

Workpackage 7

Data Analysis and Management

Report on the required hardware and software for the tomography beamline

D 7.1

August 2020



PROJECT DETAILS

PROJECT ACRONYM

BEATS

PROJECT TITLE

BEAmline for Tomography at SESAME

GRANT AGREEMENT NO:

822535

THEME

START DATE

01/01/2019

DELIVERABLE DETAILS

WORK PACKAGE: 07

EXPECTED DATE: 31/08/2020

WORK PACKAGE TITLE: DATA HANDLING AND MANAGEMENT

DELIVERABLE TITLE: REPORT ON THE REQUIRED
HARDWARE AND SOFTWARE FOR THE TOMOGRAPHY
BEAMLINE

WORK PACKAGE LEADER: THE CYPRUS INSTITUTE

DELIVERABLE DESCRIPTION: REPORT

DELIVERABLE ID: D7.1

PERSON RESPONSIBLE FOR THE DELIVERABLE:
CHARALAMBOS CHRYSOSTOMOU

NATURE

- R - Report P - Prototype D - Demonstrator O - Other

DISSEMINATION LEVEL

- P - Public
 PP - Restricted to other programme participants & EC:
 RE - Restricted to a group
 CO - Confidential, only for members of the consortium

REPORT DETAILS

VERSION: 1

DATE: 31/08/2020

NUMBER OF PAGES: 14

DELIVERABLE REPORT AUTHOR(S): **CH.
CHRYSOTOMOU, MUSTAFA ALZU'BI,
G.JORI**

FOR MORE INFO PLEASE CONTACT: C.CHRYSOSTOMOU@CYI.AC.CY

STATUS

- Template Draft
 Final Released to the EC

CONTENTS

Introduction	4
Hardware definition for the beats beamline	5
Data Acquisition system	6
Beamline Control Station	6
Hybrid CPU/GPU reconstruction cluster	6
Preview and on-the-fly processing station	7
Short Term Storage server	7
Data analysis and visualization workstation	7
Network connection components	7
Software definition for the BEATS beamline	11
Synergy between BEATS and The Cyprus Institute (CYI)	13
Network connectivity	13
Storage and Computational Infrastructure	13
References	14

INTRODUCTION

Computed tomography beamlines are considered among the most massive data generation beamlines in synchrotron facilities, where one experiment can generate up to 2 Terabyte of raw and processed data. Accordingly, they require dedicated high-end computational and network resources. An efficient computed tomography data acquisition and processing pipeline produces, within minutes or hours, a large amount of raw data that must be stored by the facility and processed in order to deliver to users fully processed data sets (reconstructed 3D volumes) as soon as the experiment is completed.

The goal of Task 2 of the BEATS Work Package 7 (WP7) was to investigate and identify hardware and software requirements for BEATS. The proposal of hardware equipment and software tools detailed in the present deliverable follows guidelines and examples already established by tomography beamlines at other synchrotron facilities and was adapted to the particular needs of SESAME and of the BEATS scientific case. This report describes the computational infrastructure, hardware and software, needed to commission and operate BEATS. The design of a computing infrastructure for BEATS described in the following paragraphs is the result of consultations with partner laboratories and IT groups, in particular at the ESRF and PSI, that were conducted by the BEATS WP7 during the first 19 project months.

HARDWARE DEFINITION FOR THE BEATS BEAMLINE

The pipeline of the BEATS data acquisition, processing and storage infrastructure is illustrated in Figure 1. It is composed of hardware equipment that will be installed mainly in two locations: the BEATS control hutch and the SESAME server room. The main components of the pipeline are described in detail in the following sections together with the design principles and tentative hardware definition:

- Data Acquisition system
- Beamline Control Station
- Hybrid CPU/GPU reconstruction cluster
- Preview and on-the-fly processing station
- Short Term Storage server
- Data analysis and visualization workstation
- Network connection components

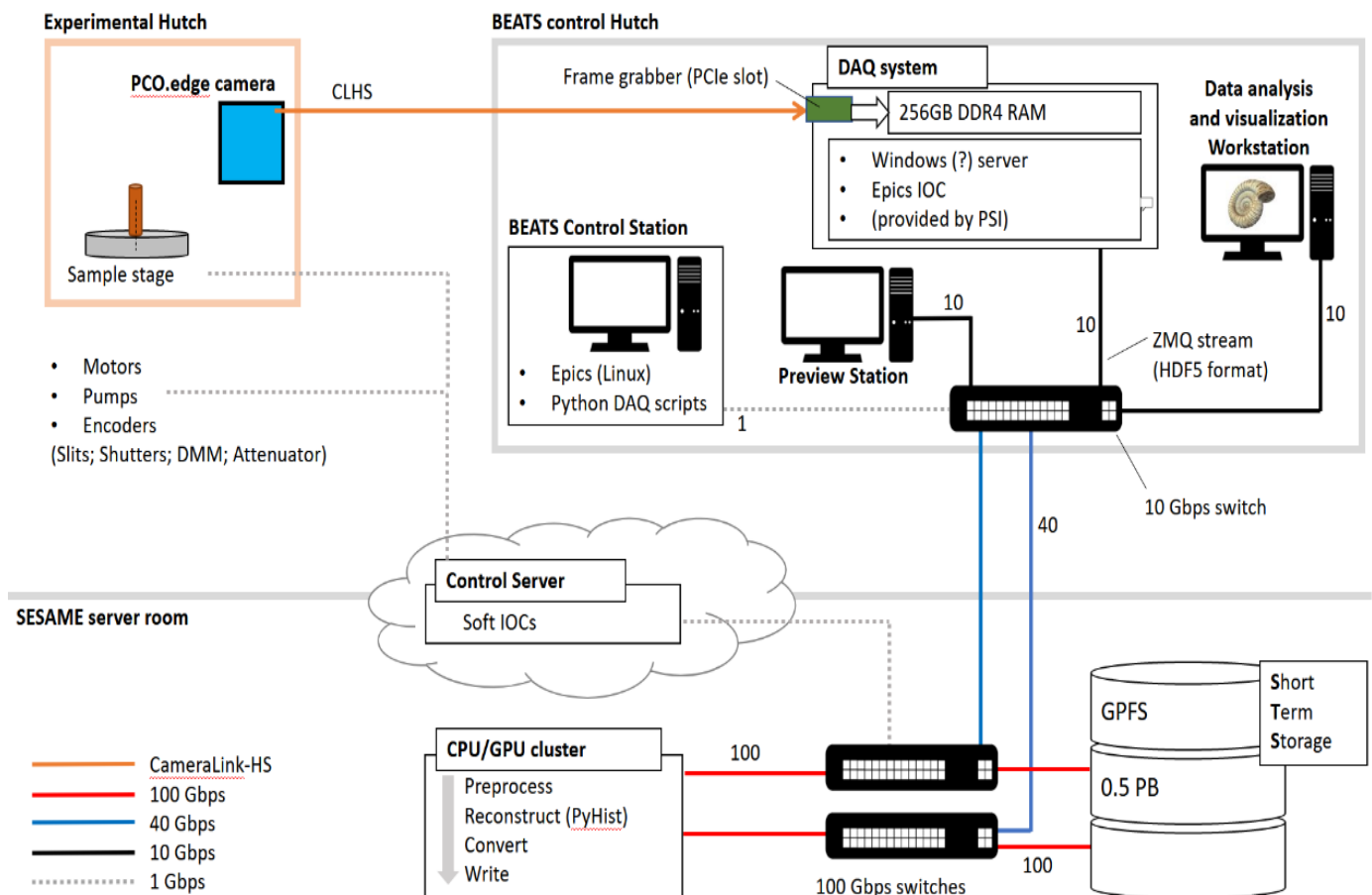


Figure 1: Computing infrastructure for the BEATS data acquisition system

Data Acquisition system

From the computational point-of-view, the process starts by collecting single frames (radiographic projections) from a high-speed sCMOS-sensor camera and transferring them to its designated destination in a Short Term Storage (STS) server. This process is realized by a backend Data Acquisition (DAQ) system located in the beamline control hutch. The DAQ system is integrated into EPICS, the Experimental Physics and Industrial Control System (Advanced Photon Source, Argonne National Laboratory, Argonne, IL 60439, USA) via a Windows or Linux server running an EPICS Input/Output Controller (IOC). EPICS is the standard used for control systems at SESAME. Raw data is transferred from the camera to the DAQ system via a Camera Link HS (CLHS) cable interface with a dedicated PCIe frame grabber. The DAQ system outputs a stream of data, which is written to the short-term storage in one single, HDF5-format, file using the ZeroMQ distributed messaging protocol (ZeroMQ, 2020). The file written to the STS contains RAW projections as well as experimental metadata compliant with the SESAME Experimental Data (SED) internal standard. The Area-Detector EPICS framework is the primary software for detector control which includes drivers, post processing plugins, data formats manipulation libraries and data visualization. PSI will provide a DAQ server for the PCO.Edge camera (PCO.Edge5.5) chosen for BEATS and integrated into EPICS. The integration of this system will be performed by the control and computing groups of SESAME. For this, consultations have already been undertaken and on-site training at PSI is foreseen. The final definition of a metadata template for the HDF5 files written to the STS will be developed during the upcoming project months in view of the latest developments at partner laboratories. The metadata will in general remain configurable by the user.

As a tentative hardware definition, the DAQ system will be equipped with one CPU with 8 cores, 256 GB of RAM, 512 M.2 SSD, 2 TB HDD, and a dedicated 10 Gbps network card.

Beamline Control Station

Existing PSI (TOMCAT beamline) control software will be utilized and adapted for the BEATS control station. The data acquisition system as well as all the beamline equipment in the experimental and optics hutch and in the front end will be accessible from here. Hardware information and settings will be accessed and controlled through EPICS within a Linux environment. Python DAQ scripts will use PyEpics: the library for Epics Channel Access for Python developed by the GSECARS team at Argonne (PyEpics, 2020). Communication between the DAQ system and beamline control station will occur through EPICS channels.

Beamline control station hardware definitions: Intel i7 9th gen., 16 GB RAM, 240 GB SATA, 1 TB HDD, Basic graphics card, 2 x 32" screen.

Hybrid CPU/GPU reconstruction cluster

The pre-processing and reconstruction of the primary input raw data (projections) will be performed on a dedicated hybrid CPU/GPU cluster using high performance (i.e. parallelized, distributed) software. The cluster can additionally be used for in-house/remote data analysis, providing a basic and appropriate set of 3D data analysis tools.

Hardware definitions: Technically, by comparing the technology available nowadays to eventual evolution within the next three years (i.e. once the beamline is in operation), we foresee the cluster to consist of 4 computing nodes, dual CPU sockets each, 18 actual cores for each CPU, 8 GB RAM for each core. On the other hand, for the computational GPUs power, two nodes will be equipped with 8 TFLOPS accelerator cards.

Preview and on-the-fly processing station

Preview of tomographic reconstructions as well as on-the-fly visualization capabilities will be offered by a further system to be installed in the control hutch. This system will be equipped with a graphics card and one or two 32" screens and will be available for the launch of preview reconstructions on a subset of the whole data. Users can interact with tomographic reconstruction with different software packages both from the command line or through a Graphical User Interface to facilitate the operation for less experienced users. As an alternative to this dedicated system, preview functionalities can be implemented on the same reconstruction hybrid cluster and accessed through remote connection from the BEATS control station.

Preview and on-the-fly processing station hardware definitions: Intel i7 9th gen., 128 GB RAM, 240 GB SATA, 4 TB HDD, 6 GB pro graphics, 2 x 32" screen.

Short Term Storage server

A high-end experimental data storage will be directly connected to the hybrid CPU/GPU cluster. This storage will be used as a centralized Short-Term Storage (STS) to keep the raw data for a duration in compliance with the SESAME data policy. The storage size and specifications for STS have been discussed during a meeting for WP7 at PSI. **As a result, the storage size was estimated to be 0.5 PB using the General Parallel File System (GPFS) file system standard.**

For the sake of storage space optimization, only raw data will be stored in the STS by default. The beamline user will have the possibility to take home the tomographic images in Tagged Image File Format (TIFF) with the desired resolution.

Data analysis and visualization workstation

SESAME and Cyl will provide tomography data analysis as a service across two different geographical locations, one at the SESAME premises (workstation situated in the beamline control hutch) and another at Cyl. The analysis and postprocessing of microtomography datasets are computationally intense and rely on high-performance 3D visualization and analysis software. High-end data analysis workstations with big RAM, powerful CPUs and GPUs are needed. The main system will be installed in the control room of BEATS. A twin system of the same kind, also dedicated to the BEATS users, can be installed in the premises of the Cyl. These workstations will be equipped with preinstalled 3D data analysis and visualization applications. According to the community needs, the BEATS workstation will be available for data visualization and analysis during user beamtime. At the same time and depending on the availability of resources, remote access to both systems can be granted for offline tomography data analysis from the user's home institute or in-house.

Data analysis and visualization workstation hardware definitions: 16 cores, 512 GB RAM, 1 TB M.2 SSD, 10 TB HDD, 10 Gbps network card, 12 GB pro graphics card, USB 3 ports, 2 X 32" screens.

Network connection components

To connect all systems, sub-systems, components and modules, a dedicated network with diverged bandwidth is needed. Currently, the backbone network speed at SESAME is 10Gbps, which is not sufficient to carry out the BEATS data handling and activities. Therefore, a modified network architecture is proposed as shown in Figure 2. The DAQ and the data analysis workstation shall be connected with a 10Gbps switch. The 10Gbps switch will be directly connected to the data centre switches over a 40Gbps uplink interconnect. The data centre switches will be mainly used to interconnect the GPU/CPU and storage nodes at 100Gbps speed, beamline switch at 40Gbps speed and the rest of the SESAME network at 10Gbps speed. **A detailed description of the minimum requirement for the network connection components is provided in Table 3.**

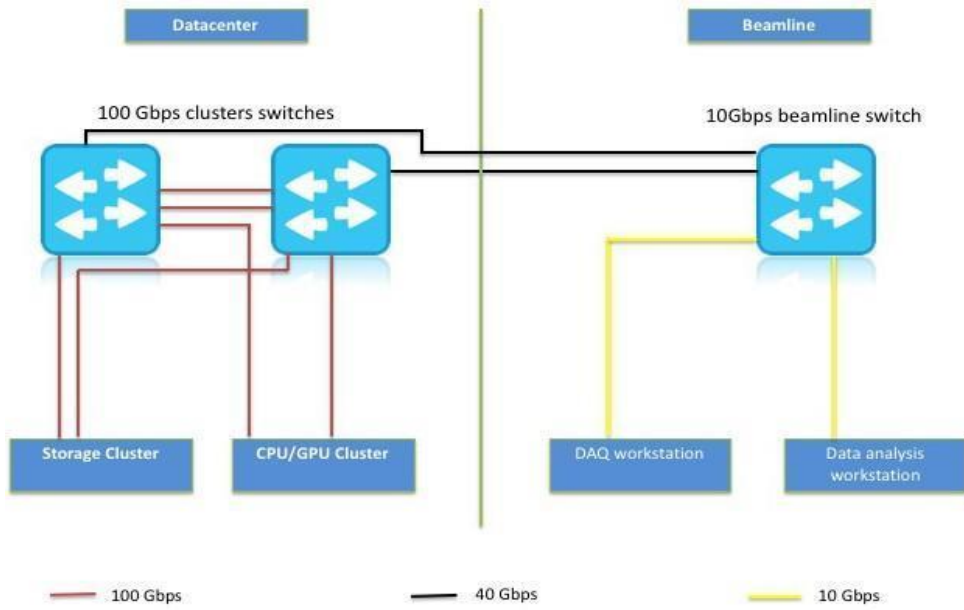


Figure 2: General network chart at beamline and data centre

Table 2 shows a list of the basic hardware required for the BEATS computational infrastructure.

Table 2: List of minimum computational resources

Component	Tech. Specs	QTY	Use Description	Category	Location
Short Term Storage server	0.5 PB, GPFS, 15-20 GB/s aggregated	1	BEATS experimental data storage (short-term storage)	Computing HW	Data Centre
Hybrid CPU/GPU cluster	4 nodes cluster, dual-socket CPU each, 18 cores/CPU, 8 GB RAM/core, 8 TFLOPS accelerator cards, Infiniband interconnect	1	μ-tomography images reconstruction, 3D volumes generation and image processing	Computing HW	Data Centre
Workstation	Intel Xeon one socket 8 cores, 256 GB RAM, 512 M.2 SSD, 2 TB HDD, 10 Gbps network card, USB 3 ports	1	Data acquisition Workstation	Computing HW	Beamline
High-end PC	Intel i7 9th gen., 16 GB RAM, 240 GB SATA, 1 TB HDD, Basic graphics card, 24" screen	1	General use PC	Computing HW	Beamline
Workstation	Intel i7 9th gen., 16 GB RAM, 240 GB SATA, 1 TB HDD, Basic graphics card, 2 x 32" screen.	1	Beamline Control Station	Computing HW	Data centre
Workstation	Intel Xeon dual socket 8 cores each, 512 GB RAM, 1 TB M.2 SSD, 10 TB HDD, 10 Gbps network card, 12 GB pro graphics card, USB 3 ports, 2 X 32" screens		Data analysis workstation	Computing HW	Experimental hall
Workstation	Intel i7 9th gen., 128 GB RAM, 240 GB SATA, 4 TB HDD, 6 GB pro graphics, 2 x 32" screen.		Preview and on-the-fly processing station	Computing HW	Experimental hall
UPS	3 KVA Online UPS	1	Power backup for the switch and other control components		Beamline
Computing and control rack	42U floor standing standard cabinet, compatible networking hardware (Depth 80cm).	2	To install the computing and control equipment at the BL station side.	Passive Components	Data centre and beamline
10 Gbps network switch	10Gbps 16, 24 or 32 ports, 40Gbps uplink	1	For the data acquisition system connected with the central storage and the computing cluster in the data centre.	Networking (Access switches)	Beamline
1 Gbps network switch	1Gbps 48 ports, 10Gbps uplink	1	For general network use, access control, PSS system, cameras, control system, vacuum components, ... etc	Networking (Access switches)	Beamline
High throughput switches (Data centre switches)	40/56 Gbps network switches	2	High available data centre switches to support 40 Gbps uplinks coming from the beamlines	Networking (Data centre switches)	Data Centre
Fibre patch panels	OM4 SM 10Gbps fibre patch panel, 4 ports, pigtails and castle included	1	Fibre patch panels	Networking (Passive components)	Beamline
Ethernet patch panel	24 ports cat 6A patch panel	2	For copper cables termination	Networking (Passive components)	Beamline
Cable guide	2U cable guide	4	to be used as cable organizer	Networking (Passive components)	Beamline
Fibre Patch Cord	OM4 MM 3 meters connector type (LC / SC)	4	To connect the 1Gbps switch with the core switches at the data centre	Networking (Passive components)	Beamline and data centre

D7.1 HW and SW for the BEATS beamline

Ethernet patch cords	Patch cord Cat.6, RJ45 unshielded LSZH 1m	48	To connect the snap-ins in the patch panel with switch, also to connect the devices in the hutches	Networking (Passive components)	Beamline
Ethernet patch cords	Patch cord Cat.6, RJ45 unshielded LSZH 5m	48	To connect different devices at the beamline with network points	Networking (Passive components)	Beamline
Ethernet patch cords	Patch cord Cat.6 A, RJ45, LSZH, 1m	10	To connect the snap-ins in the patch panel with switch, also to connect the devices in the hutches	Networking (Passive components)	Beamline
Ethernet patch cords	Patch cord Cat.6 A, RJ45, LSZH, 5m	10	To connect the snap-ins in the patch panel with switch, also to connect the devices in the hutches	Networking (Passive components)	Beamline
Fibre cables installation and termination	4 core Multi-Mode (MM) fibre cable OM4 40G, outdoor and armoured	150 m	Connect the beamline to the data centre	Networking (Passive components)	Experimental hall, BL to DC
Cat6A network cables	F/FTP AWG23 Cat 6A LSZH orange 500m reel	2	Network wiring inside the hutches	Networking (Passive components)	Beamline
Core, data centre and access switches configurations	Soft configuration	NA	to configure the security, vlans, uplinks, accessibility	Configuration	Data centre

SOFTWARE DEFINITION FOR THE BEATS BEAMLINE

All BEATS workstations will be capable of running both Windows and Linux operating systems to satisfy the needs of the beamline. The short-term storage will be running on the well-established General Parallel File System (GPFS) file system. For the Hybrid CPU/GPU reconstruction cluster SLURM will be used for scheduling the reconstructions. PSI (TOMCAT beamline) control software will be utilized and adapted to the BEATS control station. Table 3 lists the proposed software for BEATS for tomographic reconstruction and 3D tomographic data analysis and visualization.

Table 3: Preliminary tomographic reconstruction, 3D data processing and visualization software for BEATS.

#	name	URL	features	open source	license type
CT reconstruction					
	PyHST	http://ftp.esrf.fr/scisoft/PYHST2/	parallelization, computations distribution	yes	
	Nabu	http://www.silx.org/pub/nabu/doc/about.html	parallelization, computations distribution	yes	
	TOMWER	http://www.edna-site.org/pub/doc/tomwer/latest/	reconstruction workflow manager	yes	
	TomoPy	https://tomopy.readthedocs.io/en/latest/	parallelization, computations distribution	yes	
	SYRMEP Tomo Project (STP)	https://github.com/ElettraSciComp/STP-Gui	Graphical user interface	yes	
3D data processing and visualization					
	ImageJ	https://fiji.sc/		yes	
	Paraview	https://www.paraview.org/		yes	
	Dragonfly	https://www.theobjects.com/dragonfly/index.html	GPU ready; ML ready	no	Academic; single user
	3D Slicer	https://www.slicer.org/		yes	
	PALABOS	https://palabos.unige.ch/	fluid dynamics simulations; permeability	yes	
	Quanfima	https://github.com/rshkarin/quantfima	Python library for quantitative analysis of fibrous materials	yes	
	Calculix	http://www.calculix.de/	FE solver	yes	
	iMorph	http://imorph.sourceforge.net/index.html	morphological characterization of cellular materials	no	free

SYNERGY BETWEEN BEATS AND THE CYPRUS INSTITUTE (CYI)

Network connectivity

The Cyl currently has a 0.6 Gbps (gigabits per second) direct internet connection, with plans to be upgraded to 1.0 Gbps by the end of 2021. The CYI forms part of the CYNET (Cyprus Research and Academic Network), and this Cypriot NREN (National Research and Education Network) is in turn a national member of the pan-European GÉANT network. SESAME forms part of the JUNet (Jordanian Universities Research and Learning Network), and this Jordanian NREN is in turn a member of ASREN (Arab States Research and Education Network). Both networks, the GÉANT through CYNET and ASREN through JUNet, will be utilized to optimize the transfer of data between the Cyl and SESAME.

Storage and Computational Infrastructure

The Cyl currently operates the CY-TERA supercomputer with peak / sustained performance of around 600 TFlop/s, and 33 nodes are available (17 nodes with two Intel Xeon Gold 6248 each and 16 GPU nodes with two Intel Xeon Gold 6248 and 4 NVIDIA Tesla V100 SXM2 32 GB each). The memory on the compute nodes is 96GB and on the GPU nodes 192GB. The system also has a HDR100 Infiniband and 135TB of flash storage for temporary storage and more than 20PB for long term storage.

Two access modalities to the storage and computational infrastructure will be available within the BEATS project: preparatory and production access. Preparatory (on-demand) access will serve for the development and testing of codes required for the project, and will be granted to Cyl and SESAME staff. Production access to the computational resources will be allocated to the users of SESAME for the implementation of their scientific projects and for the analysis of data collected at BEATS. Every six months, an estimation will be established for the expected users for computational resources and storage, and the available resources will be adjusted.

More information on the two access modes can be found at: http://bit.do/cyi_access

REFERENCES

“The ZeroMQ Project.” n.d. GitHub. Accessed June 27, 2020. <https://github.com/zeromq>.

“PyEpics Overview — Epics Channel Access for Python.” n.d. Accessed July 1, 2020. <https://cars9.uchicago.edu/software/python/pyepics3/overview.html>.