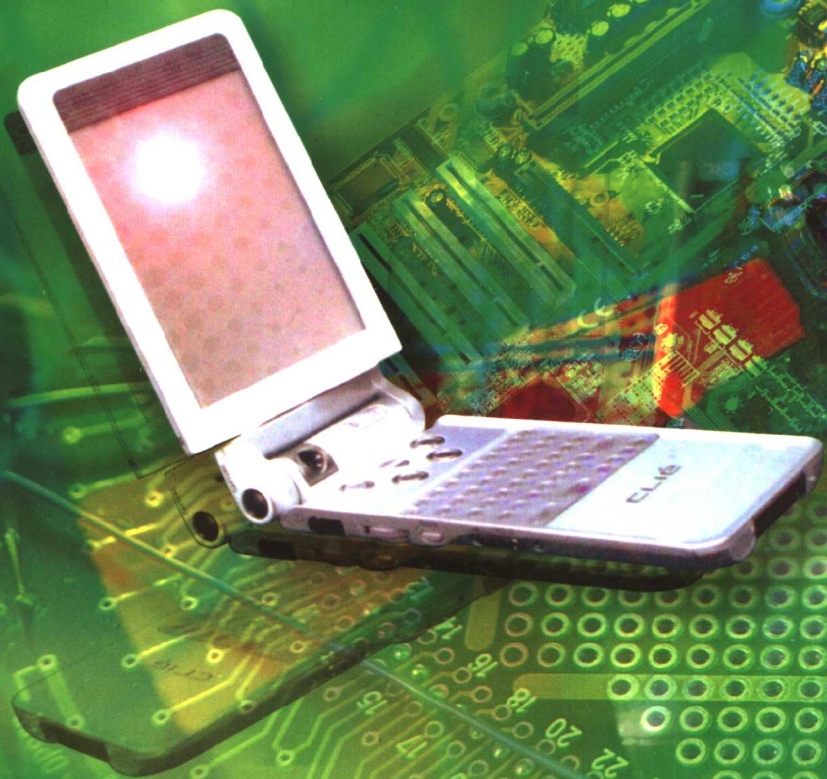


深入嵌入式 Java 虚拟机 Inside KVM

科 技 新 贵

探砵工作室 著



中国铁道出版社
CHINA RAILWAY PUBLISHING HOUSE



深入嵌入式 Java 虚拟机

探研工作室 著

中国铁道出版社

2003·北京

(京)新登字 063 号

北京市版权局著作权合同登记号: 01-2003-0990 号

版 权 声 明

本书中文繁体字版由台湾学贯行销股份有限公司出版。本书中文简体字版经台湾学贯行销股份有限公司授权由中国铁道出版社出版。任何单位或个人未经出版者书面允许不得以任何手段复制或抄袭本书内容。

本书贴有学贯行销激光防伪标签,无标签者不得销售。版权所有,侵权必究。

图书在版编目(CIP)数据

深入嵌入式 Java 虚拟机/探砂工作室著. —北京:中国铁道出版社, 2003. 4

ISBN 7-113-05214-2

I. 深… II. 探… III. JAVA 语言—程序设计 IV. TP312

中国版本图书馆 CIP 数据核字(2003)第 028001 号

书 名: 深入嵌入式 Java 虚拟机

作 者: 探砂工作室

出版发行: 中国铁道出版社(100054, 北京市宣武区右安门西街 8 号)

策划编辑: 严晓舟 郭毅鹏

责任编辑: 苏 茜 刘 颖 刘文龙

封面设计: 孙天昭

印 刷: 河北省遵化市胶印厂

开 本: 787×1092 1/18 印张: 22.25 字数: 444 千

版 本: 2003 年 5 月第 1 版 2003 年 5 月第 1 次印刷

印 数: 1~5000 册

书 号: ISBN 7-113-05214-2/TP·926

定 价: 36.00 元

版权所有 侵权必究

凡购买铁道版的图书,如有缺页、倒页、脱页者,请与本社计算机图书批销部调换。

出版说明

近几年来，由于移动终端设备的蓬勃发展，各式各样的硬件设备、软件平台都被开发出来加入这场 21 世纪的移动通讯大战。在硬件设备上有 Intel 公司的 StrongARM 系列，Motorola 公司的 Dragon ball 系列等；软件平台有著名的 Linux 操作系统，Microsoft 的 pocket PC，Accelerated Technology 公司的 Nucleus PLUS 等。一个程序设计员如果想要写出能够横跨这么多平台的应用程序来，是一件不容易的事。因此，本身具有跨平台特性的 Java 程序语言就成为目前在手机、PDA 等嵌入式系统中越来越受欢迎的热门软件平台了。

在本书中，作者以嵌入式平台上最常见、最合适的 Java 虚拟机—KVM 为核心，讲述 Java 虚拟机的内部实际架构，深入探讨实际程序代码的编写原理，并以丰富的图形来向各位读者阐述 Java 虚拟机的奥秘与秘密。作者还融汇了过去在实际工作中的调试经验和测试结果，向读者解开调整 KVM 效率的秘密，希望读者凭借本书来一窥嵌入式 Java 虚拟机的奥秘。

本书由学贯行销股份有限公司提供版权，经中国铁道出版社计算机图书中心审选，彭林、贾英茂、游广志、刘宇、朱远波、张新东、张琦等同志完成了本书的整稿工作，陈兰芳、崔仙翠、程瑞芬等同志完成了本书的编排工作。书中难免有疏漏之处，诚请各位专家和读者批评指正，我们也会在适当时间进行修订和补充，并发布在天勤网站：<http://www.tqbooks.net> “图书修订”栏目中。

中国铁道出版社

2003 年 4 月

目 录

第 1 章 Java 与 KVM 的关系	1
1-1 Java 的概念.....	2
1-2 Java 平台结构.....	2
1-3 Java 的特色.....	4
1-4 Java 的世界.....	5
1-5 企业版 J2EE.....	6
1-6 标准版 J2SE.....	7
1-7 微型版 J2ME.....	9
1-7-1 J2ME 的架构.....	11
1-7-2 J2ME 的特性.....	11
1-7-3 目前的 J2ME 环境.....	12
1-8 K Virtual Machine.....	15
1-8-1 为什么要用 KVM.....	19
1-8-2 KVM 移植技术.....	20
1-9 Java Card.....	27
1-10 小结.....	28
第 2 章 Java 类文件格式	29
2-1 magic number.....	31
2-2 minor_version, major_version.....	31
2-3 constant_pool_count, constant_pool[].....	32
2-3-1 CONSTANT_Class.....	33
2-3-2 CONSTANT_Fieldref.....	33
2-3-3 CONSTANT_Methodref.....	34
2-3-4 CONSTANT_InterfaceMethodref.....	34
2-3-5 CONSTANT_String.....	35
2-3-6 CONSTANT_Integer.....	35
2-3-7 CONSTANT_Float.....	35
2-3-8 CONSTANT_Long.....	36

2-3-9	CONSTANT_Double	36
2-3-10	CONSTANT_NameAndType.....	37
2-3-11	CONSTANT_Utf8.....	39
2-4	access_flags.....	40
2-5	this_class	41
2-6	super_class.....	41
2-7	interfaces_count, interfaces[].....	42
2-8	fields_count, fields[].....	43
2-9	methods_count, methods[]	45
2-10	attributes_count, attributes[].....	47
2-10-1	SourceFile attribute.....	48
2-10-2	ConstantValue attribute	49
2-10-3	Code attribute.....	50
2-10-4	Exceptions attribute	53
2-10-5	InnerClasses attribute.....	54
2-10-6	Synthetic attribute	56
2-10-7	LineNumberTable attribute.....	56
2-10-8	LocalVariableTable attribute	58
2-10-9	Deprecated attribute.....	59
2-11	HelloWorld_simple.class.....	60
2-12	小结	70
第 3 章 KVM 执行时所用的 class 与 instance 结构		71
3-1	执行时期的 instance 结构.....	72
3-2	执行时期的 class 结构	78
3-3	执行时, class 与 class 之间的关系.....	80
3-4	加载 Java 类的过程.....	83
3-4-1	loadClassfileInternal().....	90
3-4-2	loadClassfileHelper()	97
3-5	小结.....	103
第 4 章 KVM 执行时所用的数据结构		105
4-1	The pc Register.....	106
4-2	Java heap	106

4-2-1	allocateHeap()	111
4-2-2	InitializeHeap()	112
4-2-3	callocPermanentObject()	115
4-3	Java execution stack	118
4-4	Runtime Constant Pool	120
4-5	Frames	133
4-5-1	Local variable array	139
4-5-2	Operand stacks	140
4-5-3	Dynamic linking	141
4-5-4	Method 正常结束执行的情况	142
4-5-5	Method 不正常结束执行的情况	142
4-6	小结	142
第 5 章	KVM 内部的 Interpreter	145
5-1	KVM 运行时间取代 byte code 的机制	148
5-2	Split infrequent byte codes	156
5-3	在 KVM 内与 interpreter 相关的文件	158
5-3-1	bytecodes.c	158
5-3-2	execute.c	163
5-4	更进一步加快 interpreter 的速度: 使用汇编语言	169
5-5	小结	171
第 6 章	Methods invokation	173
6-1	从 Java 类文件中加载 method 属性到内存中	176
6-1-1	loadOneMethod()	178
6-1-2	getUTF8String(POINTERLIST_HANDLE, unsigned short)	185
6-1-3	verifyMethodFlags()	185
6-1-4	verifyName(const char*, enum verifyName_type, bool_t)	189
6-1-5	skipOverFieldType(const char*, bool_t, unsigned short)	194
6-1-6	skipOverFieldName(const char*, bool_t, unsigned short)	198
6-1-7	change_Name_to_Key()	202
6-1-8	verifyMethodType()	208
6-1-9	loadMethodAttributes()	210
6-1-10	loadCodeAttribute()	215

6-1-11	loadExceptionHandlers()	219
6-2	调用 Method	224
6-2-1	pushFrame()	224
6-3	小结	238
第 7 章	Exceptions	239
7-1	何时会抛出 exceptions	240
7-2	KVM 如何抛出 exceptions	241
7-2-1	raiseException()	241
7-2-2	raiseExceptionMsg()	243
7-2-3	fatalVMError()	244
7-2-4	fatalError()	244
7-2-5	throwException()	248
7-3	处理一个 exception	253
7-3-1	findHandler()	254
7-4	小结	256
第 8 章	Garbage Collection	257
8-1	Tracing garbage collection	259
8-2	Copying garbage collection	260
8-3	KVM1.0 版所使用的 garbage collection 机制	261
8-4	目前的 KVM 版本所使用的 garbage collection 机制	262
8-4-1	Object header	262
8-4-2	Free list 以及 available memory chunk	265
8-4-3	Compacting	270
8-4-4	KVM 内的 garbage collector 函数	275
8-4-5	Temporary root	278
8-4-6	Global root	280
8-5	小结	281
第 9 章	编写 KVM 的 native methods	283
9-1	一个简单的范例	286
9-1-1	Java 端的处理	287
9-1-2	C 程序端需要作的处理	291

9-2	在 KVM 内, 对 native code 的编写有帮助的 functions.....	293
9-3	Java code 与 native code 之间传递参数的方法	294
9-3-1	导入 primitive data type 类型的参数或返回值	295
9-3-2	导入 instance 或 array 类型的参数或返回值	300
9-3-3	导入 instance 类型的参数或返回值.....	300
9-3-4	导入 array 类型的参数或返回值	303
9-4	抓取 Java class 内某个 variable 的方法	311
9-4-1	在 instance method 中抓取 instance variable 的方法.....	311
9-4-2	在 instance method 中抓取 static variable 的方法	313
9-4-3	在 static method 中抓取 static variable 的方法.....	313
9-5	在 native code 中如何抛出 exception	315
9-5-1	三种主要的 native functions 用来抛出 exceptions.....	315
9-5-2	其他能抛出 exceptions 的 native functions.....	316
9-5-3	KVM 内部已经定义好的 exception class 字符串	317
9-6	Synchronized block 的改写	318
9-6-1	Java 的 synchronization.....	318
9-6-2	Monitor 的 notify 队列与等候队列.....	320
9-6-3	在 native code 中如何完成 Java 的 synchronized 关键字.....	321
9-6-4	KVM 提供了如下的函数可供我们来使用	323
9-6-5	在 native code 中如何完成 Java 中的 notify() 函数	323
9-7	Garbage collection.....	324
9-7-1	handle.....	326
9-7-2	Temporary root	327
9-7-3	Global root	329
9-8	小结.....	329

第 10 章 The Technology Compatibility Kit(TCK)

10-1	JavaTest 简介	334
10-2	利用 JavaTest 来进行验证工作的架构	338
10-3	配合 JavaTest 来执行 Java 程序之前的准备操作	339
10-4	告诉 JavaTest 你的每一个测试用 Java 程序的详细信息	340
10-5	执行 JavaTest.....	342
10-6	小结.....	345

Chapter

1

Java 与 KVM 的关系



KVM 的全名叫做 Kilobyte Virtual Machine，是由这三个英文字头的一个字母所串接而成的。它是 Sun Microsystems 所编写的一套 Java 虚拟机，属于 Sun 的 J2ME 这一个 Java 版本的一部分。如果读者从来没有接触过 Java 的话，绝对不会了解上面所出现的一些专有名词，甚至对很多已经接触过 Java 的人来说，上述的一些专有名词还是会有所模糊。因此接下来笔者将对整个 Java 世界的架构与环境作一个概略性的说明。

1-1 Java 的概念

1991 年 Sun Microsystems Inc. 的绿色计划 (Green Project) 原本使用 C, C++ 来开发消费性电子产品，但后来发现这些语言的功能与效果，无法达到程序设计师们的理想目标，加上那个时候互联网开始蓬勃的发展，于是乎 Sun 就与一群有志之士开始着手设计一套完全面向对象，而且不受平台限制的语言。设计小组原本是以公司外面的一棵橡树为名，把这一套全新发展的计算机语言称为 oak，但是后来发现这个名字已经为其他人所使用，最后设计小组突发奇想，以开会时的咖啡厅来为这一套新开发的计算机语言来命名，这就是 Java 的由来，所以我们时常看到其图标是一杯热滚滚的咖啡，如图 1-1 所示。



图 1-1 Java 的咖啡杯图样

1-2 Java 平台结构

其实 Java 是个什么样的东西，我想对那些已经接触过 Java 的读者应该多多少少有一些基本的印象，本质上它是一种由美国升阳计算机公司 (Sun Microsystems, Inc.) 所研发的软件技术，整体平台结构如图 1-2 所示：

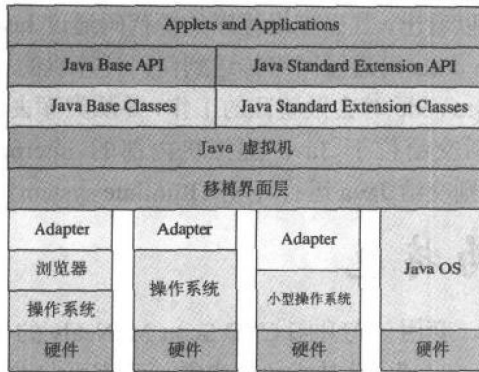


图 1-2 Java 系统结构图

由上图中我们可以看出，Java 的整体架构主要由两个部分所构成，即虚拟机 (Virtual Machine) 和应用程序接口 (API)。读者可以把虚拟机看成是一套虚拟的计算机，有一个标准的规格可以用软件或硬件来实现，Sun 开放其虚拟机的源代码，其他的如 IBM 也有自己编写的虚拟机。

而位于 Java 虚拟机下层的移植接口层和 Adapter 则是为了使 Java 虚拟机能够方便地移植到不同的操作系统之上而开发出来的。另外，位于 Java 虚拟机之上的 Java API 部分则包含了基本型的 API 类与标准延伸的 API 类，前者如标准 I/O，网络功能，GUI 等，后者则是除了前者之外，另外再加上去的应用类。

Java 的运行时期流程如图 1-3 所示：

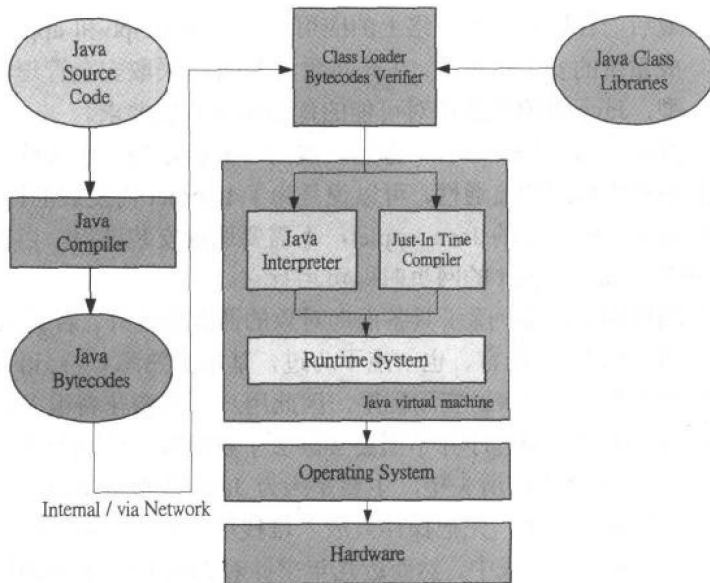


图 1-3 Java 应用运行流程

由上图中我们可以看出，首先应用程序的源代码通过 Java 编译器转换为特殊的 byte code，而 byte code 可通过网络传送到任何计算机的 Java 虚拟机去运行。程序加载器在加载 byte code 时会做验证的工作，并从类库去动态加载所需使用的 Java 类 (class)，而后交给位于 Java 虚拟机内部的 interpreter 或 Just-In-Time Compiler 解释，最后则交给 Java 运行环境 (Runtime system) 去运行。

1-3 Java 的特色

最大的特色优点，套用一句升阳 CEO Scott McNealy 的话：“Write once, run anywhere on anything safely”。从技术角度来看，我们可归纳出如下的几个特性：

1. 跨平台的语言：众所皆知，Java 语言是可以跨平台运行的，只要在 Java 虚拟机 (Java Virtual Machine) 支持的平台上，甚至在读者自己的平台上也可以自行将 Java 虚拟机移植上去，就如同 PalmOS 上的 K Virtual Machine 一样，让 Java 应用程序可以在其机器上运行。
2. 极佳的安全性：Java 具有多个层次的安全防护措施，可以阻挡病毒与其他的入侵行为，以避免造成 Java 系统内的不稳定。其安全性措施大致可归纳为三类：
 - 第一，Java 编译器内建有 byte code 验证器，在 Java 源代码编译成 Java byte code 时会做安全性检查，例如运算溢出 (Overflow) 等。
 - 第二为 Class Loader 在加载类时会指定其安全属性。
 - 最后一种则是有关网络上的应用，有一些 script 和 applet 可能会造成系统的安全性疑虑，在此问题上 Java 是采取安全管理器的沙箱模型，Java 虚拟机会针对可能的危险做出认证要求。
3. 适合网络应用：Java 语言一开始发展的目的就是为了解决软件在公共网络上不同平台间的流通性，可以说是为了 Internet 而发展的计算机语言，最明显的如网页上的 Java applet，不需要通过安装过程，当浏览器一开启内嵌有 Java applet 的网页时即可运行它。
4. 容易编写应用程序：除了具备面向对象的四大特性外，若说 Java 语言是一种简化的 C++ 语言，也一点不为过，基本上熟悉 C++ 语言的程序设计员就有办法写 Java 应用程序，因此用户不需要花费很多的时间去学习，而且编写应用程序并不需要实际去了解硬件平台与操作系统。

用户常对 Java 的效率有所失望，主要是因为 Java 虚拟机本身是以堆栈机器 (Stack Machine) 的概念作成，先把操作数放入堆栈中，再一一由堆栈顶端拿出来作运算，运算结果再放进堆栈中。然而近几年来针对 Java 效率不高的问题已有较佳的解决方案，以下笔者举出几个近几年来用来改进 Java 效率的常见方案：

- ◆ Just in time Compiler (JIT compiler) 针对 byte code 需运行时才去翻译。
- ◆ 用硬件芯片来实际操作出 Java 虚拟机内部的 interpreter, 使其 (interpreter) 负担最重的部分由硬件去解决, 这就是我们常听到的 Java chip, 如 JavaSoft 与 Sun Microelectronic 的 PicoJava, MicroJava 和 UltraJava 芯片; 国内工研院也有类似的研发成果; 另外国外也有很多公司在进行相关的研发, 例如 ARM 积极开发中的 Jazelle 技术让 ARM 微处理器可以直接运行 Java byte code, 以及一家叫做 Nazomi 的公司也有一套相当成熟的 Java Chip 产品了。
- ◆ Sun 的 HotSpot 技术采用在新的 Java 虚拟机上。

1-4 Java 的世界

Java 发展至今, 其应用已经推广到各种平台, 尤其是在互联网盛行的情况下, Java applet 和 application 几乎都可以通过网络来下载运行, 其他的应用更是不胜枚举。虽然 Sun 早期只维护一套 Java 的运行环境。但在近年对 Java 的使用益趋广泛的同时, 针对不同的应用平台如果都套用同一个 Java 运行环境的话, 就显得不合适了。因此, Sun 将最近所开发的 Java 运行环境区分成 4 大版本, 分别是应用于服务器 (Server) 上的 J2EE (Java 2 Enterprise Edition), 应用于一般个人计算机 (Personal Computer) 上的 J2SE (Java 2 Standard Edition), 应用于小型设备 (Mobile Device) 上的 J2ME (Java 2 Micro Edition) 以及应用在 Smart Card 上的 Java Card。图 1-4 即是一个区分出这四大 Java 版本所应用平台的表示图:

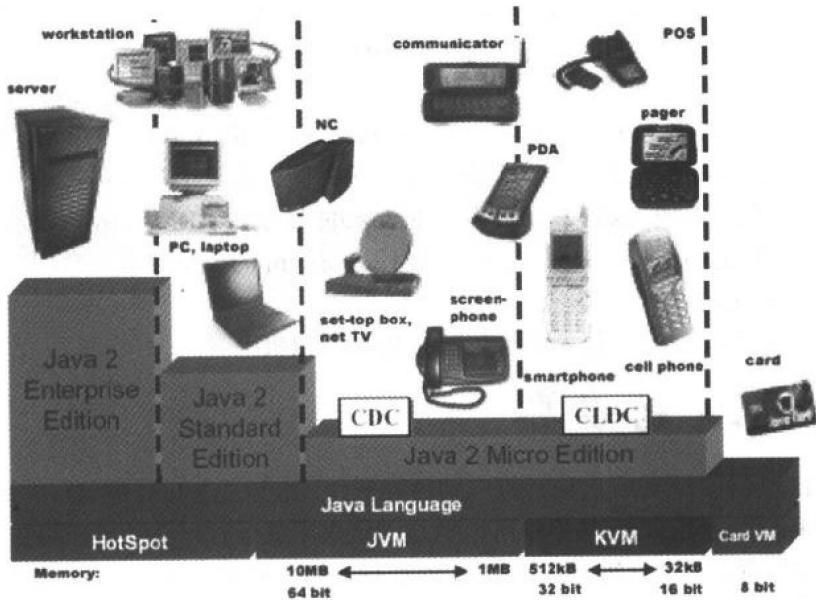


图 1-4 Sun Microsystems Inc. 的 Java 2 Editions 平台和目标市场 (文件来源: <http://java.sun.com>)

1-5 企业版 J2EE

J2EE 技术在于协助企业发展服务器的功能, 例如电子商务、网络内容服务等领域, 它的两个典型应用如图 1-5 所示:

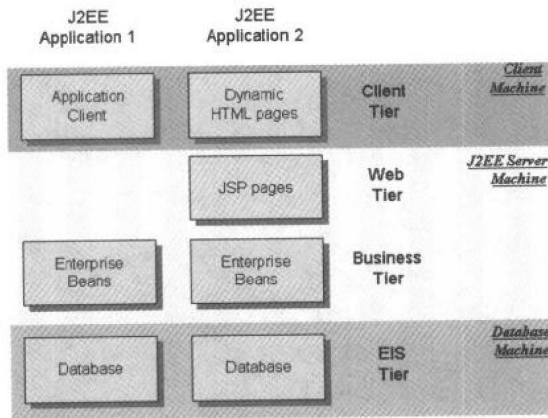


图 1-5 Sun J2EE 平台中, 采取 Multitiered 架构的应用

由上图中我们可以看出, J2EE 的应用通常都是三个层次架构, 其前端, 也就是客户端 (Client), 支持 Dynamic HTML、XML 或 WML。而下载的内容, 则可

能是以 Java Server Page 技术来做成的动态网页或 Servlets，也可以是运行在嵌入于网页浏览器内 JVM 上的 Java Applet，以提供实时信息等服务。用户可以通过网络浏览器来存取，或是通过其他在虚拟机上运行的 Java Application，通常以 AWT (Abstract Window Toolkits) 作成的图形化接口软件。

在 Server 机器端，主要提供 Enterprise JavaBean Components，用来提供客户端和服务器的文件流。通常为了方便重复使用，在设计编写 JavaBean 的组件时，则必须遵守其架构规范。通常 J2EE 的应用上会设计成 Thin Client 的架构，也就是将 load 较重的工作，如数据库搜寻，交由服务器去运行，其他负载轻的工作交由 Client 端运行。在商业上用到的 JavaBean 包含有三类：

- ◆ Session Beans 以 Client 的工作运行完成与否来决定文件去留。
- ◆ Entity Beans 则必须保证 Client 或 Server Shutdown 时文件存储于数据库的完整性。
- ◆ Message-Driven Beans 则结合 Session Beans 和 Java Message Service 机制达到组件之间稳定的同步和异步的信息沟通。

最后，J2EE 的应用软件必须与企业信息系统 (Enterprise information System) 连接，而企业信息系统则可能包含有 Enterprise Infrastructure 系统软件，如 ERP (Enterprise Resource Planning)，Mainframe Transaction Systems 和 Database 等。Sun 目前提供下载教育用或展示用的 J2 SDK EE 版、API、服务器和相关的数据库系统，支持的平台有 Solaris、Windows NT、Windows 2000 和 Linux。其他提供有 J2EE 软件的公司还包括如 BEA、Sun iPlanet、HP、Sybase、Borland、SilverStream 等，事实上 J2EE 是目前 Java 的应用上最成功的领域之一。

1-6 标准版 J2SE

J2SE (Java 2 Platform Standard Edition) 是 Sun 用来在桌上型计算机 (PC 或 Workstation) 上发展与运行 Java 程序的 Java 版本。在这个版本中所包含的组件则是定义在 J2SE 的规格书上，有运行时期组件如 Applet、Java 虚拟机、运行时期环境、类加载器、Just-In-Time Compiler 和支持类等。在笔者完成本书的时候，Sun 的 J2SE 最新版本为 1.4 版。图 1-6 就是一个用来代表 J2SE 1.4 版内部架构的方块图：

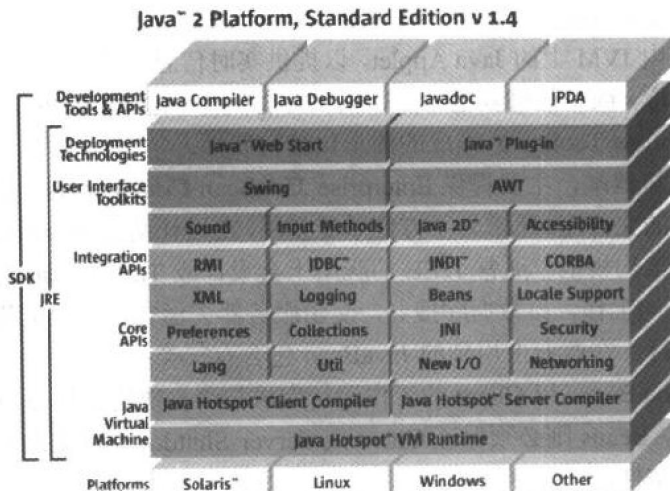


图 1-6 Sun J2SE 1.4 版结构图（文件来源：<http://java.sun.com/>）

J2SE 1.4 版相对于之前的 1.3 版来说，新的 J2SE 1.4 版除了持续加强运行效率和系统可靠度之外，还提高了对 XML、CORBA、IPv6 和 JDBC 等技术的连接性支持，增加了在传输上的安全性加解密技术，改善了效率与弹性等以往为人所诟病的地方，提供给 Client 与 Server 的程序开发人员更多的开发环境。图 1-7 为笔者在运行其 Java 2D 范例程序的效果：

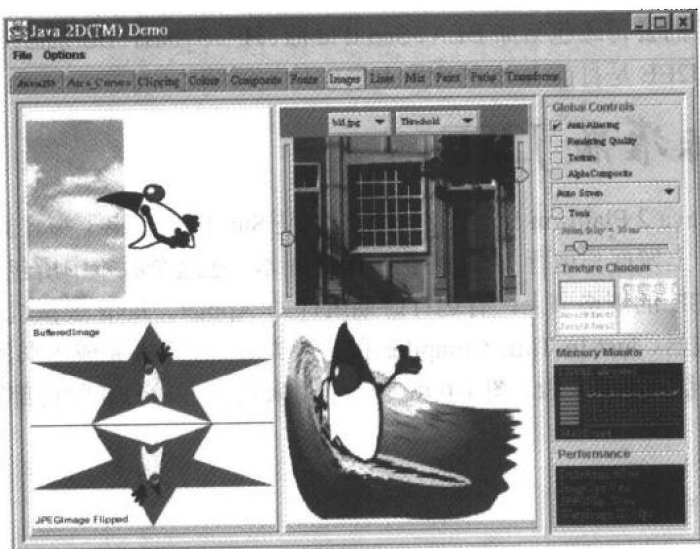


图 1-7 Sun J2SE 上面一个 2D 绘图程序范例

也许因为 Java 本身的效率问题，造成在 2D/3D 等多媒体应用环境中并不突