

EIDA WFCatalog Service

Luca Trani and the EIDA Team

- Provides a well defined API to query for seismic waveform metadata (including QC)
- Enables continuous waveforms discovery based on metadata => no unnecessary data downloads
- Preliminary analysis and processing of seismic waveforms moved to data centers => big advantage for users who can quickly browse data and related features
- More than just QC... the waveform metadata catalog offers functionalities to discover the best waveform data matching a variety of criteria combining multiple quality metrics and metadata extracted from MSEED files
- Based on a set of clear and well defined quality metrics which together with the API have been discussed and agreed among the major data centers in Europe

Example of query:

<http://orfeus-eu.org/ws/wfcatalog/query?starttime=2015-06-01&endtime=2015-06-02&sta=HGN&cha=BHZ>

```
[{"net":"NL","sta":"HGN","cha":"BHZ","loc":"02","sample_rate":40.0,"record_length":512,"quality":"D","num_gaps":0,"num_overlaps":0,"gaps_len":0,"overlaps_len":0,"sample_max":4319,"sample_min":-3115,"sample_rms":1039.382142,"sample_mean":2563.56,"sample_stddev":1016.847836,"ms_glitches":0,"ms_amplifier_saturation":0,"ms_digital_filter_charging":1,"ms_digitizer_clipping":0,"ms_missing_padded_data":0,"ms_spikes":1,"ms_suspect_time_tag":0,"ms_telemetry_sync_error":0,"ms_calibration_signal":0,"ms_event_begin":0,"ms_event_end":0,"ms_event_in_progress":0,"ms_timing_correction":0,"ms_clock_locked":0,"percent_availability":100.0,"start_time":"2015-05-..."}]
```

QC metrics presented and discussed in the FDSN WGII meeting

The selection of metrics and their definitions adopted in the WFCatalog result from the integration and harmonisation of:

- metrics produced within the NERA EC project and discussed in EIDA with broad consensus among the major european data centers
- metrics currently present in MUSTANG

Metric Name	Description
num gaps	Duration of gaps
num overlaps	Duration of overlaps
sample max	Maximum sample value
sample rms	RMS sample value
...	...

EIDA provides a complete package to extract and manage QC and wf metadata including:

- formal definitions and service specification
- software stack for computation of metrics
- framework for wf metadata management: including repository and web service layer.

The specification of the WFCatalog is the result of a joint collaboration among EIDA partners

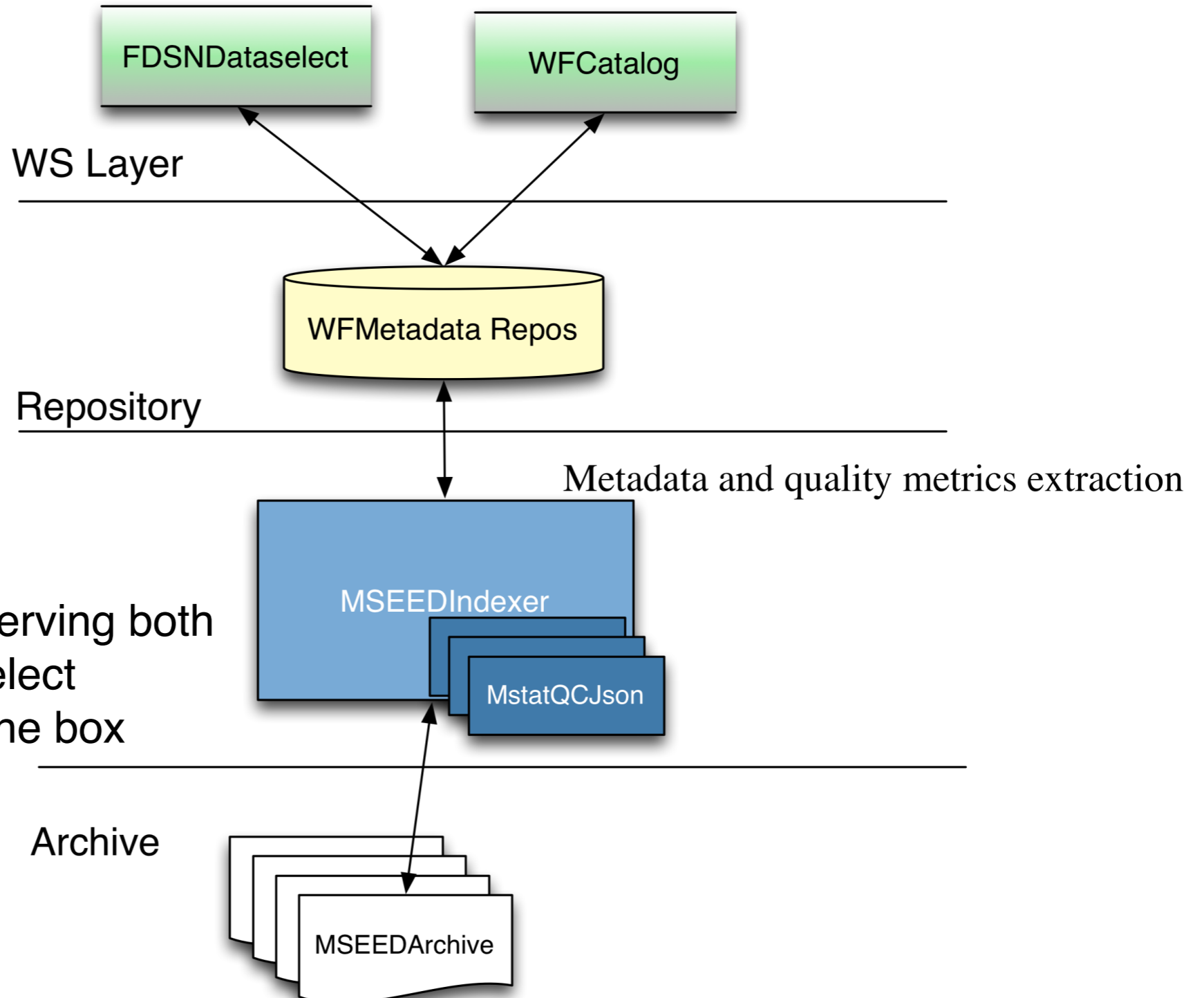
The design has been guided by the following principles:

- Maximise compliance with existing FDSN standard services like dataselect
- Facilitate compliance/interoperability with existing systems like MUSTANG
- Provide flexibility to address different use cases
- Enable future extensions e.g: number of metrics and different granularities
- Enable integration within the broader EIDA NG architecture

Prototype service currently in operation/testing at different datacenters

- scalable and flexible architecture based on a few key components and selected technologies

- “big data” ready



@ODC Metadata repository serving both WFCatalog and FDSN Dataselect
=> filtering extensions out of the box

Technological investigation started in 2012

Experiments carried on with different DB technologies: MonetDB, MySql, NoSQL
(Cassandra, CouchDB, **MongoDB**)

Main requirements: performance, scalability, efficient big data handling, flexibility and extensibility

Chosen technology addresses efficiently the desired requirements : flexible, scalable, robust and highly performant also in single node mode.

Other features:

- efficient caching system
- schema-less => collections can be dynamically modified
- multi language support
- built-in json API
- powerful indexing features
- support for high availability and clustering
- active community and support