Extern Objects in P4: an ROHC Compression Case Study

Jeferson Santiago da Silva, François-Raymond Boyer, Laurent-Olivier Chiquette and J.M. Pierre Langlois
Computer and Software Engineering Department
Polytechnique Montréal, Québec, Canada
{jeferson.santiago-da-silva, francois-r.boyer, laurent-olivier.chiquette, pierre.langlois}@polymtl.ca

ABSTRACT
P4 is an emergent packet-processing language where the user can describe how the packets are to be processed in a switching element. This paper presents a way to implement advanced operations that are not directly supported in P4. In this work, two different ways to add extensions to P4 are presented: i) using new native primitives and ii) using extern instances. As a case study, an ROHC entity was implemented and invoked in a P4 program. The tests show that both methods are effective, with similar performance in terms of average packet latency. However, extern instances utilization suggests being more suitable for target-specific switching applications, where the manufacturer/vendor can specify its own specific operations without changes in the P4 semantics, exporting them to P4 as a standard vendor-specific library.

CCS Concepts
• Networks → Programming interfaces; Network architectures;

Keywords
Software-Defined Networking; P4; Packet-Processing Languages

1. INTRODUCTION
Software-Defined Networking (SDN) promises to fill the gaps of adaptability and scalability of current rigid network implementations. SDN networks decouple the data and the control planes, attempting to reduce the abstraction of the low-level implementations. Thus, in SDN-based networks, a major challenge is to represent the data/control flow implementation in a generic and multiplatform fashion. Several high-level packet processing languages have been developed in recent years [7, 9, 10, 11] to address this issue.
In this work, we have explored the functionalities of P4 (Programming Protocol-Independent Packet Processors) [4] - a reconfigurable, multi-platform, protocol- and target-independent packet processing language. P4 describes how the packets are to be processed in a forwarding element (FE) such as switches and routers. In P4, it is possible to implement the OpenFlow protocol [18], since each P4 program is able to specify which set of headers are analyzed in FE. This adaptability characteristic makes P4 suitable for SDN-based networks, once new protocols can be supported at the FE without changing the underneath hardware device.

The main goal of this work is to propose new extensions in P4. In our work we have presented two different ways for adding new P4 commands: i) using new native P4 primitives, and ii) using extern method calls. Our implementation was integrated and tested using the P4 switch model.

An ROHC entity - a header compression scheme used in wireless networks - was used as a case study of new extensions to P4. The tasks and resources analysis used in header compression have driven the implementation of new extensions to P4 since the current language constructs were not able to describe the entire process. In this work, we have attempted to reuse as much as possible the P4 infrastructure in order to make our design as portable as possible. The additional constructs were implemented in the C/C++ language and they were called directly in the P4 program by extending the P4 constructs. The basis for this work are the P4 implementation [6] and an open-source Linux ROHC library [2]. Our implementation is open-source and can be downloaded at GitHub [1].

1https://github.com/engjefersonsantiago/p4-programs
The rest of this paper is organized as follows: Section 2 presents a review of the literature, the methodology adopted in this paper is drawn in Section 3, the P4 support for the ROHC scheme is presented in Section 4, Section 5 shows the experimental results, and, Section 6 presents the conclusions and future works.

2. RELATED WORKS

SDN-based networks have emerged in recent years as a viable solution to deploy switched networks worldwide. Since in SDN the control and data planes are decoupled, the forwarding elements in this sort of networks become a set of actions applied to the packets based on pre-defined matching rules, which are configured by a centralized controller.

A forwarding engine uses a set of match+action tables in order to apply the routing rules for a given packet. These matching operations are normally performed over the header fields to discover the running protocol as well as the machine addresses involved in the communication. Therefore, an efficient way to describe the forwarding model has become extremely necessary.

Packet processing languages attempt to address this question. Song [16] presents the POF language: a protocol-oblivious packet processing language targeting network processors. As an evolution of the previous work, Song [17] uses the concept of abstract forwarding model to expand the POF supportability for various underneath architectures.

In [4], Bosshart proposes the P4 language, a protocol-independent packet processing language. P4 also defines its abstract forwarding model, as shown in Figure 1. A P4 program consists of a parser state machine (PSM) followed by a set of match+action tables, respectively in the ingress and egress pipeline, where the processing flow is controlled by an imperative control program.

The state transitions in the PSM are driven by the analysis of the header fields. In P4, the headers are defined in terms of their fields, avoiding error-prone bit-level manipulation. The supported matching operations are exact, range, ternary and wildcard. The P4 actions are procedures composed by native P4 primitives. The control program, in turn, defines in which order the tables are to be applied according to imperative statements.

Due to its simplicity, P4 has gained popularity as packet processing language for forwarding elements in both academia and industry [1, 11, 13, 12].

At the time of this writing, there are no known works in the literature addressing new extensions to P4 and its integration with external proprietary modules. Sivaraman [15] proposes a case study for the utilization of P4 switches for data center applications, where the author maps P4 deficiencies and proposes new primitives actions for P4. However, the new primitives are placed within the P4 switch core, not utilizing external libraries. Integrating P4 to extern standard libraries can represent a relevant step to represent more complex packet processing operations at the user abstraction level.

3. METHODOLOGY

This section presents the methodology applied in this work. In P4 it is possible to add new primitives in two different ways: as a new native primitive or through extern instances. Each way is presented in the following sections.

3.1 P4 Extensions: New Native Primitives

The P4 consortium provides a behavioural implementation of its abstract switch model described in the C++ language. The switch model is configured through a JSON file, which is generated by the P4 frontend compiler. The JSON file describes all the components declared in the P4 program. It also includes the PSM and the table dependency graph. The actions are also represented in the JSON file, including the native P4 primitives. The native primitives set is defined as a separate JSON array, which feeds the P4 frontend compiler. This JSON array has a correspondent description for the primitive actions in the P4 switch model.

To add new native primitives to P4, the user must edit the primitive JSON description, by adding the new primitive, including the primitive parameters. The behaviour model, in turn, must support the new primitive. Each P4 switch target has its own set of primitives, defined by the P4 team in a file named primitives.cpf. This code implements the behaviour of all supported P4 primitives, described as C++ functors. Then, to add the new primitive, a new operator must be included in this file, which should have the same name as the new P4 primitive, followed by the macro BM_REGISTER_PRIMITIVE(my_p4_primitive).

It is important to note that adding new native primitives to P4 incurs an increase in switch complexity. All

2Source code: https://github.com/engjefersonsantiago/behavioral-model/tree/master/targets/simple-switch
3.2 P4 Extensions: Extern Instances

The P4 team has already supported extern methods on the switch model and P4 compiler, but not formally released. However, the JSON file generation was not implemented and therefore not integrated with the P4 compiler. An implementation proposal in the gen_json.py - which is part of the P4 frontend compiler - was developed in this work for automatic JSON arrays generation based on the P4 extern description.

In Listing 1 the P4 reserved word `extern_type` defines the type of the extern instance. The keyword `attribute` refers to extern type initialization parameters, which at the time of this writing were just hexadecimal integers. To finalize the extern type definition, the list of the supported extern functions are defined by the word `method`, which can include several parameters, such as: metadatas, header fields and constant integer values. The extern instance is declared, as shown in the listing, using the keyword `extern` followed by the `extern_type` and then the extern instance name. The extern declaration also assigns the initialization attributes for the extern type. The extern methods are accessed by a given instance using an object-based notation: `extern_instance_name.method_a(method_parameters ...)`, which is called within a P4 action. Listing 2 represents the equivalent JSON array of the declared extern P4 object.

For proper operation of the extern methods, the target architectures must include the external modules. As the P4 behavioural model is a software-based switch, the extern types must be instantiated in the switch core. The external module - in this case, a C++ class - uses the P4 behavioural model action class to register the extern methods as new P4 primitives in the switch. Therefore, when these extern actions are reached in a given switch pipeline, the P4 model knows it is an external call, and invokes the correct extern method. The Listing 3 shows equivalent C++ class for the extern type declared above in P4. Both ingress and egress pipelines are able to access the extern methods.

The extern class is declared as a child class of the `ExternType`, provided by the P4 behavioural model, and it is compiled along with the target switch. The macros `BM_EXTERN_ATTRIBUTES` and `BM_EXTERN_ATTRIBUTE_ADD` are used to link the C++ initialization variables with those that are defined in the P4 declaration.
ation and passed to the switch model as a JSON array. The \texttt{init} function is not a P4 exported method, it is mandatory and it is used to initialize the stateful classes and variables of the extern instance. The list of methods follows the \texttt{init} function definition. The extern type is registered in the switch model through the macro \texttt{BN\_REGISTER\_EXTERN} and the methods are set as new P4 primitives by the macro \texttt{BN\_REGISTER\_EXTERN\_METHOD}. A dummy function is then declared and it must be called in the target, as extern, to avoid linker errors.

Aiming to facilitate the manipulation of the extern types with P4, a template script generator was developed. The script receives as parameters a configuration file, where the user can specify the extern characteristics, such as extern type name, extern initialization attributes and the methods that will be exported as P4 primitives. As outcome, the template script generates a C++ template to be instantiated in the switch target and a P4 description that specifies the extern types for the P4 frontend compiler.

4. ROHC SUPPORT IN P4

In this section, we present RObust Header Compression (ROHC) and how to implement its operations in a P4-compatible switch.

Streaming, real-time and on-demand services have occupied a large portion of the network bandwidth due to the increasing number of users utilizing those applications in recent years. Voice over IP (VoIP) is an example of a real-time application that demands high-bandwidth allocation in order to guarantee high Quality of Service (QoS) indicators. Real-time (RT) applications have one factor in common: the useful payload size for each packet is normally much smaller than the length of the combined header fields which are necessary for delivering a packet.

The average payload size of a Real-time Protocol (RTP) packet is normally around 20 bytes. This protocol operates over the User Datagram Protocol (UDP), that operates over either the Internet Protocol Version 4 (IPv4) or Version 6 (IPv6). The IPv4 overhead is at least 40 bytes, while the overhead for IPv6 is up to 60 bytes, which means an efficiency of 33\% and 25\%, for IPv4 and IPv6, respectively. These real-time services, in many cases, are used through mobile devices connected to 3G/4G networks. In mobile networks, such as Long-Term Evolution (LTE), the low efficiency becomes a system bottleneck, since the available bandwidth between the mobile equipment and the base-station is very limited.

Header compression schemes address this efficiency issue by not sending with each packet redundant or easily predictable header information. Considering a flow of packets belonging to the same application, many header fields can be considered static, for instance, IP addresses and TCP/UDP ports, and these fields are not transmitted at all. Other fields change in a predictable way, such as sequence numbers, and the compressor entity can just transmit the difference between the current and the previous packet belonging to the same flow. Fields such as checksums and Cyclic-Redundancy Check (CRC) cannot be compressed and should be transmitted as is. Figure 2 illustrates a data transfer using a header compressing scheme over an Ethernet network.

Many header compression schemes have been standardized in the last 20 years. In this work, we have focused on the RObust Header Compressor (ROHC) scheme \cite{9} since it is used as a basis for header compression in LTE networks. Drafts for next-generation cellular communication, such as 5G, bet that mobile network equipments will explore concepts of SDN and NFV (Network Functions Virtualization), including base-stations and forwarding elements \cite{5,10}. In this context, the integration of header compression schemes and the SDN networks becomes more and more relevant.

Packet header compression/decompression involves several steps. Among them we highlight packet profiling and data context maintenance. Profiling is related to which type of compression technique will be applied to the headers, depending on the protocols involved in the communication. The ROHC standard defines up to 15 profiles supported for header compression. Profiling packets in P4 can be done through the parser state machine (PSM), where a profile identifier is assigned according to the received protocols.

Context maintenance is a more complex task, where new packet flows must be assigned to a new data context while similar data flows are analyzed in order to determine whether they belong - or not - to an existent context. This context manipulation involves the matching operation of several header fields, depending on the profile. P4 allows the user to perform several match types in tables, including exact match. However, P4 does not support table self-updating, which is necessary for context creation. Moreover, stateful processing is still immature with P4, limited to registers whose implementation does not support matching operations. Due to these limitations, in this work the context tables and matching engines were kept outside the P4 implementation, copying the current packet to the ROHC compressor and decompressor entities.

![Figure 2: ROHC compressor scheme.](image-url)
As explained in the previous sections, P4 is a language that describes how a packet is to be processed in a forwarding element, such as a router or a switch. P4 - by itself - does not allow modifications of the content of a payload within a packet, which is necessary for packet header compression. Since the size of the compressed header is unknown until starting decompressing, it should be treated as part of the payload since P4 is just able to extract fixed-size headers or those headers whose the size is stored in a field of the current extracted header.

Then, to perform both header compression and decompression, the packet payload should be modified and manipulated in the P4 behavioural model, a C++ implementation of a software switch. To support the ROHC schemes, the open-source ROHC library was compiled as a static library and imported into the C++ switch model. This P4 forwarding model describes basic switch primitives that are linked to the available P4 primitives. P4 calls are mapped into the switch procedures by the P4 compiler, which generates a JSON file containing the packet processing flow and the actions called by the P4 program. In this work, three new primitives were added: ROHC compression, ROHC decompression and modify and resubmit a packet.

The compression operation starts during the PSM - what was described in P4 - when a packet arrives at the switch. The PSM sweeps all headers in order to define which ROHC profile is going to be compressed, storing it into a packet metadata. This metadata is then used by the new P4 primitive - defined in P4 as rohc_comp_header. The headers supported by the ROHC algorithm are translated into an acceptable ROHC data structure - in this case, an array of chars - and then it is passed to the ROHC compressor entity. The compressor engine performs the operation, returning a byte stream, that are placed at the beginning of the packet payload. The original packet headers are invalidated and they are not transmitted. The last step in the compression task is to differentiate the ROHC packets at the lowest layer protocol. As this work is based on Ethernet networks, an unused Ethertype value was assigned for ROHC packets for proper packet identification.

Similar to the compressor, the decompression is triggered by the PSM, where the parser identifies the special Ethertype. Once identified, the decompression takes place. A very similar data structure conversion is applied to the packet payload before passing it to the decompressor engine, named rohc_decomp_header in P4. The decompressor then performs full headers recovery, placing them at the beginning of the payload field. The last steps in the compression are to remove the former compressed header and to assign a valid Ethertype value in the Ethernet header, in this case, 0x0800 (IP protocol version 4).

Since P4 describes how a packet is processed and forwarded, the operations of header compression and decompression in a packet have no value if the forwarding rules cannot be applied to these packets. Let’s consider a decompressing operation: i) a packet arrives at the switch; ii) the packet is identified as an ROHC packet; iii) let’s assume the packet is correctly decompressed; iv) the decompressed packet is processed and forwarded according to the packet rules; and v) the packet is transmitted. However, for step iv), the packet headers must be part of the parser graph in order to be correctly extracted. To handle this, it is necessary to re-parse this new packet and the solution adopted in this case was to send it back to ingress parser. For this, a new P4 primitive was created - modify_and_resubmit - instead of using the existing recirculate primitive, aiming at reducing the packet latency since the packet recirculation is applied at the end of the egress pipeline.

For the modify_and_resubmit primitive, the ingress pipeline was modified in the P4 switch model. The existing resubmit primitive forwards back a packet to the parser. However, all the modifications performed on the packet in the ingress pipeline are not applied. To allow the packet modifications, this implementation added a deparser entity before resubmitting a packet, that takes the updated header fields and the packet payload, serialize them into a stream of bytes. For proper packet resubmission, the packet length was also updated and stored in the switch standard metadata. Figure 3 shows the proposed switch model.

5. EXPERIMENTAL RESULTS

In order to validate the proposed method, an emulated three-node network was created. The simple network is composed of two hosts connected to each other through a P4 switch, according to Figure 3. A Python script was used to send packets from host A to host B, passing through the switch. The packets sent by host A are compressed. The switch then uncompresses the packets, performs the forwarding procedure, recompresses the packets and finally, it forwards the packets to B. As emulation platform, a 64-bit Ubuntu OS version 14.04.1 was used, running on a dual-core AMD processor @ 1.5 GHz, 64 kB of L1/i/d cache, 512 kB of L2 cache and 3 GB of RAM memory.

The new P4 primitives were tested through the fol-
following scenarios:
- using new native primitives;
- using extern methods.

The native primitives were included directly in the P4 behavioural model, while the extern methods were instantiated in the P4 soft-switch. Both scenarios were tested in two different conditions. One of them was using the current recirculate method of P4 to send the packets back to the parser state machine. In the second condition, the recirculate action was replaced by the new modify and resubmit primitive. It is important to highlight that both recirculate and modify and resubmit actions are equivalent in terms of functionality, the difference between them resides pipeline stage the action is applied.

For testing purposes, a set of 10,000 compressed ROHC packets was generated, which were sent from host A to host B. The original uncompressed packets have 74-bytes length, including 20-bytes of RTP payload. Table 1 presents the average computing time (ACT) for the compression and decompression operations using both scenarios. It is clear the ROHC compressing performance has approximately the same results regardless of the adopted scenario. The results shown in Table 1 were obtained using the new modify and resubmit primitive since the recirculation method has no impact on the measures of isolated primitive latencies.

### Table 1: ROHC Compressing Execution Time

| Operation  | Extern Object | ACT [µsec] |
|------------|---------------|------------|
| Compression| √             | 1586 ± 1   |
| Decompression| √            | 1263 ± 1   |
| Decompression| √             | 1610 ± 1   |
| Decompression| √             | 1313 ± 1   |

The average packet latency (APL) is presented in Table 2. The considered packet latency metric is the time taken to receive an uncompressed packet, perform the packet modifications, recompress and forward it to the destination host.

Considering the data from Table 2, the latency is lower when using the native primitive with the modify and resubmit method. The egress pipeline processing latency can explain the latency reduction of the scenarios that utilize the new modify and resubmit primitive since this primitive action is applied in the ingress pipeline and then the packet does not traverse the egress pipeline before being sent back to the parser.

Even though native primitive scenarios achieve higher performance than those using extern objects, it introduces compatibility and modularity problem. Besides, an intrinsic drawback in this scenario is that any modification in the new primitive requires a recompilation of the whole set of primitives since they belong to the same C++ file in the switch model. Proposing extensions through this method is not feasible for the P4 language maintenance since at each new primitive a new P4 keyword is created and the P4 compiler must know what to do with this new keyword. The use of extern objects, in turn, presents a modular and scalable way to describe specific features with P4 using target-specific P4 methods defined in a standard library.

### 6. CONCLUSION AND FUTURE WORK

In this work, we presented how to add new extensions to P4. Two different methods were presented: by adding new native primitives and using extern methods. Both methods were tested using the simple switch target provided by the P4 Consortium and they showed effective. As a case study, an ROHC compressing entity was used due to the relevance of compressing techniques in the field of mobile cellular communications, where the bandwidth availability is limited. Both compressor and decompressor engines were integrated into a P4 program, where the user can decide whether using or not the header compression scheme.

We highlight in this work the use of extern instances in P4 and the possibility of its integration with proprietary libraries, which the objective of keeping the natural simplicity of the language. Through extern methods, the switch manufacturers can develop their more advanced operations the way it is more convenient for their applications: a software implementation or a hardware accelerator, for instance.

As future work, we aim to explore other P4 compatible switch platforms and expand the use of extern instances on these platforms. It is also an objective of future research the utilization of hardware accelerators described as extern methods in P4 in order to perform timing critical operations, such as packet classification and deep packet inspection.

### 7. ACKNOWLEDGMENTS

This work was carried out with the support of the National Council for Scientific and Technological Development (CNPq) - Brazil and Ericsson Canada.
8. REFERENCES

[1] S. Abdi, U. Aftab, G. Bailey, B. Boughzala, F. Dewal, S. Parsazad, and E. Tremblay. Pfpsim: A programmable forwarding plane simulator. In Proceedings of the 2016 Symposium on Architectures for Networking and Communications Systems, ANCS ’16, pages 55–60, New York, NY, USA, 2016. ACM.

[2] D. Barvaux. rohc. https://github.com/didier-barvaux/rohc, 2016.

[3] N. Bonelli, S. Giordano, G. Procissi, and L. Abeni. A purely functional approach to packet processing. In Proceedings of the Tenth ACM/IEEE Symposium on Architectures for Networking and Communications Systems, ANCS ’14, pages 219–230, New York, NY, USA, 2014. ACM.

[4] P. Bosshart, D. Daly, G. Gibb, M. Izzard, N. McKeown, J. Rexford, C. Schlesinger, D. Talayco, A. Vahdat, G. Varghese, and D. Walker. P4: Programming protocol-independent packet processors. SIGCOMM Comput. Commun. Rev., 44(3):87–95, July 2014.

[5] A. Bradai, K. Singh, T. Ahmed, and T. Rasheed. Cellular software defined networking: a framework. IEEE Communications Magazine, 53(6):36–43, June 2015.

[6] P. Consortium. P4. https://github.com/p4lang, 2016.

[7] R. Duncan and P. Jungck. packetc language for high performance packet processing. In High Performance Computing and Communications, 2009. HPCC ’09. 11th IEEE International Conference on, pages 450–457, June 2009.

[8] N. Foster, A. Guha, M. Reitblatt, A. Story, M. J. Freedman, N. P. Katta, C. Monsanto, J. Reich, J. Rexford, C. Schlesinger, D. Walker, and R. Harrison. Languages for software-defined networks. IEEE Communications Magazine, 51(2):128–134, February 2013.

[9] N. Foster, R. Harrison, M. J. Freedman, C. Monsanto, J. Rexford, A. Story, and D. Walker. Frenetic: A network programming language. SIGPLAN Not., 46(9):279–291, Sept. 2011.

[10] L.-E. Jonsson, G. Pelletier, and K. Sandlund. The ROBust Header Compression (ROHC) Framework. RFC 4995, Internet Engineering Task Force, July 2007.

[11] P. Li and Y. Luo. P4gpu: Accelerate packet processing of a p4 program with a cpu-gpu heterogeneous architecture. In Proceedings of the 2016 Symposium on Architectures for Networking and Communications Systems, ANCS ’16, pages 125–126, New York, NY, USA, 2016. ACM.

[12] B. Pfaff, J. Pettit, T. Koponen, E. J. Jackson, A. Zhou, J. Rajahalme, J. Gross, A. Wang, J. Stringer, P. Shelar, K. Amidon, and M. Casado. The design and implementation of open vswitch. In Proceedings of the 12th USENIX Conference on Networked Systems Design and Implementation, NSDI’15, pages 117–130, Berkeley, CA, USA, 2015. USENIX Association.

[13] C. Ramirez-Perez and V. Ramos. Sdn meets sdr in self-organizing networks: fitting the pieces of network management. IEEE Communications Magazine, 54(1):48–57, January 2016.

[14] M. Shahbaz, S. Choi, B. Pfaff, C. Kim, N. Feamster, N. McKeown, and J. Rexford. Pisces: A programmable, protocol-independent software switch. In Proceedings of the 2016 Conference on ACM SIGCOMM 2016 Conference, SIGCOMM ’16, pages 525–538, New York, NY, USA, 2016. ACM.

[15] A. Sivaraman, C. Kim, R. Krishnamoorthy, A. Dixit, and M. Budiu. Dc.p4: Programming the forwarding plane of a data-center switch. In Proceedings of the 1st ACM SIGCOMM Symposium on Software Defined Networking Research, SOSR ’15, pages 2:1–2:8, New York, NY, USA, 2015. ACM.

[16] H. Song. Protocol-oblivious forwarding: Unleash the power of sdn through a future-proof forwarding plane. In Proceedings of the Second ACM SIGCOMM Workshop on Hot Topics in Software Defined Networking, HotSDN ’13, pages 127–132, New York, NY, USA, 2013. ACM.

[17] H. Song, J. Gong, H. Chen, and J. Dustzadeh. Unified pof programming for diversified sdn data plane devices. ICNS 2015, page 106, 2015.

[18] The Open Networking Foundation. OpenFlow Switch Specification, Dec. 2014.

[19] The P4 Language Consortium. The P4 Language Specification, Jan. 2016.