A secure network overlay for tracking and enforcement of data transaction rules

Ralph Koning, Reginald Cushing Lu Zhang, Cees de Laat, Paola Grosso

University of Amsterdam - System and networking lab (SNE), Science Park 904, 1098XH Amsterdam, The Netherlands

{*r.koning,r.s.cushing,l.zhang2,delaat,p.grosso*}*@uva.nl*

Abstract: Digital Data Marketplaces allow to securely share data between competing parties. To maintain data sovereignty in such environments we translate market transactions into audited, secure network connections which enforce policies and track data exchanges.

OCIS codes: 060.0060, 060.4250

1. Overview

Data sharing across organisational boundaries has the potential to open up new insights, as well as to create novel business opportunities. Digital Data Marketplace (DDMs) [1] [2] are emerging as architectures to support this mode of interactions. They rely on the participating parties agreeing on the permissible operations (*market transactions*) and expressing them into actionable contracts and policies. We call the allowed interactions patterns *archetypes*. Archetypes express the level of trust [3] and the underlying alliances between parties in a DDM. The archetype, inherently, influences the model of computation of applications which in turn influences the infrastructure setup to allow such applications to work. In this demo we investigate how to setup such infrastructures.

DDM archetypes define the rules of interactions between parties. A DDM must therefore ensure that members of the system do abide by these rules; and its effectiveness lies in its ability to guarantee compliance and control the operations performed in it. We distinguish between two types of controls a DDM can apply: enforcement and auditing. ”Control through Enforcement” are the set of DDM rules that are technologically enforceable e.g. identity can be enforced through cryptography. ”Control through Auditing” applies to rules that are not easily enforceable through technology and that require monitoring and auditing procedure to guarantee better control of the system and the applications running on it.

Parties are brought together on an overlay network which allows for connectivity between distributed parties while providing control points that can be used to enforce policies. Applications on the overlay are decomposed to a set of workflow transactions which in turn can be translated to network connections. Applications, implicitly, also embody the desired collaboration model. Choosing a collaboration model for an application is an area of research addressed by Zhang et al. [4] which explores the matching of applications to different archetypes allowed by DDMs.

The overlay network is composed of nodes with different functionalities. A node on the network is addressed by its unique public key (ed25519). Using public keys as an addressing scheme introduces several important concepts of the overlay. Firstly, any node can sign its actions on the network and the signatures can be easily verified by other nodes on the network since the address doubles as the verification key. Secondly, nodes can sign each other thus creating a chains of trust on the network. This creates the notion of a node owned by another node which is useful when nodes need to verify who owns what e.g. data.

Different node types serve different functionalities on the network. *Domain* nodes provide the highest level of trust on the network. The keys of the domain nodes are used to sign other nodes. This procedure allows us to identify the ownership of every node on the network. *Auditor* nodes provide a second level of trust. Their role is to sign activity on the network. Each auditor belongs to a domain and acts independently from the other domains. They compare activity on the network to their internal policy and issue a signature if the activity id compliant to their policy. Auditors also monitor and log activity on the control layer and signal alerts when illicit behaviour is detected. Each auditor maintains logs in a hash chain with each auditors cross-validating each others hashes. This creates a web of temper proof audit trails.

Cask nodes are responsible for holding and publishing data collections. A cask is an abstract concept used to identify and address datasets on the network. *Bucket* nodes are nodes used to transfer data between casks. The function of buckets is to create transient private connections between two nodes on the overlay using the address keys as a bases to setup encryption. Whereas casks are persistent data stores, buckets are transient and are created and destroyed on demand during the execution of an application. Infrastructure-wise, buckets are mapped to the

underlay endpoints that expose the buckets. In our approach, a bucket is realised as a Docker container. The container provides a layer of isolation on the host system and only allows access to the agreed upon data and compute resources. By default such a container is isolated from the Network. A bucket controller running on the host system facilitates the life-cycle of buckets and the set-up of VPN connections between buckets of which the end point interfaces directly assigned into the bucket when a transaction is initiated. *Planner* nodes encapsulate the applications running on the network. Planners are essentially single workflow coordinators. They must coordinate with the auditors to get enough signatures so that they can influence the control layer to effect the transactions.

Execution of an application entails the following high-level steps; first a planner gets an access grant/s from the auditor/s to join the network, next the planner will announce the execution plan to the auditors which will check and sign each transaction separately. The planner will then coordinate with different casks and bucket controllers to create the necessary VPN connection for every transaction. At every step the control layer checks for auditor signatures while auditors monitor the activity of the control layer to verify the activity conforms to the agreed application.

2. Innovation

Coordinating multi-domain applications is a challenge that transcends the traditional network and software stacks. Such applications are restricted by multilateral agreements between parties. Distributed multi-domain infrastructures that can enforce, audit and execute such applications are yet to be realised. Our work is a step towards such infrastructures. Through our research we discovered that the traditional approach of control and data planes are not enough to guarantee such an infrastructure and that a third plane is needed to create trust between the different domains. Furthermore, since not all policies can be enforced, an audit plane is needed to maintain an audit trail of activity. For this plane we introduced the concept of a network of auditors that sign actions which the control plane uses to evaluate its risk of taking actions such as exposing data endpoints. The importance of an audit plane means it needs to be an integral part of the network alongside the control and data plane. Here is where we propose and overlay network that brings these 3 planes together. It is also important to note that the different planes traverse different administrative domains and it is only through the careful interactions between the control plane and the audit plane that a data plane can exists between the different administrative domains. From a Software Defined Network (SDN) viewpoint, what we are proposing is a distributed multi-domain SDN, Figure 1, approach where each administrative domain is viewed as a single SDN domain while the audit plane creates a hyper-layer to allow multiple SDN domains to interact.

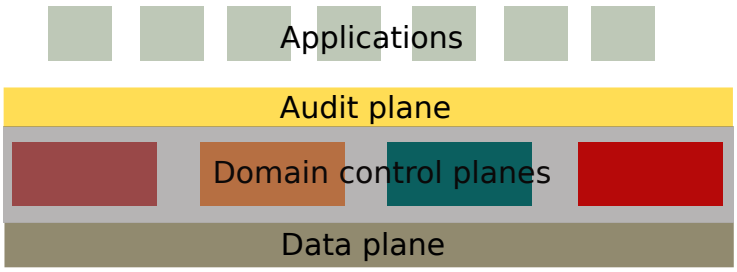

Fig. 1: Different planes in a distributed multi-domain infrastructure for running applications under multilateral agreements and policies.

3. Relevance to OFC community

Introducing uses cases for optical networks such as DDM will help attendees to understand how optical systems can be deployed to create high performance trusted environments. The audit plane, alongside the network and control plane, is an novel and interesting concept to discuss with the OFC community. Currently the auditing is limited to our software-defined-approach and its associated VPN links. In order to provide more guarantees for safely transporting the data, our future work includes the auditing of the complete communications stack, all the way down to the optical layer. By demonstrating our work (Fig. 2) to the community we hope to identify the existing technologies and the future innovations that are needed to realise this goal. Ultimately, the outcome of our work, the Digital Data Marketplace, can be used to facilitate a trusted approach to run approved algorithms on analytic and telemetry data from different organisations. The DDM is a novel architecture to address risks associated with data sharing, such as data leaks and theft of intellectual property. By establishing the trust and reducing the risk of data-misuse a DDM can provide the confidence required for novel business opportunities that can benefit service delivery, provide new insights and stimulate innovation.

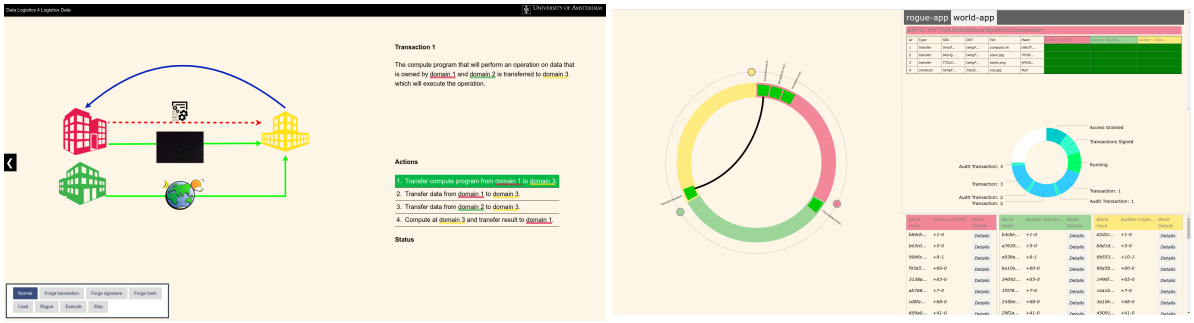

Fig. 2: Screenshots of the demonstration during Super Computing 2019: The left image shows the application running on the infrastructure and portrays the information sent on the data-plane. In the far left corner there are UI controls to control the various aspects of the auditing capabilities. The right picture shows the control-plane and auditing-plane. The circle on the left shows three domains and a VPN connection between two buckets. The right part shows the signed transactions, a circle with application/auditing progress and the signed audit trails for each of the domains.

4. Acknowledgements

This work is done within the Dutch NWO Research project 'Data Logistics for Logistics Data' (DL4LD, www.dl4ld.net), supported by the Dutch Top consortia for Knowledge and Innovation 'Institute for Advanced Logistics' (TKI Dinalog, www.dinalog.nl) of the Ministry of Economy and Environment in The Netherlands and the Dutch Commit-to-Data initiative (https://commit2data.nl/). The authors are grateful for invaluable contributions from Mr. Rodney G. Wilson and Mr. Marc Lyonnais of Ciena Corporation's External Research Programme.

References

1. S. Cisneros-Cabrera, A. Ramzan, P. Sampaio, and N. Mehandjiev, "Digital marketplaces for industry 4.0: a survey and gap analysis," in *Working Conference on Virtual Enterprises,* (Springer, 2017), pp. 18–27.
2. A. Zerdick, K. Schrape, A. Artope, K. Goldhammer, U. T. Lange, E. Vierkant, E. Lopez-Escobar, and R. Silverstone, *E-conomics: Strategies for the Digital Marketplace* (Springer Science & Business Media, 2013).
3. A. Deljoo, T. van Engers, R. Koning, L. Gommans, and C. de Laat, "Towards trustworthy information sharing by creating cyber security alliances," in *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE),* (IEEE, 2018), pp. 1506–1510.
4. L. Zhang, R. Cushing, L. Gommans, C. De Laat, and P. Grosso, "Modeling of collaboration archetypes in digital market places," IEEE Access **7**, 102689–102700 (2019).