

关于银河麒麟操作系统的说明

国防科技大学计算机学院银河麒麟项目组

2006 年 2 月

最近,本论坛中有不少网友对银河麒麟的知识产权问题提出了一些质疑,特别是提到了银河麒麟和 Linux、FreeBSD 的关系问题。首先,银河麒麟开发组感谢大家对银河麒麟的关注和关心,并就此机会对一些相关问题做出说明。

一、银河麒麟操作系统的发展历史

银河麒麟操作系统采用的是一种层次式的内核结构,该结构介于单一模块内核结构(如 Linux)和微内核结构(如 CMU 大学的 Mach)之间。这种层次式结构从逻辑上来看,主要是由具有 Mach 风格的基本内核层、具有 BSD 风格的系统服务层和具有 Windows 界面风格的桌面环境组成,前两层在核态运行。在系统服务层中,银河麒麟操作系统实现了支持 POSIX 接口和 LSB(Linux Standard Base)接口的 Linux 应用兼容层,从而保证 Linux 应用程序的目标代码在银河麒麟上能够直接执行。银河麒麟之所以采用这种层次式的内核结构,既有历史的原因,也有发展角度考虑的因素。

国防科大计算机学院从 1992 年开始,银河巨型机的体系结构从向量机巨型机向大规模并行计算机转变,因此银河巨型机操作系统的技术路线,也从专用的向量巨型机操作系统,转变为采用 Unix 操作系统。由于巨型机采用大规模并行的体系结构,一台巨型机通常包含成百上千个 CPU,一般分为计算节点、I/O 节点和服务节点。计算节点主要完成计算任务,功能单一,但对降低系统开销的要求十分高,最适合的操作系统是精简内核的操作系统。I/O 节点主要承担系统的输入输出工作,从操作系统角度,可以理解为一组专用的 I/O 功能服务器,最适合的操作系统是一个定制剪裁的 Unix 操作系统。服务节点提供用户上机所需的一切环境,最合适的操作系统是一个完整的 Unix 操作系统,上层软件的支持要丰富。为了适应不同节点的功能需求,当时的银河机操作系统就采用了这种层次式内核结构,基本内核层以 CMU 大学的 Mach 为原型 www.cs.cmu.edu/afs/cs/project/mach/public/www/mach.html,并结合 Open Software Foundation 的 OSF/1 操作系统内核技术,在当时 128 个 MIPS R4000 CPU 上实现了全 64 位的操作系统基本内核层,包括存储管理、并行优化调度、Mach 特有的 Port 通讯机制等。同时课题组还利用 CMU 大学提供的 BSD Server 实现了系统服务层,但在实际测试中发现当时的 BSD Server 过于简单,而且还存在大量的 Bug,只是提供了 Mach 内核和 BSD Server 对接技术,离实际应用还存在较大距离,而当时国际上 USL(Unix System Laboratory)的 UnixSVR4.2 操作系统已经产生了相当大的影响,并且具有丰富的应用软件,为此课题组与英国 Unisoft 公司合作,一起完成了 UnixSVR4.2 的移植工作,并最终自己实现了 Mach 的基本内核层与 BSD 的系统服务层对接。这样银河操作系统内核中包含基于 Mach 的基本内核和基于 BSD 的系统服务,计算节点运行基本内核,I/O 节点运行基本内核和部分 IO 功能服务,服务节点运行一个完整的操作系统。

2002 年课题组承担 863 “服务器操作系统内核”项目时,对操作系统内核采用何种技术路线也进行了深刻分析和调研,并对当时能够获得的操作系统都进行了评测。由于 Linux 内核采用 GPL 规则,极大地限制了课题组的选择,虽然当时课题组核心开发人员也到加拿大渥太华参加 Linux 的内核峰会,但在 Linux 内核技术上没有主导权,不具备在 Linux 内核中发展自己特色的实力,并且军方出于安全的需要,为了保护安全实现方面的一些特色技术,也强烈要求银河麒麟不能采用开源的形式。FreeBSD 操作系统因为配套应用不丰富、使用界面不好、对多处理器的支持能力较弱等因素,也没有选用。与此同时,课题组发现与我们银

河操作系统采用同样技术思路的 Apple 公司的 Darwin 内核技术正在不断兴起 (<http://www.opensource.apple.com>), 已成为 Mac OS X 的内核, 具有非常鲜明的技术特色, 它的基本内核基于 Mach3.0, 系统服务主要基于 FreeBSD 内核, 桌面环境是 Apple 公司自己开发的, 因为是面向桌面领域, Darwin 没有强调安全性、高可用性和实时性等特点。这种通过基本内核层和系统服务层来提供操作系统运行环境的作法, 以前 OSF/1 采用过, Windows NT 的早期实现中也曾经采用过, 微软还曾设想通过不同的系统服务层同时实现 Unix 和 Windows。此外军方的许多定制装备, 尽管不需要一个完整的复杂的操作系统环境, 但对系统的安全性评估要求很高, 对实时性的要求也很高, 采取层次式内核结构能较好地确保基本内核的高安全性和强实时性。基于以上种种因素的考虑, 银河麒麟操作系统仍旧选择了层次式内核结构这个技术路线。

二、银河麒麟操作系统的主要工作

一个操作系统版本, 并非仅仅是一个内核, 还包含大量的配套工具软件等。银河麒麟操作系统的主要研发工作, 除了层次式内核结构外, 还在 Linux 目标代码兼容、安全、性能、高可用、实时、环境友好性等方面进行了大量的研发工作。在承担的 863 课题验收时, 科技部曾组织总装武器装备论证中心 (代表军口) 和中国软件评测中心 (代表民口), 对银河麒麟进行了长达 2 个月的全面详细测试, 既发现了银河麒麟操作系统存在的一些 Bug, 也充分肯定了银河麒麟操作系统在性能、安全性、高可用性、实时性方面的工作成效。

1、基本内核层和系统服务层的改进。近年来随着硬件的不断更新换代, 原来的基本内核层已经无法满足新的硬件体系结构, 为此我们在处理器亲和的调度优化、基于 RDMA 的 Mach 的 Port 通信机制、基于对象的存储管理等方面进行了大量改造和升级。此外基于 Unix SVR4.2 的系统服务层对 SMP 对称多处理器的支持能力较差, 并由于 Unix 影响力的逐年下降, 应用支持的丰富性也在不断下降, 课题组通过评测和分析, 认为当时正在研发中的 FreeBSD 5.0 具有比 Unix SVR4.2 更好的发展势头, 特别是 SMPng 项目的开展, 为 FreeBSD 5.0 支持 SMP 对称多处理器系统奠定了良好的基础, 因此银河麒麟操作系统的系统服务层从 SVR4.2 升级到当时正在研发中的 FreeBSD 5.0。由于银河麒麟的系统服务层是基于 FreeBSD 5.0, 而内核显示出来的绝大部分信息 (dmesg) 是系统服务层向外提供的, 并且银河麒麟的根文件系统采用了 BSD 的 UFS 文件系统 (UFS 的性能不如 Ext3, 但因为 Ext3 是 GPL 规则的, 做根文件系统会违背 GPL 规则), 设备命名管理也采用了 BSD 的模式, 所以银河麒麟的对外风格类似于 BSD。

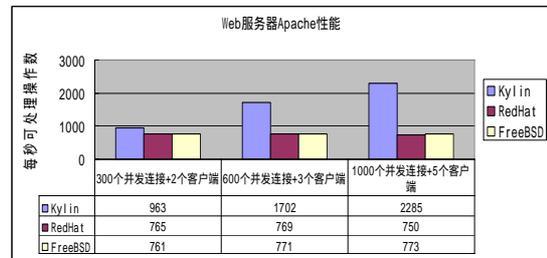
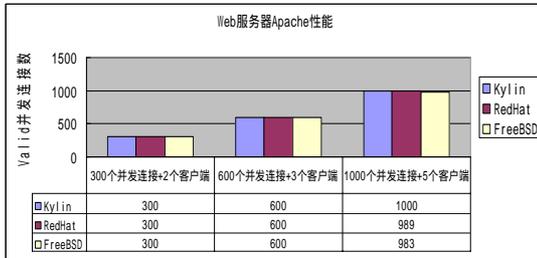
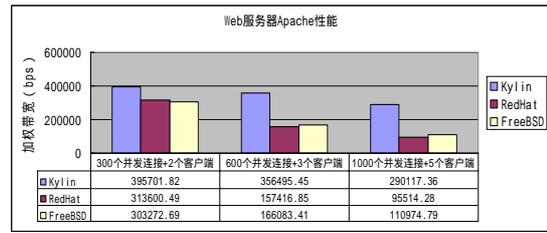
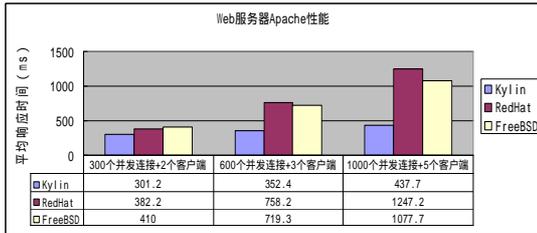
2、Linux 目标代码兼容层的实现。一个操作系统没有配套的丰富工具软件和应用软件, 是不可能市场上生存的。目前以 Linux 为代表的开源软件已经构成一个完整的软件体系, 为了充分利用 Linux 上丰富的工具软件和应用软件, 银河麒麟操作系统采用了和 Linux 目标代码兼容的技术路线, 完成了内核中的 Linux 目标代码兼容层的研制, 向用户提供了标准的 POSIX 接口, 符合 LSB 规范, 从而保证 Linux 的应用能够在银河麒麟上无缝执行。这样, 银河麒麟的核外软件, 大部分是直接采用 Linux 的核外软件。由于 Linux 版本众多, 内核发展也非常迅速, 期望全面兼容 Linux 不同发行版和不同内核版本是不现实的。为了比较全面地兼容 Linux 应用, 银河麒麟开发组设计实现的 Linux 兼容层以符合 LSB 规范和 UNIX'03 国际标准为目标, 通过了美国自由标准组织 (Open Group Standard) 的 LSB 1.3 认证, 具体详见 http://www.opengroup.org/lsb/cert/cert_prodlst.tpl?CALLER=index.tpl, 其兼容能力超过 Solaris 的 Janus、IBM 的 K42 和 Monterey 以及 FreeBSD 的 Linux 二进制兼容模式, 是国际上首个通过 LSB 认证的非 Linux 内核操作系统。至于 uname 信息显示包含 "Linux", 主要是出于支撑银河麒麟中部分 Linux 应用软件包编译的需要, 具体的解释已经在 "银河麒麟版本说明" 论坛版块中说明。

3、结构化保护级安全机制和特色安全机制的实现。针对传统操作系统的安全问题，银河麒麟操作系统实现了内核与密码机制的融合，操作系统和应用的一体化安全，进程保护、进程能力控制和进程行为控制的进程安全控制等；通过“智能卡+用户口令”的强化用户身份认证机制，保证了操作系统的身份鉴别是严密的、安全的；SAK(安全注意键)机制保证了为用户和 TCB 之间建立一条安全交互通路，防止恶意用户通过特洛伊木马之类的程序窃取用户的口令；可以区分“自身、用户、自身组、其他组、其他”等不同情形的更细粒度自主访问控制；信息访问的主体和客体的安全级进行多级划分，保证了保证高安全级信息不能转换为低安全级信息,保证低安全级不能访问到高安全级信息，防止了黑客入侵后窃取敏感信息；可信数据认证机制将可信数据/程序同不可信的数据/程序隔离，从而确保了病毒、木马等恶意程序无法执行，保证系统的机密性、完整性；能力机制确定用户能力、进程能力和文件系统能力，实现了管理员分权，在应用存在漏洞时，最大程度减少了对系统的破坏；强大的安全审计为维护系统的安全提供了有效的依据。目前银河麒麟操作系统安全版已经通过公安部计算机信息系统安全产品质量监督检验中心认证,是国内唯一通过第四级结构化保护级检验的操作系统；通过中国人民解放军军信息安全测评认证中心的认证，是军内唯一军 B + 级安全认证的操作系统；通过中国人民解放军电子工程学院安全攻击测试，在抗缓冲区溢出攻击、抗权限获取和提升、拒绝服务攻击的防护能力方面比 RedHat Linux 有显著提升。



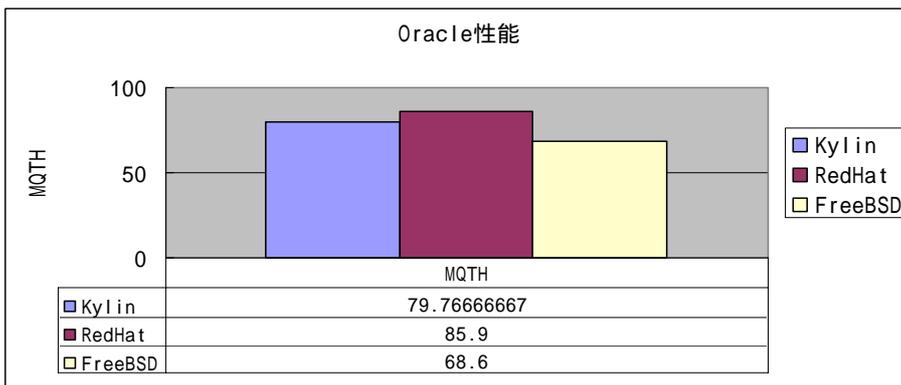
4、性能的优化和改进。银河麒麟服务器操作系统为了满足网络化应用高性能的需求，除了在层次式内核结构方面的优化工作外，还在内核级多线程 Web 服务器、银河麒麟的虚拟服务器和优化的用户级多线程库等方面进行了针对性优化设计。内核级多线程 web 服务器——Keta 采用了软流水体系结构，将传统 web 服务请求处理流程细化为若干流水栈，并使用受控缓存区 (Managed Buffer) 作为自洽的任务控制数据的存储体，通过数据缓存 (Data Cache) 方便数据的共享，提高请求的处理速度。银河麒麟虚拟服务器——KVS，采用抽象三层体系结构，在 IP 层实现了单系统映像，并提供实际服务的服务器集群和稳定、一致的文件存储服务系统，从而满足银河麒麟在集群系统中的负载平衡能力。优化的用户级多线程库——Kyth，采用两级调度方式，支持 1 对 1 和多对多线程模型，应用层可通过系统调用方式指导和引发内核级多线程调度，内核也可通过事件机制指导用户级线程调度，该线程库不同于 Linux 下的 glibc 和 NPTL，也不同于 FreeBSD 下的 libc 和 libkse，并且完全兼容 POSIX pthreads 规范，是银河麒麟实现 Linux 目标代码兼容和高效支持多线程应用的基础。中国人民解放军总装备部武器装备论证中心 (军方) 和赛迪中国软件评测中心 (民口) 分别对银河麒麟服务器操作系统的性能同 Linux 及 FreeBSD 进行了对比测试，其中最重要的 Web 服务性能测试和 TPC-C 数据库性能测试结果如下：

- Web 服务性能：采用 SpecWeb99 测试工具
测试结论：在 Keta 支持下，Kylin2.0 要明显优于 RedHat9.0 和 FreeBSD5.3



➤ 数据库性能方面：数据库均使用 Oracle 9i，采用 TPC-C 测试工具

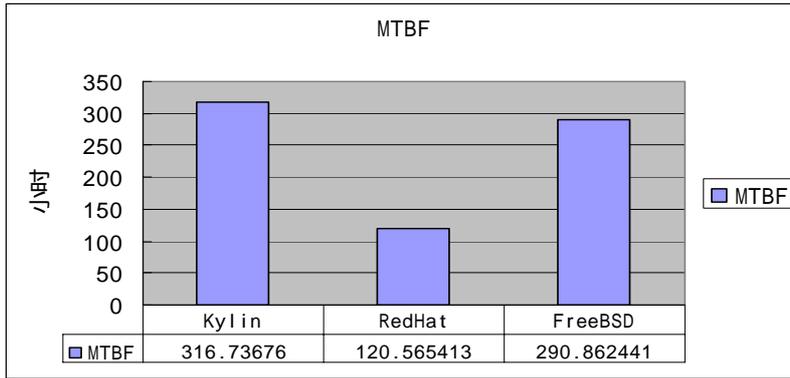
测试结论：Kylin 上 Oracle 性能比 RedHat Linux 差（主要原因是 Oracle 做了一些针对 Linux 的优化，而这些优化在银河麒麟上不能发挥效果），比 FreeBSD 5.3 强（主要原因是 FreeBSD 对 Oracle 的支持较差造成的）



5、高可用子系统的实现。计算机系统的可靠性很大程度上取决于操作系统的可靠性。减少可能导致系统失败的潜在因素可以有效提高系统的可靠性。提高可靠性还意味着增加了两次故障之间的间隔时间。可靠性和系统一起顽强对抗着可能出现的故障，当系统出现死机之后，良好的可靠性能够保证系统迅速恢复。银河麒麟操作系统采用内核级 Checkpoint、驱动程序加固、动态分区等技术提供了系统可靠性。中国人民解放军总装备部武器装备论证中心和中国软件评测中心分别对银河麒麟服务器操作系统可靠性同 Linux 及 FreeBSD 进行了对比测试，现将部分测试结果提供如下：

➤ 选用了世界上广泛应用的 Yamada、Ohba、Osaki 延迟 S 型可靠性增长模型测试系统的 MTBF（平均失效间隔时间）

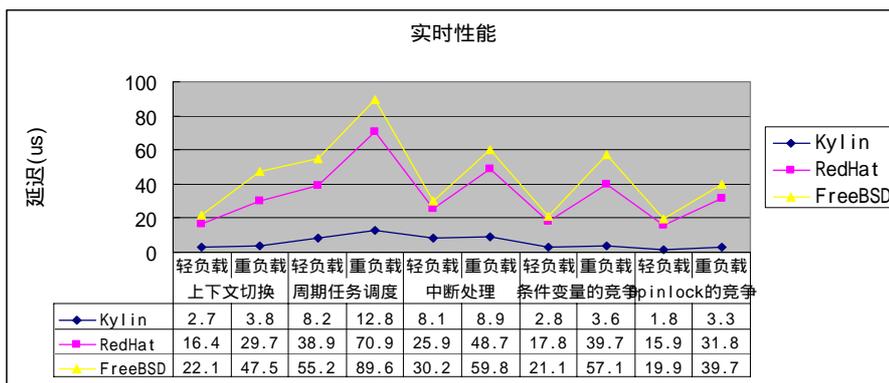
测试结论：Kylin2.0 的 MTBF 要明显大于 RedHat9.0 和 FreeBSD5.3



6、强实时的实现。最大中断禁止时间、最大任务切换时间等是衡量一个实时操作系统性能的关键指标，银河麒麟服务器操作系统采用高精度的实时时钟、快速的进程/线程切换、迅速的中断响应、具有超时处理功能等技术，实现了实时功能。中国人民解放军总装备部武器装备论证中心和中国软件评测中心分别对银河麒麟服务器操作系统实时性同 Linux 及 FreeBSD 进行了对比测试

➤ 采用 OpenGroup 的 UNIX'03 实时测试集

测试结论：Kylin2.0 的实时性能要明显优于 RedHat9.0 和 FreeBSD5.3



以上只是部分对比测试结果，应该说，Kylin 在整体性能上与国际上开源的 Linux 和 FreeBSD 已经具备一定的比较能力，但是 Kylin 在 I/O 性能方面同 Linux 和 FreeBSD (UFS2 + SoftUpdate) 相比还存在不小的差距，这也是银河麒麟开发组目前正在进行的工作之一。同国际上商用的一些服务器操作系统的对比测试，如 Windows 2003 Server、Solaris 10、IBM AIX5L、HP UX11i，由于条件限制（许多商业版 UNIX 服务器操作系统都是运行在专门的 CPU 平台之上的），还没有进行。

7、易用性方面的改进。银河麒麟操作系统在系统安装和桌面环境的易用性方面做了大量工作。它实现了简单易用的安装系统，支持图形化安装和字符化安装两种安装方式，支持光盘安装和 NFS 安装两种安装途径。它采用 X Window 作为基本图形环境，采用 Gnome 作为桌面环境，实现了 Windows 风格的控制面板，能够自动加载本机器上的 Windows 分区，能够自动加载 usb 磁盘设备，这些都使得 Windows 用户能够更加平滑地过渡到银河麒麟服务器操作系统。应该说，在这方面，与 Windows 相比，需要改进的地方还很多；但与 FreeBSD 相比，优势还是很明显的；与 Linux 相比，整体水平相当。

三、银河麒麟网站上公开发布的版本是银河麒麟操作系统的简化版

银河麒麟服务器操作系统在设计之初，主要目标是在国防、金融、电子政务等关键行业的服务器应用，因此，银河麒麟操作系统开发组针对这些应用需求在服务器操作系统的性能、安全、高可用和强实时等关键技术方面进行了大量的工作，由于涉及到特定应用领域在安全保密方面的限制，我们在网站上公开发布的是银河麒麟操作系统裁减后的简化版本，该版本能基本满足各种非敏感行业的应用需求。。如果用户在服务器应用领域有特定的要求，请通过网站提供的渠道同我们取得联系。

十多年的操作系统研究实践，我们品尝过失败痛苦，也体验过成功的喜悦。我们深知，操作系统研究是一项复杂而艰巨的工作，一个操作系统从问世到完善，需要经过长时间的磨练，银河麒麟操作系统也不例外。我们充分理解大家希望国产操作系统能够早日与国外商用操作系统分庭抗礼的急迫心情。目前，银河麒麟服务器操作系统相对已有多年历史的 Darwin、Linux 和 FreeBSD 而言，已经形成了自己的特色，但在许多方面还存在一定的差距，我们唯有加倍努力地工作，才能不辜负大家的期望。我们希望，随着版本研发的不断推进，麒麟能更加完善，最终形成能够和国际主流服务器操作系统相抗衡的实力。我们将始终牢记银河麒麟开发组的座右铭——“踏踏实实做麒麟”，为大家奉献更新更好的产品。

希望上述说明能够让广大的网友和银河麒麟的用户能够更深入地了解银河麒麟操作系统，欢迎大家提出宝贵意见和建议。再一次感谢大家对银河麒麟的关心和支持。