T4P4S & PIE

Towards an AQM Evaluation Testbed with P4 and DPDK

Sándor Laki and Péter Vörös
ELTE Eötvös Loránd University, Budapest, Hungary

T4P4S: A multi-target P4 compiler framework

T4P4S Compiler

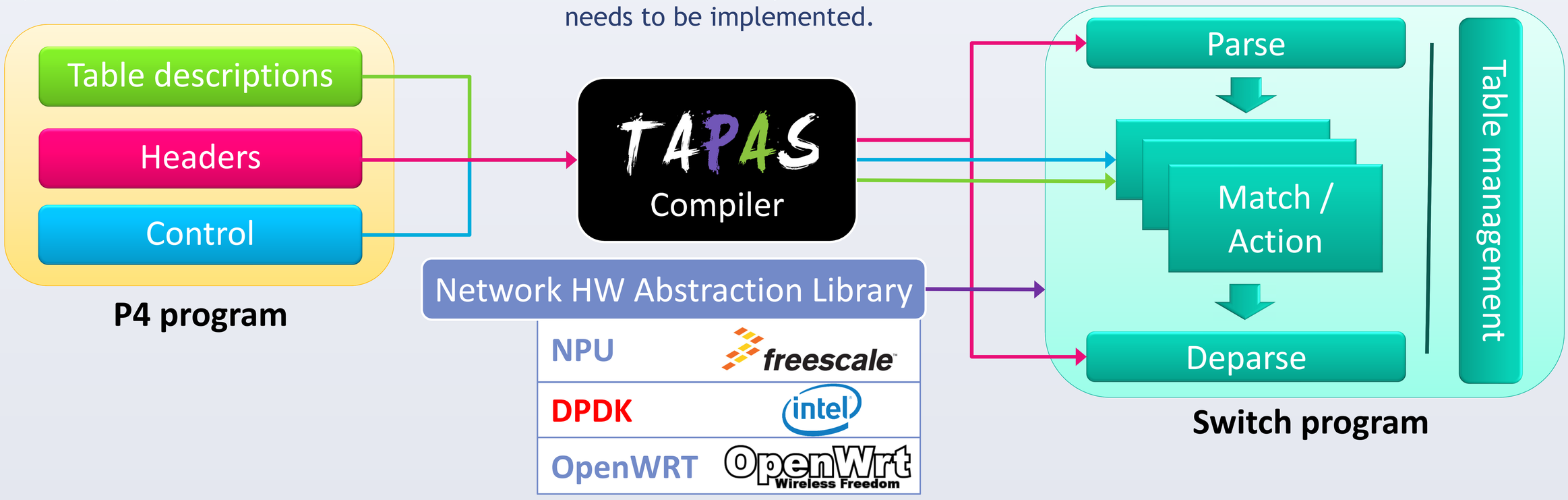
Our retargetable compiler (T4P4S - Translator for P4 Switches) turns a P4 code into a target independent C core program running on the top of a Network Hardware Abstraction Library (NetHAL).

NetHAL

Hardware dependent operations are separated to the Network Hardware Abstraction Library (NetHAL) which improves portability: to support a new architecture, only a new NetHAL needs to be implemented.

Switch program

To run the core program on a specific hardware the appropriate NetHAL needs to be linked. The compiled switch program then parse incoming packets, apply match-action rules and deparse messages before egressing.



AQM – Active Queue Management

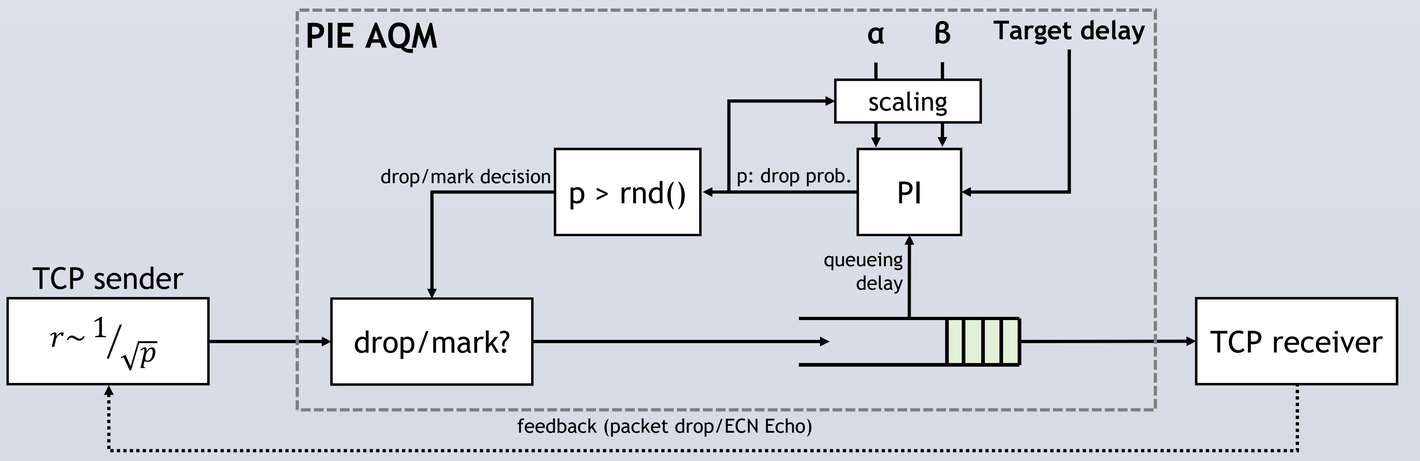
Problem: Standard loss-based TCP's congestion control plus large unmanaged buffers in Internet routers, switches, device drivers,... (a.k.a bufferbloat)

Cause: Latency issues for interactive/multimedia applications

Solution: AQM tries to signal the onset of congestion by dropping or marking packets.

AQM goals: **1)** Maintain low average queue/latency, 2) Allow occasional packet bursts, 3) Break synchronization among TCP flows

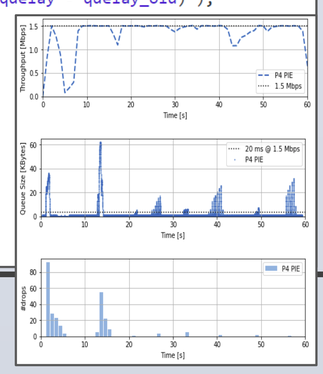
PIE AQM – How to describe drop policies in P4?



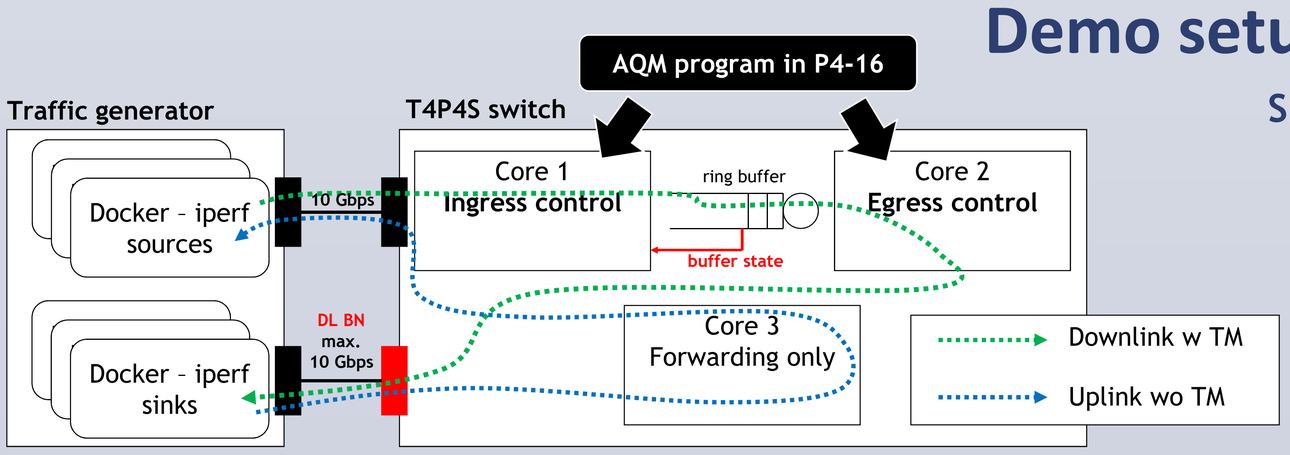

```
prob_reg.read(prob, 0);
time_next_reg.read(time_next, 0);
now = stdmeta.timestamp;
qdelay = stdmeta.qlatency;

if ( now >= time_next ) {
   /* update probability */
   qdelay_reg.read(qdelay_old, 0);
   prob_reg.read(prob, 0);
   delta = 0;

   delta = (int<64>) ( cAlpha * (qdelay - cTarget) );
   delta = delta + (int<64>) ( cBeta * (qdelay - qdelay_old) );
   delta = delta >> 8;
   if (prob < cMaxProb/1000) {
       delta = delta >> 5;
   }
   else if (prob < cMaxProb/100) {
       delta = delta >> 3;
   }
   else if (prob < cMaxProb/10) {
       delta = delta >> 1;
```

Complete P4 code

http://tiny.cc/pl0o5y



Demo setup



Small testbed deployed at our university with 2 nodes

- AMD Ryzen Threadripper 1900X
- Intel Corporation 82599ES 10-Gigabit Dual port NIC
- T4P4S + DPDK (https://github.com/P4ELTE/t4p4s)
- **AQM applied in Downlink (DL) Direction**
- **Emulated bottleneck (BN) at the outgoing link of T4P4S switch in DL direction**
- **Uplink traffic is not affected by AQM**

Eötvös Loránd University

http://p4.elte.hu - info@p4.elte.hu

communication networks laboratory

„Supported by the ÚNKP-18-4 New National Excellence Program of the Ministry of Human Capacities."