



2025 AI+ Development  
Digital Summit

# AI+ 研发数字峰会

拥抱AI 重塑研发

05/23-24 | 上海站



# 2025 AI+研发数字峰会

拥抱AI 重塑研发 AI+ Development Digital Summit

下一站预告

08/08-09 | 北京站

11/14-15 | 深圳站



查看会议详情

## 北京站论坛设置

大模型和 AI 应用评测

智能存储与检索技术

下一代知识工程

AI+ 金融业务创新

智能需求工程

智能体与研发效率工具

AI 产品运营与出海策略

大模型安全与对齐

大模型应用开发框架与实践

智能体经济 (Agentic Economy)

智能测试工具的开发与应用

具身智能与机器人

代码生成及其改进

AI+ 新能源汽车

AI 前沿技术探索与实践

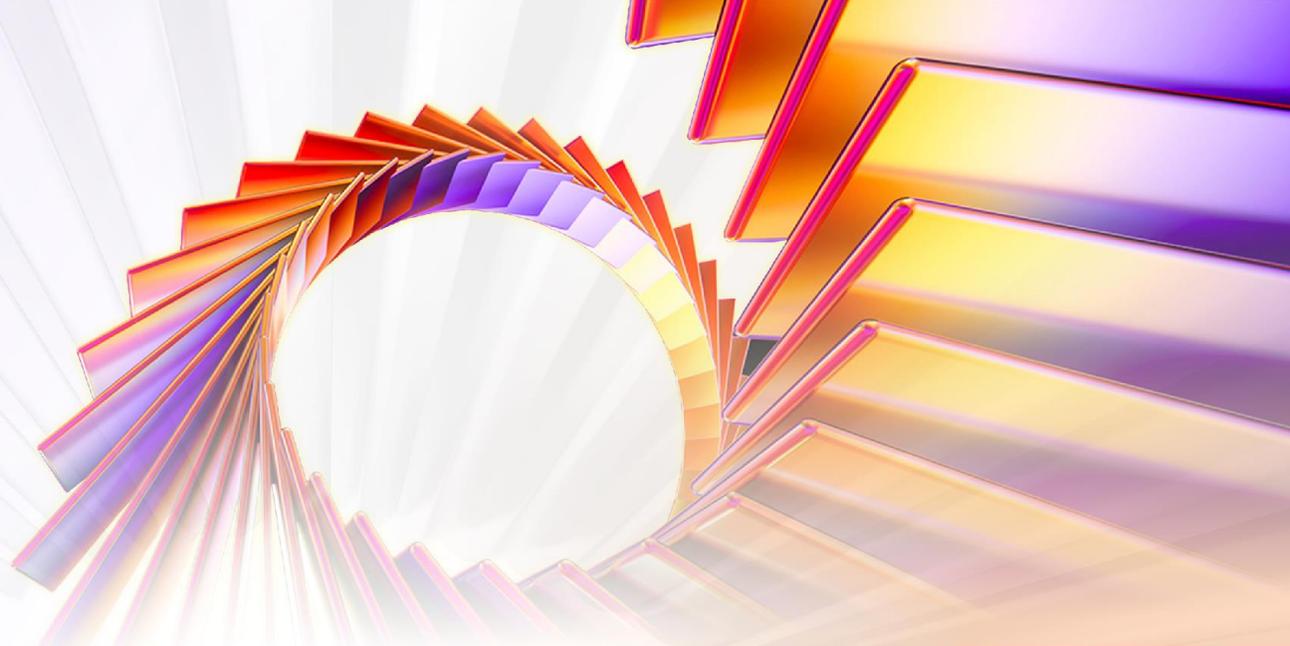


| 05/23-24 | 上海站

**2025** AI+ Development  
Digital Summit

# AI+研发数字峰会

拥抱AI 重塑研发



# 时序数据迎来 AI 智能体时代

陶建辉 | TDengine



## 陶建辉

TDengine 创始人 & CEO

---

开源软件 TDengine 主要作者，2024 年“中国计算机学会(CCF)杰出工程师奖”获得者。1994年中国科大毕业后，到美国留学，1997年起，先后在芝加哥Motorola、3Com等公司从事移动互联网的研发工作。2008年初回到北京创办和信，后被联发科收购。2013年初创办快乐妈咪，后被太平洋网络收购。2017年5月创办涛思数据，专注物联网、工业互联网大数据的处理，产品TDengine开源后，在GitHub全球趋势排行榜上多日排名第一。涛思数据已获红杉、GGV、经纬、明势资本等多家机构的近7000万美元的投资。

# PART 01

# 令人振奋的 AI 时代

- 01 需 5 到 10 年才能获取的知识，现在一周就够
- 02 特有或私域的知识，可以依靠 RAG 技术来解决
- 03 LLM 能击败 80% 以上的普通人
- 04 LLM 的推理能力还在提升



01 训练成本和推理成本已经大幅下降

02 LLM 的成本还会再降低

03 无需 AI 专业知识，就可使用

04 AI 走进工业，成本将不再是问题



# ▶ 工业 AI 智能体时代

- 01 从采集的数据感知运行状态
- 02 基于数据自主做分析
- 03 基于分析结果，给出决策建议
- 04 每人拥有自己的实时数据分析助理
- 05 Self-Learning Data Platform



# PART 02

## TDgpt

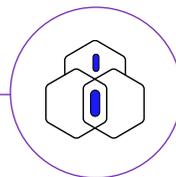
# 时序数据分析 AI 智能体

# ▶ TDgpt 时序数据分析 AI 智能体

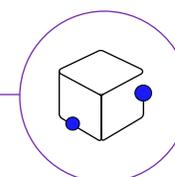
TDgpt 是 TDengine 内置的时序数据分析 AI 智能体，具备时序数据预测、异常检测、补齐和分类功能。它能无缝对接各种时序基础模型、大语言模型、机器学习及传统统计算法，并支持算法动态切换，用户只需通过一条 SQL 语句即可使用。此外，通过开放的 SDK，开发者能将自己开发的 AI 算法或模型轻松集成到 TDgpt，立即为 TDengine 全行业用户所使用。



通过 SQL 提供时序数据预测、异常检测、补齐和分类的功能



支持各种时序模型、大语言模型机器学习和传统算法的动态切换



通过 SDK，任何新的算法、模型都可无缝集成，且无需修改应用



## 样列表

列名称	类型	说明
ts	timestamp	主时间戳列
i32	int32	4 字节整数, 设备测量值 metric

## 样例数据

```
taos> select * from foo;
      ts                |      i32      |
=====|=====|
2020-01-01 00:00:12.681 |      13      |
2020-01-01 00:00:13.727 |      14      |
2020-01-01 00:00:14.378 |       8      |
2020-01-01 00:00:15.774 |      10      |
2020-01-01 00:00:16.170 |      16      |
2020-01-01 00:00:17.558 |      26      |
2020-01-01 00:00:18.938 |      32      |
2020-01-01 00:00:19.308 |      27      |
```

## SQL 预测函数定义

```
FORECAST(column_expr, option_expr)

option_expr: {"
  algo=expr1
  [,wncheck=1|0]
  [,conf=conf_val]
  [,every=every_val]
  [,rows=rows_val]
  [,start=start_ts_val]
  [,expr2]
"}

```

1. **column\_expr**: 预测的时序数据列, 只支持数值类型列输入。
2. **options**: 预测函数的参数。字符串类型, 其中使用 K=V 方式调用算法及相关参数。采用逗号分隔的 K=V 字符串表示, 其中的字符串不需要使用单引号、双引号、或转义号等符号, 不能使用中文及其他宽字符。预测支持 conf, every, rows, start, rows 几个控制参数, 其含义如下:



## 参数说明

参数	含义	默认值
algo	预测分析使用的算法	holtwinters
wncheck	白噪声 (white noise data) 检查	默认值为 1, 0 表示不进行检查
conf	预测数据的置信区间范围, 取值范围 [0, 100]	95
every	预测数据的采样间隔	输入数据的采样间隔
start	预测结果的开始时间戳	输入数据最后一个时间戳加上一个采样间隔时间区间
rows	预测结果的记录数	10

1. 预测查询结果新增三个伪列, 具体如下: **\_FROWTS**: 预测结果的时间戳、**\_FLOW**: 置信区间下界、**\_FHIGH**: 置信区间上界, 对于没有置信区间的预测算法, 其置信区间同预测结果
2. 更改参数 **START**: 返回预测结果的起始时间, 改变起始时间不会影响返回的预测数值, 只影响起始时间。

## 示例

```
-- 使用 arima 算法进行预测, 预测结果是 10 条记录 (默认值), 数据进行白噪声检查, 默认置信区间 95%.  
SELECT _flow, _fhigh, _frowts, FORECAST(i32, "algo=arima")  
FROM foo;
```

```
-- 使用 arima 算法进行预测, 输入数据的是周期数据, 每 10 个采样点是一个周期, 返回置信区间是95%的上下。  
SELECT _flow, _fhigh, _frowts, FORECAST(i32, "algo=arima,alpha=95,period=10,wncheck=0")  
FROM foo;
```

```
taos> select _flow, _fhigh, _frowts, forecast(i32) from foo;  
=====
```

_flow	_fhigh	_frowts	forecast(i32)
10.5286684	41.8038254	2020-01-01 00:01:35.000	26
-21.9861946	83.3938904	2020-01-01 00:01:36.000	30
-78.5686035	144.6729126	2020-01-01 00:01:37.000	33
-154.9797363	230.3057709	2020-01-01 00:01:38.000	37
-253.9852905	337.6083984	2020-01-01 00:01:39.000	41
-375.7857971	466.4594727	2020-01-01 00:01:40.000	45
-514.8043823	622.4426270	2020-01-01 00:01:41.000	53
-680.6343994	796.2861328	2020-01-01 00:01:42.000	57
-868.4956665	992.8603516	2020-01-01 00:01:43.000	62
-1076.1566162	1214.4498291	2020-01-01 00:01:44.000	69

3. **EVERY**: 可以与输入数据的采样频率不同。采样频率只能低于或等于输入数据采样频率, 不能高于输入数据的采样频率。
4. 对于某些不需要计算置信区间的算法, 即使指定了置信区间, 返回的结果中其上下界退化成为一个点。



## SQL 异常检测函数定义

```
ANOMALY_WINDOW(column_name, option_expr)
```

```
option_expr: {"  
  algo=expr1  
  [,wncheck=1|0]  
  [,expr2]  
"}
```

--- 使用 iqr 算法进行异常检测, 检测列 i32 列。

```
SELECT _wstart, _wend, SUM(i32)  
FROM foo  
ANOMALY_WINDOW(i32, "algo=iqr");
```

--- 使用 ksigma 算法进行异常检测, 输入参数 k 值为 2, 检测列 i32 列

```
SELECT _wstart, _wend, SUM(i32)  
FROM foo  
ANOMALY_WINDOW(i32, "algo=ksigma,k=2");
```

```
taos> SELECT _wstart, _wend, count(*) FROM foo ANOMAYL_WINDOW(i32);  
  _wstart          |          _wend          | count(*) |  
=====|=====|=====|  
  2020-01-01 00:00:16.000 | 2020-01-01 00:00:17.000 |         2 |  
Query OK, 1 row(s) in set (0.028946s)
```

1. column\_name: 进行时序数据异常检测的输入数据列, 当前只支持单列, 且只能是数值类型
2. options: 字符串。其中使用 K=V 调用异常检测算法及与算法相关的参数。采用逗号分隔的 K=V 字符串表示, 其中的字符串不需要使用单引号、双引号、或转义号等符号, 不能使用中文及其他宽字符。例如: algo=ksigma, k=2 表示进行异常检测的算法是 ksigma, 该算法接受的输入参数是 2。
3. 异常检测的结果可以作为外层查询的子查询输入, 在 SELECT 子句中使用的聚合函数或标量函数与其他类型的窗口查询相同。
4. 输入数据默认进行白噪声检查, 如果输入数据是白噪声, 将不会有任何 (异常) 窗口信息返回。

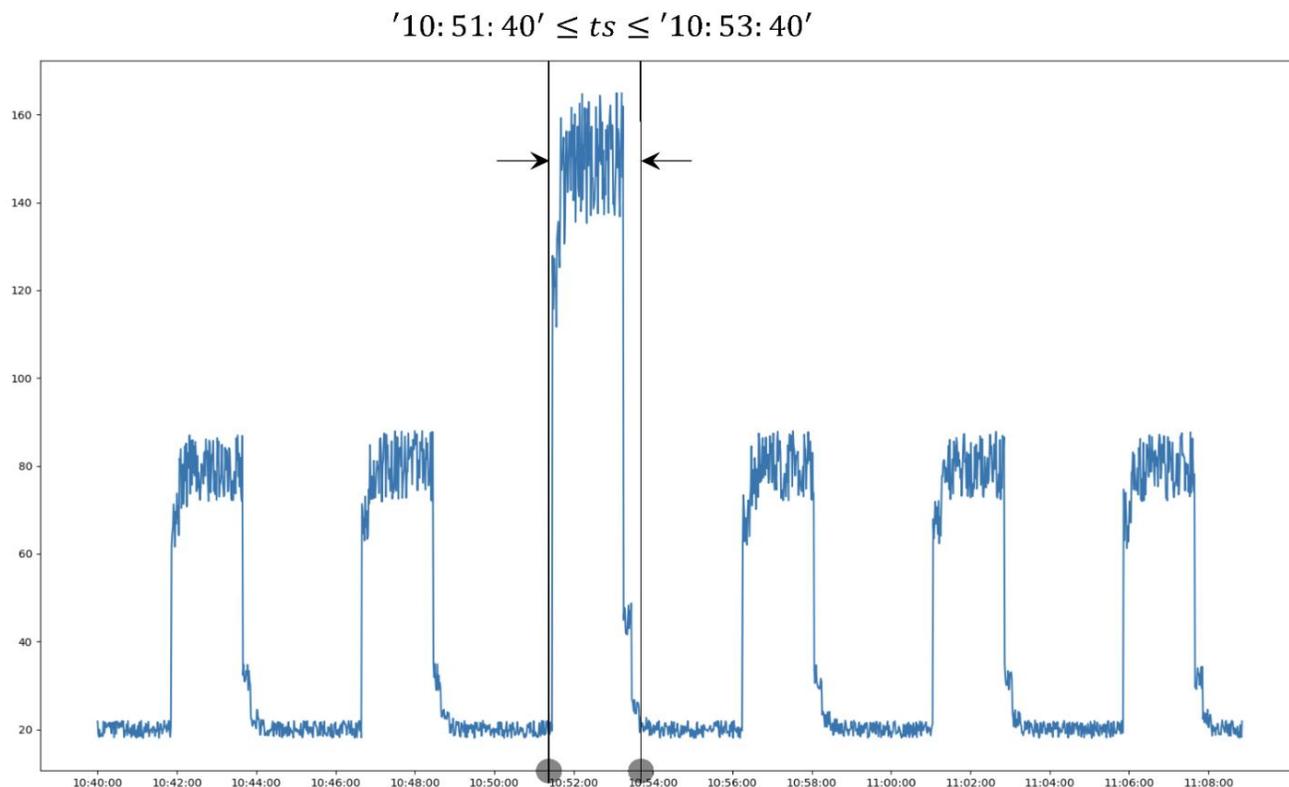


# 时序数据异常检测

—— 使用异常检测算法 IQR 对输入列 col\_val 进行异常检测。同时输出异常窗口的起始时间、结束时间、以及异常窗口内 col 列的和。

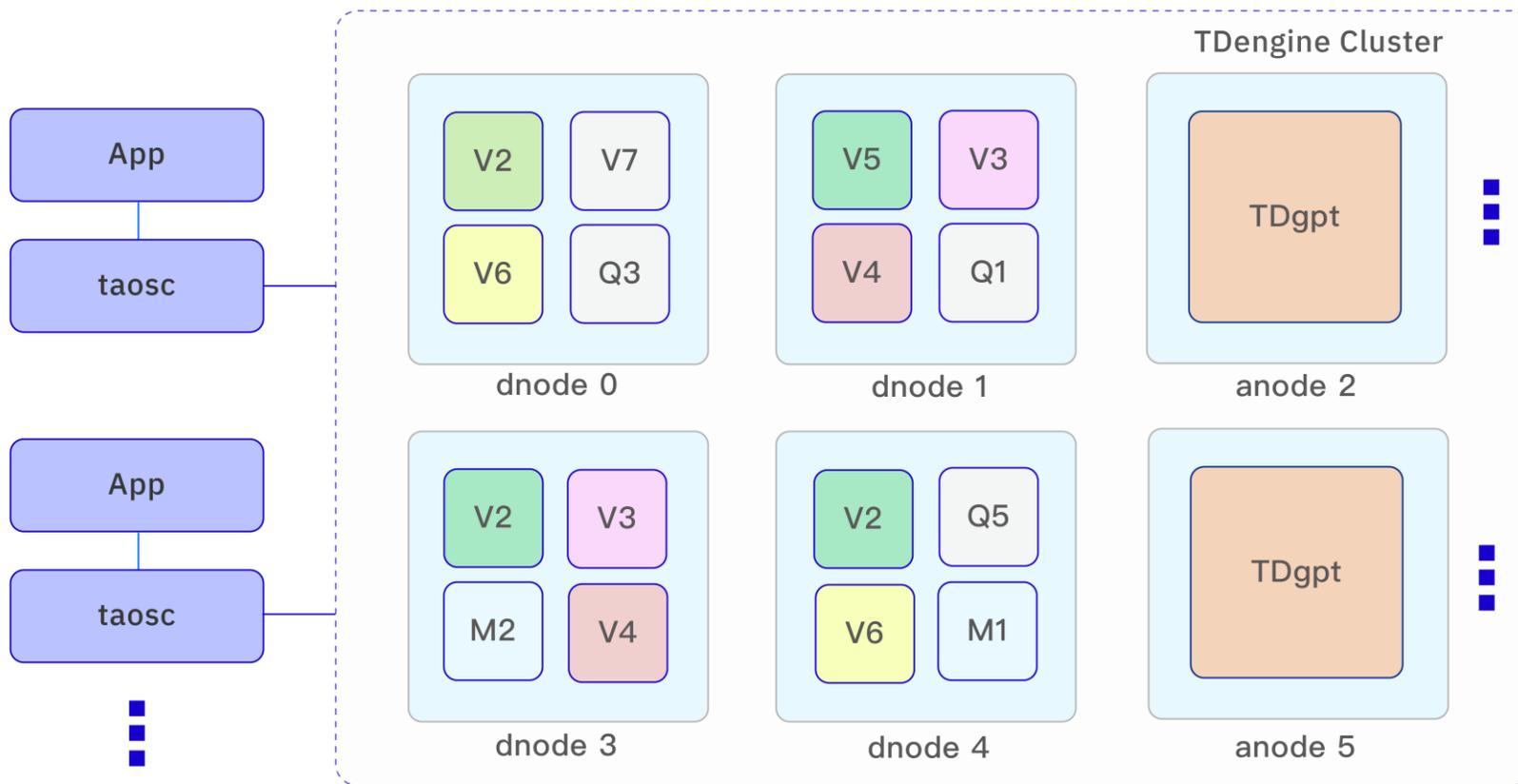
```
SELECT _wstart, _wend, SUM(col)
FROM foo
ANOMALY_WINDOW(col_val, "algo=iqr");
```

如下图所示，Anode 将返回时序数据异常窗口 [10:51:30, 10:53:40]



1. 用户可以针对异常窗口内的时序数据进行查询聚合、变换处理等操作。
2. 异常检测可以通过 where 条件限制其应用的时间范围。
3. 目前仅支持基于查询的异常检测，与流式计算的整合预计将在 Q2 完成。

# TDengine with TDgpt



基于 TDengine 3.3.6.0 版本

- 在 TDengine 已有的 Vnode、Mnode、Qnode 节点基础上增加了新的节点 Anode, 并且是独立的进程。
- Anode: TDgpt 执行分析的节点, 可以存在多个。每个 Anode 承载不同的模型和算法, 并向 Mnode 注册。



## 时序预测模型

模型名称	发布时间
ARIMA	已发布
HoltWinters	已发布
Exponential Smoothing (ES)	2025.Q1
Prophet	2025.Q1
XGBoost	2025.Q2
LightGBM.	2025.Q2
ETS	2025.Q3
MSTL	2025.Q3
Theta	2025.Q4
LSTM	2025.Q1
Deep-AR	2025.Q1
N-BEATS	2025.Q2
N-Hits	2025.Q2
Patch-TST	2025.Q3
Temporal Fusion Transformer	2025.Q3
TimesNet	2025.Q4
Multilayer Perceptron (MLP)	2025.Q4

## 异常检测模型

模型名称	发布时间
Interquartile Range (IQR)	已发布
K-sigma	已发布
Grubbs	已发布
SHESD	2025.Q2
Isolation Forest	2025.Q2
One-Class SVM	2025.Q3
Prophet	2025.Q3
AutoEncoder	已发布
Local Outlier Factor (LOF)	已发布
DBSCAN	2025.Q1
KNN	2025.Q2
PCA	2025.Q2
STL	2025.Q3
CUSUM	2025.Q3
PELT (Pruned Exact Linear Time)	2025.Q4
Gaussian Process Regression	2025.Q4
...	2025.Q4



# ▶ TDgpt 支持的时序基础模型

大语言模型能力很强，但工业时序数据有其特点，直接用它来做时序数据预测，效果并不理想，而且工业对成本要求更加苛刻，因此各大公司采用 LLM 的 Transformer 架构，推出时序基础模型。TDgpt 支持如下时序模型：

- 01 **TDtsfm:** TDengine time series foundation model by TAOS Data
- 02 **Uni2TS:** Unified Training of Universal Time Series Transformers by SalesForce
- 03 **Chronos:** Learning the Language of Time Series by Amazon
- 04 **TimesFM:** Time Series Foundation Model by Google
- 05 **Time-MoE:** Billion-Scale Time Series Foundation Models by Xiaohongshu



## Anode的目录构成

```
.
├── bin
├── cfg
├── lib
│   └── taosanalytics
│       ├── algo
│       │   ├── ad
│       │   └── fc
│       ├── misc
│       └── test
├── log -> /var/log/taos/taosanode
├── model -> /var/lib/taos/taosanode/model
└── venv -> /var/lib/taos/taosanode/venv
```

目录	说明
taosanalytics	源代码目录, 其下包含了算法具体保存目录 algo, 放置杂项目录 misc, 单元测试和集成测试目录 test。 algo 目录下 ad 保存异常检测算法代码, fc 目录保存预测算法代码
venv	Python 虚拟环境
model	放置针对数据集完成的训练模型
cfg	配置文件目录

## 定义了两类算法模版

### AbstractForecastService

- Execute 方法返回字典对象
- 提供一系列默认参数

### AbstractAnomalyDetectionService

- Execute 方法返回 List 对象
- 无默认参数

基于上述模版实现模型并放置于对应的目录下, 便可以被 Anode 发现并注册到 TDengine 集群使用。



# PART 03

## TDgpt 亮点

# 零学习门槛

## 仅需一行 SQL

01

通过两个 SQL 函数, forecast 与 anomaly\_window, 即可完成时序数据预测与异常检测。

02

无需改动应用程序, SQL 语句里指定算法名字, 即可动态切换算法、大模型。

03

无需编程经验、无需了解算法、大语言模型、时序模型的实现, 真正零门槛使用。



# 零初期投入 立即享受 AI 红利

01

TDengine 以及 TDgpt 执行代码，下载即可使用。更可免费注册云服务，立即体验。

02

提供内置通用模型，并提供广泛的合作伙伴解决方案接入，让您马上拥有时序数据预测与异常检测的能力。

03

10 分钟内就能完成时序预测与异常检测模型的上线，选择最适合自己的场景的模型和算法，加快决策。



# 零时间差

# 最新算法轻松集成

01

通过开放的 SDK，开发者能将自己开发的或开源的 AI 算法或模型轻松集成到 TDgpt。

02

对于新的算法、模型，应用系统无需做任何调整，降低研发投入。

03

算法、模型开发者可以聚焦研发，无需考虑 data engineering 的工作。



# TDgpt 核心代码开源

---

<https://github.com/taosdata/TDEngine>

# ▶▶ 立即体验 TDgpt

场景1

用电预测

场景2

发电预测

场景3

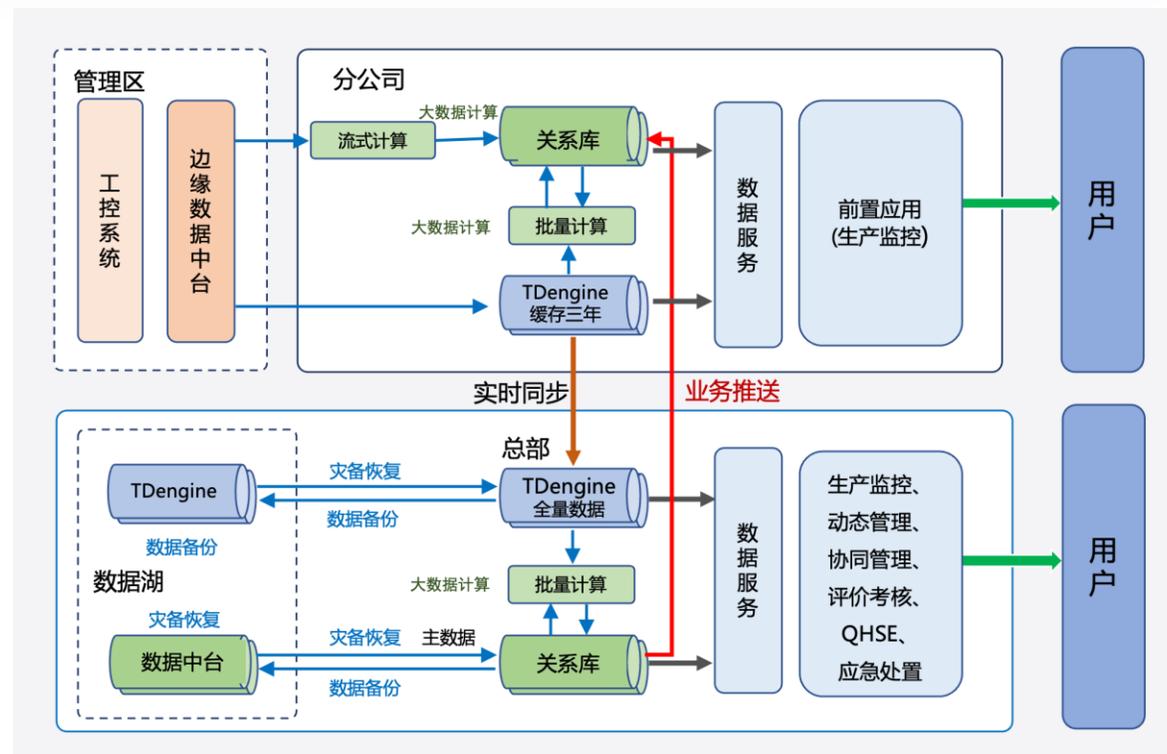
运维预测

访问地址: [www.taosdata.com/tdgpt](http://www.taosdata.com/tdgpt)

# 中石化 PCS 系统架构和场景

油田 PCS 实时数据由管理区 SCADA 汇聚至分公司时序数据库存储（存储三年），经分布式集群同步技术将数据汇聚至中心节点时序库（永久存储）；业务数据存储在关系库中，由中心节点统一管控，根据分公司应用需求向下推送。

架构中关系库和时序库数据在分公司、总部层面分别入湖，实现企业数据湖和总部数据湖的同步。结合数据湖的定位，综合考虑 PCS 系统对实时数据的高频访问和完整性需求，采取实时数据先汇聚至中心节点，由中心节点进行分节点的集中管控，再同步至数据湖，由数据湖对外（除 PCS）提供统一数据服务的路线。



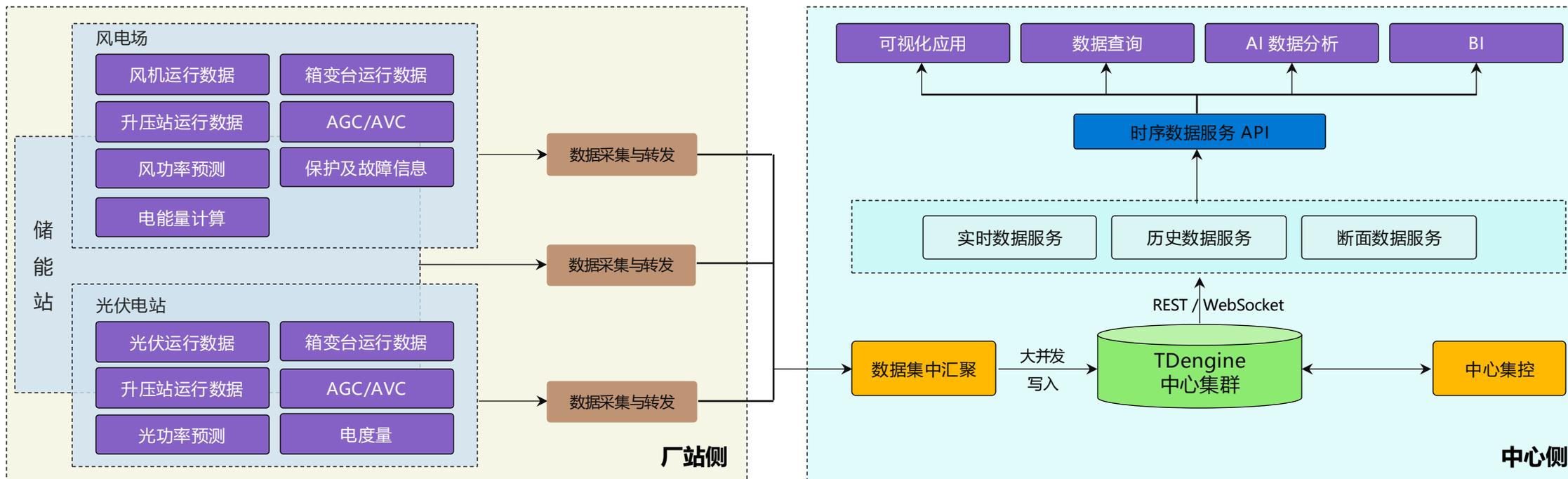
# 三峡新能源风光储集控中心

## 项目现状

- 已接入 300 座左右风光储场站，每两天新增一座接入
- 每年接入数据约 800TB~1PB

## 客户核心关切

- 高性能的实时+历史库
- 水平扩展、存储冷热分级
- 高可用
- 运行稳定可靠



- 01 7月发布时序数据补全、时序数据分类的功能
- 02 集成更多的时序数据开源模型和算法
- 03 不断提升自有的时序数据基础模型 TDtsfm 的竞争力



# PART 04

## 未来展望

# ▶ 未来展望：All in AI

- 01 持续优化涛思数据自研的时序大模型
- 02 用 AI 智能体重构工业大数据平台，计划 7 月发布
- 03 坚定地走开源的道路



# ▶ 期待大家

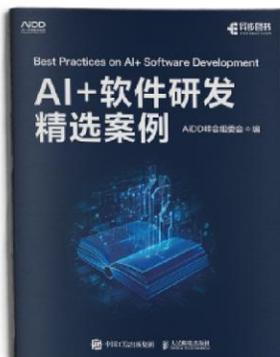
- 01 去 [taosdata.com](https://taosdata.com) 下载 TDengine, 下载 TDgpt
- 02 去 GitHub 查看 TDgpt 源代码, 自己编译、测试、修改、运行
- 03 如果有特殊场景想试用 TDgpt, 可以联系我们





# Make Industrial Data Accessible, Affordable and Valuable

## 参与调研您将优先获得



AiDD定制版  
《AI+软件研发精选案例》



专属学习顾问  
1对1需求对接

# AiDD会后小调研

AiDD峰会致力于协助企业利用AI技术深化计算机对现实世界的理解，推动研发进入智能化和数字化的新时代。作为峰会的重要共建者，您的真知灼见对我们至关重要。衷心感谢您的参与支持！

# 2025 AI+研发数字峰会

## 拥抱 AI 重塑研发



扫码参与调研

# 科技生态圈峰会 + 深度研习

—1000+ 技术团队的选择



**K+峰会** **敦煌站**  
**K+ 思考周®研习社**  
时间: 2025.08.29-30

**K+峰会** **上海站**  
**K+ 金融专场**  
时间: 2025.09.26-27

**K+峰会** **香港站**  
**K+ 思考周®研习社**  
时间: 2025.11.17-18



K+峰会详情



**AiDD峰会** **上海站**  
**AI+研发数字峰会**  
时间: 2025.05.23-24

**AiDD峰会** **北京站**  
**AI+研发数字峰会**  
时间: 2025.08.08-09

**AiDD峰会** **深圳站**  
**AI+研发数字峰会**  
时间: 2025.11.14-15



AiDD峰会详情



2025 AI+研发数字峰会  
AI+ Development Digital Summit

**感谢聆听!**

扫码领取会议PPT资料

