

## Kepler - Bug #5685

### data isn't always chunked properly

08/17/2012 04:29 PM - Derik Barseghian

<b>Status:</b>	New	<b>Start date:</b>	08/17/2012
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assignee:</b>	Derik Barseghian	<b>% Done:</b>	0%
<b>Category:</b>	sensor-view	<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>	sensor-view-1.x.y	<b>Spent time:</b>	0.00 hour
<b>Bugzilla-Id:</b>	5685		
<b>Description</b>			
<p>Jing ran the workflow in Windows XP today, and it produced at least one datapackage with data from different sampling rates. See: <a href="http://dev2.nceas.ucsb.edu/knb/metacat?action=read&amp;qformat=default&amp;sessionid=&amp;docid=doc.1345243176506661893.1&amp;displaymodule=entity&amp;entitytype=dataTable&amp;entityindex=1">http://dev2.nceas.ucsb.edu/knb/metacat?action=read&amp;qformat=default&amp;sessionid=&amp;docid=doc.1345243176506661893.1&amp;displaymodule=entity&amp;entitytype=dataTable&amp;entityindex=1</a></p> <p>Example change section from data from above link:</p> <pre>2012-08-16 22:18:14 13.28463459 2012-08-16 22:18:44 13.284519196 2012-08-16 22:19:14 13.284427643 2012-08-16 22:19:44 13.284352303 2012-08-16 22:20:19 13.284294128 2012-08-16 22:38:14 13.28584671 2012-08-16 22:38:15 13.28584671 2012-08-16 22:38:16 13.291329384</pre> <p>These were the results Jing got from the workflow for this particular sensor:</p> <p>Sensor Name: gpp-data/CR800_Batt_Volt Document URL: <a href="http://dev2.nceas.ucsb.edu/knb/metacat?action=read&amp;qformat=default&amp;docid=doc.1345243176506661893.1">http://dev2.nceas.ucsb.edu/knb/metacat?action=read&amp;qformat=default&amp;docid=doc.1345243176506661893.1</a> Time Range: 2012-08-16 00:10:09 ~ 2012-08-16 22:38:32 Number of Records: 2680</p> <p>Sensor Name: gpp-data/CR800_Batt_Volt Document URL: <a href="http://dev2.nceas.ucsb.edu/knb/metacat?action=read&amp;qformat=default&amp;docid=doc.1345243308947314582.1">http://dev2.nceas.ucsb.edu/knb/metacat?action=read&amp;qformat=default&amp;docid=doc.1345243308947314582.1</a> Time Range: 2012-08-16 22:38:34 ~ 2012-08-16 22:39:54 Number of Records: 75</p> <p>Sensor Name: gpp-data/CR800_Batt_Volt Document URL: <a href="http://dev2.nceas.ucsb.edu/knb/metacat?action=read&amp;qformat=default&amp;docid=doc.134524344776683166.1">http://dev2.nceas.ucsb.edu/knb/metacat?action=read&amp;qformat=default&amp;docid=doc.134524344776683166.1</a> Time Range: 2012-08-16 22:39:55 ~ 2012-08-17 22:37:21 Number of Records: 8186</p>			

### History

#### #1 - 08/24/2012 06:23 PM - Derik Barseghian

I thought about this some more last week, and remembered that I saw some errors when changing sensor sampling rates from my windows box to do w/ the metadata channels. (And I believe the site layout running on my other machine stopped when this happened). The sampling rate change took effect, but I suspect the metadata entries didn't make it through to DT when this happened. So my current thinking is the chunking problem isn't the archive workflow's fault, but an issue w/ metadata changes not always making it from the sensor actor into the DataTurbine metadata channel. I planned to dump the DT's metadata channel to verify, but unfortunately today when running the archival workflow for the first time against the DT with a week's worth of data in it, DT crashed.

I do have the DT archive though (attached), so I should be able to get it reloaded to verify in the future.

#### #2 - 08/24/2012 06:27 PM - Derik Barseghian

RBNN archive containing suspected missing metadata entries for various sampling rate changes. Get DT to load this, use the DatToDT script I wrote to dump the metadata channel, and compare entries with the data channel.

#### #3 - 08/25/2012 02:17 AM - Derik Barseghian

Doesn't look like trolling through the archive will be necessary. The same problem of a datapackage containing two data with two different sampling rates occurred tonight. I did not receive any errors when adjusting sensor sampling rates in Kepler. Here are the recents of the last archive workflow run:

-----

Sensor Name: gpp-data/CR800\_Batt\_Volt

Document URL:

<http://dev2.nceas.ucsb.edu/knb/metacat?action=read&qformat=default&docid=doc.13458810777906caa4227-8260-449b-8fff-11c07741dcc4.1>

Time Range: 2012-08-25 02:21:35 ~ 2012-08-25 02:48:58

Number of Records: 1641

Sensor Name: gpp-data/CR800\_Batt\_Volt

Document URL:

<http://dev2.nceas.ucsb.edu/knb/metacat?action=read&qformat=default&docid=doc.1345881206554a70e25dc-e965-4430-9b17-f99499aff969.1>

Time Range: 2012-08-25 02:49:30 ~ 2012-08-25 07:50:36

Number of Records: 602

Sensor Name: gpp-data/CR800\_sq311\_1

Document URL:

<http://dev2.nceas.ucsb.edu/knb/metacat?action=read&qformat=default&docid=doc.1345881330868470d3007-65e6-46c3-997d-5d27e37f15ee.1>

Time Range: 2012-08-25 02:21:34 ~ 2012-08-25 02:50:50

Number of Records: 826

Sensor Name: gpp-data/CR800\_sq311\_1

Document URL:

<http://dev2.nceas.ucsb.edu/knb/metacat?action=read&qformat=default&docid=doc.13458814576456ba0d71b-1962-4034-9ee0-6bc8fb41c29b.1>

Time Range: 2012-08-25 03:01:05 ~ 2012-08-25 07:41:06

Number of Records: 29

Sensor Name: gpp-data/CR800\_sq311\_2

Document URL:

<http://dev2.nceas.ucsb.edu/knb/metacat?action=read&qformat=default&docid=doc.13458815819617734014f-b2ab-4bd2-85d4-b7a059ca2aa5.1>

Time Range: 2012-08-25 02:21:33 ~ 2012-08-25 07:50:36

Number of Records: 356

-----

If you look at the raw data for sq311\_2, you'll see data at 30s, and then an un-smooth change to 60s:

```

2012-08-25 02:47:33  0.67334365845
2012-08-25 02:48:03  0.67333245277
2012-08-25 02:48:33  0.67334610224
2012-08-25 02:50:30  0.67336404324
2012-08-25 02:51:35  0.67335760593
2012-08-25 02:52:35  0.67335271835

```

Below is the result of dumping the metadata channels from the DT. You can see CR800\_sq311\_2 does only have one metadata entry.

someData.length:3

times.length:3

i:0 someData<sup>0</sup>:CR800\_Batt\_Volt altitude=0.000000,coefficients=,conversion-type=no

conversion,daq-method=,isOn=true,latitude=34.412291,longitude=-119.842335,measurement-unit=Volts,sampleMethod=average,samples-per-measurement=1,samplingPeriod=1,sensor-make=Campbell Scientific,sensor-measurement=,sensor-model=,serial-number=

i:0 times<sup>0</sup>:1.345870971714E9

i:0 someData<sup>1</sup>:CR800\_Batt\_Volt altitude=0.000000,coefficients=,conversion-type=no

conversion,daq-method=,isOn=true,latitude=34.412291,longitude=-119.842335,measurement-unit=Volts,sampleMethod=average,samples-per-measurement=1,samplingPeriod=30,sensor-make=Campbell Scientific,sensor-measurement=,sensor-model=,serial-number=

i:0 times<sup>1</sup>:1.345888139714E9

i:0 someData<sup>2</sup>:CR800\_Batt\_Volt altitude=0.000000,coefficients=,conversion-type=no

conversion,daq-method=,isOn=true,latitude=34.412291,longitude=-119.842335,measurement-unit=Volts,sampleMethod=average,samples-per-measurement=1,samplingPeriod=30,sensor-make=Campbell Scientific,sensor-measurement=,sensor-model=,serial-number=

i:0 times<sup>2</sup>:1.345888200714E9

someData.length:1

times.length:1

i:1 someData<sup>0</sup>:CR800\_sq311\_1 altitude=0.000000,coefficients=,conversion-type=no

conversion,daq-method=,isOn=true,latitude=34.412291,longitude=-119.842335,measurement-unit=mV,sampleMethod=average,samples-per-measurement=1,samplingPeriod=2,sensor-make=Apogee Instruments,sensor-measurement=Photosynthetic Photon Flux (PPF),sensor-model=SQ-311 (sun),serial-number=1612

i:1 times<sup>0</sup>:1.345870973714E9

someData.length:1

times.length:1

i:2 someData<sup>0</sup>:CR800\_sq311\_2 altitude=0.000000,coefficients=,conversion-type=no

conversion,daq-method=,isOn=true,latitude=34.412291,longitude=-119.842335,measurement-unit=mV,sampleMethod=average,samples-per-measurement=1,samplingPeriod=30,sensor-make=Apogee Instruments,sensor-measurement=Photosynthetic Photon Flux (PPF),sensor-model=SQ-311 (sun),serial-number=1609

i:2 times<sup>0</sup>:1.345870973714E9

-----

So at least we can rule out the archival workflow.

We should look at how the Sensor actor is inserting metadata entries to DT, if it can verify they get inserted, etc.

**#4 - 08/25/2012 02:45 AM - Derik Barseghian**

Sensor actor appears to be properly setting span metadata:

```
[run] SpanControl.setMetadataForSensor(Batt_Volt,CR800,samplingPeriod,10)
```

```
[run] SpanControl.setMetadataForSensor got result from _sendCommand:2012-08-25T09:37:39.714Z OK: Channel: CR800_Batt_Volt,measurement-period={10.000000}
```

Need to check that span is properly outputting these, and that they're then picked up by spanToDT.

**#5 - 08/25/2012 02:54 AM - Derik Barseghian**

(The above was an example of a metadata change that didn't make it through to DT.)

(In reply to comment [#4](#))

**#6 - 08/30/2012 05:34 PM - Derik Barseghian**

Moving to 1.x.y target.

**#7 - 03/27/2013 02:31 PM - Redmine Admin**

Original Bugzilla ID was 5685

**Files**

---

rbnb_archive.tgz	184 KB	08/25/2012	Derik Barseghian
------------------	--------	------------	------------------