

- 拦截代理是网络流量中的代理服务器，它会拦截从客户端发往服务器的请求，并提供一个缓存响应，或对其进行转发。由于拦截代理拦截了从客户端到服务器的流量，所以它会收到发送给 Web 服务器的部分 URI（参见图 6-15d）。⁹

6.5.4 代理既可以处理代理请求，也可以处理服务器请求

由于将流量重定向到代理服务器的方式有所不同，通用的代理服务器既应该支持请求报文中的完整 URI，也应该支持部分 URI。如果是显式的代理请求，代理就应该使用完整 URI，如果是 Web 服务器请求，就应该使用部分 URI 和虚拟 Host 首部。

146

使用完整和部分 URI 的规则如下所示。

- 如果提供的是完整 URI，代理就应该使用这个完整 URI。
- 如果提供的是部分 URI，而且有 Host 首部，就应该用 Host 首部来确定原始服务器的名字和端口号。
- 如果提供的是部分 URI，而且没有 Host 首部，就要用其他方法来确定原始服务器：
 - 如果代理是代表原始服务器的替代物，可以用真实服务器的地址和端口号来配置代理；
 - 如果流量被拦截了，而且拦截者也可以提供原始的 IP 地址和端口，代理就可以使用拦截技术提供的 IP 地址和端口号（参见第 20 章）；
 - 如果所有方法都失败了，代理没有足够的信息来确定原始服务器，就必须返回一条错误报文（通常是建议用户升级到支持 Host 首部的现代浏览器）。¹⁰

6.5.5 转发过程中对 URI 的修改

代理服务器要在转发报文时修改请求 URI 的话，需要特别小心。对 URI 的微小修改，甚至是看起来无害的修改，都可能给下游服务器带来一些互操作性问题。

尤其是，现在已知有些代理会在将 URI 转发给下一跳节点之前将 URI “规范”为标准格式。有些看起来无害的转换行为，比如用显式的 “:80” 来取代默认的 HTTP 端口，或者用适当的换码转义符来取代非法的保留字符以校正 URI，就可能造成互操作性问题。

总之，代理服务器要尽量宽容一些。它们的目标不是成为强制实现严格协议一致性

注 9：在某些情况下，拦截代理可能也会拦截客户端到代理的流量，在这种情况下，拦截代理可能会收到完整 URI，并需要对其进行处理。由于显式代理的通信端口通常与 HTTP 使用的端口有所不同（通常是 8080 而不是 80），而且拦截代理通常只对端口 80 进行拦截，所以这种情况并不会经常发生。

注 10：不应该经常这么做。因为用户会收到之前从未收到过的神秘错误页面。