

读者网络访问控制统计网关的原理及设计

——以南京工业职业技术学院为例



南京工业职业技术学院

张迎春

目录

引、融入示范校建设，参与院校建设工作

一、当前数据库日志统计现状

二、技术原理

三、统计网关部署，改变读者习惯的一些经验

四、一些有趣的功能



融入示范校建设，参与院校建设工作

期刊读者访问统计信息有什么作用？

1、可以分析学校里面是教授更努力，还是副教授、讲师更乐于做研究。

这可以为学校领导制定各项方针政策时，提供依据，科学决策，避免拍脑袋决定！



融入示范校建设，参与院校建设工作

期刊读者访问统计信息有什么作用？

2、根据师生研究习惯，为院校文化建设提供决策。

例如：老师是喜欢在学校做研究，还是喜欢在家专研学问。若在校，则需要提供加强院校文化建设，为广大教师提供良好的3C研究环境。



当前数据库统计信息现状带来的问题

访问量商家弄虚作假

访问量统计不准，不符合国际标准，未考虑中国国情

统计数据太粗放

数据库采购决策支持

无法提供个性化服务

期刊读者访问统计网关，包括访问读者详细信息的访问日志



1.1、访问量商家弄虚作假

提供给商家多个ip段，只有两个ip段是出口地址，其他多个ip段是内网地址。

暑假后，商家提供的日志中，包括多个不会出现的ip地址。



1.2、访问量统计不准

不符合国际协议，国际count协议，30秒内只能算一次有效下载。

未考虑中国网络情况，延时更严重，读者会点击多次，故商家统计数据不精确。

210.28.96.74	高职院校对农村富余劳动力转移培训的作用及模式	中国学术期刊 网络出版总库	2	432-433	caj下载	0.00	0.00	2013-5-31 16:18:50
210.28.96.74	高职院校对农村富余劳动力转移培训的作用及模式	中国学术期刊 网络出版总库	2	432-433	caj下载	0.00	0.00	2013-5-31 16:18:45
210.28.96.74	高职院校对农村富余劳动力转移培训的作用及模式	中国学术期刊 网络出版总库	2	432-433	caj下载	0.00	0.00	2013-5-31 16:18:35
210.28.96.74	高职院校对农村富余劳动力转移培训的作用及模式	中国学术期刊 网络出版总库	2	432-433	caj下载	0.00	0.00	2013-5-31 16:18:19

1.3、统计数据太粗放

只有ip，不知道是哪些读者在使用这些资源

有了读者的详细信息，我们可以做哪些分析工作？

- 1.大学生、研究生、博士生谁更爱学习
- 2.讲师、副教授、教授，谁更有研究精神
- 3.年轻的教师爱专研，还是年长的教授爱专研



1.4、数据库采购决策支持

试用数据库，缺乏目标读者的使用反馈，只能根据厂商反馈的访问量、抽样调查，问问身边读者的反馈。

例如新东方，反响最好的是一位财务处的老师，给子女学习用的。

数据库购买后，缺乏使用情况的详细日志，无法分析使用情况。

哪些读者在使用？学生or老师？

哪个年级的学生在使用？大一or大四？



1.5、无法提供个性化服务

例如学术馆员，与所服务的教授、专家缺乏共同语言，不知当前其研究课题，更不了解其查阅的资料。

如何为系部教师提供服务呢？



远文读者访问统计网关

1. 可以统计文章篇名、读者帐号
2. 根据读者帐号，从汇文系统，导出读者详细信息

pdf. d. cnki. net	2013-4-25	我国高等院校旅游管理专业教学模式创新研究_曾琳. pdf	114. 222. 145. 151	109F013143	校外	国际教育学院	旅管1015	杨XX	M
pdf. d. cnki. net	2013-4-25	高职院校旅游管理专业的实践教学_周义龙. pdf	114. 222. 145. 151	109F013143	校外	国际教育学院	旅管1015	杨XX	M
pdf. d. cnki. net	2013-4-25	农林院校旅游管理专业建设的几点思考_卢宏亮. pdf	114. 222. 145. 151	109F013143	校外	国际教育学院	旅管1015	杨XX	M
pdf. d. cnki. net	2013-4-25	关于旅游管理本科专业课程体系建设的再思考_蒋长春. pdf	114. 222. 145. 151	109F013143	校外	国际教育学院	旅管1015	杨XX	M
pdf. d. cnki. net	2013-4-25	高职院校旅游管理专业学生就业情况分析_刘冰. pdf	114. 222. 145. 151	109F013143	校外	国际教育学院	旅管1015	杨XX	M
pdf. d. cnki. net	2013-4-25	冲击载荷对飞机起落架螺纹连接的影响_缪宏. pdf	210. 28. 98. 22	210. 28. 98. 22	校内	-	-	-	-
pdf. d. cnki. net	2013-4-25	面向微机械应用的CVD金刚石微齿轮的制备_王红军. pdf	210. 28. 98. 22	210. 28. 98. 22	校内	-	-	-	-



二、技术细节

原理简介

其他方案对比

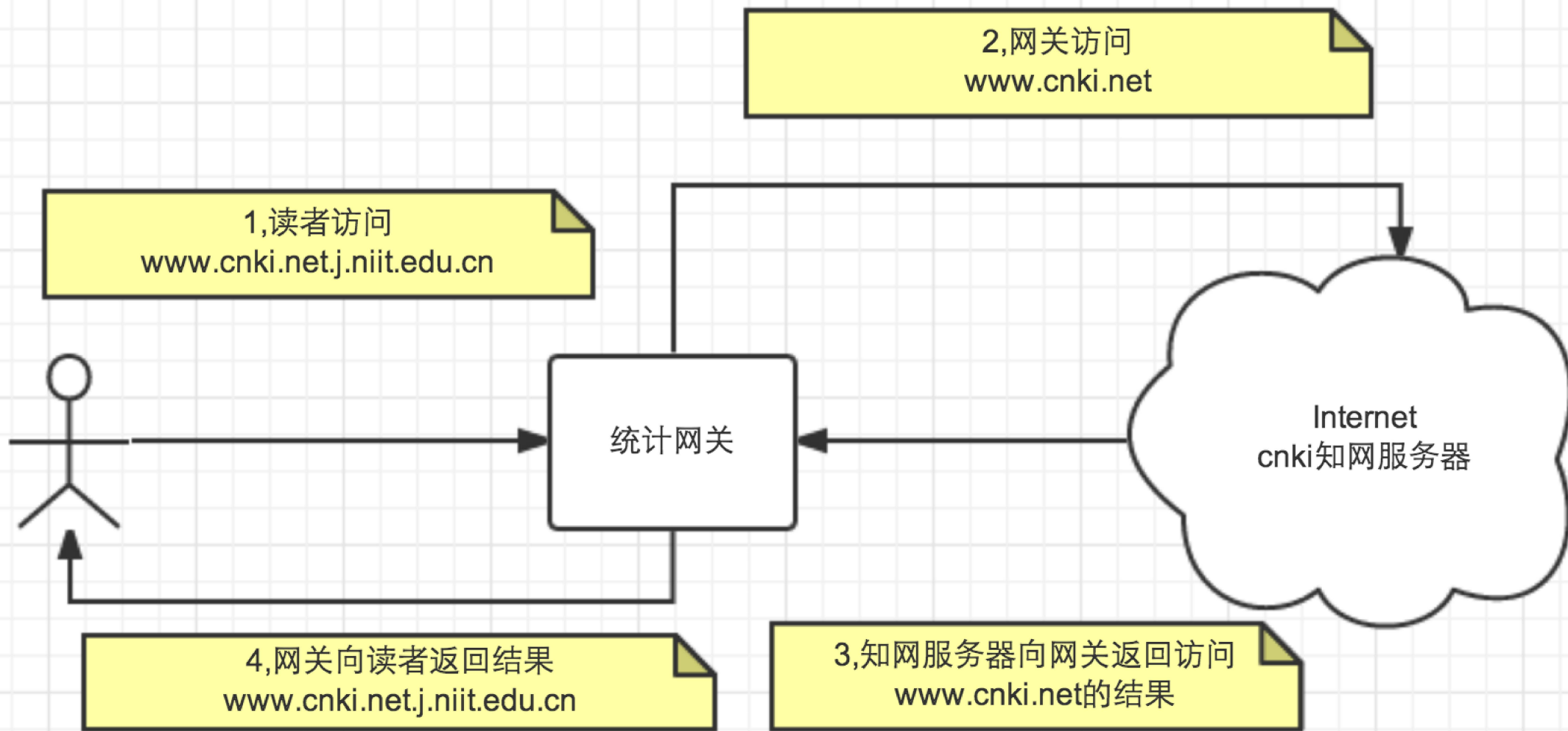
性能问题（简单集群、LVS集群）

网络中心只开放端口，没有外网ip地址

身份认证集成



2.1 原理简介：URL重写



2.2技术对比： 作为远程访问系统，

	统计网关	硬件vpn
新东方流媒体	Y	N
汇文帐号集成	Y	必须导入
易用性	简单，无需安装任何插件	需要安装插件，易被拦截
平板、多系统支持	完全支持	仅支持windows，不支持iPad、Surface等平板

	统计网关	iReader	EZProxy	易瑞
新东方流媒体	Y	Y	N	Y
平板、多系统	Y	N	Y	Y
汇文帐号集成	Y	必须导入	Y	必须导入

2.3技术对比： 访问统计网关系统

	统计网关	复旦某公司 (通过核心路由抓包)
性能可靠性	高，能记录所有读者访问	低，会丢包
记录读者帐号	支持	不支持
记录文章篇名	支持	不支持
日志准确性	100%，读者下载的日志必定被记录	会丢包，数据库商家不认可其统计日志



2.4.1性能

- 基于linux下apache开发，性能可靠，网易、QQ、雅虎等公司初期www服务器都是基于apache。

(性能预估)

未优化，2 G内存，i3 cpu，1000个并发

优化后，16G内存，至强4核cpu，6000个并发



2.4.2服务器叠加

- 多台服务器直接扩展，不同服务器代理不同资源。

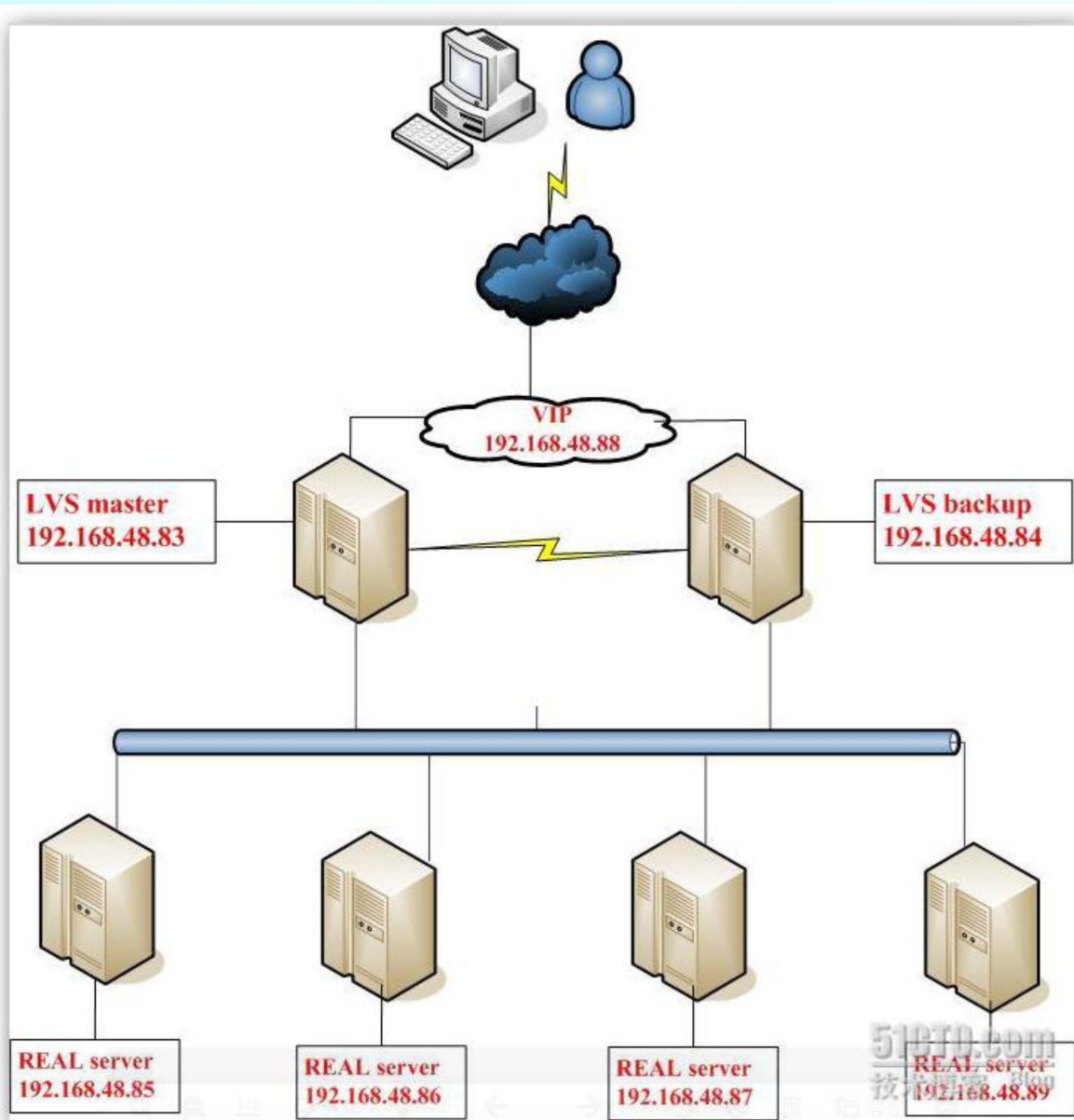
服务器一：cnki

服务器二：万方、维普



2.4.3 LVS 负载均衡

1. 设置多台统计网关服务器，由前台服务器统一调度
2. 如果性能不够，可直接增加新服务器



2.5端口转发

问：网络中心只开放端口，没有外网ip地址，可以部署吗？

答：完全可以，服务器只有内网ip，放在学校服务器后面。

无锡学院就是这样部署的，服务器真实ip为
192.168.22.188，服务端口为80；学校的出口ip为：
218.90.185.212，转发的端口为81



2.6身份认证集成

问：可以采用哪些身份认证方式？

答：可以采用任何可以想象到的认证方式。借鉴EZProxy的认证方法，提供URL外置认证，后台设置一个认证网址，发送读者帐号、密码过去，只要返回正确的结果就可用了。

理论上，这种方式可以支持LDAP、Radius、数据库等多种方式。汇文系统读者集成就是采用的外置URL认证，读者忘记密码了，到流通部修改即可！



三、我校部署经历

1. 作为校外访问代理，作为校外访问电子资源的唯一方式
2. 采用ip认证，代理部分校内流量。校内可以通过ip认证，既可以直接访问电子资源，也可以通过统计网关访问电子资源
3. 集成到校园一卡通单点登录系统
4. 寒暑假后，校内ip认证，也必须通过统计网关访问电子资源
5. 全部采用帐号登录，必须经过一卡通单点登录，或者直接输入帐号登录

总结：以上三步至少需要经过一个假期，最快一学期过度时间。



有趣的功能： 无处不在的 参考咨询



www.cnki.net.j.niit.edu.cn

cnki 中国知网 cnki.net

文献 期刊 博硕士 会议 报纸 外文文献 年鉴 百科

文献全部分类 全文 访问的是我们的代理服务器

期刊 博硕 会议 报纸 图片 年鉴 统计数据 作者投稿 采编 评价 引文 企业 医院 农业 政府
专利 成果 标准 法规 古籍 外文 工具书 学术不端 独家 优先出版 科研 法律 教育 城建

资源总库
海量中外文献，一站式检索

国际文献总库
14,000多家国外出版社的文献

行业知识服务平台

210.28.98.99读者，您好 关闭
您正在使用图书馆vpn358远程访问系统。
图书馆

点击咨询

知网动态
德古意特出版社期刊、图书加盟CNKI学术搜索
第三届"中国知网杯"全国中小学数字阅读活动开始
高校服务第三季"CNKI技术参考类资源体验月"启幕
全国高校"CNKI工具类资源体验月"活动，再添新品
商务印书馆·精品工具书数据库预发布，欢迎试用
全国少年儿童数字阅读推广月活动，火热进行中！

这里嵌入了tq的参考咨询，还可以...
中国党政系列专业期刊

仿百度维
普：定制
的谷歌学
术，下载
的论文也
会被记录



scholar.google.com.j.niit.edu.cn

1, 通过vpn358访问谷歌学术

南京工业职业技术学院图书馆

Google
学术搜索

2, 可以自定义本校的谷歌学术, 彰显图书馆无处不在的服务理念

3, 随处可见的图书馆参考咨询服务

210.26.102.22读者, 您好 关闭

您正在使用图书馆vpn358远程访问系统。

图书馆

站在巨人的肩膀上

点击咨询

定制豆瓣：仿清华豆瓣

book.douban.com.j.niit.edu.cn/subject/24153181/

豆瓣读书

书名、作者、ISBN

分类浏览 作者 阅读 购书单

1、通过代理访问豆瓣

从前我死去的家

作者: [日] 东野圭吾
出版社: 南海出版公司
原作名: むかし僕が死んだ家
译者: 李盈春
出版年: 2013-7-1
页数: 216
定价: 25.00元
装帧: 平装
丛书: 新经典文库·东野圭吾作品
ISBN: 9787544265065

想读 在读 读过 评价: ☆☆☆☆☆

写笔记 写书评 加入购书单 分享到

内容简介

★★★★☆ 7.5 (276人评价)

★★★★★	22.1%
★★★★☆	42.8%
★★★★☆	31.9%
★★★★☆	2.2%
★★★★☆	1.1%

2、插入我们自己的个性化代码

到南京工业职业技术学院图书馆检索本书

在哪儿买这本书?

亚马逊 (RMB 17.80)
京东商城 (RMB 17.80)
文轩网 (RMB 17.50)
北发图书网 (RMB 16.80)
当当网 (RMB 18.80)
China-pub (RMB 20.00)
查看8家网店价格 (RMB 16.80起)

多本比价, 批量购买 加入购书单

技术对于图书馆的出路？

图书馆的书是买的，电子资源大部分也是买的，网站是外包给公司的，技术在图书馆的定位是什么？计算机修理？接网线？

论技术，图书馆技术人，水平绝对不如商业公司！

Google开源项目：安卓、Chrome浏览器

百度开源项目：七巧板JS库、UEditor

阿里淘宝开源项目：LVS集群、Tengine高性能web服务器、
PhpWind论坛



数据驱动图书馆未来

这些公司可以开源自己的所有技术，但绝不会开放用户数据，腾讯绝不会开放QQ用户的好友关系、淘宝绝不会开放用户购买商品的信息！

故，图书馆最有价值者，不是购买的图书、不是采购的各项电子资源，只要有钱都能买到；读者在图书馆的访问数据，才是核心竞争力！



Thank You

数据驱动图书馆未来

