Yannan YUAN, Fei QIN, Jiankang LIU, Yuanyuan WANG, Jianan CAI, Xiang PAN, Dajie JIANG, 2025. A unified data collection framework based on the data plane for 6G. *Frontiers of Information Technology & Electronic Engineering*, 26(2):293-300.
https://doi.org/10.1631/FITEE.2400247

A unified data collection framework based on the data plane for 6G

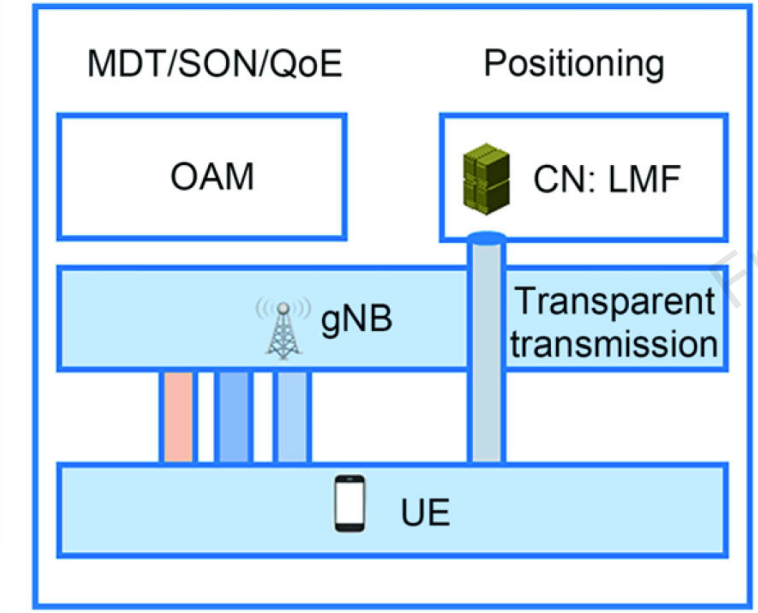
Key words: 6G unified data collection framework; Data plane; Protocol stack; Two-sided data collection mode; Digital twin network

Yannan Yuan
E-mail: yannan.yuan@vivo.com
ORCID: https://orcid.org/0000-0003-1892-612X

Motivation and gap analysis

Data collection methods for 5G

- Fragmented solutions for different use cases, leading to high standardization overheads and UE/NF duplicate data collection
- One-sided data collection mode results in a lack of data providers

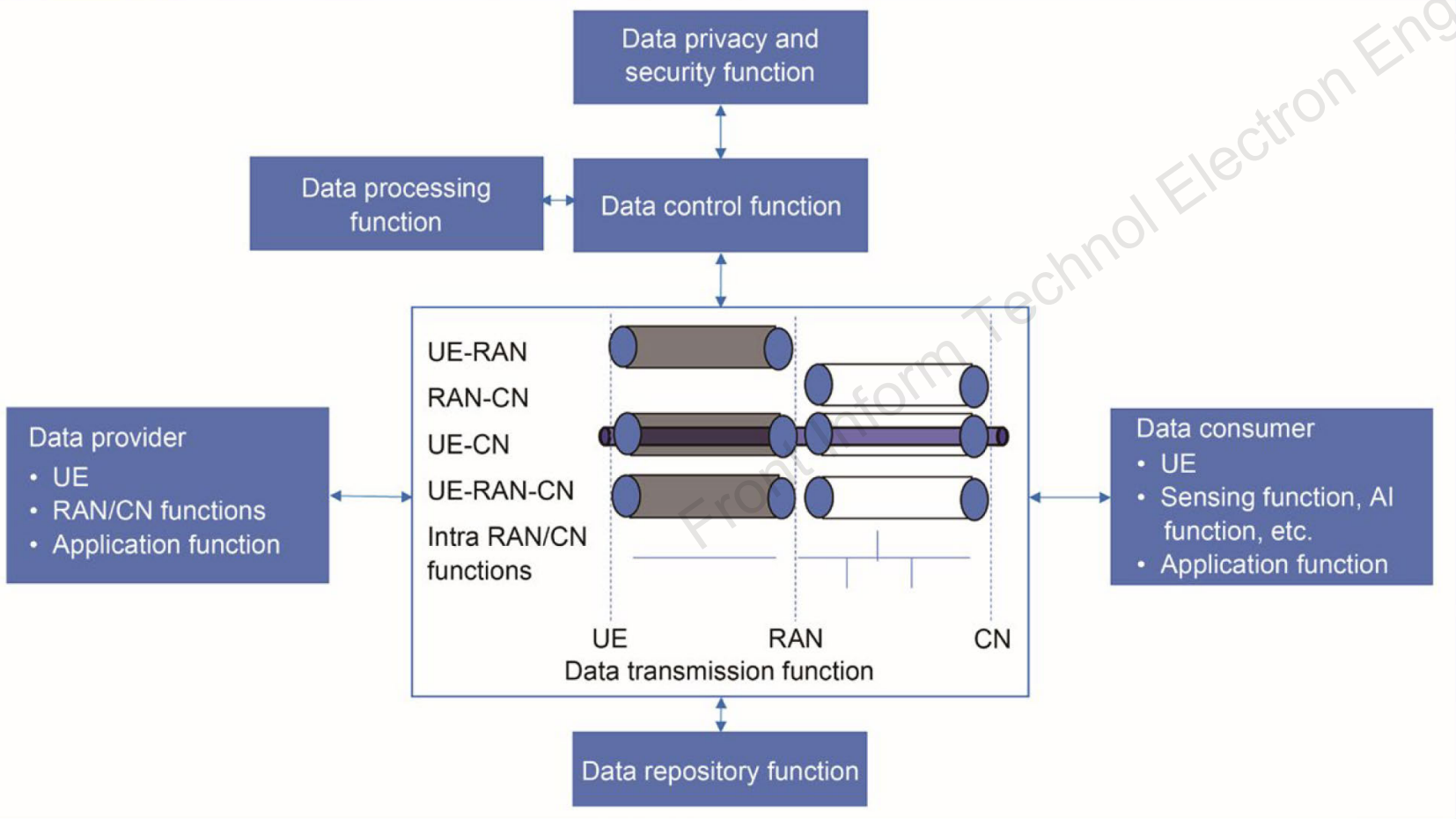


New requirements of 6G

- New data collection requirements arise from 6G, such as sensing data and AI data, which are characterized by large data volume and flexible termination nodes.
- Real-time data collection is one of the key elements of digital twin networks.
- The data collection methods based on CP are not suitable for collecting a large amount of data, such as those needed by 6G native AI, ISAC, and DT networks.
- The efficiency of data collection methods based on CP also needs to be improved, e.g., data transmission.

Architecture of the data plane for 6G

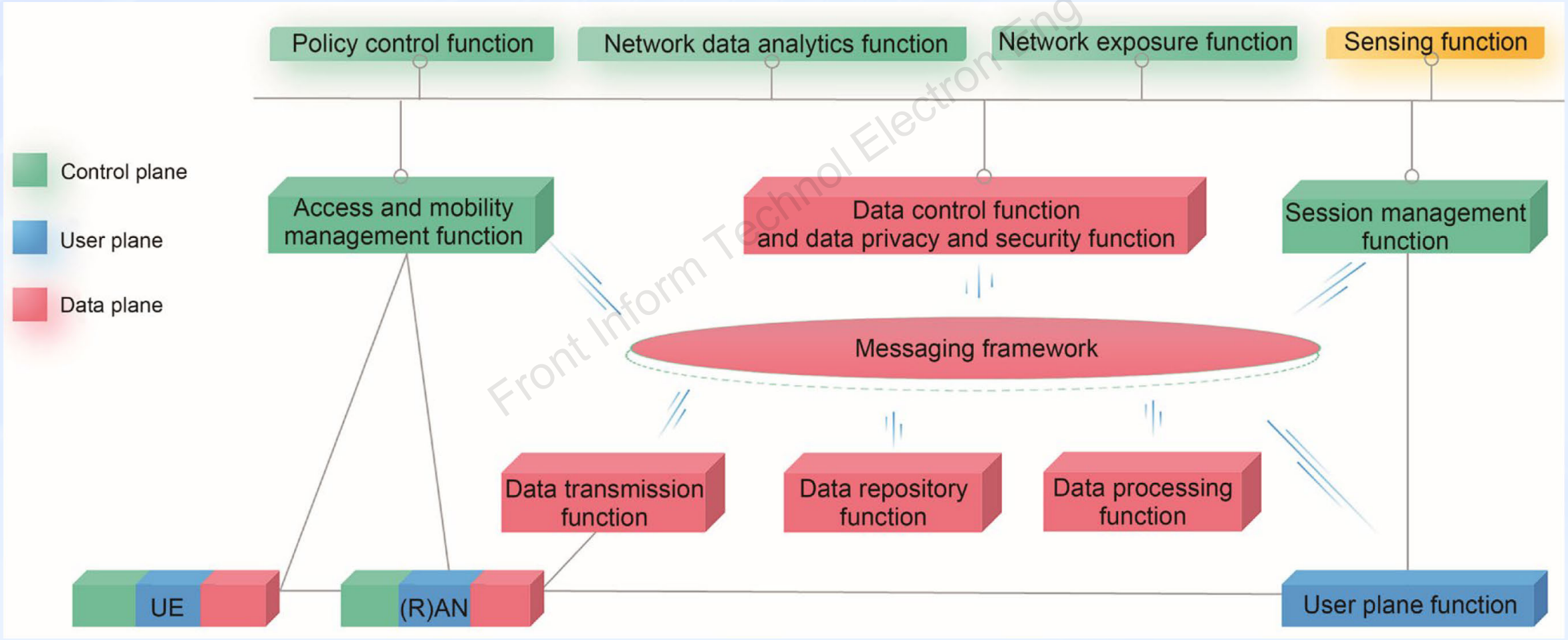
The logical functions of 6G DP include a data control function (DCF), data privacy and security function (DPSF), data repository function (DRF), data transmission function (DTF), data processing function (DProcF), data provider (DPro), and data consumer (DCons).



DCF is used to support data collection coordination and to configure data service, data transmission (e.g., establishing, modifying, and releasing the bearer of DP), and data processing.

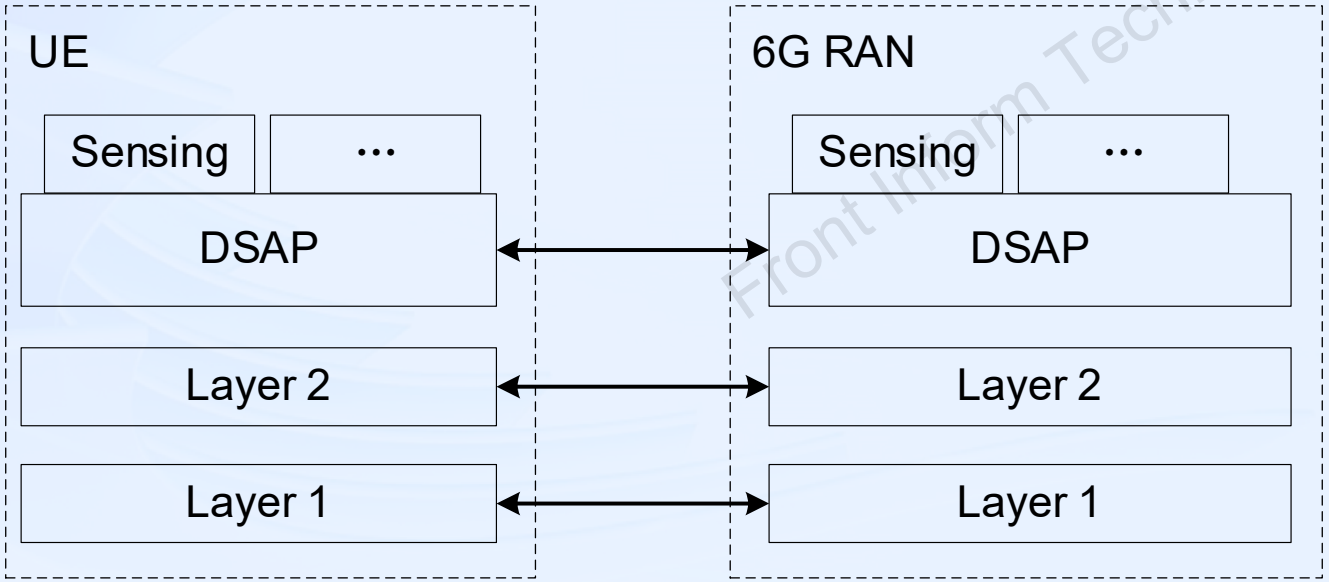
6G system architecture with the data plane

- An example of the 6G system architecture introduces the data functions. Enhancements (e.g., messaging framework) are proposed to support the transmission of a large amount of data within CN.
- RAN and UE need to introduce peer-to-peer data functions to support the new features of 6G.



Protocol stack of the data plane for 6G

- The DP protocol stack for the air interface includes a data service application protocol (DSAP) that comprises DCF, DPSF, DRF, DTF, and DProcF. DSAP establishes appropriate DP radio bearers (DPRBs) through RRC based on data collection requirements.
- DPRBs have flexible priority characteristics. The priority can be set according to the data volume, quality of data transmission, and traffic load of the user plane (UP). This makes DPRBs more suitable for collecting a large amount of data compared to signaling radio bearers.



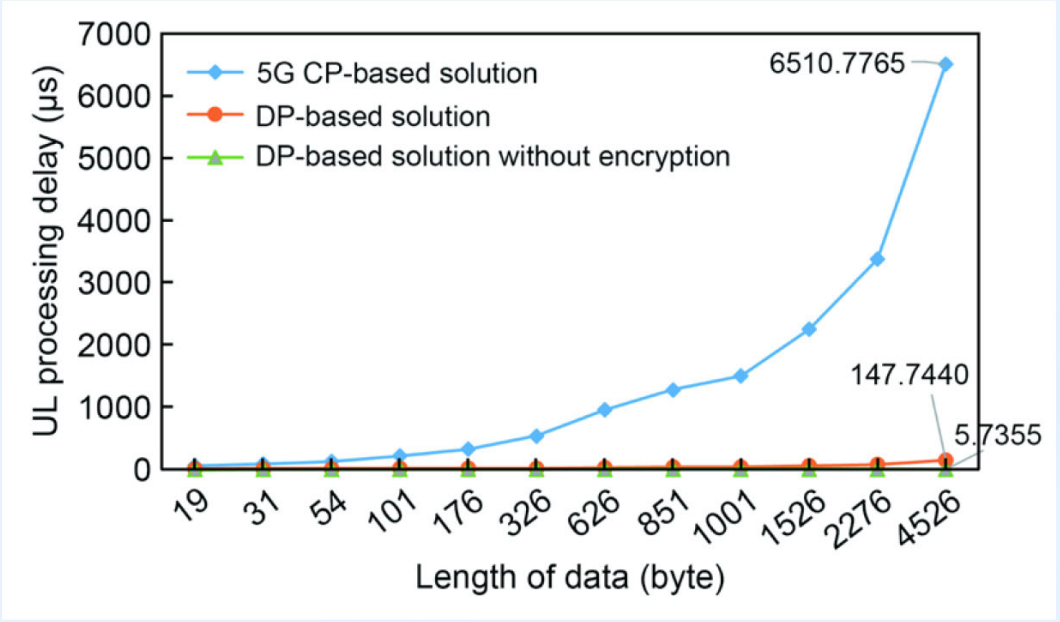
- The DP protocol stack achieves efficient data transmission through optimizations at DSAP, L2, and L1.
- For DSAP, data compression is supported to reduce data volume.
- For L2, simplified protocol functions are used, such as unified data storage, to reduce header overhead and on-demand encryption/decryption.
- For L1, efficient physical layer technologies like joint source–channel coding and semantic communication are adopted to improve spectral efficiency.

Test and performance analysis

- The efficiency of the DP protocol stack is verified using a 6G UE prototype. The UL processing delay is tested under different data lengths and user plane traffic loads. Results show that the DP-based solution has significantly lower processing delay compared to the existing CP-based solution.
- For example, when the data length is 19 bytes, the average UL processing delay of DP is about 22% of the CP-based solution. As the data length increases, the efficiency improvement of the DP protocol stack becomes more significant.
- The processing time of ASN.1 and ProtoBuf is also tested. ProtoBuf shows shorter processing time and is more suitable for a large amount of data, while ASN.1 is suitable for short messages with a small data volume.

Table 1 Parameters of the UE prototype

Parameter	Value/Description
CPU (@frequency)	Arm A53 (@1 GHz)
Memory	4 GB
ASN.1 tool/version	Objective/V773
ProtoBuf tool/version	protoc/3.6.1
Encryption algorithm	AES128



Two-sided data collection mode

Increasing Data Provider Willingness

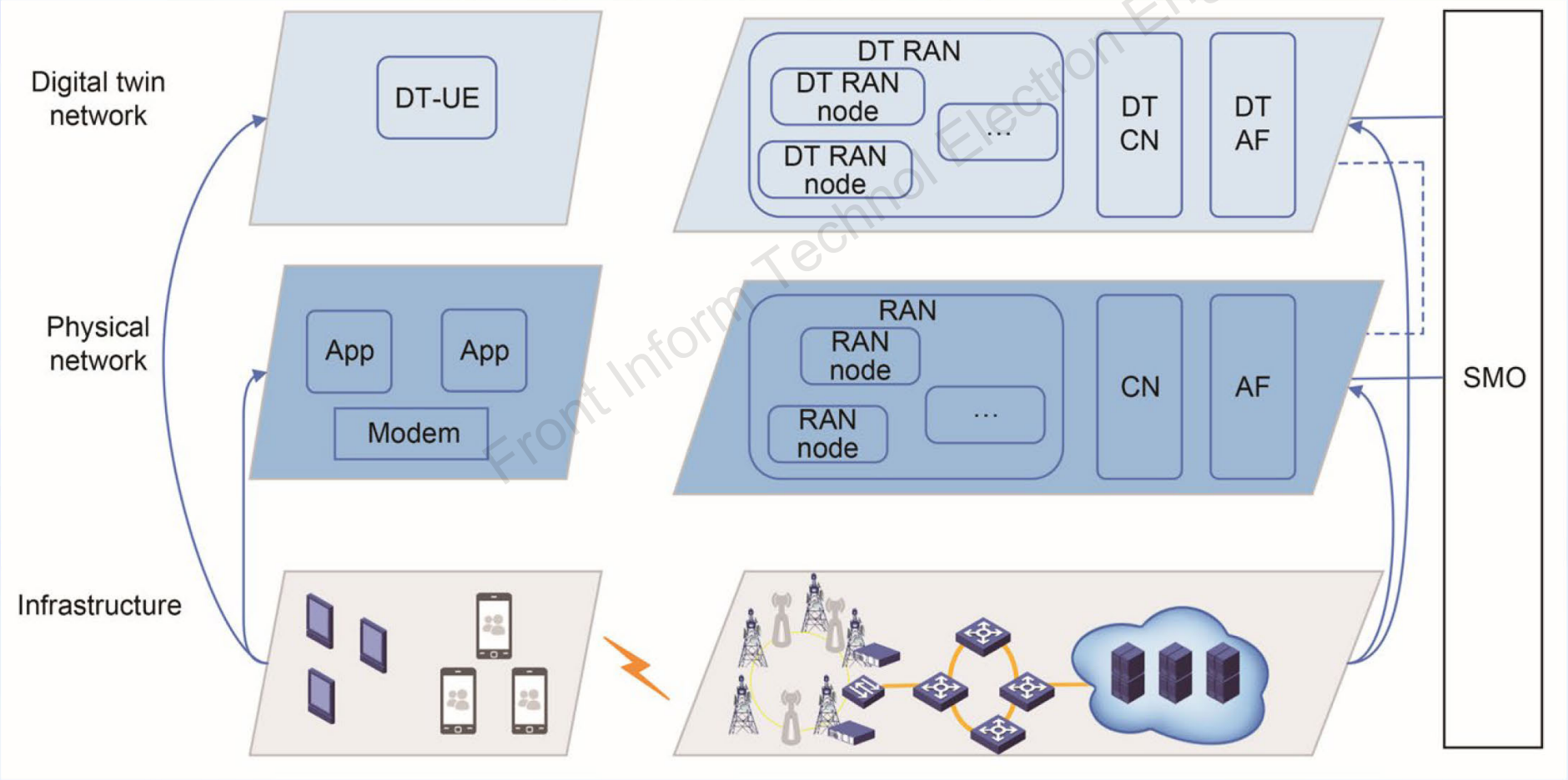
- The unified data collection framework supports both one-sided and two-sided data collection modes.
- The proposed two-sided data collection mode allows both data collection nodes and data providers to obtain the required data or rewards. Potential rewards include the number of times to obtain network data, quality of service (QoS) guarantee and priority promotion, data quota, minutes of voice, etc.
- For example, in data collection between RAN and UE, the network can indicate the data that UE can obtain during the collection process. This encourages UE to provide user consent and accept the data request.
- The two-sided mode also reduces signaling message overhead and uses only one DPRB.

Enabling Digital Twin Networks

- The two-sided data collection mode enables the possibility of DT UE. UE can provide user consent to be used as DT UE based on usage habits.
- DT UE can reuse a large amount of UE in the physical network to reduce the infrastructure resource overhead of the DT network. This allows for high performance in terms of similarity and prediction accuracy.
- The privacy and security of DT UE are also considered, with isolation between DT UE and physical UE to support decoupling in aspects such as UE identifier, location information, and context information.

A digital twin network solution using the existing infrastructure of the physical network

- The network can configure DT UE on the physical UE, creating an end-to-end DT network based on the existing infrastructure.



Conclusions and future work

Values of the Unified Data Collection Framework

- The proposed unified data collection framework based on the data plane provides a solution to meet the ubiquitous data collection requirements of 6G.
- The DP protocol stack, consisting of DSAP, simplified L2, and optimized L1, has been validated to reduce processing time and improve efficiency for a large amount of data.
- The two-sided data collection mode increases the willingness of data providers to provide data, enabling data applications such as sensing, AI, and digital twin networks.

Future Research Directions

- Some of the rewards (e.g., data or voice quotas) in the two-sided data collection mode may involve billing systems and business models that need to be further validated and refined in commercial networks.
- We will continue to simulate and analyze the performance of joint source–channel coding and semantic communication for 6G data collection.