

Infiniband 网络架构下 RTI 通信机制研究

邢 驰^{1,2}, 李伯虎^{1,3,4}

(1. 北京航空航天大学自动化科学与电气工程学院, 北京 100191; 2. 北京仿真中心, 北京 100854;
3. 北京市复杂产品先进制造系统工程技术研究中心, 北京 100854; 4. 复杂产品智能制造系统技术国家重点实验室, 北京 100854)

摘 要: 复杂系统的协同仿真中需要运行支撑软件 RTI (Run Time Infrastructure) 来解决异构模型、异构仿真软件间的数据交互的问题. 但 RTI 的 TCP/IP 通信机制却无法使得 HPC (High Performance Computer) 的高速网络 Infiniband (IB) 在仿真中发挥最大的优势. 针对这一问题, 本文提出在 IB 网络架构下基于 RDMA (Remote Direct Memory Access) 通信机制对 RTI 进行优化, 并以开源 HLA 项目 CERTI 软件为基础, 研制运行在 IB 网络下的 IB-CERTI 软件, 最后在不同网络环境下进行对比实验, 实验结果证明了 IB-CERTI 软件在仿真通信中的高效性, 特别是仿真成员间的交互数据量越大, 越能提高仿真数据传输效率.

关键词: Infiniband 网络; RDMA; CERTI; 高层体系结构; 高性能计算机

中图分类号: TP319 **文献标识码:** A **文章编号:** 0372-2112 (2016)02-0327-07

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2016.02.012

Research on RTI Communication Mechanism on Infiniband Network Architecture

XING Chi^{1,2}, LI Bo-hu^{1,3,4}

(1. College of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China; 2. Beijing Simulation Center, Beijing 100854, China; 3. Beijing Complex Product Advanced Manufacturing Engineering Research Center, Beijing 100854, China;
4. State Key Laboratory of Intelligent Manufacturing System Technology, Beijing 100854, China)

Abstract: Collaborative simulation of complex systems need Run Time Infrastructure (RTI) software to solve the problem of data exchange between heterogeneous models and heterogeneous simulation software. But the TCP/IP communication mechanism of RTI cannot make HPC High Speed Network in the simulation which is the strongest advantages of Infiniband (IB). To solve this problem, this research designs the RTI network optimization based on RDMA communication mechanism, which through the module decomposition and network optimization of CERTI software in open source HLA project, made the IB-CERTI software which run on IB. Eventually, in the experiment of comparison RTI software network performance under different network environment, the experiment results show the efficiency of the IB-CERTI software under simulation communication, especially the larger amount of interactive data between simulation federates, the more efficiency in improving simulation data transmission.

Key words: Infiniband network; remote direct memory access (RDMA); CERTI; high level architecture (HLA); high performance computer (HPC)

1 引言

仿真规模不断扩大、系统复杂度提高, 复杂系统^[1,2]迫切需要仿真支撑环境提供更高的计算能力和网络通信能力以及实时性来满足仿真需求, 传统的计算机硬件和网络架构已经无法支撑仿真技术的迅速发展. 随着高性能计算机和高速通信网络技术的发展, 高性能计算机 HPC 成为 HLA (High Level Architecture) 新的硬件环境,

HPC-HLA 也成为了仿真技术的研究新热点.

HPC 多核的特点以及 HPC 节点之间的高速通信网络能够为 HLA 仿真提供多种的通信环境和通信方式^[3], 如在一个计算节点内可采用共享内存方式、计算节点间通过 IB 网络连接、连接半实物仿真器以及外部子系统则可使用以太网. 目前比较成熟的 RTI 软件, 均支持 TCP/IP 通信机制, 此外也有部分商业公司或研究人员将共享内存式的通信机制加入到 RTI 软件中, 如

Richard Fujimoto 等在 SMP (Symmetric Multi-Processor) 的硬件环境下为 FDK 增加了共享内存通信^[4];法国国家航天航空研究中心基于 CERTI 实现了共享内存通信的 SHM-CERTI,并用于飞机的研制^[5];国防科技大学姚益平教授团队针对高性能计算平台的高性能层次式通信环境,提出了 SACM (Self-Adaptive Communication Mechanism) 机制^[6] (共享网络/IB 网络/以太网自适应 RTI 通信机制),使得基于 HPC 的不同网络通信环境下仿真系统能够自适应选择通信方式进行互联;作者所在团队的研究人员张智慧研究实现了基于共享内存的 COSIM-RTI-SHM^[7]. 综上可见,HLA 被移植到 HPC 上后,研究人员将 RTI 的底层通信作为一个新的研究领域,针对高性能计算机的 RTI 通信机制的研究大多是基于共享内存机制的,即为单计算节点之内,多核之间的通信方式,对于计算节点之间,基于 IB 网络的 RTI 通信机制的研究还尚未有较为成熟的研究成果.

而基于 IB^[8,9] 的 RDMA 协议在数据传输过程中能够避免 4 次数据拷贝(发送和接收各两次),大大减少传输过程中消耗在数据复制上的时间,从而能够缩短仿真运行的时间. 本文的研究内容是基于 IB 网络架构下的 RTI 通信机制,通过对开源 HLA 项目 CERTI 软件的分析和网络通信的优化,实现了 IB-CERTI 软件,并编写测试程序后在不同网络环境下进行多次的测试,给出了测试结果以及对结果进行分析,最后给出研究结论.

2 Infiniband 架构与 RDMA 通信机制

2.1 Infiniband 架构

IB^[10] 是一种硬件网络架构,它是支持多并发链接的转换线缆技术,被用于高性能计算中心的节点之间的通信,计算机与存储器之间的通信等,其拥有高带宽、低延迟、低 CPU 开销等特点,不同的配置下使用不同的编码方式,能够达到不同的速度和网络延迟,数据交互能力可达到每端口 2.5Gb/s 或 10Gb/s(根据硬件配置的不同).

IB 上能够运行多种不同的协议,主流的如 IPoIB (Internet Protocol over InfiniBand), RDS (Reliable Datagram Sockets) 等,如图 1 所示在单个计算机节点内,需要交互的数据在不同的协议下如何被准备好,而后通过硬件网络板卡发送,其中阴影部 TCP/IP 协议不运行在 IB 架构下,只用来作为其他协议的对比说明.

TCP/IP 协议:程序数据通过 TCP/IP 将需要交互的数据拷贝到 NIC(Network Interface Card)准备发送.

IPoIB 协议:目的是为了兼容基于 TCP/IP 的软件程序. 传统软件的通信部分几乎都是基于 TCP/IP 的,为了能够在 IB 网络下运行,IPoIB 才被提出,此部分对用户完全透明,用户可以将基于 TCP/IP 的程序不做任何改动运行在 IB 网络下,以享用 IB 的高带宽能力. 同样,

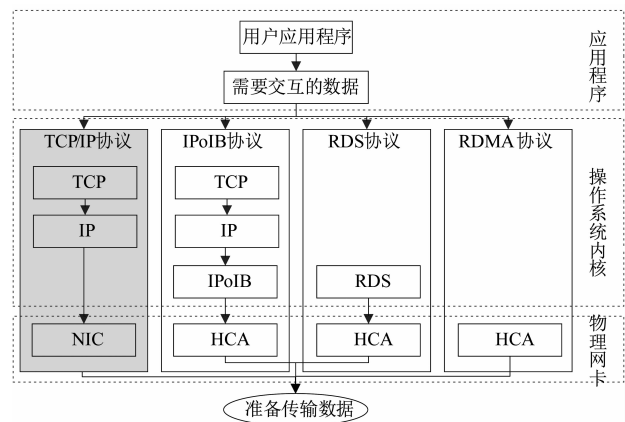


图1 IB网络中的不同协议运行示意图

RTI 也能够在不进行任何修改的情况下运行在具有 IB 网络的高性能计算机上. 但 IPoIB 的核心还是 TCP/IP, 所以通信效率受到其制约.

RDS 协议: Oracle 公司基于 IPC (Inter Process Communication) 的协议, 目的是替代效率低下的 TCP/IP 协议. TCP/IP 协议本身复杂, 需要较大的网络开销, RDS 能够减少 CPU 参与通信的开销, 并能够减少网络延迟.

RDMA 协议:一种绕开 OS (Operating System) 的协议, 在数据准备与数据发送过程中完全不需要操作系统内核参与.

通过以上对比, IPoIB, RDS 等协议需要操作系统参与, 并且要做数据拷贝操作, 而 RDMA 完全不需要操作系统参与, 无需数据拷贝, 能够释放 CPU, 使通信性能最大化. 因此本文选择 RDMA 作为 Infiniband 下 RTI 运行的通信机制.

2.2 Infiniband 的 RDMA 通信机制

2.2.1 RDMA

在 IB 的架构中, 能够支持远程直接内存访问 (RDMA). DMA (Direct memory access) 技术是设备能够直接访问主存储器, 而不经 CPU 的介入, 同理 RDMA (Remote DMA) 是一个设备能够直接访问远程计算机上的内存, 而不会中断该系统上的 CPU 的运行进程.

2.2.2 RDMA 与 TCP/IP

TCP/IP 和 RDMA 的通信机制如图 2 所示.

从 TCP/IP^[11] 的通信机制可以看出, 在整个数据传输过程中, 有以下多次数据的拷贝: ①从应用程序到系统内核的拷贝; ②从系统内核到网卡设备数据缓冲区的拷贝. 此外, 网络层由于网络最大传输单元 (MTU) 的限制对传输层传来的数据包进行分片以及对分片的重新组合引起的拷贝. 数据的拷贝需要 CPU 进行参与, 并且数据拷贝的速度跟 CPU 的核心内存与 CPU 的时钟频率有直接关系. 同时, 因为 IPoIB 协议的根本核心也是运行着 TCP/IP 协议, 因此在 IB 架构下, 其只利用了 IB 架构的带宽, 并没有从通信原理上提高通信的效率.

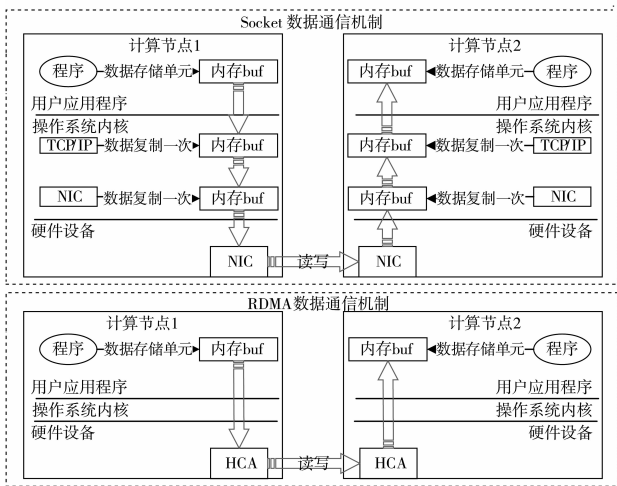


图2 TCP/IP与RDMA的通信机制对比

从 RDMA 通信机制来看,数据能够在无 CPU 和操作系统的参与下,通过 HCA 在两个独立节点之间完成数据传送,并且从数据发送到接收整个过程零拷贝。

综上, RDMA 因为没有进行数据多次复制具有低延迟的特性,同时操作系统与 CPU 不参与数据拷贝,能够使 CPU 完全专注于计算,此外高带宽以及减少内存带宽的瓶颈等这些特点,能够极大的提高网络通信的速率和效率。

将传统以 TCP/IP 通信机制的 RTI 软件升级为 RDMA 通讯机制,以支持复杂系统在 HPC 上求解与协同仿真运行,是提高仿真运行速度、缩短仿真运行耗时的重要途径。为了验证 RDMA 通信机制下, RTI 能够在 HPC 上提升仿真效率,本文选择开源 HLA 项目 CERTI 软件进行优化,并给出验证的结果与分析。

3 基于 RDMA 通信机制的 IB-RTI

通过理论分析可得出基于 RDMA 通信机制的 RTI,能够在 Infiniband 网络架构下极大的发挥高速网络的特性,以提高仿真效率、缩短仿真时间。本章中选择开源 RTI 软件 CERTI 作为优化对象,进行验证。

3.1 CERTI 结构及接口分析

在优化 CERTI 通信底层之前,需要对 CERTI 的运行机制以及各模块的包含关系梳理清楚,然后重构有关网络通信的代码,在保证各模块代码的完整功能的前提下,只优化网络通信部分。在这里着重分析 RTIG 和 RTIA 的内部层次结构和运行机制^[12,13]。

3.1.1 RTIG

相较于 RTIA, RTIG 的代码结构和运行流程较为简单,只负责读取 fed 文件后进行联邦创建,并且监听个联邦成员所发的消息,进行相应的处理和分发。RTIG 的运行流程如图 3 所示。

RTIG 的主要功能调用如图 4 所示, RTIG 在进行网络通信时,调用了 LibCERTI 中的 SocketTCP 和 SocketUDP

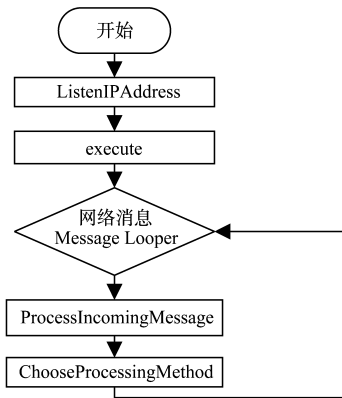


图3 RTIG运行流程图

etUDP 两个功能函数,完成了网络通信的整个操作。本文的研究内容就是在不改变 RTIG 的运行流程下,通过对通信部分代码进行修改和替换来优化网络通信,即修改图 4 中阴影部分的功能模块和功能函数。

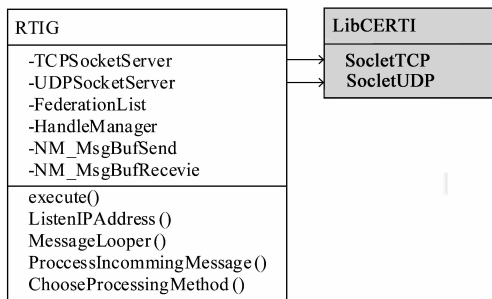


图4 RTIG模块调用LibCERTI包接口

3.1.2 RTIA

RTIA 是联邦运行中最复杂的部分,要完成六大类的服务,并且在时间管理服务的协调下,与其他五类服务进行协作,保证联邦成员的运行,包括对联邦成员属性的更新,所有权转移等操作,将完成后的操作通过网络通信告知 RTIG,同步所有状态,即完成 RTI 中的回调功能。运行流程如图 5 所示。

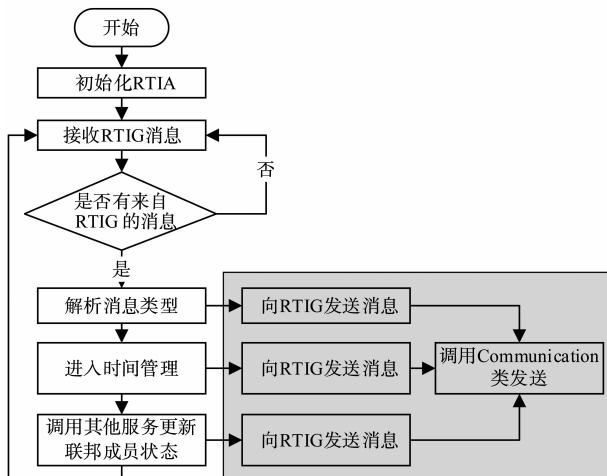


图5 RTIA运行流程图

RTIA 的主要功能接口调用如图 6 所示,上述的流程的调用方式可以清晰的看出,而在图 5 和图 6 中阴影部分,则是需要进行网络优化的部分.

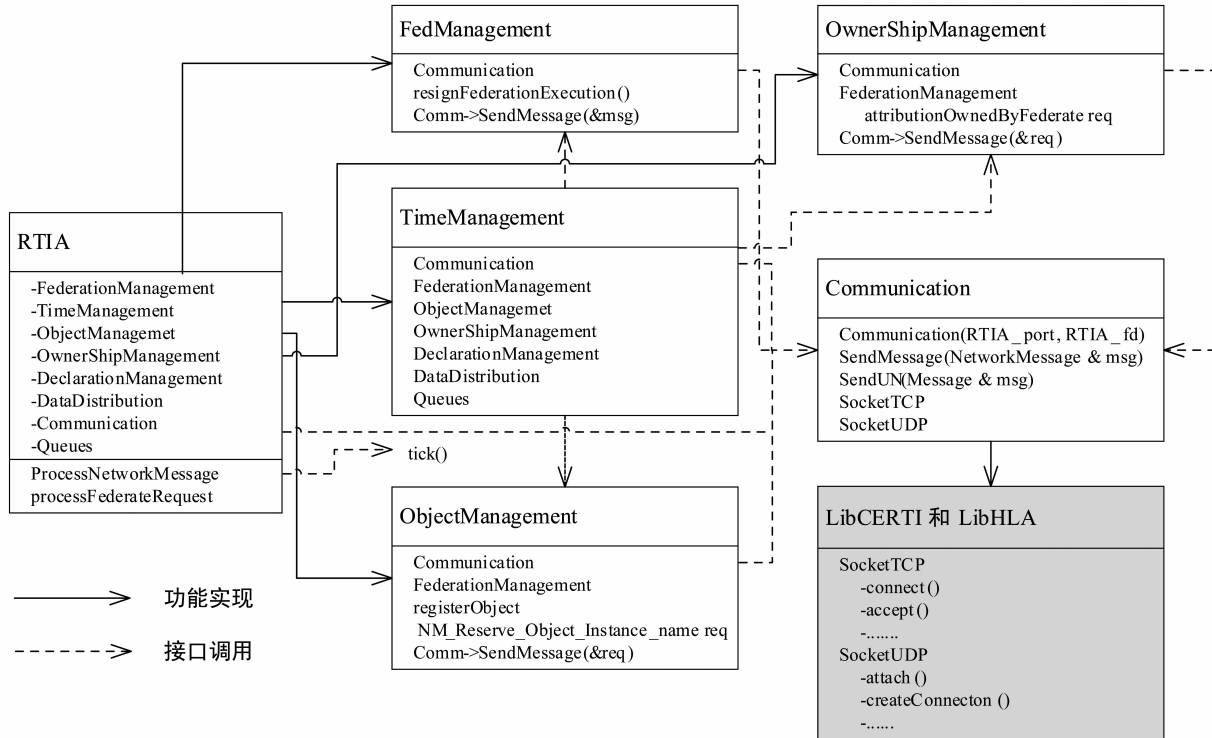


图6 RTIA模块的调用层次接口

3.2 IB-RTI 网络优化

3.2.1 数据结构和消息格式

为了在优化网络通信接口时,不修改其他各程序模块,而且能够保证正常运行,同时确保消息通过 RDMA 协议传输,数据的结构和消息格式定义如下,其中消息内容部分定义了数据发送的地址(主机位置 + 端口)、要发送/接受的数据缓存以及数据的长度;资源设置定义了跟 RDMA 相关的资源标识.

```

struct context
{
    /* 消息内容 */
    char * server_name;
    char * server_port;
    int msg_length;
    char * send_buf;
    char * recv_buf;
    /* 资源设置 */
    struct rdma_cm_id * id;
    struct rdma_cm_id * listen_id;
    struct ibv_mr * send_mr;
    struct ibv_mr * recv_mr;
};
    
```

在仿真运行过程中,所有参与网络通信的数据均封装成上述消息格式,保证在接口调用的过程中不影响其他程序模块的运行.同时,动态的数据大小能够动态的申请 RDMA 交互数据时所需要的内存,能够保证在有多多个邦员在大量数据交互时不会因为临时存储内存的浪费而造成的通信效率的影响.

3.2.2 IB-CERTI 网络主程序

IB-CERTI 网络优化的主要工作是分析 CERTI 的运行流程以及调用关系后,将程序中所涉及的底层基于 TCP/IP 协议的通信全部替换,本文的原型系统主要替换了 LibCERTI 和 LibHLA 中的 SocketTCP 的部分.对照 TCP/IP 建立连接和收发数据的流程,用 RDMA 编码进行相同功能的替换^[14,15],在程序中建立 Server 端和 Client 端的代码,具体流程如图 7 所示,其所实现功能与 TCP/IP 一样.

RDMA 的 Server 和 Client 的建立过程类似,与 TCP/IP 不同之处在于 RDMA 要自行注册操作缓存,用于存放数据,并设置 QPs(消息队列对),作为数据交换的信道,在使用完成后要及时释放缓存,以免在进行大量 RDMA 通信时,大量内存被占用后未被及时释放,造成缓存耗尽而无法继续传输数据.

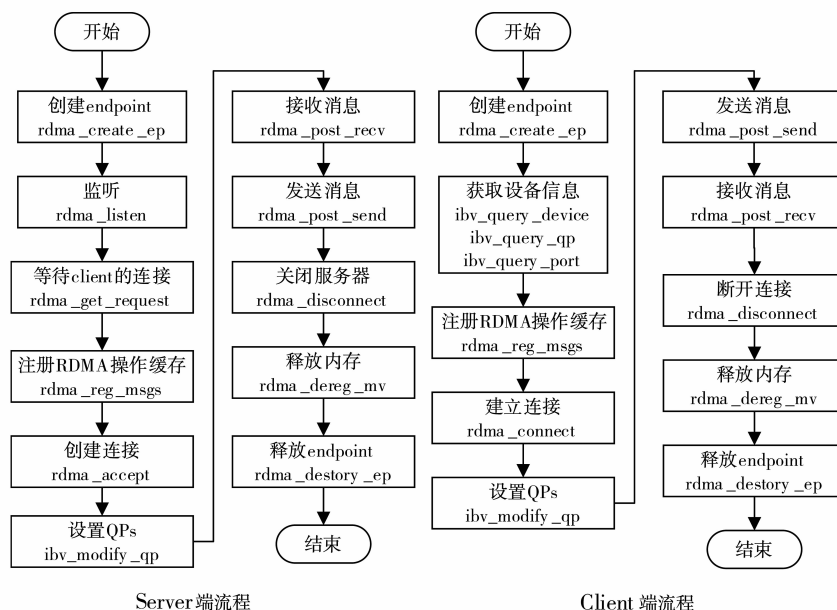


图7 IB-CERTI网络通信部分流程图

4 实验结果与分析

RTI 的测试分为功能测试与性能测试^[6], 功能测试目的是测试 RTI 的六大类服务以及其支持服务的 130 个接口服务; 而性能测试是为了测试 RTI 的性能的优劣, 其中包括: 最大传输率、属性延迟、消息丢失率、时间推进速率、tick() 速率、交互延迟、注册速率、属性所有权转移速率等。而本文主要测试 RTI 软件的网络性能, 故测试内容为最大传输率, 即对象属性更新/反射最大速率。

4.1 实验环境

本软件是在开源 RTI 项目 CERTI 上设计并实现的基于 IB 网络环境下的 RDMA 通信机制的 IB-CERTI 软件。

在 HLA 仿真程序运行中, 模型间的互操作是由对象属性值更新服务来完成的, 在没有时间管理服务参与仿真运行的情况下, 对象属性值更新的速率直接反应 RTI 的通信性能, 因此测试系统中联邦成员、对象之间没有时间管理参与, 只有对象属性更新服务等参与运行。

实验环境由 3 个高性能计算节点和 3 台 PC 机组成, 硬件环境和软件操作系统的配置如表 1 所示。

表 1 实验环境配置

计算机	高性能仿真计算机
硬件环境	高性能计算机计算节点 * 3 台
操作系统	Cent OS 6.2 64bit
IB 网络环境	Mellanox ConnectX-3 Pro

4.2 测试程序

测试系统^[5-7] 为一个“消息传递游戏”, 游戏规则是: 总共有两个 Recorder, 首先由 R1 发送一个消息给 R2, 然后 R2 在收到消息后变更接收到的累计次数, 并给 R1 回复一个消息, 同样 R1 也记录下所接收消息的累计次数, 即成员 R1 首选向 RTI 发出属性更新的请求, RTI 将更新的属性反射给成员 R2, 待 R2 收到反射后则向 RTI 发出属性更新的请求, 同样 RTI 将更新的属性反射给 R1。为了测试结果足够的精确, 将上述的过程循环 1 万次作为一个仿真运行, 其仿真运行的时间即为属性更新的延迟时间, 并且每一个仿真运行进行 12 次, 对每次运行获得的延迟时间做统计均值滤波 (12 次运算结果中, 去掉最大值和最小值, 然后做 10 次均值), 得到的结果误差最小, 最后将得到的延迟时间除以属性更新的次数 (即 $t/2 \times 10000$) 所得结果即为实验结果。此外, 不断改变发送消息数据大小, 从 32 字节到 4096 字节进行变化, 以此来观察在不同的数据包情况下, RTI 网络通信的性能。

4.3 结果与分析

测试结果如表 2 所示。

为了方便对比, 将上述数据的结果以图表分析。如图 8 所示。

(1) 在 TCP/IP 通信机制下, 传输数据小于 512 字节时, 其延迟时间与 IPoIB 与 RDMA 方式下的结果相当, 甚至在 256 字节之内, 其延迟时间要低于 IPoIB 的方式, 可见 TCP/IP 通信机制下, 端与端之间建立连接的速度要高于 IPoIB 方式。但数据大小增加时, TCP/IP 下仿真延迟时间剧增, 在数据大小为 4096 字节时, 延迟

时间约是 IPoIB 下的 4.5 倍, RDMA 下的 5.6 倍.

(2) 由于 IB 硬件结构网络高带宽的特性, 在数据包大小为 4096 之内时, 以 IPoIB 和 RDMA 两种通信所得到的延迟时间相差不大, RDMA 的延迟时间约为 IPoIB 的 80%, 而且随着数据包的增大, 延迟时间增加幅度很小. 特别指出在 IPoIB 下, 当数据包在 2048 时, 数据传输的延迟时间最小, 约为其他数据包耗时的一半. 这是由于 IB 默认设置的 MTU 为 2048, 发送和接收过程不需要进行分片和重组, 避免了数据包分片和重组所需要的时间, 因此在整个测试过程中 2048 的数据测试中网络通信效率最高.

表 2 实验结果

测试环境与软件版本			
计算节点	高效能仿真计算机		
操作系统	Linux		
网络	以太网	IB 网络	
通信方式	TCP/IP	IPoIB	RDMA
RTI 软件版本	CERTI	CERTI	IB - CERTI
仿真交互数据 (单位: 字节)	仿真运行耗时(单位: 微秒)		
32	85.71068	93.54015	81.44376
64	89.37458	94.84318	81.42545
128	88.25904	94.1126	81.66186
256	92.41795	96.23186	81.76574
512	100.5807	96.7671	82.30809
1024	134.9559	98.7378	83.0242
2048	445.2021	50.24509	83.52714
4096	468.6704	102.8956	84.02103

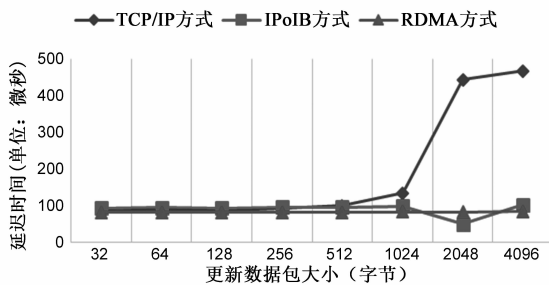


图8 不同通信机制下RTI端到端属性更新延迟

此外, 将三种通信方式下的多次测试结果以图表形式分析, 如图 9, 图 10, 图 11 所示.

在 TCP/IP 方式下, 多次实验结果的延迟时间波动相对来说比较大, 说明在端到端的属性更新时存在着数据交互的网络延迟, 数据包越大时, 网络延迟越频繁; 相反 RDMA 下, 多次测量的延迟时间变化波动不大, 可见 IB 网络低延迟的特性使得仿真运行的效率

更高.

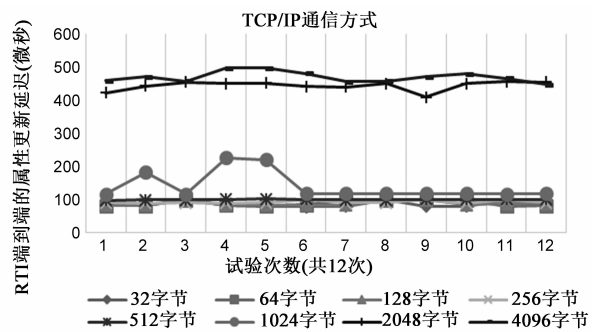


图9 TCP/IP通信机制下改变数据包大小的测试结果统计

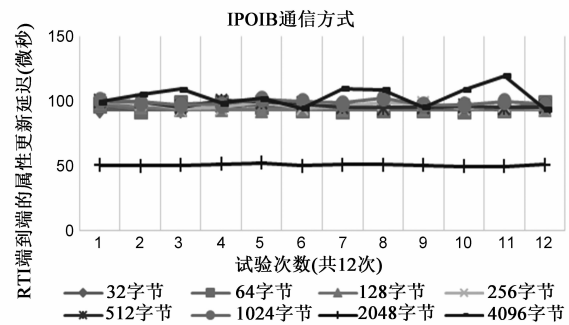


图10 IPoIB通信机制下改变数据包大小的测试结果统计

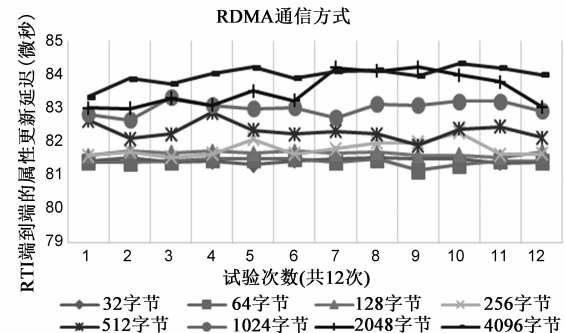


图11 RDMA通信机制下改变数据包大小的测试结果统计

上述实验表明, 在仿真运行过程中, 相比传统以太网下的 TCP/IP 通信机制, IB 网络数据的交互能力更强, 延迟时间更小, 尤其是 RDMA 机制下, 网络数据交互的效率更高, 从而仿真运行的效率更高. 实验结果初步验证了 IB-CERTI 软件实现结构, 及其数据交互机制的高效性.

5 结束语

本文提出了基于 Infiniband 高速通信网络架构下的 RTI 通信的优化方法, 对开源 HLA 项目软件 CERTI 进行分析, 定义了传输数据的数据结构, 并且以 RDMA 的通信方式优化了 CERTI 的通信部分, 完成了 IB-CERTI 软件原型系统的研发. 最后, 在 TCP/IP 协议, IPoIB 协议以及 RDMA 协议下分别进行了实验, 实验结果表

明,基于 RDMA 协议的通信网络能够有效的提高仿真数据传输效率,从而缩短仿真运行的时间。

RDMA 通信协议不但能运行在 Infiniband 网络架构下,同样也能够运行在 Ethernet 网络下,而高性能计算机的仿真资源有限,目前多数的仿真仍以传统的计算机硬件和网络架构为基础,因此将 RDMA 协议应用于传统的 Ethernet 网络,用于提高仿真效率势在必行。

下一步的研究工作:在复杂系统的仿真运行中,大量联邦成员和仿真实体存在时,在频繁的数据交互的情况下,对消息的延迟以及网络带宽承载能力的进一步实验和研究。

参考文献

- [1] Li B H, Chai X, Li T, et al. Research on High-Performance Modeling and Simulation for Complex Systems[M]. Concepts and Methodologies for Modeling and Simulation Springer International Publishing, 2015. 45 - 66.
- [2] Li B H, Li T, Hou B, et al. Research on high-efficiency simulation technology for complex system[A]. Proceedings of the 2011 Grand Challenges on Modeling and Simulation Conference[C]. Society for Modeling & Simulation International, 2011. 285 - 289.
- [3] 曹政, 王达伟, 刘新春, 等. 曙光 5000 高性能计算机多播网络的设计[J]. 电子学报, 2011, 39(2): 481 - 488.
CAO Zheng, WANG Da-wei, LIU Xin-chun, et al. Design of multicast network of dawnning 5000 high performance computer[J]. Acta Electronica Sinica, 2011, 39(2): 481 - 488. (in Chinese)
- [4] Fujimoto R, McLean T, Perumalla K, et al. Design of high performance RTI software[A]. Distributed Simulation and Real-Time Applications, 2000[C]. 2000. 89 - 96.
- [5] Adelantado M. HP-CERTI: Towards a High Performance, High Availability Open Source RTI for Composable Simulations[M]. Orlando, USA; IEEE, 2004. 1 - 9.
- [6] 梁洪波, 柳林, 姚益平, 等. 高性能 RTI 自适应通信机制研究[J]. 国防科技大学学报, 2012, 34(3): 148 - 153.
LIANG Hongbo, LIU Lin, YAO Yiping et al. Research on self-adaptive communication mechanism for high performance RTI[J]. Journal of National University of Defense Technology, 2012, 34(3): 148 - 153. (in Chinese)
- [7] 张智慧, 李伯虎, 柴旭东, 等. 一种面向共享内存环境的高性能 HLA/RTI 原型系统[J]. 系统仿真学报, 2014, 26(2): 315 - 322.
ZHANG Zhi-hui, LI Bo-hu, CHAI Xu-dong, et al. HP-HLA/RTI prototype oriented on shared memory environment[J]. Journal of System Simulation, 2014, 26(2): 315 - 322. (in Chinese)
- [8] InfinBand™ Trade Associate. InfinBand™ architecture, specification volumn1 release 1.1[S/OL]. <http://www.infinibandta.org/specs>, 2006.
- [9] Pfister G F. An introduction to the infiniband architecture[J]. High Performance Mass Storage and Parallel I/O, 2001, 42: 617 - 632.
- [10] 陈凯, 白英彩. 网络存储技术及发展趋势[J]. 电子学报, 2002; 30(S1), 1928 - 1932.
CHEN Kai, BAI Yingcai. Technology and trend of network storage[J]. Acta Electronica Sinica, 2002; 30(S1), 1928 - 1932. (in Chinese)
- [11] 樊秀梅, 单志广, 张宝贤, 陈辉. 容迟网络体系结构及其关键技术研究[J]. 电子学报, 2008, 36(1): 161 - 170.
FAN Xiu-meil, SHAN Zhi-guang, ZHANG Bao-xian, CHEN Hui. State-of-the art of the architecture and techniques for delay-tolerant networks[J]. Acta Electronica Sinica, 2008, 36(1): 161 - 170. (in Chinese)
- [12] Noulard E, Rousselot J Y. CERTI, An open source RTI, why and how[R]. Spring Simulation Interoperability Workshop, 2009.
- [13] Bréholée B. CERTI: Evolutions of the ONERA RTI Prototype[R]. Proceedings of the Fall Simulation Interoperability Workshop, Orlando, 2002.
- [14] Mellanox Technologies, Inc. . RDMA Aware Programming User Manual Rev1. 2[EB/OL]. <http://www.mellanox.com>, 2010-01.
- [15] Mellanox Technologies, Inc. . ConnectX Family Programmer's Reference Manual[EB/OL]. <http://www.mellanox.com>, 2003-06-20.

作者简介



邢 驰 男, 1984 年生于陕西省杨陵区. 北京航空航天大学自动化科学与电气工程学院博士研究生. 研究方向为云仿真、云制造、高效能仿真.
E-mail: xingchi09@163.com



李伯虎 男, 1938 年生于上海市. 中国工程院院士, 博导. 研究方向为云制造、网络化建模仿真等.
E-mail: bohuli100@aliyun.com