

地球系统模式 CAS-ESM 2.0 性能评估与分析^①



王天一^{1,2}, 姜金荣¹, 迟学斌¹, 张贺³, 何卷雄³, 郝卉群¹

¹(中国科学院 计算机网络信息中心 高性能计算部, 北京 100190)

²(中国科学院大学, 北京 100049)

³(中国科学院 大气物理研究所 国际气候与环境科学中心, 北京 100029)

通讯作者: 王天一, E-mail: wangty@sccas.cn

摘要: 地球系统模式是研究气候变化、进行地球系统建模的重要软件。中科院地球系统模式 CAS-ESM (Chinese Academy of Sciences-Earth System Model) 是中科院大气所发展的进行地球系统模拟的高性能计算应用软件, 目前已经发布了 2.0 版本, 其模拟性能一直是制约其发展的关键因素之一。为了对 CAS-ESM 2.0 进行性能评估和分析, 将 CAS-ESM 2.0 移植到中科院高性能计算系统“元”和“地球系统数值模拟装置”原型系统这两大高性能计算平台上, 开展了耦合数值模拟试验。试验结果显示, CAS-ESM 2.0 存在受平台影响的性能差异, 大气模式的运行时间占比最高, 超过了其他分模式的总和, 部分分模式存在可扩展性问题。然后对试验结果进行了进一步的分析, 发现大气模式的性能瓶颈主要是由通信造成的。因对 CAS-ESM 2.0 的后续研发工作中, CAS-ESM 的跨平台优化、大气模式的性能优化与并行算法改进、分模式的可扩展性应该是研究的重点之一。

关键词: 地球系统模式; 高性能计算; CAS-ESM; 性能评估; 耦合试验

引用格式: 王天一, 姜金荣, 迟学斌, 张贺, 何卷雄, 郝卉群. 地球系统模式 CAS-ESM 2.0 性能评估与分析. 计算机系统应用, 2021, 30(6):9-17. <http://www.c-s-a.org.cn/1003-3254/7960.html>

Performance Evaluation and Analysis of CAS-ESM 2.0

WANG Tian-Yi^{1,2}, JIANG Jin-Rong¹, CHI Xue-Bin¹, ZHANG He³, HE Juan-Xiong³, HAO Hui-Qun¹

¹(High Performance Computing, Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China)

²(University of Chinese Academy of Sciences, Beijing 100049, China)

³(International Center for Climate and Environment Sciences, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China)

Abstract: An earth system model is the important software for researching climate changes and modeling the earth system. Chinese Academy of Sciences-Earth System Model (CAS-ESM) is the high-performance computing software for earth system simulation developed by Institute of Atmospheric Physics (IAP), CAS. Currently, IAP has released CAS-ESM 2.0, while simulation performance is always one of the critical factors restricting its development. To evaluate and analyze the performance of CAS-ESM 2.0, we install CAS-ESM 2.0 to two high-performance computing platforms, “Yuan” and “Earth System Numerical Simulation Device”, for coupled numerical simulation. The experiment results show that CAS-ESM 2.0 has different performances affected by these two platforms; the atmosphere model has the highest proportion of running time, exceeding the sum of other component models; some component models perform poorly in scalability. Further analysis reveals that the atmosphere model is mainly restrained by ineffective communication. Future research and development of CAS-ESM 2.0 will focus on cross-platform optimization,

① 基金项目: 国家重点研发计划 (2016YFB0200800); 中国科学院信息化专项 (XXH13506-302); 国家自然科学基金重点项目 (41931183)

Foundation item: National Key Research and Development Program of China (2016YFB0200800); Special Project of Informatization of Chinese Academy of Sciences (XXH13506-302); Key Program of National Natural Science Foundation of China (41931183)

收稿时间: 2020-10-16; 修改时间: 2020-11-18; 采用时间: 2020-11-24; csa 在线出版时间: 2021-06-01

atmosphere model upgrading, parallel algorithm improvement, and scalability of component models.

Key words: earth system model; high performance computing; CAS-ESM; performance evaluation; coupling experiment

1 引言

随着科技的进步,人类活动与地球环境的联系越来越密切^[1],逐步衍生出了全球气候变化(global change science)这门新兴学科^[2].在早期的全球变化科学研究中,首先提出了气候系统的概念.气候系统主要以地球流体作为研究主体,主要研究大气环流、海洋环流、陆地表层和海冰的物理过程,后来扩展到研究生态与环境系统,形成了地球系统模式^[3].世界上的各个国家和地区如中国、美国、英国、欧盟、日本等国家均已建立了各自的地球系统模式^[4].

中科院地球系统模式 CAS-ESM 是中国科学院大气物理研究所发展的、我国自主研发的地球系统模式,是进行地球科学研究的重要工具. CAS-ESM 2.0 是中科院地球系统模式的最新版本,参加了第 6 次国际耦合模式比较计划(CMIP6)^[5,6].

作为高性能计算应用软件,地球系统模式的性能是制约其发展的重要因素之一. CAS-ESM 2.0 在 CAS-ESM 1.0 的基础上在大气分模式采用了新的自适应滤波与三维剖分并行计算算法,在高纬地区采用了跳点差分计算格式;在海洋分模式也采用了多种性能优化手段,其可扩展性与计算效率比 CAS-ESM 1.0 有了明显的提升.因而有必要对 CAS-ESM 2.0 进行性能评估和分析,发现其目前仍然存在的性能瓶颈,为后续的性能优化工作提供理论支持和发展方向.

本文将 CAS-ESM 2.0 分别部署在“元”和“地球系统数值模拟装置”两大高性能计算平台上,分别开展耦合数值模拟试验,避免性能测试结果受到平台硬件与软件环境的影响.然后对模拟结果进行评估,分析了 CAS-ESM 2.0 的计算速度、可扩展性以及耦合数值模拟中各分模式的表现,找到了 CAS-ESM 2.0 仍然存在的性能瓶颈.

2 CAS-ESM 2.0 简介

CAS-ESM^[7,8] 是中国科学院大气物理研究所(IAP)基于 CESM 1.0^[9] 版本开发的. CAS-ESM 是一个用于模拟地球气候系统的耦合地球系统模式.图 1 是地球系统模式 CAS-ESM 2.0 与各分模式的结构图, CAS-

ESM 主要由 7 个单独的分模式组成,包括大气、海洋、陆面、陆冰、海冰、大气化学分模式和区域分模式. 7 个单独的分模式通过耦合器结合在一起^[10,11].通过对编译运行系统进行配置,建立和组装可以模拟和运行的试验.接下来将对 CAS-ESM 2.0 所使用的耦合器与主要分模式进行简要介绍.

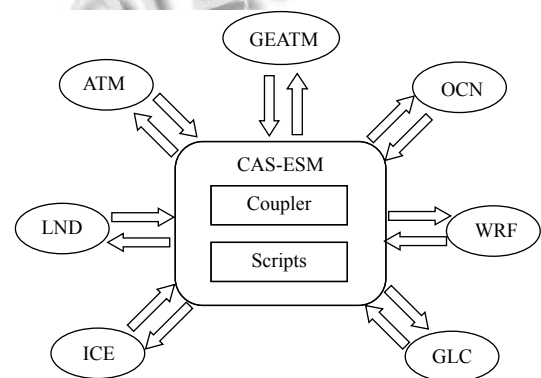


图 1 CAS-ESM 2.0 及各分模式结构

2.1 耦合器与耦合试验

耦合器是地球系统模式的关键组件之一.从图 1 中可以看出,各分模式之间的数据交换主要通过耦合器来完成.由于各分模式模拟所需的数据类型、网格分辨率以及处理器划分方式各不相同,因而在进行耦合数值模拟试验时,需要通过耦合器来完成数据类型转换、网格分辨率的映射以及处理器划分的处理.

地球系统模式 CAS-ESM 使用的耦合器是在 CESM 的 CPL7 耦合器的基础上发展起来的.美国 NCAR 研发的 CPL 耦合器主要应用于 NCAR 提出的“共同气候系统模式(Community Climate System Model, CCSM)”计划中^[12],随着 CCSM 系统的发展壮大,其成为了目前国际上应用最为广泛的耦合器之一.在此基础上, CAS-ESM 对 CPL7 进行了改进,自主研发了适用于大气模式、区域模式与大气化学模式的三维耦合算法^[13].

耦合数值模拟试验通过耦合器对各分模式进行统一调度和管理,根据各分模式的耦合频次按时进行分模式之间的数据传输,从而可以模拟整个地球系统变化的复杂试验.本文完成的耦合数值模拟试验采用了

目前应用最为广泛的1850年工业革命前的初始场数据进行模拟,使用大气模式AGCM、陆面模式CoLM、海洋模式LICOM和海冰模式CICE进行耦合.本文将对这几个分模式进行简要介绍.

2.2 主要分模式简介

2.2.1 大气模式 IAP AGCM5

大气模式 IAP AGCM 是中国科学院大气物理研究所发展的大气环流模式^[14,15], CAS-ESM 2.0 使用的版本为 IAP AGCM5. 与 IAP AGCM4 相比, IAP AGCM5 对动力框架和并行计算进行了改进, 采用了新的自适应滤波与三维剖分并行计算算法, 在高纬地区采用了跳点差分计算格式; 对物理过程参数化方案也做了大量工作, 包括深对流参数化、浅对流参数化^[16]、宏观云物理过程参数化、微观云物理过程参数化^[17], 重力波和湍流等.

2.2.2 海洋模式 LICOM

海洋模式 LICOM 是中国科学院大气物理研究所发展的海洋环流模式^[18,19], CAS-ESM 2.0 使用的版本为 LICOM2.0. 在 CAS-ESM 2.0 中, LICOM 在海气通量交换物理过程上, 提出了盐度边界条件^[20]. LICOM 在国际上首次使用了实盐通量, 认为海盐通量应是虚盐通量和实盐通量之和, 且在虚盐通量中考虑了淡水通量和海表盐度的相关关系, 替代了国际上常用的气候态常数来表征海表盐度.

2.2.3 陆面模式 CoLM

陆面模式 CoLM 是北京师范大学发展的陆面模式^[21]. 在 CAS-ESM 2.0 中, CoLM 引入了多层土壤碳方案, 可以更好地描述土壤碳的累积和分解效应. 另外, 基于风蚀物理过程, 建立了新的起沙过程计算方案, 更好地模拟了全球主要的沙尘分布带.

2.2.4 海冰模式 CICE

海冰模式 CICE 是美国国家大气研究中心 (NCAR) 发展的海冰模式^[22,23], CAS-ESM 2.0 使用的版本为 CICE4.0. 在 CAS-ESM 2.0 中, 对海冰表面反照率参数化方案及海冰模式辐射传输过程方案进行了完善, 模拟结果显示海冰在密度和厚度上都更接近观测数据^[24].

3 性能评估试验与结果分析

3.1 运行环境

本文分别将 CAS-ESM 部署在“元”高性能计算系统和“地球系统数值模拟装置”原型系统两个超级计算平台进行数值模拟试验.

中国科学院高性能计算系统“元”于2014年11月正式启用, 该计算系统目前已有CPU总共5400核, 共有270台曙光CB60-G16双路刀片, 整体峰值性能达120.96 Tflops, 每台刀片计算节点配置2颗Intel E5-2680 V2 (Ivy Bridge | 10C | 2.8 GHz) 处理器, 每颗处理器10核, 64/128 GB DDR3 ECC 1866 MHz 内存, 拥有总容量1.486 PB, 可用容量为1.189 PB的高性能工作区存储系统. 其计算存储网络采用FDR InfiniBand 全线速互联, 每颗处理器的网络带宽为56 GB.

“地球系统数值模拟装置”是峰值计算能力位于国际前列的地球系统数值模拟的专用装置, 在体系结构、计算密度、存储架构、能效比等各方面均处于目前国内领先水平. 其原型系统目前拥有600台计算节点, 整体峰值浮点性能达到1.8 Pflops, 每台计算节点配备2颗我国自主研发的Hygon处理器, 每颗处理器32核心, 主频2 GHz. 每台计算节点配备256 GB DDR4 2666 内存. 其计算存储网络采用Intel OPA 专用计算网络方案, 每颗处理器的网络带宽为100 GB.

3.2 试验设置

地球系统模式 CAS-ESM 是进行气候数值模拟和预测的高性能计算软件系统. 其模拟结果容易受到计算环境和平台的影响, 因而我们分别在高性能计算系统“元”和“地球系统数值模拟装置”原型系统两个高性能计算平台进行数值模拟试验, 测试其计算速度、并行效率、可扩展性等性能指标, 同时相同核数的模拟试验测试5次, 取模拟时间的最小值, 以消除计算环境和平台的影响.

我们首先在高性能计算系统“元”上进行了地球系统模式 CAS-ESM 的耦合数值模拟试验, 大气模式使用 AGCM4.2, 陆面模式使用 CoLM, 海洋模式使用 LICOM, 海冰模式使用 CICE, 其中 AGCM 和 CoLM 的网格分辨率为 1.4 度, LICOM 和 CICE 的网格分辨率为 1 度, 模拟时间为 1 模式月 (31 模式天), 分别使用 16 核、32 核、64 核、128 核、256 核、512 核、1024 开展模拟试验.

然后将 CAS-ESM 移植到“地球系统数值模拟装置”原型系统, 进行跨平台的耦合数值模拟试验, 其他试验设置与“元”上的模拟试验完全一致.

3.3 高性能计算系统“元”模拟试验与结果分析

按照前文中所述的试验设置, 我们在高性能计算系统“元”上开展了 CAS-ESM 的耦合数值模拟试验. 图2为运行时间随核数变化图. 从图2中可以看出, 随

随着核数的增加, CAS-ESM 耦合数值模拟试验的运行时间逐步减少, 到 1024 核为止具有比较好的可扩展性, 需要进行更大规模的试验进一步验证. 但目前受限于 CAS-ESM 2.0 分模式的并行算法设计与本文试验的计算规模, 还无法使用 2048 核进行耦合试验. 然后将运行时间换算为地球系统中常用的模拟速度, 得到图 3 的结果. 从图 3 中可以看出, 随着核数的增加, CAS-ESM 耦合数值模拟速度不断加快, 至 1024 核能达到 15.8 模式年/天, 且模拟速度仍然呈加速趋势. 对于本文使用的模式分辨率以及耦合数值模拟试验来说, 15.8 模式年/天的模拟速度已经可以满足这类试验与研究问题的要求.

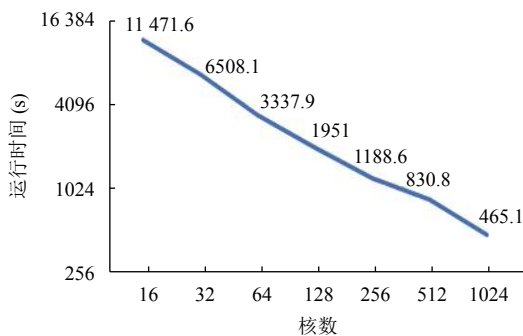


图2 在“元”上的运行时间随核数变化图

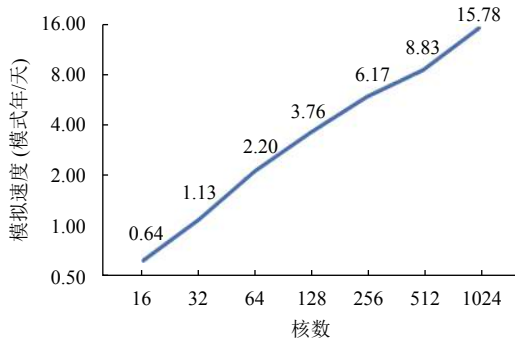


图3 CAS-ESM 在“元”上的模拟速度

为了探究 CAS-ESM 在“元”上进行耦合数值模拟试验的模拟性能变化趋势, 我们制作了如图 4 所示的加速比图. 从图 4 中可以看出, 整体来说 CAS-ESM 2.0 在“元”上的加速比较好, 在达到 512 核时, 加速比有一定程度的下降, 但下降趋势并不明显, 在达到 1024 核时, 加速比又有一定程度的回升, 说明在 512 核时加速比的下降应该是受到了高性能计算平台环境的影响, 试验结果存在一定程度的误差. 在达到 1024 核时, 加速

比可以达到 24.7, 具有良好的加速效果. 去除通信开销影响, 根据 16 核-32 核的加速比变化情况, 理想情况的 1024 核加速比约为 54.4, 与试验结果有一定的差距, 说明通信开销随着核数的增加也在不断增长, 是影响 CAS-ESM 2.0 计算效率、可扩展性的因素之一.

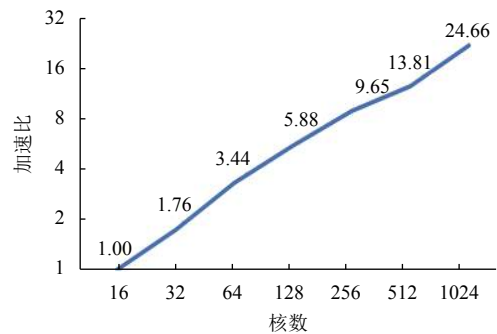


图4 CAS-ESM 在“元”上的加速比变化图

为了进一步探究 CAS-ESM 耦合数值模拟试验中, 各分模式的模拟性能情况, 我们统计了各分模式的运行时间. 表 1 为各分模式的运行时间随核数变化表, 表中的运行时间单位均为秒, 平均占比为该分模式运行时间占总运行时间百分比的平均值. 从整体来看, 大气分模式 AGCM 的平均运行时间所占比重较大, 达到了 78.21%, 超过了其他分模式的运行时间之和. 之后按照平均占比从大到小依次为海冰模式 CICE、海洋模式 LICOM、陆面模式 CoLM. CPL 运行时间为耦合器运行时间, 从表 1 中来看, 除各分模式外, 耦合器也是平均占比较大的模块之一. 陆面模式 CoLM 最多仅支持使用 128 核进行模拟, 因而从表中可以看出, 虽然其平均占比不大, 但使用 1024 核时, 其运行时间占总运行时间的比重可以达到 6.8% 左右, 远高于其平均占比 2.9%, 因而应该提高 CoLM 模式的可扩展性. 另外, 海洋模式 LICOM 与海冰模式 CICE 受限于其计算规模与并行算法, 目前 CAS-ESM 2.0 仅支持最多使用 480 核进行模拟, 因而在核数达到 512 核之后, 运行时间不再下降. 大气分模式随着核数的增加, 其运行时间不断减少, 具有较好的可扩展性.

为了进一步探究 AGCM 分模式运行时间占比较高的原因, 我们使用 1024 核耦合数值模拟试验的结果, 对 IAP AGCM 主要函数进行了运行时间占比情况的统计, 如图 5. 图例中的名称均为大气分模式 AGCM 调用的函数名. 从图 5 中可以看出, IAP AGCM 主要函数中, stepon_run3 函数运行时间占比最多, 达到了 58.13%,

其次是 phys_run1 函数, 达到了 27.72%, phys_run2 函数也能达到 5.38%.

表1 在“元”上各分模式运行时间随核数变化 (单位: s)

模式	16	32	64	128	256	512	1024	平均占比(%)
CPL	741.5	168.3	117.7	94.1	73.4	37	21.9	4.68
CICE	1276.9	217.9	139.7	165.3	57.3	35.2	34	6.21
CoLM	199.5	104.7	56.7	30.8	31.5	33.3	31.6	2.87
AGCM	8066.2	5542.6	2765.6	1500.1	944.1	672.5	334.4	78.21
LICOM	368.1	240.2	124.7	74.8	46.8	36.8	35	4.34
总时间	11471.6	6508.1	3337.9	1951	1188.6	830.8	465.1	

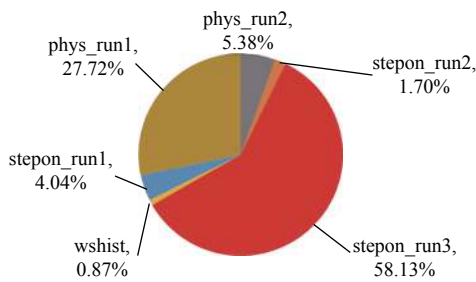


图5 AGCM 主要函数在“元”上运行时间占比情况图

由于 stepon_run3 函数在大气模式 AGCM 的运行时间占比最高, 其运行速度直接影响到大气模式的模拟速度, 甚至影响到整个地球系统模式的模拟速度, 因而有必要对 stepon_run3 函数的模拟结果进行进一步的研究与分析, 发现目前影响 stepon_run3 函数运行速度的关键和原因. 图6是在“元”上 stepon_run3 函数主要调用函数的运行时间占比情况, 图例中的名称为调用的函数名, other 表示除图例所示函数外的其他函数和计算过程.

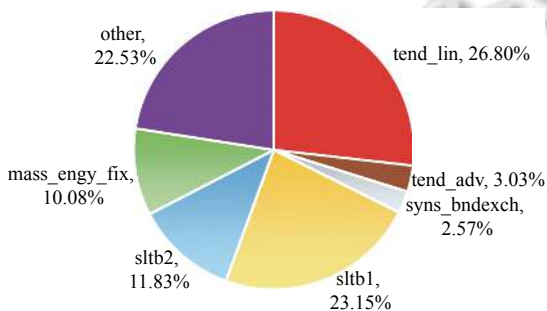


图6 在“元”上 stepon_run3 主要调用函数运行时间占比情况

图6中的 tend_lin、tend_adv、sltb1、sltb2 函数均包含大量的聚合通信, 而 mass_engy_fix 函数和 other 部分也包含大量的点对点通信. 另外 sync_bndexch 为聚合通信 MPI_Barrier 函数的调用. 因而 stepon_run3

函数的运行时间受通信时间、效率的影响较大. 对于目前的 CAS-ESM 2.0 来说, 如果能对大气分模式的 stepon_run3 函数进行并行算法的优化与重新设计, 降低通信时间开销, 应该能大大提高 CAS-ESM 的计算效率与模拟速度.

3.4 “地球系统数值模拟装置”原型系统模拟试验

为了排除高性能计算系统硬件和软件环境的影响, 我们将 CAS-ESM 移植到“地球系统数值模拟装置”原型系统上, 进行了同样的模拟试验. 图7为运行时间随核数变化图. 从图7中可以看出, 随着核数的增加, CAS-ESM 耦合数值模拟试验的运行时间逐步减少; 当核数达到 1024 核的时候, 运行时间有的减少程度并不明显, 说明遇到了性能瓶颈. 然后将运行时间换算为地球系统中常用的模拟速度, 得到图8的结果.

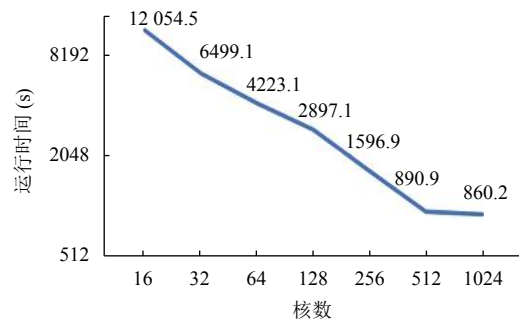


图7 在“地球系统数值模拟装置”上的运行时间随核数变化图

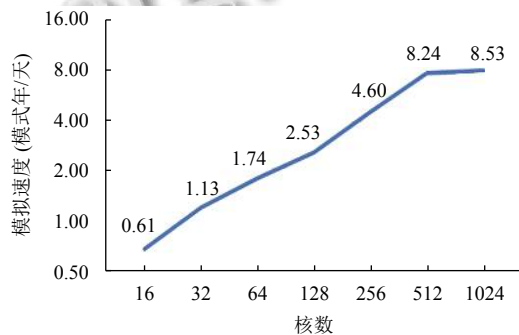


图8 CAS-ESM 在“地球系统数值模拟装置”上的模拟速度

从图8中可以看出, 随着核数的增加, CAS-ESM 耦合数值模拟速度不断加快, 当核数达到 512 核时, 模拟速度能达到 8.2 模式年/天, 当核数达到 1024 核时, 模拟速度能达到 8.5 模式年/天, 其加速效果已不明显. 与“元”上的耦合数值模拟试验结果相比, CAS-ESM 2.0 在“地球系统数值模拟装置”上的最大模拟速度与平均

模拟速度均低于在“元”上的模拟速度.从“地球系统数值模拟装置”的硬件指标来看,尽管其整体峰值浮点性能高于“元”的整体峰值浮点性能,处理器在主频、单核浮点性能上都低于“元”的处理器,这也应该是造成两个平台模拟速度差异的原因.

为了探究 CAS-ESM 在“地球系统数值模拟装置”上进行耦合数值模拟试验的模拟性能变化趋势,我们制作了如图 9 所示的加速比图.从图 9 中可以看出,在达到 32 核时,加速比的增长趋势有一定程度的下降,但下降趋势并不明显;在达到 128 核时,加速比的增长趋势又有一定程度的上升,说明 32 核时出现的变化情况应该是由高性能计算环境造成的误差;在达 1024 核之后,加速比有明显的下降,同样说明遇到了性能瓶颈.达到 1024 核时的加速比约为 14.0,与去除通信开销,根据 16-32 核计算的理想加速比相差较大,说明在“地球系统数值模拟装置”上,通信开销的影响比在“元”上更大,应该进一步的研究与分析.

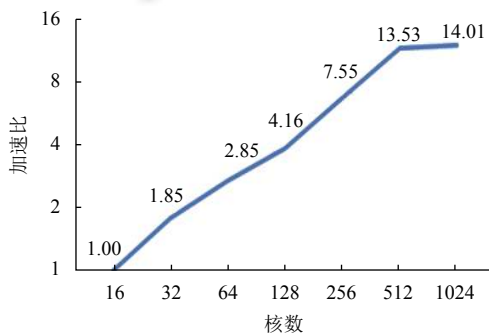


图 9 CAS-ESM 在“地球系统数值模拟装置”上的加速比变化图

为了进一步探究 CAS-ESM 在“地球系统数值模拟装置”上耦合数值模拟试验中,各分模式的模拟性能情况,我们统计了各分模式的运行时间.表 2 为各分模式的运行时间随核数变化表.从整体来看,与在“元”上的模拟试验结果一致,大气分模式 AGCM 的平均运行时间所占比重较大,达到了 81.59%,之后按照平均占比从大到小仍然为海冰模式 CICE、海洋模式 LICOM、陆面模式 CoLM.耦合器 CPL 的运行时间占比情况仍然达到了 3.19%.从各分模式运行时间随核数的变化来看,随着核数的增加,CoLM 模式受限于其可扩展性的问题,达到 128 核之后运行时间无法进一步减少.海洋模式 LICOM 和海冰模式 CICE 受限于其计算规模,最多仅能使用 480 和,因而在达到 512 核之前,其加速

效果较好,但达到 1024 核时,其运行时间已无法继续减少.耦合器和大气模式可扩展性较好,但当核数达到 1024 核时,其加速效果已出现比较明显的下降,且大气模式受影响程度较大.前文已经提到,由于大气模式中存在较多的聚合通信与点对点通信,因而初步判断加速效果的降低是由于通信造成的.

表 2 在“地球系统数值模拟装置”上各分模式运行时间随核数变化(单位: s)

模式	16	32	64	128	256	512	1024	平均占比(%)
CPL	120.5	108.3	122	134.6	84.5	34.2	25.6	3.19%
CICE	341.8	208.5	187	308.3	73.6	44.4	44.6	5.13
CoLM	281.9	142.7	72	36.6	37.1	36.9	33.3	2.55
AGCM	10492.9	5671.4	3512.4	2194.1	1241.4	712.4	689.7	81.59
LICOM	392.5	234.6	163.5	115.7	96.6	46.9	56.3	4.66
总时间	12054.5	6499.1	4223.1	2897.1	1596.9	890.9	860.2	

由于大气模式运行占比最高,且其达到 1024 核时加速效果下降明显,为了进一步探究 AGCM 分模式运行时间占比较高和加速效果下降的原因,我们使用“地球系统数值模拟装置”上 1024 核耦合数值模拟试验的结果,对 IAP AGCM 主要函数进行了运行时间占比情况的统计,如图 10.从图 10 中可以看出,与在“元”上的试验结果一致,IAP AGCM 主要函数中,仍然是 stepon_run3 函数运行时间占比最多,达到了 69.65%,其次仍然是 phys_run1 函数,达到了 15.31%,但与“元”上的模拟结果相比,phys_run1 函数的运行占比有所下降,而 stepon_run3 函数的运行占比有所上升,其他函数的运行占比变化不大.因而很可能是由于 stepon3 函数导致了“地球系统数值模拟装置”与“元”上的大气模式 AGCM 运行时间、加速比与可扩展性的差异.

为了探究在“地球系统数值模拟装置”上大气模式 AGCM 运行时间较长、可扩展性存在问题的原因.我们使用 1024 核的模拟结果对 stepon_run3 函数进一步统计和分析,图 11 是 stepon_run3 调用的其他函数运行时间占 stepon_run3 运行时间的比重.与“元”相比,stepon_run3 函数的主要调用函数运行占比情况基本一致,因而应该是 stepon_run3 调用的函数整体出现了计算时间的增加.影响 stepon_run3 函数效率的主要因素仍然是通信,而影响通信的直接因素就是网络带宽.从两个高性能计算平台的网络带宽来看,在“元”上每颗处理器 20 核,配备网络带宽为 56 GB,其单核网络带宽为 5.6 GB;在“地球系统数值模拟装置”上,每颗处理

器 32 核, 配备网络带宽 100 GB, 其单核网络带宽为 3.1 GB. 因此在两个高性能计算平台上, CAS-ESM 表现出的可扩展性差异很可能是由于网络带宽不同造成的.

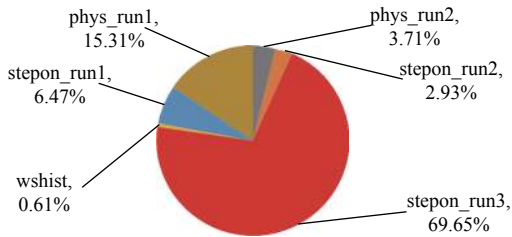


图 10 AGCM 主要函数在“地球系统数值模拟装置”上的运行时间占比情况图

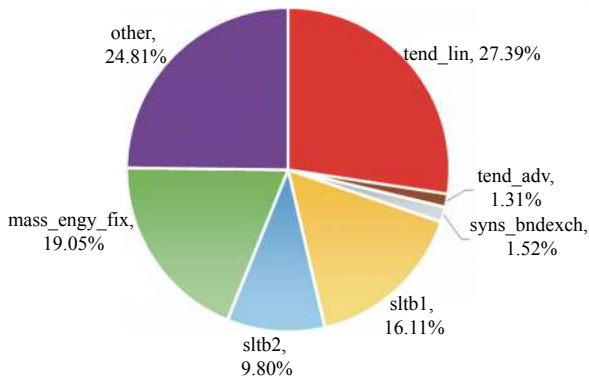


图 11 Stepon_run3 主要调用函数运行时间占比情况图

3.5 试验小结

前文使用 CAS-ESM 在中科院高性能计算系统“元”和“地球系统数值模拟装置”原型系统上分别进行了数值模拟试验并对实验结果做了简要分析. 通过分析发现, CoLM 模式和 LICOM 模式的可扩展性, 大气模式 IAP AGCM 的 stepon_run3 函数的通信开销, 是影响 CAS-ESM 在两个平台上模拟性能的主要因素. 其中通过提高 CoLM 模式的可扩展性, 根据表 1 和表 2 中的平均时间占比, 在理想情况下可以使 CAS-ESM 在 1024 核的模拟时间减少约 3.9%; 提高 LICOM 模式的可扩展性, 在理想情况下可以使 CAS-ESM 在 1024 核的模拟时间减少约 2.3%; IAP AGCM 的 stepon_run3 函数模拟时间占总模拟时间的比例达到约 45%, 如果采用各类算法优化手段, 减少 stepon_run3 函数的通信开销, 在理想情况下应该能起到最明显的加速效果.

根据以往高分辨率大气模式与海洋模式的优化经验, 对 stepon_run3 函数的通信优化, 同样可以使用高纬度的纬向跳点差分计算、三维剖分并行算法等优化

手段^[25-28]. 本文使用的 1.4 度分辨率的 IAP AGCM 仍然使用了二维剖分并行算法, 如果增加纬向 (X 方向) 的剖分, 可以提升 IAP AGCM 的并行度, 减少各进程负责的数据大小^[29], 从而降低 stepon_run3 函数的计算量与通信量. 另外, 由于 stepon_run3 函数存在大量的垂直方向 (Z 方向) 通信, 可以通过优化三维剖分算法, 尽可能使垂直方向数据位于同一处理器上, 从而大大减少通信量, 提高模拟效率.

4 结论与展望

CAS-ESM 地球系统模式是一个进行地球系统模拟的复杂的高性能应用软件. 为了对 CAS-ESM 2.0 进行性能评估, 我们分别在中科院高性能计算系统“元”和“地球系统数值模拟装置”原型系统上开展了 CAS-ESM 的耦合数值模拟试验.

结果显示, 在两个高性能计算平台上, CAS-ESM 表现出了不同的可扩展性. CAS-ESM 在“元”上的模拟在相同核数的绝对速度与可扩展性上, 都要优于在“地球系统数值模拟装置”的模拟结果. 由于高性能计算平台的硬件指标中, 单核浮点性能、主频以及单核网络带宽, “元”都要优于“地球系统数值模拟装置”, 因而模拟结果也符合计算环境的实际情况. 在后续的 CAS-ESM 发展、地球系统研究与气候模拟工作中, 应该选择单核浮点性能、主频较高, 网络通信带宽更高的高性能计算平台开展耦合试验. 在最新的 2020 中国高性能计算机 top10 榜单中, 第 2 名的“天河二号”超级计算机的单核浮点性能最高, 达到了 235 Gflops; 第 4 名的同方“青城之光”单核主频最高, 达到了 3.1 GHz; 第 5 名的联想深腾 8800 系列的网络通信带宽最大, 达到了单核 1.25 GB. 上述 3 个平台应该是今后较为适合 CAS-ESM 进行研究与试验的高性能计算平台.

另外综合两个平台耦合数值模拟试验的结果显示, 大气模式的运行时间占总运行时间的比重都是最高的, 且超过了其他模式的总和. 因而我们对结果进行了深入分析, 发现大气模式主要函数中, stepon_run3 函数在两个平台的模拟结果中, 都是运行占比最多的函数. 因而在“地球系统数值模拟装置”上继续对 stepon_run3 函数进行了深入挖掘和分析, 发现其主要的性能瓶颈是由于通信造成的. 因而在后续的 CAS-ESM 研发工作中, 需要对大气模式尤其是 stepon_run3 函数中的并行通信算法开展进一步的性能优化工作.

在各分模式的可扩展性方面,目前陆面模式 CoLM、海洋模式 LICOM 与海冰模式 CICE 都存在一定的问题,为了满足耦合模式的进一步性能优化需要,提高这些分模式的可扩展性也是研究的重点之一.本文采用的大气分模式垂直方向为 30 层,无法进行 2048 核及以上更大规模的耦合数值模拟试验.而高分辨率的分模式,垂直方向 35 层的大气模式,模拟结果还存在一定的问题.为了进行更大规模的耦合数值模拟试验,在后续工作中应该对分模式的并行算法进行优化和重新设计,使其可以使用更多的处理器,进行更大规模的性能测试与分析.

参考文献

- 1 曾庆存,林朝晖.地球系统动力学模式和模拟研究的进展.地球科学进展,2010,25(1):1-6.
- 2 叶笃正,符淙斌,董文杰,等.全球变化科学领域的若干研究进展.大气科学,2003,27(4):435-450.
- 3 Meehl GA, Covey C, Delworth T, *et al.* THE WCRP CMIP3 multimodel dataset: A new era in climate change research. Bulletin of the American Meteorological Society, 2007, 88(9): 1383-1394. [doi: 10.1175/BAMS-88-9-1383]
- 4 王斌,周天军,俞永强,等.地球系统模式发展展望.气象学报,2008,66(6):857-869.
- 5 周天军,邹立维,陈晓龙.第六次国际耦合模式比较计划(CMIP6)评述.气候变化研究进展,2019,15(5):445-456. [doi: 10.12006/j.issn.1673-1719.2019.193]
- 6 Eyring V, Bony S, Meehl GA, *et al.* Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geoscientific Model Development, 2016, 9(5): 1937-1958. [doi: 10.5194/gmd-9-1937-2016]
- 7 孙泓川,周广庆,曾庆存. IAP 第四代大气环流模式的气候系统模式模拟性能评估.大气科学,2012,36(2):215-233.
- 8 Dong X, Su TH, Wang J, *et al.* Decadal variation of the Aleutian low-icelandic low seesaw simulated by a climate system model (CAS-ESM-C). Atmospheric and Oceanic Science Letters, 2014, 7(2): 110-114. [doi: 10.1080/16742834.2014.11447144]
- 9 Vertenstein M, Craig T, Middleton A, *et al.* CESM1.0. 4 User's Guide. USA: NCAR, 2011.
- 10 Lin ZH, Yu Z, Zhang H, *et al.* Quantifying the attribution of model bias in simulating summer hot days in China with IAP AGCM 4.1. Atmospheric and Oceanic Science Letters, 2016, 9(6): 436-442. [doi: 10.1080/16742834.2016.1232585]
- 11 Adeniyi MO, Lin Z, Zhang H. Evaluation of the performance of IAP-AGCM4.1 in simulating the climate of West Africa. Theoretical and Applied Climatology, 2019, 136(3-4): 1419-1434.
- 12 辛晓歌,吴统文,张洁,等. BCC 模式及其开展的 CMIP6 试验介绍.气候变化研究进展,2019,15(5):533-539. [doi: 10.12006/j.issn.1673-1719.2019.039]
- 13 Wang YZ, Jiang JR, He JX. Coupling methods of global climate models and regional climate models. High Technology Letters, 2017, 23(1): 90-95.
- 14 Zhang H, Zhang MH, Zeng QC. Sensitivity of simulated climate to two atmospheric models: Interpretation of differences between dry models and moist models. Monthly Weather Review, 2013, 141(5): 1558-1576. [doi: 10.1175/mwr-d-11-00367.1]
- 15 Zhang H, Lin Z H, Zeng Q C. The computational scheme and the test for dynamical framework of IAP AGCM-4. Chinese Journal of Atmospheric Sciences, 2009, 33(6): 1267-1285.
- 16 Li YY, Zhang MH. The role of shallow convection over the Tibetan Plateau. Journal of Climate, 2017, 30(15): 5791-5803. [doi: 10.1175/JCLI-D-16-0599.1]
- 17 Xie XN, Zhang H, Liu XD, *et al.* Role of microphysical parameterizations with droplet relative dispersion in IAP AGCM 4.1. Advances in Atmospheric Sciences, 2018, 35(2): 248-259. [doi: 10.1007/s00376-017-7083-5]
- 18 Liu HL, Lin PF, Yu YQ, *et al.* The baseline evaluation of LASG/IAP climate system ocean model (LICOM) version 2. Acta Meteorologica Sinica, 2012, 26(3): 318-329. [doi: 10.1007/s13351-012-0305-y]
- 19 Huang WY, Wang B, Yu YQ, *et al.* Improvements in LICOM2. Part I: Vertical mixing. Journal of Atmospheric and Oceanic Technology, 2014, 31(2): 531-544. [doi: 10.1175/JTECH-D-13-00065.1]
- 20 Jin JB, Zeng QC, Wu L, *et al.* Formulation of a new ocean salinity boundary condition and impact on the simulated climate of an oceanic general circulation model. Science China Earth Sciences, 2017, 60(3): 491-500. [doi: 10.1007/s11430-016-9004-4]
- 21 Meng CL, Li ZL, Zhan X, *et al.* Land surface temperature data assimilation and its impact on evapotranspiration estimates from the Common Land Model. Water Resources Research, 2009, 45(2): W02421.
- 22 Hunke EC, Hebert DA, Lecomte O. Level-ice melt ponds in the Los Alamos sea ice model, CICE. Ocean Modelling, 2013, 71: 26-42. [doi: 10.1016/j.ocemod.2012.11.008]
- 23 Urrego-Blanco JR, Urban NM, Hunke EC, *et al.* Uncertainty

- quantification and global sensitivity analysis of the Los Alamos sea ice model. *Journal of Geophysical Research: Oceans*, 2016, 121(4): 2709–2732. [doi: [10.1002/2015JC011558](https://doi.org/10.1002/2015JC011558)]
- 24 Song MR, Liu JP. The role of diminishing Arctic sea ice in increased winter snowfall over northern high-latitude continents in a warming environment. *Acta Oceanologica Sinica*, 2017, 36(8): 34–41. [doi: [10.1007/s13131-017-1021-3](https://doi.org/10.1007/s13131-017-1021-3)]
- 25 Wang YZ, Jiang JR, Zhang H, *et al.* A scalable parallel algorithm for atmospheric general circulation models on a multi-core cluster. *Future Generation Computer Systems*, 2017, 72: 1–10. [doi: [10.1016/j.future.2017.02.008](https://doi.org/10.1016/j.future.2017.02.008)]
- 26 Xiao JM, Li SG, Wu BD, *et al.* Communication-avoiding for dynamical core of atmospheric general circulation model. *Proceedings of the 47th International Conference on Parallel Processing*. Eugene, OR, USA. 2018. 1–10.
- 27 Wu BD, Li SG, Cao H, *et al.* AGCM3D: A highly scalable finite-difference dynamical core of atmospheric general circulation model based on 3D decomposition. 2018 IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS). Singapore. 2018. 355–364.
- 28 Cao H, Yuan L, Zhang H, *et al.* A highly efficient dynamical core of atmospheric general circulation model based on Leap-format. 2020 IEEE International Parallel and Distributed Processing Symposium (IPDPS). New Orleans, LA, USA. 2020. 95–104.
- 29 周广庆, 张云泉, 姜金荣, 等. 地球系统模式 CAS-ESM. *数据与计算发展前沿*, 2020, 2(1): 38–54.