Integration bestehender IP-basierter Autorisierung und Abrechnung in Shibboleth-basierte Föderationen

Sebastian Rieger

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG)
Am Fassberg
37075 Göttingen
sebastian.rieger@gwdg.de

Abstract: Insbesondere im wissenschaftlichen Umfeld stellen Verlage und Bibliotheken in den letzten Jahren vermehrt Ihre Zugriffskontrolle für die von Ihnen bereitgestellten Ressourcen von IP-basierten auf föderative Verfahren um. Ausgehend von Entwicklungen im Rahmen des Internet2 in den USA hat sich für föderative Authentifizierungs- und Autorisierungsverfahren im wissenschaftlichen Umfeld das auf der Security Assertion Markup Language (SAML) basierende Shibboleth etabliert. Shibboleth ermöglicht durch den föderativen Ansatz ein Single Sign-On über unterschiedliche Web-Ressourcen innerhalb einer Föderation. Allerdings umfasst weder Shibboleth noch der SAML-Standard explizite Funktionen für die Unterstützung des Accounting bzw. der Abrechnung von Zugriffen. Eine differenzierte Abrechnung ist jedoch vor Allem dann erforderlich, wenn innerhalb einer Föderation unterschiedliche Organisationen existieren (z.B. unterschiedliche Einrichtungen, die die Leistungen einer gemeinsamen Bibliothek in Anspruch nehmen). Die folgenden Abschnitte stellen eine Lösung vor, die im Rahmen der Realisierung einer Shibboleth Föderation für die 80 Institute der Max-Planck-Gesellschaft (MPG-AAI) in Bezug auf die Integration von IP-basierten und föderativen Abrechnungs- und Autorisierungsverfahren erstellt wurde. Durch die vorgestellte Implementierung wird die Integration von Verlagen, die nach wie vor eine IP-basierte Autorisierung durchführen, in die Föderation möglich ohne dabei die differenzierte Abrechnung der einzelnen Institute der Max-Planck-Gesellschaft einzuschränken. Dies ermöglicht eine sanfte Migration hin zu föderativen Authentifizierungs- und Autorisierungsverfahren innerhalb der Max-Planck-Gesellschaft.

1 Zugriffsschutz auf Web-Ressourcen bei Verlagen

In der Vergangenheit wurde der Zugriff auf Web-Ressourcen wissenschaftlicher Verlage insbesondere durch die Prüfung der IP-Adresse des Clients von dem aus der Zugriff erfolgt geschützt [Mike04]. Benutzer, die über eine vom jeweiligen Verlag akzeptierte IP-Adresse verfügten wurden durch diese Adresse gleichermaßen authentifiziert und autorisiert. Um die Authentifizierung und Autorisierung an unterschiedlichen Web-Seiten unabhängig von der IP-Adresse einheitlich zu realisieren, wurden in den letzten Jahren unterschiedliche Verfahren entwickelt. Sie lassen sich in föderative und benutzerzentrierte Verfahren einteilen [Rieg09].

Föderative Verfahren basierend dabei in der Regel auf dem Security Assertion Markup Language (SAML) Standard [SAML]. Eine insbesondere in wissenschaftlichen IT-Infrastrukturen weit verbreitete Implementierung des SAML-Standards bildet Shibboleth [MCHK04]. Wesentlicher Treiber hinter der Einführung von Shibboleth sind und waren dabei auch Bibliotheken und Verlage, die Ihren Nutzern einen Zugriff unabhängig von deren aktueller IP-Adresse erlauben wollten, ohne dabei für jeden Verlag eine separate Anmeldung bzw. Benutzerkonten zu benötigen [vasc]. Die Verwendung separater Benutzerkonten für die Benutzer bei den Verlagen ist nicht zuletzt aufgrund der hohen Fluktuation in wissenschaftlichen Umgebungen nicht realisierbar [RiNe07]. Eine geeignete Lösung bieten dezentrale Authentifizierungsverfahren, wie z.B. die föderative Authentifizierung, bei denen direkt die Benutzerkonten der Heimatinstitute der Anwender verwendet werden können. Obwohl einige Verlage bereits auf föderative Authentifizierungsverfahren umgestellt haben, verwendet die Mehrzahl weiterhin eine IP-Adressbasierte Zugangskontrolle [Mike04]. Dies ist vorrangig in der Komplexität von föderativen im Vergleich zu IP-basierten Verfahren begründet. Um den Benutzern innerhalb einer Föderation auch diese Anbieter bzw. Verlage zugänglich zu machen, wurden unterschiedliche Proxy Lösungen für Bibliotheken (wie z.B. der OCLC EZproxy [EZp]) um föderative Authentifizierung und Autorisierung erweitert. Alle Benutzer des Proxy erhalten hierbei innerhalb der Föderation dessen IP-Adresse als Quell-Adresse beim Zugriff auf die Verlage. Mit dieser einzelnen IP-Adresse ist auf der Seite der Verlage keine differenzierte Autorisierung und Abrechnung, z.B. von unterschiedlichen Instituten, die den Proxy verwenden, möglich. Dieses Paper beschreibt eine Lösung, die für die 80 Institute der Max-Planck-Gesellschaft innerhalb von deren MPG-AAI Föderation [MPAAI] realisiert wurde, um die genannten Einschränkungen zu adressieren, und eine institutsbezogene Autorisierung und Abrechnung zu erlauben. Die Grundlage für den Zugriff auf die Ressourcen liefern hierbei Verträge zwischen den Verlagen und der Max-Planck Digital Library (MPDL). Gemeinsam mit dem Rechenzentrum Garching (RZG) betreibt die Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG) einen Proxy, der die Benutzer der MPG-AAI entsprechend ihrer Institute auf separate IP-Adressen abbildet und so der MPDL eine differenzierte Abrechnung sowie die Erhebung institutsbezogener Nutzungsstatistiken erlaubt.

1.1 IP-basierte Zugriffskontrolle für den Zugriff auf Web-Ressourcen

Anhand der Kontrolle der Quell-IP-Adresse, die der Web-Client (bzw. Web-Browser) des Anwenders verwendet, lässt sich eine einfache Zugriffskontrolle realisieren. Wissenschaftliche Einrichtungen verfügen in der Regel über einen gesonderten IP-Adressbereich (bzw. IP Subnetz), anhand dessen alle Benutzer des Instituts eindeutig identifiziert werden können. Die Autorisierung der Zugriffe sowie deren Abrechnung kann daher auf der Seite der Anbieter anhand der Quell-IP-Adresse durchgeführt werden. Abbildung 1 zeigt ein Beispiel für diese konventionelle Differenzierung von Zugriffen unterschiedlicher Institutionen auf Web-Ressourcen, die derzeit noch häufig von Bibliotheken und Verlagen verwendet wird [Eggl08]. Hierbei greift Benutzer i1.b1, der Angehöriger des Instituts i1 ist, welches das IP-Subnetz 192.168.0/24 verwendet (im Rahmen dieses Papers werden private Internet-Adressen für die Beispiele verwendet), auf eine Web-Ressource, die vom Verlag v1 angeboten wird, zu.

Der Verlag verwendet die Quell-IP-Adresse des HTTP-Requests, um zu entscheiden, ob der Benutzer für den Zugriff autorisiert ist. Üblicherweise werden hierfür die IP-Adressbereiche der aus Sicht des Verlags zugriffsberechtigten Institutionen beim Anbieter in entsprechenden Tabellen hinterlegt. Erfolgt danach ein Zugriff des Benutzers i2.b1, welcher dem Institut i2 angehört, auf Ressourcen des Verlags v1, kann dieser aufgrund der Quell-IP-Adresse 192.168.1.11 dem Institut i2 zugeordnet und so die unterschiedlichen Institute der beiden Benutzer i1.b1 und i2.b1 differenziert werden. So kann der Verlag v2 z.B. anhand der Quell-IP-Adresse Benutzer des Instituts i1 zulassen und Zugriffe von Benutzern des Instituts i2 verweigern.

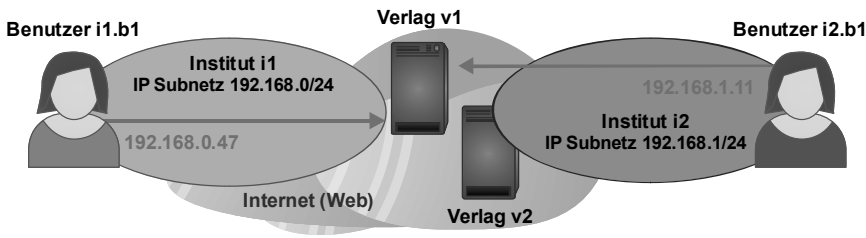

Abbildung 1: Differenzierung der Quell-IP-Adressen beim Zugriff auf Web-Ressourcen.

Während die in Abbildung 1 gezeigte Zugriffskontrolle für die Verlage sehr einfach zu implementieren ist, birgt das Verfahren einige Nachteile. Zum Einen können Quell-IP-Adressen gefälscht bzw. durch zusätzliche Gateways / Proxy-Lösungen etc. manipuliert werden (z.B. IP Spoofing). Durch die Auswertung der Quell-IP-Adresse kann somit keine Authentifizierung gewährleistet werden. Zum Anderen müssen die Benutzer bzw. deren Clients sich innerhalb des Subnetzes ihres jeweiligen Instituts befinden, damit ein Zugriff auf die Verlage erfolgen kann. Häufig werden Lösungen wie Proxy-Server oder VPN, die den Benutzern auch einen mobilen Zugriff auf das Subnetz ihres Instituts erlauben würden, in den Verträgen von den Verlagen ausgeschlossen, sofern keine zusätzlichen Maßnahmen für die Gewährleistung der Authentizität der Benutzer ergriffen werden. Diese Einschränkungen stehen den Anforderungen nach Mobilität innerhalb von wissenschaftlichen IT-Infrastrukturen z.B. durch deren räumliche Verteilung (vgl. virtuelle Organisationen oder weltweit kooperierende Forschungsprojekte) entgegen [RiNe07]. Ein zusätzliches Problem entsteht, wenn unterschiedliche Institutionen ein gemeinsames Subnetz verwenden (z.B. bedingt durch eine Kooperation in Bezug auf die Internet-Anbindung zwischen Universitäten und Forschungsinstituten eines Standorts).

1.2 Föderative Authentifizierung und Autorisierung

Im vorherigen Abschnitt wurden unterschiedliche Nachteile von IP-basierten Zugriffskontrollverfahren beschrieben. Um diese Probleme zu adressieren wurden in den letzten Jahren dezentrale Authentifizierungs- und Autorisierungsmechanismen eingeführt. Hierbei kann zwischen föderativem und benutzerzentriertem Identity Management unterschieden werden [Rieg09]. Einige Verlage haben, wie im Abschnitt 1 beschrieben, ihre Zugriffskontrolle bereits auf föderative Authentifizierungsverfahren (z.B. das SAML-basierte Shibboleth) umgestellt.

Um föderative Authentifizierungsverfahren zu unterstützen, implementieren die Verlage einen sog. Service Provider (SP) [Morg04] und schließen sich damit einer oder mehreren Föderationen an, die die Benutzer bzw. Kunden der Verlage enthalten. Auf der anderen Seite betreiben die Institute hierfür ihrerseits sog. Identity Provider (IdP) [Morg04], die ihren Benutzern den Zugriff auf SPs innerhalb der Föderation erlauben. Beispielsweise betreibt der DFN-Verein eine Föderation (DFN-AAI), der sich bereits einige Verlage angeschlossen haben [DV], und an die auch die MPG-AAI angebunden ist.

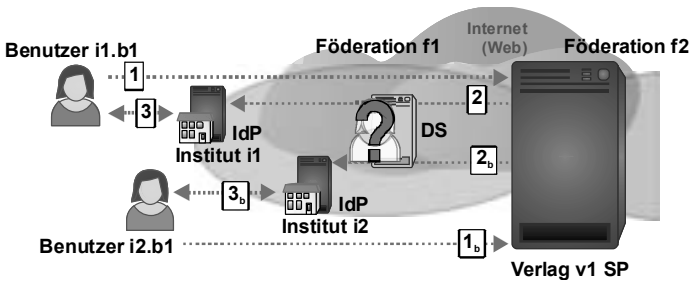

Abbildung 2: Zugriff auf Verlage innerhalb einer Föderation aus unterschiedlichen Instituten.

Abbildung 2 zeigt den Verlag v1, der sich den Föderationen f1 und f2 als Service Provider (SP) angeschlossen hat. In dem oben gezeigten Beispiel enthält die Föderation f1 zusätzlich die Institute i1 und i2 bzw. deren Benutzer (z.B. i1.b1 und i2.b1). Greift der Benutzer i1.b1 auf vom Verlag v1 angebotene Ressourcen zu (1), so wird er von dessen SP zum Discovery Service (DS) der Föderation f1 umgeleitet (2), der die Lokalisierung des Heimatinstituts übernimmt. Nachdem der Benutzer sein Heimatinstitut am DS ausgewählt hat, leitet ihn dieser an dessen IdP weiter [Morg04]. Anschließend erfolgt die Anmeldung des Benutzers am IdP (3). Nach erfolgreicher Authentifizierung erstellt der IdP ein digital signiertes Ticket (sog. Assertion), mit dem er den Benutzer an den SP des Verlags v1 zu der ursprünglich ausgewählten Ressource weiterleitet. Zusammen mit der Assertion kann der IdP dem SP hierbei Attribute übermitteln, die für die Autorisierung verwendet werden. Beispielsweise kann eines dieser Attribute verwendet werden, um einen eindeutigen Bezeichner für das Heimatinstitut des Benutzers zu übermitteln. Der SP ist dann anhand dieses Attributs in der Lage die Institute zu differenzieren und z.B. spezielle Ressourcen nur einem bestimmten Institut anzubieten. Durch die Verwendung föderativer Authentifizierungs- und Autorisierungsverfahren kann der Zugriff auf die Verlage unabhängig von der Quell-IP-Adresse des Web-Clients erfolgen, den der Benutzer aktuell verwendet. Damit adressieren föderative Verfahren ein zentrales Problem der in Abschnitt 1.1 dargestellten IP-basierten Zugriffskontrollverfahren, ohne zusätzliche Benutzerkonten oder deren Synchronisation über die Verlage hinweg zu erfordern. Durch den SAML-Standard werden zusätzlich unterschiedliche Implementierungen bzw. Software-Lösungen auf der Seite der Verlage ermöglicht. Allerdings basiert die Zugriffskontrolle bei der Mehrzahl der Verlage, wie im vorherigen Abschnitt geschildert, noch auf der Auswertung der Quell-IP-Adresse. Eine einheitliche föderative Authentifizierung und Autorisierung, unabhängig von der Quell-IP-Adresse der Benutzer, kann somit derzeit für wissenschaftliche IT-Infrastrukturen nicht realisiert werden. Ein weiteres Problem bildet die fehlende Standardisierung von Attributwerten in Föderationen.

Während z.B. in der eduPerson sowie der dfnEduPerson [DEP] feste Schemata für die Definition der Attribute existieren, können IdP und SP unabhängig davon unterschiedliche Attributnamen und insbesondere Attributwerte z.B. für die Differenzierung unterschiedlicher Institute der Benutzer verwenden.

1.3 Accounting

Sowohl Shibboleth als auch der zugrundeliegende SAML-Standard wurden für die Vereinheitlichung der Authentifizierung und Autorisierung nicht jedoch der Abbrechnung (Accounting) entwickelt. Die fehlende Standardisierung von Attributen und Attributwerten für das Accounting führt auf der einen Seite zu Problemen bei den Verlagen, wenn die IdPs von deren Kunden jeweils unterschiedliche Attribute bzw. Attributwerte für die Abrechnung an die Verlage übertragen. Auf der anderen Seite ist es auch für die Institute aufwändig für einzelne Verlage unterschiedliche Attribute für die Abrechnung (z.B. eine eindeutige Kennzeichnung des Instituts) zu konfigurieren und zu verwenden. Die geschilderten Nachteile gelten insbesondere für Bibliotheken, die Nutzern unterschiedlicher Einrichtungen (z.B. bedingt durch die Angliederung an Universitäten und Forschungsinstituten) Zugang zu Diensten bzw. Verlagen in einer Föderation anbieten wollen. Hierbei müssen die Bibliotheken die Nutzung durch die jeweilige Einrichtung getrennt abrechnen. Häufig werden beispielsweise Nutzungsstatistiken erstellt, die der Bibliothek neben statistischen Analysen auch ermöglichen, Zugriffe auf Web-Ressourcen für einzelne Einrichtungen gesondert in Rechnung zu stellen. Zusätzlich können Vorgaben aus den Verträgen zwischen Bibliotheken und Verlagen für die getrennte Erfassung der Zugriffe einzelner Einrichtungen existieren.

Um differenzierte Nutzungsstatistiken für die angebundenen Einrichtungen zu erstellen, sind die Bibliotheken dabei auf die Übermittlung von Zugriffszahlen durch die Verlage angewiesen. Die Bibliotheken können dann als Abrechnungsstelle zwischen Verlagen und Benutzern agieren. Da für föderative Authentifizierungsverfahren keine standardisierten Accounting-Verfahren existieren, verwenden einige Verlage proprietäre Lösungen für die Übermittlung der Zugriffszahlen an die Bibliotheken. Andere verwenden weiterhin eine IP-basierte Abrechnung, die anhand der Quell-IP-Adresse des Benutzers ermittelt werden. Diese sind jedoch erneut ungeeignet, sobald die Benutzer z.B. mobil aus unterschiedlichen IP-Subnetzen Zugriff auf die Web-Ressourcen der Verlage erhalten sollen. Bibliotheken und angebundene Einrichtungen fordern daher eine benutzerbezogene Abrechnung unabhängig von der aktuellen IP-Adresse des Web-Clients.

2 Verwendung von Web-Proxy Lösungen in Föderationen

Wie im Abschnitt 1.2 erläutert, wird die IP-basierte Zugriffskontrolle nicht zuletzt aufgrund der geringeren Komplexität in Bezug auf die Implementierung nach wie vor von einer Vielzahl von Verlagen verwendet. Um Benutzern innerhalb einer Föderation Zugriff auf diese Verlage unabhängig von deren aktueller IP-Adresse zu ermöglichen, wurden verschiedene Proxy-Lösungen entwickelt. Innerhalb der Max-Planck-Gesellschaft wird hierfür der OCLC EZproxy [EZp] verwendet.

Die Abbildung 3 zeigt die Verwendung des Proxy-Servers für Verlage, deren Zugriffskontrolle nach wie vor auf der Quell-IP-Adresse des zugreifenden Web-Clients basiert, innerhalb einer Föderation.

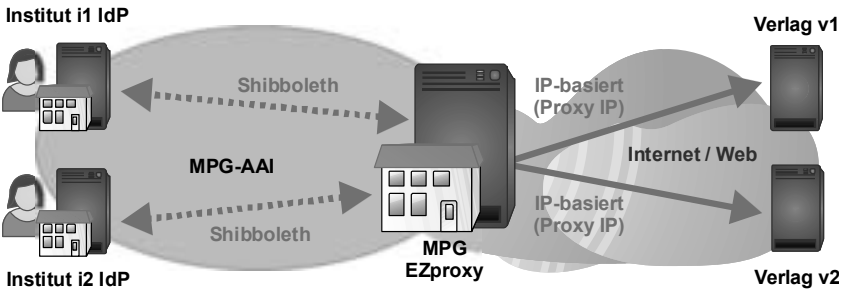

Abbildung 3: Einsatz eines Proxy-Servers für den Zugriff auf Verlage in einer Föderation

Die Abbildung zeigt einen Ausschnitt der MPG-AAI Föderation der Max-Planck Gesellschaft. In dem gezeigten Beispiel betreiben die Institute i1 und i2 für die Authentifizierung und Autorisierung Ihrer Benutzer einen Shibboleth IdP innerhalb der Föderation. Die Benutzer können sich an dem Web-Proxy „MPG EZproxy“ über die IdPs ihrer Institute authentifizieren und nach erfolgreicher Autorisierung auf die Ressourcen der Verlage v1 und v2 zugreifen. Der EZproxy bildet hierbei die Funktion eines Reverse Proxy für die Web-Ressourcen der Verlage. Benutzer können z.B. über die Adresse http://verlag-v1.proxy.aai.mpg.de auf den Verlag v1 zugreifen. Hierbei leitet der Proxy die Anfrage unter Verwendung seiner eigenen IP-Adresse als HTTP Request an den Verlag weiter. Der Verlag sendet anschließend die HTTP Response an den Proxy, der diese zurück an den Benutzer leitet. Da für die Autorisierung und das Accounting auf der Seite der Verlage die Quell-IP-Adresse des Proxy ausgewertet wird, können die Benutzer unabhängig von Ihrer aktuellen IP-Adresse Zugriff auf die bereitgestellten Ressourcen nehmen. Außerdem können die Benutzer alle Dienste der Föderation, z.B. Verlage die bereits eine föderative Authentifizierung unterstützen sowie den EZproxy, ohne separate Anmeldung nutzen. Durch die Authentifizierung und Autorisierung am EZproxy erfüllt dieses Single Sign-On auch die Anforderungen der Verlage in Bezug auf den Schutz der Ressourcen. Häufig schließt dies auch zusätzliche Anforderungen, wie z.B. maximale Gültigkeitszeiträume für Benutzer-Accounts bzw. eine zeitnahe Sperrung von Benutzer-Accounts, mit ein, die innerhalb der Policy der Föderation für alle Teilnehmer verbindlich vorgegeben werden. Die skizzierte Lösung erlaubt eine sanfte Migration zu föderativen Authentifizierungsverfahren, ohne den Benutzern den Zugriff auf Verlage, deren Zugriffskontrolle noch auf der Auswertung der Quell-IP-Adresse basiert, außerhalb des IP-Subnetzes ihres Instituts zu verweigern.

2.1 Realisierung einer mandantenfähigen Web-Proxy-Lösung für Föderationen

Während die im vorherigen Abschnitt beschriebenen Reverse-Proxy Server das Problem der Integration von Verlagen mit IP-basierter Zugriffskontrolle in Föderationen lösen, erlauben sie jedoch keine differenzierte Abrechnung. Alle Benutzer erhalten beim Zugriff auf die an den Proxy angebundenen Verlage die gleiche IP-Adresse.

Auf der Seite der Verlage kann nur die IP-Adresse des Proxy Servers für die Autorisierung und Abrechnung verwendet werden. Existieren innerhalb einer Föderation unterschiedliche Einrichtungen, die den Proxy nutzen, wie z.B. unterschiedliche Forschungseinrichtungen und Universitäten innerhalb der DFN-AAI, so können auf der Seite der Verlage keine einrichtungsbezogenen Zugriffe erlaubt oder abgerechnet werden. Dies ist, wie bereits in Abschnitt 1.3 erläutert, inakzeptabel für Bibliotheken, die die Verrechnung ihrer Leistungen anhand von Nutzungsstatistiken einzelner angebundener Einrichtungen durchführen. Eine mögliche Lösung für die differenzierte Abrechnung einzelner Nutzer der im vorherigen Abschnitt beschriebenen Web-Proxy Server innerhalb einer Föderation könnte die Übermittlung der vom IdP an den Web-Proxy als SP übertragenen Attribute sein. Beispielsweise könnte ein Attribut für die Institutszugehörigkeit auf eine HTTP Header Variable abgebildet werden, die der Proxy in den Requests an die Verlage überträgt. Allerdings würde dieses Verfahren erneut eine Standardisierung der übermittelten Variablen über unterschiedliche Verlagen und Institute hinweg erfordern. Ebenfalls wären Anpassungen auf der Seite der Verlage für die Auswertung der Variablen erforderlich. Um unabhängig von der aktuellen IP-Adresse des Web-Clients der Benutzer einen Zugriff auf alle innerhalb der Max-Planck-Gesellschaft verwendeten Verlage zu erlauben, wurde die in Abbildung 3 vorgestellte Reverse Proxy-Lösung, wie in Abbildung 4 dargestellt, um einen zusätzlichen Forward Proxy erweitert.

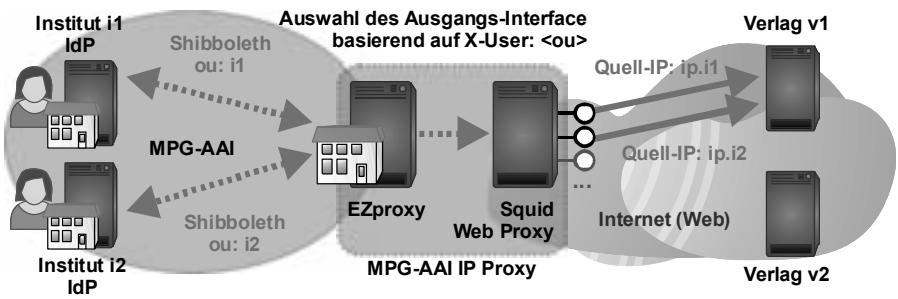

Abbildung 4: Kombination von Reverse und Forward Proxy für eine differenzierte Abrechnung und Autorisierung

Hierfür wurde ein zusätzlicher Squid Web Proxy [Squid] eingerichtet, über den der EZproxy seine ausgehenden HTTP Request sendet. Greift ein Benutzer auf eine Ressource des Verlags v1 (z.B. http://verlag-v1.proxy.aai.mpg.de) zu, so muss er sich erneut am EZproxy mittels Shibboleth Login am IdP seines Heimatinstituts authentifizieren. Dabei übermittelt jedes Institut eine eindeutige Institutskennung (Domain des Instituts z.B. institut-i1.mpg.de) in Form des OU Attributs. Der EZproxy kann dieses Attribut für die Autorisierung einzelner Institute an bestimmten Ressourcen verwenden. Für den Zugriff auf den Verlag sendet der EZproxy die Anfrage nach erfolgreicher Authentifizierung und Autorisierung an den Squid Proxy. Der EZproxy wurde hierbei erweitert, so dass der Header dieses HTTP Requests einen Parameter X-User mit der zuvor vom IdP erhaltenen Institutskennung (z.B. X-User: institut-i1.mpg.de) beinhaltet. Am Squid Proxy wurden für alle angeschlossenen Institute virtuelle Ausgangs-Interfaces mit separaten IP-Adressen eingerichtet.

Anhand des in den eingehenden HTTP Requests empfangenen X-User Headers wählt der Squid Proxy für jedes Institut ein individuelles Ausgangs-Interface und damit die Quell-IP-Adresse (tcp_outgoing_address) für den anschließenden HTTP Request an die Verlage. Verlage, die ihre Zugriffskontrolle und Abrechnung nach wie vor anhand der eingehenden IP-Adresse durchführen können so weiterhin, neben anderen die bereits föderative Verfahren einsetzen, verwendet werden. Die Max-Planck Digital Library erhält zusätzlich weiterhin institutsbasierte Nutzungsstatistiken von den Verlagen. Trotzdem können die die Benutzer unabhängig von deren aktueller IP-Adresse auch außerhalb des Subnetzes ihres Heimatinstituts alle für Ihre Forschung relevanten Verlage verwenden. Durch die Shibboleth-basierte Authentifizierung am EZproxy werden zusätzlich alle Sicherheits-anforderungen der Verlage erfüllt.

2.2 Fehlertoleranz und Lastverteilung über mehrere Standorte

Um einen ausfallsicheren und performanten Dienst der im vorherigen Abschnitt vorgestellten zentralen Proxy-Lösung für alle 80 Max-Planck-Institute zu realisieren, wurden mehrere „MPG-AAI IP Proxy“ Instanzen realisiert. Diese wurden über zwei Standorte, an denen Rechenzentren der Max-Planck-Gesellschaft existieren, verteilt. Ein Standort bildet die Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG). Den zweiten bildet das Rechenzentrum Garching (RZG). Beide Rechenzentren betreiben jeweils zwei „MPG-AAI IP Proxy“ Instanzen. Für die Realisierung der Ausfallsicherheit über beide Standorte hinweg wurde ein nginx-basierter [nginx] Load Balancer (LB) als zusätzlicher Reverse Proxy vor den Instanzen der Rechenzentren installiert. Der nginx Server erhält dabei alle Zugriffe der Benutzer auf proxy.aai.mpg.de. Sobald eine „MPG-AAI IP Proxy“ Instanz nicht in der Lage ist, die eingehenden Requests der Benutzer zu beantworten, leitet der nginx Server diese an eine andere Instanz um. Zusätzlich werden die Requests über die beiden Instanzen am jeweiligen Standort verteilt und so eine Lastverteilung erzielt. Durch die Anmeldung am IdP können bei einem Ausfall einer Instanz alle Anfragen von zuvor angemeldeten Benutzern auf eine andere Instanz umgeleitet werden, ohne ein erneutes Login der Benutzer zu erfordern.

Da eine IP-Adresse nur einmalig im jeweiligen Subnetz vergeben werden kann, besitzen die teilnehmenden Institute an den Standorten jeweils zwei IP-Adressen, wie in Abbildung 5 gezeigt. Ohne diese zusätzliche IP-Adresse wäre an einem einzelnen Standort keine Lastverteilung realisierbar. Die Anfragen eines Instituts würden alle über die gleiche Ausgangs-IP-Adresse eines einzelnen Squid Proxy laufen. Greift beispielsweise ein Benutzer des Instituts i1 auf verlag-v1.proxy.aai.mpg.de zu, so wird er z.B. an die Instanz „MPG-AAI IP Proxy Göttingen-1“ verwiesen, und erhält im Beispiel die Ausgangs-IP-Adresse 172.16.0.20. Ein Benutzer desselben Instituts der danach auf verlag-v1.proxy.aai.mpg.de zugreift, wird beispielsweise an „MPG-AAI IP Proxy Göttingen-2“ geleitet, und erhält die Ausgangs-IP-Adresse 172.16.0.21. Fällt die Instanz „MPG-AAI IP Proxy Göttingen-1“ aus, so kann der erste Benutzer direkt auf die Instanz „MPG-AAI IP Proxy Göttingen-2“ umgeleitet werden. Der EZproxy leitet ihn dabei, wie in Abbildung 4 gezeigt, an den IdP seines Heimatinstituts um, an dem er bereits eine Sitzung aufgebaut hat.

Daher muss sich der Benutzer nicht erneut anmelden und kann trotz des Ausfalls der ersten Instanz weiter mit den vom Verlag v1 bereitgestellten Ressourcen arbeiten. Durch den nginx Server werden so sowohl Ausfälle eines Standorts als auch Wartungsarbeiten innerhalb eines Rechenzentrums adressiert.

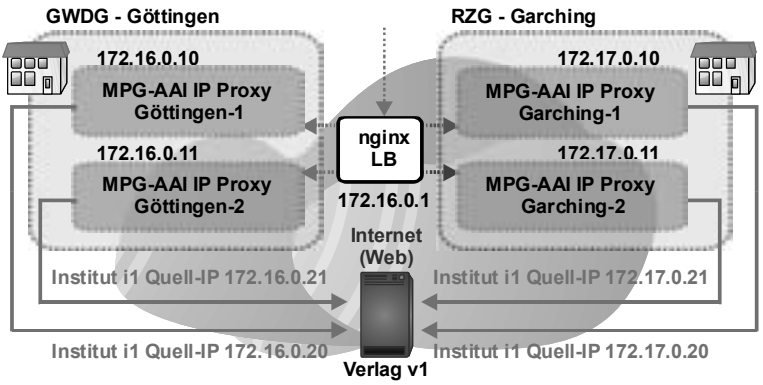

Abbildung 5: Lastverteilung und Ausfallsicherheit durch mehrere "MPG IP Proxy" Instanzen

3 Fazit und Ausblick

Die in diesem Paper vorgestellte Lösung erlaubt eine Integration von föderativen und IP-basierten Autorisierungs- und Abrechnungsverfahren innerhalb der MPG-AAI Föderation der Max-Planck-Gesellschaft. Dadurch wird eine sanfte Migration zu föderativen bzw. SAML-basierten Zugriffskontrollverfahren über alle von den Max-Planck-Instituten benötigten Verlagen ermöglicht. Für die Benutzer wird so, unabhängig von der IP-Adresse des verwendeten Web-Clients, ein Zugriff auf die Ressourcen der Verlage realisiert. Die Benutzer erhalten auf diese Weise ein Single Sign-On über IP- und bereits föderativ autorisierende Verlage. Für die Verlage und die Max-Planck Digital-Library als Bibliotheksdienstleister der Max-Planck-Gesellschaft wird darüber hinaus weiterhin eine differenzierte Abrechnung und Autorisierung der einzelnen Institute gewährleistet. Durch die im Abschnitt 2.2 vorgestellten verteilten Instanzen an den Standorten Garching und Göttingen, werden sowohl Performanz-Engpässe als auch Ausfälle minimiert, und so das Risiko des Proxy als zentrale Fehlerquelle reduziert. Ein Nachteil der Lösung besteht allerdings in dem Verbrauch von IP-Adressen für die teilnehmenden Institute. Neben diesem Nachteil erfordert auch die derzeit fehlende Anpassung der vom Proxy ausgelieferten Inhalte (z.B. werden Links in RSS Feeds nicht automatisch auf das Format http://<verlag>.ezproxy.aai.mpg.de umgesetzt), sowie die Terminierung von HTTPS Sitzungen am Proxy nach wie vor langfristig eine vollständige Migration hin zu vollständig föderativen Autorisierungs- und Abrechnungsverfahren. Aus diesem Grund arbeitet die Max-Planck-Gesellschaft zusammen mit anderen Forschungseinrichtungen derzeit an einer Erweiterung der bestehenden Standards für föderative Authentifizierung und Autorisierung (basierend auf SAML und insbesondere Shibboleth) um Accounting Attribute. Ein Vorschlag für ein geeignetes Accounting-Attribut eduPersonUsageSubset wurde bereits erstellt [EPUS].

Dieser soll nun gemeinsam mit den Entwicklern von Shibboleth sowie mit der Directory Working Group des Internet2 Middleware Architecture Committee for Education (MACE-Dir) im Hinblick auf eine mögliche Integration in die eduPerson Spezifikation diskutiert werden. Neben diesem Vorschlag existiert z.B. mit dem dfnEduPersonCostCenter [DEP] Attribut der dfnEduPerson des DFN-Vereins ein weiterer Lösungsansatz für die differenzierte Abrechnung unterschiedlicher Einrichtungen innerhalb einer Shibboleth-Föderation. Welche Erweiterung zukünftig für das Accounting in Shibboleth-basierten Föderationen verwendet werden sollte, hängt vorrangig von der Akzeptanz durch die Verlage ab. Unabhängig von der konkreten Realisierung des Accountings ist jedoch vom momentanen Standpunkt nicht absehbar ab wann alle Verlage ihre IP-basierte Zugriffskontrolle abgelöst haben. Gemeinsam mit dem bereits existierenden IdP Proxy [Rieg09] der Max-Planck-Gesellschaft ermöglicht der MPG-AAI IP Proxy bis dahin sowohl die Integration von Instituten als auch von Verlagen, die bislang noch nicht in der Lage sind Shibboleth bzw. SAML zu unterstützen.

Literaturverzeichnis

[Eggl08] Eggleston, H.: Introduction to Electronic Resources and Remote Access Issues, http://www.escholarship.org/uc/item/0hc172sp, abgerufen am 14.1.2010.

[DEP] DFN-AAI Technische und organisatorische Voraussetzungen – Attribute für den Bereich E-Learning,https://www.aai.dfn.de/fileadmin/documents/attributes/200811/DFN-AAI_E-Learning-Attribute_V.1.0.pdf, abgerufen am: 14.1.2010.

[DV] DFN-AAI Einfacher Zugang zu geschützten Ressourcen – Service-Provider, https:// www.aai.dfn.de/verzeichnis/service-provider/, abgerufen am 14.1.2010.

[EPUS] Egger, M.; Palzenberger, M.; Rieger, S.; Schier, H.: eduPersonIUsageSubset https://idp.rzg.mpg.de/mediawiki/images/2/21/Discrimination-Attribute.doc, abgerufen am 24.3.2010.

[EZp] OCLC EZproxy authentication and access software, http:// www.oclc.org/ezproxy/, abgerufen am 14.1.2010.

[Mike04] Mikesell, B. L.: Anything, Anytime, Anywhere: Proxy Servers, Shibboleth, and the Dream of the Digital Library. In: (Mahoney, P. B., Hrsg.): Proceedings of The Eleventh Off-Campus Library Services Conference, The Haworth Information Press 2004; S. 315-326.

[Morg04] Morgan, R. L.; Cantor, S.; Hoehn, W.; Klingenstein, K.: Federated Security: The Shibboleth Approach, EDUCAUSE Quarterly, Vol. 27, 2004, S. 12-17.

[MPAAI] MPG: MPG-AAI, https://aai.mpg.de, abgerufen am: 14.1.2010.

[nginx] HTTP reverse proxy server, http://nginx.org/en/, abgerufen am 14.1.2010.

[RiNe07] Rieger, S.; Neumair, B.: Towards usable and reasonable Identity Management in heterogeneous IT infrastructures. In: Proceedings of the 10th IFIP/IEEE International Conference on Integrated Network Management, 2007, S. 560-574.

[Rieg09] Rieger, S.: Benutzerzentrierte Lokalisierung für den Einsatz in Shibboleth-basierten Föderationen. In (Müller, P.; Neumair, B.; Dreo Rodosek, G., Hrsg.): Proc. 2. DFN-Forum Kommunikationstechnologien, München 2009. Gesellschaft für Informatik , Bonn, 2009; S. 13-22.

[SAML] OASIS: Security Services (SAML) TC, http://www.oasis-open.org/committees/tc_home .php?wg_abbrev=security, abgerufen am: 14.1.2010.

[Squid] Squid Optimising Web Delivery, http://www.squid-cache.org/, abgerufen am 14.1.2010.

[vasc] Verteilte Authentifizierung, Autorisierung und Rechteverwaltung (AAR), http://aar. vascoda.de, abgerufen am: 14.1.2010.