

《End-to-End arguments in system design》阅读报告

计 01 2020010869 容逸朗

这篇文章主要探讨了计算机系统设计的两个问题。即如何合理地划分功能，以及如何决定功能在系统中所处的层级。文章通过讨论不同实现方式的性能及代价，引出了不同系统功能设计的取舍之道。

首先，考虑文件传输问题的例子。在传输的过程中需要考虑到硬件错误、软件错误、网络丢包、系统崩溃等一系列问题。为了数据的稳定传输，一种直观的方法是通过错误检测、增加冗余等手段使得网络传输过程的错误率尽量低。但无论网络层的可靠性有多高，也不能保证传输过程中的其他环节不会出现错误。此时仍然需要通过端对端的校验来确保文件传输正确无误。

如果实现可靠传输的功能是放在系统中较低的层级实现时，会使得所有应用都需要负担更多代价。但是，并非所有的应用程式都需要很高的传输稳定性。例如一些语音通信软件或是视频网站，这些应用对于传输延迟的要求远远高于对稳定性的需求，即使是出现了错误，也不会过度影响用户体验。因此简化网络设计，把一部分验证功能放在系统中尽量靠近应用层的位置实现是可以使整体系统的效率最大化的。

随着世界发展，网络系统的体量只会越来越庞大。此时网络分层及功能简化不但可以降低维护系统的难度及支出，提升系统整体性能，同时也方便了后续网络系统的升级。尽管未来对于通信安全的要求可能会更高，但是该文章提到的端对端原则的开放性思想仍有实用价值。