



中国Elasticsearch 开发者调查报告

2019年12月

阿里巴巴搜索推荐事业部

阿里云开发者社区 | Elastic技术社区

关注我们



Elastic 微信公众号

关注阿里技术



扫码关注「阿里技术」获取更多资讯

加入交流群



钉钉扫码加入 Elasticsearch 中文技术群

更多电子书



扫码获取更多技术电子书

关注我们



Elastic 中文社区公众号

推荐语

Elastic and Alibaba look forward to hearing from you, our local community. We're always seeking to better support our users and Open Source in China. The Elastic community is a great learning and communication environment. It helps people at all levels to learn Elastic technologies and answer questions.

——Elastic Global Community Advocate Lead, Mark Walkom

拥抱开源，贡献开源，一直是阿里巴巴对于技术社区的承诺。Elasticsearch 是开源技术和应用价值相结合的优秀例子。在阿里巴巴和 Elastic 开源社区的合作当中，我们希望通过数据化地展示调研的结果，不断深化和开源社区的互动，探讨更加有效的合作和支持模式，提供技术开发、行业实践和职业发展的多重价值。

——阿里巴巴开源技术委员会负责人，贾扬清

序言

2012 年 Elasticsearch 首个版本发布，经过 7 年多的更新迭代，Elastic Stack 生态已经日渐成熟，Elastic 软件在国内拥有越来越多的用户，应用 Elasticsearch 的开发人群也不断扩大。

为了深入了解 Elasticsearch 开发者群体的现状，2019 年 11 月，Elastic 技术社区、阿里云 Elasticsearch 技术团队和阿里云开发者社区三方联合发起了 Elasticsearch 开发者调研活动。

在调研过程中，1186 位开发者完成了调研问卷，18 位技术专家完成了专题访问。参与调研的开发者们从不同角度，分享了个人特征、社会属性、技术能力、从业经历等信息；在 Elasticsearch 技术的应用、行业实践、职业发展等方面，忠实的提供了反馈。

作为首次在国内发起的针对 Elasticsearch 开发者群体的行业性调研报告，我们希望能给更多从业者们提供关于自身职业、行业以及技术应用的参照，也希望有更多的开发者关注并进入这一特定技术领域，共同推动 Elastic Stack 相关技术持续向前发展。

目录

| | |
|-------------|-----------|
| 职业篇 | 6 |
| 开发者画像 | 6 |
| 职业发展之路 | 9 |
| 技能进阶之路 | 13 |
| 技术篇 | 19 |
| 技术应用现状 | 19 |
| 最佳实践 | 24 |
| 技术前瞻 | 39 |
| 社区篇 | 43 |
| 开发者与社区 | 43 |
| 关于报告 | 46 |
| 重要说明 | 46 |
| 联合出品人 | 47 |
| 专家团 | 47 |

职业篇

开发者画像

占比 TOP1 的人群，占比 28.6%



男，年龄 26 ~ 30 岁，居住在北京，在望京或西二旗的互联网公司从事后端开发工作，年收入 30 ~ 50 万，其中超 6 成比例了解 Elasticsearch 在 3 年以上，是 Elasticsearch 技术兴起的中坚力量。

一个人在战斗的 Elasticsearch 开发者，占比 24.9%



所在公司从事 Elasticsearch 开发和维护的人员的只有一人的开发者，超 9 成比例在 500 人以下的初创型公司，似乎他们拥有更全能的技术能力，其中的 12% 是全栈开发者，而且他们的年龄大部分都没有超过 30 岁。

35 岁以后的“焦虑”人群，占比 7.3%



“35 岁人群”大多数分布到北京、深圳和上海三个城市，年收入超过 100 万的占 3 成，在大型互联网公司中担任中高层技术管理人员或技术专家职位的超过 6 成。

提交过代码到 Elastic 开源项目的开发者，占比 3.7%



他们是一群拥有丰富 Elasticsearch 经验的开发者，超 6 成的开发者管理着 TB 或 PB 级数据，他们所在的公司也大多数拥有 6 人以上的 Elasticsearch 开发人群。这也反映出，丰富实战经验和团队协作更能快速提升技能。

TOP3 职业特征

| | | | |
|--------------|-------------------|------------------------|------------------|
| 工作城市 TOP3 | 北京 32.5% | 深圳 19.5% | 杭州 10.8% |
| 所属行业 TOP3 | 互联网与软件 78.6% | 金融 9.7% | 教育 2.2% |
| 公司规模 TOP3 | 101 ~ 500人 22% | 1001 ~ 10000人 20.1% | 10000人以上 9.6% |

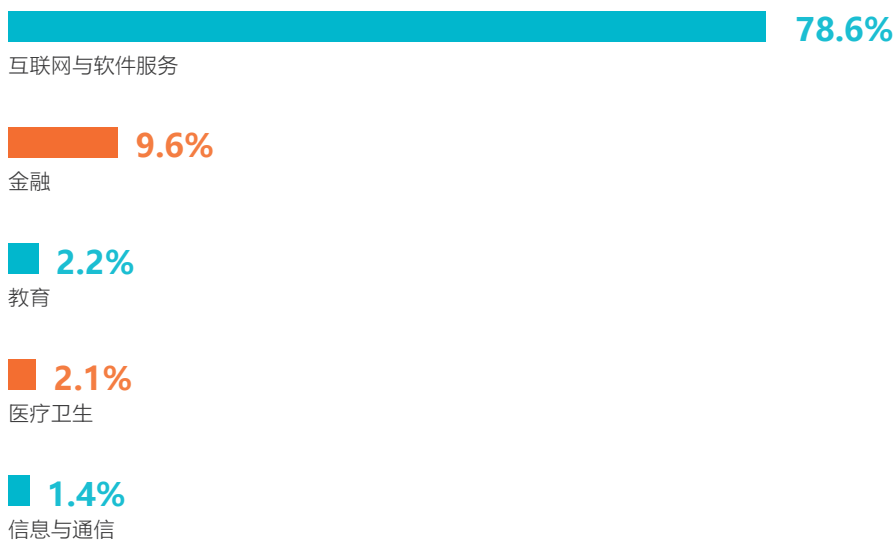
北京是 Elasticsearch 开发者“大本营”

北京是 Elasticsearch 开发者分布最多的城市，其次是深圳和杭州，其他城市的开发者则非常分散。



「互联网与软件业」是开发者密集行业，「金融业」潜力大

Elasticsearch 开发者的行业集中度非常高，「互联网与软件业」以 78.6% 的占比，成为“当之无愧的王者”。随着「金融业」对海量数据搜索分析需求的日益增加，排名第二，有越来越多的金融企业开始实施搜索或日志中台项目。



职业发展之路

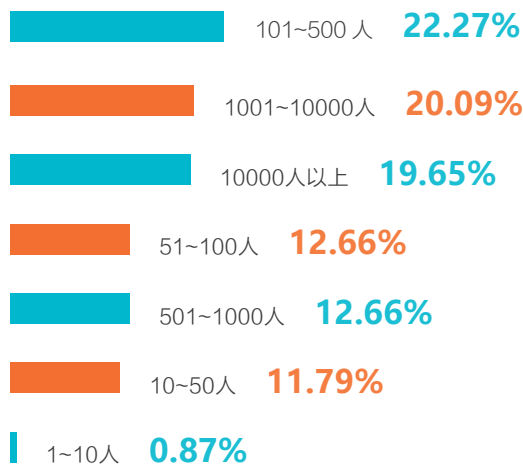
越来越多的企业使用 Elastic Stack

自从 2012 年发布首个 Elasticsearch 版本以来，到现在已形成了完整的 Elastic Stack 技术栈 (Elasticsearch、Kibana、Beats 和 Logstash)。开发者可以利用 Elastic Stack 开发出灵活的软件，能够在搜索、日志记录、安全防护中，实时的大规模使用和分析数据。

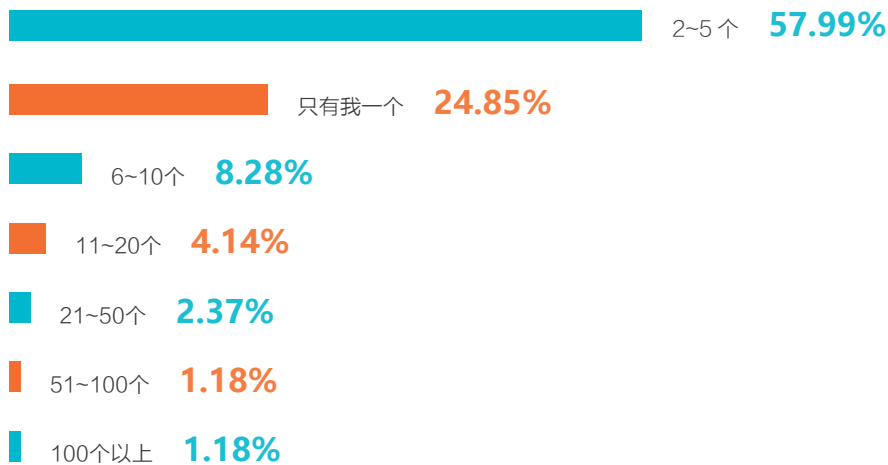
Elastic 技术社区拥有了越来越多的开发者和用户，并在国内的 8 个城市定期开展线下技术沙龙活动。国内的企业用户包括，阿里巴巴、蚂蚁金服、腾讯、字节跳动、滴滴出行、网易、vivo、携程、虎牙、新东方等大量知名企业。

中大型企业的 Elasticsearch 相关工作机会更多

中大型企业拥有更多 Elasticsearch 相关岗位，且能付出更高的薪水，因此吸引到更专业的 Elasticsearch 开发者。



小而精的 Elasticsearch 技术团队



各类型企业中 Elasticsearch 团队呈现小而精的形态。特别是在 500 人以下的初创型企业中，需要 Elasticsearch 开发者成为身兼数职的技术多面手。

来自 Elastic 社区技术大咖的经验之谈

问题 1：与 Elasticsearch 相关工作机会会有哪些？



杨振涛

vivo 搜索引擎架构师

第一类，垂直搜索业务的开发者。就我个人而言，我是从 1.0 版本开始接触，最早做垂直搜索业务时，Elasticsearch 能够帮忙我们快速上手搜索业务。如果想在搜索业务上持续精进，深入掌握 Elasticsearch 是很好的方式。Elasticsearch 是一个成熟的开源项目，底层使用到 Lucene，从它的项目设计和源代码可以学习到很多分布式系统知识，包括各种线程池的设计以及集群多个节点间的协同，也为搜索从业者提供了一个范本；即使业务规模大到一定程度，不再使用 Elasticsearch，但是自主开发的系统原理和 Elasticsearch 也可能是基本一致的。

另一类 Elasticsearch 开发者，已经分化到不同的业务领域，典型的比如安全分析、日志分析、源代码搜索等等。那么他们对 Elasticsearch 的了解是阶段性的，Elasticsearch 作为一个搜索解决方案，从数据索引到检索查询以及最后的排序，都能够很好的支撑。如果开发者希望他的业务有更好的提升，还得持续加深研究，甚至可以考虑考取 Elastic 官方认证，来给自身职业进行加分。

还有一类开发者，当所在公司的 Elasticsearch 集群发展到一定规模后，出于资源使用率和集群稳定性等方面的考虑，可能会考虑自研 Elasticsearch 服务化平台，对所有 Elasticsearch 集群进行管控和资源使用率优化。对于这一类开发者，我们建议他们对 Elasticsearch 底层做更深入的研究，可能需要为业务方做定制化的插件开发，甚至做一些基础研发，替换掉官方的某些模块。

问题 2: Elastic 相关技术领域职业前景如何?



吴晓刚
携程旅行系统研发总监

Elasticsearch 在分布式搜索、实时数据分析、监控和安全等领域有非常多的成功案例。近两年来，国内各大公司基于 Elasticsearch 开发出商业化产品或内部技术平台数不胜数，也越来越成熟。在未来几年 Elastic Stack 也会像目前流行的 Flink 和 Kubernetes 等技术一样，加速发展，应用范围会越来越广。我已经能感受到，相关技术的架构、咨询、开发及运维人员在人力资源市场上的需求正变得越来越旺盛。



戚俊
银杏谷资本总裁技术助理

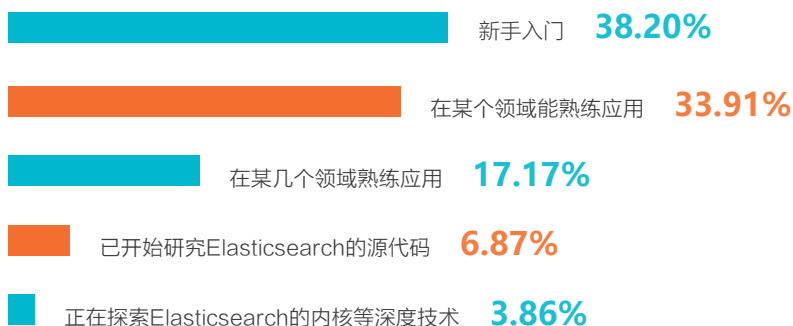
我认为围绕 Elasticsearch 的创业机会会有两条线吧。

一条线朝着做数据分析的角度出发，往上走借助 Elasticsearch 去做数据分析应用。另一条线往下走，帮助企业对接数据到 Elasticsearch。我们发现目前的 Elasticsearch 的数据链路是比较复杂的。一类是都是周期性的去抓非实时数据，比如 Logstash、阿里开源的 DataX 等。另一类是实时数据，现在已经有能力使用不同方法覆盖 Elasticsearch 数据对接的诸多场景，但是缺乏能够覆盖大部分场景的统一数据管道工具。开源产品学习和部署没有成本，但使用和后期维护的隐形成本很高，会逐渐侵蚀掉企业对开源产品的耐性。如果大家围绕这一点去打造开源或商业产品，集成到 Elasticsearch 里，我相信会有很多商业机会。

技能进阶之路

Elastic Stack 技术栈技能掌握情况统计

入门新手、熟练掌握者与领域专家大约各占三分之一。



入门新手



9 成的新手是年龄在 30 岁以下的初中级工程师，从事开发工作在 5 年以下，主要来自后端开发和运维两个技术领域，首次了解 Elasticsearch 在 1 年左右。

熟练掌握者



绝大多数熟练掌握者是高级工程师或技术专家，其中超过 3 成来自于大数据开发领域，超过 4 成使用 Elasticsearch 管理着 PB 级的数据。

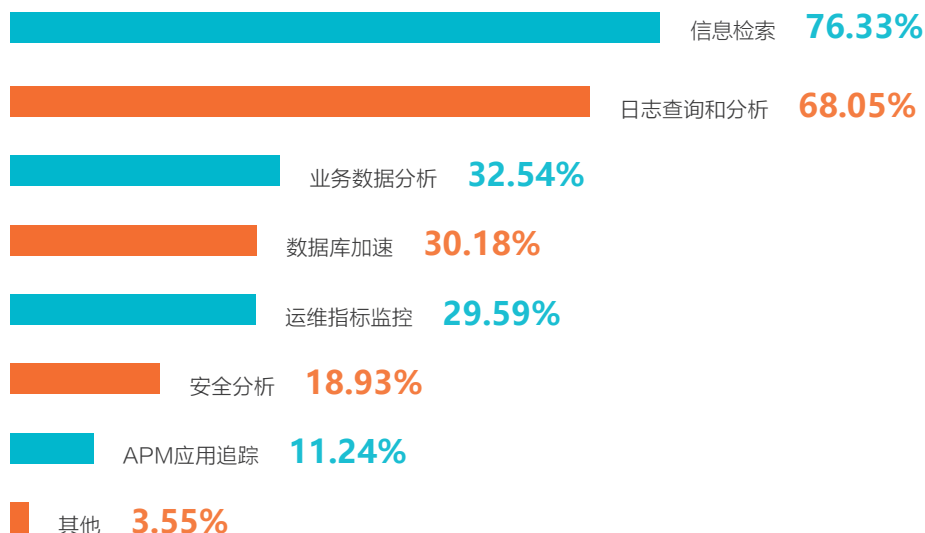
领域专家



超过 4 成的领域专家来自于大型互联网企业，从事的业务领域涵盖后端开发、运维、安全等多个方向。其中的 4% 从 Elasticsearch 第一个版本就开始使用，大约 4 成向 Elastic 社区提交过代码。

Elastic Stack 各个应用领域的技能要求

「搜索」和「日志分析」是最热门的应用领域



典型 Elasticsearch 开发者的技能模型

搜索专家



1. 3 年以上互联网技术研发工作经验，深刻理解搜索业务的核心
2. 熟悉 Elastic 开源产品（如 Elasticsearch、Kibana、Logstash、Beats）的使用方法，了解常见优化方案
3. 熟悉 Elasticsearch / Lucene 工作原理；熟悉 Java，有 JVM 性能调优经验

日志专家



1. 熟悉 Linux 系统、常用 Linux 命令，熟练掌握 Java / Go / Python / C++ 任意一种语言

2. 熟悉 batch / 实时计算领域包括 Hadoop、Spark、Flink 等分布式框架
3. 熟悉索引服务 Elasticsearch & Apache Lucene & Solr 二次开发、应用和调优

运维专家



1. 3 年以上专职 Elasticsearch 运维开发相关工作经验
2. 精通 Elasticsearch 技术体系、架构、设计思想
3. 有提交过 Elasticsearch 相关 patch 或进行功能扩展开发经验
4. 有运维过节点数超 300，每天日志量超 1 亿条经验者

学习 Elastic Stack 的路径

新手如何入门



杨铭毅

CSDN 知名技术博主
《死磕 Elasticsearch》专栏作者

第一：看第一手资料，推荐新手详细阅读官方中、英文文档《Elasticsearch 权威指南》，并在实践中加以运用。

(《Elasticsearch 权威指南》中文版：<https://www.elastic.co/guide/cn/elasticsearch/guide/current/index.html>)

第二：“教是最好的学”，多在 Elastic 中文社区中交流切磋，一方面：多看别人实战中遇到的问题，你的疑惑可能早有人遇到过并给出了详尽的解决方案；另一方面：尝试去回复别人的问题，一点点多互动，时间长了肯定会有提升。也推荐去官方英文社区中去提问，我曾经发现了 Elastic 官方文档的一个 bug，并提交到了英文社区，当天就得到了官方的回复，并且在 7.4+ 版进行了问题修复。

第三：逐步梳理出全局的知识体系、避免“只见树木、不见森林”，对于自己不理解的知识点要“死磕”底层原理，“站在巨人的肩上”，一点点探究直到通透理解。

**阮一鸣**

eBay Pronto 平台技术负责人

《Elasticsearch 核心技术与实战》极客时间技术专栏作者

学习 Elasticsearch 的正确方法是广度优先，而非深度优先。你首先需要了解 Elasticsearch 的所有概念，理解 Elasticsearch 提供了哪些能力去解决相关的问题。不要过早的深入每一个 API 的具体细节，从而避免陷入见树木不见森林的困境。

随后，你需要学习 Elasticsearch 分布式架构的原理。了解 document 在创建，更新以及查询时的内部具体过程；掌握数据建模的相关最佳实践等一系列知识。只有对这些细节有了清晰的认识，后续才能知道如何进行性能优化。

对于新手入门，我在极客时间开设了《Elasticsearch 核心技术与实战》的视频课程，将知识点，最佳实践，项目实战拆解到 100 个短视频中进行讲解。不到半年的时间，已经有 1 万 6 千个订阅用户。一方面可以看出 Elasticsearch 在技术社区的火爆，一方面也可以看到大家对这种学习方式的认可。

关于进阶提高



张超

奇安信高级研发工程师

《Elasticsearch 源码解析与优化实战》书籍作者

针对从事 Elasticsearch 底层技术开发的开发者来说，需要深入了解搜索引擎以及分布式系统的理论。因为 Elasticsearch 自己实现了像 Zookeeper 一样的协调系统，也像 HDFS 一样做了分布式存储，同时还要去做聚合计算，而且还做了类似 HBase 的 LSM，所以它涉及到的知识点还是挺多的。

对于使用 Elasticsearch 的那些业务方来说，更多的是以正确的方式使用 Elasticsearch，这是避免出问题的主要因素。同时熟悉一些其他的查询引擎，了解这些查询引擎擅长做什么，不擅长做什么。当有业务上的查询需求时，知道如何去选择合适的查询引擎。比如，查询需要出很多报表，那么 OLAP 引擎更合适，因为使用 Elasticsearch 做大范围聚合代价大一些。



魏彬

上海普翔 CTO

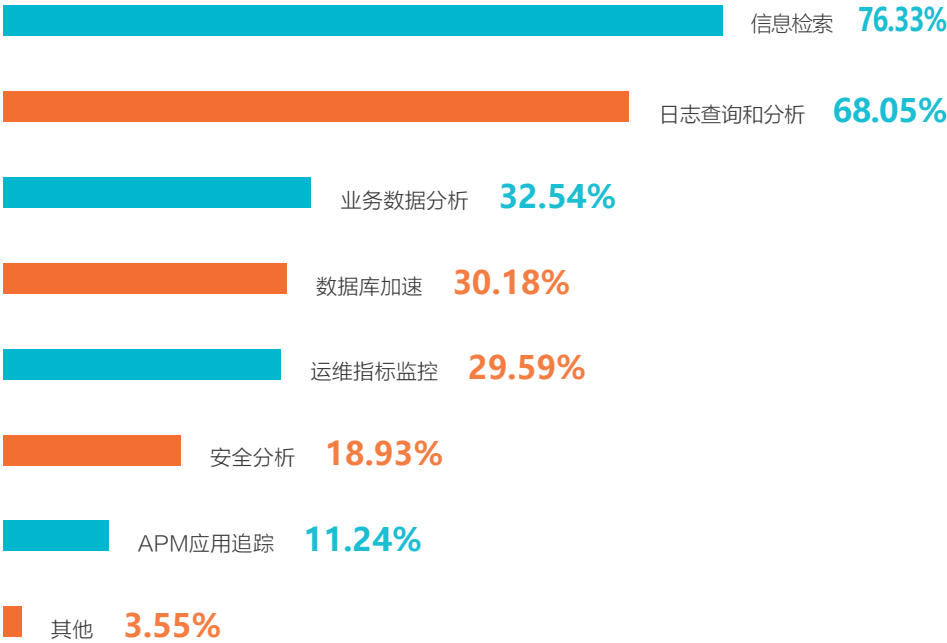
中国首位 Elastic 认证工程师

建议大家去参加 Elastic 认证工程师的考试，它可以让你更系统化地去学习 Elasticsearch 的相关知识。Elastic Certified Engineer 认证考试是非常有难度的，因为它不同于一般的考试，全部题目都是上机的实战操作题目，而且评分是基于 Performance，没有标准答案，同一个题目可能有多种解决方案，但方案之间会有优劣。在 Elastic 产品应用越来越广泛的今天，这无疑会成为你的竞争优势。

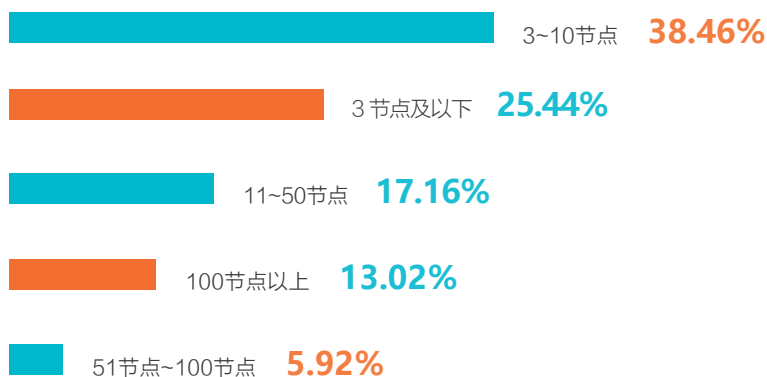
技术篇

技术应用现状

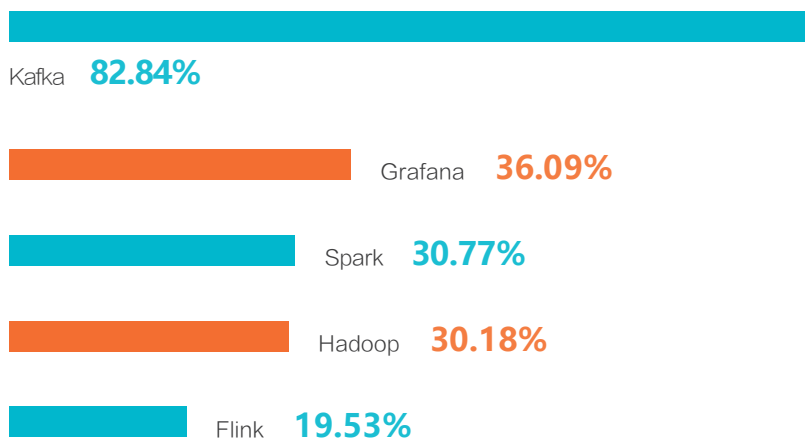
搜索、日志和数据分析是 TOP3 的应用场景



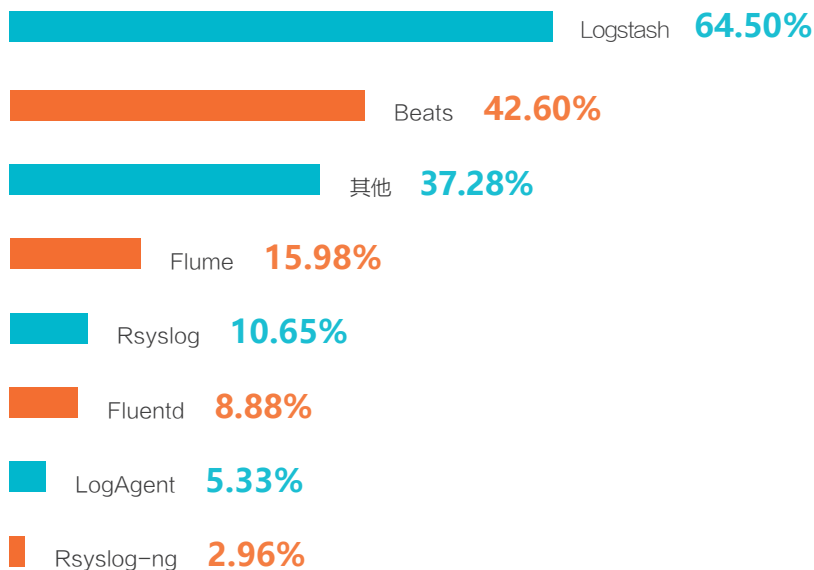
超 6 成开发者维护着 10 节点以下的 Elasticsearch 集群



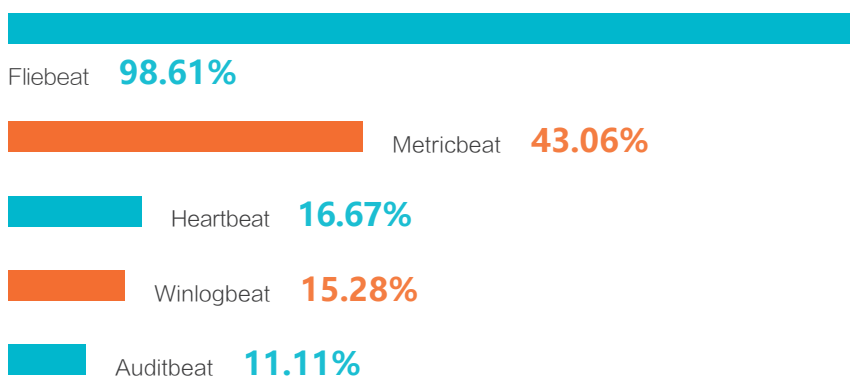
和 Elasticsearch 最“亲密”的技术组件 TOP5，Kafka 是 Elasticsearch 的最佳搭档



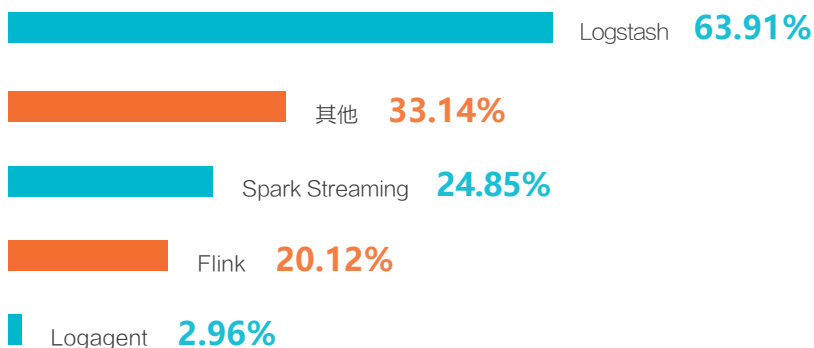
Logstash 是最受欢迎的数据采集工具



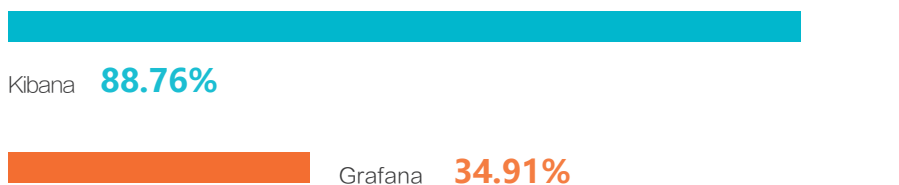
在 Beats 系列里，Filebeat 是最受欢迎的组件



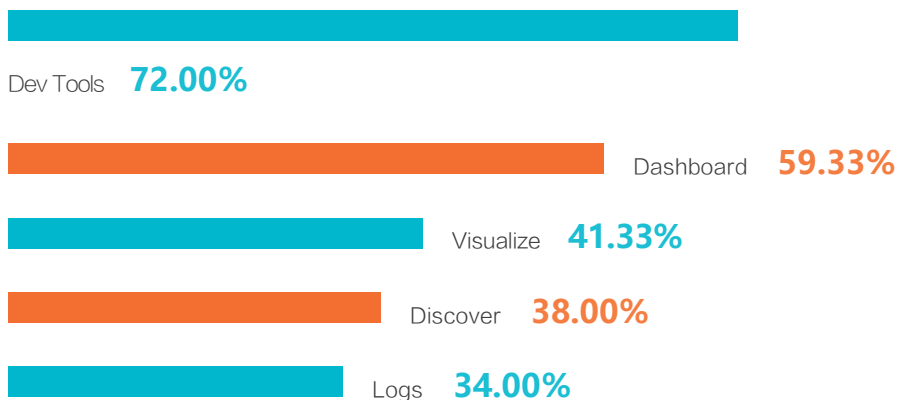
Logstash 依然是最受欢迎的数据处理工具，但 Spark Streaming 和 Flink 的热度上升非常快



Kibana 和 Grafana 是最受欢迎的可视化选择



Kibana 还有待开发者的深度探索，有很多宝藏级功能尚未被发掘



Elasticsearch 上云的趋势明显，特别是使用 Elasticsearch 托管云服务的比例显著，有超 4 成的用户采用了混搭的方式



在物理服务器上自建 Elasticsearch



在公有云虚拟服务器上自建Elasticsearch



使用公有云 Elasticsearch 托管服务



其他

最佳实践

13 位资深开发者讲述了 Elasticsearch 在各自领域的应用实践和技术经验总结。

Elastic Stack 平台化最佳实践

- 阿里云：超大规模 Elasticsearch 云托管服务的内核优化实践
- 携程：从日志分析平台到综合性 Elasticsearch 管理平台
- eBay：超大规模 Elasticsearch 托管平台 Pronto
- 网易：Elasticsearch 通用平台服务，向 SaaS 层面延展
- 虎牙：Elasticsearch 资源统一交付的 PaaS 平台
- 蚂蚁金服：ZSearch 服务中台的内核优化实践

Elastic Stack 在搜索领域的应用

- vivo：Elasticsearch 在手机垂直搜索业务的应用创新
- 贝壳找房：房产交易平台的搜索核心业务实践
- 企查查：Elasticsearch 在企业信息查询领域的应用

Elastic Stack 在安全业务上的应用

- 阿里安全：Elastic Stack 在安全业务中的实践
- 新东方：Elasticsearch 在安全分析领域的应用

Elasticsearch 在数据查询领域的应用

- 同程：Elasticsearch 作为 NoSQL 数据库的替代方案
- 众安保险：基于 Elasticsearch 的海量保单数据查询加速

阿里云：超大规模 Elasticsearch 云托管服务的内核优化实践



郭瑞杰

阿里巴巴资深技术专家

超大规模的 Elastic Stack 公有云托管服务

阿里云 Elasticsearch 的目标是解决云上用户多样化的数据分析和搜索需求，在保证原生 ELK 体验的基础上，针对性优化内核性能，并融入了诸如 Elasticsearch 集群智能诊断、向量检索、达摩院 NLP 分词等独有特性，给予用户灵活、稳定、高性能的服务。

离线部分内核优化：在 / 离线分离技术

ElasticBuild 作为 Elasticsearch 的 BuildService，解决离线全量的性能问题，做到在离线分离。它是基于实时计算引擎的离线索引构建，并利用 NRTCach 实现内存索引合并，以及 Checkpoint 取代 Translog 等内核深度优化技术，使得全量性能相比于在线得到了极大的提升。

在线部分内核优化：计算与存储分离

让集群整体位于同一 VPC 内单可用区（单机房）高速网络环境下，以集群级别实现计算存储分离，整个集群共享一套文件存储服务，用于存放用户的索引数据。每个计算节点有自己的小规格本地存储，用于存放计算节点的本地状态。依赖阿里云文件存储系统，能够保证数据可靠性以及与容量成正比的吞吐性能。并以 plugin 方式集成到社区 Elasticsearch 版本，不会影响现有业务逻辑。借助计算存储分离架构的解决方案，帮助用户至少节省 50% 的存储成本，索引构建性能提升 95.8%，扩容时间从分钟级降低至秒级。

携程：从日志分析平台到综合性 Elasticsearch 管理平台



吴晓刚

携程旅行系统研发总监

平台现状以及未来发展

大约从 2012 年开始，经过调研，我们发现在国外比较活跃的 ELK 技术栈能够简单快速的构建大规模集中式日志平台。因此，我们把它引入到携程。经过多年不断发展，Elasticsearch 在携程不仅用于构建日志平台，还被大量用于各业务线的搜索和数据分析系统开发。比如在搜索方面，可以很好的解决搜索的数据规模化问题，并支持灵活定制，因此在过去几年在携程的垂直搜索业务领域也得到越来越多的应用。目前我们的 Elasticsearch 集群，已经发展到 300 多个，运维问题变得至关重要。我们团队在运维方面拥有较多经验，所以专门组建了两到三人的 Elasticsearch 运维专家组，负责携程 Elasticsearch 集群管理平台的开发，以及线上 300 多个集群的运维和对内技术咨询工作。目前更多专注于底层平台化，未来也会朝着中台化方向进行演进，让业务开发者在搜索系统或数据分析系统使用上更加简单易用。

日志平台的数据治理问题

日志平台存在的最大痛点还是数据治理问题。ELK 做为数据平台，是缺乏数据治理能力的。当上线初期，数据规模还不是很大的时候，数据治理这方面需求还不是很迫切。随着业务发展，接入的日志类型和数据增量快速增加，数据治理的痛点就会逐步凸显。通常 Kibana 无法满足用户所有个性化的数据可视化需求，所以我们还开放了 API 接口给业务方获取数据，去做个性化的分析。API 接口一旦开放，当数据规模达到一定程度，查询、聚合或写入等操作不当的话，经常会对集群性能和稳定性造成较大影响。这些问题都需要通过数据治理的方式来解决。所以，我们在日志平台的基础上做一些日志数据治理工作，比如标准化规范化数据获取方式。

eBay: 超大规模的 Elasticsearch 托管平台 Pronto



阮一鸣

eBay Pronto 平台技术负责人

超大规模的 Elasticsearch 托管平台 Pronto

Pronto 是 eBay 内部 Elasticsearch as Service 的管理平台。Pronto 平台目前管理着 eBay 内部近 200 个不同版本的 Elasticsearch 的集群，节点总规模超过 6000 个。应用场景涵盖了近实时搜索，日志分析，指标监控，风险控制等不同的领域。例如，eBay 收购的市场数据分析平台 terapeak 和 eBay 内部的统一监控平台（UMP）都是采用 Pronto 托管的 Elasticsearch 集群进行构建。对于集群的 SLA，针对网站用户的应用，我们会确保响应时间在 100 毫秒以内。而针对监控类的应用，单个集群每天的数据写入量达到几十个 TB。

跨 Elasticsearch 版本的索引生命周期管理

为了更好的帮助客户对集群进行管理，我们开发了一系列的工具。例如，索引的生命周期管理工具。对于时序性数据，在 Elasticsearch 推出 ILM 之前，社区通常采用 Elasticsearch 提供的 Curator 工具去做索引数据的管理。除了需要针对不同的 Elasticsearch 版本进行开发，在高可用方面，Curator 并没有提供相关的解决方案。通过 Pronto 自研的索引生命周期管理工具，我们为索引预设了不同的生命周期，用户可以通过界面设置不同的 Policy，例如将数据迁移到 warm 节点后将索引设置成 read-only，并进行 force-merge；针对文档总数或分片的尺寸进行 rollup，支持定期对索引进行删除等操作。我们的工具还支持高可用，支持用户以 Self-Service 的方式，对 5.x 到 7.x 不同版本的集群进行索引的管理。

在自动运维和机器学习方向的探索

为了减少运维成本，我们开发了自动化管理平台 LOM。通过不断丰富规则，对集群进行全自动的自我修复。针对用户的一些不正确的使用：例如文档的 Dynamic Mapping 设定导致单个文档字段总数过多，索引分片数据分布不均衡，副本分片设置不当，Pronto Diagnostics Tool 都能够发现这些潜在问题并及时给用户发送错误诊断报告，提供相关的解决方法。另外，我们还对 Elasticsearch 和 Kibana 进行了扩展。例如 Pronto 和 eBay 内部的 AI 团队合作，通过开发插件的方式，集成机器学习能力，为内部用户提供了开箱即用的时序性数据的异常检测能力。

网易: Elasticsearch 通用平台服务, 向 SaaS 层面延展



林邦骏
网易资深工程师

Elastic Stack 平台服务

我们的 Elastic Stack 平台服务包括了全公司不同业务部门和项目的日志采集、数据统一, 并提供统一的管理界面让业务方去管理这些数据。在 Logstash 数据消费方面, 提供集中化的配置管理管理平台, 业务方可以进行诸如配置更新、调试和部署等集中化的操作和管理。在 Elasticsearch 这一部分, 充分考虑了数据摄入的多样性, 为不同部门或项目, 提供一套完整的后台控制系统, 包括 Elasticsearch 节点调度、节点后台操控、组件监控告警、mapping 更新管理等功能。我们还负责 ELK 的外围组件开发。比如 Kibana 的源码修改, 增强权限管理能力, 以符合公司多项目并行管理的需求。

SaaS 化平台

我们的 Elasticsearch 平台服务分为两个部分: 第一个部份是公共部分, 通用的日志采集、处理和展示。第二个部分是 SaaS 化平台, 如果业务方的精细化程度更高, 或者业务需求更明确, 我们会通过 SaaS 化平台提供独立的集群为他们服务, 并提供更完整的部署、监控告警体系。平台上层的业务除了日志分析以外, 还有舆情搜索、安全异常检测、基于 ELK 的告警服务等业务。

对 Kibana 的优化

我们面对的业务方组织复杂繁多, 需要具备索引级的权限管控能力。Elasticsearch 现阶段的权限控制功能, 不是很符合我们的应用场景, 所以我们自己构建了权限控制系统, 比如在 Kibana 上面构建一个可视化面板, 通过项目切换去管理不同的索引。

虎牙: Elasticsearch 资源统一交付的 PaaS 平台



张波

虎牙直播运维研发架构师

Elasticsearch 在虎牙的应用情况

Elasticsearch 在我们这里还是作为一个中台化的平台，帮助多个业务线接入搜索，通过 Elasticsearch 满足在海量大数据场景下面的实时查询需求。

Elasticsearch 资源统一交付的 PaaS 平台

首先我们做了一个 Elasticsearch 资源统一交付的 PaaS 平台，并提供了 Elasticsearch 服务代理层，实现 Elasticsearch 多集群路由和多地数据同步等问题以及多活方案屏蔽底层跨集群数据迁移。另外，我们是有统一日志平台，Elasticsearch 只是其中的一种存储引擎。整个平台我们正规划向微服务下沉。虽然我们使用了 Elasticsearch 代理层，但原生方式去使用 Elasticsearch 的成本还是特别高，不是每一个开发者都了解 Elasticsearch，而且 Elasticsearch 学习门槛相对较高。甚至有些开发者对 Elasticsearch 理解，还停留在 Elasticsearch 是类似 MySQL 的另一种数据库引擎。所以我们中间件团队的愿景是降低用户的使用门槛，用户只需要知道 Elasticsearch 的 API 是如何使用的，不需要去了解 Elasticsearch 本身是什么样的。我们提供给查询模板，简化用户配置，屏蔽查询后边的复杂参数配置。

Elasticsearch 集群迁移做到业务无感知

当一个集群出现问题时，它还不能像 MySQL 那样有比较完备的主备方案，让它很平滑的迁移，集群迁移很难做到对业务无感知。需要先将数据迁移到新的集群，迁移过程不太友好，所以我们开发了类似 Elasticsearch 业务代理层的系统去规避这类问题。

蚂蚁金服: ZSearch 服务中台的内核优化实践



马华标

蚂蚁金服高级技术专家

Elasticsearch 在蚂蚁金服的应用

ZSearch 的设计目标是为广大开发者提供一个云端的 Elasticsearch, 使用者无需关心容量预估、集群部署、参数调优、线上维护等一切与使用无关的事情, 同时又享受到专业的引擎优化与使用指导。

ZSearch 的内核优化实践

在金融类场景, 业务对数据的可用性和容灾能力有着很高的需求, 蚂蚁 Elasticsearch 内核团队为此研发了 Cross Data Center Replication (XDCR), 一款基于 translog 原生复制协议的跨集群同步插件, 目前为数十个核心业务提供深圳 - 上海的异地多活灾备服务, 该插件比社区的 CCR 早半年, 踩过不少坑, 现在还在不断完善中。自动化运维中台“在开发人员较少的情况下, 提供了大规模维护集群的能力, 保证了业务服务过程中集群的容灾、调度、部署、重新发布、监控报警的集中化管理。强化后的内核提供了更为稳定的 API, 典型的应用如蚂蚁金服的官网搜索, 能力服务市场, 安全类的决策日志、舆情分析等。在垂直化搜索能力方面, 独有的向量检索, 以及 Smart Search 实现的用户算法自定义能力, 增加了搜索场景的丰富性。

版本无感知平滑升级

利用 Elasticsearch 本身版本迭代的兼容性规则, 启用自己写的 Block 的方式将老节点上的实时数据灌入到我们的架构队列中的 ESQ, 并启动一个升级后的新节点回灌数据, 不断的通过起一个新节点, 关闭老节点的方式, 实现用户无感知的平滑升级。

新东方: Elasticsearch 在安全分析领域的应用



姜江

新东方运维架构组负责人

Elasticsearch 安全能力的演进—从数据源接入到实时感知

新东方对于 Elasticsearch 安全能力的布局是从 2017 年开始的，主要分为三个阶段；第一阶段是对各个应用及业务线的数据源接入；第二阶段实现单维度指标的异常监测、分析与报警；第三阶段是多指标交叉分析实现对安全攻击、病毒、漏洞等的实时感知。我们实现了数据库审计、前端 WAF、AccessLog、DNS、邮件的数据源接入，随着数据源的完善，通过算法对数据异常的判定，结合机器学习能力，逐步强化业务指标上的报警能力，从而减少安全团队的工作量。最终希望能通过统一的大屏界面实时反馈，清晰的查看到哪个点或者部分出现了安全问题，满足我们对安全能力的初步设想。

Elasticsearch 从 DNS 访问异常发现安全问题

我们通过 DNS 的访问记录发现从外部网盘到内网数据库服务器的可疑流量，随后发现该区服务器上有异常传输数据的痕迹，并在 DNS 上的访问数据中出现大的波动。我们通过搭建 Elasticsearch 的数据安全分析，提前发现并阻止了挖矿病毒的爆发，还是蛮有意思的一件事情。当然还有一些像暴力破解、高频度的修改、查询密码的情况也时有出现，也是通过对“攻击”留下的日志数据，通过 Elasticsearch 分析后得到的结果。

阿里安全: Elastic Stack 在安全业务中的实践



魏超杰

阿里巴巴高级技术专家

Elasticsearch 在高并发安全搜索场景下的应用

安全中的搜索业务是我们最大的使用场景，比如对结构化后的日志数据进行全文检索，对关键状态数据等进行多条件的检索等。我们使用 ELK 构建了业务检测系统，能够在海量数据中查找出伪造终端，而使用 Redis 缓存数据库是没办法解决的。另外在短期内产生大量日志写入的场景，也是使用的 ELK，我们构建了一个高性能的 Elasticsearch 集群，能够支持每秒数十万的写入，支撑大促活动期间的峰值流量写入。

Kibana: QueryBuilder 插件和 Label 插件

在两年前，我们开始了对 Kibana 的改造。当时业务痛点是我们需要将 Kibana 与公司的账户体系和元数据管理能够打通，主要基于应用层面的插件集去改造。随着业务的演进，要求我们对官方特性进行替换，比如鼠标悬停事件改造等。我们在 6.0 版本之上，做了三个比较大的功能：第一个我们构建了 QueryBuilder 插件，用户不需要写 SQL，不需要深入理解数据规则，就可以通过组合条件把数据搜索出来，并且启用了搜索记忆功能，方便用户再次使用，而且还可以记忆存储多个条件，这样对用户易用性有很大提升。这个功能也已经集成到阿里云 Elasticsearch。第二个是打标插件 (Label) 能够可视化的对数据进行修改操作，不需要在控制台写 DSL。对某些搜索出来的数据，希望下次搜索时不再命中，就可以使用 Label 插件，批量打上标签“已处理”。第三个是 Kibana 的千人千面，不同的用户看到不同 Panel，保存不同的搜索条件和 Dashboard。

vivo: Elasticsearch 在手机垂直搜索业务的应用创新



杨振涛

vivo 搜索引擎架构师

Elasticsearch 作为垂搜索引擎的抉择

从 2014 年开始，我们在 vivo 开始做垂直搜索业务，典型场景包括：应用商店、游戏中心、官方商城、铃声、壁纸和主题等，因为手机上有很多内容资源需要分发给用户，搜索是用户获取资源的非常重要的方式。在安卓生态蓬勃发展的时期，有两三百万 APP，搜索功能对用户体验就至关重要，对游戏或 APP 开发者来说，搜索对提升他们内容的分发也是很重要的渠道，搜索在整个平台上的分发量超过了 8 成。所以我们从那个时候开始发力做垂直搜索。现在回过头来看，庆幸当时非常明智的选择了 Elasticsearch，如果当时选择了其他技术，现在很可能会经历一个替换的过程。

在垂搜领域对 Elasticsearch 做定制化改造

vivo 的垂直搜索业务刚开始用 Elasticsearch 时，从 0 到 1 的过程没有太大挑战，很快就能实现垂搜的基本功能。但是一旦实现了这个功能之后，搜索引擎需要持续优化迭代，大家会发现诸如分词效果不好，数据召回率或准确率不够等方面的问题。数据索引、检索以及排序这三个方面都有做深入的干预。因此把 Elasticsearch 作为一个底层快速检索和粗排工具，利用 Elasticsearch 做首轮召回，然后再通过后续其他系统做精排（二次排序），实现更精细的业务逻辑。这是我们遇到的第一大挑战，我相信在业界的垂搜项目，大概率会经历这样一个过程。经历过这样的定制化改造后，就可以根据用户点击行为等评价指标做持续性优化。

贝壳找房：房产交易平台的搜索核心业务实践



高攀

贝壳找房搜索平台负责人

贝壳找房的核心搜索业务介绍

贝壳找房顾名思义就是帮助大家找房子，所以很核心的一点就是房产信息的搜索。大家无论是打开贝壳的首页，还是链家的首页，无论是 PC 站的还是 APP 端的，核心的内容都是二手房、新房、租房的搜索，这些搜索的底层都是目前我们搜索平台所支撑的。除了这几个最核心的业务，搜索平台还支撑了公司其他三百多个业务线的底层搜索，比如客源、推荐等，每天处理查询 10 亿多次。

搜索中台的稳定性优化

业务挑战一个在于稳定性方面：贝壳的 Elasticsearch 使用场景和很多公司不太一样，很多公司的 Elasticsearch 搜索不一定是核心业务，比如用于日志分析、数据分析、运维监控，但我们服务于首页最核心的业务，所以对稳定性的要求非常高，目前我们的 SLA 至少需要满足 99.99%，未来希望能达到 5 个 9，所以我们在稳定性方面做了很多工作，比如 Elasticsearch 集群的双机房互备，主备集群自动熔断降级、主备一键分流，以及网关层也有相应的限流和熔断降级策略等等。另一个是数据同步方面：搜索一般白天都是业务查询的高峰期，但是也有业务会在白天做大批量的数据变更，然后这些变更需要及时的同步到 Elasticsearch，如果同步过快，就会造成 Elasticsearch 集群的压力增加，从而影响到查询延时；如果同步太慢，又会影响到数据的实时性。所以需要我们对实时和批量做区分，并对写入速度进行权衡。我们最终通过使用多个 kafka 消息队列来对数据写入异步化和并行化，同时对生产者和消费者都进行限速。并且对于那些数据频繁变更的业务，我们在写入 Elasticsearch 之前提前进行小批量数据的预合并，从而降低 Elasticsearch 的写入压力，保证查询性能的稳定。

企查查: Elasticsearch 在企业信息查询领域的应用



范兆明

企查查搜索引擎专家

Elasticsearch 在企查查的应用：搜索、日志和安全分析

企查查提供全国企业信息查询服务，截止到今年年初，企查查全球商业数据库涵盖了 200 多个国家，4 亿多个企业主体，并且参与搜索的文档总数超过了 20 亿。

搜索作为企查查的核心入口，目前全部通过 Elasticsearch 完成，高并发、大数据给我们搜索团队带来很大的挑战，Elasticsearch 不仅很好的承载了全文搜索的工作。而且通过 ELK 日志收集平台，准实时的分析用户行为，配合 Redis 等中间件建立起一套完善的防刷机制。通过 Elasticsearch 机器学习的功能，完善了一套搜索热点预警和搜索指数查询平台。

搜索业务上的技术改进

Logstash 数据更新慢，导入效率差，后来是通过 Kafka 队列实时消费解决实时性的问题；并通过自研插件解决了数据量大的问题，提升了数据更新速度。

目前我们搜索指数、搜索热点相关的业务，基于 Elasticsearch 实现类似搜索热点事件的排序。我们通过 Elasticsearch 的一些机器学习基本的基线去做热点的管控，我们通过机器学习的方法，把一些无效的请求剔除掉，获得核心的请求数据，从中找到一些热点指数的数据。

同程：Elasticsearch 作为 NoSQL 数据库的替代方案



刘帅

同程旅游搜索工程师

Elasticsearch Platform as a Service

我们维护了超过 40 个 Elasticsearch 集群，总计 400 多个节点，服务于同程旅游数十个 BU 的 120 多个业务。作为 Elasticsearch 平台服务，为公司的各个业务 BU 提供分布式的结构化与非结构化数据的检索和分析服务。这是一种 PaaS (Platform as a Service) 服务，当业务方明确功能和性能方面的搜索需求后，我们来确定背后需要的 Elasticsearch 服务规模，比如业务方要求 5000 的并发，P99 在 100ms 以内，需要多少机器和何种配置来满足这样的需求。

Elasticsearch 作为 NoSQL 数据库的替代方案

作为 Redis 数据库的兜底方案。作为同程这样的大型 OTA 互联网公司，对 Redis 的依赖是非常强的，但我们并没有把 Redis 作为缓存，而是作为内存数据库来使用，所以会把 Redis 扩展到上 TB 级别。如果数据再往上，就有击穿的可能，我们会把 Elasticsearch 作为最底层的支撑。Elasticsearch 在大数据量的情况下，查询性能和扩展性都很好，并且可以自由变换数据检索的条件组合，是一种高性能 NoSQL 数据库的替代方案。所以大部分的业务 BU 把 Elasticsearch 的高可用特性推到了底层，在上层并没有做过多的架构。

高并发场景下如何保证 P99 的 Latency

在双 11 大促活动等高并发场景下，如果 latency 如果控制得不好，吞吐量肯定是不行的。以同程业务为例，比如微信端首页的服务，我们都是要求 5000 并发，latency 要求 50ms，理论上要控制在 40ms 以内，这对我们这边是一个挑战，需要将 GC 控制在 10ms 以内，留下 10ms 给分片结果汇聚，20ms 给 Lucene 进行底层查询，这样 P100 才能达到 50ms。

众安保险：基于 Elasticsearch 的海量保单数据查询加速



李其霖

众安保险搜索引擎专家

海量保单数据查询的解决方案

因为业务上保单搜索数据量大，会遇到查询性能的问题，业务上需要能在海量的数据下快速的查到结果，要求在毫秒级别从上百亿的保单里找到想要的。查询场景也非常多维度的，从时间维度、聚合、嵌套查询等等，因此性能问题比较重要。我们的解决方案主要是通过冷热分离，冷的数据比如上百亿保单中有部分保单状态是终止的，在业务流程里不大可能被用到，而有效保单也就是还没有终止的保单作为热数据分离出来，在业务流程中用一些更好的配置，提高热数据的查询性能，提升数据查询的吞吐率。

如何保证大规模集群的稳定性

在性能这块之外，之前也遇到诸如稳定性方面的包括网络、故障等问题，导致 Elasticsearch 集群和服务不是很稳定，我们需要实时给客户提供一个稳定的服务。这方面我们使用了阿里云 Elasticsearch 服务，利用多可用区容灾，包括日志监控、网络监控以及一些报警规则的设置，来实时的进行监控，提升整个系统的稳定性。在权限管控这块采用 proxy 的方式，所有的业务方通过代理接进来。在代理之上做了权限、限额和熔断，还包括数据的路由，其实对于业务方不需要知道它的数据在哪个 Elasticsearch 集群上，这样的话也方便做一些迁移。当某个集群不稳定，我们可以及时做数据的迁移从而保证可用性，包括对重要的数据做一些灾备，通过两个集群提供服务，这样一个集群出现问题的时候，服务也不会受到影响。

技术前瞻

Elastic Stack 的技术发展趋势和演进方向



朱杰
Elastic 资深技术顾问

Elastic Stack 的技术演进历程

Elasticsearch 从全文搜索引擎演进成为通用搜索引擎经历了 8 年多的时间，最早期的时候，Elasticsearch 只是一个反向索引，无法实现指标分析等功能，在进行聚合时也会面临性能问题。在后来的演进中增加了列存储，这对于排序和聚合产生了巨大影响，不仅加快了执行速度，还降低了资源消耗。仅仅有这些还远远不够，在 2014 年的时候，Elasticsearch 推出了聚合框架，此时它能够用于代替很多 SQL 分析场景。在 2016 年，Elasticsearch 增加了 BDK 树并且做了稀疏数据的优化，为数字搜索优化了数据结构，加快了分析，并且降低了存储空间。在 2018 年的时候，Elasticsearch 加入了数据上卷特性，将时序性数据聚合到大的时间框内，能够节省磁盘空间。经过这么多年的演进，Elasticsearch 从全文搜索引擎变成了通用的搜索引擎。

卓越的查询能力驱动更加丰富的场景应用

Elasticsearch 的查询能力目前在业界处于领先地位。用户可以将精确查询、全文查询以及地理位置的查询写到一条语句里面进行查询，之后还可以进一步进行聚合计算。正是因为强大的查询能力，Elasticsearch 的应用场景空前繁荣，能够看到最多的是代替了很多数据库查询的加速层，其次还用于用户画像、聚合报表、订单查询以及大数据风控等。

给 Elasticsearch 注入人工智能的灵魂

Elastic 的机器学习能力目前已经赋能到所提供的各种应用之中，比如 APM、业务分析、日志、指标、安全等。无监督的好处在于不需要人工去标注数据，计算和判断的执行速度都会快很多。在 Elasticsearch 7.2 版本和 7.3 版本中，推出了 Data Frame 功能，能够将原始数据的索引经过转换聚合到另外一个更加适合分析的索引中，为后续的机器学习提供了更多的能力。未来，Elasticsearch 也可能提供有监督的机器学习能力。

Beats+Logstash: 更丰富的数据生态

目前，Beats 的生态非常活跃，广大开发者也可以将自己新的想法提交到 Beats 生态中，让更多的人来使用。作为老牌插件的 Logstash 仍然在持续发展当中，未来，Logstash 倾向于成为做数据转化和清洗的中间件。无论是对于 Logstash 还是 Beats 而言，都需要进行集中管理，后续版本中也会提供更多这一方面的能力。

Kibana 的平台化战略

Kibana 的平台化战略会是未来的发展方向，首先 Kibana 会变成一个可视化分析平台，这也是目前已经实现的功能。此外，Kibana 还会演变成一个解决方案的平台，使得用户能够基于 Kibana 构建各种各样的行业解决方案。最后，Kibana 还会成为运维平台，能够监控各种各样的业务指标。

拥抱 Kubernetes

拥抱 Kubernetes 是必然趋势。社区中也有大量的 Elastic Stack 运行在 Kubernetes 之上，因此作为 Elastic 公司而言，提供 Kubernetes 的能力是最为合适的。Elastic 的 ECK (Elastic Cloud on Kubernetes) 目前提供了管理和监控集群、方便升级新版本、弹性伸缩以及定时备份的能力，未来将会进一步拥抱 Kubernetes 生态，期待 ECK 的进一步发展。

Elastic Stack 技术前瞻



吴晓刚
携程旅行系统研发总监

Elastic Stack：更全面的一体化监控运维平台

我觉得 Elastic Stack 未来可能成为整个监控运维领域事实标准。在监控领域，需要收集和处理三类数据：性能数据 (Metrics)、日志数据 (Logs)、应用调用链数据 (Trace)。以往，大家在这三个方面都是使用不同产品去做的。比如性能指标数据，传统行业使用 Zabbix 比较多，互联网使用新兴的 Prometheus、OpenFalon 等。而日志方面，又自成一体，大家有不同的方案，比如商业产品 Splunk、开源 ELK。APM 工具更是百花齐放。这三类数据通常相互割裂互成孤岛，而实际上，我们线上运维通常希望这些数据能够打通和关联起来。Elastic 有个比较好的思路，借助 Elastic Common Schema (ECS) 等规范化工具，将性能数据 (Metrics)、日志数据 (Logs)、应用调用链数据 (Trace) 三类数据标准化，整合到一起，形成一体化的监控解决方案。运用这样的产品，运维工程师可以非常自然的在不同数据之间进行跳转和关联分析。目前 Elastic 已经陆续推出了 Metrics、APM 等方案，但在规模化应用场景下还是存在底层存储成本方面的问题，对于海量 metric 和 trace 数据，存储代价还是很高的。当然 Elastic 也作出了努力，比如针对 metrics 推出了 Rollup 功能，类似准实时的流式计算，在数据写入阶段通过预置规则，对数据做实时聚合后再存储，可以实时把数据做聚合后再存储，以减少存储开销。当未来 Elastic 能够更好的解决数据存储成本问题后，可能成为优秀的一体化监控运维产品。另外，Elastic 引入了机器学习，对监控数据做智能化的异常检测和预测，顺应了监控系统 AI 化的发展趋。

**戚俊**

银杏谷资本总裁技术助理

Elasticsearch 作为数据分析底层引擎

我认为 Elasticsearch 有三个很鲜明的特性：首先非结构化数据存储的理念，注定它可以对接无数种数据源。其次得益于检索性能，Elasticsearch 在分析领域有很好的表现。再次，有很多不太重要的数据，还没有必要进入到数据库进行存储，但又需要存储到某个地方，并且在某个场景下能够及时方便的用起来，而不是压缩成文件放在 OSS 或硬盘上。三个特性就决定了它的应用场景会非常广泛。比如大家会认为 Elasticsearch 是数据库的补充，在 MySQL 上加一层 Elasticsearch，既不破坏原有业务场景和程序链路，但又能显著提升业务性能和实时性。我认为 Elasticsearch 是有成为数据分析底层仓储性质组件的能力的。特别是 Elasticsearch 具备分布式和高可用特性的，利用它做底层存储承载去做分析，是非常有前景的。

**魏超杰**

阿里巴巴高级技术专家

SIEM 等安全领域是 Elastic Stack 天然最佳场景

Elastic 最近发布了 SIEM 和 Endpoint Security 等安全方案，我觉得 Elastic 在安全领域开发出更多的解决方案是很合理的，这也是它最适合的场景，安全是企业普适性的需求。我们也希望官方能推出更多新的安全特性和成熟应用让我们去了解和使用的。我们希望 ELK 拥有更成熟的安全框架方案，借助他们可以快速搭建安全业务。

社区篇

开发者与社区

参与调研的开发者参与社区活动的情况

 **42.06%**

偶尔到 Elastic 中文社区浏览或发帖

 **35.19%**

参加过Elastic中文社区线下Meetup

 **26.61%**

经常到 Elastic 中文社区浏览或发帖

 **20.17%**

从未参加过任何社区活动

 **11.59%**

参加过年度 Elastic 开发者大会

Elastic 中文社区

Elastic 中文社区是由 Medcl (曾勇) 于 2013 年 1 月创建的社区交流网站 (<http://elasticsearch.cn>)。包括了「提问讨论」、「活动分享」、「资料下载」等版块。下面是社区上的第一个帖子：



《Elasticsearch权威指南》中文版开始翻译

本来想自己翻译这本书的，但是发现章节的数量已经超出能力范围，每天时间也有限，在不断的努力以及几个热心人的参与下，完成了6章的翻译。如果你对Elasticsearch感兴趣或者在工作中有用，那就关注我们的翻译。有兴趣的也可以fork后提交 Pull Request。当然，我觉得前6章入门已经够用，如果你想学习也可以关注我们的翻译。

Github地址：<https://github.com/looly/elasticsearch-cn>

[继续阅读 »](#)

[Elasticsearch权威指南](#)

looly 发表于：2015-03-10 09:32 [评论 \(17\)](#)

Elastic Meetup

Elastic Meetup 由 Elastic 官方在技术社区中所支持的、定期举办的线下交流活动，主要围绕 Elastic 的开源产品（Elasticsearch、Logstash、Kibana 和 Beats）及周边技术，探讨在搜索、数据实时分析、日志分析、安全等领域的实践与应用。迄今已成功举办了 30 多次，超过 130 位嘉宾进行了精彩分享，超过 5000 人次到现场参与。

和区域社区负责人聊聊社区共建



杨振涛

vivo 搜索引擎架构师

Elastic 深圳社区负责人

线下 Meetup 的目标受众

来听分享的人以新手和入门者居多，分享者大多是非常有经验的，所以我们希望实现两边的价值交换和双向互动，因为对于分享者来说，他也希望了解到别人在类似的问题上的处理经验，或者说类似实践上的更多解决方案。

区域社区的运营经验分享

一方面我们会获得国外社区的一些经验，包括国外社区讨论的内容和组织形式，另一方面国内官方的技术布道师也会给予我们很多支持，包括官方的小礼品和最新宣传资料等；此外，我们的本地化组织还推崇以运营思维做社区治理，我们希望明确活动的人群定位和价值输出定位，比如说我们定位是给 Elasticsearch 相关的开发者提供一个交流的平台，那就要尽量迎合开发者想交流这一核心诉求。比如策划一次本地的 Meetup 活动，我们前期需要有统一的信息发布，中间需要进行持续的、有节奏的、有计划性的传播和推广，最后现场有很好的实施和以及后期完整的活动总结，这样以来活动的连贯性会比较好，在活动的用户体验方面让大家会感受到社区工作的精细和用心程度，对社区的信任度以及好感度提升，并且我们会在活动中引导和呼吁大家，社区成为什么样子，是由我们自己来决定的，大家可以积极的参与进来，比如部分开发者所在公司也会对活动提供一些力所能及的支持，包括提供抽奖的礼品、会务志愿者以及活动场地等。

关于报告

重要说明

- 本报告问卷数据调研时间为 2019 年 11 月 16 日至 2019 年 12 月 3 日。
- 本次调研分为定量研究、定性研究以及专家访谈，所有调研样本均为开发者本人和相关从业者，总计成功回收 1186 份，并完成对 18 位 Elastic 技术社区专家的专题采访。
- 本报告中提及的开发者 and 受访者均为中国 Elasticsearch 相关技术的开发者。
- 本报告版权归 Elastic 技术社区和阿里云开发者社区所有，如有相关引用需获得授权。
- 感谢在 Elastic 技术社区和阿里云开发者社区参与问卷调研的开发者们，以及接受采访的社区专家。

联合出品人

Elastic 技术社区 刘征

阿里巴巴搜索推荐事业部 郭瑞杰

阿里云开发者社区 郭雪梅

专家团 (按姓氏排名)

邓万禧 (阿里巴巴)

高 攀 (贝壳找房)

姜 江 (新东方)

李其霖 (众安保险)

刘 帅 (同程)

潘禹丞 (阿里巴巴)

马华标 (蚂蚁金服)

戚 俊 (银杏谷资本)

是 溪 (阿里巴巴)

魏 彬 (普翔科技)

吴晓刚 (携程)

王 刚 (Elastic)

杨振涛 (vivo)

张 波 (虎牙直播)

朱 杰 (Elastic)

范兆明 (企查查)

葛丽丽 (Elastic)

郭宏娟 (阿里巴巴)

林邦骏 (网易)

李若冰 (阿里巴巴)

杨铭毅 (Elasticsearch 技术博主)

钱雨欣 (阿里巴巴)

阮一鸣 (eBay)

魏超杰 (阿里巴巴)

吴世龙 (阿里巴巴)

王佳玲 (阿里巴巴)

曾 勇 (Elastic)

张 超 (奇安信)

赵弘扬 (阿里巴巴)

关注我们



Elastic 微信公众号

关注阿里技术



扫码关注「阿里技术」获取更多资讯

加入交流群



钉钉扫码加入 Elasticsearch 中文技术群

更多电子书



扫码获取更多技术电子书

关注我们



Elastic 中文社区公众号

