Network Working Group
Internet-Draft
Intended status: Informational
Expires: September 13, 2012

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

March 12, 2012

Requirements for Message Access Control
draft-freeman-plasma-requirements-01

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 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 doing/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.

Status of this Memo

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 [RFC2634] 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 then 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 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, GPG/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 [RFC2634] 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 criteria of 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. These criteria are defined in terms of attributes about the subject requesting access. Examples of the types of attributes would include what roles the subject is assigned to (Role Based Access Control) or one or more attributes about the subject (Attribute Based Access Control). While an ESS Security Label provide a standardized representation of a policy identity, it does not define any methods of obtaining the necessary attributes or policy description in order to enforce the policy. Standards now exist that enable the transport of subject attributes [SAML-

overview]. Adoption of these subject attribute protocols would allow a rich set of access control policies to be supported by S/MIME in line with other applications.

An ESS Security label is a signed attribute of a SignedData object which indicates the access control policy for the message. The fact that this is a signed attribute protects the integrity of the label and provides a tamper evedent binding of the label to the message but does not protect the confidentiality of the body. 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 users 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 this is 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 unauthorized under the policy to 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 some form of authentication credential but the credentials are not X.509 certificates and who therefore cannot use S/MIME. Standard based services (e.g. [SAML-overview])are now available which abstract the specifics of an authentication technology used 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 technologoies to be able to use S/MIME to secure Email. It also allows for new authentication technology to be deployed without impacting the core S/MIME protocol at the expense of adding a third party to the transaction.

1.3. Vocabulary

Mail User Agent (MUA) is 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.

Mail Transfer Agent (MTA) is a program that transfers email from one computer to another. An MTA implements both the sending and receiving of email.

A Cryptographic Lock Box is a per recipient data structure which holds a content encryption key encrypted for a specific user. CMS implements lock boxes as RecipientInfo structures.

Early Binding is the concept of creating the cryptographic lock box for a recipient at the time the message is sent. (See to Late Binding).

Late Binding is 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)

Content Encryption Key (CEK) is a key used to encrypt protected end user data. (See Key Encryption Key)

Key Encryption Key (KEK) is a key used to encrypt a cryptographic key, often a CEK. (See Content Encryption Key)

Authentication denotes:-

The sender is able to establish to a known level of confidence the identity of the recipient or
The recipient is able to establish to a known level of confidence the identity of the sender.

Confidential denotes that a message has been protected to a known level of confidence so that the contents are not decipherable by unauthorized users.

Integrity Protected denotes that 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.

Front End Attribute Exchange is when subject attributes are relayed from the issuer to the relying party by the subject

Back End Attribute Exchange is when subject attributes are directly sent from the issuer to the relying party

Role Based Access Control (RBAC) is access control based on the assignment of authorizations to abstract job function or principle known as a role. Subjects are then allowed to assume one or more roles based on their job needs. A role is distinct from a group because a group is a collection of subjects which has no intrinsic authorizations.

Attribute Based Access Control (ABAC) is access control based attributes of the subject. An ABAC policy specifies which attributes are needed to authorize access to a resource. These attributes may be provided by the subject as part of the access request or discovered by the access control engine based on relationships it has with attribute sources such as a directory or personnel database.

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. While CMS allows for other types of security credentials to be used, S/MIME exclusively uses X.509 certificates [RFC5750] for the security credentials used 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 the 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

- o The recipient's repository may not be accessible to the sender e.g. it's behind a firewall

- o The sender may not implement the algorithms supported by the recipient

- o 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. The use of a mailing list also means that the originating sender has no chance to perform any sender side filtering on who should receive an Email based on the recipient's attributes as they do not know the full list of the recipients.

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 checks before access to the data should be granted. In many inter-organization collaboration scenarios it's impossible for the sender to satisfy the access checks on behalf of the 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 [RFC2634]. 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 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 course 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. If 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 be removed by any agent that has key access to the message (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 course 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, and updates to the policy may be eradicate. For cross cooperate 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 recipient 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 is 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.

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 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 HMAC 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 may be 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 [RFC2634] thou 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 any specific combination of identity proofing, credential technology, authentication technology from the authorization policy. 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.

3. Use Case Scenarios

This section documents some use case scenarios the new protocol aims to support.

Author Note: Add legal document signing scenarios

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 certificates difficult to obtain and then use. One of the possible use cases of PLASMA is to try and deal with this. The details of the use case are therefore: Alice wants to send an Email message to Bob that is secure so that eavesdroppers cannot read it. Bob however has never obtained an X.509 certificate for this purpose. 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 encrypted 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 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 data. This would include doctor, patient; bank, account holder; Medical insurance, insured person. Two examples are presented here.

3.2.1 Bank Statement Email

In the first example, a bank (The Bank of Alice) has determined that it will be using Email to distribute statements to its customers (Bob). The information is confidential, so any channel of communication they 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 assurance. The same 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 includes the statement of balances and activity either as an attachment or as the body of the message.

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

3. The statement mailer for Alice knows the protections to apply based on the policy; it knows to encrypt and integrity-protect protects the message and what level of assurance required for the 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 his Bank. Bob can verify the message has not been altered as it is signed by the his Bank. Bob uses his on-line banking credentials to prove his identity 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 customer. Thus messages dealing with loan applications and changes in bank policies can be sent out in the same manner, potentially using slightly 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 uses his on-line credentials when obtaining the policy description blob sent with the message there is a degree of assurance that the bank has similar to using web-based messaging today that it was Bob who sent the message.

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, so any channel of communication she selects must protect Bob's privacy. 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, the individual who is making care decisions for Bob (i.e. Bob's parent), and the 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 message readers authenticate with a level 2 or above level of identity assurance.

(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. (a) 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 as 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 could be 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 Author note: (add alternate recipient's as well)

Early in the relationship between two companies, it is frequently necessary to exchange sensitive information. This needs to occur before the relationship has matured to the point that a formal relationship is reflected through a 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 example, the IT staff might need to provide cross certifications and exposure of certificate repositories.

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 authenticates with a level 2 or above level of identity assurance

(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 sensible 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 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 relationship 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 legal 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 derives from their role or function. Also many regulated environments need to be able to verify the information for a 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 users contribute information to the Email thread.

Company Foo has been awarded a contract to build some equipment (Program X). The equipment is covered by export control. Company Bar is a 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 export control requirements. Company Foo also set up separate rules to cover the protection of its intellectual property contributed to Program A. Company Bar also sets up its own policies to protect 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.

Frank is an employee of Company Foo. He has been assigned as a team leader on Program X and an individual contributor onProgram Y. 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 Program X. When Frank sends the Email with Program X regulated content he must ensure

compliance with the export control policies. When frank sends program directions as team lead, recipient's need to verify his authority and for compliance the message need to be able to be verified for 10 years. If Frank includes Company Foo intellectual property, he must also ensure compliance with his corporate IP protection policies. When Frank sends a Program X Email he must ensure the following objectives:

(a) Only recipients who meet the Program X policy and or or Company Foo's intellectual property protection policy can read the Email
(b) To comply with policy as team lead, Frank must sign the message with certificate to indicate the signature message is legally binding e.g. it does not just indicate the identity of the sends and protects the integrity of the message.
(c) The message is also signed to indicate originator signature complies with the policy and the originator had presented the necessary claims and the allows the message to be verified for at least 10 years.
(d) The recipients authenticates with an acceptable level of assurance (i.e. level 3 or above)
(e) Recipients present any other attributes about themselves necessary to verify compliance with the applicable policies (theirs program assignment, nationality, professional or industry certifications)

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

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

(h) Recipients are made aware that the message is a Program X Email and the contents can only be shared with other Program X workers.

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 Z team lead business context which is appropriate for his work.
(3) Frank selects the Program X team lead and Company Foo IP policies from the list of available policies.

The Email client knows the protections to apply to the Email; the

message needs to be encrypted and integrity-protect the message using a certificate which is legally binding, it also has a signature which allows the message to be verified for at least 10 years; the level of assurance required for the recipients identity and what recipient attributes are necessary to access the message.

(4) Frank clinks the send Email button. The client signs the Email using Franks smart card and a certificate indicating the signature is legally binding. The client obtains a second signature on the message to indicate Franks authority to send the message and allow it to be verified for at least 10 years. 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.

(5) Grace receives the Email as sees it is a secure message from Frank. Grace's client provides the attributes necessary to comply with the policy which includes her level 3 encryption certificate to the PDP.

(6) Once Grace has shown she passes the policy requirements, the PDP releases the message CEK to grace using her level 3 encryption certificate.

(7) Grace uses her smart card to open the message. She sees the message is 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 messages. Grace includes some information which is Company Bar's IP so add 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 because it contains information from Company Foo and Company Bar both companies wants consistent policy enforcement on its information.

3.5 Delegation of Access to Email

There are a number of times when either others are given access to a recipeints mailbox or Email is forwarded to other recipients based on

recipients 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.

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 area 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 participates, 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.   workers.

The sequence of events Frank 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 business context which is appropriate for her work. Hanna selects a Patient Consultation context.

(3) Hanna selects the Patient Referral and Patient Consent decree policies from the list of available policies.

The Email client knows the protections 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 clinks the send Email button. The client signs the Email using Hanna smart card. The Client then encrypted the message and sends it to his Email server.

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

(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 PDP.

(6) Once Ida has shown she passes the policy requirements, the PDP 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 Sensitive Patient Data policies

3.7 Regulated Email Compliance Verification

TBD

3.8 Email Pipeline Inspection

TBD add pre-auth and regular access

Organizations have a huge incentive to want to some level of inspection on all mails 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.

Company Foo receives Email from the Internet. Company Foo has a policy to scan inbound Email with a view to remove inappropriate or malicious content. The scanning of inbound Email for Company Foo can happen on their own servers or it may be outsourced to a third party. This works fine as long as the Email is not encrypted, however in that event the server will have a policy on how to deal with encrypted mail. For some companies, encrypted mail will be passed through and virus detection software on the recipient's client will be relied on. In other circumstances, the decryption key used by the recipient is shared with the gateway software so that it can decrypt the message.

The ability to decrypt and check the content for malicious content is highly desirable even when a PLASMA encrypted Email message is encountered. The methods that this can be dealt with using on of the following methods:

1. The scanning server authenticates to the policy server as the entity doing virus and malware scanning. If the policy has specific attributes that allow for access to be granted to such a scanning service, the appropriate decryption keys will be released and the server will scan the mail and take appropriate action.

2. The policy server is configured with information about the server took 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 (lockbox) to the sender of the message for direct inclusion in the recipient info list. This allows normal processing by the scanning software without the necessity to stop and query the policy server for keying information at a cost of needing wider configuration.

3. 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 the 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 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.

3.9.1. 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 protections 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.

3.9.2 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 a instant message 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 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 want 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 protections 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 hte 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 PDP.

(6) Once Grace has shown she passes the policy requirements, the PDP 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 requremnts and sees the the messages from Frank and Grace.

4. General Data Model

This work is modeled on a well established set of Actors for policy enforcement [RFC3198] [XACML-core].

4.1 Vocabulary

These terms are the same as used by RFC3198. While the roles are fundamentally the same, there are some minor differences in the responsibilities of each actor with models such as XACML.

These terms are taken included for the convenience of the reader:

Policy Administration Point (PAP): The system entity that creates policies or policy sets. The policies define the rules, their conditions and actions associated with the policy.

Policy Publication Point (PPP): A service where policies are published.

Policy Decision Point (PDP): A service that is able to interpret the policy rules authored by a PAP and published by a PPP using information supplied by a PIP to renders decision requests from a PEP.

Policy Information Point (PIP): A service with issues assertions about subjects or the subject's 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 PDP. Attributes can be distributed directly between the PIP and the PDP (Backbend Attribute Exchange;BAE). Alternatively attributes may be distributed via the PEP (Front End Attribute Exchange; FAE) There are two types of PIP based on the types of attribute the PIP would assert about the subject. A 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 Enforcement Point (PEP): The service responsible for making policy decision requests to the PDP. In this model the access control is enforced by the PDP by its control of decryption keys.The PEP enforces any obligations the PDP may require such as signing or encryption of the data, generating audit events etc.

We additional make use of the following terms:

Policy Publication: The act of publishing a policy or policy update from the PAP to the Policy Repository. The process of policy publication is out of scope for this document.

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. For simplicity we use SAML as the format for both requesting and receiving attributes and would suggest the use of the SAML 2.0 Assertion Query and Request Protocol as one method for requesting the necessary attributes. The attributes can be requested either by the PEP (front end attribute exchange) or the PDP (back end attribute exchange).

Content Protection Request/Response: The protocol to be run by the PEP 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 to be run by the PEP to obtain the permissions and information needed to decode and access a data block with appropriate labeling. This protocol is part of the work to be defined by this group.

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.

Role: A role is a policy set that has an associated textual name. A role in this context is not to be confused with a rule in role-based policies, while the concepts are similar they are not identical.

Role Set: A collection of one or more roles.

Policy: The policy has two basic forms. A human readable form which defines a set of high level requirements. These human readable policies may be, for example, in the form of legislation,, or a legal contract. These high level policies are then translated into technical policies for implementation. Policies may stipulate many forms of requirements such as data protection, access control, integrity, data origination, data retention, etc.

Policy Set: A collection of one or more policies. The policy set

may also defines the logical relationship between the policies

Policy Identifier: Is 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 Label: The data structure which holds one or more policy identifiers and their logical relationship.

```
                            ------------------
                            |                |
                            |     Policy     |
                            | Administration |
                            |     Point      |
                            |                |
                            ------------------
                                   |
                                   |  Publish
                                   v  Policy
                                   v
                           -----------------
                           |               |
                           |    Policy     |
   -----------------       |   Publication |        -----------------
   |               |       |    Point      |        |               |
   |   Policy      |       |               |        |   Policy      |
   |  Information  |       -----------------        |  Information  |
   |   Point       |                |               |   Point       |
   |               |                |  Read         |               |
   -----------------                v  Policy       -----------------
        |  |                        v                       |  |
        |  |Issue          -----------------      Issue     |  |
        |  |Attributes     |               |      Attributes|  |
        |  |(BAE)          |               |      (BAE)     |  |
        |  -------------->>|     Policy    |<<---------------  |
        |                  |    Decision   |                   |
        |                  |     Point     |                   |
        |  -------------->>|               |<<-----------      |
        |  |Protect        |               |  Consume   |      |
        |  |Content        -----------------  Content   |      |
        |  |Request+                          Request+  |      |
        |  |Attributes                        Attributes|      |
        |  |(FEE)                             (FEE)     |      |
        v  |                                            v      v
        v  |                                            v      v
    -----------------                               -----------------
    |               |                               |               |
    |   Content     |           Distribute          |   Content     |
    |  Creation     |           Content             |  Consumption  |
    |   Policy      | ---------------------------->>|   Policy      |
    |  Enforcement  |                               |  Enforcement  |
    |   Point       |                               |   Point       |
    |               |                               |               |
    -----------------                               -----------------
```

Figure 1 General Scheme for Publishing and Consuming Protected Content

For the ESS security labels model, it is generally assumed that the PEP and PDP are coexistent on a single computer system (with the mail user agent (MUA)). There is no explicit reason that this is required, the PEP could be with the MUA on the user's system, but it could make a remote call to a PDP as is designed in this model.

For the purpose of the PLASMA work, it is desirable that the PEP and PDP 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 or policy represent, logic or implementation method. It also allows for a greater degree of control of the PDP by an organization as all of the PDP resources more directly under it's control and independent of the data storage location.

The content creation request protocol includes the discovery of the set of roles and thereby the set of policies that a content creator will be able to assert. This is based on role assignments where a subject may be assigned to multiple roles and therefore have 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 set for that role. Role assignment is dynamic rather than static, as such the discovery needs to be done on a regular basis. Policy selection during content creation does not have to be manual. A PEP may have sufficient context to be able to select the role and policies sets for the subject.

The model allows the content creation PEP to discover the role assignments from multiple PDP which would allow the subject to assert based on roles from within their organization and from any partner organization due to cross organization collaboration. The PDP's who are authoritative for the role assignment for a subject may be different from the PDP who are authoritative for enforcement of a policy set in question.

Policy rules processing and distribution is complex so the PEP in this model does not require policy rules to be distributed to the PEP. The PEP just needs opaque references to the polices and defers all decisions to the PDP. The use of policy references also minimizes any policy maintenance issues due to policy updates. The PEP can be required to carry out obligations of the policy such as specific encryption requirements such as key size or algorithm; or data integrity requirements such as signing or HMACing content.

The model is designed to be applicable to any data e.g. Email, documents, databases etc. This is to facilitate consistent policy enforcement for data across multiple applications. Another objective is to not require the PEP to have access to the plain text content in

order to be able to make decision requests to the PDP. The policy decision is complex so the PEP in this model just uses policy pointers or labels to indicate policy applicable to content. The Content consuming PEP dynamically discover the PDP's who are authoritative for the protected content in question.

The PDP makes its decisions based on the requested action from the PEP, 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 direly between the PIP and the PDP (Back end Attribute Exchange) or indirectly via the PEP (Front end Attribute Exchange) or both.

```
    ---------------          ---------------         ---------------
    |             |          |             |         |             |
    |  Policy     |          |  Policy     |         |  Policy     |
    | Decision    |          | Decision    |         | Decision    |
    |  Point      |          |  Point      |         |  Point      |
    |             |          |             |         |             |
    ---------------          ---------------         ---------------
           |                        |                       |
           |                 T      |      T                |
           |                 TTTTTTT|TTTTTTT                |
           V                        V                       V
           V                        V                       V
    ---------------          ---------------         ---------------
    |             |          |             |         |             |
    |  Policy     |          |  Policy     |         |  Policy     |
    | Enforcement |          | Enforcement |         | Enforcement |
    |  Point      |          |  Point      |         |  Point      |
    |             |          |             |         |             |
    ---------------          ---------------         ---------------
           |                        |                       |
    T      |     T                  |                       |
    TTTTTTT|TTTTTT                  |                       |
           V                        V                       V
           V                        V                       V
    ---------------          ---------------         ---------------
    |             |          |             |         |             |
    |  End        |          |  End        |         |  End        |
    |  User       |          |  User       |         |  User       |
    | Application |          | Application |         | Application |
    |             |          |             |         |             |
    ---------------          ---------------         ---------------
          (a)                       (b)                     (c)
```

Figure 2 Options Full Trust With Clear Text 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 PEP. This is the model for current standard access control e.g. XACML [XACML-core]. In 2a, the PEP has full access to the clear 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 act as a PEP so to be fully trusted with clear text data which is impossible.

Figure 2b shows the full trust line between the PDP and the PEP. In

2b, the PEP only has cipher text data. THe data is encrypted with a content encryption key (CEK) and the PDP has the CEK. THe PDP releases the CEK to the end user application when access is granted. This mode is viable for secure Email as either the MTA or the MUA can act as a PEP.

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 PDP and the PDP 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.

4.2 Overview

The model is applicable to any type data (Email, documents, databases etc). This document and the subsequent protocols are focused on the content creation and consumption elements of the document. The policy authoring, policy repository and decision logic modules are matters beyond the scope of this document. It is important to note that these blocks are logical entities and as such can be combined physically in different configurations.

o This model uses name based binding between the resource and the policy. When information is created, it is encrypted and a list of policies that must be enforce by the PDP is bound to the protected information

o The model is fundamentally an Attribute-Based Access Control (ABAC) model. Access is granted to information based on attributes of the subject. Any subject that can prove they possess the necessary attributes to meet all the necessary policies is granted access to the information.

o Access does not require the subject provide their orthonym. Subjects could be anonymous or use pseudonymous.

o The subject is required to bind the attributes to the channel with the relying party to a level of assurance as required by the relying party. If the PDP only requires low assurance, bearer token over SSL would be suitable. If the PDP requires higher assurance, then holder of key tokens over SSL would be suitable.

o 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 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 unforgeable token of compliance with a policy. The token can be used with any resource tagged with the same policy.

o When a subject tries to access the information they must present tokens to the PDP to prove they meet the required policies. It is not required under this model that the content consuming subject necessarily authenticate to the PDP. The access request may simply present a series of attributes to prove the subjects compliance with the policies along with a key such as an X.509 encryption certificate which the PDP can use to protect the DEK.

o This mode has a baseline of a secure transport between the PEP and the PDP. One of the decisions the PDP has to make is the level of assurance on the release of the DEK to the subject. For example the PDP can release a clear text DEK over the secure transport to the PEP. Alternatively the could require the production of a high assurance X.509 encryption certificate as a subject attribute to generate an encrypted DEK.

o When a subject creates information, this model uses roles as a means to mange the set of policies a subject is allowed to assert on information they create. The PAP publish a set of policies associated with each role. Multiple PAP can publish policies for a role. The PIP would provide information on the set of roles a subject is assigned to. The PDP issues a role token to a subject for each role they belong to based on the information provided by the PIP. Each role token contains the aggregation of all the policies from all the PAPs for each role. Each role policy set is a hint to the subject on the specific list of policies to pick from when creating information for each role. The act of aggregation by the PDP allows the role token to contain project wide policies used by all subjects across a collaborative project as well as organization specific policies applicable to the role.

o The role token is used by the subject to authorize the creation of content with specific policies. The PDP 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.

Author Note: Clarify requirements for exchange of KEK

4.2.1 Policy Data Binding

There are three ways to bind policy to data.

o 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 PDP compares the access request to the policy on the data itself.

o By name. 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 PDP finds the policy based on the identifier and then compares the access request to the referenced policy.

o By description. This is where the policy has a target description in terms of characteristics of the sets of data resources the policy applies to. When an access request is made, the set of policies are evaluated at run time to determine the set of policies to apply. For example when you author a XACML policy, you also define a target for the policy. When an access request is made to the data, 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 compares the access request to the identified policy.

The chief strength of binding policy by value is its simplicity. The policy is local to the data can can easily and quickly be read. The chief weakness in binding policy by value is maintaining policy over time. 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 description is it can be applied to data without impacting the storage format. The chief weakness in binding by description is the reliability of the matching. Any matching process must have a false positive and falsie negative rate. This rate has 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 documents and files have a poor set of 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. Its also impossible to detect if a policy is missing from the policy store which again would mean incorrect policy enforcement

This 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 my 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 wit the data..

4.3 Content Creation Workflow

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

(1) The content creation PEP is configured with the set PIP's and PDP's it trusts.
(2) The content creation PEP summits a request to all the trusted PDPs 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 PDP either via front-end (related to the PDP from the PIP via the subject) or back-end (direct exchange between the PDP and hte PIP) processing.

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

Now the PEP is bootstrapped with a list of BCs and for each BC a list of policies associated with each BC. Now the PEP is ready to create content. When the user wants to release protected content, they use the following sequence of events

(i) The user creates the new content

(ii) The user select the appropriate business context for the content, then selects one or more policies applicable to the content

(iii) The PEP encrypts the content with one or more locally generated CEKs

(iv) The PEP submits the CEK, the set of requires policies to be applied and the hash of the encrypted content to the PDP. THe CEK can be a raw key or a CEK key encrypted by a KEK if the user does not want the PDP to have the ability to access the plain text data.

(v) The PDP generates the encrypted metadata which contains the list of policies and the CEKs. The metadata is encrypted by the PDP for itself. The PDP includes a URL for itself and the hash of the protected content as authenticated attributes then signs the encrypted metadata.

(vi) The PDP returns the metadata to the PEP

(vii) The PEP attaches the PDP metadata to the protected content and distributes the content.

4.4 Content Consumption Workflow

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

(A) The PEP 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 PDP

(B) The PEP 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 PEP forwards the signed metadata and requests a read token from the PDP using the location in the authenticated attribute in the metadata

(D) The PDP decrypts the metadata, de-references the policy pointers and determines the set of access rules based on the policy published by the PAP. The PDP then determines the set of subject attributes it needs to evaluate the access rules. The PDP can the use PIP is has relationships with to query attributers about the subject. The list of attributes the PDP is missing is then returned to the PEP

(E) The PEP obtains the missing attributes requested by the PDP and sends them to the PDP

(F) Once the PDP has a complete set of attributes, and the attribute values match those required under the access policy, the PDP releases the CEK to the PEP along with a TTL which defines how long the PEP can use the CEK before it must discard the CEK and reapply for access.

(G) Once the PEP has the CEK it decrypts the information. It caches the CEK until the TTL expires.

4.5 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 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 type types of policy.

4.5.1 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 IdP would be capable of supporting basic policy due to their simplicity and basic attribute set.

4.5.2 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 and can be either primordial or derived attributes or both. Multiple policies have a logical relationship e.g. they can be AND or ORed together. It is not expected that all Plasma clients support advances policy.

5. Message Protection Requirements

5.1. General Requirements

Protected content MUST be where the content is confidential, integrity protected AND provides data origination.

Every authentication has a level of assurance associated with it depending on attributes such as the identity checks made about the subject and the authentication technology used. 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 PEP and PDP, therefore the specifics of how sender or recipient achieve the required level of identity assurance MUST be abstracted from the PDP and PEP 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.

Access to the plain text of the content MUST only be provided after the recipient has either provided suitable valid attributes to the PDP or the PDP was able to find attributes about recipient directly from a PIP, to satisfy the policy as defined by the sender (See section 2.1.1)

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 PDP MUST be abstracted from both the sender and recipients i.e. the PEP 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 PEP 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 PEP 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 PDP 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 SSL. 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 PDP (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 PDP to define when the PEP MUST discard the message CEK and submit a new access request to the PDP. The TTL value MUST be based on the message policy and optional attributes about the content consumer and their environment.

The PDP 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 PDP 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 PDP MUST be capable of generating audit events associated with access to protected content.

5.1.1 Email Specific General Requirements

It MUST be possible for domains to publish keys for boundary inspection agents. This allows senders to pre-authorize these agents 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 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. need rewording to multiple basic policies

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

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

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

Editorial Comments

[anchor21] JLS: Are these really the terms that we want to be using? I normally use data origination rather than authenticated. It would be assumed that the data origination is being attested to by a middle man for a sender w/o signature capability rather than authentication being a correct term.

[anchor22] JLS: We need to talk about what operation you are getting this level of assurance for. and who you are authenticating to.

[anchor23] JLS: Same text should apply for senders?

[anchor24] JLS: What does assurance level mean here? Are we talking about security levels or authentication levels or something else? Are levels required to define a set of requirements. I.e. An assurance level defines: Authentication requirements, confidentiality requirements (other).

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.

[RFC2634] Hoffman, P. Ed., "Enhanced Security Services for S/MIME", RFC 2634, June 1999.

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

[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