

allinea

SCALE TO NEW HEIGHTS

***Debugging and Optimizing  
Applications for  
Multicore MPP Architectures***

- **Parallelism in HPC and multicore**
- **Products in more depth**
  - DDT (Distributed Debugging Tool)
  - OPT (Optimization and Profiling Tool)
- **Conclusions**

# Parallelism in HPC

- **Parallelism is established in HPC**
  - Reality since 80s
- **Several established programming models**
  - Vectorization (e.g. Cray, NEC, FJS)
  - Multithreading → PThreads
  - Directives and pragmas → OpenMP
  - Message passing → MPI
  - Task parallelism (e.g. CORBA, PVM, SOAP)
  - .. and mixed programming (e.g. MPI and OpenMP)

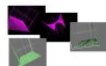
# Parallel Software is Complicated

- **Multithreaded, multiprocess code**
  - The usual issues: bugs, speed ...
  - ... now add communication, synchronization, race conditions, deadlock, scalability ....
  - Visual complexity scales ... as do the bugs
- **Hardware / Software capability misaligned**
  - Significant gap between user requirements and application capabilities

- **A mature, powerful and highly intuitive tool**
  - Traditional focus has been HPC
- **Cross-platform support**
  - Linux, Solaris (Sparc, x86-64), CLE, AIX
  - GNU, Absoft, IBM, Intel, PGI, PathScale, Sun compilers
  - Cell, EM64T, x86-64, IA64, Power, UltraSparc, Cray XT4/5 architectures
  - Across all MPI and OpenMP implementations
  - Support for all major scheduling systems

**allinea ddt**

the distributed debugging tool

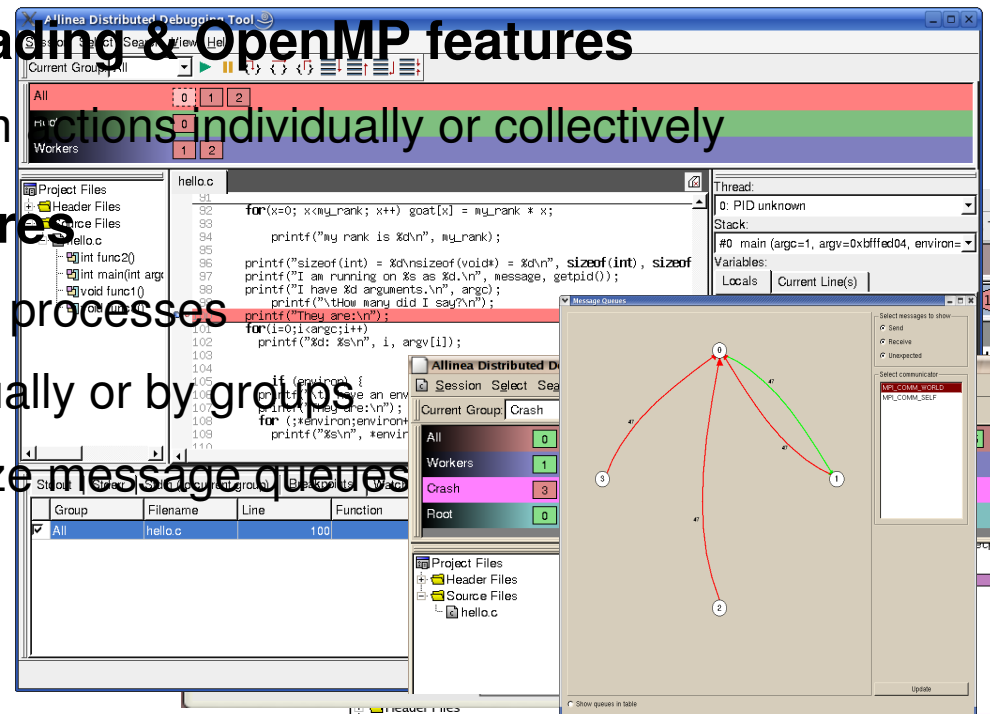
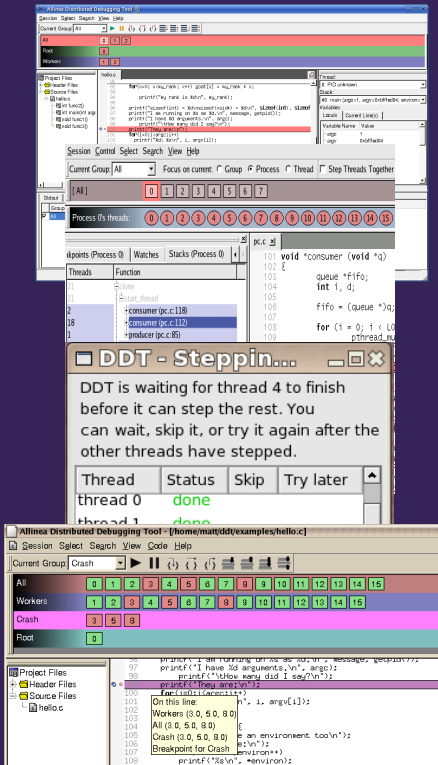


sales@allinea.com  
support@allinea.com

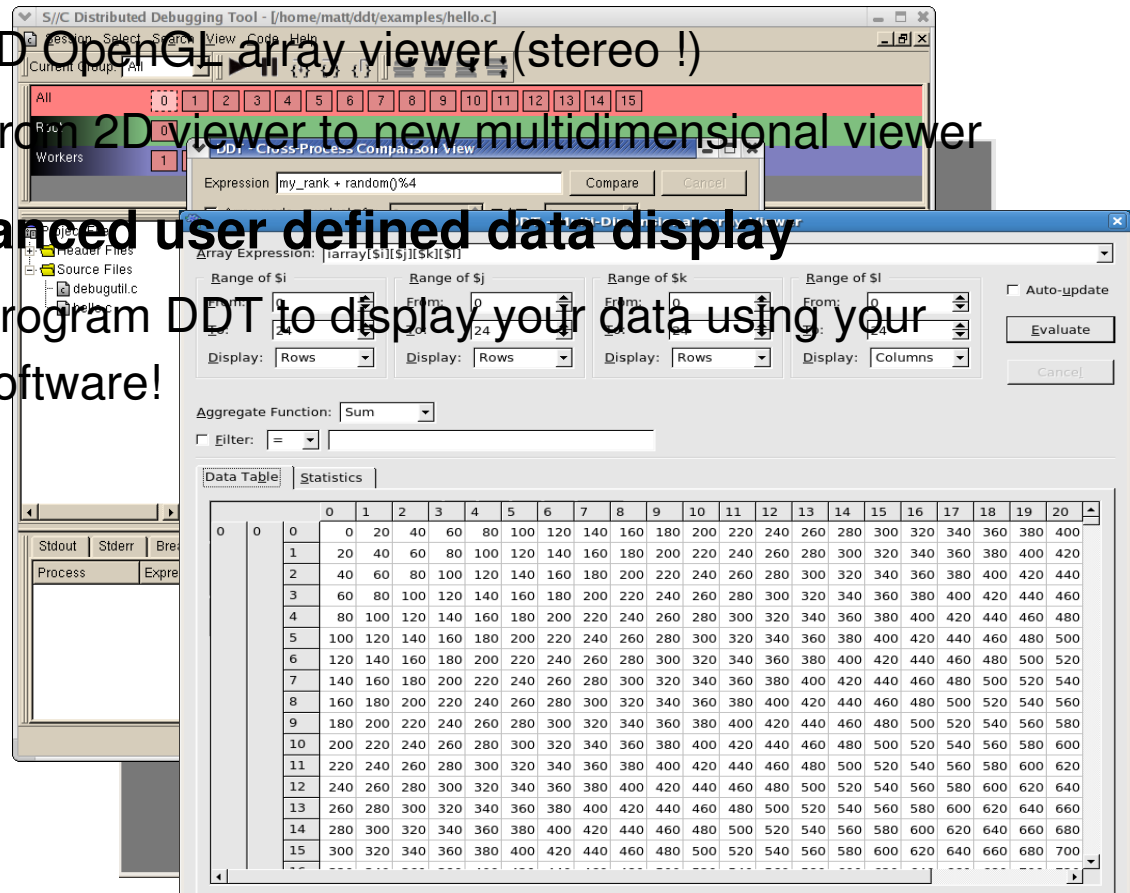
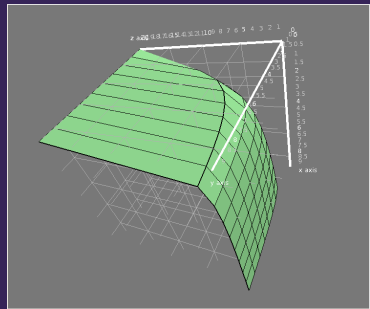
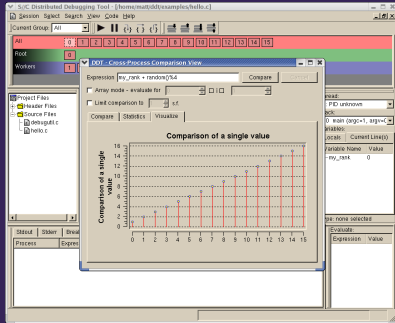
**allinea**  
SCALE TO NEW HEIGHTS

# Features for every model

- **Scalar features**
  - Advanced C++ support including STL, namespaces, virtual functions and templates
  - Advanced Fortran 90, 95 and 2003 support including modules, allocatable data, pointers and derived types
- **Multithreading & OpenMP features**
  - Perform operations individually or collectively
- **MPI features**
  - Control processes
  - individually or by groups
  - Visualize message queues



- Cross process/thread comparison
- Visualize multidimensional data
  - 3D OpenGL array viewer (stereo !)
  - From 2D viewer to new multidimensional viewer
- Advanced user defined data display
  - Program DDT to display your data using your software!



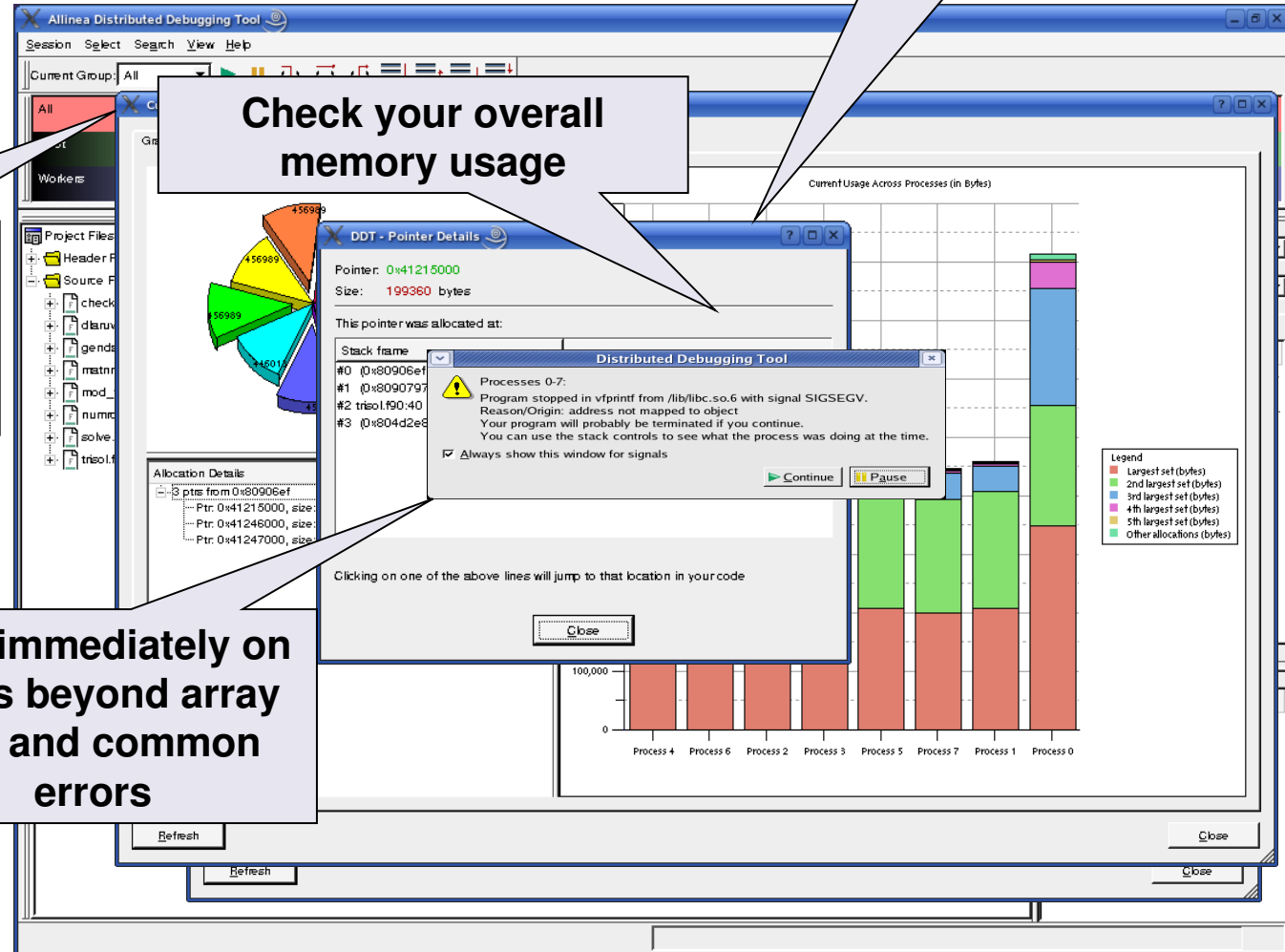
# Memory Debugging

**Locate where memory was allocated**

**Check your overall memory usage**

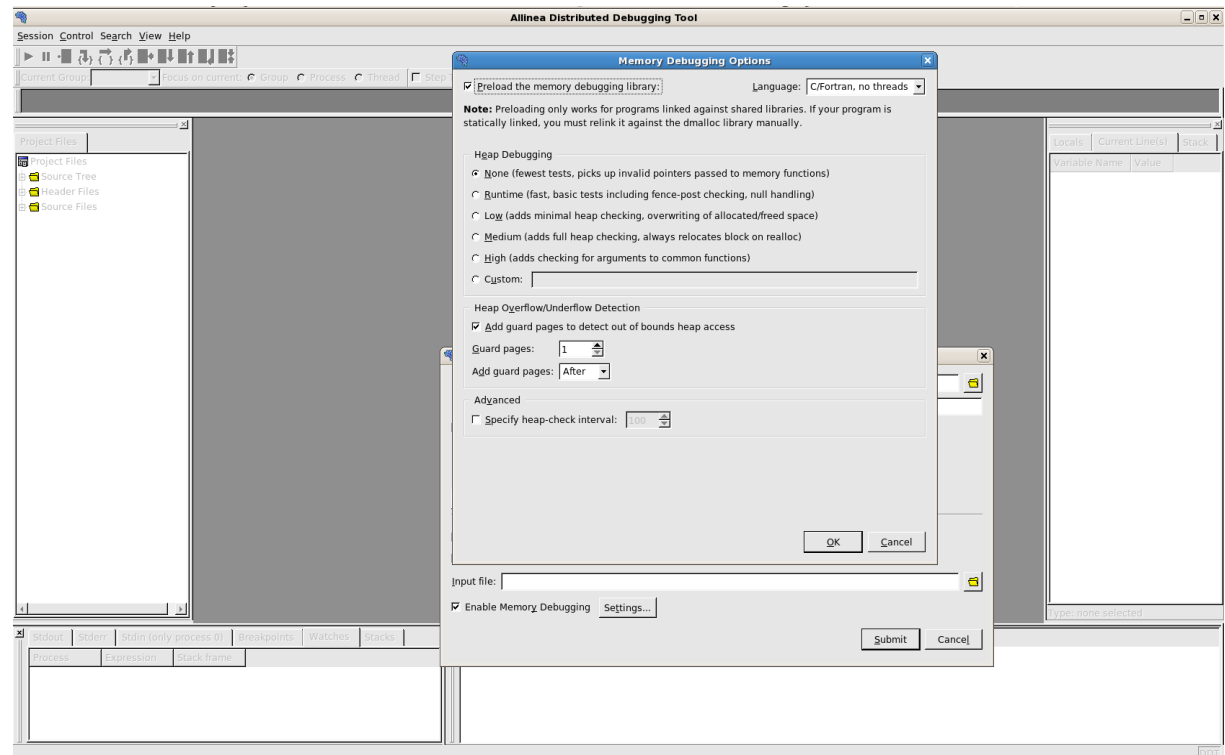
**Check your current memory usage and where memory was allocated**

**Stop immediately on reads beyond array end and common errors**





- **Simple to use**
  - Tick check box and use the default
  - or choose your own advanced settings
  - no recompile necessary



All	1000 processes (0-999)	Paused: 999	Running: 1	Enabled: 0
Root	1 process (0)	Paused: 1	Running: 1	Enabled: 1
Workers	999 processes (1-999)	Paused: 999	Running: 0	Enabled: 0
User Defined	411 processes (67-131, 224-509, 940-999)	Paused: 411	Running: 0	Enabled: 0
Show processes	Currently selected:	67		

Stdout	Stderr	Stdin (only process 0)	Breakpoints	Watches	Stacks
Procs					
4					-main (matrix_mult.cpp:160)
2					-calculate<double> (matrix_mult.cpp:55)
1					__gnu_cxx::normal_iterator<double const*, std::vector<double, std::allocator<double>>>
2					-calculate<double> (matrix_mult.cpp:59)
1					__gnu_cxx::normal_iterator<double const*, std::vector<double, std::allocator<double>>>
1					__gnu_cxx::normal_iterator<double const*, std::vector<double, std::allocator<double>>>

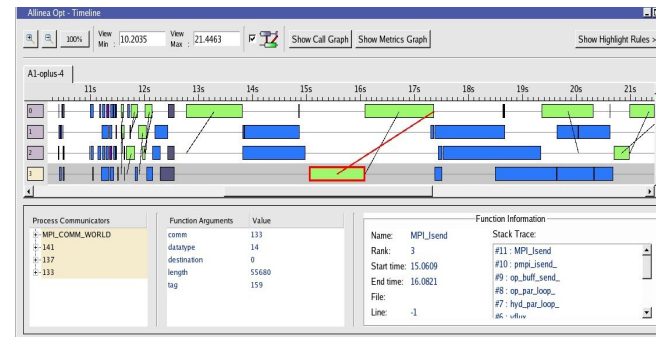
- **DDT 2.0 - released April 2007**
  - Multithreading, easy to control threads
  - Improved memory debugging
- **DDT 2.1 - released August 2007**
  - New message box when processes stop
  - Faster, improved multidimensional array viewing
  - New breakpoint setting box
  - Cell BE support
- **DDT 2.2 – released November 2007**
  - New scalable process groups
  - Enhanced parallel stack view
  - Faster message queue debugging
  - Cray XT4/5 support

# New Architectures

- **DDT NEC SX8 (Q4 2007)**
  - Port of DDT backend to vector platform
  - Remote launch facility
- **DDTLite (Q1 2008)**
  - Simplify development on the Microsoft® platform
  - Bringing features from DDT into Visual Studio®
- **DDT IBM Blue Gene/L (Q1 2008)**
  - Blue Gene/P to follow Q2 2008

- **Traditional tracers**

- Timelines: good for watching messages and memory accesses to pick out problems visually
- But not (currently....) scalable!



- **Can log everything but...**

- Vast quantities of data are generated
  - But is it really necessary?
  - Analysis becomes an expert task

# ...Keep It Simple

- **Focus is the key!**
  - Too much visual information can be confusing
  - Good parallel tools should simplify things
  - Tools should target the areas which cause problems
  - Directing the user towards the problem points...
- **OPT embraces a 'top-down' focused approach**
  - See the “big picture” first – call graph
  - Drill down successively for more information..
  - Don't drown users in too much data
  - Mixture of sampling and selective tracing

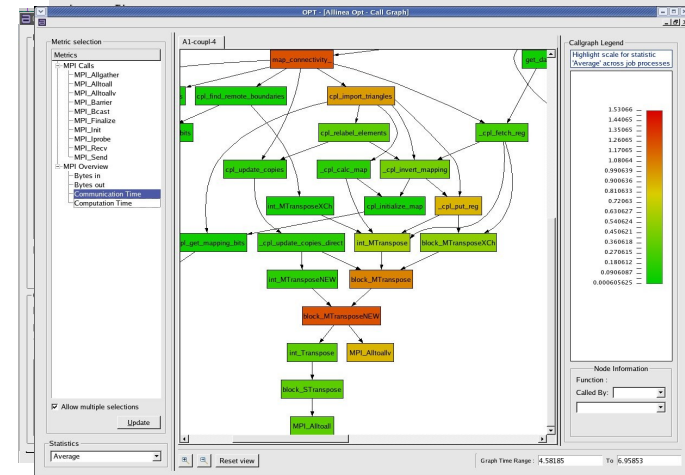


- **Traditional features**

- Timeline shows local problems with sends/receives paired up

- **...and new features**

- Callgraph shows problems
  - Linked with timeline
  - Statistical values
  - Highlights problem functions
  - Aggregated gprof timing data
- Charts show distribution of performance
  - Across processes
  - And statistical views (bar/pie)



**allinea opt**  
the optimization and profiling tool

sales@allinea.com  
support@allinea.com

**allinea**  
SCALE TO NEW HEIGHTS

# How OPT works...

- **Under the hood**
  - Database back end
    - If data is vast, database should handle it
    - Easy to optimize when necessary
    - New capabilities are just new queries
  - Client/Server architecture
    - Web services interface
    - Thin client with small memory footprint
    - Client pulls only information required over network
    - A real GRID tool
- **MPIs, compilers....**
  - Most MPIs, all compilers

# OPT Interface

**Manage Timeline view session**

**Compare different jobs**

**Metric selection**

- MPI Comm\_rank
- MPI Comm\_size
- MPI\_Finalize
- MPI\_Init
- MPI Overview
- Bytes in
- Bytes out
- Communication ...
- Computation Time
- rusage
- ru\_majflt
- ru\_minflt

Allow multiple selections

Update

**Options**

Summarize job by: Mean

Stacked graph

Normalise

Edit order:

- Communication Time
- Computation Time

Statistic: Average

Summarized metric value per job [ job(x) / metric value(y) ]

Job	Computation Time	Communication Time
test - 1 procs	1.46479	0.15678
test - 2 procs	0.8765285	0.4847915
test - 3 procs	0.6506143333	0.6360356667
test - 4 procs	0.53198575	0.66090425
test - 5 procs	0.4656352	1.0069048
test - 6 procs	0.4153381667	1.0596418333

Legend

- Computation Time
- Communication Time



# New Features 2007

- **OPT 1.3.1 - released July 2007**
  - Flat profile in call graph
  - Pruned call graph
  - C++ name demangling
  - Local Server feature
    - Simplify program start up
  - Addition of function level profiling
    - Adding gprof information to callgraph
- **OPT 1.3.3 – released November 2007**
  - Auto creation of jobs from command line
  - New “time range” widget
  - Support for mpich2 and Bull MPI2

# New Features coming...

- **Support for scalar code**
- **Support for IBM Cell BE**
- **Sampler as alternative to gprof**
  - No requirement for recompilation
  - User tunable sample frequency
- **Resource usage time-line**
  - Memory, User time, CPU time
  - Stacked for similar metric
  - Additional graph for different metrics

# Conclusions

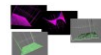
- **HPC and MPP require:**
  - Peta-flop tools.....
  - .....and highly scalable tools
- **Mainstream tools driven by multicore**
  - OpenMP and Threads
  - Task parallelism with modest data parallelism
- **Allinea can lever its expertise in HPC**
  - Offers tried and tested capabilities
  - Market leading multicore/MPP software tools

# allinea

## SCALE TO NEW HEIGHTS

***Thank you***

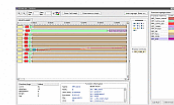
**allinea ddt**  
the distributed debugging tool



[sales@allinea.com](mailto:sales@allinea.com)  
[support@allinea.com](mailto:support@allinea.com)

**allinea**  
SCALE TO NEW HEIGHTS

**allinea opt**  
the optimization and profiling tool



[sales@allinea.com](mailto:sales@allinea.com)  
[support@allinea.com](mailto:support@allinea.com)

**allinea**  
SCALE TO NEW HEIGHTS