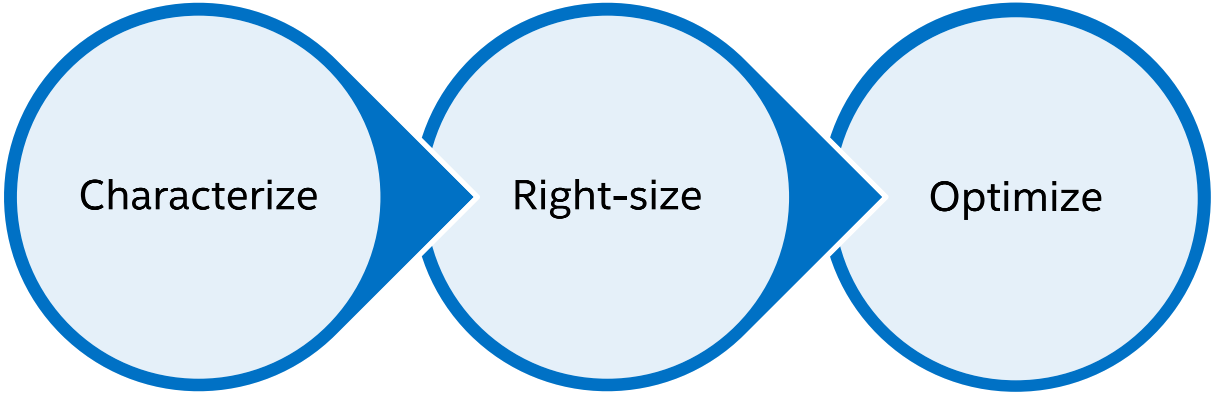
OPTIMIZE SYSTEM CONFIGURATIONS AND WORKLOADS FOR INTEL® OPTANE™ DC PERSISTENT MEMORY

Jackson Marusarz
Senior Technical Consulting Engineer
Compute Performance & Developer Products Division

3-STEPS FOR OPTIMIZING COMPLEX WORKLOADS



- What is my workload doing
- Where are the bottlenecks

- What configuration changes should I make

- How do I get the most out of my platform

THE LONG & SHORT OF PERFORMANCE ANALYSIS

GET THE BIG PICTURE FIRST WITH A SNAPSHOT OR PLATFORM PROFILER

	Snapshot Quickly size potential performance gain. Run a test "during a coffee break".	**In-Depth** Advanced collection & analysis. Insight for effective optimization.
Application Focus • HPC App developer focus • 1 app running during test	VTune Amplifier's **Application Performance Snapshot** L🕑	**VTune Amplifier** • Many profiles S-M🕑 **Intel Advisor** • Vectorization S🕑 **ITAC** • MPI Optimization S-L🕑
System Focus • Deployed system focus • Full system load test	VTune Amplifier's **Storage Performance Snapshot** L🕑	**VTune Amplifier** - System-wide sampling S-M🕑 - Platform Profiler L🕑

Maximum collection times: L🕑=long (hours) M🕑=medium (minutes) S🕑=short (seconds-few minutes)

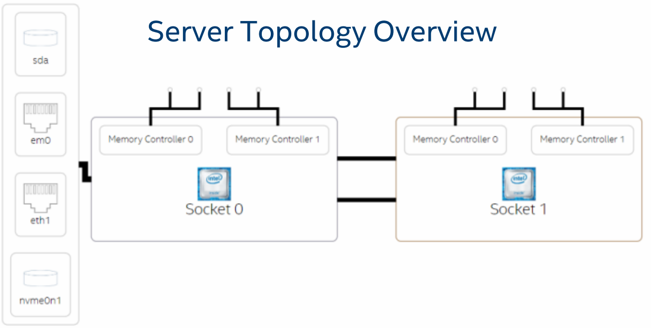
INTEL® VTUNE™ AMPLIFIER - PLATFORM PROFILER

Targets infrastructure and software architects

Longer runs – system-wide data

Interactive Topology Diagrams

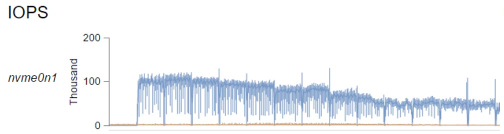
- System configuration
- Memory channel configurations

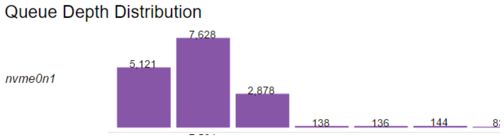


Performance metrics

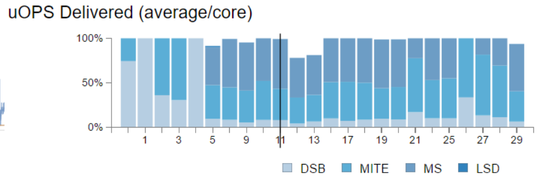
- Low overhead (targeting < 1%) – coarse grain
- Sampling OS and HW performance counters
- Extended capture (min. to hours)
- Open Data model with RESTful API for easy analysis by scripts

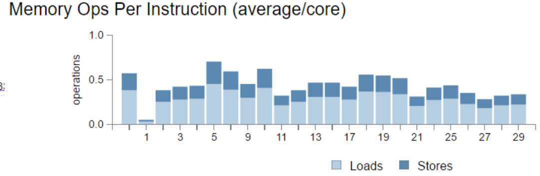
Timelines and Histograms





Core to Core Comparisons





INTEL® VTUNE™ AMPLIFIER PLATFORM PROFILER

A guided tour

DEMO

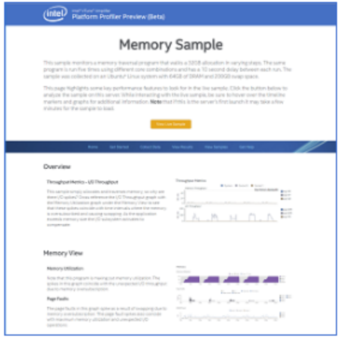
See Platform Profiler in Action

Matrix Multiplication

This sample was collected using an application derived from the Intel® Math Kernel Library Matrix Multiplication C Sample. The sample is not heavily CPU-bound and demonstrates efficient memory usage and throughput. However, the application is only running on a few cores and could benefit from parallelization. The sample was collected on an Ubuntu* Linux system with 64GB of DRAM and 36 cores.

File Copy

This sample was collected while continuously copying 30+ randomly sized files for five minutes on socket 1 followed by five minutes on socket 0. It shows a marked increase in efficiency when using socket 0 and demonstrates that applications that use a lot of file operations can benefit from using the socket closest to storage. The sample was collected on an Ubuntu* Linux system with two sockets and one NVMe* disk.



Memory Sample

This sample monitors a memory traversal program that walks a 32GB allocation in varying steps. The same program is run five times using different core combinations and has a 10 second delay between each run. The sample was collected on an Ubuntu* Linux system with 64GB of DRAM and 200GB swap space.

COLLECTING DATA

Data Collector

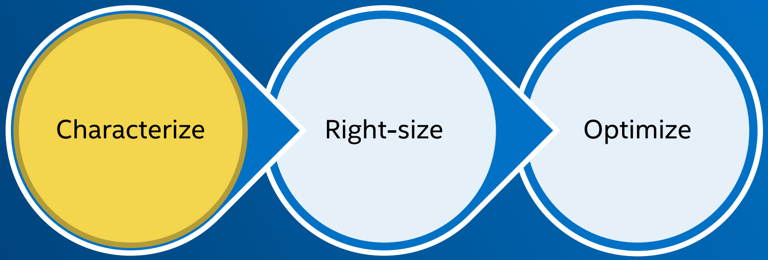
Platform Profiler Analysis

Data Collection “Targets”

- Windows*
- Linux*

Analysis and visualization

- Web-based UI
- Works best with Chrome



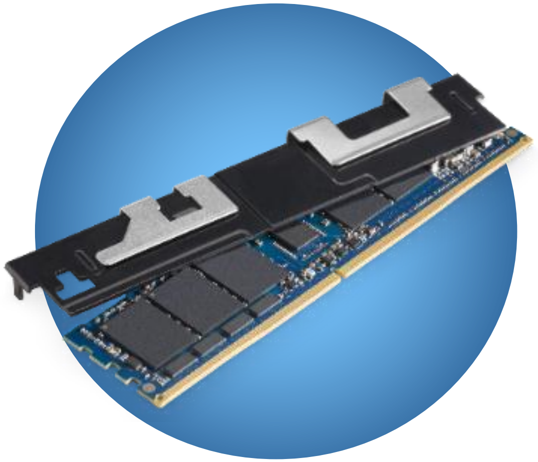
IDENTIFYING OPPORTUNITIES

Characterizing workloads for Intel® Optane™ DC Persistent Memory

THE BEST OF BOTH WORLDS WITH INTEL® OPTANE™ DC PERSISTENT MEMORY

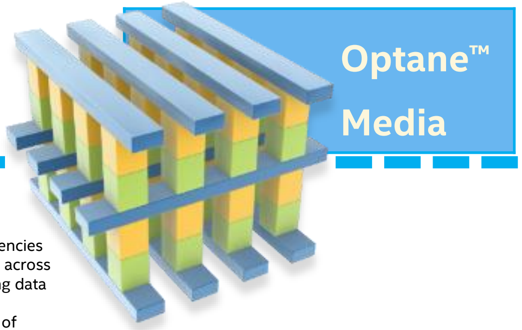
DRAM ATTRIBUTES

*Perf*ormance comparable to DRAM at *low latencies*[1]



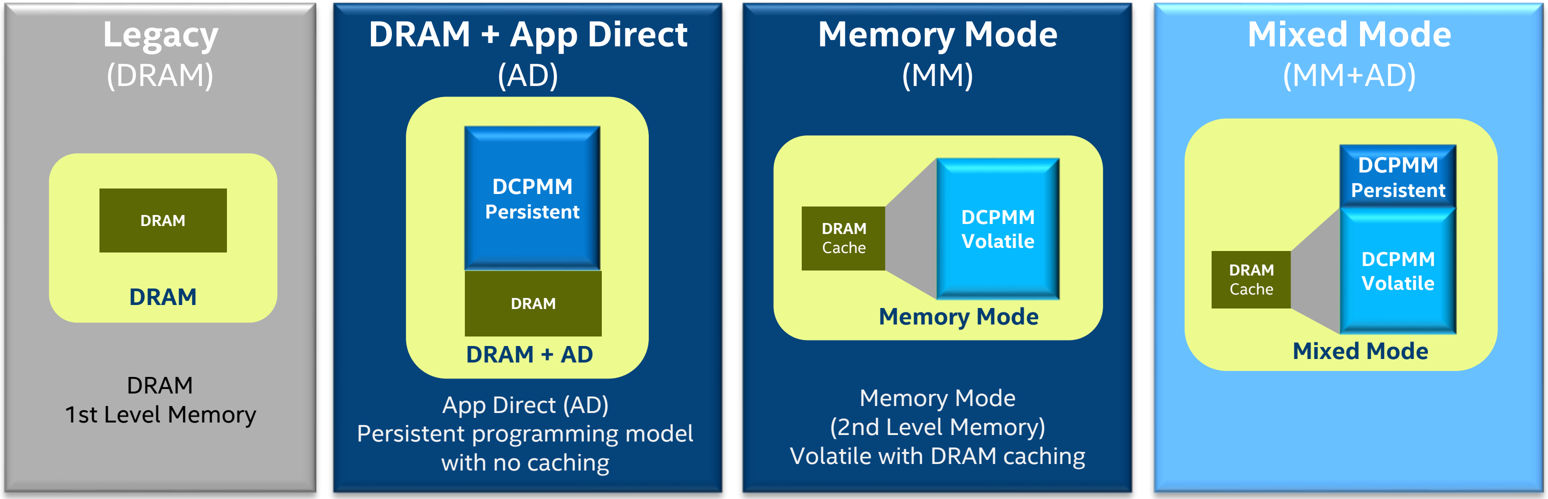
NAND SSD ATTRIBUTES

Data persistence with higher capacity than DRAM[2]



1. "*Fast performance comparable to DRAM*" - Intel persistent memory is expected to perform at latencies near DDR4 DRAM. Benchmarks and proof points forthcoming. "*low latencies*" - Data transferred across the memory bus causes latencies to be orders of magnitude lower when compared to transferring data across PCIe or I/O bus' to NAND/Hard Disk. Benchmarks and proof points forthcoming.
2. Intel persistent memory offers 3 different capacities – 128GB, 256GB, 512GB. Individual DIMMs of DDR4 DRAM max out at 256GB.

PROGRAMMING MODELS



Volatile	Persistent or Volatile	Volatile	Volatile or Mixed

WHAT TO LOOK FOR

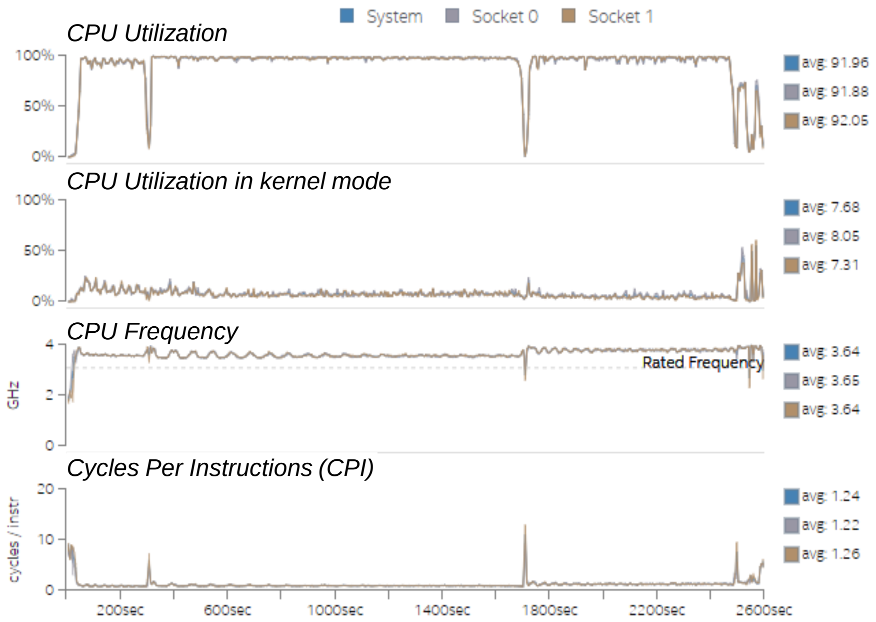
Memory and storage behavior

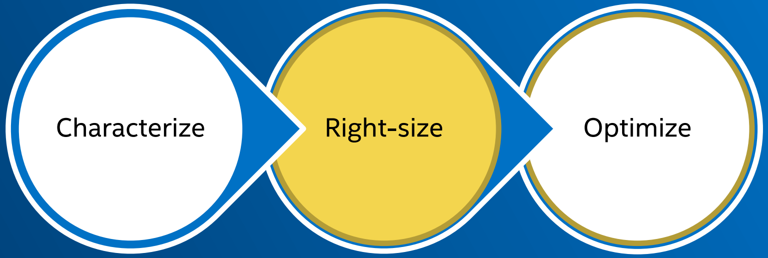
- Memory utilization
- Page faults
- Memory read/write ratios
- Memory throughput
- Disk activity

WHAT TO LOOK FOR

Performance impact

- CPU utilization (user/kernel)
- Cycles per instructions (CPI)
- Non-uniform memory architecture (NUMA) and inter-socket traffic

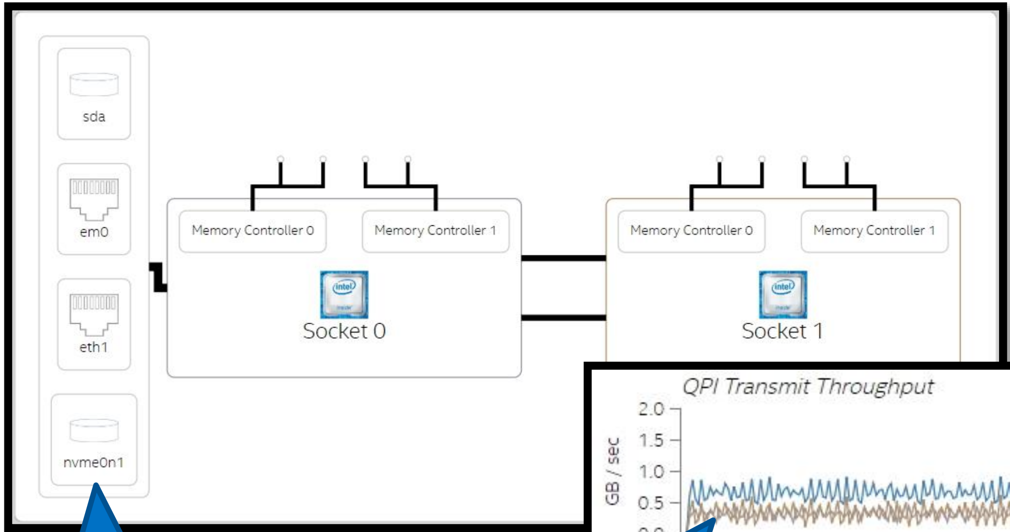


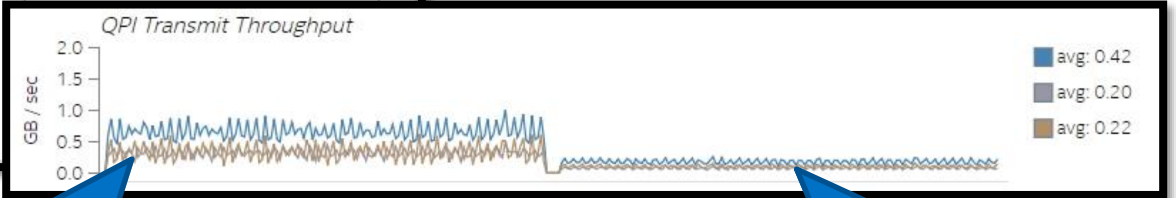


UNDERSTANDING CONFIGURATIONS AND PERFORMANCE

Finding the right system for your workloads

CONFIGURATION MATTERS!





5% performance drop when running from Socket 1 which is “far” from the storage device

Storage device connected to Socket 0

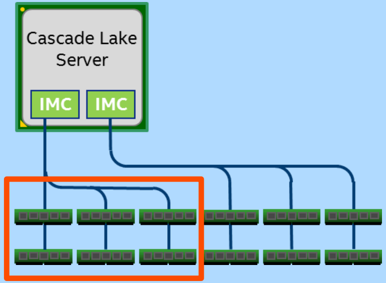
Interconnect traffic from an application running on Socket 1 doing IO

Interconnect traffic from an application running on Socket 0 doing IO

WHICH CONFIGURATION IS BEST FOR YOUR APP?

Tradeoff: DRAM / DCPMM bandwidth and cost

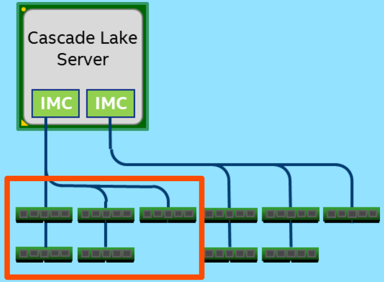
2-2-2



12 slots per CPU

Max memory capacity and bandwidth

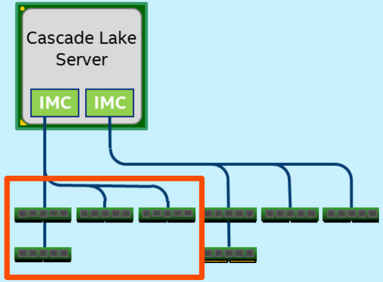
2-2-1



10 slots per CPU

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

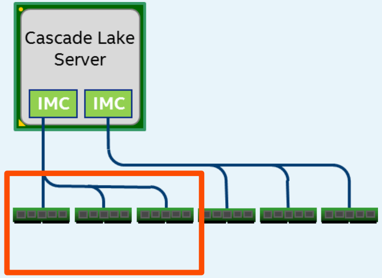
2-1-1*



8 slots per CPU

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

1-1-1

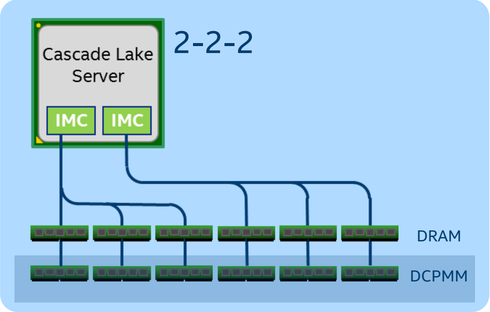


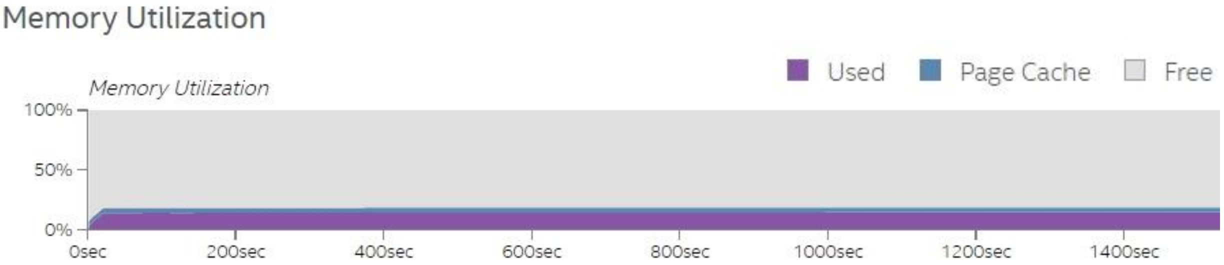
6 slots per CPU

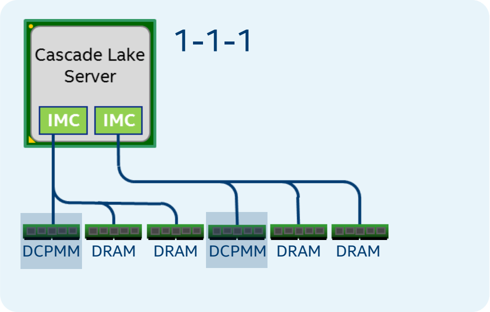
Least number of DIMM slots to utilize max memory bandwidth

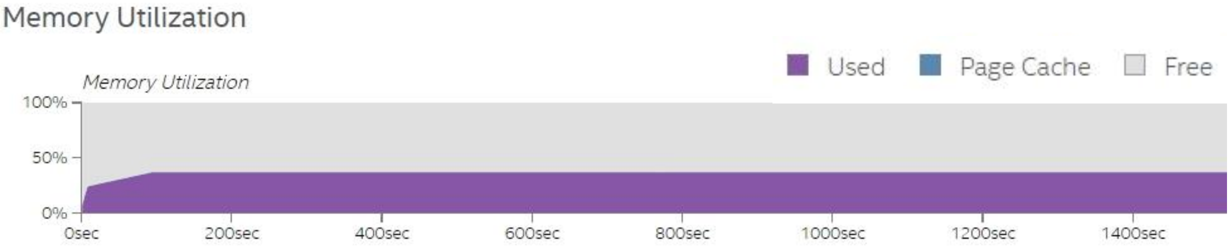
* No difference on functionality or performance when 2nd DIMM slot is in channel 0, 1 or 2 for that integrated memory controller (IMC)
† DIMM slots shown. While DRAM DIMMs can populate all slots shown, DCPMM is only populated in slot closest to CPU in each channel.

MEMORY UTILIZATION

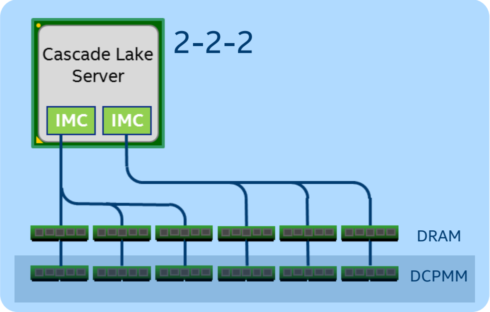


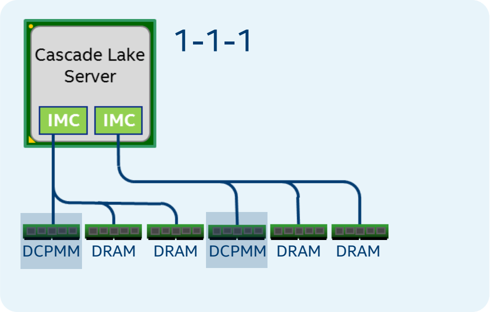


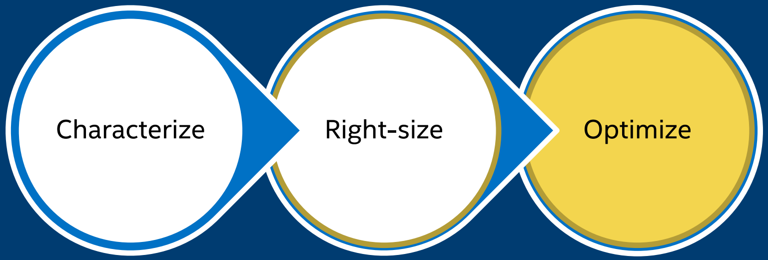




MEMORY THROUGHPUT







UNDERSTANDING CONFIGURATIONS AND PERFORMANCE

Finding the right system for your workloads

PERFORMANCE METRICS

Memory throughput

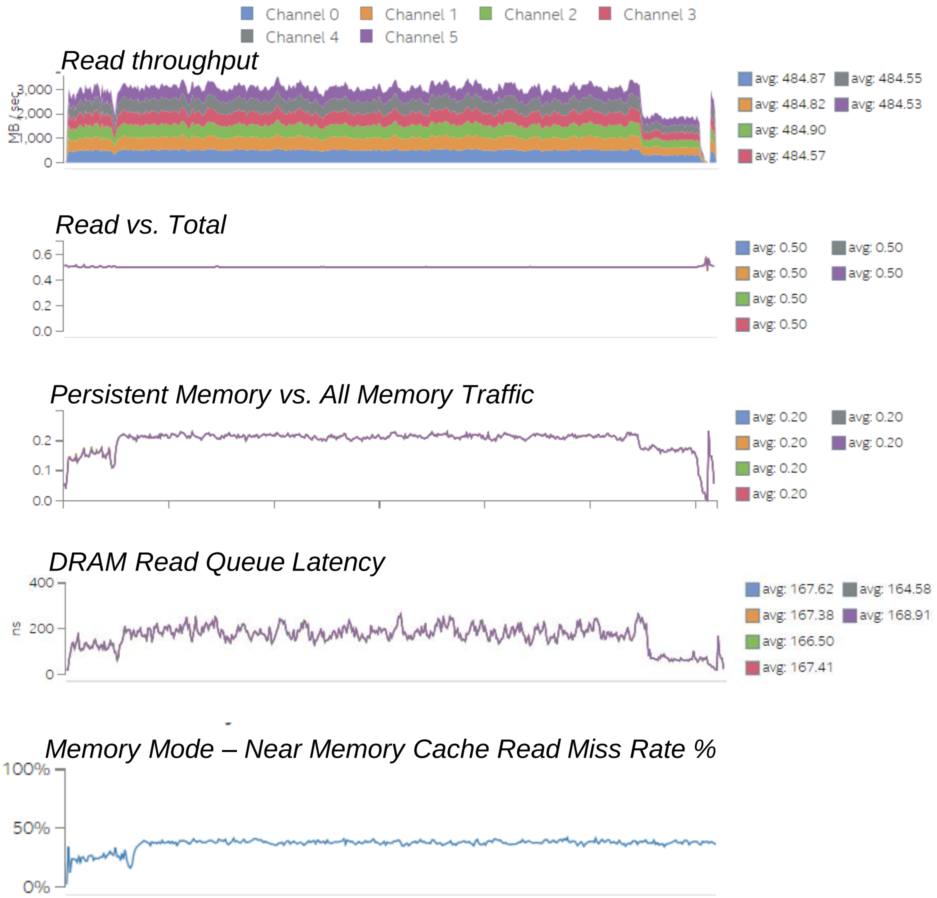
- Read/write throughput
- Read/write ratio
- Persistent vs. volatile

Memory latency

- DRAM read/write latencies
- Persistent memory read/write latencies

Memory Mode

- Near memory cache miss rate

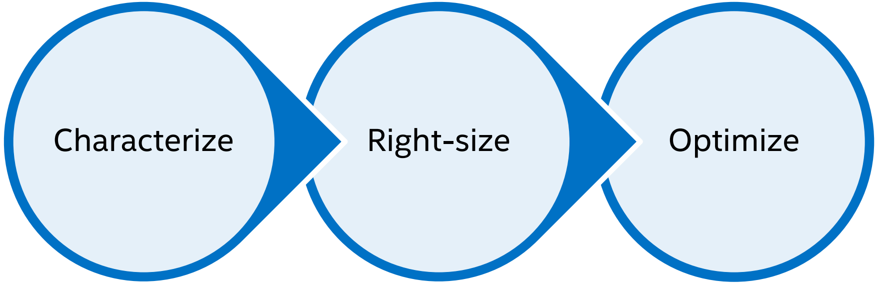


SUMMARY

What’s next

SUMMARY

Have a plan



Use Intel® VTune™ Amplifier – Platform Profiler

- To get useful insights into system behavior and performance
- To identify opportunities and optimize for Intel® Optane™ DC Persistent Memory

DOWNLOAD THE TOOLS

Software tools for Intel® Optane™ DC Persistent Memory
Free downloads and technical articles

- software.intel.com/persistent-memory/tools

Intel® VTune™ Amplifier – Performance Profiler

- software.intel.com/vtune

Intel® Inspector – Persistence and Thread Debugger

- software.intel.com/inspector

LEGAL DISCLAIMER & OPTIMIZATION NOTICE

Optimization Notice

Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

Notice revision #20110804

intel®