P7 (P4 Programmable Patch Panel): an instant 100G emulated network testbed in a pizza box

Burst Talk contributors: Fabricio E. Rodríguez Cesén (UNICAMP, Brazil), Marcos Felipe Schwarz (RNP, Brazil), Christian Esteve Rothenberg (UNICAMP, Brazil).

Abstract

Options to validate a network topology, including the link metrics, are traditionally based on virtual environments (e.g., Mininet), limiting the experiments with transmission speeds over 10Gbps. With the programmability that P4 brings to the network and the capabilities of new generation hardware supporting the PSA (Portable Switch Architecture) and TNA (Tofino Native Architecture) P4 architectures, it is possible to define emulation characteristics of the links and represent a network topology with high fidelity and computation power using a single physical P4 switch (e.g., Tofino). It is possible to emulate network topologies using recirculations, port configurations, different match+action tables, and even DAC cables. What is more, we can connect physical servers to inject traffic to the topology (see Figure 1). P7 comes out to the front from other network emulators [2] [1] [3], exploiting the limitations and gaps and bringing a friendly and robust environment for networking experiments.

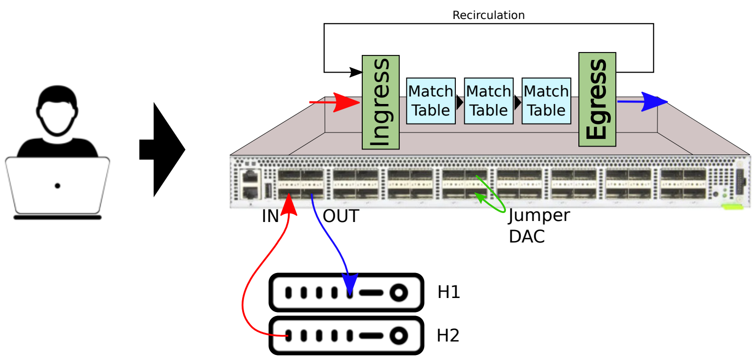

Figure 1: Pipeline

The emulation of the network link characteristics includes simple metrics (e.g., Connectivity, Latency [ms], bandwidth [Mbps]) and advance (e.g., Packet loss [%], Jitter [ms], Reordering [%]). These metrics are implemented in a P4/TNA environment to create a realistic network topology emulation, turning experiments in SDN testbeds rich and embracing. A list of characteristics that can be implemented in P7 and the different approaches that can be used to this end are detailed in table 1.

Table 1: P7 Link characteristics and implementation approaches

Link Connectivity	Jumper cabling with/without Vlan Tag Intern Recirculation + internal Tag
Latency [ms]	Using an internal timer TM + Pipelines recirculation
Jitter [ms]	Hash to determine recirculation times Lookup table with mathematical functions
Packet loss [%]	Hash to determine the probability to discard packets Random function to determine the probability to discard packets Realistic definition of packet loss
Re-ordering	TNA TM features Targeted probabilistic packet (ID) recirculation within a flow
Bandwidth	Rate limit TNA TM feature Port configuration and mapping

We can use P4/TNA characteristics such as packet generation and In-band Network Telemetry (INT) solutions. We can generate traffic of packets at rates up to 100G in each pipeline. It is possible to use this application as a service inside P7 (e.g., generation of background traffic, custom traffic generation). In addition, It is also possible to include native INT support inside the network. We can use this advantage to track different metrics, such as queue occupancy, and have a complete map of the network, including the state of all the components inside the emulated topology, to provide deeper visibility of user traffic performance across the distributed P7 testbed.

This demo features P7, a network emulator tool to test and validate network topologies and their characteristics in a 100G pizza box. Figure 2 shows an example of a topology that can be emulated using P7. In this topology, each link has en specific characteristic (i.e., BW, latency, packet loss). Physical servers connected to Tofino represent the hosts connected to the network. Figure 3 shows the representation of the topology in P7 and figure 4 shows where the P4 code will be focused (links characteristics). The link characteristics are a central part of the P4 program, and each link needs to be treated as an individual entity inside the P4 code with different configurations. Using recirculations and internal TAGs, it is possible to create the logic of all the different links, as long as the packets' paths. Figure 5 presents and overview of how the user defines a topology within P7 and then creates and generates the necessary files to run in a Tofino HW. We use Stratum to facilitate the ports and tables configuration. P7 will be released as an open source tool with all the documentation and examples to create and emulate different network topologies.

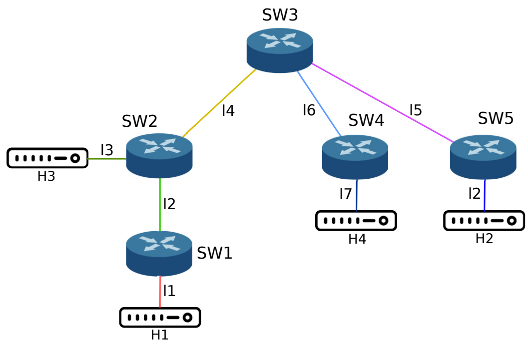

Figure 2: Network topology.

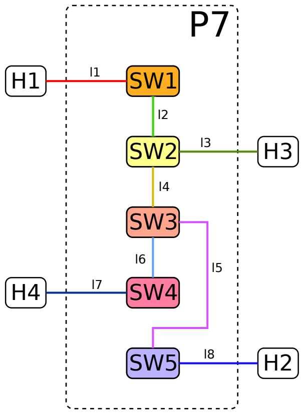

Figure 3: P7 topology.

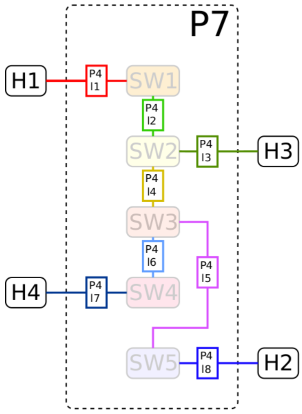

Figure 4: Links in P4.

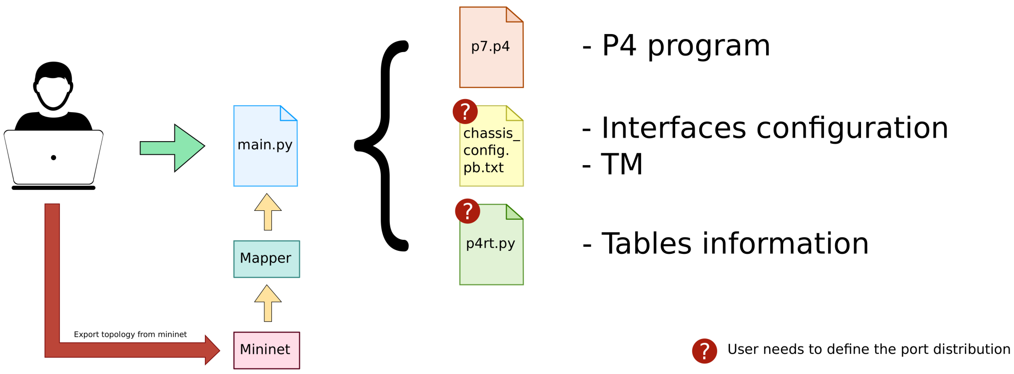

Figure 5: P7 global architecture

We need to consider some constraints and limitations, such as the limited memory, buffers, stages in Tofino, as long as the maximum size of the topology of network elements, to maintain the fidelity of the emulation. All in all, P7 will enable cost-effective advanced P4 programmable networking proposals by providing realistic network emulation capabilities starting with a pizza box deployment to large-scale distributed deployments. A tool that allows the user to emulate a network topology with realistic metrics is interesting in all the environments, including scientific research that requires rates up to 100Gbps for the evaluation tests. In addition, being a programmable high-fidelity testbed, P7 will facilitate repeatable and reproducible research. Together with the P7 emulation mechanisms, background traffic generation to generate congestion events in addition to INT will unveil novel insights into network-assisted congestion control and QoS possibilities.

References

[1] J. Cao, Y. Zhou, Y. Liu, M. Xu, and Y. Zhou. Turbonet: Faithfully emulating networks with programmable switches. In *2020 IEEE 28th International Conference on Network Protocols (ICNP)*, pages 1–11, 2020.

[2] P. G. Kannan, A. Soltani, M. C. Chan, and E.-C. Chang. BNV: Enabling scalable network experimentation through bare-metal network virtualization. In *11th USENIX CSET 18*, Baltimore, MD, Aug. 2018. USENIX.

[3] H. H. Liu, Y. Zhu, J. Padhye, J. Cao, S. Tallapragada, N. P. Lopes, A. Rybalchenko, G. Lu, and L. Yuan. Crystalnet: Faithfully emulating large production networks. In *26th SOSP*, NY, USA, 2017. ACM.