



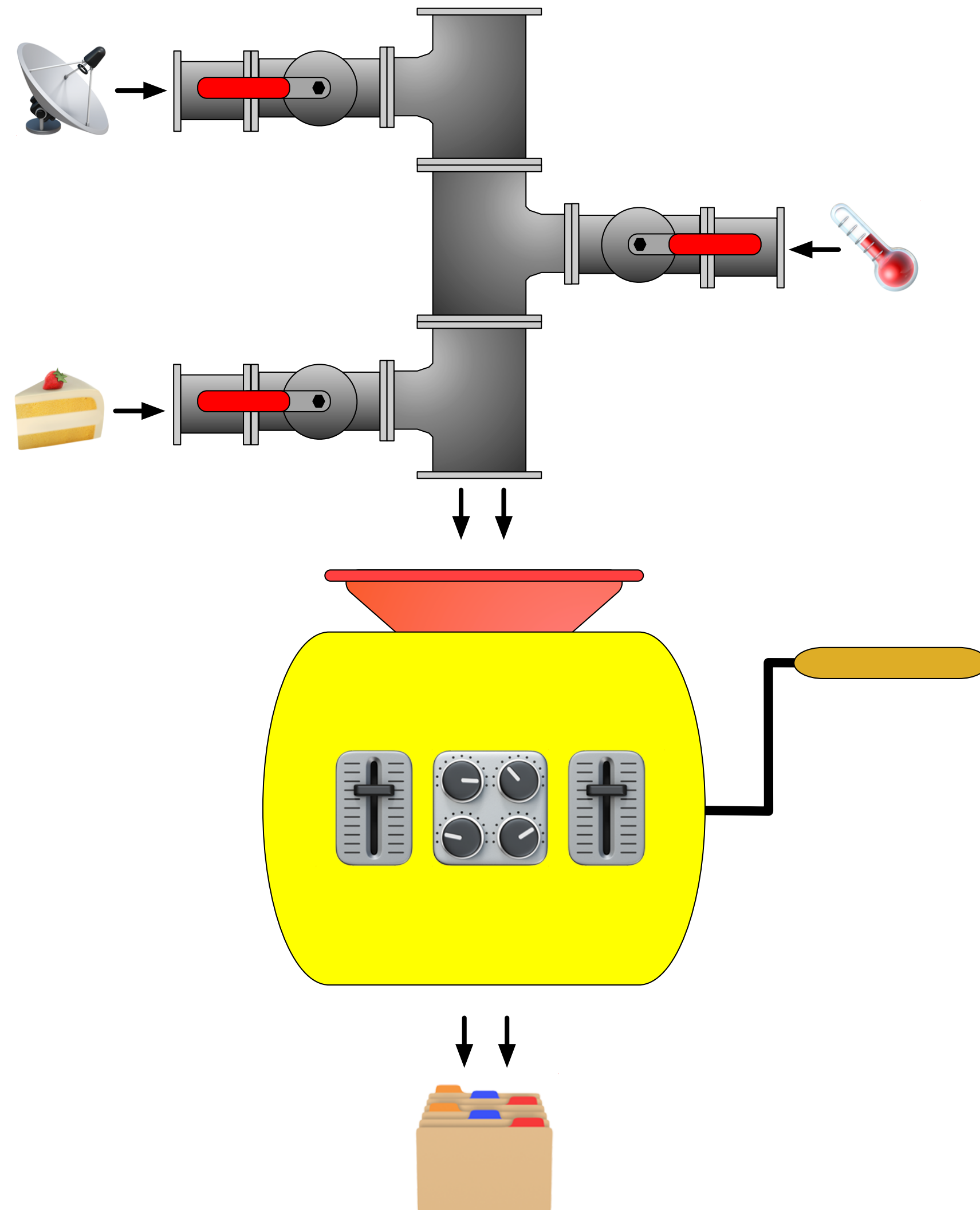
EUROPEAN  
SPALLATION  
SOURCE



# Managing High-Performance Data Flows and File Structures

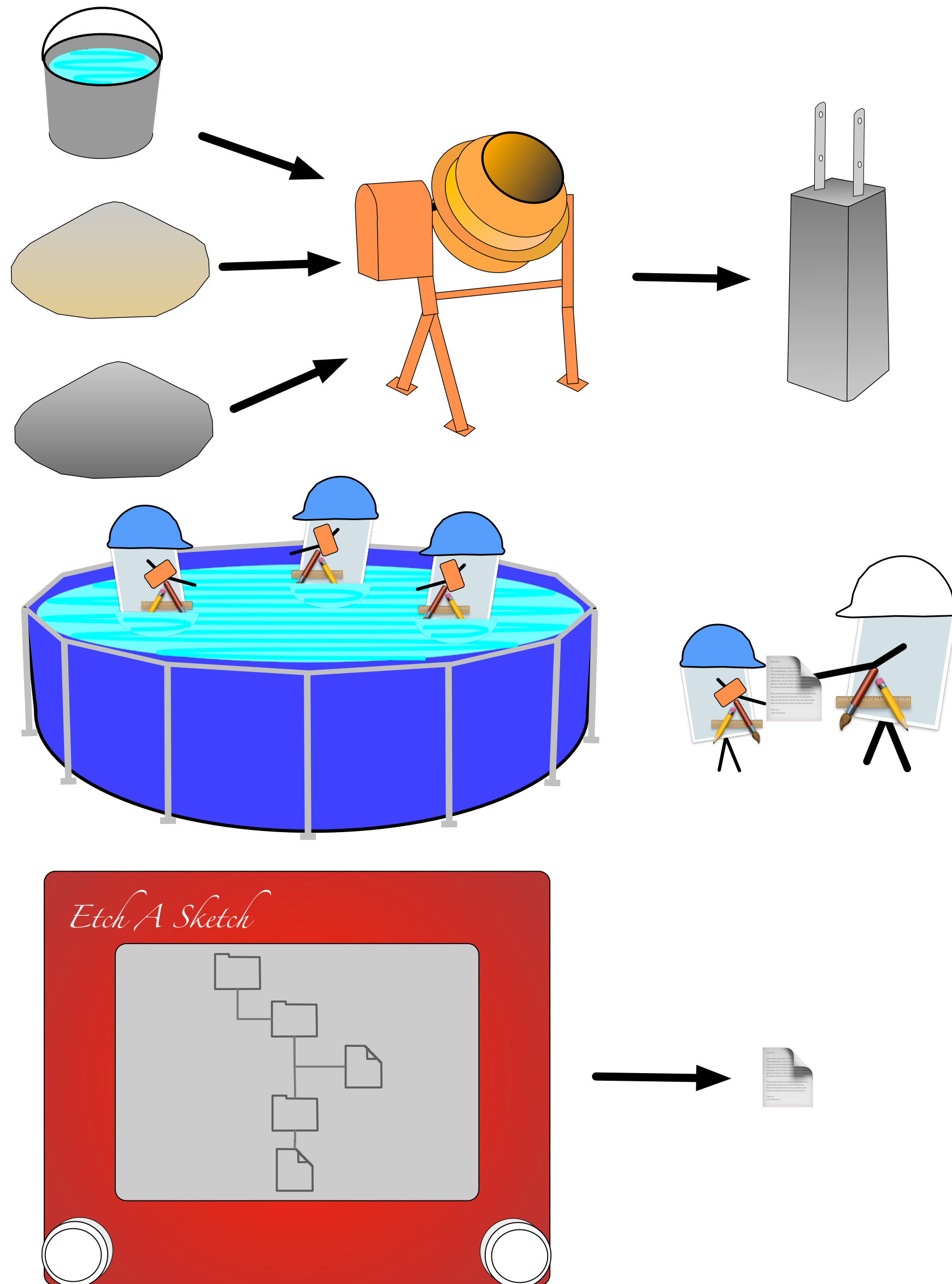
Jonas Måns Christer Nilsson, Tobias Stefan Richter (ESS),  
Jack Richard Harper (STFC), Matthew Jones (Tessella)

# Problem: Create a Data Storage Pipeline with These Requirements



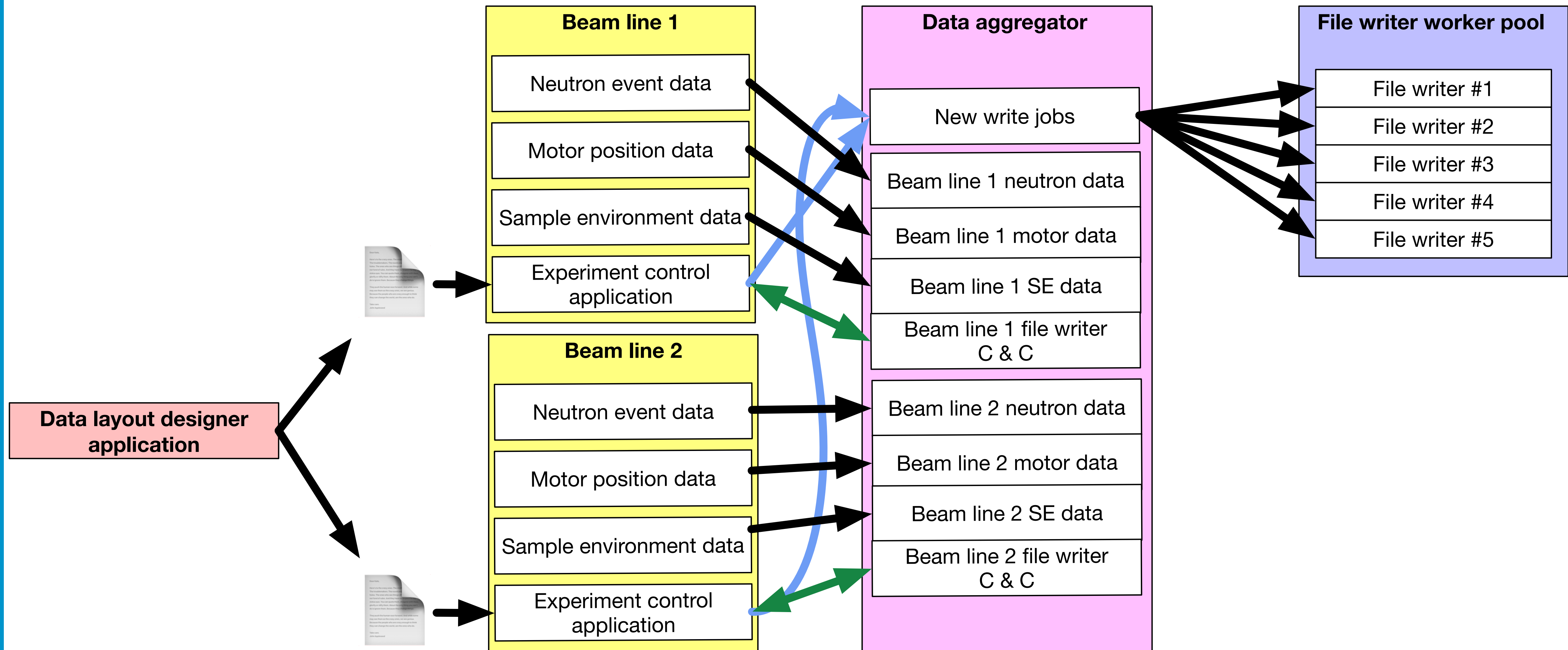
- Support many disparate data sources
- High performance
- Easily (re-)configurable

# Solution(s)



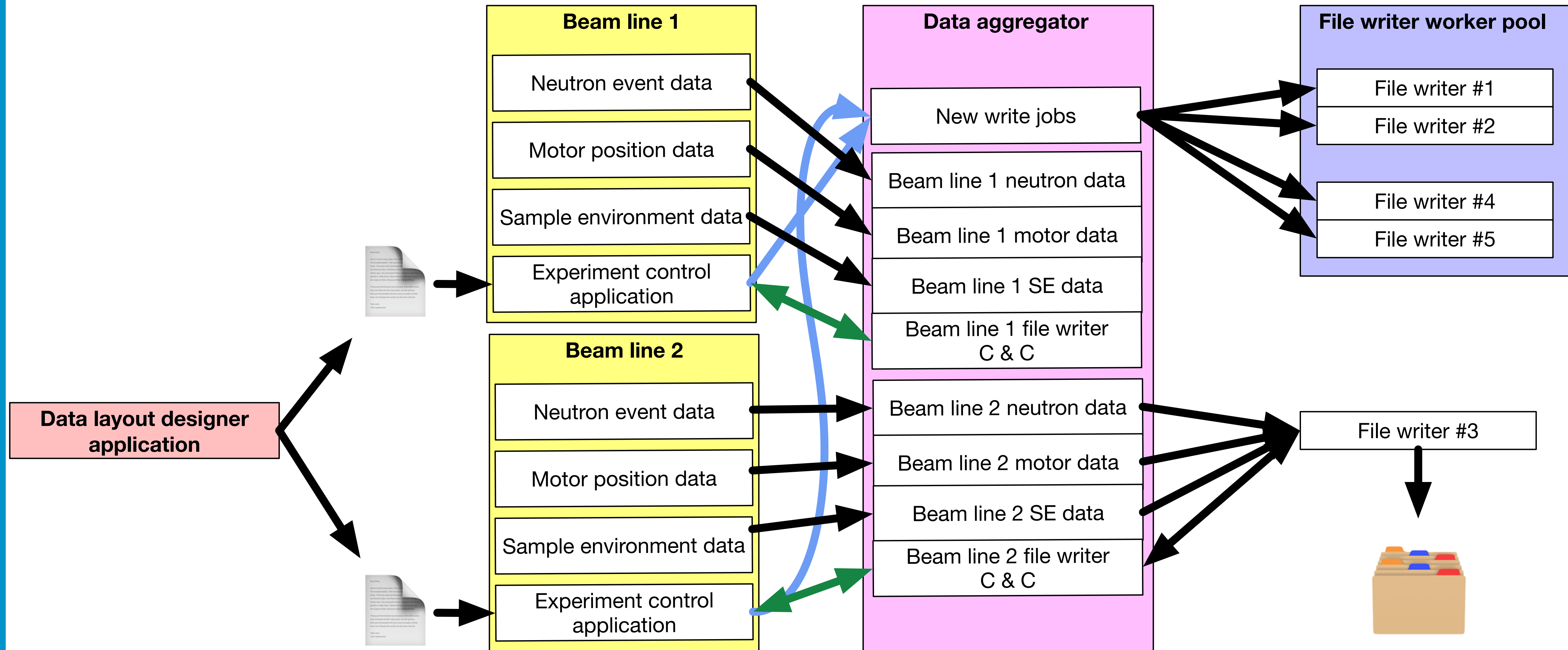
- Use a data aggregator
- Have a pool of file-writing workers
- Dynamically configure the file-writing worker for its task
- Make it easy to create or modify file-layout configurations

# The Pipeline in Detail

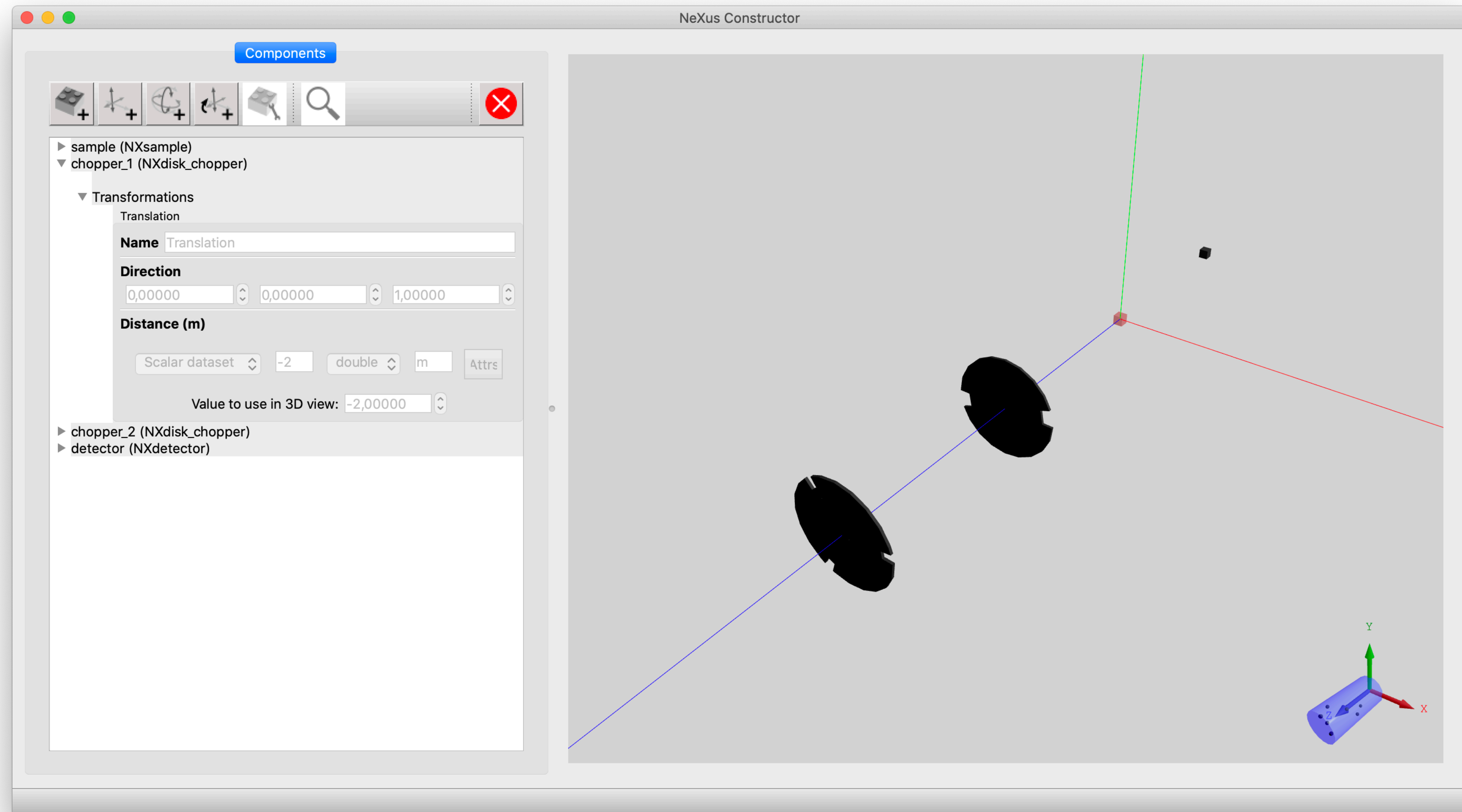




# The Pipeline in Detail



# User Interface



# User Interface



NICOS - guest at localhost:1301

Instrument: YMIR Experiment: Service mode Status: IDLE

Setup

Instrument interaction

Batch file generation

Detector Image

History

Logs

>> FileWriterControl.stop\_job()

+ New Cmd ▶ Dry run ▶ Run

Output Scan Plot Detector Image Script Status

[10:43:57] >>> [system 2021-10-18 10:43:57] setting up NICOS  
[10:43:57] loading setup 'system' (system setup)  
[10:43:57] importing module 'nicos.commands.standard'...  
[10:43:59] importing module 'nicos\_ess.commands.epics'...  
[10:43:59] loading setup 'startup' (NICOS startup setup)  
[10:43:59] Cache : now connected to localhost:14869  
[10:43:59] creating device 'Exp' (The current experiment)...  
[10:43:59] creating device 'Sample' (The currently used sample)...  
[10:44:01] creating device 'FileWriterControl' (Control for the file-writer)...  
[10:44:01] creating device 'NexusStructure' (Provides the NeXus structure)...  
[10:44:01] NexusStructure: nicos\_ess.ymir.nexus.nexus\_templates  
[10:44:01] NexusStructure: Finished importing nexus\_templates  
[10:44:01] creating device 'FileWriterStatus' (Status of the file-writer)...  
[10:44:01] creating device 'Space' (The amount of free space for storing data)...  
[10:44:01] creating device 'YMIR' (instrument object)...  
[10:44:01] creating device 'conssink'...  
[10:44:01] creating device 'daemonsink'...  
[10:44:01] creating device 'filesink'...  
[10:44:01] creating device 'liveview'...  
[10:44:01] setups loaded: startup  
[10:44:01] checking master status...  
[10:44:01] switched to master mode  
[10:44:13] >>> [guest 2021-10-18 10:44:13] FileWriterControl.start\_job()  
[10:44:14] FileWriterStatus: starting to write job 9205aae2-2fef-11ec-9122-8c85907b2108

Devices

Enter search expression

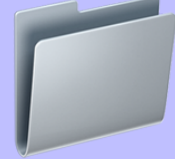
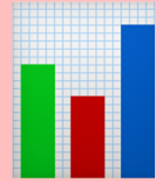



☐ show only the errors

Name	Value	Status
system		
Exp		
lastpoint	10	
lastscan	0	
FileWriterStatus		writing 49e8e328-2fef-11ec-b780-8c85
Sample		
Space	138.693...	138.69 GiB free
YMIR		



# File Layout and Source Configuration



	Directory	{ "name": "entry", "type": "group", "attributes": { "NX_class": "NXentry" }, "children": [ { "module": "f142", "config": { "source": "source_name_of_data", "topic": "channel_name_of_data" }}, { "module": "dataset", "config": { "name": "static_dataset", "values": [ 1, 2, 3, 4] } }, { "type": "link", "name": "some_link", "target": "./static_dataset" } ] }
	Meta data / attribute	
	Data stream	
	Static dataset	
	Link	

# Technology Stack



*NeXus*

- NeXuS constructor: <https://github.com/ess-dmsc/nexus-creator>
- NICOS: <https://nicos-controls.org>
- File-writer-control: <https://github.com/ess-dmsc/file-writer-control>
- Kafka: <https://kafka.apache.org>
- File-writer: <https://github.com/ess-dmsc/kafka-to-nexus>
- NeXuS standard: <https://www.nexusformat.org>
- HDF5 file format: <https://www.hdfgroup.org>