

面向分布式交互应用的事件完全序问题研究

李永, 陆伟

(盐城师范学院 信息工程学院, 江苏 盐城 224002)

摘要: 针对分布式交互应用(DIA)中的事件完全序问题进行研究,在 DIA 中,发生在两个不同节点上的事件如何排序,如何判断某个事件当前是否可以提交处理是非常关键的。由于网络传输时延的异构,不同 DIA 节点接收到事件的顺序是不一样的,DIA 节点通常不能把事件的接收顺序作为处理顺序,也不能直接将已接收到的最小时间戳的事件作为当前需处理的事件,因为它无法判断是否有更小时间戳事件仍在网上传输,还未接收到。对此,提出了基于周期采样和事件序列号的 DIA 一致性控制方法(CCM_SE),CCM_SE 包括周期采样机制和事件序列号机制两个方面。CCM_SE 为了确保事件完全序一致性,不同节点上发生的事件通过时间戳来排序,并通过事件序列号机制来及时判断一个事件是否可以提交执行,以提高 DIA 事件的响应性。模拟仿真表明,CCM_SE 可扩展性好,响应性好,并且能有效减少 DIA 中不一致现象的发生。

关键词: 分布式交互应用;事件完全序;周期采样;事件序列号

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2019)04-0072-06

doi:10.3969/j.issn.1673-629X.2019.04.015

Research on Event Total Order Problem for Distributed Interactive Application

LI Yong, LU Wei

(School of Information Engineering, Yancheng Teachers University, Yancheng 224002, China)

Abstract: The event total order consistency problem in distributed interactive application (DIA) is studied. How to sort events occurred on two different nodes and how to determine whether an event may submit the current process is critical in DIA. Due to the heterogeneity of network transmission delay, the order of events received by different DIA nodes is different. DIA nodes generally cannot take the received order of events as the processing order, nor can they directly take the received minimum timestamp event as the current event to be processed, because it cannot determine whether there are smaller timestamp events still transmitted on the network and not yet received. Therefore, we propose a DIA consistency control method (CCM_SE) based on periodic sampling and event sequence number, which includes two aspects: periodic sampling mechanism and event sequence number mechanism. In order to ensure event total order in CCM_SE, events occurring on different nodes are sorted by timestamps, and the event sequence number mechanism is used to determine whether an event can be submitted for execution to improve the responsiveness of the DIA events. Simulation shows that the CCM_SE has better scalability and response, and can effectively reduce the occurrence of inconsistency in DIA.

Key words: distributed interactive application; event total order; periodic sampling; event sequence number

1 概述

随着计算机网络的飞速发展,各种计算机应用不再局限于单机的或独立的系统,而是通过网络进行连接,以分布式的形式开展各种应用,地理上分散的多个终端用户节点能够在一个共享的统一网络虚拟环境中通过分布式交互和协作完成特定任务。这类应用称为分布式交互应用(distributed interactive application, DIA),是近年来的研究热点^[1-4]。目前分布式交互应

用主要体现在分布式虚拟环境、分布式交互仿真、协同设计与制造、军事仿真系统、网络在线游戏、视频会议等方面。

DIA 中事件序一致性问题解决的好坏会直接影响其可用性,在 DIA 中所有参与同一应用的端节点应该具有实体状态的一致性和事件处理顺序的一致性。DIA 中的事件一致性往往是有延时约束一致性,因此,良好的 DIA 一致性控制方法主要体现在两个方面:一

收稿日期:2018-04-13

修回日期:2018-08-16

网络出版时间:2018-12-20

基金项目:国家自然科学基金(61603326);盐城师范学院自然科学基金(14YCKL016)

作者简介:李永(1978-),男,博士,副教授,研究方向为计算机网络体系结构、分布式计算。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20181220.1001.010.html>

致性和响应性。一致性能够保证 DIA 在功能上的正确性,实时响应性能够给用户良好的 DIA 体验。由于存在网络传输延迟,各节点间的事件消息不一定能及时到达,会产生消息序错乱等情况,进而导致理解歧义、违背期望、因果颠倒^[5]等问题,严重影响交互应用的正常运行;并且一致性问题还会影响系统的公平性和正确性,使得最终应用得到的结果不可信;因此,能否有效解决交互应用中各节点间的一致性问题,已经成为了制约分布式交互应用系统进一步发展的关键。近年来,针对一致性问题的研究一直是 DIA 的热点,许多一致性控制方法相继提出。

Mauve 提出了本地滞后方法^[6-7],事件在发送节点产生后要延迟一段时间才能在本地产地执行,其延迟时间最好不要小于发送节点和所有接收节点间通信延迟的最大值;该方法通过牺牲系统的响应能力来增强系统的一致性,是对一致性和响应性的折衷,故该方法受系统整体的通讯情况影响较大。Thomas 提出了基于锁机制的同步技术^[8],该方法强制阻塞节点的时钟推进,为了避免了不一致现象的发生,当所有节点完成当前时间步的事件操作后才进入下一个时间步。Qin X 提出了基于间距一致的延迟一致性^[9-10],对于系统中全局事件在各节点的延迟时间由接收节点决定;对于某个接收节点来说,其接收的全部来自异地的消息都延迟相同的时间,而发送节点在发出事件后在本地产地不经过延迟立即执行,使系统各节点获得较好的响应能力,但该方法会造成不一致现象的发生。例如,如果两个节点同时向对方发送消息,由于发送方在发送消息后不延迟直接执行该事件,因此两个消息在这两个节点上的执行顺序刚好相反。总之,该方法采用只延迟异地事件而不延迟本地事件的方法来提高系统的响应性,以牺牲系统功能来达到提高性能的方法,无法真正达到事件完全序一致性,会导致不一致现象的发生。Jefferson 提出了 Time warp^[11],DIA 节点对接收到的事件立即执行,当出现不一致时,就通过回滚操作纠正 DIA 的错误;该方法在网络条件不好时性能很差。Roberts 提出了预测时间管理技术^[12],在 DIA 事件发生前对其进行预测,并通知其他节点该事件即将发生的时间和内容,当预测准确时该技术能够提高系统的性能;当预测错误时,执行回滚操作纠正错误的结果。但只有少数 DIA 事件能够预测,且预测准确率^[13]。

综上,已经提出的许多 DIA 一致性控制方法,在一定程度上能够满足 DIA 的需求,但由于 DIA 一致性问题本身的难度和复杂性,针对 DIA 一致性问题的研究还有很大的探索空间。文中研究的目的是在确保 DIA 事件完全序一致性的前提下,提高系统响应性,增强 DIA 的用户体验。提出了面向 DIA 的基于周期采

样和事件序列号的一致性控制方法(consistency control method based on periodic sampling and event sequence number,CCM_SE),CCM_SE 包括周期采样和事件序列号机制两个方面,其中周期采样用来确保 DIA 的一致性,事件序列号机制用来及时确定事件的可处理时刻,以提高 DIA 事件的响应性。

2 DIA 事件完全序一致性问题描述

在 DIA 中,地理上分散的节点通过网络进行连接,不同节点间事件的发生只能通过消息的发送、传输、接收和提交等实现节点间的交互。消息主要由两部分构成:事件内容和一致性控制信息。对事件消息处理的关键步骤是确定节点上接收的事件消息何时被提交执行,接收事件的 DIA 节点通过处理接收到消息的一致性控制信息,确保事件能够在该节点以正确的顺序执行。

在 DIA 中发生在两个不同节点上的事件如何排序,如何判断某个事件当前是否可以提交处理是非常关键的。由于网络传输时延的异构,不同 DIA 节点接收到事件的顺序是不一样的,一个 DIA 节点显然不能把事件的接收顺序作为处理顺序,也不能将已接收到的最小时间戳的事件作为当前需处理的事件,因为它无法判断是否有更小时间戳事件仍在网上传输,还未接收到。为了突出文中的研究内容,这里在时钟同步的前提下研究事件完全序一致性问题。CCM_SE 采用统一时钟来刻画不同节点上事件发生的时间,维护事件间的先后顺序。同时,进一步结合事件序列号机制及时确定事件的可处理时刻,提高 DIA 的响应性。CCM_SE 事件完全序一致性问题描述如下:

V 表示 DIA 内节点的集合; E 表示 DIA 中事件的集合。

对于任意的事件 $e_m \in E$,则 $g(e_m)$ 表示事件 e_m 的产生节点; $R(e_m)$ 表示能够接收到事件 e_m 的节点集合; $F(e_m) = \{g(e_m)\} \cup R(e_m)$ 表示可感知到事件 e_m 的节点集合。

对于任意的节点 $v_i, v_j \in V$,则 $tg_i(e_m)$ 表示事件 e_m 在 v_i 上的产生时间; $tr_j(e_m)$ 表示节点 v_j 接收到事件 e_m 的时间; $te_j(e_m)$ 表示事件 e_m 在 v_j 上开始执行的时间。 d_{ij} 表示节点 $v_i, v_j \in V$ 间的通信延时。

在 DIA 中,节点间只能通过收发消息来通信,消息传输延迟的动态性使事件的天然先后关系在各个节点失去了必然性。因此,针对分 DIA 中的事件定义一种排序关系,即“发生在先”关系,用于描述真实世界中的因果事件先后关系^[14]。

定义 1:“发生在先”关系,满足下面的两个条件之一。

(1) 若 a 、 b 是同一 DIA 节点上发生的两个事件, 且 a 在 b 之前发生, 则 $a \rightarrow b$;

(2) 若 a 是一个发送消息事件, b 是另一个 DIA 节点相应的接收消息事件, 则 $a \rightarrow b$ 。

对于不同于 a 、 b 的任意事件 c , 由“happened before”发生在先关系, 可以得到: 若 $a \rightarrow b$ 且 $b \rightarrow c$, 则 $a \rightarrow c$ 。

定义 2: 设 e_m 、 e_n 是 DIA 中的两个事件, 若 e_m 、 e_n 满足 $\neg(e_m \rightarrow e_n) \wedge \neg(e_n \rightarrow e_m)$, 则称 e_m 、 e_n 是 DIA 中的两个并发事件, 记为 $e_m \parallel e_n$ 。

定义 3: 事件完全序一致性, 即 DIA 中不同节点按相同的顺序执行事件:

$$\forall e_m, e_n \in E; v_i, v_j \in F(e_m) \cap F(e_n); te_i(e_m) \leq te_i(e_n) \rightarrow te_j(e_m) \leq te_j(e_n) \quad (1)$$

定义 4: 事件开始执行时间一致性, 即任意一个事件在不同节点上的开始执行时间都相同。

$$\forall e_m \in E; v_i, v_j \in F(e_m); te_i(e_m) \leq te_j(e_m) \quad (2)$$

事件开始执行时间一致性是要求最为严格的一致性。

定义 5: 事件的响应时间, 即事件从触发到开始提交执行的间隔时间。

$$\forall e_m \in E; v_i = g(e_m); rt_i(e_m) = te_i(e_m) - tg_i(e_m) \quad (3)$$

定义 6: 节点的响应时间, 即某个节点上产生的所有事件的响应时间的平均值。

$$\forall e_m \in E; v_i = g(e_m); rt_i = \text{avg}[te_i(e_m) - tg_i(e_m)] \quad (4)$$

定义 7: 系统的响应时间, 即系统中所有节点响应时间的平均值。

$$RT_{\text{system}} = \frac{\sum_{\forall v_i \in V} rt_i}{|V|} \quad (5)$$

3 周期采样机制

针对 DIA 的事件完全序一致性控制问题, 提出的事件完全序一致性控制方法 CCM_SE 主要包括周期采样机制和事件序列号机制。下面介绍 CCM_SE 的周期采样, 并分析其一致性。

在 CCM_SE 中, 各节点采用时钟同步机制进行时钟同步, 使得不同 DIA 节点上触发的事件可以通过事件时间戳来确定事件的发生顺序。CCM_SE 的周期采样机制把 DIA 中的时间划分为固定的时间长度 T , 即采样周期。采样周期 T 表明多长时间更新一次 DIA 的状态, 采样周期的频率 $1/T$ 决定 DIA 状态的更新速度。在实际应用中, 以 25 帧每秒的频率 ($T = 40 \text{ ms}$) 更新 DIA 状态对用户来说已经在视觉上很连贯了^[15]。这里通常取固定的 T ($T \leq 40 \text{ ms}$)。CCM_SE 的周期采样机制为了保证一致性, 一个 DIA 节点在第 i 个采样周期 $[t_i - T, t_i]$ 之内, 提交执行发生在 $[t_i - T - \Delta, t_i - \Delta]$ 时间间隔内的所有事件。CCM_SE 的周期采样机制的原理如图 1 所示。

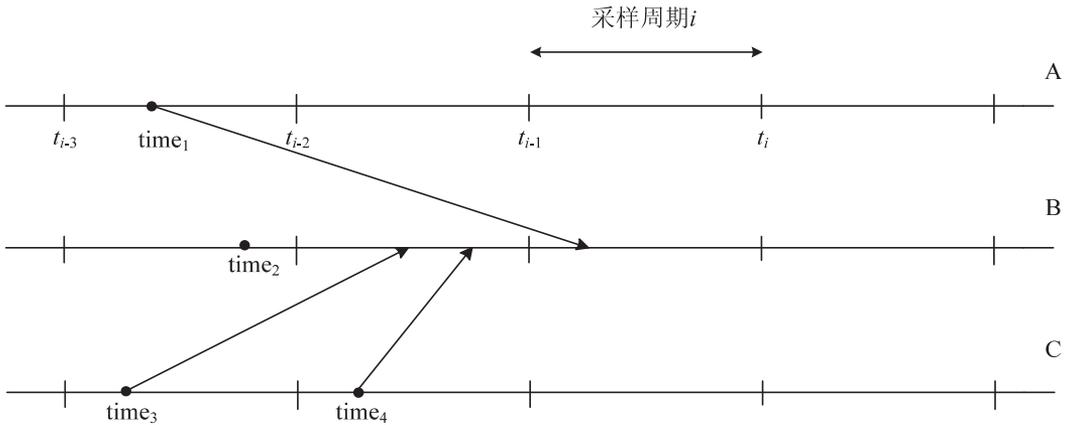


图 1 CCM_SE 周期采样机制原理

图 1 中的三个 DIA 节点 A、B、C, 节点 A 在时间 $time_1$ 触发了事件 e_{A1} , 节点 B 在时间 $time_2$ 触发了事件 e_{B2} , 节点 C 在时间 $time_3$ 和 $time_4$ 触发了事件 e_{C3} 、 e_{C4} , 节点 B 在第 i 个采样周期 $[t_i - T, t_i]$ 将按序提交执行发生在 $[t_i - T - \Delta, t_i - \Delta]$ 时间间隔内的事件 e_{A1} 、 e_{B2} 和 e_{C3} , 对于已接收到的 e_{C4} 将不会执行。

在 CCM_SE 的周期采样机制中, 为了抵消 DIA 节点间因网络延时不同造成的一致性影响, 一个 DIA 节点在第 i 个采样周期 $[t_i - T, t_i]$ 内, 提交执行发生在

$[t_i - T - \Delta, t_i - \Delta]$ 时间间隔内的所有事件。可见 DIA 节点触发一个事件后, 不是立即执行该事件, 而是延迟一定时间 Δ 后再提交执行。事件滞后时间 Δ 后再提交执行, 可以消除或显著减小所有 DIA 节点间的不一致现象。

下面利用上面对于 DIA 事件完全序一致性问题的形式化描述和定义, 分析 CCM_SE 的周期采样机制的一致性。

定理 1: 在 CCM_SE 的周期采样中, 当 Δ 的取值满

足式6时,CCM_SE的周期采样能够达到定义4中描述的“事件开始执行时间一致性”。

$$\forall e_m \in E, v_i = g(e_m) : rt_i = \max_{v_j \in R(e_m)} (d_{ij}) \quad (6)$$

$$\Delta \geq \max \{ rt_i \}, v_i \in V$$

证明:由 $\forall e_m \in E, v_i = g(e_m) : rt_i = \max_{v_j \in R(e_m)} (d_{ij})$ 可知,只有当一个节点 v_i 触发的事件 e_m 在该节点的响应时间 rt_i 等于其与所有接收节点间传输延时的最大值 $\max_{v_j \in R(e_m)} (d_{ij})$ 时,才能保证事件 e_m 在任意可以感知到该事件的节点上的开始执行时,事件 e_m 已被所有能够接收到事件 e_m 的节点 ($v_j \in R(e_m)$) 接收到。

在 CCM_SE 的周期采样机制中,一个 DIA 节点在第 i 个采样周期 $[t_i - T, t_i]$ 内,提交执行发生在 $[t_i - T - \Delta, t_i - \Delta]$ 时间间隔内的所有事件。当 $\Delta \geq \max \{ rt_i \}, v_i \in V$ 时,使得在周期 $[t_i - T - \Delta, t_i - \Delta]$ 内发生的任何事件都能够被传输到所有接收节点,从而实现在能够感知某个事件的所有节点上,同时执行该事件,定理1得证。证毕。

“事件开始执行时间一致性”是要求最为严格的一致性。因此,周期采样机制能够实现 DIA 的事件完全序一致性。在 CCM_SE 的周期采样机制中,每个节点维护自己的 DIA 状态的改变,DIA 状态的改变与事件的提交执行相对应。事件的推进是 DIA 的基础,它由“事件接收”和“事件处理”两个过程构成。CCM_SE 的周期采样为了保证交互的正确性,一个 DIA 节点在第 i 个采样周期 $[t_i - T, t_i]$ 内,提交执行发生在 $[t_i - T - \Delta, t_i - \Delta]$ 时间间隔内的所有事件。

4 基于周期采样和事件序列号的一致性控制方法 CCM_SE

针对 DIA 的事件完全序一致性问题,若只采用周期采样机制,在第 i 个采样周期 $[t_i - T, t_i]$ 将按序提交执行发生在 $[t_i - T - \Delta, t_i - \Delta]$ 时间间隔内的事件,该方法存在把所有的事件都滞后执行的缺点,即没有对接收到的事件进行判断是否可以及时提交执行。对一些已经接收到的满足可执行条件的事件,由于它们不是发生在 $[t_i - T - \Delta, t_i - \Delta]$ 间隔内,而得不到执行。因此这里引入事件序列号机制,在周期采样机制中结合事件序列号来判定一个事件何时可以提交执行。

事件序列号机制主要用来判断来自同一个节点的两个事件是不是连续事件,即若节点 v_i 接收到来自节点 v_j 的两个事件 e_{j1} 和 e_{j2} ,需要判定是否在 e_{j1} 和 e_{j2} 之间 v_j 是否触发了其他事件如 e_{j3} ,只是由于网络延迟的原因,节点 v_i 还没有接收到 e_{j3} 。

若 e_i, e_j 分别是 DIA 节点 v_i, v_j 上产生的两个事件,产生事件的时间戳分别为 t_i, t_j ,且 $t_i < t_j$,则在所

有 DIA 节点上 e_i, e_j 的处理顺序都必须为 $e_i \rightarrow e_j$ 。若 e_i 是 v_i 上正处理的事件, e_j 是 v_i 上未处理事件中时间戳最小的事件,则 v_i 处理完 e_i 后能否立刻处理 e_j ? 如何判断是否存在一个事件 e_k 且 $t_i < t_k < t_j$,该事件仍在传输。若由同一节点 v_i 产生的两事件分别记为 $e_{i,p}, e_{i,q}$,其时间戳对应为 $t_{i,p}, t_{i,q}$,在节点 v_i 处理完 $e_{i,p}$ 之后,需要判断 v_i 是否产生了一个事件 $e_{i,r}$ ($t_{i,p} < t_{i,r} < t_{i,q}$)。

定义8: DIA 事件序列号。若把节点 v_i 生成的第一个事件的序列号设置为1,把在节点 v_i 发生的事件序列号为 p 的事件记为 $e_{i,p}$,则 $e_{i,p}$ 的下一个事件的序列号为 $p+1$,该事件记为 $e_{i,p+1}$ 。

可见,同一节点 v_i 上触发事件的先后关系可用事件序列号来描述。基于事件序列号可以判断 v_i 在处理完 $e_{i,p}$ 后,是否能接着处理 $e_{i,q}$ 。

(1) 若 $q > p+1$,则在事件序列号为 $p+1, \dots, q-1$ 的事件处理之前,不能处理 $e_{i,q}$;

(2) 若 $q = p+1$,则在 v_i 上不存在时间戳为 $t_{i,r}$ 的事件 $e_{i,r}$,使得 $t_{i,p} < t_{i,r} < t_{i,q}$ 。

定义9: DIA 事件的属性除了包含物理操作,还要具有以下属性:

(1) 产生事件的 DIA 节点;

(2) 事件在 DIA 节点上产生时的事件序列号;

(3) 事件在 DIA 节点上产生时的时间戳。

定义10: 已处理事件的序列号向量。 v_i 为任意一个 DIA 节点,其上已处理事件的序列号向量为 $(D_1, D_2, \dots, D_j, \dots, D_n)$,其中, $D_j, j = 1, 2, \dots, n$ 表示从 DIA 节点 v_j 传输到 v_i 的事件中, v_i 已处理的最大的事件序列号。

定义11: 未处理事件的序列号向量。 v_i 为任意一个 DIA 节点,其上未处理事件的序列号向量为 $(D'_1, D'_2, \dots, D'_j, \dots, D'_n)$,其中 $D'_j, j = 1, 2, \dots, n$ 的计算如下:若 v_i 已接收到 v_j 生成的序列号为 D_j+1 的事件,则 $D'_j = D_j + 1$; 否则 $D'_j = \emptyset$ 。

定理2: v_i 为任意一个 DIA 的节点, $(D'_1, D'_2, \dots, D'_j, \dots, D'_n)$ 为 v_i 未处理事件的序列号向量, e 为 $(D'_1, D'_2, \dots, D'_j, \dots, D'_n)$ 中对应事件时间戳最小事件,则当所有 $D'_j = \emptyset, j = 1, 2, \dots, n$ 时, e 就是当前应处理事件。

证明:假设 e' 为 $v_j, j \in \{1, 2, \dots, n\}$ 上生成的任一事件, $e' \neq e$ 而且 e' 没有在 v_i 上被处理。 $(D_1, D_2, \dots, D_j, \dots, D_n)$ 为 v_i 的已处理事件的序列号向量,因 v_i 上的事件按时间戳顺序进行处理,故未处理事件 e' 的时间戳大于 $(D_1, D_2, \dots, D_j, \dots, D_n)$ 对应的所有已处理事件的时间戳;若 e' 为 $(D'_1, D'_2, \dots, D'_j, \dots, D'_n)$ 对应的事件,由于 e 为 $(D'_1, D'_2, \dots, D'_j, \dots, D'_n)$ 中对应事件时

间戳最小的事件,故 e' 的时间戳大于 e 的时间戳。

若 e' 不在 $(D'_1, D'_2, \dots, D'_j, \dots, D'_n)$ 对应的事件中, 则 e' 的序列号大于 D'_j 对应事件的序列号, 即事件 e' 的时间戳大于 D'_j 对应事件的时间戳, 而 D'_j 对应事件的时间戳又大于 e 的时间戳, 故 e' 的时间戳大于 e 的时间戳。综上, v_i 上任一未被处理的事件的时间戳都大于 e 的时间戳, 故 e 为当前应处理的事件, 定理 2 得证。证毕。

由定理 2 可知, 对任一 DIA 节点 v_i , 当未处理事件的序列号向量为 $(D'_1, D'_2, \dots, D'_j, \dots, D'_n)$ 的所有值不为 \emptyset 时, 可处理 $(D'_1, D'_2, \dots, D'_j, \dots, D'_n)$ 对应事件中时间戳最小的事件; 而当 $(D'_1, D'_2, \dots, D'_j, \dots, D'_n)$ 中至少有一个为 \emptyset 时, 如 $D'_j = \emptyset$, 表明在 v_i 处理完 v_j 发来的事件后, v_j 的下一个序列号事件 v_i 还未接收到, 而该下一个序列号事件的时间戳有可能小于 v_i 上所有未处理事件的最小时间戳, 故要等待所有 D'_j 不为 \emptyset 。

基于事件序列号机制可以及时判定一个事件是否可以提交执行, 一个发生在 $[t_i - T - \Delta, t_i - \Delta]$ 时间间隔的事件, 并非一定要滞后在 $[t_i - T, t_i]$ 内被提交执行。当然如果一个发生在 $[t_i - T - \Delta, t_i - \Delta]$ 时间间隔的事件, 在 Δ 取值比较小的情况下, 会造成在 $[t_i - T, t_i]$ 内仍然没有达到。此种情况, 在采样周期末, 可以对已接收到且未提交执行的发生在 $[t_i - T - \Delta, t_i - \Delta]$ 时间间隔的事件进行提交执行, 以便更新用户节点的状态, 但同时将在以后的时间里对迟到事件的到来采用修复机制。对于 DIA 交互性控制修复机制的研究将在以后进行。

5 仿真模拟与分析

由于 DIA 的全部到全部路由具实时性、高带宽等特点, 为每一个需要发送数据的 DIA 节点都以它为根构造一棵数据分发树, 费用开销太大, 而所有需要发送数据的 DIA 节点都基于单棵共享树进行数据分发树, 又会造成流量集中, DIA 延时无法保障。因此, 文中采用多棵共享树来分发 DIA 数据, 构造了基于多共享树的 DIA 全部到全部路由问题模型 DARP, 并提出了求解 DARP 问题的禁忌遗传算法 DARP_TGA。

为了验证上面提出的 DIA 事件完全序一致性控制方法 CCM_SE 的有效性, 需要确定仿真模拟的实验环境和相关参数的设置。这里采用通用的拓扑产生器 Brite 生成网络拓扑数据, 模拟底层的物理网络, 采用随机模型 Waxman 作为实验拓扑的生成算法, 生成具有实际网络特点的拓扑图。在多组实验中, 将提出的 CCM_SE 与文献[6-7]中的本地滞后(local lag, LL)以及文献[9-10]中的延迟一致性控制(delayed consis-

ency, DC) 在 DIA 事件不一致百分比和 DIA 平均响应时间(即定义 5 的“系统的响应时间”)上进行对比, 得到的实验结果分别如图 2 和图 3 所示。

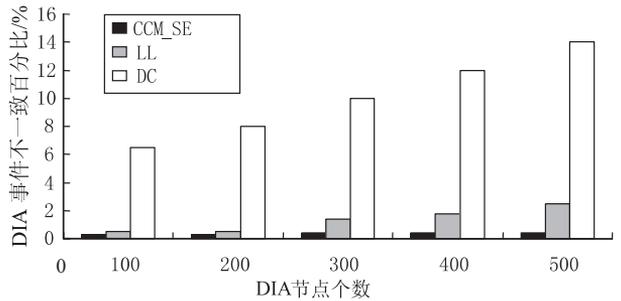


图 2 CCM_SE 的一致性控制效果

LL 为了保证节点间的状态一致性, 事件在发送节点产生后要延迟一段时间才能在本地执行, 其延迟时间大小很难确定, 因为 DIA 节点间通信延迟是动态的; LL 通过牺牲系统的响应能力来增强系统的一致性, 故该方法受系统整体的通讯情况影响较大, 在 DIA 规模小和网络稳定的情况下, 事件完全序一致性较好, 但该方法事件平均响应时间不好。延迟一致性控制 DC 采用只延迟异地事件而不延迟本地事件的方法来提高系统的响应性, 属于以牺牲系统功能来达到提高 DIA 响应性的方法, 故事件平均响应时间较好, 但通常 DC 无法满足消息序的一致性, 会导致不一致的现象的发生。

由图 2 可知, 由于 DC 无法保证事件发生的完全序关系, 事件不一致现象相对较多; 而文中提出的 CCM_SE 和 LL 的事件一致性控制较好, CCM_SE 和 LL 在节点规模小时事件不一致百分比相近, 但当节点规模增大时, 由于 LL 由发送节点确定延迟值, 受 DIA 网络的规模和网络延时的动态性影响较大, 从而导致 LL 中事件不一致现象增多。由图 3 可知, CCM_SE 和 DC 的事件平均响应时间相差不大, 都远远好于 LL 的事件平均响应时间, 其原因在于 CCM_SE 中引入了事件序列号机制, 能够及时判定满足执行条件的事件尽快提交执行, 而 DC 不延迟本地事件的执行, 所以事件的事件平均响应时间小。

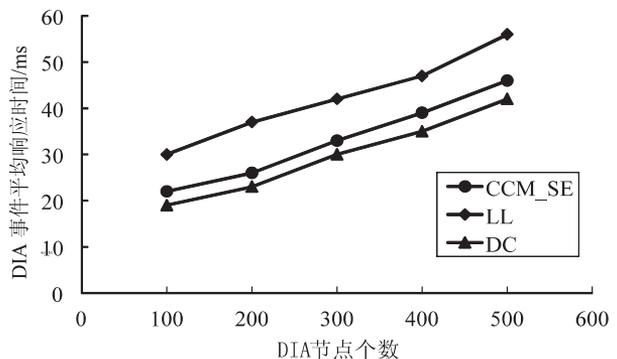


图 3 CCM_SE 的事件平均响应时间

6 结束语

在研究现状的基础上,提出了基于周期采样和事件序列号的 DIA 事件完全序一致性控制方法 CCM_SE, CCM_SE 包括周期采样机制和事件序列号机制两个方面。其中周期采样可以确保 DIA 的一致性;事件序列号机制可以及时确定事件的可处理时刻,一个发生在 $[t_i - T - \Delta, t_i - \Delta]$ 时间间隔的事件,并非一定要滞后在 $[t_i - T, t_i]$ 内被提交执行。仿真结果表明,该方法在确保 DIA 事件完全序一致的前提下,提高了系统响应性,增强了 DIA 的用户体验。

参考文献:

- [1] 郑晶晶,张晶,武继刚. 分布式交互应用中服务器放置问题的启发式算法[J]. 计算机科学,2015,42(7):95-98.
- [2] ZHANG Lu, TANG Xueyan, HE Bingsheng. Analysis of minimum interaction time for continuous distributed interactive computing[J]. IEEE Transactions on Parallel & Distributed Systems,2017,28(2):401-415.
- [3] 刘冬,赵跃龙. 实时在线交互应用的体系结构与状态一致性[J]. 计算机工程与科学,2014,36(3):381-387.
- [4] 李春洪,叶保留,顾铁成,等. 适用于大规模分布式交互仿真的应用层组播研究[J]. 系统仿真学报,2004,16(6):1271-1274.
- [5] 张伟,周航军,彭宇行,等. 分布式交互仿真中的异步时钟一致性控制方法[J]. 软件学报,2010,21(6):1208-1219.
- [6] MAUVE M, VOGEL J, HILT V, et al. Local-lag and time-warp: providing consistency for replicated continuous applications[J]. IEEE Transactions on Multimedia,2004,6(1):47-57.
- [7] SITHU M, ISHIBASHI Y, FUKUSHIMA N. Effects of dynamic local lag control on sound synchronization and interactivity in joint musical performance [J]. ITE Transactions on Media Technology & Applications,2014,2(4):299-309.
- [8] THOMAS A F. RING: a client-server system for multiuser virtual environments [C]//Proceedings of the 1995 symposium on interactive 3D graphics. Monterey, California, USA: ACM, 1995:85-92.
- [9] QIN X. Delayed consistency model for distributed interactive systems with real-time continuous media [J]. Journal of Software,2002,13(6):1029-1039.
- [10] LI Yusen, CAI Wentong. Consistency-aware zone mapping and client assignment in multi-server distributed virtual environments [J]. IEEE Transactions on Parallel & Distributed Systems,2015,26(6):1570-1579.
- [11] ANTONACCI F, PELLEGRINI A, QUAGLIA F. Consistent and efficient output-streams management in optimistic simulation platforms [C]//ACM SIGSIM conference on principles of advanced discrete simulation. Montr@al, Québec, Canada: ACM, 2013:315-326.
- [12] KENNY A, MCLOONE S, WARD T. Controlling entity state updates to maintain remote consistency within a distributed interactive application [J]. ACM Transactions on Internet Technology,2009,9(4):15.
- [13] DELANEY D, WARD T, MCLOONE S. Reducing update packets in distributed interactive applications using a hybrid model [C]//16th international conference on parallel and distributed computing systems. Reno, Nevada, USA: ISCA, 2003:417-422.
- [14] 周航军,张伟,彭宇行,等. 异步 DVE 系统中生命周期约束下的因果一致性控制方法研究[J]. 计算机研究与发展,2011,48(9):1768-1780.
- [15] DIOT C, GAUTIER L. Distributed architecture for multiplayer interactive applications on the Internet [J]. IEEE Network, 1999,13(4):6-15.