

# 异构机群上高效可扩展的 Motif 发现并行算法

李锦钟 诚

(广西大学计算机与电子信息学院 南宁 530004)

**摘要** 在节点具有不同计算速度、不同通信能力的异构机群系统上,分别建立求解  $l \leq 16$  和  $l > 16$  的 Motif 发现问题的最优序列分配模型,在此基础上设计实现融合投票和统一投影-邻居阈值思想的 Motif 发现并行算法。实验结果表明,给出的基于最优序列分配策略的 Motif 发现并行算法具有良好的加速和可扩展性,优于采用平均分配策略的 Motif 发现并行算法。

**关键词** DNA 序列, Motif 发现, 并行算法, 异构机群, 数据分配模型

**中图分类号** TP338.6, TP301.6 **文献标识码** A

## Efficient and Scalable Parallel Algorithm for Motif Finding on Heterogeneous Cluster Systems

LI Jin ZHONG Cheng

(School of Computer and Electronics and Information, Guangxi University, Nanning 530004, China)

**Abstract** The optimal sequence distribution models for solving Motif finding with length  $l \leq 16$  and  $l > 16$  were constructed respectively and a parallel algorithm to find Motif combining voting algorithm with uniform projection and neighbourhood thresholding algorithm was implemented on the heterogeneous cluster that the processor nodes have different computing speed and distinct communication capability. Experimental results show that the parallel Motif finding algorithm using optimal sequence distribution strategy has good speedup and scalability, and it is superior to the parallel algorithm using even distribution strategy.

**Keywords** DNA sequence, Motif finding, Parallel algorithm, Heterogeneous clusters, Data distribution model

## 1 引言

Motif 是指 DNA 序列中长度较短、具有保守功能的序列片段<sup>[1]</sup>。从 DNA 序列查找 Motif 的过程称为 Motif 发现。Motif 发现有助于了解序列的功能和阐明序列之间的进化关系,并在识别 DNA 序列的调控信号方面有着重要的应用。

Motif 发现的一个重要问题是植入  $(l, d)$  Motif 发现<sup>[1]</sup>: 给定长度分别为  $n$  的样本 DNA 序列集合  $S = \{s_1, s_2, \dots, s_t\}$ , 序列中元素都来自字符集合  $\Sigma = \{A, C, G, T\}$ , 已知每个序列  $s_i$  都可能包含有一个或多个长度为  $l$  的模体  $M$  的变体  $M_i$  (称之为模体实例),  $M_i$  由  $M$  进化而来且  $M_i$  与  $M$  的海明距离不大于  $d, 1 \leq i \leq t$ , 要求从  $S$  中找出模体  $M$ 。

自从 Motif 的概念被提出,为了从遗传密码中得到规律,人们提出了一些 Motif 发现串行算法。文献[2]的研究结果表明,当  $l \leq 15$  时,串行投票算法都能在合理时间内找到最优 Motif,但是当  $l > 15$  时,由于其所耗费时间较长和内存空间较多,在实际应用中将变得不可用。文献[2]引入启发式随机投影策略设计 Motif 发现算法,该算法能快速查找出结果,但不一定是最优的结果,成功率下降。文献[3]将统一投影和基于邻居桶聚集遴选提炼的策略结合起来,形成 Uniform Projection and Neighbourhood Thresholding (UPNT) 算法,进一步提升了随机投影算法的性能,具有更强的适应性,但其成功

率比不上诸如投票算法这样的精确性查找方法。

文献[4]提出了在机群上实现的适合在科学数据中挖掘 Motif 的并行算法。文献[5]设计了针对蛋白质序列的隐马尔可夫模型搜索的硬件加速器,使用流水线策略,在 FPGA 平台上实现了 Motif 发现算法。文献[6]建立一个抽象模型 PAMM,在此模型上开发基于 MEME 算法的执行工具 MA-MA,并在采用 MISD 并行的硬件芯片 PMC(模式匹配芯片)上实现该工具。文献[7]给出了在 CellMatrix 结构上设计的生物序列 Motif 发现算法。文献[8]采用均匀分配后缀群策略、并行淘汰和归并方法,在同构机群系统上设计实现了一种 Motif 发现并行投票算法。

为了能在有限的时间和条件内,使解决 Motif 发现问题的成功率尽可能高,而运行时间更少、效率更高,本文融合投票和 UPNT 算法的思想,研究设计在节点计算能力和通信能力不同的异构机群系统<sup>[9,10]</sup>上实现高效可扩展的 Motif 发现并行算法。

## 2 异构机群上序列分配与 Motif 并行发现

### 2.1 序列分配模型

设异构机群系统的处理机节点集合为  $\{P_1, P_2, \dots, P_p\}$ ,  $p$  为处理机个数,由主节点  $P_1$  存储长度分别为  $n$  的  $t$  个序列,并负责分配序列给  $p$  个从节点。  $P_1$  和每个从节点  $P_i$  之间有

到稿日期:2011-04-11 返修日期:2011-07-22 本文受国家自然科学基金(60563003)和广西大学拔尖创新团队建设计划(211-3-1-45)资助。

李锦(1978-),男,硕士,主要研究方向为并行计算, E-mail: ljj\_2002@163.com; 钟诚(1964-),男,博士,教授,博士生导师,主要研究方向为并行分布计算。

一条通信链路  $L_i$ , 设在链路  $L_i$  上启动一次通信所需时间为  $G_i$ , 传送一个字符所需时间为  $g_i$ , 则在链路  $L_i$  上传送  $x$  个字符所需时间为  $G_i + xg_i, i=2 \sim p$ 。

融合投票和 UPNT 算法的 Motif 发现算法 V-UPNT 的思想是: 当  $l \leq 16$  时, 使用精确查找的投票方法来发现 Motif; 当  $l > 16$  时, 使用 UPNT 算法来发现 Motif。下面讨论如何建立数据序列分配模型。

### (1) Motif 长度 $l \leq 16$ 的情形

主节点  $P_1$  首先广播  $t$  个序列给各个从节点, 所用时间为  $G_i + tng_i$ ; 然后依据后缀长度  $l'$  将  $4l'$  个长度为  $l'$  的序列划分成  $4l'$  个群, 并将这  $4l'$  个群分配给  $p$  个节点, 每个节点所分配得到的群个数分别为  $q_1, q_2, \dots, q_p, P_1$  也参与计算。各个节点在本地调用投票算法查找出 Motif, 最后主节点  $P_1$  从各个从节点收集最终结果。

设节点处理一个群所需要的时间为  $u_i$ , 并行求解 Motif 发现问题的完成时间为  $T_{\max}$ 。在节点具有不同的计算速度、通信能力的异构机群系统上, 为了提高处理机的利用率, 减少从节点的等待时间, 将并行求解  $l \leq 16$  的 Motif 发现问题抽象为如下线性规划模型:

$$\begin{cases} \sum_{i=1}^p q_i = 4^{l'} \\ G_i + tng_i + q_i u_i \leq T_{\max}, i=1 \sim p \\ q_i \geq 0 \end{cases} \quad (1)$$

式(1)的线性规划模型可以利用 Lingo、Matlab、Mathematica 等软件求解出分配给各个处理机节点的群个数  $q_i$  和并行求解 Motif 发现问题的完成时间  $T_{\max}$ 。

### (2) Motif 长度 $l > 16$ 的情形

主节点  $P_1$  首先对  $t$  个序列进行预处理: 构建投影序列  $T_1, T_2, \dots, T_M$ ; 计算哈希表  $H$  中每个桶的  $S_i$  值, 并统计出  $S_i$  值超过门限的桶; 然后将所有超过门限的桶划分为  $p$  个子群, 每个子群相对应的桶个数分别记为  $a_1, a_2, \dots, a_p$ 。由  $P_1$  分配子群桶给  $p$  个节点, 分配完成后  $P_1$  参与计算。

设  $P_1$  将每个桶分配给节点  $P_i$  在链路上所需的传送时间为  $C_i$ , 对每个桶进行 EM 提炼所需时间为  $w_i$ 。由于从节点返回的是最终 Motif, 通信量很小, 因此可以忽略返回结果所花费的时间。设需要进行 EM 提炼的桶的总规模为  $N$ , 并行求解 Motif 发现的完成时间为  $T_{\max}$ 。主节点  $P_1$  向从节点  $P_i$  传送子群桶数据所需时间为  $G_i + a_i C_i$ , 节点  $P_i$  执行 EM 提炼所需时间为  $a_i w_i, i=1 \sim p$ 。

分配序列顺序按照  $P_1, P_2, \dots, P_p$  的次序进行, 异构机群系统上序列分配和各个节点执行 EM 提炼的过程如图 1 所示。

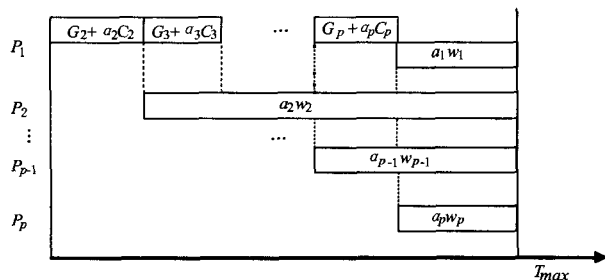


图 1 序列分配和节点提炼 EM 过程

设  $T_i$  为  $P_i$  完成 Motif 查找的时间,  $T_{\max}$  为所有处理机完

成 Motif 查找的时间, 则有  $T_{\max} = \max\{T_i, i=1 \sim p\}$ 。为了提高处理机节点提炼 EM 的利用率, 减少从节点等待时间, 将并行求解  $l > 16$  的 Motif 发现问题抽象为如下线性规划问题:

$$\begin{cases} \sum_{i=1}^p a_i = N \\ \sum_{j=1}^i (G_j + a_j C_j) + a_i w_i \leq T_{\max}, i=1 \sim p \\ a_i \geq 0 \end{cases} \quad (2)$$

## 2.2 算法描述与分析

### 算法 1 异构机群系统上 Motif 发现并行算法

输入:  $t$  条植入  $(l, d)$  样本序列的集合, Motif 长度  $l$ , 最大汉明距离  $d$ , 处理机数  $p$

输出: Motif  $M$

Begin

1. if  $l \leq 16$  then

2. 处理机  $P_i$  播送 {样本序列,  $l, d$ } 给  $P_i, 1 \leq i \leq p$ ;

3. 将  $4^{l'}$  个群划分成  $p$  个长度为  $q_1, q_2, \dots, q_p$  的子群;

4. for all  $P_i$  where  $1 \leq i \leq p$  par-do

5. for  $k=1$  to  $q_i$  do

6. 执行串行投票算法<sup>[2]</sup>;

7. 合并  $k$  个  $C$  (候选 Motif 集);

endfor

endfor

8.  $P_1$  收集并归并各  $P_i$  中的  $C$ ;

endif

else

begin

9. 构建一个以所有  $j$ -tuples 位置为索引的哈希表  $C$ ;

10. 根据  $C$  中选出最小计数的  $j$ -tuples 的一个并集依次构建总数  $M = l! / (l-d)!$  的投影序列  $T_1, \dots, T_M$ ;

11. for  $i=0$  到限定的最大投影数  $N_{\max}$  do

12. 构建大小为  $t \times l$  (所有  $j$ -tuples 的数目) 的哈希表  $H$ ;

13.  $P_i$  将所有  $t(n-l+1)$  个  $l$ -mers 根据投影映射到  $H$ ;

14. 计算哈希表  $H$  中每个桶的  $S_i$  值;

15. 统计  $S_i$  值超过门限的桶, 并将其分成  $p$  个群, 第  $i$  个群大小为  $a_i$ , 将第  $i$  个群发送给  $P_i, 1 \leq i \leq p$ ;

16. for all  $P_i$  where  $1 \leq i \leq p$  par-do

17. for  $P_i$  中的每个桶 do

18. EM 提炼;

19. if 当前提炼桶得到的 Motif 计分  $\geq$  最佳 Motif 计分 then  
用当前 Motif 替换最佳 Motif;

endfor

endfor

20. 清空哈希表  $H$ ;

21. if 最佳 Motif 计分  $\geq t$  then break;

22. if 已选用的投影总数  $\geq M$  then break;

endfor

23. return motif 一致序列及预测用例;

end

End

设发送 1 个字符需要 1 个单位的通信时间, 当 Motif 长度  $l \leq 16$  时, 数据通信只发生在开始阶段,  $P_1$  广播  $t$  个长度为  $n$  的序列给各个处理机, 共需要  $O(tn)$  的通信时间; 当 Motif 长度  $l > 16$  时, 需要提炼的桶所包含的  $l$ -mers 之和在最坏情况下为  $(3k+1)(t(n-l+1)/4^k)^{[11]}$ , 因此需要  $O(l(3k+1)(t(n-l+1)/4^k))$  的通信时间; 在各个处理机完成 Motif 查找时, 只返回最终结果, 通信时间为  $O(l)$ 。

### 3 实验

#### 3.1 实验环境和实验参数

实验平台由 3 种类型的 PC 计算机节点(见表 1)构成一个采用 100Mbps 以太网连接的异构机群系统,其中各个处理机节点运行 Red Hat Linux 9.0。

表 1 异构机群中各节点的配置参数信息

节点类型	CPU	内存容量
IBM	Pentium IV 2.0GHz	256MB
IBM	Pentium IV 2.4GHz	512MB
HP	Pentium IV 3.0GHz	512MB

选择一台 HP Pentium IV 3.0GHz 机器作为主节点。通过对主处理机和从处理机之间的各条通信链路进行多次通信实验测试,找出传送字节数和通信时间之间的关系。基于这些实验数据进行拟合计算得到  $G$  和  $g$  的值,如表 2 所列。

表 2 异构机群节点的通信性能参数信息

从处理机节点类型	$G$ (微秒)	$g$ (微秒)
IBM Pentium IV 2.0GHz	352	0.098
IBM Pentium IV 2.4GHz	203	0.088
HP Pentium IV 3.0GHz	173	0.086

实验中,通过在每种类型的处理机上反复运行 Motif 发现的 V-UPNT 算法,测出计算任务不同与计算时间的关系,然后进行拟合,得到  $u$  和  $w$  的值。

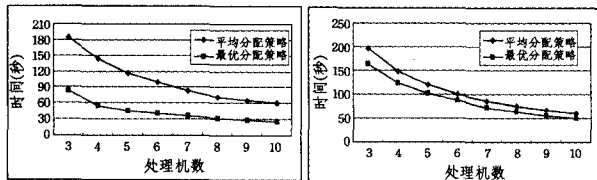
按 Motif 发现问题的标准测试参数,所有的输入样本包含  $t=20$  条序列,每条序列长度均为 600。每条序列中的每个碱基(A,C,G,T)以同样的概率独立地随机生成,Motif  $M$  也随机生成,并在输入系列中随机挑选一个位置把  $M$  的变体( $d$  个位置发生了突变)植入。

对于  $l > 16$  的情形,使用了 UPNT 算法的思想,按照 Buhler 和 Tompa<sup>[12]</sup>所计算出来的在 95%置信度条件下需要的最大投影数目来设定迭代次数, $k$  的大小均设为 7,门限值依据随机投影算法、统一投影算法和 UPNT 算法都设定为 4。

实验采用 MPI 和 C++ 语言编程实现 Motif 发现并行算法。

#### 3.2 实验结果与分析

图 2 给出处理机数变化而  $(l,d)$  固定时,分别按平均分配序列策略和按式(1)和式(2)建立的最优分配序列策略运行 Motif 发现并行算法所需的时间。其中图 2(a)为把  $(l,d)$  固定为(15,4)的挑战性问题,图 2(b)为把  $(l,d)$  固定为(19,6)的问题,这两种问题分别代表  $l < 16$  时使用精确查找 Motif 和  $l > 16$  时使用启发式查找的 Motif 发现并行算法运行的情况。



(a)  $(l,d)=(15,4)$

(b)  $(l,d)=(19,6)$

图 2 处理机增加时 Motif 发现并行算法运行时间

图 2 的结果表明,采取式(1)和式(2)建立的最优分配策略运行 Motif 发现并行算法所需的时间均比按平均分配策略运行 Motif 发现并行算法所需的时间少。

图 3 给出固定处理机为 6 台而  $(l,d)$  的规模发生变化时,分别按平均分配序列策略和按式(1)和式(2)建立的最优分配序列策略运行 Motif 发现并行算法所需的时间。

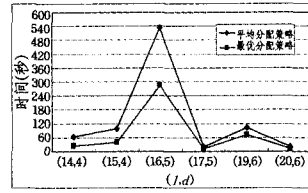


图 3  $(l,d)$  规模变化时 Motif 发现并行算法的运行时间

从图 3 可以看到,当机群中处理机数固定时,对于不同的问题规模,本文给出的基于最优分配策略的 Motif 发现并行算法运行时间要少于基于平均分配策略的 Motif 发现并行算法运行时间。对于  $l \leq 16$  的情形,由于使用精确查找思想,随着  $(l,d)$  规模的增大,运行时间也迅速增加,说明当  $l$  继续增大时再使用精确查找方法会由于运行时间过长而变得不现实。对于  $l > 16$  的情形,由于使用启发式搜索思想,运行时间较少,且运行时间随着  $(l,d)$  规模的不同而出现变化。

图 4 给出了  $(l,d)$  分别为(15,4)和(19,6),当处理机节点增加时,基于最优分配策略的 Motif 发现并行算法的加速比。

图 4 的结果表明,对于  $(l,d)$  分别为(15,4)和(19,6)的问题,随着处理机数的增加,Motif 发现并行算法获得的加速比也越来越大。但随着处理机数不断增加,加速比的增长幅度逐步放慢,这是因为此时分配给每个处理机的序列数据量逐渐减少,而处理机增加带来的额外通信开销相对增大。

图 5 给出了  $(l,d)$  分别为(15,4)和(19,6),基于最优分配策略的 Motif 发现并行算法中串行分量比例  $f$  随处理机数  $p$  增加的变化情况。

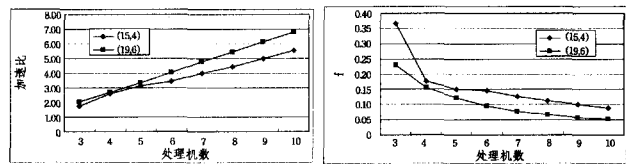


图 4 处理机增加时 Motif 发现并行算法的加速比

图 5 Motif 发现并行算法中串行分量比例  $f$  随处理机增加的变化情况

从图 5 可以看出,对于固定的问题规模,随着处理机数  $p$  的增加, $f$  逐渐减少,可见基于最优序列分配策略的 Motif 发现并行算法具有较好的可扩展性(Scalability)。

**结束语** 在异构机群系统上设计实现高效的 Motif 发现并行算法的关键是要建立自适应的最优序列分配模型,使之能够根据 Motif 长度  $l$  的不同,划分后缀群或任务桶,各个处理机节点同时结束并行查找 Motif 工作。文中给出的基于序列最优分配策略的 Motif 发现并行算法具有良好的加速和可扩展性,能够有效求解  $l < 16$  和  $l > 16$  的 Motif 发现问题。

#### 参考文献

- [1] Pevzner P, Sze S H. Combinatorial approaches to finding subtle signals in DNA sequences[C]//Proc 8th Int Conf Intel Syst Mol Biol. San Diego, CA, USA. AAAI press, 2000; 269-278
- [2] Chin F, Leung H. Voting Algorithms for Discovering Long Motifs[C]//Proc. of Asia-Pacific Bio-informatics Conference. London, 2005; 261-271
- [3] 王新建,杨德,陈建二. 基于统一投影和邻居桶聚集提炼策略的

基序查找算法[J]. 小型微型计算机系统, 2007, 28(11): 1963-1967

- [4] Wang C, Parthasarathy S. Parallel Algorithms for Mining Frequent Structural Motifs in Scientific Data[C]//Proc. of 18th annual international conference on Supercomputing, Malo, France, ACM ICS04, 2004: 31-40
- [5] Maddimsetty R P, Buhler J. Accelerator Design for Protein Sequence HMM Search[C]//Proc. of 20th Annual International Conference on Supercomputing, ACM ICS06, Cairns, Queensland, Australia, 2006: 288-296
- [6] Sandve G K, Nedland M. Accelerating Motif Discovery: Motif Matching on Parallel Hardware [C]//Proc. of WABI 2006. LN-BI 4175, Zurich, Switzerland, 2006: 197-206

- [7] 沈一飞, 陈国良, 张强峰. 基于纳米计算结构上的生物序列模式发现算法[J]. 小型微型计算机系统, 2007, 28(4): 635-639
- [8] 李锦, 钟诚, 卿海军. 基于分群策略的机群系统上 Motif 发现并行算法[J]. 计算机工程与科学, 2008, 30(11): 126-128
- [9] 钟诚, 李显宁. 异构机群系统上带返回信息的可分负载多轮调度算法[J]. 计算机研究与发展, 2008, 45(S1): 99-104
- [10] 钟诚, 范大娟. 异构机群系统上基于多轮分配方式的近似串匹配并行算法[J]. 计算机研究与发展, 2008, 45(S1): 105-112
- [11] Wang Jian-xin, Yang De. UPNT: Uniform Projection and Neighbourhood Thresholding method for motif discovery [J]. Bioinformatics Research and Applications, 2008, 4(1): 96-106
- [12] Buhler J, Tompa M. Finding Motifs using random projections [J]. J Comput Biol, 2002, 9(2): 225-242

(上接第 267 页)

中科院软件所与国家超级计算天津中心合作, 在“天河一号”上测试其自主开发的全隐式数值模拟程序, 通过使用尽量多的 CPU 核(最多至 82944 核)来进行强/弱可扩展性测试, 并尝试了数百亿网格未知数的问题规模。该测试采用 Williamson 标准测试集中的 Isolated mountain 算例。在强可扩展性测试中, 计算网格数目不变, 随着 CPU 核数增多, 求解时间相应减少, 图 9 给出了强可扩展性的测试结果, 程序从 4608 核扩展至 82944 核, 并行效率约为 60%。在弱可扩展性测试中, 固定每个 CPU 上的计算网格规模以及时间步长; 如果采用传统显式算法, 受稳定性条件的限制, 若 CPU 核数增加, 则总网格规模相应增加, 由于时间步长无法固定而造成计算步数增多, 导致计算时间随之增长; 但采用全隐式方法后, 随着 CPU 核数由 36 增至 57600(1600 倍), 计算步数无需增长, 但由于算法复杂度的关系, 计算时间缓慢增长(约 6.3 倍), 见图 10<sup>[11]</sup>。

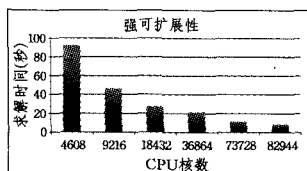


图 9 强可扩展测试结果

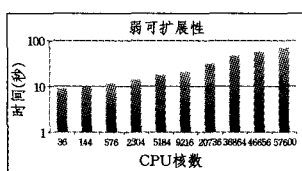


图 10 弱可扩展测试结果

### 3.8 地球外核热流动数值模拟

地球外核普遍被认为是由高热的液态金属构成的, 地磁发电机理论认为地球磁场由地球外核的热对流运动而产生。研究地球外核的热流动, 有助于加深对地球以及磁场形成的了解, 并有助于地外行星的探索。地球外核的热流动过程, 可以采用快速旋转球壳内的不可压 Navier-Stokes 方程来描述, 由于其非线性强、物理参数极端, 数值模拟有极大难度。

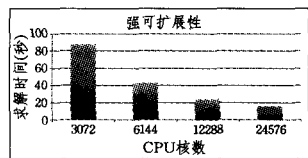


图 11 强可扩展测试结果

中科院软件所与国家超级计算天津中心合作, 在“天河一号”上对其自主开发的地球外核热流动数值模拟程序进行强可扩展性测试, 见图 11。测试采用 Christensen 测试集中的 0 号算例, 拟定常状态时的模拟结果。程序从 3072 核扩展至 24576 核, 并行效率达到 87%<sup>[12]</sup>。

**结束语** 本文通过对石油地震数据处理、飞行器流场模拟、生物分子动力学模拟、磁约束聚变数值模拟、湍流数值模拟、晶体硅分子动力学模拟、全球大气浅水波全隐式数值模拟、地球外核热流动数值模拟等应用进行了测试。测试结果表明, “天河一号”系统运行稳定, 可用性高、可扩展性强, 网络性能优于目前商用主流互连网络系统, 全系统整体性能具有一定优势。

下一步工作在于, 开展通用大规模科学计算软件的移植与优化, 充分发掘“天河一号”性能, 支持重点高性能计算应用领域; 基于“天河一号”CPU 与 GPU 异构协同并行架构, 面向重要科学计算应用领域, 与合作者共同开展 GPU 应用程序开发和 CPU+GPU 异构并行算法设计研究, 探索 GPU 以及 CPU+GPU 协同并行在高性能计算应用领域的发展趋势。

### 参考文献

- [1] BGP INC. China National Petroleum Coprorarion, Tianhe-1A” GeoEast-lighting Application Report
- [2] SINOPEC Geophysical Research Institute. “Tianhe-1A” iCluster Application Report
- [3] Institute of Mechanics, Chinese Academy of Sciences. “Tianhe-1A” Aircraft Flow Field Simulation CCFD software Application-Report
- [4] Shanghai Institute of Materia Medica, Chinese Academy of Sciences. “Tianhe-1A” Molecular Dynamic Simulation Application Report
- [5] <http://www.ks.uiuc.edu/Research/STMV/>
- [6] Liu Q, Jin Z, He P, et al. Performance Analysis of Large-Scale NAMD on DeepComp 7000. e-Science, 2010, 11: 83-89
- [7] Ethier S, Tang W M, Lin Z. Gyrokinetic particle-cell simulations of plasma microturbulence on advanced computing platforms [J]. J. Phys. : Conf. Series, 2005, 16
- [8] GTC Home Page[EB/OL]. <http://phoenix.ps.uci.edu/GTC/>
- [9] College of Engineering PEKING UNIVERSITY. “Tianhe-1A” Numerical Simulation of Turbulent Flow Application Report
- [10] Ge W, Hou C, Xu J, et al. Trans-scale Simulation of Silicon Deposition Process on Tianhe-1A. Submitted to Supercomputing, Seattle, USA, Nov. 2011
- [11] Institute of Software, Chinese Academy of Sciences “Tianhe-1A” Fully Implicit Simulation of Global Atmospheric Shallow Wave Application Report
- [12] Institute of Software, Chinese Academy of Sciences “Tianhe-1A” Heat Flow Simulation of Earth Outer Core Application Report