HTTP Data Federation

Preliminary Plan based on
Discussions with Peter and Brian

HTTP Data Federation

Goals: Demonstrate the ability of using HTTP data federations in a manner analogous to today's AAA infrastructure. Improve CMSSW's HTTP support in case we see increased use of HTTP as a transport protocol.

Aim for a system that can saturate 80 Gbps.

Testbed Setup

The Tier2 facilities at Caltech and several high performant computer systems connected in a high speed network infrastructure.

Monitoring instrumentation with high sampling rate for the IO traffic, disk IO , CPU and the Client/Server software , ….

Simulate long RTT connections using TCP flow control tools

Develop multi-thread test clients to simulate a large number of concurrent real clients with different access patterns .

Compare on the same hardware infrastructure
- **http data access**
- **dav**
- **xrootd direct and via http**

Software Technologies to be evaluated and where it is possible to be improved

libdavix-based StorageFactory (see what extensions may need to be developed)

Xrootd – using the http interface

DAV File systems

Web Servers (Apache , LWAN , NGINX, Lighttpd ….)

Learn from other technologies like Dropbox , Red Hat Storage, ….

Considerations for high performance HTTP data access

Asynchronous IO (epoll like handler for many IO concurrent operations)

Limit the number of parallel disk IO operations (depends on the type of storage systems)

Caching strategies

Map vectorized sequential IO requests into bulk operation

Recovery in case of IO errors (metalink) ?

Considerations for high performance HTTP data access

Asynchronous IO (epoll like handler for many IO concurrent operations)

Limit the number of parallel disk IO operations (depends on the type of storage systems)

Parallel stream in http 2.

Caching strategies

Map vectorized sequential IO requests into bulk operation

Recovery in case of IO errors (metalink) ?

Questions , Suggestions ?