

解疑：为什么通过网络传任何文件都要验证

TCP，是带有数据反馈的协议，它是安全的，使用 TCP 传输数据，我们不需要验证。

数据的产生方，诸如 RAR,7z,zip,文件 copy，他们会经过内存，而内存会受到 CPU 内部合并了南北桥的内存控制器影响，这会直接导致我们的数据产生时，发生错误，即使是一个 8bit 位的数据错误，也会造成严重后果，因为有数据噪音，对于小数据来说，产生数据时出错的可能性很小，对于大数据，到达数十 GB 体量以后，数据出错的可能性会很大。

在硬件方便，服务器设备，提供了 ECC 这类 mem test 的硬件技术，这是降低错误率的一种方法。另一方面，主板和 CPU，本身也带有对内存数据降噪的功能，但是，它们并不能和 ECC 划上直接的等号，因为很多 CPU 和主板，为了追求更高的频率，往往会关闭降噪这类功能，要不，信仰党每次超完频，就不会做内存和 CPU 的烤机测试。因为终端设备，往往都不是 ecc，更加不会是带有统一内存降噪的总线机制。

经过长时间的各方验证，在设计 ZServer4D 的 BigStream 支持时，特意留有一个带有 MD5 验证的机制，同时，对于虚列包，DataFrameEngine，内部都有数据验证机制。

或许，你会说，既然数据都错了，系统还不崩啊。其实原因是这样的：操作系统占用的内存其实并不大，而我们打包发布，传输大文件，会使用大量的缓存，其体谅远远超过操作系统内核，当内存使用的体谅达到几十甚至数百 GB 时，一个 bit 的错误往往会引发大量的连锁反应，最后是雪崩效应。

做大数据，或则手上的业务涉及到大文件，大型数据的传输，加个 md5,crc，或许影响性能，但是，它更安全，我选择更安全的做法。

对于 ZServer4D 在 BigStream 使用 MD5 的做法就此解释完毕。

By.qq600585

2019-4