



## Tools for OpenMP Programming

PPCES 2018

Tim Cramer

## ■ ThreadSanitizer

→ Overview

## ■ Intel Inspector XE

→ Overview

→ Live Demo

## ■ Intel VTune Amplifier XE

→ Overview

→ Live Demo

# Race Condition

---

- **Data Race: the typical OpenMP programming error, when:**
  - two or more threads access the same memory location, and
  - at least one of these accesses is a write, and
  - the accesses are not protected by locks or critical regions, and
  - the accesses are not synchronized, e.g. by a barrier.
- **Non-deterministic occurrence: e.g. the sequence of the execution of parallel loop iterations is non-deterministic and may change from run to run**
- **In many cases *private* clauses, *barriers* or *critical regions* are missing**
- **Data races are hard to find using a traditional debugger**
  - Use the Intel Inspector XE or the ThreadSanitizer

## ■ ThreadSanitizer

→ Overview

## ■ Intel Inspector XE

→ Overview

→ Live Demo

## ■ Intel VTune Amplifier XE

→ Overview

→ Live Demo

# OpenMP Correctness checking: ThreadSanitizer – Overview

---

- Correctness checking for threaded applications
- Integrated in clang and gcc compiler
- Low runtime overhead: 2x – 15x
- Used to find data races in browsers like Chrome and Firefox

# OpenMP Correctness checking: ThreadSanitizer – Usage

module load archer gcc/6

- Compile the program with clang compiler:

C

```
clang -fsanitize=thread -fopenmp -g myprog.c -o myprog
```

C++

```
clang++ -fsanitize=thread -fopenmp -g myprog.cpp  
-o myprog
```

Fortran

```
gfortran -fsanitize=thread -fopenmp -g myprog.f -c  
clang -fsanitize=thread -fopenmp -lgfortran myprog.o  
-o myprog
```

- Execute:

```
OMP_NUM_THREADS=4 ./myprog
```

- Understand and correct the detected threading errors
- Edit the source code
- Repeat until no errors reported



# OpenMP Correctness checking: ThreadSanitizer – Result Summary

```
1 #include <stdio.h>
2
3 int main(int argc, char **argv) {
4     int a = 0;
5     #pragma omp parallel
6     {
7         if (a < 100) {
8             #pragma omp critical
9             a++;
10        }
11    }
12 }
```

**WARNING: ThreadSanitizer: data race**

• Read of size 4 at 0x7ffffffdcdc by thread T2:

#0 .omp\_outlined. race.c:7  
(race+0x0000004a6dce)  
#1 \_\_kmp\_invoke\_microtask <null>  
(libomp\_tsan.so)

• Previous write of size 4 at 0x7ffffffdcdc by main thread:

#0 .omp\_outlined. race.c:9  
(race+0x0000004a6e2c)  
#1 \_\_kmp\_invoke\_microtask <null>  
(libomp\_tsan.so)

## ■ ThreadSanitizer

→ Overview

## ■ Intel Inspector XE

→ Overview

→ Live Demo

## ■ Intel VTune Amplifier XE

→ Overview

→ Live Demo



## ■ Detection of

- Memory Errors
- Deadlocks
- Data Races

## ■ Support for

- Linux (32bit and 64bit) and Windows (32bit and 64bit)
- WIN32-Threads, Posix-Threads, Intel Threading Building Blocks and OpenMP

## ■ Features

- Binary instrumentation gives full functionality
- Independent stand-alone GUI for Windows and Linux

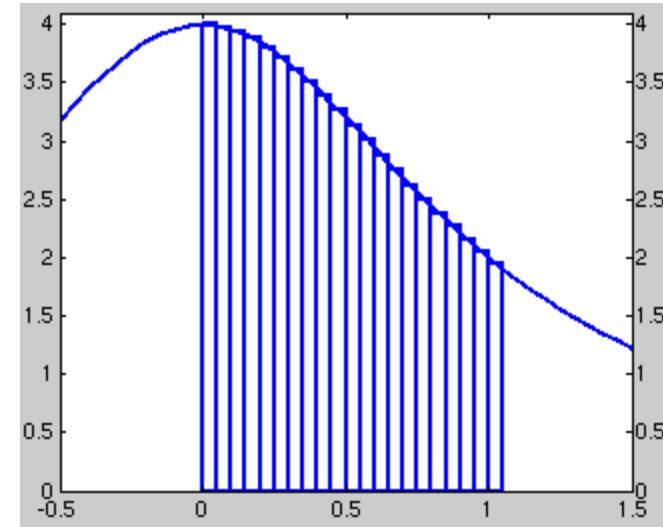
# PI Example Code

```
double f(double x)
{
    return (4.0 / (1.0 + x*x));
}
```

$$\pi = \int_0^1 \frac{4}{1+x^2}$$

```
double CalcPi (int n)
{
    const double fH = 1.0 / (double) n;
    double fSum = 0.0;
    double fX;
    int i;
```

```
#pragma omp parallel for private(fX,i) reduction(+:fSum)
    for (i = 0; i < n; i++)
    {
        fX = fH * ((double)i + 0.5);
        fSum += f(fX);
    }
    return fH * fSum;
}
```



# PI Example Code

---

```
double f(double x)
{
    return (4.0 / (1.0 + x*x));
}
```

```
double CalcPi (int n)
{
    const double fH = 1.0 / (double) n;
    double fSum = 0.0;
    double fX;
    int i;
```

```
#pragma omp parallel for private(fX,i) reduction(+:fSum)
for (i = 0; i < n; i++)
{
    fX = fH * ((double)i + 0.5);
    fSum += f(fX);
}
return fH * fSum;
}
```

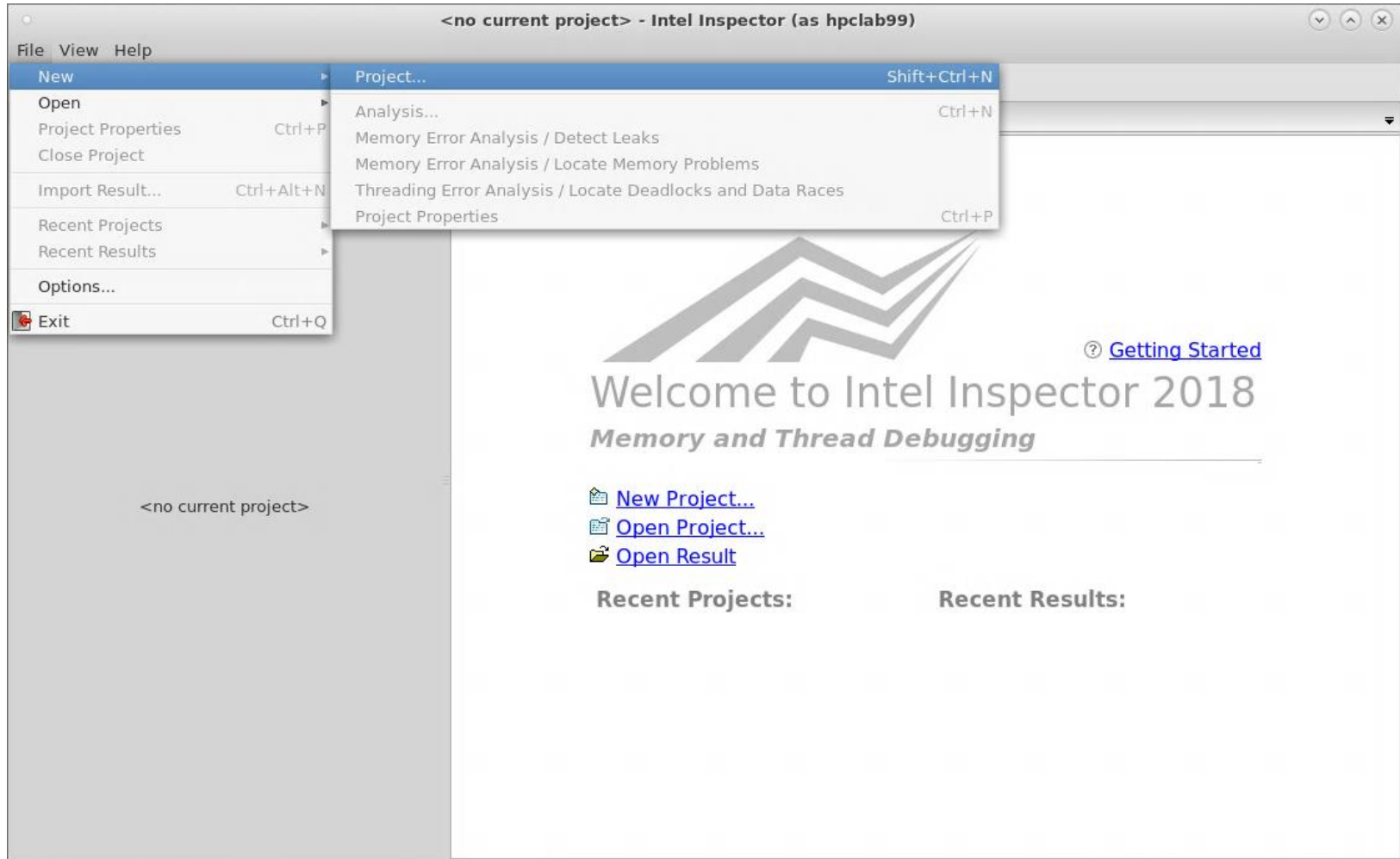
What if we  
would have  
forgotten  
this?

# Live Demo

Intel Inspector XE

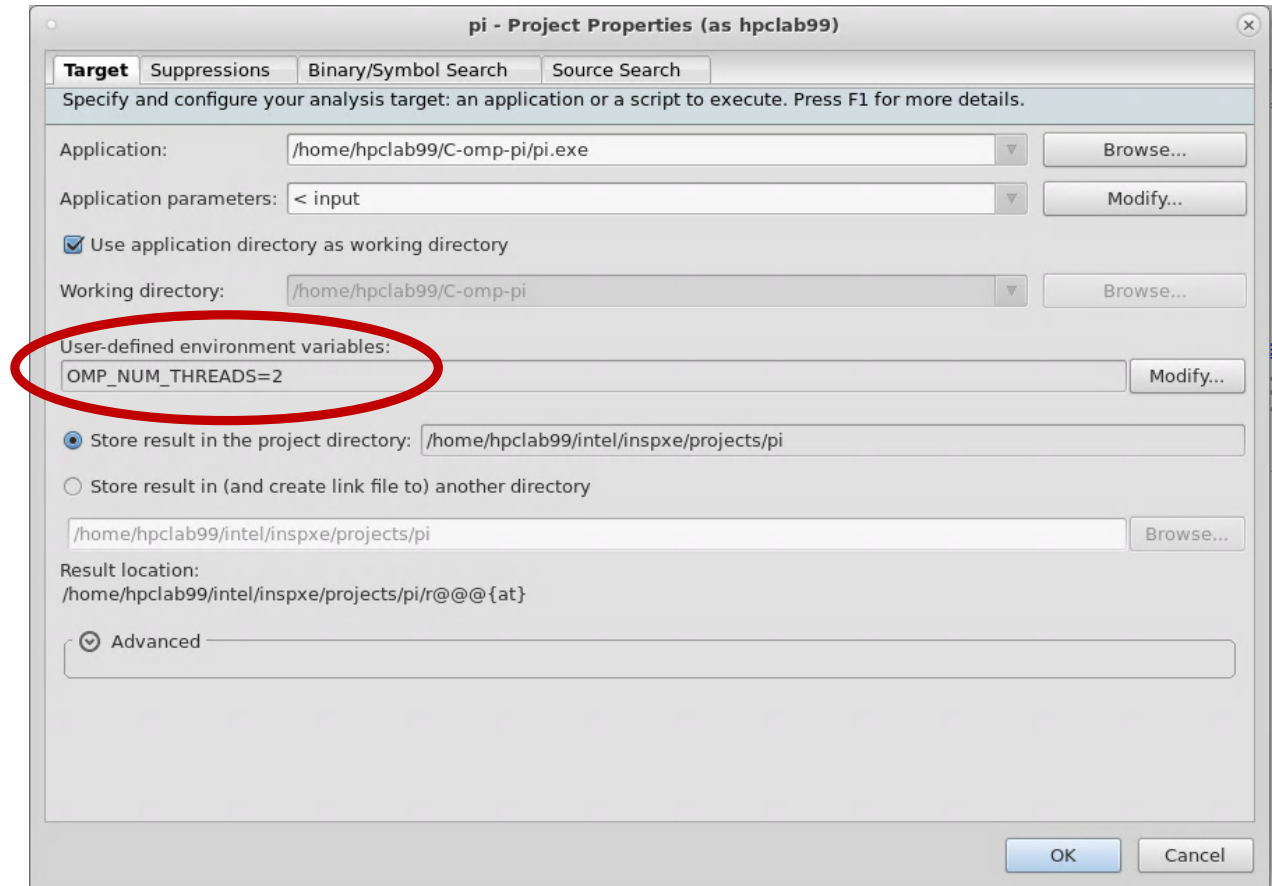
# Inspector XE – Create Project

```
$ module load intelixe ; inspxe-gui
```



# Inspector XE – Create Project

- ensure that multiple threads are used
- choose a small dataset (really!), execution time can increase 10X – 1000X

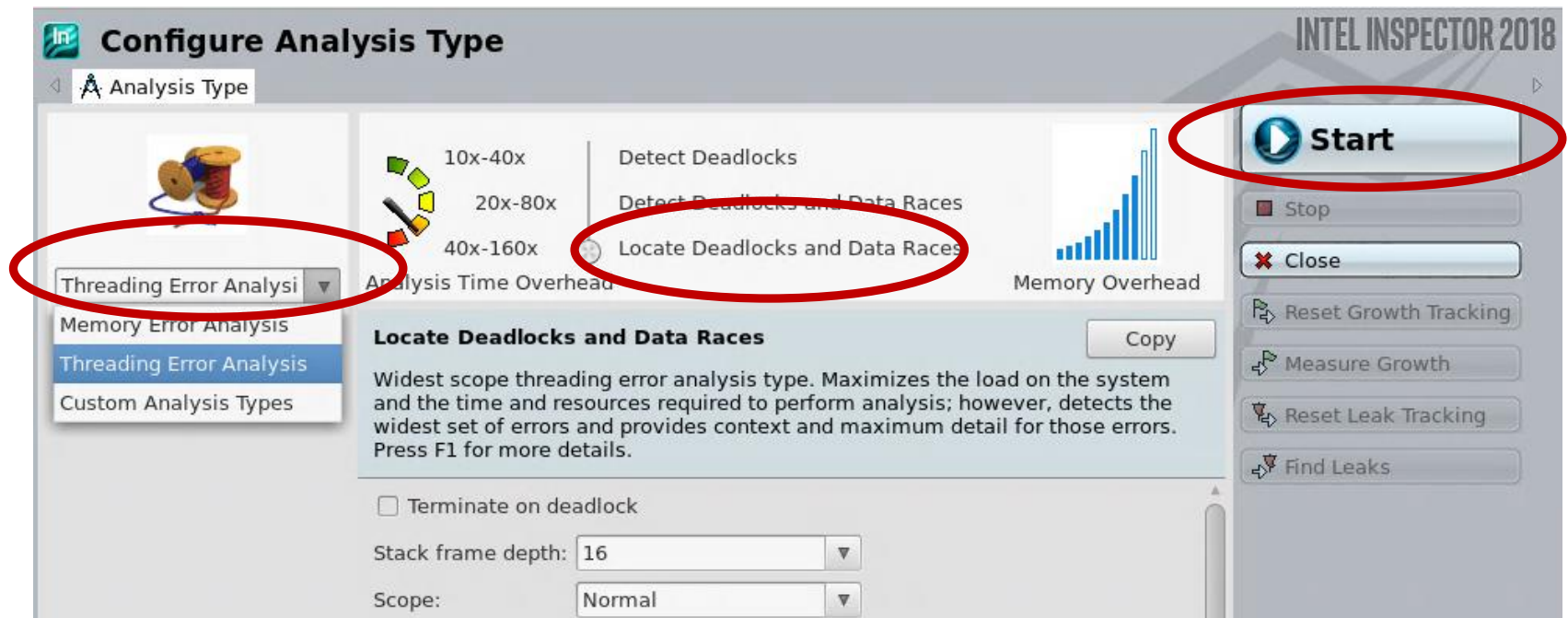


# Inspector XE – Configure Analysis

## Threading Error Analysis Modes

1. Detect Deadlocks
2. Detect Deadlocks and Data Races
3. Locate Deadlocks and Data Races

more details,  
more overhead





# Inspector XE – Results

- 1 detected problems
- 2 filters
- 3 code location
- 4 Timeline

The screenshot displays the Intel Inspector 2018 interface. The main window is titled "/home/hpclab99/intel/inspxe/projects/pi - Intel Inspector (as hpclab99)". The "Locate Deadlocks and Data Races" section is active, showing a list of detected problems. A yellow circle '1' highlights the "Problems" table, which lists three data race issues. A yellow circle '2' highlights the "Filters" panel on the right, which shows the current filter settings. A yellow circle '3' highlights the "Code Locations: Data race" panel, which shows the source code for the detected data race. A yellow circle '4' highlights the "Timeline" panel, which shows the execution timeline of the program.

**Problems**

ID	Type	Sources	Modules	State
P1	Data race	pi.c	pi.exe	New
	Data race	pi.c:72	pi.exe	New
	Data race	pi.c:72	pi.exe	New

**Filters**

Severity	Count
Error	1 item(s)

**Type**

Type	Count
Data race	1 item(s)

**Source**

Source	Count
pi.c	1 item(s)

**Module**

Module	Count
pi.exe	1 item(s)

**State**

State	Count
New	1 item(s)

**Suppressed**

Suppressed	Count
------------	-------

**Code Locations: Data race**

Description	Source	Function	Module	Variable
Read	pi.c:72	CalcPi	pi.exe	
<pre>70 { 71     fX = fH * ((double)i + 0; 72     fSum += f(fX); 73 } 74 return fH * fSum;</pre>				
Write	pi.c:72	CalcPi	pi.exe	
<pre>70 { 71     fX = fH * ((double)i + 0; 72     fSum += f(fX); 73 } 74 return fH * fSum;</pre>				

**Timeline**

Thread	Start	End
OMP Master Thread #0 (23581)		
OMP Worker Thread #1 (23717)		

# Inspector XE – Results

- 1 Source Code producing the issue – double click opens an editor
- 2 Corresponding Call Stack

The image displays the Intel Inspector 2018 interface showing a data race between two threads. The top panel, titled "Data race", shows the "Read - Thread OMP Master Thread #0 (23581) (pi.exe!CalcPi - pi.c:72)". The bottom panel shows the "Write - Thread OMP Worker Thread #1 (23717) (pi.exe!CalcPi - pi.c:72)". Both panels show the same source code snippet from pi.c, with the line `fSum += f(fX);` highlighted in blue. A yellow circle with the number "1" is placed over this line in both panels. To the right of the source code, a "Call Stack" window is visible, showing the call stack for each thread. A yellow circle with the number "2" is placed over the call stack in both panels. The call stack for the Master Thread shows `pi.exe!CalcPi - pi.c:72`, `pi.exe!CalcPi - pi.c:68`, and `pi.exe!_start`. The call stack for the Worker Thread shows `pi.exe!CalcPi - pi.c:72`.

**Data race**

Target Analysis Type Collection Log Summary Sources

**Read - Thread OMP Master Thread #0 (23581) (pi.exe!CalcPi - pi.c:72)**

pi.c Disassembly (pi.exe!0x111f)

```
67 //#pragma omp parallel for private(i, fX) reduction(+:fSum)
68 #pragma omp parallel for private(i, fX)
69   for (i = iRank; i < n; i += iNumProcs)
70   {
71       fX = fH * ((double)i + 0.5);
72       fSum += f(fX);
73   }
74   return fH * fSum;
75 }
76
```

**Call Stack**

- pi.exe!CalcPi - pi.c:72
- pi.exe!CalcPi - pi.c:68
- pi.exe!\_start

**Write - Thread OMP Worker Thread #1 (23717) (pi.exe!CalcPi - pi.c:72)**

pi.c Disassembly (pi.exe!0x1395)

```
67 //#pragma omp parallel for private(i, fX) reduction(+:fSum)
68 #pragma omp parallel for private(i, fX)
69   for (i = iRank; i < n; i += iNumProcs)
70   {
71       fX = fH * ((double)i + 0.5);
72       fSum += f(fX);
73   }
74   return fH * fSum;
75 }
76
```

**Call Stack**

- pi.exe!CalcPi - pi.c:72

## Inspector XE – Results

- 1 Source Code producing the issue – double click opens an editor
- 2 Corresponding Call Stack

The missing reduction is detected.

The screenshot displays the Visual Studio IDE with two threads running a parallel pi calculation. Both threads are at line 72 of `pi.c`, which contains the statement `fSum += f(fX);`. A green box in the top right corner indicates that a reduction is detected. The threads are labeled as follows:

- Thread 0 (Master Thread):** Read - Thread OMP Master Thread #0 (23581) (pi.exe!CalcPi - pi.c:72)
- Thread 1 (Worker Thread):** Write - Thread OMP Worker Thread #1 (23717) (pi.exe!CalcPi - pi.c:72)

The disassembly view for both threads shows the following code:

```

67  //#pragma omp parallel for private(i, fX) reduction(+:fSum)
68  #pragma omp parallel for private(i, fX)
69      for (i = iRank; i < n; i += iNumProcs)
70      {
71          fX = fH * ((double)i + 0.5);
72          fSum += f(fX);
73      }
74      return fH * fSum;
75  }
76

```

The call stack for both threads shows the following frames:

- Thread 0: pi.exe!CalcPi - pi.c:72, pi.exe!CalcPi - pi.c:68, pi.exe!\_start
- Thread 1: pi.exe!CalcPi - pi.c:72

The threads are highlighted with yellow circles containing the number 1, and the call stack frames are highlighted with yellow circles containing the number 2.

# Command Line Tool – inspxe-cl

---

## Threading Error Analysis Modes

1. Detect Deadlocks (ti1)
2. Detect Deadlocks and Data Races (ti2)
3. Locate Deadlocks and Data Races (ti3)

**\$ inspxe-cl -collect ti3 -- pi.exe < input**

Data collection without GUI allows to use batch jobs.

**\$ inspxe-cl -report problems ...**

Viewing results in text mode is helpful, when remote connections are slow.

## ■ ThreadSanitizer

→ Overview

## ■ Intel Inspector XE

→ Overview

→ Live Demo

## ■ Intel VTune Amplifier XE

→ Overview

→ Live Demo

## ■ Performance Analyses for

- Serial Applications
- Shared Memory Parallel Applications

## ■ Sampling-based measurements

## ■ Features:

- Hot Spot Analysis
- Concurrency Analysis
- Wait
- Hardware Performance Counter Support

## ■ Prerequisites:

- `ssh -Y login-t`
- `module load intelvtune`

## ■ Application

- Use representative data set
- Build with debug information (-g) AND optimization (-O3)
- Set environment variables (e.g., `OMP_NUM_THREADS`, `OMP_PROC_BIND`, `OMP_PLACES`, etc.)

# Stream

- Standard Benchmark to measure memory performance.
- Version is parallelized with OpenMP.

Measures Memory bandwidth for:

$y=x$  (copy)

$y=s*x$  (scale)

$y=x+z$  (add)

$y=x+s*z$  (triad)

```
#pragma omp parallel for  
for (j=0; j<N; j++)  
    b[j] = scalar*c[j];
```

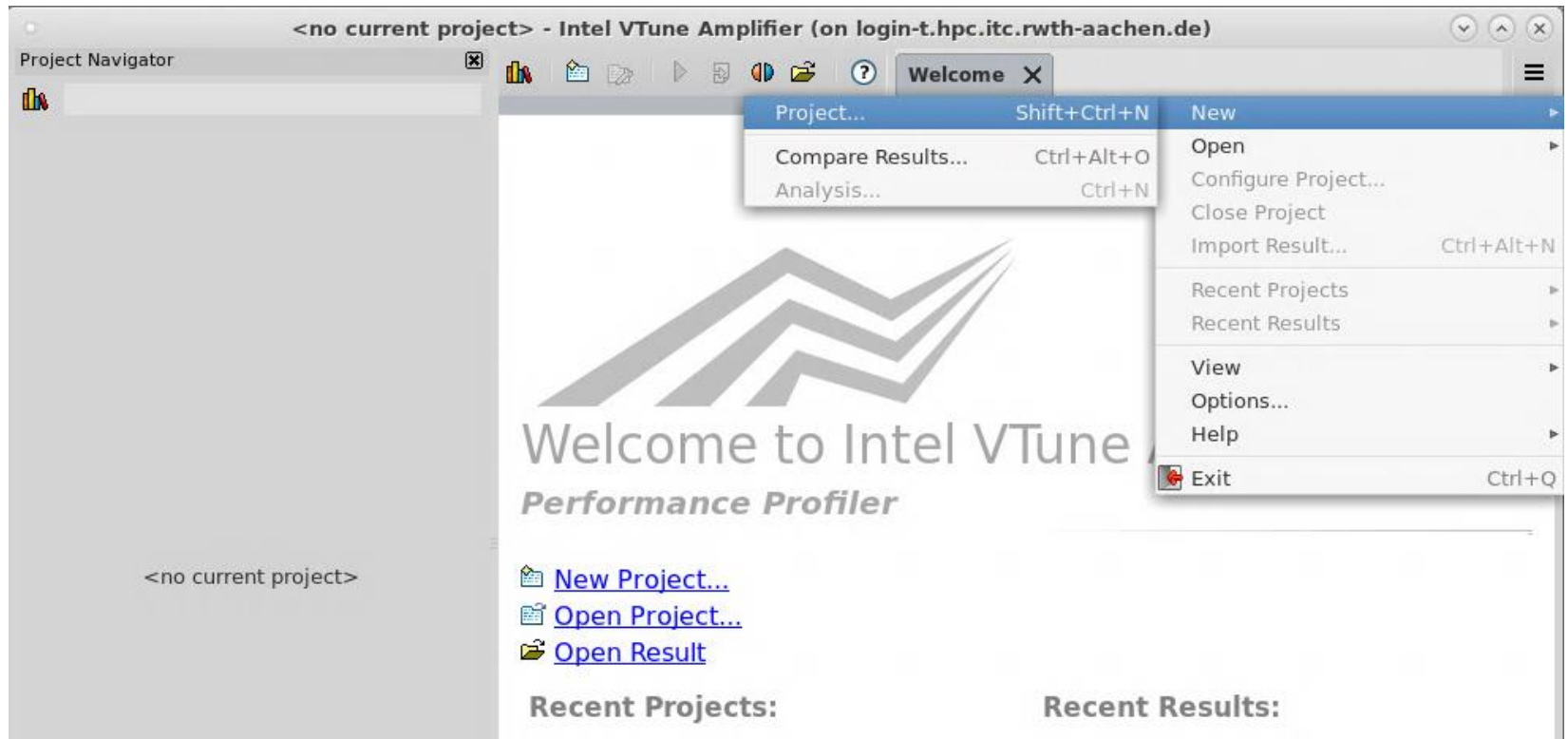
for double vectors  $x,y,z$  and scalar double value  $s$

Function	Rate (MB/s)	Avg time	Min time	Max
time				
Copy:	121755.3877	0.0119	0.0118	0.0122
Scale:	122280.4397	0.0118	0.0118	0.0119
Add:	124754.8422	0.0173	0.0173	0.0176
Triad:	124797.8048	0.0174	0.0173	0.0178



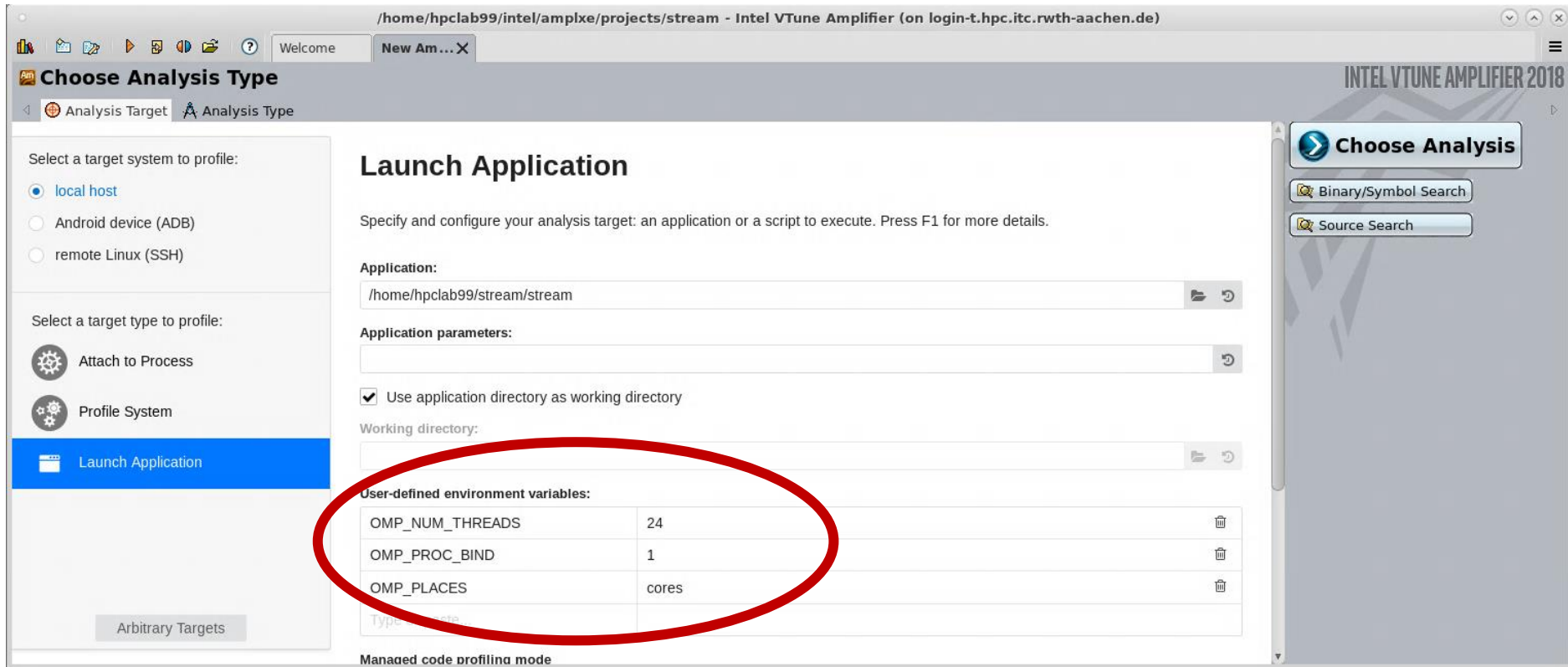
# Intel VTune Amplifier XE – Create Project

- `$ module load intelvtune; amplxe-gui`
- Create a Project in the same way as with the inspector.
- Executable should be build with optimization.
- Use a reasonable sized data set.



# Intel VTune Amplifier XE – Create Project

## ■ Set environment variables



# Amplifier XE – Algorithm Analysis

## Different Analysis Types

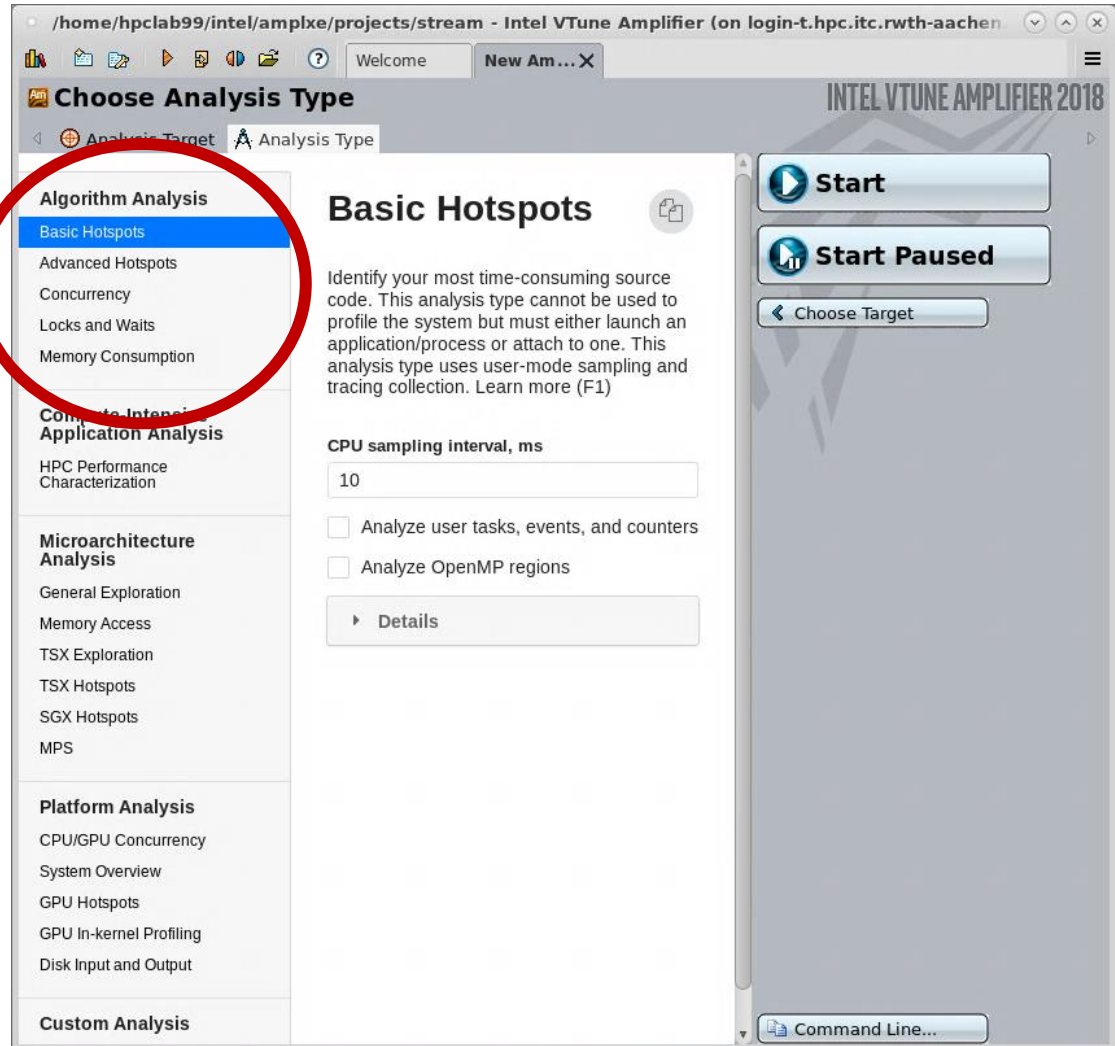
**Basic Hotspots:** Identify most time-consuming source code

**Advanced Hotspots:** Uses additional OS kernel support and analyses the CPI (Cycles Per Instruction) metric

**Concurrency:** Identify synchronization overhead and potential candidates for parallelism

**Locks and Waits:** Identify waiting times and discover how this affects your application performance

**Memory consumption**

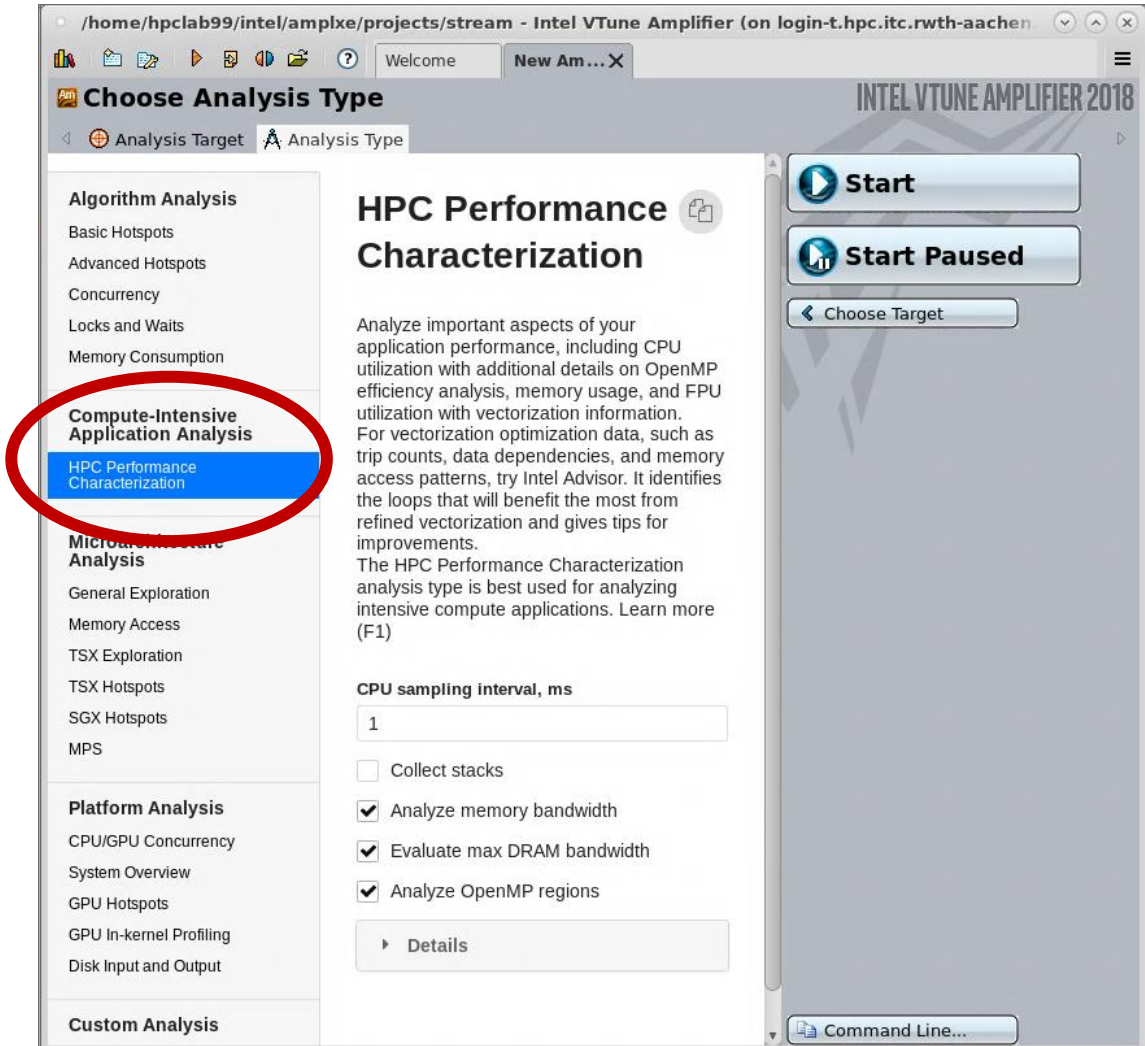


# Amplifier XE – HPC Performance Characterization

## Different Analysis Types

### HPC Performance Characterization: Analyze

- CPU utilization
- OpenMP efficiency
- memory usage
- FPU utilization
- vectorization

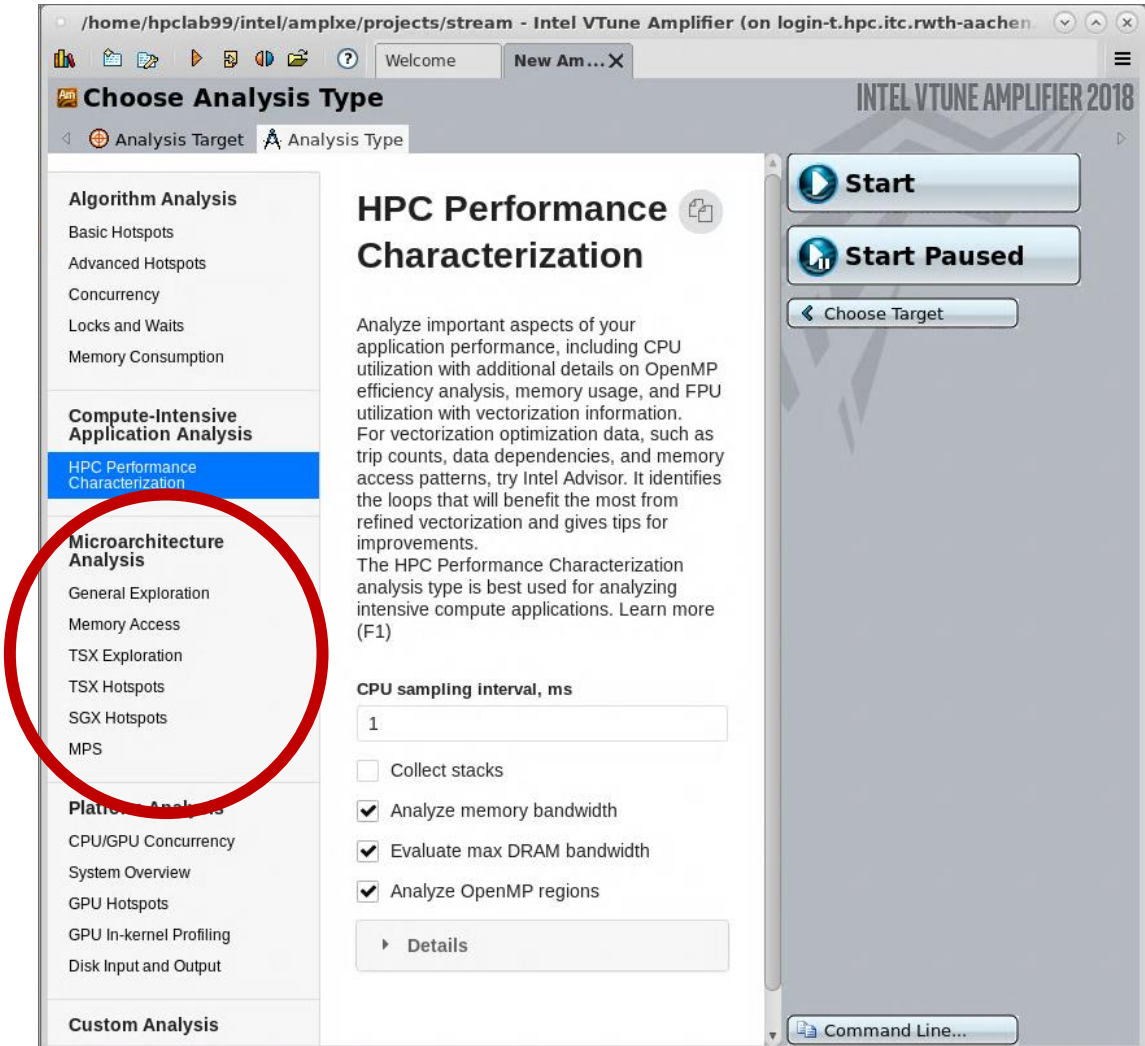


# Amplifier XE – Microarchitecture Analysis

## Different Analysis Types


### Microarchitecture Analysis:

- Based on hardware sampling
- Identify memory access related issues (e.g., NUMA effects)
- Analyze Intel Transactional Synchronization Extensions (TSX)
- MPI Performance Snapshot (MPS)



# ~~Live Demo~~

Intel VTune Amplifier XE



**Warning:** Intel Vtune Amplifier 2018 will crash the machine, if HW performance counters are used. Please load module `intelvtune/XE2017-u00` instead.



# Amplifier XE – Hotspot Analysis

**Basic Hotspots** Hotspots by CPU Usage viewpoint (change)

Collection Log Analysis Target Analysis Type Summary Bottom-up Caller/Callee Top-dov

Elapsed Time: 3.261s

CPU Time: 77.169s

Total Thread Count: 24

Paused Time: 0s

**Top Hotspots**

This section lists the most active functions in your application. Optimizing these hotspot functions typically results in improving overall application performance.

Function	Module	CPU Time
main_V\$omp\$parallel_for@241	stream	20.939s
main_V\$omp\$parallel_for@251	stream	20.391s
__intel_avx_rep_memcpy	libintlc.so.5	14.363s
main_V\$omp\$parallel_for@231	stream	13.897s
__kmp_fork_barrier	libiomp5.so	6.118s
[Others]		1.462s

**CPU Usage Histogram**

**Collection and Platform Info**

This section provides information about this collection, including result set size and collection platform data.

Application Command Line: /home/hpclub99/stream/stream

Environment Variables: OMP\_NUM\_THREADS=24; OMP\_PROC\_BIND=1; OMP\_PLACES=cores

Operating System: 3.10.0-693.17.1.el7.x86\_64 NAME="CentOS Linux" VERSION="7 (Core)" ID="centos" ID\_LIKE="rhel fedora" VERSION\_ID="7" PRETTY\_NAME="CentOS Linux 7 (Core)" ANSI\_COLOR="0;31" CPE\_NAME="cpe:/o:centos:centos:7" HOME\_URL="https://www.centos.org/" BUG\_REPORT\_URL="https://bugs.centos.org/" CENTOS\_MANTISBT\_PROJECT="CentOS-7" CENTOS\_MANTISBT\_PROJECT\_VERSION="7" REDHAT\_SUPPORT\_PRODUCT="centos" REDHAT\_SUPPORT\_PRODUCT\_VERSION="7"

Computer Name: login-t.hpc.itc.rwth-aachen.de

Result Size: 4 MB

Collection start time: 13:11:09 07/03/2018 UTC

Collection stop time: 13:11:12 07/03/2018 UTC

Collector Type: User-mode sampling and tracing

## Summary:

- 1 General Timing Information
- 2 Top Hotspots
- 3 Platform Information

Tool marks locations which might have a performance issue with a red flag.



# Amplifier XE – Hotspot Analysis

## CPU Usage Histogram

- Show CPU Utilization

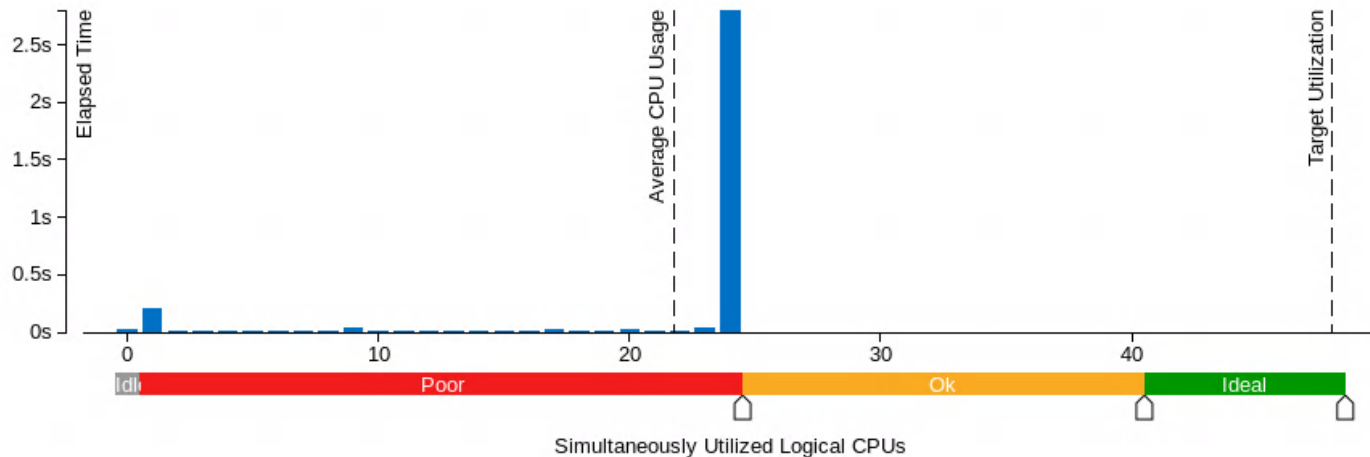
- Tool counts logical cores

→ Hyperthreads are not used in most HPC application

→ Use slider to change

### ⌵ CPU Usage Histogram 📄

This histogram displays a percentage of the wall time the specific number of CPUs were running simultaneously. Spin and Overhead time adds to the Idle CPU usage value.



# Amplifier XE – Hotspot Analysis

## CPU Usage Histogram

- Show CPU Utilization

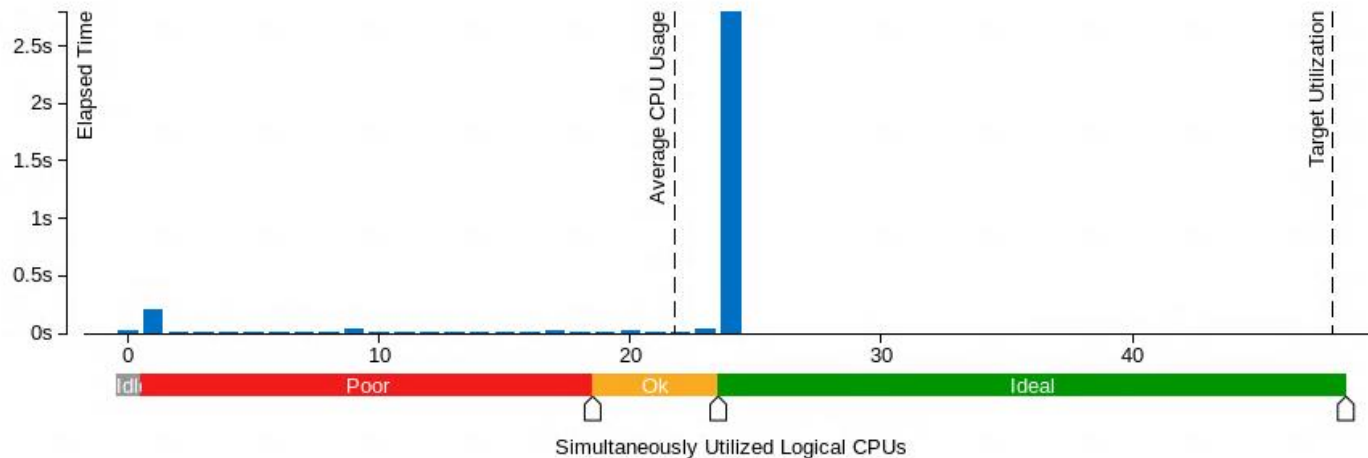
- Tool counts logical cores

→ Hyperthreads are not used in most HPC application

→ Use slider to change

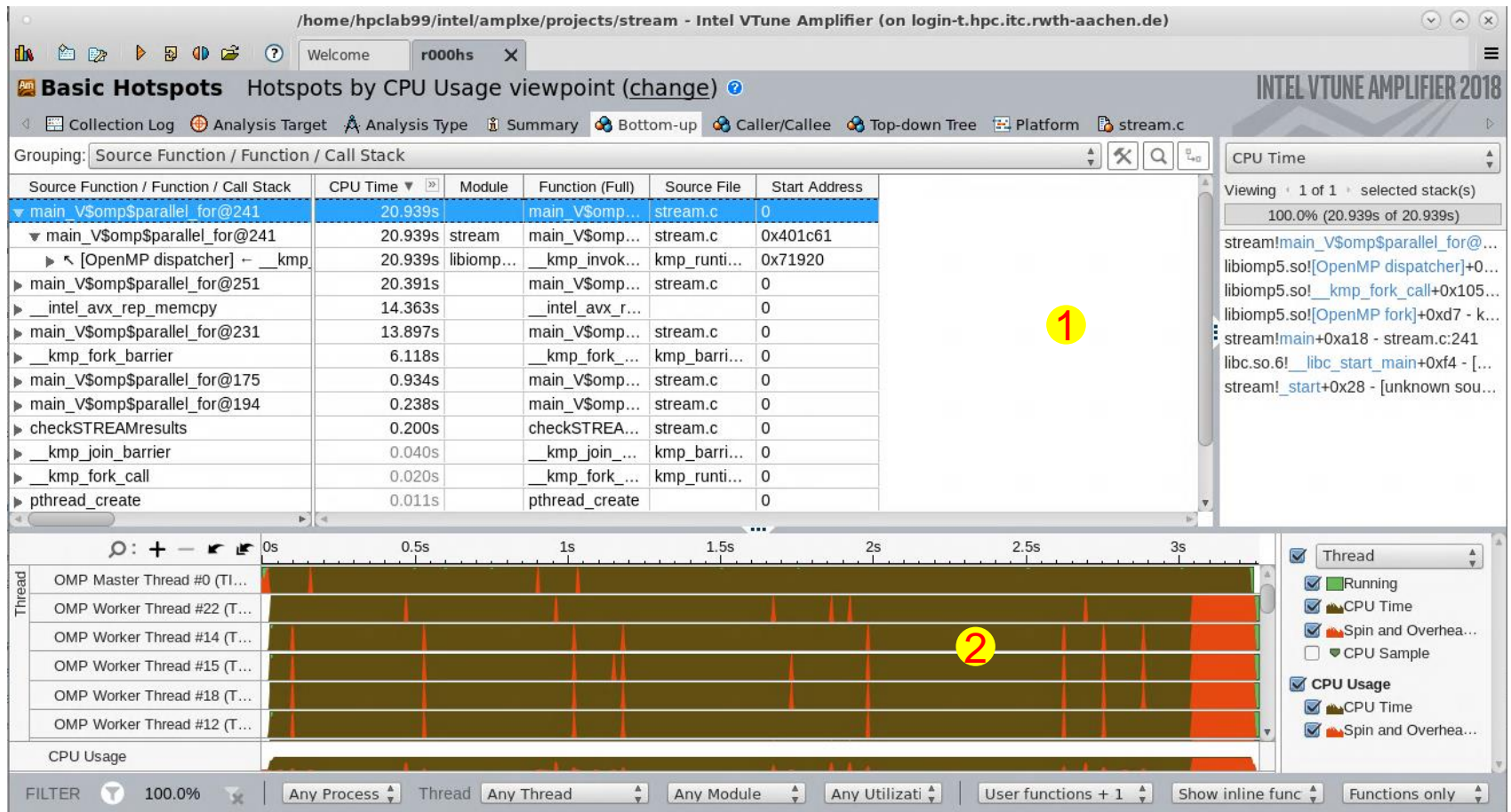
### ⌵ CPU Usage Histogram

This histogram displays a percentage of the wall time the specific number of CPUs were running simultaneously. Spin and Overhead time adds to the Idle CPU usage value.



# Amplifier XE – Hotspot Analysis

- 1 Function Summary
- 2 Timeline View

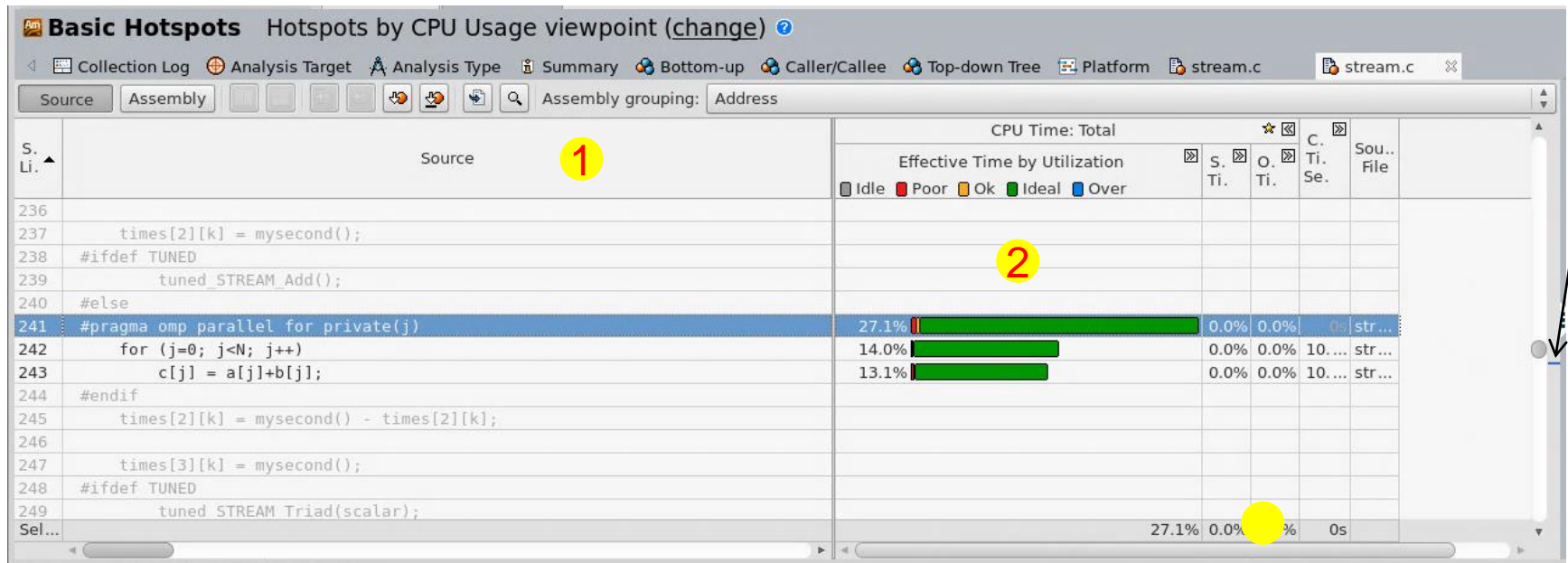


# Amplifier XE – Hotspot Analysis

Double clicking on a function opens source code view.

- 1 Source Code View (only if compiled with -g)
- 2 CPU Utilization (would be all “poor”, if slider in histogram is unmodified)

Hotspots



# Hardware Performance Counters

## ■ Hardware Counters of our Intel Nehalem Processor:

[illegible]

## L1I.HITS:

Counts all instruction fetches that hit the L1 instruction cache.

```

S.STLB_HIT,DTLB_LOAD_MISSES,PDE_MIS,
OPS_ISSUED.STALLED_CYCLE,UOPS_ISSUED.FUSED,
RETIRED.UNCA,FP_COMP_OPS_EXE.X87,
E_SING,FP_COMP_OPS_EXE.SSE.D0U,
T_128.SHUFFLE_MOV_LOAD_DISPATCH.RS,
PED,INST_QUEUE_WRITE_CYCLES,LSO_OVERFLOW,
TS.FETCHES,L2_RQSTS.PREFETCH_HIT,
M_,L2_DATA_RQSTS.DEMAND.ME,
STATE,L2_WRITE.RFO.S.STATE,
WRITE.LOCK.HIT,L2_WRITE.LOCK.MESI,
CACHE_LD_1.STATE,L1D_CACHE_LD_S.STATE,
HIT,L1D_CACHE_LOCKS.STATE,
TLB_HIT,DTLB_MISSES,PDE_MISS,
CT,L1D_CACHE_PREFETCH_LOCK,
MISSES.WALK_COMPLETE,ILD_STALL_LCP,ILD_STALL_MRU,
EXEC.RETURN_NEA,BR_INST_EXEC.DIRECT_NEAR,
EXEC.NON_CALLS,BR_MISP_EXEC.RETURN_NEA,
RESOURCE_STALLS.RS_FULL_RESOURCE_STALLS.STORE,
TLB_FLUSH.OFFCORE_REQUESTS.L1D_WR,
S,UOPS_EXECUTED.PORT015,
RE_RESPONSE_1,INST_RETIRED.ANY_P,
M.MEM_ORDE,MACHINE_CLEAR.SMC,
RED.PACKED,SSEX_UOPS_RETIRED.SCALAR,
HIT,MEM_LOAD_RETIRED.L3_UN,
MX,FP_MMX_TRANS.ANY,MACRO_INSTS.DECODED,
S.SCOREBOARD,RAT_STALLS.ANY,
L2_TRANSACTIONS.LOAD,L2_TRANSACTIONS.RFO,
IN.S.STATE,L2_LINES.IN.E.STATE,L2_LINES.IN.ANY,
FULL_STALL_CYCLES,FP_ASSIST.ALL,
PUID,SIMD_INT_64.PACKED.ARITH,
EMPTY,NOT_EMPTY,UINC_GQ_CYCLES.NOT_EMPTY,
WRITE_TRAC,UINC_GQ_ALLOC.PEER_PROBE,
_,UINC_GQ_DATA.TO_L3,UINC_GQ_DATA_TO_CORES,
TO_LOCAL_H,UINC_SNP_RESP.TO_REMOTE,
REMOTE,UINC_L3_HITS.READ,UINC_L3_HITS.WRITE,
STATE,UINC_L3_LINES.IN.S.STATE,
UINC_L3_LINES.OUT.F.STATE,
LOCAL_,UINC_QHL_REQUESTS.LOCAL_,
S,NOT_EMPTY,UINC_QHL_OCCUPANCY.IOH,
FLICT_CYCLES,UINC_QHL_CONFLICT_CYCLES.,
ULL.WRI,UINC_QMC_NORMAL_FULL.WRI,
C.FULL.WRITE.C,UINC_QMC_BUSY.READ.CH0,
_,UINC_QMC_OCCUPANCY.CH1,
DRMAL_READS.C,UINC_QMC_NORMAL_READS.C,
_PRIORITY_RE,UINC_QMC_CRITICAL_PRIORITY,
ULL.CH2,UINC_QMC_WRITES.FULL.ANY,
QMC_CANCEL.ANY,UINC_QMC_PRIORITY_UPDATE,
ALLED_SINGL,UINC_QPI_TX_STALLED_SINGL,
LLED_MULTI,UINC_QPI_TX_STALLED_MULTI,
DER.BUSY.LI,UINC_QPI_TX_HEADER.BUSY.LI,
DRAM_PAGE_CLOSE.CH1,UINC_DRAM_PAGE_CLOSE.CH2,
.CH1,UINC_DRAM_READ.CAS.AUTO,
CAS.AUTO,UINC_DRAM_WRITE.CAS.CH2,

```

BR MISP EXEC.COND:

Counts the number of mispredicted conditional near branch instructions executed, but not necessarily retired.

## Derived Metrics

### ■ **Clock cycles per Instructions (CPI)**

- CPI indicates if the application is utilizing the CPU or not
- Take care: Doing “something” does not always mean doing “something useful”.

### ■ **Floating Point Operations per second (FLOPS)**

- How many arithmetic operations are done per second?
- Floating Point operations are normally really computing and for some algorithms the number of floating point operations needed can be determined.

# Amplifier XE – Hardware Counter

## ① CPI rate (Clock cycles per instruction):

- In theory modern processors can finish 4 instructions in 1 cycle, so a CPI rate of 0.25 is possible
- A value between 0.25 and 1 is often considered as good for HPC applications
- Determine with “Advanced Hotspot” analysis

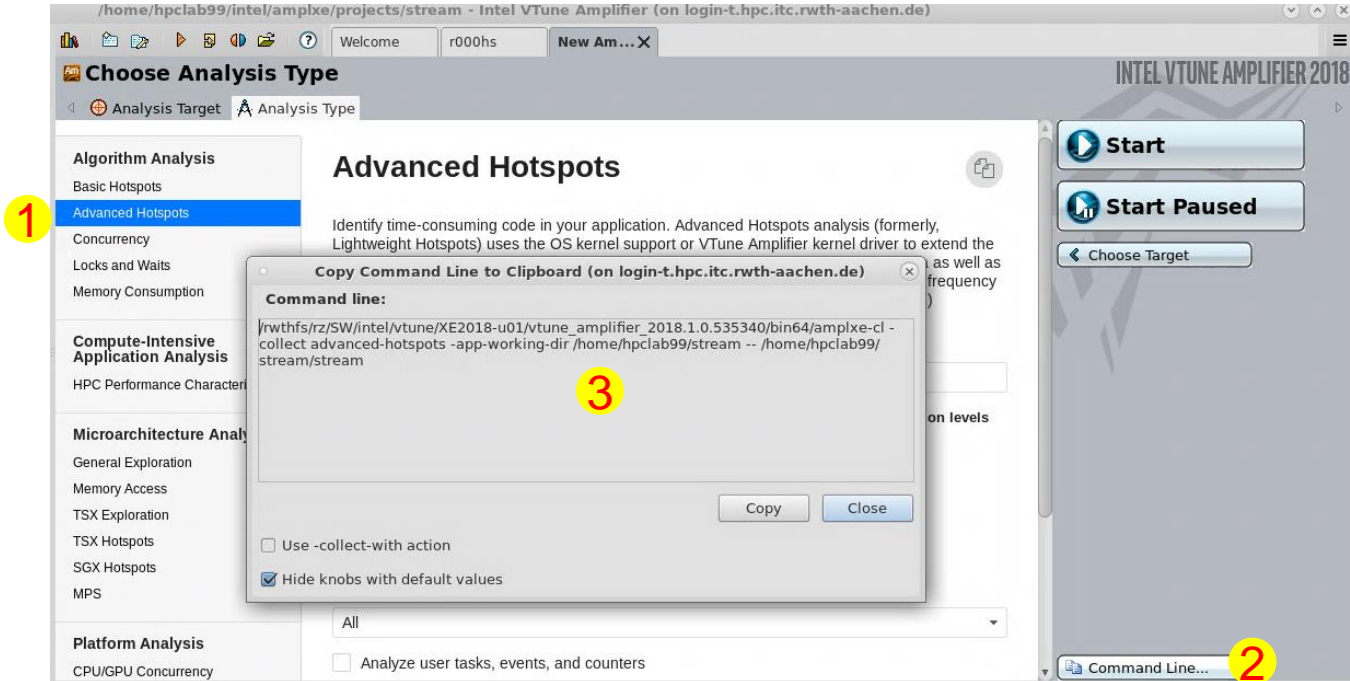
⌵	<b>Elapsed Time <sup>?</sup>: 3.327s</b>	
⌵	<b>CPU Time <sup>?</sup>:</b>	<b>76.810s</b>
	<u>Instructions Retired:</u>	45,579,600,000
①	<u>CPI Rate <sup>?</sup>:</u>	4.213 
	<u>CPU Frequency Ratio <sup>?</sup>:</u>	1.136
	<u>Total Thread Count:</u>	24
	<u>Paused Time <sup>?</sup>:</u>	0s



# VTune Amplifier in Command Line

## ■ Real-world analysis should be executed in the batch system

- 1 → Choose analysis type
- 2 → Push “Command Line...”
- 3 → Copy & paste command line and use in batch script



## Correctness:

- Data Races are very hard to find, since they do not show up every program run.
- Intel Inspector XE or ThreadSanitizer help a lot in finding these errors.
- Use really small datasets, since the runtime increases significantly.
- If possible use non optimized code.

## Performance:

- Start with simple performance measurements like hotspots analyses and then focus on these hot spots.
- In OpenMP applications analyze the waiting time of threads. Is the waiting time balanced?
- Hardware counters might help for a better understanding of an application, but they might be hard to interpret.