Intel® Processors and Processor Cores based on Golden Cove Microarchitecture

Instruction Throughput and Latency README

March 2022

Revision 2.0

350392-002US

Revision History		
Document ID	**Description**	**Date**
350392-001US	Initial Release	Feb-22
350392-002US	The csv and json files have been updated to correct some instructions and add additional instructions.	Mar-22

Notices & Disclaimers

Intel technologies may require enabled hardware, software or service activation.

No product or component can be absolutely secure.

Your costs and results may vary.

You may not use or facilitate the use of this document in connection with any infringement or other legal analysis concerning Intel products described herein. You agree to grant Intel a non-exclusive, royalty-free license to any patent claim thereafter drafted which includes subject matter disclosed herein.

All product plans and roadmaps are subject to change without notice.

The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

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.

Code names are used by Intel to identify products, technologies, or services that are in development and not publicly available. These are not "commercial" names and not intended to function as trademarks.

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document, with the sole exception that you may publish an unmodified copy. You may create software implementations based on this document and in compliance with the foregoing that are intended to execute on the Intel product(s) referenced in this document. No rights are granted to create modifications or derivatives of this document.

README

Review all information in this README file prior to using the "Throughput and Latency" csv or JSON file.

csv file format:
Iform - "Iform" is the XED term for variants of instructions. Please consider them experimental and subject to change.
Xed url: https://intelxed.github.io/

Example:
ANDNPD_XMMxuq_MEMxuq
Instruction name: ANDNPD
Instruction sources: src0 XMM, src1 MEM

Regsize - source register size
Mask - this instruction does not use a mask (k register)

JSON file format:
This format contains additional information that helps identify the instruction and understand the iform:

"instruction": "unpcklps xmm12, xmm13"
"uniq_key": "450f14e5"
"iform": "UNPCKLPS_XMMps_XMMq"
uniq_key: the actual opcode bytes used to encode the instruction by the compiler

To print all the instructions in the JSON file:

jq '.[].instruction' tpt_lat-glc-client.json

To print all the instructions in the JSON file with "isa_extension": "FMA":

jq '.[] | select(.isa_extension=="FMA")' tpt_lat-glc-client.json

Throughput definition –

The number of clock cycles required to wait before the issue ports are free to accept the same instruction again.
This number can be lower than 1, e.g., 0.5 or 0.33, indicating that multiple instructions could be executed in parallel in a given cycle.
For instructions that execute at allocation, 1/alloc_width (1/6 for Golden Cove microarchitecture) was used as throughput in the tables.

Latency definition -

The number of clock cycles that are required for the CPU to complete the execution of all of the µops that form an instruction.

The file contains partial throughput / latency data:
a) Not all instructions appear in the database.
b) Not all instructions that do appear have both their throughput and their latency specified.
If you require data that is missing, we suggest using the values available in the previous generation throughput and latency document (available here: https://software.intel.com/en-us/articles/intel-sdm#optimization) as a reasonable approximation.

Note that actual throughput measured on Intel processors may vary by up to 0.1 cycles. Also, latency may sometimes vary due to dynamic microarchitectural conditions, and in this case we report the average rounded to the nearest integer value. In all cases we assume data read from DCU or written to store-buffers.