

11 | 谷歌SRE运维模式解读

2018-01-12 赵成

赵成的运维体系管理课

[进入课程 >](#)



讲述：黄洲君

时长 08:24 大小 3.85M



前面我和你分享了一些关于运维组织架构和协作模式转型的内容，为了便于我们更加全面地了解先进的运维模式，今天我们再来谈一下谷歌的 SRE（Site Reliability Engineer）。同时，也期望你能在我们介绍的这些运维模式中找到一些共通点，只有找到这些共通点，才能更深刻地理解，并借鉴到真正对我们有用的东西。

专栏的第一篇文章我们介绍了 Netflix 的 NoOps 模式。这个模式并不意味着不存在任何运维工作，只是 Netflix 将这些事情更紧密地融入到了日常的开发工作中，又做得非常极致，所以并没有很明显地体现出来。

但是，谷歌的 SRE 却是一个真实具体的岗位，也有明晰的岗位职责。从借鉴意义上来讲，SRE 可以给我们提供更好的学习思路和样板。

SRE 这个概念，我应该是 2014 年下半年的时候听到的。当时可接触的资料和信息有限，只知道是谷歌对运维岗位的定义，负责稳定性保障，就没有更多其他的认识了。

后来，有越来越多在谷歌工作或接触过这个岗位的专家开始在公开演讲中分享这个概念。同时，《SRE: Google 运维解密》，这本由多名谷歌 SRE 亲笔撰写的图书也开始在国内广泛流传，让我们对很多细节有了更加细致的了解。

SRE 岗位的定位

首先，SRE 关注的目标不是 Operation（运维），而是 Engineering（工程），是一个“**通过软件工程的方式开发自动化系统来替代重复和手工操作**”的岗位。我们从 SRE 这本书的前面几个章节，可以看到谷歌不断强调 SRE 的工程能力。我简要摘取几段：

Common to all SREs is the belief in and aptitude for developing software systems to solve complex problems.

所有的 SRE 团队成员都必须非常愿意，也非常相信用软件工程方法可以解决复杂的运维问题。

By design, it is crucial that SRE teams are focused on engineering.

SRE 模型成功的关键在于对工程的关注。

SRE is what happens when you ask a software engineer to design an operations team.

SRE 就是让软件工程师来设计一个新型运维团队的结果。

与之相对应的，还有一个很有意思的地方，整本书中提到 Operation 的地方并不多，而且大多以这样的词汇出现：Operation load, Operation overload, Traditional/Manual/Toil/Repetitive Operation Works。你可以仔细体会一下，这些大多就是传统的纯人工操作模式下的一些典型词汇。

我们可以看到，从一开始，谷歌就没把 SRE 定义为纯操作类运维的岗位，也正是谷歌换了一个思路，从另外一个维度来解决运维问题，才把运维做到了另一个境界。

SRE 岗位的职责

书中对 SRE 的职责定义比较明确，**负责可用性、时延、性能、效率、变更管理、监控、应急响应和容量管理等相关的工作**。如果站在价值呈现的角度，我觉得可以用两个词来总结，就是“**效率**”和“**稳定**”。

接下来，详细说下我的理解和分析。

SRE 的能力模型，不仅仅是技术上的，还有产品设计、标准规范制定、事后复盘总结归纳这些技术运营能力，同时还需要良好的沟通协作能力，这个就属于**职场软技能**。

SRE，直译过来是网站稳定性工程师。表面看是做稳定的，但是我觉得更好的一种理解方式是，**以稳定性为目标，围绕着稳定这个核心，负责可用性、时延、性能、效率、变更管理、监控、应急响应和容量管理等相关的工作。**

分解一下，这里主要有“管理”和“技术”两方面的事情要做。

管理体系上，涉及服务质量指标（SLI、SLA、SLO）、发布规则、变更规则、应急响应机制、On-Call、事后复盘机制等一系列配套的管理规范和标准制定等。

技术体系上，以支持和实现上述标准和规范为目标，涉及自动化、发布、监控、问题定位、容量定位，最终以电子流程串联各个环节，做到事件的闭环。

可以看到技术上的平台和系统是用来支撑管理手段的。谷歌的运维其实并没有单独去提自动化、发布、监控等内容，而是通过稳定性这个核心目标，把这些事情全部串联在一起，同时又得到了效率上的提升。

我们来看几个主要的系统。

自动化。是为了减少人为的、频繁的、重复的线上操作，以大大减少因人为失误造成的故障，同时提升效率。比如谷歌内部大名鼎鼎的 Borg 系统，可以随时随地实现无感知的服务迁移。现在，它的开源版本，已然成为业界容器编排体系标准的 Kubernetes。

持续交付。谷歌非常重视持续交付。由于它的需求迭代速度非常快，再加上是全球最复杂的分布式系统，所以就更加需要完善的发布系统。

问题定位。这块跟监控相关但又有不同。我看到谷歌 SRE 这本书中并没有提到太多 Tracing 的内容，更多的是讲监控和问题管理层面的跟踪机制。其实，关于问题定位，谷歌的 Dapper 大名鼎鼎，功能很强大，国内外很多跟踪系统和思路都参考了 Dapper 的理论。这块也是为了能够快速定位问题，保障稳定而产生的，国内分享的大多关于全链路跟踪和分析、限流降级、开关和预案系统、强弱依赖等都属于这个范畴，我认为这块应该更准确地定义为分布式服务治理相关的内容。

各类分布式系统。如分布式锁、分布式文件、分布式数据库，我们熟知的谷歌三大分布式论文，就是这些分布式系统的优秀代表，也正是这三大论文，开启了业界分布式架构理念的落地。

这些系统大都是以稳定性为导向，同时带动了日常运维效率的大幅度提升，有了监控和全链路这样的问题发现和定位手段，也大大提升了我们对故障处理和问题定位的效率。容量管理，不仅仅可以保障容量充足，还能最大程度地保障资源分配的合理性，尽可能减少浪费，对于成本管控也大有好处。所以，围绕着稳定性这个核心目标，不仅达到了稳定的目的，还获得了高效的运维效率。

所以，SRE 的理念通过稳定性这个核心点，将整个运维体系要做的事情非常系统紧密地整合起来，而不是一个个孤立的运维系统。所以，**SRE 是一个岗位，但更是一种运维理念和方法论。**

如何借鉴和落地

在国外，SRE 岗位的薪资，和 SWE 软件开发工程师相比，要平均高出 25%。从实际的岗位要求上看，SRE 的技能要求也要比 SWE 更高、更全面，所以这样的人才比较紧缺。即使在国外，除了谷歌、Facebook 以及 Netflix 这样超一流的科技公司能够招聘到，或者内部培养出较多这样的人才，其它公司的 SRE、PE 或者应用运维也无法完全达到上面的要求。

在国内，就更难一些，那怎么做呢？一个思路就是我们之前讲组织协作模式转型时提到的，就是要依靠团队的力量、发挥团队的力量，如果是单个团队不能完成的事情，就跨团队协调完成。所以，SRE 岗位的要求很高，但是我们可以靠团队中具备不同能力的人协作，共同达成 SRE 的职责和目标。

最后，还是我反复强调的观点，**要想做好运维，就得跳出运维的局限，要站在全局的角度，站在价值呈现的角度，站在如何能够发挥出整体技术架构运维能力的角度，来重新理解和定义运维才可以。**

通过今天的内容，你对于 SRE 有什么新的理解或者疑问？结合前面的内容，你能够挖掘出哪些共通点呢？欢迎你留言与我讨论。

如果今天的内容对你有帮助，也欢迎你分享给身边的朋友，我们下期见！


赵成的运维体系管理课

带你直击运维的本质

赵成

美丽联合集团技术
服务经理



新版升级：点击「 请朋友读」，10位好友免费读，邀请订阅更有**现金**奖励。

© 版权归极客邦科技所有，未经许可不得传播售卖。页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 10 | 如何打造运维组织架构？

下一篇 12 | 从谷歌CRE谈起，运维如何培养服务意识？

精选留言 (7)

写留言



luoweiro

2018-01-12

9

赞博主精辟剖析！谈谈个人思考：

巨大流量对稳定性指标的冲击其实是量变引发了质变。不管是软件工程还是线上运维其实自然是一体，谷歌SRE其实更是无边界支撑线上业务的模式的实践。...

展开

作者回复: 





宵伯特

2018-01-14

👍 2

国内的情况大概是因为大多数的公司组织对于运维的职能概念停留在软件后期维护的层面，对于敏捷开发和devops也是一知半解，认知层次和视角都需要进一步的提升。好在如今的技术交流和组织沟通都是自由的，即便是随着技术的发展和产品规模的扩大，也会有越来越多的组织团队意识到该职位的重要性。



朱雯

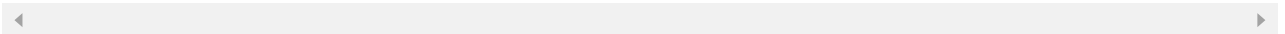
2018-01-14

👍 2

在一家创业公司做运维，现在做的事情是维护公司网络和报警。在您的专栏中提到的这些东西感觉和我差距太远了，是否因为道行太浅的缘故？

作者回复: 后面我有一篇文章介绍个人能力建设方面的，到时可以看一下。

就留言中你提到的工作而言，如果非常例行化，个人感觉没有什么提升的话，建议还是尝试一些更有挑战性的工作。



Penn

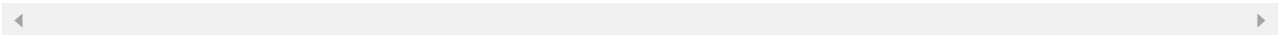
2018-01-12

👍 2

理念是道，实践是术，能落地的方案最有力，期待实践部分的讲解。

展开 ▾

作者回复: 后面会有，可以继续关注，也欢迎多留言讨论。



三宝

2018-02-07

👍 1

Site Reliability Engineering

展开 ▾



kevinsu

2019-05-16

👍

运维需要不断挑战自己，不给自己设限制

展开 ∨



孙志宇

2019-03-31



收获好大，重新定义了我对运维的理解，谢谢老师！

展开 ∨