Usage environment of Azure CycleCloud and benchmark test results on virtual machines

Toru Nagai [1)], Shuichi Gojuki [2)], Masatoshi Kawai [1)], Takahiro Katagiri [1)], Tetsuya Hoshino [1)]
1) Information Technology Center, Nagoya Univ.
2) Microsoft Japan Co., Ltd.

Abstract We conducted a performance measurement of virtual machines targeting Microsoft Azure with the main purpose of investigating the usage environment of public clouds. Specifically, we used Azure CycleCloud, which is specialized for HPC usage environment and executed various benchmark programs on virtual machines. We will report on the usage environment of Azure CycleCloud and the benchmark test results on virtual machines and discuss the collaboration between supercomputer systems and public clouds.

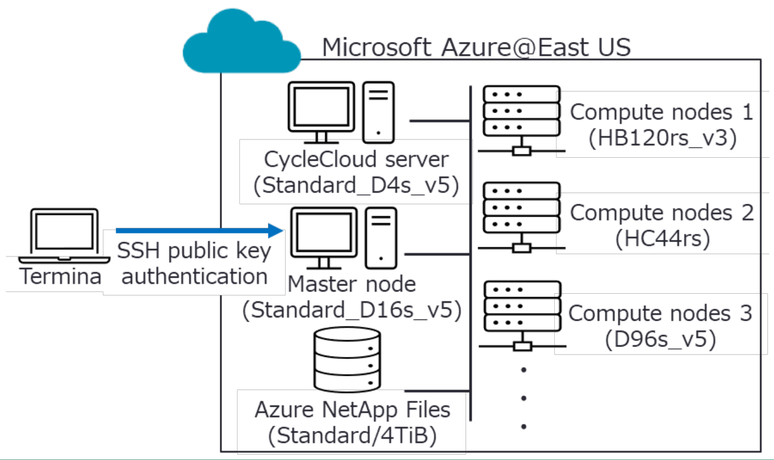
Azure CycleCloud

- building HPC environment
  - ‘VM size‘ refers to properties or resources assigned to a virtual machine such as CPU, main memory, GPU, and network bandwidth.

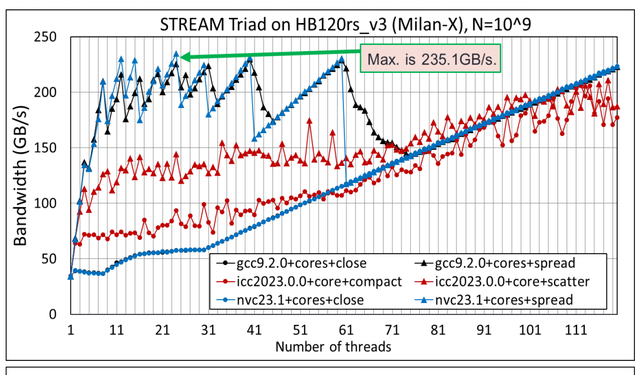
VM size	CPU	# of sockets	# of cores	# of HW threads	L3 cache size	RAM (GiB)	InfiniBand
HB120-16rs_v3	AMD EPYC 7V73X (Milan-X)	2	16	16	768MB x 2 sockets	448	200Gbps HDR
HB120rs_v3	AMD EPYC 7V73X (Milan-X)	2	120	120	768MB x 2 sockets	448	200Gbps HDR
HC44rs	Intel Xeon Platinum 8168 (Sky Lake)	2	44	44	33MB x 2 sockets	352	100Gbps EDR
D96s_v5	Intel Xeon Platinum 8370C (Ice Lake)	2	48	96	48MB x 2 sockets	384	—

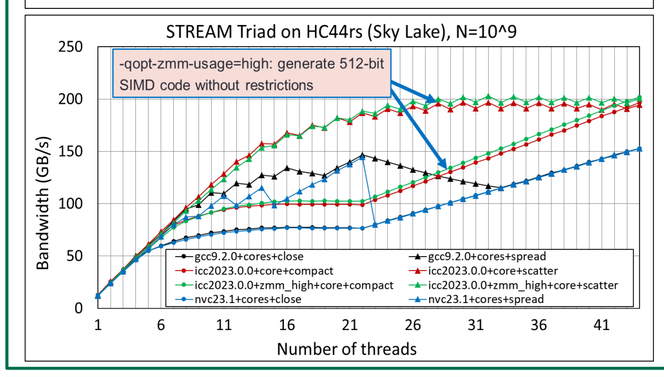
  - Based on the VM size, it orchestrates VMs and storage and deploys a job scheduler-configured cluster to an existing network.
  - It also installs specified software at the environment construction.
- managing the system
  - It provides functions of user management, resource management, job management and accounting management for the HPC environment.

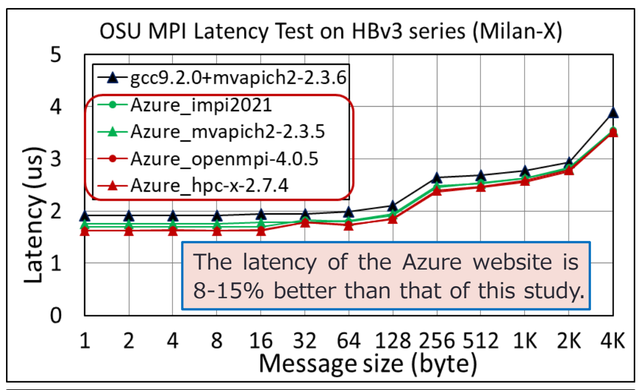
- Flow of job execution
  1. The user connects to the master node from a local terminal via SSH public key authentication.
  2. The user creates a job script on the master node and submits it to the scheduler.
  3. It automatically deploys VMs with the requested VM size collaborating with the scheduler.
  4. The job is executed.
  5. When the job is finished, the VMs are automatically stopped.

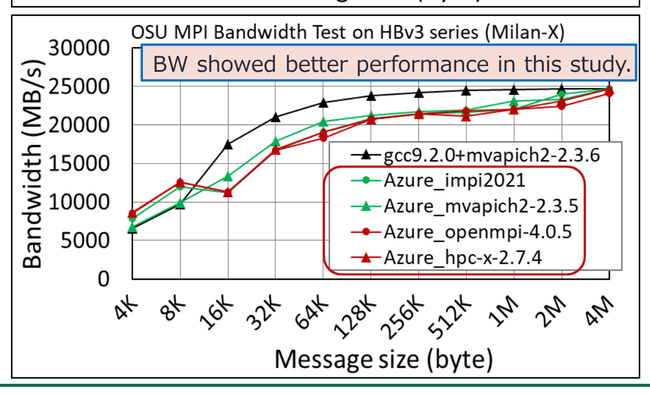


Benchmark test results





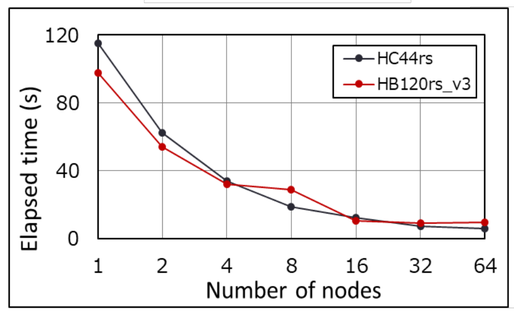




BICCG (Krylov subspace method with a block incomplete Cholesky preconditioner) solver as a practical application

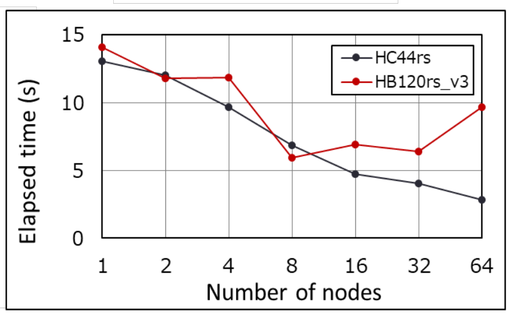
- Good performance in scaling for large problems

Poisson, DoF: 33M



The application is latency-sensitive due to small message sizes. Based on the results above, it is speculated that HC44rs has a latency advantage over HB120rs_v3.

G3_circuit, DoF: 1.5M



The variability in HB120rs_v3 in the figure above may have been significantly influenced by the network distance between the nodes deployed.

Collaboration between supercomputer centers & public clouds

Case 1: Cloud burst, or users can utilize the system without being aware of whether their submitted jobs are being executed in an on-premises environment or in a public cloud.
Case 2: Users actively utilize the public cloud.

#	Techinical requirements	Case 1	Case 2
1	Executables and shell scripts created on premises can be used as is in the public cloud.	OK	OK
2	Input files are automatically transferred to the public cloud.	OK	OK
3	Outputs can be obtained on premise.	OK	OK
4	Results are consistent between the two.	OK	OK
5	Security equivalent to or higher than on premise	OK	OK
6	Accounting information is available.	OK	OK
7	The execution time is approximately the same or less.	NG	OK

- #1-3: It is OK, only if the aligned software execution environments on both the on-premise and public cloud, and each other’s file systems mutually mounted over the network are available.
- #7: Optimization is necessary.
- The user manual is essential.

Pricing

Converting one dollar as 145 yen

VM size	CPU	# of HW threads	NVIDIA GPU	Inter-connect between nodes	Fee (yen/hour)			
					Pay as you go		Spot	
					East US	East Japan	East US	East Japan
HB120rs_v3	AMD EPYC 7V73X (Milan-X)	120	–	200Gbps HDR	522.0	756.9	52.2	75.7
HC44rs	Intel Xeon Platinum 8168 (Sky Lake)	44	–	100Gbps EDR	459.4	666.1	45.9	66.6
NC96ads_A100_v4	AMD EPYC 7V12 (Milan)	96	A100 x 4	–	2130.3	3088.9	672.3	1202.8
NC24s_v3	Intel Xeon E5-2690 v4 (Broadwell)	24	V100 x 4	–	1774.8	2432.5	826.7	852.6
D96s_v5	Intel Xeon Platinum 8370C (Ice Lake)	96	–	–	668.2	863.0	109.6	87.2
Flow Type I	Fujitsu A64fx	48	–	Tofu D	31.0			
Flow Type II	Intel Xeon Gold 6230 (Cascade Lake)	40	V100 x 4	200Gbps EDR	38.8			
Flow Cloud	Intel Xeon Gold 6230 (Cascade Lake)	80	–	100Gbps EDR	44.3			

- While the usability is not good, it seems possible to use the spot (excluding GPU usage).
- There is also room to consider a plan that allows us to use VMs cheaper than pay-as-you-go by committing to use for a certain period and contracting.