User meeting about shared use computers (Mar. 24, 2020) Q&A

Date: March 24, 2020, from 2 pm to 4 pm
Place: 2F, Large meeting room, Welfare and Conference Building, WAKO
Video conference: Webex

1. Outline of Data Science Infrastructure

- Is it possible to use the command line API for Mirantis as well as the web interface?

  → It is possible.

- Regarding the "Operation schedule of shared use computers" on p. 2, is it correct to think that it has already been decided that some kind of system will be introduced as a successor to the supercomputer Hokusai BW? If undecided, when will the decision be made on whether or not the system will be implemented? Where will the results be announced?

  → The introduction of the successor system to BW has not yet been decided. Given the amount of time it will take for the process to take place, we will have to make a decision somewhere in FY2020. The final decision will be made at the IT Strategic Planning Committee, so you will be able to know whether to introduce the system or not and the outline of the system configuration in the meeting minutes.

- The data science infrastructure HOKUSAI SS does not support massively parallel computation, but the proposal by 15 interested persons on Aug. 2018 do not deny support for massively parallel computation, and I think it is important for RIKEN to support massively parallel computation. We would like to see support for massively parallel computation in BW and beyond, based on the past performance of massively parallel computation and the background of the proposal.

  → We will receive it as a valuable opinion.

2. Current plans for shared use computers and beneficiary pays after the next fiscal year

Q&A about the Data Science Infrastructure HOKUSAI SS

- Can global IP be accessed from the outside of RIKEN.

  → It is possible.

- Is it available a high level of security settings, such as denying access from outside, when handling human genome data?

  → By default, it cannot be accessed from the Internet. If you want to make it accessible from outside, you need to set it up separately. Think of it as the same as the current general VLANs and intra at RIKEN.

- Is it possible to rent multiple VMs in one lab?

→ It can be done in tenant, since resources (number of cores and amount of memory) are paid out to tenant units. The maximum amount of resources for each tenant will be discussed in the future.

- How can we use it efficiently to process large amounts of data?。

  →The file servers are all connected at 100GbE, and the CPU farm is also connected at 10GbE to the final node. It is unlikely that data will be blocked and will not flow to the compute nodes.

- Is SSD of CPU Farm available to users?

  →40GB of SSD area is allocated per core. A VM image is created on the SSD, and the remaining area can be used as a temporary file storage area.

- In the case of tenant use, is it possible to connect from outside via SSH?

  → By default, the data science infrastructure cannot be connected via ssh from outside RIKEN without a VPN. BW can be connected via ssh from outside RIKEN as before.

- How can I use this service as a replacement for the Riken Cloud CMS testing service?

  →You're going to get a VM up and running a web server. However, this system will be shut down once or twice a year due to planned power outages and maintenance. The SLA for this system will be shown in the future, so please decide whether you want to use it or not. In addition, the provisional CMS test service will run until the end of FY2020, and we are considering how to provide web functions, including CMS services, over the long term before that time.

Q&A about HOKUSAI BW

- Will you tell us how much other people charge and they are doing high priority jobs?

  → It is under consideration.

- In the usual year, about 120% of the total available resources were allocated at the first project review. Is the percentage the same?

  → At present, we are planning to allocate about 120%, which is the same as usual.

- Can you give us a more specific picture of how tape storage is used?

  → In the future, tapes will be used for long-term data that cannot be directly accessed by users. Cold media is necessary from the viewpoint of preserving research data as RIKEN. However, we are still considering what kind of data will be stored and how it will be used.

- Until now, the upper limit of available resources was 10% for one project, but will that limit continues?

  →It has not been decided yet, but the plan is to basically eliminate it. There is a major direction to widen the frontage of users and lower the threshold.

Q&A about usage fees

- Is that an annual fee of about 100 million yen?

  → It is an annual limit.

- Is the unit of purchase for BW's priority use fee per month?

  → It is purchased on a core hour basis.

- In FY2020, there will be easing measures for users in the past, but what will be the operation policy and fee structure for BW operation after 2021? Is the 15% of the total contract only for the duration of the easing measures, and then the user fee increases after that?

  → The operation of BW in FY2021 is basically the same as the latter half of FY2020, and a usage fee will be introduced for priority execution without project review. Also, the 15% of usage fees is continuous, not easing measure. At present, subsidies for young researchers and budding issues and strategic issues are left blank.

- Is there a way to use the resources purchased at the end of the fiscal year for the following year?

  → We have examined various ways, but it is basically difficult to exceed the fiscal year both in the case of institutional funding and in the case of external funds. We have consulted with the external funds office, but there are different difficulties at each external fund. We want to continue look into this.

- Will this usage fee list continue for six years?

  → Unless the management policy of RIKEN changes, we basically expect this fee list to continue.