Business Document Exchange Architecture - BEDA

This document is modified on basis of "The PEDRI - an executive summary" written for the PEPPOL project.

Version 0.3

March 30th 2011

Index

Summary

An instances of a Business Document Exchange Architecture (BDEA) is called a Business Document Exchange Infrastructure (BDX-Infrastructure). A BDX-Infrastructure lets service providers exchange business documents on behalf of their customers (typically private companies or public sector organizations). This model is referred to at the 4-corner model because it lets business partners (two or the four corners) exchange business documents via independent service providers (the remaining two corners). The heart of the BDX-Architecture is a simple, robust, scalable and secure addressing and discovery mechanism. It supports service providers and their customers in discovering business partners and their capabilities.

*The vision of **BEDA** is that:*

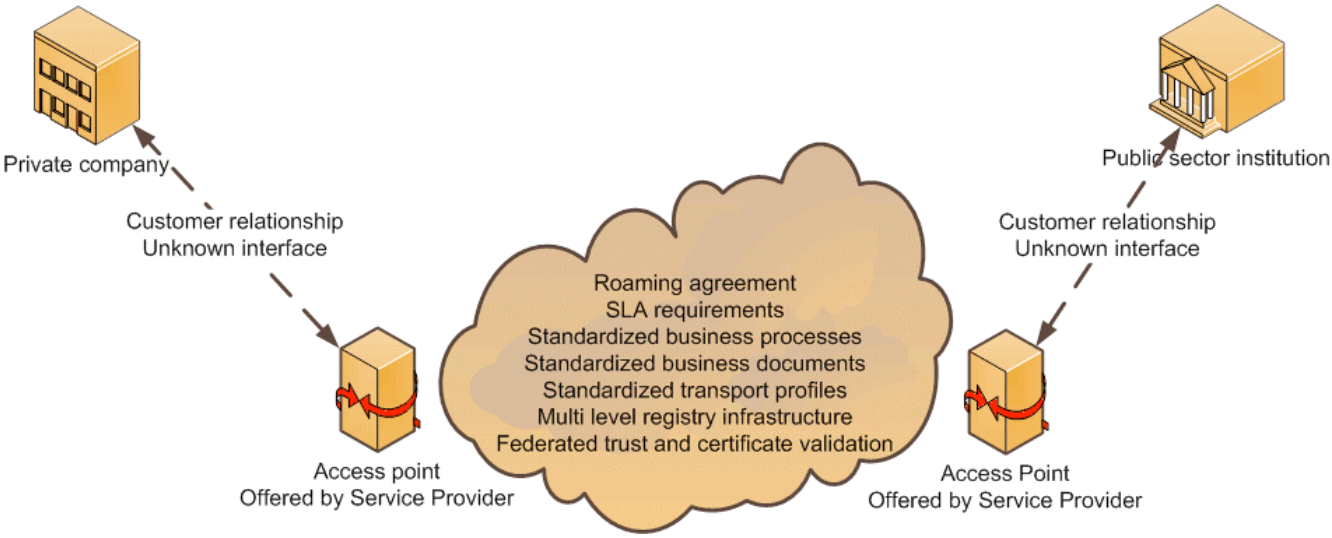
Global exchange of business documents between any two business partners should be as easy as exchanging emails.

Purpose

The purpose of this document is to give an overview of the intended architecture of a BDX-Infrastructure. And furthermore to engage the industry in discussions that can lead to a common understanding and final architecture.

Introduction

It is widely recognized that the 4-corner model is the right model for secure and reliable exchange of business documents between business partners.



Characteristics (simplified)_

* Agreed upon standards for transport

    - - open or proprietary

* Perhaps - agreed upon standards for content

* Freedom to choose service provider

A complete architecture for instances of the 4-corner model must address interoperability at several levels:

1. Organizational interoperability: Alignment of business processes and business models
2. Legal interoperability: Alignment of legislative frameworks ensures a level legal playing field
3. Semantic interoperability: The content and semantic meaning of business documents
4. Technical interoperability: Transport protocols, security, trust

The biggest obstacles are that are hindering that the 4-corner model in spreading to a comprehensive global infrastructure are:

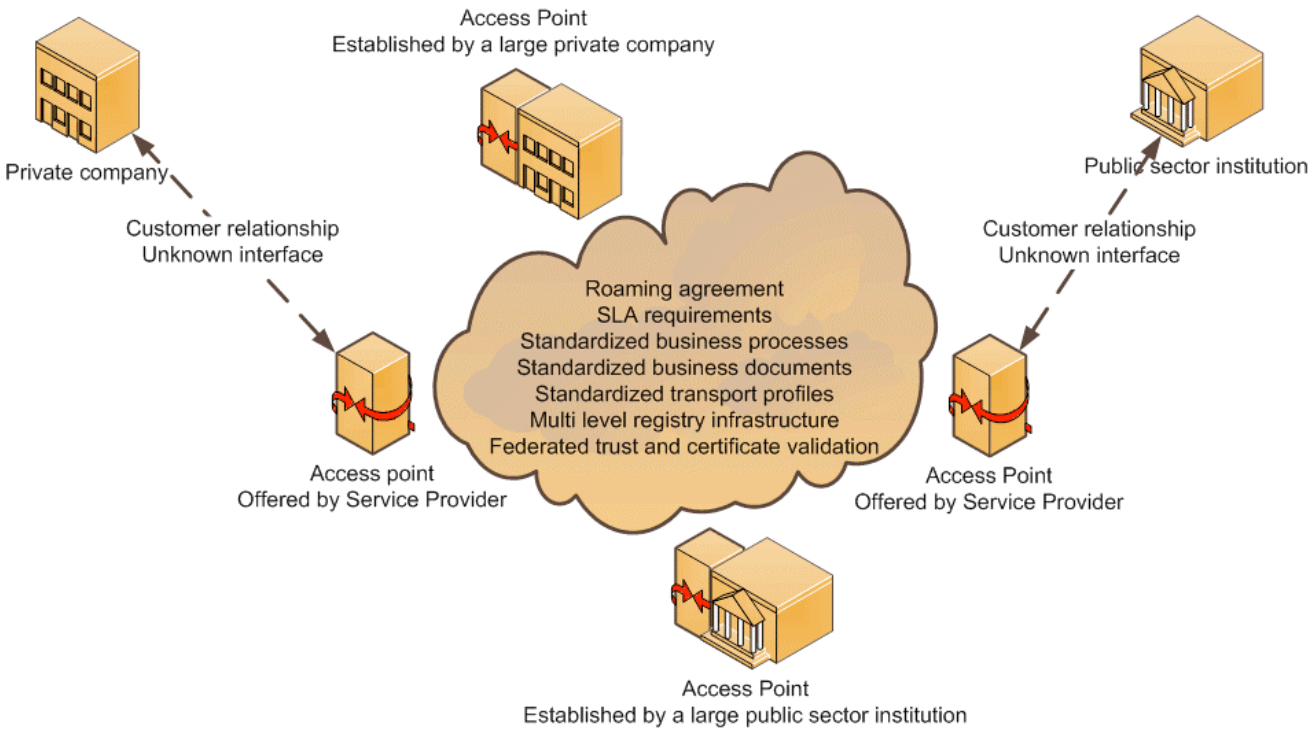
Lack of interoperable business models. Interconnection between two service providers may ruin the market for other service providers.

Establishing trust between service providers
- Addressing recipients and their capabilities between service providers

[more discussion here]

Architecture

A BDX-Infrastructure is a network of connected service providers and their customers. A private company or a public sector institution can be connected to a BDX-infrastructure through an existing service provider or by acting as their own service provider. This section gives an overview of the various components.



Overview

The BDX Architecture describes how a group of Service Providers can exchange business documents on behalf or their customers in a 4-corner model. The BDX Architecture is basically a federated architecture where each Service Provider is responsible for their own services. It is intentionally designed to be very robust and with very little centralized infrastructure.

A ***BDX Infrastructure Instance*** is a practical application of the BDX Architecture between a group of Service Providers.

The architecture consists of two major orthogonal components:

- An ***Addressing Infrastructure*** which allows Service Providers to locate receivers of business documents belonging to other Service Providers and to discover the capabilities of these receivers.
- A ***Transport Infrastructure*** which allows Service Providers to exchange business documents using one or more transport protocols.

The Addressing Infrastructure consists of two distinct components:

- The ***Service Metadata Locator (SML)*** is the only centralized component in the the BDX Architecture. Its sole purpose it to allow Service Providers to expose and discover the relationships between Identifiers (belonging to customers) and Service Provider. The SML is analog to the mechanism which allows people to switch telephone provider and still keep their telephone number. A request to the SML with an identifier as parameter will give a pointer to another service provided by the Service Provider serving the customer to whom the identifier belongs (se below).
- Each Service Provider is responsible for exposing further information related to each identifier they have exposed in the SML. Information about the capabilities associated with an identifier is exposed in a ***Service Metadata Publisher (SMP)***.

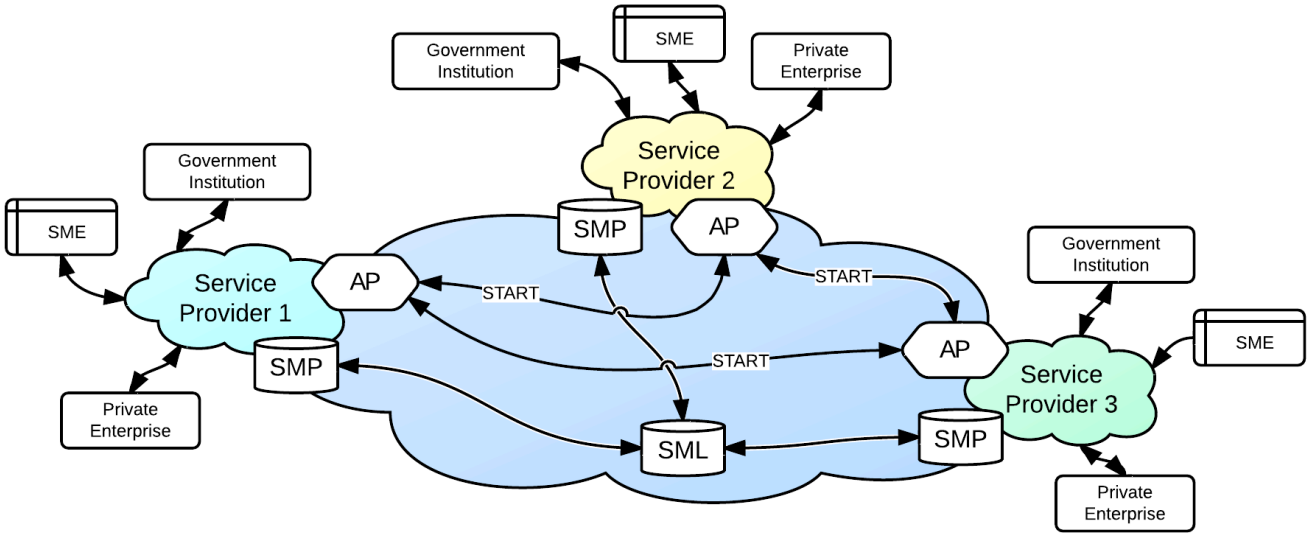
The Addressing infrastructure can answer questions like the following:

Which service provider serves a customer using the Global Location Number (GLN) 5060259470004 as identifier and what business processes does the customer support and what transport protocols can be used for exchanging business documents within these processes?

The Transport Infrastructure is very simple in its nature. Each Service Provider is responsible for operating a so called ***Access Point*** which can send an receive business documents from other Service Providers. An Access Point is basically a piece of middleware that supports a number of transport protocols.

A key requirement in a BDX Infrastructure Instance is that all Access Points must support a default transport protocol. The default transport is the common denominator that can be used for exchanging any business document between any two Service Providers. Service Providers are free to support other transport protocols.

START (Secure Trusted Asynchronous Reliable Transport) is a profile of secure, trusted, asynchronous and reliable messaging. START is a candidate to a default profile used by BDX Infrastructure Instances and designed to support exchange of business documents between Service Providers in the 4-corner model.



Addressing Infrastructure

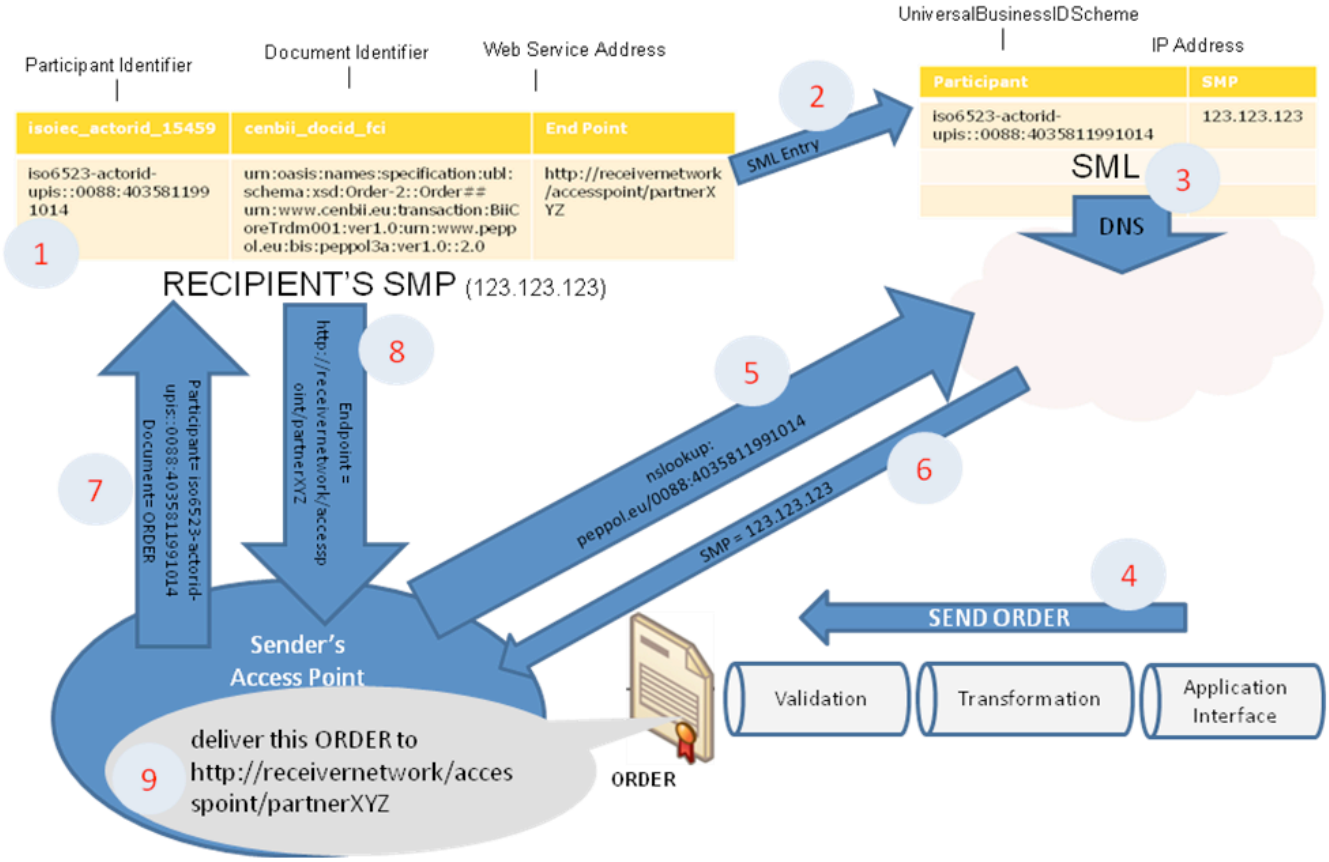
The heart of the BDX-Architecture is a network of federated addressing registries. These registries store information about the capabilities for exchange of business documents for the connected private companies and public sector institutions from different countries.

The BDX adressing mechanism is a 2-layer indirect adressing lookup. It utilizes the DNS system for easily identifying the "service meta data publisher" for a specific identifier, thus spreading the load away from a singe point of failure (and away from a single point of administrative bottleneck)

This network of e-business registries provides the users/interactors of BDX-Infrastructure with reliable trustworthy information about the participating businesses, the business processes they support and the individual business documents exchanged in each business process..

The private companies and public sector institutions are registered with a globally unique identifier as the key for accessing the information in the registry. This identifier is used for looking up the the receivers address (at which the business documents can be sent to). I.e. the identifier is similar to a postal address and is the key for obtaining information in the registry.

A number of different identifiers can be used as keys in the registries.



Using the identifier a number of questions can be asked to the registry. E.g.

- To what address can I send my business document?
- Can I expect an "order change" in return when I send company Y an electronic order?
- What transport standard should I be using when sending an electronic catalogue to company Y?

These questions are off course not asked by humans. The registry is primarily accessed by middleware products that need information about how and where they should send an electronic business document.

The registries are run and managed by various service providers connected to BDX-Infrastructure.

A model for re-using existing identifier schemes

The key to success for an interoperability infrastructure, is not just being able to bridge different transport networks, but also being able to create a common adressing-scheme, that enables interoperability of existing identifier schemes.

BDX adresses this challenge, by defining a UPIS meta-model for the interoperability identifers.

UPIS = Universal Participant Identifier Scheme.

The scheme indicates the format follow identifier format:

Participant identifiers logically consist of a scheme identifier and the participant identifier itself.

{type identifier}:{participant identifier}.

The type identifier is 4-digit number indicating the type of participant identifier, such as GLN, DUNS, CVR or another scheme.

Note that the {type identifier} part of the participant identifier is not equivalent to the scheme of the identifier - the {type identifier} indicates the type of participant identifier, whereas the scheme identifier indicates the format and semantics of an identifier string.

An example of such an identifer: 0010:5798000000001 .

Using the UPIS, participants can keep their existing identifiers, and make them available across the BDX network by adding the UPIS type identifier.

In BDX we re-use the existing code-list for initial set of identifier schemes developed by PEPPOL.

schemeID	schemeAgencyName	Numeric ISO6523 code
GLN	GS1	0088
DUNS	Dun & Bradstreet	0062
IBAN	S.W.I.F.T. Society for Worldwide Interbank Financial Telecommunications s.c.	0021
ISO6523	ISO (International Organization for Standardization)	0028

DK:CPR	Danish Ministry of the Interior and Health	9901
DK:CVR	The Danish Commerce and Companies Agency	9902
DK:P	The Danish Commerce and Companies Agency	0096
DK:SE	Danish Ministry of Taxation, Central Customs and Tax Administration	9904
DK:VANS	Danish VANS providers	9905
IT:VAT	Ufficio responsabile gestione partite IVA	9906
IT:CF	TAX Authority	9907
IT:FTI	Ediforum Italia	0097
IT:SIA	Società Interbancaria per l'Automazione	0135
IT:SECETI	Servizi Centralizzati SECETI	0142
NO:ORGNR	Enhetsregisteret ved Bronnoysundregisterne	9908
NO:VAT	Enhetsregisteret ved Bronnoysundregisterne	9909
HU:VAT		9910
SE:ORGNR		0007
FI:OVT	Finnish tax board	0037
EU:VAT	National ministries of Economy	9912
EU:REID	Business Registers Network	9913
FR:SIRET	INSEE: National Institute for statistics and Economic studies	0009
AT:VAT	Österreichische Umsatzsteuer-Identifikationsnummer	9914
AT:GOV	Österreichisches Verwaltungs bzw. Organisationskennzeichen	9915
AT:CID	Firmenidentifikationsnummer der Statistik Austria	9916
IS:KT	Icelandic National Registry	9917

Trust and security

The BUSDOX specifications must provide the foundations for secure, reliable infrastructures. To achieve this goal, the specifications address communications security between the various infrastructure components in the architecture e.g. Access Points, Service Metadata Publishers and the Service Metadata Locator. The security properties obtained are authentication, integrity and confidentiality.

The specifications do not deal with security between the first and last legs of the four corner model, i.e. between the sender and the sender Access Point or between the recipient Access Point and the recipient. This allows BUSDOX to be open-ended and connect existing, heterogeneous infrastructures such as existing VANS.

Furthermore, the specifications do deliberately not deal with end-to-end security properties e.g. obtained by signing or encrypting the payload documents transported by the infrastructure. This can be added as an additional layer on top of the BUSDOX infrastructure according specific business requirements.

Authentication is achieved using digital signatures and digital certificates. For example, the START specification requires the sender to sign the SOAP message and include its certificate in the header as a security token. The specifications are deliberately left open on how to establish trust in the certificates such that different models can be realized at run time. As an example, the PEPPOL project has defined that all operators of infrastructure components shall obtain a PEPPOL certificate from a central authority – and that these certificates are the only trusted certificates in the infrastructure. This provides a centralized trust model with the possibility for tight control and governance. Other more federated or distributed trust models are possible using different trust policies.

One final aspect of the specifications worth mentioning is the following: even though BUSDOX does not specify how an Access Point communicates with their senders/receivers, it requires that senders are authenticated by their Access Point (mechanism not specified) and that the Access Point assert the sender identity and authentication assurance level by issuing a SAML Assertion that is sent along with the message. The Access Point may choose to delegate this part to a trusted Security Token Service which will then become the issuer of the assertion.

Access points

An access point is a gateway that allows an existing infrastructure to connect to an BDX-Instance. An access point consists of two sides:

- A local side facing the chosen middleware solution for the local infrastructure,
- A an BDX-Instance side facing the standardized an BDX-Instance interfaces.

The local side is build on the basis of commercial and/or open source off-the-shelf software (i.e. message oriented middleware), whereas the an BDX-Instance side is implementing a standardized adaptor developed for the an BDX-Instance specifications.

Access points should have the ability to use existing services (such as PKI services etc.) and may use a shared registry infrastructure, depending on registry infrastructure mechanisms scheduled to be defined in a near future.

Transport protocols [Mikkel]

[Here we will explain that different transport protocols can be used. It is important that any BDX-Infrastructure has a default transport protocol which must be supported by all Access Points (service providers).

- Access point to access point is significant here

Addressing Registry infrastructure

The fundamental requirement for a registry infrastructure is that it must allow a sender or gateway to discover the recipient endpoint in order to deliver a business document to the recipient - whether the recipient endpoint is an access point, an proxy chosen by the recipient, or a technical endpoint set up by the final recipient - and that discovery is a machine-to-machine process.

In order to discover this endpoint, the sender or gateway should be assumed to know as little as possible information about the recipient or intermediaries between the sender and the recipient, apart from the information that is common to all or most business documents exchanged. This may be as little as:

- A business level identifier, representing the recipient organization
- The type of the business level identifier
- A country code representing the country of the recipient
- The type of document that the sender wants to send

The fundamental requirements for the operation of the registry infrastructure is

- That it must be scalable
- It should be as decentralized as possible
- Every recipient endpoint in the infrastructure must be discoverable for every sender

It would be desirable for the registry infrastructure to support independent, commercially driven registry operators, that may connect into the registry infrastructure within some legal agreement

framework. Assuming these properties of the infrastructure, it should be expected that a large number of independent registries will be operated.

Discoverability

A fundamental challenge of the registry infrastructure is *discoverability*, i.e. to ensure that

- Registries are discoverable
- Recipient endpoints are discoverable

In a BDX registry structure, there are two kind of registry entities in the registry infrastructure:

1. **Service Metadata Locator** - which hold references to other registries in the infrastructure
2. **Service Metadata Publisher** - which hold information on recipient endpoints and their properties, or *endpoint metadata*.

In a decentralized registry structure, in order to find a recipient endpoint, the sender (or a sender proxy) must find a suitable *addressing registry* to look up the recipient endpoint, based on metadata known about the recipient.

For the discovery of *e-business registries*, the only information available for the sender is the business ID and / or country code.

Discovering recipient endpoints

When the relevant *Service Metadata Publisher* has been found, the sender can perform a lookup based on

The business identifier and identifier type
The type of document
Various required capabilities of the endpoint, e.g. transport profile and semantic properties

Governance

References

Terminology

Access Point - connection point between existing infrastructure and the BDX interoperability infrastructure

Addressing registry - contains information about the e-procurement capabilities for exchange of business documents for the an BDX-Instance-connected private companies and public sector institutions from different countries.

Addressing registry operator - A private company or a public sector institution hosting an addressing registry. The customers of the operator can be registered in the registry.

Business document - A business document is an electronic document used to communicate business transactions between companies and private sector institutions. Examples of business documents within the procurement domain are electronic product catalogues, orders and invoices.