

基于龙芯的数据科学工作环境

李舰

2019 中国 R 语言会议（上海）

华东师范大学

2019 年 12 月 21 日

目 录

1 可信计算与自主可控

- 可信计算简介
- 自主可控平台

2 龙芯简介

3 数据科学工作环境

目 录

- 1 可信计算与自主可控
 - 可信计算简介
 - 自主可控平台
- 2 龙芯简介
- 3 数据科学工作环境

信息安全与可信计算

● 信息安全 (Information Security)

- 信息安全是指为信息系统建立和采取的技术及管理的安全保护，即保护信息系统中的硬件、软件和数据信息资源，不因偶然或恶意的原因遭到破坏、更改、泄漏，使信息系统连续可靠地正常运行，信息交换和共享服务正常有序不被中断。
- 信息安全是一门涉及计算机技术、网络技术、通信技术、密码技术、软件技术、应用数学、数论、信息论等多种科学的综合性学科。
- 信息安全主要包括密码技术、身份认证技术、防火墙技术、VPN 技术、安全扫描技术、入侵检测技术、可信计算技术。

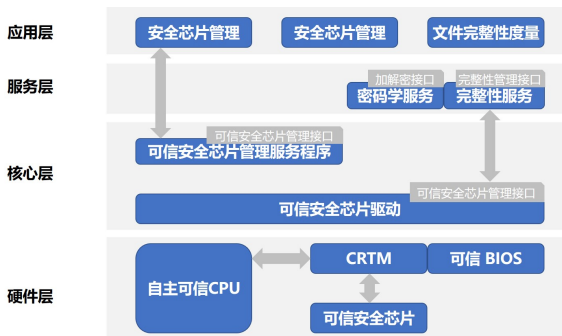
● 可信计算 (Trusted Computing)

- 可信计算技术是指在信息系统的硬件架构上添加可信计算安全芯片模块及相应可信软件，以构建一个操作系统体系之外的可信计算安全平台。
- 传统信息安全系统以防外和封堵为特征，难以应对目前主要源自内部的安全威胁，而可信计算技术在硬件平台引进安全芯片架构，可以对各种不安全因素实现主动防御。

可信计算平台体系结构

● 软硬件体系

- 硬件层由可信安全芯片、可信度量核心根（CRTM）和可信 BIOS 组成，是可信计算平台的基础。
- 软件层包括内核层、服务层和应用层，内核层与安全芯片交互，服务层以操作系统服务的形式存在，应用层提供可信安全服务支持（例如文件可信加密、软件可信运行）。



可信计算技术简介

● 密码技术

- 可信计算平台基于密码技术，实现平台自身的完整性、身份可信性和数据安全性等安全功能。
- 常见的密码算法有 DES、AES、SM4 等对称密码，RSA、SM2 等非对称密码（公钥密码），以及 SHA1、SM3 等哈希密码。

● 可信度量技术

- 可信度量技术基于完整性度量，负责度量计算平台的可信状态，包括每一个组件、每一个硬件及软件的可信状态。
- 常见的可信度量技术包括 IMA、PRIMA、DynIMA 等。

● 信任链技术

- 基本思想是从一个初始的信任根（通过安全芯片从物理层面建立）出发，在计算环境的每一次转换时，如果信任可以通过传递保持下去不被破坏，那么计算环境始终可信。

● 远程证明技术

- 将本地计算平台的信任扩展到远程计算平台的技术，包括平台身份证明和平台完整性状态的证明。

可信计算的技术发展

● 国际情况

- 20 世纪 90 年代中期，国外一些计算机厂商开始提出在硬件层内置安全模块的方案，是可信计算的雏形。
- 1999 年，在 CMU 和 NASA 的发起下，英特尔、IBM、惠普、微软等企业加入，成立了可信计算平台联盟（TCPA）。
- 2003 年时，TCPA 发展出了 200 多个会员，改名为可信计算组（TCG），并逐步建立起 TPM 标准，到 2016 年时，TPM 已经成为许多终端设备的标配。

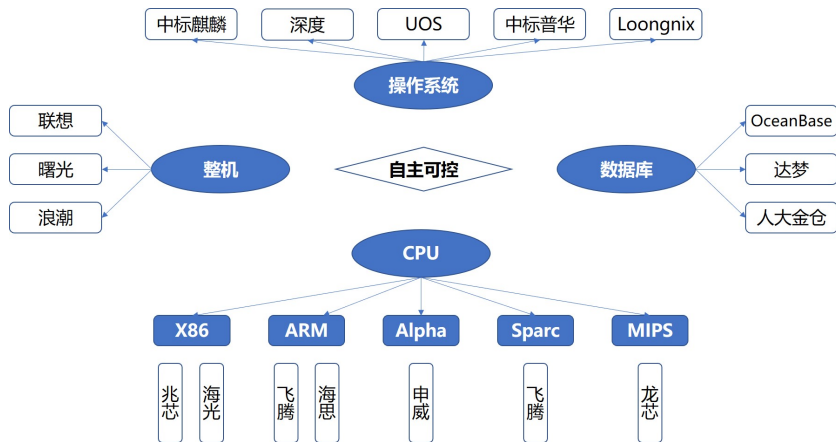
● 国内情况

- 2000 年武汉瑞达和武汉大学合作，借鉴 TCG，提出一套安全可信技术方案，2004 年通过国家鉴定，2005 年联想和兆日的 TPM 芯片研制成功。
- 2006 年中国可信计算工作组（TCMU）成立，制定出以可信密码模块（TCM）为核心的系列标准，2007 年，联想、同方、瑞达等厂商开发出基于该标准的产品。
- 2008 年之后，TCM 产品开始规模上市，获得了很多领域的高度认可。到 2010 年已建立起基本完整的产业体系。

目 录

- 1 可信计算与自主可控
 - 可信计算简介
 - 自主可控平台
- 2 龙芯简介
- 3 数据科学工作环境

自主可控产品现状



自主可控 CPU 发展模式

● 完全独立自主模式

- 龙芯获得 MIPS 永久授权的同时，对其指令集进行了大量的扩展与优化，发展成为龙芯自己的 Loongson 处理器，实现了完全自主的发展。
- 申威在 Alpha 指令集基础上进行了相应扩展与完善工作，独立自主地构建了自己的体系。

● 授权指令集模式

- 飞腾购买了 ARM64 架构的永久授权，可以自主研发。
- 海思购买了 ARMv8 架构的永久授权，可以自主研发。

● 外商合作模式

- 兆芯有威盛的资本（港澳台外资），得到了威盛的 X86 授权，可以自主生产 X86 处理器。
- 海光和 AMD 成立了合资公司，通过复杂的合作模式生产 Zen 架构处理器。
- 展讯和 Intel 合作生产 X86 处理器。

国产 Linux 简介

● 中标软件

- 中标软件有限公司成立于 2003 年，大股东是中国电子下属的中国软件，是操作系统产品专业化研发与推广企业。
- 旗下拥有“中标麒麟”、“中标普华”、“中标凌巧”三大产品品牌，在很多关键领域有广泛的应用。
- 中标麒麟以操作系统技术为核心，安全可信为特色；中标普华以办公软件为核心；中标凌巧以移动终端操作为核心。

● 深度 Deepin

- 由武汉深之度科技有限公司开发和发行，是一款非常受欢迎的 Linux 发行版，具有很强的国际影响力。
- 其口号为“免除新手痛苦，节约老手时间”，在易用和美观方面非常突出。

● 其他操作系统

- 中兴的新支点（NewStart）操作系统实力强大，可以支持 ARM、PowerPC、MIPS、X86、龙芯等不同指令集的芯片。
- 南京诚迈科技的悟空 OS 在智能硬件上有广泛的使用。

UOS 统一操作系统

● UOS 与通信软件

- 2019 年 5 月，中国电子、武汉深之度、南京诚迈、中兴、华为等操作系统巨头开始联合成立统一操作系统的筹备组。
- 2019 年 11 月 14，统信软件有限公司在北京成立，致力于打造中国操作系统新生态，已经和各大厂商开展了适配合作。
- 目前官网 (<https://www.chinauos.com/>) 开放了内测，主要基于深度系统。2019 年 12 月 13 日发布了龙芯版。



目 录

1 可信计算与自主可控

2 龙芯简介

- 产品及生态
- 安装和使用

3 数据科学工作环境

目 录

- 1 可信计算与自主可控
- 2 龙芯简介
 - 产品及生态
 - 安装和使用
- 3 数据科学工作环境

龙芯简介

● 龙芯中科

- 中科院计算所从 2001 年开始研制龙芯系列处理器，经过十多年的积累与发展，于 2010 年由中国科学院和北京市政府共同牵头出资，正式成立龙芯中科技术有限公司，旨在将龙芯处理器的研发成果产业化。
- 龙芯中科致力于龙芯系列 CPU 设计、生产、销售和服务，主要产品包括面向行业应用的专用小 CPU、面向工控和终端应用的中 CPU、以及面向桌面与服务器应用的大 CPU。

● 龙芯的历程

- 2001 年 5 月，在中科院计算所知识创新工程的支持下，龙芯课题组正式成立。
- 2002 年 8 月，我国首款通用 CPU 龙芯 1 号流片成功，此后不断升级。2009 年 9 月，首款四核 CPU 龙芯 3A 流片成功。
- 2010 年 4 月，龙芯中科技术有限公司成立，龙芯正式从研发走向产业化。
- 2017 年 4 月，龙芯 3A3000/3B3000 实现量产并推广应用。
- 2019 年 12 月 24 日，龙芯即将发布 3A4000 的信息。

龙芯的应用

● 可信计算平台

- 使用龙芯 CPU，可以避免 X86 泄露敏感信息的风险。
- 使用国产 TCM 安全芯片。
- 使用国产的 SM 系列密码。
- 软件方面，使用国产的操作系统，此外很多国内厂商提供了广泛的应用支持，比如金山、360、QQ 等。

● 行业应用情况

- 从 2015 年 4 月开始，龙芯的抗辐照芯片在北斗卫星上稳定运行。2018 年 4 月，龙芯中科收到上海微小卫星工程中心的感谢信，宣告了龙芯产品在北斗组网卫星中得到了成功应用。
- 2019 年 6 月，域名国家工程研究中心（ZDNS）宣布推出首款搭载国产龙芯 CPU 的域名服务器。
- 这些年来，龙芯 CPU 在很多关键领域的服务器和个人电脑中发挥了重要作用。
- 龙芯总裁胡伟武在采访中表示，2020 年龙芯 CPU 会推向市场，做到真正民用。

龙芯的软件生态

● 龙芯开源社区

- 龙芯的长远目标是建立一个 Wintel (Windows+Intel) 和 AA (ARM+Android) 之外的独立技术生态体系, 目前的官方项目发布在龙芯开源社区: <http://www.loongnix.org>。

● 操作系统

- 官方提供了社区维护的开源系统 Longnix, 并维护了一个衍生的 Fedora 系统, 是目前性能最好的系统。
- 普通用户推荐使用 UOS、Deepin 等国产操作系统, 易用性更好, 比较接近 Windows 的使用习惯。
- 其他主流 Linux 系统例如 Unbutu、Debian 都有官方编译的版本。很多系统的 MIPS 版本也可以直接使用。

● 应用软件

- 龙芯官方维护了常见的开发工具, 例如 Java、Tomcat、C/C++、MySQL、MongoDB、Firefox、Qt、OpenGL、PHP、Python、Ruby、Go、Docker、Hadoop、Eclipse。
- Longnix 系统还内置了 WPS Office、GIMP、金山词典、微信、QQ、百度地图、钉钉等应用工具。

龙芯的用户

● RMS 的选择

- 自由软件之父 RMS (Richard Matthew Stallman) 很长一段时间内使用的唯一电脑就是一台龙芯龙梦笔记本。
- 因为 Intel 的 BIOS 不开源，所以 RMS 怀疑存在后门，有可能被监控，在性能和自由之间他选择了后者——完全开源的龙芯电脑。



Richard Stallman

Freedom campaigner

Posted January 23, 2010 in [developer, linux](#)

Who are you, and what do you do?

I'm Richard Stallman, founder of the Free Software Movement. I campaign for computer users' freedom -- for instance, your freedom to control the software you use, to redistribute the software to others. Software that respects the user's freedom is what we call [free software](#).

In 1983 I announced the plan to develop a complete free operating system called GNU. The system that millions of people use, and often refer to as "Linux", is a variant of [the GNU system](#).

What hardware do you use?

I am using a [Lemote Yeelong](#), a netbook with a Loongson chip and a 9-inch display. This is my only computer, and I use it all the time. I chose it because I can run it with 100% free software even at the BIOS level.

目 录

- 1 可信计算与自主可控
- 2 龙芯简介
 - 产品及生态
 - 安装和使用
- 3 数据科学工作环境

电脑版本

● 龙芯的个人电脑

- 2004 年，龙芯和江苏梦兰组建了一个产业化基地，用中科梦兰的名义进行产业化尝试。2008 年中科梦兰基于龙芯 2E 芯片发布了第一款福珑迷你电脑。
- 2009 年 1 月 9 日，首家龙芯体验店在北京开业，主打两款基于龙芯 2F 芯片的产品，逸珑 8089A（512M 内存，2GB SLC 硬盘）和逸珑 8089B（1G 内存，160G 硬盘）。
- 2015 年 1 月龙芯进行了最后一批 8089B 的团购，从此该系列从市场消失。
- 龙芯 3A3000 发布后，在官方的龙芯俱乐部 (<http://www.openloongson.cn/>) 发售一款台式机和一款笔记本，由南京龙众创芯电子科技有限公司生产。

● 以 8089B 示例

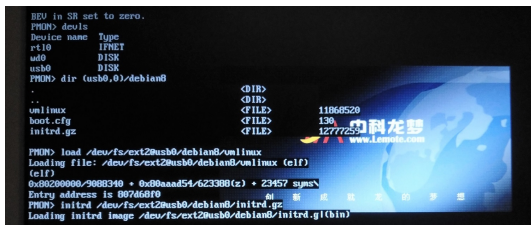
- 我们以 8089B 为例，介绍龙芯系统的安装和使用，越往后的电脑兼容性越好。
- 该电脑自带梦兰原装的系统，可以替换成主流的 Linux 发行版，推荐 Debian 8.11，目前还能提供网络更新支持。

基于 U 盘安装

● 制作启动 U 盘并安装

- 在 <http://archive.debian.org/debian/dists/jessie/main/installer-mipsel/current/images/loongson-2f/netboot/> 下载启动文件 `vmlinux-3.16.0-6-loongson-2f`、`boot.cfg`、`initrd.gz`，复制到 Fat 格式的 U 盘。
- U 盘插上电脑后重启，长按“Del”键，将会进入 PMON 命令行，输入以下命令加载，进入系统安装流程：

```
load /dev/fat/disk@usb0/vmlinux-3.16.0-6-loongson-2f
initrd /dev/fat/disk@usb0/initrd.gz
```

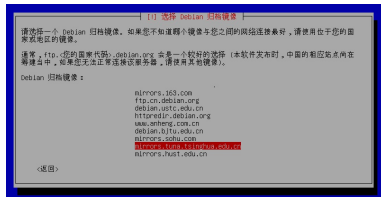


```
PMON in SR set to zero.
PMON> devls
Device name  Type
rt10        IFNET
wd0         DISK
usb0        DISK
PMON> dir (usb0,0)/debian8
.
..
vmlinux      11868520
boot.cfg     130
initrd.gz    12777259
PMON> load /dev/fs/ext2@usb0/debian8/vmlinux
Loading file: /dev/fs/ext2@usb0/debian8/vmlinux (elf)
(elf)
0x00200000/90808340 + 0x00aaad54/623308(z) + 23457 symtab
Entry address is 007d68f0
PMON> initrd /dev/fs/ext2@usb0/debian8/initrd.gz
Loading initrd image /dev/fs/ext2@usb0/debian8/initrd.gz(bin)
```

安装 Debian

● 根据提示默认安装

- 连接网络后根据提示默认安装即可，可以选择中文安装，就不会有中文编码的问题，安装成功后再将系统设置为英文。
- 在“归档镜像”中选择“手动输入信息”，第一次输入“mirrors.163.com”，然后是“[/deb.debian.org/debian/](http://deb.debian.org/debian)”，无误的话才能正常进行网络更新。



- 安装的最后可以选择轻量化的桌面“LXDE”。初次登录时图形界面可能有问题，需要使用“ctrl+alt+F1”进入终端，运行以下命令卸载显卡驱动（然后重启即可）：

```
apt-get remove xserver-xorg-video-siliconmotion
```

目 录

① 可信计算与自主可控

② 龙芯简介

③ 数据科学工作环境

- 分析工具
- 编辑器和开发环境
- 排版工具

目 录

① 可信计算与自主可控

② 龙芯简介

③ 数据科学工作环境

- 分析工具
- 编辑器和开发环境
- 排版工具

安装 R

● 安装资源

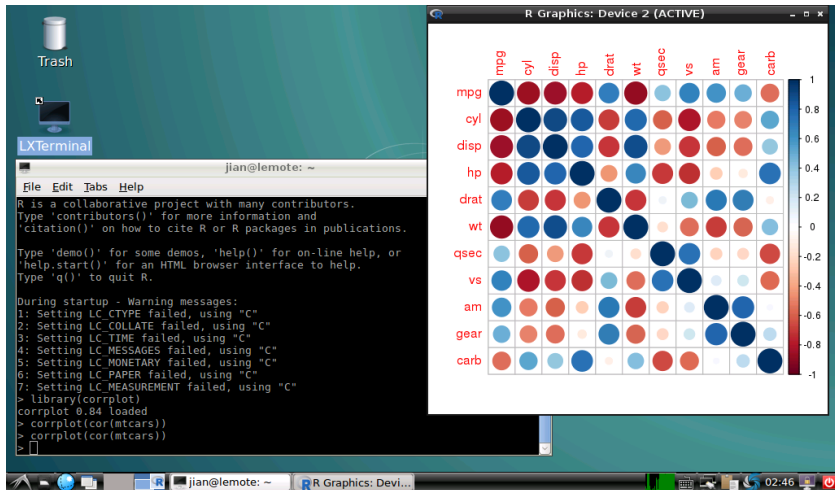
- 很多版本的 Linux 都可以直接通过类似 `apt-get` 的方式从网络源中安装。
- 在本例 Debian8.11 的 8089B 电脑中，直接安装会报错，这是龙芯系统容易出现的问题，通常编译安装可以解决，如果仍然解决不了，稍作修改再编译即可（这是龙芯开发者的重点工作）。
- 在 R 的官网下载某个版本的源文件即可，例如“`R-3.5.3.tar.gz`”。

● 安装步骤

- 首先通过 `apt-get` 的方式安装 `gcc`、`gfortran`、`libbz2-dev`、`liblzma-dev`、`libreadline6-dev`、`libxt-dev` 等依赖库。
- 然后解压缩源文件包，进入文件夹之后运行如下命令：

```
./configure --prefix=~/.R/R-3.5.3 --enable-R-shlib  
make  
make install
```

R 语言示例



目 录

① 可信计算与自主可控

② 龙芯简介

③ 数据科学工作环境

- 分析工具
- 编辑器和开发环境
- 排版工具

常用开发环境

● RStudio

- RStudio 是目前最受欢迎的 R 语言集成开发环境，也支持 Python，是一款主流的数据科学专用 IDE。
- 目前没有直接的龙芯版，需要自行编译。

● Eclipse

- 工程领域流行的 IDE，无论是 Java、C++ 还是 R、Python，基本上都是工程开发的首选。
- 开源免费，功能强大而复杂，基于 Java 环境，可以深入到软件开发的工程管理。
- 可以非常自如地和 Git/SVN 等版本管理工具、ANT/Maven 等构建工具以及很多工程应用的工具整合。

● Emacs+ESS

- Emacs 是公认的编辑神器。
- ESS(Emacs Speaks Statistics) 是 Emacs 下的统计分析模块，集成了 R 和 Julia，提供了一个交互式的开发环境。

安装 Emacs 和 ESS

● 安装 Emacs

- 当前版本的系统包含了 Emacs24 的源，可以直接安装，但打开时会报错。
- 建议编译安装，到官网下载源文件之后进入资源文件夹，然后运行以下代码进行安装：

```
./autogen.sh  
./configure --prefix=/home/jian/lib/emacs24  
make  
sudo make install
```

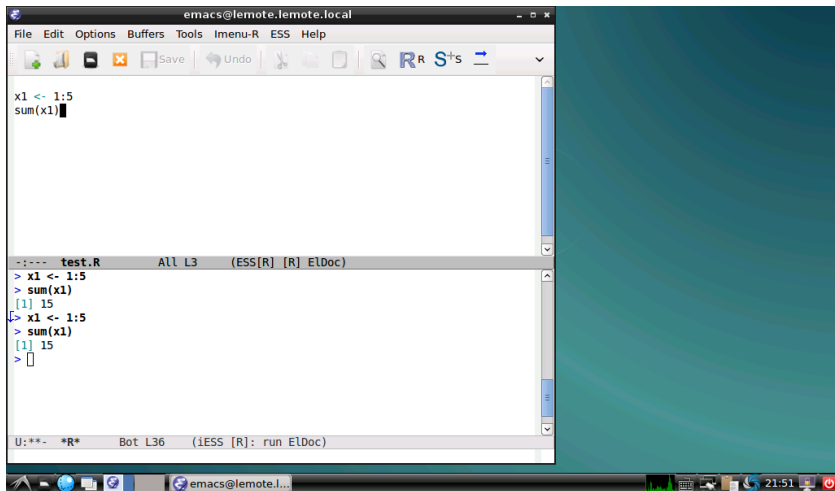
● 安装 ESS

- 直接网络安装也会报错，需要到 ESS 官网下载源文件，解压后进入资源文件夹，然后安装：

```
make  
make install
```

- 安装成功后重启 Emacs，打开后缀名为 **.R** 的文件后会自动高亮显示，并在菜单栏出现 R 操作的相关图标。

ESS 示例



The screenshot displays the Emacs ESS (Emacs Speaks Statistics) interface. The main window shows the R prompt and the following code:

```
x1 <- 1:5  
sum(x1)
```

The output of the R session is shown in the bottom panel:

```
> x1 <- 1:5  
> sum(x1)  
[1] 15  
> x1 <- 1:5  
> sum(x1)  
[1] 15  
>
```

The status bar at the bottom indicates the current buffer is `test.R` and the session is running on `Bot L36`.

目 录

① 可信计算与自主可控

② 龙芯简介

③ 数据科学工作环境

- 分析工具
- 编辑器和开发环境
- 排版工具

L^AT_EX 简介

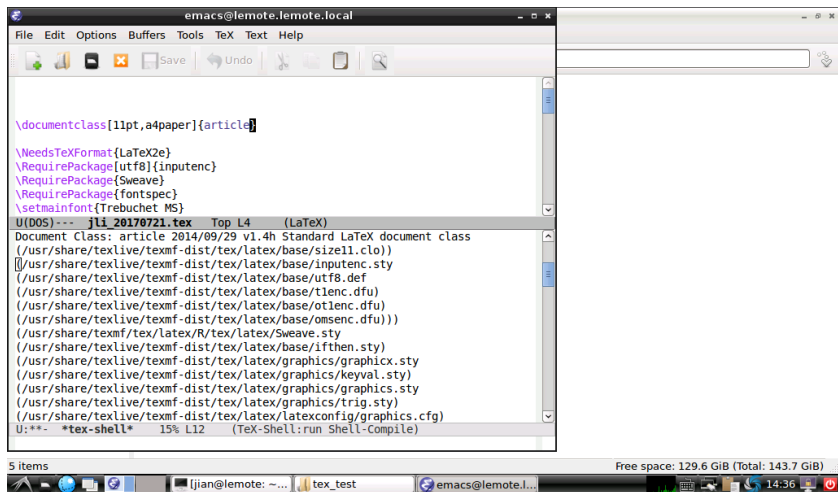
● T_EX

- T_EX 是 Donald E. Knuth 编写的一个以排版文章及数学公式为目标的计算机程序；
- 1977 年，Knuth 开始编写 T_EX 排版系统引擎；
- 现在的 T_EX 系统发布于 1982 年，在 1989 年又稍做改进；
- T_EX 发音为“Tech”，其中“ch”和德语“Ach”及苏格兰语“Loch”中的“ch”类似。

● L^AT_EX

- L^AT_EX 是一个宏集，它使用一个预先定义好的专业版面，可以使作者们高质量的排版和打印他们的作品；
- L^AT_EX 最初由 Leslie Lamport 编写，它使用 T_EX 程序作为排版引擎。现在 L^AT_EX 由 Frank Mittelbach 负责维护；
- L^AT_EX 的发音为“Lay-tech”或“Lah-tech”；
- 在一个 L^AT_EX 环境中，L^AT_EX 充当了图书设计者的角色，而 T_EX 则是其排版者。

TeX Live 示例



The screenshot shows an Emacs editor window titled 'emacs@lemote.lemote.local'. The menu bar includes File, Edit, Options, Buffers, Tools, TeX, Text, and Help. The toolbar contains icons for file operations and editing. The main text area displays LaTeX source code for a document class. Below the text area is a terminal window showing the output of the LaTeX compilation process.

```
\documentclass[11pt,a4paper]{article}

\NeedsTeXFormat{LaTeX2e}
\RequirePackage[utf8]{inputenc}
\RequirePackage{Sweave}
\RequirePackage{fontspec}
\setmainfont{Trebuchet MS}

U(DOS)--- jli_20170721.tex Top L4 (LaTeX)
Document Class: article 2014/09/29 v1.4h Standard LaTeX document class
(/usr/share/texlive/texmf-dist/tex/latex/base/size11.clo))
(/usr/share/texlive/texmf-dist/tex/latex/base/inputenc.sty
(/usr/share/texlive/texmf-dist/tex/latex/base/utf8.def
(/usr/share/texlive/texmf-dist/tex/latex/base/tlenc.dfu)
(/usr/share/texlive/texmf-dist/tex/latex/base/otlenc.dfu)
(/usr/share/texlive/texmf-dist/tex/latex/base/omsenc.dfu)))
(/usr/share/texlive/texmf-dist/tex/latex/R/tex/latex/Sweave.sty
(/usr/share/texlive/texmf-dist/tex/latex/base/iffthen.sty)
(/usr/share/texlive/texmf-dist/tex/latex/graphics/graphicx.sty
(/usr/share/texlive/texmf-dist/tex/latex/graphics/keyval.sty)
(/usr/share/texlive/texmf-dist/tex/latex/graphics/graphics.sty
(/usr/share/texlive/texmf-dist/tex/latex/graphics/trig.sty)
(/usr/share/texlive/texmf-dist/tex/latex/latexconfig/graphics.cfg)
U:** *tex-shell* 15% L12 (TeX-Shell:run Shell-Compile)
```

At the bottom of the Emacs window, the status bar shows '5 items' on the left and 'Free space: 129.6 GiB (Total: 143.7 GiB)' on the right. The taskbar at the very bottom shows the system clock as 14:36 and the Emacs window title as 'emacs@lemote.I...'. The desktop background is orange.

Thank you!