

Monitoring PanDA usage of FAX

Ilija Vukotic

Valeri Fine

Paul Nilsson



FAX usage by PanDA



The first production use case for FAX is so called FAX failover.

- In case pilot fails to get stage input files from 2 tries, it will try to get them from FAX.
- To failover or not is controlled by a AGIS queue variable “failoverToFAX”. How to set it is described here:
https://twiki.cern.ch/twiki/bin/view/AtlasComputing/FAXconfigAGIS#Enabling_FAX_failover
- As it is always beneficial, we would suggest all the sites and all the queues to have it on.
 - Currently all US sites, large part of UK sites and a few continental EU site
 - There is no need to contribute to FAX in order to use it.
- We need to verify that service runs and what it contributes so comes monitoring



Getting the information



- Pilot for each job collects:
 - Number of files transferred through FAX
 - Number of files transferred in regular way
 - Number of bytes transferred through FAX
 - Number of bytes transferred in regular way
 - Duration of FAX transfers
- Information is curl-ed to PandaLogger



Presenting information



Default way <http://pandamon.atlascloud.org/logsummary>
not exactly human friendly.

Summary of Panda logged incidents, last 6 hours

Category	Type	Level	Count
panda.mon.bamboo	taskbrokerage	INFO	449
		DEBUG	2407
		WARNING	930
		ERROR	34
panda.mon.prod	update	INFO	33
	submit	INFO	2020
	analy_brokerage	INFO	20447
	pd2p	INFO	8890
	brokerage	WARNING	4199
	killJobs	INFO	603
	registerSubscription	INFO	14574
panda.mon.fax	broker_util	ERROR	3
	pilotLog	INFO	52834
	reassignJobs	INFO	560
	FAXrecovery	INFO	386

2013-10-07 21:12:42	"WithFAX":1,"WithoutFAX":0,"bytesWithFAX":4502241443,"bytesWithoutFAX":0,"timeToCopy":111
2013-10-07 21:13:14	"WithFAX":1,"WithoutFAX":103,"bytesWithFAX":557324,"bytesWithoutFAX":99667874,"timeToCopy":6572
2013-10-07 21:13:42	"WithFAX":1,"WithoutFAX":0,"bytesWithFAX":20780683,"bytesWithoutFAX":0,"timeToCopy":1170
2013-10-07 21:13:54	"WithFAX":1,"WithoutFAX":103,"bytesWithFAX":718537,"bytesWithoutFAX":115493747,"timeToCopy":6728
2013-10-07 21:14:57	"WithFAX":1,"WithoutFAX":0,"bytesWithFAX":20780683,"bytesWithoutFAX":0,"timeToCopy":1336
2013-10-07 21:14:58	"WithFAX":1,"WithoutFAX":0,"bytesWithFAX":1007380417,"bytesWithoutFAX":0,"timeToCopy":1
2013-10-07 21:15:51	"WithFAX":1,"WithoutFAX":0,"bytesWithFAX":3000882460,"bytesWithoutFAX":0,"timeToCopy":1
2013-10-07 21:16:23	"WithFAX":1,"WithoutFAX":0,"bytesWithFAX":3421227359,"bytesWithoutFAX":0,"timeToCopy":1



Presenting information



- New entry in side menu for all things FAX
- Dedicated Panda Monitor pages
- Thanks to Valeri

PanDA Monitor
Times are in UTC

- ▶ Jobs
- Search
- ▶ Jedi Tasks
- ▼ FAX

Summary:
[Fail over](#)
[Configuration](#)

Panda report on jobs failovers to FAX over last 100 hours

Record count: 1967

Show 50 entries Search:

Site	Jobs	WithFAX [files]	WithoutFAX [files]	WithFAX [GB]	WithoutFAX [GB]			
FR: IN2P3-LPSC	1	1	10	0.23	4.90			
PandaID	Time	WithFAX	WithoutFAX	WithFAX [GB]	WithoutFAX [GB]	Status	User	
1943857996	2013-10-04 13:39:32	1	10	244778122	5263533303	finished	sascha.mehlhase@cern.ch	
IT: ANALY_INFN-T1				1833	2273	2147	281.48	251.48
US: AGLT2_SL6				4	7	0	1.04	0.00
US: ANALY_MWT2_SL6				88	88	4665	8.52	341.79
US: ANALY_SWT2_CPB				1	1	16	0.44	8.97
US: OU_OCHEP_SWT2				40	40	359	25.01	158.56

Showing 1 to 6 of 6 entries



Presenting information – to be done



- Currently data older than 3 days are deleted. We want it kept
- Have a way to select/filter-out different users – currently HC jobs filtered out in code.
- Reorganize code/db to be faster
- Add historic plots for total / per site values.



Separating other FAX use cases



- overflow - when Panda decides that site is too busy and decides to send job elsewhere while reading from "busy" site via FAX. Should start in roughly one month
- user directly gives FAX filenames in input dataset. Already possible using prun jobs.

This requires

- Changes in pilot
 - In order to separate use cases
 - To collect other information
 - where the data comes from
 - all the data that contributed to making the overflow decision
- New Panda Monitor pages

