The Georgia Tech Research Institute (GTRI)
Information Technology & Telecommunications Laboratory
250 14th Street, NW • Atlanta, Georgia 30332
(404) 407-8956 • Internet: www.gtri.gatech.edu

Project Report

Global Federation Identity and Privilege Management (GFIPM) Security Interoperability Demonstration

Report Submitted: *<August 30, 2007>*

Submitted by: The Georgia Tech Research Institute, GTRI

Acknowledgements 5
Executive Summary 7
1 Background 13
1.1 Problem Statement 13
1.2 Project Initiation and Funding 14
1.3 Project Objectives 14
2 Federated Identity and Privilege Management Concept 15
2.1 An Example of the Problem 15
2.2 The FIPM Concept 16
2.2.1 Identity Provider (IDP) 17
2.2.2 Service Provider (SP) 17
2.2.3 Common Metadata Attribute Language 18
2.3 The FIPM Value Proposition 19
2.4 GFIPM: Applying FIPM to the Justice Community 21
3 Project Overview 21
3.1 Approach 21
3.1.1 Use of Existing Vetted Participant Subscribers and Information 22
3.1.2 Use of Existing Participant Authentication Mechanisms and Policies 23
3.1.3 Use of Production Web-Based Resources 23
3.1.4 Use of Security Assertion Markup Language (SAML) version 1.1 23
3.1.5 Use of the Shibboleth Open Source SAML Middleware 24
3.1.6 Use of the Internet as a Backbone for Connecting Participants 25
3.1.7 Implementation of a Common Metadata Model for Assertions 26
3.2 General Descriptions of Project Participants 26
3.2.1 Criminal Information Sharing Alliance (CISA) 27
3.2.2 Pennsylvania Justice Network (JNET) 28
3.2.3 Regional Information Sharing Systems (RISS) 29
3.2.4 Georgia Tech Research Institute (GTRI) 30
3.3 Assumptions and Constraints 31
3.3.1 Leverage NIEM 31
3.3.2 No Sharing of Intelligence Information during Initial Phase 31
3.3.3 Transition to SAML 2.0 32
3.3.4 User-to-Application Use Case 32
4 Architecture 32
4.1 IDP Structure 32
4.1.1 Shibboleth IDP Middleware Module 33
4.1.2 Web Servlet Container 33
4.1.3 IDP Integration Points 34
4.1.3.1 Single Sign-On (SSO) Integration Point 34
4.1.3.2 Attribute Authority (AA) Integration Point 34
4.1.4 Web Single Sign-On (SSO) System 35
4.1.5 Attribute Data Store 35
4.1.6 Attribute Authority (AA) Connector and Assertion Generator 35
4.2 SP Structure 37
4.2.1 Web Server 37

4.2.2 Shibboleth SP Middleware 38
4.2.3 Protected Resource Integration Point 38
4.2.4 Optional GFIPM-Enabled Proxy/Portal Service 39
4.2.5 Protected Resources 41
4.3 SAML Usage Profile for GFIPM 41
4.3.1 Basic SAML Concepts 42
4.3.1.1 Authentication Statements 42
4.3.1.2 Attribute Statements 43
4.3.1.3 SAML Assertions 44
4.3.2 SAML Components Used by GFIPM 45
4.3.2.1 Shibboleth/SAML Web Browser Single Sign-On (SSO) Profile 45
4.3.2.2 SAML Assertion Query Profile 47
4.4 Where-Are-You-From (WAYF) Service 47
4.5 Federation Trust Fabric (Federated Entities Metadata) 48
4.6 GFIPM Metadata 50
4.6.1 Purpose of the Metadata 50
4.6.2 Metadata Framework 50
4.6.2.1 Conceptual Model Layer 52
4.6.2.2 Federation Profile Layer 53
4.6.2.3 Federation Profile Instance Layer 53
4.6.2.4 The SAML Assertion Layer 53
4.6.3 Overview of Metadata Content 55
4.6.3.1 Basic Terminology 55
4.6.3.2 GFIPM User Assertion Metadata Contents 56
4.6.3.3 GFIPM Entity Assertion Metadata Contents 56
4.6.3.4 Noteworthy Metadata Elements 57
4.6.4 Metadata Development Process 58
4.7 Miscellaneous 60
4.7.1 Firewall Considerations 60
4.7.2 Web Browser Considerations 61
5 Project Execution and Timeline 61
5.1 Project Initiation 62
5.2 GFIPM Infrastructure Establishment 64
5.3 Participant IDP/SP Development and Integration 68
5.4 User Testing, Evaluation, and Infrastructure Refinement 72
6 Pilot Federation 75
6.1 Reference Implementation 76
6.1.1 Reference IDPs 76
6.1.2 Reference SPs 77
6.1.3 Where-Are-You-From (WAYF) Service 78
6.2 Criminal Information Sharing Alliance (CISA) 78
6.2.1 CISA IDP 78
6.2.2 CISA SP 78
6.2.3 CISA Resources 79
6.2.4 CISA Users 80
6.3 Pennsylvania Justice Network (JNET) 81

6.3.1 JNET IDP 81
6.3.2 JNET SP 81
6.3.3 JNET Resources 82
6.3.4 JNET Users 83
6.4 Regional Information Sharing Systems (RISS) 85
6.4.1 RISS IDP 85
6.4.2 RISS SP 85
6.4.3 RISS Resources 85
6.4.4 RISS Users 85
6.5 GFIPM Pilot Federation User Demographics 86
6.6 GFIPM User Portal and Resource Directory 88
6.7 Sample Screen Shots of an SSO Transaction 88
7 Lessons Learned and Conclusions 93
7.1 GFIPM Concept 94
7.2 Business Case for GFIPM (When and Why to Join) 95
7.3 Metadata and Infrastructure 97
7.4 Federation Enablement of Resources 100
7.4.1 Resource Integration Profiles 101
7.4.2 Resource Integration Techniques 102
7.4.3 Profiles and Techniques for Existing Resources 104
7.5 Usability and User Support 106
7.6 Governance and Operational Support 107
8 Next Steps 110
8.1 GFIPM Standards Development, Validation, and Vetting 111
8.2 Expansion of Participants and Development of Production-Level Capabilities. 111
8.3 Establishment of Federation Governance Structure 112
Appendix A: Related Initiatives and Standards 113
A.1 E-Authentication 113
A.2 Law Enforcement Information Sharing Program (LEISP) 113
A.3 National Information Exchange Model (NIEM) 114
A.4 eXtensible Access Control Markup Language (XACML) 114
A.5 Service Provisioning Markup Language (SPML) 114
A.6 Global Technical Privacy Task Team 115
A.7 DHS Project on Authority-Based Access Control (ABAC) 115

Acknowledgements

The U.S. Department of Justice's Global Justice Information Sharing Initiative (Global) serves as a Federal Advisory Committee to the U.S. Attorney General on critical justice information sharing initiatives. Global promotes standards-based electronic information exchange to provide justice and public safety communities with timely, accurate, complete, and accessible information in a secure and trusted environment. Global is administered by the U.S. Department of Justice, Office of Justice Programs, and Bureau of Justice Assistance.

Global aids its member organizations and the people they serve through a series of important initiatives. These include the facilitation of Global working groups. The Global Security Working Group (GSWG) is one of four Global working groups covering critical topics such as intelligence, privacy, security, and standards. Ms. Chelle Uecker serves as the current chair of the GSWG.

The Global Federated Identity and Privilege Management (GFIPM) Security Interoperability Demonstration was initiated and overseen by the GSWG, and funded through a partnership between the Bureau of Justice Assistance (BJA), National Institute of Justice, and the Department of Homeland Security (DHS) Office of Chief Information Officer. Organizations that participated in the project included the Georgia Tech Research Institute (GTRI), the Criminal Information Sharing Alliance (CISA), the Pennsylvania Justice Network (JNET), and the Regional Information Sharing Systems Network (RISS). Funding for project management, engineering and development, and technical assistance was provided to GTRI through grants from BJA and DHS. Funding for other participants was provided directly to the participants from BJA in separate grants.

The *GFIPM Security Interoperability Demonstration Final Report* was developed primarily by the Georgia Tech Research Institute, with contributions from CISA, JNET, and RISS. The following individuals are acknowledged for their contributions to the GFIPM project and to this report.

- John Wandelt (GTRI) – GFIPM Project Director
- Matt Moyer (GTRI) – GFIPM Technical Lead for GTRI
- Jeff Krug (GTRI) – Research Scientist
- Stefan Roth (GTRI) – CISA Technical Representative
- Glen Gillum (CISA) – Director, Criminal Information Sharing Alliance
- John Davenport (JNET) – JNET Deputy Architect
- Joe Riggione (JNET) – JNET Project Manager
- James Dyche (JNET) – JNET Architecture Manager
- Darrell Candis (JNET) – JNET Security Architect
- Larry Maloney (RISS) – RISS Senior Project Manager
- Michael McDaniel (RISS) – RISS Applications Development Group

In addition, the following individuals are acknowledged for their contributions as GFIPM project liaisons and reviewers of this report.

- Christina Rogers, California Department of Justice – Global Security Architecture Committee
- Martin Smith, Department of Homeland Security
- John Ruegg, Los Angeles County Information Systems Advisory Body – Liaison Global Security Working Group
- Mrs. Chelle Uecker, National Association for Court Management – Chair of Global Security Working Group
- Chris Traver, Senior Policy Advisor, Bureau Justice Assistance

For additional information on the Global Information Sharing and GFIPM Initiatives please see it.ojp.gov and it.ojp.gov/GFIPM.

Executive Summary

Within the U.S. justice and law enforcement community, there are many recognized sensitive-but-unclassified (SBU) networks and information systems. Each of these systems constitutes an investment in technology, governance structures, and trust relationships designed to support the secure sharing of information with a target user base. In addition, each system is designed to comply with specific laws or policies regarding security and privacy of the information to be shared. As these systems are mostly independent of each other, in most cases there is little or no interoperability (technological or otherwise) between one system and another. Therefore, while each system provides value to its users and its business domain, on a wider scale the information sharing landscape resembles a collection of closed silos more than an open environment. This situation presents a major problem, because effective law enforcement today requires that the right individuals have access to the authorized resources they need regardless of where the user or resource resides within the enterprise.

To get around this problem, users typically subscribe to each system or resource separately, as needed for their job functions, by undertaking multiple system and user registration processes. This results in users having to manage multiple security credentials (certificates, usernames, passwords, etc.) This arrangement is tedious, time consuming, and frustrating for users, and it scales poorly as the number of resources increases. From an administrative standpoint, this solution is unattractive because it requires vetting ("Who are you?"), permissioning ("What can you access?"), and credentialing ("How do I know it's you?") processes to occur many times for each user. This is expensive for resource owners, as multifactor credentials and high-assurance vetting processes quickly become too costly and inefficient. As a result, these processes tend to be managed poorly, time consuming, with strong multifactor credentialing sometimes replaced by weaker single-factor credentialing, and vetting often limited to one-time, relatively low-assurance processes rather than ongoing, high-assurance processes.

In light of these problems, it is clear that the entire justice and law enforcement community stands to benefit from a solution that allows user management processes (vetting, permissioning, and credentialing) to be efficiently and effectively leveraged and reused across trust domains, thereby facilitating interoperability between SBU information systems. The Global Federated Identity and Privilege Management (GFIPM) Security Interoperability Demonstration Project has investigated the concept of federated identity and privilege management (FIPM) as a candidate solution. FIPM seeks to connect existing law enforcement and justice organizations with each other at a high level in a business arrangement called a federation, and to use the federation as a basis for building trust relationships that facilitate cross-trust-domain interoperability at a lower level between resources and users and ultimately provide users with access to the data they need to be effective.

Fundamental to the FIPM concept is the separation of responsibility for managing users and user identities from the responsibility for managing resources. An organizational entity that manages users and user identities is called an identity provider (IDP), and an organizational entity that manages resources is called a service provider (SP). A federation consists of one or more IDPs and one or more SPs that join together in a business arrangement for a common goal, which in the case of the GFIPM demo project is to achieve more cost-efficient, timely, and seamless information sharing across trust boundaries. At the implementation level, a federation is supported by several basic security technologies, including cryptographic trust fabric, federated identity management standards and middleware, and a metadata model for securely exchanging information about users.

Within a federation, an IDP performs basic user management tasks such as vetting, credentialing, and authentication. Vetting and credentialing are performed once per user rather than every time a user needs access to a new resource. This provides many benefits to organizations, including a huge savings in vetting and credentialing costs, higher-quality vetting (since it is performed by the user's home organization), and better credentialing (since the reduced cost is often enough to enable the use of multifactor credentialing systems for users.) It is also better for users, since each user needs only one credential to achieve single sign-on (SSO) to federation resources.

In a federation arrangement, an SP maintains complete control of its resources and performs basic resource management tasks, including definition and enforcement of resource access requirements and access control policies. To enforce access requirements and policies, SPs rely on IDPs to provide a set of informational attributes (metadata) about their users known as a *GFIPM user assertion*. GFIPM establishes a standard, well-defined set of metadata about users, so that all organizations have a common framework (semantics and representation) for describing basic user information such as identity, certifications, memberships, affiliations, and contact information. This metadata pertains to users, but is never asserted directly by users. Instead, it is asserted on the user's behalf by the user's IDP, so it can be trusted by SPs within the context of underlying federation-level trust arrangements.

The GFIPM demo project was initiated by the Global Security Working Group in 2005 in response to the National Criminal Intelligence Sharing Plan (NCISP), and it was funded jointly by the Bureau of Justice Assistance (BJA), National Institute of Justice (NIJ), and Department of Homeland Security (DHS) Office of Chief Information Officer (OCIO). The initial demonstration phase of the project included three existing networks within the justice community: the Criminal Information Sharing Alliance Network (CISAnet), the Pennsylvania Justice Network (JNET), and the Regional Information Sharing Systems Network (RISSNET). Project management, engineering, and technical assistance for the demo project were provided by Georgia Tech Research Institute (GTRI).

The demo project sought to achieve the following specific objectives.

- Demonstrate that registered subscribers of one federation participant can access the Web-based resources of another federation participant without the requirement to register with more than one federation participant.
- Demonstrate that federation participants can retain control over their resources (applications, databases, etc.) and make local dissemination and access control decisions based on a shared set of standard user attributes.
- Demonstrate single sign-on (SSO) across federation Web applications.
- Demonstrate federated authentication and authorization between disparate local technologies and vendor implementations.
- Establish a baseline of common attributes for identity and authorization assertions.

The following points highlight the key aspects of the project approach and scope.

- Use of existing vetted participant subscribers and associated information stores, security and authentication mechanisms, policies, and production resources. Explore and validate the GFIPM concept in a real production environment. Identify roadblocks and collect lessons learned in working through the technology, business, and policy related challenges.
- Focus initially on the law enforcement domain and the associated information sharing security requirements, resources, and user base.
- Focus initially on the user-to-application (Web browser to Web application) use case rather than system-to-system or other use cases.
- Develop, agree upon, and use a standard set of attributes with well-defined semantics and an XML syntax based on the National Information Exchange Model (NIEM) standard.
- Leverage the Internet communications backbone for all network traffic.
- Use standard browser Secure Socket Layer (SSL)/Transport Layer Security (TLS) for session establishment and encryption. Require no special software on the client side.
- Leverage the existing Security Assertion Markup Language (SAML) federated identity standard, and an open source implementation of SAML, to minimize software development and implementation costs for participants. (Note that none of the participants had a SAML federated identity capability in place at the time the project began.)
- Extend and configure existing open source products to provide the base GFIPM federation infrastructure. Provide simple well-defined integration points for IDPs and SPs. Participants would use these integration points to enable their existing legacy infrastructure and join the federation as an IDP, SP, or both.
- Allow SPs to make all authorization decisions for their resources based on metadata passed in SAML assertions.

The GFIPM project has produced several valuable outcomes and deliverables including a set of draft GFIPM standards, an initial set of freely available GFIPM federation

infrastructure middleware and tools, extensive documentation including lessons learned and technical point papers, and most importantly an operational GFIPM pilot federation.

The pilot federation employs the GFIPM concept and technologies to actively enable the sharing of information between real world justice and public safety users and production resources. This federation provides real operational value today for both users and resource owners. The SBU resources being shared currently fall into the following categories: criminal investigative, criminal history, criminal justice, counter-terrorism, and general government. Additionally, participants are committed to sharing criminal intelligence information once formal governance processes have been put in place. Demographics from the current user base include the following categories: sworn law enforcement, criminal intelligence, counter-terrorism, probation and correction, and other justice and public safety support missions. One of the key value propositions with the GFIPM approach is that users are effectively added "in bulk" when an IDP is brought into the federation instead of one user at a time. Even with the limited number of IDPs established during the pilot phase, three currently, more than 170,000 state, local, and federal users potentially have access to federation resources today.

One of the key specifications which has resulted as part of the project is the GFIPM metadata standard used to exchange attributes about users within the pilot federation. The metadata standard used within the federation is considered a draft and will undergo some refinement. However, it has been created over a multi-year development and vetting process with input from multiple information sharing organizations and domain experts, so it represents the thinking of many stakeholders throughout the justice community.

The project also produced a wealth of experience in the form of documented lessons learned about the process of implementing and leveraging GFIPM in a real-world information sharing system. The following points represent a very brief summary of the lessons learned during the GFIPM demo project.

- The basic GFIPM concept has proven to be viable, efficient and effective. It is possible to provide existing vetted users with cross-trust-boundary access to existing resources by leveraging existing user management mechanisms and providing resource owners with metadata attributes about users for access control purposes. The technology required to enable this functionality is sufficiently secure, user-friendly, and has no negative impact on user-perceived performance.

- The business case for joining a federation is straightforward. As with any interconnected network, early adopters will tend to benefit less than late adopters relative to the cost they expend to join; however, the value of a federation will increase for all participants with each additional user or resource. For IDPs, joining a federation depends on whether the federation can provide the IDP's users with a sufficient quantity and quality of resources that are appropriate and valuable to those users. For SPs, joining a federation depends on whether the federation can provide the SP's resources with a sufficient quantity of appropriate

users. When deciding how to join a federation, both IDPs and SPs can choose from among several business-level arrangements, including a direct connection, indirect connection through a broker, or indirect connection by means of inter-federation connectivity between two federations.

- User Identity metadata proved to be a very valuable part of the demonstration project. Participants found that, at least within the limited scope and number of participants in the demo project, it was possible to reach agreement on metadata attributes required for identification and authorization. Also, in most cases the required user metadata attributes were either already being collected and stored by agencies or could be derived as needed based on local policies.

- From both a technology and a cost perspective, it is feasible to federation-enable various types of legacy resources so that federation users can access them via the federation infrastructure. Many enablement techniques exist, each with specific advantages and disadvantages, and certain enablement techniques are better suited to certain resource types. One very useful federation enablement technique is to set up a GFIPM-aware proxy that interprets federation user metadata and implements part or all of the access control policy for federation users on behalf of a non-federation-aware resource.

- As the GFIPM project grows and matures, usability and user support will become increasingly important to the success of the project. Basic value-added services, such as a federation-wide registry and search capability ("GFIPM Google"), will become necessary tools for users. Also, federation-wide troubleshooting, help desk, and user training capabilities will become necessary as the user base increases and the range of available resources becomes larger.

- As it grows and matures, the GFIPM federation will require a formal governance structure as well as a basic operations management structure. A formal governance structure with participation from the federation membership will need to be established to develop and lifecycle manage federation policies and procedures. Also, ongoing operational support to carry out day-to-day processes and procedures related the federation will be required.

Having established the basic GFIPM concept is viable, implemented an operational GFIPM pilot federation, and developed an initial set of standards, tools, and documentation, the GFIPM initiative is focusing on the next phase of the project. This involves moving beyond a demonstration and pilot capability towards a wide-scale operational system for information sharing. To that end, the following activities are planned for the next phase of the GFIPM project.

- Establish of a Global GFIPM Delivery Team to serve as an active collaborative liaison between the GFIPM project and Global. This team will provide guidance

to the project and bring GFIPM deliverables, products, standards, and issues into the appropriate Global committees and processes.

- Establish a formal governance structure for the GFIPM federation to replace the existing ad hoc structure that participants established to facilitate the demonstration project.
- Update, validate, and vet GFIPM standards, including the SAML usage profile employed by GFIPM and the GFIPM user metadata model.
- Migrate from SAML 1.1 to SAML 2.0 and begin offering support for COTS products within the federation.
- Provide tools, assistance, and documentation to reduce the time and cost required for future participants to join the federation.
- Extend GFIPM concepts and standards to support the Global Justice Reference Architecture, which addresses the system-to-system or service-oriented architecture (SOA) use case.
- Expand the pilot federation by adding new participants, users, and resources.
- Establish and test inter-federation data exchanges with other federations, including the DOJ Law Enforcement Information Sharing Program (LEISP) federated identity management pilot project.

1 Background

As identified by the Global National Criminal Intelligence Sharing Plan (NCISP), the Markle Report, the 9/11 Commission, and Executive Order 13356, there are many recognized sensitive but unclassified (SBU) networks and information systems that support substantial investments in technology, governance structures, and trust relationships throughout the country and which are not interoperable. The sharing of justice, intelligence, and terrorism information is critical to rule of law and protection of the nation. One of the primary impediments to secure electronic information exchange and system interoperability is that of identity and privilege management – namely, making sure the right individuals have access to the authorized resources they need regardless of where they reside in the enterprise. This challenge is not limited to the intelligence community; it is applicable to justice and public safety at large.

Today justice practitioners must subscribe to multiple registration processes and manage multiple security mechanisms and passwords to get access to the resources they need. With an increasing demand for secure information sharing between federal, state, and local agencies, this approach is becoming increasingly unmanageable from a security and administrative perspective. It is also frustrating to users, costly, and unable to scale to meet the information sharing vision.

Overcoming the factors impairing systems interoperability is simpler if the entities involved in information sharing transactions have established trusted working relationships with each other. The problem then becomes a technical challenge that emphasizes the need for an identity and privilege management service that can be used to apply authentication and access controls within the disparate systems and networks desiring to make their resources "sharable".

1.1 Problem Statement

Achieving secure information sharing is imperative to improve the operational efficiency and effectiveness of organizations and agencies involved in carrying out their respective responsibilities to preserve, protect, and promote the public safety of the nation and its citizenry. A national consensus has emerged around the recognition of an urgent necessity to break down traditional barriers between justice and public safety organizations and make appropriate information more readily available to agencies with legitimate needs for information in support of improved criminal justice practices.

Throughout the country there are many recognized sensitive but unclassified (SBU) networks and information systems that support substantial investments in technology, governance structures, and trust relationships. Two of the primary impediments to electronic information exchange and systems interoperability between these networks and information systems today are incompatible technologies and policies specific to identity, authentication, and authorization. Factors that affect and often impede systems interoperability include the following:

- Numerous autonomous agencies;
- Multiple trust domains;
- Heterogeneous environments;
- Varied governance structures;
- Significant investment in legacy environments;
- Inconsistent or non-existent security policies and procedures;
- Disparate and incompatible security mechanisms.

Any candidate solution to the problem of information sharing for law enforcement must consider the entire list above. For each item in the list, the candidate solution must either address the issue directly, or at least not preclude the item from being addressed otherwise by another solution.

1.2 Project Initiation and Funding

Under the auspices of the Global Information Sharing Initiative, the Global Security Architecture Committee (GSAC) was established to develop standard security architecture in furtherance of the recommendations in the National Criminal Intelligence Sharing Plan. The concept of a federation based on emerging standards tailored for the justice and public safety domain received strong consensus within the group and has received growing interest from several federal, state, regional, and local agencies. As a result, an issue-focused working group consisting of committee members representing existing operational state and local intelligence and anti-terrorist information systems was formed to discuss strategies for moving forward. In addition, a demonstration project titled the Global Federated Identity and Privilege Management (GFIPM) Security Interoperability Demonstration was initiated. Although there are many organizations in the public safety domain that need and want to overcome the problem at hand, the programs that participated in the initial phase of the GFIPM project include the Criminal Information Sharing Alliance Network (CISAnet), the Regional Information Sharing Systems Network (RISSNET), and the Pennsylvania Justice Network (JNET). This initiative was funded through a partnership between the Bureau of Justice Assistance (BJA), National Institute of Justice (NIJ), and the Department of Homeland Security (DHS) Office of Chief Information Officer (OCIO). Funding for project management, engineering and development, and technical assistance was provided to the Georgia Tech Research Institute (GTRI) through grants from BJA and DHS. Funding to the three GFIPM participants was provided directly to the participants from BJA in separate grants.

1.3 Project Objectives

The focus of the demonstration project was to: (1) achieve a “quick win”, (2) capture “lessons learned”, and (3) lay a foundation of experience on which participants, stakeholders, and sponsors can assess the value and feasibility of a security and information-sharing architecture based on federated identity and privilege management. Specific objectives included the following.

1. Demonstrate that registered subscribers of one federation participant can access the web-based resources of another federation participant without the requirement to register with more than one federation participant.
2. Demonstrate that federation participants can retain control over their resources (applications, databases, etc.) and make local dissemination and access control decisions based on a shared set of standard subscriber attributes.
3. Demonstrate single sign-on (SSO) across federation web applications.
4. Demonstrate federated authentication and authorization between disparate local technologies and vendor implementations.
5. Establish a baseline of common attributes for identity and authorization assertions.

All of the above initial objectives have been successfully accomplished. As a result of this success, a number of lessons learned (Section 7.0) and next steps (Section 8.0) have been identified to move from this initial demonstration capability to a fully specified operational GFIPM framework supporting the secure information sharing needs of federal, state, local, and tribal agencies.

2 Federated Identity and Privilege Management Concept

This section introduces the concept of Federated Identity and Privilege Management (FIPM) and describes why it is powerful technology for implementing an information sharing capability in domains such as law enforcement.

2.1 An Example of the Problem

Suppose that two law enforcement information networks, such as CISA and JNET, decide to begin an information sharing program with each other to provide their users with access to more information. For simplicity, also suppose that each organization has 100 users that need access to the new information that is to be shared. Without FIPM technology, it would be necessary for CISA to create and manage a user accounts for every user from JNET, so each JNET user could access the available CISAnet resources. Similarly, JNET would need to provide an account for every user from CISAnet. Multiple practical challenges would arise in this scenario. For example, the IT staff for each organization would need to manage user accounts for all 100 of its own users, plus the 100 users from the other organization, for a total of 200 user accounts – essentially doubling the resources required for user account management. This includes items such as the cost and time required for vetting new users, issuing credentials (e.g. passwords, certificates, etc.) to new users, and keeping user account information (including personal information such as phone numbers, as well as various roles and permissions) up-to-date. It also may include costs related to deploying and managing authentication infrastructure for the new users. In addition, each user would need to keep track of his/her login credentials for two organizations. This places a credential management burden on users, and can lead to undesirable credential management practices, such as writing passwords on post-it notes, choosing easy-to-guess passwords, or failing adequately secure hardware authentication tokens when they are not in use.

Despite the undesirable consequences that it can cause, this approach might be acceptable as long as only two organizations are involved in the information sharing program. But what would happen if five – or even fifty – additional organizations wanted to participate in this program? And what would happen if each organization had not 100 users, but 1,000 or 10,000? The benefits of the information sharing program would quickly begin to pale in comparison with the costs, as the users and IT departments within each organization struggle to manage the complexity of the system. Clearly, for this type of wide-scale information sharing program to work, there must be a better way for IT staff and users to cope with the increased complexity of a multi-organization information sharing environment. Federated identity and privilege management (FIPM) provides a solution to this problem.

2.2 The FIPM Concept

Section 2.1 provides a classic example of the type of environment in which traditional user account management practices begin to break down. Users and administrators begin to suffer as the size and complexity of the information sharing environment increase and expand across multiple organizations. Fortunately, there is a solution to this problem. ***Federated Identity Management*** (FIM) consists of the technologies, tools, and techniques that allow organizations to reuse their members' digital identities with peer and partner organizations in an arrangement called a ***federation***. A federation is a group of two or more trusted partners with business and technical agreements which allow a user from one federation partner (participating agency) to seamlessly access resources from another partner in a secure and trustworthy manner. FIM provides a standardized method for agencies to provide information services to trusted users that they do not directly manage. The identities from one enterprise domain, or ***identity provider***, are granted access to the services of another enterprise, or ***service provider***.

The GFIPM project leverages and extends the concept of Federated Identity Management into the broader concept of ***Federated Identity and Privilege Management*** (FIPM). FIPM encompasses all of the features and capabilities of FIM, as well as a new set of capabilities not found in the basic FIM concept. The distinguishing feature of FIPM is that it incorporates the idea of a rich metadata model about users. More specifically, in FIPM a well-defined set of trusted metadata attributes about locally authenticated users can be securely exchanged between identity providers and service providers. These attributes are also authoritative, i.e. they come from a trusted third party (often called an ***attribute authority***) and are not self-attested by the user.[1] The primary purpose of metadata in FIPM is to facilitate the integration of legacy applications into a federation by helping to meet the applications' usage requirements for federation users. Most legacy applications have basic usage requirements that fall into the following categories:

[1] The concept of an entity's ***authority to assert*** a metadata attribute on behalf of a user is of critical importance to the challenge of appropriate sharing of information within the justice community. This issue is of primary importance to the DHS Project on Authority-Based Access Control (ABAC). See Appendix A for more information about the ABAC project.

- ***Terms of Use*** – The application may require that a user agree to specific terms of use prior to using it.
- ***Provisioning*** – The application may require that a user register a local account with it before using it.
- ***Inter-Session Persistence*** – The application may need to maintain state information about a user from one session to another.
- ***Identification*** – The application may need to know the user's identity at all times while the user is using it.
- ***Access Control*** – The application may impose certain access restrictions based on some combination of the user characteristics, such as: rank, certifications, role, or some other important personal characteristics.
- ***Auditing*** – The application may log all actions performed by a user in an audit log for review, compliance, etc.
- ***Personalization*** – The application may need to maintain miscellaneous personal data about a user for the purpose of delivering certain features. For example, locality information would help the application deliver a list of alerts or bulletins that specifically pertain to a user's region or locality.

The following subsections present more detail about specific aspects of the FIPM concept.

2.2.1 Identity Provider (IDP)

There are three basic components of a FIPM system. The first is an ***identity provider*** (IDP). An IDP is the authoritative entity responsible for authenticating an end user and asserting an identity for that user in a trusted fashion to trusted partners. The identity provider is responsible for the typical identity management life-cycle functions such as user vetting, credentialing, collecting, validating, and maintaining user information, provisioning, and password/token management. These responsibilities may be fulfilled with existing locally accepted security mechanisms and tools. In the passport illustration, a citizen's home government is the identity provider responsible for validating the true identity of the citizen. There can be an arbitrary number of IDPs in a federation, and each IDP can manage an arbitrary number of user accounts, subject to practical system constraints. An IDP is sometimes called by alternate names, such as ***credential service provider*** (CSP); however, this report consistently uses the terms "identity provider" or "IDP".

2.2.2 Service Provider (SP)

The second key component of a FIPM system is a ***service provider*** (SP). An SP is responsible for managing access to applications, services, and other resources used by federation users. To do this, it relies on IDPs to assert information about users, leaving the SP to manage access control and dissemination based on the trusted set of attributes it receives for each user. There can be an arbitrary number of SPs in a federation, and each SP can manage an arbitrary number of resources, subject to practical system constraints.

2.2.3 Common Metadata Attribute Language

The third noteworthy component in a FIPM system is a common attribute language that can be used between IDPs and SPs for the purpose of communicating information about users. The separation of user account management and resource management into two different components (IDP and SP) represents a fundamental paradigm shift away from a centralized system administration model (in which a single system manages both user accounts and resources) and towards a more distributed, federated management model. Separating the responsibilities of an IDP from those of an SP can provide some valuable benefits, as have already been discussed. This separation also comes at a price, however. In a centralized model, information about user accounts is available in a local user account database and can be used as needed for the purpose of user identification, access control, and auditing. Also, in a centralized model, applications tend to operate under implicit assumptions that are true because of local policies or out-of-band processes. For example, consider a local application for which only sworn law enforcement officers are permitted to gain access. If the local account provisioning processes or policies for such an application allow accounts to be provisioned only for sworn officers, then the application itself can assume that each user is a sworn officer. But in a federated model, where the application cannot implicitly assume that all users are sworn officers, such implicit assumptions are no longer acceptable. Instead, the critical information about each user must be made explicit. This is why FIPM system requires a common attribute language.

The GFIPM concept is based on a set of well-defined attributes called ***metadata. Metadata is*** associated with an authenticated subject (typically a user) and passed from IDPs to SPs in trusted Security Assertion Markup Language (SAML) assertions. IDPs are responsible for authenticating their subscribers and constructing SAML attribute assertions based on information that they can either collect and maintain directly or access via a trusted relationship with one or more attribute authorities. SPs provide business logic which can implement their user identification, access control, and auditing policies based on a trusted set of attributes received from an IDP.

The GFIPM metadata initiative is working on the task of designing a standard set of attributes for the purpose of implementing this privilege management concept. These attributes contain information which can be used to identify the end user, as well as the user's contact information, organizational affiliations, title and job function, the strength and quality of the user's electronic identity, and the strength of the authentication method used to authenticate the user to his IDP. Additionally, a set of authorization context attributes are defined based on some general data categories (e.g. Counter-Terrorism, Intelligence, Criminal History, etc.) These attributes are asserted by IDPs based on the user's job function and a minimal set of well-understood eligibility requirements. For example, it is expected that an IDP will assert an attribute for access to a category of information on a user's behalf if and only if the user has access to that category of information within his/her local system or agency. A federation-wide MOU would establish the minimal criteria (eligibility requirements) for an IDP to assert a given attribute for one of its subscribers. SPs use these attributes as a contributing piece of trusted information in making access control and dissemination decisions. Note that this

model does ***not*** prescribe a strict set of privileges that must be honored by all SPs federation-wide; in the GFIPM model, SPs maintain complete control over their access control policies at all times, and may make any access control decision regardless of what information is presented by an IDP on a user's behalf. It is expected that as the GFIPM initiative matures, additional metadata attributes and user profiles will emerge to address specific business and policy needs.

2.3 The FIPM Value Proposition

As illustrated in Section 2.1, the cost and complexity of identity administration in today's cross-organization information sharing environments are primarily due to a single cause: providing access to a user for a service or an application means giving the user an account within the service or application-specific repository. The fundamental practice of creating and managing user accounts leads to various administration, single sign-on, and compliance issues. The basic federated identity management (FIM) concept addresses these issues elegantly, and also provides other valuable benefits, as summarized in Table 1. [2]

FIM Benefit	Description
User Convenience	Users can access multiple services using a common set of credentials, making it easier to sign on and access applications and to manage account information.
Interoperability	By specifying the security standards and framework, applications can adopt security profile specifications for authentication and authorization processes.
Information Sharing	Federation facilitates information sharing about an individual's identity by reducing the overall work required to maintain connections and reduce the friction among multiple domains.
Privacy	Federated domains can reduce the propagation of personal identity information, reduce the redundant capture and storage of personal identity information, and depersonalize data exchanges across domains.
Security	Federations can improve the security of local identity information and data in service provider and service consumer applications.

Table 1: Summary of Benefits Provided by the Basic FIM Concept

As an extension of FIM, the federated identity and privilege management (FIPM) concept provides all of the benefits described in Table 1, along with the following additional benefits that basic FIM does not provide. Note that all of the FIPM-specific benefits arise directly out of the sharing of a rich set of trusted metadata about users.

[2] From Mike Neuenschwander and Dan Blum, *Federating a Distributed World: Asserting Next-Generation Identity Standards*, Version 1.0, April 15, 2005, Burton Group.

1. ***Level of Assurance for Identity and Authentication*** – Trusted metadata about a user can contain information about that user's identity, such as which identity-proofing techniques were used when the user established his account with his IDP. In addition, the metadata can contain information about the authentication method that the user employed when he authenticated with his IDP for the current session. Each of these data points can be valuable for applications that require a minimum standard for identity proofing and/or authentication of users.
2. ***Dynamic Account Provisioning*** – For applications and services that require each user to have a local account, FIPM can facilitate the process of account registration and provisioning via trusted metadata about the user. This benefit is significant because it can eliminate the need for manual intervention and/or out-of-band communication between the user and the application owner during the account provisioning process.
3. ***Dynamic Account Updates*** – In addition to enabling dynamic account provisioning, FIPM can also help ensure that important changes to user data (such as a change in the user's rank or contact information) are made available from IDPs to applications in a timely manner, thereby helping to increase the likelihood that a user's local account within an application is fresh and up-to-date.
4. ***Federated Privilege Management*** – The FIPM concept can facilitate the decision-making process of applications with regards to access control. If trusted metadata about a user contains information about that user's certifications, job function, role, or privileges within his local organization, then applications in the federation can use that information to inform their decisions about which permissions they should grant to the user.
5. ***Rich Audit Logs*** – FIPM can provide applications with metadata containing extensive personal contact information about users, including a user's phone number, physical location, supervisor's name, etc. All of this information can be used within applications to build very comprehensive audit logs containing not only the user's identity and actions, but other relevant information about the user as well. Some applications may require this additional information for the purpose of compliance with local policies or laws.
6. ***Personalization of the User Experience*** – Information contained in metadata can be used to improve the user's experience when using an application. Specific attributes, such as the user's physical location or his organizational memberships, may be useful for certain applications that display information in a localized or regionalized manner. It may not be possible to foresee all of the ways that metadata can be used for personalization at this time; however, it is clear that metadata can have value not only for cost savings and simplification of user account management, but also for the user experience.

To summarize, the FIPM concept provides all of the benefits of basic FIM, plus several additional benefits that FIM does not provide, to deliver significantly more value to federation members than they could derive through FIM alone.

2.4 GFIPM: Applying FIPM to the Justice Community

Sections 2.2 and 2.3 describe the basic FIPM concept and discuss the benefits that are made possible because of it. The GFIPM project sought to leverage the FIPM concept for the benefit of the justice community, as follows:

1. Connect existing law enforcement users to a federation infrastructure via IDPs, leveraging the existing user account vetting procedures, authentication systems, and LDAP directories.
2. Connect existing web-based law enforcement applications and data resources to the federation infrastructure via SPs.
3. Leverage existing FIM standards, plus the FIPM concept of rich metadata about federation users, to connect users with applications while providing benefits such as single sign-on, dynamic account provisioning, federated privilege management, enhanced auditing capability, and personalization of the user experience, all in the context of a secure, standards-based federation environment.

The remainder of this report describes in detail the process of carrying out the actions listed above, including the basic strategy, project participants, technical architecture, project execution timeline, current federation structure, lessons learned, conclusions, and next steps.

3 Project Overview

This section begins to describe the GFIPM demonstration project in more detail. First it reviews the basic approach taken, including specific technologies and choices that were made. Second, it describes all of the major participants in the project. Finally, it identifies all of the critical assumptions and constraints that impacted the project.

3.1 Approach

The objective of the GFIPM demonstration project was to develop and test an identity and privilege management service that could be used to apply authentication and access controls by disparate systems and networks desiring to make their resources “sharable”. The desired outcome of the project was to demonstrate a universal (implementation-independent and non-vendor specific) mechanism through which trusted metadata assertions could be shared and used to apply authentication and access controls. The following key points characterize the basic approach employed by the project:

- Use of existing vetted participant subscribers and associated information stores, security and authentication mechanisms, policies, and production resources. Explore and validate the GFIPM concept in a real production environment. Identify roadblocks and collect lessons learned in working through the technology, business, and policy related challenges.
- Focus initially on the law enforcement domain and the associated information sharing security requirements, resources, and user base.

- Focus initially on the user-to-application (web browser to web application) use case rather than system-to-system or other use cases.
- Develop, agree upon, and use a standard set of attributes with well-defined semantics and an XML syntax based on the NIEM standard.
- Leverage the Internet communications backbone for all network traffic.
- Use standard browser Secure Socket Layer (SSL)\Transport Layer Security (TLS) for session establishment and encryption. Require no special software on the client side.
- Leverage the existing Security Assertion Markup Language (SAML) federated identity standard, and an open source implementation of SAML, to minimize software development and implementation costs for participants. (Note that none of the participants had a SAML federated identity capability in place at the time the project began.)
- Extend and configure existing open source products to provide the base GFIPM federation infrastructure. Provide simple well-defined integration points for IDPs and SPs. Participants would use this integration points to enable their existing legacy infrastructure and join as an IDP, SP, or both.
- Allow SPs to make all authorization decisions for their resources based on metadata passed in SAML assertions.

Figure 1 provides an illustration of the basic approach of the GFIPM demo project.

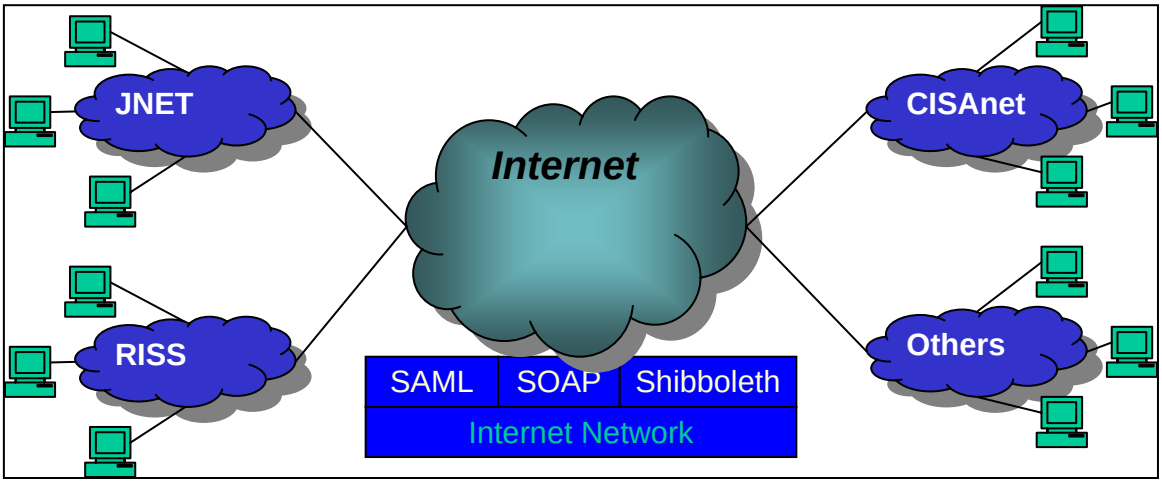

Figure 1: Basic Approach of the GFIPM Demonstration Project

Additional discussion regarding each of these key aspects of the approach is provided below.

3.1.1 Use of Existing Vetted Participant Subscribers and Information

To fulfill the project objectives it was necessary to collect lessons learned and ensure viability of the GFIPM concept with existing vetted federation participant subscriber bases. Each of the three demonstration participants (CISA, JNET, and RISS) selected a small initial subset of users from their existing subscriber base to participate. These subscribers had already been vetted in accordance with their local policies and procedures

by their respective agencies/systems. Information (attributes) about the users had already been collected, stored, and maintained in some database or directory. The demonstration project leverages this subscriber vetting and information for use across the federation – first for the initial set of subscribers, and ultimately for the entire federation participant subscriber base. One of the key objectives of the project was to evaluate the acceptance and suitability of subscriber vetting and information stores as well as identify shortfalls and rough level of effort for use in the federation. One of the challenges that emerged due to this decision was participants needed to arrange for selected users to actually use the GFIPM federation systems after they were operational. This required that participants create and execute detailed plans regarding the practicalities of rolling out a new service, such as awareness-building, training, etc. Numerous lessons learned have resulted from the choice to use real users in the project. More detailed information about this topic is available in Section 7 of this report.

3.1.2 Use of Existing Participant Authentication Mechanisms and Policies

Leveraging existing account management mechanisms in the project also presented some challenges. The most significant challenge was that project participants had to connect their existing user authentication mechanisms to their GFIPM IDP middleware, so users could leverage their existing login credentials in the federation. Another challenge for participants was to connect their existing user databases (typically LDAP directories) to their GFIPM IDP software, so that the IDP software could leverage existing data about users for the purpose of constructing GFIPM assertions.

3.1.3 Use of Production Web-Based Resources

Using production resources in the project presented some challenges during the project. One set of challenges stems from the fact that production resources contain sensitive data for which there is policy in place to govern its dissemination. These policies address who the data may be released to, as well as the conditions for release. Accordingly, the process of bringing these resources online involved much care and due diligence, at both the policy and technical level. Another set of challenges regarding the use of production resources arose out of the decision to use a common metadata model for privilege management within the federation. None of the resources that were brought online in the federation had been designed with the concept of metadata-based privilege management in place. This led to many important lessons learned regarding both policy decisions and technical strategies. Many critical lessons learned have been captured for this topic. They are available in Section 7 of this report.

3.1.4 Use of Security Assertion Markup Language (SAML) version 1.1

Sections 2.2.1 and 2.2.2 of this report introduce two fundamental concepts in federated identity: the identity provider (IDP) and the service provider (SP). An IDP handles user management, including authentication and user metadata. An SP handles the management of access to protected resources based on information given to it by an IDP. To perform their respective roles, an IDP and an SP need to communicate with each other, and the most widely accepted standard through which this communication occurs is the Security Assertion Markup Language (SAML). SAML was developed by the

Security Services Technical Committee (SSTC) of the Organization for the Advancement of Structured Information Standards (OASIS), and it is an XML-based framework for communicating user authentication, entitlement, and attribute information. As its name suggests, SAML allows business entities to make assertions regarding the identity, attributes, and entitlements of a subject (an entity that is often a human user) to other entities, such as a partner company or another enterprise application.

The most recent version of the SAML specification is version 2.0. SAML 2.0 was approved by OASIS in March 2005; however, at the time that the GFIPM project began (in late 2005), SAML 2.0 was not yet widely supported by commercial vendors of federated identity management middleware products. None of the project participants had a pre-existing SAML capability in place when the project began, and one of the project's goals was to minimize costs to federation participants (and avoid the lengthy procurement cycles that are typically required to obtain commercial software). Therefore, the participants chose to use a free, non-commercial implementation of SAML for the demo project federation. At the time that the project began, there was only one mature open source SAML middleware implementation: Shibboleth. But Shibboleth version 1.3 (which was the latest available version of Shibboleth as of late 2005) does not contain support for SAML 2.0. For these reasons, the participants chose to begin the GFIPM project using SAML 1.1 (which is supported by Shibboleth 1.3) and eventually migrate to SAML 2.0 at the appropriate time. At the time of this writing (June 2007), the GFIPM demo federation is still using Shibboleth 1.3 and SAML 1.1; however, it is anticipated that the GFIPM federation will move to SAML 2.0 in the very near future. The topic of Shibboleth is covered in more detail in the next section.

3.1.5 Use of the Shibboleth Open Source SAML Middleware

As was briefly discussed in the previous section, project participants chose to use software from the Shibboleth project as the federated identity endpoint middleware during the GFIPM project. This decision was motivated by several factors, some of which have been previously noted.

1. It was critical to help each federation participant develop a SAML capability quickly and inexpensively, and as a freely available open source package, Shibboleth was ideal for this purpose.
2. It was important to ensure the GFIPM project would remain vendor-neutral and not dependent on one or a few commercial software vendors. The intent was to standardize on open standards (e.g. SAML), but not on specific implementations of those standards.
3. There were concerns about operational interoperability between the participants' middleware implementations. It is well-known in the SAML community that earlier versions of SAML (up to and including version 1.1 – the latest SAML version that was widely supported when the demo project began) have not always been implemented in a standards-compliant fashion. Rather than asking each participant to choose its preferred SAML implementation and then spend valuable project time and resources trying to achieve low-level protocol interoperability between them, it was a better use of time and resources to standardize the entire

federation on one SAML implementation and move forward from a common starting point across the federation.

The remainder of this section provides a basic overview of the Shibboleth project and software.

The Shibboleth project is run by the Middleware Architecture Committee for Education (MACE), which is a part of the Internet2 Middleware Initiative. Its focus is to develop architectures, policy structures, practical technologies, and an open source implementation to support inter-institutional sharing of web resources subject to access controls. The project is geared primarily towards applications in the domain of higher education; however, the Shibboleth software itself does not contain any features that preclude its use in other federated identity application domains.

The Shibboleth software's greatest asset is its ability to insulate administrators, developers, and end-users from the low-level details of SAML, thereby making the process of developing and deploying a Shibboleth-based federated identity system relatively simple and fast. Shibboleth delivers this value via three powerful and very flexible system integration points: the single sign-on (SSO) integration point, the attribute authority integration point, and the protected resource integration point. Minimal protocol-level knowledge of SAML is required to deploy a Shibboleth system. All three Shibboleth integration points are described in more detail in Section 4 of this report.

3.1.6 Use of the Internet as a Backbone for Connecting Participants

One of the primary value propositions for federated identity is the low cost of participation and ubiquitous access gained by leveraging existing mechanisms and infrastructure. To minimize cost and maximize accessibility, the GFIPM project leveraged use of the Internet for inter-participant communications, including the following types of transactions:

1. Browser-to-SSO data transfer for the purpose of user authentication with an IDP;
2. Browser-to-SP data transfer for the purpose of accessing protected resources via web applications and portals;
3. IDP-to-SP data transfer for the purpose of sharing certain SAML assertions;
4. IDP-to-Browser-to-SP data transfer for the purpose of sharing other SAML assertions.

All transactions were secured using standard browser Secure Socket Layer (SSL)\Transport Layer Security (TLS) for session establishment and encryption. This required no special software on the client side keeping operational cost and support issues to a minimum. Existing participants' private intranets and back-end networks remained in place and were used by participants to route and respond to federation requests as necessary.

3.1.7 Implementation of a Common Metadata Model for Assertions

As described briefly in Section 2.2.3, project participants developed and used a well-defined standard metadata model for passing information about users from IDPs to SPs. The decision to use a common metadata model for this purpose takes the scope of the GFIPM demo project beyond merely federated identity management, because it provides for the ability to make a rich set of metadata about a user available to each application accessed by that user. The GFIPM concept requires a well defined set of metadata attributes to be shared in assertions between IDPs and SPs for the following purposes.

1. ***Identification/Authentication***—Attributes are needed to communicate the identities of end users and their associated authentication contexts. Who is the end user? How did his IDP vet his identity during the account creation process? And how did he authenticate?
2. ***Dynamic Account Provisioning and Updates*** – For an application that requires each user to register for a local account, attributes captured and delivered by IDPs can facilitate the account provisioning and maintenance process. Does this user already have a local account? Has any of the information stored about him changed recently?
3. ***Privilege Management***—Attributes captured by identity providers (IDPs) can assist service providers (SPs) in making authorization decisions. What certifications, clearances, job functions, local privileges, and organizational affiliations are associated with the end user that can serve as the basis for authorization decisions?
4. ***Audit***—Attributes are required for the purposes of auditing systems, system access, use, and legal compliance of data practices.
5. ***Personalization***—Attributes can enable local systems to feature specialized services, regionalization, or special-interest characteristics of their local software (e.g., regional news or alerts, SIG information, display, and tool settings or preferences).

The GFIPM metadata did not exist prior to the start of the demonstration project. Throughout the duration of the project, the participants collaborated to develop the metadata standard which is currently in place today. More information about the GFIPM metadata can be found in Section 4.6 of this report. In addition, the GFIPM Metadata Package 0.4 (see Appendix B) contains a complete description of the process by which the GFIPM metadata model was developed.

3.2 General Descriptions of Project Participants

The GFIPM demonstration project included four participants: CISA, JNET, RISS, and GTRI. The first three participated in the role of federation members, while GTRI's role was to provide project leadership, technical direction, and engineering services as needed. This section contains a general description of each organization that participated in the GFIPM demonstration project. For additional information about the federation members (CISA, JNET, and RISS), including implementation details about their federation

infrastructure components and the types of users and resources that they brought to the federation, please see Section 6.

3.2.1 Criminal Information Sharing Alliance (CISA)

The purpose of the Criminal Information Sharing Alliance (CISA), is to operate, maintain, enhance, and expand the Criminal Information Sharing Alliance network (CISAnet). CISA was established with Department of Defense funding to improve the collection, use and sharing of law enforcement information among the states of Alabama, Arizona, California, Georgia, Idaho, Louisiana, Mississippi, New Mexico, Oklahoma and Texas, the El Paso Intelligence Center (EPIC) and the six Regional Information Sharing Systems (RISS).

Current open source intelligence estimates that over 73% of all cocaine and 52% of all illegal drugs entering the United States pass through US/Mexican border. One of the key principles of the National Drug Control Strategy is an aggressive and coordinated law enforcement effort. This domestic effort includes border controls, investigations of drug trafficking organizations, and the collection and dissemination of drug law enforcement information. CISAnet supports the National Drug Control Strategy by improving the collection, use and sharing of criminal information among federal, regional, state, and local law enforcement agencies. It enhances secure information sharing while guaranteeing the privacy rights of individuals and organizations. CISAnet leverages automation to provide integrated technical support to the front line in the nation's fight against illegal drug activity and violent crime. In addition, CISA supports federal, regional, state and local multi-jurisdictional task forces which play a significant role in reducing drug availability.

CISA provides a multi-state model for secure law enforcement information sharing. It allows all participants to maintain control of their own data; preserves participant investments in state systems, improves officer safety and efficiency; and reduces cost for future participants by using standards-based, open systems technology. As such, the system was designed and has been implemented to accept the data sources of additional states, federal agencies, and other participants at a minimal cost.

The Criminal Information Sharing Alliance (CISA), incorporated in the state of Maryland, was established as a tax-exempt organization under IRS Code Section 501(c)(3). In addition to supporting the daily operation, maintenance, and expansion of CISAnet, CISA educates Department of Public Safety, Department of Justice, and other law enforcement personnel on the efficiencies of CISAnet and the importance of information sharing. The ability to collect, use and share information directly affects the successful prosecution of law enforcement strategies at all levels. The timeliness of this information is critical, not only to the success of a particular investigation, but to a law enforcement officer's safety. CISA's use of industry and government standards has reduced overall costs (R&D, O&M, etc.) and the government's burden by reducing the development of multiple parallel efforts and eliminating more costly proprietary systems. CISA, working with the Departments of Defense and Justice will assist in the

coordination and integration of other law enforcement information sharing efforts already in progress.

CISA's participation in the GFIPM demo project was in the role of both an identity provider (for connecting existing CISA users to the federation) and a service provider (for connecting existing CISA resources to the federation). CISA's responsibilities in the project included implementation of a CISAnet identity provider connector, implementation of a CISAnet service provider connector, reviewing and providing input to the demonstration scenario, and technical documentation as required. GTRI (see Section 3.2.4) serves as the technical integrator for CISAnet. Funding for CISA's participation in the project was provided by the U.S. Department of Justice.

3.2.2 Pennsylvania Justice Network (JNET)

The Pennsylvania Justice Network (JNET) is the Commonwealth's primary public safety and criminal justice information broker. JNET's integrated justice portal provides a common online environment for authorized users to access public safety and criminal justice information. This critical information comes from various contributing municipal, county, state, and federal agencies.

JNET's secure web portal provides access to over 31,000 practitioners throughout the Commonwealth's sixty-seven counties as well as federal and state agencies. JNET provides these practitioners with the ability to conduct secure investigations in a web-based environment.

Access to JNET's secure web portal is dependent upon policy, secure connectivity, and role-based entitlements. Practitioners with proper security credentials and authority have access to information that historically took days or weeks to obtain through legacy, paper-driven, and sometimes manual-based business processes.

One-time data entry has improved the effectiveness of participating agencies, and has significantly improved data accuracy throughout the Commonwealth's criminal justice system. Information entered into a records management system at the onset of an investigation can now follow the offender throughout their criminal justice tract. As offenders pass through the gateway of justice through post-sentencing supervision, offender information flows in concert with the offender's progression.

JNET's secure messaging infrastructure allows for the secure transfer of information between agency systems and users. This data exchange and event messaging model provides stakeholders with the ability to maintain ownership and control of their data systems, and has elevated county and state agency data availability. This currently includes participating in electronic data exchange, and subscribing to real-time event messaging services.

JNET's participation in the GFIPM demo project was in the role of both an identity provider (for connecting existing JNET users to the federation) and a service provider (for connecting existing JNET resources to the federation). JNET's responsibilities in the

project included implementation of a JNET identity provider connector, implementation of a JNET service provider connector, reviewing and providing input to the demonstration scenario, and technical documentation as required. Funding for JNET's participation in the project was provided by the U.S. Department of Justice.

3.2.3 Regional Information Sharing Systems (RISS)

The Regional Information Sharing Systems (RISS) Program is an innovative, federally funded law enforcement support program composed of six Regional Intelligence Centers that coordinate efforts regionally, nationally, and internationally to:

- Share intelligence information;
- Promote public and officer safety;
- Respond to natural disasters;
- Identify, target, and remove criminal conspirators;
- Combat terrorist activity, illegal drug trafficking, organized criminal activity, criminal gangs, and violent crime.

On national-scope issues, the six regional Centers initiate joint, cross-Center efforts, coordinating and cooperating as one body.

Since its inception, RISS membership has grown to serve more than 7,300 law enforcement and criminal justice agencies representing over 750,000 sworn officers. Membership includes local, county, state, federal, and tribal law enforcement member agencies in all 50 states, the District of Columbia, U.S. territories, Australia, Canada, and England. For most of the services that RISS provides, agencies must join their regional RISS Center through an application process established by the Center. RISS is making a growing list of information services (e.g., RISS Automated Trusted Information Exchange™ (ATIX) and The RISS National Gang Database (RISSGang)) available to a broader community of law enforcement, emergency services, and critical infrastructure constituents who will be vetted individually through a regional RISS Center.

Traditional support services provided by RISS to its law enforcement member agencies are:

- Information sharing resources;
- Analytical services;
- Loan of specialized investigative equipment;
- Confidential funds;
- Training classes;
- Technical assistance.

RISS operates a secure sensitive but unclassified nationwide communications and information sharing intranet, known as RISSNET™, which employs Web technology to provide controlled access to a variety of sensitive information resources within the RISS

and other node agency environments to those individuals who have successfully completed the RISS identification and approval (vetting) process.

The RISS Automated Trusted Information Exchange™ (RISS ATIX) Program was created to provide secure interagency communications and information sharing resources for the exchange of law enforcement and public safety-related information among executive and official staff from governmental or nongovernmental entities involved with planning and implementing prevention, response, mitigation, and recovery efforts regarding terrorism, disasters, or other law enforcement and public safety strategic and tactical response efforts.

RISS's participation in the GFIPM demo project was in the role of both an identity provider (for connecting existing RISS users to the federation) and a service provider (for connecting existing RISS resources to the federation). RISS's responsibilities in the project included implementation of a RISSNET identity provider connector, implementation of a RISSNET service provider connector, reviewing and providing input to the demonstration scenario, and technical documentation as required. The RISS Office of Information Technology provides application development and integration support for RISS. Funding for JNET's participation in the project was provided by the U.S. Department of Justice.

3.2.4 Georgia Tech Research Institute (GTRI)

Georgia Tech Research Institute (GTRI) is the nonprofit applied research arm of the Georgia Institute of Technology in Atlanta, GA. GTRI is committed to creating solutions through innovation, on time and on budget. GTRI assists clients in federal, state, local and international government agencies, industrial firms, academic institutions and private organizations. More than 70 percent of GTRI's research personnel hold advanced degrees, and all are committed to an independent, unbiased approach to solving problems. GTRI also has a close association with colleagues at the Georgia Institute of Technology ("Georgia Tech") who can contribute additional talent and knowledge for meeting technological and engineering challenges.

GTRI has extensive experience in providing database software and information systems technical support to the law enforcement and justice communities. GTRI designed, built, and has maintained a secure law enforcement intelligence information sharing system for over 10 years. GTRI assists the State of Georgia criminal justice community and the Georgia Bureau of Investigation. It also directs and operates the Criminal Justice Science and Technology Center. Furthermore, GTRI participates in the GLOBAL Infrastructure and Standards Working Group.

For the last several years under sponsorship of the Office of Justice Programs (OJP), the DOJ Office of Chief Information Officer (OCIO), Department of Homeland Security, the FBI NDEX PMO, and guidance from the Global XML Structure Task Force, and NIEM PMO and governance bodies (NTAC, NBAC), GTRI has provided specialized technical support for the design, implementation, training, maintenance, and evolution of the Global Justice XML Data Model (GJXDM), the National Information Exchange Model

(NIEM), the Global Federated Identity and Privilege Management (GFIPM) Security Demonstration, the Law Enforcement Information Sharing Program (LEISP) Exchange Specifications (LEXS), and the FBI National Data Exchange XML Specifications. GTRI has participated in the development of the GFIPM concept and project since its inception, and has served in a project management and technical leadership role throughout the implementation of the current GFIPM federation. GTRI has led the design and implementation of all associated GFIPM specifications, tools, training materials, and GFIPM websites.

GTRI's participation in the GFIPM demo project was in the role of technical program management, integration, and assistance. GTRI's responsibilities in the project included development, maintenance, and tracking of project plan and deliverables; implementation/modification of core federation software components (middleware); conducting technical meetings and conference calls as required; establishment of test infrastructure, test plan, and facilitation of integration testing; development of technical documentation including interface specifications, presentations, and final report; providing integration support and assistance for all participants; facilitation of the collection, documentation, and reconciliation of identity and privilege attributes; and serving as the technical liaison to the GJXDM and NIEM initiatives. Funding for GTRI's participation in the project was provided jointly by the Department of Justice and the Department of Homeland Security.

3.3 Assumptions and Constraints

This section of the report describes key assumptions and constraints that have shaped and guided the GFIPM project.

3.3.1 Leverage NIEM

At the start of the GFIPM project, it was decided that the GFIPM metadata model would leverage the National Information Exchange Model (NIEM) content and architectural framework. NIEM is a Federal, State, Local, and Tribal interagency initiative providing a foundation for seamless information exchange. It leverages the data exchange standards efforts successfully implemented by the Global Justice Information Sharing Initiative ("Global") and extends the Global Justice XML Data Model (GJXDM) to facilitate timely, secure information sharing across the whole of the justice, public safety, emergency and disaster management, intelligence, and homeland security enterprise. Given the work and success of the GJXDM and NIEM data modeling efforts, it is a logical choice to leverage and reuse these specifications in describing the GFIPM metadata. The advantage of leveraging the NIEM specification is it inherently makes the GFIPM metadata model immediately more applicable to other domains and systems, rather than focused only on criminal justice users and systems.

3.3.2 No Sharing of Intelligence Information during Initial Phase

For the purpose of the initial phase of the GFIPM demonstration project, it was agreed upon by all participants that no intelligence information would be shared over the federation infrastructure The participants wanted to restrict the information shared

during the initial project phase to reduce the risk associated with sharing critically sensitive information prior to gaining confidence in the technologies and processes inherent to federated identity and privilege management as implemented in GFIPM.

3.3.3 Transition to SAML 2.0

As previously described in Section 3.1.4, SAML 1.1 was selected for use in the GFIPM demo project due to the limitations of SAML 2.0 support in the commercial and open source marketplace when the project began. But one of the primary assumptions at the start of the GFIPM project was that the GFIPM federation, if successful, would transition to SAML 2.0 as the SAML standard matured and gained traction in the community. The participants' ability to meet project objectives (as stated in Section 1.3) and capture valuable lessons learned (as described later in Section 7) were largely not dependent on the version of SAML used for this demonstration phase. As of June 2007, GFIPM participants are in the early stages of investigating the transition to SAML 2.0 and establishing the necessary infrastructure to support interoperability testing using SAML 2.0 COTS products and the GFIPM concept and standards.

3.3.4 User-to-Application Use Case

Three distinct use case scenarios have been identified by the Global Security Architecture Committee as applicable to information sharing between federal, state, regional, local, and tribal agencies. They are as follows.

1. System-to-System Connectivity, (e.g. Service-Oriented Architecture)
2. User-to-Application Connectivity (e.g. Browser-to-Web Service)
3. User-to-User/System-to-User Messaging (e.g. Instant Messaging)

Although each use case scenario is pertinent and must eventually be addressed within the concept of a federation, scenario #2 above (user-to-application connectivity) was chosen to be the sole focus of the GFIPM demonstration project. This decision was made for several reasons. First, the user-application use case required the least amount of integration. Second, all of the initial project participants had web applications in place that could be shared within a federation infrastructure.

4 Architecture

This section of the report documents the basic architecture of the federation at a technical level. It includes all major software components that were required to construct the pilot federation infrastructure, and it also highlights specific topics of concern that required significant effort or special consideration by participants while deploying the federation.

4.1 IDP Structure

Each IDP that was deployed in the pilot federation consists of several components. Figure 2 illustrates the IDP structure, and the details of each component in the figure are discussed in the following subsections.

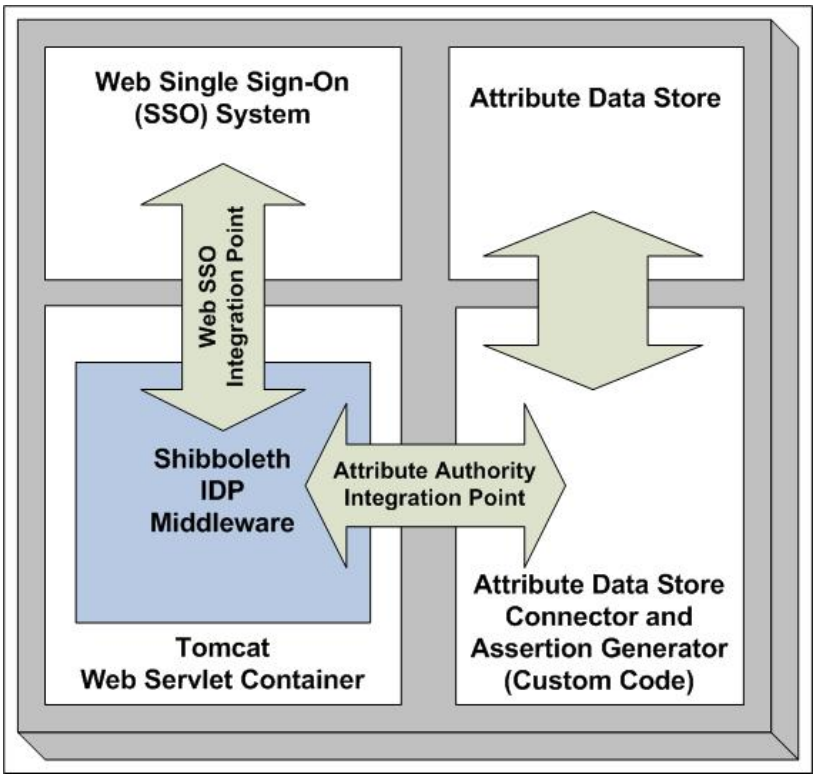

Figure 2: GFIPM Identity Provider (IDP) Structure

4.1.1 Shibboleth IDP Middleware Module

The Shibboleth IDP middleware, depicted by a blue box in Figure 2, is implemented as a Java servlet and runs within a web servlet container. It was developed by the Internet 2 project using the OpenSAML open source implementation of SAML. The Shibboleth middleware is not very useful as a stand-alone system, because it consists of a set of interfaces, called ***integration points***, which must be integrated with other system components for the IDP to work. (See Section 4.1.3 for more detail on these integration points.) The Shibboleth IDP module handles the processing of incoming SAML messages from SPs, as well as the creation of outgoing SAML messages to SPs. In addition, it manages signing and encryption of all outgoing SAML messages, as well as signature verification and decryption of all incoming SAML messages. The connection-level SSL/TLS encryption at the IDP is handled by the servlet container in which the Shibboleth IDP middleware runs. See Section 4.3 for details about the SAML protocols and messages used within the GFIPM federation.

4.1.2 Web Servlet Container

As discussed in the previous section, a web servlet container is required on a GFIPM IDP to run the Shibboleth IDP middleware module. Many such web servlet containers are available for use, but participants chose to use the freely available Tomcat open source servlet container. Tomcat was chosen for several reasons. First, Tomcat is fully compatible with the Shibboleth IDP middleware, i.e. it runs the Shibboleth code without any problems. Second, Tomcat supports connection-level encryption using both SSL and TLS. Third, Tomcat supports client certificate authentication of browsers with support

for certificate revocation lists (CRLs). This feature was important to several participants, as will be discussed in Section 6. Fourth, Tomcat runs on both of the major operating system platforms (Microsoft Windows and Red Hat Enterprise Linux) that GFIPM participants wanted to use for IDPs in the federation. Other servlet containers would have probably also met the needs of participants for the purpose of the GFIPM project; however, since Tomcat is very popular and it met the requirements of the federation, other servlet container alternatives were not investigated.

4.1.3 IDP Integration Points

There are two IDP integration points for any SAML 1.1 IDP. This includes not only Shibboleth IDPs, but any other IDP middleware package as well. One is the ***Single Sign-On (SSO) Integration Point***, and the other is the ***Attribute Authority (AA) Integration Point***. Both of these integration points are discussed in the following subsections.

4.1.3.1 Single Sign-On (SSO) Integration Point

The purpose of the Single Sign-On (SSO) Integration Point is to allow the Shibboleth IDP middleware to construct SAML authentication statements and send them to SPs to state that a user has authenticated properly within the federation. The integration point consists of a secure (HTTPS) URL that is to be visited by a browser during the SSO process as part of the Shibboleth/SAML Web Browser SSO Profile. (See Section 4.3 for more detail about this.) When browser visits the secure URL, it triggers the Shibboleth IDP middleware to generate a SAML assertion with an authentication statement. The integration aspect of this URL is very straightforward. First, the secure URL must be protected by some access control mechanism such that a web browser cannot visit the URL until after it has performed an authentication transaction with the SSO system that is protecting the URL. Second, at the time that a browser visits the SSO URL, the Shibboleth IDP middleware must be told the identity of the user that has authenticated, so that Shibboleth knows which user the authentication statement is for. This can occur via an environment variable within the web servlet container, if the servlet container has access to the information, or it can occur via an HTTP header that can be read by the Shibboleth IDP middleware.

4.1.3.2 Attribute Authority (AA) Integration Point

The purpose of the Attribute Authority (AA) Integration Point is to allow the Shibboleth IDP middleware to construct SAML attribute statements containing information about users and send those statements to SPs that request them. The integration point consists of a set of software modules or functions that operate according to a callback-style paradigm. These modules are called ***connectors***. The purpose of a connector is to construct an attribute value for a specified SAML attribute and deliver value to the Shibboleth IDP middleware so it can be packaged into a SAML attribute statement. (See Section 4.3 for detail about SAML attribute statements.) The Shibboleth IDP middleware is packaged with several default connectors that are capable of constructing SAML attribute values by performing very simple operations such as LDAP repository lookups or ODBC queries. But these standard connectors could not generate GFIPM metadata

assertions, so GTRI engineers designed and built a custom connector package. This custom code is discussed in more detail in Section 4.1.6.

4.1.4 Web Single Sign-On (SSO) System

The Web Single Sign-On (SSO) System integrates with the Shibboleth IDP middleware via the SSO Integration Point (discussed previously in Section 4.1.3.1) and provides the basis for the Shibboleth IDP middleware to generate SAML authentication statements about users. In order to realize the single sign-on benefit of federated identity management, an SSO system is usually a pre-existing component that is already used for authenticating users for other purposes. Details about the SSO systems used by federation participants can be found in Section 6.

There are very few limitations on an SSO system; the only major constraint is that it be capable of integrating properly with Shibboleth. The simplest type of SSO system to integrate with Shibboleth is one that (1) physically resides on the machine where the Shibboleth IDP middleware resides (this is called "co-location") and (2) uses an authentication module that can integrate with the servlet container in which the Shibboleth IDP middleware module is running. In this type of configuration, the SSO system would pass identifying information about each authenticating user to the Shibboleth middleware via the servlet container environment. It is also possible to configure the Shibboleth IDP to use a non-co-located SSO system, as long as the SSO system is capable of securely proxying all web traffic to the SSO URL properly. This non-co-located SSO system integration model is not currently used by any GFIPM participants; however, it has been tested by GTRI and verified to work properly.

4.1.5 Attribute Data Store

The Attribute Store integrates with the Shibboleth IDP middleware via the Attribute Authority Integration Point (discussed previously in Section 4.1.3.2) and provides a source of trusted data about users that can be used to construct GFIPM assertions. Any component that acts as an Attribute Data Store is essentially a database. There are virtually no limitations on the Attribute Data Store in terms of how it stores attributes; however, it must store attributes in a fashion that allows for attribute queries based on a user ID or some other key that can be understood by the Shibboleth IDP and maps uniquely to a specific user. Typically, an organization will want to connect an IDP to an existing user attribute repository – often an LDAP repository or an Active Directory database. But it is also possible to use an ODBC or SQL database, a flat file on the local machine's filesystem, or any other repository, via custom Java code. Details about the Attribute Data Store systems used by project participants in the GFIPM federation are available in Section 6.

4.1.6 Attribute Authority (AA) Connector and Assertion Generator

As discussed previously in Section 4.1.3.2, it was necessary during the GFIPM project to create a custom AA connector to generate SAML attribute values containing GFIPM metadata. This custom code was designed to conform to the Shibboleth IDP middleware's callback interface for attribute authority connectors. It works as follows.

1. The Shibboleth IDP middleware invokes the connector to request an attribute value for a specified user account.
2. The connector performs LDAP lookups of appropriate attribute values for the specified user in the Attribute Data Store, based on a configuration file.
3. The connector performs data transformations on the retrieved LDAP data as needed, and packages the data in an XML structure that is conformant to the GFIPM metadata standard. Most of these transformations are specified using XSLT and must be uniquely configured by each participant according to its needs.
4. The connector passes the GFIPM-conformant XML back to the Shibboleth IDP middleware to be packaged as an attribute value within a SAML attribute statement.

Figure 3 illustrates the four-step process outlined above in the context of the basic IDP structure that was presented in Section 4.1.

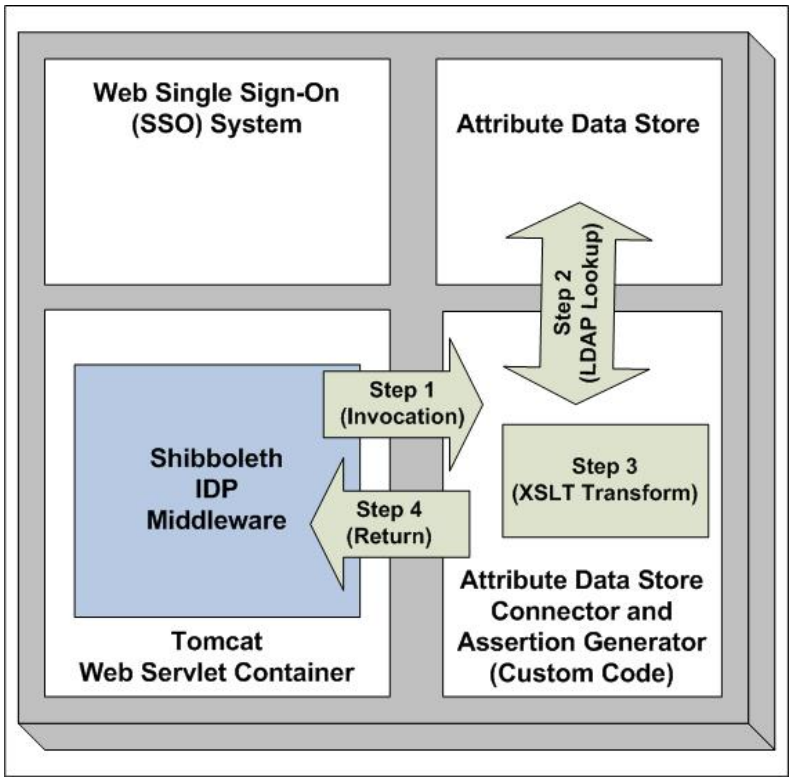

Figure 3: Process of Generating a GFIPM Assertion

The custom connector was written in Java so that it could integrate easily with the Shibboleth IDP middleware, which is also written in Java. GTRI wrote the Java code; however, since each participant needed to pull raw attribute data out of an LDAP repository with a structure unique to that organization, it was necessary for each participant to create some custom XSLT style sheets for use with the custom connector.

4.2 SP Structure

Just as in an IDP, there are several basic components in each SP within the GFIPM federation that are noteworthy. Figure 4 illustrates the basic SP architecture used by GFIPM participants. Details of each component in the diagram are discussed in the subsections below.

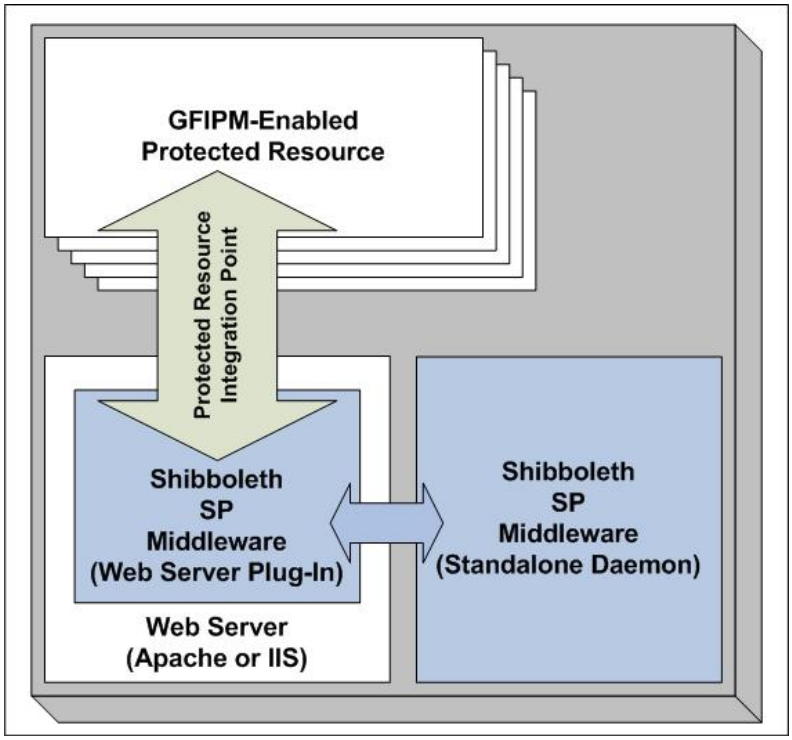

Figure 4: Basic GFIPM Service Provider (SP) Structure

4.2.1 Web Server

Shibboleth enables basic federated identity management functionality for web-based resources; not surprisingly, therefore, a Shibboleth SP must contain a web server. The web server is responsible for serving sensitive web-based resources to users that request them, subject to access controls and other usage policies that may exist. Resources may be served either directly or via a reverse-proxy-based architecture. (See Section 4.2.4 for discussion about the rationale for a reverse-proxy-based architecture and the basic structure of a reverse proxy.) In addition to serving web-based resources, the web server is responsible for handling SSL/TLS encryption of HTTPS traffic, including incoming SAML messages that are sent to the server from IDPs via HTTPS. The web server component is integrated with Shibboleth SP middleware and relies on the Shibboleth module to handle basic SAML operations, as discussed in the next section.

GFIPM participants have successfully used two different web servers at their SPs: Microsoft Internet Information Services (IIS) and Apache. All IIS deployments in the GFIPM federation have been in a Microsoft Windows environment, and all Apache deployments have been in a Red Hat Enterprise Linux (RHEL) environment. It may be

possible to deploy a Shibboleth-based SP using a web server other than IIS or Apache, but this issue was not explored during the GFIPM project, as it was not necessary.

4.2.2 Shibboleth SP Middleware

The Shibboleth SP middleware, depicted by a pair of blue boxes in Figure 4, is implemented in C++ and consists of a pair of components (a web server extension module and a stand-alone daemon) that work in tandem. The web server extension module runs as a plug-in for Apache, and as an ISAPI filter on Microsoft IIS. Like the Shibboleth IDP middleware, it was developed by the Internet 2 project using the OpenSAML open source implementation of SAML. The Shibboleth SP middleware must be integrated with sensitive web-based resources via an integration point that allows it to protect them. (See Section 4.2.3 for more detail on this integration point.) Shibboleth SP middleware handles the processing of incoming SAML messages from IDPs, as well as the creation of outgoing SAML messages to IDPs. In addition, it manages the signing and encryption of all outgoing SAML messages, as well as signature verification and decryption of all incoming SAML messages. Additional information about the SAML protocols and messages used within the GFIPM federation is available in Section 4.3.

4.2.3 Protected Resource Integration Point

There is only one basic SP integration point for any SAML 1.1 SP. (This includes not only Shibboleth SPs, but also any other SP middleware package.) It is called the ***Protected Resource Integration Point***, and it enables the Shibboleth SP middleware to protect or help protect sensitive web-based resources that are to be shared within the GFIPM federation. There are two ways in which the Shibboleth SP middleware can be integrated with sensitive resources. Each method is discussed below.

- ***Integration Method #1*** – The first integration method involves the use of access control logic that is built into the Shibboleth SP middleware itself. When using this integration method, it is possible to construct simple access control policies that permit or deny access to specific URLs served by the web server based on logic involving the SAML attribute values presented for a user. At the implementation level, this integration method is as simple as encoding the appropriate access control logic in a static Shibboleth configuration file on the SP host. This integration method works only when the following two conditions apply. First, the granularity of access control for a resource protected in this manner is limited to static URLs. In other words, it is not possible to make access control decisions that depend on HTTP query variables (such as CGI parameters) using this technique. Second, the SAML attribute values used to make access control decisions in this manner must be relatively simple, e.g. limited to simple values such as a name or an email address. It is generally not practical to use this integration method when access control decisions must be made based on very complex SAML attribute values, such as XML-encoded GFIPM metadata instances. Therefore, this integration method is of limited usefulness within the GFIPM federation.

- ***Integration Method #2*** – The second integration method involves passing SAML attribute data directly from the Shibboleth SP middleware to a protected resource and allowing the resource itself to make its own access control decisions based on the data. SAML attribute data is passed to an application via the following mechanism. The Shibboleth middleware inserts each SAML attribute value into an HTTP header that is passed to a resource at the same time that the resource receives notification of an HTTP request from the web server. (At the implementation level, the resource typically receives all of its HTTP headers – including the headers constructed by the Shibboleth SP middleware – via environment values in its local environment.) After receiving the SAML attribute values, the resource can examine and process them as needed and make decisions accordingly.

SPs in the GFIPM pilot federation use a combination of both integration methods, as follows. First, basic access to each portal is controlled using Shibboleth's native access control, with the following simple policy: "A user may access resources on this SP if and only if he has authenticated with an IDP in the federation." Then, a more fine-grained access control policy is implemented by the resources themselves, or by a proxy/portal service that controls access to the resources, using SAML attribute values passed from the Shibboleth SP middleware. As previously mentioned in Section 4.2.1, there are important considerations that influence the decision to serve a sensitive resource in the GFIPM federation directly or via a reverse proxy arrangement. The next section begins to discuss this topic in more detail.

4.2.4 Optional GFIPM-Enabled Proxy/Portal Service

The previous section of this report discussed the Protected Resource Integration Point through which a sensitive resource can interact with the Shibboleth SP middleware within a web server. It is through this integration point that a resource can receive GFIPM metadata assertions for processing. Of course, a GFIPM assertion is not useful to a resource unless the resource has the ability to understand it, process it, and use it as needed for purposes such as user identification, access control, and auditing. A resource that has this ability is called ***Federation-Enabled*** or ***GFIPM-Enabled***.

Most of the resources that are useful to federation users tend to be complex applications that contain their own logic, rather than simple static web pages. And for most of these applications, it is difficult or impossible to modify the application's source code, often due to circumstances surrounding the ownership of the resource. (Several of the participants in the GFIPM project act as umbrella organizations and manage networks containing resources that are owned by other organizations. For example, JNET manages access to data from the Pennsylvania Department of Corrections; however, JNET does not own that data.) If a resource is not already federation-enabled (as is the case for nearly all resources), and its source code cannot be modified for federation-enablement, then the resource cannot exist natively within the GFIPM federation, so it must be made available to federation users via a ***GFIPM-Enabled Proxy/Portal Service***. As its name implies, such a service exists to provide proxied access to a non-federation-enabled resource in a manner that allows the resource to exist in the GFIPM federation (and be

available to federation users) without the resource having to be modified. The process of determining the need for such a service and implementing a working prototype of it constitutes one of the most important lessons learned during the GFIPM demo project. The topic of federation-enablement of resources is discussed at length in Section 7. The remainder of this section is concerned with the basic internal structure of the service.

Figure 5 shows the GFIPM SP structure that has already been depicted in Figure 4; however, it also illustrates how the GFIPM-Enabled Proxy/Portal Service fits into the SP architecture. Note that the proxy/portal service integrates directly with the Shibboleth SP middleware, and that it relieves non-GFIPM-enabled resources from having to integrate with Shibboleth. It also relieves resources of the burden of processing GFIPM metadata assertions.

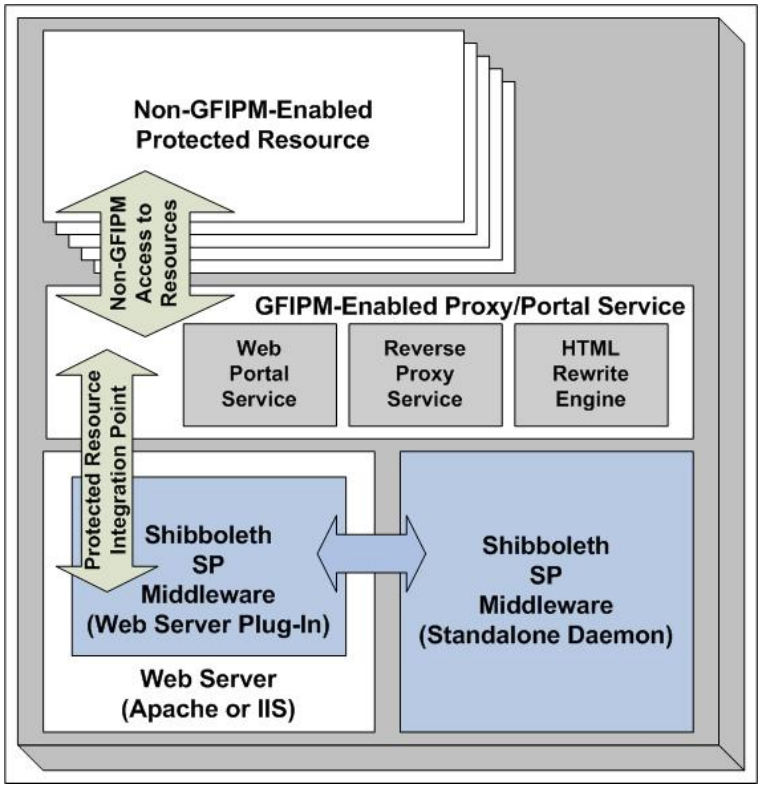

Figure 5: GFIPM Service Provider (SP) Structure with a GFIPM-Enabled Portal/Proxy Service

As illustrated in Figure 5, a GFIPM-Enabled Proxy/Portal Service includes the following logical components.

- ***Web Portal Service*** – This is a federation-aware application that consumes GFIPM assertions and translates their content into access privileges for the proxied resources. Using these privileges, it defines network-level access control policies that are to be enforced by the reverse proxy service (see below). The Web Portal Service also presents a web-based user interface for accessing the

proxied resources. Finally, when necessary it may also handle the auditing of access attempts to the proxied resources.

- ***Reverse Proxy Service*** – This is a network-level service that allows traffic to flow from a user's browser, through the SP's web server, to protected resources. It enforces network-level access control based on instructions that it receives from the Web Portal Service. (The technical details concerning how access control policy is communicated from the Web Portal Service to the Reverse Proxy Service are implementation-dependent and outside the scope of this report.) Finally, when necessary, the Reverse Proxy Service handles authentication to protected resources in a manner that the resources understand.
- ***HTML Rewrite Engine*** – One of the challenging but necessary aspects of reverse-proxying web content is rewriting the HTML, JavaScript, and other content served by the proxied resource, so that to the user's browser the content appears to have come directly from the proxy. This process typically involves modifying all URLs within the content that refer back to the proxied web server, so that they refer to the reverse proxy server.

GTRI has developed a low-cost solution for service providers that need to deploy a GFIPM-Enabled Proxy/Portal Service capability. The solution leverages a relatively inexpensive COTS reverse proxy product and HTML rewrite engine, called EZproxy, and also contains custom code. CISA is currently using this GTRI-developed solution to serve non-federation-enabled resources to GFIPM users. JNET has also developed its own GFIPM-Enabled proxy/portal capability using entirely custom code, and is serving non-federation-aware resources with it.

4.2.5 Protected Resources

A protected resource is the actual sensitive content that needs to be protected by the SP infrastructure. As implied by the discussion in the previous two sections, all protected resources fall into two categories: GFIPM-enabled resources, that natively understand GFIPM metadata, and non-GFIPM-enabled resources that must be proxied. Most of the protected resources that are currently in the GFIPM federation are non-GFIPM-enabled, and it is expected that nearly all resources that will be added to the federation in the future will also be non-GFIPM-enabled. Section 7 contains extensive discussion of lessons learned during the project on the topic of integrating non-GFIPM-enabled resources into the GFIPM federation.

4.3 SAML Usage Profile for GFIPM

This section presents an overview of the SAML Usage Profile for GFIPM, which consists of the SAML standard currently being used in the GFIPM pilot federation. Section 4.3.1 introduces some basic concepts that are central to the SAML standard, and Section 4.3.2 describes the parts of SAML that are currently used in GFIPM.

4.3.1 Basic SAML Concepts

The Security Assertion Markup Language (SAML) is a large, comprehensive standard for data and transactions related to federated identity management. SAML addresses concepts such as the following:

- ***Statements*** – in which claims can be made about users;
- ***Assertions*** – protocol messages containing statements;
- ***Protocols*** – which define the patterns of interaction through which assertions are delivered to their recipients;
- ***Bindings*** – which define how SAML protocols are carried out in terms of lower-level protocols such as SOAP and HTTP;
- ***Profiles*** – which incorporate specific protocols and bindings to achieve a specific federated identity management function.

One of the fundamental building blocks of SAML is the concept of a ***statement***. There are three types of SAML statements:

- ***Authentication Statement*** – used by an IDP to state that a specific user has performed an authentication transaction;
- ***Attribute Statement*** – used by an IDP to state that a specific set of attributes are associated with a specific user;
- ***Authorization Decision Statement*** – used by an IDP to state that a specific user is permitted to perform a specific action on a specific resources, given a specific set of evidence.

GFIPM uses only authentication statements and attribute statements. Authorization decision statements are not used, because they violate one of the basic tenets of the GFIPM concept – specifically, each SP must be permitted to make all access control decisions about the resources it manages.

4.3.1.1 Authentication Statements

An authentication statement is used by an IDP to claim that a specified user has successfully authenticated with an SSO system at the IDP. It can include additional supplementary information such as the time of the authentication event, the IP address of the user's computer, and the type of authentication method used. Figure 6 contains a sample SAML 1.1 authentication statement that could have been sent within the GFIPM federation.

```
<AuthenticationStatement
   xmlns="urn:oasis:names:tc:SAML:1.0:assertion"
   AuthenticationInstant="2006-06-26T19:16:03.717Z"
   AuthenticationMethod="urn:oasis:names:tc:SAML:1.0:am:unspecified">
 <Subject>
  <NameIdentifier
    Format="urn:mace:shibboleth:1.0:nameIdentifier"
    NameQualifier="global:gfipm:linuxrefidp">
     _51d57480b10a61e4fb987ba6b98ea9c6
  </NameIdentifier>
  <SubjectConfirmation>
   <ConfirmationMethod>urn:oasis:names:tc:SAML:1.0:cm:bearer</ConfirmationMethod>
  </SubjectConfirmation>
 </Subject>
 <SubjectLocality IPAddress="10.50.14.239"/>
</AuthenticationStatement>
```

Figure 6: Sample SAML 1.1 Authentication Statement

Note that in this authentication statement, the subject of the statement (the user for whom the statement was made) is not identified using any clearly identifiable name. Specifically, the text string "_51d57480b10a61e4fb987ba6b98ea9c6" is used to identify the subject. This is due to a Shibboleth-specific feature that allows for subjects to remain anonymous when possible for the sake of privacy. In the GFIPM environment, this anonymous subject identifier is of no consequence, because an SP would also receive an attribute statement pertaining to this subject, and the attribute statement would contain personally identifying information about the user.

4.3.1.2 Attribute Statements

An attribute statement is used by an IDP to claim that certain attributes and attribute values pertain to a specified user. Figure 7 contains a sample SAML 1.1 attribute statement. Note that the attribute statement pertains to a subject with the identifier "_51d57480b10a61e4fb987ba6b98ea9c6" – the same identifier used in the authentication statement in Figure 6. This common subject identifier provides the means through which an SP can associate specific attributes with a specific authentication event, IP address, etc. In Figure 7, the statement claims that for the attribute "FederationPersonName", this subject has the value "George Burdell", and that for the attribute "UserID", this subject has the value "gburdell".

```
<AttributeStatement>
 <Subject>
  <NameIdentifier
    Format="urn:mace:shibboleth:1.0:nameIdentifier"
    NameQualifier="global:gfipm:linuxrefidp">
    _51d57480b10a61e4fb987ba6b98ea9c6
  </NameIdentifier>
 </Subject>
 <Attribute AttributeName="FederationPersonName"
            AttributeNamespace="urn:mace:shibboleth:1.0:attributeNamespace:uri">
  <AttributeValue>George Burdell</AttributeValue>
 </Attribute>
 <Attribute AttributeName="UserID"
            AttributeNamespace="urn:mace:shibboleth:1.0:attributeNamespace:uri">
  <AttributeValue>gburdell</AttributeValue>
 </Attribute>
</AttributeStatement>
```

Figure 7: Sample SAML 1.1 Attribute Statement

Note that the example in Figure 7 uses very few attributes and also uses very simple attribute values. When an IDP is sending only a few simple attributes to an SP, it is straightforward to use a technique like this and encode the attributes in SAML by defining one SAML attribute name for each logical attribute value. But in the GFIPM federation, IDPs send a large volume of attribute data that conforms to a large, complex data model for which there is an XML schema. In this case, the approach used here (using one logical attribute per SAML attribute) may not work well. The topic of encoding GFIPM metadata assertions inside SAML attribute statements is covered in more depth in Section 4.6.

4.3.1.3 SAML Assertions

A SAML assertion is a message sent from an IDP to an SP containing statements such as those described in the previous two sections. Typically, a SAML assertion that contains an authentication statement is called an ***authentication assertion***, and a SAML assertion that contains an attribute statement is called an ***attribute assertion***. It is possible, under certain SAML profiles, for a SAML assertion to contain multiple statements; however, this does not occur in the GFIPM federation.

Figure 8 contains a sample SAML 1.1 attribute assertion. This assertion is simply the attribute statement from Figure 7 with the appropriate assertion context around it.

```
<Assertion
   xmlns="urn:oasis:names:tc:SAML:1.0:assertion"
   AssertionID="_19ef0cd871a089eb5f8d262fa8813885"
   IssueInstant="2006-06-26T19:16:33.369Z"
   Issuer="global:gfipm:linuxrefidp"
   MajorVersion="1" MinorVersion="1">
 <Conditions NotBefore="2006-06-26T19:16:33.369Z" NotOnOrAfter="2006-06-
26T19:46:33.369Z">
  <AudienceRestrictionCondition>
   <Audience>global:gfipm:linuxrefsp</Audience>
  </AudienceRestrictionCondition>
 </Conditions>
 <AttributeStatement>
  <Subject>
   <NameIdentifier
     Format="urn:mace:shibboleth:1.0:nameIdentifier"
     NameQualifier="global:gfipm:linuxrefidp">
     _51d57480b10a61e4fb987ba6b98ea9c6
   </NameIdentifier>
  </Subject>
  <Attribute AttributeName="FederationPersonName"
             AttributeNamespace="urn:mace:shibboleth:1.0:attributeNamespace:uri">
    <AttributeValue>George Burdell</AttributeValue>
  </Attribute>
  <Attribute AttributeName="UserID"
             AttributeNamespace="urn:mace:shibboleth:1.0:attributeNamespace:uri">
    <AttributeValue>gburdell</AttributeValue>
  </Attribute>
 </AttributeStatement>
</Assertion>
```

Figure 8: Sample SAML 1.1 Attribute Assertion

Note that the assertion in Figure 8 contains supporting detail that provides context for the attribute statement that it contains. In this example, the assertion provides the identifier

of the IDP that issued it (using "Issuer"), the time at which it was issued, (using "IssueInstant"), the version of SAML to which it conforms (using "MajorVersion" and "MinorVersion"), the period of time during which it is valid (using "NotBefore" and "NotOnOrAfter"), and the intended recipient (SP) of the assertion (using "Audience").

4.3.2 SAML Components Used by GFIPM

As described in Section 4.3.1, SAML consists of formal specifications for a wide range of structures related to federated identity management. Some of the simpler SAML constructs – statements and assertions – have already been introduced with examples in previous sections. At a higher level, SAML also contains protocols, bindings, and profiles. For the purpose of defining what subset of SAML 1.1 is used by GFIPM, it suffices to list the SAML profiles that are used. The GFIPM federation uses two SAML 1.1 profiles: the Shibboleth/SAML Web Browser SSO Profile and the SAML Assertion Query Profile. These SAML profiles were chosen because they constitute the subset of SAML 1.1 used by Shibboleth 1.3 and required for basic federated identity management using a Shibboleth-based infrastructure.[3] Each of these SAML profiles is discussed in detail in the following subsections.

4.3.2.1 Shibboleth/SAML Web Browser Single Sign-On (SSO) Profile

The first SAML 1.1 profile used in the GFIPM federation is the Shibboleth/SAML Web Browser SSO Profile. This profile consists of a complete specification of SAML protocols and bindings by which a user can be directed through the SSO process in a very user-friendly and seamless manner. The profile is not technically part of the SAML 1.1 specification; however, it is implemented and supported by Shibboleth as an extension to SAML 1.1. The choice of this specific SSO profile carries the implication that non-Shibboleth SAML 1.1 systems cannot currently interoperate with the GFIPM federation. But this lack of interoperability is not a serious problem, because (1) the GFIPM participants plan to move the federation to SAML 2.0 in the very near future, and the SSO profile used in the GFIPM federation has been adopted into the SAML 2.0 specification as the standard SAML 2.0 web browser SSO profile. In other words, after the federation adopts SAML 2.0, this interoperability problem will resolve itself.

Figure 9 contains a message sequence chart depicting the basic message flow that occurs during a transaction in the SSO profile used by the federation.

[3] Note that the specific SAML Usage Profile for GFIPM that is presented here is not intended to serve as an official usage profile recommendation or standard for the project; rather, it merely represents the current usage of SAML within the GFIPM federation. It is expected that this usage profile will change as the GFIPM project moves from SAML 1.1 to SAML 2.0 in the near future.

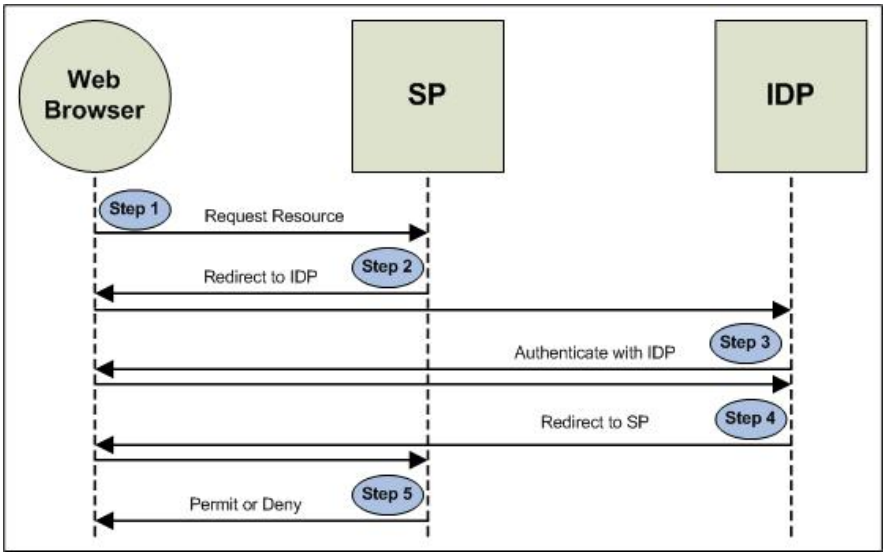

Figure 9: Shibboleth/SAML Web Browser SSO Profile

The profile works as follows, with each numbered item in the following list corresponding to a step in Figure 9.

1. The user attempts to access a protected resource at an SP by visiting the URL of the resource with his browser. If the user's browser has not already established a session with the SP via this profile, then the SP initiates the profile in Step 2.
2. The SP redirects the user's browser to the user's IDP – specifically to the SSO integration point URL on the IDP – via an HTTP redirect. Details about how the SP determines which IDP to use at this point in the transaction are outside the scope of this SAML profile; however, there are several practical solutions to this problem. One such solution, and the solution used in GFIPM, is to use a Where-Are-You-From (WAYF) service. This is discussed in Section 4.4.
3. When the user arrives at his IDP, he is asked to authenticate with the IDP's SSO system if he has not already done so. If the user has already authenticated and established a valid session with the IDP, then this step is unnecessary.
4. The IDP generates a SAML assertion containing an authentication statement. This statement vouches for the user as having authenticated properly with the IDP. The IDP then redirects the user's browser back to the SP – specifically to the SP's assertion consumer service – and directs the browser to perform an HTTP post operation with the contents of the SAML assertion. Through this post operation, the SAML assertion travels from the IDP to the SP via the browser.
5. The SP either grants or denies access to the original URL (protected resource) that the user tried to access.

All steps of this profile occur using an HTTP or HTTPS binding, except for Step 3 (user authentication with the IDP). The details of Step 3 are outside the scope of the profile; however, in practice, this step usually occurs over HTTP or HTTPS as well. Screen shots that illustrate an SSO transaction using this SAML profile from the perspective of the end user are available in Section 6.7.

4.3.2.2 SAML Assertion Query Profile

The second SAML 1.1 profile used in the GFIPM federation is the SAML Assertion Query Profile. This profile consists of a complete protocol and binding specification by which an SP can query an IDP to receive an attribute statement about a user. Transactions of this profile typically occur after an SP and IDP have engaged in the Shibboleth/SAML Web Browser SSO Profile to authenticate a user and deliver an authentication statement for that user to the SP. The Assertion Query Profile uses a direct channel between the SP and IDP. This direct channel from SP to IDP is often called a "back channel", because unlike the indirect channel between IDP and SP in the SSO profile, the channel used in the Assertion Query Profile does not rely on the browser as an intermediary between the SP and IDP.

The Assertion Query Profile is very simple, and works as follows.

1. The SP makes an attribute query using the SAML Assertion Query Protocol, using a SOAP binding over a secure HTTP (HTTPS) channel.
2. The IDP sends an assertion containing an attribute statement over the same channel used for the request.

In this profile, the SP tells the IDP which user is the subject of the requested attribute statement, based on a subject ID from the authentication statement for that user. This connection between authentication statement and attribute statement was previously highlighted in Section 4.3.1.2.

4.4 Where-Are-You-From (WAYF) Service

In Step 2 of the Shibboleth/SAML Web Browser SSO Profile discussed in Section 4.3.2.1, it is necessary for the SP to redirect a user's browser to the appropriate IDP where the user can authenticate and have an authentication assertion sent to the SP on his behalf. But the process through which the SP chooses which IDP to use during this transaction is outside the scope of the SAML standard. This is a problem for which there is currently no ideal solution. However, several acceptable, practical solutions to the problem do exist. One of the most popular techniques for solving the problem is to use a Where-Are-You-From (WAYF) service. A WAYF service offloads the task of choosing an IDP from the SP to the user by inserting an additional step in the Shibboleth/SAML Web Browser SSO Profile. During an SSO transaction, instead of redirecting the user's browser directly to an IDP, the SP redirects the browser to the WAYF service. Then the WAYF service presents the user with an input box (e.g. a pull-down selector) through which the user can tell the WAYF which IDP he wants to use. Finally, based on the user's input, the WAYF redirects the browser to the appropriate IDP. Figure 10 illustrates how a WAYF service fits into the SSO profile. The steps in the SSO profile transaction in Figure 10 are identical to those in Figure 9, except for Step 2, which has been modified as follows.

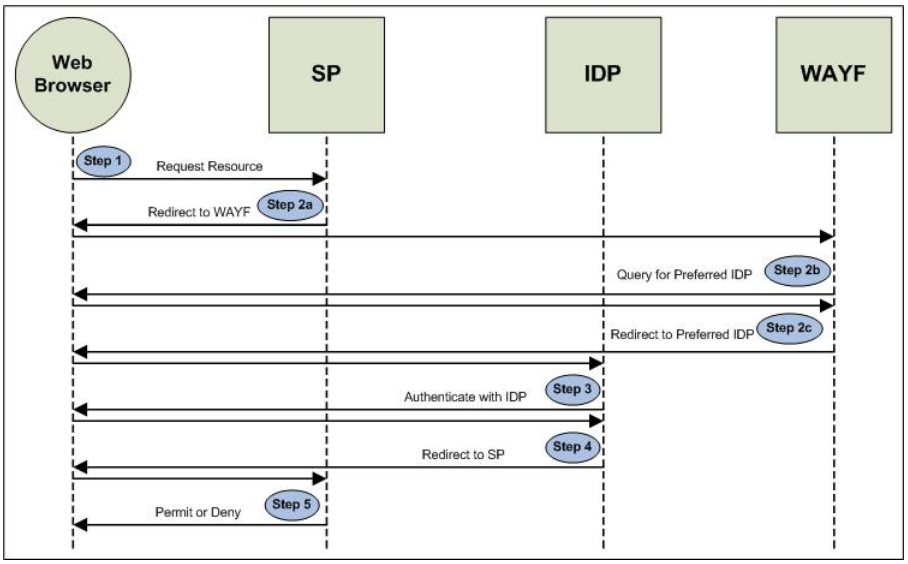

Figure 10: Modified Shibboleth/SAML Web SSO Profile with a Where-Are-You-From (WAYF) Service

2a. The SP redirects the user's browser to the WAYF service.
2b. The WAYF service queries the user for his preferred IDP.
2c. The WAYF service redirects the user's browser to the preferred IDP.

The WAYF service used in the GFIPM federation is a modified version of an open source WAYF application originally created by the Swiss Education & Research Foundation (SWITCH). It is implemented in PHP, and was modified slightly by GTRI to suit the needs of GFIPM. Section 6.7 contains a screen show that shows how the WAYF service appears from the end user's perspective during an SSO transaction in the GFIPM federation.

4.5 Federation Trust Fabric (Federated Entities Metadata)

Section 4.3 and its subsections presented the SAML standard as it is used in GFIPM. Note, however, that every concept in the SAML standard is posited on the existence of trust between IDPs and SPs in a federation at the protocol level. In other words, before IDPs and SPs can securely and privately communicate, they need to be able to trust in two assumptions: (1) that messages sent can be read only by the intended recipient, and (2) that messages received actually came from the presumed sender. To achieve this goal, a federation must employ a cryptographic "trust fabric". In the GFIPM federation, this trust fabric takes the form of ***SAML Federated Entities Metadata***. This metadata is *not the same as the GFIPM assertion metadata* that has already been discussed throughout the report. Rather, it is a document containing critical trust information (metadata) about each federated entity (IDP, SP, or WAYF service) in the GFIPM federation. Table 2 provides a summary of the information that is included in this document for each federated entity.

	Component	Description/Purpose
For All Entities	**Entity Descriptor**	Contains the name of the federated entity.
	Organization	Contains information about the organization that manages this federated entity. Includes at least the organization name.
	Contact Person	Contains contact information for a person who is in charge of this federated entity. Includes first name, last name, and email address.
For IDPs Only	**IDP SSO Descriptor**	Describes the single sign-on service endpoint at this IDP by giving it a unique name, binding its key descriptor to its URL, and listing the SSO protocols that it supports.
	SSO Key Descriptor	Contains the public key used for the single sign-on service endpoint at this IDP.
	SSO URL	Contains the URL for the single sign-on service endpoint at this IDP.
	AA Descriptor	Describes the attribute authority endpoint at this IDP by giving it a unique name, binding its key descriptor to its URL, and listing the attribute query protocols that it supports.
	AA Key Descriptor	Contains the public key used by the attribute authority endpoint at this IDP to sign attribute assertions.
	AA URL	Contains the URL for the attribute authority service at this IDP.
For SPs Only	**SP SSO Descriptor**	Describes the endpoint used during the SSO protocol at this SP by giving it a unique name, binding its key descriptor to its URL, and listing various options about the SSO protocols that it supports.
	Key Descriptor	Contains the public key used by the SP to sign requests in both the SSO protocol and the attribute query protocol.
	Assertion Consumer Service URL	Contains the URL to which IDPs must send authentication assertions during SSO transactions with this SP.

Table 2: Information Provided in the SAML 2.0 Federated Entities Metadata Document (Trust Fabric) for Each Federated Entity

There are several additional points about this trust fabric that are important enough to warrant mention in this report.

1. The federated entities metadata document itself is signed by a key that is trusted by all federation participants. The signing key for this document is the lynchpin for all trust among IDPs and SPs within the federation. The key must be explicitly trusted by each IDP and SP in the federation, or else complete protocol-level trust cannot exist. Therefore, it is of utmost importance for the federation to take adequate measures to protect this signing key. Any compromise of this key could lead to a catastrophic compromise of trust within the federation.
2. The federated entities metadata document format used within the GFIPM federation is in conformance with the SAML 2.0 standard. This document format it is used in GFIPM, even though the federation does not use SAML 2.0, because the version of Shibboleth used in the federation supports this particular aspect of SAML 2.0.
3. One of the challenges that must be addressed within any SAML-based federation is the generation and dissemination of trust fabric updates to all federated entities in the federation in a timely manner. These updates must happen if, for example, an IDP or SP joins or leaves the federation. The SAML 2.0 specification for federated entities metadata does not prescribe any specific mechanisms by which to distribute new versions of the trust fabric document. Up to this point in the project, GFIPM participants have dealt with this issue in a manual fashion, e.g.

via email. Updates to the trust fabric have been very infrequent, since the pilot federation's size has remained constant throughout the demonstration project, with no new IDPs or SPs.

4. The federation trust fabric described here is only one of the levels at which cryptographic trust must exist within the federation. There are others as well. For example, each IDP and SP in the federation must actively listen for HTTP requests on a secure (SSL/TLS) channel. (Both IDPs and SPs interact with users' browsers via SSL/TLS, and in addition, IDPs and SPs interact with each other via SSL/TLS.) It is therefore necessary for each IDP and SP to have an SSL/TLS server certificate within a public key infrastructure (PKI) that is trusted by browsers. There are several methods by which this problem can be solved, and each has its benefits and drawbacks. Note that this issue is not unique to federations; it applies to all web services that leverage SSL or TLS. Another level at which cryptographic trust must exist in the federation is between IDP and browser if users of that IDP authenticate with it using client certificate authentication. The implementation details of such a PKI are likely to differ for each federation member, and are not important for the purpose of this discussion.

4.6 GFIPM Metadata

One of the most important components of the GFIPM architecture is the GFIPM metadata. This topic has already been described briefly in Sections 2.2.3 and 3.1.7. This section expands on that description and provides more detail about the purpose, structure, contents, and development process of the GFIPM metadata. The information in this report pertains specifically to version 0.4 of the GFIPM metadata, which was released in March 2007 and is currently the most recent version of the metadata to be released.

4.6.1 Purpose of the Metadata

A significant part of the GFIPM demo project has been concerned with the development of a standard metadata model through which to pass information about federation users and other federation entities from IDPs to SPs. This aspect of the GFIPM project is what most clearly distinguishes GFIPM from other similar federated identity projects in the law enforcement space. By developing a common data model that describes users, and allowing an IDP to share that data with an SP at the point in time when a user accesses the SP's services, the GFIPM concept facilitates the integration of legacy applications into a federation by helping to meet the applications' usage requirements for federation users. As discussed in Section 2.2, the usage requirements for applications typically fall into several basic categories, including *terms of use*, *provisioning*, *inter-session persistence*, *user identification*, *access control*, *auditing*, and *personalization*. One of the principal lessons learned during the GFIPM project is that federation enablement of legacy applications is crucial to the success of any information sharing federation. See Section 7.4 for more detail about federation enablement of applications.

4.6.2 Metadata Framework

The GFIPM metadata model was developed in accordance with two overriding requirements that reflect the intended usage of the metadata in the future. First, the

metadata model was intended not only for the current user-to-application use case, but also for other use cases such as a service-oriented architecture. (See Section 3.3.4 for more details about the user-to-application use case and other use cases.) Therefore, the entire federated user conceptual model had to be reusable across these various use cases within the justice domain. The second requirement was that the metadata model needed to meet the basic business requirements of federation members without being so large and unwieldy that is was unusable. To this end, participants developed a multi-layer GFIPM metadata framework that defines flexible and reusable concepts of a federated user and a federated entity for federated identity and privilege management, while meeting the two requirements stated above. This section and its subsections describe the framework in detail.

The metadata framework has the following primary objectives:

1. Leverage the existing GJXDM and NIEM data modeling concepts, principles, architecture, and content (semantics and structure).
2. Leverage existing federated identity standards, especially SAML. Support for other standards and versions is anticipated in the future.
3. Separate the identification and modeling of GFIPM Metadata from the encoding and transport of that metadata between federation participants in a SAML assertion. This allows for parallel efforts of reconciling and consensus on the business aspects of the metadata with that of specific technical details.

Figure 11 depicts the four layers of the GFIPM metadata framework.

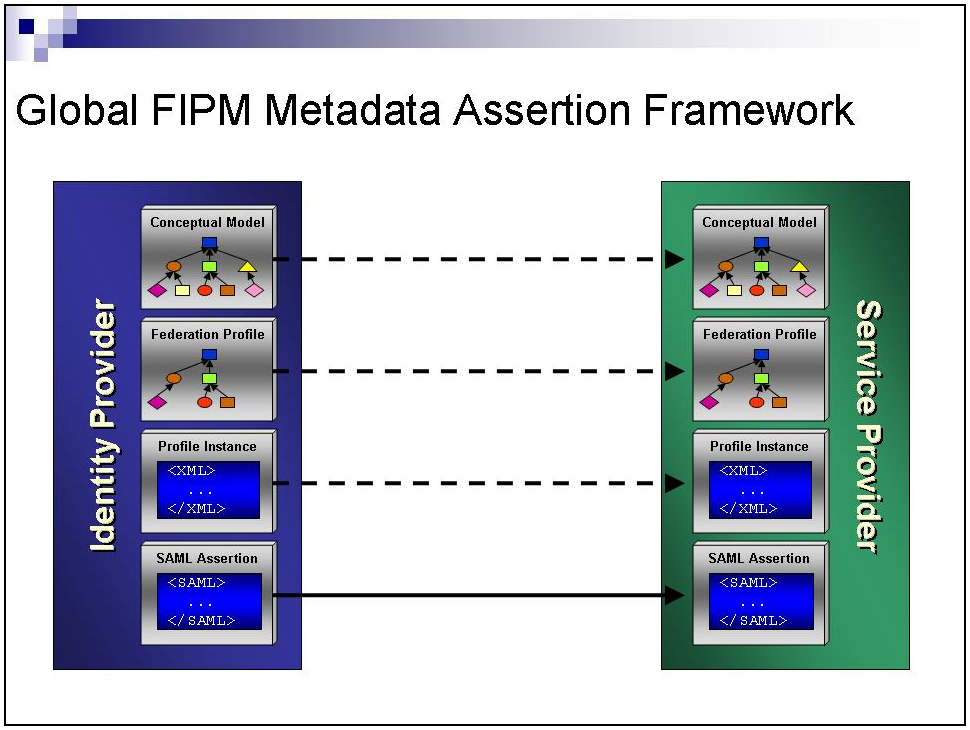

Figure 11: The GFIPM Metadata Framework

Each layer's purpose and representation is explained in the following subsections.

4.6.2.1 Conceptual Model Layer

At the highest layer the framework is the ***Conceptual Model***. A well-defined, standardized conceptual data model of a federated user (or entity) is essential to the GFIPM concept. Without a standardized data model, even simple concepts, such as "name", "employment", and "job title", can introduce ambiguity when they are used and shared across organizational boundaries. The conceptual model provides clear structure, semantics, and relationships of properties associated with a federated user. It is independent of the underlying transport protocol and representation used to move federated user metadata between IDPs and SPs.

In addition to the concept of a federated user, the conceptual model also defines the concept of a federated entity, which is any non-human object (e.g. hardware, software, service, etc.) that exists within a federation and requires a federated identity.

As previously stated, the GFIPM Metadata leverages the GJXDM/NIEM data modeling standard as the base vocabulary and naming and design rules in the data modeling effort for describing the conceptual model and building the associated schemas. However, neither GJXDM nor NIEM currently includes the concept of a federated user or a federated entity; therefore, these concepts must be defined here. It is expected that the GFIPM Metadata defined as part of this effort will be reconciled and potentially added to NIEM in the future.

The following principles were applied in the construction of the conceptual model.

1. ***Optional and Over-Inclusive***: The conceptual model is designed to act as a superset of the federated user or federated entity model that a typical federation would adopt for its use. It includes many concepts that may have limited or no applicability in one federation and be critical concepts in another federation. By being optional and over-inclusive, the GFIPM conceptual model can address the data modeling requirements of many different federations, thereby achieving maximum flexibility and reusability.
2. ***Leverage GJXDM and NIEM***: The conceptual model leverages existing GJXDM and NIEM data model standards as much as possible. This helps to minimize the development effort for the federated user and federated entity data models, leverage the existing knowledge base, and promote interoperability with existing systems and tools that have been built around the GJXDM and NIEM standards.
3. ***Leverage Existing Standards***: Many of the data requirements for the federated user and federated entity conceptual models have been identified in existing standards which address the generic federated identity and privilege management problem. Where attributes have been identified as part of a broader industry standard, they are referenced in the context of the existing standard rather than being redefined.

4. ***Supplement Existing Standards Where Necessary***: The GJXDM and NIEM standards form a critical part of the federated user and federated entity models; however, they do not include all of the concepts that are necessary for federated identity and privilege management. The federated user and federated entity models supplement GJXDM and NIEM by defining their own objects that represent additions to the GJXDM and NIEM vocabularies.

The conceptual model layer is currently specified via a set of documents that lay out its content and structure: a Microsoft Word document that illustrates the conceptual model from a high-level perspective, and a Microsoft Excel spreadsheet that presents a thorough overview of the conceptual model in greater detail, including a definition for each element and references to the appropriate NIEM, GJXDM, and SAML structures where necessary.

4.6.2.2 Federation Profile Layer

The ***Federation Profile*** layer allows the conceptual model to be subsetted and constrained for a given federation implementation. It serves as an adapter layer, allowing a specific federation to distill the general federated user and federated entity models down to an essential supported set. In addition to specifying a subset of the larger model, a federation profile can specify constraints and mandatory elements. The federation profile layer is specified as a set of NIEM constraint and subset schemas.

4.6.2.3 Federation Profile Instance Layer

The third layer of the architectural framework is the ***Federation Profile Instance*** layer. While the first two layers of the architecture (the conceptual model and the federation profile) are essentially schemas that define the structure of data objects, the federation profile instance layer is an actual data object (payload). It takes the form of an XML document that conforms to the XML schema of a specific federation profile. A profile instance is generated by an IDP that has firsthand knowledge of the personal attributes of a specific user or entity. The IDP builds it from data in a local attribute store, such as an LDAP directory or an ODBC database. The profile instance is intended for consumption by an SP, and is to be used for supporting identification, authentication, privilege management, auditing, personalization, and possibly account provisioning. Implicit in the federation concept is that an SP may use the profile instance as it sees fit. It may choose to use all of the data elements within the profile instance, or it may use very few of the elements, ignoring certain data elements in a profile instance, even though they may have been designated as mandatory by the federation profile for that federation. The goal of this layer is to allow an IDP to convey information about a federated user to an SP. How the SP chooses to use that information is outside the scope of the profile instance.

4.6.2.4 The SAML Assertion Layer

The fourth and final layer of the framework is the ***SAML Assertion*** layer which defines GFIPM assertions. A SAML assertion acts as a transport mechanism for the Federation Profile XML Instance on its journey from an IDP to an SP. SAML assertions can carry

attribute statements, which state facts about a user in the form of name/value pairs. The Federation Profile XML Instance can be encoded within one or more SAML attributes and transported in a SAML assertion.

The specification of rules for encoding an XML profile instance in a GFIPM SAML assertion has been a topic of investigation during the development of this framework. Several encoding strategies are possible, each with associated advantages and disadvantages. Lessons learned regarding this issue are currently being captured as participants conduct experiments to determine which encoding strategy is most appropriate for GFIPM. Many factors affect the choice of encoding strategy; the following list describes the most important ones.

1. A SAML attribute statement can contain any number of attributes, and in theory, these attributes can contain any type of data that can legally reside within an XML element. But in practice, some SAML COTS products may have limitations on what types of data they can reliably support within SAML attributes. Also, some SAML COTS products may have size limitations on SAML attribute values, or even on entire SAML statements.
2. Information about a federated user in the conceptual model requires rich context and relationships to be represented. NIEM provides the necessary constructs for this to be precisely represented in complex XML types. SAML attribute values allow for complex XML types but provide no method of linking context or associations between individual attribute/value pairs. It is also desirable to that the conceptual model of a federated user be consistent across the federation regardless of user-to-application or system-to-system (SOA) scenario. From the standpoint of simplicity, it is advantageous to keep the XML structure of a profile instance intact during transit within a SAML attribute statement. Attempting to break it up into multiple parts or translate it into a set of simpler structures is not desirable from this standpoint; however, as stated above, COTS product limitations may dictate that such steps be taken anyway.

The results of this research will be reported to the community when this work is complete.

Table 3 summarizes each of the four layers of the metadata framework.

Layer	Description

Conceptual Model	• Abstract conceptual model of information about a federation user or a federated entity • Provides consistent semantics and formal basis for sharing information about users and other entities in a federation • Optional and over-inclusive, defines a superset of well-defined attributes pertinent to the GFIPM concept • Represented by a set of overview documents
Federation Profile	• A profile of the conceptual model that addresses the needs of a specific federation instance • Places subset and constraint rules on the abstract federated user and federated entity models as needed • Represented by a set of NIEM conforming subset and constraint schemas.
Federation Profile Instance	• XML instance that conforms to a specific federation profile • Encapsulates the metadata (data payload) for a specific authenticated federation user or federated entity conforming to the federation profile schema
SAML Assertion	• Acts as the transport mechanism for the XML instance between an Identity Provider and a Service Provider

Table 3: Summary of GFIPM Metadata Framework Layers

4.6.3 Overview of Metadata Content

This section of the report provides a high-level overview of the basic content and structure of the GFIPM metadata for building an assertion for ***federated users*** (referred to as GFIPM User Assertion Metadata) and ***federated entities*** (referred to as GFIPM Entity Assertion Metadata). References to "GFIPM User Assertion" or "GFIPM Entity Assertion" refer to the actual assertion on the wire, at which time it is an encoding of the metadata in an assertion transport protocol such as SAML. Additional detailed views, definitions, and schema artifacts can be found in the GFIPM Metadata, version 0.4 (see Appendix B).

4.6.3.1 Basic Terminology

This section provides some basic terms that are useful for understanding the metadata model that follows in the next sections of the report.

- ***Identity*** – A unique name corresponding to a "real world" user or entity. Since the legal names of persons are not necessarily unique, the identity of a person must include sufficient additional information (for example an address, or some

unique identifier such as an employee or account number) to make the complete name unique with in the domain of an Identity Provider and federation.
- ***Credential*** – An object that authoritatively binds an identity (and optionally, additional attributes) to a token possessed and controlled by a person or entity.
- ***Token*** – Something that the claimant (user/entity) possesses and controls (typically a key or password) used to authenticate the claimant's identity.
- ***Electronic Identity*** – A unique electronic representation of a user's identity issued by an Identity Provider which is security mechanism specific. Embodies the user identity, credential, token, and security mechanism and practices used.
- ***Assertion*** – A statement from a verifier (Identity Provider) to a relying party (Service Provider) that contains identity information about a user or entity. Assertions may also contain other verified attributes.

4.6.3.2 GFIPM User Assertion Metadata Contents

The GFIPM user assertion metadata constitutes a statement from an IDP to an SP that contains identity information and other attributes about an authenticated user in the federation. A typical user assertion contains the following categories of information about a user.

- ***Identification***
- ***Certifications and Memberships***
- ***Contact Information***
- ***Organizational Affiliations***
- ***Authorization Context***
- ***Electronic Identity***
- ***Authentication Context***

Please see the GFIPM Metadata Package 0.4 (see Appendix B) for more detailed information about the specific metadata elements in a GFIPM user assertion.

4.6.3.3 GFIPM Entity Assertion Metadata Contents

The GFIPM entity assertion metadata constitutes a statement from an IDP to an SP that contains identity information and other attributes about an authenticated application, service, or device in the federation. A typical entity assertion contains the following categories of information about an entity.

- ***Identification***
- ***Contact Information***
- ***Organizational Affiliations***
- ***Authorization Context***
- ***Electronic Identity***
- ***Authentication Context***

Please see the GFIPM Metadata Package 0.4 (see Appendix B) for more detailed information about the specific metadata elements in a GFIPM entity assertion.

4.6.3.4 Noteworthy Metadata Elements

The previous two sections provide a high-level overview of the GFIPM metadata model from the perspective of content. This section highlights several specific metadata elements that bear mention due to their importance or uniqueness within the GFIPM federation.

1. ***Federation ID*** – This is a globally unique and persistent identifier for a user or entity. It must conform to a standard format that includes a federation part and a local part. The federation part consists of the IDP name, and the local part is in a format specified by the IDP.
2. ***Local User ID*** – This is a locally significant, persistent, and unique ID (such as a local logon ID) that identifies a user or entity within an IDP.
3. ***Sworn Law Enforcement Officer (SLEO) Indicator*** – This is Boolean attribute that indicates whether a user is a SLEO. It is used often by federation applications for access control decisions.
4. ***Public Safety Officer Indicator*** – This is another Boolean attribute – similar to the SLEO Indicator – that indicates whether a user is a public safety officer (e.g. EMT, firefighter, etc.)
5. ***User Home Data Access Privileges*** – This is a set of Boolean attributes defining whether a user has access to certain categories of information within his home organization. Information categories include criminal intelligence data, criminal history data, criminal investigative data, etc.
6. ***Identity Provider*** – This is a string that identifies which IDP a user belongs to. The string identifying a user's IDP is the same as the federation part of that user's Federation ID. (See the first item in this list above.)
7. ***Electronic Identity Type*** – This is a codified attribute that describes the type of electronic identity that the user has with his IDP, e.g. software certificate, hardware certificate, username/password, etc.
8. ***Identity Proofing Assurance Level*** – This is a codified attribute that describes the level of assurance during the process of vetting the user's identity when he obtained his electronic identity with his IDP. The attribute can take on four values – Level 1 through Level 4 – which correspond to the four identity proofing levels defined in the NIST Electronic Authentication Guideline.
9. ***Electronic Identity Assurance Level*** – This is a codified attribute that describes the level of assurance during the transaction by which the user authenticated with his IDP. The attribute can take on four values – Level 1 through Level 4 – which correspond to the four electronic authentication assurance levels defined in the NIST Electronic Authentication Guideline.
10. ***Authentication Event Date/Time*** – This attribute contains the date and time at which the user authenticated with his IDP.
11. ***Authenticated Client IP Address*** – This attribute contains the IP address of the host from which the user's web browser authenticated with his IDP.

4.6.4 Metadata Development Process

The development process for the GFIPM Assertion and the testing of the GFIPM model was based on a limited scope. The primary focus of the development process was on the collection of attributes (metadata) required to support the GFIPM use cases and specify federated users and federated entities in accordance with known and applicable industry standards. The scope of the development process was initially limited to responses provided by GSAC survey participants.

The first level of development was the identification and collection of metadata based on a survey of GSAC members and the systems that they represent. This initial set of metadata was grouped and harmonized among the independent responses and then mapped to NIEM 0.3 as the base vocabulary. This resulted in a strawman set of metadata which was then vetted by the entire GSAC, a separate GSAC GFIPM tiger team, and the DOJ/DHS GFPIM demonstration project participants. This resulted in the GFIPM Metadata Package Version 0.2, which was the first version of the metadata to be used in the demonstration project federation. Lessons learned from this project have been captured and incorporated into the GFIPM Metadata Package Version 0.4, which is the current version and provides the basis for further development and expanded vetting.

The next level of development process will build this metadata set into the form of a technology standardized "assertion" format (i.e. SAML) resulting in a GFIPM Assertion. Several different techniques for encoding GFIPM metadata into SAML assertions have been identified and documented as part of the GFIPM Metadata Package Version 0.4. Lessons learned from the demonstration project and feedback from the broader community will lead to specific recommendations and standards for the GFIPM Assertion.

The distinction of the attributes available within the GFIPM Metadata is specific to the requirements for describing either a user or a federation entity. In other words, the GFIPM Metadata supports both the necessary attributes for system-to-system, system-to-user, and/or user-to-user contexts for information sharing. The profile, or use case, of either creating a federation entity assertion or a user assertion are subsetted within the GFIPM Metadata. Separately, the GFIPM Assertion specification will detail the attributes and requirements for SAML encoding, binding, and assertion transport for either assertion use case. However, beyond the context of the GFIPM Metadata, it should be noted that a comprehensive collection of all security metadata requirements needed for the justice or national information sharing community including privacy, Service-Oriented Architecture (SOA), networking, other layers of the security stack, and a comprehensive security process, were considered outside the scope of this initial survey and draft specifications.

Table 4 provides a summary of the metadata development process.

Process	Description	Status/Due
Call for Data Requirements	An initial set of data requirements was solicited from the Global Security Architecture Committee participants via	***Done Aug 2005***

	a simple set of questions. Many of the GSAC participants represent systems and networks with embedded identity and privilege management requirements in operation today.	
Collect and Compile Survey Results	Submissions were transformed to a consistent machine readable format for manipulation and analysis.	***Done Sept 2005***
Validate Data Requirement Submissions for Completeness and Clarity	All submitted data requirements were inspected for association with unambiguous, semantically precise definitions. Also, representation and content were determined for each data requirement (e.g. numeric, text, code, etc.). NIEM core representation types provided the basis for typing data components.	***Done Oct 2005***
Supplement with Attributes from Applicable Standards as Required	The following standards were considered for the extraction of data requirements: ▪ Security Assertion Markup Language (SAML) 2.0; ▪ Liberty Alliance Identity Service Interface Specifications (ID-SIS) 1.0; ▪ Liberty Alliance Identity Federation Framework (ID-FF) 2.0; ▪ Internet Engineering Task Force (IETF) inetOrgPerson.	***Done Oct 2005***
Logically Group Attributes into Categories	Analysis of the submitted data requirements drove the process of categorizing and aggregating attributes.	***Done Nov 2005***
Develop Strawman	A strawman was developed from the superset of candidate attributes from survey submissions and applicable standards for each of the categories defined above.	***Done Dec 2005***
Vet and Refine Strawman	The initial vetting of the strawman was through the Global Security Architecture Committee members. Survey participants validated mappings between their submissions and the resultant strawman for semantic consistency. Feedback was solicited and incorporated. The strawman was updated and provided as a recommendation by the GSAC for review and vetting to a broader Global community.	***Done Jan 2006***
Harmonize with current version of GJXDM/NIEM	Once consensus has been reached on the data requirements, semantics, and representation of the GFIPM Security Assertion, it will be semantically and structurally harmonized with the current version of the GJXDM/NIEM. Metadata which is currently in the GJXDM/NIEM will be mapped while new data requirements will be submitted for inclusion in a future version of the GJXDM/NIEM. This will be an ongoing activity until the convergence and stabilization of NIEM and GJXDM. Data requirements and associated attributes derived from existing standards (e.g. SAML) will be referenced in the GFIPM Assertion for completeness but not duplicated in the GJXDM/NIEM.	***Ongoing Thru 2007*** ***Initial: Feb 2006***
Advanced Vetting of the GFIPM Metadata	A GSAC representative "Tiger Team" was established for continued vetting and design process of the GFIPM metadata and assertion specification development. An initial vetting session was conducted in March 06. Broader vetting of the GFIPM metadata is required prior to making a full recommendation for implementation. The GSWG and Global community will continue to serve as the vehicle for this expanded vetting of the GFIPM.	***Ongoing Thru 2007*** ***Initial: March 06***
Incorporate	The current version of the GFIPM Metadata is 0.4.	***Ongoing***

Feedback and Iteratively Refine and Publish "Draft" GFIPM Metadata Packages	Based on feedback from the vetting process and lessons learned from the DHS/DOJ demonstration process, additional versions of the GFIPM Metadata package will be published.	***Thru 2007*** ***Initial: April 2006***
Develop and Vet GFIPM Assertion Specification	A set of alternatives for encoding the GFIPM Metadata in SAML, along with pros and cons, have been identified and documented. The GFIPM demonstration project participants have reviewed and deliberated on these alternatives and selected one of these alternatives for use in the demonstration project. Lessons learned will be captured from the demonstration project leading to further specification and recommendations for the GFIPM Assertion. Additionally, specific encoding techniques have implications with regard to COTS product support. Some limited COTS testing is being conducted as part of the demonstration product. Lessons learned will be captured and provided to the GSAC/GSWG for consideration.	***Ongoing Thru 2007*** ***Initial: September 2006***

Table 4: Overview of GFIPM Metadata Development Process

4.7 Miscellaneous

This section describes several miscellaneous but noteworthy architectural details about the GFIPM federation. Section 4.7.1 describes some implications of the GFIPM architecture for firewall administration, and Section 4.7.2 lists some basic requirements of web browsers used within the federation.

4.7.1 Firewall Considerations

The GFIPM demo federation does not require the use of any network protocols other than HTTP over SSL/TLS (also known as HTTPS). As a rule, federation network traffic travels over the standard HTTPS port (443). There is, however, one non-standard firewall issue that caused some administrative challenges during the deployment of the federation infrastructure. An IDP in the GFIPM federation is required to listen on two HTTPS ports simultaneously: one port for handling browser authentication/SSO transactions as part of the SAML/Shibboleth Web Browser SSO Profile (see Section 4.3.2.1), and another port for handling back-channel attribute query transactions with SPs as part of the SAML Attribute Query Profile (see Section 4.3.2.2). Under certain circumstances – specifically if the IDP is using client certificate authentication for its SSO system – it is not desirable for the IDP to accept connections from both browsers and SPs on the same port.[4] Therefore, the firewall protecting the IDP must make two HTTPS ports available. Typically, a non-standard port such as port 8443 is used to handle back-channel traffic from SPs; however, the choice of port number for back-channel communication is arbitrary. Note that this issue is related to the specific SAML usage profile employed by GFIPM, and is not affected by the choice of SAML implementation.

[4] If an IDP were to accept connections from both browsers and SPs on the same SSL channel, then the client certificates used by browsers and SPs to authenticate with the IDP would need to be managed within the same PKI. Configuring a PKI in this manner would be considered unwise by most PKI administrators.

4.7.2 Web Browser Considerations

Typically, the choice of web browser is a personal decision made by a user according to personal preferences and their end device constraints (e.g. mobile access). In an ideal scenario, GFIPM would be able to extend this liberty to federation users; however, in reality, this is often not possible. Throughout the demo project, participants have gained several important insights into the web browser options that are available to federation users. These insights are listed here.

1. Any web browser used in GFIPM must be able to support HTTPS (HTTP over SSL and TLS), as well as HTTP redirection. All reasonably modern browsers (released in the last five years) can do this. However, as participants discovered during the project, not all modern browsers are configured by default to use TLS. While this issue is relatively easy to solve in the browser (by simply changing the browser's configuration), it can cause usability problems nevertheless, because the problem typically manifests as a cryptic web server error that is not easily identifiable as a browser configuration problem. This problem specifically affects Internet Explorer (IE) version 6. The simplest work-around for the problem is for the user to upgrade to IE version 7, which is configured to use TLS by default.
2. The only other constraints imposed by the GFIPM federation are specific to the limitations of certain applications within the federation. For example, if a specific application required the use of Internet Explorer (IE) by its users prior to its becoming federation-enabled, then it will almost certainly require IE for federation users as well.
3. Participants may place further constraints on browsers. For example, CISA users must use IE for authenticating with the CISA client certificate SSO system at CISA's IDP.
4. Interoperability problems may arise if the browser required by a user's IDP is Browser X (e.g. IE) and the browser required by an SP that the user wishes to access is Browser Y (e.g. Mozilla Firefox). But this scenario has not happened yet. Typically, if an application requires a specific browser, then the browser required is IE. There are no known instances in the federation in which an application has a browser-specific requirement for a browser other than IE.

5 Project Execution and Timeline

This section documents activities, decisions, and outcomes of the demonstration project over the period of performance. This chronological walk-through is intended to provide an understanding of how the project was managed, as well as a context for interpretation of lessons learned presented in Section 7. Minutes, presentations, and whitepapers from regularly held project conference calls and face-to-face meetings have been used as the basis for the content of this section and are maintained on the GFIPM participant project website. Links to this material can be found in Appendix B.

Table 5 contains a timeline of the demonstration project. The remainder of this section describes each major phase of the project in more detail.

Phase	Activities	Dates
1. Project Initiation	▪ Develop project plan (delivered 15 Dec 2005). ▪ Develop and achieve consensus on demonstration scenario. ▪ Develop and submit participant grant requests. ▪ Develop technical objectives. ▪ Develop draft GFIPM attribute data dictionary for demonstration. ▪ Collect and assess participant platform requirements. ▪ Conduct project participant conference calls.	***15 Oct 2005 to 31 Dec 2005***
2. Infrastructure Establishment	▪ GFIPM Middleware implementation. ▪ Begin development of participant federation service provider web portals. ▪ Finalize GFIPM attribute data dictionary for project. ▪ Establish GFIPM testbed. ▪ Develop IDP/SP Interface Specification / Integration Guide. ▪ Conduct project participant conference calls. ▪ Develop class on Shibboleth and SAML to present at face-to-face meeting. ▪ Conduct face-to-face meeting (held April 12th and 13th in Atlanta).	***1 Jan 2006 to 30 Apr 2006***
3. Participant IDP/SP Development and Integration	▪ Complete development of participant federation service provider web portals. ▪ CISA IDP and SP interface development. ▪ JNET IDP and SP interface development. ▪ RISS IDP and SP interface development. ▪ GTRI provided technical assistance to IDP / SP interface developers. ▪ Continue to refine and vet GFIPM metadata. ▪ Identification and approval of participant shared resources. ▪ Development of business rules in terms of GFIPM metadata for participant shared resources. ▪ Develop test plan. ▪ Conduct integration and testing with project personnel. ▪ Conduct conference calls with project participants to discuss and address issues.	***1 May 2006 to 31 Jan 2007***
4. User Testing, Evaluation, and Infrastructure Refinement	▪ Identify participant users and conduct field testing ▪ Collect feedback and lessons learned ▪ Second GFIPM Face to Face with participants to develop lessons learned, next steps, final report content (held Feb 12th and 13th in Atlanta). ▪ Incorporate refinements as appropriate. ▪ Build SAML 2.0 and Shibboleth 2.0 testbed. ▪ Conduct demonstrations for stakeholder and funding agencies. ▪ Identify next steps. ▪ Develop final report.	***1 Feb 2007 to 30 Jun 2007***

Table 5: GFIPM Demonstration Project Timeline

5.1 Project Initiation

The first phase of the project (Project Initiation) began in October 2005 and ran through December 2005. During this phase, participants focused on setting up and defining

important aspects of the project. The following bullet list summarizes the highlights of Phase 1.

- ***Project Initiation Conference Call*** – The first project participant conference call was held on 11/3/2005. Conference calls have been ongoing throughout the project since then, approximately on a biweekly basis. There have been about 30 participant conference calls during the project.

- ***Submission of Participant Grant Applications*** – GTRI helped participants write grant applications for their part of the funding by providing them with a template grant submission package.

- ***Development of Functional and Technical Objectives*** – Participants developed and built consensus on a set of functional and technical objectives to be achieved during the project.

- ***Development of Demonstration Scenario*** – Participants developed and built consensus on a demonstration scenario. Possible options included:

    1. A portal-based approach in which each participant created a portal through which federation users could access its resources;
    2. A task-based approach in which the entire federation would be focused on solving a specific business problem;
    3. A community-based approach in which the entire federation would be focused on solving a problem for a specific group of users.

    After some discussion, participants chose the portal-based approach, as it was the most realistic for the time frame and scope of the project. Also, participants chose to focus specifically on law enforcement information sharing for sensitive-but-unclassified (SBU) data. As a result of this decision, and the realization that participants would need to proxy applications that they could not physically touch or alter, GTRI began at this time to investigate techniques for reverse proxying web applications so their content could be accessed via federation-aware portals.

- ***Development of Project Timeline*** – Participants developed and built consensus on a basic timeline for the project.

- ***Development of Participant Website*** – GTRI developed a basic web site for participants to use throughout the project. This web site is called the GFIPM Participant Portal, and it is still maintained by GTRI. It contains all written materials, including software and documentation, developed during the project. It is available at http://gfipm.net/participants/, but it is password-protected and accessible only by project participants.

- ***Analysis of Platform Requirements*** – GTRI performed an analysis of participants' required operating system and web server platform support for the

project. This analysis guided and informed later decisions regarding the development of tools and technical documentation for the project.

- ***Choice of Shibboleth 1.3 and SAML 1.1*** – Participants decided to use Shibboleth 1.3 and SAML 1.1 as the federated identity middleware infrastructure for the project. Sections 3.1.4 and 3.1.5 contain a summary discussion about why these choices were made.

- ***Initial Development of SAML and Shibboleth Training Materials*** – GTRI began to develop training materials on SAML and Shibboleth that would be used in Phase 2 of the project at the first GFIPM face-to-face meeting.

- ***Initial Development of Metadata Attribute Dictionary*** – Participants began developing a straw man metadata attribute dictionary that would eventually evolve into the GFIPM metadata model that currently exists. Early discussions about the metadata were focused on how the attributes would be used (authorization, auditing, etc.), how the metadata would be defined and developed, and what data categories would be supported by the metadata. (A decision was made to begin by using the data categories used by CISA: criminal intelligence, criminal history, criminal investigative, etc.) Also, participants agreed at this time that it would be beneficial to eventually roll the GFIPM metadata into the NIEM standard.

- ***Submission of Formal Project Plan*** – Participants developed a formal project plan and submitted it to DHS.

- ***Initial Experimentation with Reverse Proxy Techniques*** – During initial exploration of the federation concept, participants realized that many potential federation resources and applications would not be modifiable by participants. Bringing those resources online in the federation would require some type of proxy solution that would make them available to federation users while not requiring that the resources themselves be modified. GTRI began to conduct experiments with reverse proxy solutions at this time to determine the scope challenge represented by this issue. The reverse proxy solution that was eventually developed is described in Section 4.2.4 of this report.

5.2 GFIPM Infrastructure Establishment

Phase 2 of the project focused on the establishment of the basic federation infrastructure. This phase ran from January 2006 through April 2006. Highlights of this phase include the following.

- ***Metadata Tiger Team Meeting*** – GTRI hosted a tiger team meeting on March 7-8, 2006 to review and flesh out the metadata straw man model. The team included GTRI personnel and domain experts from the justice community. By the end of the meeting, the tiger team had created a fairly robust conceptual metadata

model. Based on this model, GTRI asked GFIPM participants to begin working out an appropriate federation profile for the demo project.

- ***Continued Development of Metadata Attribute Dictionary*** – Participants continued to develop the straw man metadata attribute dictionary from the previous project phase. Specific actions and decisions that occurred relative to the metadata model during this phase of the project included the following.

    1. Participants realized that reaching agreement on common definitions for federation user roles would be very difficult, and that it would be better to avoid the use of roles in favor of commonly understood, basic attributes about users within the metadata model.

    2. Participants decided to augment the data category aspect of the metadata model by adding action-based permissions such as search/query, write, submit/post, and edit for each category. In addition, permission to perform each action was divided into two parts: "on behalf of self" and "on behalf of organization". During the process of discussing this data model, it became clear that in the federation's metadata model, these permissions must act as suggestions from a user's home organization, and not as federation-wide facts that must be honored by all SPs. An SP must always be free to choose its own access control policies, and the metadata model is intended to support the implementation of those policies where necessary.

    3. It became clear to participants that the metadata model would grow very large over time, and that many attributes in the model would be optional or unnecessary for certain applications or communities. One idea for handling this issue was to create a multi-layer metadata model in which there is a large set of optional attributes and a process by which the set can be narrowed into a subset that applies to a specific purpose. This idea evolved into the current model of a conceptual model and a federation profile, as presented in Section 4.6.2.

    4. Participants began discussion about how to encode GFIPM metadata in SAML for the purpose of passing it from an IDP to an SP. It was clear even at this phase of the project that SAML encoding of metadata was a complex topic without any obvious best solution, and this question is still under investigation at the time of this writing (June 2007). Section 4.6.2.4 provides some details about the specifics of this situation and helps to illustrate why it is a challenging issue.

    5. Participants decided that the metadata strawman from GFIPM would be presented to Global as a recommendation, with the assumption that Global would improve and expand upon it as needed.

- ***Completion of GFIPM User Assertion 0.1*** – Participants completed version 0.1 of the GFIPM metadata model (which was at that time called the "GFIPM User Assertion") and submitted it to Global Security Working Group (GSWG) in March 2006. The model included the four-layer architecture (conceptual model, federation profile, profile instance, and SAML assertion) that is described in detail in Section 4.6.2, and it also addressed the seven basic types of metadata about users that are identified in Section 4.6.3.2.

- ***Development of a Technical Implementation Plan*** – Participants developed and built consensus around a four-stage plan for implementing the federation infrastructure at a technical level. The four stages of the plan were as follows.

    1. Deployment of Basic Infrastructure with Shibboleth Middleware
    2. Integration of Live Resources and User Bases
    3. Migration to SAML 2.0
    4. Interoperability Testing with COTS SAML 2.0 Products

    Note that this technical implementation plan is not the same as the high-level GFIPM project plan, which is summarized throughout this section and illustrated in Table 5. Stages 1 and 2 of the technical implementation plan do, however, correspond to phases 1 and 2 of the GFIPM project plan. Also note that the implementation plan was based on the assumption that a new version of Shibboleth software, featuring support for SAML 2.0, would be available in mid-late 2006. But since the delivery of the next-generation Shibboleth software has slipped by about 12 months, only the first two stages of the implementation were completed.

- ***Development of Testbed Infrastructure*** – As part of the technical assistance process, GTRI developed and deployed an initial testbed infrastructure for the federation. This included two reference IDPs (one on a Windows platform and one on Red Hat Enterprise Linux (RHEL)), two reference SPs (again, one on Windows and one on RHEL), and a Where-Are-You-From (WAYF) service. These components were made available to participants for testing purposes as they worked through Stage 1 of the technical implementation plan.

- ***Completion of Stage 1 of the Implementation Plan*** – GTRI developed instructions for participants to follow during Stage 1 of the implementation plan, and participants began deploying their basic IDP and SP software. By mid-April, most of the Stage 1 implementation work was complete. At that point, the entire federation consisted of stubs, as nobody had yet begun to work on connecting live users or resources to it. At this point, participants began to realize that the task of performing the necessary integration work (which included connecting existing authentication systems, LDAP directories, and protected resources to the federation middleware) was inherent to any SAML federation and depended little on the choice of SAML middleware.

- ***Initiation of Stage 2 of the Implementation Plan*** – As discussed previously, participants created a basic implementation plan for building out the GFIPM federation infrastructure. During this phase of the project, GTRI began to develop implementation instructions describing how to integrate existing authentication systems, LDAP directories, and protected resources with the Shibboleth middleware that had already been implemented.

- ***GFIPM Face-to-Face Meeting*** – Participants planned and held the first GFIPM face-to-face meeting. The meeting was hosted by GTRI in Atlanta on April 12-13, 2006, and it focused on issues surrounding the implementation of the federation infrastructure and integration of the infrastructure with existing user bases and applications. During the meeting, GTRI presented a set of training materials on SAML, Shibboleth, and IDP/SP integration points. These materials had been developed prior to the meeting, and they served as a starting point for technical discussions over the two-day session.

- ***Decisions about Resources to be Shared*** – Participants began to explore the question of what resources they would share with federation users via their portals, and also which of their own users they would bring into the federation. One of the major challenges of this process was that some participants (specifically CISA and JNET) act as brokers for data that they do not own; therefore, they had to obtain permission to share data with the federation before they could commit to share it.

- ***Exploration of FIPS 140-2 and TLS*** – Participants explored issues related to the use of the Internet as the connectivity backbone of the federation. It became clear during this process that details such as strength of authentication mechanism and strength of data encryption would be important factors in deciding what resources a user is permitted to access. At this time the participants began to address questions related to making the federation compliant with FIPS 140-2 cryptographic specifications, with regards to the establishment of cryptographic session keys. It was eventually determined (1) that FIPS 140-2 compliance could be achieved if all IDPs, SPs, and web browsers in the federation used TLS encryption, and (2) that it was technically possible and feasible to implement TLS encryption for all IDPs, SPs, and web browsers.

- ***Exploration of Non-Standard Firewall Configuration Issues*** – Participants discovered a potential challenge related to the use of HTTPS on a non-standard port for the IDP's back channel. The issue, in brief, is that under certain circumstances an IDP must be configured to accept HTTPS connections on two different ports, to accommodate both its single sign-on endpoint and its attribute authority endpoint. (See Section 4.7.1 for additional background information about this issue.) Each participant conducted discussions with their internal IT security managers to determine whether it would be possible to open non-standard ports in their firewalls for this purpose. JNET and RISS both encountered internal resistance to this question, so as a work-around, the decided to configure their

IDPs to accept HTTPS connections on port 443 (the standard port for HTTPS) and port 80 (which is typically used for HTTP connections, but which was unused for their IDP hosts). This solution is not ideal, but it is sufficient for the purposes of the demo project.

- ***Initial Development of Integration Point APIs/Toolkits*** – GTRI began initial design discussions about an API or toolkit that would assist participants in building metadata instances on their IDPs from local (LDAP) data repositories, as well as simplifying the parsing and interpretation of the GFIPM metadata on their SPs for consumption by protected resources and applications. Subsequent discussion with participants confirmed that such a toolkit would be of value for IDPs, but not for SPs.

- ***Continued Experimentation with Reverse Proxy Techniques*** – GTRI continued to conduct experiments with various web proxy configurations for the purpose of developing a proxy solution that would allow a wide range of integration options for participants to choose from when they connect their systems to the federation. Section 4.2.4 of this report outlines the basic architecture of the solution that was eventually developed.

5.3 Participant IDP/SP Development and Integration

The focus of phase 3 of the project was the integration of existing user bases and live resources with the basic federation infrastructure that had been previously developed. This phase of the project began in May 2006 and ran through January 2007. Highlights of this project phase are summarized below.

- ***Completion of GFIPM Metadata 0.2*** – Participants collaborated to further refine the GFIPM metadata model, and GTRI subsequently released the updated model as version 0.2. This version of the metadata was used as the federation-wide metadata standard from the time of its release in June 2006 until it was replaced by metadata version 0.4 in March 2007.

- ***Completion of Stage 1 Pairwise Testing*** – As a final step of the first stage of the technical implementation plan, GTRI coordinated a series of pairwise tests between each IDP and SP within the federation to make sure that basic SAML-level connectivity was working properly for all IDPs and SPs. At this point, the federation consisted of five IDPs and five SPs, plus one WAYF service. Figure 12 in Section 6 illustrates the structure of the federation at this point.

- ***Completion of Stage 2 Deployment Kit*** – Based on initial design discussions and participant feedback discussed previously, GTRI completed the stage 2 deployment kit (including software and documentation) for participants to use during the process of bringing online live users and resources in the federation. Most of the content in this deployment kit consists of software and instructions related to the creation of GFIPM metadata XML structures at an IDP. For the

other two integration points – SSO system and protected resources – GTRI was unable to provide any software tools. The SSO integration point is relatively simple, so the participants didn't need much help with it. And protected resources are so diverse and complex that GTRI could not do much to help participants other than to offer assistance on an as-needed basis.

- ***Development of GFIPM User Portal*** – During this phase it became clear that for usability's sake, federation users would need some index or directory of available resources, as well as a central location in which to find documentation such as frequently asked questions. To fill this need, GTRI collected information about the resources that each participant planned to offer to federation users. GTRI also collected data about the access control policies that participants were planning to enforce on their resources. This information has been consolidated and is available online at the GFIPM User Portal, which is at http://gfipm.net/users/.

- ***Investigation of Metadata 0.2 Problems*** – During the process of federation-enabling protected resources, participants began to realize that some of the definitions of metadata elements were not as clear as they needed to be for the federation. For example, the ***Sworn Law Enforcement Officer (SLEO) Indicator*** is a Boolean metadata element that indicates whether a user is a SLEO, but the metadata version 0.2 did not contain any clear guidance as to when it is acceptable for an IDP to assert this indicator to be true for a user. Other examples of definitional ambiguity were identified as well, and this realization led to the eventual development of metadata versions 0.3 and 0.4, which contained successively more clarity in their specifications.

- ***Development of Consensus on the GFIPM Concept*** – Having acquired enough experience to understand the implications of the basic federation concept that had been chosen, participants built a consensus around the opinion that the GFIPM concept is a sound approach to information sharing in the justice domain. This basic approach was approved in September 2006 at the GSWG meeting.

- ***Development of GFIPM Metadata 0.3*** – GTRI released metadata version 0.3 in September 2006. It was an incremental improvement over version 0.2, designed to be more readily understood by external audiences. This version was disseminated to the GSWG, but was never used within the GFIPM federation.

- ***Development and Implementation of Stage 2 Test Plan*** – As stage 2 of the technical implementation plan came to an end, it became clear that another round of comprehensive pairwise testing between participants would be advantageous as a means of discovering and correcting errors prior to making the federation available to users. Towards this end, GTRI developed a stage 2 test plan, and participants implemented the plan over the course of several months. The first step of the test plan was to manually validate metadata assertion instances generated by each participant's IDP to ensure that they conformed to the

federation profile XML schema. The second step of the plan was to verify that each resource's access control policy was operating as advertised.

- ***Pre-Planning for User Rollout*** – Near the end of this project phase, participants began to make plans for bringing users online, providing assistance to them when necessary, and collecting feedback from them about their experiences. GTRI assisted in this process by developing template documents for training materials and questionnaires that could be used with federation users. The actual user rollout did not occur until phase 4 of the project, and results of user testing can be found in Sections 6.2.4, 6.3.4, and 6.4.4 of this report.

- ***Investigation of Resource Control Limitations in a Federated Environment*** – During the process of federation-enabling protected resources, some participants needed to secure permission from resource owners before making resources available to federation users. This process was challenging at a political level for two reasons. First, the federation concept is a new idea that has not yet gained wide acceptance in the justice community. Second, the federated approach can make resources available to a very large user base in a short amount of time, and this act of rapidly scaling up a resource's user base can cause concerns for the resource owner in terms of help desk costs, information leakage risks, etc. Through this process, participants have learned valuable lessons about how to present the federation concept to resource owners in the best possible light.

- ***Investigation of Policy Limitations on Testing of Resources*** – During the process of performing pairwise tests on live federation resources, it became clear that engineers would be in violation of participants' information sharing policies if they performed tests on certain resources. Even authorized users would be in violation of policy if they performed certain queries without having a legitimate law enforcement-related reason for doing so. These strict regulations caused some delays in the stage 2 testing process, and are likely to continue to be a source of friction during future tests. The most effective means for getting around these limitations are to (1) set up a non-live test resource that behaves like the live resource and use it for testing in lieu of the real thing, or (2) arrange with the resource owner to perform a specific set of operations or queries that will be recognized and treated as tests.

- ***GFIPM Bugzilla Bug Tracking System*** – GTRI set up a Bugzilla bug tracking system for participants' technical staff to use for collaboration on outstanding problems and issues. This system is online at https://bugs.gfipm.gtri.gatech.edu/, but it is available only to project participants at this time.

- ***Investigation of TLS Support Problems in Internet Explorer*** – As participants worked through the stage 2 test plan, an issue emerged concerning the use of TLS in the Microsoft Internet Explorer version 6 (IE6) web browser. As discussed in Section 5.2 under the item heading "Exploration of FIPS 140-2 and TLS", it is important that support for TLS be available throughout the federation in all web

servers (IDPs and SPs) and in all web browsers. Careful investigation of this specific issue revealed that while IE6 does contain support for TLS, it is not configured by default to support TLS connections. This could become a problem, since most federation users are likely to use IE; however, there is a simple workaround for the problem: install Internet Explorer version 7, which is configured by default to handle TLS connections. GTRI wrote a short document containing lessons learned about this issue.

- ***Initial Discussion of Metadata Changes to Accommodate Resource Needs*** – As CISA began the process of defining its access control policies for the resources that it would bring online in the federation, it became clear that the current metadata model (version 0.3) was inadequate for providing CISA's SP with the information that would be required by some of CISA's resources. Specifically, CISA's resources required information about an IDP's level of assurance for a user's electronic authentication, as well as the IDP's level of assurance in the identity proofing technique used when a user received his electronic account credentials. Discussion about these shortcomings led to the development of a set of requirements for metadata changes so that the GFIPM metadata would be capable of expressing this information. This change request eventually resulted in metadata version 0.4.

- ***Development of Metadata Compatibility Matrix*** – As participants started to define access control policies for their resources and make these policies available to each other, questions arose about how much "compatibility" would exist between federation users and federation resources. In other words, for a given resource with a given access control policy, how many IDPs in the federation would be able to assert the required metadata attributes required by it for their users? This question is of critical importance to the federation, because if a basic level of compatibility does not exist between the attributes that an IDP can provide and that attributes that an SP requires, then that resource is unusable by a significant number of federation users. This issue led GTRI to create a ***metadata compatibility matrix***, which is a table containing a pairwise analysis of this basic question across each pair of an IDP and a resource.

- ***Investigation of Security Vulnerabilities in the Federation*** – During the process of bringing protected resources online, it became clear that having reference IDPs in the federation along with live resources was risky from the standpoint of data leakage. In their role as reference IDP implementations, the Windows Reference IDP and the Red Hat Enterprise Linux Reference IDP were intended to be semi-public and available for all participants to use during testing. Credentials for test accounts on the reference IDPs had already been made widely available, and there was serious concern that having these reference IDPs and test accounts in a position to potentially access live sensitive data constituted a major risk. For this reason, GTRI took the reference IDPs offline in fall 2006. This led to the realization that the GFIPM project needed to have an entire reference federation in which to perform testing without affecting live resources or live users. Details

about the initial deployment of this reference federation are available in Section 5.4.

- ***Discussions with Shibboleth 2.0 Development Team*** – During this phase, GTRI began a series of discussions with the Shibboleth 2.0 development team regarding the timeline for the development and release of Shibboleth 2.0. (With support for the SAML 2.0 standard, Shibboleth 2.0 is anticipated to provide a relatively clean upgrade path for the GFIPM federation in the process of moving from SAML 1.1 to SAML 2.0. Therefore, it has been important to GFIPM participants since the GFIPM project began.) In an effort to help expedite the Shibboleth 2.0 development cycle as much as possible, GTRI offered to assist the Shibboleth development team in development and/or testing. The Shibboleth team accepted the offer to help with testing, and that testing is currently ongoing, with Shibboleth 2.0 expected to be released in Fall 2007.

- ***Discussions with Potential New Federation Members*** – During this phase, participants began thinking beyond the current demonstration project and towards a fully operational federation. Towards this end, GTRI, GFIPM participants, and the Global GFIPM Delivery Team, began discussions with several potential new federation members as part of an initial outreach initiative for the purpose of growing and leveraging the federation beyond its current demonstration status. More details about potential new federation participants are available in Section 5.4.

- ***Investigation of Federated User Support Challenges*** – As participants began to think ahead to a point at which the federation would be available to many users, they realized the large scope of potential challenges related to providing effective user support in a federated model. A federated support model must leverage the user support infrastructure already in place for each participant, but also enable support personnel to collaborate when necessary to address federation specific issues. Section 7.5 contains some initial insights into this challenge.

- ***Investigation of Auditing for Proxied Resources*** – One requirement each participant needed to satisfy when bringing live resources online was how to capture audit data about each event in which a user accessed a resource. This presented a challenge for proxied, non-federation-aware resources, because generating a complete end-to-end audit log for a proxied resource requires reconciliation of audit logs from the proxy with audit logs from the resource itself. More information about this topic is available in Section 7.4.

5.4 User Testing, Evaluation, and Infrastructure Refinement

The fourth and final phase of the GFIPM demo project included user testing, evaluation, and refinement of the federation infrastructure, as well as discussion and planning related to next steps of the project. Phase 4 of the project began in February 2007 and ended in June 2007. This section summarizes the highlights of this phase.

- ***Second Face-to-Face Meeting*** – Participants met for another face-to-face meeting in Atlanta on February 12-13, 2007, to kick off the final stage of the demo project. Items addressed during the meeting included user testing, SAML encoding alternatives for GFIPM metadata, and a thorough discussion of lessons learned during the project. Much of the output of this meeting can be seen in Section 7 of this report.

- ***Initial Deployment of GFIPM Reference Federation*** – GTRI began to build the basic infrastructure for the GFIPM Reference Federation, which is a SAML 2.0 testbed to be used for the purpose of performing tests and experiments without disturbing the live GFIPM federation. When complete, the reference federation will contain reference implementations of SAML 2.0 IDPs and SPs, as well as various COTS implementations of IDPs, SPs, and other software that requires testing as needed. Issues to be addressed and tested within the reference federation in the near term include SAML-level interoperability testing among various COTS products and questions related to GFIPM metadata encoding within SAML. Longer-term plans for the reference federation will be driven by the needs of federation members.

- ***Participation in Shibboleth 2.0 Development*** – GTRI began participating in the development process for Shibboleth 2.0, which will deliver SAML 2.0 support in a freely available, open source middleware package. GTRI's role in the project has been in a testing capacity, performing basic functionality tests of early components and versions of Shibboleth 2.0. Testing is still ongoing as of June 2007, and Shibboleth 2.0 is expected to be generally available in fall 2007. After it becomes available, GFIPM participants will begin the process of upgrading the federation infrastructure to Shibboleth 2.0 and SAML 2.0.

- ***Investigation of COTS SAML Products*** – GTRI began to investigate several COTS SAML products for the purpose validating GFIPM concepts and specifications as well as providing guidance to current and future GFIPM participants about the feasibility of using certain products within the federation. Issues to explore for each COTS product include the following:

    1. Its ability to interoperate with Shibboleth and other COTS products at a SAML level;
    2. Its ability to integrate with participants' SSO systems;
    3. Its ability to support the generation of GFIPM metadata XML structures dynamically via custom code plug-ins or APIs;
    4. Its ability to support the consumption of GFIPM metadata XML structures by federation-aware resources at an SP;
    5. Its ability to support various GFIPM metadata encoding strategies within its SAML payload.

The scope of issues to be addressed in the COTS testing process is very large, and there is still have much work remaining on this topic. COTS products that are currently being tested include Ping Identity PingFederate, Novell Access Manager, and Microsoft SharePoint with third party enable products.

- ***GFIPM Delivery Team Formation*** – On February 26-27, 2007, GTRI hosted a meeting with Global Security Working Group (GSWG) personnel to discuss the question of how to establish a GFIPM Delivery Team. The team will consist of representatives from each GFIPM participant, as well as representatives from Global. Its objectives will be to span the divide between Global and the GFIPM project, and to determine when and how to bring various GFIPM work products into the Global forum.

- ***Release of Metadata 0.4 and Investigation of the Lockstep Upgrade Issue*** – As discussed in the previous section (Section 5.3), in phase 3 of the project discussions began about metadata changes required by CISA for some of its resources. These discussions continued into phase 4 of the project, and eventually resulted in the release of GFIPM Metadata 0.4 in March 2007. Following the release of the new metadata version, participants began the process of upgrading all IDPs and SPs in the federation to use version 0.4. This was the first time that the federation had performed an upgrade from one metadata version to another, and during the upgrade process it became clear that a "lockstep" metadata upgrade process (in which each IDP and SP must upgrade to a new metadata model at exactly the same time) was very difficult – probably to the point of being an unacceptable method of metadata upgrade. Several solutions have been proposed for this problem, but at this point the issue is still open.

- ***Initial Development of Metadata 0.5*** – Participants began the process of developing the next version of the GFIPM metadata, which is tentatively being called version 0.5. When complete, this version is expected to contain several notable improvements over version 0.4, including the following:

    1. Obvious errors, including spelling errors and clear inconsistencies, will be corrected.
    2. Code tables will be introduced for metadata elements as necessary.
    3. Element definitions will be improved and expanded.
    4. Clear usage guidance will be provided where necessary, to help IDPs understand the circumstances under which it is acceptable to assert specific elements.
    5. Lessons learned from the demo project will be incorporated where necessary.
    6. The entire data model will be reconciled and based on the latest version of NIEM (NIEM 2.0 scheduled for release at the end of July 2007).

  It is expected that version 0.5 of the metadata will be released for review by late fall 2007.

- ***Initial User Testing*** – Participants began the process of recruiting and training users for the purpose of using the GFIPM federation and providing feedback about it. This first phase of user testing has been limited in scope; however, it has still yielded valuable feedback about the federation. See Sections 6.2.4, 6.3.4, and 6.4.4 for more information about the user testing done by users within CISA, JNET, and RISS, respectively.

- ***Initial Discussions about Federation Growth Strategy*** – Participants began considering the issue of how to best grow the GFIPM federation, in terms of both users and resources. During the discussion, it became clear that there exists a chicken-and-egg problem related to bringing online new users and new resources. On one hand, users and IDPs are typically not interested in joining the federation unless it can provide them with access to useful resources. On the other hand, it is difficult to justify the effort and expense required to bring new resources online if there is not a large user base ready to use them. In fact, this issue is even more subtle than the above statements would indicate, because for any given SBU application, its potential base of eligible users in the federation is not the entire federation user population, but rather the subset of federation users that would be permitted to access it based on the resource's access control policies and the users' credentials. This issue is discussed in more detail in Section 7.4.

- ***Discussions with Potential New Members*** – GTRI and the Global Delivery Team began discussions with potential new federation members. While several agencies have expressed interest in the benefits of participating in a GFIPM federation, the following potential new members and collaborators represent a short list for initial focus: Los Angeles County, CA; the Automated Regional Justice Information System (ARJIS), the Homeland Security Information Network (HSIN), and the Law Enforcement Information Sharing Program (LEISP). Discussions with all of these organizations are still ongoing at the time of this writing.

- ***Definition of Next Steps*** – Participants collaborated to begin defining next steps for the project as it moves beyond a demonstration phase and into the role of an operational federation for information sharing.

- ***Development of Final Report*** – Participants collaborated to write this report – the final report for the demonstration phase of the GFIPM project.

6 Pilot Federation

This section of the report describes the pilot federation that was built throughout the demonstration project. The basic structure of the pilot federation is illustrated in Figure 12.

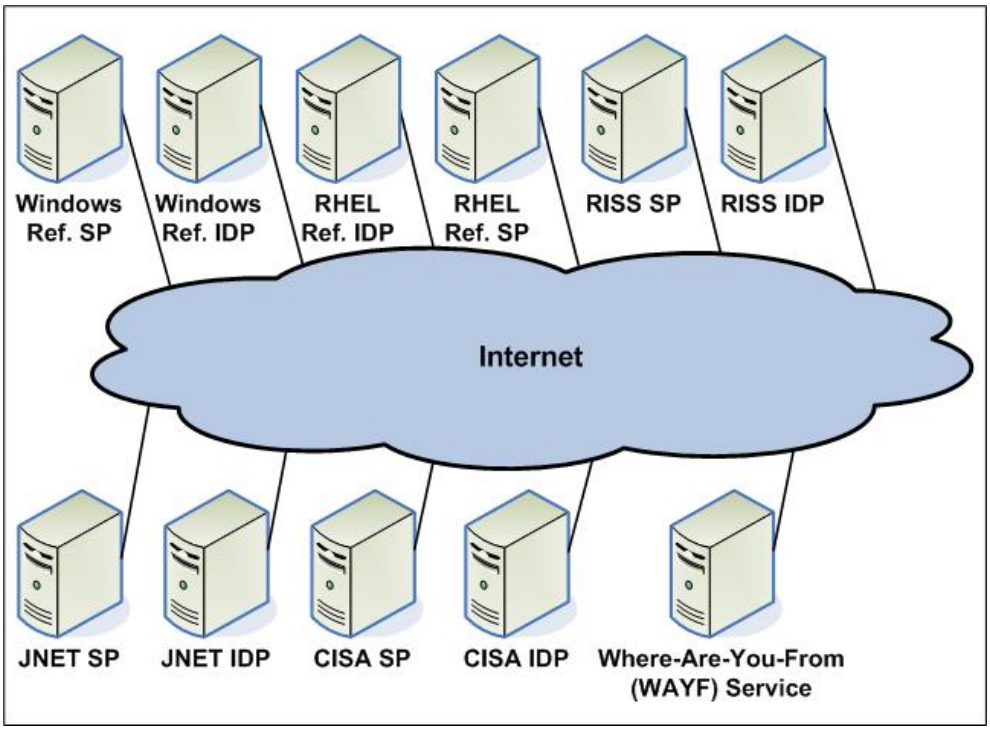

Figure 12: The GFIPM Pilot Federation

6.1 Reference Implementation

The first parts of the pilot federation to be built were reference implementations of each major functional component. These reference components were deployed by GTRI, and they served two purposes. First, the process of deploying the reference components served as a valuable learning experience and a source of documentation artifacts that were later used by other participants. Second, the reference components themselves served as a base on which the federation participants could bootstrap their components into the federation. Some of these reference components are still used in the federation. Each reference component is discussed individually in the subsections below.

6.1.1 Reference IDPs

GTRI deployed two reference IDPs in the pilot federation. Both IDPs were based on the Shibboleth 1.3 implementation of SAML 1.1. One of the reference IDPs was deployed on a Microsoft Windows platform, and the other was deployed on Red Hat Enterprise Linux (RHEL). There is no functional difference between a Shibboleth IDP running on Windows and one running on RHEL; however, the deployment processes for a Shibboleth IDP on each platform are different enough to merit the task of working through each and documenting them separately. During and after the deployment process, GTRI created a detailed set of instructions for deploying a Shibboleth IDP on each platform.

As discussed previously in Section 4.1.3, deploying an IDP in a federation requires that two integration issues be addressed. The first of these issues involves the integration of the IDP with a user authentication system (the Single Sign-On Integration Point), and the

other involves connecting the IDP to an attribute repository (the Attribute Authority Integration Point). During the deployment of the reference IDPs, the Windows-based reference IDP was integrated with a username/password authentication system, and the RHEL-based reference IDP was integrated with a PKI-based client certificate authentication system. At the Attribute Authority Integration Point, both reference IDPs were connected to a reference LDAP repository.

Both reference IDPs were very useful during the initial process of building out the pilot federation's infrastructure. Credentials for test accounts on each reference IDP were provided to participants, and these test accounts proved to be a valuable resource for participants as they brought their SPs online. After the pilot federation became more mature, however, it became clear that the reference IDPs were a double-edged sword: they provided a convenient means for reliably testing SPs for functional correctness, but they also provided a means for accessing sensitive resources. Due to this risk, GTRI decided to take the reference IDPs offline when participants began to bring live resources online in the pilot federation. This issue has been discussed previously in Section 5.3.

6.1.2 Reference SPs

In addition to deploying reference IDPs, GTRI also deployed two reference SPs in the pilot federation. Both SPs were based on the Shibboleth 1.3 implementation of SAML 1.1. As with the reference IDPs, one of the reference SPs was deployed on a Microsoft Windows platform, and the other was deployed on RHEL. Again, as with the reference IDPs, there is no functional difference between a Shibboleth SP running on Windows and one running on RHEL; however, the deployment processes for a Shibboleth SP on each platform are different enough to merit the task of working through each and documenting them separately. During and after the deployment process, GTRI created a detailed set of instructions for deploying a Shibboleth SP on each platform.

The integration work required for an SP involves setting up the SP to provide protected access to sensitive resources. During the deployment of the reference SPs, GTRI created some simple HTML and PHP pages to serve as protected resources. These pages serve two important purposes. First, they help participants debug various problems with their IDPs at the SAML configuration level, and second, they allow for careful inspection of the GFIPM metadata that an IDP sends to a reference SP. This feature has been very valuable in helping participants to identify and correct problems related to the generation of metadata by their IDPs.

As with the reference IDPs, participants found the reference SPs to be useful during the deployment process for their infrastructure. Participants were able to test their IDPs by attempting to access resources on the reference SPs. The RHEL-based reference SP is still online as of June 2007, serving as needed in a test capacity, but the Windows-based reference SP has been taken offline, because there is no need to have both reference SPs online anymore. After initially deploying reference SPs on both platforms and verifying that the Shibboleth software could function properly in both environments, there was no further need for keeping two reference SPs online simultaneously.

6.1.3 Where-Are-You-From (WAYF) Service

The final reference component in the pilot federation is the Where-Are-You-From (WAYF) service. As discussed in Section 4.4, a WAYF service provides a convenient means for a user to specify which IDP he would like to use for single sign-on within the federation. The pilot federation currently uses only one WAYF service, which is managed by GTRI; however, there is no inherent limitation on the number of WAYF services that a federation can use. For purposes of scalability and reliability, it may be necessary in the future to deploy multiple WAYF services within the federation, or to deploy an alternate means of allowing users to specify their IDP.

6.2 Criminal Information Sharing Alliance (CISA)

This section describes the work done by CISA in the pilot federation, including implementation decisions and current status of CISA's federation infrastructure as of the end of the demonstration project (Spring 2007).

6.2.1 CISA IDP

CISA's IDP implementation consists of a Shibboleth IDP software installation running as a Java servlet inside a Tomcat servlet container on a RHEL (version 4) platform. CISA's IDP is connected to an LDAP repository, which is used to generate metadata assertions on behalf of CISA users. The LDAP repository used by CISA is the Critical Path Directory Server 4.2, which is a full X.500 directory server with an LDAP interface. For user authentication, CISA's IDP uses the client certificate authentication (CCA) capability of the Apache web server and Internet Explorer with the Microsoft Cryptographic Application Programming Interface (CAPI) on the user's PC. CISA's CCA implementation uses the pre-existing CISA PKI with a certificate revocation list (CRL). The CRL is loaded daily onto the IDP, along with the CISA root certificate, to allow CISA users to authenticate directly to the CISA IDP.

6.2.2 CISA SP

CISA's basic SP implementation consists of a Shibboleth SP software installation running in conjunction with the Apache web server on a RHEL platform. CISA's protected resources fall into two categories: federation-aware resources (which are capable of natively consuming GFIPM metadata and reacting accordingly), and non-federation-aware resources (which cannot natively consume GFIPM metadata). Federation-aware resources have either been designed specifically for the GFIPM federation environment (with built-in support for processing GFIPM metadata), or have been modified to be able to exist in a GFIPM environment. In either case, a federation-aware resource can be served to federation users directly through an Apache/Shibboleth environment. The only CISA resource that is currently federation-aware is the CISAnet Federated Query Tool (CFQT). All other CISA resources are non-federation-aware, and therefore need to be connected to the federation via a federation-aware proxy application. CISA built such a proxy application and it's discussed in the following paragraphs.

Often it is cost-prohibitive or simply not possible to modify an existing resource so that it is federation-aware. As a result, nearly all existing resources in the GFIPM pilot

federation are non-federation-aware resources that have been made available via a federation-aware "reverse proxy" solution. A federation-aware reverse proxy, as discussed in Section 4.2.4, provides proxied access to non-federation-aware resources in a manner that is consistent with the access control and auditing policies of the organization managing the resource. For reverse proxying, CISA uses EZproxy, a third-party software package that has been integrated with Shibboleth and Apache. EZproxy is an inexpensive commercial product that performs robust reverse proxying. CISA developed custom software to translate GFIPM metadata assertions into "tickets" that conform to EZproxy's authentication and authorization model. All of CISA's available resources in the pilot federation, with the exception of the CFQT, are served to federation users via this EZproxy-based reverse proxy architecture.

6.2.3 CISA Resources

CISA has made the following resources available to federation users through its SP portal:

1. ***Arizona Counter-Terrorism Information Center (ACTIC)*** – Provides access to the fusion center. Functions include retrieval of public domain information documents, retrieval of information and links relating to counter-terrorism, opportunity to provide tips or leads regarding ongoing counter-terrorism events.
2. ***Arizona Sex Offender Information Center*** – Provides access to data regarding sex offenders in Arizona who meet the threshold criteria for public disclosure. The sex offender information center maintains basic identifying information, sex offender category, and registered address for Arizona's 11,000 registered sex offenders.
3. ***Arizona Amber Alert*** – Provides information about ongoing Amber Alerts (child abduction cases) for Arizona.
4. ***Georgia Bureau of Investigation Sex Offender Registry*** – This registry holds information pertaining to sex offenders who have been released from prison or youth detention facility, placed on probation or parole, or have a supervised release date after July 1, 1996.
5. ***Oklahoma State Bureau of Investigation Officer Safety Bulletin*** – The OSBI Officer Safety Bulletin has been compiled from many different sources in an effort to provide law enforcement personnel with up to date information on issues regarding undercover drug stings, violent disturbances, ideological extremist individuals and groups, weapons, hidden storage spaces, and numerous other officer safety concerns.
6. ***Texas Criminal Law Enforcement Online (CLEO)*** – The Criminal Law Enforcement Online (CLEO) is a secure website restricted to the criminal justice community and hosted by the Texas Department of Public Safety (TX DPS) Criminal Intelligence Service. CLEO participants may post information on a message board or submit information to the CLEO Administrator to be posted in a specific area of the website such as Alerts, Bulletins, Announcements, Profiles, News or Calendar. The website is used to disseminate a wide variety of information including unsolved crimes, wanted fugitives, career criminals, trends, concealment methods, officer safety alerts and more.

7. ***California Joint Regional Information Exchange System (JRIES)*** – Read-Only Data for Law Enforcement Officers and Public Safety employees.
8. ***New Mexico Complete Arrest Information (CAI)*** – Consists of fingerprint supported arrest record information and non-fingerprint supported information (corrections downloads, missing person information, sex offender registration information, FBI downloaded information for New Mexico) maintained by the NM State CJIS.
9. ***New Mexico Incident Based Reporting System (NMIBRS)*** – New Mexico State Police offense/incident report information.
10. ***New Mexico Sex Offender Registration*** – New Mexico Sex Offender Registration information maintained by the Department of Public Safety.
11. ***New Mexico Missing Person & Unidentified Bodies*** – Information obtained from law enforcement agencies throughout the State of New Mexico. Provides information pertaining to Missing Individuals or Unidentified Bodies/Body Parts.
12. ***New Mexico Field Interview (FI)*** – Field interview information captured by NM State Police.
13. ***New Mexico Law Enforcement Information Network with Corrections (LINC)*** – Information downloaded from the New Mexico Corrections database.
14. ***New Mexico Criminal Law Enforcement Reporting and Information System (CLERIS)*** – Information from the Criminal Law Enforcement Reporting and Information System (CLERIS).
15. ***Arizona Criminal Investigative Database*** – Application providing retrieval of information from the Arizona Criminal Investigative Database. This database holds Arizona criminal investigative reports, Arizona field interview reports, the ACTIC Watch Log, and other investigative related information.
16. ***Criminal Law Enforcement Reporting and Information System (CLERIS)*** – Information from the Texas DPS Criminal Law Enforcement Reporting and Information System.

Additional information about each resource listed here is available at the GFIPM User Portal, which is available online at http://gfipm.net/users/.

6.2.4 CISA Users

CISA performed the following actions to bring its local users into the GFIPM federation. During the CISA Strategy Session and Board of Directors meeting on November 28-30, 2006, the CISA leadership requested that each CISAnet participant provide CISA with the names of 7-10 analysts/officers who could assist with GFIPM operational testing. By December 12, 2006, a list of 53 test users had been compiled, and on the following day a conference call was conducted to ensure that that all test users were aware of the access procedures and the tasks that they were to perform. Testing began in January.

During the testing period (which ran from January to May 2007), CISA asked its test users for feedback on issues, problems and successes that they may have encountered, and by May 24, 2007, 16 responses had been received.

Regarding **access and ease of use**, 12 of 16 users responded positively by saying they found value in gaining access to data sources outside of CISAnet without having to be registered in the resource provider's system. All 12 of these users liked the fact that a single sign-on transaction would allow them access to CISAnet, JNET and RISS resources. Three users experienced difficulty with the GFIPM sign-on process; however, this was due to a CISAnet registration issue and was unrelated to the GFIPM concept.

Regarding **resources available** via the federation, eight users responded positively by saying they found the resources available from the JNET very useful. Six users responded that the JNET resources were easy to use and understand. Four users expressed concern that there was little value in the RISS resources provided. Five users noted that there were too few resources available to justify the federation's operational use. Two officers found it difficult to determine which resources were available for use.

6.3 Pennsylvania Justice Network (JNET)

This section describes the work done by JNET in the pilot federation, including implementation decisions and current status of JNET's federation infrastructure as of the end of the demonstration project (Spring 2007).

6.3.1 JNET IDP

JNET's IDP implementation consists of a Shibboleth IDP software installation running as a Java servlet inside a Tomcat servlet container on a Windows platform. For user authentication, JNET's IDP uses the client certificate authentication (CCA) capability of Tomcat. JNET's CCA implementation uses a pre-existing JNET Public Key Infrastructure (PKI). JNET created and developed this PKI in 1997, and it has served JNET well in being able to authenticate users with a higher level of assurance than username/password authentication can provide. (The PKI has enabled cross-political and cross-government security trust among government agencies within Pennsylvania.) JNET performs CRL checking for this PKI and is looking to implement OCSP in the near future. The JNET IDP is connected to an LDAP repository, which is used to generate metadata assertions on behalf of JNET users.

6.3.2 JNET SP

JNET's SP implementation consists of a Shibboleth SP software installation running in conjunction with Microsoft IIS on a Windows platform. JNET has faced many of the same issues that CISA has encountered regarding non-federation-aware resources. To deal with this challenge, JNET has implemented a custom reverse proxy solution similar to CISA's, except that rather than incorporating an existing COTS proxy application, JNET developed its reverse proxy capability internally as custom code. The JNET proxy provided for the necessary aspects of the demonstration and has provided a valuable environment for JNET to learn lessons related to reverse-proxying and federation-enablement of legacy applications. JNET's goal in creating this proxy service was to find a "best-fit" proxy solution which would work for the GFIPM demonstration, as well as a timely solution with the latest technology that would promote capabilities to existing JNET applications. As such, JNET created a solution which would work with the

demonstration project and meet the project timelines. JNET plans to evaluate COTS proxy solutions in the near future based upon basic GFIPM requirements and the requirements of JNET agencies' internal applications.

While JNET's current SP implementation provides a functioning solution for the GFIPM demonstration project, it is not optimal in providing a supportable environment beyond a demonstration/pilot environment. JNET will need to implement a more supportable proxy solution as the GFIPM federation scales up from demonstration/pilot into an operational mode. JNET needs to replace the GFIPM custom software with a supportable and flexible technology which integrates with JNET's existing security solution set.

Within JNET's proxy solution, GFIPM users were linked to backend legacy resources using a mapping strategy in which users were classified into groups based on their credentials. Then each group was given specific access privileges on each backend application via a pre-provisioned account. Authentication from the proxy to backend resources was accomplished using proxy certificates that leveraged the authentication mechanisms of the legacy applications. This federation enablement strategy (reverse proxying with secondary authentication and group-based access) was used because JNET found that it would be easier than reengineering the legacy resources to accept Shibboleth credentials. Further discussion of this particular federation enablement strategy for resources can be found in Section 7.4.

One other unique aspect of the JNET SP worthy of mention here is that JNET has implemented a terms-of-use web page that all GFIPM users encounter prior to accessing JNET's GFIPM resources. The page presents JNET's terms of use to the user and requires that the user click a button to explicitly indicate acceptance of the terms prior to allowing the user to access any JNET resources. The specific purpose of this page is to present the user with identifying information about himself that the JNET SP has received from the user's IDP in a GFIPM assertion and ask the user to explicitly indicate that the identifying information is correct, i.e. that the user really is who his IDP says he is. This feature is important for JNET from the standpoint of accountability and non-repudiation for actions taken by users while using JNET resources.

6.3.3 JNET Resources

JNET has made the following resources available to federation users through its SP portal.

1. ***Pennsylvania Department of Corrections Intake/Exit Photos*** – From the Commonwealth of Pennsylvania Department of Correction.
2. ***Pennsylvania Arrest Warrants Outstanding for Parolees who failed to report (Absconders)*** – These persons have had warrants issued for their arrest due to not reporting for supervision as required by conditions of parole. All were under state supervision after conviction and sentencing for serious crimes involving prison sentences at state facilities or were assigned specifically to state supervision during their probation or parole period.

3. ***Pennsylvania State Prisoner Locator*** – Prisoners in Commonwealth facilities can be found on this site. Persons confined in local county prisons are not on this site.
4. ***Pennsylvania Criminal Trial Case Information*** – Case, scheduling, bail, and warrant information is available based on Pennsylvania State Identification Number (SID). (See Department of Corrections above.) This criminal trial case information is available for the entry level arraignment courts in Pennsylvania. Included are the charges stemming from the arrest as well as bail and warrants issued on the case.
5. ***Pennsylvania Arrest Warrants Outstanding for Failure to Pay Child Support*** – When arrest warrants are issued for individuals who have failed to pay child support they are added to this site. The details of any Pennsylvania warrant served for failure to pay child support can be found on this site. The service provides a last-name search capability, and search results include the defendant's full name, SSN, driver's license number, date of birth, sex, domestic relations case id, date of warrant, issuing county, city, zip code, and contact phone and information for the applicable domestic relations enforcement officer.
6. ***Pennsylvania Amber Alert*** – The details of an Amber Alert activation in Pennsylvania are found on this site.

Additional information about each resource listed here is available at the GFIPM User Portal, which is available online at http://gfipm.net/users/.

6.3.4 JNET Users

JNET performed the following actions to bring its local users into the GFIPM federation. First, on November 2, 2006, JNET invited 123 users to participate in the GFIPM federation. Invitations were sent via email. A follow-up email was sent to interested users about six weeks later, in mid-December 2006, to provide quick-start procedures and example federation use case scenarios. There were no face-to-face meetings with potential participants. As of June 2007, JNET has approximately 60 users participating in the GFIPM federation.

Since the majority of the GFIPM resources that were made available by CISA and RISS contained data that was beneficial to law enforcement, most of the JNET users selected for participation in the federation were law enforcement personnel. The JNET users represented in the GFIPM demonstration were selected from all levels of Pennsylvania government, including local, county and state levels. Types of JNET users represented in the demonstration project included the following;

- County Adult Probation Supervisors
- State Adult Probation Supervisors
- State Probation Officers
- Local Law Enforcement (Chiefs, Detectives, Lieutenants, Officers, and Sergeants)
- D.A. Office Staff Members
- Community Service Officers
- Domestic Relations Enforcement Officers
- Emergency Management Chiefs

- Terminal Access Control (TAC) Officers

Figure 13 contains a pie chart illustrating the relative numbers of JNET test users who participated, by user type.

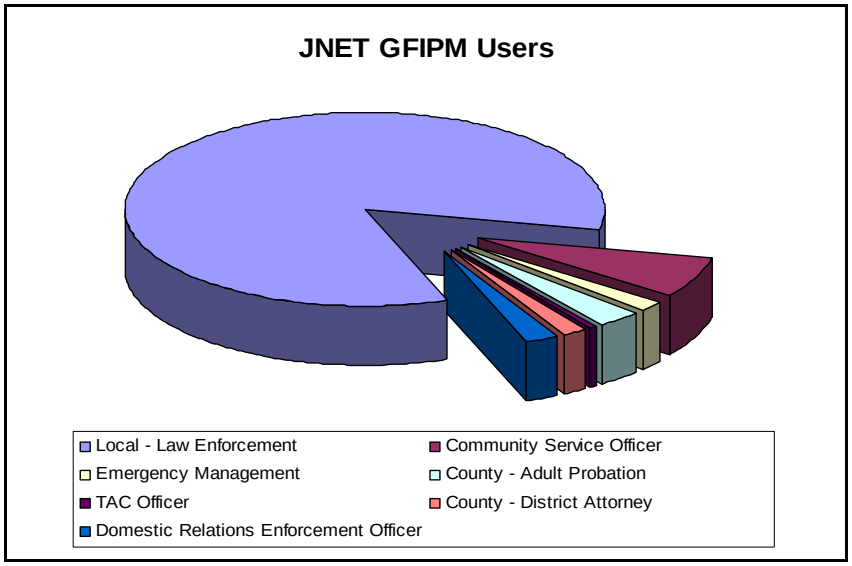

Figure 13: JNET User Participation in GFIPM by User Type

In June 2007, JNET sent a questionnaire to all of its participating GFIPM users to collect feedback about their experiences using the federation. The seven users who responded represent the following demographics:

- State Probation/Parole Officer;
- County Chief of Detectives;
- Local Police Lieutenant;
- Local Police Sergeant;
- Local Chief of Police;
- Sworn Law Enforcement Officers with local criminal history privileges but not criminal intelligence privileges.

Participants responded positively to the following features and aspects of the GFIPM federation:

- Ability to access sensitive data via the web using TLS encryption, rather than having to be at the office or in a cruiser;
- Ability to access resources outside Pennsylvania without having to get a new user account from other agencies;
- Ability to access data from multiple databases with one sign-on transaction.

In addition, most respondents answered favorably when asked whether resources in the demo federation were useful, and most respondents found that non-JNET resources were

easy to use and understand without needing any training other than the instructions found within the resources themselves.

The most significant concerns raised by JNET's test users were related to the usefulness of the content available through the federation. This issue is already well-understood by JNET and the other federation participants, and will be addressed as the federation moves from a demonstration phase into an operational phase.

6.4 Regional Information Sharing Systems (RISS)

This section describes the work done by RISS in the pilot federation, including implementation decisions and current status of RISS's federation infrastructure as of the end of the demonstration project (Spring 2007).

6.4.1 RISS IDP

RISS's IDP implementation consists of a Shibboleth IDP software installation running as a Java servlet inside a Tomcat servlet container, which is running in conjunction with Microsoft Internet Information Services (IIS) on a Windows platform. The SAML metadata information is generated using the GFIPM User Assertion LDAP Connector to retrieve user information from Microsoft Active Directory. The connector transforms an XML representation of the user with an XSL document to produce a GFIPM user assertion. Like CISA and JNET, RISS is using client certificate authentication (CCA) on its IDP. The connector resolves the user's identity from the certificate that the user presents

6.4.2 RISS SP

RISS's SP implementation consists of a Shibboleth SP software installation running on a Windows platform in conjunction with Microsoft IIS. Currently, the RISS SP provides access to only static resources, not applications; therefore, RISS has not needed to deal with the reverse proxy issue. RISS anticipates that it may be able to offer more resources to GFIPM users after the federation has moved to SAML 2.0 and begun to integrate COTS products into the infrastructure.

6.4.3 RISS Resources

RISS has made the following resources available to federation users through its SP portal.

1. ***HSIN Counter-Terrorism Briefs, Reports, and Documents***
2. ***RISS Counter-Terrorism Briefs, Reports, and Documents***

Additional information about each resource listed here is available at the GFIPM User Portal, which is available online at http://gfipm.net/users/.

6.4.4 RISS Users

RISS performed user testing within GFIPM using five individuals who have a broad range of local permissions within RISS. Three of the individuals are active staff members of the Regional Organized Crime Information Center (ROCIC) and two are

from RISS Office of Information Technology (OIT). Individual permissions ranged from very limited (virtually no local privileges) to very broad (e.g. Criminal Intelligence Analyst with NCIC privileges). None of the RISS test users were sworn law enforcement officers. The test users were given the information available at the GFIPM User Portal and shown how to access the various resources available within the federation.

Test user feedback was gathered using the "GFIPM Resource Test Report" feedback form provided by GTRI. Each report was initially completed by a network administrator whose role was granting access to the RISS IDP. This information included each test user's name, organization, certifications, and available federation permissions. The test users then navigated through the available federation resources with guidance from the network administrator.

The overall consensus from the RISS test users was positive, since they understood that they would have access to resources from various entities in the federation without having to log-in multiple times with different username/password combinations. The users also identified several concerns, which are listed below.

- Test users had limited knowledge about the resources that were available from other federation members.
- Many of the resources that would be relevant to criminal investigations were available only to sworn law enforcement officers.
- Some of the resources available through the federation are already available via the Internet without requiring any credentials.
- Some test users were confused when they encountered the Where-Are-You-From service and were asked to choose their identity provider.
- Some test users were frustrated by the inability to request access to federation resources when their permissions and certifications were insufficient for access according to the resource owner's access requirements.

Many of the concerns cited here support the notion that the federation would benefit from a central directory or registry of available resources, including basic access requirements for each resource and a clear description of the process for requesting access to resources that are inaccessible to the user. (Some resources will be inaccessible because the user does not have the requisite permissions, but others will be inaccessible simply because the user's IDP does not provide the appropriate metadata about the user.) This topic is further discussed in Section 7.

6.5 GFIPM Pilot Federation User Demographics

Demographics from the current user base include the following categories: sworn law enforcement, criminal intelligence, counter-terrorism, probation and correction, and other justice and public safety support missions. One of the key value propositions with the GFIPM approach is that users are effectively added "in bulk" when an IDP is brought into the federation instead of one user at a time. Even with the limited number of IDPs

established during the pilot phase, three currently, more than 170,000 state, local, and federal users potentially have access to federation resources today. Although only a relatively small number of users have participated during the evaluation phase, access could be expanded to the full user base of the current IDPs without additional integration. However, certain operational, training, and support issues would need to be considered (see *Section 7 Lessons Learned and Conclusions*) prior to this expansion. Figure 14 breaks down the potential GFIPM user base into types of users. The categories and totals presented are not mutually exclusive; rather they are based on what IDPs would assert for their existing user base. For example, an IDP may assert sworn law enforcement officer, criminal intelligence, and counter terrorism for a user with a job function of criminal intelligence analyst. These numbers are reflected in the counts depicted.

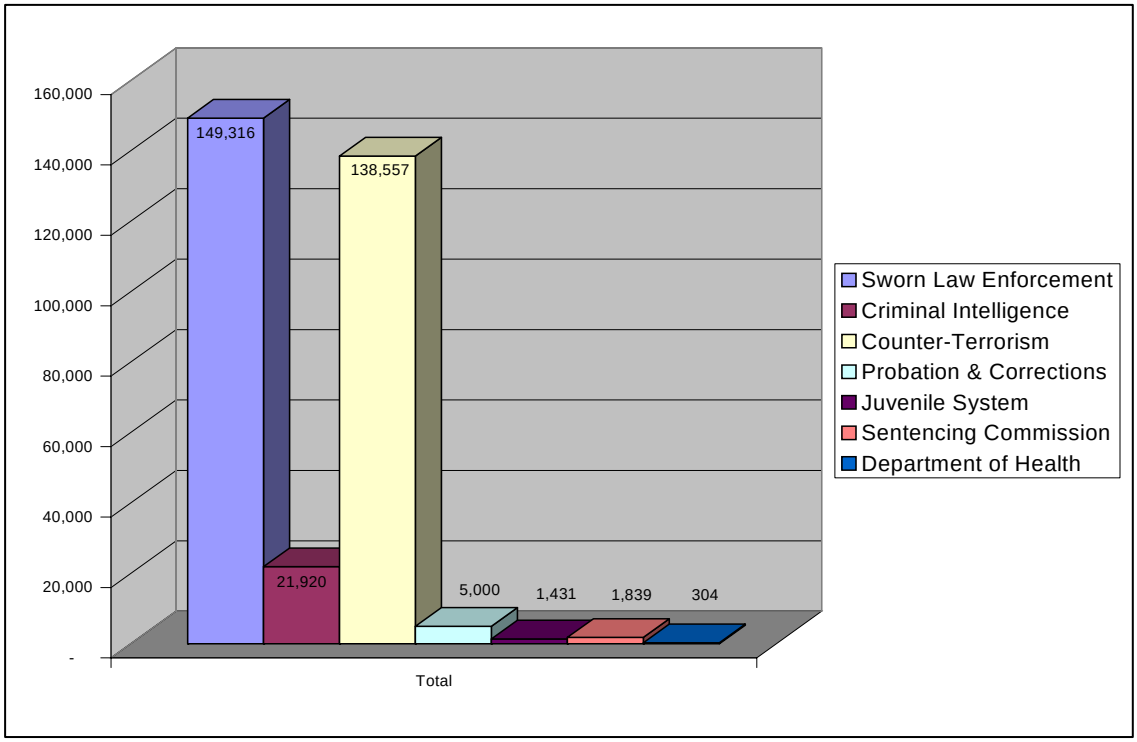

Figure 14: GFIPM Pilot Federation User Demographics

Figure 15 provides the breakdown of potential GFIPM users based on the existing IDPs who are federal, state, and local.

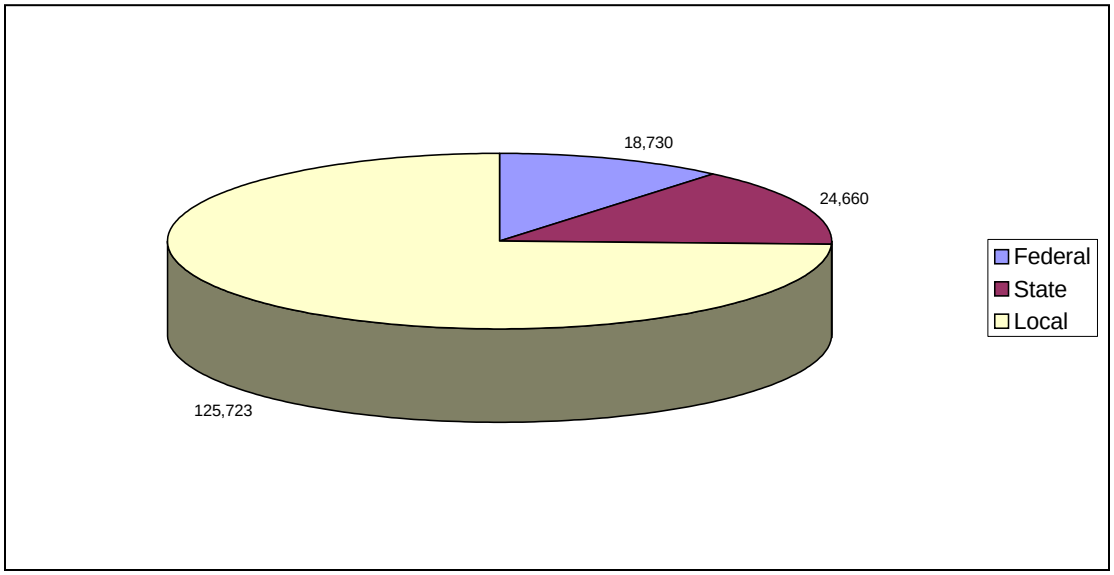

Figure 15: GFIPM Pilot Federation User Demographics (Federal, State, Local)

6.6 GFIPM User Portal and Resource Directory

During the demonstration project, it became apparent that even if many valuable resources were available online for federation users, they would be of little value if users did not know they existed or could not navigate to them in a web browser. This realization led to the development of the GFIPM User Portal, which serves as a resource directory for all available resources in the federation, as well as a source of basic content for users about the federation concept. The portal is publicly accessible at http://gfipm.net/users/. This resource directory is currently maintained manually by GTRI; however, it is expected that as the federation matures, the directory will evolve into a registry and its management will become more automated.

6.7 Sample Screen Shots of an SSO Transaction

This section illustrates a typical use case in which a user attempts to access a protected resource and is redirected through a single-sign-on (SSO) transaction before being granted access.

1. User visits an unprotected portal page. In this case, the user is visiting the CISA portal. To begin the authentication process and gain access to the protected resources, the user must click on the link marked “Login to CISAnet Protected Resources”. Figure 16 illustrates this step of the transaction.

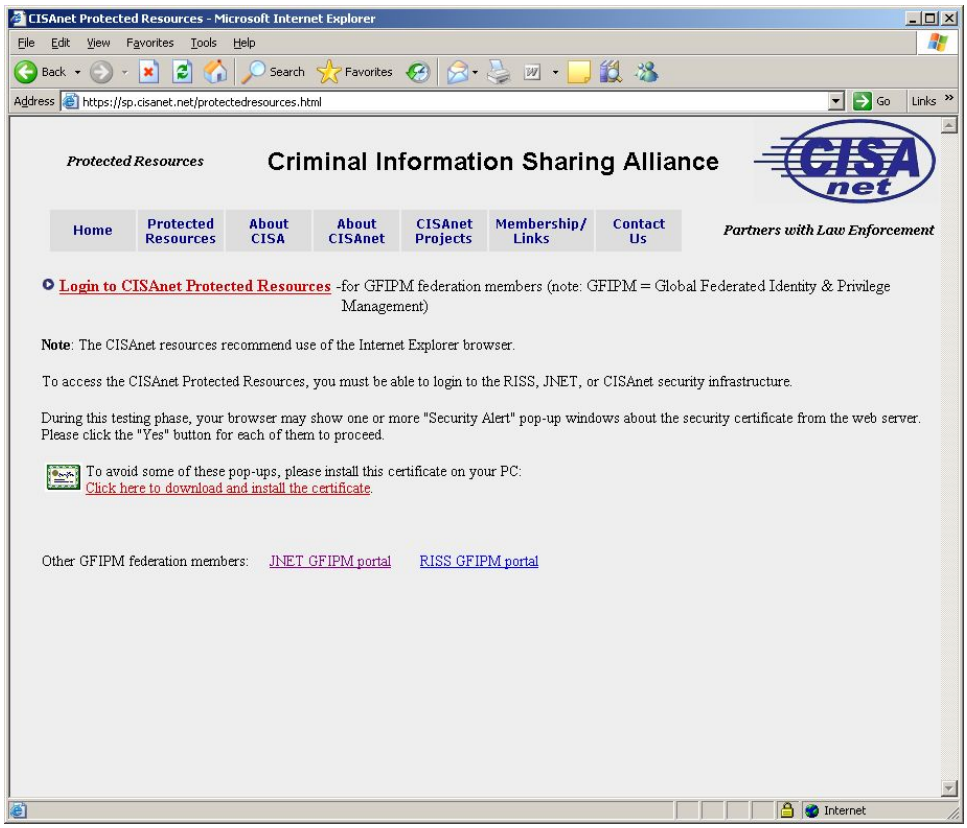

Figure 16: Step 1 of an SSO Transaction in GFIPM

2. User is redirected to the federation's Where-Are-You-From (WAYF) page. At this point, the user must select his home organization. This user's home organization is CISA. Figure 17 illustrates this step of the transaction.

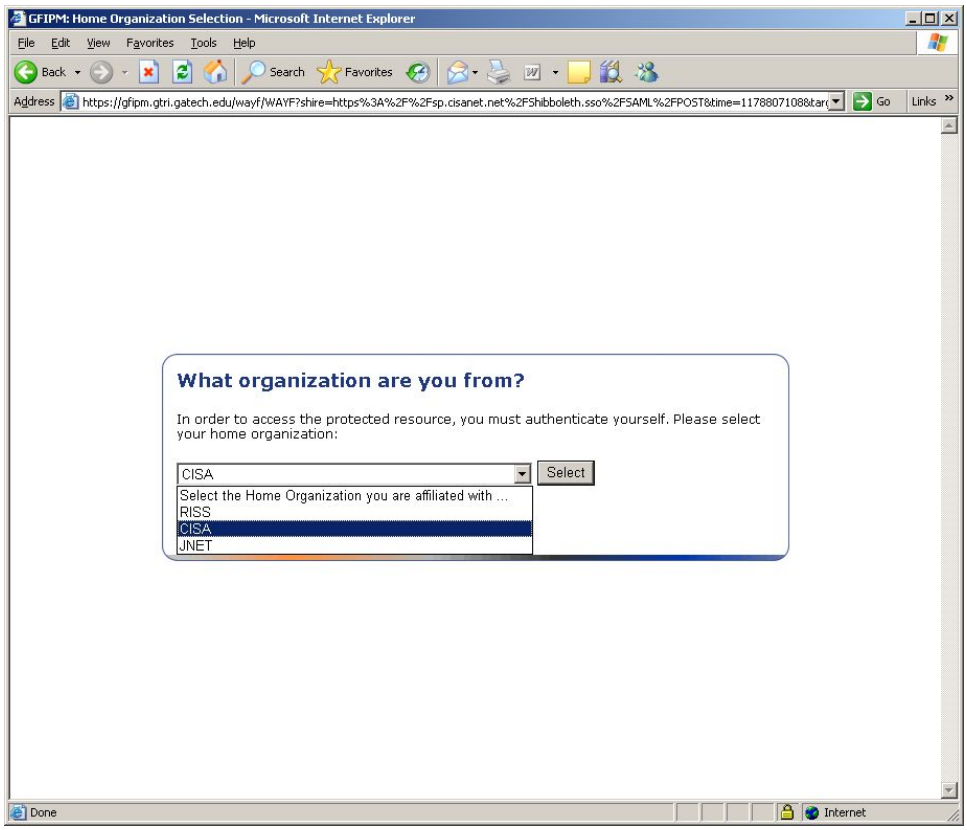

Figure 17: Step 2 of an SSO Transaction in GFIPM

3. User is redirected to his home organization's IDP and goes through the authentication process. In this case, the user's IDP uses two-factor authentication with a hardware-based security token and password. Figures 18 and 19 illustrate this step of the transaction.

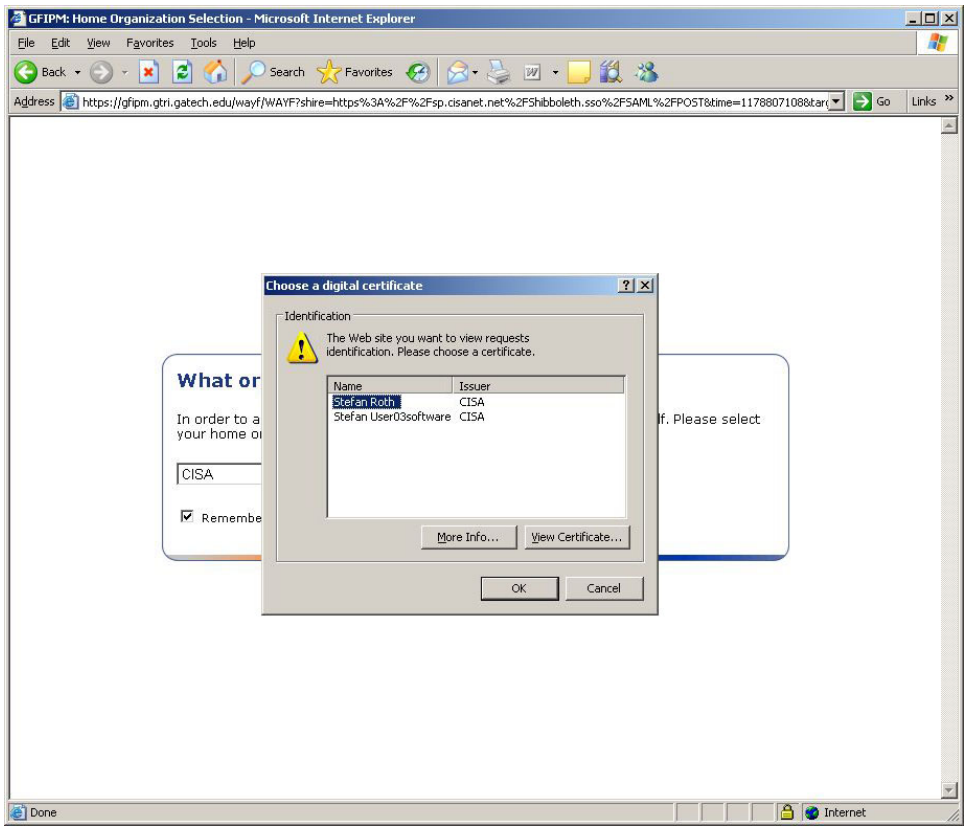

Figure 18: Step 3 of an SSO Transaction in GFIPM

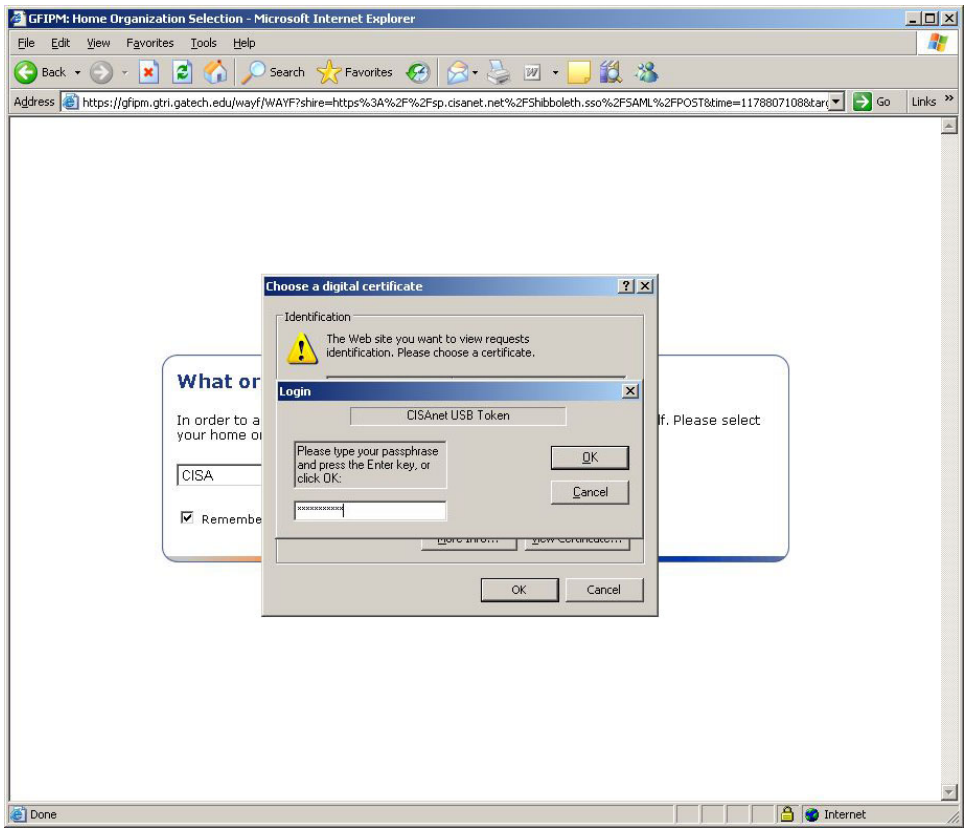

Figure 19: Step 3 of an SSO Transaction in GFIPM

4. User is redirected to the protected resources that he wanted to see. On the CISA portal, protected resources are listed in the table at the bottom of the web page. The portal provides links for resources that are available to this user, based on his permissions. (For example, see the row for "Arizona Amber Alert".) It also indicates resources that are not available to this user, by marking them "denied". (See the row for "Arizona Criminal Investigative Database".) This step of the transaction is illustrated in Figure 20.

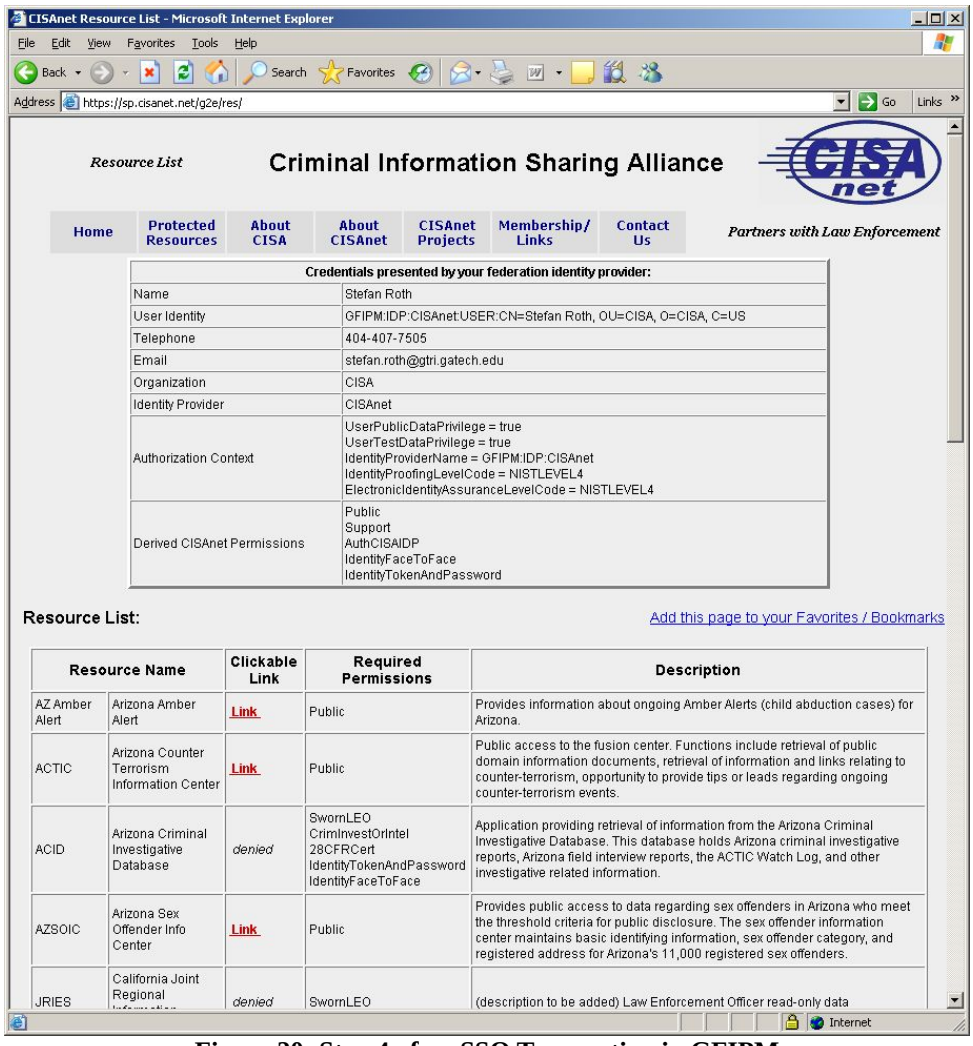

Figure 20: Step 4 of an SSO Transaction in GFIPM

The sequence of screens a user sees may vary slightly from what is shown here, depending on the user's home organization or the resources that he is trying to access. However, an SSO transaction for GFIPM users always follows the same basic progression that has been illustrated here.

7 Lessons Learned and Conclusions

Throughout the demonstration project there were countless lessons learned ranging from major issues such as federation governance to small technical details such as how to configure a specific reverse proxy solution on a specific platform. GTRI attempted to capture and catalog as many of these lessons as possible, and this section of the report contains a summary of the most important lessons and conclusions that were captured during the project. For ease of readability, the section is divided into subsections based on major topics.

7.1 GFIPM Concept

This section contains lessons and conclusions related to the basic GFIPM concept: federated identity and privilege management using a common metadata model to transmit information about users from IDPs to SPs.

1. The basic GFIPM concept – federated identity and privilege management for inter-agency law enforcement information sharing – proved to be viable and worked as expected. The demonstration project was able to securely connect existing CISA, JNET, and RISS users to participant production web-based resources over the Internet. GFIPM metadata was used by participant IDPs and SPs to share authentication and authorization information within SAML assertions in a secure and trusted manner. This was accomplished by leveraging currently existing participant authentication mechanisms, LDAP data stores, web-based resources, and participant subscriber bases.

2. The GFIPM concept enables the shift from vetting a large numbers of individual users across the federation to vetting a much smaller number of IDPs and their policies and practices. Individual user vetting is the responsibility of each IDP. Existing trust relationships between local organizations (IDPs) and their vetted users were leveraged in the federation through the trust relationships established between participating organizations (IDPs and SPs). The trust established between participating organizations served as the basis for trust between federation users and organizations. See Section 7.6 for additional discussion of lessons related to federation governance and trust management.

3. Federated identity and privilege management (FIPM) based on the sharing of a rich set of trusted attributes about federated users between IDPs and SPs can enable several important benefits beyond that which simple federated identity management (FIM) can provide. These benefits stem from FIPM's ability to facilitate the integration of legacy applications into the federation by helping to fulfill those applications' usage requirements for users. But to achieve the benefits, FIPM also requires that a common metadata model be defined, vetted, and deeply integrated into many parts of the federation infrastructure. Section 7.3 of this report contains important lessons learned regarding the metadata, and Section 7.4 contains lessons learned regarding the federation-enablement of resources in relation to the metadata model.

4. The GFIPM concept provides a sufficient level of security for sharing sensitive-but-unclassified (SBU) resources via the Internet. The level of security provided is "sufficient" in the sense that it is consistent with existing security policies and practices for users who access existing legacy applications. In addition, the FIPM concept provides a level of security that is equal to or better than the level of security provided in a legacy environment, in the following ways.

a. Metadata attributes about federated users are attested in a cryptographically secure manner by IDPs that have been vetted in accordance with federation-wide vetting procedures, and can be trusted by SPs accordingly.
b. User information attested by an IDP tends to be both fresher and more trustworthy than information that is self-attested by the user. It is fresher in the sense that updates to the user's contact information, certifications, or other credentials will tend to be reflected in a timely manner by that user's IDP. In contrast, if users are responsible for updating their account profiles with many applications, then the updates may not happen in a timely manner, or may not happen at all. Information attested by an IDP is more trustworthy in the sense that it constitutes a statement by an authoritative third party source.
c. TLS is accepted as a sufficiently secure protocol for the establishment and exchange of cryptographic session keys for protecting sensitive-but-unclassified information in the justice domain, and all IDPs, SPs, and browsers in the GFIPM federation can be configured to use TLS.
d. The SAML standard itself is widely accepted and has been thoroughly reviewed by the security community. Its wide use ensures that it will continue to receive careful scrutiny as time goes on, thereby helping to ensure its continued security.

5. There is no significant difference in performance, as measured by responsiveness of resources from the end-user's perspective, between the federated GFIPM environment and a traditional non-federated information sharing environment.

7.2 Business Case for GFIPM (When and Why to Join)

This section contains lessons and conclusions that relate to the business case for GFIPM. It addresses the question of when to join the federation from the perspective of an IDP or an SP, as well as how to grow the federation from the perspective of existing federation participants.

1. The GFIPM concept provides maximal value when it leverages a large number of users and resources in support of a common data sharing mission. Just like any other network (including the Internet), its value to both users and resource providers increases as the number of participants increases. By implication, the value of the GFIPM federation concept to early adopters may not initially be very high. But as the federation grows to include many constituencies (users and resources), its value to those constituencies will grow also.

2. For an SP, the value in joining the GFIPM federation depends on the federation's ability to provide a user base that is both large enough and demographically suited to justify the cost outlay of connecting the SP's resources to the federation. For this reason, it is critical that the federation develop processes, mechanisms, and tools for collecting information about the demographic makeup of the federation's

user base across certain basic dimensions such as job function (e.g. sworn law enforcement officer, first responder, etc.), certifications (e.g. 28-CFR, NCIC Hotfile, etc.), and geographic locality (by city, county, state, or region). This information is vital to facilitate outreach and recruitment of new resources into the federation. Additional discussion about the topic of bringing resources into the federation can be found in Section 7.4 of this report.

3. For an IDP, the value in joining the GFIPM federation depends on the federation's ability to provide the IDP's user base with a set of resources for which the added value can justify the cost outlay of connecting that user base to the federation. For this reason, it is critical that the federation develop processes, mechanisms, and tools for collecting information about the resources that are available in the federation. This information may take the form of a browsable resource registry or directory containing essential information about each federation resource. It may also include usage statistics for specific resources, as well as success stories or case studies that highlight the value that federation resources have provided to the current user base. This information is vital to facilitate outreach and recruitment of new IDPs and users into the federation.

4. There are many business arrangements through which an organization can enable its users and/or resources to participate in the GFIPM federation. Here are some sample arrangements.

    a. Join the federation and directly connect an IDP and/or SP to it. (CISA, RISS, and JNET have already done this.)

    b. Join an organization that is already a member of the federation, and connect to the federation via its IDP and/or SP. (Organizations such as the Georgia Bureau of Investigation and the Pennsylvania Dept. of Corrections have already done this through their affiliation with CISA and JNET, respectively.)

    c. Connect to the federation through a broker service that acts as an IDP or SP on a fee-based basis. (This approach is not currently being used within GFIPM; however, it appears to have value. The LEISP project is experimenting with this concept. See Appendix A for more information about LEISP.)

    Note - this list is not intended to be exhaustive. Also, each participation arrangement is likely to have unique benefits and drawbacks in terms of cost, implementation time, and freedom of choice regarding policy and technical decisions.

5. The GFIPM concept does not impose or imply any level of information-sharing exclusivity among its users. Participants may – and probably will – join multiple federations based on pre-existing business relationships and cost/benefit

considerations related to their business needs. In addition, the GFIPM federation itself may choose to connect its users and resources with the users and resources of other information-sharing federations. At the time of this writing, GFIPM participants are planning to hold discussions with representatives from the LEISP project to discuss this topic.

7.3 Metadata and Infrastructure

This section contains lessons and conclusions concerning the metadata model and other basic infrastructure required to support the GFIPM concept.

1. It is possible to define a metadata model that can adequately describe federated users for the purpose of ensuring proper identification, authorization, access control, and auditing during the process of accessing sensitive resources within a federated environment. The metadata can be securely generated by each user's home organization and electronically transmitted to federation resources in a manner that is secure and trustworthy.

2. The metadata model required to support federated identity and privilege management imposed some key requirements and constraints on participants during the demo project, specifically in these four areas.

    a. ***Standards and Representation*** – Participants were required to reach agreement on definitional issues such as what metadata elements to include in the conceptual metadata model, how to precisely define each metadata element in the conceptual model, what subset of the conceptual model to include in the federation profile, how to encode the metadata model in an XML-based data structure, and how to encode the resulting XML structure within a SAML assertion for transport from an IDP to an SP. All of these decisions were further constrained by the decision to leverage NIEM content and structure, where possible. It became clear during the project that this decision-making process – the "metadata lifecycle" process – would benefit from some basic structure and oversight. Section 7.6 discusses this concept further in the context of federation governance.

    b. ***Construction of a Profile Instance*** – Participants worked in conjunction with GTRI to develop a suite of tools that could be used to generate a metadata profile instance for an individual user at an IDP. This process exposed several important lessons related to constructing metadata instances at IDPs, the most important of which are listed here.

        i. Most of the required metadata attributes about users could be gathered directly from attributes in local LDAP repositories and converted into GFIPM metadata. Additional metadata was able to

be derived from local LDAP attributes or inferred from organizational policies.

ii. There are some categories of metadata elements that IDPs simply cannot provide. One example of such a category is application-specific information, such as user preferences that pertain to a specific resource. Another example category is transaction-specific information, such as a case number that might be required during a resource query.

iii. In the current federation, the attributes for user assertions are provided by a single logical attribute authority (AA), which is collocated and tightly integrated with the IDP. There are conceivable use cases, however, in which it would be beneficial for user attributes to be gathered from multiple AAs. For example, attributes associated with certain certifications (28 CFR, NCIC, etc.) or security clearances might be acquired through an AA considered authoritative for that particular attribute. Support for multiple AAs to be associated with a single user identity would further distribute the federation and require changes at many levels of the federation infrastructure, including the SAML standard, SAML middleware implementation, and the GFIPM metadata standard. This scenario is currently being explored by members of the Shibboleth development team and OASIS.

c. ***Transport via SAML*** – Participants worked in conjunction with GTRI to develop candidate solutions for transporting GFIPM metadata profile instances from IDPs to SPs via SAML. This process involved studying the capabilities of both the SAML standard and SAML implementations to find acceptable methods of encoding GFIPM metadata XML within SAML assertions. The solution currently used within the GFIPM federation works by encoding a GFIPM profile instance as a single XML document and placing the entire document inside the attribute value for one SAML attribute inside a SAML attribute statement. This encoding technique has worked within the Shibboleth-based demo federation; however, more research is required to investigate how well the current solution will work with other SAML products, and/or to develop and investigate new candidate encoding solutions as needed. This is an important next step for the project as GFIPM moves beyond a demo federation and towards an operational federation. See Section 4.6.2.4 for more detail about the challenges of choosing a suitable SAML encoding strategy for GFIPM metadata assertions.

d. ***Consumption of a Profile Instance*** – During the process of federation-enabling their protected resources, participants needed to develop methods of connecting applications that did not have any inherent knowledge of the GFIPM concept or metadata into the federation in a manner that leveraged

the metadata model as much as possible to meet each application's basic requirements. This topic is discussed in more detail in Section 7.4.

3. Demo project participants performed an assessment of application requirements in terms of metadata attributes, and factored out a small set of primitive attributes to be used for access control. This set consisted primarily of widely accepted certifications and definitions from the broad law enforcement community, e.g. "Sworn Law Enforcement Officer", "NCIC Hotfile Certified", "28-CFR Certified", etc. Attempts to further refine these attributes into a set of federation-wide user roles for access control (e.g. "Patrol Officer", "Detective") were not as successful, because participants found it difficult to reach agreement on role definitions at this time. Therefore, the participants chose to exchange the more primitive attributes with each other in metadata and use them as each participant saw fit for local access control purposes within applications.

4. Participants found that the lack of standard code tables for some metadata elements caused problems in local applications where that metadata was used for access control purposes. It became clear that any metadata upon which machine decisions are made (including, but not necessarily limited to, access control and privilege metadata) must be codified to maximize each application's ability to interpret it correctly. This has been identified as an important next step of the project.

5. Participants recognize that the current metadata model needs further refinement in at least the following dimensions (previously discussed in Section 5.4);

    - Correction of obvious errors, including spelling errors and clear inconsistencies;
    - Introduction of code tables for metadata elements as necessary;
    - Improvement of element definitions where necessary;
    - Addition of clear usage guidance regarding use of metadata in federation where necessary;
    - Expansion of current metadata based on broader base of IDPs and SPs;
    - Review of current metadata against policy requirements;
    - Development of a small set of standard metadata profiles to foster adoption and interoperability;
    - Incorporation of lessons learned from the demo project where necessary;
    - Harmonization with the latest version of NIEM (version 2.0).

    It is expected that the next version of the GFIPM metadata standard (tentatively named version 0.5 and tentatively scheduled for release in fall 2007) will address each item in the above list. Also, an important next step regarding the metadata is to augment the metadata model itself so that it eventually grows to meet the needs of the broader justice community. Towards this goal, the federation wants to seek out new members that have substantially different metadata requirements than those of the current membership. In addition, several other metadata definition

efforts within the justice community are being considered for inclusion or reconciliation with the GFIPM metadata model. See the discussion of the DHS ABAC project and the Global Technical Privacy Task Team in Appendix A for more details about this.

7.4 Federation Enablement of Resources

For the GFIPM concept to succeed, it is critical that techniques exist through which a wide range of legacy resources in the law enforcement domain can be made available to federation users. The process of making a resource available to federation users is called ***federation enablement***. As the federation grows and gains wider adoption, resource owners will look to the GFIPM model to help them realize the basic value proposition of federated identity and privilege management: to achieve resource sharing with a large base of established users and partners who would normally not have access to their resources, while keeping costs low, providing a simplified and improved user experience (via single sign-on), and providing better security and privacy protection for users' personal data (via the reduction or elimination of redundant data capture and storage processes). But achieving federation enablement for a wide range of legacy resources can be challenging in that they can be very diverse in many aspects, including application architecture, implementation platforms and vendor products, type and structure of resource, application functionality, support model, security and access policies, etc. Many insights and lessons about federation enablement of resources have been gained throughout the demo project during the process of federation-enabling a handful of existing CISA, JNET, and RISS resources. They are discussed in this section of the report.

Applications and resources tend to have usage requirements that must be met by all of their users. Most usage requirements fall into the following categories.

- ***Terms of Use*** – The application may require that a user agree to specific terms of use prior to using it.
- ***Provisioning*** – The application may require that a user register a local account with it before using it.
- ***Inter-Session Persistence*** – The application may need to maintain state about the user from one session to another.
- ***Identification*** – The application may need to know the user's identity at all times while the user is using it.
- ***Access Control*** – The application may impose certain access restrictions based on some combination of the user's rank, certifications, role, or some other important personal characteristics.
- ***Auditing*** – The application may log all actions performed by a user in an audit log for review, compliance, etc.
- ***Personalization*** – The application may need to maintain miscellaneous personal data about a user for the purpose of delivering certain features. For example, locality information would help the application deliver a list of alerts or bulletins that specifically pertain to a user's region or locality.

It became clear during the demo project that the GFIPM concept will work only if resource owners are permitted to maintain control over the usage requirements of their resources within the federation and are not forced to modify the requirements in a manner they find unacceptable. GFIPM must allow a wide range of existing resources and applications to be federation-enabled and made available to federation users in a manner that fulfills the usage requirements of those applications. The GFIPM concept provides many valuable tools that help to simplify federation enablement of resources while allowing those resources to still meet their usage requirements.

- Federation-wide policy-level agreements and MOUs can form the basis of interagency trust which can be layered with additional bi-lateral or community agreements as required.
- The basic SAML-based infrastructure provides a standard means of authentication/identification of users and the convenience of SSO.
- The GFIPM metadata provides detailed personal information about individual federation users, including identification, contact information, affiliations, memberships, certifications, and basic data access privileges within the user's home organization. This information can be trusted because it comes to the resource from an authoritative source: the user's IDP.

The following subsections serve to help prospective federation members better understand the basic federation enablement options that are available to them for various categories of resources.

7.4.1 Resource Integration Profiles

The following ***resource integration profiles*** are based on common categories of resources and applications. They are intended to help resource owners better understand the level of effort required to federation-enable specific types of resources. Note that a specific resource may or may not fit neatly into a specific integration profile. The intent of this section is not to prescribe an exhaustive set of resource classes, but rather to provide enough detail about the critical differences between resources to illuminate the important issues that must be addressed when federation-enabling them.

Profile 1: Read-Only Content without Individual User Accounts

A resource that fits into Profile 1 would tend to have the following characteristics:

- Used for dissemination of information;
- Does not require a unique pre-provisioned user account for each user;
- May require the user's identity and contact information for auditing purposes;
- Requires some basic information about the user for access control;
- Does not require personalization data;
- Has no persistence requirement.

Profile 2: Resource with Individual User Accounts and Dynamic Provisioning

A resource in Profile 2 typically has the following characteristics:

- Its provisioning requirement can be met by GFIPM metadata and leverage the IDP user vetting without the need for any additional out-of-band communication or user vetting during the provisioning process;
- It requires the user's identity and contact information for auditing purposes;
- It requires information about the user at account provisioning time for provisioning the account's access control permissions;
- It may require personalization data;
- It has a requirement for persistence of user account information between sessions.

Profile 3: Resource with Individual User Accounts and Pre-Provisioning

A resource that fits into Profile 3 has these characteristics:

- It requires a unique pre-provisioned user account for each user;
- Its provisioning requirement ***cannot*** be met by GFIPM metadata and IDP vetting alone, as it requires out-of-band communication to facilitate a direct relationship with the user during the provisioning process, but GFIPM can provide single sign-on functionality with an account linking capability for it after the provisioning process is complete;
- It requires the user's identity and contact information for auditing purposes;
- It requires information about the user at account provisioning time for provisioning the account's access control permissions;
- It may require personalization data;
- It has a requirement for persistence of user account information between sessions.

The next section discusses the actual techniques that can be used for federation-enablement of resources that fit these integration profiles.

7.4.2 Resource Integration Techniques

The following ***resource integration techniques*** are presented to illustrate how many of the current federation resources were federation-enabled. They are intended to help resource owners better understand the types of techniques that can be used to federation-enable specific resources. Note - this list is not exhaustive and that any given resource may or may not work with any of the integration techniques described here.

Technique 1: GFIPM-Aware Reverse Proxy with No Secondary Authorization

Integration Technique 1 works best for a resource that meets the following criteria:

1. It is not natively GFIPM-aware;
2. It does not require local user accounts or user authentication as a prerequisite to access;

3. It does not require or have any notion of inter-session persistence or personalization;
4. It cannot be modified at the source code level due to policy restrictions or technical impracticalities.

In this integration technique, access to the resource is provided for federation members via network configuration by a reverse proxy service that is GFIPM-aware. (See Section 4.2.4 for additional discussion about this type of proxy.) Also, the proxy service must implement access control and auditing if they are required by the application.

Technique 2: GFIPM-Aware Reverse Proxy with Secondary Authorization

Integration Technique 2 applies to any resource that meets these criteria:

1. It is not natively GFIPM-aware
2. It does require local user accounts for access
3. It cannot be modified at the source code level due to policy restrictions or technical impracticalities.
4. Its access policy may require that each federated user have a unique account, or it may allow multiple users to share a group account.
5. It may or may not require inter-session persistence for an account, and it may or may not have any personalization requirements.

In this integration technique, access to the resource is provided for federation members via a reverse proxy architecture (as discussed in Section 4.2.4) in which the proxy maps each GFIPM user into a corresponding user account for the proxied resource. The mapping from GFIPM user to back-end resource account may be many-to-one (using local accounts at the resource as group accounts for GFIPM users) or one-to-one (using individual accounts). The proxy must authenticate the user to the proxied resource using the appropriate account.

If access to the resource is permitted via group accounts, then a typical configuration would consist of relatively few local accounts on the resource (e.g. the number of group accounts in the application), with each account corresponding to an access class, group, or role that would be used by many federation users. In this configuration, since the resource knows only group accounts, it has no way to know which specific GFIPM user is accessing it at any given time. Therefore, to implement precise end-to-end auditing, it is necessary to combine the proxy's audit logs with the resource's audit logs. Both JNET and CISA have successfully enabled federation-resources using group accounts in this manner.

If access to the resource requires each GFIPM user to have a unique back-end account with the resource, then it is necessary to provision individual user accounts for access. The most elegant approach to solving this problem is to allow user accounts to be provisioned via an account sign-up page on the back-end resource, and configure the proxy to populate the resource's account provisioning (sign-up) page with data extracted

from a GFIPM user assertion. If this type of dynamic account provisioning is not possible (for either policy-related or technical reasons), then the proxy can act as an account-linking bridge between GFIPM federated user accounts and local back-end accounts that have been pre-provisioned out-of-band. In either scenario, the proxy must somehow know how to map each GFIPM user ID to a back-end user ID on the resource. This may be accomplished in several ways, but the most straightforward technique is for the proxy to maintain a database or map from one domain to the other. In some cases where accounts follow a regular form, a set of rules or algorithmic mapping may be possible.

Technique 3: Native GFIPM Enablement

If it is possible and practical to modify the source code of a resource during the federation enablement process, then the resource can be configured so it natively understands GFIPM metadata and can exist in the GFIPM environment without the aid of a proxy. If the resource is a newly developed application, then it may be advantageous to design and build the application such that it fundamentally understands GFIPM metadata as its primary internal model for fulfilling its requirements related to identifying the user, enforcing access control rules, and auditing access. But if the application already exists, then it may be necessary to take an alternate approach and develop a GFIPM-aware module that exists within the application and serves to translate information from GFIPM metadata into the native user model that the application uses. Issues related to account provisioning (dynamic or out-of-band) are the same for this technique as they are in the discussion above for Technique 2.

7.4.3 Profiles and Techniques for Existing Resources

To provide a more concrete perspective on the discussion about integration profiles and integration techniques in the previous two sections, this section of the report contains summary information about how resources currently in the federation were federation-enabled. Table 6 lists each resource currently in the federation, along with its integration profile and the integration technique used to integrate it with the federation.

	Resource Name	Integration Profile	Integration Technique
CISA	Arizona Counter-Terrorism Information Center (ACTIC)	1	1
	Arizona Sex Offender Information Center	1	1
	Arizona Amber Alert	1	1
	Georgia Bureau of Investigation Sex Offender Registry	1	1
	Oklahoma State Bureau of Investigation Officer Safety Bulletin	3	2
	Texas Criminal Law Enforcement Online (CLEO)	3	2
	California Joint Regional Information Exchange System (JRIES)	3	2
	CISAnet Federated Query Tool • New Mexico Complete Arrest Information	2	3

	(CAI) • New Mexico Incident Based Reporting System (NMIBRS) • New Mexico Sex Offender Registration • New Mexico Missing Person & Unidentified Bodies • New Mexico Field Interview (FI) • New Mexico Law Enforcement Information Network with Corrections (LINC) • New Mexico Criminal Law Enforcement Reporting and Information System (CLERIS) • Arizona Criminal Investigative Database • Texas Criminal Law Enforcement Reporting and Information System (CLERIS)		
JNET	Pennsylvania Department of Corrections Intake/Exit Photos	3	2
	Pennsylvania Arrest Warrants Outstanding for Parolees who failed to report (Absconders)	3	2
	Pennsylvania State Prisoner Locator	3	2
	Pennsylvania Criminal Trial Case Information	3	2
	Pennsylvania Arrest Warrants Outstanding for Failure to Pay Child Support	3	2
	Pennsylvania Amber Alert	3	2
	GFIPM Lessons Learned	1	3
RISS	HSIN Counter-Terrorism Briefs, Reports, and Documents	1	3
	RISS Counter-Terrorism Briefs, Reports, and Documents	1	3

Table 6: Integration Profiles and Integration Techniques for Currently Existing Resources in the GFIPM Federation

Table 7 provides a concise summary of the information in Table 6. It provides a count of current federation resources by integration technique and by integration profile.

	Profile 1	Profile 2	Profile 3	TOTAL
Technique 1	4	0	0	4
Technique 2	0	0	9	9
Technique 3	3	1	0	4
TOTAL	7	1	9	17

Table 7: Summary Count of Federation Resources by Integration Profile and Integration Technique

It is important to note that while Table 7 provides some insight into the breakdown of GFIPM resources for federation enablement purposes, it is not necessarily representative of the broader set of information sharing resources in the justice community. For

federation growth and outreach purposes, it is important that the federation gain some insight into how well the current set of integration techniques can accommodate the broad range of resources that are potentially applicable to GFIPM. This information should become clearer as the federation begins to grow and expand beyond its current state.

7.5 Usability and User Support

The following insights and lessons apply to the topics of usability and user support, both of which are critical for wide-scale adoption and acceptance of the federation by users.

1. For usability within the federation, users need to be able to discover and access resources that are meaningful to their job functions. Naturally, as discussed in Section 7.2, it is critical that these resources actually exist in the federation. But beyond this, it is important that users know the resources are available. The most straightforward way to implement this functionality is via a resource directory or registry that describes each resource, provides a link to access the resource, and lists the resource's access requirements. For the purposes of the GFIPM demonstration project, GTRI implemented such a directory which is accessible through a web browser; however, as discussed in Section 7.6, the resource directory will need to be managed and overseen by the federation's governance structure as the federation grows and matures.

2. Additional tools and value-added products, such as federated search services, resource portals, and brokers, can be layered on top of the current federation infrastructure to provide more functionality or a better user experience. Like the resource directory, these services will usually need to be managed within the federation's governance structure. See Section 7.6 for additional discussion about this topic.

3. As the federation grows, it will be necessary to implement user support procedures that leverage existing user support resources (e.g. participant help desks, user training, etc.) that federation participants already have in place for their local users. When problems are encountered by users in accessing federation resources, collaboration between federation partners will be necessary for identification and resolution. Additionally, providing complex applications requiring special training to an extended federation user base will need to be considered in the deployment decision by SPs. Additional investigation into this topic is a critical next step for the project. See Section 7.6 for additional discussion about this.

4. Basic usability features and considerations, such as centrally available online user support resources, will serve to increase usability and reduce the cost burden on the federation's user support infrastructure. A basic GFIPM User Portal is already online at http://gfipm.net/users/ for this purpose. Expanding and improving the portal as needed is an important next step for the project.

5. Even though SPs control their own resources (as previously discussed in Section 7.4 and elsewhere in this report), for usability the federation needs resource owners to cooperate enough to ensure a reasonable level of consistency across the federation in terms of access requirements for specific types of data. Different access requirements for the same type of data (e.g. intelligence, counter terrorism, etc) between federation participants could lead to confusion among federation users and become a barrier to information sharing.

7.6 Governance and Operational Support

Throughout the demonstration project, it became clear that if the GFIPM federation is to thrive and grow into a wide-scale operational system, it must be effectively managed and governed. This section of the report summarizes the most important lessons learned about the governance and operational support infrastructure that needs to exist for GFIPM as the federation grows beyond the demo phase.

1. Several important governance-related support functions have been identified for the purpose of keeping the GFIPM federation online and working smoothly. These include policy management, operations, engineering, technical assistance, and outreach. Figure 21 illustrates and summarizes the role of each support function.

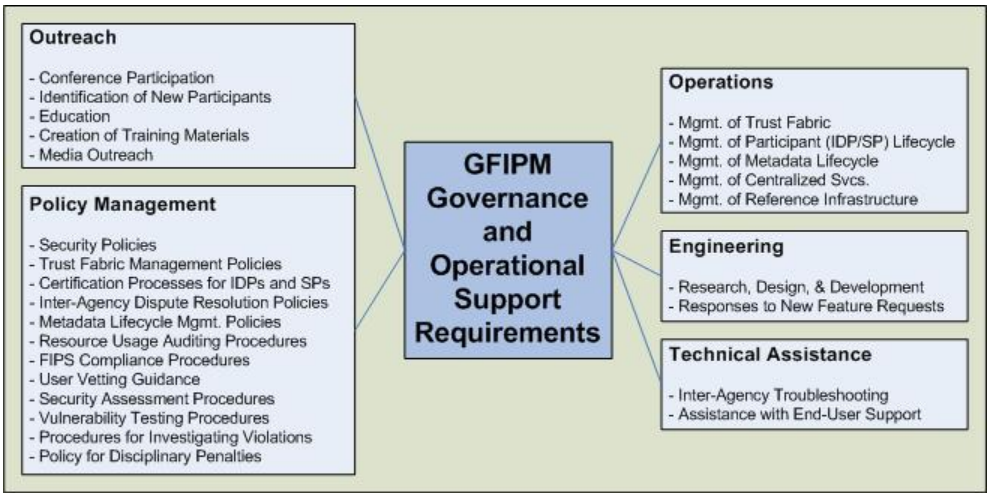

Figure 21: GFIPM Governance and Operational Support Requirements

2. A ***policy management*** capability is required to develop policies and procedures to govern the behavior of other operational support functions, as well as memoranda of understanding for basic inter-agency trust and other critical issues. As outlined in figure 19, the scope of federation-wide policies and procedures includes, but is not limited to, the following items:

    - Security Policies;
    - Trust Fabric Management Policies;

- Certification Processes for IDPs and SPs;
- Inter-Agency Dispute Resolution Policies;
- Metadata Lifecycle Mgmt. Policies;
- Resource Usage Auditing Procedures;
- FIPS Compliance Procedures;
- User Vetting Guidance;
- Security Assessment Procedures;
- Vulnerability Testing Procedures;
- Procedures for Investigating Violations;
- Policy for Disciplinary Penalties.

These policies should establish a minimal set of baseline requirements for participation in the federation, and may be superseded by more stringent peer-to-peer policies as needed.

3. An ***operations management*** capability is required to oversee and carry out day-to-day processes and procedures related to the federation trust fabric, participant lifecycle process, reference infrastructure, metadata lifecycle process, and centralized services. Each of these topics is discussed in more depth below.

    - ***Management of Federation Trust Fabric*** – As discussed in Section 4.5, the GFIPM federation relies on a cryptographic trust fabric document that conforms to the SAML 2.0 standard. This metadata includes public keys for all SAML protocol endpoints in the federation (at both IDPs and SPs), and is digitally signed so that IDPs and SPs can establish trust with each other by simply validating signatures against public key certificates in the trust fabric document. To ensure that this cryptographic trust is valid and grounded in reality, the federation must securely manage the signing key used to sign the trust fabric document. In addition, the federation must ensure that new versions of the trust fabric document are disseminated to IDPs and SPs as needed whenever a security-critical event occurs. Examples of security-critical events include the addition of a new IDP or SP to the federation, removal of an existing IDP or SP from the federation, or expiration of the current trust fabric document.

    - ***Management of the Participant Lifecycle Process*** – When new IDPs and SPs join the federation, there are policy and technical functions that must be fulfilled. These may include, but are not necessarily limited to the following: (1) an IDP/SP certification process including a security audit and/or vulnerability testing, (2) interoperability testing with the federation's reference infrastructure, and (3) a formal review of the new IDP's or SP's security policies. Similarly, when an IDP or SP leaves the federation, certain actions may need to be performed on the federation's behalf. These actions may differ depending on the reason why the IDP or SP is leaving the federation.

- ***Reference Infrastructure Management*** – A federation test/reference infrastructure capability is important for new IDPs and SPs to use as a working reference for debugging prior to joining the live federation. It is also useful for experimenting with new techniques for resource proxying, SAML metadata encoding, and anything else that is too risky or too difficult to do in the live federation. (Section 5.4 contains more information about the GFIPM Reference Federation.) For this test environment to have lasting value for the federation, it must be managed and maintained properly.

- ***Metadata Lifecycle Management*** – As discussed in Section 7.3 of this report, one of the implications of having a federation-wide standard metadata model is that the model must be maintained and managed. As the federation grows and evolves, members will require that various changes and additions be made to the metadata standard. This is a political process that requires procedural oversight. It is the job of the GFIPM governance body to define the precise rules and procedures for managing the metadata. But regardless of what specific metadata management policies are chosen, it is clear that the following questions must be addressed.

    - How does a proposed metadata change become part of the metadata used by the federation?
    - How are conflicts resolved between participants?
    - Is there a voting process involved? Is there a pre-defined cycle of change involving proposals, requests for comments, etc.? For example, does the federation always release a new metadata version every six months, or does the federation wait until "enough" changes have accumulated? How many changes are "enough"?

  After a set of metadata changes have been accepted and a new metadata model has been agreed upon in concept, it must be made into a new metadata package including schemas, definitions, assertion guidance, and other documentation as needed. Then the metadata must be officially released and made available for federation members to use. The release of a new metadata version will bring about another set of questions and issues, such as the following:

    - It is likely that not all IDPs and SPs will want to upgrade on Day 1 when a new metadata version is released. How does the federation accommodate this? Can multiple versions of the metadata be used simultaneously within the federation? If so, how are they managed so that they can coexist without causing confusion?
    - Will old versions of metadata be phased out over time? How will the phase-out process occur?

- If metadata releases occur infrequently, then how can the federation accommodate new resources and SPs that require metadata changes? Can the metadata model support the notion of schema extensibility to handle temporary deficiencies in its robustness or pairwise agreements that require specialized metadata schema extensions between participants?

Most of the questions raised here represent open issues that require careful consideration. Clearly, there must be a management and support capability in place so that the federation can effectively manage the metadata lifecycle process.

- ***Management of Current and Future Centralized Services*** – At the time of this writing (June 2007), the only centralized services in the GFIPM federation are the WAYF service and the GFIPM User Portal (which contains a directory of federation resources). But as the federation grows, it may be necessary to develop and deploy additional centralized services, as discussed in Section 7.5. All such centralized services must be managed by the federation.

4. An ***engineering*** capability is required to help the federation address longer-term research, design, and development tasks related to important technology issues, as well as for managing and responding to bugs and engineering change requests for new feature.

5. A ***technical assistance*** capability is required for the purpose of managing interagency technical troubleshooting and problem resolution, issue tracking, and end-user help desk support. The federation needs to provide a basic user support infrastructure, and such a capability needs to leverage the user support infrastructure that already has been put in place provided by each federation participant.

6. An ***outreach*** capability is required for the purpose of identifying new participants, overseeing the education process (for end users, help desk staff, system administrators, etc.), and marketing via participation in conferences and media.

8 Next Steps

Sections 1-7 of this report focus on past activities from the demonstration phase of the GFIPM initiative. This section looks forward and provides a brief outline of the next steps for the project. The next steps focus on three critical activities for the federation: continuing to develop the GFIPM standards, expanding the federation to include more participants and production-level capabilities, and establishing a federation governance structure. Additionally, GFIPM will be extended and evaluated to support other use cases (see Section 3.3.4). Each of these major categories of activity is discussed in its own subsection here.

8.1 GFIPM Standards Development, Validation, and Vetting

As the GFIPM federation expands and becomes more widely known throughout the justice community, the standards on which it is based will come under increasing scrutiny. Therefore, these standards must be well-defined and appropriate for the task of justice information sharing. Towards that end, the following activities have been identified as appropriate next steps in the GFIPM standards maturation process.

- Incorporate lessons learned and recognized shortfalls in the continued development of the GFIPM metadata model through version 0.5 and beyond.
- Expand the GFIPM metadata to support additional privileges, data classifications, privacy requirements, and alignment with other identified standards, including those defined by the DHS ABAC project and the Global Technical Privacy Task Team. (See Appendix A for more information about these initiatives.)
- Update GFIPM standards to conform to NIEM 2.0 and then submit GFIPM metadata content for incorporation into the NIEM standard.
- Develop and document SAML 2.0 profile and encoding standards for GFIPM metadata assertions.
- Begin the process of further refining metadata into a set of common usage profiles that address the needs of most legacy applications that are or may become candidates for joining the federation.
- Develop and document standards for the establishment and maintenance of the federation trust fabric and infrastructure, including SAML 2.0 federated entities metadata, certificates for federation-wide trusted signing keys, and related information as needed.
- Extend GFIPM concepts and standards to support Global's Justice Reference Architecture, which encompasses the service-oriented architecture (SOA) and addresses the system-to-system use case for justice information sharing.
- Facilitate the vetting of all standards through an iterative process. (Many of the critical components of this process have been identified in Section 7.6 of this report.)
- Validate all standards and products used in the GFIPM federation from a technical standpoint. At a minimum, this includes the SAML Usage Profile for GFIPM and the SAML implementations (both COTS and open source) used in GFIPM.
- Incorporate metadata and feedback from activities in Section 8.2.

8.2 Expansion of Participants and Development of Production-Level Capabilities

Whereas the previous section addressed standards-related tasks that must be completed during the next phase of the GFIPM project, this section identifies tasks that will serve to help the federation grow and mature in terms of operational legitimacy in the eyes of users and prospective participants as well as validating the standards developed in Section 8.1 in operational use. The following activities have been identified.

- Incorporate lessons learned that have been identified by demo project participants in the areas of security, usability, interoperability, implementation, administration, performance, scalability, etc.
- Upgrade the GFIPM federation to SAML 2.0 and the latest GFIPM metadata standard (version 0.5) after it is released.
- Perform interoperability tests with multiple COTS commercial products.
- Define and develop mechanisms to add new federation partners (IDPs or SPs) to the GFIPM production environment.
- Develop testing procedures and tools to facilitate interoperability testing and problem identification and resolution.
- Provide tools, assistance, and documentation to reduce the time and cost of joining the federation.
- Expand the federation participant base by 6-8 agencies selected on the basis of additional users that they can bring into the federation and additional potential metadata that could be discovered.
- Demonstrate GFIPM in support of Global's Justice Reference Architecture service-oriented architecture (SOA) system-to-system use case for justice information sharing between federation participants.
- Establish and test inter-federation data exchanges with the LEISP identity management pilot federation. (See Appendix A for more information about the LEISP program.)
- Develop processes, mechanisms, and tools for collecting information about the demographic makeup of the federation's user base across certain basic dimensions such as job function, certifications, and geographic locality, for the purpose of outreach and recruitment of new resources into the federation.
- Develop processes, mechanisms, and tools for collecting information about the resources that are available in the federation, usage statistics for specific resources, and success stories or case studies highlighting the value that federation resources have provided to the current user base, for the purpose of outreach and recruitment of new IDPs and users into the federation.

8.3 Establishment of Federation Governance Structure

The final activity that has been identified as a next step for GFIPM is to begin the process of establishing a federation governance structure. Section 7.6 identifies the basic responsibilities of such a governance organization, and it also implies certain additional functions (e.g. engineering, operations management, and technical assistance) that are necessary for the federation's ongoing operation. All of these functions flow out of the governance structure, and need to be defined and staffed accordingly.

Appendix A: Related Initiatives and Standards

This appendix contains brief descriptions of several initiatives and standards considered by GFIPM participants to be important to the future of the GFIPM program. Each item is introduced and described in the context of its relationship with GFIPM.

A.1 E-Authentication

The E-Authentication Initiative aims to provide trusted, secure, standards-based authentication architecture to support Federal E-Government applications and initiatives. This approach intends to provide a uniform process for establishing electronic identity, thereby eliminating the need for each initiative to develop a redundant solution for the verification of identity and electronic signatures. E-Authentication's distributed architecture will also allow citizens and businesses to use non-government issued credentials to conduct transactions with the government.

A.2 Law Enforcement Information Sharing Program (LEISP)

The LEISP Federated Identity Management pilot, initiated in 2006, is under the direction of the U.S. Department of Justice (DOJ) Office of the Chief Information Officer (CIO), with funding provided by the Program Manager for the Information Sharing Environment (PM-ISE), housed within the Office of the Director of National Intelligence (ODNI). The project was initiated in response to the need for a comprehensive strategy for authentication for the LEISP to address participation of autonomous federal, state, local and regional systems and their existing user bases. The pilot is focused on the sharing of law enforcement and counterterrorism information with an initial set of pilot participants, including select users and applications from DOJ, DHS, FBI (LEO), and RISS.

The LEISP initiative is similar to GFIPM in that it employs the concept of Federated Identity Management as a cornerstone of its secure information sharing strategy; however, there are some important distinctions in focus and approach between LEISP and GFIPM.

1. The LEISP pilot's initial objective is to address the authentication and identification of federated users; leveraging existing vetting and providing a single sign-on capability and facilitating the linking of a federated user to an existing application account. A minimal set of metadata has been defined and agreed upon by federation participants for that purpose. GFIPM addresses these aspects as well but has the additional objective of providing a framework for privilege management based on rich set of metadata about federated users. The LEISP approach could support such a framework once consensus on such metadata is reached.
2. The LEISP FIdM pilot implements a trusted broker architecture in which identity providers and service providers establish a direct trust relationship with an intermediary. The implementation of the federation is facilitated through a central

services contract associated with the trusted broker for integration of new federation participant identity providers and service providers. Identity providers do not make assertions directly to service providers; instead they communicate with service providers through the intermediary, called a trusted broker. GFIPM employs the standard SAML 2.0 distributed peer-to-peer trust model to create a cryptographic (PKI-based) trust fabric directly between federation identity providers and service providers; the public keys for all providers are published in the federation's SAML 2.0 metadata. Providers trust all hosts within the SAML 2.0 metadata file. The SAML 2.0 metadata is digitally signed by a trusted source, so that identity providers and service providers can update and validate this configuration file in an automated fashion.

A.3 National Information Exchange Model (NIEM)

NIEM is a federal, state, local, and tribal interagency initiative providing a foundation for seamless information exchange (www.niem.gov). It leverages the data exchange standards efforts successfully implemented by the Global Justice Information Sharing Initiative (Global) and extends the Global Justice XML Data Model (GJXDM) to facilitate timely, secure information sharing across the whole of the justice, public safety, emergency and disaster management, intelligence, and homeland security enterprise. Given the work and success of the GJXDM and NIEM data modeling efforts, it is important to leverage and reuse these specifications in describing the GFIPM metadata. The advantage of the NIEM specification is that it inherently makes the model immediately more applicable to other domains and systems, rather than focused on criminal justice users and systems.

A.4 eXtensible Access Control Markup Language (XACML)

XACML is an OASIS open standard for specifying access control policies in XML against objects that are themselves defined in XML. It defines an access control policy language as well as a data flow model describing how policies are to be interpreted and enforced. It is anticipated that XACML will become an important tool in the eventual realization of Global's goal of facilitating information sharing throughout the law enforcement community. As Global begins to adopt a common metadata model (such as the GFIPM model), it will become necessary to exchange information about the access control and privacy policies of protected resources in terms of the elements in that metadata model. XACML can and likely will provide a standards-based infrastructure for exchanging this information.

A.5 Service Provisioning Markup Language (SPML)

SPML is an XML-based framework for exchanging user, resource and service provisioning information between cooperating organizations. In this context, provisioning consists of the automation of all steps required to manage (setup, amend, and revoke) user or system access entitlements or data relative to electronically published services. The goal of SPML is to allow organizations to securely and quickly set up user interfaces for web services and applications, by letting enterprise platforms such as web portals, application servers, and service centers generate provisioning requests within and

across organizations. This can lead to automation of user or system access and entitlement rights to electronic services across diverse IT infrastructures. As the GFIPM project and the entire Global initiative matures and grows to encompass a large, interconnected network of services and users, SPML may become a useful tool in the management of provisioning for access to services across the network.

A.6 Global Technical Privacy Task Team

The Global Technical Privacy Task Team is a technical standards group that is working on the task of translating privacy and requirements into technical specifications and standards. It focuses on identification, specification, and addressing standards for privacy for the Global Justice Information Sharing Initiative (Global). The scope for the tasks and deliverables of this task team covers the entire justice domain, including local, state, regional, tribal, and federal organizations.

One of the responsibilities of this team is to identify a set of privacy-related metadata that apply to information sharing transactions within the law enforcement community. It is anticipated that the metadata identified in this effort will be included as necessary in a future version of either the GFIPM metadata standard.

A.7 DHS Project on Authority-Based Access Control (ABAC)

The DHS ABAC initiative seeks to identify a set of base attributes that are critical for authorization decisions for current and future government information sharing systems, and also to identify the authoritative source of each critical attribute. The goal of the project's work is to identify attributes that can enable rule-based access control to be implemented based on the most current information available about a user at the time of an access attempt. The long-term vision of the project is for each attribute to be queried as needed from the attribute authority that is deemed to be authoritative for that attribute, thereby relieving applications of the need to constantly keep access control lists current.

The overarching vision of the ABAC project is similar to GFIPM in the sense that both projects seek to relieve applications and resources of the burden of directly managing data about users. In addition, the efforts of the ABAC project will provide valuable input to the GFIPM metadata model in terms of specific attributes that are of critical importance for government information sharing. It is expected that the ABAC work will be incorporated into the GFIPM metadata model in the near future.