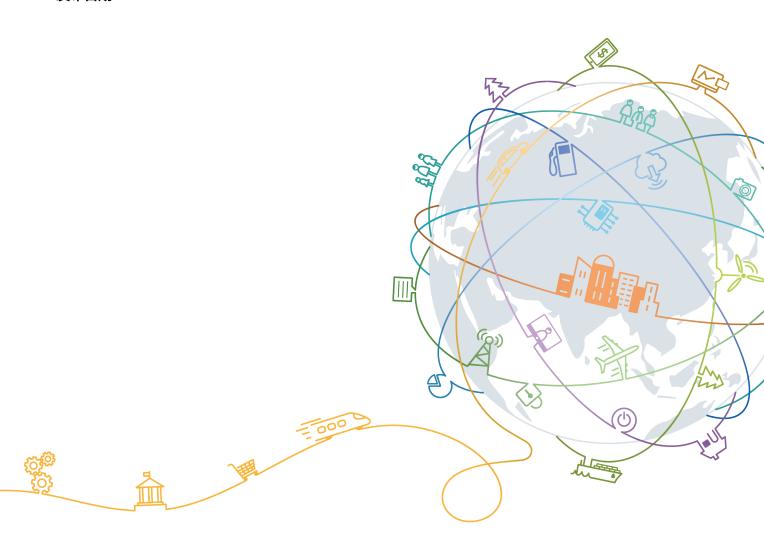
什么是 Telemetry? 如何监控 Telemetry 数据?

文档版本 01

发布日期 2020-12-29





版权所有 © 华为技术有限公司 2020。 保留一切权利。

非经本公司书面许可,任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部,并不得以任何形式传播。

商标声明



HUAWE 和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标,由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束,本文档中描述的全部或部分产品、服务或 特性可能不在您的购买或使用范围之内。除非合同另有约定,华为公司对本文档内容不做任何明示或默示的声 明或保证。

由于产品版本升级或其他原因,本文档内容会不定期进行更新。除非另有约定,本文档仅作为使用指导,本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为技术有限公司

地址: 深圳市龙岗区坂田华为总部办公楼 邮编: 518129

网址: https://e.huawei.com

目录

1	什么是 Telemetry	. 1
2	为什么使用 Telemtry 监控网络	. 2
3	Telemetry 系统架构及处理流程	.4
4	配置订阅	. 7
5	Telemetry 监控的数据	. 8
6		12

1 什么是 Telemetry

Telemetry是一项远程的从物理设备或虚拟设备上高速采集数据的技术。设备通过推模式(Push Mode)周期性的主动向采集器上送设备的接口流量统计、CPU或内存数据等信息,相对传统拉模式(Pull Mode)的一问一答式交互,提供了更实时更高速的数据采集功能。

2 为什么使用 Telemtry 监控网络

随着SDN网络的设备规模日益增大,承载的业务越来越多,用户对SDN网络的智能运维提出了更高的要求,包括监控数据的采样周期拥有更高的精度以便及时检测和快速调整微突发流量,同时监控过程要对设备自身功能和性能影响小以便提高设备和网络的利用率。

传统网络监控方式(如SNMP get和CLI),因存在如下不足,管理效率越来越低,已不能满足用户需求的演进:

- 通过拉模式来获取设备的监控数据,不能监控大量网络节点,限制了网络增长。
- 监控数据的采样周期一般为分钟级,只能依靠加大查询频度来提升获取数据的准确度,但是这样会导致网络节点CPU利用率高而影响设备的正常功能。
- 由于网络传输时延的存在,监控到的网络节点数据并不准确。

因此,面对大规模、高性能的网络监控需求,用户需要一种新的网络监控方式。 Telemetry技术可以满足用户要求,支持智能运维系统管理更多的设备、监控数据拥有 更高精度和更加实时、监控过程对设备自身功能和性能影响小,为网络问题的快速定 位、网络质量优化调整提供了最重要的大数据基础,将网络质量分析转换为大数据分 析,有力的支撑了智能运维的需要。Telemetry与传统网络监控方式的对比如表2-1所示。

表 2-1 Telemetry 与传统网络监控方式的对比

对比项	Telemetry	SNMP Get	SNMP Trap	CLI	SYSLOG
工作模式	推模式	拉模式	推模式	拉模式	推模式
精度	亚秒级	分钟级	秒级	分钟级	秒级

对比项	Telemetry	SNMP Get	SNMP Trap	CLI	SYSLOG
是否结构化	YANG模型 定义结构	MIB定义结构	MIB定义结构	非结构化	非结构化

山 说明

如<mark>表2-1</mark>所示,SNMP Trap和SYSLOG虽然是推模式的,但是其推送的数据范围有限,仅是告警或者事件,对于类似接口流量等监控数据不能采集上送。

3 Telemetry 系统架构及处理流程

系统架构

Telemetry是一个闭环的自动化运维系统,分为OSS侧和设备侧,由网络设备、采集器、分析器和控制器等部件组成,如图3-1所示。

- OSS侧:由采集器、分析器、控制器组成。进行数据的收集、存储、应用分析及 控制。
- 设备侧:由网络设备组成。将采样的数据按照编码格式进行编码,并且使用传输协议进行数据传输。

其中,网络设备、采集器、分析器、控制器既可以使用第三方的系统,也可以使用华为的系统。华为Telemetry系统中,网络设备对应的产品为: CloudEngine交换机,采集器、分析器对应的产品为: iMaster NCE-FabricInsight,控制器对应的产品为: iMaster NCE-Fabric。

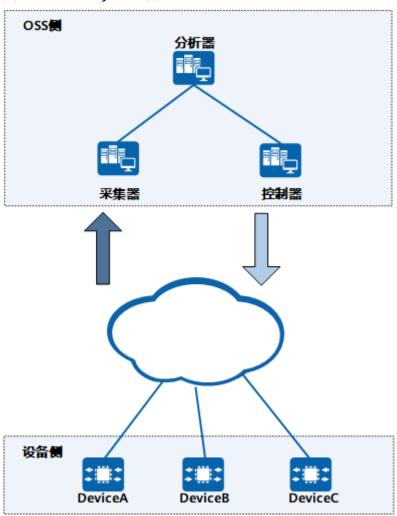


图 3-1 Telemetry 系统架构

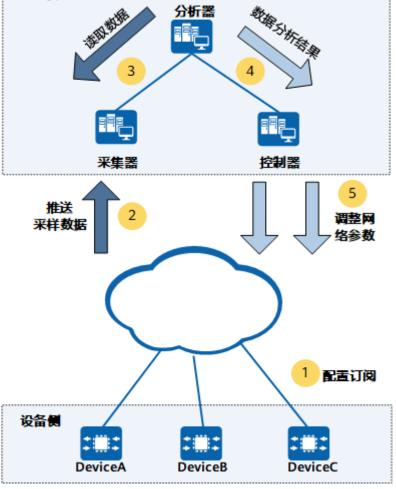
Telemetry 处理流程

Telemetry需要OSS侧和设备侧协同运作,完成整体的Telemetry需要五个操作步骤顺序执行,如图3-2所示:

- 1. 配置订阅:订阅数据源,完成数据采集。目前支持如下两种方式:
 - 静态订阅:通过命令行配置订阅数据源,完成数据采集。Telemetry静态订阅 往往用于粗粒度的数据采集。
 - 动态订阅:通过命令行配置gRPC(google Remote Procedure Call Protocol,Google远程过程调用协议)服务的相关配置后,由采集器下发动 态配置到设备,完成数据采集。
- 2. 推送采样数据:网络设备依据控制器的配置要求,将采集完成的数据,上报给采集器进行接收和存储。
- 3. 读取数据:分析器读取采集器存储的采样数据。
- 4. 分析数据:分析器分析读取到的采样数据,并将分析结果发给控制器,便于控制器对网络进行配置管理,及时调优网络。
- 5. 调整网络参数:控制器将网络需要调整的配置下发给网络设备;配置下发生效后,新的采样数据又会上报到采集器,此时Telemetry OSS侧可以分析调优后的网络效果是否符合预期,直到调优完成后,整个业务流程形成闭环。

OSS∰

图 3-2 Telemetry 业务处理流程



4 配置订阅

在CloudEngine侧通过命令行配置订阅数据源,完成数据采集。目前支持如下两种方式:

- 静态订阅:通过命令行配置订阅数据源,完成数据采集。当前华为iMaster NCE-FabricInsight分析器通过静态订阅方式进行CPU、内容等数据的采集。
- 动态订阅:通过命令行配置gRPC服务的相关配置后,由采集器下发动态配置到设备,完成数据采集。当前华为iMaster NCE-FabricInsight分析器通过动态订阅方式进行路由表项等数据的采集。

具体配置可以参考《CloudEngine配置指南-网络管理与监控-Telemetry配置》。

5 Telemetry 监控的数据

华为iMaster NCE-FabricInsight分析器,主要采集和分析数据中心网络中的应用和流的信息,以及网络设备上报的Telemetry信息,并可视化展现给用户,帮助用户快速完成问题定界和排障。



目前支持采集和分析的数据为:

- 设备:设备级性能指标信息,包括: CPU利用率、内存利用率、防火墙IPV4会话数。
- 单板:单板级性能指标信息,包括: CPU利用率、内存利用率、防火墙IPV4会话数、ARP利用率、FIBv4利用率、FIBv6利用率、MAC利用率。
- 芯片:芯片级性能指标信息,包括:TCAM利用率。
- 接口:接口级性能指标信息,包括:接收字节数、发送字节数、接收包数、发送包数、接收丢包数、发送丢包数、单播/组播/广播接收包数、单播/组播/广播发送包数、接收错误包数、发送错误包数、接收带宽占用率、发送带宽占用率、ECN报文数等。
- 队列:接口队列级性能指标信息,包括:队列缓存、已使用Headroom缓存、接收PFC反压帧数、发送PFC反压帧数、PFC死锁监控次数、PFC死锁恢复次数、已使用Guaranteed缓存。
- 光链路:光链路级性能指标信息。
- Al Fabric: RoCE网络性能指标信息。从带宽利用率、PFC反压帧数、PFC死锁次数、ECN报文数等性能指标维度,展示使能Al ECN带来的收益。

设备级性能指标信息

以CPU利用率为例,显示如下信息:

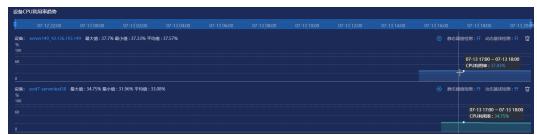
● 设备CPU利用率分布

圆环图展示CPU利用率超阈值的设备和正常的设备分布情况。矩阵图展示每个设备的CPU利用率指标概要信息。



● 设备CPU利用率趋势

通过Portlet直观展示指定时间窗内各设备的CPU利用率趋势。将鼠标置于某个时间点,可查看该时间点各设备的CPU利用率。



单板级性能指标信息

查看单板性能指标概览信息,以CPU利用率为例:

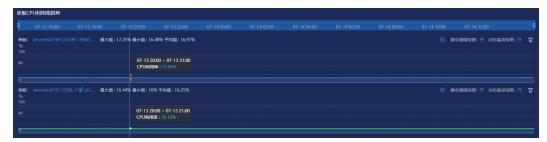
● 单板CPU利用率分布

圆环图展示CPU利用率超阈值的单板和正常的单板分布情况。矩阵图展示每个单板的CPU利用率指标概要信息。



● 单板CPU利用率趋势

通过Portlet直观展示指定时间窗内各单板的CPU利用率趋势。



芯片级性能指标信息

查看芯片性能指标概览信息,以TCAM利用率为例:

● 芯片TCAM利用率分布

圆环图展示TCAM利用率超阈值的芯片和正常的芯片分布情况。矩阵图展示每个芯片的TCAM利用率指标概要信息。



● 芯片TCAM利用率趋势

通过Portlet直观展示指定时间窗内各芯片的TCAM利用率趋势。



接口级性能指标信息

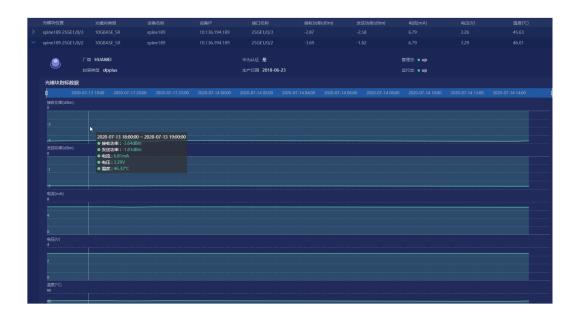
接口性能指标,包括:接收字节数、发送字节数、接收包数、发送包数、接收丢包数、发送丢包数、单播/组播/广播接收包数、单播/组播/广播发送包数、接收错误包数、发送错误包数、接收带宽占用率、发送带宽占用率、ECN报文数等。

接口队列级性能指标信息

接口队列性能指标,包括:队列缓存、已使用Headroom缓存、接收PFC反压帧数、发送PFC反压帧数、PFC死锁监控次数、PFC死锁恢复次数、已使用Guaranteed缓存。

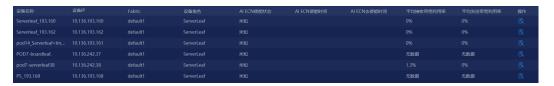
光链路性能指标信息

光链路的列表中展示了所有光模块的性能指标,展开可查看对应光模块的基本属性和 各种性能指标的详细信息。



AI Fabric 性能指标信息

RoCE网络设备统计列表中,展示了RoCE网络中各设备的ECN性能指标信息。



单击某个设备操作列的,可查看对应的AI Fabric性能指标详情。

AI ECN使能收益对比:通过对比图从带宽利用率、PFC反压帧数、PFC死锁次数、ECN报文数等性能指标进行对比,展示使能AI ECN给RoCE网络设备带来的收益。



6相关参考

《CloudEngine配置指南-网络管理与监控》

《iMaster NCE-FabricInsight 产品文档》