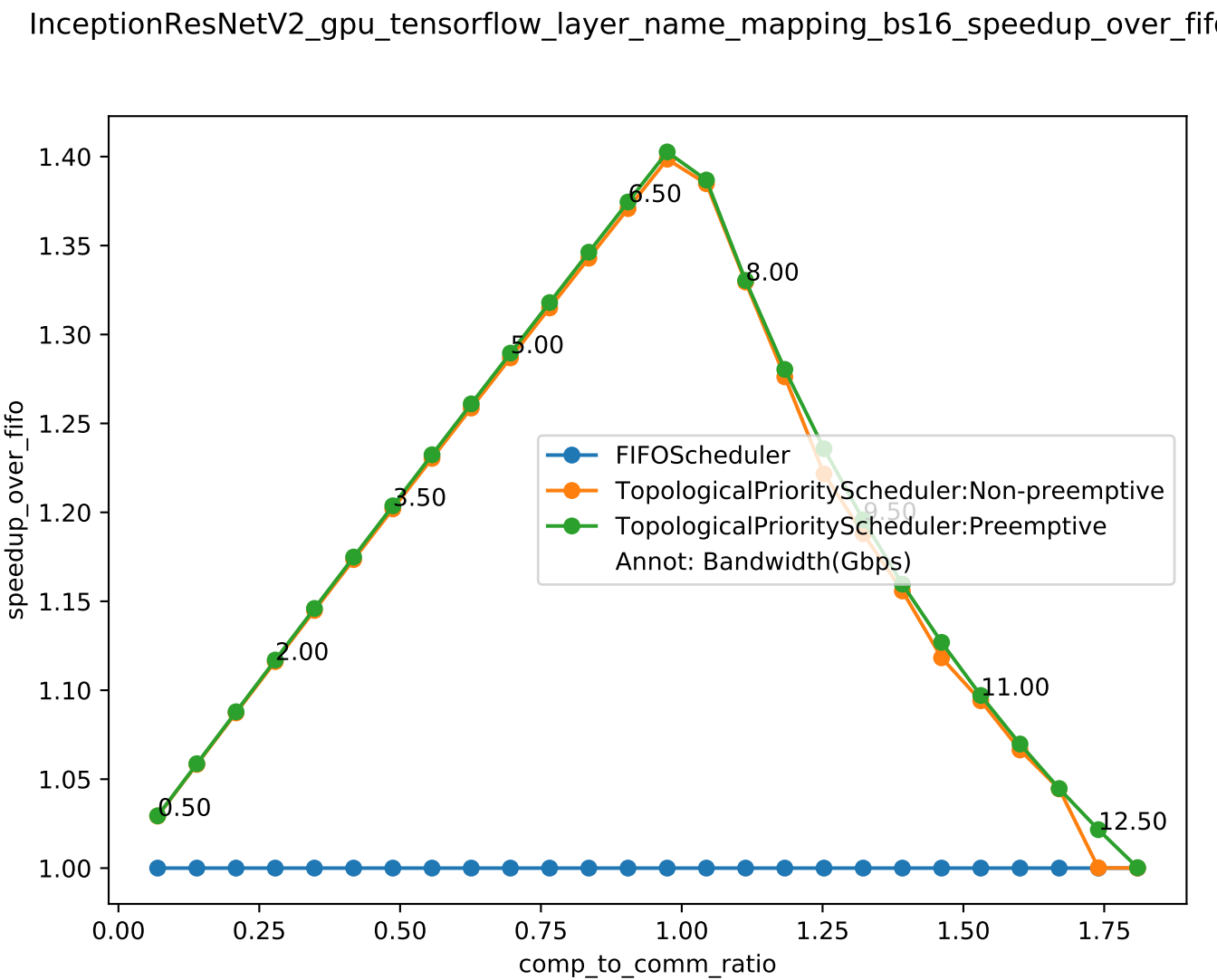
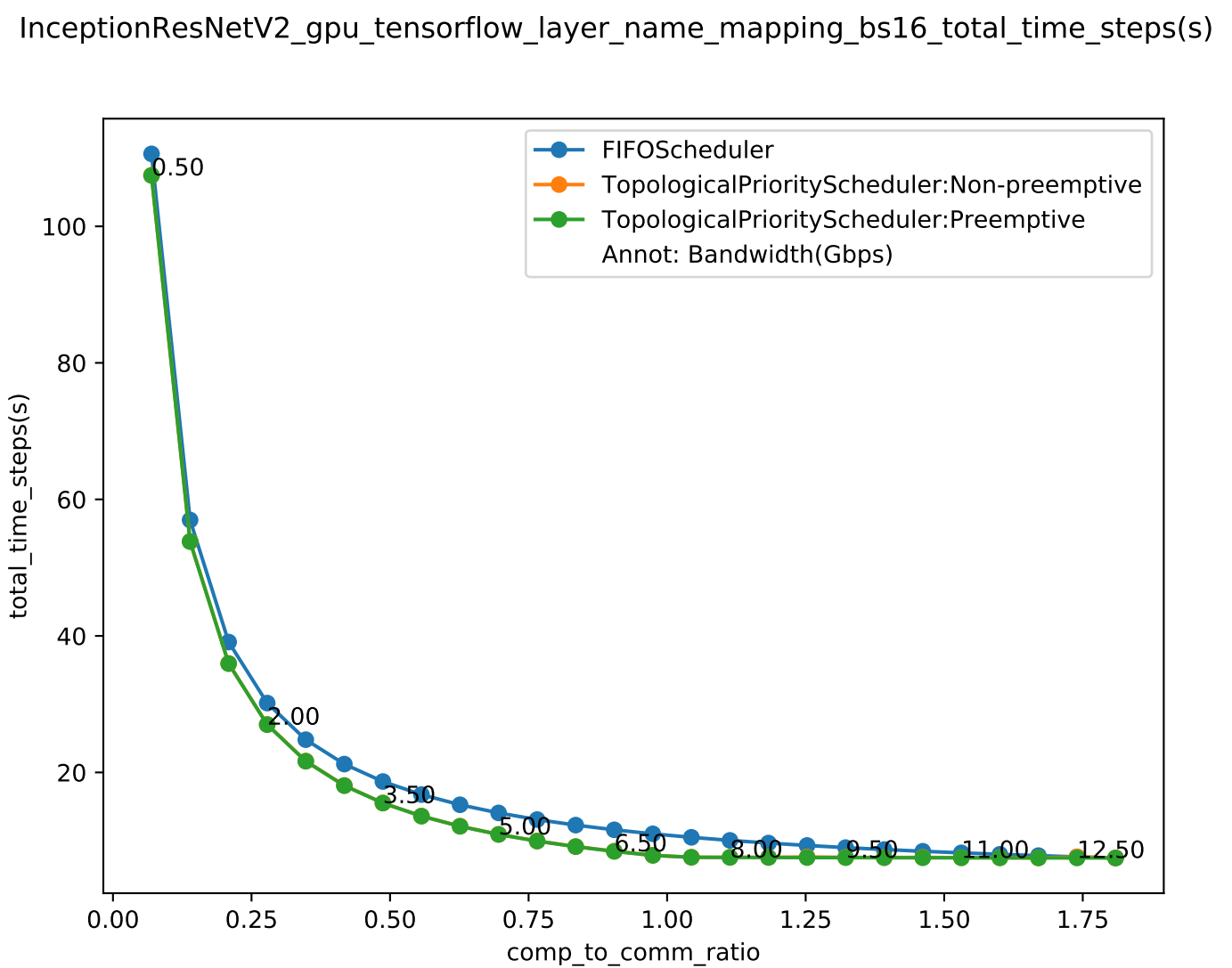
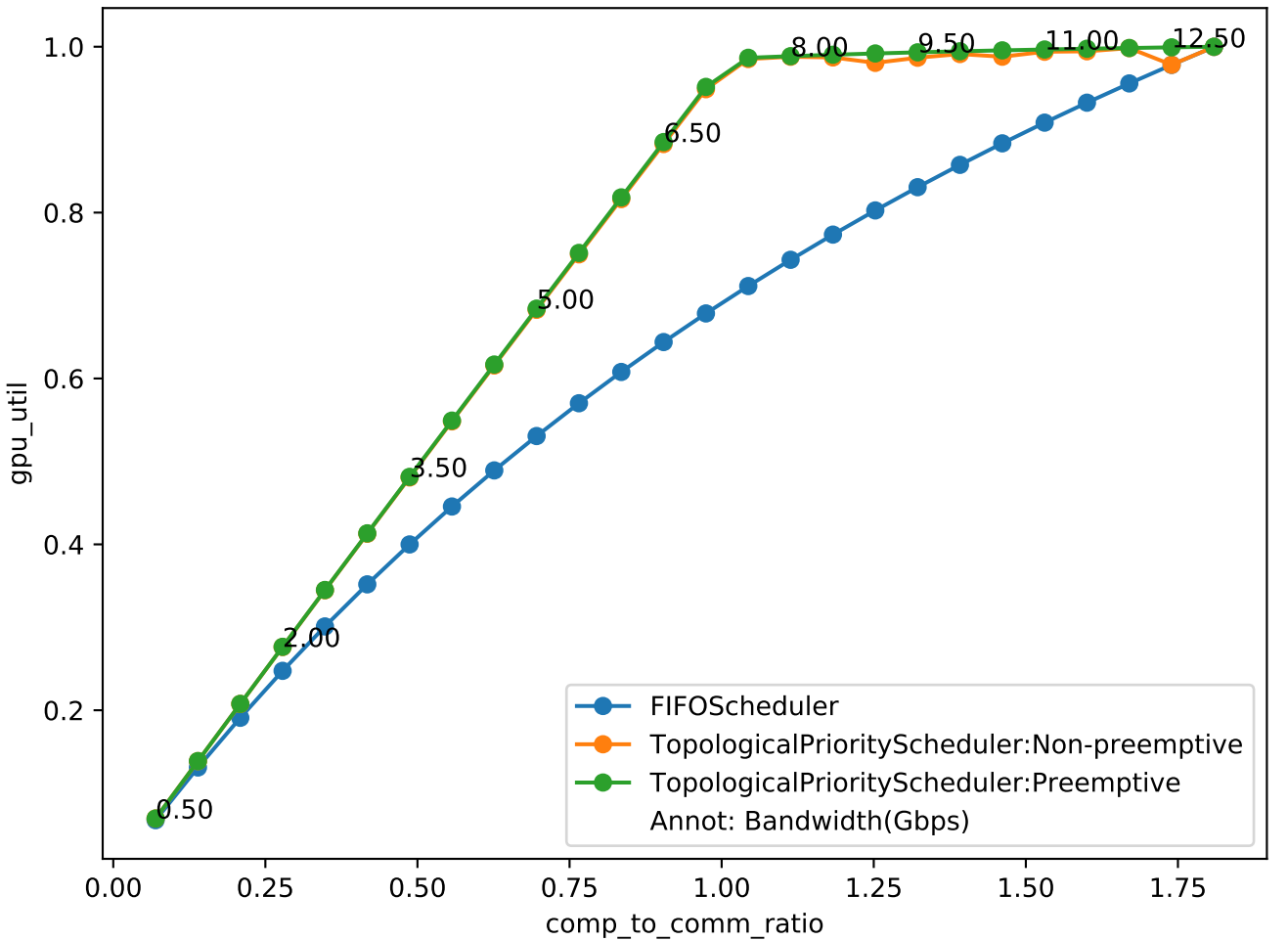
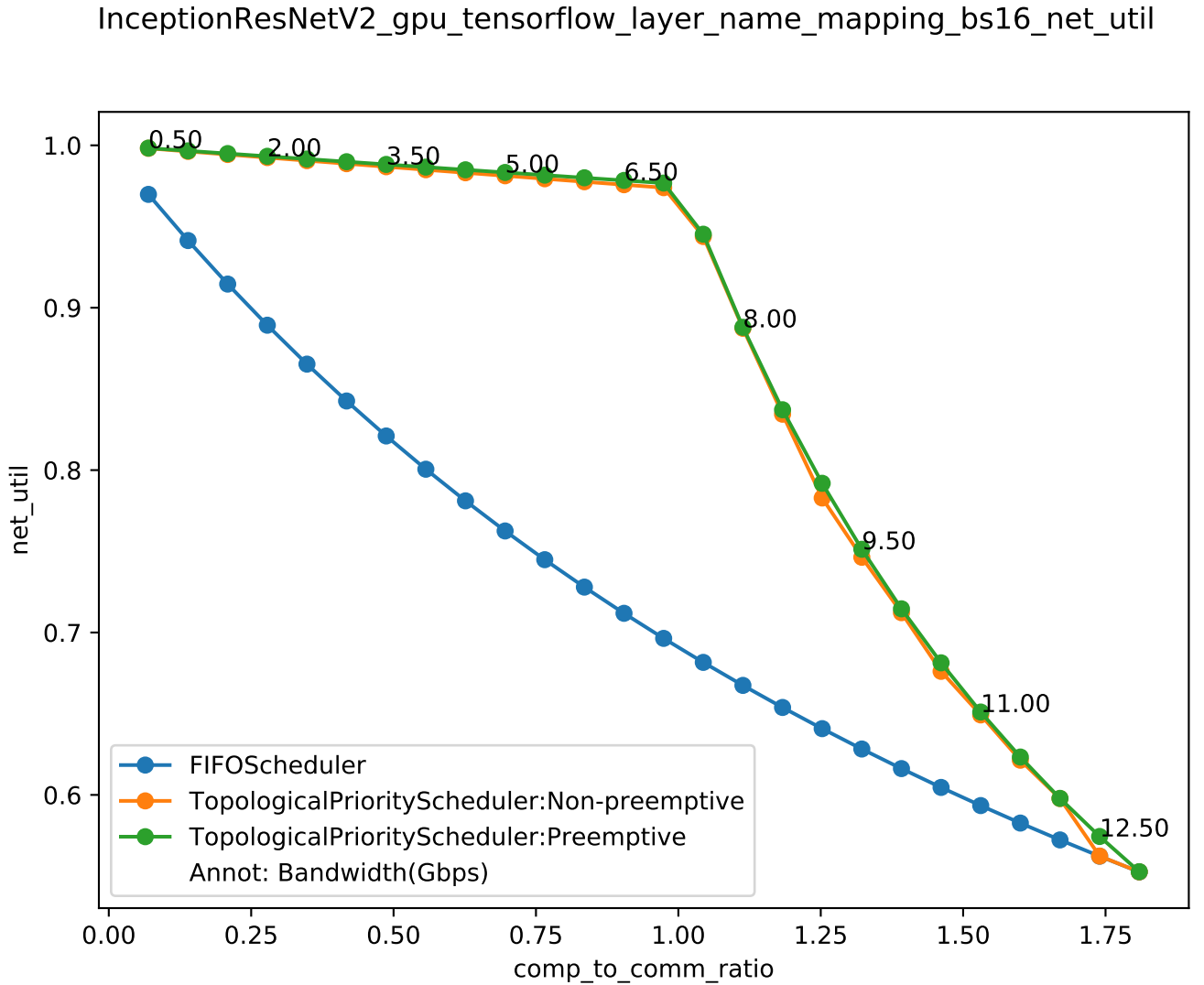
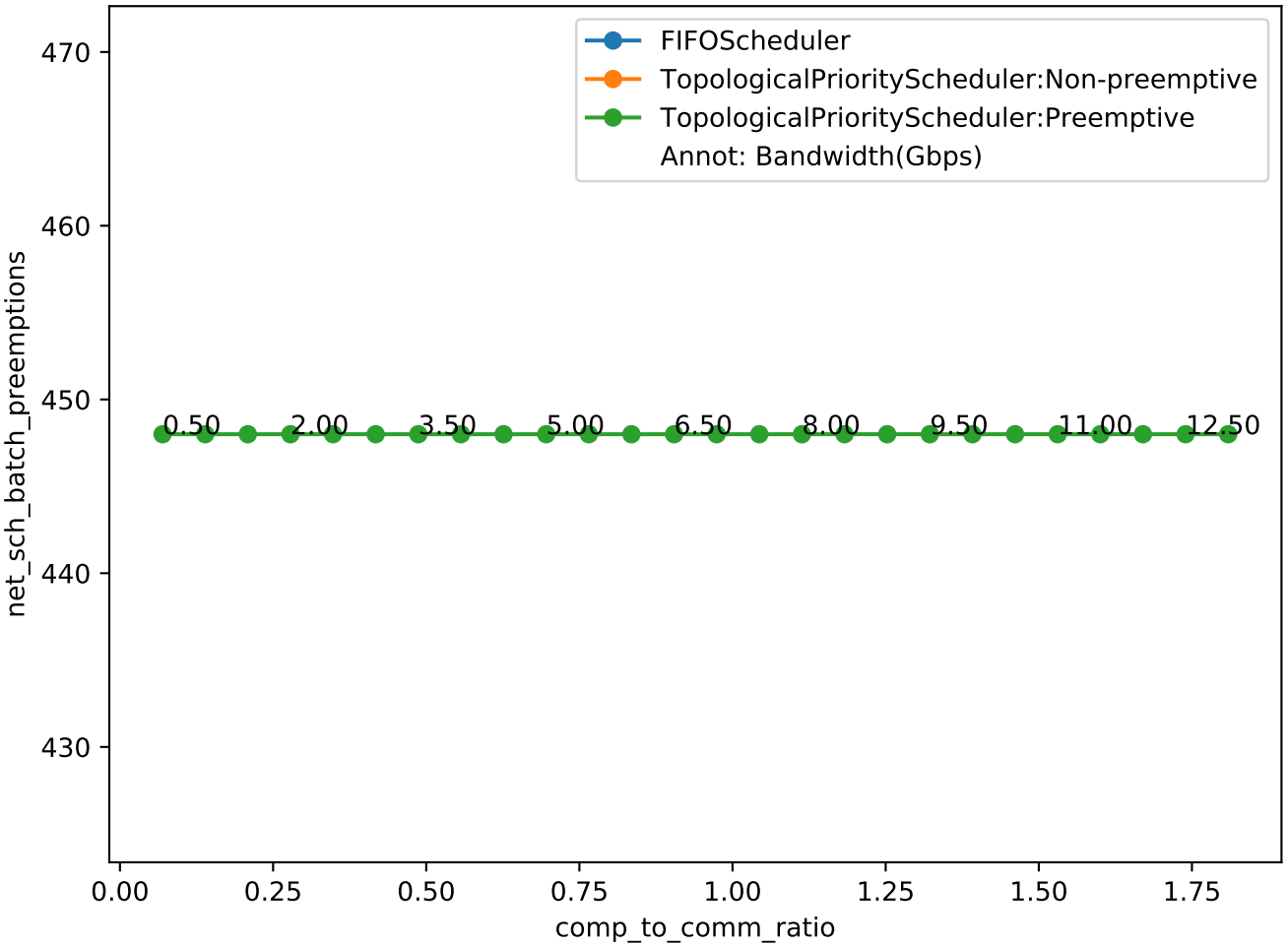

InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_speedup_over_fifo
speedup_over_fifo
comp_to_comm_ratio
1.40
1.35
1.30
1.25
1.20
1.15
1.10
1.05
1.00
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
FIFOScheduler
TopologicalPriorityScheduler:Non-preemptive
TopologicalPriorityScheduler:Preemptive
Annot: Bandwidth(Gbps)
0.50
2.00
3.50
5.00
6.50
8.00
9.50
11.00
12.50


InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_total_time_steps(s)
FIFOScheduler
TopologicalPriorityScheduler:Non-preemptive
TopologicalPriorityScheduler:Preemptive
Annot: Bandwidth(Gbps)
0.50
2.00
3.50
5.00
6.50
8.00
9.50
11.00
12.50
total_time_steps(s)
comp_to_comm_ratio
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
20
40
60
80
100

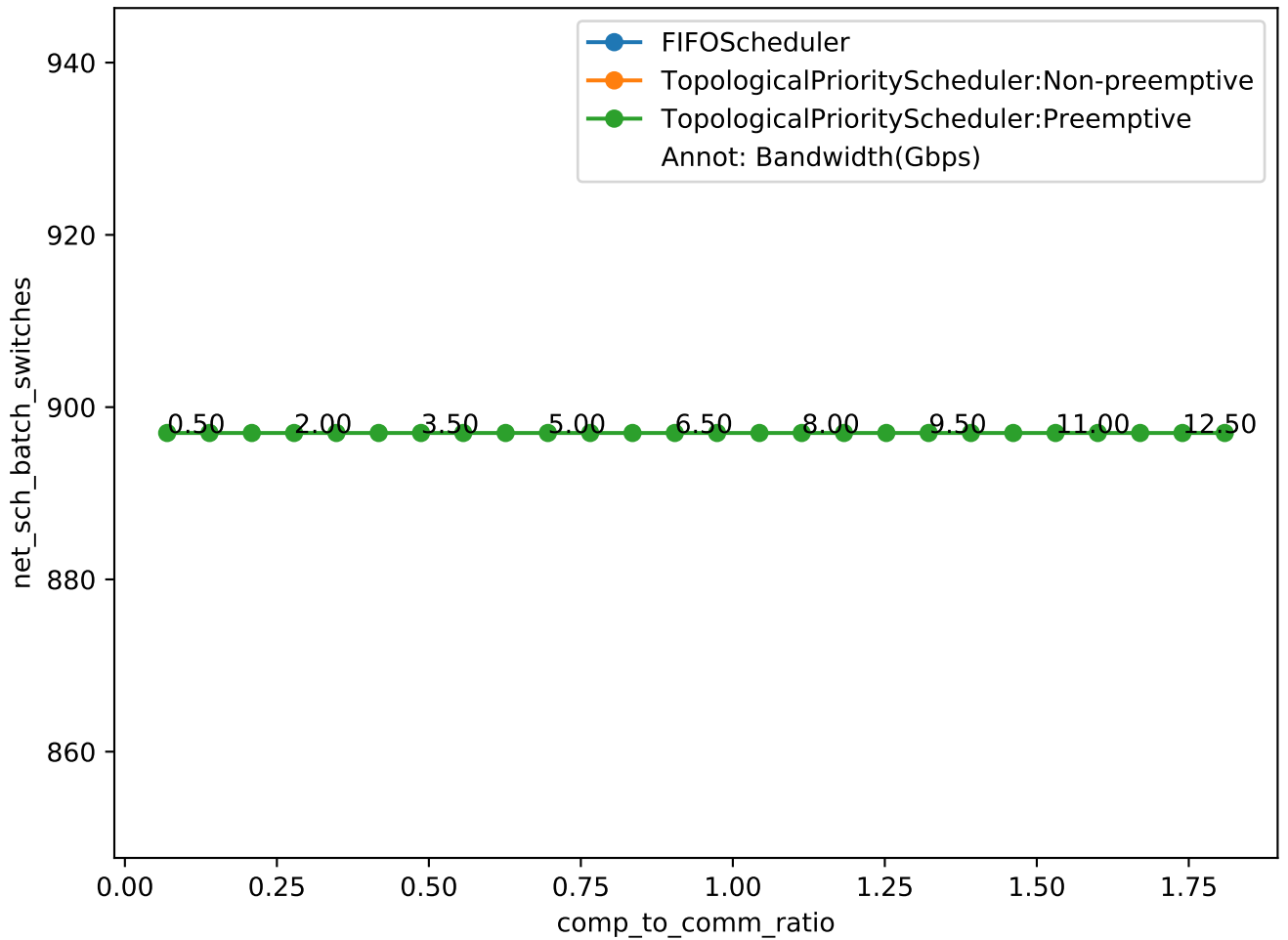

InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_gpu_util
gpu_util
comp_to_comm_ratio
1.0
0.8
0.6
0.4
0.2
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
0.50
2.00
3.50
5.00
6.50
8.00
9.50
11.00
12.50
FIFOScheduler
TopologicalPriorityScheduler:Non-preemptive
TopologicalPriorityScheduler:Preemptive
Annot: Bandwidth(Gbps)

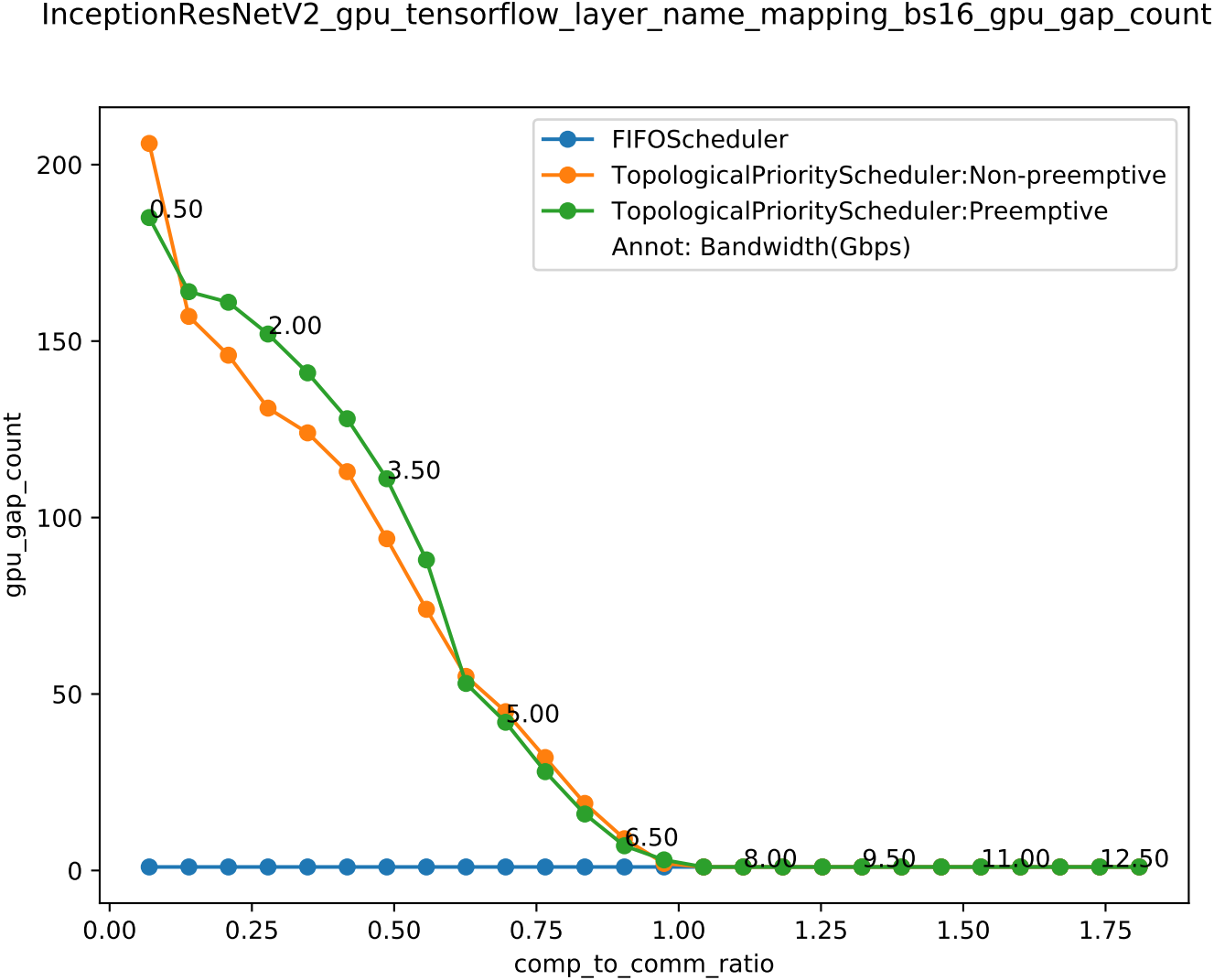

InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_net_util
net_util
comp_to_comm_ratio
1.0
0.9
0.8
0.7
0.6
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
0.50
2.00
3.50
5.00
6.50
8.00
9.50
11.00
12.50
FIFOScheduler
TopologicalPriorityScheduler:Non-preemptive
TopologicalPriorityScheduler:Preemptive
Annot: Bandwidth(Gbps)

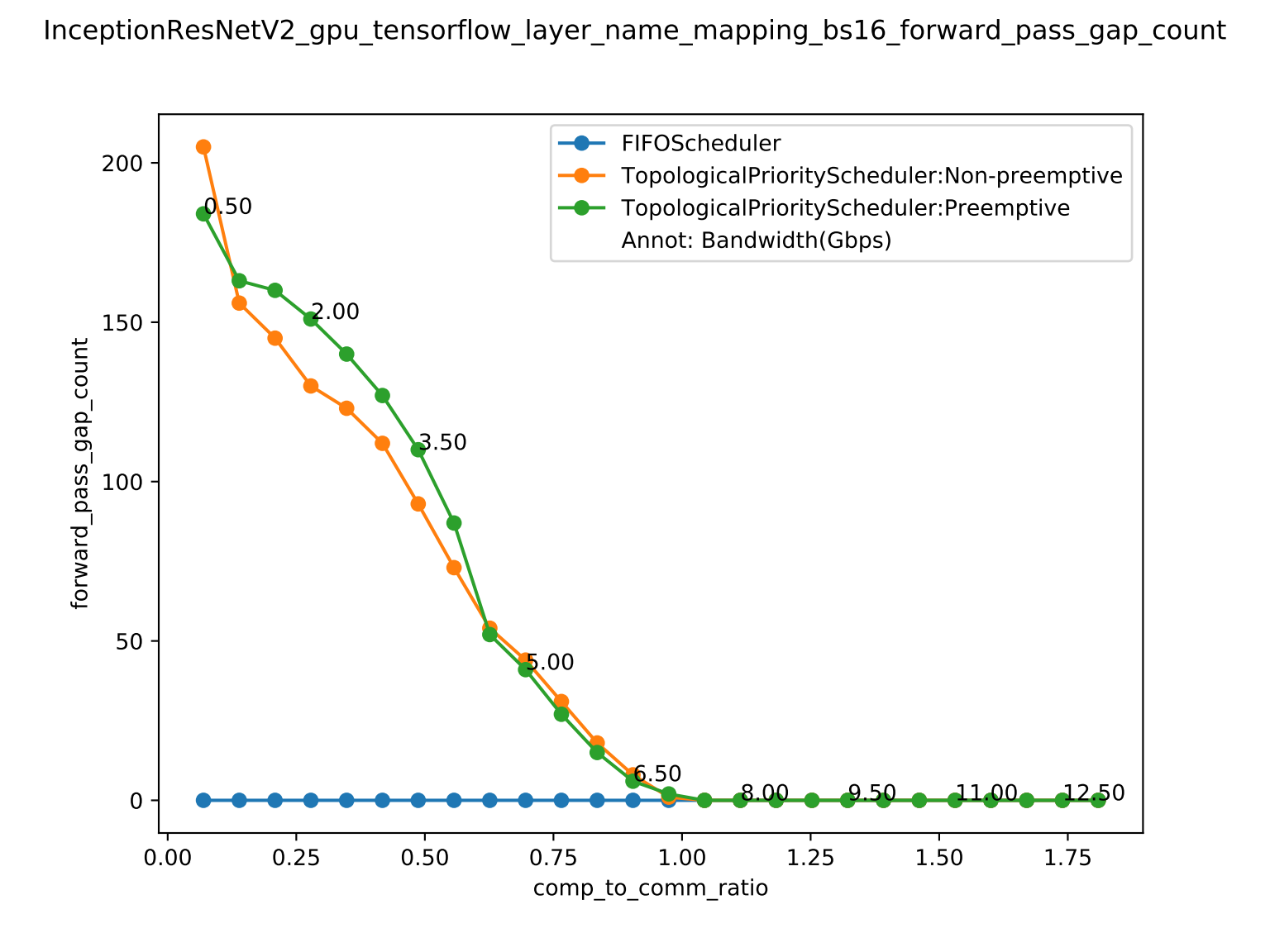
InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_net_sch_batch_preemptions

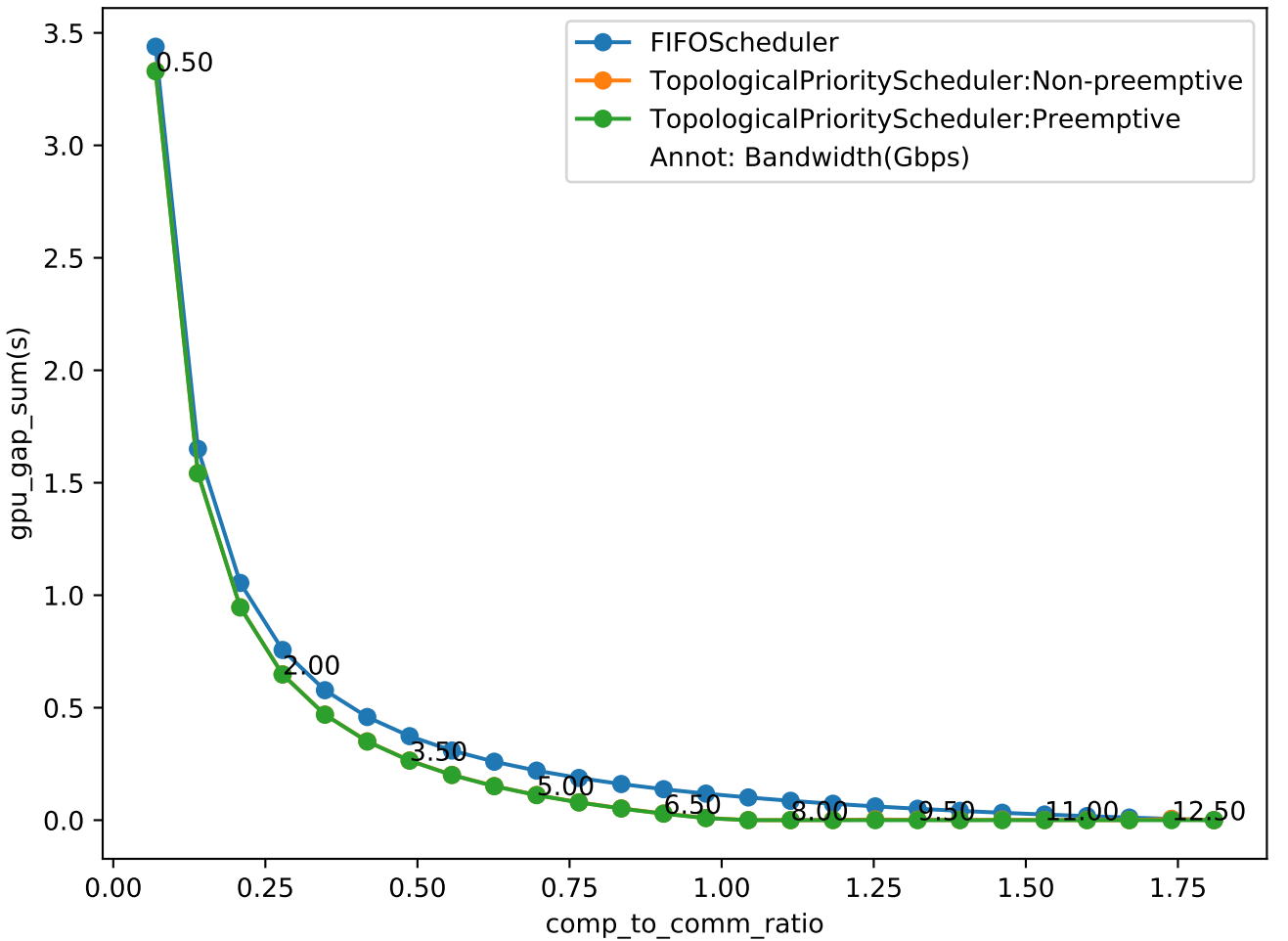


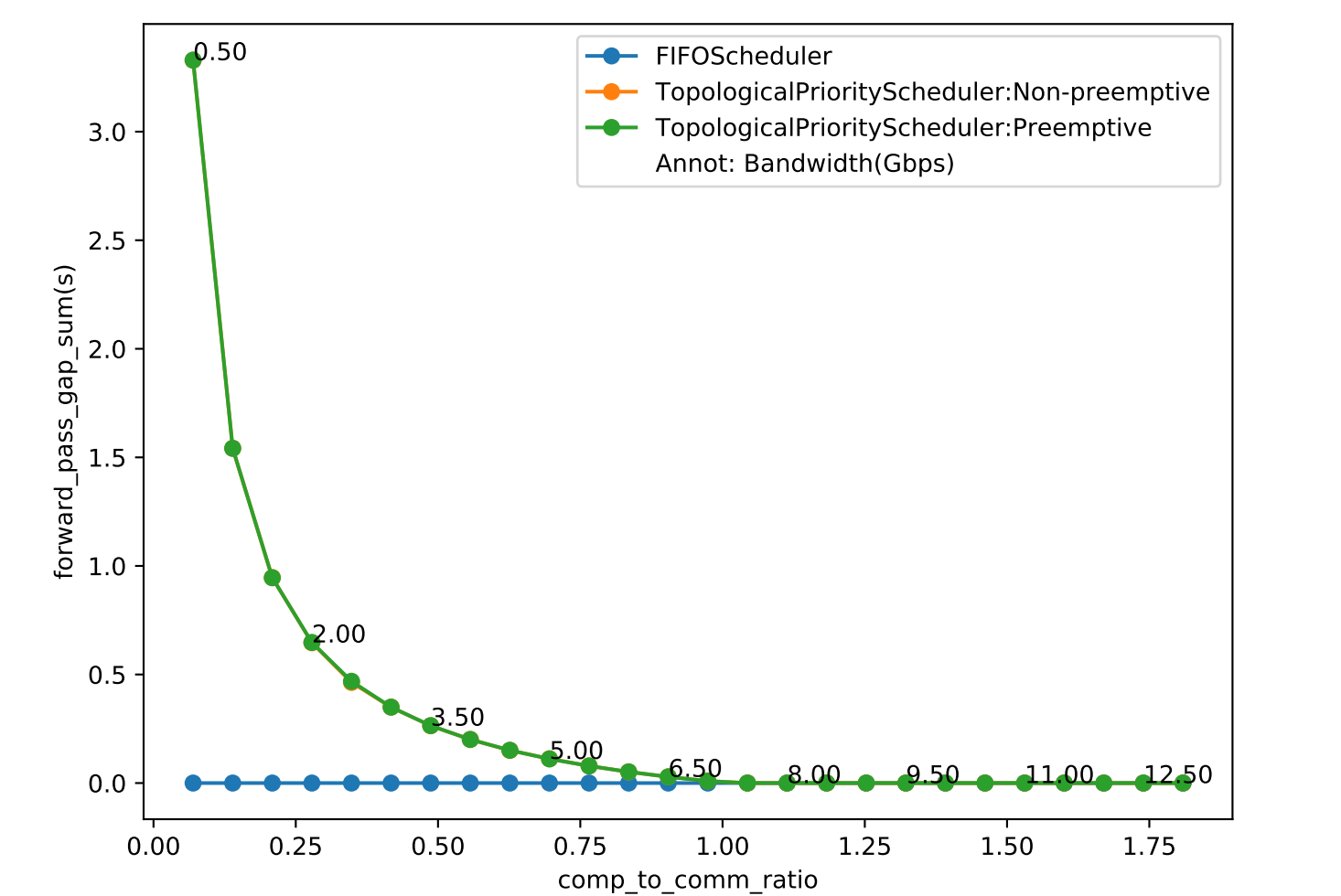
InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_net_sch_batch_switches

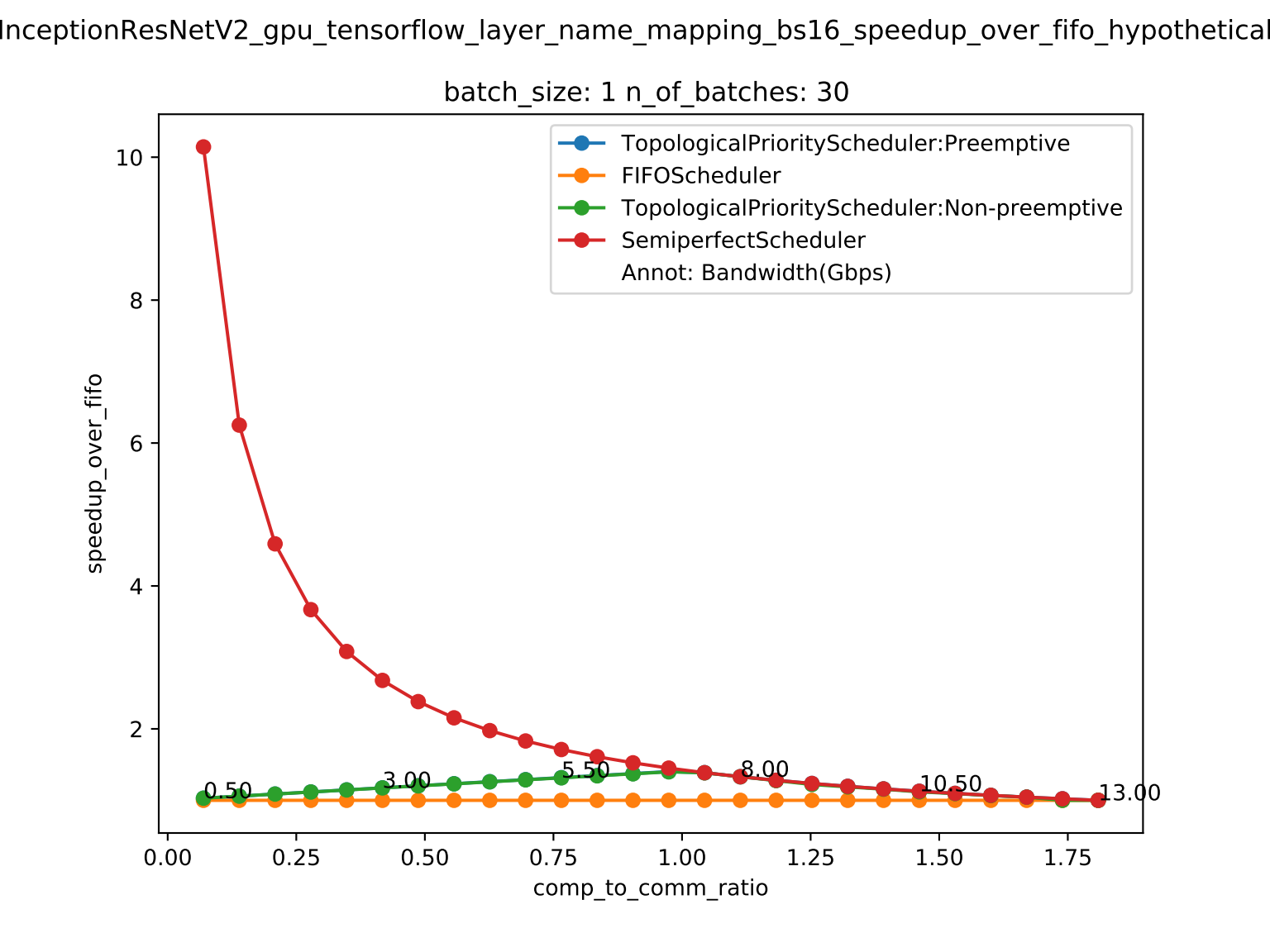


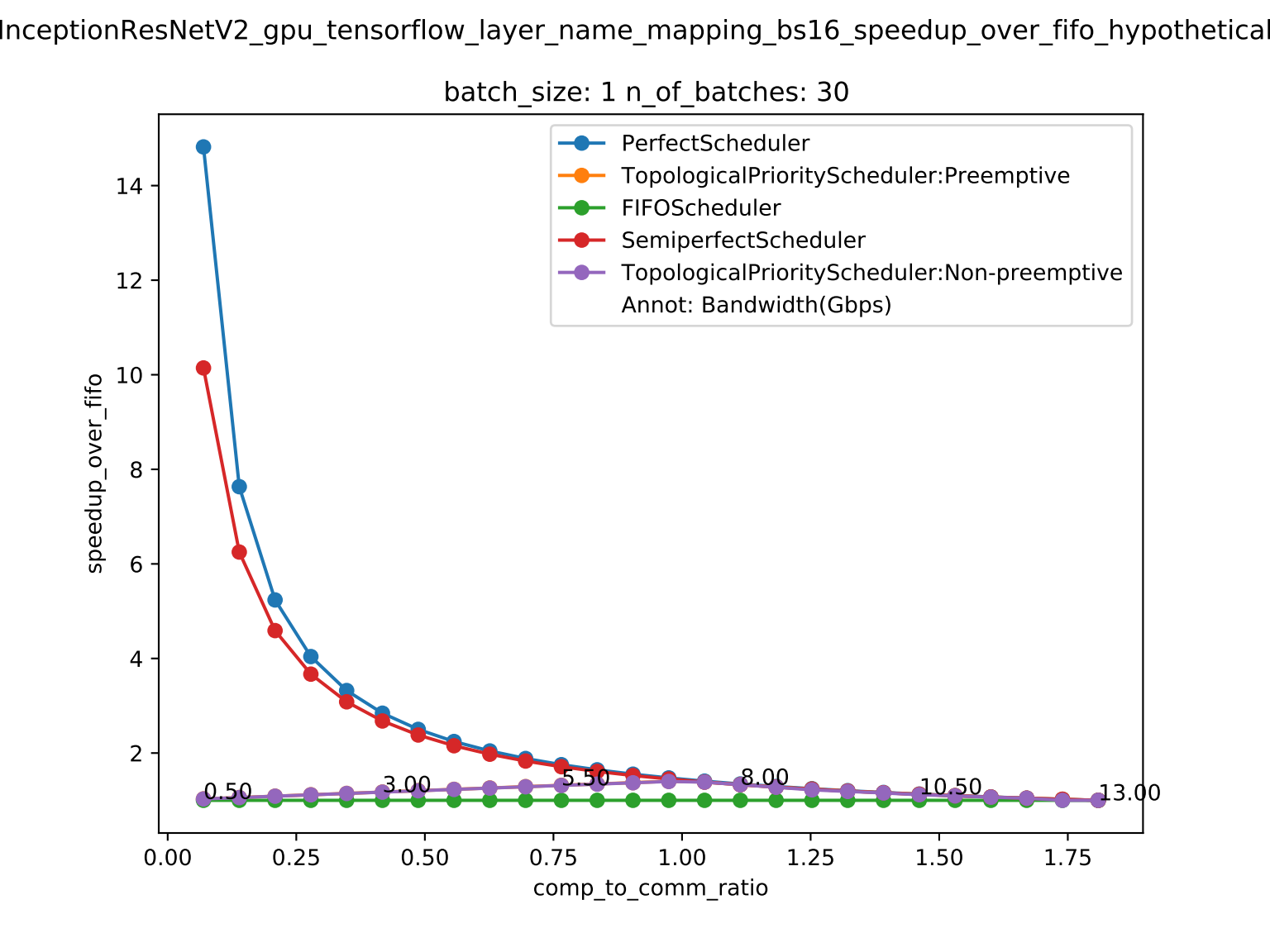

InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_gpu_gap_count
FIFOScheduler
TopologicalPriorityScheduler:Non-preemptive
TopologicalPriorityScheduler:Preemptive
Annot: Bandwidth(Gbps)
0.50
2.00
3.50
5.00
6.50
8.00
9.50
11.00
12.50
gpu_gap_count
comp_to_comm_ratio
0
50
100
150
200
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75


InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_forward_pass_gap_count
FIFOScheduler
TopologicalPriorityScheduler:Non-preemptive
TopologicalPriorityScheduler:Preemptive
Annot: Bandwidth(Gbps)
0.50
2.00
3.50
5.00
6.50
8.00
9.50
11.00
12.50
forward_pass_gap_count
comp_to_comm_ratio
0
50
100
150
200
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75

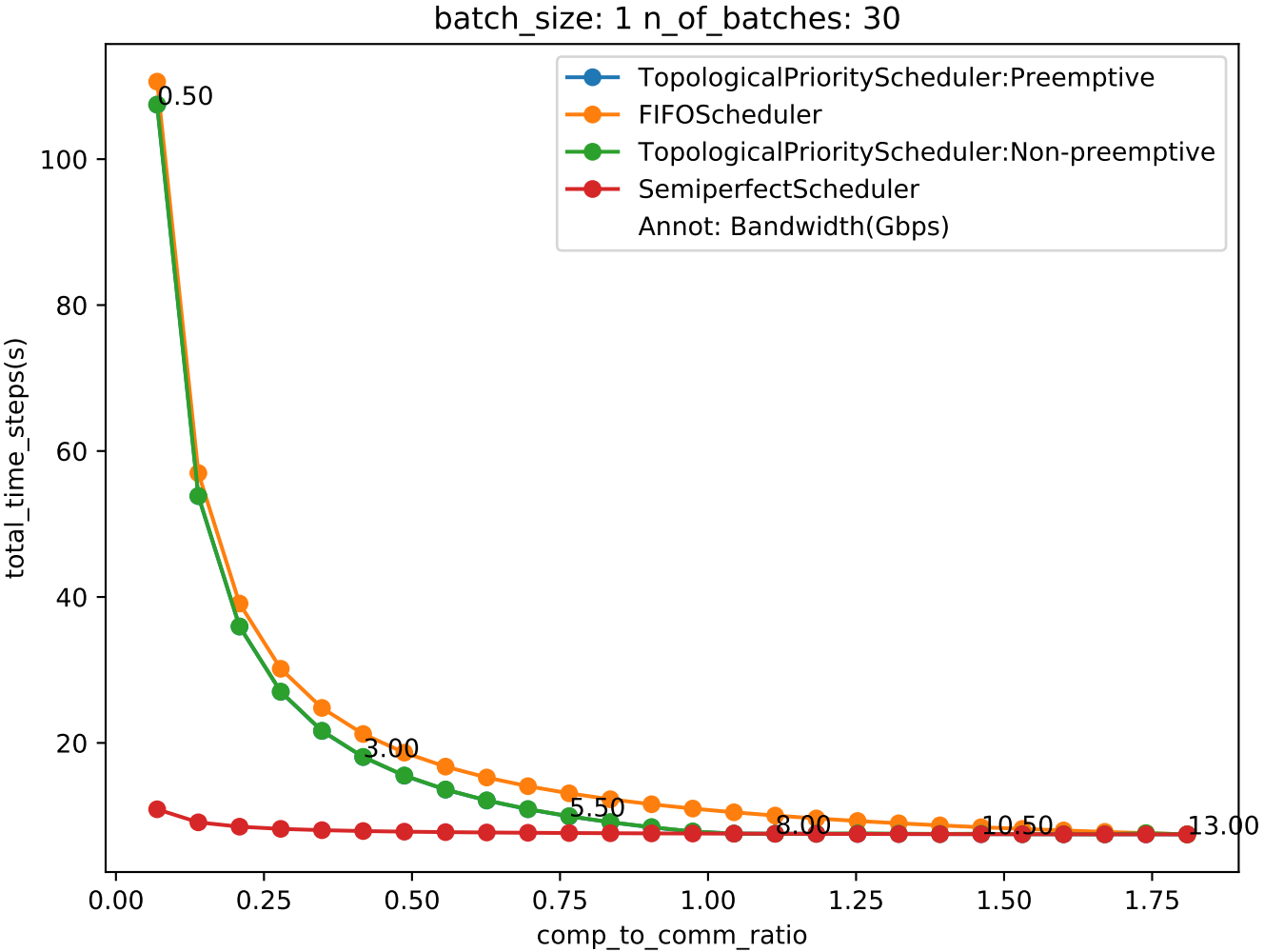

InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_gpu_gap_sum(s)
FIFOScheduler
TopologicalPriorityScheduler:Non-preemptive
TopologicalPriorityScheduler:Preemptive
Annot: Bandwidth(Gbps)
0.50
2.00
3.50
5.00
6.50
8.00
9.50
11.00
12.50
gpu_gap_sum(s)
comp_to_comm_ratio
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75


InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_forward_pass_gap_sum(s)
FIFOScheduler
TopologicalPriorityScheduler:Non-preemptive
TopologicalPriorityScheduler:Preemptive
Annot: Bandwidth(Gbps)
0.50
2.00
3.50
5.00
6.50
8.00
9.50
11.00
12.50
forward_pass_gap_sum(s)
comp_to_comm_ratio
0.0
0.5
1.0
1.5
2.0
2.5
3.0
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75

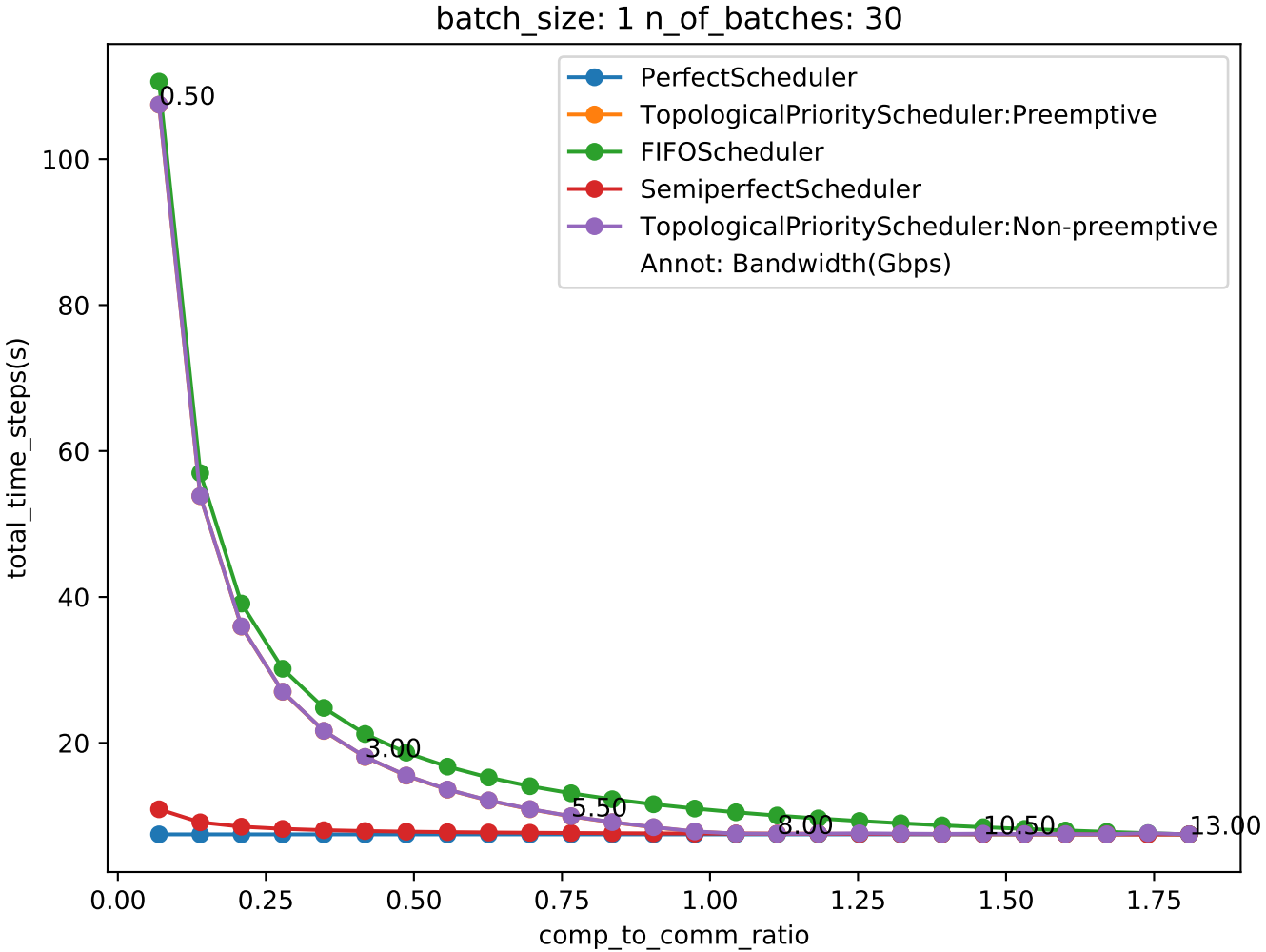

InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_speedup_over_fifo_hypothetical
batch_size: 1 n_of_batches: 30
TopologicalPriorityScheduler:Preemptive
FIFOScheduler
TopologicalPriorityScheduler:Non-preemptive
SemiperfectScheduler
Annot: Bandwidth(Gbps)
speedup_over_fifo
comp_to_comm_ratio
0.50
3.00
5.50
8.00
10.50
13.00

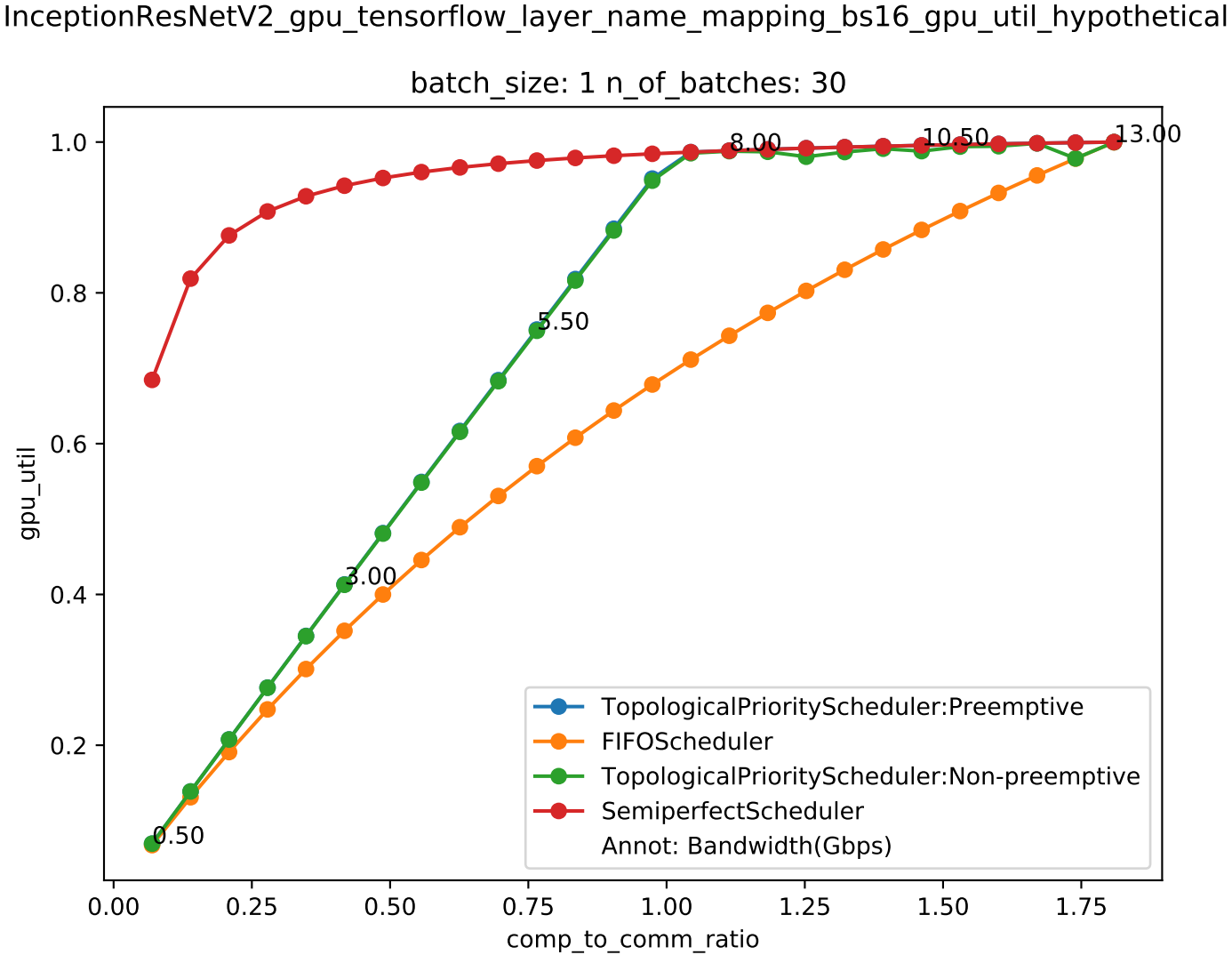

InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_speedup_over_fifo_hypothetical
batch_size: 1 n_of_batches: 30
PerfectScheduler
TopologicalPriorityScheduler:Preemptive
FIFOScheduler
SemiperfectScheduler
TopologicalPriorityScheduler:Non-preemptive
Annot: Bandwidth(Gbps)
speedup_over_fifo
comp_to_comm_ratio
0.50
3.00
5.50
8.00
10.50
13.00

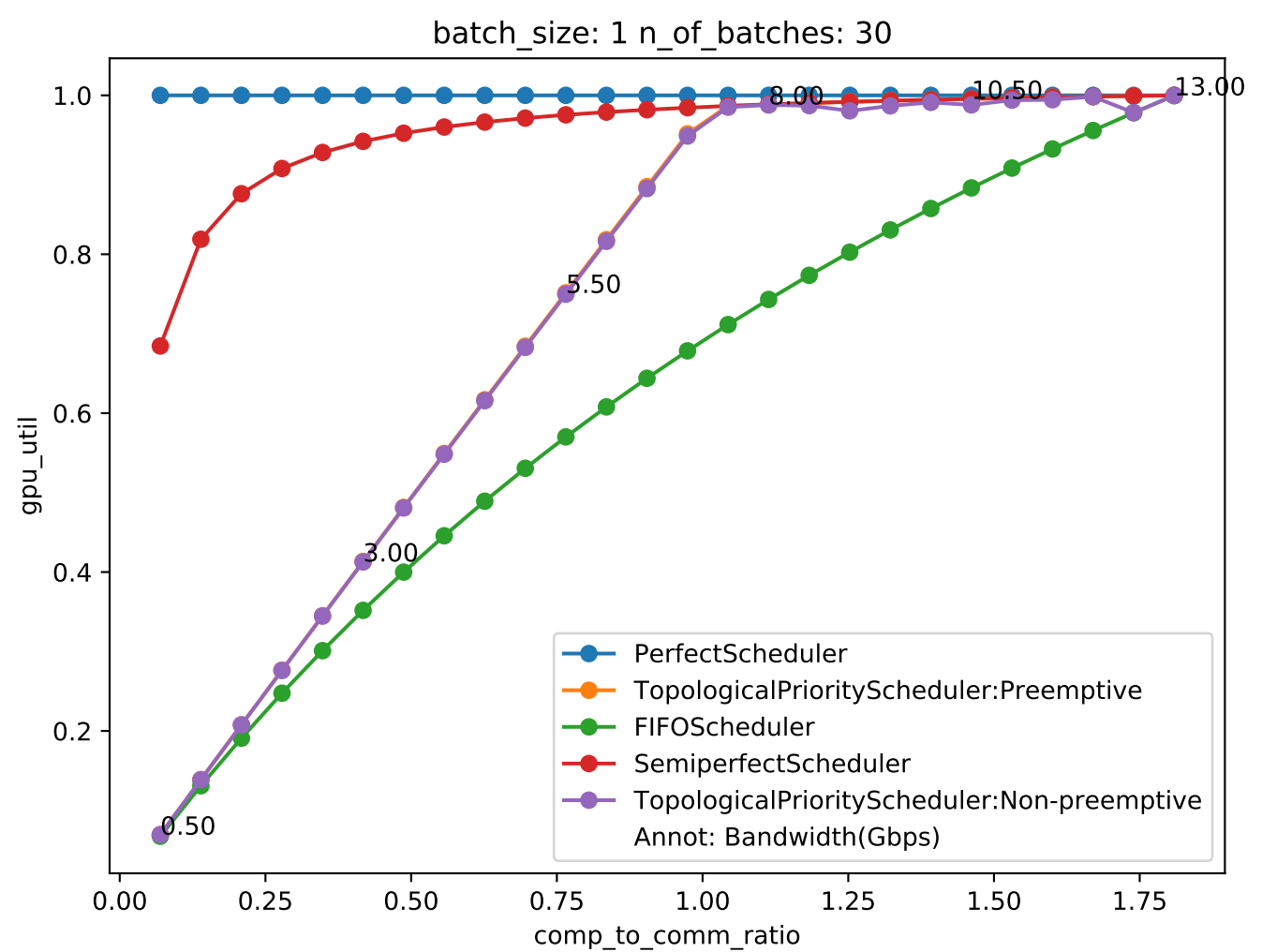
nceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_total_time_steps(s)_hypothetica

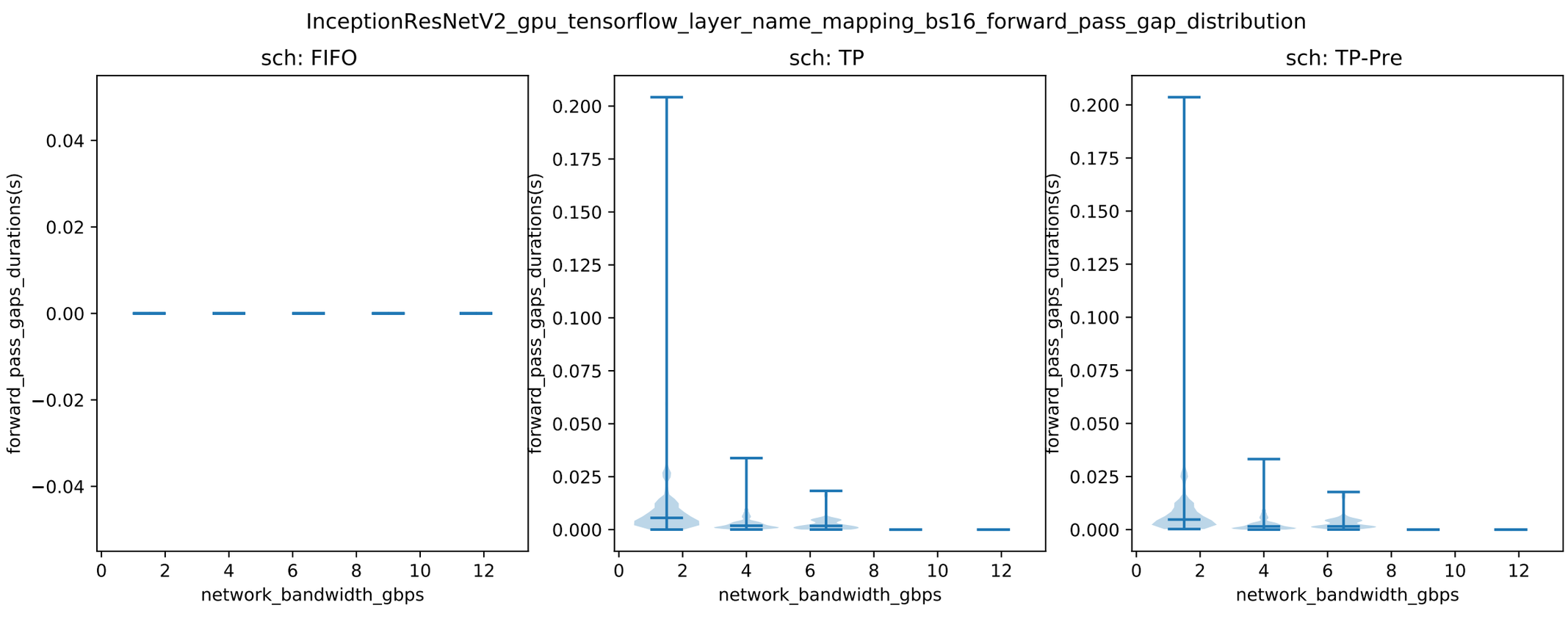


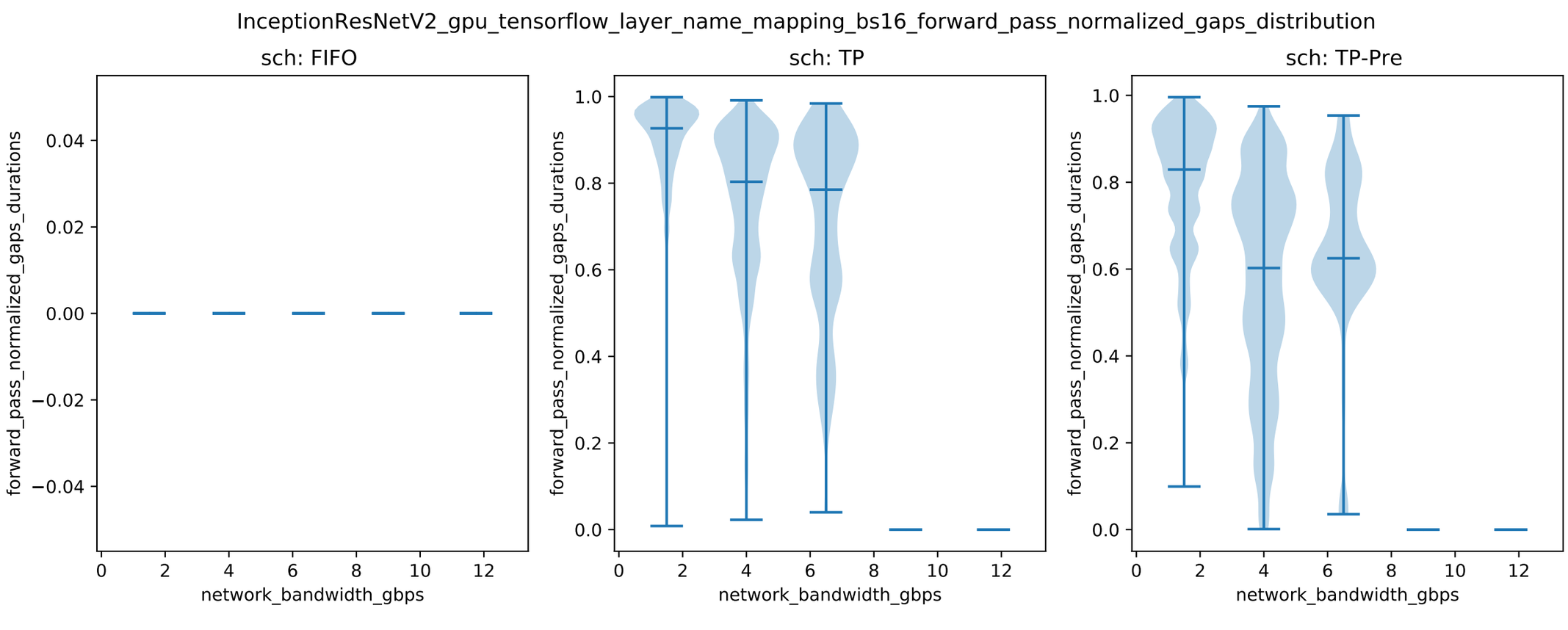
nceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_total_time_steps(s)_hypothetica

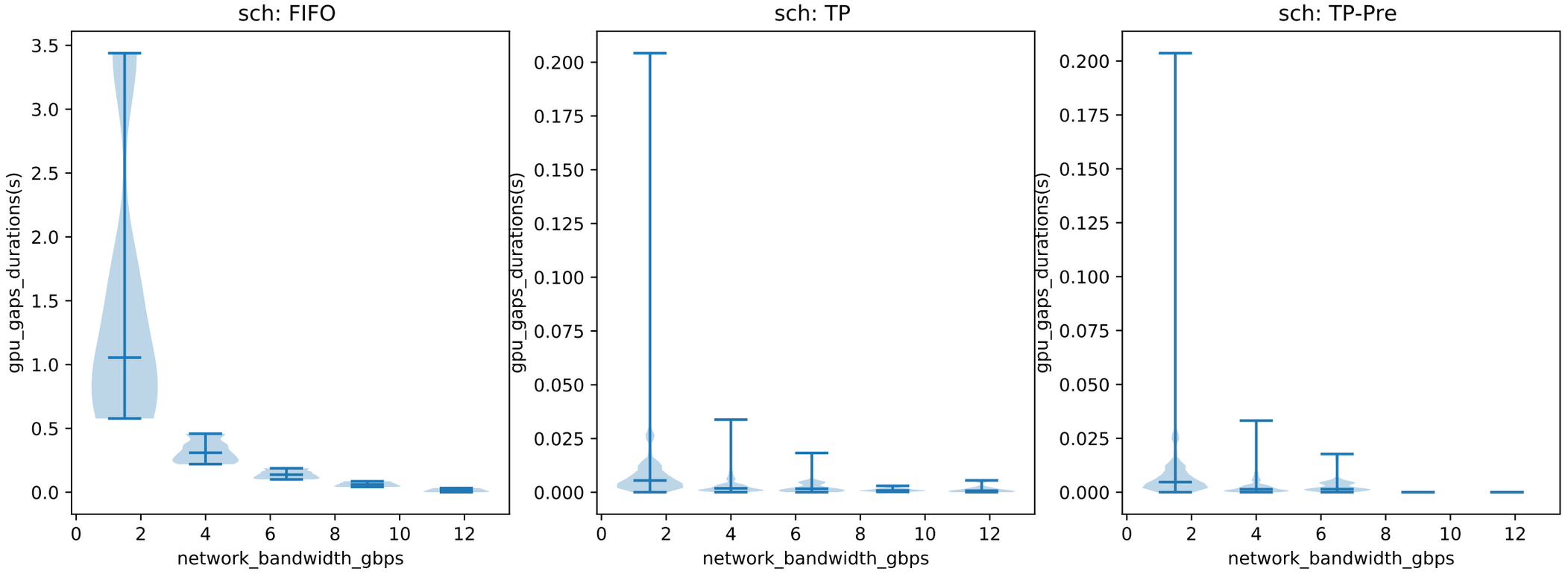


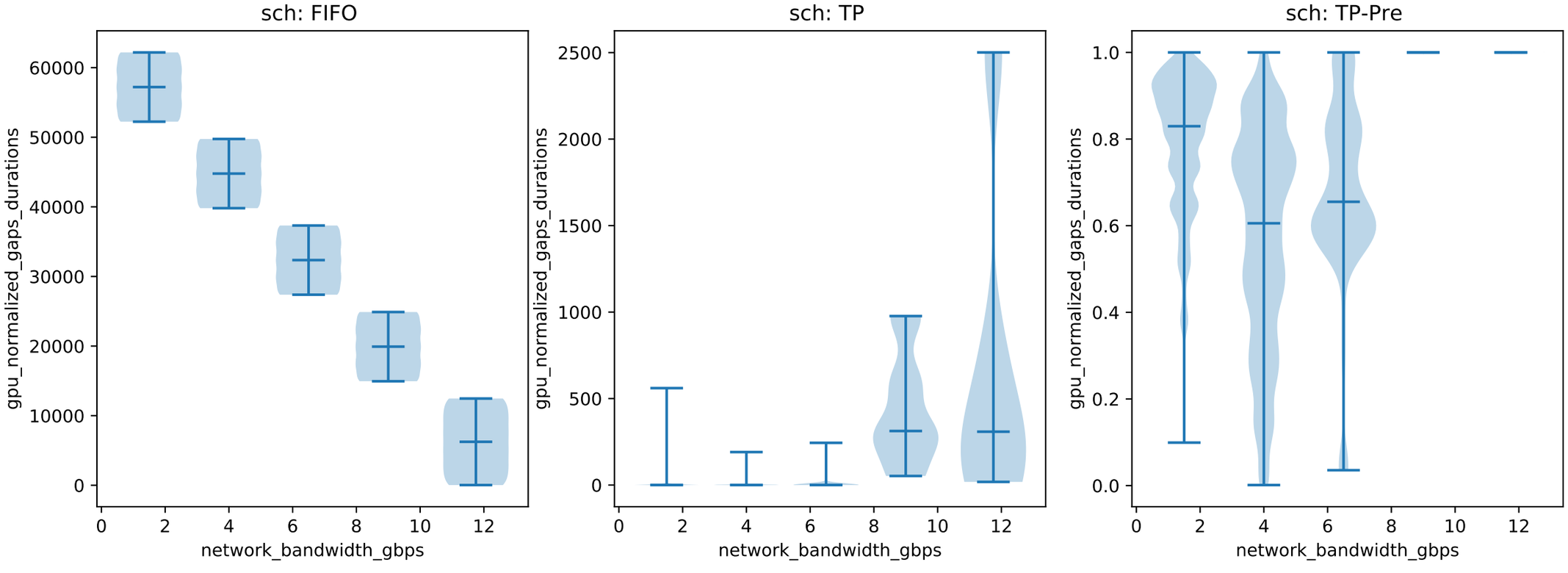

InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_gpu_util_hypothetical
batch_size: 1 n_of_batches: 30
gpu_util
comp_to_comm_ratio
0.50
3.00
5.50
8.00
10.50
13.00
TopologicalPriorityScheduler:Preemptive
FIFOScheduler
TopologicalPriorityScheduler:Non-preemptive
SemiperfectScheduler
Annot: Bandwidth(Gbps)

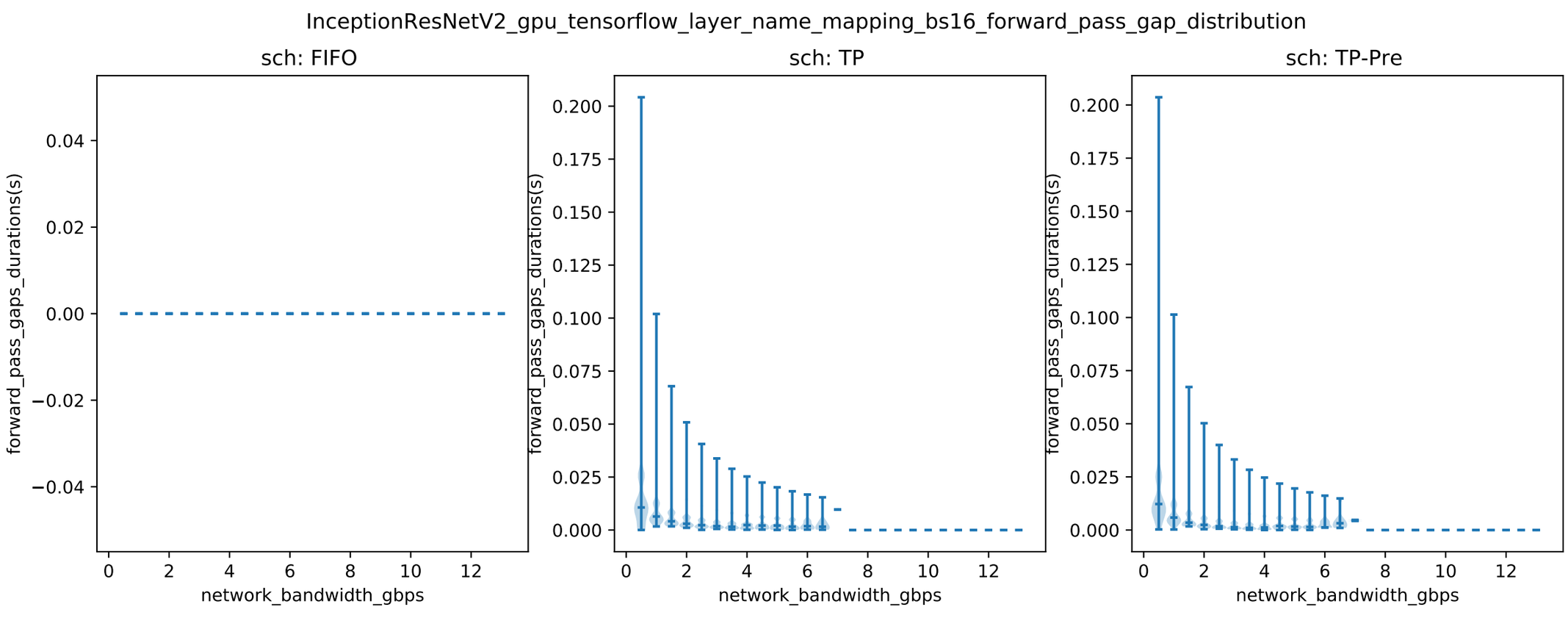

InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_gpu_util_hypothetical
batch_size: 1 n_of_batches: 30
gpu_util
comp_to_comm_ratio
1.0
0.8
0.6
0.4
0.2
0.00
0.25
0.50
0.75
1.00
1.25
1.50
1.75
0.50
3.00
5.50
8.00
10.50
13.00
PerfectScheduler
TopologicalPriorityScheduler:Preemptive
FIFOScheduler
SemiperfectScheduler
TopologicalPriorityScheduler:Non-preemptive
Annot: Bandwidth(Gbps)

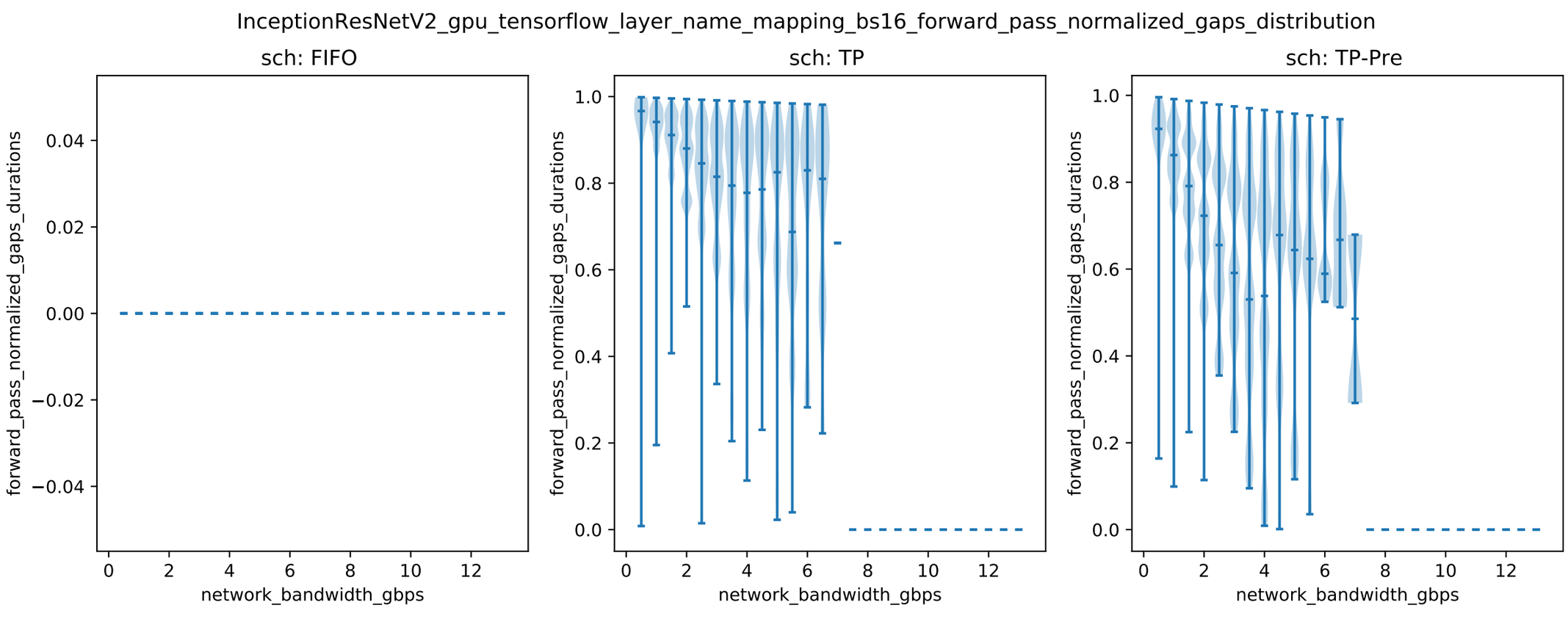

InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_forward_pass_gap_distribution
sch: FIFO
sch: TP
sch: TP-Pre
forward_pass_gaps_durations(s)
network_bandwidth_gbps
0.04
0.02
0.00
−0.02
−0.04
0.200
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
0
2
4
6
8
10
12

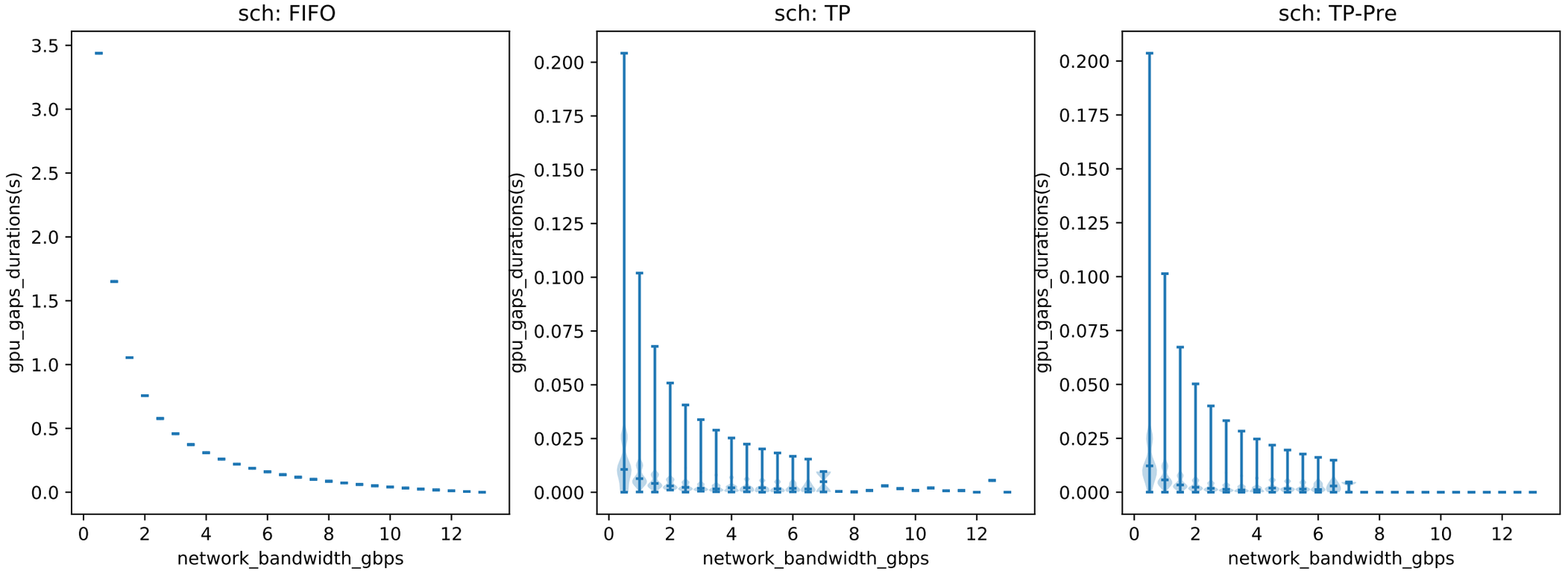

InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_forward_pass_normalized_gaps_distribution
sch: FIFO
sch: TP
sch: TP-Pre
forward_pass_normalized_gaps_durations
network_bandwidth_gbps
0.04
0.02
0.00
−0.02
−0.04
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
12

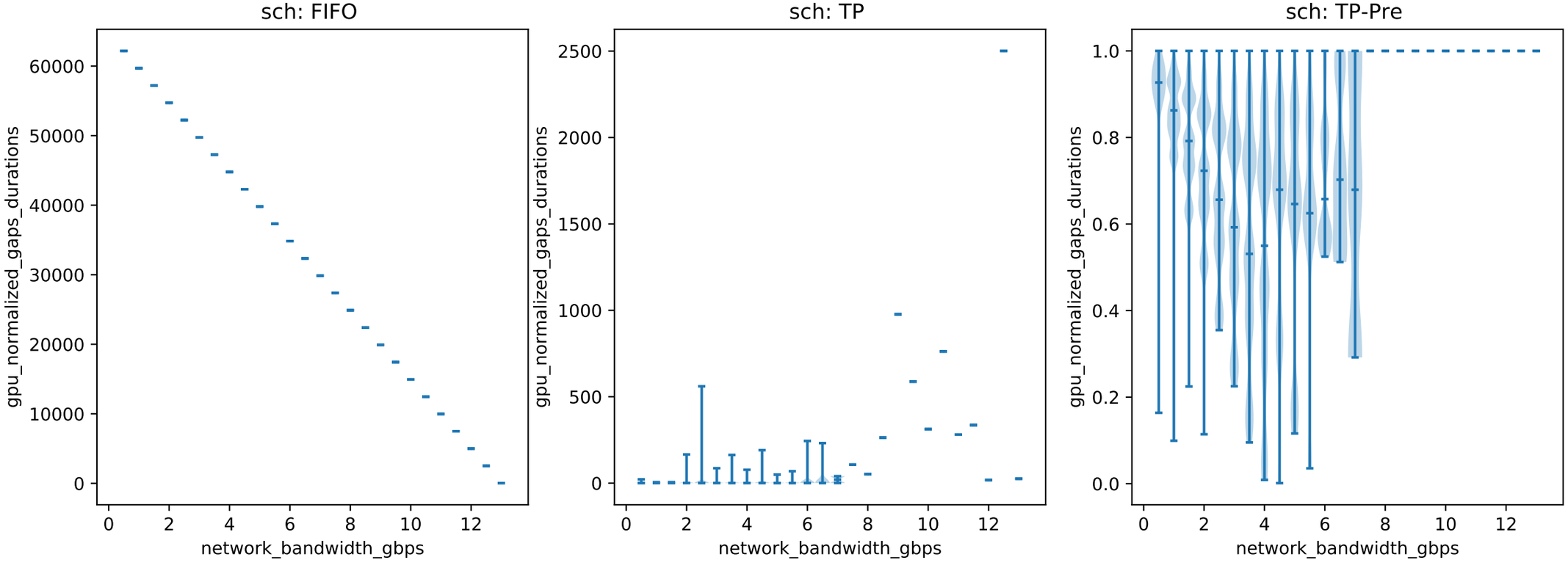

InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_gpu_gap_distribution
sch: FIFO
sch: TP
sch: TP-Pre
gpu_gaps_durations(s)
network_bandwidth_gbps
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.175
0.200
0
2
4
6
8
10
12


InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_gpu_normalized_gaps_distribution
sch: FIFO
sch: TP
sch: TP-Pre
gpu_normalized_gaps_durations
network_bandwidth_gbps
0
10000
20000
30000
40000
50000
60000
500
1000
1500
2000
2500
0.0
0.2
0.4
0.6
0.8
1.0
0
2
4
6
8
10
12


InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_forward_pass_gap_distribution
sch: FIFO
sch: TP
sch: TP-Pre
forward_pass_gaps_durations(s)
network_bandwidth_gbps
0.04
0.02
0.00
−0.02
−0.04
0.200
0.175
0.150
0.125
0.100
0.075
0.050
0.025
0.000
0
2
4
6
8
10
12


InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_forward_pass_normalized_gaps_distribution
sch: FIFO
sch: TP
sch: TP-Pre
forward_pass_normalized_gaps_durations
network_bandwidth_gbps
0.04
0.02
0.00
−0.02
−0.04
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
12


InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_gpu_gap_distribution
sch: FIFO
sch: TP
sch: TP-Pre
gpu_gaps_durations(s)
network_bandwidth_gbps
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
0.000
0.025
0.050
0.075
0.100
0.125
0.150
0.175
0.200
0
2
4
6
8
10
12


InceptionResNetV2_gpu_tensorflow_layer_name_mapping_bs16_gpu_normalized_gaps_distribution
sch: FIFO
sch: TP
sch: TP-Pre
gpu_normalized_gaps_durations
network_bandwidth_gbps
0
10000
20000
30000
40000
50000
60000
500
1000
1500
2000
2500
0.0
0.2
0.4
0.6
0.8
1.0
0
2
4
6
8
10
12