迈向智能世界白皮书2023

数据通信

网络加速AI，AI改变网络

构建万物互联的智能世界

序言

人工智能正在重塑整个人类社会。我们预计人工智能的行业渗透率将从2021年的7%增长到26年的30%。随着大模型加速行业智能化的转型，到2030年渗透率将超过50%，提升10倍。人工智能的快速发展，将进一步推动行业数字化转型，为网络创新提供了新机遇。

企业加速上云，广域网络敏捷性和安全性亟需提升：

一方面，全球企业上云比例已高达70%，混合分布式多云成为主流，上云和云间流量快速增长，企业希望通过一张弹性、敏捷的网络灵活、按需的连接多云以释放云端效率。另一方面，能源、交通、金融等传统行业进入快速云化转型期，对网络提出差异化的承载诉求，网络需要提供基于业务诉求的定制化质量保障能力。为做好企业业务云化的有效支撑，网络需要向弹性敏捷，安全可靠演进。

AI算力激增，带来数据中心网络变革：

ChatGPT等激发了AI的快速普及和增长，到2026年，AI行业渗透率将达到30%。从2023到2030年，AI算力将增长500倍。AI训练所用的计算量呈指数增长，带宽需求平均每3.5个月便会翻倍，远超摩尔定律定义的18个月。AI算力激增将带来全球数据中心网络建设需求增长以及网络技术变革。0.1%的网络丢包会带来50%的计算性能下降，如果要100%释放算力，需要构建高吞吐、零阻塞的数据中心网络。

园区网络代际升级，进入体验为王的时代：

行业数字化转型需要建立高速、稳定的园区网络环境。园区网络连接范围正在快速扩大，从办公到生产，从联接人到联接物，未来5年，园区接入终端数将增加3倍，园区网络需要提供泛在的网络连接，并基于办公业务和生产业务提供隔离能力；园区业务正在快速变革，移动办公和视频会议成为园区的两大主流发展趋势，园区内80%的流量将是音视频流量，园区网络进入体验为王的时代。这些都需要对原有网络进行升级换代，例如，从Wi-Fi 4/5到Wi-Fi 6/7，从千兆到万兆接入。

网络复杂性急剧提升，智能化加速网络自治：

随着云计算和物联网等技术的不断发展应用，万物互联、万物感知、万物智能的智能社会逐步推进。企业网络从辅助办公到支撑生产，从静态配置到按需调整，从单域管理到全网协同，网络的边界不断拓宽，网络的质量属性日益增强，网络的运维模式正在发生质的改变。AI应用于网络可以帮助网络突破人工运维的效率瓶颈，让网络具备高度的自动化和智能化能力，即实现网络的自动驾驶，为企业数字化业务创新和敏捷运营铺就基石。

网络攻击无处不在，需构筑一体化安全防御体系：

随着业务上云，传统网络边界被打破，给网络安全带来了更大的挑战和更多的不确定性。2022年，85%的企业都经历过网络攻击，全球网络攻击数增长42%，平均每11s就会发生一次勒索攻击。网络攻击可直接引起企业业务中断，敏感数据泄露，甚至巨额的经济损失。为有效防范网络攻击，建立“云网边端”一体化的安全防御体系是关键。

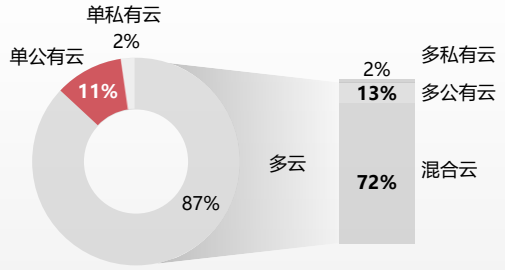
数据通信产业正在变革，网络提升AI训练效率，让算力无所不及，AI也将彻底改变网络，让智能无处不在。

目录

多云成为企业数字化新常态

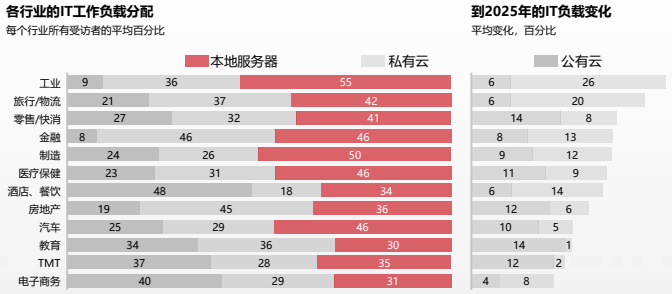
- **企业多云新常态：**企业出于成本节约、数据安全、不同云间的技术整合等多方面考虑，开始逐渐从过去的上IaaS、PaaS和SaaS，发展到私有云、行业云、边缘云、分布式云等各种形态，企业基于自身业务诉求，灵活利用各种不同形态的云服务，将正确的工作匹配到正确的云服务，将多项工作负载分散在不同云平台上运行，就是企业上多云。Flexera 2023 云业务报告显示，在受访企业中，已有87%的企业正在使用多云服务，多云成为企业数字化转型新常态；
- **传统行业上云加速：**随着云计算、大数据和人工智能的发展，云服务作为助推业务创新、企业升级的动力覆盖到了越来越多的领域。企业为了紧追技术变革，寻求新的发展机遇，纷纷向“云”敞开了怀抱，云服务的实践者从互联网行业开始逐步向工业、教育、医疗、政府、能源、金融等传统非数字原生行业迈进，数字化转型的带动效应开始显现。以工业领域为例，麦肯锡预计，到2025年，传统行业的业务上云比例将大幅提升；
- **行业上云以分布式混合多云为主：**行业上云发生在企业不断追求更高的效率、性价比和业务增长的背景下，而混合多云结合了公有云和私有云的优点，既保障了企业数据的安全性，又提供灵活的云架构，受到企业的青睐。同时，大量新兴业务应用需要海量数据分析和计算能力，多层级、分布式的云计算建设模式成为行业的主流方向，如金融行业“两地三中心”，能源行业的分布式数据中心叠加公有云服务的架构等；

2023年，87%受访企业已采用多云



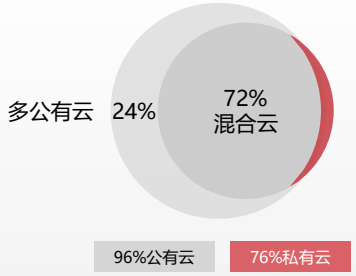
来源：Flexera 2023 State of the Cloud Report

到2025年，行业业务上云比例将大幅提升



来源：麦肯锡2021年中国云计算调研

72%的受访客户使用混合云



来源：Flexera 2023 State of the Cloud Report

匹配多云战略，企业多云网络建设进入高峰期

网络成为多云战略核心需求：

云业务的开展需要强大的网络能力支撑，网络资源的优化同样要借鉴云计算的理念。随着多云战略的快速落地，异构联接、复杂的网络管理、E2E业务体验保障以及安全防护，都要求网络基础设施要更好的适应云计算应用的需求，并能更好的优化网络结构，以确保网络满足行业云业务要求。

行业数字化专网建设进入高峰期：

- **金融：**金融业务规模快速增长和分布式架构转型，对承载金融业务的广域网络提出了新需求。金融广域网作为金融云和金融网点间的连接通道，是金融业务高效、稳定运行的基石。金融机构以满足多地多中心及分支机构的互联互通需求为基础，正在加速构建高速、智能、弹性的广域网络。在中国，已经有包括中国建设银行、交通银行等20多家金融机构进行多云网络重构；
- **能源：**多地多中心的分布式多云架构成为能源行业云化战略，生产、管理、经营数据全面上云，要求网络、算力以及数据的高效灵活调度。数字化生产，前端数据采集和后端实时智能化分析，前后端联动实现智能化作业，要求能源数据网提供确定性体验的网络保障能力。电力、油气等能源领域已经基于数字化发展需求，开始进行多云网络的建设；
- **政府：**政务云是一朵物理分散、逻辑集中的云，面向全国各地各部门提供统一的云服务。政府多云网络打通各政府部门的壁垒，实现资源融通，满足不同部门不同业务的差异化承载诉求，实现多级联动、服务智能、集约建设、全面覆盖；

2019年到2023年，中国已有300+行业客户进行多云网络建设

招商银行 CHINA MERCHANTS BANK　交通银行 BANK OF COMMUNICATIONS　广东政务服务网　国家管网 PipeChina

中国银行 BANK OF CHINA　中国建设银行 China Construction Bank　PICC 中国人民保险集团公司　中国农业银行 AGRICULTURAL BANK OF CHINA

江苏省大数据管理中心　中信银行 CHINA CITIC BANK　中国光大银行 Bank　中国南方电网

…

行业多云网络特征

弹性敏捷：多云按需连接，多云协同

业务隔离：数据安全，确定性体验保障

实时可视：网络、业务、体验多维可视

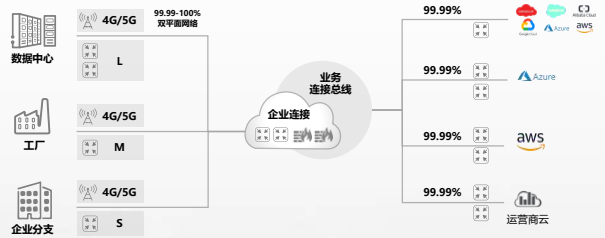
运营商抓住企业上云风口，多种模式展开多云网络创新

传统运营商网络无法满足行业多云业务诉求：随着架构在网络基础上的云计算及其应用的快速发展，云计算对于网络的要求正从简单的提供专线接入向弹性敏捷、业务隔离、体验可靠的多云网络演进。但传统运营商网络注重建设和运维，在业务网络开通速度，灵活调整，智能化等方面，运营商传统大网无法满足企业需求。

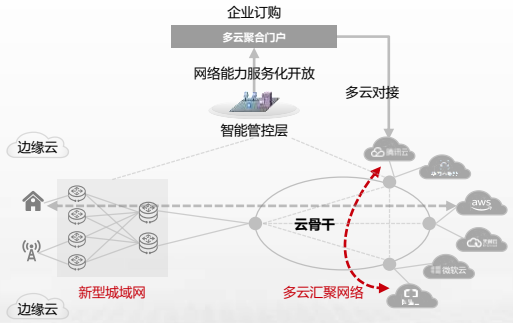
运营商开始进行多云网络创新：运营商拥有庞大的网络基础设施，这对运营商来讲既是优势，可以基于网络基础设施面向企业提供各类业务，但同时也成为一种负担，当业务出现新诉求时，运营商因为庞大的网络资产负担，网络转型需要逐步进行。

- **上云专线模式，叠加套餐增加盈利：**传统组网或互联网专线向上云专线和多云互联专线转变，部署SD-WAN等实现上云专线敏捷开通。基于POP资源池，提供任意接入方式的网关能力。POP点通过传统专线实现与云的按需连接。SD-WAN + 专线连接到多云，并不是真正的多云网络，无法提供端到端的连接、安全、可视等能力。
- **多云汇聚模式，抢占统一采购入口：**运营商新建多云汇聚骨干网络，实现本地不同云资源的预连接，并基于Overlay网络，提供一站式多云连接及增值服务。运营商提供从企业到骨干网的端到端SRv6能力，实现企业多云网络的自动化编排，以及多云协同。运营商建设多云汇聚平台，通过与第三方云平台的API接口对接，实现第三方云服务的转售，使能商业模式创新；
- **行业专网模式，价值客户体验保障：**行业办公、生产上云，金融、政府、教育等高价值行业出于安全性的考虑，业务要求和公用业务隔离，业务体验对网络质量要求高，不能因为网络中的其他突发业务导致体验受损。运营商为了满足行业客户需求，在SRv6云骨干基础上部署网络切片，或基于客户诉求建设行业物理专网。上云业务SLA通过对租户网络业务流量和性能指标的实时分析呈现，一方面让租户可实时感知自己专线的服务质量，支撑SLA商业变现。同时，可通过集中监控各租户的业务SLA，及时发现上云网络流量、性能指标劣化等事件，进行提前优化和有针对性的维护，提升云业务体验；

模式1：增加上云专线，捆绑套餐增加盈利



模式2：建设多云汇聚平台，统一多云入口



模式3：高价值客户建设专属网络，确保数据可靠

关键特征1：弹性敏捷，多云算存资源可调度

流量突发业务对传统专线的商业模式提出挑战：传统上云专线供应模式为固定时间内固定带宽，无法应对企业临时性大带宽业务：购买专线带宽不足，导致业务体验受损，或者长期维护大带宽专线，导致成本过高；临时性大带宽业务包括大带宽实时通信和周期性数据搬迁两类业务。大带宽实时通信即时性强，主要由事件造成，持续时间数小时或几天不等，无法通过随意拉长通信时间解决带宽不足问题，只能损失即时的通信体验；周期性数据搬迁业务实时性不强但总耗时有要求；

弹性上云专线，满足企业潮汐业务带宽诉求：弹性计算在云计算领域已经相当成熟，从消费者角度看弹性服务带来的是满足其任务诉求的最优成本的服务产品，真正实现PAYG，既满足任何业务场景要求又买的起、不浪费的高性价比产品，从供应商角度看，弹性服务本质是资源的高效管理，使其发挥最大效能；弹性上云专线引入云计算的弹性理念，将网络带宽资源池化，并基于对租户业务流量进行实时感知与预测，实现全网带宽资源的灵活调度，保障上云专线业务体验；

弹性专线保障用户体验，助力网络资源变现：通过弹性专线，企业在保留一个固定带宽专线的同时，基于业务情况，临时增加带宽或购买流量包，并在临时增加带宽或流量包上提供和固定带宽专线一致的质量保障，真正实现按需购买、按量付费。运营商则可以充分运用空闲带宽资源，最大化网络价值；

弹性网络架构

弹性网络服务层
弹性网络服务目录及开放API

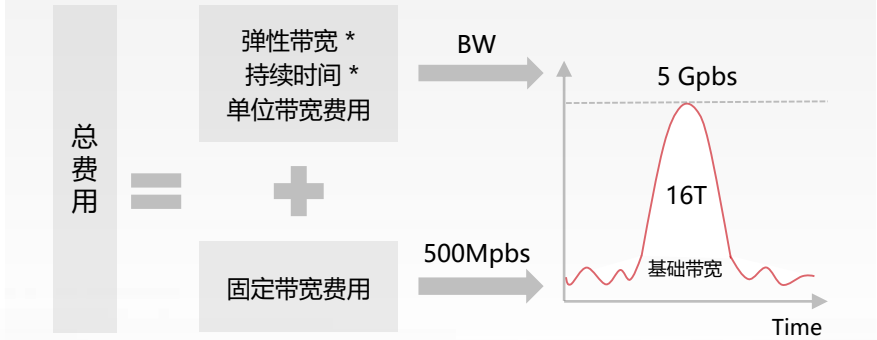
流量决策调度层
全网和节点两级流量均衡调度

弹性资源抽象层
节点级带宽颗粒抽象和管理

弹性网络4大特征

- **M级带宽颗粒池化**
- **分钟级租户流量预测**
- **秒级智能决策和调度**
- **秒级弹性带宽调整**

弹性网络使能用户基于使用量按需购买弹性流量包

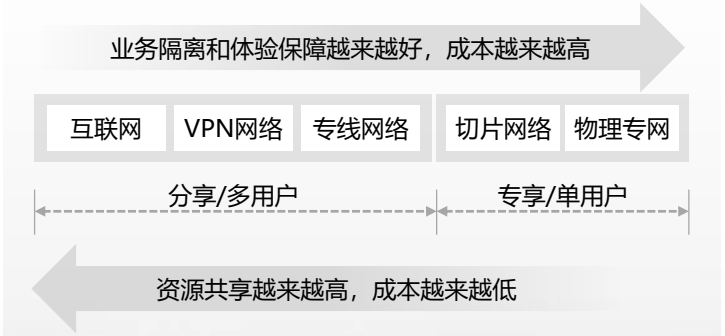


关键特征2：业务隔离，关键上云业务质量可保障

行业业务上云，对安全隔离、可靠性提出更高要求。为了在同一张网络上满足不同业务的安全隔离和差异化保障需求，业界提出网络切片的概念。

- **资源与安全隔离：**IP网络切片隔离的目的，一方面是从服务质量的角度，需要控制和避免某个切片中的业务突发或异常流量影响到同一网络中的其他切片，做到不同网络切片内的业务之间互不影响。这一点对于垂直行业尤其重要，如智能电网，这类行业对于时延、抖动等方面的要求十分严苛，无法容忍其他业务对其业务性能的影响。另一方面是从安全性角度，某个IP网络切片中的信息不希望被其他用户访问或者获取，这时需要为不同切片之间提供有效的安全隔离措施，如金融、政府等专线业务；
- **差异化SLA保障：**网络切片使运营商从单一的流量售卖服务，逐步向面对不同行业、不同业务提供差异化服务进行转变，以切片商品的方式为租户提供差异化服务。按需、定制、差异化的服务将是未来运营商提供业务的主要模式，也是运营商新的价值增长点；
- **高可靠保障：**高价值业务和uRLLC业务要求IP网络提供高可用性网络，毫秒级故障恢复已经成为IP网络的基础要求。基于SRv6的网络切片提供针对IP网络中任意故障点的本地保护技术，如TI-LFA（Topology-Independent Loop-free Alternate，与拓扑无关的无环路备份路径）、中间节点保护等，利用这些技术可以极大地提高保护成功率，增强IP网络切片的可靠性。并且，各网络切片内的链路故障倒换能够控制在切片内进行，不影响其他业务切片；

切片提供最高的资源隔离和体验保障



网络切片为行业提供专网式上云体验

灵活分片，差异化SLA

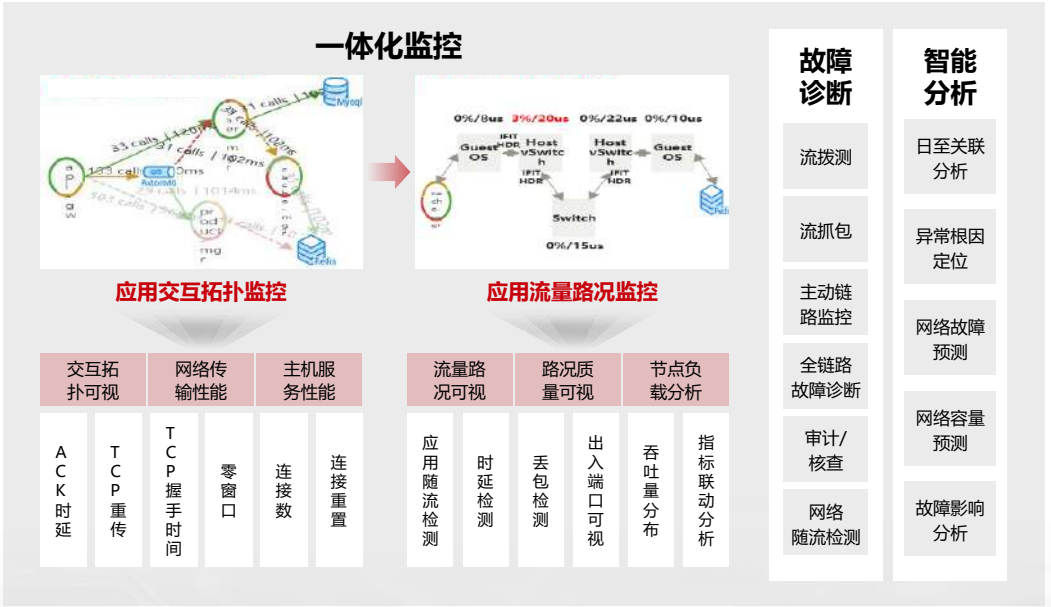
- 基于固定带宽或收敛比的端到端网络分片
- 不支持分片的局部网络可以直接透传

一网多用，体验可保障

- 分片即子网，资源独享
- 分片内SRv6灵活计算最优业务路径

关键特征3：实时可视，端到端业务质量可监控

- **网络不可视，运维效率低：**企业网络的复杂性将指呈数级增加：由于混合办公的趋势，互联分支增多，接入位置也随之增多；办公网与物联网融合，联接数量激增；云化与新应用对网络性能的要求更高、变更频繁；网络设备的种类多、厂家多，设备管理量规模化扩大；网络保障从基于联接到基于体验，要求更高。与此同时，运维保障工程师的数量却不会线性增加，甚至不增加，这就意味着要用少量的人去做更多的事情。因此，网络运维的痛点更加凸显出来，没有一张统一的视图感知企业网络的健康状态，用户网络体验差，故障投诉多，异常恢复的效率低等，远远跟不上企业数字化转型的步伐；
- **多维可视化，实时感知网络变化：**网络可视化拥有实时、动态、高清的全网资源可视能力，通过大数据计算引擎、AI、搜索算法、路由仿真和验证算法等关键技术，实现多维可视、路径导航、搜索定位、确定性应用体验保障等，提供网络质量实时可视、定界定位和自愈能力，帮助客户从传统的静态拓扑运维模式切换到动态高清的电子地图运维模式，即通过网络数字地图来看直观感知网络，大大提高网络运维效率；



行动建议：针对多云网络建设，聚焦弹性敏捷、业务隔离、实时可视

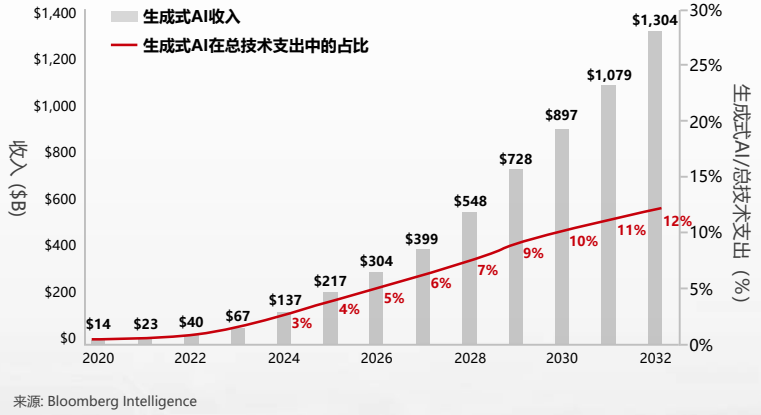
增加云网投资	顺应当前行业云网诉求高涨趋势，升级上云网络、建设多云生态等，逐步推动云、网深度协同，降低业务上云难度，提升上云业务体验；
持续探索商业模式创新	运营商通过转售第三方合作伙伴云服务、提供弹性专线能力，一方面可以更好的满足企业客户上云对网络的诉求，还可以更好的释放网络资源优势，实现营收增长；
积极引用网络新技术	SRv6、网络切片等网络技术，在简化上云网络复杂度，保障上云业务体验等方面均取得效果，企业和运营商在进行云网协同建设过程中，应该考虑引入新技术，享受技术红利；
增加网络可视化能力	数字孪生已经在行业中得到广泛应用，网络数字孪生已经成为业界共识，网络数字孪生结合人工智能、大数据等技术，实现业务体验劣化、网络故障等问题的预测，主动实施网络优化， 可以让企业和运营商更好的掌握网络、业务质量。

目录

AIGC催生万亿产业市场新价值，全球算力基础设施建设提速

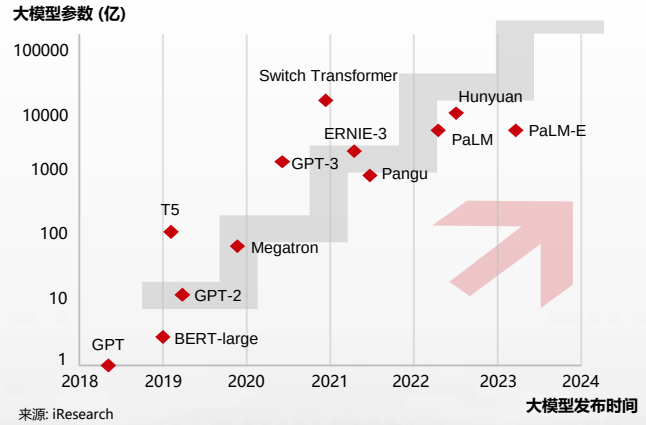
- **大模型爆发，AIGC时代到来：**2022年11月，随着Open AI发布ChatGPT，AI产业迅速进入以大模型为技术支撑的AIGC（Artificial Intelligence Generated Content，简称生成式AI）时代，开辟了人类生产交互的新纪元，也引爆了AI时代的内容生产力革命。根据Bloomberg Intelligence的最新报告显示，到2032年，生成式AI市场的营业收入从2022年的400亿美元将增长到1.3万亿美元，年均复合增长率达到42%；
- **AIGC将在全行业引发深度变革：**AIGC正在加速渗透到各行各业，但总体而言，AIGC主要影响内容创作和人机交互，行业线上化程度和内容在价值链中的占比越高，AIGC对其颠覆效应越明显。比如电商、游戏和广告行业线上化程度高，且内容质量直接决定其价值创造，因此AIGC应用在这些领域能够产生最大化的价值；
- **全球算力基础设施建设提速：**2018年6月Open AI的GPT模型参数量已经突破1.17亿，模型参数量开始亿级别的飞跃式发展，平均每3-4个月即呈现翻倍态势，由此带来的训练算力需求也“水涨船高”。计算力指数平均每提高1点，国家的数字经济和GDP将分别增长3.5‰和1.8‰，算力正成为影响国家综合实力的关键要素，算力基础设施建设成为国家数字经济高质量发展的战略举措。IDC数据显示，全球企业在AI基础设施及服务的投资，有望到2025年突破2000亿，增幅远超企业数字化转型（DX）和国内生产总值（GDP）；

生成式AI市场收入预测（2020年-2032年）



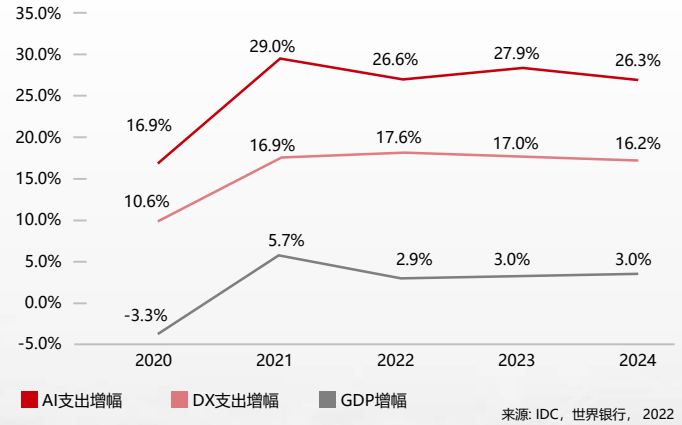
来源: Bloomberg Intelligence

全球大模型参数量变化趋势



来源: iResearch

全球AI支出增幅远超数字化转型及GDP

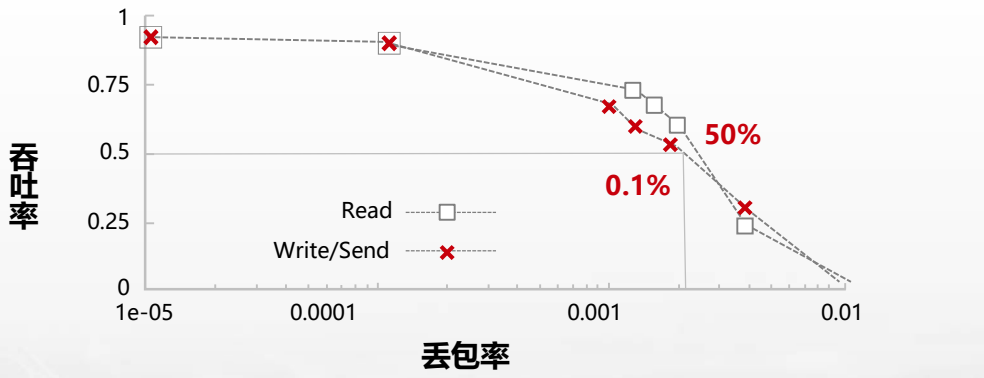


来源: IDC，世界银行，2022

网络性能决定算力效率，传统网络无法满足AI需求

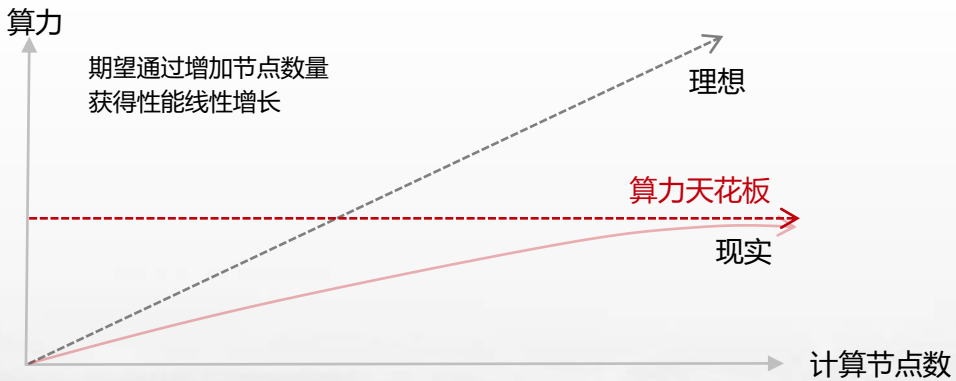
- **传统数据中心网络不能满足AI数据中心要求：** IDC报告显示，当前主流数据中心以太网占比超过95%，但传统以太网在AI训练等场景下，吞吐量、时延及避免丢包等方面的表现并不出色。众所周知，由于人工智能应用程序的通信方式会对网络造成很大负担，这给 CPU 和 GPU 服务器以及将这些系统连接到一起的现有底层网络基础设施带来了新的挑战。在如今的AI时代，AI训练过程中不能出现任何数据的丢失，而传统以太网络具有“天然丢包”的特性，已经不适应AI时代数据中心的需求；
- **万卡算力集群，需要超大规模的网络：**为了能够更快推出AI大模型，同时又满足参数和token数十倍的增长需求，GPU集群规模已经从千卡走向万卡，例如OpenAI GPT4使用上万张GPU卡训练1.8万亿参数。这就需要一张大规模的训练网络支撑如此庞大的算卡间无阻塞互联；
- **万亿参数模型，需要超高吞吐的网络：**大模型采用分布式训练方法来提高训练质量和速度，海量的参数分布于多个服务器的多个GPU之上，需要用到成千上万个GPU来训练数十TB级甚至更大的数据，大量GPU之间的通信容易出现由于网络负载分担不均而导致的网络吞吐下降，从而引发AI训练性能整体下降；
- **长稳训练，需要极致可靠的网络：**大模型训练是一个复杂的系统工程，从数据准备，模型预训练到模型训练，系统稳定运行十分重要，而网络基础设施是长稳训练的关键。某个千亿大模型总训练时长65天，期间由于故障引起的重启达到50多次，真正的训练时长只有33天，平均无故障时间（MTBF）仅为1.3天。AI大模型训练时间长，中断次数多，亟需通过提升网络健壮性，确保训练高效可靠的进行；

0.1%的丢包会造成50%的算力损失



来源: Congestion Control for Large-Scale RDMA Deployments

网络性能带来算力天花板，投资收益严重失衡



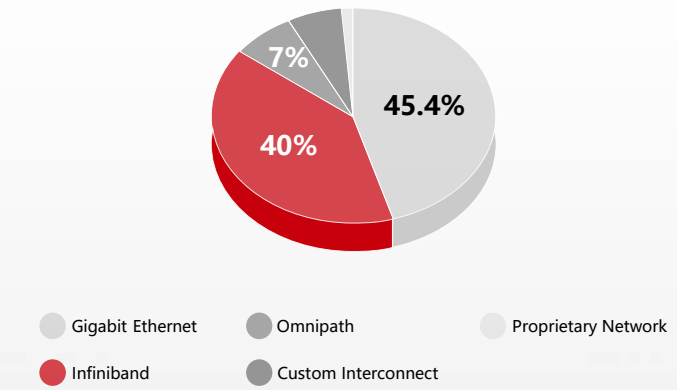
以太网技术持续创新，推动AI数据中心从封闭走向开放

- **产业积极布局：**2023年7月，Linux 基金会联合多家厂商成立超以太网联盟 (Ultra Ethernet Consortium，UEC)，旨在提高数据传输速度和网络性能，以更好地适应人工智能和HPC工作负载的更高要求。UEC主席表示该项目构建在以太网技术之上，因为它是行业内持久、灵活且适应性强的基本网络技术的最佳例子；
- **用户广泛使用：**Infiniband具有高带宽、低时延转发的特点，传统的HPC网络通常采用其进行组网。但InfiniBand架构封闭，可扩展性不足，网络部署和维护成本高。随着以太网技术的发展，以太网在HPC和人工智能领域的应用规模持续扩大。最新TOP500统计的数据显示，全球HPC TOP500使用以太网互联的比例达45.5%，超过Infiniband。基于以太网技术的RoCE网络也被广泛被用于大模型计算集群，如鹏程·神农、华为·盘古、百度·文心等；
- **创新方案接连落地：**业界基于以太网技术持续开展创新，自2018年10月华为率先发布AI Fabric极速以太网方案以来，多个主流厂商积极推进技术攻关，并陆续推出用于HPC和人工智能领域的高速互联产品和方案；

主流玩家持续发布基于以太网技术创新的新产品

时间	厂商	事件
2018年10月	华为	发布AI Fabric极速以太网
2020年8月	HPE	HPC以太网互联技术Slingshot
2022年4月	浪潮	发布基于RoCE的无损以太网解决方案
2023年5月	NVIDIA	发布高性能以太网架构Spectrum-X
2023年7月	微软、博通、AMD、Intel等	联合成立超以太网联盟UEC

HPC TOP500 以太网占比超过Infiniband

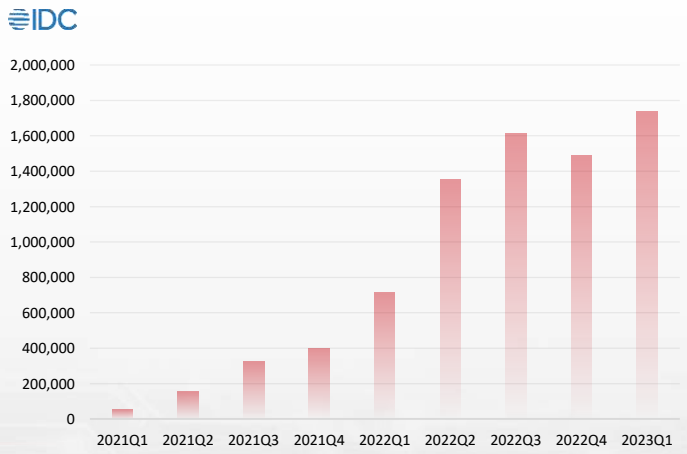


大模型广泛采用基于以太网技术的RoCE网络

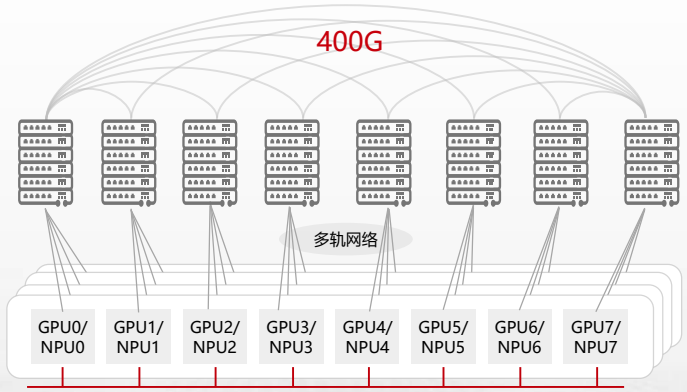
400GE交换机进入批量部署阶段，支撑超宽极简架构

- **生态开放，技术快速发展：**以太网一直是开放的生态，这为网络技术的迭代奠定了良好的环境基础。经过40 多年的发展，其速率已从10Mbps快速演进到 400Gbps，为人工智能场景下海量数据的高速流动提供了超宽的通道。IDC数据显示，在2021年到2023年的两年内，400GE端口的发货数量复合增长率达46%，2023年第一季度400GE端口发货量高达173万；
- **从标准到产品，产业成熟：**2013年，400G的以太网标准工作正式启动；2017年，IEEE 802.3bs以太网定义标准被批准，预示着400GE标准全面成熟。当前主流厂商均能够提供400GE交换机，华为在2019年已经率先发布了业界首款面向AI时代的最高密的400GE数据中心交换机CloudEngine 16800；
- **400GE构建超宽极简架构：**AI集群当前广泛采用200G/400G的高性能网卡，400G接入和互联需求凸显。基于大带宽的以太网交换机可以构建灵活的网络架构，满足用户不同业务场景的组网需求。其中，多轨网络架构和CLOS网络架构成为业界两种重要的选项。在分布式AI训练场景，多轨网络架构只需要建多个独立的网络平面，把同号卡连接起来。相较于传统架构，可以有效减少网络层级，降低数据转发跳数，降低建网成本；另一种是通常采用两层CLOS网络架构，网络的上下行收敛比需要满足1:1，这种架构的优点是通用性和可扩展性更强，可支持更大规模的组网需求；

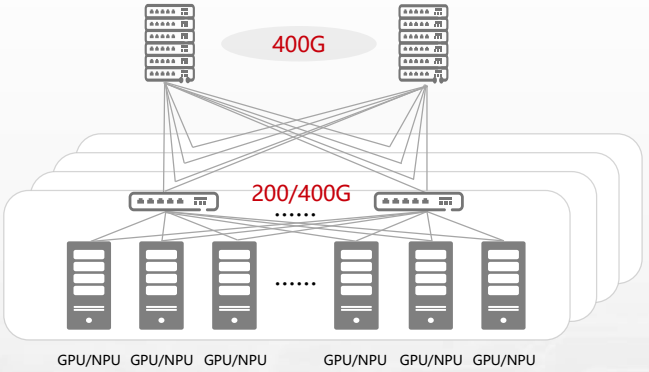
400GE端口发货数量统计报告



单层多轨网络架构，减少网络层级和建网成本



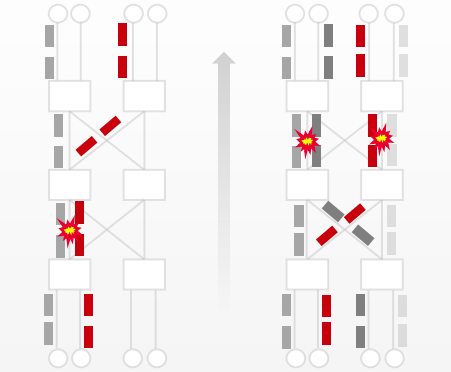
CLOS网络架构，通用性和扩展性强



从零丢包到零阻塞，提升AI训练效率

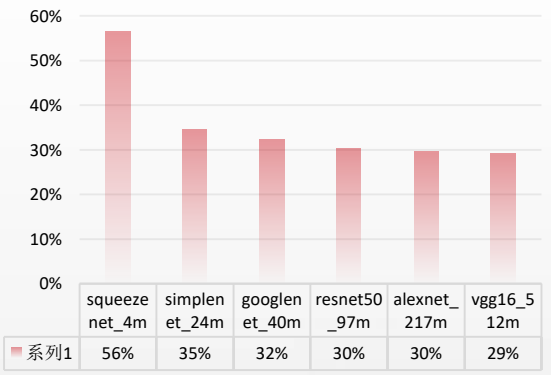
- **网络吞吐是AI训练效率的关键：**目前业界主流厂家已能够很好的解决以太丢包问题，但零丢包问题只是智算网络的基础。除此之外，还需要进一步提升网络的吞吐量。因为从技术上说，几乎所有的网络传输都有一个固有的问题，就是同一条连接在网络内要避免乱序，一旦发生乱序，在接收端就会触发重传逻辑导致降速。AI训练大流为主（100MB~几GB），流数量少，单流通信数据量大，基于传统模式进行负载均衡会导致网络节点仅站在自身视角将流量选路发送，会出现流量分布不均，常年吞吐率较低；每个周期内最慢的一条流到达后，才能进行下一轮通信，性能取决于最慢的流。在没有实现全局负载均衡的网络中，整体通信效率为30%~50%左右，这说明有一半的网络性能没有被使用，也就意味着整个集群的算力使用率仅仅为30%~50%左右；
- **网络级负载分担提升网络吞吐：**为了提升网络吞吐量，业界主流玩家的优化思路基本一致，即要想使RoCE网络适配大模型AI训练的需求，需要针对端、网和协议进行深度协同以及适配，实现整网负载均衡和90%以上的高吞吐性能，才能实现通信效率提升。目前华为通过网络级负载均衡（NSLB）技术，通过网络控制器和AI调度器协同，可根据整网交换机节点流拥塞状态和全网拓扑进行全局算路，并根据AI调度器分配的训练任务获取通信关系矩阵，结合通信库和网络拓扑、带宽、拥塞情况识别出最优路径，自动下发至网络交换机，业务流根据统一规划路径进行传输，整网吞吐可提升至90%以上；

传统负载均衡流量分布不均会导致网络拥塞，影响训练速度



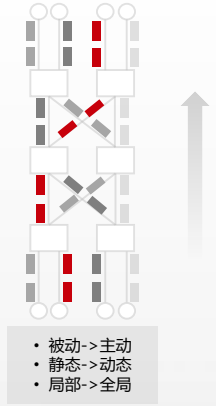
本地冲突
被动静态哈希，负载分担不均，leaf上行拥塞

全局冲突
无法预见下一跳冲突，spine下行拥塞

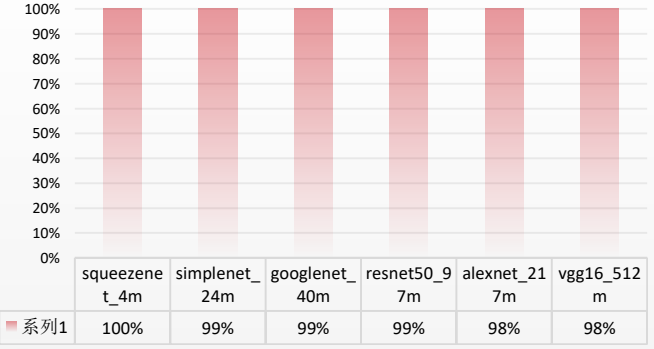


	squeezenet_4m	simplenet_24m	googlenet_40m	resnet50_97m	alexnet_217m	vgg16_512m
系列1	56%	35%	32%	30%	30%	29%

华为网络级负载均衡技术整网吞吐提升至90%以上



- 被动->主动
- 静态->动态
- 局部->全局



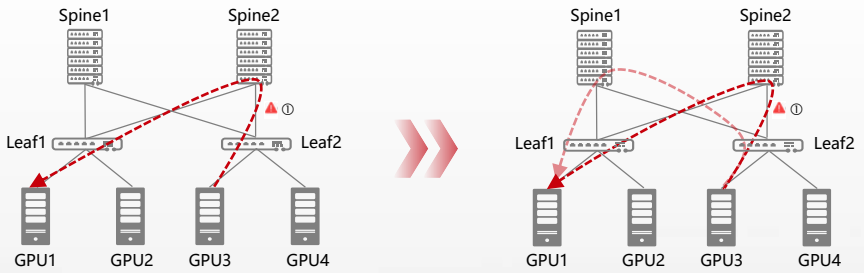
	squeezenet_4m	simplenet_24m	googlenet_40m	resnet50_97m	alexnet_217m	vgg16_512m
系列1	100%	99%	99%	99%	98%	98%

结合AI流量特征，实现全网动态负载均衡，网络无拥塞、满吞吐，支持多任务、多租户

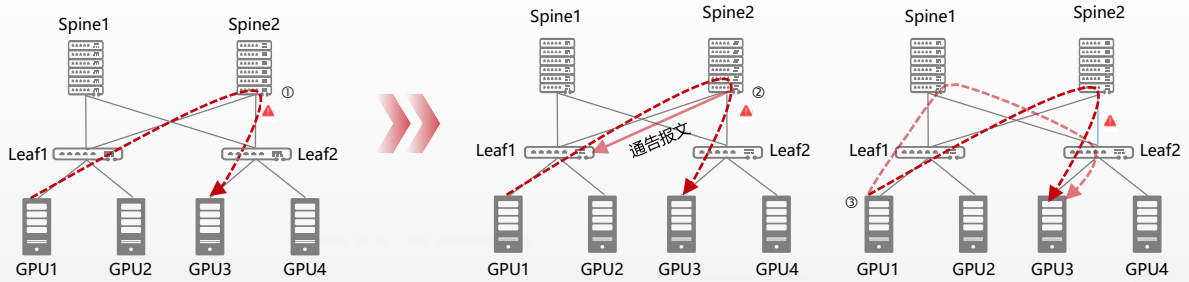
故障收敛突破毫秒，保障集群稳定可靠

- **网络高可靠是集群系统稳定的基础：**AI大模型依托的智算中心网络是业务流量的核心枢纽，其稳定性直接关系到整个集群系统的稳定性。一方面，一个网络节点的故障可能会影响数十个甚至更多计算节点的连通性，网络故障域大。另一方面，与单个GPU或者服务器容易被隔离不同，网络作为集群共享资源，性能波动会导致所有计算资源的利用率受到影响，具有放大效应。因此，在大模型的训练过程中，确保网络的持续稳定至关重要，提升网络的故障恢复能力和运维效率成为当前亟待解决的问题之一；
- **网络高可靠的技术创新方向：**
  - **(1) 硬件快速感知，亚毫秒级故障恢复：**AI训练场景里每次主机间通信任务时间在毫秒级，如果依靠传统的路由收敛方式，通过感知端口状态、路由收敛、转发路径切换等操作完成流量从故障链路到备用链路的收敛，时间一般在秒级，中断多轮AI主机通信，极大地影响了AI效率。针对这个问题，一种优化策略是利用数据面快速收敛技术，提供基于数据面的故障快速感知、本地快速收敛或远程快速收敛等能力，实现故障链路亚毫秒级快速切换，训练任务无感知；
  - **(2) 训前智能自检，训中智能运维：**据统计，90%的高性能网络故障是由配置错误导致。随着AI训练集群规模不断扩大，进一步增大了配置的复杂度。通过算网协同机制，设计符合 AI 场景的网络模型，完成网络配置的自动化生成、自动化下发和自动检测，被普遍认为是AI集群稳定交付的重要技术。此外，AI大模型具有流量大、周期短的特征，传统的轮询和报文采样机制无法支持AI网络流量的指标可视化，整网被视为一个黑盒。通过毫秒级的网络性能测量、网络与计算协同的集合通信性能测量，实现业务可视化、质差分析与故障定界；并联合集群计算运维平台统一调度，实现网络故障快速闭环，是行业探索的另一个重要方向；

本地快速收敛



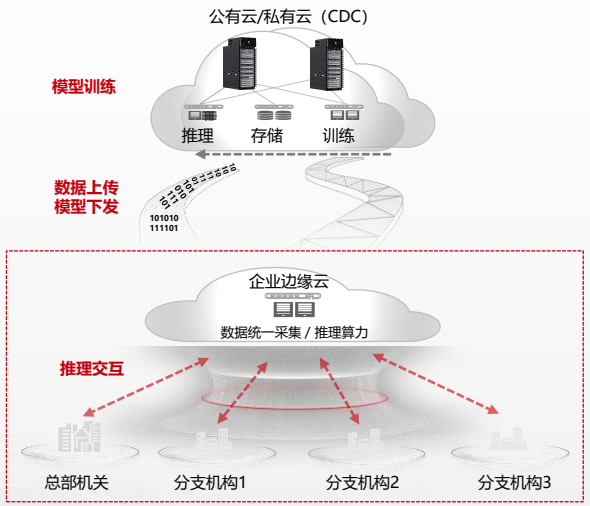
远端快速收敛



广域网络向弹性智能演进，加快AI推理速度

- **AI推理给网络带来新需求：** 随着AI技术的发展和行业智能化的深入，网络不但要联接人、联接物、支持传统应用，还要支持好大模型训练、分发、推理、迭代等智能应用的全流程。针对广域网络来说，一方面，模型云上训练，云下推理带来了海量数据流转，需要网络具备大带宽高吞吐的能力。另一方面，随着AI推理海量终端与应用走进企业核心生产系统，带来应用数量百倍增长，不同的AI应用对网络有不同的要求。比如工业园区网络中AOI机器视觉质检要求实时推理交互，软件包下载要求高峰值带宽，视频会议要求稳定带宽。网络如何提供更加精细化、差异化的体验保障成为新的挑战；
- **400GE/800GE构筑弹性智能广域网：** 目前行业正在探索利用400GE/800GE构建超宽网络，并通过网络与终端和计算协同、智能调度算法对应用进行智能感知和分析，准确预测网络的流量变化趋势，从而根据不同的应用类型智能地优化网络资源，提前消除网络拥塞，保障海量训练数据的高效流转 ，满足应用对网络时延、带宽等差异化服务保障的需求；

PB级训练数据上传，TB级模型文件下发
带来海量广域数据传递需求



AI进入生产系统，应用百倍增长
网络面临时延保障和吞吐不足双重挑战

弹性智能的广域网络
加速推理交互

行动建议：将开放的超高吞吐、极致可靠的AI网络带到每个数据中心

选择开放的以太网络技术	在构建数据中心网络时，应充分考虑技术的可扩展性， 以满足AI应用的不断增长。开放的以太网技术可以根据不同的业务场景和算力需求，灵活组网并兼容多样化算力接入。此外，使用开放技术可以避免被单一供应商锁定，从而增加议价权和选择权；
引入400G大带宽网络	网络性能已经成为决定AI训练效率的关键因素，当前主流GPU服务器网卡接口已经到达200G/400G，应当构建400GE的互联网络，并具备向800GE演进的能力，以支撑AI训练海量数据的高效传输；同时建议引入网络级负载均衡等领先的技术，构建超宽极简的无阻塞网络，提升网络有效吞吐和AI训练效率;
积极推进网络自动化、智能化	随着大模型的参数越来越多，网络规模也将成倍增长，随之而来的网络部署和运维管理复杂度指数级增长。数据中心网络应该尽可能地自动化、智能化，包括网络部署、配置管理、故障自愈等，在增强网络可靠性的基础上，将大幅提升企业的运营效率并使能全新的智能化业务;
关注算网协同保障应用差异化体验	AI网络是覆盖云、网、边、端全场景的端到端网络，包含数据中心网络、广域网络以及覆盖边和端的网络。算网协同被普遍认为是支撑实现AI大模型从训练到推理，从专用到通用的关键技术。通过实时感知应用，保障关键应用的差异化体验，加速推理实时交互。

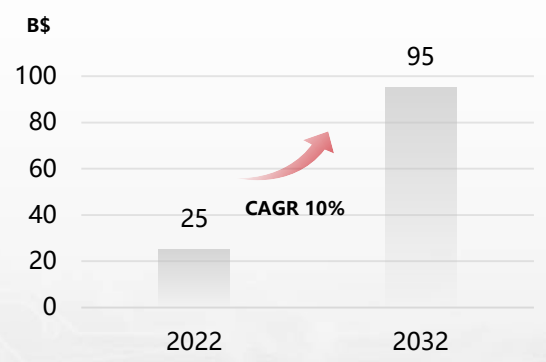
目录

新兴业务涌现驱动园区网络升级，更好支撑企业数字化转型

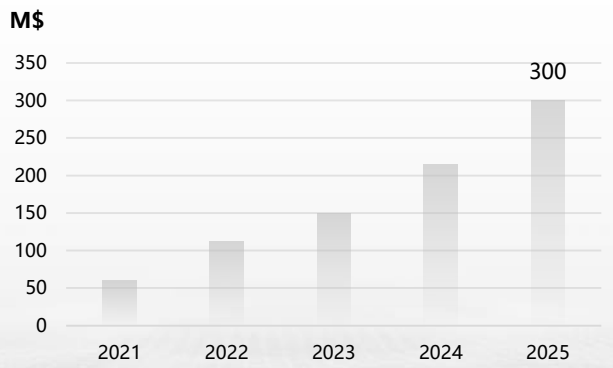
随着企业数字化转型加速，新兴业务及应用不断涌现，在提升企业办公及生产效率的同时也对园区网络提出了新的要求，目前看到主要的驱动力包括：

- **视频会议盛行：**视频会议成为企业远程沟通和混合办公的重要工具，预计全球每年将以10%增长，到2032年将达到950亿美元。以华为为例：视频会议连接着员工及上下游合作伙伴近40万用户，覆盖170个国家、1000多个办公点，每天最高峰在线用户达6万人，每个月有60多万场会议召开，视频传输的质量直接决定着企业内外部的沟通及交流效率；
- **物联应用海量部署：**除了传统办公类终端，企业数字化将催生海量的物联类终端，包括资产管理、电子价签/标签、高精定位、智能仪表、环境感应器等。以零售业为例，大量超市开始采用支持远程实时刷新价格的电子价签替换传统纸质价签，预计到2025全球电子价签市场规模将超过30亿美金。物联应用及设备的快速普及，会进一步增加企业IT系统的复杂度；
- **用户终端即将升级：**6GHz频谱的陆续发放推动Wi-Fi终端升级，截至 2023年上半年，支持6GHz的终端已达到2064款，同比2022年上半年增长了260%，这其中有67款设备支持新一代Wi-Fi 7，包括22款手机、30个路由器和网关、11个接入点以及4台笔记本电脑。终端的升级换代也将驱动企业加速WLAN无线网络升级换代，以更好适配新型终端的接入；
- **沉浸式体验等新兴业务逐步兴起：**算力以及视频显示技术的不断成熟，也在催生沉浸式应用的不断涌现，包括全息投影、裸眼3D、元宇宙办公等逐步在企业得到应用。如果企业考虑在近期尝试新技术创新，对于园区网络也会带来冲击；

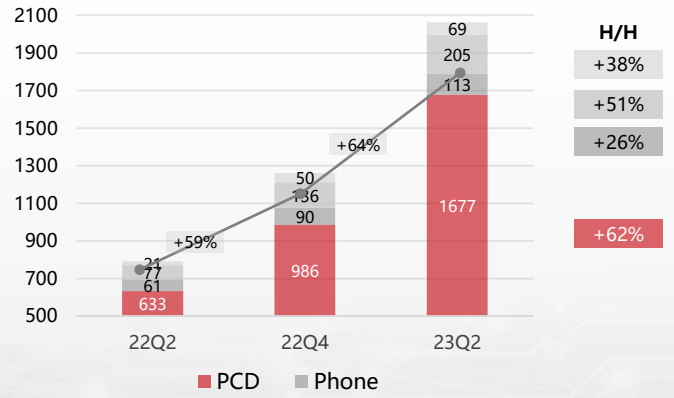
企业视频会议市场预测



2021~2025全球电子价签市场预测



支持6Ghz频段的终端快速普及

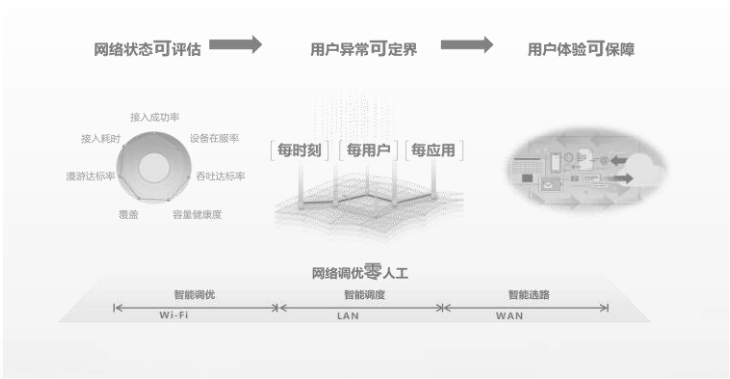


企业园区建网理念升级，打造面向未来的高品质园区网络

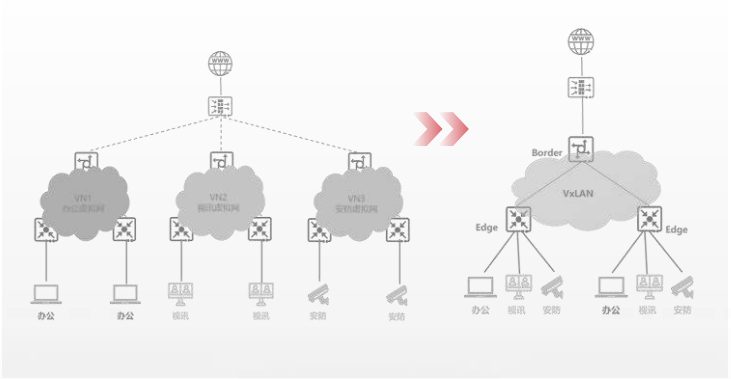
随着应用的多元化发展，企业网络承载的业务数量和类型快速增加，不同业务对带宽、时延、安全等诉求有所差异，需要改变传统的建网思路，全力构建以用户体验为中心的高品质园区网络；

- **从带宽驱动到体验保障：**园区网络的业务从传统电脑办公为主演进到高清视频会议为主，办公应用也正从本地走向云端，对园区网络带宽和时延提出更高要求，网络带宽是有成本的，因此我们要提升带宽效率，探索基于应用感知的差异化服务，保障用户体验；
- **从专网专用到融合承载：**园区网络中业务类型快速增加，专网专用的建网模式导致综合建网及运维成本高，网络资源利用率低，且会导致信息孤岛，无法实现企业数据的自由流转。通过网络虚拟化实现多重业务的融合承载成为刚需，为企业客户降本增效；
- **从网络自建到网络服务化：**传统园区网络的规划，部署及运维通常在本地开展，且依赖于IT运维人员的经验和专业技能，无论是建设还是运维效率都比较低下；我们需要基于云的自动化网络规划部署及智能化的运维方式，提升IT运维效率，让企业聚焦于自身业务发展；
- **从通信连接到通感一体：**在企业无线网络中实现通信和感知的功能，通信辅助感知，感知辅助通信，实现通信与智能生活的融合，推动各行各业的绿色发展；

从网络状态可评估到用户体验可保障



从多张网络到一张融合园区网络

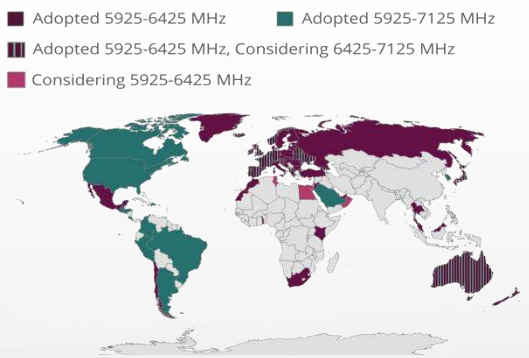


基于云化园区网络管控

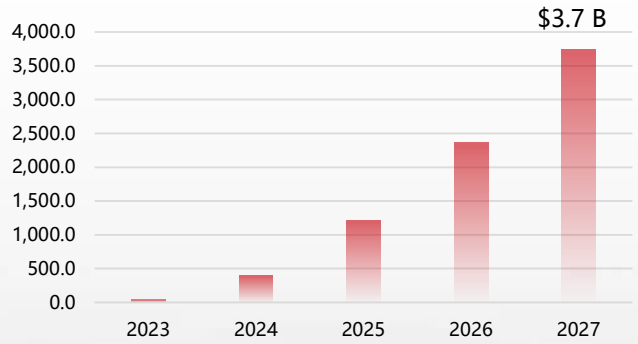
WLAN进入Wi-Fi 7时代，园区无线网络全面提速

- **Wi-Fi 7已经从标准定义进入商业阶段：**Wi-Fi 7标准—802.11BE已于23年7月完成Draft4.0的发布，预计将于24年Q1定稿并发布；芯片方面包括高通，博通，MTK，英特尔等厂家从2022年开始发布了Wi-Fi 7芯片。在频谱方面，全球54个国家已发布6GHz频谱给Wi-Fi，覆盖欧洲、亚太、中东、拉美等地区。华为将在23年9月正式发布业界首款企业级 Wi-Fi 7 ，其他大多数制造商将在2024 年将Wi-Fi 7产品推向市场；
- **企业级Wi-Fi 7市场即将进入快速发展期：**根据Gartner预测，预计到2027年，企业级Wi-Fi 7 AP的出货量将达到1240万台，相当于AP总量的 27%；尤其在制造行业市场中超过30%的组织将升级到Wi-Fi 7，这将为他们的业务流程引入更多用例。与此同时，Wi-Fi 7与时间敏感网络（TSN)的组合在带宽以及可靠性方面全面提升，支持制造和仓储中的关键业务流程，有助于新一代Wi-Fi技术的加速采用；
- **Wi-Fi升级将同步带动有线网络升级：**Wi-Fi 7的峰值速率超过10Gbps，2.5GE将成为Wi-Fi 7 AP最低要求，这意味着传统千兆交换机已经无法支撑新一代企业无线网络的带宽需求，这意味着接入交换机正在从GE朝着2.5/5GE升级，进而带动25GE汇聚，100GE核心的市场发货；IDC发布的市场份额数据显示2022年2.5/5GE端口发货量同比增长108%，25GE同比增长78%，100GE同比增长62%，同时预计在未来几年仍将保持高速增长势头，推动企业园区网络全面进入万兆时代；

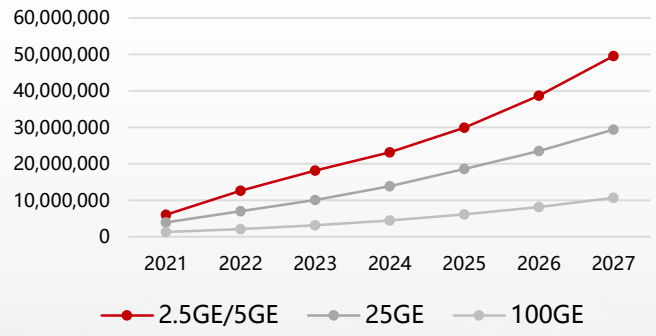
全球已经有超过50个国家发布6 GHz频谱



全球企业级Wi-Fi 7市场收入预测



园区多速率/25GE/100GE端口市场发货预测



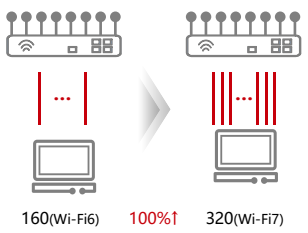
从带宽到可靠性，Wi-Fi 7加速海量行业场景化应用

- **更高带宽：**相比Wi-Fi 6，Wi-Fi 7在2.4G/5Ghz频段基础上新增支持6GHz频段，在降低信号干扰的同时提供更宽的频谱资源，160Mhz连续组网成为可能，配合4096-QAM将用户带宽提升2.4倍，轻松满足4K视频，AOI高清质检，车载软件灌装，AR/VR等高带宽诉求；
- **更低时延：**相比Wi-Fi 6，Wi-Fi 7基于Multi-RU特性实现空口RU的灵活组合，为单用户分配多个RU，提升空口资源使用效率，可将用户平均降低时延25%以上，特别适合高品质办公场景，为高清视频会议，交互式办公，云端多媒体渲染等时延敏感业务提供更好保障；
- **更高可靠性：**在链路可靠性及用户体验保障方面，Wi-Fi 7也取得了较大改进，引入Multi-Link Operation特性，让终端与AP之间可同时建立多条数据连接（2.4Ghz，5Ghz及6GHz），三条链路可同时收发数据，增加链路带宽；也可收发相同数据（多发选收），提升链路的可靠性；也可支持基于应用识别的数据链路匹配，实现差异化体验保障；这将为AGV智能仓储，柔性制造提供更优选择；
- **适用场景：**Wi-Fi 7可广泛适用于终端无线化改造场景，如智能生产线、智能仓储，工业终端控制，车辆路测，未来的元宇宙等；

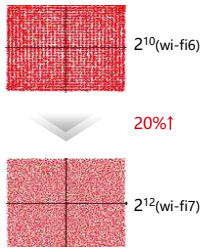
更大无线带宽

单终端峰值达到5Gpbs

Spectrum Bandwidth:
160MHz-> 320MHz

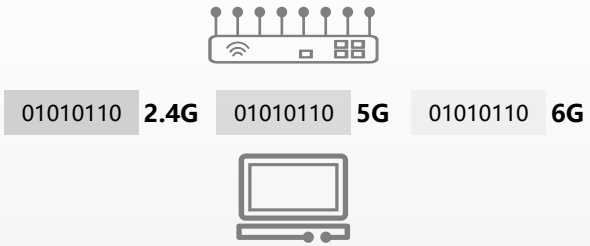


QAM:1024->4096



更高链路可靠

单链路 升级为 多链路



工作模式1	工作模式2	工作模式3
高带宽，同时收发数据	高可靠，数据多发选收	优体验，应用差异化传输

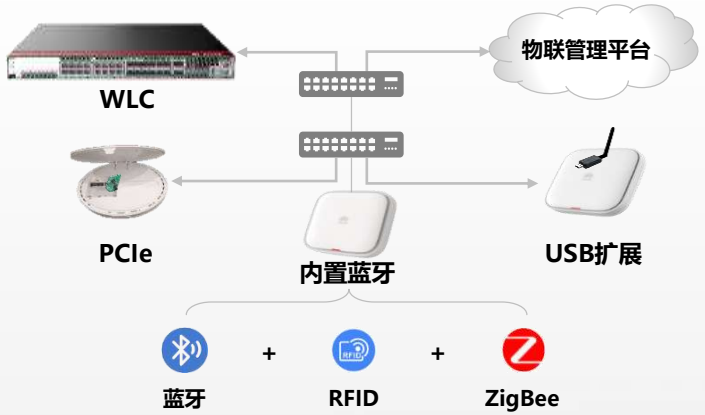
催生广阔应用场景

元宇宙	制造AOI	工业控制
		
高带宽低时延 10Gbps+, 毫秒级	高带宽 7~8 Gbps (站点)	低时延 5毫秒
远程医疗	**仓库**	**AR/VR教育**
		
低时延、高可靠性 5ms，多链路	高可靠性 多链路，0中断	高带宽低时延 1Gbps+,<5毫秒

从多张网到一张网，融合网络架构实现企业网络投资最优

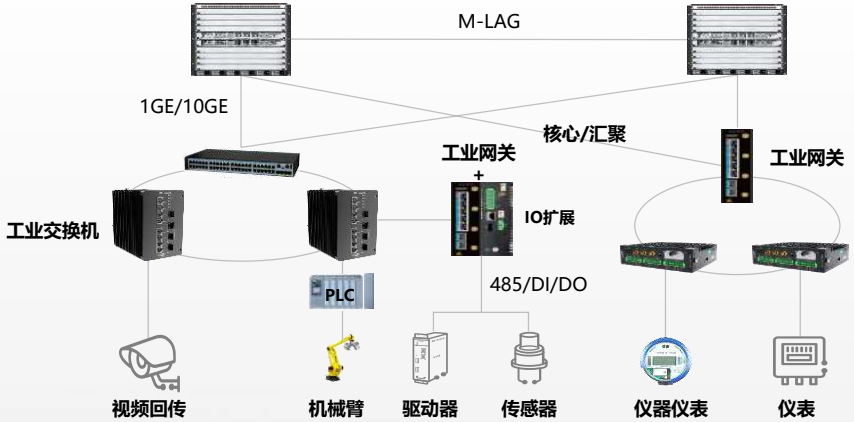
- **Wi-Fi&IoT融合：**园区内IoT应用的出现使得企业IoT基站采购、部署及运维诉求激增，考虑到IoT基站和AP的工作原理类似，且WLAN作为全无线时代首选接入方式，AP已遍布在企业办公及生产区域，基于WLAN AP进行IoT应用扩展让成为更优选择，当前主要实现方式有PCIe插卡，USB dongle，内置蓝牙等实现方式，该方案可大幅降低综合建网和运维成本；
- **生产网和办公网融合：**园区通常存在办公、视讯，安防，生产，IoT等多种业务，专网专用的建网模式综合成本高，占用大量机房空间，产生大量综合布线，但网络利用率通常在5%以下；随着VxLAN被引入到园区网络，多业务的融合承载成为可能，由一张物理网络同时承载多中业务，同时通过应用识别、网络切片等技术，为不同业务提供差异化的优先级调度，保障业务质量和用户体验；
- **适用场景：**未来可在教育，医疗，零售，大企业等行业的物联应用、多业务融合承载场景，广泛部署和应用；

IoT与Wi-Fi 融合组网



共用WLAN AP作为接入点、共享有线回传资源，建网成本节省30%以上

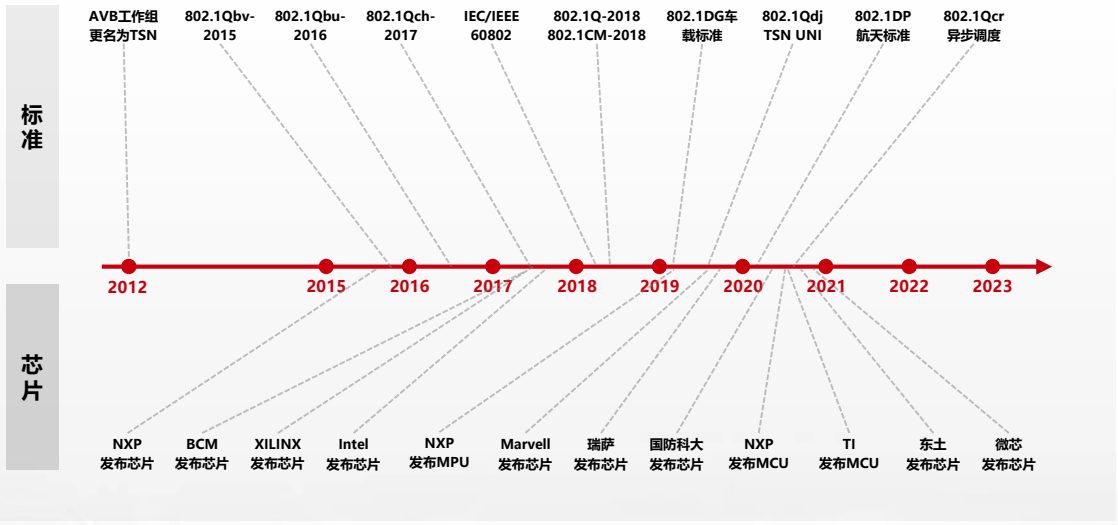
生产与办公 融合网络

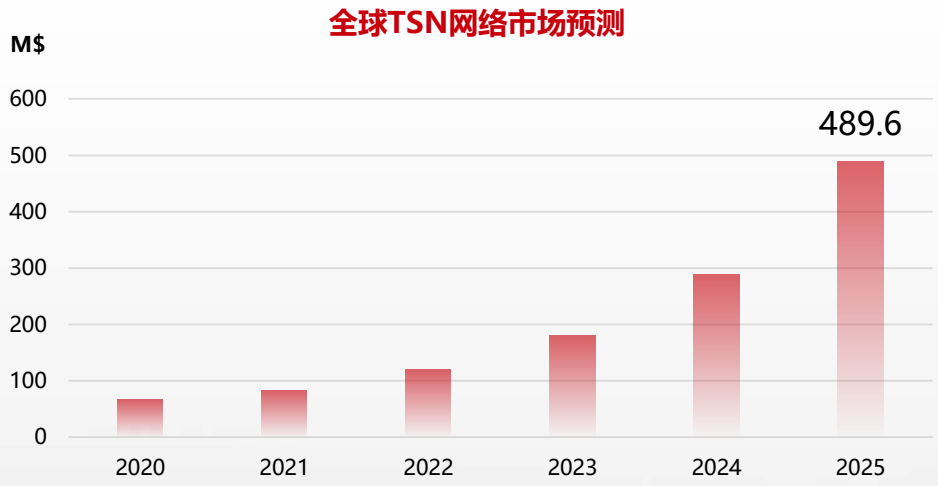


由一张物理网络承载多种业务，彼此安全隔离，综合建网成本节省50%以上

从不确定到确定性时延，时间敏感网络（TSN）实现IT/OT融合

- **TSN技术飞速发展：**当前制造行业主流连接方式是使用工业总线和工业以太。传统的IP/Ethernet网络，虽然具有开放性好、互通性好、产业成熟、带宽大、成本低等优势，但是其网络服务是统计复用的、尽力而为的，不能提供行业所需要的时延确定性保障。而基于场景设计的工业以太网络，通过特定的方法实现了时延有界，但是其互通性、可扩展性较差，且使用专用的软硬件，对于用户而言成本更高。时间敏感网络（TSN）是一项兼具传统以太和工业以太两者的优点的技术，为用户提供低成本、大带宽、支持统计复用的网络基础设施，解决各种总线、工业以太协议互通难的问题；又具备提供有界时延、极低时延、自动化网络配置、高可靠性等性能优势；
- **TSN商用进程加速：**TSN支持采用周期性网络传输机制，借助时间同步+精准调度可为业务提供微秒级的确定性时延保障，满足生产、制造、交通场景下时间敏感型业务的传输诉求；当前，TSN技术标准（有界时延、资源管理、时间同步、高可靠性等）完成了发布，众多芯片厂家陆续推出了满足标准的芯片，包括华为在内的10+设备厂家发布了产品及方案，异厂家的互联互通得到充分验证，在北美及中国已有商用部署案例。同时，Wi-Fi 7在带宽、时延及可靠性方面优势明显，可为TSN网络提供无线侧的灵活性和扩展性，通过Wi-Fi 7无线+TSN有线结合为工业自动化、机器人等应用提供更多可能，加速商用进程；

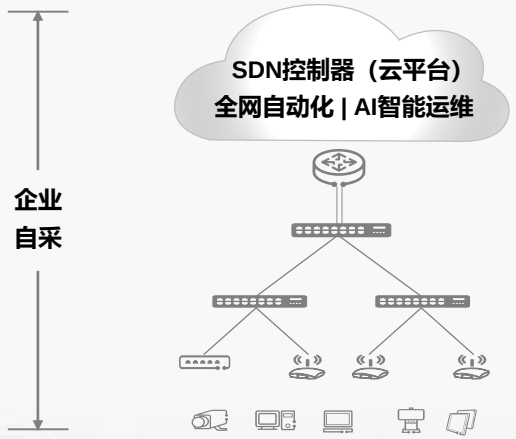




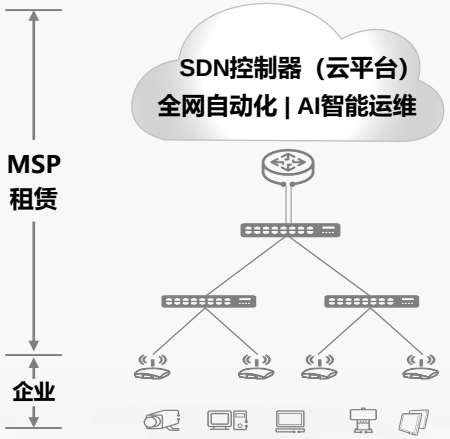
从设备采购到网络即服务（NaaS），企业数字化创新开启加速度

- **越来越多的企业从采购设备转向采购服务：**过去，企业通过一次性购买硬件、软件、许可和服务来建设园区网络，这些硬件、软件、许可证和服务有时会被打包在一起，同时网络运营管理也要求IT团队具备专业的数据通信的规划，部署及运维管理优化能力，以确保园区网络稳定及安全运行，由于受限于已购买的硬件和基础设施，传统网络的灵活性通常较低。随着数字化业务的深入，网络复杂度与日俱增，企业对于灵活创新的需求越来越旺盛，基于云管理的模式，网络即服务应运而生，其具有高度的灵活性，可以定制网络配置并选择所需的特定服务，实现敏捷运营、服务定制和灵活计费模式，来支持复杂的网络和多云环境；
- **网络即服务（NaaS）的优势：**首先企业无需购买、拥有或维护网络基础设施，即可运营和控制网络，可以根据需求扩大或缩小规模，快速部署服务，并降低或消除与硬件相关的成本。 其次在使用过程中允许企业通过按期订购和按用量计费的灵活商业模式，财务的灵活性可以使企业能够在快速变化的业务环境中灵活选择服务模式。此外NaaS服务还可以保持软件的实时更新以及增强网络安全特性，加快业务创新速度和降低安全漏洞造成的风险；
- **市场趋势：**NaaS模式近年来保持着快速增长势头，根据 Mordor Intelligence的预测，从2023 年至2027年，NaaS市场的复合年增长率将达到34.5%，预计到2027年全球园区NaaS服务市场收入将超过6亿美金；

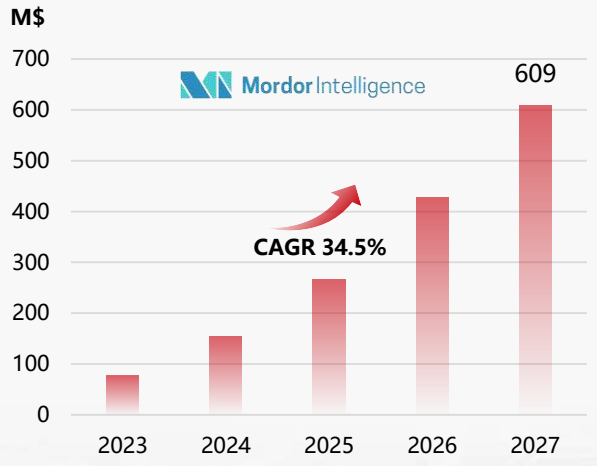
传统模式：企业自己采购设备



NaaS模式：企业向MSP购买服务



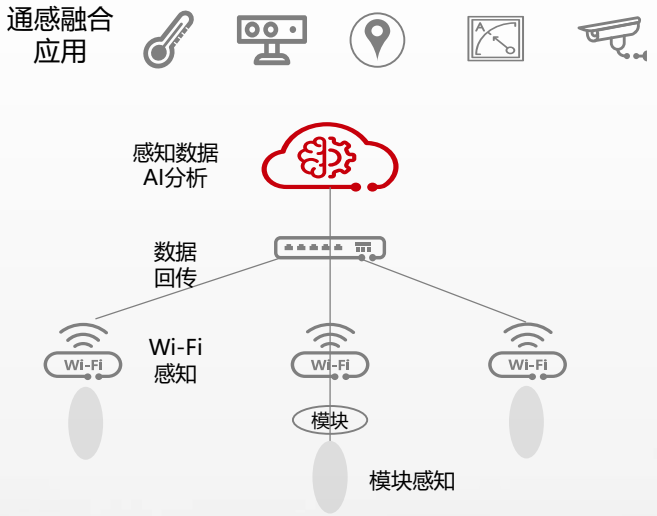
园区NaaS市场快速增长



从网络通信到通感一体，赋予数据通信网络感知世界的能力

- **Wi-Fi 设备走向智能感知：**Wi-Fi不仅可以用于通信，还可以用于感知，Wi-Fi 传感使用 Wi-Fi 波来检测运动和存在，然后应用机器学习算法以促进高级应用。可将无线设备转变为传感器，能够通过无线的方式来进行高精度的身体定位和动作识别，并把识别的结果转换为指令，并实时传回控制系统，计算人和物体物理空间中信号的干扰和反弹收集有关人和物体的数据，已建立的 Wi-Fi 设备将成为用于确定特定区域内人和物体位置和网络交互的一部分；
- **Wi-Fi感知的标准进展：**2020年9月IEEE 802.11bf标准组成立，开启通感一体化标准化研究的序幕，它不是用于数据通信，而是用于传感。23年1月发布标准规范Draft1.0，7月发布Draft 2.0，计划11月发布Draft 3.0，2024年1月发布Draft 4.0，届时将有早期的产品及方案问世，而正式标准将在2025年发布；
- **开启广阔应用场景：**通感融合，提高无线系统的性能和效率，为更多新应用场景提供可能，如实现高精度生理性摔倒检测实现康养监护，存在检测实现节能减碳；

Wi-Fi网络走向感知



毫米波

驱动广阔应用场景创新


低空安防


智慧交通


健康监测


导航/跟踪


智能电力


手势感知

...

- 室内感知
- 家庭安全
- 音频追踪
- 存储感知
- 家居控制
- 手势识别
- 生物感知
- 人脸识别
- 距离态势检测
- 摔打检测
- 远程诊疗
- 喷嚏感知
- 车内感知

行动建议：数字化应用的推进速度远超想象，改变园区网络建设理念

超宽接入，物联融合

企业存量Wi-Fi 5 AP面临设备老化及过保风险，如今换代升级已迫在眉睫；新增办公生产区域或无线化改造场景，WLAN成为刚需，如AOI高清质检，车载平台灌装升级，AGV智能仓储等。无论是WLAN换代或WLAN新建场景，推荐选用Wi-Fi 7，为用户提供带宽倍增能力的同时，提供更低时延，更高可靠性；IoT共站场景，推荐采用IoT融合AP，为企业节省综合建网成本；

一网多用，体验保障

在园区接入推荐多速率交换机，满足高性能Wi-Fi 6/7大带宽回传需求，同时为有线终端提供超千兆接入服务；园区汇聚推荐高密25GE交换机，园区核心推荐100GE交换机，从而打造10G接入/25G汇聚/100GE核心的全无线办公网络，为用户提供万兆极速体验；园区拥有多种业务时，推荐多业务融合承载方案，在一张物理网络上为多种业务提供差异化策略，保证用户业务体验，提升园区网络资源使用效率，降低网络部署成本；

云化管理，智能运维

推荐采用SDN控制器实现对园区有线及无线网络的统一管理和控制，业务配置自动化发放，提升IT运维人员网络规划及部署效率；同时基于Telemetry实现网络、设备、用户及应用的实时可视，发生故障后快速定位定界，智能分析根因并完成故障处置，简化园区网络日常管理运维及故障排查的难度，从而提升园区内网络用户的满意度；

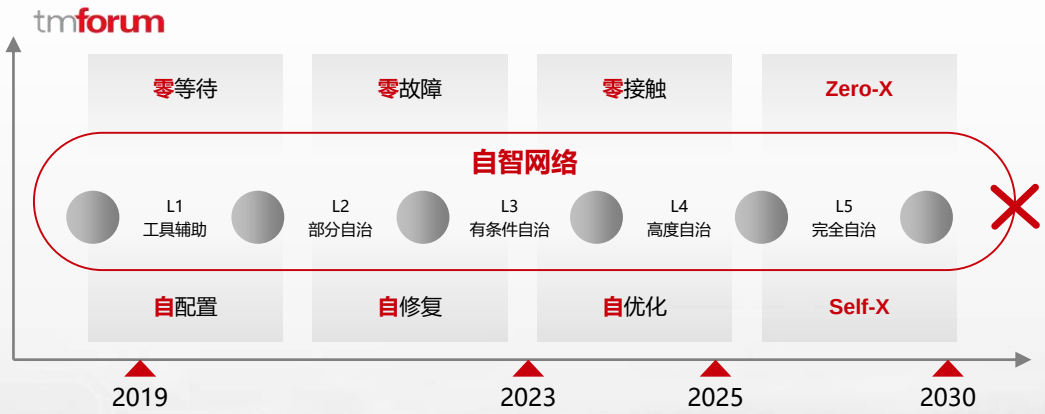
- 在大型网络场景（含多分支），企业有独立的IT运维团队，推荐自建云管理平台（控制器）进行自身网络的日常运维和管理；
- 在中型网络场景，企业网络投资紧张且无运维能力，推荐采用Naas模式完成网络建设及运维托管，节省初期建网成本，降低投资风险。

目录

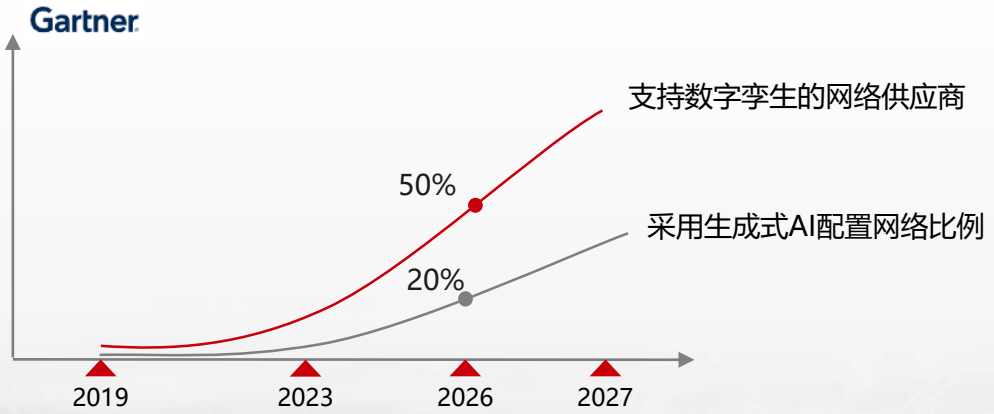
人工智能快速发展，网络智能化的规模部署拐点已至

- **网络智能化的定义和发展：**网络智能化是指通过实时数据收集、关联、预测来快速发现和隔离问题，使网络人员不需要深入的配置和故障排除技能来管理网络，人工智能在应对不断增长的网络复杂性方面发挥着越来越重要的作用，提供了巨大的潜力，可以颠覆长期存在的传统网络运营，从而大幅提高生产力。电信管理论坛(tmforum)从2019年开始引入网络自动驾驶概念，将通信网络分为L1~L5五个等级，目标是实现“完全自治网络”，目前业界的通信网络大致处于L2~L3的水平；
- **网络智能化面临的挑战：**网络智能化，有效数据是关键，过去由于网元设备自身的分析能力有限，运维人员难以从海量日志、告警信息中获取有价值的网络建议，即便是有经验的网络工程师，很多时候也难以输入准确的意图；另外一方面，即便是有可靠的网络建议，由于缺乏全方位的可视手段，运维人员也担心风险，在未看到实际数据以及效果之前，并不愿意完全信任人工智能。因此迄今为止人工智能网络应用大部分还是针对某些具体问题的点状应用，难以形成大规模系统级，整个人工智能网络的采用率不到10%；
- **网络智能化即将进入快速发展期：**近几年来，算力的快速进步催生了数字孪生和各种大模型广泛应用，数据的有效性以及对于业务影响的可视程度得到了一个全面的提升，这将推动网络智能化即将进入一个快速部署的阶段。根据Gartner的预测，到2026年，50%的网络供应商将在其解决方案中提供数字孪生功能，20%的初始网络配置将采用生成式人工智能技术，而到2027年，使用人工智能功能来自动化网络运营的企业会从现在的10%提升到90%，人工智能本身的快速发展，推动网络智能化打开规模部署的大门；

L5网络自治系统定义及时间轴



人工智能网络创新趋势洞察



数据通信网络智能化进入系统级，数字孪生和生成式AI成为关键

- **系统级网络智能化是广泛部署的核心：**AI的应用在各行各业都需要一个过程，通常分为三个层次，第一层是点级别解决方案：AI解决非常具体的问题，用于改进现有流程且可独立部署，不改变系统，包括告警压缩、Wi-Fi体验保障以及站点自动开通等都属于这个层级；第二层是应用级解决方案：AI解决一系列问题，使能独立可部署的新流程，也不改变系统，数据中心网络中的知识图谱自动识别、定位并解决问题属于这一类，能够适应部分场景，但无法做到整个系统重构。第三层是系统级解决方案：AI能够同时改进多个现有流程，或者通过改变相互依赖的流程使能多个新流程。在数据通信网络领域，数字孪生以及生成式AI属于具备重构多个流程的技术变革，已经开始进入系统级解决方案的阶段，这也是能够支撑网络智能化广泛部署的关键；
- **系统级解决方案1：**数字孪生，从2018年开始，基于数字孪生技术的网络数字地图开始在现网中开始逐步应用，最初从园区和数据中心网络，主要是提供网络可视以及提供部分网络故障修复，如今数字地图已经开始广域应用，把网络数字孪生体作为网络的基础运维平台来实现低成本试错、加快创新迭代、提高网络智能运维水平；
- **系统级解决方案2：**生成式人工智能，2022年底，随着ChatGPT将生成式 AI 推向前沿，网络大模型迎来快速发展期，网络智能化将越来越多地包含生成式人工智能，它可以根据人类输入创建详细的配置和故障排除程序，而无需明确的模板，目标是赋予其业务意图转换成网络需求的关键能力，打造业务意图引擎；

点级：现有流程某个环节增强

故障处理

告警事件压缩环节：基于规则→AI模型（改进现有流程）

站点开通

查验环节：人工查验→AI查验（改进现有流程）

Wi-Fi体验保障

WiFi调优：人工调整→基于AI的一键调优（改进现有流程）

……

应用级：基于数据+AI构建新流程

AI知识图谱，DCN网络故障闭环

基于数据+AI使能的新流程

运维助手

原来靠人工查→现在基于外线助手APP实时给出用户影响、故障具体位置、恢复方案(新应用流程)

基于网络数字仿真的倒换

基于数字孪生仿真重新设计的倒换新流程

……

系统级：基于数据+AI同时重构多个流程

网络数字地图，业务全息可视

基于数字孪生+智能算法重构网络可视流程

生成式AI构筑网络意图引擎

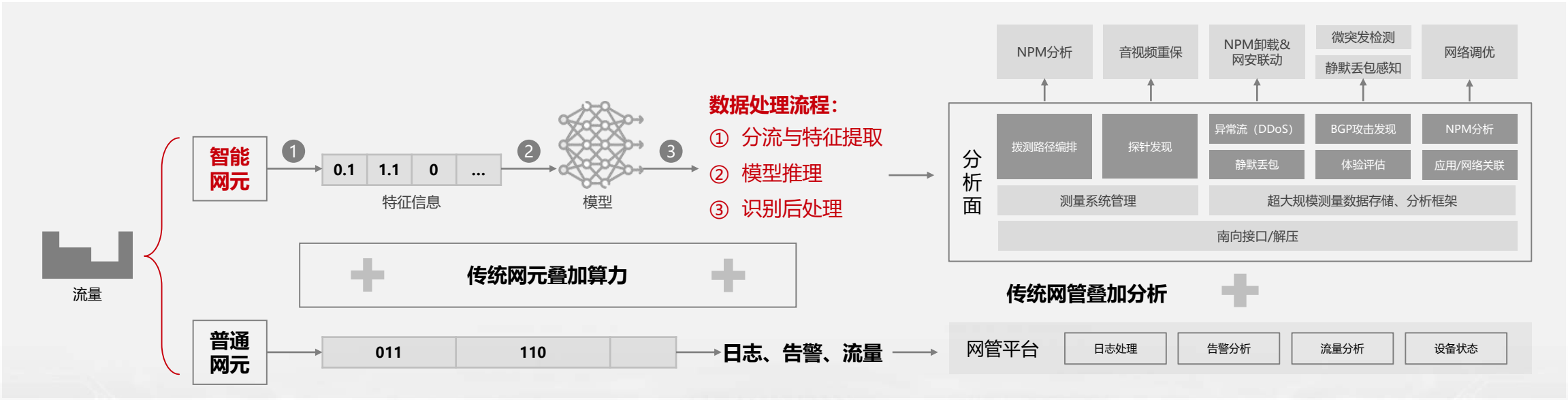
网络大模型(NetGPT)，重构网络意图交互流程

……

AI商业应用的三个层次

从接受意图到生成意图，智能网元是网络智能化的前提

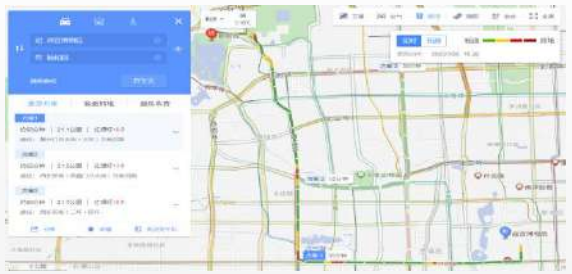
- **智能网元是网络智能化的前提：**数据是实现网络智能化的关键要素，网络数据主要有设备网元产生，如果网络设备单纯只是日志以及告警，这些数据对于运维人员而言，在大部分时间也很难形成准确的网络意图，导致被动运维，如果网元不智能化，整个网络是难以智能化的，因此网络智能化成为必然趋势，其重要意义在于，可以让网元从接受意图到生成意图网元，设备做自身的分析，多维度数据信息收集、预处理、上报，网络根据实际的流量变化，生成人分析不出来的意图，为AI提供数据基座；
- **网元智能化的关键价值：**网元的智能化可以通过深度学习模型，基于报文流行为特征，对数据流进行分类，并根据推理结果对流做出处置动作，提升设备关键能力，如业务保障能力，及安全检测能力。以金融骨干网场景为例，应用流量异常会对高价值业务造成冲击，导致业务出现故障，但这种网络流量的异常是完全可以通过设备自行发现并上报风险的，网络设备通过智能化升级，提供应用流量异常感知服务，异常流秒级识别，支撑流量限速与疏导闭环，保障高价值业务SLA性能。此外在安全防护DDOS攻击方面，路由器秒级DDoS攻击检测仅依赖流速变化，无法感知攻击细类，且对现网链路故障多切一等特定场景可能存在误报；设备自身协议攻击防护能力弱，易被攻陷，而智能化设备可以针对过路流量提供攻击细类感知能力，秒级触发上报，与清洗设备快速闭环；



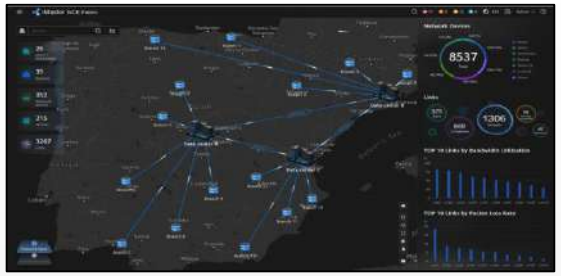
从多维可视到优化仿真，数字孪生地图开始规模应用

- **网络数字地图全面应用：**网络数字孪生体作为物理网络设施的数字镜像，与物理网络具有几乎相同的网络拓扑、业务及流量数据模型，是真实物理网络全生命周期、多维度的精细化副本，可以为网络运维提供真实网络的数字化验证环境；相比传统的仿真技术，网络数字地图不只是静态的网络快照，不但可以根据网络的状态实时更新；与AI技术相结合，能自我学习，使得网络的数字孪生可根据在线预验证反馈的结果自我演进，具备更高的真实性和可靠性；
- **技术趋势1：**从单维可视到多维可视，业务网络关联感知：普通的网管系统更多是针对网络质量单维度的可视化，但实际运维中更经常出现的是业务发生故障但无法找到网络故障，或者网络无法快速自证清白，以某银行为例，上百个应用与上万个网元之间没有建立可视的关联，导致运维效率低下，因此网络数字孪生不仅仅需要够构筑网络层面的可视，还要基于网络数字孪生引擎构建物理网络统一模型的数字化副本，面向运维场景以多领域、多维度仿真验证为基础，支撑网络规建维优活动，包括应用内、应用间互访关系呈现业务异常秒感知；
- **技术趋势2：**从离线仿真到实时仿真，降低网络变更风险：网络仿真按照场景分两大类，过去主要是以用于网络规划的离线仿真为主，但随着网络自动化水平越来越高，网络故障、自助申请一类的业务闭环周期短，期望在分钟级、秒级实现，因此实时仿真技术开始成为热点，根据网元配置数据模拟设备路由协议控制面和转发面行为，精准生成网元协议路由表，全局路由表，基于路由表项进行分析，完成对网络影响分析验证；

交通数字地图和导航



网络数字地图



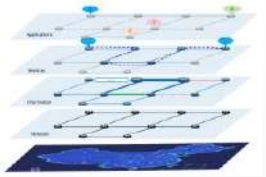
Orange Spain 实践

- 网络全息可视：全网SLA可视
- 网络性能最优：P3测试网络时延降低30%
- 网络自治优化：路径调优时间 3个月→3分钟

持续升级

多维度可视

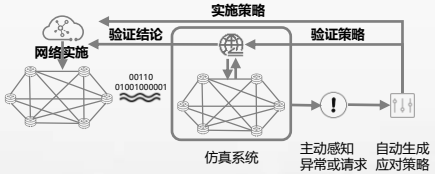
- 全网一图可视、
- 实时拓扑还原
- 应用网络互视



应用
业务
信息
网络

网络实时仿真

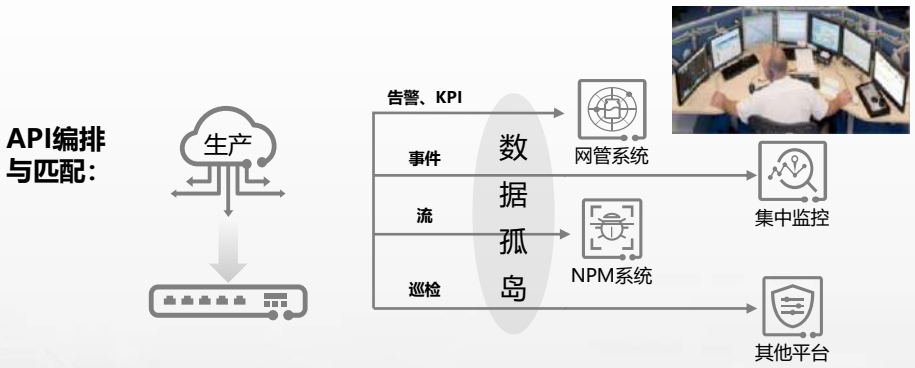
- 路径一键导航
- 多云变更仿真
- 全息主动感知



生成式人工智能为网络大模型（NetGPT）发展带来机遇

- **生成式人工智能驱动网络大模型发展：**大模型”百模千态“的个性化发展为通信网络大模型的出现提供了绝佳的机遇，应用层大模型必须经过网络才能和终端用户连接起来，而通信网络具备“瘦腰”的特性，因此要求存在一个通信网络大模型（NetGPT），使用统一的大模型处理多类网络业务，减少业务复杂度，是未来趋势；当前网络大模型还处于起步阶段，应用场景还在初步探索中，目前主要有大模型赋能知识管理平台以及交互式网络意图引擎处理等几个场景，面向未来，网络大模型可能会发展到在云边协同部署，这样有助于有效编排异构分布式通信和计算资源，是网络大模型NetGPT发挥重大作用的关键一步；
- **探索场景1：知识管理平台，大幅提高交互问答准确度，**传统的AI问答助手基于知识图谱技术，如果用户提问的关键字存在偏差，那么系统回答的准确率就比较低，大模型训练基于网络构建数通领域语料数据集，以及NetGPT大模型知识扩展重训练+任务对齐微调，精准理解用户意图，提升交互效率，突破通用大模型 + 专用小模型的有效协同机制， 应用于智能问答、网络体验保障的场景，特性类回答率从20%提升到80%；
- **探索场景2：人工意图引擎，全面理解网络意图，**过去虽然有网管平台，但是网络运维工具多，界面入口多；处理问题时，不知道从何入手，应用/业务流数据与网络数据割裂，缺乏关联，依赖专家经验来协同。网络大模型可以识别用户意图，区分类型，精准理解语义并完成网络需求转化，同时根据维护人员输入创建详细的配置和故障排除程序，而无需明确的模板，提升运维效率，降低运维团队的专业依赖度；

As-Is：用户上手难，故障排查复杂



To-Be：智能识别意图，推荐闭环措施

行动建议：将AI技术与人类智慧结合，加快创新速度，提升网络运营效率

部署智能化的网络设备	优先选择具备一定算力能力或者具备算力扩展能力的网络设备，对于网络智能化而言，单纯依靠网络控制器以及分析组件肯定是不够的，设备的智能可以实现从静态的推理和被动的响应，变为动态的分析和主动的推荐，这无论对于实现网络全局可视以及网络自治都是非常必要的；
采用数字孪生技术实现全面网络可视	对于数据通信网络而言，使用数字孪生技术首先实现网络全面可视，是网络智能化的第一步，有了数字孪生网络这个平台，现网实施的调整、维护、优化等变更操作，都可以先在数字孪生网络中进行充分的试验和验证，并通过其反馈来不断的评估、修正、优化操作方案，最大限度降低对真实网络带来的冲击；同时数字孪生网络还会实时记录网络的数字孪生体的状态和行为，支持对历史的追溯和回放，从而能在不影响网络运营的情况下完成预验证，极大地降低试错成本；
加速生成式AI应用创新	保持对于网络大模型的研究以及应用进展的关注，随着生成式人工智能对于行业的拓展速度不断加快，可以考虑尝试在智能问答、网络配置指导以及运维交互方面尝试引入一些创新，引入有远见的厂商和解决方案，将生成式AI扩展到本地网络领域，实现创新的价值；
拥抱AI从现在开始	网络智能化对于提高网络可用性、性能和运营效率有明显的效果，随着AI技术的不断进步，应用的价值已经开始显现，规模部署开始进入拐点。AI的本身并不是全面取代人，而是更好地辅助人，人工智能网络将全面提高网络可用性、优化效率、提高性能，使用相同或者更少的资源做更多的事情。

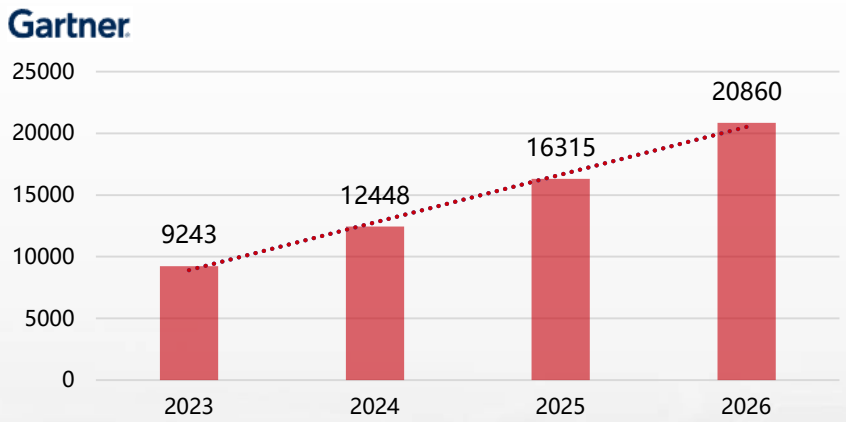
目录

企业上云和混合办公打破安全边界，网安一体化协同防御成为主流选择

- **企业上云打破防御边界：**传统的网络安全体系架构是以企业内部为中心，在企业边界构筑层层防御体系以确保数据安全，然而随着企业上云逐渐兴起，网络从传统封闭架构走向多云多分支互联，企业的边界被打破，网络安全的风险暴露面增多，同时面临生产高可靠、高安全的诉求。原先的集中的服务访问与安全体系显得越来越无效和繁琐。传统网络安全技术无法处理网络外围面临的日益高级的威胁和漏洞。随着外部访问的加速，企业需要实施高级访问控制，以确保具有处理相关网络安全需求和风险的能力；
- **混合办公增加安全风险：**企业员工不再局限于固定场所办公，混合办公成为常态，这意味着员工可能会在任意时间以及任意地点通过无法保障安全的互联网接入到公司网络，原有的基于企业局域网边界的安全架构不再有效，这对于企业数据安全的防御带来了新的挑战，企业需要考虑如何随时随地保障员工安全接入到企业总部以及多云平台；
- **以SASE为代表的网安融合成为趋势：**为了应对这些转变，以零信任为基础的SASE（安全访问服务边缘）成为趋势，带来分支网络和安全融合的新服务演进。SASE通过从单一云交付平台提供多种融合网络和安全即服务功能，如零信任网络访问 、云访问安全代理 、安全 Web 网关 、防火墙和SD-WAN等，可以提供通过任何网络、任何位置或者设备为任何应用提供安全且无缝的连接；根据Gartner预测，2026年将有80%企业会采用SASE方案进行架构组网改造，市场空间达到210亿美元；

企业上云以及混合办公对于网络安全架构的影响

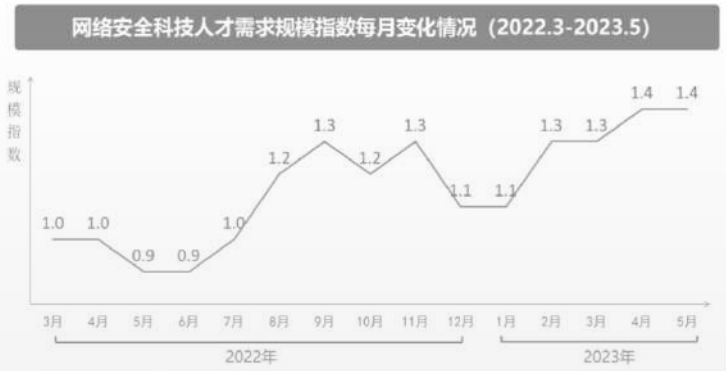
SASE 最终用户支出预测（百万美金）



网络安全建设成本高、专业人才匮乏，网络安全迈向云服务是必然趋势

- **网络安全建设成本高：**当前，大多数企业安全建设仍是在发生信息安全事件后才进行信息安全的资源投入，属于“事件驱动型”和“项目驱动型”，遇到相关问题采购相应设备，由于缺少体系化的考虑和设计，最终导致安全设备极度依赖设备堆叠部署解决某个具体安全事件问题，重复建设严重，设备成本过大；
- **网络安全专业能力匮乏：**网络攻击的日趋复杂对企业的IT人员提出了很高的技术要求，既要了解攻击手法、又要精通防御手段，以及安全数据分析和高执行力。对于一些中小企业或单位，专业安全人才的人力成本过高，这就造成了大部分中小企业或单位安全管理缺失、面对安全事件束手无策的局面；此外安全威胁事件的发生是没有时间规律的，即使部署了安全设备，由于缺乏专业的安全管理流程和安全事件预警机制，运维人员无法快速发现攻击威胁，并及时响应；
- **网络安全向云服务迈进：**对大多数企业来说，受制于时间、资金、人才、流程等方面的缺失，要建立一个全天候的安全服务团队是不现实的，并非所有企业都能自建现代化安全运营中心，通过持续运营获得网络动态防御能力，安全云服务是一种网络安全创新模式，通过云端为企业提供持续进化的安全防护能力和一站式服务，具有实施技术难度低、服务成本低、不需要专业网络安全人员等优势，补足中小企业的安全短板；同时通过云端服务联动和持续进化，可以实现一处检出、全局免疫，网络安全防护能力也有很大优势。根据调查，85%小企业愿意采用云管理、云服务的方式，为数字化转型构筑ICT特别是安全基础设施；

2023年中国网络安全人才需求增长40%



《2023网络安全人才市场状况研究报告》

从传统驻场服务到基于云的网络安全服务架构

安全产品 + 安全驻场

现场L1级别人员技能有欠缺，
联动后端L2人员效率低、纯手工实现服务交付

- 5*8安全运维，部分日志分析
- 主动发现部分安全问题
- 7*24被动应急，基于远程指导

安全产品 + 云端服务

专业安全专家与服务人员，技能储备强、
实操经验足，借助云端平台实现高效服务交付

- 7*24安全运维，全量日志分析
- 主动发现基于流量全量安全问题
- 7*24主动应急，威胁情报触发

勒索攻击成为常态，建设纵深防御体系以筑牢堤坝变得尤为重要

- **勒索软件攻击成为常态：**近年来，勒索攻击事件层出不穷，已经对全球政府、金融、教育、医疗、制造、能源等关键领域造成严重影响，在某些事件中，攻击者挟持关键基础设施索要高额赎金，甚至可能影响国家的正常运行。迄今为止，勒索软件平均导致的业务中断达到16天，每11秒就有一个组织遭受勒索攻击，最大一笔勒索赎金高达7000万美元。大公司企业拥有庞大而复杂的数字基础设施，已成为勒索软件网络犯罪分子的主要目标之一。IDC报告显示，全球35%的组织经历了3-4起勒索软件事件，一次成功的勒索攻击，平均要求缴纳赎金约15万美元，平均造成5天业务中断；
- **防护体系趋向纵深防御：**勒索攻击方法和勒索变种类型不断演进，传统的数据备份、网络边界防护设备和依靠特征检测的传统杀毒软件已经基本失效。并且勒索病毒变种数量呈指数上升，从2021年H2的5400种增长到2022年H1的10666种，增长了98%；勒索软件的加密速度和窃取权限的速度非常快，留给管理员处置的时间窗口期非常短，勒索最快渗透系统获取权限时间是45分钟，而平均加密10万个文件的加密速度仅为43分钟，另外最新一代勒索软件攻击的目标是备份系统、设备和虚拟机，经常导致被攻击后超过46%交付赎金的组织，最终也无法完全恢复数据。新业务变化和新型威胁频发让安全防护变得越来越专业，越来越复杂，同时也需要更智能的安全防护手段，安全防护产品叠加向基础设施网络可信演变，纵深防御体系建设成为企业投资热点；

勒索软件威胁持续升级：病毒变种快，业务中断久、受攻击频繁

57倍

2021年因勒索软件损失
200亿美金
是2015年的57倍
预计2031年2650亿美金

Source from: Cybersecurity Ventures

16天

勒索软件攻击平均导致
16个工作日的系统停机

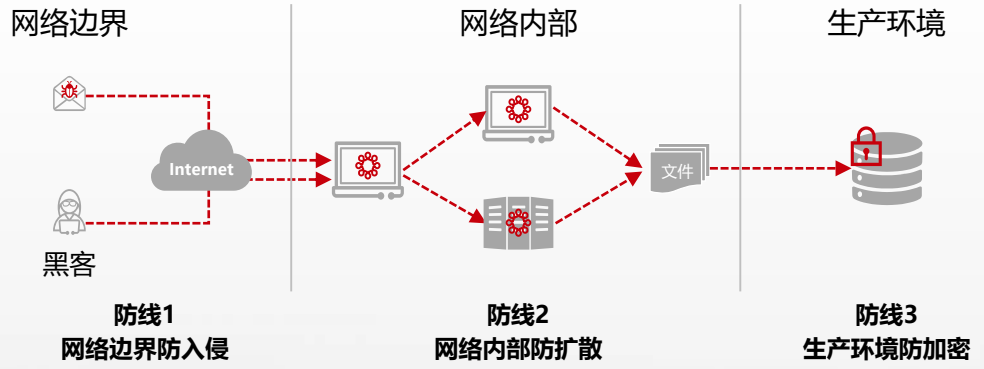
Source from: ZDNet

11s

2021年，每11秒就有
一个组织遭受勒索软件攻击，
预计到2031年，每2秒就有
新攻击

Source from: Cybersecurity Ventures

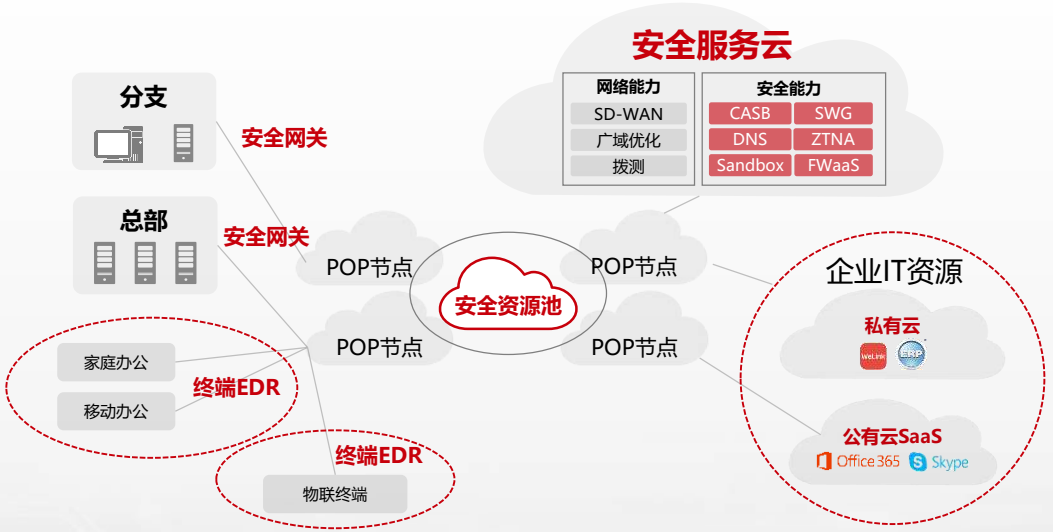
从勒索攻击看企业纵深防御体系构筑



从网安分离到网安一体，融合架构全面提升企业整体安全态势

- **网络安全融合架构体系：**对于企业而言，网络的主要需求包括分支上网、分支和总部互联、分支访问SaaS服务等场景。企业纷纷迁移到云，越来越多的员工采用移动办公的今天，导致大量用户、设备、应用程序和数据位于企业数据中心和企业网络之外。网络和安全一体化融合架构，把网络能力和安全防护能力部署在对应网络节点，通过软件定义，实现灵活，分布式的Overlay逻辑网络，将安全防护能力应用到实体就近的位置，通过运营大脑的协同，提供统一策略、统一安全态势感知，以满足企业各种场景下得网络安全互联需求；
- **网安融合技术优势：**相较于传统网络安全架构，以SASE为代表的网安融合具有零信任访问、云原生架构、支持所有边缘、全球分布四大特点，更能适应企业对于云上应用服务与云化网络安全产品需求的增长，这四个特点也是网络、安全融合部署的体现。SASE可提供整体网络和安全服务，建立分支上网、分支访问总部、分支上云的overlay灵活组网能力，建立基于“overlay连接”的端到端管理机制和安全防护措施，摆脱物理网络的限制，复杂网络简单化，以身份为中心的SASE提供泛在防御能力。通过集中运营服务简化策略管理、安全事件处置，为客户提供简单、高效、安全、稳定的网络接入和业务部署体验；

基于SASE的网安融合架构以及用户价值



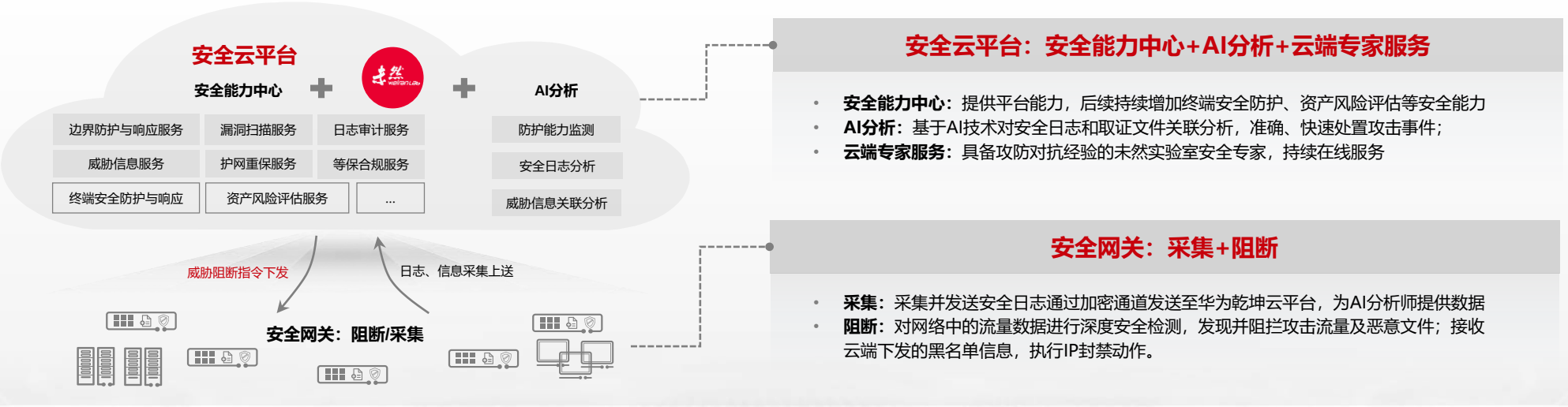
- **降低复杂性和成本**：单一服务提供商，减少分支结构边界的物理或虚拟设备及减少代理数量；
- **性能/延迟的改善**：SASE供应商提供遍布全球的POP点，能够优化接入时延和优化选路；
- **运营费用低**：企业不在在受硬件扩容和EOL的更新设备烦扰，同时可针对新威胁快速提供防护能力，而无需关注特征库升级等；
- **支持零信任**：使用多种威胁信号和上下文信号来确保对内部资源和互联网的安全访问；
- **提升网络和网络安全人员的效率：**单一平台来构建企业安全战略；

从本地防御到云端协同，安全专业能力持续升级

基于云服务的网络安全体系：该体系架构由基于安全云服务平台以及在客户本地网络的安全网关防护节点构成，实现云端服务+本地设备联动，构建简单、高效、易用的安全云服务方案。云边端分工明确，安全网关发挥本地实时防护优势，仅将安全日志信息及攻击取证数据上云，安全云平台发挥算力及威胁信息优势，关联分析，全面检测；

- **全天候动态变化防护：**通过在互联网出口位置部署安全检测设备，本地实时流量检测，云端威胁情报实时更新，专家模型结合AI算法实现海量日志的智能聚合分析，用动态变化的安全防护能力，应对动态变化的安全威胁，用全自动化的威胁分析处置能力，威胁秒级判定；
- **全自动威胁实时阻断：**云端分析发现外部攻击源后，自动下发安全策略，联动本地硬件盒子，分钟级封禁外部攻击，用全自动化的威胁分析处置能力，应对专业复杂的安全分析和人工处置，企业基本可以零人力成本投入分析和使用安全能力；
- **安全服务可按需订阅：**云端安全服务种类多样，订阅更多的服务持续演进，云端漏洞扫描和日志审计服务等新的安全服务能力不断更新，可以持续增强安全防护能力，企业可以完全按照实际需求订阅使用，避免初期大量无效的网络安全投资；

华为乾坤网络安全云服务架构



安全云平台：安全能力中心+AI分析+云端专家服务

- **安全能力中心：**提供平台能力，后续持续增加终端安全防护、资产风险评估等安全能力
- **AI分析：**基于AI技术对安全日志和取证文件关联分析，准确、快速处置攻击事件；
- **云端专家服务：**具备攻防对抗经验的未然实验室安全专家，持续在线服务

安全网关：采集+阻断

- **采集：**采集并发送安全日志通过加密通道发送至华为乾坤云平台，为AI分析师提供数据
- **阻断：**对网络中的流量数据进行深度安全检测，发现并阻拦攻击流量及恶意文件；接收云端下发的黑名单信息，执行IP封禁动作。

从单点防御到多级联动，构建智能化全流程防勒索防护体系

- **全流程防御策略：**基于勒索病毒的攻击流程，勒索软件的防范也需要从传统的防范策略演变到全流程防御策略。首先，事前网络边界防入侵阶段，做好边界的隔离、防攻击、防病毒及恶意文件检测，基本能防住70%的攻击；其次，事中横向扩散阶段，通过边界隔离、自动化的安全分析和处置，可以防住20%的攻击；最后，事后勒索加密阶段，通过存储备份、隔离区建设，剩余10%的攻击也能防护住了。但这三个阶段的防护成本和效果差异很大，网络层的安全防护成本最低，效果最好，做好网络层的防入侵和防扩散，可以防住90%以上的勒索攻击，就好比建一道“防盗门”，让黑客没那么容易进来；
- **网络存储多层联动：**通过存储、网络等基础设施的结合，采用多层次、端到端的有效防护，可提供抵御勒索软件的最佳防御。网络与存储多层检测及联动的数据保护，通过有效的攻击前预防、攻击时的精准检测及响应和攻击后快速恢复，使勒索攻击防护从被动响应向主动防御转变，帮助用户及时发现并拦截勒索攻击，保护数据不被非法加密和窃取，在必要时还可快速安全恢复数据，全方位构建防勒索安全防护体系。网存联动勒索攻击防护可以实现事前、事中、事后全流程覆盖，具有攻击识别准、威胁防护全、数据恢复快三大特征；

边界防御为主，内部防护辅助

企业纵深防御体系建设

行动建议：网络安全成为核心要素，加快安全防护理念和技术演进

将网络安全置于业务核心位置

“发生事件后再考虑网络安全”的被动防御已经过时，要想让数字化转型稳定开展，就需要将网络安全建设被视为数字化业务发展的伙伴，并得到公司的战略支持。这也意味着，制定网络安全计划不仅仅是为了防止各种网络安全攻击事件发生，更是为了提高企业有效承担数字化发展风险的能力；

采用基于云服务的网络安全模式，快速提升安全能力

网络安全牵一发而动全身。对于非网络安全或IT专业人员来管理公司网络空间，或者安全投资预算有限的公司，建议选择安全云服务的方式，为企业数字化转型保驾护航；

拥抱网络和安全一体化创新，统一网络和安全管理策略

由于连接性增加、SaaS和云应用程序得到广泛使用，企业组织的安全攻击面持续变大，公司需要更广泛的可见性和统一策略来持续监控威胁和风险暴露情况。企业组织需要构筑一体化的防护系统，体系化的开展并管理对威胁的检测、调查和响应工作，让安全运营团队全面了解风险和潜在影响。

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。

Bring digital to every person, home and organization for a fully connected, intelligent world.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.

HUAWEI