

The Open High Throughput Computing Content Delivery Network

Tuesday, 10 July 2018 11:30 (15 minutes)

LHC experiments make extensive use of Web proxy caches, especially for software distribution via the CernVM File System and for conditions data via the Frontier Distributed Database Caching system. Since many jobs read the same data, cache hit rates are high and hence most of the traffic flows efficiently over Local Area Networks. However, it is not always possible to have local Web caches, particularly for opportunistic cases where experiments have little control over site services. The Open High Throughput Computing (HTC) Content Delivery Network (CDN), openhtc.io, aims to address this by using Web proxy caches from a commercial CDN provider. Cloudflare provides a simple interface for registering DNS aliases of any Web server and does reverse proxy Web caching on those aliases. The openhtc.io domain is hosted on Cloudflare's free tier CDN which has no bandwidth limit and makes use of data centers throughout the world, so the average performance for clients is much improved compared to reading from CERN or a Tier 1. The load on WLCG servers is also greatly reduced. WLCG Web Proxy Auto Discovery is used to select local Web caches when they are available and otherwise use openhtc.io caching. This paper describes the Open HTC CDN in detail and provides initial results from its use for USCMS opportunistic computing and LHC@Home.

Primary author: DYKSTRA, Dave (Fermi National Accelerator Lab. (US))

Co-authors: BOCKELMAN, Brian Paul (University of Nebraska Lincoln (US)); BLOMER, Jakob (CERN); FIELD, Laurence (CERN)

Presenter: DYKSTRA, Dave (Fermi National Accelerator Lab. (US))

Session Classification: T4 - Data handling

Track Classification: Track 4 - Data Handling