

## 【第57话：今天一名同学被问到的CAP和BASE】

---

Hello 小伙伴们，这节课给大家带来一个在分布式系统中非常重要的两个概念，这两个概念也是面试过程中问到分布式相关时常见的问题，它们就是“CAP定理和BASE理论”。

我们先来说一说CAP定理。CAP定义称为分布式一致性定理。是指在一个分布式系统中Consistency(一致性)、Availability(可用性)、Partition tolerance(分区容错性)，最多同时满足其中两个，三者不可兼得。

一致性 (C Consistency)

在分布式系统中，是否立即达到数据同步的效果（平时多说的强一致性）。在分布式系统一般都是最终会一致的。如果请求时，整个分布式系统同步后才返回结果，叫做强一致性（满足一致性）。如果先返回结果，在一定时间后才实现一致性就叫做弱一致性，就不满足C的要求。

例如：在Zookeeper中，每次请求Zookeeper集群时，必须等待消息同步后才会响应结果，这就满足一致性。

可用性 (A Availability)

在集群中一部分节点出现故障后，整个集群是否还能响应客户端请求。

例如：在Eureka集群中，即使集群中某些节点挂掉了，整个集群依然能够正常运行，这就满足可用性。

而在Zookeeper中，如果Leader挂掉了，必须等待重新选举后集群才可用，这就不满足可用性。

分区容错性 (P Partition tolerance)

在分布式系统中，是否可以在有限的时间内达到数据一致的效果，如果因为网络等问题最终没有达到一致性，这时称为出现分区错误，此时就必须在A和C中做选择。

然后我们在说一说BASE理论。

是指Basically Available（基本可用）、Soft state（软状态）和Eventually consistent（最终一致性）三个短语的缩写。

BASE理论是对CAP中一致性和可用性权衡的结果，是基于CAP演化而来的。

BASE理论核心思想：即使无法做到强一致性，每个应用都可以根据自身业务特点，采用适当的方式达到最终一致性。

基本可用 (BA)

是指在分布式系统中出现不可知故障的时候，允许损失部分可用性。此处要注意：损失部分可用性，不代表整个系统不可用。

例如：

(1) 可以增加响应时间。由之前0.5秒，在出现故障的时候变成1~2秒。

(2) 由于一些特殊原因，使网站访问流量激增，为了保证整个系统的稳定性，部分访问者可能被引导到降级页面中。

软状态 (S)

是指系统中数据允许存在中间状态（软状态），并认为这个状态是不影响系统的可用性的。通俗解释：允许分布式节点之间存在同步延迟。

例如：

在Eureka集群中数据同步时就存在软状态。

### 最终一致性

允许整个系统中数据在经过一定时间后，最终能达到整个系统的一致性。但是这个时间绝对不可以过长。

强一致性要求系统接收请求后，整个系统必须达到一致性效果，才会响应结果。

最终一致性是弱一致性的特例。满足最终一致性的系统在响应给用户结果时整个系统可能是没有达到一致性的，但是最终一定会达到一致性效果的。