面向大规模智能监控系统的算力共享网络解决方案

作者 吉莉 熊重驰



# 摘要

面向大规模智能监控系统的算力共享网络是面向大规模智能监控系统发展的主旋律，基于边缘计算的算力共享网络在未来大规模视频监控预警市场占绝对优势。随着“智能管道 智慧管网”的迅速推进，将安防监控与运营商的网络、边缘计算、云技术相结合，为视屏监控云网融合、云边协同带来可能。

智能监控系统作为实现“智能管道 智慧管网”系统的重要组成部分受到了社会各界的重视。为提高监控系统效率，降低风险漏报率，面向大规模智能监控系统的算力共享网络已经成为迫切需求。本文聚焦于大规模智能监控系统所面临的设备出现异常导致的计算能力不足、计算核心故障导致的风险漏报等问题。

针对设备出现异常（电力不足、过热降频以及计算任务激增）导致的计算能力不足，提出一种算力感知网络的架构，围绕着周围算力盈余的节点分配计算资源，克服传统部署方式在设备出现异常导致的计算核心无法发挥全部性能而产生的计算能力不足的问题。通过服务器向周围区域计算性能正常的节点发送区域内因设备异常导致无法正常工作的节点的视频数据，实现区域计算节点任务的统一调度和管理，实现基于区域和视频源的视频计算资源部署和分配模式。受影响的计算节点只需将摄像机记录的信息片段发送到服务器，有效解决了因设备计算能力不足导致的风险漏报问题。

针对计算核心故障（死锁、过热死机以及硬件故障）导致的风险漏报问题，提出两种重启电源恢复设备的策略。

1. 由于计算核心出现的死锁、过热死机等故障具有可自恢复特点，提出一种使用GSM手动控制的断电开关器实现设备断电重启的方法。当设备计算核心故障无法响应时，GSM手动控制的断电开关器会自动识别计算核心的故障类型（做不到）并选择对应的解决方案。
2. 除了使用GSM手动控制的断电开关器实现设备断电重启的方法，我们还提出了给电源开关加装蓝牙模块，由蓝牙模块来控制断电开关器的方法。该策略同样也能实现第一种方案中所实现的根据故障类型自动选择解决方案的需求。

以上两种方案都能在保证整个监控系统稳定的同时最大减少风险漏报的问题。本文基于以上提出的计算架构与方法，探讨面向大规模智能监控系统的算力共享网络的可行性、有效性和实用性。

# 课题研究背景及意义

长输油气管道安防主要依靠人力巡检巡查，西气东输每年投入的人力巡护成本较高，为了响应国家油气体制改革要求，西气东输公司率先开展了“智能管道 智慧管网”的探索与建设，并在管道智能监控感知威胁态势方面取得了实质性的突破。视频监控技术作为实现“智能管道 智慧管网”的重要技术手段也越来越受到社会各界的重视，同时在各个行业得到了广泛的应用。2010年，为了顺利、安全地举办世博会，上海市大规模增加监控网络的范围和密度，仅在浦东新区就建设了包含一万多个监控点位的高清监控系统；截至2011年，黑龙江投入8亿多建设资金用于监控系统建设，共建设近千个监控中心，二十多万个监控点；武汉市为推进“平安城市”建设，建设了覆盖跨江大桥、主干道和主城区城道口等重点部位的监控系统。

长输油气管道安防主要依靠人力巡检巡查，西气东输每年投入的人力巡护成本较高，为了响应国家油气体制改革要求，西气东输公司率先开展了“智能管道 智慧管网”的探索与建设，并在管道智能监控感知威胁态势方面取得了实质性的突破。视频监控技术作为实现“智能管道 智慧管网”的重要技术手段也越来越受到社会各界的重视，同时在各个行业得到了广泛的应用。2010年，为了顺利、安全地举办世博会，上海市大规模增加监控网络的范围和密度，仅在浦东新区就建设了包含一万多个监控点位的高清监控系统；截至2011年，黑龙江投入8亿多建设资金用于监控系统建设，共建设近千个监控中心，二十多万个监控点；武汉市为推进“平安城市”建设，建设了覆盖跨江大桥、主干道和主城区城道口等重点部位的监控系统。

视频监控系统的发展主要分为以下三个阶段：

第一阶段：以模拟摄像机及模拟录像机为主的模拟视频监控系统，这种监控系统多是以摄像机、录像机、分割器为核心，通过模拟信号输出视频，传输距离短，图像质量差，目前已逐步退出市场。

第二阶段：基于模拟摄像机和数字录像机的视频监控系统，采用模拟信号传输视频，数字方式处理与贮存，属于过渡阶段。

第三阶段：基于网络摄像机的全数字视频监控系统，以数字方式传输视频，传输距离长，图像质量好，目前正得到广泛应用。

以上三代视频监控系统，主要解决视频捕捉、传输、存储等问题，并不具备视频职能分析的功能，真正起到安全防范的作用则需要人为观察视频从而做出安全评估。一种方法是实时观看，建设一个监控中心，将监控点所有的视频流导入监控中心，同时配备多块显示屏来播放视频画面，由工作人员通过观看屏幕上的监控点的视频，事后进行调阅，本质上还是以人观看的方式来查找感兴趣的信息。

然而人工监控具有本身固有的缺陷，例如：消耗大量人力、易使人产生疲劳、可扩展差等等。尤其随着“智能管道 智慧管网”监控系统规模不断扩大，面对数以万计的监控点和海量视频数据，以人工为核心的监控系统越来越无法满足需求。如何解决人工监测带来的种种问题，加强监控系统的事前预警和事后的检索，成为当前摆在“智能管道 智慧管网”发展前的重要问题。

监控界提出了智能视频监控(Intelligent Video Surveillance, IVS)的概念。智能监控即使用计算机强大的计算能力以及图像分析技术对传统的监控视频进行处理、分析和理解，过滤无用信息。目前智能视频监控在计算机视觉领域中备受关注。

相比于传统监控系统，智能监控系统具有以下优势：

1. 全天候的监控能力：使用计算机代替人来监控视频画面，实现对监控画面的不间断分析，彻底改变以往由监控人员对监控画面进行监视和分析的方式。
2. 快速的响应能力：通过设置某些可疑行为的识别规则，在其发生危害之前通知相关人员采取相应措施，为潜在威胁做好准备工作。
3. 扩展视频资源的应用领域：传统视频监控系统只能获取单一的视频数据，而智能监控系统可以提取视频中的各种信息，为其他应用提供支持。

随着“智能管道 智慧管网”监控系统规模的不断扩大，面向大规模监控系统的视频计算对智能监控系统提出了新的挑战：

1. 监控设备出现电力不足、过热降频以及计算任务激增等状况所导致的算力不足问题。视频监控节点通常部署在野外，由太阳能供电，而监控视频流数据量大，计算核心需要在稳定的工作电压下才能高效运算，如果出现电量不足、散热不良导致过热降频或者短时间内计算任务激增等情况，计算核心将无法发挥全部性能以负荷繁重的计算任务，这将导致设备反应缓慢，遇到风险时无法及时汇报。如何解决监控设备出现故障导致算力不足的问题，是智能监控系统需要解决的首要问题；
2. 计算核心故障导致的风险漏报问题。视频监控摄像头每分每秒都在向计算核心输出大量的视频流数据，计算核心任务十分繁重。当计算核心出现故障（如死锁、过热和硬件故障）时，该区域的风险监控将瘫痪，此时若存在风险，势必会出现风险漏报问题，后果不堪设想。如何保证智能监控系统的稳定性，杜绝因计算核心故障而导致的风险漏报也是亟待解决的问题；

本文针对以上两个问题进行研究，有利于实现智能监控，推进“智能管道 智慧管网”建设， 对公共安全、社会稳定等具有重要意义。

# 国内外发展研究现状

参考文献

第一个版本

基于边缘计算的智能城市实时监控视频分析目前引起了从业者和研究人员的广泛关注，大量的视频数据被提交到互联网。根据原始视频内容，如何删除冗余视频帧、正确分割视频序列并减少不必要的计算资源消耗是一项具有挑战性的工作。X. Zhang等人提出了一种检测和定位具有运动场形状的视频异常的方法，S. Ding等人利用改进的 Harris-Laplace 时空兴趣点从大视频中识别有趣的片段，然后从中选择关键帧。时空兴趣点可以准确地附着在检测到的目标周围，兴趣区域构建算法可以快速准确地定位包含目标的候选区域。由于图像分类和目标识别的巨大成功，特别是视频监控、目标计数和目标检测，深度学习受到了广泛关注。相比之下，Y. Fang等人提出了空间和时间特征图的计算，然后检测视频显著性。由于资源限制（能源消耗、计算和内存），在智能终端设备上部署这些具有低延迟要求的大型、强大的视频任务仍然是一个挑战。因此，考虑将这些计算任务转移到更强大的边缘服务器或云。然而，云计算模型不适用于需要短期执行的边缘服务，因为将任务卸载到云中心增加了网络往返传输延迟，应用服务请求不会很快得到响应，但是充分利用云中强大的计算和内存资源将减少总响应时间。由于边缘节点和边缘服务器离用户近，可以快速响应用户请求，成为首选的帮手。在边缘服务器上运行计算密集型任务时，需要对多个终端设备资源进行有效管理。目标是平衡精度、能耗、延迟和负载平衡的性能参数。 VideoStorm引入了这些权衡，以在满足准确性和延迟目标的前提下为每个请求选择正确的配置。例如，在 Chameleon中，配置在流视频输入期间在线更新。然而，基于任务划分的不同模型分割点会导致不同的计算延迟。因此，有必要选择合理的切割策略，以最大限度地发挥端边甚至云协同的优势。在将任务卸载到边缘服务器时，我们可以在边缘进行数据预处理，减少冗余、带宽、延迟和对云中心的依赖，同时提高视频分析的效率。为了减少带宽消耗，有学者提出了端边云协同架构和模型压缩，以消除不同环境下的数据传输。例如M. Song等人提出只有从边缘设备推断出的数据传输到云端进行再训练，以减少数据传输，建议在不影响精度的情况下去除冗余数据以减少数据传输。Glimpse 将所有 DL 计算任务迁移到最近的边缘服务器，同时使用更新检测来删除应该卸载的相机帧。如果检测没有变化，将在本地进行帧跟踪。这种过滤增强了系统的处理能力并使移动设备上的实时目标检测成为可能。Vigil提出了一种分布式架构，它巧妙地利用边缘和云之间的处理任务来减少视频监控中的带宽消耗。类似地，VideoEdge提出了边缘和云的分层架构来处理摄像机流，以便在多个目标和约束之间实现更好的权衡，公平地分配资源。我们在智能视频监测系统中提出采用算力感知网络（CAN，Computing-aware Networking），该网络通过对多个边缘节点算力资源的合理分配和调度， 综合考虑实时的网络资源状况和计算资源状况，通过网络灵活匹配、动态调度，将动态调度能力调至最优节点，让网络支持提供动态的服务来保证边缘计算节点正常工作。

第二个版本

作为解决多级算力资源（云计算、边缘计算以及端计算）并存情况下资源统一供给问题的一种新型网络技术方案，算力网络通过网络控制面（如集中式控制器、分布式路由协议等）分发服务节点的算力、存储、算法等资源信息，并结合网络信息和用户需求，提供计算、存储、网络等资源的分发、关联、交易与调配，从而实现整网资源的最优化配置和使用 [1]。

算力网络的产生与边缘计算息息相关，它可以重点解决资源节点泛在化后的两个重要问题：用户体验一致性和服务灵活动态部署[2]。首先，算力网络可以解决用户体验一致性的问题：用户无须关心各类基础资源（算力、存储等）的位置和部署状态，通过网络即可协同调度各类资源，保证用户的一致体验；其次，算力网络可以解决服务灵活动态部署的问题：基于用户的服务等级协议（SLA）需求，综合考虑实时的网络、算力、存储等多维资源状况，通过网络灵活匹配与动态调度，将业务流量动态调度至最优资源节点。

从 2019 年初至今，业界对算力网络的研究仅有两年的时间。算力网络巨大的潜在需求却掀起了业界的波澜。目前，三大运营商、各厂商以及学术机构纷纷开始研究算力网络。

2020 年 6 月，CCSA TC614 成立了算力网络特别工作组，依托联盟的平台和资源，联合多方力量，共推、共创算力网络产业影响力，构建算力网络生态圈。2020 年 11 月， 中国联通成立了中国联通算力网络产业技术联盟，将在“联接 + 计算”领域和全产业链合作伙伴携手并进，共建算力网络生态，推动商业落地，共享转型成果。

中国主流运营商还先后发布了《中国联通算力网络白皮书》《算力感知网络技术白皮书》《算力网络架构与技术体系白皮书》等。

在各方的不懈努力下，算力网络的标准化工作取得了进展：在 ITU-T、互联网工程任务组（IETF）、宽带论坛（BBF）、ETSI、CCSA 等全球标准组织中，已立项相关的国际标准 9 项、中国标准 4 项。

* 在 ITU-T SG13 组，中国电信、中国移动、中国联通、华为等单位分别从算力网络架构、算力感知网络相关技术等方面推进了 Y.CPN-arch 标准、Y.CAN 系列标准的制定。
* ITU-T SG11 组启动了 Q.CPN 标准（算力网络的信令需求）与 Q.BNG-INC 标准（算力网络边界网关的信令要求）的制定等工作。
* 在 IETF，华为等撰写了 Computing First Network 系列文稿，研究算力路由协议；
* BBF 启动 了“Metro Computing Network（SD-466）”，专门研究算力网络在城域网中的应用。
* ETSI 提出 了“NFV support for network function connectivity extensions（NFV-EVE020）”。该方案以内容转发网络（CFN）为基础，研究 NFV的计算和网络集成相结合的网络功能连接扩展方案。
* CCSA TC3 目前已经完成《算力网络需求与架构》的研究报告和面向全网的算力感知网络关键技术研究。2021年4月 TC3 全会形成了算力网络系列行业标准的立项，包括算力网络总体技术要求、算力网络标识解析技术要求、算力网络路由协议要求、算力网络控制器技术要求、算力网络交易平台技术要求和算力网络开放能力研究等工作。

第三个版本

算力网络伴随着计算产业与网络产业的结合而兴起，一年多来，在标准制定、生态建设以及产业活动方面均取得了一系列成果。

标准制定方面，国内三大运营商中国移动、中国电信与中国联通分别在ITU-TSG11 与SG13组立项了Y.CPN，Y.CAN和Q.CPN等系列标准，在IETF开展了Computing First Network Framework 等系列研究，华为联合国内运营商在ETSI和宽带论坛（BBF）也启动了包括NWI、城域算网在内的多个项目。在中国通信标准化协会(CCSA)，“算力网络需求与架构”、“算力感知网络关键技术研究”、“云边协同的承载网控制系统架构及技术要求”等研究工作也在有序开展。

生态建设方面，国内未来数据网络研究的重要组织网络5.0产业联盟专门成立了“算力网络特设工作组"，MEC 领域的多个开源组织也发起了KubeEdge、Edge-Gallery 等开源项目，2019年底，中国联通、中国移动和边缘计算网络产业联盟（ECNI）均发布了算力网络领域相关白皮书，进一步阐述了算网融合的重要观点。

产业活动方面，2019年中国电信与中国移动均已完成算力网络领域的实验室原型验证，并在GSMA巴塞罗那展、ITU-T和GNTC相关展会上发布成果。中国联通研发的算力网络服务平台，也已在积极推进试点工作。

展望未来，算力网络已经成为国内IMT-2030推进组 6G网络研究的重要课题之一,不论是从IPv6网络技术、云原生技术和可编程芯片技术的发展来看，还是从新基建政策的导向来看，都给予了构建在算力网络之上的新商业模式尝试的空间。未来两年,是算力网络在业界由概念普及转向试点验证的重要节点，中国联通算力网络研究团队经过和业界合作伙伴的深入探讨，特在此白皮书后续章节中明确中国联通算力网络的架构和关键技术，希望给业界更多的启示并作为后续研究持续推进的基础。

# 课题需求分析

在大规模智能视频监控系统的背景下，虽然现存网络利用边缘计算技术，缓解了云中心计算延迟高、带宽不足、安全性低的问题，实现了“就地、就近”提供服务的功能，但是当智能边缘计算设备存在不稳定情况（如：产生死机或掉线等），依然会对整体网络造成巨大影响。如果出现以上情况，很有可能会造成该终端节点所连接摄像机中视频数据的丢失，那么对于该监控区域来说，必定存在一定风险，因为智能边缘计算设备无法进行实时的检测和报警工作。

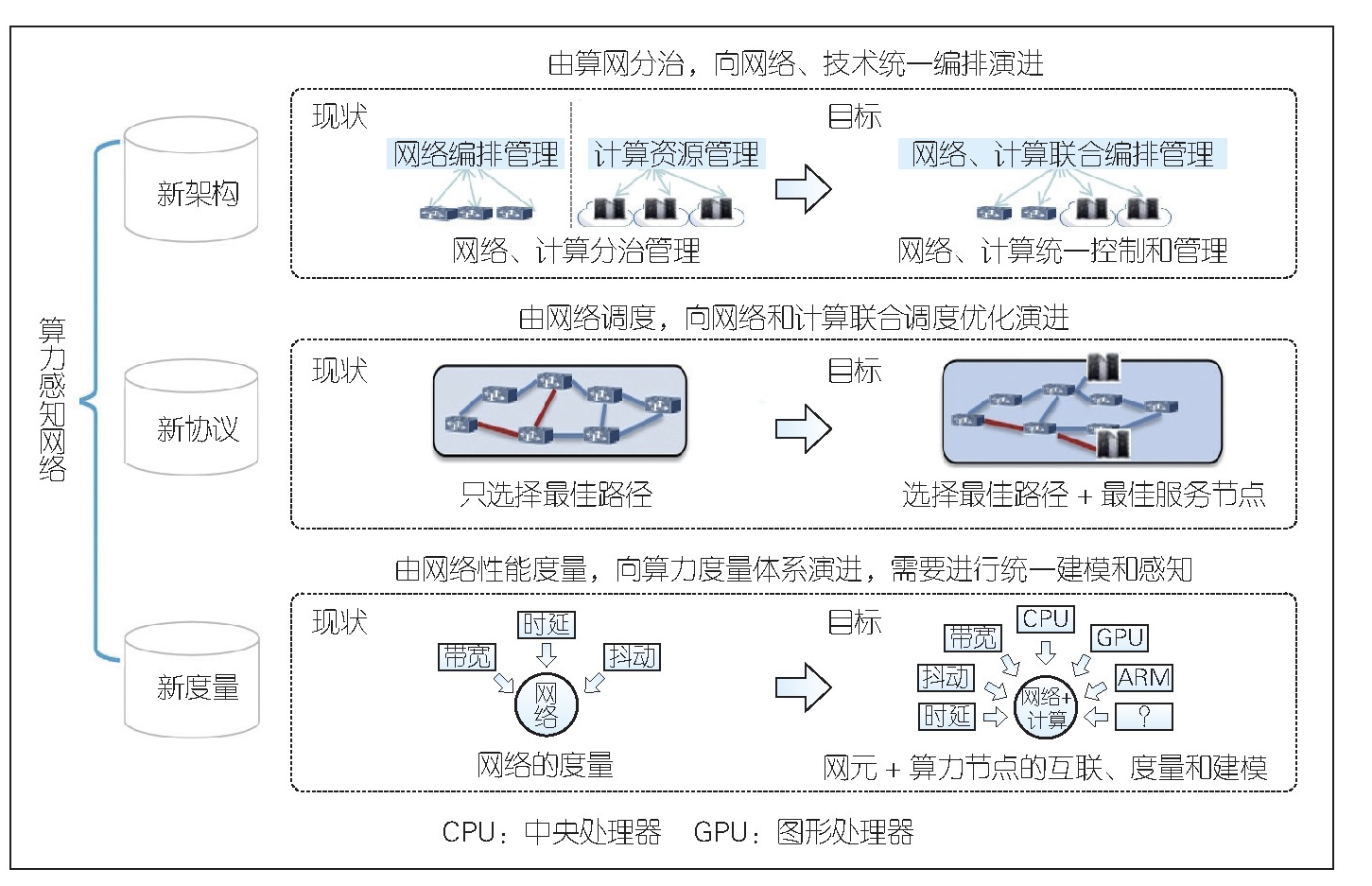
除此以外，考虑到一些智能边缘计算设备因为部署在野外，采用太阳能板进行供电，其面临各种由环境所带来的问题。如连续阴雨天气导致无法为太阳能板持续提供能量、夏季长时间阳光暴晒导致计算核心过热降频以及遇上风沙或大雾天气导致计算任务激增等问题，那么边缘计算节点将面临着计算能力不足的问题。在这种情况下，无法对所收集到的视频数据提供正常的处理，往往会出现对异常情况的漏报。针对以上智能边缘计算设备的不稳定性和计算能力不足问题，我们提出以下需求分析：

1. 规避计算能力不足及漏报风险：当边缘计算节点的电压没有达到计算核心的工作电压或者计算核心因为天气或者计算任务激增导致过热降频时，将导致计算能力下降，可以借助算力共享方法，使得当前节点只需执行少量本地任务，利用丢帧策略将剩余视频数据任务传输给距离当前节点最近的边缘设备帮助计算，要求这些边缘设备是连接电源的稳定节点。除此之外，当边缘计算节点的能量几乎耗尽，无法支持自我计算的情况下，可能会因为设备宕机而产生漏报风险，那么此时，边缘计算节点只需要维持传输功能，将所有的视频数据信息传输给最近的稳定节点进行计算，以避免漏报风险。
2. 减少后勤工作人员的修理成本：从物理安全方面来说，因为边缘计算节点往往部署在室外环境，缺少一定的保护措施。当物理设备收到由外部环境所带来的无可预知的干扰或破坏时，通常会导致设备死机，故障等问题。通常在这种情况下，后勤工作人员会来到现场进行维修，维修的耗时以及人工的耗费不可避免的增加了维护成本。但由于部分问题（例如：死锁，过热等）可以通过重启电源来解决，那么此时我们就需要一个自动识别故障类型并自动选择相应的处理方案，用以最快恢复故障节点和最大节省维修成本。

目标是解决设备死机问题（虽然没有办法从根源上解决，但是能在出现死机以后有应对策略），而不是减少人员成本

# 解决方案

针对计算能力不足问题，我们提出采用算力感知网（CAN）进行解决。为了实现泛在计算和服务的感知、互联和协同调度，CAN从三个角度改变了现存边缘计算能力不足问题：（1）针对当前将网络和计算资源分别管理，CAN能够将网络和计算能力相融合，进行统一控制与管理；（2）针对当前网络设备通过路由表信息选择最佳路径，CAN同时考虑网络设备路由信息及边缘节点的计算资源，旨在选择最佳路径与最佳服务节点；（3）针对当前利用带宽、时延和抖动等指标度量网络性能，CAN还会考虑CPU、GPU和ARM处理器等终端设备性能。（详见下图）



要不要补充CAN的五个层次（背景知识）

利用CAN，我们将无处不在的网络连接为基础，通过服务的自动化部署、最优路由和负载均衡，构建算力感知的全新的网络基础设施，最终实现网络的无所不达及算力无处不在。CAN的实施主要分为三个模块：

（1）算力度量和建模

对于异构计算类型的资源，将采用基于算力统一的度量体系，通过算力资源进行统一抽象描述，形成算力能力模板，对算力路由、算力设备管理、算力计费等提供标准的算力度量规则。首先，异构硬件设备通过统一的算力度量和建模，实现对现场CPU和GPU等异构物理资源的统一资源描述，从而可以有效地提供计算服务。其次，考虑到计算过程受算法的影响，需要对部署在边缘计算节点上的时间序列模型所需的算力进行度量，更有效地了解应用调用算法所需的算力，从而服务于应用。最后，由于边缘计算节点的算力需求有大有小，因此需要把设备的需求映射为实际所需的算力资源，从而可以使网络更充分有效地感知终端设备的需求，提高交互的效率。

（2）算力路由

利用算力路由，通过对网络、计算、存储等多维资源、服务的感知与通告，从而实现“网络+计算”的联合调度。算力路由层包括算力路由控制技术和算力路由转发技术，这两种技术可以实现边缘节点请求的按需调度。

算力路由控制通告算力节点的信息并生成算力拓扑，生成算力感知的新型路由表来表示边缘计算节点的剩余计算能力。算力路由控制基于计算能力不足的节点需求，生成动态、按需的算力调度策略，实现算力感知的算网协同调度，帮助该节点完成相关任务。算力路由转发通过路由协议实现对网络环境进行感知，通过网络路径和算力的结合，推选最优的节点提供服务。

（3）算力管理

对于边缘计算节点，需要进行统一的管理，以实现对网络和算力的监测和控制。通过对不同计算类型进行统一抽象描述，算网管理层能够形成算力能力模板，从而为算力设备的操作维护管理（OAM，Operation Administration and Maintenance）等提供标准的算力度量规则；算力注册需实现对算力节点的注册、更新和注销，并对相应的路由通告策略进行管理；算力OAM需实现对算力资源层的算力性能监测控制、算力计费管理等。

针对计算设备的不稳定性，我们提出一种使用GSM手动控制的断电开关器实现设备断电重启的方法，该方法可以在设备出现故障的时候，采用远程控制方式进行设备开关机，而管理人员不必到达故障节点所在现场。Xxx（如何检测死机？严所只说了死机需要去现场，那他们又是怎么知道的呢），当某边缘计算节点频繁出现故障并且需要进行重启时，我们需要将每次重启的设备节点以及故障时间进行记录。如果在一个月的时间内，被记录的次数过多，那么说明该节点可能受到硬件性能或其他因素的影响，此时需要管理人员到边缘计算节点所在现场进行调研，从根源上解决该节点的问题，以减少重启的次数，帮助保护设备硬件并且让其更加稳定地进行计算。

* **如果节点死机**，那么远程控制系统中该节点的颜色会标记为灰色（暂且当作是这样）。对于这样的节点，因为节点是有电的，所以GSM手动控制开关器不会电话提醒，所以我们可以利用GSM去进行设备关机，再启动开机。
* 但是有个问题，这里远程怎样主动去提醒用户，利用前端自动提醒用户功能？
* **如果节点没电（可能会牵连到GSM）**，那么GSM会打电话给用户进行提醒。用户可以等到太阳能电板充电用GSM开关或者到现场处理。因为没电的话，GSM会打电话提醒用户（发短信），所以我们可以把断电的次数做记录，如果短期内多次接收到断电提醒，那么代表该节点的供电系统是存在巨大隐患的。为了解决该节点的供电问题，我们可以为其多增加几块太阳能板或者更换该节点的电池。

当设备计算核心故障无法响应时（做不到），GSM手动控制的断电开关器会以推送的方式来通知管理人员，管理人员可以依照需要选择是否重启设备电源，远程恢复设备的正常运行。当设备在重启电源多次后都没有恢复正常运行，GSM手动控制的断电开关器会判断设备出现了硬件故障，即设备进入了不可恢复式死机的状态，

这时GSM手动控制的断电开关器会给后勤维护人员发送维修推送，（做不到）由负责后勤维护的工作人员来接管恢复工作。

除了使用GSM手动控制的断电开关器实现设备断电重启的方法，我们还提出了给电源开关加装蓝牙模块（可以不必，问了电话卡不用月租，或者蓝牙比较便宜就不用GSM了），由蓝牙模块来控制断电开关器的方法。该策略同样也能实现GSM方案中所实现的重启设备电源与发送维修推送的需求。

# 参考文献

[1]雷波,陈运清.边缘计算与算力网络——5G+AI 时代的新型算力平台与网络连接[M].北京:电子工业出版社,2020

[2]中国移动.算力感知网络技术白皮书[R].2019