

```

{
  "source" :
  {
    "pubsub" :
    {
      "project_id" : "ncbi-sandbox-blast",
      "subscript_id" : "spark-test-subscript",
      "use" : "true"
    }
  },

  "blastjni" :
  {
    "num_db_limit" : 15,
    "manifest_root" : "dbs_50M.json"
  },

  "result" :
  {
    "asn1" :
    {
      "bucket" : "blastgcp-pipeline-test"
    },
    "status" :
    {
      "bucket" : "blastgcp-pipeline-test"
    }
  },

  "spark" :
  {
    "num_executors" : 15,
    "num_executor_cores" : 2
  }
}

```

minimal ini.json for spark-blast

where to get requests from

what databases to use
and a limit in case of smaller test-clusters

where to write the results to

dimensions of the cluster

```

{
  "databases": [
    { "type": "nucleotide", "manifest": "gs://nt_50mb_chunks/nt_50M.20180502_1.dbmanifest.json" },
    { "type": "protein", "manifest": "gs://nr_50mb_chunks/nr_50M.20180502_1.dbmanifest.json" }
  ]
}

```

dbs_50M.json

```

<?xml version="1.0"?>
<allocations>

  <pool name="default">
    <schedulingMode>FAIR</schedulingMode>
    <weight>1</weight>
    <minShare>2</minShare>
  </pool>

  <pool name="production">
    <schedulingMode>FAIR</schedulingMode>
    <weight>1</weight>
    <minShare>2</minShare>
  </pool>

  <pool name="test">
    <schedulingMode>FIFO</schedulingMode>
    <weight>2</weight>
    <minShare>3</minShare>
  </pool>

</allocations>

```

pooles.xml