合同编号：（2026）MZ-058

福田区智能算力服务合同

项目名称：福田区智能算力服务采购项目

甲　　方：深圳市福田区民意速办和智慧城市建设中心

乙　　方：中国联合网络通信有限公司深圳市分公司

甲方：深圳市福田区民意速办和智慧城市建设中心

地址：深圳市福田区深南大道1006号国际创新中心F座3楼

联系人：[redacted]

联系电话：[redacted]

乙方：中国联合网络通信有限公司深圳市分公司

地址：深圳市福田区福田街道4005号深南大道中国联通大厦

联系人：[redacted]

联系电话：[redacted]

根据深圳公共资源交易中心（深圳交易集团有限公司福田分公司）FTCG2025000032A号项目结果，中国联合网络通信有限公司深圳市分公司单位为中标人。按照《中华人民共和国民法典》（第三编 合同）和《深圳经济特区政府采购条例》，经深圳市（以下简称甲方）和单位（以下简称乙方）协商，就甲方委托乙方承担 **福田区智能算力服务**，达成以下合同条款：

第一条 项目概况

项目名称：福田区智能算力服务采购项目

项目内容：采购所需政务智能算力服务的总体要求：一是AI视频智能分析算力，能支撑不少于2500路，每路不少于8个算法的并发视频分析；二是大模型政务智能算力，为一系列AI+政务应用提供高性能计算算力支撑。具体要求详见附件一《算力技术要求》。

服务期限：本项目服务期限最长为36个月，合同一年一签，如采购方对履约情况不满意，可以不再续约。服务期内因政策原因需要终止服务的，

招标人可以提前一个月通知供应商解除合同。

服务时间：2026 年 6 月 3 日至 2027 年 6 月 2 日

合同价款：本项目合同支付上限为项目招标预算 11,750,000 元（大写：壹仟壹佰柒拾伍万元整），含一切税、费。本合同项下服务的单价为固定价格，不因通货膨胀调整。但若服务期内因市场行情、服务内容变更等因素导致费用标准发生变化，双方可协商调整单价。本合同费用涵盖乙方为履行本合同所需的全部成本，包括但不限于：

（1）实施本项目算力服务相关的设备仪器购置、运维费用；

（2）设备运输、车辆租赁及本地化部署至甲方指定机房的费用；

（3）裸光纤部署服务（含主备）：从甲方指定机房至福田区委办公大楼指定机房，采用直连方式保障高带宽、低延迟的稳定传输，并符合行业标准；

（4）技术服务：包括月度设备使用率报告编制、协助甲方完成技术整改及其他技术支持；

（5）以各种方式寄送技术资料到甲方办公室所发生的费用；

（6）乙方履行本合同义务所发生的一切费用和支出。

支付方式：分期支付（详见合同第十五条）。

第二条　部署要求

1、采用本地化部署方式，智算服务部署到甲方指定机房，作为本地资源池扩展的形式，均直接复用政务云网络安全和密码应用安全防护能力；

2、提供裸光纤部署服务（含主备）：从甲方指定机房到福田区委办公大楼指定机房，采用裸光纤直连的方式实现高带宽、低延迟的稳定传输。部署过程严格遵循行业标准，保障链路质量与可靠性；

3、AI 训推一体机配套服务：若采购多台（含相同和不同型号，如 AI

训推一体机1和2）设备，须完成集群部署。须接入福田区政务云管理。提供配套软件平台，提供数据集管理、模型管理及微调、模型部署及量化等训练大模型操作功能，支持工作流搭建、知识库搭建、用户管理等功能；

4、 其他合同未明示的相关工作。

第三条 时间要求及阶段成果

1、乙方须在合同签订后1个月内提供具体服务方案，具备本地化部署智能算力服务，由甲方根据业务使用需求自主调度。

第四条 服务资料归属

1、所有提交给甲方的服务文件及相关的资料的最后文本，包括为履行技术服务范围所编制的图纸、计划和证明资料等，都属于甲方的财产，乙方在提交给甲方之前应将上述资料进行整理归类和编制索引。

2、乙方未经甲方的书面同意，不得将上述资料用于与本服务项目之外的任何项目。

3、合同履行完毕，未经甲方的书面同意，乙方不得保存在履行合同过程中所获得或接触到的任何内部数据资料。

第五条 甲方的义务

1、负责与本服务项目有关的第三方的协调，提供开展服务工作的外部条件。

2、向乙方提供与本项目服务工作有关的资料。

第六条 乙方的义务

1、应按照招标文件、投标文件要求按期完成本项目服务工作。

2、负责组织项目的实施，保证工程进度和工作质量，并满足验收相关标准。

3、向甲方提交本合同要求提供的所有资料等纸质件各一套，电子版文

件各一套。

4、在履行合同期间或合同规定期限内，不得泄露与本合同规定业务活动有关的保密资料。

第七条　甲方的权利

1、有权向乙方询问工作进展情况及相关的内容。

2、有权阐述对具体问题的意见和建议。

3、当甲方认定乙方人员不按合同履行其职责，或与第三人串通给甲方造成经济损失的，甲方有权要求更换人员，直至终止合同并要求乙方承担相应的赔偿责任。

第八条　乙方的权利

1、乙方在本项目服务过程中，如甲方提供的资料不明确时可向甲方提出书面报告。

2、乙方在本项目服务过程中，有权对第三方提出与本服务业务有关的问题进行核对或查问。

3、乙方在本项目服务过程中，有到工程现场勘察的权利。

第九条　甲方的责任

1、应当履行本合同约定的义务，如有违反则应当承担违约责任，赔偿给乙方造成的损失。

2、甲方向乙方提出赔偿要求不能成立时，则应补偿由于该赔偿或其他要求所导致乙方的各种费用的支出。

第十条　乙方的责任

1、乙方的责任期即本合同有效期。如因非乙方的责任造成进度的推迟或延误而超过约定的日期，双方应进一步约定相应延长合同有效期。

2、乙方的责任期内，应当履行本合同中约定的义务，因乙方的单方过

失造成的经济损失，应当向甲方进行赔偿。

3、乙方对甲方或第三方所提出的问题不能及时核对或答复，导致合同不能全部或部分履行，乙方应承担责任。

4、乙方向甲方提出赔偿要求不能成立时，则应补偿由于该赔偿或其他要求所导致甲方的各种费用的支出。

第十一条　人员要求

项目服务团队成员不少于11人提供配套服务，包括项目负责人1人、技术负责人1人、安全工程师1人、资源管理专员1人、售后负责人1人、实施和运维团队不少于6人。

序号	岗位	职责
1	项目负责人（1人）	统筹项目整体进度，协调资源，监督智能算力服务器部署及运维，确保项目按时交付并满足客户需求。负责日常技术管理、人员管理、内外部沟通协调等相关管理工作。
2	技术负责人（1人）	制定智能算力服务器部署的技术方案，指导部署与调试，解决技术难题，确保系统性能及稳定性。
3	安全工程师（1人）	部署安全防护措施，监控系统安全，处理安全事件，保障智能算力服务安全运行。
4	资源管理专员（1人）	负责智能算力服务器资源规划、云平台部署及性能优化，确保云环境稳定运行，支持业务需求。
5	售后负责人（1人）	建立售后服务体系，及时响应客户问题，提供运维支持，确保智能算力服务持续稳定运行。
6	实施和运维团队(不少于6人)	负责项目实施和服务期内的运维工作。具体负责服务器安装、系统部署、日常监控及故障处理等等，服务期内高质量执行运维任务，保障智能算力服务高效运行。

第十二条　乙方服务要求

1、乙方应配备中标项目所需的足够数量的智能算力服务，具备本地化部署智能算力服务，由甲方根据业务使用需求自主调度。具体要求详见附件一《算力技术要求》。

2、乙方应于每月第五个工作日前，按行业通用标准编制上月《设备使用率分析报告》（含使用数据统计图表、效能分析及异常情况说明），以书面形式提交至甲方指定联系人审核备案。

3、乙方应配合甲方提出的合理整改要求，在接到甲方指令后十日内完成设备调试、参数调整或系统升级等技术整改工作。

第十三条　保密要求

1、由甲方收集的、开发的、整理的、复制的、研究的和准备的与本合同项下工作有关的所有资料在提供给乙方时，均被视为保密的，不得泄漏给除甲方或其指定的代表之外的任何人、企业或公司，不管本合同因何种原因终止，本条款一直约束乙方。

2、乙方在履行合同过程中所获得或接触到的任何内部数据资料，未经甲方同意，不得向第三方透露。

3、乙方实施项目的一切程序都应符合国家安全、保密的有关规定和标准。

4、乙方参加项目的有关人员均需同甲方签订保密协议。

第十四条　验收

1、算力服务、其余文件和工作由用户组织有关技术人员根据国家和行业有关规范、规程、标准和用户需求直接验收。

2、验收依据为招标文件、投标文件，国家和行业有关规范、规程和标准。

第十五条　付款方式

本项目合同支付上限为人民币大写壹仟壹佰柒拾伍万元整，小写¥11750000.00元（含税）。具体根据甲方实际申请使用的算力服务计算费用计算，算力服务单价详见附件二《算力服务目录单价表》。

1）预付款：合同生效且服务方案通过甲方审批后，支付合同总金额的25%作为预付款，即人民币大写贰佰玖拾叁万柒仟伍佰元整，小写¥2937500.00元（含税）。

2）结算款：项目服务期间，视实际情况，按季度开展结算及付款。费用按照算力服务单价*实际使用数量进行结算，算力服务单价详见附件二《算力服务目录单价表》，算力服务结算款优先从已支付的预付款中扣除。若已支付的预付款不足时，由乙方提出付款申请。

3）支付方式：待甲方财政资金到账且甲方收到乙方提供的合法、有效、等额的含税发票及费用清单后，甲方负责办理相关付款资料向乙方支付费用。

4）费用调整：若在合同执行过程中，因市场行情、服务内容变更等因素导致费用标准发生变化，需合同双方另行签署书面价格更改协议，双方应在季度结算时根据实际情况进行调整，并在结算明细中注明调整原因及金额。

备注：如果后续年度经人大审议通过的部门预算中，该采购项目预算金额较提前采购计划金额发生变化的，双方可根据相关规定签订补充协议或终止协议执行。

第十六条　争议解决办法

执行本合同发生的争议，由甲乙双方协商解决，如协商不成的，应提交甲方所在地人民法院诉讼解决。

第十七条　风险责任

1、乙方应完全地按照招标文件的要求和乙方投标文件的承诺完成本项目，出于自身财务、技术、人力等原因导致项目失败的，应承担全部责任。

2、乙方在实施荷载试验过程中应对自身的安全生产负责，若非因甲方原因发生的各种事故甲方不承担任何责任。

第十八条　违约责任

1、因乙方原因，未能按规定时间完成有关工作的，每延误一天，甲方可在支付合同余款中扣除合同价款万分之一。

2、由于乙方原因造成试验成果质量低劣，不能满足大纲要求时，应继续完善试验工作，其费用由乙方承担。

3、乙方交付的成果经验收不合格，应于7日内无条件修改，费用由乙方自行承担，在甲方要求整改后再次验收不合格的，甲方有权解除合同、要求乙方返还甲方已支付的合同款项，并有权要求乙方按合同总额 2 %支付违约金。

4、若甲方发现乙方派出的试验服务人员或提供的试验仪器设备不符合合同要求，乙方应在3天之内按要求派出人员或提供满足投标文件承诺的仪器设备，否则甲方有权终止合同，并保留追究乙方责任及要求赔偿损失的权利。

5、乙方或其工作人员违反本合同约定的保密义务，甲方有权要求乙方按合同总额 2 %支付违约金；造成不良影响或对甲方造成损失的，甲方有权要求乙方消除影响，承担赔偿责任，并有权解除合同。

6、因乙方提供的服务成果受到侵权指控或者引发法律纠纷，影响甲方使用服务成果或者导致合同目的不能实现的，甲方有权要求乙方按合同总

额 2 %支付违约金，并有权解除合同。

第十九条　合同组成部分

1、本协议书与下列文件一起构成合同文件，如下述文件之间有任何抵触、矛盾或歧义，应按以下顺序解释:

（1）政府采购合同协议书及其变更、补充协议

（2）政府采购合同专用条款

（3）政府采购合同通用条款

（4）中标（成交）通知书

（5）投标（响应）文件

（6）采购文件

（7）有关技术文件，图纸

（8）国家法律、行政法规和规章制度规定或合同约定的作为合同组成部分的其他文件。

2、下列文件均为本合同的组成部分:

（1）招标文件、答疑及补充通知;

（2）投标文件;

（3）本合同执行中共同签署的补充与修正文件。

第二十条　争议的解决

本合同未尽事宜，双方协商解决，并以书面形式补充。协商不成的，任何一方均可以向甲方所在地人民法院起诉。

第二十一条　通知条款

（一）在本合同有效期内，因法律、法规、政策的变化，或任一方丧失履行本合同的资格或能力影响本合同履行的，该方应承担在合理时间内通知其他各方的义务。

（二）合同各方同意，与本合同有关的任何通知，以书面方式送达方为有效。书面形式包括但不限于：传真、快递、电子邮件。

（三）通知送达下列地点或传至下列传真号码或发至下列电子邮箱，无论送达或无法送达或拒收均视为有效送达：

甲　方：深圳市福田区民意速办和智慧城市建设中心

地　址：深圳市福田区深南大道 1006 号国际创新中心 F 座 3 楼

收件人：[redacted]　　邮　　编：518000

电　话：[redacted]　　传真号码：\

电子邮箱：

乙　方：　中国联合网络通信有限公司深圳市分公司

地　址：深圳市福田区福田街道 4005 号深南大道中国联通大厦

收件人：[redacted]　　邮　　编：518000

电　话：[redacted]　　传真号码：\

电子邮箱：[redacted]

（四）任何一方约定的联系方式发生变更，应提前 3 个工作日通知对方，并以书面方式告知对方变更后的联系方式。因一方前述约定联系方式问题或一方联系方式变更未履行提前书面告知义务等原因导致对方所发送的通知等文件无法送达、送达被拒收、邮件返回的，视为该等通知已送达相对方，且相对方已知悉通知及文件所记载之内容。

（五）特别约定：本合同约定的地址为争议解决时人民法院或仲裁机构的法律文书送达地址，人民法院或仲裁机构的诉讼文书（含裁判文书）向合同任何一方当事人的上述地址送达的，无论送达或拒收均视为有效送达。

（六）如上述地址未约定的，以双方当事人的注册地址作为送达地址。

（七）本合同送达条款与争议解决条款均为独立条款，不受合同整体

或其他条款的效力的影响。

第二十二条 合同的生效与份数

本合同壹式肆份，甲方执贰份，乙方执贰份，均具同等法律效力。本合同自双方法人代表签字（盖章）认可之日起生效。

本合同未尽事宜，双方友好协商，达成解决方案，经双方签字后，可作为本合同的有效附件。

甲方（盖章）：深圳市福田区民意速办和智慧城市建设中心

法定代表人（或委托代理人）：[redacted]

日期：2026年 6 月 3 日

乙方（盖章）：中国联合网络通信有限公司深圳市分公司

法定代表人（或委托代理人）：[redacted]

日期：2026年 6 月 3 日

附件一 算力技术要求

序号	指标项		算力技术要求
1	算力卡 1	CPU	每台物理机配置为 2 颗 CPU，每 CPU 为 48 物理核，主频为 2.4GHz
2		内存容量	每台物理机配置为 32 个内存,每个内存为 64GB
3		配套存储	每台物理机配置为 2*3.2TB · NVMe · SSD 缓存盘
4		AI 智算卡	▲每台物理机配置为 8 张智算卡，单卡显存为 80GB，智算卡半精度(FP16)为 989TFLOPS
5		卡间互联能力	▲节点内卡间互联带宽为 400GB/s，跨节点卡间互联 RDMA 带宽为 400Gb/s
6		部署形式	支持裸金属部署，支持如训练、推理、渲染等不同场景不同性能，满足业务全场景需求
7	算力卡 2	CPU	每台物理机配置为 2 颗 CPU，每 CPU 为 96 个物理核，主频为 2.6GHz
8		内存容量	每台物理机配置为 24 个内存，每个内存为 96GB，内存类型 DDR5
9		配套存储	每台物理机配置为 4*6.4TB NVMe SSD 硬盘。
10		AI 智算卡	▲每台物理机配置为 8 张智算卡，单卡显存为 96GB，智算卡半精度（FP16）为 145TFLOPS
11		卡间互联能力	▲卡间互联带宽为 900GB/s,跨节点卡间互联 RDMA 带宽为 400Gb/s
12		部署形式	支持裸金属部署，支持如训练、推理、渲染等不同场景不同性能，满足业务全场景需求。
13	算力卡 3	CPU	每台虚拟机配置为 32 核 vCPU，主频为 2.6GHz
14		内存容量	每台虚拟机内存为 96GB，内存类型 DDR5
15		配套存储	每卡配套存储为 640GB
16		AI 智算卡	▲每台虚拟机配置为 1 张智算卡，单卡显存为 48GB，智算卡半精度（FP16）为 110TFLOPS
17		卡间互联能力	▲卡间互联带宽为 64Gbs
18		部署形式	支持不同如推理、渲染等不同场景不同性能，满足业务全场景需求。
19	算力卡 4	CPU	每台虚拟机配置为 10 核 vCPU，主频为 2.5GHz
20		内存容量	每台虚拟机内存为 40GB，内存类型 DDR4
21		配套存储	每卡配套存储为 640GB
22		AI 智算卡	▲每台虚拟机配置为 1 张智算卡，单卡显存为 32GB，智算卡半精度（FP16）为 125TFLOPS
23		卡间互联能力	▲卡间互联带宽为 300GB/s
24		部署形式	支持如推理、渲染等不同场景不同性能，满足业务全场景需求。
25	算力卡 5	CPU	每台物理机配置为 2 颗 CPU，每 CPU 为 40 物理核，主频为

			2.5GHz
26		内存容量	每台物理机配置为 32 个内存,每个内存为 64GB
27		配套存储	每台物理机配置为 2*3.2TB · NVMe · SSD 缓存盘
28		AI 智算卡	▲每台物理机配置为 8 张智算卡，单卡显存为 64GB，智算卡半精度(FP16)为 310TFLOPS
29		卡间互联能力	▲节点内任意两卡间互联带宽为 392GB/s，跨节点卡间任意两卡间互联带宽为 200Gb/s
30		部署形式	支持如训练、推理、渲染等不同场景不同性能，满足业务全场景需求
31	算力卡 6	CPU	每台物理机配置为 2 颗 CPU,每 CPU 为 32 物理核,主频为 2.6GHz
32		内存容量	每台物理机配置为 4 个内存，每个内存为 32GB
33		配套存储	每台物理机配置为 1*8TB SATA 缓存盘
34		AI 智算卡	▲每台物理机配置为 6 张智算卡，单卡显存为 8GB，智算卡(INT8)为 10.9TOPS，采用国产数据流架构芯片，芯片利用率即实测 TOPS 与峰值 TOPS 的比高于 60%；单张计算板卡工作功耗低于 58W；
35		部署形式	支持 AI 算法推理等场景
36	存储指标要求	整体需求	文件存储应提供可扩展的共享文件存储服务，满足与计算服务器，GPU 服务等服务搭配使用。需搭载标准的 NFS 文件系统访问协议，为多个实例或其他计算服务提供共享的数据源，支持弹性容量和性能的扩展。满足现有应用无需修改即可挂载使用。
37		功能参数	提供成熟的存储系统，按照使用容量计费，无需手动扩容。
38			支持快照和定期快照策略的方式对文件系统进行数据保护。
39			支持在控制台上，直接将云文件存储挂载至云服务器的指定目录
40			支持不低于 500 个计算实例同时访问文件存储
41		支持协议	支持 NFS v4.x 访问协议,支持 NFS 和 CIFS 协议交叉互访，支持 Linux 和 Windows 客户端共同访问同一个文件系统
42		带宽	支持带宽性能随文件系统容量线性提升，每增加 1TB，带宽提升高于 100MB/s，最大不超过 240GB/s。
43		IOPS	支持 IOPS 性能随文件系统容量线性提升，每增加 1TB，IOPS 提升高于 1300，最大不超过 3120000。
44		时延	4K 随机读延迟为 2ms，4K 随机写延迟 1~3ms
45		文件数量	支持文件数量随文件系统容量线性提升，每增长 1TB，文件数量支持增长高于 400 万个，最大可支持百亿级别
46		RDMA 组网	支持到每个计算节点的组网带宽为 200Gb
47		CPU	每台物理机配置为 2 颗 CPU，每 CPU 为 12 物理核，主频为 2.0GHz

48	AI 训推一体机 1	内存容量	每台物理机配置为 8 个内存,每个内存为 64GB
49		配套存储	每台物理机配置为 2*3.84TB · NVMe · SSD 缓存盘
50		AI 智算卡	▲每台物理机配置为 2 张智算卡，单卡显存为 48GB，总智算卡半精度(FP16)为 130TFLOPS。每台物理机支持训练和推理参数量为 320 亿（32B）的 AI 模型，且在搭载参数量为 320 亿（32B）的 AI 大模型情况下，tokens 总吞吐率为 1536tokens/s，并发数为 192。
51		AI 扩容卡	每台物理机配置为 2 张扩容卡，单卡可扩大显存 2TB
52		卡间互联能力	▲节点内卡间互联带宽为 64GB/s
53		部署形式	1.支持裸金属部署，支持如训练、推理等不同场景不同性能。2.支持 DeepSeek、千问等主流开源大模型。3.支持对模型文件通过参数设置及微调、输入专属数据集训练，生成新的专属领域大模型。
54	AI 训推一体机 2	CPU	每台物理机配置为 2 颗 CPU，每 CPU 为 12 物理核，主频为 2.0GHz
55		内存容量	每台物理机配置为 8 个内存,每个内存为 64GB
56		配套存储	每台物理机配置为 2*3.84TB · NVMe · SSD 缓存盘
57		AI 智算卡	▲每台物理机配置为 4 张智算卡，单卡显存为 48GB，总智算卡半精度(FP16)为 270TFLOPS。每台物理机支持训练和推理参数量为 700 亿（70B）的 AI 模型，且在搭载参数量为 700 亿（70B）的 AI 大模型情况下，tokens 总吞吐率为 2236tokens/s，并发数为 192。
58		AI 扩容卡	每台物理机配置为 2 张扩容卡，单卡可扩大显存 2TB
59		卡间互联能力	▲节点内卡间互联带宽为 64GB/s
60		部署形式	1.支持裸金属部署，支持如训练、推理等不同场景不同性能。2.支持 DeepSeek、千问等主流开源大模型。3.支持对模型文件通过参数设置及微调、输入专属数据集训练，生成新的专属领域大模型。

附件二 算力服务目录单价表

序号	智能服务（算力卡）	服务配置	计量单位	单卡限价	部署方式
1	算力卡1	显存≥80GB，智算卡半精度(FP16)≥989TFLOPS，配套CPU≥10核，内存≥256GB，裸存储≥800GB	元/卡/月	12796.58	本地化部署
2	算力卡2	显存≥96GB，智算卡半精度(FP16)≥145TFLOPS，配套CPU≥12核，内存≥128GB，存储≥1024GB	元/卡/月	4940.00	本地化部署
3	算力卡3	显存≥48GB，智算卡半精度(FP16)≥110TFLOPS，配套CPU≥16核，内存≥96GB，存储≥640GB	元/卡/月	1827.80	本地化部署
4	算力卡4	显存≥32GB，智算卡半精度(FP16)≥125TFLOPS，配套CPU≥16核，内存≥48GB，存储≥640GB	元/卡/月	2779.24	本地化部署
5	算力卡5	显存≥64GB，智算卡半精度(FP16)≥310TFLOPS，配套CPU≥24核，内存≥256GB，存储≥800GB	元/卡/月	4787.85	本地化部署
6	算力卡6	显存≥8GB，智算卡算力(INT8)≥10.9TOPS，配套CPU≥10核，内存≥21GB，存储≥1TB	元/卡/月	1647.00	本地化部署
7	高性能存储	100G	元/月	79.04	本地化部署
8	AI训推一体机1	显存≥96GB，总智算卡半精度(FP16)≥130TFLOPS。支持训练和推理参数量≥320亿（32B）的AI大模型，且在搭载参数量≥320亿（32B）AI大模型情况下，tokens总吞吐率≥1500tokens/s，并发数≥180。	元/台/月	11757.20	本地化部署
9	AI训推一体机2	显存≥192GB，总智算卡半精度(FP16)≥270TFLOPS。支持训练和推理参数量≥700亿（70B）的AI大模型，且在搭载参数量≥700亿（70B）AI大模型情况下，tokens总吞吐率≥2000tokens/s，并发数≥180。	元/台/月	13930.80	本地化部署