DeepFlow®

混合云网络监控诊断方案

云杉网络

目录

1. 前言

经过十多年的发展，企业在 IT 基础设施以及云原生的业务应用上稳步推进。上云业务规模增加，混合云中网络变得更为复杂，企业对业务安全的诉求、行业主管部门监管的要求有增无减。本方案介绍如何面向新型企业混合云环境，将网络保障与应用业务紧密结合建设有效的网络监控、分析系统。

By 2024, 50% of network operations teams will be required to rearchitect their network monitoring stack, due to the impact of hybrid networking, which will be a significant increase from 20% in 2019. "Market Guide for Network Performance Monitoring and Diagnostics 2020.04"

2. 混合云网络监控面临的新挑战

网络监控诊断并不是一个新的领域，其伴着网络的发展始终存在。企业 IT 基础设施部门对此也并不陌生，但在混合云、云原生环境中，网络监控诊断面临新的挑战。主要集中在：

资源池内的网络监控诊断

在现有 IT 环境中，网络不再是仅仅由物理网络设备组成，在资源池中，是由软件编写的虚拟交换机，运行在计算节点上实现逻辑计算资源间的网络交换，通过网络编排，具备虚拟私有云（*VPC，Virtual Private Cloud*），网络功能服务链（*SFC，Service Function Chaining*）、微分段（*Micro-Segmentation*）等功能、承载业务应用。

在数据中心，资源池内的东西向流量占据大额比例，越来越多的业务应用迁移至混合云、云原生环境中，在微服务架构中，服务间的网络监控是业务保障中重要部分；在容器网络中，POD 间的网络流量迫切需要工具手段进行监控保障。

网络与业务应用紧密结合

网络团队的日常工作越来越多地关注、涉及应用业务，与系统部门的交界面也在融合，简单、孤立的 IP 视角不再满足目前的工作需求。

网络监控架构需要弹性扩展

通过资源池化，计算、网络等资源实现了可弹性扩展，面向整个企业不断发展、增加的网络资源池，网络监控架构同样需要可扩展的架构设计。混合云环境中，网络规模宏大且资源池类型繁多，需要考虑多数据中心的整体方案，避免针对不同需求重复安装探针，分散建设、分散管理的情况。虚拟交换机不再是物理网络设备，其数量相等于计算节点数量，与物理链路的采集点相比，数量是几个数量级的增长。此外，虚拟化及容器资源池动态性很强，尤其是容器，其资源随应用需求变化频繁发生迁移、切换或回收，流量采集策略、流量分发策略也要随着变化进行迁移或释放。

在构建整体网络监控诊断方案时，应充分考虑需要监控、优化的业务，分布在哪些链路、区域以及资源池，平台可以分阶段进行部署，但要具备扩展和统一管理能力。

对现网环境的影响

应尽可能地避免对现有云环境的影响，在已经投入生产的环境中，可能存在未规划且独立的流量监控平面；逻辑 CPU 按用途完全划分；已经部署和应用不同网络虚拟化产品的方案等情况。在进行流量采集部署时，需要满足平滑部署且保证业务不间断，同时，有机制保障对计算资源的消耗限制。

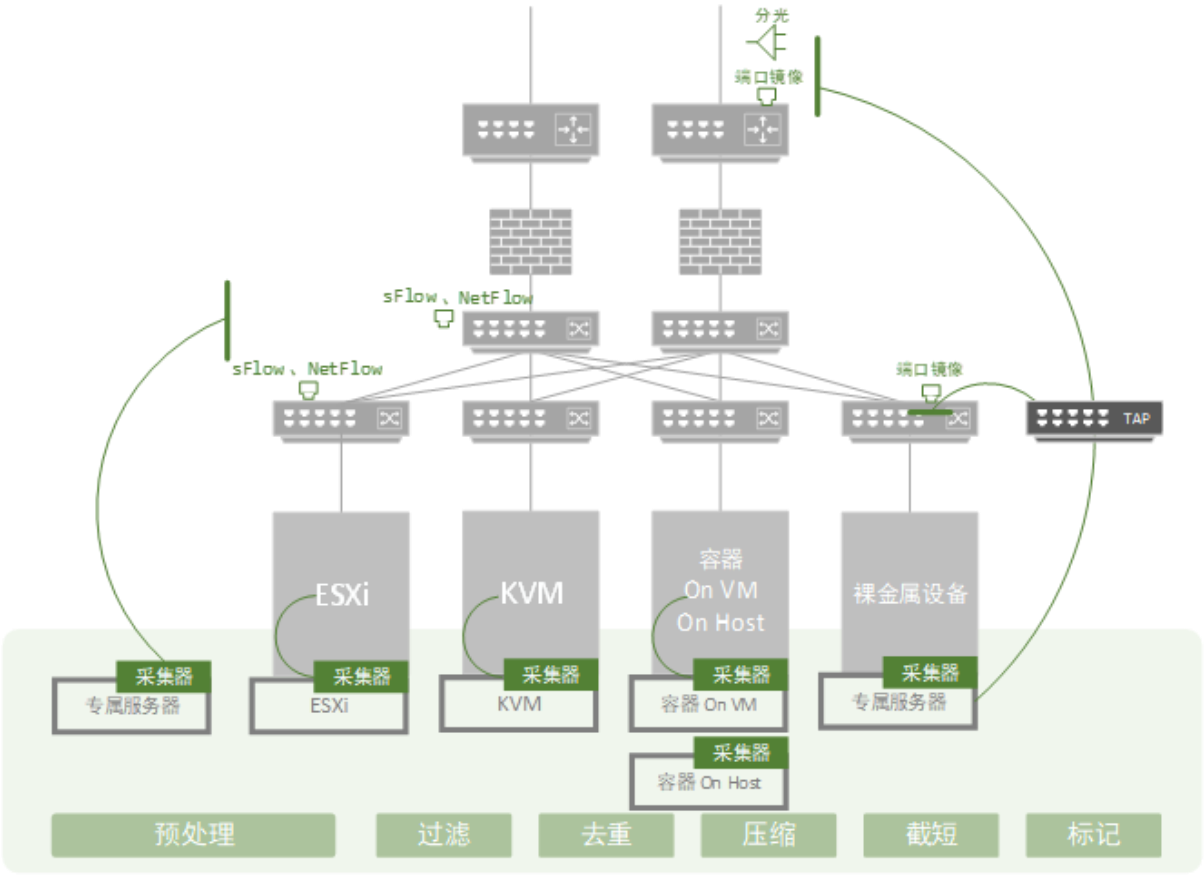
此外，流量采集系统的部署也要保证对已有的物理网络分流镜像有能力进行兼容或平滑切换，并可以对接已有的分析工具。

3. 混合云流量监控诊断方案

本方案的目标是为企业混合云、云原生环境建设可扩展的网络流量监控诊断平台，应对 KVM、ESXi、容器等各类资源池内的网络性能分析。支持 IPv4、IPv6 协议环境，紧密结合业务应用，实现对全网流量监控。整体方案由 DeepFlow®的采集器、控制器以及数据节点组成。

采集器 网络流量获取及预处理

在生产环境中，获取混合云、云原生环境的网络流、数据包并非易事，基于分布式架构，实现对各类型资源池以及物理网络的流量采集处理抽象层，如图：



在可扩展网络监控框架下，流量采集与后端监控分析实现解耦，在采集侧，各类型号的 DeepFlow® 流量采集器为全网流量采集方案提供基础捕获能力，支持物理网络、KVM、ESXi、容器等资源池网络环境，详细方案可参考云杉网络“混合云全网流量采集与分发方案”。

控制器 平台控制中枢

对于多数据中心、多云异构的混合云基础设施，面临着采集器的数量巨大的问题，如在容器环境中，单一资源池拥有 50-100 台物理计算节点，运行 10000-20000 个 Pod 单元。

同时，对混合云的整体网络监控也需要提供统一的服务提供点。控制器是整个平台的控制中枢，以控制器集群方式实现对平台的管理控制平面的扩展，分为主控制器、备控制器、从控制器，可按照部署要求进行选择。

主控制器：整个 DeepFlow®平台的控制中枢和提供对外交互、服务的接口。部署后的 DeepFlow®平台中只有一台主控制器，主控制器所在的区域称之为主区域。

备控制器：与主控制器的功能完全一致，当主控制器出现宕机或不能提供服务的其他故障时，自动切换到备控制器。在没有备控制器的情况下，DeepFlow®控制器集群没有高可用能力。整个 DeepFlow®中只有一台备控制器，且必须和主控制器在同一个区域中，并共享一个用于提供外部服务的虚 IP 地址。

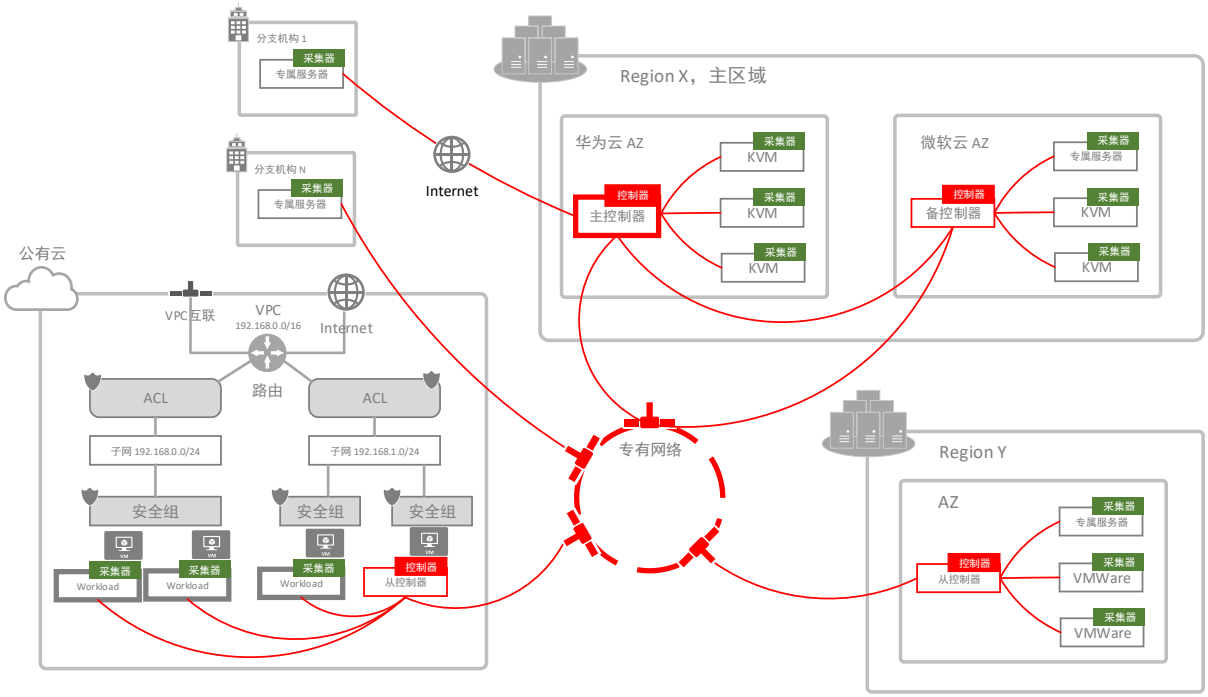
从控制器：负责控制所在区域（*Region*）或可用区（*AZ，Available Zone*）中的采集器及数据节点，将主控制器的策略和云平台资源信息同步至所有的采集器和数据节点。除主、备控制器所属的区域，每个区域中至少部署一台从控制器，同一个区域的多台从控制器之间可以实现负载均衡和高可用。

在多点的部署环境中，首先指定主区域（*Region*），主控制器存在于主区域中，当需要启用主控制器高可用功能时，主区域内应部署多台控制器，通过心跳保证控制器间的状态同步，及时启动主、备控制器选举。选举产生主控制器后，为整体流量管理平台提供控制入口。除主区域外的其他区域控制器为从控制器，不参与主控制器选举。

在区域中可以划分多个可用区（*AZ，Available Zone*），通常以可用区为单元，由单一控制器独立控制可用区内的各类型采集器，对本地采集器进行采集策略、分发策略、预处理策略下发。多区域间可通过专线网络进行控制通信，主要包括管理、策略等通信。

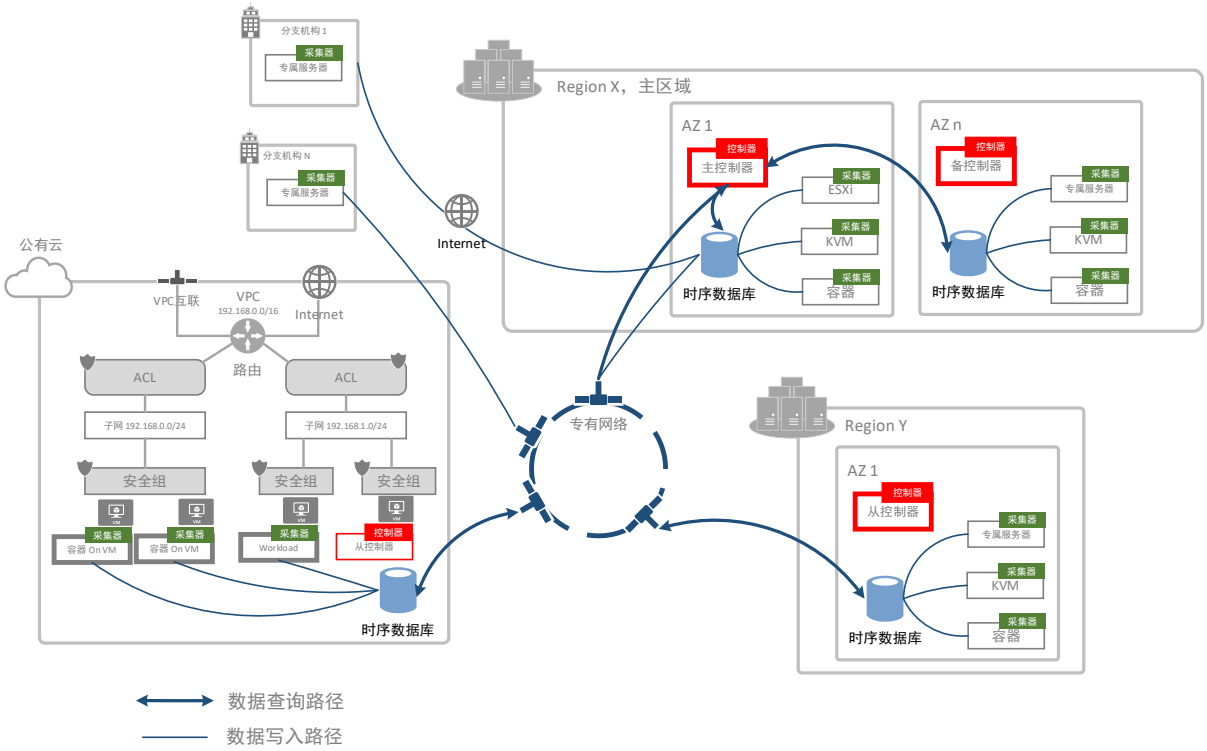
通常在有分支机构的环境下，数量相对数据中心较多，主要是请求服务的流量，其区域内没有服务端，需要流量数据主要是构建网络整体状况以及业务端到端网络性能分析。不需要独立部署控制器，可以按实际情况，将采集器划分在附近区域的控制器管理下。

公有云环境中，控制器部署在虚拟机中，管理范围内的采集器。



数据节点 高性能网络时序数据库

网络流量数据是典型的时间序列数据，同时具备相应的网络特性。满足网络监控诊断要求，需要具备对所存储的网络数据进行分组聚合，提供高性能查询能力，展示网络趋势、规律、异常等。数据节点分布式运行在网络时序数据库，为平台提供时序数据的快速写入、持久化、多纬度的聚合查询等基本功能。



时序数据库（*TSDB，Time Series Database*）是用于存储监控数据的专用数据库，通过对监控数据在时间维度上的压缩存储降低写入开销。

写入特征：由于网络通信的端到端特性，一个万台服务器的环境中产生的系统监控数据每秒写入量级为 O(N)，但每秒产生的网络数据取决于相互通信的服务器数量，极端情况下可能达到 O(N^2)。若将通信时的协议、端口号也进行记录，还会导致监控数据对象进一步升高，因此用于记录网络监控数据的时序数据库所需具备的首要特性是亿级别数据对象的支持能力。此外，云环境中所固有的弹性也要求时序数据库需要支持弹性伸缩。

查询特征：除了常规的查询某个 IP 地址以外，对云环境中资源的网络监控还要求能从各种维度进行查询，这需要对监控数据添加不同维度的属性。例如资源池维度的区域、可用区；虚拟化维度的宿主机、虚拟机；容器维度的节点、POD；应用维度的业务、资源组等。另外，对于属性的查询还需要拥有 IP 段、CIDR、服务端口段等范围查询能力。最后，云环境中资源数量的增多，要求故障诊断不能再依靠总量、峰值、均值等简单的统计数据，时序数据库应当提供更丰富的统计和计算能力，如中值、概率分布、信息熵、方差等。

网络数据可视化

企业通过 DeepFlow®平台已经掌握了混合云环境中的网络数据，将数据在监控诊断领域有效地展示及使用是本方案的重点。经过长期对网络工作的跟踪，结合混合云、容器等新型环境网络特点，将以下列几个视角归类网络数据。

按数据类型

原始数据包（*Packet*）

对重点业务，需要周期性保时保序地存储完整的数据包，或者在异常情况下，有方式手段地进行数据包抓取，供深度的协议、载荷内容（*Payload*）分析。

网络流数据（*Flow*）

完整记录全网网络流信息，除源目的 MAC 地址、源目的 IP 地址、源目的端口号、协议等七元组信息外，还包括隧道、TCP 状态、出入流量、延时等百余种类型信息。供获取全网网络状态以及各类指标热点分布、流特征等类型分析。

统计数据

通过简单网络管理协议（*SNMP，Simple Network Management Protocol*）、遥测（Telemetry）技术、平台自身进行统计等方式获取，通常针对物理网络设备的数据获取，与采集器获取的数据进行比对。

按属性维度

资源池： 区域（*Region*）、可用区（*AZ，Available Zone*）

虚拟化：宿主机（*Host*）、虚拟机（*VM，Virtual Machine*）

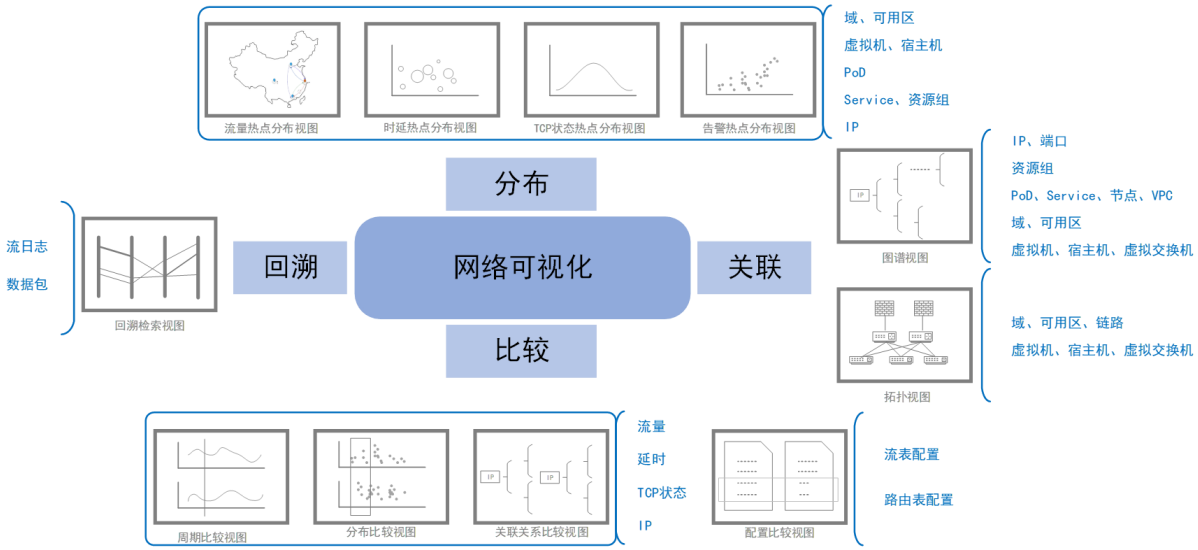
容器：容器节点（*Node*）、容器 Pod

应用相关：业务，资源组

网络：IP、VPC（*Virtual Private Cloud*）、子网

按展示方式

对数据进行目标明确的搜索展示，需要根据使用场景有选择的对各类型数据进行展示，主要分为分布、关联、对比及回溯几个重点方式。



分布类

在分布类视图中，主要解决流量、延时、TCP 状态、告警等历史周期数据、瞬时数据在地域、虚拟化平台、容器平台、应用等维度的分布，快速对应热点，调整分配计划。如在流量热点分布视图中，快速全面掌握不同地域间多个可用域的服务访问压力、可用域间专线链路状态，可从此视图向其他视图对比切换。

关联类

对虚拟网络达到监控诊断的目的，需要对各维度间的实例，所存在的多对多关系进行合理、快速展示。以任何一种维度实例快速查询相关联的维度实例，为整个容器环境构建网络图谱，实现可视化、网络图谱查询、有效提高在动态性突出的资源池内网络中快速定位故障、实现对资源池内网络的可控可管。

比较类

在比较类视图中，包含周期比较、分布比较、关联关系比较视图，其主要目的是快速找不同，找差异。以周期比较视图为例，获取同一时间周期的流量、数据包、延时的折线图，以时间轴推进来比较是在分析网络异常时常用的比较类视图。

网络虚拟化中，涉及到繁多网络配置信息，包括安全组策略、流表等，配置比较视图能快速对比配置，找出配置差异，在排障过程中节省宝贵时间。

回溯类

回溯类视图通过平行坐标图或桑基图，为网络运维团队提供了对历史网络流日志数据完整多维度视图回溯的能力，并可获取相关的原始数据包。

按监控指标

描述网络状态、性能的各类指标，主要包括吞吐量、时延、异常、传输状态等核心指标。

网络的点线面

在网络监控诊断的过程中，不同岗位，不同阶段运用不同的操作以及有不同的关注点，DeepFlow®通过点、线、面的操作划分，为混合云所涉及到的庞大网络提供完整的、丰满的监控保障。

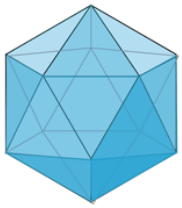
点

网包 Packet

线

网流 Flow

端到端 End-to-End



面

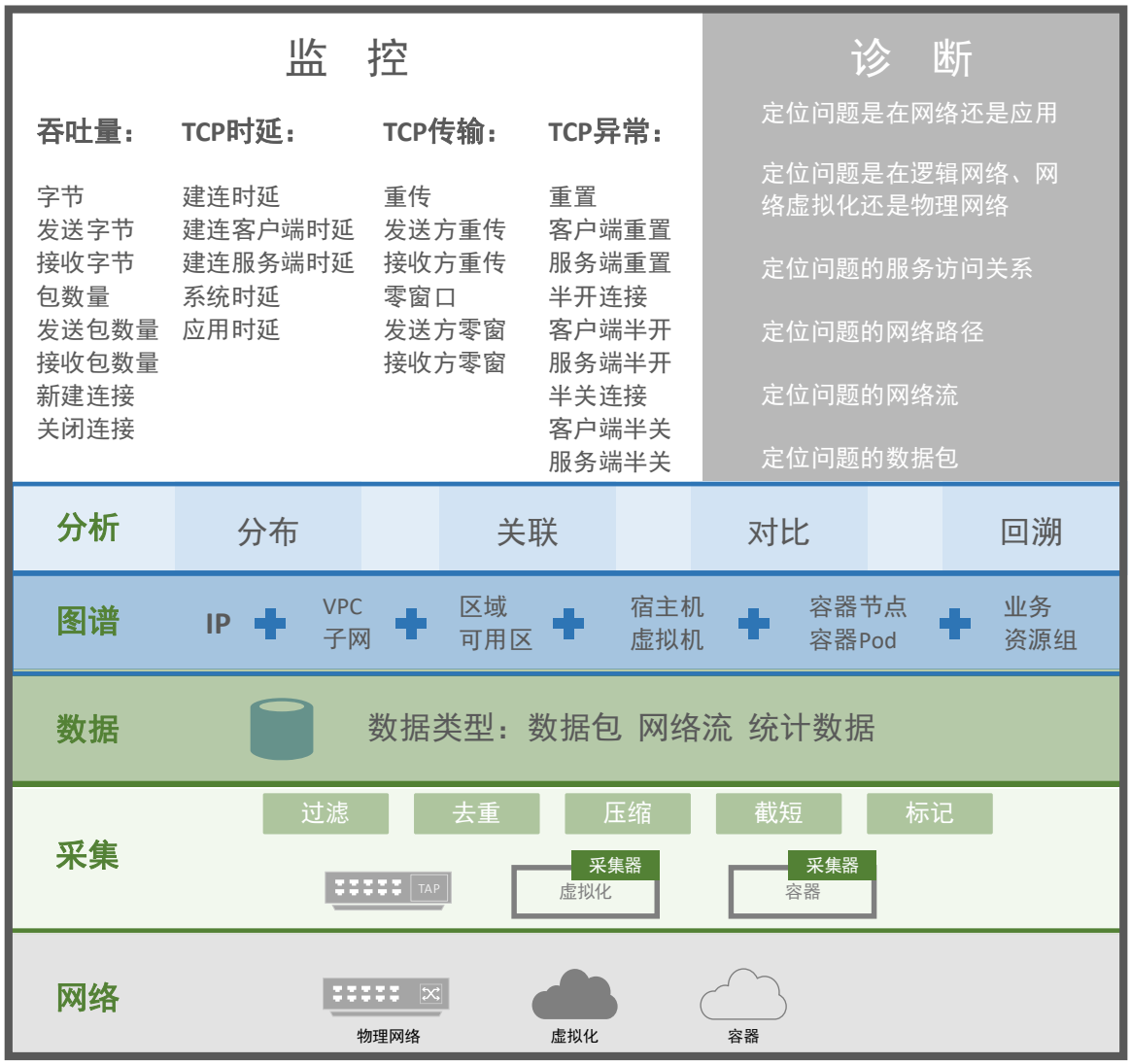
网络全景图 View

在网络监控所涉及到的分析对象中，将最终钻取到的具体数据包看做是“点”，提供深入详细的协议识别，异常排查证据。

完整的业务访问路径，一条网络流则是由多个“点”连成的“线”。在路径中包括链路、防火墙、负载均衡、服务等信息，是网络保障业务稳定运行的关键视角。流信息包含了关键的网络元数据。

“点”与“线”在传统的网络监控方案中很常见，在混合云环境中，只要能采集到相应的现网流量，展现及分析都变化不大。但在多地数据中心、资源池化，并且涉及公有云资源、专线链路的 IT 环境中，市场上缺少全局的网络状态视角，这并不是一个单纯的统计汇总视角，而是一张关联 IaaS 资源、PaaS 资源、服务应用的知识图谱。

知识图谱包含网络所涉及的对象实体映射关系，显示结构拓扑与现网流量的一系列不同视角的视图展示，应用可视化技术描述资源实体及搜索技术提供更深度更广度的搜索交互

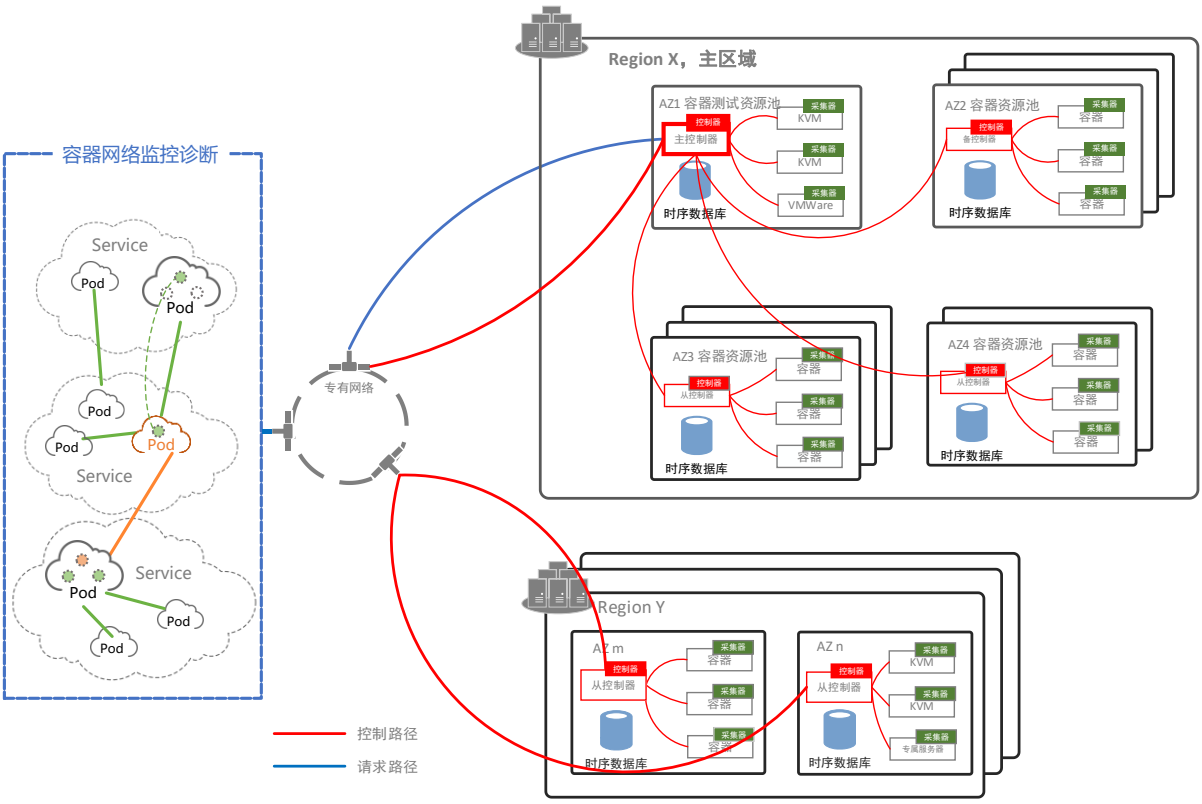


在混合云网络中，通过有效的网络流量采集、数据分类存储以及网络“面”、“线”、“点”的紧密结合，以应对企业网络监控诊断的挑战。

容器网络监控诊断场景

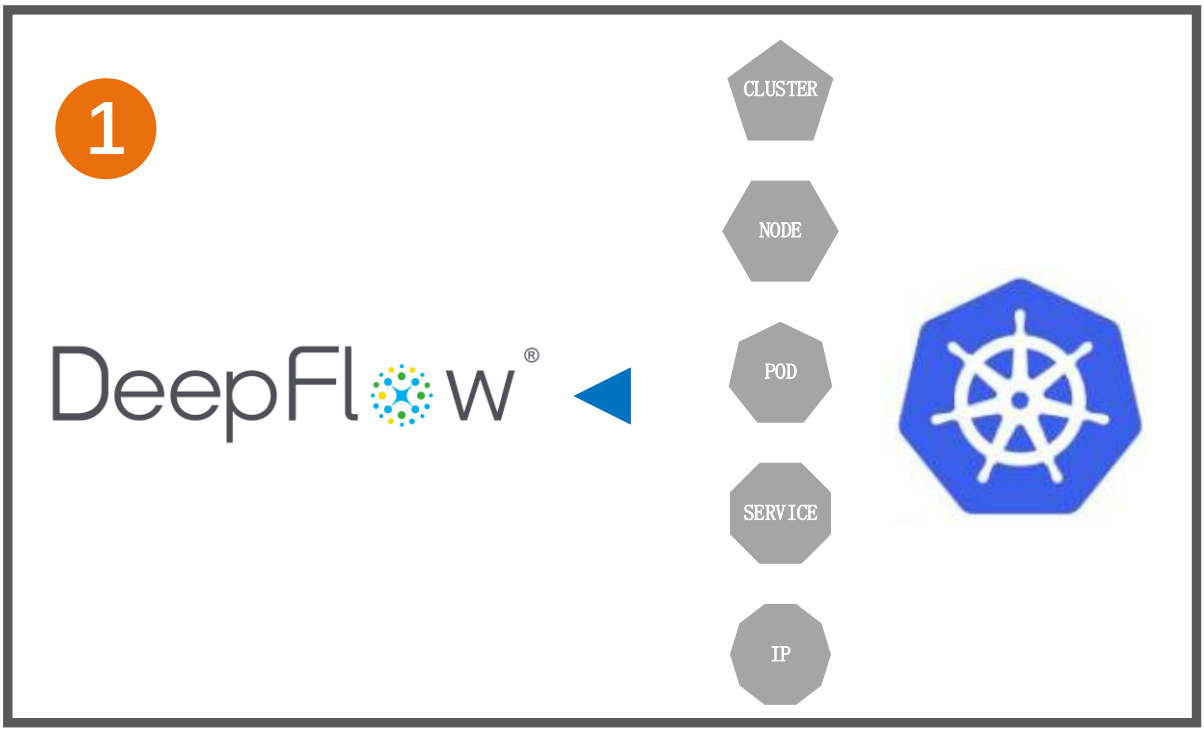
云原生系统发展迅速，大量企业已经将容器环境投入至生产，有效的容器网络保障是迫切的需求。本节重点结合 DeepFlow®客户处的应用实践，以 Kubernetes 环境为例，解决网络监控诊断的难题。

在容器环境部署及管理方案中，对于系统监控报警会更多地关注开源项目 Prometheus （https://prometheus.io），客户尝试其结合 Grafana（https://grafana.com/）、Zabbix（https://www.zabbix.com）以解决容器网络监控保障的难题，虽然涉及到一部分网络指标，但对于深入的网络需求，以及规模扩容后，对采集器控制、采集精度、关联分析等有更高的要求，以上是存在瓶颈的。

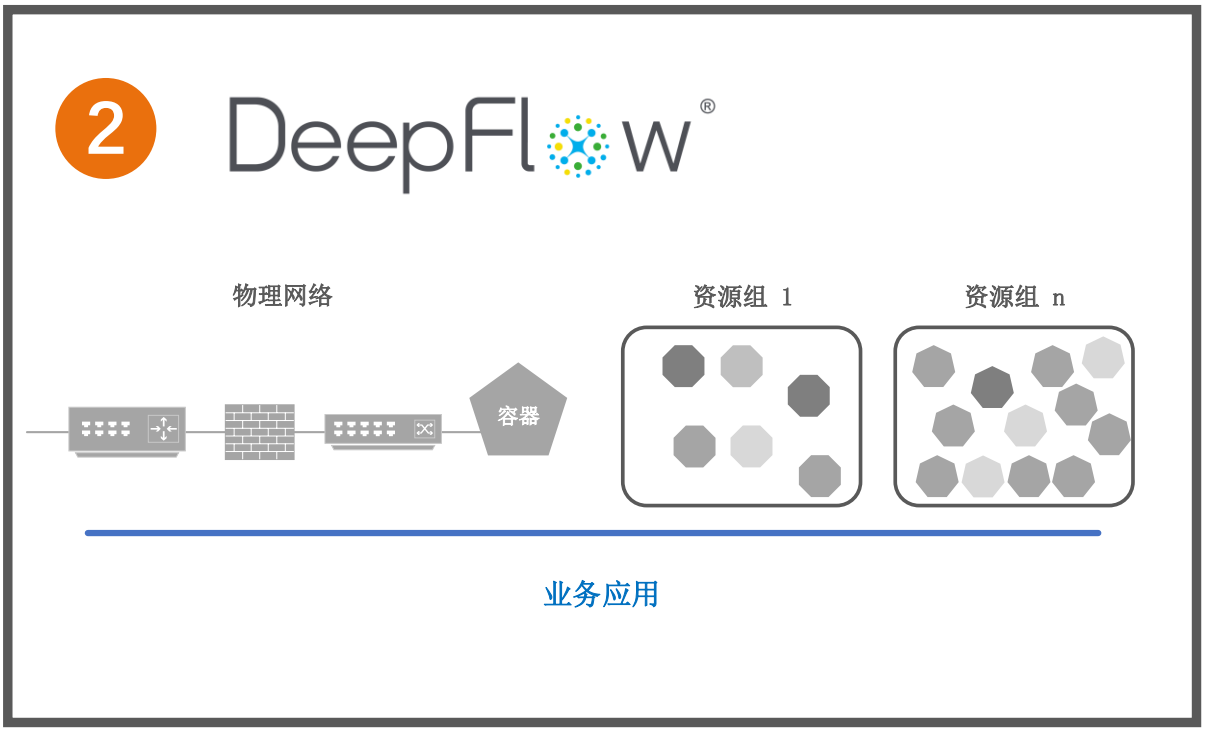


如上图所示，容器网络流量通过采集器进行获取，支持虚拟机，物理服务器做为计算节点的部署环境，物理网络流量由分光镜像获取，统一存入分布式时序数据库中。对于容器环境，以 Pod 为单元获取网络流量，获取全网流量数据，支持以秒粒度查询分析。单一控制器最大支持 2000 个采集点，控制器集群可扩展至 50 台主、备、从控制器。采集器部署在容器计算节点上，通常每个节点运行 100-200 个 Pod，进行流量获取及预处理。

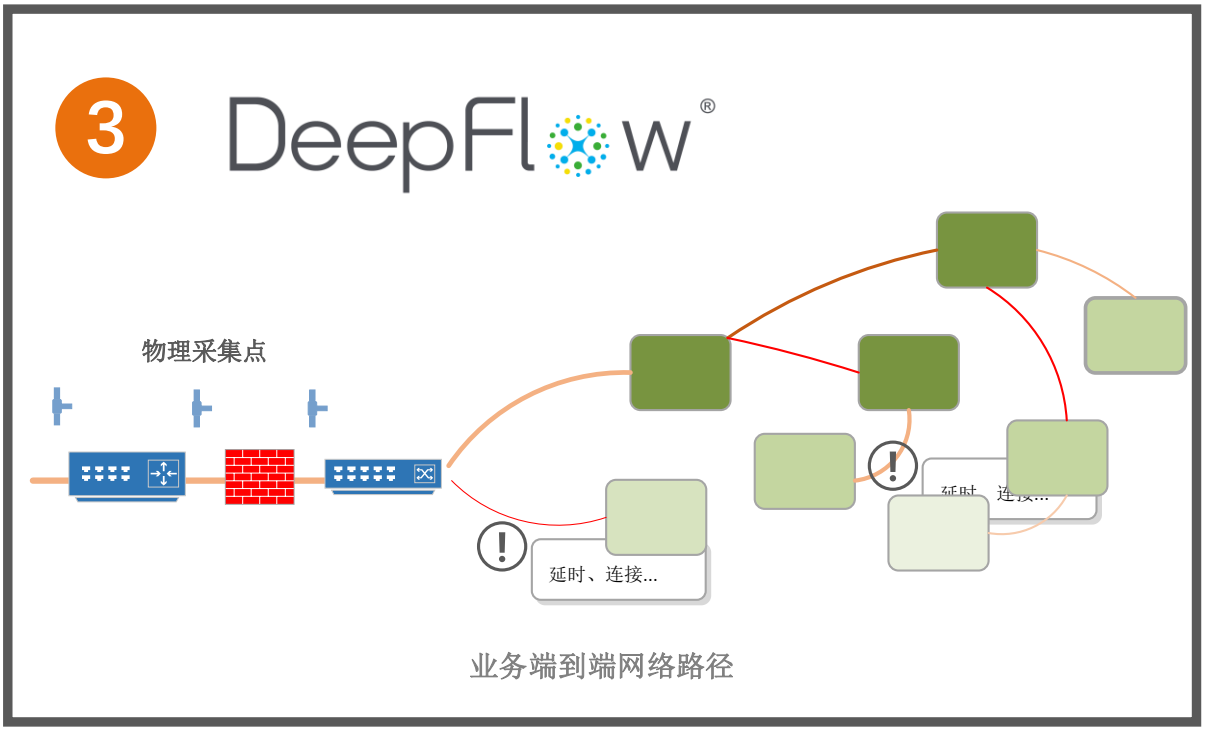
容器网络的监控诊断，最大的挑战是其固有的波动特性，全网的流量数据以及网络知识图谱保证对全网状态的可视化，此外针对重点业务应用，需要将其纳入视图进行持续关注。包含完整的容器资源、应用、网络几个维度才能完整绘制出业务的端到端网络路径。DeepFlow®平台通过对容器平台（如 Kubernetes）进行 API 对接，主动学习容器环境中的相关信息，包括集群（Cluster）、节点（Node）、Pod、服务（Service）、Ingress 等。



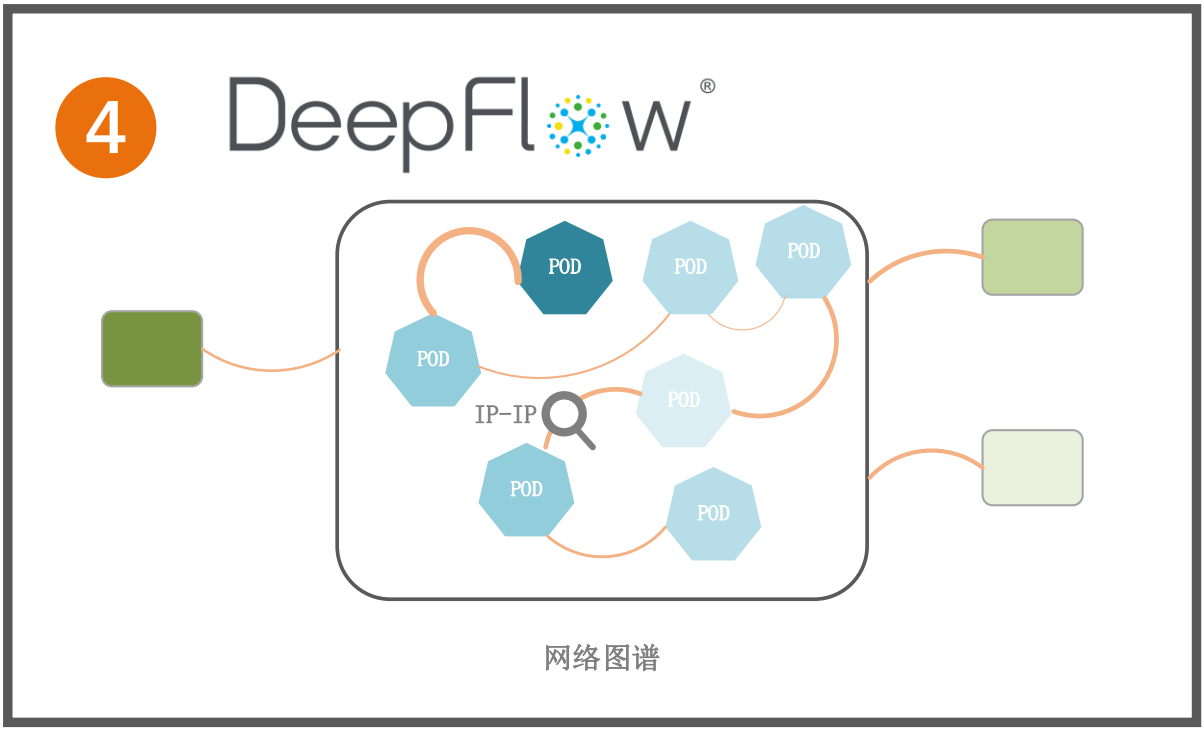
在“业务画像”功能中创建业务，并加入相关的资源组、链路，描述业务应用的网络访问路径。不同于传统物理网络中通过探针位置定位网络路径，在池化后的网络环境中，部署在节点上的采集点面向整个资源池，涵盖所有的 Pod 资源，而且，业务所涉及的 Pod 也并不是固定在某一节点之上的。“业务画像”通过“资源组”，归类 IP、功能服务、路径、链路等，来描述完整的业务应用所涉及的资源。



通过业务资源的描述，资源池内的流量将按此规则过滤，实现业务应用端到端访问的网络监控与诊断。 在整条路径中，分段排查网络状态，快速缩小问题范围，定位异常原因。



网络图谱中，区域、节点、POD、IP 等多维度的网络状态查询展示，不断缩小范围，回溯定位网络流、数据包进行分析取证。



网络图谱

部署

整体方案主要涉及采集器、控制器、数据节点三部分，在完成规划整体方案后，可分区域、分资源池按阶段投入建设，最终使企业具备对混合云基础设施全网络监控诊断能力，保障应用业务稳定运行。

由于大部分企业已经具备对传统物理网络的监控能力，通常重点选择 KVM、容器资源池网络为第一阶段建设，重点解决资源池内网络东西向流量“黑盒”不可见的问题，实现资源池网络可视化，提高运维排障效率，保证网络服务等级协议。

第二阶段纳入更多资源池，与新建扩容的资源池同步部署，接入物理网络中分光镜像流量，实现对整体数据中心网络监控。

第三阶段面向混合云中的公有云资源，对运行其上的网络进行监控，部署采集器，具备对 Workload 或其上容器流量采集能力，完成对混合云 IT 环境网络整体监控管理。

对于已经运行的混合云环境，可以在不影响生产环境运行的情况下部署实施，网络规划上将 DeepFlow®平台所涉及的管理、监控分发平面复用在已有的网络平面中，通常可以复用已经存在的网络管理平面。

对于整体规划的方案，建议对整体混合云规划独立的网络监控平面，对于混合云的监管流量统一、独立地进行管理。另外，采集器对计算能力的要求，可以根据处理流量、资源情况进行整体规划，对单一采集器最低可配置 1vCPU、128M 的资源使用。

方案优势

流量采集先进

DeepFlow®全网采集方案中，主要是以采集器技术实现流量采集，采集器支持 KVM、VMWare、容器等型号，以进程形态部署安装，最大程度上避免对生产交换平面的干扰，不存在与生产平面交换机流表冲突的风险，同时在操作系统上继承进程级保护优势，实现整体系统稳定。

分布式处理系统

资源池内东西向流量大，对网络数据包进行分布式预处理，避免集中单点且资源投入大，写入分布式高新能时序数据库，提升查询展示效率。

场景全规模大

整体方案是基于分布式设计模型以及多地域管理，可以充分保障资源池规模弹性扩展，整体系统可管理 10 万台采集器，涵盖虚拟化、容器、公有云资源池。

可管理性

平台主控制器是管理员统一集中的采集、分发策略配置、监控诊断交互入口，同时具备对所有采集器状态管理能力。各类操作贴近资源池特性，支持虚拟机名称、子网、集群、容器 POD 等多维度进行。资源存在迁移、回收、重新部署等场景，策略跟随保障采集能力在动态环境下的持续执行。平台整合物理网络、资源池内网络特点，为企业在云数据中心提供出众的网络监控诊断能力。

网络全景图

补齐企业对虚拟化、容器网络管理短板，落地资源池内网络监控诊断方案，关联网络、资源池、应用等维度，绘制网络图谱，提供运维有效手段，增强企业对新型基础设施网络管理能力。

总结

DeepFlow®混合云网络监控诊断解决方案为企业在混合云、云原生等新型 IT 基础设施环境演进过程中，补齐网络监控空白，应对云原生特点，紧密结合业务，向网络智能运维迈进。避免重复投入，重复安装，紧密结合应用业务，解决实际网络监控难题，也为企业规划整体运维、安全平台补齐资源池内网络保障这一缺失板块。本方案已应用于金融、运营商等客户 IT 环境中。

了解更多信息

专业的售前技术支持及商务合作，协助您选择最合适的解决方案

详询：400-9696-121

网址：www.yunshan.net

北京云杉世纪网络科技有限公司

北京市海淀区成府路 28 号优盛大厦 A 座 1209