

# Metadata DB service: current status

Andrey Mikhailov

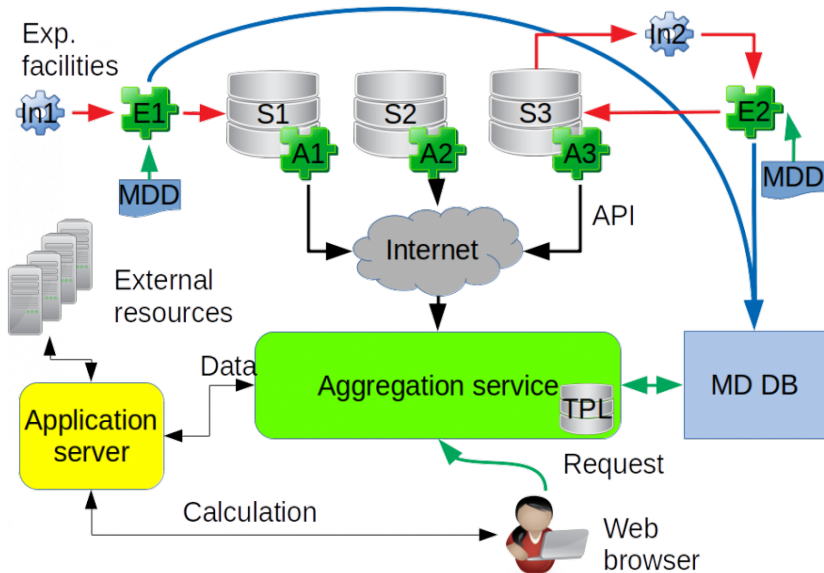
[mikhailov@icc.ru](mailto:mikhailov@icc.ru)

Matrosov Institute for System Dynamics and  
Control Theory of Siberian Branch of Russian Academy of Sciences

Irkutsk, Russia

<http://idstu.irk.ru>

# APPDS Architecture

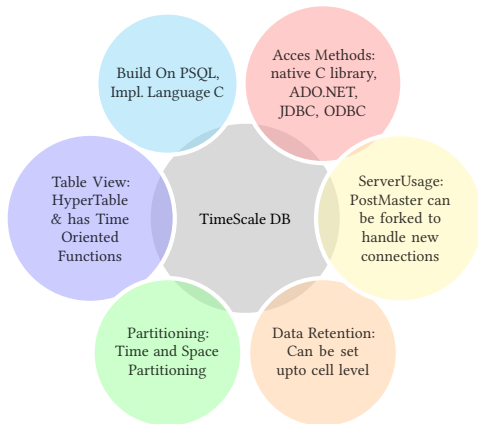


# Development stack



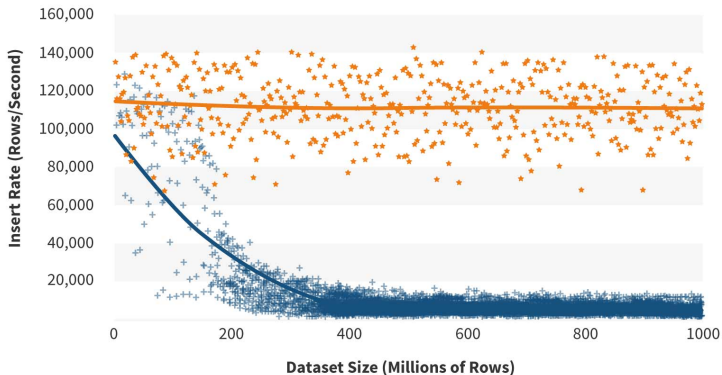
- **Python** - main development language
- **TimeScale** - Database
- **Kaitai** used for metadata extractors
- **GraphQL** - API
- **Git** - version control system
- **Docker** - software container

**TimeScale DB** - A Time-Series database specifically designed for scale, ease of use, complex queries



# TimescaleDB vs. PostgreSQL

## Ingest Rate: PostgreSQL vs. Timescale



+ PostgreSQL

★ TimescaleDB

Insert batch size: 10,000 Rows

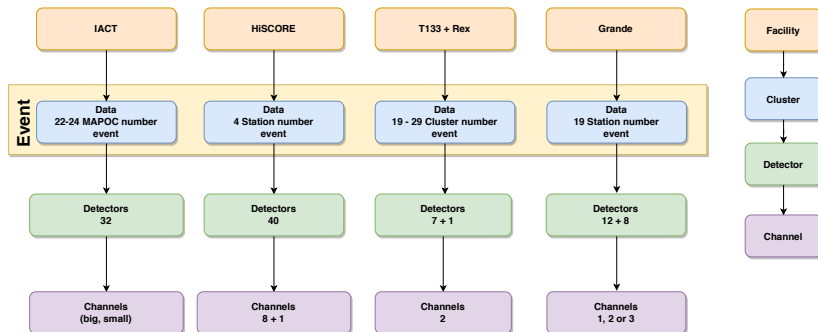
Cache: 16 GB Memory

Final avg. throughput: 5k (PG) vs. 111k (TS)

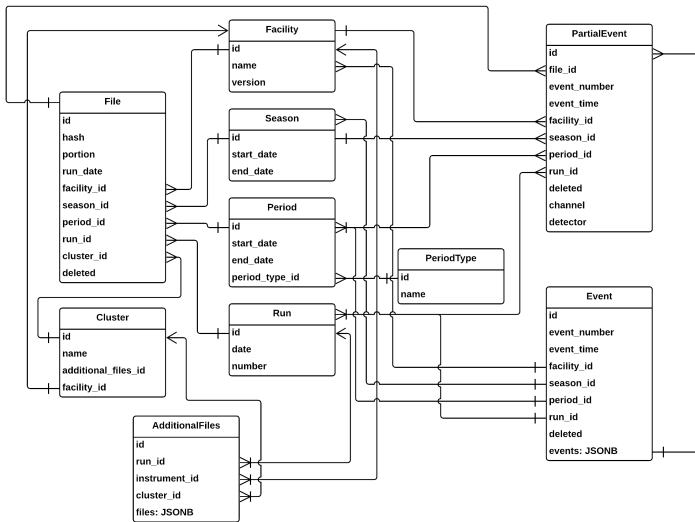
Time to load 1B rows: 37.9h (PG) vs. 2.6h (TS)

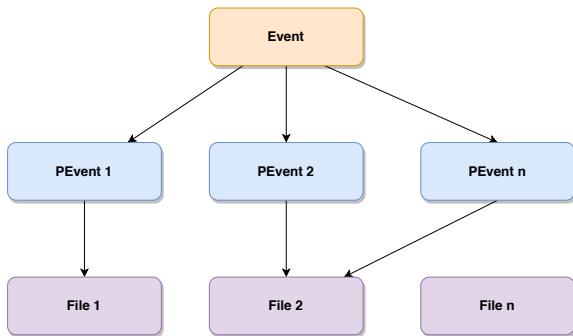
# Facilities metadata hierarchy

- **T133 Event** = 33, 25, 17 or 9 packages per event from all clusters, 2 packages for 1 channel + 1 info package
- **GRANDE Event** = 2 packages from all clusters
- **HiSCORE Event** = 1 package from all detectors
- **IACT Event** = 1 package from all clusters



# Metadata service DB scheme concept







# Extractor

```
def extract_facility(file_path):
    fields = file_path.split('/')
    for field in fields:
        if 'DATA' in field:
            source_pattern = 'DATA(_|\ -) (IACT|HiSCORE|T133\+
                             GRANDE|GRANDE\+ReX)'
            compiled_pattern = re.compile(source_pattern)
            m = compiled_pattern.match(field)
            ...
    return None
def extract_cluster(file_path):
    ...
def process(file_path):
    facility = extract_facility(file_path)
    cluster = extract_cluster(file_path)
    ...
    return {
        'facility': facility,
        ...,
        'cluster': cluster
    }
```

- 1 Metadata schema implementation on TimeScaleDB
- 2 IACT metadata extractor implementation
- 3 Filling metadata database
- 4 HiSCORE, GRANDE, T133 metadata extractors implementation
- 5 Filling metadata database

# Metadata DB service: current status

Andrey Mikhailov

[mikhailov@icc.ru](mailto:mikhailov@icc.ru)

Matrosov Institute for System Dynamics and  
Control Theory of Siberian Branch of Russian Academy of Sciences

Irkutsk, Russia

<http://idstu.irk.ru>