

44 | 互联网架构模板：“平台”技术

2018-08-07 李运华 来自北京

《从0开始学架构》



当业务规模比较小、系统复杂度不高时，运维、测试、数据分析、管理等支撑功能主要由各系统或者团队独立完成。随着业务规模越来越大，系统复杂度越来越高，子系统数量越来越多，如果继续采取各自为政的方式来实现这些支撑功能，会发现重复工作非常多。因此我们自然而然就会想到将这些支撑功能做成平台，避免重复造轮子，减少不规范带来的沟通和协作成本。

今天，我就来聊聊互联网架构模板的“平台”技术。由于每个平台本身都是一个庞大的体系，专栏只是介绍一下平台的核心职责和关键设计点，具体细节就不详细展开了。

运维平台

运维平台核心的职责分为四大块：配置、部署、监控、应急，每个职责对应系统生命周期的一个阶段，如下图所示：



配置：主要负责资源的管理。例如，机器管理、IP 地址管理、虚拟机管理等。

部署：主要负责将系统发布到线上。例如，包管理、灰度发布管理、回滚等。

监控：主要负责收集系统上线运行后的相关数据并进行监控，以便及时发现问题。

应急：主要负责系统出故障后的处理。例如，停止程序、下线故障机器、切换 IP 等。

运维平台的核心设计要素是“四化”：标准化、平台化、自动化、可视化。

1. 标准化

需要制定运维标准，规范配置管理、部署流程、监控指标、应急能力等，各系统按照运维标准来实现，避免不同的系统不同的处理方式。标准化是运维平台的基础，**没有标准化就没有运维平台。**

如果某个系统就是无法改造自己来满足运维标准，那该怎么办呢？常见的做法是不改造系统，由中间方来完成规范适配。例如，某个系统对外提供了 RESTful 接口的方式来查询当前的性能指标，而运维标准是性能数据通过日志定时上报，那么就可以写一个定时程序访问 RESTful 接口获取性能数据，然后转换为日志上报到运维平台。

2. 平台化

传统的手工运维方式需要投入大量人力，效率低，容易出错，因此需要在运维标准化的基础上，将运维的相关操作都集成到运维平台中，通过运维平台来完成运维工作。

运维平台的好处有：

可以将运维标准固化到平台中，无须运维人员死记硬背运维标准。

运维平台提供简单方便的操作，相比之下人工操作低效且容易出错。

运维平台是可复用的，一套运维平台可以支撑几百上千个业务系统。

3. 自动化

传统手工运维方式效率低下的一个主要原因就是要执行大量重复的操作，运维平台可以将这些重复操作固化下来，由系统自动完成。

例如，一次手工部署需要登录机器、上传包、解压包、备份旧系统、覆盖旧系统、启动新系统，这个过程中需要执行大量的重复或者类似的操作。有了运维平台后，平台需要提供自动化的能力，完成上述操作，部署人员只需要在最开始单击“开始部署”按钮，系统部署完成后通知部署人员即可。

类似的还有监控，有了运维平台后，运维平台可以实时收集数据并进行初步分析，当发现数据异常时自动发出告警，无须运维人员盯着数据看，或者写一大堆“grep + awk + sed”来分析日志才能发现问题。

4. 可视化

运维平台有非常多的数据，如果全部通过人工去查询数据再来判断，则效率很低。尤其是在故障应急时，时间就是生命，处理问题都是争分夺秒，能减少 1 分钟的时间就可能挽回几十万元的损失，可视化的主要目的就是为了提升数据查看效率。

可视化的原理和汽车仪表盘类似，如果只是一连串的数字显示在屏幕上，相信大部分人一看到一连串的数字，第一感觉是眼花，而且也很难将数据与具体的情况联系起来。而有了仪表盘后，通过仪表盘的指针偏离幅度及指针指向的区域颜色，能够一目了然地看出当前的状态是低速、中速还是高速。

可视化相比简单的数据罗列，具备下面这些优点：

能够直观地看到数据的相关属性，例如，汽车仪表盘中的数据最小值是 0，最大是 100，单位是 MPH。

能够将数据的含义展示出来，例如汽车仪表盘中不同速度的颜色指示。

能够将关联数据整合一起展示，例如汽车仪表盘的速度和里程。

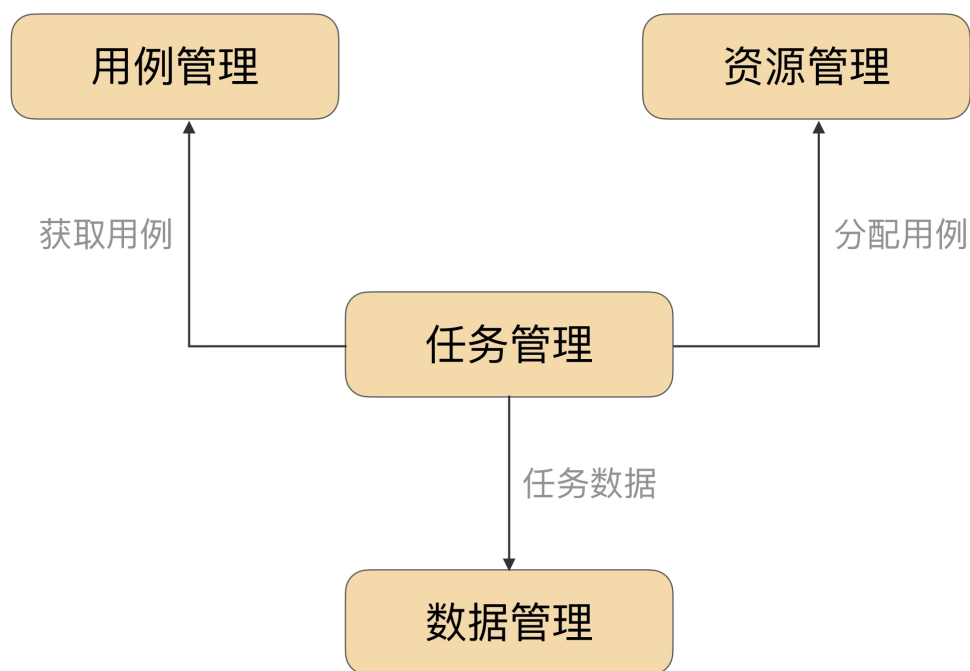


测试平台

测试平台核心的职责当然就是测试了，包括单元测试、集成测试、接口测试、性能测试等，都可以在测试平台来完成。

测试平台的核心目的是提升测试效率，从而提升产品质量，其设计关键就是自动化。传统的测试方式是测试人员手工执行测试用例，测试效率低，重复的工作多。通过测试平台提供的自动化能力，测试用例能够重复执行，无须人工参与，大大提升了测试效率。

为了达到“自动化”的目标，测试平台的基本架构如下图所示：



1. 用例管理

测试自动化的主要手段就是通过脚本或者代码来进行测试，例如单元测试用例是代码、接口测试用例可以用 Python 来写、可靠性测试用例可以用 Shell 来写。为了能够重复执行这些测试用例，测试平台需要将用例管理起来，管理的维度包括业务、系统、测试类型、用例代码。例如，网购业务的订单系统的接口测试用例。

2. 资源管理

测试用例要放到具体的运行环境中才能真正执行，运行环境包括硬件（服务器、手机、平板电脑等）、软件（操作系统、数据库、Java 虚拟机等）、业务系统（被测试的系统）。

除了性能测试，一般的自动化测试对性能要求不高，所以为了提升资源利用率，大部分的测试平台都会使用虚拟技术来充分利用硬件资源，如虚拟机、Docker 等技术。

3. 任务管理

任务管理的主要职责是将测试用例分配到具体的资源上执行，跟踪任务的执行情况。任务管理是测试平台设计的核心，它将测试平台的各个部分串联起来从而完成自动化测试。

4. 数据管理

测试任务执行完成后，需要记录各种相关的数据（例如，执行时间、执行结果、用例执行期间的 CPU、内存占用情况等），这些数据具备下面这些作用：

展现当前用例的执行情况。

作为历史数据，方便后续的测试与历史数据进行对比，从而发现明显的变化趋势。例如，某个版本后单元测试覆盖率从 90% 下降到 70%。

作为大数据的一部分，可以基于测试的任务数据进行一些数据挖掘。例如，某个业务一年执行了 10000 个用例测试，另外一个业务只执行了 1000 个用例测试，两个业务规模和复杂度差不多，为何差异这么大？

数据平台

数据平台的核心职责主要包括三部分：数据管理、数据分析和数据应用。每一部分又包含更多的细分领域，详细的数据平台架构如下图所示：



1. 数据管理

数据管理包含数据采集、数据存储、数据访问和数据安全四个核心职责，是数据平台的基础功能。

数据采集：从业务系统搜集各类数据。例如，日志、用户行为、业务数据等，将这些数据传送到数据平台。

数据存储：将从业务系统采集的数据存储到数据平台，用于后续数据分析。

数据访问：负责对外提供各种协议用于读写数据。例如，SQL、Hive、Key-Value 等读写协议。

数据安全：通常情况下数据平台都是多个业务共享的，部分业务敏感数据需要加以保护，防止被其他业务读取甚至修改，因此需要设计数据安全策略来保护数据。

2. 数据分析

数据分析包括数据统计、数据挖掘、机器学习、深度学习等几个细分领域。

数据统计：根据原始数据统计出相关的总览数据。例如，PV、UV、交易额等。

数据挖掘：数据挖掘这个概念本身含义可以很广，为了与机器学习和深度学习区分开，这里的数据挖掘主要是指传统的数据挖掘方式。例如，有经验的数据分析人员基于数据仓库构建一系列规则来对数据进行分析从而发现一些隐含的规律、现象、问题等，经典的数据挖掘案例就是沃尔玛的啤酒与尿布的关联关系的发现。

机器学习、深度学习：机器学习和深度学习属于数据挖掘的一种具体实现方式，由于其实现方式与传统的数据挖掘方式差异较大，因此数据平台在实现机器学习和深度学习时，需要针对机器学习和深度学习独立进行设计。

3. 数据应用

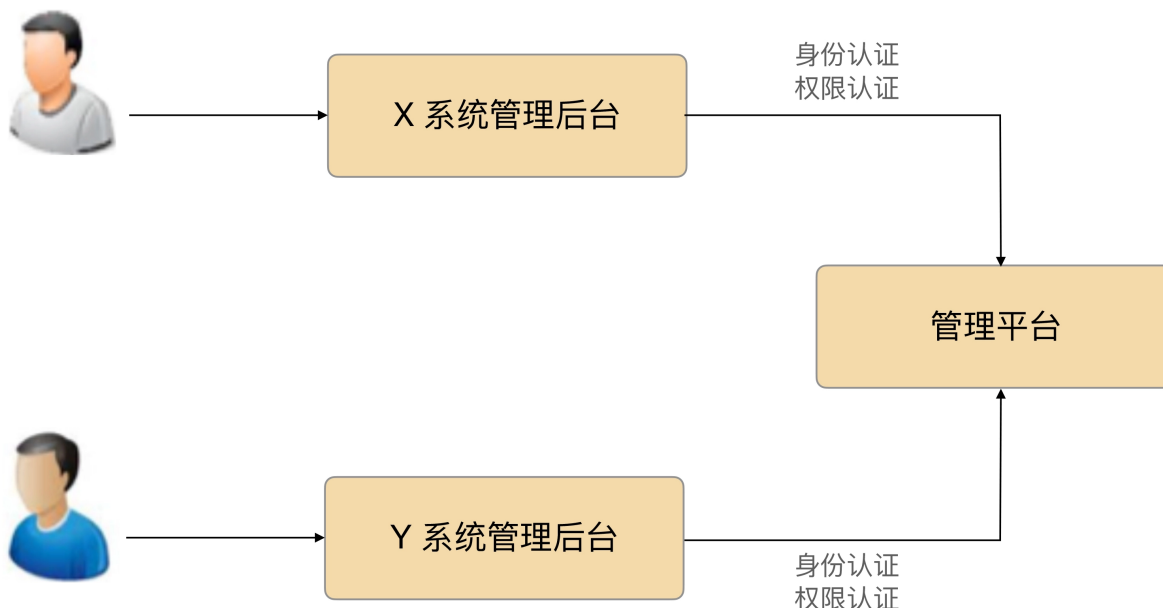
数据应用很广泛，既包括在线业务，也包括离线业务。例如，推荐、广告等属于在线应用，报表、欺诈检测、异常检测等属于离线应用。

数据应用能够发挥价值的前提是需要有“大数据”，只有当数据的规模达到一定程度，基于数据的分析、挖掘才能发现有价值的规律、现象、问题等。如果数据没有达到一定规模，通常情况下做好数据统计就足够了，尤其是很多初创企业，无须一开始就参考 BAT 来构建自己的数据平台。

管理平台

管理平台的核心职责就是**权限管理**，无论是业务系统（例如，淘宝网）、中间件系统（例如，消息队列 Kafka），还是平台系统（例如，运维平台），都需要进行管理。如果每个系统都自己来实现权限管理，效率太低，重复工作很多，因此需要统一的管理平台来管理所有的系统的权限。

权限管理主要分为两部分：身份认证、权限控制，其基本架构如下图所示。



1. 身份认证

确定当前的操作人员身份，防止非法人员进入系统。例如，不允许匿名用户进入系统。为了避免每个系统都自己来管理用户，通常情况下都会使用企业账号来做统一认证和登录。

2. 权限控制

根据操作人员的身份确定操作权限，防止未经授权的人员进行操作。例如，不允许研发人员进入财务系统查看别人的工资。

小结

今天我为你讲了互联网企业常见的平台以及基本的平台功能，希望对你有所帮助。

这就是今天的全部内容，留一道思考题给你吧，运维平台或者测试平台，有的公司是由中间件团队负责开发，有的是运维和测试团队自己开发，你觉得两种方式各有什么优缺点，分别适用什么场景呢？

欢迎你把答案写到留言区，和我一起讨论。相信经过深度思考的回答，也会让你对知识的理解更加深刻。（编辑乱入：精彩的留言有机会获得丰厚福利哦！）

© 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。

精选留言 (28)



天使

2018-08-09

jira+gitlab+jenkins+nexus+bearychat 最简单的DevOps 平台。如果将生产环境完全交给运维团队的话，个人觉得这个应该可以称为开发平台。输入的是需求，输出的是各种工件。

作者回复：这个可以算开发平台👍👍👍



👍 38



文竹

2018-08-26

运维和测试平台由中间件团队开发：

优点：平台架构有保障，代码质量高，开发效率高

缺点：前期业务沟通成本大

适用场景：运维和测试开发能力弱。

运维和测试平台的由运维和测试人员开发：

优点：前期沟通成本低

缺点：技术能力弱，开发效率低

场景：运维和测试开发能力强

作者回复：逻辑很清晰👍

共 4 条评论 >

👍 27



Freedom

2018-08-07

为啥没有产品设计平台，开发平台

作者回复: 这个问题非常有意思, 先说产品平台, 实际上是有的, 其他例如商务, HR也有平台, 但因为专栏聚焦技术, 且我对这些平台没有太多理解, 所以没有讲。

再说开发平台, 为何我们说数据平台, 运维平台等, 但不说开发平台呢? 我理解是运维平台, 数据平台, 测试平台, 管理平台, 产品平台, 商务平台等, 这些平台都是“管理”平台, 运维平台是管理机器和系统, 数据平台是管理数据……以此类推, 但开发平台如果说管理的话就是代码, 但这明显跟我们讲的分层技术栈不是一个概念, 所以从整个一个公司的技术架构来讲, 一般不会说“开发平台”, 但其实狭义的开发平台是存在的, 例如maven+git就可以算开发平台, 完成代码和包管理

共 2 条评论 >

👍 17



小胖狗

2018-08-07

如果运维系统让中间件团队开发

- 1.中间件团队需要去理解运维方的需求, 他们本身可能并不熟悉运维。
- 2.像阿里的中间件团队, 看他们的技术博客得知, 貌似只专注中间件。

让运维开发:

- 1.运维人员只需将其日常操作平台化即可, 能更好的解决运维人员的系统。
- 2.当然, 这种情况下, 运维团队需要形成一定的规模和能力。

作者回复: 写不了代码的运维不是好的开发 😊😊👍👍



👍 15



旭东(Frank)

2018-08-09

平台这种需要领导层认可和推动, 否则只能在作坊的沼泽里苦苦挣扎。靠开发工程师个人和运维工程师来推动, 很是痛苦。

作者回复: 非常正确, 需要一个好的CTO, 不然很难落地



👍 10



钱

2019-09-04

课后思考及问题

运维平台或者测试平台, 有的公司是由中间件团队负责开发, 有的是运维和测试团队自己开

发，你觉得两种方式各有什么优缺点，分别适用什么场景呢？

我们公司的运维平台是中间件团队开发的，性能测试平台是测试团队自己开发的。

中间件团队的开发

优点：问题少，规范，统一

缺点：体验稍差，问题修复慢一些

适用场景：大厂有中间件团队，需求多，测试或运维研发有困难

自己的团队开发自己使用的平台

优点：体验更好，问题修复响应更快

缺点：代码bug多一些，

适用场景：运维测试研发能力强，有时间及精力做和维护

感觉架构实践环节讲的内容大而广，比较靠上层，增长见闻，辅助写PPT可以😊，具体到要做一个东西找最佳实践是找不到的！，老师为啥这么安排？

作者回复：给架构师参考用的，尤其是当你需要规划整个公司的技术架构的时候，知道一个全貌和基本的范围更重要

共 2 条评论 >

👍 5



孙振超

2018-10-03

最近几期的内容，每一个小主题都可以独立成一个专栏来讲了，在这里只能简要做个介绍。

对于课后作业，中间件团队来做的优点：平台的性能、可用性、扩展性、伸缩性等非功能性属性会好不少；缺点是在功能性需求上，易用性和需求的响应速度会差些。

运维或者测试团队自己开发的话优点是：功能完善性好，交互界面符合一线同学的使用习惯。实际上，虽然有些公司也有测试开发工程师和运维开发工程师，但真正的开发水平和开发工程师还是有一些差距，因而缺点可能是开发效率差些，使用的技术也会老些，系统的性能和稳定性差。

作者回复：篇幅只能告诉大家一个公司的总体技术架构，知道总体技术架构后，你再按照架构设计的方法论来实现各个系统就可以了😊

共 2 条评论 >

👍 3



2018-08-30

数据平台，初创公司，针对业务数据的一些报表统计，是动态查询好，还是抓取业务数据统一存储！例如：数据库是mysql 业务表有多个，要实现报表统计，需要关联多张表，这时候会存在性能问题，如果是独立报表统计的表，然后抓取数据存储统计，这时候就会发生数据一致性问题。老师有什么好的建议吗？

作者回复: 抓取业务数据统一存储好一些，因为数据一致性不影响报表整体准确性，几条或者几十条数据不一致没关系，如果大规模不一致那就要重跑报表



👍 3



那迦树

2020-03-26

个人觉得平台或者中台，在大公司才能发展起来，小公司很难开展，毕竟业务受限

作者回复: 是的，小公司就开源全家桶组合就差不多了，发展业务是最优先的



👍 2



feifei

2018-08-13

中间件团队开发运维或测试平台，这个优势是平台具有很强的通用性，在性能和可靠性上较好，单针对单系统来说，缺少很多针对性的功能，功能上来说就是满足80%,系统的可复用性好！这一般适用于公司开发公司统一的平台

测试自己开发的平台正好相反，功能100%，但平台就真对单系统，不能复用或复用很小，而且系统的性能和可靠性一般，适用于小业务系统



👍 2



开胃

2018-08-12

中间件团队开发出来的平台一般是通用型的，性能高且易扩展的基础平台，而运维团队和测试团队更加偏于自身的痛点去设计开发



👍 2



hello

2018-08-07

运维和测试的技术能力没有中间件强，开发效率低。但是中间件团队对运维和测试的痛点需要

沟通交流才能理解。如果运维和测试技术OK或者中间件团队对痛点理解OK谁做都一样，就看谁时间多。

作者回复: 很难做到痛点理解一样，我们有运维开发的运维平台，也有研发开发的运维平台，前者架构设计不行或者没有架构设计，很难扩展，系统不稳定；后者很操作难用😞



👍 2



IT生涯路漫漫

2020-10-02

觉得还少了一个风控平台

作者回复: 一般涉及金融和交易的业务有风控平台



👍 1



慕士塔格

2020-05-05

数据平台的设计可以更详细讲讲吗？或者推荐些资料。其中数据挖掘和机器学习在架构上的区别，或者选一两个实际例子讲讲

作者回复: 最近有本书《数据中台》可以看看



👍 1



brant

2020-04-08

老师请教一个问题，你是怎么定义什么是平台的。然后你觉得应该什么时候开始建设平台的

作者回复: 平台就是将不同业务都会用到的功能提取出来



👍 1



蛤蟆不好找

2018-08-13

关注点不同，所能设计的产品也会有重点跟非重点的区别，中间件可能更关注的是功能的实现，重点在于技术，
运维团队可能关注的是平台的运行稳定以及硬件方面的性能

测试团队在于平台本身功能点的覆盖情况，
所以由专门的团队来处理，让后其他的队伍提需求

作者回复: 优势互补，但只有大公司才有能力成立独立的团队来负责各种管理平台的开发



空档滑行

2018-08-07

1.开发人员关注的技术点不一样，中间件开发人员关心更多的是性能并发这些，对运维整体业务可能了解欠缺一些。运维和测试更偏向业务一点，对中间件关键的功能点可能不会理解很深
2.运维的kpi是降低成本，提高效率，是能够从数字体现出来的。合理的利用平台和中间件其实是很好的降成本的方式，比如消息队列消息包多大最合适，topic应该怎么划分最合理。硬币总有正反面，两个结合起来可能是最能发挥价值的

作者回复: 小公司结合可以优势互补，大公司一般有专门的运维开发，测试开发



Kevin

2022-12-06 来自广东

个人认为，除了打造这些平台，怎样让这些平台能关联互相支持的运行起来更重要。尤其是公司规模越来越大，it团队人员越来越多，系统越来越复杂，需要打通开发到-测试到运维一体化。数据做好支撑服务。开发根据数据标准产生数据，数据指标同时服务于开发。

作者回复: DevOps可以用来串起这些平台技术



杜秀清

2022-09-05 来自日本

这章偏向研发管理了

作者回复: 只能点到为止





coconut

2022-06-19

感觉 中间件团队开发给运维用的平台时，运维人员应该充当产品经理的角色

作者回复: 是的，不然不知道运维具体在什么场景下需要什么样的运维系统

