

虚拟多端口的短信分发系统设计与实现

蒋 鹏¹, 黄声勇²

(1. 广东技术师范学院, 广东 广州 510665;
2. 广东电网公司汕尾供电局, 广东 汕尾 516600)

摘 要:针对虚拟短信端口的下发问题,文中设计并实现了一种基于多端口的动态负载均衡和多线程策略相结合的短信分发系统。首先提出了基于中国电信 ISAG 的 MSDS 系统总体架构,针对大数据量短信分发,接着引入动态负载均衡和基于 MDB 异步并发多线程相结合的分发策略,最后给出了基于 EJB 和 Web Service 的系统实现以及详细的测试结果。通过实践结果表明,该系统是可以确保短信分发应用的高并发性、实时性和负载均衡的,并且可以很好地解决不同平台应用的问题。

关键词:排队论; $M/M/c$; 负载均衡; 实时; 多端口

中图分类号: TP302.1

文献标识码: A

文章编号: 1673-629X(2014)12-0221-05

doi: 10.3969/j.issn.1673-629X.2014.12.052

Design and Implementation of SMS Dispatch System of Virtual Multi-port

JIANG Peng¹, HUANG Sheng-yong²

(1. Guangdong Polytechnic Normal University, Guangzhou 510665, China;
2. Shanwei Power Supply Bureau of Power Grid, Shanwei 516600, China)

Abstract: Dedicated to the dispatch of SMS based on virtual port, design and realize a SMS dispatch system with dynamic load balancing based on multi-port and multithreading strategy in this paper. Firstly, present the whole structure of the MSDS based on China Telecom ISAG. Then, in view of the large amount of data message, introduce the distribution strategy combined dynamic load balancing with asynchronous concurrent multi thread based on MDB. Finally, the system implementation based on EJB and Web Service is given and a detailed test results are put forward. Through the practice, it is concluded that the system can ensure the high concurrency, real-time performance and load balancing for SMS application, and can solve the problem of application of different platform.

Key words: queuing theory; $M/M/c$; load balancing; real-time; multi-port

1 概 述

目前,分发短信(Short Message Service, SMS)的应用系统可以通过专线接入网关、Model 接入及虚拟短信端口接入三种方式接入无线网络^[1],其中 SMPP (Short Message Peer to Peer) 为短消息点对点通讯协议,例如中国电信的 SMGP 协议《短信网关通信协议》^[2],就是以国际标准 SMPP 协议为基础,结合自身特点而设计的一套企业标准。

专线接入网关的短信下发方式进行 SMS 发送的业务量高,但实现细节复杂;Model 接入方式收发速率低,实时性和并发性也比较差,整体下发性能也不

高^[3];虚拟短信端口(虚拟端口)接入方式灵活性大、业务量适中,其直接利用组件接口和网络而无需掌握复杂的底层编码细节就可以接入短信网关,这样不但便于虚拟营运商进行短信应用系统的开发,而且节省了开发时间和共享了硬件资源,所以虚拟端口接入方式成为当前企业短信分发系统的主流应用。

虽然虚拟端口成为主流的接入方式,但当前的短信系统都是基于虚拟单端口的方式,却未见虚拟多端口的接入方式,然而单端口的短信下发速率是有上限的,为了更充分地保证短信应用的大发送量和高并发量,确保像无线城市、政府网络中心以及运营商综合服

务平台等应用场景的短信实时性,文中提出了基于虚拟多端口的短信分发系统(Multi-port SMS Dispatch System,MSDS)的设计与实现解决方案。设计了 MSDS 系统的总体架构,通过引入动态负载均衡策略和多线程并发与同步控制机制,以期把大量的短信在最短的时间内进行分发,从而保证了 SMS 实时性和系统的利用率。

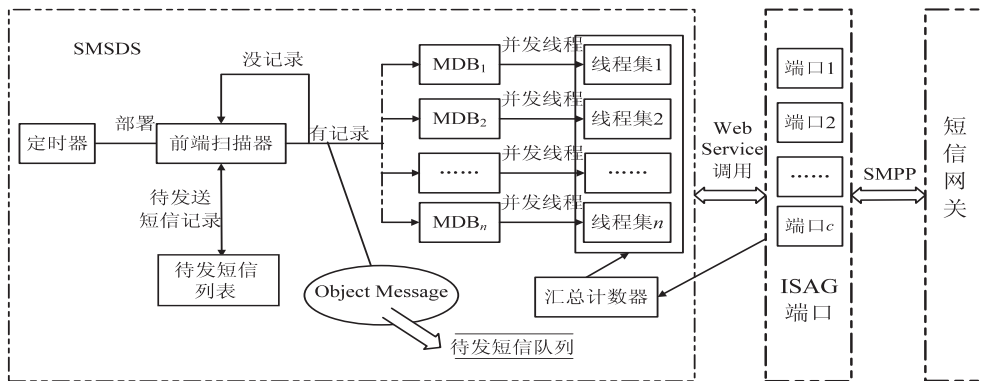


图 1 MSDS 总体设计架构

在上述的架构中,如果要保证每一条短信都是随机的和实时处理的,则需要引入多线程处理^[6]。MSDS 的处理流程是:部署好系统后,前端扫描器在系统设定好的间隔时限不断地被全局定时器触发,扫描待发短信列表,当依照指定规则扫描到有待发的短信记录,把每条记录封装成为对象消息 Object Message,并放入扫描器的待发短信队列^[7],并同时触发消息驱动 Bean (Message Driver Bean, MDB)^[8]去处理, MDB 的数目 1, 2, ..., n 代表异步并发 n 种线程集并发处理,每种线程集包含相同的处理线程,例如 MDB_1 异步并发 2 个短信实体组装线程,而 MDB_2 是异步并发 15 个提交 IS-AG 短信端口线程;最后由提交线程根据汇总计数器进行动态的负载均衡任务派遣到各个短信端口所对应的缓冲队列中,并调用 Web Service 把待发短信任务提交 ISAG 端口;一次扫描到放入待发队列中的所有任务分发完毕后,再启动下一次扫描分发,整个分发过程就是这样轮询。每个 MDB 所并发的处理线程数目可以是不同的,这样提高了系统的灵活性和扩展性,例如对于组装短信实体的线程数目可以少点,对于提交 IS-AG 端口的线程数目应该多一些,因为组装短信实体的速率非常快;而且 MDB 异步并发多线程是可以同时并发多种线程,无需等待某线程处理完后再执行另一线程,提高了系统的并发性。

3 MSDS 的短信分发策略

3.1 动态负载均衡策略

3.1.1 问题描述

在 MSDS 中,设定了两个假设:

2 MSDS 的系统总体架构

文中提出的 MSDS 的系统架构是在中国电信 IS-AG^[4] (Integrated Service Access Gateway,综合业务接入网关)的虚拟短信多端口的基础上,将短信的号码和内容封装成为指定的短信实体,通过系统端口所提供的 Web Service^[5] 接口推送到短信网关。具体系统总体架构如图 1 所示。

(1)系统的容量为无穷大,即待发短信队列可以容纳无限数的待发短信;

(2)当有待发短信到达队列时,有空闲的端口立即进行服务。

在实际应用中,队列可以容纳的数目是有限的,当有大量待发短信到达时,将导致输入消息在前端被阻塞,而产生丢包的问题;调度任务需要提供实际的判断信息给空闲的短信端口,尽可能让每个端口负载近似相同,提高端口资源的利用率。

已有研究表明^[9-11],静态负载均衡调度和动态负载均衡调度都可以提高集群系统的性能,在实际应用的负载均衡环境中要考虑延迟和计算负载耗时^[12],没有哪种算法是最优的,能达到负载效果和符合系统自身的算法才是最好的。

3.1.2 策略描述

针对以上问题,提出基于端口队长实时计算的动态负载均衡策略来解决问题^[13]。其数学模型如图 2 所示。引入一个缓冲队列 Q 和 Q 队列长度的计数器到每个端口中, Q 主要是解决 MSDS 中输入队列可能溢出的问题和及时接收来自输入队列的待发短信;用每个 Q 相对应的队列长度计数器来标记每个端口的负载情况,计数器值与端口的负载成正比,保证将任务派遣给计数器值小的端口。

具体工作原理如下:

(1)在任务启动时,将每个端口缓冲队列 Q 的计数器初始化为 0;

(2)通过输入队列按端口的顺序把待发短信推送给计数器为 0 的端口;

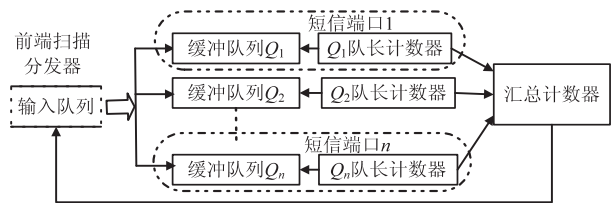


图 2 基于端口队长动态负载均衡的数学模型

(3)系统运行 $t(t > 0)$ 时间后,缓冲队列 Q 的计数器将根据接收的待发短信和端口发送的待发短信,发生变化,并把该缓冲队列的存放数目(计数器值)发送给汇总计数器,汇总计数器把汇总信息反馈给前端扫描分发器,接着输入队列根据 n 个缓冲队列的汇总信息,把待发短信发往计数器值最低的端口,使每个端口的负载近似相同。这样,输入队列在进行待发短信分发时,其不需要进行复杂的计算,只需检查汇总计数器的信息即可,降低输入队列的分发时间,实现了负载均衡,提高了端口的处理效率。

3.1.3 动态负载均衡的实现

根据 ISAG 短信端口的配置规范描述,其配置片段主要包括 Web Service 服务端 URL 地址(短信端口)、Web Service 客户端 URL 地址(MSDS)、特服号、发送短信端口接入号码以及密码等,其多端口就体现在多个 ISAG 短信端口配置片段,要体现多端口就必须使每个配置片段的 Web Service 服务端^[14]、短信端口接入号码都不同,而且短信端口接入号码可以唯一标识一个短信端口。经过短信端口配置片段的信息分析,在短信分发模块设置多个缓冲队列和短信端口配置片段一一对应,其队列的大小就是该端口当前的负载情况。

汇总计数器就是在提交短信端口时实时触发线程去读取各个队列的大小,并在某一时刻保存各个队列大小的取值。

根据汇总计数器、缓冲队列以及短信端口的对应关系,进行动态负载均衡的过程如下所示:

```
synchronized(待发短信队列){
while(待发短信队列不为空){
读取每个缓冲队列的大小;
把当前每个缓冲队列的大小存放到汇总计数器;
if(缓冲队列的大小都一样){
把短信实体封装到第一个端口并提交到该 ISAG 端口;}
else{获取缓冲队列最小所对应的端口信息;
if(几个缓冲队列的最小值都一样){
把短信实体封装到其中的第一个端口并提交到该 ISAG 端口;}
else{把短信实体封装到该端口并提交到该 ISAG 端口;}}
```

```
}
删除待发短信队列中已提交 ISAG 端口的短信对象;
}
```

3.2 多线程的设计与同步控制策略

利用有线程同步机制和调度策略来确保整个系统的运行效率,使得 SMSDSM 中同一时刻存在着多个线程在并发运行。

前端扫描器的输入队列中存放的每个待发短信消息称为单元任务(Unit Task, UT),根据短信的业务特点,在输入过程中有获取手机号码(Numbers Split Thread, NST)、组装短信实体(Package SMS Thread, PST)、提交短信端口 ISAG(Submit ISAG Thread, SIT)及检查汇总计数器(Check Total Calculator, CTC)四种线程在并发运行。线程的并发处理就必须得提供相对应的同步安全控制机制,表现在共用同一个待发短信队列中的线程要同步,同一个 UT(一条待发记录)的 NST、PST、SIT 三个线程要同步,另外,每个 UT 的 SIT 线程在提交 ISAG 时,都要执行 CTC 线程,此时由于 CTC 是实时进行同一个汇总计数器读取的,所以也要进行同步控制。线程同步控制模型如图 3 所示。

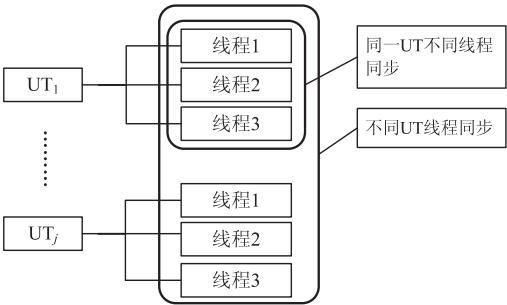


图 3 多线程的同步控制模型

采用新增变量和共享资源加锁的机制来保证线程同步,即“时间和空间”相结合的方式,解决了大量浪费空间和线程高阻塞情况的问题。不同 UT 的 NST、PST、SIT 要对同一个待发短信队列进行操作,而同一 UT 这三个线程只对该 UT 进行操作,另外 CTC 是对同一个汇总计数器进行操作,所以针对不同 UT(即同一个待发短信队列)的 NST、PST、SIT 线程同步采用同步语句块 synchronized 来维持,而同一 UT 的线程同步采用 java. lang. ThreadLocal 进行控制,另外,CTC 线程的同步也采用 synchronized 进行控制。

4 系统的实现

根据上述多线程策略的描述,MSDS 系统的 MDB 只有三个,分别是触发线程 NST、PST 及 SIT;另外根据实际的项目需要,设置该系统只有四个短信端口,所以

必须得有四个缓冲队列。利用 UML 的动态建模机制,设计出短信分时的活动图如图 4 所示。当扫到有待发记录时,把每个待发记录封装成为对象消息放入待发短信队列,并同时输入消息到三个 MDB 所对应的 JMS 消息队列中,每个 MDB 的 onMessage 方法响应并异步触发 NST、PST、SIT 三种线程,三种线程各自独立运行。一般是 NST 先把号码串分割好每一个的下发号码放入手机号码列表中,然后 PST 把每个手机号码及其对应的短信内容封装成为短信实体,放入短信实体列表中,最后 SIT 启用 CTC 线程检测汇总计数器,并根据虚拟短信端口缓冲队列的检测情况进行动态的下发任务负载均衡,同时调用 Web Service 接口提交到 IS-AG 短信端口,整个 SMSDS 系统就是重复这样的分发过程。

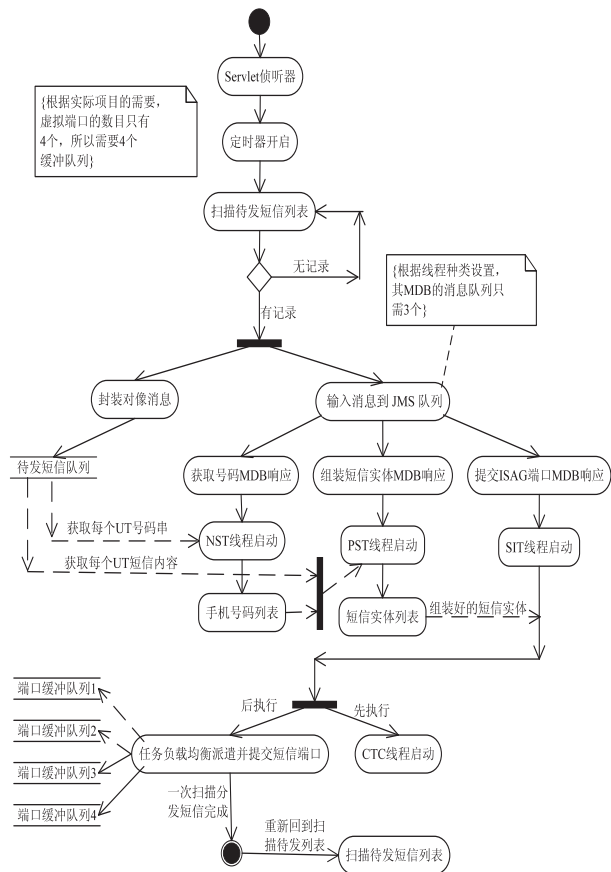


图 4 MSDS 短信分发的活动图

根据 MSDS 系统架构、系统所使用的动态负载均衡策略、多线程机制以及短信分发活动图,利用 EJB、Web Service、JNDI 及 JMS 技术实现了 MSDS。

5 系统的性能测试分析

5.1 单端口与多端口的测试比较分析

为了证明基于虚拟多端口的 MSDS 系统更具有分发短信的高效率,与单端口的短信分发系统进行测试比较。为了使实验数据具有可比性,四端口的 SMSDS

和单端口短信系统分别独立地在同样硬件环境和软件系统环境中运行,而且是分发同一批短信(短信内容和号码个数一样)。单端口和四端口的短信分发系统部署在一主机上运行,而 ISAG 短信端口的 Web Service 服务端部署在另外的四台主机中^[15],如果是单端口的情况,只需其中一台即可。

实验进行了单端口和四端口分别发送 20、100、1 000 及 60 000 个号码的短信分发测试,记录发送耗时和发送速率。为了更直观清晰地表达出发送短信数量与单端口或四端口之间的发送速率的关系,给出图 5,标识在上方的数据例如 0.97、1.8、6.12 及 19.77 分别是四端口的 SMSDS 系统发送 20、100、1 000 及 60 000 个手机号码的短信发送速率。

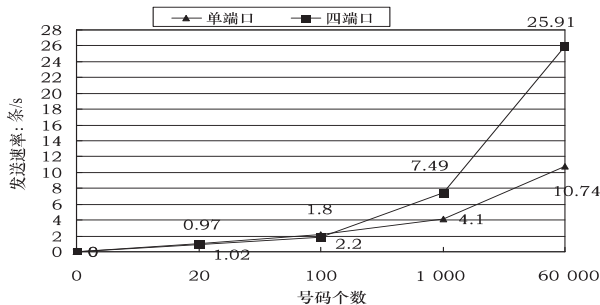


图 5 单端口与四端口的发送速率比较

根据上述实验数据的记录情况,四端口的 SMSDS 比起单端口短信分发系统,在发送号码个数少于大概 300 时,发送速率(条/s)稍微低一点,例如在发送 100 条短信时,单端口是 2.2 条/s,而四端口是 1.8 条/s;但是在发送量超过 300 条时,四端口比起单端口在发送速率方面具有明显的优势,例如发送 60 000 条短信时,四端口的发送速率是 25.91,而单端口是 10.74,可以达到 2 倍以上的速率优势。原因是四端口系统所进行的动态负载均衡计算和派遣需要耗时,在单端口系统下,线程都在排队等待一个短信端口进行服务,当发送量大时,线程都在排队阻塞;而四端口系统有四个短信端口进行任务均衡服务,线程的排队顾客等待接受服务的时间短一些。所以在高发送量和并发量的情况下,多端口的短信分发调度系统更能保证短信应用的实时性。

5.2 多端口的动态负载均衡效果测试

在上述进行四端口的 MSDS 测试时,记录了四次发送时四个服务台(四台 Web Service 服务端的受理机)发送短信数目的情况,获取每个端口在发送短信总数为 20、100、1 000 及 60 000 时所完成发送短信的数目,如图 6 所示,从而反映了短信分发时四个端口的负载均衡情况^[16]。

从上述的负载均衡情况来看,以发送 60 000 条短信为例,四个短信端口分别完成发送短信的数目为

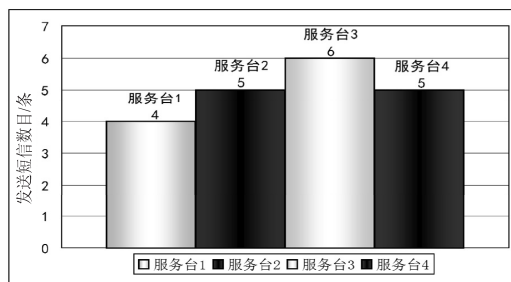
15 007、14 965、14 987 以及 15 041,则每个端口所被分派的发送短信数目都接近平均数 15 000,可以得出四个端口接收到的任务负载比较均衡,即短信发送的任务负载能够被动态地均衡调度到各个端口去处理,这说明了系统使用的动态负载均衡调度策略起到了实际效果,提高了系统的整体利用率和并发处理能力。

衡计算和任务派遣,需要极短的耗时,所以对于发送少量(小于 300 条)的短信,采用该系统并未取得明显的分发速率优势,但对于大发送量的短信分发具有明显的分发速率优势。

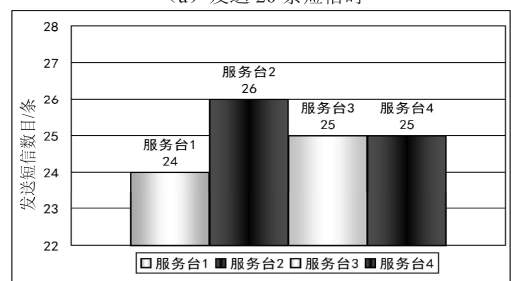
文中设计的 MSDS 方案在企业短信系统中具有一定的实际指导意义;另外,文中实际采用 J2EE 架构,利用 EJB 组件规范中的 MDB 来支持消息队列,可以很好地解决不同平台应用的问题。

参考文献:

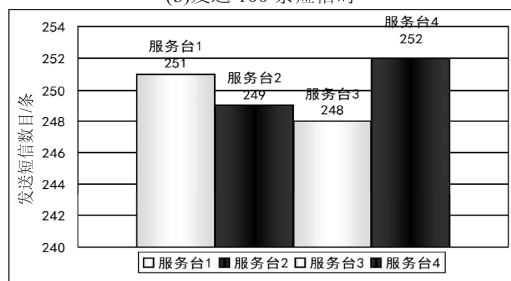
- [1] 中国移动通信公司. 中国移动通信信息资源站实体与互联网短消息网关接口协议[S]. 2001.
- [2] 中国电信集团公司企业标准. 短消息网关(SMG3.0)协议[S]. 2005.
- [3] 沈 斌,李兴国,钟金宏,等. 基于多队列和多线程的短信实时并发控制算法[J]. 计算机工程,2008,34(8):62-65.
- [4] 中国电信集团公司. 中国电信综合业务接入网关(ISAG)开放接口技术规范 RC V1.1.0[S]. 2008.
- [5] 饶 元,冯博琴. 新网络体系结构—Web Services 研究综述[J]. 计算机科学,2004,31(5):1-4.
- [6] Wang Zhentao, Guo Ziyang, Wang Zhenqi. Analysis of multi-task scheduling based on SMS[C]//Proc of 2008 international conference on computer science and software engineering. Wuhan:IEEE,2008:1086-1089.
- [7] 陈申燕,曹 旻. 多层关联规则挖掘算法的研究及应用[J]. 计算机工程与设计,2010,31(4):885-888.
- [8] 栗松涛. Java EJB 应用程序设计[M]. 北京:机械工业出版社,2008.
- [9] 曾东海,刘 海,金士尧. 集群负载调度算法性能评价[J]. 计算机工程,2006,32(11):78-79.
- [10] 金 杉,李 秋. 基于多目标遗传算法的动态负载均衡方案[J]. 计算机工程与科学,2013,35(12):102-106.
- [11] 李长志. 集群服务系统负载均衡原理的分析与实现[J]. 重庆邮电大学学报:自然科学版,2004,16(6):84-87.
- [12] Meng Qingyang, Qiao Jianzhong, Liu Jun, et al. A dynamic load balancing method based on stability analysis[C]//Proc of 2008 international symposium on computer science and computational technology. Shanghai:IEEE,2008:404-408.
- [13] 李程程,张永胜,李 静,等. 一种简单的 Web 服务安全通信模型研究[J]. 计算机技术与发展,2010,20(9):157-160.
- [14] Microsoft, IBM, VeriSign[EB/OL]. 2008. <http://www.ibm.com/developerworks/webservices/library/WS-Security/>.
- [15] W3C Recommendation. XML encryption syntax and processing[EB/OL]. 2002. <http://www.w3.org/TR/2002/REC-XMLe-enc-20021210/>.
- [16] 郭 静,祝永志,王延玲. 基于 MPI 的动态负载均衡算法的研究[J]. 计算机技术与发展,2009,19(5):150-153.



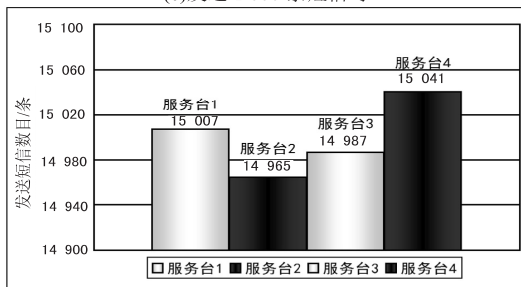
(a) 发送 20 条短信时



(b) 发送 100 条短信时



(c) 发送 1 000 条短信时



(d) 发送 60 000 条短信时

图 6 四服务台的负载均衡情况

6 结束语

文中提出将利用动态负载均衡策略和多线程机制引入到短信多端口的 MSDS 中,可以为短信应用系统进行大发送量和高并发量的短信发送需求提供实际保障,确保短信应用的实时性。当然由于文中的 MSDS 系统在进行短信分发时,要进行基于多端口的负载均

虚拟多端口的短信分发系统设计与实现

作者：[蒋鹏](#)，[黄声勇](#)，[JIANG Peng](#)，[HUANG Sheng-yong](#)
作者单位：[蒋鹏, JIANG Peng\(广东技术师范学院, 广东 广州, 510665\)](#)，[黄声勇, HUANG Sheng-yong\(广东电网公司汕尾供电局, 广东 汕尾, 516600\)](#)
刊名：[计算机技术与发展](#)[ISTIC](#)
英文刊名：[Computer Technology and Development](#)
年，卷(期)：2014(12)

引用本文格式：[蒋鹏. 黄声勇. JIANG Peng. HUANG Sheng-yong 虚拟多端口的短信分发系统设计与实现\[期刊论文\]-计算机技术与发展 2014\(12\)](#)