A World-wide DataBridge Supported By A Commercial Cloud Provider

• How to reliably collect large output files from volunteers
  • Who can be thousands of miles away from the central storage location
• Larger distances results in a higher RTTs
  • And an increased chance of packet loss during transmission
    • Thus lowering the transfer performance
  • Logs from CMS@home show that network related error codes occur more frequently as the distance from CERN increases
• Optimise the global transfer of files by offering the volunteers a closer upload endpoint to reduce latency
• Use the DataBridge to federate upload endpoints provided by a commercial provider (AWS S3 buckets) and redirect upload requests to an endpoint that is closest to the volunteer
  • Also offloads scalability and operational responsibility of the upload endpoints to the commercial provider
A World-wide DataBridge Supported By A Commercial Cloud Provider
A World-wide DataBridge Supported By A Commercial Cloud Provider

Average File Transfer Time Compared To File Size

Transfer time (s)

File size (MB)

Direct Upload  Redirected Upload