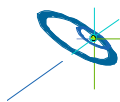


Parallel Debugging

Matthias Müller, Pavel Neytchev

University of Stuttgart
High-Performance Computing-Center Stuttgart (HLRS)
www.hlrs.de

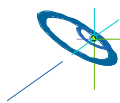


Parallel Debugging and TotalView M. Müller, P. Neytchev
Hochleistungsrechenzentrum Stuttgart



Outline

- Motivation
- Approaches and Tools
- Totalview

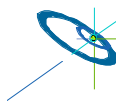


Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 2 Hochleistungsrechenzentrum Stuttgart



Problems of Parallel Programming

- All problems of serial programming
- Additional problems:
 - Increased difficulty to verify correctness of program
 - Increased difficulty to debug N parallel processes
 - New parallel problems:
 - **deadlocks**
 - **race conditions**
 - **irreproducibility**

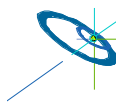


Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 3 Höchstleistungsrechenzentrum Stuttgart

H L R I S 

What is a debugger?

- Common Misconception:
A debugger is a tool to find and remove bugs
- A debugger does:
 - tell you where the program crashed
 - help to gain a better understanding of the program and what is going on
- Consequence:
 - A debugger does not help much if your program does not crash, e.g. just gives wrong results
 - Avoid using a debugger as far as possible.
 - Use it as last resort.

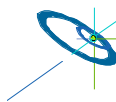


Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 4 Höchstleistungsrechenzentrum Stuttgart

H L R I S 

Avoiding debuggers

- Think about a verbose execution mode of your program
- Use a careful/paranoid programming style
 - check invariants and pre-requisites (assert(m>=0), assert(v<c))
- Use the debugging/assertion techniques of the compiler
 - use debug flags (-g), warnings (-Wall)
 - array bound checks in Fortran
 - use memory debug libraries (-lefence)

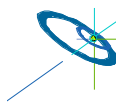


Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 5 Höchstleistungsrechenzentrum Stuttgart

H L R I S 

Avoiding Debuggers

- Write portable programs
 - it avoids future problems
 - **architectures/platforms have a short life**
 - **all compilers and libraries have bugs**
 - **all languages and standards include implementation defined behavior**
 - running on different platforms and architectures significantly increases the reliability
- Use verification tools for parallel programming like assure

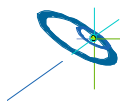


Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 6 Höchstleistungsrechenzentrum Stuttgart

H L R I S 

Parallel Debuggers

- Most vendor debuggers have some support
- gdb has basic support for threads
- Debugging MPI programs with a “scalar” debugger is hard but possible
 - mpich supports debugging with gdb attached to one process
 - manual attaching to the processes is possible
- totalview is a good but expensive tool



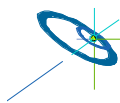
Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 7 Höchstleistungsrechenzentrum Stuttgart



TOTALVIEW

Matthias Müller

University of Stuttgart
High-Performance Computing-Center Stuttgart (HLRS)
www.hlrs.de

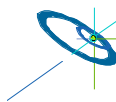


Parallel Debugging and TotalView Müller, Neytchev
Höchstleistungsrechenzentrum Stuttgart



What is TotalView?

- Parallel debugger
- Source level debugging for C, C++, F77, F90, HPF
- **MPI, OpenMP**, Pthreads, PVM, shmem
- SMPs, MPPs, PVPs, Clusters
- Available on all major Unix Platforms and most Supercomputers
- GUI (independent of platform, exception Cray T3E)
 - totalview 4.x based on tcl/tk
 - totalview 5.x based on Motif

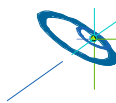


Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 9 Höchstleistungsrechenzentrum Stuttgart

H L R I S 

Availability of TotalView

- [Compaq Digital Alpha](#)
- [HP-UX](#)
- [IBM RS6000 and SP Power](#)
- [SGI MIPS](#)
- [Sun SPARC SunOS 5](#)
- [Linux Intel IA32 \(RedHat\)](#)
- [Linux Alpha \(RedHat\)](#)
- Cray T3E by Cray
- Hitachi SR2201 by SofTek, SR8000
- NEC SX-4 (V 3.x) by SofTek, SX-5 in beta-test



Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 10 Höchstleistungsrechenzentrum Stuttgart

H L R I S 

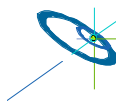
Availability of TotalView at HWW

Platform	Availability	Remarks
Volvox	Yes	V 5 and V4
Hitachi SR2201	Yes	V 3.7.9
Hitachi SR8000	Yes	V 4.0
Cray T3E	Yes	Cray 3.0.0
NEC SX-4	Yes	V 3.7.9
NEC SX-5	Yes	Betatest
SGI Onyx	No	Use cvd
HP N-Class	Yes	V 4.1 and V5
IBM SP	Yes	V 4.1
Cray SV1	Yes	Cray 3.0.0

Development
Platforms

More information:

www.hlr.de/organization/par/services/tools/debugger/totalview



Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 11 Höchstleistungsrechenzentrum Stuttgart

H L R I S

Availability of TotalView at University Stuttgart

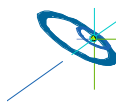
2 user 8 CPU Floating License for University Stuttgart:

1. Download Software from <http://www.etnus.com>

2. Set environment variable for license.

LM_LICENSE_FILE=7244@servint1.rus.uni-stuttgart.de

More information about campus licenses available at
<http://www.hlr.de/organization/par/services/tools/campus>

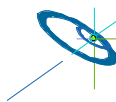


Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 12 Höchstleistungsrechenzentrum Stuttgart

H L R I S

TotalView usage at HLRS

- set USE_TOTALVIEW in your login scripts
- CRAY T3E: set USE_PROG_ENV
- Compile with -g compiler switch
CRAY T3E: compiler switch -G
- command name: **totalview**



Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 13 Höchstleistungsrechenzentrum Stuttgart

HLRS

Starting TotalView

On a new process:

```
% totalview myprog -a arguments to myprog
```

To debug MPI programs:

```
% totalview mpirun -a -nprocs 3 myprog
```

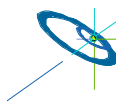
```
% mpirun -tv -np 3 myprog
```

To debug IBM POE programs:

```
% totalview poe -a myprog [args]
```

To debug CRAY T3E programs:

```
% totalview -X #procs myprog [args]
```

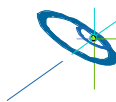


Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 14 Höchstleistungsrechenzentrum Stuttgart

HLRS

Totalview on Hitachi SR8000

- Compilation:
 - f90 -g
 - cc -g
 - KCC -g --backend -tv
- OpenMP
 - f90 -g -omp -procnum=8
 - cc -g -omp -parallel=1 -O2
- MPI
 - mpirun -tv

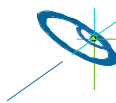


Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 15 Höchstleistungsrechenzentrum Stuttgart

H L R I S 

Totalview on HPN

- Compilation:
 - f90 -g
 - cc -g
 - KCC -g
- OpenMP
 - guidef90 -g
 - guidec -g
 - guidec++ -g
- MPI
 - mpirun -np #procs -tv ./a.out

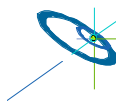


Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 16 Höchstleistungsrechenzentrum Stuttgart

H L R I S 

Totalview Exercise: Basic Look & Feel

- Log into hwwhpn.hww.de
- use bash as shell
- Change into directory
~/TOTALVIEW/#NR/TOTALVIEW/SIMPLE
- Compile `calc_pi_{f90,c,cc}.{f90,c,cc}`
- Start totalview with `totalview executable`

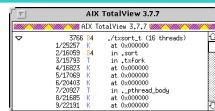


Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 17 Höchstleistungsrechenzentrum Stuttgart

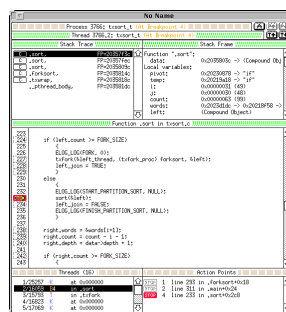


TotalView Windows

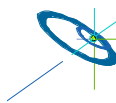
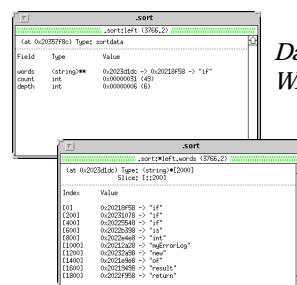
Root Window



Process Window



Data Windows

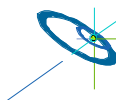
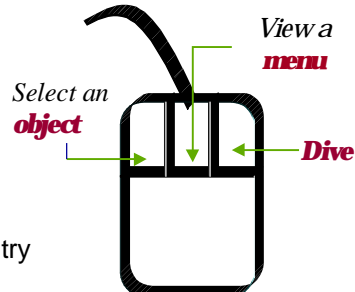


Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 18 Höchstleistungsrechenzentrum Stuttgart



Totalview Mouse Buttons

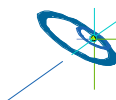
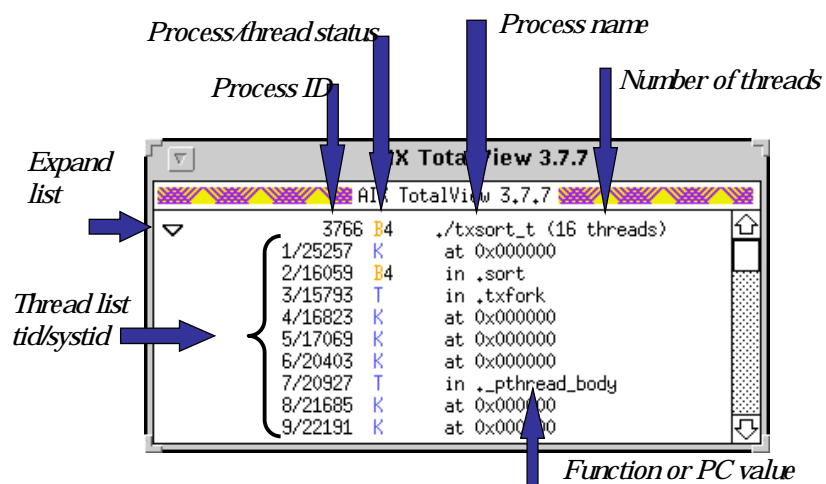
- **Left button is Select:**
 - Chooses an item of interest,
 - Starts editing a item
- **Middle button is Menu:**
 - Raises a menu of actions
 - All menus have a **Help** (^?) entry
- **Right button is Dive:**
 - Gets more information about an item
 - **Shift+Dive** forces open a new window



Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 19 Höchstleistungsrechenzentrum Stuttgart

HLRS

Totalview Main Window



Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 20 Höchstleistungsrechenzentrum Stuttgart

HLRS

TotalView Process Window

The screenshot shows the TotalView Process Window for a process named 'tsort_t'. The window is divided into several panes:

- Stack Trace pane:** Shows the call stack with frames for 'sort', 'forksort', and 'main'.
- Source pane:** Displays the source code for the function 'sort' in 'tsort.c'. The code includes a recursive sorting function with a 'fork' call.
- Thread pane:** Shows a list of threads, including 'in _lap_sens_wait', 'in _signal_wait', 'in _lap_sens_wait', and 'in sort'.
- Action Points pane:** Shows a list of action points, including '1 line 230 in sort+0x30c'.

Arrows point to the following components:

- Process/thread motion buttons:** Located at the top right of the window.
- Local variables for the selected frame:** Located on the right side of the window, showing variables like 'pivot', 'i', 'count', 'word', and 'left'.

Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 21 Höchstleistungsrechenzentrum Stuttgart

H L R I S

TotalView Source Pane

Gridded box is a possible site for a breakpoint

Current function and source file

Select to set one

Current point of execution

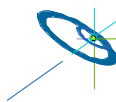
- **Dive** on a source word to get more information
- **Select** a line to use **Run to selection** command
- **Select** or **dive** on a line number to set an action point

Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 22 Höchstleistungsrechenzentrum Stuttgart

H L R I S

Parallel Debugging - Philosophy

- By default, TotalView places processes in groups
 - Program Group - Includes parent and all related processes
 - Share Group - Only processes that share the same source code
- Command can act on single process or share group
 - halt process (h) , halt group (H)
 - next step process (n), next step group (N)
 - go process (g), go group (G)

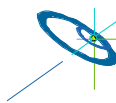


Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 23 Höchstleistungsrechenzentrum Stuttgart

H L R I S 

Totalview Exercise: Debug simple program

- Run calc_pi inside totalview:
 - Check where the program crashes
- analyze core file with totalview
 - run calc_pi
 - execute `totalview calc_pi core`
- for advanced users: choose another programming paradigm:
 - MPI, OpenMP, MPI+OpenMP

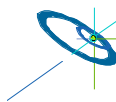


Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 24 Höchstleistungsrechenzentrum Stuttgart

H L R I S 

Totalview support for debugging MPI

- Special support for MPI is available depending on your MPI library:
 - display message queue state of a process
- Supported MPI implementations:
 - mpich v1.1.0 or later (use -debug in configure)
 - HP MPI v1.6
 - Compaq MPI >v1.7
 - IBM, release >2.2 of Parallel Environment, threaded version of MPI
 - SGI MPI v1.3 or later



Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 25 Höchstleistungsrechenzentrum Stuttgart



TotalView MPI Message Queue Window

Communicator name and info

Non-blocking receive operations

Unmatched incoming messages

Non-blocking send operations

- Dive on source or target to refocus Process window
- Dive on buffer to see message contents

testsome.0

Message State for "testsome.0" (1288,1)

MPI_COMM_WORLD

Comm_size 3

Comm_rank 0

Pending receives (0)

Status Pending

Source 2 (testsome.2)

Tag 0x00000000 (0)

User Buffer 0x000605c0 -> 0x00000000 (0)

Buffer Length 0x00000014 (20)

Unexpected messages (0)

Status Complete

Source 2 (testsome.2)

Tag 0x00000002 (2)

System Buffer 0x00000000

Buffer Length 0x00000000 (0)

Received Length 0x00000000 (0)

Non-blocking sends (0)

Status Complete

Target 2 (testsome.2)

Tag 0x00000000 (0)

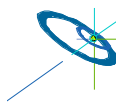
Buffer 0x000605a0 -> 0x00000001 (1)

Buffer Length 0x00000014 (20)

MPI_COMM_WORLD_collective

Comm_size 3

Comm_rank 0

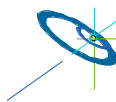


Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 26 Höchstleistungsrechenzentrum Stuttgart



TotalView Exercise: Parallel program

- Example in TOTALVIEW/MPI:
 - `deadlock_{c,cc,f90}.{c,cc,f90}`
 - start program with `mpirun -tv -np 2 a.out`
 - interrupt execution after “deadlock”
 - try to find the reason for the deadlock and fix it
- For advanced users:
 - `pending_{c,cc,f90}.{c,cc,f90}`
 - try to find pending message by setting breakpoint at `MPI_Finalize`

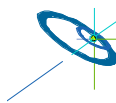


Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 27 Höchstleistungsrechenzentrum Stuttgart



TotalView more information

- <http://www.etnus.com/products/totalview/index.html>
- <http://www.hlrs.de/organization/par/services/tools/debugger/totalview>
 - User Guide
 - Installation Guide
 - CLI Guide
 - Powerpoint Tutorial
- CRAY T3E: Online Documentation at <http://www.hlrs.de/platforms/crayt3e>



Parallel Debugging and TotalView M. Müller, P. Neytchev
Slide 28 Höchstleistungsrechenzentrum Stuttgart

