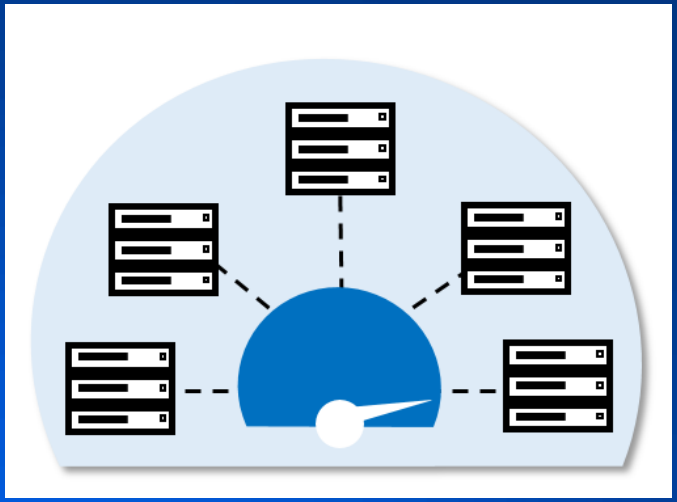
CDRH HPC Clusters



CDRH HPC Team

Stuart Barkley
Rusif Eyvazli
Nadia Guimont
Fu-Jyh Luo
Mike Mikailov
Nicholas Petrick

HPC Computational Environments

- CDRH Betsy/Bluefin Environments
  - General purpose HPCs to support a wide array of projects
    - Web-based bioinformatics analysis (Galaxy)
    - Artificial intelligence/machine learning
    - Genomics, next-generation sequence analysis and alignment,
    - Modeling and simulation
    - Statistical analysis and more

Betsy Cluster

Windows/Mac/
Linux workstation

FDA internal network

Login node
(betsy.fda.gov)

Grid engine

7 "Bigbox" nodes
40 cores each
256 GB RAM

74 "Betsy" nodes
8 cores each
24/96 GB RAM

4 "SuperMicro Monster" nodes
40 cores each
2 TB RAM per node

44 "SuperMicro Standard" nodes
24 cores each
512 GB RAM per node

"GPU Cluster" 32 nodes
1 GPU, 2 CPUs,
24 cores, 512 GB RAM
per node

10 Gb Ethernet

40 Gb Infiniband

100 Gb Intel OPA

100 Gb IB

201 "Blue Meadow" nodes
8 cores each
24 GB RAM

1.2 PB
DDN GPFS

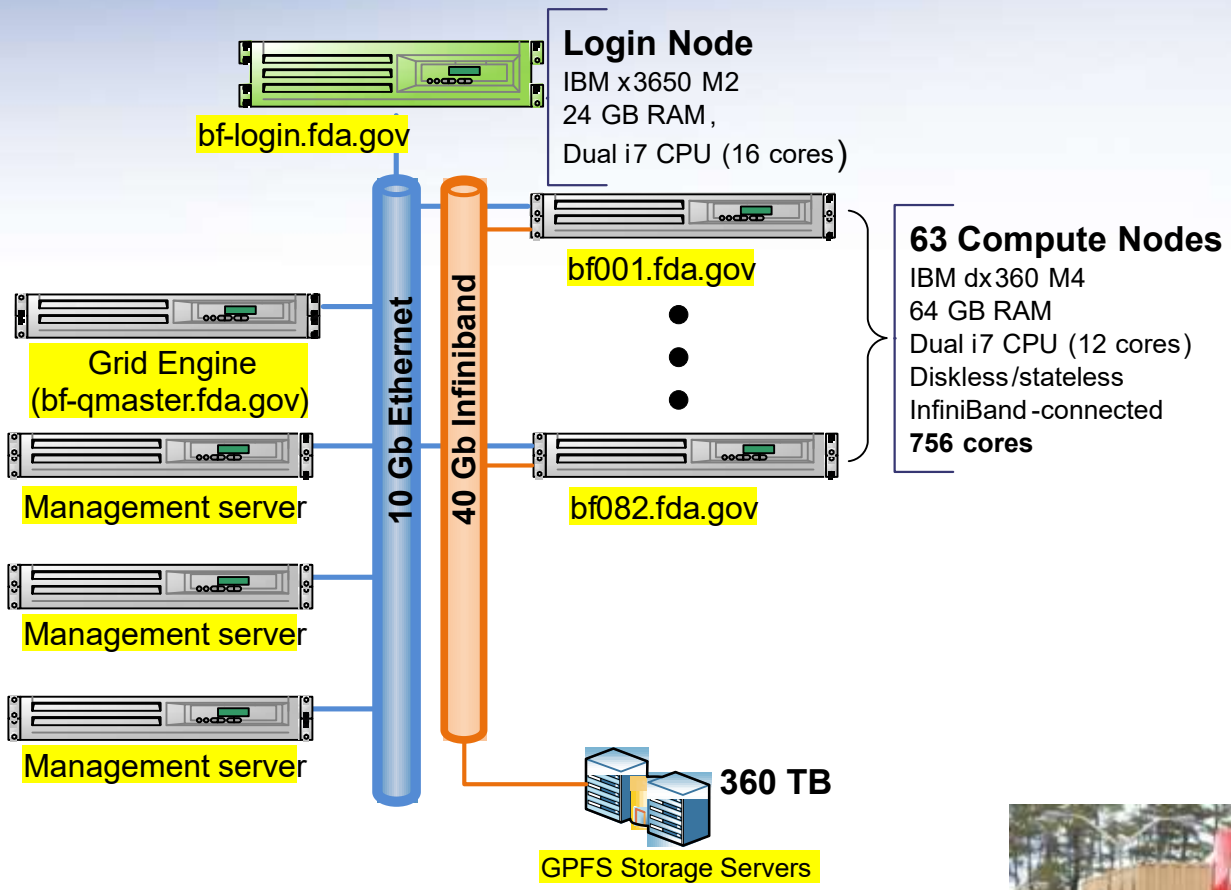
Current Production Storage

1.5 PB
DDN GPFS

Additional Capacity Storage

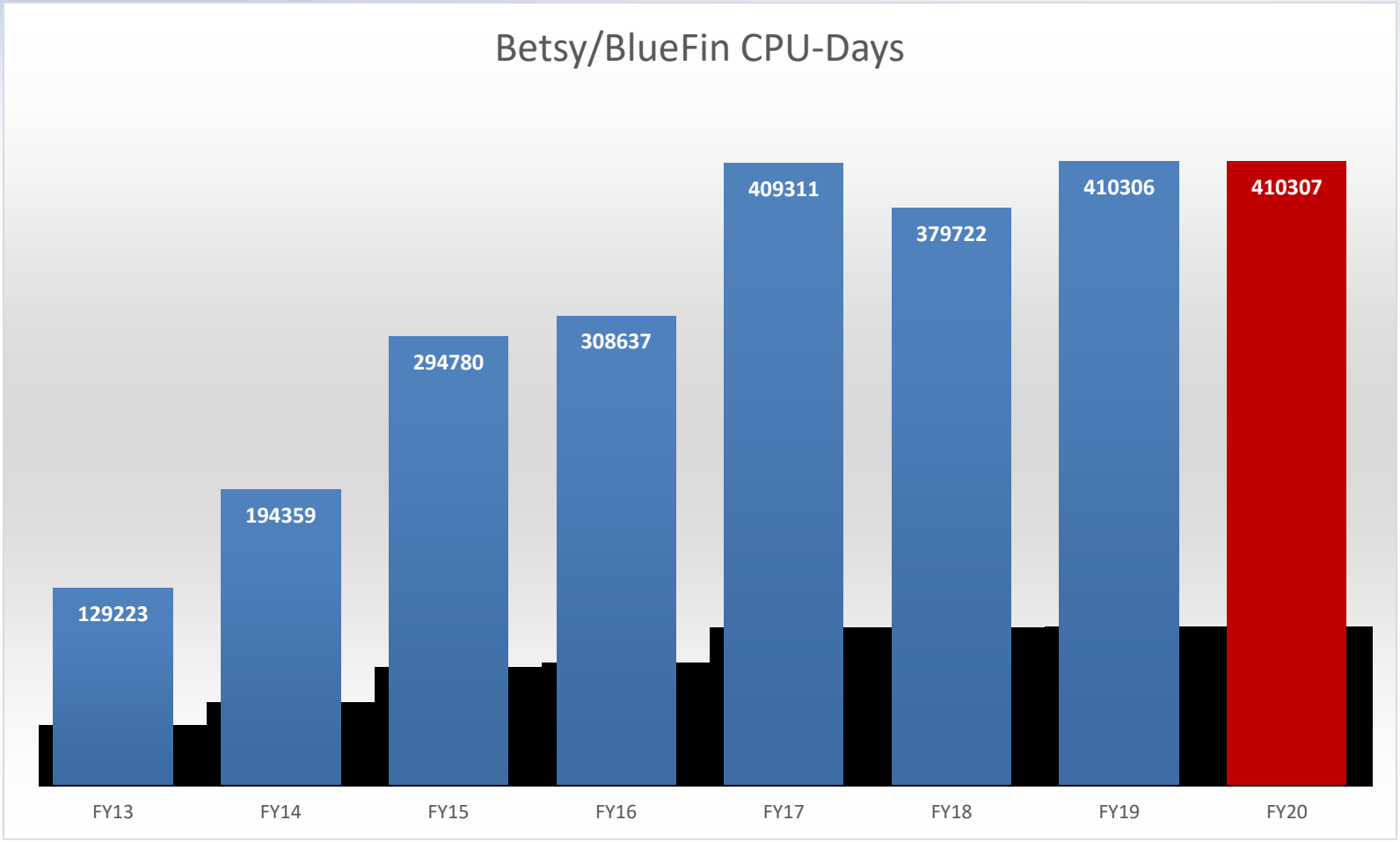
Active Nodes	Nodes	Core/node	GPU/node	GB/node
Bigbox	7	40	0	256
Besty	74	8	0	24
SM "Monster"	4	40	0	2000
SM Standard	44	24	0	512
Blue Meadow	201	8	0	24
GPU Cluster	32	24	1	512
Summed Totals	362	4,464	32	55,304

Bluefin Cluster



- Bluefin is a “mobile” computing platform
  - Currently located at CVM facility in Beltsville

Betsy/Bluefin Utilization



*FY20 utilization is 10/2019-6/2024 only

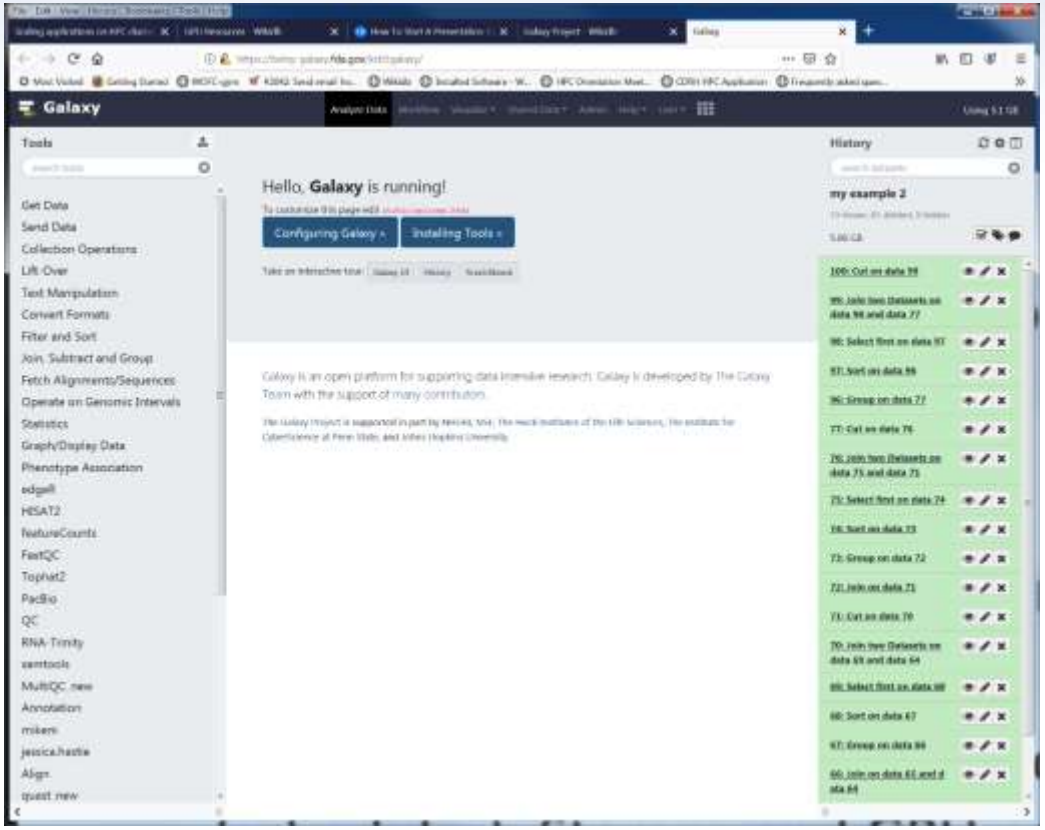
Scientific Applications

300 FDA-approved applications:

- **Computational Chemistry:** Amber, Autodock, Desmond, LAMMPS, …
- **Computational Fluid Dynamics:** OpenFoam
- **Image Analysis:** CellProfiler, Freesurfer, FSL, ImageJ …
- **Linkage/Phylogenetics:** CD-HIT, GARLI, Mothur, PHYLOSHOP, POY, QIIME …
- **Mass Spectrometry:** TPP (PeptideProphet, ProteinProphet, ASAPRatio, …)
- **Mathematical/Statistics/Modeling & Simulation:** R, SAS, Chaste, FluTE, GridMathematica, …
- **Next-Generation Sequencing:** ABySS, ALLPATHS-LG, Bcl2fastq, Bedtools, Bowtie, Bowtie2, Bwa, CCMpred, Cufflinks, GATK, GNUMAP, HIVE, Kraken, Mira, Picard, Samtools, Tophat
- **Sequence Analysis & Alignment:** Blast, Blat, Cortex, Dfam, Exonerate, Geneious, GotoBLAS2, HMMER, Jalview, JELLYFISH, Mauve, mpiBLAST, Mugsy, MUMmer, MUSCLE, NovoAlign, ParSNP, Pfam, PRINSEQ, ScalaBLAST, SGA Assembler, SNP Pipeline, SPAdes, SUPERFAMILY, Tablet
- **Multi-paradigm environments**: Parallel MATLAB, Octave, R, Galaxy, …

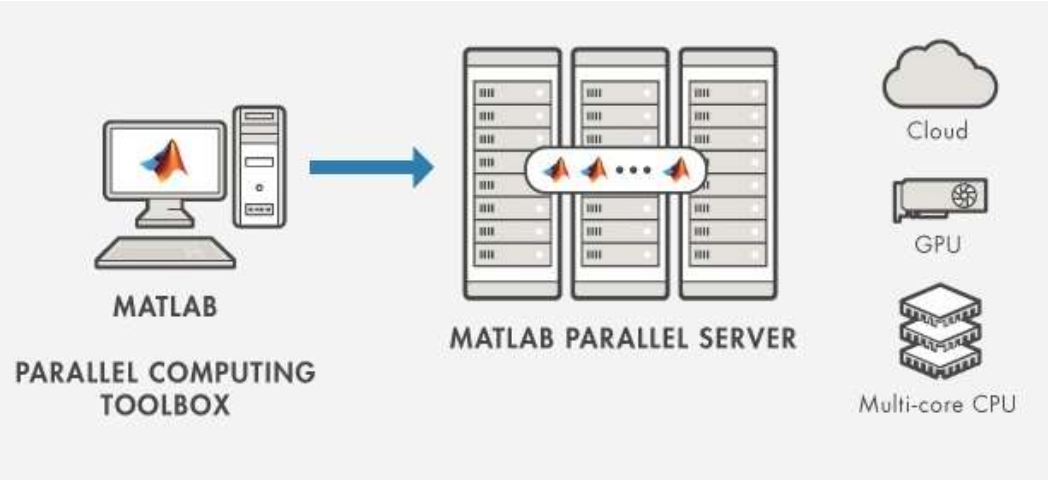
Galaxy Web Annalistic Platform

- Open-source platform
  - Enables researchers without informatics expertise to perform computational analyses through web
    - Users upload data and define/run analysis pipelines
  - Galaxy Tutorials
    - Galaxy 101. How to create workflows.
    - Uploading data - How to get data into Galaxy.
    - Learn Galaxy



MATLAB Parallel Server

- MATLAB Parallel Server
  - Lets users scale MATLAB® and Simulink® programs to HPC cluster
  - Runs programs/ simulations as scheduled applications on cluster
  - Desktop license profile dynamically enabled on cluster, so no need to supply MATLAB licenses for cluster



HPC Application Scaling

Mike Mikailov
CDRH/OSEL/DIDSR

Process Scaling

Before parallelization
Single run

```
max=2000
for ( i in 1:max)
{
   [computations]
}

[summarization]

[final results]
```

After parallelization
Many parallel runs

```
i=1
[computations]

[partial results-1]
```

...

```
i=2000
[computations]

[partial results-2000]
```

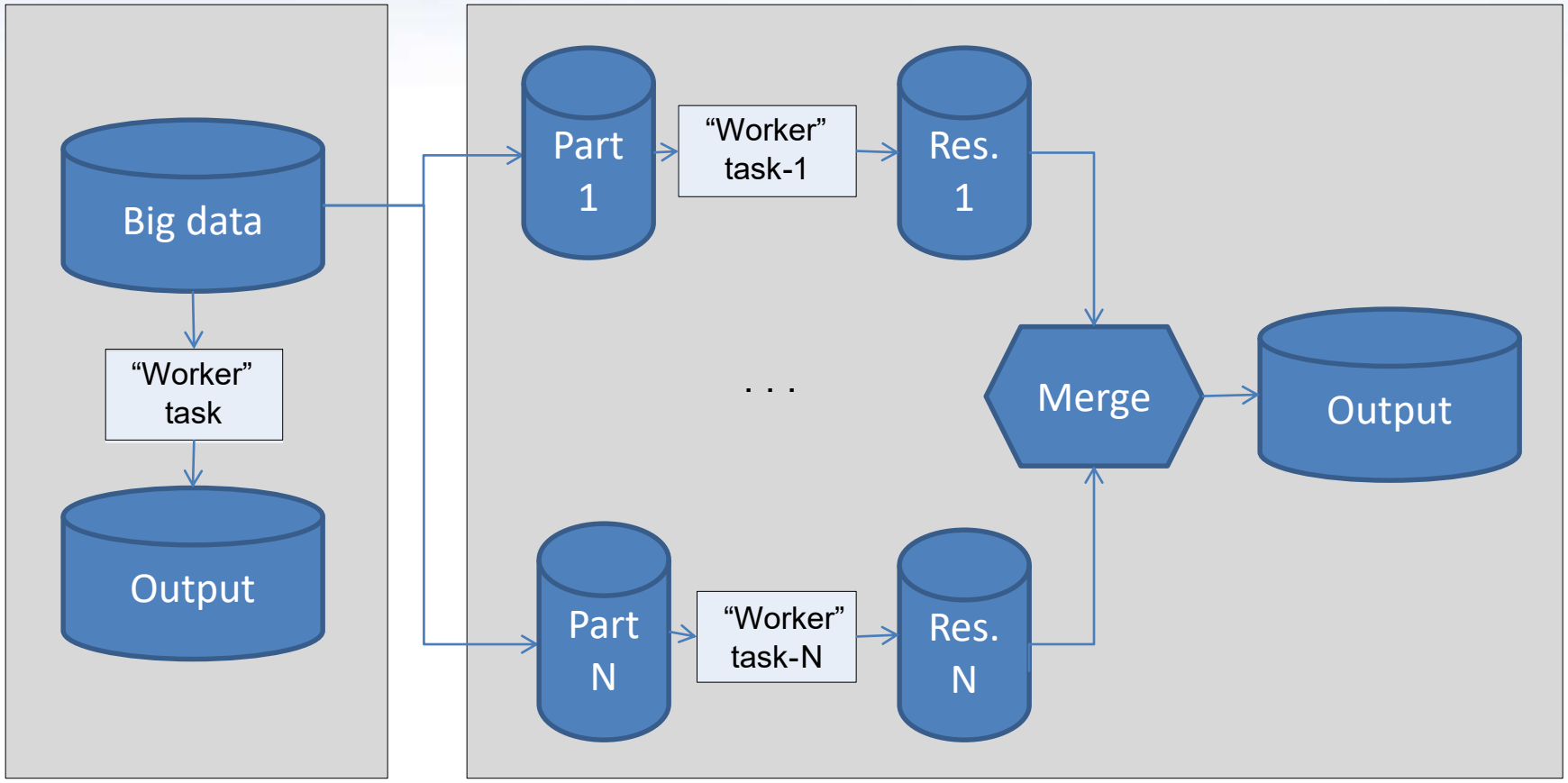
Concatenation, summarization
within minutes

```
[partial results
concatenation]

[summarization]

[final results]
```

Data Scaling

Before scaling:
Time needed T

After scaling: Time needed ~T/N



Scaling Techniques

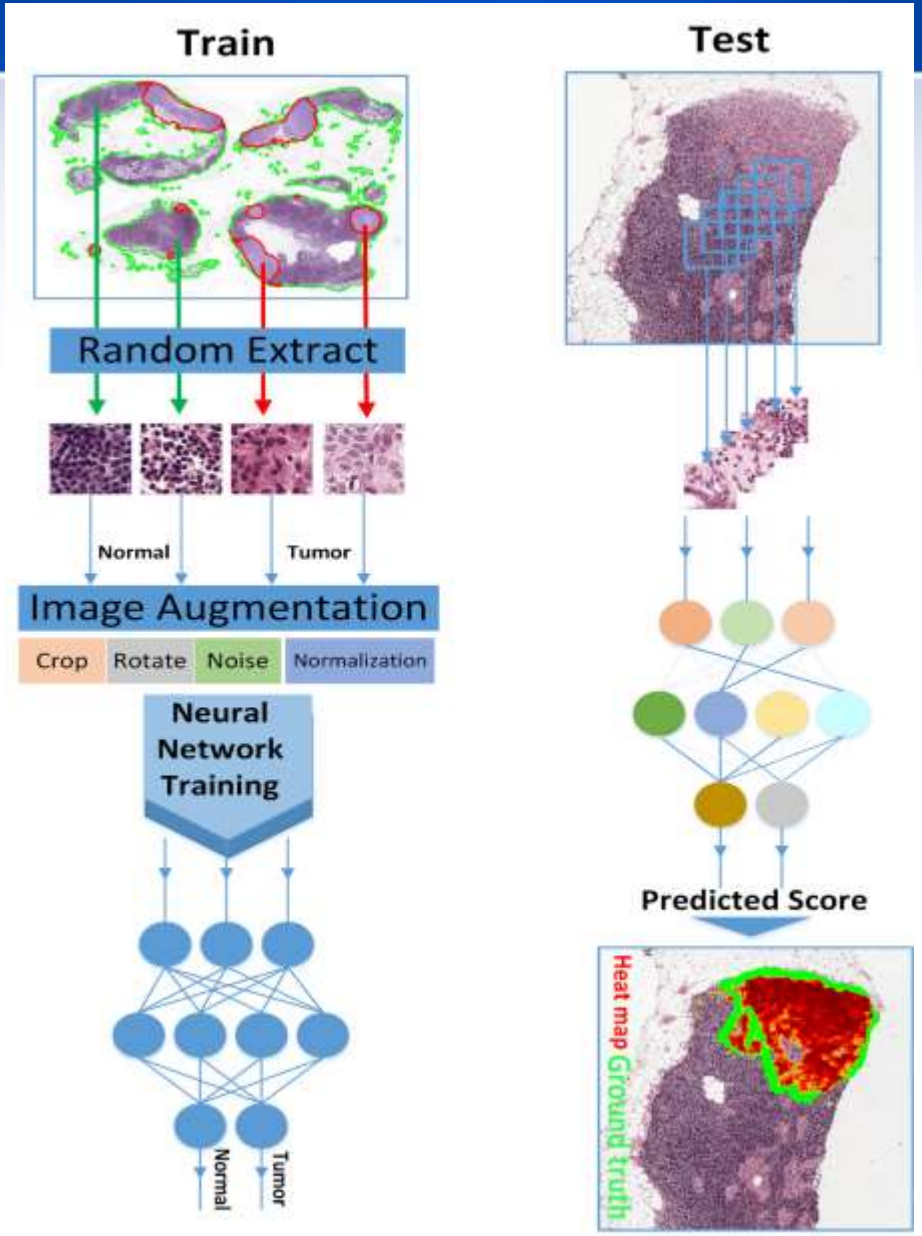
Scaling Technique	Advantages	Disadvantages
Multi-threading, OpenMP	• Multiple parallel threads within a node	• Scaling is limited to cores on one node
MPI	• Multiple parallel threads across one or more nodes	• Overhead for I/O coordination and load balancing • In practice, all requested resources must be available to start • No checkpointing • Cannot exceed max capacity of the cluster
Scientific workflows, MapReduce, Spark, Hadoop	• Scalable computational or data manipulation tasks on one or more nodes	• Does not offer integrated approach for scaling multi-level nested loops or random number generation
Single loop parallelization	• Multiple parallel tasks on one or more nodes	• Does not parallelize multilevel nested loops
Array-based parallelization	• Multiple parallel tasks across one or more nodes	• Overhead for setup/convergence phases

HPC DLNN Project

Weizhe Li, Weijie Chen, Mike Mikailov
CDRH/OSEL/DIDSR

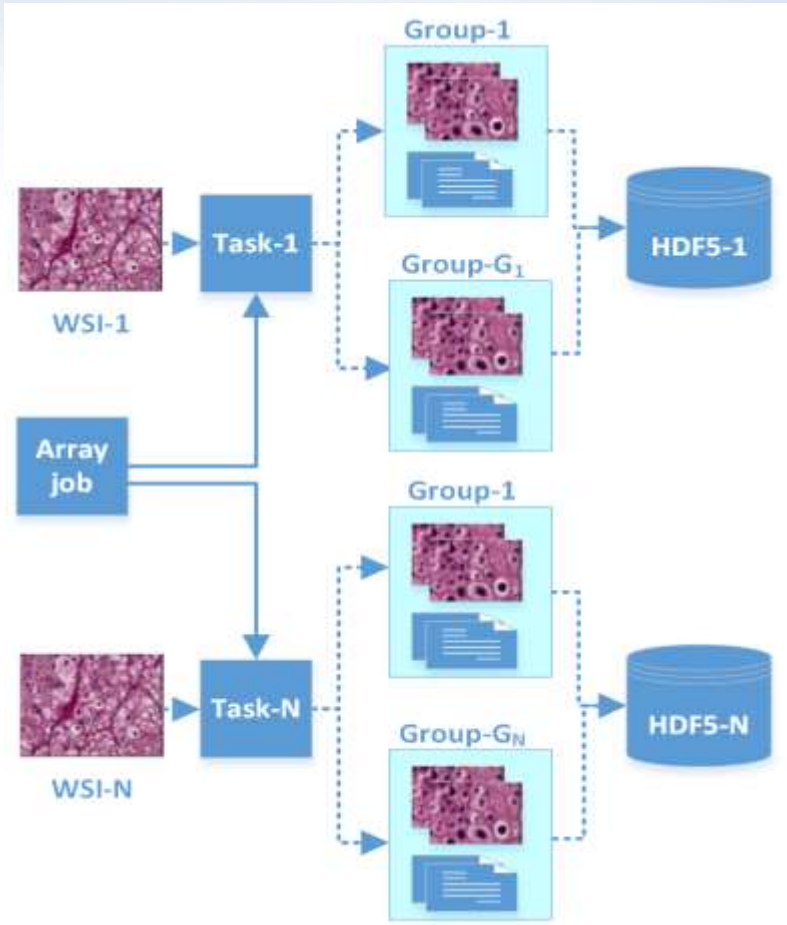
DLNN AI/ML with WSI data

- DLNN pipeline for digital pathology
- Training
  - Normal/tumor image patches randomly extracted
    - Normal (green)
    - Tumor (red)
  - Patches used to optimize NN using GPUs
  - HPC implementation generates pixel-wise heatmap via a sliding window



DLNN AI/ML with WSI data

- Scaling DDLN testing on HPC
  - Reformat/group patches into HDF5 file format for improved parallel I/O
  - Job splitting/scaling for parallel HPC implementations



DLNN AI/ML with WSI data