

ỨNG DỤNG PHÂN TÁN

CSE702063-1-3-24(04)

ThS. Nguyễn Thành Trung



NHÓM 15

TRẦN BÁ TÀI-22010012

Với chủ đề :

influxdb - kho dữ liệu
có thể mở rộng cho số
liệu, sự kiện và phân
tích thời gian thực

Đề tài :

Xây dựng hệ thống
giám sát và cảnh báo
thời tiết sử dụng
InfluxDB và Python



INFLUXDB

Tổng quan dự án
đã chọn

- Tổng quan về Influx
- Mục đích sử dụng
- Chức năng
- Ứng dụng thực tế
- Cài đặt influx

Phát triển và triển khai
kỹ thuật mới

- Mô hình hệ thống
- Lưu trữ dữ liệu
- Cảnh báo qua mail
- Xây dựng giao diện web
- Triển khai

Kết luận và hướng
phát triển

- Kết luận
- Hướng phát triển





TỔNG QUAN

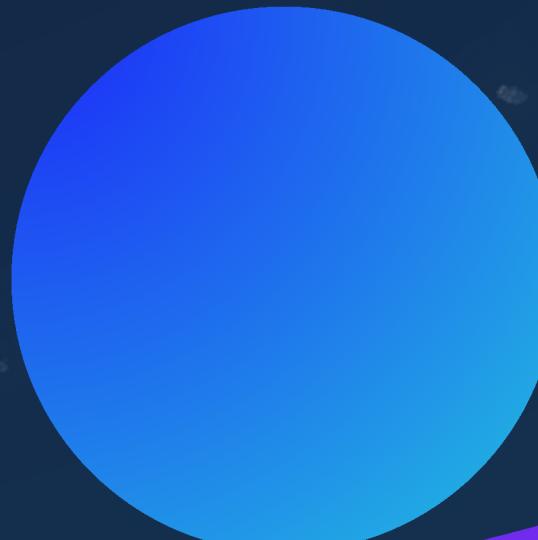
- InfluxDB là cơ sở dữ liệu chuỗi thời gian mã nguồn mở do InfluxData phát triển.
- Viết bằng Go, hỗ trợ HTTP, UDP.
- Là một phần của hệ sinh thái gồm:
 - → Telegraf (thu thập)
 - → Chronograf (trực quan)
 - → Kapacitor (xử lý thời gian thực)





MÔ HÌNH DỮ LIỆU

- Measurement: Nhóm dữ liệu (như bảng SQL)
- Tags: Metadata (có index)
- Fields: Dữ liệu chính (không index)
- Timestamp: Mốc thời gian chính xác
- → Một series = Measurement + Tag + Field + Timestamp



MỤC ĐÍCH , CHỨC NĂNG & ỨNG DỤNG

- Ghi dữ liệu thời gian thực từ cảm biến, hệ thống, tài chính...
- Truy vấn nhanh theo thời gian
- Lưu trữ tối ưu, hỗ trợ cloud

- Dữ liệu thời gian chuyên biệt
- Truy vấn với InfluxQL, Flux
- Quản lý: Retention, Sharding
- Tích hợp: Grafana, Python, Java...
- Hiệu suất cao, nén tốt



influxdb

- Giám sát hệ thống: Server, app, container
- IoT – công nghiệp: Nhà máy, cảm biến, năng lượng
- Tài chính: Giao dịch, phân tích dữ liệu
- Y tế: Theo dõi bệnh nhân, nghiên cứu lâm sàng

CÀI ĐẶT INFLUXDB



influxdb

influxdb

Tất cả Hình ảnh Video Mua sắm Tin tức Video ngắn Web Thủ

 InfluxData <https://www.influxdata.com> · Dịch trang này

InfluxDB Time Series Data Platform | InfluxData

InfluxDB is a G2 Leader in Time Series. "InfluxDB is a strong database built specifically for time series data. It has made working with such data seamless and ..."

InfluxData Downloads

InfluxDB is a comprehensive, open source time series toolkit for ...

InfluxDB OSS 2.x

InfluxDB is a comprehensive, open source time series toolkit for managing and analyzing time series data. The InfluxDB platform includes a command-line interface, a web-based UI, and a set of monitoring and alerting tools, and an ecosystem with client libraries and powerful integrations with the tools you love, including Grafana.

Info InfluxDB 3 is now generally available. For new users, we recommend starting there.

Platform Windows Binaries (64-bit) - using PowerShell **Version** InfluxDB v2.7.12

SHA256: <adb952c67eaeeaa742d81899367ec95defa26b47a0fa99dc773ab4de822e9872d>

```
wget https://download.influxdata.com/influxdb/releases/v2.7.12/influxdb2-2.7.12-windows.zip -UseBasicParsing -OutFile influxdb2-2.7.12-windows.zip
Expand-Archive .\influxdb2-2.7.12-windows.zip -DestinationPath 'C:\Program Files\InfluxData\influxdb'
```

[Documentation](#) [Release Notes](#) [Register to get early access to InfluxDB 3.0](#)

> This PC > Tайдzz (C:) > Users > Tai > influxdb2-2.7.12-windows

Name	Date modified	Type	Size
influxd	5/20/2025 10:53 PM	Application	111,327 KB
LICENSE	5/20/2025 10:53 PM	File	2 KB
README	5/20/2025 10:53 PM	Markdown Source ...	12 KB

C:\Users\Tai\influxdb2-2.7.12->

```
e\\autogen\\2", "index_version": "tsil", "duration": "64.999ms"}  
2025-06-04T22:36:39.539335Z info Finished loading shard, current progress 100.0% shards (1 / 1). {"log_id": "0wwHH600000", "service": "storage-engine"}  
2025-06-04T22:36:39.540324Z info Open store (end) {"log_id": "0wwHH600000", "service": "storage-engine", "service": "store", "op_name": "tsdb_open", "op_event": "end", "op_elapsed": "124.957ms"}  
2025-06-04T22:36:39.540597Z info Starting retention policy enforcement service {"log_id": "0wwHH600000", "service": "retention", "check_interval": "30m"}  
2025-06-04T22:36:39.541712Z info Starting precreation service {"log_id": "0wwHH600000", "service": "shard-precreation", "check_interval": "10m", "advance_period": "30m"}  
2025-06-04T22:36:39.549623Z info Starting query controller {"log_id": "0wwHH600000", "service": "storage-reports", "concurrency_quota": 1024, "initial_memory_bytes_quota_per_query": 9223372036854775807, "memory_bytes_quota_per_query": 9223372036854775807, "max_memory_bytes": 0, "queue_size": 1024}  
2025-06-04T22:36:39.560588Z info Configuring InfluxQL statement executor (zeros indicate unlimited). {"log_id": "0wwHH600000", "max_select_point": 0, "max_select_series": 0, "max_select_buckets": 0}  
2025-06-04T22:36:39.574693Z info Starting {"log_id": "0wwHH600000", "service": "telemetry", "interval": "8h"}  
2025-06-04T22:36:39.576694Z info Listening {"log_id": "0wwHH600000", "service": "tcp-listener", "transport": "http", "addr": ":8086", "port": 8086}  
2025-06-04T22:36:46.772277Z info Unauthorized {"log_id": "0wwHH600000", "error": "token required"}  
2025-06-04T23:06:39.557564Z info Retention policy deletion check (start) {"log_id": "0wwHH600000", "service": "retention", "op_name": "retention_delete_check", "op_event": "start"}  
2025-06-04T23:06:39.567085Z info Pruning shard groups after retention check (start) {"log_id": "0wwHH600000", "service": "retention", "op_name": "retention_delete_check", "op_name": "retention_prune_shard_groups", "op_event": "start"}  
2025-06-04T23:06:39.568197Z info Pruning shard groups after retention check (end) {"log_id": "0wwHH600000", "service": "retention", "op_name": "retention_delete_check", "op_name": "retention_prune_shard_groups", "op_event": "end", "op_elapsed": "7.049ms"}  
2025-06-04T23:06:39.572484Z info Retention policy deletion check (end) {"log_id": "0wwHH600000", "service": "retention", "op_name": "retention_delete_check", "op_event": "end", "op_elapsed": "18.774ms"}
```



PHÁT TRIỂN VÀ TRIỂN KHAI KỸ THUẬT MỚI

Mô hình hệ thống tổng quát

- Nguồn dữ liệu: API OpenWeatherMap (dữ liệu thời tiết thời gian thực)
- Collector: Script Python thu thập dữ liệu định kỳ
- Lưu trữ: InfluxDB lưu nhiệt độ, độ ẩm, áp suất, gió,...
- Cảnh báo: Module kiểm tra ngưỡng và gửi email cảnh báo
- Giao diện Web: Flask + Chart.js hiển thị biểu đồ lịch sử

Thu thập & lưu trữ dữ liệu

- Script `weather_to_influx.py`:
- → Gọi API thời tiết (Hà Nội)
- → Trích xuất dữ liệu cần thiết
- → Ghi vào InfluxDB (measurement = "weather")
- Thu thập định kỳ bằng schedule hoặc chạy vòng lặp



PHÁT TRIỂN VÀ TRIỂN KHAI KỸ THUẬT MỚI

Cảnh báo thời tiết qua Email

- Script alert_email.py:
- → Truy vấn dữ liệu mới nhất từ InfluxDB
- → So sánh ngưỡng cảnh báo (ví dụ: nhiệt độ > 30°C)
- → Gửi email nếu vượt ngưỡng
- Tự động hóa bằng schedule hoặc threading.Timer

Giao diện Web

- Flask tạo giao diện web đơn giản
- Truy vấn dữ liệu InfluxDB
- Biểu đồ hiển thị bằng Chart.js
- Người dùng dễ dàng theo dõi thời tiết lịch sử

Tự động hóa toàn hệ thống:

- → Thu thập, cảnh báo, hiển thị đều chạy nền
- Flask Web: triển khai trên localhost, dễ sử dụng

KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN



influxdb

Kết luận

- Hệ thống đã xây dựng thành công:
- Thu thập dữ liệu thời tiết từ API OpenWeatherMap
- Lưu trữ bằng InfluxDB (thời gian thực, hiệu quả)
- Cảnh báo tự động qua email khi điều kiện bất thường
- Giao diện web hiển thị lịch sử thời tiết bằng biểu đồ
- Hoạt động ổn định trong môi trường cục bộ
- Kết hợp hiệu quả: Python + InfluxDB + Flask
- Giúp sinh viên:
- Hiểu ứng dụng cơ sở dữ liệu thời gian thực
- Rèn luyện kỹ năng tích hợp hệ thống

Hướng phát triển

- Mở rộng vùng theo dõi: Cho phép chọn nhiều thành phố
- Giao diện nâng cao: Dashboard phân tích, nhiều biểu đồ
- Cảnh báo đa kênh: Thêm Telegram, Zalo, SMS
- Triển khai lên Cloud: AWS, Heroku để chạy 24/7
- Phân quyền người dùng: Tài khoản, ngưỡng riêng, cảnh báo cá nhân
- Ứng dụng đa thiết bị: Tương thích điện thoại, máy tính bảng
- ➡ Hệ thống có tiềm năng áp dụng trong nông nghiệp thông minh, cảnh báo thiên tai, hệ thống IoT môi trường.



CẢM ƠN THẦY VÀ CÁC BẠN
ĐÃ LẮNG NGHE