

CORAL Tools Breakout Group Notes

What will Coral be?

- Oak Ridge and LLNL: IBM Power, Nvidia Volta, Mellanox IB interconnect
 - Fewer nodes than Titan
 - Power9 + multiple GPUs per node
 - Shared memory between GPU and CPU in these systems
- Argonne Cray Knights Landing (Theta) then Knights Hill (Aurora)
- These systems will have NVRAM
 - Tools would like some space in NVRAM

Coral Tools Working Group

- Focused on IBM system at present
 - Lots of NDA material at present, but becoming more open over time
 - Challenge: open up the NDA window enough to get the right people involved
- CORAL points of contact for mailing list
 - Michael Brim for IBM system
 - Scott Parker Intel system?
 - High-level Intel contacts for the tools community?
 - Perhaps Tony Luck or Bob Wisniewski

Programming model for IBM systems

- Using OpenACC now
- Plan to move to OpenMP 4 with targets as is possible.
- IBM OpenMP runtime + LLVM backend from nvidia.
 - See A. E. EICHENBERGER AND K. O'BRIEN. Experimenting with low-overhead OpenMP runtime on IBM Blue Gene/Q . IBM J. RES. & DEV. VOL. 57 NO. 1/2 PAPER 8 JANUARY/MARCH 2013
 - LOMP is substantially better than other OpenMP runtimes.
 - There is needed evolution for GPUs.

Performance tool interfaces for IBM Summit and Sierra platforms

- NVIDIA is providing the CUPTI interface.
 - Activity API is the new asynchronous API
 - newer versions won't deadlock when used from signal handlers
- OMPT and OMP Target will be wrappers around the Activity API.

Supporting CORAL systems with open source tools will be a challenge

- IBM will deliver APIs with input and guidance from tools working group
- IBM not interested in delivering higher level tools.
- If monitoring tools are going to be OpenSource then there will need to be funding to develop those capabilities and to setup a timeline.
 - Measurement challenges when nodes are working at 40TF and so we will need to do data reduction on the fly.
 - Today's data collection may not be up to the speed of these nodes.
- How to extract insight from this sea of information

- detect outliers
- detect insufficient use of parallelism
- Thus labs will have to find operational money to make these tools.

Things that performance tools will need

- Monitor computation on the cards.
- Data movement within and across nodes. nvlLink and nest. Nest is uncore on Power. We have been told that there will be something for nvlLink. Also OmniPath on intel.
- Information about interconnect performance.

Debugging: OMPD

- OMPD is currently being refined in a collaboration between Livermore and Rogue Wave.
- Need to get Alinea on board for DDT.
- Problem: currently no support for OMP target.

Test environments?

- Current test environment HW doesn't have stable SW stack.
- Testing develop platform should be available in late 2016 - early 2017 so that the tools developers can build stuff.
 - Test system at OLCF late 2016 will have NVIDIA Pascal GPU (with sampling measurement)
 - Livermore also will have one.

Need data for current which are NDA but are needed for OpenSource

- Platform Analysis Library
- Trace library

Power monitoring and adaptation?

- There will also be power monitoring APIs written into the contract as well.
 - There will be coarse grained power monitoring on a system level
- Should there be power consumption requirements for users?
- It would be useful to have power aware schedulers but they will not likely be available in time for these systems.
- Advanced runtimes being evolved for exascale will probably need to have introspection hooks.

Performance Data Analysis

- Analyzing performance data from Summit and Sierra will be data and computation intensive
 - It may be useful to employ big-data approaches for performance analysis
 - use hadoop or similar?
- Will need a cluster for high level analysis and viz of performance data from CORAL systems
 - No current plans for dedicated systems.
 - The system might be a good analysis system for itself.

- Can the system be used for applications and data analysis concurrently?
- The main users will also need to allocate time and support for analysis of performance data

Need regression tests for Summit/Sierra Software stack

- BG/Q experience
 - PAMI still has data races that cause trouble with sampling based tools
 - XL OpenMP has data races as well that cause trouble with sampling based tools
- Suggestion: need regression tests for system acceptance that exercise software stack with sampling

Summary

- There are both process issues and technical issues that need to be resolved
- We won't be able recognize all the technical issues until we get some of the process issues resolved
- Regression tests and test plans into the Tools Working group for Coral are important.
- Open source tools won't magically appear for CORAL systems without investment in their development