ARTICLE

Off-campus Access to Licensed Online Resources through Shibboleth

Francis Jayakanth, Ananda T. Byrappa, and Raja Visvanathan

ABSTRACT

Institutions of advanced education and research, through their libraries, invest substantially in licensed online resources. Only authorized users of an institution are entitled to access licensed online resources. Seamless on-campus access to licensed resources happens mostly through Internet Protocol (IP) address authentication. Increasingly, licensed online resources are accessed by authorized users from off-campus locations as well. Libraries will, therefore, need to ensure seamless off-campus access to authorized users. Libraries have been using various technologies, including proxy server or virtual private network (VPN) server or single sign-on, to facilitate seamless off-campus access to licensed resources. In this paper, authors share their experience in setting up a Shibboleth-based single sign-on (SSO) access management system at the JRD Tata Memorial Library, Indian Institute of Science, to enable authorized users of the institute to seamlessly access licensed online resources from off-campus locations.

INTRODUCTION

The Internet has both necessitated and offered options for libraries to enable remote access to an organization's licensed online content—journals, e-books, technical standards, bibliographical and full-text databases, and more. In the absence of such an option for remote access, faculty, students, and researchers have limited and constrained access to the licensed online content from off-campus locations.

As scholarly resources transitioned from print to online in the mid-1990s, libraries and their vendors had to start identifying user affiliations in order to grant access to licensed online resources to the authorized users of an institution. The IP address was an obvious mechanism to do that. Allowing or denying access to online resources based on a user's IP address was simple, it worked, and, in the absence of practical alternatives, it became the universal means of authentication for gaining access to licensed library content.[1] To facilitate seamless access to licensed online resources from off-campus sites, libraries have been using various technologies including proxy server or VPN server or remote desktop gateway or federated identity management or a combination of the said technologies.

In our institute, the on-campus IP-based access to the licensed content is supplemented by VPN technology for off-campus access. The COVID-19 pandemic has necessitated academic and scientific staff work from home, which demands smooth and seamless access to the organization's licensed content. The sudden surge in demand for seamless off-campus access to the licensed online resources had an impact on the institute's VPN server. Also, not all authorized users of the

Francis Jayakanth (francis@iisc.ac.in) is Scientific Officer, J.R.D. Tata Memorial Library, Indian Institute of Science. **Ananda T. Byrappa** (anandtb@iisc.ac.in) is Librarian, J.R.D. Tata Memorial Library, Indian Institute of Science. **Raja Visvanathan** (raja@inflibnet.ac.in) is Scientist C (Computer Science), INFLIBNET Centre, Gandhinagar, India.

https://doi.org/10.6017/ital.v40i2.12589

institute are entitled to get VPN access. To mitigate the situation, the library, therefore, had to explore a secure, reliable, and cost-effective solution to facilitate seamless off-campus access to all the licensed online resources to all the authorized users of the institute. After exploring the possibilities, the library decided to implement a single sign-on solution based on Shibboleth. Shibboleth software implements the Security Assertion Markup Language (SAML) protocol, separating the functions of authentication (undertaken by the library or university, which knows its community of end users) and authorization (undertaken by the resource provider, which knows which libraries have licenses for their users to access the resource in question).[2]

ABOUT THE INDIAN INSTITUTE OF SCIENCE (IISC)

The Indian Institute of Science (IISc, or "the Institute") was established in 1909 by a visionary partnership between the industrialist Jamsetji Nusserwanji Tata, the Maharaja of Mysore, and the Government of India. Over the 109 years since its establishment, IISc has become the premier institute for advanced scientific and technological research and education in India. Since its inception, the Institute has laid a balanced emphasis on the pursuit of fundamental knowledge in science and engineering, and the application of its research findings for industrial and social benefit. During 2017–18, the Institute initiated the practice of undergoing international peer academic reviews over a 5-year cycle. Each year, a small team of invited international experts reviews a set of departments. The experts spend 3 to 4 days at the Institute. During this period, they interact closely with the faculty and students of these departments and tour the facilities, aiming to assess the academic work against international benchmarks.

IISc has topped the Ministry of Human Resource Development (MHRD), Government of India's NIRF (National Institutional Ranking Framework) rankings not only in the university's category but also overall among all ranked institutions. Times Higher Education has placed IISc at the 8th position in its Small University Rankings (that is, among universities with fewer than 5,000 students), at the 13th position in its ranking of universities in the Emerging Economies, and in the range 91–100 in its World Reputation Rankings. In the QS World University Rankings, IISc is ranked 170. In the same ranking system, on the metric of Citations per Faculty, IISc is placed in second position.

IISc publishes about 3,000 papers per year in Scopus and Web of Science indexed journals and conferences and, each year, the Institute awards around 400 PhD degrees.

About the IISc Library

JRD Tata Memorial Library (https://www.library.iisc.ac.in), popularly known as the Indian Institute of Science Library, is one of the best science and technology libraries in India. Started in 1911, as one of the first three departments in the Institute, it has become a precious national resource center in the field of science and technology. The library receives annually a grant of 10-12% of the total budget of the Institute. The library spends about 95% of its budget toward periodical subscriptions, which is unparalleled in this part of the globe. With a collection of nearly 500,000 volumes of books, periodicals, technical reports and standards, the JRD Tata Memorial Library is one of the finest in the country. Currently, it subscribes to over 13,000 current periodicals. The library also maintains the IISc's research publications repository, ePrints@IISc (http://eprints.iisc.ac.in), and its theses and dissertations repository, etd@IISc (https://etd.iisc.ac.in).

OFF-CAMPUS ACCESS TO LICENSED ONLINE RESOURCES

In a typical research library, licensed scholarly resources comprise research databases, electronic journals, e-books, standards, and more. A library licenses these resources through publishers/vendors. These license agreements limit access to the resources to the authorized users of an institute. In our case, authorized users include faculty members, enrolled students, current staff, contractual staff, and walk-in users to the library.

Seamless access to the licensed resources from on-campus sites is predominantly IP-address authenticated, which is a simple and efficient model for users physically located on the institute campus. These users expect a similar experience while accessing licensed online resources from off-campus locations. Therefore, the challenge to the libraries is to ensure that such off-campus accesses are secure, seamless, and restricted to authorized users of an institute. Libraries have been using various technologies including proxy servers, VPN servers, or single sign-on to facilitate seamless off-campus access to licensed resources. Our institute has been using VPN technology to enable off-campus access to licensed online resources.

A virtual private network (VPN) is a service offered by many organizations to its members to enable them to remotely connect to the organization's private network. A VPN extends a private network across a public network and allows users to send and receive data across shared or public networks as if their computing devices were directly connected to the private network. Applications running across a VPN may therefore benefit from the functionality, security, and management of the private network. Encryption is common, although not an inherent, part of a VPN connection.[3]

In our Institute, faculty members and students are provided access to the VPN service when their Institute email address is created. Users follow four steps to use a VPN client to get connected to the campus network:

- Install VPN client software on their computer system. Cisco AnyConnect (https://www.cisco.com/c/en/us/products/collateral/security/anyconnect-secure-mobility-client/at-a-glance-c45-578609.html) is one such software.
- Start the VPN client software every time there is a need to connect to the private network.
- Enter the address of the Institute's VPN server, and click Connect in the AnyConnect window.
- Log in to the VPN server using their institutional email credentials.

An authorized user of the Institute can use any of the IP authenticated network services, including the licensed online resources, after a successful login to the VPN server. The VPN technology has been serving the purpose well, but the service is, by default, available only to the institute's faculty and students. Other categories of employees such as project assistants, project associates, research assistants, post-doctoral fellows, and others, who constitute a good percentage of IISc staff, are provided VPN access on a case-by-case basis. During the COVID-19 lockdown, the library received several enquiries about accessing the online resources from off-campus sites. Realizing the importance of the situation, the library quickly assessed the various possibilities for facilitating seamless off-campus access to the subscribed online resources apart from the VPN-based access. Federated access through Shibboleth identity provider (IdP) service emerged as a possible solution to facilitate seamless off-campus access to the entire Institute community.

FEDERATED ACCESS

Federated access is a model for access control in which authentication and authorization are separated and handled by different parties. If a user wishes to access a resource controlled by a service provider (SP), the user logs in via an identity provider (IdP). More complex forms of federated access involve the use of attributes (information about the user passed from the IdP to the SP, which can be used to make access decisions) and can include extra services such as trust federations and discovery services (where the user selects which IdP to use to connect to the SP).[4]

Examples of this federated access model include Shibboleth and OpenAthens. Shibboleth is open-source software that offers single sign-on infrastructure. OpenAthens is a commercial product delivered as a cloud-based solution. It supports many of the same standards as Shibboleth. So, an institution could pay and join the OpenAthens federation, which will provide technical support to set up, integrate, and operationalize federated access using OpenAthens. We decided to go with Shibboleth for the following reasons:

- To avoid the recurring cost associated with the OpenAthens solution.
- The existence of a Shibboleth-based INFED federation in the country. INFED manages the trust between the participating institutions and publishers (http://infed.inflibnet.ac.in/).
- INFED is part of the eduGain inter-federation, which enables our users to gain access to the resources of federations of other countries.

WHAT IS SHIBBOLETH?

Shibboleth is a standards-based, open-source software package for web single sign-on across or within organizational boundaries. It allows sites to make informed authorization decisions for individual access of protected online resources in a privacy-preserving manner.

The Shibboleth software implements widely used federated identity standards, principally the OASIS Security Assertion Markup Language (SAML), to provide a federated single sign-on and attribute exchange framework. A user authenticates with their organizational credentials, and the organization (or identity provider) passes the minimal identity information necessary to the service provider to enable an authorization decision. Shibboleth also provides extended privacy functionality allowing a user and their home site to control the attributes released to each application (https://www.shibboleth.net/index/).

Shibboleth has two major components: (1) an identity provider (IdP), and (2) a service provider (SP). The IdP supplies required authorizations and attributes about the users to the service providers (for example, publishers). The service providers make use of the information about the users sent by the IdP to make decisions on whether to allow or deny access to their resources.

INTERACTION BETWEEN A SHIBBOLETH IDENTITY PROVIDER AND SERVICE PROVIDER.

When a user attempts to access licensed content on the service provider's platform, the service provider generates an authentication request and then directs the request and the user to the user's IdP server. The IdP prompts for the login credentials. In our setup, the IdP server communicates the login credentials to the Institute's active directory (AD) using the secure Lightweight Directory Access Protocol (LDAP).

AD is a directory service provided by Microsoft. In a directory service, objects (such as a user, a group, a computer, a printer, or a shared folder) are arranged in a hierarchical manner facilitating easy access to the objects. Organizations primarily use AD to perform authentication and authorization. Once the authenticity of a user is verified, AD helps in determining if a user is authorized to use a specific resource or service. Access is granted to a user only if the user checks out on both counts.

The AD authenticates a user, and the response is sent back to the IdP along with the required attributes. The IdP then releases only the required set of attributes to the service provider. Based on the IdP attributes, which is nothing but a user's entitlement, the SP grants access to the resource. Figure 1 illustrates the functioning of the two components of Shibboleth.

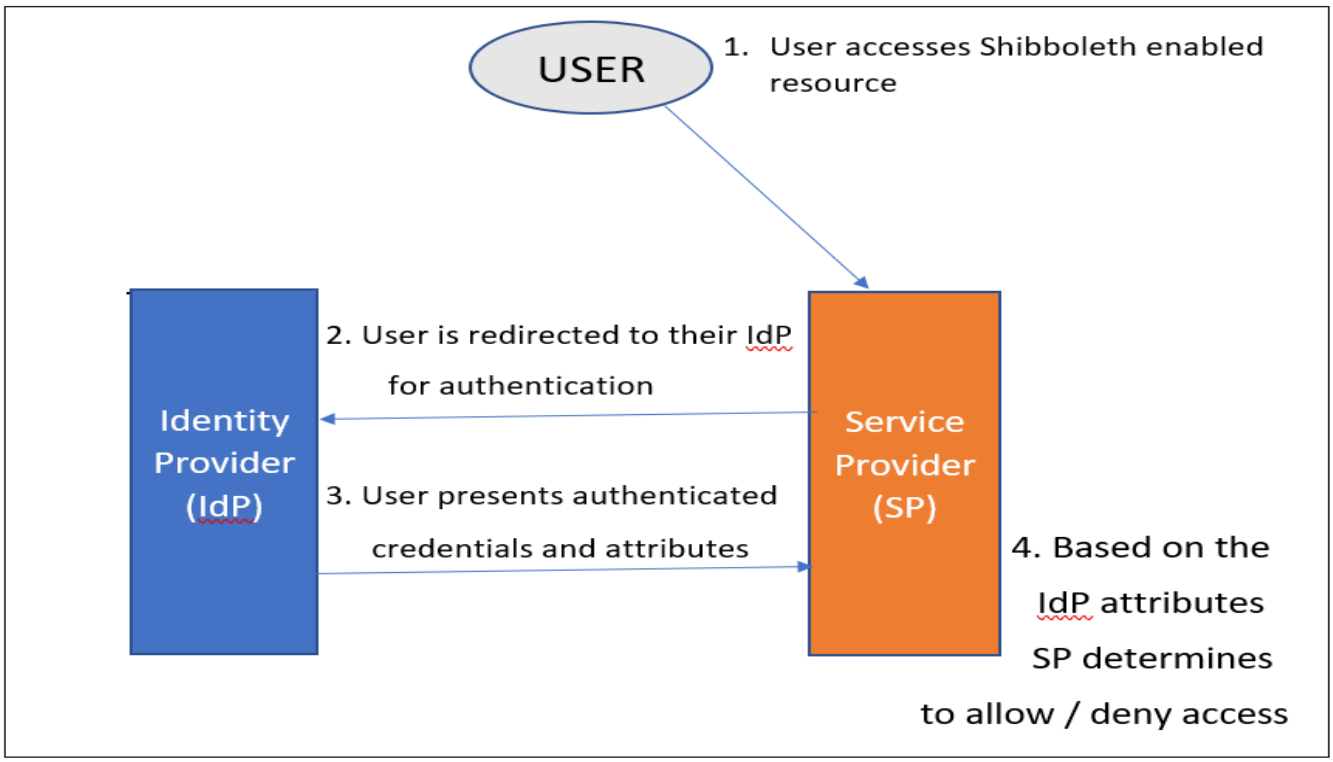

Figure 1. A Shibboleth workflow involving a user, identity provider, and service provider.

IDENTITY FEDERATION

The interaction between a service provider and identity provider happens based on mutual trust. The trust is established by providing IdP metadata as encrypted keys and the IdP URL that the SP uses to send and request information from the IdP. The exchange of metadata between IdP and SP can be informal if an institution licenses online resources from only a few publishers. However, research libraries license content from hundreds of SPs. Therefore, the role of federations is significant. In the absence of a federation, each identity provider and service provider must individually communicate with each other about their existence and configuration, as illustrated in figure 2.

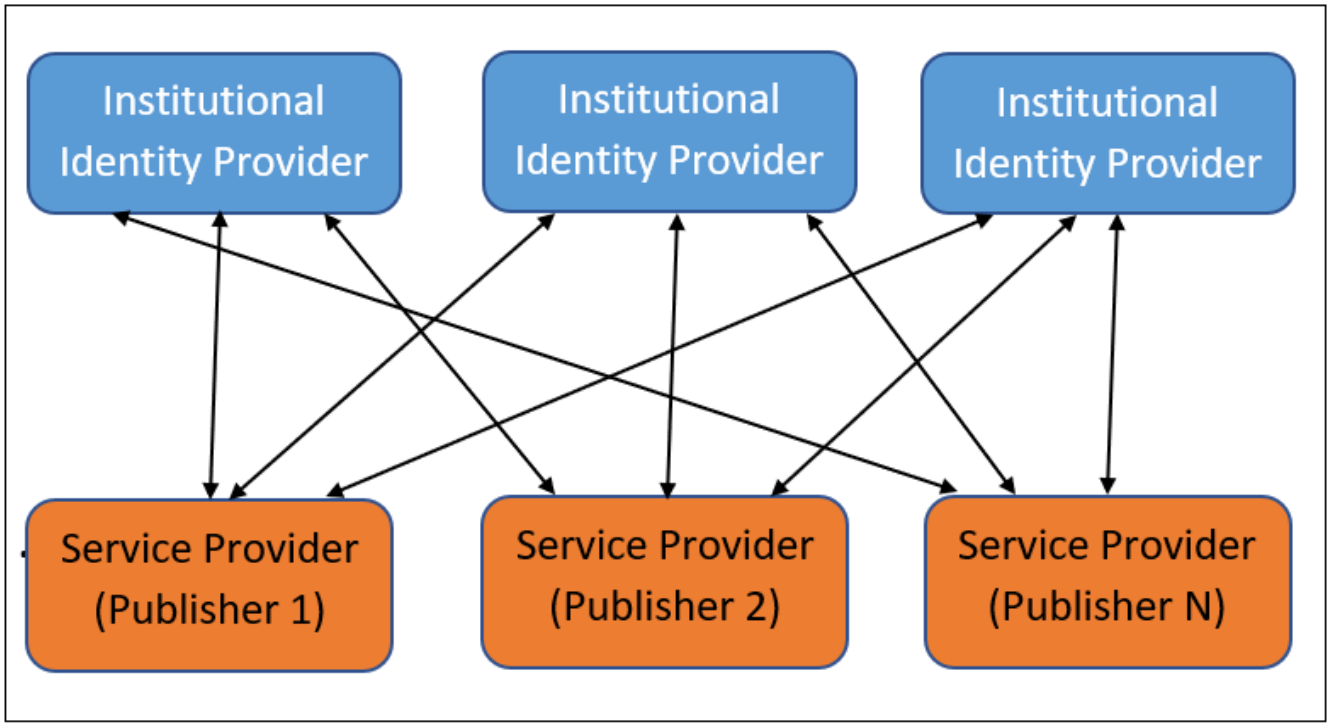

Figure 2. Individual communication between IdPs and SPs.

A federation is merely a list of metadata entries aggregated from their member IdPs and their SPs. Our Institute is a member of INFED (Information and Library Network Access Management Federation). INFED was established as a centralized agency to coordinate with member institutions in the process of implementing user authentication and access control mechanism across all member institutions. INFED manages the trust relationship between the IdPs and SPs (publishers) in India. Therefore, individual IdPs that intend to facilitate access to subscribed online resources through Shibboleth will share their metadata with INFED. INFED, in turn, will share the metadata of the IdPs with respective service providers, as illustrated in figure 3. Other regions have their federations. For example, n the US, InCommon (https://www.incommon.org/) serves as the federation, and in the UK, it is the UK Access Management Federation (http://www.ukfederation.org.uk/).

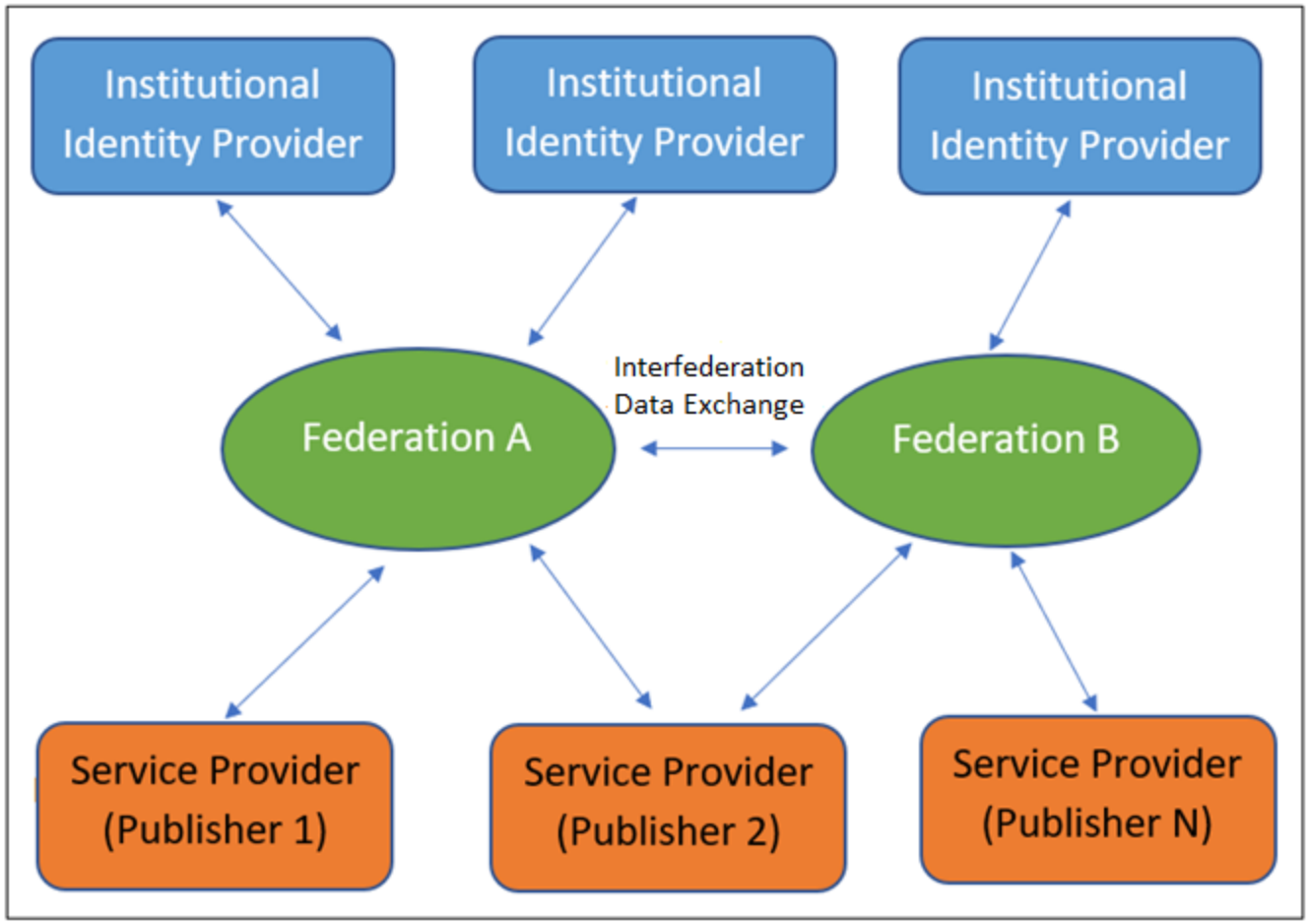

Figure 3. Role of a federation as a trust manager between IdPs and SPs.

HOW DOES ONE GAIN ACCESS TO SHIBBOLETH-ENABLED RESOURCES?

A federation manages the trust between identity providers and service providers. The SPs enable Shibboleth-based access to subscribed resources to the IdPs based on the metadata shared by a federation. Once the SPs allow access, users can access such resources by using the institutional login option via the Athens/Shibboleth link found on the service provider's platform. Alternatively, a library can create a simple HTML page listing all the Shibboleth-enabled licensed resources, as shown in figure 4.

Figure 4. Partial screenshot of Shibboleth-enabled resources of our Institute.

Each of the links in figure 4 is a WAYFless URL. A WAYFless URL is specific to an institution (IdP), and it enables users of that institution to gain federated access to a service or resource in a way that bypasses Where Are you From (WAYF), or the institutional login (discovery service) steps on the SP's platform. Since the institutional login or the discovery service step can be confusing to end users, WAYFless links to the resources will facilitate an improved end-user experience in accessing licensed resources. A user needs to follow a link from the list of resources. The link will take the user to the SP. The SP will redirect the user to the IdP server for authentication. After successful authentication, the user will gain access to the resource.

There are two ways to get a WAYFless URL to a service: (1) The service provider can share the URL or (2) One can make use of a WAYFless URL generator service like WUGEN (https://wugen.ukfederation.org.uk/wugen/login.xhtml).

BENEFITS OF SHIBBOLETH-BASED ACCESS

Shibboleth-based single sign-on can effectively address several requirements of the libraries in ensuring secure and seamless on-campus and off-campus access to subscribed online resources. There are other benefits of Shibboleth-based SSO:

1. It is open-source software that provides single sign-on infrastructure.
2. It enables organizations to use their existing user authentication mechanism to facilitate seamless access to licensed online resources.
3. Being a single sign-on system, for the end users, it eliminates the need to have individual credentials for each online resource.
4. It uses security assertion mark-up language (SAML) to securely transfer information about the process of authentication and authorization.
5. It is used by most of the publishers, who facilitate Shibboleth-based access through Shibboleth federations.
6. It requires a formal federation as a trusted interface between the institutions as an identity provider (IdP) and publishers as service providers (SP) thereby ensuring the use of uniform standards and protocols while transmitting attributes of authorised users to publishers. INFLIBNET's access management federation, INFED, plays this role (https://parichay.inflibnet.ac.in/objectives.php).

IDP SERVER CONFIGURATION

We installed the Shibboleth IdP software version 3.3.2 on a virtual machine on the Azure platform. The VM system is configured with two virtual CPUs, 4 GB of RAM, 300 GiB of OS disk (standard HDD), and Ubuntu Linux OS version 18.04.4 LTS.

Coordination with the organization's network support team is essential. The network support team handles the domain name service resolution of the IdP server and facilitates the IdP server to communicate with the organization's active directory and to open non-standard communication ports on the IdP server.

SHIBBOLETH IDP USAGE STATISTICS

The INFED team has developed a beta version of the usage analysis tool called INFEDStat to analyse the use of federated access to gain access to licensed resources. We have implemented the tool on the IdP server. Figure 5 shows the redacted screenshot of the INFEDStat dashboard. It shows

- Date-wise usage details of logged-in users along with IP address, time logged in, and the publishers' platforms accessed,
- Number of times the publishers' platforms were accessed during a specific period,
- Number of times users logged in for a specific period,
- Unique users for a specific period, and
- Unique publishers accessed during a specific period.

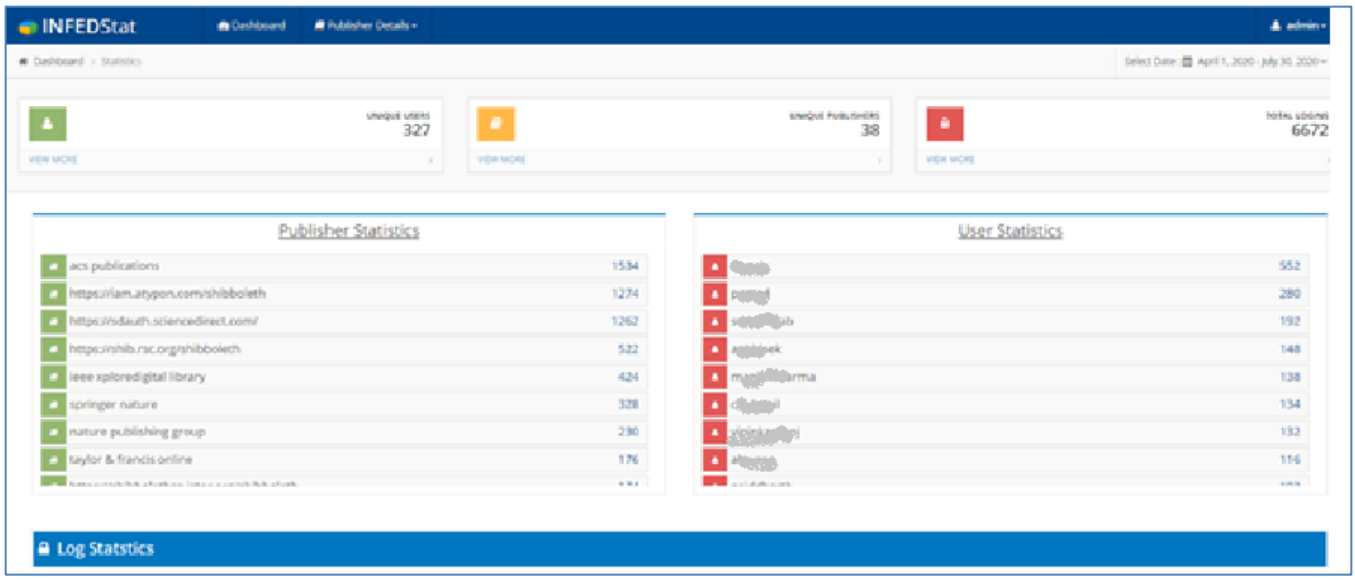

Figure 5. IdP usage dashboard.

CONCLUSIONS

The implementation of federated access to subscribed online resources has ensured that all the authorized users of the Institute can access almost all the licensed resources from wherever they are. The COUNTER 5 usage analysis of subscribed resources for the period of January 2020 to October 2020 indicates that usage of online resources has increased by nearly 20 percent over the last year for the same period. The enhanced use could be partly because of ease of accessing online resources facilitated by federated access. To assess the reasons for enhanced usage of online resources, the library is planning to conduct a survey to understand how convenient and useful federated access to online resources has been especially while being off campus.

Federated access through single sign-on is useful not just for accessing licensed online resources. A typical research library offers various other services to its users, including the institutional repository service, learning management system, online catalogue, etc. The library intends to integrate such services with SSO, thereby freeing the end users from service-specific credentials.

ENDNOTES

[1] Thomas Dowling, "We Have Outgrown IP Authentication," *Journal of Electronic Resources Librarianship* 32, no. 1 (2020): 39–46, https://doi.org/10.1080/1941126X.2019.1709738.

[2] John Paschoud, "Shibboleth and SAML: At Last, a Viable Global Standard for Resource Access Management," *New Review of Information Networking* 10, no. 2 (2004): 147–60, https://doi.org/10.1080/13614570500053874.

[3] Andrew G. Mason, ed., *Cisco Secure Virtual Private Network* (Cisco Press, 2001): 7, https://www.ciscopress.com/store/cisco-secure-virtual-private-networks-9781587050336.

[4] Masha Garibyan, Simon McLeish, and John Paschoud, "Current Access Management Technologies," in *Access and Identity Management for Libraries: Controlling Access to Online Information* (London, UK: Facet Publishing, 2014): 31–38.