

DDT Training at Keeneland

David Lecomber david@allinea.com





- Introduction to Allinea and DDT
- Getting started with DDT on Keeneland
- A worked example
- Getting help



Allinea Software

- Software tools company since 2001
 - Allinea DDT the scalable parallel debugger
 - Allinea OPT the optimization tool for MPI and non-MPI
 - Users at all scales at 1 to 100,000 cores and above
 - World's only Petascale debugger!

Simplifying the challenge of multi- and many-core development

- Bugs at scale need a debugger at scale
 - ... until recently debuggers limited to ~4,000-8,000 cores
- Bugs on GPUs need a debugger for GPUs
 - ... until recently GPU software couldn't be debugged

alinea Some clients and partners









































Climate and Weather















Energy











Electronic Design Automation







Academic

Over 200 universities



Collaborations



Partnership to develop Petascale debugger with NVIDIA support



Partnership to develop Petascale/ Exascale tools and standards



Partnership on Full Scale debugging on IBM Blue Gene /P & /Q



Allinea DDT is "*Debugger of Choice*" on NERSC 5 and NERSC 6 and first implementation on CRAY XE6



Partnership with CEA French Atomic Energy Authority on scalable programming and CUDA



Partnership on Keeneland project to help solving software challenges introduced by mixed architectures



Debugging kernels

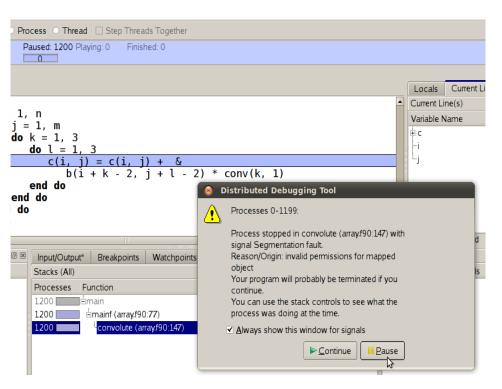
Set breakpoints

- Automatically stop on kernel launch
- Stop at a line of CUDA code
- Hover the mouse for more information
- Step a warp
 - Follow the logic of individual threads through the kernel
- ... or just run through to a crash

```
1
  Threads:
                           0 Thread 15 15
  CUDA Threads
                                                                                Grid size: 3
Project Files
                        edge.cu 🔀
                           18 \overline{\{1, 2, 1\}}
Project Files
                           19 {0, 0, 0},
 Source Tree
                           20 {-1,-2,-1}
 Header Files
                           21 };
 Source Files
                           23 /// 2D convolution filter using global memory
                               _global__ void conv2d_global(uchar4* in, uchar4* out, int
                                  int x = threadIdx.x + blockIdx.x * BLOCK_SIZE;
                                  int y = threadIdx.y + blockIdx.y * BLOCK SIZE;
                                  int index = (y * width + x);
                                  float4 output = make float4(0.0f,0.0f,0.0f,in[index].w
                                  #pragma unroll
                                  for(int cy=-1; cy<=1; ++cy)</pre>
                                      #pragma unroll
                                      for(int cx=-1; cx<=1; ++cx)</pre>
```



Handling regular bugs

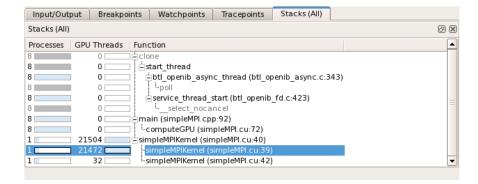


- Immediate stop on crash
 - Segmentation fault, or other memory problems
 - Abort, exit, error handlers
 - CUDA errors
- Scalable handling of error messages
- Leaps to the problem
 - Source code highlighted
 - Affected processes shown
 - Process stacks displayed clearly in parallel



Simplifying kernel scale

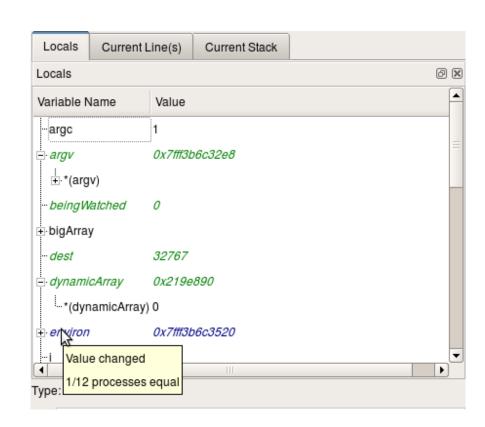
- View all threads in parallel stack view
 - At one glance, see all GPU and CPU threads together
 - Links with thread selection
 - Pick a tree node to select one of the CUDA threads at that location
- Full MPI support
 - See GPU and CPU threads from multiple nodes





Finding the cause

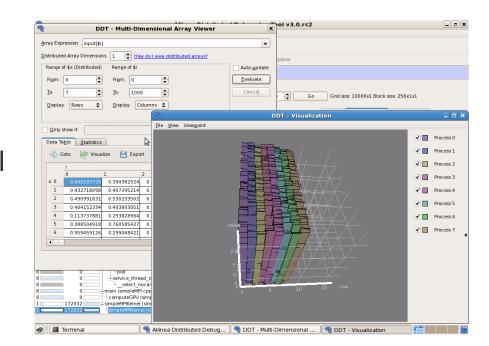
- Full data browsing
 - Local and current line(s) variables
 - Show variables relevant to current position
 - Drag in the source code for more
 - C, C++, F90
 - Shows memory type for CUDA
 - register, shared, ...
- Smart Highlighting
 - Scalable and fast automatic comparison and change detection





A typical CUDA error

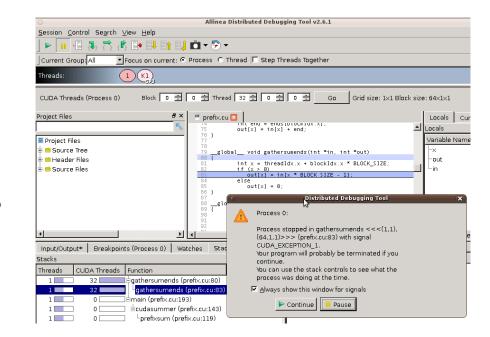
- Grid/block size problems
 - Incorrect dimensions lead to incomplete results
- Easy to see with DDT
 - Use the multi-dimensional array viewer to look at data and find the rough edges
 - 3D display and filtering support
 - New: displays data from multiple processes





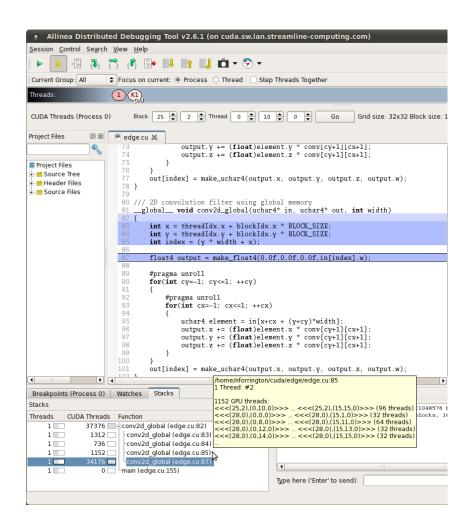
A more serious error

- Grid/block size problems
 - Sometimes lead to worse consequences: overwriting good values, or crashing
 - Bugs will often trigger
 "CUDA memcheck" errors
 - A mode of execution similar to DDT's beyond bound checks for CPU programming





Latest DDT CUDA news



- SDK 3.1 and SDK 3.2
 - Available now in DDT
 - Multi-device support
 - Fermi and Tesla support
 - CUDA Memcheck support for memory errors
 - MPI and CUDA support for GPU clusters
 - Breakpoints, thread control, and data evaluation
 - Stop on kernel launch
- SDK 4.0 in progress
 - C++ kernel debugging

Using DDT on Keeneland

- DDT 2.6.1 is installed and is default
 - To add DDT to your session:
 - % module load ddt
 - Presently requires CUDA 3.2
 - % module load cuda/3.2
- Recompile your application to support debugging
 - More than ordinary "-g" requires "-G" as well

```
% nvcc -g -G prog.cu -o prog
```

Make copy of the sample and build it

```
% cp ~dlecomber/prefix.cu ~
% nvcc -G -g prefix.cu -o prefix
```

- MPI applications supported too but not today's session
 - "Just" add "-G -g" to the nvcc usages
 - ~dlecomber/SimpleMPI/src/SimpleMPI has a Makefile that works feel free to use it later!

Using DDT on Keeneland

- Now to start debugging
 - Load DDT:
 - % ddt
 - Choose "Run and Debug"
 - Select "prefix" using the file browser
 - The prefix example is not MPI but it is CUDA
 - Tick "Run without MPI support"
 - Untick "Run without CUDA support"
 - Click "Submit"
 - NOTE DDT submits your job to the queue including for scalar jobs on Keeneland – by default
 - and now for the real debugging!

Getting Help



- Today's example is available on our website
 - http://www.allinea.com/ Downloads/Whitepapers
- Tutorial material covers other topics
 - Good exercises for learning more about debugging MPI and what DDT can do
 - Memory debugging
 - Multi-dimensional arrays
 - Tips and tricks for getting the most out of the debugger
 - Fully worked through examples with additional exercises
 - Located at: ~dlecomber/Training on Keeneland system
- ... Allinea is here to help
 - support@allinea.com is ready for you!