

02 | 分布式系统的指标：啥是分布式的三国

2019-09-23 聂鹏程 来自北京

《分布式技术原理与算法解析》



你好，我是聂鹏程。

在上一篇文章中，通过对分布式发展历程的学习，我们对分布式技术有了一个整体印象。接下来，我们就再来看看可以用哪些指标去具体地衡量一个分布式系统。如果你已经对分布式系统的指标了解得很清楚了，可以直接跳过这篇文章，学习下一讲的内容。

分布式系统的指标

从分布式技术的起源可以看出，分布式系统的出现就是为了用廉价的、普通的机器解决单个计算机处理复杂、大规模数据和任务时存在的性能问题、资源瓶颈问题，以及可用性和可扩展性问题。换句话说，分布式的目的是**用更多的机器，处理更多的数据和更复杂的任务**。

由此可以看出，**性能、资源、可用性和可扩展性**是分布式系统的重要指标。没错，它们就是分布式系统的“三国”。接下来，我们一起来看看这几个指标吧。

性能 (Performance)

性能指标，主要用于衡量一个系统处理各种任务的能力。无论是分布式系统还是单机系统，都会对性能有所要求。

不同的系统、服务要达成的目的不同，关注的性能自然也不尽相同，甚至是相互矛盾。常见的性能指标，包括吞吐量 (Throughput)、响应时间 (Response Time) 和完成时间 (Turnaround Time)。

吞吐量指的是，系统在一定时间内可以处理的任务数。这个指标可以非常直接地体现一个系统的性能，就好比在客户非常多的情况下，要评判一个银行柜台职员의 办事效率，你可以统计一下他在 1 个小时内接待了多少客户。常见的吞吐量指标有 QPS (Queries Per Second)、TPS (Transactions Per Second) 和 BPS (Bits Per Second)。

QPS，即查询数每秒，用于衡量一个系统每秒处理的查询数。这个指标通常用于读操作，越高说明对读操作的支持越好。所以，我们在设计一个分布式系统的时候，如果应用主要是读操作，那么需要重点考虑如何提高 QPS，来支持高频的读操作。

TPS，即事务数每秒，用于衡量一个系统每秒处理的事务数。这个指标通常对应于写操作，越高说明对写操作的支持越好。我们在设计一个分布式系统的时候，如果应用主要是写操作，那么需要重点考虑如何提高 TPS，来支持高频写操作。

BPS，即比特数每秒，用于衡量一个系统每秒处理的数据量。对于一些网络系统、数据管理系统，我们不能简单地按照请求数或事务数来衡量其性能。因为请求与请求、事务与事务之间也存在着很大的差异，比方说，有的事务大需要写入更多的数据。那么在这种情况下，BPS 更能客观地反映系统的吞吐量。

响应时间指的是，系统响应一个请求或输入需要花费的时间。响应时间直接影响到用户体验，对于时延敏感的业务非常重要。比如用户搜索导航，特别是用户边开车边搜索的时候，如果响应时间很长，就会直接导致用户走错路。

完成时间指的是，系统真正完成一个请求或处理需要花费的时间。任务并行（也叫作任务分布式）模式出现的其中一个目的，就是缩短整个任务的完成时间。特别是需要计算海量数据或处理大规模任务时，用户对完成时间的感受非常明显。

资源占用 (Resource Usage)

资源占用指的是，一个系统提供正常能力需要占用的硬件资源，比如 CPU、内存、硬盘等。

一个系统在没有任何负载时的资源占用，叫做**空载资源占用**，体现了这个系统自身的资源占用情况。比如，你在手机上安装一个 App，安装的时候通常会提示你有多少 KB，这就是该 App 的空载硬盘资源占用。对于同样的功能，空载资源占用越少，说明系统设计越优秀，越容易被用户接受。

一个系统满额负载时的资源占用，叫做**满载资源占用**，体现了这个系统全力运行时占用资源的情况，也体现了系统的处理能力。同样的硬件配置上，运行的业务越多，资源占用越少，说明这个系统设计得越好。

可用性 (Availability)

可用性，通常指的是系统在面对各种异常时可以正确提供服务的能力。可用性是分布式系统的一项重要指标，衡量了系统的鲁棒性，是系统容错能力的体现。

系统的可用性可以用**系统停止服务的时间与总的时间之比衡量**。假设一个网站总的运行时间是 24 小时，在 24 小时内，如果网站故障导致不可用的时间是 4 个小时，那么系统的可用性就是 $4/24=0.167$ ，也就是 0.167 的比例不可用，或者说 0.833 的比例可用。

除此之外，系统的可用性还可以用**某功能的失败次数与总的请求次数之比来衡量**，比如对网站请求 1000 次，其中有 10 次请求失败，那么可用性就是 99%。

你可能经常在一个系统的宣传语中见到或听到 3 个 9（或 3N, 3 Nines）、5 个 9（或 9N, 9 Nines）。这些宣传语中所说的 3 个 9、5 个 9，实际上就是系统厂商对可用性的一种标榜，表明该系统可以在 99.9% 或 99.999% 的时间里能对外无故障地提供服务。

讲到了可用性，你可能还会想到一个非常近似的术语：可靠性 (Reliability)。那**可靠性和可用性有什么区别呢？**

可靠性通常用来表示一个系统完全不出故障的概率，更多地用在硬件领域。而**可用性**则更多的是指在允许部分组件失效的情况下，一个系统对外仍能正常提供服务的概率。

杰夫·迪恩 (Jeff Dean) 曾在 Google I/O 大会上透露：谷歌一个基于 1000 台通用计算机的集群，一年之内就有 1000+ 硬盘会出现故障。由于现在比较常见的分布式系统基本上都是基于通用计算机的，这就意味着在这些系统中无法实现真正的可靠，所以我们会在一些场合见到可靠性和可用性交换使用的情况。

可扩展性 (Scalability)

可扩展性，指的是分布式系统通过扩展集群机器规模提高系统性能 (吞吐量、响应时间、完成时间)、存储容量、计算能力的特性，是分布式系统的特有性质。

分布式系统的设计初衷，就是利用集群多机的能力处理单机无法解决的问题。然而，完成某一具体任务所需要的机器数目，即集群规模，取决于单个机器的性能和任务的要求。

当任务的需求随着具体业务不断提高时，除了升级系统的性能做垂直 / 纵向扩展外，另一个做法就是通过增加机器的方式去水平 / 横向扩展系统规模。

这里垂直 / 纵向扩展指的是，增加单机的硬件能力，比如 CPU 增强、内存增大等；水平 / 横向扩展指的就是，增加计算机数量。好的分布式系统总是在追求“线性扩展性”，也就是说系统的某一指标可以随着集群中的机器数量呈线性增长。

衡量系统可扩展性的常见指标是加速比 (Speedup)，也就是一个系统进行扩展后相对扩展前的性能提升。

如果你的扩展目标是为了提高系统吞吐量，则可以用扩展后和扩展前的系统吞吐量之比进行衡量。

如果你的目标是为了缩短完成时间，则可以用扩展前和扩展后的完成时间之比进行衡量。

不同场景下分布式系统的指标

我们都希望自己的分布式系统是高性能、高可用、高扩展和低资源占用的。但出于硬件成本、开发效率等因素的约束，我们无法在性能、可用性、可靠性和资源占用做到面面俱到。因此，在不同的业务场景中，设计者们需要有所取舍。

接下来，我带你一起看一下典型的电商、IoT、电信、HPC（高性能计算）、大数据、云计算、区块链等业务或系统对不同指标的诉求。

电商系统。对于一个电商系统而言，系统设计者最看重的是吞吐量，为了处理更多的用户访问或订单业务，甚至不惜牺牲一些硬件成本。

IoT。对于一个 IoT 系统而言，设计者最看重的是资源占用指标，因为在一些功能极简的 IoT 设备上 RAM、ROM 的可用资源通常都是 KB 级的。

电信业务。对于电信业务而言，最重要的无疑是响应时间、完成时间，以及可用性。因为，你在打电话时不希望你的声音半天才被对方听到，也不希望半天才听到对方的回应，更不希望你的电话无法拨出。

HPC。HPC 系统最显著的特点是任务执行时间极长，一个天体物理任务的分析和计算通常耗时数周甚至数月。因此，通过水平扩展来提高系统的加速比，是 HPC 系统设计者需要关注的。

大数据。大数据任务的处理时间可能相对 HPC 系统来讲比较短，但常见的完成时间也达到了小时级，所以扩展性也是大数据系统首先要考虑的。

云计算。对于一个云计算系统而言，常见任务是虚拟主机或容器的创建、资源调整、销毁等操作，如何减少这些操作的完成时间，从而提升用户体验是设计者们要重点关注的。另外，云计算系统本质上卖的是资源，那么降低系统本身的资源开销，也是系统设计的重中之重。

区块链。区块链的吞吐量比较低，比特币的 TPS 只有 7 次每秒，单平均一次交易的确认就需要 10 分钟左右，因此吞吐量和完成时间通常是区块链系统设计者的首要目标。

	吞吐量	响应时间	完成时间	资源占用	可用性	可扩展性
电商系统	极高	中	中	高	高	高
IoT	低	较短	较短	极低	中	中
电信业务	高	短	短	低	极高	高
HPC	较高	中	极长	较高	高	极高
大数据	中	中	长	高	高	高
云计算	高	中	较长	较高	高	高
区块链	低	中	长	高	极高	低

不同业务、系统对各个指标的要求

总结与思考

按照不同维度，分布式系统的指标可以分为性能、资源占用、可用性、可扩展性这四大类。我们自然希望自己的系统，是高性能、高可用、高扩展和低资源占用的，但考虑到硬件成本、开发效率等因素，必须要在设计不同的系统、业务时有所取舍。

所以，我又和你分析了典型的电商、IoT、电信、HPC（高性能计算）、大数据、云计算、区块链等业务或系统的不同诉求，进而得出了系统设计者需要关注哪些指标。你在设计其他类型的系统时，可以按照这个思路进行取舍。

我在文中提到了，分布式系统的指标之间会存在一些冲突或约束。那你不妨思考一下：我们今天讲解的指标中，哪些指标之间是相互制约、相互冲突的，它们又是如何制约的呢？

我是聂鹏程，感谢你的收听，欢迎你在评论区给我留言分享你的观点，也欢迎你把这篇文章分享给更多的朋友一起阅读。我们下期再会！

精选留言 (36)



亦知码蚁

2019-09-23

软件测试是这么写的 可伸缩性和可拓展性

可伸缩性翻译自 Scalability，指的是通过简单地增加硬件配置而使服务处理能力呈线性增长的能力。最简单直观的例子，就是通过在应用服务器集群中增加更多的节点，来提高整个集群的处理能力。

而可扩展性翻译自 Extensibility，指的是网站的架构设计能够快速适应需求的变化，当需要增加新的功能实现时，对原有架构不需要做修改或者做很少的修改就能够快速满足新的业务需求。

作者回复：从你的留言看出，你的知识面确实比较广。

在分布式领域将Scalability翻译成可扩展性的情况确实会更多一些，特别是在工业界。从可扩展性和可伸缩性在搜索引擎中，分别检索出来的结果数可以窥见一斑。

即便是严肃的学术界，我们也会看到可扩展性与可伸缩性互换表示Scalability的场景，比方说经典的两本分布式教材中，《分布式系统概念与设计》使用了可伸缩性的翻译（见原书1.5.4节），而《分布式系统原理与范型》使用了可扩展性的翻译（见原书1.2.4节）。

这种现象告诉我们不能教条的去照搬书上的东西，必须既要结合具体语境去理解问题，更要结合具体的问题去寻找最佳方案。

Robic，谢谢你的抛砖引玉！

共 16 条评论 >

👍 50



开心小毛

2019-09-24

离开响应时间的要求是无法衡量QPS的，例如一个每秒处理100个查寻且99%响应时间200秒的系统，同样可以每秒处理1000个查询且99%响应时间1秒。如果继续放松对响应时间的要

求，每秒处理查询数峰值可能到达5000，但99%响应时间已经到了无法忍受的程度，所以QPS不能被定为5000。

作者回复：首先非常感谢你的留言。肯定是像你说的这样，离开了约束去谈任何指标都是没有意义的。这也是我为什么在文章中谈到各种指标会相互制约、相互冲突的原因。所以，你的回复也是在一定程度上回复了我在文末留下的思考题。

在这里，我替所有的订阅者感谢你精彩的回复！！

共 3 条评论 >

👍 25



开心小毛

2019-09-24

把QPS解释成每秒处理的查询数是有问题的，应该解释为可确保某给定响应时间下的每秒到达的查询数的上限。例如，某单核系统在99%分位200ms响应时间的要求下，系统最多能承受1000个查询，则QPS为1K，而不是5。

作者回复：再次感谢你的留言。正如刚才我讲到的：离开了约束去谈任何指标都是没有意义的。在解释BPS的时候，我也提到了查询和查询，事务和事务之间也是不尽相同的。不尽相同的地方也就包括你所说的响应时间。其实影响QPS的不仅仅只是响应时间，还包括其它的约束，如资源占用等。为了不过多影响阅读体验，我并没有把所有的约束放在定义中，而是选择把这个点作为一个课后习题。

在这里，我再次替所有的订阅者感谢你精彩的回复！！

也欢迎你再次思考文末的问题。期待你更精彩的留言！



👍 16



Geek_f6f02b

2019-10-04

CAP原理就是讲一个系统无法同时满足一致性，高可用性与分布式。至多只能满足其中2项。单机就是CA一致性且高可用（时间上的高可用，没有同步数据耗时，相同时间可以处理更多问题），如今的分布式肯定是要有P，那么剩下的只能选CP跟AP了。像cdn服务就属于AP高可用且为分布式，像pxc数据库就属于CP一致性且分布式，不知道这样理解对不对？很明显你要满足一致性且分布式就不可能满足高可用，因为同步数据肯定要耗时。其它也一样，不是在于技术是否能实现，而是这个是内在矛盾，不可能通过技术解决，只能取舍。



👍 6



leslie

2019-09-24

老师今天的東西里面是不是有疏漏啊？吞吐量里面其实就有一对有冲突的：QPS和TPS是很难做到没有没有冲突的；存储中间件/数据系统中大多数都有这种问题，RMDDB是最典型的，即使是做了读写分离同样无法解决，故而才会从过去的模式升级出一主多从、双主、、、各种数据模型；甚至我们的硬件磁盘都是读性能远强于写性能。吞吐量内部就有一对互相约束的指标。

资源占用和可用性其实是有约束的：物联网中的这对其实在MQ上就非常非常明显，不然就不至于需要特意去使用MQTT协议；9月24号《消息队列高手课》中刚好强调了这点。

其它就暂时不知：可能学习学习的还不够深入吧；毕竟学无止境，觉得明白了一些懂了一些发现还有关联，就像之前的某些课程学习中发现学员感悟其实是另外一门同时在学的老师写的。期待老师答案的揭晓：谢谢老师的教诲。

作者回复：没有问题的哦。我在介绍性能指标刚开始就强调了“不同的系统、服务要达成的目的不同，关注的性能自然也不尽相同”，QPS、TPS作为吞吐量指标当然是性能指标。你的留言质量非常高，从一定程度上也回复了我文末的思考题。

加油！继续保持这种学习+思考+分享的习惯！期待你更多留言！



👍 6



张理查

2019-12-20

分布式技术原理与算法#Day2

分布式就是更多的机器处理更多的数据，更复杂的业务。就像算法要先知道时间复杂度与空间复杂度一样，分布式系统也有其评价体系，我们希望的是吃的少跑得快跑不死的马🐎。就是想马儿跑，又想马儿不吃草，还想马儿死不了。也就是性能好，资源占用低，高可用的可扩展系统。

性能包括了吞吐量，响应时间和完成时间。读业务关注qps,写业务关注tps，传输关注bps。资源关注cpu ram 磁盘 gpu ssd等等，用得越少越好，要看系统不跑的空载占用，也要看系统跑到极限的满载占用。

可用性就是系统的故障率，越小越可用，这里还有个可靠性的概念，可以理解为可靠性对标的硬件，但硬件的不可靠会带来系统的不可用，因此经常模糊。

可扩展性就不用说了，是分布式的初衷，但好的分布式系统是线性扩展的，即数量增加一倍，性能提升一倍，但这仅仅是理想情况。

当然不可能有上面提到的神马，一定是有取舍的，跑得快一般吃得多，吃得多容易死得早，而活得久又不能跑太快，要看关注点。

作者回复: 优秀, 总结的非常好, 这个马的比喻确实很形象!



👍 5



周涛

2019-09-26

老师, 你讲的很好, 对于我这样在门槛边上的初学有很大的指引, 我也发现分布式概念太多, 内涵很广, 如果光是听你的这样的课程, 恐怕过不了十节课, 就会陷入概念的漩涡, 请教下老师, 是否有实践性的指导, 比如什么实验平台的介绍和操作, 让我们能够动手做, 这样对我们的认知有很大的帮助。有没有这样的实际操作的平台或者小型操作指导呢?

共 5 条评论 >

👍 4



蚂蚁内推+v

2019-09-24

LOT 是什么系统能请教下吗

作者回复: 你说的应该是IoT吧? IoT是Internet of Things物联网的缩写

共 3 条评论 >

👍 3



Scarf

2019-10-04

您好! 请问, 正文中的“各系统对于各个指标的要求”这张表格, 是不是部分内容有误? 正文里说了“吞吐量和完成时间通常是区块链系统设计者的首要目标”, 但是表格里面对于区块链里, 这两个指标的要求分别是低和中, 怎么理解呢?

共 3 条评论 >

👍 2



.

2022-03-01

可用性那里5 个 9 (或 9N, 9 Nines) 应该是5N, 5Nines吧



👍 1



Gopher

2019-09-29

老师的讲解很清晰, 知识结构化良好, 很赞, 感谢!

btw: 如果把robustness翻译成“健壮性”就更好了。技术领域本就概念繁多，没必要再徒增烦恼，对新人也不友好。



zhaozp

2019-09-29

打卡文章学习：

1、分布式系统的衡量指标：性能、资源占用、可用性和可扩展性。

性能：吞吐量（QPS、TPS、BPS）、响应时间、完成时间。

资源占用：系统正常运行占用的硬件资源，比如CPU、内存、硬盘等。

可用性：系统停止时间与总时间的占比，或者某功能失败次数与总请求数据之比衡量。就是我们常说的4个9、5个9。

可扩展性：垂直扩展和水平扩展。

2、不同的分布式场景，所衡量的指标侧重点是不一样的，系统设计时需要有所取舍。

作者回复：积跬步，而终至千里！加油！



William Ning

2022-03-12

保持学习



ky

2021-10-16

脑子中出现了一个伪公式:吞吐量*响应时间=资源占用



A君

2021-03-12

分布式系统的衡量指标，qps，tps，bps，响应时间，完成时间，资源占用率，可用率，加速比等



abc

2020-10-08

在 24 小时内，如果网站故障导致不可用的时间是 4 个小时，那么系统的可用性就是 $4/24=0.167$ 。

感觉应该是20/24?



西门吹牛

2020-08-12

在这几个指标中，可用性是分布式系统首要考虑的，只要保证了可用，才能在可用的基础上进行其他 指标的提升。

响应时间和吞吐量：两者是依赖的，响应短了，吞吐量就上去了，但是响应时间的缩短，是有难度的，要考虑响应的数据的实时性问题，如果对实时性要求不高，次要业务的响应可以给核心业务让路。如果所有数据都要求强一致性，那根据CAP原则，首先C是必须保证的，A和P是不能同时满足的。一般的分布式系统都是AP就可以了。

可扩展性：好的可扩展性是建立在好的分区容错性上的，一味的增加机器，容错性不好，还是不行的。

资源占用：合理分配内存，合理设计数据资源，优化业务逻辑和系统设计是关键。



鸭先知

2020-03-29

分布式的目的是用更多的机器，处理更多的数据和更复杂的任务。高性能、低资源占用、高可用性和高可扩展性是衡量一个分布式系统是否成功的重要指标，单这四个目标难以兼顾。

作者回复：通常情况下是根据业务特征，综合考虑这四个指标



钱

2020-02-12

阅过留痕

挺好的把分布式的三围都给我们看了，下面就可以好好欣赏分布式女神的舞步啦！

根据自己压测的经验，我们最看重的指标主要是TP99+TPS+CPU+MER，马儿比喻很棒！



linker

2020-01-20

qps与tps.可用性与可靠性他们互相制约

