BSAS Macro

TRUE BATCH PROCESSING FROM SAS EG USING LOAD SHARING FACILITY (LSF)

ADAM HENDRICKS, SR. PRINCIPAL SYSTEMS ANALYST, GENERAL DYNAMICS FEDERAL CIVILIAN HEALTH
DEREK GRITTMANN, SR. PRINCIPAL SYSTEMS ANALYST, GENERAL DYNAMICS FEDERAL CIVILIAN HEALTH
The Centers for Medicare & Medicaid Services (CMS) launched the Chronic Conditions Data Warehouse (CCW), a research database, in response to the Medicare Modernization Act of 2003 (MMA).

Section 723 of the MMA outlined a plan to improve the quality of care and reduce the cost of care for chronically ill Medicare beneficiaries. In addition to chronic conditions, the CCW supports health policy analysis and other CMS initiatives.

Managed by HealthAPT (joint venture between NewWave Technologies and General Dynamics Information Technology), the CMS Virtual Research Data Center (VRDC) within the CCW is an alternative solution that provides timelier access to Medicare and Medicaid program data in a more efficient and cost effective manner.

The CCW VRDC provides federal agencies, health policy analysts, and researchers with Medicare, Medicaid, assessment, and Part D data linked by beneficiary enabling longitudinal analysis across the continuum of care. With over 30 data sources, the CCW VRDC has SAN capacity of 4 PBs and contains over 322 billion records (1 billion records added monthly) for 115+ million beneficiaries.
As a trusted health solutions organization for more than 30 years, General Dynamics Federal Civilian Health, part of General Dynamics Information Technology, provides end-to-end solutions and professional services to health organizations in the defense, federal civilian government, state and local government, commercial and international sectors.

Healthcare management services include:
- Database architecture
- Data access and dissemination
- Applications
- Security
- Data science services
### Differences Between True Batch Job vs SAS EG on LSF

<table>
<thead>
<tr>
<th>What is a “True Batch Job”</th>
<th>SAS EG on LSF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not interactive.</td>
<td>Interactive front end.</td>
</tr>
<tr>
<td>Not child process of interactive process that launched it.</td>
<td>Uses only one server grid as exec host by default.</td>
</tr>
<tr>
<td>Completely independent of interactive process that launched it.</td>
<td>Has background processes that are child processes of interactive session.</td>
</tr>
<tr>
<td>I/O environment same as that of interactive process that launched it.</td>
<td>Can leverage multiple servers on grid but only with complex SAS wrapper code. These are still not independent batch sessions. If the interactive parent session ends; so do the multiple grid job child sessions.</td>
</tr>
</tbody>
</table>
SAS EG True Batch Job on LSF

- SAS code is saved under file system available to EG.
- Batch job uses full Load Sharing Facility (LSF) grid capabilities.
- Batch job is completely independent of SAS EG session.
- Results of batch job are available in same directory as SAS program launched in batch.
- Depending on how EG environment is set up, no change in SAS code may be necessary to run job in batch.

Note: Any macro or script for true batch in EG will work identically on the command line.
Grid Configuration for True Batch

Second grid metadata server with XCMD turned on or sasgrid & WorkspacesServer.sh script edited to give XCMD capability to users approved for true batch processing.

Example of WorkspaceServer.sh modification:

```bash
cmd="$SAS_COMMAND $USERMODS_OPTIONS"
arg='-noxcmd'
batch_group_found=`/usr/bin/groups $USER | grep 'sas_batch' | wc -l`

if [[ $batch_group_found -ne 0 ]]
then
  arg="-xcmd"
fi

cmd="$cmd $arg"
```
Flow of True Batch Job from EG

1. Get Run Directory
2. Build Launch SASGSUB Script
3. Launch SASGSUB Script
4. Custom SASGSUB Wrapper Function
5. LSF SASGSUB Job
6. Loop Subprocess
7. SASGSUB Wrapper Log in Run Directory w/ completion status.
8. Job Ends

Trash

Self-Deleting Script

SAS Job Log/List In Run Directory

Copy Log/List Updates to Run Directory

SAS Job Log/List In gridwork
CCW SAS Program

File: bsasl.sas

Date: Tue Jun 2 17:03:56 EDT 2015

Programmer: Adam Hendricks, adam.hendricks@gidt.com

Description:

Launches LSF batch jobs from two externally stored SAS programs.

;

options ps=999 missing=' ' nocenter nodate nonumber errorabend mprint;

Start ;

%bsas(bsas/pgms/test.sas)

x 'sleep 30';

%bsas(bsas/pgms/oratest.sas)

End of Program ;
* CCW SAS Program
*
* File: bsaslook.sas
*
* Date: Tue Jun  2 17:03:56 EDT 2015
*
* Programmer: _, Adam Hendricks, adam.hendricks@gdit.com
*
* Description:
*
* Lists LSF batch jobs currently running.
*
*;

options ps=999 missing="" nocrunch nodate nonumber errorabend mprint;

* Start *;

%bsaslook

* End of Program *;
<table>
<thead>
<tr>
<th>Grid ID</th>
<th>Job ID</th>
<th>Job Status</th>
<th>Job Queue</th>
<th>From Host</th>
<th>Host</th>
<th>Execute</th>
<th>Job Name</th>
<th>Job Submitted</th>
<th>Job Last Poll Time</th>
<th>Job Last Poll Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>272868</td>
<td>RUN</td>
<td>interactive_pgp</td>
<td>SAS</td>
<td>test01</td>
<td>SAS</td>
<td>09Nov2018</td>
<td>15:52</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
</tr>
<tr>
<td>272919</td>
<td>RUN</td>
<td>normal</td>
<td>test02</td>
<td>09Nov2018</td>
<td>15:52</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272920</td>
<td>RUN</td>
<td>normal</td>
<td>test03</td>
<td>09Nov2018</td>
<td>15:52</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272921</td>
<td>RUN</td>
<td>extract</td>
<td>test04</td>
<td>09Nov2018</td>
<td>15:52</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272922</td>
<td>RUN</td>
<td>priority-extract</td>
<td>test05</td>
<td>09Nov2018</td>
<td>15:52</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272923</td>
<td>RUN</td>
<td>normal</td>
<td>test06</td>
<td>09Nov2018</td>
<td>15:53</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272924</td>
<td>RUN</td>
<td>normal</td>
<td>test07</td>
<td>09Nov2018</td>
<td>15:53</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272925</td>
<td>RUN</td>
<td>extract</td>
<td>test08</td>
<td>09Nov2018</td>
<td>15:53</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272926</td>
<td>RUN</td>
<td>priority-extract</td>
<td>test09</td>
<td>09Nov2018</td>
<td>15:53</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272927</td>
<td>RUN</td>
<td>normal</td>
<td>test10</td>
<td>09Nov2018</td>
<td>15:53</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272928</td>
<td>RUN</td>
<td>extract</td>
<td>test11</td>
<td>09Nov2018</td>
<td>15:53</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272929</td>
<td>RUN</td>
<td>priority-extract</td>
<td>test12</td>
<td>09Nov2018</td>
<td>15:53</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272930</td>
<td>RUN</td>
<td>normal</td>
<td>test13</td>
<td>09Nov2018</td>
<td>15:53</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272932</td>
<td>RUN</td>
<td>extract</td>
<td>test15</td>
<td>09Nov2018</td>
<td>15:53</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272933</td>
<td>RUN</td>
<td>priority-extract</td>
<td>test16</td>
<td>09Nov2018</td>
<td>15:53</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272934</td>
<td>RUN</td>
<td>normal</td>
<td>test17</td>
<td>09Nov2018</td>
<td>15:53</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272935</td>
<td>RUN</td>
<td>normal</td>
<td>test18</td>
<td>09Nov2018</td>
<td>15:54</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272936</td>
<td>RUN</td>
<td>extract</td>
<td>test19</td>
<td>09Nov2018</td>
<td>15:54</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272937</td>
<td>RUN</td>
<td>priority-extract</td>
<td>test20</td>
<td>09Nov2018</td>
<td>15:54</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272938</td>
<td>RUN</td>
<td>normal</td>
<td>test21</td>
<td>09Nov2018</td>
<td>15:54</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272939</td>
<td>RUN</td>
<td>normal</td>
<td>test22</td>
<td>09Nov2018</td>
<td>15:54</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272940</td>
<td>RUN</td>
<td>extract</td>
<td>test23</td>
<td>09Nov2018</td>
<td>15:54</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272941</td>
<td>RUN</td>
<td>priority-extract</td>
<td>test24</td>
<td>09Nov2018</td>
<td>15:54</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272942</td>
<td>RUN</td>
<td>normal</td>
<td>test25</td>
<td>09Nov2018</td>
<td>15:54</td>
<td>09Nov2018</td>
<td>12:58</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N = 26
Questions?

Adam Hendricks, Sr. Principal Systems Analyst, General Dynamics Federal Civilian Health
  ◦ Adam.Hendricks@gdit.com

Derek Grittmann, Sr. Principal Systems Analyst, General Dynamics Federal Civilian Health
  ◦ Derek.Grittmann@gdit.com