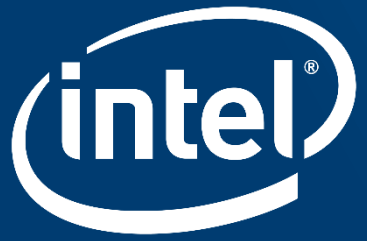


ISA-L® UPDATE & USERCASE SHARING

SPDK China Summit 2018

Xiaodong Liu (Xiaodong.liu@intel.com)

Chunyang Hui (Chunyang.hui@intel.com)

Legal Disclaimer

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration.

No computer system can be absolutely secure.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. For more complete information about performance and benchmark results, visit http://www.intel.com/benchmarks .

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit http://www.intel.com/benchmarks .

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.

Intel® Advanced Vector Extensions (Intel® AVX)* provides higher throughput to certain processor operations. Due to varying processor power characteristics, utilizing AVX instructions may cause a) some parts to operate at less than the rated frequency and b) some parts with Intel® Turbo Boost Technology 2.0 to not achieve any or maximum turbo frequencies. Performance varies depending on hardware, software, and system configuration and you can learn more at http://www.intel.com/go/turbo.

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.

Cost reduction scenarios described are intended as examples of how a given Intel-based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.

Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.

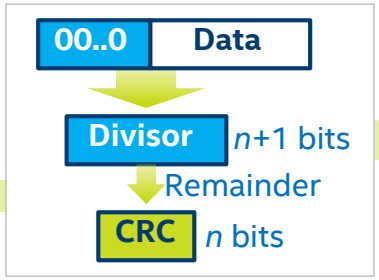
❑ Features Update & Usercase Sharing

- Encryption
- Stitching
- CRC64
- Compression
- Multi-buffer Hash

❑ State of the Project

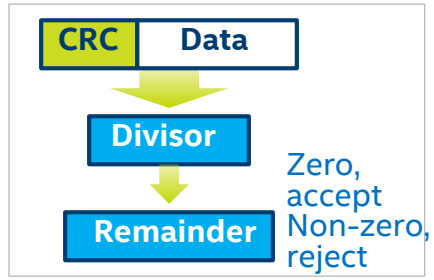
Intel® ISA-L **Functions**

XOR (RAID 5), **P+Q** (RAID 6), Reed-Solomon **Erasure Code**



Sender

CRC-T10, **CRC-IEEE** (802.3), **CRC32-iSCSI**, **CRC64**



Receiver

DATA PROTECTION

DATA INTEGRITY

PERFORMANCE OPTIMIZING

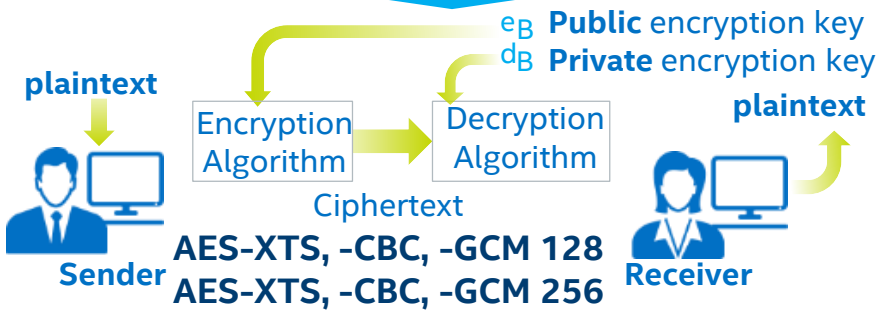
CRYPTOGRAPHIC HASHING

Multi-Buffer: SHA-1, SHA-256, SHA-512, MD5
Multi-Hash: SHA1, SHA1+Murmur

COMPRESSION "DEFLATE"

IGZIP: Faster DEFLATE (zlib)
Compression & Decompression

ENCRYPTION



AES-XTS, -CBC, -GCM 128
AES-XTS, -CBC, -GCM 256

ENCRYPTION

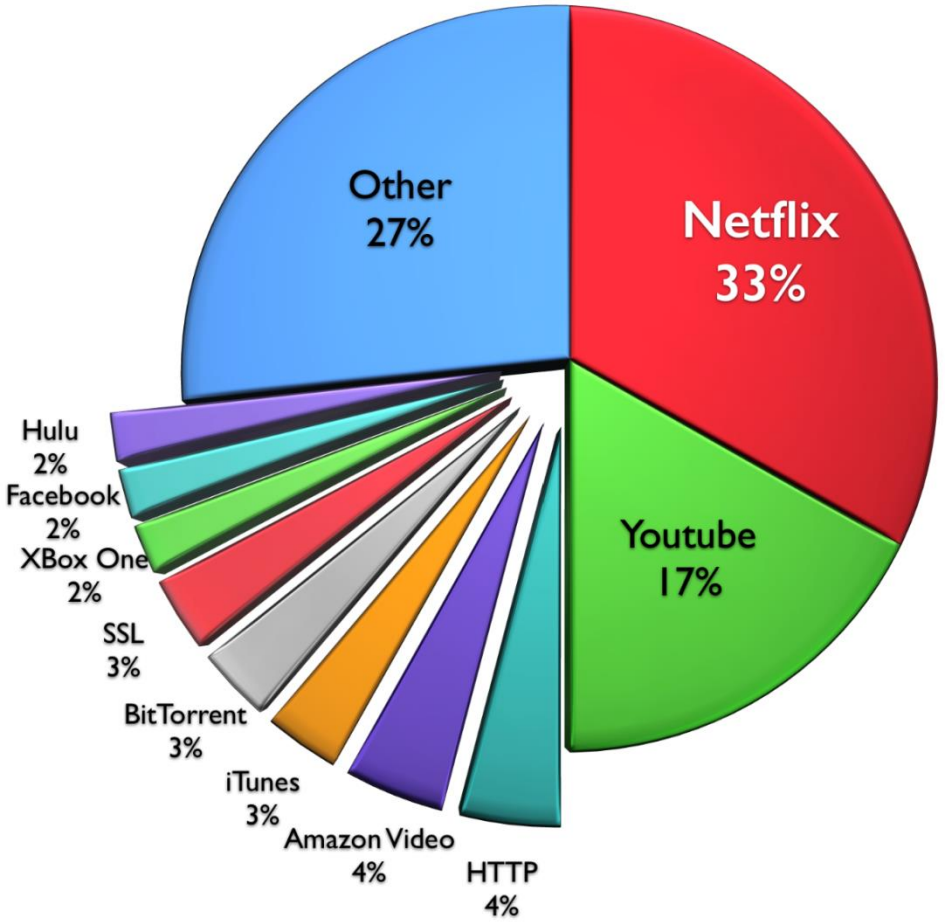
Encryption Cycle/Byte Performance on the Intel® Xeon® Processor Scalable Family

(cache cold cycle/byte)

ISA-L Function	Intel® Xeon® Platinum 8180 Processor @ 2.5 GHz 1 Socket			
	ISA-L		OpenSSL 1.0.2j	
	Cycle/Byte Performance (lower is better)	Single Core Throughput (higher is better)	Cycle/Byte Performance (lower is better)	Single Core Throughput (higher is better)
AES-XTS 128	0.64	3.8 GB/s	0.64	3.8 GB/s
AES-XTS 256	0.88	2.7 GB/s	0.89	2.7 GB/s
AES-CBC 128 Decode	0.64	3.8 GB/s	0.64	3.8 GB/s
AES-CBC 192 Decode	0.76	3.2 GB/s	0.76	3.2 GB/s
AES-CBC 256 Decode	0.88	2.7 GB/s	0.88	2.7 GB/s
AES-GCM 128	0.67	3.6 GB/s	1.58	1.5 GB/s
AES-GCM 256	0.89	2.7 GB/s	1.83	1.3 GB/s

** ISA-L function uses AVX512 instructions

Netflix & Intel: Background



North American Aggregate Internet Traffic
Sandvine 2016 Global Internet Phenomenon Report
https://www.sandvine.com/trends/global-internet-phenomena

Netflix pushes how many bits?

- Average of 35Tb/s all day, every day, and rising

And how do they do it?

- Built their own custom Content Delivery Network
- Vast majority of the library is served from boxes living in your local ISP/IXP
- Heterogeneous hardware, but all single socket, all FreeBSD based

How come?

- Saves vast amounts of backbone traffic
- Gives Netflix direct control at both ends of the wire
- Improves user experience

The Challenge

Design Goal:
Upgrade to 100Gbps per Open Connect Appliance

Curveball:
Add encryption (HTTPS/TLS) for streaming video to safeguard user privacy, too

Budget:
Do it cost effectively

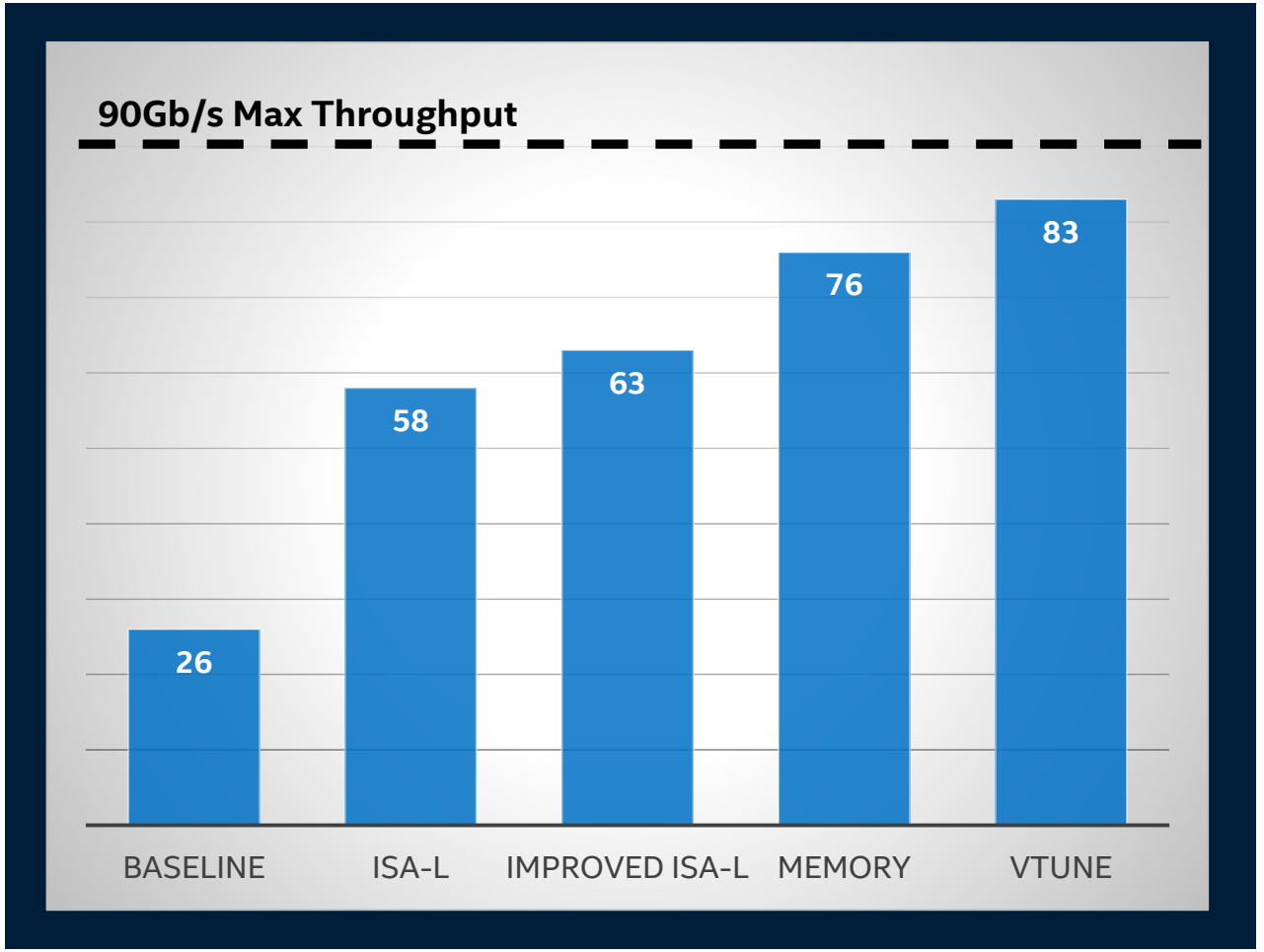
Netflix & Intel: Before and After

Started with OpenSSL
- Required compromises in their data path

Tried all the alternatives: BoringSSL, etc
- ISA-L was the fastest on the market
- Long-lived connections, only in the data path

ISA-L was tweakable
- Asked for non-temporal instructions: eureka!
- Identified the bottleneck: memory bandwidth
- Tuned the hardware
- … but it also fit the entire deployed infrastructure

Netflix* 2016 100G Flash OCA Performance



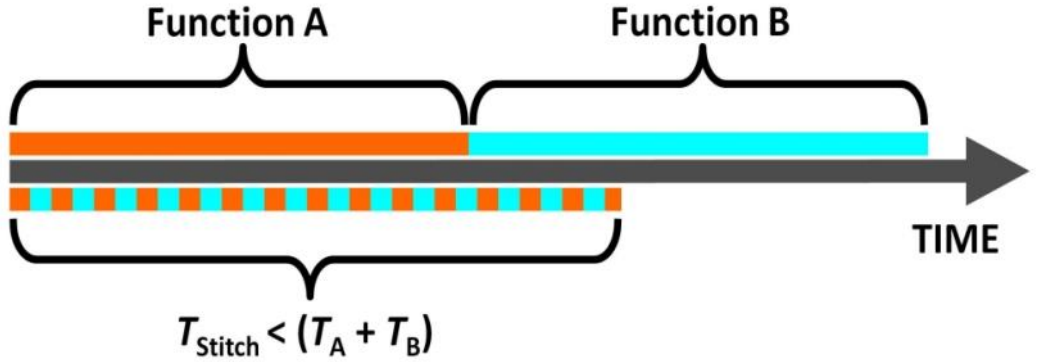
FUNCTIONS STITCHING

What is function stitching

- **Definition:**
  A method to interleave instructions from pairs of functions to maximize execution efficiency of the cores.

- **Advantages:**
  - In turn allows the processor core to execute instructions from both algorithms at the same time, makes better use of the execution resources, and results in a lower total execution time.
  - May also offer second-order benefits, such as only requiring data to be fetched from memory/caches once rather than twice.

Figure 1. Stitching Two Functions



Usercases

“ We were able to run with the new T10PI calculate and move routines you created. The API works well for us, and so does the routine. With this routine, we improved performance (in our usage) from ~3.0 GB/s to ~5.1 GB/s! ”

-- Data from a famous storage solution company in US

USING CRC64

CRC Update

- Definition: A Cyclic Redundancy Check (CRC) is the remainder, or residue, of binary division of a potentially long message, by a CRC polynomial. This technique is ubiquitously employed in communication and storage applications due to its effectiveness at detecting errors.

- Types
  - CRC16: T10 standard
  - CRC32: IEEE standard, gzip standard (IEEE reflected), iSCSI
  - CRC64: ECMA-182 standard (reflected/normal), ISO standard (reflected/normal) …

- ISA-L Specific
  - Self defined polynomial supported
  - 64 bit is as fast as 32 bit CRC
  - Instruction level supported for CRC32 iSCSI

Cycle/Byte Performance on the Intel® Xeon® Processor Scalable Family

(cache cold cycle/byte)

ISA-L Function	Intel® Xeon® Platinum 8180 Processor @ 2.5 GHz 1 Socket	
	ISA-L	
	Cycle/Byte Performance (lower is better)	Single Core Throughput (higher is better)
CRC16 T10	0.22	11.0 GB/s
CRC32 IEEE (802.3)	0.22	11.1 GB/s
CRC32 iSCSI	0.18	13.8 GB/s
CRC32 GZIP Reflective	0.24	10.3 GB/s
CRC64 Normal	0.22	11.0 GB/s
CRC64 Reflective	0.19	12.9 GB/s

** ISA-L function uses AVX512 instructions

CRC64 Usage Case

The Intel team helped this company by writing custom assemblers that worked with its generic polynomial implementation, which later was incorporated it into their software,

"The result of this was a reduction of 33% in the latency for the temp which has direct impact in the job performance. The performance of the iSCSI implementation of CRC also improved by 2.9x and it provided a 7.9x improvement in performance of CRC64."

-- data from a famous world wide software company

COMPRESSION

IGZIP: Fast Deflate

DEFLATE (aka zlib, gzip, pkzip, etc)

- Lossless compression, ubiquitous adoption
- Web servers & clients, MS Windows, DirectX & Office, Android apps, png images, pdf files.

Fast Compression

Novel, fully zlib-compatible implementation:

- Level 0-3 optimized for throughput
- 5X greater throughput than zlib-1, equal to lz4 and lzo
- 6% better compression ratio than lz4 and lzo with same throughput
- semi-dynamic compression allows granular tuning of compression ratio and throughput

Optimized Decompression

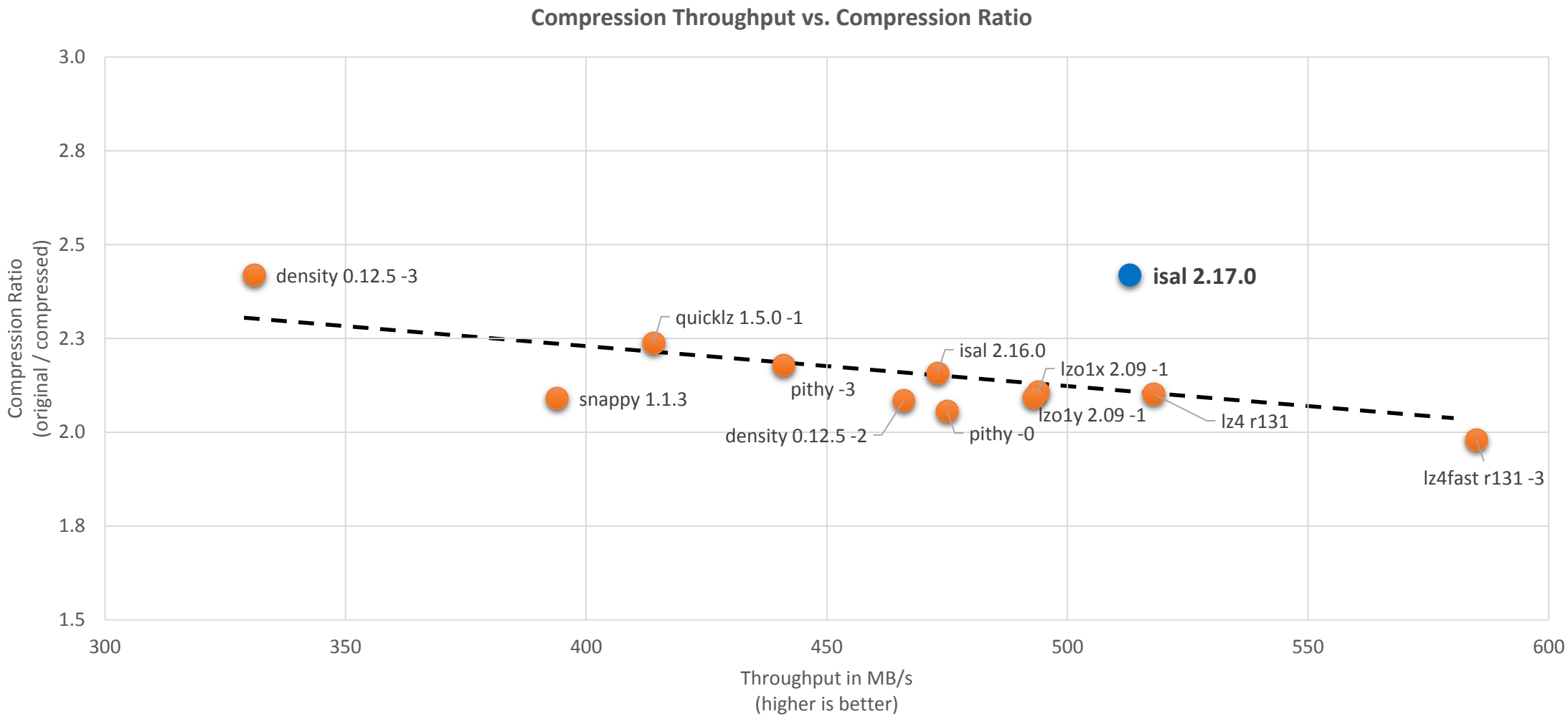
- >2X throughput vs. zlib, equal to lzo
- Fully compatible with zlib and gzip archives

Compressor name	Compression Throughput (MB/s)	Ratio
lz4fast r131 -3	585	50.51
lz4 r131	518	47.59
isal 2.17.0	513	41.35
snappy 1.1.3	394	47.83
zlib 1.2.8 -1	82	36.45

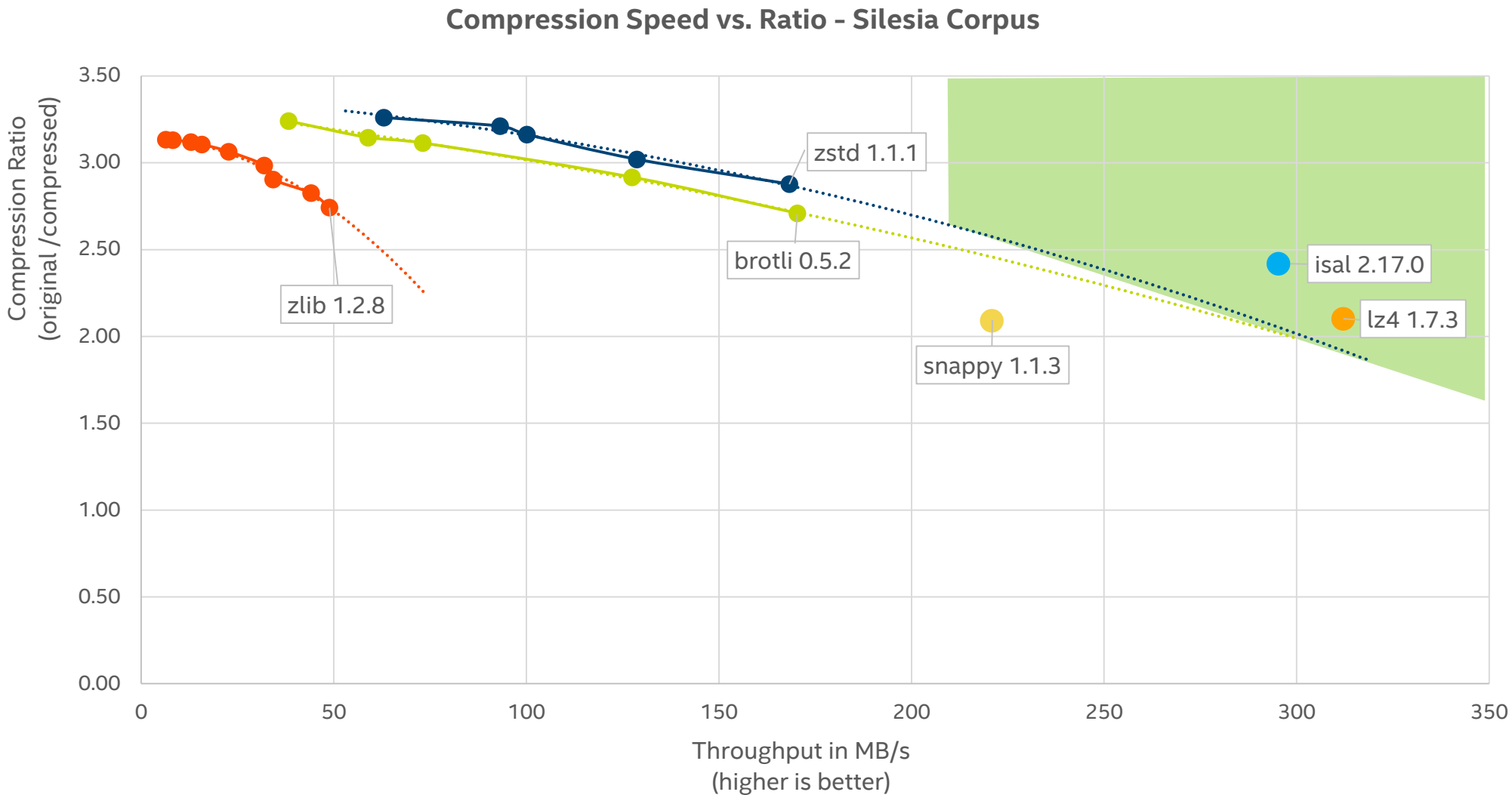
Compressor name	Decompression Throughput (MB/s)	Ratio
brotli 0.4.0 -0	318	36.91
isal 2.17.0	643	41.35
lz4 r131	2484	47.59
snappy 1.1.3	1335	47.83
zlib 1.2.8 -1	289	36.45
lzo1b 2.09 -1	671	45.78

Preliminary Results generated with lzbench v1.2 via https://github.com/inikep/lzbench
using Silesia Corpus on Intel® Xeon® E5-2650v3 (Haswell) CPU , turbo boost enabled, 16GB DDR4

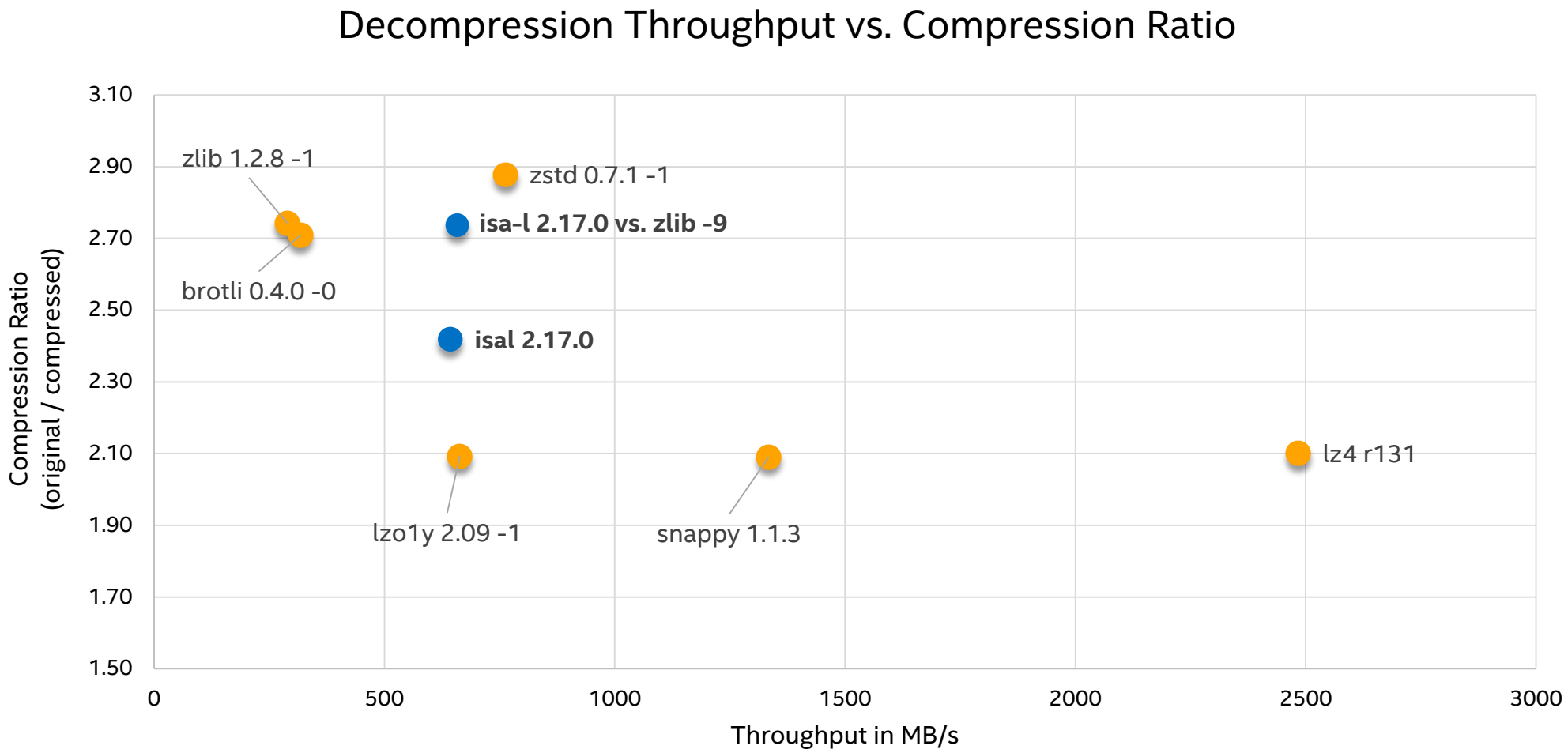
Algorithms Comparison: **Compression**



Algorithm Comparison: zlib, zstd, brotli...



Algorithms Comparison: **Decompression**



Cycle/Byte Performance on the Intel® Xeon® Processor Scalable Family

(cache cold cycle/byte)

Intel® Xeon® Platinum 8180 Processor @ 2.5 GHz 1 Socket				
ISA-L Function	ISA-L			
	Corpus	Cycle/Byte Performance (lower is better)	Single Core Throughput (higher is better)	Compression Ratio
Compress – Stateless Level 0	Calgary	8.21	304 MB/s	40.52
Compress – Stateless level 0	Silesia	7.03	355 MB/s	41.35
Compress – Stateless Level 1	Calgary	8.54	292 MB/s	37.51
Compress – Stateless Level 1	Silesia	7.25	344 MB/s	36.86
Compress – Stateless Level 2	-	-	-	-
Compress – Stateless Level 3	-	-	-	-
Decompress "Inflate"	Calgary	6.23	400 MB/s	N/A
Decompress "Inflate"	Silesia	5.30	471 MB/s	N/A

Cycle/Byte Performance (lower is better)	Single CoreThroughput (higher is better)
zlib 1.2.11 - Deflate	
51.80 CC WT AVE ratio 39.24%	48 MB/s
49.29 Silesia WT AVE ratio 38.33%	50 MB/s
zlib 1.2.11 - Inflate	
12.69 CC WT AVE	197 MB/s
12.25 Silesia WT AVE	204 MB/s

** ISA-L function uses AVX512 instructions

Case Study: Genome Analysis Tool Kit (GATK)

Genomics Data

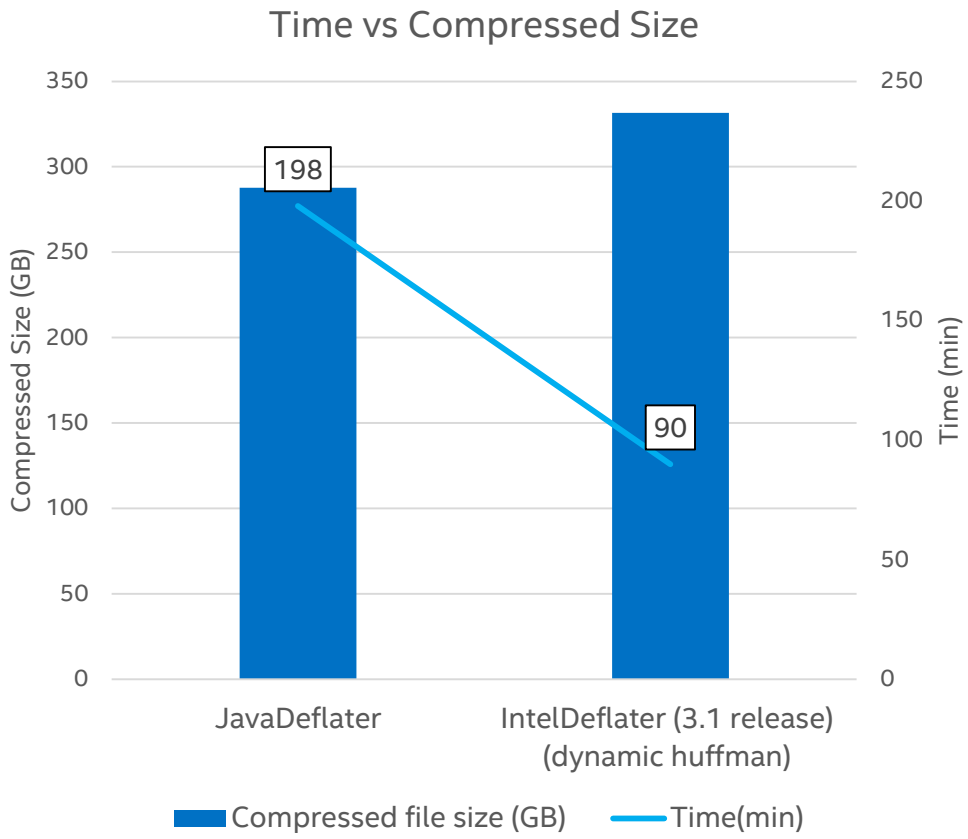
- Huge volumes of data: 100s of GB is minimum
- Cancer Cloud: to come up with a cure, for each given cancer type, need 100k – 1M patients with 2TB of genomics per patient = 2 Exabytes per cancer

Economics

- Handling that volume is hard!
- DEFLATE does a great job on sequenced genomes
- Industry tools reliant on zlib, usually Java

GATK Integration

- Handling that volume is hard!
- DEFLATE does a great job on sequenced genomes
- Industry tools reliant on zlib, usually Java



Libz Replacement

- Goal

  After installing IGZIP, applications which use ZLIB for compression and decompression can use IGZIP for better performance without modification of your code.

- Status

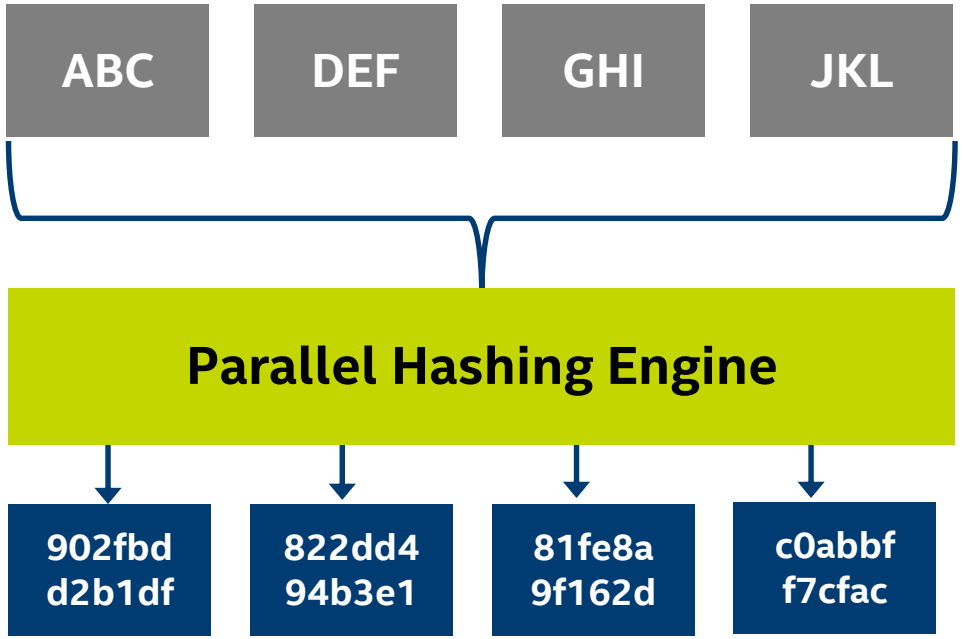
  Under development.

CRYPTOGRAPHIC HASHING

Multibuffer Hash

Vectorized hashes!

- Uses Intel® Advanced Vector Extensions 512 (Intel® AVX-512)
- MD5, SHA1, SHA2-256, SHA2-512
- Asynchronous interface
- “Four for one”



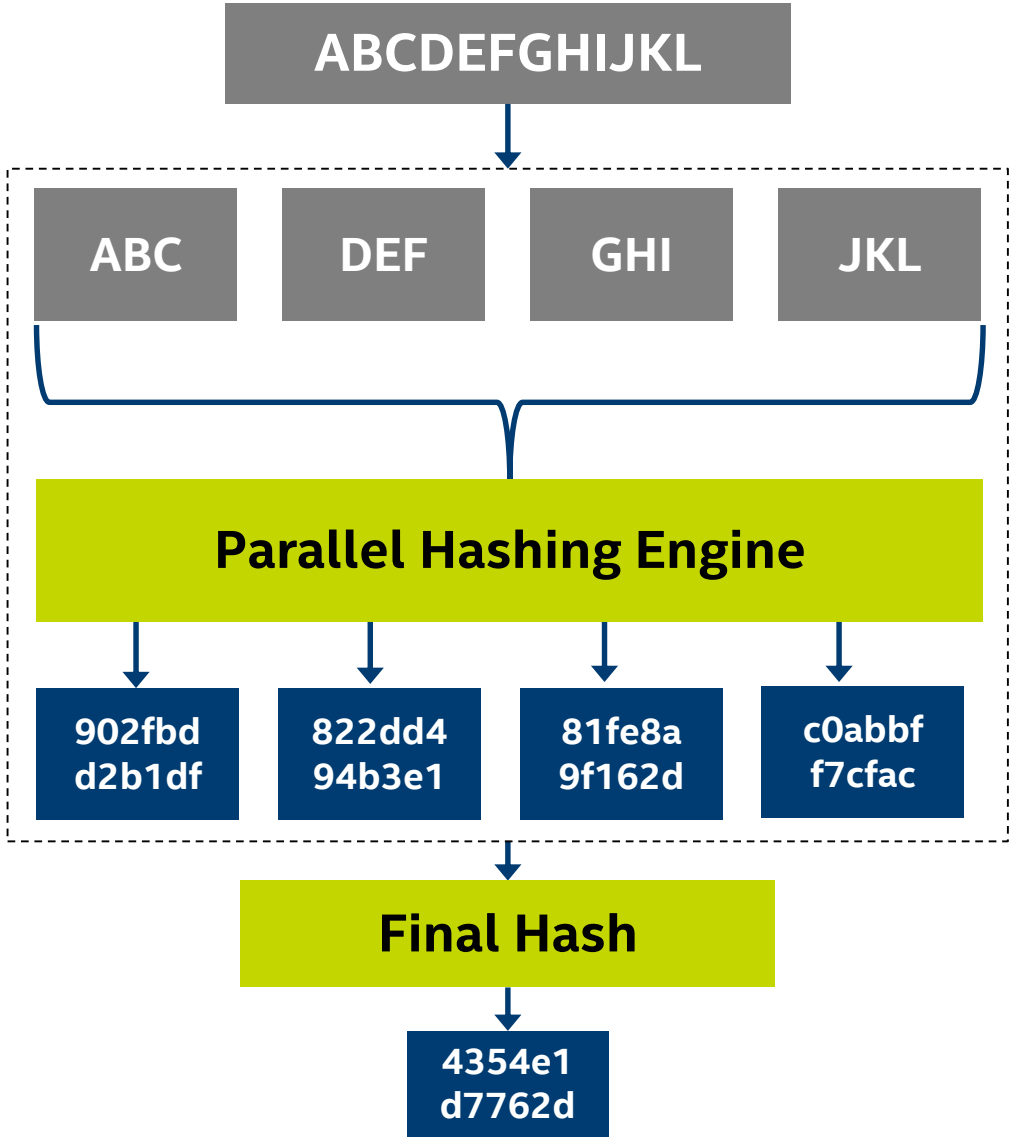
Multihash

What is ISA-L Multihash?

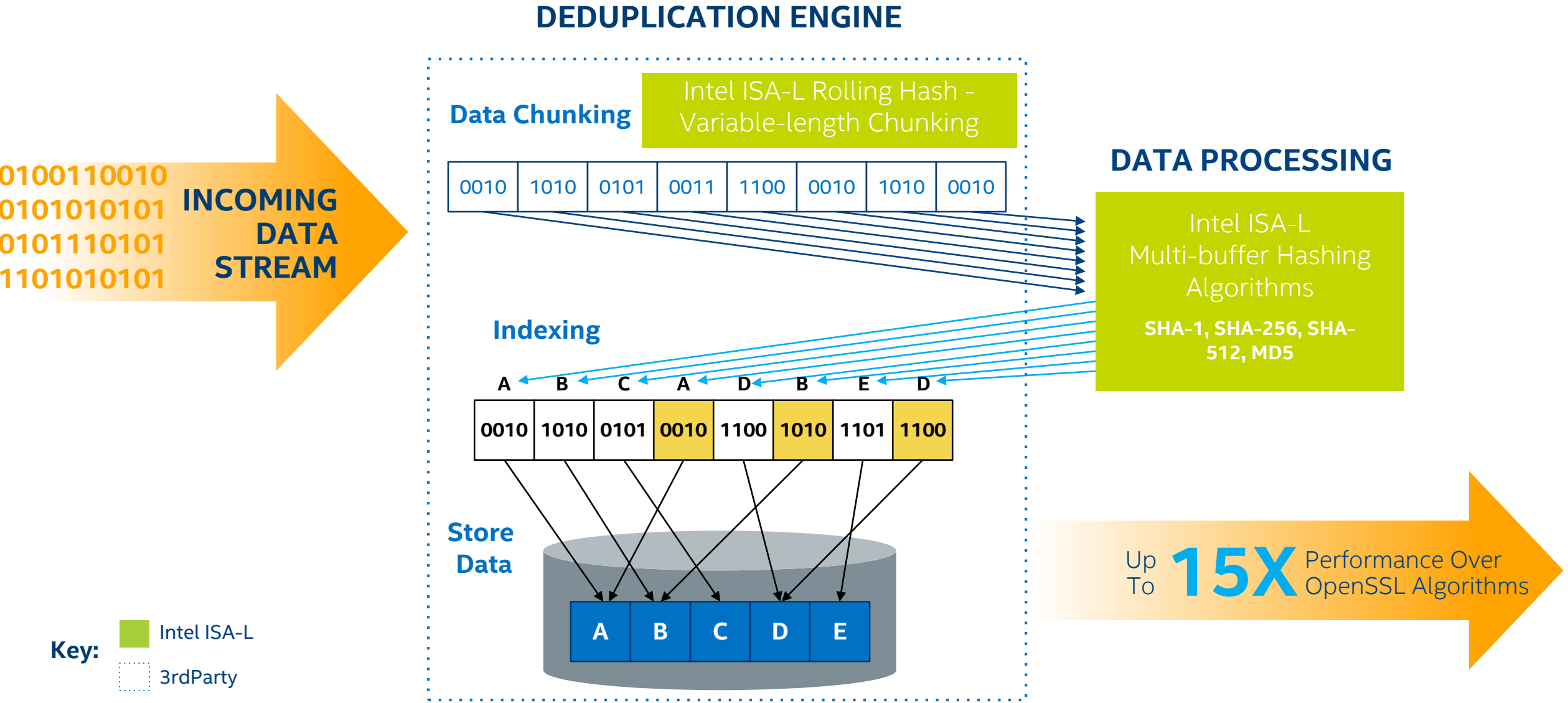
- Synchronous interface
- SHA1 != SHA1

Use Cases

- Data integrity
- Encryption
- Deduplication



Hashing Usage: Data Deduplication Optimizations



Cycle/Byte Performance on the Intel® Xeon® Processor Scalable Family

(cache cold cycle/byte)

ISA-L Function	Intel® Xeon® Platinum 8180 Processor @ 2.5 GHz 1 Socket			
	ISA-L		OpenSSL 1.0.2j	
	Cycle/Byte Performance (lower is better)	Single Core Throughput (higher is better)	Cycle/Byte Performance (lower is better)	Single Core Throughput (higher is better)
Rolling Hash 32 bit	4.12	606 MB/s	-	-
Rolling Hash 64 bit	2.53	988 MB/s	-	-
Multihash SHA-1**	0.42	5.8 GB/s	-	-
Multihash SHA-1 Murmur**	0.63	3.8 GB/s	-	-
Multihash SHA-256**	0.82	2.9 GB/s	-	-
Multibuffer SHA-1**	0.45	5.4 GB/s	4.13	605 MB/s
Multibuffer SHA-256**	0.88	2.7 GB/s	11.56	216 MB/s
Multibuffer SHA-512**	1.08	2.2 GB/s	7.48	334 MB/s
Multibuffer MD5**	0.25	9.7 GB/s	4.99	401 MB/s

Up to 9X bandwidth boost

Up to 13X bandwidth boost

Up to 7X bandwidth boost

Up to 15X bandwidth boost

** ISA-L function uses AVX512 instructions

Multi-buffer Hash with SHA-NI

- **What is SHA_NI:**

The Intel® SHA Extensions are a family of seven Streaming SIMD Extensions (SSE) based instructions that are used together to accelerate the performance of processing SHA-1 and SHA-256 on Intel®Architecture processors.

- **Update:**
  - sha1_ni_x1 -> sha1_ni_x2
  - Sha256_ni_x1 -> sha256_ni_x2

- **Supported platform:** Denverton®, Skylake-Xeon® , to be continued…

- **Status:** Recently released and no official performance data yet.

Instruction	Op 1	Op 2	Op 3	Opcode
SHA1 New Instructions				
SHA1RNDS4	xmm (rw)	xmm/m128 (r)	imm8	0F 3A CC /r ib
SHA1NEXTE	xmm (rw)	xmm/m128 (r)	NA	0F 38 C8 /r
SHA1MSG1	xmm (rw)	xmm/m128 (r)	NA	0F 38 C9 /r
SHA1MSG2	xmm (rw)	xmm/m128 (r)	NA	0F 38 CA /r
SHA256 New Instructions				
SHA256RNDS2	xmm (rw)	xmm/m128 (r)	<xmm0> (implicit)	0F 38 CB /r
SHA256MSG1	xmm (rw)	xmm/m128 (r)	NA	0F 38 CC /r
SHA256MSG2	xmm (rw)	xmm/m128 (r)	NA	0F 38 CD /r

Table 1: Intel® SHA Extensions Definitions (rw – Read/Write, r – Read Only)

STATE OF ISA-L

Community of ISA-L

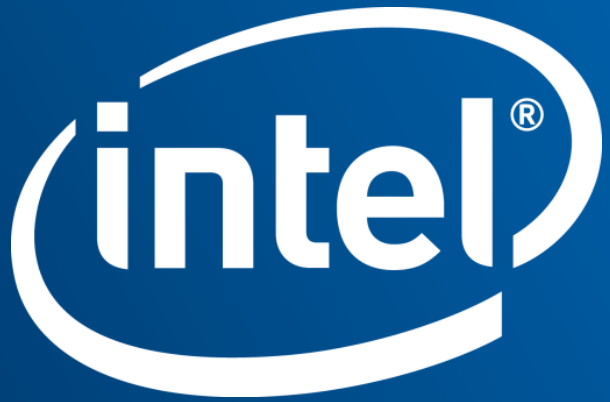
- **Open Source Projects at Github:**
  - https://github.com/01org/isa-l
  - https://github.com/01org/isa-1_crypto
- **License:** Algorithms are available under BSD license.
- **To use it:** see the included Getting Started Guide, API Guide, and C language reference applications
- **Feedback & Notification**
  - Github Issues Report
  - Subscribe to ISA-L: https://lists.01.org/mailman/listinfo/isal

SUMMARIZE

❑ Features Update & Usercase Sharing

- Encryption
- Stitching – CRC & copy
- CRC64
- Compression – Level 3 supported
- Multi-buffer Hash – sha_ni

❑ State of the Project


intel

PERFORMANCE TABLE

Cycle/Byte Performance on the Intel® Xeon® Processor Scalable Family

(cache cold cycle/byte)

ISA-L Function	Intel® Xeon® Platinum 8180 Processor @ 2.5 GHz 1 Socket			
	ISA-L		OpenSSL 1.0.2j	
	Cycle/Byte Performance (lower is better)	Single Core Throughput (higher is better)	Cycle/Byte Performance (lower is better)	Single Core Throughput (higher is better)
Rolling Hash 32 bit	4.12	606 MB/s	-	-
Rolling Hash 64 bit	2.53	988 MB/s	-	-
Multihash SHA-1**	0.42	5.8 GB/s	-	-
Multihash SHA-1 Murmur**	0.63	3.8 GB/s	-	-
Multihash SHA-256**	0.82	2.9 GB/s	-	-
Multibuffer SHA-1**	0.45	5.4 GB/s	4.13	605 MB/s
Multibuffer SHA-256**	0.88	2.7 GB/s	11.56	216 MB/s
Multibuffer SHA-512**	1.08	2.2 GB/s	7.48	334 MB/s
Multibuffer MD5**	0.25	9.7 GB/s	4.99	401 MB/s
AES-XTS 128	0.64	3.8 GB/s	0.64	3.8 GB/s
AES-XTS 256	0.88	2.7 GB/s	0.89	2.7 GB/s
AES-CBC 128 Decode	0.64	3.8 GB/s	0.64	3.8 GB/s
AES-CBC 192 Decode	0.76	3.2 GB/s	0.76	3.2 GB/s
AES-CBC 256 Decode	0.88	2.7 GB/s	0.88	2.7 GB/s
AES-GCM 128	0.67	3.6 GB/s	1.58	1.5 GB/s
AES-GCM 256	0.89	2.7 GB/s	1.83	1.3 GB/s

Up to 13X bandwidth boost

Up to 15X bandwidth boost

** ISA-L function uses AVX512 instructions

Cycle/Byte Performance on the Intel® Xeon® Processor Scalable Family

(cache cold cycle/byte)

ISA-L Function	Intel® Xeon® Platinum 8180 Processor @ 2.5 GHz 1 Socket			
	ISA-L		OpenSSL 1.0.2j	
	Cycle/Byte Performance (lower is better)	Single Core Throughput (higher is better)	Cycle/Byte Performance (lower is better)	Single Core Throughput (higher is better)
PQ Gen (16+2)**	0.10	23.4 GB/s	-	-
XOR Gen (16+1)**	0.10	24.2 GB/s	-	-
Reed Solomon EC (10+4)**	0.19	12.7 GB/s	-	-
CRC16 T10	0.22	11.0 GB/s	-	-
CRC32 IEEE (802.3)	0.22	11.1 GB/s	-	-
CRC32 iSCSI	0.18	13.8 GB/s	-	-
CRC32 GZIP Reflective	0.24	10.3 GB/s	-	-
CRC64 Normal	0.22	11.0 GB/s		
CRC64 Reflective	0.19	12.9 GB/s		
Compress - Stateless	*Level 0* 8.21 CC WT AVE ratio 40.52 7.03 Silesia WT AVE ratio 41.35 *Level 1* 8.54 CC WT AVE ratio 37.51 7.25 Silesia WT AVE ratio 36.86	*Level 0* 304 MB/s 355 MB/s *Level 1* 292 MB/s 344 MB/s		
Decompress "Inflate"	6.23 CC WT AVE 5.30 Silesia WT AVE	400 MB/s 471 MB/s		

Cycle/Byte Performance (lower is better)	Single CoreThroughput (higher is better)
zlib 1.2.11 - Deflate	
51.80 CC WT AVE ratio 39.24% 49.29 Silesia WT AVE ratio 38.33%	48 MB/s 50 MB/s
zlib 1.2.11 - Inflate	
12.69 CC WT AVE 12.25 Silesia WT AVE	197 MB/s 204 MB/s

** ISA-L function uses AVX512 instructions

Cycle/Byte Performance on the Intel® Xeon® Processor Scalable Family

(cache cold cycle/byte)

ISA-L Function	Intel® Xeon® Platinum 8180 Processor @ 2.5 GHz 1 Socket			
	ISA-L		zlib 1.2.11	
	Cycle/Byte Performance (lower is better)	Single Core Throughput (higher is better)		
Compress - Stateless	*Level 0* 8.21 CC WT AVE ratio 40.52 7.03 Silesia WT AVE ratio 41.35 *Level 1* 8.54 CC WT AVE ratio 37.51 7.25 Silesia WT AVE ratio 36.86	*Level 0* 304 MB/s 355 MB/s *Level 1* 292 MB/s 344 MB/s	51.80 CC WT AVE ratio 39.24% 49.29 Silesia WT AVE ratio 38.33%	
Decompress "Inflate"	6.23 CC WT AVE 5.30 Silesia WT AVE	400 MB/s 471 MB/s		

Cycle/Byte Performance (lower is better)	Single CoreThroughput (higher is better)
zlib 1.2.11 - Deflate	
51.80 CC WT AVE ratio 39.24% 49.29 Silesia WT AVE ratio 38.33%	48 MB/s 50 MB/s
zlib 1.2.11 - Inflate	
12.69 CC WT AVE 12.25 Silesia WT AVE	197 MB/s 204 MB/s

** ISA-L function uses AVX512 instructions