

团 体 标 准

T/QGCML XXXX—XXXX

轻量级分布式计算系统

Lightweight distributed computing system

(征求意见稿)

XXXX - XX - XX 发布

XXXX - XX - XX 实施

目 次

前言	II
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 要求	1
5 特点	2
6 模块功能	2
7 运行测试	3

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由全国城市工业品贸易中心联合会提出并归口。

本文件起草单位：

本文件主要起草人：

轻量级分布式计算系统

1 范围

本文件规定了轻量级分布式计算系统的术语和定义、要求、特点、模块功能、运行测试。
本文件适用于轻量级分布式计算系统的设计及应用。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 20270 信息安全技术 网络基础安全技术要求

GB/T 20988 信息安全技术 信息系统灾难恢复规范

GB/T 22239 信息安全技术 网络安全等级保护基本要求

GB/T 25000.10 系统与软件工程 系统与软件质量要求和评价（SQuaRE） 第10部分：系统与软件质量模型

GB/T 25000.51 系统与软件工程 系统与软件质量要求和评价（SQuaRE） 第51部分：就绪可用软件产品（RUSP）的质量要求和测试细则

YD/T 5238 分布式数据计算平台设备安装工程设计规范

3 术语和定义

YD/T 5238界定的以及下列术语和定义适用于本文件。

3.1

轻量级分布式计算系统 `lightweight distributed computing system`

轻量级分布式计算系统是一种通过网络连接多个计算机，共同执行一个大型计算任务的架构。这些计算机相互协作，将计算任务分割成小块，分配给不同的机器进行处理，然后将结果合并以得出最终结论。

4 要求

4.1 性能要求

4.1.1 可靠性

应遵从GB/T 25000.10、GB/T 25000.51中的可靠性要求，在应用场景需求内，长时间连续运行时不应出现崩溃、闪退、卡死、无响应、响应迟缓等问题，准确完成多源异构数据采集、存储、管理任务，并具有容错能力。

4.1.2 易用性

应提供联机帮助，软件中各子系统用户界面风格应一致，软件应易学易用。

4.1.3 可扩展性

结构应具有可扩展性，系统应支持通过增加服务器或换用处理能力更强的服务器的方式对系统进行扩展。组成系统的每个逻辑单元都可独立于其他单元进行升级。

4.1.4 系统实时性

系统实时性应符合下列要求：

- a) 界面响应时间小于 3s;
- b) 数据查询时间小于 15s。

4.2 信息安全

应符合GB/T 20270、GB/T 20988、GB/T 22239的相关要求。

4.3 备份和恢复

备份和恢复要求包括但不限于:

- a) 宜有数据备份机制, 并对备份数据进行保护;
- b) 在使用恢复的数据前应校验其可用性、完整性;
- c) 被非正常退出, 或其他应用程序异常导致关机或者重启, 再次启动时, 功能应能恢复正常使用。

5 特点

系统充分利用分布式计算的特点, 将数据散列到不同的机器上, 减轻每台机器的负担并提高计算速度。同时, 为了支持多次迭代的计算, 系统会将每次计算的数据保存在内存中, 而非磁盘上, 从而减少了每次迭代时磁盘的读写时间, 提高了计算的效率。

6 模块功能

6.1 任务调度与分配

- 6.1.1 系统能够接收计算任务, 并根据集群中各个节点的计算能力、负载状态和网络条件等因素, 智能地将任务分配到最适合的节点上执行。
- 6.1.2 支持动态负载均衡, 确保所有节点都能够得到充分利用, 避免某些节点过载而其他节点空闲。

6.2 并行计算与加速

- 6.2.1 通过将大型计算任务分解为多个较小的子任务, 并在多个节点上并行执行这些子任务, 从而显著加快计算速度。
- 6.2.2 支持数据并行和任务并行, 根据计算任务的特点选择最合适的并行策略。

6.3 容错与恢复

- 6.3.1 系统能够自动检测和处理节点故障, 确保计算任务在节点故障时不会丢失数据或中断执行。
- 6.3.2 通过数据备份和复制、容错编码等技术, 保证数据的一致性和完整性。

6.4 资源管理与监控

- 6.4.1 监控集群中各个节点的状态、负载、网络延迟等关键指标, 确保系统始终运行在最佳状态。
- 6.4.2 提供资源管理功能, 如节点添加、删除、升级等, 以满足不同规模和需求的计算任务。

6.5 编程接口与框架

- 6.5.1 提供简单易用的编程接口和框架, 使得开发者能够方便地编写和部署分布式计算任务。
- 6.5.2 支持多种编程语言和开发框架, 降低开发门槛和成本。

6.6 安全性与隔离性

- 6.6.1 提供安全认证和授权机制, 确保只有授权的用户才能访问和执行计算任务。
- 6.6.2 支持计算任务之间的隔离性, 确保不同任务之间的数据和资源不会被互相干扰或窃取。

6.7 可扩展性与灵活性

- 6.7.1 系统能够支持从几个节点到数千个节点的扩展, 满足不同规模和需求的计算任务。
- 6.7.2 提供灵活的配置选项和扩展接口, 使得系统能够适应不同的应用场景和计算需求。

6.8 日志记录与审计

6.8.1 记录计算任务的执行日志和性能数据，便于分析和优化系统性能。

6.8.2 提供审计功能，确保系统操作的合规性和可追溯性。

7 运行测试

7.1 测试范围

根据需求分析说明书中对功能性需求以及非功能性需求的描述，确定此次的测试范围。

7.2 功能性需求测试范围

功能性需求测试的范围包括：

- a) 综合监测；
- b) 安全管理；
- c) 大数据分析；
- d) 系统管理需求模块。

7.3 非功能性需求测试范围

非功能性需求测试的范围包括：

- a) 性能测试需求：测试系统基本且常用的功能以及对响应时间要求严格的功能模块；
- b) 可靠性测试需求：运行稳定性、屏蔽用户操作错误、错误提示的准确性以及故障异常恢复能力；
- c) 易用性测试需求：操作界面符合标准和规范，系统整体功能的直观性、一致性、正确性及可理解性。

7.4 测试方法

使用黑盒测试方法，Bug跟踪管理工具，定位问题抓包工具，覆盖所有功能需求对其进行等价类划分、边界值分析、错误推测等各类测试策略测试，确保功能的实现满足系统需求要求。

7.5 性能测试

利用HP LoadRunner软件，结合参数化方法实现多用户的并发登录，使用虚拟用户并发来模拟实际用户对业务系统施加压力，查看各操作场景响应时间。

7.6 安装调试

现场安装调试软件、拟定培训材料，进行相应的前期培训，及时记录交付、安装过程中系统出现的问题。