Product Insights
Intel® Optane™ DC persistent memory (DCPMM) & FUJITSU Server PRIMERGY

BIG MEMORY BREAKTHROUGH FOR THE BIGGEST DATA CHALLENGES

CONTENT

INTRODUCTION 2
FUJITSU, DCPMM AND NEXTGENIO PROJECT 2
BENEFITS 3
OPERATION MODES OF DCPMM 3
USE CASES 5
SPECIFICATION 5
SERVER SUPPORT 6
IMPLEMENTATION REQUIREMENTS 6
PERFORMANCE 6
INTRODUCTION

Every day, the amount of data created across the world is exploding to new levels. Businesses thrive on this data to make critical decisions, gain new insights, and differentiate services. The demand for memory capacity is growing at an insatiable rate and there is a need to keep larger amounts of data closer to the CPU.

Intel® Optane™ DC persistent memory (DCPMM) represents a groundbreaking technology innovation that delivers a unique combination of affordable large memory capacity and persistence (non-volatility). DCPMM memory provides a new memory technology tier for FUJITSU Server PRIMERGY and PRIMEQUEST systems configured with 2nd Generation Intel® Xeon® Scalable processors.

The technology that dominates traditional main memory, DRAM, is fast to access, but small, expensive, and volatile. Storage is large, cheap, persistent, but is slow to access. There is a huge latency and bandwidth penalty as you jump from RAM based memory and disk based storage. The ever increasing amount of data and the need to access more of it quickly have further magnified the gap. Intel's breakthrough product, Intel® Optane™ DC persistent memory, is disrupting the traditional memory-storage hierarchy by creating a new tier to fill the memory-storage gap providing greater overall performance, efficiency, and affordability.

FUJITSU, DCPMM AND NEXTGENIO PROJECT

NEXTGenIO is an R&D project funded by the European Commission that is develop solutions to high performance computing’s (HPC) I/O and data challenges. The consortium partners are EPCC, Intel, Fujitsu, Technische Universität Dresden, Barcelona Supercomputing Center, the European Centre for Medium-Range Weather Forecasts, Arm (formerly Allinea) and Arctur.

Using a requirements-driven co-design process involving all the stakeholders of the NEXTGenIO project, Fujitsu and Intel developed a prototype system using DCPMM memory. As key output of the project, a 34-node cluster has been delivered in June 2019 by Fujitsu and installed at the University of Edinburgh’s Advanced Computing Facility, where it is now in the final the testing and integration phase before going into full production during the summer.

The new PRIMERGY and PRIMEQUEST servers configured with 2nd Generation Intel® Xeon® Scalable processors benefit from Fujitsu’s pioneering work in the NEXTGenIO project. The success of this EU project made it possible for Fujitsu to introduce these unprecedented I/O capabilities to industry standard x86 PRIMERGY and PRIMEQUEST servers.
**BENEFITS**

- **SSD meets RAM:** The new 3D-XPoint structure combines the advantages of RAM and NAND-SSD. The DCPMM modules offer higher data throughput and faster access times than SSDs. In contrast to SSDs, the performance is significantly higher with a full data carrier.

- **Lower Cost:** The new DCPMM modules offer significantly lower cost/GB than conventional DRAMs. Due to the possible server consolidation (up to 2.5x more memory per server compared to DRAM only), software license and support costs can also be reduced.

- **Data Persistence:** With DCPMMs, database applications have the ability to store large amounts of data completely in persistent memory and not on disk. A further advantage is that there is no loss of data in the event of a power failure or restart of the system, and the recovery times after power failures are significantly faster.

- **Responsive data query:** If your processor needs files from a mechanical hard drive, it must wait a few milliseconds for it to continue working. SSDs react faster, especially when the processor requests many bits at the same time. After a little delay, the data arrives in quick succession. Intel® Optane™, on the other hand, always delivers memory in the shortest possible time.

---

**OPERATION MODES OF DCPMM**

The DCPMM modules operate in three different modes:

1. **Memory Mode (MM)** - 100% of DCPMM memory acts as system memory
   In this mode, DCPMMs act as volatile system memory, while DRAM DIMMs act as cache. Only DCPMM capacity is displayed as system memory in this mode. The total displayed volatile system memory in this mode is the sum of DCPMM capacity.

2. **App Direct Mode (AD)** - 0% of DCPMM memory acts as system memory
   In this mode, DCPMMs act as independent and persistent memory resources directly accessible by specific applications, and DRAM DIMMs act as system memory. The total displayed volatile system memory in this mode is the sum of DRAM DIMM capacity.

3. **Mixed Mode** - 1-99% of DCPMM memory acts as system memory
   In this mode, some percentage of DCPMM capacity is directly accessible to specific applications (App direct), while the rest serves as system memory. The App direct part of DCPMM is displayed as persistent memory, while the rest of DCPMM capacity is displayed as system memory. DRAM DIMMs act as cache in this mode. The total displayed volatile system memory in this mode is the DCPMM capacity that is invested in volatile system memory.
Memory Mode (MM): When configured for Memory Mode, the applications and operating system perceives a pool of volatile memory, no differently than it does today on DRAM-only systems. In this mode, no specific persistent memory programming is required in the applications, and the data will not be saved in the event of a power loss. In Memory Mode, the DRAM acts as a cache for the most frequently-accessed data, while the Intel® Optane™ DC Persistent Memory provides large memory capacity. Cache management operations are handled by the Intel® Xeon® Scalable processor’s memory controller. When data is requested from memory, the memory controller first checks the DRAM cache, and if the data is present, the response latency is identical to DRAM. If the data is not in the DRAM cache, it is read from the Intel® Optane™ DC Persistent Memory with slightly longer latency. Virtualization can benefit from Intel® Optane™ DC persistent memory in Memory Mode because there is larger memory capacity which provides more VMs and more memory per VM at a lower cost compared to DRAM. Workloads that are I/O bound can also benefit from using Memory Mode as the Intel® Optane™ DC persistent memory provides larger memory capacity which supports larger databases and at a lower cost compared to DRAM. Memory Mode seamlessly brings large memory capacity at affordable cost points to legacy applications.

App Direct Mode: In App Direct Mode, software and applications have the ability to talk directly to the Intel® Optane™ DC persistent memory, which reduces complexity in the stack. There is the option of having App Direct Mode use legacy storage APIs. This allows it to act like an SSD and can boot an OS. The operating system sees Intel® Optane™ DC persistent memory and DRAM as two separate pools of memory. It is persistent like storage, byte addressable like memory, cache coherent which extends the usage of persistent memory outside the local node, and consistent low latency supporting larger datasets. The power of persistent memory adds business resilience to systems with faster restart times because data is retained even during power cycles. Memory bound workloads benefit from Intel® Optane™ DC persistence with its large capacity and higher endurance and greater bandwidth compared to NAND SSDs. In-memory databases, in-memory analytics frameworks and ultrafast storage applications are good examples of workloads that greatly benefit from using App Direct Mode.

Mixed Mode: A sub-set of App Direct, can be provisioned so that some of the Intel® Optane™ DC persistent memory is in Memory Mode and the remaining is in App Direct Mode. In Mixed Mode, applications can take advantage of high performance storage without the latency of moving data to and from the I/O bus.
USE CASES
DCPMM offers advantages for a variety of applications:
Memory Mode seamlessly brings large memory capacity at affordable cost points to legacy applications. Virtualized database deployments and big-data analytics applications are great candidates for Memory Mode.
In-memory databases, in-memory analytics frameworks and ultrafast storage applications are good examples of workloads that greatly benefit from using App Direct Mode.

SPECIFICATION
Intel® Optane™ DC persistent memory uses the same form factor as traditional memory DIMMs (DRAM), but offers higher capacity, greater affordability, and data persistence. DCPMM modules are available in capacities of 128GB, 256GB, and 512GB. The module fits into standard DDR4 DIMM slots of servers equipped with selected 2nd Generation Intel® Xeon® Gold, Platinum and Silver processor SKUs.
- DCPMMs are installed in standard memory slots in supported servers
- Supported on Gold, Platinum and selected Silver Cascade Lake-SP SKUs
- 2666 MT/s memory bus speed. Any 2933 MT/s DDR4 DIMMs installed will also operate at 2666 MT/s
- Optional data encryption using AES 256-bit encryption
- Double Device Data Correction (DDDC)
SERVER SUPPORT
A maximum of 6 slots are available for DCPMM modules per CPU socket (please see relevant system configurator for details):

<table>
<thead>
<tr>
<th>System</th>
<th>Memory Slots per CPU</th>
<th>Memory Slots per System</th>
<th>DCPMM Slots per CPU</th>
<th>DCPMM Slots per System</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRIMERGY TX2550 M5</td>
<td>6</td>
<td>12</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>PRIMERGY EX25x0 M5</td>
<td>8</td>
<td>16</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>PRIMERGY RX2530 M5</td>
<td>12</td>
<td>24</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>PRIMERGY RX2540 M5</td>
<td>12</td>
<td>24</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>PRIMERGY RX4770 M5</td>
<td>12</td>
<td>48</td>
<td>6</td>
<td>24</td>
</tr>
<tr>
<td>PRIMEQUEST 3800B2</td>
<td>12</td>
<td>96</td>
<td>6</td>
<td>48</td>
</tr>
<tr>
<td>PRIMEQUEST 3800E2</td>
<td>12</td>
<td>96</td>
<td>6</td>
<td>48</td>
</tr>
</tbody>
</table>

IMPLEMENTATION REQUIREMENTS
- DCPMMs require second generation Intel Xeon Scalable Family processors. First generation Xeon Scalable processors are not supported.
- Platinum processors, Gold processors and selected Silver processors support DCPMM.
- Only one type of DCPMM is allowed per system.
- Only one DCPMM package and one Memory package is allowed per CPU.
- Every CPU has to have the same DCPMM & Memory configuration.
- Maximum 6 DCPMMs per processor (install 1 in each memory channel).
- If a channel is equipped with DCPMM and DRAM, the DCPMM must always be plugged into the slot near the CPU (Slot 1).
- When using DCPMM, IMC#0 must be equipped with at least one DCPMM and one DRAM.
- Depending on the operating mode of the DCPMM, the following rules apply according to Intel:
  - App Direct Mode (AD): no additional restrictions.
  - Memory Mode (MM): DRAM:DCPMM = 1:4 to 1:16.
  - MM+AD mode: no 3DS DRAMs.

PERFORMANCE
ISC'19, Fujitsu and Intel, together with ECMWF, EPCC, Arctur and the other partners of the NEXTGenIO project, disclosed the latest breakthrough performance results using Intel Optane DC persistent memory across various supercomputing applications:
- The European Centre for Medium-Range Weather Forecasts (ECMWF) achieved 10-times higher bandwidth when its Fields Database, which holds the meteorological data for medium-range weather forecasts, was stored in persistent memory and distributed across multiple computing nodes. Use of compute notes equipped with Intel Optane DC persistent memory accelerated ECMWF’s global weather forecasts and a reduced number of I/O nodes needed to run its models.
- The Arctur HPC center, in partnership with Barcelona Supercomputer Center, achieved a 2x3 speed-up in simulating 3D models of an electric lightweight aircraft, which reduced its OpenFOAM runtimes by 50% on 16 nodes.
- EPCC achieved 2-times higher throughput on the CASTEP material science application when running its code on computing nodes equipped with Intel Optane DC persistent memory, accelerating material science research across multiple domains.
INTEL® OPTANE™ DC PERSISTENT MEMORY
CHANGING THE MEMORY AND STORAGE PARADIGM

10X HIGHER BANDWIDTH
INTO RESULTS DATABASE
VS. CONVENTIONAL STORAGE SYSTEMS
ACCENETERS GLOBAL WEATHER FORECASTING ...
WITH FEWER I/O NODES

2X SPEED UP
VS. CONVENTIONAL STORAGE SYSTEMS
REDUCING SIGNIFICANT WRITE OVERHEAD INTO FILE SYSTEM
REDUCING OPENFOAM
RUNTIME BY 50%

2X HIGHER THROUGHPUT
ON SAME NUMBER OF NODES
VS. DDR BASED SYSTEMS
ACCELERATING MATERIAL
SCIENCE RESEARCH
ACROSS MULTIPLE DOMAINS

Performance results based on testing by EPCC. See slide 13 for system configuration details as provided by EPCC.
The NEXTGenIO project is funded by the European Union's Horizon 2020 Research and Innovation program under Grant Agreement no. 671951