Evaluation of Topology-based OpenFlow Controller for Multipath TCP on PRAGMA-ENT

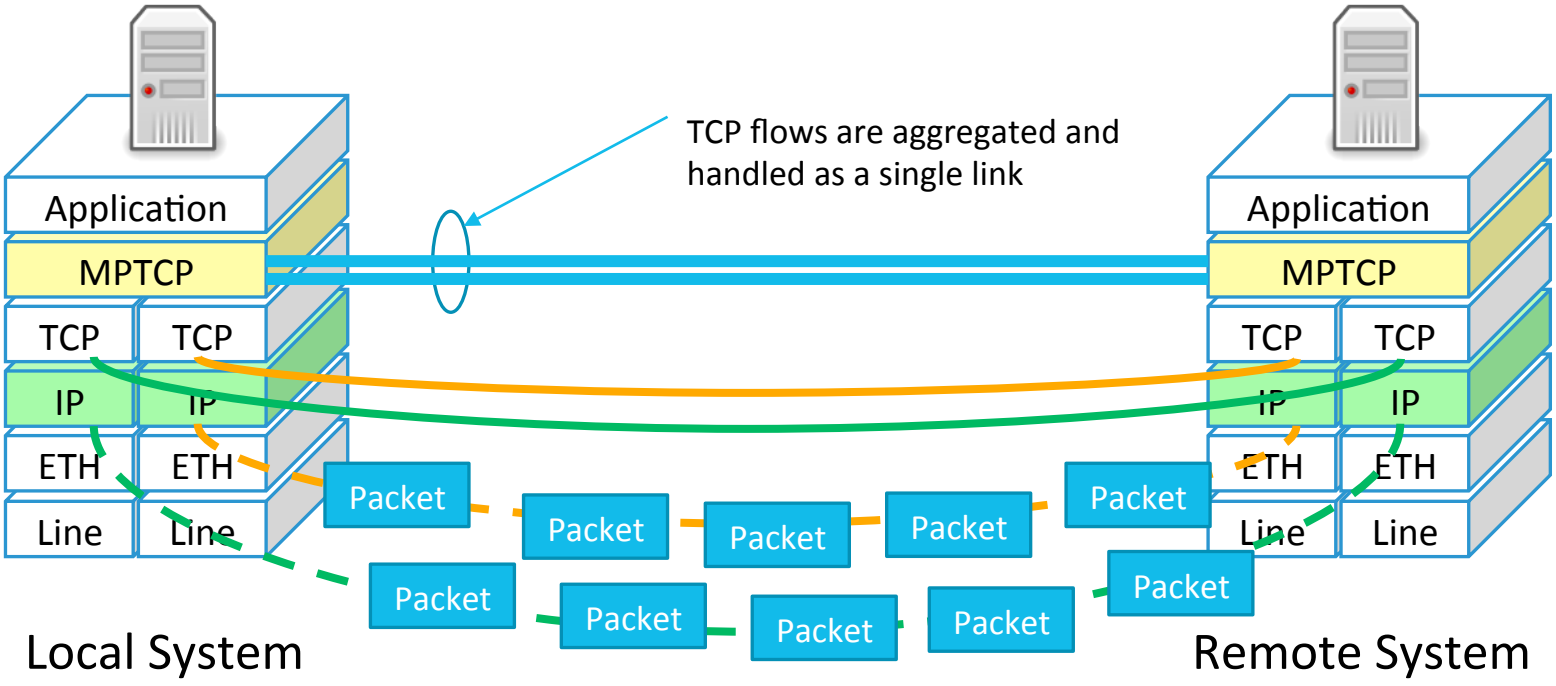
Chawanat Nakasan, Yasuhiro Watashiba,
Kohei Ichikawa, Hajimu Iida,
Putchong Uthayopas*

Nara Institute of Science and Technology
*Kasetsart University

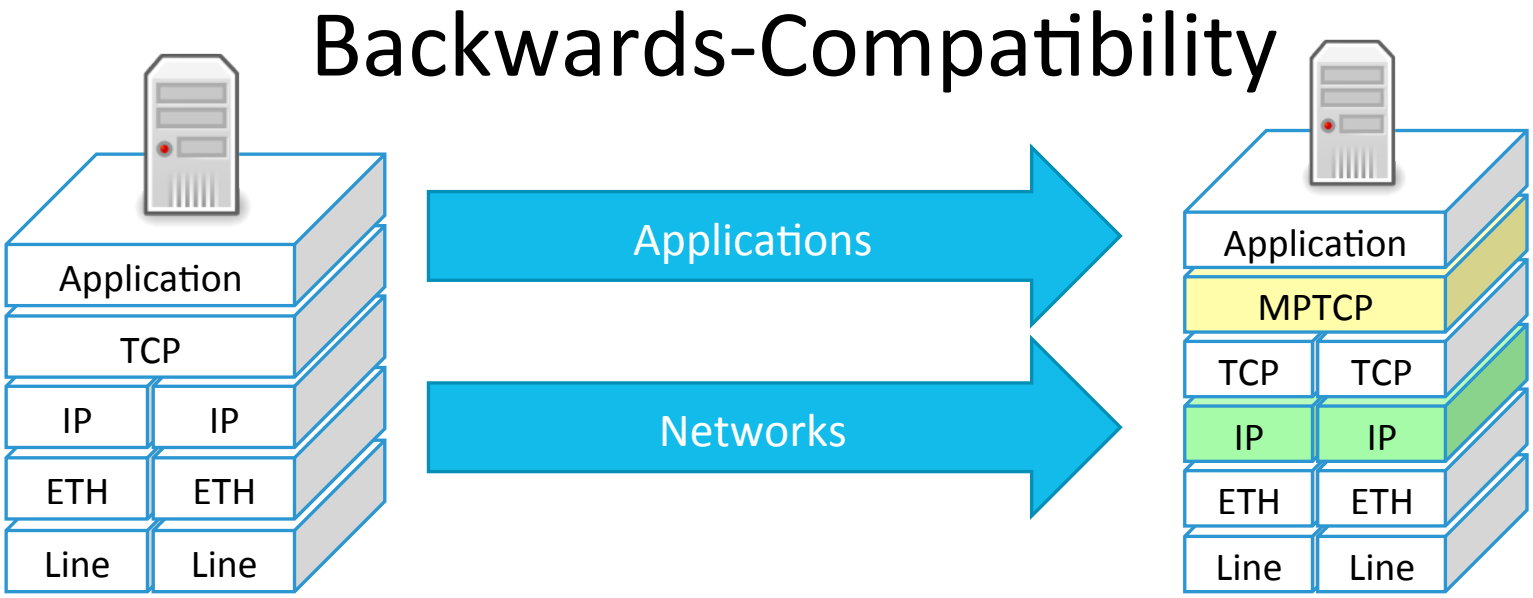
2015/10/07, PRAGMA 29 Lightning Talk @Universitas Indonesia

MPTCP aggregates TCP so application needs one socket.

Everything else is transparently handled by MPTCP in the OS kernel.

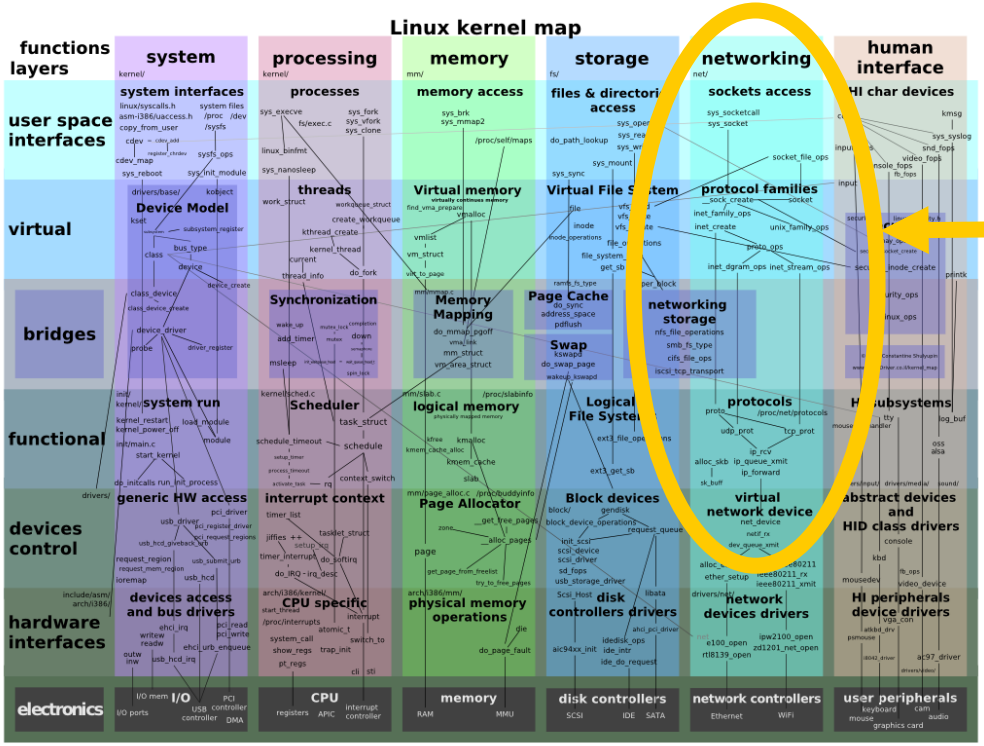


Benefits of MPTCP



Limitation: MPTCP needs OS modification

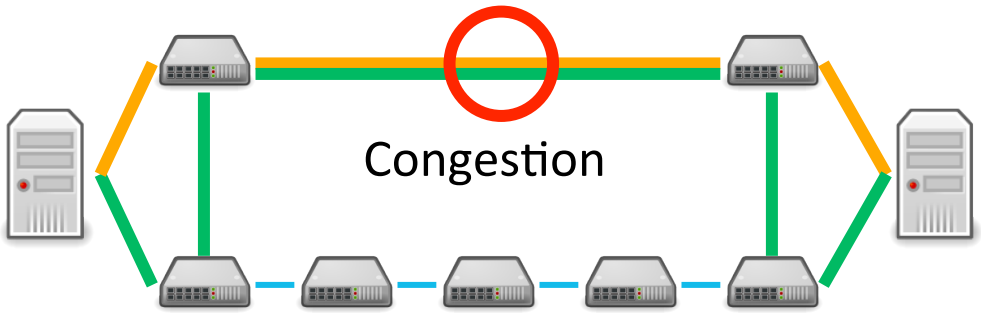
MPTCP must be implemented as a separate kernel, not even a module. Support is confined usually into the active development community.



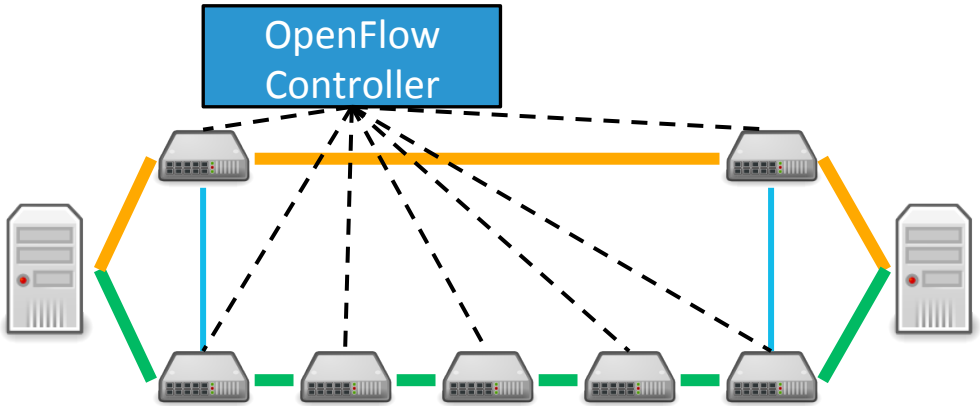
Around here

https://upload.wikimedia.org/wikipedia/commons/5/5b/Linux_kernel_map.png

Limitation:
Cannot route itself, may not be fully efficient with some networks



Congestion caused by traditional shortest-path routing



Multipath routing can avoid congestion

Algorithms

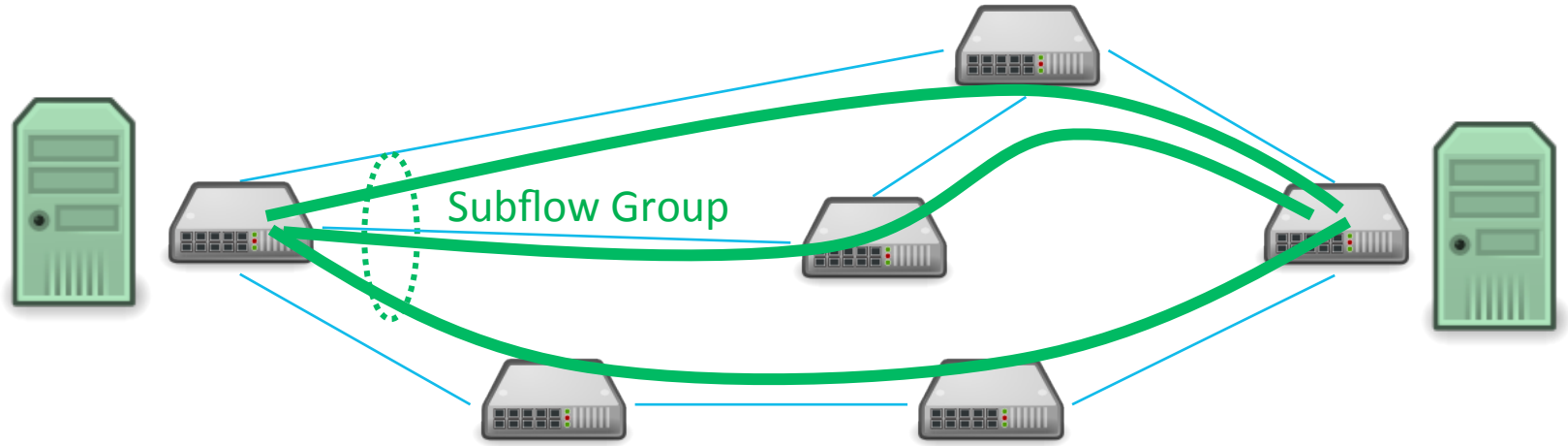
Routing

Flow Management

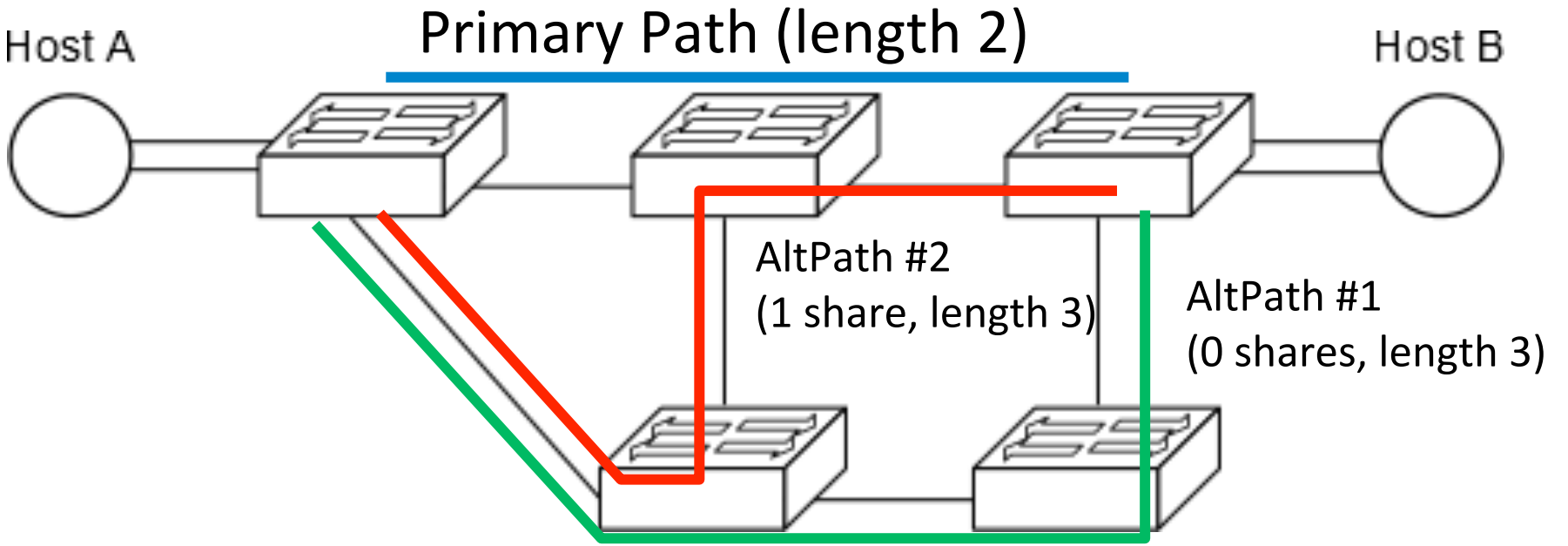
Goals

Flow management determines which subflow belongs to which MPTCP session

Routing algorithm distribute flows to different paths, minimizing path conflict that leads to bottlenecks.



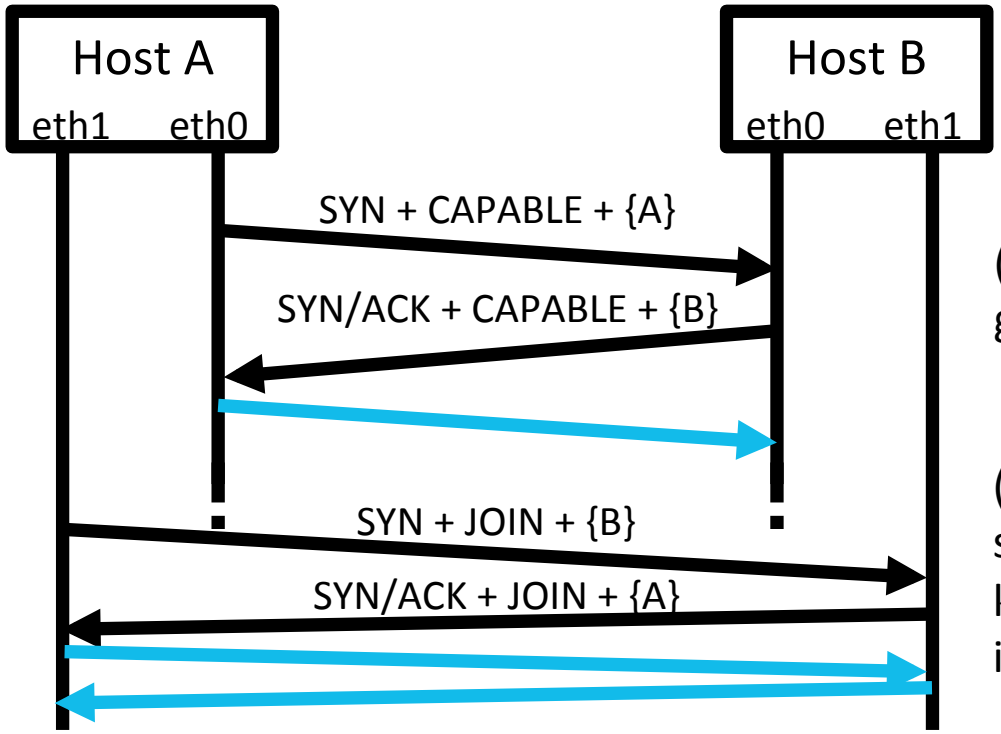
Routing Mechanism is based on shortest path



Sample routing results from our algorithm

The TCP Handshake

One of the most well-known topics about TCP is handshake. MPTCP expands on that by adding more information.



(1) Host introduce themselves by giving out its own keys.

(2) When hosts want to create another subflow, it sends the other party's key to indicate which MPTCP session it wants to talk to.

{X} = X's MPTCP session key

pending_capable

From	To	key	pathset

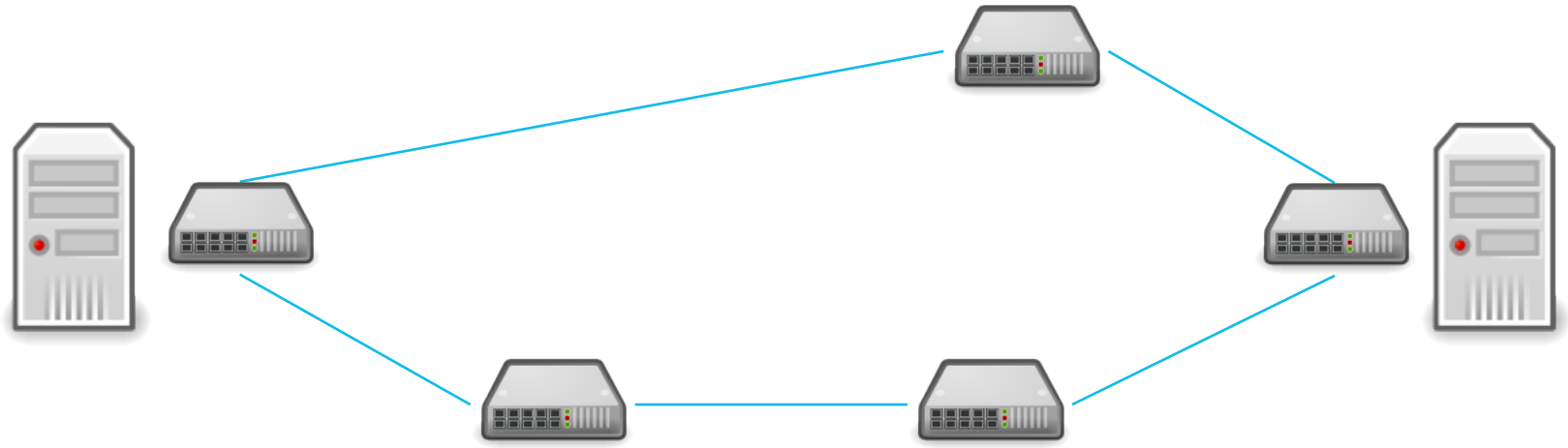
pending_join

From	To	key	pathset

established

From	To		

X = X's IP address & Port
{X} = X's MPTCP Hash/Key



Host A

Host B

Routing

Flow Management

pending_capable

From	To	key	pathset

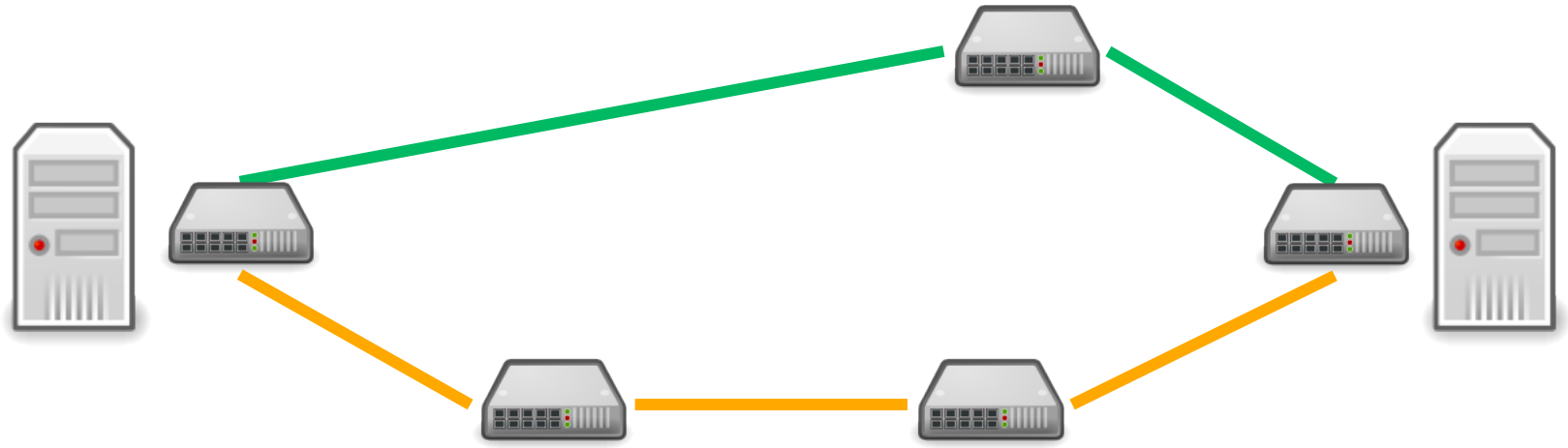
pending_join

From	To	key	pathset
A	B	{B}	A=>B[1]
B	A	{A}	B=>A[1]

established

From	To	Pathset
{A}	{B}	A=>B[0]
{B}	{A}	B=>A[0]

X = X's IP address & Port
{X} = X's MPTCP Hash/Key



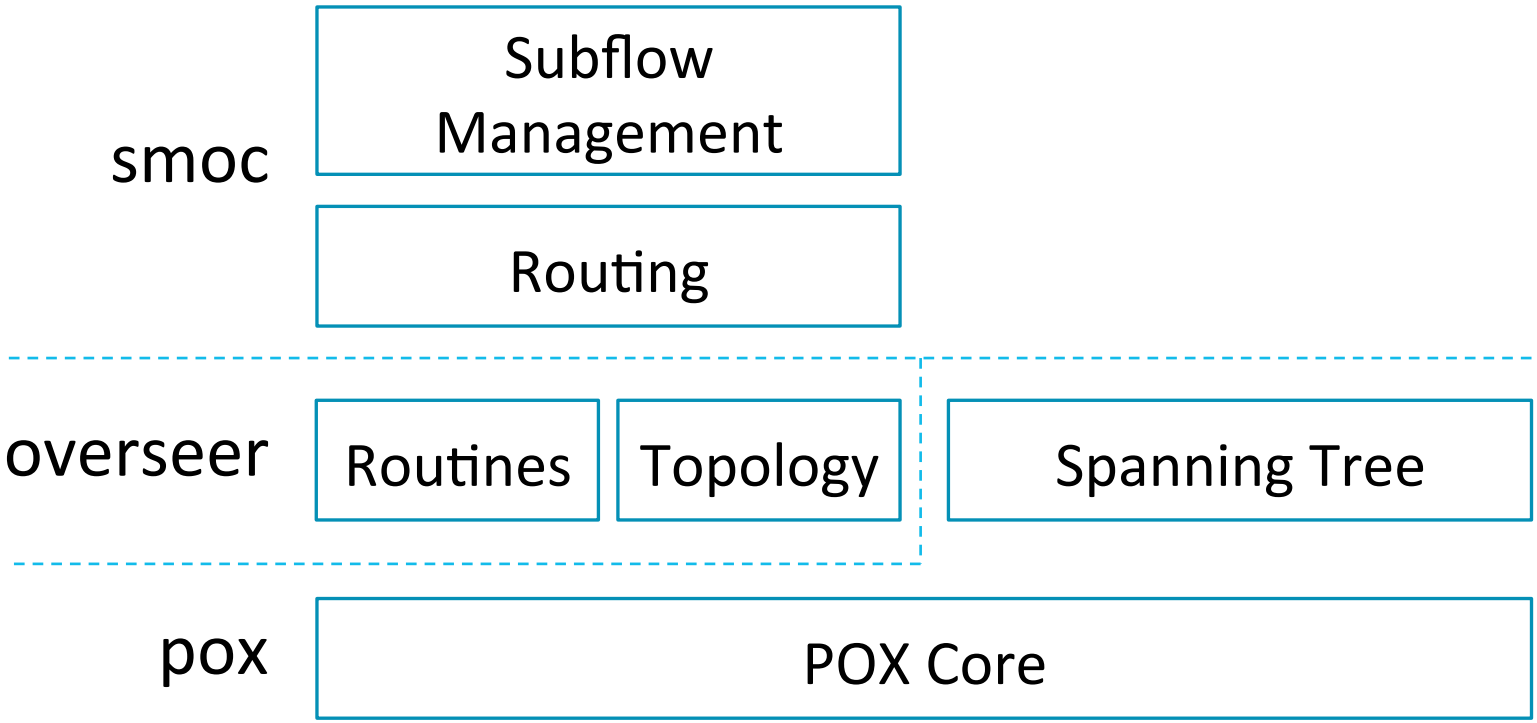
Host A

Host B

Routing

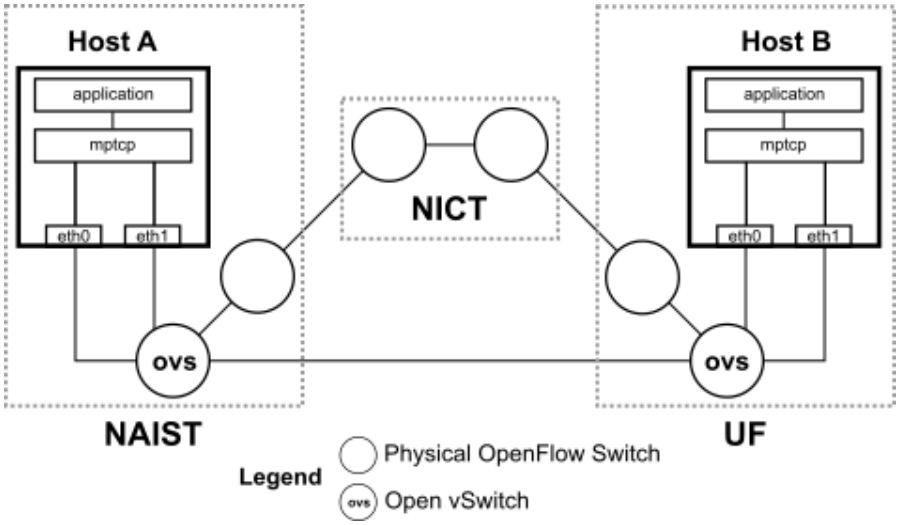
Flow Management

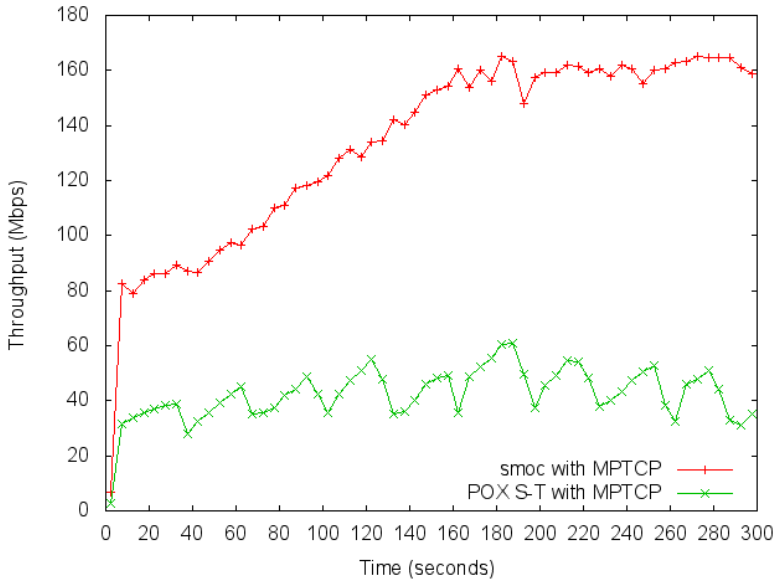
smoc:
Simple Multipath OpenFlow Controller



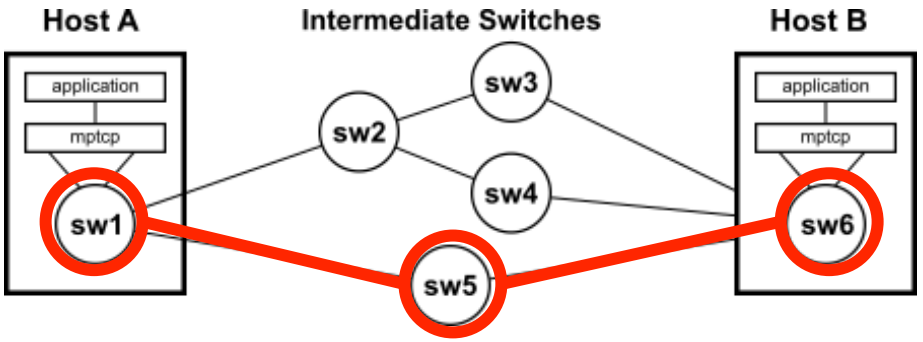
※ networkx is used for additional path calculation

Experiment in PRAGMA Experimental Network Testbed



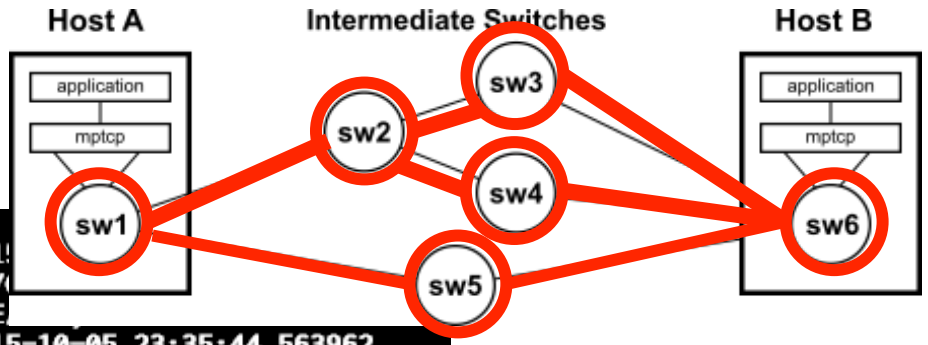


smoc w/o MPTCP



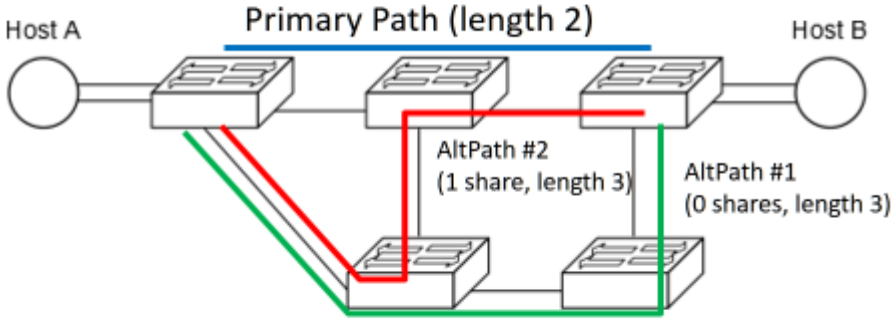

```
[mupoxstats             ] Pinging Switches
[mupoxstats             ] 46-8c-03-80-da-49 @ 2015-10-05 23:33:04.385336
[mupoxstats             ]   35: RX 816330 B/0 E/0 D; TX 103564262 B/0 E/0 D
[mupoxstats             ]   48: RX 103564220 B/0 E/0 D; TX 816416 B/0 E/0 D
[mupoxstats             ] fe-6a-e0-c6-e8-4b @ 2015-10-05 23:33:04.390581
[mupoxstats             ] 2e-69-92-64-a6-42 @ 2015-10-05 23:33:04.392020
[mupoxstats             ]   49: RX 102965694 B/0 E/0 D; TX 816352 B/0 E/0 D
[mupoxstats             ]   52: RX 816416 B/0 E/0 D; TX 102965758 B/0 E/0 D
[mupoxstats             ] fa-23-7e-4f-3a-4f @ 2015-10-05 23:33:04.395016
[mupoxstats             ]   38: RX 816416 B/0 E/0 D; TX 103331834 B/0 E/0 D
[mupoxstats             ]   37: RX 103331770 B/0 E/0 D; TX 816416 B/0 E/0 D
[mupoxstats             ] fe-09-79-cb-63-42 @ 2015-10-05 23:33:04.397678
[mupoxstats             ] 36-1a-38-0a-7d-48 @ 2015-10-05 23:33:04.399173
```

smoc with MPTCP




```
[mupoxstats              ] Pinging Switches
[mupoxstats              ] fe-6a-e0-c6-e8-4b @ 201
[mupoxstats              ]   52: RX 1452230 B/0 E/
[mupoxstats              ]   51: RX 65893132 B/0 E
[mupoxstats              ] fa-23-7e-4f-3a-4f @ 2015-10-05 23:35:44.563962
[mupoxstats              ]   38: RX 2836030 B/0 E/0 D; TX 125046848 B/0 E/0 D
[mupoxstats              ]   37: RX 125046848 B/0 E/0 D; TX 2835966 B/0 E/0 D
[mupoxstats              ] 46-8c-03-80-da-49 @ 2015-10-05 23:35:44.564975
[mupoxstats              ]   47: RX 59146400 B/0 E/0 D; TX 1321990 B/0 E/0 D
[mupoxstats              ]   46: RX 65888768 B/0 E/0 D; TX 1452156 B/0 E/0 D
[mupoxstats              ]   35: RX 5609792 B/0 E/0 D; TX 250078808 B/0 E/0 D
[mupoxstats              ]   48: RX 125048324 B/0 E/0 D; TX 2836030 B/0 E/0 D
[mupoxstats              ] fe-09-79-cb-63-42 @ 2015-10-05 23:35:44.567417
[mupoxstats              ]   38: RX 1321916 B/0 E/0 D; TX 59144924 B/0 E/0 D
[mupoxstats              ]   37: RX 59144860 B/0 E/0 D; TX 1321916 B/0 E/0 D
[mupoxstats              ] 36-1a-38-0a-7d-48 @ 2015-10-05 23:35:44.569871
[mupoxstats              ]   62: RX 1322064 B/0 E/0 D; TX 59141908 B/0 E/0 D
[mupoxstats              ]   60: RX 125036516 B/0 E/0 D; TX 2774102 B/0 E/0 D
[mupoxstats              ]   61: RX 1452230 B/0 E/0 D; TX 65894608 B/0 E/0 D
[mupoxstats              ] 2e-69-92-64-a6-42 @ 2015-10-05 23:35:44.573156
[mupoxstats              ]   51: RX 2773954 B/0 E/0 D; TX 125094080 B/0 E/0 D
[mupoxstats              ]   49: RX 250202664 B/0 E/0 D; TX 5609994 B/0 E/0 D
[mupoxstats              ]   52: RX 2835966 B/0 E/0 D; TX 125104412 B/0 E/0 D
```

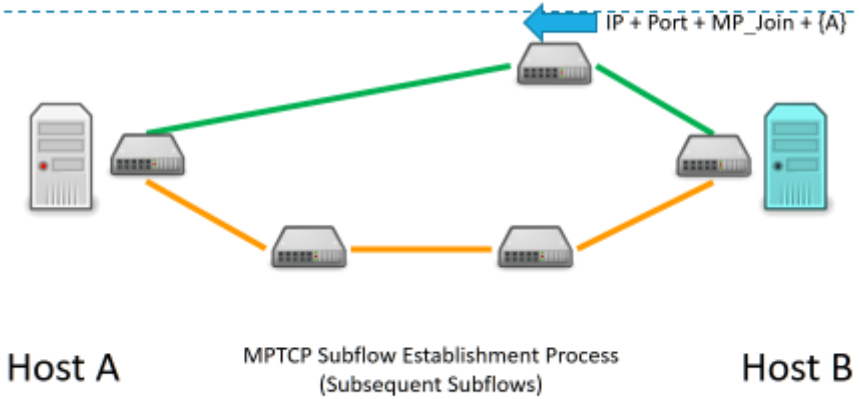
Routing Mechanism is based on shortest path



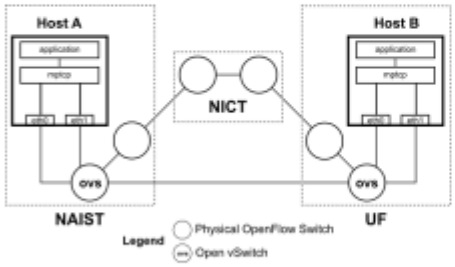
Sample routing results from our algorithm

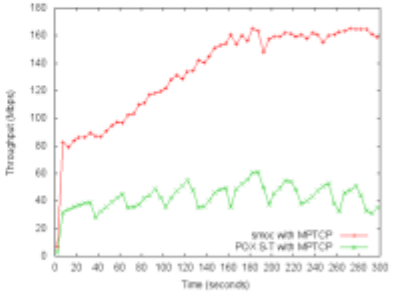
pending_capable				pending_join				established		
From	To	key	pathset	From	To	key	pathset	From	To	Pathset
				A	B	{B}	A=>B[1]	{A}	{B}	A=>B[0]
				B	A	{A}	B=>A[1]	{B}	{A}	B=>A[0]

X = X's IP address & Port
{X} = X's MPTCP Hash/Key



Experiment in PRAGMA Experimental Network Testbed





Canned Results: smoc with MPTCP