Ondersteuning van library walk-ins in SURFconext

Utrecht, december 2017
Versienummer: 1.0

Colofon

Ondersteuning van library walk-ins in SURFconext
P3 – Betrouwbare en Veilige Omgeving – Project Trust & Identity
Work Package 1 – Next Generation Trust & Identity

SURF
Postbus 19035
NL-3501 DA Utrecht
T +31 88 787 30 00

info@surf.nl
www.surf.nl

Auteurs
Bas Zoetekouw (SURFnet)

Reviewers
Pieter van der Meulen (SURFnet), Arnout Terpstra (SURFnet), Femke Morsch (SURFnet), Pim Slot (SURFmarket)

Projectleider
Michiel Schok (SURFnet)

december 2017

Inhoudsopgave

Samenvatting

Nu wetenschappelijke tijdschriften vrijwel exclusief online verschijnen en uitgevers de toegang tot deze tijdschriften steeds vaker afschermen met federatieve logins, hebben bibliotheekbezoekers zonder instellingsaccount geen toegang meer tot wetenschappelijk tijdschriften. Het faciliteren van zulke gebruikers behoort echter wel tot de taak van de universitaire bibliotheken.

In dit document beschrijven we een mogelijk oplossingsrichting voor de kortere termijn. In deze oplossing krijgen bibliotheekbezoekers toegang via een anonieme identity provider, die echter alleen beschikbaar is vanaf de IP-adressen van de bibliotheken.

Voor een dergelijke oplossing ingericht kan worden, moet er echter wel overeenstemming zijn tussen alle partijen (uitgevers, instellingen en bibliotheken) dat ene dergelijke oplossing de problemen rondom bibliotheekbezoekers oplost.

1. Introductie

Van oudsher hebben universiteitsbibliotheken de taak om, naast de populatie van de eigen instelling, ook personen van buiten de instelling te bedienen. Dat gaat dan over het uitlenen van boeken, maar expliciet ook de toegang tot wetenschappelijke tijdschriften. Tot enkele jaren geleden was dat geen probleem: de tijdschriften waren grotendeels fysiek aanwezig in de bibliotheken, en konden dus net als de normale collectie worden ingezien en gekopieerd, ook door personen die niet aan de universiteit verbonden waren.

Nu de wetenschappelijke tijdschriften vrijwel exclusief online verschijnen, begint hier een probleem te ontstaan. De initiële oplossing die voor autorisatie van de content wordt gebruikt, is meestal een afscherming op basis van IP-adres, waarmee typisch de hele campus van een instelling (inclusief de bibliotheken) toegang verkrijgt tot wetenschappelijk journals. In dit scenario hebben bibliotheekbezoekers nog steeds op eenvoudige wijze toegang tot de content via een lokale PC in de bibliotheek.

IP-gebaseerde toegang heeft echter een aantal nadelen. Zo is het bijhouden van de IP-reeksen administratief lastig - SURFmarket moet wijzigingen in de reeksen van een instelling afstemmen met alle individuele uitgevers. Verder zijn er voor diverse usecases (thuiswerkende medewerkers, etc.) aparte technisch voorzieningen nodig, zoals VPN's, remote-desktoptoegang of EZproxy-software.

Om deze redenen zien we de afgelopen jaren steeds meer uitgevers overstappen op federatieve login via SURFconext. Voor reguliere instellingsgebruikers is dit een vooruitgang: zij hebben geen aparte login meer nodig bij portals van uitgevers en kunnen ook vanaf locaties buiten hun instelling tijdschriften raadplegen.

Voor de externe bezoekers van bibliotheken ontstaat nu echter een probleem. Zij hebben geen instellingsaccount, en hebben dus geen toegang meer tot wetenschappelijke tijdschriften die via SURFconext ontsloten zijn; ook niet als ze zich fysiek in de bibliotheek bevinden.

De universiteitsbibliotheken zijn op basis van (onder andere) dit argument terughoudend bij de overgang van IP-afscherming naar federatieve logins; zie bijvoorbeeld paragraaf 3.5.5 van AARC-deliverable DJRA1.1:

> *The UKB, however, identified a number of potential stumbling blocks to the implementation of federated access. These are, for example: [...] So called "walk-by users*", such as citizen scientists, are not able to access academic content. With IP-based access, they are able to access content as long as they reside at the campus*.
>
> *: bij SURFnet gebruiken we hiervoor de term walk-in

Het is daarom opportuun om dit probleem binnen de context van SURFconext op te pakken. In dit document leggen we uit hoe we vanuit SURFnet Trust & Identity deze oplossing voor ons zien.

2. Technische oplossingen

Allereerst moeten we constateren dat het model van universiteitsbibliotheken om niet-instellingsgebruikers alleen te bedienen als ze zich fysiek in de bibliotheek bevinden, niet erg toekomstvast is. Het zou uiterst vreemd zijn dat in een wereld waar alles online kan, iemand om een wetenschappelijk tijdschrift te raadplegen, fysiek een universiteitscampus moet bezoeken. Het is echter vooralsnog onduidelijk hoe de universiteitsbibliotheken dit bredere probleem willen oplossen; wellicht worden alle tijdschriften in de toekomst Open Access, of wellicht dat in de (iets minder verre) toekomst SURF dit soort usecases door middel van een levenslange 'EduID' of door het beschikbaar maken van inlogmiddelen van de overheid (zoals DigiD) binnen SURFconext zou kunnen faciliteren. Voor nu is dit echter buiten scope.

In de tussentijd ligt er echter nog wel een probleem: hoe kunnen we de groep van externe bibliotheekgebruikers in een federatieve wereld toch faciliteren? Een voor de hand liggende oplossing is om een gast-IdP in SURFconext te introduceren die toegang verleent op basis van het IP-adres van de gebruiker.

Hoewel zo'n IP-gebaseerde IdP natuurlijk niet de hele problematiek oplost, en nog steeds een administratie van IP-adressen vereist, biedt hij wel aanzienlijke voordelen boven klassieke IP-based access. Er is één overzichtelijke manier van inloggen: via SURFconext, en uitgevers hoeven geen aparte IP-afscherming meer te onderhouden. Bovendien wordt de administratie ook voor SURF en de instellingen veel eenvoudiger: SURFmarket hoeft niet langer een wijziging van IP-reeksen bij alle uitgevers door te voeren, maar kan op een centrale plek de reeksen administreren. Je zou zelfs het beheer van de IP-reeksen als self-service bij de universiteiten in beheer kunnen geven, of de administratie direct kunnen koppelen aan de IP-registratie van SURFinternet.

3. Implementatie

Bovenstaande oplossing is binnen het EU-project AARC verder uitgewerkt, en daar ook technisch geïmplementeerd in Shibboleth. Daarnaast is er binnen SURFnet een PoC-implementatie in SimpleSAMLphp gebouwd.

Binnen AARC is een uitgebreide pilot uitgevoerd voor library walk-ins. Deze pilot, inclusief usecases en user journeys, staat beschreven op de AARC wiki. Als onderdeel van de pilot is een IP-gebaseerde IdP geïmplementeerd op basis van Shibboleth. Deze omvat naast de IdP ook een self-service managementportal, waar instellingen zelf hun IP-reeksen kunnen instellen. De implementatie wordt ook op de AARC wiki beschreven.

De implementatie die binnen SURFnet is gemaakt, is op basis van SimpleSAMLphp. Deze implementeert slechts een IdP, en niet een self-service management portal. De IP-reeksen moeten hier door de beheerder van de IdP worden bijgehouden. De code is beschikbaar op GitHub.

4. Integratie in SURFconext

Beide implementaties van de IP-gebaseerde IdP zijn technisch eenvoudig te implementeren in SURFconext. De IdP zelf kan zonder problemen gekoppeld worden aan SURFconext en worden geactiveerd voor de uitgevers die deze loginmogelijkheid toestaan. De self-serviceportal zal als SP aan SURFconext moeten worden gekoppeld, en beschikbaar moeten worden gemaakt voor een groep beheerders, bijvoorbeeld door middel van een SAB-rol.

Qua policy ligt het maken van de aansluiting echter wel lastiger. SURFconext is in het algemeen terughoudend met het aansluiten van gast-IdPs, en deze specifieke gast-IdP geeft ook nog eens identiteiten vrij die niet direct te herleiden zijn tot een persoon. Dat past niet in de huidige policy van SURFconext, en om dit te kunnen invoeren moeten dus aparte afspraken gemaakt worden, zowel met uitgevers die deze gebruikers moeten toelaten als met de instellingen die op deze manier toegang willen gaan verlenen.

Het lijkt daarom alleen zinvol om hier daadwerkelijk mee aan de slag te gaan als er een duidelijke businesscase is, en er zich concrete partijen aandienen die deze route willen gaan gebruiken. Op dit moment lijken de partijen niet erg enthousiast: de universiteitsbibliotheken zijn eigenlijk heel tevreden met de bestaande oplossing (IP-gebaseerde whitelisting aan de kant van de uitgevers), en uitgevers zijn tamelijk huiverig om hun systemen aan te passen (om bijvoorbeeld met minder attributen om te kunnen gaan), zeker als dat contractueel niet is vastgelegd.

Als blijkt dat alle partijen wat dit betreft op een lijn zitten en deze oplossing omarmen, kan worden gekeken wie een dergelijke IdP zou moeten aanbieden. Dat zou SURFconext kunnen zijn, SURFmarket, of wellicht een commerciële partij als OCLC (aanbieder van onder meer EZproxy).

5. Conclusie

In dit document is kort beschreven wat de problemen zijn met de huidige manier van het verlenen van toegang voor walk-in users bij universiteitsbibliotheken. Omdat de universiteitsbibliotheken op basis van (onder andere) dit argument terughoudend zijn bij de implementatie van federatieve logins voor wetenschappelijk tijdschriften en online databases, lijkt het zinvol om deze problemen binnen de SURFconext-federatie op te pakken.

We hebben laten zien (zowel binnen AARC als binnen Trust & Identity) dat een IdP die (anonieme) toegang verleent op basis van het IP-adres van de gebruiker, technisch een oplossing kan zijn voor de library-walkin-problematiek. Om zo'n IdP binnen SURFconext beschikbaar te maken, is wel nog wel enig werk nodig op policyniveau.

6. Aanbevelingen

Onze aanbeveling is om nu eerst de businesscase voor deze oplossing te borgen in nieuw te onderhandelen contracten met content-leveranciers. Als blijkt dat zowel uitgevers als universiteitsbibliotheken het eens kunnen worden over de oplossing van een IP-gebaseerde IdP, kan deze op een relatief korte termijn voor SURFconext worden ingericht.

Daarnaast bevelen we aan om in de lopende innovatietrajecten binnen SURFnet deze casus expliciet mee te nemen. Het ligt voor de hand dat in de trajecten rond het inloggen met externe identiteiten (zoals DigiD) of het faciliteren van externe gebruikers, de casus van bibliotheekbezoekers op de langere termijn gemakkelijk als een special case van een algemenere faciliteit zou kunnen worden opgelost.