

Thema:	Portale im Portalverbund: Verarbeitung von Namensräumen und Cookies beim Umschreiben von URLs
Datum:	2004-09-28
Autor:	Rainer Hörbe
Version:	2.0 Ersetzt das Dokument „Namespace- und Cookie-Handling“ vom 29. Jänner 2004

Problemstellung

Ein Portal wird im Portalverbundprotokoll der Version 1.x im Allgemeinen als Gateway im Sinne der HTTP-Spezifikation [RFC 2616] implementiert, was auch als "Reverse Proxy" bezeichnet wird¹. Aus der Sicht des HTTP-Clients ist ein Gateway keine Zwischenstation, sondern der endgültige Kommunikationspartner. Dem entsprechend sind Namensraum und Adressierung auf das Portal bezogen. Ein PV-Portal unterscheidet sich von einem gewöhnlichen Reverse Proxy durch zwei wesentliche Merkmale: Die Funktion als Authentifizierungs- und Autorisierungsproxy einerseits und das Mapping von URLs auf mehrere Anwendungsportale andererseits. URL-Mapping bedeutet, dass der Pfad-Teil des URLs entscheidet, zu welchem Server ein Request weiter geleitet wird. Dadurch entsteht am Reverse Proxy ein gemeinsamer Namensraum der Anwendungen. Diese Funktion ist verantwortlich für ein spezielles Problem bei der Verarbeitung von Cookies, das hier besprochen werden soll.

¹ Die Implementierung als Reverse Proxy in Portalen ist allerdings nur eine Konvention. Eine Implementierung als Forward Proxy über das HTTP-Proxy Protokoll wäre eleganter und würde die Probleme der Namensräume lösen. Allerdings geht das im Falle von Stammportalen nur bei Open-Source Proxies (Squid, Delegated), und nicht für proprietäre Produkte von MS/Novell/Sun etc.

Reverse Proxy und URL-Umschreiben

Da ein Reverse Proxy seine eigene Adresse auf die Zieladresse umschreibt, müssen alle Adress-Referenzen im Datenstrom ebenfalls geändert werden. Adress-Referenzen können in folgenden Teilen eines HTTP-Datenstromes enthalten sein:

- Request
 1. Authentication Header (z.B. Basic Authentication)
 2. Content
 3. Cookies (Domain- und Path-Parameter nur bei RFC 2109-Cookies)
 4. Host Header
- Response
 5. Content
 6. Set-Cookie Header (Domain- und Path-Attribut)
 7. Location Header

Ad 1.) Der Authentifizierungs-Header ist im Kontext des PVP irrelevant.

Ad 2.) Absolute Adressen innerhalb von Anwendungen sollten unbedingt vermieden werden, da ein Umschreiben von Adressen im Content aufwändig und nur bedingt realisierbar ist. (URLs in Javascript, PDF, ..).

Ad 3.) Der HTTP-Header SET-COOKIE muss umgeschrieben werden, wenn die Attribute PATH oder DOMAIN gesetzt sind.

Ad 4.) Der HTTP-Header HOST muss immer entsprechend der Regeln für das URL-Mapping umgeschrieben werden.

Ad 5.) siehe ad 2.)

Ad 6.) explizite Angaben im Domain oder Path-Attribut müssen umgeschrieben werden.

URL-Mapping bei Stammportalen

Im PVP ist spezifiziert, dass sich Anwendungen nicht nur im Host-Namen des Anwendungsportals unterscheiden müssen, sondern auch der Pfad mit der Domäne des Portalbetreibers und der Anwendungskennung qualifiziert ist. Aufgrund dessen kann ein Stammportal Requests den Anwendungsportalen eindeutig zuordnen und URLs entsprechend zuordnen, z.B.:

Der Request am Stammportal

`https://sterz.stlg.gv.at/xyz.gv.at/APP1/servlet/start`

wird umgeschrieben auf

`https://xyz.gv.at/xyz.gv.at/APP1/servlet/start`

Verarbeitung des HOST-Headers

Der Host-Header des Requests ist immer auf den Namen des Hosts umzuschreiben, auf den der Request weiter geleitet wird.

Verarbeitung des LOCATION-Headers

In Responses, die einen Location-Headern setzen (z.B. bei HTTP-302) wird der Hostname des Portals eingesetzt, das den zum Response gehörigen Request erzeugt hat.

Verarbeitung von Cookies

Domain-Attribut

Wird das Domain-Attribut im Set-Cookie Header nicht gesetzt, dann braucht es vom Reverse Proxy nicht umgeschrieben werden. Andernfalls müssen Cookie-Domains bei Responses so umgeschrieben werden, dass sie bei einem darauf folgenden Request einerseits vom Browser an das Portal übergeben werden und andererseits das Stammportal die Domäne wieder korrekt zurücksetzen kann. Dafür sind wiederum die Regeln des URL-Mappings anzuwenden.

Path-Attribut

Das Path-Attribut ist analog zum Domain-Attribut zu verarbeiten.

Lokale Cookies im Stammportal

Wenn im Stammportal ein Cookie erzeugt wird, etwas JSESSIONID zur Verwaltung der Benutzersession, und für den gleichen URL von der Anwendung ein Cookie des gleichen Namens erzeugt wird, müssen die Cookies durch unterschiedliche PATH-Attribut qualifiziert werden, etwa indem für das Stammportal-Cookie explizit PATH=/ gesetzt wird. Alternativ kann ein eindeutiger Name für das Cookie (z.B. XXX.GV.AT-SESSIONID) verwendet werden.

Referenzen

RFC 2109 HTTP State Management Mechanism