

# AutoPyFactory

Jose Caballero, John Hover

**Joint US ATLAS S&C / Physics Support Meeting**  
Lawrence Berkeley National Laboratory

# Pilot submission behavior

Why my site is not getting pilots?

There is no nqueue anymore.

No jobs in PanDA → no pilots.

Why my site is not getting more pilots  
if there are thousand of jobs in PanDA?

APF makes a decision based on a chain of Sched plugins. Final decision is based on the combination of all of them, including each one configuration.

With n Sched plugins, the number of possibilities is:

$$N = \sum_{i=1}^n \frac{n!}{(n-i)!}$$

There are 13 plugins. Calculating N is left as exercise to the reader.

# Current configuration:

schedplugin = Ready, Scale, MaxPerCycle, MinPerCycle, StatusTest, StatusOffline, MaxPending

sched.scale.factor = 0.25

sched.minpercycle.minimum = 1

sched.maxpercycle.maximum = 100

sched.maxpending.maximum = 10

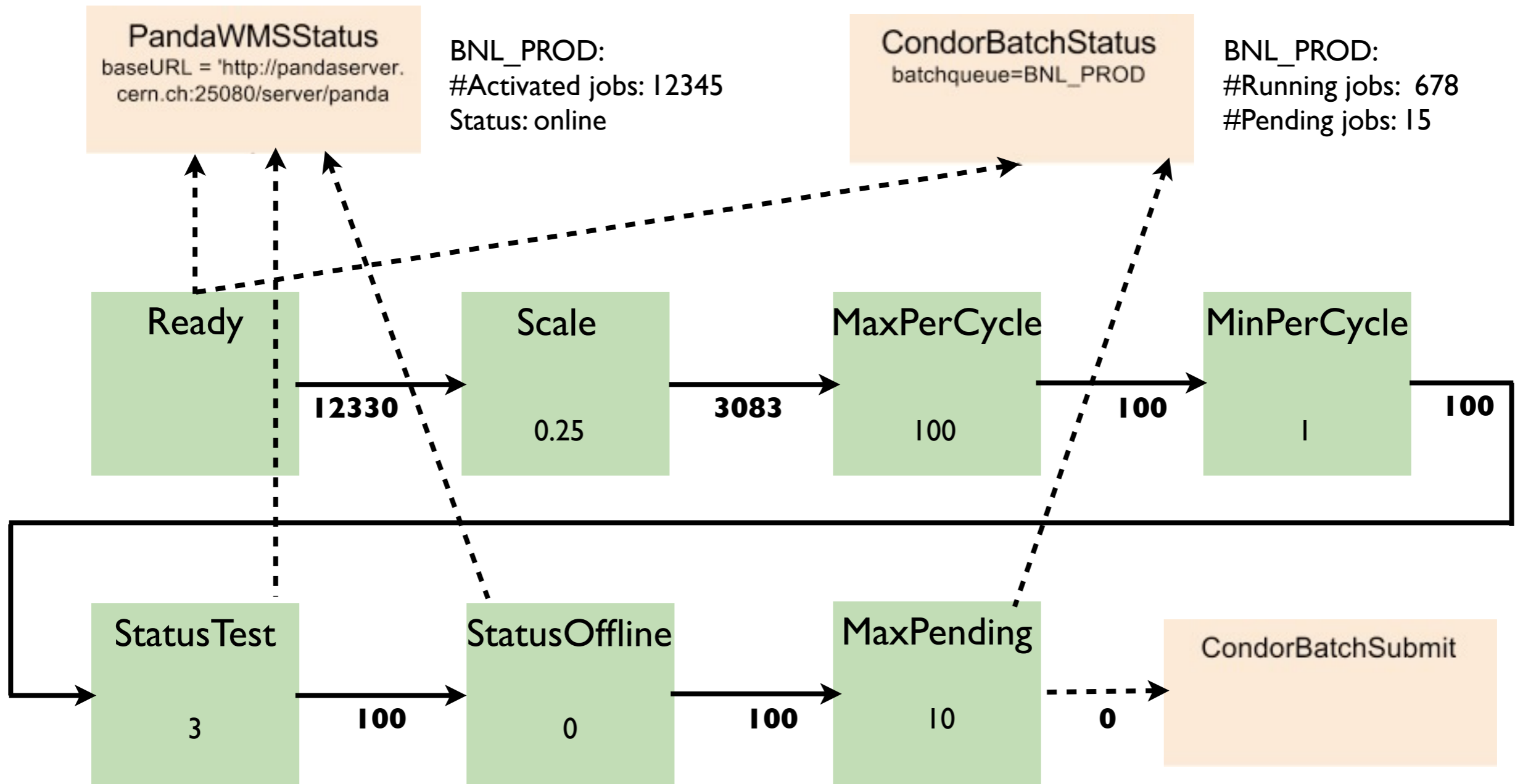
sched.statustest.allowed = True

sched.statustest.pilots = 3

sched.statusoffline.allowed = True

sched.statusoffline.pilots = 0

# Current configuration:



## Right number of idle pilots?

- Total farm size
- Timefloor

## Latency in HTC?



# Factory configuration

# Typical configuration:

```
[ANALY_BNL_SHORT-gridgk02]
```

```
batchqueue = ANALY_BNL_SHORT-condor
```

```
wmsqueue = ANALY_BNL_SHORT
```

```
globusrsl.gram5.queue = analy.short
```

```
batchsubmit.condorgt5.gridresource = gridgk02.racf.bnl.gov/jobmanager-condor
```

```
executable.arguments = %(executable.defaultarguments)s -u user
```

# Special configurations:

batchsubmit.condorgt5.environ = APP=/cluster/grid/app TMP=/scratch DATA=/cluster/grid/data  
clusterid=\$(Cluster)

batchsubmit.condorgt5.environ = TMP=/scratch/ DATA=/atlasgrid/osg-data

batchsubmit.condorgt5.environ = TMP=/scratch/ DATA=/n/data I /ATLAS/osg/data

batchsubmit.condorgt5.environ = TMP=/scratch/ DATA=/n/data I /ATLAS/osg/data/

batchsubmit.condorgt5.environ = TMP=/scratch DATA=/nfs/slac/g/grid/osg/data

batchsubmit.condorgt5.environ = TMP=/tmp DATA=/atlas/data08/OSG/DATA clusterid=\$(Cluster)

# Special configurations:

```
globusrsl.gram5.globusrsladd = (condorsubmit=('+AccountingGroup' "\"group_atlasanaly.usatlas I\""))
```

```
globusrsl.gram5.globusrsladd = (condorsubmit=('+AccountingGroup' "\"group_atlasprod.usatlas I\""))
```

```
globusrsl.gram5.globusrsladd = (condorsubmit=('Requirements' '(Cpus==8)') ('+MCORE' 'True') ('+AccountingGroup' "\"group_atlasmcore.usatlas I\""))
```

```
globusrsl.gram5.globusrsladd = (xcount=16)
```

```
globusrsl.gram5.xcount = 8
```

```
globusrsl.gram5.count = 8
```

# Special configurations:

globusrsl.gram5.maxwalltime = 15830  
globusrsl.gram5.maxWallTime = 2880  
globusrsl.gram5.maxWallTime = 4500  
globusrsl.gram5.maxWallTime = 60  
globusrsl.gram5.maxwalltime = 63350  
globusrsl.gram5.maxwalltime = 960  
globusrsl.gram5.globusrsladd = (maxWallTime=15830)  
globusrsl.gram5.globusrsladd = (maxWallTime=180)  
  
globusrsl.gram5.globusrsladd = (maxTime=7200)

# Special configurations:

atlas_analysis	atlas_prod	cvmfs_test_q
ANAL_QUE	ANAL_QUE_1HR	ANAL_QUE_LMEM
analy	analy.long	analy.short
analysis	atlas_analy_q	ATLAS_Analysis
ATLAS_MCORE	ATLAS_Production	hep
hep_killable	mapr	mc8
mcore	mp8	multi_core_q
prod	PROD_QUE	PROD_QUE_LMEM
PROD_QUE_MP8	prodtest	production
rcf	short	Tier3Test

# Backward compatibility

- Hardcode default behavior for new features:  
makes APF more obscure
- Write an script to update the configuration files on  
every update
- New project every time there are new features:  
APF3-1.0.0, APF4-1.0.0, ...