Atitit 网络协议的几种格式 二进制 纯文本+命令符 结构化文本

目录

[1. 分类 1](#_Toc288)

[1.1. 二进制字节块格式、文本格式和结构化文本格式3种典型的数据格式。 1](#_Toc20658)

[1.2. 实际上，数据比如C#的Linq。 1](#_Toc18778)

[2. 二进制字节块格式 2](#_Toc12535)

[3. 文本格式 2](#_Toc11855)

[3.1. 文本流的管道处理 2](#_Toc24298)

[3.2. 空格+命令符 msn协议 2](#_Toc17318)

[4. 结构化文本格式 2](#_Toc8317)

[4.1. Markdown 2](#_Toc23748)

[4.2. 语言源码类 3](#_Toc5053)

[4.3. Xml json类 3](#_Toc16087)

# 分类

## 二进制字节块格式、文本格式和结构化文本格式3种典型的数据格式。

## 实际上，数据比如C#的Linq。

**3、XMPP通过TCP传什么了?**

传输的是与即时通讯相关的指令。在以前这些命令要么用2进制的形式发送(比如QQ)，要么用纯文本指令加空格加参数加换行苻的方式发送(比如MSN)。而XMPP传输的即时通讯指令的逻辑与以往相仿，只是协议的形式变成了XML格式的纯文本。这不但使得解析容易了，人也容易阅读了，方便了开发和查错。而XMPP的核心部分就是一个在网络上分片断发送XML的流协议。这个流协议是XMPP的即时通讯指令的传递基础，也是一个非常重要的可以被进一步利用的网络基础协议。所以可以说，XMPP用TCP传的是XML流。

# 二进制字节块格式

在偏底层的操作系统、嵌入式和通信系统中，二进制的字节块是最常见的一种数据格式。二进制数据布局紧凑和接近机器的特点使得它常常作为系统间通信或系统文件的数据格式。但一般高级语言不方便直接和0101打交道，而是基于记录、结构体和类等结构化表示操作数据，这就存在着在底层的二进制字节块和高层的结构化数据之间的转换问题。

# 文本格式

## 文本流的管道处理

文本格式是另一种十分常见的数据格式。《Unix编程艺术》中是这样描述文本格式的："Text streams are a valuable universal format because they're easy for human beings to read, write, and edit without specialized tools ”。基于文本流的管道处理是一种备受赞誉的Unix风格。Shell可以通过管道把各种功能单一的命令串联起来，让文本流在管道上流动，因而Shell语言具有很好的文本数据亲和力。许多文本数据处理任务Bash都可以一行搞定，这就是Hacker们酷爱的One Liner风格。

下面我们来看两个用Bash进行文本处理的例子：

1. 统计当前目录下的gz文件数目： ls –l \*.gz | wc –l

## 空格+命令符 msn协议

# 结构化文本格式

## ****Markdown****

## 语言源码类

## Xml json类

XML是最近十几年来流行起来的一种通用（半）结构化的文本数据交换格式。XML除具有一般文本格式的优点外，还具有表达复杂的层次信息的优势，所以它至诞生以来就被大量用于配置文件和各种Web Service中。现代程序设计基本都少不了了XML打交道，不过在C++、Java和C#集中静态类型语言中处理XML却并不是一件十分轻松的事情。我们先来看一个Java解析和构建下面这个XML的例子

编程语言与数据的亲和力 - 漫漫字节|漫漫编程

Atitit ****数据交换语言之道 attilax著 xml json yml****