



Intel® Xeon Phi™ Workshop 3-Dimensional Finite Difference (3DFD) MPI code

Victor Getmanskiy
middle software engineer
Singularis Lab, LLC

MPI launcher commands

- module load launcher/mpiexec mic
- mpirun \
-n 1 -host hostname1 -env I_MPI_PIN_DOMAIN=socket -env \
OMP_NUM_THREADS=14 ./\${binfullname} :\
-n 1 -host hostname1 -env I_MPI_PIN_DOMAIN=socket -env \
OMP_NUM_THREADS=14 ./\${binfullname} :\
-n 1 -host nknc-mic1 -env OMP_NUM_THREADS=240
./\${binfullname}.MIC :\
-n 1 -host nknc-mic2 -env OMP_NUM_THREADS=240
./\${binfullname}.MIC

MPI launcher commands

```
mpirun ${TRACE_OPTION} \  
    -n 1 -host $nknc-mic1 \  
-env LD_LIBRARY_PATH $COMPILER_ROOT/lib/mic/ \  
-env OMP_NUM_THREADS $MIC_NUM_THREADS \  
-env KMP_AFFINITY $MIC_AFFINITY \  
-env KMP_LIBRARY $MIC_LIBRARY \  
{binfullname}.MIC $n1 $n2 $n3 \  
$MIC_NUM_THREADS $nit \  
$mic_blk1 $mic_blk2 $mic_blk3 \  
$factor_speed_xphi $factor_speed_xeon : \  
    -n 1 -host $nknc-mic2
```

...

Prerequisites

1. Go to directory ~/PhiWorkshop

2. Allocate 2 mic nodes

```
salloc -N 2 --partition=mic
```

3. Substitute hostlist in scripts

```
run_1_node.sh
```

```
run_2_nodes.sh
```

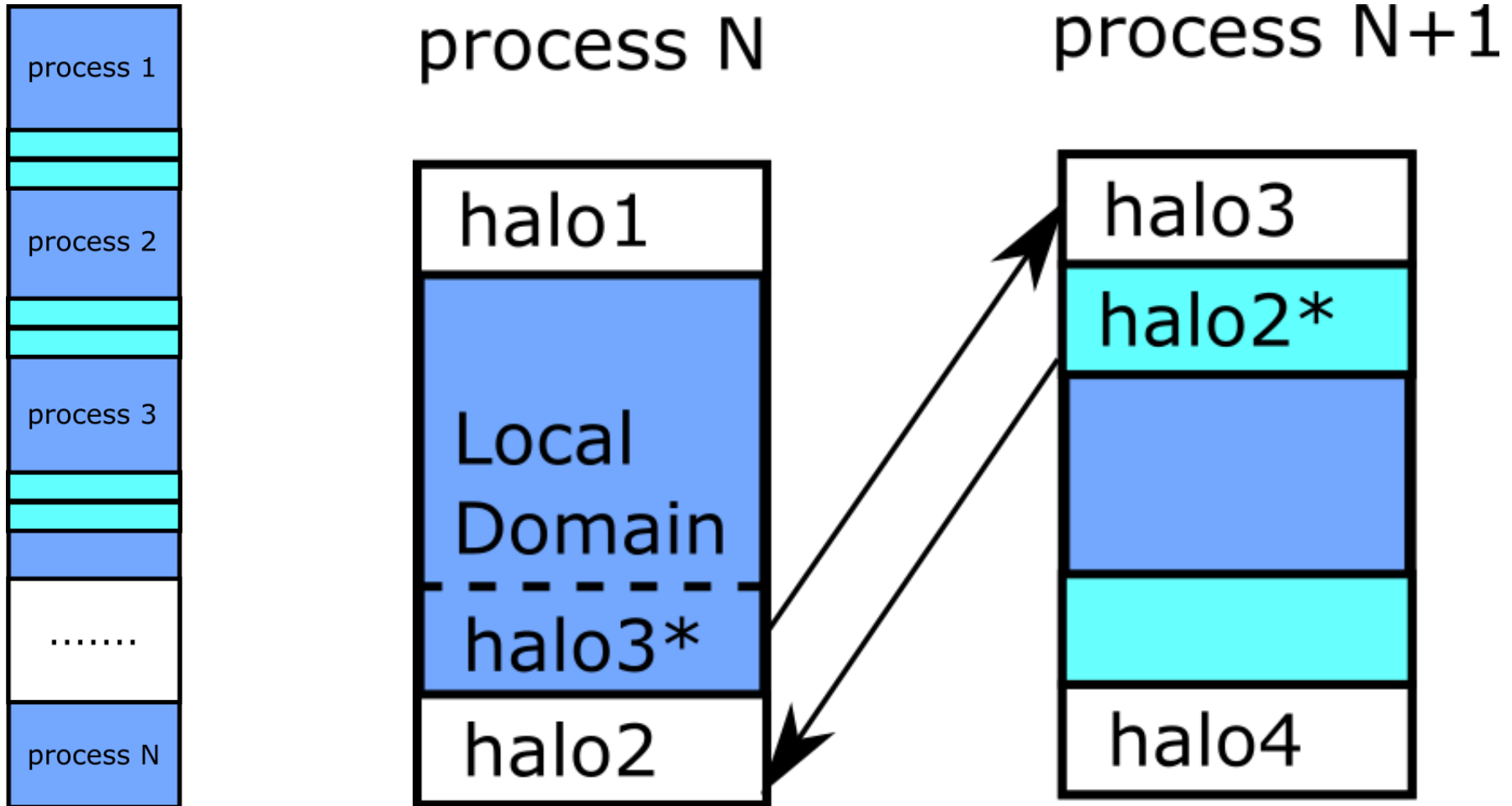
```
run_2nodes_double_size.sh
```

```
echo Executing ${binfullname}.MIC  
mpirun -hostlist "nknc-mic66,nknc-mic67" -n 2 -env LD_LIBRARY_PATH $COMPILER_ROOT/lib/mic/ -env OMP_NUM_THREADS
```

4. Load modules for direct launch of MPI application:

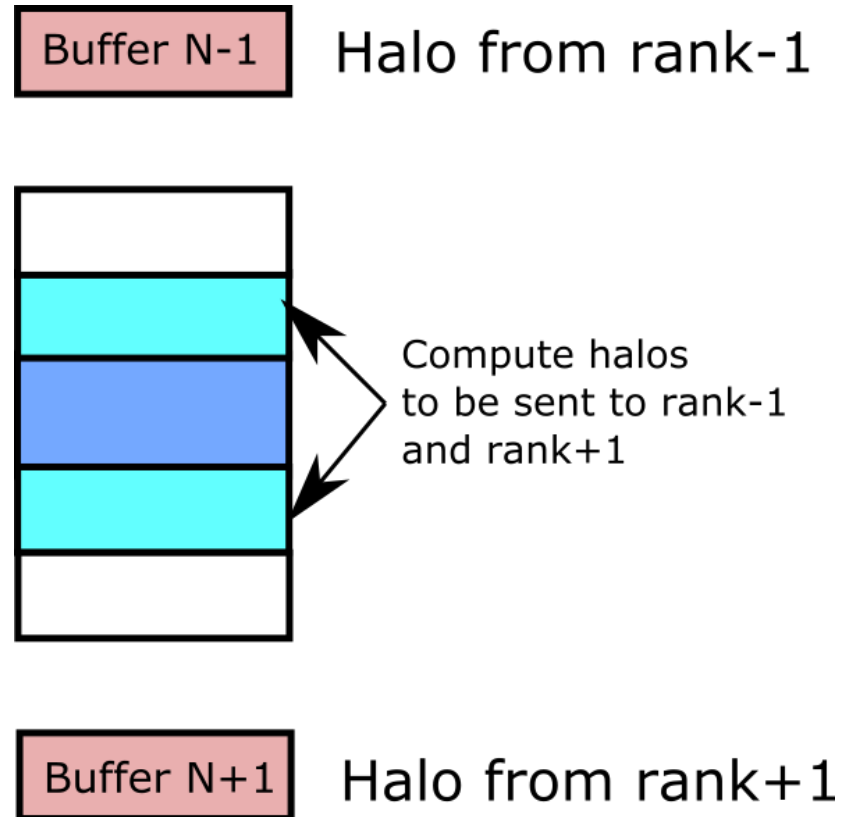
```
module load launcher/mpiexec mic
```

Subdomain configuration



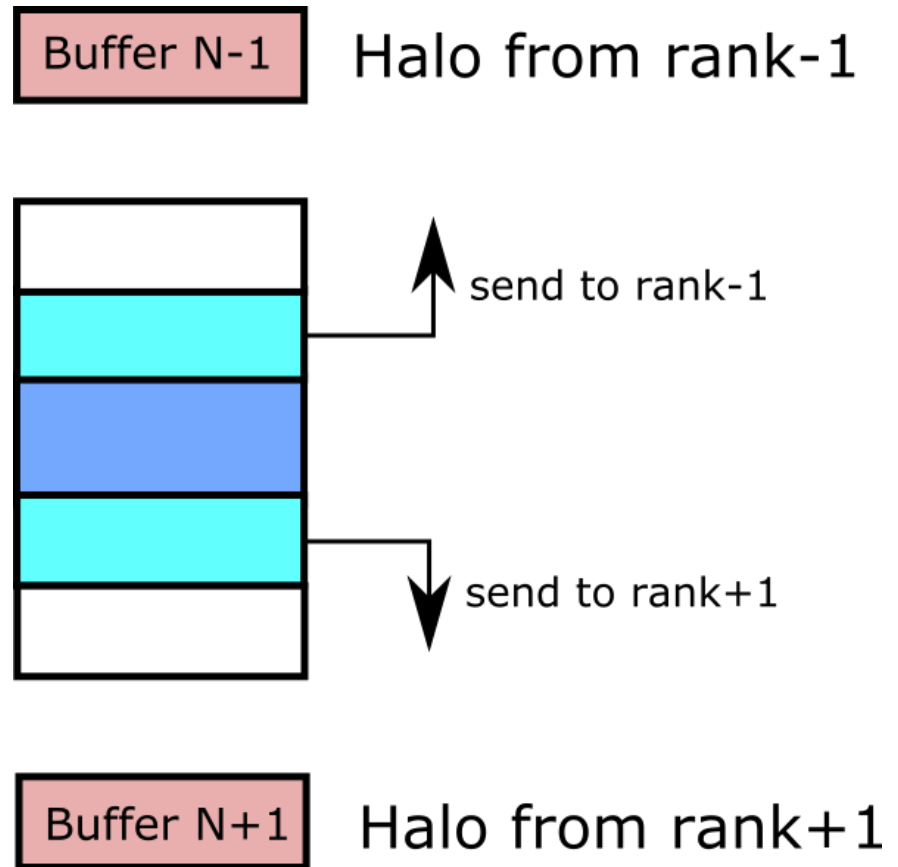
Compute inner halos

```
MPI_Irecv(BufferPrev,...,rank-1,...,&request[0]);  
MPI_Irecv(BufferPrev,...,rank+1,...,&request[1]);  
  
{compute halos to be sent to rank-1 and rank+1}
```



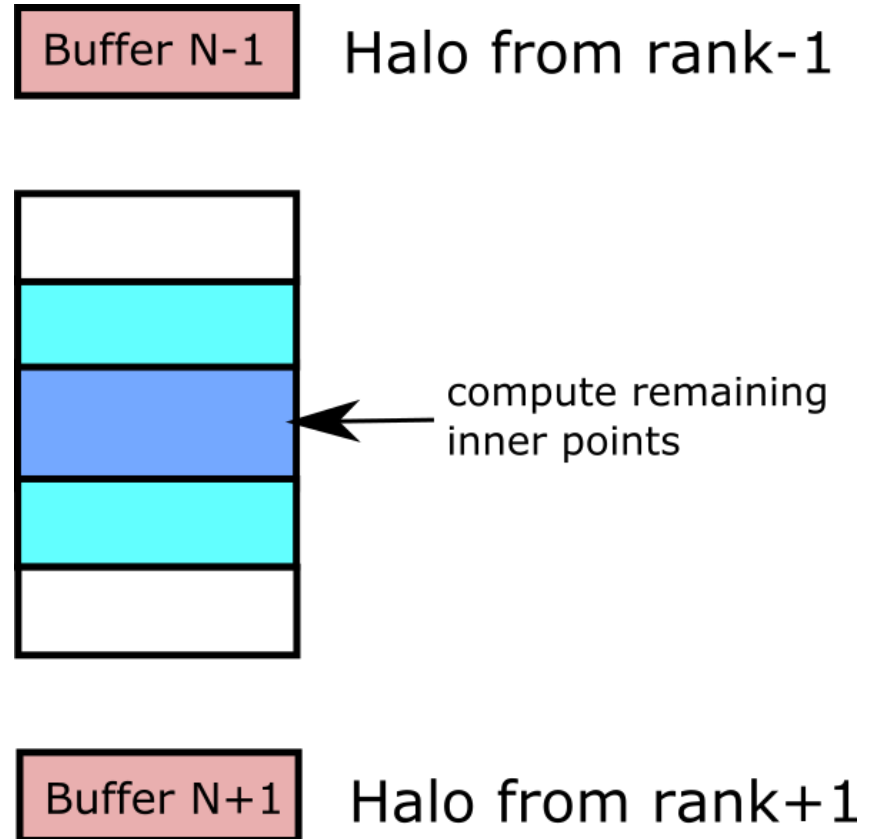
Asynchronous send computed halos to adjacent subdomains

```
MPI_Irecv(BufferPrev,...,rank-1,...,&request[0]);  
MPI_Irecv(BufferPrev,...,rank+1,...,&request[1]);  
  
{compute halos to be sent to rank-1 and rank+1}  
  
MPI_Isend(DataForPrev,...,rank-1,...,&request[2]);  
MPI_Isend(DataForNext,...,rank+1,...,&request[3]);
```



Compute inner fields independent from halos

```
MPI_Irecv(BufferPrev, ..., rank-1, ..., &request[0]);  
MPI_Irecv(BufferPrev, ..., rank+1, ..., &request[1]);  
  
{compute halos to be sent to rank-1 and rank+1}  
  
MPI_Isend(DataForPrev, ..., rank-1, ..., &request[2]);  
MPI_Isend(DataForNext, ..., rank+1, ..., &request[3]);  
  
{compute remaining inner points}
```



Replace halos computed in adjacent subdomains

```
MPI_Irecv(BufferPrev, ..., rank-1, ..., &request[0]);
MPI_Irecv(BufferPrev, ..., rank+1, ..., &request[1]);

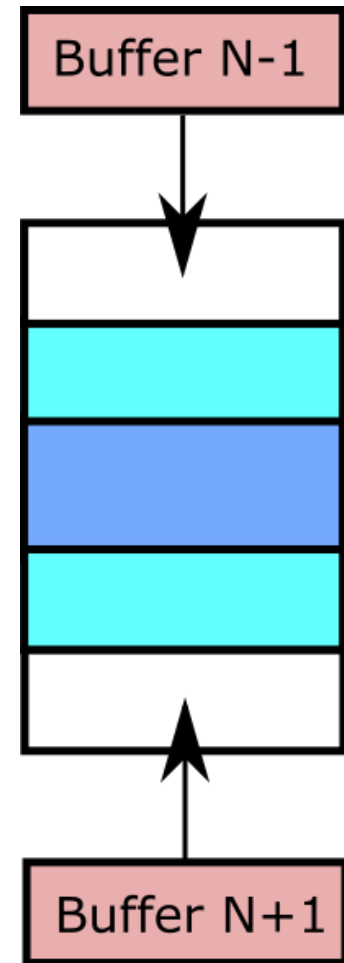
{compute halos to be sent to rank-1 and rank+1}

MPI_Isend(DataForPrev, ..., rank-1, ..., &request[2]);
MPI_Isend(DataForNext, ..., rank+1, ..., &request[3]);

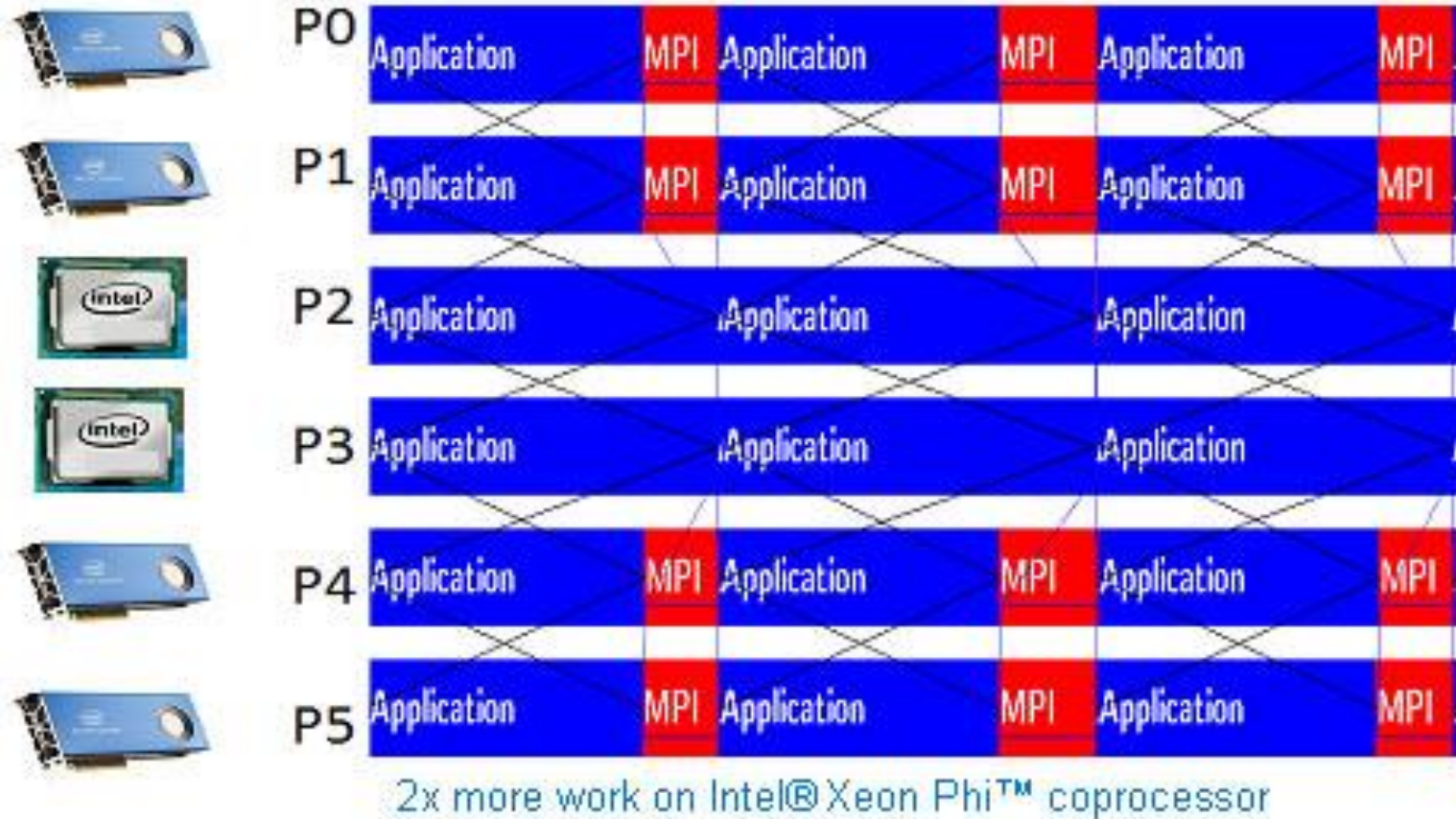
{compute remaining inner points}

MPI_Waitall(4, request, ...);

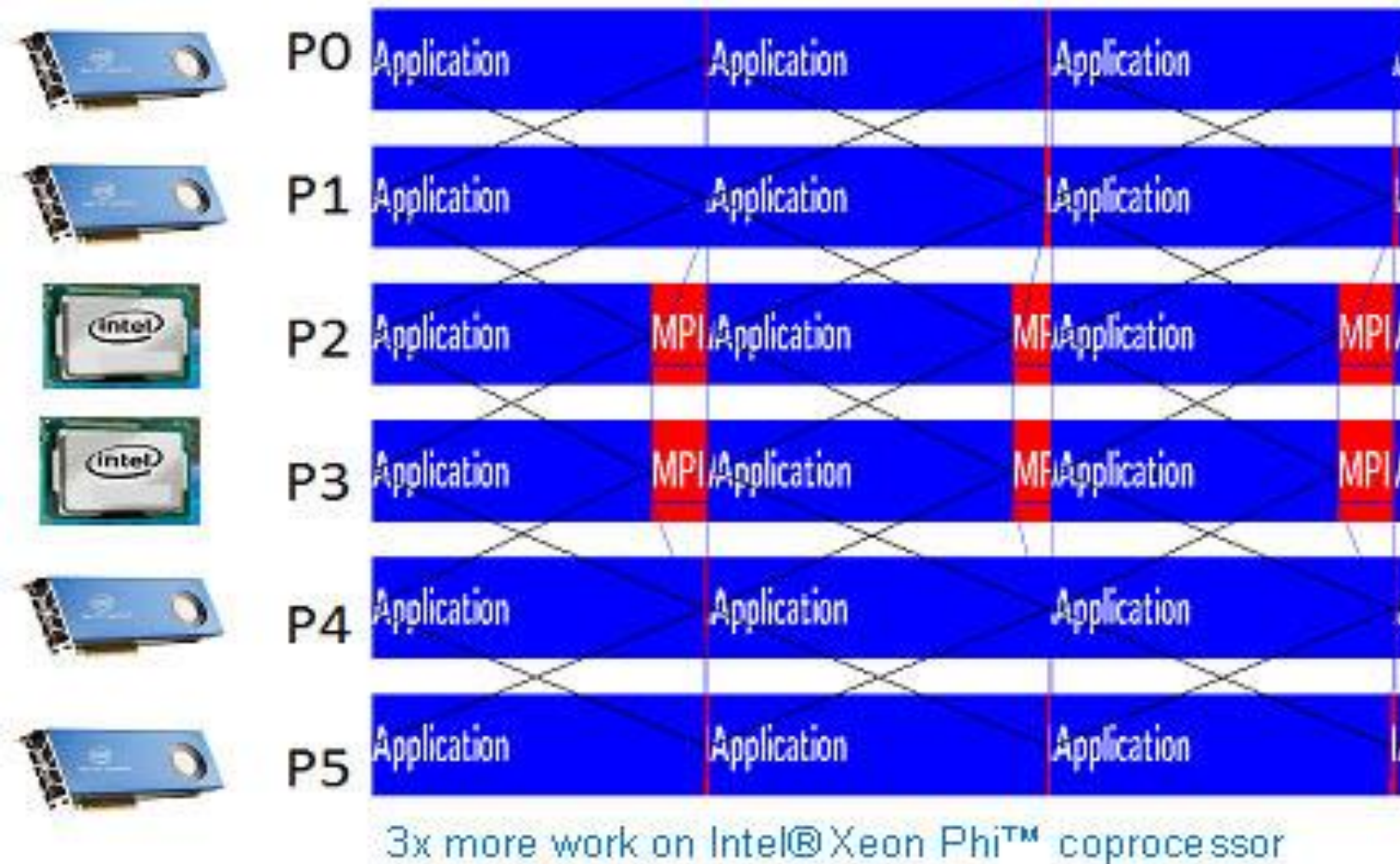
Copy Halo fom (rank-1) to Buffer from (rank-1)
Copy Halo fom (rank+1) to Buffer from (rank+1)
```



MPI trace of the first launch with ratio of 2 assumption



MPI trace of the second launch with ratio of 3 assumption



Optimal launch MPI trace with ratio of 2.6

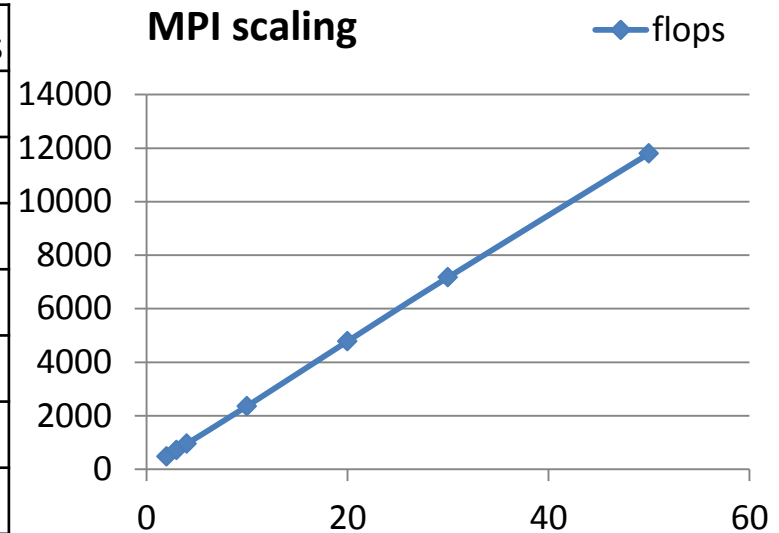


2.6x more work on Intel® Xeon Phi™ coprocessor

Scaling

Scaling is linear for stencil code

n Phi	time	mpoints	flops	points/flops
2	21,16	7752,6	473	16,39
3	21,26	11594	707	16,39
4	21,14	15562	949	16,39
10	21,33	38634	2356	16,40
20	21,03	78363	4780	16,39
30	21,04	117541	7170	16,39
50	21,32	193350,5	11794	16,39



Speedup

n Phi	Time	Mpoints	Flops	points/flops	Grid size
1	20,95	3897,04	237,72	16,39	5Gbyte
2	10,63	7680,5	468,51	16,39	5GByte
2	21,06	7792,8	475,36	16,39	10GByte

Conclusion

Task has a perfect scalability and speedup on [RSC PetaStream®](#)

References

Original work:

Finite Differences on Heterogeneous Distributed Systems

Submitted by [Leonardo B. \(Intel\)](#), [Cédric ANDREOLLI \(Intel\)](#), [Philippe T. \(Intel\)](#) on September 18, 2015 // <https://software.intel.com/en-us/articles/finite-differences-on-heterogeneous-distributed-systems>

Thanks for attention!

Victor Getmanskiy
middle software engineer
Singularis Lab, LLC

victor.getmasnkiy@singularis-lab.com

Dmitry Kryzhanovsky
CEO at Singularis Lab, LLC

dmitry.kryzhanovsky@singularis-lab.com



<https://www.singularis-lab.com/en.html>