南京大學

研 究 生 毕 业 论 文

（申请硕士学位）

论　文　题　目　支持事务一致性的微服务动态更新系统实现

作　者　姓　名　王东宇

学科、专业方向　计算机科学与技术

研　究　方　向　软件方法学

指　导　教　师　曹春 教授

2021 年 **4** 月 **7** 日

学　　号：**MG1833071**

论文答辩日期：**2021** 年 **6** 月 **1** 日

指 导 教 师：　　（签字）

Implementation of a Microservice System Supporting Dynamic Update With Transaction Consistency

by

Dongyu Wang

Supervised by

Professor Chun Cao

A dissertation submitted to
the graduate school of Nanjing University
in partial fulfilment of the requirements for the degree of
MASTER
in
Computer Science and Technology

Department of Computer Science and Technology
Nanjing University

April 15, 2021

南京大学研究生毕业论文中文摘要首页用纸

毕业论文题目：支持事务一致性的微服务动态更新系统实现

计算机科学与技术 专业 2018 级硕士生姓名：王东宇

指导教师（姓名、职称）：曹春 教授

摘　　要

关键词：小世界理论；网络模型；数据中心

南京大学研究生毕业论文英文摘要首页用纸

THESIS: Implementation of a Microservice System Supporting Dynamic Update With Transaction Consistency

SPECIALIZATION: Computer Science and Technology

POSTGRADUATE: Dongyu Wang

MENTOR: Professor Chun Cao

Abstract

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

keywords: Small World, Network Model, Data Center

目　　次

目　　次 ····· **v**

插图清单 ····· **ix**

附表清单 ····· **xi**

1　绪言 ····· **1**

1.1　研究背景 ····· 1

1.1.1　微服务架构 ····· 1

1.1.2　软件动态更新技术 ····· 2

1.2　研究现状 ····· 2

1.3　本文工作 ····· 4

1.4　本文组织 ····· 4

2　相关工作 ····· **7**

2.1　软件动态更新技术 ····· 7

2.1.1　软件系统实例 ····· 8

2.1.2　Quiescence 算法 ····· 10

2.1.3　Tranquility 算法 ····· 11

2.1.4　Version Consistency 算法 ····· 12

2.2　微服务架构 ····· 15

2.2.1　Spring ····· 16

2.2.2　Dubbo ····· 17

2.2.3　Service Mesh ····· 18

2.2.4　微服务架构与动态更新 ····· 19

2.3　本章小结 ····· 20

3 支持事务一致性的微服务动态更新方法 ····· 21
3.1 需求与挑战 ····· 21
3.2 支持微服务动态更新的方法 ····· 22
3.2.1 事务模型与依赖管理 ····· 23
3.2.2 服务生命周期 ····· 25
3.3 本章小结 ····· 27

4 支持动态更新的微服务系统实现 ····· 29
4.1 系统设计概述 ····· 29
4.2 系统扩展实现 ····· 30
4.2.1 事务依赖管理器 ····· 30
4.2.2 服务生命周期管理器 ····· 31
4.2.3 更新控制模块 ····· 38
4.3 系统模块交互 ····· 38
4.4 本章小结 ····· 40

5 案例研究与实验评估 ····· 41
5.1 案例介绍 ····· 41
5.2 实验环境与设计 ····· 43
5.2.1 实验环境 ····· 43
5.2.2 实验设计 ····· 43
5.3 实验分析 ····· 44
5.3.1 安全性 (Safety) ····· 44
5.3.2 及时性 (Timeliness) ····· 45
5.3.3 干扰性 (Disruption) ····· 45
5.4 实验结论 ····· 47
5.5 本章小结 ····· 48

6 总结与展望 ····· 49
6.1 工作总结 ····· 49
6.2 研究展望 ····· 50

致 谢 ····· 51

参考文献 **53**

简历与科研成果 **59**

学位论文出版授权书 **61**

插图清单

2-1 示例系统静态依赖图 …… 9
2-2 示例系统时序图 …… 9
2-3 Quiescence 节点状态转移图 …… 11
2-4 系统运行时动态依赖关系示例 …… 14
2-5 Spring 架构图 …… 16
2-6 Dubbo 服务调用关系图 …… 18
2-7 Service Mesh 架构 …… 19

3-1 事务状态转移图 …… 24
3-2 服务生命周期 …… 27

4-1 Istio 架构图 …… 31
4-2 微服务实例模型 …… 32
4-3 控制器模式 …… 32
4-4 初始状态下的系统 …… 34
4-5 更新准备前的系统 …… 34
4-6 版本更新中的系统 …… 35
4-7 维护 PAST_SET 集合 …… 36
4-8 阻塞访问 …… 36
4-9 同步状态 …… 37
4-10 等待撤销 …… 38
4-11 版本更新时序图 …… 39
4-12 版本撤销时序图 …… 39

5-1 台风系统静态依赖关系图 …… 42
5-2 台风系统时序图 …… 42
5-3 及时性实验结果 …… 46
5-4 总干扰性实验结果 …… 47

5-5 平均干扰性实验结果 ····· 47

附表清单

2-1　动态更新算法比较 …… 15

5-1　安全性实验结果…… 45

第一章　绪言

1.1　研究背景

1.1.1　微服务架构

在软件技术的发展过程中，诞生了多样化的软件架构和技术。随着软件系统规模的扩大、功能的追加扩展，传统的单体应用 (Monolithic) 将应用程序的所有功能打包成一个独立的单元，最终将成为一个庞然大物，变得越来越复杂，逻辑耦合严重，难以理解。因此为了满足用户对于一定规模的软件的快速开发集成、提高可扩展性等要求，必须将其迁移到微服务架构中来[1–4]。微服务架构 (Microservice Architecture) 是一种基于一组独立部署运行的小型服务来构建应用的架构方法，服务间使用与语言无关的轻量级的通信协议进行通信[5,6]。与传统的单体应用相比，微服务架构可以帮助我们对应用进行有效地拆分，缩短软件开发周期，降低软件维护成本，满足了软件程序对于快速持续集成和持续交付的需求[7,8]。由于前述的相关特性，微服务架构开始在诸多公司内部得到广泛的应用，包括 Amazon[1]、Netflix[2]、Spotify[3]等，成熟的微服务框架有 Spring[4]、Dubbo[5]。

微服务架构的基本思想是将应用软件拆分为若干个较小的服务，不同的服务可由不同的团队使用不同的技术栈进行开发，从而达到解耦和降低复杂度的目的[5]。微服务架构的开发流程中的关键步骤在于对服务的拆分。微服务在进行拆分时需要遵循服务独立、数据一致、高内聚低耦合等原则，并且随着需求的迭代，服务的拆分需要持续的进行优化。

从开发过程来看，微服务架构尽量采用去中心化的管理机制，服务主要围绕应用系统的业务能力来进行开发。从软件架构来看，一个微服务架构的系统需要满足一系列的条件和原则，而不仅仅是使用了某个微服务框架即可。因此

[1]https://gigaom.com/2011/10/12/419-the-biggest-thing-amazon-got-right-the-platform/
[2]https://www.nginx.com/blog/Microservices-at-netflix-architectural-best-practices/
[3]https://www.infoq.com/presentations/linkedin-Microservices-urn/
[4]https://spring.io/projects
[5]https://dubbo.apache.org/zh/

微服务架构更多地被看成一种架构风格，而不是一种具体的架构。

现有的关于微服务架构的研究较多关注于服务敏捷开发、服务治理等技术，而对于如何保证微服务在运行时正确地进行更新，相关的技术研究较少。

1.1.2 软件动态更新技术

Internet 的发展推动了软件技术的快速发展，而随着软件外部环境和用户需求的不断变化，已部署的软件同样需要不断地进行更新迭代，来达到修复 bug、增加功能、适应需求变化等多方面的目的。传统的软件更新通常需要停止旧版本软件的运行，然后完成软件的更新，再重新进行新版本软件的部署。显然，这种更新方法会引起一定时间内的系统中断，但某些服务提供商需要为用户提供全天候不间断的服务，对于企业服务，系统的中断将会导致经济效益的损失；而在航空和医疗领域，系统的中断将可能威胁到人的生命安全[9–11]。因此，此类软件系统不能接受传统的更新方式，需要在不中断软件正常运行的情况对其进行更新，保证系统服务的持续可用，即所谓的软件动态更新[12]。然而，软件的动态更新并不容易实现，它在安全性[1]、及时性、低干扰性等多方面面临着较大的挑战。

从软件动态更新的粒度来看，基于软件的动态更新可大致分为三个层面：

- **业务过程层面** 借助可变过程模型建模和过程迁移技术，可以将过程模型的变化迅速反应到过程的动态执行上；
- **程序代码层面** 在程序代码层面，对新旧版本程序进行差别分析，将旧版本程序的状态动态迁移到新版本中，并按照新版本的程序继续执行；
- **软件服务层面** 在构成软件服务的实例层面，研究面向动态更新的服务容器等，实现软件系统的动态更新和重配置；

本文主要关注于从软件服务层面对软件进行动态更新。

1.2 研究现状

目前在工业界较为常见的微服务动态更新部署方式有：蓝绿部署[13]、金丝雀部署[14,15] 和滚动发布。其中蓝绿部署利用冗余的硬件设备来加载新版本的

[1]本文中的安全性和一致性均指系统在动态更新过程的前后都能正确运行

软件服务，属于基于硬件的动态更新技术，这种方式时间和价格成本较高；另外两种则属于基于软件的动态更新技术，在旧版本实例运行时同时运行新版本服务实例，利用划分用户流量的方式，将用户请求逐渐地转发往新版本服务实例，直至将系统中所有服务实例变为新版本，完成更新。

对于基于微服务的动态更新，文献[16] 中实现的 JRO 基于指定的目标体系结构来完成微服务的自动更新部署，但局限于特定的编程语言 Jolie[17]。文献[18] 描述了一个支持微服务演化的模型，该模型涵盖了微服务架构与功能相关内容，确保微服务完成重构的一致性。文献[19] 提出了一种描述相关 PaaS 站点信息及其内部微服务部署信息的架构模型，运维人员通过指定更新策略来对特定的架构模型进行更新操作，最终反映到具体的服务中。这些方法实现了微服务的更新替换，但都缺乏对更新过程中运行时事务的考虑，无法保证动态更新前后系统运行的安全性。

如何选择安全的更新点来保证动态更新前后系统运行的安全性，是动态更新中至关重要的问题，一致性要求系统中的运行时事务在动态更新前后都不会出现正确性问题[20]。在这方面的研究已经存在较多的研究成果。文献[12,21] 中的 CONIC 系统，提出了 Quiescence 更新算法，该算法给出了动态更新过程中对构件的管理包括增加、删除和替换等操作，而且定义了 Quiescence 状态并将其作为可以执行动态更新的安全时机。但为使系统到达 Quiescence 状态，算法要求系统中断所有外来的业务处理直至目标更新完成，对系统的干扰极大。因此在文献[22] 中放松了对安全点的定义，提出了 Tranquility 动态更新算法。但对于某些具体应用，Tranquility 算法虽然大大降低了更新对系统造成的干扰，它仅仅保证了事务的局部而非全局一致性，因此导致它所定义的安全点可能并不安全。综合前述两种算法在动态更新过程中的问题，文献[23,24] 提出了 Version Consistency 算法，该算法利用运行时动态依赖关系来对安全点进行定义，在保证动态更新过程的安全性的同时较大程度地减少了对系统带来的干扰，在安全性和干扰性等方面取得了较好的平衡。

想要将动态更新应用于实际的微服务系统中，尚存在诸多问题。主要原因在于当前的动态更新算法通常都建立在特定的系统假设和约束条件之上，当系统的具体运行环境泛化时，相关的框架系统缺乏对动态更新的支撑。现有的微服务系统对动态更新支持的不足主要体现为：

- **安全性不足** 无法保证在系统中某个目标服务在动态更新过程中，正在运行时事务和将要发生的事务的一致性；

- **及时性不足** 当收到动态更新请求后，系统需要长时间的操作才能完成，无法快速及时地完成目标服务的动态更新；
- **干扰性过大** 系统在进行动态更新的过程中，需要长时间地中断目标服务，对目标服务造成的干扰性过大；

1.3 本文工作

随着微服务开发架构的流行，微服务系统变得多元且复杂，而现有的微服务架构并未对服务的动态更新提供相关的支持。本文针对上述动态更新技术中存在的问题，提出了一种支持事务一致性的微服务动态更新的实现方法。该方法利用 Service Mesh[25] 作为支撑微服务系统的底层架构，在其基础之上进行模型的扩展，为系统中服务的动态更新提供支持，对实际场景中不同类型的微服务系统均具有透明性。整体来看，本文的工作主要包括：

- 分析当前主流微服务系统框架的基本功能和原理，得出现有框架技术在动态更新支撑方面的不足。基于此，提出一种支持事务一致性的微服务动态更新的实现方法，在事务模型、动态依赖管理和服务生命周期三个方面进行了相应的扩展，为具体的动态更新过程提供运行时支持。
- 基于上述的实现方法，在 Service Mesh 的开源实现框架 Istio[1]之上完成了模型的扩展，添加系统对动态更新的支持。该系统不仅保证动态更新过程的安全性、及时性和低干扰性，而且实现的扩展模块对用户保持透明，与具体的服务业务逻辑解耦开。
- 利用实际的应用案例，整合上述实现方法并进行实验评估，使用不同的更新算法，主要对动态更新过程中各个算法的安全性 (Safety)、及时性 (Timeliness)、干扰性 (Disruption) 进行了性能比较。

1.4 本文组织

本文内容组织如下：

[1]Istio home page: https://istio.io/

第二章从软件动态更新和已有的微服务框架对动态更新的支持两个方面对相关工作进行介绍。在软件动态更新方面，主要介绍已有的动态更新算法，包括 Quiescence、Tranquility 和 Version Consistency；在微服务框架支持方面，重点介绍常见主流的微服务框架，主要包括 Spring、Dubbo 和相关的 Service Mesh 技术架构，以及相关的云平台对微服务架构在动态更新方面的支持。

第三章给出一种支持事务一致性的微服务动态更新实现方法。首先针对相关微服务架构对动态更新支持方面的不足与需求进行了探讨，然后基于此讨论，在事务模型、动态依赖管理和服务生命周期三个方面的扩展进行了详细的阐述，并选择将 Service Mesh 架构作为底层架构，完成模型的扩展。

第四章介绍支持事务一致性的微服务动态更新系统实现。首先，给出系统实现的设计目标和准则。然后结合上述的扩展方案，在具体的开源实现框架 Istio 之上对相关的模型进行了模块化开发，分别详细介绍了各个扩展模块的内部结构设计与实现，以及系统运行时的模块交互关系。

第五章案例分析与性能评估。结合具体的应用案例，通过实验来对不同的更新算法在安全性、及时性和干扰性三个方面进行性能的比较评估，并对实验结果进行分析。

第六章对本文的工作进行总结，并对未来的研究工作进行展望。

第二章　相关工作

本文主要研究针对分布式微服务系统的一致性动态更新，与本文的相关工作包括: 软件动态更新的相关算法，以及现有主流的微服务框架，和相关微服务架构在动态更新方面的支持。下面将具体从这两个方面展开介绍。

2.1　软件动态更新技术

随着网络环境和用户需求的不断变化，运行时的软件系统无法在设计时就考虑到所有的功能，因此软件系统总是需要不断地进行动态更新演化，以满足多元的功能和性能需求[26]。相比较与传统的离线更新方式，动态软件更新(DSU)[27] 要求在软件系统不停止服务的状态下，使用新版本的软件来取代目标软件的旧版本，从而使软件系统达到增加功能或修复错误等目的。在动态更新的过程中，软件系统必须要确保程序始终处于运行状态，保证用户的请求不会被阻塞，而且需要满足一定的约束条件。因此软件系统的动态更新变得十分困难，面临诸多挑战[28]:

- **安全性 (Safety)** 动态更新需要选择合适安全的更新时机，保证更新发生的过程中以及更新完成后，系统中运行的请求都不会出现不一致的情况。
- **及时性 (Timeliness)** 系统从收到动态更新请求到更新完成所耗费的时间，要求动态更新操作应尽快完成。
- **干扰性 (Disruption)** 系统在处理动态更新请求的状态下，相较于没有动态更新请求的状态下，对正常处理业务请求所额外造成的干扰程度，即响应时间的增加。要求动态更新算法所带来的干扰性应尽量减小。

动态更新的基础是保证安全性，关键在于如何确定安全的更新时间点。最基本的要求是，当目标更新服务没有处理请求，同时也不存在于任何一个用户请求的分布式调用中，那么此服务处于一个安全的更新时间点。为了更好地对相关算法进行描述，我们在此引入事务的相关概念。

定义 2-1 (事务) 一个事务表示在某一服务上执行且在有限时间内结束的一系列操作。在某一服务上运行的事务 T 可以向其它服务发起调用，从而在目标服务生成一个新的事务 T'，前后两个事务 T 和 T' 我们分别称为父事务和子事务，记为 $sub(T, T')$。由系统外部调用生成的事务称为根事务。事务 T 运行所在的节点记为 h_T。

定义 2-2 (分布式事务) 根事务及其所有的子事务统称为一个分布式事务，使用对应的根事务符号 T 进行标识。分布式事务 T 包括的所有拓展事务集 $ext(T)$ 表示为：$ext(T) = \{x | x = T \vee sub(T, x)\}$

定义 2-3 (动态更新的安全性) 一次动态更新满足安全性，当且仅当不存在一个分布式事务的某两个子事务在同一个微服务的不同版本上运行，即只能全由微服务的旧版本进行处理或全由微服务的新版本进行处理。

为了解决上述动态更新所面临的挑战，现有工作已有不少的研究。本节主要介绍三种：Krammer 和 Magee 提出了 Quiescence[12] 算法。但 Quiescence 算法对系统造成的干扰极大，Vandewoude 等则在其基础上提出了 Tranquility[22] 算法。针对 Tranquility 只能保证局部一致性而不是全局一致性的问题，马晓星等提出了 Version Consistency[23] 算法，此算法通过维护服务运行时的动态依赖关系，对安全点提出了新的定义，从而较好地满足了动态更新过程中的对于安全性的要求。

2.1.1 软件系统实例

为了更好地阐述说明这些动态更新算法在实际更新时的原理，在具体介绍三种算法之前，本节先引入一个简单但不失一般性的图像拍摄转换系统，静态依赖图如 2-1。该系统由四个服务组成，分别是 Portal、Camera、Viewer 和 Convertor。用户可通过 Portal 发起一次图像拍摄请求，收到请求的 Portal 转而调用 Camera 进行具体的拍摄，拍摄所得结果经 Convertor 转换后返回。然后 Portal 请求 Viewer 服务，Viewer 服务则进一步请求 Convertor，Convertor 服务负责将转换后的数据进行解码并返回。为了加快传输转换速度，Convertor 负责将来自 Camera 的数据编码，相应的，将来自 Viewer 的数据解码。

系统运行的时序图如 2-2。由用户调用在 Portal 服务上首先开始根事务 T_0，T_0 在 Camera 上发起子事务并得到拍摄数据，为了能够快速传输数据，

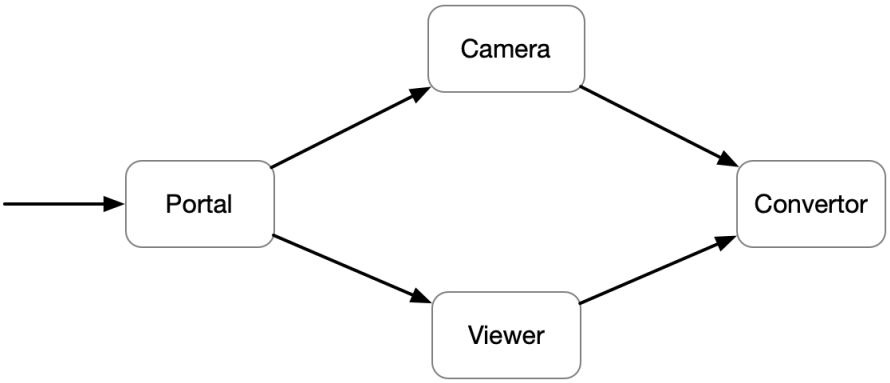

图 2-1: 示例系统静态依赖图

Camera 请求 Convertor 对数据进行转换，即子事务 T_2。Camera 将得到的内部表示形式的数据返回给 Portal。Portal 则在 Viewer 上发起子事务 T_3，并向其传输数据。Viewer 在 Convertor 上发起新的子事务 T_4，并将数据进一步传输，最终将解码后的数据结果返回，整个根事务 T_0 结束。

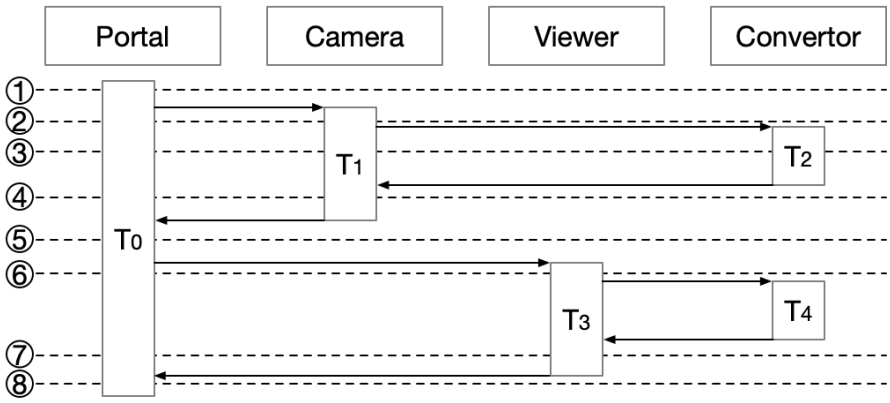

图 2-2: 示例系统时序图

在该软件系统实例中，假设希望提高数据的传输转换速度，系统需要对 Convertor 服务进行动态更新。若动态更新操作发生在时刻④和⑥之间，那么动态更新前的事务由旧版本的 Convertor 进行编码，而更新完成之后的事务由新版本的 Convertor 进行解码。若新旧版本的编解码算法不可兼容，便会导致系统出现不一致的情形。因此为满足安全性的需求，动态更新算法需要保证，不存在一个分布式事务的某两个子事务在同一个微服务的不同版本上运行。

2.1.2 Quiescence 算法

对于一个分布式系统，通常由若干个节点组成，它们之间相互请求调用从而形成相应的依赖关系。此处节点为泛化代称，在具体的应用场景中，节点有具体的表现形态。当我们需要对此分布式系统中的某个节点执行动态更新操作时，包括节点的创建、删除等，便有可能导致系统的静态依赖关系或目标节点的内部状态发生变化，从而进一步地影响运行时事务的正确性。因此，为了保证动态更新过程中系统的安全性，Kramer 和 Magee[12] 提出了 Quiescence 动态更新算法，算法中重点为节点定义了 Active 和 Passive 状态：

- **Active** 处于该状态的节点可以主动地发起事务、接受事务并为其他事务服务。
- **Passive** 处于该状态的节点不能主动地发起新事务，可以继续接受并为其他事务提供服务，但需要保证：(1) 该节点当前不处于自身发起的事务中 (2) 该节点不会发起新的事务。

当某个节点需要执行动态更新操作时，节点的状态可以在 Active 和 Passive 之间进行转换，具体的转换关系如图 2-3。由前述的 Passive 状态定义可知，即使节点处于 Passive 状态下，执行动态更新也不能保证系统的安全性，因为此时节点有可能正在处理由其它节点所发起的事务。因此作者提出了更严格且保证安全性的约束，即 Quiescence 状态：

- 该节点当前没有参与到自身发起的事务之中
- 该节点不会发起新的事务
- 该节点当前未为其他事务提供服务
- 该节点无需为其他节点已经发起或者是将来会发起的事务服务

前述对于 Quiescence 状态的定义中，前两点与 Passive 状态定义一致，要求当前节点处于 Passive 状态。而后两点要求本节点不能为其他节点已经发起或者是将来会发起的事务服务，即表示依赖于本节点的其他节点也应处于 Passive 状态。若某一节点 Q 想要达到 Quiescence 状态，那么对于集合 Passive Set $PS(Q) = \{Q\} \cup \{$ 直接依赖于 Q 的节点集合 $\}$，只有 $PS(Q)$ 中的所有节点均

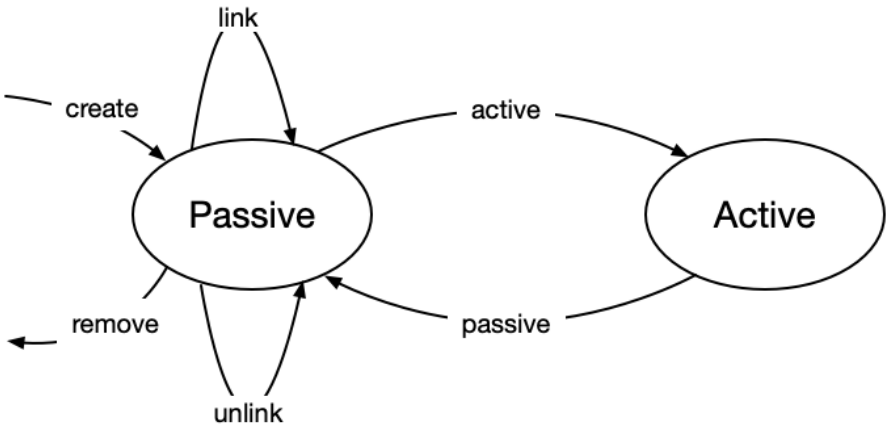

图 2-3: Quiescence 节点状态转移图

达到 Passive 状态，节点 Q 才能达到 Quiescence 状态。

对于 $PS(Q)$ 的定义仅仅考虑了一层依赖关系，更为全面的定义为集合 Enlarged Passive Set $EPS(Q)$：

- $PS(Q)$ 中的所有节点均属于 $EPS(Q)$
- 所有依赖于 $PS(Q)$ 的节点

综上，若某一节点 Q 想要达到 Quiescence 状态，只有当 $EPS(Q)$ 中所有的节点均达到 Passive 状态时，节点 Q 才能达到 Quiescence 状态。

在图2-1的示例中，对应的 Convertor 服务需要等到 Portal、Camera、Viewer 服务上运行的事务全部结束后才可达到 Quiescence 状态。可以发现，此时的 Portal、Camera、Viewer 服务均处于 Passive 状态，不能发起事务对外提供服务。因此 Quiescence 算法虽然可以确保动态更新过程中的安全性，但同时给系统带来了极大的干扰[29]。

2.1.3 Tranquility 算法

Quiescence 算法对于安全点的定义过于严格，给运行中的系统带来了极大的干扰。针对此问题，Vandewoude 等在 Quiescence 的基础上提出了 Tranquility 算法[22]，并利用该算法在 DRACO[30] 中间件上实现了一个动态更新系统。Tranquility 算法中继续沿用了 Quiescence 中事务、Active 状态和 Passive 状态的相关定义，重新为节点定义了 Tranquility 状态。节点想要达到 Tranquility 状态需要满足条件：

- 该节点当前没有参与到自身发起的事务中
- 该节点不会发起新的事务
- 该节点当前没有处理任何事务
- 该节点的相邻节点不存在过去使用过此节点并且将来还将使用此节点的事务。

在 Tranquility 的定义中，前两点同样与 Passive 状态定义一致。而后两点定义均弱于 Quiescence 的定义，因此满足 Quiescence 状态的节点必然满足 Tranquility 状态，但反之则未必。不同点在于，Tranquility 算法只要求目标节点处于 Passive 状态，且与其相邻的节点也处于 Passive 状态即可。它不需要考虑运行于其它更多节点的事务运行情况，只要相邻的节点不存在使用过目标节点且将来还会使用目标节点的事务，目标节点就可以达到 Tranquility 状态。

在图 2-2中，当 Camera 将处理结果返回给 Portal 后，系统运行到时刻⑤，此时 Camera、Viewer 和 Convertor 均处于 Passive 状态，因此 Convertor 满足 Tranquility 状态的定义，系统判断可进行更新。但如果此时对 Convertor 进行版本更新，便有可能因为新版本的 Convertor 与旧版本存在不可兼容性而出现系统不一致的问题。因此 Tranquility 算法相比较于 Quiescence 算法，虽然可以减少对系统造成的干扰，但是无法保证动态更新的过程中系统的安全性问题。

2.1.4 Version Consistency 算法

已有的一些动态更新算法或依据静态依赖信息[12]，或依据动态依赖信息[22,31,32] 来保证本地一致性。而马晓星等提出了 Version Consistency 算法[23] 则对版本一致性的概念进行定义，用于解决了分布式系统下分布式事务的全局一致性问题。

具体来说，Version Consistency 中对于一次动态更新可使用如下的元组进行表示：$< \Sigma, \omega, \omega', T, s >$。其中，$\Sigma$ 表示系统的静态依赖配置，ω 表示旧版本的目标集合 (即需要动态更新的节点集合)，ω' 表示新版本的目标节点集合，T 表示系统状态转移工具，s 则表示执行动态更新前对应的系统状态。因此，系统执行动态更新后，对应的状态可表示为：$s' = T(s)$。

定义 2-4 (版本一致性) 对于某一动态更新配置 $< \Sigma, \omega, \omega', T, s >$，事务 T 是满足版本一致性，当且仅当 $\nexists T_1, T_2 \in ext(T) \mid h_{T_1} \in \omega \wedge h_{T_2} \in \omega'$。若当前动态更新配置 Σ 中所有的事务都满足版本一致性，那么称此动态更新是版本一致的。

以图2-2为示例，系统在时刻②之前和时刻⑦之后，对应的 $ext(T_0)$ 中的事务均满足2-4的定义，因此都可以进行动态更新。但若系统选择在时刻⑤对 Convertor 进行动态更新，此时 $h_{T_2} = Convertor, h_{T_4} = Convertor'$，且 $T_2, T_4 \in ext(T_0)$，显然不满足2-4的定义，此次动态更新将使得系统出现版本不一致的情况。

版本一致性虽然定义了动态更新的条件，但该条件在实际应用中需要全局中心化的信息，很难进行操作和检测判断。因此 Version Consistency 算法使用节点运行时的动态依赖关系，很好地解决了更新状态条件难以判断的问题。在 Version Consistency 算法中，动态依赖关系由具体的动态依赖边进行定义表示：

定义 2-5 (动态依赖边) 使用 $C \xrightarrow[T]{future} C'$ 表示在根事务 T 中，节点 C 将来可能会用到节点 C'；使用 $C \xrightarrow[T]{past} C'$ 表示在根事务 T 中，节点 C 过去曾经使用过节点 C'。

系统在运行时，将依据每一个根事务 T 的运行状态，动态地在节点间执行依赖边的添加和删除操作。同时要求在对依赖边进行操作的过程中，应满足如下的约束：

定义 2-6 (配置有效性) 使用动态依赖边来表示节点间的动态依赖关系信息是正确有效的，当且仅当对于运行时 $future/past$ 边的创建和删除操作满足以下规则：

- 事务 T 从开始到结束，T 所在的节点 C 必须存在边 $C \xrightarrow[root(T)]{future} C$ 和 $C \xrightarrow[root(T)]{past} C$，其中 $root(T)$ 表示事务 T 的根事务；
- 如果节点 C 和 C' 之间没有相应的静态依赖关系，那么不能建立 C 到 C' 的 $future/past$ 边；
- 对于边 $C \xrightarrow[T]{future} C'$，必须在 $ext(T)$ 的第一个子事务开始前建立，且必须在确定 C 不会再在 C' 上发起子事务后才能够被删除；

- 对于边 $C \xrightarrow[T]{past} C'$，必须在由 C 在 C' 上发起的属于 $ext(T)$ 的子事务结束之前建立；

基于动态依赖边，Version Consistency 算法提出了 Freeness 更新判断条件，使得系统只需针对目标更新节点上相关动态依赖边的知识，就可以很方便地判断当前时刻目标更新节点是否满足动态更新的条件。

定义 2-7 (Freeness) 对于某个配置 Σ，如果当前不存在同时指向节点 C(或者节点集合 ω) 的边 $C' \xrightarrow[T]{future} C$ 和边 $C' \xrightarrow[T]{past} C$，则称该节点 C(或者节点集合 ω) 对根事务 T 是 *Freeness* 的。若 C(或者节点集合 ω) 对于此配置 Σ 中的所有事务均是 *Freeness* 的，则称 C(或者节点集合 ω) 在配置 Σ 中是 *Freeness* 的。

图 2-4展示了系统中某个根事务运行时的动态依赖关系，其中 $(a),(b),(c),(d)$ 分别对应图 2-2中的时刻①、时刻③、时刻⑤、时刻⑦。利用定义 2-7可以发现，Convertor 在时刻①和时刻 178 满足 Freeness，不存在标记为 T_0 的一对 *future* 和 *past* 边指向 Convertor，可进行动态更新；而在时刻③和时刻⑤则同时存在标记为 T_0 的 *future* 和 *past* 边指向 Convertor，因此 Convertor 不可以进行动态更新。

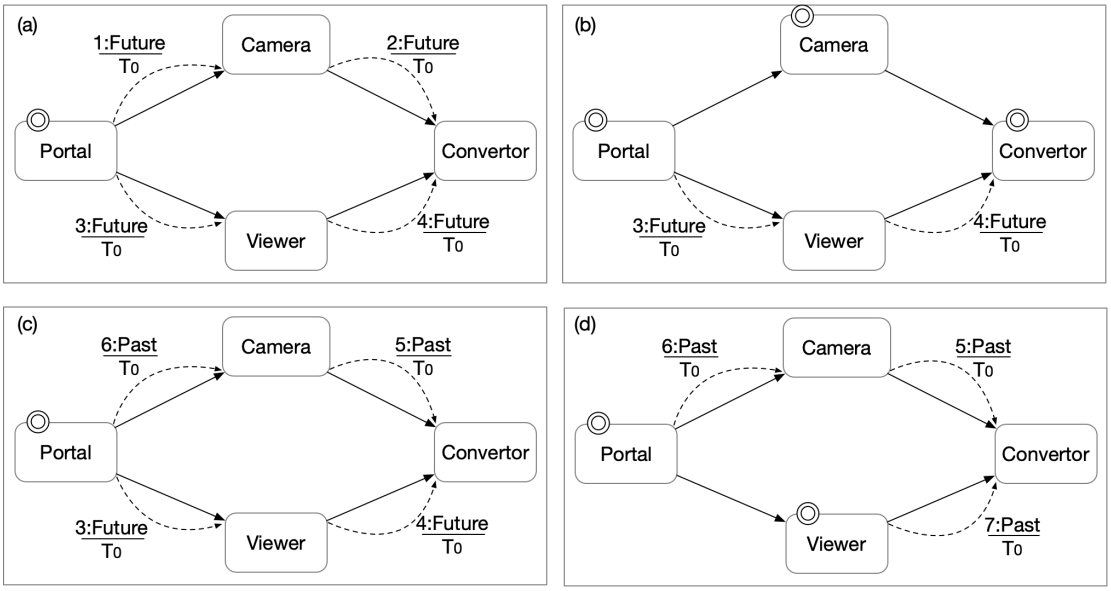

图 2-4: 系统运行时动态依赖关系示例

上面介绍的三种具体软件动态更新算法分别对安全更新点给出了自己的定

表 2-1: 动态更新算法比较

更新算法	安全性	及时性	干扰性
Quiescence	安全	慢	高
Tranquility	不安全	快	低
Version Consistency	安全	快	低

义，即节点分别需要满足 Quiescence 状态、Tranquility 状态和 Freeness 状态。判断的条件和使用的方式也略有差别，例如 Quiescence 算法主要利用的是事务和节点间的静态依赖信息，而 Version Consistency 在其基础上进一步地使用了节点间运行时刻的动态依赖信息，使得对于安全点的判断更为精准。对于前述的动态更新所面临的诸多挑战，如安全性 (Safety)、及时性 (Timeliness) 和干扰性 (Disruption)，三种算法的表现也不尽相同。具体对比如表 2-1所示。

2.2 微服务架构

基于微服务理念进行分布式应用程序的开发，已经成为多数平台的通用指导模式[33,34]。应用微服务架构的关键之一在于对复杂的业务系统进行拆分，使得拆分出来的服务能够相互保持逻辑上的独立，从而保证服务可由不同的团队进行负责，快速地完成开发、部署和更新等操作。拆分步骤必然导致系统中的模块数量变多，实际上是将原先业务系统内部维护的复杂度，转换成了模块与模块之间的通信与管理监控的复杂度。

随着微服务架构的发展，为了简化开发流程，出现了典型的类库，如 Netflix OSS[35] 套件，和相关的开发框架，如 Spring 和 Dubbo。框架实现了分布式系统通信的相关通用功能，包括服务发现、负载均衡等，使得开发人员通过编写较少的代码，即可完成分布式微服务系统的开发，有效地提高了开发人员的工作效率。使用相关的开发框架虽然简易，但也存在不少的局限性问题。因此，本节将主要介绍较为主流的微服务框架：Spring、Dubbo，以及处理服务间请求的基础架构层技术：服务网格 (Service Mesh)。

2.2.1 Spring

Spring 是一个于 2003 年兴起的轻量级 Java 开发框架，由 Rod Johnson 提出，为了解决企业应用开发的复杂而诞生[36]。Spring 框架如图 2-5所示，是一种分层架构，共分为如下具体的模块：

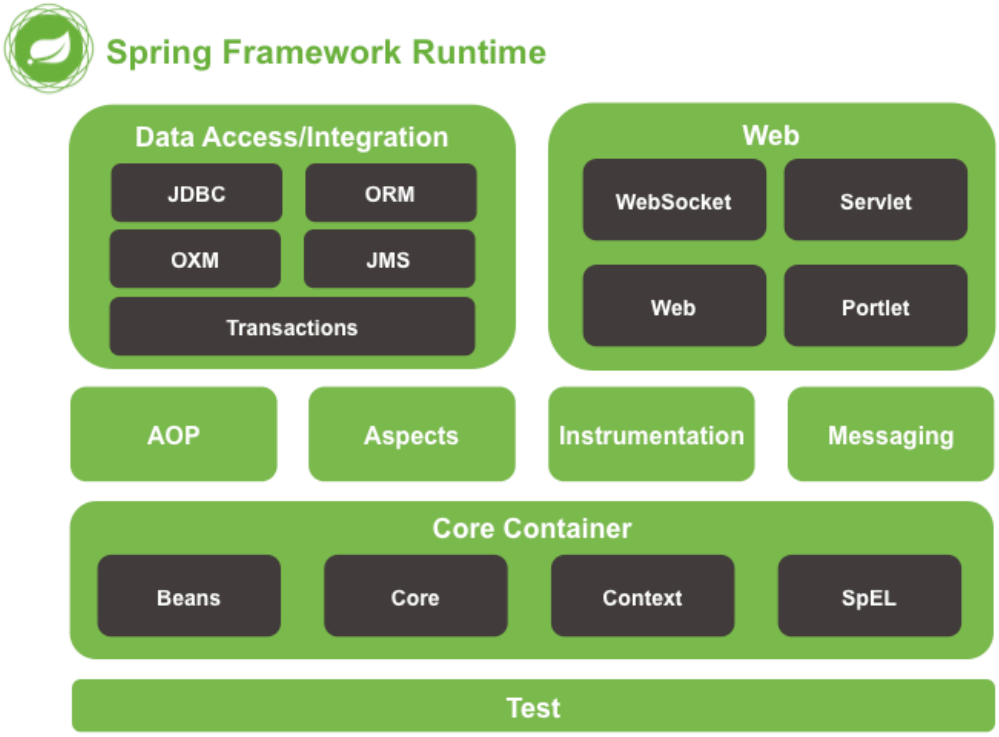

图 2-5: Spring 架构图

- **Core Container** 核心容器提供 Spring 的基本功能，使用 BeanFactory 以控制反转 (IOC) 的模式来进行依赖注入，以 bean 的形式组织管理 Java 程序中的各个组件及其关系。
- **Data Access/Integration** 提供与多种数据框架集成的相关功能，包括 jdbc 框架等，实现事务控制的相关模块。
- **Web** 提供基础的 Web 集成功能、相关的架构模块和与前端进行双向通信的协议支持。
- **AOP** 提供多种 AOP 实现方式的支持，包括 AspectJ 等。

•**Instrumentation** 提供了类工具和类加载器的相关实现的支持。

•**Messaging** 对消息架构和协议的支持。

•**Test** 提供对单元测试和集成测试的支持。

分层架构允许开发人员自由地选择需要的组件模块，同时又为 J2EE 应用程序开发提供完整的集成框架。作为一个较为成熟的开源开发框架，Spring 的核心在于控制反转 (Inverse of Control, IOC) 和面向切面编程 (Aspect Oriented Programming, AOP) 两部分。其中，控制反转的思想在于利用 Spring 核心容器进行依赖注入，一方面实现了资源的可配置，另一方面很好地降低了使用资源双方的耦合度；面向切面编程则使得开发人员不修改框架源代码，却能够实现想要的扩展功能，并形成可复用模块。概括来说，Spring 的优点主要包括：

- **轻量级** Spring 的体积和管理开销都很小。
- **简化开发** 将对象间的依赖关系交由 Spring 控制，降低了业务对象替换时的复杂度，避免硬编码所造成的过度耦合。
- **支持 AOP** 允许将通用任务如日志等进行集中式模块管理。
- **测试** 可以方便地使用非容器依赖的编程方式来完成测试工作。
- **集成** 提供了对各种成熟框架的直接支持，降低了开发人员的使用难度。

2.2.2 Dubbo

Dubbo 是 Alibaba 开源的 Java 分布式服务框架，致力于提供高效的 RPC 远程服务调用方案以及服务治理方案，被广泛地应用于 Alibaba 的各方面业务中[37]。Dubbo 的架构与 Spring 类似，同样属于分层架构，使得各个层之间的耦合度尽量地减少。具体的服务调用关系流程如图 2-6所示：

- 服务提供方运行容器并向注册中心进行注册；
- 服务调用方向注册中心进行订阅，获取服务提供方的地址信息并进行调用；

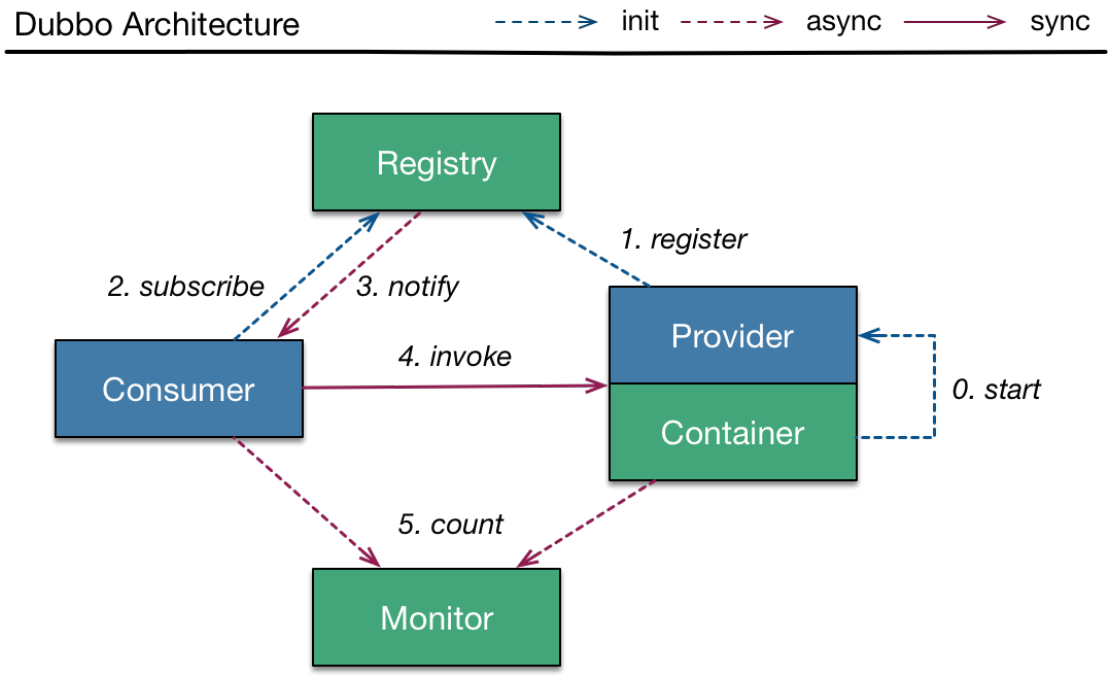

图 2-6: Dubbo 服务调用关系图

•Monitor 监控模块负责异步地接收调用过程中的相关信息。

Dubbo 主要提供了服务注册和服务调用的相关功能，服务注册方面使用的自动注册和管理机制，使得服务调用方不需要硬编码服务方地址，提供多种注册中心的选择。服务调用方面透明化远程方法的调用，自动为服务调用方提供软负载均衡及容错机制。Dubbo 架构具有连通性、健壮性和伸缩性等特点，有助于增强服务治理，使得传统的单体应用程序可以逐步平滑地重构为可扩展的分布式微服务架构。

2.2.3 Service Mesh

虽然开发框架本身对用户屏蔽了绝大多数通用功能的实现细节，但开发人员同样需要耗费大量时间与精力去学习掌握框架本身。同时，开发框架通常局限于一种或几种特定的编程语言，业务服务往往得随着框架服务的升级而升级。

Mogan 于 2017 年在其文章中提出 Service Mesh 的概念[38]。Service Mesh 考虑使用边车模式 (Sidecar)，为每一个服务实例部署一个对应的 Sidecar 实例，开发人员只需关注核心的业务逻辑代码，而由 Sidecar 实例来负责实现服务发现、

负载均衡、认证授权等通用性功能[25,39]。Service Mesh 的架构图如图 2-7所示：

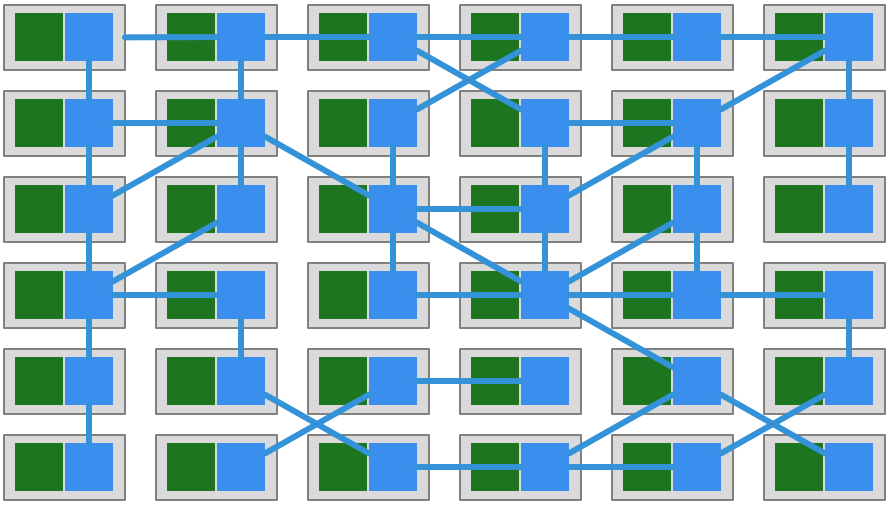

图 2-7: Service Mesh 架构

随着微服务数量的增加，部署的 Sidecar 实例也逐渐增加，服务与 Sidecar、Sidecar 之间的通信便形成了一个错综复杂的网络，也即是服务网格 (Service Mesh) 名称的由来。Service Mesh 插入的 Sidecar 实例通常为轻量级的网络代理，资源占用消耗较小，对部署的服务实例保持透明性。此外，Service Mesh 将原本通过开发框架实现的通用性功能，从应用服务代码中解耦出来，降低耦合度的同时也可保证 Service Mesh 的组件可以独立于业务服务单独升级。

2.2.4 微服务架构与动态更新

随着微服务架构的流行，相关动态更新的工作也有了较多的关注，目前已经出现了不少的研究成果。

在 Spring 框架中，支持两种热部署 (Hot Deploy) 的方式，热部署主要希望在修改应用代码时，不重启停止应用即可完成代码的更新。此功能类似与动态更新 (动态配置)，但是在具体的实现上，热部署主要基于 Java 的动态类加载机制，不能很好地定义安全点，无法保证应用更新过程中的正确性和一致性。例如在 Spring 框架中，Spring Load 是一个用于在 JVM 运行时重新加载类文件更改的 JVM 代理，允许用户动态对某个方法字段等进行添加、修改、删除。当使用 Spring Load 启动程序后，系统会自动监视对应的文件，当目标文件发生改动时，系统便会重新加载类文件，无需重启服务。Devtools 的功能类似，原理在于使用两个类加载器：不改变的类包括第三方 jar，由 base 类加载器加载，而

目标正在开发的类由 restart 类加载器加载。当检查到类发现变化时，应用进行重启，restart 类加载器直接丢弃重建，而 base 类加载器则继续留用。由于 base 类加载器已经可用且已填充，因此本次应用的重启会比普通的重启速度要快得多。这两种热部署的实现方式均考虑的是目标文件的相关变化，但并不关心更新过程中运行时事务的相关状态，因此无法保证更新的正确性。

微服务架构具有耦合度低、独立部署扩展等特性，因此它通常在相关的分布式云服务平台上进行部署运行并使用其底层框架的能力[18]。例如 Spinnaker[1], AWS CodeDeploy[2], IBM UrbanCode[3]都支持对部署于异构 Paas 平台上的微服务进行更新部署。更新的执行主要依靠运维人员对底层操作进行编写，并形成一系列的流水线操作，最后以脚本的形式来运行。这种更新方法具有较大的局限性，一方面它要求运维人员撰写出正确且有效的运行脚本，这通常需要耗费较长的调试时间；另一方面，脚本的执行通常不具有幂等性，如果基于脚本的更新过程出现异常的话，往往需要回滚并重启整个更新过程，并小心翼翼地保存相关的状态。push2cloud[4]则运行对更新的目标系统结构进行描述，并支持部署在其上的微服务进行更新。但目前仅支持单站点的微服务部署，更新策略同样需要定义成对应平台的底层流水线操作。

2.3 本章小结

本章主要介绍了软件动态更新的相关算法，以及现有主流的微服务框架、相关微服务架构对动态更新方面的支持。在软件动态更新的相关算法中，本章重点介绍了 Quiescence、Tranquility 和 Version Consistency 三种算法，它们分别在各自的算法中对动态更新的安全点进行定义，以及安全点的相关判断条件。同时还对这三种算法在安全性 (Safety)、及时性 (Timeliness) 和干扰性 (Disruption) 进行了具体的比较。而在现有主流的微服务框架中，主要介绍了 Spring、Dubbo 框架，以及新一代的 Service Mesh 技术，对其架构和相关特性进行了说明。最后介绍了 Spring 框架对于热部署机制的两种支持方式，以及相关分布式云服务平台对微服务架构在动态更新方面的支持。

[1]Spinnaker. https://www.spinnaker.io/
[2]AWS CodeDeploy. https://aws.amazon.com/codedeploy/
[3]IBM UrbanCode. https://developer.ibm.com/urbancode/
[4]Push2Cloud website. https://push2.cloud/

第三章　支持事务一致性的微服务动态更新方法

动态更新在实际应用场景中有着较大的需求，但当前的微服务架构对更新过程中的安全性考虑不足，缺乏对服务正确且快速更新的支持。本章首先对微服务动态更新的挑战进行讨论，分析当前的微服务架构在动态更新支撑方面的不足。然后基于上述讨论与分析，针对性地在事务模型、动态依赖管理和服务生命周期三个方面提出了相应的扩展和实现方案，并选择在服务网格 Service Mesh 之上进行相关模型的扩展，给出一种支持事务一致性的微服务动态更新实现方法 VCIM(Version Consistency In Microservices)。

3.1　需求与挑战

为了在微服务架构之上，支持相关的动态更新算法，我们需要考虑：

- **事务模型**

  前述动态更新算法中均与事务模型相关，依据前述对于事务的定义 2-1，是指服务上执行的一系列有意义的动作。而当前相关的微服务框架所支持的事务主要指传统数据库中的 ACID(Atomicity, Consistency, Isolation, Durability) 事务，如 Spring 框架中的 @Transactional 注解，无法为更新算法中的事务提供有效的语义支持。事务是服务运行时处理请求的基础，事务的状态一定程度上反映当前服务所处的状态，是动态更新算法判断系统状态的重要依据。因此，我们需要为微服务系统定义合适的事务状态模型，为更新算法提供运行时事务信息。

- **动态依赖管理**

  服务间的静态依赖关系可以保守地对系统运行时进行描述，但是无法保证动态更新的及时性。基于运行时服务间的动态依赖信息，动态更新算法可以有效地对安全更新点进行判断，在保证动态更新一致性的前提下

，及时地完成服务的更新。因此，我们需要维护服务间的动态依赖信息，为更新算法提供基础保障。

- **服务生命周期**

  现有的微服务系统通常仅考虑服务的基本状态，包括初始状态、运行状态、终止状态等。显然单一的运行状态并不能很好地刻画目标服务在整个更新过程中的状态。具体来说，除了相关的基本状态，运行状态的服务从收到更新请求到最后更新结束，还需包括更多拥有特定意义的状态，如更新开始状态、版本切换状态、结束撤销状态等。因此，我们需要对更细粒度的服务生命周期及其状态跳转进行定义，保证更新前后的一致性和持续性。

与此同时，支持动态更新的微服务架构还面临着不少挑战。该架构还需尽量保证：

- 动态更新开始前，用户可以对目标更新服务进行指定，且更新前后需保证系统的安全性 (定义 2–3)，不会出现因更新而导致的不一致现象。
- 能够尽快地完成动态更新，有较好的及时性，且更新过程不会对整个系统，包括其它非更新服务带来较大的干扰，保持良好的用户体验。
- 内部的相关模块具有较低的耦合度，从而保证不会对基于此架构开发的应用造成较大的影响，可以适配实际不同应用场景下的微服务系统，具有较好的兼容性。

基于上述的讨论和分析，为使当前的微服务系统支撑具体的动态更新算法，我们需要对事务模型、动态依赖管理、服务生命周期等进行扩展实现，并将其整合到相应的微服务架构中。

3.2 支持微服务动态更新的方法

实际应用场景下，不同的微服务系统将由不同的团队来进行开发，在通信方式、编程语言、运行环境等方面都存在较大的差异。而上述讨论均是针对微服务系统中将要扩展的模块来进行，与具体的微服务业务逻辑无关，应保证两者之间的松耦合性。因此，我们选择 Service Mesh 支撑微服务的底层架构，在

其上进行相关模块的实现扩展和整合，给出一种支持事务一致性的微服务动态更新实现方法 VCIM(Version Consistency In Microservices)。

使用 Service Mesh 作为微服务的底层实现架构的优势主要体现在：

- Service Mesh 使用流量代理的方式来对具体的微服务实例进行管理，因此相关的扩展模块可以使用相同的方式来进行实现，在保证透明性的同时，开发人员的业务代码具有较小的侵入性。
- Service Mesh 允许开发人员选择自身偏好的编码语言，只需保证微服务间的网络可达性即可，可适配于实际应用场景下不同类型的微服务系统。
- Service Mesh 架构可保证微服务与动态更新的管理模块解耦开，两者都可独立地完成升级。

3.2.1 事务模型与依赖管理

运行时事务信息和服务间的动态依赖信息是更新算法判断服务能否被更新的基础依据。动态依赖信息需要同特定的事务相绑定才有具体的表征意义，因此 VCIM 扩展实现了事务依赖管理器，统一负责对运行时事务和动态依赖进行管理。

VCIM 首先对事务的状态进行了定义，如图3-1所示。具体来说，相关的状态包括：初始化、运行、依赖变更、结束。

- **初始化** 当服务接收到外来请求并被流量代理所拦截时，生成相应的事务，且事务进入初始化状态。此时的事务还未被真正的服务实例处理，请求将被进一步地转发。
- **运行** 当服务实例开始执行由流量代理转发而来的事务时，事务进入运行状态，表示事务开始执行具体的业务逻辑。
- **依赖变更** 事务在运行时有可能对一个或多个依赖服务发起远程调用，并在其之上发起相应的子事务。当远程调用返回时，服务间的动态依赖关系便可能发生变化，当前服务已经使用过该依赖服务，以及当前服务在将来是否还会调用该依赖服务。处于依赖变更状态的事务仍处于运行状态，若后续还存在相关的远程调用，则此事务状态保持不变，否则，等

待事务执行完成，将转入结束状态。

- **结束** 当事务执行完成时，将结果经流量代理返回给具体的调用者。此时事务进入结束状态，表示一次事务完整的执行。

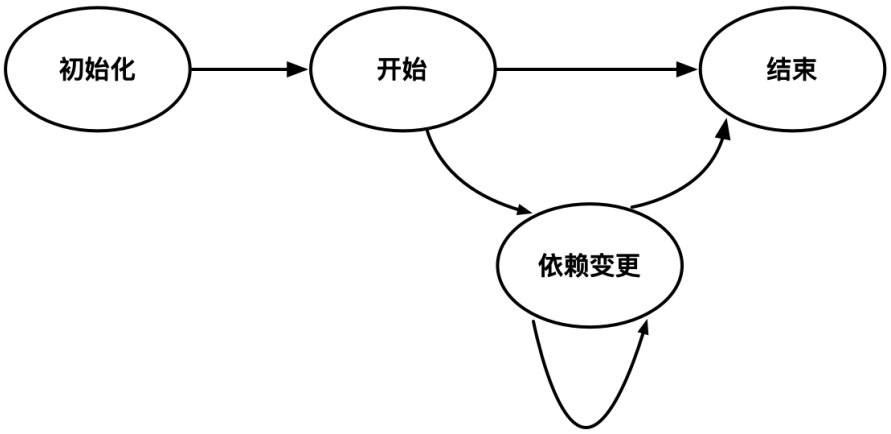

图 3-1: 事务状态转移图

分布式事务状态的定义与事务状态的定义保持一致，不同之处在于它由其所有的子事务的状态来共同表示。这些子事务的状态在分布式服务实例的本地进行维护，不存在中心化的存储。当系统在需要相关信息时，每个服务可借助于分布式事务标识信息，向依赖于自己的相关服务发起请求，获得全局的分布式事务信息状态。

事务运行时状态影响服务间的动态依赖关系，反之服务间的动态依赖关系也需要具体的事务来进行标识，两者密切相关，相互耦合。VCIM 通过对具体的程序进行分析，得到相对应的事务状态自动机，并利用此事务状态机来进行动态依赖信息的获取[40]。当本地事务运行时，事务的依赖会发生相应的变化，从而反映出服务间依赖关系的变化。该事务状态自动机将分析出在当前状态下，事务曾经调用过的服务集合 PAST_SET，以及将来有可能需要调用的服务集合 FUTURE_SET，依据此信息形成动态依赖关系，并在本地进行管理。

虽然 VCIM 从服务启动后就在不断地维护相关的事务和动态依赖信息，但可以看到 VCIM 对事务模型的定义较为精简，每个服务只需知晓和维护其本地事务的详细信息和对应的分布式事务标识，而无需在服务间进行事务上下文的同步工作，尽可能地减少对全局事务信息的维护，从而避免了服务间通信所带来的额外开销，并且达到了对相关模块减负的目的。

3.2.2 服务生命周期

Service Mesh 架构并未对运行于之上的服务的生命周期进行相关的约束，VCIM 需要对系统中服务的生命周期模型进行更细粒度的扩展定义。该生命周期模型应包括动态更新过程中以及结束后服务可能的状态，以及状态间如何完成跳转，形成一次完整的更新闭环流程，为系统提供多次连续更新的能力。因此，如何对生命周期模型进行合适的定义，是服务正确进行动态更新的关键因素。本节将对动态更新过程中，服务的生命周期模型进行具体的定义，包括相关的服务状态及跳转关系。

VCIM 允许用户在执行动态更新前指定需要更新的服务，称为目标更新服务，该服务可能对应于多个具体的服务实例。此处扩展的服务生命周期模型主要是针对于服务运行时 (Running) 状态的扩展，这里服务的状态将代表系统对外暴露的服务运行情况，是所有服务实例运行状态的总称，而不是特指某一个服务实例的运行状态。扩展后的服务生命周期模型如图 3-2所示，具体包括了：

- **Normal** 当服务初始启动就绪，或者某次更新完成后，服务将进入 Normal 状态。处于 Normal 状态下的服务和未扩展前正常运行状态下的服务一致，系统仅存在 (使用) 一个版本的服务实例来处理外来请求。

- **Deployed** 当用户想要对某个服务进行动态更新时，首先需要上线相应的新版本服务实例并等待其全部就绪，此时服务进入 Deployed 状态。此时系统中对应的服务包括新旧多个版本实例，但请求仅由旧版本的服务实例来进行处理。若此时未能确保所有新版本服务实例就绪，则后续随着服务状态的转变，将请求发往未就绪的实例时，便会导致异常，返回错误结果。

- **Prepared** 服务在完成 Deployed 状态之后，用户可指定目标更新服务，并向系统发送动态更新请求。如何正确地为用户请求选择合适的服务实例，取决于流量代理模块中的相关配置。为了能够保证流量在新旧版本的服务实例间进行切换，VCIM 首先需要对相关服务的流量代理模块添加相关的路由规则，完成更新的准备工作，服务进入 Prepared 状态。

- **Shifted** 当前述动态更新的准备工作完成后，VCIM 更新相关服务的流量

代理模块中的相关路由规则，进行流量的切换，服务进入 Shifted 状态。当存在外来请求向目标更新服务发起调用时，若该请求对应的分布式事务曾经在目标更新服务上发起过子事务，那么请求将被转发往旧版本的服务实例；否则，请求将被转发往新版本的服务实例。显然，处于此状态中的服务对于用户请求的处理满足定义 2-3中对于动态更新的安全性要求。

- **Freeing** 为了避免旧版本服务实例对于系统资源持续占有所造成的浪费，用户向目标更新服务发起撤销旧版本的请求，服务进入 Freeing 状态。在此状态下，目标更新服务将向依赖于自己的相关服务发起请求，并获得相关的事务和动态依赖信息。同时结合本地所维护的相关信息，若所有在目标更新服务上发起过子事务的分布式事务均不会再次向目标更新服务发起请求，同时目标更新服务的所有旧版本服务实例当前都没有在执行事务。则目标更新务满足定义 2-7中的 Freeness 条件，旧版本服务实例可以被正确撤销。

- **Updated** 处于 Freeing 状态的服务在满足 Freeness 条件后，便可执行旧版本实例的撤销操作，服务进入到 Updated 状态。在此状态下，VCIM 将对相关服务的流量代理模块进行路由规则的重配置和相关事务的标记清理工作。目标更新服务完成更新，回到 Normal 状态，完成一次完整的动态更新闭环流程。

在前述动态更新的过程中，VCIM 仅关注目标更新服务的状态变化情况。在特定状态下，目标更新服务需要向依赖于自身的相关服务发起请求，以获得动态依赖信息，但并不对相关服务的状态进行限定。这意味着使用 VCIM 方法扩展的服务生命周期模型具有较好的隔离性，将不同服务间的状态解耦开，有效地提高了更新的效率。

由于普通的微服务通常具有无状态、即插即用等特性，因此其生命周期较为简单，无法对动态更新中的相关语义进行理解。VCIM 通过对服务的生命周期模型进行扩展，很好地满足了动态更新过程中对于目标服务不同阶段的更新状态的需求。此外，动态更新中具体的版本更新切换算法和旧版本撤销算法，可在生命周期模型中的相应状态中进行具体的实现。

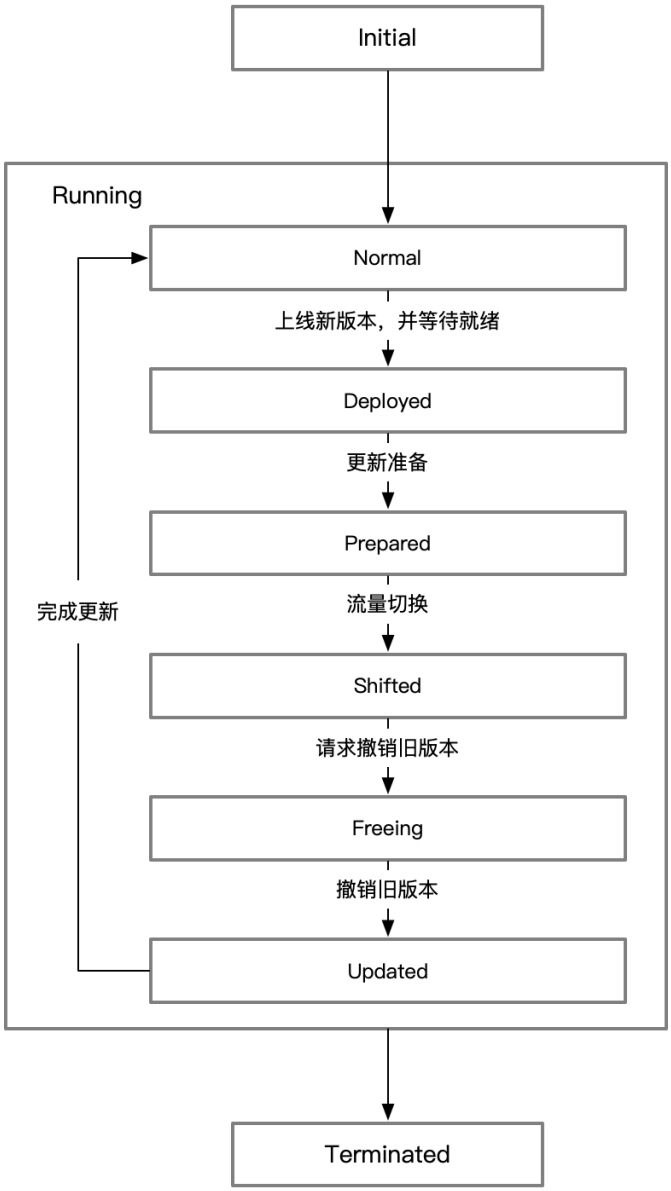

图 3-2: 服务生命周期

3.3 本章小结

本章主要对支持事务一致性的微服务动态更新方法 VCIM 进行了介绍。首先，对当前微服务架构支持动态更新所面临的挑战进行了探讨，对其不足和需求进行了具体的分析。然后针对相应的问题，有效地在事务模型、动态依赖管理和服务生命周期三个方面提出了具体的扩展方案。基于上述模型在 Service Mesh 上的扩展，给出对应的实现方法 VCIM。下一章将结合具体的微服务系统示例，对 VCIM 的具体系统实现进行详细的阐述和说明。

第四章　支持动态更新的微服务系统实现

在上一章中，VCIM 具体从事务模型、动态依赖管理和服务生命周期三个方面，对支持动态更新的微服务架构进行了扩展。本章将根据第三章中提出的 VCIM 方法，结合具体的微服务系统示例，介绍如何在开源框架 Istio[1]之上进行具体设计，实现支持动态更新算法的微服务系统。

4.1　系统设计概述

根据章节 3中的介绍，为使 Service Mesh 架构支持服务的动态更新，这里考虑使用与流量代理类似的方式，对事务和依赖管理进行扩展。同时，为了使用户可以方便地对系统执行动态更新等操作，在具体的应用系统之外，定义了额外的服务生命周期模型和更新控制模块，为应用系统提供相应的管理功能，实现了模块间耦合度的降低。

概括来说，支持动态更新的微服务系统设计目标和准则包括：

- 兼容当前 Service Mesh 以及微服务的相关标准。在完成动态更新功能的支持后，对任何正常开发部署的微服务应用都不应造成影响。
- 添加的相关模块应保证动态更新过程的安全性、及时性和低干扰性。即要求动态更新过程可以高效的完成，同时不会对系统中正常业务的请求造成干扰，破坏系统的一致性。
- 提供相应的动态更新管理接口，允许用户对系统中服务状态进行查询，和发起动态更新请求等操作，系统依据服务的相关状态，自动完成动态更新操作。

[1]Istio home page: https://istio.io/

4.2 系统扩展实现

本节将基于上一章提出的扩展方案，对相关的扩展模块进行实现，并且为用户提供相关的动态更新管理接口。下面将详细阐述事务依赖管理器模块 *TxDepManager*、服务生命周期管理器模块 *SvcLifecycleManager*，以及更新控制模块 *UpdateController* 的内部结构设计。

4.2.1 事务依赖管理器

为了使微服务系统能够对运行时事务提供相关的语义支持，维护服务间的动态依赖关系，定义了事务依赖管理器模块，并选择 Service Mesh 作为支撑微服务的底层实现架构。当前主流的 Service Mesh 实现方案是 Google、IBM 及 Lyft 合作开发的 Istio 开源框架[41,42]。Service Mesh 架构为每一个微服务实例均部署特定的网络代理，整个系统引入了大量独立运行的单机代理服务。为了更好地保障系统的整体稳定性，Istio 使用集中式的控制面板，提供了统一的管理运维入口，具体架构图如图4-1所示，其中：

- **控制平面 (Control Plane)** 负责监听底层平台数据源，作为配置中心，将实时的配置信息动态地发往数据平面，控制整个系统的通信等各方面行为；
- **数据平面 (Data Plane)** 自动注入的所有服务实例的流量代理总称，负责连接到控制平面并接收相关配置信息，然后对请求按配置进行处理，是配置规则的实际执行者；

为实现 VCIM 中对于事务与依赖的扩展模型，我们选择在数据平面对相关模型提供支持。扩展后的某个具体的微服务实例如图4-2所示，其中包括了：

- **流量代理** 拦截并接管所有的外来请求，生成相应的本地事务并添加或继承特定的分布式事务标识。流量代理模块一方面负责将请求分发至相应的业务逻辑模块，另一方面负责对前述的事务上下文信息进行传递。
- **事务依赖管理器** 负责事务与依赖的管理，保证在事务依赖发生变更时，对相应的服务集合 PAST_SET 和 FUTURE_SET 等信息进行更新保存，为 VCIM 提供运行时事务信息和动态依赖信息。

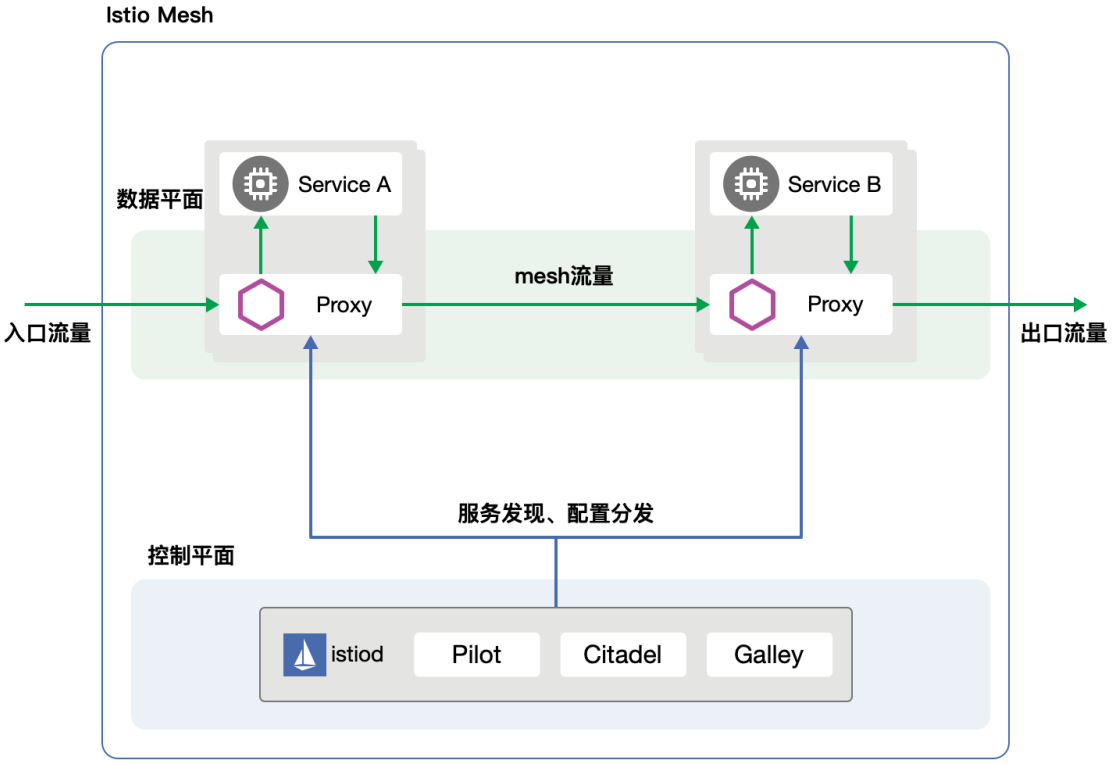

图 4-1: Istio 架构图

- **业务逻辑** 由服务的开发人员完成编写，从流量代理模块接收请求，执行真正的业务逻辑并返回结果，再由流量代理模块将结果进一步地返回给调用者。

其中，流量代理模块和事务依赖管理器通过 UDS[1](Unix Domain Socket) 进行通信，UDS 通过指定同一主机上的 socket 文件，不需要经过网络协议栈即可完成数据的传输，效率较高。同时相关的数据总是以单向的形式发送，因此不会对流量代理模块中的请求处理分发流程造成阻塞，保证了对系统具有较低的干扰性。与此同时，流量代理模块对业务逻辑保持透明，为其屏蔽了调用方的相关细节，开发人员只需关注于业务代码，可以很好地与其它模块解耦开。

4.2.2 服务生命周期管理器

想要对运行于 Istio 框架之上的服务进行动态更新，需要对原先服务的生命周期模型进行了扩展，定义了对应的服务生命周期管理器，提供监听服务状态、管理服务生命周期的基本功能。

[1]https://en.wikipedia.org/wiki/Unix_domain_socket

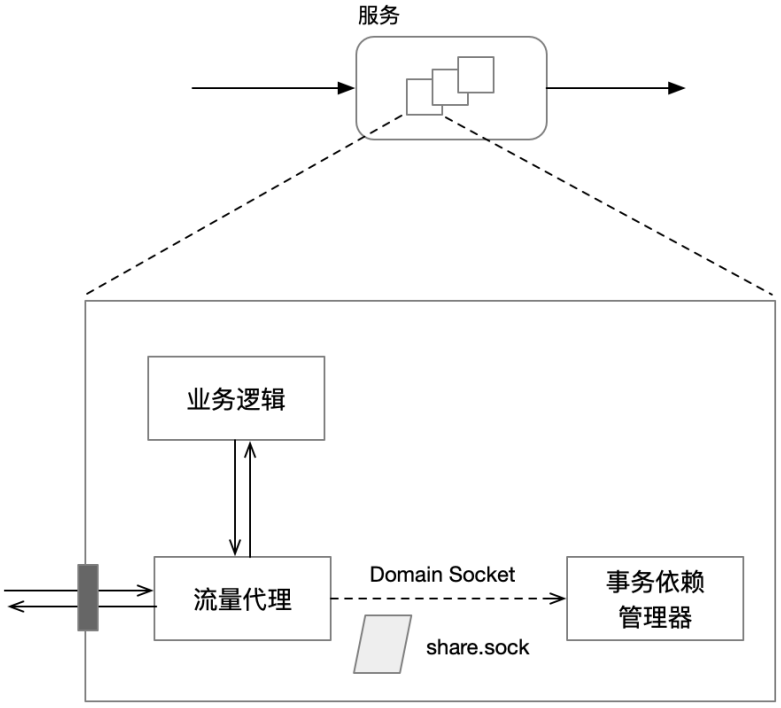

图 4-2: 微服务实例模型

当前 Istio 框架的运行环境与容器编排系统 Kubernetess[1]相耦合，服务以容器的形式运行在其之上。因此，服务生命周期管理器的实现主要利用 Kubernetes 中提供的自定义资源 CRD(Custom Resource Definition) 功能，其实现原理称为控制器模式，如图 4-3所示。用户通过声明式命令对期望状态进行描述，控制器对比期望状态与系统实际的运行状态，如果有出入则通过自定义的调谐逻辑 (Reconcile) 驱使实际状态调整为期望状态。

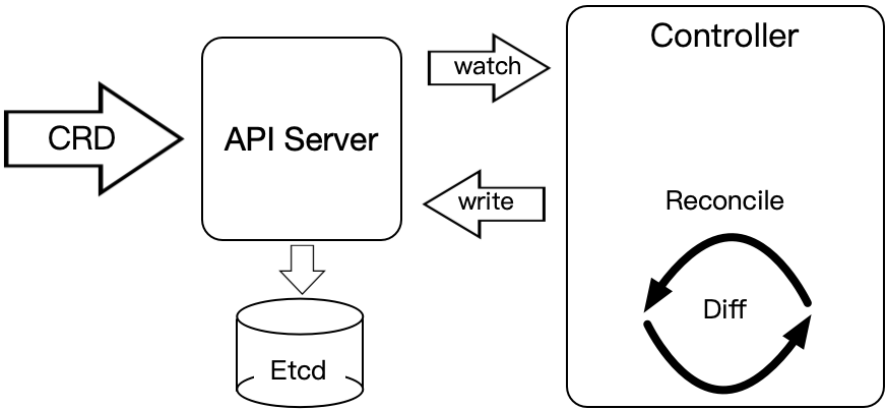

图 4-3: 控制器模式

[1]Kubernetes home page: https://kubernetes.io/

通过对系统中服务的概念进行抽象，定义出与服务生命周期相匹配的自定义资源，称之为托管服务 MS(Managed Service)。托管服务是一个抽象资源，每个托管服务包含了服务名称、服务的路由版本和当前状态等信息，负责对系统中的某一服务进行管理，对应服务在系统中可能存在多个版本的运行实例。当用户主动发起对托管服务的状态变更请求，如切换新版本、撤销旧版本等请求时，托管服务的控制器将监听到相应的变化，然后通过更新系统中的相关资源配置，驱使服务向目标状态进行跳转；若出现异常，调谐逻辑将重新执行，直至服务的状态跳转成功。

托管服务资源的定义很好地与服务生命周期的概念完成了匹配，对外屏蔽了内部服务的细节信息，抽象出对应的服务运行状态，对内又可以很好地对各个服务实例的生命周期进行管理。

由前一章节对服务生命周期模型的讨论可知，用户对目标更新服务的动态更新操作与服务当前所处的状态存在着密不可分的关系。因此，我们将动态更新的相关算法整合到托管服务的调谐逻辑之中，在完成动态更新中关键的版本更新和撤销操作的同时，驱使服务状态进行相应的跳转。下面将具体对实现的版本更新和版本撤销算法进行介绍，为方便说明，这里仍以2.1.1节中引入的系统为例。

4.2.2.1 版本更新

为了保证动态更新过程中的安全性，即满足定义2-3中的要求，更新算法首先需要对运行中的事务所涉及的服务版本进行标识。因此，我们在头信息中引入自定义字段 x-version 来进行具体的标识。当调用方向服务发起请求并返回时，利用流量代理模块的拦截功能，在其头信息中添加上带有 x-version 的服务版本信息，作为事务上下文信息进行返回。随后调用方在接收处理结果时，从头信息提取出事务上下文信息并进行更新本地信息，后续发起远程调用时将利用更新后的信息进行传递，从而完成了对运行时事务的版本标识工作。

如图4-4所示，表示 Convertor 服务在初始版本为 v1 时，系统的正常运行状态。对于所有向 Convertor 服务发起调用的请求，在其返回头信息中添加 x-version: convertor-v1 字段，表征该事务针对于目标更新服务 Convertor 被标识为 v1 版本。这里将 Convertor 服务作为目标更新服务，因此仅关注与 Convertor 服务相关的字段信息，其它非目标更新服务可类推。

基于运行时事务版本的正确标识，我们可进一步实现服务版本的正确更

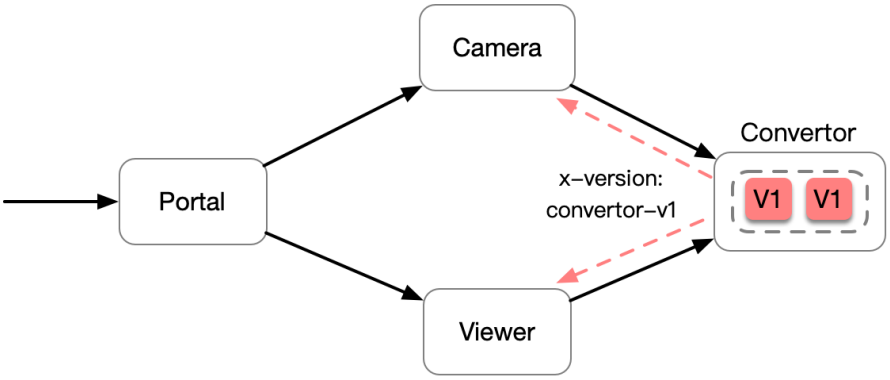

图 4-4: 初始状态下的系统

新。服务被选定为目标更新服务之后，用户需要对新版本进行上线部署工作，服务进入 Deployed 状态，假定对应的新旧版本标识分别为 v2、v1。此时，对流量代理模块中的路由规则进行修改，添加头信息匹配规则：若本次请求头信息中的 x-version 字段包含特定的版本信息，则将该请求转发往特定版本；否则，默认转发往旧版本，同时在返回头信息中添加对应字段。对应系统的运行状态如图 4-5所示，虽然系统中的 v2 版本服务实例已经就绪，但请求全部由 v1 版本的服务实例进行处理。

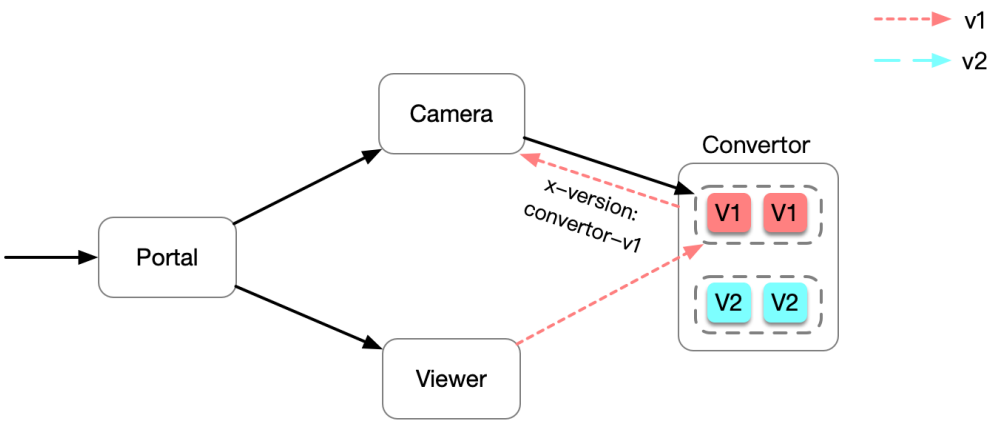

图 4-5: 更新准备前的系统

前述的更新准备工作完成后，将原先路由规则中默认转发往旧版本修改为默认转发往新版本。对应系统的运行状态如图 4-6所示，考虑对于任意一个分布式事务 T，在规则生效之后，向 Convertor 服务发起请求，将包括如下几种情况：

①若分布式事务 T 在运行的过程中还未在 Convertor 服务上发起过子事务，则其请求头信息中将不包含相关的 x-version 字段。依据生效后的路由规则，此

分布式事务 T 对于 Convertor 服务的请求，将被转发往新版本；

②若分布式事务 T 在运行的过程中曾经在 Convertor 服务上发起过子事务，则由前述规则可知，其请求头信息中必然包含了相关的 x-version 字段。依据生效后的路由规则，此分布式事务 T 对于 Convertor 服务的请求，将被转发往与 x-version 字段具有相同标识的版本实例中；

综上，保证了任一分布式事务 T 在运行过程中，不会在目标更新服务 Convertor 的不同版本实例上发起子事务，即满足了动态更新算法对于安全性的要求，同时也使得新生成的分布式事务可以较快地使用到新版本的服务。

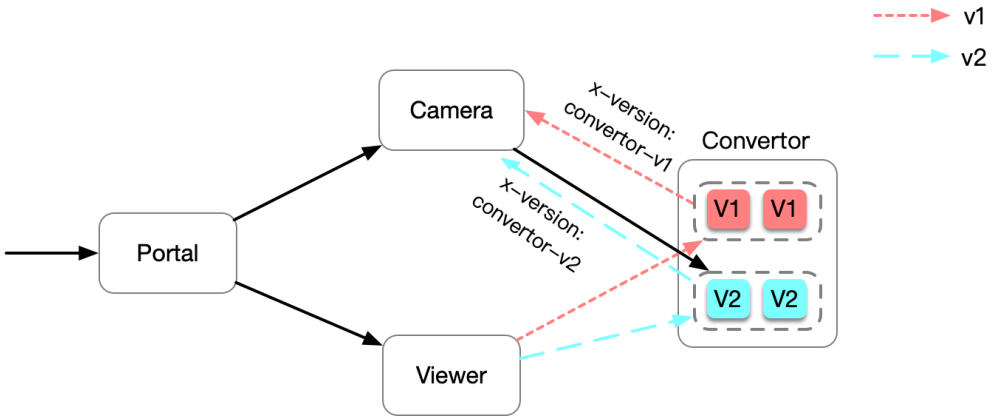

图 4-6: 版本更新中的系统

4.2.2.2 版本撤销

前述的版本更新算法在保证动态更新安全性的基础上，很好地完成了目标更新服务新旧版本请求的正确导向。在此基础上，当收到用户的旧版本撤销请求时，要求版本撤销算法在保证用户请求不会出错的情况下，尽快完成旧版本实例的撤销，及时释放其占用的系统资源。

在正常状态下，当服务收到请求生成事务时，事务依赖管理器将负责对 PAST_SET 集合进行更新，保存所有曾经在此服务上发起过子事务的分布式事务的标识，如图 4-7所示。

版本撤销算法的关键在于判断：

- 旧版本服务实例不会再为未结束的事务提供服务
- 旧版本服务实例当前未执行任何事务

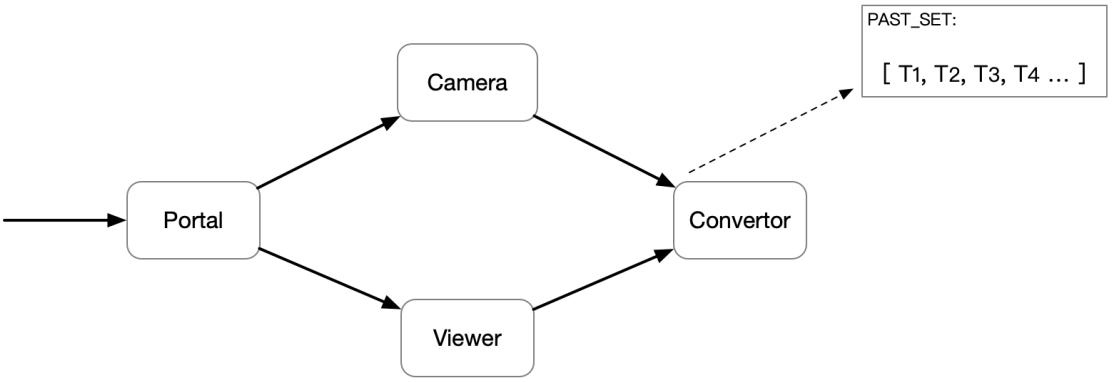

图 4-7: 维护 PAST_SET 集合

应用文献[23] 中所提出的 Freeness 更新判断条件 2-7，利用前述的事务依赖管理器维护的 PAST_SET 和 FUTURE_SET 集合，实现了对应的版本撤销算法，具体包括如下步骤：

1 **阻塞访问**：当收到旧版本撤销请求时，系统首先阻塞所有向旧版本服务实例发起调用的新生成的分布式事务，即其头信息中不存在标识版本信息的 x-version 字段。而对于曾经在旧版本服务实例上发起过的子事务的分布式事务，则允许通过。被阻塞的请求将在调用方执行重试，依据前述版本更新算法中路由规则的定义，本次请求将会被转发往新版本的服务实例中，保证了用户请求依然能够被正常地处理，而且不会再有新的分布式事务标识添加到相关的事务依赖管理器所维护的 PAST_SET 集合中，对应的系统状态如图 4-8所示；

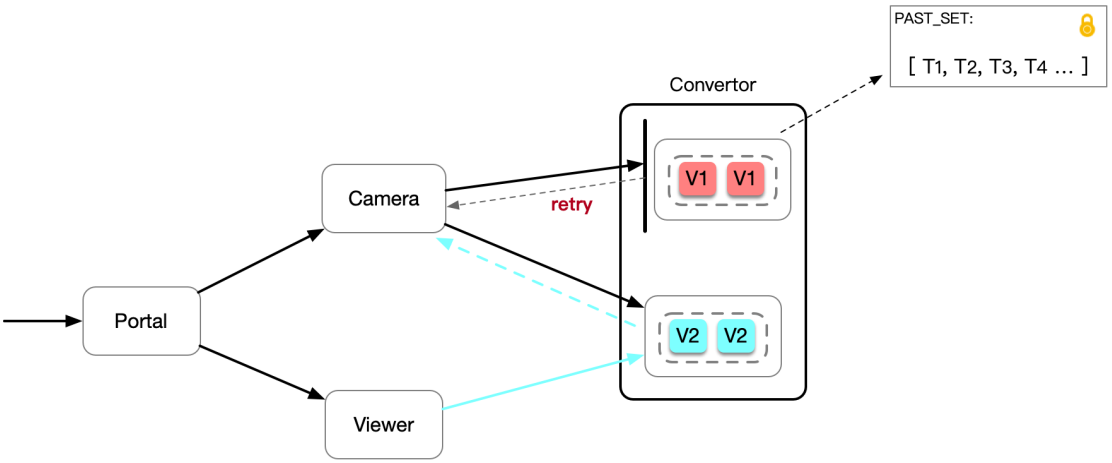

图 4-8: 阻塞访问

2 **同步状态**：由前述可知，分布式事务将来是否会调用目标更新服务的状态信息，保存于系统内的各个服务实例中。目标更新服务可利用本地维护的 PAST_SET 集合，向依赖于自身的相关服务发起请求，同步相关服务的事务状态信息，便可得到相应的 FUTURE_SET 集合。此集合代表所有将来会向目标更新服务发起请求的分布式事务标识集合，同步过程如图 4-9所示，其中的红色虚线段表示具体请求过程：

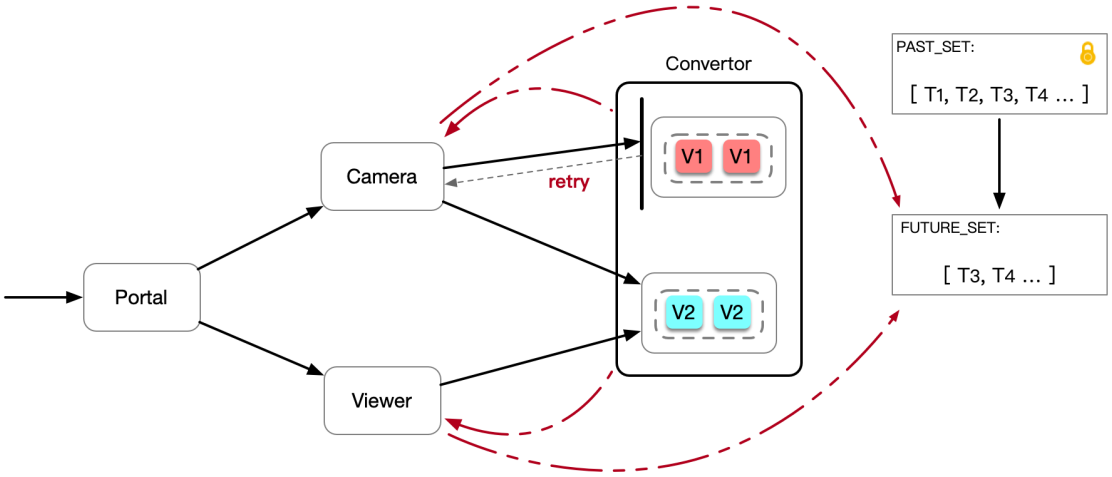

图 4-9: 同步状态

3 **等待撤销**：基于前一步骤所得到的 FUTURE_SET 集合，并结合实例本地事务依赖管理器所维护的相关信息，撤销算法如图 4-10所示，可执行进一步的判断：若接收到的 FUTURE_SET 集合为空，表示针对当前所有调用过旧版本服务实例的分布式事务，均不会再向目标更新服务发起请求，因此只需等待当前旧版本实例上的事务执行结束，即可执行撤销操作，算法执行结束；否则，即表示当前系统中的分布式事务在将来仍会使用到旧版本的服务实例，此时撤销算法流程将进入等待状态，待所有 FUTURE_SET 集合中的分布式事务均完成调用后，对 PAST_SET 集合进行更新，并重新执行前一步骤，以重新同步相关的分布式事务状态。

版本撤销算法利用本地维护的相关事务依赖信息，通过相关的同步请求来收集到具体的分布式事务状态信息，正确判断服务能否被撤销，保证撤销操作的及时性，而且不会阻塞正常的业务请求，对系统的干扰性较小。

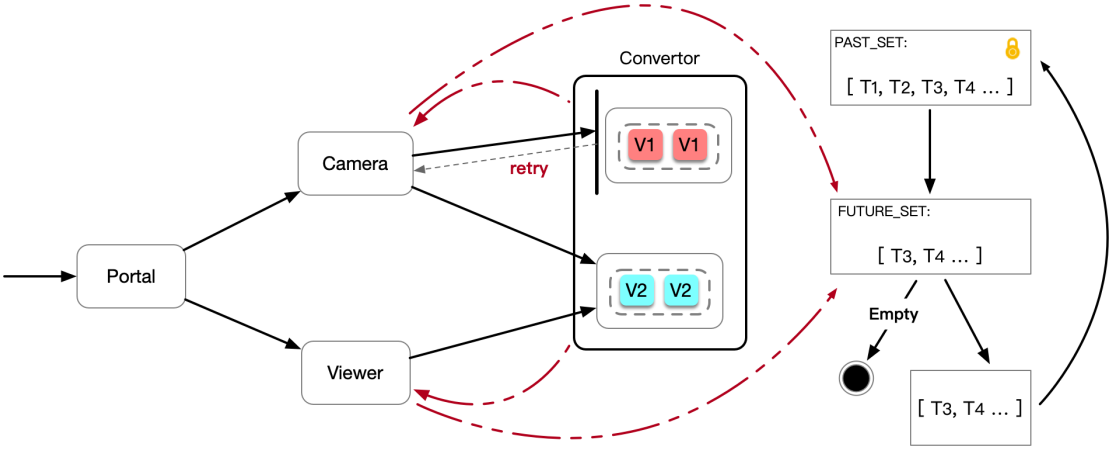

图 4-10: 等待撤销

4.2.3 更新控制模块

前述的相关模块为系统提供了支持动态更新的核心能力，为使系统具有更好的操作性，定义了更新控制模块。

更新控制模块作为系统与用户的交互工具，屏蔽系统内部服务的详细运行信息的同时，对外暴露相关的管理接口。用户可通过接口来进行服务状态的查询，以及发起相关的动态更新请求，系统可自动完成动态更新过程，无需额外的人工操作。

4.3 系统模块交互

上一节具体对扩展模块的内部结构及每个模块的功能进行了介绍，本节将重点描述系统运行时各个模块之间的交互关系。

图4-11展示了版本更新过程的时序图。用户在完成更新前的相关准备工作之后，向更新控制模块发起更新请求。更新控制模块向运行平台的 *API Server* 发送相关资源的修改请求，而服务生命周期管理器将监听到相应的资源变化，进入相应的调谐逻辑。调谐逻辑中通过对路由规则进行更新，利用 Istio 框架前述的配置分发功能，对所有相关服务实例的流量代理模块进行行为配置，以驱使服务的状态完成跳转。

图4-12则展示了版本撤销过程的时序图。同样由用户向更新控制模块发起撤销请求，再由服务生命周期管理器进行相应的处理。此时服务生命周期管理器将进一步与目标更新服务的事务依赖管理模块进行交互，执行前述的版本撤

销算法，使得服务达到目标 Freeness 条件。

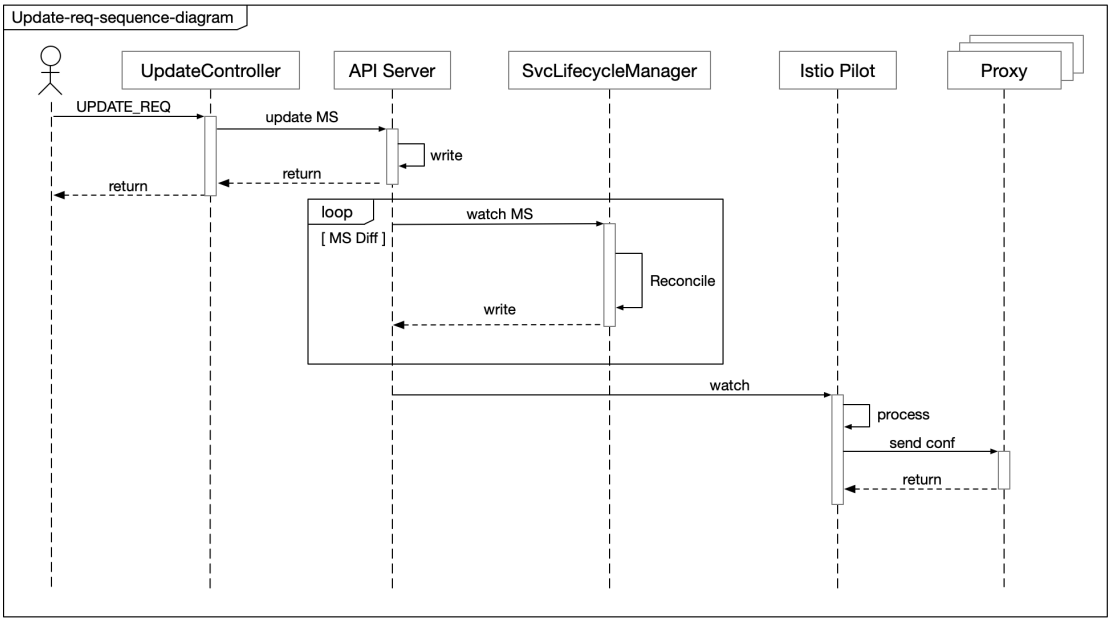

图 4-11: 版本更新时序图

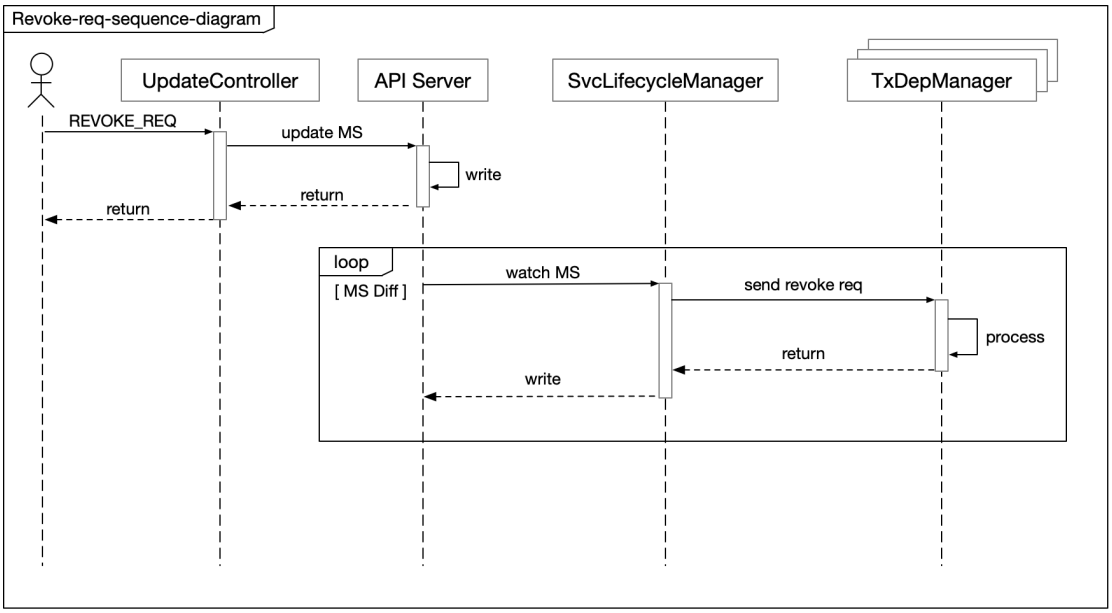

图 4-12: 版本撤销时序图

4.4 本章小结

本章基于前一章节所提出的 VCIM 方法，实现支持动态更新的微服务系统。首先，我们对系统的设计实现进行了概述，对系统的设计目标和准则进行了说明。然后本章对三个扩展模块的实现及其模块内部结构进行了详细的介绍，其中包括事务依赖管理器如何对事务进行拦截和管理，为系统支持动态更新提供基础信息；服务生命周期管理器如何对服务的状态进行扩展，版本更新和版本撤销算法如何完成整合，保证动态更新过程的一致性；更新控制模块如何完成用户与系统间的解耦，为用户提供相关的更新管理功能。最后，重点介绍了系统运行时各个模块间的交互关系。

第五章　案例研究与实验评估

为了验证前述的微服务动态更新解决方案是否满足动态更新过程的相关准则，本章将利用一个实际场景下的分布式微服务系统进行实验。实验将对不同的动态更新算法，在安全性、及时性和干扰性三个方面的性能进行评估和比较。

5.1　案例介绍

这里使用的分布式微服务系统，是一个面向开发日的台风演示系统。该系统模拟 2018 年的第 18 号台风 Rumbia，将台风从生成发展到最后结束的运行轨迹进行展示，同时系统通过对雨情和风情的相关服务进行查询，整合了台风及其途径城市的详细信息，为用户提供展示。具体来说，用户可通过该演示系统，直观地看到当前台风的运行轨迹。当台风的影响范围波及某个具体的城市时，系统将对所在城市的相关雨情和风情服务进行查询，并进行展示预警。该系统的静态依赖关系如图 5-1所示：

台风演示系统中，部分主要服务提供的功能如下：

- **Backend** 服务负责整合包括台风强度、位置等详细信息，以及途径影响城市的具体信息。
- **RainController** 服务负责整合所有台风波及城市具体的雨情信息。
- **WindController** 服务负责整合所有台风波及城市具体的风情信息。
- **Typhoon** 服务负责提供台风当前的详细信息，经纬度位置、强度和速度等。
- **xx-Rain(Wind)** 服务负责对应 xx 城市的雨情 (风情) 信息。

服务间的交互时序图如图 5-2所示 (已省略去部分非关键服务)：

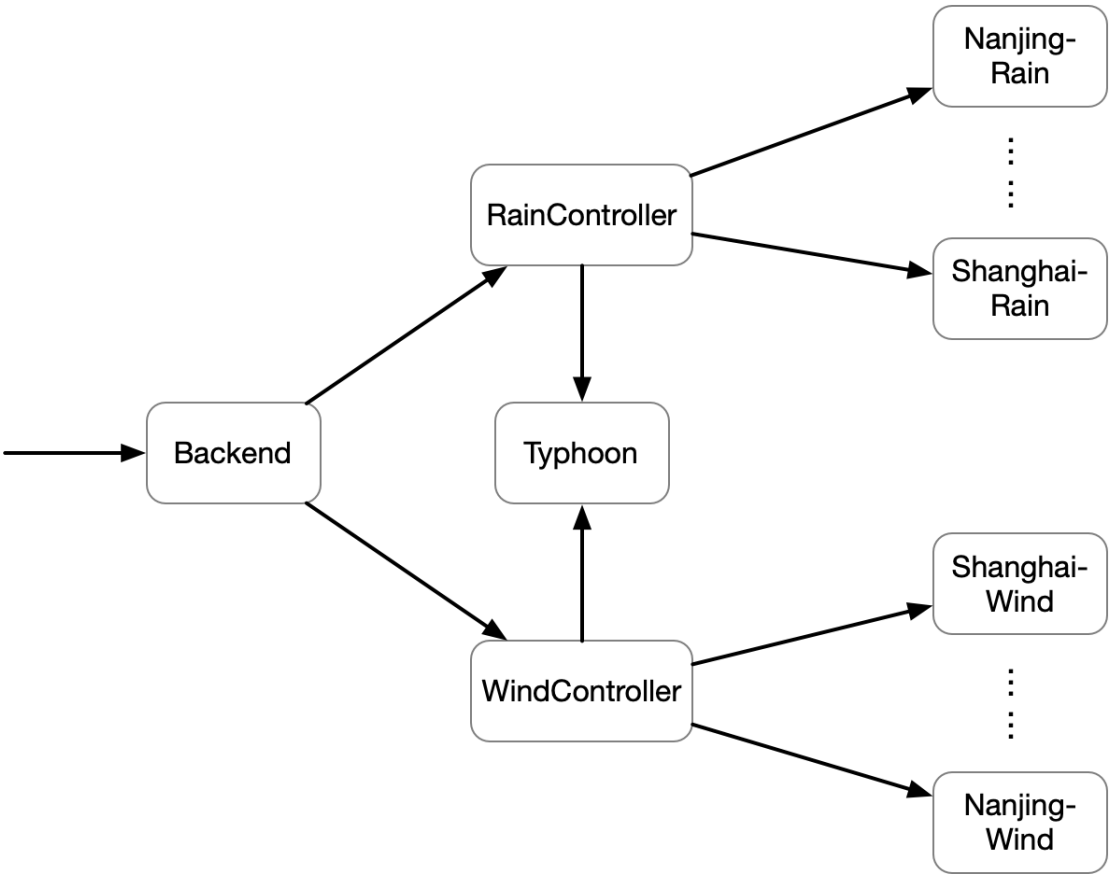

图 5-1: 台风系统静态依赖关系图

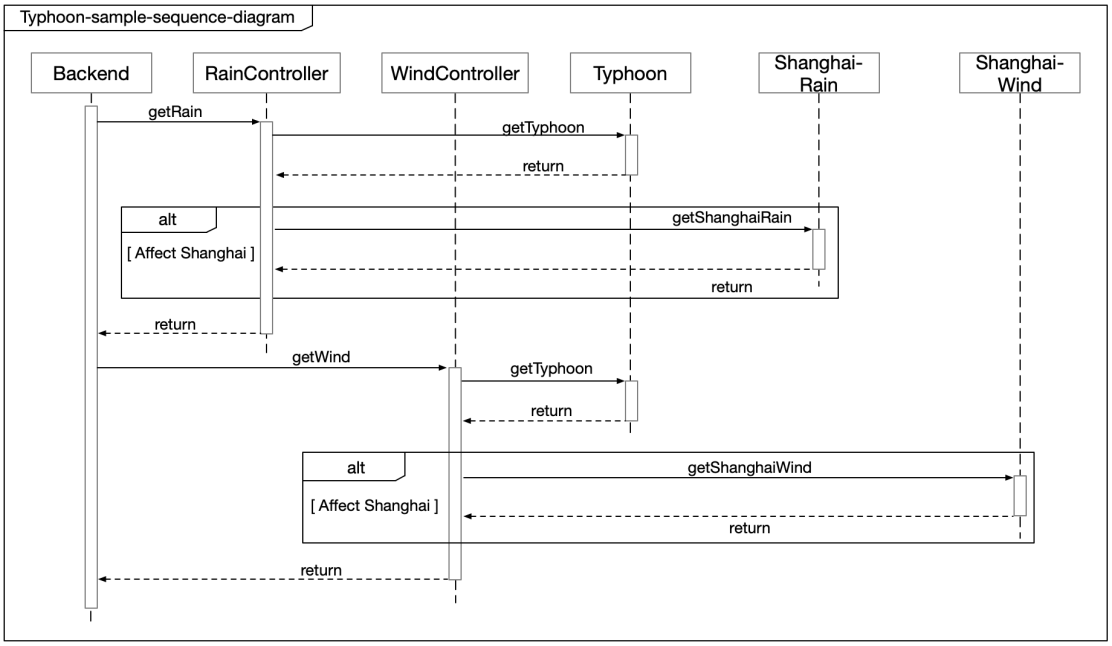

图 5-2: 台风系统时序图

5.2 实验环境与设计

5.2.1 实验环境

我们将使用上述的台风演示系统来对本文实现的动态更新技术进行评估，为了更好地模拟真实场景下的分布式微服务系统，本次实验将台风演示系统中的 16 个服务部署于 Kubernetes 系统之上，由其进行容器运行时的编排，具体的环境配置为：

物理节点机器硬件配置：24 核 2.5GHz CPU，64GB 内存，1T 硬盘，千兆以太网卡

软件环境配置：Ubuntu 20.04.1 LTS，Kubernetes v1.17.14

5.2.2 实验设计

在台风演示系统的实际运行中，系统开发人员希望能够在完成代码的改动之后，不停止系统运行的前提下，对对应的服务进行动态更新，如 Typhoon 服务可能需要通过更新来更新数据或修复错误。本章所进行的实验主要是针对 Typhoon 服务进行动态更新。

服务间的静态依赖关系如5-1所示，Typhoon 服务在实验案例中处于整个调用层次的第三层，可在其之上发起子事务的服务包括 RainController 和 WindController，整个分布式事务的根事务在 Backend 服务上发起，因此四个服务的调用关系近似形成一个“菱形”的关系。对于 Quiescence 算法来说，当 Typhoon 服务需要进行动态更新时，需要保证 EPS(Typhoon) 集合中的所有服务均属于 Passive 状态，即 Backend、RainController、WindController 都必须达到 Passive 状态，显然这将对系统造成很大的干扰，并且更新所耗费的时间也会很长，对应于动态更新的干扰性和及时性。而对于 Version Consistency 算法来说，只需依据 Typhoon 服务上事务的相关信息以及同步相关服务上事务的运行状态，即可完成动态更新。因此选择 Typhoon 服务进行动态更新可以有效地评估不同更新算法的及时性和干扰性。另一方面，对于常见的微服务部署方案来说，其考虑新版本服务的快速部署更新，对应的算法为直接切换版本算法 (Direct)。若 Typhoon 服务需要进行动态更新，系统会在新版本服务就绪后，直接进行版本切换更新，显然，若更新的时机发生在 RainController 发生之后，且在开始调用 WindController 之前，那么本次更新将导致系统出现不一致的情

况，而 Quiescence 算法和 Version Consistency 则不会出现不一致的情况。因此对 Typhoon 服务进行更新对验证各个算法的安全性具有实际意义。

在具体的实验设置方面，将 < 算法，请求间隔，实验目标 > 称为一组实验，其中算法包括 Quiescence、Version Consistency 和相应的基准算法；请求间隔包括 1500ms、750ms、500ms、300ms、150ms，请求间隔表示向系统发起请求的时间间隔，用来表征系统的负载状况，请求间隔越小说明系统负载越大；实验目标则包括了安全性、及时性和干扰性三个方面。

针对每一组实验，将分别独立地进行 5 次。每次实验将依据对应的请求间隔，不断地向系统发起正常的用户请求，一共发送 60 次，同时在第 20 个请求处，向系统中的目标更新服务发起动态更新请求。每次实验仅在第 20 个请求处，额外向系统发起一次动态更新请求，此时可以保证系统中已经存在多个运行时分布式事务，可以更好地模拟真实状态下的分布式微服务系统，确保实验的合理性和公平性。

5.3 实验分析

在实际的应用系统中，当运维人员尝试对其中的服务进行动态更新时，动态更新的前后系统的运行时事务是否安全、更新所需要耗费的时间以及整个更新过程对系统会造成多大的干扰，这些都是其评价动态更新算法的重要标准。因此本章的实验将从安全性 (Safety)、及时性 (Timeliness) 和干扰性 (Disruption) 三个方面，进行具体的评测。

5.3.1 安全性 (Safety)

安全性表示在服务进行动态更新的过程前后，系统处理正在运行时的事务和后续将要发生的事务都应满足一致性，即逻辑正确性。不同的动态更新算法对于安全更新点的定义各不相同，因此对于安全性的满足也并不一致。

安全性实验针对前述的三种动态更新算法，分别按照对应的实验设置来进行。例如实验 <Direct，1500ms，安全性 > 表示使用直接切换版本算法进行安全性的实验，模拟的用户请求间隔为 1500ms。实验中记录的结果为动态更新过程中出现不一致的用户请求数量，用来表示该算法是否满足安全性。实验结果如表 5-1所示。

从表中数据可以看出，直接切换版本算法并未对更新的安全点进行考虑，

表 5-1: 安全性实验结果

请求间隔 (ms)	1500	750	500	300	150	100
Direct	5	8	12	16	33	50
Quiescence	0	0	0	0	0	0
VCIM	0	0	0	0	0	0

不能保证动态更新的安全性，且不一致的数量随系统负载的增加而增加。如在请求间隔为 500ms 时，系统中共出现 12 次不一致的用户请求数量。而 Quiescence 算法和本文中实现的 VCIM 算法在实验中并没有出现不一致的请求，可以保证动态更新的安全性，实验结果和算法的相关理论分析相吻合。

由于直接切换版本算法无法保证动态更新的安全性，因此在后续及时性和干扰性的实验中，我们将仅对 Quiescence 算法和实现的 VCIM 算法进行实验。

5.3.2 及时性 (Timeliness)

及时性表示服务从收到动态更新请求到真正完成动态更新所需要的时间。及时性用于反映动态更新算法的效率，及时性越高对应的服务到达安全更新点所耗费的时间越短，即可以更快地完成动态更新。

在及时性实验中，对于每一组实验都分别记录下服务收到动态更新请求的时间，记为开始时间 $time_0$，以及对应完成动态更新的时间，记为结束时间 $time_1$，则该组及时性实验的计算结果为 $time_1 - time_0$。每组实验将独立重复多次进行，通过对多次实验结果进行求和平均，便可得到该组实验的及时性实验结果。

实验结果如图5-3所示。从图中可以看出，两种算法更新目标服务所需要耗费的时间都随系统负载的增加而增加。对于同一系统负载的情况下，Quiescence 算法要求的安全点较为严格，更新所耗费的时间较长。而本文中实现的 VCIM 算法所耗费的更新时间则要少一些，优于 Quiescence 算法。

5.3.3 干扰性 (Disruption)

干扰性表示应用动态更新算法给系统正常处理用户请求所导致的中断程度，即相较于没有动态更新请求下，动态更新过程对系统处理用户请求额外增加的时间。某种动态更新算法的干扰性越小，表明该算法更为透明高效。

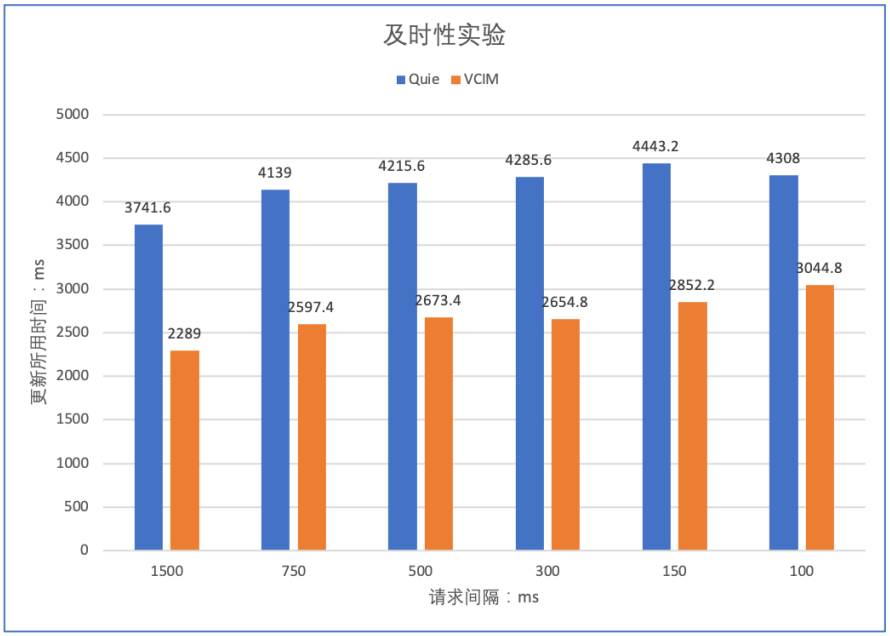

图 5-3: 及时性实验结果

干扰性实验的结果可包括：总干扰性 (Total Disruption)，平均干扰性 (Average Disruption)。

由于需要计算动态更新算法导致额外增加的时间，因此干扰性实验只对受更新影响的分布式事务进行计算。对于某一个分布式事务，其受更新影响当前仅当对应的事务生命周期与服务进行更新的时间段相互重叠。在进行干扰性实验时，首先进行基准组实验，即在没有更新请求的状态下，系统正常处理用户请求的响应时间，重复多次后计算得到平均的响应时间，记为 $time_0$。然后对于使用特定更新算法的第 j 次实验中，统计所有受更新影响的用户请求 (分布式事务) 的数量，记为 $count_j$，以及这些请求对应的响应时间，记为 $time_i$。该组实验共进行多次，对应的总干扰性计算结果为：$TotalDisruption = \sum_{j=1}^{5} \sum_{i=1}^{count_j} (time_i - count_j \cdot time_0)$，对应的平均干扰性计算结果为：$AverageDisruption = TotalDisruption / (\sum_{j=1}^{5} count_j)$，其中 $count_j$ 表示在第 j 次实验中，受更新影响的事务的总数量。

总干扰性和平均干扰性的实验结果分别如图 5-4、图 5-5所示。从干扰性实验结果可以明显看出，VCIM 算法可以显著降低更新过程对系统的影响，明显优于 Quiescence 算法。

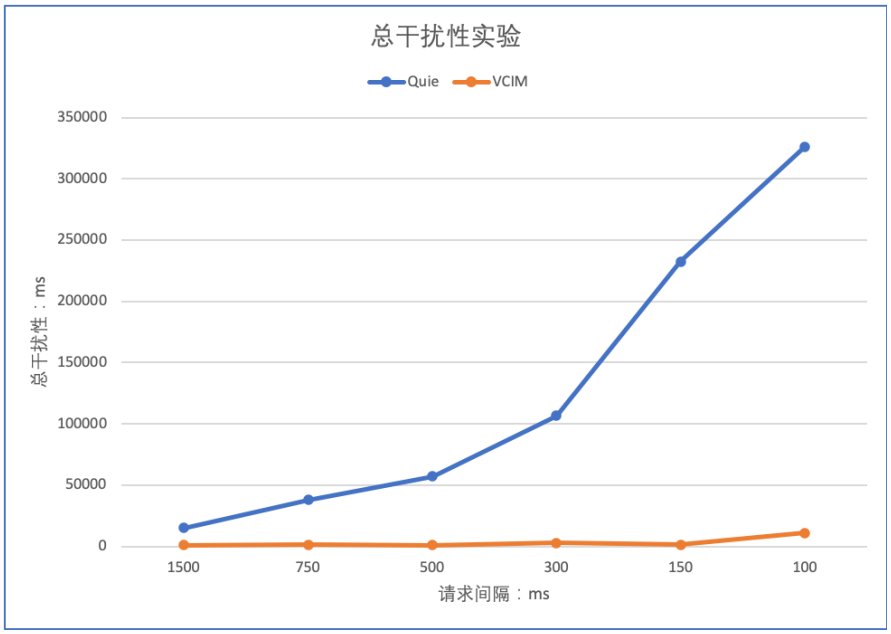

图 5-4: 总干扰性实验结果

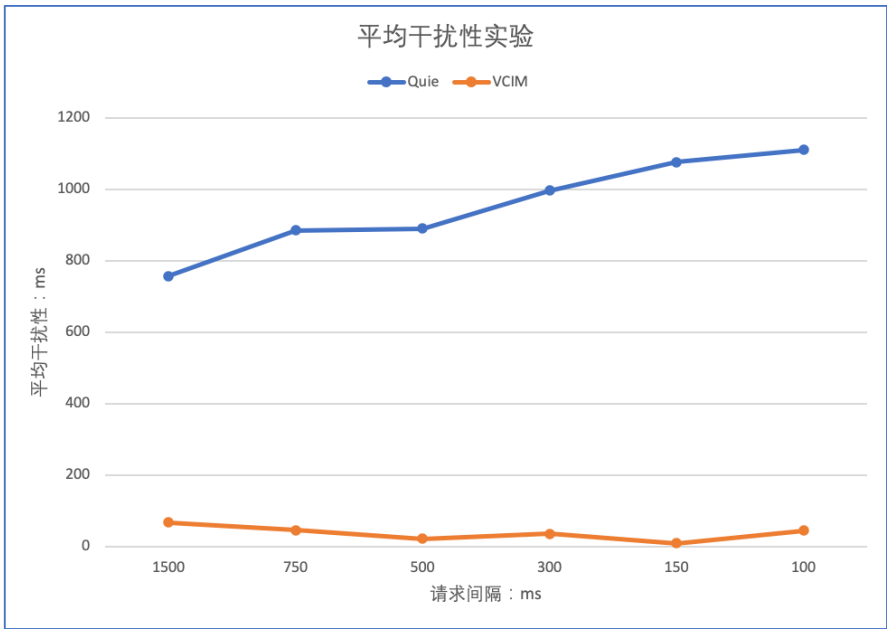

图 5-5: 平均干扰性实验结果

5.4 实验结论

通过上述的实验，我们对各个动态更新算法在安全性、及时性、干扰性三方面进行了性能对比：

安全性： Quiescence 算法和 VCIM 算法都保证了动态更新过程中的安全

性，不会导致系统出现不一致的情况。而直接切换版本算法对应于常见的蓝绿部署和金丝雀部署方案，主要目标是版本的快速更新迭代，并未对安全更新点进行考虑，显然将导致更新过程中系统有可能会出现不一致的情况。因此，在安全性方面，Quiescence 算法和 VCIM 算法要优于直接切换版本算法。

及时性：Quiescence 算法利用静态依赖信息，更新需要等待相关事务全部结束，而 VCIM 算法则利用动态依赖信息，只需保证相关事务不会再次调用到目标更新服务即可。因此在及时性方面，VCIM 算法要优于 Quiescence 算法。

干扰性：由于 Quiescence 算法更新时对相关服务的状态具有较强的约束性，导致对系统造成较大的干扰性。而 VCIM 算法使用流量代理接管用户请求的方式，不需要阻塞事务的运行，可以降低对系统造成的干扰。因此在干扰性方面，VCIM 算法要优于 Quiescence 算法。

5.5 本章小结

本章首先对一个实际场景下的分布式微服务系统进行介绍和分析，然后基于此案例进行了相应的动态更新实验。实验通过分析并选取了系统中合适的目标更新服务，重点对各个动态更新算法在安全性、及时性和干扰性三个方面进行了比较，并对实验结果进行了解释说明。

第六章　总结与展望

6.1　工作总结

本文首先探讨了基于微服务架构的软件开发技术在企业级应用中的重要性。但随着互联网和软件技术的发展，软件运行环境和用户的需求也在不断地变化，微服务系统变得多元复杂且难以维护，因此微服务系统对于动态更新的支持显得尤为重要。本文首先对现有的软件动态更新算法以及它们在安全性、及时性和干扰性方面的表现进行了介绍，并讨论了当前微服务框架和架构在动态更新方面的支持。接着，针对当前微服务架构在动态更新支撑方面的不足，本文对支持动态更新的分布式微服务系统所面临的需求与挑战进行了分析，对应在事务模型、动态依赖管理以及服务生命周期三个方面提出了具体的扩展模型，并选择了 Service Mesh 作为底层实现架构，给出了一种支持事务一致性的微服务动态更新方法。然后依据此方法，在具体的 Istio 框架上完成了系统的设计与实现。最终，通过实际的应用案例对此动态更新方法和相关更新算法进行了实验对比，验证了此方法在更新过程中的安全性、及时性和低干扰性。

本文的主要贡献总结如下：

- 分析当前主流微服务系统框架在动态更新支撑方面的不足，针对性地提出了一种支持事务一致性的微服务动态更新的实现方法。该方法选择在 Service Mesh 架构上，对事务模型、动态依赖管理和服务生命周期三个方面进行了具体的扩展，为具体的动态更新过程提供运行时支持。

- 基于上述的实现方法，在 Service Mesh 的开源实现框架 Istio[1]之上完成了模型的扩展，并且提供了相关的动态更新管理接口，支持服务运行时的动态更新。同时系统保证动态更新过程的安全性、及时性和低干扰性，模块之间具有较低的耦合度。

- 利用实际的应用案例，应用上述实现方法并进行实验评估，使用不同的

[1]Istio home page: https://istio.io/

更新算法，主要对动态更新过程中各个算法的安全性 (Safety)、及时性 (Timeliness)、干扰性 (Disruption) 进行了性能比较。

6.2 研究展望

本文探讨了微服务架构在动态更新支持方面的不足，给出了一种支持事务一致性的微服务动态更新实现方法，并且在具体的开源框架 Istio 上进行了模型的扩展和实现，并通过具体的实验验证了此方法的可行性和有效性。然而，还有很多具体的工作需要进一步的研究：

1. 对于真实网络环境下的分布式微服务系统，消息的发送与处理都有可能出现异常，但本文所定义的相关模型均基于如下假设：消息能够正确无异常地完成传输。在未来的工作中，我们将对相关的异常情况进行考虑，对模型进行完善。

2. 在真实运行环境中，节点可能在任意时刻出现崩溃，若目标更新服务运行于之上则更新过程将出现错误。在后续工作中，我们将考虑更新过程发生错误的回滚机制，使得系统在动态更新方面具有更好的健壮性。

致　　谢

时光荏苒，研究生三年的学习时光已经接近尾声，在此我想对我的父母，我的老师和同学们表达由衷的谢意。

感谢我的导师曹春教授。在我读研期间，曹老师一直关心我的科研和生活。本文从选题、成文到修改都离不开曹老师的耐心指导。曹老师对待学术严谨执着，待人和善，平时非常关心学生的生活状况和身体健康。他的谆谆教诲使我受益匪浅。

感谢吕建教授、马晓星教授、陶先平教授、徐峰教授、许畅教授、黄宇教授、胡昊副教授、余萍副教授、徐经纬副教授、姚远副教授、张建莹老师等所有关心和帮助过我的老师。

感谢实验室的所有同学，良好的学习氛围是我顺利完成科研项目的保障，能够和你们一起学习和工作让我感到非常荣幸。

感谢同寝的邱圣广和张文明同学，你们的陪伴是我宝贵的精神财富。

最后向我的家人和女朋友致以最诚挚的感谢。感谢父母对我的养育之恩和无条件的支持，使我可以全身心的投入到科研中；感谢我的女朋友对我的关心和鼓励，让我一路上都有人同行，充满喜悦。

参考文献

[1] BUCCHIARONE A, DRAGONI N, DUSTDAR S, et al. From monolithic to microservices: An experience report from the banking domain[J]. Ieee Software, 2018, 35(3): 50–55.

[2] TAIBI D, LENARDUZZI V, PAHL C. Processes, motivations, and issues for migrating to microservices architectures: An empirical investigation[J]. IEEE Cloud Computing, 2017, 4(5): 22–32.

[3] AL-DEBAGY O, MARTINEK P. A comparative review of microservices and monolithic architectures[C] // 2018 IEEE 18th International Symposium on Computational Intelligence and Informatics (CINTI). 2018: 000149–000154.

[4] 马晓星, 刘譞哲, 谢冰, et al. 软件开发方法发展回顾与展望 [J/OL]. 软件学报, 2019, 1: 3–21.
http://www.jos.org.cn/1000-9825/5650.htm.

[5] LEWI J, FOWLER M. Microservices[EB/OL]. 2014 [online].
https://www.martinfowler.com/articles/microservices.html.

[6] DRAGONI N, GIALLORENZO S, LAFUENTE A L, et al. Microservices: yesterday, today, and tomorrow[J]. Present and ulterior software engineering, 2017: 195–216.

[7] BALALAIE A, HEYDARNOORI A, JAMSHIDI P. Microservices architecture enables devops: Migration to a cloud-native architecture[J]. Ieee Software, 2016, 33(3): 42–52.

[8] TAIBI D, LENARDUZZI V, PAHL C. Architectural Patterns for Microservices: A Systematic Mapping Study.[C] // CLOSER. 2018: 221–232.

[9] NEAMTIU I, DUMITRAŞ T. Cloud software upgrades: Challenges and opportunities[C] // 2011 International Workshop on the Maintenance and Evolution of Service-Oriented and Cloud-Based Systems. 2011 : 1 – 10.

[10] BARESI L, GUINEA S, LA MANNA V P. Consistent runtime evolution of service-based business processes[C] // 2014 IEEE/IFIP Conference on Software Architecture. 2014 : 77 – 86.

[11] ELLIOT S. DevOps and the cost of downtime: Fortune 1000 best practice metrics quantified[J]. International Data Corporation (IDC), 2014.

[12] KRAMER J, MAGEE J. The evolving philosophers problem: Dynamic change management[J]. IEEE Transactions on software engineering, 1990, 16(11) : 1293 – 1306.

[13] FOWLER M. BlueGreenDeployment[J]. WWW, Available: http://martinfowler.com/bliki/BlueGreenDeployment.html, 2010.

[14] SATO D. Canary update strategies[J]. WWW, Available: https://martinfowler.com/bliki/CanaryRelease.html, 2014.

[15] TARVO A, SWEENEY P F, MITCHELL N, et al. CanaryAdvisor: a statistical-based tool for canary testing[C] // Proceedings of the 2015 International Symposium on Software Testing and Analysis. 2015 : 418 – 422.

[16] GABBRIELLI M, GIALLORENZO S, GUIDI C, et al. Self-reconfiguring microservices[G] // Theory and Practice of Formal Methods. [S.l.] : Springer, 2016 : 194 – 210.

[17] Jolie. Official Web Site.[EB/OL]. [online]. https://www.jolie-lang.org/.

[18] SAMPAIO A R, KADIYALA H, HU B, et al. Supporting microservice evolution[C] // 2017 IEEE International Conference on Software Maintenance and Evolution (ICSME). 2017 : 539 – 543.

[19] BOYER F, ETCHEVERS X, DE PALMA N, et al. Architecture-based automated updates of distributed microservices[C] // International Conference on Service-Oriented Computing. 2018 : 21 – 36.

[20] SEGAL M E, FRIEDER O. On-the-fly program modification: Systems for dynamic updating[J]. IEEE software, 1993, 10(2) : 53 – 65.

[21] MAGEE J, KRAMER J. Dynamic structure in software architectures[J]. ACM SIGSOFT Software Engineering Notes, 1996, 21(6) : 3 – 14.

[22] VANDEWOUDE Y, EBRAERT P, BERBERS Y, et al. Tranquility: A low disruptive alternative to quiescence for ensuring safe dynamic updates[J]. IEEE Transactions on Software Engineering, 2007, 33(12) : 856 – 868.

[23] MA X, BARESI L, GHEZZI C, et al. Version-consistent dynamic reconfiguration of component-based distributed systems[C] // Proceedings of the 19th ACM SIGSOFT symposium and the 13th European conference on Foundations of software engineering. 2011 : 245 – 255.

[24] BARESI L, GHEZZI C, MA X, et al. Efficient dynamic updates of distributed components through version consistency[J]. IEEE Transactions on Software Engineering, 2016, 43(4) : 340 – 358.

[25] LI W, LEMIEUX Y, GAO J, et al. Service mesh: Challenges, state of the art, and future research opportunities[C] // 2019 IEEE International Conference on Service-Oriented System Engineering (SOSE). 2019 : 122 – 1225.

[26] 王怀民, 史佩昌, 丁博, et al. 软件服务的在线演化 [D]. [S.l.] : [s.n.], 2011.

[27] HICKS M, NETTLES S. Dynamic software updating[J]. ACM Transactions on Programming Languages and Systems (TOPLAS), 2005, 27(6) : 1049 – 1096.

[28] MENS T, WERMELINGER M, DUCASSE S, et al. Challenges in software evolution[C] // Eighth International Workshop on Principles of Software Evolution (IWPSE'05). 2005 : 13 – 22.

[29] ARNOLD R S. Software change impact analysis[M]. [S.l.] : IEEE Computer Society Press, 1996.

[30] VANDEWOUDE Y, RIGOLE P, URTING D, et al. Draco: An adaptive runtime environment for components[J]. Appendix of the EMPRESS deliverable for Run-time Evolution and Dynamic (Re) configuration of Components, 2003.

[31] BIDAN C, ISSARNY V, SARIDAKIS T, et al. A dynamic reconfiguration service for CORBA[C] // Proceedings. Fourth International Conference on Configurable Distributed Systems (Cat. No. 98EX159). 1998 : 35 – 42.

[32] CHEN X, SIMONS M. A component framework for dynamic reconfiguration of distributed systems[C] // International Working Conference on Component Deployment. 2002 : 82 – 96.

[33] NEWMAN S. Building microservices: designing fine-grained systems[M]. [S.l.] : " O'Reilly Media, Inc.", 2015.

[34] NADAREISHVILI I, MITRA R, MCLARTY M, et al. Microservice architecture: aligning principles, practices, and culture[M]. [S.l.] : " O'Reilly Media, Inc.", 2016.

[35] Netflix Open Source Software.[EB/OL]. [online]. http://netflix.github.io/.

[36] JOHNSON R, HOELLER J, DONALD K, et al. The spring framework–reference documentation[J]. interface, 2004, 21 : 27.

[37] REN R, MA J, SUI X, et al. D 2 P: a distributed deadline propagation approach to tolerate long-tail latency in datacenters[C] // Proceedings of 5th Asia-Pacific Workshop on Systems. 2014 : 1 – 6.

[38] MORGAN W. What'sa service mesh? And why do I need one[J]. Tarkistettu, 2017, 27 : 2019.

[39] EL MALKI A, ZDUN U. Guiding architectural decision making on service mesh based microservice architectures[C] // European Conference on Software Architecture. 2019 : 3 – 19.

[40] SU P, CAO C, MA X, et al. Automated management of dynamic component dependency for runtime system reconfiguration[C] // 2013 20th Asia-Pacific Software Engineering Conference (APSEC) : Vol 1. 2013 : 450 – 458.

[41] SHEIKH O, DIKALEH S, MISTRY D, et al. Modernize digital applications with microservices management using the istio service mesh[C] // Proceedings of the 28th Annual International Conference on Computer Science and Software Engineering. 2018 : 359 – 360.

[42] WANG Y, MA D. Developing a process in architecting microservice infrastructure with Docker, Kubernetes, and Istio[J]. arXiv preprint arXiv:1911.02275, 2019.

简历与科研成果

基本信息

王东宇，男，汉族，1996 年 3 月出生，海南省临高人。

教育背景

2018 年 9 月 — 2021 年 6 月 南京大学计算机科学与技术系 硕士
2014 年 9 月 — 2018 年 6 月 南京大学计算机科学与技术系 本科

攻读硕士学位期间完成的学术成果

1. Xiaobao Wei, Jinnan Chen, "Voting-on-Grid Clustering for Secure Localization in Wireless Sensor Networks," in *Proc. IEEE International Conference on Communications (ICC) 2010*, May. 2010.
2. Xiaobao Wei, Shiba Mao, Jinnan Chen, "Protecting Source Location Privacy in Wireless Sensor Networks with Data Aggregation," in *Proc. 6th International Conference on Ubiquitous Intelligence and Computing (UIC) 2009*, Oct. 2009.

攻读硕士学位期间参与的科研课题

1. 国家自然科学基金面上项目 "无线传感器网络在知识获取过程中的若干安全问题研究"（课题年限 2010 年 1 月 — 2012 年 12 月），负责位置相关安全问题的研究。
2. 江苏省知识创新工程重要方向项目下属课题 "下一代移动通信安全机制研究"（课题年限 2010 年 1 月 — 2010 年 12 月），负责 LTE/SAE 认证相关的安全问题研究。

学位论文出版授权书

本人完全同意《中国优秀博硕士学位论文全文数据库出版章程》（以下简称“章程”），愿意将本人的学位论文提交“中国学术期刊（光盘版）电子杂志社”在《中国博士学位论文全文数据库》、《中国优秀硕士学位论文全文数据库》中全文发表。《中国博士学位论文全文数据库》、《中国优秀硕士学位论文全文数据库》可以以电子、网络及其他数字媒体形式公开出版，并同意编入《中国知识资源总库》，在《中国博硕士学位论文评价数据库》中使用和在互联网上传播，同意按“章程”规定享受相关权益。

作者签名：_____

_____年_____月_____日

论文题名	支持事务一致性的微服务动态更新系统实现				
研究生学号	MG1833071	所在院系	计算机科学与技术系	学位年度	2018
论文级别	☑ 硕士　☐ 硕士专业学位 ☐ 博士　☐ 博士专业学位　（请在方框内画勾）				
作者电话	18362926576		作者 Email	wdongyu@outlook.com	
第一导师姓名	曹春 教授		导师电话	18951679203	

论文涉密情况：

☐ 不保密

☑ 保密，保密期：_____年_____月_____日 至 _____年_____月_____日

注：请将该授权书填写后装订在学位论文最后一页（南大封面）。