ICONA: Inter Cluster Onos Network Application*

Matteo Gerola[†], Michele Santuari[†], Elio Salvadori[†], Stefano Salsano[§],
Pier Luigi Ventre[‡], Mauro Campanella[‡], Francesco Lombardo[§], Giuseppe Siracusano[§]

[†]Create-Net, [§]CNIT / Univ. of Rome Tor Vergata, [‡]Consortium GARR

Abstract—**Several Network Operating Systems have been proposed in the last few years for Software Defined Networks; however, only few of them are offering resiliency, scalability and high availability required for production environments. In our demonstration we present a geographically distributed SDN Control Plane, called ICONA, build on top of the Open Networking Operating System (ONOS) and designed to meet the aforementioned Service Providers requirements. During the demo, that runs inside the GEANT OpenFlow pan-european testbed, we show how a Service Provider engineer can easily manage and monitor the network, deploy some services and how ICONA can automatically recover from Control and Data planes failures.**

Keywords—*Software Defined Networking, Open Source, Geographical Distributed Control Plane, Open Networking Operating System*

Fig. 1: Distributed NOS architecture

I. INTRODUCTION

Since the beginning of the Software Defined Networking (SDN) revolution, one of the major concerns stated by Service and Cloud Providers has been the Control Plane reliability, scalability and availability [2] [3]. Indeed, a standard IP/MPLS network, is surely complex to configure and debug, but offers natively fast resiliency in case of hardware failures. The SDN architecture, that split Data and Control Planes, allows an easier design and management, but moves the intelligence from the physical devices to a Network Operating System (NOS), that is in charge of all the forwarding decisions. For this reason, the NOS can't be a single point of failure in production environments, such as Service Provider's (SP) networks. However, only few SDN NOSes are focusing on resiliency aspects (DISCO [4], Kandoo [5], HyperFlow [6] and ONIX [7]), and only one of them, called Open Networking Operating System (ONOS) [8], offers an open-source stable implementation.

ONOS provides a distributed Network Operating System with a logically centralized network view. Scalability is provided by the partition of the network resources controlled by different ONOS instances. The control plane distribution offers also fault tolerance: each device registers to multiple ONOS instances. In case of fault, a backup controller gains the control of the switches and notifies the others through a distributed registry. The data plane itself is resilient: when a link or a network device fails, ONOS automatically provides traffic rerouting. ONOS cluster is designed to operate with high bandwidth and limited latency between its instances (under 5 ms of latency and 10 Gbps of bandwidth). Usually these requirements are met in a single data center, but a distributed control plane should be able to manage complex geographical topologies, 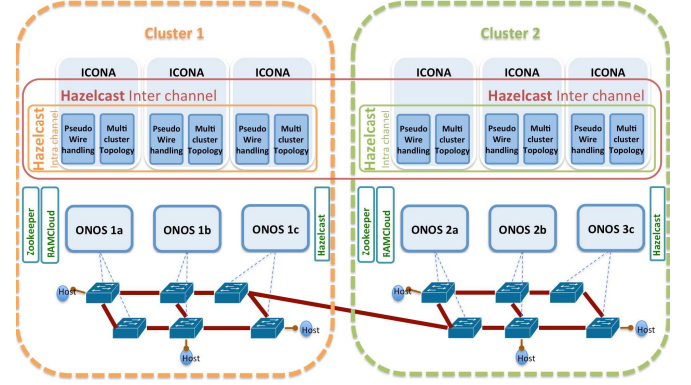where latency can reach easily 50 ms. The Inter Cluster ONOS Network Application (ICONA), presented in this paper, aims at designing and implementing a new architecture on top of ONOS to provide such fundamental feature.

II. ICONA ARCHITECTURE

ICONA partitions the SP's network in geographical regions, each one managed by a different cluster of ONOS controllers. In this scenario, each ONOS cluster will provide both scalability and resiliency to a small geographical region, while several clusters will use a publish-subscribe event manager to share topology information, monitoring events and operators requests. The Control Plane is geographically distributed among different data centers, each one controlling the near portion of the physical network. Figure 1 shows the aforementioned architecture where each cluster is located in a different data center (e.g. ONOS1a, ONOS1b and ONOS1c are instances co-located). To offer the same reliability, also ICONA application runs on top of each ONOS instance. Inside each cluster, a master application is elected, using a distributed registry. The ICONA master is in charge of sharing information and applying decision, while all the backups are aware of the network status, and can replace the master in case of failure.

In a geographical architecture, that covers thousands of square kilometers, a key point is the type and amount of information that the different segment of the control plane have to share, in order to minimize the traffic while maximizing the benefits in terms of: i) offering an up-to-date view of the network, including status of the nodes, ii) configuring the network devices and iii) reacting to failures both in Data and Control Plane without disrupting customer's traffic. In our initial implementation, we have considered only few scenarios, such as L2 pseudo-wire tunnels and MPLS VPN

overlay networks, that are key services in a SP network. However, the application can be easily extended to provide other functionalities.

The communication between different ICONA instances, both intra and inter-cluster, is multicast and is based on Hazelcast [9]: one channel is devoted to the intra communication (e.g. same ONOS cluster) and another one for the inter clusters messages. To offer a partial view of its network portion, each cluster shares with the others the following information through the inter-channel:

- Inter-cluster links (IL): an IL is the physical connection between two clusters. ICONA implements an enhanced version of the ONOS discovery mechanism, based on the Link Layer Discovery Protocol (LLDP). Each IL is shared with all the clusters with some metrics, such as the link delay and the available bandwidth.

- End-point (EP) list: an EP defines the interconnection between the customer's gateway router and the SP network. Each cluster shares the list of its EPs and the metrics (bandwidth, delay) between these EPs and all the clusters ILs.

With these information, each remote ICONA application can compute the overall metrics and choose the best path between two EPs, no matter if they belong to its portion of the network or not. The channel is also used to manage (instantiate, modify and delete) network paths between EPs, that are translated into OpenFlow [10] MPLS-based flows and installed in the physical devices.

In case of a failure in the data-plane, the hybrid architecture defined by ONOS and ICONA presents two different situations: (i) intra-cluster and (ii) inter-cluster failure. In the first case, ONOS, by design, takes care of rerouting all the paths involved in the failure. Otherwise, ICONA handles the inter-clusters failures (e.g. it involves an IL or one of the two nodes that sits at the edge of the IL). For each IL, a backup link (BIL), completely decoupled from the primary one, is pre-installed in the data plane, and, in case of failure, all the traffic crossing the IL is rerouted to the BIL by each ICONA edge cluster, without the need to wait for remote instances to share any information. Then, after the rerouting is completed, the new path is notified to all the remote clusters, in order to share the same network status.

III. DEMONSTRATION

The objective of the demonstration is to show how ICONA can offer a geographically distributed SDN Control Plane designed to improve reliability without worsen the performances. To offer a real experiment, the proof of concept runs on top of the GEANT GOFF OpenFlow facility [11], a distributed testbed interconnecting different EU countries (Nederland, Germany, Austria, UK, Croatia). Figure 2 shows both the physical SDN infrastructure (both Data and Control Planes) and one of the customer's overlay network. The Control Plane is composed of thee ONOS clusters, for an overall of six instances running both ONOS and ICONA. Each clusters controls a subset of the physical network, while ICONA provides the glue that offers a single control interface to the network
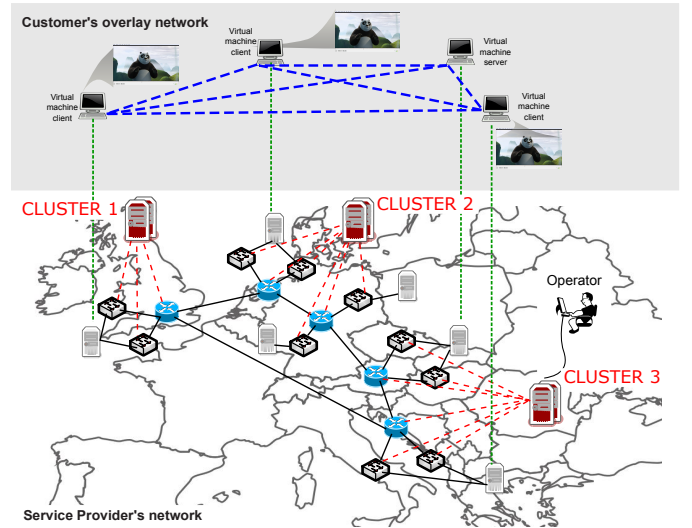

Fig. 2: Pilot setup

operator. During the demo the operator creates a couple of services, such as pseudo-wire links, that interconnect virtual machines located in different countries, using the ICONA dashboard. This operation is accomplished in a matter of hundred of milliseconds, so the machines are able to ping outright. Finally, the operator demonstrates how ICONA and ONOS can react to failures both in the Control and Data planes. To maximize the impact of the demo, a multimedia stream flowing from a server to multiple clients is shown. First two ONOS instances, with ICONA, are shutdown, and then a physical backbone link is cut off the network. After a short downtime, the video flows again smoothly.

REFERENCES

[1] "Geant project - http://www.geant.net/pages/default.aspx."

[2] D. Levin, A. Wundsam, B. Heller, N. Handigol, and A. Feldmann, "Logically centralized? state distribution trade-offs in software defined networks," in *HotSDN 2012*.

[3] S. H. Yeganeh, A. Tootoonchian, and Y. Ganjali, "On scalability of software-defined networking," in *IEEE Communications Magazine 51 (2), pp. 136-141*, 2013.

[4] K. Phemius, M. Bouet, and J. Leguay, "Disco: Distributed sdn controllers in a multi-domain environment," in *NOMS 2014*.

[5] Y. H. Y. Soheil, "Kandoo: a framework for efficient and scalable offloading of control applications," in *HotSDN 2012*.

[6] Y. G. A. Tootoonchian, "Hyperflow: a distributed control plane for openflow," in *2010 internet network management conference on Research on enterprise networking, pp. 3-3*, 2013.

[7] T. Koponen, M. Casado, N. Gude, J. Stribling, L. Poutievski, M. Zhu, R. Ramanathan, Y. Iwata, H. Inoue, T. Hama, and S. Shenker, "Onix: A distributed control platform for large-scale production networks," in *Proceedings of the 9th USENIX Conference on Operating Systems Design and Implementation*, 2010.

[8] P. Berde, M. Gerola, J. Hart, Y. Higuchi, M. Kobayashi, T. Koide, B. Lantz, B. O'Connor, P. Radoslavov, W. Snow, and G. Parulkar, "Onos: towards an open, distributed sdn os," in *HotSDN 2014*.

[9] "Hazelcast - http://www.hazelcast.com."

[10] "OpenFlow Switch Specification version 1.5.0." [Online]. Available: https://www.opennetworking.org/images/stories/downloads/sdn-resources/onf-specifications/openflow/openflow-switch-v1.5.0.noipr.pdf

[11] "Goff geant openflow testbed – https://openflow.geant.net/."