



A Real-Time Archiving Framework from EPICS to Time-Series Databases for Fusion Plant Data

Guang Yang,

guangyang@mail.ustc.edu.cn

25th IEEE Real Time Conference - La Biodola, Elba, Italy

25-29 May, 2026



EAST

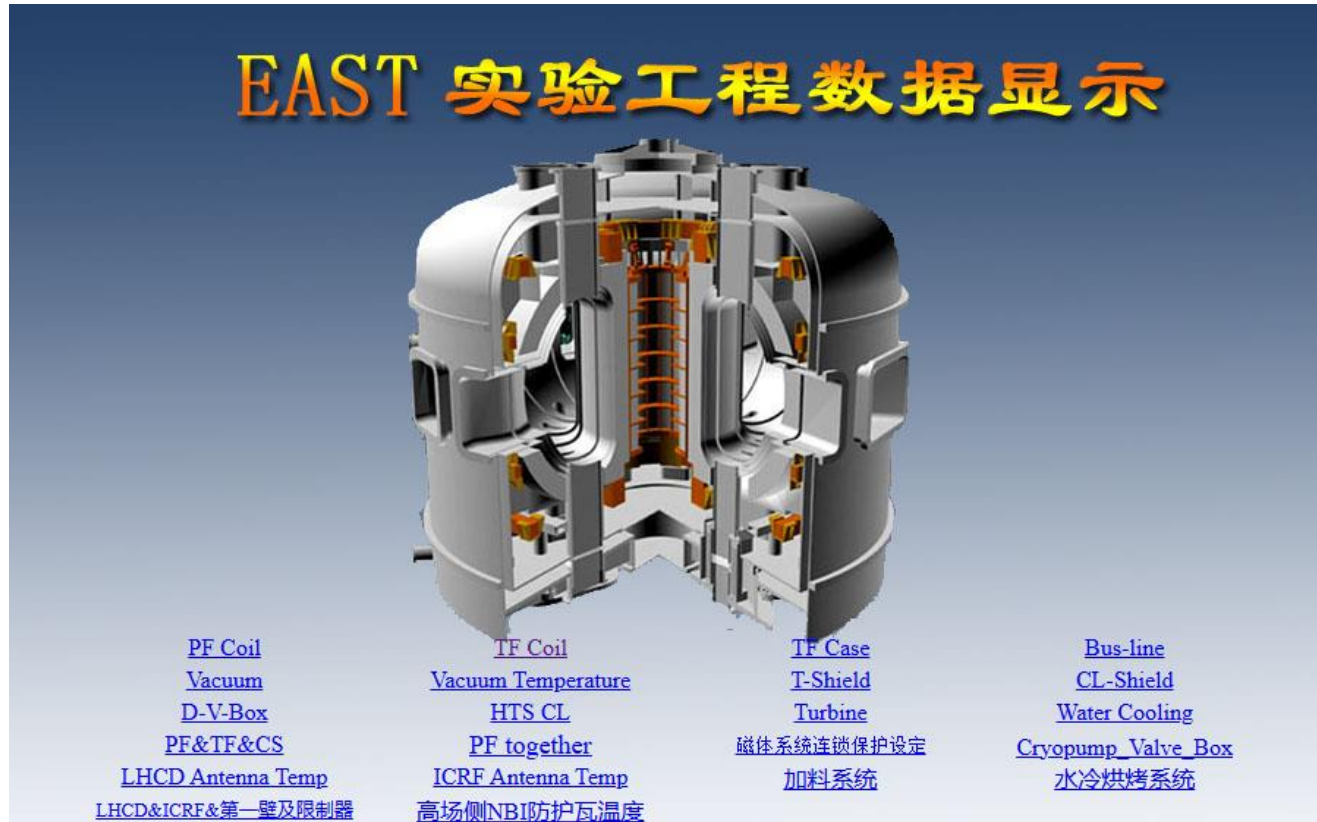
outline

- 1. Introduction**
- 2. Design of the framework**
- 3. Implementation of the system**
- 4. Test**
- 5. Summary and plan**

Introduction

Background

EAST rely on large-scale distributed control systems to supervise subsystem with the continuous growth of plant data.



EAST Plant Data System

Image from: <https://epics-controls.org>



ITER Fusion Device

ITER ("The Way" in Latin) is one of the most ambitious energy projects in the...



China Spallation Neutron Source (CSNS)

The China Spallation Neutron Source (CSNS), operated by the Institute of High Energy Physics



KSTAR Fusion Device

KSTAR (Korean Superconducting Tokamak Advanced Research), the most recent among global tokamak facilities, is a...



Beijing Electron Positron Collider (BEPC-II)

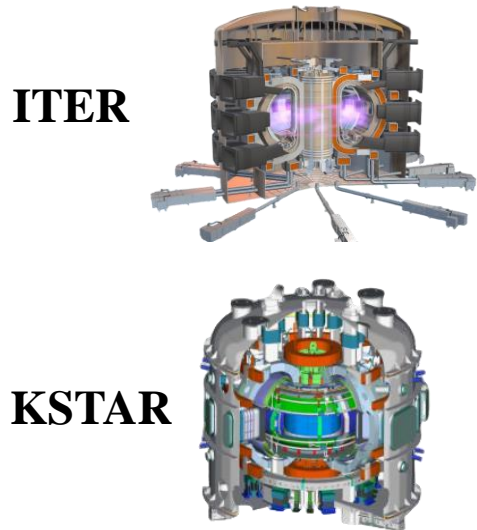
BEPC-II is the upgrade project of Beijing Electron Positron Collider (BEPC), located at the Institute...

EPICS Widely used in: EAST, ITER, KSTAR, China Spallation Neutron Source (CSNS), Beijing Electron Positron Collider (BEPC-II).

Introduction

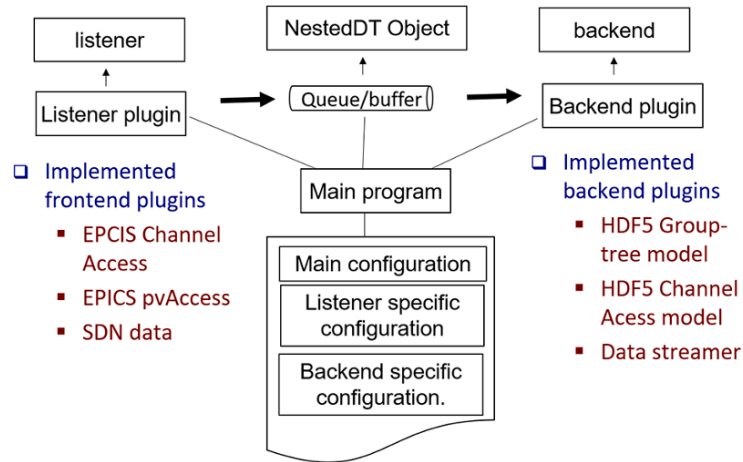
Work Motivation and Existing Archivers

The volume, velocity, and diversity of EAST plant data will continue to grow as more equipment is instrumented and higher sampling rates are adopted in future tokamaks.

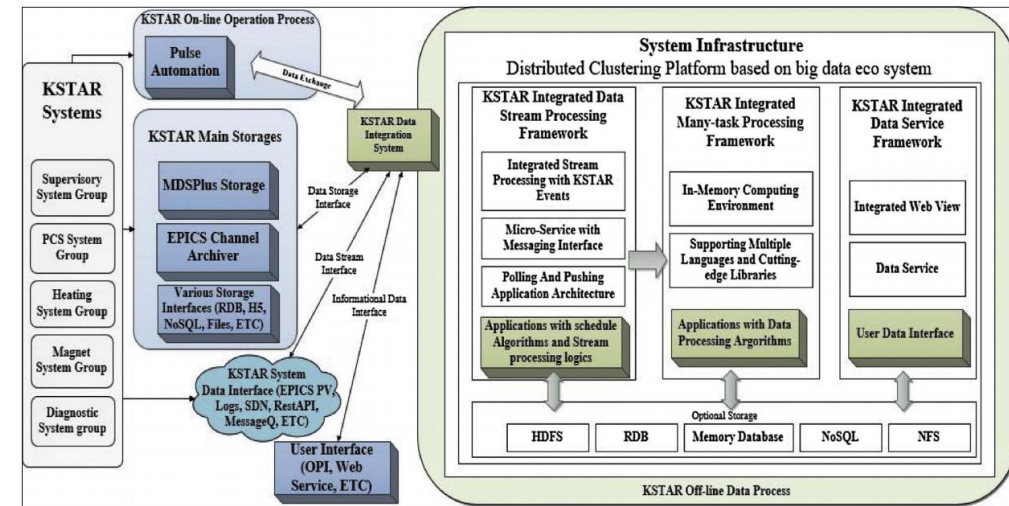


ITER

KSTAR



ITER Data archiver service architecture



KSTAR Data System

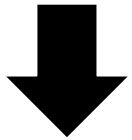
Varying data scales, access patterns, existing ecosystems, and engineering decisions across different devices call for distinct solutions.

[1]Castro et al., 2023 R. Castro, Y. Makushok, L. Abadie, B. Bauvir, A. Neto, J. Vega, Storing EPICS process variables in HDF5 files for ITER, Fusion Engineering and Design
 [2]Tak et al., 2018 T. Tak, J. Hong, K. Park, W. Lee, T. Lee, H. Han, G. Kwon, J. Park, Conceptual design of new data integration and process system for KSTAR data scheduling, Fusion Engineering and Design.

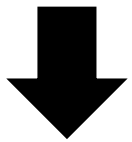
Introduction

Work Targets

Future fusion devices



Fusion plant data



TimeSeries DataBase

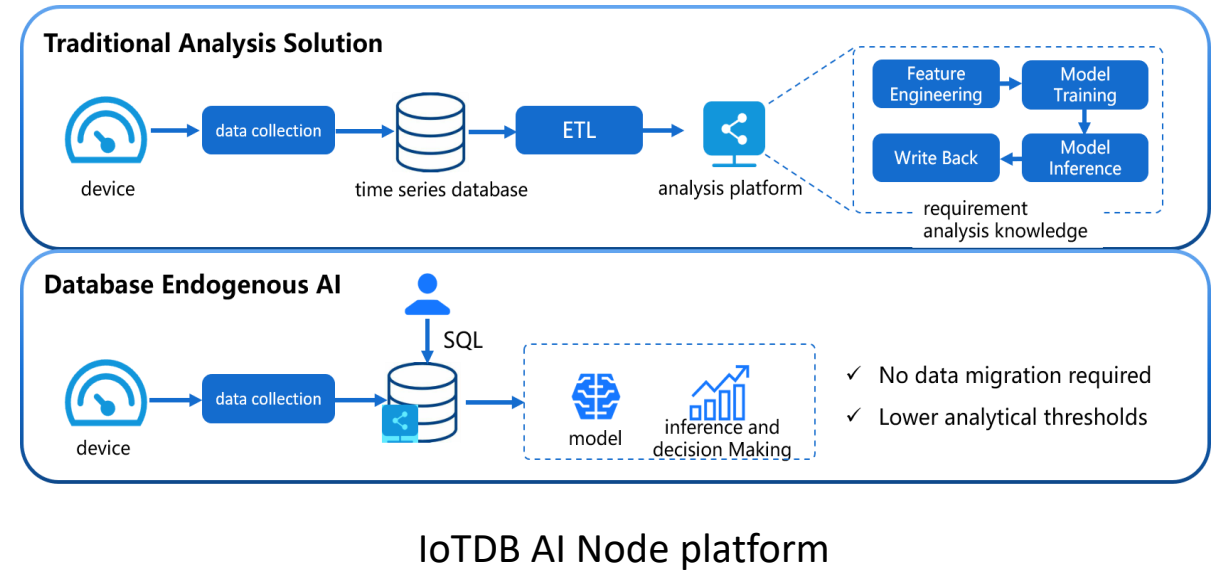
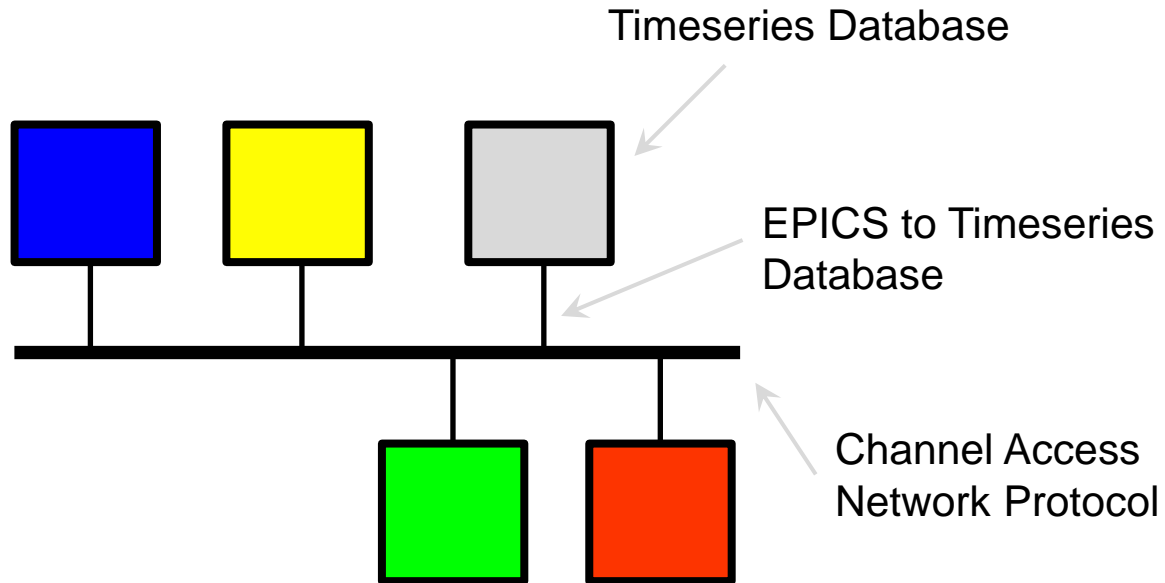
- Stronger large-scale query performance
- Integrated AI training platform.

- Generated by sensors and control systems
- Orderd by timestamp

- Purpose-built for high-throughput, time-ordered data.
- Native support for: column store, compression, down-sampling, range query, distributed clusters.
- Representative systems: InfluxDB, TimescaleDB, Apache IoTDB.

Design of the framework

From EPICS to Timeseries Database

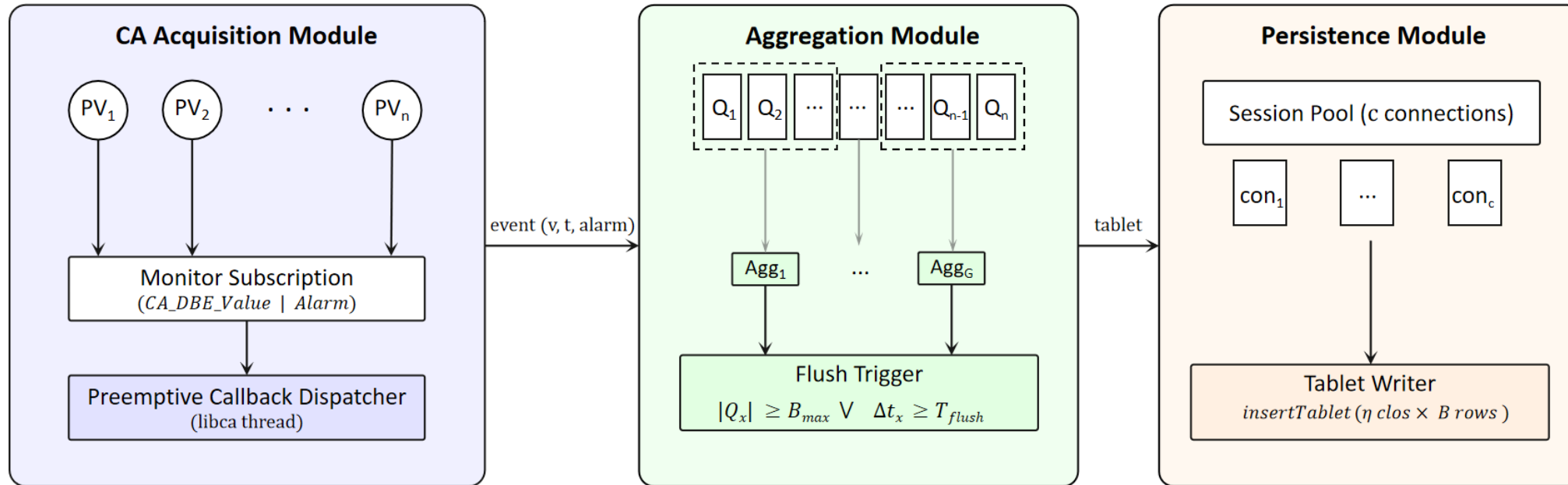


- EPICS side: PVs, distributed IOCs, Channel Access network.
- TSDB side: column-oriented, multi-device range queries, AI integration.
- Gap: Channel Access is event-driven; TSDB is batch-oriented.
- Need a bridge that converts CA monitor events into batched Tablet writes — efficiently, reliably, and at scale.

C. Wang, X. Huang, J. Qiao et al., "Apache IoTDB: Time-series database for internet of things," Proceedings of the VLDB Endowment, vol. 13, no. 12, pp. 2901–2904, 2020

Design of the framework

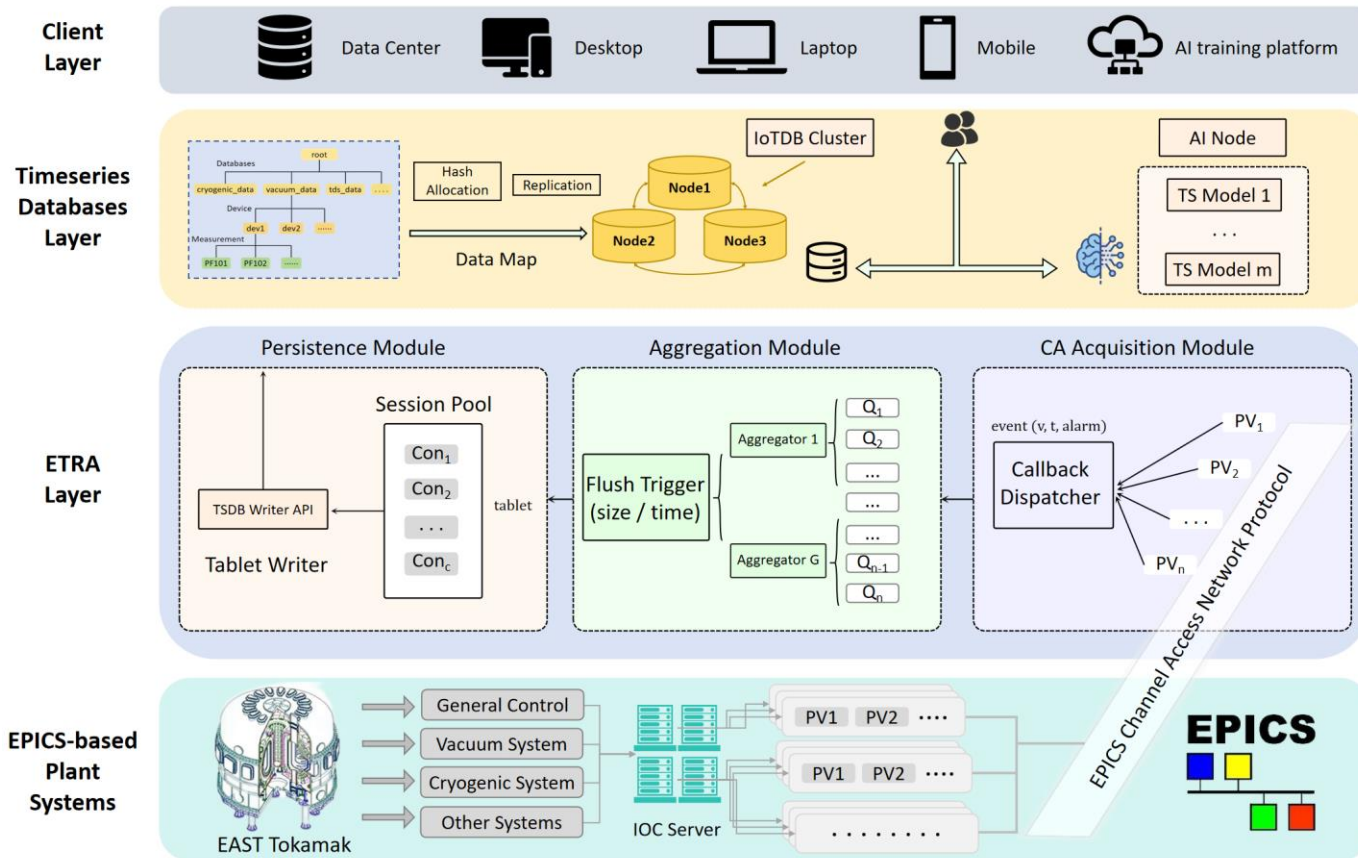
EPICS To Timeseries Database Architecture (ETRA)



- CA Acquisition Module — $libca$ preemptive callback dispatcher; one event channel per PV.
- Aggregation Module — PVs partitioned into G groups; flush trigger by size or time.
- Persistence Module — Session pool of c connections; batched Tablet write.
- Pipeline: event \rightarrow channel \rightarrow aggregator \rightarrow Tablet \rightarrow TSDB.

Implementation of the system

The EAST Plant system based on ETRA



- EPICS plant systems: EAST tokamak, vacuum, cryogenic, general control; IOC servers; PVs published via EPICS Channel Access Network Protocol.

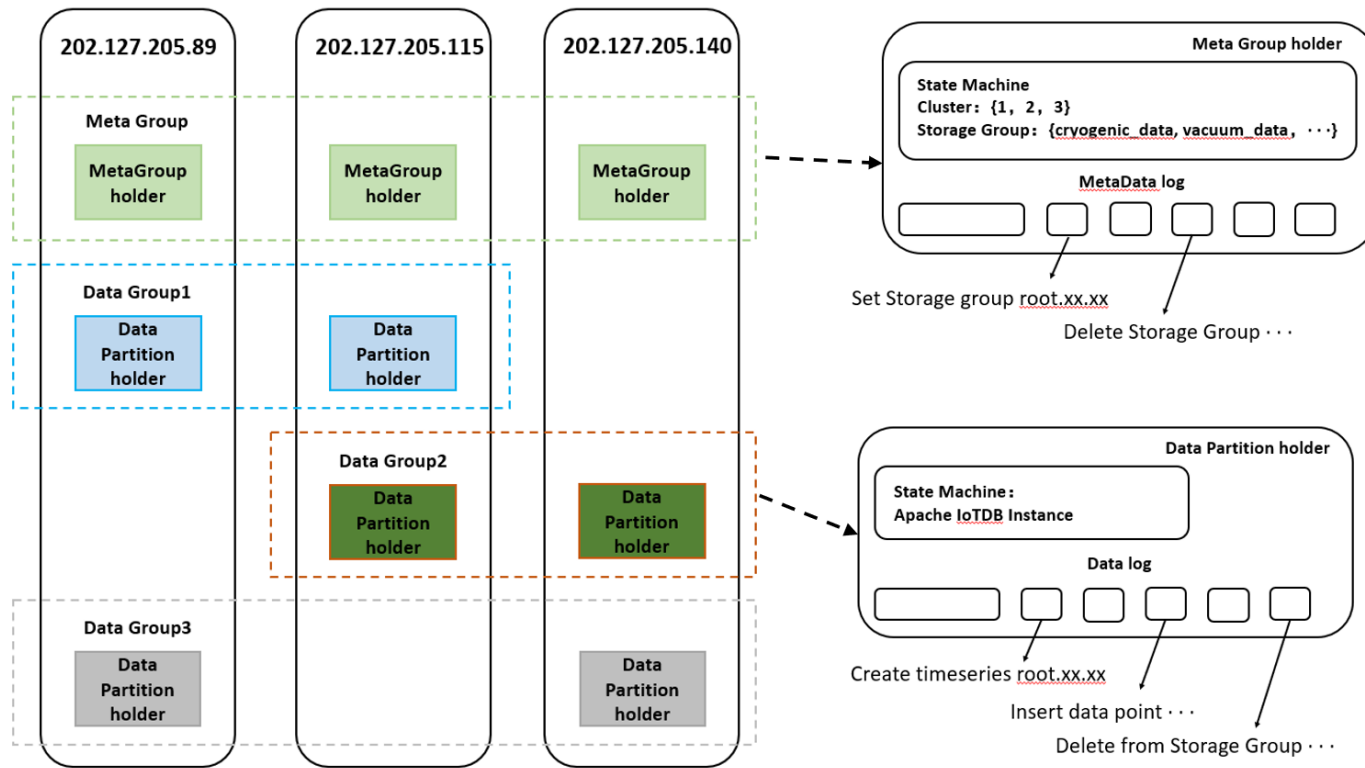
- ETRA layer: CA Acquisition, Aggregation, Persistence modules.

- TSDB layer: IoTDB cluster with hash allocation and replication; AINode for ML inference.

- Client layer: data center, desktop, mobile, AI training platform.

Implementation of the system

The EAST distribute data cluster based on IoTDB



Timeseries Database cluster Logic Design

- Metadata:

Each system's metadata is kept in three copies across different servers.

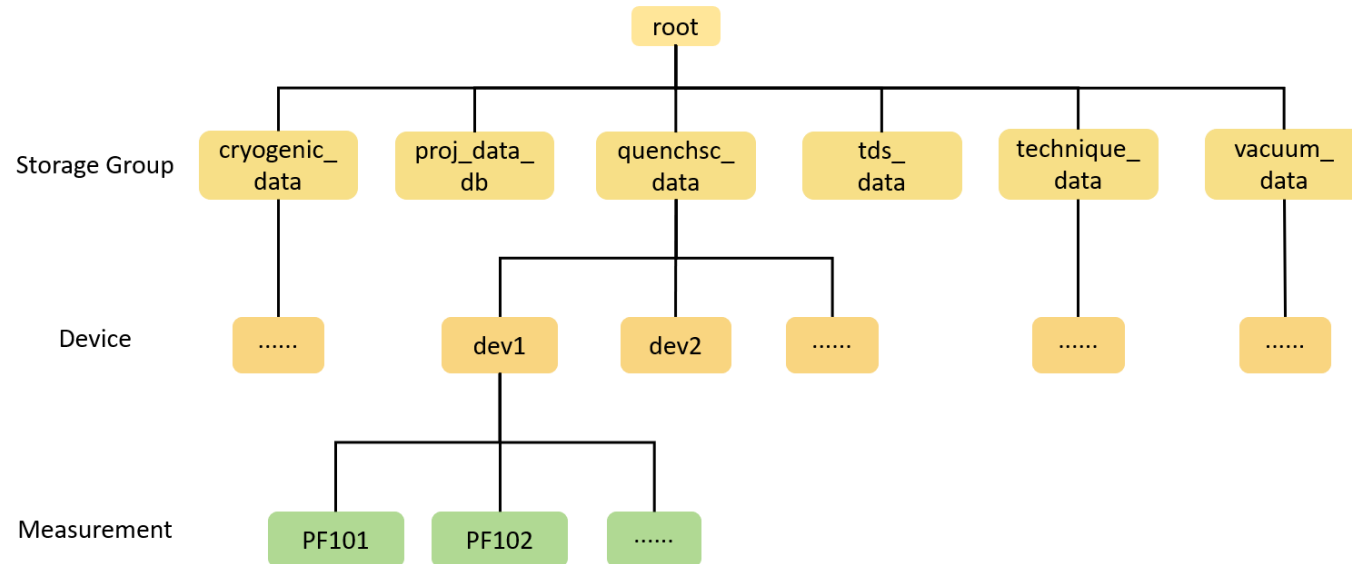
- Plant Data:

Split the EAST plant data into three parts and store them on different servers.

Copy the EAST plant data, store it on different servers according to the principle of hash allocation.

Implementation of the system

Data mapping



Data mapping From EAST Plant data to IoTDB

- System: The system in the EAST plant corresponds to the storage group level in IoTDB, which is the first level under the root node.

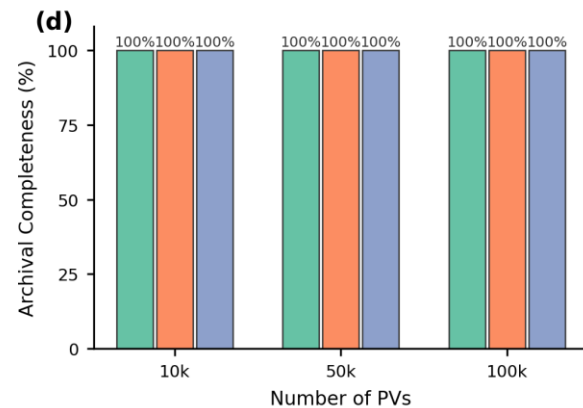
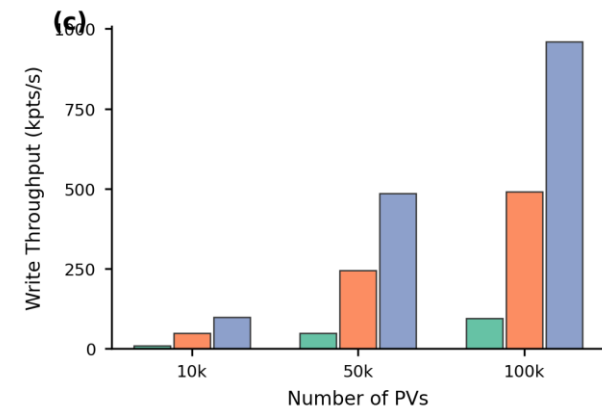
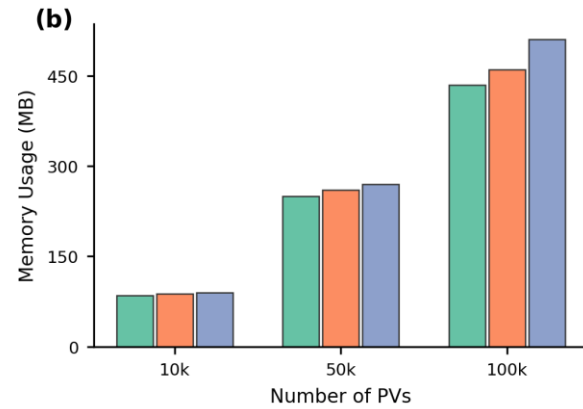
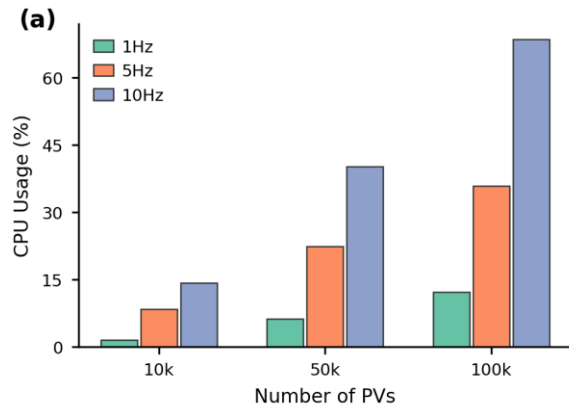
- Device: The devices in the EAST plant corresponds to the device level in IoTDB, each device contains 100 sensor (IoTDB measurement).

- Sensor : In the EAST plant data, sensors correspond to measurements level in IoTDB, which are the leaves of the tree structure.

Test

ETRA Performance Test (Write Throughput, CPU, Memory, Completeness)

Test Data: EAST cryogenic system and vacuum system plant data.



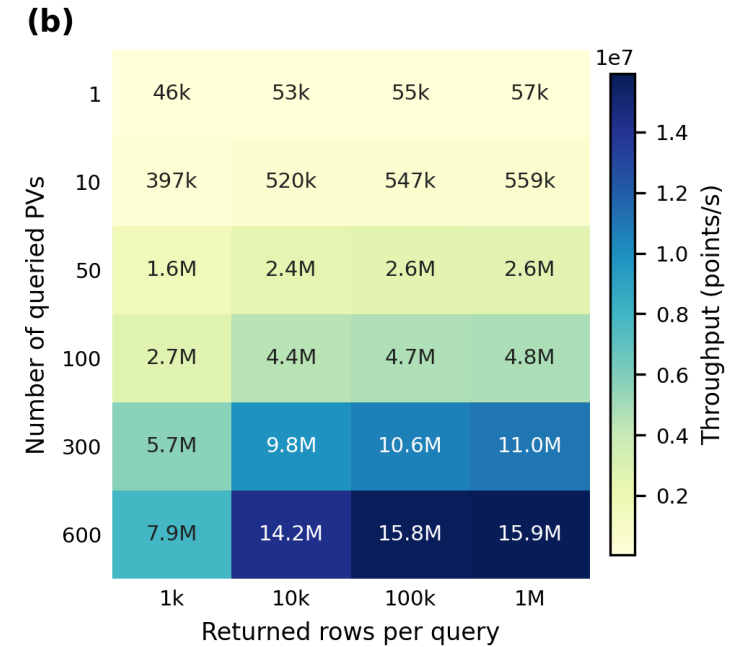
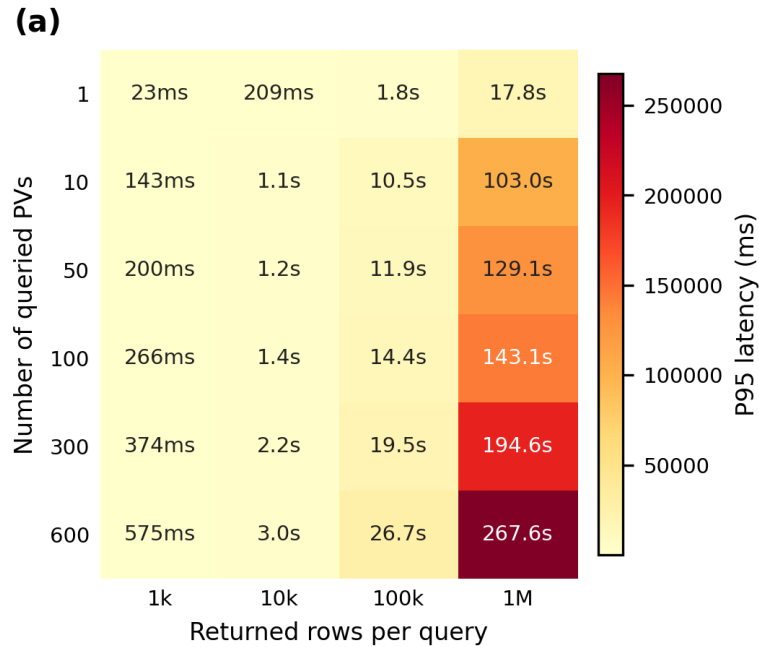
- Matrix: PVs \in {10k, 50k, 100k} \times frequency \in {1, 5, 10 Hz} for 300 s.
- (a) CPU usage stays under \sim 70% even at 100k PV \times 10 Hz.
- (b) Memory usage linear with PV count, \leq \sim 500 MB.
- (c) Write throughput: up to \sim 1 M points/s at 100k \times 10 Hz.
- (d) Archival completeness = 100% across all 9 cells — zero point loss.

Test

ETRA Performance Test (Range Query Performance)

Test Data:

EAST cryogenic system and vacuum system plant data.

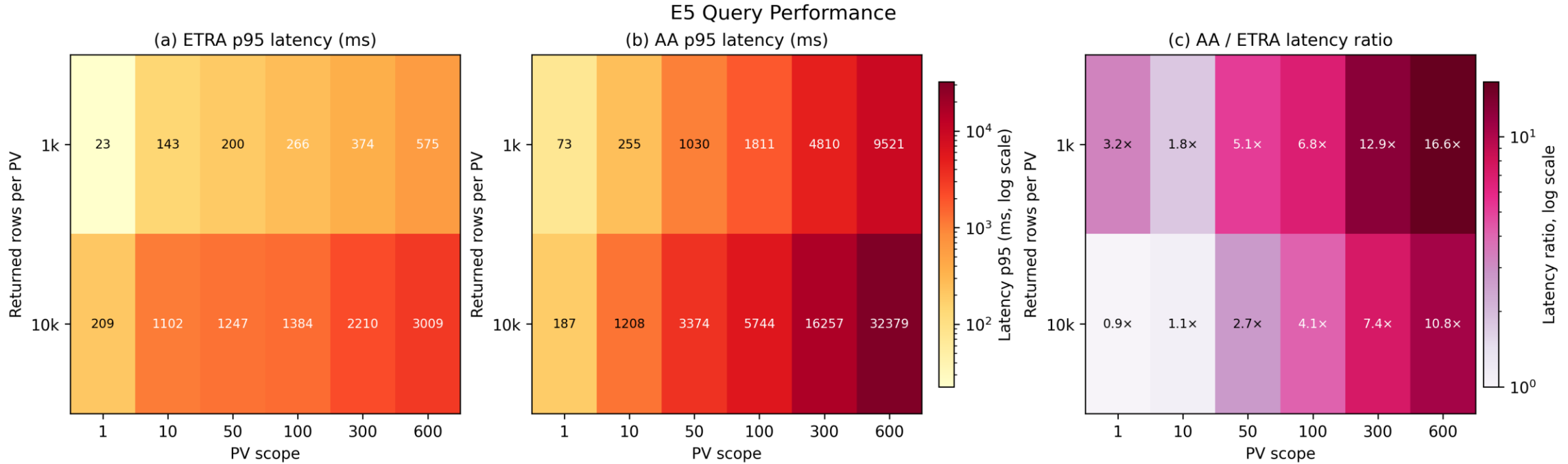


- Workload: 600 cryogenic PVs, real EAST archive.
- Matrix: returned rows per query $\in \{1k, 10k, 100k, 1M\} \times$ PV scope $\in \{1 \dots 600\}$.
- (a) p95 latency: scales smoothly from 23 ms (1 PV, 1k rows) to 267 s (600 PV, 1M rows).
- (b) Throughput: peaks at ~ 15.9 M points/s for 600 PV \times 1M rows.
- Multi-PV range queries scale near-linearly without per-PV fan-out cost.

Test

Comparative Test

Query performance test compare with EPICS Archiver Appliance



- Matched returned data volume: 1k and 10k rows per PV, 1–600 PV scopes.
- (a) ETRA p95 latency: 23 ms – 3.0 s.
- (b) AA p95 latency: 73 ms – 32.4 s.
- (c) Latency ratio: AA is up to 16.6x slower at 600 PV × 1k rows; up to 10.8x slower at 600 PV × 10k rows.

Test

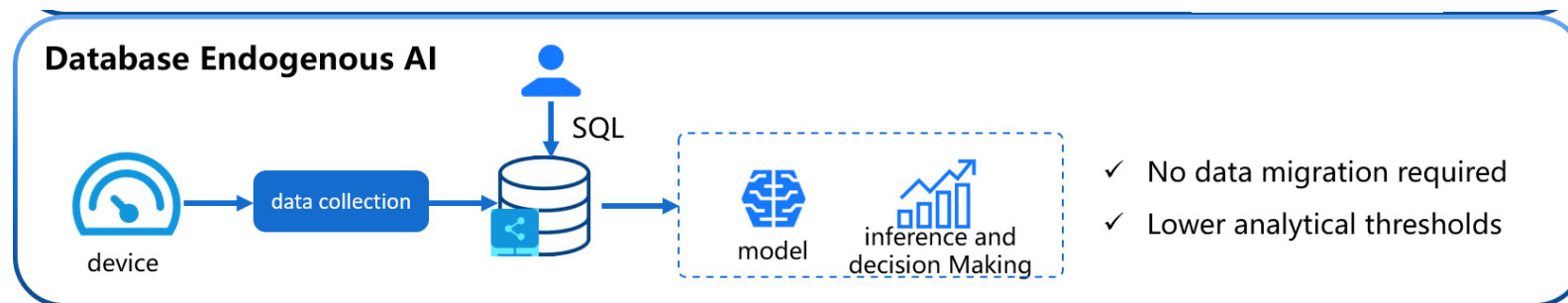
AI Node Train Test

Timeseries-Model { Multi-step forecasting: Timer-XL, Sundial.
One-step forecasting: navie_forecaster, exponential_smoothing, holtwinters, stl_forecaster
Anomaly detection: stray

Test Data

EAST Plant Data, Cryogenic system, TI2226, 10000 Timestamp.

Method of testing

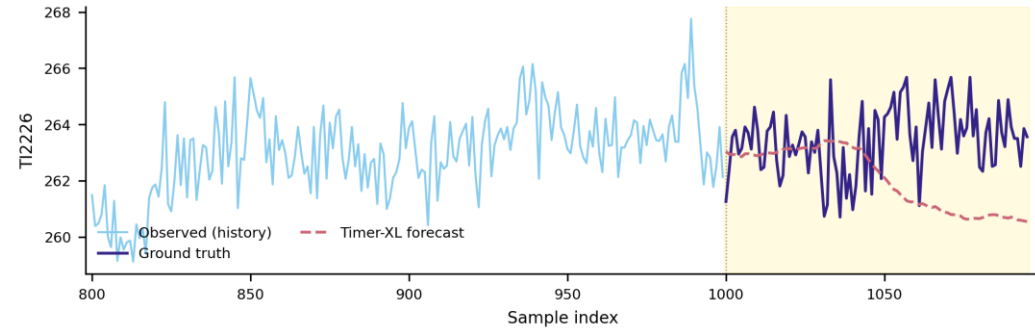


- [1] Timer- Generative Pre-trained Transformers Are Large Time Series Models, Yong Liu, Haoran Zhang, Chenyu Li, Xiangdong Huang, Jianmin Wang, Mingsheng Long.
- [2] TIMER-XL- LONG-CONTEXT TRANSFORMERS FOR UNIFIED TIME SERIES FORECASTING ,Yong Liu, Guo Qin, Xiangdong Huang, Jianmin Wang, Mingsheng Long.
- [3] Sundial- A Family of Highly Capable Time Series Foundation Models, Yong Liu, Guo Qin, Zhiyuan Shi, Zhi Chen, Caiyin Yang, Xiangdong Huang, Jianmin Wang, Mingsheng Long, ICML 2025 spotlight.

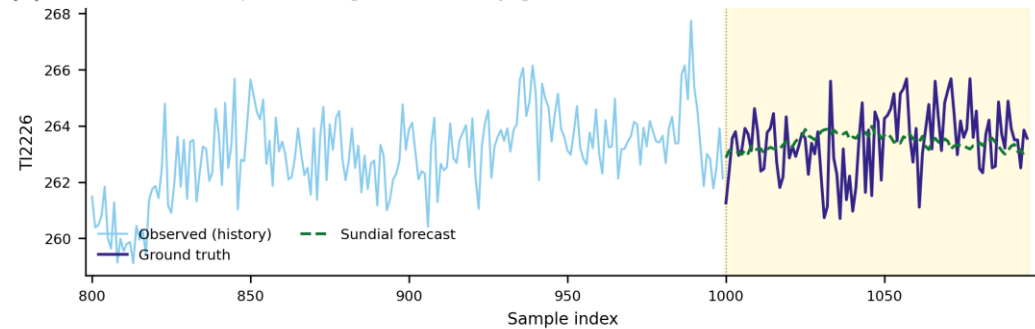
Test

AI Node Train Test

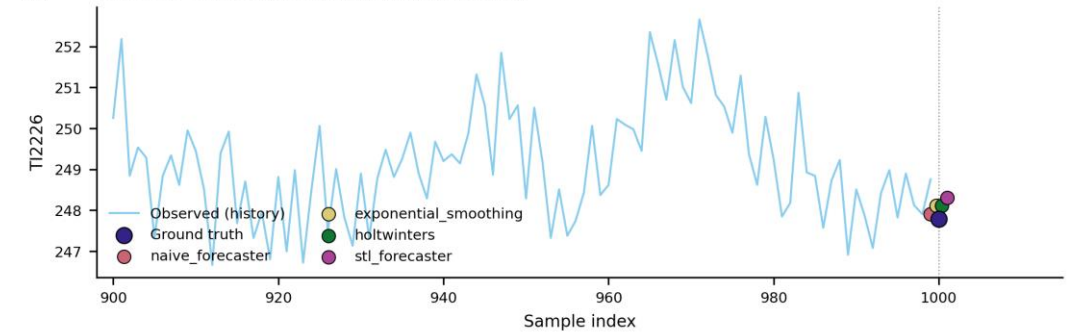
(a) Timer-XL multi-step forecasting on archived cryogenic data



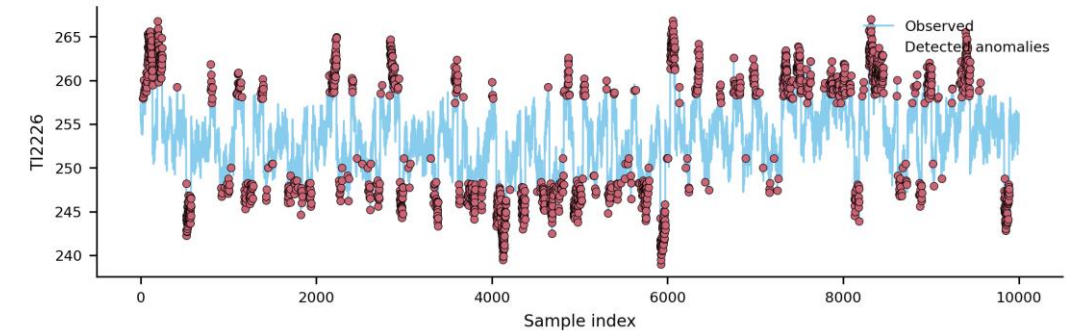
(b) Sundial multi-step forecasting on archived cryogenic data



(c) One-step-ahead forecasting (classical models)



(d) Anomaly detection using Stray



- (a) Timer-XL multi-step forecast: drifts toward mean over 80-step horizon.
- (b) Sundial multi-step forecast: tracks trend, smooths high-frequency noise.
- (c) Classical one-step models (naive, exp-smoothing, Holt-Winters, STL): all converge near ground truth.
- (d) Stray anomaly detection: identifies excursions outside normal cryogenic operating band.

Summary and plan

Summary and Future work

Summary:

- ETRA bridges EPICS Channel Access and a Timeseries Database.
- By leveraging the large-scale query capability of TSDBs, ETRA achieves strong performance for EAST plant data archiving and analysis.
- Native AI workloads (forecasting, anomaly detection) supported via IoTDB AINode.

Future work:

- Adaptive flush policy based on PV update statistics.
- Design time series models more suitable for EAST plant data.
- Deployment validation on full EAST production workload.



Thank you for your attention !

Guang Yang,
guangyang@mail.ustc.edu.cn
25th IEEE Real Time Conference - La Biodola, Elba, Italy
May 25-29, 2026



EAST