

文章编号:1001-9081(2009)07-1974-04

基于 Linux 的内存式 WebGIS 设计与实现

周狄波^{1,2,3,4,5}, 王迪峰^{1,2,3}

(1. 卫星海洋环境动力学国家重点实验室, 杭州 310012; 2. 中国科学院 上海技术物理研究所, 上海 200083;

3. 国家海洋局 第二海洋研究所, 杭州 310012; 4. 中国科学院 研究生院, 北京 100049;

5. 杭州市信息化办公室, 杭州 310026)

(zhoudibo@sina.com)

摘要:从计算机存储系统的角度,提出以内存取代硬盘作为运行系统的存储介质,解决磁盘 I/O 瓶颈,以提升 WebGIS 响应速度的技术思路。通过分析 Linux 操作系统上 ram disk、ramfs、tmpfs 三种内存盘技术和 Linux 操作系统的组成,针对内存的易失性问题,给出了基于 tmpfs 和 initrd 的内存式 WebGIS 实现方法,并详细阐述了内存式 WebGIS 的系统组成、系统框架和内存规划。在此基础上,采用 64 位 Debian GNU/Linux 操作系统和 MapServer WebGIS 平台,从构建硬盘 WebGIS 母系统、构建内存系统镜像文件、重构 initrd、内存分配、系统引导和 WebGIS 应用更新六个方面对内存式 WebGIS 的实现进行了阐述。实际测试和应用结果表明,应用内存存储提升 WebGIS 应用响应速度是正确可行的。

关键词:操作系统;内存;万维网地理信息系统;存储;性能

中图分类号: TP39 **文献标志码:** A

Design and implementation of in-memory WebGIS based on Linux

ZHOU Di-bo^{1,2,3,4,5}, WANG Di-feng^{1,2,3}

(1. State Key Laboratory of Satellite Ocean Environment Dynamics, Hangzhou Zhejiang 310012, China;

2. Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China;

3. The second Institute of Oceanography, State Oceanic Administration, Zhejiang Hangzhou 310012, China;

4. Graduate University of Chinese Academy of Sciences, Beijing 100049, China;

5. Hangzhou Municipal Informatization Office, Hangzhou Zhejiang 310026, China)

Abstract: From the perspective of computer storage systems, a new idea to speed up WebGIS response was presented, which used memory as data storage instead of normal hard disk to solve the I/O bottleneck problem of hard disk. To solve the volatile problem of RAM and develop in-memory WebGIS, a solution based on tmpfs and initrd was offered after analyzing three types of ramdisk techniques (ram disk, ramfs, tmpfs) in Linux and the composition of Linux operation system. The system composition, framework and memory usage strategy of the in-memory WebGIS were detailed. On this basis, based on 64-bit Debian GNU/Linux and MapServer WebGIS platform, the implementation of an in-Memory WebGIS application was described in six aspects: buliding WebGIS application on hard disk, building in-memory WebGIS application mirror image file, reconstructing initrd, planning memory, loading system and updating WebGIS application. The test result shows that using memory storage is an effective and proper way to speed up the response of a WebGIS application.

Key words: operating system; memory; WebGIS; storage; performance

0 引言

Internet 技术的快速发展促成了 Web 应用的迅速崛起,基于 Web 的地理信息系统 WebGIS 得到了广泛应用。凭借利用 Web 进行信息发布、数据共享、在线查询、交流协作的优势,WebGIS 将 GIS 通过 Web 推送到全球用户面前,不仅拓宽了地理信息资源的应用领域,也为基于 GIS 的信息整合提供了可能。随着“Google 地图”、“雅虎地图”、“微软地图”、“图行天下”等应用的纷纷涌现,WebGIS 用户与日俱增,对系统响应速度的要求也日益提高,为此,出现了从优化程序设计到采用系统集群,从服务器端解决方案到基于 AJAX 的服务器、客户机均衡计算等众多解决方案,如何提升 WebGIS 响应速度成为目前的研究热点之一。

本文从计算机存储器系统的角度提出了一种利用内存相对于硬盘的高速存取特性,以内存为存储介质,解决磁盘 I/O 瓶颈,提升 WebGIS 响应速度的技术思路,并基于 64 位的

Debian GNU/Linux 操作系统和 MapServer WebGIS 平台在配备了 8G 内存的服务器上进行了实践。结果表明,性能较之当前普遍应用的基于硬盘的 WebGIS 有较大提升。

1 内存存储

现代计算机系统普遍采用称为存储器层次结构的方法来组织存储器系统,通过在高速低容量和低速高容量的存储设备之间插入相对高速的存储器,并配合相关调度算法,在性能和价格方面取得折中。

根据《深入理解计算机系统》中的观点:“如果程序需要的数据存储在 CPU 寄存器中,那么在执行期间,在零个周期内就能访问到它们。如果存储在高速缓存中,需要 1~10 个周期。如果存储在主存(内存)中,需要 50~100 个周期。而如果存储在磁盘上,需要大约 20 000 000 个周期!”。随着多核处理器盛行,CPU 性能与硬盘性能差距不断扩大,磁盘 I/O 速度对系统整体速度的拖累已日益显著,迫切需要高速存储

收稿日期:2009-01-15;修回日期:2009-03-10。

作者简介:周狄波(1977-),男,浙江宁波人,系统分析师,博士研究生,主要研究方向:电磁场与微波技术、RS 与 GIS 应用;王迪峰(1978-),男,福建厦门人,助理研究员,博士研究生,主要研究方向:海洋航空遥感、遥感系统集成。

设备来解决磁盘 I/O 瓶颈问题。从访问时间看,CPU 寄存器和高速缓存速度最快,但主要为运算服务,容量有限,而内存相对于硬盘有着明显的速度优势,即使在考虑了数据调度的系统开销以后,这种优势仍然值得期待,因此,以内存作为高速存储设备来提高系统响应速度从理论上讲应该是可行的。目前,国外已有厂商开展了这方面的尝试。英国 HyperOS 公司的产品 HyperDrive 4 经 Tom's Hardware 测试,被认为是世界上最快的硬盘,其技术原理主要就是利用内存取代硬盘作为存储设备。由于采用了其他技术,HyperDrive 4 售价昂贵。

2 Linux 内存盘

在 Linux 中,允许将一部分内存当作硬盘分区使用,以提高数据存取速度,即通常所谓的内存盘机制。由于 Linux 的虚拟文件系统(Virtul File System,VFS)隐藏了各种不同硬件的具体细节,为所有设备提供了统一的接口进行数据交换,应用程序通常无需修改即可运行内存盘之上。当前 Linux 上常见的内存盘技术包括 ram disk、ramfs 和 tmpfs。其中,ram disk 是较早实现的内存盘技术,作为一个基于内存的块设备,它必须经格式化后挂载使用,且大小固定,局限性较大。ramfs 解决了 ram disk 的大小固定问题,是一个可动态改变大小的基于内存的文件系统,应用较广。源自于 ramfs 的 tmpfs 继承了 ramfs 的特性,不仅可根据文件大小自动收缩,减少对物理内存的占用,又能使用交换分区,还能对所使用的内存数量进行限制,一般被认为是 Linux 上最好的内存盘技术。tmpfs 由 Linux 内核 2.4 以上版本支持,并需在编译内核时选择“虚拟内存文件系统支持”。

3 系统设计

3.1 设计思想

tmpfs 为使用内存盘提供了系统级的实现技术。要构建内存式 WebGIS,提升系统响应速度,简单的做法是通过 tmpfs 在内存空间开辟一块区域,建立内存盘存放 WebGIS 应用。事实上,一些数据库管理系统厂商也是采取类似技术来获得高速缓存机制,提升对数据库的访问效率。但这样做的缺点也是显而易见的,应用是一个整体,WebGIS 放到了内存,操作系统仍在硬盘上,WebGIS 应用是不可能脱离操作系统调用的,这一“短板”的存在,将直接影响应用整体的效率。因此,要构建真正的内存式 WebGIS,必须把操作系统、GIS 相关数据也放到内存上。得益于芯片技术的高速发展,目前内存价格已快速下降到可承受的范围之内,仅就内存本身而言,经济上的可行性已初步具备。

3.2 技术难点

尽管内存可提供优秀的高速存取性能,但它不具备硬盘那样的断电持久保存数据能力,易失性问题给系统设计带来了巨大的挑战。内存本身并不支持在系统冷启时取代硬盘进行系统引导,因为此时它上面并没有操作系统存在。同样的问题在系统重启或断电后,一样存在。尽管类 UNIX 的设计思想和内存保护模式,使得 Linux 在系统稳定性和免重启维护方面相当优秀,可对于实际系统的运行而言,停机检修不可避免。HyperDrive 4 采取硬件存储技术来解决内存的易失性问题,但价格昂贵,阻碍了应用的普及。探索一条廉价可行的技术路线成为本文研究的关键。

3.3 解决方案

严格意义上讲,Linux 只是一个操作系统内核,而非一个

操作系统,需要在内核之上,增加实用工具、shell 和文件结构等内容才构成 Linux 操作系统(图1)。由于大部分的这些工具和内容都来源于 GNU 项目,而内核又是统一的,Linux 操作系统的开放性、层次化和模块化特征非常明显,在运作机制上也比较一致。目前基于 Linux 2.6 内核开发的 Debian、ubuntu、Fedora 等 Linux 发行版虽然有着不一样的具体实现,但从系统开机到挂载文件系统、完成启动、提供系统服务的核心流程基本还是一致的。

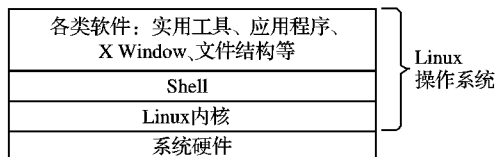


图1 Linux 操作系统构成

系统加电后,首先找到启动引导器(如 Grub、LILO 等),由启动引导器加载内核,内核在内存中制造一个 rootfs 作为临时空间,并将 initrd 作为一个系统挂载到 rootfs 上激活,然后执行 initrd 中的 init 文件,进行主要文件系统挂载、初始设备文件建立、设备探测、模块加载、真实文件系统挂载并切换等操作,最后执行实体操作系统中的/sbin/init,完成后继续启动,进入服务提供状态(图2)。

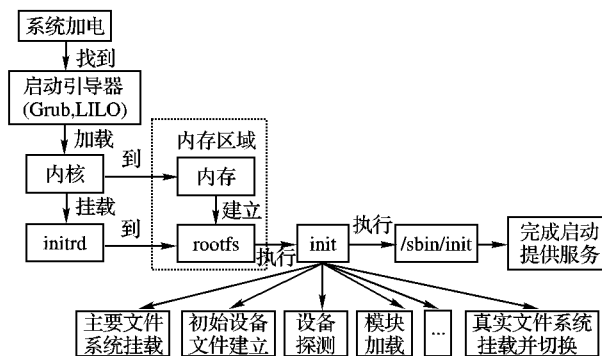


图2 Linux 操作系统开机运行机制

不难发现,在 Linux 操作系统的启动过程中,init 文件起了关键作用,它实现了从基于内存的 rootfs 向实体操作系统的切换,是从 initrd 挂载到切入实体操作系统,执行/sbin/init 之前程序流程控制的中心。因此,可以通过重写 initrd 的 init 文件和其他相关文件的办法,在 init 文件的执行过程中,建立基于 tmpfs 的内存盘,将其挂载为真实文件系统的根目录,并把硬盘上的 WebGIS 应用和操作系统一起复制到内存盘,使系统完全基于内存运行。由于 init 文件本身是 initrd 的组成部分,运行时所需的模块、工具、文件系统、设备目录等环境要素主要由 initrd 提供,为此需重构 initrd,建立符合 init 文件执行需要的环境。

3.4 系统组成

整个内存式 WebGIS 由操作系统、WebGIS 平台、Web 服务器和相关应用数据组成。操作系统平台选用了 Debian GNU/Linux。Debian 是迄今为止最遵循 GNU 规范的 Linux 操作系统,除了 Linux 内核组件,系统的其余部分主要由 GNU 工程编写和提供的程序组成,层次清晰,接口标准,稳定性高,开放性好,便于从底层进行扩展和定制。

WebGIS 平台选用了开源软件 MapServer。MapServer 由美国明尼苏达大学开发,并得到 NASA 支持,可运行于 Linux, Unix, Windows, Mac OS X 等多种平台之上。MapServer 支持 OGC(Open Geospatial Consortium)标准,包括 Web Map Service

(WMS) 和 Web Feature Service (WFS), 能通过 PostGIS 与 PostgreSQL 数据库集成, 通过 MyGIS 与 MySQL 数据库集成, 通过 OGR 和 GDAL 库访问众多的矢量和栅格空间数据, 并能以 PHP, Python, Perl, Ruby, Java, C# 等多种语言进行应用开发和定制。根据加拿大 Refractions Research 公司 Brock Anderson 的测试, MapServer 在 WMS 的系统速度和 OGC 标准兼容性方面超过了业界巨头 ESRI 的 ArcIMS 地图发布服务器。

Web 服务器采用了 Linux 下久负盛名的开源 Web 服务器 Apache, 稳定、强大、安全、可扩展性好。

3.5 系统框架

系统框架按流行的三层架构设计, 分为数据服务层、业务逻辑层和应用表现层(图 3)。业务逻辑层和数据服务层均承载于基于内存存储的 Debian GNU/Linux 操作系统环境之上。当应用表现层的 Web 浏览器发出地图服务请求后, Apache Web 服务器接收请求, 转发给 MapServer 处理。MapServer 支持以 CGI(FastCGI) 或 MapScript 模式运行, 收到服务请求后, 根据地图配置文件 Mapfile(一个用于 MapServer 应用的结构化文本配置文件, 定义了地图的区域、图层, 投影、符号、数据源、地图影像输出位置和地图显示外观等内容)和服务端端的业务逻辑, 通过数据访问接口, 读取地理空间数据和属性数据, 执行相关查询、统计操作, 将抽取到的地理空间信息通过 GD Library 转化成 png(Portable Network Graphics)格式的地图图像, 连同其他服务结果, 以 mapfile 中定义的模板文件为样本, 制作成 html 格式的地图服务结果页面, 由 Apache Web 服务器将页面返回给 Web 浏览器。

图 3 中的 GD Library、zlib、libpng 和 FreeType 是本系统中 MapServer 正常工作的必要组件。其中 GD Library 是一个开放源码的图形图像例程库, 它向上为 MapServer 提供调用服务, 向下调用 zlib 进行图像压缩, 调用 libpng 生成 png 图像, 调用 FreeType 生成字体, 屏蔽了具体的实现细节。同时 libpng 库本身也需要调用 zlib。

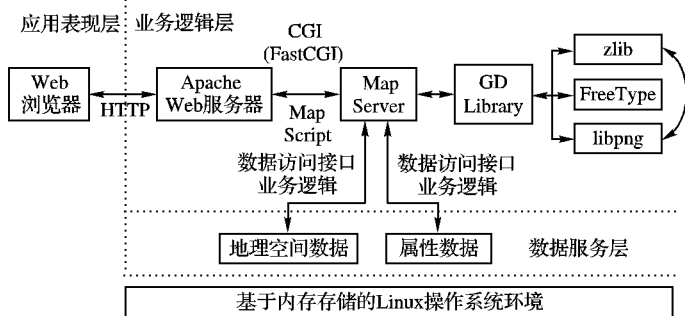


图 3 系统框架

3.6 内存规划

研究将服务器的物理内存划分为系统存储内存和系统应用内存, 系统存储内存用作 tmpfs 内存盘, 系统应用内存用于系统正常应用。系统存储内存和系统应用内存的具体容量分割根据服务器物理内存配置和应用需求决定。

4 系统实现与测试

4.1 构建硬盘 WebGIS 母系统

在服务器硬盘上构建 Debian 基本系统, 根据“开放最少的服务, 实现最高的效率和最大的安全”原则, 通过剪裁 Linux 操作系统中不必要的服务, 保持系统轻量, 实现提高系统稳定

性和可维护性, 减少停机、重启概率的目标。按 Apache Web 服务器、MapServer 地图发布平台、地图数据、地图应用的顺序安装、配置相应软件, 确保运行正常。因 Debian 4.0 apt 源中的 MapServer 版本过老, 采取了下载 MapServer 及相关组件源代码, 自行在 Debian 平台上编译的办法构建 WebGIS 平台, 有关软件版本如表 1 所示。

表 1 软件版本

软件	版本
Debian GNU/Linux 4.0	64 位(内核版本: 2.6.18-6-amd64, 默认支持 tmpfs)
zlib	1.2.3
FreeType	2.3.6
libpng	1.2.29
Gd Library	2.0.35
MapServer	5.0.3

鉴于软件之间的依赖性关系, 采取的编译安装顺序为 zlib、FreeType、libpng、Gd Library、MapServer。同时, 需指定 64 位平台上编译的参数 CFLAGS = “-O3 -fPIC”。

4.2 构建内存系统镜像文件

建立母系统的备份系统, 修改其中 fstab 文件对根目录挂载点的描述, 使之符合内存系统的需要。以 tar 文件格式打包备份系统为内存系统的镜像文件, 并删除备份系统。

4.3 重构 initrd

Debian 4.0 的 initrd 采用 cpio 格式, 为此, 以 gunzip 和 cpio 指令解压 initrd, 在生成的文件和目录中为硬盘文件系统的临时挂载点建立临时目录; 修改其中的 init 和相关文件, 去掉对硬盘文件系统正常挂载部分的代码, 编写指令, 执行如下操作:

- 1) 挂载硬盘文件系统到临时目录上;
- 2) 建立基于 tmpfs 的内存盘, 将其挂载为真实文件系统的根目录;
- 3) 搜索临时目录上的内存系统镜像文件, 将其释放到 tmpfs 内存盘上;
- 4) 卸载临时目录;
- 5) 执行内存盘上的 /sbin/init 进行后续的正常启动。

以上修改完成后, 使用 find、cpio 和 gzip 指令生成新的 initrd, 命名为 initrd.toram。

4.4 内存分配

本项研究的服务器内存配置为 8 GB 容量, 操作系统、WebGIS 应用和相应数据的合计数数据量超出 3 GB, 考虑到系统日志文件、临时文件等的开销和保险系数, 设定 tmpfs 内存盘的最大可用内存为 5 GB。

4.5 系统引导

修改启动引导器 Grub 的菜单配置文件, 增加启动到内存的系统引导项, 并将其中的 initrd 参数指向重构生成的 initrd.toram, 以便引导内存系统, 母系统仍通过原有启动项引导。

4.6 WebGIS 应用更新

WebGIS 应用的更新通过母系统实现, 母系统更新完成后按构建内存系统镜像文件的步骤重新打包, 更新运行于 tmpfs 内存盘之上的 WebGIS 应用镜像文件。

4.7 测试

系统测试采用了业界著名的文件系统 benchmark 工具

iozone,通过全面测试比较 tmpfs 内存盘与系统硬盘的数据存取性能。数据表明,基于 tmpfs 的内存盘以明显优势胜出,特别是在“Write、Re-write、Random Write、Fwrite、Frewrite”等执行写操作的方面,而以 time 指令跟踪 iozone 的测试时间,发现在 tmpfs 内存盘上完成所有测试所需的时间仅为硬盘的 1/16,速度优势十分明显。在 tmpfs 内存盘上构筑的 WebGIS 应用经杭州市政务信息资源共享及业务协同项目实际运行测试,证明在并发用户较多、数据读写频繁的情况下,较之普通基于硬盘的 WebGIS 在响应速度和用户体验上有明显提升。可见,应用内存存储解决磁盘 I/O 瓶颈,提升 WebGIS 的响应速度,在理论和实践上都是正确可行的。由于三层架构支持多层分布式应用,业务逻辑层和数据服务层可分别部署于基于内存存储的不同服务器之上,因此,系统不仅扩展性强,而且可支持海量数据应用。debian 最高可支持 64 TiB 的物理内存,存储容量已不再是内存存储应用的“短板”。

5 结语

磁盘 I/O 瓶颈是目前系统应用面临的普遍问题,在分布式应用不断扩展,中央处理器与硬盘性能差距日益扩大,内存价格大幅下降的今天,研发廉价可行的高速内存存储应用技术,具有现实意义。本文提出的内存式 WebGIS 解决方案不仅适用于 GIS 领域,也可推广、应用到其他领域,所不同的只是专业平台(如地图服务器)的变化。鉴于内存的易失性,稳定持续的电力供给十分重要,本研究在实现中采用 UPS 保证电力供应的稳定,在推广应用到其他领域时,也应对系统的电力供应给予足够重视。

参考文献:

[1] BRYANT R E, O'HALLRON D. 深入理解计算机系统[M]. 龚奕

利,雷迎春,译. 修订版. 北京:中国电力出版社,2004.

- [2] SCHMID P. HyperDrive 4 Redefines Solid State Storage: HyperDrive 4-the fastest hard disk in the world[EB/OL]. [2008-11-16]. <http://www.tomshardware.com/reviews/hyperdrive-4-redefines-solid-state-storage>, 1719. html.
- [3] 邱世华. Linux 操作系统之奥秘[M]. 北京:电子工业出版社,2008.
- [4] KROPLA B. Beginning MapServer: Open source GIS development [M]. Berkeley: Apress, Inc, 2005.
- [5] MCCARTY B. Learning debian GNU / Linux [M]. Sebastopol: O'Reilly, 1999.
- [6] 李大治. Linux2.6 内核的 Initrd 机制解析[EB/OL]. [2008-11-16]. <http://www.ibm.com/developerworks/cn/Linux/l-k26initrd/index.html>.
- [7] MCKENNA J, FAWCETT D, BUTLER H. An introduction to MapServer[EB/OL]. [2008-11-16]. <http://www.mapserver.org/introduction.html#mapserver-overview>.
- [8] ANDERSON B. A comparison of ArcIMS to MapServer[EB/OL]. [2008-11-16]. <http://www.refractions.net/expertise/whitepapers/mapserver-arcims/mapserver-vs-arcims.pdf>.
- [9] LANDLEY R. Introducing initramfs, a new model for initial RAM disks[EB/OL]. [2008-11-16]. <http://www.linuxdevices.com/articles/AT4017834659.html>.
- [10] MapServer 网站[EB/OL]. [2008-11-16]. <http://mapserver.org/>.
- [11] LANDLEY R. ramfs-rootfs-initramfs. txt[EB/OL]. [2008-11-16]. <http://www.kernel.org/doc/Documentation/filesystems/ramfs-rootfs-initramfs.txt>.
- [12] ROHLAND C, DICKINS H. tmpfs. txt[EB/OL]. [2008-11-16]. <http://www.kernel.org/doc/Documentation/filesystems/tmpfs.txt>.

(上接第 1965 页)

经过准确调整后的模型可以用来模拟发动机在瞬态工况下的放热性能,为发动机热管理系统的仿真及冷却系统零部件的优化提供可靠的热边界条件。

5 结语

依据四质量调节原理,发动机热管理瞬态模型可以有效地模拟发动机冷却水以及润滑油温度在发动机瞬态工况下的变化情况,为发动机及整车热管理系统提供可靠的放热边界条件,综上所述可得如下结论:

1) 基于一维流动传热理论建立的发动机热管理模型可以有效地模拟发动机向冷却介质释放的热量,从而进行整个热管理系统的仿真、匹配和优化;

2) 发动机瞬态模型的建立,使其传热过程更接近实际车辆行驶状态,更具实用性,对于某种机型,该瞬态模型需要专门的瞬态实验数据进行校核,校核正确后,该发动机的瞬态模型就可以用于该种机型的整个热管理系统中,从而进行整个系统的模拟;

3) 发动机热管理瞬态模型中特性参数的调整过程是一种模型设置方法,本文主要介绍了这种方法的运用过程,该过程提出了发动机热瞬态过程的一种开发思想;

4) KULI 软件采用发动机四质量块模型的模式能够较准确的模拟发动机的瞬态放热特性。通过以实验为依据的瞬态模型参数设置,其精度能够满足要求,可靠性得到保证;

5) 发动机热管理的模拟需要大量的实验数据进行校核,这就需要企业前期投入很大精力进行实验数据的收集与整理,但这也是发动机热管理发展的一个必然趋势,目前国内通用、一汽等几家大型汽车公司已经开始进行热管理相关方面研发的投入。

参考文献:

- [1] MAHMOUD K G, LOIBNER E, KRAMMER J. Integrated 1-d tools for modeling vehicle thermal management system [EB/OL]. [2008-10-25]. <http://www.sae.org/technical/papers/2004-01-3406>.
- [2] KITANOSKI F, PUNTIGAM P, KOZEK M, *et al.* An engine heat transfer model for comprehensive thermal simulations [EB/OL]. [2008-10-15] <http://www.sae.org/technical/papers/2006-01-0882>.
- [3] PUNTIGAM W, ALMBAUER R, HAGER J. Transient co-simulation of compressive vehicle models by time dependent coupling [EB/OL]. [2008-10-15]. <http://www.sae.org/technical/papers/2006-01-1604>.
- [4] CHOUKROUN A, CHANFREAU M. Automatic control of electronic actuators for an optimized engine cooling thermal management [EB/OL]. [2008-10-25]. <http://www.sae.org/technical/papers/2001-01-1758>.
- [5] GINGER M. Optimierung eines fahrzeugkühlsystems im instationar-en betrieb aufgrund der anforderungen des thermomanagements [R]. Frankfurt am Main: Forschungsvereinigung Verbrennungskraftmaschinen, 2004.