

Adaptive Data Transfers that Utilize Policies for Resource Sharing

Junmin Gu¹, Robert Schuler², **Ann L. Chervenak²**, Alex Sim¹

¹Lawrence Berkeley National Laboratory

²University of Southern California Information Sciences Institute



Motivation

Scientific data are being collected at unprecedented volumes and rates

The success of large scientific collaborations requires distributed data access with improved latency, reliability

ADAPT Project (Adaptive Data Access and Policy-Driven Transfers)

- Goal: to develop and deploy a general-purpose data access framework for scientific collaborations
- Policy Service and Adaptive Data Transfer library
- Provide framework for the current data movement practices to achieve better data movement performance through adaptivity and policy-based resource allocation

ADAPT Project

Provides fine-grained and adaptive transfer performance

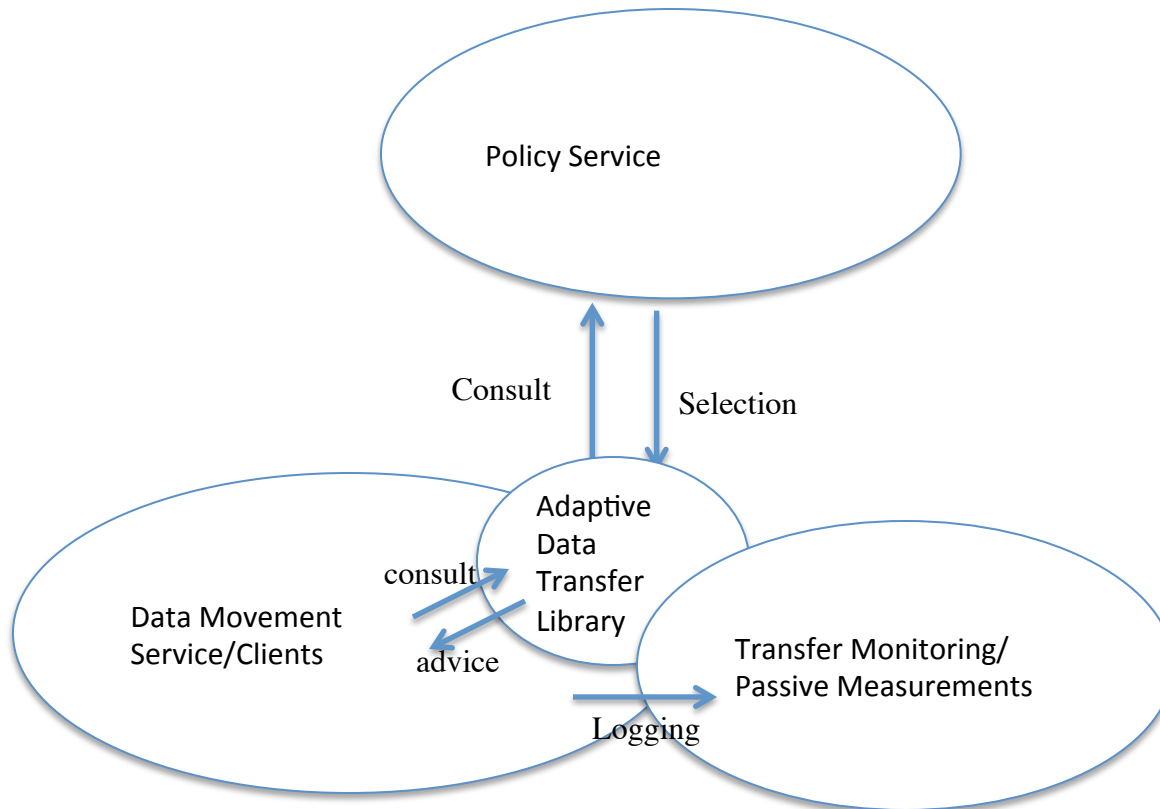
- Select transfer properties based on past performance, available resources
- Adapt properties when observed performance changes due to dynamic load on storage, network, other resources

Based on passive performance monitoring of transfers

Use of site and Virtual Organization policies for resource sharing

- Regarding replication, resource allocation
- Priorities for resources, users
- Balance user requirements for data access with load on resources

System Overview: Component Interactions



- Adaptive Data Transfer (ADT) Library
- Policy Service (PS)

System Overview: Components

- Adaptive Data Transfer (ADT) Library
 - ADT library makes an advice on transfer parameters with a consult from Policy Service (PS)
 - Number of parallel streams, concurrency, buffer size
 - Data movement service or client tools can use the recommended transfer parameters
 - For long-running, multi-file transfers, ADT library:
 - Adapts transfer parameters to accommodate changes in resource availability
 - Makes a new advice with a consult from the PS
 - Existing data movement service/clients can use the ADT library and PS via a well-defined API
 - Reference implementation and use case with srm-copy

System Overview: Components

Policy Service (PS)

- Suggests transfer parameter values based on:
 - Past performance measurements stored in the **perfSONAR**-based measurement archive (if available)
 - Available system resources
 - Maintains allocated resources
 - Site or Virtual Organization policies, such as:
 - Default parallelism for transfers
 - Maximum number of streams between source, destination
 - Resource or user priorities

Implementation

Adaptive Data Transfer (ADT) library

- Periodically, gradually adjust number of concurrent streams up to optimal point
- Example with srm-copy data movement client

Policy Service (PS)

- Generates advice based on historical information, VO & site policies, knowledge of resources allocated
- Stand-alone Policy Service
 - Easier to share policy, resource allocation state across clients, nodes
- Policy library implemented using Drools open source policy engine

Greedy Allocation Policy

- Policy sets a threshold for number of streams or bandwidth allocated between {source, destination} sites
- When a request for transfer advice arrives at PS:
 - Check Resource Allocation Log to determine streams/ bandwidth that have already been allocated
 - If unallocated streams/bandwidth remain below the threshold to satisfy the new request, then return advice to allocate requested resources
 - Otherwise, allocate resources up to threshold
 - Once threshold is reached:
 - Refuse additional requests, or
 - Allocate low level of resources to avoid starvation

Evaluation

Compare performance of adaptive data transfers with unmodified srm-copy client performance

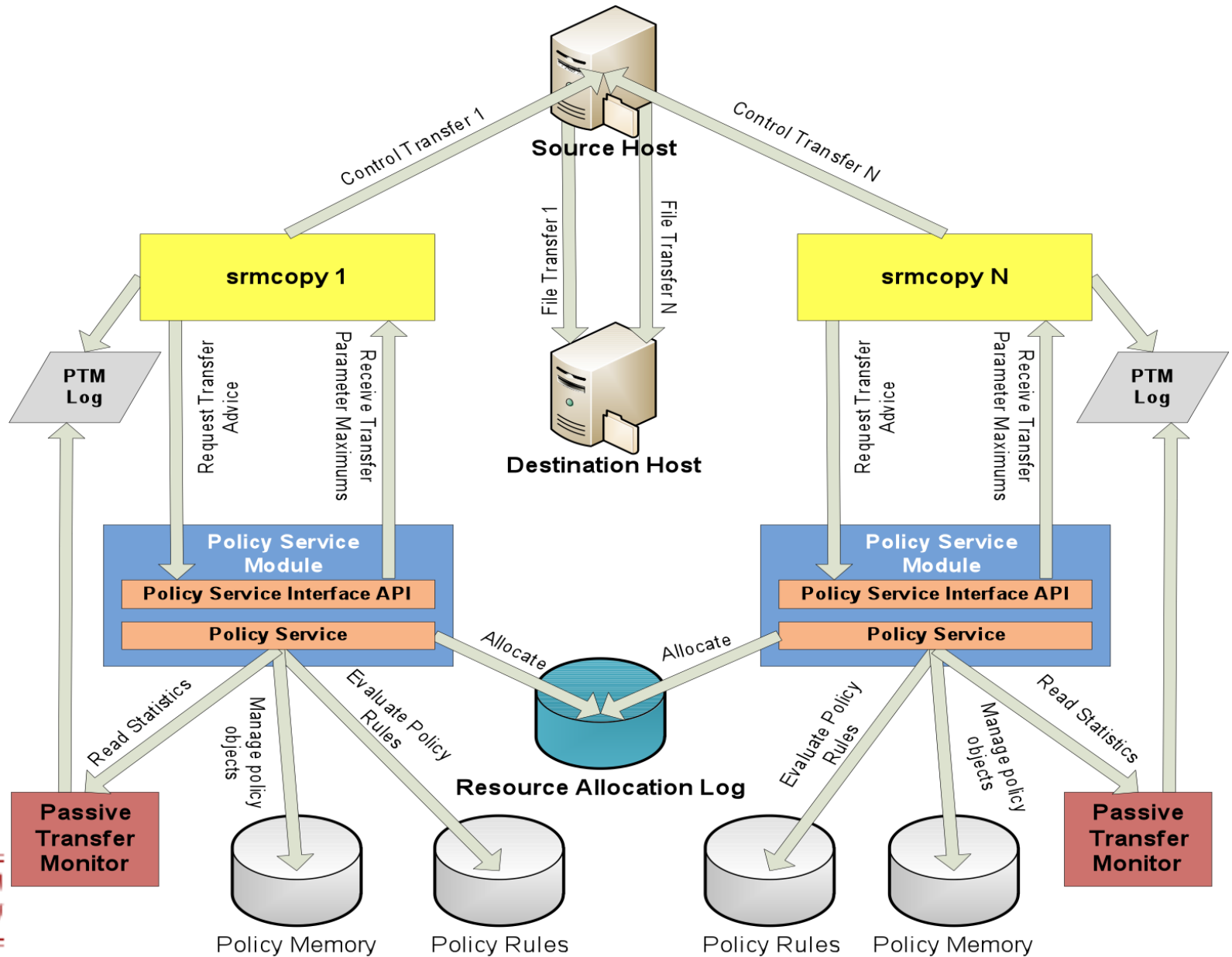
Transfer data from LBNL in Berkeley, CA to Open Science Grid site at University of Nebraska at Lincoln (UNL), 1 Gbit/sec network

6 srm-copy clients performing multi-file transfers: 205 GBytes

Threshold 100 streams; each client requests 80 streams; minimum allocation 4 streams

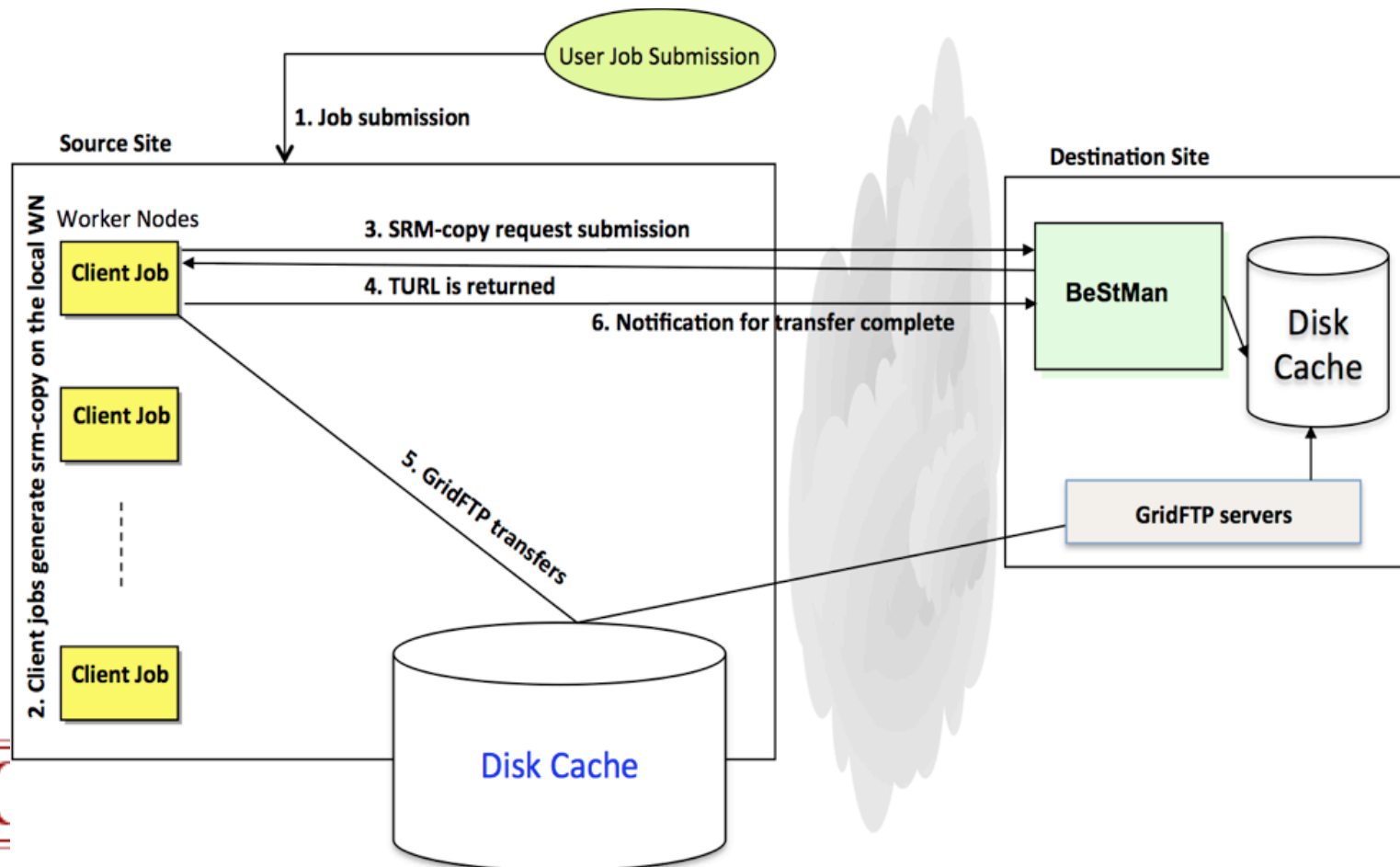
Client	Total data transferred	Total files	File sizes
1	45749 MB	84	45x957MB, 39x132MB
2	43571 MB	80	40x957MB, 40x132MB
3	21786 MB	40	20x957MB, 20x132MB
4	33768 MB	62	31x957MB, 31x132MB
5	31589 MB	58	29x957MB, 29x132MB
6	33768 MB	62	31x957MB, 31x132MB

Evaluation Testbed

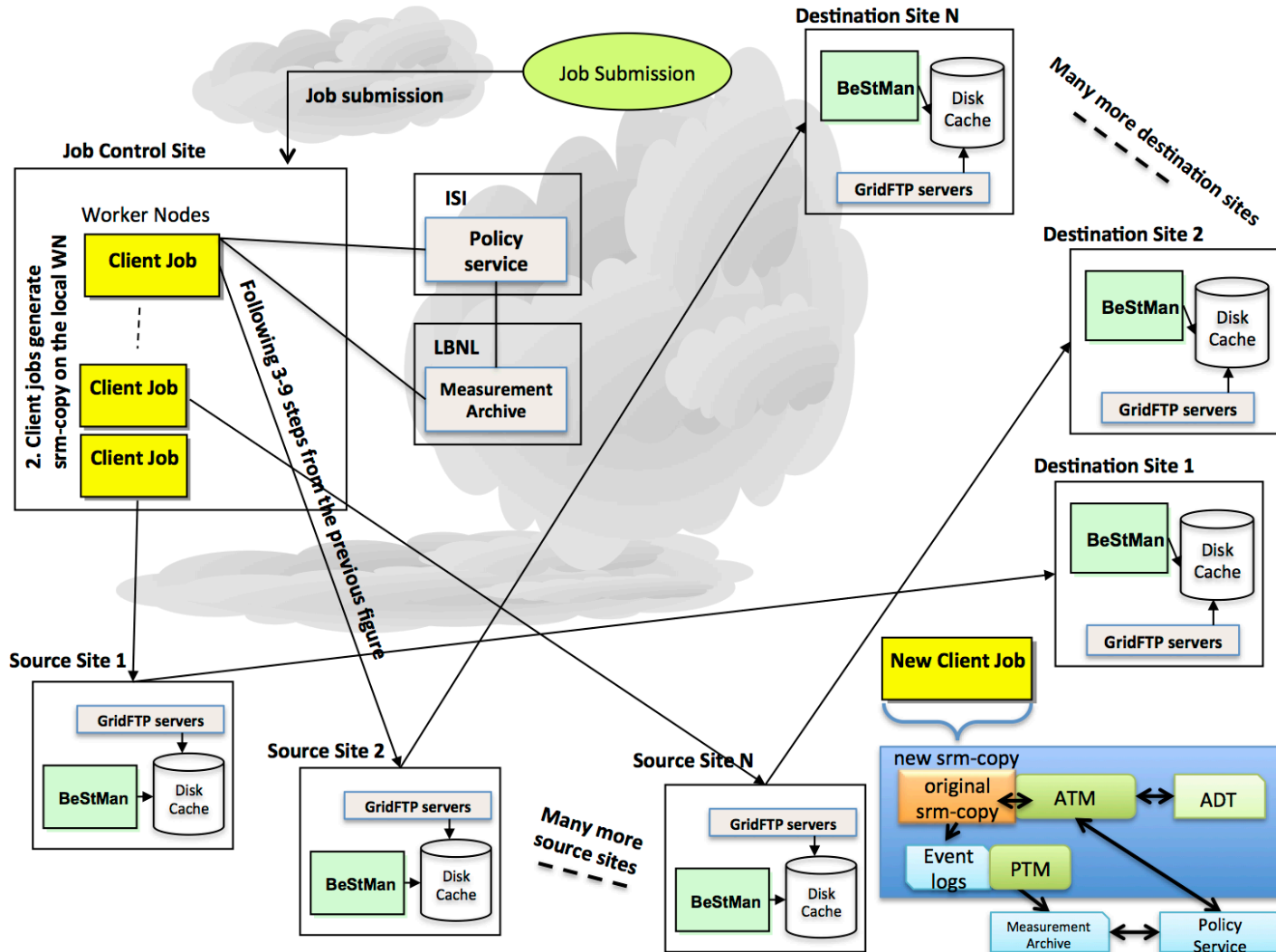


Experiment Represents a Common Use Case (1)

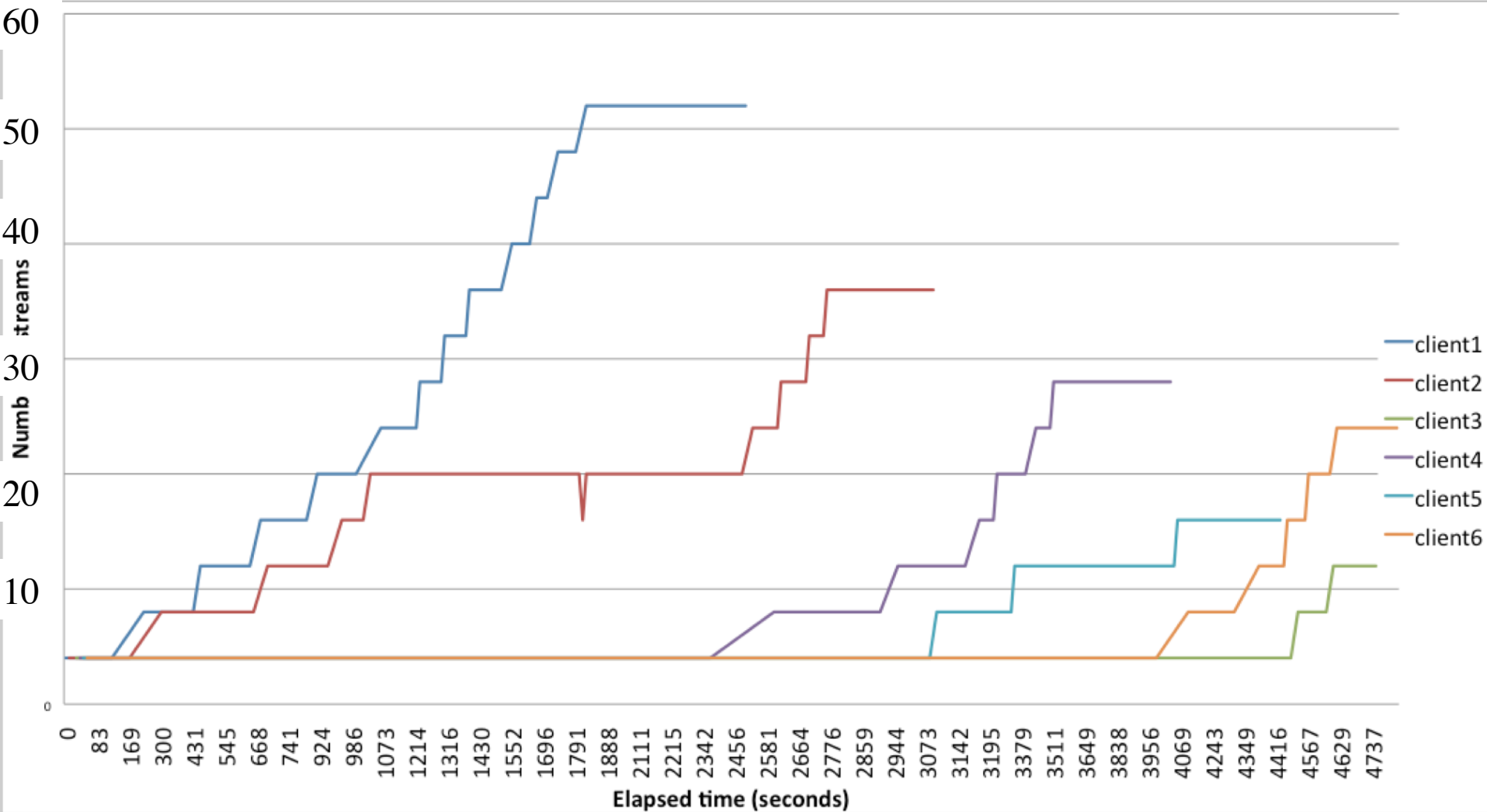
Users want to run analysis on an Open Science Grid site
Must first stage data from a remote location



Common Use Case (2)

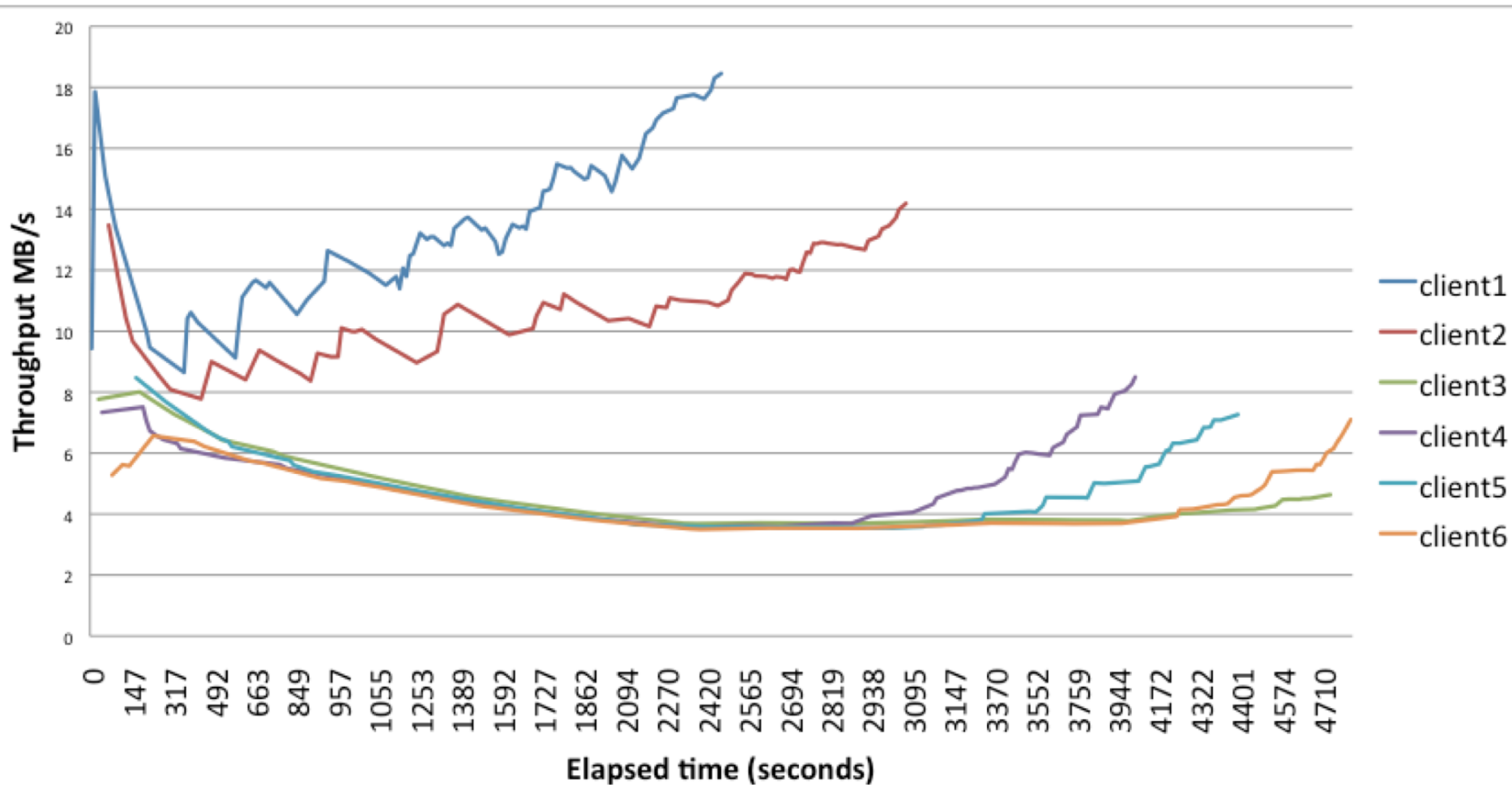


Performance with Adaptation: Number of Streams Allocated vs. Elapsed Time



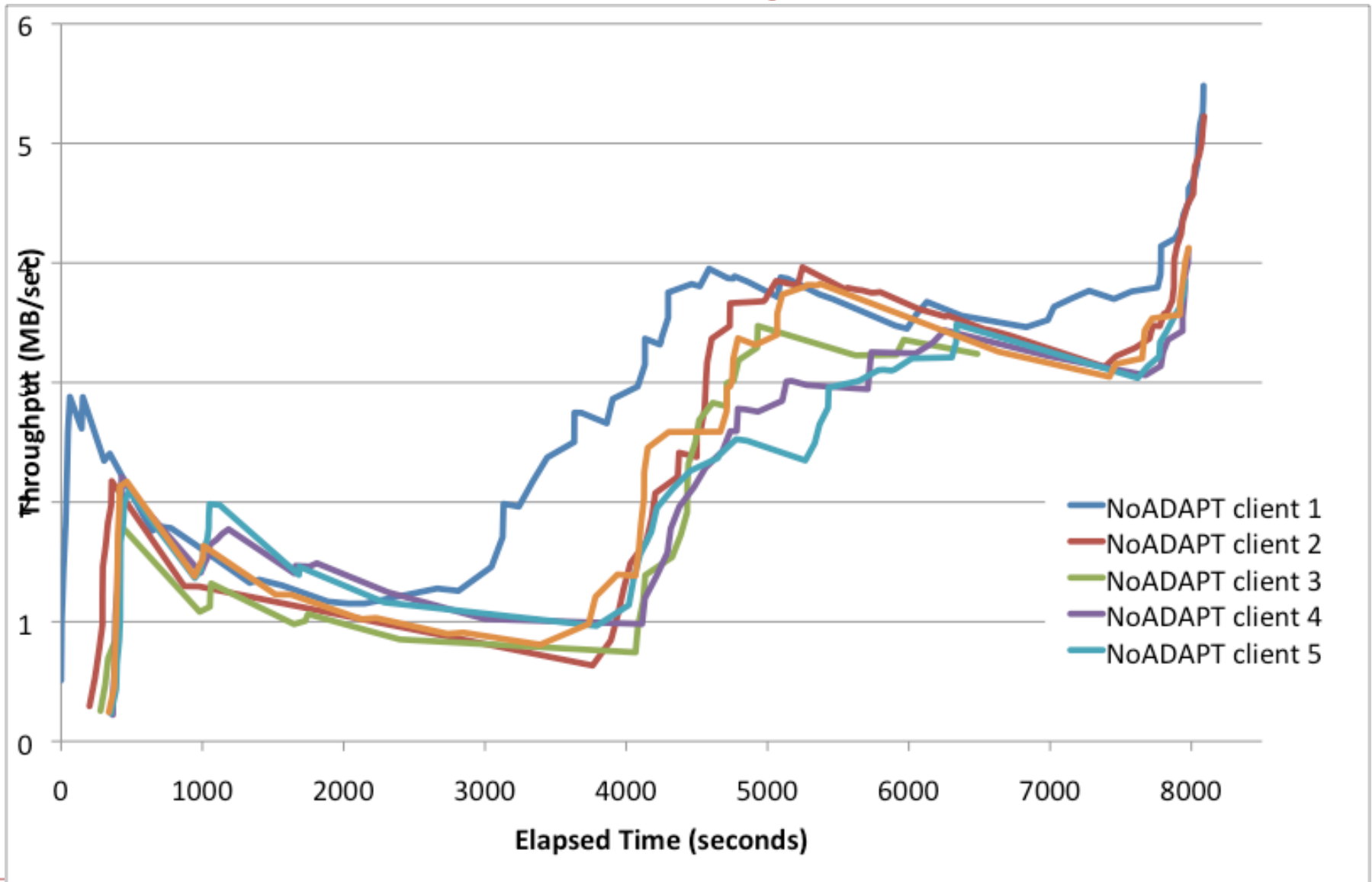
First two clients gradually increase allocation to 80, 20 streams. Other clients get minimum level of 4 streams, then adjust

With Adaptation: Throughput vs. Elapsed Time



Bandwidth increases as additional streams are allocated
All clients achieve at least 3.7 MBytes/sec throughput

Without Adaptation: Throughput vs. Elapsed Time



Network resources become oversubscribed
Overall took 30% longer to complete data transfers

Summary

Plans for the near future:

- Enhancing Adaptive Transfer
- Deploying standalone Policy Service
- Integrating perfSONAR as Measurement Archive
- Exploring richer policies for managing resources, adaptation
- Work with application communities to deploy and evaluate ADAPT software