

GPU support in MetaCentrum

Miroslav Ruda

CESNET

April, 2013



Two GPU clusters in Czech grid

- nodes with 2xNVIDIA GeForce GTX 465, 4xTesla M2090
- third cluster based on Kepler K20 scheduled this year
- national grid based on Torque, nodes not visible in EGI
- used by user-developed applications, Matlab, tools from computational chemistry, . . .

Torque

- large Torque modifications (not related to GPU)
 - scheduling, various types of resources, distributed setup
 - <http://www.metacentrum.cz/en/devel/torque/>
- GPU resource defined for nodes, usage similar to CPU
 - -lnodes=1:gpu=2
 - handled by standard scheduler+server logic
 - type of GPU card as regular node property

Modifications needed on MOM - granting access to users

- three possible solutions discussed
 - set compute-exclusive mode
 - fails for users accessing card from two processes
 - in prologue/epilogue set access right to `/dev/nvidia[X]`
 - easy, elegant, no changes to code
 - problems with more than one job of the same user (ordering of cards can change during the job)
 - set `CUDA_VISIBLE_DEVICES`
 - cannot be done in prologue, MOM patch
 - user can overwrite it
 - no interference between two jobs of the same user
 - currently used in production
- dedicated queue for jobs requiring GPU cards
 - better priority on GPU nodes
 - no(t-yet) control of real GPU usage