**Scrapy**

**配置管理总结**

小组成员：

郭炜锋

王益飞

胡勇

武丁泽宇

版本变更历史

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 版本 | 提交日期 | 主要编制人 | 审核人 | 版本说明 |
| 1.0 | 20170414 | 郭炜锋 | 王益飞  胡勇  武丁泽宇 | 配置管理总结v1.0 |
| 2.0 | 20170614 | 郭炜锋 | 王益飞  胡勇  武丁泽宇 | 配置管理总结 v2.0 |
|  |  |  |  |  |
|  |  |  |  |  |

# 配置管理总结

小组一共在GitHub上有252次commit。与项目实际相关的commit统计如下1-6章节的表格：

## 会议记录

|  |  |  |  |
| --- | --- | --- | --- |
| 序号 | 文档名称 | Commit时间 | 提交人 |
| 1 | 工作日志 | 20170317 | 郭炜锋 |
| 2 | 工作日志 | 20170320 | 郭炜锋 |
| 3 | 工作日志更新至3.29 | 20170329 | 郭炜锋 |
| 4 | 工作日志更新至4.5 | 20170405 | 郭炜锋 |
| 5 | 工作日志更新至4.7 | 20170408 | 郭炜锋 |
| 6 | 工作日志更新至4.14 | 20170414 | 郭炜锋 |
| 7 | 工作日志更新至4.16 | 20170417 | 郭炜锋 |
| 8 | 工作日志更新至4.18 | 20170419 | 郭炜锋 |
| 9 | 工作日志更新至4.21 | 20170422 | 郭炜锋 |
| 10 | 工作日志更新 | 20170428 | 郭炜锋 |
| 11 | 工作日志更新至5.2 | 20170502 | 郭炜锋 |
| 12 | 工作日志更新至 5.9 | 20170509 | 郭炜锋 |
| 13 | 工作日志更新至5.22 | 20170522 | 郭炜锋 |
| 14 | 工作日志更新至 5.26 | 20170531 | 郭炜锋 |
| 15 | 工作日志更新至6.5 | 20170606 | 郭炜锋 |
| 16 | 工作日志更新至6.9 | 20170613 | 郭炜锋 |
| 共16次commit | | | |

## 工作日志

|  |  |  |  |
| --- | --- | --- | --- |
| 序号 | 文档名称 | Commit时间 | 提交人 |
| 1 | 个人工作日志 | 20170421 | 王益飞 |
| 2 | 个人工作日志 | 20170422 | 郭炜锋 |
| 3 | 个人工作日志 | 20170426 | 胡勇 |
| 4 | 个人工作日志 | 20170426 | 武丁泽宇 |
| 5 | 个人工作日志 | 20170427 | 郭炜锋 |
| 6 | 个人工作日志 | 20170505 | 郭炜锋 |
| 7 | 个人工作日志 | 20170519 | 郭炜锋 |
| 8 | 个人工作日志 | 20170522 | 王益飞 |
| 9 | 个人工作日志 | 20170522 | 王益飞 |
| 10 | 个人工作日志 | 20170522 | 王益飞 |
| 11 | 个人工作日志 | 20170522 | 郭炜锋 |
| 12 | 个人工作日志 | 20170524 | 胡勇 |
| 13 | 个人工作日志 | 20170601 | 郭炜锋 |
| 14 | 个人工作日志 | 20170429 | 武丁泽宇 |
| 15 | 个人工作日志 | 20170505 | 胡勇 |
| 16 | 个人工作日志 | 20170512 | 胡勇 |
| 17 | 个人工作日志 | 20170607 | 王益飞 |
| 18 | 个人工作日志 | 20170602 | 王益飞 |
| 19 | 个人工作日志 | 20170607 | 郭炜锋 |
| 共19次commit | | | |

## 问题清单

|  |  |  |  |
| --- | --- | --- | --- |
| 序号 | 文档名称 | Commit时间 | 提交人 |
| 1 | 项目计划问与答 | 20170317 | 郭炜锋 |
| 2 | 评审意见汇总 | 20170320 | 郭炜锋 |
| 3 | AB组评审单 | 20170329 | 郭炜锋 |
| 4 | 需求意见问答 | 20170329 | 郭炜锋 |
| 5 | 需求互评 | 20170330 | 胡勇 |
| 6 | 评审意见汇总 | 20170330 | 郭炜锋 |
| 7 | AB组评审 | 20170330 | 武丁泽宇 |
| 8 | 评审意见汇总 | 20170330 | 郭炜锋 |
| 9 | 评审汇总单 | 20170409 | 胡勇 |
| 10 | 需求规格说明书V1.3问题清单 | 20170410 | 郭炜锋 |
| 11 | 需求规格说明书V1.6问题清单 | 20170425 | 郭炜锋 |
| 12 | AB组评审单 | 20170425 | 郭炜锋 |
| 13 | AB组评审单 | 20170425 | 胡勇 |
| 14 | 需求规格说明书V1.6问题清单汇总 | 20170426 | 郭炜锋 |
| 15 | 软件需求规格说明书问题清单汇总 | 20170509 | 郭炜锋 |
| 16 | AB组软件测试评审 | 20170524 | 王益飞 |
| 17 | AB组软件测试评审 | 20170525 | 郭炜锋 |
| 18 | AB组软件测试评审 | 20170525 | 胡勇 |
| 19 | 添加测试检查单 | 20170522 | 王益飞 |
| 20 | 需求规格说明书检查单 | 20170508 | 王益飞 |
| 21 | AB组软件测试评审 | 20170601 | 郭炜锋 |
| 22 | 测试需求规格说明书检查单 | 20170607 | 王益飞 |
| 共22次commit | | | |

## 项目ppt

|  |  |  |  |
| --- | --- | --- | --- |
| 序号 | 文档名称 | Commit时间 | 提交人 |
| 1 | 项目计划ppt | 20170317 | 郭炜锋 |
| 2 | 往期ppt | 20170417 | 胡勇 |
| 3 | 需求ppt | 20170328 | 胡勇 |
| 4 | 0428ppt | 20170428 | 胡勇 |
| 5 | 0505ppt | 20170505 | 胡勇 |
| 6 | 0519ppt | 20170519 | 胡勇 |
| 7 | 0526ppt | 20170526 | 胡勇 |
| 8 | 0602ppt | 20170602 | 胡勇 |
| 共8次commit | | | |

## 项目提交文档

|  |  |  |  |
| --- | --- | --- | --- |
| 序号 | 文档名称 | Commit时间 | 提交人 |
| 1 | 项目计划书 | 20170317 | 郭炜锋 |
| 2 | 需求规格说明书v1.00 | 20170320 | 武丁泽宇 |
| 3 | 需求规格说明书 v1.11 | 20170330 | 武丁泽宇 |
| 4 | 需求规格说明书 v1.30 | 20170406 | 武丁泽宇 |
| 5 | 需求规格说明书 v1.40 | 20170408 | 胡勇 |
| 6 | 需求规格说明书 v1.40 | 20170408 | 武丁泽宇 |
| 7 | 管理配置总结说明书v1.0 | 20170414 | 郭炜锋 |
| 8 | 需求规格说明书 v1.50 | 20170416 | 郭炜锋 |
| 9 | 需求规格说明书 v1.20 | 20170404 | 武丁泽宇 |
| 10 | 需求规格说明书 v1.60 | 20170422 | 武丁泽宇 |
| 11 | 需求规格说明书 v1.70 | 20170426 | 武丁泽宇 |
| 12 | 需求规格说明书 v1.80 | 20170428 | 胡勇 |
| 13 | 个人工作总结提交 | 20170502 | 郭炜锋 |
| 14 | 实验三反爬部分改进 | 20170502 | 郭炜锋 |
| 15 | 实验三过滤器改进 | 20170502 | 王益飞 |
| 16 | 需求规格说明书 v1.75 | 20170503 | 武丁泽宇 |
| 17 | 需求说明书修改记录简表 | 20170503 | 武丁泽宇 |
| 18 | 搭建爬虫与测试环境 | 20170503 | 胡勇 |
| 19 | 爬虫改进部分\_基于网页的UI | 20170503 | 胡勇 |
| 20 | 个人工作总结 | 20170508 | 王益飞 |
| 21 | 改进部分2测试需求分析 | 20170522 | 郭炜锋 |
| 23 | 改进部分2测试用例和对照表 | 20170522 | 郭炜锋 |
| 24 | 测试需求规格说明书2.0 | 20170523 | 武丁泽宇 |
| 25 | 扩展测试需求与实践 | 20170523 | 胡勇 |
| 26 | 软件需求规格说明书v1.90 | 20170511 | 武丁泽宇 |
| 27 | 测试需求规格说明书 v1.0 | 20170513 | 武丁泽宇 |
| 28 | 测试需求规格说明书3.0 | 20170529 | 武丁泽宇 |
| 29 | 软件需求规格说明书 v1.95 | 20170529 | 武丁泽宇 |
| 30 | B组部分测试报告 | 20170601 | 郭炜锋 |
| 31 | B组部分测试报告 | 20170601 | 胡勇 |
| 32 | B组部分测试报告 | 20170602 | 王益飞 |
| 33 | B组测试报告汇总 | 20170607 | 郭炜锋 |
| 34 | 软件需求汇总表 | 20170607 | 郭炜锋 |
| 35 | 测试需求及用例汇总 | 20170607 | 武丁泽宇 |
| 36 | 需求汇总 | 20170607 | 武丁泽宇 |
| 37 | 类别汇总 | 20170607 | 武丁泽宇 |
| 共37次commit | | | |

## 项目计划进度

|  |  |  |  |
| --- | --- | --- | --- |
| 序号 | 文档名称 | Commit时间 | 提交人 |
| 1 | 项目计划 | 20170401 | 王益飞 |
| 2 | 项目计划 | 20170404 | 王益飞 |
| 3 | 项目计划 1.2 | 20170406 | 王益飞 |
| 4 | 项目计划 1.2 | 20170407 | 王益飞 |
| 5 | 项目计划 1.2 | 20170416 | 王益飞 |
| 6 | 项目计划进度 1.2 | 20170418 | 王益飞 |
| 7 | 项目计划进度 1.2 | 20170421 | 王益飞 |
| 8 | 项目计划进度 1.2 | 20170426 | 王益飞 |
| 9 | 项目计划进度 1.2 | 20170502 | 王益飞 |
| 10 | 项目计划进度 1.2 | 20170513 | 王益飞 |
| 11 | 项目计划进度 | 20170524 | 王益飞 |
| 12 | 项目计划进度 | 20170601 | 王益飞 |
| 13 | 项目计划进度 | 20170607 | 王益飞 |
| 共13次commit | | | |

## 规格说明书变更

下面表格是对需求规格说明的版本变更：

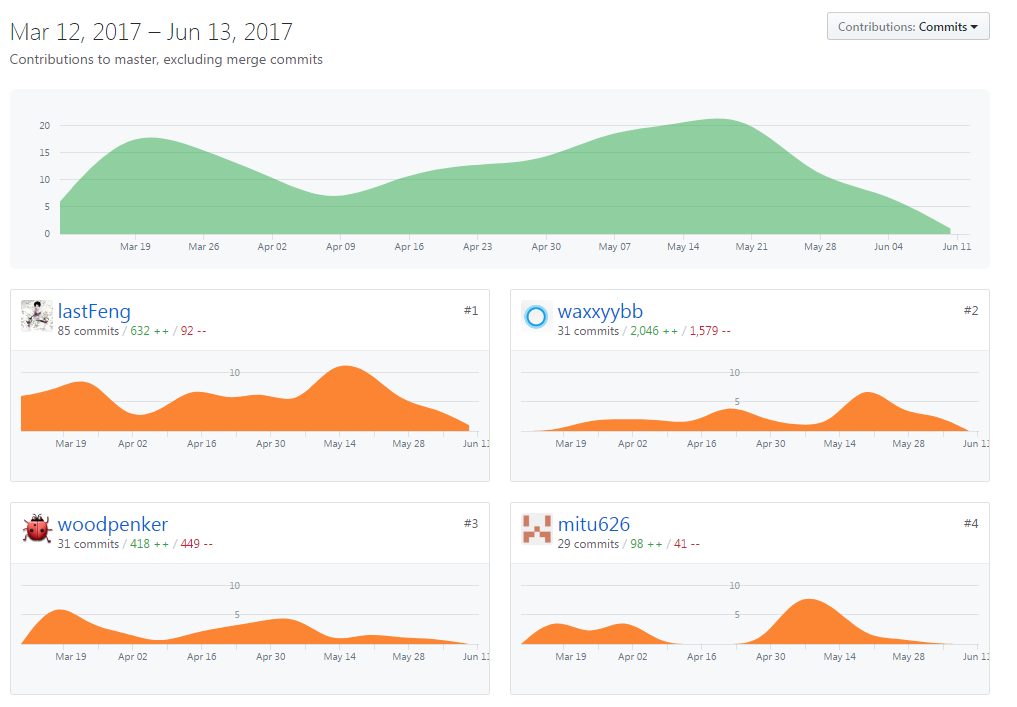
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 版本号 | 日期 | 所修改章节 | 修改说明 | 修改人 | 审核人 |
| 1.00 | 2017.3.21 |  | 完成第一版 | 武丁泽宇 | 胡勇 |
| 1.10 | 2017.3.29 | 4 | 修改第4章内容 | 武丁泽宇 | 胡勇 |
| 1.11 | 2017.3.29 | 4 、5 | 修改4章内容,删除5章 | 武丁泽宇 | 胡勇 |
| 1.20 | 2017.4.4 | 1、2 、3、 4 | 按照修改建议进行修订 | 武丁泽宇 | 胡勇 |
| 1.30 | 2017.4.5 | 3 | 完善业务需求和用例图 | 武丁泽宇 | 胡勇 |
| 1.40 | 2017.4.9 | 全文 | 排版修改 | 胡勇 | 胡勇 |
| 1.50 | 2017.4.16 | 1.3、全文字体、全文格式 | 将1.3定义使用表格表述，将正文的格式统一 | 郭炜锋 | 胡勇 |
| 1.60 | 2017.4.23 | 全文修改 | 根据老师批注进行修改 | 武丁泽宇 | 胡勇 |
| 1.70 | 2017.4.26 | 全文修改 | 根据各组评审意见进行修改和调整 | 武丁泽宇 | 胡勇 |
| 1.75 | 2017.5.3 | 批注章节修改增加第5章 | 按照批注进行修改完善 | 武丁泽宇 | 胡勇 |
| 1.80 | 2017.5.4 | 全文 | 格式化文本格式 | 胡勇 | 胡勇 |
| 1.90 | 2017.5.11 | 全文 | 修改RUCM和用例图 | 武丁泽宇 | 胡勇 |
| 1.95 | 2017.5.29 | 第5章 | 增加改进需求 | 武丁泽宇 | 胡勇 |

接下里啊是对于每个版本变更中的具体原因简表：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 修改位置 | 对应问题 | 提出人 | 修改说明 | 修改说明概要 |
| 1 | 1.引言 | 只说了“功能需求”，但实际上内容包括其他非功能需求等内容 | 刘少凡  黄飞 | 修改 | 将非功能性需求内容单独列出来 |
| 2 | 3.1.1业务需求 | “scrapy设计出了图1所示的框架来满足设计需求”出现了图1，但是图1实际的位置离此文字 ，不方便阅读 | 黄飞  刘少凡 | 删除 | 将此句话删除 |
| 3 | 多处 | 很多自然段没有首行缩进 | 黄飞 | 修改 | 进行文档格式化 |
| 4 | 3.4  3.5 | 出现了很多具体的类名和函数名，在需求分析阶段不应该出现具体的类名和函数名，可用文字描述的方式替代 | 黄飞 | 删除 | 删除这些具体的名称 |
| 5 | 3.3 | 用例图中的用例和RUCM的不能很好的对应，中文名称不一致或者一个是中文另一个是英文 | 黄飞 | 不作处理 | 两者本来就是一个意思，根据个人习惯使用，两者都可以使用，没必要修改。 |
| 6 | 多处 | 字体是方正书宋\_GBK建议使用宋体 | 黄飞 | 修改 | 进行文档格式化 |
| 7 | 第3章 | 没有非功能模块的章节，建议将所有非功能需求集中放在非功能需求的章节中 | 黄飞 | 修改 | 同问题1处修改 |
| 8 | 4.1 | 内存要求建议改为“256MB以上” | 黄飞 | 修改 | 修改此句话表述为“256MB以上” |
| 9 | 文档修改记录 | 表格内容有的居中，有的不居中 | 刘少凡 | 修改 | 同问题3处修改 |
| 10 | 2.1节 | 第一句没有主语，容易引起歧义，建议明确说明是Scrapy的目标，而非你们小组的目标 | 刘少凡 | 修改 | 修改成“scrapy的目标” |
| 11 | 3.1.2 | “其他软件功能需求”建议改为“软件的其他功能需求”以消除歧义 | 刘少凡 | 不作处理 | 此处问题由于问题1处修改而不再存在 |
| 12 | 3.1.2 | 第11、14条等高扩展性的内容是否应该归为“非功能性需求”？ | 刘少凡 | 不作处理 | 此处问题由于问题1处修改而不再存在 |
| 13 | 3.3 | 用例图的图与题注“图2”不挨着 | 刘少凡 | 修改 | 将标注与图放在一起 |
| 14 | 3.5.2 | 该段文字分层了…… | 刘少凡 | 修改 | 将分段符号删除 |
| 15 | 1.3 | 后文出现的一些英文简写在此处没有说明，比如XML、JSON、CSV等 | 刘少凡 | 增加 | 增加后文中出现的英文简写的说明 |
| 16 | 文档修改记录 | 没有审核人 | 温元祯 | 增加 | 增加文档审核人一栏 |
| 17 | 第5页 | 第5页是空白页 | 温元祯 | 删除 | 删除多余的空白页 |
| 18 | 1.引言 | 第二句本文详细描述了功能需求。个人认为需求说明不只有功能需求。 | 温元祯 | 修改 | 将此句话改为“功能、非功能性、业务、拓展和改进需求” |
| 19 | 3.3 | 用例图配图标题位置不对 | 温元祯 | 修改 | 将位置调整一致 |
| 20 | 3.6 | 配图没有图片的标题 | 温元祯 | 不作处理 | 未找到问题出处 |
| 21 | 目录之后 | 多了一页空白 | 李岳檑 | 删除 | 同问题17的处理 |
| 22 | 图2 | 图2的标题放到了文字下面 | 李岳檑 | 修改 | 同问题13的处理 |
| 23 | 多处 | 三级目录字体加粗 | 李岳檑 | 修改 | 进行文档格式化 |
| 24 | 多处 | 用Word打开格式不对 | 谭伟良 | 不作处理 | 我们没有使用word编写，故格式不对，应该以pdf格式为准。 |
| 25 | 首页 | 首页不在同一个文档上 | 谭伟良 | 不作处理 | 我们没有使用word编写，故格式不对，应该以pdf格式为准。 |
| 26 | 正文第4页 | 缩进有问题 | 谭伟良 | 修改 | 进行文档格式化 |
| 27 | 多处 | 没有非功能需求 | 谭伟良 | 修改 | 同问题1处修改 |
| 28 | 多处 | 字体不规范 | 王春柳 | 修改 | 进行文档格式化 |
| 29 | 3.3 | 用例图没有与rucm图一一对应 | 王春柳 | 修改 | 修改相应的RUCM图 |
| 30 | 3.3 | 用例图的画法上存在问题 | 王春柳 | 修改 | 修改用例图 |
| 31 | 第七页 | 框架图说明看着有点乱 | 王春柳 | 修改 | 调整文字格式和位置 |
| 32 | 3.1.2 | 非功能性需求混杂在功能需求里面 | 王春柳 | 修改 | 同问题1处修改 |
| 33 | 多处 | 标点错误以及多处空白行 | 刘超 | 修改 | 相应的错误标点和空白行修改成正确的 |
| 34 | 2.1 | “的”和“地”使用错误 | 刘超 | 修改 | 改正错误的词语 |
| 35 | 3.1.1 | 语句错误“开发人员面临的共同” | 刘超 | 修改 | 修改此处语句为“它” |
| 36 | 3.1.3 | “非功能”部分属于该软件的“功能性”特征 | 刘超 | 修改 | 同问题1处修改 |
| 37 | 3.1.3 | 语句错误：“提供数据导出功能，并提供…” | 刘超 | 修改 | 修改此处语句的表达 |
| 38 | 3.1.3 | 带格式的两端对齐 | 刘超 | 修改 | 进行文档格式化 |
| 39 | 3.1.4 | 需要细化。建议在增加第5章，具体说明 | 刘超 | 新增 | 新增第5章，将原来3.1.4内容移到此处，并增加了详细的设计文字描述。 |
| 40 | 3.2 | 标注scrapy框架图的来源。 | 刘超 | 新增 | 添加scrapy框架图的来源 |
| 41 | 3.3 | 1．“xx中间件”、“xx提取器”：应改为对其”目的“（或功能）的表示，如  调度器(Scheduler)：（页面爬取）任务调度  下载器(Downloader)：页面下载（包括转发给引擎）  2．文件、图像、文字后面都应加上”存储“（动词）。因为”用例“是指一个交互过程（动作）。  3．数据存储与文件存储、图像存储、文字存储之间应该是《包含》还是《扩展》？  4．对于《扩展》，应当定义“扩展点”  5．用例之间的《包含》关系，描述的不是系统的模块之间调用或者“包含”关系。 | 刘超 | 修改 | 同问题30处修改 |
| 42 | 3.3 | 应该对用例图中的每一个用例，采用RUCM模板，给出详细 的定义和描述！  RUCM是对系统（或“模块”）的每一个用例的具体内含的详细描述。  用例名，不应当是系统或模块（组件、中间件）的名字，而应当是对其所要支持或要实现的“目的”的定义，是参与者通过与其得“一次交互”而实现的（参与者需要达到的）一个目的，所以用例的名称通常应当是“动作性的名词（组）”，如爬取数据或数据爬取等。 | 刘超 | 修改 | 同问题30处修改 |
| 43 | 3.3 | 这些模块的用例，与Scrapy（图2）的用例之间的关系？ | 刘超 | 修改 | 同问题29/30处修改 |
| 44 | 3.3 | 所有图都有类似问题。需要重新讨论修改。 | 刘超 | 修改 | 同问题29/30处修改 |
| 45 | 3.3 | 需要认真学习用例模型和RUCM的含义和要求。  1．Brief Description是对“目的”或“交互行为”的描述。  要注意的是，爬虫的使用者的目的是“爬取数据”（用例），而爬虫的开发者的目的是“使用爬虫（Spider）”或者“构建爬虫软件”（用例），所以，他们所用到的用例是不同的，描述自然也是不同的，例如，  用例“爬取数据”的描述可以是：根据爬虫的使用者指定的URL，递归地搜索其页面及其中链接的所有其它URL，获取并解析所有这些页面内容，…（如何终止？）；  用例 “使用爬虫”的描述应当是：使用爬虫（引擎？），以支持爬取指定URL（及其链接）的所有页面数据。  2．Precondition：用例的启动条件，而非“对应模块”的启动条件。  3．Dependency:《include》或者《extend》的用例。如果有，则在steps中，应该使用。  4．Steps：本用例中的步骤，是对spider内部执行过程的描述，而非对上述用例的描述。应该如何描述上述两个用例的事件流？每个步骤，都应当用“简单句“或RUCM规定的句式。  5．Postcondition:在Steps中应当有“读取URL”(或在Preconditio中有对输入参数URL的描述)。Postcondition应当为：对指定URL的数据爬取成功。  6．Specific Alt。 Flow：一般是针对“异常事件处理“的。而”RFS 1“要求的登陆等，属于正常的、必需的步骤”。此外，“RFS 1,2,3,4”，应对不同类型的异常，分别说明。 | 刘超 | 修改 | 同问题29/30处修改 |

## 管理自评

1. 下图是小组四人提交时间以及提交次数的图表：



1. 为了使项目具有合理的组织框架，小组在前人的优秀案例中，将之前不合理的目录结构修改成合理的目录结构，同时添加了相应的说明文档，这也非常感谢小组的王益飞同志的努力：



1. 为了防止版本变更过程中出现命名混乱，我们统一了对本次软件工程综合实验文档的命名方法。

## 总结

在本次实验的配置管理过程中运用的工具是GitHub，GitHub的使用极大提高了小组对文档管理的效率，也提供了对于数据提交的简单统计。但是在使用过程中，由于小组内部的提交中没有及时的沟通好，会导致提交的内容被覆盖或者被误删的情况，在使用过此配置管理工具之后，觉得可以由小组中的其中一个组员进行集中的提交，因为GitHub的提交仅仅只需要几分钟的时间，更多的工作量在于其中文档的编写以及相关的工作。