

Instructions for Implementing WLCG Site Network Information

The WLCG Monitoring Task Force is conducting a campaign to describe and instrument the network for the largest WLCG sites in advance of the upcoming WLCG Network Data Challenge 2024(DC24). This document provides a high level view of what site administrators will need to do in response to the campaign.

Summary of Required Tasks

There are two primary components for the network site information that need to be provided:

- A description of the site's network architecture, components and configuration
- Monitoring of the total site IN and OUT traffic, updated every 60 seconds and accessible from CERN MONIT

We are using CERN GitLab to host the relevant instructions, example code and markdown templates at <https://gitlab.cern.ch/wlcg-doma/site-network-information/> Site administrators will need access to this project to allow them to create and maintain their site network information.

Site Network Description

Sites will need to create a site network description following our markdown template at <https://gitlab.cern.ch/wlcg-doma/site-network-information/-/blob/master/Template/SitePageTemplate.md> Instructions:

- Copy the SitePageTemplate.md to the SitePages directory (<https://gitlab.cern.ch/wlcg-doma/site-network-information/-/tree/master/SitePages>), naming the file as RCSITE .md (replace RCSITE with your site's WLCG-CRIC "RC Site" name)
- Edit the RCSITE .md file filling in all the MANDATORY sections as well as any OPTIONAL sections you are willing to provide. Note this file will allow others to help your site identify problems and suggest improvements, so adding more information will allow more effective help.
- Publish the completed file at some URL (web server), accessible to others in WLCG. There are various tools that convert markdown to html (see information on `pandoc` in <https://gitlab.cern.ch/wlcg-doma/site-network-information/-/blob/master/README.md>)
- Last step is to register the URL for your site network description into WLCG-CRIC NetSite Info URL. For sites with more than one NetSite, we suggest describing all net site details in one markdown file and using the same Info URL in each NetSite.
 - Example: AGLT2 RC Site has two NetSites: US-AGLT2 Michigan State University and US-AGLT2 University of Michigan but both use the same Info URL: <https://wlcg-cric.cern.ch/core/netsite/detail/US-AGLT2%20Michigan%20State%20University/> and <https://wlcg-cric.cern.ch/core/netsite/detail/US-AGLT2%20University%20of%20Michigan/>

The intent is that sites own the information they provide and will regularly update information as their network evolves. In the above instructions, we are assuming sites will provide a web server that will host their completed site network description. In the future we may be able to provide a WLCG web server that can utilize the RCSITE .md files in the SitePages directory.

Site Traffic Monitoring

The second component for site network information is enabling tracking of the total network traffic IN and OUT of EACH NetSite. A site's connectivity can vary significantly across WLCG, from single WAN (wide area network) connections to many connections of various bandwidth. For DC24 we want sites to add up all input and output traffic and publish suitably formatted JSON with those numbers every 60 seconds.

- First step is to identify all the network interfaces that connect the site to the wide area network
- Second step is to programmatically gather the IN and OUT statistics for all interfaces, summing up IN and OUT to get totals. Example code using Python3 and SNMP are available in <https://gitlab.cern.ch/wlcg-doma/site-network-information/-/tree/master/WLCG-site-snmp>
- Third step is publish the statistics in the right JSON format every 60 seconds at an accessible URL for CERN MONIT.
- Last step is to register the JSON statistics URL in WLCG CRIC for the NetSite in the Monitoring URL

Sites should verify that the MONIT dashboard starts showing their data within about 24 hours. See https://monit-opensearch.cern.ch/dashboards/goto/f0607fb8528ce6b7c9a336aef74be40b?security_tenant=global

Additional information

There is additional information on the project GitLab web pages (<https://gitlab.cern.ch/-/ide/project/wlcg-doma/site-network-information>) as well as previous presentations on this topic:

- WLCG Operations presentation <https://docs.google.com/presentation/d/1sB4xPJPLoLbNnfV0mfwfXmHXzX8IZ8VtC4TSqDXWUMw/edit?usp=sharing>
- CHEP 2023 Poster <https://docs.google.com/presentation/d/1yzD3Gm6Ph8IAGf3c0WupL3m4L96zT61mASx5dVpBt-Q/edit?usp=sharing>

Questions can be directed to the WLCG Monitoring Task Force (CERN egroup WLCGMon-TF).