4th Year Project

Distributed Caching in Datacenter Switches

Project Number: p-2022-093

Students: Anna Axalrod, Shir Granit.

Supervisors: Prof. Chen Avin, Dr. Gabriel Scalosub.

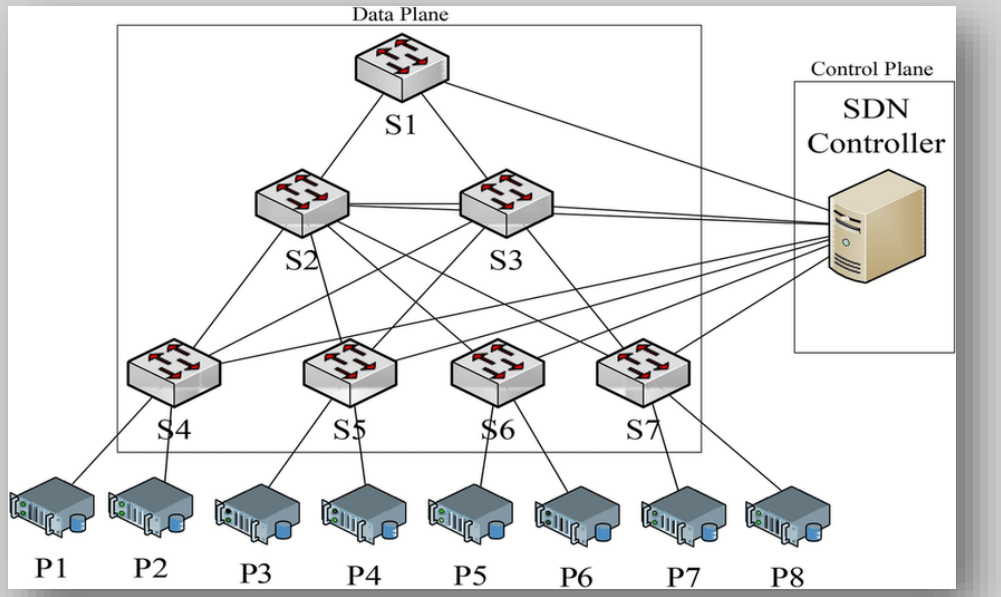
Introduction

- Datacenters rely on **high connectivity**.
- Switches are required to store an **enormous amount** of traffic rules.
- Storing all forwarding policies is **not possible** due to **memory limitations**.

An illustrative image of a data center
Server farm (cybrain; iStock by Getty Images)

The problem

- What happens to **a packet with no forwarding rule**?
- SDN: forward to a **controller**
  - The controller contains **all** forwarding rules.
- Causes a **degradation in the response time.**

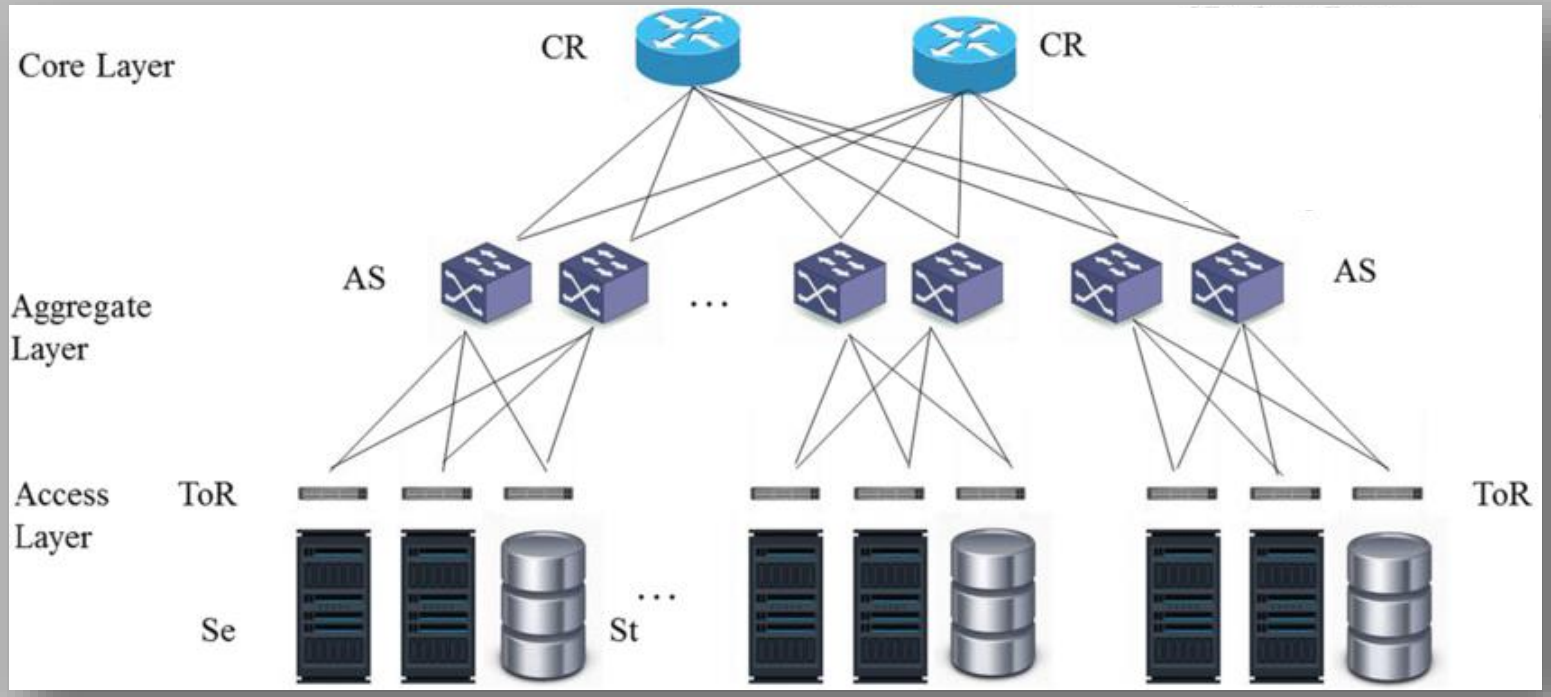


Software-defined load-balanced data center: design, implementation and performance analysis. (From: Montazerolghaem, Ahmadreza., 2021)

Project Goals

- Response time is an **essential Key Performance Indicator (KPI)** in datacenters.
- Find a solution that **avoids multiple accesses** to the controller.
- Method: **cache mechanism** for storing relevant **subset** of forwarding rules.

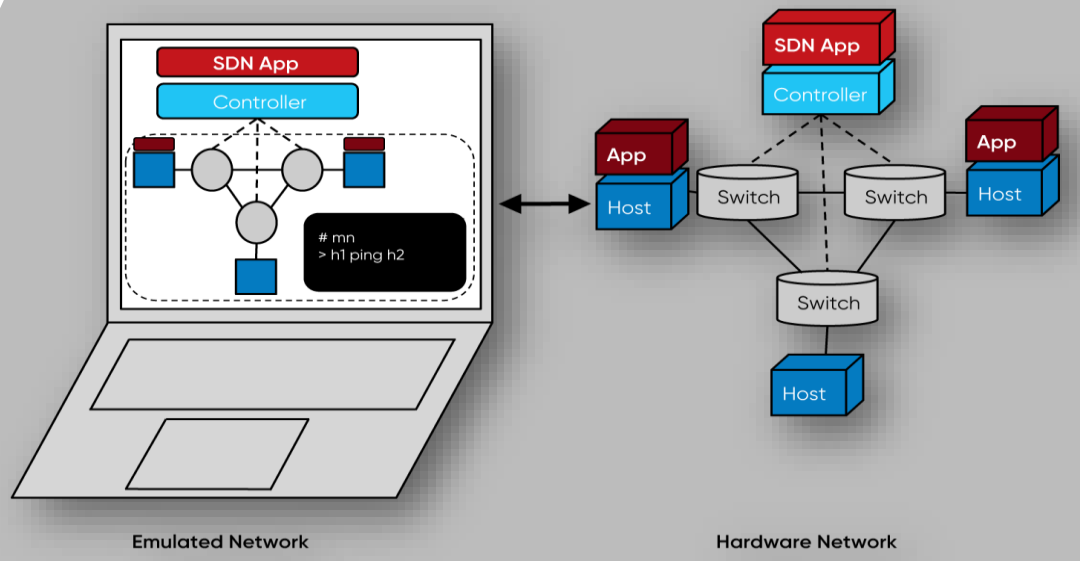
Datacenter 3-tier Network Architecture



Datacenter 3-tier network architecture
(From: Energy, network, and application-aware virtual machine placement model in SDN-enabled large scale cloud data centers, 2021)

- CR: Core Router
- AS: Aggregation Switch
- ToR: Top of Rack Switch
- Se: Server
- St: Storage Node

Mininet Emulation Environment

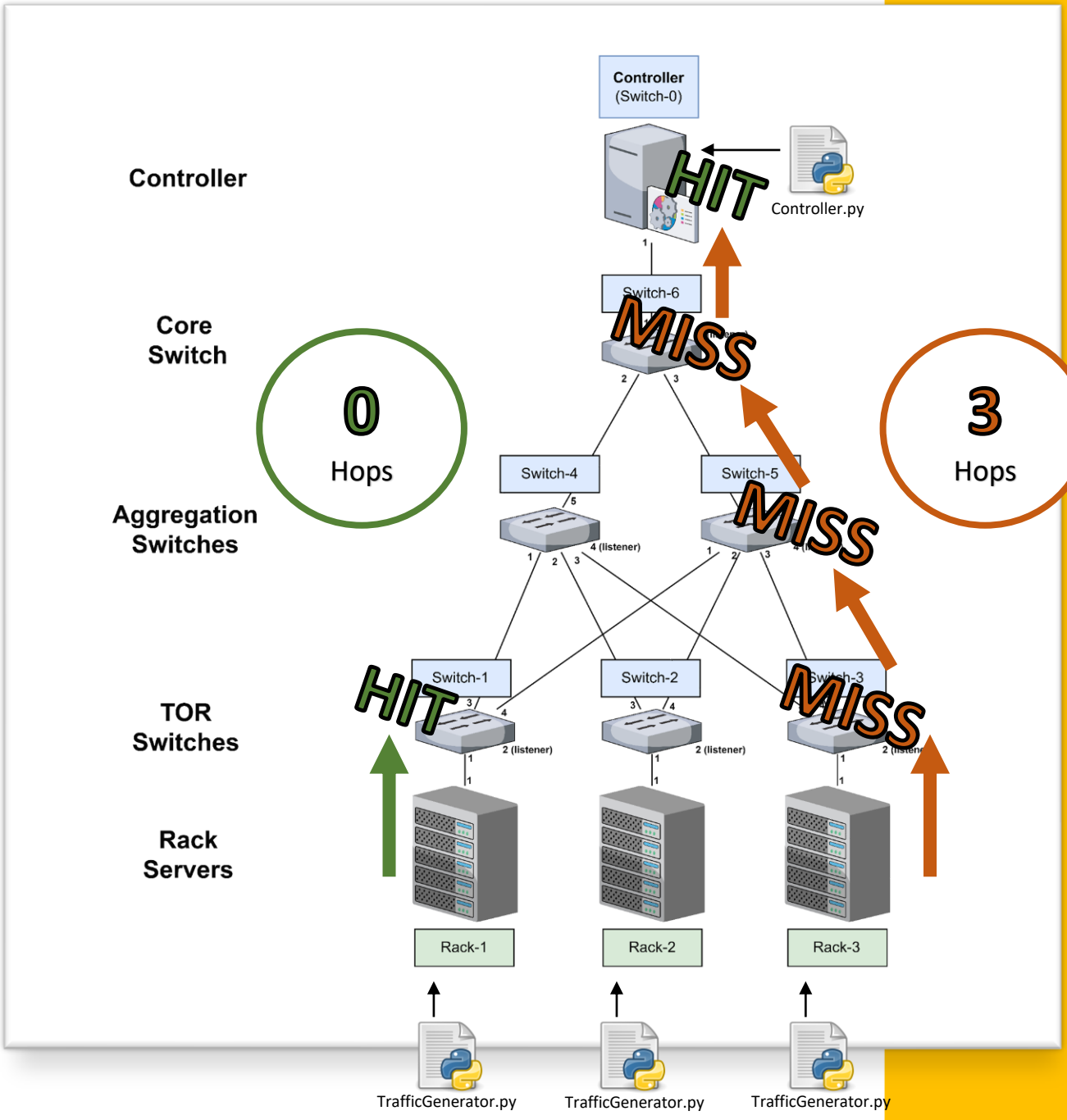


From opennetworking.org/mininet/

- Emulation ≠ Simulation
- Set up a basic **SDN DC topology**
- **Programable** switches – BMv2
- Controller-switch communication: **p4runtime** API protocol

Our Emulation Topology

- Mimic a datacnter 3-tier topology.
- Generating high-rate traffic.
- Best case - packet gets a **'hit' in ToR** switch's cache.
- Worst case - packet gets a **'miss' in every cache** until it reaches the controller.



Solution Concept

- So how do we reduce the average number of hops?
  - Insert the rules in lower hierarchy layer.
- How can we decide when to insert the new rules?
  - **Hit threshold**.

Evaluation Setup

- Eviction mechanism for cache – **LRU**.
- A **constant total cache size** in the network - 90 rules.
  - Over all switch caches
- Traffic: includes **30 shared** destinations and **10 unique** destinations for each rack.

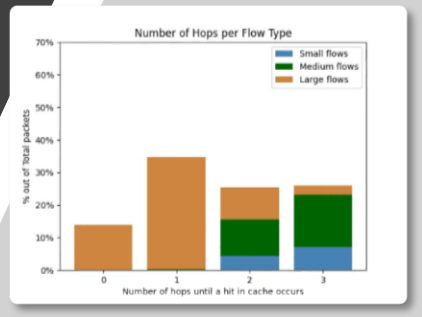
Experiments Setup

- SDN controller **monitors** the traffic in the network.
- Conducting **4** experiments:

Experiment #	Name	Cache Sizes (total: 90)			Threshold Bar	
		ToR Switches (3)	Agg. Switches (2)	Core Switch (1)	Agg. Switches	Core Switch
1	Baseline	15	15	15	10	20
2	Lower TH	15	15	15	7	3
3	Uneq. Cache	10	20	20	7	3
4	**Optimized**	10	20	20	5	3

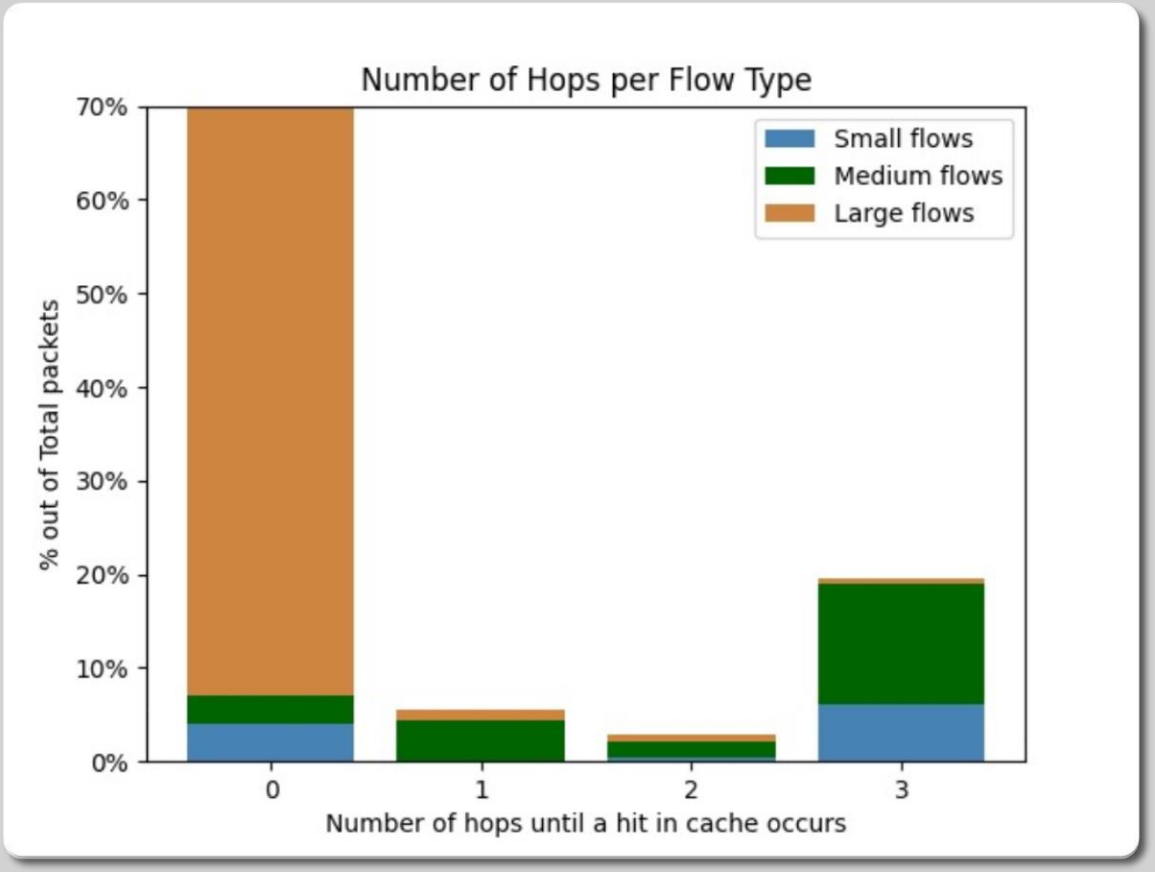
Results

Baseline



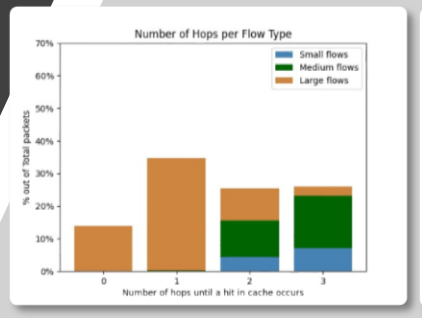
#	Name	Cache Sizes (total: 90)			Threshold Bar		Average Number of Hops
		ToR Switches (3)	Agg. Switches (2)	Core Switch (1)	Agg. Switches	Core Switch	
1	Baseline	15	15	15	10	20	**1.63**
2	**Lower TH**	**15**	**15**	**15**	**7**	**3**	**0.69**
3	Uneq. Cache	10	20	20	7	3	
4	Optimized	10	20	20	5	3	

Lower Threshold

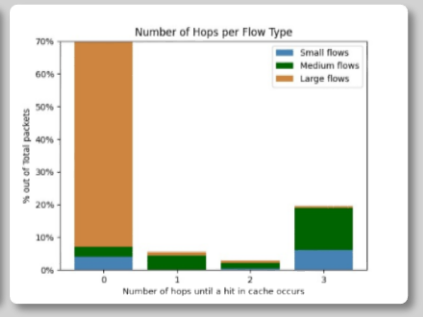


Results

Baseline

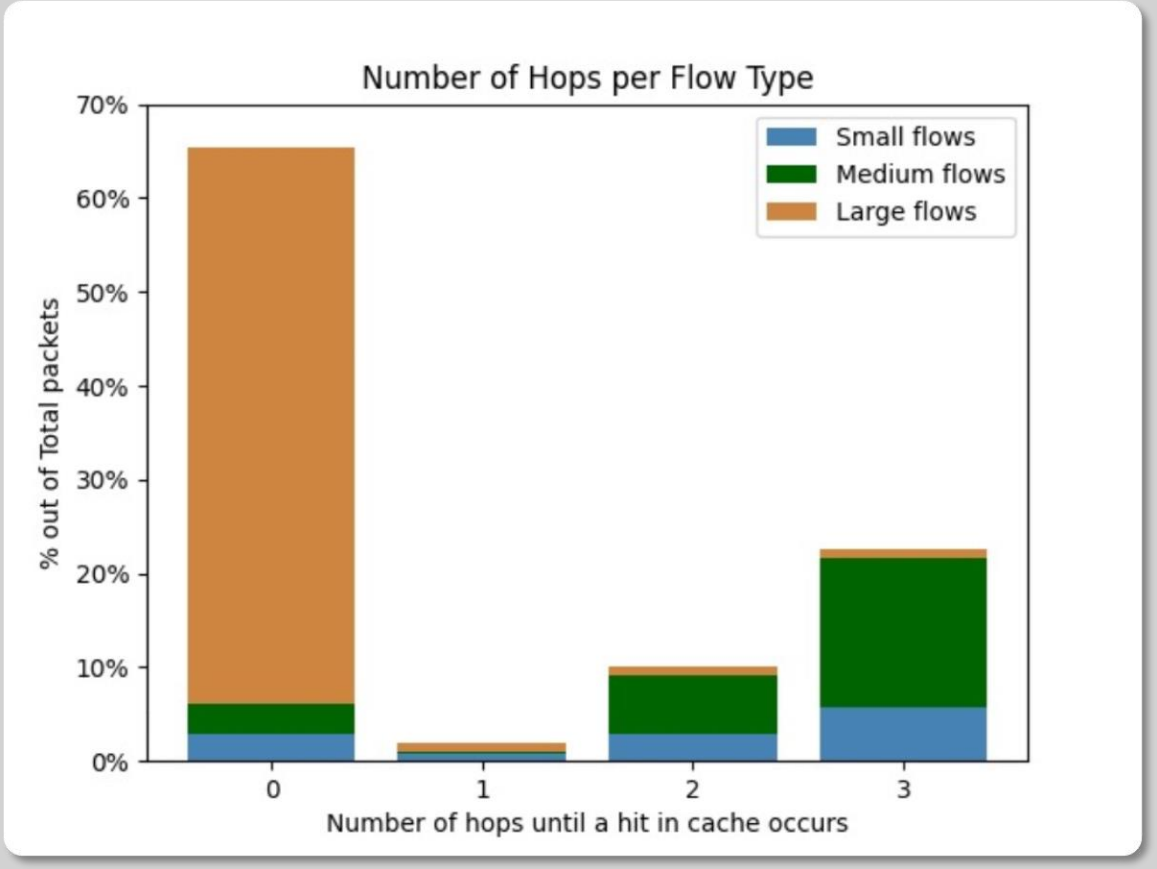


Lower Threshold



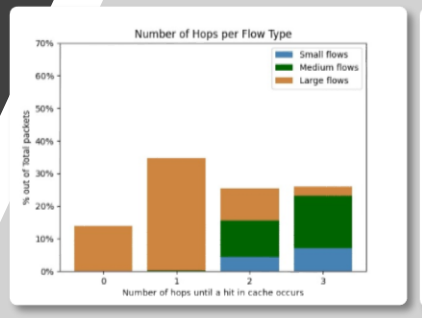
#	Name	Cache Sizes (total: 90)			Threshold Bar		Average Number of Hops
		ToR Switches (3)	Agg. Switches (2)	Core Switch (1)	Agg. Switches	Core Switch	
1	Baseline	15	15	15	10	20	**1.63**
2	Lower TH	15	15	15	7	3	**0.69**
3	**Uneq. Cache**	**10**	**20**	**20**	**7**	**3**	**0.89**
4	Optimized	10	20	20	5	3	

Unequal Cache

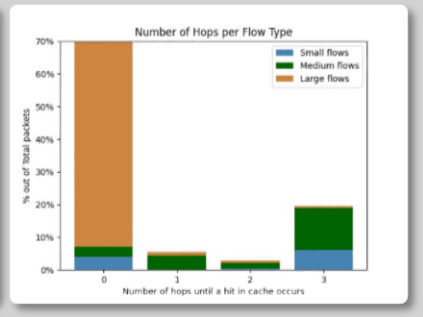


Results

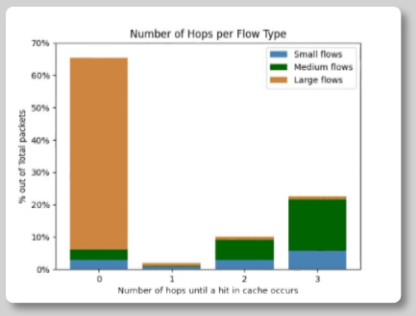
Baseline



Lower Threshold

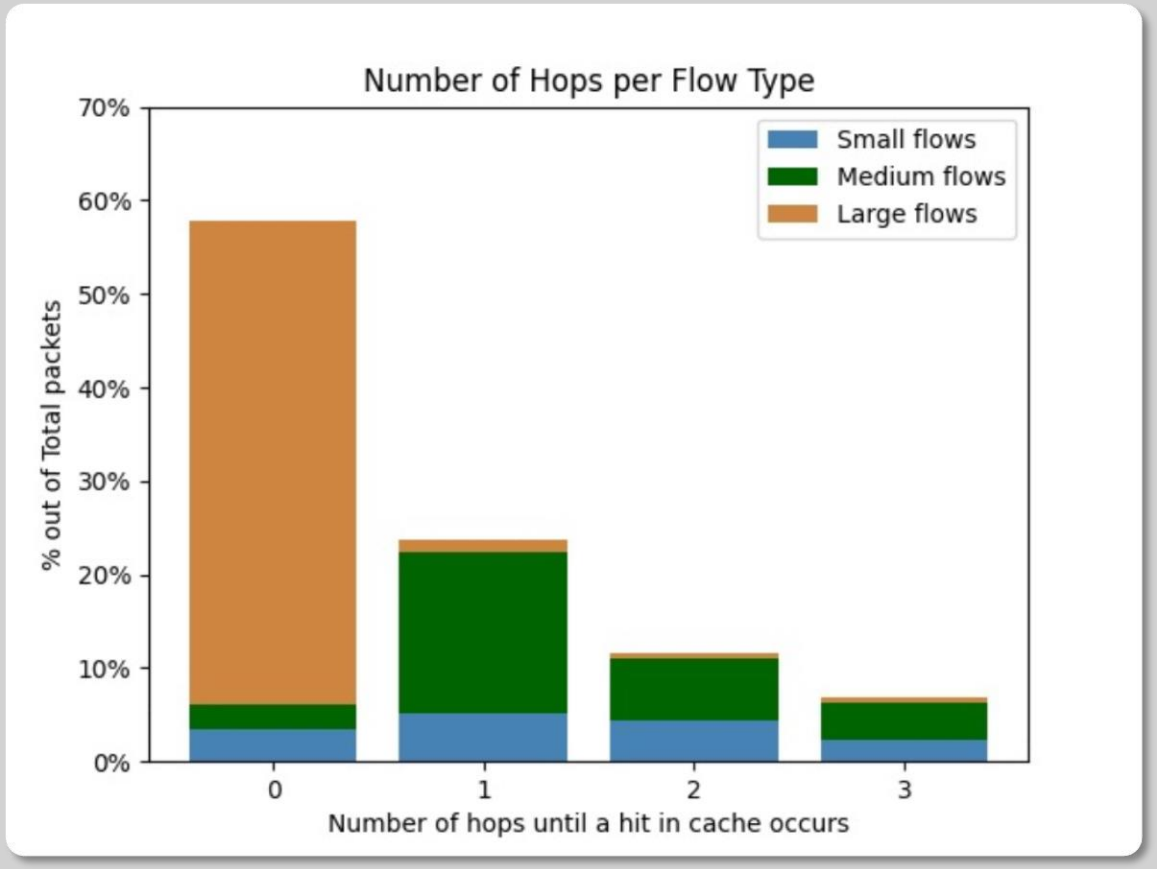


Unequal Cache



#	Name	Cache Sizes (total: 90)			Threshold Bar		Average Number of Hops
		ToR Switches (3)	Agg. Switches (2)	Core Switch (1)	Agg. Switches	Core Switch	
1	Baseline	15	15	15	10	20	**1.63**
2	Lower TH	15	15	15	7	3	**0.69**
3	Uneq. Cache	10	20	20	7	3	**0.89**
4	**Optimized**	**10**	**20**	**20**	**5**	**3**	**0.67**

Optimized



Challenges

- The ability to emulate a real DC – **Scalability**.
- Implementation within the P4 framework.
- Mininet emulation environment in a **virtual machine** on a PC with 4 cores.

Conclusion

- Caching mechanisms in datacenter switches **improves performance**.
  - Evaluating parameter – average number of hops ~ average response time.
- Non-trivial solution
- Evaluation environment should be chosen with care.