Network Working Group
Internet-Draft
Intended status: Informational
Expires: December 14, 2013

T. Freeman
Microsoft Corp.
J. Schaad
Soaring Hawk Consulting
P. Patterson
Carillon Information Security Inc

June 12, 2013

Requirements for Message Access Control
draft-freeman-plasma-requirements-06

Abstract

There are many situations where organizations want to protect information with robust access control, either for implementation of intellectual property right protections, enforcement of contractual confidentiality agreements or because of legal regulations. The Enhanced Security Services (ESS) for S/MIME defines an access control mechanism for email which is enforced by the recipient's client after decryption of the message. The ESS mechanism therefore is dependent on the correct access policy configuration of every recipient's client. This mechanism also provides full access to the data to all recipients prior to the access control check, this is considered to be inadequate due to the difficulty in demonstrating policy compliance.

This document lays out the deficiencies of the current ESS security label, and presents requirements for a new model for providing access control to messages where the access check is performed prior to message content decryption. This new model also does not require policy configuration on the client to simplify deployment and compliance verification.

The proposed model additionally provides a method where non-X.509 certificate credentials can be used for encryption/decryption of S/MIME messages.

This Internet-Draft will expire on January 20, 2012. 99

Copyright Notice

Table of Contents

1 Introduction

The S/MIME (Secure/Multipurpose Internet Mail Extensions) standard [RFC5652] today provides digital signatures (for message integrity and data origination) and encryption (for data confidentiality). The Enhanced Security Services (ESS) for S/MIME [RFC5035] provides for additional services including security labels (eSSSecurityLabel) which represent the access control policy. The label is a signed attribute in the signed data block of a message. The recipient of the message is responsible for checking that the recipient has a legitimate right to see the message based on the label. This type of security labeling is similar to that of stamping "Top Secret" on the cover of a document. It relies on the reader to not open and read the document when the policy is discovered.

The Cryptographic Message Syntax (CMS) [RFC5652] allows for a variety of different types of lock boxes to be applied to an encrypted message. This allows for a variety of different type of security mechanisms to be used by the sender and the recipient to process the message. However the S/MIME standard is currently solely based on X.509 certificates. This means anyone without an X.509 certificate is unable to leverage the S/MIME protocol for securing Email. The vast majority of users on the Internet have other forms of credentials (passwords, one time passwords, PGP keys etc.).

1.1 Data Access Control

There are many situations where organizations want to include information which is subject to regulatory or other complex access control policy in Email. Regulated information requires some form of robust access control to protect the confidentiality of the information. While ESS for S/MIME [RFC5035] defines an access control mechanism for S/MIME (eSSSecurityLabel), it is an extremely weak form of access control as the recipient is responsible for the enforcement and is given access to the data even if they fail to meet the access criteria as defined by the label.

An access control policy defines a set of criteria and evaluation logic that must be satisfied in order to grant access to the information. This criteria can be defined in terms of group membership if the policy is a conventional Discretionary Access Control (DAC) policy. If the policy is a Role Based Access Control Policy (RBAC) they are defined in terms of what role the subject needs to belong to. If the policy is an Attribute Based Access Control (ABAC) policy it is defined in terms of attributes about the subject, their device or environment, their intended action on or use of the information and the resource. Examples of the types of attributes would include attributes about the subject such as their

employers identity, their nationality, citizenship etc., or attributes about their device such as its name, boot state. Standards now exist that enable the transport of attributes [SAML-overview].

An ESS Security label is a signed attribute of a SignedData object which indicates the access control policy for the message. While an ESS Security Label provides a standardized representation of an access policy identifier, it does not define any methods of obtaining the necessary information to satisfy the policy or policy description in order to enforce the policy. The fact that this is a signed attribute protects the integrity of the ESS label and provides a tamper evident binding of the label to the message but does not by itself protect the confidentiality of the message. At the point where you learn the access control policy to enforce on the data you already have access to the data. While the signature provides a tamper evident integrity for the label over the clear text, it is not tamper proof because it is susceptible to unauthorized removal if you only have a SignedData message, i.e. any Message Transport Agent (MTA) in the path can remove a signature layer of a SignedData message therefore altering the access control data. Encrypting the signed message protects the confidentiality of the data and protects the SignedData from tampering from anyone unable to decrypt the message. However encrypting the message means that no intermediate agent can enforce the label policy and it does not protect the label from any entity who has the ability to decrypt the message.

From a regulatory enforcement perspective, ESS labels are an extremely weak form of access control because cryptographic access to the data is given before the access check. The correct enforcement of the access check is dependent on the configuration of every recipient's Email client. Since the cryptographic access is granted before the access checks, there is no cryptographic impediment for a recipient who is able to decrypt the data but unauthorized under the policy, to ignore the policy and access the data. A stronger enforcement model is needed for regulatory control for Email where cryptographic access is only granted after the access check is successful.

1.2 Encrypted E-Mail Using Web-based Credentials

There are many users on the Internet today who have forms of authentication credential other than X.509 certificates. S/MIME today can only use X.509 certificates to protect the confidentiality or the data origination authentication of the messages. This means the many users without X.509 certificates cannot use S/MIME. Standard based services (e.g. [SAML-overview])are now available which abstract the specifics of an authentication technology used to identify a subject

from the application itself (S/MIME in this case). Adoption of this abstraction model would enable a broader set of authentication technologies to be able to use S/MIME to secure Email for confidentiality or data origination authentication. It also allows for new authentication technology to be deployed without impacting the core protocol.

1.3 Vocabulary

Some of these terms are the same as used by RFC3198. While the Plasma actors are fundamentally the same as rfc3198, there are some minor differences in the responsibilities of each actor with models such as XACML[XACML-core].

Attribute Based Access Control (ABAC)	Where the policy is specified by the set of attributes, their values and any relationship between attributes required to authorize an action on a resource. These attributes may be provided by the subject as part of the decision request (Front End Attribute Exchange) or discovered by the policy decision service itself (Back End Attribute Exchange). The policy for example may require attributes about the subject, their device or environment, a resource or the intended use of the information.
Back End Attribute Exchange (BEE)	When attributes are directly sent from the attribute issuer to the PDEP.
Capability Based Access Control (CBAC)	Where access control is via a communicable, unforgeable token. A capability token is a protected object which, by virtue of its possession by a subject, grants that subject the capability.
Cipher text	Plain text which has been processed by an encryption algorithm to render it unreadable by a program or human without the appropriate cryptographic key.
Confidential	A message has been protected to a known level of confidence so that the contents are not decipherable by unauthorized users.
Content Encryption Key (CEK)	A key used to encrypt protected end user data. (See Key Encryption Key)
Cryptographic Lock Box	A data structure which holds a CEK encrypted

	for a specific user. CMS implements Cryptographic Lock Boxes as RecipientInfo structures.
Data Origination Authentication	Enables the recipient to verify that data or messages have not been tampered with in transit and that the originator is the expected sender.
Decision Requester (DR)	The service responsible for making policy decision requests to the PDEP. In this model the policy decision is enforced by the PDEP by its control of cryptographic keys. The DR enforces any obligations the PDEP may require such as signing or encryption of the data, generating audit events etc. An DR is distinct from a Policy Enforcement Point in other models such as XACML in that an DR is not by default trusted with the clear text data. Policy enforcement is performed by the PDEP. An DR may establish trust by presentation of attributes about itself and its environment to show it is trustworthy.
Early Binding	The concept of creating the cryptographic lock box for a recipient at the time the message is sent. (See Late Binding).
Front End Attribute Exchange (FEE)	When subject attributes are relayed from the attribute issuer to the PDEP party via the Plasma client.
Integrity Protected	A recipient of a message can determine to a known level of confidence that a message has not been modified between the time that it was created and it was received by the recipient.
Key Encryption Key (KEK)	A key used to encrypt another cryptographic key, often a CEK. (See Content Encryption Key)
Late Binding	The concept of creating the cryptographic lock box for a recipient when the recipient attempts to decrypt the message. Late binding has a potential downside because the sender cannot know what symmetric algorithms the recipient supports which can lead to interoperability issues. (See Early Binding)

Mail Transfer Agent (MTA)	A program that transfers email from one computer to another. An MTA implements both the sending and receiving of email.
Mail User Agent (MUA)	A program or service used to manage a user's email. The MUA may be a program run on the users computer or a Web service accessed via the users browser.
Orthonym	The correct or legal name of a place or person or thing. (See Pseudonym)
Plain text	The information in a state which can be directly read by a human or an appropriate application.
Policy	(1) A statement in a human language which defines a course of action by an individual or organization. These statements may be in the form of legislation, regulation, a legal contract or organization goals. (2) Technical controls for implementation of the human readable policies. Policies may stipulate many forms of technical controls requirements such as data protection, access control, data integrity, data origination, data retention, etc.
Policy Administration Point (PAP)	The system entity that creates, maintains and publishes policies or policy collections. The policies define the rules, their conditions and actions associated with the policy.
Policy Collection	A collection of one or more policies which is associated with a role. The policy collection may also defines the logical relationship between the policies.
Policy Decision and Enforcement Point (PDEP)	The system entity that evaluates the policy criteria published by a PAP, using attributes supplied by a PIP to render decisions from request made by DRs.
Policy Identifier	The tag that is used to identify a policy. For the purposes of our document we are focusing on two different types of policy identifiers. Object Identifiers (OIDs) are what are currently used in many security policy systems

	and are the only method of policy identification supported by ESS security labels. Additionally we will support URIs as policy identifiers as they provide a more user friendly method of uniquely identify a policy and allow discovery of the policy.
Policy Information Point (PIP)	A service with issues assertions for example about a subject, their device or environment e.g. a SAML Security Token Service. This model supports both front end and back end exchange of assertions between the PIP and the PDEP. Attributes can be distributed directly between the PIP and the PDEP (Backbend Attribute Exchange;BAE). Alternatively attributes may be distributed via the DR (Front End Attribute Exchange; FAE) There are two types of PIP based on the types of attribute the PIP would assert about the subject. An Identity Provider (IdP) PIP will issue authentication attributes e.g. information about how and when the subject authenticated to the IdP. An IdP may also issue attributes about the subject themselves e.g. their full name, age or citizenship. An attribute provider (AtP)PIP only issues attributes about the subject or the subject's environment.
Policy Label	The data structure which holds one or more policy identifiers and their logical relationship.
Pseudonym	A name that a person or group assumes for a particular purpose, which differs from their original or true name. (see Orthonym)
Role	An abstract subject which has a series of authorizations assigned to it. Users are assigned to roles to perform the duties of the role. Users typically select a role to perform a function to disambiguate which role they are currently functioning as. A role is distinct from a group because a group is a collection of subjects which has no intrinsic authorizations.
Role Based Access Control (RBAC)	Access control based on the assignment of a role. Subjects are then allowed to assume one or more roles based on their job needs as

for as long as their job requires.

We additionally make use of the following terms:

Attribute Request/Issuance	The act of a client requesting and obtaining a set of attributes for a subject. The issuance of attributes will itself be controlled by policy and thus recursively embeds this same picture in that process. The attributes can be requested either by the AR (front end attribute exchange) or the PDEP (back end attribute exchange).
Content Protection Request/Response	The protocol to be run by the DR to get the set of decisions and information required to successfully create and encode a data block with appropriate labeling. This protocol is part of the work to be defined by this group.
Content Consumption Request/Response	The protocol run by a DR to obtain the permissions and information needed to decode and access data with appropriate labeling.
Content Distribution	Can be any of a number of methods by which the content is transmitted from the Content Creator to the Content Consumer. These methods include, but are limited to: Email, FTP, XMPP, HTTP and SneakerNet.

1.4 Keywords

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119.

2 Background

The S/MIME standard [RFC5751] provides a method to send and receive secure MIME messages. S/MIME uses CMS[RFC5652] as the means to protect the message. While CMS allows for many types of security credentials to be used, S/MIME exclusively [RFC5750] uses X.509 certificates [RFC5280] for the security credentials for signing and encryption operations. S/MIME uses an early binding mechanism for encryption keys where the sender needs to discover the public key for each recipient of an encrypted message before it can be sent. This requires the sender to maintain a cache of all potential recipient certificates (e.g. in a personal address book) and/or have the

ability to find an acceptable certificate for every recipient from a repository at message creation. This key management model has limited the use of S/MIME for encryption for a variety of reasons. For example:

- The recipient may not have an X.509 encryption certificate
- The sender may not have received a signed Email with the recipient certificate
- The recipient may not have an available repository
- The sender may be unaware of the location of the recipient's repository
- The recipient's repository may not be accessible to the sender e.g. it's behind a firewall
- The sender may not have a valid certificate path to a trust anchor for the recipients certificate

If one or more recipient certificates are missing, then the sender is left with a stark choice: send the message unencrypted or remove the recipients without certificates from the message.

The use of secure mailing lists has the ability to provide some relief to the problem. The original sender does not need to know the appropriate encryption information for all of the recipients of the mailing list, just for the mailing list itself. It can thus be thought of as a form of late-binding of recipient information for the originating sender. However it is still early-binding encryption for the mail list agent; as it needs to perform all of the gathering and processing of certificate information for every recipient that the agent will relay the message to.

In many regulated environments end-to-end confidentiality between sender and recipients by itself is not enough. The regulatory policy requires some form of access control check before access to the data is granted. In many inter-organization collaboration scenarios it's impossible for the sender to satisfy the access checks on behalf of all recipients since they don't have, and frequently should not have access to, all the recipient's attributes because to do so may be a breach of the recipients privacy. Indeed to release the attributes to the sender may require that the sender's attributes first be released to the recipient's attributes provider. It's a fundamental tenet of good security practice that users should control the release of data about themselves.

2.1. ESS Security Labels

Security labels are an optional security service for S/MIME defined in Enhanced Security Services for S/MIME [RFC5035]. The ESS security label allows classification of the sensitivity of the message contents using a hierarchical taxonomy in terms of the impact of unauthorized disclosure of the information [RFC3114]. The security label can also indicate access control such as full time employees only or US nationals only. ESS security labels are authenticated attributes of a CMS signer-info structure in a SignedData object. The label when applied to signed clear text data provides the access control decisions for the plain text. If applied to cipher text such as the outer layer of a triple wrapped S/MIME message the label is used for coarse grained optimization such as routing.

2.1.1. Problems With ESS Security Labels

ESS Security Labels have been found to have a number of limitations.

1. When the label is on the innermost content, access to the plain text is provided to the recipient (in some form) independent of the label evaluation as it will be processed for the purpose of hash computation as part of signature validation. Depending on how a triple wrapped message is processed by the recipient's CMS code, the inner content may be processed for signature validation even before the outer signature is validated. This would happen for a stream based CMS processor which starts processing inner-layers immediately rather than finishing processing of each layer and caching the intermediate results.

2. Labels applied can be removed in transit. If a signed layer is seen then it can be removed by any agent that processes the message (such as a Message Transit Agent). If the label is protected by an encryption layer then it can only be removed by any agent that has a decryption key (Encryption Mail List agents or Spam Filtering software would be two such examples).

3. Policies are identified by Object Identifiers. This makes for a small tight encoding, but it does not provide any mechanism for an Email client to discover how to enforce a new access control policy if the message contains a policy the client is unaware of. This provides a stark choice: ignore the access control policy and grant access to the message or block access to the message. Object identifiers also do not provide a good display name for a user so that they could manually find and download a new policy.

4. The current ESS standard only allows for a single policy label in a message, no standardized method of composing multiple policy

labels together has been defined. This is adequate for coarse grained policy binding to express a limited set of choices such as with information sensitivity which typically provides a hierarchy of 3-5 choices. Many data sets need to be subject to multiple access control policies. For instance, a message may contain information that is both propriety and export controlled. Trying to represent combinations of policies via a single policy label would lead to an exponential growth in the number of policy labels.

5. ESS Labels do not provide for any auditing of who has been granted accessed the message. All policy evaluation is local to the recipient's machine, no centralized logging of access to the message can be performed

6. Enforcement of the policy occurs on the recipient's machine, the compliance with the policy is dependent on the state of the configuration of every receiving agent. The policy is enforce by whatever module is located on the user's system. For cross corporate systems, this means that the policy provided by Company A must be installed on Company B machines, or Company B must install a policy that Company A will accept as being equivalent to their own policy enforcement module. Additionally any time that a new version of the policy module is rolled out; there will be a time lag before every recipients machine will have the updated module. This makes policy compliance practically impossible in anything but a small closed environment.

7. Access to the message cannot be granted or removed after the message has been sent. Therefore if a recipient has a designated alternate recipient they will not be able to read the message. Also if the sender subsequently learns one of the recipients was in error, they cannot correct the mistake.

2.2. Access Control and the Web

A prerequisite for many web transactions is the disclosure of attributes about the subject such as name, age, Email address, physical location, address, credit card number, social security number etc. Some attributes lend themselves to easy verification but many do not. An assertion of an Email address can be verified by sending an Email to the address containing a secret ephemeral challenge. Subsequent demonstration of knowledge of the ephemeral challenge verifies the Email address assertion. Other assertions such as "this is my credit card account number" are not easily verified. The fact that it is a valid credit card number can be verified but not the binding to the subject attempting to use it. Where a claim is not easily verified it is often combined with other

assertions under the assumption that knowledge of this larger data set verifies all the assertions in the data set. If you know the account number, billing address, etc., 'of course' you must be the account holder. This is a very weak form of verification as is often demonstrated by the growth of identity theft; much of this bigger data set is often publicly available via social networks or easily guessed e.g. the most popular professions for a parent is dead or retired. Many of the assertions which are harder to verify are based on government issued documents such as a birth certificates, driver's license, identity card or passport. This requires an exchange of the documents between the relying party and the subject. For a small number of high value transactions (e.g. opening a new account) with relying parties that have widespread physical presence (e.g. a bank or Post Office) this is acceptable because the applicant can present themselves with the required documentation in person. However,web based relying parties cannot perform an in person exchange of documents to verify information on government issued documents. The approach taken with such relying parties is to have trusted assertion providers where the assertion provider can perform an in person exchange of documents with the subject then vouch for the set of assertions they have verified.

SAML [SAML-core] defines an XML framework for describing and exchanging attributes about subjects. The entity making the assertions about the subject is known as the assertion provider, the entity consuming the assertions is known as the relying party. The well-known scenarios for using SAML are:

- Single Sign On across systems on different platform technology
- Federated Identity between business partners
- Web Services and other standards e.g. SOAP based protocols

The critical difference between SAML and pure authentication protocols such as mutually authenticated TLS is that SAML is able to exchange the rich and variable set of assertions which are necessary for authorizing transactions. X.509 certificates can exchange a limited and fixed set of identity assertions such as an x.500 distinguished name, Email address, Kerberos principal name, etc. SAML is able to do this in addition to an extensible set of other assertions about the subject such as: date of birth, business sign-off limits levels, etc. SAML additionally defines a number of query/response style profiles [SAML-QUERY] that allow for a relying party to specify the type of attributes that are required to evaluate a policy. It is a matter of local policy on the SAML identity provider what attributes to release about the subject to the relying

party.

SAML also abstracts the details of the authentication protocol from the relying party. The assertion provider can use a broad range of authentication mechanisms such as passwords, one time passwords, biometrics, X.509 certificates, etc., without impacting the relying party. The assertion provider can include the details of the authentication mechanism or its strength using an established strength of authentication scale such as NIST SP800-63-1 [SP800-63-1]. The relying party can then inspect the claims about how or how strongly the subject authenticated to the identity provider to determine if it complies with its access policy. Low value transactions can use simple short lived assertions where possession of the assertion alone is considered acceptable for the transaction risk. These are known as Bearer assertions. Higher value transactions can require proof of possession keys (either symmetric or asymmetric cryptographic keys) where the subject demonstrates knowledge of a cryptographic secret to the relying party via a MAC or digital signature. These are defined by the SAML specification as Holder of Key assertions. The subject has to demonstrate possession of the key to the relying party. Holder of key assertions can be either symmetric or asymmetric keys.

2.3. Information Asset Protection

Information Asset Protection (IAP) is a concept developed by the Transglobal Secure Collaboration Program (TSCP), a working group comprised of the major players in the western Aerospace and Defense industry. The industry is highly regulated and operates in an environment with many policies governing the access to information assets. These policies are motivated by the desire to protect intellectual property, the confidentiality of information, or are imposed by government regulators such as the US International Traffic in Arms Regulations (ITAR) from the US Department of State. They apply to the information assets in whatever form the asset may take and are independent of the application used to create the information. These policies take many forms, e.g. verification the recipient has demonstrated a need to know the information because they are working on a specific project, that they have passed the appropriate background and nationality checks, or that they have signed the appropriate non-disclosure agreement. What is needed is a policy driven information centric protection where the applicable policies either is manually or automatically attached to the information and based on the policy the system understands what access control and data protection is necessary.

Email is an application widely used in the Aerospace and Defense industry. S/MIME is widely used today and provides sender to

recipient confidentiality. This protects the contents of the message from discloser to unauthorized third parties e.g. while it is in transit between MTA's or while at rest in a MTA message queue or recipient's mailbox. However it does not impose any finer grained access control such as those required by many policies. S/MIME does define an extension mechanism for access control via an ESS security label [RFC5035] though this mechanism has drawbacks (see above).

2.4. Authentication Assurance Frameworks

A number of organizations have created taxonomies to define the possible levels of identity assurance for electronic authentication. The objective of the framework is to provide a simple abstraction the details of

- Identity proofing and registration of subjects
- Tokens used by subjects for providing electronic identity
- The token management mechanisms
- Protocols used for subject to use tokens to authenticate to an identity provider
- Protocols used by subjects to authenticate and pass attributes to a relying party

These frameworks have been drafted by industry organizations [lib-iaf][kan-iaf] and governments [SP800-63-1]. While all of these frameworks may not agree on every aspect, at a macro level they do exhibit many similarities. A common theme in many is the adoption of a small number of levels of identity assurance, typically between 3-5. A simplified description of the levels is:

Level 1 Negligible confidence in the asserted identity

Level 2 Some confidence in the asserted identity

Level 3 Significant confidence in the asserted identity

Level 4 High confidence in the asserted identity

The framework defines broad characteristics in the area of identity proofing, credential type and management, identity provider authentication and relying party authentication.

2.5 Electronic Signatures: Authentication vs. Authorization

Electronic signatures on email are used today to show data origination so only authentication is required. However with transactions that are legally or regulatory significant, authentication alone if frequently insufficient. Policy requires other factors to be considered to ensure the transaction meets policy requirements.

- The state of the system generating the signature
- An indication of the signers intent
- Attributes about the signer to indicate for example, job function in the company, job assignments professional qualifications, signing authority etc.

Many organizations would like email based work flows to be an option for these transactions.

3. Use Case Scenarios

This section documents some email based use case scenarios the new protocol aims to support. Also included are some related scenarios where the same underlying theme of consistent policy enforcement equally applies.

3.1 Consumer to Consumer Secure Email

One of the issues that is stopping the use of secure Email in personal mail is the fact that consumers find X.509 certificates difficult to obtain and then use - especially across a set of devices (phone, tablet, workstation). One of the possible use cases of Plasma is to try and deal with this by removing the dependency on X.509 certificates. The details of the use case are therefore: Alice wants to send an Email message to Bob that contains sensitive, personal data so she is concerted to ensure only Bob can read it. Bob has a strong credential he can use to identity himself, but is is not an X.509 certificate. Alice needs to ensure the following:

(a) Only Bob can read the Email.

(b) Bob has the ability to verify the Email is from Alice.

(c) Bob has the ability to verify the Email message has not been modified since Alice sent it.

The sequence of events could be as follows:

1. Alice composes the Email to Bob.

2. Alice's Email client allows here to classify the Email. Alice classifies the Email using Personal Communication which is a basic policy provided by her ISP.

3. Alice's Email client knows the protections to apply to a Personal communication; it knows to encrypt and sign the message.

4. The protected Email is able to flow securely and seamlessly through existing Email infrastructure to Bob. The data is protected while in transit or at rest.

5. Bob receives the Email and sees that it is a secure message. Bob can verify that the secure message has not been altered. Bob attempts to open and decrypt the Email. If Bob is on the same ISP as Alice, then the same username/password as he uses to get his Email to obtain the needed keys. If Bob is on an ISP that is federated with Alice's ISP then an infrastructure such as SAML, OpenID, OAUTH or ABFAB could be used to validate Bob's identity and allow the needed decryption keys to be released.

3.2. Business to Consumer Secure Email

There are many examples of business to consumer secure Email scenarios where the Email could potentially contain sensitive medical or financial data. This would include doctor, patient; bank, account holder; Medical insurance, insured person; mortgage broker, customer. Two examples are presented here.

3.2.1 Bank Statement Email

A bank (The Bank of Foo) has determined that it will be using Email to distribute statements to its customers (Bob). The information is confidential, so any channel of communication the Bank selects must protect Bob's privacy. The bank needs to ensure the following:

(a) Only Bob (or additional owners of the account) can read the Email

(b) Bob authenticates with a sufficient level of identity assurance. The same identity assurance authentication level used to do on-line banking would be considered sufficient

(c) Bob can verify the statement is from his bank

(d) Bob can verify the statement has not been modified since his bank sent it.

The sequence of events would be as follows:

1. As part of routine end of the month processing, the Bank composes an Email to Bob. They include the statement of balances and activity either as an attachment or as the body of the message.

2. The statement mailer for the Bank of Foo has been configured to use a specific policy on the Email.

3. The statement mailer for the Bank of Foo knows the protections to apply based on the policy; it knows to encrypt and integrity-protect protect the message and what level of assurance required for the recipients identity

4. The protected email is able to flow securely and seamlessly through existing email infrastructure to Bob. the data is protected while in transit or at rest.

5. Bob receives the email as sees it is a secure message from the Bank of Foo. Bob can verify the message has not been altered as it is signed by the his Bank. Bob uses the same credential as he would for on-line banking to prove his identity to the email system and obtain the keys necessary to decrypt the message.

The same process could be used for any messages sent between the bank and its customers. Thus, messages dealing with loan applications and changes in bank policies can be sent out in the same manner potentially using different policies. In some of these cases it might be in the bank's interests to record in an audit trail if and when the keys were handed out on some Emails. For a statement, the Bank would not expect a reply to occur, however for other types of messages it should be possible for Bob to reply under the same level of protection. If Bob is able to use the same credentials when sending a message, to the one he uses for access the banks web site then the bank has the same assurance of the message sender identity.

3.2.2 Doctor-Patient Communications

In the second example, let's say that Alice is a doctor and has received test results for her patient Bob. This information is confidential and regulated, so any channel of communication she selects must protect Bob's privacy and comply with regulatory requirements. Alice elects to use Email to reach Bob quickly with news of the results. In this respect it is similar to the previous use case; however there are some additional complications that might need to be dealt with as well. Depending on who Bob is and where is currently is there are additional people that may also need to be automatically informed of the same information, or need to have the

ability to access the contents of the message. Examples of these would be Bob's spouse, an individual who is making care decisions for Bob (i.e. Bob's parent), and an individual in charge of dealing with Bob's day-to-day health care (i.e. a charge nurse in a hospital or a visiting nurse). All of these people may have the same need to know as Bob. There is also the possibility that some parts of the message may need to be released to some individuals but not to others. As an example, the mail message could contain a prescription, that specific portion of the message may need to be read by Bob's pharmacist. Alice needs to ensure the following:

(a) Only authorized individuals can read the Email. However, the definition of authorized will vary with the content of the message and thus the policy applied. (General health issues will certainly be treated differently than mental health issues, even by a General Practitioner.)

(b) The Bob is required to authenticate with a identity assurance level 2 or above level.

(c) The Bob can verify the Email is from Alice.

(d) The Bob can verify the Email has not been modified after Alice sent it.

The sequence of events would be as follows:

1. Alice composes the Email to Bob. She includes some comments and suggestions for Bob and attaches the test results.

2. Alice's Email client allows her to classify the Email. Alice classifies the Email as a Doctor-Patient communication. As a side effect of classifying the Email message, the policy may suggest or mandate additional individuals that the communication should be addressed to.

3. Alice's Email client knows the protections to apply to Doctor-Patient communication; it knows to encrypt and integrity-protect the message.

4. The protected Email is able to flow securely and seamlessly through existing Email infrastructure to Bob. The data is protected while in transit or at rest.

5. Bob receives the Email and sees it is a secure message from Alice. Bob can verify the message has not been altered. Bob attempts to opens the Email. Bob provides a Level 2 password to retrieve the necessary decryption keys. After Bob has proved his

identity, he is able to read the Email.

There are number of different places where the identity provider for Bob could live. The first is at Alice's office, Bob already has a face-to-face relationship with Alice and the credential could be setup in her office. A second could be Bob's insurance provider. Bob has a relationship with his insurance provider as does Alice, thus it can serve as an trusted identity provider to healthcare providers. A third location could be a federation of doctors in an area, potentially with other health providers (such as hospitals and convalescent centers), Bob has setup an identity with Alice, but if he gets referred to Charlie by Alice for some procedures, Charlie would not need to setup a new identity for Bob but instead could just refer to Alice for the necessary identity proof. Many of these types of situations are dealt with by [I-D.ietf-abfab-arch].

There are a number of other additional services that could be provided by the policy system. One example would be that if the information was time critical, if Bob does not access his message within a given time period, the policy server could notify Alice of this fact so that an alternate method of communication can be attempted with the same information.

3.3 Business to Business Ad-Hoc Email

Early in the relationship between two companies, it is frequently necessary to exchange sensitive information as a preliminary to a more formal business relationship e.g. contract negotiations. This needs to occur before the relationship has matured to the point that a formal relationship is reflected through a specific legal agreement. Business owners need the agility to interact with potential partners without having to engage their respective IT staffs as a prerequisite of the communication.

As an example, Charlie works for Company Foo. He has just met Dave from Company Bar to discuss the prospect of a potential new business opportunity. Following the meeting, Charlie wants to send Dave some sensitive information relating to the new business opportunity. When Charlie sends the Email to Dave with the sensitive content, he must ensure the following objectives:

(a) Only Dave can read the Email

(b) Dave is required to authenticate with an identity assurance level 2 or above

(c) The Dave can verify the Email is from Charlie

(d) The Dave can verify the Email has not been tampered with

(e) Charlie may also need to keep a record of the fact that Dave accessed the message and when it was done.

The sequence of events Charlie would use is as follows:

1. Charlie composes the Email to Dave. He include some sensitive information relating to potential terms and conditions for the new contract that Foo and Bar would sign to form a partnership for the business opportunity.

2. Charlie's Email client allows him to classify the Email. He classifies the Email as an Ad-hoc pre-contractual communication.

3. Charlie's client knows the protections to apply to Ad-hoc pre-contractual communication; it knows to encrypt and integrity-protect the message and the level of assurance required for the recipients identity.

4. The protected Email is able to flow securely and seamlessly through existing Email infrastructure to the recipients (Dave in this case). The data is protected while in transit or at rest.

5. Dave receives the Email as sees it is a secure message from Charlie. (Charlie requires level 2, Dave uses a password) Dave is able to prove his identity to the level of assurance requested by Charlie so is able to read the Email. The organization Dave work for has an identity service which he uses to prove his identity for Charlie's Email. Dave opens the Email.

If Dave or his delegate replies to the Email from Charlie, the new message inherits the policy from the original messages so the entire message thread has the same policy. The policy also applies to messages forwarded by Dave because it contains information from Charlie and Company Foo wants consistent policy enforcement on its information.

3.4 Business to Business Regulated Email

As business relationships mature they often result in a formal contractual agreement to work together. Contractual agreements would define a number of work areas and deliverables. These deliverables may be subject to multiple corporate and/or regulatory policies for access control, authentication and integrity. Some classes of Email may have information which is legally binding or the sender needs to demonstrate authorization to send some types of message where authority to send the message is derived from their role or function.

Also many regulated environments need to be able to verify the information for an extended period - well beyond the typical lifetime of a users certificate. The set of policies applicable to an Email is potentially subject to change as the different user's contribute information to the Email thread.

3.4.1 Regulated Email Requiring a Confidentiality Policy

Company Foo has been awarded a contract to build some equipment (Program X). The equipment is covered by export control which requires information only be released to authorized recipients under the terms of the export control license. Company Bar is a foreign subcontractor to company Foo working on Program X. Company Foo sets up some business rules for access to program X data to ensure compliance with the export control license requirements. Company Foo also set up separate rules to cover the confidentiality of its intellectual property contributed to Program X. Company Bar also sets up its own policies to protect the confidentiality its own intellectual property it contributes to Program X. As part of the agreement between Foo and Bar, they have agreed to mutually respect each other's policies.

Confidentiality policies can change over time. It is important to be able to implement the changes without the need to update the data itself to reflect the change as finding all instances of the data in an intrinsically impossible problem to solve.

Frank is an employee of Company Foo. He has been assigned as a design team leader on Program X and as an individual contributor on Program X integration. Frank wants to send some mail as a team leader to colleagues working on Program X in both Companies Foo and Bar.

Grace is an employee of Company Bar. She has also been assigned to the design team of Program X.

When Frank sends the Email with Program X regulated content he must ensure compliance with the export control policies. When Frank sends a Program X email he must ensure recipients are authorized to read the contents to ensure Company Foo remains in compliance with its export control license.

If Frank also includes Company Foo intellectual property in an email, he must also ensure recipients are authorized to read the intellectual property contents.

When Grace receives a Program X email, she must provide attributes about herself to prove compliance with the export control policy.

If the email also contains Company Foo intellectual property, she must also provide attributes to show she is authorized to read the information under the agreement between Company Foo and Company Bar.

If Grace sends an email with Company Bar intellectual property, she must ensure recipients are authorized to read the contents under the agreement between Company Bar and Company Foo.

When Frank sends a Program X Email he must ensure the following objectives:

(a) Only recipients who meet the Program X policy and/or Company Foo's intellectual property protection policy can read the Email
(b) Recipients authenticate with a identity assurance level of level 3 or above
(c) Recipients present all other attributes about themselves necessary to verify compliance with the applicable policies (their program assignment, nationality, professional or industry certifications, etc.)

(d) Recipients can verify the Email is from Frank to the level of identity assurance as defined by the message policy (i.e. level 3 or above)

(e) Recipients can verify the Email has not been tampered with the level of identity assurance as defined by the message policy

(f) Recipients are made aware that the message is a Program X Email and the contents can only be shared with other Program X workers and/or the message contains Company Foo's intellectual property

The sequence of events Frank would use is as follows:

(1) Frank composes the Email and includes a Program X distribution list as a recipient. He include some information relation to Program X. Frank also includes some information which is Company Foo's Intellectual Property.
(2) Frank's Email client allows him to select the Program X role. The client then allows Frank to select from a set of policies appropriate for Program X.
(3) Frank selects the Program X content and Company Foo IP policies from the list of available policies.
(4) The Email client knows to encrypt the message, the key size and algorithm to use to use; that the message needs to be signed with a level 3 or above certificate.
(5) Frank clicks the send Email button. The client signs the Email using his smart card and a certificate indicating the signature.

The Client then encrypts the message and obtains data from a server that will enforce the access control requirements for Frank, and sends it to his Email server.

The Email is able to flow securely and seamlessly through existing Email infrastructure recipients of the distribution list. Grace is on the distribution list so receives the Email from Frank.

(6) Grace receives the Email, Grace's client provides the attributes necessary to comply with the policy which includes her level 3 encryption certificate to the PDEP.
(7) Once Grace has shown she passes the policy requirements, the PDEP releases the message CEK to Grace using her level 3 encryption certificate.
(8) Grace uses her smart card to open the message. She sees the message is signed by Frank and marked with both the Program X and Company Foo IP policies

If Grace replies to the Email from Frank, the new message inherits the policy from the original message. If Grace includes some information which is Company Bar's IP she also adds her companies IP protection policy requirements to the message.

Frank receives the reply from Grace. Frank is able to prove his identity to the level requested by Grace and provides the requested attributes about himself to satisfy both the Program X export control, the Company Foo IP protection policies as well as the Company Bar IP protection policies. Frank opens the Email.

The policy also applies to messages forwarded by Frank and Grace because they contain information from Company Foo and Company Bar and both companies wants consistent policy enforcement on their information.

After some time, Company Bar fails an audit to show they are complying which all the requirements for Program X. As a result, Company Foo updates its policies for Program X to remove company Bar as an approved to access Program X data. Grace will no longer be able to access the Program X email as she can no longer satisfy the Program X policy requirements.

3.4.2 Regulated Email Requiring an Integrity Policy

Company Foo has been awarded a contract to build some equipment (Program X). This equipment is regulated by the National Aviation Authority (NAA) for Company Foo. The NAA requires strict procedures at a number of significant events for Program X such as in the design, and maintenance of the Project X (e.g. when a design is

complete and released to manufacturing). The sign of process requires personal be suitability qualified and that the documentation needs to be maintained for the service life of the project (25 years for Program X)

Company Foo has instigated an email based sign off procedure to simplify sign off and reduce costs. It also has authored a policy for compliance with the NAA requirements. At the appropriate time, signoff email is sent to the designated program members. Recipients apply the NAA policy when they reply to the sign off request message.

Frank is the lead on the Program X design team. They have a design which they believe can be released to the integration team. Frank initiates the sign off process for the design.

Grace is one of the sign-off design team members for Program X. She receives the sign off email. Grace responds and applies the sign-off signature policy to the email. The policy requires Grace to authenticate with the required level of assurance, present attributes about herself, her work effort assignments and professional qualifications to demonstrate compliance with the policy to send the message. The message is signed to indicate Grace passed the policy.

When Frank initiates a Program X Sign Off Email the system must ensure the following objectives:

(a) Frank was authenticated to the level of identity assurance under the policy to initiate the sign off process
(b) Frank possessed the necessary attributes as required by policy to initiate the sign off process
(c) The contents of the email are accurate to the level of integrity assurance required by the policy
(d) Frank was fully aware and intended to initiate the sign off process
(e) The state of Franks system was known to the level of assurance required under the policy to be free from agents which might interfere with the sign off process
(f) Recipients can easily confirm over the lifetime as required by the policy that the sign off procss passed the policy without having to know specifics of what the policy entailed.

The sequence of events Grace would use is as follows:

(1) Grace receives the sign off request email.
(2) Grace replies to the email and completes the form data in the email to show she is approving the sign-off

(3) Grace clicks the send button to send the email
(4) Grace receives a sign-off confirmation dialogue before the email is sent where she is able to confirm her intent is to approve the sign off of the component.

Grace's system submits the decision request to send the sign-off email. Her system is asked to provide data about Grace and the state of her system, and the data being authenticated. If Grace's request passes the policy, her system receives a signed statement the message passes the policy which is attached to the email and the message sent.

3.5 Delegation of Access to Email

There are a number of times when others are given access to a recipient's mailbox or Email is forwarded to other recipients based on recipient's rules. This may be a long standing relationship such as when an assistant is given access to an executives mailbox. Alternatively it may be a temporary relationship due to short term needs (e.g. to cover for a vacation). There are also organizational role mailboxes where the recipient is a role and one or more users are assigned to the role.

Grace is going on vacation. While Grace is away, Brian will act as a delegate for Grace. Grace configures a mailbox rule to forward Program X Email to Brian for the duration of her vacation. Brian is able to satisfy the policy requirements for the Program X Email as outlined above and is therefore able to open the protected Email sent to Grace. Frank does not need to take any actions to allow Brian to access the Email.

3.6 Regulated Industry Email

Some organizations work in areas which are intrinsically subject to policy such as regulatory policy e.g. healthcare. In such environments the policies are often tied to the roles of the participants, the institution they are working at and the subject of the exchange.

Hanna is a primary care physician working for FooBar Healthcare. She has a patient which she is referring to a specialist Ida for further investigations. Ida works at the Bar Hospital. Hanna needs to send the relevant patient notes, test results and comments to Ida. Hanna knows she needs to comply with the confidentiality regulations and needs to respect her patients consent decree for the privacy of their Healthcare information. When Hanna sends the referral message she must ensure:

(a) Only recipients who meet the healthcare regulatory policy and the patients consent decree can read the Email
(b) The message has the appropriate level of integrity and data origination as required by the policies
(c) The recipients authenticate with an acceptable level of assurance (i.e. level 3 or above)
(d) Recipients present attributes about themselves necessary to verify compliance with the policies (e.g. their professional qualification, professional registration, affiliated healthcare facility and department)
(e) Recipients can verify the Email is from the sender (Hanna) to the level of assurance as defined by the message policy (i.e. level 3 or above)
(f) Recipients can verify the Email has not been tampered with the level of assurance as defined by the message policy
(g) Recipients are made aware that the message is a Patient referral and contains sensitive patient data.

The sequence of events Hanna would use is as follows:

(1) Hanna composes the Email and includes Ida as a recipient. She includes the patient information, test results and comments in the Email
(2) Hanna's Email client allows her to select a policy which is appropriate for her work.
(3) Hanna selects the Patient Referral and Patient Consent decree policies from the list of available policies.

The Email client knows the protection to apply to the Email; to encrypt and integrity-protect the message, the level of assurance required for the recipient's identity and what recipient attributes are necessary to access the message.

(4) Hanna clicks the send Email button. The client signs the Email using Hanna smart card. The client then encrypts the message and sends it to the Email server.

The Email is able to flow securely and seamlessly through existing Email infrastructure to recipients of the distribution list. Ida is on the distribution list so receives the Email from Hanna.

(5) Ida receives the Email as sees it is a secure message from Hanna. Ida's client provides the attributes necessary to comply with the policy which includes her level 3 encryption certificate to the PDEP.
(6) Once Ida has shown she passes the policy requirements, the PDEP releases the message CEK to Ida using her level 3 encryption certificate.

(7) Ida uses her smart card to open the message. She sees the message is marked with both the Patient Referral and Patient Consent Data policies

3.7 Email Compliance Verification

Verification is an essential part of compliance. Verification may be conducted by internal staff or external auditors. The verification need to confirm that the policy rules are being enforced. Auditing relies on the generation of artifacts to capture information about events. Typically this is does via some form of logging. A challenge here is that for distributed system the set of logs which completely describes the transaction are scattered across many systems so consistency of the audit settings and correlating all the audit data is problematic. Another consideration is accurately capturing only the set of desired data i.e. accurately targeting the set of events that needs to be logged

Jerry is the compliance officer for Company Foo. He has a procedure for ensuring compliance for Program X. The procedure defines what to log and when to audit access to Program X data. Jerry has tools to collect the audit data and run analysis to verify the polices are being followed.

The sequence of events Jerry would use is as follows:

(1) Jerry configures an audit obligation for access to Program X data. The obligation defines the set of attributes to capture when Program X data is accessed. The obligation is part of the Program X policy. Part of the Program X policy is the set of PDEPs which can process policy decisions on Program X data.

(2) Jerry configures his audit log collection to download Program X audit log entries from the designated PDEPs.

(3) Jerry also has an audit confirmation tool which pings the PDEPs for access to Program X data. Jerry's audit log analysis tool looks for these pings to confirm that auditing is taking place as expected.

3.8 Email Pipeline Inspection

Organizations have a huge incentive to inspect emails entering or leaving the organization. This is desired for many different reasons. Inspection of mail leaving an organization is targeted towards making sure that it does not leak confidential information. It also behooves them to check that they are not a source of malicious content or spam. Inbound mail is checked primarily for malicious content, phishing attempts as well as spam. For domains

with a high volume of messages there is a strong need to process email with minimal overhead. Such domains may mandate that they be pre-authorized to process an email due to the overhead a per-message call to an external service would add to message processing.

Company Foo has a policy to scan all inbound and outbound email to ensure it is free from malware. Company Foo also want to ensure email is not spam. Company Foo can own their scanning servers or it may be outsourced to a third party service. Company Foo wants to ensure that its policy of scanning also applies to encrypted email.

The ability to decrypt and check the content for malicious content is highly desirable. There are a number of methods that can accomplish this:

1. When a Company Foo client requests to send a Plasma email, the PDEP is able to check to see if the policy allows email content inspection by MTA for this policy, and if it does, that Company Foo has an outbound email scanning, and that the scanning servers meet the policy requirements. It is able to pre-authorize the Company Foo email scanning servers to access the email.

2. The scanning MTA authenticates to the PDEP as an entity doing virus and malware scanning on a protected message. If the PDEP has specific policy that allows for access to such a scanning service, the appropriate decryption keys will be released and the server will scan the mail and take appropriate action.

3. The policy server is configured with information about various gateways (both internal and external) and has certificates for the known gateways. The policy server can then return a normal X.509 recipient info structure (cryptographic lockbox) to the sender of the message for direct inclusion in the recipient info list of the message. This allows normal S/MIME processing by the scanning software without the necessity to stop and query the PDEP server for keys for specific messages.

4. If the scanning server cannot gain access to the decrypted content using one of the two proceeding methods, it either passes the encrypted mail on to the recipient(s) without scanning it or it rejects the mail. This decision is based on local policy. If the message is passed to the recipient, then the necessary scanning either will not be done or needs to be done on the client's system after the message has been decrypted.

3.9 Distribution List Expansion

A distribution list (DL) is a function of an MTA that allows a user to send an email to a group of recipients without having to address all the recipients individually. Membership of the DL may be confidential so the sender may not know all the recipients. The DL may be maintained by an external organization. Since a DL is identified by an email address, the user may be unaware they are sending to a DL.

Plasma polices may have the list of recipients as a parameter therefor the fact the message is being process by the distribution list means the MTA processing the message needs to update the policy to allow the new recipients to access the message. Organizations may also require inbound scanning of email and have therefore published keys to enable pre-authentication of the MTA by the sender to expedite processing. For both scenarios the DL MTA has to notify the Plasma server that it is adding recipients to the message and supply the list of new recipients. The Plasma server can then take appropriate action on the message token and return an updated token if required.

3.10 Scalable Decision Making

Collaboration involves working with partners and suppliers. These collaborations may be short or long lived, with small or very large number of participants. Organizations therefore need flexibility in deployment and scaling. Organizations do not want to be locked into having to provide capacity themselves. Senders would be happy to delegate decisions to partners providing those decisions use the sets of rules they define for their data. Likewise partners would be happy to leverage their local decision capacity providing they don't have to duplicate rules themselves, and can simply and easily use policies published by their partners. Also organization may want to use cloud based decision services as a cost effective way to add capacity and to be able to respond to transient capacity fluctuations.

Company Foo has been awarded a contract to build some equipment (Program X). The equipment is covered by export control which requires information only be released to authorized recipients under the terms of the export control license. Company Bar is a foreign subcontractor to company Foo working on Program X. Company Foo sets up some business rules for access to program X data to ensure compliance with the export control license requirements. Company Foo also set up separate rules to cover the confidentiality of its intellectual property contributed to Program X. Company Bar also sets up its own policies to protect the confidentiality its own intellectual property it contributes to Program X. As part of the agreement between Foo and Bar, they have agreed to mutually respect each other's policies.

The Program Managers for Program X at Companies Foo and Bar agree a series of roles which are used to manage personal and their assigned policy groups. The policy administrators for Company Foo and Bar respectively publishes the roles and a policy collection for each role. There are rules associated with the policy collection, for example every roles uses the Program X policies published by Company Foo. Employees from Company Foo also get the company Foo Intellectual Property polices for that roles, whereas employees from company Bar get the company Bar intellectual property polices for Program X. Company Foo has also decided to allow enforcement of Program X policies by decision engines in both Company Foo and Company Bar. Company Foo has also decided to use a cloud decision engine for Program X to allow lower cost capacity and scaling. Company Foo is able to add new instances of the cloud decision services as the program scales up and more uses start working on the program. Each decision engine dynamically discovers the policies it needs from the set published by Company Foo and Company Bar. Both company Foo and Company Bar can add new polices to the policy collections at any time and they are dynamically discovered by all the policy decision engines

3.11 Related scenarios

There are other scenarios which are related to the Email cases because they would be subject to the same policy requirements. Email allows users to create content and transport it to a set of recipients. You can perform similar actions with other formats such as documents and instant messages. Policy is agnostic to the underlying technology therefore if an organization has a policy relating to a type of information, then that policy would apply to the same content in an Email, a document an instant message, etc.

3.11.1. Document Protection

This scenario is very similar to 3.4 and 3.6 above. The difference is that the information being generated is in the form of a document not an Email. It could be as part of an ad-hoc sharing or a regulated sharing or information.

Frank is an employee of Company Foo. He has been assigned to Program X. Grace is an employee of Company Bar. She has been assigned to Program X. Frank creates a document for the program. He also includes some Company Foo IP in the document. When Frank creates the document he must ensure compliance with export control regulations and his corporate IP protection policies. Frank must ensure:

1. Only users who meet the Program X policy or Company Foo's intellectual property protection policy can open the document

2. Users authenticates with an acceptable level of assurance as defined by the set of policies applied to the document

3. Users present any other attributes about themselves necessary to verify compliance with the applicable policies.

4. Users can verify who the author was to an acceptable level of assurance as defined by the document policy

5. Users can verify the document has not been tampered with to an acceptable level of assurance as defined by the document policy

6. They can also tell it is a Program X document and the contents can only be shared with other Program X workers.

Frank creates a document for Program X. He include some information relation to Program X. Frank also includes some information which is Company Foo's IP.

Franks word processor client allows him to classify the document. Frank classifies the document as Program X and Company Foo proprietary information.

The word processor client knows the protection to apply to the document; to encrypt and integrity-protect the document, the level of assurance required for the users identity and what user attributes are necessary to access the document.

The document is able to be published on a cloud based Web portal. The document is protected while in transit to the portal or at rest on the portal. The document is also protected on any backup or replica of the portal data. Frank does not to worry about where on the portal he publishes the document. He can make the most appropriate choose based on the project and the document content.

Grace sees the document on the portal and tries to open the document. Grace is able to prove her identity to the level requested by Frank and provides the requested attributes about herself to satisfy both the Program X export control and the Company Foo IP protection policies. Grace opens the document.

If Grace edits the document and includes some information which is Company Bar's IP so adds her companies IP protection policy requirements to the document. Grace saves the updated document to the same location on the portal.

Frank sees that Grace has updated the document on the portal. Frank is able to prove his identity to the level requested by both the

Company Foo and Company Bar policies and provides the requested attributes about himself to satisfy both the Program X export control, the Company Foo IP protection policies as well as the Company Bar IP protection policies. Frank opens the document.

3.11.2 Instant Message Protection

This scenario is very similar to 3.4 and 3.6 above. The difference is that the information being generated is in the form of an instant message not an Email. It could be as part of an ad-hoc sharing or a regulated sharing of information.

Frank is an employee of Company Foo. He has been assigned to Program X. Grace and Hank are employees of Company Bar and also has been assigned to Program X. Frank want to discuss an urgent topic with Grace and Hank. The topic necessitates discussion of Company Foo IP. Because of the urgency, Frank wants to use IM. Frank must ensure:

(a) Only users who meet the Program X policy or Company Foo's intellectual property protection policy can join the IM session

(b) Users authenticates with an acceptable level of assurance as defined by the set of policies applied to the IM session

(c) Users present any other attributes about themselves necessary to verify compliance with the applicable policies.

(d) Users can verify who IM initiator was to an acceptable level of assurance as defined by the session policy

(e) Users can verify the IM data has not been tampered with to an acceptable level of assurance as defined by the session policy

(f) They can also tell the session is a Program X session and the contents can only be shared with other Program X workers.

The sequence of events Frank would use is as follows:

(1) Frank initiate the IM session and includes Grace as a participant.
(2) Frank's IM client allows him to select a role a role which is appropriate for the session. Frank then selects a Program X and Company Foo IP policies for the session.

The IM client knows the protection to apply to the IM session; to end to end encrypt and integrity-protect the session, the level of assurance required for participant's identity and what participant's attributes are necessary to join the session.

The IM is able to flow securely and seamlessly through existing IM infrastructure to session participants. Grace a session participant so her client attempts to join the IM session with Frank. Hank is in a meeting so does not join the IM session at that time.

(5) Grace receives the IM and sees it is a secure IM from Frank. Grace's client provides the attributes necessary to comply with the policy which includes her level 3 encryption certificate to the PDEP.
(6) Once Grace has shown she passes the policy requirements, the PDEP releases the IM session CEK to Grace using her level 3 encryption certificate.
(7) Grace uses her smart card to open the IM session. She sees the from Frank is marked with both the Program X and Company Foo IP policies

(8) Grace composes a response to Frank's question and hits send

(9) When Hank's meeting is finished, he joins the IM session because he to passes the policy requirements and sees the messages from Frank and Grace.

4. Plasma Data Centric Security Model

A common theme from these scenarios is the need to closely tie the information asset to the set of technical controls via the data owners policies in such a way as it is possible to consistently apply the technical controls across a broad set of applications (not just email); for a broad set of users; (not just those within an organization) and in a broad set of environments. Assumptions based on closed world enterprise security models are increasingly breaking down. Perimeter security continues to diminish in relevance, and focus need to be shifted to self protecting data as opposed to protecting the machines that store such data. The binding between the data and the applicable polices needs to happen as close to the data creation time as possible so ad-hoc trust decisions are not required.

The delivery of the documented use cases will require the integration of many existing and some new protocols. In order to ensure the right overall direction for Plasma as each part of the work proceeds, a high level data model is documented here to act as a guide. While this is technically informative to the developments of each individual component, it is normative to the work overall.

This Data Centric Security model is based on a well established set of actors for policy enforcement used elsewhere [RFC3198] [XACML-core].

Figure 1 shows the relationship between the actors.

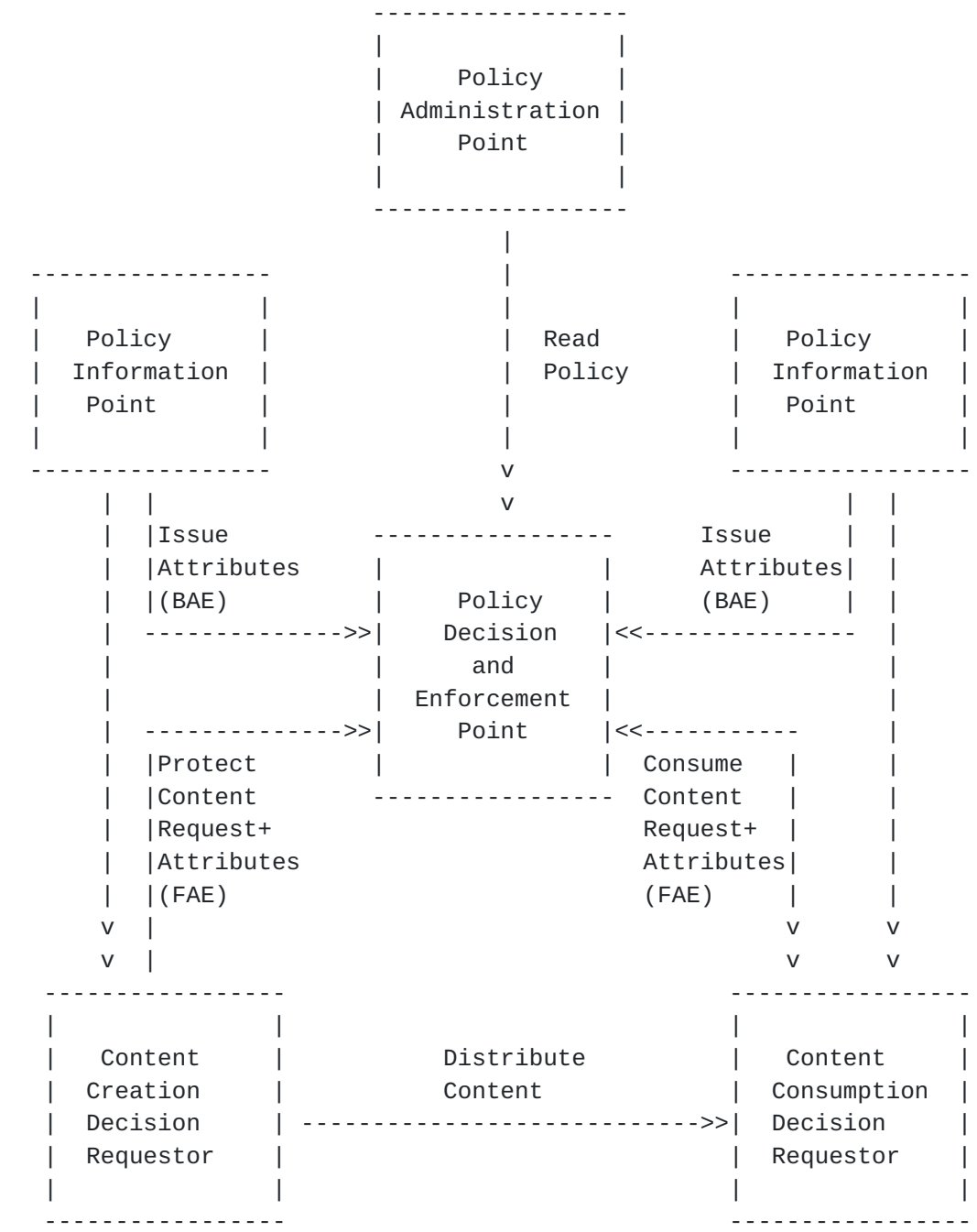


Figure 1 General Scheme for Publishing and Consuming Protected Content

The Plasma model is applicable to any type data (Email, documents, databases, IM, VoIP etc). This is to facilitate consistent policy enforcement for data across multiple applications. Another objective is to not require the data holder to have access to the plain text data in order to be able to make decision requests to the PDEP. The policy decision is complex so the content creation DR in Plasma just

uses policy pointers or labels to indicate the set of policies applicable to content. The content consuming DR dynamically discovers the PDEP's who are authoritative for the decisions on protected content in question. The PDEP's dynamically discover the specifics of a policy from a PAP using the policy references. The specifics of policy authoring, policy decision logic modules are matters beyond the scope of this document. It is important to note that the actors in this model are logical entities and as such can be combined physically in different configurations.

- The Plasma model uses references to bind the data and the policy. When information is created, it is encrypted and a list of policies that must be enforce by the PDEP is bound to the protected data.

- The Plasma model is an Attribute-Based Access Control (ABAC) model where the ABAC policy is specified in terms of a set of attributes, their values and relationships. The policy may specify attributes about the subject, their device or environment, or attributes about a resource.

- The ABAC policy does not require the subject provide their O rthonym. Subjects could be anonymous or pseudonymous. What is required is the presentation of a set of attributes that satisfies the policy.

- The requestor can be required to bind the supplies attributes to the channel with the PDEP to a level of assurance as required by the PDEP. If the PDEP only requires low assurance, bearer token over TLS would be suitable. If the PDEP requires higher assurance, then holder of key tokens over TLS would be required where the token key is bound to the TLS channel.

- This model also supports Capability-Based Access Control (CBAC) where security tokens represent a capability to meet a policy. Once a subject has proven compliance with a policy, they can be issued a capability token. The client can subsequently present this capability token in lieu of a token or tokens with the set of subject attributes. The net result is the model can transition to a Capability Based Access Control because the capability token is an un-forgeable token of compliance with a policy. The token can be used with any resource tagged with the same policy.

- Plasma has a baseline of a secure transport between the DR and the PDEP. One of the decisions the PDEP has to make is the level of assurance on the release of the CEK to the subject. For example the PDEP can release a clear text CEK over the secure transport to the DR. Alternatively the could require the production of a high

assurance X.509 encryption certificate as a subject attribute to generate an encrypted CEK.

For the purpose of the Plasma work, it is desirable that the DR and PDEP be clearly defined as separate services which may be on separate systems. This allows for a generalization of the model and makes it less dependent on any specific deployment model, policy representation or implementation method. It also allows for a greater degree of control of the PDEP by an organization such that it is possible to keep as all of the PDEP resources directly under it's control and independent of the data storage location.

The base set of information for a Plasma client is as follows:

- The address of one or more IdP able to issue identity attributes to the subject
- A means to authenticate to the IdP(s)and issue attributes to the subject
- The address of zero or more AtP(s) able to issue additional attributes to the subject
- The address of one or more Plasma PDEPs able to issue role tokens to the subject to initiate Plasma policy discovery.

From this base set of data, the subject is able to authenticate to each Plasma PDEP in turn using the identity token from the IdP and discover the set of assigned roles. Each role has a set of policies which can be applied to data. A subject may be assigned to multiple roles and therefore has the ability to select the most appropriate role for the content being created. Once a role is selected the subject is able to select from the policy collection for that role. Role assignment is dynamic so the role discovery needs to be done on a regular (but not frequent) basis. Policy selection during content creation can be either manual or automatic. A DR may have sufficient context to be able to select the role and policies for the subject or have some rules that facilitate policy selection.

The model allows the content creation DR to discover the role assignments from multiple PDEP which would allow the subject to asset policies based on roles from within their organization and from any partner organization due to cross organization collaboration. The PDEP's who are authoritative for the role assignment for a subject may be different from the PDEP who are authoritative for enforcement

of a policy collection in question. The DR uses the role token to authenticate the content creation request. The PDEP will check the requested list of policies for the information is a subset of the policies in the role token. If the set of policies are a subset of the policies in the role, then it will issue the metadata token to be attached to the protected information.

Policy rules processing and distribution is complex so Plasma model does not require policy rules to be distributed to the DR. The DR just uses opaque references to the policies. The references are bound to the content to reflect the set of policies that are applicable to the data in such a way as they will travel with the data. The use of policy references minimizes any policy maintenance issues due to policy updates. The DR can be required to carry out obligations of the policy such as specific encryption requirements e.g.key size or algorithm; data integrity requirements or creating audit records.

The PDEP makes its decisions based on the requested action from the DR, the policy requirements from the PAP and the information from the PIP about the subject and the subjects environment. The information about the subject may be exchanged directly between the PIP and the PDEP (Back end Attribute Exchange) or indirectly via the DR (Front end Attribute Exchange) or both.

There is no guarantee that Identity and Attribute providers will consistently use the same name to identity a specific attributes or attribute data. For example they may use different schemas to identify an email address or use localized names to describe job functions or roles. These kinds of values may be standardized within communities of interest, but not globally across all identity and attribute providers. Therefore it is necessary to canonicalize the attribute names and values before processing by the policy. The attribute name and value mapping is part of the policy data set i.e. it is in addition to the policy processing rules.

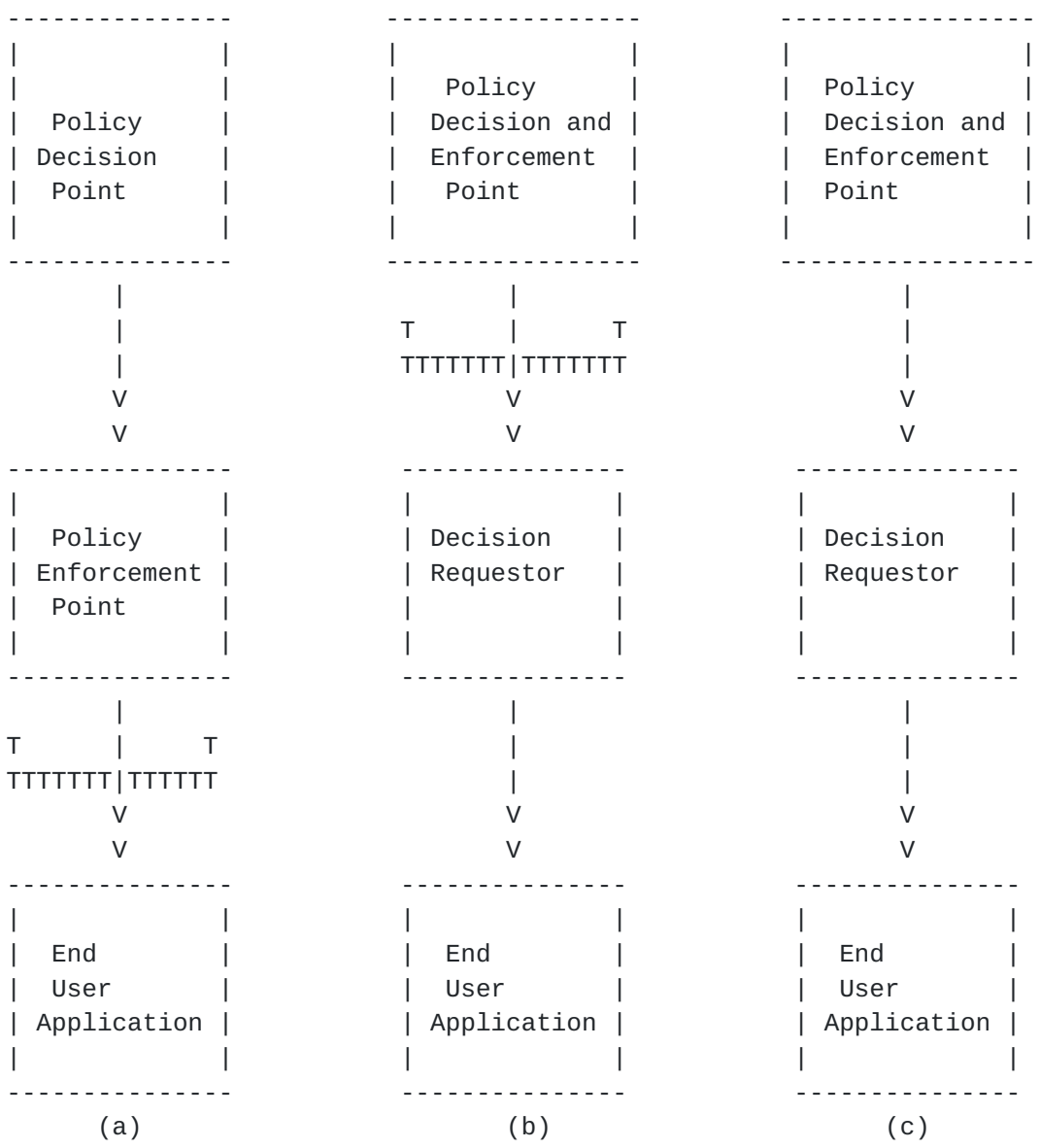


Figure 2 Options For Trusted Actors With Data.

When drawing a line where the actors in the model are full trusted with the clear text data there are three possibilities (see figure 2).

Figure 2a shows the full trust line between the user application and the Policy Enforcement Point(PEP). This is the model for current standard access control e.g. XACML [XACML-core]. In 2a, the PEP has full access to the plain text data. It makes decision requests to the PDP and if the decision is allow the PEP releases the data to the application. To use fig 2a for secure Email would require every MTA and MUA to be fully trusted with plain text data which is impossible.

Figure 2b shows the full trust line between the PDEP and the DR. In 2b, the DR only has cipher text data. The data is encrypted with a content encryption key (CEK) and the PDEP has the CEK. The PDEP releases the CEK to the end user application when access is granted so the application can recover the plain text. This mode is viable for secure Email as it does not require the MTA to be trusted with the plain text data and either the MTA or MUA can act as a DR.

In figure 2c, no actor is given full trust. When the data is encrypted, the CEK is encrypted for each recipient just as S/MIME does today. The encrypted CEKs are given to the PDEP and the PDEP releases the encrypted CEK when access is granted. This mode is also viable for secure Email as the sender can use either conventional Public key cryptography or Identity Based Encryption[RFC5408] to protect the CEK for each recipient.

4.1 Plasma Client/Server Key Exchange Level of Assurance

There are a number of mechanisms by which a client and servers can exchange CEKs. As a baseline, Plasma is establishing a secure transport between the client and server via TLS. However the client may be a proxy acting on behalf of the subject, therefore transporting a clear text CEK over the TLS transport would expose the key to the proxy. There also may be a proxy at the server which is terminating the TLS transports and forwarding the requests to another server which would mean a clear text CEK over the transport would be exposed to the server proxy. Policies may require a higher level of assurance that the CEK is not exposed to unauthorized principals. This requires encrypting the CEK for the subject before transport. This would require the client or the server to provide a public key to the other party to be used to protect the CEK before sending over the secure transport.

4.2 Policy Data Binding

There are three ways to bind policy to data.

- By value. This is where a copy of the machine readable rule set is directly associated with the data e.g. where a file system has a Access Control List for the file or directory or where a rights management agent has embedded a copy of the policy expressed in a policy expression language in the rights protects data. When an access request is made to the data, the PDEP compares the access request to the policy on the data itself.

- By reference. This is where a reference to the policy is directly associated with the data. e.g. a URI or a URN which identifies the policy to be enforced or points to where the policy is published.

For example with S/MIME the ESS label identifies the applicable policy by an OID. When an access request is made to the data, the PDEP finds the policy based on the identifier and then compares the access request to the referenced policy.

o By inference. This is where the policy has a target description in terms of resource attributes the policy applies to. When an access request is made, a set of attributes describing the resource which is the subject of the access request are included in the request by a PEP. The PDP then compares the resource attributes to the set of target descriptions of the policies in its policy store to determine the set of policies to apply to the request. For example when you author a XACML policy, you also define a target description in terms of the attributes of the resource for the policy. When an access request is made, the PDP finds the policy using the set of attributes of the resource looking for any policies that match the target description associated with the policy. It then processes the access request using the identified policy set.

The chief strength of binding policy by value is its simplicity. The policy is local to the data can easily and quickly be read. The chief weakness in binding policy by value is maintaining policy over time as binding by value results in the policy being replicated for every instance of data the policy is applied to. Many policies have a multi-year life span and during the course of that time there is a very high probability that the policy would need to be updated. Given the high number of copies, it has proven to be an very costly and imperfect process both from an enforcement and audit perspective. This process is complicated by the fact that because only the result is stored and not an identifier, it is hard to identify the policy which has to be updated.

The chief strength of binding by names is once bound to the data the association with the policy travels with the data. The chief weakness in binding by name is it requires the reference to be strongly bound to the data. This is possible using cryptography but then process of persisting the binding impacts the storage format. This can break backwards compatibility.

The chief strength of binding by inference is it can often be applied to data without impacting the storage format providing the data already has resource attributes such as with a SQL table. The chief weakness in binding by inference is the reliability of the matching in part due to the assumption the necessary policy is in the policy store. Any matching process must have a false positive and false negative rate. These rates have to be evaluated on a case by case basis over time as it can change making compliance expensive. The set

of available attributes also varies with different data types e.g. structured database information has a rich set of attributes whereas unstructured data such as documents and files have a poor set of resource attributes. This inconsistently over available attributes impacts matching reliability. The resultant set of policies for a policy target is also dependent on the correctness of the set of policies evaluated. It's also impossible to detect if a policy is missing from the policy store which again would mean incorrect policy enforcement

The Plasma model is choosing to use binding by name because we need to encrypt the data which means we will impacting the storage format anyway which negates the main weakness of binding by name. We get the reliability of policy enforcement which is independent of location and we get low maintenance since we are only storing a reference to the policy and not the policy with the data.

4.3 Content Creation Workflow

The Content Creation DR bootstraps itself via the following sequence of events:

(1) The content creation DR is configured with the set PIP's and PDEP's it trusts.
(2) The content creation DR summits a request to all the trusted PDEPs for the set of roles it allows for the subject. The subject is authenticated and authorized for the roles via attributes from the PIP. The PIP attributes can be obtained by the PDEP either via front-end (related to the PDEP from the PIP via the subject) or back-end (direct exchange between the PDEP and the PIP) processing.

(3) The content creation DR receives a list of roles the PDEP can configured for the subject
(4) The DR submits a request for the policy collection for each role. Additional attributes may be required from the PIP to authorize the release of the role token.

Now the DR is bootstrapped with a list of roles and for each role, a policy set . Now the DR is ready to create content. When the user wants to create protected content, they use the following sequence of events

(i) The user creates the new content
(ii) The user select the appropriate role for the content, then selects one or more policies from the policy set that are applicable to the content

(iii) The DR encrypts the content with one or more locally generated CEKs
(iv) The DR submits the CEK(s), the set of requires policies to be applied and the hash of the encrypted content to the PDEP. The CEK can be a raw key or a CEK key encrypted by a KEK if the policy does not want the PDEP to have the ability to access the plain text data.
(v) The PDEP generates the encrypted metadata which contains the list of policies and the CEKs. The metadata is encrypted by the PDEP for itself. The PDEP includes a URL for itself and the hash of the protected content as signed authenticated attributes then signs the encrypted metadata.
(vi) The PDEP returns the metadata to the DR
(vii) The DR attaches the PDEP metadata to the protected content and distributes the content.

4.4 Content Consumption Workflow

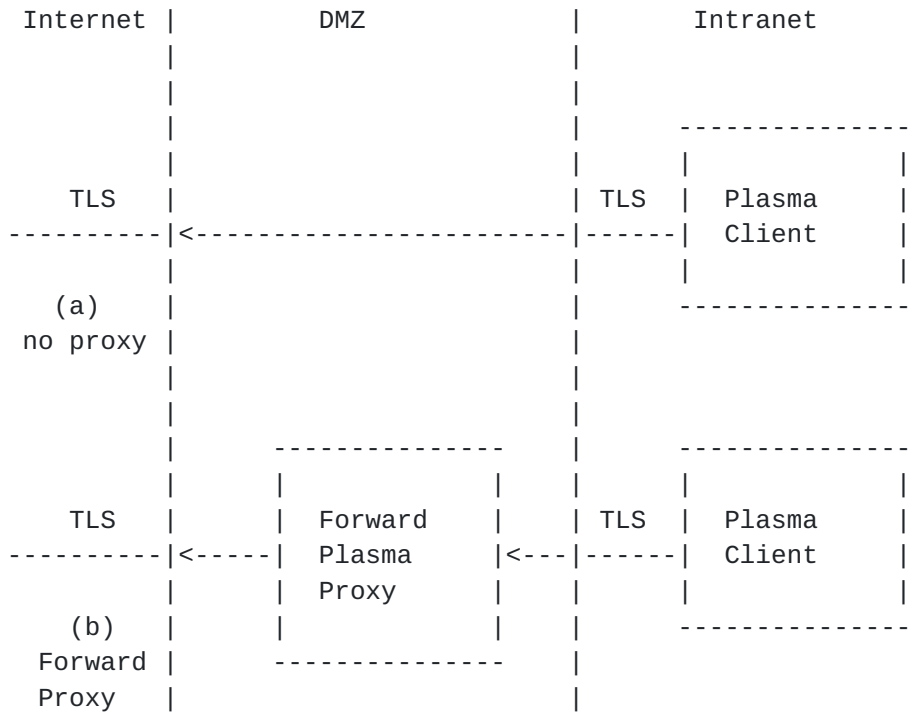
When a user wants to open some protected content they would follow the following workflow.

(A) The DR verifies the certificate in the signed metadata then determines via local policy if it want to process the protected information based on the identity of the PDEP
(B) The DR verifies the signature on the metadata token and the binding to the encrypted data by hashing the encrypted information and comparing it to the authenticated attribute in the metadata
(C) The DR forwards the signed metadata and requests a read token for the content from the PDEP using the address of the PDEP in the authenticated attribute in the metadata
(D) The PDEP decrypts the metadata, de-references the policy pointers and determines the set of access rules based on the policy published by the PAP. The PDEP then determines the set of attributes it needs to evaluate the access rules. The PDEP can the use PIP is has direct relationships with to query attributers about the subject. If the the PDEP is missing attributes it need to process the policy, it returns a list of the missing attributes to the DR
(E) If the DR receives a list of missing attributes from the PDEP, it obtains the missing attributes requested by the PDEP and sends them to the PDEP in a new read token request.
(F) Once the PDEP has a complete set of attributes, and the attribute values match those required under the access policy, the PDEP releases the CEK to the DR along with a TTL which defines how long the DR can use the CEK before it must discard the CEK and reapply for access.
(G) Once the DR has the CEK it decrypts the information. It caches the CEK until the TTL expires.

4.5 Plasma Proxy Servers

There are two separate use cases for Plasma Proxy servers. The forward proxy use case where the Plasma client needs to connect to a Plasma server outside of its organization and the reverse proxy use case where the Plasma client outside an organization, need to connect to a Plasma server.

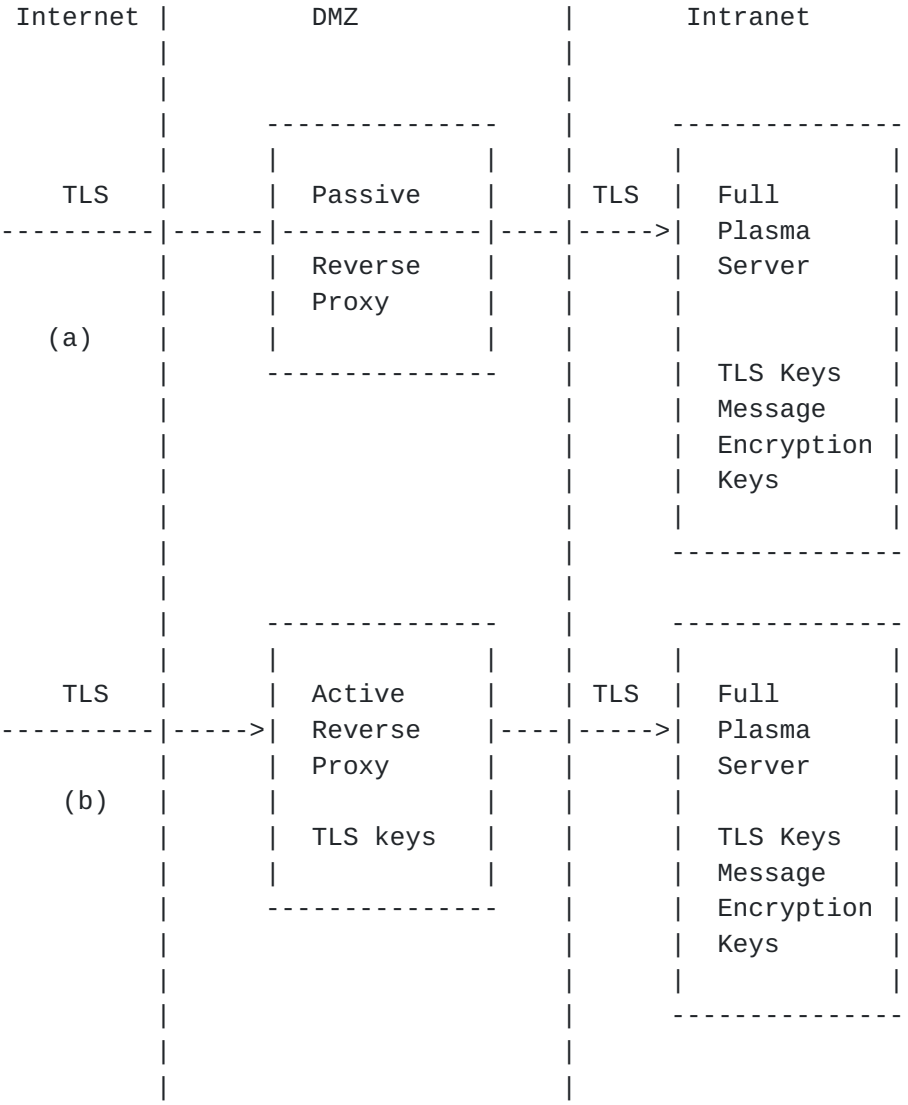
A recipient has no control over senders creating Plasma email and sending to them. Malicious sender can craft harmful payloads and protect it in a Plasma envelope. Therefore Plasma recipients need a policy to determine the set of Plasma servers they are willing to interact with. This can be a local policy which is pushed down to every Plasma client. An alternate approach is to have a forward proxy manage the policy on behalf of the Plasma recipient. A forward proxy would eliminate the need to push policy about the set of trusted Plasma server by mediating the connection requests from the Plasma recipients to the Plasma servers. The forward proxy could be a server belonging to the client organization or a cloud service.



Forward Plasma Proxy

Since the Plasma service has sensitive cryptographic keys used to protect the message CEKs, it would be unwise to host those directly connected to the Internet. However, Plasma servers will need to be Internet addressable to Plasma requests from DR's outside the organization. The simplest possible configuration would be to have a

passive reverse proxy in front of the Plasma server. Since Plasma is using TLS with channel binding, the passive proxy has a limited function and would be only able filter based on IP addresses. The Plasma protocol is a series of request-response messages, so an active reverse proxy can be implemented like other store and forward message based services (e.g. SMTP). The Internet facing proxy server would terminate the TLS from the external DRs, ensure DR can authenticate the TLS connection. Because the active proxy terminates the TLS session, it can scan submitted messages to ensure they are not malformed and are free from malicious content before relaying messages to a full Plasma server further inside the network for processing of the request.



4.6 Policy Types

Policies range from very simple to very complex. Policies have

dependencies not only on the technical implementation of the software but on the range of attributes a PIP would issue to subjects. This is likely constrained by the physical procedures a PIP would support to capture and verify the information about the subject. To manage this range of requirements, this model uses two type types of policy.

4.6.1 Basic Policy

Basic policy is intended to be universally usable by using a small fixed set of attributes. For example, basic policy is intended to be equivalent to sending encrypted Email with S/MIME today. It is a simple policy that authenticated recipients of the Email get access to the message. Its intended target is simple scenarios involving consumers and small businesses who are using public PIP which issue a limited set of attributes. It is expected that all Plasma clients and commercial IdPs would be capable of supporting basic policy due to their simplicity and basic attribute set required by basic polices. As the available set of attributes increases over time, later standards may define more basic polices which a bigger set of attributes types.

4.6.2 Advanced Policy

Advanced policy is intended to be used where one or more arbitrary policies are required on the content. It is intended to target more complex scenarios such as content with regulated information or content subject to other organization and contractual policies. The input set of attributes is defined by the policies are in theory unbounded and can be either primordial such as date of birth or derived attributes such as age or both. In practice, advanced polices are constrained by the set of attributes available under the IdP Trust Framework for the subjects. A data object may require multiple policies and any instance of multiple policies requires a logical relationship between the polices e.g. they can be AND-ed or OR-ed together. It is not expected that all Plasma clients support the rich set of attributes necessary for advances policy.

5. Message Protection Requirements

5.1. General Requirements

Confidentiality policy protected data MUST be where the data is protected from unauthorized disclosure, integrity protected from unauthorized alteration AND provides data origination authentication so that recipients know who created the data.

Integrity protected is where the data MUST be integrity protected AND provide data origination authentication and recipients are NOT

allowed to alter the data.

Every authentication has a level of identity assurance associated with it depending on attributes such as the identity checks made about the subject and the authentication technology used by the subject. The authentication of content creator and content consumers MUST support the multiple levels of identity assurance framework. (see scenarios 3.1, 3.2, 3.3 and 3.4)

The specifics of every possible authentication mechanism or every detail about how the subject's identity was proofed by the IdP cannot be known to the DR and PDEP, therefore the specifics of how sender or recipient achieve the required level of identity assurance MUST be abstracted from the PDEP and DR by use of a simple numeric scale (e,g, 0-4, or 1-6) liked to an identity assurance framework identifier which defines the specifics of how to derive the LoA.(See section 3.1, 3.2, 3.3 and 3.4)

Access policies are complex and subject to change over time. For this reason, policies MUST be identified by reference rather than inclusion of the actual policy with the data so the policy change can be implemented without updating the data. (See section 3.4.1)

Access to the plaintext of the content MUST only be provided after the recipient has either provided suitable valid attributes to the PDEP or the PDEP was able to find attributes about recipient directly from a PIP, to satisfy the policy as defined by the sender (See section 3.1 3.2, 3.3, 3.4.1, 3.5)

The sender MUST be provided with a list of policies applicable to content they create and scoped to their current role i.e. what tasks they are currently assigned to deliver(see scenarios 3.1, 3.2, 3.3).

The specifics of the access control policy used by the PDEP MUST be abstracted from both the sender and recipients i.e. the DR MUST NOT make the access control decision or need specifics of the access policy(see scenarios 3.1, 3.2, 3.3 and 3.4).

Content consumers DR MUST receive authenticated attributes of the identity of the creator, the level of identity assurance of the creator and the cryptographic fingerprint of the original content so the DR can confirm who created the content and that the content has not been altered (see section 3.1, 3.2, 3.3 and 3.4)

The key exchange between content creator and content consumer and the PDEP MUST support multiple levels of assurance so an appropriate strength of mechanism can be selected based on the level of assurance required. For example, for low assurance situations this could be via

a plan text CEK over a secure transport such as TLS. For high assurance situations recipient MAY be required to provide a suitable key exchange key such as an X.509 certificate to encrypt the CEK. (See scenarios 3.3 and 3.4)

The level of key exchange assurance requited MUST be selected by the sender and enforced by the PDEP (See section 3.1, 3.2, 3.3 and 3.4).

If the content consumers is unable to initially comply with the content creators policy, they MUST be able resolve any issues by getting the suitable credentials or attributes and gain access to the content without intervention from the content creator.

A time-to-live MUST be provided to content consumers when access is granted by the PDEP to define when the DR MUST discard the message CEK and submit a new access request to the PDEP. The TTL value MUST be based on the message policy and optional attributes about the content consumer and their environment.

The PDEP MUST be stateless for processing policy requests from content creators and consumers with respect to any instance of protected content. It MUST be possible to have multiple instances of a PDEP service and load balance requests across all instances of the service transparently to the client and not require synchronization of state about requests between instances of the service.

A PDEP MUST be capable of generating audit events associated with access to protected content using policy defined by the PAP.

5.1.1 Email Specific General Requirements

It MUST be possible for domains to publish keys and attributes about the boundary inspection agents. This allows senders to pre-authorize the inspection agents of recipients for access to the message. It MUST be possible for boundary inspection agents to request access to protected messages which have not been preauthorized by the sender.

It MUST be possible for MTAs to request access to protected messages which have not been preauthorized by the sender (see section 3.5).

5.2. Basic Policy Requirements

The use of Basic Policy MUST be backwards compatible with existing S/MIME. A sender's agent MAY discover some recipient's certificates and create recipient info structures using the existing standard (unless specifically forbidden by the selected policy). A sender's agent MAY elect to use this mechanism for recipients for whom keys cannot be discovered.

One Basic Policy is to be defined by this work. The Basic to map to NIST 800-63-1. This process does not preclude other Basic Policies to be defined by other groups or even within the context of the IETF.

When using Basic Policy, the sending agent MUST define which basic policy is required and the list of recipients.

Basic policy MUST support multiple levels of identity assurance. The levels of identity assurance MUST map to an existing identity authentication assurance framework e.g. to NIST 800-63-1 or equivalent.

A sender using Basic policy MUST be able to send protected messages without discovering any recipient's encryption key.

Using basic policy MUST NOT require bilateral agreements between sender and recipients a priori to sending the message.

5.2.1 Email Specific Basic Policy Requirements

The use of Basic Policy MUST be backwards compatible with existing S/MIME.

A sender's agent MAY discover some recipient's certificates and create recipient info structures as per the existing S/MIME standard and elect to use the new mechanism for recipients it cannot discover keys for rather than remove the recipient's without certificates.

5.3. Advanced Policy Requirements

A basic policy MAY be combined with advanced policies

It MUST be possible to apply one or more Advanced Policies to a protected content. Where 2 or more policies are applied to protected content, the logical relationship between the policies MUST also be expressed i.e. are the policies a logical AND or a logical OR. (See section 3.3)

An advanced policy MAY require attributes about:

- The content consumer
- The device the content consumer is using
- The environment of the device is attempting to access the protected content from
- The content being accessed

Advances policy MUST support an extensible list of obligations on the

content creator where use of the policy requires some specific action on the part of the content creator e.g. sign content with 2 factor smart card and/or that the signature is legally binding, or the message needs to be verified for an extended period(see scenarios 3.3 and 3.4).

Advanced policies must support the ability to verify the content for an extended period (10 or more years)

6. IANA Considerations

This document describes the requirements for message access control. As such no action by IANA is necessary for this document

7. Security Considerations

Authentication by itself is not a good trust indicator for users. Authentication raises the level of assurance the identity is correct but does not address whether the identity is trustworthy or noteworthy to the recipient. Authentication should be coupled with some form of reputation e.g. the domain is on a white list or is not or a black list. Malicious actors may attempt to "legitimize" a message if an indication of authentication is not coupled with some form of reputation.

Malicious actors could attempt to use encrypted Email as a way to bypass existing message pipeline controls or to mine information from a domain. Domain should have sufficient granularity of policy to handle situations where their Email pipeline agents have not been authorized to inspect the contents.

It must be possible for a third party to, upon correctly presenting a legitimate legal justification, to recover the content of a message. This includes the Sender's and Recipient's companies for business continuity purposes, as well as Law Enforcement. If the entity requesting the information and the entity controlling the access are in different jurisdictions, then the process would be subject to some form of rendition.

The use of a security label type that requires the recipient of a message to query a PDEP in order to obtain the contents of a message opens an additional method for adversaries to confirm that an Email address does or does not exist. Additionally it allows for a new channel for materials to be delivered to the recipient's mail processor that is not checked for malware or viruses by the standard mail scanning methods in place. For these reasons recipient processing systems need to implement the following counter-measures:

1) The pointer to the PDEP MUST be checked against some policy before attempting to query the PDEP for a policy decision. 2) Care MUST be taken when processing the responses from a PDEP check that they are well-formed and meet local policy before using the responses.

Editorial Comments

Appendix A. References

A.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3198] Westerinen et. al., "Terminology for Policy Based Management", November 2001.

[RFC5035] Schaad, J., "Enhanced Security Services (ESS) Update", August 2007.

[RFC5280] Cooper, D, et al, "Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile", RFC 5280, May 2008

[RFC5652] Housley, R., "Cryptographic Message Syntax (CMS)", RFC 5652, September 2009.

[RFC5750] Ramsdell, B. and S. Turner, "Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.2 Certificate Handling", RFC 5750, January 2010.

[RFC5751] Ramsdell B., Turner S., "Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.2 Message Specification", January 2010

[SAML-core] OASIS, Assertions and Protocols for the Security Assertion Markup Language (SAML) Version 2.0, March 2005

[sp800-63-1] NIST SP 800-63-1 "Electronic Authentication Guideline", December 2008

A.2. Informative References

[bc-iaf] Province of British Columbia; Electronic Credential And Authentication Standard, version 1.0

[kan-iaf] Kantara Initiative; Identity Assurance Framework: 4 Assurance Levels, version 2.0

[lib- iaf] Liberty Alliance; Liberty Identity Assurance Framework, version 1.1

[RFC3114] Nicolls, W., "Implementing Company Classification Policy with the S/MIME Security Label", RFC 3114, May 2002.

[RFC5408] Appenzeller, G., "Identity-Based Encryption Architecture and Supporting Data Structures", RFC5408, January 2009.

[SAML-over] OASIS, Security Assertion Markup Language (SAML) Version 2.0 Technical Overview

[XACML-core] OASIS, eXtensible Access Control Markup Language (XACML) Version 3.0 Core Specification

Appendix B Authors' Addresses

Trevor Freeman

Microsoft Corp.

Email: trevorf@microsoft.com

Jim Schaad

Soaring Hawk Consulting

Email: ietf@augustcellars.com

Patrick Patterson

Carillon Information Security Inc

Email: ppatterson@carillon.ca

Appendix C Document Change History

Added general data model (section 4) Added regulated industry Email scenario (section 3.4 Split requirements into general requirements and Email specific requirements Cleaned up scenarios to differentiate requirements and workflow Fixed multiple document nits from Jim Schaad Need a paRAGRAPH on namespaces to deal with SAML attributes of different names with same semantics Don't use a proxyf5 tls offload. Define a proxy in arch Made changes from XACML TC to better align terminology Cleaned up integrity scenario Merged the two vocabulary sections into a single section Added LOA for key exchange Added the forward proxy to the architecture addressed [anchor21] comment by clarifying text in paragraph 1.2 Addressed [anchor22] comment by clarifying text in paragraph 2.4 Addresses [anchor23] comments by clarifying text in section 4. Addresses [anchor24] comment by clarifying text in section 3. Completed the scalable policy decision making scenario (3.10) Added Email compliance scenario (3.7)