

解决方案简介

英特尔® Smart Edge
云原生解决方案
5G
边缘计算

intel®

借力英特尔® Smart Edge，基于云原生解决方案的 灵雀云 ACP 5G 专网版本获得多维度优化加速



“灵雀云 ACP 5G 专网版本的构建，旨在通过引入 5G 网络能力，为各类边缘应用提供更为强大的、基于云边协同的处理、传输和分析平台。而英特尔® Smart Edge 以及其它英特尔软硬件产品与技术的加入，不仅为新版本提供了多样化的能力注入，也使 FEC、QAT 等关键功能获得了有效加速。”

陈恺
创始人兼 CTO
灵雀云

得益于云计算技术和 5G 网络的高速发展，越来越多的商用化智能应用正通过边缘部署的方式来解决其在处理性能、实时性和安全性等方面的高要求。而不断演进的云原生技术也为边缘节点提供了良好的编排、管理、运维和能力交付解决方案。

为了让基于“5G 网络 + 云原生”的融合方案更为高效，并为边缘场景中的应用提供经加速和优化的 5G 网络功能，云原生解决方案提供商灵雀云 (Alauda) 正与英特尔一起，通过引入英特尔® Smart Edge Open 软件工具套件，打造全新的灵雀云 ACP 5G 专网版本。

新的 5G 专网版本不仅集成了 5G 核心网与无线接入网服务、视频分析服务、存储服务以及零信任安全服务等多项能力，还基于多种硬件加速卡，配合英特尔® Smart Edge 软件工具套件，向 Kubernetes (以下简称 K8s) 提供了面向前向纠错 (Forward Error Correction, FEC)、动态设备个性化 (Dynamic Device Personalization, DDP)、QAT (Quick Assist Technology) 等技术的加速插件，为用户的边缘应用提供了面向“5G 网络 + 云原生”的可观能力加速。

背景概述: 更多边缘场景需要 5G 网络与云原生技术相互融合

随着云服务成为越来越多行业的下一代核心 IT 基础设施，一系列融合人工智能 (Artificial Intelligence, AI)、大数据、物联网等前沿技术的创新应用，如智能制造、元宇宙、数字地铁、数字养殖以及车联网等，正依托云服务在灵活性、可扩展性和成本等方面的优势获得快速发展，从而为企业的数字化转型和业务变革提供更多助力。

不断扩大的应用规模和不断深入的商用化程度, 使用户在应用的处理效率、实时性、可靠性及安全性等方面也有了更高的要求。而通过分布式架构等方式, 将各类云服务能力由中心下沉到边缘, 并通过“云边协同”模式来提供高效处理能力, 从而最大程度降低时延、提升实时性的边缘计算解决方案正获得越来越多的关注。

与此同时, 新一代的 5G 网络也在边缘计算解决方案中起到更为重要的作用, 除了能为应用提供高品质的网络服务能力之外, 其开放、灵活的网络架构也为其提供了良好的网络“底座”。各类充分解耦的 5G 网络功能, 如集中控制单元 (Centralized Unit, CU) / 分布式单元 (Distributed Unit, DU)、接入和移动性管理功能 (Access and Mobility Management Function, AMF)、会话管理功能 (Session Management Function, SMF) 以及用户平面功能 (User Plane Function, UPF) 等网元, 都可以下沉部署至应用场景附近, 为应用就近提供数据上下行传输、存储分析、安全认证等能力。

5G 网络带来的性能提升, 以及更多云服务能力向边缘的下沉, 也意味着用户需要有效解决中心节点与众多边缘节点之间的统一编排、集群管理、快速交付、边缘自治以及弹性伸缩等问题。而日益受到认可的云原生技术则是解决以上问题的有效手段。通过容器、微服务等云原生技术, 用户可以大幅简化边缘节点的构建工作, 轻松实现对边缘应用的统一管理、对边缘集群资源的集中调度, 达成应用的快速开发和交付。

可以看到, 5G 网络与云原生技术都将在未来边缘计算场景中起到不可或缺的作用, 但要将在 5G 网络与云原生技术有机融于一体并不容易。与中心节点相比, 边缘节点在处理性能上面临的性能压力更大、资源分配上受到的限制更广、潜在的安全隐患也更多。尤其在具有高带宽、广连接和低时延特性的 5G 网络服务中, 传统云服务平台所提供的能力无疑也将面临巨大的挑战。

因此两者的融合需要业内领先企业的强强联手, 发挥各自优势, 通过协同创新来推进。作为企业级云原生解决方案提供商, 灵雀云拥有全球领先的超大规模云原生平台开发、运维和管理经验, 同时也是中国最早将 Kubernetes 产品化的厂商之一。如图 1 所示, 基于完备的云原生技术生态体系和丰富的云原生方案积累, 灵雀云正通过旗下企业级云原生全栈私有云平台 ACP, 助力用户获得更为完善的解决方案和更为优质的全栈服务。

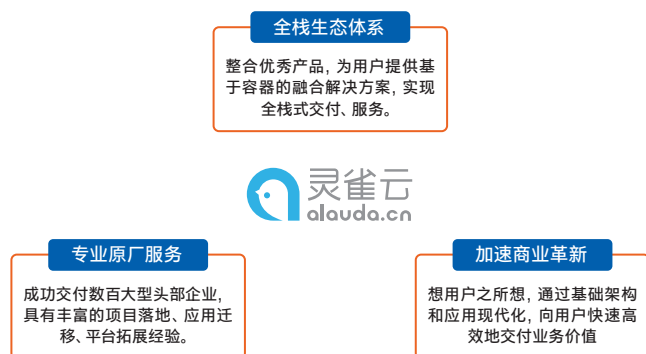


图 1 为用户提供云原生全栈解决方案的灵雀云

在灵雀云看来, 传统面向边缘部署的云服务组件中, 往往缺乏为边缘集群提供的用于在单个编排域下部署 5G 网络功能的工具。而在 5G 网络和边缘计算技术领域都具有雄厚实力的英特尔, 无疑可补齐这一关键环节。为此, 灵雀云正与英特尔一起, 借助英特尔® Smart Edge 以及相应的硬件加速卡, 实施“5G 网络 + 云原生”融合策略, 打造可部署于一系列边缘计算应用场景的灵雀云 ACP 5G 专网版本。

解决方案: 集成英特尔® Smart Edge, 灵雀云 ACP 5G 专网版本获得多维度优化与加速

英特尔® Smart Edge 是英特尔为了更快更低成本打造边缘方案而推出的软件套件, 可以简化边缘端网络配置、边缘设备使用 and 应用程序部署的复杂性, 它加快边缘解决方案的开发。而这些边缘解决方案将网络与人工智能、媒体处理和安全工作负载集成在一起, 并为认证 Kubernetes 云原生堆栈支持的常见用例提供优化的参考解决方案。

Intel® Smart Edge Architecture

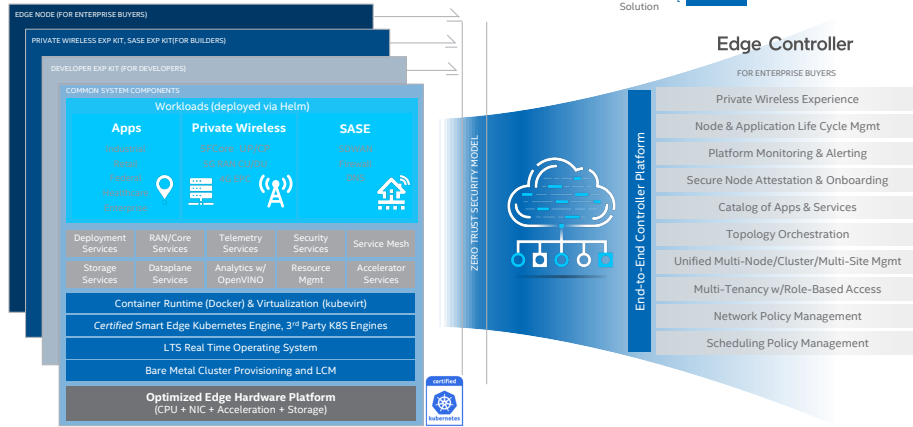


图 2 英特尔® Smart Edge 架构图

集成英特尔® Smart Edge 模块的灵雀云 ACP 5G 专网版本如图 3 所示。整个方案基于云原生技术构建,可分为 ACP 云端平台和 ACP 边缘节点两部分。其中,云端平台拥有统一云边应用生命周期管理、分布式异构边缘节点统一管控、5G 核心网控制面、统一运维与管理等组件,用于实现对边缘节点集中监测、统一管控和统一运维。用户可基于云端平台来进行边缘节点镜像的集中管理和下发,从而实现网络和边缘业务的统一部署,提升应用的发布、管理和运维效率。

而在以英特尔® 架构处理器、以太网网络适配器、存储设备以及各类硬件加速器为基础构建的各个边缘节点中,灵雀云也为其提供了边缘应用自治、5G 无线接入网、5G UPF 网元、站点自动注册 / 断线重连与切换以及弹性伸缩等组件,用于保障一系列应用,如智能制造、车联网等在边缘环境的高效运行,即便出现云边断网也不会造成业务中断,同时还能实现弹性伸缩、故障自愈等边缘自治能力。

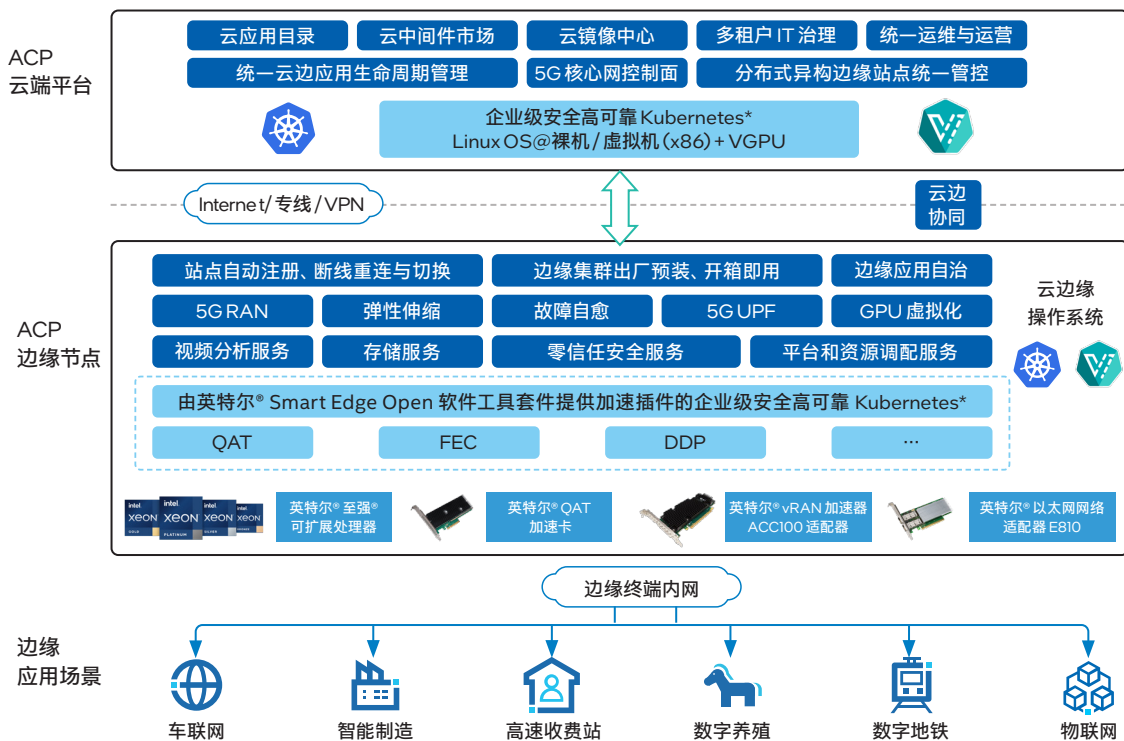


图 3 集成英特尔® Smart Edge 模块的灵雀云 ACP 5G 专网版本架构

云端平台与众多边缘节点之间通过互联网、专线或 VPN 连接起来, 构建出良好的云边协同能力, 包括:

- 高速分发用于边缘侧网络功能和边缘业务部署的镜像;
- 边缘应用批量下发部署、配置、更新与管控;
- 边缘管理、运维数据的实时上报, 对多组网环境的适配。

为了让新方案在边缘场景中为用户应用提供更多处理分析、数据存储、安全可信等方面的优势技术支持, 灵雀云也与英特尔一起引入诸多软硬件技术, 为新方案提供了一系列服务能力, 包括:

- **视频分析服务:** 能提供基于通用处理器处理或加速, 并由英特尔® 发行版 OpenVINO™ 工具套件提供支持的视频分析微服务;
- **存储服务:** 基于英特尔® 架构, 为在单个或集群边缘节点上的原生边缘应用提供高吞吐、低时延的数据存储能力;
- **零信任安全服务:** 对任意状态的数据, 都能提供完备的平台认证、完整性保护和可信执行能力, 并保证设备登录时的安全性和数据隐私;
- **平台和资源调配服务:** 在 BIOS 和固件中提供底层平台的配置和管理能力, 可在裸金属或虚拟机中自动安装操作系统和软件堆栈;
- **5G 核心网与无线接入网服务:** 可动态重配置 5G 核心网与无线接入网, 为 5G 无线接入网的 IO、功能加速、编排和功能提供支持;

除上述服务能力之外, 更为重要的是, 英特尔还注重与灵雀云的云原生技术能力相结合, 向 K8s 提供了各类面向 FEC、DDP、QAT 等技术的加速插件, 包括:

- 英特尔® vRAN 加速器 ACC100 适配器及 SR-IOV 前向纠错操作器 (SR-IOV FEC Operator);
- 英特尔® QAT 加速卡及英特尔® QAT 技术;

- 英特尔® 以太网网络适配器 E810 (以下简称“E810 网卡”) 及英特尔® DDP 技术。

上述产品与技术的加入, 能让边缘应用可以轻松、便捷地在基于 K8s 的云原生解决方案中调用这些能力, 从而为用户提供了面向“5G 网络 + 云原生”的能力优化与加速。

■ FEC 能力卸载

在边缘部署的网络功能中, 有很多属于密集型计算处理负载, 典型如 DU 网元中的 FEC 功能。FEC 功能是 5G 网络中重要的数据传输差错控制方式, 其原理是数据在发送时加入一定冗余信息后编码传输, 接收时解码并进行检查, 如发现错误则重新传输。

如使用通用处理器来执行 FEC 功能所需的加解密、错误信息检查等工作负载, 当业务量急剧增加时, FEC 负载就会大量占用处理器资源而降低网元的整体工作效率。而使用专门的硬件加速器来将 FEC 功能从处理器中卸载可以有效应对这一问题。但传统的 K8s 等云原生解决方案, 通常不对此类硬件加速器提供“即开即用”支持, 用户通常需要在平台上进行复杂而繁琐的配置过程, 才能让应用成功调用到上述硬件加速器。

为此, 英特尔为新版本方案提供了英特尔® vRAN 加速器 ACC100 适配器与 SR-IOV FEC Operator 的软硬件组合, 来助力用户获得有效的 FEC 能力卸载。在英特尔® Smart Edge Open PWEK (Private Wireless Experience Kit) 中, SR-IOV FEC Operator 可以协调和管理 K8s 集群中的一系列英特尔 vRAN FEC 硬件加速器 (如英特尔® vRAN 加速器 ACC100 适配器) 所提供的加速资源, 其能以云原生的方式, 将这些加速资源绑定到数据平面开发套件 (Data Plane Development Kit, DPDK) 驱动程序中的 SR-IOV 虚拟功能 (Virtual Functions, VF), 使之可作为 K8s 的自定义资源被边缘应用便捷、灵活地发现和调用。

■ 加解密 / 压缩能力加速

作为一种更为安全、可靠的无线通信技术, 5G 网络的数据交互涉及大量的加解密和解压缩工作负载。通常而言, 这些工作负载都消耗巨大, 以 5G 网络常用的 GTP-U (User Plane Part of GTP, GPRS 用户平面部分) 协议为例, 对其进行封包、解包就需要消耗大量的处理器资源。

作为领先的数据加解密、压缩和安全技术, 英特尔® QAT 技术能有效增强各个应用场景中数据的安全性和压缩性能。目前已有多款英特尔® 架构处理器和板卡设备中内置了英特尔® QAT 技术 (如英特尔® 至强® 可扩展处理器家族、英特尔® 至强® 处理器 D 家族等), 同时英特尔还提供了专门的英特尔® QAT 加速卡产品 (如英特尔® QAT 加速卡 8960 / 8970 等)。

为了让用户能在灵雀云 ACP 5G 专网版本中方便地调用英特尔® QAT 加速卡, 英特尔® Smart Edge Open 软件工具套件也为其提供了适配能力。这支持新版本方案首先通过云原生技术, 基于 K8s 提供的设备插件框架, 来为 CU / DU 等网元部署英特尔® QAT 加速卡的插件, 使加速卡能为边缘节点所用。

当边缘节点的英特尔® QAT 加速卡资源可用后, 新方案进一步通过其提供的 3 个 SR-IOV 物理功能 (PF) 来创建所需虚拟功能 (VF), 即将加速卡的 PF 绑定到 DPDK IGB_UIO 用户空间驱动程序。然后创建多个 VF 并绑定到 DPDK 用户空间驱动程序, 并最终将各个用于加速的 VF 分配到不同的边缘应用工作负载中去。

■ 提供数据转发性能加速

数据转发也是灵雀云 ACP 5G 专网版本中各个 5G 网元的核心功能之一, 为有效应对未来边缘应用中的海量数据转发需求, 新方案也引入 E810 网卡、英特尔® DDP 技术等来提供加速能力。

作为英特尔® 以太网网络适配器 800 系列的最新产品, E810 网卡不但能提供高达 200 Gbps 数据吞吐能力 (单一适配器), 更能灵活地支持新的协议或用户专有协议, 为丰富的边缘应用提供来自英特尔的应用设备队列 (Application Device Queues, ADQ)、支持 iWARP 或 RoCEv2 协议的远程直接内存访问 (Remote Direct Memory Access, RDMA) 以及 DDP 等先进技术。

这其中, 英特尔® DDP 技术能通过加载固件配置文件 (profile) 来实现对特定通信网络协议, 如 PPPOE、GTP-U 以及 L2TP 等的解析, 按需重配置报文处理引擎, 从而让特定场景中的数据处理变得更为高效。

例如在 5G UPF 网元中, 处理的主要报文类型是 GTP (上行) 报文和 TCP / IP (下行) 报文。但由于 GTP 报文的内层报头无法被普通网卡所解析, 造成 GTP 报文不能利用英特尔® 以太网 Flow Director (FDIR) 或 Receive Side Scaling (RSS) 技术来将散列 (Hash) 数据分流到不同队列, 并绑定到不同处理器核上执行并行处理。

为此, 处理器就需要额外使用一部分资源来完成上述 GTP 报文的内层报头解析工作, 这无疑会增加系统的整体开销。而内置英特尔® DDP 技术的 E810 网卡, 则能帮助新方案将这一工作负载从处理器卸载到网卡, 实现分发功能的硬件卸载, 在大幅提高性能的同时也降低了系统的整体处理时延。

收益与展望

为验证集成英特尔® Smart Edge 模块的灵雀云 ACP 5G 专网版本在面向边缘的部署实践时, 能否获得预期的效果, 灵雀云与英特尔一起在多种边缘场景中对新方案开展了多方位且缜密的测试。

在面向 5G 专网接入性能的验证测试中, 5G 核心网和接入单元 (Access Unit, AU) 设备的核心配置如下:

- **5G 核心网设备:** 具备 12 个核心的英特尔® 架构处理器、8Gi 内存、4Gi 的大内存页;
- **AU 设备:** 具备 40 个核心的英特尔® 架构处理器、24Gi 内存、20Gi 的大内存页;

基于以上配置, 5G 专网的往返时延仅为 8 毫秒¹, 峰值吞吐量可达理论值的 94% 以上², 端到端连接可保持 2 天以上且在服务器重启后可自动恢复。在时隙配比为 8D2U_644 的下行优先配置中, 上行速率能达到 1667Mbps。而在时隙配比为

1D3U 的上行优先配置中, 上行速率可达 655Mbps。这一结果令人满意。

实践证明新的 5G 专网版本方案在包括边缘 AI 应用、视频通话在内等多个场景中都可获得良好的性能表现, 达成了双方合作的预期。

面向未来, 灵雀云与英特尔还计划在云原生、5G 网络以及边缘计算等领域继续展开深入的技术合作, 将更多维度和层面的技术方案融入灵雀云云原生解决方案中, 从而为用户打造更为丰富、更具系统化的云服务能力, 使之成为企业数字化革新的助力。



^{1,2} 测试组 处理器: 双路英特尔® 至强® 金牌 6338N 处理器, 96 内核 / 128 线程, 主频 2.20GHz, 动态加速频率 3.50GHz, 超线程开启, 睿频关闭; 内存: 251GB (16GB 3200MT/s DDR4); 网络适配器: 10GbE 英特尔® 以太网适配器 X710-DA2; 存储: 英特尔® 固态硬盘 D3-S45101.92TB, 操作系统: CentOS 7.9, 内核版本: 3.10.0-1127.19.1.rt56.1116.el7.x86_64

实际性能受使用情况、配置和其他因素的差异影响。更多信息请见 www.Intel.com/PerformanceIndex

本文并未 (明示或默示、或通过禁止反言或以其他方式) 授予任何知识产权许可。英特尔未做出任何明示和默示的保证, 包括但不限于, 关于适销性、适合特定目的及不侵权的默示保证, 以及在履约过程、交易过程或贸易惯例中引起的任何保证。

英特尔运营所需的任何商品和服务预测仅供讨论。就与本文中公布的预测, 英特尔不负有任何购买责任。本文中提供的所有信息可在不通知的情况下随时发生变更。关于英特尔最新的产品规格和路线图, 请联系您的英特尔代表。

英特尔技术特性和优势取决于系统配置, 并可能需要支持的硬件、软件或服务得以激活。产品性能会基于系统配置有所变化。没有任何产品或组件是绝对安全的。更多信息请从原始设备制造商或零售商处获得, 或请见 intel.com

在特定系统的特殊测试中测试组件性能。硬件、软件或配置差异将影响实际性能。当您考虑采购时, 请查阅其他信息来源评估性能。关于性能和基准测试程序结果的更多信息, 请访问 www.intel.com/benchmarks

英特尔并不控制或审计第三方数据。请您自行审核该内容、咨询其他来源, 并确认提及数据是否准确。

英特尔、英特尔标识以及其他英特尔商标是英特尔公司或其子公司在美国和/或其他国家的商标。文中涉及的其它名称及品牌属于各自所有者资产。