Assistant Professorship of Embedded Systems and Internet of Things
Department of Electrical and Computer Engineering
Technical University of Munich

Quantitative Performance Comparison of Various Individual and Combined Traffic Shapers in Time-Sensitive Networking

Luxi Zhao
Paul Pop
Sebastian Steinhorst

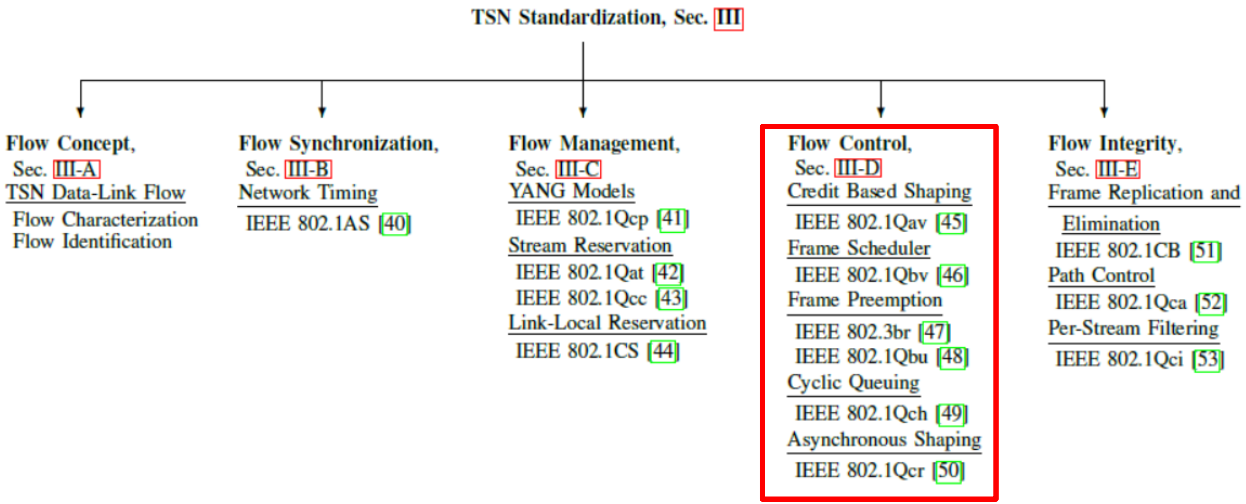
Necessity and Contributions

Necessity

- Set of substandards (flow control):
  802.1Qbv – Time Aware Shaper (TAS);
  802.1Qav – Credit Based Shaper (CBS);
  802.1Qcr – Asynchronous Traffic Shaper (ATS);
  802.1Q – 2005 – Strict Priority (SP);
  Combinations …
- Independent studies;
  No quantitative comparison;
  Proper shapers selection – tricky

Contributions

- Tutorial of NC-based analysis for TSN;
- Two new combined architectures (TAS+ATS+CBS, TAS+ATS+SP); extend NC approach;
- Plenty of quantitative comparison → surprising but interesting results;
- Provide a basis, select the suitable TSN traffic shapers.

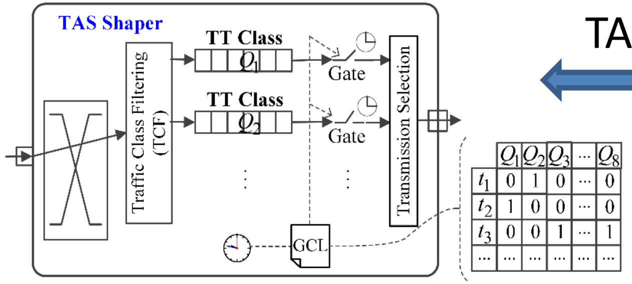


Architecture – Individual Traffic Shapers

Evaluation Parameters

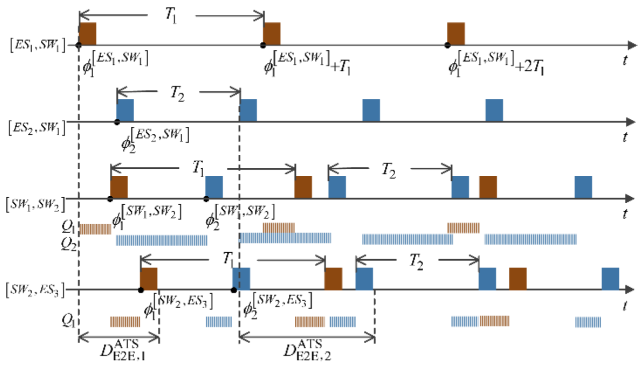
- Schedulability – End-to-end latency bound
- Buffer size without frame loss – Backlog bound
- Stable Communication – Jitter bound

1. 802.1Qbv
2. Scheduling Synthesis
3. 2016, [1] S. S. Craciunas et al.; [2] P. Pop et al.



TAS

- IEEE 802.1Qbv – Time Aware Shaper (TAS);
- Global network clock synchronization (IEEE 802.1ASrev);
- Time-Triggered communication – GCL synthesis – Schedulability guarantee;
- GCL synthesis – NP-complete problem [1], [2].



GCL – Gantt Chart

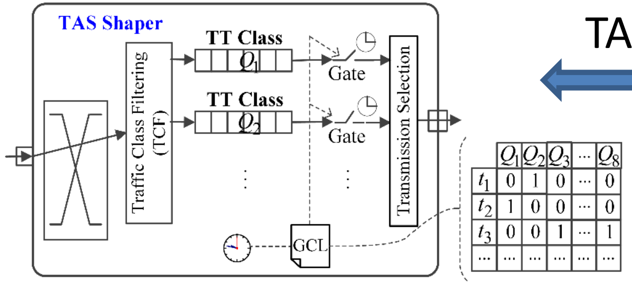
Individual Traffic Shapers

Architecture – Individual Traffic Shapers

Evaluation Parameters

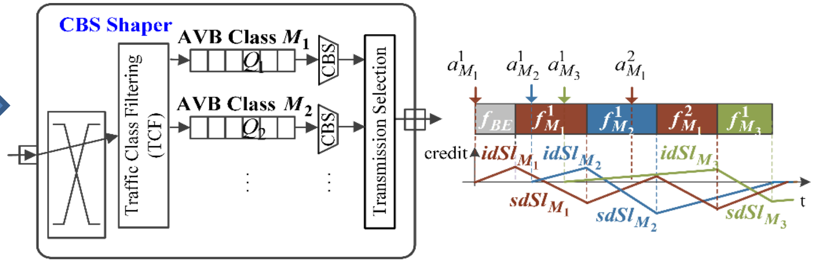
- Schedulability – End-to-end latency bound
- Buffer size without frame loss – Backlog bound
- Stable Communication – Jitter bound

TAS



1. 802.1Qbv
2. Scheduling Synthesis
3. 2016, [1] S. S. Craciunas et al.; [2] P. Pop et al.

CBS

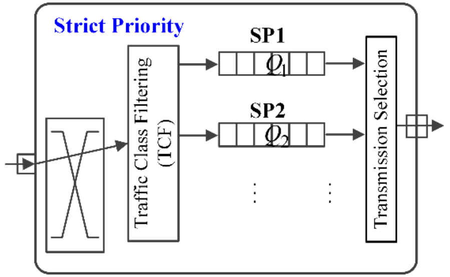


(a) Architecture (b) CBS forwarding frames example

1. 802.1Qav
2. Network Calculus
3. 2014, [3] J. A. R. De Azua et al.

- IEEE 802.1Qav – Credit Based Shaper (CBS);
- Allocate the bandwidth reservation for different classes (priority)
- CBS algorithm – credit value (idleSlope / sendSlope) – non-work conserving;
- Schedulability guarantee – Network Calculus [3];

SP



1. 802.1Q - 2005
2. Network Calculus
3. 2003, [4] J. Schmitt et. al.

- IEEE 802.1Q - 2005 – Strict Priority (SP);
- Low priority traffic can transmit only when the high priority queue is empty;
- Schedulability guarantee – Network Calculus [4];

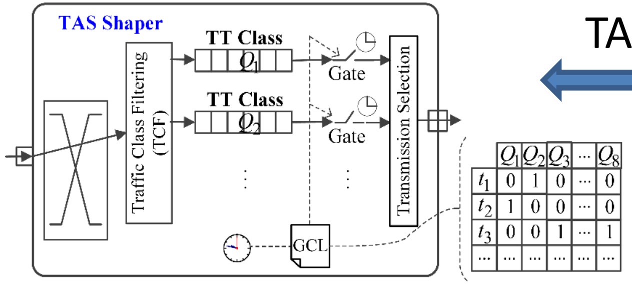
Individual Traffic Shapers

Architecture – Individual Traffic Shapers

Evaluation Parameters

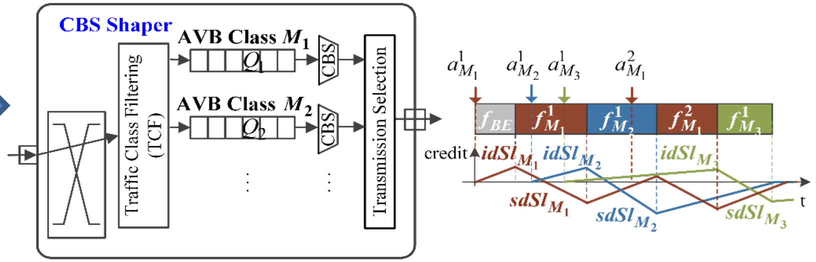
- Schedulability – End-to-end latency bound
- Buffer size without frame loss – Backlog bound
- Stable Communication – Jitter bound

TAS



1. 802.1Qbv
2. Scheduling Synthesis
3. 2016, [1] S. S. Craciunas et al.; [2] P. Pop et al.

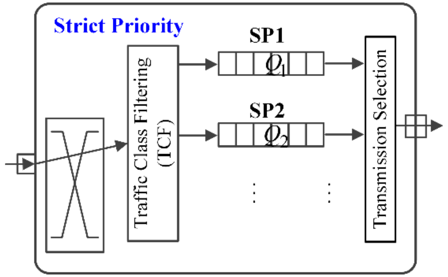
CBS



(a) Architecture (b) CBS forwarding frames example

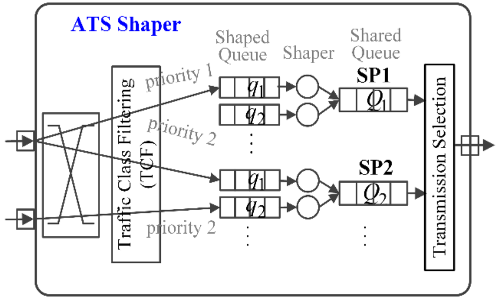
1. 802.1Qav
2. Network Calculus
3. 2014, [3] J. A. R. De Azua et al.

SP



1. 802.1Q - 2005
2. Network Calculus
3. 2003, [4] J. Schmitt et. al.

ATS (+SP)



1. 802.1Qcr
2. Closed-form Formula
3. 2016, [5] J. Specht et al. 2018, [6] J. Y. Le Boudec

- IEEE 802.1Qcr – Asynchronous Traffic Shaper (ATS);
- Asynchronous transmission, local clock;
- Two hierarchies of queues – shaped queue & shared queue;
- ATS – interleaved regulator – avoid burstiness cascades;
- Schedulability guarantee – Closed-form Formula [5], Network Calculus [6].

Individual Traffic Shapers

Evaluation – Individual Traffic Shapers

- Synthetic test cases – SRM, MR, MM, ST, MT
  - Each topology – 100 TCes;
  - Frame size – minimum (64 bytes) ~ maximum (1522 bytes);
  - Period (periodic) / Min time interval (sporadic) – T={1000, 2000, 5000, 10000} (µm);
  - 1 priority;
  - GCLs for TAS, Route – existing work [2];
  - Physical link rate 100 Mb/s.

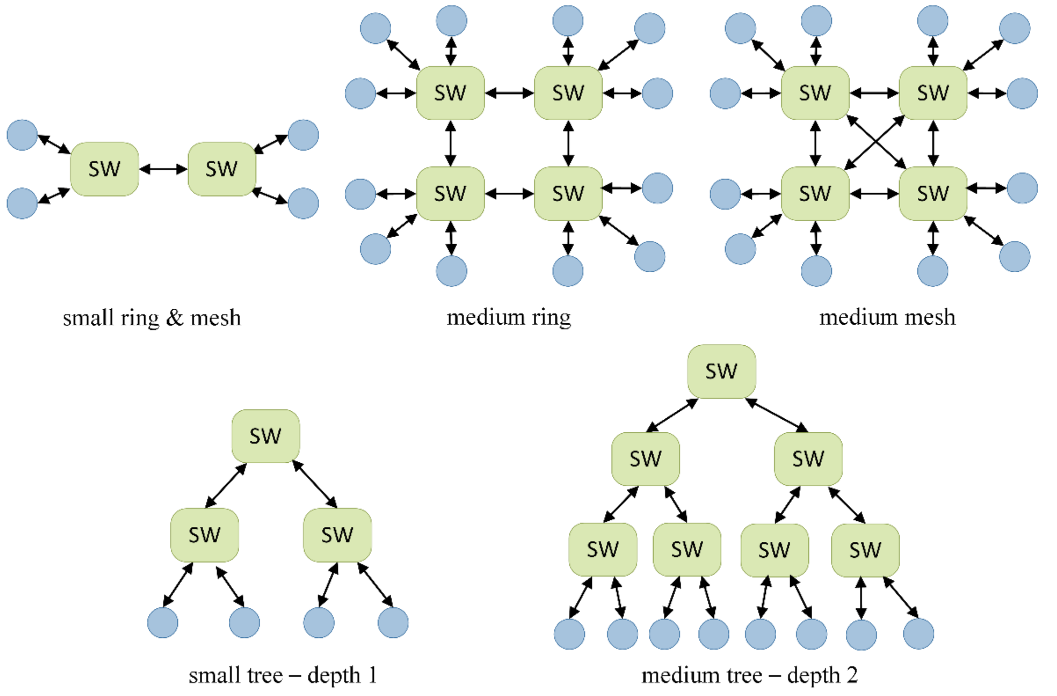
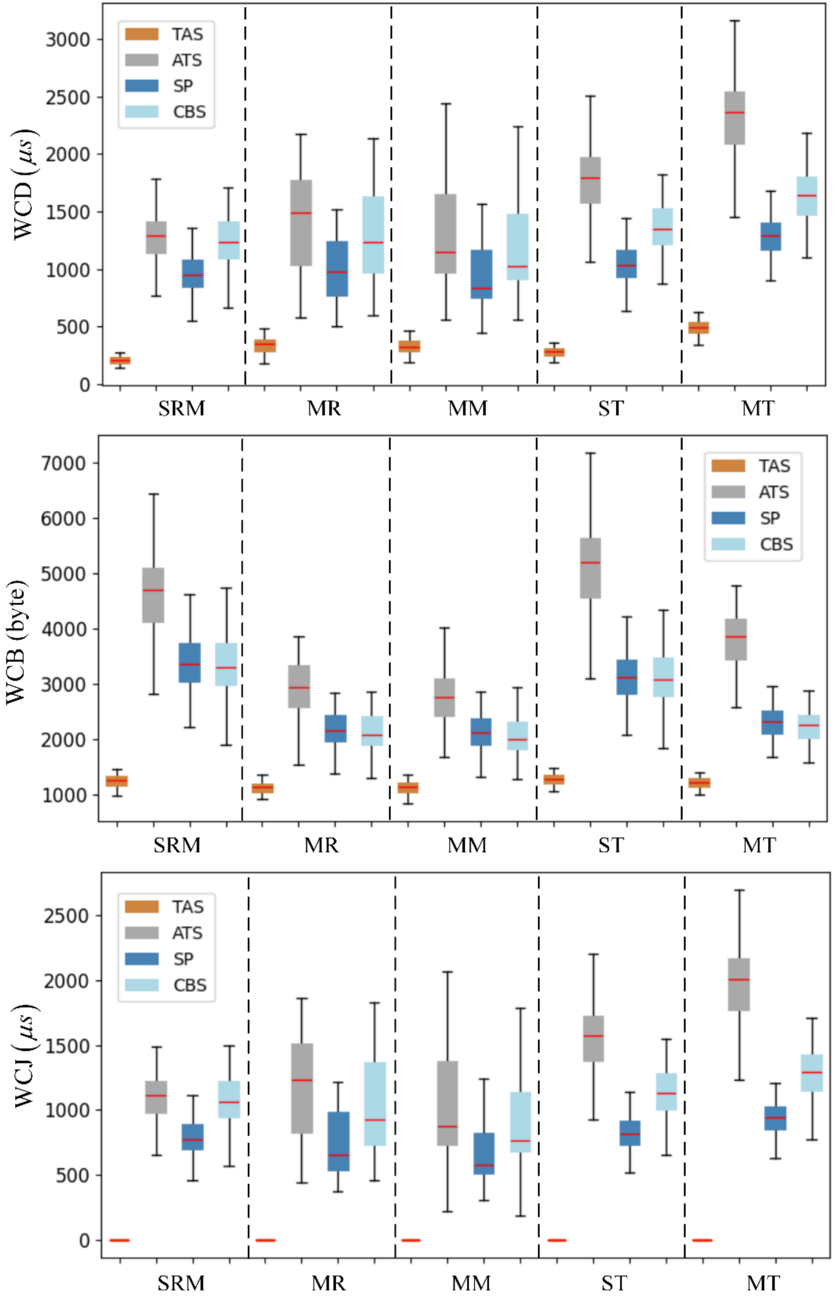


TABLE II
STATISTICAL HOPS AND TRAFFIC LOAD FOR 100 TEST CASES

	SRM	MR	MM	ST	MT
Average Hops	2.7	4.2	3.8	3.5	5.5
Average Traffic Load	28.9%	20.5%	17.4%	29.0%	19.7%
Max Traffic Load	47%	40%	38%	47%	30%
Min Traffic Load	13%	8%	6%	13%	10%

Evaluation – Individual Traffic Shapers (1)

- **Results**
  - Each TC,
    1) End-to-end latency upper bounds – flow;
    2) Backlog upper bounds – egress port;
    3) Jitter bounds – flow.
  - Figure: each TC – metric – average value;
    100 TC – 100 dots - box plot.
- **Comments**
  - Different topologies – similar trends while comparing different traffic shapers;
  - TAS performs the best – latency, backlog, jitter;
  - …



End-to-end latency

Backlog

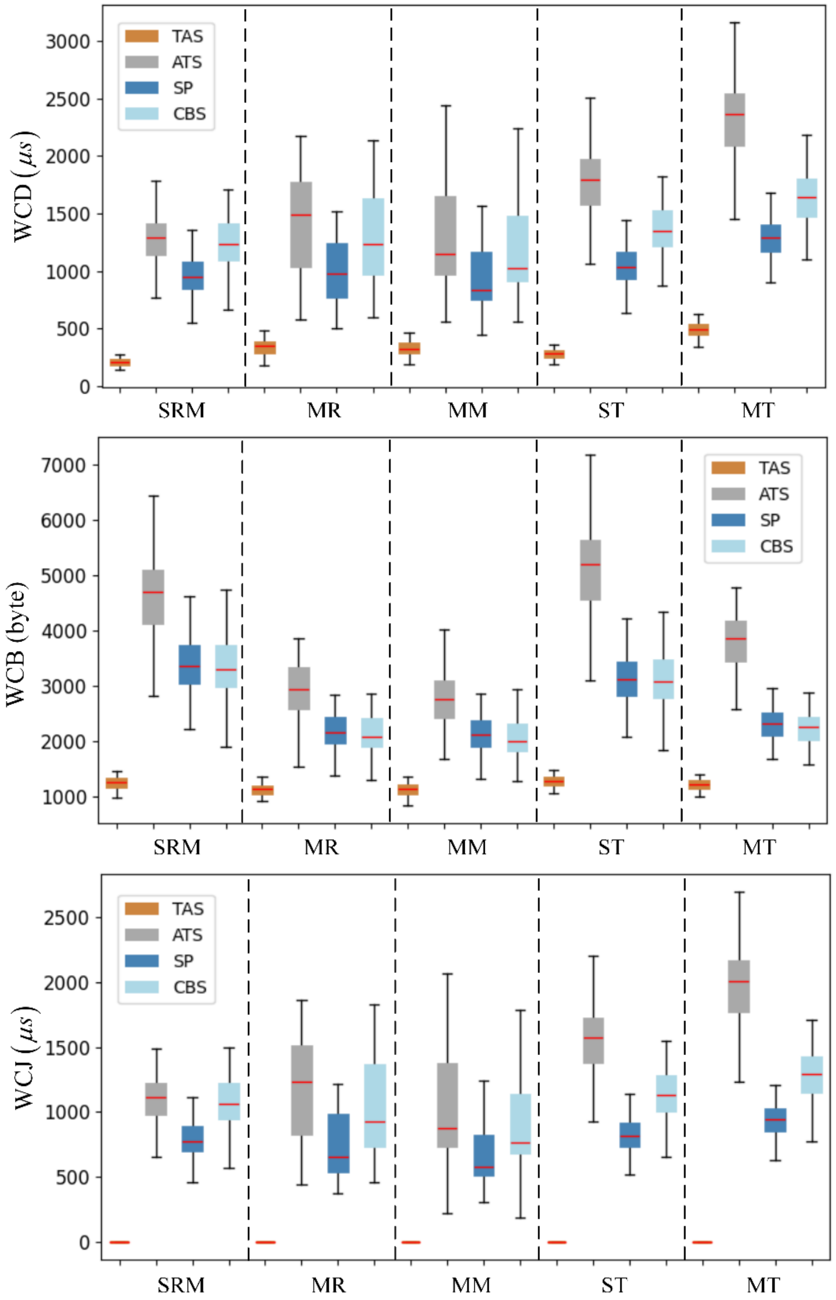
Jitter

Figure 1

Evaluation – Individual Traffic Shapers (1)

- **Results**
  - Table: X_i metric value for flow f_i;
    $\bar{X} = aver(X_i) = aver\left(\left(X_i^{Y_1} - X_i^{Y_2}\right)/X_i^{Y_2}\right)$
- **Comments**
  - latency & jitter – SP<CBS<ATS;
    backlog – CBS<SP<ATS;
  - Advantage of ATS ↓ ← concentration of flows ↑
    ← number of hops ↑
  - …

		(CBS – SP) /SP	(ATS – SP) /SP	(ATS – CBS) /CBS
Average WCD	SRM	30.7%	34.1%	2.8%
	MR	28.2%	43.9%	12.5%
	MM	26.2%	35.7%	7.7%
	ST	31.2%	72.3%	31.4%
	MT	27.3%	79.1%	40.8%
Average WCB	SRM	-1.3%	38.0%	39.9%
	MR	-2.5%	34.7%	38.1%
	MM	-3.2%	31.0%	35.4%
	ST	-0.8%	65.0%	66.5%
	MT	-3.3%	64.4%	70.0%
Average WCJ	SRM	37.1%	41.4%	3.4%
	MR	38.0%	60.7%	16.6%
	MM	30.3%	43.0%	10.4%
	ST	39.5%	91.9%	37.7%
	MT	37.3%	108.3%	51.9%



End-to-end latency

Backlog

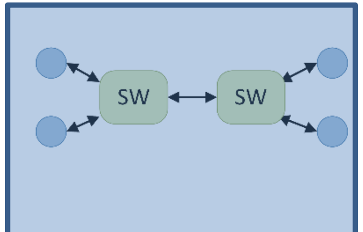
Jitter

Figure 1

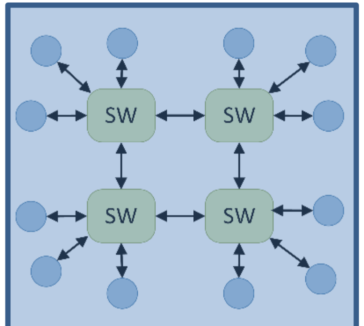
Evaluation – Individual Traffic Shapers (1)

Comments

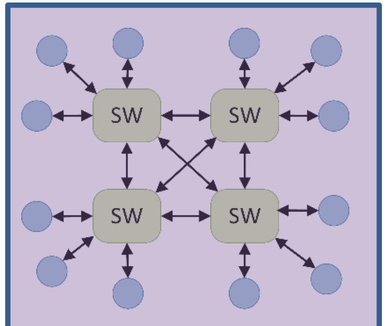
- ATS positive effect ↓ ← concentration of flows ↑
  ← number of hops ↑
- For example
  1. Flows concentration: MR > MM →
     ATS positive effect: MR < MM
  2. Number of hops: ST > SRM →
     ATS positive effect: ST < SRM



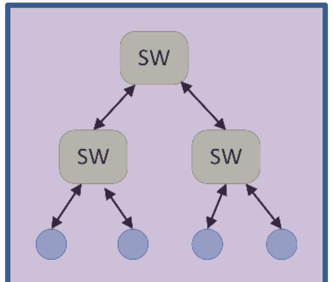
small ring & mesh



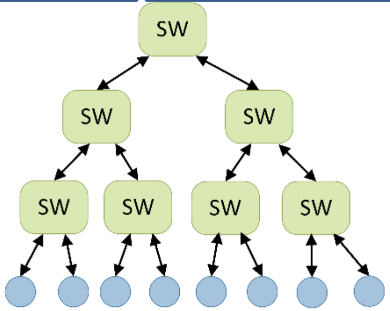
medium ring



medium mesh



small tree – depth 1



medium tree – depth 2

		[illegible]	(ATS – SP) /SP	[illegible]
Average WCD	SRM	[illegible]	34.1%	[illegible]
	MR	[illegible]	43.9%	[illegible]
	MM	[illegible]	35.7%	[illegible]
	ST	[illegible]	72.3%	[illegible]
	MT	[illegible]	79.1%	[illegible]
Average WCB	SRM	[illegible]	38.0%	[illegible]
	MR	[illegible]	34.7%	[illegible]
	MM	[illegible]	31.0%	[illegible]
	ST	[illegible]	65.0%	[illegible]
	MT	[illegible]	64.4%	[illegible]
Average WCJ	SRM	[illegible]	41.4%	[illegible]
	MR	[illegible]	60.7%	[illegible]
	MM	[illegible]	43.0%	[illegible]
	ST	[illegible]	91.9%	[illegible]
	MT	[illegible]	108.3%	[illegible]

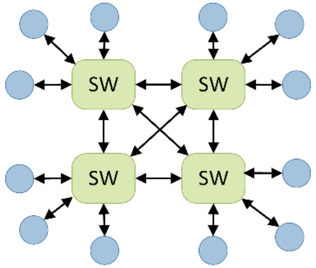
Table 1

TABLE II
STATISTICAL HOPS AND TRAFFIC LOAD FOR 100 TEST CASES

	SRM	MR	MM	ST	MT
Average Hops	2.7	4.2	3.8	3.5	5.5
Average Traffic Load	28.9%	20.5%	17.4%	29.0%	19.7%
Max Traffic Load	47%	40%	38%	47%	30%
Min Traffic Load	13%	8%	6%	13%	10%

Evaluation – Individual Traffic Shapers (2)

- **Synthetic test cases – MM**
  - **Case 1 – 1 priority**
  - Average traffic load: 10% ~ 90%
  - Each traffic load – 20 Tces



medium mesh

- **Results**
  - $X = aver(X_i) = aver\left(\left(X_i^{SP} - X_i^{ATS}\right)/X_i^{ATS}\right)$
- **Comments**
  - End-to-end latency bounds:
    – Average traffic load ↑ → comparison percentage X ↑ → ATS positive effect ↑;
    – Average traffic load < 70% – SP performs better than ATS;
  - Backlog bounds:
    – Average traffic load < 70% – unobvious change;
    – ATS performance is always inferior to SP.
  - …

ATS vs. without ATS (i.e., SP)

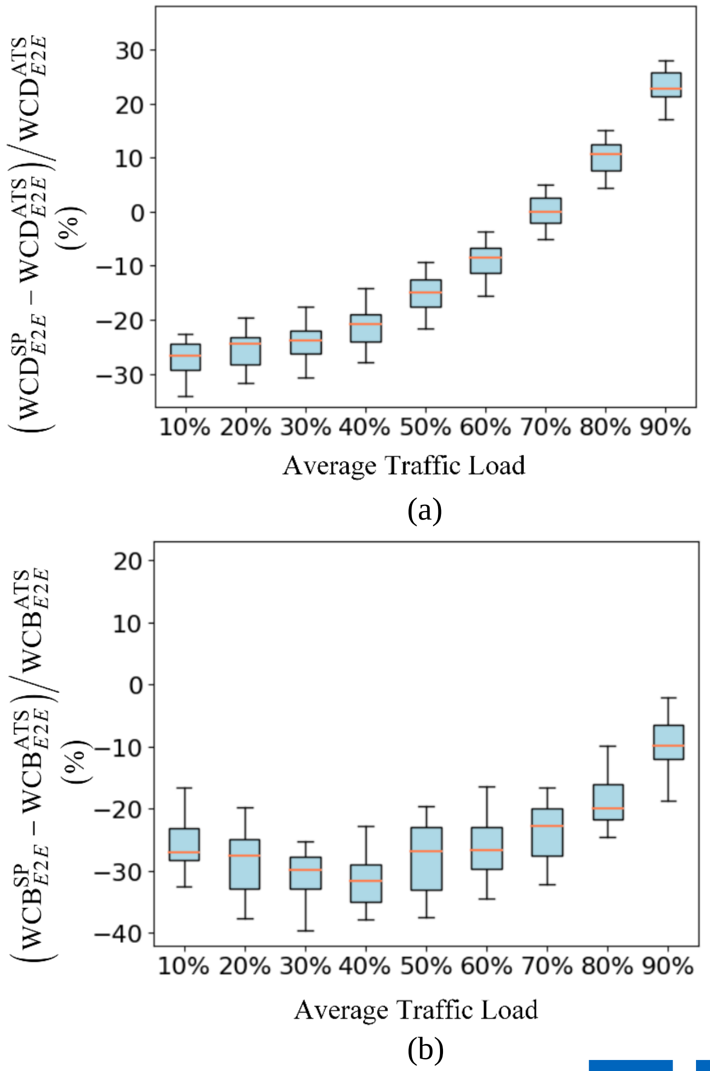
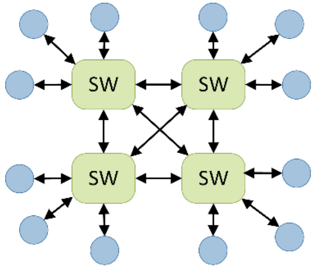


Figure 2

Evaluation – Individual Traffic Shapers (3)

- **Synthetic test cases – MM**
  - **Case 2 – 2 priorities + BE**
  - Average traffic load: 10% ~ 90%;
    High/low: 50% of overall traffic load
  - Each traffic load – 20 Tces



medium mesh

- **Results**
  - $X = aver(X_i) = aver\left(\left(X_i^{SP} - X_i^{ATS}\right)/X_i^{ATS}\right)$
- **Comments**
  - End-to-end latency bounds:
    <u>High priority</u>:
    – ≈ Top 40% traffic load with single priority;
    – ATS no positive effect on high-priority traffic.
    <u>Low-priority</u>:
    – ATS positive effect on low-priority traffic ← average overall traffic load > 30%.
  - …

ATS vs. without ATS (i.e., SP)

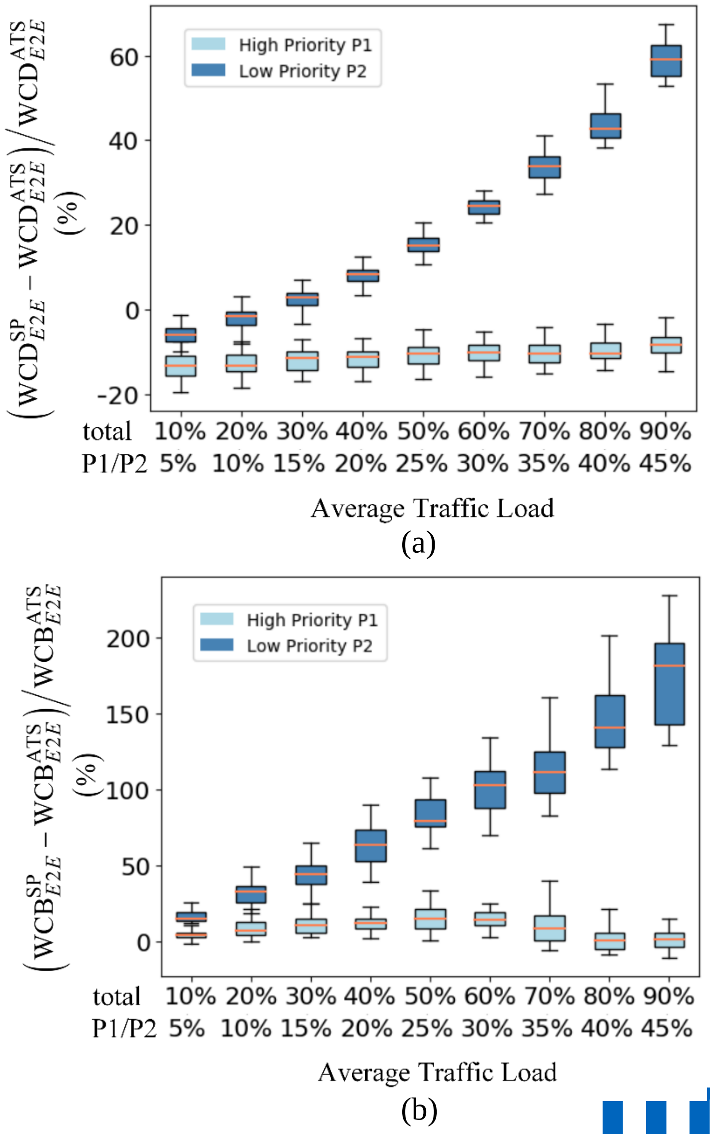
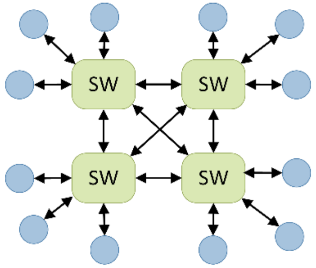


Figure 3

Evaluation – Individual Traffic Shapers (3)

- Synthetic test cases – MM
  - **Case 3 – 2 priorities + BE**
  - Average traffic load: 10% ~ 90%;
    High/low: 50% of overall traffic load
  - Each traffic load – 20 Tces
- Results
  - $X = aver(X_i) = aver\left(\left(X_i^{SP} - X_i^{ATS}\right)/X_i^{ATS}\right)$
- Comments
  - Backlog bounds:
    <u>High priority</u>:
    – ≈ Top 40% traffic load with single priority + BE;
    <u>Low-priority</u>:
    – ATS positive effect on low-priority traffic.



medium mesh

ATS vs. without ATS (i.e., SP)

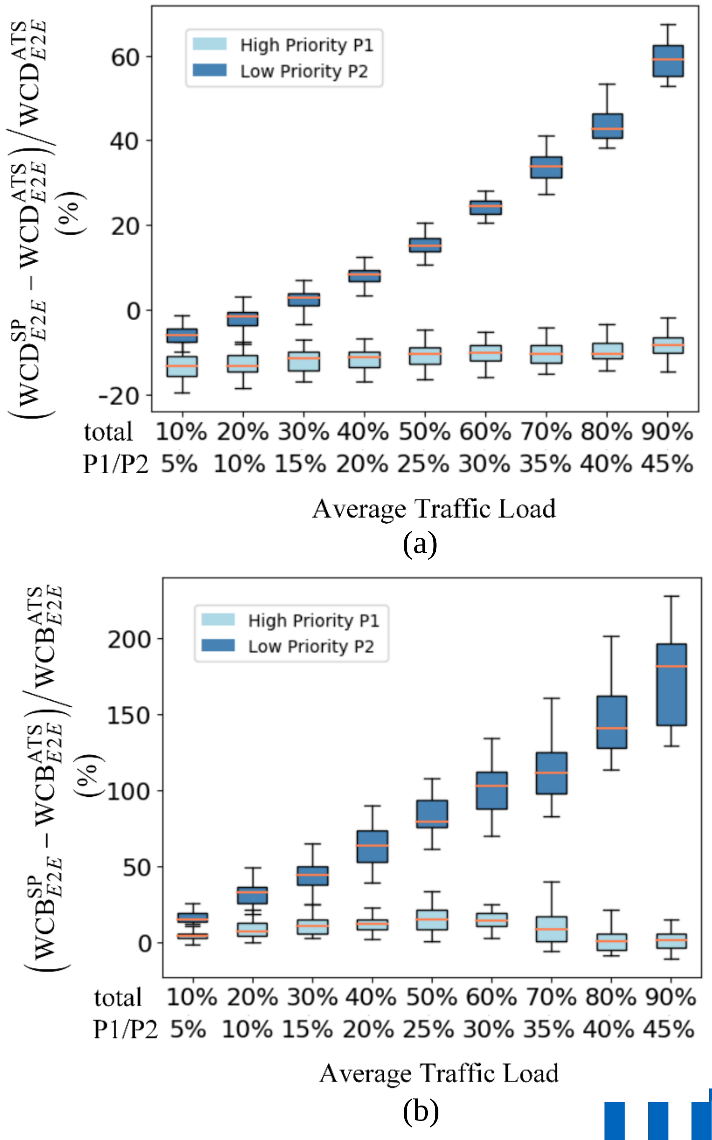
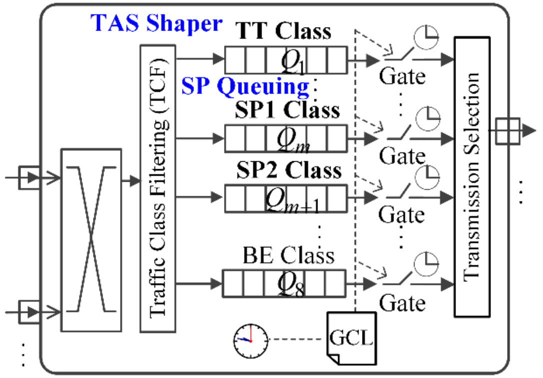


Figure 3

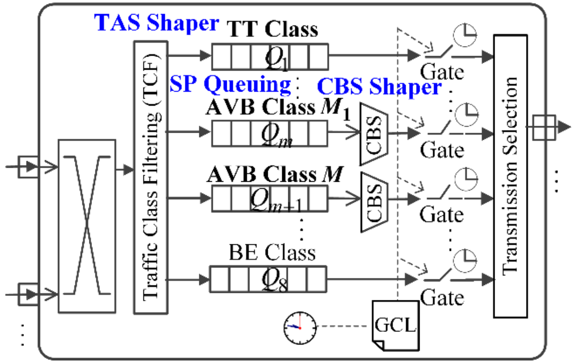
Architecture – Combined Traffic Shapers

1. 802.1Qbv+802.1Q - 2005
2. Network Calculus
3. 2017, [7] L. Zhao et. al.



TAS+SP

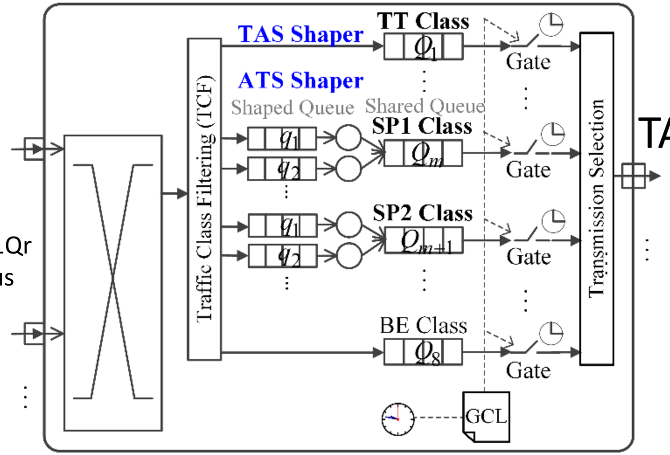
TAS+CBS



1. 802.1Qbv+802.1Qav
2. Network Calculus
3. 2016, [8], [9] L. Zhao et. al.

NEW

1. 802.1Qbv+802.1Qr
2. Network Calculus

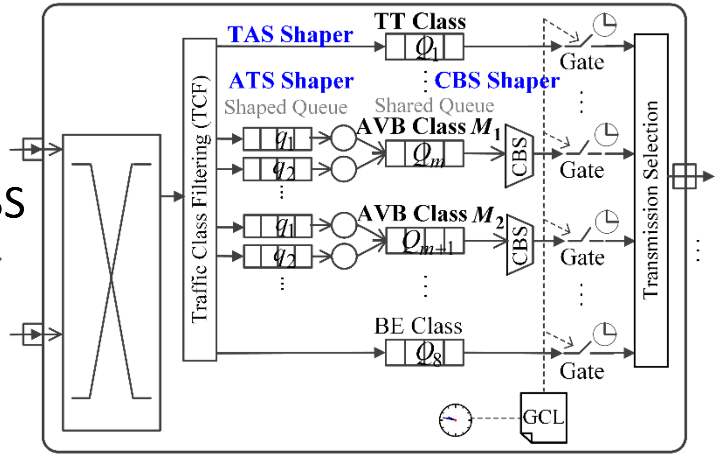


NEW

TAS+**ATS**(+SP)

NEW

TAS+**ATS**+CBS



NEW

1. 802.1Qbv+802.1Qcr +802.1Qav

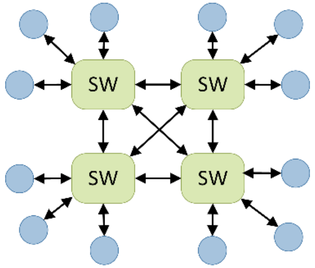
- TAS outperforms than all the others (latency, backlog, jitter); Scalability problem.
- Promising combination model: TAS+X

- New: TAS+ATS(+SP); TAS+ATS+CBS
- TAS+SP vs. TAS+ATS(+SP)
  TAS+CBS vs. TAS+ATS+CBS

Combined Traffic Shapers

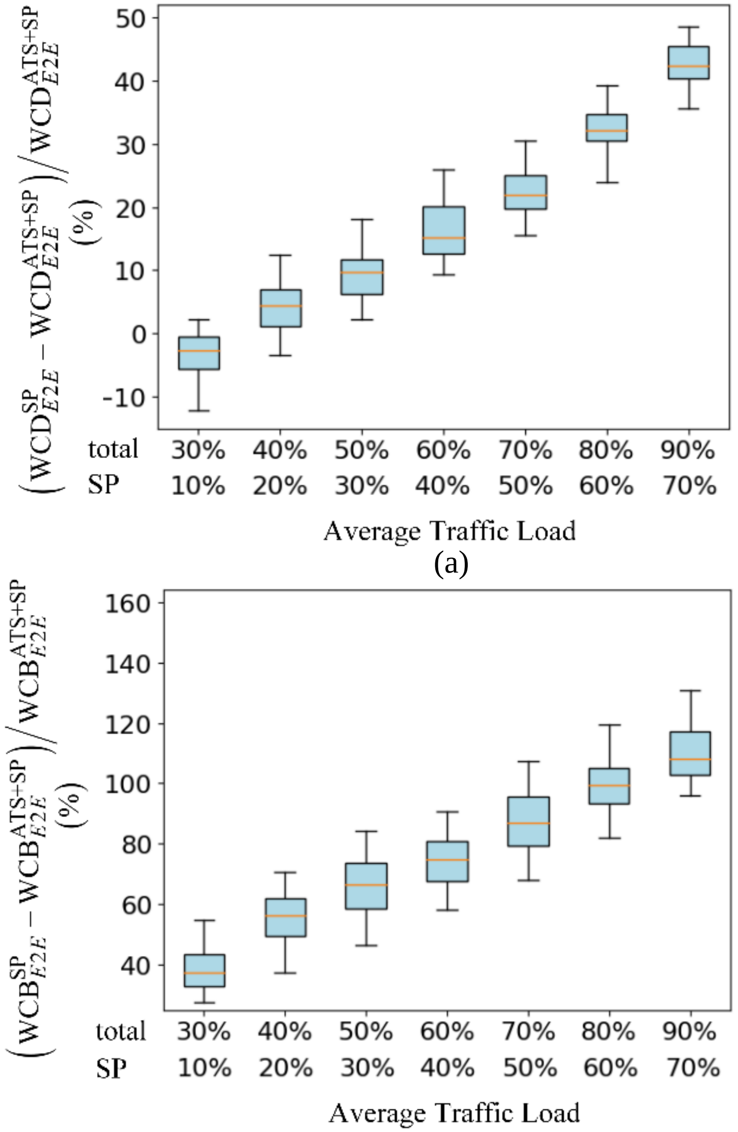
Evaluation – Combined Traffic Shapers (1)

- **Synthetic test cases – MM**
  - **Case 1 – TT traffic load: 20%;**
    – SP average traffic load: 10% ~ 70%
  - Each traffic load – 20 TCes
- **Results**
  - $X = aver(X_i) = aver\left(\left(X_i^{SP} - X_i^{ATS+SP}\right)/X_i^{ATS+SP}\right)$
- **Comments**
  - With the influence of TT traffic (TAS)
  - End-to-end latency bounds:
    – Average traffic load ↑ – ATS positive effect ↑;
    – ATS positive effect ← average overall traffic load > 40%;
  - Backlog bounds:
    – ATS positive effect ← average overall traffic load > 30%;



medium mesh

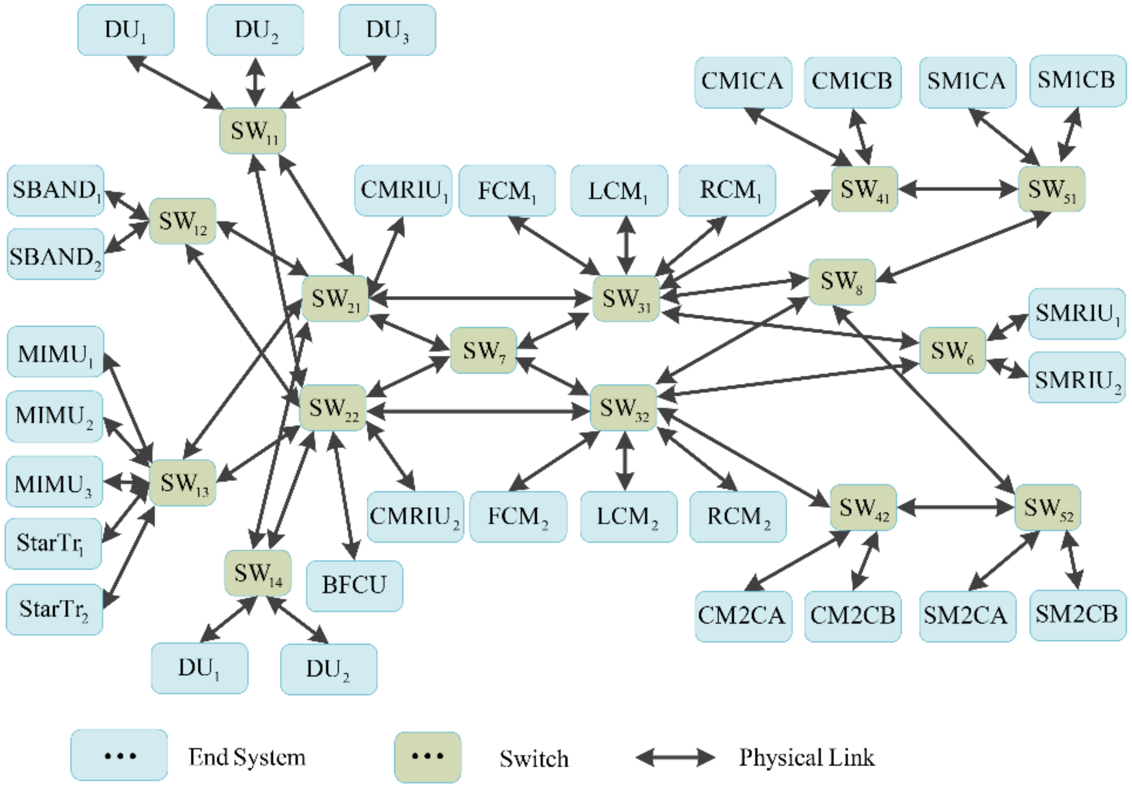
TAS+SP vs. TAS+ATS (+SP) – TT 20%



(b)
Figure 4

Evaluation – Combined Traffic Shapers (2)

- **Realistic Test Cases – Orion CEV**
  - 31 ESes, 15 SWs, 188 dataflow routes, 100 Mbps link rate;
  - 99 TT flows (TAS), 87 rate constraint flows with the same priority → SP flows / AVB flows (CBS);
  - TT traffic load in network
    → 1.5% on average & 5.5% in maximum.
    Overall traffic load in network
    → 3.5% on average & 10% in maximum.
  - IdleSlope for AVB is set to 75% (default);
- **Results**
  - $100 \times ln(X), X = (WCD, WCB);$
  - Sorted in increasing order by results (WCD, WCB).



Orion CEV

Evaluation – Combined Traffic Shapers (2)

- **Comments**
  - End-to-end latency bounds: – ATS no positive effect.
  - Backlog bounds: – ATS positive effect.
  - ← Average overall traffic load (TT, SP/AVB) low;
  - → Consistent with results in Fig. 4.

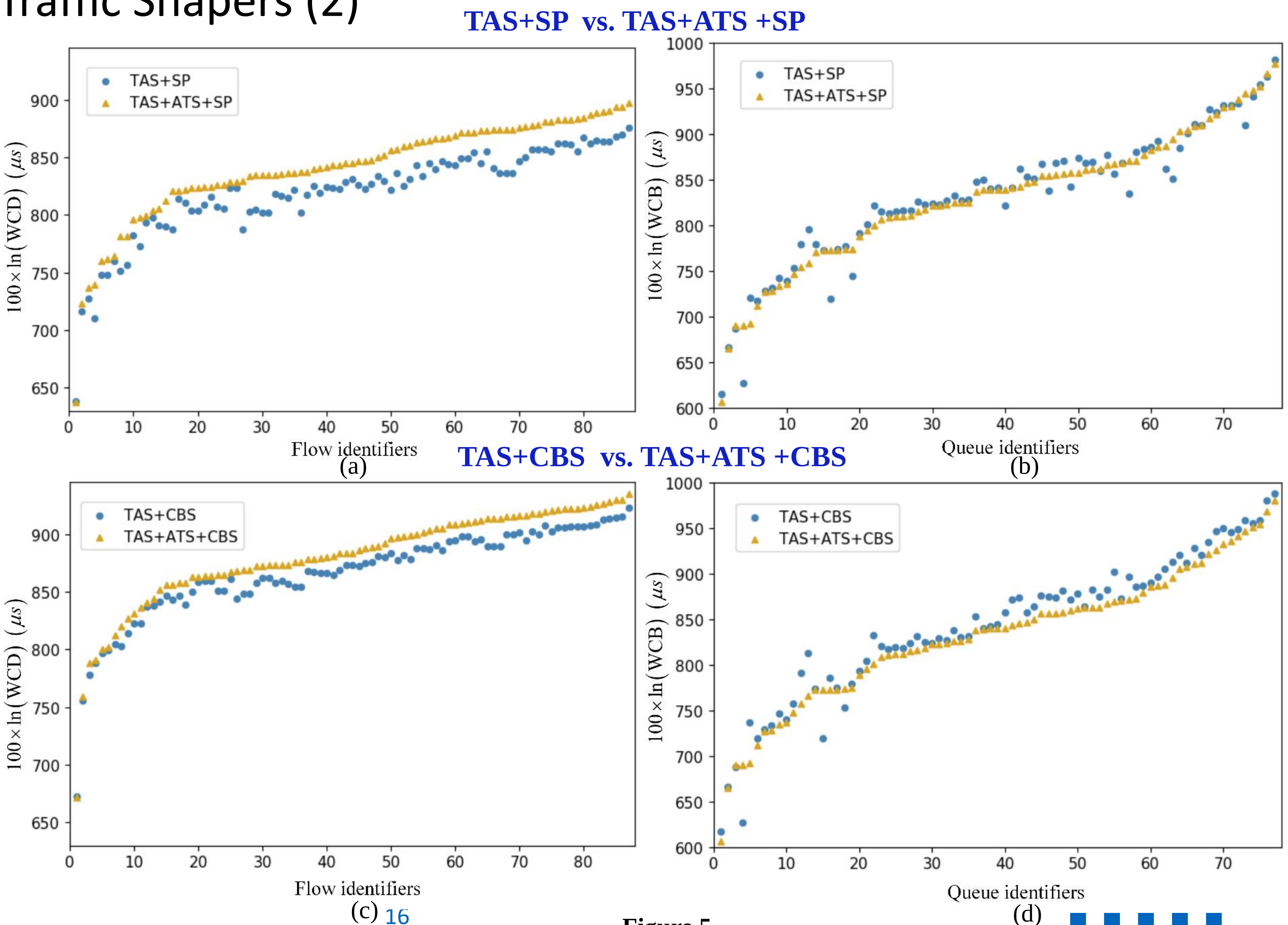
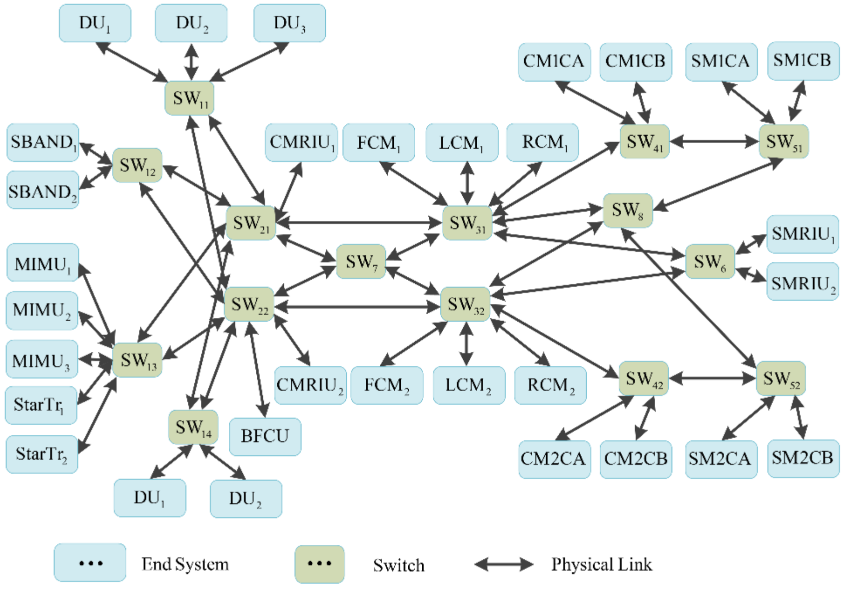


Figure 5

Evaluation – Combined Traffic Shapers (3)

- ## Realistic Test Cases – Orion CEV
  - Increase traffic load:
    TT traffic load in network
    → 15% on average & 54% in maximum.
    Overall traffic load in network
    → 25% on average & 69% in maximum.
  - 4 priorities, 25 flows of P1, 25 flows of P2, 24 flows of P3, 13 flows of P4;
  - IdleSlope ← actual bandwidth utilization,

$$idSl_i = operIdleSlope(P_i) \cdot \frac{OperCycleTime}{GateOpenTime}$$



Orion CEV

- ## Results
  - $100 \times ln(X), X = (WCD, WCB)$;
  - Sorted in increasing order by results (WCD, WCB).

Evaluation – Combined Traffic Shapers (3)

Comments

- End-to-end latency bounds:
  Backlog bounds:
  – ATS positive effect.
- ATS positive effect
  TAS+ATS+CBS > TAS+ATS+SP
  ← service capability AVB < SP.
- In combination of ATS
  → performance of SP & CBS get closer to each other.

TAS+SP vs. TAS+ATS +SP

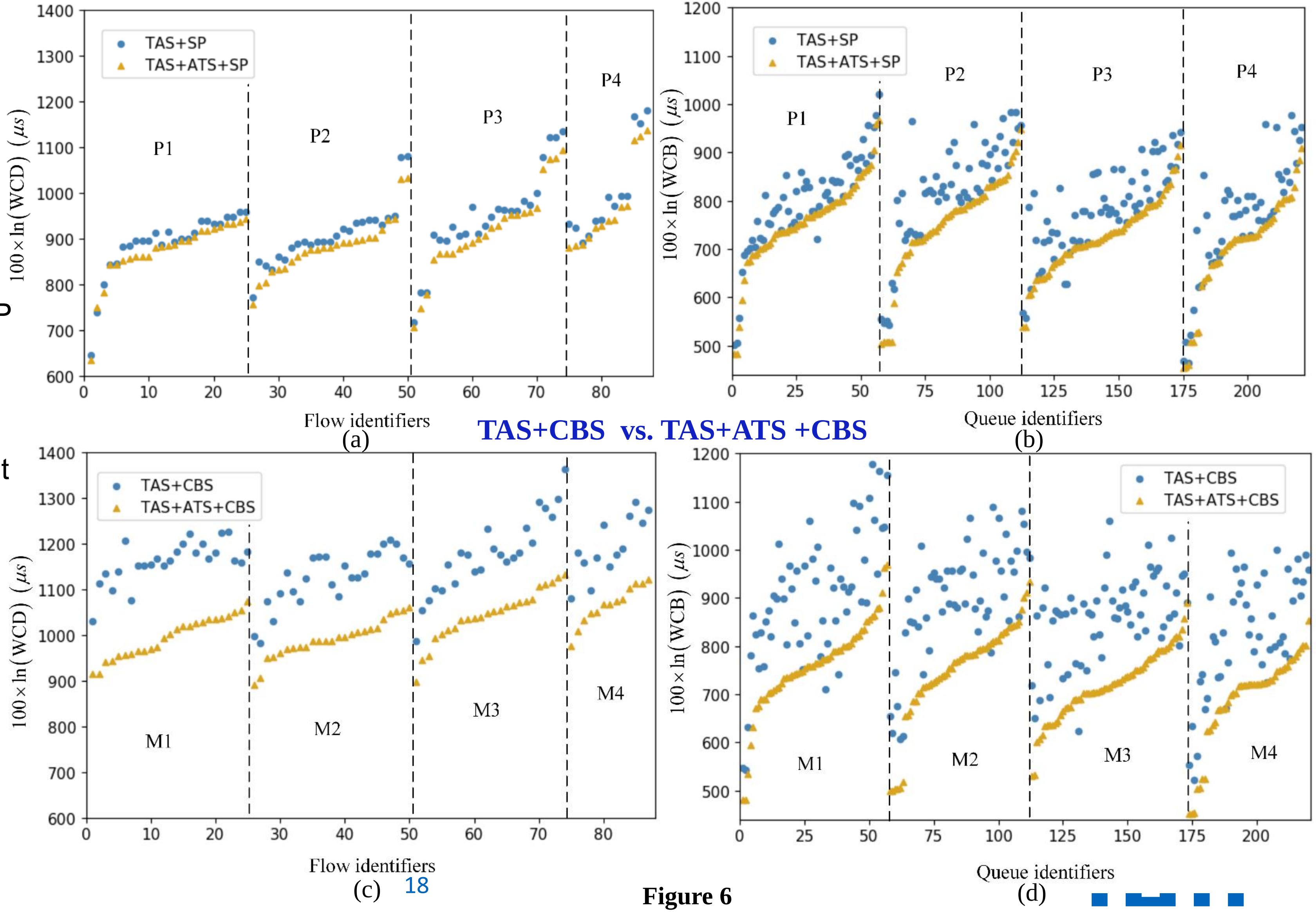


Figure 6

Conclusion

- SP vs. CBS
  - SP is more beneficial to the transmission delay of high-priority traffic, while CBS can specify bandwidth reservation for each priority traffic;
  - Due to the credit controlling by CBS, the long-term rate of AVB traffic arrival is reduced, backlog bounds of AVB traffic are possible lower than SP traffic.
- ATS vs. SP, CBS
  - ATS has limited advantages for high-priority traffic;
  - Only when the average traffic load of high priority traffic exceeds a certain value (around 70% in MM for example), ATS can show its superiority;
  - The positive effect of ATS on low priority traffic is more obvious.
- TAS vs. SP, CBS, ATS
  - TAS implement flow-based TT scheduling, has the highest performance with ultra low latency, jitter and backlog;
  - TAS requires the synthesis of optimized GCLs, to which is difficult to scale to large networks with many flows.
- TAS+ATS+X vs. TAS+X
  - Combined use of ATS with TAS will make ATS play a more active role, of which the effect is similar to the reshaping impact of ATS used individually on low priority traffic.
  - TAS will maintain unchanged its advantages of ultra low latency and jitter.

Bibliography

[1] S. S. Craciunas, R. S. Oliver, M. Chmelik, and W. Steiner.
Scheduling real-time communication in IEEE 802.1 Qbv time sensitive networks.
In 24th International Conf. on Real-Time Networks and Systems (RTNS), 2016.

[2] P. Pop, M. L. Raagaard, S. S. Craciunas, and W. Steiner.
Design optimisation of cyber-physical distributed systems using IEEE time-sensitive networks.
IET Cyber-Physical Systems: Theory \& Applications, 1(1), 2016.

[3] J. A. R. De Azua, and M. Boyer.
Complete modelling of AVB in network calculus framework.
In 22nd International Conf. on Real-Time Networks and Systems, 2014.

[4] J. Schmitt, P. Hurley, M. Hollick, and R. Steinmetz.
Per-flow guarantees under class-based priority queueing.
In IEEE Global Telecommunications Conference (GLOBECOM'03), 2003.

[5] J. Specht, and S. Samii.
Urgency-based scheduler for time-sensitive switched ethernet networks.
In 28th Euromicro Conf. on Real-Time Systems (ECRTS), 2016.

[6] J. Y. Le Boudec
A theory of traffic regulators for deterministic networks with application to interleaved regulators.
IEEE/ACM Transactions on Networking, 2018.

[7] L. Zhao, P. Pop, Q. Li, J. Chen, and H. Xiong,
Timing analysis of rate-constrained traffic in TTEthernet using network calculus.
Real-Time Systems, 53(2), 2017.

[8] L. Zhao, P. Pop, Z. Zheng, and Q. Li.
Timing analysis of AVB traffic in TSN networks using network calculus.
In IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS), 2018.

[9] L. Zhao, P. Pop, Z. Zheng, H. Daigmorte, and M. Boyer.
Latency Analysis of Multiple Classes of AVB Traffic in TSN with Standard Credit Behavior using Network Calculus.
IEEE Transactions on Industrial Electronics, 2020.

Thank you!