原来链接：https://www.openstack.org/assets/edge/OpenStack-EdgeWhitepaper-v3-online.pdf

openstack

CLOUD EDGE COMPUTING：

Beyond the Data Center

云边缘计算：

不止于数据中心

原来链接：https://www.openstack.org/assets/edge/OpenStack-EdgeWhitepaper-v3-online.pdf

Table of Contents 目录

Introduction 4
概述 4
What is Cloud Edge Computing? 7
什么是云边缘计算？ 7
A Deeper Exploration of Edge Computing Considerations 9
边缘计算的深入探索 9
Exploring Characteristics 14
特征 14
Use Cases 16
用例 16
Scenarios 21
场景 21
Challenges 25
挑战 25
Conclusion and Call to Action 30
结论和呼吁 30

原文贡献者：
CONTRIBUTORS：

Beth Cohen, Distinguished Member of Technical Staff, Verizon
Gnanavelkandan Kathirvel, Director – Cloud Strategy & Architecture, AT&T and Board of Directors, OpenStack Foundation (OSF)
Gregory Katsaros, Senior System Architect, Inmarsat
Adrien Lebre, Ass. Prof., IMT Atlantique/Inria/LS2N, France
Andrew Mitry, Sr. Distinguished Engineer, Walmart Labs
Christopher Price, President, Ericsson Software Technology
Paul-André Raymond, SVP Technology, B.Yond
Alex Reznik, Enterprise Architect, HPE and Chair, ETSI ISG MEC
Pasi Vaananen, Systems Architect, NFV, Red Hat
Ildiko Vansca, Ecosystem Technical Lead, OpenStack Foundation (OSF)
Ian Wells, Distinguished Engineer, Cisco

TECHNICAL WRITER
Brian E. Whitaker, Founder, Zettabyte Content

中文翻译：肖　晴　绿盟科技

中文校审：江国龙　绿盟科技

　　　　　刘文懋　绿盟科技

Introduction

概述

For over a decade, centralized cloud computing has been considered a standard IT delivery platform. Though cloud computing is ubiquitous, emerging requirements and workloads are beginning to expose its limitations.

十多年来，集中式的云计算一直被认为是标准的 IT 交付方式。尽管云计算的普及程度很高，但新兴的需求和工作负载开始暴露出其局限性。

With its strong data center centric view, where compute and storage resources are relatively plentiful and centralized, little or no thought was ever given to the optimization of the supporting hypervisor and management platform footprint.

由于云计算这种模式会将数据集中存储到数据中心，而数据中心里的计算和存储资源相对丰富和集中，所以几乎没有人考虑对 Hypervisor 和管理平台进行优化。

Few cloud developers seriously considered the requirements needed to support resource-constrained nodes reachable only over unreliable or bandwidth-limited network connections, or thought about the needs of applications that demand very high bandwidth, low latency, or widespread compute capacity across many sites.

很少有云计算的开发人员认真考虑，在仅可以通过不可靠或者带宽受限的网络连接情况下，支持受限资源的访问问题；或者很少考虑那些需要高带宽、低延迟或跨多个站点的广泛计算资源的应用程序的需求。

New applications, services, and workloads increasingly demand a different kind of architecture, one that's built to directly support a distributed infrastructure.

新的应用、服务和工作负载越来越需要全新的架构进行支撑，这种架构需要能够直接支持分布式的基础设施资源。

New requirements for availability and cloud capability at remote sites are needed to support both today's requirements (retail data analytics, network services) and tomorrow's innovations (smart cities, AR/VR).

为了支持当今的需求（零售数据分析、网络服务）和未来的创新应用（智能城市、AR/VR），这些都对远程站点的可用性和云能力提出新的要求。

The maturity, robustness, flexibility, and simplicity of cloud now needs to be extended across multiple sites and networks in order to cope with evolving demands.

因此，云计算需要在其成熟性、健壮性、灵活性和简单性的基础上，实现跨越多个站点和网络的扩展，以应对不断变化的需求。

Recently companies have begun to apply the simplified administration and flexibility of cloud computing architectures to distributed infrastructures that span across multiple sites and networks.

近期,很多公司已开始将云计算的这种简单和灵活的架构应用于跨越多个站点和网络的分布式基础设施中。

Organizations have an emerging need to take cloud capabilities across WAN networks and into increasingly smaller deployments out at the network edge. Though this approach is in its early days, it is becoming clear that many emerging use cases and scenarios would benefit from distributed architectures.

很多企业希望能够通过 WAN 网络实现云计算的功能，并在网络边缘实现越来越轻量的部署。虽然这种方式还处于初期阶段，但很明显，许多新的应用和场景都将从分布式计算架构中受益。

In this paper, we explore this emerging need. It has been called many names: distributed cloud, fog computing, 4th generation data centers, but for the purposes of this document, we will stick with a common, easily understood term—cloud edge computing.

在本文中，我们会来探讨这个新兴的需求。它有许多名字：分布式云计算、雾计算、第四代数据中心，但在本文中，我们将使用一个常见的，易于理解的术语——云边缘计算。

The OSF Edge Computing Group sees this evolution of cloud computing as very exciting, though we recognize that OpenStack's ability to support cloud edge computing is in its early days.

尽管我们认为 OpenStack 在支持云边缘计算上的能力还处于早期阶段，但 OSF 边缘计算小组非常看好这种云计算的发展方式。

Based on initial community interest expressed at the OpenStack Summit Boston, a two-day workshop was held in September 2017 that brought together over 200 users and developers to start the hard work of defining relevant use cases and considering the tools and architectures needed to support them.

基于社区在波士顿 OpenStack 峰会上表达的初步兴趣，2017 年 9 月举办了为期两天的研讨会，这次研讨会汇集了 200 多名用户和开发人员，开始艰难地定义相关用例并考虑支持这些用例所需的工具和架构。

Proof of concepts have been done and the community has a few early deployments in place. The OSF Edge Computing Group has now taken up the challenge to describe fundamental requirements of a fully functional edge computing cloud infrastructure.

社区已经完成了概念验证，并且有了一些早期部署。OSF 边缘计算小组现在已经准备好了构建全功能云边缘计算基础设施的挑战。

In this document, we aim to accomplish several important tasks:

在本文中，我们旨在完成几项重要任务：

1. Cultivate a conversation around cloud edge computing, including some basic definitions, stimulating interest and engagement from the open source community.

   围绕云边缘计算展开，包括一些基本定义，激发开源社区的兴趣和参与度。

2. Guide both the broader open source and OpenStack communities in developing tools and standards needed for broad adoption.

   引导更广泛的开源和 OpenStack 社区，来开发所需的工具和标准。

3. Explore how current tools, standards and architectures may need to change to accommodate this distributed cloud model.

   探索如何通过改变当前的工具、标准和架构，来适应这种分布式云计算模型。

There is much work to be done to achieve our goals, and we welcome and encourage the entire open source community to join in both the effort and the opportunity of creating or adapting tools to meet the new requirements of cloud edge computing.

为实现我们的目标需要做许多工作，我们欢迎并鼓励整个开源社区的加入，以满足云边缘计算的新需求。

What is Cloud Edge Computing?

什么是云边缘计算？

It is worth highlighting that many overlapping and sometimes conflicting definitions of edge computing exist—edge computing means many things to many people. But for our purposes, the most mature view of edge computing is that it is offering application developers and service providers cloud computing capabilities, as well as an IT service environment at the edge of a network.

值得强调的是，目前边缘计算存在许多重叠甚至互相矛盾的定义，不同的人对边缘计算对有着不同的理解。但就我们而言，我们希望描述出一个对边缘计算最成熟的说法：它为应用开发人员和服务提供商提供云计算，以及位于网络边缘的 IT 环境。

The aim is to deliver compute, storage, and bandwidth much closer to data inputs and/or end users. An edge computing environment is characterized by potentially high latency among all the sites and low and unreliable bandwidth—alongside distinctive service delivery and application functionality possibilities that cannot be met with a pool of centralized cloud resources in distant data centers.

其目标是使计算、存储和网络资源更接近数据产生端或最终用户。边缘计算环境的特点是所有站点之间可能存在高延迟、低速、网络不可靠，伴随着云端的集中式资源池计算方式无法满足特定的服务交付和应用功能。

By moving some or all of the processing functions closer to the end user or data collection point, cloud edge computing can mitigate the effects of widely distributed sites by minimizing the effect of latency on the applications.

因此通过将部分或全部处理程序，转移到最终用户或数据收集点，减少应用之间数据计算和传输的延迟，云边缘计算可以有效减轻广泛分布站点所带来的影响。

Edge computing first emerged by virtualizing network services over WAN networks, taking a step away from the data center. The initial use cases were driven by a desire to leverage a platform that delivered the flexibility and simple tools that cloud computing users have become accustomed to.

边缘计算首先通过在 WAN 网络中的虚拟化网络服务里实现与数据中心的分离。最初的应用案例是希望提供给云计算用户一个熟悉的架构，从而建立一个灵活简单的平台工具。

As new edge computing capabilities emerge, we see a changing paradigm for computing—one that is no longer necessarily bound by the need to build centralized data centers.

随着新的边缘计算能力的出现，我们看到了计算模式的变化——不再需要构建集中的数据中心。

Instead, for certain applications, cloud edge computing is taking the lessons of virtualization and cloud computing and creating the capability to have potentially thousands of massively distributed nodes that can be applied to diverse use cases, such as industrial IoT or even far-flung monitoring networks for tracking real time water resource usage over thousands, or millions, of locations.

相反，对于某些应用来讲，云边缘计算正在借鉴虚拟化和云计算的经验，并有可能创建数千个可应用于各种用例的大规模分布式站点的能力，例如工业物联网，甚至是用于跟踪数千甚至数百万个地点的水资源使用情况的实时监控网络。

Many proprietary and open source edge computing capabilities already exist without relying on distributed cloud—some vendors refer to this as “device edge.” Components of this approach include elements such as IoT gateways or NFV appliances.

现在许多专有和开源边缘计算功能可以不依赖于分布式云而存在——一些供应商将此称为“边缘设备”。此方法的组件包括诸如：物联网网关或 NFV 设备等。

But increasingly, applications need the versatility of cloud at the edge, although the tools and architectures needed to build distributed edge infrastructures are still in their infancy. Our view is that the market will continue to demand better capabilities for cloud edge computing.

但随着时间的推移，应用程序需要云端具备多功能性，尽管构建分布式边缘基础设施所需的工具和架构仍处于初级阶段。但我们认为，市场将会持续对云边缘计算的能力提出更高的要求。

Edge computing capabilities include, but are not limited to:

边缘计算的功能包括但不限于：

- A consistent operating paradigm across diverse infrastructures.

  贯穿不同基础设施的一致性操作模式

- The ability to perform in a massively distributed (think thousands of global locations) environment.

  支持大规模分布式（例如全球数千个地点）环境的能力

- The need to deliver network services to customers located at globally distributed remote locations.

  向位于全球分布式远程位置的客户提供网络服务

- Application integration, orchestration and service delivery requirements.

  应用集成、编排和服务交付

- Hardware limitations and cost constraints.

  硬件限制和成本限制

- Limited or intermittent network connections.

  有限或间歇的网络连接

- Methods to address applications with strict low latency requirements (AR/VR, voice, and so forth).

  AR/VR、语音等应用的严格低延迟要求

- Geofencing and requirements for keeping sensitive private data local.

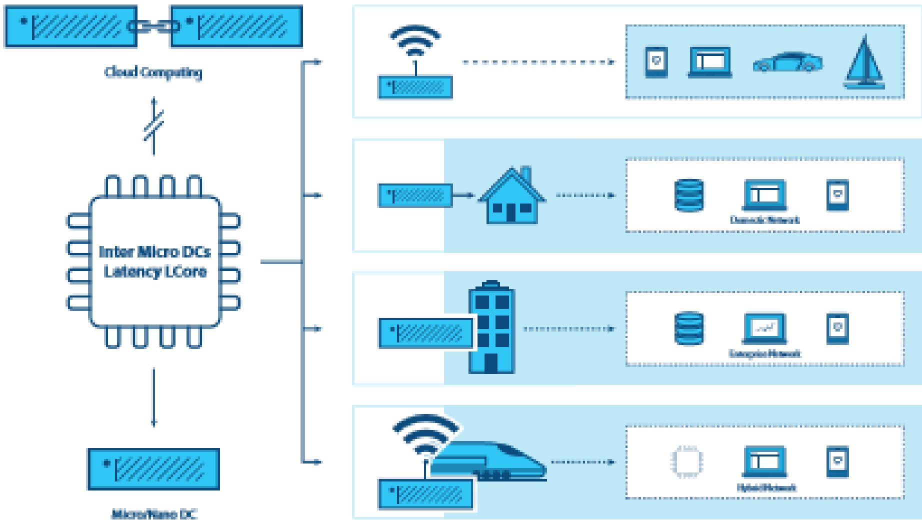
  对敏感私人数据进行本地保护的要求

A Deeper Exploration of Edge Computing Considerations

边缘计算的深入探索

The "edge" in edge computing refers to the outskirts of an administrative domain, as close as possible to discrete data sources or end users. This concept applies to telecom networks, to large enterprises with distributed points of presence such as retail, or to other applications, in particular in the context of IoT.

边缘计算中的"边缘"指的是管理域的边界，尽可能接近离散数据源或最终用户。这一概念适用于运营商网络、分布式存在的大型企业，如零售业或者是物联网领域的其他应用。



One of the characteristics of edge computing is that the application is strongly associated with the edge location.

边缘计算的特点之一是应用程序与边缘位置紧密相关。

For telecoms, “the edge” would refer to a point close to the end user but controlled by the provider, potentially having some elements of workloads running on end user devices. For large enterprises, “the edge” is the point where the application, service or workload is used (e.g. a retail store or a factory).

对于运营商来讲，“边缘”指的是靠近最终用户但是由运营商控制的站点，可能会在用户终端设备上运行一些负载。对于大型企业来说，“边缘”是一个正在运行的应用程序、服务或工作负载的地方（例如零售商店或工厂）。

For the purposes of this definition, the edge is not an end device with extremely limited capacity for supporting even a minimal cloud architecture, such as an IoT or sensor device. This is an important consideration, because many discussions of edge computing do not make that distinction.

从这个定义来看，边缘不是容量非常有限、甚至不足以支持一个极小云架构的终端设备，如物联网或传感器设备。这需要重点关注，因为许多关于边缘计算的讨论没有对此区分。

Edge computing is similar to data center computing in that:

边缘计算与数据中心计算的相似之处在于：

- It includes compute, storage and networking resources.

它们都包括计算、存储和网络资源

- Its resources may be shared by many users and many applications.

  它们的资源可以被许多用户和许多应用程序共享

- It benefits from virtualization and abstraction of the resource pool.

  它们都得益于虚拟化技术和抽象的资源池

- It benefits from the ability to leverage commodity hardware.

  它们都得益于商用硬件

- It uses APIs to support interoperability.

  它们都使用 API 来支持互操作性

Edge computing differs from computing in large data centers in that:

边缘计算与大型数据中心的计算不同之处在于：

- Edge sites are as close as possible to end users. They improve the experience over high latency and unreliable connections.

  边缘站点尽可能靠近最终用户。在高延迟和不可靠网络连接情况下提升了用户体验。

- May require specialized hardware, such as GPU/FPGA platforms for AR/VR functionality.

  可能需要专用硬件，例如：用于提高 AR/VR 功能的 GPU/FPGA 平台。

- Edge can scale to large numbers of sites, distributed in distinct locations.

  边缘可以扩展到大量的站点，分布在不同的位置。

- An edge site's location and the identity of the access links it terminates are significant. An application that needs to run close to its users, needs to be in the right part of the edge. It is common for the application location to matter in edge computing.

  边缘站点的位置以及它终止访问认证链接非常重要。应用程序需要运行在靠近用户边缘的正确位置。应用程序运行位置的重要性，在边缘计算中具有普遍性。

- The entire pool of sites can be considered to be dynamic. Because of their physical separation, edge sites will, in some cases, be connected to each other and the core with WAN connections. Edge sites will join and leave the pool of infrastructure over time.

  整个站点池应该是动态的。因为它们在物理上是分离的，在某些情况下，边缘站点之间、或者边缘站点与核心站点之间通过 WAN 相互连接。随着时间的推

移，边缘站点可以随时加入或离开资源池。

- Edge sites are remote and potentially unmanned, and therefore must be administered remotely. Tools need to support intermittent network access to the site.

  边缘站点通常是位于远端，可能无法被人操控，因此必须能够对其进行远程管理。管理工具需要能够通过不稳定的网络对边缘站点进行访问。

- Edge supports large differences in site size and scale, from data center scale down to a single device.

  边缘站点支持在大小和规模上的巨大差异，大到数据中心、小到单个设备。

- Edge sites may be resource constrained; adding capacity to an existing site is restricted due to space or power requirements.

  边缘站点通常会受到资源限制；由于空间或电力要求，现有站点中添加资源往往会受到限制。

- Multi-tenancy on a massive scale is required for some of the use cases.

  在一些场景中，需要具备实现大规模多租户的功能。

- Isolation of edge computing from data center clouds may be required to ensure that compromises in the “external cloud” domain cannot impact services.

  边缘计算可能需要与云数据中心隔离，以确保“外部云”域不会影响正常服务。

The concept of edge computing must cover both the edge site (e.g. the compute, network and storage infrastructure), but also the applications (workloads) that run on it.

边缘计算的概念必须涵盖边缘站点（例如计算、网络和存储资源），还要包括运行在其上的应用程序（工作负载）。

It is worth noting that any applications in an edge computing environment could potentially leverage any or all of the capabilities provided by a cloud— compute, block storage, object storage, virtual networking, bare metal, or containers.

值得注意的是，边缘计算环境中的任何应用程序都可能利用云计算系统提供的任意或全部功能，比如，块存储、对象存储、虚拟网络、裸机或容器。

The essential features that define and separate edge computing from cloud computing are:

边缘计算区别于云计算的基本特征是：

- The ability to support a dynamic pool of multiple potentially widely distributed sites

  能够支持多个站点分布广泛的动态资源池

- Potentially unreliable network connections, and

  可能存在不可靠的网络连接

- The likelihood of difficult-to-resolve resource constraints at sites across the network.

  可能在通过多网络连接站点时，遇到难以解决的资源限制。

Exploring Characteristics

特征

So what do we know so far about edge computing characteristics, use cases, and scenarios?

到目前为止，我们对于边缘计算的特点、用例和场景知道些什么呢？

The defining need that drives cloud edge computing is the need for service delivery to be closer to users or end-point data sources. Edge computing environments will work in conjunction with core capacity, but aim to deliver an improved end user experience without putting unreasonable demands on connectivity to the core. Improvements result from:

驱动云边缘计算的决定性需求是：服务交付需要更接近用户或终端数据源。边缘计算环境将与核心计算能力结合使用，在不向核心计算能力发出不合理连接请求的前提下，改善终端用户体验，改进有以下几点：

1. Reducing latency: The latency to the end user could be lower than it would be if the compute was farther away—making, for instance, responsive remote desktops possible, or successful AR, or better gaming.

1. 降低延迟：最终用户感受到的延迟比依赖中心计算更低——例如，有效实现响应式远程桌面，AR 或实现更好的游戏体验。

2. Mitigating bandwidth limits: The ability to move workloads closer to the end users or data collection points reduces the effect of limited bandwidth at a site. This is especially useful if the service on the edge node reduces the need to transmit large amounts of data to the core for processing, as is often the case with IoT and NFV workloads. Data reduction and local processing can be translated into both more responsive applications and reduces the cost of transporting terabytes of data over long distances.

2. **打破带宽限制**：将工作负载转移到靠近最终用户或数据收集点，可以降低带宽不足带来的负面影响。特别重要的是有了边缘站点服务，传输到远端计算中心去处理的数据将大量减少。IoT 和 NFV 场景通常就是这种情况。数据缩减和本地处理可以推动交互性更高的应用发展，同时降低远距离传输 TB 级数据的成本。

But there are tradeoffs. To deliver edge computing, it is necessary to vastly increase the number of deployments. This institutes a significant challenge to widespread edge deployments. If managing a single cloud takes a team of ten, how can an organization cope with hundreds or even thousands of small clouds? Some requirements include：

但这仍需我们去权衡。应用边缘计算，就必须大量增加部署数量。广泛的边缘部署是我们面临的重大挑战。如果管理一个单独的云需要十人团队，那么应对数百甚至数千个小型云呢？需要的部分要求包括：

1. Standardization and infrastructure consistency are needed. Each location has to be similar; a known quantity.

1. 需要标准化和统一的环境。每个位置必须相近，已知数量。

2. Manageability needs to be automated; deployment, replacement and any recoverable failures should be simple and straightforward.

2. 自动化管理；部署、替换和任何恢复故障的方法都应该简洁明了。

3. Simple, cost-effective plans need to be laid for when hardware fails.

3.当硬件出现故障时，需要制定简单而经济高效的应对计划。

4. Locally fault-tolerant designs might be important, particularly in environments that are remote or unreachable—zero touch infrastructure is desirable. This is a question that balances the cost of buying and running redundant hardware against the cost of outages and emergency repairs. Considerations include:

4.本地的容错设计尤为重要，特别是在远程或无法访问的环境中——如零接触基础环境。这也是一个需要权衡购买、运行大量硬件的成本和中断、紧急修复成本的问题。需要考虑的因素包括：

a. Do these locations need to be self-sufficient?

a. 这些地点是否需要自给自足？

b. If a location has a failure, no one is going to be onsite to fix it, and local spares are unlikely.

b.如果一个地点出现故障，没有人在现场解决问题，并且不太可能具备本地

备件。

c. Does it need to tolerate failures? And if it does, how long is it going to be before someone will be available to repair it—two hours, a week, a month?

c. 是否需要容错机制？如果确实如此，在有人可以修复它之前需要多长时间——两个小时，一周，一个月？

5. Maintainability needs to be straightforward—untrained technicians perform manual repairs and replacements, while a skilled remote administrator re-installs or maintains software.

5. 维护操作需要简单明了——未经培训的技术人员可以进行修理和替换，而熟练的远程管理员可以重新安装或维护软件。

6. Physical designs may need a complete rethink. Most edge computing environments won't be ideal—limited power, dirt, humidity and vibration have to be considered.

6. 需要对物理设计进行彻底的考虑。大多数边缘计算的环境都将是不理想的——必须考虑有限的能源、灰尘、湿度和振动。

Use Cases

用例

There are probably dozens of ways to characterize use cases and this paper is too short to provide an exhaustive list. But here are some examples to help clarify thinking and highlight opportunities for collaboration.

可能有几十种方式来描述案例，本文因篇幅问题不再一一罗列。但这里有一些例子将有助于理清思路并凸显合作的机会。

Four major categories of workload requirements that benefit from a distributed architecture are analytics, compliance, security, and NFV.

受益于分布式架构的四大类工作负载分别是：分析、合规、安全和 NFV。

DATA COLLECTION AND ANALYTICS

数据收集和分析

IoT, where data is often collected from a large network of microsites, is an example of an application that benefits from the edge computing model. Sending masses of data over often limited network connections to an analytics engine located in a centralized data center is counterproductive; it may not be responsive enough, could contribute to excessive latency, and wastes precious bandwidth.

IoT 通常是从大规模的站点收集数据，是受益于边缘计算模型的典型示例。一般情况下，将大量数据通过有限的网络连接发送到集中式数据中心的分析引擎进行分析，往往会适得其反。因为它可能没有足够的响应能力，可能会导致过度的延迟，并浪费宝贵的带宽。

Since edge devices can also produce terabytes of data, taking the analytics closer to the source of the data on the edge can be more cost-effective by analyzing data near the source and only sending small batches of condensed information back to the centralized systems. There is a tradeoff here—balancing the cost of transporting data to the core against losing some information.

由于边缘设备也可能产生 TB 级的数据，因此将分析引擎靠近数据源并仅将少量的精简信息发送回中枢系统，使数据分析更接近边缘数据源，从而更具成本效益。这里也需要在数据传输到中枢系统的成本和信息缺失程度成本之间进行权衡。

SECURITY

安全

Unfortunately, as edge devices proliferate－－including mobile handsets and IoT sensors－－new attack vectors are emerging that take advantage of the proliferation of endpoints. Edge computing offers the ability to move security elements closer to the originating source of attack, enables higher performance security applications, and increases the number of layers that help defend the core against breaches and risk.

不幸的是，随着边缘设备的激增—包括移动手持设备和物联网传感器—正在涌现利用端点扩散的新攻击媒介。边缘计算需要具有安全组件移近原始攻击源的能力，启动更高性能的安全应用，并增加分层数量来防御核心漏洞和风险。

COMPLIANCE REQUIREMENTS

合规要求

Compliance covers a broad range of requirements, ranging from geofencing, data sovereignty, and copyright enforcement. Restricting access to data based on geography and political boundaries, limiting data streams depending on copyright limitations, and storing data in places with specific regulations are all achievable and enforceable with edge computing infrastructure.

合规性涵盖广泛的要求，包括地理防护、数据主权和版权强制。基于地理和政治边界对数据进行访问控制、基于版权要求限制数据流、以及将数据存储在具有特殊规定的地方。这些都是对边缘计算基础设施的要求。

NETWORK FUNCTION VIRTUALIZATION (NFV)

网络功能虚拟化（NFV）

Network Function Virtualization (NFV) is at its heart the quintessential edge computing application because it provides infrastructure functionality. Telecom operators are looking to transform their service delivery models by running virtual network functions as part of, or layered on top of, an edge computing infrastructure. To maximize efficiency and minimize cost/complexity, running NFV on edge computing infrastructure makes sense.

网络功能虚拟化（NFV）是典型的边缘计算应用，因为它提供了基础设施功能。电信运营商正在通过将虚拟网络功能作为边缘计算基础设施的一部分或分层来运行，从而改变他们的服务交付模式。在边缘计算基础设施上运行 NFV 是有意义的，可以最大限度地提高效率并降低成本/复杂性，。

REAL-TIME

即时性

Real-time applications, such as AR/VR, connected cars, telemedicine, tactile internet Industry 4.0 and smart cities, are unable to tolerate more than a few milliseconds of latency and can be extremely sensitive to jitter, or latency variation. As an example, connected cars will require low latency and high bandwidth, and depend on computation and content caching near the user, making edge capacity a necessity. In many scenarios, particularly where closed-loop automation is used to maintain high availability, response times in tens of milliseconds are needed, and cannot be met without edge computing infrastructure.

诸如 AR/VR、车联网、远程医疗、触觉互联网、工业 4.0 和智慧城市等实时应用程序无法承受超过几毫秒的延迟，并且可能对抖动或延迟变化非常敏感。例如，车联网可能需要低延迟和高带宽，并依赖靠近用户的计算和缓存，这些要求使得使边缘能力成为了必需品。在许多情况下，特别是在使用闭环自动化来维持高可

用性的情况下，需要几十毫秒的响应时间，这在没有边缘计算架构的情况下是无法满足的。

IMMERSIVE

拟真

Edge computing expands bandwidth capabilities, unlocking the potential of new immersive applications. Some of these include AR/VR, 4K video, and 360° imaging for verticals like healthcare. Caching and optimizing content at the edge is already becoming a necessity since protocols like TCP don't respond well to sudden changes in radio network traffic. Edge computing infrastructure, tied into real-time access to radio/network information can reduce stalls and delays in video by up to 20% during peak viewing hours, and can also vary the video feed bitrate based on radio conditions.

边缘计算扩展了带宽能力，释放了新的拟真式应用的潜力。其中一些包括 AR/VR、4K 视频、以及像医疗保健这种垂直领域的 360° 成像。由于像 TCP 这样的协议无法对无线网络流量的突变进行及时响应，因此边缘缓存和优化数据已经成为一种必然手段。将边缘计算基础设施应用到实时访问无线电/网络信息交互时，可以在高峰观看时间内减少高达 20%的视频停顿和延迟，并且还可以根据无线电条件实时改变视频的输入比特率。

NETWORK EFFICIENCY

网络效率

Many applications are not sensitive to latency and do not require large amounts of nearby compute or storage capacity, so they could theoretically run in a centralized cloud, but the bandwidth requirements and/or compute requirements may still make edge computing a more efficient approach. Some of these workloads are common today, including video surveillance and IoT gateways, while others, including facial recognition and vehicle number plate recognition, are emerging capabilities.

许多应用程序对延迟不敏感，不需要大量的就近计算或存储资源，因此理论上它们可以在集中式的云中运行，但对带宽和计算资源有要求的应用，使用边缘计算将成为更有效的方法。部分应用在今天很常见，包括视频监控和物联网网关，而其他一些应用（包括面部识别和车牌识别）则是新兴功能。

With many of these, the edge computing infrastructure not only reduces bandwidth requirements, but can also provide a platform for functions that enable the value of the application—for example, video surveillance motion detection and threat recognition. In many of these applications, 90% of the data is routine and irrelevant, so sending it to a centralized cloud is prohibitively expensive and wasteful of often scarce network bandwidth. It makes more sense to sort the data at the edge for anomalies and changes, and only report on the actionable data.

有了这些，边缘计算基础设施不仅可以降低带宽需求，还可以为应用提供有价值的功能平台——例如视频监控运动检测和威胁识别。在这些应用中，90%的数据是常规且不相关的，因此将其发送到集中式的云计算中心成本较高，同时还浪费了稀缺的带宽资源。因此，在边缘层对一些数据进行处理，仅向云中心报告一些必要的数据，显得更具意义。

SELF-CONTAINED AND AUTONOMOUS SITE OPERATIONS

自主体系和独立运营

Many environments, even today, have limited, unreliable or unpredictable connectivity. These could include transportation (planes, buses, ships), mining operations (oil rigs, pipelines, mines), power infrastructure (wind farms, solar power plants), and even environments that should typically have good connectivity, like stores.

即使在今天，仍有许多场所网络连接受限、不可靠或不可预知。这些场所包括交通运输工具（飞机、巴士、船舶），采矿作业区（石油钻井平台、管道、矿山），电力基础设施（风力发电场、太阳能发电厂），甚至包括通常情况下具有良好网络连接性的仓库。

Edge computing neatly supports such environments by allowing sites to remain semi-autonomous and functional when needed or when the network connectivity is not available. The best example of this approach is the need for retail locations to maintain their point of sales (POS) systems, even when there is temporarily no network connectivity.

边缘计算能够巧妙地支撑这样的环境，允许站点在需要时或网络连接不可用时保持半自治和功能可用。这种方法的最好例子就是零售店，即使暂时没有网络连接时，也仍维护销售点 POS 系统正常运行。

PRIVACY

保密性

Enterprises may have needs for edge computing capacity depending on workloads, connectivity limits and privacy. For example, medical applications that need to anonymize personal health information (PHI) before sending it to the cloud could do this utilizing edge computing infrastructure.

企业可能会因为考虑到工作负载、网络连接限制和隐私等因素，而选择边缘计算。例如，医疗应用程序在将个人健康信息（PHI）发送到云端之前，需要进行匿名化处理，边缘计算可以用以实现这一目的。

Another way to look at requirements that would benefit from cloud edge computing is by the type of company that would deploy them. Operator applications are workloads put on edge computing infrastructure that is built and managed by operators—telecommunications companies, for example. Third-party applications are built by organizations to run on existing edge infrastructure, in order to leverage others' edge computing infrastructure. It is worth noting that any applications could leverage any or all of the capabilities provided by a cloud— compute, block storage, object storage, virtual networking, bare metal, or containers.

此外，可以按照不同公司部署类型来判断从云边缘计算获益的方式。例如，操作员应用程序是将工作负载放在由运营商（电信公司）构建和管理的边缘计算基础架构上的。第三方应用程序是由其它组织构建的，在现有的边缘计算基础设施上运行，以便利用他人的边缘计算基础资源。值得注意的是，任何应用程序都可以影响云中提供的计算、块存储、对象存储、虚拟网络、裸机或容器等任意或全部功能。

Scenarios

场景

The basic characteristic of the edge computing paradigm is that the infrastructure is located closer to the end user, that the scale of site distribution is high and that the edge nodes are connected by WAN network connections. Examining a few scenarios in additional depth helps us evaluate current capabilities that map to the use case, as well as highlighting weaknesses and opportunities for improvement.

边缘计算的基本特征是基础设施位置更接近最终用户，站点分布规模较高，边缘站点通过 WAN 网络进行连接。对场景的进一步验证有助于我们评估映射到用例的当前功能，找出不足并加以改进。

1. Retail/finance/remote location "cloud in a box": Edge computing infrastructure that supports a suite of applications customized to the specific company or industry vertical.

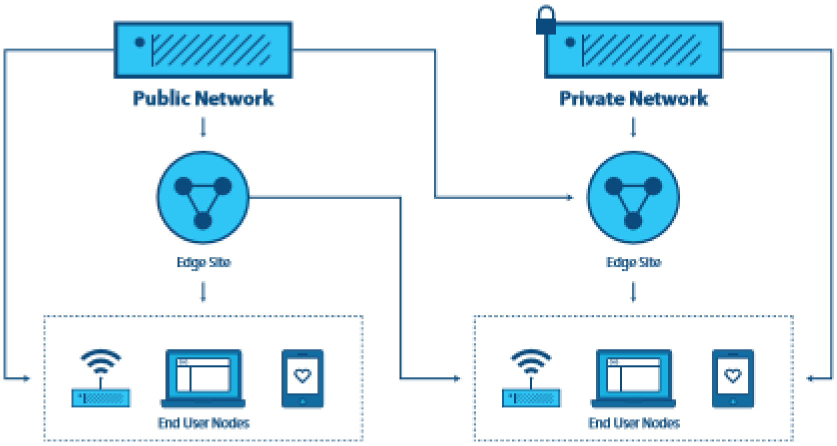
1. 零售/金融/远程位置“盒中云”：是支持针对特定公司或行业，垂直定制的一套应用程序的边缘计算框架。

Often used by the enterprise, edge computing infrastructure, ultimately coupled together into distributed infrastructure, to reduce the hardware footprint, standardize deployments at many sites, deliver greater flexibility to replace applications located at the edge (and to have the same application running uniformly in all nodes irrespective of HW), boost resiliency, and address concerns about intermittent WAN connections. Caching content or providing compute, storage, and networking for self-contained applications are obvious uses for edge computing in settings with limited connectivity.

边缘计算基础架构往往由企业使用，它耦合到分布式基础架构中，以减少硬件占用空间，在众多站点实现标准化部署，为位于边缘的应用程序提供更大的灵活性（无论硬件如何，都可以在所有站点中运行相同的应用程序），提高弹性，解决间歇性 WAN 连接问题。在连接受限的情况下，边缘计算经常使用内容缓存或为自主体系的应用程序提供计算、存储和网络服务等方法进行响应。

2. Mobile connectivity: Mobile/wireless networks are likely to be a common environmental element for cloud edge computing, as mobile networks will remain characterized by limited and unpredictable bandwidth, at least until 5G becomes widely available. Applications such as augmented reality for remote repair and telemedicine, IoT devices for capturing utility (water, gas, electric, facilities management) data, inventory, supply chain and transportation solutions, smart cities, smart roads and remote security applications will all rely on the mobile network to greater or lesser degrees. They will all benefit from edge computing’s ability to move workloads closer to the end user.

2. 移动连接：在 5G 变得广泛应用前，移动网络依然存在着带宽有限和不可预知的特征，移动/无线网络仍是云边缘计算的一个常见环境因素。例如用于远程维修的远程医疗、增强现实、用于捕捉公用事业（水，气，电，设施管理）数据物联网设备、库存、供应链和运输解决方案、智能城市、智能道路和远程安全应用等应用都或多或少的依赖于移动网络。它们都将受益于边缘计算这种将工作负载转移到靠近最终用户的能力。



3. Network-as-a-Service (NaaS): Coming from the need to deliver an identical network service application experience in radically different environments, the NaaS use case requires both a small footprint of its distributed platform at the edges, and strong centralized management tools that cross over unreliable or limited WAN network connections in support of the services out on the edge.

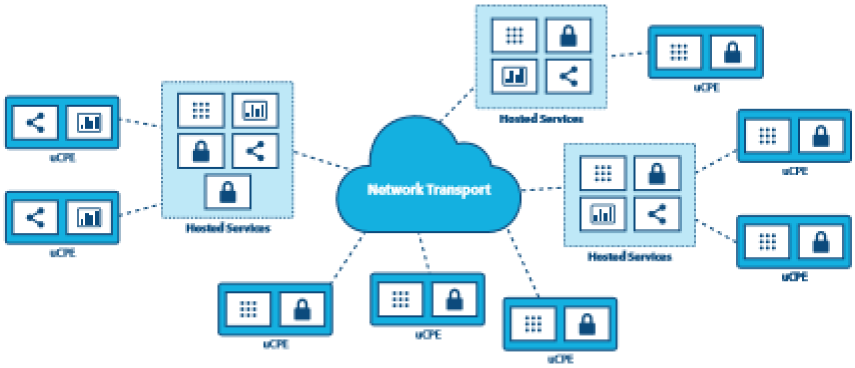
3. 网络即服务（NaaS）：由于需要在完全不同的环境中提供相同的网络服务，因此 NaaS 用例只能占据位于边缘处小部分的分布式平台，而且需要能够通过不可靠或有限 WAN 网络实现集中管理工具的连接，来支持边缘服务。

The main characteristics of this scenario are: small hardware footprint, moving (changing network connections) and constantly changing workloads, hybrid locations of data and applications. This is one of the cases that needs infrastructure to support micro nodes—small doses of compute in non-traditional packages (not all 19in rack in a cooled data center).

这种场景的主要特点是：硬件资源少、网络连接和工作负载不断变化、数据和应用程序的混合位置。这是基础设施需要支持微计算站点的情况之一——并非在传统数据中心的机架上。

NaaS will require support for thousands or tens of thousands of nodes at the edge and must support mesh and/or hierarchical architectures as well as on demand sites that might spin up as they are needed and shutdown when they are done. APIs and GUIs will have to change to reflect that large numbers of compute nodes will have different locations instead of being present in the same data center.

NaaS 需要在边缘有数千或数万个站点支持，并且这些站点必须支持网格、分层体系结构以及按需分配，这些站点可能会在需要时启动、在完成后关闭。API 和 GUI 将不得不改变，反映出大量计算站点处于不同位置，而不是存在于一个数据中心中。



4. Universal Customer Premises Equipment (uCPE): This scenario, already being deployed today, demands support for appliance-sized hardware footprints and is characterized by limited network connections with generally stable workloads requiring high availability. It also requires a method of supporting hybrid locations of data and applications across hundreds or thousands of nodes and scaling existing uCPE deployments will be an emerging requirement.

4. 通用客户端设备（uCPE）：这种场景现在已经有部署，支持家电大小的设备，其特点是网络连接资源有限、有高可用需求的稳定运行工作负载。它还需要一种支持数百或数千个站点的数据和应用程序的混合部署的方法，并且扩展现有的 uCPE 部署将成为一个新兴需求。

This is particularly applicable to NFV applications where different sites might need a different set of service chained applications, or sites with a different set of required applications that still need to work in concert. Mesh or hierarchical architectures would need to be supported with localized capacity and the need to store and forward data processing due to intermittent network connections. Self-healing and self-administration combined with the ability to remotely administer the node are musts.

这尤其适用于 NFV 应用程序，在这些应用程序中，不同的站点可能需要不同的服务链应用集合，或者具有不同的应用程序集合的站点仍然可以协同工作。网格或分层体系结构需要支持本地化能力，并且能够在间歇性的网络连接下实现数据的存储和转发处理。自我修复、自我管理、远程管理站点的能力是必须的。

5. Satellite enabled communication (SATCOM): This scenario is characterized by numerous capable terminal devices, often distributed to the most remote and harsh conditions. At the same time, it makes sense to utilize these distributed platforms for hosting services, especially considering the extremely high latency, limited

bandwidth and the cost of over-the-satellite communications. Specific examples of such use cases might include vessels (from fishing boats to tanker ships), aircrafts, oil rigs, mining operations or military grade infrastructure.

5. 卫星通信（SATCOM）：该场景的特点是有着多种功能强大的终端设备，而且这些设备通常分布在最偏远和恶劣的环境中。与此同时，将这些分布式平台用于托管服务是有意义的，特别是考虑到极高的延迟，有限的带宽和卫星通信的成本。这种用例的具体例子可能包括船只（从渔船到油船），飞机，石油钻井平台，采矿作业或军事级别的基础设施。

Challenges

挑战

Though there are plenty of examples of edge deployments already in progress around the world, widespread adoption will require new ways of thinking to solve emerging and already existing challenges and limitations.

尽管世界各地已经有很多边缘计算部署的例子，但广泛的部署和应用，还需要采用新的思维方式来解决新兴和已有的挑战和限制。

We have established that the edge computing platform has to be, by design, much more fault tolerant and robust than a traditional data center centric cloud, both in terms of the hardware as well as the platform services that support the application lifecycle.

我们已经搭建的边缘计算平台，不论在硬件还是支持应用程序生命周期平台服务方面，都比以数据中心为中心的云更具容错性和可靠性。

We cannot assume that such edge use cases will have the maintenance and support facilities that standard data center infrastructure does. Zero touch provisioning, automation, and autonomous orchestration in all infrastructure and platform stacks are crucial requirements in these scenarios.

我们不能假定这种边缘用例也会具有标准数据中心基础架构所具有的运维和支持能力。这些场景中所有基础设施和平台堆栈都要满足零接触配置、自动化、自动编排的需求。

But there are other challenges that need to be taken under consideration.

但是还需要考虑其他的挑战。

For one, edge resource management systems should deliver a set of high-level mechanisms whose assembly results in a system capable of operating and using a geo-distributed IaaS infrastructure relying on WAN interconnects.

例如，边缘计算的资源管理系统应该依靠 WAN 网络的互连操作，使用地理上分布式的 IaaS 基础设施。

In other words, the challenge is to revise (and extend when needed) IaaS core services in order to deal with aforementioned edge specifics — network disconnections/bandwidth, limited capacities in terms of compute and storage, unmanned deployments, and so forth.

换句话说，挑战在于修改（并在需要时扩展）IaaS 核心服务，以便处理上述边缘计算的一些细节——网络中断/带宽、计算和存储资源限制、无法手动部署等等。

Some foreseeable needs include:

一些可预见的需求包括：

- A virtual-machine/container/bare-metal manager in charge of managing machine/container lifecycle (configuration, scheduling, deployment, suspend/ resume, and shutdown).

  生命周期管理。负责管理虚拟机、容器、裸机的生命周期（配置，调度，部署，挂起/恢复和关闭）。

- An image manager in charge of template files (a.k.a. virtual-machine/ container images).

镜像管理。负责模板文件（也称为虚拟机/容器镜像）的管理。

- A network manager in charge of providing connectivity to the infrastructure: virtual networks and external access for users.

  网络管理。负责连接基础设施，包括用户的虚拟网络和外部访问。

- A storage manager, providing storage services to edge applications.

  存储管理。为边缘应用程序提供存储服务。

- Administrative tools, providing user interfaces to operate and use the dispersed infrastructure.

  使用管理。提供操作和使用分布式基础设施的用户界面。

These needs are relatively obvious and could likely be met by leveraging and adapting existing projects. But other needs for edge computing are more challenging. These include, but are not limited to:

以上这些需求相对较为明显，可以通过利用和调整现有项目来满足。但其他边缘计算的需求更具挑战性。这些包括但不限于：

- Addressing storage latency over WAN connections.

  应对通过 WAN 网络连接导致的存储延迟。

- Reinforced security at the edge—monitoring the physical and application integrity of each site, with the ability to autonomously enable corrective actions when necessary.

  增强边缘站点安全性——监控每个站点的物理和应用完整性，并在必要时自动启用修正措施。

- Monitoring resource utilization across all nodes simultaneously.

  同时监控所有站点上的资源利用率。

- Orchestration tools that manage and coordinate many edge sites and workloads, potentially leading toward a peering control plane or “self-organizing edge”.

管理和协调众多边缘站点和工作负载的编排工具，可能导致对等控制平面或“自边缘组织”。

- Orchestration of a federation of edge platforms (or cloud-of-clouds) has to be explored and introduced to the IaaS core services.

  探索多边缘计算平台（或多云之云）的编排，并将其引入到 IaaS 核心服务中。

  - Automated edge commission/decommission operations, including initial software deployment and upgrades of the resource management system’s components.

    边缘站点使用/释放操作的自动化，包括软件初始化部署和资源管理系统组件的升级。

  - Automated data and workload relocations—load balancing across geographyically distributed hardware.

    数据和工作负载更换位置的自动化——跨地域分布的硬件之间负载均衡。

- Some form of synchronization of abstract state propagation should be needed at the “core” of the infrastructure to cope with discontinuous network links.

  在基础设施的“核心”需要一种同步机制，应对不连续的网络链路。

- New ways to deal with network partitioning issues due to limited connectivity—coping with short disconnections and long disconnections alike.

  需要新的方法来处理，由于连接受限导致的网络分区问题——应对短暂断开连接和长时间断开连接。

- Tools to manage edge application life cycles, including：

  管理边缘应用程序生命周期的工具，包括：

  - The definition of advanced placement constraints in order to cope with latency requirements of application components.

    定义高级部署位置的约定，以便应对应用程序组件的延迟需求。

  - The provisioning/scheduling of applications in order to satisfy placement requirements (initial placement).

提供/安排应用程序需满足部署位置要求（初始放置）。

- Data and workload relocations according to internal/external events (mobility use-cases, failures, performance considerations, and so forth).

  根据内部/外部事件（移动用例，故障，性能考虑等）重新分配数据和工作负载。

- Integration location awareness: Not all edge deployments will require the same application at the same moment. Location and demand awareness are a likely need.

  集成定位感知：并非所有边缘部署都需要在同一时间使用同一个应用程序。定位和需求感知可能是需要的。

- Discrete hardware with limited resources and limited ability to expand at the remote site needs to be taken into consideration when designing both the overall architecture at the macro level and the administrative tools. The concept of being able to grab remote resources on demand from other sites, either neighbors over a mesh network or from core elements in a hierarchical network, means that fluctuations in local demand can be met without inefficiency in hardware deployments.

  在设计宏观层面的整体架构和管理工具时，需要考虑资源的有限性和远程站点扩展能力的局限性。能够从其它站点按需获取远程的资源——无论是相邻的站点还是分层网络中的核心站点——这个意味着可以在不降低硬件部署效率的情况下，满足本地需求的波动。

Conclusion and Call to Action

结论和呼吁

Edge computing is not and should not be limited to just the components and architectures of OpenStack, but there are some reasons that OpenStack is particularly attractive as a platform for cloud edge computing.

边缘计算不是也不应该仅限于 OpenStack 的组件和架构，但 OpenStack 作为云边缘计算平台具有其独特的魅力。

The OSF Edge Computing Group is asking the open source community to begin exploring these challenges and possibilities. We recognize that there is work to be done to achieve our goals of creating the tools to meet these new requirements. We welcome and encourage the entire open source community to join in the opportunity to define and develop cloud edge computing. You can find more information about the group activities on the OpenStack Edge Computing web page.

OSF 边缘计算小组正在呼吁开源社区探寻这些挑战和更多的可能性。我们意识到要创造出满新需求的工具，我们还有很多工作要做。我们欢迎并鼓励整个开源社区加入定义和开发云边缘计算的中来。您可以在 OpenStack 边缘计算网页上找到有关活动的更多信息。