



# AKKA

## *Insights into the Technical Aspects of the InGeoCloudS platform*

### *ElasticDB and DataImport Services*

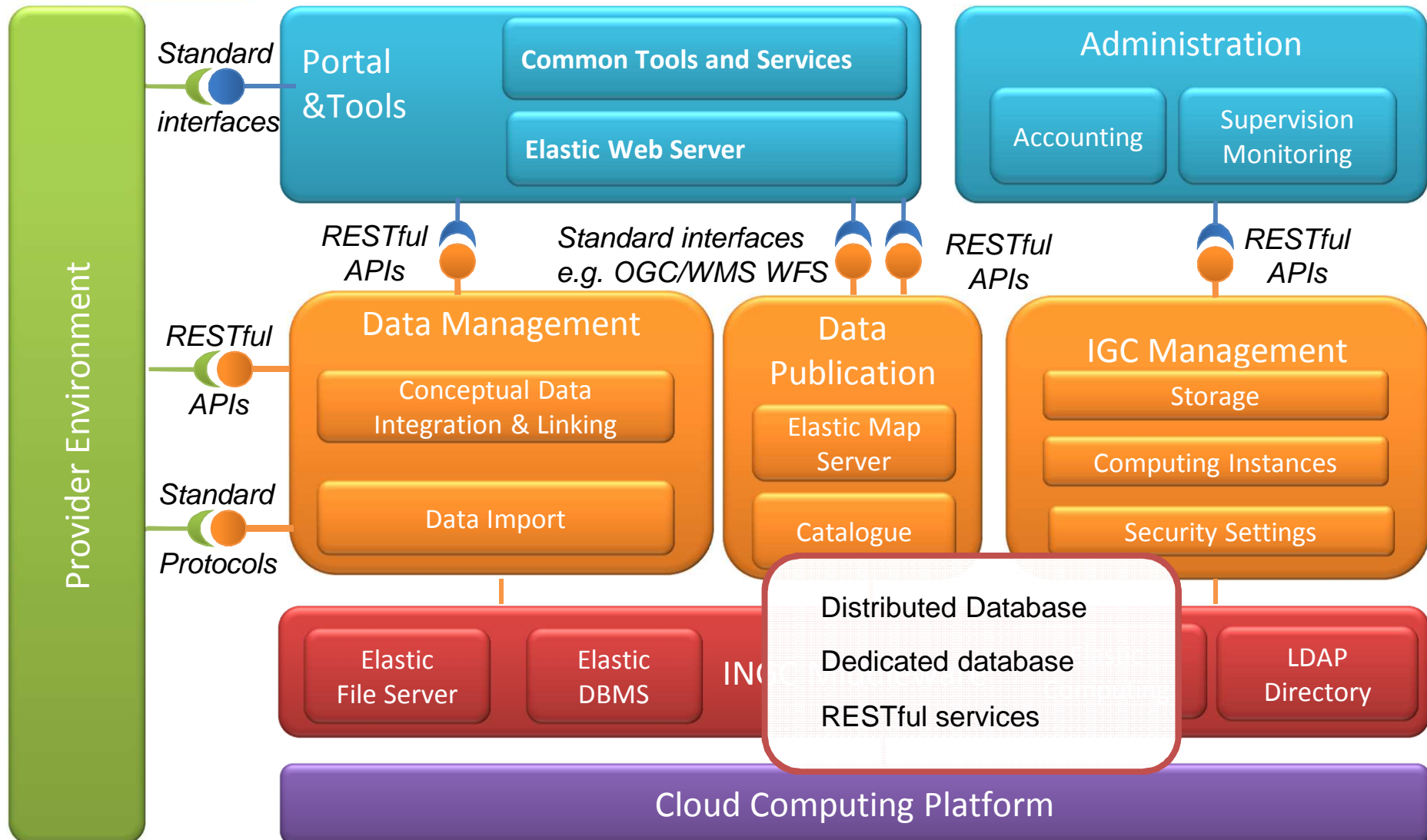
Experts Workshop

Nov. 21st, 2013



**InGeoCloudS**  
Inspired GEOdata CLOUD Services

# ElasticDB



- Elastic database system:
  - Scalable storage capacity,
  - High availability of data and performance.
- Relies on PostgreSQL/PostGIS and Pgpool-II:
  - Most used solution by data providers, and in GIS domain
  - Replication of data: Hot-Standby with Streaming replication
  - Load balancing by Pgpool-II.
- Scalable solution in the cloud:
  - Number of EC2 instances adjusted according to the needs



**InGeoCloudS**  
Inspired GEOdata CLOUD Services



# ElasticDB

**Load Balancing**

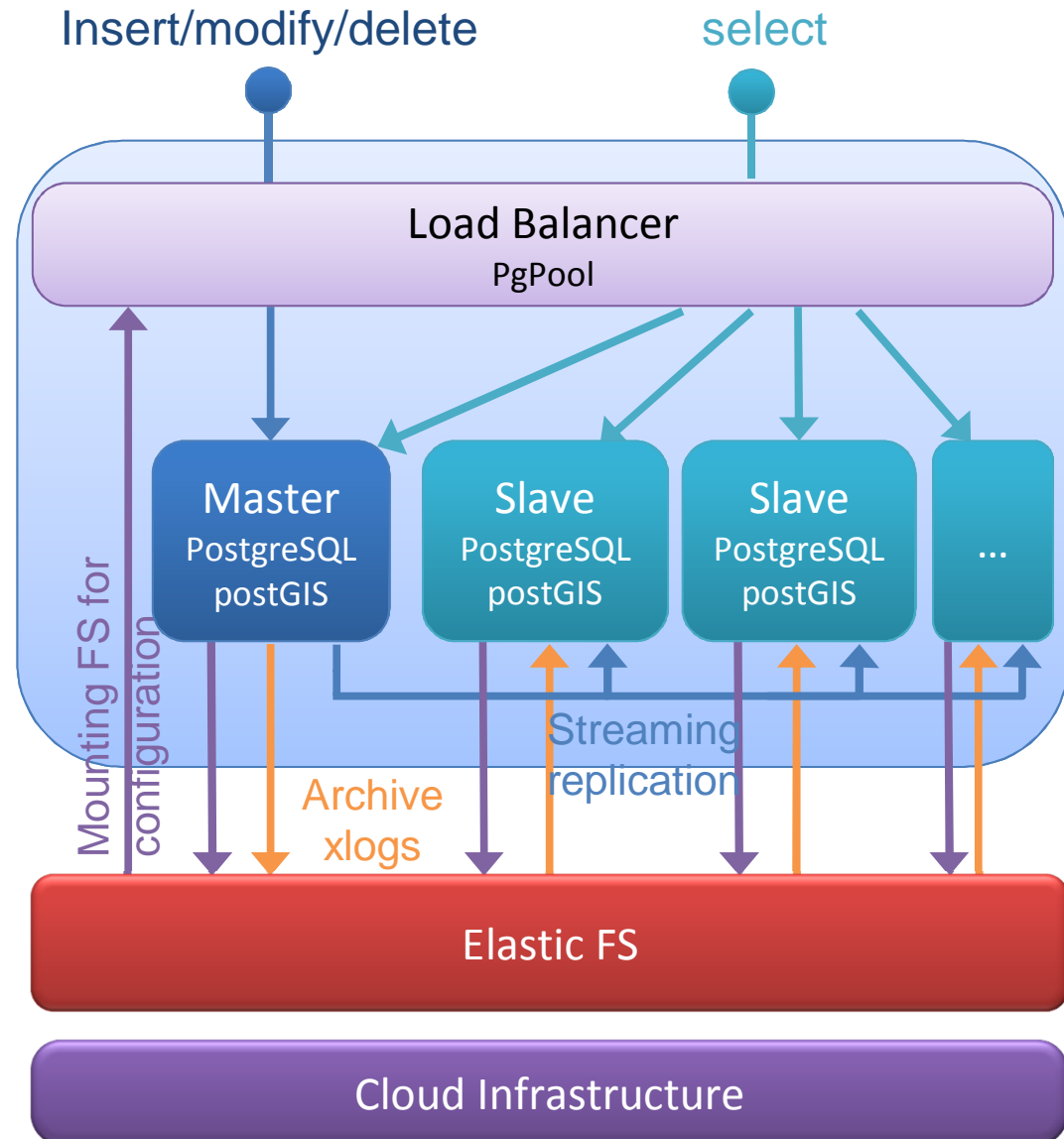
- Modification on master.
- Select on any node.

**Replication of Data**

- Master create Archive logs,
- New slaves first read Archive logs,
- Then register for streaming replication,
- Master send in live modifications to registered slaves

**Management and Monitoring**

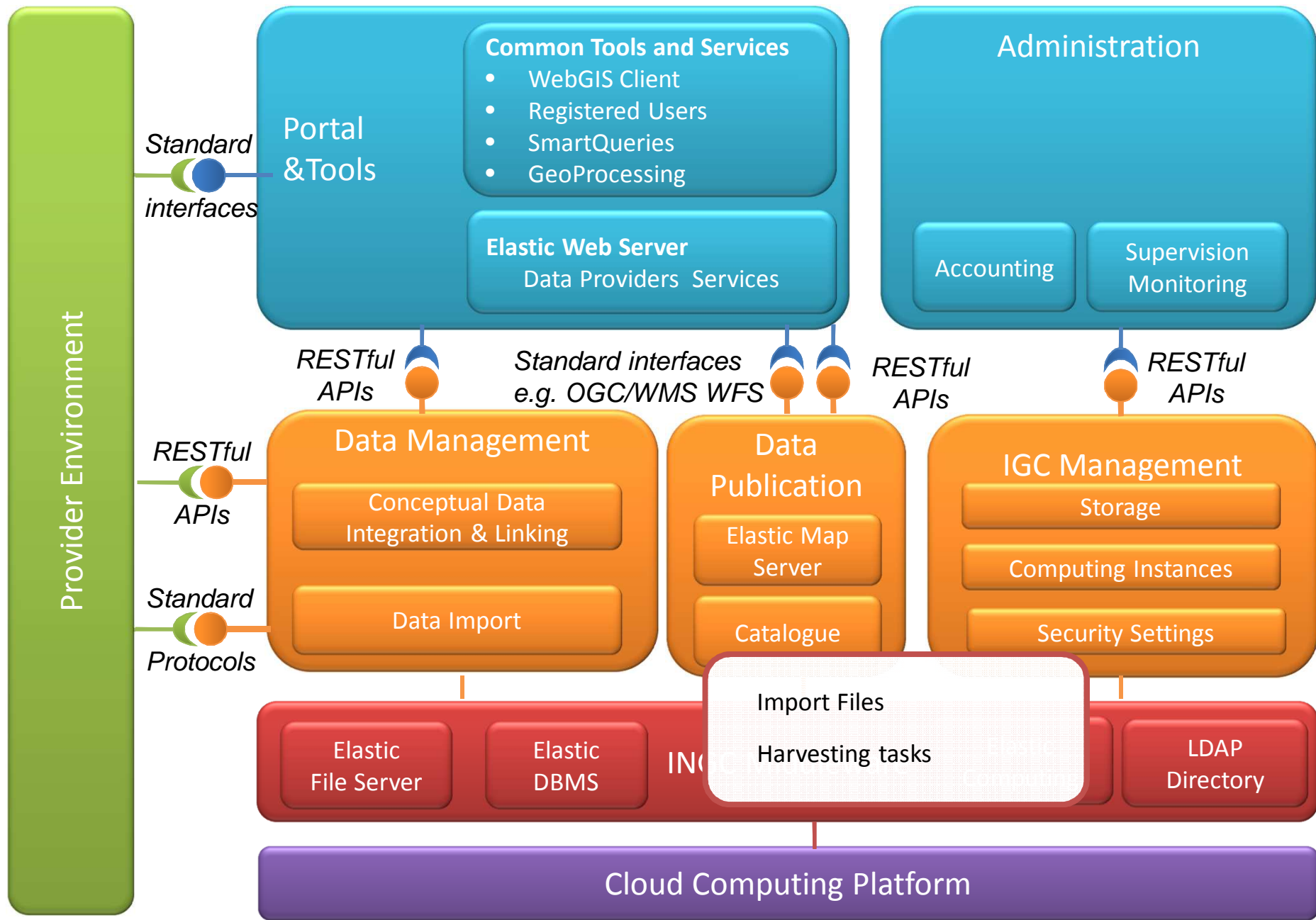
- Shared configuration



- Automatic failover of the master:
  - Faulty master is removed from the pgpool servers,
  - 1<sup>st</sup> slave becomes the new master.
  - Other slaves register to the new master for streaming replication.
- Automatic failover of the Pgpool:
  - Relies on Watchdog feature of Pgpool: 2 instances (1 active and 1 backup), and 1 IP address shared by the 2 instances.
  - The backup Pgpool monitors the active Pgpool,
  - The backup Pgpool becomes active when the first one fails,

- RESTful services for instance management:
  - [Start/stop the ElasticDB \(Administrator\)](#)
  - [Add a new slave or a new pgpool \(Administrator\)](#)
- RESTful services for monitoring:
  - [Check the ElasticDB health \(Administrator\)](#)
  - [Get detailed status on the ElasticDB \(Administrator\)](#)
- RESTful services for database maintenance:
  - [Backup/restore the entire database \(Administrator\)](#)
  - [Rebuild indexes and optimize disk space \(Administrator\)](#)

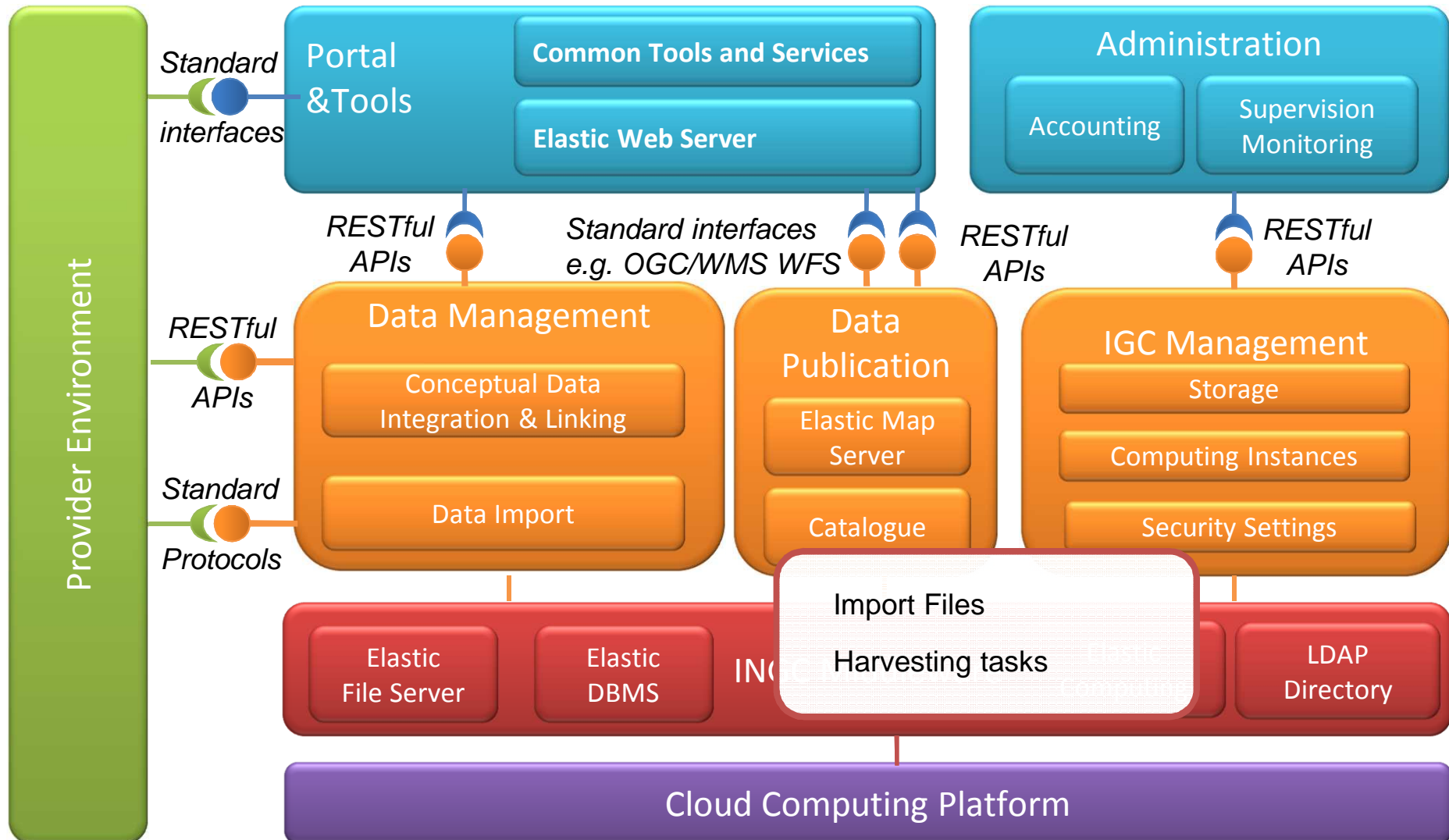
- RESTful services for database management:
  - Create/delete a dedicated database (and dedicated users) for a provider (Administrator)
  - Get information on a provider's database (Administrator, Data Provider)
  - Backup/restore a provider's database (Administrator, Data Provider)
- RESTful services for database access:
  - Get connection information on a provider's database (Administrator, Data provider)
  - Get connection information on entire database (Administrator)





**InGeoCloudS**  
Inspired GEOdata CLOUD Services

# Data Import



# Import Files

- RESTful services to import files in workspace:
  - Upload/download entire or part of a file (Data Provider)\*
  - Create/List a folder (Data Provider)\*
  - Copy/move/delete a file or a folder (Data Provider)\*
  - Get metadata on a file or a folder (Data Provider)\*
  - Archive/restore files or folders (Data Provider)\*
  - [Get a media link to a file or a folder \(Data Provider\)](#)
- Import Files application provided the Data Provider Toolkit\*

(\* ) not yet available

# Import Files

- Media link supported protocols:
  - FTP and FTPS,
  - We plan to support SFTP and SCP.
- Media links are secured:
  - Period of validity is about 24 hours,
  - Authentication is required,
  - User is rooted to its workspace.

# Harvesting

- RESTful services to manage harvesting tasks:
  - CRUD operation on tasks (Data Provider)
- Harvesting application provided the Data Provider Toolkit