



中华人民共和国气象行业标准

QX/T 148—2020
代替 QX/T 148—2011

气象领域高性能计算机系统 测试与评估规范

Specification for high performance computer system
test and evaluation in the meteorological field

2020-07-31 发布

2020-12-01 实施

中 国 气 象 局 发 布

目 次

前言	III
引言	IV
1 范围	1
2 术语和定义	1
3 测试规范	1
3.1 测试要求	1
3.2 测试内容	2
3.3 测试准备	3
3.4 测试结果	4
3.5 测试说明	4
4 评估规范	5
4.1 定性评估	5
4.2 定量评估	6
4.3 测试评估报告	8
附录 A(资料性附录) 高性能计算机系统总体测试说明	9
附录 B(资料性附录) 高性能计算机系统分项测试说明	21
附录 C(资料性附录) 高性能计算机系统测试评分示例	25
附录 D(资料性附录) 高性能计算机系统测试评估报告大纲	28

前 言

本标准按照 GB/T 1.1—2009 给出的规则起草。

本标准代替 QX/T 148—2011《气象领域高性能计算机系统测试与评估规范》。与 QX/T 148—2011 相比,除编辑性修改外,主要技术变化如下:

- 修改了标准适用范围,由“采购”扩展为“采购或租用”(见第 1 章,2011 年版的第 1 章);
- 修改了高性能计算机系统术语定义(见 2.1,2011 年版的 2.1);
- 增加了加速比术语和定义(见 2.3);
- 将核心测试术语改为基准测试术语(见 2.5,2011 年版的 2.4);
- 删除了再现性测试术语和定义(见 2011 年版的 2.7);
- 增加了输入输出(I/O)及消息传递接口(MPI)通信性能测试(见 3.2.3);
- 删除了断点/重起和分时调度功能测试(见 2011 年版的 3.2.3);
- 增加了作业管理调度等功能测试(见 3.2.4);
- 增加了合理性评估指标(见 4.1.4);
- 删除了管理软件效率分析(见 2011 年版的 4.2.5);
- 增加了定量评估方法(见 4.2.5);
- 修改了《高性能计算机系统总体测试说明》(参见附录 A,2011 年版的附录 A)、《高性能计算机系统分项测试说明》(参见附录 B,2011 年版的附录 B、C),增加了《高性能计算机系统测试评分示例》(参见附录 C)、《高性能计算机系统测试评估报告大纲》(参见附录 D)。

本标准由全国气象基本信息标准化技术委员会(SAC/TC 346)提出并归口。

本标准起草单位:国家气象信息中心。

本标准主要起草人:魏敏、孙婧、沈瑜、李娟、肖华东、王彬、洪文董、曹燕、田浩。

本标准所代替标准的历次版本发布情况为:

- QX/T 148—2011。

引 言

QX/T 148—2011《气象领域高性能计算机系统测试与评估规范》规定了气象领域高性能计算机在采购过程中的测试要求、内容和方法,以及对测试结果的评估方法。随着气象数值模式及高性能计算技术的发展,气象数值模式对高性能计算机系统的需求不断变化,高性能计算机系统的规模不断扩展,系统复杂性不断增加,高性能计算机系统测试结果评估需求进一步精细化。

为更全面、规范地对高性能计算机系统进行测试与评估,特对 QX/T 148—2011《气象领域高性能计算机系统测试与评估规范》进行修订。

气象领域高性能计算机系统测试与评估规范

1 范围

本标准规定了气象行业运行气象数值模式的高性能计算机系统的测试与评估规范。
本标准适用于气象行业采购或租用运行气象数值模式的高性能计算机系统的测试与评估。

2 术语和定义

下列术语和定义适用于本文件。

2.1

高性能计算机系统 high performance computer system; HPCS

由一定数量高性能计算节点、高速低延迟互连网络和大容量存储子系统及配套软件构成的,以科学与工程计算为主要应用目标的大规模并行计算机系统。

2.2

峰值性能 peak performance

高性能计算机系统的最高理论性能值。

2.3

加速比 speedup ratio

给定气象数值模式程序在高性能计算机系统单节点或处理器核(CPU核)上的运行时间与在多个这种节点或处理器核(CPU核)上的运行时间之比。

2.4

应用测试 application test

使用实际业务或科研气象数值模式程序对高性能计算机系统进行的测试。

2.5

基准测试 benchmark test

选择国际上通用、公开的基准测试程序,测试高性能计算机系统的相关性能。

2.6

非优化测试 un-optimized test

不对程序源代码进行运算性能优化型修改,并可得到合理结果的测试。

2.7

优化测试 optimized test

对程序源代码进行运算性能优化型修改,并可得到合理结果的测试。

3 测试规范

3.1 测试要求

3.1.1 筛选气象数值模式

应从已有的业务或科研气象数值模式中筛选出稳定运行的具有计算、访存、通信及 I/O 等特点的

模式。

3.1.2 设定测试时效要求

根据业务或科研气象数值模式的运行时效要求来设定模式测试的运行时间要求,也可用缩短运行时间的要求来测更大规模的系统。

3.1.3 测试用例

应对选定的气象数值模式程序设置约束条件,确定测试数据及计算规模,形成测试用例。

3.1.4 测试用机

应使用一套具有完整硬件配置、完整软件配置,且配置已全部安装的独立的高性能计算机为测试用机。测试用机应是 UNIX 或 LINUX 环境和 64 位精度及以上的高性能计算机系统,测试节点的计算单元应满配置,内存容量可按需配置。

3.1.5 综合测试

应采用应用测试、基准测试和功能测试等方法,对高性能计算机系统的计算性能、内存性能、内部互联网络性能、I/O 性能、可靠性及软件功能等进行全面测试。对于租用的远程访问的高性能计算机系统应对访问带宽和访问性能等进行测试。

3.1.6 测试方法

应以应用测试为主、基准测试为辅。应用测试可分非优化和优化两种方式。各项测试均应由厂商或服务商自测试。

3.1.7 运行方式

所有测试相关的脚本应以批作业的方式运行。

3.2 测试内容

3.2.1 测试题目

一道测试题目(Test)可对应一个气象数值模式程序的测试或对应某个基准程序的测试,也可对应多个气象数值模式程序的组合测试或多个基准程序的组合测试。测试题目以 Test1,Test2,Test3 等顺序编号。

3.2.2 应用测试

3.2.2.1 系统配置测试

在规定时间内运行完给定气象数值模式程序所需的 CPU 核资源、内存资源等最小配置。

3.2.2.2 加速比测试

加速比测试中节点数或 CPU 核数系列值的选取可根据采购或租用系统的规模调整,宜至少选取 8 个阶梯,尽量接近或大于等于采购或租用系统的节点数或 CPU 核数。

3.2.2.3 性能测试

性能测试应包含:

- 单 CPU 核应用性能测试；
- 单节点应用性能测试；
- 多节点应用性能测试。

3.2.3 基准测试

基准测试应包含：

- I/O 性能测试；
- MPI 通信性能测试；
- 内存带宽测试。

3.2.4 功能测试

功能测试可包含：

- 作业管理调度测试；
- 软件开发工具测试；
- 科学计算库测试。

3.3 测试准备

3.3.1 程序、数据及相关文档准备

包括以下内容：

- a) 气象数值模式程序及相关数据准备应在 UNIX 或 LINUX 操作系统环境下进行，生成的目录结构应以测试程序名开始，下级目录为源程序、数据；

示例：

全球与区域同化预报系统 (GRAPES) 程序目录结构：

GRAPES/src/

GRAPES/data/

- b) 可在存放气象数值模式程序及相关数据主目录下用 tar 命令打包、压缩，并以该程序名命名；

示例：

GRAPES 程序文件名：

GRAPES.tar.gz 或 GRAPES.tar.Z

- c) 应准备《高性能计算机系统总体测试说明》《高性能计算机系统分项测试说明》等文档；

- d) 基准测试程序、相关数据与文档可由厂商或服务商自行获取。

3.3.2 介质准备

包括以下步骤：

- a) 介质选择：存储介质宜选轻便易传递的移动硬盘或光盘。
- b) 数据记录：把在 UNIX 或 LINUX 环境下准备的气象数值模式程序、数据及文档传输到移动存储介质。存储介质制作之后应进行可用性读出检查。
- c) 贴标签：存储介质制作完成后，应统一对每件介质编号、贴标签。标志样例可包括以下 3 行：
- 编号：单位名称缩写名-1(1/4)；
 - 模式名称：GRAPES.tar.Z；
 - 制作日期：××××年××月××日。

制作完成后应加密封标志，加盖公章。

3.4 测试结果

3.4.1 输出结果

包括以下步骤：

a) 筛选输出结果。应包括以下内容：

- 测试过程所使用的所有源程序、头文件、库文件、目标文件、可执行文件、输入输出数据、配置脚本、作业提交脚本、标准输出、标准错误输出及日志文件等；
- 优化时修改过的最终版本的源文件，不应包括优化过程中的中间文件；
- 应对作业主要部分的开始和结束打上墙钟时间标志；
- 测试的结果文件。

b) 存放路径与打包：在程序测试运行过程中所使用的气象数值模式目录层次不变，为区分不同厂商或服务商各测试题目非优化和优化结果，应在原模式目录层次增加若干目录。可按示例的目录结构存放，打包返回结果。

示例：

厂商或服务商 A 的目录结构：

A/Test1/GRAPES/un-optimized/

A/Test1/GRAPES/optimized/

上述路径表明是厂商或服务商 A 基于 GRAPES 进行的 Test1 测试结果，分为优化和非优化两种。

可打包成 A-Test1-GRAPES.tar.gz 文件，表明是厂商或服务商 A 完成的 Test1 测试结果，采用 tar 命令打包并压缩后的文件。

3.4.2 填写测试报告表

各测试题目应分别对应一份测试报告表，内容包含测试结果和相应的测试环境。由采购或租用单位提供，由测试人填写并签字，标明日期，厂商或服务商盖章，用纸质形式和电子文档(PDF 或 WORD 格式)提交。

3.4.3 撰写分析报告

应包括对测试的总体描述，测试环境和关键系统参数调整对性能的影响，测试题目自身存在的问题，优化过程中主要修改部分，测试结果分析等。基于部分实测结果进行推算的测试题目，其结果分析中应给出推算理由、推算方法、公式和说明。

应包含保证结果的真实性和测试的可再现性，推算结果视同承诺的文字。

应使用中文，用纸质形式和电子文档(PDF 或 WORD 格式)提交，纸质形式文档应有撰写人签字，并标明日期，厂商或服务商盖章。

3.5 测试说明

3.5.1 总体测试说明

应准备一份总体测试说明，说明的书写应清楚、严谨。说明的大纲宜包含：概述、说明文档与程序、测试要求、测试结果、测试题目和联系方式等。总体测试说明参见附录 A。

3.5.2 分项测试说明

分项测试说明可包括但不限于以下内容：

a) 气象数值模式名称，主要参数，程序目录结构等；

- b) 编译、链接源程序,修改文件参数,运行程序的方法等;
- c) 模式输入、输出数据存放目录及其文件集,数据格式等;
- d) 模式运行结果,结果合理性检查方法等。

分项测试说明参见附录 B。

4 评估规范

4.1 定性评估

4.1.1 完成情况

应对所有厂商或服务商完成的测试及结果提交情况进行统计,填写测试完成情况定性评估表,见表 1。该表应涵盖所有参测厂商或服务商和所有测试题目,表项可按实际情况增减。

表 1 测试完成情况定性评估表

测试题目		厂商或服务商			
		测试完成情况	测试方法正确性	测试结果合理性	优化百分比 %
Test1	非优化				—
	优化				
测试完成情况:√ 完成测试,○未完成测试; 测试方法正确性:√ 正确,× 错误,? 部分正确,○ 未完成测试; 测试结果合理性:√ 合理,× 不合理,? 部分合理,○ 未完成测试; 优化百分比:优化后相对优化前计算性能提升百分比,○ 未完成测试。					

4.1.2 测试环境分析

对测试环境进行分析,至少应包括以下内容:

- a) 测试系统是 UNIX 或 LINUX;
- b) 测试系统精度是 64 位或 64 位以上;
- c) 测试系统 CPU、节点、内部网络互联、存储系统等硬件配置与采购或租用系统差异;
- d) 测试系统操作系统、编译器、作业管理调度等软件配置与采购或租用系统差异。

4.1.3 测试方法正确性分析

对测试方法进行分析,至少应包括以下内容:

- a) 测试使用采购或租用方提供(或指定来源)的源代码;
- b) 测试使用采购或租用方提供(或指定来源)的数据;
- c) 非优化测试没有对程序源代码进行运算性能优化型修改;
- d) 优化测试真正做过源代码优化,使用的预编译器、编译选项及链接库;
- e) 没有通过减少输出数据而减少墙钟时间等;
- f) 功能测试按规定的方法测试;
- g) 测试的真伪性,检查批作业运行的标准输出、运行结果等;
- h) 测试的完成程度,测试结果是全部实测、全部推算或部分实测部分推算。

4.1.4 测试结果合理性分析

4.1.4.1 模式运行结果的合理性

测试过程中,因修改了部分程序代码(例如在优化方式测试时)或调整了气象数值模式的参数,可能造成模式运行结果的差异或预报预测的错误。可通过运算结果与标准结果的典型气象要素场分布形势和相关系数等模式模拟结果合理性指标来检验。

4.1.4.2 输出结果被人为修改的检验

气象数值模式运行过程中,标准输出、标准错误输出等文件输出时具有关联的时间标志序列。当输出结果(如敏感的运行时间等)被人为修改,其文件的时间属性会发生变化。可通过对模式运行输出的系列文件的创建、修改等时间关联性来检查。

4.1.5 问题清单

应在问题清单上列出妨碍定量评估的问题,以测试题目为单位,逐个厂商或服务商分析登记,见表2。该表应涵盖所有参测厂商或服务商和所有测试题目,表项可按实际情况增减。

表 2 测试结果问题清单

测试题目	厂商或服务商
Test1	

4.2 定量评估

4.2.1 总则

表3至表5中各表项可按实际情况增减。

4.2.2 配置分析

对涉及测量最小配置规模的测试题目,宜统计分析CPU核数(或CPU数、或节点数)、单CPU核峰值性能及内存配置和总功耗等内容,计算得到峰值性能,内存容量和每瓦峰值性能,见表3。

表 3 最小系统配置统计表

测试题目	厂商或服务商							
	运行时间 s	CPU核数	单CPU核 峰值性能 GFLOPS	峰值性能 GFLOPS	单CPU核 内存配置 GB	内存容量 GB	总功耗 W	性能功耗比 GFLOPS/W
Test1	非优化							
	优化							
峰值性能=单CPU核峰值性能×CPU核数; 内存容量=单CPU核内存配置×CPU核数; 性能功耗比=峰值性能/总功耗。								

4.2.3 测量时间分析

对涉及测量时间的测试题目,应统计分析作业使用的编译选项和运行时间等内容,见表4。

表4 运行时间统计表

测试题目		厂商或服务商	
		编译选项	运行时间 s
Test1	非优化		
	优化		

4.2.4 加速比分析

对涉及加速比的测试题目,应统计分析不同并行规模作业运行时间,计算加速比,见表5。表中计算单元可为节点或CPU核,根据实际情况确定。

表5 加速比统计表

测试题目		厂商或服务商		
		运行时间 s	峰值性能 GFLOPS	加速比
CPU核	128			
	256			
	512			
	1024			
	2048			
	4096			
	8192			
	16384			

4.2.5 评分方法

每厂商或服务商测试满分可分为测试总体情况分值和测试题目分值两部分,见式(1)。

$$S = S_s + S_t \dots\dots\dots(1)$$

式中:

S ——测试满分, $S \geq 0$;

S_s ——测试总体情况满分, $S_s \geq 0$;

S_t ——测试题目满分, $S_t \geq 0$ 。

测试总体情况应主要考虑测试厂商或服务商是否按照要求完成测试及提交文档情况,可分为“优”“良”“一般”3个级别,为3个级别设定分值,按此顺序分值依次递减,“优”级别分值最高,“一般”级别分值最低。根据厂商或服务商实际测试情况给出测试总体情况得分 $S_{sa} (0 \leq S_{sa} \leq S_s)$ 。

测试题目得分应为厂商或服务商各测试题目总得分,见式(2)。

$$S_{ta} = \sum_{i=1}^n (S_{bi} + (S_{mi} - S_{bi}) \times R_i) \dots\dots\dots(2)$$

式中:

- S_{ta} —— 测试题目总得分, $0 \leq S_{ta} \leq S_t$;
- n —— 测试题目总数, $n > 0$ 的整数;
- i —— 测试题目个数, $i \leq n$ 的正整数;
- S_{bi} —— 第 i 题目基础分, $0 \leq S_{bi} \leq S_{mi}$;
- S_{mi} —— 第 i 题目满分, $S_{mi} > 0$ 且 $\sum_{i=1}^n S_{mi} = S_t$;
- R_i —— 第 i 题目得分因子, $0 \leq R_i \leq 1$ 。

测试总得分应为厂商或服务商测试总体情况得分与测试题目总得分之和,见式(3)。

$$S_a = S_{sa} + S_{ta} \dots\dots\dots(3)$$

式中:

- S_a —— 测试总得分, $0 \leq S_a \leq S$;
- S_{sa} —— 测试总体情况得分, $0 \leq S_{sa} \leq S_s$ 。

按要求完成测试即可获得该题目基础分。测试未完成、测试不符合测试要求或测试结果不合理该测试题目不得分。

各部分分值分配,基础分分值设置及得分因子的计算方法应根据实际测试需求及测试内容确定。测试评分示例参见附录 C。

4.3 测试评估报告

测试评估报告应包括概述、定性评估、定量评估和总体评估,但不限于这些内容。测试评估报告大纲参见附录 D。

附 录 A
(资料性附录)
高性能计算机系统总体测试说明

A.1 概述

测试分为应用测试、基准测试和功能测试三类。参与测试的厂商或服务商依据测试说明按照要求完成各项测试,并提交测试结果及测试文档。采购或租用单位提供气象数值模式源程序、数据和文档。基准测试的测试程序和数据由厂商或服务商自行获取。所有测试工作由参测厂商或服务商独立完成。测试文档采用中文。

A.2 说明文档、程序及数据

A.2.1 说明文档

包括测试整体情况说明及各气象数值模式程序说明:

- a) 总体测试说明;
- b) GRAPES 模式程序说明;
- c) 天气研究与预报模式(WRF)程序说明;
- d) 北京气候中心气候系统模式(BCC_CSM)程序说明;
- e) 北京气候中心大气环流模式(BCC_AGCM)程序说明。

A.2.2 程序及数据

包括测试所用各气象数值模式程序及数据:

- a) Test1-a-GRAPES_GFS. tar. gz; GRAPES 全球模式;
- b) Test1-b-GRAPES_MESO. tar. gz; GRAPES 区域模式;
- c) Test1-c-WRF. tar. gz; WRF 模式;
- d) Test1-d-BCC_CSM. tar. gz; BCC_CSM 模式;
- e) Test1-e-BCC_AGCM. tar. gz; BCC_AGCM 模式。

A.3 测试要求

所有测试在与厂商或服务商即将提供的采购或租用系统相同或相近的系统配置规模环境下运行,一般是 UNIX 或 LINUX 环境和 64 位或以上精度的机器。

应用测试应采用采购或租用单位提供的气象数值模式源程序和数据。所有输入和输出数据应存储于全局共享文件系统。程序运行时应以单作业方式占用节点。所有测试相关的脚本应以批作业的方式运行,不用交互方式,如只能用交互方式运行,应说明原因。

应用测试可分为非优化和优化两种方式,应提交非优化测试结果,否则此测试题目视为无效。可对应用测试题目进行优化,同时提交优化测试结果。具体如下:

- a) 非优化测试代码修改应遵循下列条件:
 - 1) 不影响模式的预报预测结果;
 - 2) 不影响代码的通用性,修改的代码重新运行能产生再现的结果,即使用完全相同的输入数

据和参数,作业重新运行能得到二进制位逐位相同的结果;

- 3) 各模式的参数修改限制参见各自的说明文件;
 - 4) 在代码和脚本文件中所做修改的地方,标注 UNOPTMOD 字符串。
- b) 优化测试代码修改应遵循上述非优化测试 1)–3) 条件与以下条件:
- 1) 编译前源代码可使用预编译器;
 - 2) 为指导编译器完成某种功能而在源代码中插入编译指导语句;
 - 3) 各模式说明文件中指明的可修改部分;
 - 4) 所做的代码修改不能影响程序的可读性和代码的可移植性;
 - 5) 在代码和脚本文件中做修改的地方,标注 OPTMOD 字符串,并提供汇集所有修改内容及修改目的的描述性电子文档。

应用测试未标明“实测”的要求,可进行推算,推算应在实测结果基础上进行,推算结果视为正式承诺,对于推算的结果,应给出具体推算方法和公式及相关必要说明。功能测试应实测,不允许推算。

对未按照测试要求完成的测试题目,即使提交了测试结果,该测试题目视为未完成。气象数值模式各测试结果应具有合理性,采购或租用单位对测试结果进行合理性检查。

厂商或服务商应确保提交的所有数据、文档内容真实、可靠。各项测试结果应在采购或租用系统均可再现。

测试结果数据应采用移动存储设备存储,该设备应能在指定环境识别,需提供硬件连接方式及系统挂载命令。如测试结果数据不能够被正确读取,则视为未提交数据。

A.4 测试结果

A.4.1 输出结果

A.4.1.1 内容

应包括但不限于以下内容:

- a) 实际作业运行时所使用的源程序和编译后对应的目标程序,库函数,可执行文件,编译链接选项配置文件,参数文件,数据文件及日志文件等;
- b) 优化及非优化程序源代码及编译链接选项配置文件;
- c) 提交作业脚本和作业主要部分的开始和结束墙钟时间日志;
- d) 运行过程中各进程的标准输出日志、标准错误输出日志和作业运行信息;
- e) 每道测试题目每个并行规模测试结果文件。

A.4.1.2 文件说明

为保证测试结果数据的正确读取,应对每一道测试题目的每一个测试结果文件说明其存放目录、文件格式(二进制文件应明确大端或小端存储)及读取方法,对各测试结果数据来源进行明确说明。

A.4.1.3 存放路径与打包

在测试过程中所使用的气象数值模式目录层次不变,为区分不同厂商或服务商各测试题目非优化和优化结果,应在原模式目录层次增加若干目录。可按下述目录结构存放,以厂商或服务商 A 为例:

厂商或服务商 A 的目录结构:

A/Test1/GRAPES/un-optimized/

A/Test1/GRAPES/optimized/

上述路径表明是厂商或服务商 A 基于 GRAPES 进行的 Test1 测试结果,分成优化和非优化两种。

打包成 A-Test1-GRAPES.tar.gz 返回结果,这表明是厂商或服务商 A 完成的 Test1 测试结果。如果文件太大,可分成几个包,如 A-Test1-GRAPES-1.tar.gz, A-Test1-GRAPES-2.tar.gz 等。

A.4.2 填写测试报告表

各测试题目分别对应一份测试报告表,包含测试结果和相应的测试环境。由测试人填写并签字,标明日期,厂商或服务商盖章,用纸质形式和电子文档提交。

A.4.3 撰写分析报告

A.4.3.1 概述

主要包括对测试情况的总体描述,运行环境和关键系统参数调整对性能的影响,优化过程中主要修改部分,结果的性能分析等。用纸质形式和电子文档提供,纸质形式文档要有撰写人签字,厂商或服务商盖章。

厂商或服务商应保证测试结果的真实性和测试的可再现性,明确说明推算结果作为正式承诺。

A.4.3.2 测试系统信息

厂商或服务商应针对每道测试题目提交相关测试系统的详细信息。

A.4.3.3 测试结果分析

应包括但不限于以下内容:

- a) 详细说明测试系统与采购或租用系统的区别;
- b) 比较分析所有优化代码和原始代码的差别;
- c) 详细说明使用到的脚本、命令和提供的各类文档;
- d) 详细说明使用到的优化方法、工具及软件等;
- e) 分析测试结果,包括图形、表格和文字表述等;
- f) 对测试结果进行推算的测试题目,给出实测结果、推算理由、推算公式和推算可行性分析。

A.5 测试题目

测试题目包括 Test1、Test2 和 Test3。

Test1 为应用测试,可包括但不限于表 A.1 的内容。

Test2 为基准测试,可包括但不限于表 A.2 的内容。

Test3 为作业管理调度功能测试,可包括但不限于表 A.3 的内容。

表 A.1 应用测试题目

序号	测试题目	测试名称	测试要求
1	Test1-a	GRAPES_GFS 模式测试	完成分辨率为 0.25 度的 8 d 预报,4096 核实测所需最短时间,同时应推算同等并行规模在采购或租用系统的最短时间。
2	Test1-b	GRAPES_MESO 模式测试	完成分辨率为 0.03 度的 1 d 预报,8192 核实测所需最短时间,同时应推算同等并行规模在采购或租用系统的最短时间。

表 A.1 应用测试题目(续)

序号	测试题目	测试名称	测试要求
3	Test1-c	WRF 模式测试	完成分辨率为 3 km 的 12 h 预报,每小时一次输出,每次输出单独一个文件。应提供 128、256、512、1024、2048、4096、8192、16384 核的最短时间,建议测试到尽可能多的核数,并尽量提供密集测试结果。同时应推算同等并行规模在采购或租用系统的最短时间。
4	Test1-d	BCC_CSM 模式测试	在 1 d 内完成 10 个模式年的积分,至少使用 512 核测试所需的最短时间,要求至少输出连续积分 5 个模式年的实测结果数据。同时应推算在采购或租用系统的最短时间。
5	Test1-e	BCC_AGCM 模式测试	在 1 d 内完成 5 个模式年的积分,至少使用 512 核测试所需的最短时间,要求至少输出连续积分 1 个模式年的实测结果数据。同时应推算在采购或租用系统的最短时间。
注:“核”指处理器核(CPU 核)。			

表 A.2 基准测试题目

序号	测试题目	测试名称	测试要求
1	Test2-a	I/O 性能测试 (IOzone)	测试要求实测,针对系统 I/O 通路设计,选择具有代表性的计算节点(如通路最长,中等,最短),依次在每节点采用单进程完成测试。给出节点选择依据,填写测试表格。应推算各测试结果在采购或租用系统的对应结果。
			测试要求实测,针对系统 I/O 通路设计,分别选择 64、128、256、512 进程(每进程运行在一个核),每节点满配运行,并发执行测试。给出节点选择依据,填写测试表格。应推算各测试结果在采购或租用系统的对应结果。
2	Test2-b	MPI 通信性能测试 (IMB)	测试要求实测,针对系统计算网络通信拓扑结构,完成 MPI 各类点到点通信测试,包括一个节点内、跳数为 1、跳数为 3 直至跳数最大节点等,测试 MPI 带宽和延迟。给出节点选择依据,填写测试表格。应推算各测试结果在采购或租用系统的对应结果。
			测试要求实测,针对系统计算网络通信拓扑结构,至少对 64、128、256、512 进程(每进程运行在一个核),每节点满配运行,完成 MPI 各类集合通信测试,测试 MPI 的带宽和延迟。给出节点选择依据,填写测试表格。应推算各测试结果在采购或租用系统的对应结果。
3	Test2-c	内存性能测试 (Stream)	测试要求实测,系统每节点配置 n 个核,完成 $1/n/4/n/2/n$ 个核规模测试,测试所用数组的大小分别设为所用到最后一级 cache(如 L3)总和的 2、4、6、8 倍,给出详细说明,填写测试表格。应推算各测试结果在采购或租用系统的对应结果。
注:“核”指处理器核(CPU 核)。			

表 A.3 作业管理调度功能测试题目

序号	测试题目	测试名称	测试要求
1	Test3-a	高优先级作业抢占功能测试(采用 Test1-a 中 GRAPES_GFS 模式和 Test1-b 中 GRAPES_MESO 模式)	提交低优先级的 GRAPES_GFS 模式作业之后,提交高优先级的 GRAPES_MESO 模式作业抢占 GRAPES_GFS 模式作业计算资源,GRAPES_GFS 模式作业被挂起(保留在内存),GRAPES_MESO 模式作业运行结束后,GRAPES_GFS 模式作业可正常继续运行。
2	Test3-b		高优先级作业抢占低优先级作业资源,作业被抢占后能够自动重新提交的功能。
3	Test3-c	多级作业抢占功能测试(采用 Test1-a 中 GRAPES_GFS 模式和 Test1-b 中 GRAPES_MESO 模式)	测试不同优先级队列的作业按照级别进行计算资源挂起方式抢占的功能。 如有 A、B、C 三个队列,优先级顺序为 $A > B > C$ 。多个研发作业同时运行(C 队列),当较高级别的作业提交(B 队列)后,在资源不满足时抢占研发作业(C 队列)的资源,当最高级别作业提交(A 队列)后,在资源不满足时可以抢占较高级别的作业(B 队列)和研发作业(C 队列)的资源,目标是保证最高优先级作业的时效性。在高级别作业结束后,低级别作业恢复运行。

A.6 联系方式

分别给出不同测试题目联系人信息,包括姓名、联系电话及邮箱等。

A.7 测试报告表

图 A.1 给出系统环境配置表格式,可包括但不限于表中内容。图 A.2—图 A.6 给出 Test1-a 至 Test1-e 测试报告表格式,可包括但不限于表中内容。

管理节点	服务器型号		
	主板芯片组型号		
	CPU	型号	
		数量	
	内存	型号	
		数量	
	磁盘	型号	
		数量	
	计算网络接口	型号	
		数量	
	管理网络接口	型号	
		数量	

图 A.1 系统环境配置表

登录节点	服务器型号		
	主板芯片组型号		
	CPU	型号	
		数量	
	内存	型号	
		数量	
	磁盘	型号	
		数量	
	计算网络接口	型号	
		数量	
管理网络接口	型号		
	数量		
计算节点	服务器型号		
	主板芯片组型号		
	CPU	型号	
		数量	
	内存	型号	
		数量	
	磁盘	型号	
		数量	
	计算网络接口	型号	
		数量	
管理网络接口	型号		
	数量		
存储节点	服务器型号		
	主板芯片组型号		
	CPU	型号	
		数量	
	内存	型号	
		数量	
	磁盘	型号	
		数量	
	计算网络接口	型号	
		数量	
管理网络接口	型号		
	数量		

图 A.1 系统环境配置表(续)

计算网络	网络类型	
	网络拓扑	
	阻塞比	
	任意两个节点最远跳数	
	核心交换机型号	
	核心交换机数量	
管理网络	网络类型	
	网络拓扑	
	核心交换机型号	
	核心交换机数量	
存储网络	网络类型	
	网络拓扑	
	核心交换机型号	
	核心交换机数量	
监控网络	网络类型	
	网络拓扑	
	核心交换机型号	
	核心交换机数量	
软件	操作系统版本	
	操作系统内核版本	
	并行文件系统版本	
	作业调度系统版本	
	集群管理软件版本	
	系统监控软件版本	
	并行环境版本	
	编译调试环境版本	
	数学函数库版本	
散热方式	计算节点	
	管理节点	
	登录节点	
	存储节点	
	存储系统	
	网络交换机	
能耗	功耗	
	能效	
测试系统安装地点		
测试系统安装时间		

图 A.1 系统环境配置表(续)

报告表 1: Test1-a 测试结果

第 1 部分: Test1-a 运行时间

测试题目	测试系统			采购或租用系统	
	采用处理器核 (CPU 核)数 (理论峰值)	非优化实测时间 (s)	优化实测时间 (s)	采用处理器核 (CPU 核)数 (理论峰值)	推算时间 (s)
GRAPES_GFS					

第 2 部分: Test1-a 运行环境

机型	
CPU	
标量时钟/向量时钟(GHz)	
每时钟指令数	
单处理器核(CPU 核)峰值性能(GFLOPS)	
每 CPU 核数	
单 CPU 峰值性能(GFLOPS)	
总节点数	
每节点 CPU 数	
缓存与寄存器大小	
每节点实际使用内存(GB)	
每节点内存配置(GB)	
互联网速(GB/s)	
互联网硬件延迟	
操作系统版本	
编译器版本	
编译命令及选项	

厂商或服务商: _____

测试者姓名: _____

测试者签名: _____

时间: _____

图 A.2 Test1-a 测试报告表

报告表 2: Test1-b 测试结果

第 1 部分: Test1-b 运行时间

测试题目	测试系统			采购或租用系统	
	采用处理器核 (CPU 核)数 (理论峰值)	非优化实测时间 (s)	优化实测时间 (s)	采用处理器核 (CPU 核)数 (理论峰值)	推算时间 (s)
GRAPES_MESO					

第 2 部分: Test1-b 运行环境

机型	
CPU	
标量时钟/向量时钟(GHz)	
每时钟指令数	
单处理器核(CPU 核)峰值性能(GFLOPS)	
每 CPU 核数	
单 CPU 峰值性能(GFLOPS)	
总节点数	
每节点 CPU 数	
缓存与寄存器大小	
每节点实际使用内存(GB)	
每节点内存配置(GB)	
互联网速(GB/s)	
互联网硬件延迟	
操作系统版本	
编译器版本	
编译命令及选项	

厂商或服务商: _____

测试者姓名: _____

测试者签名: _____

时间: _____

图 A.3 Test1-b 测试报告表

报告表 3: Test1-c 测试结果
第 1 部分: Test1-c 加速比

采用处理器核 (CPU 核)数	测试系统			采购或租用系统		
	理论峰值 (GFLOPS)	运行时间 (s)	加速比	理论峰值 (GFLOPS)	运行时间 (s)	加速比
128						
256						
512						
1024						
2048						
4096						
8192						
16384						

第 2 部分: Test1-c 运行环境

机型	
CPU	
标量时钟/向量时钟(GHz)	
每时钟指令数	
单处理器核(CPU 核)峰值性能(GFLOPS)	
每 CPU 核数	
单 CPU 峰值性能(GFLOPS)	
总节点数	
每节点 CPU 数	
缓存与寄存器大小	
每节点实际使用内存(GB)	
每节点内存配置(GB)	
互联网速(GB/s)	
互联网硬件延迟	
操作系统版本	
编译器版本	
编译命令及选项	

厂商或服务商: _____

测试者姓名: _____

测试者签名: _____

时间: _____

图 A.4 Test1-c 测试报告表

报告表 4: Test1-d 测试结果
第 1 部分: Test1-d 运行时间

BCC_CSM							
非优化实测	总处理器核(CPU 核)数 (理论峰值)	处理器核(CPU 核)数最优配置					非优化实测时间 (s)
		atm	ocn	lnd	ice	cpl	
优化实测	总处理器核(CPU 核)数 (理论峰值)	处理器核(CPU 核)数最优配置					优化实测时间 (s)
		atm	ocn	lnd	ice	cpl	
测试系统 推算	总处理器核(CPU 核)数 (理论峰值)	处理器核(CPU 核)数最优配置					测试系统推算时间 (s)
		atm	ocn	lnd	ice	cpl	
采购或租用 系统推算	总处理器核(CPU 核)数 (理论峰值)	处理器核(CPU 核)数最优配置(可以不填)					采购或租用系统推算时间 (s)
		atm	ocn	lnd	ice	cpl	

第 2 部分: Test1-d 运行环境

机型	
CPU	
标量时钟/向量时钟(GHz)	
每时钟指令数	
单处理器核(CPU 核)峰值性能(GFLOPS)	
每 CPU 核数	
单 CPU 峰值性能(GFLOPS)	
总节点数	
每节点 CPU 数	
缓存与寄存器大小	
每节点实际使用内存(GB)	
每节点内存配置(GB)	
互联网速(GB/s)	
互联网硬件延迟	
操作系统版本	
编译器版本	
编译命令及选项	

厂商或服务商: _____
 测试者姓名: _____
 测试者签名: _____
 时间: _____

图 A.5 Test1-d 测试报告表

报告表 5: Test1-e 测试结果

第 1 部分: Test1-e 运行时间

测试题目	测试系统				采购或租用系统	
	采用处理器核 (CPU 核)数 (理论峰值)	非优化实测时间 (s)	优化实测时间 (s)	推算时间 (s)	采用处理器核 (CPU 核)数 (理论峰值)	推算时间 (s)
BCC_AGCM						

第 2 部分: Test1-e 运行环境

机型	
CPU	
标量时钟/向量时钟(GHz)	
每时钟指令数	
单处理器核(CPU 核)峰值性能(GFLOPS)	
每 CPU 核数	
单 CPU 峰值性能(GFLOPS)	
总节点数	
每节点 CPU 数	
缓存与寄存器大小	
每节点实际使用内存(GB)	
每节点内存配置(GB)	
互联网速(GB/s)	
互联网硬件延迟	
操作系统版本	
编译器版本	
编译命令及选项	

厂商或服务商: _____

测试者姓名: _____

测试者签名: _____

时间: _____

图 A.6 Test1-e 测试报告表

附 录 B
(资料性附录)
高性能计算机系统分项测试说明

B.1 GRAPES 模式

B.1.1 概述

GRAPES 模式测试包括全球与区域预报模式测试。PPI 是全球预报模式的并行驱动,支持单、双精度。

B.1.2 全球预报模式

B.1.2.1 编译 PPI

进入 PPI/src 目录;根据需要修改 Makefile 文件;执行 make 命令编译 PPI。

B.1.2.2 编译模式

进入 MODEL/src 目录;根据需要修改 configure、grapes_r4、INTEL 文件、physics/CoLM/build/Makefile 文件和 physics/RRTMG/build/Makefile 文件;执行 compile.sh INTEL r4 命令编译模式;编译成功后在 MODEL/RUN 目录下生成 grapes_global_4.exe 可执行文件。

B.1.2.3 运行模式

进入 RUN_0.25 目录;链接可执行文件 ln-s ../RUN/grapes_global_4.exe./grapes_global.exe;修改 namelist.input 文件的 nproc_x 和 nproc_y 参数,二者相乘与 MPI 进程数相等;运行 grapes_global.exe。

B.1.2.4 数据文件说明

全球预报模式数据为大端二进制数据。

B.1.2.5 模式运行结果

返回全部源程序、编译选项文件、计算结果文件 modelvar2013050106_2304 和日志文件 std.out.0000。结果文件为单精度二进制流文件。测试运行时间以 std.out.0000 中最后的模式运行时间为准,即以“in main model use $\times\times\times\times.\times\times\times\times\times\times\times\times$ second”一行为准。

modelvar2013050106_2304-org 为标准计算结果,是大端二进制流文件, std.out.0000-org 为标准日志文件。编译运行 cor.f,读取 modelvar2013050106_2304-org 和 modelvar2013050106_2304,选取其中部分信息计算相关系数,若相关系数 ≥ 0.98 为通过。可根据实际需要对程序进行修改,如读取文件名、文件路径、大小端等,但计算的要素场不可变更。

B.1.3 区域预报模式

B.1.3.1 编译和运行步骤

进入 grapes_model-no-rsl-new 目录;根据需要修改 config/configure、grapes、INTEL.p 文件;执行 ./comp.sh INTEL 命令编译模式,编译成功后在 run 目录下生成 grapes.exe 可执行文件;运行

grapes.exe。

B.1.3.2 数据文件说明

有限区域预报模式数据为小端二进制数据。

B.1.3.3 模式运行结果

返回全部源程序、编译选项文件、计算结果文件 `rmf.hgra2016052300024.grb2` 和日志文件。结果文件为双精度二进制流文件。测试运行时间以作业日志文件输出的“all_time=××××××”为准。

`rmf.hgra2016052300024.grb2-from-yh` 为标准计算结果, `rsurm-xxxxxx.out-2048-2880` 为标准日志文件。编译运行 `cor.f`, 读取 `rmf.hgra2016052300024.grb2-from-yh` 和 `rmf.hgra2016052300024.grb2`, 选取其中部分信息计算相关系数, 若相关系数 ≥ 0.98 为通过。可根据实际需要对程序进行修改, 如读取文件名、文件路径、大小端等, 但计算的要素场不可变更。

B.2 WRF 模式

B.2.1 概述

WRF 模式水平分辨率采用 3 km, 水平范围基本覆盖中国地区。

B.2.2 编译和运行步骤

进入 WRFV3 目录; 设置 `NETCDF`、`WFIO_NCD_LARGE_FILE_SUPPORT` 环境变量; 修改 `configure.wrf` 配置文件; 执行 `configure` 命令配置模式; 执行 `compile em_real 1 > log.txt 2 > &.1` 命令编译模式, 编译成功后在 `main` 目录下生成 `real.exe` 和 `wrf.exe` 两个可执行文件; 进入 `WRFV3/test/em_real` 目录, 执行 `llsubmit wrf.cmd` 命令提交作业。

B.2.3 数据文件说明

模式初始场文件名为 `wrfinput_d01`, 边界条件文件名为 `wrfbdy_d01`, 均为 `NETCDF` 格式。

B.2.4 模式运行结果

模式预报结果时间由 2011_06_13_00 时到 2011_06_13_12 时, 预报时效共 12 h, 模式预报结果文件格式为 `NETCDF` 格式。返回 WRF V3.3 的完整目录, 包含所有源程序、编译选项文件、目标文件、库函数、可执行文件及模式每次运行的 `namelist.input`、`namelist.output`、`rsl.out.*`、`rsl.error.*`、作业提交脚本、执行信息、预报结果和日志等文件。运行时间以日志文件 `rsl.out.0000` 文件尾部的“Analysis spending time: ××××× seconds”为准。

用于对比结果的文件为 `Test1-c-WRF-result.tar.gz`。编辑 `grads` 数据描述文件 `control.ctl` 和 `experiment.ctl`。进入 `control.ctl` 和 `experiment.ctl` 所在目录, 在 `grads` 环境中运行 `pressure.gs` 脚本程序, 计算两个结果文件中的部分场(p+pb)的相关统计量, 生成 `pressure.gif` 文件, 若相关系数 ≥ 0.98 为通过。可根据实际需要对程序进行修改, 如读取文件名、文件路径、大小端等, 但计算的要素场不可变更。

B.3 BCC_CSM 模式

B.3.1 概述

北京气候中心气候系统模式 BCC_CSM 包含 5 个主要模块: 大气分量模式、海洋分量模式、陆面分

量模式、海冰分量模式和耦合器。

B.3.2 编译和运行步骤

BCC_CSM 模式编译设置在各个分量模式的脚本文件中,编译参数在 models/bld/目录的 Makefile 文件中指定。修改 BCC_CSM/bccesm. build. csh 脚本,设置 MODEL_PATH、DATA_PATH 和 WORK_PATH 参数,执行 bccesm. build. csh 脚本编译模式。修改 BCC_CSM/bccesm. submit. csh 脚本,设置 MODEL_PATH、DATA_PATH 和 WORK_PATH 参数,执行 bccesm. submit. csh 脚本提交作业。

B.3.3 数据文件说明

模式运行所需要的数据文件在 BCC_CSM/INIDATA/目录,atm、ocn、lnd、ice 和 cpl 子目录分别存放大气、海洋、陆面、海冰和耦合器需要的数据,数据为 NETCDF 格式或 ASCII 文本格式。

B.3.4 模式运行结果

模式运行输出结果存放在 output 目录,后缀为“.err”和“.out”的文件记录模式的部分输出信息。atm、ocn、lnd、ice 和 cpl 子目录分别存放各分量模式的输出结果,每个目录下文件名中包含“.out”和“.log.”的文件为输出的日志文件。模式输出结果需要返回的数据如表 B.1 所示。

表 B.1 BCC_CSM 模式输出结果需要返回数据

大气	海洋	陆面	海冰	耦合器
timing. * 文件	无	无	无	无
至少连续积分 5 年月平均模式结果(*h0*.nc)	至少连续积分 5 年模拟结果(默认月平均)	无	无	至少连续积分 5 年模拟结果(默认月平均)
参数文件	参数文件	参数文件	参数文件	参数文件
日志文件	日志文件	日志文件	日志文件	日志文件
目标文件	目标文件	目标文件	目标文件	目标文件

atm 子目录文件名中包含“.h0.”的 NETCDF 文件为大气模式的主要运行结果,还需要返回 atm. stdin、timing. *、atm. log. * 等文件。ocn 子目录返回 ocean_****_***.nc、*_table、*.out、*.log. * 文件。测试运行时间以大气模式分量下文件 timing. 0 中总的模式运行时间为准,即以“total ×××××××× ×××× ××××”一行中的 Wallclock 时间为准。

用于对比结果的相关文件存放于压缩文件 Test1-d-BCC_CSM-result. tar. gz 中,可主要对比分量模式 atm 的运行结果。执行 plot_test1-d. ncl 脚本,读取标准结果和测试结果,选取其中部分信息计算相关系数,若相关系数 ≥ 0.98 为通过。可根据实际需要对程序进行修改,如读取文件名、文件路径等,但计算的要素场不可变更。

B.4 BCC_AGCM 模式

B.4.1 概述

北京气候中心大气环流模式 BCC_AGCM 包含大气和陆面两个分量模式。

B.4.2 编译和运行步骤

修改 test 目录 t266.csh 脚本,设置 INC_NETCDF、LIB_NETCDF、CSMDATA、camroot 环境变量,设置 case、runtype、wrkdir、blkdir、rundir 参数,执行 gmake 命令编译,编译过程中在与个例有关的运行目录下建立子目录 bld,在其中生成可执行文件 bccam。在运行子目录 rundir 下生成 namelist 文件,执行 llsubmit bccam.cmd 命令提交作业。

B.4.3 数据文件说明

模式运行所需要的数据在 inputdata 目录,数据格式以 NETCDF 为主。

B.4.4 模式运行结果

模式运行输出结果存放在 output 目录,后缀为“.err”和“.out”的文件记录模式的部分输出信息。文件名中包含“.log.”的文件为输出的日志文件。模式输出结果需要返回的数据如表 B.2 所示。

表 B.2 BCC_AGCM 模式输出结果需要返回数据

timing.* 文件
至少连续积分 1 年的月平均结果(*h0*.nc)
参数文件
日志文件
目标文件

文件名中包含“.h0.”的 NETCDF 格式文件为主要运行结果,还需要返回 atm.stdin、timing.*、atm.log.* 等文件。测试运行时间以文件 timing.0 中总的模式运行时间为准,即以“total ××××××××××××××××”一行中的 Wallclock 时间为准。

用于对比结果的相关文件存放于压缩文件 Test1-e-BCC_AGCM-result.tar.gz 中。执行 plot_test1-e.ncl 脚本,读取标准结果和测试结果,选取其中部分信息计算相关系数,若相关系数 ≥ 0.98 为通过。可根据实际需要程序进行修改,如读取文件名、文件路径等,但计算的要素场不可变更。

B.5 I/O 性能测试

IOZone 测试程序可由 <http://www.iozone.org> 网址获得。

B.6 MPI 通信性能测试

IMB 测试程序可由 <https://software.intel.com/en-us/articles/intel-mpi-benchmarks> 网址获得。

B.7 内存带宽测试

STREAM 测试程序可由 <http://www.cs.virginia.edu/stream/> 网址获得。

附 录 C
(资料性附录)

高性能计算机系统测试评分示例

C.1 概述

结合附录 A 中表 A.1 至表 A.3 测试题目相关内容,假设有 A、B 和 C 三家厂商或服务商参加测试,对分值权重设置及评分方法进行举例说明。

C.2 分值设置示例

测试满分 100 分,包括测试总体情况分值和测试题目分值两部分。依据测试的重要程度,设定权重与分值。如表 C.1 所示。

表 C.1 分值设置

类别			权重/%	分值			
				基础分	满分		
测试总体情况分值 S_0			3	—	3		
测试 题目 分值 S_i	应用测试	Test1-a	测试系统结果	9	1	9	
			采购或租用系统推算结果	9	1	9	
		Test1-b	测试系统结果	8	1	8	
			采购或租用系统推算结果	8	1	8	
		Test1-c	测试系统结果	7	1	7	
			采购或租用系统推算结果	7	1	7	
		Test1-d	测试系统结果	7	1	7	
			采购或租用系统推算结果	9	1	9	
		Test1-e	测试系统结果	7	1	7	
			采购或租用系统推算结果	9	1	9	
		合计			80	10	80
		基准测试	Test2-a	测试系统结果	2	0.5	2
	采购或租用系统推算结果			2	0.5	2	
	Test2-b		测试系统结果	2	0.5	2	
			采购或租用系统推算结果	2	0.5	2	
	Test2-c		测试系统结果	2	0.5	2	
			采购或租用系统推算结果	2	0.5	2	
	合计			12	3	12	
	作业管理调度 功能测试		Test3-a	测试系统结果	1	0	1
			Test3-b	测试系统结果	2	0	2
Test3-c		测试系统结果	2	0	2		
合计			5	0	5		
总计			100	13	100		

C.3 测试总体情况评分示例

测试总体情况满分 3 分,参照 4.2.5 评分方法,分“优”“良”“一般”3 个级别,“优”为 3 分,“良”为 2 分,“一般”为 1 分。根据测试总体情况,厂商或服务商 A 完成情况最好得 3 分,厂商或服务商 B 完成情况一般得 1 分,厂商或服务商 C 完成情况较好得 2 分。如表 C.2 所示。

表 C.2 测试总体情况评分表

类别	分值	厂商或服务商 A	厂商或服务商 B	厂商或服务商 C
优	3	√	—	—
良	2	—	—	√
一般	1	—	√	—

C.4 测试题目评分示例

以应用测试测试题目 Test1-a 为例,参照 4.2.5 评分方法,计算厂商或服务商 A、B 和 C 该项测试题目得分。

对厂商或服务商 A、B 和 C 的 Test1-a 的测试结果进行分类统计。如表 C.3 所示。

表 C.3 Test1-a 测试结果统计表

类别	厂商或服务商 A		厂商或服务商 B		厂商或服务商 C	
	CPU 核数	最短时间/s	CPU 核数	最短时间/s	CPU 核数	最短时间/s
测试系统结果	4096	913.35	4096	896.98	4096	940.97
采购或租用系统 推算结果	4096	802.45	4096	809.21	4096	865.19

根据采购或租用需求,对 Test1-a 测试题目设定得分因子,见式(C.1)。

$$R = \frac{T_{\min}}{T} \dots\dots\dots(C.1)$$

式中:

R —— Test1-a 测试题目得分因子;

T_{\min} —— Test1-a 测试题目所有厂商或服务商某类测试结果中的最短时间;

T —— 某厂商或服务商 Test1-a 测试题目与 T_{\min} 同类测试结果的时间。

测试系统结果中,厂商或服务商 B 时间最短,即 $T_{\min} = 896.98$,

依据 4.2.5 中公式(2)可得

厂商或服务商 A Test1-a 测试题目测试系统结果得分:

$$1 + (9 - 1) \times (896.98 / 913.35) = 8.857;$$

厂商或服务商 B Test1-a 测试题目测试系统结果得分:

$$1 + (9 - 1) \times (896.98 / 896.98) = 9;$$

厂商或服务商 C Test1-a 测试题目测试系统结果得分:

$$1 + (9 - 1) \times (896.98 / 940.97) = 8.626。$$

采购或租用系统推算结果中,厂商或服务商 A 时间最短,即 $T_{\min} = 802.45$,

厂商或服务商 A Test1-a 测试题目采购或租用系统推算结果得分:

$$1 + (9 - 1) \times (802.45 / 802.45) = 9;$$

厂商或服务商 B Test1-a 测试题目采购或租用系统推算结果得分:

$$1 + (9 - 1) \times (802.45 / 809.21) = 8.933;$$

厂商或服务商 C Test1-a 测试题目采购或租用系统推算结果得分:

$$1 + (9 - 1) \times (802.45 / 865.19) = 8.420。$$

各厂商或服务商 Test1-a 测试系统结果与采购或租用系统推算结果之和即为该厂商或服务商 Test1-a 测试题目得分。

$$\text{厂商或服务商 A Test1-a 测试题目得分: } 8.857 + 9 = 17.857;$$

$$\text{厂商或服务商 B Test1-a 测试题目得分: } 9 + 8.933 = 17.933;$$

$$\text{厂商或服务商 C Test1-a 测试题目得分: } 8.626 + 8.420 = 17.046。$$

以此类推,可得厂商或服务商各测试题目得分,厂商或服务商各测试题目得分之和即为该厂商或服务商测试题目总得分,该厂商或服务商测试总体情况得分与测试题目总得分之和即为测试总得分。

附 录 D
(资料性附录)

高性能计算机系统测试评估报告大纲

D.1 概述

测试情况概要性描述,包括参测厂商或服务商、分析评估的测试题目等。

D.2 定性评估

每道测试题目测试结果的定性分析评估。

D.3 定量评估

每道测试题目测试结果的定性分析评估。

D.4 总体评估

综合各测试题目、各厂商或服务商的评估结果得出总体评估。

中华人民共和国
气象行业标准
气象领域高性能计算机系统测试与评估规范
QX/T 148—2020

*

气象出版社出版发行
北京市海淀区中关村南大街46号
邮政编码:100081
网址:<http://www.qxcbs.com>
发行部:010-68408042
北京建宏印刷有限公司印刷

*

开本:880 mm×1230 mm 1/16 印张:2.25 字数:67.5千字
2020年8月第1版 2020年8月第1次印刷

*

书号:135029-6175 定价:45.00元

如有印装差错 由本社发行部调换
版权专有 侵权必究
举报电话:(010)68406301