

# Simulated Errors

Paul Nilsson (BNL)

# Historical error codes

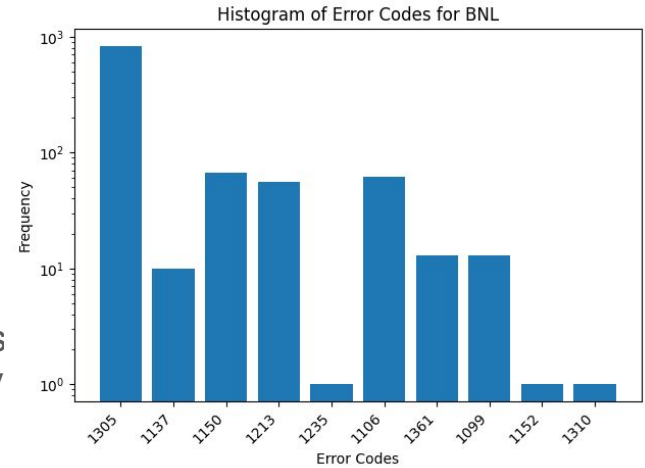
- 5-10% of all grid jobs fail for different reasons
  - Need to mirror this for the SimGrid simulation to be realistic
- Simulate errors in SimGrid using real data from PanDA
  - Relevant categories depending on how deep the simulation will go; errors on the simulated “worker node” easiest to start with
- Data source: (pilot errors from Tania)
  - [last 24h]  
<https://bigpanda.cern.ch/errors/?json&hours=24&limit=10000000&fields=errsBySite>
  - [historical period]  
[https://bigpanda.cern.ch/errors/?json&limit=10000000&fields=errsBySite&date\\_from=2024-12-01T00:00:00&date\\_to=2024-12-02T00:00:00](https://bigpanda.cern.ch/errors/?json&limit=10000000&fields=errsBySite&date_from=2024-12-01T00:00:00&date_to=2024-12-02T00:00:00)
- Ref: Pilot error codes wiki (up to date)
  - <https://github.com/PanDAWMS/pilot3/wiki/Error-Codes>

```
    "1137": 1  
  },  
  "BNL": {  
    "1305": 831,  
    "1137": 10,  
    "1150": 67,  
    "1213": 56,  
    "1235": 1,  
    "1106": 62,  
    "1361": 13,  
    "1099": 13,  
    "1152": 1,  
    "1310": 1  
  },  
  "BNL_OPP": {  
    "1201": 1250,  
  }
```

Example of filtered data with pilot errors

# Randomization of errors

- In SimGrid
  - When a given job runs, simulation should decide if job should finish or fail
  - If failure, draw random error from distribution
    - Depending on which error it is, set job info accordingly (stage-in error e.g., will not spend any CPU, or very little)
    - Also affects simulated data transfers (no output files will be produced when there is a stage-in error, only the job log will be staged out)
  - Need to know total number of jobs for same time period (add to JSON)
    - Can be added to “error code histogram” for small optimization (“error code”=0 means no error)
- Compilable stand-alone demonstration code
  - <https://github.com/PalNilsson/random-errors>
  - Using real data from PanDA
  - To be added to SimGrid when we know exactly how to handle errors there (not trivial/obvious)



Example of filtered data with pilot errors