university of groningen

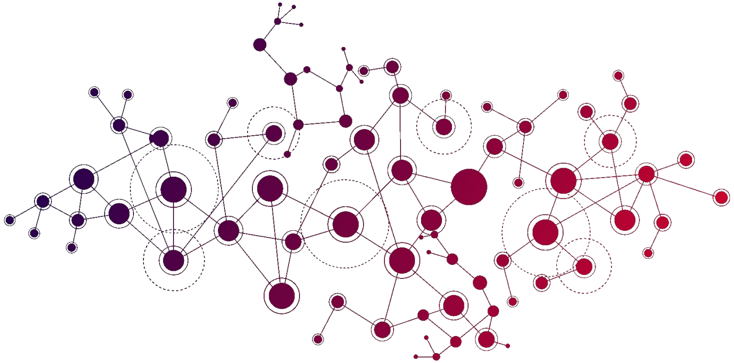
faculty of science and engineering

bernoulli institute

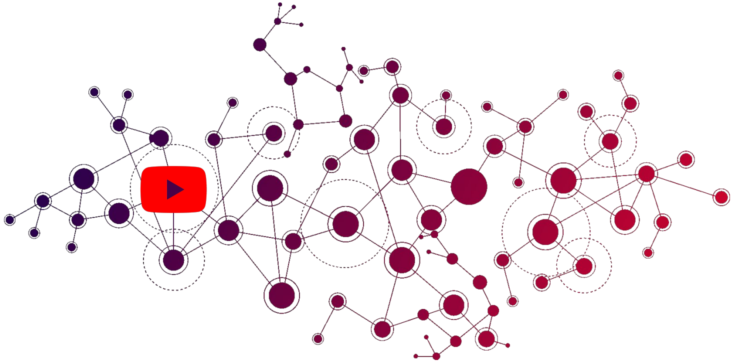
STATEFUL DATA ANALYTICS OVER PROGRAMMING MODELS OF NETWORKS

Author : Gheorghe Pojoga
Supervisors : Prof. Dr. Boris Koldehofe, Bochra Boughzala

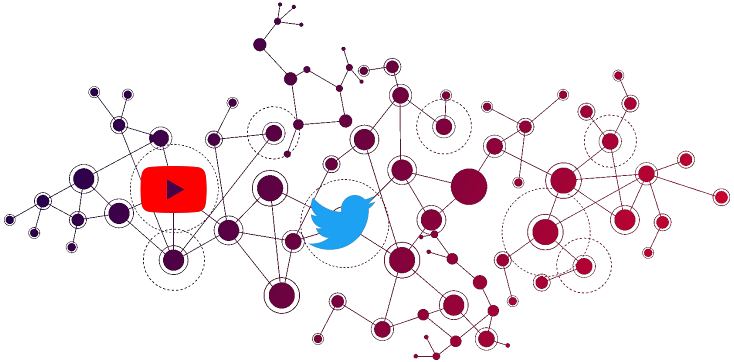
Networking - a way of life



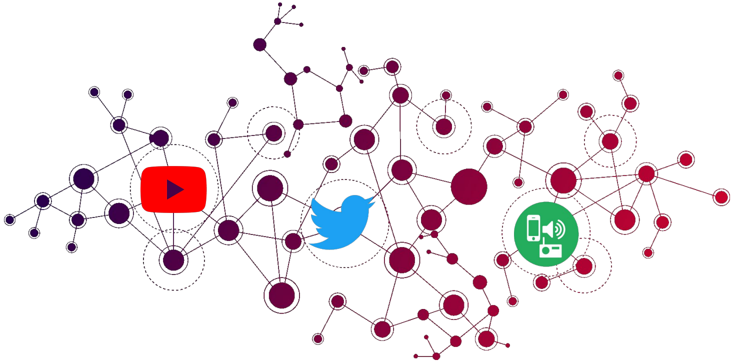
Networking - a way of life



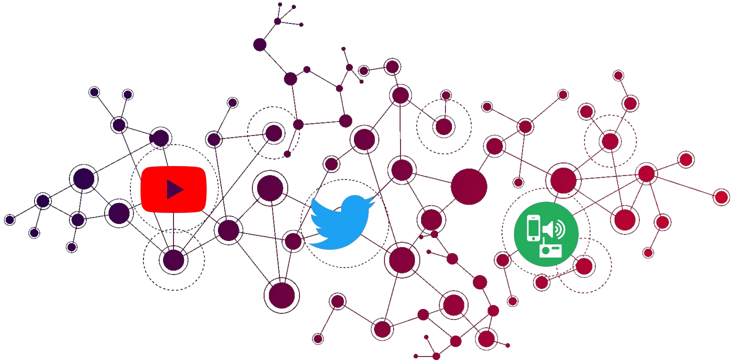
Networking - a way of life



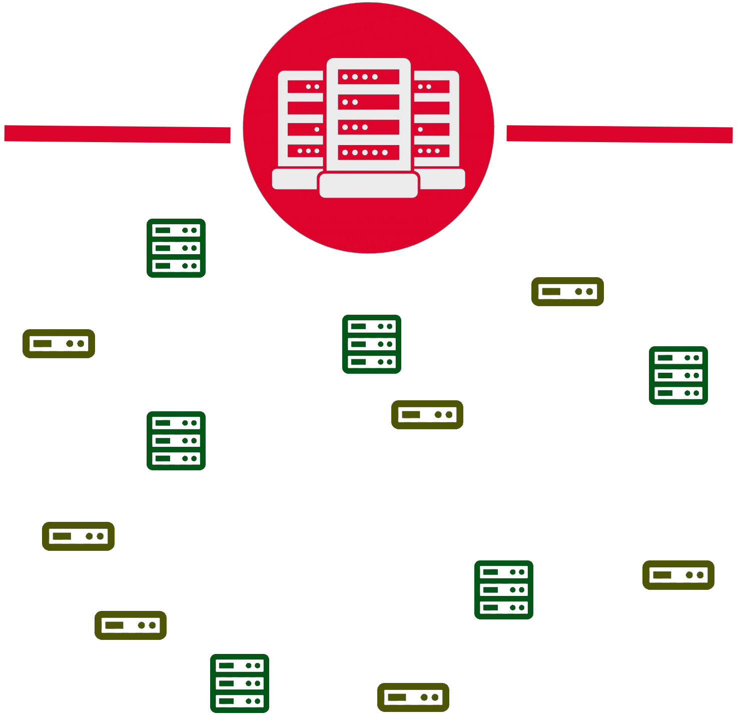
Networking - a way of life

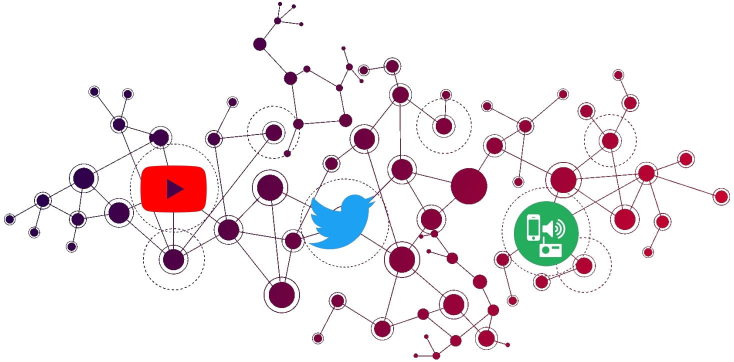


Networking - a way of life

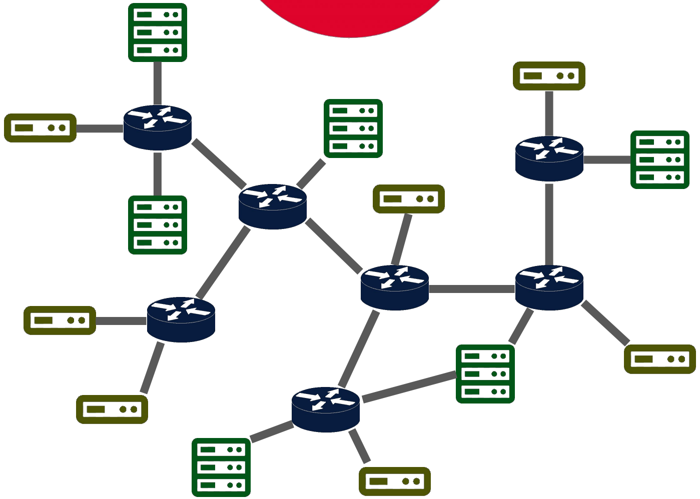


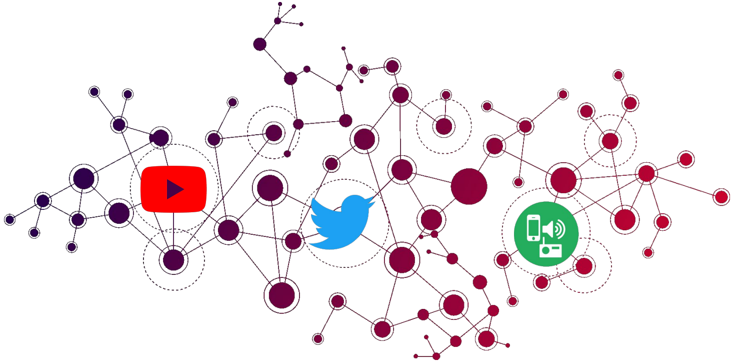
Networking - a way of life



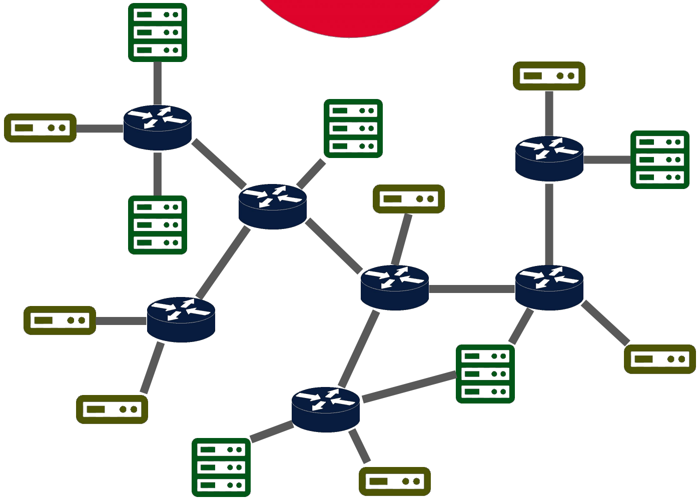


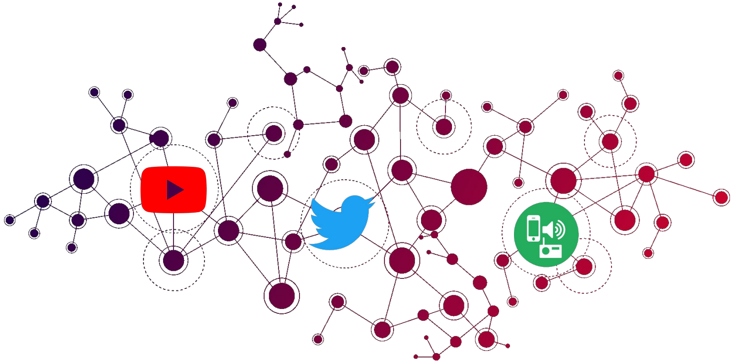
Networking - a way of life



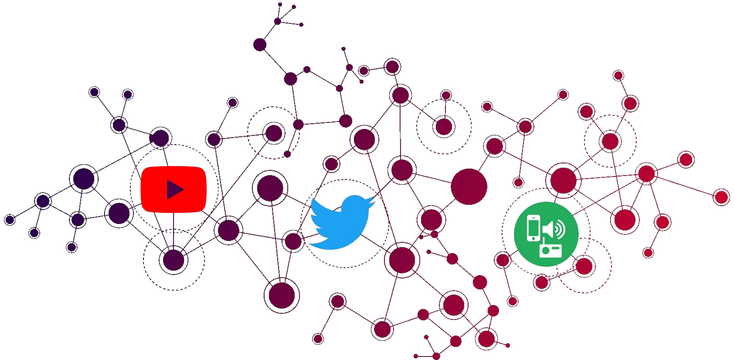


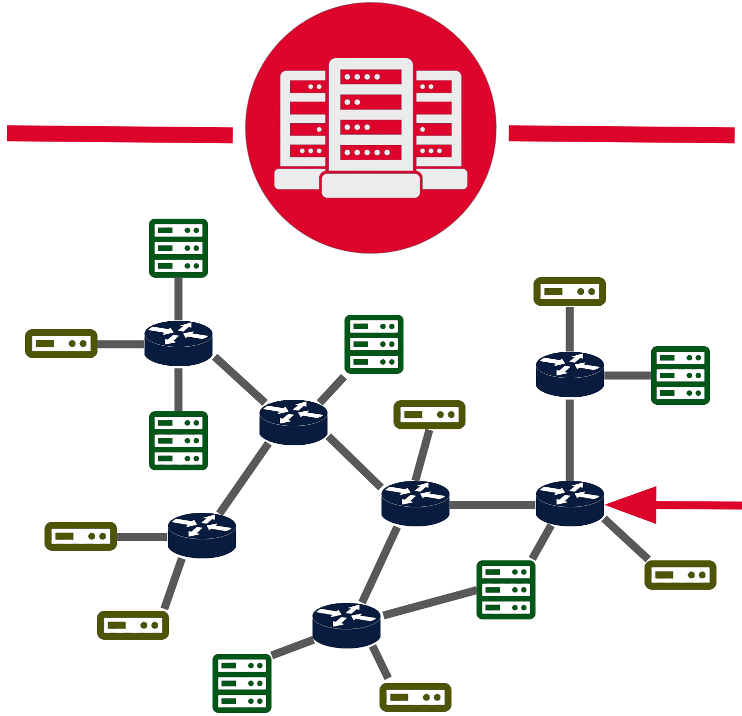
Networking - a way of life

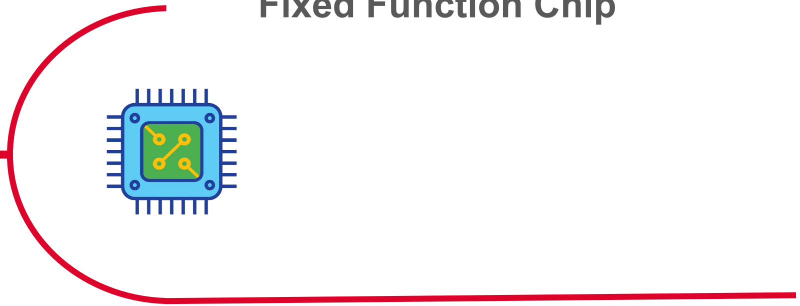




Networking - a way of life

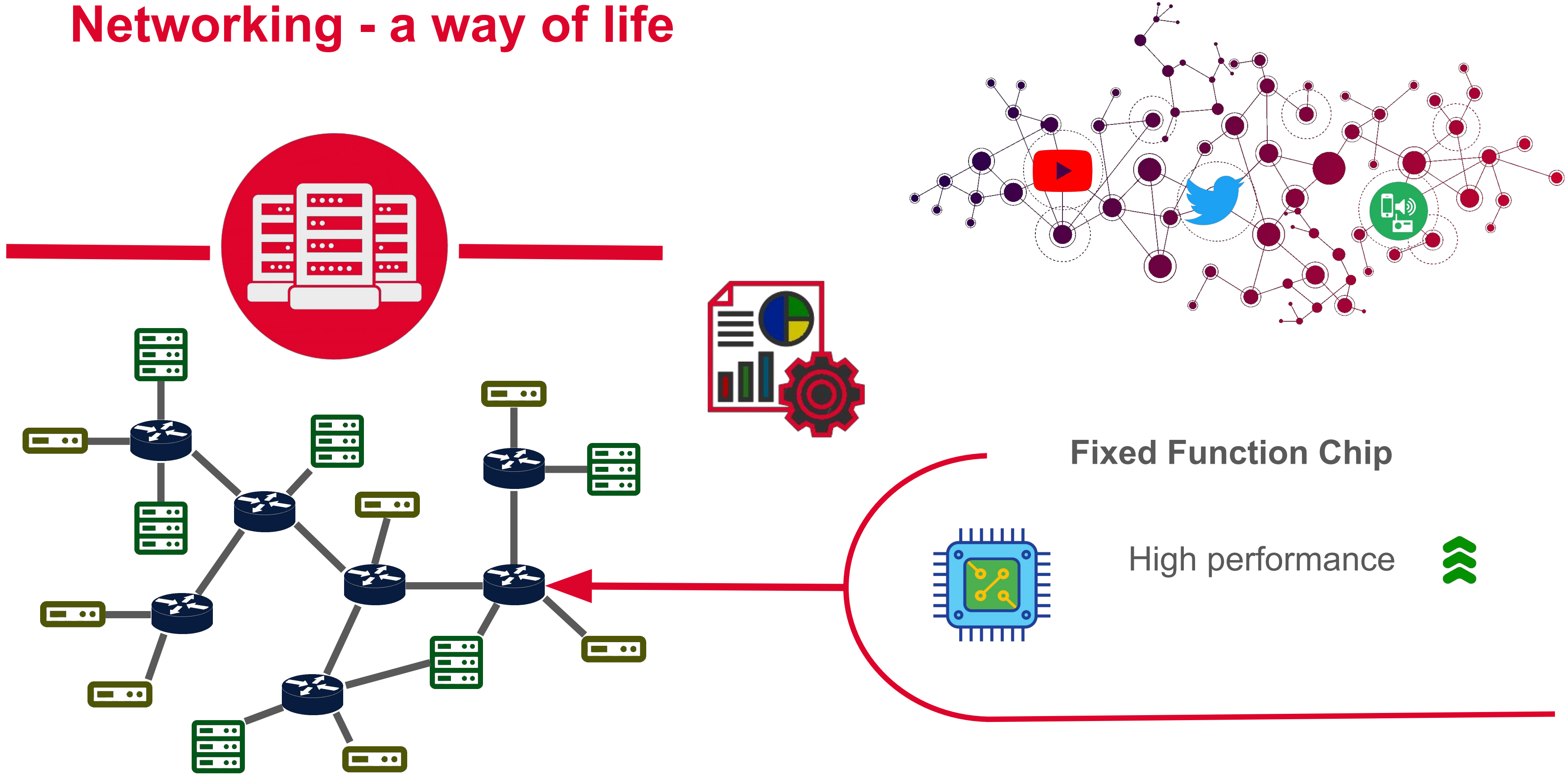






Fixed Function Chip

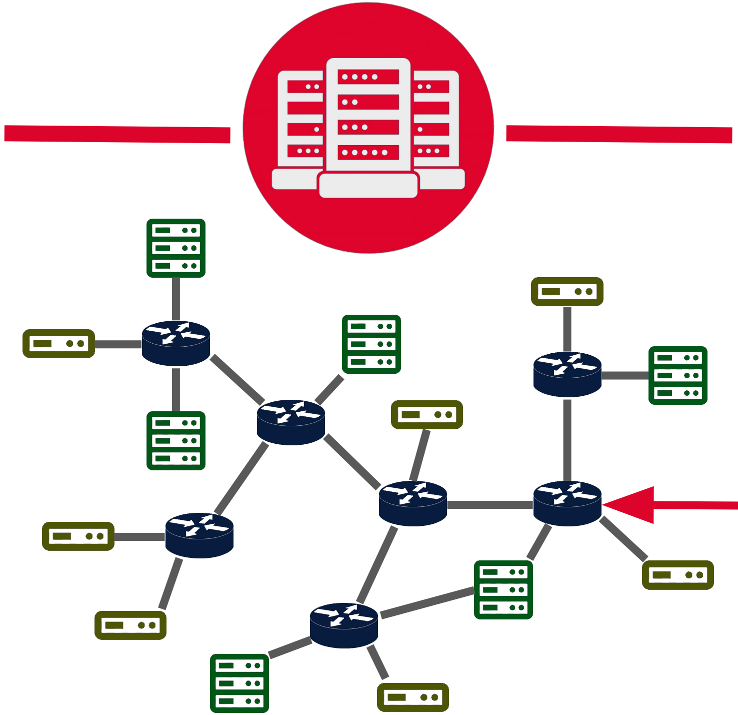
Networking - a way of life

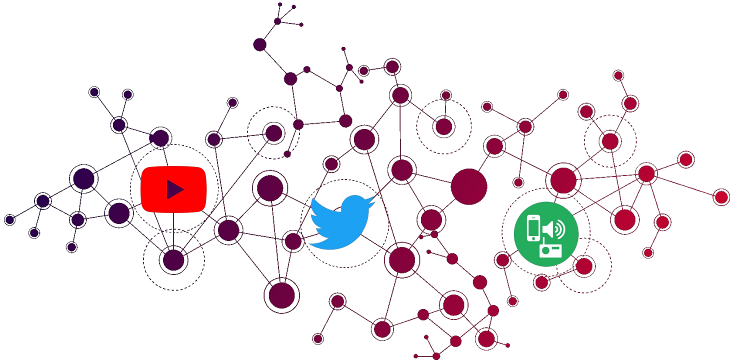


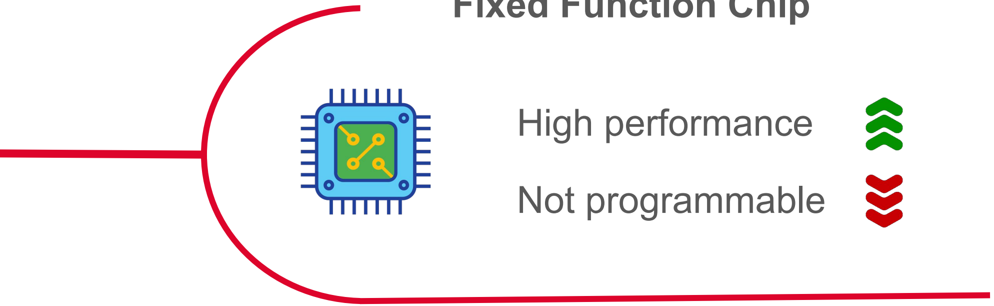
Fixed Function Chip

High performance

Networking - a way of life







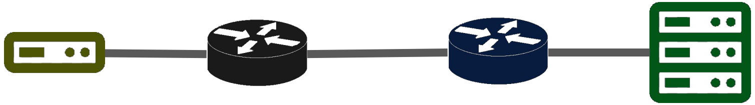
Fixed Function Chip

High performance

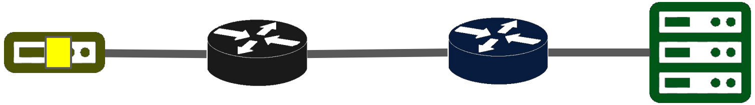
Not programmable

Programmable Networks !

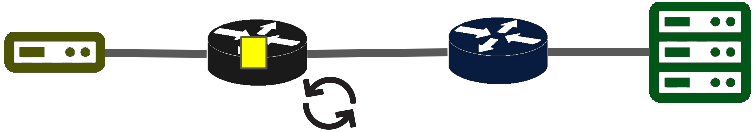
Programmable Networks !



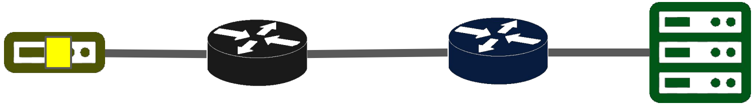
Programmable Networks !



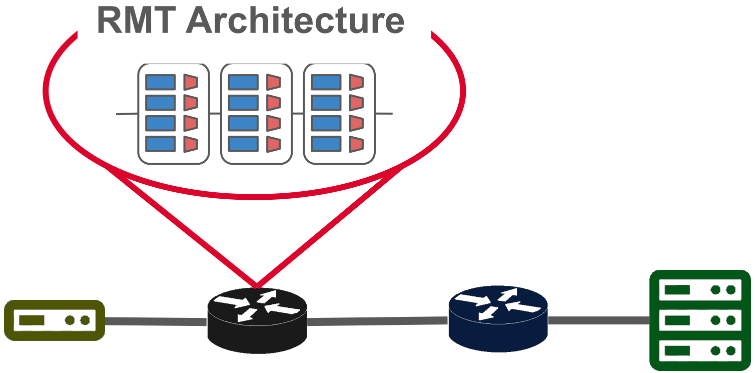
Programmable Networks !



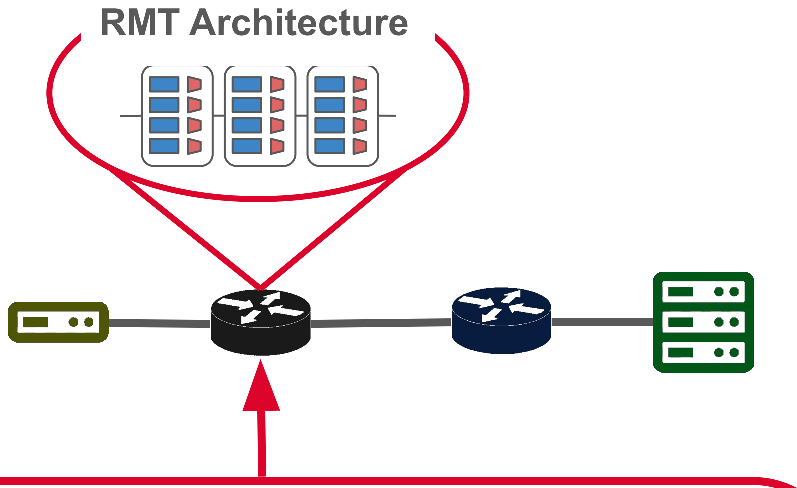
Programmable Networks !



Programmable Networks !

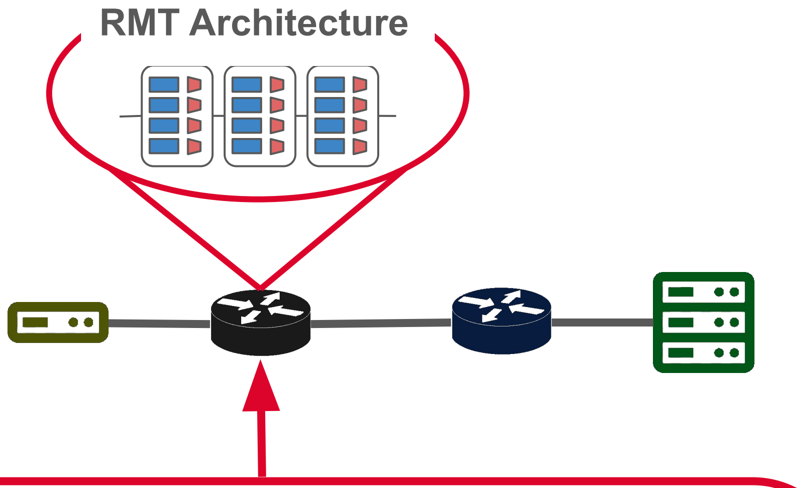


Programmable Networks !



High performance

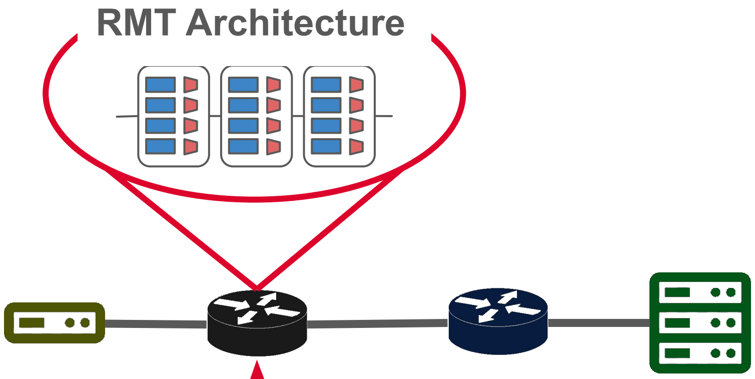
Programmable Networks !



High performance

Programmability

Programmable Networks !

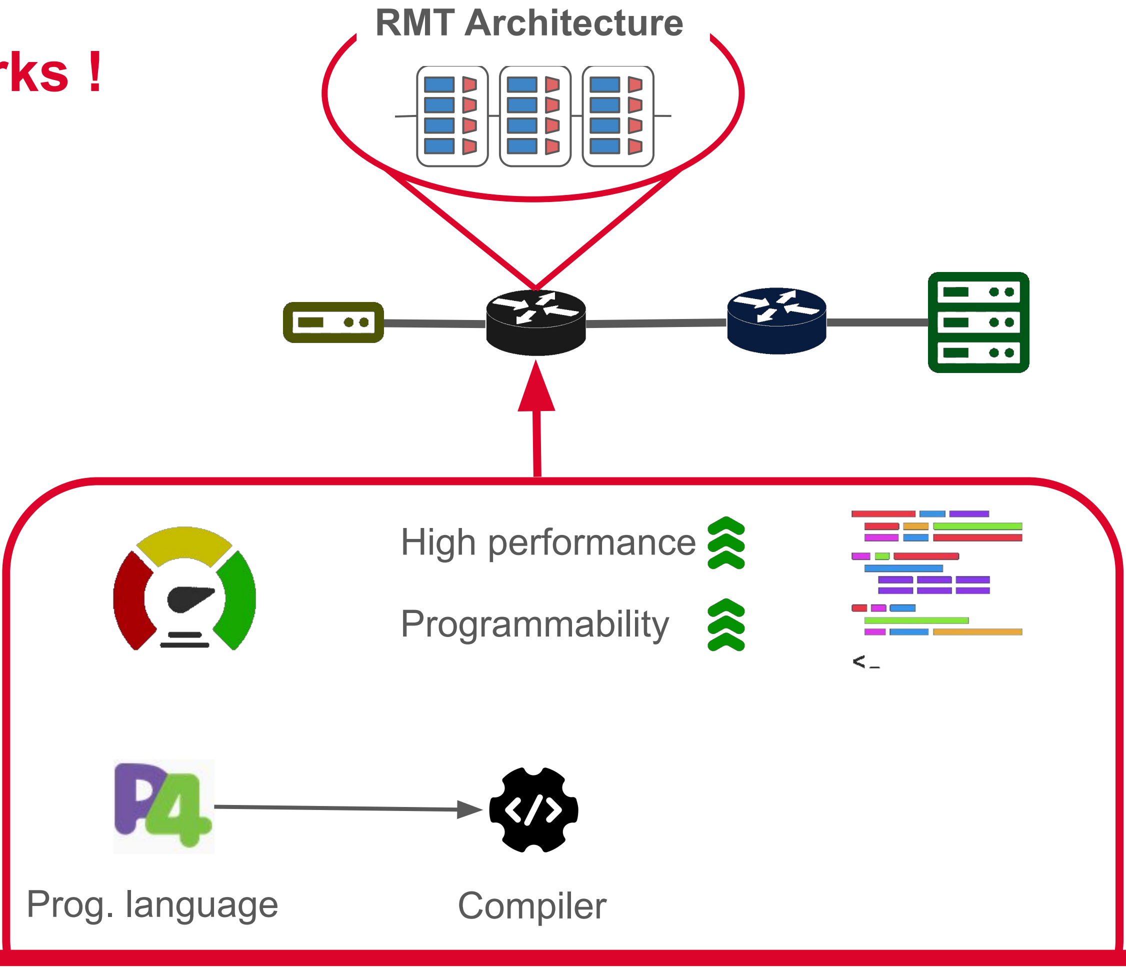


High performance ⏫

Programmability ⏫

Prog. language

Programmable Networks !



RMT Architecture

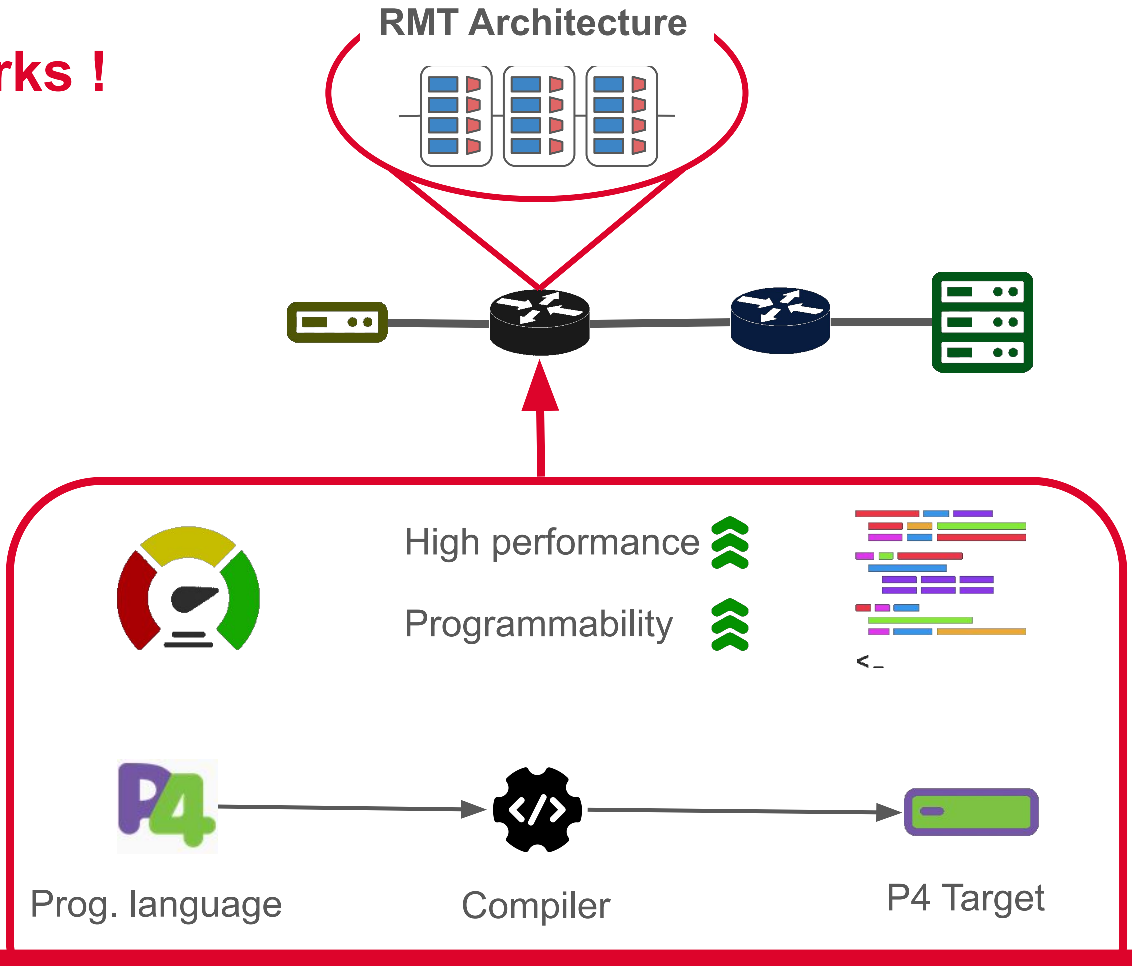
High performance

Programmability

Prog. language

Compiler

Programmable Networks !



RMT Architecture

High performance

Programmability

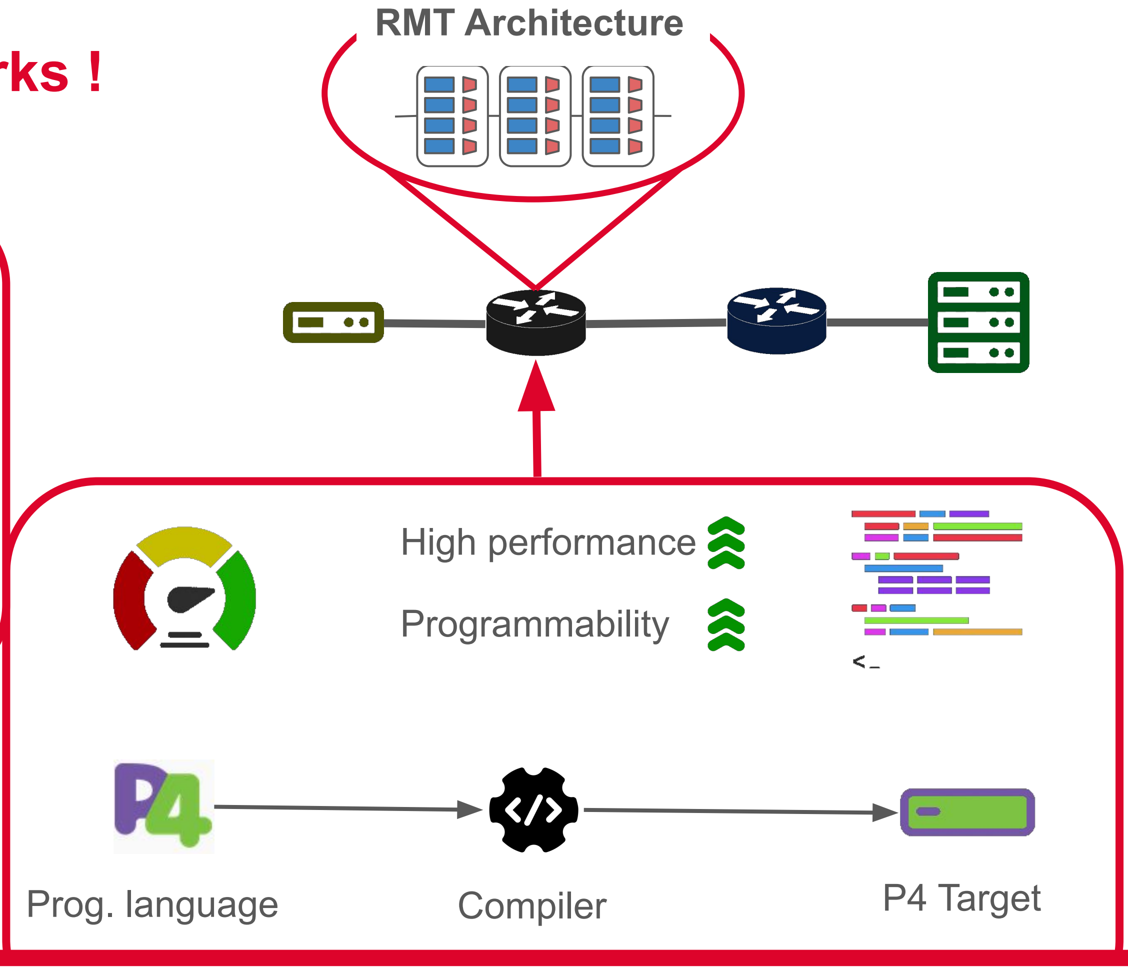
Prog. language

Compiler

P4 Target

Programmable Networks !

Software Defined Networking



RMT Architecture

High performance

Programmability

Prog. language

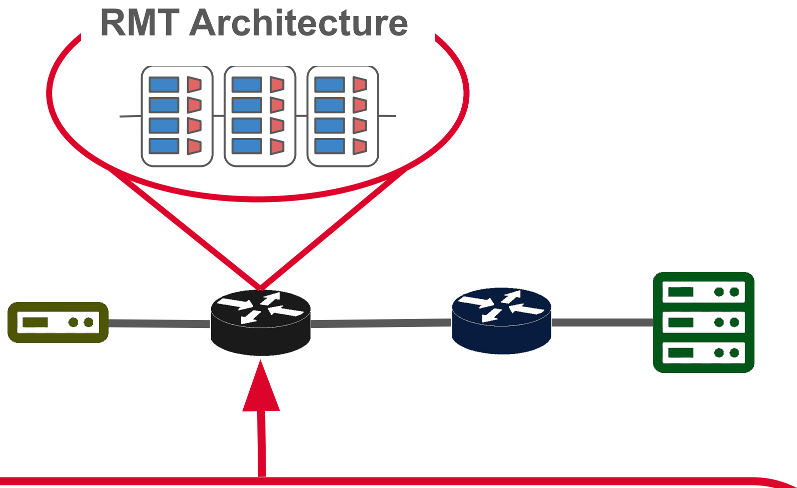
Compiler

P4 Target

Programmable Networks !

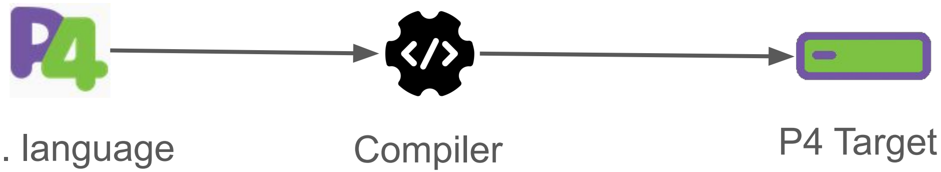
Software Defined Networking

Data Plane | Control Plane

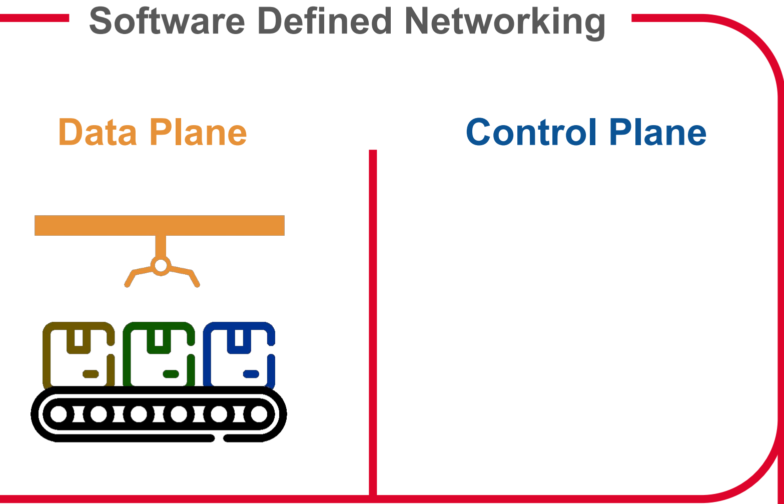


High performance

Programmability



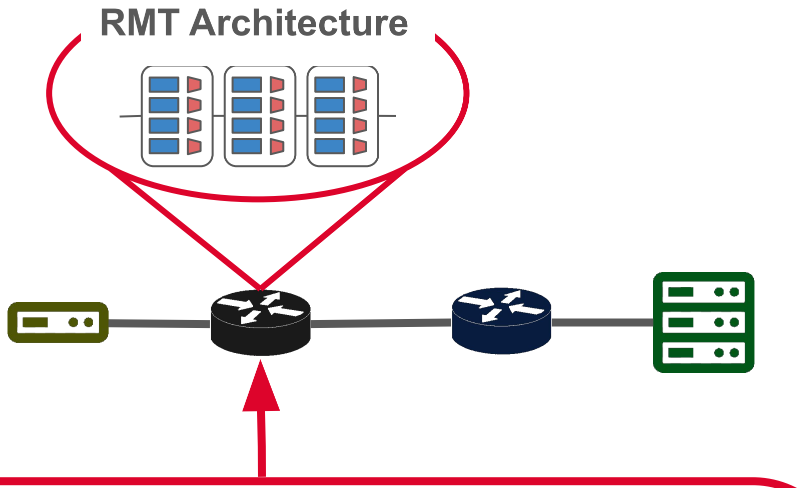
Programmable Networks !



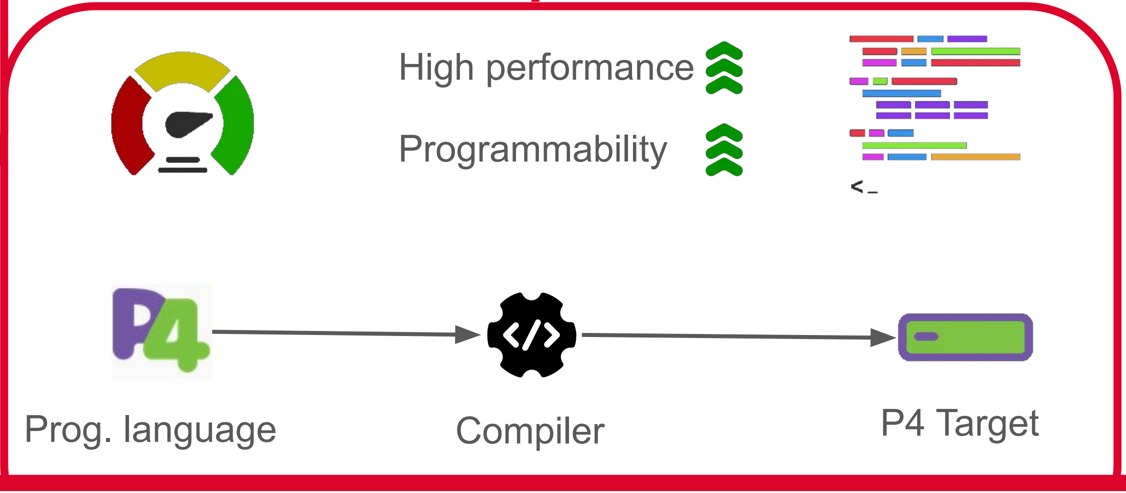
Software Defined Networking

Data Plane

Control Plane



RMT Architecture



High performance

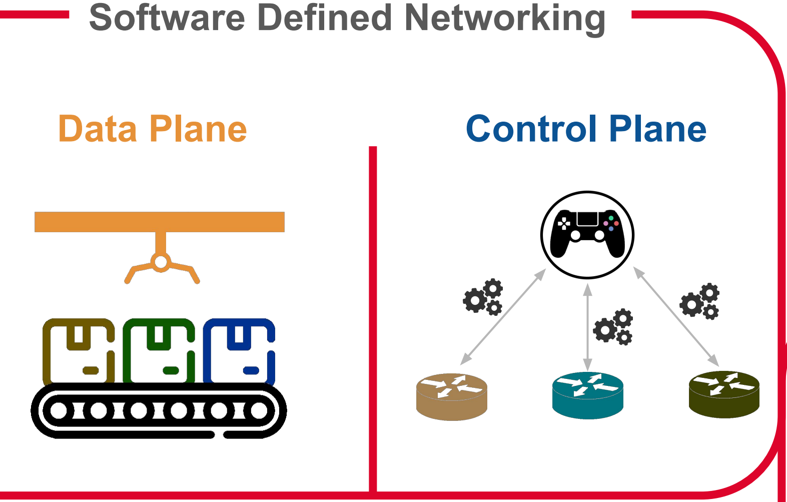
Programmability

Prog. language

Compiler

P4 Target

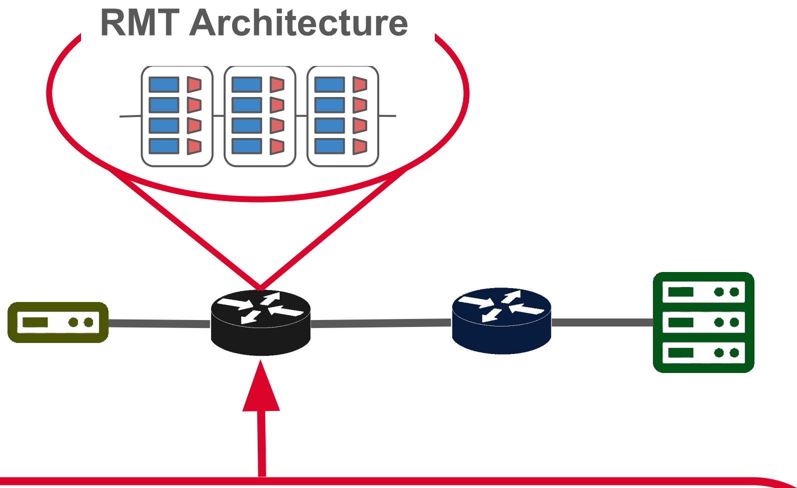
Programmable Networks !



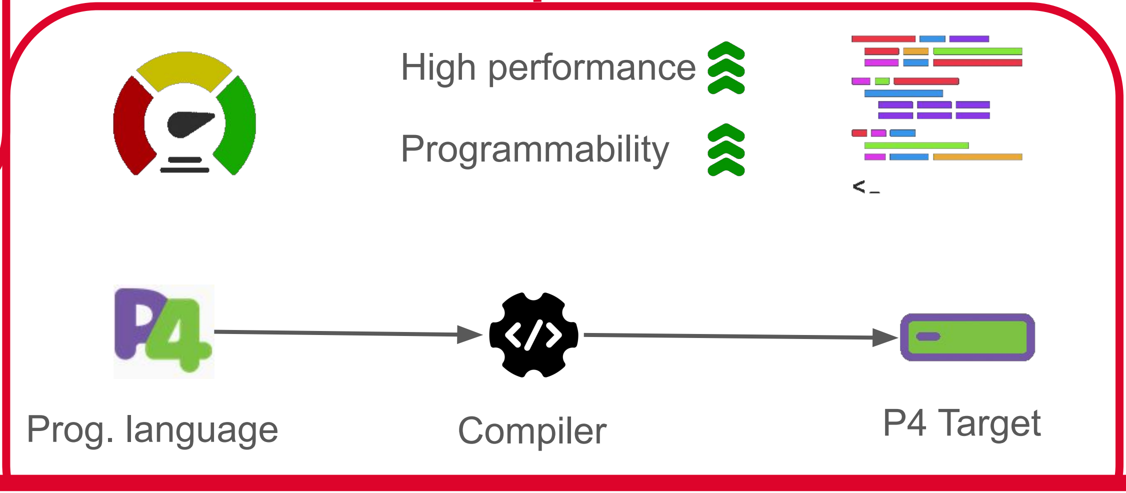
Software Defined Networking

Data Plane

Control Plane



RMT Architecture



High performance

Programmability

Prog. language

Compiler

P4 Target

P4 Constraints

P4 Constraints

P4 ≠ C | Java

P4 Constraints

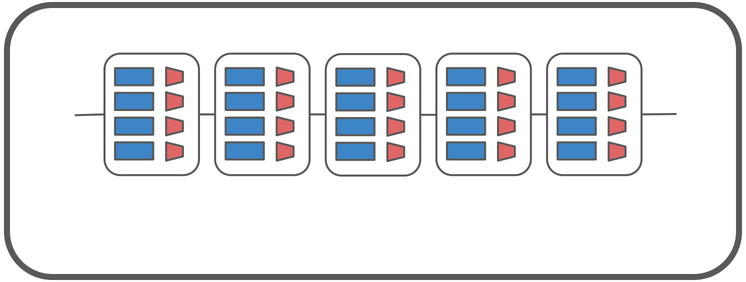
P4 ≠ C | Java

- ☒ Floating Point Numbers
- ☒ Loops

P4 Constraints

P4 ≠ C | Java

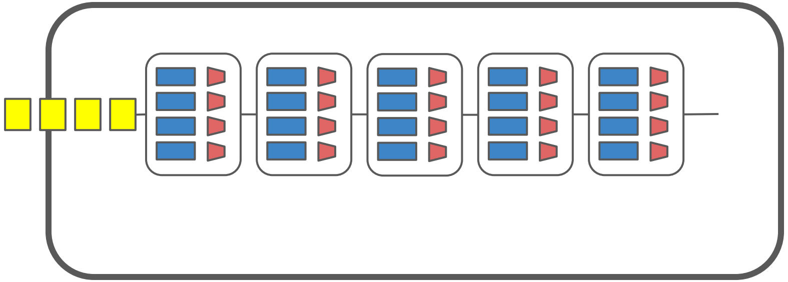
- ☒ Floating Point Numbers
- ☒ Loops



P4 Constraints

P4 ≠ C | Java

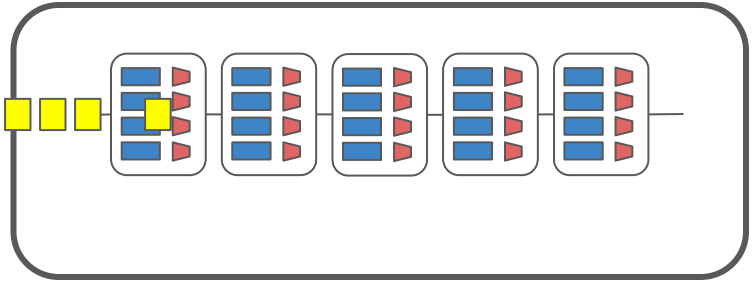
- ❌ Floating Point Numbers
- ❌ Loops



P4 Constraints

P4 ≠ C | Java

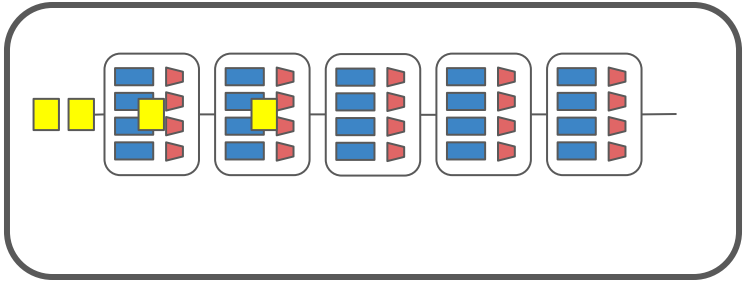
- ☒ Floating Point Numbers
- ☒ Loops



P4 Constraints

P4 ≠ C | Java

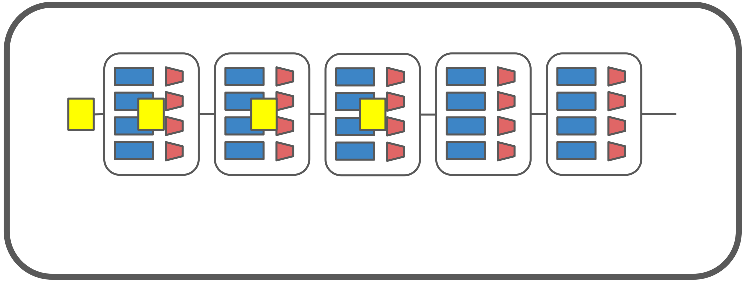
- ✗ Floating Point Numbers
- ✗ Loops



P4 Constraints

P4 ≠ C | Java

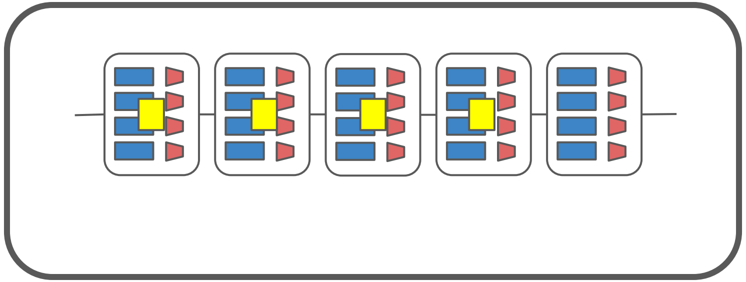
- ☒ Floating Point Numbers
- ☒ Loops



P4 Constraints

P4 ≠ C | Java

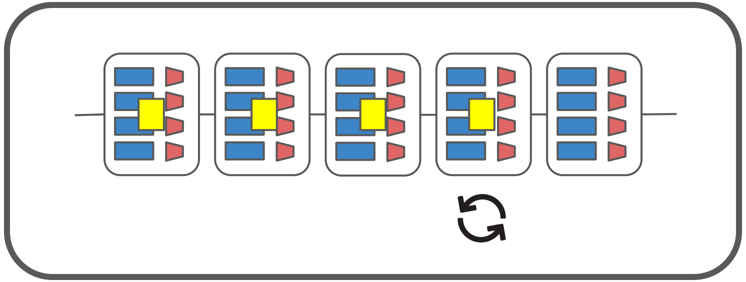
- ☒ Floating Point Numbers
- ☒ Loops



P4 Constraints

P4 ≠ C | Java

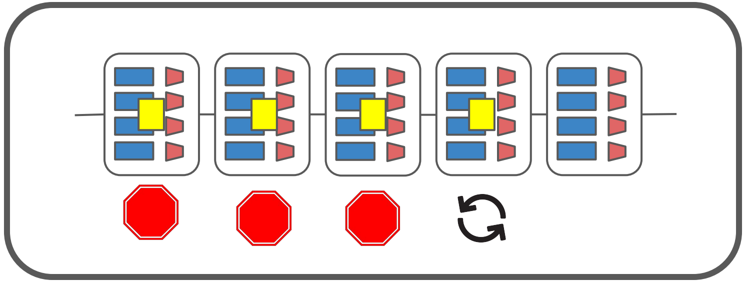
- ☒ Floating Point Numbers
- ☒ Loops



P4 Constraints

P4 ≠ C | Java

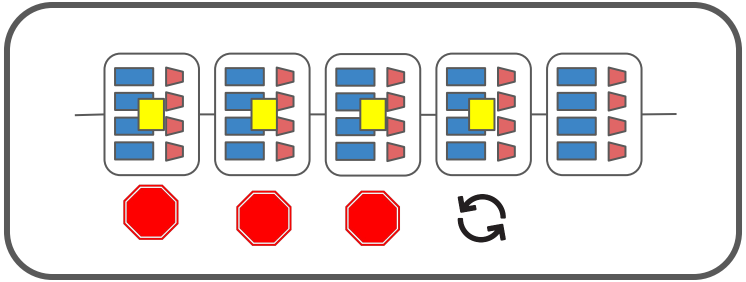
- ☒ Floating Point Numbers
- ☒ Loops



P4 Constraints

P4 ≠ C | Java

- ☒ Floating Point Numbers
- ☒ Loops



Limitations :

- **Constant time complexity is required**
- **Limited amount of memory**

Research Questions

Research Questions

What are the acceleration capabilities stateful data analytics inside the network ?

Research Questions

What are the acceleration capabilities stateful data analytics inside the network ?

What are the limitation of the in-network computing model ?

Research Questions

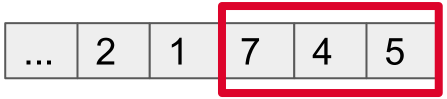
What are the acceleration capabilities stateful data analytics inside the network ?

What are the limitation of the in-network computing model ?

Is it viable to offload stateful end-host operators onto network devices ?

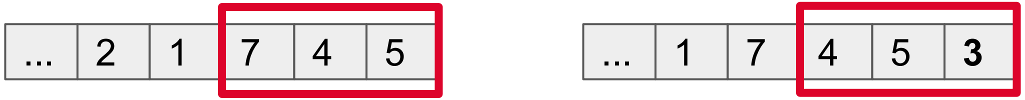
Sliding Window Average

Sliding Window Average



...	2	1	7	4	5

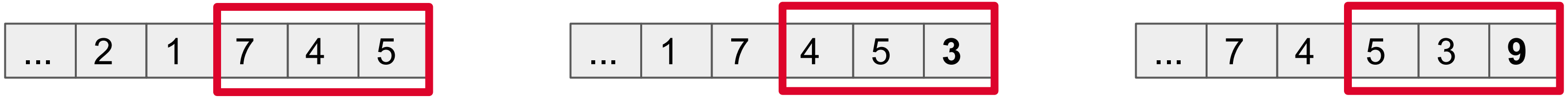
Sliding Window Average



...	2	1	7	4	5

...	1	7	4	5	**3**

Sliding Window Average

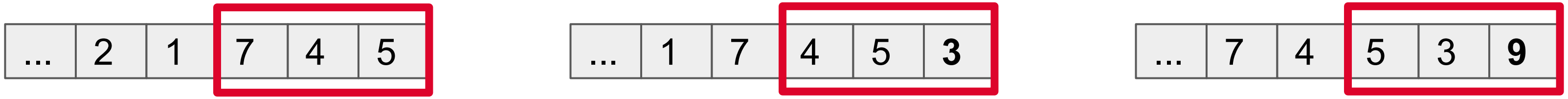


...	2	1	7	4	5

...	1	7	4	5	**3**

...	7	4	5	3	**9**

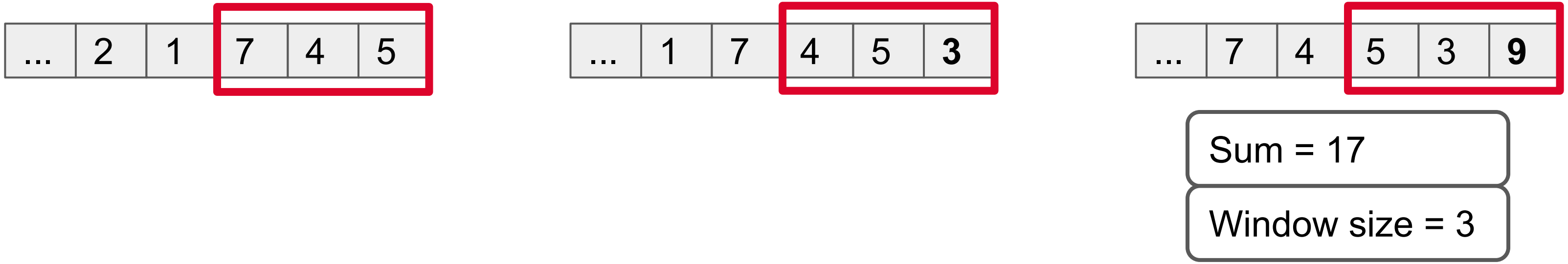
Sliding Window Average



P4-target Constraints

- ☒ **Division at runtime**
- ☒ **Floating point numbers**

Sliding Window Average



...	2	1	7	4	5

...	1	7	4	5	**3**

...	7	4	5	3	**9**

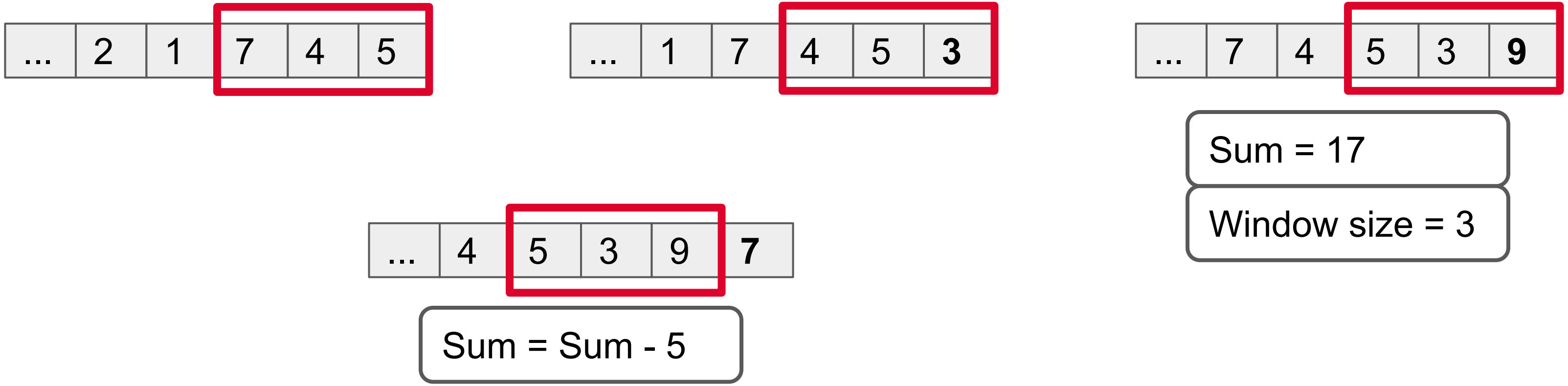
Sum = 17

Window size = 3

P4-target Constraints

- ☒ **Division at runtime**
- ☒ **Floating point numbers**

Sliding Window Average



Sum = 17

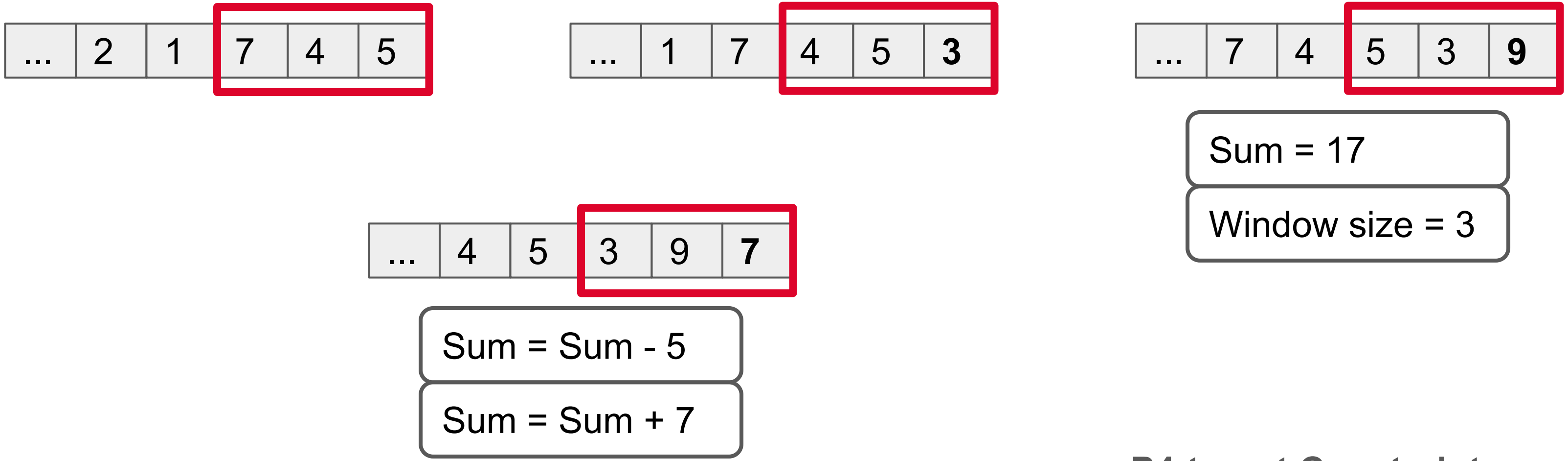
Window size = 3

Sum = Sum - 5

P4-target Constraints

- ☒ **Division at runtime**
- ☒ **Floating point numbers**

Sliding Window Average



... | 2 | 1 | 7 | 4 | 5

... | 1 | 7 | 4 | 5 | **3**

... | 7 | 4 | 5 | 3 | **9**

Sum = 17

Window size = 3

... | 4 | 5 | 3 | 9 | **7**

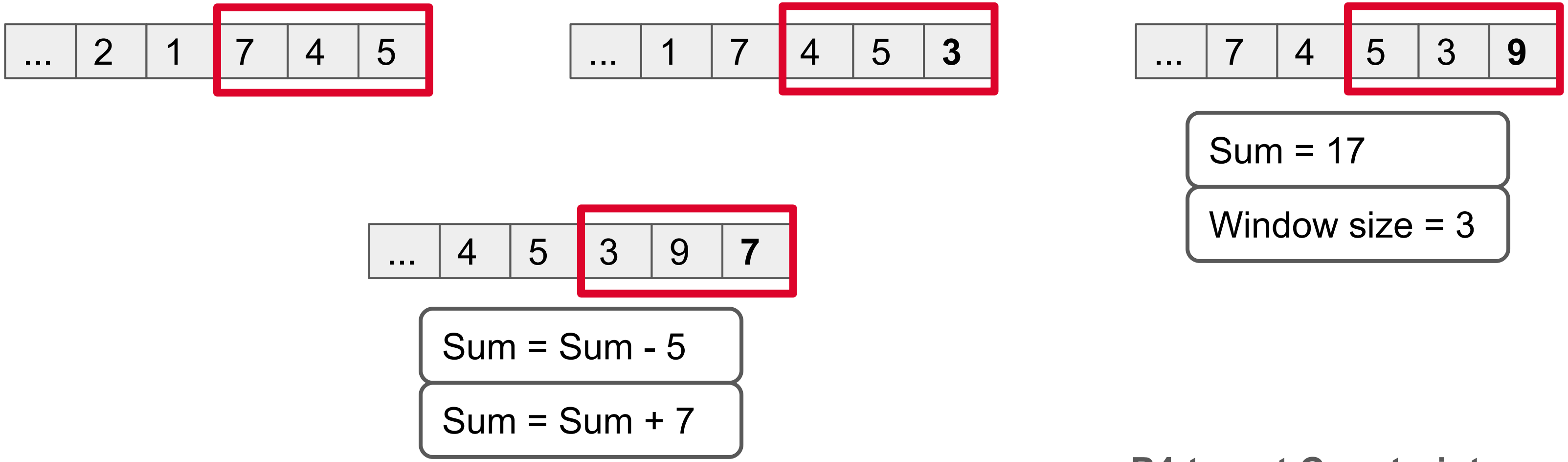
Sum = Sum - 5

Sum = Sum + 7

P4-target Constraints

- ☒ **Division at runtime**
- ☒ **Floating point numbers**

Sliding Window Average



...	2	1	7	4	5

...	1	7	4	5	**3**

...	7	4	5	3	**9**

Sum = 17

Window size = 3

...	4	5	3	9	**7**

Sum = Sum - 5

Sum = Sum + 7

Time Complexity : O(1)

Memory Complexity : O(n), n - the size of the window

P4-target Constraints

- ☒ **Division at runtime**
- ☒ **Floating point numbers**

Median

Median

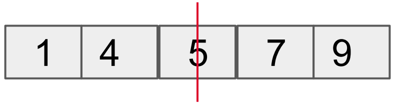
Constraints:

- Only integers are accepted
- The set of possible values must be known in advance

Median

Constraints:

- Only integers are accepted
- The set of possible values must be known in advance

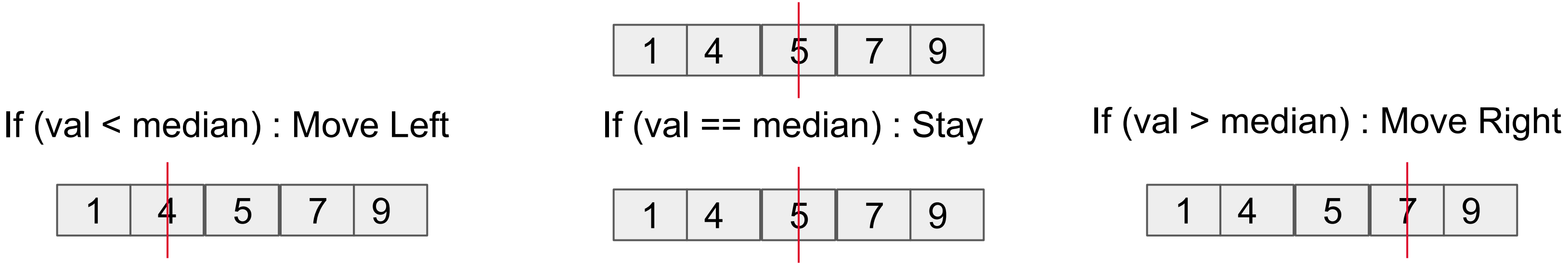


1	4	5	7	9

Median

Constraints:

- Only integers are accepted
- The set of possible values must be known in advance



If (val < median) : Move Left

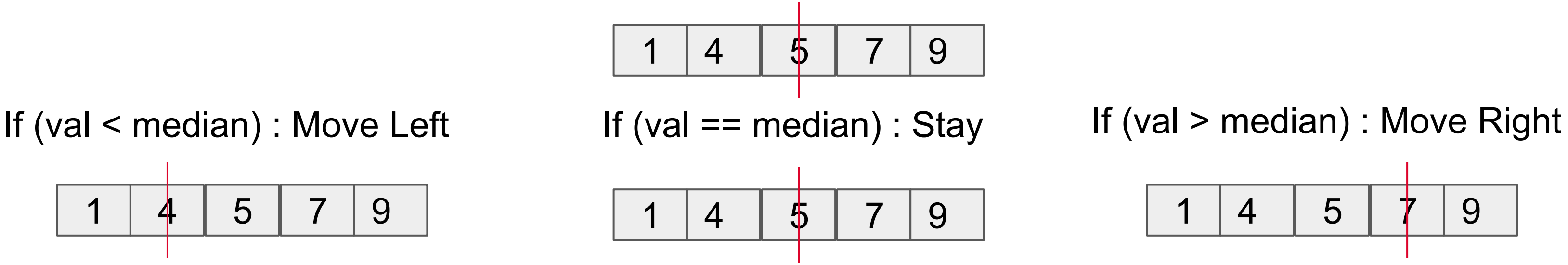
If (val == median) : Stay

If (val > median) : Move Right

Median

Constraints:

- Only integers are accepted
- The set of possible values must be known in advance



1 | 4 | 5 | 7 | 9

If (val < median) : Move Left

1 | 4 | 5 | 7 | 9

If (val == median) : Stay

1 | 4 | 5 | 7 | 9

If (val > median) : Move Right

1 | 4 | 5 | 7 | 9

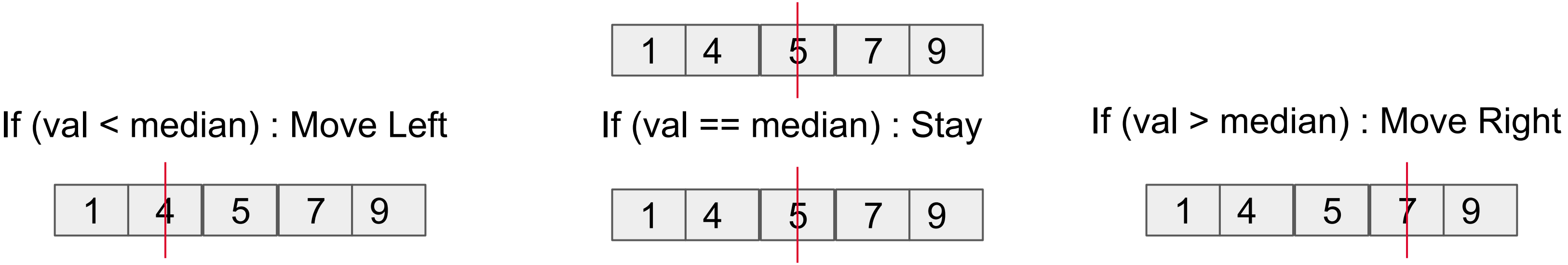
P4-target Constraints

- ✗ Nontrivial operations
- ✗ Bit shift of more than 8 bits

Median

Constraints:

- **Only integers are accepted**
- **The set of possible values must be known in advance**



1	4	5	7	9

If (val < median) : Move Left

1	4	5	7	9

If (val == median) : Stay

1	4	5	7	9

If (val > median) : Move Right

1	4	5	7	9

Time Complexity : O(1)

Memory Complexity : O(n), n - the number of possible values

P4-target Constraints

- ☒ **Nontrivial operations**
- ☒ **Bit shift of more than 8 bits**

Performance Comparison

Performance Comparison

In-Network Computing P4

End-Host Computing DPDK DATA PLANE DEVELOPMENT KIT

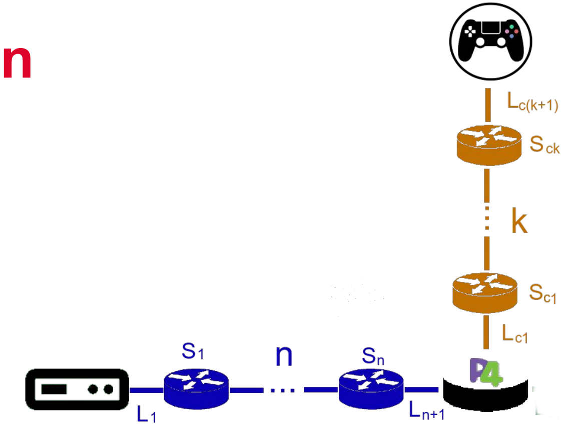
Performance Comparison

In-Network Computing

End-Host Computing

Create / Remove a flow (Latency)

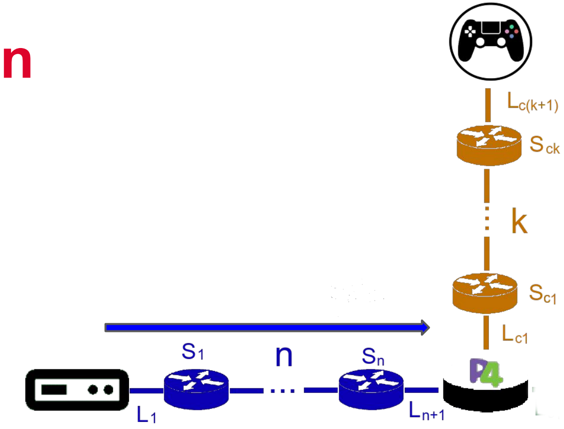
P4



Performance Comparison

In-Network Computing

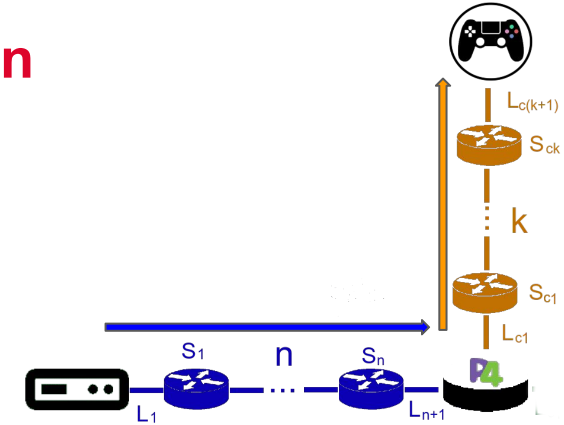
End-Host Computing



Create / Remove a flow (Latency)

P4 $\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right)$

Performance Comparison



In-Network Computing P4

End-Host Computing DPDK DATA PLANE DEVELOPMENT KIT

Create / Remove a flow (Latency)

P4

$$\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + \left(\sum_{i=1}^{k+1} L_{ci} + \sum_{i=1}^{k} S_{ci}\right)$$

Performance Comparison

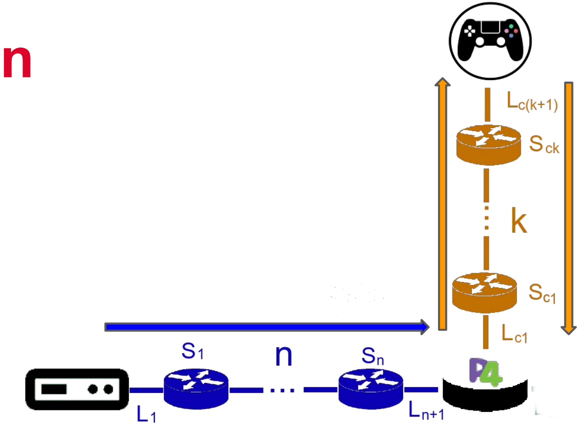
In-Network Computing

End-Host Computing

Create / Remove a flow (Latency)

P4

$$\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 2\left(\sum_{i=1}^{k+1} L_{ci} + \sum_{i=1}^{k} S_{ci}\right) + T_{controller} + T_{P4OS}$$



Performance Comparison

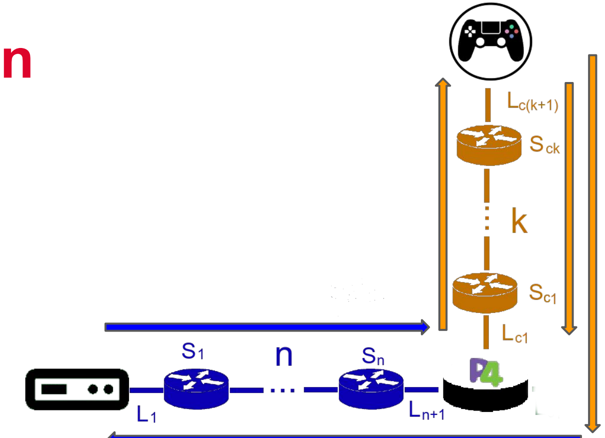
In-Network Computing P4

End-Host Computing DPDK DATA PLANE DEVELOPMENT KIT

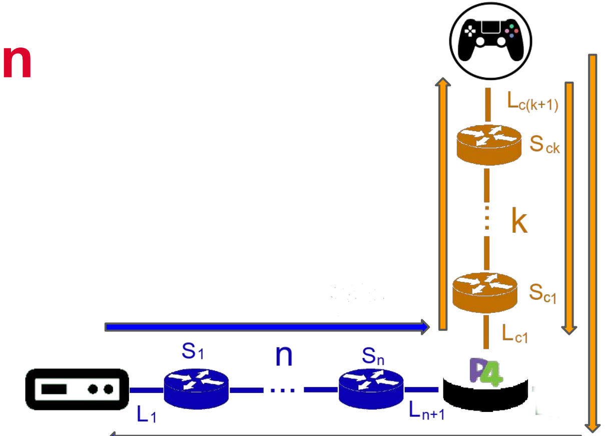
Create / Remove a flow (Latency)

P4

$$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 3\left(\sum_{i = 1}^{k + 1} L_{ci} + \sum_{i=1}^{k} S_{ci}\right) + T_{controller} + T_{P4OS}$$



Performance Comparison



In-Network Computing

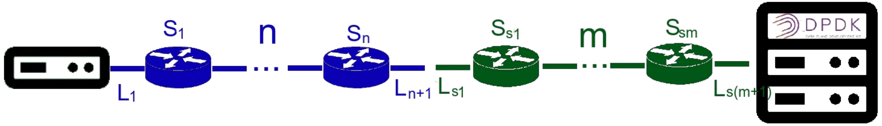
End-Host Computing

Create / Remove a flow (Latency)

P4

$$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 3\left(\sum_{i=1}^{k+1} L_{ci} + \sum_{i=1}^{k} S_{ci}\right) + T_{controller} + T_{P4OS}$$

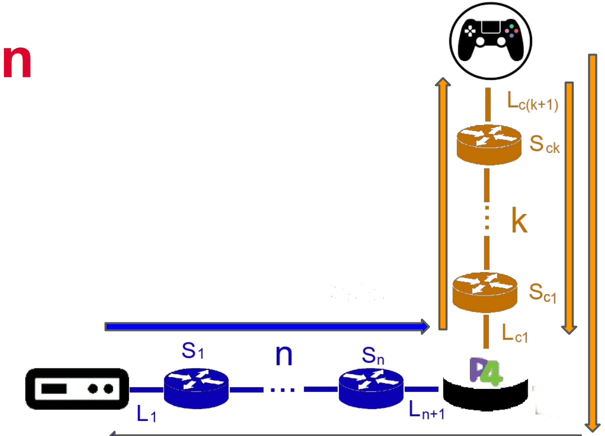
DPDK



Performance Comparison

In-Network Computing P4

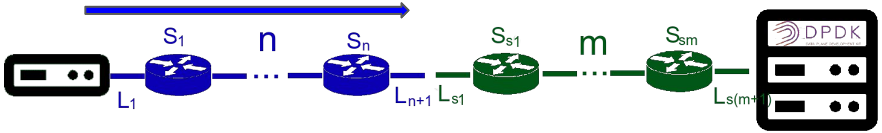
End-Host Computing DPDK DATA PLANE DEVELOPMENT KIT



Create / Remove a flow (Latency)

P4
$$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 3\left(\sum_{i=1}^{k+1} L_{ci} + \sum_{i=1}^{k} S_{ci}\right) + T_{controller} + T_{P4OS}$$

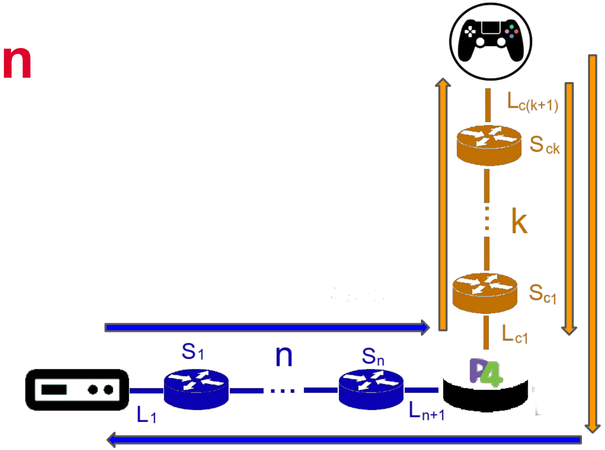
DPDK
$$\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right)$$



Performance Comparison

In-Network Computing

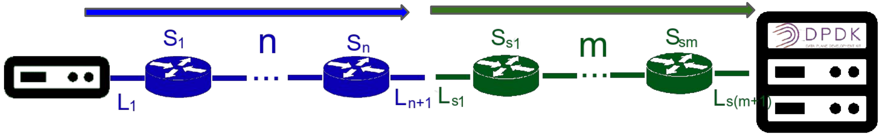
End-Host Computing



Create / Remove a flow (Latency)

P4 $$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 3\left(\sum_{i=1}^{k+1} L_{ci} + \sum_{i=1}^{k} S_{ci}\right) + T_{controller} + T_{P4OS}$$

DPDK $$\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + \left(\sum_{i=1}^{m+1} L_{Si} + \sum_{i=1}^{m} S_{Si}\right) + T_{DPDK}$$



Performance Comparison

In-Network Computing

End-Host Computing

DPDK
DATA PLANE DEVELOPMENT KIT

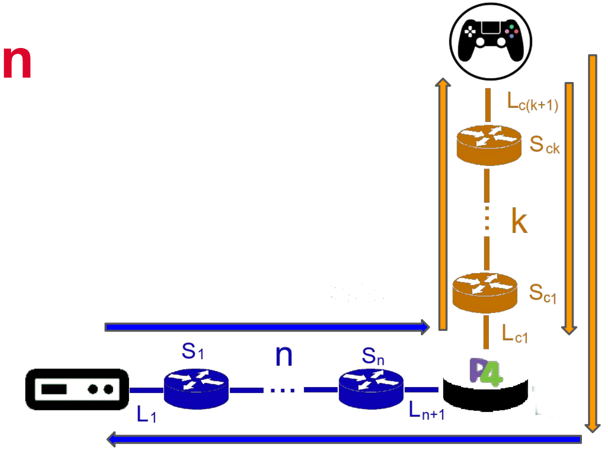
Create / Remove a flow (Latency)

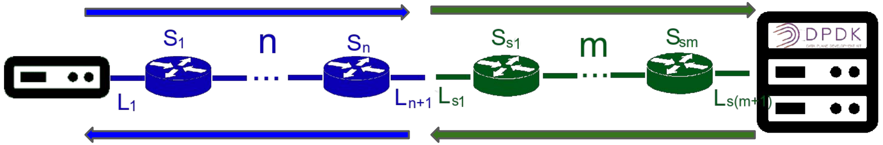
P4

$$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 3\left(\sum_{i=1}^{k+1} L_{ci} + \sum_{i=1}^{k} S_{ci}\right) + T_{controller} + T_{P4OS}$$

DPDK

$$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 2\left(\sum_{i=1}^{m+1} L_{Si} + \sum_{i=1}^{m} S_{Si}\right) + T_{DPDK}$$





Performance Comparison

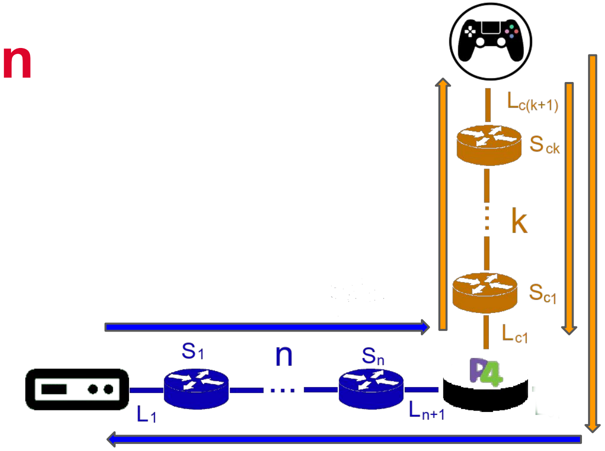
In-Network Computing

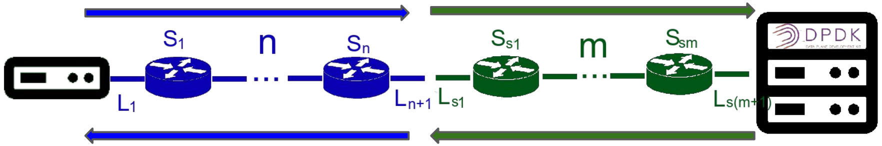
End-Host Computing

Create / Remove a flow (Latency)

⏬ P4 $$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 3\left(\sum_{i=1}^{k+1} L_{ci} + \sum_{i=1}^{k} S_{ci}\right) + T_{controller} + T_{P4OS}$$

⏫ DPDK $$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 2\left(\sum_{i=1}^{m+1} L_{Si} + \sum_{i=1}^{m} S_{Si}\right) + T_{DPDK}$$

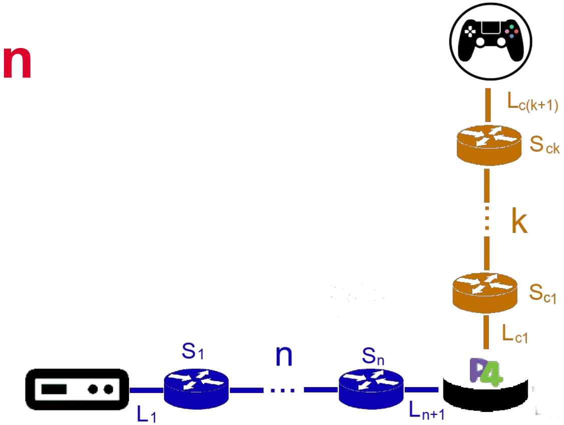




Performance Comparison

In-Network Computing

End-Host Computing



Create / Remove a flow (Latency)

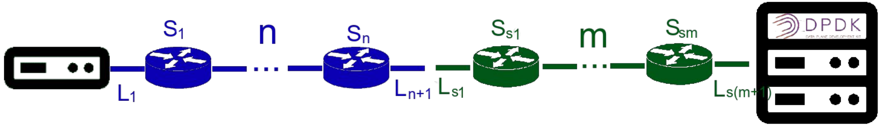
P4

$$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 3\left(\sum_{i=1}^{k+1} L_{ci} + \sum_{i=1}^{k} S_{ci}\right) + T_{controller} + T_{P4OS}$$

DPDK

$$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 2\left(\sum_{i=1}^{m+1} L_{Si} + \sum_{i=1}^{m} S_{Si}\right) + T_{DPDK}$$

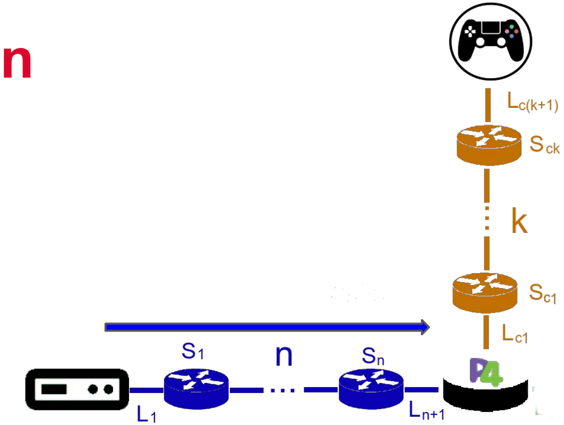
Process a packet (Latency)



Performance Comparison

In-Network Computing

End-Host Computing



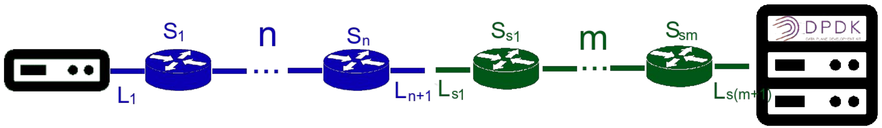
Create / Remove a flow (Latency)

⏬ P4

$$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 3\left(\sum_{i=1}^{k+1} L_{ci} + \sum_{i=1}^{k} S_{ci}\right) + T_{controller} + T_{P4OS}$$

⏫ DPDK

$$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 2\left(\sum_{i=1}^{m+1} L_{Si} + \sum_{i=1}^{m} S_{Si}\right) + T_{DPDK}$$



Process a packet (Latency)

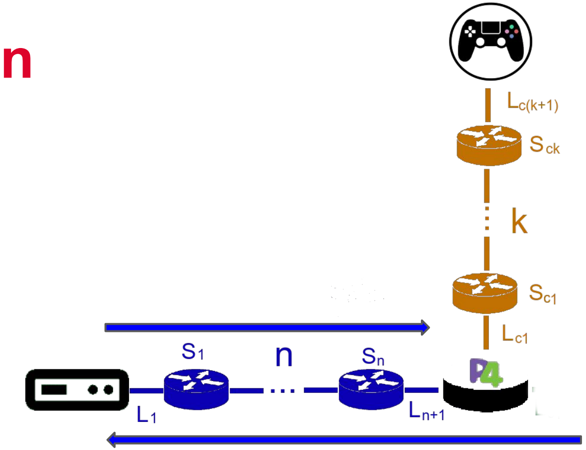
P4

$$\left(\sum_{i=1}^{n+1} L_i + \sum_{i=i}^{n} S_i\right) + T_{P4}$$

Performance Comparison

In-Network Computing P4

End-Host Computing DPDK DATA PLANE DEVELOPMENT KIT



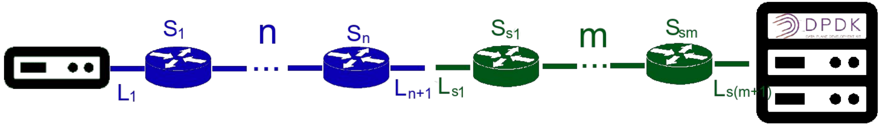
Create / Remove a flow (Latency)

⏬ P4 $$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 3\left(\sum_{i=1}^{k+1} L_{ci} + \sum_{i=1}^{k} S_{ci}\right) + T_{controller} + T_{P4OS}$$

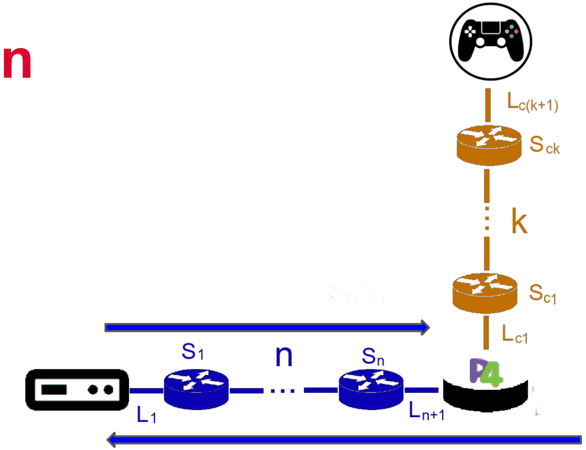
⏫ DPDK $$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 2\left(\sum_{i=1}^{m+1} L_{Si} + \sum_{i=1}^{m} S_{Si}\right) + T_{DPDK}$$

Process a packet (Latency)

P4 $$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=i}^{n} S_i\right) + T_{P4}$$



Performance Comparison



In-Network Computing

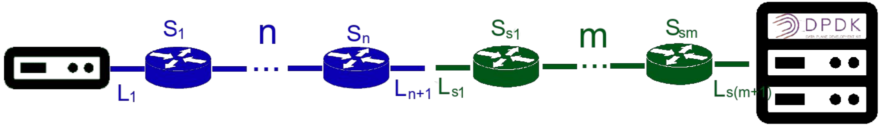
End-Host Computing

Create / Remove a flow (Latency)

P4 $$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 3\left(\sum_{i=1}^{k+1} L_{ci} + \sum_{i=1}^{k} S_{ci}\right) + T_{controller} + T_{P4OS}$$

DPDK $$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 2\left(\sum_{i=1}^{m+1} L_{Si} + \sum_{i=1}^{m} S_{Si}\right) + T_{DPDK}$$

Process a packet (Latency)



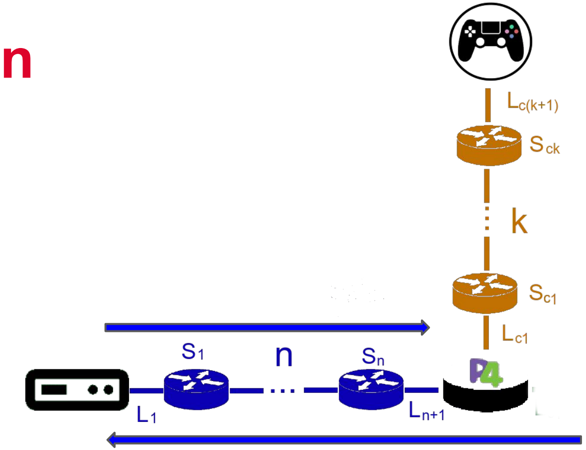
P4 $$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=i}^{n} S_i\right) + T_{P4}$$

DPDK

Performance Comparison

In-Network Computing

End-Host Computing



Create / Remove a flow (Latency)

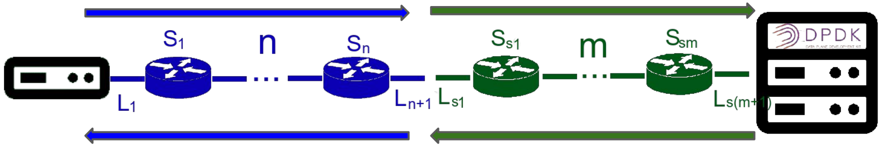
P4 $$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 3\left(\sum_{i=1}^{k+1} L_{ci} + \sum_{i=1}^{k} S_{ci}\right) + T_{controller} + T_{P4OS}$$

DPDK $$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 2\left(\sum_{i=1}^{m+1} L_{Si} + \sum_{i=1}^{m} S_{Si}\right) + T_{DPDK}$$

Process a packet (Latency)

P4 $$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=i}^{n} S_i\right) + T_{P4}$$

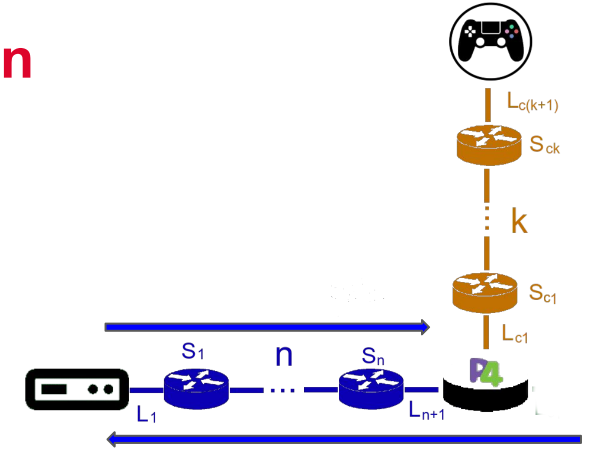
DPDK $$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 2\left(\sum_{i=1}^{m+1} L_{Si} + \sum_{i=1}^{m} S_{Si}\right) + T_{DPDK}$$



Performance Comparison

In-Network Computing

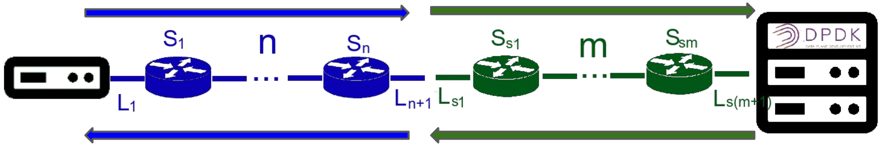
End-Host Computing



Create / Remove a flow (Latency)

⏬ P4 $$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 3\left(\sum_{i=1}^{k+1} L_{ci} + \sum_{i=1}^{k} S_{ci}\right) + T_{controller} + T_{P4OS}$$

⏫ DPDK $$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 2\left(\sum_{i=1}^{m+1} L_{Si} + \sum_{i=1}^{m} S_{Si}\right) + T_{DPDK}$$



Process a packet (Latency)

⏫ P4 $$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=i}^{n} S_i\right) + T_{P4}$$

⏬ DPDK $$2\left(\sum_{i=1}^{n+1} L_i + \sum_{i=1}^{n} S_i\right) + 2\left(\sum_{i=1}^{m+1} L_{Si} + \sum_{i=1}^{m} S_{Si}\right) + T_{DPDK}$$

Conclusion

Conclusion

Elephant Flows : in-network computing

Conclusion

Elephant Flows : in-network computing

Mice Flows : end-host computing

Conclusion

Elephant Flows : in-network computing

Mice Flows : end-host computing

	P4 <=>	DPDK
Performance	✓	⚠
Flexibility	⚠	✓

Conclusion

Elephant Flows : in-network computing

Mice Flows : end-host computing

	P4 <=>	DPDK
Performance	✓	⚠
Flexibility	⚠	✓

P4 Constraints :

- Lack of loops
- Lack of support for floats

P4-Target Constraints :

- Limited memory
- Lack of support for non-trivial operations, such as
    - Multiplication
    - Division
    - Arbitrary bit shifting

Conclusion

Elephant Flows : in-network computing

Mice Flows : end-host computing

	P4 <=>	DPDK
Performance	✓	⚠
Flexibility	⚠	✓

P4 Constraints :

- Lack of loops
- Lack of support for floats

P4-Target Constraints :

- Limited memory
- Lack of support for non-trivial operations, such as
  - Multiplication
  - Division
  - Arbitrary bit shifting

Viability depends on :

- The ability of the operator to conform to the P4 and P4-target constraints
- The topology of the network
- The size of the flows