

大多数 URL 方案的 URL 语法都建立在这个由 9 部分构成的通用格式上：

```
<scheme>://<user>:<password>@<host>:<port>/<path>;<params>?<query>#<frag>
```

几乎没有哪个 URL 中包含了所有这些组件。URL 最重要的 3 个部分是方案 (scheme)、主机 (host) 和路径 (path)。表 2-1 对各种组件进行了总结。

表2-1 通用URL组件

组 件	描 述	默 认 值
方案	访问服务器以获取资源时要使用哪种协议	无
用户	某些方案访问资源时需要的用户名	匿名
密码	用户名后面可能要包含的密码，中间由冒号 (:) 分隔	<E-mail 地址>
主机	资源宿主服务器的主机名或点分 IP 地址	无
端口	资源宿主服务器正在监听的端口号。很多方案都有默认端口号 (HTTP 的默认端口号为 80)	每个方案特有
路径	服务器上资源的本地名，由一个斜杠 (/) 将其与前面的 URL 组件分隔开来。路径组件的语法是与服务器和方案有关的 (本章稍后会讲到 URL 路径可以分为若干个段，每段都可以有其特有的组件。)	无
参数	某些方案会用这个组件来指定输入参数。参数为名 / 值对。URL 中可以包含多个参数字段，它们相互之间以及与路径的其余部分之间用分号 (;) 分隔	无
查询	某些方案会用这个组件传递参数以激活应用程序 (比如数据库、公告板、搜索引擎以及其他因特网网关)。查询组件的内容没有通用格式。用字符 “?” 将其与 URL 的其余部分分隔开来	无
片段	一小片或一部分资源的名字。引用对象时，不会将 frag 字段传送给服务器；这个字段是在客户端内部使用的。通过字符 “#” 将其与 URL 的其余部分分隔开来	无

比如，我们来看看 URL：http://www.joes-hardware.com:80/index.html，其方案是 http，主机为 www.joes-hardware.com，端口是 80，路径为 /index.html。

26

2.2.1 方案——使用什么协议

方案实际上是规定如何访问指定资源的主要标识符，它会告诉负责解析 URL 的应用程序应该使用什么协议。在我们这个简单的 HTTP URL 中所使用的方案就是 http。

方案组件必须以一个字母符号开始，由第一个 “:” 符号将其与 URL 的其余部分分隔开来。方案名是大小写无关的，因此 URL “http://www.joes-hardware.com” 和 “HTTP://www.joes-hardware.com” 是等价的。