

## Towards the Future Supercomputing Services at the Cybermedia Center

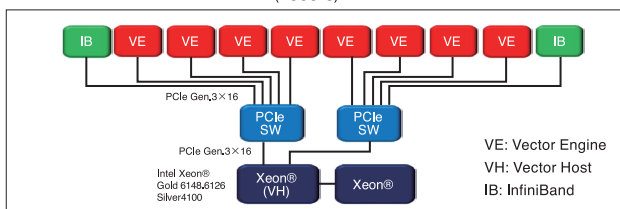
### Tuning on SX-Aurora TSUBASA

The current flagship supercomputing system at the Cybermedia Center is NEC SX-ACE system. Now we are in the process of investigating SX-Aurora TSUBASA as a candidate processor for the next supercomputing system.

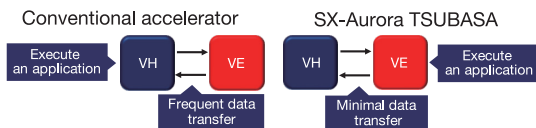
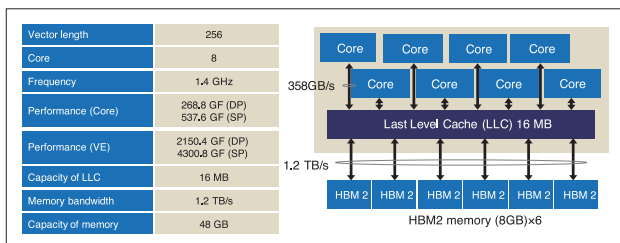
\* SX-Aurora TSUBASA

- a vector processor from NEC.
- deal with 256 elements with 1 directive.

Architecture of SX-Aurora TSUBASA (A300-8)



Architecture of VE



### Lesson learnt: Things to keep in mind to improve the performance of applications on SX-Aurora TSUBASA

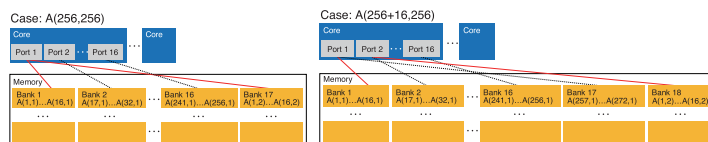
The following three considerations affect on application performance.

- 256 or multiples of 256 iterations: SX-Aurora TSUBASA can deal with 256 elements.
- sufficient margins: avoiding of cpu port conflict increases the performance for non sequential access.
- sequential access: avoiding of non sequential access increases the performance.

Example code

```
REAL, DIMENSION (256+16, 256) :: A, B
DO j=1,256
DO i=1,256
A(j, i) = B(i, j)
```

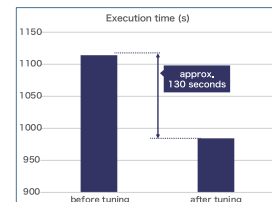
If array A doesn't have sufficient margins in the example code, load/store requests are concentrated on a specific port.



### An application example

Radiation Fluid simulation code running on SX-Aurora TSUBASA was accelerated as follows.

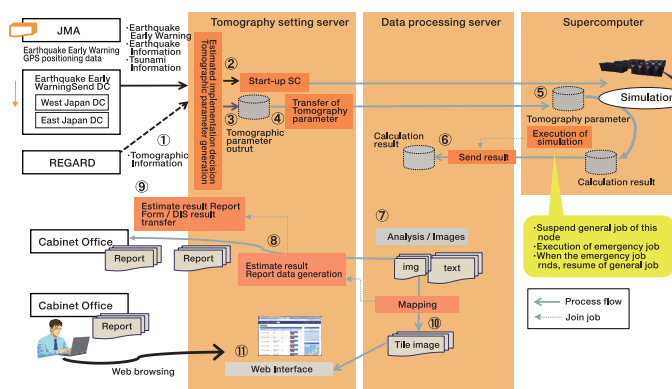
	Execution time (s)
before tuning	1114.085353
after tuning	984.334410



### Reproduction of a system that executes emergency jobs

Tsunami inundation damage estimation system is now operated using two SX-ACE systems at Tohoku University and Osaka University. The system completes Tsunami inundation damage estimation in 30 minutes after a large-scale earthquake that may trigger Tsunami happens. The current coverage of the system is from the Izu Peninsula to Osumi Peninsula.

### Tsunami inundation damage estimation system



- OS: SUPER-UX (UNIX SystemV + NEC Extension)
- RHEL7 (Red Hat Enterprise Linux7)
- Batch system: NQSII

### Experiment

We conducted an experiment to investigate whether Slurm, OSS scheduler, can be used as an alternative of NQSII or not. In the experiment the following time was measured when we switched a running job to an emergency job.

- A: Switching time from a low priority job to an emergency job
- B: Switching time from the emergency job to the low priority job



### Result

suspend/resume	
A	0.4 s
B	0.5 s

requeue	
A	0.4 s
B	100 s

The functionality with which the system stops running jobs, runs this emergency job and then recovers the previous running jobs to the original status heavily depends on NEC NQSII/JM, the proprietary job scheduler proprietary from NEC.

A. Musa, et al, "Real-time Tsunami Inundation Forecast System for Tsunami Disaster Prevention and Mitigation." The Journal of Supercomputing, vol. 74, pp.3093-3113, 2018. [DOI:10.1007/s11227-018-2363-0]

### Investigation of Slurm

Currently, Slurm supports the following three modes for suspending, executing and recovering of jobs. We are investigating whether the upper two can be used as an alternative of NQSII.

- Suspend low priority job, Execute emergency job, Resume low priority job
- Cancel low priority job, Execute emergency job, Requeue low priority job
- Stop at the checkpoint of the low priority job, Execute the emergency job, Restart from the checkpoint of the low priority job