

# MD Catalog - integration with external data storages. Current state

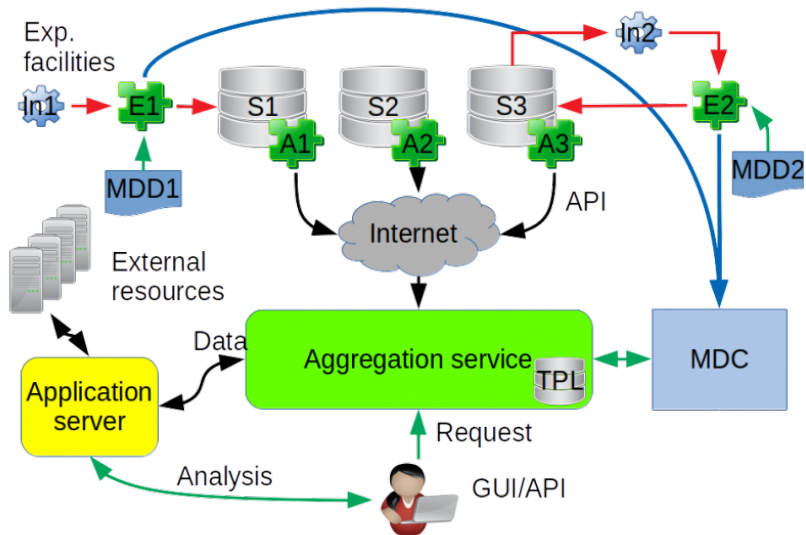
Andrey Mikhailov

[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



Main ideas that are embedded in the architecture:

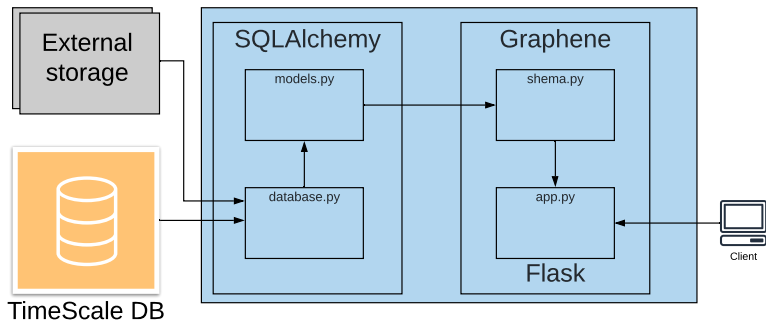
- lack of intervention in local storage, adapters are used for data access
- user requests are processed on a dedicated server based on metadata only

MDC – Metadata catalog. This is a service which supports two main functions:

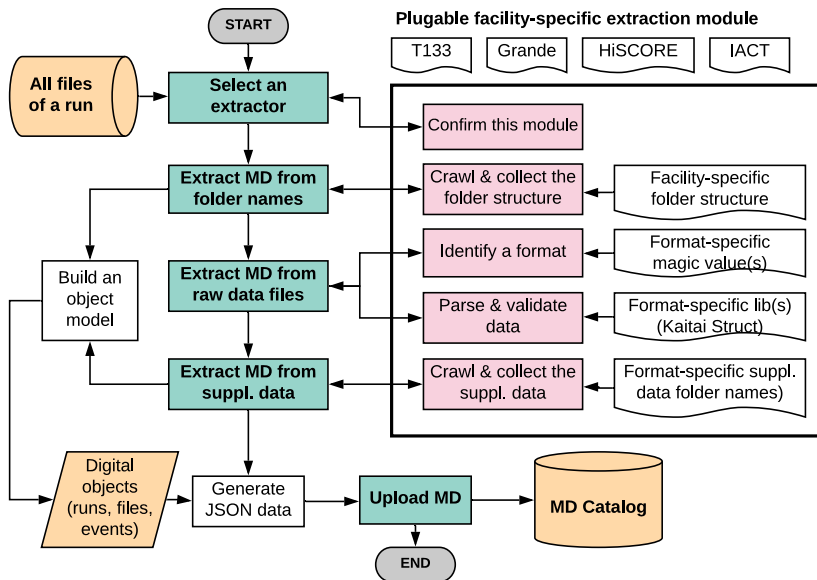
- register collected metadata
- process the user requests for data

This is the single place where the location of necessary data define taking into an account user request

# MDC architecture



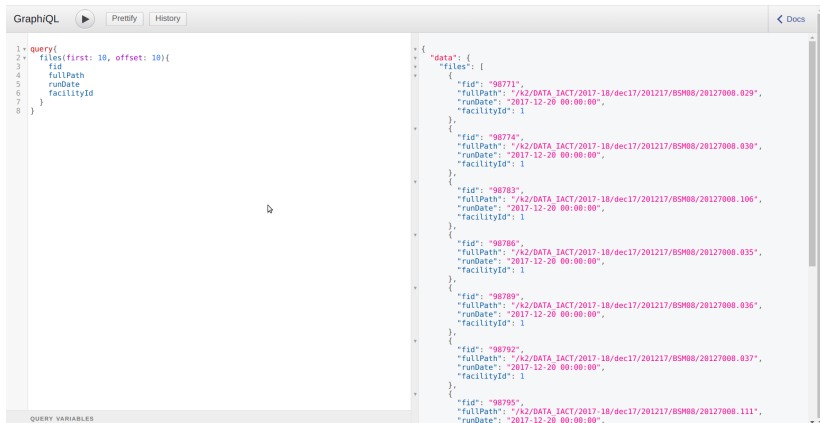
# If raw data is accessed



Extractors for **HiScore** and **IAC**T had implemented

- IACT - 65 528
- HiSCORE - 7 674

```
files (  
  fid: Int  
  facilityId: Int  
  startTime: DateTime  
  endTime: DateTime  
  first: Int  
  offset: Int  
  weatherId: Int  
  trackingSourceId: Int  
  filters: [String]  
): [File]
```



The screenshot shows the GraphiQL interface with a query on the left and its JSON response on the right. The query is:

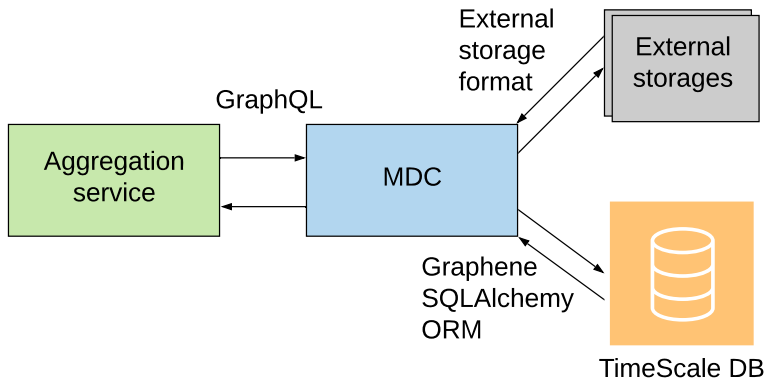
```
1 query{
2   files(first: 10, offset: 10){
3     fid
4     fullPath
5     runDate
6     facilityId
7   }
8 }
```

The response is a JSON object with a "data" field containing an array of file objects:

```
{
  "data": {
    "files": [
      {
        "fid": "98771",
        "fullPath": "/k2/DATA_IAC2/2017-18/dec17/201217/BSM08/20127008.029",
        "runDate": "2017-12-20 00:00:00",
        "facilityId": 1
      },
      {
        "fid": "98774",
        "fullPath": "/k2/DATA_IAC2/2017-18/dec17/201217/BSM08/20127008.030",
        "runDate": "2017-12-20 00:00:00",
        "facilityId": 1
      },
      {
        "fid": "98783",
        "fullPath": "/k2/DATA_IAC2/2017-18/dec17/201217/BSM08/20127008.106",
        "runDate": "2017-12-20 00:00:00",
        "facilityId": 1
      },
      {
        "fid": "98786",
        "fullPath": "/k2/DATA_IAC2/2017-18/dec17/201217/BSM08/20127008.035",
        "runDate": "2017-12-20 00:00:00",
        "facilityId": 1
      },
      {
        "fid": "98789",
        "fullPath": "/k2/DATA_IAC2/2017-18/dec17/201217/BSM08/20127008.036",
        "runDate": "2017-12-20 00:00:00",
        "facilityId": 1
      },
      {
        "fid": "98792",
        "fullPath": "/k2/DATA_IAC2/2017-18/dec17/201217/BSM08/20127008.037",
        "runDate": "2017-12-20 00:00:00",
        "facilityId": 1
      },
      {
        "fid": "98795",
        "fullPath": "/k2/DATA_IAC2/2017-18/dec17/201217/BSM08/20127008.111",
        "runDate": "2017-12-20 00:00:00",
        "facilityId": 1
      }
    ]
  }
}
```

At the bottom left, there is a section labeled "QUERY VARIABLES" which is currently empty.

# If raw data isn't accessed



MDC know all available sources and external storage API formats



# Request. GraphQL to JSON-RPC 2.0

```
query{
  files(
    facilityId: 5,
    startTime: "2017-12-20T00:00:00"
    endTime: "2017-12-21T00:00:00"
  ){
    url
    facilityId
    __typename
  }
}
```

```
{
  "id": "cf919bb3-7028-4238-9438-2601cbfde3a6",
  "jsonrpc": "2.0",
  "method": "new_task",
  "parameters": {
    "type": [ 'kascade' ],
    "datetime_min": 1256481198,
    "datetime_max": 1256481201
  }
}
```

# Response. JSON-RPC 2.0 to GraphQL

```
{
  "jsonrpc": "2.0",
  "result": {
    "task_id": 111
  },
  "id": "cf919bb3-7028-4238-9438-2601cbfde3a6"
}
```

```
{
  "url": "http://141.52.67.147:55000/api/data/cf919bb3-7028-4238-9438-2601cbfde3a6.txt.bz2",
  "facilityId": 5,
  "__typename": "File"
}
```

- Extractors for IACT and HiScore are implemented
- Metadata extracted and stored
  - IACT - 65 528
  - HiSCORE - 7 674
- API for querying is implemented
- MDC integrated with KCDC storage

# MD Catalog - integration with external data storages. Current state

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>