IFDATA Handling

Marc Mengel

# Introduction

Need an API for experiments to use to do local data handling related tasks (reading and writing). Make local vs. remote file access transparent (e.g. the system determines whether to use cpn or gridFTP). Solves the problem of getting a file deposited into some storage system by SAM or some other service onto the worker or desktop/laptop node.

Also adds logging for monitoring purposes that is compatible with the REX monitoring system. In the current system, there is no tracking of such actions.

Embodies all “local” file transfer solutions and our recommendations.

Is built for SAMWeb.

Create interfaces to other languages (python, java, …) with SWIG.

# Tasks

## Design exists

A “Swiss army knife” like design is documented in CS-docdb-4669 “IFDATA Handling Layer Design”

## Implementation

Prototype implementation is in ups/upd and redmine. IFDHC (IF Data Handling Client) at v0.2.

## Test v0.2

Test with jobsub, dagnabit, and in a “get next file” wrapper script. Coupled to dagnabit roll out

People: Dennis, Marc. Duration: two weeks. Started: in progress

## Design and implement a flexible configuration system

How will the initial storage tool be determined? (user choice? Experiment policy?). People: Marc, Art. Duration: two weeks. Start mid May.

## Complete implementation and complete integration with jobsub

Fix problems found in v2. Add logging for monitoring. Complete integration with the different storage tools. People: Marc, Dennis. Duration: one month. Started: End of May.

## Write ART framework Service for interfacing to IFDH

Write a service for ART to interface ART to IFDH. Have a separate sub-package for this to avoid direct coupling. People: Marc, Adam. Duration: two weeks. Start: May

## Determine how to handle case of no file handling clients (possible add-on)

For laptops, how to handle the case if no file handling client (like srmcp) is installed.

People: Marc. Duration: two weeks. Start: ?

## Determine how IFDATA Handling integrates with a system for storing files back to user

People: Marc, Robert, Andrew, Adam, Art, Dennis, Mike

Duration: two weeks. Start in June.

## Add-on: Need a common CLI for IFDATA, IFBEAMDATA, and NUCONDB?

These packages share the same backend. Do we want to add CLI commands for the other packages

People: Marc, Andrew. Duration: two weeks. Start: End of June

# Milestones

Integrate IFDH with job submission and job wrapper script for production. Deploy – June 15

Integrate with experiments to use IFDH (includes ART) – July 31

# Risks

## Coupled to SAM & ART integration project

If the ART project slows down, this project has to wait to deploy.

## Scaling problems with SAMWeb & Logging

When in wide use, the SAMWeb and logging servers can’t handle the load and, for some reason, do not scale.

# Conclusions

When complete and coupled with SAM we’ll have a really nice end-to-end file delivery system.