

分类号 TP391???

学号 17060062

U D C 681???

密级 公开

工学博士学位论文

基于自然语言模型的文件访问模式分析与文件预取方法研究

博士生姓名 陈辉

学科专业 计算机科学与技术

研究方向 高性能计算

指导教师 周文强 研究员

国防科技大学研究生院

二〇一九年十月

An NLP Model-Based Mechanism for I/O Access Pattern Analysis and File Prefetching

Candidate: **Hui Chen**

Supervisor: **Prof. Enqiang Zhou**

A dissertation

Submitted in partial fulfillment of the requirements

for the degree of **Ph.D of Engineering**

in **Computer Science and Technology**

Graduate School of National University of Defense Technology

Changsha, Hunan, P. R. China

October, 2019

独 创 性 声 明

本人声明所呈交的学位论文是我本人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表和撰写过的研究成果，也不包含为获得国防科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文题目：基于自然语言模型的文件访问模式分析与文件预取方法研究

学位论文作者签名：_____ 日期：_____ 年 _____ 月 _____ 日

学位论文版权使用授权书

本人完全了解国防科技大学有关保留、使用学位论文的规定。本人授权国防科技大学可以保留并向国家有关部门或机构送交论文的复印件和电子文档，允许论文被查阅和借阅；可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密学位论文在解密后适用本授权书。）

学位论文题目：基于自然语言模型的文件访问模式分析与文件预取方法研究

学位论文作者签名：_____ 日期：_____ 年 _____ 月 _____ 日

作者指导教师签名：_____ 日期：_____ 年 _____ 月 _____ 日

目 录

摘 要	i
ABSTRACT	ii
第一章 绪论	1
1.1 研究背景	1
1.2 国内外研究现状	1
1.3 本文主要工作	1
1.4 论文结构	1
第二章 分布式文件系统缓存预取研究	2
2.1 分布式文件系统层次结构	2
2.1.1 GlusterFS 整体架构	2
2.2 文件缓存预取技术	2
2.2.1 缓存	2
2.2.2 主动预取	3
2.3 基于 GlusterFS Tiering 功能的缓存管理模块设计	3
2.3.1 Trace 模块	3
2.3.2 访问模式识别模型建立	3
2.3.3 运行时缓存管理	3
2.4 本章小结	3
第三章 基于自然语言处理模型的文件访问模式识别	4
3.1 基于词向量模型的文件名向量化	4
3.1.1 词向量概述	4
3.1.2 基于子词模型的文件名向量化	4
3.2 基于循环神经网络的文件访问模式分析	4
3.2.1 循环神经网络及其在时间序列分析中的应用	4
3.2.2 采用 GRNN (Gated Recurrent Neural Network) 的文件访问模式分析模型	4
3.3 本章小结	4
第四章 实验设计与分析	5
4.1 实验环境与工作负载	5
4.1.1 存储服务器集群搭建	5
4.1.2 工作负载选取	5

4.2	实验数据采集与预处理	5
4.2.1	文件名向量化	5
4.2.2	文件访问日志记录与预处理	5
4.3	循环神经网络模型训练与仿真测试	5
4.4	实验结果与分析	5
4.5	本章小结	5
第五章	总结	6
5.1	本文工作总结	6
5.2	未来工作展望	6
致谢	7
参考文献	8
附录 A	模板提供的希腊字母命令列表	9

表 目 录

图 目 录

摘 要

关键词: 分布式文件系统; 文件预取; 词向量; 循环神经网络

ABSTRACT

Many applications spend a large proportion of the execution time to access files. To narrow the increasing gap between computing and I/O performance, several optimization techniques were adopted, such as data prefetching and data layout optimization. However, the effectiveness of these optimization processes heavily depends on the understanding of the I/O behavior. Traditionally, spatial locality and temporal locality are mainly considered for data prefetching and scheduling policy. Whereas for most real-world workloads, the file access pattern is hard to capture.

For the goal of deeply and intelligently understanding the I/O access pattern of modern applications, and efficiently optimizing the performance of current file systems, we propose a new mechanism to embed file names to vectors and train a GRNN (Gated Recurrent Neural Network) to provide policies for file prefetching and cache replacing.

Key Words: Distributed File Systems, File Prefetching, Word Embedding, RNN

符号使用说明

HPC	高性能计算 (High Performance Computing)
cluster	集群
Itanium	安腾
SMP	对称多处理
API	应用程序编程接口
PI	聚酰亚胺
MPI	聚酰亚胺模型化合物, N- 苯基邻苯酰亚胺
PBI	聚苯并咪唑
MPBI	聚苯并咪唑模型化合物, N- 苯基苯并咪唑
PY	聚吡咙
PMDA-BDA	均苯四酸二酐与联苯四胺合成的聚吡咙薄膜
ΔG	活化自由能 (Activation Free Energy)
χ	传输系数 (Transmission Coefficient)
E	能量
m	质量
c	光速
P	概率
T	时间
v	速度

第一章 绪论

1.1 研究背景

1.2 国内外研究现状

1.3 本文主要工作

1.4 论文结构

第二章 分布式文件系统缓存预取研究

2.1 分布式文件系统层次结构

描述主流分布式系统在层次结构上的共性，分层存储带来的优点与伴随的性能鸿沟。腾讯云社区 构建存储系统时需要基于成本和性能来考虑，因此存储系统通常采用多层不同性价比的存储器件组成存储层次结构。构建高效合理的存储层次结构，可以在保证系统性能的前提下，降低系统能耗和构建成本，利用数据访问局部性原理，可以从两个方面对存储层次结构进行优化。从提高性能的角度，可以通过分析应用特征，识别热点数据并对其进行缓存或预取，通过高效的缓存预取算法和合理的缓存容量配比，以提高访问性能。从降低成本的角度，采用信息生命周期管理方法，将访问频率低的冷数据迁移到低速廉价存储设备上，可以在小幅牺牲系统整体性能的基础上，大幅降低系统的构建成本和能耗。以 GlusterFS 为例，介绍其层次架构。引用相关数据图表来说明不同层存储之间的性能差异

2.1.1 GlusterFS 整体架构

GlusterFS[1] 是一个开源、可扩展的分布式文件系统，目前被广泛应用于各类商业存储服务器集群，其主要特性如下：

- 全局命名空间。
- 集群存储管理。
- 模块化的层次机构。
- 内置 replication and geo-replication 特性。
- 自修复功能。
- 高效负载均衡。

2.2 文件缓存预取技术

接上节内容，引出缓存与预取技术在降低访问延迟方面的巨大作用

2.2.1 缓存

被动缓存介绍

2.2.2 主动预取

简述主动预取的定义与缓存的区别。

文件预取的定义

预取 (prefetching), 也称为预分页 (prepaging) 或预读 (read-ahead), 是操作系统数据读取过程中的重要优化方法 [2] [3] [4]。它通过隐藏或减少对非缓存数据的访问延迟来补充传统的缓存策略 (例如 LRU)。其目标是预测将来的数据访问, 并在请求数据之前使其在内存中可用。

与缓存技术的被动数据迁移不同, 预取技术的关键在于对即将访问的数据内容和生存周期进行主动预测。

- **内容预测:** 与被动缓存不同, 主动预取需要预测程序下一阶段可能访问的数据, 是对数据访问的空间局部性的扩充。准确的预测将极大地降低访问延迟, 而错误预测将会引发浪费传输带宽, 挤占缓存空间等负面影响。
- **时机预测:** 与缓存机制中的时间局部性类似, 主动预取需要针对缓存数据的时效性和生命周期建立有效的评估。非缓存数据的及时预取有助于提高命中率, 而清除短期内不再读取的数据将提高缓存空间的利用率。

预取的主要流程

1. 针对特定负载, 提取访问日志;
2. 分析访问日志, 对该负载的文件访问模式进行抽象表达;
3. 将负载的访问模式作为依据, 引导文件系统进行主动预取。

2.3 基于 GlusterFS Tiering 功能的缓存管理模块设计

2.3.1 Trace 模块

2.3.2 访问模式识别模型建立

2.3.3 运行时缓存管理

2.4 本章小结

第三章 基于自然语言处理模型的文件访问模式识别

3.1 基于词向量模型的文件名向量化

3.1.1 词向量概述

3.1.2 基于子词模型的文件名向量化

p.1

词向量是将单词在向量空间中分布式表示的自然语言模型，该类模型可帮助学习算法通过对相似单词进行分组来在自然语言处理任务中实现更好的性能。最早使用单词表示法的方法可以追溯到 1986 年 Rumelhart, Hinton 和 Williams 的工作。此想法此后已成功应用于统计语言建模 [1]。后续工作包括自动语音识别和机器翻译 [14、7] 以及各种 NLP 任务 [2、20、15、3、18、19、9] 的应用。

Mikolov 等于 2013 年引入了 Skip-gram 模型，这是一种从大量非结构化文本数据中学习单词的高质量向量表示的有效方法。与大多数以前使用的用于学习单词向量的神经网络体系结构不同，Skip-gram 模型的训练不涉及密集矩阵乘法，因此训练过程非常高效：经过优化的单机实现可以在一天中培训超过 1000 亿个单词。

More details about Skip-gram.

p.1 通过为每个单词分配不同的向量，学习此类表示形式的流行模型会忽略单词的形态。这是一个限制，特别是对于具有大量词汇和许多稀有单词的语言。在本文中，我们提出了一种基于跳过图模型的新方法，其中每个单词都表示为一包字符 n-gram。向量表示与每个字符 n-gram 相关；单词被表示为这些表示的总和。我们的方法快速，可以快速在大型语料库上训练模型，并允许我们为未出现在训练数据中的单词计算单词表示。我们在词的相似性和类比任务上用九种不同的语言评估词的表示形式。通过与最近提出的形态词表示法进行比较，我们表明，我们的向量在这些任务上达到了最先进的性能。

3.2 基于循环神经网络的文件访问模式分析

3.2.1 循环神经网络及其在时间序列分析中的应用

3.2.2 采用 GRNN (Gated Recurrent Neural Network) 的文件访问模式分析模型

3.3 本章小结

第四章 实验设计与分析

4.1 实验环境与工作负载

4.1.1 存储服务器集群搭建

4.1.2 工作负载选取

编译。网页访问

4.2 实验数据采集与预处理

4.2.1 文件名向量化

4.2.2 文件访问日志记录与预处理

4.3 循环神经网络模型训练与仿真测试

4.4 实验结果与分析

4.5 本章小结

第五章 总结

5.1 本文工作总结

5.2 未来工作展望

致 谢

衷心感谢导师 xxx 教授和 xxx 副教授对本人的精心指导。他们的言传身教将使我终生受益。

感谢 NUDTPAPER，它的存在让我的论文写作轻松自在了许多，让我的论文格式规整漂亮了许多。

参考文献

- [1] GlusterFS: A scalable network filesystem. <http://www.gluster.org/>.
- [2] Griffioen J, Appleton R. Reducing File System Latency using a Predictive Approach. [C]. In USENIX summer. 1994: 197–207.
- [3] Amer A, Long D D, Burns R C. Group-based management of distributed file caches [C]. In Proceedings 22nd International Conference on Distributed Computing Systems. 2002: 525–534.
- [4] Nanopoulos A, Katsaros D, Manolopoulos Y. A data mining algorithm for generalized web prefetching [J]. IEEE Transactions on Knowledge & Data Engineering. 2003 (5): 1155–1169.

附录 A 模板提供的希腊字母命令列表

大写希腊字母:

Γ \Gamma	Λ \Lambda	Σ \Sigma	Ψ \Psi
Δ \Delta	Ξ \Xi	Υ \Upsilon	Ω \Omega
Θ \Theta	Π \Pi	Φ \Phi	
Γ \varGamma	Λ \varLambda	Σ \varSigma	Ψ \varPsi
Δ \varDelta	Ξ \varXi	Υ \varUpsilon	Ω \varOmega
Θ \varTheta	Π \varPi	Φ \varPhi	

小写希腊字母:

α \alpha	θ \theta	o o	τ \tau
β \beta	ϑ \vartheta	π \pi	υ \upsilon
γ \gamma	ι \iota	ϖ \varpi	ϕ \phi
δ \delta	κ \kappa	ρ \rho	φ \varphi
ϵ \epsilon	λ \lambda	ϱ \varrho	χ \chi
ε \varepsilon	μ \mu	σ \sigma	ψ \psi
ζ \zeta	ν \nu	ς \varsigma	ω \omega
η \eta	ξ \xi	\varkappa \varkappa	\digamma
α \upalpha	θ \uptheta	o \mathrm{o}	τ \uptau
β \upbeta	ϑ \upvartheta	π \uppi	υ \upupsilon
γ \upgamma	ι \upiota	ϖ \upvarpi	ϕ \upphi
δ \updelta	κ \upkappa	ρ \uprho	φ \upvarphi
ϵ \upepsilon	λ \uplambda	ϱ \upvarrho	χ \upchi
ε \upvarepsilon	μ \upmu	σ \upsigma	ψ \uppsi
ζ \upzeta	ν \upnu	ς \upvarsigma	ω \upomega
η \upeta	ξ \upxi		

希腊字母属于数学符号类别, 请用\bm命令加粗, 其余向量、矩阵可用\mathbf。