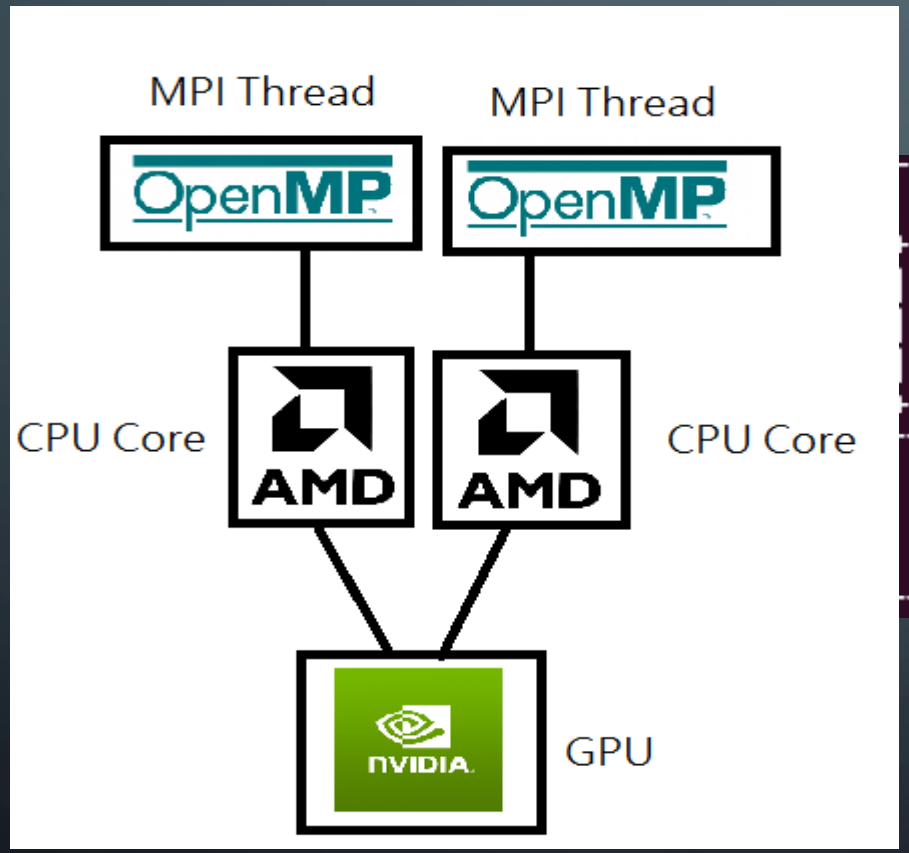
LAMMPS @ SC22

NATIONAL TSING HUA UNIVERSITY @ SC22

OUTLINE

- Running multiple CPU processes per 1 GPU
- Using Intel package on AMD nodes
- Comparison between NGC and Compile from source
- Hyperthreading of AMD vs Intel
- Scalability on Oracle Cloud Infrastructure
- Comparison between GPU acceleration package and KOKKOS package

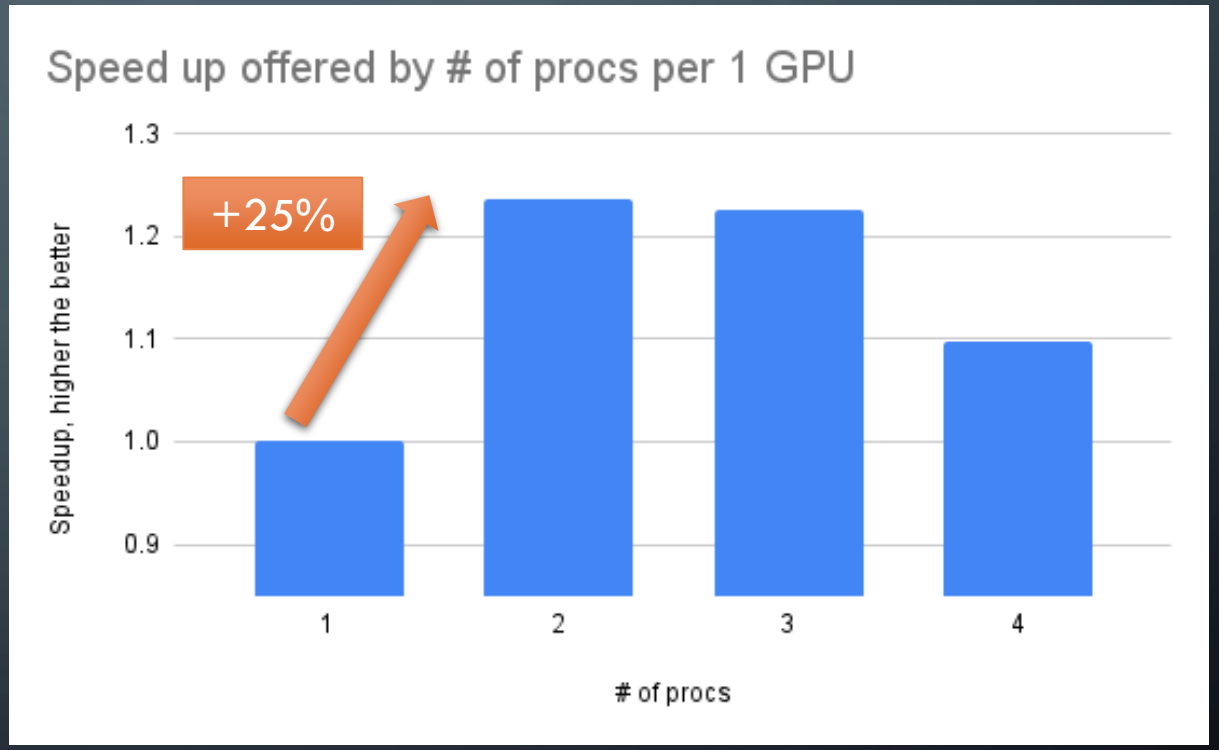
GPU UNDER UTILIZATION & PROCS PER GPU



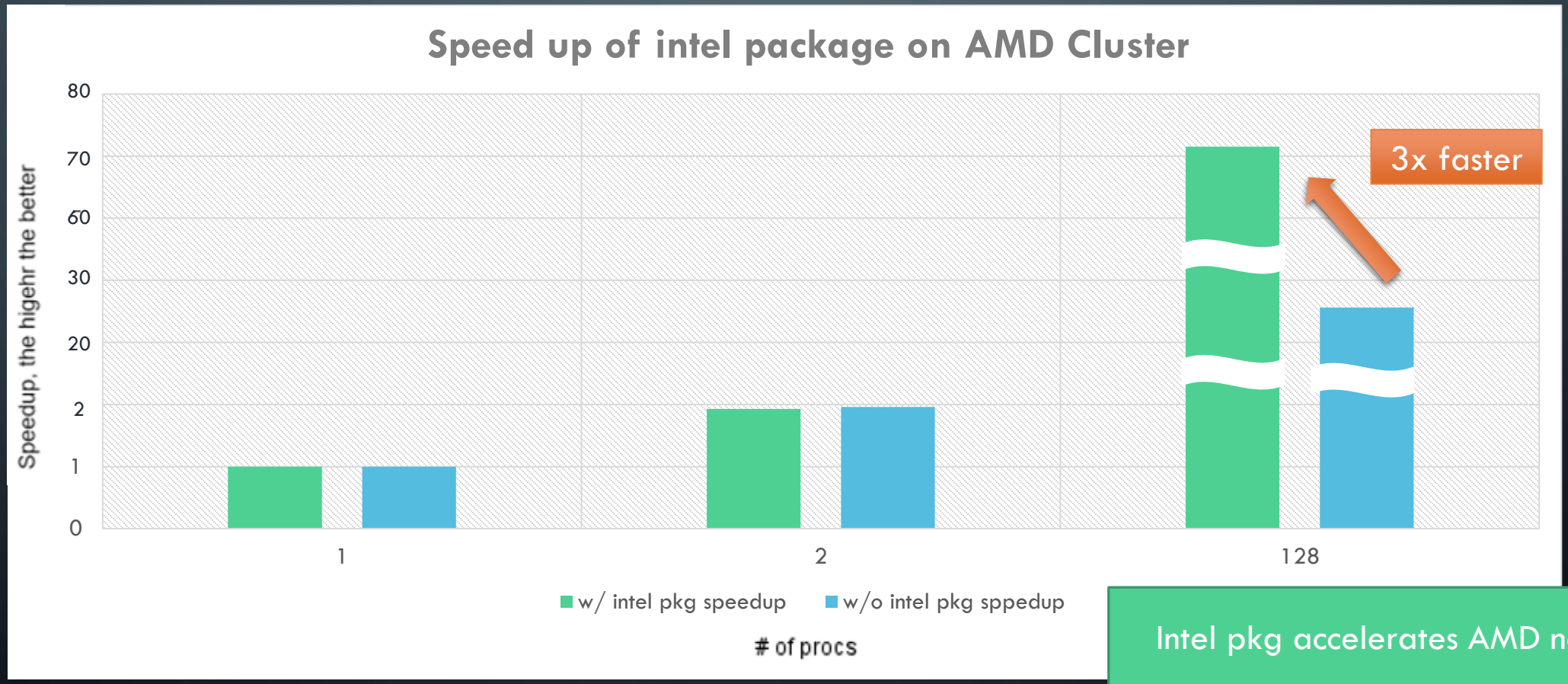

```
+-----------------------------------------------------------------------------+
| Version: 450.119.04   CUDA Version: 11.3                                    |
|-------------------------------+----------------------+----------------------+
|  Bus-Id            Disp.A     | Volatile Uncorr. ECC |
|          Memory-Usage         | GPU-Util  Compute M. |
|                               |                  MIG M. |
|===============================+======================|
|-------------------------------+----------------------+
| 00000000:3E:00.0 Off          |                    0 |
|    482MiB / 32510MiB          |     73%      Default |
|                               |                  N/A |
+-------------------------------+----------------------+
```

INTERPRETATION & STRATEGY

- Use 2 procs per GPU



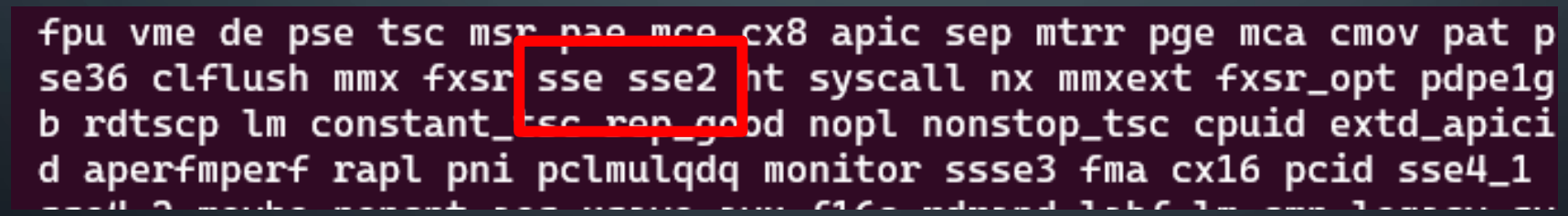
INTEL PACKAGE @ AMD CLUSTER



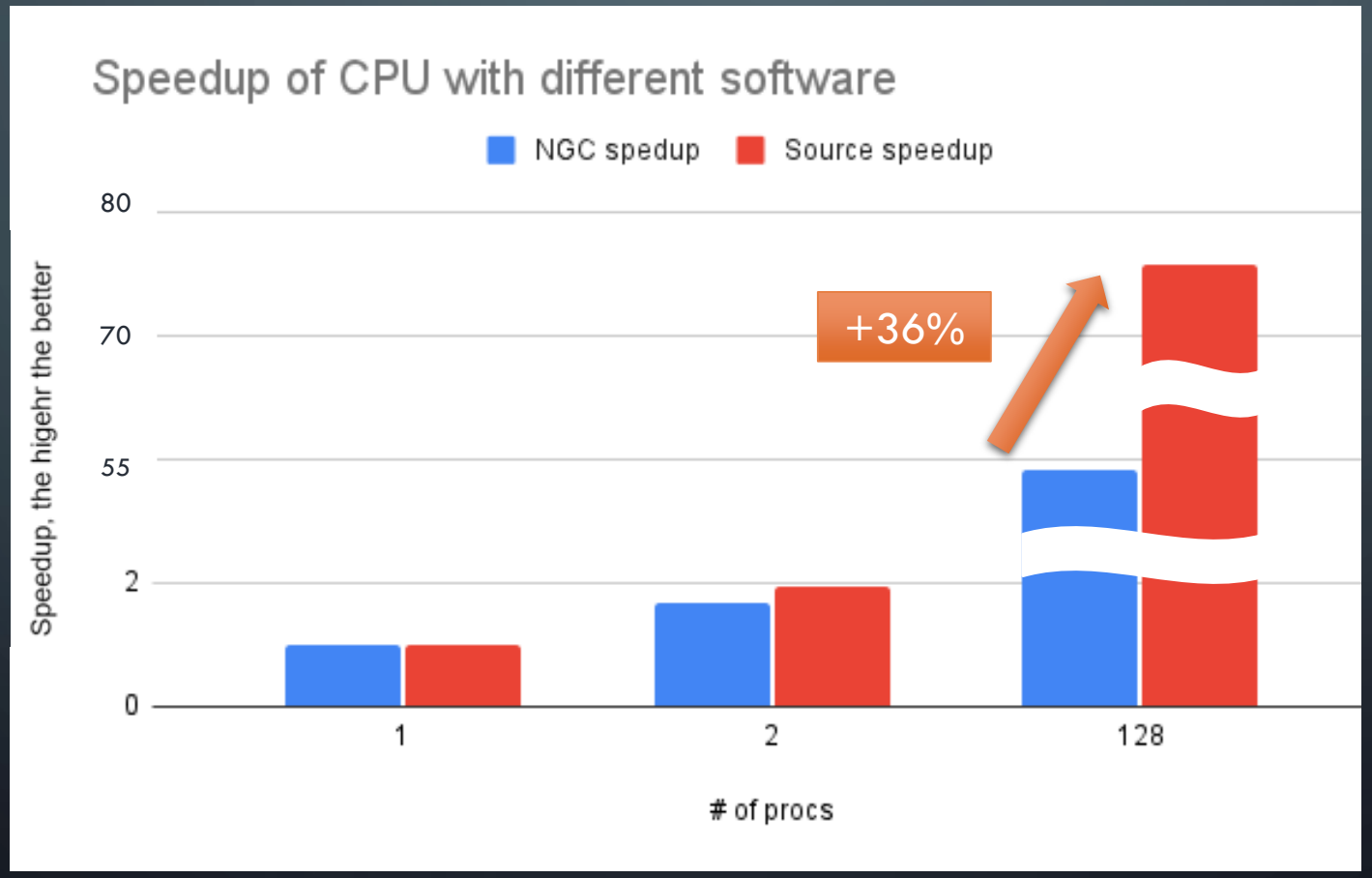
Intel pkg accelerates AMD nodes !!!

INTERPRETATION & STRATEGY

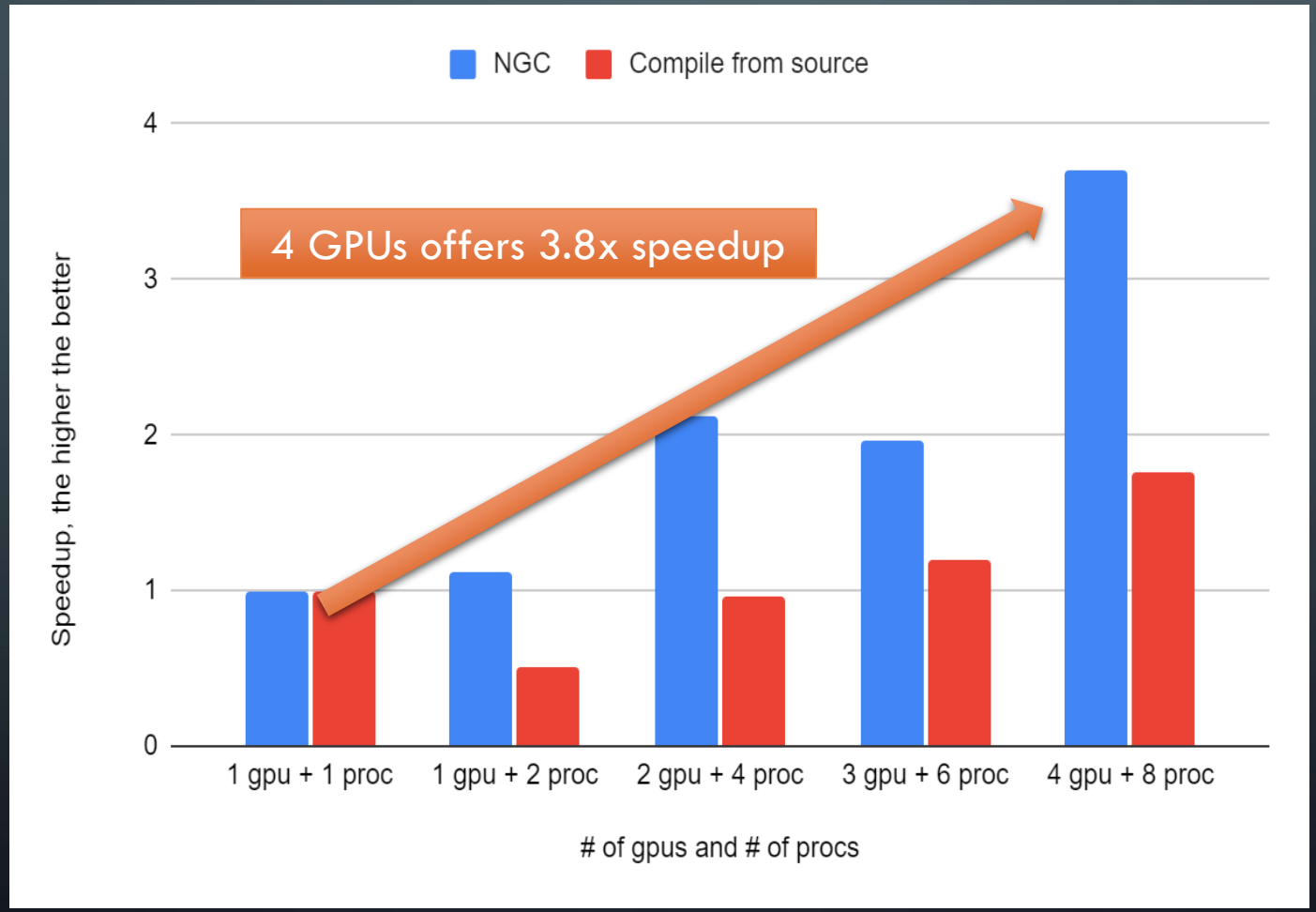
- Uses vectorization of Intel (AVX, SSE) to speedup
- AMD also supports vectorization (SSE) for speedups
- Use Intel Package to speedup our AMD cluster



NGC & COMPILE FROM SOURCE @ CPU



NGC & COMPILE FROM SOURCE @ GPU



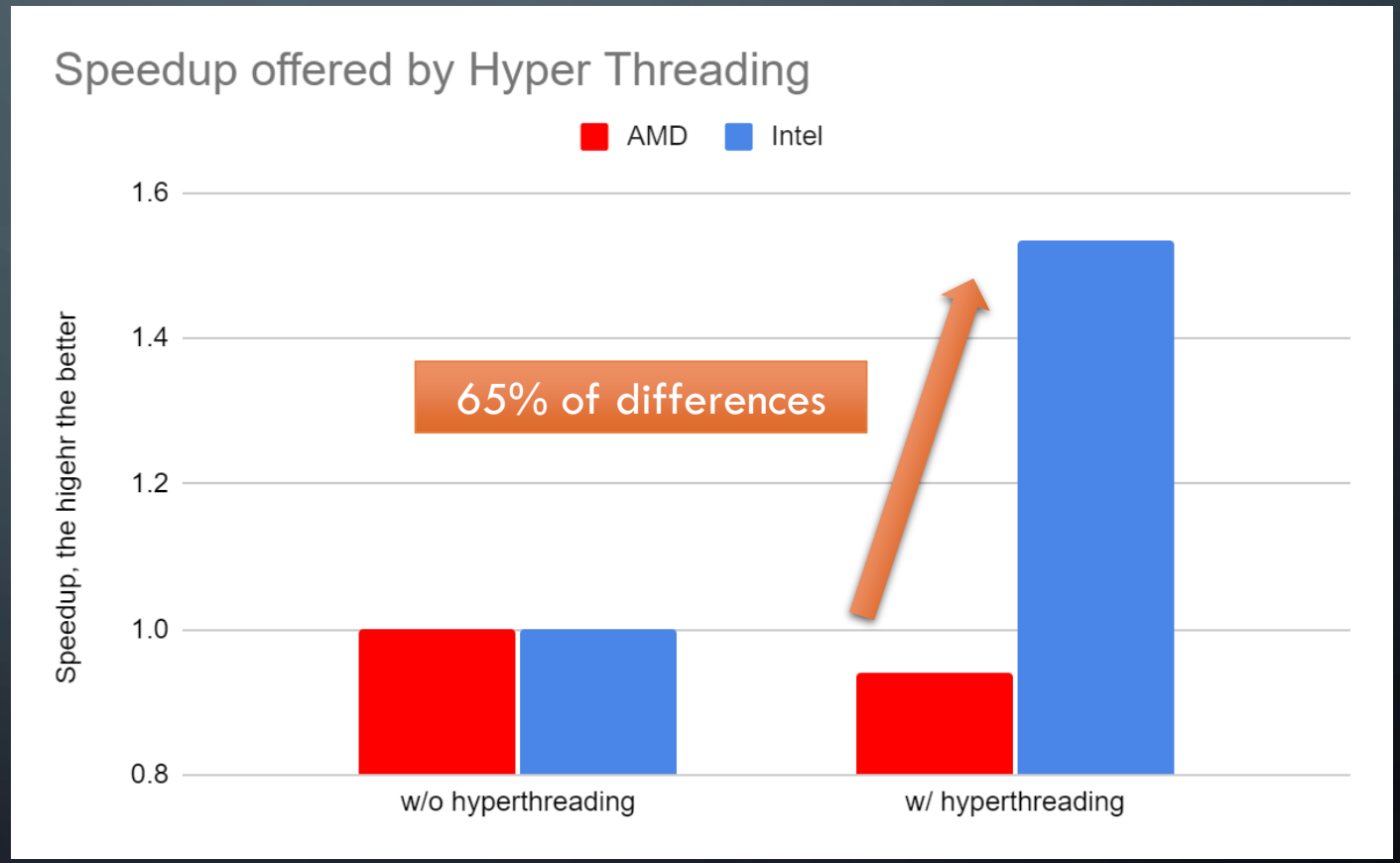
STRATEGY & INTERPRETATION

Problem size	Hardware	Software
Small problem	1 GPU	Compile from source
Medium problem	2 GPUs w/ NvLink	NGC
Big problem	4 GPUs w/ NvLink	NGC
All problems	CPU	Compile from source

- NGC incurs overhead due to virtualization
- Compile from source always works better on CPU

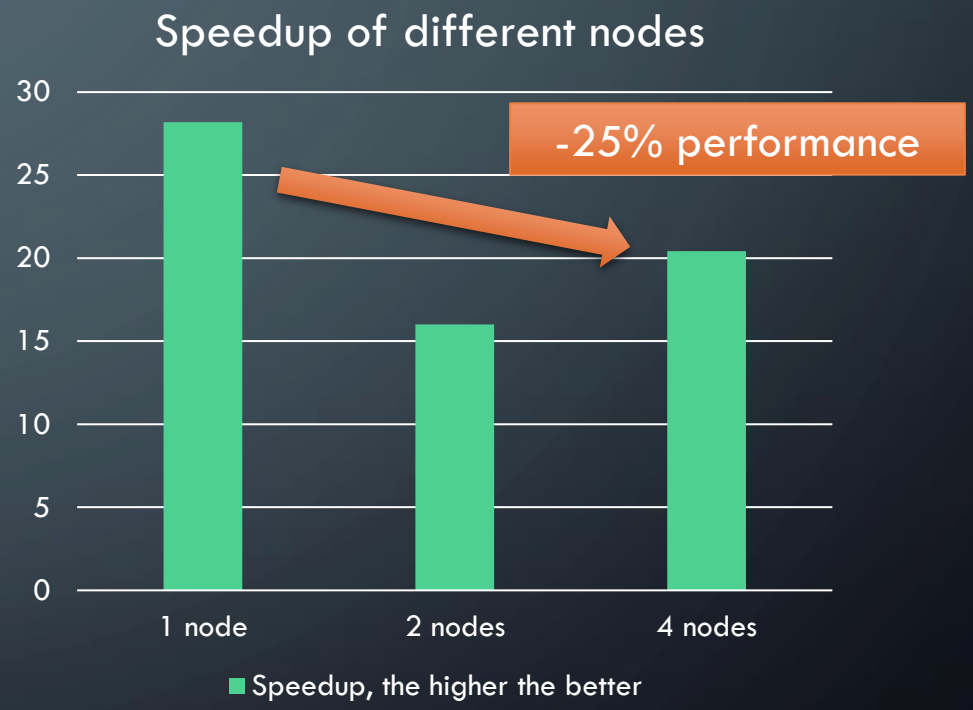
HYPERTHREADING @ AMD, INTEL

- Use HT on Intel only
- Intel HT > AMD HT

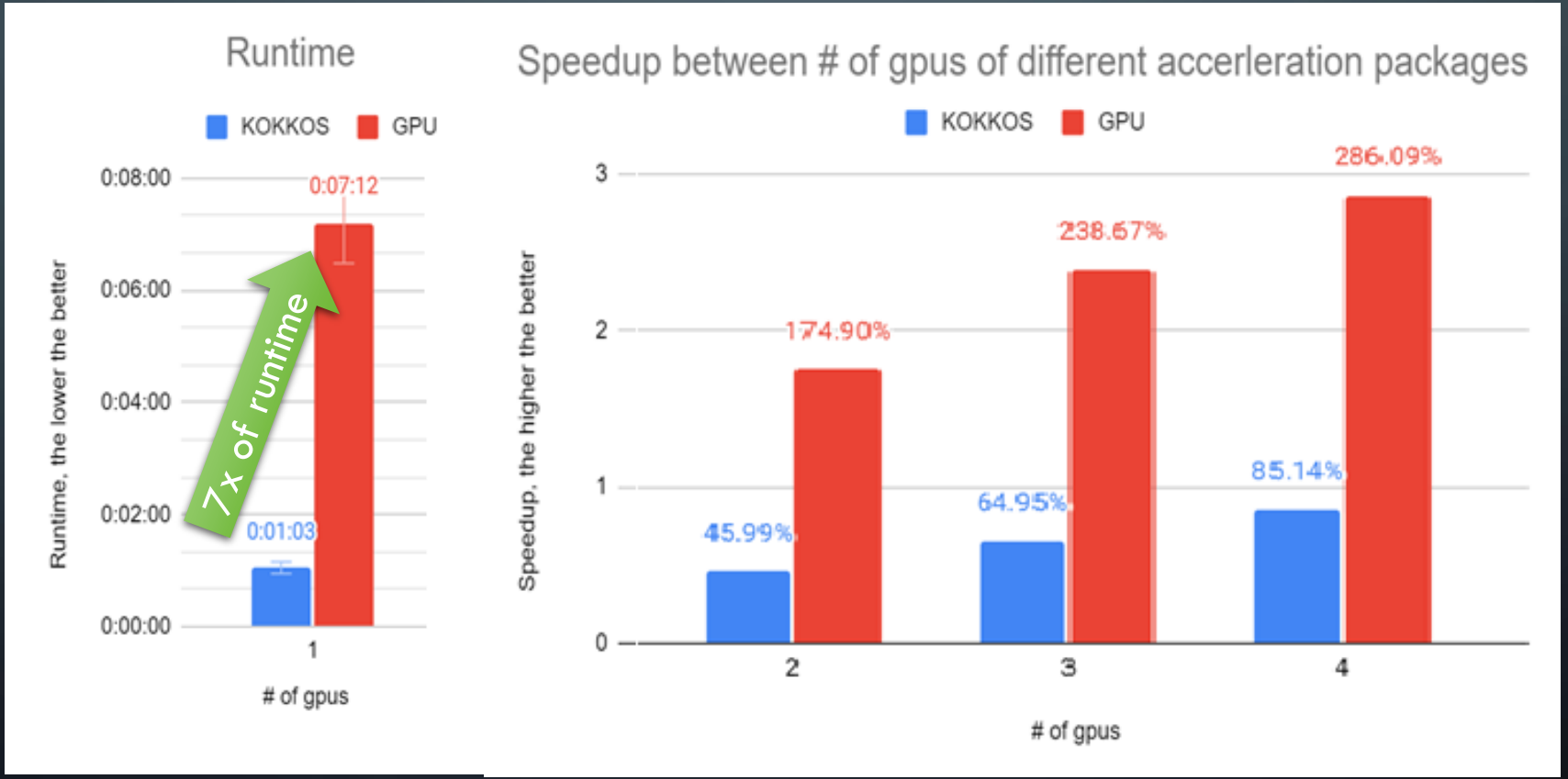


OCI SCALABILITY

- OCI nodes are connected by ethernet
- Run intranode task for OCI only
- Communication overhead is significant



KOKKOS AND GPU



INTERPRETATION & STRATEGY

	GPU	KOKKOS
Purpose	General purpose	Specific purpose
Hardware	NV, Intel, AMD, OpenGL	Nvidia, Intel Phi
Scalability	✓	
Absolute speed		✓
Special hardware support		✓

- Personal clusters = Nvidia GPU
- Use KOKKOS

LENNARD JONES & EAM & ML-SNAP

Benchmarks	GPU node	CPU node	Potential model
Lennard Jones	✓	✓	Mathematics
EAM	✓	✓	Mathematics
ML-Snap	✓		Machine Learning

INTERPRETATION & STRATEGY

- Machine learning requires tensor operations
  - Better for GPU nodes
  - GPU provides numerous CUDA cores for tensor operations
- Don’t put ML-Snap tasks on CPU nodes