SCIENCE CHINA
Information Sciences

• **LETTER** •

September 2026, Vol. 69, Iss. 9, 190310:1–190310:2
https://doi.org/10.1007/s11432-025-5018-8

Special Topic: User-Centric Intelligent Networking

B-CUBE: a burst-aware 5G-TSN architecture and its large-scale mixed-flow scheduling

Yi XUE, Jie JIA*, Haoyang MAO, Jian CHEN, Wenlong HE & Xingwei WANG

School of Computer Science and Engineering, Northeastern University, Shenyang 110819, China

Received 10 October 2025/Revised 16 March 2026/Accepted 22 June 2026/Published online 26 August 2026

Citation Xue Y, Jia J, Mao H Y, et al. B-CUBE: a burst-aware 5G-TSN architecture and its large-scale mixed-flow scheduling. Sci China Inf Sci, 2026, 69(9): 190310, https://doi.org/10.1007/s11432-025-5018-8

The rapid development of the Industrial Internet of Things (IIoT) necessitates the integration of 5G and time-sensitive networking (5G-TSN) [1]. However, conventional 5G-TSN planar architectures are limited to single-mode (wired or wireless) access and transmission at each node, severely hindering networking flexibility and timely response to bursting flows. Meanwhile, cyclic queuing and forwarding (CQF) [2] and its enhanced variant, multi-CQF, struggle with unpredictable burst arrivals and latency tailing control, which impede deterministic and real-time transmission in large-scale mixed-flow scenarios. To address these issues, we propose B-CUBE, a novel burst-aware 5G-TSN architecture that equips each industrial node with co-deployed TSN and 5G system (5GS) bridges to enable wired, wireless, and dual-mode transmission, and designs a set of efficient cross-domain mixed-flow scheduling algorithms to achieve high scheduling success rates with low computational complexity for large-scale mixed flows in industrial scenarios.

Architecture of B-CUBE. As shown in Figure 1, B-CUBE departs from the conventional 5G-TSN architecture where 5GS bridges are only deployed at the edge, instead equipping each industrial node (e.g., industrial workshop or control center) with both a TSN bridge and a 5GS bridge, with industrial devices connecting to their respective local bridges via dual wired-wireless interfaces. This forms a dual-topology network with independent yet isomorphic wired and 5G domains that support collaborative resource scheduling, fundamentally breaking the single-mode access limitation of traditional architectures and enabling three flexible cross-domain transmission modes for data flows.

- Wired transmission. All data frames are completely forwarded by TSN bridges in the wired domain.
- Wireless transmission. All data frames are completely forwarded by 5GS bridges in the 5G domain.
- Dual-mode transmission. All data frames are forwarded collaboratively through the wired and 5G domains according to the splitting ratio.

For mixed-flow transmission, B-CUBE adopts a differentiated and burst-aware scheduling strategy tailored to the characteristics of time-triggered (TT) flows and bursting flows. TT flows, which typically transmit periodic real-time monitoring or control data, are flexibly configured via the traffic splitting mechanism. For example, as shown in Figure 1, a TT flow from node 1 can be split and transmitted synchronously over the 5G domain path ① and wired domain path ②, improving resource utilization and network capacity. Each TT flow $f_i \in \mathcal{F}$ is defined as

$$f_i = \{\text{src, dst, period, size, dd, offset, sr, path}\}, \tag{1}$$

where $f_i.\text{sr} \in [0,1]$ denotes the splitting ratio for TSN transmission ($f_i.\text{sr} = 1$ for pure wired transmission, $f_i.\text{sr} = 0$ for pure wireless transmission, and $0 < f_i.\text{sr} < 1$ for dual-mode transmission), and the transmission path of TT flow fragments traverses the TSN and 5GS bridges of the same nodes respectively. Bursting flows, which handle unexpected control or safety-related data with stricter end-to-end latency requirements, are assigned a fixed wireless transmission mode and cached in uplink burst-aware queues of 5GS bridges. For instance, as shown in Figure 1, bursting flows generated at node 3 are directly forwarded through the 5G domain to their destinations to ensure timely response. Each bursting flow $\bar{f}_i \in \bar{\mathcal{F}}$ is defined as

$$\bar{f}_i = \{\text{gt, src, dst, size, dd, path}\}, \tag{2}$$

where $\bar{f}_i.\text{gt}$ is the unknown generation time of $\bar{f}_i$ before occurrence.

Cross-domain mixed-flow scheduling. We formulate the cross-domain mixed-flow scheduling problem in B-CUBE as a bilevel optimization problem (P0), with the core objective of maximizing the age of information (AoI)-weighted number of schedulable flows under deterministic transmission constraints

$$\text{(P0)} \quad \max_{\mathcal{B}^k(t),\mathcal{F}.\text{offset},\mathcal{F}.\text{sr},\mathcal{S},\bar{\mathcal{S}}} O(\mathcal{S},\bar{\mathcal{S}}) \tag{3}$$

$$\text{s.t. } \{\mathcal{F}.\text{offset}, \ \mathcal{F}.\text{sr}\} \in \mathop{\arg\max}_{\mathcal{F}.\text{offset}, \ \mathcal{F}.\text{sr}} \big\{O(\mathcal{F})|c1, \ c2, \ c3\big\},$$

$$S_i \in \{0,1\}, \ \bar{S}_i \in \{0,1\}, \tag{3a}$$

$$\sum_{m\in\mathcal{M}^k} B_m^k(t) \leqslant B^{\text{total}}, \tag{3b}$$

$$B_m^k(t) \geqslant B^{\min}, \tag{3c}$$

where $O(\mathcal{S},\bar{\mathcal{S}}) = \sum_{t\in\mathcal{T}}\sum_{\hat{b}_k\in\mathcal{B}_{5\text{GS}}}\sum_{m\in\mathcal{M}^k}(\sum_{S_i\in\mathcal{S}_m^k(t)} S_i + \alpha_m^k(t)\times\sum_{\bar{S}_i\in\bar{\mathcal{S}}_m^k(t)}\bar{S}_i)$, $O(\mathcal{F}) = \sum_{i=1}^{N^F}\big(f_i.\text{period}\times f_i.\text{size}\times f_i.\text{sr}\big)$, $c1: \ 0 \leqslant f_i.\text{offset} \leqslant \text{offset}_i^{\max}$, $c2: \ \sum_{i=1}^{N^F}\Big(f_i.\text{sr}\times f_i.\text{size}\times A_k^e(f_i,t)\Big) \leqslant Q^{\text{CQF}}$, $c3: \ 0 \leqslant f_i.\text{sr} \leqslant 1$. Considering the black

* Corresponding author (email: jiajie@mail.neu.edu.cn)

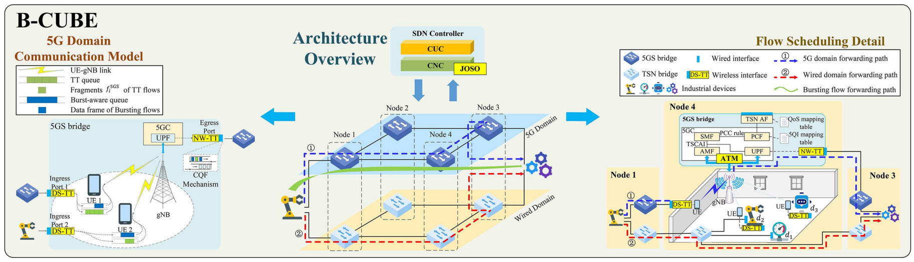

Figure 1 (Color online) Architecture of B-CUBE.

box of wireless scheduling and the fact that no linking variables are included in problem (P0) [3], we decouple the bilevel problem into wired and wireless scheduling subproblems. For more details on the problem definitions, please see Appendixes A and B.

Joint offset and splitting optimization. For wired scheduling, we propose the joint offset and splitting optimization (JOSO) algorithm, which targets the mixed-integer nonlinear programming (MINLP) nature of the wired scheduling subproblem, a key challenge in 5G-TSN mixed-flow scheduling due to its high computational complexity and non-convexity. To address this, JOSO unifies TT flow offsets and splitting ratios into a joint optimization vector $\boldsymbol{x} \in \mathbb{R}^{N\times 1}$, where $N = \sum_{i=1}^{N^F}(\text{offset}_i^{\max}+1)$. The dimension of $\boldsymbol{x}$ represents all feasible offset scheduling strategies, and each element value x_n of $\boldsymbol{x}$ represents the splitting rule on the corresponding offset setting, then the f_i.offset and f_i.sr for each TT flow can be determined by a specific subset $\boldsymbol{x}_i$ of $\boldsymbol{x}$. To guarantee the uniqueness of solutions, we introduce an l_0-norm constraint for the elements of x_i, which can be given by

$$\sum_{x_n \in \boldsymbol{x}_i} ||x_n||_0 \leqslant 1, \tag{4}$$

which can equivalently replace constraints $c1$ and $c3$. Since the l_0-norm is non-smooth and non-convex, we adopt smoothed l_0-norm (SL0) [4] approximation to render it tractable, which can be given by $g_\delta(x_n) \triangleq 1-\exp(-\frac{x_n}{\delta})$ $(\forall x_n \in \boldsymbol{x})$, where $\delta > 0$ is a parameter that controls the smoothness of the approximation. Then, constraint (4) can be relaxed as $\sum_{x_n\in\boldsymbol{x}_i} g_\delta(x_n) \leqslant 1$. We further apply the majorization-minimization (MM) method to construct a convex upper bound of $g_\delta(x_n)$ via first-order Taylor expansion, which can be given by $\hat{g}_\delta(x_n) \triangleq g_\delta(x_n^{(l-1)}) + \frac{1}{\delta}\exp(-\frac{x_n^{(l-1)}}{\delta})(x_n - x_n^{(l-1)})$. Therefore, constraint (4) can be reformulated as

$$\sum_{x_n \in \boldsymbol{x}_i} \hat{g}_\delta(x_n) \leqslant 1. \tag{5}$$

This transformation converts the original non-convex MINLP problem into a solvable convex form, reducing computational complexity to enable efficient scheduling in large-scale scenarios.

AoI-weighted transmittable flow ratio maximization. For online wireless scheduling, we propose the AoI-weighted transmittable flow ratio maximization (ATM) algorithm, which is designed to address the real-time nature of wireless scheduling, where bursty traffic with strict latency requirements and dynamic wireless channel conditions demand fast, adaptive resource allocation. Its core innovation lies in resolving two critical challenges: binary scheduling state constraints (which introduce combinatorial complexity) and the need to prioritize bursting flows. To eliminate integer constraints of binary scheduling states, ATM maps discrete states to continuous transmittable flow ratios $a_m^k(t) \in [0,1]$ for TT flows and $\bar{a}_m^k(t) \in [0,1]$ for bursting flows, respectively, simplifying the problem into a convex optimization form that can be solved in real time. Consequently, the problem can be transformed into an equivalent convex form, with its Lagrangian function defined as shown in (6); then we can obtain the optimal solution via KKT conditions

$$\begin{aligned}\mathcal{L} = \sum_{m=1}^{2M^k} & n_m \times a_m + \beta_m(1-a_m) + \lambda_m a_m \\ & + \xi_m\left(B_m^k(t) - e_m \times a_m - e_{m'} \times a_{m'}\right) \\ & + \rho\left(B^{\text{total}} - \textstyle\sum_{m\in\mathcal{M}^k} B_m^k(t)\right).\end{aligned} \tag{6}$$

Experiments. Experiments on three industrial topologies (ring, linear, mixed) show B-CUBE achieves 92.39% average scheduling success rate (53.20% higher than conventional CQF), 1.71% bursting flow frame loss rate, and 250 µs jitter boundary. JOSO and ATM outperform baselines in success rate and time cost, ensuring efficient large-scale mixed-flow transmission. For more details on the experiments, please see Appendix D.

Conclusion. B-CUBE, featuring co-deployed TSN and 5GS bridges per industrial node, breaks the single-mode limitation of traditional 5G-TSN architectures, supporting wired, wireless, and dual-mode transmission. It adopts differentiated scheduling: TT flows achieve cross-domain load balancing via traffic splitting, while bursting flows use fixed wireless transmission for real-time response, compatible with 3GPP and IEEE 802.1 standards. Formulating cross-domain scheduling as a bilevel optimization problem, JOSO reduces complexity by converting non-convex MINLP to convex form, and ATM enables real-time wireless resource allocation with AoI prioritization. Experiments validate B-CUBEs superior performance in large-scale mixed-flow scheduling, lowering burst frame loss and latency.

Acknowledgements This work was supported in part by National Natural Science Foundation of China (Grants Nos. 62132004, 62572113, 62572109), Aeronautical Science Foundation of China (Grant No. 20230026050001), Central Government Guided Local Science and Technology Development Fund Project of Liaoning (Grant No. 2025JH6/101100009), Young and Middle-Aged Leading Talents in Technological Innovation of Shenyang (Grant No. RC231173), and Shenyang Science and Technology Plan Fund Project (Grant No. 23-503-6-17).

Supporting information Appendixes A–E. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Val I, Seijo S, Torrego R, et al. IEEE 802.1AS clock synchronization performance evaluation of an integrated wired-wireless TSN architecture. IEEE Trans Ind Inf, 2022, 18: 2986–2999
2 IEEE. Standard for local and metropolitan area networks-media access control (MAC) bridges and virtual bridged local area networks amendment: cyclic queuing and forwarding. IEEE 802.1Qch Std, 2017. https://1.ieee802.org/tsn/802-1qch/
3 Kleinert T, Labbé M, Ljubić I, et al. A survey on mixed-integer programming techniques in bilevel optimization. EURO J Comput Optim, 2021, 9: 100007
4 Tao M, Chen E, Zhou H, et al. Content-centric sparse multicast beamforming for cache-enabled cloud RAN. IEEE Trans Wireless Commun, 2016, 15: 6118–6131