Tutorial: Finding Hotspots

Intel® VTune™ Amplifier for Linux* OS

Fortran Sample Application Code

Legal Information

Contents

Legal Information

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.

Intel disclaims all express and implied warranties, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, as well as any warranty arising from course of performance, course of dealing, or usage in trade.

This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest forecast, schedule, specifications and roadmaps.

The products and services described may contain defects or errors which may cause deviations from published specifications. Current characterized errata are available on request.

Intel processor numbers are not a measure of performance. Processor numbers differentiate features within each processor family, not across different processor families. Go to: Learn About Intel® Processor Numbers

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.

Cilk, Intel, the Intel logo, Intel Atom, Intel Core, Intel Inside, Intel NetBurst, Intel SpeedStep, Intel vPro, Intel Xeon Phi, Intel XScale, Itanium, MMX, Pentium, Thunderbolt, Ultrabook, VTune and Xeon are trademarks of Intel Corporation in the U.S. and/or other countries.

*Other names and brands may be claimed as the property of others.

Microsoft, Windows, and the Windows logo are trademarks, or registered trademarks of Microsoft Corporation in the United States and/or other countries.

Overview


Discover how to use the Basic Hotspots, Concurrency, and Locks and Waits analysis types of the Intel® VTune™ Amplifier to identify *hotspots* - the most time-consuming program units, understand how effectively your code is using available cores, and discover causes of ineffective utilization.

About This Tutorial	This tutorial uses the sample `nqueens_parallel` application and guides you through basic workflow steps required to analyze the code for hotspots, parallelism, and locks.
Estimated Duration	10-15 minutes.
Learning Objectives	After you complete this tutorial, you should be able to: • Choose an analysis target. • Choose an analysis type. • Run the Basic Hotspots analysis to locate most time-consuming functions in an application. • Analyze the function call flow and threads. • Analyze the source code to locate the most time-critical code lines. • Run the Concurrency analysis to identify function candidates for parallelization. • Run the Locks and Waits analysis to identify where synchronization objects spent too much CPU time waiting. • Compare results before and after optimization.
More Resources	• Intel VTune Amplifier tutorials (HTML, PDF): https://software.intel.com/en-us/articles/intel-vtune-amplifier-tutorials/ • Intel VTune Amplifier support page: https://software.intel.com/en-us/intel-vtune-amplifier-xe-support/ • Intel Parallel Studio XE support page: https://software.intel.com/en-us/intel-parallel-studio-xe/

1 Navigation Quick Start

Intel® VTune™ Amplifier provides information on code performance for users developing serial and multithreaded applications on Windows*, Linux*, and OS X* operating systems. VTune Amplifier helps you analyze algorithm choices and identify where and how your application can benefit from available hardware resources.

VTune Amplifier XE Access

VTune Amplifier installation includes shell scripts that you can run in your terminal window to set up environment variables:

1. From the installation directory, type `source amplxe-vars.sh`.

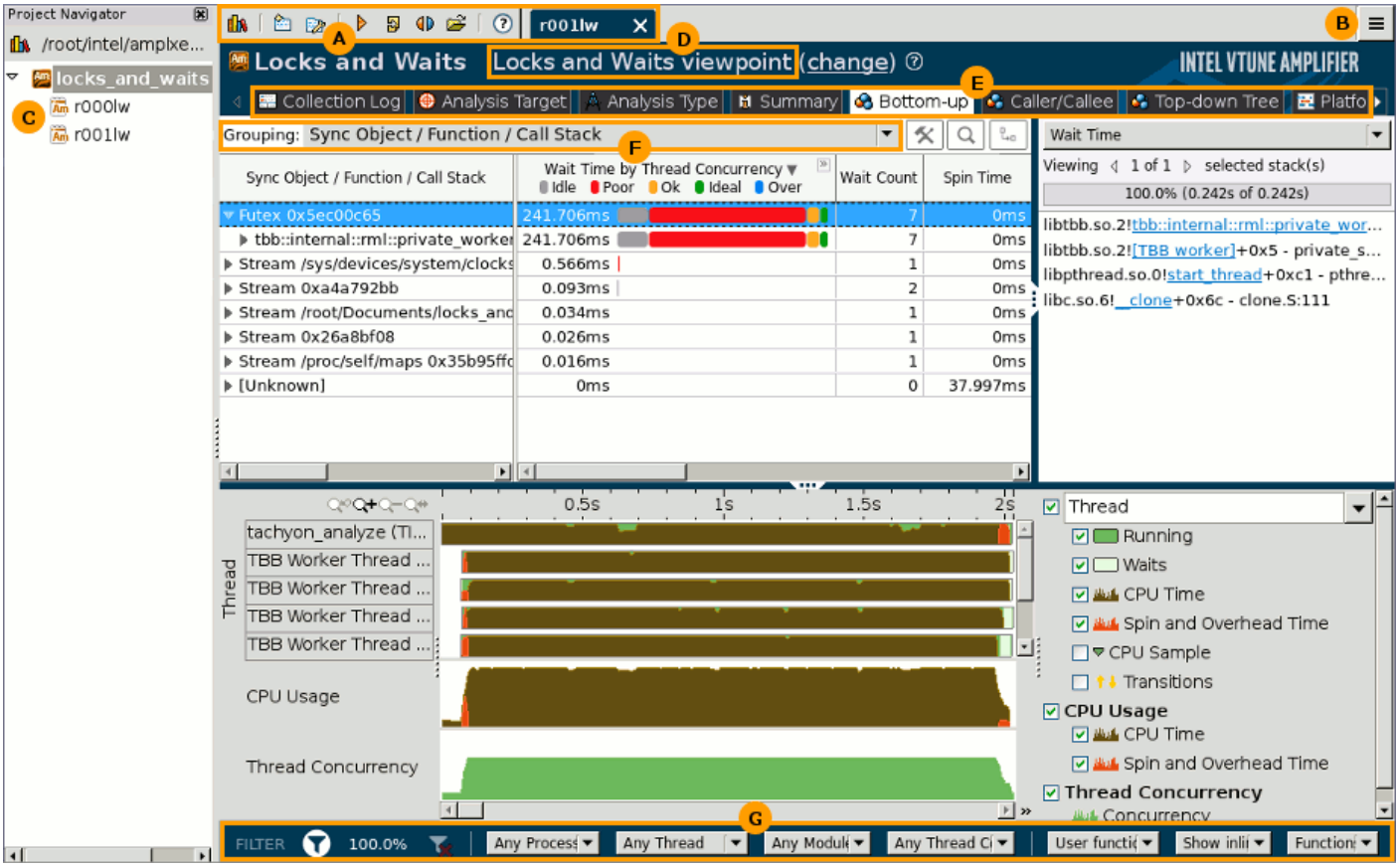
   This script sets the PATH environment variable that specifies locations of the product graphical user interface utility and command line utility.

   NOTE
   The default installation directory is:

   - For root users: `/opt/intel/vtune_amplifier_xe_version`
   - For non-root users: `$HOME/intel/vtune_amplifier_xe_version`

2. Type `amplxe-gui` to launch the product graphical interface.

VTune Amplifier GUI



A Configure and manage projects and results, and launch new analyses from the primary toolbar. Click the **Configure Project** button on this toolbar and use the **Analysis Target** tab to manage result file locations. Newly completed and opened analysis results along with result comparisons appear in the results tab for easy navigation.

B Use the VTune Amplifier menu to control result collection, define and view project properties, and set various options.

C The **Project Navigator** provides an iconic representation of your projects and analysis results. Click the **Project Navigator** button on the toolbar to enable/disable the **Project Navigator**.

D Click the **(change)** link to select a *viewpoint*, a preset configuration of windows/panes for an analysis result. For each analysis type, you can switch among several viewpoints to focus on particular performance metrics. Click the yellow question mark icon to read the viewpoint description.

E Switch between window tabs to explore the analysis type configuration options and collected data provided by the selected viewpoint.

F Use the **Grouping** drop-down menu to choose a granularity level for grouping data in the grid.

G Use the filter toolbar to filter out the result data according to the selected categories.

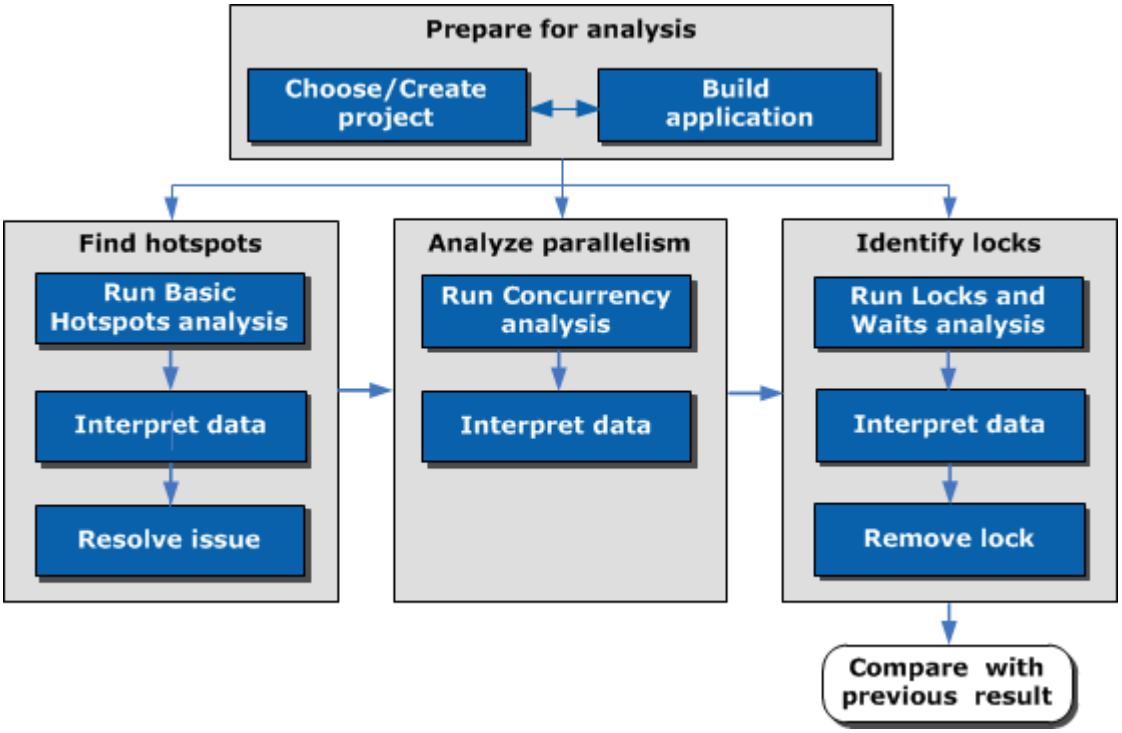
See Also

Click here for more Getting Started Tutorials

2

Finding Hotspots

You can use the Intel® VTune™ Amplifier to identify and analyze hotspot functions in your serial or parallel application by performing a series of steps in a workflow. This tutorial guides you through these workflow steps while using a sample multithreaded application named `nqueens_parallel`.



Step 1: Prepare for analysis	Build an application to analyze for hotspots and create a new VTune Amplifier project
Step 2: Find hotspots	• Choose and run the Basic Hotspots analysis. • Interpret the result data. • Resolve issue.
Step 3: Analyze parallelism	• Choose and run the Concurrency analysis. • Interpret the result data.
Step 4: Identify locks	• Choose and run the Locks and Waits analysis. • Interpret the result data. • Remove lock.
Step 5: Check your work	Re-build the target, re-run the Locks and Waits analysis, and compare the result data before and after optimization.

Build Application and Create New Project

Before you start analyzing your application target for hotspots, do the following:

1. Get software tools.
2. Build application in the release mode.
3. Create a performance baseline.
4. Create a VTune Amplifier project.

Get Software Tools

You need the following tools to try tutorial steps yourself using the `nqueens_fortran` sample application:

- Intel® VTune™ Amplifier, including sample applications
- `.tgz` file extraction utility
- Supported Fortran compiler (see Release Notes for more information)

Acquire Intel VTune Amplifier

If you do not already have access to the VTune Amplifier, you can download an evaluation copy from http://software.intel.com/en-us/articles/intel-software-evaluation-center/.

Install and Set Up VTune Amplifier Sample Applications

1. Copy the `nqueens_fortran.tgz` file from the `<install-dir>/samples/<locale>/Fortran` directory to a writable directory or share on your system.

   NOTE
   The default installation path for the VTune Amplifier XE is `/opt/intel/vtune_amplifier_xe_`*`version`*. For the VTune Amplifier for Systems, the default `<`*`install_dir`*`>` is:
   - For super-users: `/opt/intel/system_studio_`*`version`*`/vtune_amplifier_for_systems`
   - For ordinary users: `$HOME/intel/system_studio_`*`version`*`/vtune_amplifier_for_systems`

2. Extract the sample from the tgz file.

   NOTE
   - Samples are non-deterministic. Your screens may vary from the screen captures shown throughout this tutorial.
   - Samples are designed only to illustrate the VTune Amplifier features; they do not represent best practices for creating code.

Build the Target in the Release Mode

Build the target in the Release mode with full optimizations, which is recommended for performance analysis. For this tutorial, Intel® Fortran Compiler is used to build the application.

1. Browse to the directory where you extracted the sample code (for example, `/home/fortran/linux`). Make sure this directory contains `Makefile`.
2. Clean up all the previous builds as follows:

   ```
   $ make clean
   ```
3. Build your target in the release mode as follows:

   ```
   $ make
   ```

The `nqueens_parallel` application is built.

Create a Performance Baseline

Run the application to create a performance baseline that will be used to identify optimization you achieve during performance tuning with the VTune Amplifier.

NOTE
Before you start the application, minimize the amount of other software running on your computer to get more accurate results.

1. Run `nqueens_parallel` with the task size of 15. For example:

   ```
   $ ./nqueens_parallel 15
   ```

   ```
   Starting nqueens solver for size 15 with 1 thread(s)
   Number of solutions: 2279184
   Correct Result!
   Calculations took 256710ms.
   ```

2. Note the execution time displayed in the shell window caption. In the example above, the execution time is 256710 milliseconds.

NOTE

- Run the application several times, note the execution time for each run, and use the average number. This helps to minimize skewed results due to transient system activity.
- The screenshots and execution time data provided in this tutorial are created on a system with 4 CPU cores. Your data may vary depending on the number and type of CPU cores on your system.

Create a VTune Amplifier Project

1. Set the `EDITOR` or `VISUAL` environment variable to associate your source files with the code editor (like emacs, vi, vim, gedit, and so on). For example:

   ```
   $ export EDITOR=gedit
   ```

2. Run the `amplxe-gui` script launching the VTune Amplifier GUI.
3. Create a new project via **New > Project...**.
4. Specify the project name `nqueens` that will be used as the project directory name and click **Create Project**.

   VTune Amplifier creates the `tachyon` project directory under the `$HOME/intel/amplxe/projects` (for VTune Amplifier XE) or `$HOME/intel/amplsys/projects` (for VTune Amplifier for Systems) directory and opens the **Choose Target and Analysis Type** window with the **Analysis Target** tab active.
5. From the left pane, select the **local** target system from the **Accessible Targets** group. From the right pane select the **Launch Application** target type.
6. Specify and configure your target as follows:

   - For the **Application** field, browse to: `<sample_code_dir>`, for example: `/home/vtune/nqueens_fortran/linux/nqueens_parallel`.
   - In the **Application parameters** field, specify the task size for this target: 15.
7. Click **Choose Analysis** to select an analysis type.

Recap

You built the target in the Release mode, created the performance baseline, and created the VTune Amplifier project for your analysis target. Your application is ready for analysis.

Key Terms

- Baseline
- Target

Next Step

Run Basic Hotspots Analysis

Run Basic Hotspots Analysis

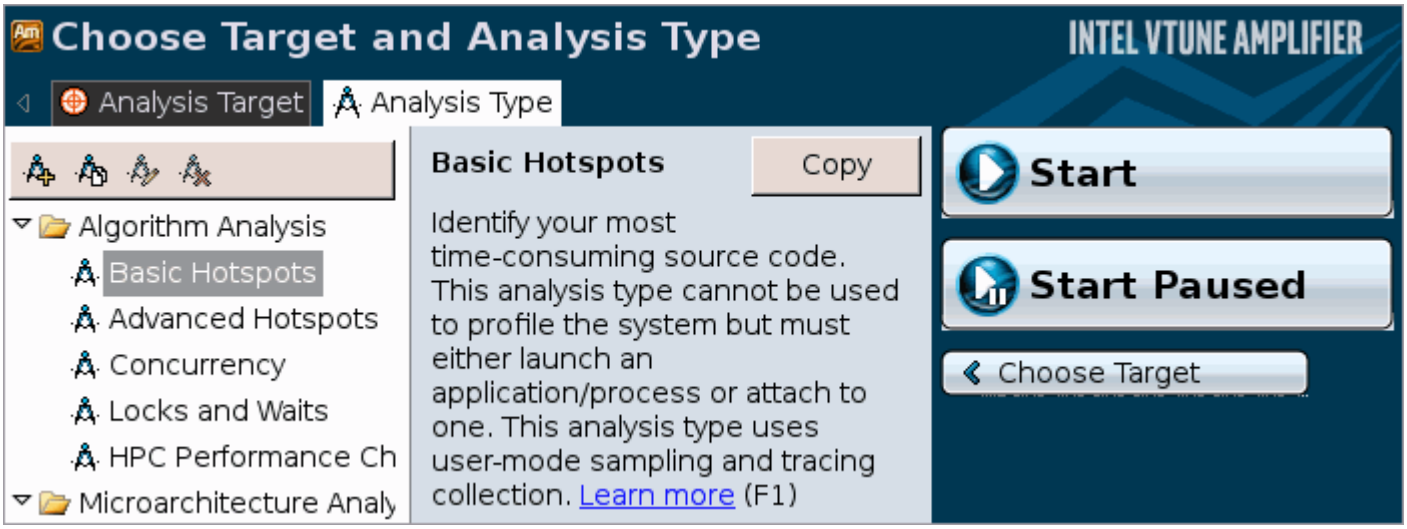
Before running an analysis, choose a configuration level to influence Intel® VTune™ Amplifier analysis scope and running time. In this tutorial, you run the Basic Hotspots analysis to identify the hotspots that took much time to execute.

To run an analysis:

1. From the VTune Amplifier toolbar, click the **New Analysis** button.

   The **New Amplifier Result** tab opens with the **Analysis Type** window active.
2. On the left pane of the **Analysis Type** window, locate the analysis tree and select **Algorithm Analysis > Basic Hotspots**.

   The right pane is updated with the predefined settings for the Basic Hotspots analysis.
3. From the right pane, select the `Analyze OpenMP regions` checkbox.
4. Click the **Start** button on the right command bar.



VTune Amplifier launches the `nqueens_parallel` application that makes calculations, displays the execution time, and exits. VTune Amplifier finalizes the collected results and opens the analysis results in the **Hotspots by CPU Usage** viewpoint.

To make sure the performance of the application is repeatable, go through the entire tuning process on the same system with a minimal amount of other software executing.

> **NOTE**
> This tutorial explains how to run an analysis from the VTune Amplifier graphical user interface (GUI). You can also use the VTune Amplifier command-line interface (`amplxe-cl` command) to run an analysis. If you run the example program from the VTune Amplifier command-line interface, specify 15 as a command argument. For more details, check the *Command-line Interface Support* section of the VTune Amplifier Help.

Key Terms

- Elapsed time
- Finalization
- Hotspot
- Viewpoint

Next Step

Interpret Results

Interpret Results

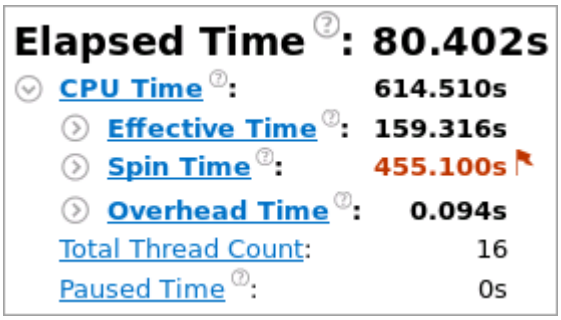
When the sample application exits, the Intel® VTune™ Amplifier finalizes the results and opens the **Hotspots by CPU Usage** viewpoint where each window or pane is configured to display code regions that consumed a lot of CPU time. To interpret the data on the sample code performance, do the following:

1. Explore application-level performance.
2. Analyze the most time-consuming functions.
3. Identify the hotspot code region.

NOTE
The screenshots and execution time data provided in this tutorial are created on a system with 4 CPU cores. Your data may vary depending on the number and type of CPU cores on your system.

Explore Application-level Performance

Start analysis with the **Summary** window that opens by default when data collection completes. To interpret the data, hover over the question mark icons to read the pop-up help and better understand what each performance metric means.



The **Elapsed Time** metric shows the duration of the collection including Paused Time. You may use this metric as one of the basic performance indicators.

Note that **CPU Time** for the sample application is equal to 614.510 seconds. It is the sum of CPU time for all application threads. **Total Thread Count** is 16, so the sample application is multi-threaded.

The `nqueens_parallel` application uses the OpenMP* threading model. VTune Amplifier analyzes performance in OpenMP parallel regions as well as serial code performance. The **OpenMP Analysis** section provides metrics based on the **Collection Time**, which is the wall time from the beginning to the end of collection, excluding Paused Time. The `nqueens_parallel` application ran serially only 0.464 seconds, which is 0.6% of Collection Time. According to the provided estimates, you can improve the efficiency of your code in parallel regions and get 40.422 seconds of performance gain (maximum estimate), which is 50.3% of Collection Time.

The **Top OpenMP Regions by Potential Gain** section displays the parallel region in the `nqueens_parallel` application that should be optimized.

Top OpenMP Regions by Potential Gain

This section lists OpenMP regions with the highest potential for performance improvement. The Potential Gain metric shows the elapsed time that could be saved if the region was optimized to have no load imbalance assuming no runtime overhead.

OpenMP Region	OpenMP Potential Gain	(%)	OpenMP Region Time
nqueens_IP_solve_omp parallel:16@/root/Documents/nqueens_fortran/linux/../src/nqueens_parallel.f90:159:163	40.421s	50.3%	79.938s

Clicking this region opens the **Bottom-up** window with the data grouped by **OpenMP Region** and detailed statistics for the hot regions.

Analyze the Most Time-consuming Functions

Click the **Bottom-up** tab to explore the **Bottom-up** pane. By default, the data in the grid is sorted by Function. You may change the grouping level using the **Grouping** drop-down menu at the top of the grid.

Click the **Effective Time bu Utilization** to sort the hotspots functions by effective time.

The `nqueens_IP_setqueen` function took 159.288 seconds to execute, ineffectively using CPU resources during all this time.

Basic Hotspots — Hotspots by CPU Usage viewpoint (change)

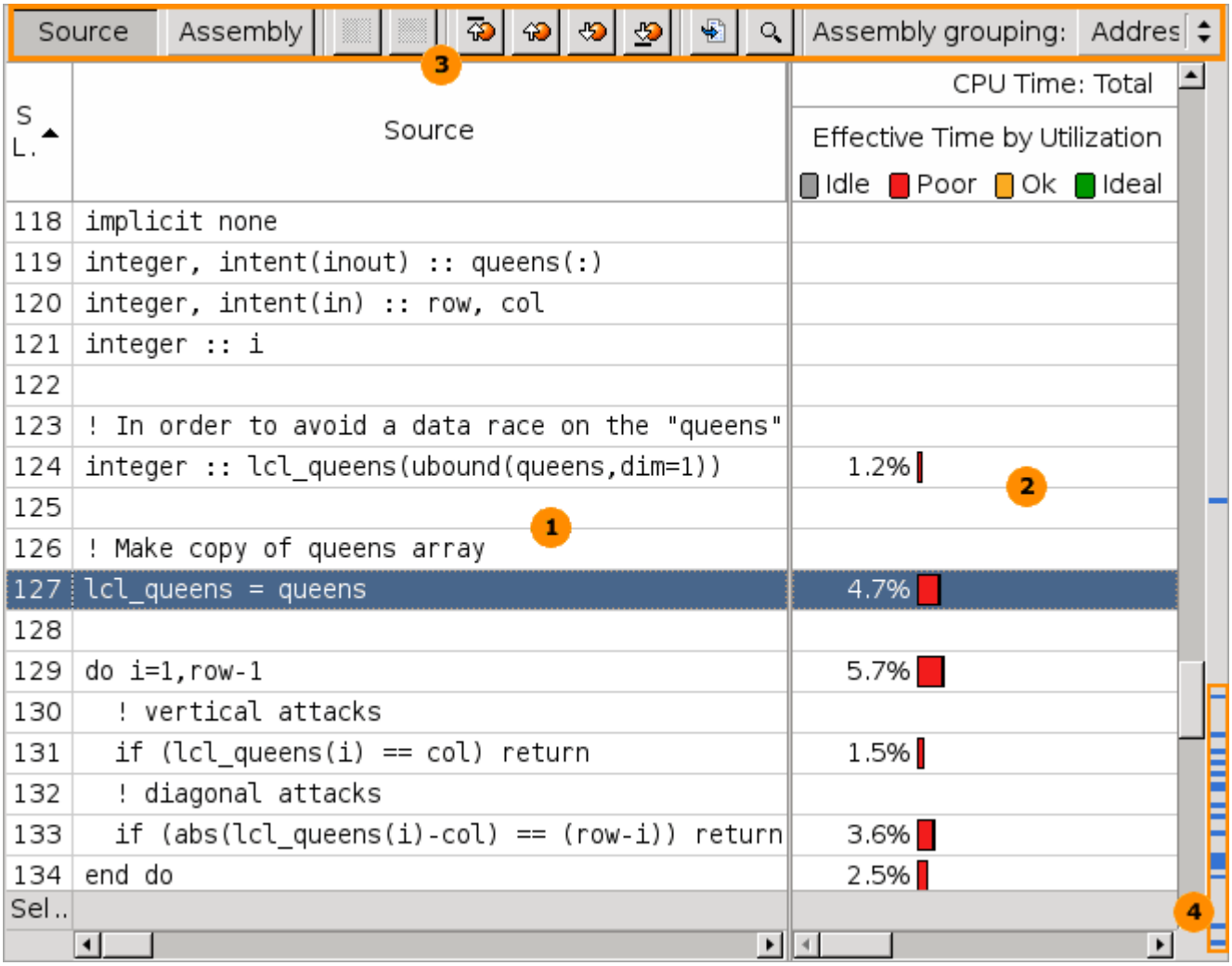
Collection Log | Analysis Target | Analysis Type | Summary | Bottom-up | Caller

Grouping: Function / Call Stack

Function / Call Stack	CPU Time: Effective Time by Utilization (Idle, Poor, Ok, Ideal, Over)	CPU Time: Spin Time	CPU Time: Overhead Time	Module
nqueens_IP_setqueen_	159.288s	0s	0s	nqueens_parallel
[Import thunk sched_yield]	0.018s	0s	0s	libiomp5.so
[Import thunk __tls_get_addr]	0.010s	0s	0s	libiomp5.so
__kmpc_critical	0s	454....	0.020s	libiomp5.so
__kmp_release_queuing_lock	0s	0.388s	0s	libiomp5.so
__kmpc_end_critical	0s	0.042s	0.032s	libiomp5.so
[Import thunk __kmp_wait_yie	0s	0.032s	0s	libiomp5.so
__kmp_get_global_thread_id_r	0s	0s	0.022s	libiomp5.so

Double-click the hotspot function to open the source and identify the most time-critical code lines.

Identify the Hotspot Code Region



The table below explains some of the features available in the **Source** window when viewing the Basic Hotspots analysis data.

1. **Source** pane displaying the source code of the application if the function symbol information is available. The hottest code line is highlighted. The source code in the **Source** pane is not editable.

   If the function symbol information is not available, the **Assembly** pane opens displaying assembler instructions for the selected hotspot function. To enable the **Source** pane, make sure to build the target properly.

2. Processor time attributed to a particular code line. If the hotspot is a system function, its time, by default, is attributed to the user function that called this system function.

3. **Source** window toolbar. Use the hotspot navigation buttons to switch between most performance-critical code lines. Hotspot navigation is based on the metric column selected as a Data of Interest. For the Basic Hotspots analysis, this is **CPU Time**. Use the **Source**/**Assembly** buttons to toggle the **Source**/**Assembly** panes (if both of them are available) on/off.

4. Heat map markers to quickly identify performance-critical code lines (*hotspots*). The bright blue markers indicate hot lines for the function you selected for analysis. Light blue markers indicate hot lines for other functions. Scroll to a marker to locate the hot code line it identifies.

By default, when you double-click the hotspot in the **Bottom-up** pane, the VTune Amplifier opens the source file related to this function with the hottest code line highlighted. For the `nqueens_IP_setqueen` function, this is the code line that is used to create a local copy of the `queens` array to avoid a data race. Click the  **Source Editor** button on the **Source** window toolbar to open the default code editor and work on optimizing the code.

> **NOTE**
> Depending on the sample code version, your source line numbers may slightly differ from the numbers provided in this tutorial.

Key Terms

- CPU time
- Viewpoint

Next Step

Resolve Issue

Resolve Issue

 You identified that the most time-consuming function is `nqueens_IP_setqueen`. If you click the  **Source Editor** button from the **Source** pane, the VTune Amplifier opens the source `nqueens_parallel.f90` file at the hotspot line in the default code editor. You see that the OpenMP* cycle is calling the recursive `setQueen` function that initializes the `queens` array. To avoid a data race, this array is copied in each thread (see line 127):

> **NOTE**
> Depending on the sample code version, your source line numbers may slightly differ from the numbers provided in this tutorial.

```
123   ! In order to avoid a data race on the "queens" array,
124   integer :: lcl_queens(ubound(queens,dim=1))
125
126   ! Make copy of queens array
127   lcl_queens = queens
128
129   do i=1,row-1
130     ! vertical attacks
131     if (lcl_queens(i) == col) return
132     ! diagonal attacks
133     if (abs(lcl_queens(i)-col) == (row-i)) return
134   end do
```

This means that the number of local copies is equal to the number of threads. Since the function is recursive, the array is also copied in every function call, which is unnecessary and creates a big overhead.

To resolve this issue, you may enable OpenMP to make a copy of the array per thread. To do this:

1. Comment out lines 124 and 127.

```
123   ! In order to avoid a data race on the "queens" array,
124   ! integer :: lcl_queens(ubound(queens,dim=1))
125
126   ! Make copy of queens array
127   ! lcl_queens = queens
```

2. Search and replace all `lcl_queens` entries with `queens`.
3. Edit line 159 to add the `PRIVATE(queens)` directive.

   This enables the OpenMP run-rime to create a private copy of the array for each thread.

```
158 ! Enable dynamic load scheduling
159 !$OMP PARALLEL DO PRIVATE(queens)
160   do i=1,size
161     ! try all positions in first row
162     call SetQueen (queens, 1, i)
163   end do
```

4. Save the changes made in the source file.
5. Browse to the directory where you extracted the sample code (for example, `/home/vtune/nqueens_fortran/linux`).
6. Rebuild your target in the release mode using the make command as follows:

   ```
   $ make clean
   ```

   ```
   $ make
   ```

   The `nqueens_parallel` application is rebuilt.
7. Run `nqueens_parallel` as follows:

   ```
   ./nqueens_parallel 15
   ```

```
Starting nqueens solver for size 15 with 16 thread(s)
Number of solutions: 2279184
Correct Result!
Calculations took 56566ms.
```

   System runs `nqueens_parallel`. Note that the execution time has reduced from 256710 ms to 56566 ms. This means that the proposed solution gives 200144 ms of CPU time reduction.

To identify other possible performance issues, you may run the Concurrency analysis and see how effectively your application is parallelized.

Next Step

Run Concurrency Analysis

Run Concurrency Analysis

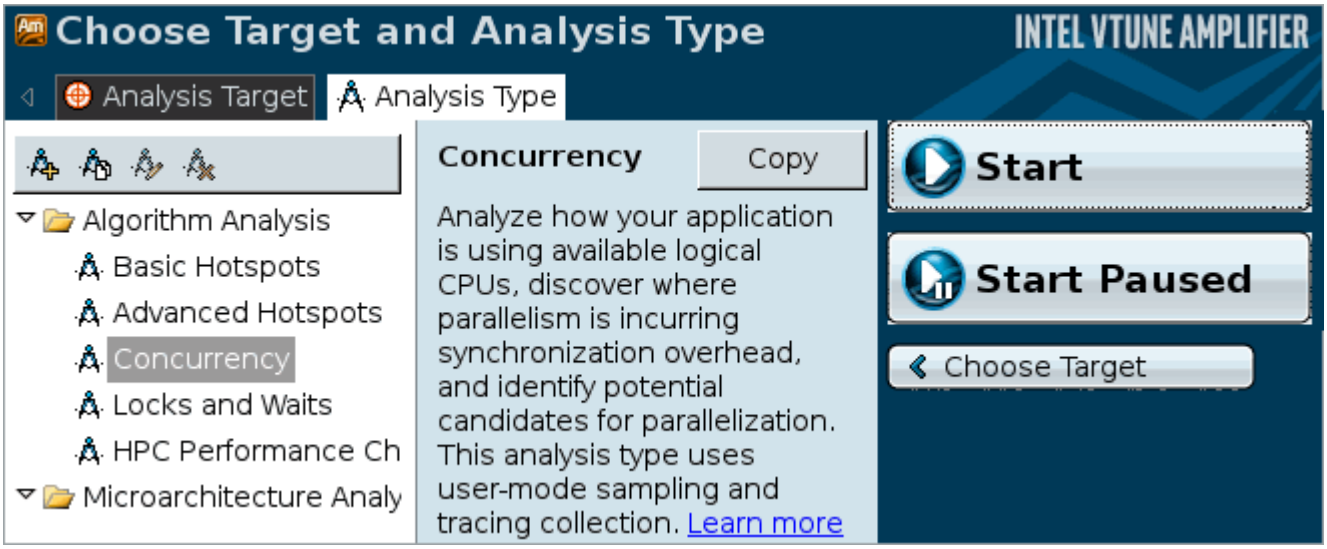
Run the Concurrency analysis to understand how effectively your application is parallelized.

To run an analysis:

1. From the VTune Amplifier toolbar, click the **New Analysis** button.

   New Amplifier Result tab opens with the **Analysis Type** window active.
2. On the left pane of the **Analysis Type** window, locate the analysis tree and select **Algorithm Analysis > Concurrency**.

   The right pane is updated with the predefined settings for the Concurrency analysis.
3. Select the **Analyze Intel runtimes and user synchronization** checkbox.
4. Click the **Start** button on the right command bar.



VTune Amplifier launches the `nqueens_parallel` application that makes calculations, displays the execution time, and exits. VTune Amplifier finalizes the collected results and opens the analysis results in the **Hotspots by CPU Usage** viewpoint.

To make sure the performance of the application is repeatable, go through the entire tuning process on the same system with a minimal amount of other software executing.

> **NOTE**
> This tutorial explains how to run an analysis from the VTune Amplifier graphical user interface (GUI). You can also use the VTune Amplifier command-line interface (`amplxe-cl` command) to run an analysis. For more details, check the *Command-line Interface Support* section of the VTune Amplifier Help.

Key Terms

- Finalization
- Viewpoint

Next Step

Interpret Concurrency Results

Interpret Concurrency Results

When the sample application exits, the Intel® VTune™ Amplifier finalizes the results and opens the **Hotspots by CPU Usage** viewpoint where each window or pane is configured to display data on application parallelism and usage of processor cores. To interpret the data on the sample code performance, do the following:

1. Explore application-level concurrency
2. Identify the most time-consuming function.

> **NOTE**
> The screenshots and execution time data provided in this tutorial are created on a system with 4 CPU cores. Your data may vary depending on the number and type of CPU cores on your system.

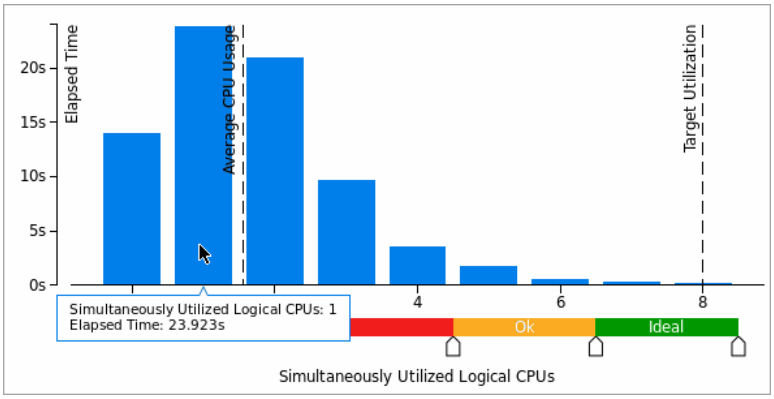
Explore Application-level Concurrency

Start analysis with the **Summary** window that opens by default when data collection completes. To interpret the data, hover over the question mark icons to read the pop-up help and better understand what each performance metric means.

You see that after optimization, the Elapsed time has increased from 80.402 seconds to 75.093 seconds.

NOTE
The Concurrency analysis adds an overhead to the application execution. The overhead often depends on the number of threads and synchronization objects used in the application. This is the reason why Elapsed time data provided in the **Summary** window may differ from the data reported after the application launch outside of the VTune Amplifier.

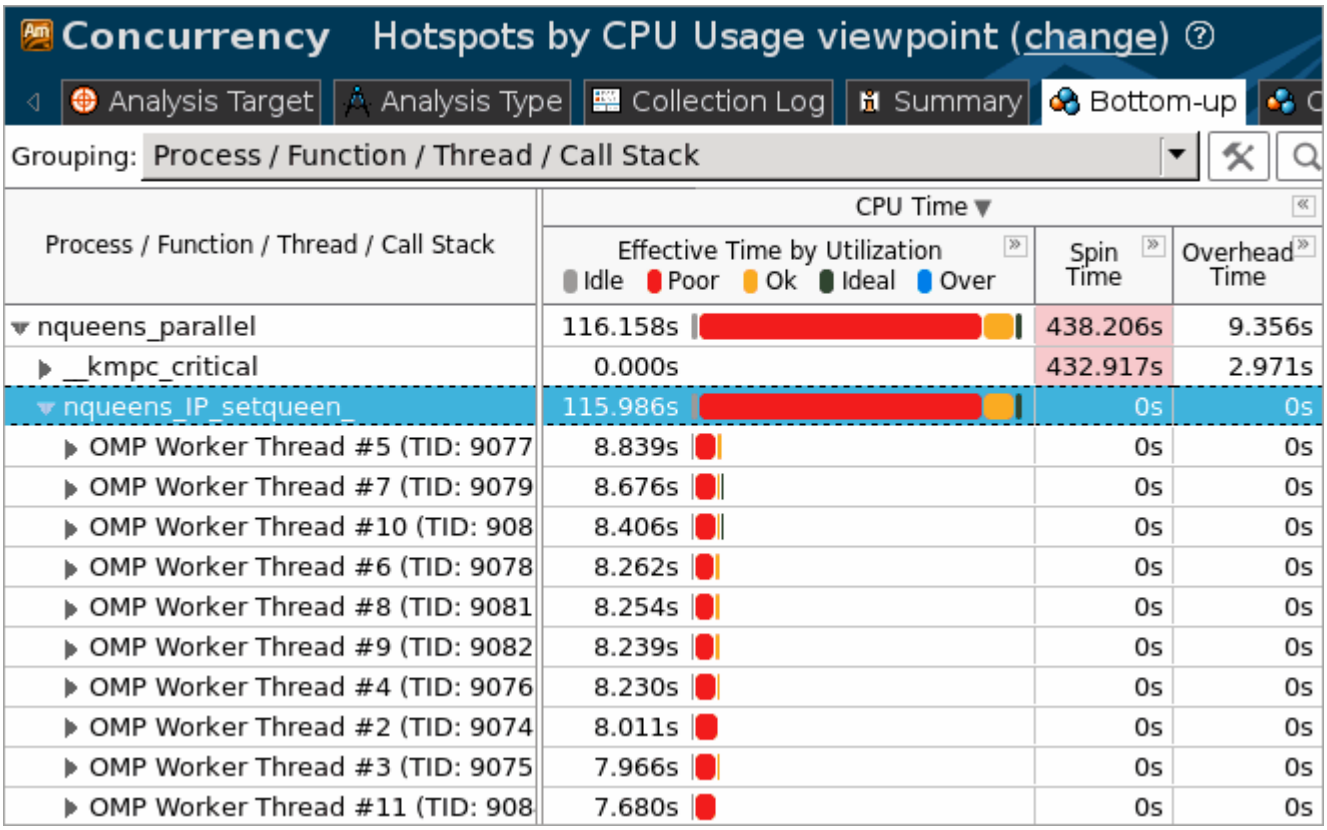
The **CPU Usage Histogram** shows that the average concurrency level of the sample application is about 2 while the target concurrency level for this application on the 4-core system is 8. If you hover over the highest bar, you see that this application has run 1 threads for almost 24 seconds, which is categorized by the VTune Amplifier as Poor processor utilization on this system.



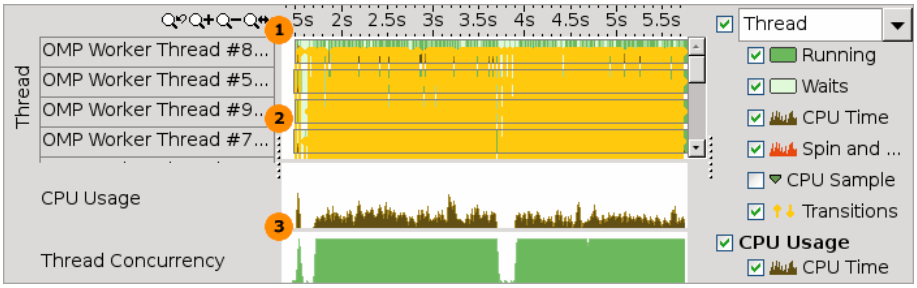
Identify the Most Time-consuming Function

Click the **Bottom-up** tab to switch to the **Bottom-up** window and analyze application performance by function. By default, the grid is sorted by the **CPU Time** metric in the descending order. Select the **Process/Function/Thread/Call Stack** grouping level from the **Grouping** menu. This granularity enables you to visualize threads where the hotspots functions were executed.

After initial optimization, the `nqueens_IP_setqueen` function is still a bottleneck. Click the arrow sign ▶ at the `nqueens_IP_setqueen` function. You see that this function's execution was parallelized among fifteen threads.



Select these threads in the grid, right-click and choose the **Filter In by Selection** context menu option. The **Timeline** pane below is updated to display data for the selected threads only.



1. **Timeline area**. When you hover over the graph element, the timeline tooltip displays the time passed since the application has been launched.

2. **Threads area** that shows the distribution of CPU time utilization per thread. Hover over a bar to see the CPU time utilization in percent for this thread at each moment of time. Dark green zones show the time threads are active. Light-green zones show the time threads were waiting.

3. **Performance metric area** that shows application performance over time by a performance metric. In the **Hotspots by CPU Usage** viewpoint, **CPU Usage** and **Thread Concurrency** metrics are used.

The CPU Usage chart shows the distribution of CPU time utilization for the whole application. Hover over a bar to see the application-level CPU time utilization in percent at the particular moment. VTune Amplifier calculates the overall **CPU Usage** metric as the sum of CPU time per each thread of the Threads area. Maximum **CPU Usage** value is equal to [number of processor cores] x 100%.

The **Thread Concurrency** chart shows the application-level thread concurrency at each moment of time. Hover over a bar to see an exact concurrency level at the particular moment.

Transitions. The execution flow between threads where one thread signals to another thread waiting to receive that signal. You may zoom in to a time region to get more detailed view of the transitions. To do this, drag and drop to select the region and right-click to select the **Zoom In on Selection** option from the context menu.

The **Timeline** pane for the sample application shows a large number of transitions between threads, which means that the threads spent noticeable time transferring execution to each other. If you uncheck the **Transitions** display option on the right, you see that workload balance is also poor since many of the threads were waiting for `OMP Worker Thread #7` to complete execution.

Run the Locks and Waits analysis to understand what prevents the sample code from effective thread concurrency and processor utilization.

Key Terms

- Thread concurrency
- Viewpoint

Next Step

Run Locks and Waits Analysis

Run Locks and Waits Analysis

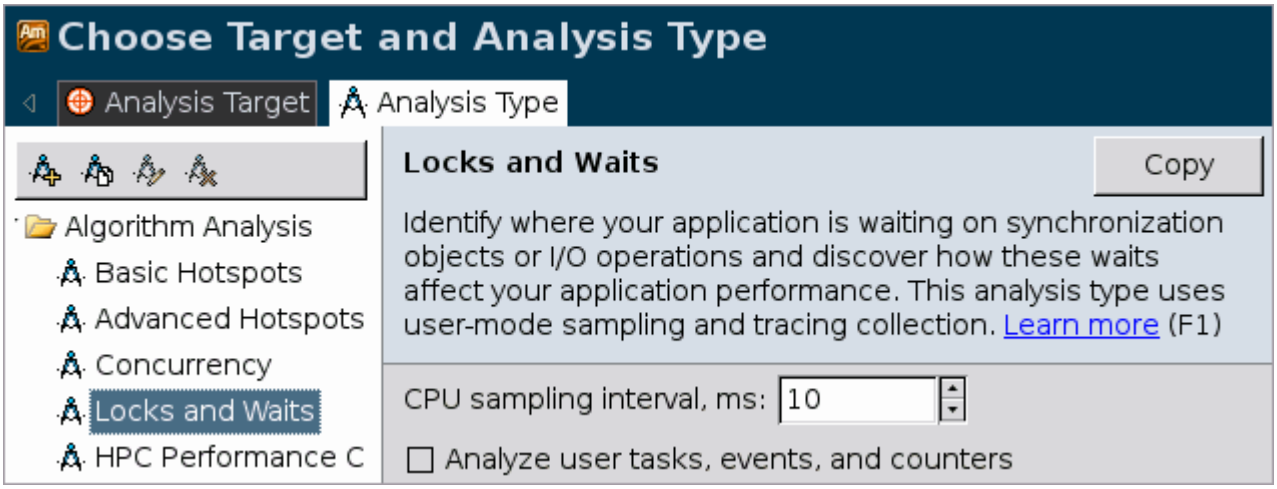
Run the Locks and Waits analysis to identify synchronization objects that caused contention and fix the problem in the source.

To run an analysis:

1. From the VTune Amplifier toolbar, click the **New Analysis** button.

   VTune Amplifier result tab opens with the **Choose Analysis Type** window active.
2. From the analysis tree on the left, select **Algorithm Analysis > Locks and Waits**.

   The right pane is updated with the default options for the Locks and Waits analysis.
3. Select the **Analyze Intel runtimes and user synchronization** checkbox.
4. Click the **Start** button on the right command bar.



VTune Amplifier launches the `nqueens_parallel` executable that makes calculations, displays the execution time, and exits. VTune Amplifier finalizes the collected data and opens the results in the **Locks and Waits** viewpoint.

NOTE

- To make sure the performance of the application is repeatable, go through the entire tuning process on the same system with a minimal amount of other software executing.
- This tutorial explains how to run an analysis from the VTune Amplifier graphical user interface (GUI). You can also use the VTune Amplifier command-line interface (`amplxe-cl` command) to run an analysis. For more details, check the *Command-line Interface Support* section of the VTune Amplifier Help.

Key Terms

- Finalization
- Viewpoint

Next Step

Interpret Locks and Waits Results

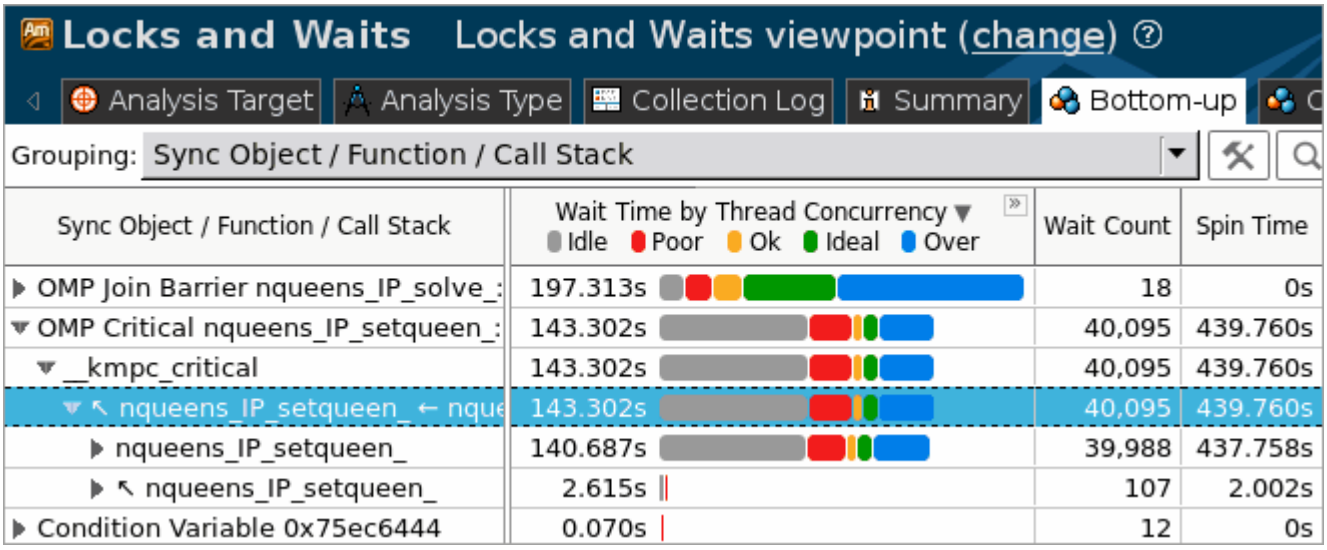
Interpret Locks and Waits Results

When the sample application exits, the Intel® VTune™ Amplifier finalizes the results and opens the Locks and Waits viewpoint that is configured to display synchronization objects sorted by Wait time. To interpret the data on the sample code performance, do the following:

1. Identify locks.
2. Analyze source code.

Identify Locks

Click the **Bottom-up** tab to open the **Bottom-up** pane.



The table below explains the type of data provided in the **Bottom-up** pane:

1. Synchronization objects that control threads in the application. The hash (unique number) appended to some names of the objects identify the stack creating this synchronization object.

2. The utilization of the processor time when a given thread waited for some event to occur. By default, the synchronization objects are sorted by **Poor** processor utilization type. Bars showing OK or Ideal utilization (orange and green) are utilizing the processors well. You should focus your optimization efforts on functions with the longest poor CPU utilization (red bars if the bar format is selected). Next, search for the longest over-utilized time (blue bars).

   This is the Data of Interest column for the Locks and Waits analysis results that is used for different types of calculations, for example: call stack contribution, percentage value on the filter toolbar.

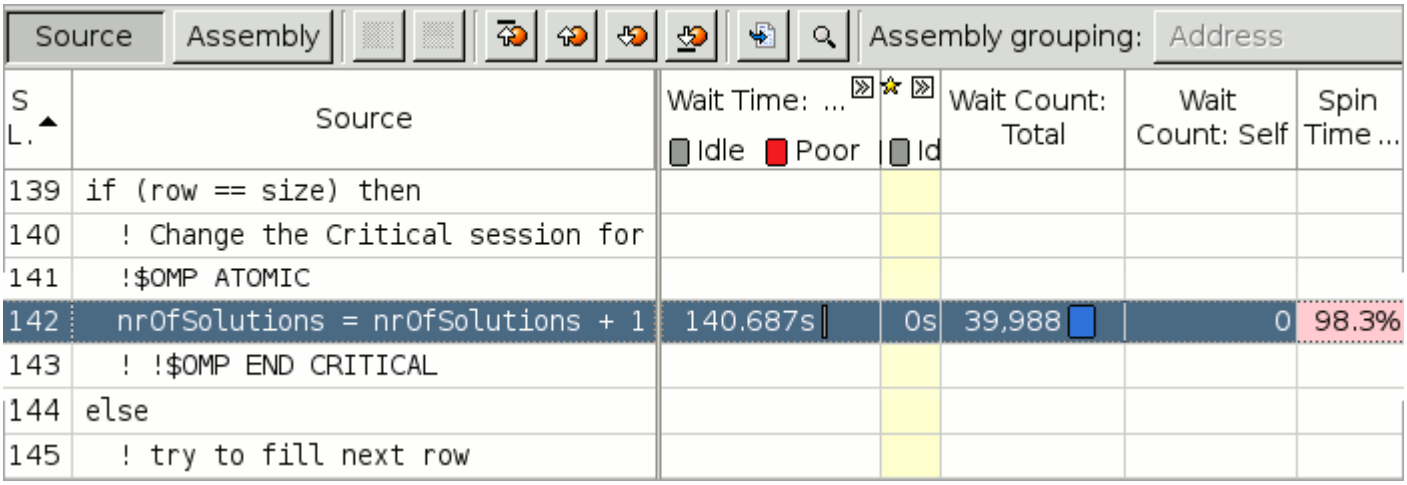
3. Number of times the corresponding system wait API was called. For a lock, it is the number of times the lock was contended and caused a wait. Usually you are recommended to focus your tuning efforts on the waits with both high Wait Time and Wait Count values, especially if they have poor utilization.

4. Wait time, during which the CPU is busy. This often occurs when a synchronization API causes the CPU to poll while the software thread is waiting. Some Spin time may be preferable to the alternative of the increased thread context switches. However, too much Spin time can reflect lost opportunity for productive work.

In the `nqueens_parallel` sample code, there are two critical wait objects, `OMP Critical nqueens_IP_setqueen` and `OMP Join Barrier`, that caused redundant synchronization and took the longest Wait time and highest Wait count. The bar indicators in the **Wait Time** column indicate that most of the time for these objects processor cores were underutilized.

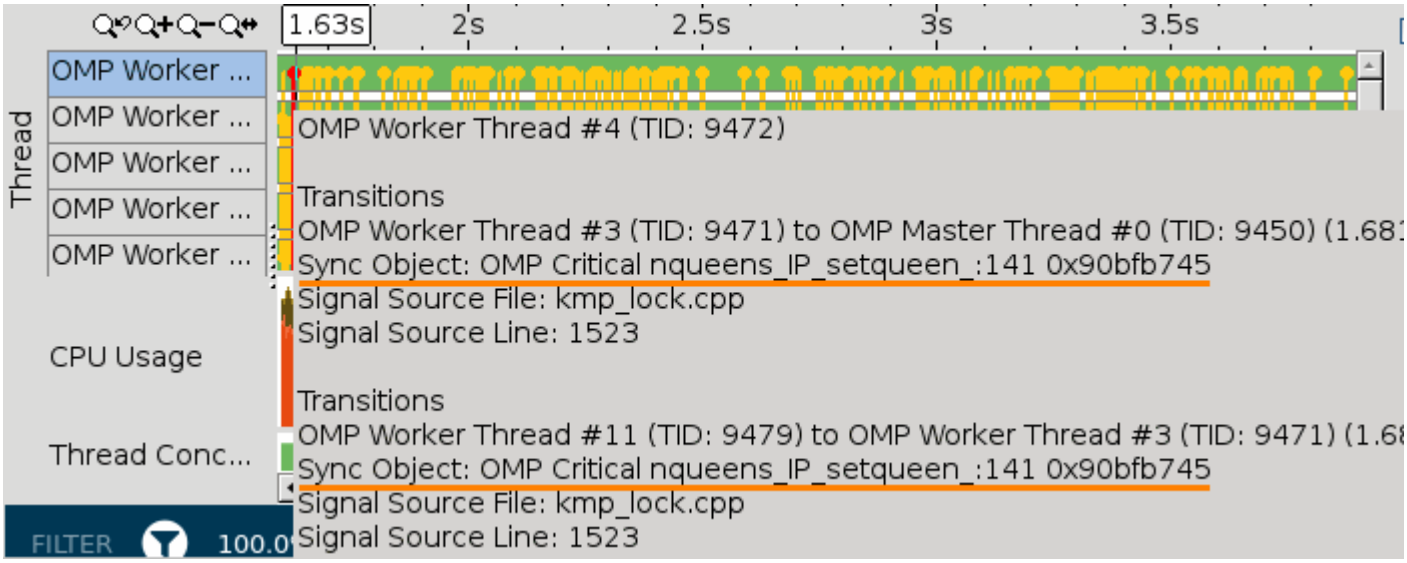
Analyze Source Code

Explore the source of the critical synchronization objects that caused significant Wait time and poor processor utilization. Double-click the `nqueens_IP_setqueen` object to analyze the source of the `setqueen` wait function. Click the button on the **Source** pane toolbar to go to the biggest hotspot code line in the function. VTune Amplifier highlights line 142 protected by the OpenMP* critical section.



The `setqueen` function was waiting for 140.687 seconds while this code line was executing. During this time, this operation was contended 39,988 times.

Hover over any transition line in the **Timeline** pane below to explore the infotip and make sure that all the transitions are caused by the `OMP Critical nqueens_IP_setqueen` critical section.



The `OMP Critical nqueens_IP_setqueen` section is the place where the application is serializing. Each thread has to wait for the critical section to be available before it can proceed. Only one thread can be in the critical section at a time.

You need to optimize the code to make it more concurrent. Click the **Source Editor** button on the **Source** window toolbar to open the code editor and optimize the code.

Key Terms

- Elapsed time
- Wait time

Next Step

Remove Lock

Remove Locks

In the Source window, you located the synchronization objects that caused significant waits while the processor cores were underutilized . To resolve the issue, do the following:

1. Open the code editor.
2. Modify the code to remove locks.
3. Recompile the project and check the result.

Open the Code Editor

NOTE
Depending on the sample code version, your source line numbers may slightly differ from the numbers provided in this tutorial.

Click the  **Source Editor** button to open the `nqueens_parallel.f90` file in your default editor:

```
139   if (row == size) then
140     ! Change the Critical session for the Atomic directive OMP ATOMIC
141     !$OMP CRITICAL
142     nrOfSolutions = nrOfSolutions + 1
143     !$OMP END CRITICAL
144   else
145     ! try to fill next row
146     do i=1,size
147       call setQueen (queens, row+1, i)
148     end do
149   end if
150 end subroutine SetQueen
```

Remove Locks

The critical section introduced in line 141 protects the global variable from a race condition in a multithreaded application but it spawns a redundant synchronization. To resolve this issue, you may replace the critical section with an atomic operation as follows:

1. Edit like 141 to replace the `OMP CRITICAL` with the `OMP ATOMIC` directive.
2. Comment out or remove line 143.

```
141     !$OMP ATOMIC
142     nrOfSolutions = nrOfSolutions + 1
143     !!$OMP END CRITICAL
```

3. Save your changes.

Recompile the Project and Check the Result

1. Browse to the directory where you extracted the sample code (for example, `/home/vtune/nqueens_fortran/linux`).
2. Rebuild your target in the release mode using the make command as follows:

```
$ make clean
$ make
```

The `nqueens_parallel` application is rebuilt.

3. Run `nqueens_parallel` as follows:

```
./nqueens_parallel 15
```

```
Starting nqueens solver for size 15 with 16 thread(s)
Number of solutions: 2279184
Correct Result!
Calculations took 14585ms.
```

System runs the `nqueens_parallel`. Note that execution time reduced from 56566 ms to 14585 ms.

Key Terms

- Wait time

Next Step

Compare with Previous Result

Compare with Previous Result

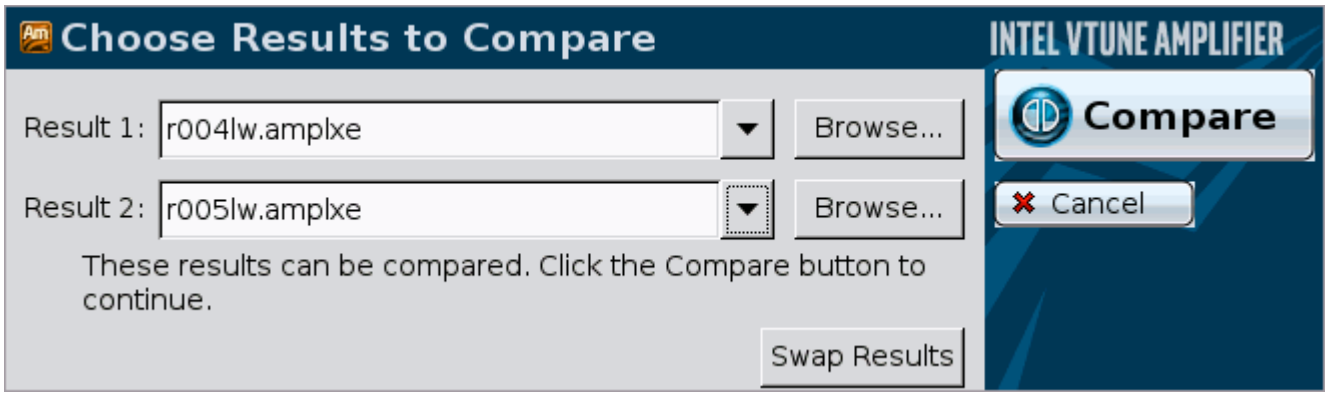
You made sure that removing the critical section gave you 41981 ms of optimization in the application execution time. To understand the impact of your changes and how the CPU utilization has changed, re-run the Locks and Waits analysis on the optimized code and compare results:

1. Compare results before and after optimization.
2. Identify the performance gain by metrics.
3. Compare timeline data.

Compare Results Before and After Optimization

1. Run the Locks and Waits analysis on the modified code.
2. Click the  **Compare Results** button on the VTune Amplifier toolbar.

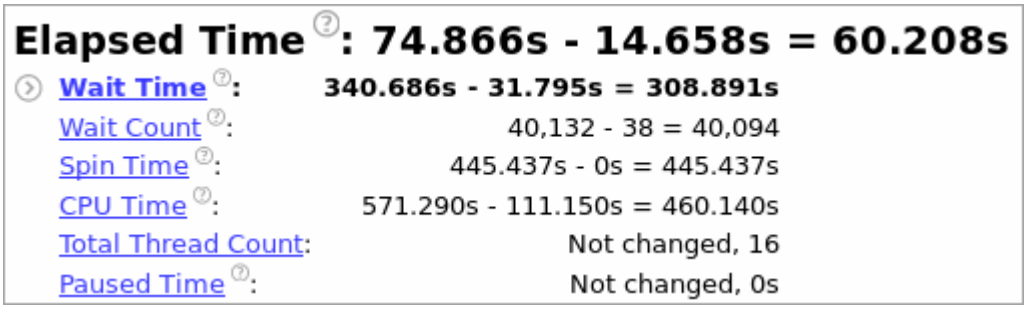
   The **Compare Results** window opens.
3. Specify the Locks and Waits analysis results you want to compare:



The **Summary** window opens providing the statistics for the difference between collected results.

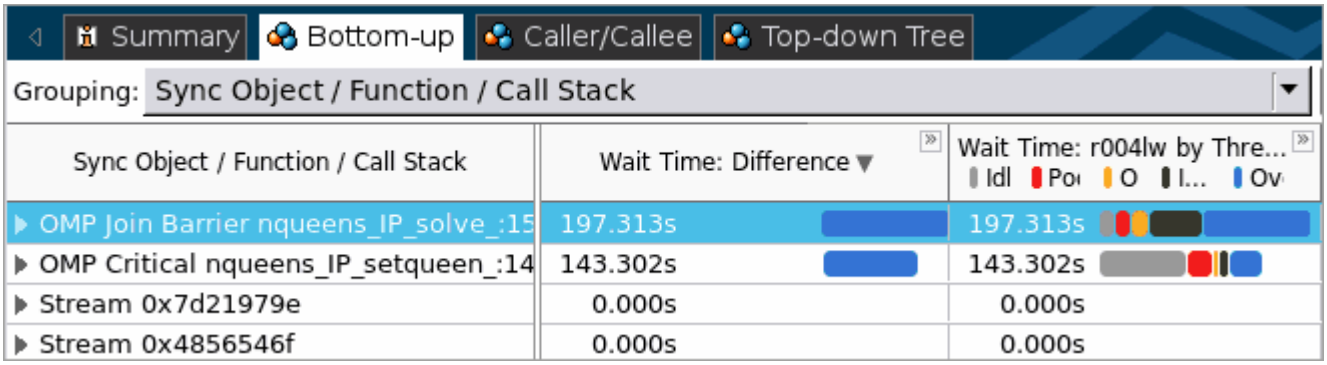
Identify the Performance Gain by Metrics

The **Result Summary** section of the **Summary** window shows that after optimization all critical metric values has reduced significantly. The **Elapsed Time** data shows the optimization of 60.208 seconds for the whole application. **Wait Time** decreased by 308.891 seconds, **Wait Count** - by 40,094.



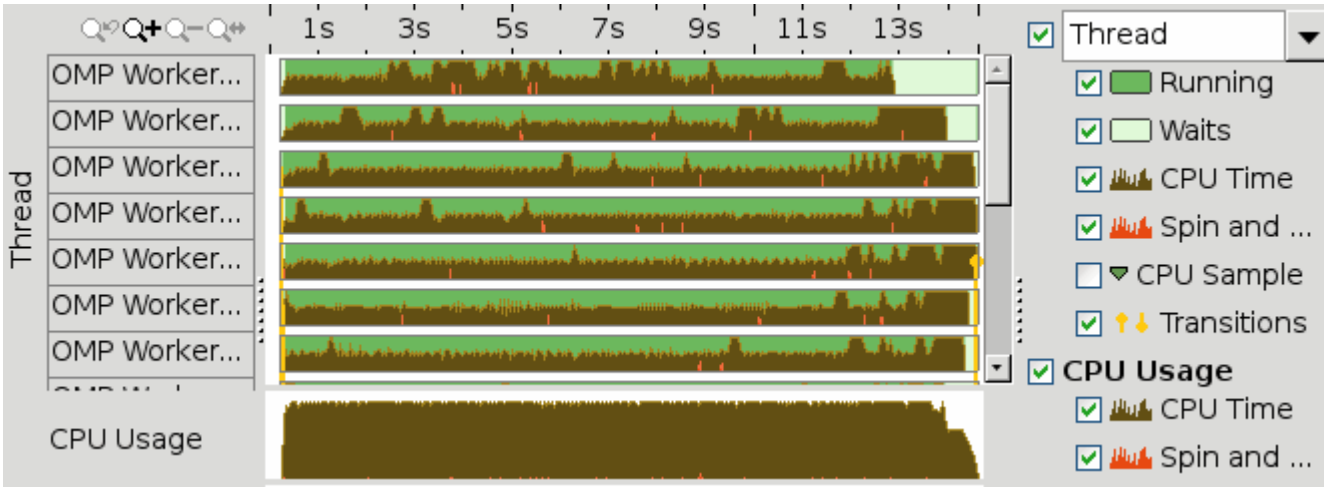
NOTE
The Locks and Waits analysis adds an overhead to the application execution. The overhead often depends on the number of threads and synchronization objects used in the application. This is the reason why Elapsed time data provided in the **Summary** window may differ from the data reported after the application launch outside of the VTune Amplifier.

In the **Bottom-up** pane, locate the OpenMP* critical section you identified as a bottleneck in your code. Since you removed it during optimization, the optimized result does not show any performance data for this synchronization object. If you collapse the **Wait Time: Difference** column by clicking the « button, you see that with the optimized result you got 143.302 seconds of optimization in Wait time.



Compare Timeline Data

Open the optimized result of the Locks and Waits analysis, open the **Bottom-up** tab, and analyze the **Timeline** pane.



The optimized result does not have multiple transitions anymore.

Compare analysis results regularly to look for regressions and to track how incremental changes to the code affect its performance. You may also want to use the VTune Amplifier command-line interface and run the `amplxe-cl` command to test your code for regressions. For more details, see the *Command-line Interface Support* section in the VTune Amplifier online help.

Key Terms

- Elapsed time
- Thread concurrency
- Wait time

3

Summary

You have completed the Finding Hotspots tutorial. Here are some important things to remember when using the Intel® VTune™ Amplifier to analyze your code for hotspots:

Step	Tutorial Recap	Key Tutorial Take-Aways
1. Prepare for analysis	You set up your environment to enable generating symbol information for your binary files, built the target, created the performance baseline, and created the VTune Amplifier project for your analysis target.	• Configure your project properties to get the most accurate results for user binaries and to analyze the performance of your application at the code line level. • Create a performance baseline to compare the application versions before and after optimization. Make sure to use the same workload for each application run. • Use the **New Amplifier Result** tab to choose and configure your analysis target. • Use the **Analysis Type** configuration window to choose, configure, and run the analysis. You can also run the analysis from command line using the `amplxe-cl` command.
2. Find hotspots	You launched the Basic Hotspots data collection that analyzed function calls and CPU time spent in each program unit of your application and identified the following hotspots: • A function that took the most CPU time and could be a good candidate for algorithm tuning. • The code section that took the most CPU time to execute.	• Start analyzing the performance of your application from the **Summary** window to explore the performance metrics for the whole application. Then, move to the **Bottom-up** window to analyze the performance per function. Focus on the *hotspots* - functions that took the most CPU time. By default, they are located at the top of the table. • Double-click the hotspot function in the **Bottom-up** pane or **Call Stack** pane to open its source code and navigate between hotspots using the **Source** window navigation buttons.
3. Eliminate hotspots	You optimized the algorithm by enabling the OpenMP* library create a private copy of the array. You rebuilt the application and got performance gain of 200144 ms.	Click the  **Source Editor** button to open your default source editor directly from the VTune Amplifier **Source** window.
4. Analyze concurrency	You launched the Concurrency analysis and identified poor thread concurrency for the whole application execution. You analyzed the timeline and identified poor thread balance: all OpenMP threads were constantly transferring	• Start your analysis with the **Summary** window. Consider the **Target** concurrency metric specified in the **CPU Usage Histogram** as your optimization goal. The **Average** metric is calculated as CPU time / Elapsed time. Use this number as another baseline for your performance measurements. The closer this number to the number of cores, the better.

Step	Tutorial Recap	Key Tutorial Take-Aways
	execution to each other and were waiting for all threads to complete execution.	• In the **Bottom-up** window, use the **Filter In by Selection** context menu option to focus on the performance-critical functions in the grid and analyze their performance over time in the **Timeline** pane.
5. Find lock	You ran the Locks and Waits analysis and identified two synchronization objects with the high Wait Time and Wait Count values and poor CPU utilization that could be locks affecting application parallelism. Your next step is to analyze the code of their wait functions.	• Use the **Analysis Type** configuration window to choose, configure, and run the analysis. For recently used analysis types, you may use the shortcuts to run a recent analysis: • From the **File** menu, select **New > [recent_analysis_type]**. • In the **Bottom-up** window, focus on the synchronization objects that under- or over-utilized the available logical CPUs and have the highest Wait time and Wait Count values. By default, the objects with the highest Wait time values show up at the top of the window.
6. Remove lock	You optimized the application execution time by removing the unnecessary critical section that caused redundant synchronization .	Double-click the most time-critical synchronization object in the **Bottom-up** pane. This opens the source code for the wait function it belongs to. Use the hotspot navigation buttons to identify the most time-critical code lines.
7. Check your work	You ran the Locks and Waits analysis on the optimized code and compared the results before and after optimization using the Compare mode of VTune Amplifier.	Perform regular regression testing by comparing analysis results before and after optimization. Click the  **Compare Results** button on the VTune Amplifier toolbar. From command line, use the `amplxe-cl` command.

Next step: Prepare your own application(s) for analysis. Then use the VTune Amplifier to find and eliminate hotspots.

4 Key Terms

baseline: A performance metric used as a basis for comparison of the application versions before and after optimization. Baseline should be measurable and reproducible.

CPU time: The amount of time a thread spends executing on a logical processor. For multiple threads, the CPU time of the threads is summed. The application CPU time is the sum of the CPU time of all the threads that run the application.

Elapsed time: The total time your target ran, calculated as follows: **Wall clock time at end of application - Wall clock time at start of application**.

finalization: A process during which the VTune Amplifier converts the collected data to a database, resolves symbol information, and pre-computes data to make further analysis more efficient and responsive.

hotspot: A section of code that took a long time to execute. Some hotspots may indicate bottlenecks and can be removed, while other hotspots inevitably take a long time to execute due to their nature.

target: A target is an executable file you analyze using the VTune Amplifier.

thread concurrency: A performance metric that helps identify how an application utilizes the processors in the system by comparing the application concurrency level (the number of active threads) and target concurrency level (by default, equal to the number of physical cores). Thread concurrency may be higher than CPU usage if threads are in the runnable state and not consuming CPU time.

Utilization Type	Default color	Description
Idle		All threads in the program are waiting - no threads are running. There can be only one node in the **Summary** chart indicating idle utilization.
Poor		Poor utilization. By default, poor utilization is when the number of threads is up to 50% of the target concurrency.
OK		Acceptable (OK) utilization. By default, OK utilization is when the number of threads is between 51-85% of the target concurrency.
Ideal		Ideal utilization. By default, ideal utilization is when the number of threads is between 86-115% of the target concurrency.
Over		Over-utilization. By default, over-utilization is when the number of threads is more than 115% of the target concurrency.

viewpoint: A preset result tab configuration that filters out the data collected during a performance analysis and enables you to focus on specific performance problems. When you select a viewpoint, you select a set of performance metrics the VTune Amplifier shows in the windows/panes of the result tab. To select the required viewpoint, click the **(change)** link and use the drop-down menu at the top of the result tab.

Wait time: The amount of time that a given thread waited for some event to occur, such as: synchronization waits and I/O waits.