# 数据交换格式

## JSON

JSON 是 JavaScript 原生格式，这意味着在 JavaScript 中处理 JSON 数据不需要任何特殊的 API 或工具包。JSON可以将JavaScript对象表示的一组数据转换为字符串，然后在函数之间传递这些字符串，或者在异步程序中将字符串从Web客户端传递到服务器端。

JSON是一种轻量级的数据交换格式，没有文件的存储形式。

JSON的对象object数据格式：（name/value），name是String类型的，value是String、number、Boolean、null、Object或数组类型。数组array是多个value的有序序列。

## XML

XML是SGML标准通用标记语言的子集，允许用户对自己的标记语言进行定义的源语言。格式统一，跨平台和语言，业界公认的标准。

## JSON与XML的比较

JSON在JavaScript中比XML有优势，一般在Web开发中不用XML作为数据交换格式，它的解析会造成复杂的编码和降低运行效率。

### 3.1 解析方法

目前，XML提供两种解析方案：DOM和SAX。

DOM需要把整个XML文件读入内存，这一点上DOM和SAX是一样的，但是XML需要考虑父节点和子节点，而JSON是键值对的集合，他比XML解析难度小很多。

SAX不需要读入整个文档就可以解析出内容，是一种逐步解析的方法，适合大规模的解析，这一点JSON做不到。

所以JSON适用于少量数据的传递，而大量数据的处理依靠XML。

### 3.2编码

JSON的语法简单，XML的语法规范；JSON编码简单，而XML更适用于标记文档。都是基于文本的，且都是用Unicode编码。

## 4 . Protocol Buffer

Protocol Buffer是一种平台中立、语言中立的可扩展的序列化结构数据方式，是Google公司开发的一种数据描述语言，类似于XML，在仅仅描述一个对象的时候可以代替XML，可用于数据存储和通信协议方面。与XML相比更简单，数据描述文件只需要原来的1/3或1/10，解析速度是原来的20到100倍，减少了二义性，使用自动工具生成数据访问类。目前针对于C++、Java、Python的开发者。

它的数据保存格式是基于二进制的，不是自描述的，必须有一个格式定义文件(.proto文件)。

### 4.1 解析工作过程

写一个.proto文件，利用protocol buffer编译器生成C++、Java、Python数据访问类，这些数据访问类提供了简单的访问数据项的方法。数据的输入输出采用的还是流（Stream）的概念。