

Experimenting TCP Performance with FABRIC

Ewa Deelman¹, Anirban Mandal², Prasanna Balaprakash⁵, Mariam Kiran⁵,
Krishnan Raghavan³, Hongwei Jin³, Cong Wang², Komal Thareja²,
Imtiaz Mahmud⁴, George Papadimitriou¹

¹University of Southern California, ²Renaissance Computing Institute,

³Argonne National Laboratory, ⁴Lawrence Berkeley National Laboratory, ⁵Oak Ridge National Laboratory

Meet the Team



Ewa Deelman
USC (Lead PI)



Anirban Mandal
RENCI (Co-PI)



Prasanna Balaprakash
ORNL(Co-PI)



Mariam Kiran
ORNL(Co-PI)



George Papadimitriou
USC



Cong Wang
RENCI



Krishnan Raghavan
ANL



Imtiaz Mahmud
LBNL

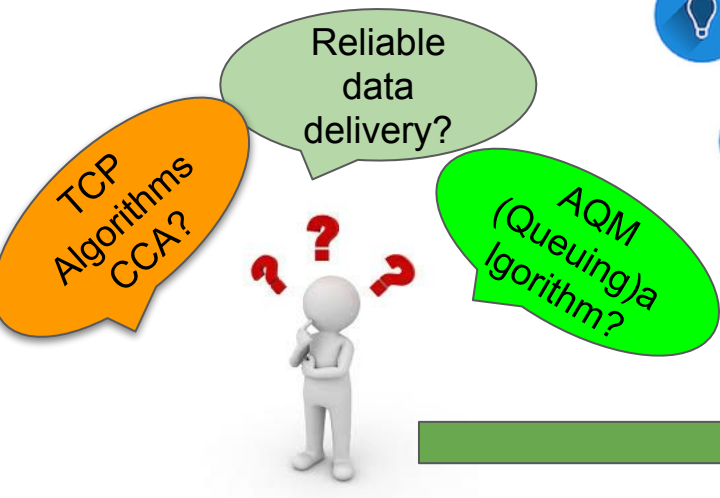


Komal Thareja
RENCI



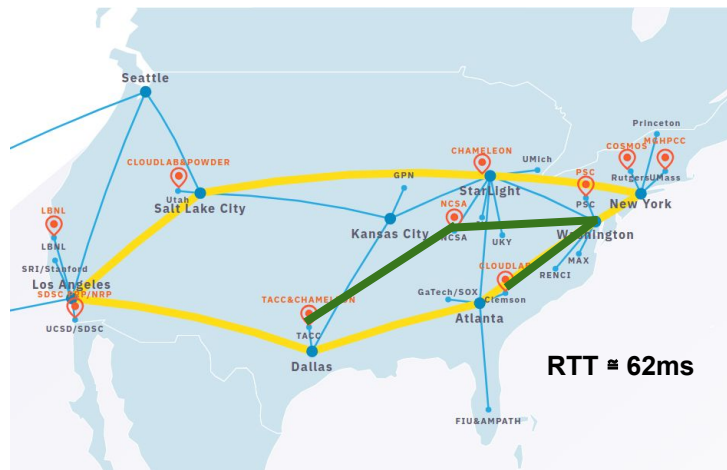
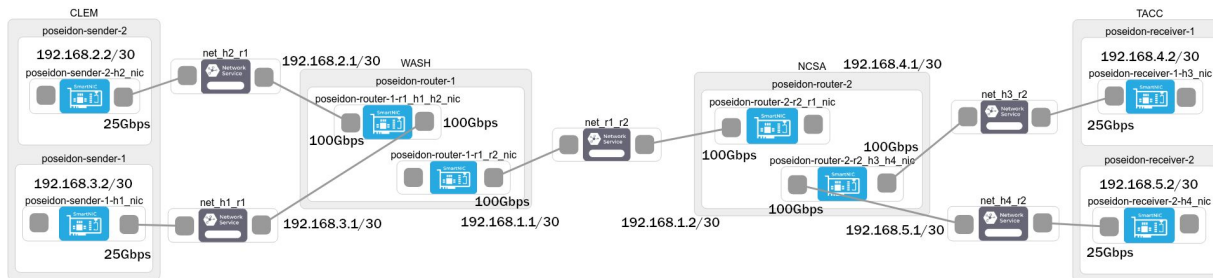
Hongwei Jin
ANL

What are we doing? Improving Network Performance for our data transfers

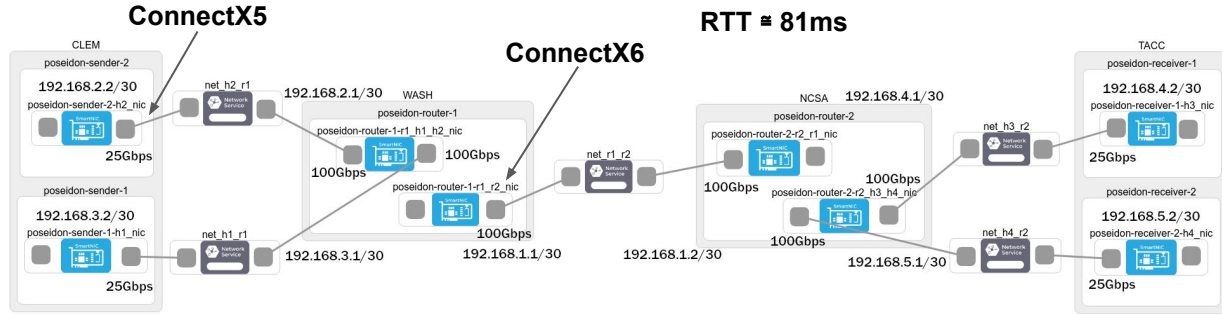


- Novel Innovations for Community:
- Interplay between Transfer and queuing in routers
 - Build dataset for transfer protocols behavior
 - Lead to a NEW transfer methods for **“better and fair Internet”**

Experimental Setup on FABRIC



Experimental Setup on FABRIC



Scenarios

CCA 1 - CCA 2	AQM	Queue Length	Bottleneck BW
BBRv1 - CUBIC	FIFO	0.5 x BDP	100 Mbps
BBRv2 - CUBIC			
BBRv3 - CUBIC			
HTCP - CUBIC	FQ CODEL	1 x BDP	500 Mbps
Reno - CUBIC			
CUBIC - CUBIC			
BBRv1 - BBRv1	RED	2 x BDP	1 Gbps
BBRv2 - BBRv2			
BBRv3 - BBRv3			
HTCP - HTCP	RED	8 x BDP	25 Gbps
Reno - Reno			

Iperf3 Configuration

Bottleneck BW	Total #Flows	iperf3 Configuration
100 Mbps	2	1 iperf3 process/node 1 stream
500 Mbps	10	5 iperf3 processes/node 1 stream each
1 Gbps	20	10 iperf3 processes/node 1 stream each
10 Gbps	200	10 iperf3 processes/node 10 parallel streams each
25 Gbps	500	25 iperf3 processes/node 10 parallel streams each

Experimental Setup on FABRIC

Pseudocode

```
foreach agm_type:
  foreach cca_config:
    foreach speed_config:
      Calculate bdp
      Calculate buffer_size list based on BDP
      foreach buffer_size:
        Apply agm and buffer_size
        for 1..5:
          for 1..num_of_procs:
            Start iperf3 servers
            Start iperf3 clients for 200 seconds
          Wait for 210 seconds
          Kill iperf3 servers
```

```
mtu_setting = 8900
rtt_setting = 0.062 #change this based on the ping reponse from earlier cells

speed_config = [
  {"tag": "100mbps", "speed": 100000000, "processes": 1, "parallel_streams": 1},
  {"tag": "500mbps", "speed": 500000000, "processes": 5, "parallel_streams": 1},
  {"tag": "1gbps", "speed": 1000000000, "processes": 10, "parallel_streams": 1},
  {"tag": "10gbps", "speed": 10000000000, "processes": 20, "parallel_streams": 5},
  {"tag": "25gbps", "speed": 25000000000, "processes": 25, "parallel_streams": 10}
]

cca_config = [
  {"cca1": "bbr", "cca2": "cubic"},
  {"cca1": "bbr2", "cca2": "cubic"},
  {"cca1": "htcp", "cca2": "cubic"},
  {"cca1": "reno", "cca2": "cubic"},
  {"cca1": "cubic", "cca2": "cubic"},
  {"cca1": "bbr", "cca2": "bbr"},
  {"cca1": "bbr2", "cca2": "bbr2"},
  {"cca1": "htcp", "cca2": "htcp"},
  {"cca1": "reno", "cca2": "reno"}
]
```



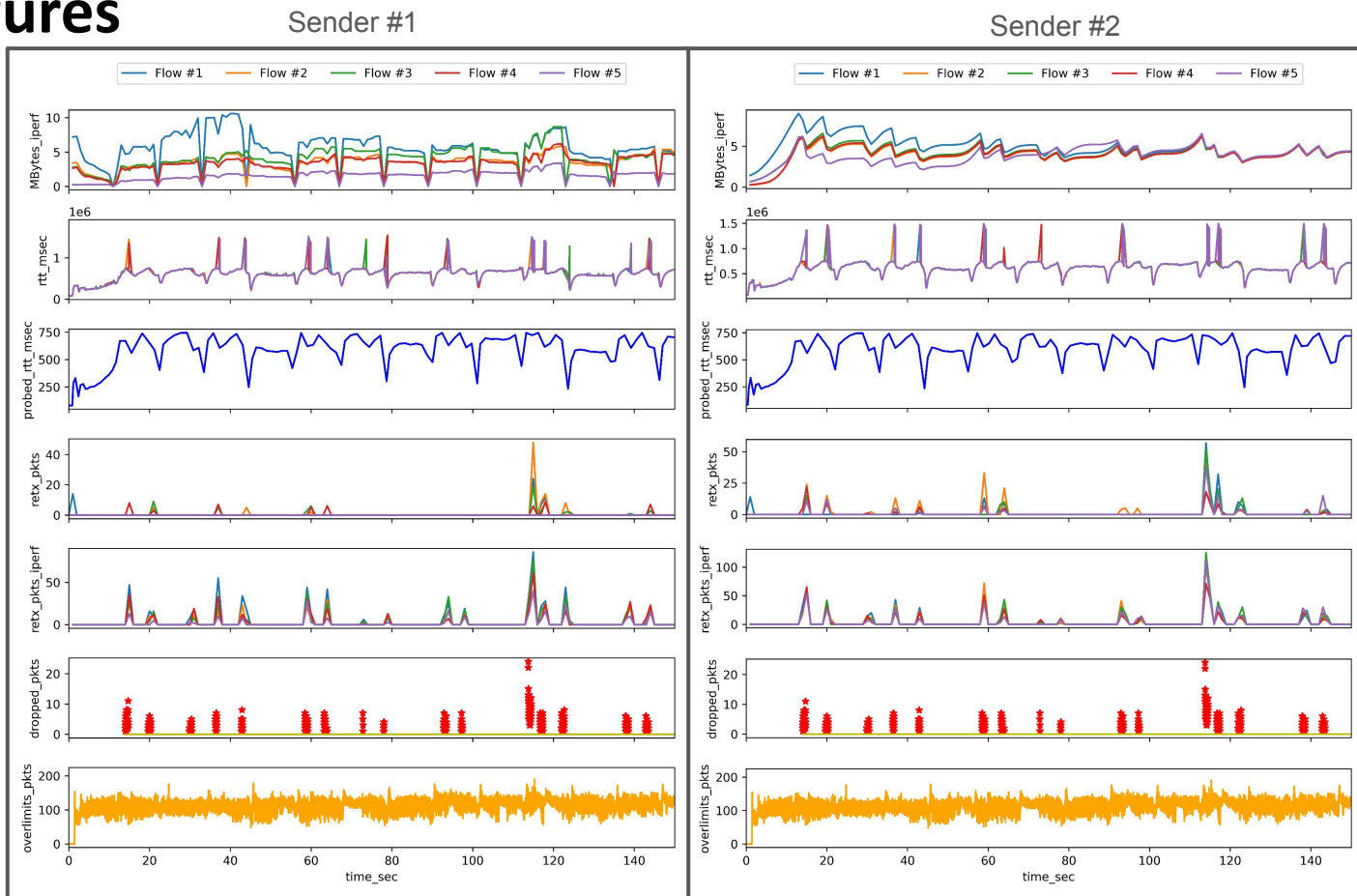
$$BDP \text{ Formula } \quad BDP = \frac{BW_{\text{bottleneck}} * RTT}{8} \text{ bytes}$$

```
iPerf3 Server CMD: iperf3 -s -p {port} -f m
iPerf3 Client CMD: iperf3 -c {server_ip} -p {port_num} -C {cca} -t 200 -f m -P {flows} -M 8900
```

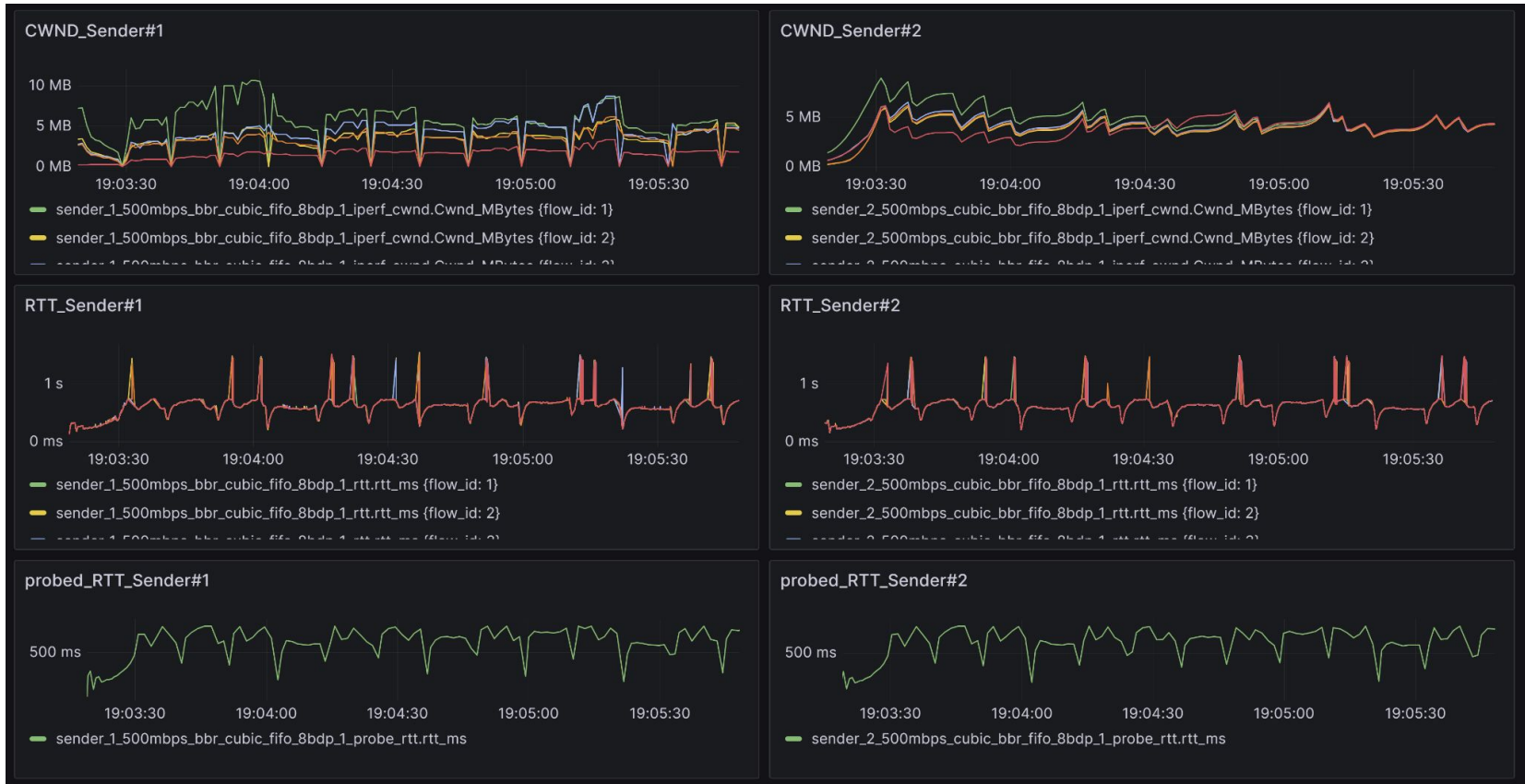
```
tc qdisc add dev {iface} root handle 1: tbf rate {speed} burst 1570000 limit {applied_bdp}
tc qdisc add dev {iface} parent 1:1 handle 10: fq_codel limit {packets} FQ_CoDel
```

Generated Figures

BBRv1 vs CUBIC
8 BDP buffer size
AQM = FIFO
Bandwidth = 500
Mbps



Grafana dashboard



Summary

- Details of Implementation and Code Access
 - Full Fabric implementation - automated code.
 - Scripts for processing pcap files and extracting necessary data.
 - A script to process iperf data and compile information for each flow across experiments.
 - Tools to create both PNG and interactive HTML visualizations for the collected data.
- Access to the gathered data, figures, and raw files.
- A grafana dashboard for easy data access and visualization.

Acknowledgements



DOE ASCR Award (DE-SC0022328): Integrated Computational and Data Infrastructure (ICDI) Program

