Recent Update

- Watson Machine Learning Community Edition (formerly PowerAI)
  - module load wmlce/1.6.1-py3.6
  - module load wmlce/1.6.1-py2.7
  - Tensorflow-v1.14
  - PyTorch-v1.0.1
  - RAPIDS cuML and cuDF upgrade to v0.7.2
Recent Update

• Open XDMoD 8.1
  • Open XDMoD is an open source tool to facilitate the management of high performance computing resources.
  • Cluster Usage Statics
  • User Top Chart
  • http://hal-xdmod.ncsa.illinois.edu:8080/
Recent Update

• HAL Slurm Wrapper Suite
  • Early release sw-v0.1
  • “swrun”
  • “swbatch”

• New partitions include:
  • gpux1 / gpux2 / gpux4 : single node job
  • gpux8 / gpux12 / gpux16 : multiple nodes job
  • cpu : multiple nodes job
  • gpu-debug : hidden queue for single node debug job

• Old partition “debug”, “solo”, “batch” will be deprecated.
Recent Update

• “swrun” Usage
  • Only 3 options
    • Partition
    • CPUs Per GPU
    • Wall Time
  
• Restrictions
  • Partitions vary by GPU number (x1, x2, x3, …)
  • CPU Per GPU (12 <= c <= 36, default 12)
  • Wall Time (1 <= t <= 72, default 24 hours)

• Default
  • gpux1(required), 12x CPUs, 18GB Memory, 1x GPU, 24 Hrs

• Full Node Example
  • swrun -q gpux4 -c 36 -t 72
Recent Update

• “swrun” Usage
  • Debug queue
    • Only need to set time to 4 hours or less on queue “gpux1/2/3/4” and “cpu”
    • Multi-nodes queue less than 4 hours still go to normal queue
Recent Update

• “swbatch” Usage
  • Only 6 options
    • Partition
    • CPUs Per GPU
    • Wall Time
    • Job name
    • Output file
    • Error file

• Restrictions
  • Partitions vary by GPU number (x1, x2, x3, …)
  • CPU Per GPU (12 <= c <= 36, default 12)
  • Wall Time (1 <= t <= 72, default 24 hours)

• Run “swbatch demo.sb”
THANK YOU FOR YOUR TIME!