PostgeSQL with R

John Dey jfdey@fredhutch.org

Feb 2, 2017

Database as a Service - DBaaS

On premises resource for hosting Postgres instances

Fully managed containerized instance of Postgres

Full root access to Postgres
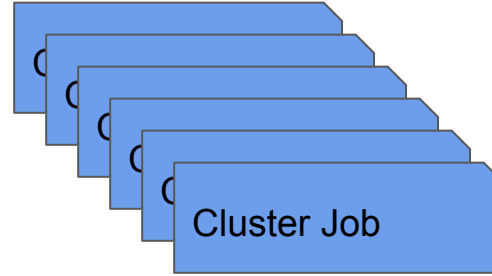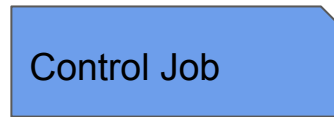
https://mydb.fredhutch.org/

Database vs. File System

Database Advantages

- Concurrent IO
- Data Access
- Management of Cluster Jobs
- Automate Workflows (pipelines)
- Configuration
- Repeatability

Manage Cluster Jobs

Control Table

Data file	Job_ID	Job_Num	Start	End	Results

Control Job

Cluster Job

Result Tables

Database Access

Keep Postgres passwords in .pgpass file located in /home/user

File permissions read write by owner only. chmod 600

Do not keep passwords in scripts or pass on the command line

.pgpass is used by PGadminIII on OSX

Postgres Language Support

Shell

    module load PostgreSQL

R

    module load R/3.3.2-foss-2016b-fh2

    library(**RPostgreSQL**)

Python

    module load Python/2.7.13-foss-2016b

    import **psycopg2**

Demo

Create DB with MyDB

Database Administration

pi.R