Extensible Supply-chain Discovery Service Problem Statement
draft-rezafard-esds-problem-statement-03

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on May 21, 2009.

Copyright Notice

Copyright (C) The IETF Trust (2008).
Abstract

This document discusses the requirements of an application layer protocol for Discovery Services. Discovery Services at its core offers to authenticated and authorized users the means to discover resources of information for a particular identifier. The information resource can be any data source which provides an interface on a network that allows for retrieval of the data. An information resource could publish its network address (reference to the resource) to a Discovery Service coupled with an identifier. Then an authenticated and authorized user could query the Discovery Service with the same identifier to receive reference information to the resources. Interfacing with the resources for actually retrieving the data is out of scope of Discovery Services; the role of Discovery Services is to enable a client to find the network addresses and types (e.g. URLs) of information resources for a particular identifier of interest.

This protocol is applicable to numerous scenarios such as track and trace in trade supply chains, where a number of independent resources may hold information about a particular object and may have an interest in selective sharing of that information, in order to improve the efficiency of the supply chain network.

However, other applications can be envisaged, as a 'bottom-up' or 'grassroots' way for generators of information content to: 1) declare that they have information or commentary on a particular topic or subject (which might be a physical object, geographic location or even an abstract concept) 2) specify a network address through which that content can be retrieved, 3) specify restrictions about the community of clients that are entitled to receive knowledge about the existence of their content or see the link. This approach can be contrasted with the top-down approach of existing web search engines that rely on crawling/spidering of content that must be already posted in the public domain before it can be indexed - and where the link information is generally made publicly available in a manner that does not discriminate between clients on an individual basis.

This document outlines a set of design concerns that an application layer protocol needs to address in order to be widely adopted and deployable on public networks

This document obsoletes "Extensible Supply-chain Discovery Service Problem Statement draft-rezafard-esds-problem-statement-02".

Comments are solicited and should be addressed to the mailing list at esds@ietf.org and/or the author(s).
1. Introduction

As the Internet applications become more and more service oriented and specialized, the demand for collaboration among these applications increases. These applications can collaborate on services or on topics. Open collaboration on services is already taking place via Universal Description Discovery and Integration (UDDI). Selective collaboration on specific topics is the problem domain that Discovery Services aims to address.

A first step towards collaboration among applications on topics would be to use a common identifier for each topic. For example a topic can be something conceptual like the name of a book, or a topic can be a particular physical object with a serialized identifier. The use of common identifiers will enable dispersed applications to inter-communicate with each other. Note that it is not the role of Discovery Services to develop such common identifiers - this is left to existing standards such as URIs and semantic web initiatives such as the Dublin Core vocabulary.

The second step for the applications would be to find other resources of information on a common topic. This is made possible when an application advertizes that it holds information on a topic. This is made possible by Discovery Services since an application can advertize to a Discovery Service that it holds information on a specific topic, by publishing a reference to resources that hold relative information on the topic of interest.

At a conceptual level, Discovery Services can be superficially regarded as being somewhat analogous to a bottom-up ‘search engine’ for an ‘Internet of things’. However, there are some fundamental differences from the paradigm of public web search engines. The serial-level information will usually be protected and will only be made available to authorized organizations, with whom the information provider has an established trust relationship. For this reason, we cannot assume that Discovery Services will be able to automatically ‘spider’ or ‘crawl’ the network to compile an index of links, since the underlying information may not be generally available.

Instead, we assume that each information provider will be able to voluntarily publish a record or registration for a particular common identifier, in order that a link can be established back to them as a potential provider of information. At the same time, also the ‘link’ information will be protected by access control policies defined by each information provider, as well as any policies that are imposed by the operator of a Discovery Service. Unlike most public search engines, there may be little or no ‘link’ information provided to non-authenticated users or directly to the general public without
authentication, although in some specific applications, citizens may be provided with indirect access to traceability information that was gathered via an authorized authenticated actor, such as the retailer or manufacturer.

This approach enables distributed management of collected information and allows each organization to restrict who can access the data; they can specify an access control policy (which is enforced by a Discovery Service) to limit visibility of the ‘link’ information - and additionally, they can specify and enforce an access control policy within their own information resource that holds the more detailed information.

1.1. Introduction: Supply chain applications as one driver

Currently, a number of industry sectors are moving towards improved track and trace systems by the use of Auto-ID technologies (such as Radio-Frequency Identification (RFID)) that improve the ability to automate collection of observations, as well as the use of unique identifiers for each individual object, to improve the granularity of information, so that it is possible to track not only at the granularity of batches or lots - but actually trace the entire life history of an individual object. This is particularly important in the fight against counterfeiting and for assuring the traceability and authenticity of foods, pharmaceuticals and other safety-critical objects, such as aircraft and automotive parts.

As well as automating the collection of such data, companies are also moving towards sharing and exchange of serial-level data (i.e. data about individually identifiable objects), in order to improve the efficiency of supply chains. This involves greater use of machine-to-machine information exchange, using standard interfaces for queries and output formats and requiring less intervention by human operators (e.g. less exchanging of spreadsheets via e-mail or FTP). However, serial-level data is commercially sensitive, since it can reveal information about stock levels and volumes and flows of goods. For this reason, most companies insist on maintaining strict control over who is granted access to their data, including the links to their data repositories or information services. The entire lifecycle information for an individual object therefore remains decentralized and fragmented across multiple actors in the supply chain or extended product lifecycle. Discovery Services will provide secure referral services that can be queried with a unique ID and return to authenticated authorized clients a list of links to multiple information providers, who can then be queried for more detailed information about the individual object.
1.2. What is ESDS?

ESDS is a protocol for infrastructure that enables track and trace applications as well as product lifecycle information systems to find multiple sources of information. Using the links provided by Discovery Services, they can then attempt to gather complete information about an individual physical object.

Several organizations may handle an individual object at different stages of its life - and each of these organizations may collect and store some information about each object. This may include observations of where the object was seen, details of operations performed on the object, as well as measurements (e.g. temperature) of the object or its environment.

Even two objects that are of the same product type and which were created in the same batch or lot may ultimately follow different paths throughout their lives and each might be handled by a completely different set of organizations.

We DO believe that each organization should be able to keep control of the data that they collect and store - and should have the ability to determine how much of this information is made available to others - and with whom. This can be achieved by requiring authentication of clients specifying fine-grained access control policies associated with the roles assigned to those authenticated clients.

We DO NOT propose that ESDS should act as an aggregator of detailed information.

Instead, ESDS is intended as a lightweight referral service, whose sole purpose is to help a client to find one or more sources of detailed information. In practice, a query to an ESDS about an object ID should return to an authenticated client a list of URLs of information resources that hold the detailed information about that particular object. The client will only receive the URLs to information resources that have authorized (e.g. via access control policies) the client to see their URL address for that specific object ID. The client can then contact each of those information resources in order to request more detailed information. This process is OUTSIDE the scope of ESDS and may be subject to further authentication and authorization checks bilaterally between the client and each individual information resource.

1.3. Terminology

There are three actors in this problem:
Client: refers to an actor that seeks (sources of) information about an individual object (or topic).
Resource: refers to an actor that holds information about an individual object (or topic) and may choose to share (some of) that information with a client (under certain conditions).
ESDS: refers to an intermediary lookup service protocol that enables a client to obtain referral information to one or more resources.
Publish: refers to the act of writing to an intermediary lookup service, Resource reference (address) information coupled with an object (or topic) identifier. The publishing actor, as the information owner, could define access control rules to restrict which clients (or client roles) are allowed to view the Resource reference (address) for a particular object (or topic).

What emerges from combinations of these actors and their relationships are:
Organization (Partner): refers to a business entity, which can be comprised of client actors or Resource actors.
Community (Supply chain): refers to a group of member organizations (partners) that may be willing to share some information among themselves.

Note that not all members of the community will share exactly the same access privileges; there can be different bilateral access privileges within a community.

Each organization can define rules that restrict which clients shall be allowed to see the address of its resource as a potentially relevant resource for a particular object identifier (or topic). Note that in some situations, the overall access control policy may be the result of a combination of rules defined by one or more organizations, as well as rules defined at a community level.

Access privileges as defined by a community and its member organizations are only applicable to the information published by ESDS. Each Resource must take its own security measures to protect its data (hosted at the published network address) from unprivileged clients.

1.4. Problem Statement

Information sharing between organizations is essential for all modern networked communities. There is a strict set of business and technical requirements that must be observed in order to enable open sharing of information between organizations.

There are resources and clients in each community. Each resource wants to advertise its knowledge (data) to interested and authorized
clients. Only deploying private networks to connect whole communities is no longer feasible, as the more efficient and economical means of connecting resources and clients is to use public network infrastructure. Currently, many industrial inter-organizational track-and-trace initiatives are deployed on public networks and the information is routed through the Internet. Because of this trend, it is highly desirable that a draft protocol should be developed. An IETF process consensus would ensure that the adopted protocol conforms to the Internet’s operational needs.

A key principle of information sharing within a community is that data ownership must be respected. This means that each organization can collect information within their own systems and is not required to route that information to any other organizations. However, they can choose to share selected information with other trusted organizations. This in turn means that the information about any individual object may be held by multiple resources in a community. Therefore a mechanism is needed in order to facilitate the gathering of this information.

A key requirement is that information accessibility is fully controllable. The confidential information must be secure and hidden from unauthorized clients and only visible to the intended clients in a community. In order to establish trust and reputation with the participating resources, unauthorized clients need to be barred from viewing, eavesdropping, or deducing information or even being aware of the existence of another organization’s involvement with a specific identifier, if the client is not authorized to be aware of this involvement.

There must be a common set of data that all resources will offer (only to authorized clients). This set must be defined and agreed upon in the final protocol. This set may include answers to who, what, when, where, and why as well as links to back-end systems. An essential feature is for the protocol to be general-purpose and extensible to support multiple applications and uses beyond the needs of trade supply chains that motivated its initial development.

Defining a bootstrapping procedure is a necessity when designing a global and autonomous network of systems. Bootstrapping procedure would be the process of locating on the network, the appropriate Discovery Service(s) for a specific identifier. Currently there are deployed supply chain systems that bootstrap via Internet’s DNS roots. One example of this is the EPCglobal Object Name Service [epc-object-naming-services] (ONS). While ONS enables an economical means of bootstrapping, improper implementations may increase the traffic on DNS roots exponentially. Since data will be routed through the Internet, the bootstrapping procedure needs to be
1.5.  Description

An Extensible Supply-chain Discovery Service (ESDS) provides a mechanism to locate one or more resources of information about a specific topic such as an individual physical object. In a trade supply chain resources of information may include the original manufacturer or supplier of the object, as well as other organizations who have handled the object at some point during its lifecycle (including repair and maintenance organizations) and even organizations (such as customs or insurance agencies) who might hold information records related to the object, even though they may have never had physical custody of the object.

This document defines the Problem Statement, Objectives and Technical Challenges related to the application layer protocol currently proposed as Extensible Supply-chain Discovery Services (ESDS). ESDS enables the publishing and querying of referral links to information services that historical events related to specific objects with attached object identifiers. The interface enables disparate applications to discovery resources of information and compile a consolidated view for an identifier. Primarily, ESDS provides referral services in a loosely coupled mechanism with granular security that enables selective visibility.
2. Internet Concerns

2.1. Public Networks and Tree-Walking Concerns

Currently, another standards body, EPCglobal, has issued a related standard referred to as ONS [epc-object-naming-services]. ONS is effectively an extended version of DNS that does not benefit from the IETF review process and, in its design, necessitates increased tree-walking. ONS specifies a reverse mapping of the EPCglobal "SGTIN" (one type of supply chain object identifier) as a hostname and allows for a reference (i.e., URL) to a manufacturer’s relevant back-end system (using a NAPTR record). The SGTIN identifier is comprised of an EPC Manager Number, an Object Class, and a Serial Number. The ONS lookup excludes the Serial Number portion of the SGTIN. However, the ONS specification has already been extended in industry pilot projects to include the Serial Number, as this enables item level lookups for tagged objects in the supply chain. Using serial level lookups, ONS could be used to indicate point to point referrals through the passage of a relevant identifier throughout the supply chain. This would require successive updates to the hierarchal ONS to indicate incremental supply chain member referrals, reducing the effectiveness of caching. Alternatively, the ONS could provide a single, static point of referral to the first or initiating supply chain member. However, even a relatively static entry, which only refers to the point of origin within the supply chain, would drive the number of public zone entries to extremely large numbers if an individual record was created for each serial number. The number of records could exceed the current IPv4 address space by several orders of magnitude. Also, since many of the tagged physical objects would neither require nor provide network connectivity, utilizing IPv6 for such objects would waste this finite address space resource unnecessarily.

One common suggestion to manage this problem is to have multiple alternate ONS roots which can be managed for separate and unrelated supply chains and/or regions. However, since there is nothing to prevent ONS from operating in the existing Internet root hierarchy, even alternate ONS providers can opt to drive traffic to the existing Internet root servers, rather than operate their own ONS roots. There has already been a pilot ONS implementation under the .aero zone (sgtin.id.ons.autoid.aero) where this phenomenon may already be observed.

It should be noted that ONS 1.0 [epc-object-naming-services] currently only specifies how to perform a lookup for a GTIN class identifier, which represents a product type. It does not currently specify that ONS should be used for lookup of fully serialized identifiers. While the number of GTIN codes may be a manageable
number, the potential number of serialized SGTIN identifiers is much larger. Already in EPCglobal Tag Data Standards, the structure of SGTIN identifiers are defined such that it is theoretically possible to create in excess of $10^{38}$ unique SGTIN identifiers for each GTIN code that is currently resolvable using ONS 1.0. Already there are some companies who create over a billion objects per year for a single GTIN class.

It was therefore a very wise decision of EPCglobal not to propose that ONS 1.0 should store a DNS record for each unique EPC, since the number of resulting DNS records could clearly overwhelm the DNS infrastructure.

However, there is currently a missing element in the EPCglobal Network architecture, namely a search service that provides authorized authenticated clients with links to the multiple organizations who hold information about the unique life history of an individual object. This missing element is essential for enabling robust track and trace applications. Because this element (‘Discovery Services’) has not yet been defined, there is a danger that some organizations may misuse the ONS design and begin creating DNS records for each fully serialized object. There is already one industry pilot where this is happening. It is therefore essential that rapid progress should be made towards a protocol for Discovery Services that avoids placing such a burden on the DNS infrastructure.

To summarize, there are two variations to the ONS specifications that can be observed in current deployments. The first variation is an extension to the ONS structured hostname to include serialized level data (serialized SGTIN) instead of just the specified class data (GTIN class). The second variance involves the tendency of the implementers to utilize the existing Internet root DNS servers instead of operating alternate ONS roots. These variances could adversely affect the stability of the public network infrastructure. The IESG and IAB overview procedures at the IETF will ensure that the protocol and architecture proposed by the ESDS Work Group will be mindful of such deviations and take counter steps to prevent it.

2.2. Bootstrapping Concerns

The bootstrapping process enables an interested and authorized client to locate an object’s Discovery Service server using only the information provided by the object identifier. Unlike DNS, where there is a known set of root servers, ESDS will have numerous roots for various communities (supply chains) operating globally. This, in turn, complicates the bootstrapping process.

A common bootstrapping scenario would be exception handling in a
trade supply chain. For example, if an object is mis-delivered, a recipient who has no pre-existing relationship to the trade supply chain, needs to obtain object ownership information and its corresponding Discovery Service server. ESDS design must aim to accommodate this scenario while respecting privacy and security considerations.

It has been suggested that ONS could be used for the bootstrapping process. However, ONS-URN encoding and decoding scheme is only defined for EPC identifiers and ESDS is being architected to support non-EPC identifiers as well. Also, ONS’s hierarchical identifiers have raised privacy and security concerns by multiple participants in some Discovery Service scenarios. While ONS can technically support multiple identifier schemes, with multiple issuing authorities, its hierarchical operation does depend on structured identifiers (for example, ManagerNumber.ObjectType.SerialNumber). These identifiers display information about the object they are representing, such as type and manufacturer and as a result could compromise the privacy of an individual using the identifier. Additionally, ONS has no authentication nor access control restricting who can query it. ESDS must be designed for serial-level lookups and must support unstructured opaque identifiers to use as lookup keys within an ESDS service. Because serial-level lookup information could potentially be subject to data mining by competitors, to reveal information about volumes and flows of goods, ESDS should fully support authentication and serial-level access controls to address concerns about privacy and confidentiality of data.

To facilitate bootstrapping and exception handling scenarios ESDS design could investigate the use of peer-to-peer lookup protocols (such as JXTA [jxta-protocols]) as an alternative to hierarchical lookup protocols such as DNS and ONS. This would keep the ESDS traffic flat and avoid walking up the Internet root hierarchy. A major concern is that the bootstrapping design must not implicitly establish monopolies in the long run. The IETF process will ensure that the resulting protocol design addresses the concerns of both the end-users and the network operators.
3. Other Standards

Discovery Services address a problem domain that is not fully covered by any existing protocols. Discovery Services enables authorized and authenticated users, to gather and retrieve links to resources of information about a topic or an object, this information is typically fragmented across multiple organizations and domains. While there are related standards such as ONS, EPCIS, and UDDI, each of these address different problem statements and domains as follows:

3.1. Object Naming Service (ONS)

ONS is a mechanism that leverages the Domain Name System (DNS) to locate sources of authoritative information and related services about a product when queried using a hostname derived from the Electronic Product Code (EPC). Its query interface as currently defined lacks access controls, authorization and authentication and this security deficiency makes it unsuitable for a Discovery Service handling commercially sensitive serial-level information. The volumes of serial-level records generated by the supply chains would place an unreasonable burden on the DNS roots if ONS were mis-used for holding ONS records for each serial number. However, Discovery Service can be used in collaboration with ONS to meet specific business requirements. [epc-object-naming-services]

3.2. EPC Information Services (EPCIS)

EPCIS targets the sharing of data within an established network of trust and relationships. Typically, a company can provide an EPCIS query interface to allow trusted organizations to retrieve some of its data. It is not the role of EPCIS to enable robust linking to information across the entire supply chain or product lifecycle, nor to perform recursive ‘proxy queries’ to retrieve data from other parties, nor to facilitate exception handling and object bootstrapping across the whole supply chain. [epc-information-services]

3.3. Universal Description Discovery and Integration (UDDI)

UDDI is a standard for description and discovery of services and the functionality that they offer. In contrast with UDDI, ESDS is concerned with locating services that provide information about a specific physical object; the ID of a physical object is the primary lookup key for ESDS, whereas UDDI is primarily queried for a particular type of service functionality desired. Unlike UDDI registries, the data held within some Discovery Services may be much more commercially sensitive. For example a supply chain Discovery Service, could reveal volumes and flows of goods in supply chains;
from a security and scalability perspective, ESDS therefore attempts to solve a more challenging problem than the problem that UDDI already solves.

3.4. Interface for Metadata Access Point (IF-MAP)

The protocol, IF-MAP, defines a simple publish/search/subscribe interface to the MAP. The MAP is a common metadata database. Within the MAP, metadata automatically aggregates with respect to common unique keys (i.e. identifiers).

While IF-MAP’s initial use-cases are in coordinated security, it was designed to meet other coordinated computing use-cases and may be appropriate for ESDS. However, IF-MAP does not provide historical information (according to section 2.2 of the IF-MAP protocol). A major use for ESDS is to maintain historical link information between the ID of a physical object (or a discussion topic or keyword) and multiple entities who have collected/generated some information about that individual object (or discussion topic or keyword) at previous times.

IF-MAP is concerned with building a graph to build associations between different identifiers and pieces of meta-data. With ESDS, the internal data model is concerned primarily with time stamped events containing associations between:

- ID of object (or keyword/topic)
- URL of resource that holds more detailed information
- [ timestamps to help with sorting into chronological order and completeness of result sets ]
- [ serviceType indicating what kind of resource to expect at the URL ]
- [ other metadata that provides additional context or is helpful for defining access control policies ]

Furthermore, ESDS may also need to handle an additional data model of the access control policies it is expected to enforce by the publishers of information.

Authorization model is outside the scope of IF-MAP (according to the footnote 3 on p49 of the IF-MAP protocol), "3 An IF-MAP server implementation may provide an authorization model, which protects data published by one IF-MAP client from being visible to another IF-MAP client. The specifics of such an authorization model are outside the scope of this specification."

The ability to enforce an authorization model based on policies defined by individual publishers is a key functional requirement for
ESDS. Without this, many end-users will not publish data to it if it cannot be protected (hidden) from other end-users with whom they do not wish to share this information.

In IF-MAP, the link is an un-named bi-directional binding relationship between two identifiers. (S.2.6.2 of IF-MAP protocol). In ESDS, there will be 1-many links: One entity (or information provider) may hold information about many different individual physical objects (or many different topics / keywords). One individual physical object (or topic/keyword) may be linked [via ESDS] to multiple entities that hold more detailed information fragments about it.

However, in supply chain applications, resource owners may prefer that the link cannot be discovered the other way, e.g. by querying a Discovery Service with the address of a resource in order to retrieve a list of IDs of all the objects it claims to hold information about. ESDS will need to provide for such directionality restrictions on how links can be discovered.
4. Related Standards Development Organizations

Discovery Services is an interesting problem with significant challenges for both end-users and service providers. To tackle these challenges, the problem must be approached from both viewpoints, to ensure that the final adopted solution meets the needs of all organizations involved.

The EPCglobal Data Discovery Joint Requirements Group (JRG) is in the process of gathering use cases and user requirements in order to define a Discovery Services standard which can meet the needs of supply chain participants and end-users. In a parallel and complementary effort, ESDS is addressing the technical challenges of Discovery Services, such as its deployment on public network infrastructures. The final goal is to share the lessons learned in a co-operative effort to forge a single standard protocol.

ESDS is chartered to produce draft proposals for the Discovery Service protocol. Similar to the manner in which the ONS standard adopted the approach and architecture developed by the IETF DNS Work Group, it is the intention of the ESDS activity that any draft protocols developed will be useful to EPCglobal and will help to accelerate the development of an EPCglobal standard for Discovery Services.
5. Objectives

This section outlines the objectives of the ESDS protocol. In an effort to convey the goal of each objective, possible solutions are provided in order to trigger discussion on the subject matter.

5.1. Flow Diagrams

Draft flow diagrams for a directory based model of sharing reference information.

Diagram showing the publish of reference information by multiple organizations:

```
+-------------------------------------------------------------------- +
|                             ESDS Server                             |
|                                                                     |
|                             Community_1                             |
+---------------------------------------------------------------------+

^                       ^                       ^
|ESDS Record Published  |ESDS Record Published  |ESDS Record Published  |
|for an individual      |for an individual      |for an individual      |
|object X               |object X               |object X               |
|(containing reference  |(containing reference  |(containing reference  |
|to sources             |to sources             |to sources             |
|of information)        |of information)        |of information)        |
|Subject to access      |Subject to access      |Subject to access      |
|control privileges     |control privileges     |control privileges     |

+------------------+    +------------------+    +------------------+
|  ESDS            |    |  ESDS            |    |  ESDS            |
|  Publish         |    |  Publish         |    |  Publish         |
|  Client          |    |  Client          |    |  Client          |
|                  |    |                  |    |                  |
|  Organization_1  |    |  Organization_2  |    |  Organization_3  |
+------------------+    +------------------+    +------------------+
```

Figure 1
Diagram showing the query client sending a request to a virtual community to retrieve list of resources for an object or topic.

Figure 2

5.2. Security

Since ESDS needs to be deployable over a public network infrastructure, issues of security and privacy are of heightened importance. Clients must be authenticated to prevent theft of information and resources must be authenticated to ensure integrity of information. The information routed over the Internet must be encrypted. It is suggested that the security model should be based on open standards, trusted by the industry.

The ESDS protocol should be decoupled from the security layer and not have embedded components specific to certain security protocol implementations. This will enable ESDS implementations to respond quickly to changes in the ever advancing security layer protocols.

5.2.1. Access Control

An ESDS should provide a resource with the ability to protect its link information in order to retain control over which clients are
allowed to read this information. Such rules may be expressed in the form of access control policies which are evaluated against the client’s authentication and authorization credentials and its role in relation to the provider of the resource, as well as other criteria such as the time elapsed since the link was published.

ESDS needs to define the granularity of information at which access control policies are specified and enforced. Whether they apply to individual events as atomic indivisible entities, or they can specify which data fields to include or omit.

A situation may arise where a client and the provider of a resource have no existing trust relationship with each other. An ESDS should allow a resource to specify multiple levels of ‘visibility’ to such a client, so that the resource either remains completely ‘silent’ or ‘invisible’, or so that an opaque handle is visible. The opaque handle should not reveal the identity of the resource in any way, but can be used to facilitate the initial negotiation between a client and a resource, such as forwarding a number of messages from the client to the provider of the resource, provided that the client specifies the handle in their message, and provided that an ESDS or associated service incorporates such a mechanism. To protect the resource provider and ESDS from additional burden, such a facility may be limited in the number of messages which are forwarded and the time window following the client’s query during which forwarding of messages is offered.

It is recommended that ESDS follows existing standards for defining access control policies, such as XACML. ESDS should study for the BRIDGE project work package 4 (Security) on recommendations for access control interface and protocols [bridge-network-confidentiality].

5.3. Publish

An ESDS must provide a mechanism whereby a resource can dynamically establish a link for an individual object (or list or range or class of objects) that points from the ESDS to the resource. The link can be expressed as a string or URI and it is helpful if this is accompanied by an indication of the type of service which can be accessed via the link (in order to distinguish between web pages, web services, EPC Information Services (EPCIS) and other communication mechanisms which might even include phone or fax numbers).

5.4. Query

An ESDS should provide a mechanism whereby a client can query the ESDS in order to retrieve a list of links to one or more resources.
Queries may be one-time queries or standing queries, and an ESDS may support either type of query, or both.

The query interface needs to define the criteria fields as well as acceptable criteria values, such as regular expressions or wildcards.

The ESDS Work Group should decide on whether to support multilayer information visibility. A query with limited access can be informed of the existence of information for a particular object, but to view the actual information, full access privileges would be required. This has particular implications for peer-to-peer searching across multiple Discovery Services. Another scenario to consider is whether visibility into the information in the extended fields of a published event, should be controlled independently of event information? (i.e. should access control policies be sufficiently expressive that they can grant or deny access to individual data fields within events, rather than simply granting or denying access to events) This has particular implications for tracking aggregation and disaggregation events and visibility into the entire lifecycle of an object.

5.5. Relabel/Aggregation/Disaggregation

To provide complete and resilient visibility across the entire life of an object or a product, it is required that changes in identifiers are maintained in the Discovery Service layer.

One scenario for relabeling can be a Discovery Service for maintenance of installed software packages. An organization with a multitude of machines, deployed anywhere on the Internet could use DS to notify machines about the availability and location of (secure and tested) updates to packages. Each machine can submit standing-queries to the DS, for all the packages that are installed on it locally. The organization operators would test and verify every new release of a package. Once a package is approved, it is placed on trusted file servers and the link to those files (and any mirror links) is published to DS along with the package name. This would result in all the machines with standing-queries to be notified of the new downloadable and verified package. Package name change is a common phenomena, specially when the version number is appended to the package name (e.g. wxWidgets libraries). So ESDS needs to provide relabeling facilities in order to be comprehensive enough to address a wide range of problem domains.

Another scenario requiring the capture of aggregation and decomposition events is the regulation of the European parliament with respect to the cold supply chain regarding the tracking of fresh meat. For example, as a cow moves from a farm into the supply chain where it is finally sold as beef to a consumer at a retail store, it
is vital that relabelling and disaggregation events can be stored within a Discovery Service in order to enable querying information about the entire lifecycle of the cow "from farm to fork" While having access control over visibility of these kind of events is desired, it is recommended that Discovery Services should not attempt to interpret or validate these events. It has been suggested that these complexities be excluded from the core of ESDS, instead facilitating them via the extensions mechanism and/or leaving such interpretation, validation and analysis to client applications that query ESDS. This approach must also give consideration to providing access control to the extension information.

5.6. Multiple Organizations

It is the intention of ESDS to provide authorized clients with the link referral information necessary to enable client applications to gather information from multiple organizations covering an object’s or a topic’s lifespan. For a product this could include link data from the design phase, through manufacture, distribution and retail, usage period (including service, repair, maintenance, overhaul) and even end-of-life phase (including collection, sorting, remanufacturing and recycling). So it could be desirable to record the current phase. Also there may exist lifecycle steps within a particular phase, a higher perspective or meta view could show that entire phase as a single step in the life of the product. For example, a product can be in a Manufacturing supply chain, then move into a Logistics supply chain, followed by a Service supply chain. Each of these particular supply chains may have their own set of lifecycle steps, but the entire lifecycle is spread across all three supply chains. ESDS should facilitate the linking of these supply chains and phases together to enable the capture of the entire lifecycle of the object.

5.7. Identifier Agnostic

To enable the grouping of information belonging to the same object, each object needs to be uniquely identifiable. Information about the object is held by a number of independent organizations that agree to use a shared identifier schema. These identifier schemas are defined and distributed by numbering authorities in each industry sector. Some numbering authorizes guarantee global uniqueness of the issued identifiers, but this is not the case all the time. Especially when there exists issued legacy numbers currently in use within the domain of a numbering authority. Also there are cases were numbering authorities for different domains have collisions in their space of issued identifiers (e.g. Animal Identification Number (AIN), National Stock Numbers (NSN), and the International Mobile Equipment Identity (IMEI) are all 15 digit identities and there are collisions
in the space of issued identities.)

There are procedures for converting identifiers to URIs which can also assure that the URI is globally unique. However although global uniqueness of identifiers is highly desirable for use within a co-operating federation of Discovery Services, it is not a requirement for each individual instance of a Discovery Service, and it would be out of scope for a Discovery Service to verify global uniqueness for identifiers used.

To define a protocol which is applicable to all scenarios, it is essential that the protocol supports various numbering authorities. The ESDS protocol should have a broad scope and support multiple identifier schemes including schemas with non-unique identifiers. The intelligent handling of identifier associated with an object is out of scope of Discovery Services and can be handled accordingly by ESDS query applications. It must also be identifier agnostic in order for Discovery Services to be embraced by all potential adopters.

5.7.1. Reusing Identifiers

There are scenarios where the same object (and its identifier) is used within a Discovery Service but in different sessions or phases. It could be the case that the authorization and access control policies differ based on attributes such as timestamps, or metadata. An identifier reuse scenario can be for reusable assets and other returnable transport items that may circulate among multiple organizations - and in many situations, it may be more economical and more feasible to tag the reusable asset than to tag its contents. This is particularly true when attaching not only a unique ID tag but also environmental sensors to the reusable asset.

Company A that provides and manages reusable assets can provide an additional service to its customers, such as company B (which borrows or leases an asset from company A), by allowing company B to access data about the asset while it carries the products from company B. At the same time, there are benefits to company A in being able to track its own assets, to balance supply and demand of assets and detect when assets are not being returned correctly, e.g. for cleaning and repair.

However, it is also very important to protect the confidentiality of the data for company B from its competitor, company C, who may happen to use the same asset at a later time.

One possible solution to this is for company A to dynamically manage access permissions to data about the asset and its contents using
time-based windows that provide its customers with visibility only to data about the asset during the time periods when it is associated with their products or their lease/use of the asset.

This could be achieved by appending additional time-based constraint qualifications to access control policies.

For example, company B could be granted visibility to track and trace data about the asset for events that occur after they take custody of the asset and before they relinquish custody of the asset. Company C might be granted access for a different time window. There should be no overlap between the time windows granted to competing companies, in order to protect the confidentiality of their data.

5.8. Extensible

The event data needs to accommodate various scenarios and their business requirements. Therefore it must be an extensible protocol which enables the collaborating organizations to communicate messages specific to their industry and terminology.

An ESDS protocol should enable the sharing of information without involving the business rules of various scenarios. This will keep the interface clean and uniform across all scenarios and ESDS services.

ESDS should give consideration to approaches for dealing with access control and visibility of extended information and protocols. One approach could be to provide one level of access to all the event information, whereby if a client can access an event information, it can retrieve the extended information. Another approach could be to provide multilayer visibility, where only clients with proper privileges can retrieve the extended information.

5.9. Policies

The Retention policy for data records in ESDS would vary based on the business and regulatory requirements. Some scenarios would need retention of only few days and some would require retention for 30 years. Some other policies with similar parameters but different interpretation are the archiving policies, purging policies, and audit policies.

5.10. Stale Links

There are situations where the link information (such as a URL) that was originally specified is no longer effective (e.g. because the provider of the resource has not taken care to maintain redirection
from the original link address to the new location of the resource when restructuring a website or moving to a different domain name). In such situations, it is desirable that a Discovery Service could provide a client with link information that is current and effective. For audit purposes, it may also be necessary for an ESDS to be able to retain and reconstruct the original (or previous) link information, even if it is no longer effective.

This may also imply that an ESDS should allow a resource provider to loosely couple the link record with the current link address, to ease migration of multiple records to a new resource address, while still providing a mechanism to recover a full audit trail of such changes of link addresses. This is particularly important when the link records are digitally signed and we wish to avoid having to "void" such records merely because they embedded a URL which is no longer in service.

5.11. Peer-to-Peer Communications

Architecting a bootstrapping policy will involve defining a protocol for peer-to-peer communication between Discovery Services (DS). The ultimate goal is to locate the target DS without dependence on hierarchical information and services such as ONS or the underlying DNS. To facilitate the peer-to-peer communication it is recommended that existing protocols and proven architectures such as JXTA [jxta-protocols] be utilized. Another candidate architecture is ECRIT’s approach to locating authoritative sources of information in a peer-to-peer network [ecrit-mapping-arc].

One point to keep in mind with DSs is that finding one target DS does not mean that the search is over. One object identifier could be acceptable to many DSs; therefore a search cannot stop once a DS is located but needs to propagate to all peers.

Care and consideration must be given to privacy and the sensitivity of information at the DS nodes. It is also critical that steps are taken to protect against increased vulnerabilities that the extra exposure that P2P brings, such as denial-of-service attacks, or data mining tricks.

Another challenge is when the target DS is successfully located. How much information on an object inquiry can that DS share with the peer DS, without jeopardizing security and privacy? Should the DSs facilitate the peer-to-peer access negotiation process or should it be done by the Certificate Authority?

A sample scenario would be a query client negotiating for access to information from an organization where the information-providing
organization does not have an established business (trust) relationship with the client and therefore chooses initially not to reveal its identity to the client. In this case, it is possible to use the opaque ‘node ref’ concept and perhaps each object can have such a ‘node ref’ handle to deal with such scenarios [bridge-discovery-services-design]. Considerations must be given to protection against data mining to discovery whether a serial number exists or has been issued.
6. Technical Challenges

ESDS should be purely an application layer protocol disassociated from implementation specifics and challenges. Below are some of the technical challenges that certain business requirements demand. However ESDS is not chartered to address these implementation challenges.

6.1. Auditability

Based on some business and regulatory requirements, auditing capabilities should be facilitated by ESDS to provide accountability if and when something goes wrong. With an auditing mechanism, record data can be tracked and the person responsible identified, thus a series of data records can be reconstructed at a later date, allowing the Discovery Service operators to prove who was responsible for which data records [bridge-security-analysis].

6.2. Responsiveness

ESDS implementations will need to be designed to accept updates in a close to real time basis from multiple providers of information across the globe. Because they store serial-level records, they will need to be sufficiently scalable to store large volumes of data, possibly up to trillions of records per year.

6.3. Availability

ESDS availability requirements would vary from business case to business case. Research by BRIDGE shows that the uptime requirements for some supply chain business cases are 99.99% year round.

It is expected that the ESDS instances will be available over a shared network that exposes them to the effects of network attacks. Furthermore, in many cases it is expected that these components should be globally reachable from the Internet and not hosted on a secure private network. Such components are also built using commonly available Operating Systems and middleware (e.g. Application Servers). Thus they are also subject to any vulnerabilities of these supporting systems.

A major security issue for shared services such as the ESDS is service availability. In particular if services are considered to be vital for businesses and organizations (e.g. "pharma e-Pedigree" or "product-authentication") ESDS needs to be able to guarantee minimal amount of service downtime due to security vulnerabilities and attacks [bridge-security-analysis].
7. Related Activities

Currently, the standards organization EPCglobal, and EU research projects BRIDGE, SMART, and PROMISE are looking into the global supply chain track-and-trace challenge. As part of their research, each of them has identified the need for Discovery Services to link resources and clients in the supply chains.

EPCglobal is beginning to gather user requirements for Discovery Services. However, EPCglobal is a paid membership organization focused on serving their own subscriber community. Although the final ratified EPCglobal standards are open and freely available to download, the subscription fees for joining the EPCglobal community may deter a significant proportion of the end-user community from directly participating in the development of their global standards.

IETF would be the ideal body to oversee the development of this protocol because of its deep knowledge of Internet infrastructure and experience with development of application layer protocols.

An IETF working group would be an inviting and open community which facilitates contribution and participation of all interested parties involved in the global supply chain as well as other potential industries with organizations enthusiastic about the benefits of collaboration with each other. The output would be released as a freely available RFC in the public domain. ESDS will make best efforts to ensure good communication with complementary activities at EPCglobal, in order to ensure that the output of ESDS is as relevant and useful as possible to a future EPCglobal standard for Discovery Services.
8. Acknowledgements

This document and the process of authoring was made possible by the excellent xml2rfc tool [RFC2629]. Credit must also be given to Scott Hollenbeck author of [RFC4930] for the organization and grouping of RFC sections.
9. Contributors

Dr. Mark Harrison  
Senior Research Associate, Distributed Information and Automation Laboratory  
Director, Auto-ID Labs at Cambridge  
Institute for Manufacturing  
University of Cambridge  
Mill Lane  
Cambridge, UK  
CB2 1RX

Phone: +44-(0)1223-338178  
Email: mark.harrison@cantab.net

Michael Young  
Senior Director, Technology  
Afilias Canada  
204-4141 Yonge Street  
Toronto, ON M2P 2A8  
CA

Phone: +1-416-646-3304  
Email: myoung@ca.afilias.info
10. IANA Considerations

This document has no actions for IANA.
11. Security Considerations

This document is a problem statement that does not by itself introduce any security issues.
12. References

12.1. Normative References


12.2. Informative References


[epcglobal-tds] EPCglobal Tag Data Standard Work Group, "EPC Tag Data

[jxta-protocols]
Author’s Address

Ali Rezafard
Afilias Canada
204-4141 Yonge Street
Toronto, ON M2P 2A8
CA

Phone: +1-416-646-3304
Email: arezafar@ca.afilias.info