

The Definitive Guide to Elasticsearch+Logstash+Kibana, Second Edition

ELK Stack 权威指南

第2版

饶琛琳◎编著

Elasticsearch+Lo
5.0版本全面更新

析解决方案，基于ELK

从基础部署到千亿级扩展方案，从性能优化到插件开发，从数据模型到源码解析，全方位解析ELK，融入了作者多年日志分析、数据挖掘的实战经验



机械工业出版社
China Machine Press

数据分析与决策
技术丛书

The Definitive Guide to Elasticsearch+Logstash+Kibana, Second Edition

ELK Stack 权威指南

第2版

饶琛琳◎编著



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

ELK Stack 权威指南 / 饶琛琳编著. —2 版. —北京: 机械工业出版社, 2017.4
(数据分析与决策技术丛书)

ISBN 978-7-111-56329-7

I. E… II. 饶… III. 数据处理软件 - 指南 IV. TP274-62

中国版本图书馆 CIP 数据核字 (2017) 第 050727 号

ELK Stack 权威指南 (第 2 版)

出版发行: 机械工业出版社 (北京市西城区百万庄大街 22 号 邮政编码: 100037)

责任编辑: 吴 怡

责任校对: 殷 虹

印 刷: 北京市荣盛彩色印刷有限公司

版 次: 2017 年 5 月第 2 版第 1 次印刷

开 本: 186mm×240mm 1/16

印 张: 26

书 号: ISBN 978-7-111-56329-7

定 价: 79.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

客服热线: (010) 88379426 88361066

投稿热线: (010) 88379604

购书热线: (010) 68326294 88379649 68995259

读者信箱: hzit@hzbook.com

版权所有 • 侵权必究

封底无防伪标均为盗版

本书法律顾问: 北京大成律师事务所 韩光 / 邹晓东

《ELK Stack 权威指南》第 1 版面世之后的这一年多时间里，ELK Stack 在 Elastic.co 公司以及社区的共同努力下飞速发展。国内外都出现了不少基于 ELK Stack 实现的日志分析产品和创业公司。ELK Stack 已经成为 DevOps 技术栈中必不可少的一部分，较大型的互联网公司甚至已经配备有专职的 ELK Stack 管理团队。

对于并不精通 ELK Stack 技术及其发展历史的人来说，过去复杂的版本对应是新手的第一道门槛。最近全新更新的 ELK Stack 各组件，统一使用 5.x 系列版本号，大大方便了新手入门。而 5.x 系列同样携带了大量崭新的特性，在日志分析、监控告警等场景，带来性能提升、管理简化、功能丰富等诸多好处。推荐广大读者积极尝试和升级。

IT 运维模式正在向数据驱动、精细化、智能化发展。这个过程中，ELK Stack 恰好是运维人员达成这个目的最方便的工具和平台。基于 ELK Stack 平台，越来越多的周边开源项目在涌现。这次再版，也进一步丰富了这些周边项目的介绍。

与第 1 版相比，第 2 版修订、删补了 180 多页内容，接近全书的一半。修改期间，怀孕的妻子一直默默陪伴左右，时不时叮嘱我注意保存。谨以此书献给她和刚出生的启舟宝贝，我爱你们！

本书章节内容

本书包括三大部分共 19 章，各部分可以独立阅读。但对于还没有大规模应用经验的新手，建议按顺序阅读全文。

第一部分 Logstash

第 1 章：入门示例。该章介绍 Logstash 及其插件的配置安装方法，自定义配置语言的设

计用途，并为不熟悉 Linux 系统管理的开发人员介绍了多种后台运行方式。

第 2 章：插件配置。该章列举 Logstash 最常用的几十种插件，通过实际示例和效果，讲解各插件的配置细节和用途。

第 3 章：场景示例。该章以最常见的运维、网络、开发和数据库场景，介绍 Logstash 处理 Nginx、Postfix、Ossec、Log4J、MySQL、Docker 等日志的最佳实践。

第 4 章：性能与监控。了解 Logstash 的性能情况一直是个难题，该章从 Logstash 设计原理和 JVM 平台本质出发，介绍几种行之有效的检测和监控方案。

第 5 章：扩展方案。该章介绍采用 Redis 和 Kafka 完成 Logstash 水平扩展的方案，同时也介绍其他几种日志收集系统与 Logstash 的配合方式。

第 6 章：Logstash 源码解析。该章解析 Logstash 源码中最重要的 Pipeline 设计，以及 Logstash::Event 的来龙去脉。

第 7 章：插件开发。该章以最常见的用户登录记录和地址库解析、Consul 数据更新等需求，实际演示 Logstash 的自定义 Filter、Input 和 Output 插件的编写，同时还涉及了插件打包的 RubyGems 规范共有 HttpClient 功能项等细节。

第 8 章：Beats。该章讲述 ELK Stack 家族新成员 Beats 生态圈各组件的使用，包括 Filebeat、packetbeat、metricbeat、winlogbeat 等内容。

第二部分 Elasticsearch

第 9 章：架构原理。该章从更高级的架构层面，介绍 Elasticsearch 分布式设计中涉及稳定性和高性能的部分原理，并由此引发相关的优化配置介绍。另外，还提供了一种针对时序数据索引的读写分离方案，适用于拥有少部分 SSD 设备的用户。

第 10 章：数据接口用例。该章介绍 Elasticsearch 的 RESTful 接口的基础知识，并针对常见的重建索引需求提供两种快速实现方案，为有 Spark 经验的读者介绍通过 Spark Streaming 接口读写 Elasticsearch 的方法。

第 11 章：性能优化。该章介绍 Elasticsearch 在日志处理场景下的读写优化知识和官方推荐的 curator 工具，其中重点介绍了 Elasticsearch 中几种不同 cache 的区别和有效场景。

第 12 章：测试和扩展方案。该章介绍 Elasticsearch 在生产环境中需要的一些周边工具，比如 Puppet 配置管理、Shield 权限管理、版本升级操作、别名切换流程设计等。新增了快照与恢复功能。

第 13 章：映射与模板的定制。该章详细介绍 Elasticsearch 中的核心类型及其对应的常见映射设置，以及如何通过动态模板简化映射定制操作的复杂度。

第 14 章：监控方案。Elasticsearch 作为一个分布式系统，也是有一定的运维难度的，因此其本身的监控也相当重要。该章介绍 Elasticsearch 自带的一系列监控接口，以及由此衍生的

多种实时或长期的监控方案。

第 15 章：Elasticsearch 在运维监控领域的其他应用。该章介绍 Elasticsearch 在运维方面的其他运用方式，包括实时过滤接口、定时报警系统设计、时序数据存储和相关性排序等。

第三部分 Kibana

第 16 章：Kibana 的产品对比。该章介绍 Kibana 3 与 Kibana 5 之间，以及它们与 Hadoop、Splunk 之间的差异，方便读者在不同场景需求下选择更正确的工具。

第 17 章：Kibana 5。该章介绍 Kibana 5 的安装部署和界面操作方式，重点介绍 Kibana 5 提供的几种可视化图表的配置细节和效果，并以几种场景的日志分析需求演示了 Kibana 5 全新的子聚合功能的效果。最后还介绍了一种采用 phantom.js 截图方式记录长期报表数据的方案。

第 18 章：Kibana 5 源码解析。该章介绍 Kibana 4 的界面实现，重点包括其内部 ORM 实现的 Counrier 类、可视化绘图的 Vislib 类等。

第 19 章：Kibana 插件开发示例。该章讲述 Kibana 最常用的插件类型二次开发实例，包括可视化效果、服务器段进程、完整 App 演示等内容。

致谢

我本人虽然接触 ELK 较早，但本身专于 Web 和 App 应用数据方面，动笔以来得到诸多朋友的帮助，在此深表感谢。此外，还要特别感谢 Elastic.co 公司的曾勇（medcl）和吴晓刚（Wood），曾勇完成 Elasticsearch 在国内的启蒙式分享，并主办 Elasticsearch 中国用户大会，吴晓刚积极帮助新用户，并最早分享了携程的 ELK 日亿级规模的实例。

目 录 *Contents*

前言

第一部分 Logstash

第 1 章 入门示例 3

- 1.1 下载安装 3
- 1.2 Hello World 4
- 1.3 配置语法 8
 - 1.3.1 语法 8
 - 1.3.2 命令行参数 10
 - 1.3.3 设置文件示例 11
- 1.4 插件安装 12
- 1.5 长期运行方式 13

第 2 章 插件配置 15

- 2.1 输入插件 15
 - 2.1.1 标准输入 16
 - 2.1.2 文件输入 17
 - 2.1.3 TCP 输入 18
 - 2.1.4 syslog 输入 19
 - 2.1.5 http_poller 抓取 21

- 2.2 编解码配置 22
 - 2.2.1 JSON 编解码 23
 - 2.2.2 多行事件编码 24
 - 2.2.3 网络流编码 26
 - 2.2.4 collectd 输入 27
- 2.3 过滤器配置 30
 - 2.3.1 date 时间处理 30
 - 2.3.2 grok 正则捕获 33
 - 2.3.3 dissect 解析 35
 - 2.3.4 GeoIP 地址查询 36
 - 2.3.5 JSON 编解码 38
 - 2.3.6 key-value 切分 38
 - 2.3.7 metrics 数值统计 40
 - 2.3.8 mutate 数据修改 41
 - 2.3.9 随心所欲的 Ruby 处理 45
 - 2.3.10 split 拆分事件 47
 - 2.3.11 交叉日志合并 48
- 2.4 输出插件 49
 - 2.4.1 输出到 Elasticsearch 49
 - 2.4.2 发送 email 54
 - 2.4.3 调用系统命令执行 54
 - 2.4.4 保存成文件 55

2.4.5 报警发送到 Nagios	56	第 4 章 性能与监控	85
2.4.6 statsd	58	4.1 性能测试	85
2.4.7 标准输出 stdout	61	4.1.1 配置示例	85
2.4.8 TCP 发送数据	62	4.1.2 使用方式	86
2.4.9 输出到 HDFS	62	4.1.3 额外的话	87
第 3 章 场景示例	64	4.2 监控方案	87
3.1 Nginx 访问日志	64	4.2.1 logstash-input-heartbeat 心跳 检测方式	88
3.1.1 grok 处理方式	64	4.2.2 JMX 启动参数方式	89
3.1.2 split 处理方式	65	4.2.3 API 方式	90
3.1.3 JSON 格式	68	第 5 章 扩展方案	94
3.1.4 syslog 方式发送	69	5.1 通过 Redis 队列扩展	95
3.2 Nginx 错误日志	69	5.1.1 读取 Redis 数据	95
3.3 Postfix 日志	71	5.1.2 采用 list 类型扩展 Logstash	96
3.4 Ossec 日志	72	5.1.3 输出到 Redis	97
3.4.1 配置所有 Ossec agent 采用 syslog 输出	72	5.2 通过 Kafka 队列扩展	98
3.4.2 配置 Logstash	72	5.2.1 Kafka 基础概念	99
3.4.3 推荐 Kibana 仪表盘	73	5.2.2 Input 配置	100
3.5 Windows 系统日志	73	5.2.3 Output 配置	101
3.5.1 采集端配置	73	5.2.4 性能	103
3.5.2 接收解析端配置	75	5.3 logstash-forwarder	103
3.6 Java 日志	77	5.3.1 Indexer 端配置	104
3.6.1 Log4J 配置	77	5.3.2 Shipper 端配置	104
3.6.2 Logstash 配置	78	5.3.3 AIX 上的 logstash-forwarder- java	106
3.6.3 异常堆栈测试验证	78	5.4 Rsyslog	107
3.6.4 JSON Event layout	79	5.4.1 常用模块介绍	107
3.7 MySQL 慢查询日志	80	5.4.2 与 Logstash 合作	109
3.8 Docker 日志	82	5.4.3 Mmexternal 模块	109
3.8.1 记录到主机磁盘	82	5.5 Nxlog	112
3.8.2 通过 logspout 收集	83		

5.6	Heka	114	8.2.1	安装部署	143
5.7	Fluentd	115	8.2.2	配置	144
5.7.1	配置示例	115	8.2.3	生成的可用字段	145
5.7.2	Fluentd 插件	117	8.3	packetbeat 抓包分析	145
5.8	Message::Passing	117	8.3.1	安装部署	146
第 6 章	Logstash 源码解析	119	8.3.2	配置示例	146
6.1	Pipeline	120	8.3.3	dashboard 效果	147
6.2	Plugins	122	8.3.4	Kibana 3 拓扑图	148
第 7 章	插件开发	125	8.4	metricbeat	150
7.1	插件格式	125	8.4.1	配置示例	152
7.2	插件的关键方法	126	8.4.2	各模块输出指标示例	152
7.3	插件打包	127	8.4.3	采集 Docker 中的指标	164
7.4	Filter 插件开发示例	128	8.5	winlogbeat	164
7.4.1	mmdb 数据库的生成方法	129			
7.4.2	LogStash::Filters::Mmdb 实现	130		第二部分 Elasticsearch	
7.4.3	logstash-filter-mmdb 打包	131	第 9 章	架构原理	169
7.5	Input 插件开发示例	132	9.1	准实时索引的实现	169
7.5.1	FileWatch 模块原理	132	9.1.1	动态更新的 Lucene 索引	169
7.5.2	LogStash::Inputs::Utmp 实现	133	9.1.2	利用磁盘缓存实现的准实时 检索	170
7.6	Output 插件开发示例	136	9.1.3	translog 提供的磁盘同步控制	171
第 8 章	Beats	138	9.2	segment merge 的影响	172
8.1	libbeat 的通用配置	138	9.2.1	归并线程配置	173
8.1.1	过滤器配置	138	9.2.2	归并策略	174
8.1.2	输出配置	139	9.2.3	forcemerge 接口	174
8.1.3	shipper 网络配置	142	9.3	routing 和 replica 的读写过程	174
8.1.4	日志配置	142	9.3.1	路由计算	175
8.1.5	运行配置	142	9.3.2	副本一致性	175
8.2	Filebeat	142	9.4	shard 的 allocate 控制	176

9.4.1	reroute 接口	178	11.4.2	shard request 缓存	207
9.4.2	分配失败原因	179	11.4.3	field_stats 接口	208
9.4.3	节点下线	180	11.5	字段数据	209
9.4.4	冷热数据的读写分离	180	11.5.1	Circuit Breaker	209
9.5	自动发现的配置	181	11.5.2	doc values	210
第 10 章	数据接口用例	183	11.6	curator 工具	212
10.1	增删改查操作	183	11.6.1	参数介绍	213
10.2	搜索请求	185	11.6.2	常用示例	214
10.2.1	全文搜索	185	11.7	profiler 调试接口	214
10.2.2	聚合请求	187	第 12 章	测试和扩展方案	217
10.2.3	pipeline 聚合	189	12.1	测试方案	217
10.2.4	搜索请求参数	191	12.2	多集群互联	220
10.3	脚本	192	12.3	puppet-elasticsearch 模块的使用	223
10.3.1	动态提交	192	12.3.1	安装和配置示例	223
10.3.2	固定文件	193	12.3.2	配置解释	224
10.3.3	其他语言	194	12.4	计划内停机升级的操作流程	224
10.4	重建索引	194	12.5	Shield 权限管理	227
10.4.1	Perl 客户端	194	12.5.1	Shield 架构	227
10.4.2	用 Logstash 重建索引	195	12.5.2	安装部署	227
10.4.3	新 reindex 接口的应用	195	12.6	searchguard 权限管理	229
10.5	Spark Streaming 交互	197	12.6.1	安装	229
第 11 章	性能优化	199	12.6.2	权限角色配置	231
11.1	bulk 提交	199	12.6.3	其他组件配置方式	233
11.1.1	bulk 大小	200	12.7	别名的应用	234
11.1.2	UDP 方式	200	12.7.1	索引更名时的无缝切换	234
11.2	gateway 配置	201	12.7.2	限制索引数据部分可读	236
11.3	集群状态维护	202	12.8	快照与恢复	237
11.4	缓存	206	12.8.1	HDFS 插件安装配置	237
11.4.1	filter 缓存	206	12.8.2	Hadoop 配置	238

12.8.3 备份操作	240	14.5 Zabbix trapper 方案	275
12.9 rollover 和 shrink 管理	240	14.5.1 安装配置	275
12.9.1 rollover 管理	240	14.5.2 模板应用	276
12.9.2 shrink 缩容	241		
12.10 ingest 节点	243	第 15 章 Elasticsearch 在运维监控	
12.10.1 创建管道流	243	领域的其他应用	278
12.10.2 测试管道流	243	15.1 Percolator 接口	278
12.10.3 处理器	244	15.2 Watcher 报警	281
第 13 章 映射与模板的定制	246	15.3 ElastAlert	284
13.1 映射的增删改查	246	15.3.1 安装	284
13.2 Elasticsearch 的核心类型	248	15.3.2 配置结构	284
13.3 自定义字段映射	249	15.3.3 扩展	286
13.3.1 精确索引	249	15.4 时序数据库	288
13.3.2 时间格式	249	15.5 Etsy 的 Kale 异常检测	290
13.3.3 多重索引	250	15.6 Grafana 可视化	291
13.4 特殊字段	250	15.6.1 安装	291
13.5 动态模板映射	251	15.6.2 配置数据源	292
13.6 索引模板	252	15.6.3 生成第一个图表	293
		15.6.4 模板功能	295
第 14 章 监控方案	254	15.6.5 在线资源	300
14.1 监控相关接口	254	15.7 Juttle 可视化	301
14.1.1 集群健康状态	254	15.7.1 安装部署	302
14.1.2 节点状态	257	15.7.2 命令行运行示例	302
14.1.3 热线程状态	264	15.7.3 可视化界面	304
14.1.4 索引状态	265	15.7.4 可视化相关指令介绍	304
14.1.5 任务管理	266		
14.1.6 cat 接口的命令行使用	268	第三部分 Kibana	
14.2 日志记录	271		
14.3 实时 bigdesk 方案	272	第 16 章 Kibana 的产品对比	309
14.4 cerebro	274	16.1 Kibana 3 的设计思路和功能	309

16.2	Kibana 5 的设计思路和功能	310	17.6.1	创建一个连接到 Elasticsearch 的索引模式	339
16.3	与 Hadoop 体系的区别	310	17.6.2	字段格式	342
16.4	Splunk 场景参考	311	17.6.3	创建一个脚本化字段	344
第 17 章	Kibana 5	312	17.6.4	设置高级参数	345
17.1	安装、配置和运行	313	17.6.5	管理已保存的搜索、可视化 和仪表盘	345
17.2	生产环境部署	314	17.7	设置 Kibana 服务器属性	346
17.2.1	Nginx 代理配置	316	17.8	常用 sub agg 示例	347
17.2.2	开启 SSL	317	17.8.1	函数堆栈链分析	347
17.3	Discover 功能	318	17.8.2	分图统计	349
17.3.1	设置时间过滤器	318	17.8.3	TopN 的时序趋势图	350
17.3.2	搜索数据	319	17.8.4	响应时间的百分占比趋势图	352
17.3.3	按字段过滤	321	17.8.5	响应时间的概率分布在不同 时段的相似度对比	353
17.3.4	过滤器的协同工作方式	321	17.9	Kibana 报表的快速实现	354
17.3.5	查看文档数据	323	17.10	timelion 应用	355
17.4	各种可视化功能	324	17.11	console 应用	357
17.4.1	area	326	第 18 章	Kibana 5 源码解析	359
17.4.2	table	329	18.1	Kibana 索引的数据结构	360
17.4.3	line	330	18.2	主页入口	361
17.4.4	Markdown	331	18.2.1	Kibana App	362
17.4.5	metric	331	18.2.2	Courier 类	367
17.4.6	pie	332	18.2.3	路径记忆功能的实现	370
17.4.7	tile map	332	18.3	Discover 解析	370
17.4.8	vertical bar	334	18.4	Visualize 解析	374
17.4.9	tagcloud	335	18.4.1	vis_types 实现	375
17.5	仪表盘功能	335	18.4.2	savedVisualizations 实现	382
17.5.1	开始	336	18.4.3	Visualize 实现	382
17.5.2	容器功能	336	18.4.4	VisEditorSidebar 实现	383
17.5.3	修改可视化	337	18.5	Dashboard 解析	384
17.5.4	修改主题风格	339			
17.6	management 功能	339			

第 19 章 Kibana 插件开发示例	388	19.3 服务器端插件示例	394
19.1 Kibana 插件	388	19.4 完整应用开发示例	398
19.1.1 部署命令	388	19.4.1 App 模块的 index.js 结构 ..	398
19.1.2 默认插件	389	19.4.2 服务器端部分	399
19.2 可视化插件示例	390	19.4.3 前台界面的 app.js	399
19.2.1 插件目录生成	390	19.4.4 页面模板	401
19.2.2 主文件及解释	391		



第一部分 *Part 1*

Logstash

- 第 1 章 入门示例
 - 第 2 章 插件配置
 - 第 3 章 场景示例
 - 第 4 章 性能与监控
 - 第 5 章 扩展方案
 - 第 6 章 Logstash 源码解析
 - 第 7 章 插件开发
 - 第 8 章 Beats
- 

Logstash is a tool for managing events and logs. You can use it to collect logs, parse them, and store them for later use (like, for searching). —http://logstash.net

Logstash 项目诞生于 2009 年 8 月 2 日。其作者是世界著名的运维工程师乔丹·西塞 (Jordan Sissel)，乔丹·西塞当时是著名虚拟主机托管商 DreamHost 的员工，还发布过非常棒的软件打包工具 fpm，并主办着一年年度的 Sysadmin Advent Calendar (advent calendar 文化源自基督教氛围浓厚的 Perl 社区，在每年圣诞来临的 12 月举办，从 12 月 1 日起至 12 月 24 日止，每天发布一篇小短文介绍主题相关技术)。

Logstash 动手很早，对比一下，Scribed 诞生于 2008 年，Flume 诞生于 2010 年，Graylog2 诞生于 2010 年，Fluentd 诞生于 2011 年。Scribed 在 2011 年进入半死不活的状态，大大激发了其他各种开源日志收集处理框架的蓬勃发展，Logstash 也从 2011 年开始进入 commit 密集期并延续至今。

作为一个系出名门的产品，Logstash 的身影多次出现在 Sysadmin Weekly 上，它和小伙伴们 Elasticsearch、Kibana 直接成为了和商业产品 Splunk 做比较的开源项目 (乔丹·西塞曾经在博客上承认设计想法来自 AWS 平台上最大的第三方日志服务商 Loggly，而 Loggly 两位创始人都是 Splunk 员工)。

2013 年，Logstash 被 Elasticsearch 公司收购，ELK Stack 正式成为官方用语。Elasticsearch 本身也是近两年最受关注的大数据项目之一，三次融资已经超过一亿美元。在 Elasticsearch 开发人员的共同努力下，Logstash 的发布机制、插件架构也愈发科学和合理。

社区文化

日志收集处理框架很多，如 Scribe 是 Facebook 出品，Flume 是 Apache 基金会项目，都算声名赫赫。但 Logstash 因乔丹·西塞的个人性格，形成了一套独特的社区文化。每一个在 Google Groups 的 Logstash-users 组里问答的人都会看到这么一句话：

Remember: if a new user has a bad time, it's a bug in Logstash.

所以，Logstash 是一个开放的、极其互助和友好的大家庭。如有问题，尽管在 Github Issue、Google Groups、Freenode#logstash Channel 上发问就好！

入门示例

什么是 Logstash？为什么要用 Logstash？怎么用 Logstash？这是本章将要介绍的内容。本章从最基础的知识着手，从以下几步介绍 Logstash 的必备知识。1) 下载安装。介绍 Logstash 软件的多种安装部署方式，并给出推荐的方式。2) 初次运行。通过 Hello World 示例，演示 Logstash 最简单的运用，解释其逻辑上的基础原理。3) 配置语法。介绍 Logstash 的 DSL 设计，Logstash 命令的运行参数。4) 插件安装。灵活和丰富的插件是 Logstash 最重要的优势。本节会介绍 Logstash 插件的安装方式。5) 长期运行方式。从初次终端测试到长期后台稳定运行，本节会介绍几种不同方案，供读者根据实际场景选择。

1.1 下载安装

1. 下载

Logstash 从 1.5 版本开始，将核心代码和插件代码完全剥离，并重构了插件架构逻辑，所有插件都以标准的 Ruby Gem 包形式发布。

下载官方软件包的方式有以下几种：

❑ 压缩包方式

<https://artifacts.elastic.co/downloads/logstash/logstash-5.1.1.tar.gz>

❑ Debian 平台

<https://artifacts.elastic.co/downloads/logstash/logstash-5.1.1.deb>

❑ Redhat 平台

<https://artifacts.elastic.co/downloads/logstash/logstash-5.1.1.rpm>

2. 安装

在上面这些包中，你可能更偏向使用 rpm、dpkg 等软件包管理工具来安装 Logstash，开发者在软件包里预定义了一些依赖。比如，logstash-5.0.2 就依赖于 jre 包。

另外，软件包里还包含有一些很有用的脚本程序，比如 /etc/init.d/logstash。

如果你必须在一些很老的操作系统上运行 Logstash，那你只能用源代码包部署了，记住要自己提前安装好 Java：

```
yum install openjdk-jre
export JAVA_HOME=/usr/java
tar zxvf logstash-5.0.2.tar.gz
```

3. 最佳实践

但是真正的建议是：如果可以，请用 Elasticsearch 官方仓库来直接安装 Logstash！

❑ Debian 平台

```
wget -qO - https://artifacts.elastic.co/GPG-KEY-elasticsearch | sudo apt-key add -
sudo apt-get install apt-transport-https
echo "deb https://artifacts.elastic.co/packages/5.x/apt stable main" | sudo tee
-a /etc/apt/sources.list.d/elasticsearch-5.x.list
sudo apt-get update && sudo apt-get install logstash
```

❑ Redhat 平台

```
sudo rpm --import
https://artifacts.elastic.co/GPG-KEY-elasticsearch
sudo cat > /etc/yum.repos.d/elk.repo <<EOF
[logstash-5.x]
name=Elastic repository for 5.x packages
baseurl=https://artifacts.elastic.co/packages/5.x/yum
gpgcheck=1
gpgkey=https://artifacts.elastic.co/GPG-KEY-elasticsearch
enabled=1
autorefresh=1
type=rpm-md
EOF
sudo yum install -y logstash
enabled=1
EOF
yum clean all
yum install logstash
```

1.2 Hello World

与绝大多数 IT 技术介绍一样，我们也以一个输出“Hello World”的形式开始学习 Logstash。