

文章编号: 1005-8451 (2012) 02-0024-03

铁路客票系统中基于 Paxos 算法的数据一致性模型研究

王红爱, 祝红光, 贾新茹

(中国铁道科学研究院 电子计算技术研究所, 100081)

摘要: 文章介绍分布式系统中解决一致性问题的 Paxos 算法, 分析铁路客票发售和预定系统 (简称客票系统) 的数据架构, 建立了基于 Paxos 算法的数据一致性模型。该模型可以解决铁路客票系统中数据的一致性和完整性问题。

关键词: 数据一致性; Paxos 算法; 客票系统; 云计算

中图分类号: U293.22 : TP39 **文献标识码:** A

Research on Paxos Algorithm-based model of data consistency in Railway Ticketing and Reservation

WANG Hong-ai, ZHU Hong-guang, JIA Xin-ru

(Institute of Computing Technologies, China Academy of Railway Sciences, Beijing 100081, China)

Abstract: This paper introduced the Paxos Algorithm of solving the consistency problems in a distributed system, analyzed the data structure of Railway Ticketing and Reservation (TRS), established the model of data consistency based on Paxos Algorithm. This model could solve the problems of data consistency and integrity in TRS.

Key words: data consistency; Paxos Algorithm; Ticketing and Reservation System (TRS); cloud computing

铁路客票发送和预定系统 (简称客票系统) 自 1996 年应用以来, 已经建成了铁道部、地区中心、车站三级分布式结构。系统拥有以市场需求为导向的客票销售体系, 满足旅客的不同需求。本系统每日与海量数据进行交互, 这些数据是客票系统发展的基础, 从应用的角度来说, 分为基础数据和业务数据。基础数据一般包括站名字典、线路字典、票价基础价率等相对稳定的数据, 由铁道部全表复制到各中心和各车站, 而业务数据一般包括存根、席位、财收等动态数据, 根据统计的要求由车站级或中心级向上级单位逐级传输。基础数据的一致性要求非常强, 需要严格一致, 关键数据的丢失会影响售票业务的正常进行。而业务数据的丢失会造成铁路局、铁道部报表的统计数据不完整, 对分析运营结果和制定运营策略产生不利影响。云计算的 Paxos 算法可以解决分布式系统中的一致性问题的。本文分析了客票系统的数据架构, 介

绍了目前客票系统保证数据一致性的技术, 建立了基于 Paxos 算法的数据一致性模型, 可以有效解决客票系统数据的一致性和完整性问题。

1 Paxos 算法介绍

Paxos 算法^[1-2]是由 Leslie Lamport 最先提出的一种基于消息传递 (Messages Passing) 的一致性算法。用来解决一个分布式系统如何就某个值 (决议) 达成一致的问题。

一个典型的场景是, 在一个分布式数据库系统中, 如果各节点的初始状态一致, 每个节点都执行相同的操作序列, 那么他们最后能得到一个一致的状态。在该算法中规定了 3 种角色 (proposers, acceptors, 和 learners) 和 2 个阶段 (accept 和 learn)。其中 proposers 提出决议, acceptors 批准决议, learners 获取并使用已经通过的决议。在这种情况下, 满足以下 3 个条件^[3]就可以保证数据的一致性:

收稿日期: 2011-10-21

作者简介: 王红爱, 助理研究员; 祝红光, 副研究员。

(1) 决议只有在被 proposers 提出以后才能批准。

(2) 每次只批准一个决议。

(3) 只有决议确定被批准后 learners 才能获取这个决议。

图 1 为 Paxos 算法的模型简图。



图 1 Paxos 算法的模型简图

Paxos 算法实现时 proposer 向 acceptors 组发起请求告知需批准的编号 n ，获取 acceptors 多数派最新一次批准的最大编号 m ，如果此次提案编号 n 大于 m ，则 acceptors 根据收到的请求承诺不再批复编号小于 n 的提案，在收到该提案时即批准该请求，此时 learner 才能获取该提案。

2 数据复制与传输管理

数据复制 (replication) 和传输管理 (DBCS) 是客票分布式数据系统的 2 种技术手段。客票系统利用 Sybase 的复制服务器将数据从主站点复制到其他站点，通过在主点建立数据复制定义，在复制点建立复制约定，自动地把对主点数据的操作通过复制机制作用到复制点数据上。数据传输管理主要完成分布式数据库之间数据的相互透明传送，从而在分布式数据库之间实现数据的一致性和完整性，属于松散一致性的数据耦合机制。

2.1 数据复制

客票系统利用的复制服务器来完成复制功能。复制数据会在整个系统中有多份数据影像，分为紧密一致性和松散一致性两种方式，紧密一致性需要相关的复制节点都做好了提交准备才去提交，而松散一致性对于事物的执行主点和复制点总是会有延迟。数据复制具有以下优点：

(1) 高可用性

如果有一个节点出错，其他节点可以接管它的工作，不会使整个系统停止工作。

(2) 提高系统的扩展能力

通过把负载分布到多个节点上，或者增加节点来提高系统的扩展能力。

(3) 高性能

复制服务器提供一种机制，复制系统是独立的，不会给各节点服务器造成负担。同时客户端可以访问就近的节点，提高了用户访问速度。

2.2 数据传输

客票系统拥有自主开发的传输中间件 (DBCS)，主要根据用户定制，从源节点数据库按照一定的规则导出数据，传输到目的节点服务器，再按照导入定义将文件导入到相应的数据库中，完成日常数据在各节点之间的传送。包括原始业务数据向上级节点传输，同级节点之间的传输，消息类数据向下级节点传输等场景。

3 客票系统客户端与服务端架构

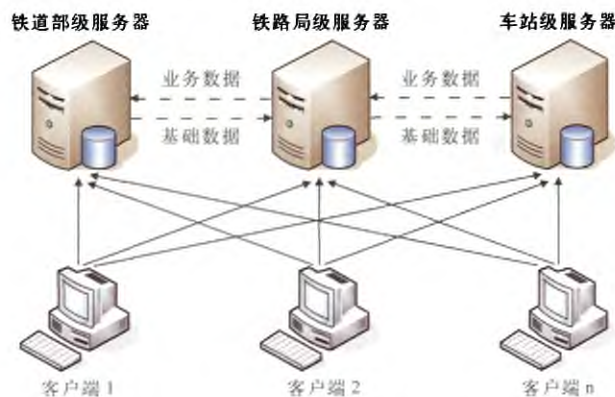


图 2 客票系统客户端与服务端的结构图

如图 2，以售票为例，当客户端要完成售票业务，需要分别向车站、铁路局、铁道部各级不同数据库发出指令，由于网络等原因，如果各服务器基础数据或者业务类数据不一致，可能会造成业务停滞，或者产生错误的结果，比如改点调令没有复制到车站，会造成票面开车时间不正确，导致旅客漏乘列车。为了使全路客票发售系统能够正常运行，数据的一致性尤为重要。在分布式系统中 Paxos 算法最大的用途就是保持多个节点数据的一致性。

客票系统是一个庞大的信息系统，根据业务的不同，有些需要采用分布式提交，提交意味着所有的操作均已完成。放弃意味着存在无法顺利

(下转 P30)

5.3 其他经济效益

售票设备数量的减少,除了节省设备购置费,还可以减少设备安装工料费,节省运维成本,节省房屋建筑面积和电源消耗。

6 结束语

通过本文的分析,可以得出如下结论:随着铁

路旅客订、购票和乘车实名制的实施,以及网络订票、电话订票和卡类支付应用于铁路客票发售和预定系统,不仅旅客受益,而且铁路客票系统可以充分利用这些社会资源,以较理想的方式实现铁路客票票制的统一,并对车站前端设备优化,以减少客票系统前端设备数量,降低部分前端设备采购成本,降低系统运维成本,经济效益十分明显。

责任编辑 杨利明

(上接 P25)

完成的操作,需要放弃,那些已执行的操作也将撤销,系统回滚到本次提交之前的状态,即必须保证提交的原子性,这是保证数据一致性的重要前提。

4 基于 Paxos 算法的数据一致性模型

数据一致性模型从根本上来说是对任务和数据存储之间的控制。正常情况下,执行读操作时,它期待该操作返回的是该数据在其最后一次写操作之后的结果。为了保证业务正常、正确地进行,需使所有服务器端的基础数据最终一致。对于基础数据,由铁道部向下级服务器逐级复制;对于业务数据,由车站向上级服务器逐级传输。本文基于 Paxos 算法提出了加强数据一致性的方案。

图3为基于 Paxos 算法的数据一致性模型,处理流程如下:

(1) 基础数据按照由上向下的复制机制,业务数据按照由下向上或同级间的传输机制。

(2) 业务维护时,会在同步控制器存储本次更新的记录(编码 n ,更新值 $newvalue$),其中 $value$ 值可以是表名加索引值,也可以是在业务维护时的校验值等。

(3) 铁道部、铁路局、车站各级数据库服务器作为节点集,通过 Paxos 算法推选出主节点。

(4) 同步控制器(客户端或工作流进程)将本次更新的记录(编码 n ,更新值 $newvalue$)向主节点发起更新申请。

(5) 主节点比较最新一次更新的记录(编码 m ,更新值 $value$),以下情况的更新是成功的:

$value = newvalue$, 且 $m < n$;

$value < > newvalue$

(6) 如果更新成功,说明新数据在某些节点

不一致,发送消息给管理端,启动快照机制或者重新复制等手段保证最终数据的一致性。

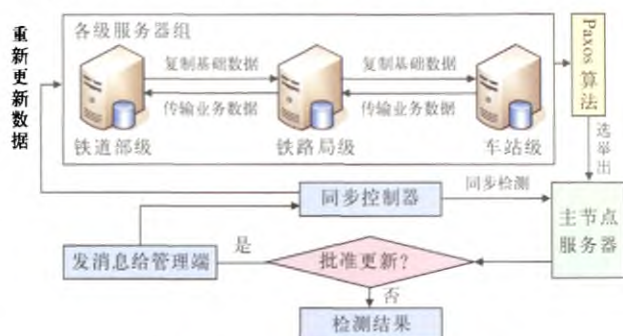


图3 基于paxos算法的数据一致性模型

5 结束语

铁路客票系统的业务在不断扩展,为给旅客提供方便快捷的服务,客票系统对业务连续性提出了更高要求,因此数据的一致性和完整性尤为重要。在云计算飞速发展的今天,其中的关键技术也被各个行业吸纳,本文首先对其中的Paxos算法进行了研究,建立了基于Paxos算法的数据一致性模型,该模型可以解决铁路客票系统中数据的一致性和完整性问题,进一步提高客票系统的业务质量。

参考文献:

- [1] LAMPORT, L. Paxos made simple[J]. ACM SIGACT News 32, 4 (2001): 18-25.
- [2] LAMPORT, L. The part-time parliament[J]. ACM TOCS 16, 2 (1998): 133-169.
- [3] 赵黎斌. 面向云存储的分布式文件系统关键技术研究[D]. 西安: 西安电子科技大学, 2011.

责任编辑 杨利明