

# 项目需求说明

## 1、项目背景

气候系统模式是研究全球气候变化的重要工具之一，已经在地球科学研究中得到广泛应用。但遗憾的是，目前模式性能尚不能完全满足气候变化模拟和预测研究的需求。例如，对参与国际耦合模式比较计划 (Coupled Models Intercomparison Project, CMIP) 的耦合模式结果的比较和评估分析表明，尽管当前气候模式在很多方面具有不错的模拟能力，但是模拟结果仍然存在较大的不确定性，主要表现在热带模拟误差、云—气溶胶—辐射气候反馈过程、区域尺度气候模拟、极端事件模拟等许多方面。大量研究指出，气候模式的不确定性主要来自次网格参数化过程，因此提高模式空间分辨率，直接分辨更多的中小尺度动力过程，从而减小对参数化过程的依赖性，是目前物理气候系统模式的重要发展趋势之一。以往的天气预报和气候模式研发经验也表明，提高模式分辨率是减小模式误差的重要途径之一。这是因为大气和海洋环流包含各种时空尺度的运动，而且不同尺度运动之间存在密切的相互作用。但是任何气候模式的空间分辨率都是有限的，为了考虑次网格尺度的过程对于可分辨尺度环流的贡献和影响，模式通常采用次网格物理参数化过程描述其对大尺度环流的影响和贡献。这些物理参数化过程一般是基于某种前提条件和有限的观测数据建立的，缺乏普适性，往往是模式不确定性的最大根源之一。因此，当模式分辨率提高之后，模式就能直接模拟出原本一些需要进行次网格参数化的过程，因而会减小模式对物理参数化过程的依赖性，对模式性能改善大多会起到正面效果。但另一方面，任何气候系统模式的空间分辨率与次网格参数化过程应该相互匹配，二者具有同等的重要性。因此高分辨率模式必须同时兼顾动力框架和次网格参数化过程的改进，才有可能取得更好的模拟效果。

自20世纪80年代末以来，中国科学院大气物理研究所大气科学和地球流体力学数值模拟国家重点实验室 (IAP/LASG) 的科学家一直致力于自主发展全球海洋环流模式。为了更好地模拟海洋中尺度涡旋和西边界流等特征，IAP/LASG学者在改进后的第四代全球海洋环流模式，即LASG气候系统海洋模式LICOM2.0 (LASG/IAP Climate system Ocean Model) 的基础上，把模式水平分辨率提高到

全球均匀的0.1度(10 km)，垂直分辨率提高到55层，并对模式的动力框架和并行方案进行必要改动之后，建立了一个准全球(模式区域为78° S~66° N，不包含北冰洋)涡分辨率海洋环流模式。

实验室建设了千万亿次科研液冷高性能计算系统，包含刀片计算节点、管理/登录节点、100G 高速计算网络、管理网络、并行存储系统、KVM 系统、机柜及配电系统、集群管理软件、并行文件系统、作业调度软件、并行开发环境等全套软硬件组成部分，并基于此建设气候资料库和应用平台。

现需提供面向“未来地球”的高分辨率数值模式研制液冷高性能计算系统安装部署所需要的数据中心场地和相关配套条件，满足高性能计算系统运行的环境要求，并提供计算、存储、网络、安全、系统软件、并行环境、应用软件、大气科学数据应用平台、气候基本数据库和个例库的运行维护与技术支持。

投标人需协助甲方安装部署的系统包括但不限于：IT 主机系统、软件系统、配电系统、冷板式液体冷却系统，提供匹配环境包括但不限于独立场地、双路高可用供电系统、备用 UPS 电源、冷冻水制冷环境、动环监控系统、可视化系统等。

投标人建立专用运维管理和技术支持团队负责对基础环境系统、主机软硬件系统、应用平台的运行维护工作，并对相关工作提供技术支持。

本次采购服务期限为 1 年。

2. 系统主要硬件清单

货物分项	主要规格	数量
高性能计算平台 ——双路计算节点	4U 机架式刀片机箱，可以支持 8 个液冷 CPU 计算刀片； 1*管理模块，集成远程 KVM 和远程虚拟媒体； 1*千兆网络交换模块（4 个 RJ45 千兆接口+2 个万兆 SFP+接口）； 支持刀箱液冷分液单元； 满配冗余电源模块、散热模块；	75
	2*Intel Xeon Gold 6132 处理器（14 核，2.6GHz）； 12*8GB DDR4 2666MHz ECC REG 内存； 1*240G 企业级 SSD 硬盘； 1*100Gb Omni-Path 高速网络端口； 2*千兆以太网接口； 主机标配 2 个 PCI-E 插槽，支持 1 个 X16 插槽、1 个半高半长标准卡	600

高性能计算平台 ——四路计算节点	4*Intel Xeon Gold 6132 处理器（14 核，2.6GHz）； 192GB DDR4 2666MHz ECC REG 内存； 2*600G 10K 12Gb SAS 硬盘； 1*100Gb Omni-Path 高速网络端口； 4*千兆以太网接口； 满配冗余电源风扇；	40
高性能计算平台 ——32 路大内存节点	基于 NUMA 架构的高性能计算系统，全局共享内存体系结构，单一操作系统映像；单一节点最大可扩充至 256 颗物理 CPU(4096CPU 核)，最大可扩充至 64TB 内存； 32*Intel Xeon E5 4627v3 处理器（10 核，2.6GHz）； 2*NVIDIA Quadro® Plex 7000 图形子系统； 80*32GB DDR4 2666MHz ECC REG 内存； 2*480GB SSD 硬盘； 1*1GB Cache PCI-E 八通道 SAS RAID 卡，支持 RAID0/1/5/10/50； 2*千兆以太网端口； 2*100Gb Omni-Path 高速网络端口； 满配冗余电源风扇；	1
高性能计算平台 ——GPU 计算节点	2U 机架式； 2*Intel Xeon Gold 6132 处理器(14 核，2.6GHz)； 2*Nvidia Tesla P100 GPU 卡； 96GB DDR4 2666MHz ECC REG 内存； 2*480G 2.5 寸 SSD 热插拔硬盘； 4*千兆以太网接口； 1*100Gb Omni-Path 高速网络端口； 满配冗余电源风扇；	2
高性能计算平台 ——KNL 计算节点	2U4 节点机架式； 1*Intel KNL 7210 处理器（64c 1.3GHz）； 6*32GB DDR4 2666MHz ECC REG 内存； 4*240G 2.5 寸 MLC SSD 热插拔硬盘； 4*千兆以太网接口； 1*100Gb Omni-Path 高速网络端口； 满配冗余电源风扇；	4
高性能计算平台 ——管理节点	2*Intel Xeon Gold 6132 处理器（14 核，2.6GHz）； 96GB DDR4 2666MHz ECC REG 内存； 4*1.2TB 10K 12G SAS 2.5 寸热插拔硬盘； 1*2GB Cache 12Gb RAID 控制器； 1*100Gb Omni-Path 高速网络端口； 4*千兆以太网接口； 满配冗余电源风扇；	2

高性能计算平台 ——登录节点	2*Intel Xeon Gold 6132 处理器（14 核，2.6GHz）； 192GB DDR4 2666MHz ECC REG 内存； 4*1.2TB 10K 12G SAS 2.5 寸热插拔硬盘； 1*2GB Cache 12Gb RAID 控制器； 1*100Gb Omni-Path 高速网络端口； 4*千兆以太网接口； 满配冗余电源风扇；	2
高性能计算平台 ——并行存储系统	ParaStor300 并行存储系统，包含 2 个索引控制器（OPara-G20）和 12 个数据控制器（OStor-G20），共提供 432 个磁盘槽位、11.52GB 缓存加速、2160TB 数据硬盘存储容量和 15GB/S 以上带宽、14 个 100Gb/S 高速数据接口，可用容量不低于硬盘容量的 80%。 每个索引控制器采用高性能 64 位处理器，64GB Cache，Raid6 模式保护的高速 Flash 存储，内嵌高性能数据索引引擎，管理客户端的并行访问，实现全局统一命名空间，并支持索引控制器以 Active-Active 集群方式运行，能够满足高 IOPS，高带宽等各类应用需求；每个数据控制器采用高性能 64 位处理器，64GB Cache，内嵌高性能数据存取引擎，并行处理所有客户端的数据访问请求，支持分级存储，并支持多个数据控制器以节点间冗余方式容错。部署 ParaStor300 并行文件系统及客户端软件，客户端软件提供光盘介质和实施服务，无 License 数量限制；五年原厂商免费技术支持。	1
高性能计算平台 ——磁盘阵列存储系统	NCS7700 高端高密度磁盘阵列存储，多控虚拟化系统。共配置 12 个控制器，6*128GB 高速缓存；48 个 16Gb FC 端口和 48 个千兆网络端口；包含 Thin-provisioning、快照、本地卷镜像；支持远程复制、数据压缩、外部虚拟化；支持 CIFS, NFS, FTP 等协议；支持 SSD/SAS/NL-SAS 混插，单台最多支持 4032 颗盘；单扩展柜 92 盘位，配置 6T 3.5 寸专用 SAS 磁盘, 7.2K rpm*672, 480G SSD*4, 容量配置为 4032TB。	1
高性能计算平台 ——光纤交换机	48 端口交换机，48 端口激活，双电源（热拔插），含 48 个 16Gb 短波 SFP，含 Web tools、Zoning 软件授权，支持级联，提供所需线缆。	2

高性能计算平台 ——共享文件系统 IO	2*Intel Xeon Gold 6132 处理器 (14 核, 2.6GHz); 96GB DDR4 2666MHz ECC REG 内存; 4*1.2TB 10K 12G SAS 2.5 寸热插拔硬盘; 1*2GB Cache 12Gb RAID 控制器; 1*100Gb Omni-Path 高速网络端口; 2*双口 16Gb HBA 卡; 4*千兆以太网接口; 满配冗余电源风扇;	8
高性能计算平台 ——高速计算存储网络	模块化 100Gb OPA 交换机, 交换模块满配, 管理模块满配, 电源模块满配, 最大可支持 720 端口, 支持全线速连接、胖树、Torus 等组网方式; 匹配不低于 673 根 100Gb OPA 光纤线缆、满足本次系统连接需求。	1
高性能计算平台 ——管理网络	48 个 10/100/1000M 自适应电接口, 复用 4 个 SFP 接口 (SFP 为千/百口), 2 个扩展槽扩展万兆接口, 1+1 冗余电源, 提供系统集成所需的千兆以太网线缆、万兆网络模块和万兆光纤线缆。	2
高性能计算平台 ——IPMI 网络	48 口 SFP+ 万兆光口交换机 (含 4 个 40GE 端口), 含冗余风扇、冗余电源, 提供系统集成所需的千兆以太网线缆、万兆网络模块和万兆光纤线缆。	2
高性能计算平台 ——IPMI 网络	48 个 10/100/1000M 自适应电接口, 复用 4 个 SFP 接口 (SFP 为千/百口), 2 个扩展槽扩展万兆接口堆叠, 1+1 冗余电源, 提供系统集成所需的千兆以太网线缆、万兆网络模块和万兆光纤线缆。	4
高性能计算平台 ——系统外部网络接入与网络安全	配备路由、流控、VPN、防火墙功能, 配置 8 个千兆电口、4 个千兆光口, 1U 机架式, 液晶监控屏; 吞吐量 10Gbps, 最大并发连接数 700 万, 每秒新建连接数 20 万; 单电源, 内存 8G, 硬盘 120G SSD/1TB SATA; 支持 IPS、防病毒和应用识别、上网行为管理、流量控制功能; 支持 IPSec VPN 和 SSL VPN, 支持链路和服务负载均衡, 支持 WEB 分类过滤和垃圾邮件功能; 支持防火墙虚拟化; 支持 IPv6 协议。产品没有用户数许可限制, 含病毒库、IPS 库、应用控制库、WEB 过滤库升级服务	1
高性能计算平台 ——控制台	Cluskvml00 1U 手动伸缩控制台 (17" 液晶显示器、鼠标、键盘、16 口切换器等), Cluskvml00 菊花链 KVM 主控端及 CIM 模块, 满足系统所需	2
高性能计算平台 ——液体冷却系统	闭式冷却塔, 型号 CCT-D70CUF, 流量 70m3/h, 1+1 冗余设计, 并机后带载所有室内 CDM; 液冷分配模块, 用于为液冷服务器高温水提供冷源换热及分配管理, 二次侧进/回水温度 35/45 度工况下、单机制冷量为 35KW, 每台 CDM 最多带载 8 箱液冷刀片服务器,	

	机柜底部机架式安装，高 4U，供电接口为 C14、PDU 取电；包含水水换热器、冗余循环水泵、控制模块、安装托轨和压力无关节阀；一次侧和二次侧进出管方式均为 DN25 卡盘连接。竖直分液单元，包含竖直分液单元 VCDU 共 2 个，每个竖直分液单元包含 8 个分液端口，每个端口形式为无滴漏快速连接器公头；液冷自控系统，PLC 控制系统，可自动调节液冷管路系统。	
其他基础设施	机柜、电气、冷却系统等其他基础设施设备	1

### 3. 技术服务内容和技术要求

序号	指标项	服务内容及技术要求
1	总体要求	
1.1	总体要求	<p>提供面向“未来地球”的高分辨率数值模式研制液冷高性能计算系统安装部署所需要的场地和相关配套条件，满足高性能计算系统运行的环境要求，并提供计算、存储、网络、安全、系统软件、并行环境、应用软件、大气科学数据应用平台、气候基本数据库和个例库的运行维护与技术支持。</p> <p>投标人需协助甲方安装部署的系统包括但不限于：IT 主机系统、软件系统、配电系统、冷板式液体冷却系统，提供匹配环境包括但不限于独立场地、双路高可用供电系统、备用 UPS 电源、冷冻水制冷环境、动环监控系统、可视化系统等。投标人建立专用运维管理和技术支持团队负责对基础环境系统、主机软硬件系统、应用数据库和数据应用平台的运行维护工作，并对相关工作提供技术支持。</p>
2	IDC 基础环境	
2.1	安装环境	独立系统设置相对独立的安装部署场地，满足全部软硬件及配套设备安装需求；室内可部署高密度的高性能计算系统，要求单个 42U 机位承载 IT 系统功耗指标不低于 20KW；室外可部署专用液冷冷却塔和冷冻水冷机。
2.2	供电需求	要求具备双路高可用供电环境，双路电源来自不同的 110KV 站
2.3	备用电源	为关键设备提供双路 UPS 备用电源
2.4	制冷需求	部署甲供冷板式液体冷却系统，提供额外冷冻水制冷环境，满足系统冷量要求。
2.5	动环监控	提供动环监控对：UPS 设备监测、通讯转换模块、精密空调设备监测、通讯转换模块、供配电监测、通讯模块、温湿度监测、智能温湿度传感器、漏水监测系统（机房）、定位漏水控制器、引出线、定位感应绳、终止端、螺杆机、冷塔、环境监控设备联网集中监控、串口服务器、监控中心（双机热备）、监

		控主机、液晶显示器、数据中心综合管理组态软件、语音卡、短信服务器、多媒体语音报警系统、电子邮件报警、控制输出模块、信号采集箱、电源、动环线缆等。实现对机柜排内空调、配电、机柜等设备以及机柜排内微环境进行实时监控和管理。
2.6	环境运维	1) 定期对机房环境设备做定期检测，其内容包括：连接接口、线缆、电源输入/输出稳定性，并形成检测报告； 2) 分析和解决日常设备运行中出现的问题，并提交故障记录与修复记录，给出安全操作建议； 3) 提供对机房环境设备的月度巡检，确保各设备运行稳定； 4) 为液冷和水冷系统、电气系统提供整年维保和运维服务； 5) 为机房及楼宇视频监控系统故障设备提供维修、更换及除尘服务； 6) 提供清洁除尘服务； 7) 提供消防系统检查及维护服务；
2.7	可视化系统	提供配套使用的可视化系统环境，并将动环监控系统、集群管理与作业调度系统、在线运维系统、气象应用等数据接入可视化系统，可同时呈现多个系统界面。
2.8	能源消耗	系统部署后在服务期限内全部水电网消耗等运行费用均由投标人承担，请按前表清单评估成本，采购人承诺系统总体功耗不高于 600KW，年用电量不超过 525.6 万度。
2.9	材料费用	系统运维期间所需的全部维修、维保和辅助材料费用，均由中标人承担。
3	计算子系统运维	
3.1	日常运维	1) 定期对服务器及终端设备做巡检，检查和清理系统 Error Log File，确保各设备运行稳定，并形成巡检报告； 2) 分析日常设备运行中出现的问题，完成由设备引起的故障修复，并提交故障记录与修复记录，给出安全操作建议； 3) 针对实际运行情况，提出合理的升级和扩容建议； 4) 提供配合安全检查服务； 5) 为关键小型机设备提供整年维保； 6) 对关键设备故障提供快速响应服务； 7) 为部分设备关键部件提供故障换新服务；
3.2	系统监控	系统状态 7*24 小时监控
3.3	系统管理	规范用户按规定使用服务器，按规定投递运行任务。按时更新系统漏洞、修复故障系统。根据业务调整系统参数。
3.4	用户管理	对所有用户进行统一认证，按要求开通、禁用账号，并配置相关权限。
3.5	软件管理	提供 Intel 编译器、GCC、Python、Perl、R、MPI 等常用公共软件。安装、维护、升级上述常用公共软件，协助用户安装需要使用的软件以及软件环境。
3.6	数据上传	将甲方送至数据中心的硬盘数据上传至用户指定目录，并修改对应文件权限。
4	存储子系统	

4.1	日常运维	1) 提供对存储及备份设备的定期巡检，确保存储及备份设备运行稳定； 2) 提供配合安全检查服务； 3) 提供设备故障修复服务，提交故障记录与修复记录，并给出安全操作建议； 4) 保证应用系统对存储设备的正常读取、写入和访问，满足应用系统对磁盘空间的使用提出的需求，备份策略设置需要按照业务数据更新频率进行动态调整； 5) 运维期间至少做一次备份数据有效性验证； 6) 对设备故障提供快速响应服务； 7) 对易损件提供备件储备；
4.2	系统监控	系统状态 7*24 小时监控，存储瓶颈分析，存储使用量预警通知。
4.3	空间管理	按需求分配、回收、扩展目录空间，统计项目、用户、部门总使用量。
4.4	系统优化	根据不同业务情况，优化存储参数，维护文件系统。
5	网络与安全系统运维及环境要求	
5.1	日常运维	1) 提供定期巡检服务，确保各网络互联设备运行稳定； 2) 提供每月一次的网络带宽使用情况报告； 3) 提供配合安全检查服务； 4) 完成由网络设备引起的故障修复，并提交故障记录与修复记录； 5) 为关键网络安全设备提供整年维保，及对相关软件提供升级服务； 6) 提供漏洞扫描特征库升级及定期扫描服务，并提供漏洞报告及修补建议；
5.2	高速 Torus 网络	进行 Torus 互连网络组网，用大小固定的交换单元互连而成一个多维环绕的格子型网，每个交换单元通过单向或双向链路与其相邻节点相连，并具有双向数据通路的连接线卡，路由选择时要求在每个方向上分别经过若干转接点。本次系统设计互连采用 6D-Torus 网络结构，系统整体计算网络分为三层。第一层为全交换网络，8 个计算节点（Node）通过全交换网络高速连接，形成一个超节点（SuperNode）。每一个超节点对外有 40 条连接，分为为 A、B、C+、C-、X+、X-、Y+、Y-、Z+、Z-。其中 A、B、C+、C- 这四条连接的带宽为：单向传输带宽为 400GBps，双向传输带宽为 800GBps。X+、X-、Y+、Y-、Z+、Z- 的连接带宽为：单向传输带宽为 400GBps，双向传输带宽为 800GBps。12 个超节点分别通过 A、B、C+、C- 网络链路连接起来，每个超节点通过 A、B、C+、C- 四条连接（每个连接 800GBps 双向传输带宽）连接到 4 个超节点，形成一个 2×2×3 的三维立方体结构，每一个硅元中的超节点都通过 X+、X-、Y+、Y-、Z+、Z- 六个连接（每个连接 800GBps 双向传输带宽），连接其前、后、左、右、上、下六个硅元中的对应的超级节点最终形成整个系统。



5.5	管理网络	维护并优化 6 台 48 端口千兆，2 台 48 端口万兆交换机。
5.6	安全环境	提供防火墙和 VPN 设备，满足系统外部用户访问需求 数据中心网络和安全基础环境满足系统等保三级测评的基本要求
6	集群管理	
6.1	调度系统	商业集群管理软件；部署 Slurm、PBS、SGE 调度系统
6.2	系统配置	完成对整个集群的检测和配置，如检测网络可用性、配置 SSH/RSH 无密码登陆、停止系统无用服务、配置 NFS、同步集群时间、同步集群用户等；能够对集群进行快速软件部署和环境配置、远程命令执行，实现远程开关机。
6.3	状态监控	直观的集群物理拓扑视图，监控包括总体及各计算节点的 CPU 平均使用率、内存剩余率、磁盘剩余率、集群整体 I/O 运行，集群节点的静态属性信息、动态属性信息，显示节点的环境参数等。
6.5	告警管理	支持实时告警管理，可定制告警的方式和发送策略；支持历史告警的查询、分析和报表统计，能够监控主要故障点信息，对外提供异常告警信息访问 API。
6.6	集群报表	提供报表系统，能按时间体现集群系统总体运行情况，可生成关联报表，并提供关系分析，导出成 PDF 和 Excel 等文件格式
6.7	集群管理	提供 IP 管理，Hosts 管理，进程管理，服务管理，远程开机/关机管理
6.8	计费功能	对用户或队列使用的资源（CPU、磁盘等）计费并生成报表。
7	数据库、个例库与数据应用平台维护	
7.1	气候基本数据库与个例库	定期对数据库进行检查，并提交检查报告，保证软件运行稳定；提供配合安全检查服务；制定合理的数据库备份策略，保证数据安全；提供数据库故障修复服务，提交故障与修复记录，并给出安全操作建议。
7.2	数据应用平台	对应用平台软件运行状态进行定期检查，保证软件运行稳定；提供配合安全检查服务；提供应用平台软件配置文件定期备份，保障应用安全；提供应用平台软件的优化检查及故障修复服务，并提交故障与修复记录，给出安全操作建议。
8	技术服务人员要求	
8.1	IT 基础设施运维	设备部署的数据中心至少具备 4 名或以上基础硬件运维工程师
8.2	数据中心基础设施运维	设备部署的数据中心至少具备 2 名或以上水电暖通运维工程师
8.3	调度系统开发与优化	配备 HPC 资深工程师，可协助完成调度策略优化，提升集群使用效率
8.4	运维开发与运营服务	配备运维软件开发与运营服务人员，具备运维和运营软件开发和维护使用经验，可协助进行软件开发和运营管理服务
9	服务标准	
9.1	服务期限	7*24 小时*365 天*3 年。

9.2	服务响应	即时响应；系统部署地现场驻场服务，发现问题即时响应；同时向用户单位所在地派驻技术支持工程师，可响应用户现场技术支持需求。
10	管理要求	
10.1	管理机制	<ol style="list-style-type: none"> <li>1) 编制运维实施进度计划并组织实施，定期向业主提供运维实施进展情况报告，负责协调解决实施过程中的各种技术问题，并做好技术把关工作；</li> <li>2) 按招标人安排，对运维成果进行定期汇报；</li> <li>3) 协调、指导和解决运维实施过程中出现的各种技术问题；</li> <li>4) 协助业主处理实施过程中出现的重大事故；</li> <li>5) 协调好各参与运维单位之间的关系；</li> <li>6) 驻场人员的日常组织、管理、协调，做好工作记录，按期向招标人报告工作进展，按招标人要求提供相关文档；</li> <li>7) 根据实施计划，定期检查其工作情况，若发现不合格事项，应及时向业主反映，并给出可行的解决方案；</li> <li>8) 遵循相应的运维管理规范，并提交相应的运维管理文档；</li> <li>9) 与业主建立及时有效的沟通机制；</li> <li>10) 计划周密，当运维实施过程中出现与目标有较大偏差时，要能及时做出调整，对运维任务中做出的调整，要及时汇报，并尊重业主的安排。</li> </ol>
10.2	文档管理	<ol style="list-style-type: none"> <li>1) 对所有服务范围内的设备和软件，建立相应的系统维护档案，并记录每次故障响应服务和预防性维护的详细信息；</li> <li>2) 所有档案信息及报告同步提供给用户接口人员；</li> <li>3) 维护合同期满时，装订汇总后与年度服务报告一起移交。</li> </ol>