

# Dynamic Job Mapping in Galaxy Simplified

Eric Enns

Public Health Agency of Canada

Presented at GCC2015

F I GB USA MEX

**ONE SIZE  
DOES NOT  
FIT ALL**



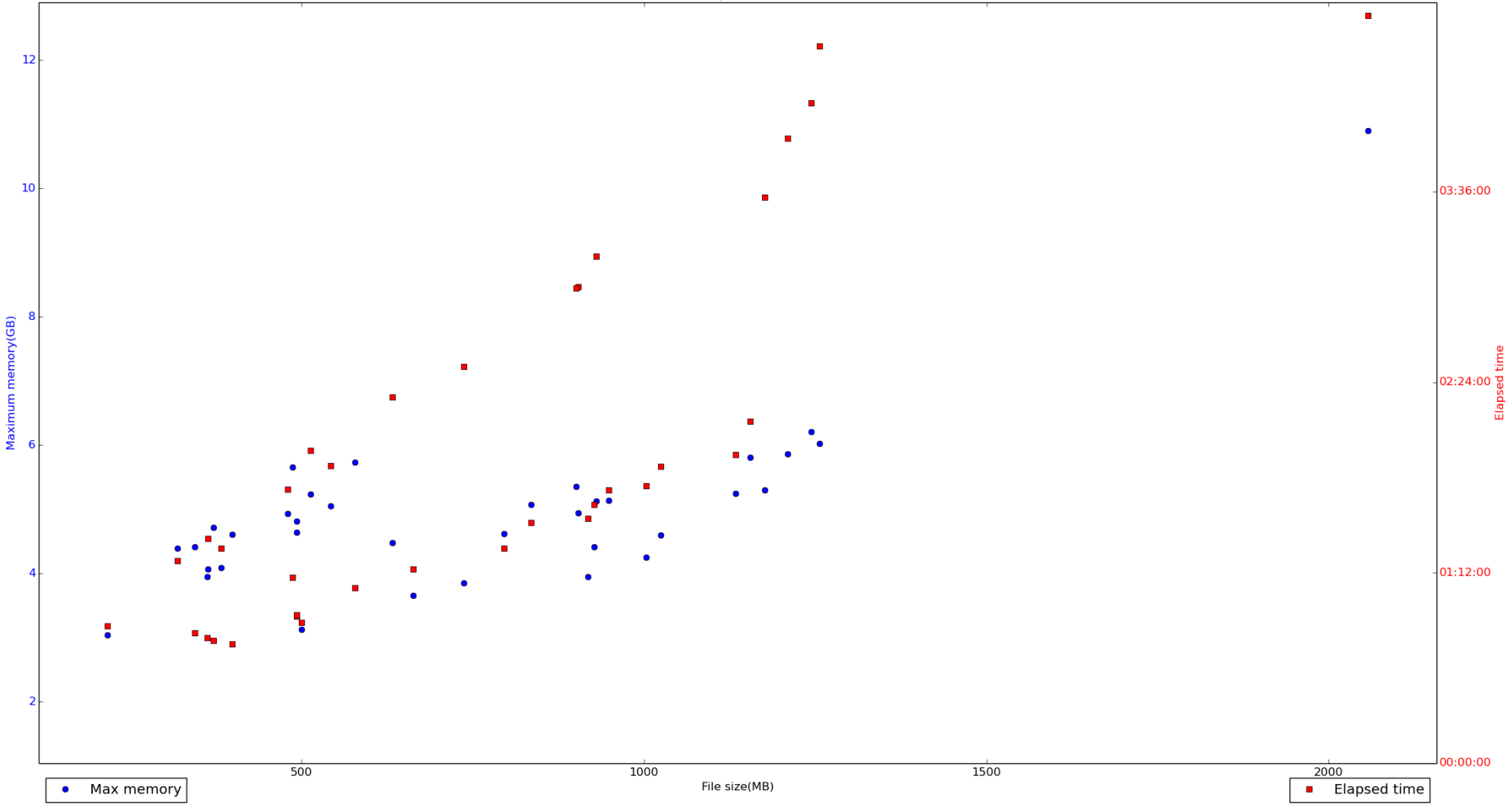
Von links bügeln/ iron inside out/  
repasser sur l'envers /只熨反面

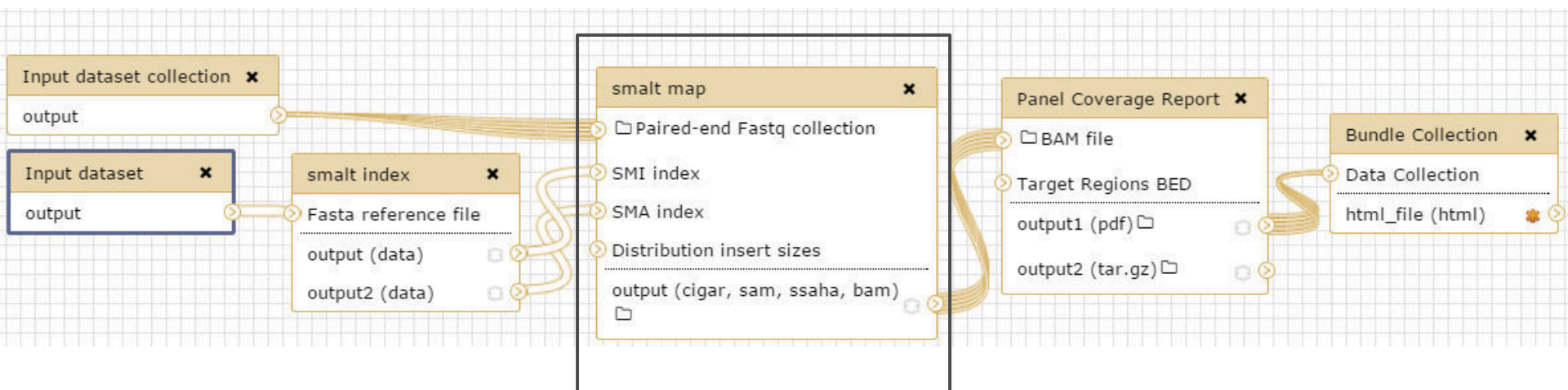


OR



spades





# jobConfig.py

- Provides `configure_job` function
- Maps jobs to destinations or adjusts params or induces failures via conditions
  - `filesize (lbound, hbound)`
  - `records (lbound, hbound)`
- Maps jobs to “`dynamic_default`” if no rule defined

# Job\_conf.xml

```
<destinations default="dynamic_job"/>
  <destination id="dynamic_job" runner="dynamic">
    <param id="type">python</param>
    <param id="function">configure_job</param>
  </destination>
  <destination id="dynamic_default" runner="drmaa">
</destinations>
```

# job\_options.xml

```
<jobs>
  <job name="spades">
    <condition type="filesize" lbound="0" hbound="10 MB"
      err_msg="Too few reads for spades to work" param="fail" />
    <condition type="filesize" lbound="4 GB" hbound="-1"
      err_msg="Too much data, shouldn't run" param="fail" />
    <condition type="filesize" lbound="10 MB" hbound="2 GB"
      param="-q test.q -pe galaxy 24 -l h_vmem=15G" />
    <condition type="filesize" lbound="2 GB" hbound="4 GB"
      param="-q test.q -pe galaxy 48 -l h_vmem=96G" />
  </job>
</jobs>
```



# Future Work

- Integrate auto re-run of failed jobs due to resource problems
  - i.e. not enough memory
- Add more condition types
  - i.e. command line arguments set

# Acknowledgements

Daniel Bouchard

Philip Mabon