安全 WEB 隧道技术

孙 琨 王艳峰 （南开大学计算机系 300071）

摘要：当前 internet 应用中，存在着企业组织的网络边界和防火墙的范围不完全一致的问题。企业的职员希望能从防火墙外部访问内部的 WEB 服务，就如同在内部网中访问一样的简便和安全，同时，防火墙能够执行充分的访问控制，记录和监听功能。本文讨论的安全 WEB 隧道技术可以用于解决这个问题。

关键词：安全 WEB 隧道 防火墙 HTTPS 安全套接口层

一、简介

安全 WEB 隧道的实现可以实现企业职员使用公共终端从防火墙外部访问内部的 WEB 服务器。此时，防火墙应对此人透明，就如同他在防火墙内部一样，同时，防火墙仍可以对攻击进行检测和监听。

主要技术特性：

1. 所有的内部 WEB 服务器可以从外部网络访问，无须特别改动。

2. 从防火墙外部访问 WEB 与从内部访问相类似，统一资源定位符（URL）并不是完全混合的，也就是说负责人在外部获得的 URL 回到防火墙内部仍然可以使用。

3. 防火墙和要访问内部网的负责人是相互认证的。它们之间的通信是防偷听的。

4. 防火墙内部认证和监听的防护机制并不发生改变，不需要新的机制。

5. 防火墙能够记录和监听流经防火墙的信息。

6. 客户端不需要安装特殊的软件，普通的浏览器就足够了。

7. 管理负载很小，执行操作十分简单。

安全 WEB 隧道技术运行在 HTTP 级。一般来讲，防火墙允许 IP 包，TCP 流，HTTP 请求和应答，或者其他协议中的数据通过。使用低层协议的访问可以提供较多的通用性，但另一方面，当使用应用层协议时，我们可以充分利用应用层的支持进行访问控制，记录和监听。当应用 HTTP 协议时，我们可以允许或限制 HTTP 请求，而不是对低层的 IP 包进行处理。

为了简化安全模型，此技术不涉及经过多个不同组织的防火墙的连接。比如，A 在 B 的防火墙内与 A 的内部网进行通信，此时 A 收发的数据均可被 B 中的节点监听到。

二、相关的技术

如今使用的通过防火墙访问内部网的技术有很多种，它们各自有其特点和缺点：

1. 许多内部网允许远程用户通过拨号线使用 IP 协议直接对其访问。虽然它们具有用户认证机制，但未加密的数据通过不安全的电话线或 ISP 的网络传输，而且不具备记录和监听的功能。拨号连线需要 Modem 和电话线，许多公用终端不能提供这些条件。

2. 安全 IP 隧道领域已经取得很大进展。一条 IP 隧道由一对包路由器组成，它们在高层协议上相互转发 IP 数据包。因为使用了加密技术，封装协议能够保证路由器的可靠性及 IP 包的安全。同时，网络管理员可以允许 IP 隧道通过防火墙，防火墙允许外部的可信用户与内部网之间进行 IP 包路由选择。

由于这种方法允许可信用户的所有 IP 包进入内部网，而我们仅需要 HTTP 包。因为工作在 IP（IP 路由）层，它不可能记录和监听所有流经防火墙的信息，所以安全漏洞不易被发现。这样一来，整个内部网的完整性依赖于可信用户在手提或家用计算机中对低层协议的配置。因为上述原因，防火墙管理员经常对 IP 隧道进行限制，另外公用终端使用 IP 隧道时，需要安装额外的软件。

3. 通过配置防火墙使指定的内部网主机可以接受 INTERNET 的连接请求。访问这些主机时，外部请求需要经过认证，比如使用 SSL 协议。实际应用中，这种方法在多数内部网中不可行。对于防火墙管理员来说，指定和维护能与外部连接的主机列表是很困难的，而且每个能接收外部直接连接的主机也成为内部网安全周边的

一部分。增加这样的主机对于安全性是有害的，而且同防火墙的目的相违背。

4. Netscape Proxy Server(v2.5)提供名为"反向代理"的机制来支持防火墙外的一个代理服务器转发 HTTP 请求给防火墙内部的一台服务器。代理服务器将 URL 请求直接映射为内部相应服务器的 URL。用户－－代理连接和代理－服务器连接都由 SSL 保护。反向代理在许多方面不能满足我们的要求。首先，防火墙外部的 URL 与其内部的 URL 不同，前者使用代理服务器的名字，而不是相应服务器的名字。例如，在防火墙内部存为标签的 URL 在防火墙外部可能无法工作。相类似的，通过防火墙检索到的主页中的 URL 可能无法在防火墙外部使用。这种方法要求在代理服务器中对每一个要被访问的内部网中的 WEB 服务器进行明确的配置。所以此方法对较大或快速增长的内部网不适用。

三、设计方案

在描述安全 WEB 隧道如何工作之前，让我们先讨论一些设想。

1. 设想

我们的技术要求客户机仅需装备标准的 WEB 浏览器(比如 IE 或 netscape Navigator)，这种 WEB 浏览器应能够使用 HTTPS，HTTPS 变量的安全依靠 SSL 进行认证和加密。SSL(Secure Sockets Layer)是 Internet 浏览器与 IIS 服务器之间的连接标准，这种协议建立在一种公共密钥的基础结构之上，它使用公共密钥证书来对客户端和服务器进行验证。SSL 利用由被信任的证书管理机构发布的公共密钥证书，它们并不需要一个在线的验证服务器。

客户端应是可信的，至少在使用时。这种设想是必不可少的，因为一个恶意的客户能够打印从内部网检索的敏感数据，我们假定客户是可信任的，而且使用的浏览器没有明显的缺陷。

2. WEB 隧道和它的操作

我们系统的核心是一台特殊的服务器。我们称之为 WEB 隧道，这台服务器控制从外部访问内部网的 WEB 页。它作为内部网边界代理处理所有指向内部网的请求。

WEB 隧道工作在防火墙的保护之下，逻辑上作为防火墙的一部分。它允许连接到防火墙内部服务器的 HTTP 连接。防火墙允许 HTTP 和 HTTPS 从外部连接到 WEB 隧道。我们不限制 WEB 隧道必需与防火墙在同一台机器上。

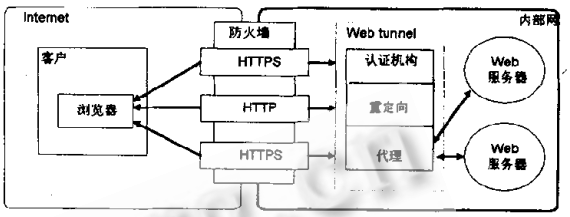

图 1　WEB 隧道结构

如图 1 所示，WEB 隧道由 3 部分组成：认证机构，重定向及代理。

(1)认证机构提供客户到代理的认证。

(2)重定向处理接收到的 HTTP 请求，将 HTTP URLS 映射到 HTTPS URLS，它在 HTTP 重定向的应答中返回客户这些 HTTPS URLS。

(3)当接到 HTTPS URLS 中的一个，客户建立一条到代理的安全连接，用于传输重定向的请求。代理向内部网转发原始的 HTTP 请求，并且在安全连接上返回给客户应答信息。

下面解释这三部分是如何工作的。假定 JOHN 是一名为 ACME 公司工作的可信的职员，当它向从内部网之外访问 ACME 的内部网时，它需要依次进行以下操作：

(1)找到一台连接 INTERNET 的计算机，计算机上装备能够使用 HTTPS 的浏览器。配置浏览器使其信任一个认证权威机构，它为 WEB 隧道登记了一个公钥证书。

(2)JOHN 配置浏览器使其对以 http://*.acme.com/ 开头的 URLS 开头的请求使用 WEB 隧道的重定向部件作为代理。

(3)如果 JOHN 事先没有认证材料，他需要通过与 WEB 隧道的认证机构进行交互来获得。

(4)假定 JOHN 想访问 ACME 防火墙内部一台服务器上的 URL，比如 http://ht.acme.com/foo.htm，JOHN 的浏览器试图通过发送一个非安全的普通 HTTP (http://ht.acme.com/foo.htm)请求给重定向机构，如果重定向接受这个请求，它使用 HTTP 重定向工具进行应答：它告诉浏览器使用新的重定向后的 URL 重新发送这个请求，这个新的 URL 使一个 HTTPS 的 URL，并且使

得浏览器与 WEB 隧道的代理相连接。

(5)当 JOHN 的浏览器请求访问 HTTPS URL 时,浏览器同代理建立一条安全的连接。浏览器通过确信 WEB 隧道的共钥证书来确认代理,代理则通过 JOHN 事先得到的认证材料来确认 JOHN。如果认证成功,则代理重新建立原始的 HTTP URL(http://ht.acme.com/foo.htm),并且作为一个普通的 HTTP 代理转发请求给防火墙。内部的 WEB 服务器使用普通 HTTP 通过安全连接将结果转发给客户端。

注意到 WEB 隧道可以获得请求和应答的明文,而且它们没有再细分成低层的难以理解的包。因此 WEB 隧道可以阅读和理解请求和应答,能够实现强大的访问控制,记录和监听。特别是防火墙的管理员可以通过配置 WEB 隧道来限制从外部访问内部网的某一部分。其实一个 WEB 隧道也就是一个应用级网关。

3.重定向

安全 WEB 隧道技术依靠一种机制利用一个重定向的 URL 把浏览器同 WEB 隧道的代理相连接。该机制可以使用多种方法实现。比较简单的方法是使重定向的 URL 和原始的 URL 基本类似,仅仅把 HTTP 改为 HTTPS;浏览器的配置使用隧道作为重定向 URL 的代理服务器;而且隧道足以假扮内部的 WEB 服务器。我们采用一种较复杂的方法,如下:

(1)重定向 URL 的协议部分为 HTTPS。

(2)重定向 URL 的主机名为 WEB 隧道中代理部件的主机名。

(3)重定向 URL 剩余的部分由原始的 URL 组成。

例如 HTTP 的 URL http://hr.acme.com/foo.htm 转换成 HTTPS 的 URL:https://tunnel.acme.com/hr.acme.com/foo.htm,当代理接收到一个重定向的 URL 的请求时,它将重定向的 URL 恢复成原始的 URL,并将请求转发给原始 URL 指向的服务器。

这种重定向 URL 的构建方法具有诱人的特性,这种 URL 可以通过防火墙传送。想象一下,浏览器通过 WEB 隧道获得了一个 WEB 主页,当浏览器显示该页时,它显示重定向的 URL 作为当前页的 URL。用户可以拷贝这个 URL,并且把它发送给防火墙内部的用户(比如通过 E-MAIL),防火墙内部的用户能够直接使用这个 URL,而不用对其进行任何编辑,使用此 URL 时会涉及到 WEB 隧道,但是代理不用认证用户,因为它能够辨别请求是来自防火墙内部的 IP。

浏览器发送给重定向机构的 HTTP 请求可能会泄露某些敏感信息。WEB 隧道对于用户泄露敏感的信息无能为力,但 WEB 隧道不应助长这种泄露。所以,重定向应该拒绝重定向敏感的 HTTP 请求。我们依靠管理策略来辨别那些敏感请求。注意,发送原始的 HTTP 请求之后,浏览器将只发送 HTTPS 的请求。对于通过 HTTPS 获得的主页中的相对 URL 请求,无须经过重定向部件,也不用进行各种限制,仅仅需要对绝对 URL 进行映射。

4.重写 HTML

当代理部件将 HTML 发送给客户端,它会将 HTML 中的内部 URLS 映射成 HTTPS URL。这种映射节省了某些通信量,也就是不再需要同重定向部件之间的交互。这时仅仅需要对绝对 URL 进行映射,对于相对 URLS 无须进行映射,因为相对 URLS 已经在 HTTPS 的上下文中进行了解释。

假设所有的内部 URL 都按照这种方式映射到 HTTPS URLS,则当浏览器从一个合适的服务器(它们的 URL 不会泄露敏感信息)开始浏览之后,所有从那一点来的 URLS 和内容在流经 Internet 时都是经过加密的。

在我们的重定向方法中,对于一些以"/"开始的相对 URL,必需进行重写。("/"表示当前主机)。当一个 HTML 页包含这样一个 URL 时,代理将它转写成一个合适的绝对的 HTTPS URL。

四、认证方式

在我们的系统里,在防火墙外的用户和用户访问的内部网的 WEB 服务器之间没有直接的认证。替代的方法是在浏览器和 WEB 隧道的代理部件之间进行认证。浏览器认证代理,以防 WEB 页欺骗以及中途攻击,并且向用户显示代理的身份。另一方面,代理需要确认用户的身份,以便决定是否接受请求。也就是说认证是相互的,而且两个方向不近相同。

我们使用 SSL 协议和 X。509 标准证书来认证代理。上面提过,浏览器必需信任签发浏览器证书的证书权威机构。许多浏览器支持证书管理,所以这种信任关系可以手工配置,代理也可以使用一个浏览器默认的证书权威机构签发的证书。

对于用户的认证有许多种方法。一个吸引人的方案是用户使用一件包含公钥和个人证书的硬件。这种硬件可以是一块智能卡或家用计算机。在这种方式下,用户可以依靠公钥和个人证书来获得认证。然而,尽管智能卡和公钥加密技术近来取得很大进展,但它们还不是很

普及。

为了在缺少私钥的情况下获得认证的材料，用户首先要和 WEB 隧道的认证机构进行交互。这种交互在用户试图通过 WEB 隧道访问 URL 时自动进行，而无须提供合法的认证材料。交互在一条 SSL 连接上进行，此时 WEB 隧道已经被用户认证，但用户还没有被认证，这条 SSL 提供信任保证。通过这条连接，用户向 WEB 隧道证明自己的身份。

1. 身份证明可以采用用户名和密码的方式，WEB 隧道可在防火墙内部维护的数据库中进行验证。

2. 用户还可以通过"问题－回答"的方式来验证身份。

作为同认证机构交互的结果，用户获得了一些认证材料，以便向代理提交。因为连接到代理的连接提供了保密性，我们系统的认证材料采用一个短密码的形式。密码提交给代理的方式可以作为一个 cookie。在这种方式下，浏览器可以存储密码，并且在需要的时候调用它，无须用户的干预。每个密码使用一段时间后失效。密码过期后，用户可以再次访问认证机构以获取一个新的密码。

五、安全保证

前几部分描述了 WEB 隧道如何工作，这部分总结它们的安全特性。这些特性的中心就是经过防火墙和 WEB 隧道的 WEB 访问具有和防火墙内部 WEB 访问相同的安全特性。另外，防火墙可以提供 HTTP 级的保护。

具体来说，当一个负责人通过隧道访问内部 WEB 服务时，可以获得以下安全特性：

1. 负责人的认证。WEB 隧道确信请求和应答是和一个可信的负责人之间进行的，所以只有可信的负责人可以访问内部的 WEB 服务。

2. WEB 隧道的认证。负责人确信一个合法的隧道在处理请求和应答。因此负责人确信请求被发送到内部网的对应的服务器中，所以应答是由相应的服务器产生的。

如果一个攻击者窜改了原始 HTTP 的重定向，则负责人可能连接到一个伪造的隧 道。如果负责人错误的向伪造隧道提供了认证材料，则伪造隧道就获得了对内部网的访问权限。所以负责人对于隧道的认证就显得十分重要。尽管 JAVA SCRIPT 代码可以重写负责人看到的地址域，普通浏览器能够提供对隧道的可靠的认证，所以负责人可以对隧道进行必需的认证。

3. 请求和应答的隐秘性。请求和应答能够防止防火墙外部的偷听。唯一的例外是第一个请求（发送给重定向），它可以被偷听者知道。这样，一些内部 URL 会泄露。隧道提供了减少这种泄露的措施，它限制能够重定向的 URL 的个数。所有其他的请求和应答都是加密的。

4. 管理控制。所有通过隧道的信息都是 HTTP 级的，能够在该级被记录和监听。而且一个管理员可以禁止外部对内部 WEB 服务的访问，也是在 HTTP 级的。

5. 在内部网中的缺省保护。在内部网中没有增加额外的安全措施，所以不能防止内部网内部的欺骗。

这些特性都不是绝对的。它们依赖于实现其的部件的安全性，比如浏览器，操作系统，用于认证的编码等等。我们将这些部件看作黑箱，自然有相应厂商和开发商在它们的安全性上进行开发。

六、总结

安全 WEB 隧道允许通过防火墙对内部网的 WEB 资源进行受控访问。它不需要在客户机上安装额外软件，也无须对内部网进行复杂的配置。它提供了认证和信任特性，允许防火墙在 HTTP 级上进行访问控制，记录，监听。因为这些优点，WEB 隧道正逐渐成为 WEB 安全方面的一种十分有用的工具。

参考文献

[1] Martin Abadi, Andrew Birrel , Raymie stata, Edward Wobber"Secure Web Tunneling" Computer networks and isdn systems 1998

[2] Derek Atkins "Internet 网络安全专业参考手册" 机械工业出版社

[3] Adrian Tang & Sophia scoggins "开放式网络和开放式系统互连" 电子工业出版社

（来稿时间：1999 年 6 月）