

《云应用中的服务质量》

图书基本信息

书名：《云应用中的服务质量》

13位ISBN编号：9787111523520

出版时间：2016-2

作者：埃里克.鲍尔

页数：247

译者：谭励,杨明华

版权说明：本站所提供下载的PDF图书仅提供预览和简介以及在线试读，请支持正版图书。

更多资源请访问：www.tushu000.com

《云应用中的服务质量》

内容概要

本书介绍了关于应用程序服务质量、云模型以及虚拟化架构缺陷的基础内容，系统地分析了应用程序服务由于云架构缺陷受到的影响，为云计算应用以及尚在开发过程中的应用提供了技术和策略方面的建议，优化其提供优质服务的能力。本书适合从事云计算、云应用设计以及软件工程行业的人士阅读，也适合作为相关专业的师生的参考书。

《云应用中的服务质量》

作者简介

Eric Bauer是阿尔卡特-朗讯的IP平台CTO的可靠性工程经理，他曾在阿尔卡特-朗讯的平台、应用以及解决方案的可靠性方面工作超过十年。在从事可靠性工程领域之前，Bauer花了二十年时间设计和开发嵌入式固件、网络操作系统、IP PBX、互联网平台以及光传输系统。BAUER获得了十多项美国专利，撰写了《云计算实战:可靠性与可用性设计》、《超越冗余：地理冗余如何才能提高计算机系统的可用性和可靠性》、《可靠性设计：信息和计算机系统》、《系统可靠性实用技术》（均由Wiley-IEEE出版社出版）等著作，并有多篇论文在《贝尔实验室技术期刊》发表。BAUER拥有康奈尔大学电子工程学士学位和普渡大学电气工程硕士学位，他住在新泽西州弗里霍尔德。

Randee Adams是阿尔卡特-朗讯的IP平台CTO的技术顾问，她花了近十年时间专注于产品的可靠性设计，曾在各种内部可靠性论坛上多次发言。ADAMS撰写了《超越冗余：地理冗余如何才能提高计算机系统的可用性和可靠性》和《云计算的可靠性和可用性》等著作。她曾作为5ESS交换机的程序员，于1979年加入贝尔实验室。ADAMS在整个公司的多个项目（例如，软件开发、故障单管理、负载管理研究、软件交付、系统工程、软件架构、软件设计、开发工具和联合风险设置）和多个功能领域（例如，数据库管理、公共信道信令、操作实施、指导和管理、可靠性和安全性）工作过。ADAMS拥有亚利桑那大学的学士学位以及伊利诺斯理工学院的计算机科学硕士学位。她住在伊利诺伊州内珀维尔。

书籍目录

关于作者

译者序

1 概述

1.1 入门

1.2 目标读者

1.3 本书组织结构

I 配置

2.应用程序服务质量

2.1 简单应用程序模型

2.2 服务边界

2.3 质量和性能的关键指标

2.4 关键应用特征

2.4.1 服务急迫性

2.4.2 应用程序交互性

2.4.3 网络传输缺陷的耐受性

2.5 应用程序服务质量指标

2.5.1 服务可用性

2.5.2 服务延迟

2.5.3 服务可靠性

2.5.4 服务可访问性

2.5.5 服务可维持性

2.5.6 服务吞吐量

2.5.7 服务时间戳精度

2.5.8 特定应用程序的服务质量度量

2.6 技术服务与支持服务

2.6.1 技术服务质量

2.6.2 支持服务质量

2.7 安全事项

3 云模型

3.1 云计算中的角色

3.2 云服务模型

3.3 云的基本特征

3.3.1 按需自助服务

3.3.2 广泛的网络访问

3.3.3 资源池

3.3.4 快速弹性

3.3.5 度量服务

3.4 简化云架构

3.4.1 应用软件

3.4.2 虚拟机服务器

3.4.3 虚拟机服务器控制器

3.4.4 云操作支持系统

3.4.5 云技术组件“即服务”

3.5 弹性度量

3.5.1 密度

3.5.2 配置间隔

3.5.3 释放间隔

- 3.5.4 向内和向外扩展
- 3.5.5 向上和向下扩展
- 3.5.6 敏捷性
- 3.5.7 转换速率和线性度
- 3.5.8 弹性加速
- 3.6 空间和区域
- 3.7 云意识
- 4 虚拟化架构缺陷
 - 4.1 服务延迟、虚拟化和云
 - 4.1.1 虚拟化和云导致的延迟变化
 - 4.1.2 虚拟化开销
 - 4.1.3 增加架构性能的可变性
 - 4.2 虚拟机故障
 - 4.3 无法交付的虚拟机配置容量
 - 4.4 交付退化的虚拟机容量
 - 4.5 尾部延迟
 - 4.6 时钟事件抖动
 - 4.7 时钟漂移
 - 4.8 失败或缓慢的虚拟机实例分配和启动
 - 4.9 虚拟化架构缺陷展望
- II 分析
- 5 应用程序冗余和云计算
 - 5.1 故障，可用性和简单建构
 - 5.2 通过虚拟化改进软件修复时间
 - 5.3 通过虚拟化改进架构修复时间
 - 5.3.1 理解硬件修复
 - 5.3.2 虚拟机修复即服务
 - 5.3.3 讨论
 - 5.4 冗余和可恢复性
 - 5.4.1 通过虚拟化改进恢复时间
 - 5.5 顺序冗余和并发冗余
 - 5.5.1 混合并发策略
 - 5.6 虚拟化缺陷对应用服务的影响
 - 5.6.1 简单架构的服务影响
 - 5.6.2 顺序冗余架构的服务影响
 - 5.6.3 并发冗余架构的服务影响
 - 5.6.4 混合并发架构的服务影响
 - 5.7 数据冗余
 - 5.7.1 数据存储策略
 - 5.7.2 数据一致性策略
 - 5.7.3 数据架构注意事项
 - 5.8 讨论
 - 5.8.1 服务质量的影响
 - 5.8.2 并发控制
 - 5.8.3 资源使用
 - 5.8.4 简易性
 - 5.8.5 其他注意事项
- 6 负载分配与均衡
 - 6.1 负载分配机制

- 6.2 负载分配策略
- 6.3 代理负载均衡器
- 6.4 非代理负载分配
- 6.5 负载分配的层次结构
- 6.6 基于云的负载均衡所面临的挑战
- 6.7 负载均衡在支持冗余方面的作用
- 6.8 负载均衡与可用区域
- 6.9 工作负载服务度量
- 6.10 操作注意事项
 - 6.10.1 负载均衡与弹性
 - 6.10.2 负载均衡与过载
 - 6.10.3 负载均衡与发布管理
- 6.11 负载均衡与应用程序服务质量
 - 6.11.1 服务可用性
 - 6.11.2 服务延迟
 - 6.11.3 服务可靠性
 - 6.11.4 服务可访问性
 - 6.11.5 服务可维持性
 - 6.11.6 服务吞吐量
 - 6.11.7 服务时间戳精度
- 7 故障容器
 - 7.1 故障容器
 - 7.1.1 故障级联
 - 7.1.2 故障容器与恢复
 - 7.1.3 故障容器与虚拟化
 - 7.2 故障点
 - 7.2.1 单点故障
 - 7.2.2 单点故障与虚拟化
 - 7.2.3 关联性和反关联性考虑
 - 7.2.4 在云计算中确保无SPOF
 - 7.2.5 无SPOF和应用程序数据
 - 7.3 极端共存解决方案
 - 7.3.1 极端共存解决方案的风险
 - 7.4 多租户与解决方案容器
- 8 容量管理
 - 8.1 工作负载变化
 - 8.2 传统容量管理
 - 8.3 传统过载控制
 - 8.4 容量管理与虚拟化
 - 8.5 云容量管理
 - 8.6 弹性存储注意事项
 - 8.7 弹性和过载
 - 8.8 操作注意事项
 - 8.9 负载拉锯
 - 8.10 一般弹性风险
 - 8.11 弹性故障场景
 - 8.11.1 弹性增长故障场景
 - 8.11.1 弹性逆增长故障场景
- 9 发布管理

- 9.1 相关术语
- 9.2 传统的软件升级策略
 - 9.2.1 软件升级需求
 - 9.2.2 维护窗口
 - 9.2.3 应用升级的客户端注意事项
 - 9.2.4 传统的离线软件升级
 - 9.2.5 传统的在线软件升级
 - 9.2.6 讨论
- 9.3 支持云的软件升级策略
 - 9.3.1 I型云支持升级策略：街区聚会
 - 9.3.2 II型云支持升级策略：每车一司机
 - 9.3.3 讨论
- 9.4 数据管理
- 9.5 软件升级中的服务编排角色
 - 9.5.1 解决方案级软件升级
- 9.6 结论
- 10 端到端考虑因素
 - 10.1 端到端服务环境
 - 10.2 三层端到端服务模型
 - 10.2.1 通过三层模型估算服务缺陷
 - 10.2.2 端到端服务可用性
 - 10.2.3 端到端服务延迟
 - 10.2.4 端到端服务可靠性
 - 10.2.5 端到端服务可访问性
 - 10.2.6 端到端服务可维持性
 - 10.2.7 端到端服务吞吐量
 - 10.2.8 端到端服务时间戳精度
 - 10.2.9 现实检查
 - 10.3 分布式和集中式的云数据中心
 - 10.3.1 集中式云数据中心
 - 10.3.2 分布式云数据中心
 - 10.3.3 服务可用性考虑
 - 10.3.4 服务延迟考虑
 - 10.3.5 服务可靠性考虑
 - 10.3.6 服务可访问性考虑
 - 10.3.7 服务可维持性考虑
 - 10.3.8 资源分配考虑
 - 10.4 多层解决方案架构
 - 10.5 灾难恢复与地理冗余
 - 10.5.1 灾难恢复目标
 - 10.5.2 地理冗余架构
 - 10.5.3 服务质量考虑
 - 10.5.4 恢复点考虑
 - 10.5.5 地理冗余和可用区域减轻灾难的影响
- III 建议
- 11 服务质量问责
 - 11.1 传统的问责
 - 11.2 云服务交付路径
 - 11.3 云问责

- 11.4 问责案例研究
 - 11.4.1 问责和技术组件
 - 11.4.2 问责和弹性
- 11.5 服务质量差距模型
 - 11.5.1 应用程序面向资源服务差距分析
 - 11.5.2 应用程序面向用户服务差距分析
- 11.6 服务水平协议
- 12 服务可用性度量
 - 12.1 服务度量概述
 - 12.2 传统服务可用性度量
 - 12.3 服务可用性度量演化
 - 12.3.1 应用演化分析
 - 12.3.2 技术组件
 - 12.3.3 存储即服务的使用
 - 12.4 硬件可靠性度量演化
 - 12.4.1 虚拟机故障生命周期
 - 12.5 弹性服务可用性度量演化
 - 12.6 发布管理服务可用性度量演化
 - 12.7 服务度量展望
- 13 应用程序服务质量需求
 - 13.1 服务可用性需求
 - 13.2 服务延迟需求
 - 13.3 服务可靠性需求
 - 13.4 服务可访问性需求
 - 13.5 服务可维持性需求
 - 13.6 服务吞吐量需求
 - 13.7 时间戳精度需求
 - 13.8 弹性需求
 - 13.9 发布管理需求
 - 13.10 灾难恢复需求
- 14 虚拟化架构度量与管理
 - 14.1 架构服务质量度量的业务环境
 - 14.2 云消费者度量选项
 - 14.3 缺陷度量策略
 - 14.3.1 虚拟机故障度量
 - 14.3.2 无法交付的虚拟机配置容量度量
 - 14.3.3 交付退化的虚拟机容量度量
 - 14.3.4 尾部延迟度量
 - 14.3.5 时钟事件抖动度量
 - 14.3.6 时钟漂移度量
 - 14.3.7 失败或缓慢的虚拟机实例分配和启动度量
 - 14.3.8 度量总结
 - 14.4 管理虚拟化架构缺陷
 - 14.4.1 最小化应用程序对架构缺陷的敏感度
 - 14.4.2 虚拟机级拥塞检测与控制
 - 14.4.3 分配更多虚拟资源容量
 - 14.4.4 终止性能欠佳的虚拟机实例
 - 14.4.5 接受性能退化
 - 14.4.6 积极主动的供应商管理

- 14.4.7 重新设定最终用户服务质量期望
- 14.4.8 SLA注意事项
- 14.4.9 更换云服务提供商
- 15 基于云的应用程序分析
 - 15.1 可靠性框图和参照分析
 - 15.2 IaaS缺陷影响分析
 - 15.3 PaaS故障影响分析
 - 15.4 工作负载分配分析
 - 15.4.1 服务质量分析
 - 15.4.2 过载控制分析
 - 15.5 反相关性分析
 - 15.6 弹性分析
 - 15.6.1 服务容量增长场景
 - 15.6.2 服务容量增长操作分析
 - 15.6.3 服务容量逆增长操作分析
 - 15.6.4 存储容量增长场景
 - 15.6.5 在线存储容量增长操作分析
 - 15.6.6 在线存储容量逆增长操作分析
 - 15.7 发布管理影响效应分析
 - 15.7.1 服务可用性影响
 - 15.7.2 服务可靠性影响
 - 15.7.3 服务可访问性影响
 - 15.7.4 服务可维持性影响
 - 15.7.5 服务吞吐量影响
 - 15.8 恢复点目标分析
 - 15.9 恢复时间目标分析
- 16 测试注意事项
 - 16.1 测试环境
 - 16.2 测试策略
 - 16.2.1 云测试平台
 - 16.2.2 用于测试的容量
 - 16.2.3 统计置信度
 - 16.2.4 服务中断时间
 - 16.3 模拟架构缺陷
 - 16.4 测试计划
 - 16.4.1 服务可靠性和延迟测试
 - 16.4.2 架构缺陷测试
 - 16.4.3 健壮性测试
 - 16.4.4 持久性/稳定性测试
 - 16.4.5 应用程序弹性测试
 - 16.4.6 升级测试
 - 16.4.7 灾难恢复测试
 - 16.4.8 极限共存测试
 - 16.4.9 PaaS技术组件测试
 - 16.4.10 自动回归测试
 - 16.4.11 构造发布测试
- 17 关键点连接与总结
 - 17.1 应用程序服务质量所面临的挑战
 - 17.2 冗余和健壮性

17.3 可伸缩性设计

17.4 可扩展性设计

17.5 故障设计

17.6 规划注意事项

17.7 传统应用的演化

17.7.1 阶段0：传统应用

17.7.2 阶段I：虚拟化架构上的高服务质量

17.7.3 阶段II：手动应用弹性

17.7.4 阶段III：自动发布管理

17.7.5 阶段IV：自动应用弹性

17.7.6 阶段V：虚拟机迁移

17.8 结束语

参考文献

缩略语

《云应用中的服务质量》

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:www.tushu000.com