945358-BIGPICTURE
Central Repository for Digital Pathology

WP2 - Infrastructure and database hosting

D2.06 - Report on hardware scaling

Lead contributor	2 - LIU
	joel.hedlund@liu.se
Other contributors	5 - CSC 6 - UU 37 - JANSSEN 38 - BAYER

Document History

Version	Date	Description
V0.1	19 Dec 2023	First Draft
V0.2	22 Dec 2023	Comments
V1.0	13 Jan 2023	Second Draft
V1.1	25 Jan 2023	Final Version

Abstract

This report describes how the repository infrastructure has been able to meet Bigpicture hardware scaling needs to date, and makes estimates of our ability to meet projected future needs.

The hardware capacity for the Bigpicture platform consists mainly of storage capacity for 3 million WSI and associated metadata, which is expected to require 4,5 Petabytes effective storage space. This data will be replicated, for high availability and disaster resilience, between two separate data centers at CSC - IT Center for Science Ltd. in Finland, and at a private cloud provider in Sweden procured by Uppsala University through a framework agreement set up by the Swedish government in a procurement for best value-for-money in secure processing of sensitive personal data for research.

There are servers and storage for receiving the data and for preprocessing. Some services are used for the processing such as providing Cytomine servers as-a-service, and as a framework to enable processing of indirectly accessible datasets and AI models.

Some additional lightweight capacity is also needed for hosting of various services such as REMS portal, discovery portal, and a planned AI algorithm repository. The intention is to use cloud or container platforms for these.

Methods

During Bigpicture Task 2.5 – Pilot operations in Phase III - Refinements, two data collection milestones of relevance to repository infrastructure hardware scaling were scheduled. The first one, MS5 - Significant amount data received (approx. 75TB) was scheduled to M15, as a first verification of technical hardware scalability. This was also designed as a stepping stone in preparation for milestone MS8 - Delivered clinical dataset via the data submission portal to the BIGPICTURE repository, which was scheduled to M18. Both these milestones were met, without encountering any technical hardware scaling issues.

For MS5, 76.5TB of pilot data was received by the repository, including clinical and non-clinical WSI and metadata from Bigpicture beneficiaries from academia, caregivers and EFPIA companies. The pilot data included Bigpicture standards compliant image and metadata formats, as well as non-compliant data and mock-up data, all thoroughly anonymized or even randomly generated in order to avoid risk of privacy breach.

For MS8, an addition of three datasets containing clinical WSIs and associated Bigpicture standards compliant metadata were submitted to the repository by Bigpicture partners Region Östergötland, Medical University of Vienna, and University Medical Center Utrecht.

This data ingestion was achieved without any hardware scaling issues being encountered. However, we expect that the pace of data submission will increase strongly in the coming period, as the repository enters Task 2.8 – Large scale archive operations in Phase IV - Scaleup (M25 onward). We are therefore making the necessary preparations, weighing opportunities and risks associated with the different procurement strategies, and taking the present global chip shortage situation into account as far as possible.

Hardware procurements are normally planned so that capacity is available when needed, and so that hardware lifetime persists throughout the lifetime of whatever project that the hardware is intended to serve. Also, hardware is usually scaled to match the growth in needs. However, this is not always possible, for example for reasons in the intersection of technology, project lifetimes, and investment economy. Larger investments are typically more cost effective, however this always carries a risk that not all capacity in the end can be used.

The Bigpicture repository infrastructure is built on object storage, with a containerized service layer.

The storage capacity consumed the major part of the hardware costs. IT systems typically have a 4-5 year lifetime, after which hardware failures become an increasing risk, which carries an increasing mitigation cost. This would suggest delaying procurement so that no replacements of entire storage systems need be carried out within the lifetime of the project. However, storage system investment costs are not linear in the capacity range that is relevant to Bigpicture, but tend to be less effective in the "sub-Petabyte" range, than in the "several Petabytes" range, which would suggest procuring storage capacity only in larger amounts. In addition to that investment costs are due to be amortized during several years of the project. Therefore, CSC's investments must be done early enough.

Preprocessing/inference capacity costs are harder to assess, as these depend largely on GPU performance needs of software codes that are still unwritten, when used together with data that is still unsubmitted. This would suggest delaying procurement of the bulk of the preprocessing/inference capacity, until these parameters are better known. However, such parameters could be challenging to correctly estimate, without having at least some capacity in place for testing and piloting purposes, using real-world codes and data. At the same time, procurement at single-server scale is impractical for overall system cohesion and maintainability, which would suggest not procuring capacity to only support small-scale pilot tests. This requires planning and testing with AI algorithm developers in WP4.

Therefore, some of the preprocessing capacity would be procured at a time point where the requirements are better known. This requires planning and testing with AI algorithm developers (WP4). Investing in a single server for a cloud service is not feasible, and thus such preprocessing capacities should be made as part of a larger investment. Operational expenses will also be taken into account.

Operational costs are also a concern, for both storage- and preprocessing/inference systems. This would suggest delaying procurements of all systems as much as possible.

The current global chip shortage is a patent concern, which carries a risk that procured systems may not be delivered in time to adequately serve the needs for which they were procured. This would suggest procuring earlier than needed in order to mitigate potential delays, and/or larger than needed to leverage economy of scale in procurement.

Some additional lightweight capacity is also needed for hosting of various services such as the REMS portal, data discovery portal, and a planned AI algorithm repository. The intention is to use cloud or container platforms for these, and the expectation is that these capacities will have negligible impact on overall infrastructure scalability and costs.

Two different approaches to ensuring cost effective and reliable capacity provisioning are being pursued by the two repository service operating organizations CSC and UU. CSC plans to procure, own and operate hardware on-premise for use in Bigpicture, while UU on the other hand, procures capacity from a private cloud provider in Sweden through a framework agreement set up by the Swedish government in a procurement for best value-for-money in secure processing of sensitive personal data for research.

Results

CSC

CSC is operating object storage service in large scale as part of its HPC services in its data center. Thus, the economy of scale allows cost efficient investments with existing contracts and service operations. To enable scaling out of pilot operations and into large scale operations, CSC procured the first batch of systems for its storage service in 2022, amounting to 0.5PB of storage. Procurement of the remaining 4PB is scheduled for 2023, so that the capacity will be ready for production and available when needed by the data contributors. This investment schedule is also intended to allow optimal lifetime overlap with the project duration.

CSC also procured two servers in 2021 for use in data submission and management services.

Uppsala University
Uppsala University (UU) has set up virtual machines and object storage to host the Swedish side of Bigpicture services, which are now used for data storage, and for piloting the Cytomine server as-a-service together with WP4, and for piloting a software platform for indirect access to AI models and datasets in collaboration with grand-challenge.org.

The UU choice to acquire capacities from a Swedish nationally procured private cloud provider has afforded an amount of increased flexibility in early procurements, as the cloud provider is able to aggregate multiple small increments in needs across multiple customers into larger and more effective medium-size hardware investments. This opportunity to increase capacity in small increments has already been useful in early scaling of storage. It is also expected to be similarly useful in early scaling of preprocessing/inference capacities to support smaller-scale piloting and testing to determine performance requirements of novel or upcoming software codes when used together with incoming real-world data, in order to advise later large scale procurements in both data centers. A potential drawback of the choice could be that the availability of GPU types and system architectures will by necessity be limited to whatever is installed at the cloud provider at the time. We are in continuous dialogue with the cloud provider, and will proactively communicate estimated GPU/architecture needs as they are identified by AI developers in WP4.

Conclusion

Bigpicture project has made initial capacity and server procurement for piloting and is ready to be expanded in 2023. Image preprocessing workflows will be designed in 2023 and the server capacity for the planned activities will be implemented.

Finnish node has procured 0,5 PB for piloting and is going to expand to 4 PB as of 2023. Sweden is using a cloud service provider and has an advantage in early and incremental scaling.

The two different approaches are expected to offer a degree of flexibility in investment planning, and allow for smaller scale piloting and testing of preprocessing/inference capacities than would otherwise be advisable, which is expected to provide better data points for planning the larger investments to address late project needs, for example in taking the next steps toward morphological data mining.