

Data Retention Practice of SACLA Data System



Published on: 1st July, 2020

Enforced on: March, 2021

Issued by: The administrator of the SACLA data system

Preface

The practice rules were established in compliance with the data retention policy enacted by the SPring-8 Data and Network Committee¹. This document describes the rules of data retention and disposal in the SACLA data system.

¹ http://dncom.spring8.or.jp/document/pdf/exp_data_policy.pdf

Revision History

Date	Description
30.6.2020	First issue

1. Data Retention & Disposal

The rules on data retention and disposal in the SACL A data system are defined for the three data categories (A, B, and C) specified in the data retention policy of the SPring-8 Data and Network Committee. Table 1 shows data retention rules for respective data typ.

A) Data collected with the facility's photon diagnostic systems, such as beam intensity monitors and wavelength monitors (Data stored in the event-synchronized database (SyncDB)).

These data are retained for 10 years since they are collected. For the first three years, the data are stored in the disk devices accessible from the SACL A HPC system. After the three years, the data are transferred to the tape archive inaccessible from the HPC system and retained for another seven years before they are deleted.

B) Raw data from SACL A official detectors (such as MPCCD, OPAL, IMPERX)

These data are retained for eight years since they are collected. The data are stored in the disk devices accessible from the HPC system for the first three years. After the three years, the data are transferred to the tape archive inaccessible from the HPC system and retained for another five years before they are deleted.

C) User-generated data

User-generated data are stored in one of the three directories below according to the users' purpose. The data retention rules for each directory are as follows.

① `/home/{account_name}`

The quota limit for this directory is 100 GiB. The data in this directory will not be deleted automatically.

② `/work/{account_name}`

This directory is intended to be used for short-term storage. Every user can have this directory. The default quota limit for this directory is 5 TiB. Each file is automatically deleted after 90 days from the date of the latest access to the file.

Each group can increase the quota limit up to 100 TiB for a month at the longest by asking SACLA HPC helpdesk if they need more quota temporarily, for example, during a beamtime.

③ `/UserData/{account_name}`

This directory is intended to be used for mid-term storage. The quota limit for the directory is 100 TiB for each user. Each file in this directory is automatically deleted after three years from the date of creation.

2. [Revision of the Data Retention Practice Rules](#)

The administrator of the SACLA data system can change the practice rules and the quota limitations according to the actual usage state of the system. When the administrator makes any changes to the practice rules, the administrator must notify the changes to users, for example, on this website and/or the SACLA HPC portal site.

Table 1. Data retention rules for respective data type.

Data type	Quota limit	Retention period	Comment
A) SyncDB	-	10 years	
B) Storage raw data	-	3 years	/b11strg /b12strg /b13strg
# After the above period, archived on tape devices	-	+5 years	
C) User-generated data			
• /home/<account_name>	100 GiB	-	User code
• /work/<account_name>	1-10 TiB	90 days	Temporary data
• /UserData/<account_name>	100 TiB	3 years	Mid-term data