

分布式系统

What

定义

什么是分布式系统：把多个物理上分离的终端融合成一个虚拟的超级终端

在一个分布式系统中，一组独立的计算机展现给用户的是一个统一的整体，就好像是一个系统似的。系统拥有多种通用的物理和逻辑资源，可以动态的分配任务，分散的物理和逻辑资源通过计算机网络实现信息交换。系统中存在一个以全局的方式管理计算机资源的分布式操作系统。

分布式系统是多个处理机通过通信线路互联而构成的松散耦合的系统。从系统中某台处理机来看，其余的处理机和相应的资源都是远程的，只有它自己的资源才是本地的。至今，对分布式系统的定义尚未形成统一的见解。一般认为，分布式系统应具有以下四个特征：

1. 分布性。分布式系统由多台计算机组成，它们在物理上是分散的。整个系统的功能是分散在各个节点上实现的，因而分布式系统具有数据处理的分布性。
2. 自治性。分布式系统中的各个节点都包含自己的处理机和内存，各自具有独立的处理数据的功能。通常，彼此在地位上是平等的，无主次之分，既能自治地进行工作，又能利用共享的通信线路来传送信息，协调任务处理。
3. 并行性。一个大的任务可以划分为若干个子任务，分别在不同的主机上执行。
4. 全局性。分布式系统中必须存在一个单一的、全局的进程通信机制，使得任何一个进程都能与其他进程通信，并且不区分本地通信与远程通信。同时，还应当有全局的保护机制。系统中所有机器上有统一的系统调用集合，它们必须适应分布式的环境。在所有 CPU 上运行同样的内核，使协调工作更加容易。

优点

1. 资源共享。若干不同的节点通过通信网络彼此互联，一个节点上的用户可以使用其他节点上的资源；可以共享远程文件，使用远程特有的硬件设备（如高速阵列处理器），以及执行其他操作。
2. 加快计算速度。如果一个特定的计算任务可以划分为若干个并行运行的子任务，则可把这些子任务分散到不同的节点上，使它们同时在这些节点上运行，从而加快计算速度。另外，分布式系统具有计算迁移功能，如果某个节点上的负载太重，则可把其中一些作业移到其他节点去执行，从而减轻该节点的负载。这种作业迁移称为负载平衡。
3. 可靠性高。分布式系统具有高可靠性。如果其中某个节点失效了，则其余的节点可以继续操作，整个系统不会因为一个或少数几个节点的故障而全体崩溃。因此，分布式系统有很好的容错性能。
4. 通信方便、快捷。分布式系统中各个节点通过一个通信网络互联在一起。通信网络由通信线路、调制解调器和通信处理器等组成，不同节点的用户可以方便地交换信息。在低层，系统之间利用传递消息的方式进行通信，这类似于单CPU系统中的消息机制。单CPU系统中所有高层的消息传递功能都可以在分布式系统中实现，如文件传递、登录、邮件、Web浏览和远程过程调用(Remote Procedure call, RPC)。
5. 便于拓展设备 可以便捷地在分布式系统中添加或者删除设备以达到增强系统功能的目的

缺点

尽管分布式系统具备众多优势，但它也有自身的缺点，主要是可用软件不足，系统软件、编程语言、应用程序以及开发工具都相对很少。此外，还存在通信网络饱和或信息丢失和网络安全问题，方便的数据共享同时意味

着机密数据容易被窃取。虽然分布式系统存在这些潜在的问题，但其优点远大于缺点，而且这些缺点也正得到克服。因此，分布式系统仍是人们研究、开发和应用的方向。

文件系统（摘取自x-gkd）

业界对于分布式文件系统的追求无非就是几个方面：

搭建部署系统操作是否繁复

传输和储存文件是否安全可靠

文件系统性能是否满足使用要求

对计算机的性能要求是否高

对文件的格式是否有限制

结合对往届项目的实现的研究，我们计划并完全实现了这样一个家庭式分布式文件系统：

高效利用分散在不同设备上的存储空间，方便多人合作办公，避免资源浪费 所有安装了客户端的机器共享、贡献存储空间和存储文件碎片，并对数据做好高效的备份；一个专门用于协调处理请求，维持系统状态的服务端；有浏览器的电脑都可访问分布式文件系统 在以上要求外，尽可能让性能相对好，安全性相对高，操作简单 在调研报告和可行性报告部分，我们已经对于我们实现思路做了分析。现在首先简要概括本项目的完成情况：

在原基于互联网的小型分布式文件系统项目的基础上，小组成员齐心协力，学习rust语言并完成了client、server端的完全改写，实现了整个分布式文件系统的逻辑架构。此外，采用前后端分离的模式，采用actix-web框架将web应用的后端用rust语言改写，并与原项目的web前端衔接，最终呈现出完整的分布式文件系统，并在浏览器端提供了用户对文件系统的管理操作。

<https://github.com/OSH-2020/x-gkd>

一些想法

1. 鸿蒙怎么用

1. 鸿蒙体量相当的大 学其一完成大作业就已经绰绰有余了 包括他的软总线 数据库 文件系统
2. 鸿蒙上提供了一些codelab开源代码，其中有不少项目与分布式相关，其中的操作和思想可以学习。

2. 现在要干什么

说实话 就现在这个讲课内容 我相信大家都还是茫然的，但是有一些事情我们可以先做。通过这些天的查阅，不难发现这么多的分布式系统基本都是java写的，我们可以先学习并熟悉java语言

3. 有什么方向？2020年的x-gkd项目是基于2017年的一个java分布式文件系统改进的，我们可以结合鸿蒙操作系统的理念，爆改他们的文件系统。比如添加融合搜索功能，改进软总线提高数据传输速度，改进数据库这种（虽然我不知道数据库是什么）