

为何你需要一个开源软件战略

作者：Pranay Ahtawat; Johannes Boyne; Dominik Herz; Florian Schmieg; Michael Stephan

当一个运动成为主流时，管理它就需要战略。1991 年 8 月，当时 21 岁的赫尔辛基大学计算机科学学生 Linus Torvalds 通过一篇 Usenet 帖子随意宣布：“我正在做一个免费的操作系统，只是个业余爱好，不会像 GNU 那样庞大和专业……”结果，他的业余爱好催生了世界上第一个免费操作系统——Linux 内核，并启动了开源软件运动。整整三十年后的今天，“开源”已成为软件开发的主要方式之一。公司广泛使用开源软件，它正日益塑造企业软件架构。开发和部署开源软件不再只是一个新奇的想法。在快速变化的数字世界中，它是一种战略必需。

开源软件无处不在，可以定义为开发人员可以检查、复制、修改和重新分发的软件。专有软件供应商仍然主导市场，但开源软件发挥着同等重要的作用。例如，2020 年，开源 Linux 支撑了公共云 75% 的工作负载，预计到 2024 年其份额将升至 85%。一些最受欢迎的软件开发技术栈——如 LAMP（Linux, Apache, MySQL, PHP）和 MEAN（MongoDB, Express.js, AngularJS, Node.js）栈——都是开源软件。去年，全球约 85% 的智能手机运行在基于开源 Linux 内核构建的开源操作系统 Android 上（见图 1）。毫不奇怪，使用开源软件的能力正迅速成为对所有软件的要求。

开源软件在商业中的应用正在飙升。高达 80% 的 IT 部门计划在未来 12 个月内增加开源软件的使用，95% 的 IT 专家认为开源已具有战略重要性（见侧边栏 1）。数字化转型的推动力——软件开发人员和数据分析师——严重依赖开源社区。他们通常更喜欢使用开源软件，尤其是作为基础，因为软件选择和筛选过程简单，且不太可能需要进行冗长的谈判。因此，这种选择使得能够快速推出和采用创新应用。除了提高上市速度外，使用开源软件还能防止供应商锁定，并显然降低成本。

第一章：未来的软件技术栈将日益开放

开源软件在未来将继续流行，创新周期只会巩固其地位。随着开源核心公司和社区推动创新，开源软件很可能在企业软件技术栈的多个层面发挥基础作用，从操作系统和编程语言到中间件和开发工具。

如今，大多数超级计算机都使用基于 Linux 的操作系统。它稳定、廉价，并且重要的是——可修改，这与 Windows 等专有系统不同。因此，很容易创建一个仅包含最必要代码的内核，去掉所有对提高性能非关键的部分。

由于同样的原因，Linux 被用作各种硬件（如服务器、台式机、智能手机）的操作系统，其份额正在增长。例如，Linux 在服务器安装中的份额已从 2017 年的 68% 上升到 2020 年的 75%。数据管理也走向开源，HashiCorp、GitLab、Datadog、Elastic、Confluent 和 Databricks 等公司已进入市场。

开源软件在容器和容器编排平台等新技术中获得了发展势头。与其他技术如 Mesos 和 Docker Swarm 一起，Kubernetes 已成为容器管理的标准之一。它协调着全球超过 50% 的容器，预计到 2024 年其份额将上升到 85%。

聪明的公司已开始识别企业技术栈每一层最佳的新兴开源技术——例如用于分析的 Spark，用于消息传递和流处理的 Kafka，用于人工智能的 PyTorch 和 TensorFlow 等。因此，闭源代码技术栈面临激烈竞争，它们必须在每一层都做到同类最佳，才有可能被选中。

DevSecOps 等软件工具也呈现类似趋势，像 Git 这样的源代码仓库，像 React 这样的库框架，以及像 Spring 这样的后端应用框架，因其在开发者中受欢迎而逐渐取代专有软件。开源软件的成功将引发一个良性循环，更多的项目涌现，吸引更多开发者，开源软件的使用将变得更加普遍。

此外，公司正在学习利用开源社区获取人才并提升自身技能。一些公司已从勉强接受使用开源软件，转变为鼓励员工参与开源项目。首席信息官和首席技术官们开始意识到，他们必须重新思考其方法，优先发展开源软件，以超越竞争对手。他们越来越多地思考：我们是否有面向 2020 年代的开源软件战略？

第二章：开源软件的崛起...与持续崛起...

开源软件在几个方面与专有软件不同。与需要付费的专有软件不同，开源软件不属于任何人。它可以免费使用，但除非公司从商业供应商处获得使用许可（如下所述），否则没有社区支持。此外，与供应商保密的专有软件源代码不同，开源软件是公开开发的，因此可以轻松测试、修改和自由分发。

由于多种原因，制定和实施开源软件战略已变得势在必行。开源软件开发人员，无论是个人还是集体，都在寻找技术问题的最佳解决方案，这使得他们创建的软件可靠、安全——并且免费。由于他们不懈的努力，软件会随着时间的推移变得更好。Linux 基金会（支持跨多个技术领域的开源）、Apache 软件基金会和 Eclipse 基金会等几个基金会促进了这一过程。他们与 AWS、Facebook、Google、IBM、Microsoft、Netflix、SAP 等数字巨头，以及 Cisco、Intel、Tesla 等硬件制造商合作，共同制定标准。他们也为项目创建与供应商无关的“家园”，为基础设施提供资金支持，协助营销，并任命委员会来做出项目的关键决策。

由开源催生的第一批初创公司出现在三十年前，提供支持服务，为第二代公司铺平了道路，这些公司在内部开发软件，但发布源代码以便社区测试和完善。采用这种“开源核心”模式（核心是开源代码，周围是专有代码）的公司，提供一个功能有限的免费产品，以及一个功能丰富的专有版本，用户必须支付订阅或许可费——本质上是一种免费增值商业模式。产品的开放程度范围很广，从拥有小封闭外壳的大开源核心（可称为“薄壳”产品），到另一个极端——小的开源核心和大的封闭外壳（也称为“厚壳”产品）。

基于开源核心模式，现在几家品牌软件供应商将开源软件和商业软件混合，并以许可的基于云的服务形式提供，通过他们提供的支持和服务实现盈利。他们提供软件的付费版本和一套捆绑服务以收取订阅费。客户获得附加组件，如仪表盘和分析工具，确保安全和性能的更新，安全认证，以及针对受监管行业的其他许可。供应商保证，即使软件过时，他们也会支持关键应用。他们还提供软件维护，协调和安装更新，甚至提供实时软件支持。他们经常作为软件选择的顾问，并培训客户的员工。通过使用这种软件即服务许可模式，大大小小的开源商业供应商都成功地扩大了市场阵地（见侧边栏 2）。

第三章：开源软件市场正在增长

从云端到边缘，开源软件如今已成为许多架构的组成部分，其使用在未来可能会增加。（见图 3。）

在云端， Linux 已成为首选云操作系统，主要是因为与专有软件相比，它具有更大的灵活性和更低的成本。它支持容器的使用，这是微服务等云原生开发构件的必备条件。由于工作负载不断上升，Linux 操作系统市场预计将以每年 7% 的速度增长，到 2024 年达到 97 亿美元。虽然 2024 年 70% 的工作负载将在云端运行（2020 年为 50%），但预计 2024 年 82% 的付费 Linux 用户将在多云环境中运行工作负载——高于 2020 年的 72%。混合云架构（同时使用内部部署和公共云基础设施）的采用率也在上升，去年超过 95% 的企业部署了混合云。

根据我们的市场研究，付费 Linux 提供商在 2020 年创造了约 33 亿美元的收入。然而，总目标市场——包括 Windows、Unix 等其他操作系统的份额以及免费 Linux 部署——超过这个数字的两倍，约为 75 亿美元。由于大多数新的应用程序和工作负载将在 Linux 上运行，现有的 Windows 和 Unix 应用程序可能会迁移到 Linux，企业 Linux 提供商将获得市场份额。在未来三年半内，他们的年增长率可能约为 12%，因此到 2024 年，他们的总收入将超过 52 亿美元。

在企业内部， 开源容器管理软件和软件定义存储最受欢迎。约 90% 使用容器的公司部署了容器管理平台。在 Kubernetes（作为容器管理行业标准崛起的开源软件）之上，红帽、VMware、SUSE Rancher 和云提供商等公司提供打包解决方案、附加组件、用户友好界面以及更新和维护等服务。他们瞄准的是一个约 22 亿美元的可寻址市场，该市场将以每年约 35% 的速度增长，到 2024 年达到 75 亿美元。而可服务市场（不包括超大规模云提供商内部基础设施中部署的容器）在 2020 年约为 7 亿美元，将以每年 53% 的速度增长，快于整体市场，到 2024 年达到 40 亿美元。

同样，软件定义存储的使用最近变得越来越重要，它允许公司通过整合内部部署和云存储来动态控制数据存储位置。虽然专有软件系统通常管理存储，但 Ceph、Gluster 和 Longhorn 等开源解决方案正获得关注。云原生、容器驱动的软件定义存储仅占 40 亿美元总可寻址市场的一小部分，预计到 2024 年将达到 62 亿美元。然而，与容器化相关的细分市场将以更快的速度增长，每年 50%，从 2020 年的 8000 万美元增长到

2024 年的 4 亿美元。

在边缘，公司未来产生的数据将比现在更多。推动边缘计算的物联网连接数量将从 2019 年的 200 亿增长到 2025 年的 410 亿，翻一番。超低延迟、可靠连接、更高安全性和成本节约等几个因素将推动对边缘计算的需求。

随着边缘计算预计在 2020 年至 2024 年间以每年 24% 的速度增长，边缘硬件、软件和服务的计算支出预计将从 2020 年的 890 亿美元急剧上升到 2024 年的 2180 亿美元。

开源软件提供商将在平台软件和基础设施软件领域找到机会，这些领域预计将从 2020 年的 100 亿美元增长到 2024 年的 249 亿美元。

这些供应商在很大程度上依赖于开源社区。除了员工，自由职业者和业余爱好者程序员也参与项目，部分是为了证明其技术能力，主要是因为对软件开发充满热情。最近，一些公司发布了他们开发的功能和适配的代码，以便将其集成到更多软件中。2020 年，超过 5600 万开发人员在领先的开源协作平台 GitHub 上列出的 1.4 亿个项目（代码仓库）上工作，贡献超过 19 亿次。Amazon、Facebook、Google、IBM、Intel、SAP 和 Microsoft 都不是开源公司，却是 GitHub 上最大的贡献者之一。

没有哪家公司比 Microsoft 更能体现对开源软件态度的转变，该公司最初曾对其发起法律战。如今，这家数字巨头广泛使用开源软件。Microsoft Azure 大部分运行在 Linux 上，并且它创建了一个兼容层——适用于 Linux 的 Windows 子系统，以便在 Windows 上原生运行 Linux 二进制可执行文件。Microsoft 还开放了 .NET（用于 Windows、Linux 和 macOS 操作系统的软件框架）、TypeScript 编程语言以及其任务自动化和配置管理框架 PowerShell 的源代码。这家数字巨头已加入开源促进会，于 2018 年以 75 亿美元收购了 GitHub（当时最大的企业软件收购），其员工积极参与 GitHub，2020 年有超过 5000 名员工为开源项目做出贡献。

第四章：优势与风险

在制定战略之前，公司应对开源软件的优缺点有一个细致入微的理解。

优势。随着时间的推移，开源软件之所以如此流行，有以下六个原因。

第一，开源软件社区庞大，技术多样化，致力于用数字技术解决问题。其精湛技艺和活力提供了优势，社区确保应用程序快速开发。问题越大，吸引的开发人员就越多，就像磁铁一样被吸引来解决它。

第二，社区采用协作式的软件开发方法，有助于推动创新。人工智能和机器学习等最新技术运行在开源软件上并非偶然。除了基础设施，开源软件还推动了最新的技术飞跃，如用于自动驾驶汽车的边缘计算。下一代支持使用高度分布式系统的超复杂企业软件，很可能也是开源的。

第三，开源软件得到大量开发者的支持。例如，自 2005 年以来，来自约 1400 家公司的超过 15,500 名开发者为 Linux 内核做出了贡献，他们每天添加 1 万行代码，使其成

为世界上发展最快的项目。

第四，与闭源软件相比，开源代码可以完全访问和定制。它通常是模块化的，因此供应商可以调整部分代码或为其添加功能，为每个业务进行定制。这是为什么开源软件在企业技术栈的任何可用层面通常都能与专有软件相媲美的另一个原因。

第五，企业级开源软件过时的风险较低，因为有社区的参与。依赖专有软件的公司面临软件被停用或随着时间的推移不得不支付更多费用的风险，而数字技术快速变化的性质放大了这种风险。

第六，对大多数传统公司来说，找到执行数字化转型的人才是挑战，因此他们可以求助于开源社区。它是人才和工具的海洋，其深度可能只有在世界上最大的软件公司中才能找到。此外，鉴于开源软件的广泛适用性，找到熟悉开源软件的开发者比找到熟悉专有软件所需特定工具的人更容易。

风险。像生活中的一切一样，使用开源软件也存在一些风险。与提供清晰法律协议的商业软件相比，从商业开源供应商处获得许可有时可能具有模糊性。

一些协议，如流行的 MIT 和 Apache 许可证，只包含关于软件再分发的最基本要求。虽然 MIT 许可证的措辞相当简单，但 Apache 2.0 许可证的条款更为详细，因此后者在为企业级部署（如 Docker、Kubernetes、Swift 和 TensorFlow）设计的大型开源项目中更受欢迎。

其他许可证，如 GNU 通用公共许可证（GNU GPL），要求免费再分发修改版本的源代码。这意味着即使包含了开源代码的专有软件也必须公开源代码，这被称为“著佐权”或“传染性许可”，会让企业担忧。公司在使用开源软件时应牢记著佐权的连锁反应。

随着开源软件在商业中的使用增长，最大的风险是没有实体对不利后果承担责任。缺乏责任追究会导致法律上的复杂性，尤其是在公司将其用于开发关键任务应用时，例如控制汽车刹车系统。公司必须学会在收获此类软件的好处与了解自己将承担法律责任之间取得平衡。

开源软件通常是安全的。开源代码允许许多双眼睛审查并确保其安全性。然而，资金不足的项目有时可能存在影响深远的安全问题。例如，OpenSSL 是一个加密软件库，被网络服务器、网站和操作系统用来安全处理密码和信用卡详细信息等敏感数据。

2014 年，OpenSSL 中发现了一个名为 Heartbleed 的漏洞，导致了安全相关的紧急事件。在 OpenSSL 社区修补 Heartbleed 之前，由于这个漏洞，互联网上五分之一的网络服务器容易受到黑客攻击。当时，只有一名全职人员从事 OpenSSL 工作。

第五章：迈向开源软件战略

软件开发人员在面对技术挑战时几乎本能地转向开源软件，因此制定一个战略来管理其在组织中的使用至关重要。

第一步是明确阐述员工可以以及不可以使用开源软件的目的。这样做将帮助员工了解在哪些领域允许他们利用此类软件以及如何选择工具，从而使得使用开源软件的组织风险处于可承受范围内。影响这些决策的关键因素是软件的流行度、维护成本及其安全程度。

每家公司都需要为使用开源软件建立治理、法律和风险结构。它必须规定是倾向于标准许可证，还是希望起草自己的许可证，以及对著佐权条款的接受程度。尽管后者可能是最公平的，但大多数公司避免使用包含著佐权要求的许可证。

根据其接纳程度，企业必须发展管理开源软件使用的能力。大多数公司会设立项目办公室，作为开源相关活动的一站式服务中心。它们协调内部围绕法律、技术和安全问题的活动，以及面向外部的活动，如营销和沟通。

其他公司则建立了卓越开源社区。他们识别组织内每个部门使用的开源软件，并促进协作以及最佳实践分享。通过促进爱好者之间的交流，并让各个职能部门分享成功故事，有助于公司充分发挥开源软件的潜力。

最后，企业应决定员工是否可以参与开源计划，无论是作为工作的一部分还是在业余时间。这在很大程度上取决于公司的雄心，但这并不像看起来那么牵强。最近，沃尔玛发布了一个开源云管理系统，埃克森美孚发布了一个帮助能源公司采用标准数据格式的开发工具包，摩根大通和富国银行投资了 Hyperledger，这是一个用于企业级区块链部署的开源软件套件。

聪明的公司将效仿他们的做法，确定他们希望从开源中获得的一系列好处，从吸引人才到增加收入。然后，他们将决定如何通过使用软件、为项目做贡献或参与基金会活动等方式来实现其目标。他们可以从参与小项目开始，例如贡献修复程序，并随着时间的推移扩大参与度。重要的是，公司可以通过构建开源生态系统来影响新兴技术标准的发展。例如，2014年，谷歌推出了一个开源容器编排系统 Kubernetes，该系统正成为容器管理的事实标准。

结语

随着“开放”理念成为21世纪社会的标志之一，并延伸到创新、研究和标准制定领域，开源软件已成为商业不可或缺的一部分。由于其发生的速度和规模，开源已成为当今战略的核心。它通过为企业提供强大的能力而无需其进行前期投资，正在改变创新的经济模式。除了为初创企业创造公平的竞争环境外，开放式创新也迫使数字巨头转向并拥抱开源软件。那些找到方法利用开源社区广度和深度的公司将节省时间和金钱，并提升其创新能力。这听起来可能有些矛盾，但现在是首席执行官们将开源软件作为为其公司获得专有优势的途径来关注的时候了。