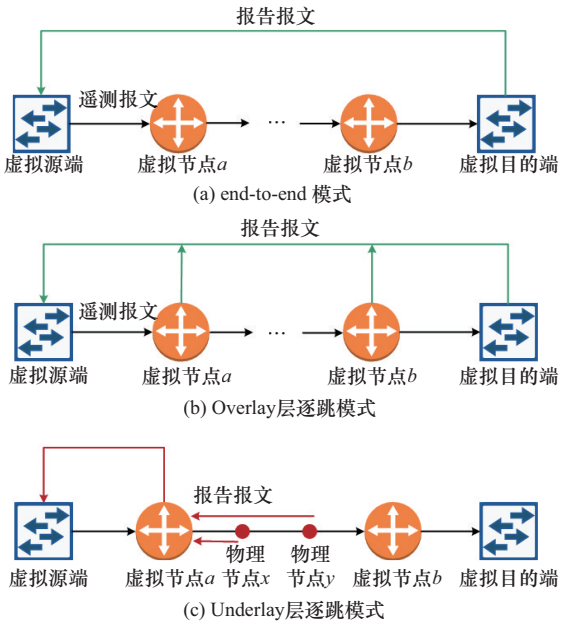


图2　测量模式

2.2　数据平面工作流程

在数据包转发过程中，vSwitch、云网关和卫星节点等设备负责基于其本地存储的转发表项进行查询，对数据包进行高速转发决策与隧道封装等，不同模式（如end-to-end模式和逐跳模式）下的查表过程如图3所示。低轨卫星与公有云融合网络的数据转发路径中至少包含以下重要的转发表。

1) VXLAN路由表：位于地面，根据VXLAN封装数据包的内部目的IP，找到正确的VPC。

2) 虚拟机—服务器映射表：位于地面，根据VXLAN封装数据包的内部目的IP，以及从VXLAN路由表获得的VPC，找到承载目标虚拟机的物理服务器地址。

3) VPC—地理信息映射表：位于地面云网关，根据目的VPC的VXLAN网络标识符（VXLAN network identifier, VNI）查询目的网关的地理信息（纬度和经度）。

4) 卫星—地理信息映射表：位于卫星上，根据卫星的ID查询卫星的实时地理位置信息（纬度和经度）。

5) 卫星路由表：位于卫星上，根据目的卫星的ID/目的IP查询下一跳节点。

2.2.1　端到端跨域遥测

图4展示了低轨卫星与公有云融合网络中数据平面测



注：Lat指纬度，Lon指经度。

图3　不同模式下的查表过程

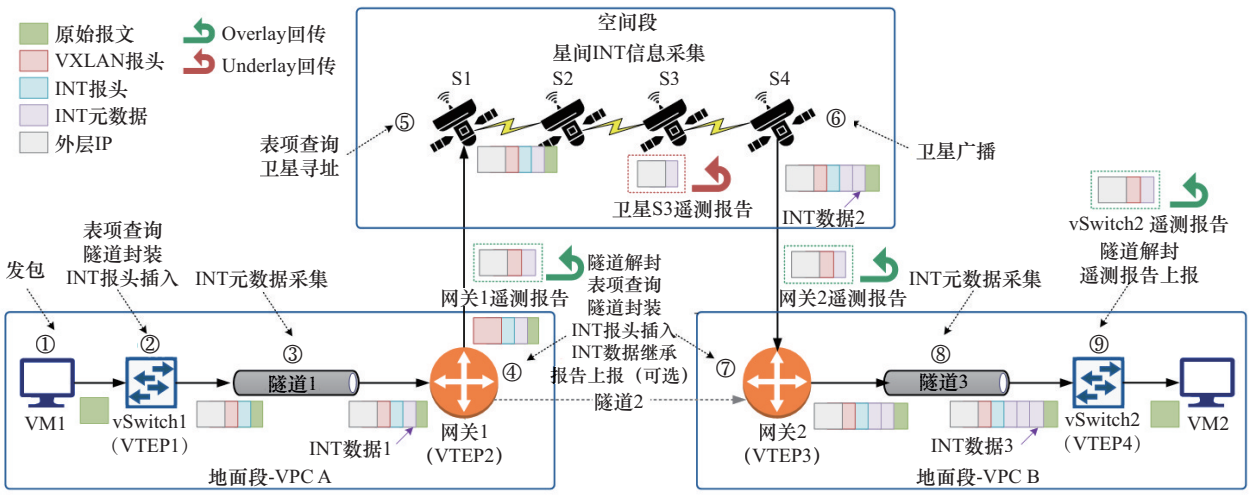


图4　数据平面测量信息采集与转发过程

量信息采集与转发过程。在端到端跨域网络遥测中，测量路径上的每个网络设备将INT元数据插入数据包，最终由接收端或指定节点生成遥测报告。以虚拟机1（VM1）向虚拟机2（VM2）发送报文为例，测量报文转

发与遥测信息采集的具体过程如下（其中的查表过程可见图3中end-to-end模式部分）。

1) VM1发送原始数据包。

2) vSwitch1接收数据包，使用VXLAN隧道（隧道1）封装。vSwitch1作为INT Source将INT报头嵌入VXLAN报头与被封装的原始数据之间作为VXLAN载荷。INT报头中设置了采集指令，用于指定中间节点插入的元数据类别，如设备ID、时间戳、输入/输出端口、队列深度等。vSwitch1将外层目的IP设置为网关1的IP地址并转发。

3) 数据包通过隧道1转发至网关1，同时沿路采集设备内部状态信息，并将采集到的INT数据存储在INT报头之后的INT元数据堆栈之中。

4) 数据包到达网关1后，网关1需要解封装原来的隧道，并重新封装报文。此过程中，网关1通过查询本地维护的VXLAN路由表获得目的VPC（即VPC B），确定VPC B需要通过卫星网络远程访问，随后通过查询VPC—地理信息映射表，得到目的网关（即网关2）的地理信息，并将其嵌入VXLAN报头，完成了跨域路径的计算与下一跳的选择。同样，将INT报头嵌入VXLAN报头与被封装的原始数据之间，上一阶段的测量数据（INT数据1）一同被继承，插入INT报头之后。随后数据包经上行链路转发至网关1上方的卫星S1。

5) 接入卫星S1通过比较VXLAN报头中携带的地理信息，查询其星上存储的卫星—地理信息映射表，将地理坐标映射为最优接入卫星ID，从而确定目的卫星S4，并将隧道的外层目的IP重写为S4的IP地址。随后，S1查询卫星路由表，底层卫星网络根据外层目的IP将数据包传递给S4，同时沿路采集遥测信息，并将采集到的遥测数据（INT数据2）存储在数据包之中。

6) S4接收到数据包后，经下行链路将数据包广播到其覆盖的地面网关（即网关2）。

7) 数据包到达网关2后，网关2需要解封原来的隧道，并使用VXLAN隧道（隧道3）重新封装报文。网关2通过查询VXLAN路由表和虚拟机—服务器映射表，获得vSwitch2的地址，并将外层目的IP设置为vSwitch2的地址。同样地，嵌入INT报头，并继承遥测数据INT数据1和INT数据2。

8) 数据包通过隧道3转发到vSwitch2，同时沿路采集设备内部状态信息，并将采集到的INT元数据（INT数据3）存储在数据包中。

9) 数据包到达vSwitch2后，vSwitch2拆除隧道，将原始报文发送到VM2。作为INT Sink，vSwitch2将遥测数据报告（含INT数据1、2、3）传输至始端交换机vSwitch1进行后续编排。

因此，基于采集的端到端遥测信息，可以初步建立Overlay至Underlay网络的跨层关联。例如，当网络发生拥塞时，可以通过上层租户对（VM IP对）查询底层路径的详细遥测数据，如设备ID、端口ID、端口时间戳等。

值得注意的是，在端到端测量过程中，数据包被多次封装与解封装，其携带的INT测量信息可能被丢弃。因此，将VXLAN报头保留位中的一位定义为继承标志位（FLAG），用于指示当前INT信息是否需要被继承，这是对标准VXLAN协议的轻量扩展。当FLAG被设为1时，表示当前INT信息需要被继承。Overlay网络节点根据数据包的携带信息识别出其为INT数据包后，可以设置此标志位。实际上，VXLAN头部设计之初就包含若干保留位及可选字段（Option），以供厂商自定义功能，业界主流云厂商（如AWS、阿里云）在其Overlay网络中均有类似实践。更重要的是，在卫星—云融合网络中，VXLAN报文作为隧道协议，其内部字段对于承载它的Underlay网络（包括传统IP路由器和卫星节点）是完全透明的。Underlay设备仅根据外层IP头进行转发，不会解析或修改VXLAN头的内部内容。此外，本系统VTEP由云厂商统一运维，可通过控制平面协商确保仅在双方支持时启用该标志位，保证与传统设备的兼容性。鉴于此，即使中间路径存在不支持该标志位的传统商用设备，它们也只会将其作为普通VXLAN流量透传，不会因此丢弃报文。只有当报文到达同样支持本方案的VTEP（如可编程云网关或虚拟交换机）时，FLAG位才会被识别并执行遥测数据继承操作。因此，标志位定义方案在现有商业网络设备体系中具备良好的兼容性与可行性，能够支撑自成一体的卫星—云融合网络遥测需求。

2.2.2　Overlay层逐跳报告回传

当网络链路中出现故障丢包时，为了定位故障点，需要启用逐跳模式。首先启动hop-by-hop Overlay模式定位故障所在的Overlay网络段。遥测报告回传过程如图5所示，其中，在hop-by-hop Overlay模式下，INT Sink为vSwitch1，INT Source为全链路中的Overlay节点（例如，本例中的网关1、网关2和vSwitch2），均需要向vSwitch1发送携带INT测量数据的报告报文。报告回传的转发过程与end-to-end模式下的转发过程类似，但是不再需要沿路逐跳封装遥测信息，并且传输方向相反。下面以vSwitch2节点报告回传为例，详细介绍其转发过程。

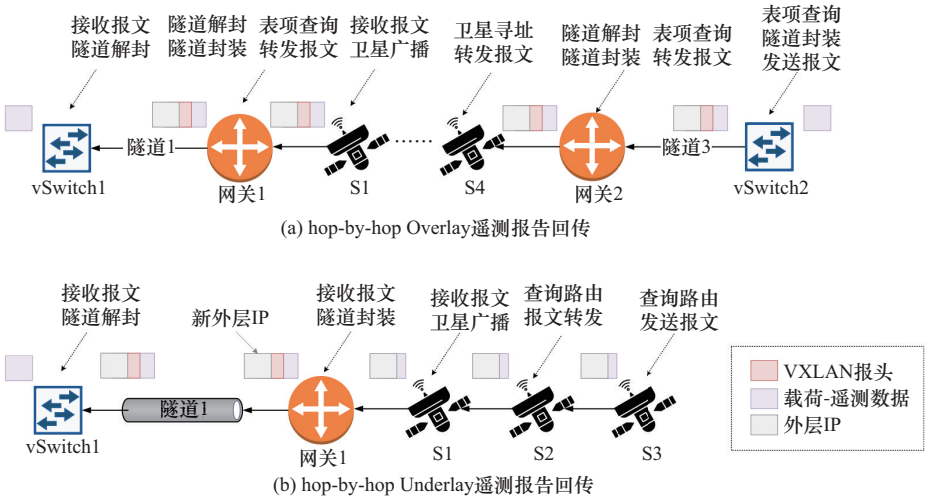


图 5　遥测报告回传过程

1) vSwitch2 发送报告数据包，载荷为 INT 遥测数据。vSwitch2 作为 INT Source，通过查询本地虚拟机—服务器映射表，确定去往目的 INT Sink（vSwitch1）的路径需要经过网关 2，因此将报告报文封装进隧道 3，将外层目的 IP 设置为网关 2 的 IP 地址并转发。

2) 数据包通过隧道 3 转发至网关 2。网关 2 解封装原来的隧道，并重新封装报文。同样地，网关 2 通过查询 VXLAN 路由表获得目的 VPC（即 VPC A），随后通过查询 VPC—地理信息映射表，得到目的网关（即网关 1）的地理信息，并将其嵌入 VXLAN 报头。随后数据包通过上行链路转发上星。

3) 接入卫星 S4 通过比较 VXLAN 报头中携带的地理信息，查询卫星—地理信息映射表，识别出最接近网关 1 的卫星（即 S1），将隧道的外层目的 IP 重写为 S1 的 IP 地址，随后转发数据包。

4) S1 接收到数据包后，将数据包广播到其覆盖的地面网关 1。

5) 数据包到达网关 1 后，网关 1 解封装原来的隧道，使用 VXLAN 隧道（隧道 1）重新封装报文，并将外层目的 IP 设置为 vSwitch1 的地址。

6) 数据包通过隧道 1 转发到 vSwitch1，vSwitch1 拆除隧道，获得 INT 数据。

2.2.3　Underlay 层逐跳报告回传

为了进一步定位网络故障的具体底层位置，启动 hop-by-hop Underlay 模式。假设故障发生在低轨卫星网络段，即卫星网络段中的每个底层卫星节点（例如，本例中的 S1、S2、S3 和 S4）均需要向 vSwitch1 发送报告报文，以卫星节点 S3 为例，报告回传过程如下。

1) 卫星节点 S3 生成报告报文，数据包构造为外层目的 IP 为 S1 的 IP 地址，内层目的 IP 为 vSwitch1 的 IP 地址，载荷为 INT 遥测数据。S3 查询卫星路由表，确定去往 S1 的路径（如 S3—S2—S1），将数据包发出至下一跳 S2。

2) S2 同样查询卫星路由表，并根据外层目的 IP 进行标准路由转发，不关心内层 IP 内容。

3) 数据包到达卫星节点 S1，S1 通过广播的方式将数据包传送至地面网关 1。

4) 网关 1 接收到数据包之后，使用 VXLAN 隧道（隧道 1）封装数据包，并通过查询 VXLAN 路由表和虚拟机—服务器映射表，根据内层 IP 获得远端 VTEP 为 vSwitch1，并将外层目的 IP 设置为 vSwitch1 的地址，查询去往 vSwitch1 的路由并发送。

5) vSwitch1 接收到数据包，拆除隧道，获得 INT 数据。

需要注意的是，在低轨卫星网络中，星间链路切换与星地链路中断可能会导致 INT 报告报文丢失，破坏测量上下文的连续性。然而，这两类事件均可通过星历精确预知：星地链路的可见窗口，星间链路的通断时刻。利用这一可预测性，控制平面利用星历预知星间链路切换时刻，以及对应时刻的星地链路（包括开始时间、持续时长），从而确定时间窗边界，在相应窗口内可视为静态拓扑。这段静态拓扑的维持时间通常可以支撑完成一次完整卫星网络测量。例如，对于星链（Starlink）[17]星座中的“壳层 3”（轨道高度 560 km，倾角 97.6°，6 个轨道面，每面 36 颗卫星），卫星过顶时间通常为 4~6 min，连续两次星间链路变化时间间隔从 1 min 到数分钟不等，

而一次卫星网络测量用时往往为秒级。

在每个窗口开始前，控制平面将针对该卫星的测量任务（如INT指令、需采集的元数据类型、报告目的地址）通过信令通道下发至卫星节点。卫星节点据此在窗口内执行数据面测量，并生成INT报告报文，从而实现跨星间与星地链路的可靠传输。

2.3　基于INT遥测信息的重路由机制

在低轨卫星与公有云融合网络中，网络路径的高动态性和异构性要求系统能够实时响应链路故障或性能劣化，以保障租户业务的服务质量。基于INT回传的遥测信息，可以通过分析实时网络状态数据，动态调整流量路径，实现负载均衡和故障规避等。基于INT遥测信息的重路由机制的核心在于利用INT提供的细粒度网络状态（如时延、丢包率、队列深度等），结合等价多路径路由（equal-cost multi-path routing, ECMP）方法，实现智能路径选择与流量调度。

2.3.1　重路由触发与决策

重路由的触发依赖控制平面对INT遥测数据的实时分析。控制平面汇聚来自数据平面节点的INT报告（包括end-to-end和逐跳模式），并提取以下关键指标：链路质量指标，如端到端时延、抖动、丢包率等；设备状态指标，如队列占用率、缓冲区溢出计数、端口错误率等；拓扑变化事件，如卫星链路切换、地面网关连接状态变化等。当满足以下条件时，系统触发重路由决策：检测到故障（如INT报告显示某条路径的丢包率持续高于阈值（如20%））或链路中断（如卫星节点失联）等；性能劣化，如端到端时延超过服务水平协议（service-level agreement, SLA）约定值，或抖动显著增加（如超过50 ms）；拥塞预警，如队列深度超过设备容量的70%，表明链路即将拥塞等。控制平面基于这些指标定位故障点或瓶颈链路，并计算替代路径。决策过程需考虑的因素包括：路径的实时质量评分（基于INT数据加权计算）、卫星星历预测的可视窗口与链路稳定性、租户SLA要求（如时延敏感型业务优先选择低时延路径）。

2.3.2　基于ECMP的动态负载均衡

传统ECMP在多条路径间静态分配流量，无法适应网络状态变化。本系统通过INT遥测增强ECMP，实现动态负载均衡。①路径状态监控，控制平面为每条ECMP路径维护健康状态表，包括当前时延、丢包率和可用带宽。状态表通过INT报告定期更新。②权重自适应调整，根据路径健康评分，动态调整ECMP哈希权重。健康路径获得更高权重，故障或拥塞路径权重降低甚至归零。权重计算公式为 $W_i = \dfrac{1}{\alpha D_i + \beta L_i + \gamma Q_i}$，其中，$D_i$为时延归一化值，$L_i$为丢包率归一化值，$Q_i$为队列深度归一化值，$\alpha$为时延权重系数、$\beta$为丢包权重系数、$\gamma$为队列深度权重系数，可根据业务类型调整。③哈希算法优化，在ECMP哈希算法中引入路径状态参数，确保新流量分配优先选择健康路径，同时保持现有流量的亲和性以避免乱序。例如，当INT报告显示卫星链路S2—S3出现拥塞预警时，控制平面评估从S1到S4的可用路径，发现原路径S1—S2—S3—S4性能下降，于是动态调整ECMP权重，将后续流量引导至备用的S1—S2—S5—S4路径，从而绕过拥塞链路，实现负载均衡（卫星网络拥塞规避场景如图6所示）。同时，系统支持按租户粒度配置ECMP策略，满足不同业务的负载均衡需求。

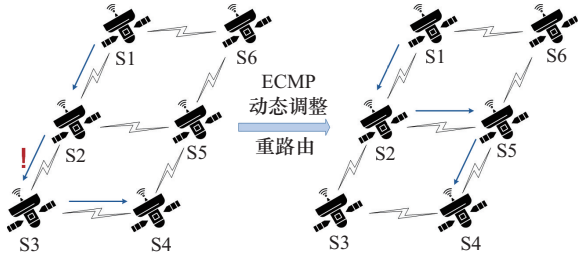


图6　卫星网络拥塞规避场景

2.3.3　重路由执行流程

重路由决策由控制平面生成并执行下发。首先，基于故障定位信息进行路径计算，利用逐跳报告精确定位故障设备或链路（如卫星节点S2队列溢出），控制平面结合网络拓扑和INT数据，计算替代路径（如绕行卫星S3）；随后，进行策略下发，控制平面封装新的路由策略消息，通过南向接口下发至相关设备（如地面网关、卫星节点），策略包括更新ECMP权重、修改转发表或调整隧道端点等；最后，执行流量切换，数据平面设备根据新策略转发流量，实现无缝重路由。对于已建立的流，系统采用“make-before-break”机制，先建立新路径再拆除旧路径，避免业务中断。卫星节点根据更新的路由表重写外层IP地址，引导流量至新路径。整个重路由过程的完成时间为秒级，可最小化对租户业务的影响。此外，系统支持渐进式权重调整，避免流量骤变导致网络振荡。

3　实验结果与分析

3.1　实验设置

为验证所提测量系统的有效性，本文基于Mininet[18]、

BMv2[19]和系统工具包（systems tool kit, STK）[20]搭建了一个网络仿真环境。该仿真环境搭建了穿越地面公有云、空间卫星网络的租户到租户的信息传输场景，完整路径为：VM1→vSwitch1→网关 1→卫星网络→网关 2→vSwitch2→VM2。地面云网络由 Mininet 网络仿真框架和 BMv2 可编程软件交换机构建，用于模拟公有云数据中心内的虚拟化网络。该部分创建了包含多个虚拟机、虚拟交换机和虚拟隧道端点的网络拓扑，模拟 VXLAN 隧道的流量封装、转发与交换行为。低轨卫星轨道由 STK 仿真生成，参考 Walker-Star 卫星星座设计，模拟低轨卫星网络的动态拓扑与星间链路特性。其中，构建了不同规模的低轨卫星星座，规模大小为 40~200 颗星。STK 负责生成高精度的卫星轨道数据（星历），并驱动网络拓扑实时演化。卫星节点间的路由与转发逻辑则通过自定义的网络仿真模块与地面网络进行对接。网络协议的数据平面报文处理由 P4 程序定义[21]，支持 INT 功能的 VXLAN 封装与解封装逻辑，确保 INT over VXLAN 协议栈的准确执行。此外，利用 Scapy 工具灵活构造符合协议格式的数据包，用于系统功能和性能的验证。

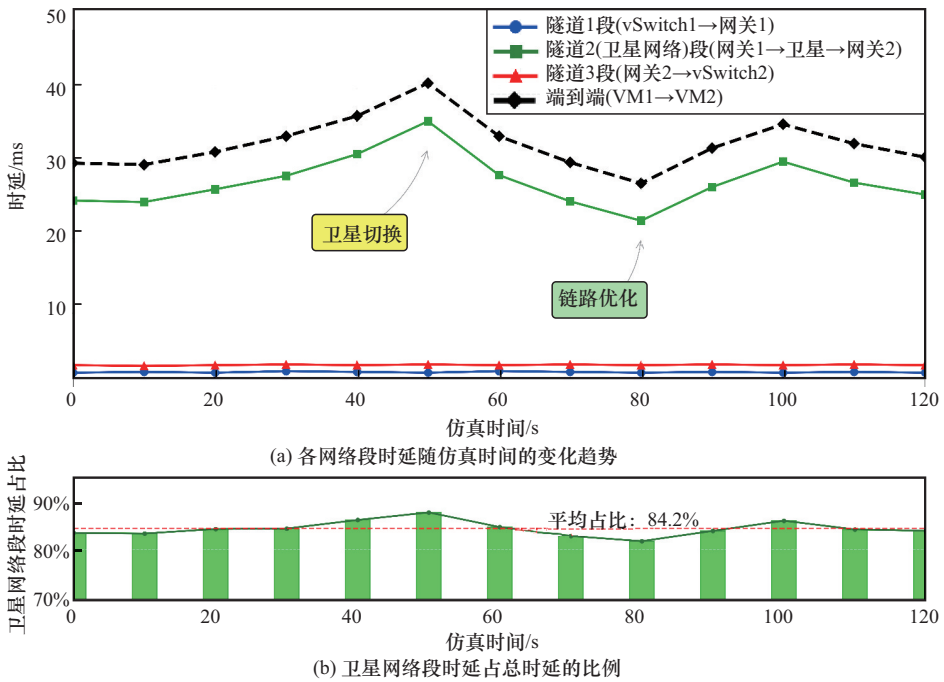
3.2　结果分析

3.2.1　遥测结果分析

基于仿真环境构建的具体场景，进行传输链路的带内网络遥测，对长时仿真中各分段时延的波动进行监测，得到各网络段时延如图 7 所示。实验获得了隧道 1 段时延、隧道 2（卫星网络）段时延、隧道 3 段时延和端到端总时延，端到端总时延与各分段时延之和高度一致，验证了 INT 数据在多次封装/解封装过程中的继承性与完整性。其中，地面网络的隧道 1 段与隧道 3 段时延保持基本稳定（2~3 ms），而卫星网络段时延呈现显著波动（22~36 ms），端到端总时延曲线与卫星网络段时延曲线高度重合，卫星网络段时延占总时延的 80%~90%，平均占比约 84.2%。结果表明，在此仿真场景中，卫星网络段可能成为网络性能瓶颈，时延波动与卫星网络动态事件直接相关。例如，仿真时间在 50 s 左右，卫星切换导致时延突增至 35.6 ms；仿真时间在 80 s 左右，经过链路优化后，时延下降至 22.3 ms。这证明了 INT 数据能够有效反映并关联底层网络拓扑变化。

3.2.2　遥测性能评估

启用 INT 遥测机制可能会对业务数据传输性能产生影响，为此，本文在 VM1 与 VM2 之间部署稳定的数据流，分别测量启用与不启用 INT 遥测机制下的时延，并剖析时延在跨域路径上的分布。遥测性能可以用时间开销来评估，时间开销定义为“启用 INT 遥测机制后的额外用时/原始业务时延×100%”。低轨卫星与公有云融合网络各段时延测量结果如图 8 所示。由图 8 可知，正常状态



(a) 各网络段时延随仿真时间的变化趋势

(b) 卫星网络段时延占总时延的比例

图 7　各网络段时延

下，启用INT遥测机制后，端到端总时延从29.6 ms增加为31.0 ms，增量小于5%，引入的额外时延开销很小，在可接受范围内。卫星网络段时延在拥塞情况下从24.5 ms增至42.8 ms，此增长主要源于卫星节点生成和处理独立INT报告报文时的处理时延。

此外，实验评估了测量方案引入的遥测开销，并选择INT-postcard[16,22]（也称INT-XD模式）作为对比基线。INT-postcard是INT规范定义的3种标准工作模式之一，代表最直接的分布式遥测数据采集思路，即在每个节点独立发送报告。本文提出的多模式机制正是在INT-postcard和INT-MD的思想基础上进行的，因此将纯粹的INT-postcard作为对比对象，能够最直观地验证“按需触发与跨域继承”机制在降低开销方面的有效性。此外，遥测开销主要分为两部分：一是数据平面开销，指生成、封装INT报告报文所消耗的额外带宽与设备转发处理资源；二是控制平面开销，指为支持动态测量和维护全局映射表所产生的信令开销。遥测系统开销如图9所示，与单纯使用INT-postcard方法相比，本文引入的数据平面和控制平面的开销均更小。与INT-postcard相比，本系统采用混合模式：默认使用低开销的end-to-end模式，仅在故障时按需触发逐跳模式。本文仅当链路故障时启用多个节点报告，而INT-postcard需要在每个节点产生独立的报告报文，监控系统需要处理更多的数据包，从而带来更多网络开销。

3.2.3　故障定位评估

为验证系统的故障诊断能力，仿真系统在不同位置注入故障，并测试故障定位效率。在实验中，受限于仿真场景，重点验证了丢包率指标，并在不同位置注入持续丢包故障。实验故障共有3种具体场景：网关故障测试（在网关1处注入异常丢包）；卫星链路故障测试（在S2—S3星间链路注入丢包）；混合故障测试（在网关、星间链路处注入丢包）。此外，在实现ECMP动态负载均衡时，为了兼顾时延敏感性与丢包敏感性，权重系数设定为α=0.5（时延权重）、β=0.3（丢包权重）、γ=0.2（队列深度权重）。多层故障定位用时随仿真时间的变化趋势如图10所示，其分别给出了总定位用时（T）、Overlay故障定位用时

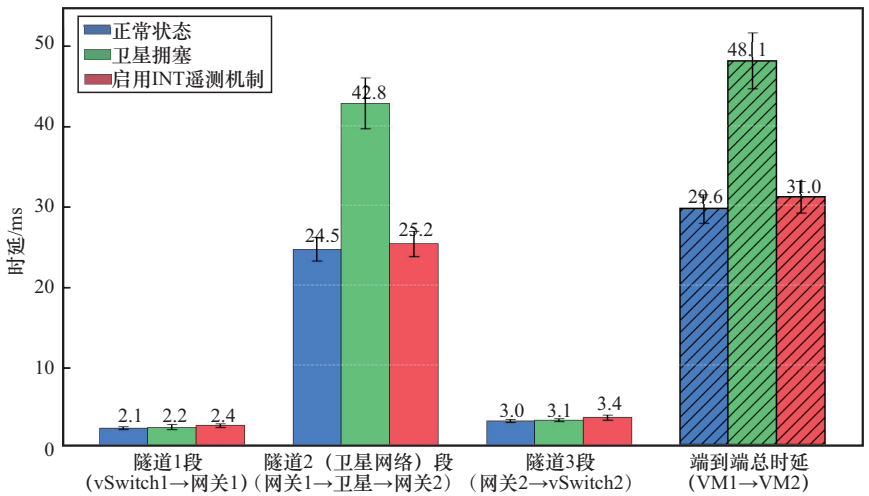


图8　低轨卫星与公有云融合网络各段时延测量结果

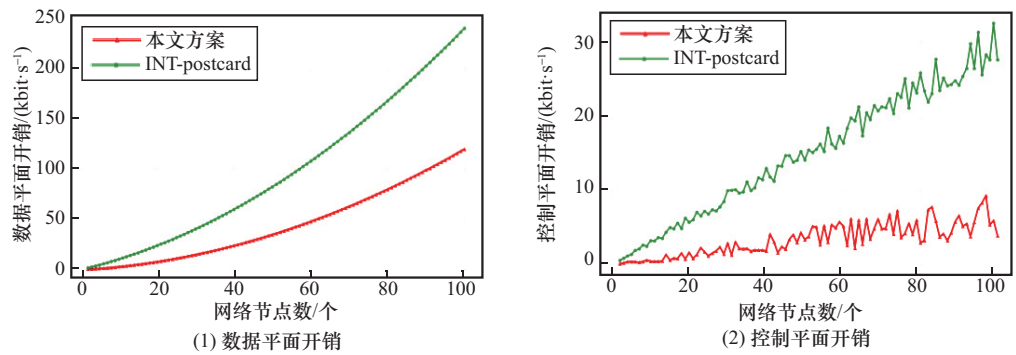


图9　遥测系统开销

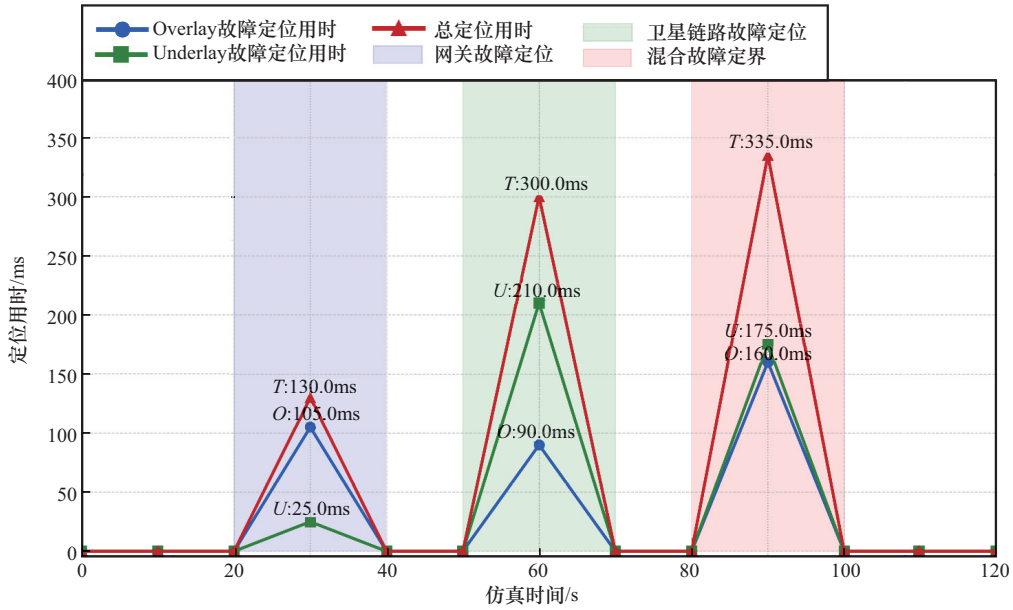


图 10　多层故障定位用时随仿真时间的变化趋势

（*O*）以及 Underlay 故障定位用时（*U*）。其中，网关故障定位最快（130.0 ms），主要由 Overlay 定位完成；卫星链路故障定位较慢（300.0 ms），Underlay 定位占主导；混合故障定位最慢（335.0 ms），需要 Overlay 和 Underlay 完整定位。实验结果证明，“端到端监测→Overlay 定位→Underlay 定位”机制有效实现了多层协同故障定位。Overlay 定位速度快，适用于快速收敛故障范围；Underlay 定位虽耗时稍长，但能实现物理位置的精准定位，两者结合实现了效率与精度的平衡。

4　结束语

随着低轨卫星网络与公有云平台的深度融合，构建精准高效的网络运维与故障诊断体系已成为支撑天地一体化云服务的关键。本文针对该融合场景下网络可观测性面临的独特挑战，提出了一套完整的跨域遥测系统：首先分析了卫星网络动态性、云网络虚拟化以及二者融合带来的多层隧道封装、跨域关联感知、动态链路回传等测量难题；在此基础上，设计了基于 INT 的遥测系统架构，详细阐述了支持数据继承的端到端测量、Overlay/Underlay 逐跳报告回传等关键工作流程，并提出了基于实时 INT 信息的智能重路由机制。跨域遥测系统实现了星地融合场景下的连续细粒度测量，通过轻量级继承机制解决了多次隧道封装导致的测量中断问题；打通了虚拟与物理网络的故障关联分析，利用多层逐跳 INT 报告机制实现了从租户业务到卫星链路的精确故障定位；设计了跨动态星地链路下的遥测信息传输机制，实现了遥测信息的有效星地传输；将实时测量与动态路由优化结合，提升了网络的自愈性与服务质量保障水平。仿真结果表明，本系统能以低于 5% 的额外时间开销，实现全链路连续测量；在故障诊断方面，能够快速完成异常感知和网络链路的精确定位。

未来工作将围绕以下方向展开：探索在星上资源受限环境下，进一步优化 INT 报告报文的压缩与聚合算法，降低传输开销，缓解星地链路压力；利用机器学习模型（如时空图神经网络）对历史与实时 INT 数据进行联合分析，防范网络拥塞、卫星负载过高等潜在风险。

参考文献

[1] AWS. Aerospace and satellite solutions: accelerate innovation from orbit to cloud[R]. 2025.

[2] IETF. Generic protocol extension for VXLAN (VXLAN-GPE): draft-ietf-nvo3-vxlan-gpe-11[S]. 2025.

[3] Kim C, Sivaraman A, Katta N, et al. In-band network telemetry via programmable dataplanes[C]//Proceedings of the ACM SIGCOMM Symposium on SDN Research (SOSR). New York: ACM, 2015.

[4] Zhang Y, Pan T, Fu Q, et al. In-band network-wide telemetry for topology-varying LEO satellite networks[C]//Proceedings of the GLOBECOM 2024-2024 IEEE Global Communications Conference. Piscataway: IEEE Press, 2025: 4479-4484.

[5] Zhang Y, Pan T, Zheng Y, et al. Topology-adaptive LEO satellite network telemetry via graph isomorphism and topology partitioning[C]//Proceedings of the 9th Asia-Pacific Workshop on Networking. New York: ACM, 2025: 93-99.

[6] Wood T, Shenoy P J, Gerber A, et al. The case for enterprise-ready virtual private clouds[C]//Proceedings of the 2009 USENIX Workshop on Hot Topics in Cloud Computing (HotCloud '09). Berkeley: USENIX Association, 2009.

[7] Case J D, Fedor M, Schoffstall M L, et al. A simple network management protocol (SNMP): RFC 1157[S]. 2025.

[8] Rasley J, Stephens B, Dixon C, et al. Planck: millisecond-scale monitoring and control for commodity networks[C]//Proceedings of the 2014 ACM Conference on SIGCOMM. New York: ACM, 2014: 407-418.

[9] Zhu Y B, Kang N X, Cao J X, et al. Packet-level telemetry in large datacenter networks[J]. ACM SIGCOMM Computer Communication Review, 2015, 45(4): 479-491.

[10] Gupta A, Harrison R, Canini M, et al. Sonata: query-driven streaming network telemetry[C]//Proceedings of the 2018 Conference of the ACM Special Interest Group on Data Communication. New York: ACM, 2018: 357-371.

[11] Guo C X, Yuan L H, Xiang D, et al. Pingmesh: a large-scale system for data center network latency measurement and analysis[C]//Proceedings of the 2015 ACM Conference on Special Interest Group on Data Communication. New York: ACM, 2015.

[12] Roy A, Bansal D, Brumley D, et al. Cloud datacenter SDN monitoring: experiences and challenges[C]//Proceedings of the Internet Measurement Conference 2018. New York: ACM, 2018: 464-470.

[13] Zhu S M, Lu J Y, Lyu B, et al. Zoonet: a proactive telemetry system for large-scale cloud networks[C]//Proceedings of the 18th International Conference on Emerging Networking Experiments and Technologies. New York: ACM, 2022: 321-336.

[14] Zhang Y, Pan T, Zheng Y, et al. VXLAN-based INT: in-band network telemetry for overlay network monitoring[C]//Proceedings of the IEEE INFOCOM 2021-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). Piscataway: IEEE Press, 2021: 1-2.

[15] Moshref M, Yu M L, Govindan R, et al. Trumpet: timely and precise triggers in data centers[C]//Proceedings of the 2016 ACM SIGCOMM Conference. New York: ACM, 2016: 129-143.

[16] The P4.org Architecture Working Group. In-band network telemetry (INT) dataplane specification, version 2.1[S]. 2025.

[17] SPACEX. Starlink[EB]. 2025.

[18] MININET. Mininet: an instant virtual network on your laptop (or other PC) - Mininet[EB]. 2025.

[19] Elbediwy M, Pontikakis B, David J P, et al. Enabling rank-based P4 programmable schedulers: requirements, implementation, and evaluation on BMv2 switches[J]. IEEE Transactions on Networking, 2025, 33(1): 299-310.

[20] ANSYS. Ansys STK digital mission engineering software[EB]. 2025.

[21] P4.org. P4 language consortium: learn, develop, standardize[EB]. 2025.

[22] IETF. Postcard-based on-path flow data telemetry: draft-song-ippm-postcard-based-telemetry-05[S]. 2019.

[作者简介]



张妍（1994– ），女，硕士，紫金山实验室工程师，主要研究方向为低轨卫星网络、云数据中心网络等。



潘恬（1987– ），男，博士，紫金山实验室骨干科研人员，主要研究方向为云数据中心网络、智算中心网络、低轨卫星网络、高速可编程网络设备等。



郑奘（1988– ），男，硕士，紫金山实验室工程师，主要研究方向为低轨卫星网络仿真、网络测量等。

吴慧文（1997– ），女，硕士，国网山东省电力公司青岛供电公司工程师，主要研究方向为电力系统通信、5G电力应用研究等。

赵晓红（1990– ），女，硕士，国网山东省电力公司电力科学研究院工程师，主要研究方向为电力通信及网络研究。

陈平平（1986– ），男，紫金山实验室高级工程师，主要研究方向为新型融合网络、网络人工智能、确定性网络和数据通信技术。