Intel® VTune™ Profiler: STATE OF THE UNION

Dr. Sri Doddapaneni

Senior Director, Developer Software
Intel Corporation

Storage Performance Development Kit (SPDK)
Persistent Memory Development Kit (PMDK)
Intel® VTune™ Profiler

Virtual Forum

AGENDA

01 **The Big Picture**
oneAPI and Intel® Parallel Studio XE

02 **oneAPI Overview**
Short description of topic

03 **Performance**
Deep Dive

04 **Conclusion and Q&A**
Thank you!

THE BIG PICTURE

- What is VTune Profiler?
- Really? You can do that?
- I have heard about oneAPI, how does that relate to VTune Profiler?
- I have heard about Parallel Studio, how does that relate to VTune Profiler?
- What are examples of uses of VTune Profiler that could transform the way we work?

PROGRAMMING CHALLENGES FOR MULTIPLE ARCHITECTURES

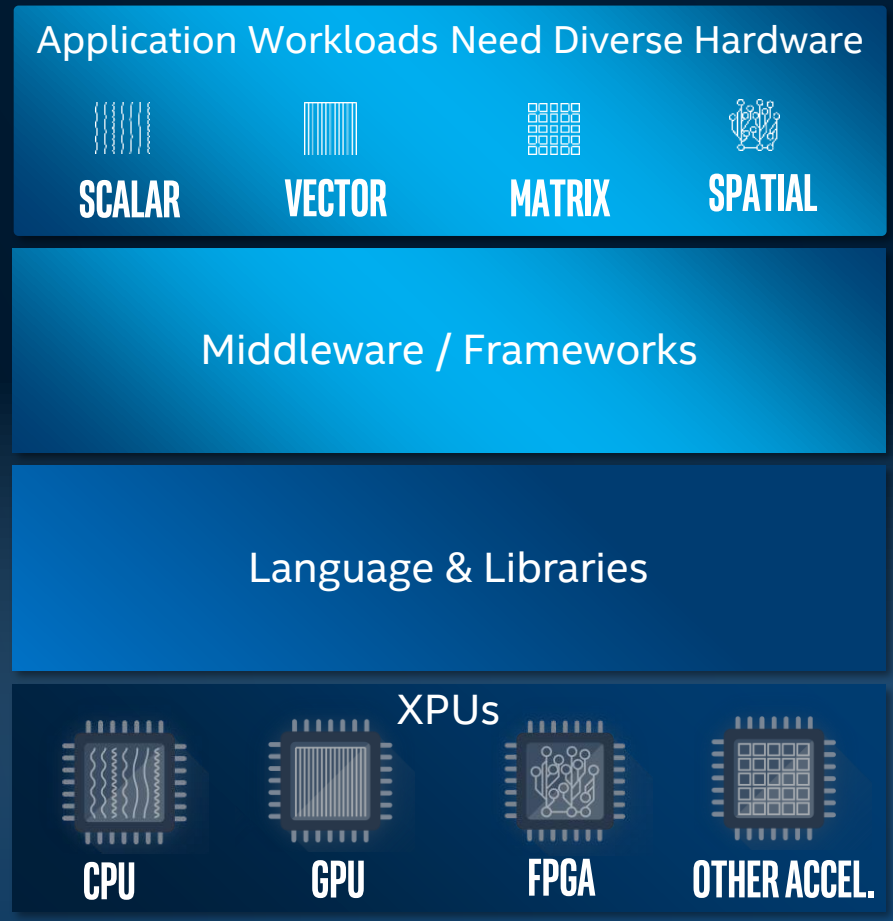
Growth in specialized workloads

Diverse set of data-centric hardware required

No common programming language or APIs

Inconsistent tool support across platforms

Each platform requires unique software investment



INTRODUCING ONEAPI

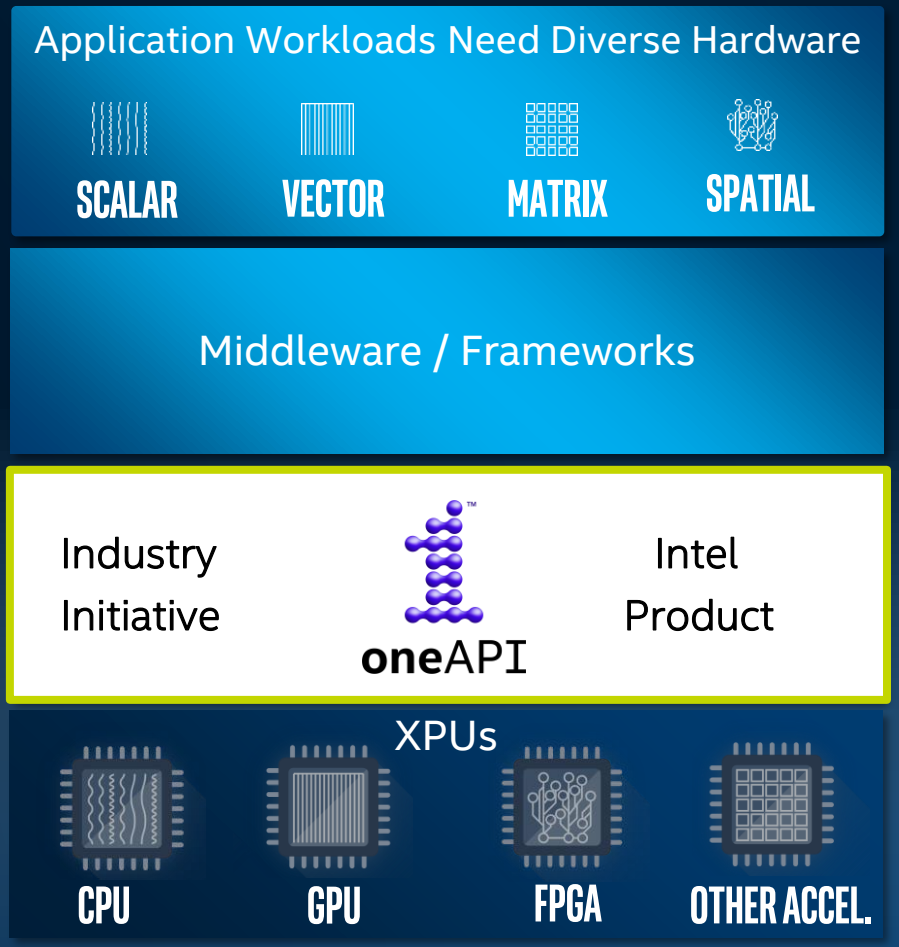
Unified programming model to simplify development across diverse architectures

Unified and simplified language and libraries for expressing parallelism

Uncompromised native high-level language performance

Based on industry standards and open specifications

Interoperable with existing HPC programming models



YES, YOU STILL GET INTEL® VTUNE™ PROFILER

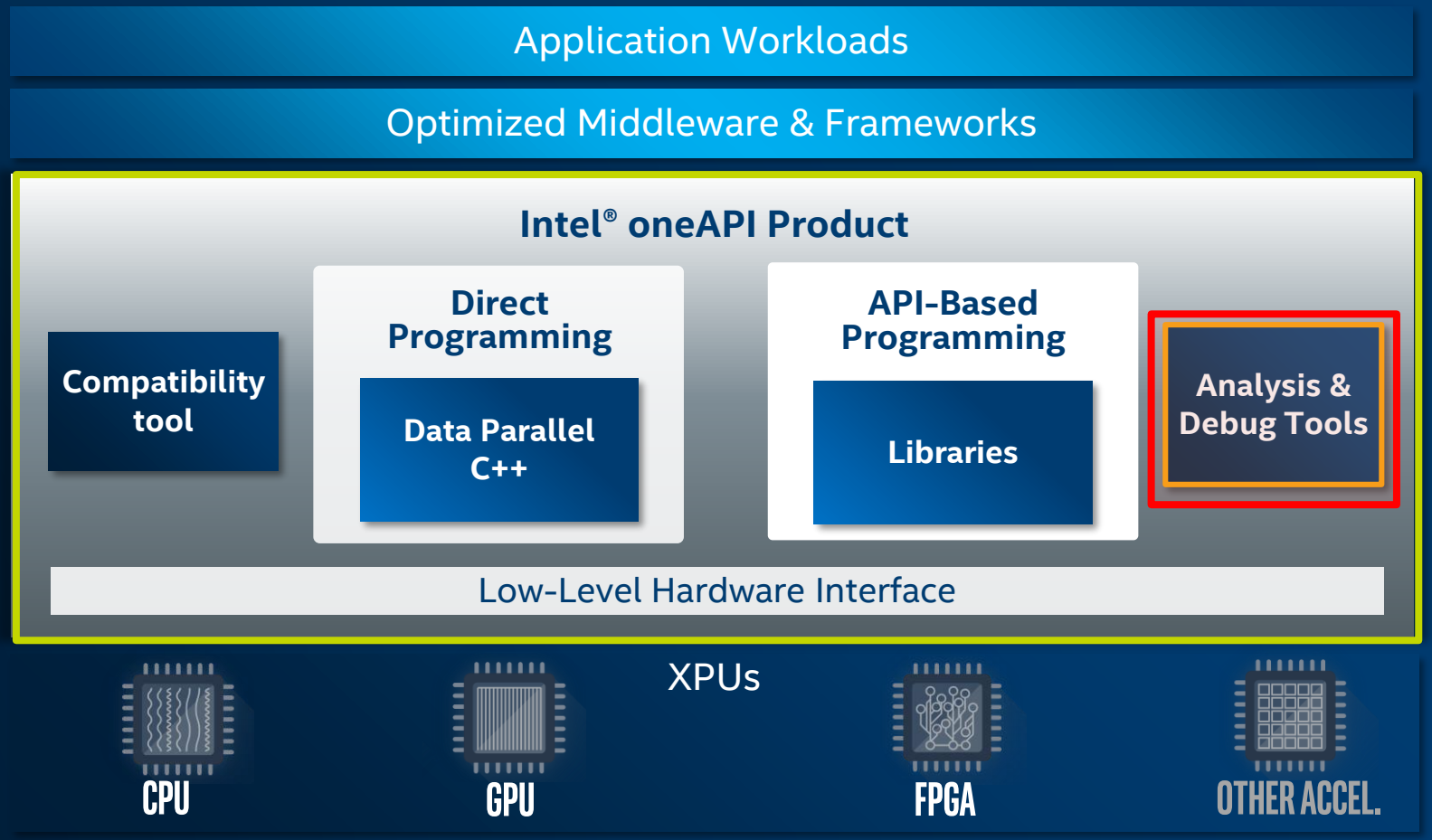
ONEAPI

Analysis Tools:

- Intel® VTune™ Profiler
- Intel® Advisor
- Intel® Inspector

Your favorite tools extended to accelerators

Beta Available Now



Visit software.intel.com/oneapi for more details

Some capabilities may differ per architecture and custom-tuning will still be required. Other accelerators to be supported in the future.

Refer to software.intel.com/articles/optimization-notice for more information regarding performance & optimization choices in Intel software products.

What's Inside Intel® Parallel Studio XE

Comprehensive Software Development Tool Suite

COMPOSER EDITION

BUILD
Compilers & Libraries

C / C++, Fortran Compilers

Intel® Math Kernel Library

Intel® Data Analytics Acceleration Library

Intel Threading Building Blocks
C++ Threading

Intel® Integrated Performance Primitives
Image, Signal & Data Processing

Intel® Distribution for Python
High Performance Python

PROFESSIONAL EDITION

ANALYZE
Analysis Tools

Intel® VTune™ Profiler
Performance Profiler

Intel® Inspector
Memory & Thread Debugger

Intel® Advisor
Vectorization Optimization
Thread Prototyping
& Flow Graph Analysis

CLUSTER EDITION

SCALE
Cluster Tools

Intel® MPI Library
Message Passing Interface Library

Intel® Trace Analyzer & Collector
MPI Tuning & Analysis

Intel® Cluster Checker
Cluster Diagnostic Expert System

Operating System: Windows, Linux, MacOS[1]

Intel® Architecture Platforms

[1]Available only in the Composer Edition.

More Resources

Training materials, cookbooks, case studies and more..

INTEL® VTUNE™ PROFILER

PERFORMANCE PROFILER

Optimize:
- application performance,
- system performance,
- system configuration
- and more...

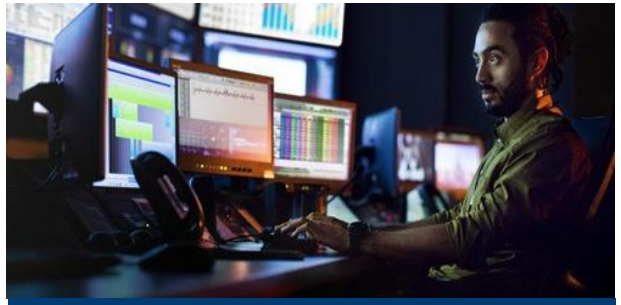
software.intel.com/vtune

INTEL® ADVISOR

DESIGN AND ANALYZE CODE FOR MODERN HARDWARE

- Roofline Analysis
- Vectorization Optimization
- Thread Prototyping
- Flow Graph Analysis
- Offload Advisor

software.intel.com/advisor



INTEL® INSPECTOR

THREAD AND MEMORY CHECKER

- Threading Checker
- Memory Checker
- Persistent Memory Checker

Software.intel.com/inspector

PERFORMANCE

Deep Dive

“If all you have is a Hammer, everything looks like a Nail …”


“The best kept secret”

KEY TAKEAWAY

Intel offers developers an extensive Tool Box … world class tools that will help developers achieve the best performance

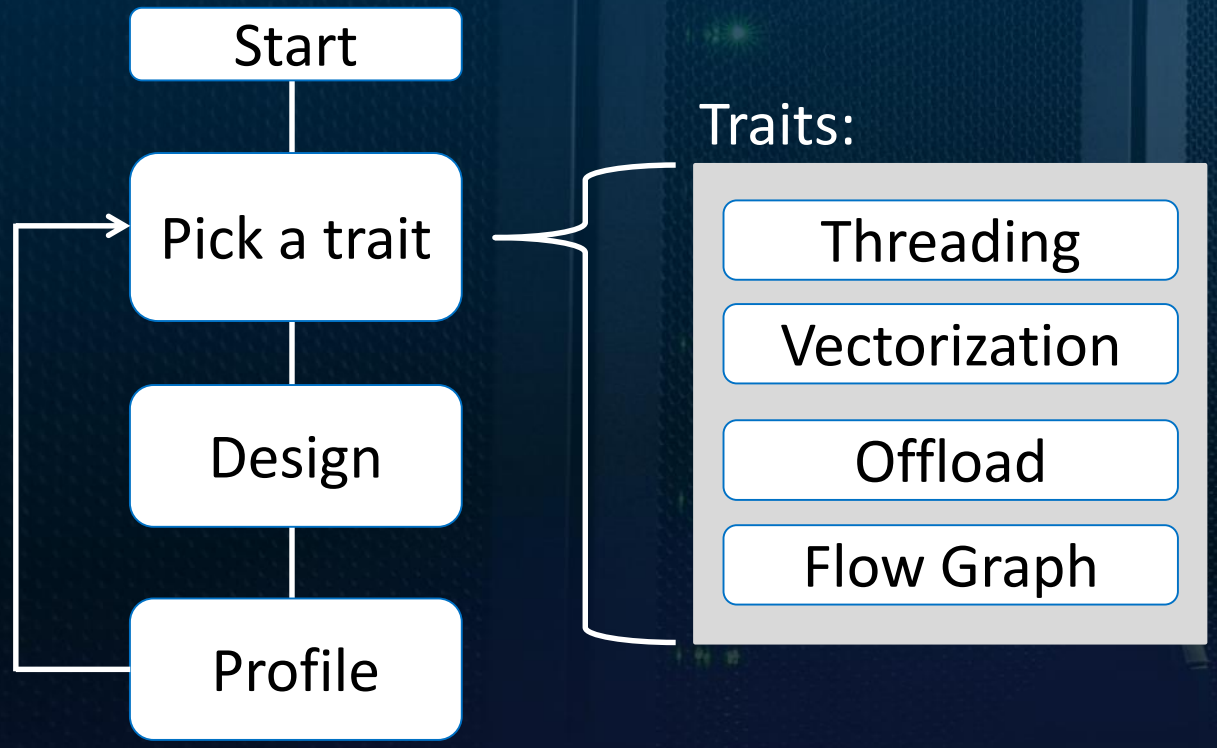
PERFORMANCE IS MANY THINGS

- Efficiency/speed of executing instruction stream in CPU cores
- Doing more per instruction (aka vector ops, loads, stores)
- Minimize wasted cycles due to waiting for cache and memory access
- Minimize wasted cycles due to IO for data from storage and network
- Utilize all the CPU cores with thread and process parallelism

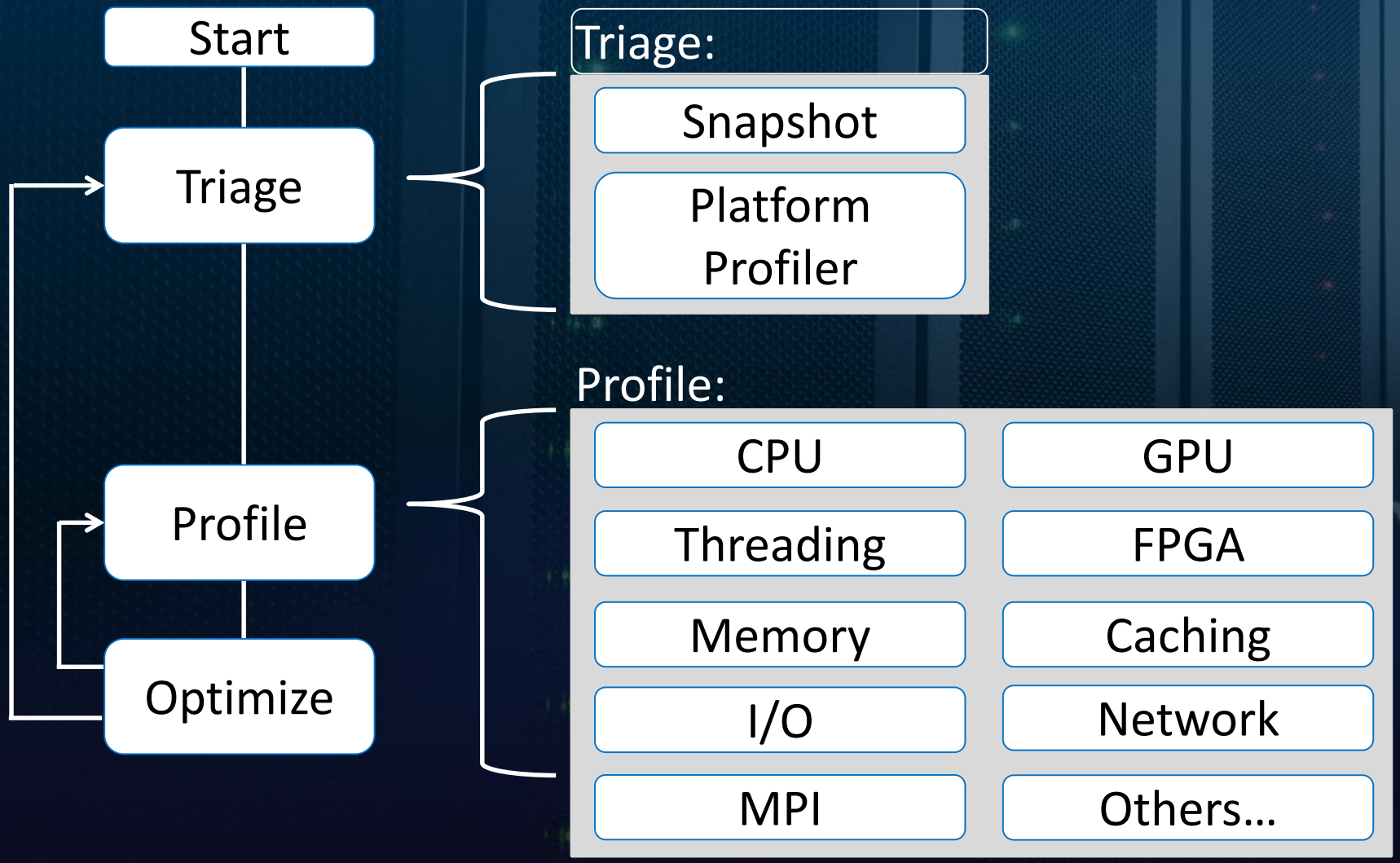
DEVELOPER ROLES

- Programmer: develops algorithms and implements in high level programming languages such as C, C++, Fortran, Open CL, Data Parallel C++
- Software Architect: architects large software applications and frameworks; specifies external behavior and functionality, and internal modularity and interfaces
- Performance Engineer: deep understanding of application performance characteristics, domain expertise in software and HW bottlenecks and dependence on HW for delivering performance goals

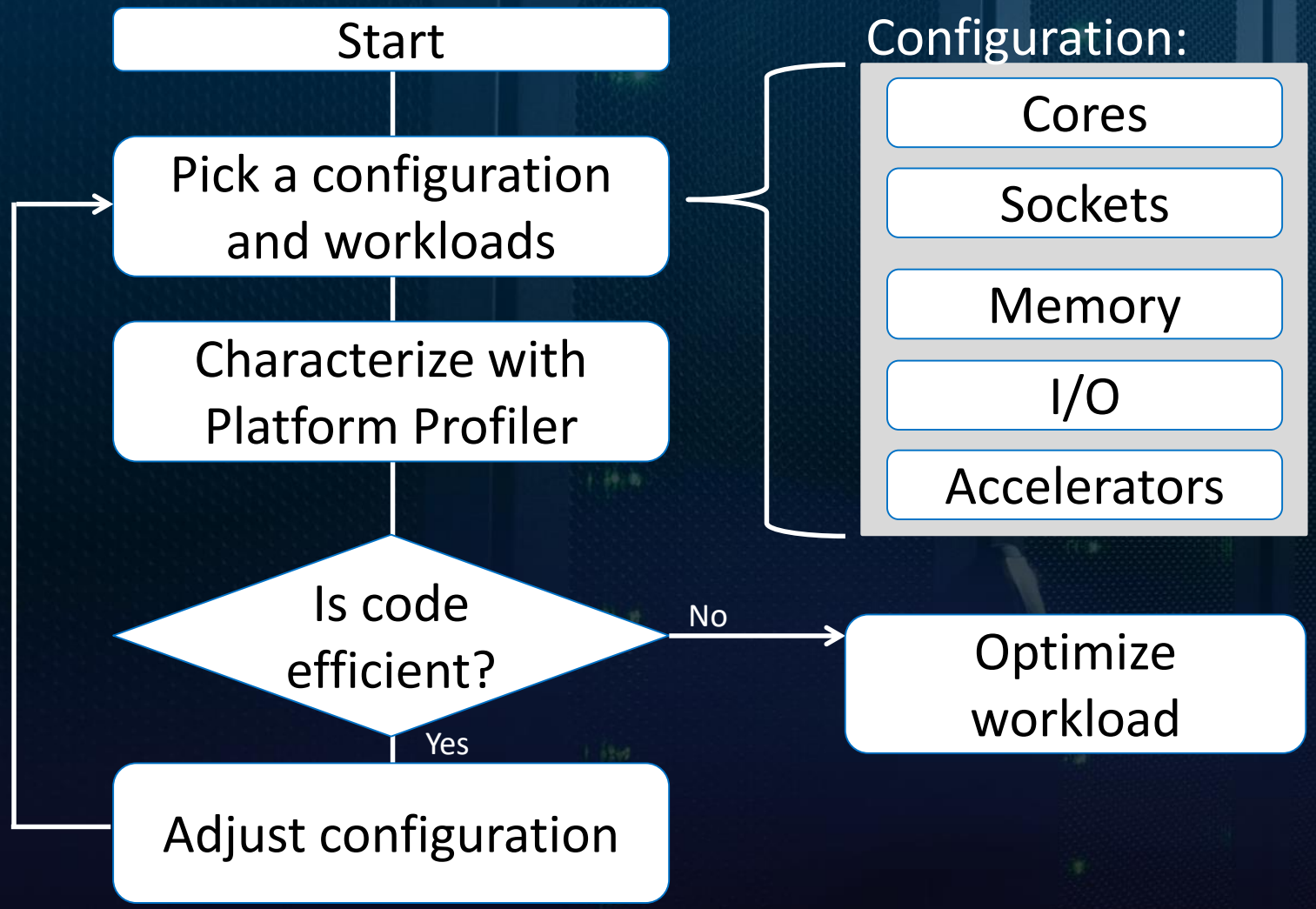
DESIGN WORKFLOW



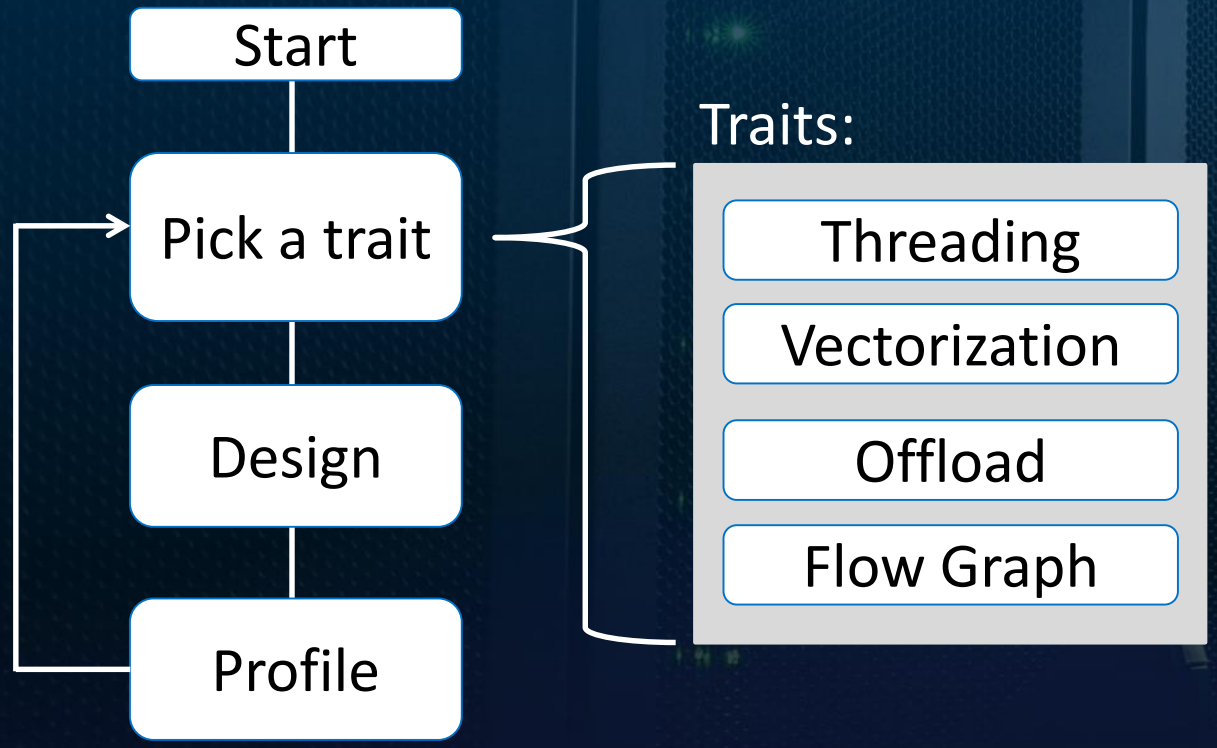
PROFILING WORKFLOW



CONFIGURATION WORKFLOW



DESIGN WORKFLOW



DESIGN IT, TUNE, DEBUG, THEN IMPLEMENT

INTEL® ADVISOR THREAD PROTOTYPING—DESIGN WITH DISRUPTING DEVELOPMENT

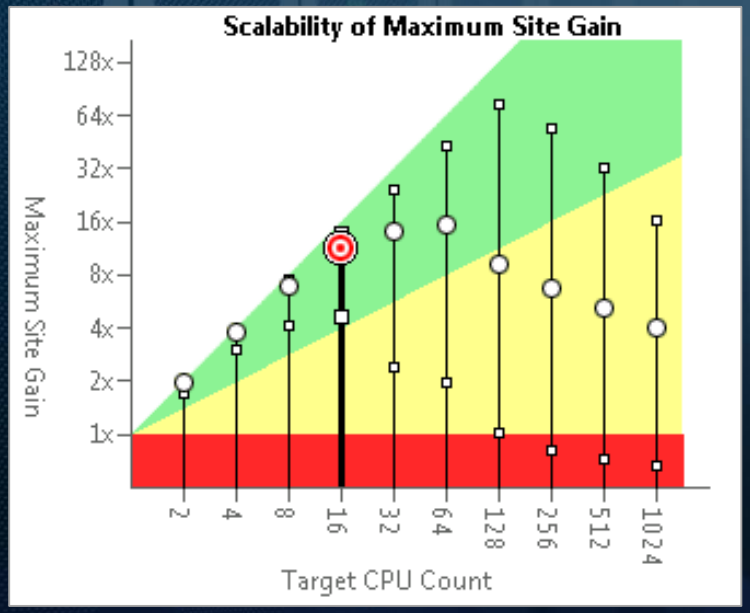
Have You

- Threaded an app, but seen little benefit?
- Hit a “scalability barrier?”
- Delayed release due to synchronization errors?

Data Driven Threading Design

- Quickly prototype multiple options
- Project scaling on larger systems
- Find synchronization errors before implementing threading
- Design without disrupting development

Add Parallelism with Less Effort, Less Risk & More Impact



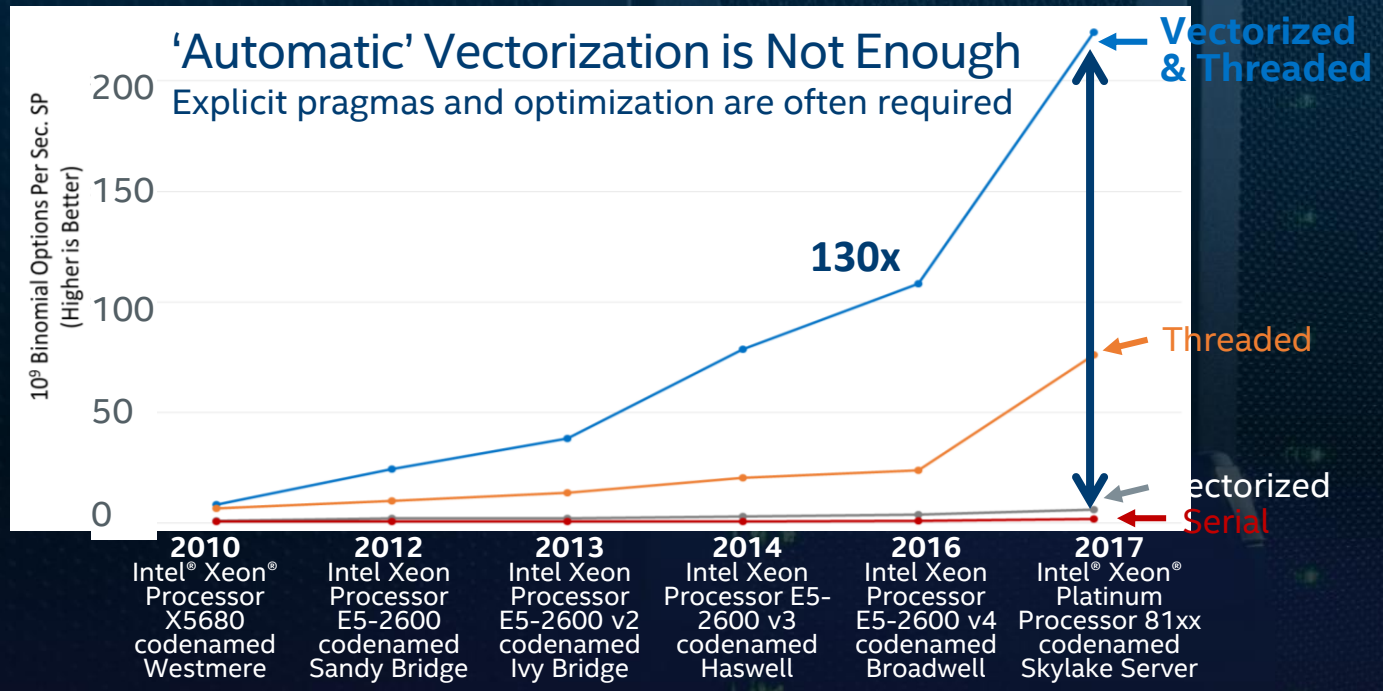
*“**Intel® Advisor** allowed us to quickly prototype ideas for parallelism, saving developer time & effort”*

Simon Hammond
Senior Technical Staff
Sandia National Laboratories

MODERNIZE YOUR CODE WITH INTEL® ADVISOR

OPTIMIZE VECTORIZATION, PROTOTYPE THREADING, CREATE AND ANALYZE FLOW GRAPHS

The Difference is Growing with Each New Hardware Generation



Modern Performant Code

- Vectorized (uses Intel® AVX-512/AVX2)
- Efficient memory access
- Threaded

Intel® Advisor

- Adds & optimizes vectorization
- Analyzes memory patterns
- Quickly prototypes threading

New

- Enhanced hierarchical roofline analysis
- Shareable HTML roofline
- Flow graph analysis
- Offload advisor

Benchmark: Binomial Options Pricing Model https://software.intel.com/en-us/articles/binomial-options-pricing-model-code-for-intel-xeon-phi-coprocessor

Benchmark results were obtained prior to implementation of recent software patches and firmware updates intended to address exploits referred to as "Spectre" and "Meltdown". Implementation of these updates may make these results inapplicable to your device or system.
For more complete information about performance and benchmark results, visit www.intel.com/benchmarks
See Vectorize & Thread or Performance Dies Configurations for 2010-2017 Benchmarks in Backup. Benchmark source: Intel Corporation.

Learn More: http: intel.ly/advisor-xe

FASTER CODE FASTER USING INTEL® ADVISOR

Vectorization

"**Intel® Advisor's Vectorization Advisor** permitted me to focus my work where it really mattered. When you have only a limited amount of time to spend on optimization, it is invaluable."

Gilles Civario
Senior Software Architect
Irish Centre for High-End Computing

"**Intel® Advisor's Vectorization Advisor** fills a gap in code performance analysis. It can guide the informed user to better exploit the vector capabilities of modern processors and coprocessors."

Dr. Luigi Iapichino
Scientific Computing Expert
Leibniz Supercomputing Centre

Threading

"**Intel® Advisor** has been extremely helpful in identifying the best pieces of code for parallelization. We can save several days of manual work by targeting the right loops and we can use Advisor to find potential thread safety issues to help avoid problems later on."

Carlos Boneti
HPC software engineer,
Schlumberger

"**Intel® Advisor** has allowed us to quickly prototype ideas for parallelism, saving developer time and effort, and has already been used to highlight subtle parallel correctness issues in complex multi-file, multi-function algorithms."

Simon Hammond
Senior Technical Staff
Sandia National Laboratories

More Reviews

"AUTOMATIC" VECTORIZATION OFTEN NOT ENOUGH

A GOOD COMPILER CAN STILL BENEFIT GREATLY FROM VECTORIZATION OPTIMIZATION

Compiler will not always vectorize

- Check for Loop Carried Dependencies using Intel® Advisor
- All clear? Force vectorization.
  C++ use: pragma simd, Fortran use: SIMD directive

Not all vectorization is efficient vectorization

- Stride of 1 is more cache efficient than stride of 2 and greater. Analyze with Intel® Advisor.
- Consider data layout changes
  Intel® SIMD Data Layout Templates can help

Benchmarks on prior slides did not all "auto vectorize." Compiler directives were used to force vectorization and get more performance.

Arrays of structures are great for intuitively organizing data, but are much less efficient than structures of arrays. Use the Intel® SIMD Data Layout Templates (Intel® SDLT) to map data into a more efficient layout for vectorization.

GET BREAKTHROUGH VECTORIZATION PERFORMANCE

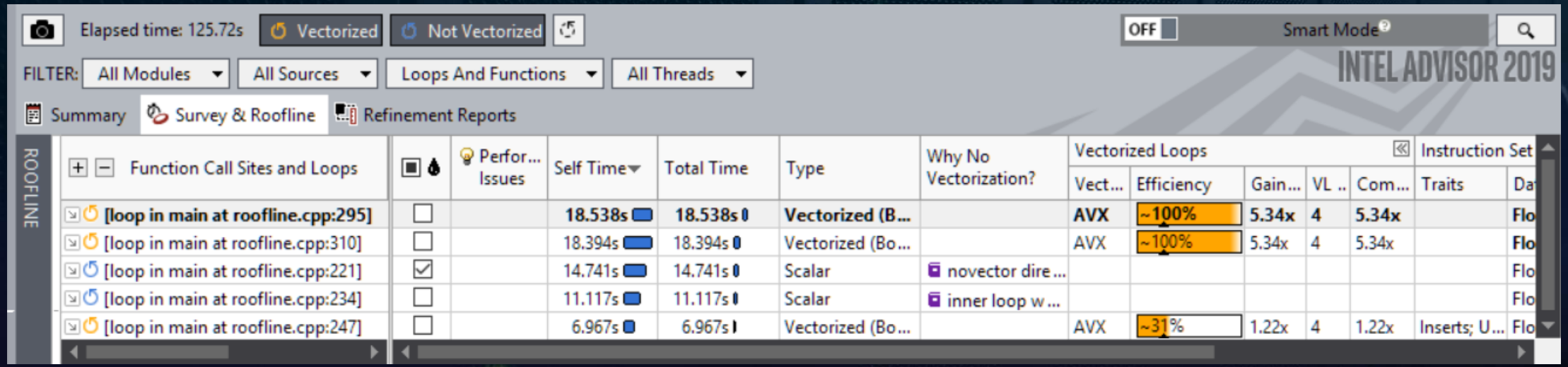
INTEL® ADVISOR—VECTORIZATION ADVISOR

Faster Vectorization Optimization

- Vectorize where it will pay off most
- Quickly ID what is blocking vectorization
- Tips for effective vectorization
- Safely force compiler vectorization
- Optimize memory stride

Data & Guidance You Need

- Compiler diagnostics + Performance Data + SIMD efficiency
- Detect problems & recommend fixes
- Loop-Carried Dependency Analysis
- Memory Access Patterns Analysis



Optimize for Intel® AVX-512 with or without access to AVX-512 hardware

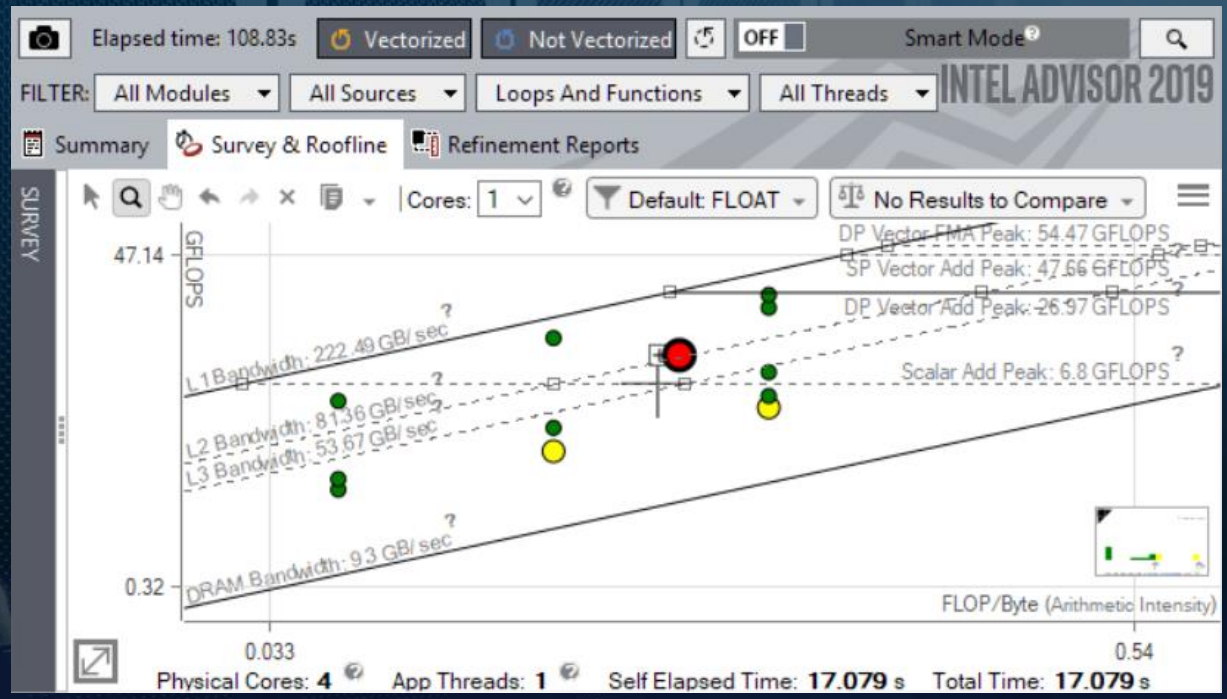
http://intel.ly/advisor-xe

FIND EFFECTIVE OPTIMIZATION STRATEGIES

INTEL® ADVISOR—CACHE-AWARE ROOFLINE ANALYSIS

Roofline Performance Insights

- Highlights poor performing loops
- Shows performance 'headroom' for each loop
  - Which can be improved
  - Which are worth improving
- Shows likely causes of bottlenecks
- Suggests next optimization steps



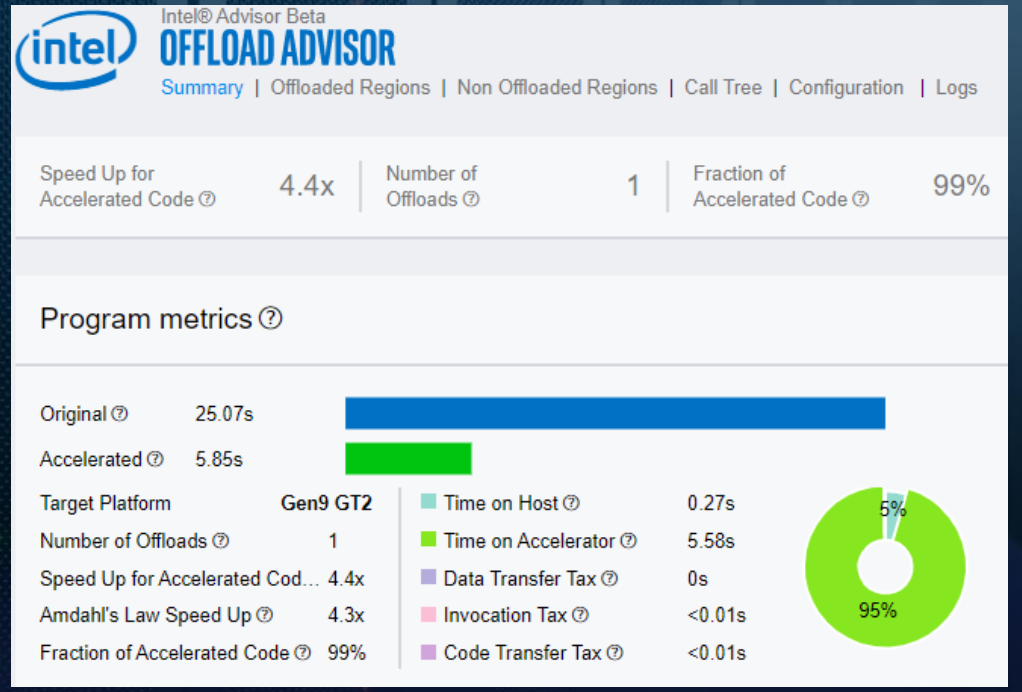
"I am enthusiastic about the new "integrated roofline" in Intel® Advisor. It is now possible to proceed with a step-by-step approach with the difficult question of memory transfers optimization & vectorization which is of major importance."

Nicolas Alferez, Software Architect
Onera – The French Aerospace Lab

EFFICIENTLY OFFLOAD TO GPU

INTEL® ADVISOR(BETA)— OFFLOAD ADVISOR

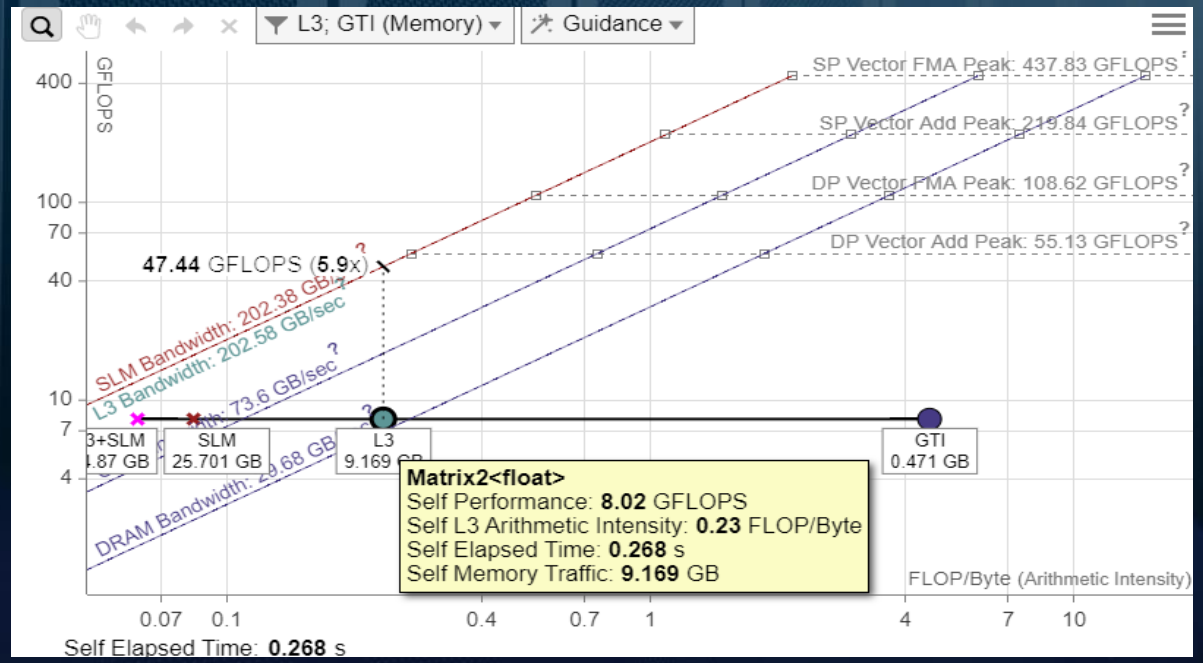
- Identify high-impact opportunities to offload
  - And areas that are not advantageous to offload
- Model performance, headroom and bottlenecks to get your code ready even before you have the GPU
- Pinpoint performance bottlenecks and key bounding factors
  - e.g. memory, cache, compute and data transfer
- No special recompile required



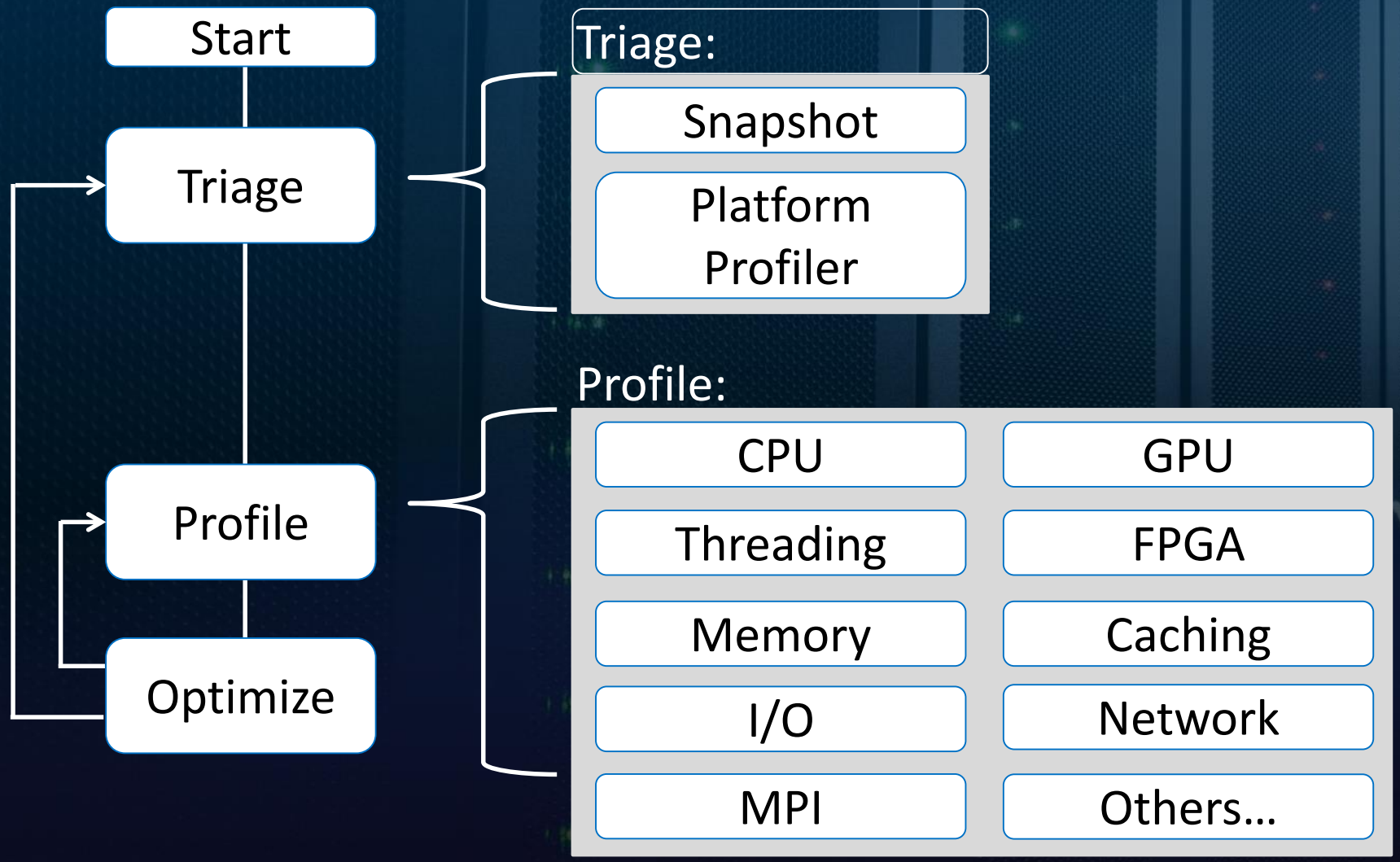
FIND EFFECTIVE OPTIMIZATION STRATEGIES

INTEL® ADVISOR (BETA)- GPU ROOFLINE

- See performance headroom against hardware limitations
- Determine optimization strategy
  - Identify which optimizations will payoff the most
  - Detect bottlenecks and their likely causes (memory bound vs. compute bound)
- Visualize optimization progress



PROFILING WORKFLOW



APPLICATION PERFORMANCE SNAPSHOT

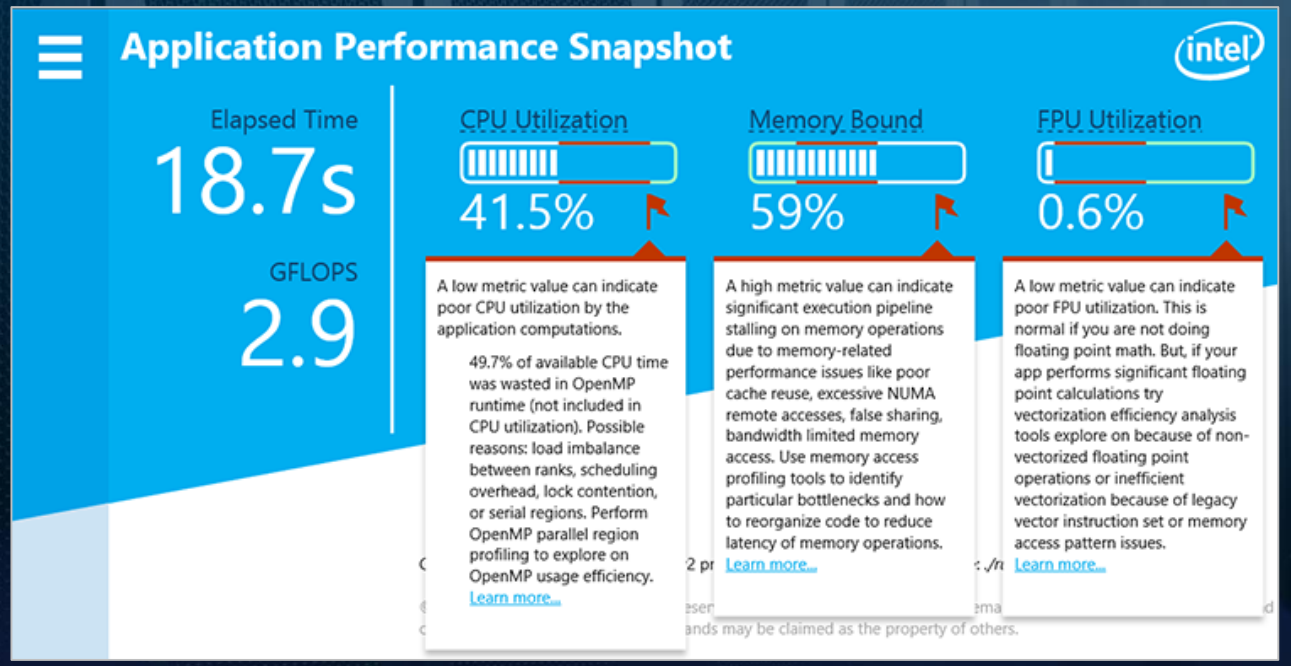
DISCOVER OPPORTUNITIES FOR BETTER PERFORMANCE WITH VECTORIZATION & THREADING

Objectives

- Simple enough to run during a coffee break
- Highlight where code modernization can help

Users

- Performance teams – fast prioritization of which apps will benefit most
- All Developers – size the potential performance gain from code modernization



Non-Objectives

- Actionable tuning data – that is another tool. Snapshot is just a fast “health” check.

Free download: http://www.intel.com/performance-snapshot

Also included with Intel® Parallel Studio and Intel® VTune™ Profiler products.

INTEL® VTUNE PROFILER

Single Thread

Optimize single-threaded performance.

Multithreaded

Effectively use all available cores.

System

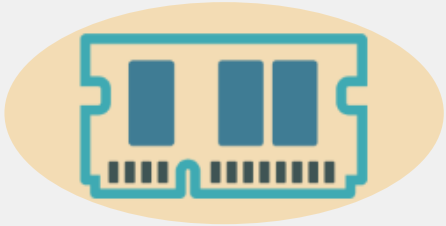
See a system-level view of application performance.

Media & OpenCL™ Applications

Deliver high-performance image and video processing pipelines.

HPC & Cloud

Access specialized, in-depth analyses for HPC and cloud computing.



Memory & Storage Management

Diagnose memory, storage, and data plane bottlenecks.

Analyze & Filter Data

Mine data for answers.

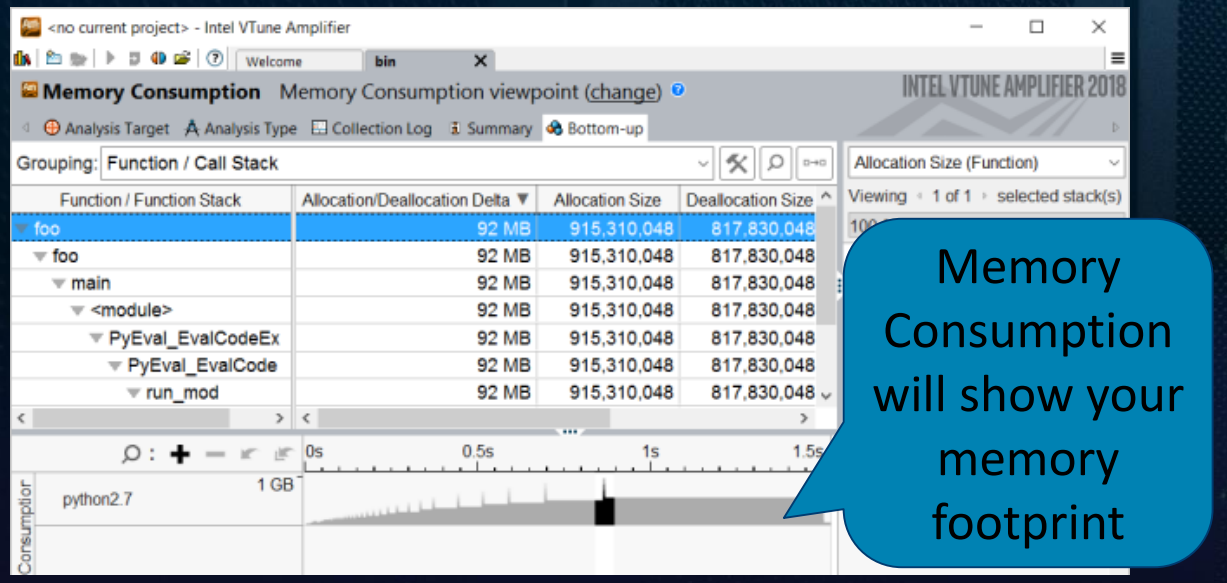
Environment

Fits your environment and workflow.

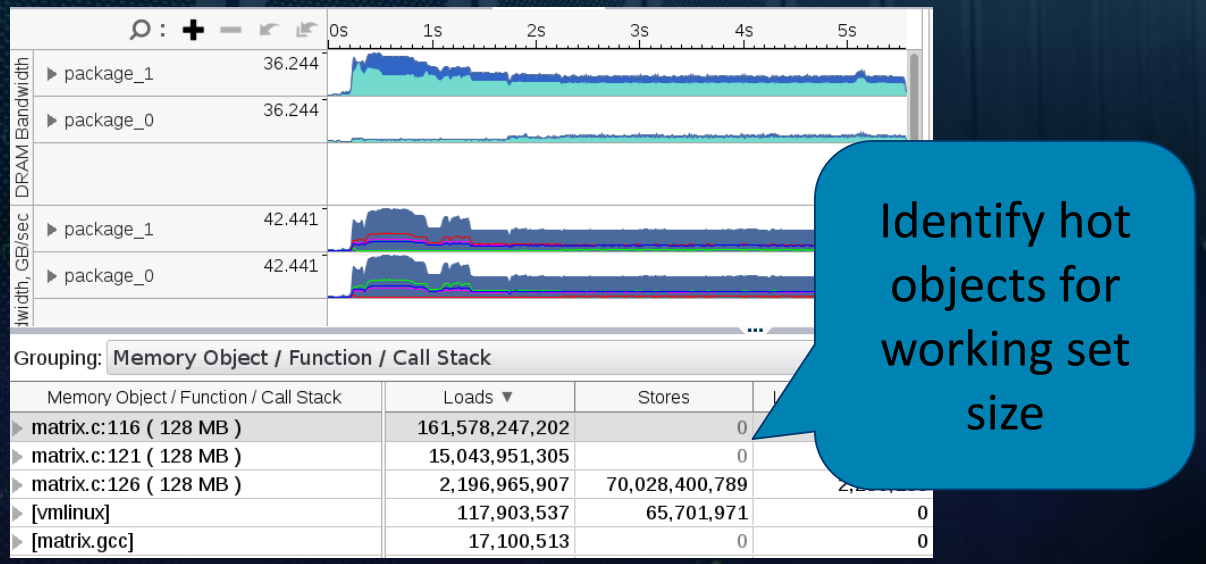
INTEL® VTUNE™ PROFILER MEMORY ANALYSIS

WILL BIGGER MEMORY HELP WITHOUT CODE CHANGES?

- Use memory in Memory Mode. Look for applications with:
  - A memory footprint larger than DRAM
  - A hot working set size smaller than DRAM

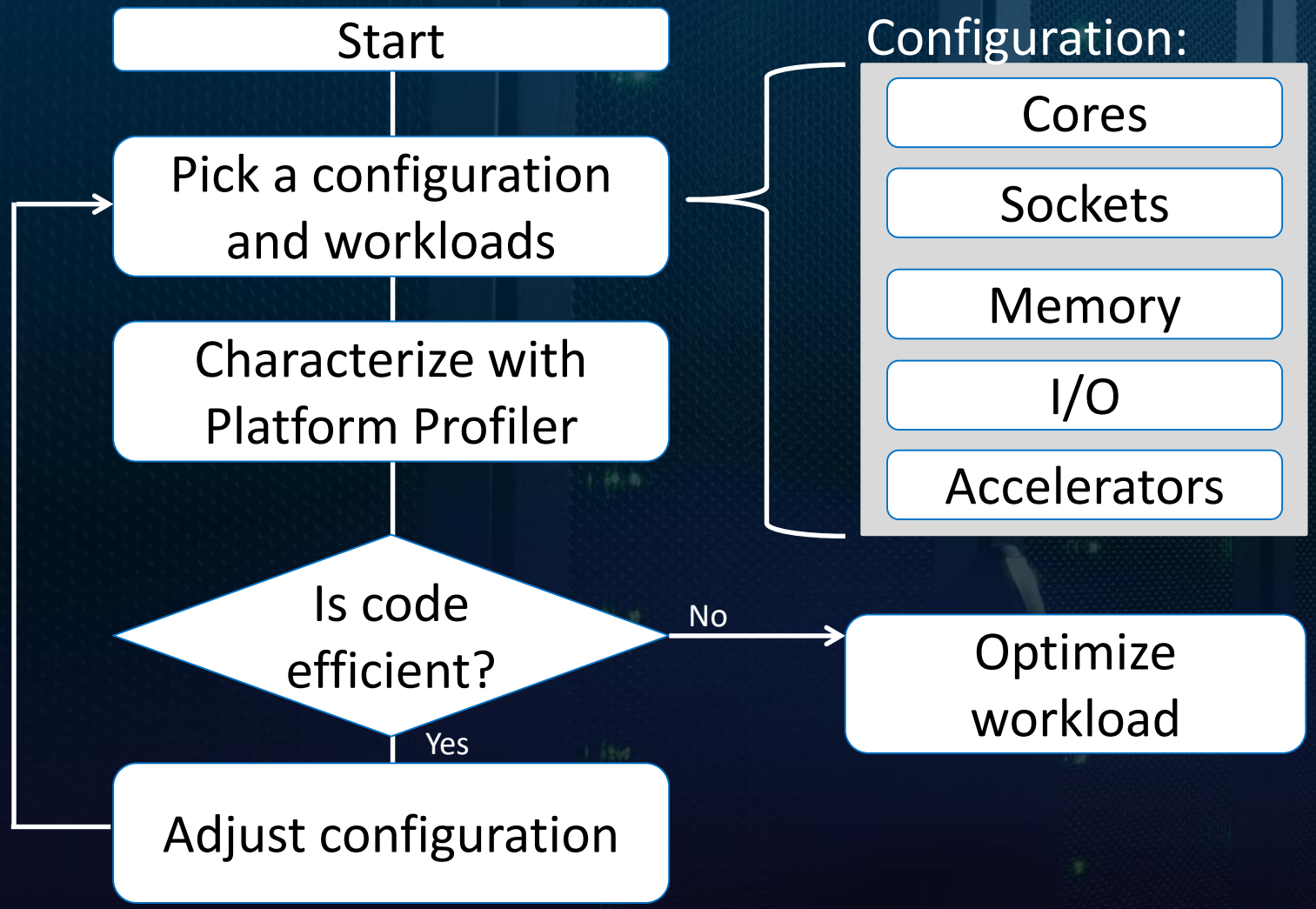


Memory Consumption Analysis



Memory Access + Dynamic Memory Object Analysis

CONFIGURATION WORKFLOW



INTEL® VTUNE PROFILER - PLATFORM PROFILER

TUNE WORKLOADS & SYSTEM CONFIGURATION

Finds

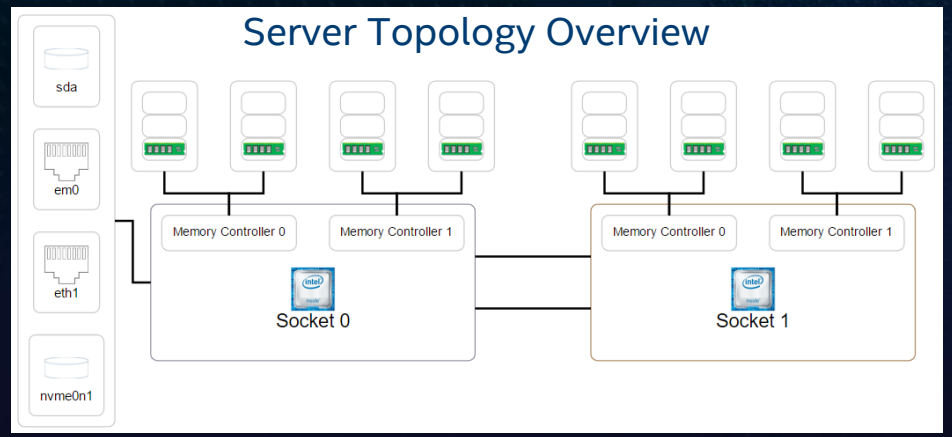
- Configuration issues
- Poorly tuned software

Target Users

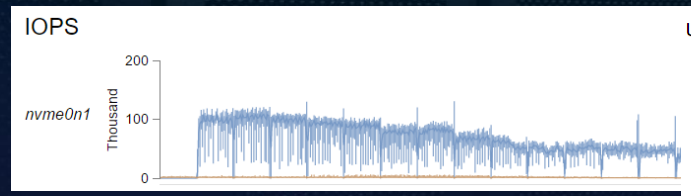
- Infrastructure Architects
- Software Architects & QA

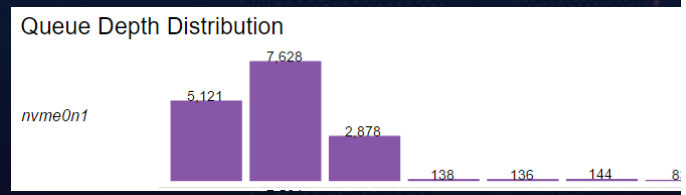
Performance Metrics

- Extended capture (minutes to hours)
- Low overhead – coarse grain metrics
- Sampling OS & hardware performance counters
- RESTful API for easy analysis by scripts

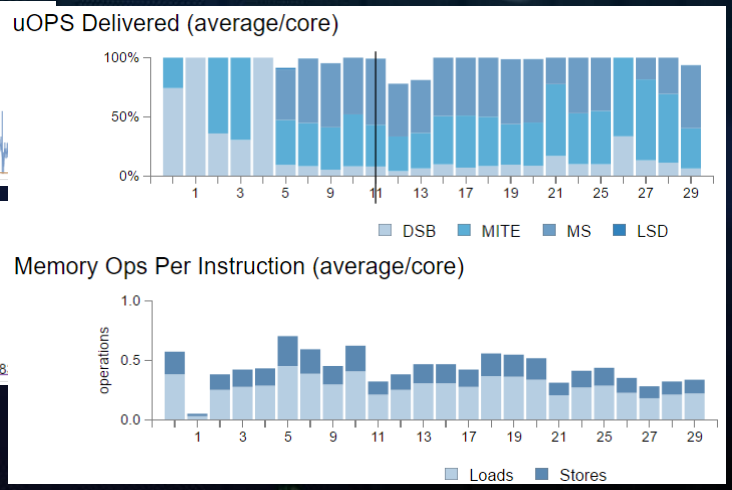


Timelines & Histograms



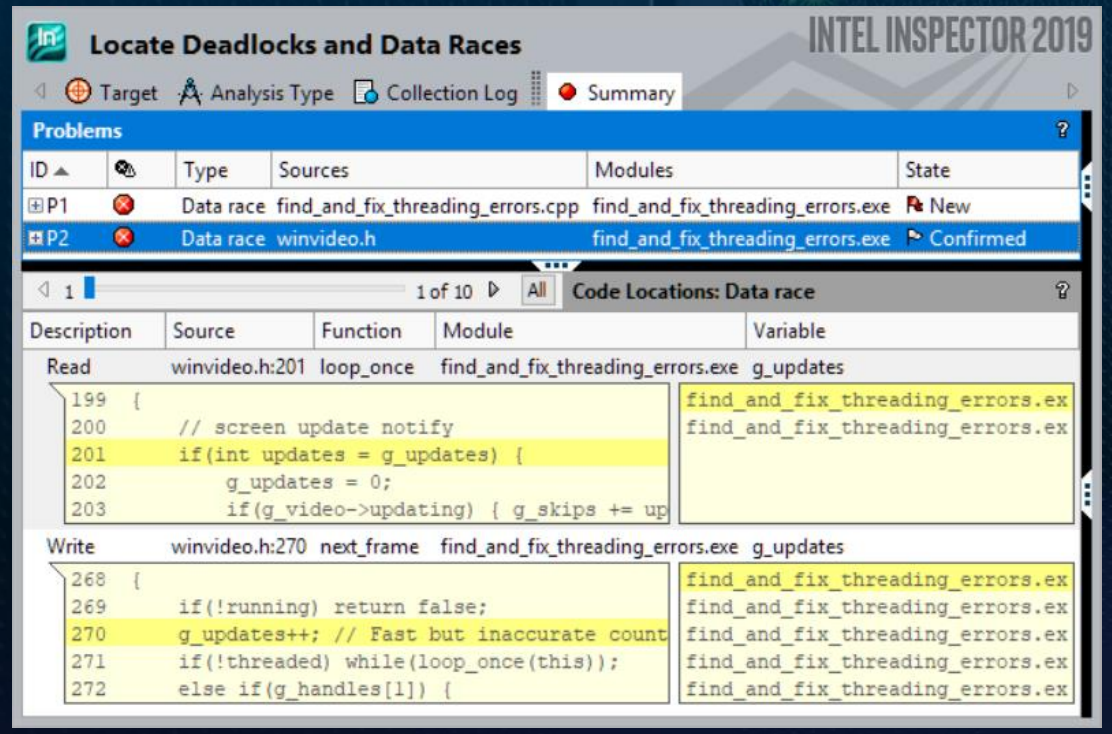


Core to Core Comparisons



DEBUG MEMORY & THREADING WITH INTEL® INSPECTOR

FIND & DEBUG MEMORY LEAKS, CORRUPTION, DATA RACES, DEADLOCKS



Correctness Tools Increase ROI by 12%-21%[1]

- Errors found earlier are less expensive to fix
- Races & deadlocks not easily reproduced
- Memory errors are hard to find without a tool

Debugger Integration Speeds Diagnosis

- Breakpoint set just before the problem
- Examine variables and threads with the debugger

What's New

Find Persistent Memory Errors

- Missing / redundant cache flushes
- Missing store fences
- Out-of-order persistent memory stores
- PMDK transaction redo logging errors

Learn More: bit.ly/intel-inspector

[1]**Cost Factors – Square Project Analysis -** *CERT: U.S. Computer Emergency Readiness Team, and Carnegie Mellon CyLab NIST: National Institute of Standards & Technology: Square Project Results*

DEVELOPER EXPERIENCE

New usages

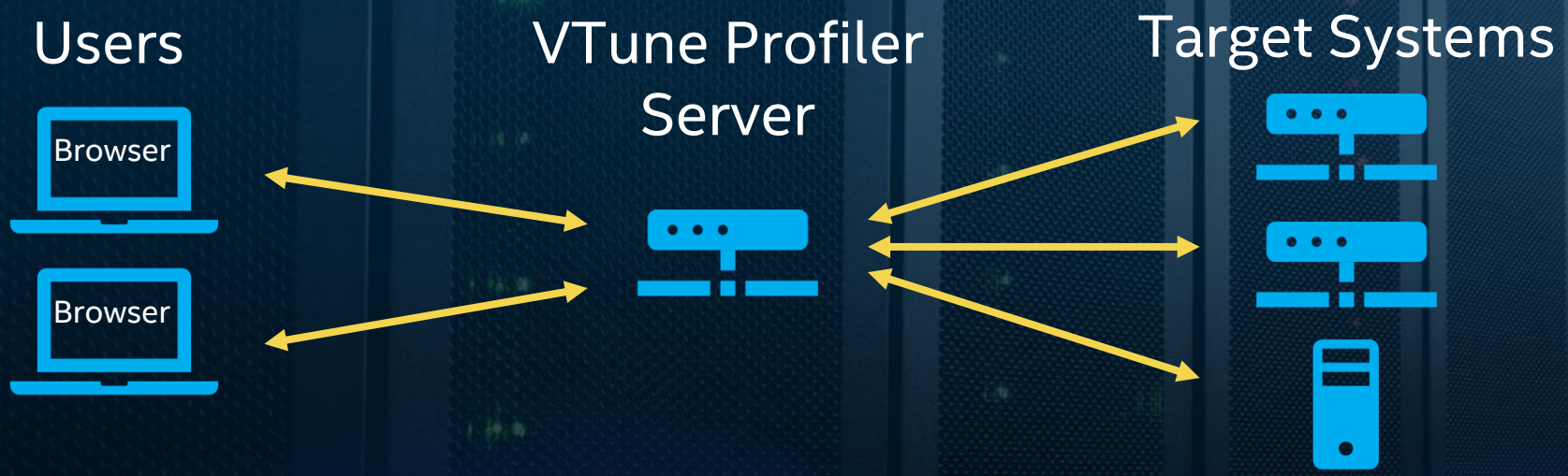
Developer productivity

Integrate Analyses

Automate & Simplify Usage

INTEL® VTUNE™ PROFILER SERVER ARCHITECTURE

JUST LAUNCH YOUR BROWSER AND GO



Easier profiling

- **Access with a web browser** – no install required by users
- **Share results** – all results available to all users with server access
- **Profile any target** – pre-install collector agent or push from server via SSH

EASIER SETUP, MORE INTELLIGIBLE RESULTS

INTEL® VTUNE™ PROFILER

Fresh, Accessible Analysis Setup

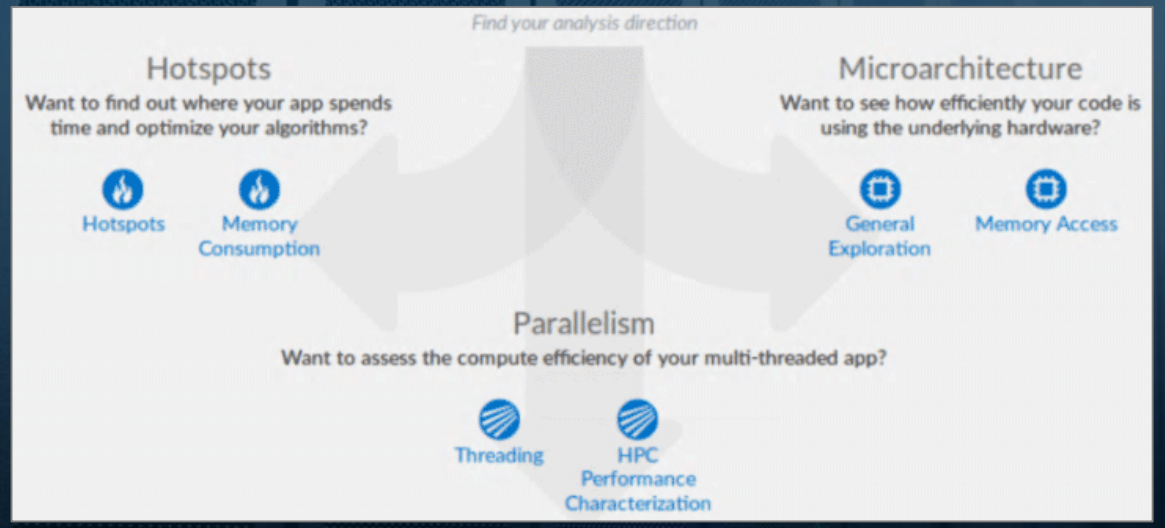
- Simplified workflow
- More familiar terminology
- More logical groupings

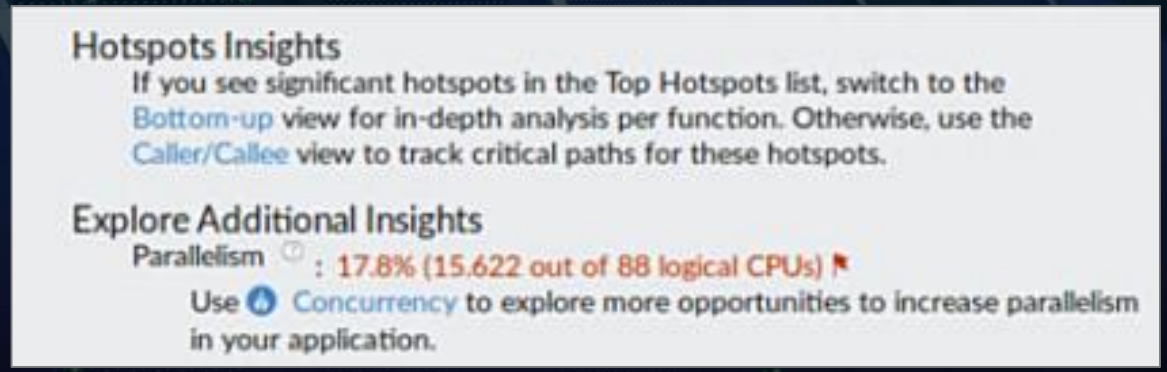
Performance Insights

- Suggestions for further analysis

Improved Displays

- New hardware pipeline display





DOWNLOADS & TECHNICAL ARTICLES

Persistent Memory Home Page

- software.intel.com/persistent-memory

Intel VTune Profiler

- Free Download
- Learn: Platform Profiler, Memory Access Programming, I/O Profiling

Intel Inspector – Persistence Inspector

- Free Download
- Learn: How to Detect Persistence Errors

CONCLUSION

Intel offers a rich and comprehensive suite of analysis capabilities to meet the needs of developing high performance software

Not every role requires use of the all capabilities; it is important to understand what to use when

We like to hear from you and help with best practices with use of tools

Q&A

Thank You!

Backup