Interprocess message specifications

V 1.0

The major revision number of this document should match the major revision number of the associated flowchart.

Task Queue Message Specification

```
{
  "userdata" :
  {
    "TaskID" : "My Task ID",
    "operations" : {},
    "??" : "??"
  }
}
```

- *Task ID* is a UUID, generated by the task-initiating application, which will be used to identify this task from start to finish. All jobs, results, and reports generated from this Task will be tagged with this ID.
- *operations* is a hash such as we provide to the management ops validator

What else should go in here?

- Something providing field geometry, etc. ?
- A *Final Action*, such as a url to post report to? Email address? Something indicating where results are sent or how the client is notified that the task is done ?

Report Gen Queue Message Specification

```
{
  "receive_queue_name" : "My Task ID"
  "userdata" :
  {
    "Task ID" : "My Task ID",
    "Job IDs" : ["JobID_0","JobID_1","JobID_...","JobID_5999"],
    "Final Action" : ""
  }
}
```

- *receive_queue_name* is an EZQ directive telling the report handler to poll the named queue to get job results. This should have the same value as *Task ID*
- *Task ID* is the same Task ID as in the Task Queue
- *Job IDs* is an array of unique job ids, each of which is a UUID generated by the GIS Query Node
- *Final Action* (or something similar). See comment in Task Queue section

Job Queue Message Specification

```
{
  "result_queue_name" : "My Task ID",
  "userdata" :
  {
    "Job ID" : "The UUID for this job",
    "mmp360 input data" : {}
  }
}
```

- *result_queue_name* is an EZQ directive telling it to place the results of this job into the named queue. This name should be set to the Task ID, as that is the name of the results queue created by GIS Query Node for this task
- *Job ID* is a UUID matching one of those in the Job IDs array of a Report Gen Queue message

- *mmp360 input data* is whatever data structure is needed for a single mmp360 Worker Node to process this task. Depending on the size of this structure, we may need to have GIS Query Node write out the job input details to S3 and place “bucket” and “key” directives outside of the userdata object. In that case, the *mmp360 input data* field should be removed from this specification.

Results Queue Message Specification

```
{
  "bucket" : "Name of an S3 bucket",
  "key" : "Key of the file containing the full results of this job",
  "userdata" :
  {
    "Job ID" : "The UUID for this job",
  }
}
```

- *bucket* is the name of the S3 bucket that workers are told to store their results in. That bucket name is set at initial configuration of the worker node AMI. We repeat it here so that EZQ can handle fetching the data automatically. If it turns out that the result data from a worker will always be fairly small (<256k), we can skip caching results in S3 and simply place them directly in the userdata portion of the message.
- *key* is the name of the specific file containing this job's results
- *Job ID* is the job id associated with these results