Scaling bio-analyses from computational clusters to grids

George Byelas
University Medical Centre Groningen, the Netherlands
IWSG-2013, Zürich, Switzerland, June 3rd, 2013
Content

- Bio-workflows
- Workflow
  - Modeling
  - Job Generation
  - Tool deployment
  - Data management
  - Workflow execution
  - Implementation detail
  - Recent developments
- Conclusion and further steps
Bio-workflows
Example: NGS alignment workflow

- **HiSeq**
- **Raw data**
  - 10-100 samples
- **Alignment workflow**
- **Result data**
  - 20 – 200 days
  - 80 – 800 GB

Per Project:
1. Aligned reads
2. QC-reports
3. SNP lists
Alignment & SNP calling workflow

31 steps, ≥ 2 days per sample

• Input
  • Analysis protocols
  • Sample DNA data
  • Reference DNA data

• Analysis
  • Scripts are generated and executed

• Output
  • Aligned DNA and QC reports
An analysis job (script) generated from a protocol

#!/bin/bash
#PBS -q test
#PBS -l nodes=1:ppn=4
#PBS -l walltime=08:00:00
#PBS -l mem=6gb
#PBS -e $GCC/test_compute/projects/batch4/intermediate/test1/err/err_test1_BwaElement1A102a_FC81D90ABXX_L7.err
#PBS -o $GCC/test_compute/projects/batch4/intermediate/test1/out/out_test1_BwaElement1A102a_FC81D90ABXX_L7.out
mkdir -p $GCC/test_compute/projects/batch4/intermediate/test1/err
mkdir -p $GCC/test_compute/projects/batch4/intermediate/test1/out
printf "test1_BwaElement1A102a_FC81D90ABXX_L7_started " >>$GCC/test_compute/projects/batch4/intermediate/test1/log_test1.txt
date "+DATE: %m/%d/%yTIME: %H:%M:%S" >>$GCC/test_compute/projects/batch4/intermediate/test1/log_test1.txt
date "+start time: %m/%d/%ytH:%M:%S" >>$GCC/test Compute/projects/batch4/intermediate/test1/
test1_BwaElement1A102a_FC81D90ABXX_L7.txt
echo running on node: `hostname` >>$GCC/test_compute/projects/batch4/intermediate/test1/
test1_BwaElement1A102a_FC81D90ABXX_L7.txt
/target/gpfs2/gcc/tools/bwa-0.5.8c_patched/bwa aln 
/target/gpfs2/gcc/resources/hg19/indices/human_g1k_v37.fa 
$GCC/test_compute/projects/batch4/rawdata/110121_I288_FC81D90ABXX_L7_HUMrutRGADIAAPE_1.fq.gz 
-t 4 
-f $GCC/test_compute/projects/batch4/intermediate/A102a_110121_I288_FC81D90ABXX_L7_HUMrutRGADIAAPE_1.fq.gz.sai
printf "test1_BwaElement1A102a_FC81D90ABXX_L7_finished " >>$GCC/test_compute/projects/batch4/intermediate/test1/log_test1.txt
date "+finish time: %m/%d/%ytH:%M:%S" >>$GCC/test_compute/projects/batch4/intermediate/test1/
test1_BwaElement1A102a_FC81D90ABXX_L7.txt
date "+DATE: %m/%d/%yTIME: %H:%M:%S" >>$GCC/test_compute/projects/batch4/intermediate/test1/log_test1.txt
Imputation workflow

- **Imputation:**

  
  \[
  \text{GWAS data} \quad \begin{array}{c}
  \text{A} \\
  \text{A} \\
  \text{A}
  \end{array}
  \begin{array}{c}
  \_ \\
  \_ \\
  \_ \\
  \end{array}
  \text{G} \quad \begin{array}{c}
  \_ \\
  \_ \\
  \_ \\
  \end{array}
  \text{T}
  \]

  
  \[
  \text{Reference data} \quad \begin{array}{c}
  \text{A} \\
  \text{C} \\
  \text{A} \\
  \text{A}
  \end{array}
  \begin{array}{c}
  \text{C} \\
  \text{A} \\
  \text{A} \\
  \_ \\
  \end{array}
  \begin{array}{c}
  \text{G} \\
  \text{T} \\
  \text{T} \\
  \_ \\
  \end{array}
  \]

  
  \[
  \text{Imputed data} \quad \begin{array}{c}
  \text{A} \\
  \text{A} \\
  \text{A}
  \end{array}
  \begin{array}{c}
  \text{C} \\
  \text{A} \\
  \text{A} \\
  \text{T} \\
  \end{array}
  \text{G} \quad \begin{array}{c}
  \text{T} \\
  \text{T} \\
  \text{T} \\
  \_ \\
  \end{array}
  \text{T}
  \]

- **Number of jobs**

  \[
  \text{\textit{number\_jobs} = \sum_{\text{chr}=1..22} \frac{\text{Length}_{\text{chr}}}{5\_\text{megabase}} \times \frac{\text{number\_of\_samples}}{500}}
  \]

- **One run:**
Bio-workflow complexity

- Many analysis steps
  - Many analysis jobs
  - Different analysis tools and their dependencies
- Large various data involved
- Heterogeneous resources
Workflow design and generation
MOLGENIS approach

- Model
- Generate
- Use

Analyses...

Species...

Projects...
- Jobs are generated from the model
- Every job has an analysis target (*e.g.* Genome region)
Generates jobs (scripts) from model described in files

Suitable for workflows (PBS cluster) and single jobs (gLite grid)
Database solution with MOLGENIS software toolkit (1)

Model (xml)

Workflow

NextGenSeq

Model organisms

Generator (java)

Use (web)

Molgenis/compute

Solexa Sequencer LIMS

Animal Observatory
Database solution with MOLGENIS software toolkit (2)

Model (xml)
- Model
  - workflow.xml / 100 loc
  - ui.xml / 25 loc
- NextGenSeq

Workflow

Generator (java)

Use (web)
- Molgenis/compute
- Solexa Sequencer LIMS
- Animal Observatory

Model organisms / 25 loc

Generate
- *.sql / 1722 loc
- *.java / 46639 loc

1 : 400!
Workflow design view in the generated Molgenis web-UI

**Compute Testing**

<table>
<thead>
<tr>
<th>Workflow Element</th>
<th>Workflow Element</th>
</tr>
</thead>
<tbody>
<tr>
<td>File ➤ Edit ➤ View</td>
<td>File ➤ Edit ➤ View</td>
</tr>
<tr>
<td>id</td>
<td>id</td>
</tr>
<tr>
<td>name</td>
<td>name</td>
</tr>
<tr>
<td>description</td>
<td>description</td>
</tr>
<tr>
<td>protocol</td>
<td>protocol</td>
</tr>
<tr>
<td>PreviousSteps</td>
<td>PreviousSteps</td>
</tr>
</tbody>
</table>

**Workflow Step:**

<table>
<thead>
<tr>
<th>id</th>
<th>name</th>
<th>Workflow</th>
<th>protocol</th>
<th>PreviousSteps</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>impute2_s00</td>
<td>workflowImpute.csv</td>
<td>preparePedMapForImpute2.ftl</td>
<td>impute2_s00</td>
</tr>
<tr>
<td>2</td>
<td>impute2_s01</td>
<td>workflowImpute.csv</td>
<td>convertPedMapToTrnTyper.ftl</td>
<td>impute2_s01</td>
</tr>
<tr>
<td>3</td>
<td>impute2_s02</td>
<td>workflowImpute.csv</td>
<td>prepareStudy.ftl</td>
<td>impute2_s02</td>
</tr>
<tr>
<td>4</td>
<td>impute2_s03</td>
<td>workflowImpute.csv</td>
<td>convertPreparedStudyPedMapToGen.ftl</td>
<td>impute2_s03</td>
</tr>
<tr>
<td>5</td>
<td>impute2_s04</td>
<td>workflowImpute.csv</td>
<td>imputeWithimpute2.ftl</td>
<td>impute2_s04</td>
</tr>
<tr>
<td>6</td>
<td>impute2_s05</td>
<td>workflowImpute.csv</td>
<td>concatImpute2ResultsPerChr.ftl</td>
<td>impute2_s05</td>
</tr>
<tr>
<td>7</td>
<td>impute2_s06</td>
<td>workflowImpute.csv</td>
<td>convertConcatImpute2ResultToPedMap.ftl</td>
<td>impute2_s06</td>
</tr>
<tr>
<td>8</td>
<td>impute2_s07</td>
<td>workflowImpute.csv</td>
<td>calculateBingleR2ForImpute2Results.ftl</td>
<td>impute2_s07</td>
</tr>
<tr>
<td>9</td>
<td>impute2_s08</td>
<td>workflowImpute.csv</td>
<td>convertImputeResultToPedMapTo836.ftl</td>
<td>impute2_s08</td>
</tr>
</tbody>
</table>

**Analysis Protocol:**

- preparePedMapForImpute2.ftl
- convertPedMapToTrnTyper.ftl
- prepareStudy.ftl
- convertPreparedStudyPedMapToGen.ftl
- imputeWithimpute2.ftl
- concatImpute2ResultsPerChr.ftl
- convertConcatImpute2ResultToPedMap.ftl
- calculateBingleR2ForImpute2Results.ftl
- convertImputeResultToPedMapTo836.ftl

**Previous Steps:**

- impute2_s00
- impute2_s01
- impute2_s02
- impute2_s03
- impute2_s04
- impute2_s05
- impute2_s06
- impute2_s07
- impute2_s08
## Workflow run-time view (analysis jobs)

### Task ID: BwaAlignLeft_run_01_1355507698165482000

**Running on node:** wn-car-004.farm.nikhef.nl

**srmpc**:

```bash
srmpc --server-mode=passive srm://srm.grid.sara.nl/pnfs/grid.sara.nl/data/bbmrni/RP2/resources/hg19/indices/human_g1k_v37.chr1.fa
file:///tmp/jobdir/28851242.stro.nikhef.nl/resources/hg19/indices/human_g1k_v37.chr1.fa
```

**Picked up JAVA_OPTIONS:** `-Xmx512M`

**server_mode** is specified as passive, setting number of streams to 1

```bash
srmpc --server-mode=passive srm://srm.grid.sara.nl/pnfs/grid.sara.nl/data/bbmrni/RP2/resources/hg19/indices/human_g1k_v37.chr1.fa.amb
file:///tmp/jobdir/28851242.stro.nikhef.nl/resources/hg19/indices/human_g1k_v37.chr1.fa.amb
```

**Picked up JAVA_OPTIONS:** `-Xmx512M`

**server_mode** is specified as passive, setting number of streams to 1

```bash
srmpc --server-mode=passive srm://srm.grid.sara.nl/pnfs/grid.sara.nl/data/bbmrni/RP2/resources/hg19/indices/human_g1k_v37.chr1.fa.amb
```

**Picked up JAVA_OPTIONS:** `-Xmx512M`

**server_mode** is specified as passive, setting number of streams to 1

```bash
srmpc --server-mode=passive srm://srm.grid.sara.nl/pnfs/grid.sara.nl/data/bbmrni/RP2/resources/hg19/indices/human_g1k_v37.chr1.fa.bwt
file:///tmp/jobdir/28851242.stro.nikhef.nl/resources/hg19/indices/human_g1k_v37.chr1.fa.bwt
```

**Picked up JAVA_OPTIONS:** `-Xmx512M`

**server_mode** is specified as passive, setting number of streams to 1

```bash
srmpc --server-mode=passive srm://srm.grid.sara.nl/pnfs/grid.sara.nl/data/bbmrni/RP2/resources/hg19/indices/human_g1k_v37.chr1.fa.bwt
```

**Picked up JAVA_OPTIONS:** `-Xmx512M`

**server_mode** is specified as passive, setting number of streams to 1

```bash
srmpc --server-mode=passive srm://srm.grid.sara.nl/pnfs/grid.sara.nl/data/bbmrni/RP2/resources/hg19/indices/human_g1k_v37.chr1.fa.pac
file:///tmp/jobdir/28851242.stro.nikhef.nl/resources/hg19/indices/human_g1k_v37.chr1.fa.pac
```

**Picked up JAVA_OPTIONS:** `-Xmx512M`

**server_mode** is specified as passive, setting number of streams to 1

```bash
srmpc --server-mode=passive srm://srm.grid.sara.nl/pnfs/grid.sara.nl/data/bbmrni/RP2/resources/hg19/indices/human_g1k_v37.chr1.fa.pac
```

**Picked up JAVA_OPTIONS:** `-Xmx512M`

**server_mode** is specified as passive, setting number of streams to 1

```bash
srmpc --server-mode=passive srm://srm.grid.sara.nl/pnfs/grid.sara.nl/data/bbmrni/RP2/resources/hg19/indices/human_g1k_v37.chr1.fa.rbwt
file:///tmp/jobdir/28851242.stro.nikhef.nl/resources/hg19/indices/human_g1k_v37.chr1.fa.rbwt
```

**Picked up JAVA_OPTIONS:** `-Xmx512M`

**server_mode** is specified as passive, setting number of streams to 1

```bash
srmpc --server-mode=passive srm://srm.grid.sara.nl/pnfs/grid.sara.nl/data/bbmrni/RP2/resources/hg19/indices/human_g1k_v37.chr1.fa.rbwt
```

**Picked up JAVA_OPTIONS:** `-Xmx512M`

**server_mode** is specified as passive, setting number of streams to 1
Failed jobs overview

<table>
<thead>
<tr>
<th>TaskID</th>
<th>Status</th>
<th>StatusTime</th>
</tr>
</thead>
<tbody>
<tr>
<td>imputation_run01_6664677232470394</td>
<td>failed</td>
<td>April 6, 2013, 07:47:18</td>
</tr>
<tr>
<td>imputation_run01_6664677232470394</td>
<td>failed</td>
<td>April 6, 2013, 14:31:48</td>
</tr>
<tr>
<td>imputation_run01_6664677232470394</td>
<td>failed</td>
<td>April 6, 2013, 19:04:43</td>
</tr>
</tbody>
</table>

Running on node: v33-45.gina.sara.nl

Error: terminate called after throwing an instance of 'std::bad_alloc'
what(): St9bad_alloc

How much memory: virtual memory (kbytes, -v) 4194304

chr: 4 from: 185000001 to: 190000001
Workflow deployment
Computational environments

- Local compute/cloud
- Cluster compute
- Local storage/cloud
- Cluster storage
- Distributed grid storages
- (Inter)national grid

Ease of use vs. redundancy & scale

Tool environment

Data environment
“Harmonized” tool management

Tool in input sandbox
“getFile('tool.zip')”

In $WORKDIR
• Download

Tool deployed as
“load module”

In $VO_BBMRI_NL_SW_DIR
• Download
• Build
• Configure

Simple download vs. On-site build deployment
Harmonized’ tool management: modules

- Build using standard ‘modulecmd’ tool
- Software should be deployed at all grid sites
- Module file should be added to all sites

```tcl
# %Module1.0
#
## bwa 0.5.8c_patched modulefile
##

proc ModulesHelp { } {
    global toolversion
    global toolname

    puts stderr "Set up the environment for $toolname version $toolversion"

    # for Tcl script use only
    set toolname bwa
    set toolversion 0.5.8c_patched
    set tooldir "\$env(VO_BBMRI_NL_SW_DIR)/tools/$toolname-$toolversion"
```
Workflow execution
• Started pilot jobs retrieve analysis jobs from Molgenis server
Workflow execution with pilots (1)

Start

Server send Pilot to scheduler

Pilot asks DB for Job to do

Is Job available in DB?

No

Pilot stops

Yes

glite-wms-job-submit \
-d $USER ... $HOME/maverick.jdl

curl ... -F status=started \ 
-F backend=ui.grid.sara.nl \ 
http://$SERVER:8080/api/pilot > script.sh
bash -l script.sh 2>&1 \ 
| tee -a log.log &

curl ... -F status=done \ 
- F log_file=@done.log \ 
http://$SERVER:8080/api/pilot
Workflow execution with pilots (3)

Pilot starts Job in background

Job reports to Pilot after execution

Pilot send Job's pulse and update to Server

Is Job's pulse received by Server?

Yes

Server check DB, if Job reported

No

Job failed

Is Job reported?

Yes

Job completed

No

while [ 1 ] ; do
  ...
  check_process "script.sh"
  CHECK_RET=$?
  if [ $CHECK_RET -eq 0 ];
  then
    ...
    curl ... -F status=nopulse \ -F log_file=@inter.log ...
  ...
  elif
  ...
  curl ... -F status=pulse \ -F log_file=@inter.log ...
  ...

Back-end independent analysis templates
//header
#MOLGENIS walltime=15:00 nodes=1 cores=4 mem=6

//tool management
module load bwa/${bwaVersion}

//data management
getFile ${indexfile}
getFile ${leftbarcodefqgz}

//template of the actual analysis
bwa aln \
${indexfile} ${leftbarcodefqgz} \n-t ${bwaaligncores} -f ${leftbwaout}

//data management
putFile ${leftbwaout}
Data transfer

- `getFile` and `putFile` are back-end specific.
- Now, we check if the files are present (cluster or `localhost`).
- Do `srm /lfn` file transfer (grid).

Input

```
getFile $WORKDIR/groups/gonl/projects/imputationBenchmarking/eQtl/hapmap2r24ceu/chr20.map
```

Generated output

```
getFile ${studyInputPedMapChr}.map
```

```bash
getRemoteLocation()
{
  ARG=$@
  myFile=$ARG
  remoteFile=srm://srm.grid.sara.nl/pnfs/grid.sara.nl/data/bbmri.nl/RP2$myFile
  echo $remoteFile
}

getFile()
{
  ARG=$@
  NUMBER="${#ARG}";
  if [ "$NUMBER" = "1" ]
    then
    myFile=$ARG
    remoteFile= getRemoteLocation $myFile
    # 1. myPath = getPath( myFile ) will strip off the file name and return the path
    mkdir -p $(dirname "$myFile")
    # 2. cp srm://remoteFile myFile
    echo "srmdcp -server_mode=passive $remoteFile file:///${myFile}"
    srmdcp -server_mode=passive $remoteFile file:///${myFile}
    chmod 755 ${myFile}
  else
    echo "Example usage: getData "$TMPDIR/datadir/myfile.txt""
    fi
}

putFile()
{
  ARG=$@
  NUMBER="${#ARG}";
  if [ "$NUMBER" = "1" ]
    then
    myFile=$ARG
    remoteFile= getRemoteLocation $myFile
    echo "srmdcp -server_mode=passive file:///${myFile} ${remoteFile}"
    srmdcp -server_mode=passive file:///${myFile} $remoteFile
    if [ $returnCode -ne 0 ]
        then
        exit 1
        fi
    else
    echo "Example usage: putData "$TMPDIR/datadir/myfile.txt"
    fi
}
//header
#MOLGENIS walltime=15:00 nodes=1 cores=4 mem=6

//tool management
module load bwa/0.5.8c_patched

//data management
getFile $WORKDIR/resources/hg19/indices/human_g1k_v37.fa
getFile $WORKDIR/groups/gcc/projects/cardio/run01/rawdata/121128_SN163_0484_AC1D3HACXX_L8_CAACCT_1.fq.gz

//template of the actual analysis
bwa aln \human_g1k_v37.fa 121128_SN163_0484_AC1D3HACXX_L8_CAACCT_1.fq.gz -t 4 \ -f 121128_SN163_0484_AC1D3HACXX_L8_CAACCT_1.bwa_align.human_g1k_v37.sai

//data management
putFile $WORKDIR/groups/gcc/projects/cardio/run01/results/121128_SN163_0484_AC1D3HACXX_L8_CAACCT_1.bwa_align.human_g1k_v37.sai
Current developments
Pilots Dashboard

- During execution

<table>
<thead>
<tr>
<th><a href="mailto:run08@ui.grid.sara.nl">run08@ui.grid.sara.nl</a></th>
<th>(May 31, 2013 11:31:31 AM)</th>
</tr>
</thead>
</table>
| Active
| Inactivate
| Stop Submitting Pilots |

- Workflow is completed

<table>
<thead>
<tr>
<th><a href="mailto:run08@ui.grid.sara.nl">run08@ui.grid.sara.nl</a></th>
<th>(May 31, 2013 11:31:31 AM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completed</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Jobs generated</th>
<th>Jobs done</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jobs ready</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Jobs running</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pilots submitted</td>
<td>17</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Jobs generated</th>
<th>Jobs done</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jobs ready</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jobs running</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pilots submitted</td>
<td>21</td>
<td>12</td>
</tr>
</tbody>
</table>
## Dash-board for jobs monitoring (work-in-progress)

<table>
<thead>
<tr>
<th>Run Number</th>
<th>Username</th>
<th>Status</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="mailto:run05@ui.grid.sara.nl">run05@ui.grid.sara.nl</a></td>
<td>Not active</td>
<td>(May 30, 2013 3:43:51 PM)</td>
<td>Jobs generated: 6, Jobs done: 0, Jobs ready: 0, Jobs running: 0, Pilots submitted: 0, Pilots started: 0</td>
</tr>
<tr>
<td><a href="mailto:run04@ui.grid.sara.nl">run04@ui.grid.sara.nl</a></td>
<td>Active</td>
<td>(May 30, 2013 3:43:33 PM)</td>
<td>Jobs generated: 6, Jobs done: 0, Jobs ready: 0, Jobs running: 0, Pilots submitted: 0, Pilots started: 0</td>
</tr>
<tr>
<td><a href="mailto:run03@ui.grid.sara.nl">run03@ui.grid.sara.nl</a></td>
<td>Active</td>
<td>(May 30, 2013 3:43:18 PM)</td>
<td>Jobs generated: 2, Jobs done: 0, Jobs ready: 4, Jobs running: 0, Pilots submitted: 0, Pilots started: 0</td>
</tr>
<tr>
<td><a href="mailto:run02@ui.grid.sara.nl">run02@ui.grid.sara.nl</a></td>
<td>Not active</td>
<td>(May 30, 2013 3:43:02 PM)</td>
<td>Jobs generated: 5, Jobs done: 0, Jobs ready: 0, Jobs running: 1, Pilots submitted: 0, Pilots started: 0</td>
</tr>
<tr>
<td><a href="mailto:run01@ui.grid.sara.nl">run01@ui.grid.sara.nl</a></td>
<td>Completed</td>
<td>(May 30, 2013 3:42:45 PM)</td>
<td>Jobs generated: 0, Jobs done: 6, Jobs ready: 0, Jobs running: 0, Pilots submitted: 0, Pilots started: 0</td>
</tr>
</tbody>
</table>
Enhancements and further steps

- What if not all parameters are known at the generation time
  - Run-time parameters passing to DB from previous steps (implemented)

- Advanced pilot management
  - Pilots re-using

- Better workflow visualization
  - Showing workflow elements and their properties

Conclusion
Conclusion

- One protocol template style that is suitable for different back-ends
- Workflow tools deployment using module system
- Hidden in scripts data management
- Workflow execution using pilot jobs
All available as open source

http://www.molgenis.org
http://www.molgenis.org/wiki/ComputeStart
h.v.byelas@gmail.com m.a.swertz@gmail.com

Thank you!

Questions?