

Table of Contents

| | |
|---------------------------------------|----------|
| PandaDynamicDataPlacement..... | 1 |
| Introduction..... | 2 |
| General Policies..... | 3 |
| Tier 1 Algorithm..... | 4 |
| Tier 2 Algorithm..... | 5 |
| Logger..... | 6 |
| Monitor..... | 7 |
| Additional Information..... | 8 |

PandaDynamicDataPlacement

Introduction

This page describes PanDA Dynamic Data Placement (PD2P).

General Policies

The following general policies apply to PD2P algorithm for making replicas of datasets at Tier 1 and Tier 2 sites.

- The dataset name prefix is `mc` or `data`
- The data type is not `RAW,RDO,HITS,ESD,EVNT`
- The `hidden` metadata is not `True`
- Destination of all PD2P datasets will be `DATADISK`
- Dataset replicas on `LOCALGROUPDISK` are ignored
- Skip `GangaRobot`, `HammerCloud`, `_dis`, `_sub` etc
- Site must have sufficient free space. The total disk size and disk usage are obtained using `DQ2.queryStorageUsage()`. The threshold is `max(totalDiskSize*5%,3TB)`.
- Site must be `analysis`, not `test`, `production` etc
- Replicas at non-online sites are ignored
- Only online sites are used

Tier 1 Algorithm

It is specified as SELECTEDT1 in PD2P log.

- Primary copies of ATLAS data will be placed at Tier 1 sites based on CREM policies
- PD2P will make additional copy at Tier 1 if:
 - ◆ no PD2P replica was made during the past week to a T1
 - ◆ $\text{int}(\log_{10}(\text{nused})) > \text{nSecondaryT1Replicas}$, where `nused` stands for how many times the dataset was used per jobset
 - ◆ $\log_{10}(\text{nused}) == 10^{*\text{int}(\log_{10}(\text{nused}))}$ (it means that another replica is created when `nused=10,100,1000,...` and not for `nused=1`)
- All T1s are used even if no T2 has `shcedconfig.cachedSE=True` in the cloud
- Location of additional 'secondary' copy is based strictly on MoU share
- If dataset is a container, each containing dataset is placed using MoU share
- immediately makes an additional copy at a Tier 2 using MoU share (SELECTEDT2_T1MOU in PD2P log)

Tier 2 Algorithm

Tier 2 Algorithm is executed independently of Tier 1 Algorithm. If one or more sites in a cloud have `shcedconfig.cachedSE=True`, the cloud is regarded as a PD2P cloud. When an analysis job is submitted to Panda, PD2P is triggered. Datasets which match the following conditions are distributed.

- The dataset was used one or more times before the job is submitted. I.e., $nUsed > 0$, where $nUsed$ is how many times the dataset was used beforehand
- No replica within PD2P T2 sites ($nSitesHaveDS=0$). Or if $nWaitingJobsets > 2$, the number of replicas should be less than $maxSitesHaveDS = \text{int}(\log_{10}(nWaitingJobs/200))$. $nWaitingJobsets$ is the number of Jobsets which have at least one waiting job. $nWaitingJobs$ is the number of waiting jobs. Both $nWaitingJobsets$ and $nWaitingJobs$ are computed per dataset when the analysis job is submitted. The maximum value of $maxSitesHaveDS$ is limited to 5. This condition is checked only for $nUsed > 1$, i.e., the second analysis job always triggers PD2P
- The number of active DQ2 subscriptions per dataset made by PD2P is not greater than 2

$maxSitesHaveDS$ is checked only for the second and subsequent analysis jobs. PD2P makes a copy at a candidate site with `shcedconfig.cachedSE=True` from any PD2P cloud by using the weight. The site must have fast FTS channel to sites which already have the replica. In PD2P log, it is specified as `SELECTEDT2_JOB` for $nUsed=1$, `SELECTEDT2_NOREP` for $nSitesHaveDS=0$, or `SELECTEDT2_WAIT` for others. Another copy is made at a Tier 2 using MoU share when $nUsed=1$ (`SELECTEDT2_T2MOU` in PD2P log).

$$W_{T2} = \frac{W}{S \times R}$$

where

- W

... The same weight for analysis brokerage described in this link

- S ... The number of subscriptions made by PD2P to the site for last 24 hours + 1
- R ... The number of replicas of the dataset in T2 sites in the same cloud + 1

A new copy tends to go to another cloud due to the second negative weight if a T2 site in a cloud already has a copy. If the input is a dataset container, one candidate is found for each constituent dataset.

Logger

PD2P Logger [↗](#)

- **action=UNSELECTEDT2** : Selection according to brokering algorithm
- **action=SELECTEDT2_T2MOU** : Selection according to a/b/c/d classification

Monitor

All the monitorings associated to PD2P and extracted from the logger informations are documented in PD2P Monitor

Additional Information

- Slides, Tier 1 PD2P Plan [↗](#)
 - Slides, T1 PD2P Justification [↗](#)
 - Slides, Dataset age in PD2P [↗](#)
 - Slides, PD2P Dataset Tagging Proposal [↗](#)
 - Slides, T2 PD2P Experience [↗](#)
-

Major updates:

-- TadashiMaeno - 15-Nov-2010 -- Revised KaushikDe - 13-Apr-2011

Responsible: TadashiMaeno

Never reviewed

This topic: PanDA > PandaDynamicDataPlacement

Topic revision: r27 - 2013-02-21 - TadashiMaeno



Copyright &© 2008-2021 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.
or Ideas, requests, problems regarding TWiki? use Discourse or Send feedback