SML federation model suggestions

eDelivery internal

Exported on 03/06/2024

Table of Contents

1 Introduction

1.1 Purpose

The purpose of the document is to provide feedback from the CEF eDelivery team to the paper: ***e-Invoice Exchange Framework: Approach to Managing a Federated Registry Services Model in a Four-Corner Network*** prepared by ***the Business Payments Coalition*** [9] (see page 10).

The comments in this document are made in good faith and with a view to the continuing evolution of the message exchange standards that can meet the diverse needs of various business areas.

1.2 CEF eDelivery background

During the past years, the CEF eDelivery has been heavily involved in many cross-border, Large Scale Projects (LSP) to promote and facilitate the electronic message exchange for various networks of participants (also called business domains) such as eInvoicing, eHealth, eJustice, registered mail delivery, ...

The main goal of the CEF eDelivery is to enable the electronic exchange of digital data and documents in an interoperable, secure, reliable, trusted, and a reusable way for businesses and public administrations in the European Union.
As a leading message exchange standards throughout the projects, the OASIS ebMS 3.0 with AS4 profile was adopted along with OASIS standards BDXL v1.0 and OASIS SMP 1.0 for dynamic discovery of the participants' message exchange capabilities.

1.3 CEF eDelivery building block

The CEF eDelivery building block is developing various messaging components. The purpose is to prototype and prove concepts that facilitate the electronic exchange of digital data and documents and then provide an EU reference thereof. By doing so, CEF eDelivery limits the costs and the burden during the initial phase of expanding digital governance processes and e-business, and thus helps establish a "shared digital market." At the same time, it gathers best practices and experiences from the past pilot projects in a form ready to be used in new initiatives. The guiding principle is to help establish message exchange networks for various business domains in a way that avoids both vendor- and business domain lock-in. When the messaging network of participants reaches a certain degree of maturity and sustainability, it is expected to become independent from the CEF services. That is why it is important that the used standards enable a smooth and low-cost transition from the CEF eDelivery service infrastructure to established business domain infrastructure.

2 CEF eDelivery challenges

In addition to the scope objective described in the BPC document [9] (see page 10), this document seeks to highlight the following challenges faced by CEF eDelivery's services/infrastructure:

- **Top DNS domain transition:** CEF eDelivery offers an SML service with the top DNS domain **.edelivery.tech.ec.europa.eu**. The purpose of the service is to facilitate the establishment of different exchange networks for each business domain. When, later on, business domains set up their own SML, they use a different DNS domain under their authority. The relocation of existing participants to the new top DNS domain presents a technical challenge with possible disruptions in the functioning of the business domain exchange network.
  A concrete example: at the moment, the CEF SML service hosts SML DNS records for approximately 2.5M OpenPEPPOL end users and approximately 200 OpenPEPPOL SMP service providers [4] (see page 10). In case the exchange network allows only one top domain at a time, all 2.5M end users and 200 SMP service providers will have to update their configuration at the same time. With this amount of participants, this type of big-bang transition to the new dedicated top DNS domain is practically impossible. To allow a smooth transition of participants between top DNS domains, a federated model which allows multiple top domains at the same time is needed.
- **Big DNS zone:** DNS software solutions are optimized according to user requirements. Most authoritative DNS servers handle between a few dozen to a hundred DNS records. The largest DNS domain is .com [1] (see page 10) [5] (see page 10) which has 155 million records. The next largest DNS domain is .de [5] (see page 10), which has 16 million records. And according to Verisign [1] (see page 10), there are about 633M domains registered on the web. Therefore, it is difficult to find reliable, efficient, and accessible (from a cost perspective) DNS software on the market that could process more than a few tens of million DNS records in a given DNS zone. If we also include DNSSEC for the records, that task is even more challenging. In the EU, with a population of 448M [6] (see page 10) and 22M enterprises [3] (see page 10), we can expect that the e-invoice exchange network can grow to over 100M participants. Since not many authority DNS solutions handle well the huge DNS zones with DNSSEC enabled, a suggestion is to allow a federated model with multiple top domains.
- **Legal restrictions of the top DNS domain owners:** DNS domains are bought or loaned from the domain registrar. The domains (including the top-level domains (TLDs)) are under the legal jurisdiction of the country in which the domain registrar has its seat. In certain domains, EU countries or EU government agencies will not use networks based on DNS infrastructure that can allow non-EU countries (directly or indirectly) to disrupt the message exchange because of political or other reasons. As illustrated by the situations described in articles [7] (see page 10) and [8] (see page 10), the .com and .us TLDs may not always be suitable for use in the EU, and, conversly, the .eu TLD may not always be suitable for the USA or Canada.

To address to these challenges, the following chapter describes two variants of the option described in the chapter Option 1: Federated Registry services using Unique Domain scheme for each registry in Appendix D - Federated Registry Approach Options of the BPC document [9] (see page 10).

3 Additional federated Registry Approach Options

3.1 CEF eDelivery Option 1.A: Federated Registry services using fallback.

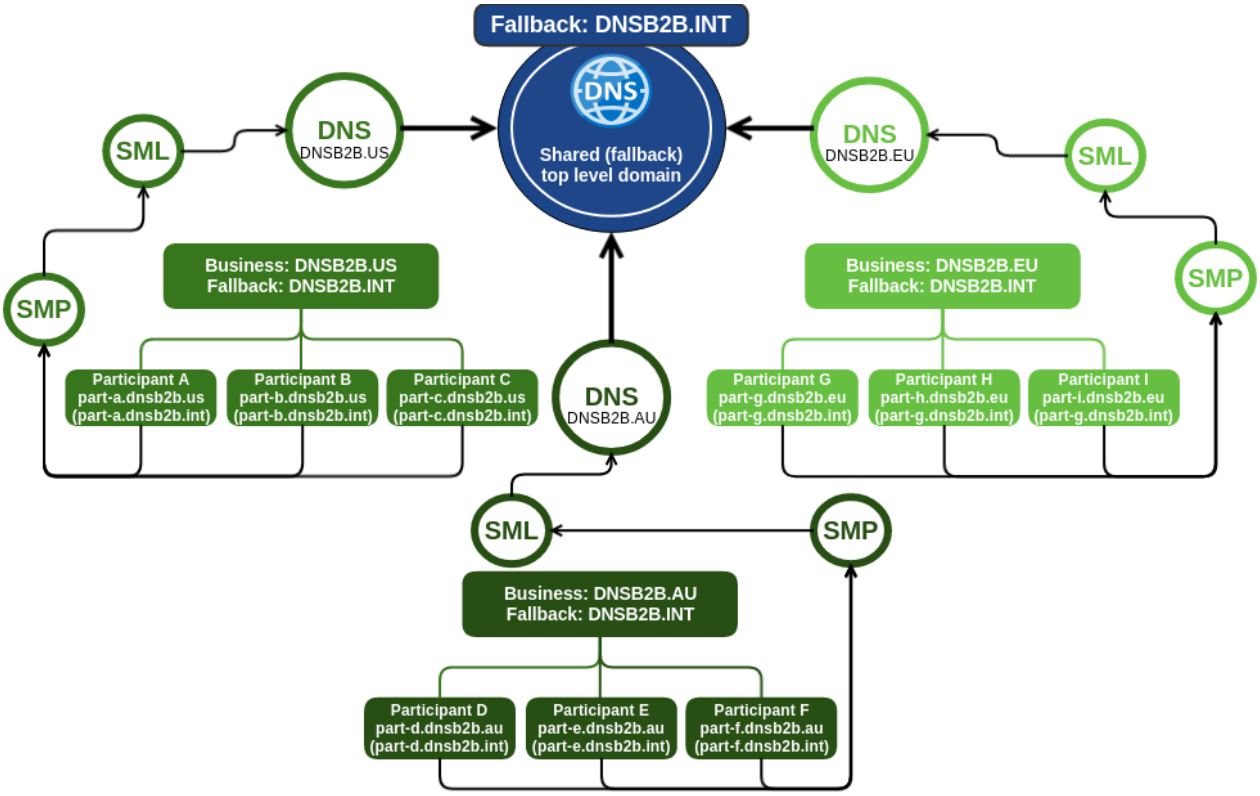
The difference with the option 1 presented in the BPC document [9] (see page 10) is that in this proposal registrars replicate participant IDs only to one "global" authoritative DNS server. The setup considers that, even in a global context, most invoices are still exchanged in the same region. The regions can each have their own trusted top domains, e.g.:

- USA: dnsb2b.us,
- Europe: dnsb2b.eu,
- Australia: dnsb2b.au
- ...

and one international fallback domain, e.g. dnsb2b.int.

Because most enterprises do not perform business globally, the regional SML service providers register participants in the regional DNS top domain. The SML service providers can also offer services (e.g., for additional costs to network participants) to register a company also in the international DNS top domain. The cost-driven approach avoids that all participants are by default published also in the international DNS top domain, and therefore that a very big zone is required at that level.

The lookup process would work by participants first trying to discover a recipient via the regional DNS top domain and, if this fails, then via the international DNS top domain.



3.1.1 Advantages:

- Smaller regional DNS zones (lower DNS maintenance costs).
- SML service providers can easily operate only the regional DNS service.
- Global discovery can be provided as additional service by the regional SML service provider.
- Technology exists today without modification.
- Technology is well known, well understood and robust, with established mechanisms.
- Small retry/fallback change to existing discovery model for the regional discovery.
- The upgrade of the discovery clients/libraries to be able to use an additional top domain is not complex.

3.1.2 Disadvantages:

- The discovery clients must be able to configure two top DNS domains: regional top domain and international top domain.
- It is not immediately clear who operates the international DNS zone.

3.1.3 Security notes

The network authority must maintain and publish a list of authorized top-level domains in the exchange network. The authorization can be granted or revoked in case of security concerns or network rules violation

to any of the SMLs providers. It is also recommended that the fallback international top domain implements the rules, enabling it to be registered with the TLD '.int'.

In case the network authority allows the same participant identifier to be registered in several regional top domains, then clear rules and/or additional means of verification should be defined by the network authority to allow participants to trust that the discovered SMP data matches the actual party they want to communicate with.

In case the network authority does not allow the same participant identifier to be registered in more than one regional top domain, then several options exist:

Security option 1: Mandatory use of "fallback/international" top domain: Each regional SML provider must register participants also in the "fallback/international" top domain. If another SML service provider already registered the participant identifier in the fallback DNS, then no other SML provider must register the same identifier in its own regional top domain. The downside of this approach is that it creates the "big DNS zone" problem at the "fallback/international" top domain.

Security option 2: Before registering an identifier, the regional SML (or SMP, or both) must verify if the hash of the identifier already exists in any other of the network's top domains with a simple DNS lookup. If so, it should return an error and not register the participant in the regional top domain.

Security option 3: The network authority sets up the directory service of all identifiers in the network. Before registering an identifier in the one of the regional top domains, the SML service provider must verify in the directory that the identifier is not already registered by another SML service provider. Only then should the SML service provider register the identifier in the directory service and add a corresponding record in the regional top domain.

3.1.4 Top DNS domain transition

In a situation where all participants' access points are configured with a (legacy) regional top domain and a fallback domain, the following steps would be necessary to achieve a DNS domain transition:

1. Participants from the legacy regional top domain are gradually moved to both the future regional domain and to the fallback domain. Note that this model works even if the network authority does not allow the same participant identifier to be registered in more than one regional top domain.
2. At the same time, participants' access points can be gradually re-configured to point to the future regional top domain and the (same) fallback domain.

This ensures that both the transition of SML data and the re-configuration of access points can take place gradually.

3.1.5 Migration of identifiers between top domains

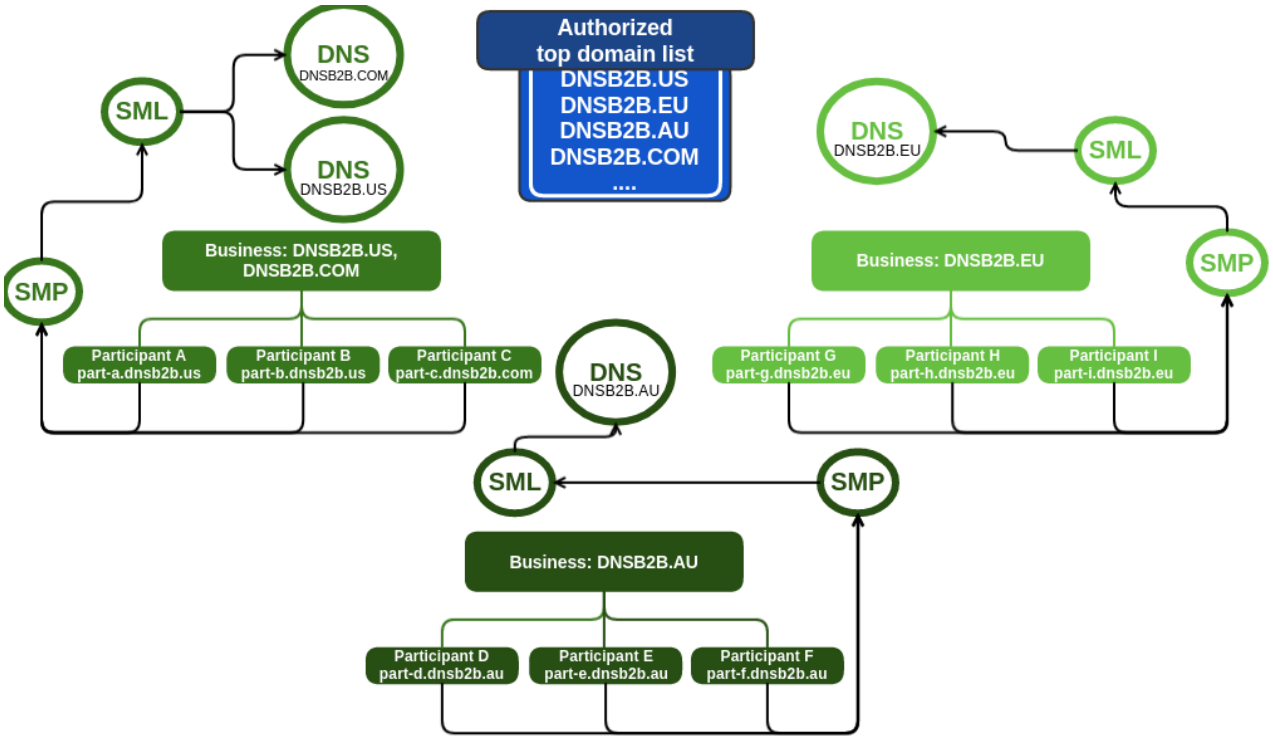
Migration of identifiers from one regional top domain to another is difficult if the network enforces a security rule that participant identifiers can be registered only once in the network.

The most straightforward approach is for participants to de-register their identifier from an existing regional SML provider and then register it with a new regional SML provider. The downside of the approach is the possible long duration of the migration process.

The alternative is to define a migration protocol that all SML service providers must support to allow SMP providers an atomic migration of participant identifiers from one regional SML provider to another regional SML provider.

3.2 CEF eDelivery Option 1.B: Federated Registry services using list of domains.

The characteristic of the option is that the messaging network does not have a "central DNS server" that would be shared by all SML service providers. Instead, the message exchange network defines a list of the authorized DNS domains that offer SML DNS services for the business domain. If dynamic discovery does not resolve one DNS domain, it can try the next one in the list, etc. The list order can be different for various regions to increase lookup efficiency. For example, in Europe, the first domain in the list would be dnsb2b.eu, since that is the most likely one to provide a hit, while in the USA, the first domain in the list would be dnsb2b.us. Also, simple logic can be implemented to choose the most probable top domain for a particular party identifier registrar.



Besides the regional static list, the software providers can implement soft logic to determine the preferable order of top domains based on the original party location or receiver party identifier scheme. For example, if party identifiers are using ISO6523, the search party client can assume the most likely top domain where the party is registered: parties starting with "iso6523-actorid-upis::0198:" are very likely registered with the EU top domain and parties beginning with "iso6523-actorid-upis::0151:" are very likely registered in Australia.

3.2.1 Advantages:

- Smaller regional DNS zones (lower DNS maintenance costs).
- SML service providers can easily operate only the regional DNS service.
- Allows for true federated registries.
- Technology exists today without modification.

- Technology is well known, well understood and robust, with established mechanisms.
- No change to existing discovery model for the regional discovery.
- Allows simple or more sophisticated implementation of top domain retry mechanism.

3.2.2 Disadvantages:

- To be useful for global network discovery, the discovery clients must be able to configure a top DNS domain list.
- An enhancement is needed for discovery clients to establish the best top domain list search order.
- When using non-global party identifiers, the same party identifier can be used in different SML service providers for different participants. To prevent this issue, the exchange network should use only party identifiers using international standards, such as ISO6523.

3.2.3 Security notes

The security concerns are the same as described in the previous chapter, with the exception that the **Security option 1** does not apply to this model.

An additional benefit of allowing participant identifiers to be registered in multiple regional SML service providers is that it allows parties to have backup records and faster discovery over the globe if they perform business globally.

3.2.4 Top DNS domain transition

The start position that all participants' access points are configured with a list of domains which contains the legacy regional top domain, the following steps would be necessary to achieve a DNS domain transition:

1. Network authority publishes an updated list of top regional domains, which also contains the new regional top domain.
2. Participants access points (automatically) updates their list of regional top domains.
3. Participants from the legacy regional top domain are gradually moved to the new regional domain. Note that this model works even if the network authority does not allow the same participant identifier to be registered in more than one regional top domain.

This ensures that both the transition of SML data and the re-configuration of access points can take place gradually.

3.2.5 Migration of identifiers between top domains

Migration of identifiers is straightforward with this model if the network authority allows participants to register their identifiers in multiple regional top domains. Participants can register identifiers in the new regional top domain and have the old entry as a backup.

The migration concerns for the case where duplicate identifier registration is forbidden are the same as described in the previous chapter.

4 References

1. The Domain Name Industry Brief (Volume 18, Issue 2)[1], Verisign, June 2021
2. Business demography statistics explained[2], Eurostat, Retrieved: June 2021
3. Annual enterprise statistics by size class for special aggregates of activities (NACE Rev. 2)[3], Retrieved: June 2021
4. Digital Service Infrastructure dashboard[4], CEF Digital, Retrieved: June 2021
5. Domain Count Statistics for TLDs[5], DomainTools, Retrieved June 2021
6. EU population in 2020[6], Eurostat, July 2020
7. ICANN Doesn't Take Down Websites[7], ICANN, Retrieved: June 2021
8. US shuts down Canadian gambling site with Verisign's help[8], Trevor Pott, Iain Thomson, Mar 2012, Retrieved: June 2021
9. e-Invoice Exchange Framework: Approach to Managing a Federated Registry Services Model in a Four-Corner Network[9], Business Payments Coalition, e-Invoice Technical Work Group, March 2021

1 https://www.verisign.com/assets/domain-name-report-Q12021.pdf
2 https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Business_demography_statistics#Active_enterprises_in_the_business_economy
3 https://ec.europa.eu/eurostat/databrowser/view/SBS_SC_SCA_R2__custom_905984/
4 https://ec.europa.eu/digital-building-blocks/wikis/display/DIGITAL/eDelivery+dashboard
5 https://research.domaintools.com/statistics/tld-counts/
6 https://ec.europa.eu/eurostat/documents/2995521/11081093/3-10072020-AP-EN.pdf
7 https://www.icann.org/en/blogs/details/icann-doesnt-take-down-websites-3-12-2010-en
8 https://www.theregister.com/2012/03/01/bodog_shut_via_verisign/
9 https://businesspaymentscoalition.org/wp-content/uploads/bpc-e-delivery-network-validation-exercise-2020.pdf