DEVELOPER TOOLS FOR OPTANE DC PERSISTENT MEMORY

Director, Intel Developer Products

INTEL® VTUNE™ AMPLIFIER – PERFORMANCE 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.

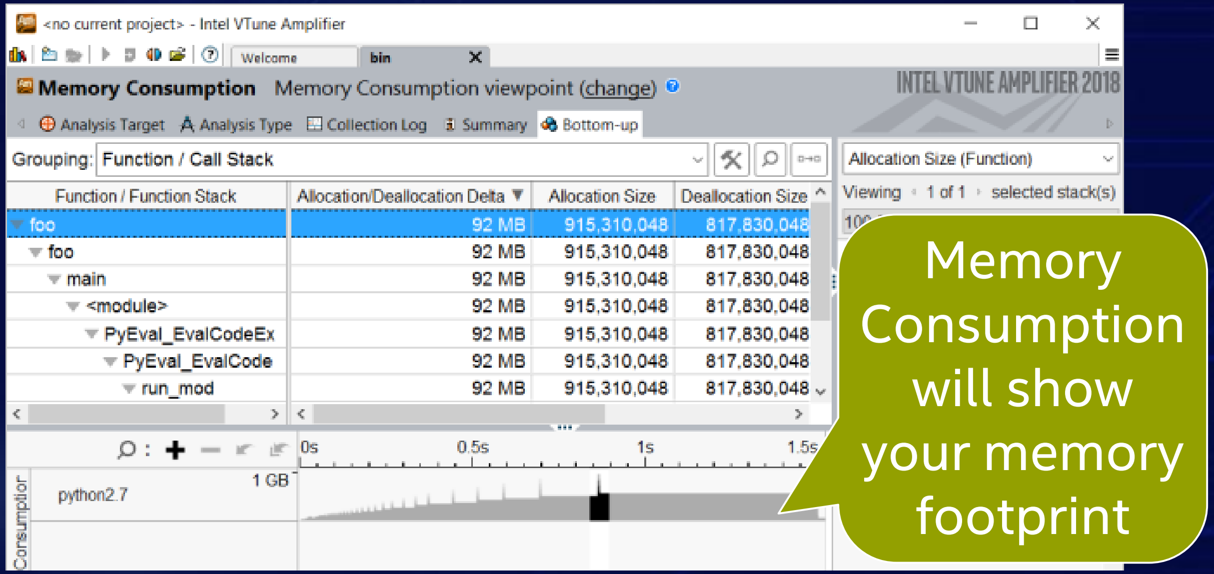
DEVELOPMENT STEPS FOR INTEL® OPTANE™ DC PERSISTENT MEMORY

1. Assess the suitability of an application
2. Pick the best memory configuration for it
3. Design it for Optane DC Persistent Memory
4. Tune its performance

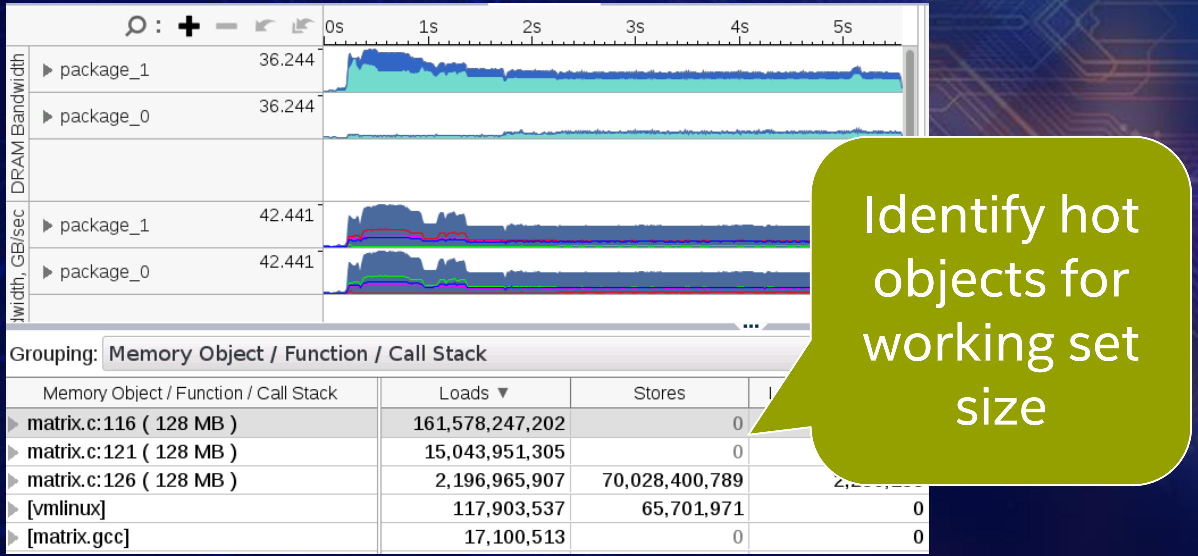
Intel® VTune™ Amplifier Memory Analysis

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

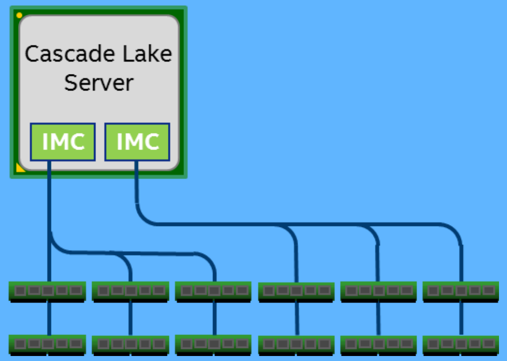


Memory Access + Dynamic Memory Object Analysis

WHICH CONFIGURATION IS BEST FOR YOUR APP?

Tradeoff: DRAM / DCPMM bandwidth and cost

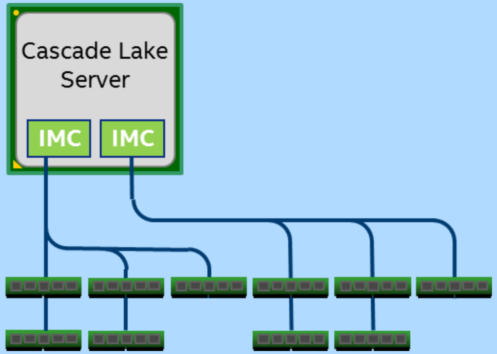
2-2-2



12 slots per CPU

Max memory capacity and bandwidth

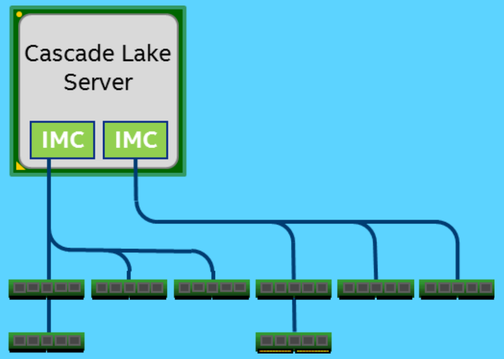
2-2-1



10 slots per CPU

Trade DDR bandwidth for smaller board real estate on the DIMM slots

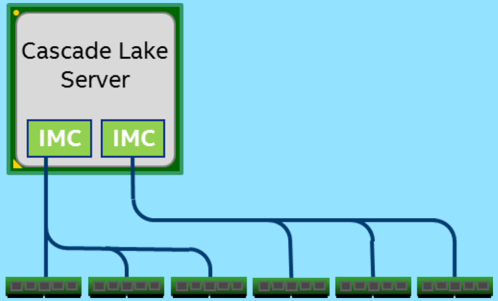
2-1-1*



8 slots per CPU

Trade DDR bandwidth for smaller board real estate on the DIMM slots

1-1-1



6 slots per CPU

Least number of DIMM slots to utilize max memory bandwidth

* No difference on functionality or performance when 2nd DIMM slot is in channel 0, 1 or 2 for that integrated memory controller (IMC)

† DIMM slots shown. While DRAM DIMMs can populate all slots shown, DCPMM is only populated in slot closest to CPU in each channel.

Intel® VTune™ Amplifier Platform Profiler

TUNE THE SYSTEM CONFIGURATION

Finds

- Configuration issues
- Opportunities for large persistent memory
- Poorly tuned software

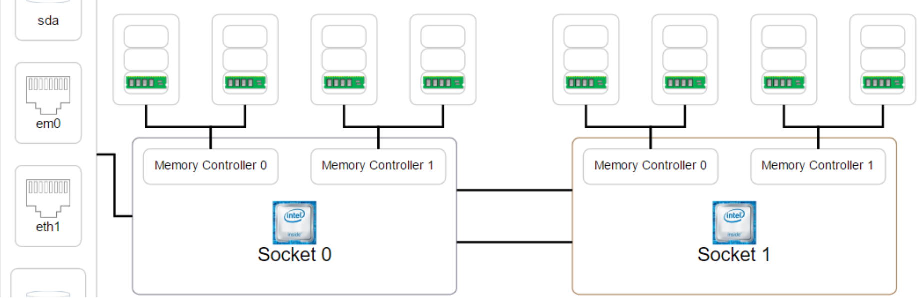
Target Users

- Infrastructure Architects
- Software Architects & QA

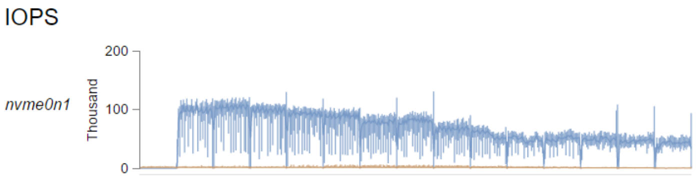
Performance Metrics

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

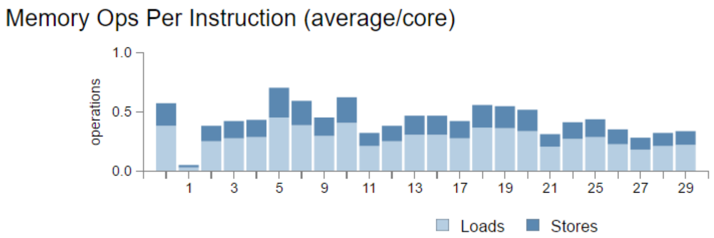
Server Topology Overview



Timelines & Histograms



Core to Core Comparisons



Intel® VTune™ Amplifier Memory Analysis

WILL BIGGER MEMORY HELP IF I MAKE CODE CHANGES?

Use app direct

Identify objects to allocate in Intel® Optane™ DC Persistent Memory

- Objects smaller than LLC – allocate in Persistent Memory (likely to be cached)
- Objects larger than DRAM – allocate in Persistent Memory (avoid swap)

Grouping: Bandwidth Domain / Bandwidth Utilization Type / Memory Object / Allocation Stack

Bandwidth Domain / Bandwidth Utilization Type / Memory Object / Allocation Stack	CPU Time	Memory Bound	Loads	Stores	LLC Miss Count	Average Latency (cycles)
DRAM, GB/sec	9.703s	64.3%	6,517,0...	4,141,26...	191,811,508	92
High	4.253s	56.8%	2,345,0...	2,111,23...	119,007,140	115
lin_stream.cpp:100 (152 MB			910,002...	887,613,...	51,453,087	119
lin_stream.cpp:99 (152 MB			826,002...	770,011,...	39,902,394	91
lin_stream.cpp:98 (152 MB			609,001...	452,206,...	27,651,659	142
[Unknown]			0	1,400,021	0	0
Medium	2.880s	70.3%	2,765,0...	981,414,...	52,853,171	83
Low	2.571s	71.6%	1,407,0...	1,048,61...	19,951,197	57

Hottest objects in Cache/DRAM

Warm objects in Persistent memory

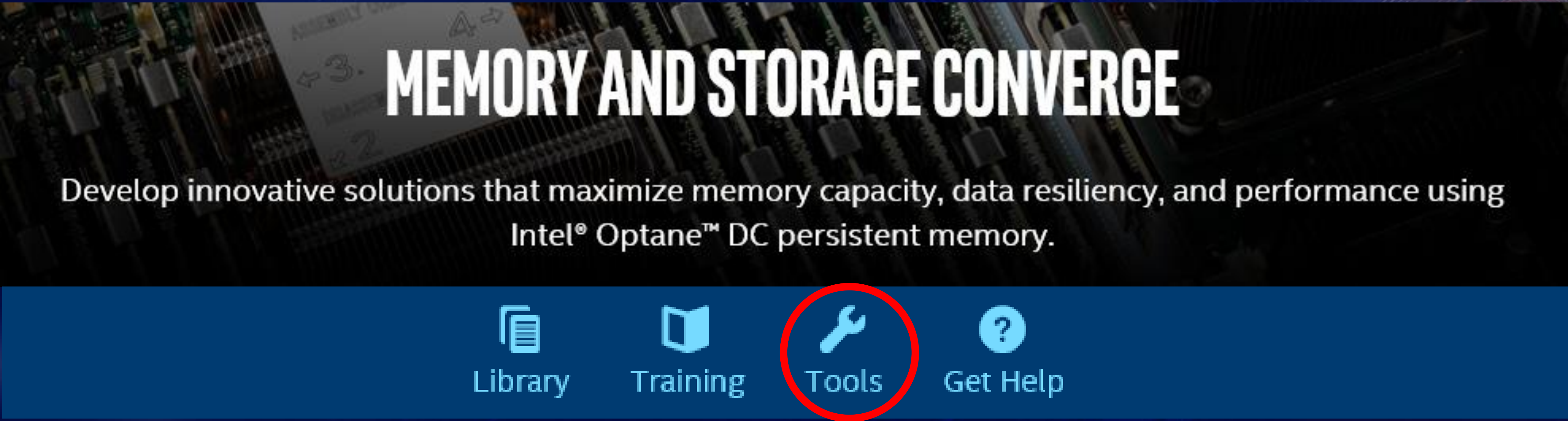
Cold objects on disk

Memory Access Analysis + Dynamic Memory Object Analysis

intel

DOWNLOADS & TECHNICAL ARTICLES

software.intel.com/persistent-memory



Click