



Mastering Splunk

精通 Splunk

机器数据处理与分析

[美] James Miller (詹姆斯·米勒) / 著

[中] 宫鑫 谢金秀 郑智超 / 译

使用 Splunk 进行高级分析，有效优化机器大数据

[PACKT]
PUBLISHING

中国工信出版集团

电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
<http://www.ghei.com.cn>

精通 Splunk

机器数据处理与分析

[美] James Miller (詹姆斯·米勒) / 著

[中] 宫鑫 谢金秀 郑智超 / 译



电子工业出版社
Publishing House of Electronics Industry
北京·BEIJING

内 容 简 介

本书内容丰富, 实战性强, 讲解细致, 理论与操作相结合, 详细介绍了 Splunk 的各项功能, 包括如何索引和搜索数据, 如何制作可视化图表, 如何创建仪表板和应用程序, 如何创建警报和数据模型等内容, 任何想通过 Splunk Enterprise 平台进行智能运维的读者都可以阅读本书。

本书力求引用信息的准确性, 但并非毫无瑕疵。对本书所带来的问题或者可能由本书直接或间接引起的问题, 无论是作者, 亦或 Packt 出版社、零售商及分销商均不承担任何责任。

本书中所提到的公司和产品商标信息均由 Packt 出版社提供, 并以大写的形式标出, 但 Packt 出版社不能保证信息的准确性。

Copyright©Packt Publishing 2014. First published in the English language under the title 'Mastering Splunk-(9781782173830)'

本书简体中文版专有翻译出版权由 Packt Publishing 授予电子工业出版社。

未经许可, 不得以任何方式复制或抄袭本书之部分或全部内容。

版权所有, 侵权必究。

版权贸易合同登记号 图字: 01-2015-6974

图书在版编目 (CIP) 数据

精通 Splunk: 机器数据处理与分析 / (美) 米勒 (Miller, J.) 著; 宫鑫, 谢金秀, 郑智超译.
北京: 电子工业出版社, 2016.1

书名原文: Mastering Splunk

ISBN 978-7-121-27676-7

I. ①精… II. ①米… ②宫… ③谢… ④郑… III. ①数据处理软件 IV. ①TP274

中国版本图书馆 CIP 数据核字 (2015) 第 284560 号

策划编辑: 董亚峰

责任编辑: 董亚峰

特约编辑: 田学清 赵海红

印 刷: 三河市华成印务有限公司

装 订: 三河市华成印务有限公司

出版发行: 电子工业出版社

北京市海淀区万寿路 173 信箱

邮编 100036

开 本: 720×1000 1/16

印张: 19.5

字数: 375 千字

版 次: 2016 年 1 月第 1 版

印 次: 2016 年 1 月第 1 次印刷

定 价: 68.00 元

凡所购买电子工业出版社图书有缺损问题, 请向购买书店调换。若书店售缺, 请与本社发行部联系, 联系及邮购电话: (010) 88254888。

质量投诉请发邮件至 zltts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

服务热线: (010) 88258888。

前言

本书不再讲述 Splunk 介绍性的话题，而是从企业建构的角度来介绍一些更深入的概念（结合实例）。本书非常具有实用性，介绍了一个思考型领导者的思维模式，这是所有 Splunk 专家都应具备的。

本书以简单易懂的方式呈现了 Splunk 所有关键的特性，使你能够轻松了解每个特性的功能和工作实例。另外，从企业的角度介绍了进行 Splunk 知识开发的关键概念。

本书涵盖内容

第 1 章 介绍 Splunk 的定义及其在企业建构路线图中的应用。在讨论哪些是该技术的标准或常规用途的同时，会介绍该技术的发展过程。最后，介绍一些 Splunk 的创新用法。

第 2 章 将结合重要的案例讲解高级搜索和技巧。

第 3 章 介绍更深入的 Splunk 表格、图表和字段使用方法，并给出一些实际的案例。

第 4 章 介绍 Splunk 查询与其工作流程。

第 5 章 首先介绍 Splunk 的默认仪表板，进而扩展说明 Splunk 商用高效仪表板的高级特性。

第 6 章 介绍索引的概念、运行方式及其之所以重要的原因。本章将一步一步地分析索引的基本概念及高级概念。

第 7 章 首先介绍 Splunk 应用程序基础知识，之后介绍 Splunk 应用程序和附件方面更深入的内容，如导航、搜索和共享。同时还会给出几个查找附加应用程序源的例子。

第 8 章 从计算机层面介绍 Splunk 技术的监测和预警功能，并将 Splunk 与其他监测工具进行对比。

第 9 章 从企业角度（或企业组织）定义和描述 Splunk 交易。

第 10 章 从企业的视角介绍 Splunk。详细介绍开发的最佳案例、命名、测试及发展前景。

附录 A 快速入门附录中，为读者提供成为 Splunk 大师的各种丰富的资源，从工程师认证系列课程到公司的网站和支持门户，以及其间的任何内容。阅读本部分同时也会让读者了解获取最新 Splunk 版本副本的过程，以及如何进行默认设置安装。最后，本部分会讨论创建有益于个人研究和练习环境的一些想法。

阅读本书的准备工作

如果你没有时间参加正式的培训，或没有时间阅读帮助文件，但仍然想掌握 Splunk 的使用方法，那么本书便是你最佳的选择。你只需要一个有 Windows 系统的笔记本、掌握 Windows 基本的使用技巧及你想研究的数据。

本书适合的阅读人群

本书会介绍许多深刻的见解及详细的、不常见的 Splunk 解决方案案例，因此无论你是否了解 Splunk 的基础知识，本书都能使你成为 Splunk 大师。

体例

在本书中，你会发现许多不同类型的文本，目的是区分不同类型的信息。下面给出了几个不同类型的例子及其所表示的意义。

文本中的编码语言、数据库表名、文件夹名称、文件名、文件扩展、路径名称、虚拟网址、用户输入及 Twitter 名称显示如下：“第一步是编辑 transforms.conf 配置文件，添加新的查找应用函数。”

一个代码块设置如下：

```
[subsearch]
maxout = 250
maxtime = 120
```

```
t11 = 400
```

当我们想将你的注意力吸引到代码块的某个特定部分时，便会将相关的语句或词汇设置成粗体：

```
lookup BUtoBUName BU as "Business Unit" OUTPUT BUName as "Business Unit Name"  
| Table Month, "Business Unit", "Business Unit Name", RFCST
```

指令行输入或输出的写法如下：

```
splunk restart
```

新词和重要的词语用粗体显示，你在屏幕、菜单或对话框中看到的词语会出现在如下文本中：“点击**设置**，然后**检索**。”

警告或要点会在这样的框中显示。

提示或诀窍会这样显示。

读者反馈

我们永远欢迎读者反馈。让我们了解您对本书的看法——不管您喜欢哪里或者不喜欢哪里都请告诉我们。读者反馈对我们继续开发能让您充分受益的书籍至关重要。

要想给我们发送反馈，只需发送邮件到 feedback@packtpub.com，并在邮件主题标注书名即可。

如果你擅长某个项目，并有兴趣编写或参与撰写一本书，请从 www.packtpub.com/authors 查阅我们的作者指南。

客户支持

你已购买了 Packt 图书，我们为你准备了物超所值的福利。

下载本书的彩色图像

同时，我们提供给你一份 PDF 文件，该文件包含本书所使用的屏幕截图或图表的彩色图像。彩色图像会帮助你更好地了解输出内容的变化。你可以从 https://www.packtpub.com/sites/default/files/downloads/3830EN_ColoredImages.pdf 下载该文件。

勘误

尽管我们千方百计确保书中内容的准确性，但错误还是不可避免。如果您在我们的某本书中发现错误，不管是文本错误还是代码错误，向我们报告该错误都会让我们非常感激。您这样做能让其他读者免受挫折，以及帮助我们在随后的版本中加以改进。如果您发现任何错误，请访问 <http://www.packtpub.com/submit-errata> 向我们报告，选择您阅读的书，点击**勘误表提交表格**链接，并输入勘误详情。您提交的勘误表一旦被核实，其将被接受，勘误表会上传到我们的网站，或添加到现有勘误表列表中，勘误列表就在该书名下的勘误表部分。前往 <http://www.packtpub.com/support> 选择您的书名，可以浏览现有的勘误表。

盗版

所有媒体都存在互联网盗版的问题。对 Packt 来说，我们严格保护版权和许可。如果您在网上遇到我公司作品的任何形式的盗版，请马上给我们提供地址或网站名称，以便我们采取补救措施。

请访问 copyright@packtpub.com 联系我们，并给出可疑盗版材料的链接。

我们会感激您协助捍卫作者的知识产权，只有这样我们才能为您提供优质内容。

疑问

如果您有关于本书的任何问题，都可以联系 questions@packtpub.com，我们会尽力解答。

译者序

面对快速变化的业务和日趋复杂的网络应用环境,不少 IT 管理者追赶变化时却举步维艰。**Splunk** 在这个技术的困境中应运而生,以灵活、敏捷的数据管理能力为企业带来一场 IT 管理的创新革命! **Splunk** 可以收集、索引、关联和监控数据,把看似枯燥的机器数据转化为价值非凡的运维智能。通过 **Splunk** 生成的直观、生动的可视化图表,用户可以获得更加深刻的见解,更好地做出决策。

本书内容丰富,实战性强,讲解细致,理论与操作相结合,详细介绍了 **Splunk** 的各项功能,包括如何索引和搜索数据,如何制作可视化图表,如何创建仪表板 and 应用程序,如何创建警报和数据模型等内容,任何想通过 **Splunk Enterprise** 平台进行智能运维的读者都可以阅读本书。

本书的翻译工作由射手学院的译者团队完成,除封面译者外,刘婷婷还承担了本书的部分翻译与审校工作。

本书的专业审校由信通网赢的郑智超完成。信通网赢是全球领先的企业通信与 400 电话服务商,郑智超先生及同事的参与,保证了本书的专业品质。

虽然译者始终谨慎动笔,仔细求证,但难免还会存在疏漏,恳请广大读者批评指正。

目 录

第 1 章 Splunk 的应用	1
1.1 Splunk 的定义	1
1.2 处理通用文件	4
1.3 保密性与安全性	5
1.4 传统应用案例	7
1.4.1 调查研究性搜索	8
1.4.2 监测	10
1.4.3 操作领域的可视化	12
1.4.4 决策支持——实时分析	14
1.5 Splunk——创新应用	17
1.5.1 客户关系管理	17
1.5.2 新兴的技术	17
1.5.3 知识发现和数据分析	18
1.5.4 灾难恢复	18
1.5.5 病毒防护	18
1.5.6 加强结构化数据	18
1.5.7 项目管理	19
1.5.8 防火墙应用程序	19
1.5.9 企业无线解决方案	19
1.5.10 Hadoop 技术	19
1.5.11 媒体评估	20
1.5.12 社交媒体	20
1.5.13 移动设备管理	20
1.6 Splunk 的实际应用	21
1.7 小结	21

第 2 章 高级搜索	22
2.1 Splunk 搜索	22
2.1.1 搜索仪表板	22
2.1.2 新建搜索仪表板	23
2.1.3 Splunk 搜索机制	23
2.1.4 Splunk 快速参考指南	24
2.1.5 请帮助我	24
2.1.6 基本优化	24
2.1.7 快速、详细或智能搜索模式	25
2.1.8 指令的分解	26
2.1.9 了解稀疏和密集搜索之间的差别	26
2.1.10 搜索运算符、指令格式和标签	26
2.1.11 处理流程	27
2.1.12 布尔表达式	28
2.1.13 将我放在“引号”内，否则需要转码	29
2.1.14 在 Splunk 中添加标签	30
2.1.15 事物型搜索	32
2.2 知识管理	33
2.3 子级搜索	35
2.3.1 子级搜索的输出设置	36
2.3.2 Search Job Inspector	37
2.4 使用参数进行搜索	38
2.5 Splunk 宏	39
2.5.1 创建自定义宏指令	40
2.5.2 使用宏	40
2.5.3 Splunk 的限制条件	41
2.6 搜索结果	42
2.6.1 Splunk 基本搜索案例	42
2.6.2 更多的格式选项	43
2.7 总结	43

第 3 章	掌握表格、图表和字段的使用方法	44
3.1	表格、图表和字段	44
3.1.1	汇总成表	45
3.1.2	以图表的形式返回搜索结果	50
3.2	Splunk 目录存放	56
3.2.1	使用 timechart 指令报表	57
3.2.2	timechart 指令需要的参数	58
3.2.3	目录存放时间范围 VS per_*函数	58
3.3	挖掘分析	59
3.3.1	挖掘分析选项	61
3.3.2	基本的分析功能	62
3.3.3	行分析	62
3.3.4	单元格分析	63
3.3.5	图表分析	65
3.3.6	图例	66
3.4	透视表	66
3.4.1	透视编辑器	68
3.4.2	使用透视元素	69
3.5	拆分	70
3.6	Column Values (列值)	71
3.7	数据透视表格式	71
3.8	一个简单的例子	72
3.9	Sparklines (走势图)	74
3.10	总结	76
第 4 章	查询 (Lookup)	77
4.1	简介	77
4.2	配置简单的字段查询 (Field Lookup)	79

4.2.1	在 Splunk 网页端中定义查询 (Lookup)	79
4.2.2	自动查询 (Automatic Lookups)	85
4.2.3	配置文件 (Configuration Files)	88
4.2.4	使用配置文件 (Configuration File) 执行查询 (Lookup) —— 示例	90
4.2.5	填充查询表 (Lookup Table)	91
4.2.6	使用 dedup (去重复指令) 处理 重复值 (Duplicate)	93
4.2.7	动态查询 (Dynamic Lookup)	94
4.2.8	使用 Splunk 网页端	95
4.2.9	使用配置文件 (Configuration File) 替代 Splunk 网页端	97
4.2.10	时基查询 (Time-based Lookup)	99
4.2.11	出现两个一样的值	103
4.3	指令 (Command) 综述	105
4.3.1	lookup (查询) 指令	105
4.3.2	inputlookup (输入查询) 指令与 outputlookup (输出查询) 指令	105
4.3.3	inputcsv (输入 csv) 指令与 outputcsv (输出 csv) 指令	106
4.4	总结	108
第 5 章	先进的仪表板	109
5.1	创建有效的仪表板	109
5.1.1	视图	110
5.1.2	面板	111
5.1.3	模块	111
5.2	表格搜索	112
5.3	返回仪表板	117
5.3.1	面板编辑器	117
5.3.2	可视化编辑器	117
5.3.3	详细学习仪表板编辑器	118
5.3.4	构建仪表板	119

5.3.5	仪表板与 XML	128
5.3.6	给我的世界上色	131
5.4	搜索详解	132
5.5	动态挖掘	137
5.6	真实、即时的解决方案	141
5.7	总结	143
第 6 章	索引库与索引	144
6.1	索引的重要性	144
6.2	什么是 Splunk 索引	145
6.2.1	事件处理	145
6.2.2	索引构成	146
6.2.3	默认索引库	147
6.3	索引、indexers 和集群	147
6.4	Splunk 索引库管理	148
6.5	处理多个索引库	150
6.5.1	创建多个索引的原因	150
6.5.2	创建和编辑 Splunk 索引	150
6.5.3	其他索引方法	153
6.5.4	使用新建索引	155
6.5.5	将所有事件纳入索引	155
6.5.6	导入具体事件	157
6.6	删除索引库及索引的数据	159
6.6.1	删除 Splunk 事件	159
6.6.2	删除数据	162
6.7	配置索引库	166
6.8	移动索引数据库	166
6.9	扩展 Splunk 索引	167

6.10	容量	167
6.11	浅谈限制	168
6.12	总结	171
第 7 章	改进 APP	172
7.1	基础应用程序	172
7.1.1	应用程序列表	173
7.1.2	安装应用程序	176
7.1.3	禁用或删除 Splunk 应用程序	179
7.2	BYO 或创建自定义应用程序	180
7.3	应用程序常见问题解答	180
7.4	Splunk 的端对端自定义	181
7.5	应用程序创建的准备工作	181
7.5.1	开启 Splunk 应用程序创建	182
7.5.2	分享应用程序	199
7.6	总结	199
第 8 章	监测与报警	200
8.1	监测内容	200
8.1.1	Recipes	202
8.1.2	将 Splunk 指向数据	202
8.1.3	监测范畴	203
8.2	高级监测	204
8.3	定位、定位、定位	204
8.4	利用转发器	205
8.5	我能使用应用程序吗	207
8.6	Splunk 中的 Windows 输入	208
8.7	开始监测	209

8.7.1	自定义数据	209
8.7.2	输入端键入	210
8.8	Splunk 用监测到的数据做什么	210
8.9	Splunk	212
8.9.1	该程序在哪	212
8.9.2	安装程序	213
8.10	查看 Splunk 部署监控器程序	216
8.11	关于预警	217
8.12	编辑预警	225
8.12.1	编辑描述	225
8.12.2	编辑权限	226
8.12.3	编辑预警类型和触发机制	226
8.12.4	编辑动作	227
8.12.5	禁用预警	228
8.12.6	复制预警	228
8.12.7	删除预警	229
8.13	Scheduled 或 real time	229
8.14	扩展功能	230
8.14.1	Splunk 加速	231
8.14.2	有效期	231
8.14.3	提要项索引	232
8.15	总结	232
第 9 章	交易型 Splunk	233
9.1	交易及交易类型	233
9.2	交易搜索	235
9.2.1	Splunk 交易案例	236
9.2.2	交易指令	237

9.2.3 交易和宏搜索	238
9.2.4 回顾搜索宏	239
9.3 交易的高级应用	243
9.3.1 设置交易类型	243
9.3.2 汇总——事件集合与关联	247
9.3.3 并发事件	247
9.3.4 避免什么——stats 而非 transaction	251
9.4 总结	253
第 10 章 Splunk-企业板块	254
10.1 基本概念	254
10.2 最佳案例	255
10.3 Splunk 知识的定义	256
10.3.1 数据解释	257
10.3.2 数据的分类	257
10.3.3 数据改进	257
10.3.4 标准化	258
10.3.5 建模	258
10.4 战略知识管理	258
10.5 用 Splunk 知识管理进行对象管理	260
10.6 为文档创建命名规范	261
10.7 测试	264
10.7.1 分享前先测试	265
10.7.2 测试级别	265
10.8 改装	268
10.9 企业视野	269
10.9.1 评估和实施	270
10.9.2 创建、使用和重复	270

10.9.3	管理和优化	270
10.9.4	更多关于愿景的信息	271
10.9.5	一种结构化方法	271
10.10	总结	272
附录 A	快速入门	273
A.1	话题	273
A.2	在哪里学习 Splunk & 如何学习 Splunk	274
A.3	认证	274
A.3.1	信息管理人员	274
A.3.2	管理员	275
A.3.3	建筑师	275
A.3.4	补充认证	275
A.4	Splunk 文件	276
A.5	www.splunk.com	277
A.6	Splunk 答复	278
A.7	Splunk 基地	278
A.8	支持门户	278
A.9	关于 Splexicon	279
A.10	关于“如何”的教程	280
A.11	用户会议、博客及新闻组	281
A.12	专业服务	281
A.13	获取 Splunk 软件	282
A.13.1	免责声明	282
A.13.2	磁盘空间要求	282
A.13.3	安装和调试	284
A.14	在环境中学习	291
A.15	总结	292