Network Scheduling in Software-defined Environments

Tal Mizrahi*◊, Yoram Moses*

*Technion – Israel Institute of Technology

◊Marvell

SDN Research Group, IRTF

IETF 96, Berlin, Germany

July 2016

Outline

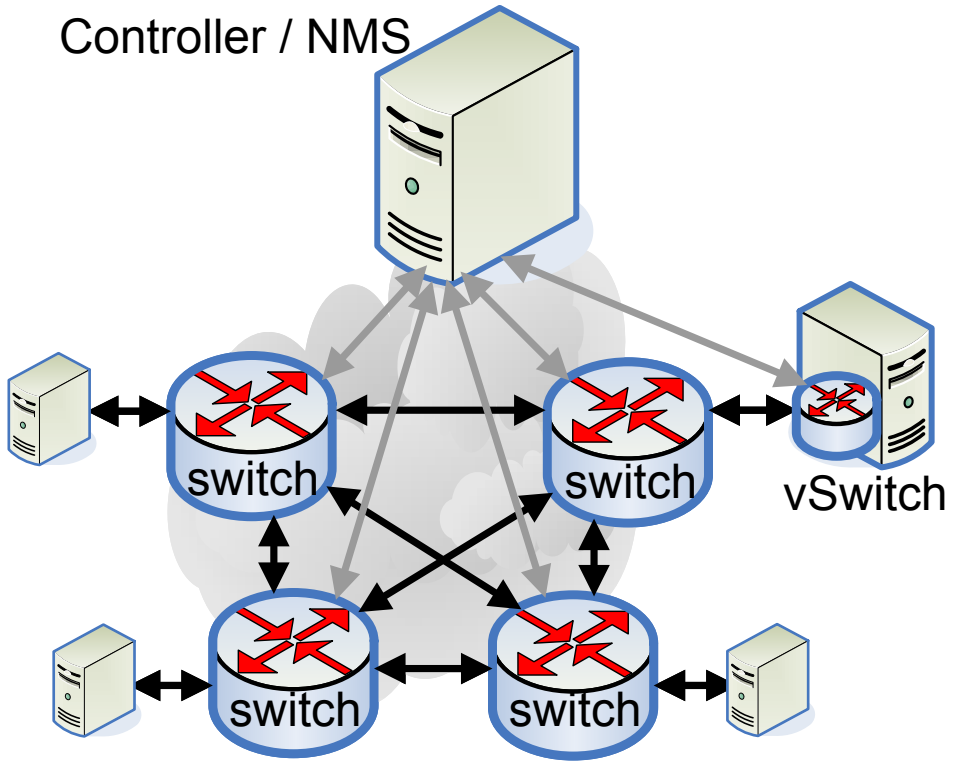
- Background
- Use cases
- Accurate scheduling
- Conclusion

Outline

- Background
- Use cases
- Accurate scheduling
- Conclusion

Software Defined & Programmable Networks

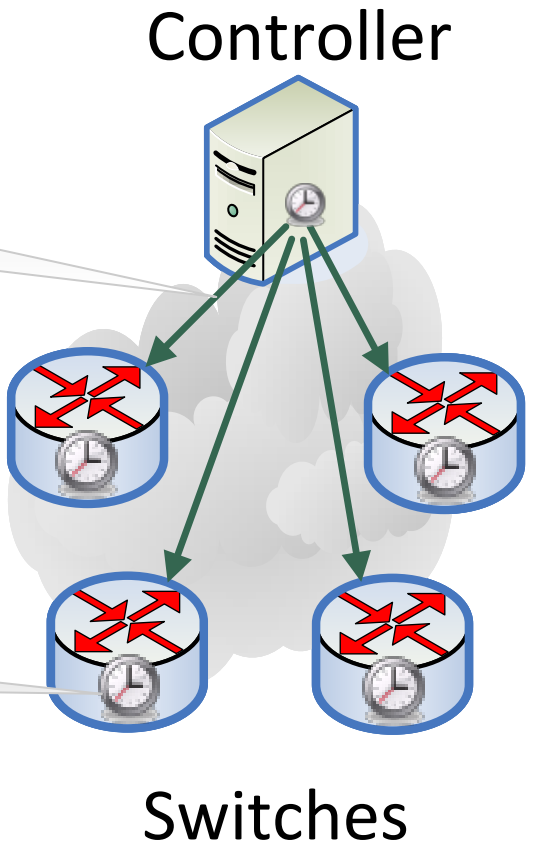
- **Centralized**
- **Flexible**
- **Dynamic**



Timed SDN at a Glance

A protocol allowing the controller to schedule network updates

Synchronized clocks

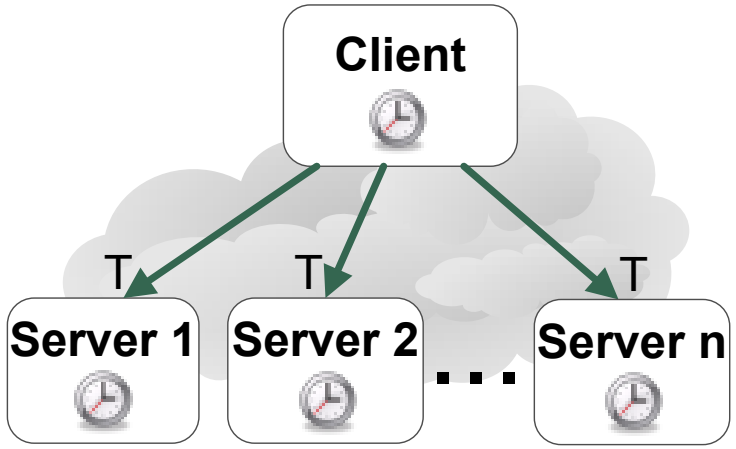


Outline

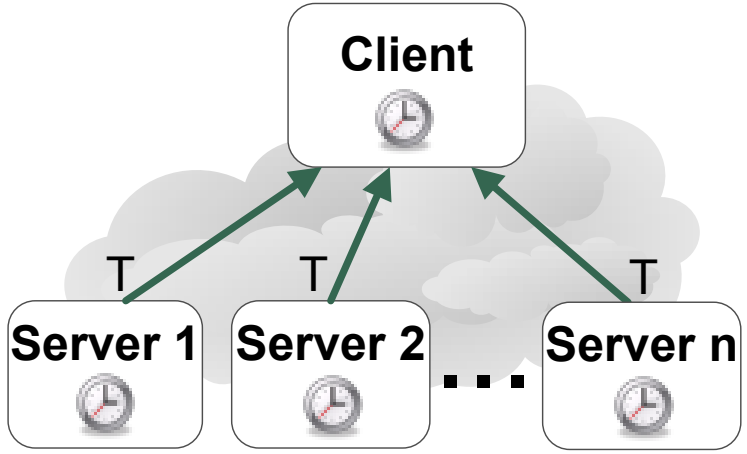
- Background
- Use cases
- Accurate scheduling
- Conclusion

Using Time in Network-managed Environments

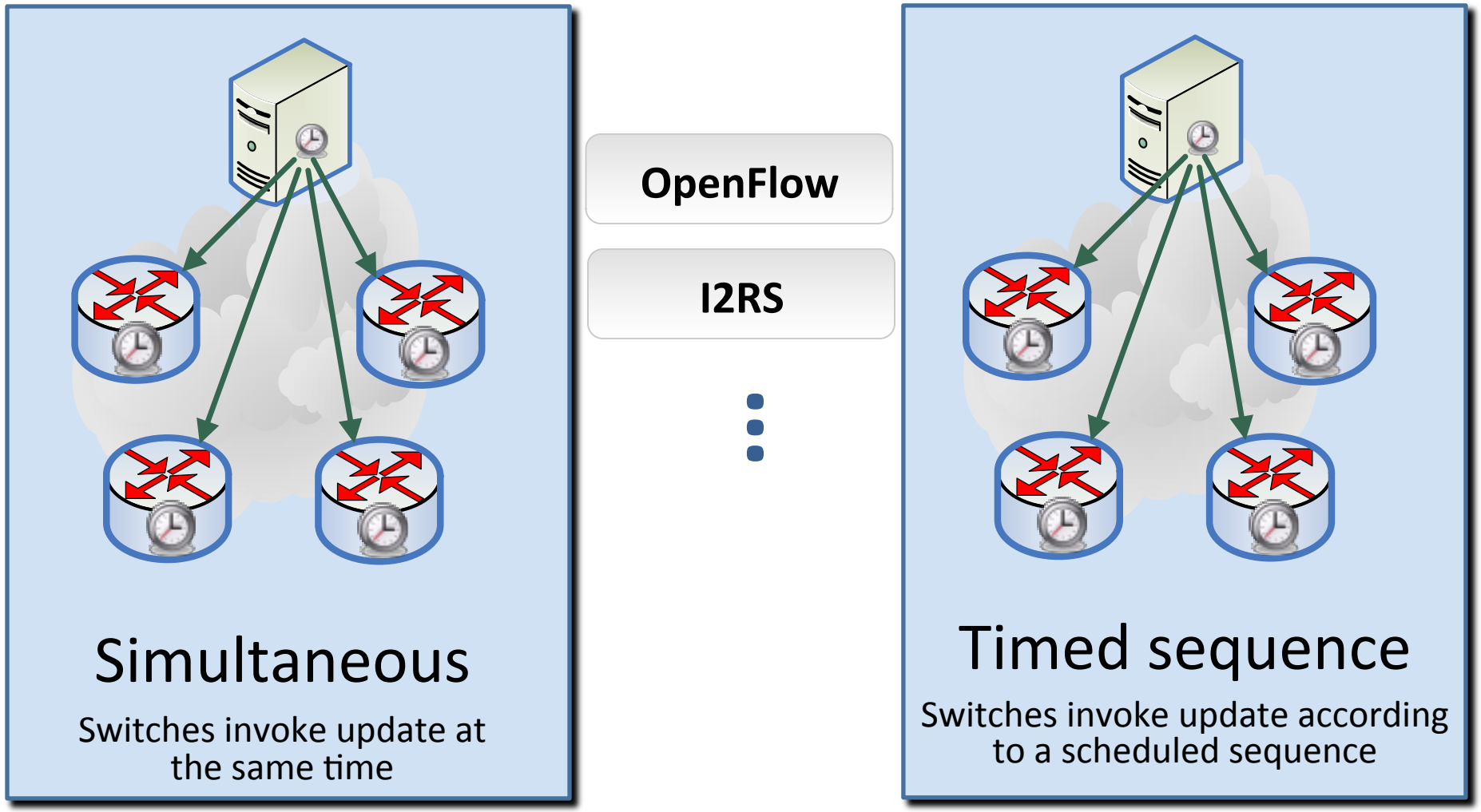
Coordinated Configuration Update



Coordinated Snapshot

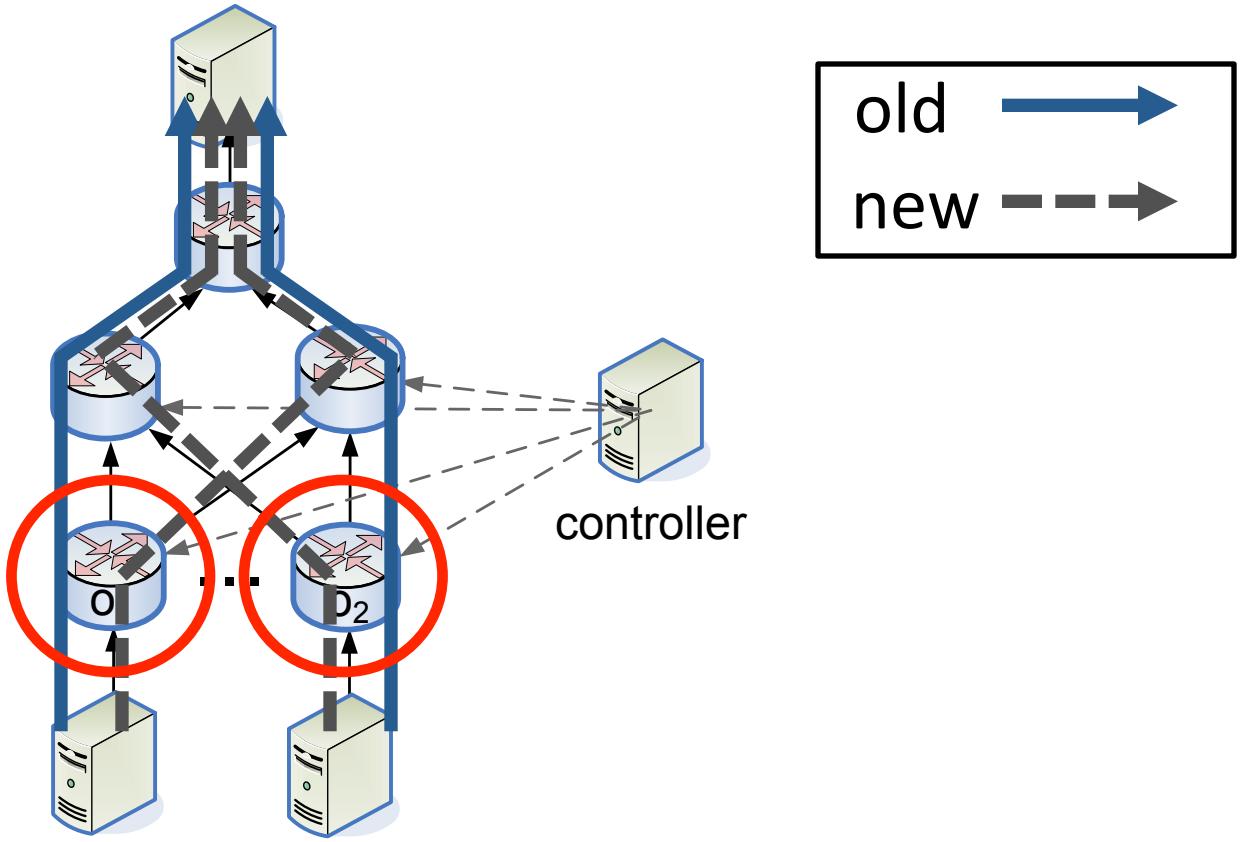


Using Time to Coordinate Path Updates



Time is Optimal for Flow Swaps [INFOCOM '16]

A key benefit of SDN:
Dynamic path allocation based on network load



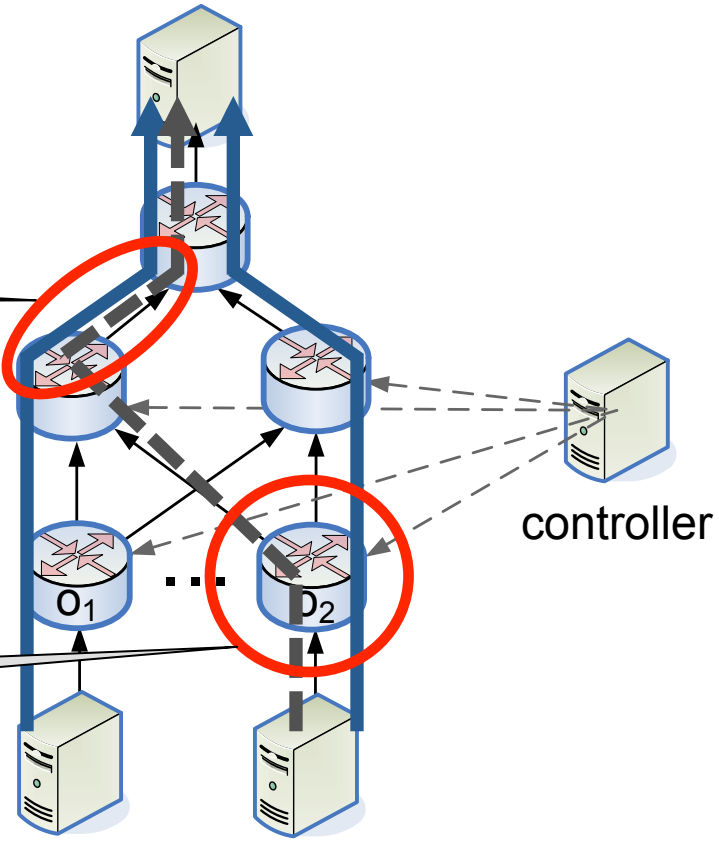
Time is Optimal for Flow Swaps [INFOCOM '16]

A key benefit of SDN:
Dynamic path allocation based on network load

If O_1 , O_2 are not updated at the same time…

Temporary congestion.

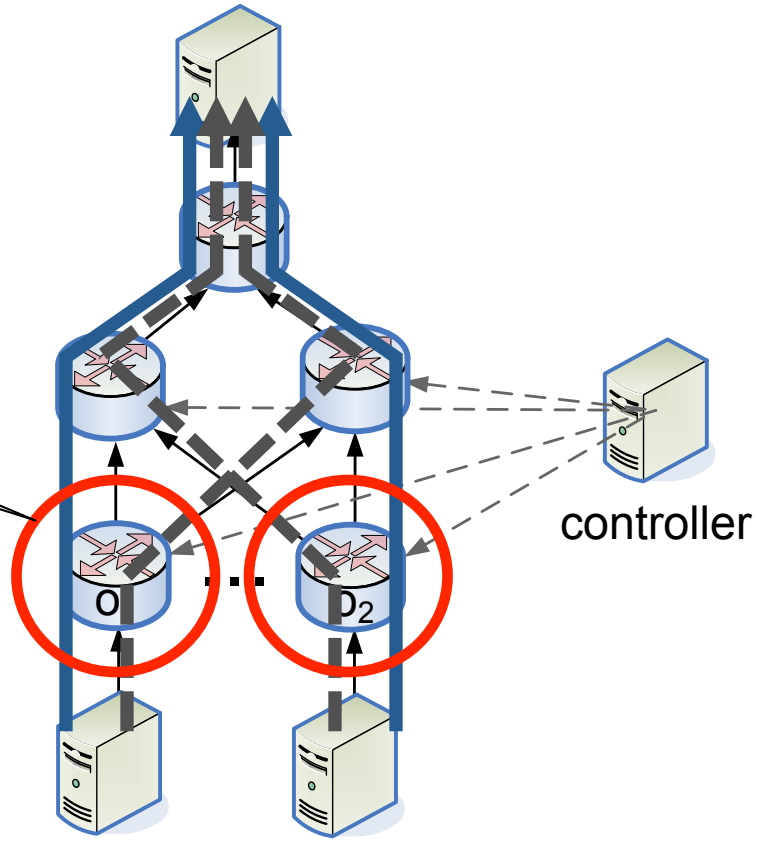
O_2 is updated first.

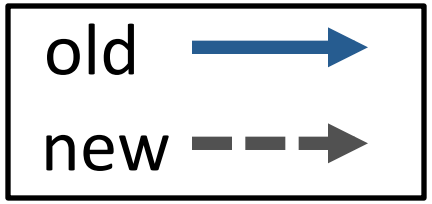


Time is Optimal for Flow Swaps [INFOCOM '16]

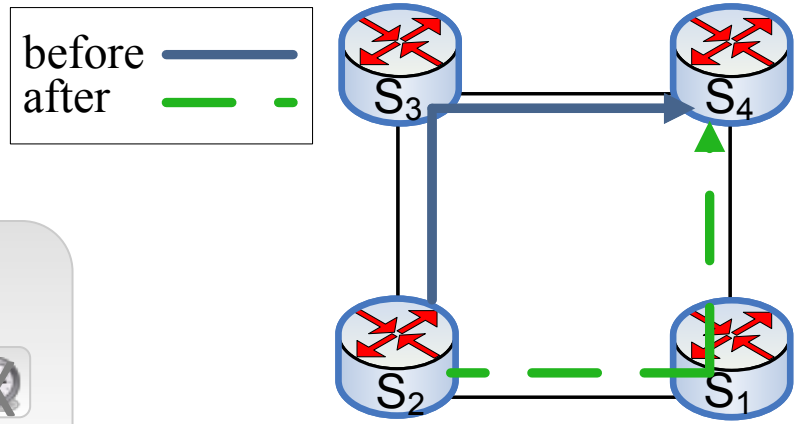
A key benefit of SDN:
Dynamic path allocation based on network load

If O_1 , O_2 are updated at the same time: No congestion!

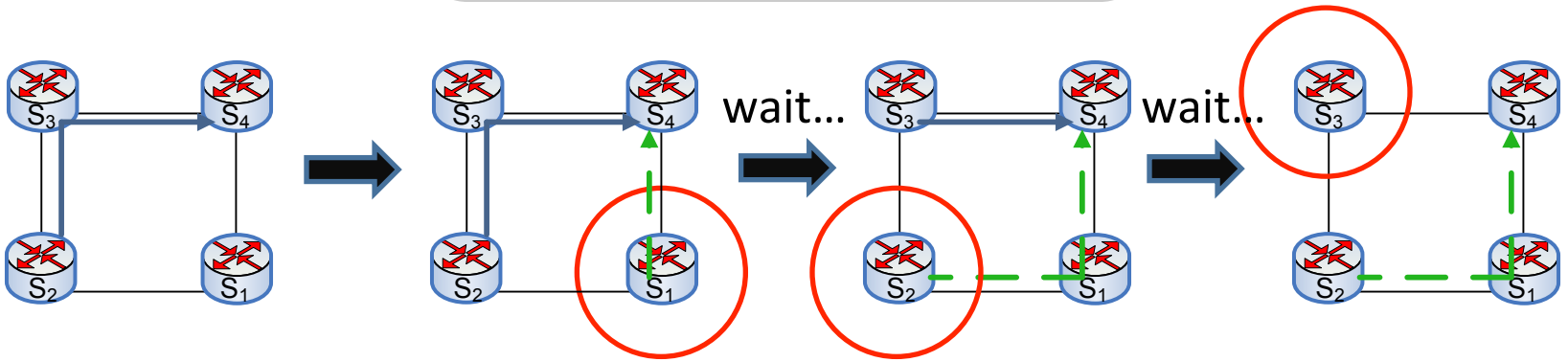




Consistent Path Update



0. The ‘before’ configuration.
1. Controller updates S_1.
2. Controller updates S_2.
3. Controller updates S_3.

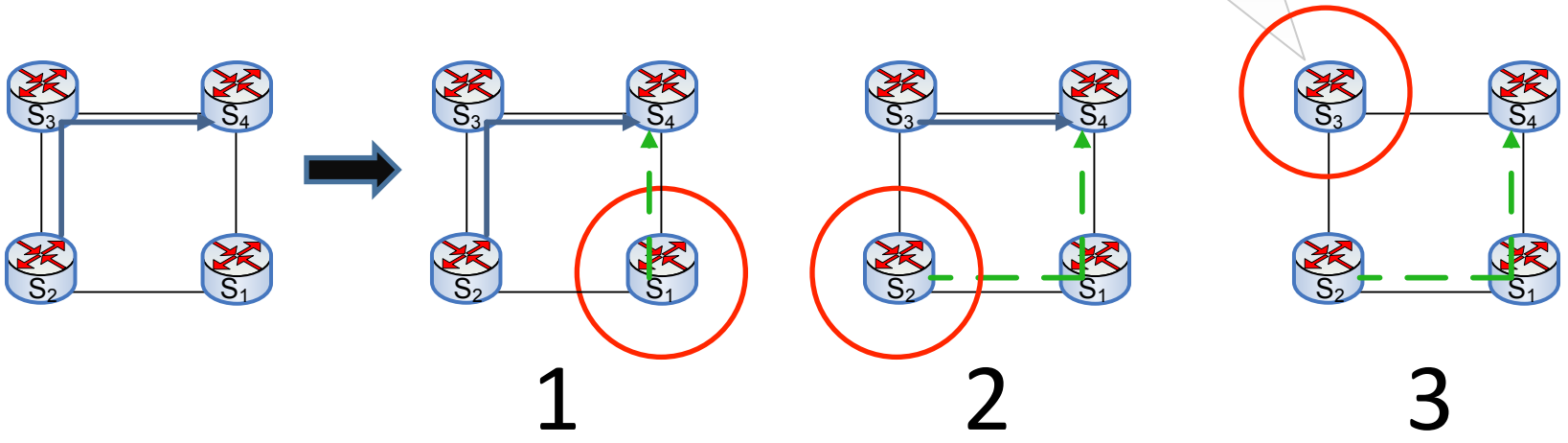


0 1 2 3

Simultaneous Updates?

En-route packets run into a ‘black hole’.

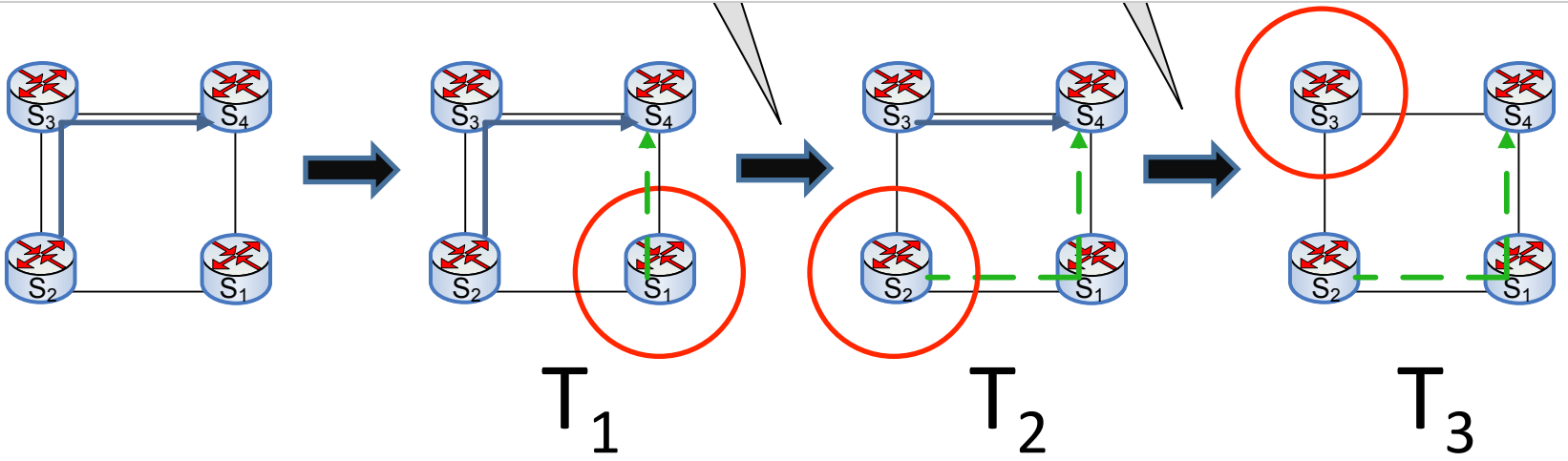
Not consistent!



Timed Multi-phase Consistent Updates [SOSR '15]

- **The controller sends timed update messages to S_1, S_2, S_3.**
- **Scheduled updates occur at times T_1, T_2, T_3.**

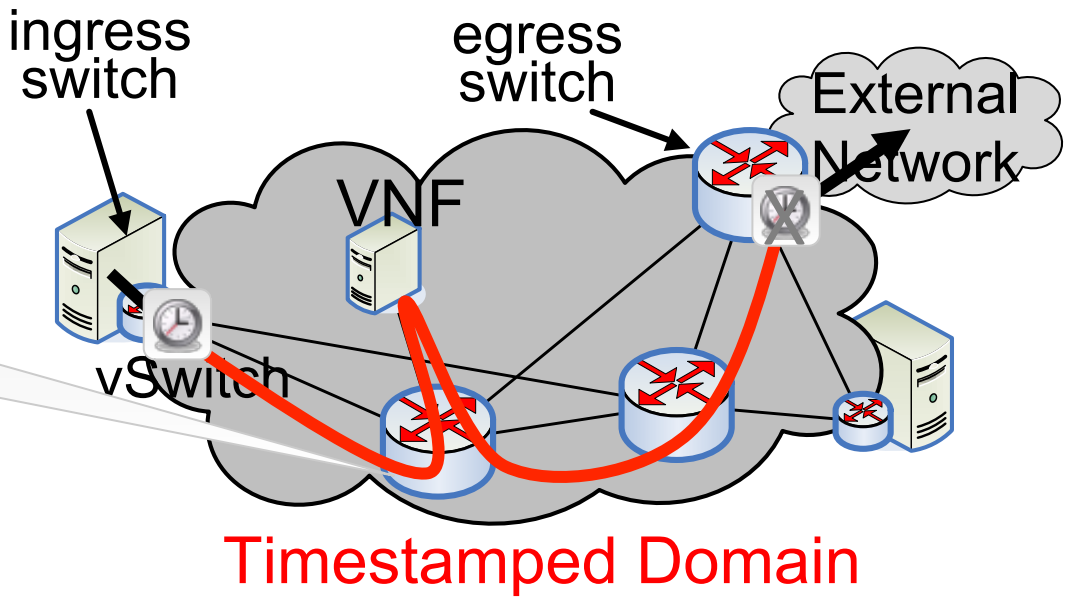
Controller does not need to wait between steps!



Using Timestamps instead of Time [SWFAN '16]

Timestamp can be used in processing of intermediate devices.

- **Delay measurement.**
- **Passive performance monitoring.**
  https://tools.ietf.org/html/draft-ietf-ippm-alt-mark
  Timestamp-based marking.
- **Policy / path decision criterion:**
  "Do action A if Timestamp≥T_0"



Timestamp is
- **Pushed by ingress switch.**
- **Removed by egress switch.**

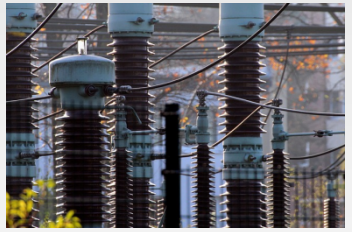
Outline

- Background
- Use cases
- Accurate scheduling
- Conclusion

Accurate Synchronization: It's Already Here

Precision Time Protocol (PTP)
[IEEE 1588 2008]
~1µsec accuracy

Mobile backhaul



Power substations

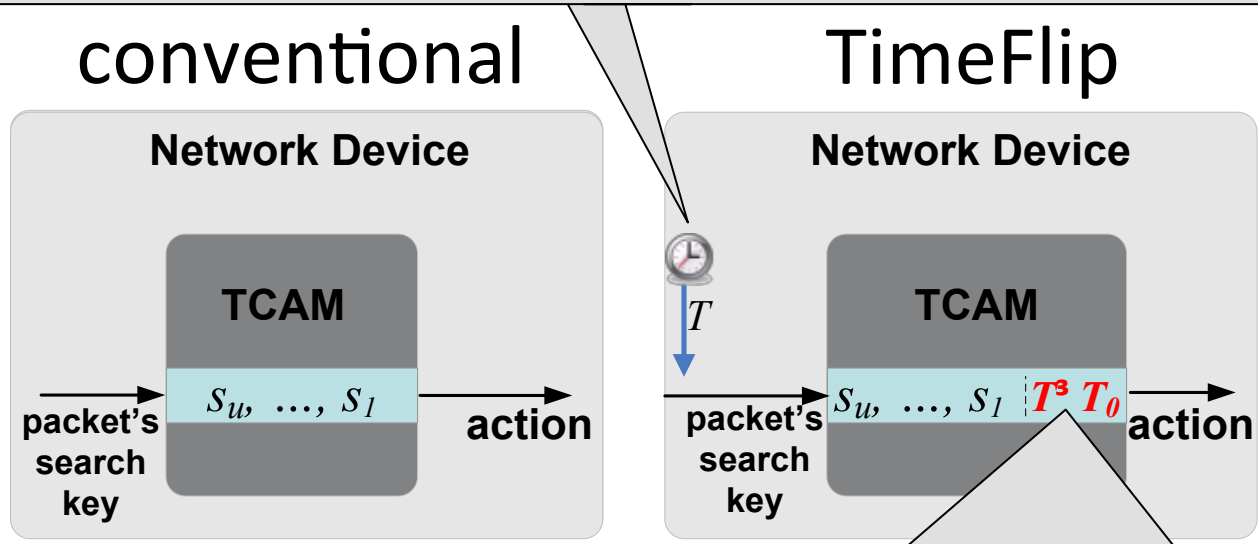
Industrial automation

Financial applications

China Mobile: over 1,000,000 PTP-enabled base stations
https://mailarchive.ietf.org/arch/attach/tictoc/pdfsY1ADO.pdf

TIMEFLIP: Timestamp-based TCAM Lookup

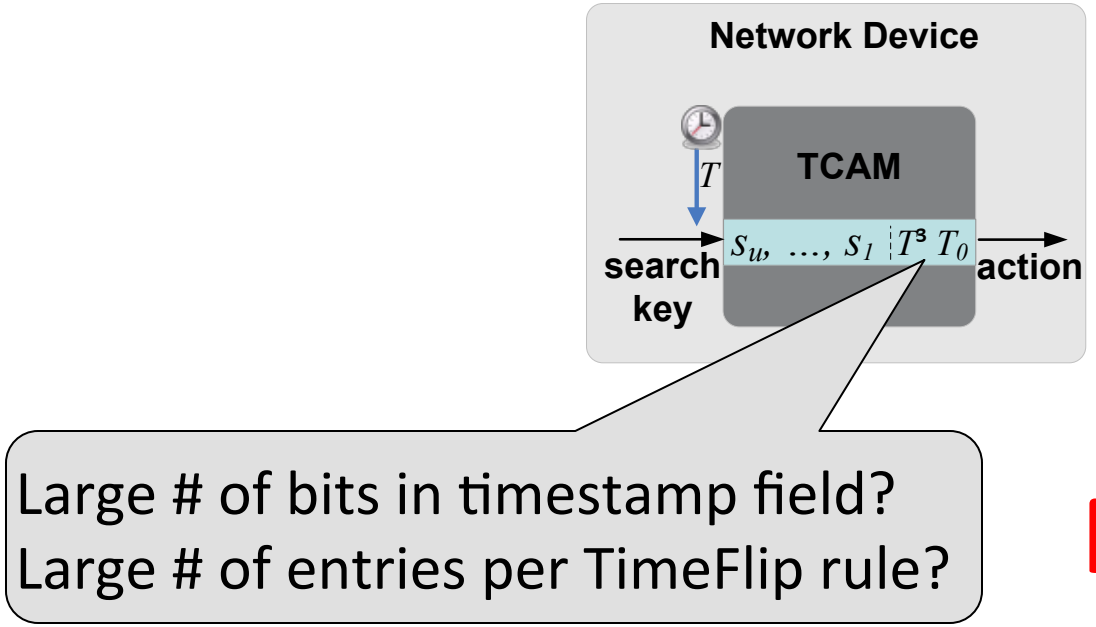
A Timestamp is attached to packet's metadata by the switch



The timestamp is used in the TCAM key – time range.

COTS switches can synchronize clocks very accurately ~ 1 µsec.
[Using IEEE 1588 Precision Time Protocol (PTP) or GPS]

The Cost of TIMEFLIP: 1 Bit / 1 Entry



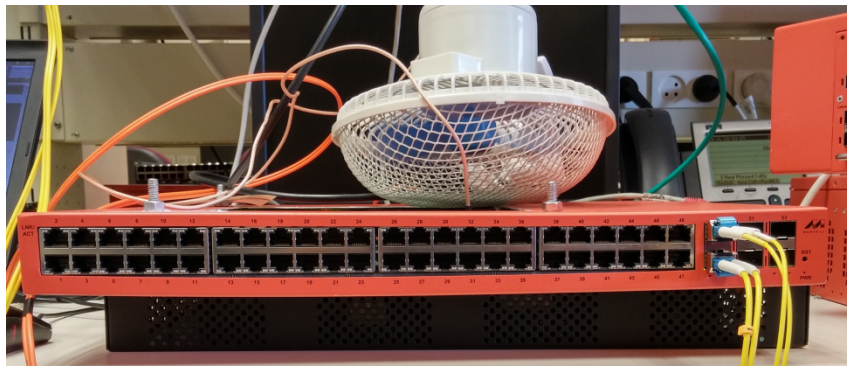
No !!

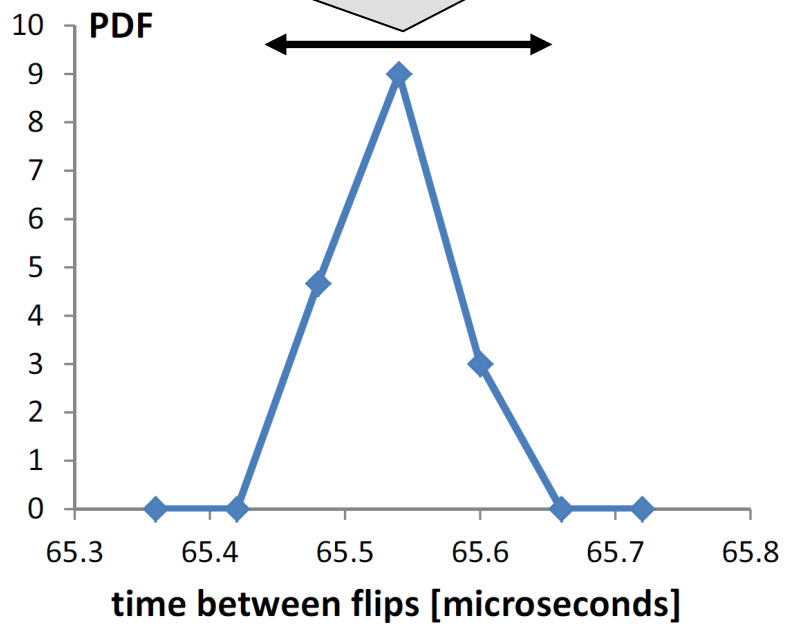
- Theorem: if $TOL \geq 2 \uparrow \lceil log \downarrow 2 (\Delta) \rceil$, then the every timestamp range requires as little as:
  - 1 bit in the timestamp field.
  - 1 TCAM entry.

TIMEFLIP: Would it Work on a Real Switch?

Yes !!

Microbenchmark: TimeFlip was tested on a Marvell 98DX4251 with a sub-microsecond accuracy.

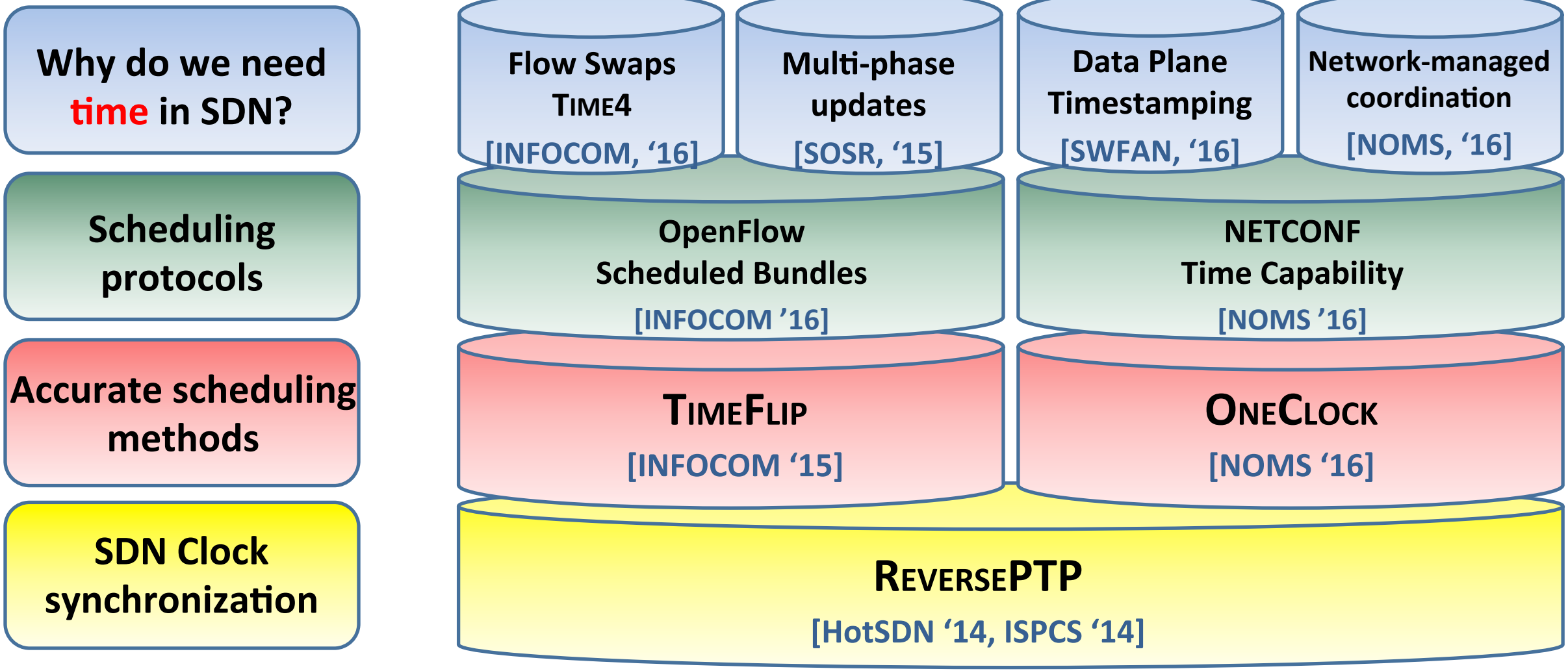




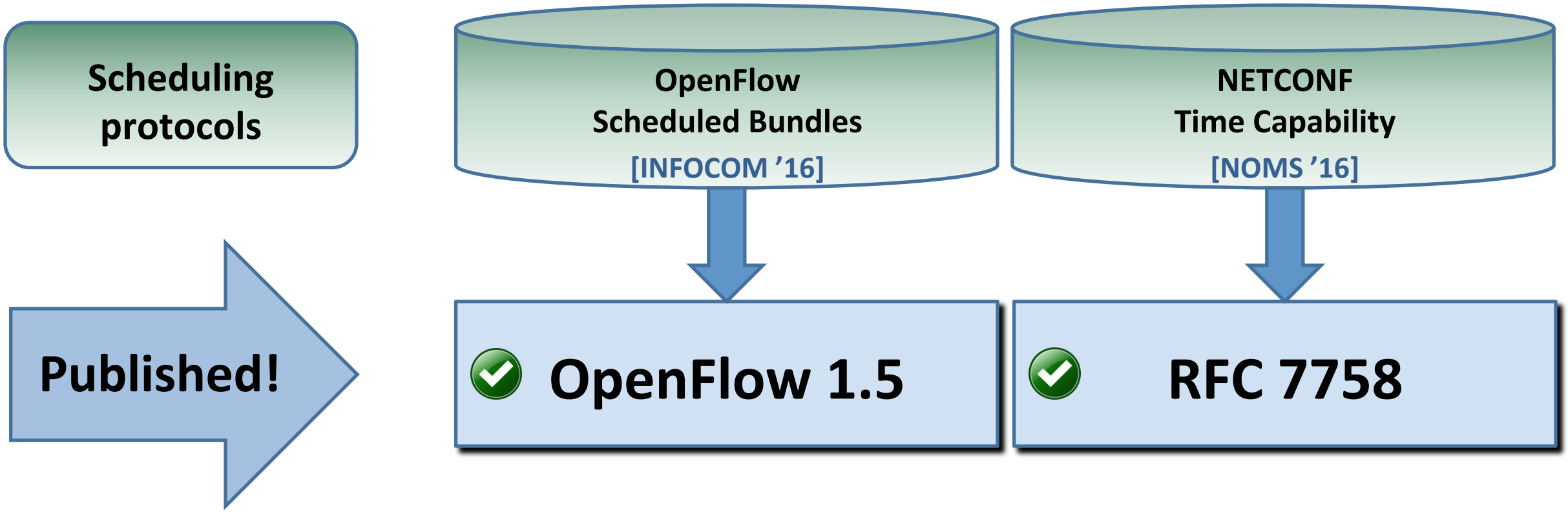
Outline

- Background
- Use cases
- Accurate scheduling
- Conclusion

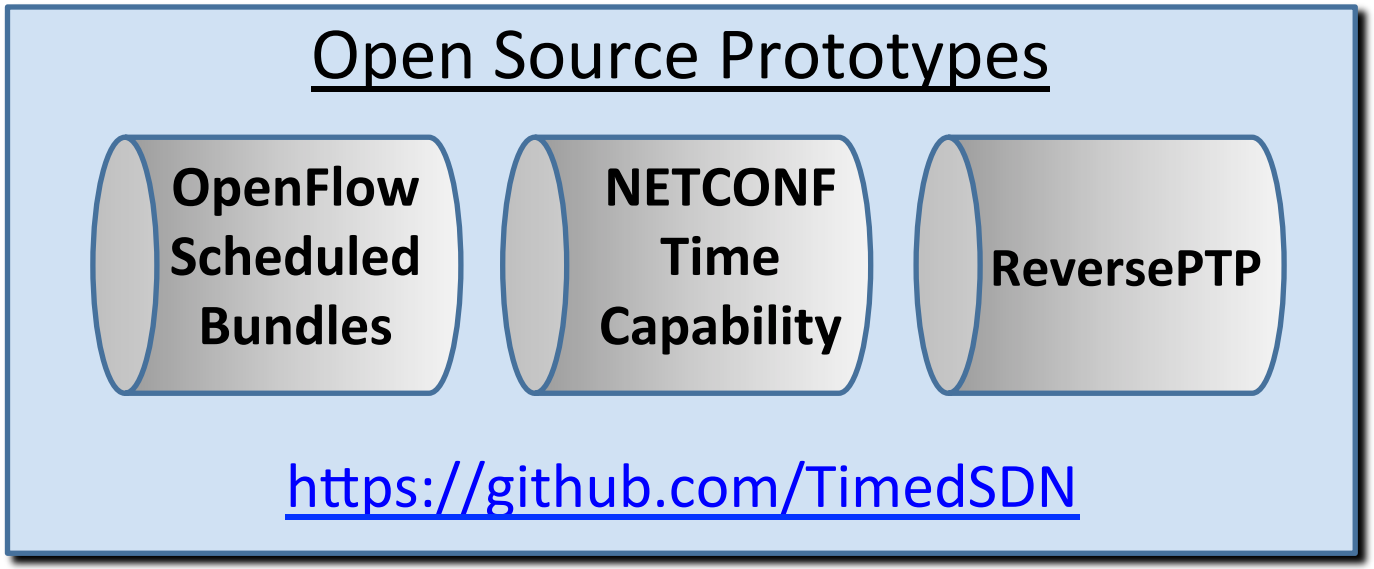
The TimedSDN Project



The TimedSDN Project – Practical Aspects

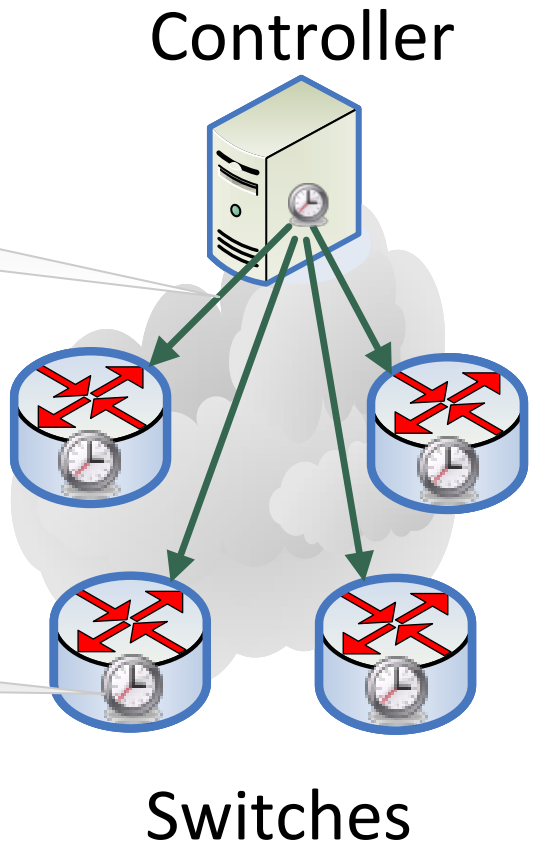


The TimedSDN Project – Practical Aspects



Timed SDN at a Glance

- A protocol allowing the controller to **schedule** network updates
- **Synchronized** clocks



The TimedSDN Project – Future Directions

Timed Updates

NETCONF
RFC 7758 ++

I2RS

...

Data plane timestamping

SFC
https://tools.ietf.org/html/draft-browne-sfc-nsh-timestamp-01

NVO3

...

Time as a network programming abstraction

P4

...

Implementation and wide-scale experiments

We would be happy to collaborate with vendors / operators who are interested!

Thanks!

The TimedSDN Project

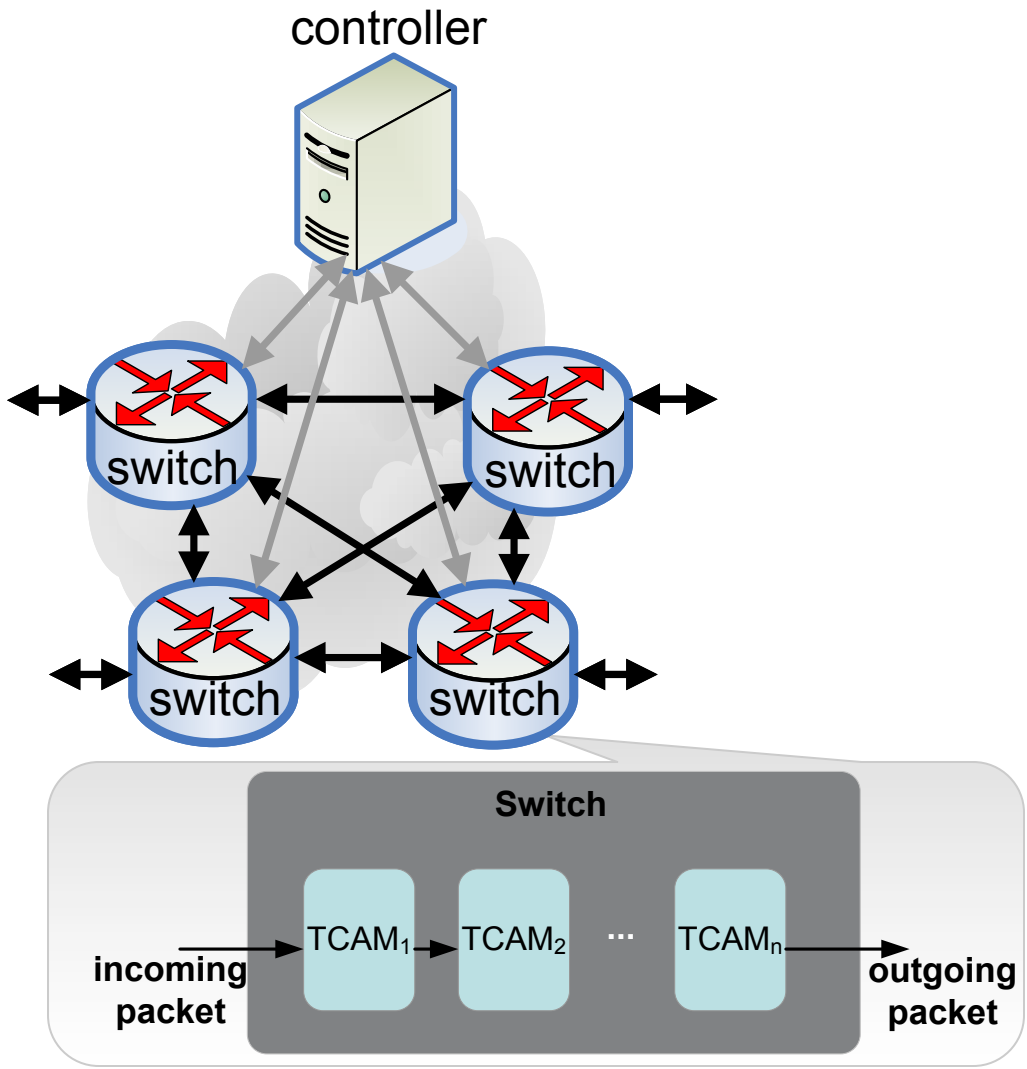
http://tx.technion.ac.il/~dew/TimedSDN.html

A Closer Look at TIMEFLIP

Further details can be found in:

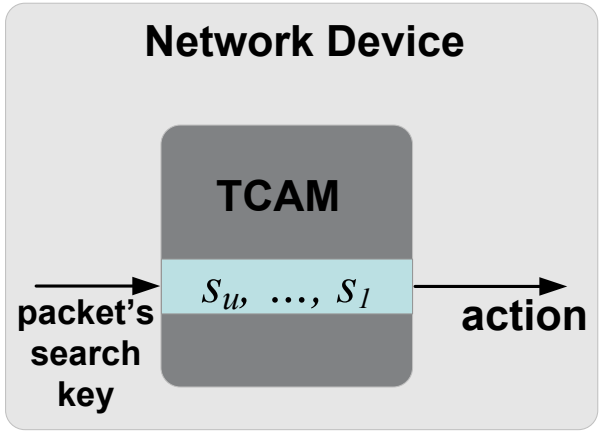
- **TimeFlip paper (http://tx.technion.ac.il/~dew/TimeFlipINFOCOM.pdf)**
- **TimeFlip presentation (http://tx.technion.ac.il/~dew/TimeFlipInfocomPres.pdf)**
- **Next few slides…**

SDN Switches



TCAM: Ternary Content Addressable Memory

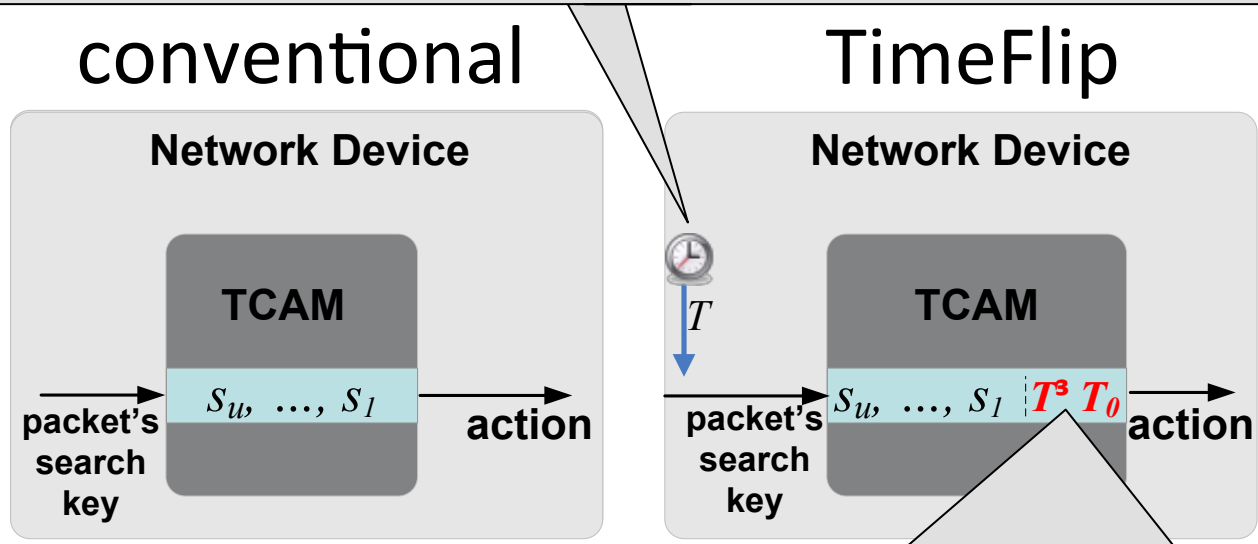
- Memory for quick searching



- Top-down search: first match "wins"
- Ternary: 0 / 1 / *
- * = don't care

TIMEFLIP: Timestamp-based TCAM Lookup

A Timestamp is attached to packet's metadata by the switch

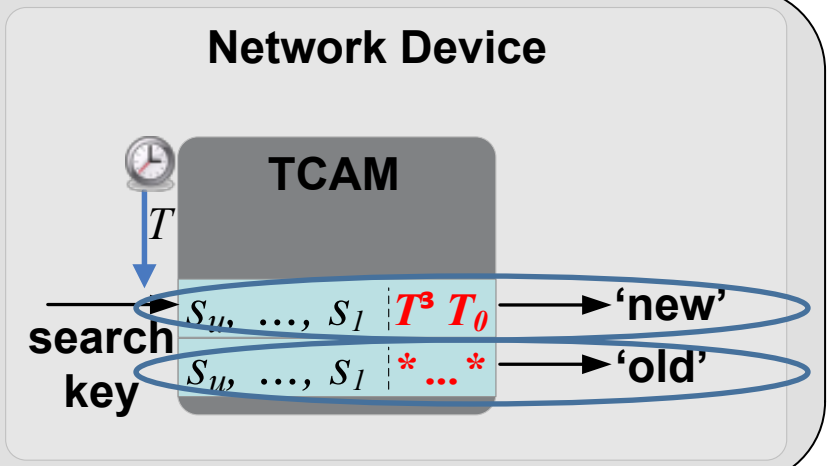


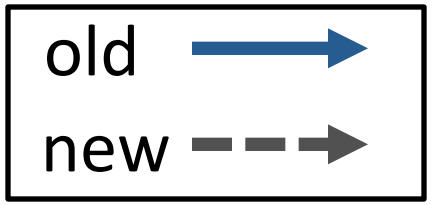
The timestamp is used in the TCAM key – time range.

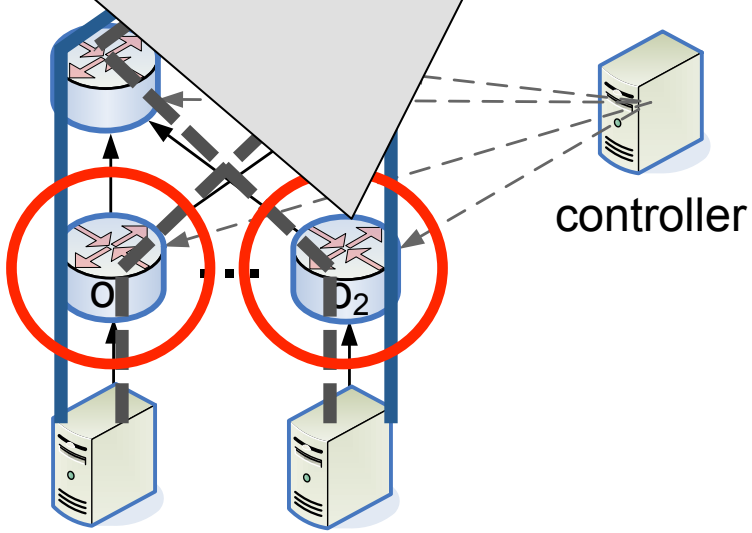
COTS switches can synchronize clocks very accurately ~ 1 µsec.
[Using IEEE 1588 Precision Time Protocol (PTP) or GPS]

Time-based Updates using TIMEFLIP

TimeFlip:
switch <u>accurately</u>
starts using ‘new’ at T_0.







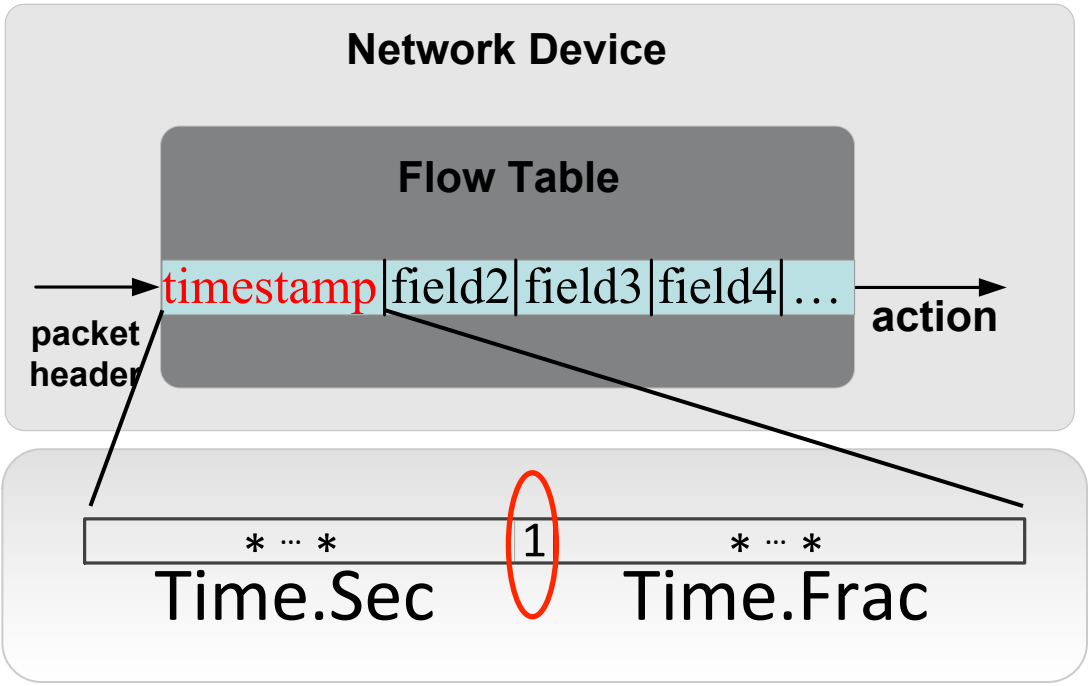
Example: Timestamp Format

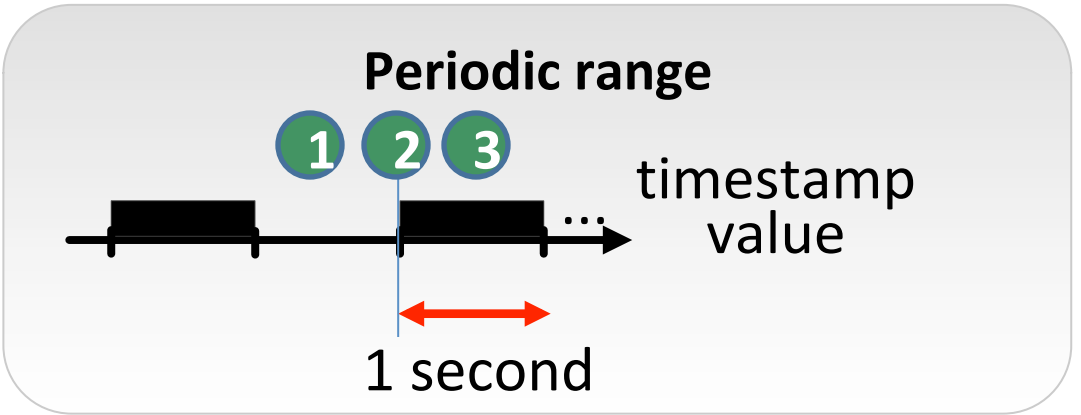
Network Time Protocol (NTP) timestamp format:

Example: TIMEFLIP in Practice

Goal: schedule an update to be performed at 2016-07-22 11:41:11 (at the beginning of the second)

Procedure: (steps also appear in the figure)
Step 1: install a TimeFlip with '1' in the lsbit of the seconds field, and 'don't care' in the rest.
Step 2: the update becomes effective exactly at the turn of the second.
Step 3: make the TCAM rule permanent by writing 'don't care' to the lsbit of the timestamp field.





References

[1] T. Mizrahi, Y. Moses, "Software Defined Networks: It's About Time", IEEE INFOCOM, 2016.

[2] T. Mizrahi, Y. Moses, "OneClock to Rule Them All: Using Time in Networked Applications", IEEE/IFIP Network Operations and Management Symposium (NOMS), 2016.

[3] T. Mizrahi, E. Saat, Y. Moses, "Timed Consistent Network Updates", ACM SIGCOMM Symposium on SDN Research (SOSR), 2015.

[4] T. Mizrahi, E. Saat, Y. Moses, "Timed Consistent Network Updates in Software Defined Networks", IEEE/ACM Transactions on Networking (ToN), 2016.

[5] T. Mizrahi, O. Rottenstreich, Y. Moses, "TimeFlip: Scheduling Network Updates with Timestamp-based TCAM Ranges", IEEE INFOCOM, 2015.

[6] T. Mizrahi, Y. Moses, "Time-based Updates in Software Defined Networks", workshop on hot topics in software defined networks (HotSDN), 2013.

[7] T. Mizrahi, Y. Moses, "On the Necessity of Time-based Updates in SDN", Open Networking Summit (ONS), 2014.

[8] T. Mizrahi, Y. Moses "Using ReversePTP to Distribute Time in Software Defined Networks", International IEEE Symposium on Precision Clock Synchronization for Measurement, Control and Communication, (ISPCS), 2014.

[9] T. Mizrahi, Y. Moses, "ReversePTP: A Software Defined Networking Approach to Clock Synchronization", workshop on hot topics in software defined networks (HotSDN), 2014.

[10] T. Mizrahi and Y. Moses, "ReversePTP: A clock synchronization scheme for software defined networks", International Journal of Network Management (IJNM), 2016.

[11] T. Mizrahi, Y. Moses, "Time4: Time for SDN", arXiv preprint arXiv:1505.03421, 2016.

[12] T. Mizrahi, Y. Moses, "The Case for Data Plane Timestamping in SDN", IEEE INFOCOM Workshop on Software-Driven Flexible and Agile Networking (SWFAN), 2016.

[13] T. Mizrahi, Y. Moses, "Time Capability in NETCONF", RFC 7758, 2016.

[14] Open Networking Foundation, OpenFlow switch specification, Version 1.5.0, 2015.

[15] Open Networking Foundation, OpenFlow extensions 1.3.x package 2, 2015.