NOKIA Bell Labs

Technical University of Munich

Towards Understanding the Performance of P4 Programmable Hardware

EuroP4 ' 19, Cambridge, UK
September 23, 2019

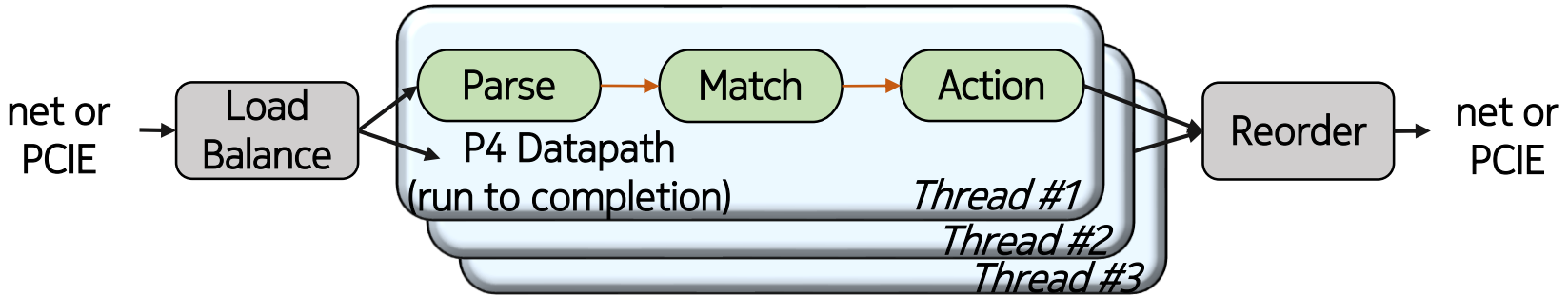
Hasanin Harkous
Michael Jarschel, Mu He, Rastin Pries, Wolfgang Kellerer

NOKIA Bell Labs

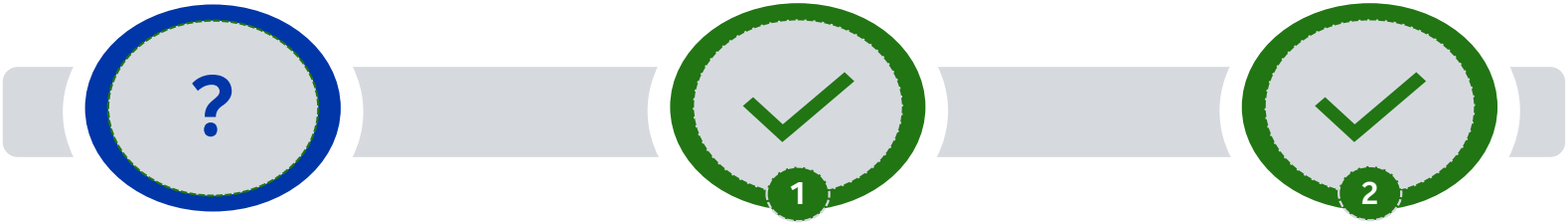
Background

- P4 abstracts the packet processing pipeline into stages:
  - Parser
  - Control Blocks
  - Deparser

- Netronome SmartNIC:



Introduction



What is the relation between packet processing latency and a certain P4 pipeline structure?

Analyze the impact of a basic set of P4 constructs on packet processing latency to derive the influential parameters.

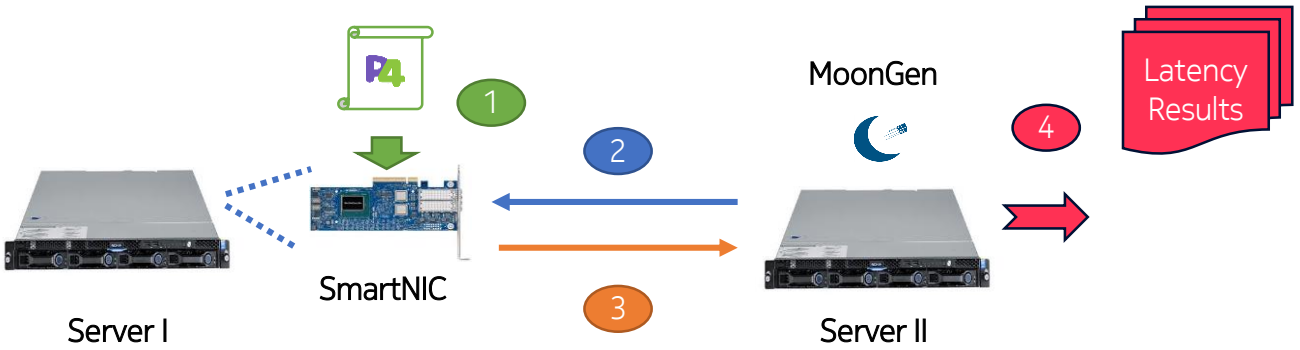
Propose a method for estimating the packet latency of P4-based network functions.

NOKIA Bell Labs

Experiments & Measurement Setup

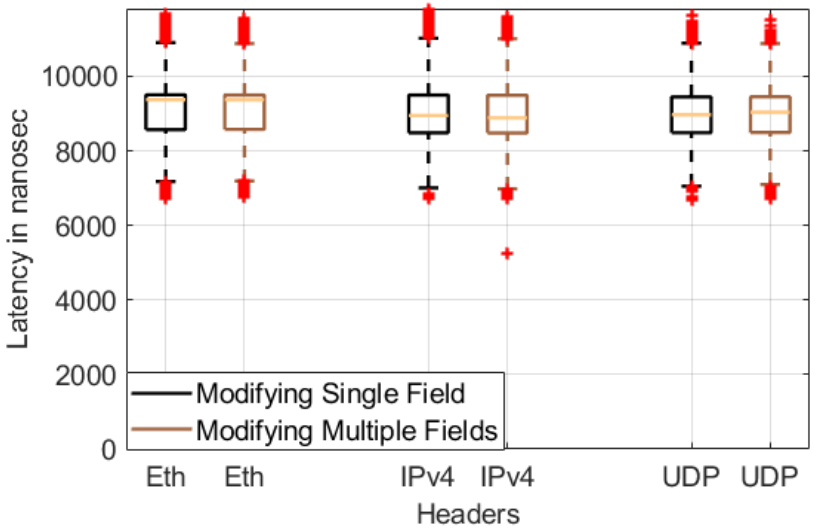
Experiments:

1) Modifying a single field of a header versus modifying multiple fields.
2) Executing arithmetic and binary operations in P4 actions.
3) Parsing and modifying a different number of headers.
4) Adding more tables into P4 pipeline.



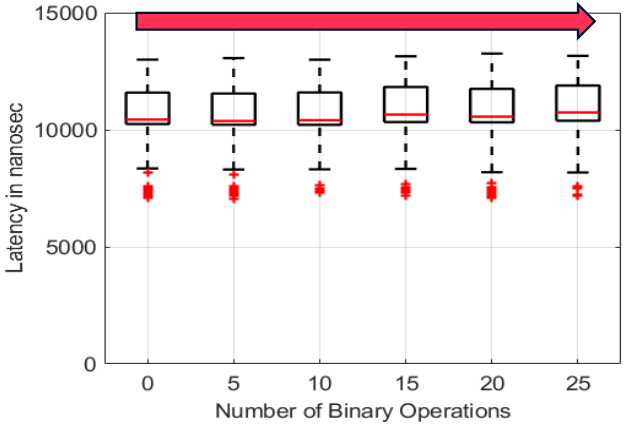
Results
Header Fields Modification

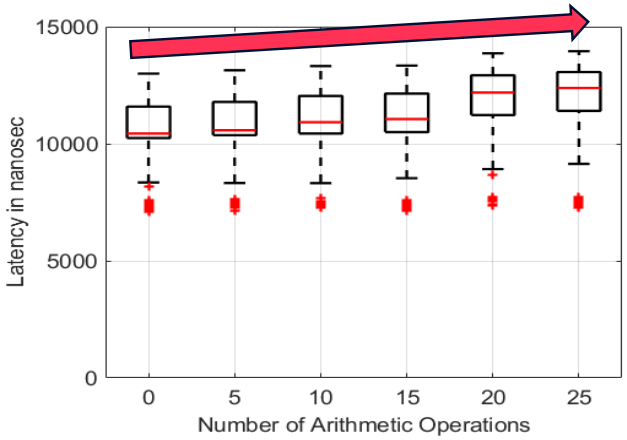
- **Design Objective:** Study the effect of modifying a different number of fields of the same header.



✓ **Observation:** Modifying a single field of a header has the same impact on latency as modifying multiple fields.

Results
Operations Execution





- **Design Objective:** Study the latency cost of applying binary and arithmetic operations.

✓ **Observation:** The latency of binary operations can always be neglected while that of arithmetic operations should be considered only if a significant number of operations is applied.

NOKIA Bell Labs

Results
Headers Parsing and Modification

- **Design Objective:** Study the impact of header parsing and header modification on the processing latency.
- Case ID : **A 1**

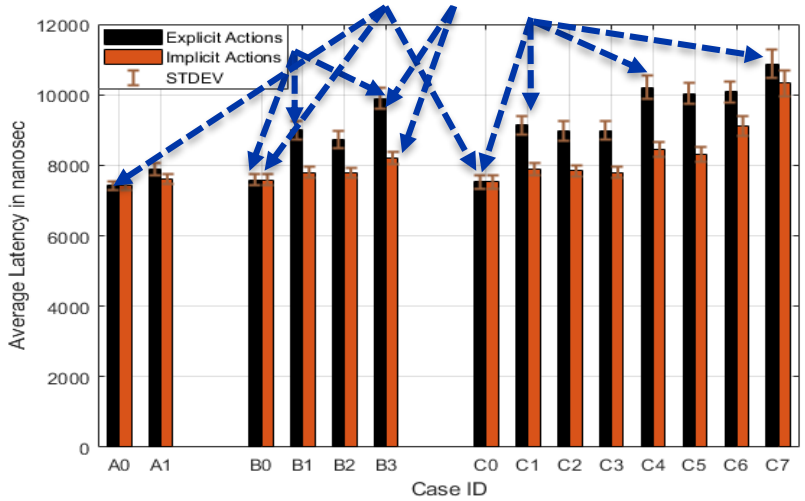
- Headers modification was examined within explicit and implicit actions.

	Case ID	Parsed Headers			Modified Headers		
		Eth	IPv4	UDP	Eth	IPv4	UDP
Eth Parsing	A0	+	-	-	-	-	-
	A1	+	-	-	+	-	-
Eth+IPv4 Parsing	B0	+	+	-	-	-	-
	B1	+	+	-	+	-	-
	B2	+	+	-	-	+	-
	B3	+	+	-	+	+	-
Eth+IPv4 + UDP Parsing	C0	+	+	+	-	-	-
	C1	+	+	+	+	-	-
	C2	+	+	+	-	+	-
	C3	+	+	+	-	-	+
	C4	+	+	+	+	+	-
	C5	+	+	+	+	-	+
	C6	+	+	+	-	+	+
	C7	+	+	+	+	+	+

NOKIA Bell Labs

Results
Headers Parsing and Modification

- ✓ **Observations:**
  - ✓ The impact of parsing additional headers is negligible.
  - ✓ Latency cost of modifying additional headers is clearly observable.
  - ✓ With identical parsing blocks, the latency varies according to the number of modified headers.
  - ✓ Explicit actions lead to more latency compared to implicit actions.



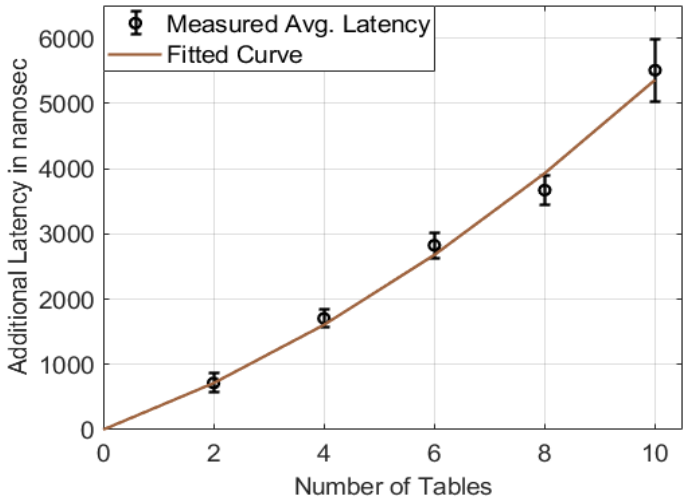
Number of Parsed Headers	1	2		3		
Number of Modified Headers	1	1	2	1	2	3
Explicit Latency (in ns)	7900	8900	9900	9000	10100	10900
Implicit Latency (in ns)	7600	7800	8200	7800	8600	10300

Table1: Average measured latency as a function of the number of parsed and modified headers.

NOKIA Bell Labs

Results
Tables Scaling

- **Design Objective:** Quantify the latency cost of adding tables into a P4 pipeline.



✓ **Observations:**

- ✓ Latency increases as more tables are added to the pipeline.
- ✓ $f(\gamma) = 22.44 \times \gamma^2 + 311.6 \times \gamma \quad \text{for } \gamma = 0, \ldots, 10$ (Eq.1)

where f (.) is the additional latency in ns, and γ is the number of added tables.

NOKIA Bell Labs

Estimation Method

Given a P4 program:

1) Extract the following parameters:
   i. α : The number of parsed headers.
   ii. β : The number of modified headers.
   iii. γ : Number of tables minus one.
2) Read the estimated latency due to headers parsing and modification from Table 1.

α	1	2		3		
β	1	1	2	1	2	3
Explicit Latency (in ns)	7900	8900	9900	9000	10100	10900
Implicit Latency (in ns)	7600	7800	**8200**	7800	8600	10300

3) Evaluate the estimated latency due to adding tables based on equation 1.

$$f(\gamma) = 22.44 \times \gamma^2 + 311.6 \times \gamma \quad \text{for } \gamma = 0, \ldots, 10 \qquad \text{(Eq.1)}$$

4) The estimated average latency equals the sum of latencies evaluated in steps (2) and (3).

Example: *L3_Forwarding*

1) Parameters:
   i. α=2,
   ii. β=2
   iii. γ=0
2) L1=8200 ns
3) f(0)=0 ns

4) $\hat{L} = 8200 + 0 = 8200ns$

NOKIA Bell Labs

Estimation Method
Validation

- Network Functions:
  - L3_forwarding: $\alpha=2$, $\beta=2$, and $\gamma=0$.
  - L3_forwarding + UDP-based Firewall: $\alpha=3$, $\beta=2$, and $\gamma=1$.

Network Function	Estimated Average Latency	Measured Average Latency	ΔL
L3_Fwd	8200 ns	8387 ns	**187 ns**
L3_Fwd+Firewall	8957 ns	9022 ns	**65 ns**

Conclusion & Future Work

Conclusion

- Identified a relationship between P4 pipeline complexity and packet latency.
- Derived influential parameters:
  - Number of parsed headers
  - Number of modified headers
  - Number of tables
- Proposed a method for estimating the packet latency and validated it.

Future Work

- Study the impact of other P4 constructs such as adding/removing headers, etc.
- Perform measurements on other P4 targets such as software switch, etc.

Technical University of Munich

Thank you for your attention!

NOKIA Bell Labs