

中华人民共和国国家标准

GB/T 42140—2022

信息技术 云计算 云操作系统 性能测试指标和度量方法

Information technology—Cloud computing—Cloud operating system
performance test metrics and measurement method

2022-12-30 发布

2023-07-01 实施

国家市场监督管理总局
国家标准化管理委员会 发布

目 次

前言	I
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 缩略语	1
5 云操作系统技术框架	2
6 云操作系统性能测试指标和度量方法	2
6.1 一般要求	2
6.2 性能效率关注点	3
6.3 性能效率测度	3
参考文献	11

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。

本文件由全国信息技术标准化技术委员会(SAC/TC 28)提出并归口。

本文件起草单位：中国建设银行股份有限公司、中国电子技术标准化研究院、清华大学、浪潮电子信息产业股份有限公司、阿里云计算有限公司、华为云计算技术有限公司、腾讯云计算(北京)有限责任公司、华云数据控股集团有限公司、天津大学、中国科学院计算技术研究所、浪潮云信息技术股份公司、上海云轴信息科技有限公司、东软集团股份有限公司、新华三技术有限公司、天翼云科技有限公司、江苏博云科技股份有限公司、云宏信息科技股份有限公司、桂林电子科技大学、平安科技(深圳)有限公司、深圳市金蝶天燕云计算股份有限公司、安超云软件有限公司、陕西省信息化工程研究院。

本文件主要起草人：金磐石、周平、常冬冬、王志鹏、史惠康、陈志峰、苏志远、朱松、吴炎冰、王升东、王荔、申思、李成栋、郭晓、张敏、刘磊、赵赫、杨贵垣、单洪博、李克秋、包云岗、李培源、饶通宇、刘娇、王伟、王勇、王鹏、林琳、白福、秦龙、赵越鹏、宋超林、武欣、杨帆、金铸、赵华、吴涛、许刚、王永霞、文钧正、何爽、赵来平、蒋德钧、潘正泰。



信息技术 云计算 云操作系统 性能测试指标和度量方法

1 范围

本文件给出了云操作系统技术框架,规定了云操作系统性能指标,并描述了对应的度量方法。

本文件适用于指导云操作系统及其基础服务的性能效率测评。

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 25000.23—2019 系统与软件工程 系统与软件质量要求和评价(SQuaRE) 第23部分:系统与软件产品质量测量

GB/T 32399—2015 信息技术 云计算 参考架构

GB/T 32400—2015 信息技术 云计算 概览与词汇

3 术语和定义

GB/T 25000.23—2019,GB/T 32400—2015界定的以及下列术语和定义适用于本文件。

3.1

云操作系统 **cloud operating system**

用于将物理设备以共享、弹性和可伸缩的方式供应和管理,提供人机界面以按需自服务可计量的方式为用户提供计算、存储和网络等基础云服务资源,可支撑 PaaS、SaaS 等服务执行的一种云计算系统软件。

3.2

测度 **measure**

以测量结果来赋值的变量。

注1:术语“测度”作为集合名词,用来指代基本测度、导出测度和指标。

注2:本文件中,当“测度”被用来量化质量特性或子特性时,它指的是一个质量测度。

4 缩略语

下列缩略语适用于本文件。

IO:输入/输出(Input Output)

PaaS:平台即服务(Platform as a Service)

SaaS:软件即服务(Software as a Service)

5 云操作系统技术框架

云操作系统技术框架包括资源抽象和调度、基础服务,以及相应的运维运营及安全能力,具体技术框架见图 1。

资源抽象和调度及基础服务指对计算、存储和网络等进行资源抽象和调度功能,主要包括:计算、存储、网络虚拟化和资源池,以及在资源层之上提供的基础服务。

运维运营指对调度和基础服务的运营和运维管理工作,主要包括:资源实例调整操作、运维操作等。

安全主要包括所有支持云操作系统所需的安全工具,包括鉴别和身份管理、授权和安全策略管理、安全防护等。

通用云服务主要包括:云操作系统之上运行的 PaaS 和 SaaS,通常包括中间件、数据库、人工智能、大数据、应用软件等服务。

基础设施主要指数据中心通常使用的设备,例如服务器、网络交换机和路由器、存储设备等,和服务器上运行的非云特有的软件,以及其他设备,例如主机操作系统、设备驱动程序、通用系统管理软件。

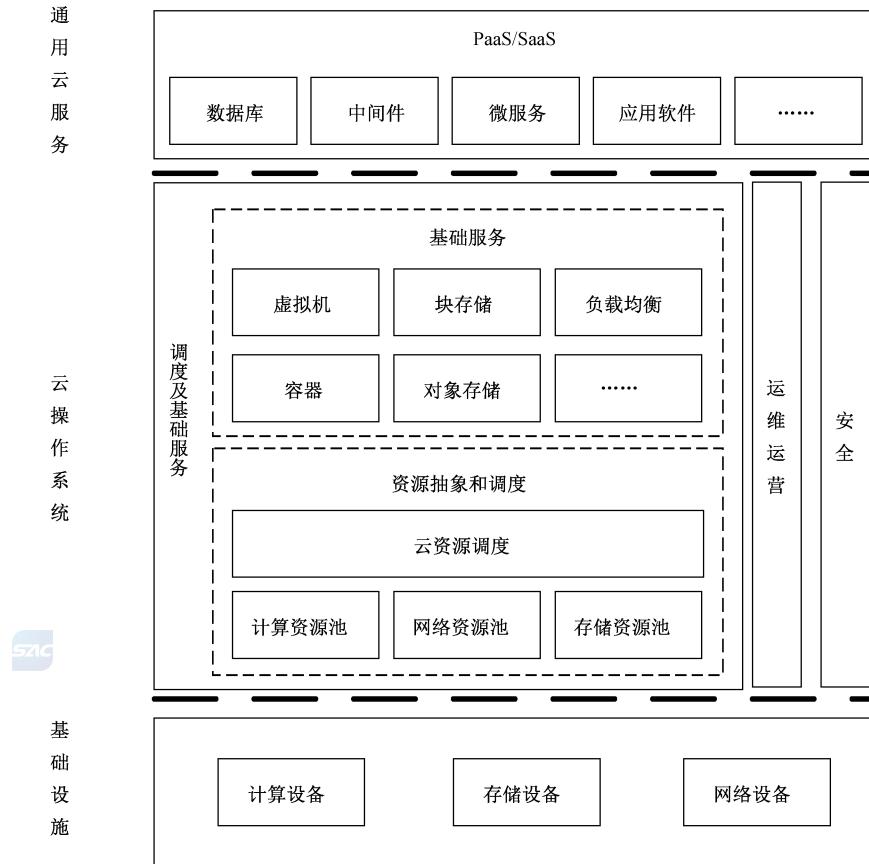


图 1 云操作系统技术框架

6 云操作系统性能测试指标和度量方法

6.1 一般要求

云操作系统性能测试指标及度量方法包括:

- a) 使用 GB/T 25000.23—2019 中 8.3 的内容,主要包括时间特性、资源利用性、容量三个方面;
- b) 结合云操作系统特点和应用需求,在 GB/T 25000.23—2019 中 8.3 的基础上,增加响应时间百分位和平均响应时间方差指标和度量方法。具体见表 1。

表 1 扩展度量指标及方法

名称	描述	测量函数	方法
响应时间百分位	系统响应时间满足规定目标的程度	$Y_i = \text{SORT}(x_i), i = (1-n)$ $Y_{\text{tail}} = y_i = \text{INT}(M * n)$ <p>式中: Y_{tail}——y_1, y_2, \dots, y_n 中的 M 分位的值; x——表示任务处理时间结果序列; Y——表示 X 序列升序排序后的结果; n——计算出的有效的服务端处理时间的总个数 </p>	测量为完成特定用户任务或系统任务而耗费的时间,并进行多次测量,按一定的速率抽取某一时间段的任务响应时间数值,将所有的数值按从小到大排序,位于列表 $N/100$ 位置的数值定义为 N th 分位数响应时间
平均响应时间方差	指系统响应一个用户任务或系统任务所用时间的平均波动程度	$S^2 = \frac{(M - x_1)^2 + \dots + (M - x_n)^2}{n}$ $M = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$ <p>式中: S^2——表示平均响应时间方差的值; n——表示任务的数量; $x_1, x_2, x_3, \dots, x_n$——表示分别响应 n 个任务所用的时间值; M——表示响应 n 个任务所用的平均时间值 </p>	测量为完成特定用户任务或系统任务而耗费的时间,并进行多次测量,计算平均响应时间方差

6.2 性能效率关注点

云操作系统质量需要综合衡量,一般包括性能、可靠性、可用性和安全性等方面,其中在性能效率方面关注以下几个方面的指标:

- a) 关注 GB/T 32399—2015 中 9.2 功能视图相关功能响应时间和资源利用性;
- b) 关注 GB/T 32399—2015 中 8.5.7 性能相关要求;
- c) 关注云操作系统提供的基础服务响应时间;
- d) 关注计算、存储、网络资源的调度处理时间;
- e) 关注云操作系统提供基础服务的调整操作响应时间,如创建虚拟机、删除虚拟机等;
- f) 关注云操作系统对资源实例等运维对象进行运维操作的响应时间;
- g) 关注云操作系统在 PaaS、SaaS 等业务负载下的响应时间、业务吞吐量以及业务本身的响应时间。

6.3 性能效率测度

6.3.1 概述

按照云操作系统的技术框架和性能效率关注点,对技术框架中的调度及基础服务、运维运营和承载的业务性能开展性能效率测试进行规定。安全部分参照国家相关信息技术安全标准执行。

6.3.2 调度及基础服务

6.3.2.1 计算资源

计算资源性能测试指标及方法见表 2。

表 2 计算资源性能测试指标及方法

序号	名称	描述	方法
1	浮点计算平均响应时间	计算资源响应一个浮点计算任务的平均时间	使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法
2	整型计算平均响应时间	计算资源响应一个整型计算任务的平均时间	使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法
3	字符串计算平均响应时间	计算资源响应一个字符串计算任务的平均时间	使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法
4	综合计算平均响应时间	计算资源响应综合计算任务(整形、浮点、字符串等计算任务)的平均时间	使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法
5	计算平均响应时间方差	计算资源响应浮点、整型等计算任务的平均响应时间方差	使用表 1 中的平均响应时间方差测量函数和方法
6	存储卷平均吞吐量	计算资源单个存储卷每秒进行数据读/写操作的字节速率	使用 GB/T 25000.23—2019 表 5 中 PTb-5-G 平均吞吐量测量函数和方法
7	网络平均吞吐量	计算资源单个虚拟网卡接收/发送数据的字节速率	使用 GB/T 25000.23—2019 表 5 中 PTb-5-G 平均吞吐量的测量函数和方法
8	内存平均吞吐量	计算资源可持续运行的内存数据传输速率	使用 GB/T 25000.23—2019 表 5 中 PTb-5-G 平均吞吐量的测量函数和方法
9	存储卷平均响应时间方差	存储卷响应一个存储管理任务所用时间的平均波动程度	使用表 1 中的平均响应时间方差测量函数和方法
10	网络平均响应时间方差	网络响应一个用户任务所用时间的平均波动程度	使用表 1 中的平均响应时间方差测量函数和方法

注 1：计算资源指提供云操作系统的计算单元，包括但不限于：虚拟机、容器、裸金属等。

注 2：计算资源性能一般按照同等单位资源进行测量。



6.3.2.2 存储资源

6.3.2.2.1 对象存储

对象存储性能测试指标及方法见表 3。

表 3 对象存储性能测试指标及方法

序号	名称	描述	方法
1	对象存储空间并发访问数	单位时间内可同时对对象存储空间进行访问的连接数	使用 GB/T 25000.23—2019 表 7 中 PCa-2-G 用户访问量的测量函数和方法
2	对象存储空间平均响应时间	单位时间内同时执行多个对象存储进行查询、读写等的平均响应时间	使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法
3	对象存储空间平均响应时间方差	对象存储空间进行一次数据读/写操作的平均响应时间方差	使用表 1 中平均响应时间方差测量函数和方法
4	对象存储空间平均响应时间百分位	对象存储空间进行一次数据读/写操作的响应时间百分位	使用表 1 中响应时间百分位测量函数和方法
5	对象存储空间读写吞吐率	对象存储空间每秒进行数据读/写操作的字节速率	使用 GB/T 25000.23—2019 表 5 中 PTb-5-G 平均吞吐量测量函数和方法
6	对象存储空间读写吞吐率方差	对象存储空间进行一次数据读/写操作的平均响应时间方差	使用表 1 中平均响应时间方差测量函数和方法
7	对象存储空间上传最大带宽	单位时间内单流上传多个对象, 系统可达的最大带宽(MBps)值	使用 GB/T 25000.23—2019 表 6 中 PRu-4-S 带宽占用率测量函数和方法
8	对象存储空间下载最大带宽	单位时间内单流下载多个对象, 系统可达的最大带宽(MBps)值	使用 GB/T 25000.23—2019 表 6 中 PRu-4-S 带宽占用率测量函数和方法

6.3.2.2.2 块存储

块存储性能测试指标及方法见表 4。

表 4 块存储性能测试指标及方法

序号	名称	描述	方法
1	存储卷平均吞吐量	云主机单个存储卷每秒进行数据读/写操作的字节速率	使用 GB/T 25000.23—2019 表 5 中 PTb-5-G 平均吞吐量测量函数和方法

表 4 块存储性能测试指标及方法 (续)

序号	名称	描述	方法
2	存储卷平均响应时间	单位时间内执行存储卷读写的平均时间	使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法
3	存储卷平均响应时间方差	存储卷响应一个存储管理任务所用时间的平均波动程度	使用表 1 中的平均响应时间方差测量函数和方法
4	存储卷响应时间百分位	存储卷响应一个存储管理任务响应时间满足规定目标的程度	使用表 1 中的响应时间百分位测量函数和方法

6.3.2.2.3 文件存储

文件存储性能测试指标及方法见表 5。

表 5 文件存储性能测试指标及方法

序号	名称	描述	方法
1	文件卷平均吞吐量	云主机单个文件卷每秒进行数据读/写操作的字节速率	使用 GB/T 25000.23—2019 表 5 中 PTb-5-G 平均吞吐量测量函数和方法
2	文件卷平均每秒读写次数	云主机单个文件卷每秒进行数据读/写操作的次数	使用 GB/T 25000.23—2019 表 5 中 PTb-5-G 平均吞吐量测量函数和方法
3	文件卷平均响应时间	单位时间内执行文件卷读写的平均时间	使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法
4	文件卷平均响应时间方差	文件卷响应一个存储管理任务所用时间的平均波动程度	使用表 1 中的平均响应时间方差测量函数和方法
5	文件卷响应时间百分位	文件卷响应一个存储管理任务响应时间满足规定目标的程度	使用表 1 中的响应时间百分位测量函数和方法

6.3.2.3 网络资源

网络资源性能测试指标及方法见表 6。

表 6 网络资源性能测试指标及方法

序号	名称	描述	方法
1	IO 吞吐率	在无帧丢失时, IO 进行数据读取所能承受的最大数据传输速率	使用 GB/T 25000.23—2019 表 5 中 PTb-5-G 平均吞吐量测量函数和方法

表 6 网络资源性能测试指标及方法 (续)

序号	名称	描述	方法
2	数据包收发量	单位时间内,在无数据包丢失时,进行数据读取所能承受的最大数据包传输量	使用 GB/T 25000.23—2019 表 5 中 PTb-5-G 平均吞吐量测量函数和方法
3	网络平均吞吐量	单位时间内,在无数据包丢失时,进行读取数据的字节速率	使用 GB/T 25000.23—2019 表 5 中 PTb-5-G 平均吞吐量的测量函数和方法
4	新建连接数	单位时间内,单一云主机新建的平均连接数	使用 GB/T 25000.23—2019 表 7 中 PCa-2-G 用户访问量的测量函数和方法
5	并发连接数	单位时间内,单一云主机同时连接的平均连接数	使用 GB/T 25000.23—2019 表 7 中 PCa-2-G 用户访问量的测量函数和方法
6	平均监听请求数(负载均衡专用)	单位时间内,负载通过超文本传输协议完成监听或查询的平均请求数是多少	使用 GB/T 25000.23—2019 表 14 中 P0p-5-S 监视能力测量函数和方法
7	网络时延	等待对网络中的存储数据访问完成所引起的延时时间	使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法
8	平均方差	响应一个用户任务所用时间的平均波动程度	使用表 1 中的平均响应时间方差测量函数和方法

6.3.3 运维运营

6.3.3.1 资源实例操作

资源实例操作性能测试指标及方法见表 7。

表 7 资源实例操作性能测试指标及方法

序号	名称	描述	方法
1	创建资源实例平均响应时间	单位时间内,创建一个资源实例的平均时间	使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法
2	删除资源实例平均响应时间	单位时间内,删除一个资源实例的平均时间	使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法
3	资源实例列表平均查询响应时间	单位时间内,进行一次资源列表查询的平均时间	使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法

表 7 资源实例操作性能测试指标及方法(续)

序号	名称	描述	方法
4	资源实例配置平均修改响应时间	单位时间内,进行一次资源实例配置修改的平均时间	使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法
5	开启资源实例平均响应时间	单位时间内,开启一个资源实例机的平均时间	使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法
6	关闭资源实例平均响应时间	单位时间内,关闭一个资源实例的平均时间	使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法
7	迁移资源实例平均响应时间	单位时间内,迁移一个资源实例的平均时间	使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法
8	资源实例调整操作平均响应时间方差	资源实例的创建、启动、迁移等调整事件的平均响应时间方差	使用表 1 中平均响应时间方差测量函数和方法
9	资源实例调整操作响应时间百分位	资源实例的创建、启动、迁移等调整事件的响应时间百分位	使用表 1 中响应时间百分位测量函数和方法
10	运维操作最大并发规模	单位时间内,支持对资源实例等运维对象进行运维操作的最大总任务数	并发规模 = $\sum_{i=1}^n$ 租户 i 并发任务总数, $i=(1 \sim n)$

注 1: 资源实例操作对象包括但不限于:虚拟机、容器、文件存储、对象存储、块存储、裸金属、负载均衡器等。

注 2: 运维操作指对已生成的资源实例的资源纳管、配置变更、定时及批量作业等维护类操作。

6.3.3.2 资源实例采集和监控

资源实例采集和监控性能测试指标及方法见表 8。

表 8 资源实例采集和监控性能测试指标及方法

序号	名称	描述	方法
1	资源实例监控最短周期	查询的资源监控指标采样的最短间隔时间周期	监控指标的两个连续监控数据的最短间隔时间
2	资源实例异常检测耗时	监控的资源实例发生真实异常后,从异常发生到收到告警所消耗的时间	使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法
3	监控数据查询响应时间	查询单一对象固定监控指标,单位时间周期监控结果数据的响应时间	使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法

表 8 资源实例采集和监控性能测试指标及方法 (续)

序号	名称	描述	方法
4	监控策略生效响应时间	单一对象固定监控指标策略的生效响应时间	使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法
5	运维工具自身运行资源使用率	单位时间内,有关采集、监控和操作等运维工具占用资源实例宿主中央处理器总平均占用率和内存平均占用率	使用 GB/T 25000.23—2019 表 6 中 PRu-1-G 中央处理器平均占用率,以及 PRu-2-G 内存平均占用率的测量函数和方法

6.3.4 业务性能

6.3.4.1 IO 密集型

IO 密集型性能测试指标及方法见表 9。

表 9 IO 密集型性能测试指标及方法

序号	名称	描述	方法
1	云操作系统执行 IO 密集型业务性能	单位时间内,云操作系统执行大数据等 IO 密集型业务时性能表现,性能包括几个维度指标内容: a) 业务吞吐量,如单位时间内完成数据库插入、读取的数量; b) 实例配置平均响应时间; c) 实例所需资源扩展平均响应时间	业务吞吐量使用 GB/T 25000.23—2019 表 5 中 PTb-5-G 平均吞吐量的测量函数和方法; 平均响应时间使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法

6.3.4.2 计算密集型

计算密集型性能测试指标及方法见表 10。

表 10 计算密集型性能测试指标及方法

序号	名称	描述	方法
1	云操作系统执行计算密集型业务性能	单位时间内,云操作系统执行人工智能、高性能计算等计算密集型业务时性能表现,性能包括几个维度指标内容: a) 业务吞吐量、任务平均响应时间; b) 实例配置平均响应时间; c) 实例所需资源弹性扩展平均响应时间	业务吞吐量使用 GB/T 25000.23—2019 表 5 中 PTb-5-G 平均吞吐量的测量函数和方法; 平均响应时间使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法

6.3.4.3 内存密集型

内存密集型性能测试指标及方法见表 11。

表 11 内存密集型性能测试指标及方法

序号	名称	描述	方法
1	云操作系统执行内存密集型业务性能	<p>单位时间内,云操作系统执行内存数据库等内存密集型业务时性能表现,性能包括几个维度指标内容:</p> <p>a) 业务吞吐量、数据库查询平均响应时间;</p> <p>b) 实例配置平均响应时间;</p> <p>c) 实例所需资源弹性扩展平均响应时间</p>	<p>业务吞吐量使用 GB/T 25000.23—2019 表 5 中 PTb-5-G 平均吞吐量的测量函数和方法</p> <p>平均响应时间使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法</p>

6.3.4.4 综合业务型

综合业务型性能测试指标及方法见表 12。

表 12 综合业务型性能测试指标及方法

序号	名称	描述	方法
1	云操作系统执行综合业务型业务性能	<p>单位时间内,云操作系统执行中间件、数据库、大数据、人工智能、高性能计算等综合业务时性能表现,性能包括几个维度指标内容:</p> <p>a) 业务吞吐量、业务平均响应时间;</p> <p>b) 实例配置平均响应时间;</p> <p>c) 实例所需资源弹性扩展平均响应时间</p>	<p>业务吞吐量使用 GB/T 25000.23—2019 表 5 中 PTb-5-G 平均吞吐量的测量函数和方法</p> <p>平均响应时间使用 GB/T 25000.23—2019 表 5 中 PTb-1-G 平均响应时间的测量函数和方法</p>

参 考 文 献

- [1] GB/T 25000.10—2016 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
-