TIES Security

Contents

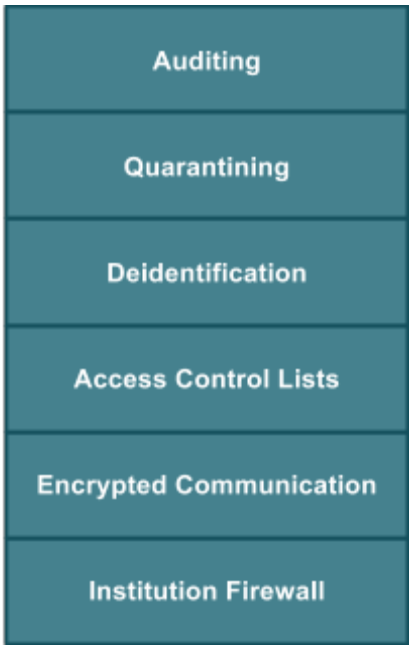
What is TIES?

TIES is a federated network of clinical document servers. Each TIES participating organization will serve up some subset of its Surgical Pathology Reports. TIES Users can search across all participating organizations and create Order Sets which are merely folders full of Pathology Reports. Since each Pathology Report represents materials like slides and FFPE tissue, researchers may use their Order Set to get real physical materials delivered for subsequent physical experimentation.

TIES is designed to allow the assembly of unique experimental cohorts. Potentially gathering rare samples from across the network.

How Does TIES Secure PHI?

The answer is in “layers”.



Deidentification

Each TIES Organization is required to purchase and install the commercial DeIDentification Software DeID. When the Organization begins to load its reports into TIES they will be automatically piped through DeID and the deidentified copy will be placed in the “public” database while the original copy is placed in the “private” database. DeID has been proven to remove of 99% of the HIPPA Safe Harbor compliant identifiers while keeping 95% of the non PHI.

Firewalls

The TIES server machine runs three servers on three ports. The Organization Systems

Group is responsible to get these configured with the appropriate level of protection. As mentioned

- **Port 80 HTTP Incoming/Outgoing** - will serve up the De-identified Path Reports to all IPs.
- **Port 90 HTTP Incoming/Outgoing** - should be opened to the Organization's Intranet. (i.e., those sub domains within the Organization encompassing potential Honest Broker access) Honest Brokers will need to see PHI in order to gather real materials and ship them to the Research Protocol.
- **Port 3306 TCP/IP Localhost** - is the default MySQL database server. This should be accessible only by the TIES server machine itself. Most default firewall configurations will have no exceptions for port 3306.
- All other firewall and distributed computing infrastructure can be configured per typical Organizational Protocol for a Virtual Machine Public Web Site.

Application Enforced Communication Security

TIES Web Applications are secure sites running the equivalent of HTTPS protocols. The difference is that TIES encrypts and digitally signs only the message content and not the message header. SSL - Secure Socket Layer provides the foundation of TIES communication security just as it does across the internet for online banking and commerce systems. TIES provides for what has come to be known as the three pillars of internet security.

Authenticity - Each user has a fully qualified distinguished name and a password required to be sufficiently strong. (i.e., letters, numbers, special characters). Message content passed between sites contain the user information embedded so each server knows exactly who is requesting a resource. All user names passwords, private and public keys are stored at the user's organization's database.

Privacy - Message bodies are encrypted using RSA 1024 asymmetric key cipher. Consequently each TIES site and user has a public/private key pair generated

by TIES at initiation time.

Integrity - Digital Signature technology assures that message content has not been tampered with while in transit. TIES uses MD5. Essentially a one way hash is generated by MD5 of the outgoing message. At pickup time the same algorithm is applied to the payload and compared to the hash sent from the caller.

GSI Grid Security Infrastructure (Java Implementation)

TIES uses the Globus Toolkit 4.2.2 Java implementation which incorporates multiple security components that establish the identity of users or services (authentication), protect communications, and determine who is allowed to perform what actions (authorization), as well as manage user credentials. The Web Services portion of GT 4.2.2 uses SOAP over HTTP for communicating messages. WS Authentication & Authorization in Java (Java WS A&A) implements the WS-Security standard and the WS-SecureConversation specification to provide message protection for SOAP messages. Features include:

- authentication of the sender
- encryption of the message
- integrity protection of the message
- replay attack protection

Java WS A&A provides a secure channel by using HTTP over SSL/TLS (HTTPS) for transporting the messages. This security mechanism supports all of the security features provided by SSL/TLS with the addition of support for X.509*Proxy Certificates*. The Authorization Framework component of Java WS A&A provides the infrastructure to process attributes and protect resource access based on access policy. It allows for

authorization policy to be configured and enforced at various levels of granularity (container, service or resource). It also provides client-side authorization to allow clients to authorize the services they access. The framework is pluggable and can be configured to use custom mechanisms for attribute collection and policy evaluation. It also provides multiple authorization module implementations; for example, support for gridmap-based authorization, a callout module that uses the SAML protocol to query an external service for an authorization decision. TIES uses role-based secure web services flavors including ***Anonymous Encrypted*** along with ***gridMap Authorization*** depending on the context of the communication***.***

Public Key Cryptography

TIES GSI uses public key cryptography (also known as asymmetric cryptography) as the
basis for its encryption functionality. The most important thing to know about public key
cryptography is that, unlike earlier cryptographic systems, it relies not on a single key (a
password or a secret "code"), but on two keys. These keys are numbers that are mathematically related in such a way that if either key is used to encrypt a message, the other key must be used to decrypt it. Also important is the fact that it is next to impossible (with our current knowledge of mathematics and available computing power) to obtain the second key from the first one and/or any messages encoded with the first key.

By making one of the keys available publicly (a public key) and keeping the other key private (a *private key*), a person can prove that he or she holds the private key simply

by encrypting a message. If the message can be decrypted using the public key, the person must have used the private key to encrypt the message.

Digital Signatures

Using public key cryptography, it is possible to digitally "sign" a piece of information. Signing information essentially means assuring a recipient of the information that the information hasn't been tampered with since it left the sender's hands.

To sign a piece of information, a mathematical hash of the information is first computed. (A hash is a condensed version of the information. The algorithm used to compute this hash must be known to the recipient of the information, but it isn't a secret.) The hash is then encrypted using the cryptography method previously described, after which it is attached it to the message.

To verify that your signed message is authentic, the recipient of the message will compute the hash of the message using the same hashing algorithm used by the sender, and will then decrypt the encrypted hash attached to the message. If the newly-computed hash and the decrypted hash match, it proves the integrity of the message. Digital signature is inherent in all TIES GSI communications.

Certificates

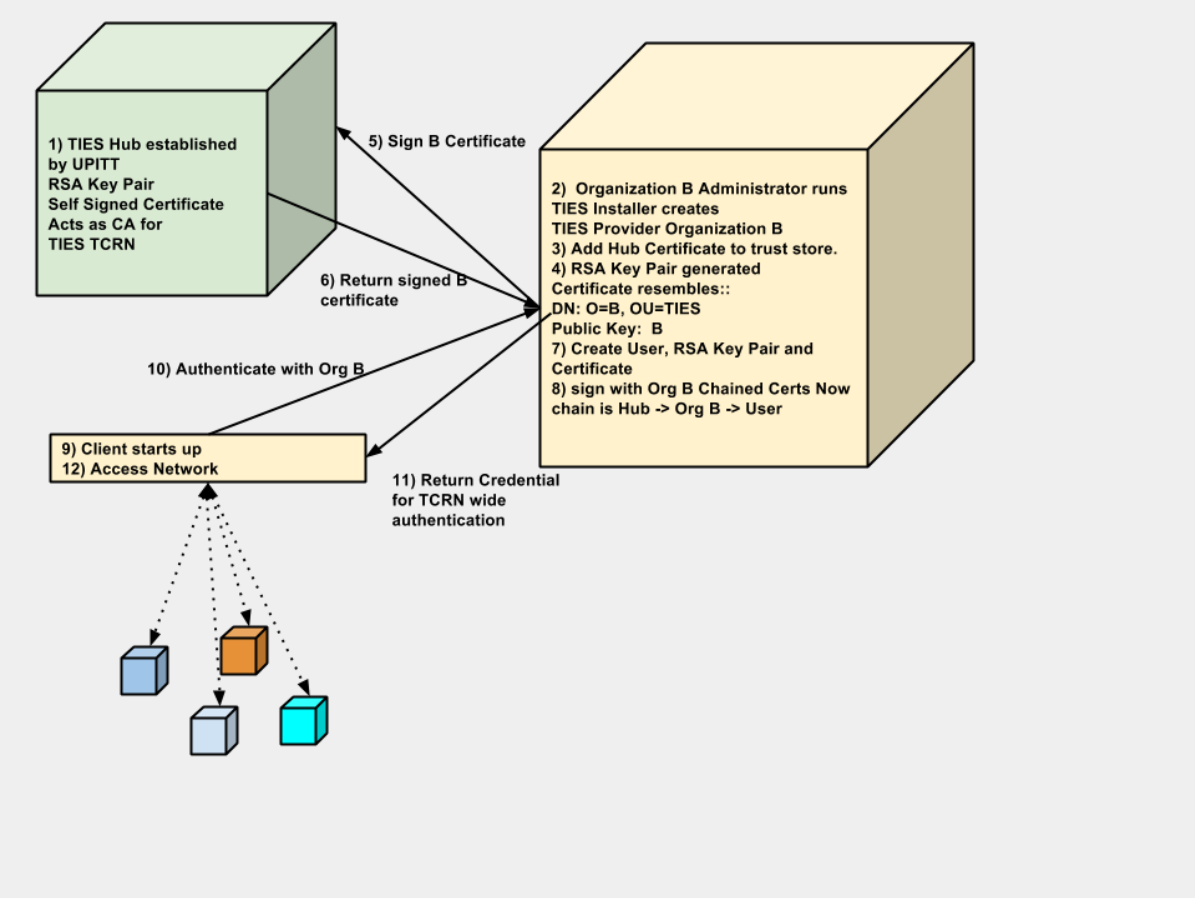
A central concept in TIES GSI authentication is the *certificate*. Every user and service on the TIES TCRN is identified via a certificate, which contains information vital to identifying and authenticating the user or service.

A TIES GSI certificate includes four primary pieces of information:

- A subject name, which identifies the person or object that the certificate represents.
- The public key belonging to the subject.
- The identity of a *Certificate Authority (CA)* that has signed the certificate to certify that the public key and the identity both belong to the subject.
- The digital signature of the named CA.

Typically a third party (a Certificate Authority CA) is used to certify the link between the public key and the subject in the certificate. In order to trust the certificate and its contents, the CA's certificate must be trusted. The link between the CA and its certificate must be established via some non-cryptographic means, or else the system is not trustworthy.

TIES TCRN establishes this link at network configuration time as illustrated by the following graphic.



GSI certificates are encoded in the X.509 certificate format, a standard data format for certificates established by the Internet Engineering Task Force (IETF). These certificates can be shared with other public key-based software, including commercial web browsers from Microsoft, Mozilla, and Google.

Application Enforced Authorization

TIES has different roles for users that provide different levels of access to the data. The Preliminary User role only allows access to aggregate numbers. The Researcher role allows access to de-identified patient data. The Honest Broker role allows access to PHI. The Document Providing Organization Administrator has the power to assign roles to users and access rights to a Research Study. Once access is granted each researcher who is a member of the study will be placed in the Providing

Organization's Access Control List. When a member of TIES wants to access documents from a local organization not their own, they must be working in context of a sanctioned Research Study.

Quarantining

TIES Researchers must accept a Data Use Agreement each time they log in. As a good TIES citizen it is each participant's responsibility to Quarantine any report that has be "under scrubbed" (i.e., has not be completely de-Identified) . Periodically the TIES Organization administrator can review the quarantined reports and explicitly scrub them by hand. Or for a systemic problem DeID Software might be adjusted by extending its local rules and lookup lists.

Auditing

TIES maintains comprehensive audit logs for all various types of activities including access to identified document in a table in the database. These records are never automatically expunged and it is up to your institution to decide when to expunge or archive these records.

The following types of activities are logged

a. Successful authentications
b. Unsuccessful authentications
c. Password change and reset
d. Searches
e. De-identified and identified document access

For each log entry the following information is captured wherever applicable

a. Activity Type
b. Date and time.
c. Userid
d. User's First and Last Name

e. User Role

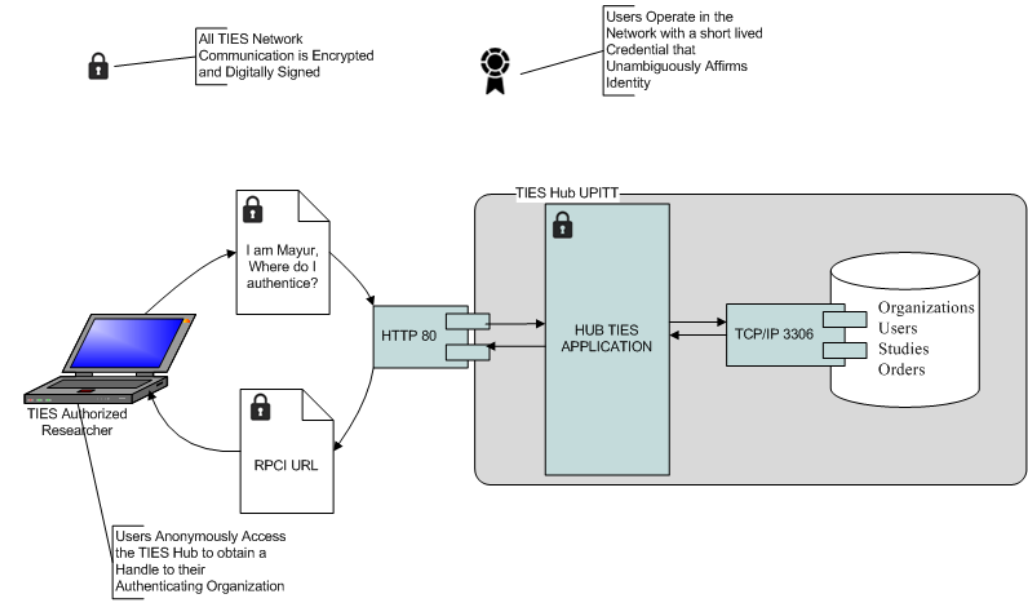
f. Protocol / Study user has logged into.

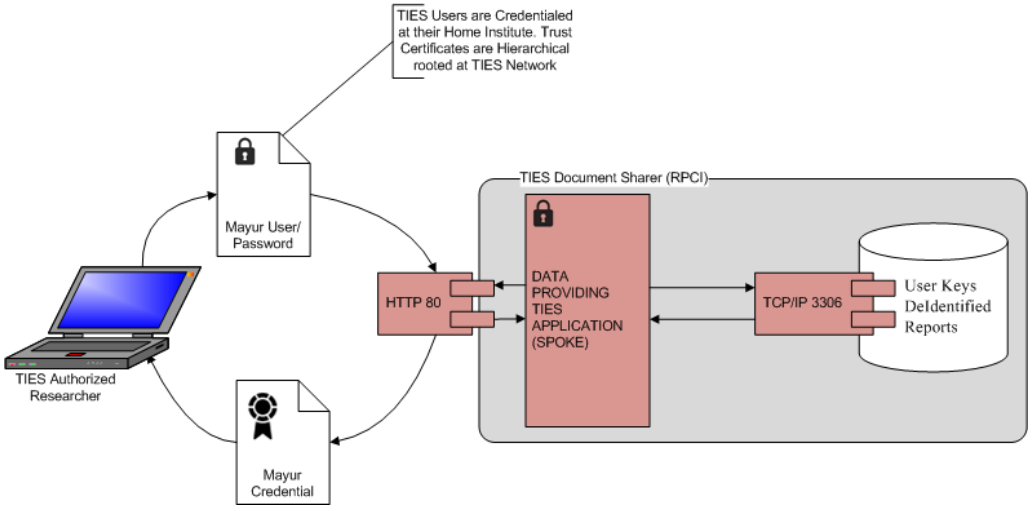
g. Any additional information pertinent to the activity type.

How Authentication Works

TIES is a network of federated document servers. TIES participants are affiliated with one and only one authentication server. That server is the TIES node for that participant's Organization. As seen in Figure 1 authentication to the TIES network is a two step process.

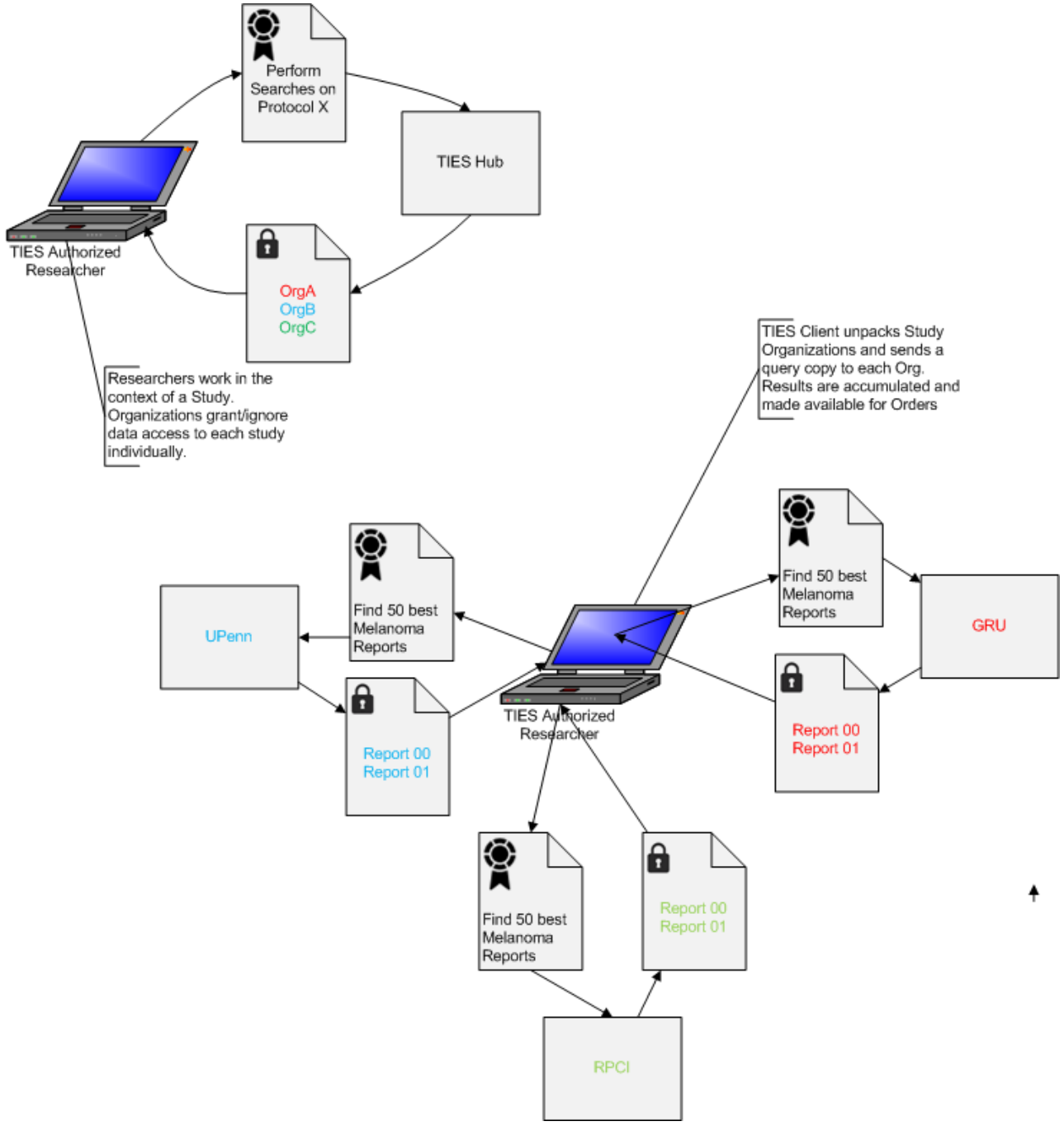
First the authenticating organization is determined and second the user authenticates to that organization. The result is a "credential" that unambiguously identifies the user during her TIES session.





How Authorization and Access Control Works

TIES Researchers work in the context of a Research Study or "Protocol". Each network Document Provider must explicitly choose to participate in a given Study. Figure 2 shows how Study based authorization works in TIES.



Appendix A Technical Documentation for Security layer

Globus Web Services Resource Framework - This is the key “GRID” technology underlying TIES. Web Services provide stateful access to “resources” which may be files, databases, or computational components.

Globus® Toolkit is a fundamental enabling technology for the "Grid," letting people share computing power, databases, and other tools securely online across corporate, institutional, and geographic boundaries without sacrificing local autonomy. The toolkit includes software services and libraries for resource monitoring, discovery, and management, plus security and file management. TIES is built upon the Java implementation of the Globus Toolkit.

The **Grid Security Infrastructure** (GSI), formerly called the **Globus** **Security Infrastructure**, is a specification for secret, tamper-proof, delegatable communication between software in a grid computing environment. Secure, authenticatable communication is enabled using asymmetric encryption.

OGSA-DAI, Open Grid Services Infrastructure Data Access and Integration. Provides Grid based access to relational database management systems. TIES uses this technology to implement some of its middleware on top of Globus.

Appendix B: Security FAQ

Are there individual accounts for all users of the system so that there is accountability for the integrity of the data within the system?

Yes

Accounts capable of expiring?

Yes

Do the passwords have complexity requirements?

Yes. Passwords must contain characters from three of the following five categories:

* Uppercase characters
* Lowercase characters
* Base 10 digits (0 through 9)
* Nonalphanumeric characters: ~!@#$%^&*_-+=`|\(){}[]:;"'<>,.?/
* Any Unicode character that is categorized as an alphabetic character but is not uppercase or lowercase.

Is access to the system granular enough so that people who fit a particular role can only see the data associated with that roll?

Yes

Does user activity in the system get logged?

Yes, the following types of activities are logged

a. Successful authentications
b. Unsuccessful authentications
c. Password change and reset
d. Searches
e. De-identified and identified document access

3. For each log entry the following information is captured wherever applicable

a. Activity Type
b. Date and time.
c. Userid
d. User's First and Last Name
e. User Role
f. Protocol / Study user has logged into.
g. Any additional information pertinent to the activity type.

Does the system have automatic backout, logout, or lockout in order to reduce the risk of patient information being left on a screen?

Yes

Are internal communication lines and external (if any) communications protected?

Yes. RSA 1024 cipher strength message level security for all communications between clients and the server.

Does all access to data occur through controlled programming to reduce the risk of back-end data changes reducing the integrity of the system?

Yes

Did the vendor supply all the required system documentation? Is it readily accessible?

Yes

Do all users have access to user level documentation?

Yes