ETSI GS PDL 013 V1.1.1 (2022-10)

Permissioned Distributed Ledger (PDL); Supporting Distributed Data Management

Disclaimer

The present document has been produced and approved by the Permissioned Distributed Ledger (PDL) ETSI Industry Specification Group (ISG) and represents the views of those members who participated in this ISG.
It does not necessarily represent the views of the entire ETSI membership.

Reference
DGS/PDL-0013_Sup_Dis_Data_Mgmt

Keywords
data management, PDL

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM**® and the GSM logo are trademarks registered and owned by the GSM Association.

Foreword

This Group Specification (GS) has been produced by ETSI Industry Specification Group (ISG) Permissioned Distributed Ledger (PDL).

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Executive summary

The present document defines requirements and functional architecture of supporting distributed data management based on Permissioned Distributed Ledger (PDL) reference architecture. This includes expanded ETSI ISG-PDL platform services for supporting distributed data management.

Introduction

The present document specifies PDL-based distributed data management. The organization of the present document is as follows. Clause 1 defines the scope of the present document. Clauses 2 and 3 provide normative and informative references and definition of terms, respectively. Clause 4 provides an overview of PDL reference architecture. Clause 5 describes distributed data management use cases and requirements. Clause 6 lists architectural requirements of PDL-based distributed data management. Clause 7 defines expanded ETSI ISG-PDL platform services for PDL-based distributed data management.

1 Scope

The present document specifies distributed data management based on PDL reference architecture. This includes:

- defining architectural requirements that are derived from distributed data management use cases including related use cases such as those described in ETSI GR PDL 009 [i.1] and ETSI GR PDL 002 [i.2];
- defining PDL-based distributed data management architecture according to PDL reference architecture as defined in ETSI GS PDL 012 [1]); and
- defining expanded ETSI ISG-PDL platform services for PDL-based distributed data management.

2 References

2.1 Normative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

Referenced documents which are not found to be publicly available in the expected location might be found at https://docbox.etsi.org/Reference.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are necessary for the application of the present document.

[1] ETSI GS PDL 012: "Permissioned Distributed ledger (PDL); Reference Architecture".

[2] ETSI GS PDL 011: "Permissioned Distributed Ledger (PDL); Specification of Requirements for Smart Contracts' architecture and security".

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] ETSI GR PDL 009 (V1.1.1): "Permissioned Distributed Ledger (PDL); Federated Data Management".

[i.2] ETSI GR PDL 002 (V1.1.1): "Permissioned Distributed Ledger (PDL); Applicability and compliance to data processing requirements".

3 Definition of terms, symbols and abbreviations

3.1 Terms

Void.

3.2 Symbols

Void.

3.3 Abbreviations

For the purposes of the present document, the following abbreviations apply:

API	Application Programming Interface
ARPS	Application Registration Platform Service
DC	Data Collector
DCC	Data Computation Controller
DCN	Data Computation Node
DCS	Data Consumer
DD	Data Discoverers
DDAPP	Distributed Data Application
DDM	Distributed Data Management
DH	Data Host
DLT	Distributed Ledger Technology
DO	Data Owner
DP	Data Provider
DPS	Discovery Platform Service
DS	Data Source
ETSI	European Telecommunications Standards Institute
ETSI ISG-PDL	ETSI Industry Specification Group for Permissioned Distributed Ledger
FCAPS	Fault, Configuration, Accounting, Performance and Security
FL	Federated Learning
GDPR	General Data Protection Regulation
GR	Group Report
GS	Group Specification
IRP	Interface Reference Point
ISG	Industry Specification Group
MARL	Multi-Agent Reinforcement Learning
MPS	Messaging Platform Service
PDL	Permissioned Distributed Ledger
RD	Requirement on Decentralization
RDCS	Requirement on Data Control and Sovereignty
RDDA	Requirement on Distributed Data Application
RDI	Requirement on Data Integrity
RDMA	Requirement on Data Management Automation
RDP	Requirement on Data Privacy
RDPV	Requirement on Data Provenance
RI	Requirement on Incentivization
RL	Reinforcement Learning
RPS	Registration Platform Service
RPSL	Requirement on Platform Service Layer
RT	Requirement on Trust
RUDLTN	Requirement on Underlying DLT Networks
SPS	Storage Platform Service
TCI	Transaction Creation Indication
TMPS	Transaction Management Platform Service

4 PDL Reference Architecture

4.1 Introduction

ETSI GS PDL 012 [1] develops a layered PDL reference architecture, which consists of five layers as illustrated in Figure 4.1-1. Each layer is designed in a manner that allows abstraction, such that it can be operated regardless of the implementation specifics of the other layers. In addition, Interface Reference Points (IRPs) are defined between different layers:

- **PDL Applications:** Various PDL-based applications leverage PDL services as provided by the below described Service Layer to interact with different DLT networks. For example, a PDL-based data sharing application utilizes a DLT network as a distributed infrastructure to enable distributed sharing. An application may also interact with external storage to store certain data that requires better privacy control or to reduce the overhead to DLT networks.

- **Application Abstract Layer:** This layer utilizes Data Model Brokers/Gateways enabling applications that allow different data models to communicate with ETSI ISG-PDL compliant platforms. This layer is located between the PDL Applications and Platform Services Layers and implemented through the Data Model Broker Platform Service where necessary.

- **Platform Services Layer,** which provides useful services for applications to support various types of applications using PDL technology. As a result, an application could leverage services from the Platform Service Layer rather than embed such services within the application itself. This reduces applications' complexity, accelerates application development and deployment, and increases interoperability. For example, the Platform Service Layer may include a Transaction Management Platform Service to facilitate transaction creation in a manner transparent to a specific PDL type (i.e. a specific deployed DLT network) and in a manner uniform across applications using such platform; this is an example of layer abstraction in its essence. Such a Transaction Management Platform Service can perform transaction transformation/adaptation between applications running on different PDL types to facilitate application operations in a complex environment.

- **DLT Abstraction Layer,** which consists of a Data Model Broker/Gateway enabling Platform services to communicate with ETSI ISG-PDL compliant PDL types regardless of the specific type of the underlying PDL. An additional functionality of such abstraction layer is to allow interoperability between different DLT types, which may differ not only in data model structure but also on consensus mechanism and smart-contract functionality. Such abstraction layer hides the differences between PDL types and provides a unified service-facing interface on the services side and a PDL specific interface on the PDL side. This layer is located between the "Techno" and the "Disco" IRPs and implemented through the Data-Model Broker Platform Service where applicable.

- **DLT Layer,** which includes various DLT networks (e.g. an implementation of a specific DLT type) and potentially the abstraction of DLT networks. While DLT networks and chains may vary in terms of consensus mechanism and smart contract format, the abstract functionality of a chain is very similar across most DLTs: Storing a distributed chain of data blocks in a tamper-resistant manner, and performing pre-programmed actions based on rules (i.e. "Smart Contracts") on all copies of the distributed chain. Yet, not all DLT types are necessarily compliant with the ETSI ISG-PDL layered architecture approach, thus the DLT layer can only include and accommodate DLT types that are compliant with said architecture.

- **Interface Reference Points (IRPs),** which define communication channels through which the functional blocks defined above communicate with each other. The IRPs are given names for reference purposes (e.g. Debka, Tango, etc.).

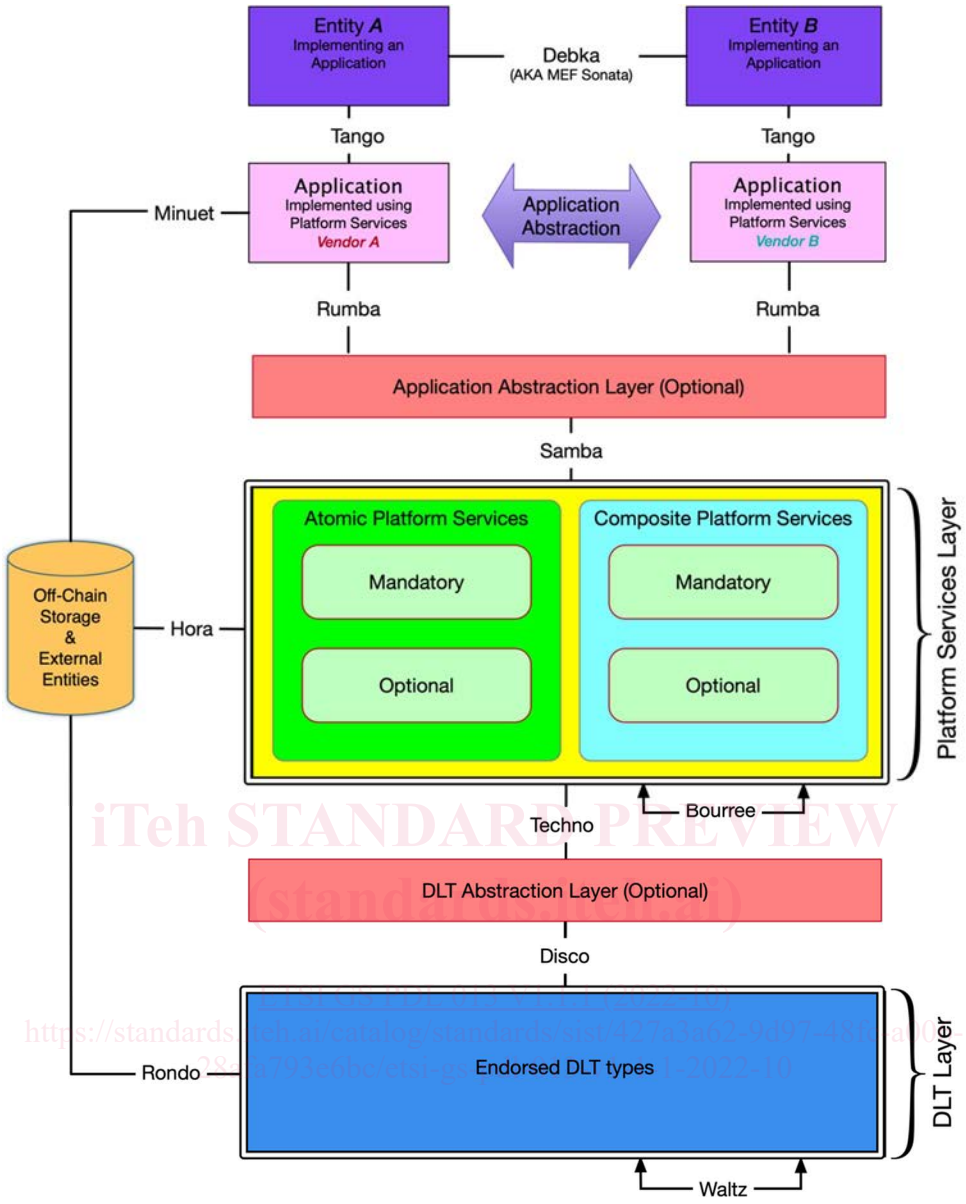

Figure 4.1-1: ETSI ISG-PDL Reference Architecture (Source: ETSI GS PDL 012 [1])

Four ETSI ISG-PDL platform categories (namely: Alpha, Bravo, Charlie, and Delta) are defined in ETSI GS PDL 012 [1]. The differences among those four categories lie in a few factors:

1) the number of involved vendors;

2) the number of supported underlying DLT technologies; and

3) the number of supported applications.

- Alpha Platforms: that are designed, developed, delivered, and integrated to all users of the said platform by a single vendor using a single DLT technology.

- Bravo Platforms: that are designed, developed, delivered, and integrated to all users of the said platform by a single vendor, but can operate using two or more underlying DLT technologies.

- Charlie Platforms: that can operate using two or more underlying DLT technologies and are designed and developed towards a specification of an application abstraction layer so that any application that supports such an abstraction layer can interface with the said platform. Moreover - the Platform Services in a Charlie Platform can be developed by multiple vendors towards the specification defined in ETSI GS PDL 012 [1].

- Delta Platforms: That use a single DLT technology and are designed and developed towards a specification of an application abstraction layer so that any Application that supports such an abstraction layer can interface with the said platform. This is, in essence, a simplified Charlie Platform that uses a single DLT type thus eliminating the DLT abstraction layer and eliminating the overheads associated with DLT interoperability.

4.2 Platform Services Layer

The Platform Services Layer hosts several types of services; details of each service are described in ETSI GS PDL 012 [1] where each such Platform Service is defined:

- The PDL Platform Services can be Atomic services or Composite services. Each such service could be Mandatory or Optional. Atomic services are self-sufficient and do not rely on other Platform services for their proper operation, while Composite services use one or more other Platform service to operate. A platform cannot function properly unless all Mandatory Platform Services are implemented therein, while Optional services may only be required for specific purposes or use-cases.

- PDL Platform Services are services and functionalities provided by the PDL platform that all applications may use. Platform Services may reuse or be built upon other Platform Services. Examples of Platform Services include: namespace, identity, location, discovery, messaging, policy, governance, security, composition, access control, concurrency storage, modelling, distributed processing, resource management, service management, transaction management, etc.

- In addition, PDL Platform Services Layer has Application Specific Platform Services that are services used by specific applications or specific groups of applications and are not needed or cannot be made useful for other applications (e.g. measurement of precipitation is useful for agriculture and weather applications but has no use for data storage applications). Such services may be implemented within the application itself, however the developer may want to contribute them and install them on the platform so the can be re-used by other applications in the future if the need arises.

Table 4.2-1 lists the Platform Services as defined in ETSI GS PDL 012 [1].

Table 4.2-1: ETSI ISG-PDL Platform Services [1]

PDL Platform Service name	Mandatory (M) or Optional (O)	Atomic (A) or Composite (C)	Short description
Namespace	M	A	Ensures that all of a given set of objects for a particular function have unique names.
Identity	M	A	Unambiguously identifies an instance of an entity from all other instances of this and other objects.
Location	O	A	Associates an object with a location.
Registration	O	A	List a managed object with authorities or registries.
Discovery	O	A	Discovery of services offered by the services layer and discovery of PDL networks.
Messaging	M	C	Enables communication between a group of entities.
Policy	O	C	Manage and control the changing and/or maintaining of the state of managed objects.
Security	M	C	A collection of services that assess, reduce, protect, and manage security risks.
Authentication	M	C	Verifies that a subject requesting to perform an operation on a target is who they say they are.
Authorization	O	C	Permitting or denying access to a target by a subject.
Cryptography	O	C	Managing protocols that prevent third parties from reading private communications.
Encryption	O	C	Encoding information using a key into an unintelligible form.
Identity Management	O	C	Access control based on the identity of an entity.
Key Management	O	C	Management of cryptographic keys.
Logging	O	C	Dynamic ingestion and collection of logs.
Governance	M	C	Rules and tools that control the behaviour and function of a PDL.
Implementation Agreements	O	C	Rules and agreements that describe how ETSI ISG-PDL Services are implemented and control the behaviour of a PDL platform.

PDL Platform Service name	Mandatory (M) or Optional (O)	Atomic (A) or Composite (C)	Short description
Governing Entity	M	C	Defines the rules and implementation agreements. Ensures compliance. Resolves conflicts where needed.
Composition	O	C	Defines who can compose new services and how such new services are composed.
Access Control	M	C	Defines who can perform which operations on which set of *target* entities.
Fault Tolerance	O	C	Defines how to handle faulty instructions.
Distribution Transparency	O	C	defines how to maintain transparency when distributing information to target entities.
Publish and Subscribe	O	C	Defines how entities publish services and subscribe to services.
Concurrency	O	C	Defines how entities handle concurrency.
Storage	M	C	A group of services related to Storage.
In Memory Storage	M	C	Data that is stored in the random access memory of a computer running an application.
File System Storage	M	C	Storage on a directly connected storage device.
On-Chain Storage	M	C	Application data that is stored in blocks on all nodes using the chain.
Off-Chain storge	O	C	Information in a digital, machine-readable medium that is not stored on the main chain.
Distributed Blockchain Storage	M	C	Storage on a Distributed Blockchain ledger.
Modelling	M	C	A group of services related to Modelling.
Information Model	M	C	Presentation of concepts of interest to platform management environment in a *technology-neutral* form as objects and relationships between objects.
Data Model	M	C	Representation of applicable concepts in a *technology-specific concrete* form.
Model Search	O	C	Enables search for specific or generic models within existing information and data models.
Model Stitching	O	C	Enables integrating multiple models or parts of models into a single model.
Topology	M	C	Allows a node to identify other nodes on the PDL and identify which nodes to communicate with when performing PDL related tasks.
Event Processing	M	C	Processes node-specific and platform-wide events as they occur.
Distributed Data Collection	O	C	Performs tasks related to collection of data that are location-independent.
Distributed Secret Sharing	O	C	Sharing of confidential data between nodes in a manner that maintains confidentiality of the data.
Resource Management	M	C	Defines how to administer and manage Resources.
Resource Discovery	O	C	Enables discovery of resources available to applications and nodes.
Resource Virtualization	O	C	Creating a virtual resource that mimics the behaviour of a physical resource.
Resource Inventory Management	O	C	Management of node-specific and platform-wide resource inventory.
Resource Admin and Management	M	C	Administration and management of node-specific and platform-wide resources.
Resource FCAPS	O	C	Resource management tasks defined by the ISO model.
Resource Composition	O	C	Creation and management of composite resources.
Platform Services Management	M	C	Defines how to administer and manage Platform Services.
Platform Service Discovery	M	C	Provides means to discover services available to applications and nodes.
Platform Service Virtualization	O	C	Creating a service using virtual resources.
Platform Service Inventory Management	O	C	Keeping track of inventory and serviceability of Platform services.
Platform Service Admin and Management	M	C	Administration and management of Platform Services through governance.

PDL Platform Service name	Mandatory (M) or Optional (O)	Atomic (A) or Composite (C)	Short description
Platform Service FCAPS	O	C	Platform Service management tasks defined by the ISO model.
Platform Service Composition	O	C	Creation and management of the composition of Composite Platform Services.
Application Management	M	C	Creation and management of Applications.
Application Composition	M	C	Composing an Application from two or more managed objects.
Application and Service Orchestration	O	C	Orchestrating multiple managed objects so they provide a desired set of behaviours.
Orchestration	O	C	Orchestration of objects, resources, services, and/or applications so that they collectively provide the desired functionality and behaviour.
Platform Exploration	O	C	Allows an application to indicate its requirements and explore whether the platform offers such service capabilities
Application Registration	O	C	Registers and lists all applications operated on a platform.
Transaction Management	O	C	Facilitates transaction related interactions between applications/services and underlying PDL networks.
Data Model Gateway/Broker	O	C	Defines tools that enable two systems with different data models to interact.
API Presentation	O	C	A specific Data Model Gateway/Broker implementation for environments that use APIs to exchange data between objects.
Application Specific Services	O	C	Serve a specific application or a group of applications but not required or used by other applications using the platform.

5 Distributed Data Management

5.1 Introduction

Distributed Data Management (DDM) is referred to operation and manipulation of data in distributed manners such as those illustrated in Figure 5.1-1. For each scenario in Figure 5.1-1, there are multiple distributed parties (referred to as data nodes); each data node has a Distributed Data Application (DDAPP), which supports a specific distributed data management task among those distributed data nodes:

- Distributed Data Discovery: Data is discovered from multiple distributed parties. Data discovery is required for distributed data collection, distributed data storage, distributed data sharing, and distributed data computation.
- Distributed Data Collection: Data is collected from multiple distributed parties.
- Distributed Data Storage: Data is stored in multiple distributed parties.
- Distributed Data Sharing: Data is distributed and shared among multiple parties.
- Distributed Data Computation: Multiple parties perform data computation in a distributed and collaborative way, for example, federated learning, distributed machine learning, and multi-party computation.