

分布式系统的定义

- **分布式系统**是由一组进程或节点组成的系统，这些进程或节点相互协作，以实现某些共同的目标。
- **进程/节点**可以是核心、处理器、物理或虚拟机、集群等。

分布式系统的特点

1. **进程没有系统的全局视图**：在协同工作的过程中，尽管各个环节都在为共同目标努力，但单个进程可能无法完全掌握整个系统的状态。
2. **故障和异步性**：特别是当故障和异步性结合时，它们常常影响问题的可解决性。
3. **并发**：分布式系统中的并发是一个主要挑战。

分布式系统中的进程崩溃

- 进程可能因各种原因崩溃，如忽略消息发送/接收、恶意或任意故障（拜占庭式故障）等。
- **通信异步**处理：网络故障（如交换机故障）可能导致消息丢失或延迟，不同的处理速度也引入了异步性。
- **故障检测的不可靠性**：依赖时间假设来检测进程故障是不可靠的。

安全性与活性

- **安全性 (Safety)**：指明不应发生任何坏事的属性。
- **活性 (Liveness)**：指明应该发生好事的属性。
- **例子**：安全性是“不说谎”，而活性是“说话”。结合起来就是“（总是）说实话”。
- 有用的分布式服务应该同时提供活性和安全性。

解决方案和正确性

- 提出解决方案并根据规范证明其正确性。
- 注意到“正确性可能是理论上的，但错误会产生实际影响”。
- **分析复杂性**：考虑时间复杂性（影响延迟）和消息复杂性（影响吞吐量）。
- **编码方式**：使用伪代码，没有特定的编程语言。

关于共识的定理

- 如果底层通信信道容易出现任意消息丢失，则无法解决两个进程之间的共同决策问题。
- 这涉及到一个基于矛盾证明的理论，表明如果存在消息丢失，那么无法保证两个进程间的有效通信和共识。

A1和A2执行消息交换的连续阶段

矛盾证明

假设算法A使用最少的通信相位

当A终止时（即{yes, yes}），存在发送消息的最后阶段，例如由A1向A2发送的消息m

A1执行的最后一个语句不能取决于A2是否接收到m（因为m可能丢失）

类似地，A2执行的最后一个语句不能依赖于m

m是无用的，可以移除

存在通信相位较少的A

