



单位代码 由教务处统一填写

学 号 3XXXXXXXX

分 类 号 TP301

密 级 不涉密

北京航空航天大学

B E I H A N G U N I V E R S I T Y

毕业设计(论文)

远程内存池的设计与实现

院 (系) 名 称 计算机学院

专 业 名 称 计算机科学与技术

学 生 姓 名 张三

指 导 教 师 王小五

200X年6月



本人声明

我声明，本论文及其研究工作是由本人在导师指导下独立完成的，在完成论文时所利用的一切资料均已在参考文献中列出。

作者： 张三

签字：

时间： 200X年Y月



远程内存池的设计与实现

学 生： 张三

指导教师： 王小五

摘 要

随着网络速度的提高，集群系统的应用越来越普遍，如何有效利用集群中空闲结点的内存以提高集群的整体性能这一问题越来越受到重视。

本文介绍了一种...

本文实现了一个 XXX 的原形系统，并对它的可用性和性能进行了测试。测试表明...

本文将首先对近年来相关的研究进行叙述，指出前人已经进行的工作和问题所在；之后，本文将介绍远程内存池的设计和远程内存池原形系统的实现；再后将给出对原形系统的测试数据，说明它的可用性和性能；最后将指出不足和改进的方向。

关键词： 第一, 第二



Remote Memory Pool—Design and Implementation

Author: Zhang San

Tutor: Wang Xiaowu

Abstract

In recent years there has been a great interest in the use of clusters. Exploiting idle memory in clusters to avoid swapping and improve the performance is taken more and more attention.

In this paper, we introduce ...

This paper describes the design and implementation of XXX, and also implements a prototype. We tested and profiled the prototype, and found that ...

This paper first introduce previous work in related fields, then present the design detail of our implementation. In section 5, it deals with the detailed performance evaluation and analysis of our designs. Then, in the last section, it draw the conclusion and state future works.

Key words: First, Second



目录

1 绪论	1
1.1 背景及目的	1
1.2 国内外相关研究	1
1.3 问题的提出	1
1.4 论文构成	1
2 技术背景	2
2.1 Linux 操作系统	2
2.2 内核模块	2
2.3 地址空间与 VMA	2
3 结构与流程	3
4 设计与实现	4
5 性能测试	5
结论	6
致谢	7
附录A XXX 源码简介	8
参考文献	8
附录B XXX 关键数据结构源码	9
B.1 connector	9
附录C XXX 的使用	10
C.1 编译	10



1 绪论

1.1 背景及目的

介绍背景及目的.

1.2 国内外相关研究

介绍国内外相关研究.

根据要求, 如果引用是句子的一部分, 如: [?]提出, ..., 就要用\cite; 否则, 如: 有些人认为不是这样的^{[2][3]}. 但是..., 就要用\upcite

1.3 问题的提出

提出问题.

1.4 论文构成

论文结构.



2 技术背景

2.1 Linux 操作系统

2.2 内核模块

2.3 地址空间与 VMA

3 结构与流程

这一部分保密.

图3.1中, ...

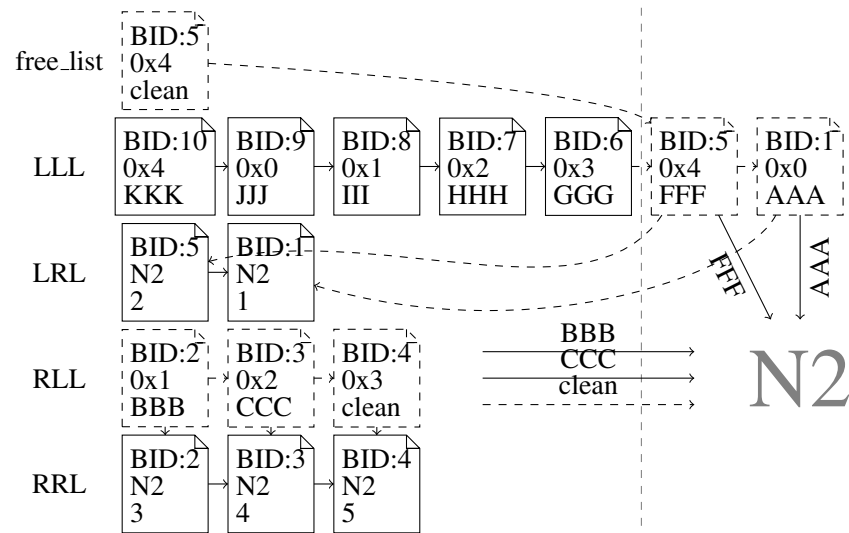


图 3.1: 示例三



4 设计与实现

这一部分保密.



5 性能测试

性能很好啊.



结论

根据要求, 结论一节是没有节号的.



致谢

首先,我要衷心感谢北京航空航天大学计算机学院四年来对我的辛勤培养。我在计算机学院四年的经历将成为我一生的宝贵财富。

然后,我要感谢....

感谢 Linus Torvalds,他无私的开放了 Linux 的源代码,为我们提供了课题和饭碗。感谢 VIM 的作者 Bram Moolenaar(虽然我没有给乌干达的可怜儿童和 Bram 捐款),他设计的编辑器是世界上最好的编辑器。感谢 $\text{T}_{\text{E}}\text{X}$ 的作者 D. E. Knuth 教授、 $\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X} 2_{\epsilon}$ 的作者 Leslie Lamport 以及 $\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X} 2_{\epsilon}$ 中无数个宏包的设计者们,他们的设计组成的这套排版系统使我可以排出精美的论文,另外在紧张的论文撰写同时给我带来了许多快乐。

最后,感谢 $\text{C}_{\text{T}}\text{E}_{\text{X}}$ 社区的所有朋友们。



参考文献

- [1] E. A. Anderson and J. M. Neeffe. An exploration of network ram[R]. Technical Report CSD-98-1000, UC Berkley, December 1994.
- [2] A. Acharya and S. Setia. Availability and utility of idle memory in workstation clusters[J]. *ACM SIGMETRICS Performance Evaluation Review*, 1999.
- [3] N.J. Boden, D. Cohen, R.E. Felderman, A.E. Kulawik, C. L. Seitz, J.N. Seizovic, and W. Su. Myrinet: A gigabit-per-second local area network[J]. *IEEE Micro*, 1995, 15(1):29–36.
- [4] Daniel P. Bovet and Marco Cesati. *Understanding the Linux Kernel*[M]. O'Reilly, 2005.
- [5] Jonathan Corbet, Alessandro Rubini, and Greg Kroah-Hartman. *Linux Device Drivers*[M]. O'Reilly, 2005.
- [6] M. J. Feeley, W. E. Morgan, F. H. Pighin, A. R. Karlin, H. M. Levy, and C. A. Thekkath. Implementing global memory management in a workstation cluster[J]. *ACM SIGOPS Operating Systems Review*, 1995, pages 201–212.
- [7] M. D. Flouris and E. P. Markatos. The network ramdisk: Using remote memory on heterogeneous nodes[J]. *Cluster Computing*, 1999, 2(4):281–293.
- [8] M. J. Frankling, M. J. Carey, and M. Livny. Global memory management in client-server dbms architectures[A]. In *Proceeding of the 18th VLDB Conference*[C]. 1992.
- [9] Jizhong Han, Dan Zhou, Xubin He, and Jinzhu Gao. I/O profiling for distributed ip storage systems[A]. In *Proceeding of The Second International Conference on Embedded Software and Systems*[C]. 2005.
- [10] X. He, Q. Yang, and M. Zhang. A caching strategy to improve iscsi performance[A]. In *Proceeding of Local Computer Networks*[C]. 2002.
- [11] M. R. Hines, M. Lewandowski, and K. Gopalan. Anemone: Adaptive network memory engine[D]. Master's thesis, Florida State University, 2003.



-
- [12] L. Iftode, K. Li, and K. Petersen. Memory servers for multicomputers[A]. In *Proceeding of the IEEE Spring COMPCON 93*[C]. 1993:538-547.
- [13] S. Koussih and S. Setia A. Acharyam. Dodo:a user-level system for exploiting idle memory in workstation clusters[A]. In *Proceeding of the Eighth IEEE International Symposium on High Performance Distributed Computing*[C]. 1999.
- [14] S. Liang, R. Notonha, and D. K. Panda. Swapping to remote memory over infiniband: An approach using a high performance network block device[J]. *IEEE Cluster Computing*, September 2005.
- [15] Robert Love. *Linux Kernel Development*[M]. Sams Publishing, 2005.
- [16] E. P. Markatos and G. Dramitions. Implementation of a reliable remote memory pager[A]. In *Proceeding of the 1996 Usenix Technical Conference*[C]. 1996.
- [17] John McCalpin. Streambenchmark. <http://www.streambench.org>.
- [18] T. Newhall, S. Finney, K. Ganchevm, and M. Spiegel. Nswap:a network swapping module for linux clusters[A]. In *Proceeding of Euro-Par'03 International Conference on Parallel and Distributed Computing*[C]. Klagenfurt, Austria, 2003.
- [19] J. Oleszkiewicz, L. Xiao, and Y. Liu. Parallel network ram: Effectively utilizing global cluster memory for large data-intensive parallel programs[A]. In *Proceeding of International Conference on Parallel Proceeding*[C]. 2004:577-592.
- [20] Fabrizio Petrini, Eitan Frachtenberg, Adolfo Hoisie, and Salvador Coll. Performance evaluation of the quadrics interconnection network[J]. *Journal of Cluster Computing*, April 2003, 6(2):125–142.
- [21] H. Tang Sun, M. Chen, and J. Fan. A scalable dynamic network memory service system[A]. In *Proceeding of High-Performance Computing in Asia-Pacific Region*[C]. 2005.
- [22] T. S. Trevisan, V. S. Costal, L. Whately, and C. L. Amorim. Distributed shared memory in kernel mode[A]. In *Proceeding of Computer Architecture and High Performance Computing*[C]. 2002.



-
- [23] L. Xiao, S. Chen, and X. Zhang. Adaptive memory allocations in clusters to handle unexpectedly large data-intensive jobs[J]. *IEEE Transactions on Parallel and Distributed Systems*, June 2004, 15(6):577–592.
- [24] The infiniband architecture. <http://www.infinibandta.org/spec>.



附录A XXX 源码简介

本附录对 XXX 的源代码做一简要介绍。

以下是 XXX 源码树的结构：

```
|-- Makefile
|-- common
|   |-- Makefile
|   |-- config.h
|   |-- debug.c
|   |-- debug.h
|   |-- fast_printk.c
|   |-- slabs.c
|   |-- slabs.h
|   |-- tty_printk.c
|   '-- ulist.h
```




附录B XXX 关键数据结构源码

本附录列出 RMP 关键数据结构的代码，以便读者理解前文提及的数据结构。

B.1 connector

connector.h 文件定义了 connector 结构。

```
1  /*
2   *
3   * connector.h
4   *
5   * ZhangSan, Mar. 20, 200X
6   *
7   * 连接抽象层是一个面向内核提供抽象连接的机制，每种类型的连接通过提供
8   * 一组 ops 和一个连接标识向核心应用提供透明的连接。
9   */
10
11 #ifndef CONNECTOR_H
12 #define CONNECTOR_H
13
14 #include "common/config.h"
15
16 #ifdef __KERNEL__
17 #include <linux/module.h>
18 #include <linux/byteorder/generic.h>
19 #include <asm/msr.h>
20 #endif
21
22 #define CONNSTRING_MAX 32
23 #define CONN_TYPE_SOCKET 1
24 #define CONN_TYPE_NOEXIST 10
25
26 /* 消息标识*/
```



附录C XXX 的使用

本附录介绍 XXX 的编译和使用。

C.1 编译

XXX 的源码经测试可以在 Kernel 2.6.11.12 和 Kernel 2.6.13 上编译通过...

获得源代码后，在根目录下执行“make”即可开始编译。正常的编译过程将有以下输出：

```
$ make
make -C manager
make[1]: Entering directory `/tmp/rmp/manager'
rm -f ../common/debug.o
```

之后会生成以下文件：

```
./common/fast_printk.ko
./common/tty_printk.ko
```