

Table of Contents

Pilot FactoryHadoop.....	1
Introduction.....	2
Use cases and specific questions.....	3
Data volume and format.....	4

Pilot Fact or yHadoop

Introduction

Autopyfactory condor logs contain useful information about the state of Grid computing fabric which are otherwise invisible to PanDA WMS. We will store these logs and analyze using the lxhadoop cluster.

Usecases and specific questions

- how much wallclock time is wasted, aggregated and per panda queue
- identify broken sites which otherwise missed by existing tools

Data volume and format

These logs are condor history logs, about 200MB each and rotated by condor. A single factory has a few of these per day so for 7 factories this is a total of 5GB per day. The content are condor classads per job with records separated by the '*** Offset' line. A sample is here [↗](#).

Major updates:

-- PeterLove - 2014-12-17

Responsible: PeterLove

Last reviewed by: **Never reviewed**

This topic: AtlasComputing > PilotFactoryHadoop

Topic revision: r2 - 2016-05-16 - IlijaVukotic



Copyright © 2008-2019 by the contributing authors. All material on this collaboration platform is the property of the contributing authors.

Ideas, requests, problems regarding TWiki? Send feedback