

文章编号: 2095-4980(2017)02-0268-05

基于 InfiniBand 的并行绘制系统中的数据压缩

彭世雄, 季玉龙, 吴志红

(四川大学 计算机学院, 四川 成都 610065)

摘 要: 为提高并行绘制系统中的帧率, 从数据压缩的角度出发, 搭建了基于 InfiniBand 的并行绘制系统。首先对并行绘制系统中数据进行分类, 根据当前帧的数据量大小设置临界值, 对达到临界值的几何数据进行无损压缩。本文研究并对比了 LZ77, LZW, LZO 三种数据压缩算法对系统的影响。实验表明, LZ77, LZW 和 LZO 算法均对系统的帧率有所提高, 而 LZO 算法对系统的帧率提高最明显且稳定性最好, 使系统帧率提高近 60%。无损数据压缩对提高并行绘制系统性能具有明显的效果。

关键词: 并行绘制系统; 数据压缩; InfiniBand 架构; LZO 算法; LZW 算法; LZ77 算法

中图分类号: TN957.2

文献标志码: A

doi: 10.11805/TKYDA201702.0268

Data compression in parallel rendering system based on InfiniBand

PENG Shixiong, JI Yulong, WU Zhihong

(College of Computer, Sichuan University, Chengdu Sichuan 610065, China)

Abstract: From the perspective of data compression, parallel rendering system based on InfiniBand is set up in order to improve the frame rate. Data classification in parallel rendering system is performed; and the thresholds are set according to the data of the current frame; the lossless compression is applied to the geometric data which reach up to the critical value. The effects of three types of data compression algorithms, LZ77, Lempel-Ziv-Welch(LZW), Lempel-Ziv-Oberhumer(LZO), are researched and compared. The experiments results show that, LZ77, LZW and LZO algorithms can improve the frame rate of system; LZO algorithm increases the frame rate of system most significantly with the best stability, and the system frame rate is increased by nearly 60%. Lossless data compression can improve the performance of parallel rendering system significantly.

Keywords: parallel rendering system; data compression; InfiniBand; Lempel-Ziv-Oberhumer algorithm; Lempel-Ziv-Welch algorithm; LZ77 algorithm

随着图形处理软硬件技术的发展, 实际应用中要求绘制系统能够处理更加精细的几何场景, 提供高分辨率、高真实感的绘制效果, 但单一硬件往往因计算资源有限, 绘制速度不能满足实际需求。所以将并行计算引入图形处理领域形成并行绘制技术, 从而提供高性能图形绘制服务。分布式并行绘制技术通过构建高性能的分布式图形处理系统, 为以上问题提供了解决办法。

在并行绘制系统中, 网络带宽容易成为瓶颈, 影响系统的性能和可扩展性^[1]。如何应用压缩技术减少数据的传输量, 缓解网络带宽的瓶颈, 是并行绘制系统研究的一个重要问题。虽然可以通过增加带宽, 使用高性能的网络设备来缓解性能瓶颈, 比如使用 InfiniBand 网络架构, 但是随着对高分辨率、高真实感的绘制效果的追求, 以及高性能网络设备的昂贵价格, 体系结构的设计和软件层面的数据优化显得尤其重要。

目前提升并行绘制系统性能的关键点主要包括数据压缩传输、负载均衡以及图形合成的优化。文献[2]针对 Chromium 绘制系统中 3-D 模型数据进行压缩, 但是只是针对特定的系统中的数据进行压缩, 有很大的局限性。文献[3]对几何指令流进行压缩, 但随着时代发展, 数据格式变换较大, 该方法很难有普遍的适用性。本文将绘制系统中的数据分类, 考虑到控制信息数据量小, 压缩算法带来开销反而会得不偿失, 因此对控制信息数据不进行压缩, 只对数据量大的几何数据进行压缩, 文中使用 3 种不同特点的无损压缩算法对系统中几何数据进行压缩, 分析了它们对基于 InfiniBand 的并行绘制系统帧率的影响。

1 相关工作

1.1 基于 InfiniBand 的并行绘制系统

InfiniBand 是一种开放的处理器及 IO 互连标准^[4],它具有高带宽、低延时的特点,目前已经成为高性能计算(High Performance Computing, HPC)互连的主流网络之一。InfiniBand 架构是一种支持多并发链接的“转换线缆”技术,在这种技术中,每种链接都可以达到 2.5 Gbps 的运行速度。并行绘制系统在运行过程中需要实时传输大量数据,对底层网络提出了较高的要求,InfiniBand 具有高带宽、低延迟等特点,正好满足这一需求。

本文设计的并行绘制系统使用 InfiniBand 作为底层网络,系统采用 sort-first 架构和冗余数据存储模式,构建于基于 GPU 的 3-D 实时应用开发平台 Hive 上,能够支持复杂场景的实时 3-D 渲染。将系统分为 3 类功能节点:服务器节点(Server Node)、绘制节点(Render Node)、显示节点(Display Node),3 类节点之间通过 InfiniBand 网络互连。图 1 为系统基本架构图。

集群并行绘制系统由一个服务器节点、若干绘制节点与若干显示节点组成,其中服务器节点作为系统的中心控制节点,不承担具体的显示、绘制任务。绘制节点接收来自服务器的场景绘制任务,完成绘制后,收集相关的绘制数据(如区域的场景图像及其相对位置、绘制开销、视口信息、帧号等)将其发回服务器。显示节点在服务器的控制下定期请求所承担显示区域的场景图像,服务器在收齐所有显示节点的数据请求后,一次性派发当前帧图像数据,显示节点在服务器发出的同步信号控制下显示图形。

1.2 常用压缩算法

本质上压缩数据是因为数据自身具有冗余性。数据压缩是利用各种算法将数据冗余减少到最少,并尽可能地减少失真,从而提高传输效率和节约存储空间。数据压缩一般分为有损压缩和无损压缩^[5],无损压缩是指重构压缩数据(还原,解压缩),而重构数据与原来数据完全一致。衡量一个压缩算法的优劣一般有压缩比、压缩速度、解压速度^[6]。

1.2.1 LZ77 算法

LZ77 算法^[7]是 1977 年 2 位以色列科学家 Jacob Ziv 和 Abraham Lempel 提出的一种基于滑动窗口缓存的技术,这个算法一般称为 LZ77,利用该算法进行数据压缩、解压的过程,就像一个窗口在原始数据中滑动过去,故也称为基于滑动窗口的自适应字典压缩方法。

解压缩的过程非常简单,只要像压缩的时候那样维护好窗口的指针,随着三元组的不断输入,在窗口中找到相应的匹配串,加上后继字符 c 输出(如果 off 和 len 都为 0,则只输出后继字符 c),即可还原输出。

LZ77 算法是基于字符串匹配的第一个有实用价值的算法,它的压缩和解压缩速度非常快,由于它限制匹配的最大长度,对于某些存在大量的极长匹配的文件来说,这种折衷算法有一定缺陷。

1.2.2 LZW 算法

LZW 算法^[8]是 LZ78 的变形,由 Lempel-Ziv-Welch 三人共同创造,用他们的名字命名,简称 LZW 算法。它采用一种先进的字典压缩,通过建立一个字符串表,用较短的代码来表示较长的字符串,实现压缩,将每个第一次出现的串放在一个串标中,用一个数字来表示串,压缩文件只存储数字,从而使图像文件的压缩效率得到较大的提高。

LZW 算法的特点非常明显,对可预见性不强的数据,具有很好的处理效果,常用于 gif 格式的图像压缩,其平均压缩比在 2:1^[9]以上,是一种很简洁但效率高的算法。

1.2.3 LZO 算法

LZO 算法是致力于解压速度的一种数据压缩算法,该算法为无损算法,实现它的是一个自由软件工具 lzop。LZO 算法的基本思想脱胎于 LZ77 算法,但是其使用了哈希表来记录滑动窗口内的字典数据,并十分精细化了 LZ77 的压缩,使用哈希表可以实现只扫描一遍数据就可以完成压缩,由于哈希表的时间复杂度为 $O(1)$,使得搜索相同的字符时间非常短。

LZO 算法的优点非常明显,由于使用哈希表,压缩和解压缩速度非常快,而且解压不需要额外的内存;LZO

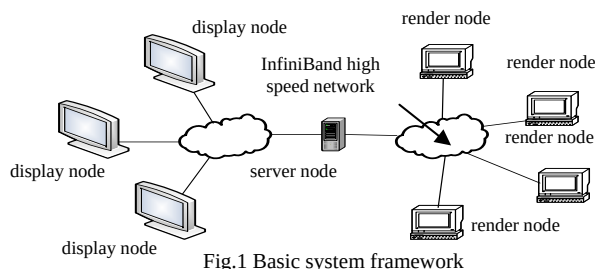


Fig.1 Basic system framework

图 1 系统基本框架

压缩是基于 Block 分块的, 这样一个大的文件可以由多个进程同时压缩, 进一步提高压缩和解压速度; 另外 LZO 算法具有很高的压缩比, 实现程序是线程安全的, 常被用于网络传输。

2 使用压缩算法加速

系统运行时, 显示节点和绘制节点间没有耦合, 互不干扰, 但是显示与绘制节点间的通信需服务器节点转发, 使得服务器节点承担的网络压力较大, 以“ $2\,560 \times 1\,440, 60\text{ fps}$ ”显示要求为例, 每秒的数据量 864 MB, 这还不包括控制信息和其他开销(相对位置、绘制开销、视口信息、帧号等)。虽然节点间采用高性能网络 InfiniBand

互连, 但是随着对更高分辨、更高实时性的要求, 底层网络需要传输的数据量将会越来越大。本文设计的并行绘制系统运行过程中各类节点间会交换大量数据, 系统数据交换如图 2 所示。

在基于 InfiniBand 的并行绘制系统中, 系统的数据主要由图像几何数据、系统的控制数据和负载均衡数据量组成, 其中系统的控制数据、负载均衡数据量并不大, 对系统的性能影响可以忽略不计, 系统中的数据主要是几何数据, 所以在该系统中只对几何数据进行压缩。图 3 为系统的时序图, 只有在 Render 节点向 Server 节点发送图形数据, Server 节点向 Display 节点发送图形数据才使用压缩算法。

场景图像是集群绘制系统中的一种临界资源, 在 Render 节点中设置工作中心的活动(Work Center Event, WCE)压缩线程, 由于系统的实时性, 绘制数据量会发生巨大变化, 而且多个绘制节点绘制任务不一, 压缩算法本身的压缩解压缩开销也同样需要考虑, 如果对所有大小的几何数据都进行压缩的话, 压缩和解压缩的损耗可能会拖累整个系统。

在 Server 节点和 Render 节点中设置一个临界值, 临界值取当前帧数据量的 1/5, 当数据量超过临界值, 将其中的几何数据进行压缩。首先 Render 节点通过 InfiniBand 网络将数据发送至 Server 节点; Server 节点中需要再将各个绘制节点的数据解压, 根据各 Display 节点区域重新划分, 判断每一幅子图像是否超过临界值, 并将超过临界值的子图像压缩后发给相应的 Display 节点; 所有 Display 节点收到子图像数据后, 如果数据经过压缩, 解压后和其他非压缩数据在 Server 节点的控制下同步显示。本系统中图像压缩流程如下:

- 1) Server 节点启动, 向 Render 节点发送任务命令, Render 节点接收到命令, 执行绘制任务, 并开启 WCE 压缩线程;
- 2) Render 节点中 WCE 线程对从 Render 节点将发送到 Server 节点中的数据进行判断, 如果是图像几何数据, 且数据量大于临界值, 则使用压缩算法, 否则不使用;
- 3) Server 节点接收到从 Render 节点发送来的数据, 启动 WCE 启动线程, 判断数据是否压缩, 解压图像数据, 并对所有图像数据合成, 根据 Display 节点数量划分图像, 对大于临界值的图像数据再压缩, 通过 InfiniBand 网络发送数据;
- 4) Display 节点接收到对应的图像数据, 解压缩, 由 Server 节点控制同步显示图像。

3 实验结果

针对前文提出的基于 InfiniBand 的并行绘制系统, 搭建了一个拥有 3 个 Render 节点、4 个 Display 节点和 1 个 Server 节点的原型系统, 对 3 种压缩算法进行了测试。所有实验节点均采用 Mellanox 的 InfiniBand 和以太网连接, 交换机为 IS5023 的 16 端口、无阻塞速率为 20 Gbps InfiniBand 光纤交换机。系统软硬件详细配置如表 1, Display 节点采用 4 台分辨率为 $1\,920 \times 1\,080$ 的 PHILIPS 显示器。

不同的算法会有不同的数据压缩率、压缩速度和解压速度。为了测试不同算法对需要传输大量数据的并行系

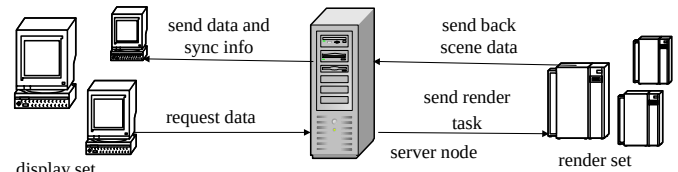


Fig.2 Data exchange in the system
图 2 系统中的数据交换

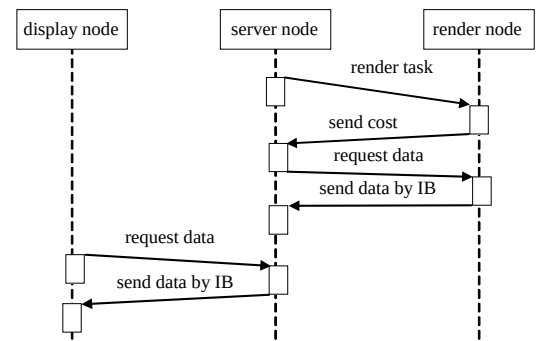


Fig.3 System sequence diagram
图 3 系统时序图

统的影响,测试中使用动态模型。另外在测试中设定取当前帧数据量的 1/5 作为临界值,超过该值,执行压缩算法,表 2 展示了不同算法对同一动态模型的测试结果。

表 1 系统的软硬件配置
Table1 Hardware and software configurations of system

node type	amount	configuration
server node	1	Win7 64-bit system, Intel(R) Xeon(R)3.07 GHz, 4.00 GB RAM,NVIDIA Quadro FX 580 0
render node	3	Win7 64-bit system, Intel(R) Xeon(R)3.07 GHz, 4.00 GB RAM,NVIDIA GTX 560
display node	4	Win7 64-bit system, Intel(R) Xeon(R)3.07 GHz, 4.00 GB RAM,NVIDIA GTX 560,1920*108 0

从表 2 中可以看出,经过压缩算法的处理后,帧率都有提升。其中 LZW 算法的压缩率最高,达到 2.54, LZO 算法的帧率提高最明显,提高 62%,这是由于 LZO 算法对几何数据具有非常快的压缩和解压缩速度,这也看出来并行绘制系统中并不能一味追求压缩率,应该综合考虑压缩率、压缩速度和解压速度。综合来看,经过 LZO 算法处理的系统效果令人满意。图 4 展示了使用压缩算法前后的帧率动态变化。

表 2 不同算法对动态模型的测试结果
Table2 Dynamic model test results for different algorithms

algorithm name	average frame rate before compression/fps	average frame rate after compression/fps	compression ratio	average compression time/ms	average decompression time/ms
LZ77	19.4	23.8	1.71	90	60
LZW	19.4	27.0	2.54	65	40
LZO	19.4	31.5	2.01	50	20

图 4 可以看出不使用压缩算法的并行绘制系统帧率较低,但是波动很小;使用 LZ77 算法和 LZW 算法的系统帧率虽然比不使用压缩算法有所提高,但是波动较大,稳定性一般;使用 LZO 算法的系统帧率提高很明显,且波动较小,在并行绘制系统中是首选算法。

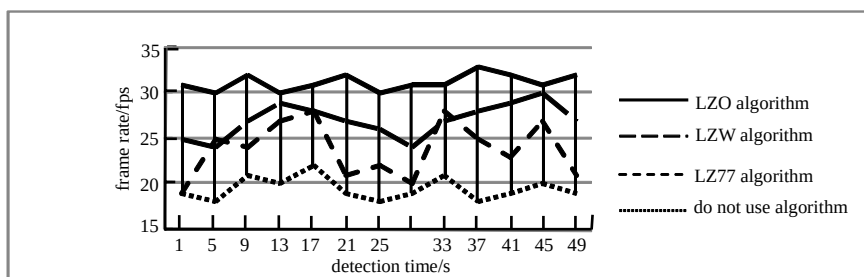


Fig.4 Dynamic changes of frame rate comparison

图 4 帧率动态变化对比

4 结论

在并行绘制系统中,网络传输一直是限制并行绘制系统性能的难题之一,本文成功地将数据压缩算法应用于并行绘制系统中,并根据数据类型和数据量大小使用压缩算法,有效缓解网络压力,使得系统的帧率大幅提高,为减少并行绘制系统的成本提供了另一种可能,具有很高的实用价值。从实验结果来看,使用压缩算法使得系统的稳定性稍微降低,帧率波动变大,这可能是压缩算法对不同帧的压缩效果变化较大,因此下一步还需要进一步提高系统的稳定性。

参考文献:

- [1] 彭浩宇. 基于 PC 集群机的并行图形绘制系统研究[D]. 杭州:浙江大学, 2006. (PENG Haoyu. Research of parallel graphics rendering system based on PC-cluster[D]. Hangzhou, China: Zhejiang University, 2006.)
- [2] 王欢,秦开怀. 并行绘制系统 Chromium 中的 3D 模型数据压缩[J]. 计算机应用, 2011,31(1):25-28. (WANG Huan, QIN Kaihui. 3D model data compression in parallel rendering system chromium[J]. Journal of Computer Applications, 2011,31(1):25-28.)
- [3] 金哲凡,杨建,石教英. 分布式并行绘制系统中几何指令流压缩的研究与实现[J]. 计算机辅助设计与图形学学报, 2002,14(9):824-828. (JIN Zhefan, YANG Jian, SHI Jiaoying. Research and implementation of geometry instruction stream compression in distributed parallel rendering system[J]. Journal of Computer-Aided Design and Computer Graphics, 2002, 14(9): 824-828.)

- [4] ALFARO F,SANCHEZ J,MENDUINA M,et al.A formal model to manage the infini-band arbitration tables providing QoS[J]. IEEE Transactions on Computers, 2007,56(8):1024–1039.
- [5] 吕高焕,吕文涛. 基于训练字典优化测量矩阵的SAR图像压缩算法[J]. 太赫兹科学与电子信息学报, 2015,13(6):930–936. (LYU Gaohuan,LYU Wentao. A SAR image compression algorithm based on optimization measurement matrix[J]. Journal of Terahertz Science and Electronic Information Technology, 2015,13(6):930–936.)
- [6] DEROOWICZ S,GRABOWSKI S. Data compression for sequencing data[J]. Algorithms Mol Biol, 2013,18(8):25–30.
- [7] 郑翠芳. 几种常用无损数据压缩算法研究[J]. 计算机技术与发展, 2011,21(9):73–76. (ZHENF Cuifang. Research of several common lossless data compression algorithms[J]. Computer Technology and Development, 2011,21(9):73–76.)
- [8] CHEN Qinghui,CHEN Xiaosong,HAN Deliang. Compression algorithm LZW on Chinese text[J]. Computer Engineering and Application, 2014,50(3):112–117.
- [9] CHIANG S W. Adaptive lossy LZW algorithm for palettised image compression[J]. Electronics Letters, 1997,33(10):852–854.

作者简介:



彭世雄(1991–), 男, 湖南省娄底市人, 在读硕士研究生, 主要研究方向为计算机图形、并行计算.email:sx123peng@163.com.

季玉龙(1978–), 男, 河北省吴桥县人, 博士, 工程师, 主要研究方向为图形图像.

吴志红(1964–), 女, 成都市人, 副教授, 硕士生导师, 主要研究方向为数字图像处理、模式识别等.

(上接第 267 页)

- [10] RACHIDI F. A review of field-to-transmission line coupling models with special emphasis to lightning-induced voltages on overhead lines[J]. IEEE Transactions on Electromagnetic Compatibility, 2012,54(4):899–911.
- [11] 林金永,张春侠. 高空核爆环境下飞行控制系统的防护方法[J]. 太赫兹科学与电子信息学报, 2012,10(6):695–701. (LIN Jinyong,ZHANG Chunxia. Protective method of flying control system under high altitude nuclear detonation[J]. Journal of Terahertz Science and Electronic Information Technology, 2012,10(6):695–701.)
- [12] 周壁华. 电磁脉冲及其工程防护[M]. 北京:国防工业出版社, 2003:410–411. (ZHOU Bihua. EMP and EMP Protection[M]. Beijing:National Defense Industry Press, 2003:410–411.)

作者简介:



赵建鹏(1991–), 男, 陕西省宝鸡市人, 在读硕士研究生, 主要研究方向为高功率微波技术、光电对抗.email:zjpzhao@163.com.

侯德亭(1963–), 男, 河南省洛阳市人, 硕士, 教授, 主要研究方向为微波技术、光电对抗技术.

胡涛(1976–), 男, 安徽省铜陵市人, 博士, 副教授, 主要研究方向为高功率微波.

祝民鹏(1992–), 男, 石家庄市人, 在读硕士研究生, 主要研究方向为高功率微波技术、光电对抗.