

服务核聚变研究领域的集群系统^①



李传冰, 李弘, 兰婷, 郑江山, 秦宏

(中国科学技术大学 物理学院 工程与应用物理系, 合肥 230026)

通讯作者: 李传冰, E-mail: xx123@mail.ustc.edu.cn

摘要: 在核聚变研究领域中, 高性能集群的应用十分广泛. 一方面需要借助集群进行高性能计算模拟托克马克中粒子的运动状态, 另一方面需要存储大量十分珍贵的放电数据以作后续的数据分析. 随着位于中国科学技术大学的大型反场箍缩磁约束聚变实验装置“Keda Torus eXperiment”(KTX) 的建设取得重要进展, KTX 实验室对于高性能计算以及实验数据存储具有高度的需求. 因此, 部署了一个高性能集群并对存储做出了安全设计. 根据 IOzone 的测试, GPFS 文件系统的冗余性机制发挥了作用, 数据的读写性能很稳定.

关键词: 核聚变; 高性能集群; xCAT; 并行文件系统; GPFS

引用格式: 李传冰, 李弘, 兰婷, 郑江山, 秦宏. 服务核聚变研究领域的集群系统. 计算机系统应用, 2019, 28(2): 49-54. <http://www.c-s-a.org.cn/1003-3254/6767.html>

High Performance Cluster for Nuclear Fusion

LI Chuan-Bing, LI Hong, LAN Ting, ZHENG Jiang-Shan, QIN Hong

(Department of Engineering and Applied Physics, School of Physical Sciences, University of Science and Technology of China, Hefei 230026, China)

Abstract: High performance clusters are applied to the study of nuclear fusion. On one hand, the simulation of particles movement in the Tokamak is rely on high performance clusters heavily. On the other hand, valuable discharge data have been stored for further researches. Keda Torus eXperiment reversed field pinch (KTX) device which is situated in University of Science and Technology of China works well, and the KTX laboratory highly demands for high performance computing and experimental data storage. Therefore, a high performance cluster is deployed and a security design for storage is proposed. Based on IOzone tests, the redundancy mechanism of the GPFS file system works, and the data read and write performance is stable.

Key words: nuclear fusion; high performance cluster; xCAT; parallel file system; GPFS

超级计算机是通过网络将数量庞大的计算节点连接起来, 以达到增强计算能力的目的. 在全球超级计算机 top 500 的榜单上, 中美两国长期交替霸占榜首. 2010 年 10 月, 天河一号二期 (天河-1A) 以峰值速度每秒 4700 万亿次, 成为当时世界上最快的超级计算机. 2013 年 6 月起至 2016 年 6 月, 最快的超级计算机是位于广州的天河二号^[1]. 2018 年 6 月 8 日, 美国橡树岭实验

室宣布开发出全球最快的超级计算机“顶点”(Summit). 理论上计算速度可以达到每秒 20 亿亿次, 超过迄今最快的中国超算“神威·太湖之光”, 有实力重夺全球第一.

高性能计算目前在社会的应用极其广泛, 航空航天、汽车制造、气象预测、芯片制造、军事情报搜集, 这些都需要极强的算力, 才能得出准确的计算结果^[2,3]. 而这些结果对工业设计, 前沿理论的推进起着重大作

① 基金项目: 安徽省自然科学基金 (1808085MA25)

Foundation item: Natural Science Foundation of Anhui Province (1808085MA25)

收稿时间: 2018-08-13; 修改时间: 2018-09-05; 采用时间: 2018-09-11; csa 在线出版时间: 2019-01-28

用. 在当今科研界, 主流的三大研究方式, 理论、实验、数值模拟, 三者相辅相成, 缺一不可. 高性能计算支撑的数值模拟, 完美的再现了无法在现实中实现的实验, 既节约了高昂的实验成本, 也避免了破坏环境生态. 目前, 在科研界内应用高性能计算最多的几个领域是核聚变、生物医药、微电子. 尤其是核聚变的研究, 在很多实验条件不成熟的情况下, 用计算机模拟是最好的研究方式, 既能够验证理论, 也能够提供对实验装置设计的指导, 在科学研究中能够提供难以估量的价值, 它的作用是理论和实验无法替代的.

现在, 有诸多 Tokamak 大科学装置都使用了超级计算机. 2010年8月, 在德国的 Forschungszentrum Jülich (FZJ) 有一个超级计算机专门为磁约束核聚变服务, 耗资 1000 万欧元, 由欧盟出资 50%, 德国出资 40%, EFDA 成员国出资 10% 组建的, 浮点运算的峰值可以达到 101 Tflops, 可以有效加快 ITER 的研究进展. 英国的卡尔汉姆聚变能研究中心 (CCFE) 为了支持 JET 和 MAST 项目也有一个超级计算机名叫 Diego, 通过 Myrinet 的网络架构将 36 台 HP 的服务器连接起来. 在印度, 等离子体研究中心的 Aditya 托克马克, 有一个用网线和光纤连接的集群, 600 台服务器, 提供网络、邮箱、域名解析等服务. 在日本, 于 2012 年至 2016 年投入使用一个计算能力强大的集群, 名叫 HELIOS, 包含有 4500 个计算节点, 72 000 个核心, 总的运算峰值可以达到 1.5 Pflops, 完全用于核聚变研究领域, 一半计算资源给日本使用, 另一半资源给欧洲使用. 可以清楚的看到, 高性能集群在托克马克中的广泛应用. 在科研探索中起着极其重要的作用, 强大的计算能力可以飞速推进数值模拟的发展.

1 KTX 装置对集群的要求

受控核聚变是人类实现最终能源的梦想, 磁约束是目前实现受控核聚变的主流技术思路之一. 磁约束有三种位形^[4-7]: ① 仿星器, ② 托克马克, ③ 反场箍缩. 反场箍缩^[8-11](Reversed Field Pinch, RFP) 是 β 值 (等离子体的压强与磁能的比值) 最大的磁约束聚变位形. 反场箍缩最主要的优点是工程实施简单, β 值高, 能够通过欧姆加热直接点火. 位于中国科学技术大学的大型反场箍缩磁约束聚变实验装置“Keda Torus eXperiment” (KTX) 实验装置采用的就是反场箍缩位形.

KTX 的数据先通过采集卡存储在本地, 之后再通过网络传输数据至集群. 采集卡分为低速采集和高速采集. 低速采集 250 KB/s, 每次采集有 900 道, 高速采集 60 MB/s, 每次采集有十几道, 放电时长 10–20 ms, 最多至 200 ms. 一年的实验月有 6 个月, 每周有 3 天实验日, 一天有 200 炮实验数据. 按照最大的原则计算, 一年高速采集的数据至多 3.5 TB 左右, 低速采集的数据最多 13 TB, 两者合计不会超过 20 TB.

KTX 对数据的存储有以下一些要求:

(1) 容错性. 如果一个存储单元出现了损坏, 可以不间断的接收新数据的存入, 还能完成原有数据的复制和转移.

(2) 冗余性. 将整个文件的两个或更多副本存储在不同的服务器上, 或将每个文件分割到多个服务器上. 从多个客户端并行访问副本.

(3) 有较快的读写性能.

(4) 兼容性好.

(5) 易于安装和配置.

KTX 实验装置的数据还需要有效的处理, 所以需要高性能集群有较高的计算能力. 快速做好数据的预处理才能高效的将计算结果演化为图像, 呈现给研究者面前, 用于做更深入的物理分析. 由于一炮数据具有多通道的特点, KTX 实验组对同时处理大批量数据具有强烈的需求, 所以对集群的并行处理能力有较高要求.

针对上述的 KTX 实验装置的计算和存储需求, 设计了集群的相关参数, 并做了相关配置和部署.

2 集群的硬件以及网络架构

为了满足核聚变研究领域的需求——可以较快的存储、读取、在线处理数据, 在集群的设计与部署中必须保证足够的计算性能以及文件系统的读写性能. CPU 核心数量多才能同时处理较多的作业, 在线同时处理 KTX 一炮多达数百道实验数据. CPU 的主频高, 运算速度快才能快速的计算数据. 文件系统读写性能佳, 才可以胜任快速存储实验数据, 也能够需要的时候, 快速地将数据从集群中读取出来.

集群平台为联想深腾 X8800 超算集群平台. 有 42 个计算刀片, 刀片式服务器的型号均为 Lenovo Think System SN550, 每个刀片有两个物理 CPU, 每个 CPU 有 16 个计算核心. 处理器都是 Intel Xeon 可扩展处理器金牌 6142, 主频 2.6 GHz. 集群的理论浮点运算

峰值可达 110 Tflops. 每个刀片有 128 GB 内存外加一块 240 GB 的企业级 SSD 硬盘. 另外还有两台机架式服务器用作管理和登陆, 型号为 Lenovo Think System SR550, 每台机架式服务器有两个物理 CPU, 每个 CPU 有 8 个核心, 主频 2.1 GHz, 处理器是 Intel Xeon 可扩展处理器金牌 4110 配置. 每个机架式服务器有 64 GB 内存, 4 块 600GB 的 SAS2.5 寸热插拔硬盘. 最后有 6 台 IO 节点, 也是机架式服务器, 型号与管理节点相同, 用来管理和负载均衡 I/O.

存储系统是 Lenovo Storage D3284, 集中式存储架构, GPFS 并行文件系统. 总容量 1.12 PB, 实际容量可利用率超过 70%. 数据盘数量达到 112 块, 系统实测总聚会 IO 带宽大于 4.0 GB/s.

网络系统有两套, 一个是计算网络, 另一个是管理网络. 计算网络采用 56 Gbps FDR Infiniband 网络, 互连所有节点 (包括计算节点、管理节点、IO 节点), 保证整个集群计算刀片点对点带宽 56 Gbps. 管理网络采用 1 Gb 以太网连接到所有节点, IPMI 监控管理网络, 能够实现全部服务器的远程监控管理.

本系统中最大的优势是采用了 Infiniband 高速网络, 用作点对点传输数据. Infiniband 高速网络有效解决了传统 I/O 总线造成的系统瓶颈^[12]. Infiniband 支持的是 RDMA 协议, 它允许应用程序直接读取或写入远程内存, 而无内核干预和内存拷贝发生, 故而有很强的读写性能^[13].

3 使用 xCAT 软件一站式部署系统

3.1 xCAT 的功能

xCAT 是一个高级的集群管理工具, xCAT 给高性能集群、云计算、数据中心提供了一体化管理服务. xCAT 的实施规模可以很大, 可以管理多达 1000 多台节点. xCAT 是开源的, 可以得到公开的参考资料. xCAT 能够十分方便地远程给高性能集群部署系统, 也能够批量下发命令给各从节点, 完成各种控制, 包括对所有从节点远程开关机. xCAT 能够实现集群的全自动化安装, 也即是基于网络的无人看守式的安装系统^[14,15].

3.2 xCAT 定义数据表

xCAT 在做系统部署前会定义一系列数据表, 存储在 xCAT 的数据库中. 为了使将来集群工作时, 各种网络能够高效、稳定地提供服务, 设计和定义好集群的网络表是很重要的. xCAT 中定义的集群网络分两种,

以太管理网络和 infiniband 高速网络. 使用 `tabedit network` 命令设置网络配置, 定义网段和子网掩码, 以及主服务器 master 的 IP. 使用 `tabedit hosts` 命令配置主机名列表和 IP 列表. 使用 `tabedit mac` 命令配置各服务器的网卡 mac 地址.

在完成一系列数据表的定义后, 对服务器进行群组划分, 便于以后分组控制, 统一执行主服务器 master 的远程命令. 给集群定义了四个组.

- (1) all, 代表全部服务器.
- (2) computeserver, 所有计算刀片构成的组.
- (3) ioserver, 所有 IO 节点构成的组.
- (4) slaves, 除了头节点 master 以外的所有服务器.

待完成数据表后, 给划分好群组的服务器导入相应的系统镜像, 给各个备选的系统镜像配置合适的分区文件, 以及需要安装的软件包, 还有安装完系统后需要配置的软件服务, 例如 ntp、rsyslog 等服务. 根据群组的设计, 集群的服务器分为三类. 第一类是两台管理节点, 一台用作集群主服务器 master, 同时作为并行作业调度器的头节点; 另一台用作登陆节点. 第二类是计算刀片, 有 42 个节点, 用来执行并行作业计算任务. 第三类是 IO 节点, 有 6 台, 用作管理和平衡负载文件系统的读写.

最后使用 `nodeset` 命令开始无人看管式远程部署系统, 每个服务器以 PXE 的方式启动, 推入系统镜像.

4 存储的安全设计

KTX 实验组不仅需要集群可以快速处理、读取数据, 还需要集群有足够的安全机制能够保障数据的安全性. KTX 一年当中只有一半的时间是实验期, 因此实验数据的产出来之不易, 一旦数据有损坏或者丢失, 对核聚变的研究都会造成不小的损失. 所以针对前文中提到的存储需要具有容错性、冗余性, 对 GPFS 并行文件系统进行了安全设计.

GPFS 文件系统本身具有数据安全性. GPFS 文件系统是一种日志文件系统, 为不同节点建立各自独立的日志, 日志中记录 metadata 的分布, 一旦节点发生故障后, 可以保证快速恢复数据. GPFS 的 fail-over 功能通过规划, 将数据分布到不同 failure group 内达到高可用性, 减少单点故障的影响. 为了保证数据可用性, GPFS 在多个 failure group 内为每个数据实例做备份^[16].

4.1 GPFS 可用性机制

GPFS 有一套可用性判断机制来保证数据完整性

与系统安全. GPFS 提供三套不同的 quorum 机制来判断系统当前的状态, 其中 File system Descriptor Quorum 是系统内置的, 不能做配置, 另外两种 Node Quorum 和 Tiebreaker Quorum 方式只能选择其中一个, 根据系统环境与可靠性分析来选择使用哪一种机制. File system Descriptor Quorum 的 File system Descriptor 描述的是文件系统信息的数据. 在几个不同的 failure-group 的磁盘上创建 GPFS 文件系统时, 文件系统的配置信息 (简称为 FD) 会被拷贝写到多个磁盘上, 以实现冗余备份. FD quorum 的机制是通过判断含有 FD 磁盘的在线数量来判断当前系统是否正常, 当超过半数的含有 FD 的磁盘掉线时, 就判断为系统故障, 文件系统就会自动关闭. Node Quorum 是通过主机状态来判断系统可用性的机制. GPFS 文件系统集群中, 可以设置多个主机节点为 quorum node. Node Quorum 的机制是通过判断 quorum node 的状态来判断系统是否正常, 当超过半数的 quorum node 在线时, 判断系统为正常, 反之, 将关闭文件系统. Tiebreaker Quorum 是通过磁盘的状态来判断系统的可用性. 可以设置系统通过监视指定的一些磁盘作为 Tiebreaker Disk. 当超过半数的 Tiebreaker Disk 掉线时, 则判断系统故障, 文件系统将会自动关闭. Tiebreaker Quorum 最多只能配置两个用来监控磁盘状态的 quorum 主机, 当 2 台 quorum 主机都宕机的话, GPFS 系统也将会关闭. 根据以上三种判断机制, GPFS 自动判断系统的状态, 当异常发生时自动关闭系统以保护系统和数据的完整性.

4.2 GPFS 可靠性分析与如何设计

基于上面阐述的 GPFS 可用性机制, 可以看出 GPFS 是通过上述的三种 quorum 机制来检查资源是否超过半数状态正常来判断系统状态好坏. 在设计 GPFS 文件系统集群的时候需要注意最好保证各种资源数都为 $2N+1$ 个 (N 是指数量), 也即数量为奇数, 来获得系统最大的可用性. 根据 FD Quorum 的设计, 一般情况下, 丢失一个 failure group 的磁盘实际不影响数据的完整性, 但是由于 FD Quorum $2N+1$ 的机制发挥作用, 文件系统仍将会关闭, 所以在创建一个 GPFS 文件系统时, 可以通过增加一个很小的本地的磁盘作为第三个 failure group. 以实现 $2N+1$ 的冗余设计. 本地的磁盘可以设置为只保存 GPFS 文件系统信息 (FD), 实际不参与数据读写. Node Quorum 机制如果采用了 $2N+1$ 个 quorum node, 那么这个系统就能容忍 N 个主机节点的离线. Tiebreaker Quorum 机制只能配置两个 quorum 主

机, 但是只要 tiebreaker 磁盘在线, 有一个 quorum 主机状态正常, 系统也能正常工作, 这意味着至少有一台 quorum 主机在线. 如果是主机节点数较多的情况, 采用 Tiebreaker Quorum 机制其可靠性不如 Node Quorum 机制.

鉴于集群的规模, 采用 Node Quorum 模式. 6 台 IO 节点是 quorum-manager node, 登陆节点是 quorum node, 6 台 IO 节点作为仲裁管理节点, 当它们损坏一半时, 才会关闭文件系统. 登陆节点则存储 GPFS 文件系统信息. 按照这种安全机制, 冗余性是很强的, 保证了 KTX 实验数据存储时的安全要求. 使用读写测试工具 IOzone, 对 GPFS 文件系统进行压力测试, 一方面是检查 GPFS 文件系统的性能, 另一方面, 检验并行文件系统的稳定性, 从结果来看, 文件系统的读、写性能在 4.9 GB/S、2.8 GB/s 左右. 测试结果放在 /gpfs 目录下, 下面为测试的结果:

```
[root@io01 gpfs]# sh sh
```

```
Command line used: ./iozone -i 0 -i 1 -+n -r 1M -s 32g -t 4 -+m nodelist -Rb output_new.xls
```

```
Output is in kBytes/sec
```

```
Time Resolution = 0.000001 seconds.
```

```
Processor cache size set to 1024 kBytes.
```

```
Processor cache line size set to 32 bytes.
```

```
File stride size set to 17 * record size.
```

```
Throughput test with 4 processes
```

```
Each process writes a 33554432 kByte file in 1024 kByte records
```

```
Test running:
```

```
Children see throughput for 4 initial writers = 2846204.53 kB/sec
```

```
Min throughput per process = 447215.25 kB/sec
```

```
Max throughput per process = 985903.31 kB/sec
```

```
Avg throughput per process = 711551.13 kB/sec
```

```
Min xfer = 15233024.00 kB
```

```
Test running:
```

```
Children see throughput for 4 readers = 4925115.38 kB/sec
```

```
Min throughput per process = 1135915.75 kB/sec
```

```
Max throughput per process = 1300317.00 kB/sec
```

```
Avg throughput per process = 1231278.84 kB/sec
```

```
Min xfer = 29333504.00 kB
```

Test cleanup:

"Throughput report Y-axis is type of test X-axis is number of processes"

"Record size = 1024 kBytes "

"Output is in kBytes/sec"

"Initial write" 2846204.53

"Rewrite" 0.00

"Read" 4925115.38

"Re-read" 0.00

iozone test complete.

Iozone 命令后接的参数-i 0 -i 1 指的是做 write/rewrite 和 read/re-read 测试, -r 1M 指定测试块大小为 1 MB, -s 32 g 指定测试文件大小为 32 GB, -t 4 指定测试时使用 4 线程, -+m nodelist 指定节点列表. 根据结果代码所示, 四线程总的写速度是 2.8 GB/s, 四线程中最快的写速度是 986 MB/s, 最慢的写速度是 447 MB/s; 四个线程总的读速度是 4.9 GB/s, 四个线程中最快的读速度是 1.3 GB/s, 最慢的读速度是 1.1 GB/s. 验证了 GPFS 文件系统有相当好的读写性能, 能够快速读写文件.

4.3 GPFS 冗余性测试

我们使用 6 台 IO 节点作为仲裁管理节点, 在 $2N+1$ 的 Node Quorum 冗余机制下, 应该能够保证在 3 台 IO 节点离线, 文件系统仍然能正常的工作. 为了验证 GPFS 文件系统的冗余性, 我们做了如下测试. 分别在有 6 台、5 台、4 台、3 台仲裁管理节点正常工作的情况下, 使用 IOzone 工具测试集群的读写速率. 测试结果如图 1 所示.

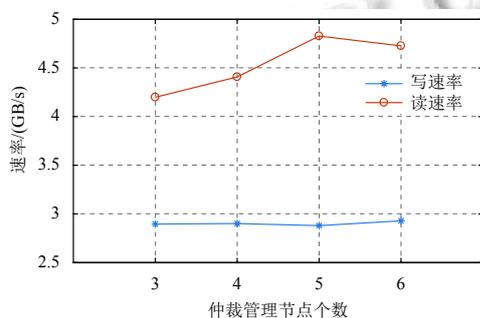


图 1 GPFS 文件系统冗余性检测

从图 1 中可以看出仲裁管理节点的个数从 6 台减少到 3 台的过程中, GPFS 文件系统的读写速率维持在 4.5 GB/s、2.8 GB/s 附近. 可见 $2N+1$ 的 Node

Quorum 机制发挥了作用, 最多在半数的仲裁管理节点失效的情况下, GPFS 文件系统仍然能够正常工作, 而且读写速率不受大的影响. GPFS 文件系统的冗余性得到了保证, 这样避免了以后在集群的使用中单点故障的问题, 数据的安全性得到了大大的增强.

5 结论与展望

为了方便核聚变的研究, 安装并部署了一个高性能集群, 考虑到 KTX 实验组的需求, 一方面是能够快速的处理、存储、读取数据, 另一方面保证数据有足够的安全保障. 据此, 对集群的设计进行了优化.

集群开启超线程后一共有 2688 个逻辑 CPU, 每个 CPU 的主频 2.6 GHz, 理论峰值浮点运算能力达到 110 Tflops, 保证了强大的计算能力. 在实际的 Linpack 测试当中发现浮点运算的峰值不理想, 经过研究发现, 是因为 CPU 的内存通道没有占满, 每个物理 CPU 有 6 通道, 但是只装了两根内存条. 为了提升浮点运算性能, 每台刀片加装 8 根同类型内存条, 一共加满到 12 根内存条. 另外集群选用 GPFS 并行文件系统, 高达 4.9 GB/s、2.8 GB/s 的读写速率, 提供了优秀的读写性能. 选用 infiniband 高速网络, 保证了较高的网络传输速度和较低的延迟率, 并使用 xCAT 管理工具对集群进行了有效稳定的管理. 设计了 KTX 实验数据存放于 GPFS 并行文件系统的安全保障机制, 保证了数据的容错性、冗余性、稳定性.

未来 KTX 实验室会将实验数据大规模存入, 数据都是 HDF5 格式的, 下一步会规划与设计快速读取基于 HDF5 数据格式的方案, 并设计和完善在线处理数据的功能.

参考文献

- 周兴铭. 高性能计算技术发展. 自然杂志, 2011, 33(5): 249-254. [doi: 10.3969/j.issn0253-9608.2011.05.001]
- 赵毅, 朱鹏, 迟学斌, 等. 浅析高性能计算应用的需求与发展. 计算机研究与发展, 2007, 44(10): 1640-1646.
- 迟学斌, 赵毅. 高性能计算技术及其应用. 中国科学院院刊, 2007, 22(4): 306-313. [doi: 10.3969/j.issn.1000-3045.2007.04.012]
- Wesson J, Campbell DJ. Tokamaks. Oxford: Oxford University Press, 2011.
- 王乃彦. 聚变能及其未来. 北京: 清华大学出版社, 2001.
- 毛文哲. KTX 装置的概念设计和 RFP 电流启动阶段的实

- 验研究[博士学位论文]. 合肥: 中国科学技术大学, 2015.
- 7 李晨光. KTX 反场箍缩装置中磁流体不稳定性的反馈控制研究与设计[博士学位论文]. 合肥: 中国科学技术大学, 2014.
 - 8 Bodin HAB, Newton AA. Reversed-field-pinch research. *Nuclear Fusion*, 1980, 20(10): 1255–1324. [doi: [10.1088/0029-5515/20/10/006](https://doi.org/10.1088/0029-5515/20/10/006)]
 - 9 Schoenberg KF, Gribble RF, Baker DA. Oscillating field current drive for reversed field pinch discharges. *Journal of Applied Physics*, 1984, 56(9): 2519–2529. [doi: [10.1063/1.334316](https://doi.org/10.1063/1.334316)]
 - 10 Sarff JS, Almagri AF, Anderson JK, *et al.* Tokamak-like confinement at a high beta and low toroidal field in the MST reversed field pinch. *Nuclear Fusion*, 2003, 43(12): 1684–1692. [doi: [10.1088/0029-5515/43/12/014](https://doi.org/10.1088/0029-5515/43/12/014)]
 - 11 Lorenzini R, Martines E, Piovesan P, *et al.* Self-organized helical equilibria as a new paradigm for ohmically heated fusion plasmas. *Nature Physics*, 2009, 5(8): 570–574. [doi: [10.1038/NPHYS1308](https://doi.org/10.1038/NPHYS1308)]
 - 12 贾志国, 赵青苹, 董小社, 等. InfiniBand: 一种新型的高速互连网络. *计算机工程与应用*, 2003, 39(9): 172–175. [doi: [10.3321/j.issn:1002-8331.2003.09.055](https://doi.org/10.3321/j.issn:1002-8331.2003.09.055)]
 - 13 刘天华, 朱宏峰, 杜梅, 等. RDMA 技术的研究与应用. *沈阳师范大学学报(自然科学版)*, 2006, 24(2): 185–188. [doi: [10.3969/j.issn.1673-5862.2006.02.016](https://doi.org/10.3969/j.issn.1673-5862.2006.02.016)]
 - 14 张雨馨. xCAT 在 IBM 高性能计算机集群系统中的应用. *中国管理信息化*, 2014, 17(21): 34–35. [doi: [10.3969/j.issn.1673-0194.2014.21.025](https://doi.org/10.3969/j.issn.1673-0194.2014.21.025)]
 - 15 吴琼, 王颖, 梁金铃, 等. Linux 环境下基于 Qt 和 xCAT 的集群部署系统. *计算机系统应用*, 2013, 22(7): 27–30. [doi: [10.3969/j.issn.1003-3254.2013.07.006](https://doi.org/10.3969/j.issn.1003-3254.2013.07.006)]
 - 16 李永盛. 基于并行文件系统的集群高可用性研究与应用[硕士学位论文]. 沈阳: 中国科学院沈阳计算技术研究所, 2008.