Remote configuration server for local rewriting proxy instances

Rewriting proxies have been traditionally used to provide off-campus access to resources subject to access control, when this control is applied by source IP-address filtering. By means of such mechanisms, users access the resource through specific URLs that the proxy rewrites back to the original one. The resource finds a connection coming from an authorized IP address and returns the content, which is rewritten by the proxy before delivering it to the requesting user in order to keep the links inside the specific proxy URL space.

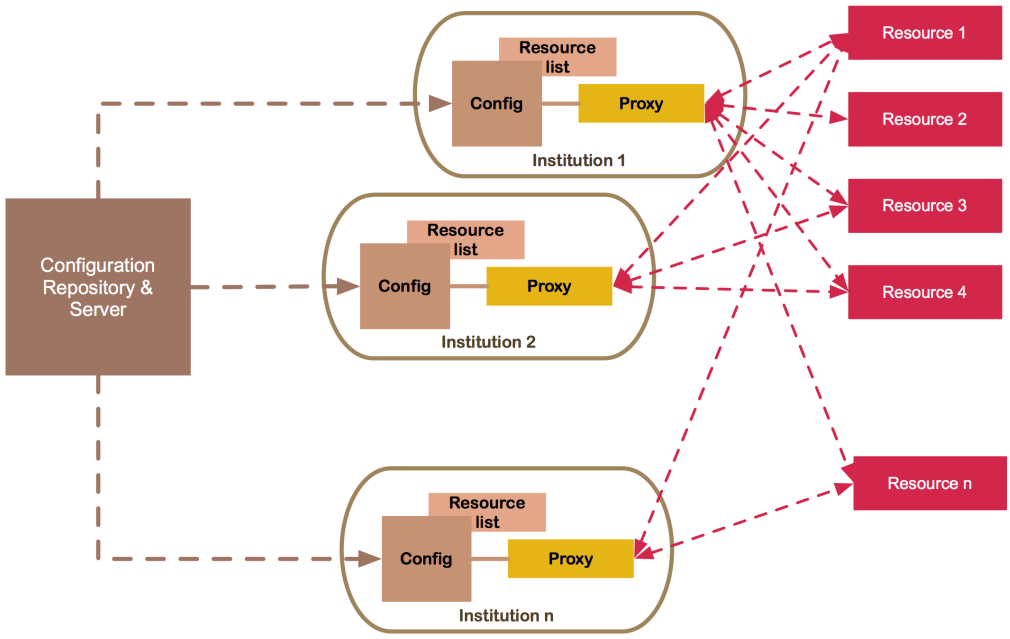
Though it can be argued that this approach predates and/or is against federated access control, it is not actually the case. First, it is obvious that in a federation such proxies will act as SPs. Second, let us consider in brief a few scenarios in which rewriting proxies can be used to enhance federation services.

1. **Link resolvers**. Many reference services in wide use today work by returning general identifiers that are converted to actual links to the resources by means of *link resolver* services. A typical example is DOI (http://www.doi.org/), widely used in the library environment. When users look for a set of references to papers or other published material, it is very likely that they get back a list of these DOIs. Library portals are able to translate these DOIs into actual URLs to the resources, according to the license agreements that the institutions have with publishers. To do so, library portals rely on link resolvers, which accept the DOI and apply their configuration to make the appropriate mapping. Most configuration rules in these resolvers are associated to IP address ranges. A rewriting proxy acting as a front-end to link resolvers could be used to show an institutional IP to the link resolver(s) and rewrite the response back into a federated direct link (usually referred as «WAYFless URLs»), thus enhancing user experience and reducing reluctance based on lack of usability when federation scenarios are intended to be applied in customary workflows.
2. **Reference managers**. This term is usually applied to plugins to document authoring systems that can be used to manage the references in a paper, allowing the authors to make direct queries and citations. These plugins typically rely on simple RESTful interfaces to servers that provide them the information back. It is difficult to make the plugins (and the servers themselves) federation-aware, so for reasons rather similar to the case above the application of rewriting proxies would enhance user experience.
3. **Legacy and firewalled servers**. Resources that, for whatever the reason, are not able to migrate to federated access control constitute the most evident application scenario for the use of federated rewriting proxies. In the case of legacy services this can be considered an interim solution, to fade out as federations become pervasive, but it is important to note that the use of rewriting proxies will translate in faster user awareness and in the perception of the federated access as a comprehensive solution. Moreover, there are other services that run inside firewalled environments, whether in institutional or group-wide "intranets", only allowing connections from those inner addresses. Proxies using compatible addresses are in the position of opening access to these services through federated identity, possibly requiring specific LoAs.
4. **Appliances with web interfaces**. Management of any kind of dedicated equipment or appliance via a web interfaces is commonplace. Many (if not all) of these appliances do not implement web servers able to deal with access method beyond IP-address filtering or HTTP BasicAuth. Rewriting proxies provide an interesting opportunity to include these management interfaces under federated access control.

This proposal considers a usage pattern in which each institution runs an independent installation of rewriting proxies. This pattern is rather common, as a local proxy is able to take into account local IP-based license particularities, show a

local IP address to some of the reference systems, and provide a closer integration with local mechanisms for information searching and retrieval.

Configuring such proxy services may become a tedious (if not complicated for certain resources) task for the staff in charge of the information system installation, as well as fall into pure reactive maintenance, only acting upon user complaints on malfunctions when the proxy rules must be adapted to a change at the resource remote site.



We propose here the implementation of a repository of proxy configuration rules, maintained by the identity federation operator, able to feed configuration data to local proxy installations. The local proxy admins will only need to establish a list of resources to keep an updated configuration for their proxy systems.

As it has been said elsewhere, configurations in this central repository would play a similar role as federation metadata do, since participating institutions would have to put a similar degree of trust in them. Moving further this analogy, a federated web interface to the configuration server could provide additional services to participant institutions, allowing them to manage particular aspects of the resources they use or plan to use, much as metadata management interfaces currently do.