

Adaptive Data Access and Policy-driven Transfers (ADAPT) Components/integration planning document for the first year deliverables

Revised on 2/7/2012

adapt@hpcrd.lbl.gov
<http://sdm.lbl.gov/adapt/>

1 Project components

1.1 Data transfer clients

SRM-copy and Bulk Data Mover (BDM) will be enhanced in the first year, and integrated with a new ATM module, a new PTM module, and ADT library. Client tools will also support the Policy module.

1.2 Adaptive Data Transfer (ADT) Library

Based on the adaptive data transfer algorithms, the adaptive data transfer library will be developed, integrated in the srm-copy, BDM and BeStMan server in this project and released as a library package for general use.

1.3 Passive Measurement Archive (PMA)

A data transfer monitoring module (Passive Transfer Monitor, PTM) will be developed and integrated in the client tools to collect the measurements passively from the client tools. PTM will provide the collected measurements to the perfSONAR based network performance archive as PMA in the 2nd and 3rd years of project.

1.4 BeStMan server

BeStMan server needs to integrate ADT library for its gridftp transfers, to collaborate with STAR and/or ALICE/LIGO experiment. Also, BeStMan server needs to support FDT service in addition to GridFTP, and should coordinate calls from srm-copy.

1.5 Policy module

Policy module comes from USC/ISI. Its design and interface specification is documented in another link on <http://sdm.lbl.gov/adapt/>.

2 Component functions

Figure 1 shows the diagram of the new enhancement to the client tools.

2.1 Adaptive Data Transfer Library

Adaptive Data Transfer (ADT) library will be developed as a general transfer adaptation library.

- ADT library is given a "current" transfer rate, and returns the number of streams to use, based on an adaptation algorithm.
- It is not going to make a transfer rate (or performance) estimation. It does not have any knowledge of the on-going performance rate information, transfer volume, or source-destination information.
- ADT library is per-process (or per-job) adaptive transfer adjustment module to maximize the data throughput for the process. It has no knowledge of other processes on the same site or out on the network.
- The ADT library will be integrated into the client tools designated in the project, as well as distributed as a software library under open source license.

2.1.1 ADT library interface

```
Int getNumberOfStreamsSuggestion ( int currentThroughput,
                                   int achievableThroughputPercentage,
                                   int RTT, // Round Trip Time
                                   int maxStreams)
```

2.2 Passive Measurement Archive

srm-copy and BDM collects the transfer rate and log it into a file as it does now.

- A new client transfer monitoring module, Passive Transfer Monitor (PTM) will be developed as an extension library of the client tools to retrieve the current transfer statistics.
- The current transfer statistics include the current transfer rate, total transferred data, the current number of transfer streams, source and destination.
- In the 2nd and 3rd years of the project, the current transfer statistics will be archived into the perfSONAR repository as Passive Measurement Archive (PMA).

2.2.1 PTM library interface

```
Long[] getStatistics ()
    Returns, in order, currentTxfRate in MB/sec,
    totalTransferredData in MB,
    currentNumberOfStreams
```

2.3 Client data transfer tools

A new connection module, Adaptive Transfer Manager (ATM) will be developed as an extension of LBNL client tools (e.g. srm-copy and BDM), and it will integrate original client tool, PTM module library, ADT library and policy module library from ISI.

2.3.1 ATM library interface

Client requests suggestions on the number of transfer streams.

```
Int getNumberOfStreamsSuggestion ()
```

2.4 Notes on the policy module

- Policy module knows about the max bandwidth from source site to destination site, max streams per process (sometimes it's limited by the hardware – e.g. memory size, bus speed, storage speed, etc), max available throughput from source site to destination site,
- Policy module knows the instant limits on max throughput per process (or per job) and max streams per process.
- When policy module is given with the data volume and source/destination sites, it would return the time and transfer estimates based on the policy and some historical transfer rates (e.g. an average of the past transfer rates), and number of streams for the client transfers based on the current limits.
- Policy module can collect the current transfer statistics from client tools through the new extension component PTM.

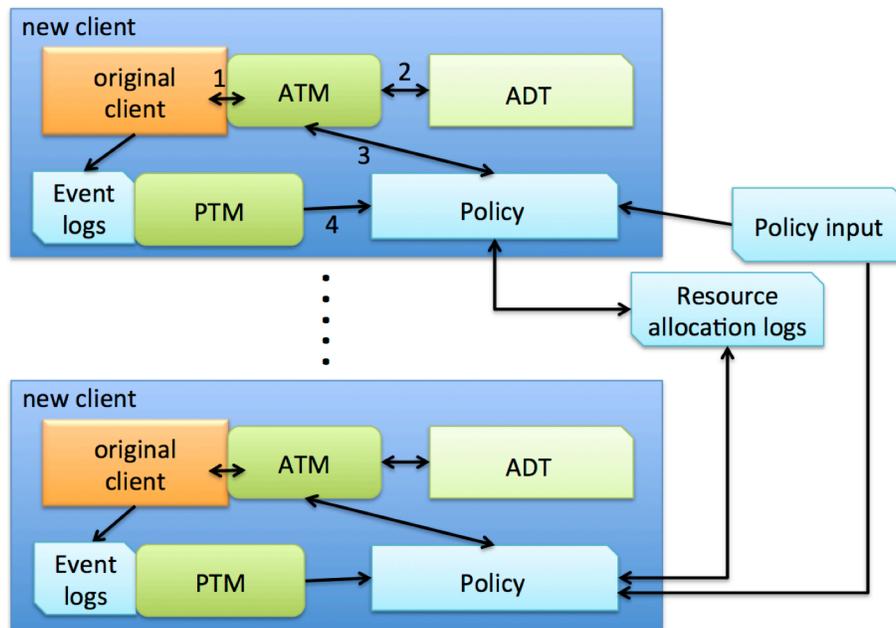


Figure 1: the interactions between the components for the first year deliverables

3 Component Interactions

3.1 Component interactions for the first year deliverables

Figure 1 shows the interactions between the components for the first year deliverables.

- srm-copy as the client tool calls time estimation to ATM via (1) with data volume and source/destination site information.
- Then ATM calls policy module via (3) to find out the estimation and number of maximum transfer streams for the client, and returns the estimation and the number of transfer streams to srm-copy.
- Policy module may collect the current transfer statistics (e.g. transfer rate, transferred data volume) from PTM via (5) that srm-copy has recorded to determine the past transfer measurements as well as measurements from on-going transfers.
- srm-copy may call ATM for adjustment on the number of transfer streams via (1) while the data transfer is in progress.
- ATM will contact policy module via (3) for the various transfer limits (e.g. instant max limit on the number of transfer streams), and call ADT for the adjustments via (2).
- The adjusted data streaming information will then go back to srm-copy as well as to policy module.
- Policy module can save the current streaming information in the “resource allocation logs” for “global” view of the total number of streams on the site. This “resource allocation logs” can be shared among policy modules in different srm-copy processes, so that policy modules can determine the “global view” on the site.

Figure 2 shows the first year deliverable idea for the srm-copy use case, and BDM use case would be similar.

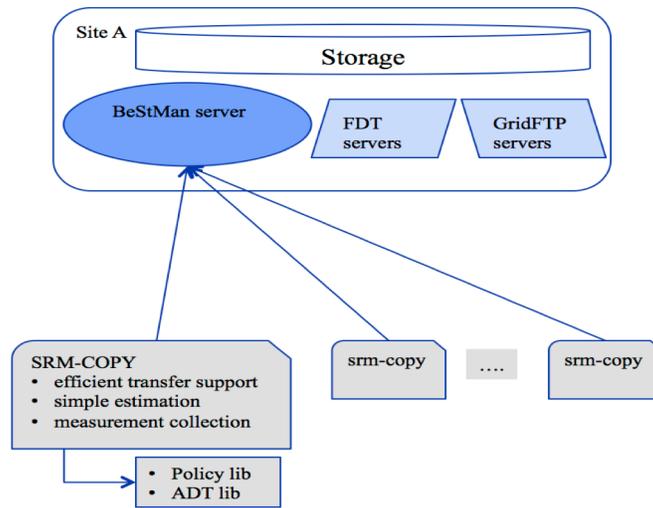


Figure 2: The first year deliverable

Figure 3 shows the use case of the enhanced srm-copy in some of OSG use cases. In these OSG use cases, data transfers by srm-copy should be adaptive to the performance. STAR experiment is shown in figure 4, and BeStMan server is the one that should have adaptive transfers for STAR.

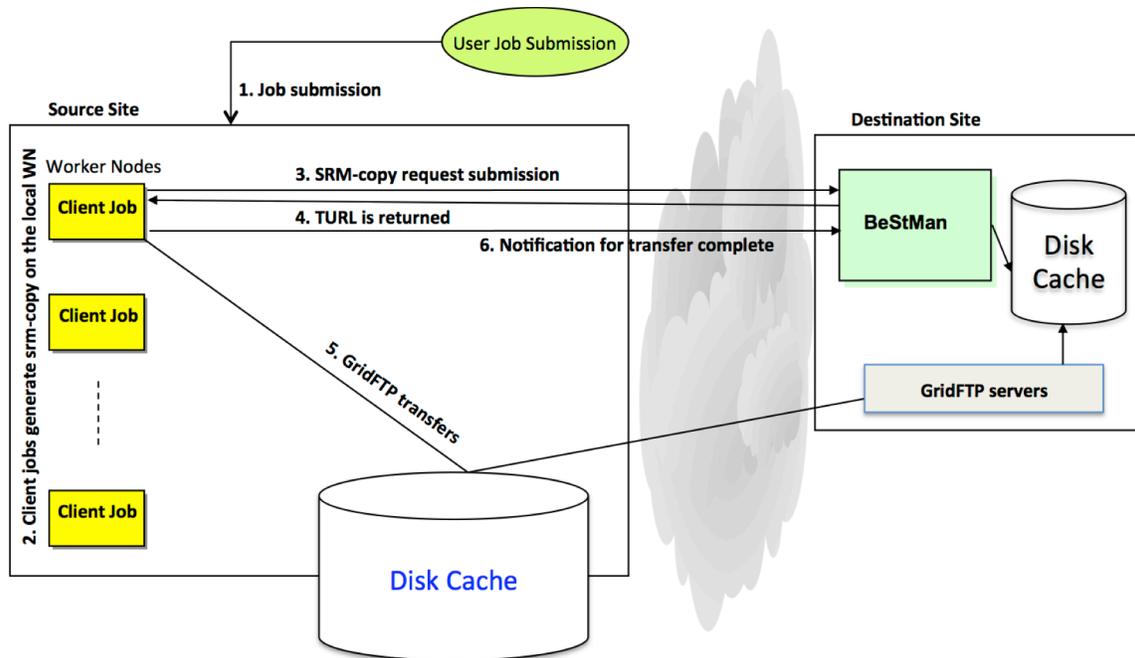


Figure 3: LIGO use case for OSG-based analysis. Source at Caltech and Destination at UNL or one of the OSG sites

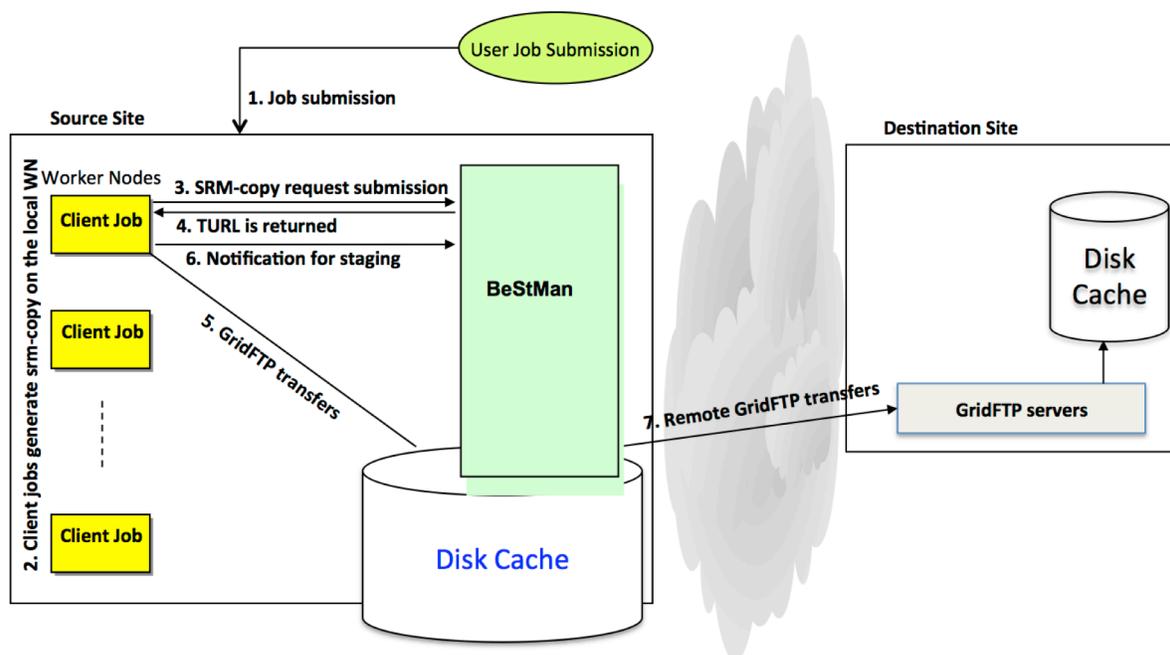


Figure 4: STAR use case at PDSF. Source at NERSC and Destination at BNL

4 Testing environment and measurement for the first year

4.1 srm-copy

4.1.1 OSG use case

Source site will be PDSF/NERSC and destination will be one of the OSG sites such as UNL. We'll have collaboration with the destination site. Running the new enhanced srm-copy jobs on PDSF worker nodes will test all aspects of the first year deliverable with GridFTP transfers. For FDT transfers, we'll use destination host on one of SDM machines and/or one of ISI machines.

4.1.2 STAR use case

Source site will be PDSF/NERSC and destination will be at one of STAR collaborating sites or at ISI for distance consideration. We'll have collaboration with the STAR experiment. Running the new enhanced srm-copy jobs on PDSF worker nodes and BeStMan server on one of the PDSFDTN nodes will test the function of adaptive data transfers with GridFTP and FDT.

4.1.3 Measurements

Transfer performances before and after the enhancements on srm-copy will be collected and compared.

4.2 BDM

Source site will be PDSF/NERSC and destination will be one of the ESGF sites. We'll have collaboration with the destination site. Running the new enhanced BDM on PCMDI node will test all aspects of the first year deliverable with IPCC CMIP5 datasets from NERSC to PCMDI over GridFTP as well as FDT protocols.

4.2.1 Measurements

Transfer performances before and after the enhancements on BDM will be collected and compared.

5 Timeline

5.1 First year deliverables

- By 6/30/2012
- All prototype implementations should be done by 4/15/2012.
 - Enhanced srm-copy and BDM, with ATM, ADT library, PTM and policy module.

- Enhanced BeStMan server

- Test data should be collected for results by 6/15/2012.

5.1.1 components development

- 1) srm-copy: FDT integration and test with a local FDT server
- 2) srm-copy: PTM development as a library module, and integration with srm-copy
- 3) srm-copy: ATM development as a library module, and integration with srm-copy
- 4) srm-copy: gridftp enhancement for connection re-use
- 5) ADT: development as a library module
- 6) Policy module: development as a library module from ISI
- 7) Integration: srm-copy to integrate with ADT and Policy module
- 8) srm-copy: gridftp stream adjustment
- 9) srm-copy: FDT stream adjustment if possible
- 10) bestman: FDT support for TURL as well as bestman transfers in full mode, and integration with ADT, gridftp connection re-use and stream adjustment

5.1.2 Measurement collection

- 1) srm-copy: test on PDSF->BNL (coordinate with Iwona and Jeff), PDSF->UNL (coordinate with Brian) and PDSF->ISI (coordinate with David) for performance without bestman server at PDSF, for the following test cases (clients running on PDSF clusters). Keep the records in a table with the detailed test info.
 - a. GridFTP rates in old srm-copy
 - b. GridFTP rates in srm-copy with ADT and adjustments
 - c. GridFTP rates in srm-copy with ADT, policy and adjustments
 - d. Test (c) with different policies
 - e. FDT rates by itself
 - f. FDT rates in srm-copy with ADT and adjustments
 - g. FDT rates in srm-copy with ADT, policy and adjustments
 - h. Test (g) with different policies
- 2) STAR use case: test on PDSF->BNL and PDSF->ISI for performance with bestman server, for the following test cases. Keep the records in a table with the detailed test info.
 - a. GridFTP rates with old bestman/srm-copy
 - b. GridFTP rates with enhanced bestman/srm-copy
 - c. FDT rates by itself
 - d. FDT rates with enhanced bestman/srm-copy