



中国科学技术大学  
University of Science and Technology of China

# Relion PWDFT pyTorch VASP HiResX PyFR GPU平台测试

测试者：张文帅，范逸，叶创超，王亮，李杰岚

通讯邮箱 [wszhang\(at\)ustc.edu.cn](mailto:wzhang@ustc.edu.cn)



测试平台

硬件测试

应用测试

用户建议

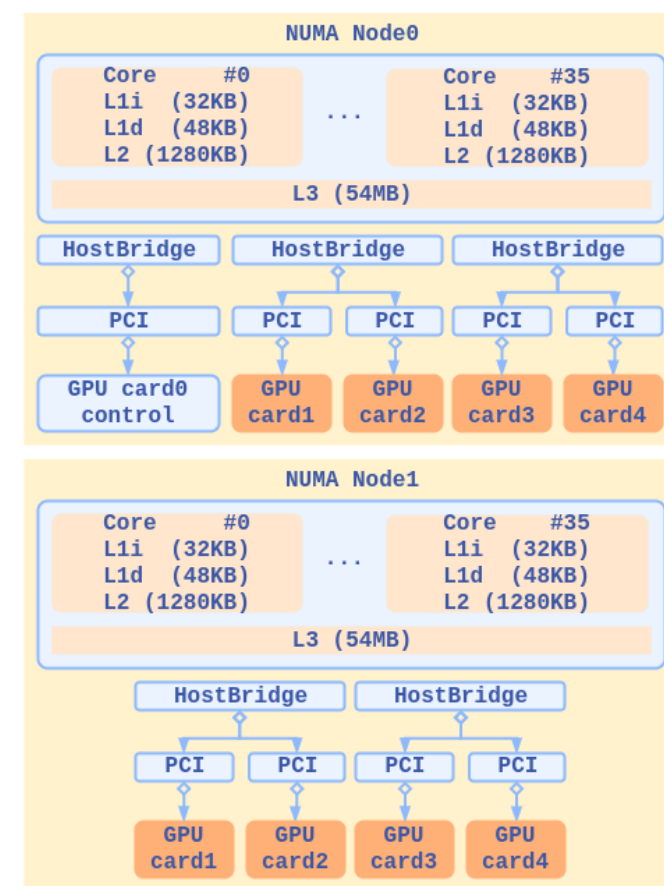
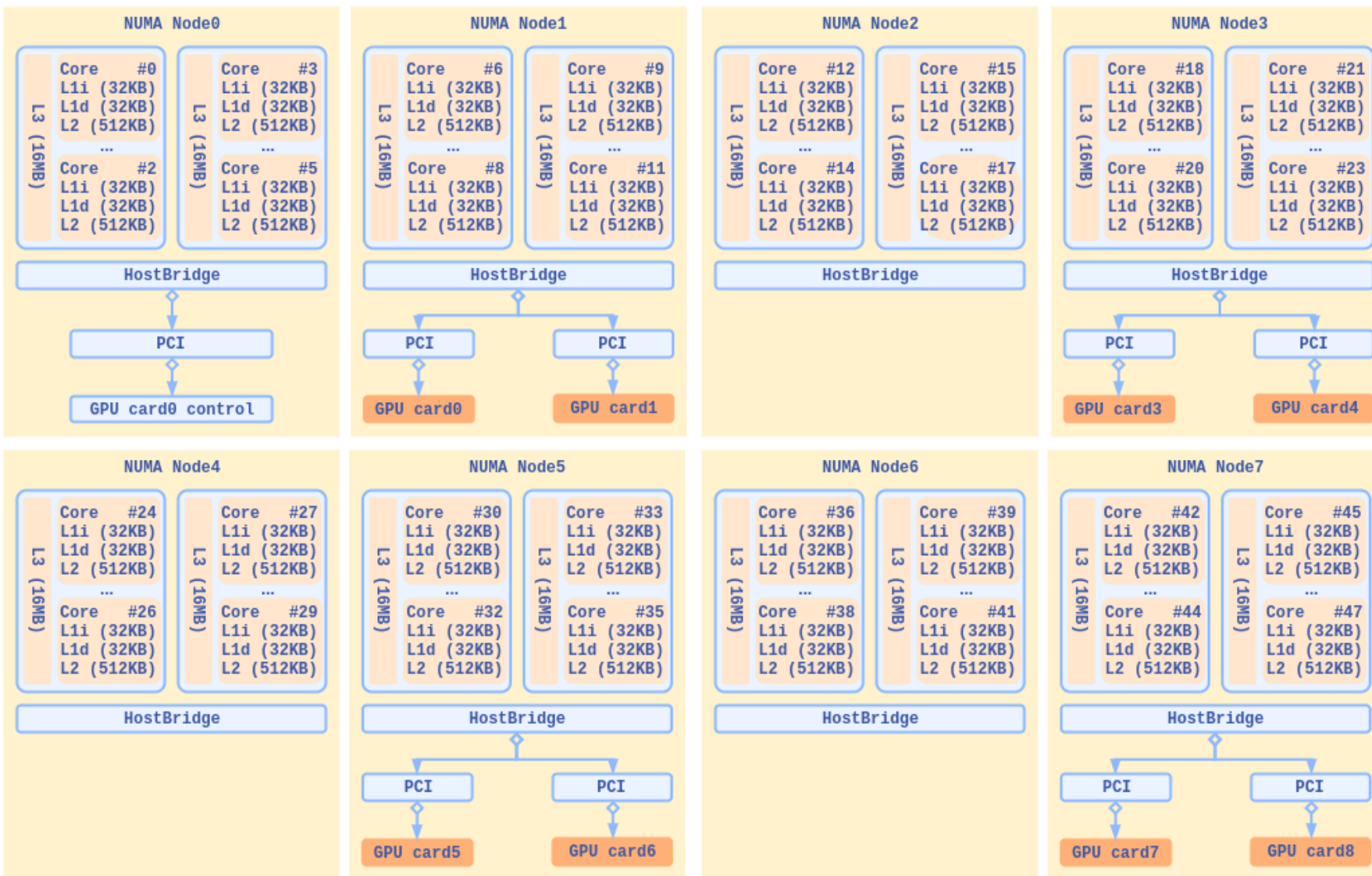
总结



配置	平台1 : A235P Host	平台2 : A280N Host	平台3: A235N Host	平台4 : I220N Host
CPU	2颗 AMD 7452 32C/64T 2.35GHz 155W	2颗 AMD 7402 EPYC 24C 2.8GHz 180W	2颗 AMD 7452 32C/64T 2.35GHz 155W	2颗 Intel 36Core 2.2GHz 270W
CPU总核数	64	48	64	72
GPU	4 颗 A100 40 GB	8 颗 HGX A100 40 GB	8 颗 HGX A100 40 GB	8 颗 HGX A100 40 GB
GPU 互联	PCI-E 4.0 ( 16 通道, 总 31.51 GB/s )	NVLink ( 600 GB/s )	NVLink ( 600 GB/s )	NVLink ( 600 GB/s )
内存	16 颗 镁光 32G DDR4 2933 RECC ( 总量 : 512 GB )	16 颗 32GB DDR4 ECC 3200MHz ( 总量 : 512 GB )	16 颗 32GB DDR4 RECC 3200MHz	16 颗 16GB
硬盘	SSD Intel NVME HDS-2VD- SSDPE2MX012T7	2块480G SATA SSD + 1块 1.92TB SATA SSD	1.92 T U.2 NVME	168GB
网卡		单口HDR 100高速IB网卡 + 双口25GE网卡		
电源	2200W带冗余	3000W电源模块 ( 3+1冗余 )		
系统	CentOS Linux release 8.0.1905 (Core)	CentOS Linux release 8.2.2004 (Core)	CentOS 8	CentOS 8.2.2004

## A280N Host

## I220N Host

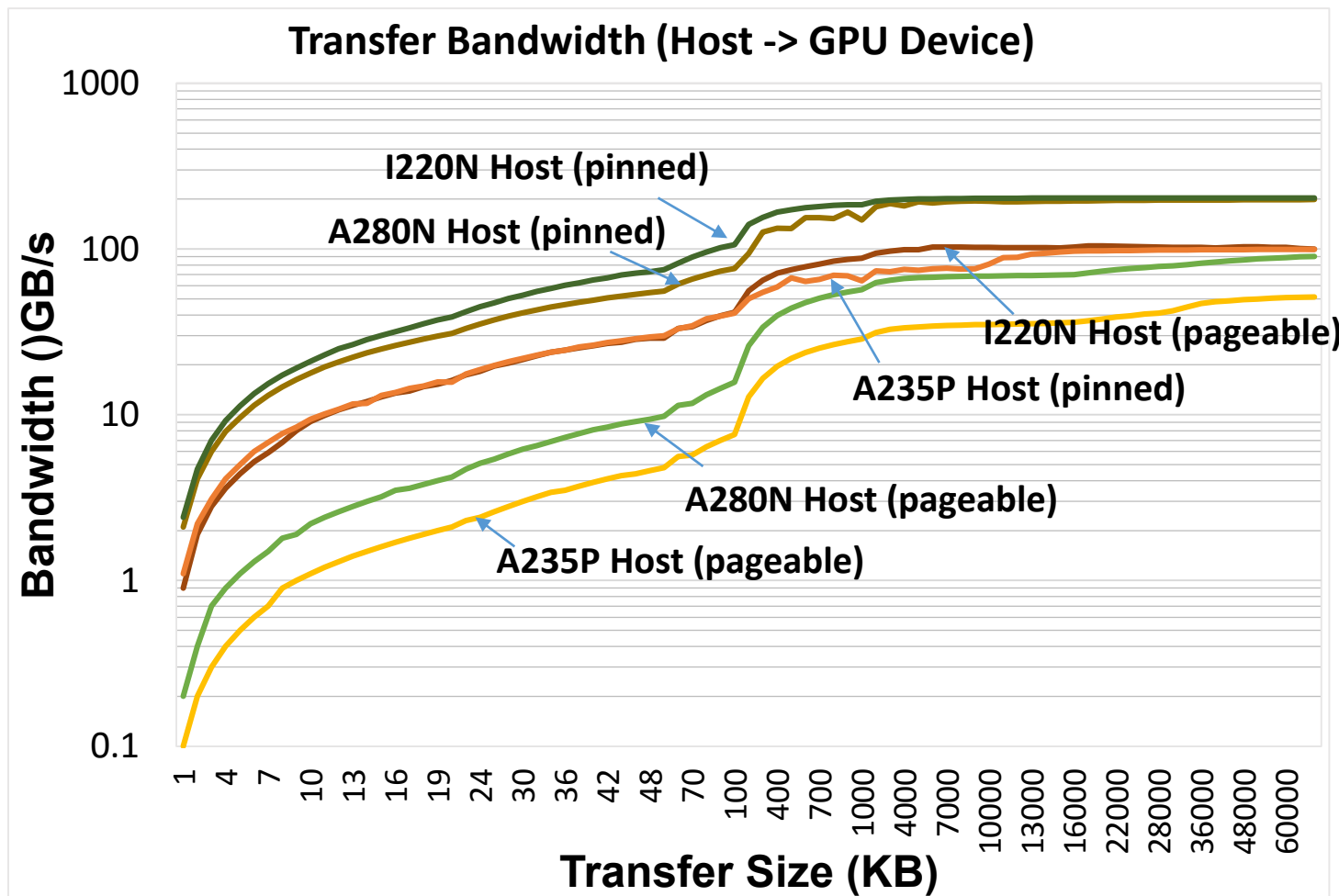


A280N Host具有8个NUMA分区，16个独立的L3 Cache；  
I220N Host具有2个NUMA分区，2个独立的L3 Cache。



# 第一部分： 硬件测试

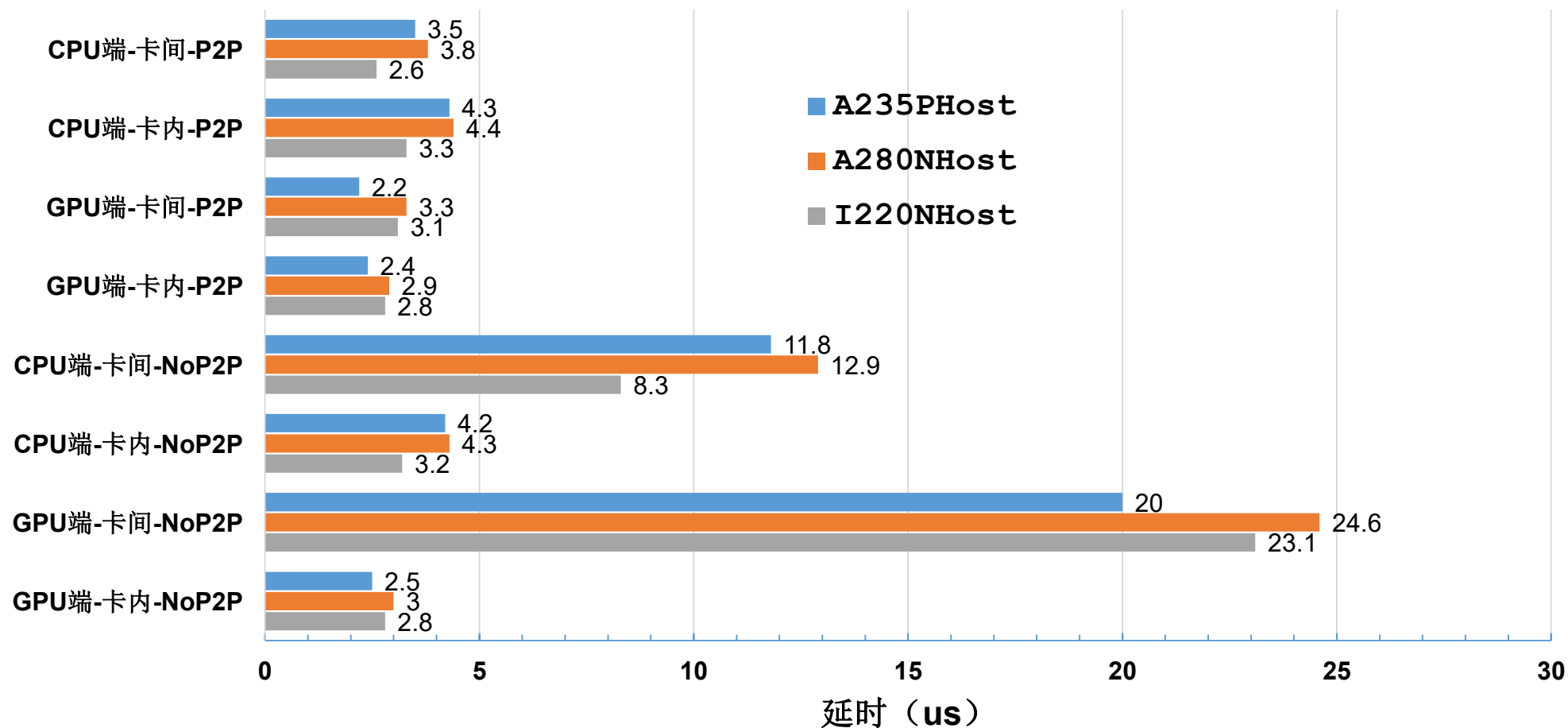
```
./bandwidthTest --device=all --memory=pinned --mode=shmoo  
./bandwidthTest --device=all --memory=pageable --mode=shmoo
```



- I220N Host性能最好
- A280N Host性能次之，且跟I220N Host差距明显。
- A235P Host性能较差
- pageable模式且Transfer Size小于100KB时，I220N Host的速度是A280N Host速度的3-4倍

对内存显存数据传输依赖较重的程序，选择I220N Host会有更好效果。

## GPU通信延时



- A235PHost与I220NHost具有明显更好的延时性能（前者经常表现更好）；
- A280NHost的表现总是最差。



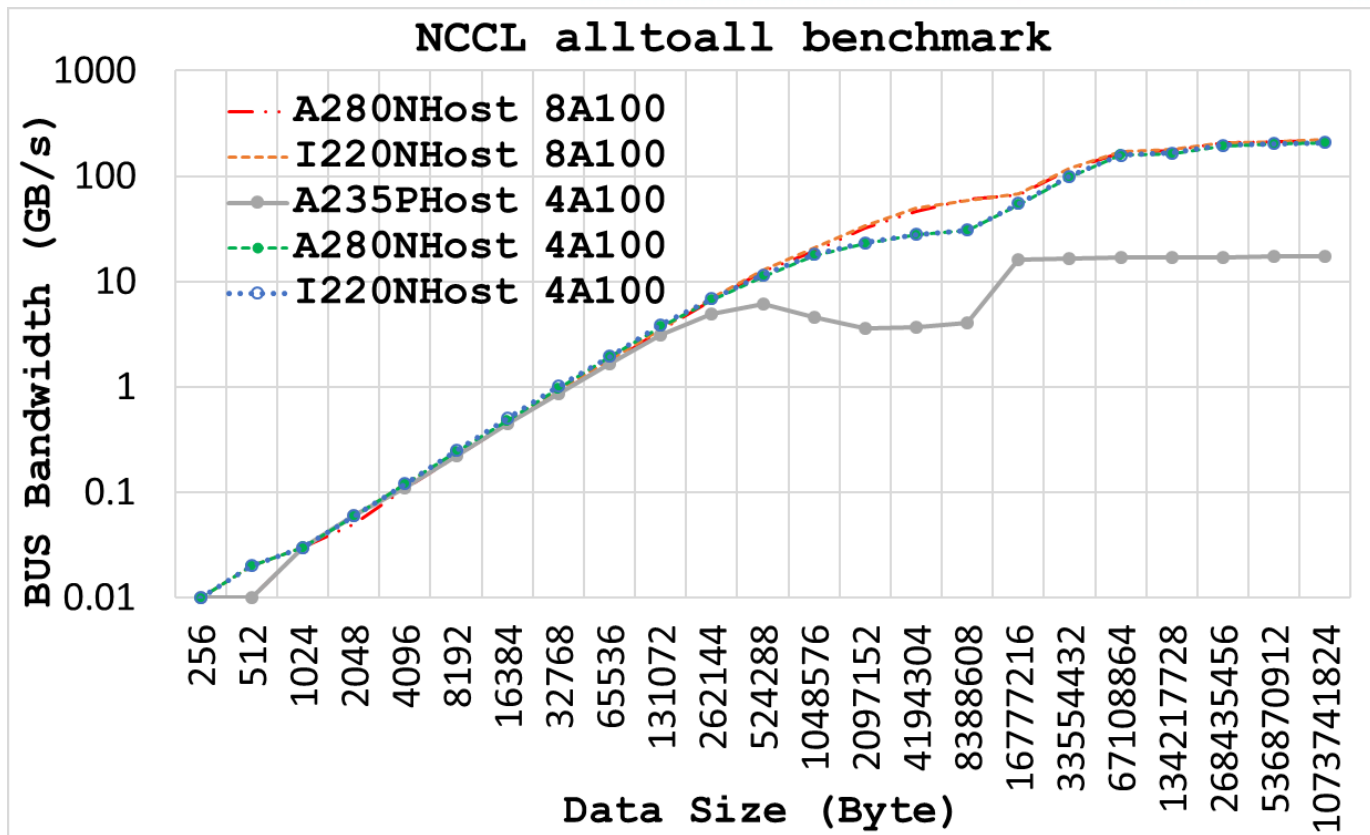
绿色表示性能较好  
蓝色表示性能一般  
红色表示性能较差

A235PHost / A280NHost / I220NHost GPU带宽 (GB/s)	
Bidirectional Bandwidth P2P=Disabled	
卡内	卡间
1305 / 1301 / 1301	29.3 / 18.4 / 26.2
Bidirectional Bandwidth P2P=Enabled	
卡内	卡间
1306 / 1296 / 1298	49.3 / 436 / 437

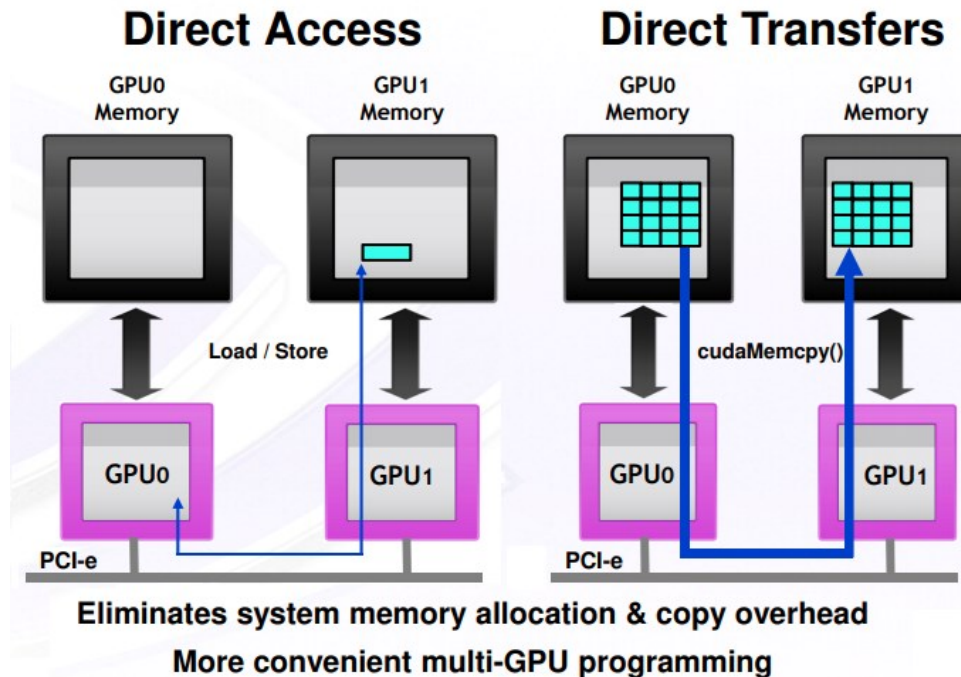
在P2P Disabled时，卡间性能最好的是A235PHost (~29.3GB/s)，其次是I220NHost (~26.2GB/s)，A280NHost性能最差 (18.4GB/s)。

在P2P Enabled时，A280NHost与I220NHost性能基本相同 (436/437 GB/s)，A235PHost (PCIe4) 性能很差 (49.3 GB/s)。

**A280NHost** 在P2P Disabled模式下的双向通信带宽速度异常(18.4 GB/s)，程序在该硬件下需要避免使用此通信模式。



## NCCL 自动使用 P2P Direct 通讯



[https://developer.download.nvidia.cn/CUDA/training/cuda\\_webinars\\_GPUDirect\\_uva.pdf](https://developer.download.nvidia.cn/CUDA/training/cuda_webinars_GPUDirect_uva.pdf)

- 在使用NCCL多卡通信库时，I220NHost与A280NHost的多卡alltoall通讯总线带宽性能基本相同
- 使用PCIe4的A235PHost多卡alltoall性能较差，与两个NVLink平台测试机的差距在10倍以上

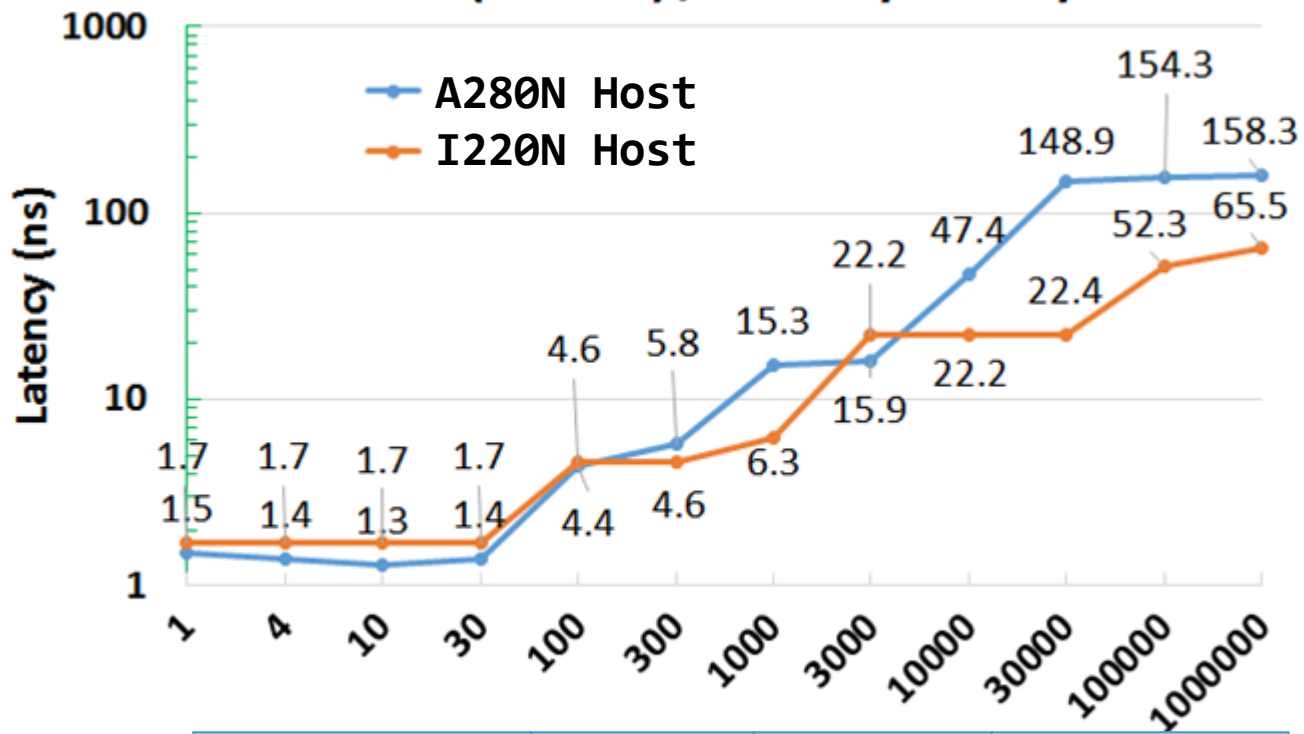
NCCL会自动使用P2P Direct通信，选择最快的连接路径，使应用发挥最佳的多卡间传输性能。因通讯绕过了CPU Host，两个NVLink平台具有基本相同的alltoall性能（其他allgather allreduce broadcast的性能类似）

# CPU缓存、内存的延时与带宽测试

## Intel(R) Memory Latency Checker v3.9

除L1缓存延时外，Intel CPU具有更好的内存(缓存)延时/带宽性能

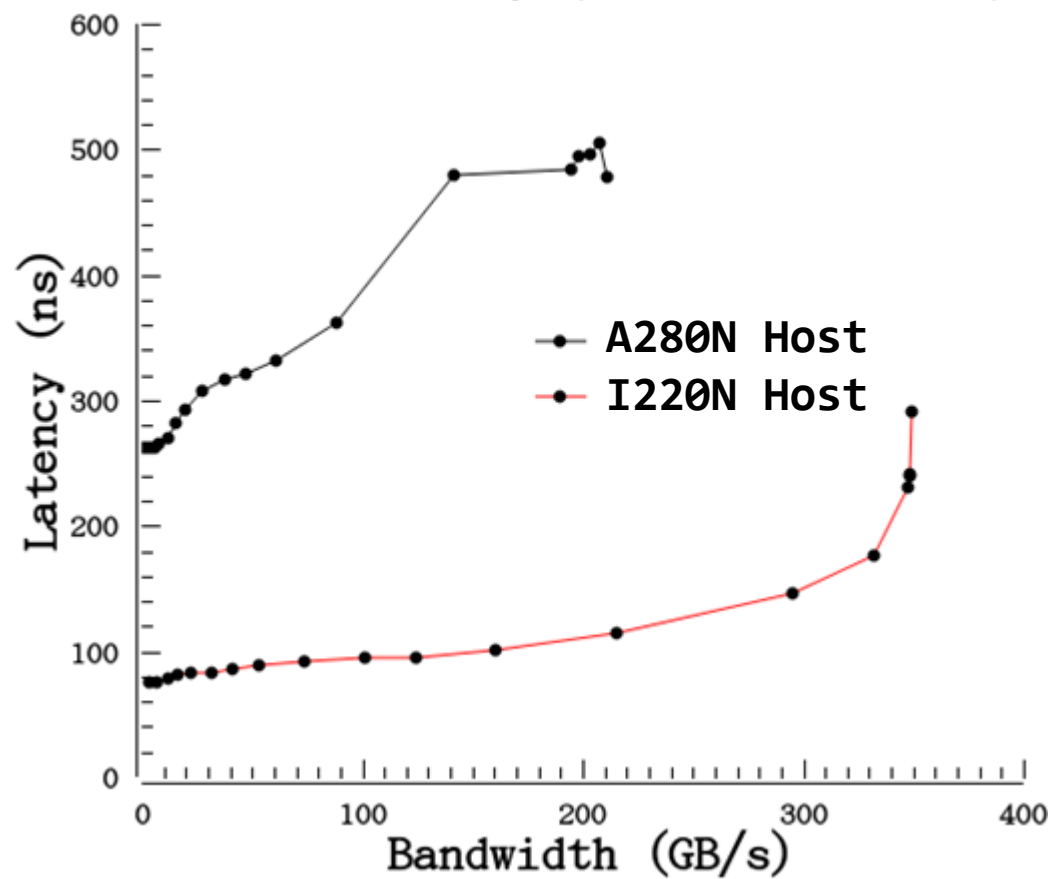
### Cache (L1 L2 L3) / Memory Latency



A280NHost:	L1	L2	L3	Memory
I220NHost:	L1	L2	L3	Memory

- A280NHost相比I220NHost具有稍低的L1 Cache延时 (~ 1.4ns vs. 1.7ns)，具有相似的L2 Cache延时。
- I220NHost具有明显更小的内存(Memory)延时(~ 60ns vs. 150ns)。

### Loaded Latency (Buffer Size: 1GB)



- A280NHost 最高负载仅能达到211GB/s，负载下的延时为500ns
- I220NHost最高负载可达到350GB/s，且该负载下的延时为290ns。

- 使用Intel CPU的I220NHost相比使用AMD CPU的A280NHost在内存延时与带宽上具有显著的性能优势，且在发起GPU指令的时候具有更快的非阻塞响应性能，其影响内存<->显存的数据传输，也影响非P2P Direct模式下的卡间通讯性能。
- 在使用NCCL卡间数据通信库时，其自动使用最优的通信路径。在P2P Direct通信模式下，主机的性能差异不影响卡间的直接通信，所以两个NVLink GPU平台具有相似的AlltoAll等通信性能，且相比PCIe4平台下的卡间通信性能具有10倍以上的优势。
- PCIe平台具有可见的卡间延时优势，且在关闭P2P Direct时与NVLink平台的卡间通信带宽差异不明显。因此，在程序编写不当或卡间通信数据（Buffer Size）较小时，PCIe平台可能与NVLink平台具有相似的性能，甚至前者会更好。



## 第二部分：应用测试

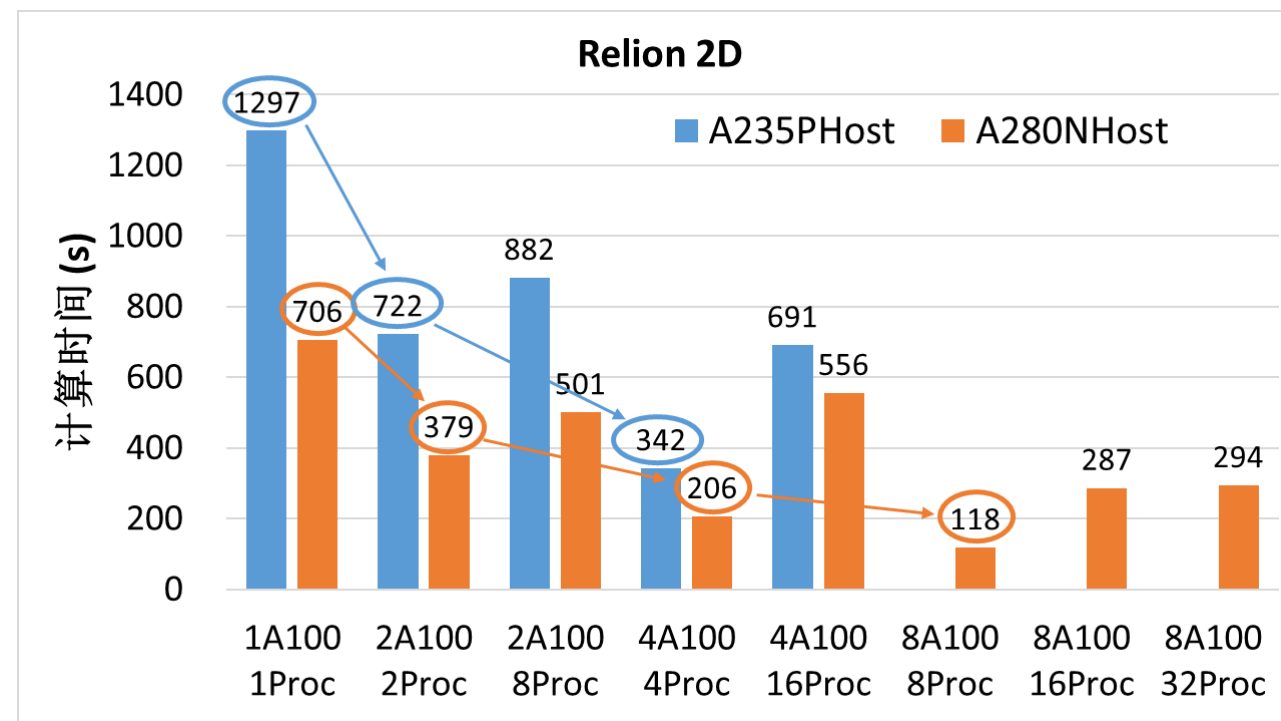
**软件配置**：RELION3.1.1，OpenMPI 4.0.2，CUDA 11.0，GCC 8.3.1，CMake 3.11.4

## 算例说明：

- 1) 对大小40~50 GB 的原始32帧movie ( 293张 ) 进行motion-correction，得到矫正位移后的micrographs；
- 2) 利用CTFFIND4 软件对micrographs进行衬度信息传递，求解方程contrast transfer function；
- 3) 对micrographs中的颗粒进行挑选，这里采用reference-free的Laplacian-of-Gaussian算法。捕获颗粒坐标后，对相应坐标位置的颗粒图形进行提取，获取约 3.6 GB大小的颗粒。
- 4) **2D测试时**，利用最大似然法对蛋白颗粒进行分类；**3D测试时**，采用最大似然法和傅里叶变换进行三维重构。
- 4) 已经提取的颗粒进行2D或3D Classification，分别采取分类数k=10与1，迭代次数iter=3，每次向内存输入16个颗粒，总颗粒数94,000个。搜索角度参数：offset\_range=5、offset\_step=2，对称性设置：sym=C1。

## 测试结果解析 :

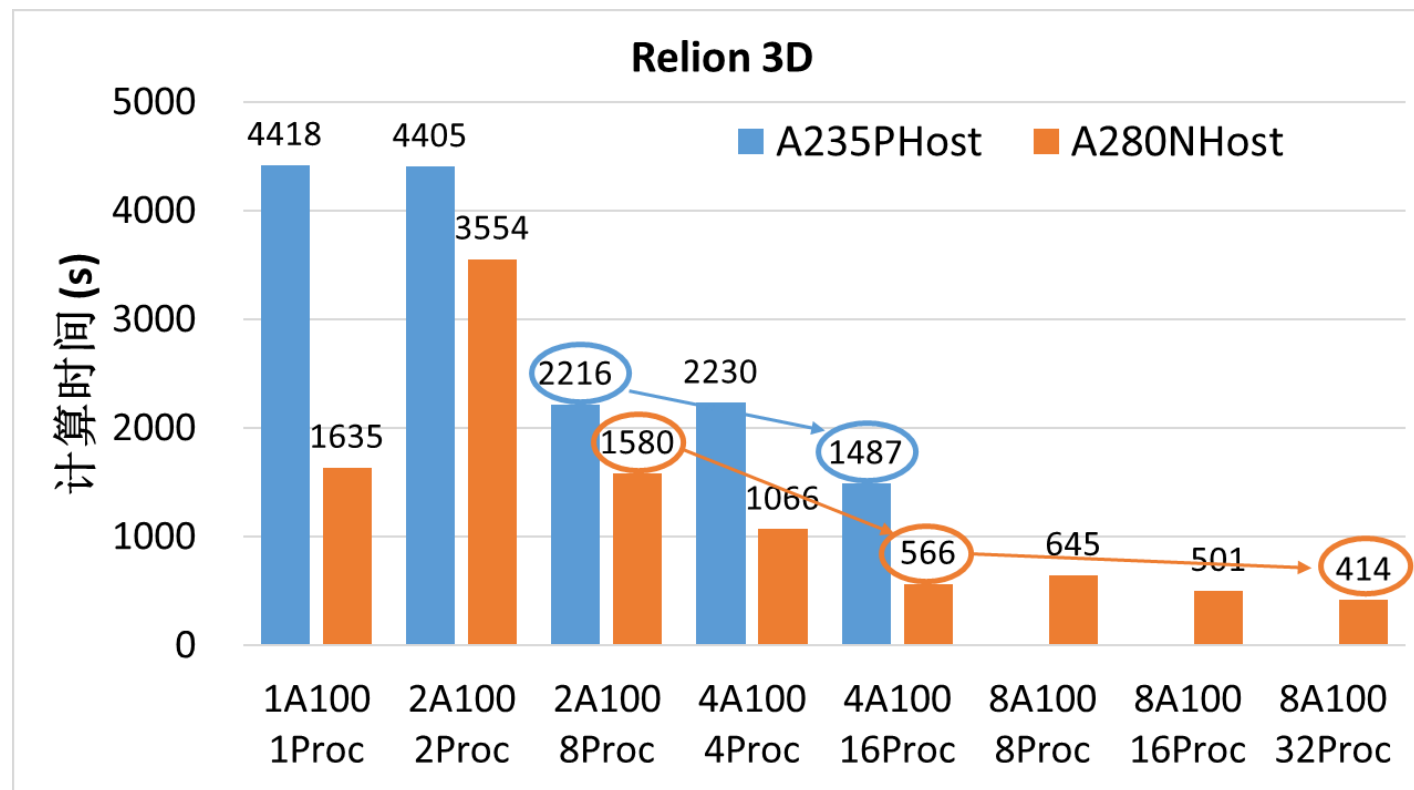
- 每卡分配单个进程时速度最快
- 4卡NVLink平台与4卡PCIe4平台下的速度比为：  
 $342 / 206 = 1.66$
- NVLink平台下，并行度从4卡扩展至8卡时，效率为：  
**87%**
- 计算量估计：**2.3卡·年**



NVLink 平台下扩展到8卡时，依然具有较高的并行效率（87%）

## 测试结果解析

- 每卡分配单个进程时速度最快
- 4卡NVLink平台与4卡PCIe4平台下的速度比为： $1487 / 566 = 2.6$
- NVLink平台下，并行度从4卡扩展至8卡时，效率为：**68%**
- 计算量估计：**2.9 卡·年**



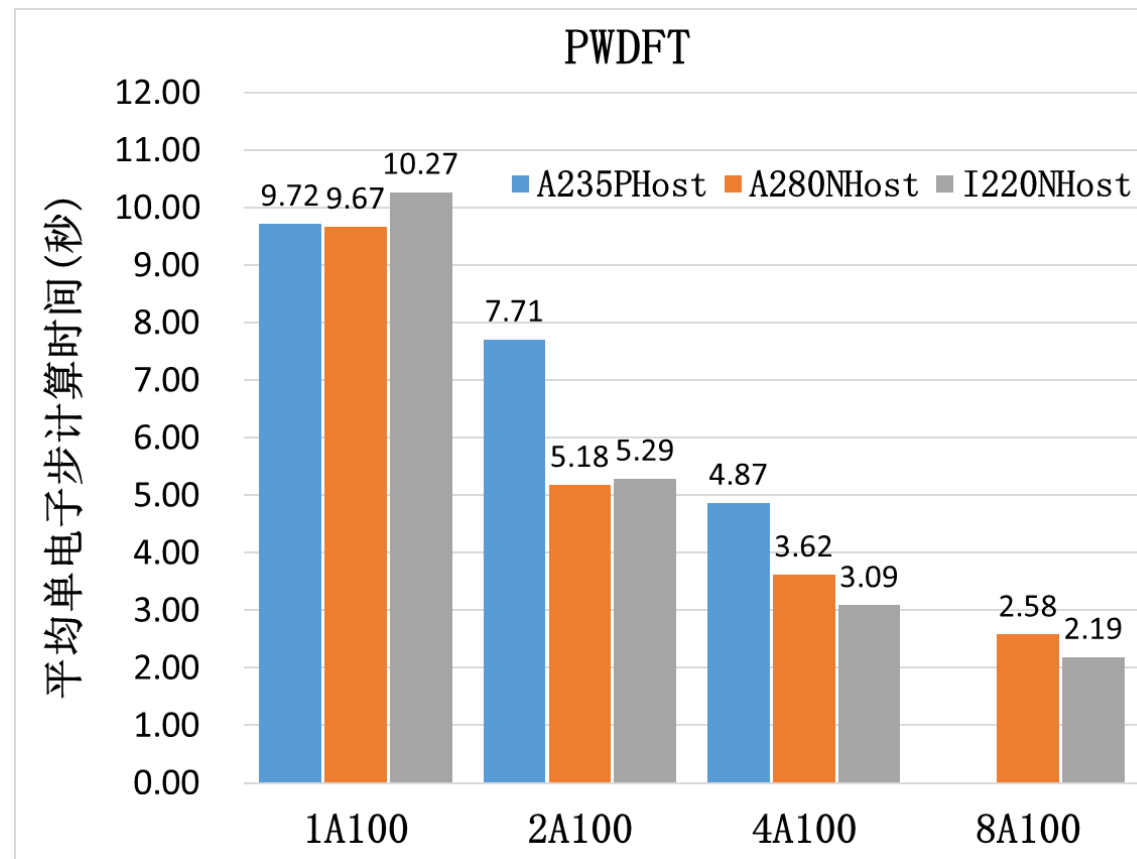
NVLink平台相比PCIe平台具有显著的计算速度优势（4卡时2.6倍）

**软件配置：** CUDA11 , FFTW3.3.8 , libxc4.2.3 , 使用Intel icc ifort 与 OpenMPI 编译。

**算例说明：** 使用具有代表性的Si512体系，截断能 10 Hartree，计算直到收敛，平均每电子步时间的计算方法是最后10 SCF步取平均，单位：秒。

## 测试结果解析：

- 多卡时，使用Intel CPU与NVLink的I220NHost具有最佳性能
- 4卡NVLink平台 ( I220NHost ) 与4卡PCIe4平台下的**速度比**为：  
 $4.87 / 3.09 = 1.57$
- NVLink平台 ( I220NHost ) 下，并行度从4卡扩展至8卡时，效率为：  
**70%**
- 计算量估计：**34 卡·年**



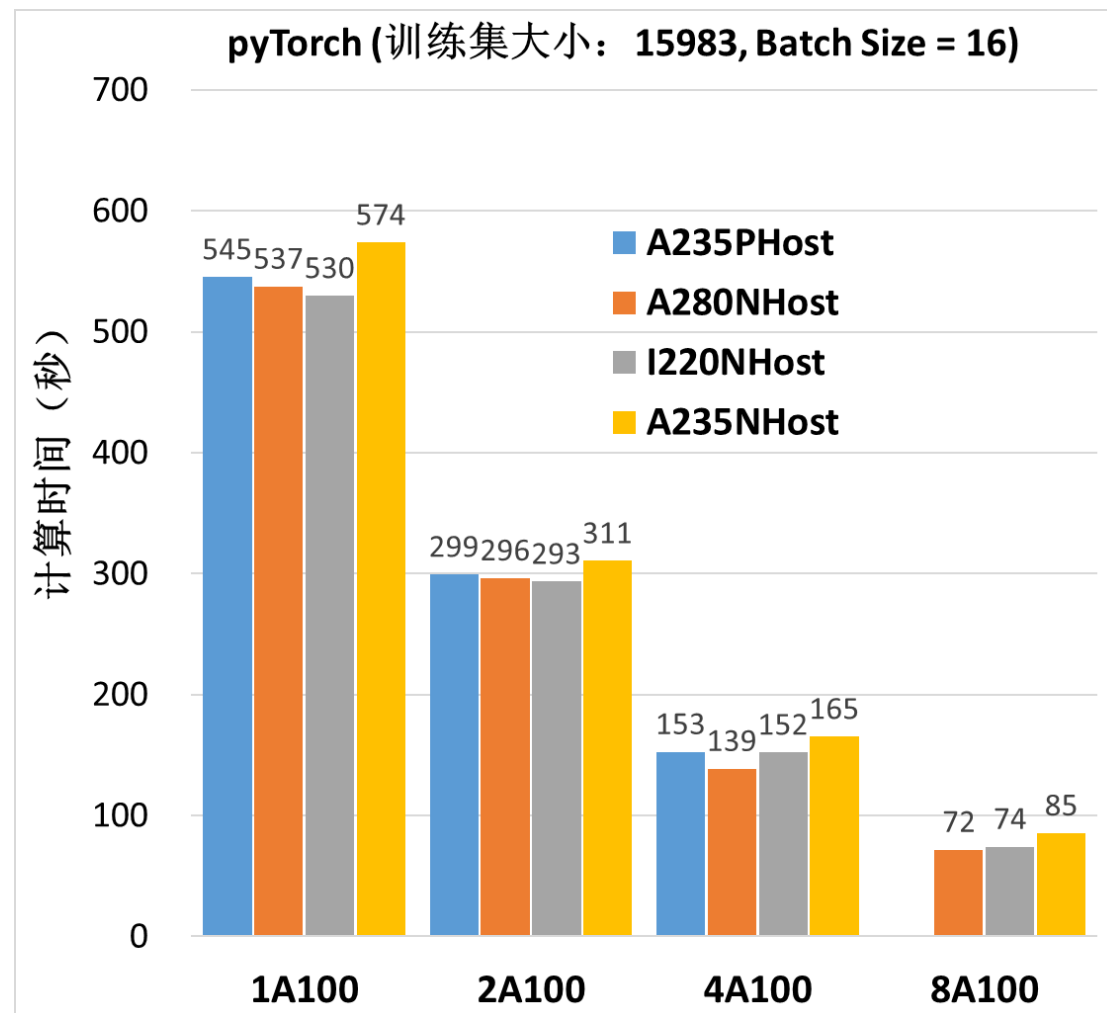
# 物化应用 pyTorch 测试：算例1

**软件配置**：pyTorch 1.7.0 , CUDA 11.0 , cuDNN 8.0.5 , NCCL 2.8.3 ,  
OpenMPI 4.0.5

**算例说明**：神经网络拟合Cu-C训练集。训练集大小：63935，参数总个数：1283331，Batch size = 16。统计每Epoch耗时（秒）。

## 测试结果解析：

- A235NHost的CPU主频较低，其性能相比A280NHost低 5%-15%。
- 4卡NVLink平台（A280NHost）与4卡PCIe4平台下的速度比为：  
 $153 / 139 = 1.10$ ，NVLinkd的效用不高。
- NVLink平台（A280NHost）下，并行度从4卡扩展至8卡时，效率为：  
96%
- 计算量估计：3.3 卡·年

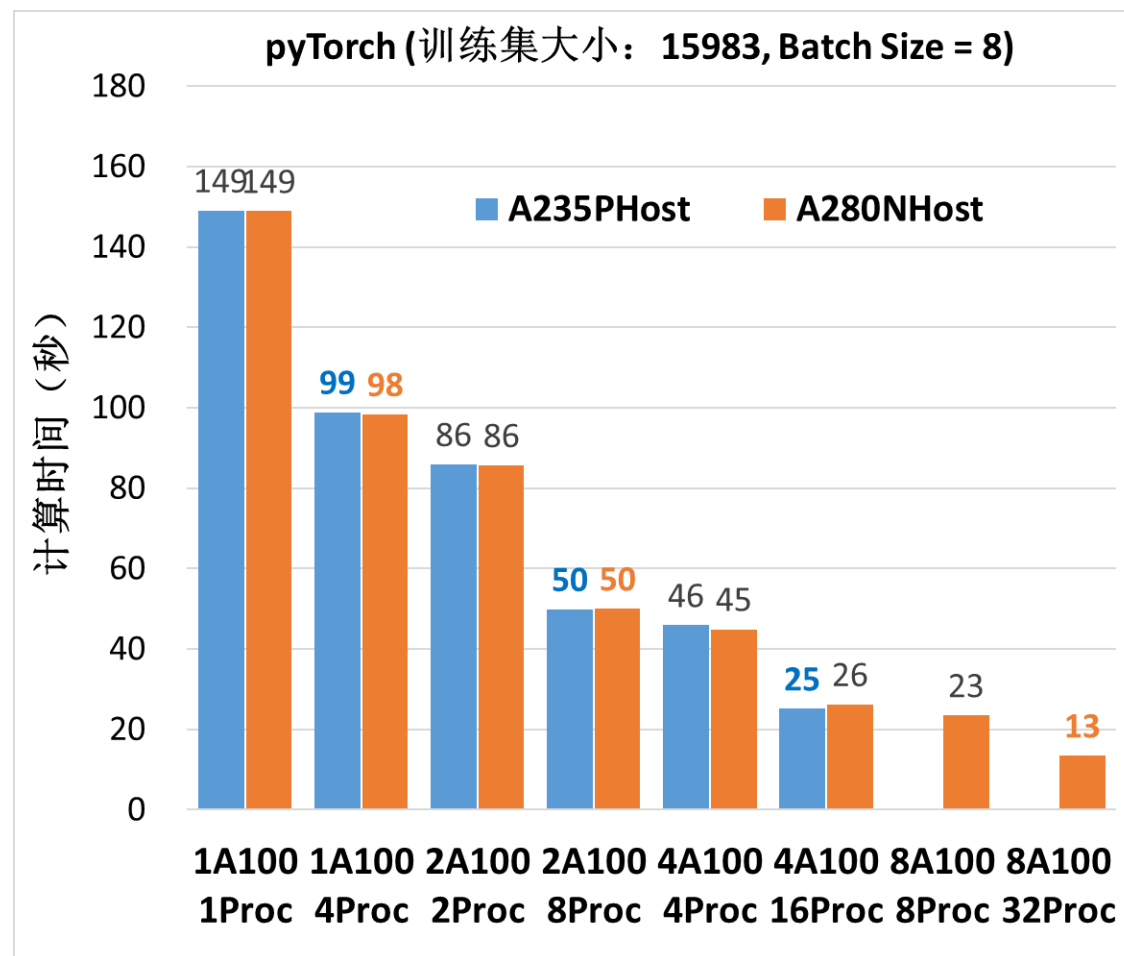


**软件配置：** pyTorch 1.7.0 , CUDA 11.0 , cuDNN 8.0.5 , NCCL 2.8.3 ,  
OpenMPI 4.0.5

**算例说明：** 神经网络拟合Cu-C训练集。训练集大小：15983，参数总个数：1283331，Batch size = 8。统计每Epoch耗时（秒）。**本算例目的是考察单卡多进程共享模式(MPS)的性能，受版本故障限制，pyTorch算例1下无法完成此测试。**

## 测试结果解析：

- 开启单卡多进程共享模式后，计算性能在1、2、4、8卡下分别加速约**1.50、1.72、1.78、1.76倍**。在8卡平台，最好需要32个以上CPU Core来支持此类计算，所有平台均可以满足。
- **NVLink硬件在此类算例下的效能较低。**

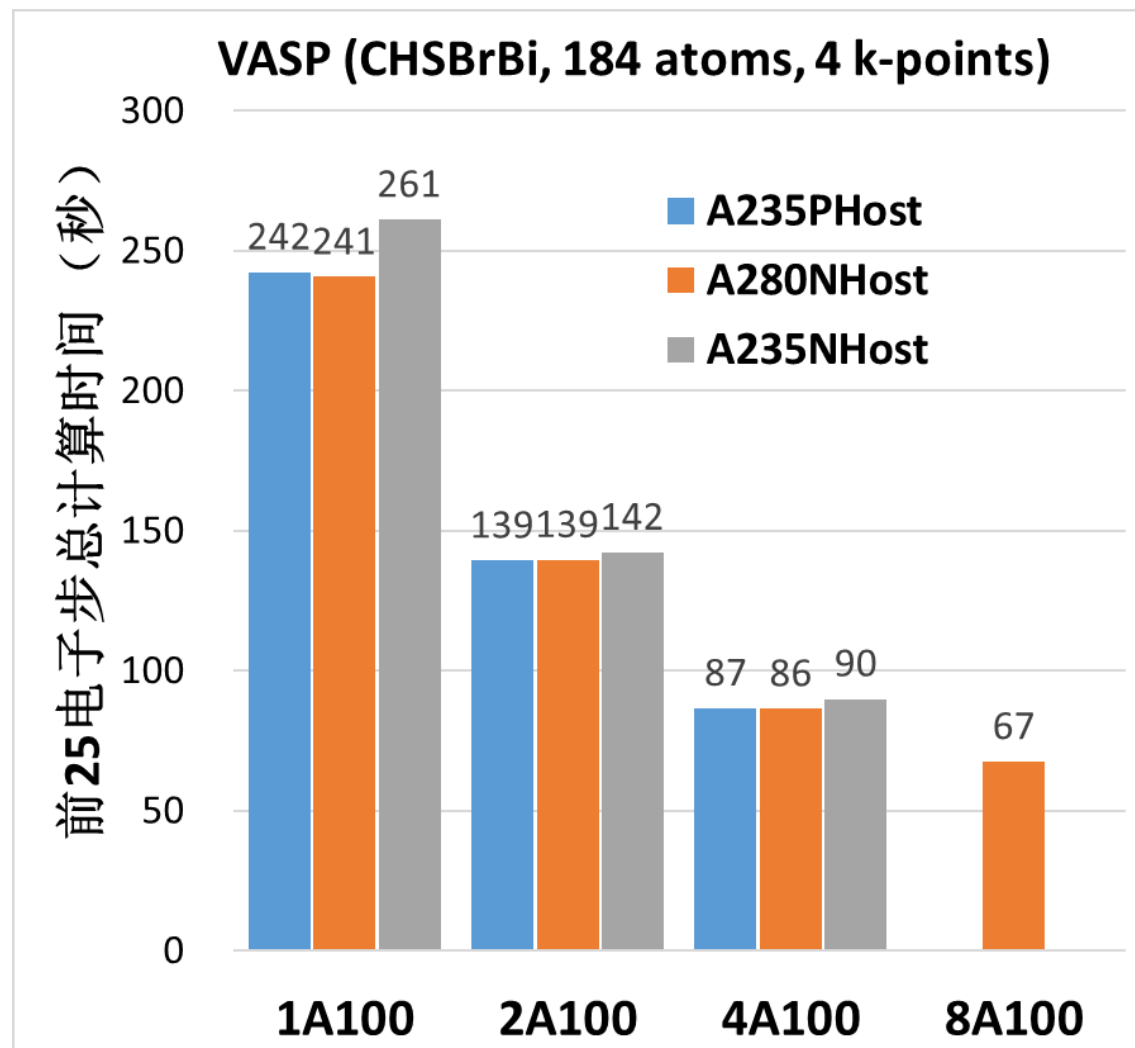


**软件配置：** VASP 6.1.0 (OpenACC) , NVHPC 20.9 (CUDA 11.0 & OpenMPI 3.1.5) , NCCL 2.8.3

**算例说明：** CHSBrBi的SCF计算，统计第一个离子步内前25个电子步总时间（秒）。ENCUT = 800 NIONS (总原子数) = 184 , NELECT (总电子数) = 720 , NKPTS (总K点数) = 4, NBANDS = 453 , NEDOS = 301, NSIM=16, KPAR = 4, ALGO = Normal (IALGO = 38), NELM = 100

## 测试结果解析：

- 同样采用AMD CPU , A235NHost比A280NHost速度慢5%-8% , 应源于前者主频低
- 多K点并行下 , 卡间通讯较少 , 此时PCIe与NVLink对速度影响不大
- 4卡NVLink平台与4卡PCIe4平台下的速度基本相同
- NVLink平台下 , 并行度从4卡扩展至8卡时 , 效率为 : 64%。此较差结果源于总K点只有4个 , 所以只能高效率的扩展到4卡计算
- 计算量估计 : 4.0 卡·年

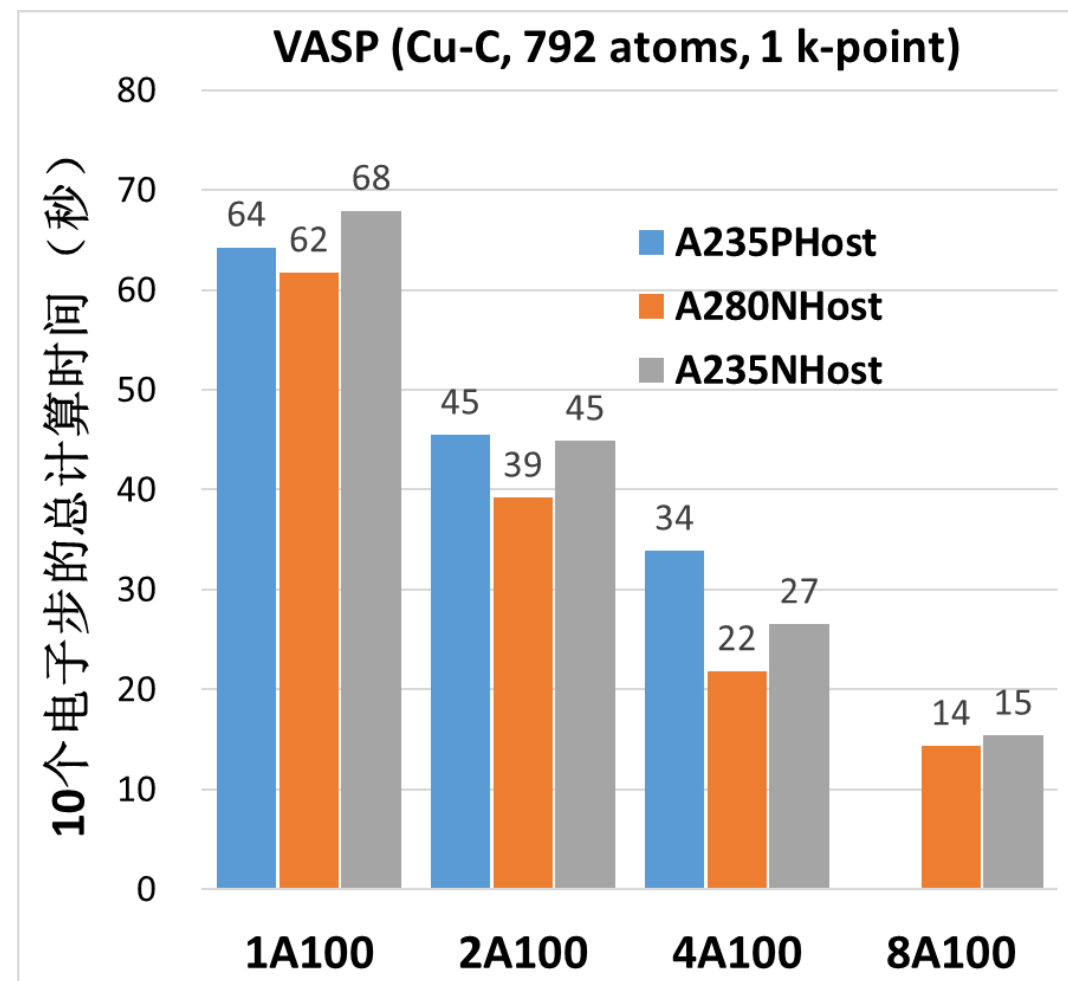


**软件配置：** VASP 6.1.0 (OpenACC) , NVHPC 20.9 (CUDA 11.0 & OpenMPI 3.1.5) , NCCL 2.8.3

**算例说明：** Cu-C的第一性原理MD , **单Gamma点** (无K点并行) , 统计10个 (前5步之后) 电子步的平均时间 (秒) 。 ENCUT = 400 NIONS ( **总原子数** ) = 792, NELECT ( **总电子数** ) = 7200 , NKPTS= 1, NBANDS = 4320 , NEDOS = 301 KPAR = 1, ALGO = Normal (IALGO = 38), NELM = 60

## 测试结果解析：

- 采用同样的内存与AMD-CPU , 主频更低的A235NHost比A280NHost速度慢**9%-18%**
- **单K点计算下, 卡间通讯较多** , 此时PCIe与NVLink对速度影响很大
- 4卡NVLink平台与4卡PCIe4平台下的速度比为： $34 / 22 = 1.55$
- A280NHost平台下, 并行度从4卡扩展至8卡时, 效率为：**79%**, 效率较高。
- 计算量估计：**3.0 卡·年**

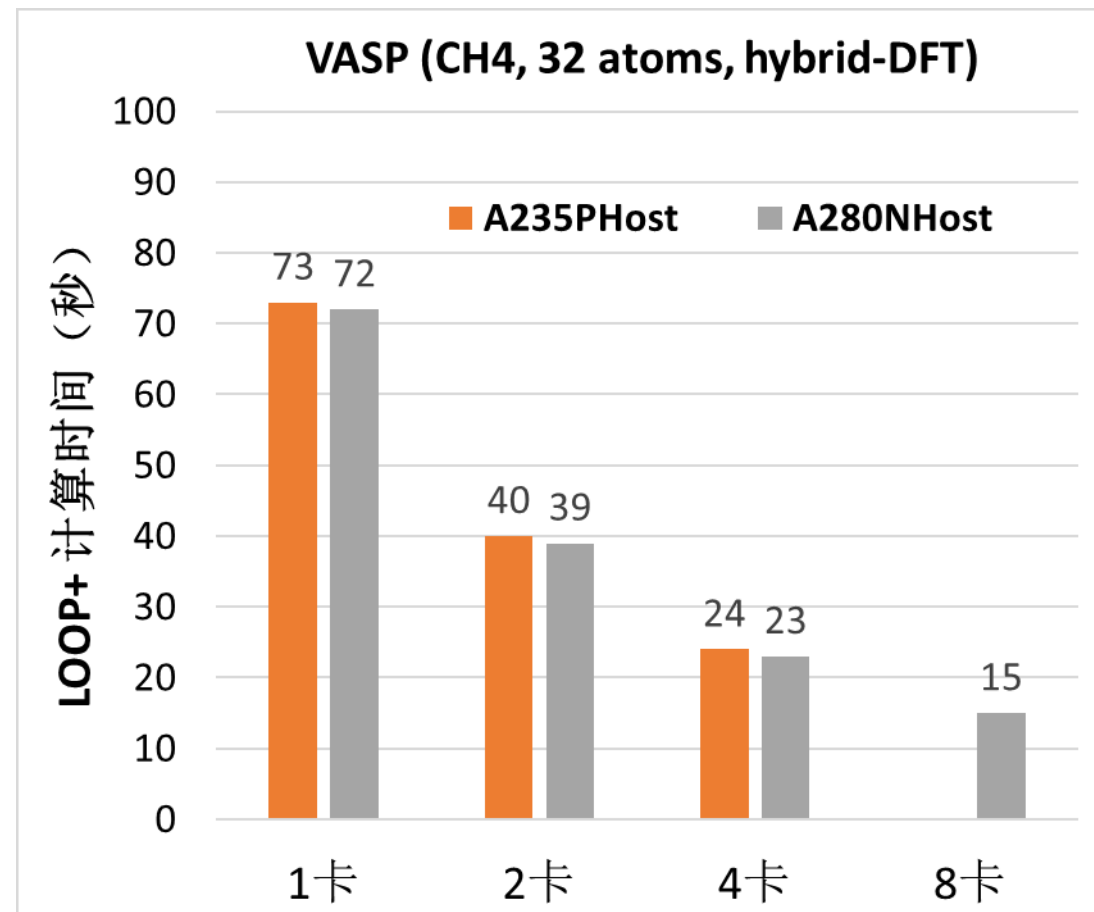


**软件配置：** VASP 6.1.0 (OpenACC) , NVHPC 20.9 (CUDA 11.0 & OpenMPI 3.1.5) , NCCL 2.8.3

**算例说明：** CH<sub>4</sub> 体系 , HybridDFT计算 , 参数 : NIONS=32 NBANDS=224 NELECT = 278 ENCUT=380 LHFCALC=.TRUE. IALGO=48 NELM=3 NKPT=5 NGX/Y/Z=48/48/98。

## 测试结果解析：

- 属于多K点并行计算，卡间通讯较少，PCIe与NVLink对速度影响很小
- 4卡NVLink平台与4卡PCIe4平台下的速度基本相同
- A280NHost平台下，并行度从4卡扩展至8卡时，效率为：**77%**，效率较高。
- 计算量估计：**2.0 卡·年**

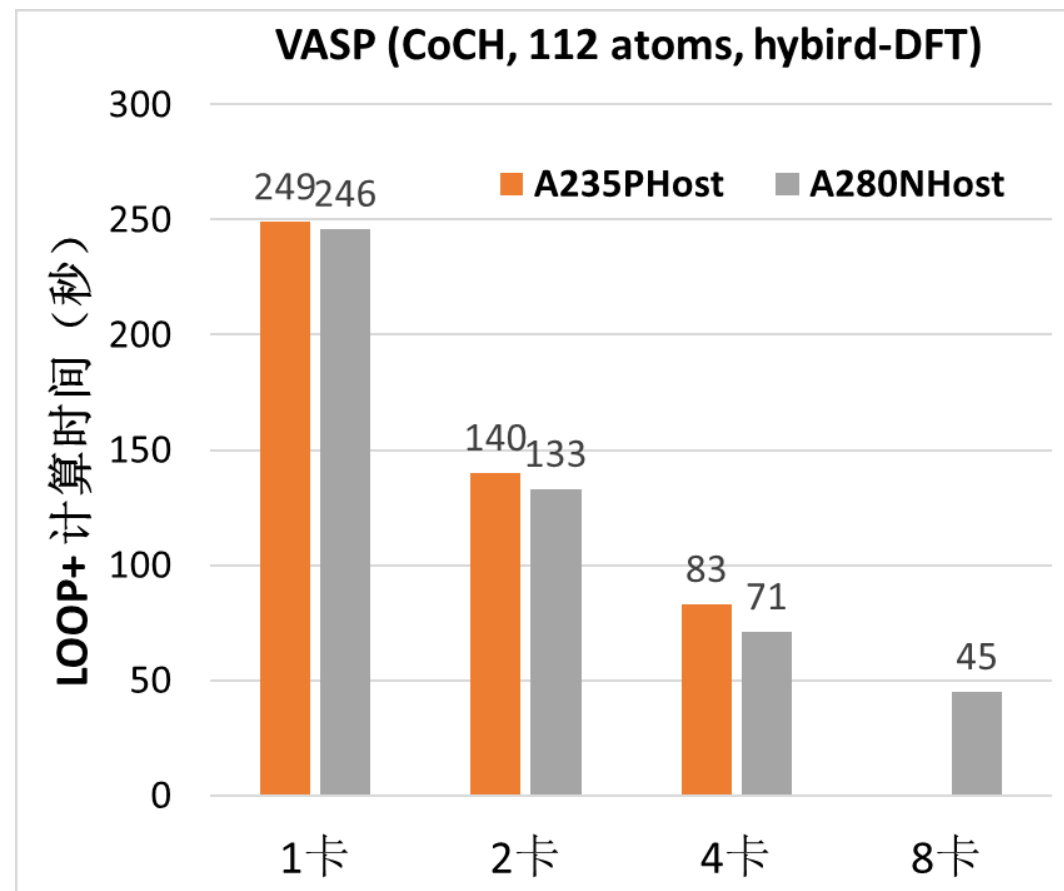


**软件配置**：VASP 6.1.0 (OpenACC) , NVHPC 20.9 (CUDA 11.0 & OpenMPI 3.1.5) , NCCL 2.8.3

**算例说明**：CoCN 体系，HybirdDFT计算，NIONS=112 NBANDS=384  
NELECT=456 ENCUT=400 LHFCALC=.TRUE. IALGO=38 NELM=3  
NKPT=2 NGX/Y/Z=108/42/128

## 测试结果解析：

- K点个数较少（2个），卡间通讯一般，PCIe与NVLink对速度具有一定的影响
- 4卡NVLink平台与4卡PCIe4平台下的速度比为： $83 / 71 = 1.17$
- A280NHost平台下，并行度从4卡扩展至8卡时，效率为：**79%**，效率较高。
- 计算量估计：**2.0 卡·年**

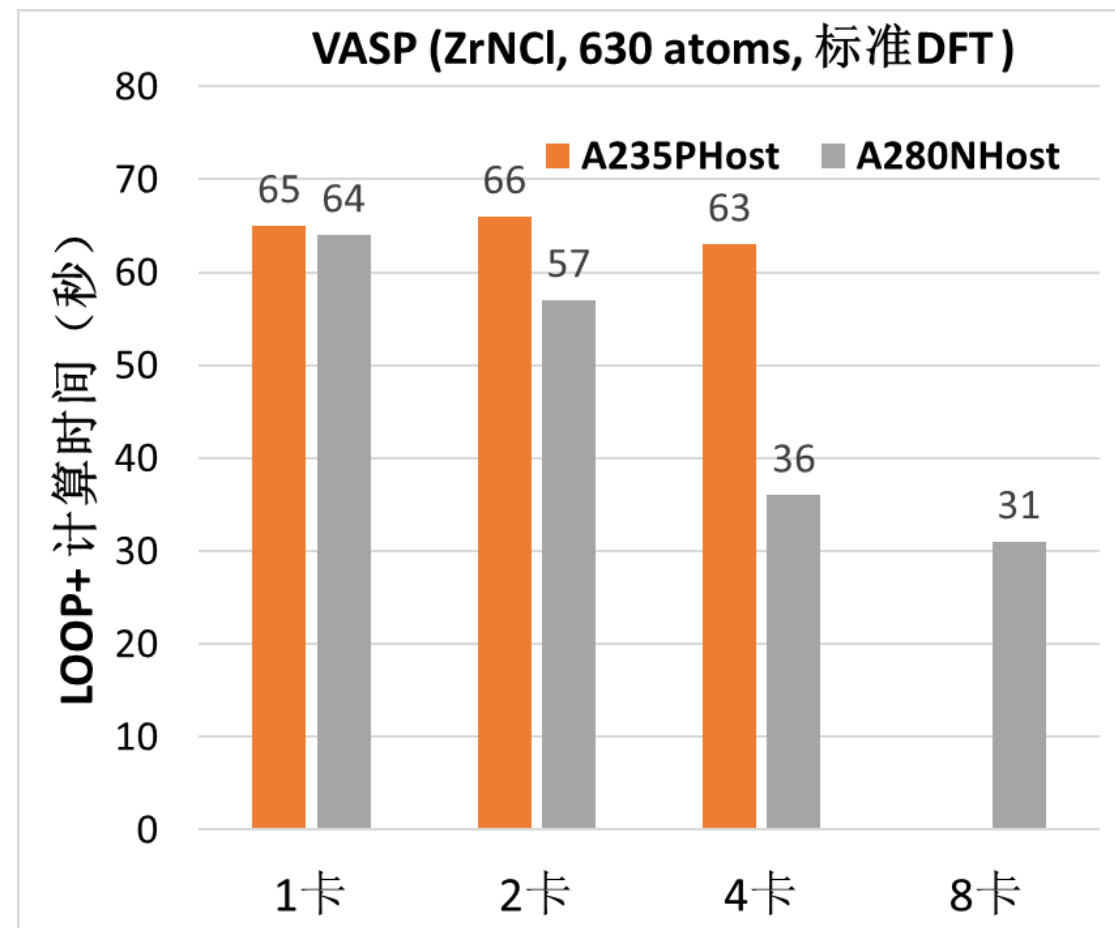


**软件配置**：VASP 6.1.0 (OpenACC) , NVHPC 20.9 (CUDA 11.0 & OpenMPI 3.1.5) , NCCL 2.8.3

**算例说明**：ZrNCl 体系，标准DFT计算，参数：NIONS=630  
NBANDS=3034 NELECT=5058 ENCUT=400 LHFCALC=.FALSE.  
IALGO=38 NELM=5 NKPT=1 NGX/Y/Z=108/108/140

## 测试结果解析：

- 单个K点，卡间通讯要求高，PCIe与NVLink对速度具有较大影响
- 4卡NVLink平台与4卡PCIe4平台下的速度比为： $63 / 36 = 1.75$
- A280NHost平台下，并行度从4卡扩展至8卡时，效率为：**58%**，效率较低。
- 计算量估计：**1.0 卡·年**



**软件配置：** GCC , Intel Fortran 2013 , CUDA-11 , OpenMPI-4.0.1(基于CUDA) , HWLOC-2.0.4

## 算例说明：

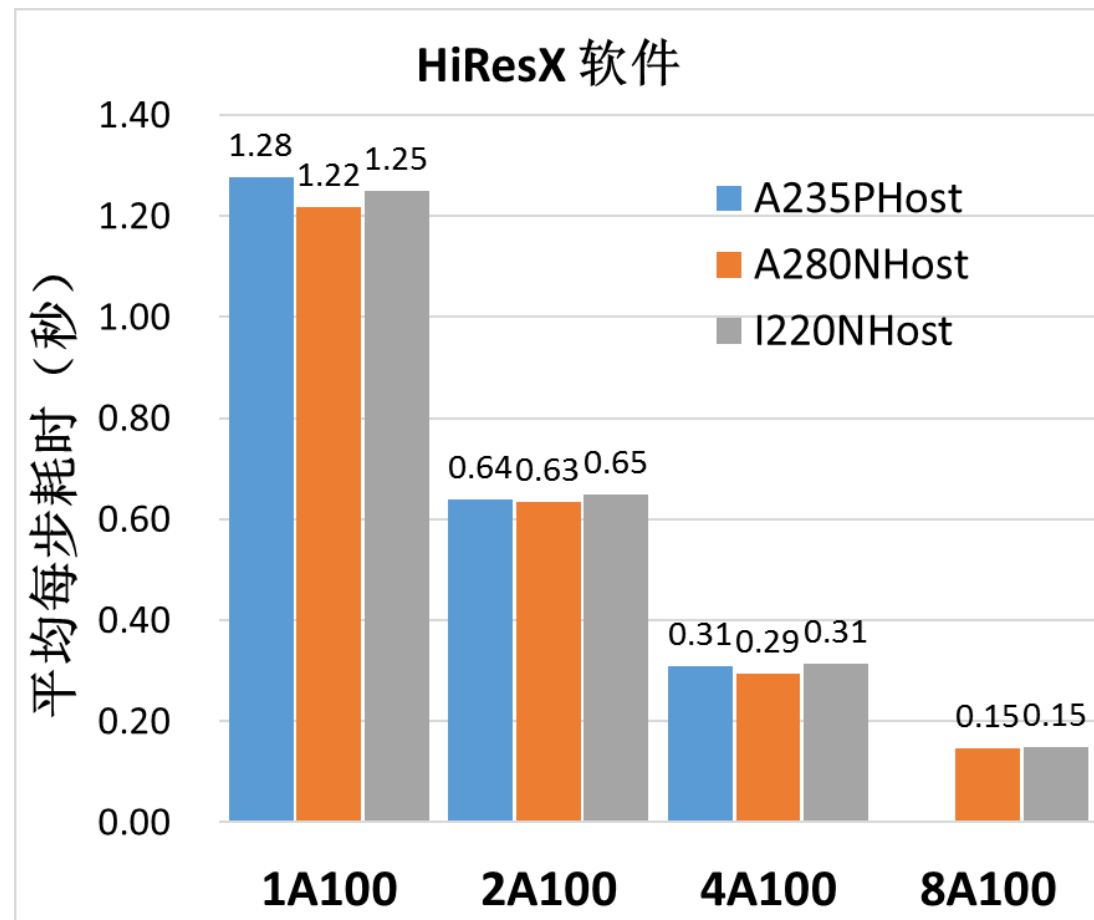
(1) 计算域为一个立方体，总共8000万网格点，求解Navier-Stokes方程，边界均为远场边界；

(2) 对流项采用5阶HWCNS格式，粘性项采用6阶中心差分格式，时间推进采用3步Runge-Kutta格式；

(3) 推进时间步数为2000步，统计平均每步的耗时。

## 测试结果解析：

- 4卡NVLink平台与4卡PCIe4平台下的速度比为**1.07**，速度基本相同。
- A280NHost平台下，并行度从4卡扩展至8卡时效率依然很高，为**97%**。
- 计算量估计：**20 卡·年**



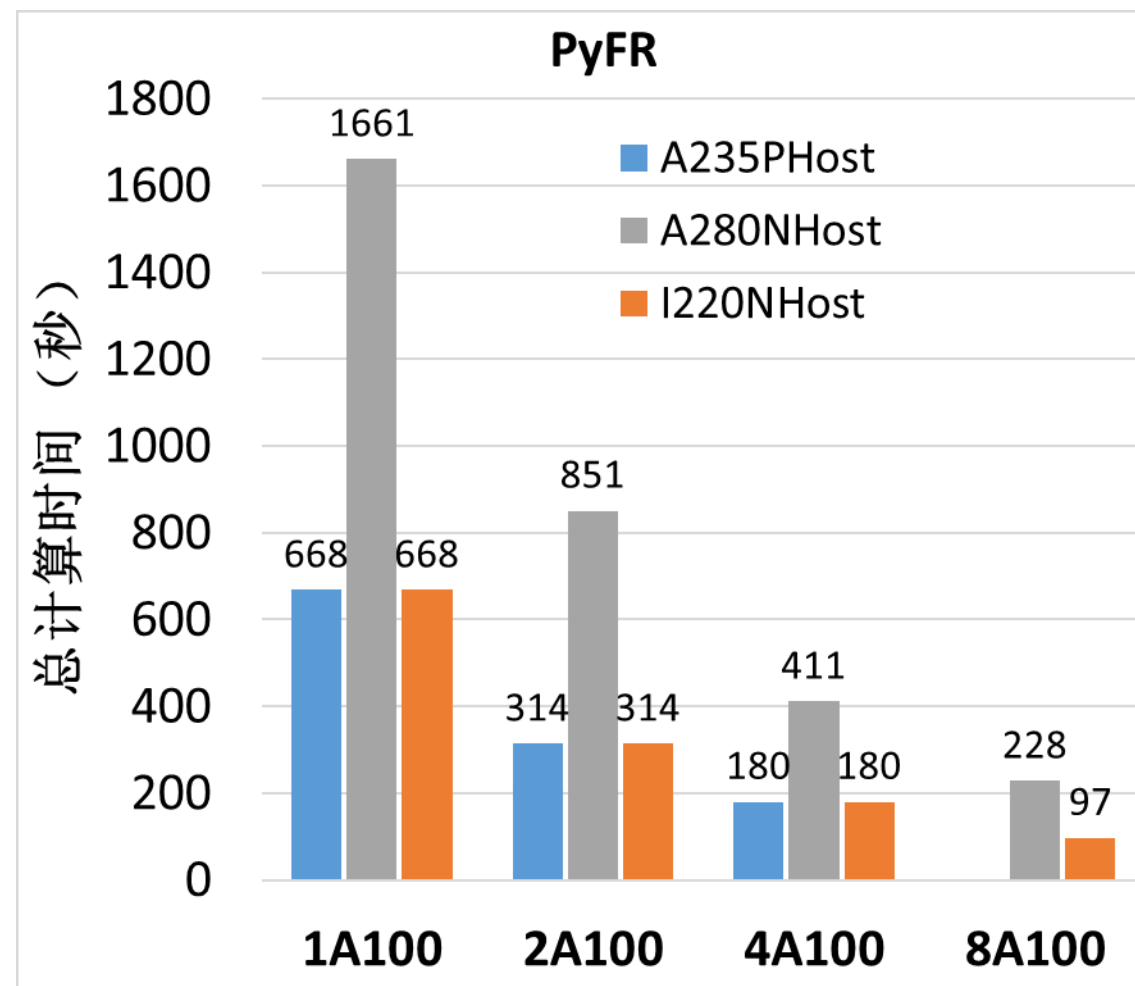
**软件配置：** GCC , CUDA-11 , OpenMPI-4.0.1(基于CUDA) , python-3

## 算例说明：

- (1) Taylor-Green Vortex标准算例测试，总自由度约为6210万；
- (2) 空间采用5阶通量重构法，时间推进采用3步Runge-Kutta格式；
- (3) 推进时间步数为1000步，统计总耗时。

## 测试结果解析：

- 使用**AMD CPU的A280NHost的计算时间异常高**，或源于硬件测试部分中展示的延时/通讯异常
- 4卡NVLink平台与4卡PCIe4平台下的速度比为**1.0**，速度基本相同。
- A280NHost平台下，并行度从4卡扩展至8卡时效率依然很高，为**93%**。
- 计算量估计：**10 卡•年**





## 第三部分： 用户建议



- 建议**选择NVLink**的GPU平台。理由：随着电镜拍照速度的提高，数据量将越来越大，更好的配置可以更好的适应未来几年数据增长带来的计算压力。
- **可以选择混搭配置单节点8卡和单节点4卡**，分别应对不同的数据计算要求，8卡并行将大幅度降低计算总时间。目前两种配置也是RELION计算平台比较常用的。

- 如果大量采购A100平台，**建议购买NVLink**的GPU平台，4卡就可以相对于PCI-E平台加速50%。不建议购买平台2（A280NHost），同样是支持NVLink的平台，平台4（I220NHost）的表现更好。
- **建议购买8卡GPU**，8卡可以**最大程度的发挥NVLink的性能**，同时也可以减少GPU节点的数量，**节约CPU**，内存等部件的成本，**减少空间的占用**。



1、建议新集群选择基于 PIC-E 做卡间互联的GPU平台？还是基于 NVLink 的GPU平台？及其理由：

(综合考虑：NVLink平台的价格约增加10%，且运行最大功耗从250W增加至400W)

目前本人的课题所需的程序对GPU通信的要求不算很大，PCI-E基本能满足。但是类似于算例4这类本身难以进一步优化通信的程序或算例应该并不少见，此外未来的课题也很有可能因为需求增长导致PCI-E不够用，故**推荐NVLink的平台**。

2、对单节点选择配置4卡或8卡GPU的建议，及其理由：

**建议单节点配置8卡**，就个人涉及到的一些研究而言：

- VASP涉及杂化泛函甚至加SOC的体系的计算，以及针对化学模拟的神经网络的拟合和使用，对资源的消耗极大，**单节点8卡可以完全用满**。
- 使用高通信需求的程序时，**单节点8卡的效率会比2个4卡节点高**。
- GPU版的程序开发时，**优化节点内的通信会比跨节点容易**。
- 配合NVIDIA MPS服务应该**可以合理调度空闲的GPU资源，避免进程之间相互影响或是资源浪费**。

## 1、建议新集群选择基于 PIC-E 做卡间互联的GPU平台？还是基于 NVLink 的GPU平台？及其理由：

(综合考虑：NVLink平台的价格约增加10%，且运行最大功耗从250W增加至400W)

- 本实验室的程序并行效果较好，测试结果表明NVLink和PCI-E导致的性能差异不明显，或源于PCIe 4.0相比上一代带宽更高。因此，**如果学校用户对GPU需求量很高，资源不够用，则建议采购更多的GPU，即采用PCI-E版本。**
- **如果GPU资源相对充足，则采购NVLink的机器对应用开发有益处。**因为应用程序的开发通常赶不上英伟达的技术更新，采购具有最新技术的设备，将有利于程序的开发升级，促进校内的应用开发进展。

## 2、对单节点选择配置4卡或8卡GPU的建议，及其理由：

- 机器内GPU密度越高越好，可以减小跨节点并行带来的延时。目前测试看，就同一个问题来说，从4卡增加到8卡，计算时间几乎是减半。故**强烈推荐高计算密度的8卡方案**，甚至经费允许和未来可能的话采用16卡方案，减少主机其他部件的采购成本，以增加GPU卡采购总数。

用户应用	用户建议
生物Relion	<ul style="list-style-type: none"><li>• 建议选择NVLink的GPU平台。</li><li>• 可以选择混搭配置单节点8卡和单节点4卡，两种配置也是RELION计算平台比较常用的。</li></ul>
物化PWDFIT	<ul style="list-style-type: none"><li>• 建议购买8卡NVLink的GPU平台，可最大程度发挥NVLink性能，并节约CPU、内存、机柜空间等。</li></ul>
物化pyTorch、VASP	<ul style="list-style-type: none"><li>• 推荐8卡NVLink的平台。部分杂化泛函计算对通信需求高，可用满8卡，效率会比2个4卡节点高。</li></ul>
流体HiResX、PyFR	<ul style="list-style-type: none"><li>• 建议尽可能的购买更多的GPU，故选择8卡 &amp; PCIe4平台</li><li>• 如GPU资源充足，则采购NVLink平台可促进开发升级</li></ul>

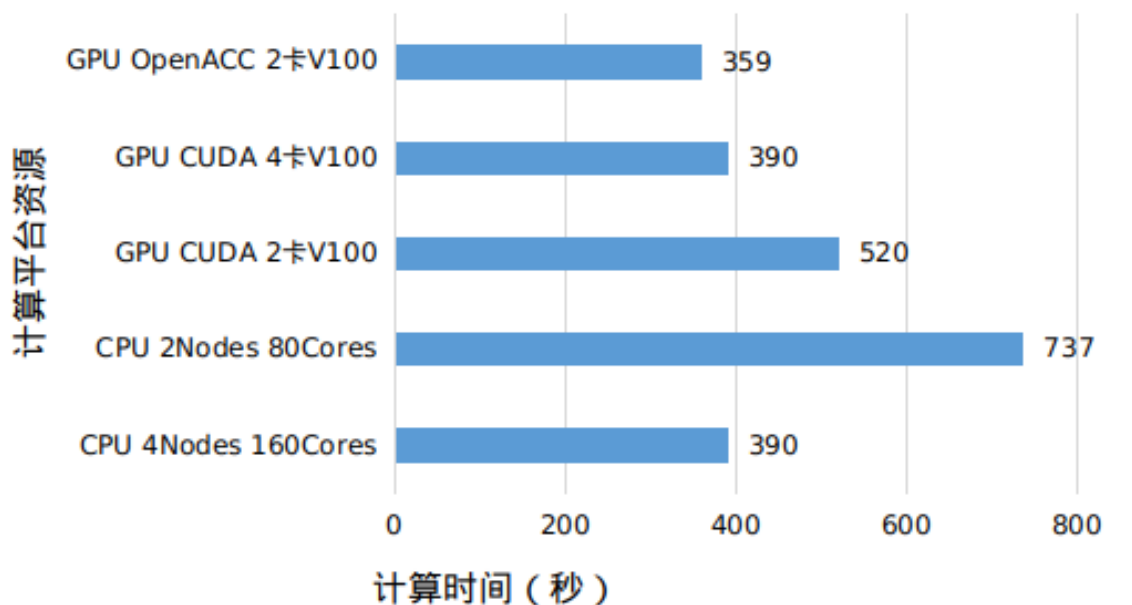
普遍建议选择8卡NVLink连接的A100 GPU节点；部分用户在少数情况下，可接受PCI-E互联的节点



## 第四部分： 总结

# 规模需求2：VASP在GPU平台的计算需求

VASP STD类型计算 CPU vs GPU 测试



测试条件：184个原子 C36 H84 S20 Br36 Bi8 体系，NSPIN = 1，4个K点，NGXYZ=140 x140 x140，开启K点并行，CPU硬件：Xeon(R) Gold 6248 CPU @ 2.50GHz。

对VASP计算占比较多的标准DFT算例的测试结果表明：

- 2卡V100性能已经高于4节点160 CPU核。
- VASP在GPU平台上的运行具有较好的性价比。
- GPU平台可以促进用户进行较大规模的VASP计算。

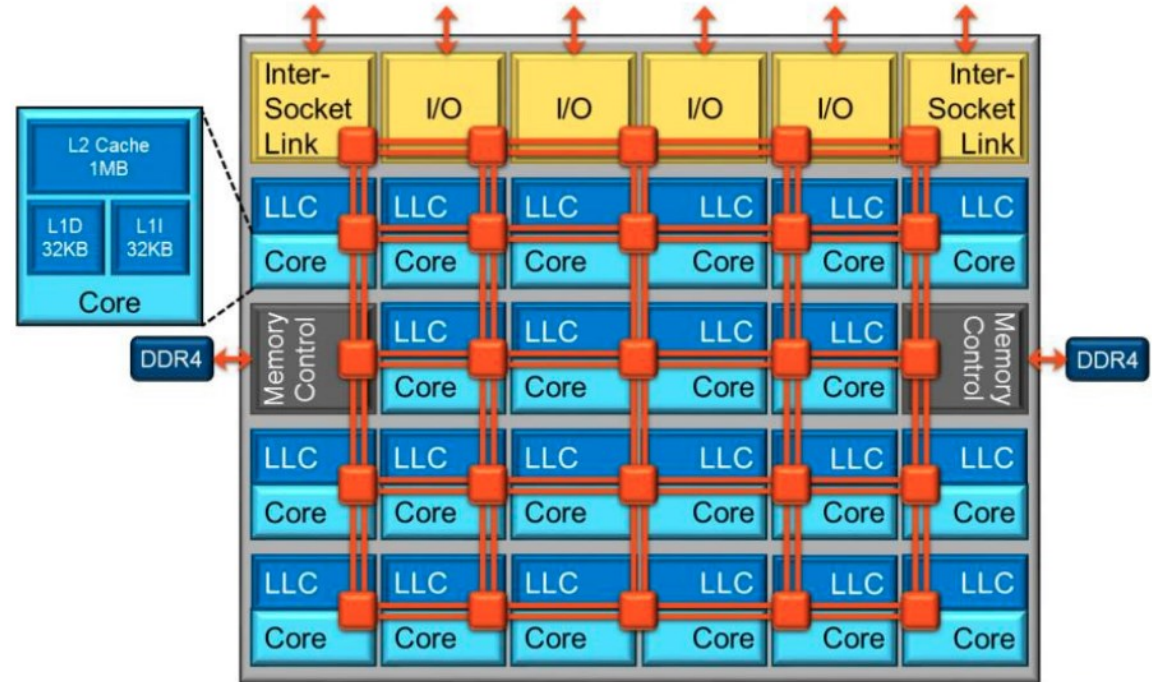
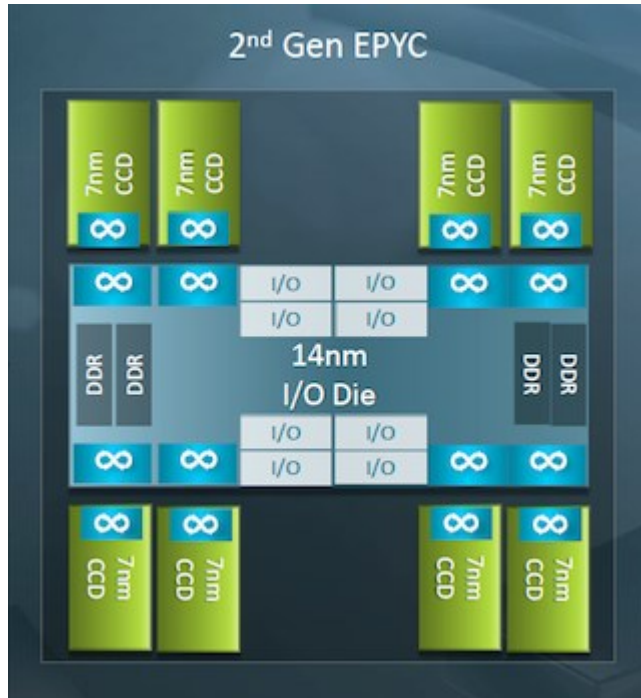
## VASP资源需求估计

- 超算中心当前VASP计算使用的CPU资源约为2万核心，考虑2卡 ~ 160 CPU核心，则相当于250张A100卡。

VASP计算在GPU下的突破性进展（性能/成本），使得几乎所有CPU作业均可以迁移到GPU集群，可保证新建集群在250卡规模左右下具备很好的利用率。

应用-算例	NVLink四卡与 PCIe4.0四卡 速度比	NVLink四卡扩展至 八卡的并行效率
生物Relion-2D算例	1.66	87%
生物Relion-3D算例	2.60	68%
物化PWDFT	1.57	70%
物化pyTorch-算例1	1.10	96%
物化pyTorch-算例2	1.02	100%
物化VASP-算例1-标准DFT	1.00	64%
物化VASP-算例2-大体系	1.55	79%
物化VASP-算例3-杂化泛函	1.00	77%
物化VASP-算例4-杂化泛函	1.17	79%
物化VASP-算例5-大体系	1.75	58%
流体HiResX	1.07	97%
流体PyFR	1.00	93%

- 相比过去，更多应用显著受益于NVLink技术，使NVLink更具性价比。
- 先进的NVLink技术可以更好的促进自有软件的研发进步
- 当前已经有较多应用算例可以高效率的扩展到8卡并行。
- 8卡可提高其他硬件（CPU、主板等）的利用效率，且降低节点间互联的需求，性价比更高。



## 选择Intel CPU的优势总结:

- NUMA 结构更简单，用户可以更容易的获得高性能；AMD CPU 所有CCD 都通过 Infinity Fabric和I/O Die 进行通信，p2p 通信效率相比Intel CPU 要低一些。
- 硬件测试表明：除L1 Cache外，Intel的延时与内存带宽都具有显著优势，且在发起GPU指令的时候具有更快的非阻塞响应性能，其影响内存<->显存的数据传输，也影响非P2P Direct模式下的卡间通讯性能。

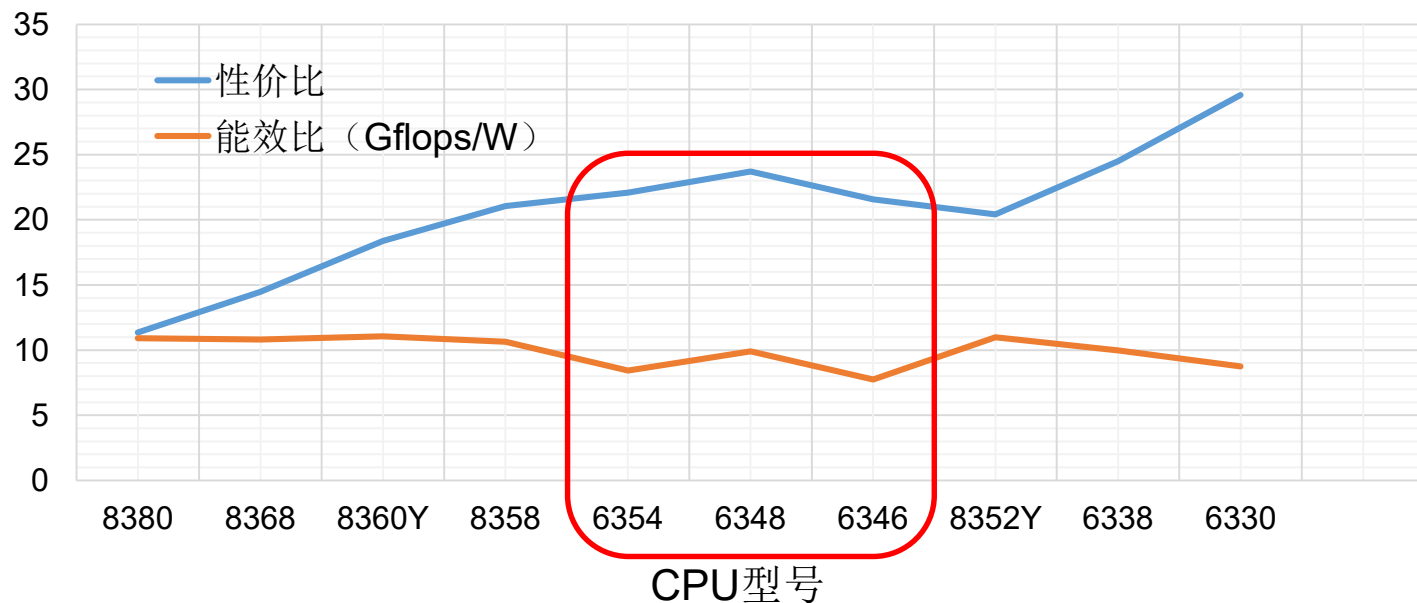
# CPU : 高主频 & 32+ Cores ( 4 Cores / GPU-Card )



中国科学技术大学  
University of Science and Technology of China

型号		功率 (W)	核心数	主频 (G)	价格 (美元)	性价比	能效比 (Gflops/W)
8380	XCC	270	40	2.3	8099	11.35942709	10.9037037
8368	XCC	270	38	2.4	6302	14.47159632	10.80888889
8360Y	XCC	250	36	2.4	4702	18.37515951	11.0592
8358	XCC	250	32	2.6	3950	21.06329114	10.6496
6354	XCC	205	18	3	2445	22.08588957	8.429268293
6348	XCC	235	28	2.6	3072	23.69791667	9.913191489
6346	XCC	205	16	3.1	2300	21.56521739	7.742439024
8352Y	XCC	205	32	2.2	3450	20.4057971	10.98926829
6338	XCC	205	32	2	2612	24.50229709	9.990243902
6330	XCC	205	28	2	1894	29.56705385	8.741463415

备注：此处只列举了4月发布的XCC型号，HCC型号预计6月以后才发布。



应用测试表明CPU核数和主频都将一定程度影响

GPU卡最佳性能的发挥：

- 高主频的CPU可以带来5%-19%的性能提升。
- 多进程共享单卡可以充分利用GPU的计算能力，当4进程共享GPU时，Relion可以达到约90%的GPU利用率。

选型规则：

- 高主频
- 32+ Cores
- 在核心数满足的情况下，优先高主频

参考范围：

**Intel IceLake 6354、6348、6346**



## • 内存

- ✓ 按照配置经验，为充分发挥硬件性能，内存通常配置为总显存的1.5~2倍。
- ✓ 总显存为 $40\text{GB} \times 8 = 320\text{GB}$ ，内存配置为512GB，按照满通道和最佳性价比配置为 $32\text{GB} \times 16$ 。

## • 高速网络

- ✓ 仅部分应用刚刚具备8卡并行能力，且资源消费能力待提高，跨节点的GPU应用需求较小，故而节点间网络可选择最高性价比的网卡硬件。

## ✓ 硬盘

- ✓ 采用无盘启动形式，配置一块1.92TB NVMe SSD盘

## • 电源

- ✓ N+N冗余电源



THANKS