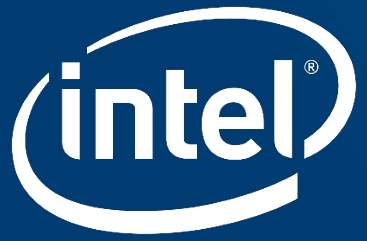


storage acceleration with ISA-L

Greg Tucker, Intel

Notices and Disclaimers

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Learn more at intel.com, or from the OEM or retailer.

No computer system can be absolutely secure.

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/performance**.

Motivation

- I/O machines -> storage processors.
- Trend to software-defined storage
- Disk speeds finally increasing
- Common set of algorithms are pervasive

ISA-L - Intelligent Storage Acceleration Library

ISA-L is a collection of optimized low-level functions targeting storage applications.

- Mostly ASM-optimized functions
- Open Source - BSD Licensed
- Portable to Linux, Windows, FreeBSD

When do you use ASM?

Almost never

When large performance advantage

When you can embed multiple version

When can span your deployments

When interface/algorithms are not brittle

Future proof from users

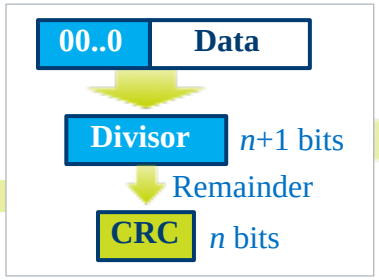
ISA-L - Intelligent Storage Acceleration Library

Includes algorithms that satisfy:

- **High CPB**: Target only the highest cycle-per-byte functions in modern storage systems.
- **Pain points**: Include only core storage algorithms where throughput and latency are the most critical factors.
- **Can Optimize**: Look for cases where hand-optimized asm or specific structural changes can have a big advantage.

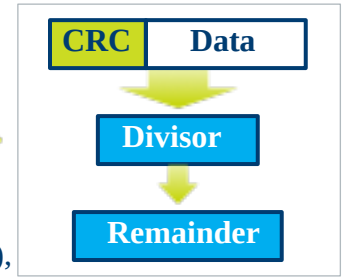
Intel® ISA-L **Functions**

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

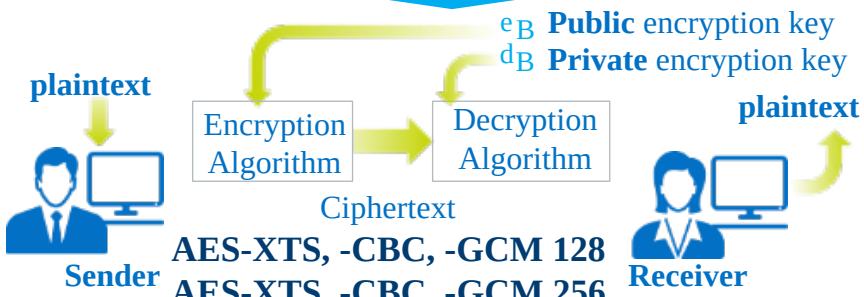
06d80e7
b0C50bs
49a509t
b49f249
24e8c8o
05x84q4

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

COMPRESSION "DEFLATE"

IGZIP: Faster DEFLATE (zlib)
Compression & Decompression

ENCRYPTION



Sender

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

Receiver

Where is ISA-L used?

Open Source Projects

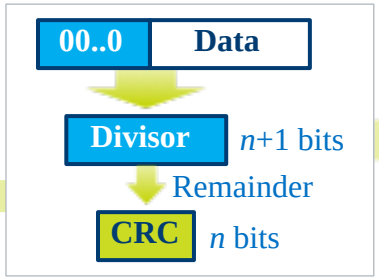
- Scale-out storage (HDFS*, Ceph* & Swift*)
- Streaming encryption
- Deduplication software
- File systems

Proprietary Projects

- Hyperscale object storage
- Deduplication & backup solutions
- Multi-cloud backup
- Low-latency scale-up appliances

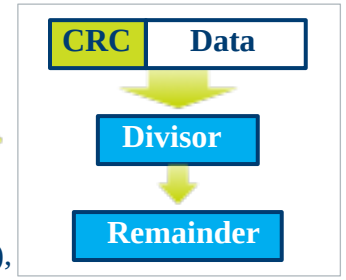
Intel® ISA-L **Functions: Compression**

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

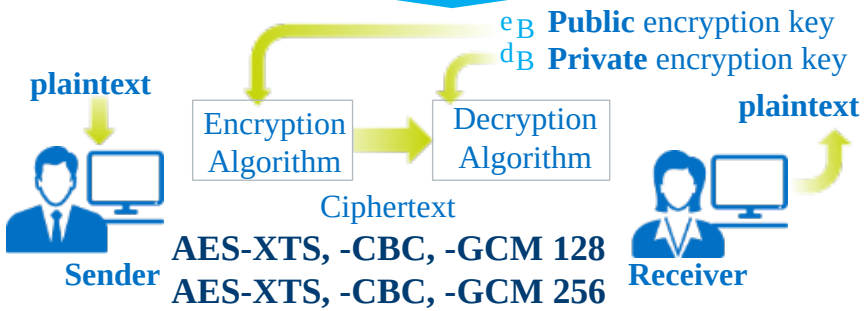
06d80e7
b0C50bs
49a509t
b49f249
24e8c8o
05x84q4

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

COMPRESSION "DEFLATE"

IGZIP: Faster DEFLATE (zlib)
Compression & Decompression

ENCRYPTION



Sender

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

Receiver

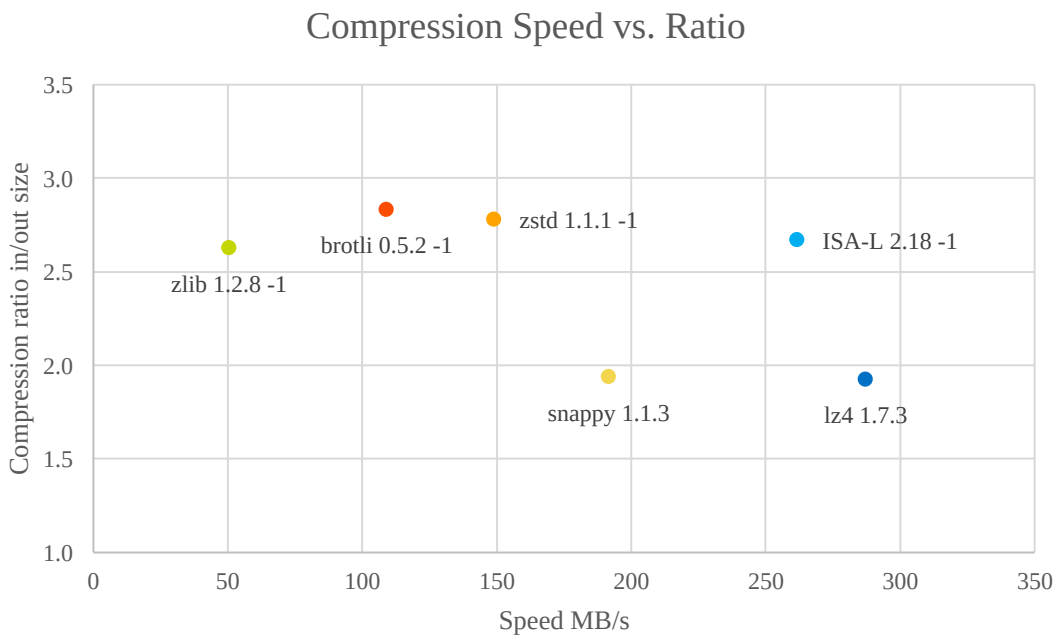
ISA-L Fast Deflate

Deflate (aka zlib, gzip, pkzip, etc)
- Lossless compression
- Ubiquitous adoption

v2.18: ISA-L Level-1 Deflate
- **5X greater throughput** than zlib -1
- **13% better compression ratio** than lz4 and lzo

v2.17: Optimized Decompression
- **>2X throughput** vs. zlib, equal to lzo
- **Fully compatible** with zlib and gzip archives

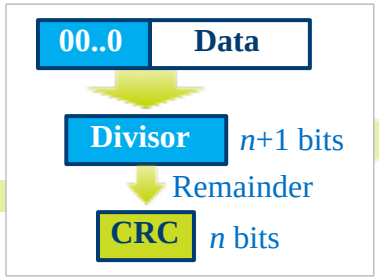
Compressor Name	Compression Throughput (MB/s)	Ratio
lz4 1.7.3	287.1	52.0%
ISA-L 2.18 -1	**261.6**	**37.5%**
snappy 1.1.3	191.6	51.6%
zstd 1.1.1 -1	149.0	36.0%
brotli 0.5.2 -1	109.0	35.3%
zlib 1.2.8 -1	50.5	38.1%



Hardware Configuration: Aztec City CRB, 2x Intel® Xeon® E5-2650v4, 4x 8GB DDR4 2400 MT/s, BIOS GRRFCRB1.86B.0276.R02.1606020546
BIOS configuration: Hyperthreading: disabled; Turbo Boost: disabled; Speed Step: disabled; P- and C-states: disabled. **Calgary Corpus, single core throughput.**

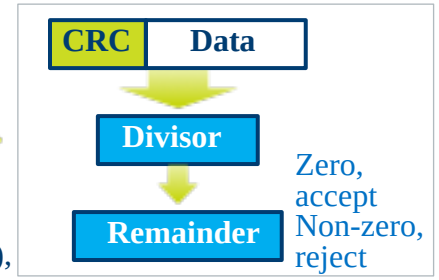
Intel® ISA-L **Functions: Hashing**

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

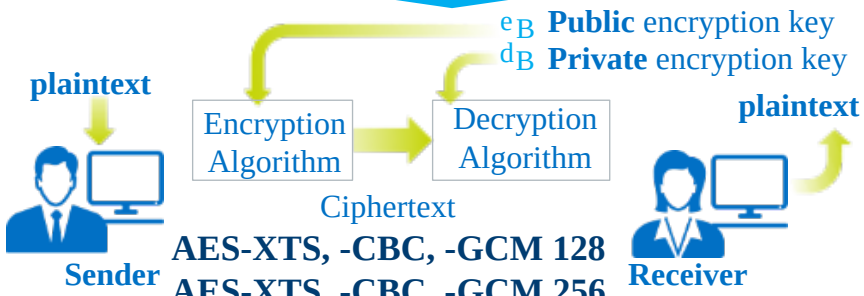
CRYPTOGRAPHIC HASHING

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

COMPRESSION "DEFLATE"

IGZIP: Faster DEFLATE (zlib)
Compression & Decompression

ENCRYPTION



Sender

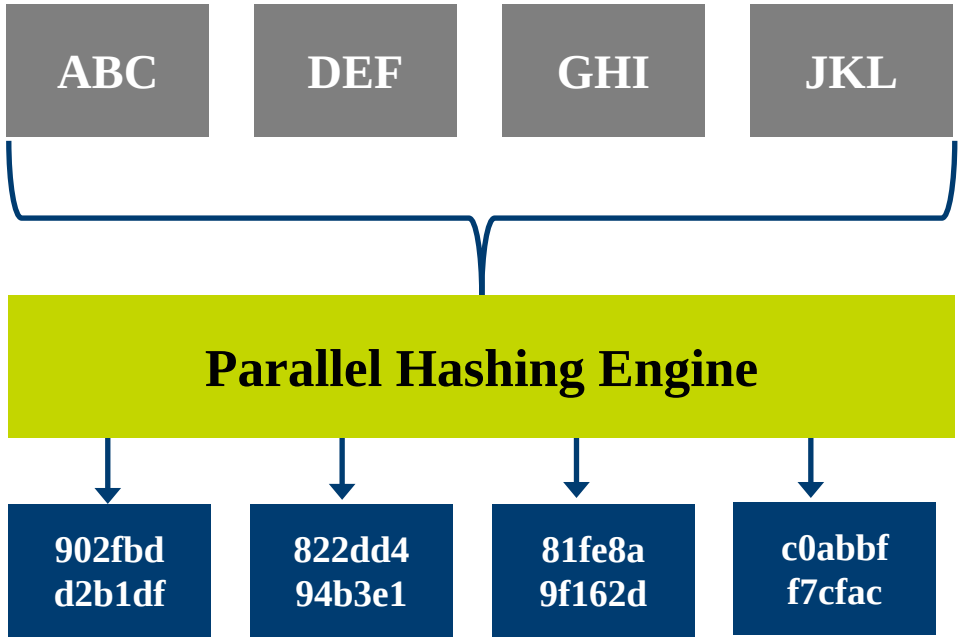
AES-XTS, -CBC, -GCM 128
AES-XTS, -CBC, -GCM 256

Receiver

Multi-buffer Hashing

Vectorized cryptographic Hashes

- Uses SSE, AVX, AVX2 or AVX512
- MD5, SHA1, SHA2-256, SHA2-512
- Asynchronous interface
- 4-32 at a time for greater throughput



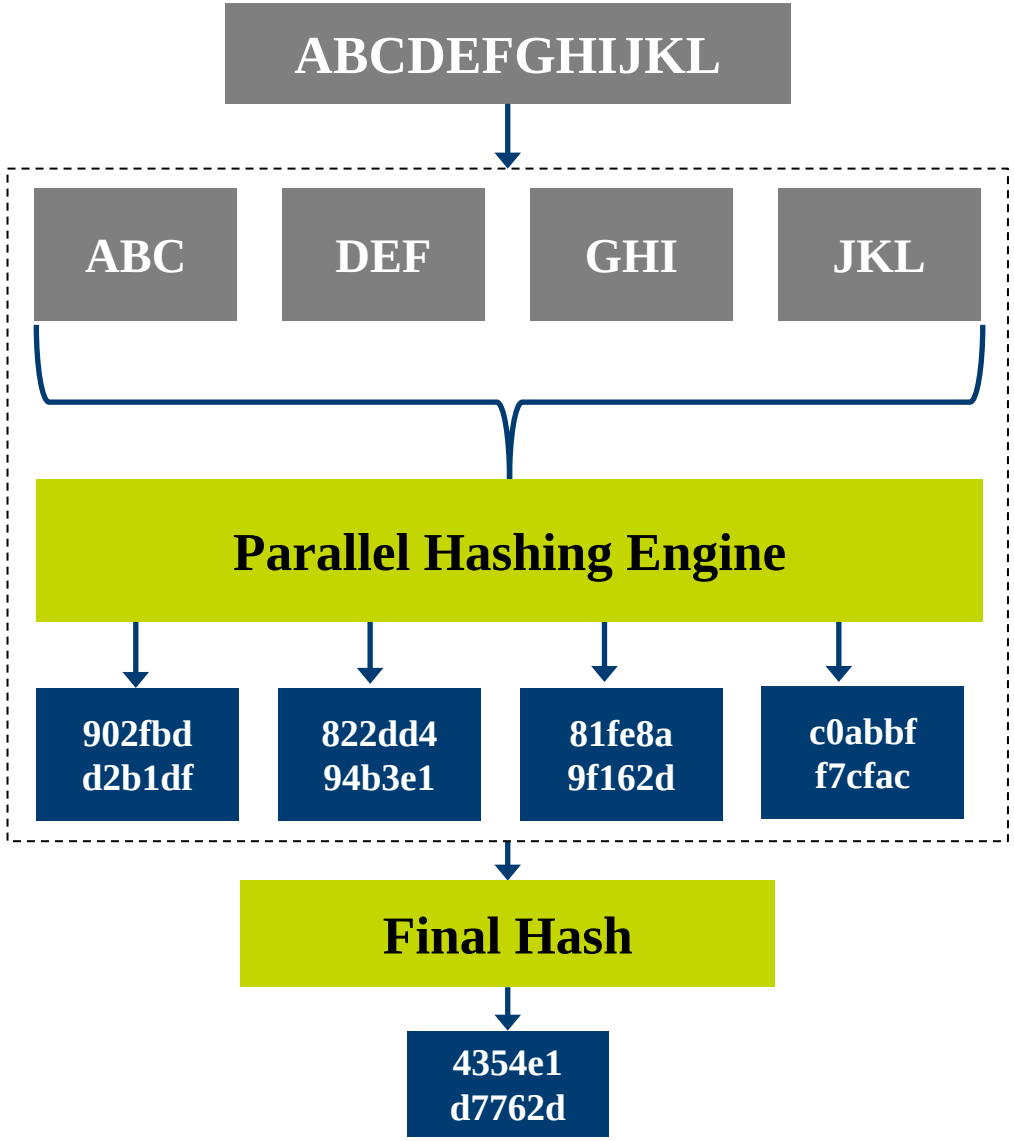
Multihash

What is ISA-L Multihash?

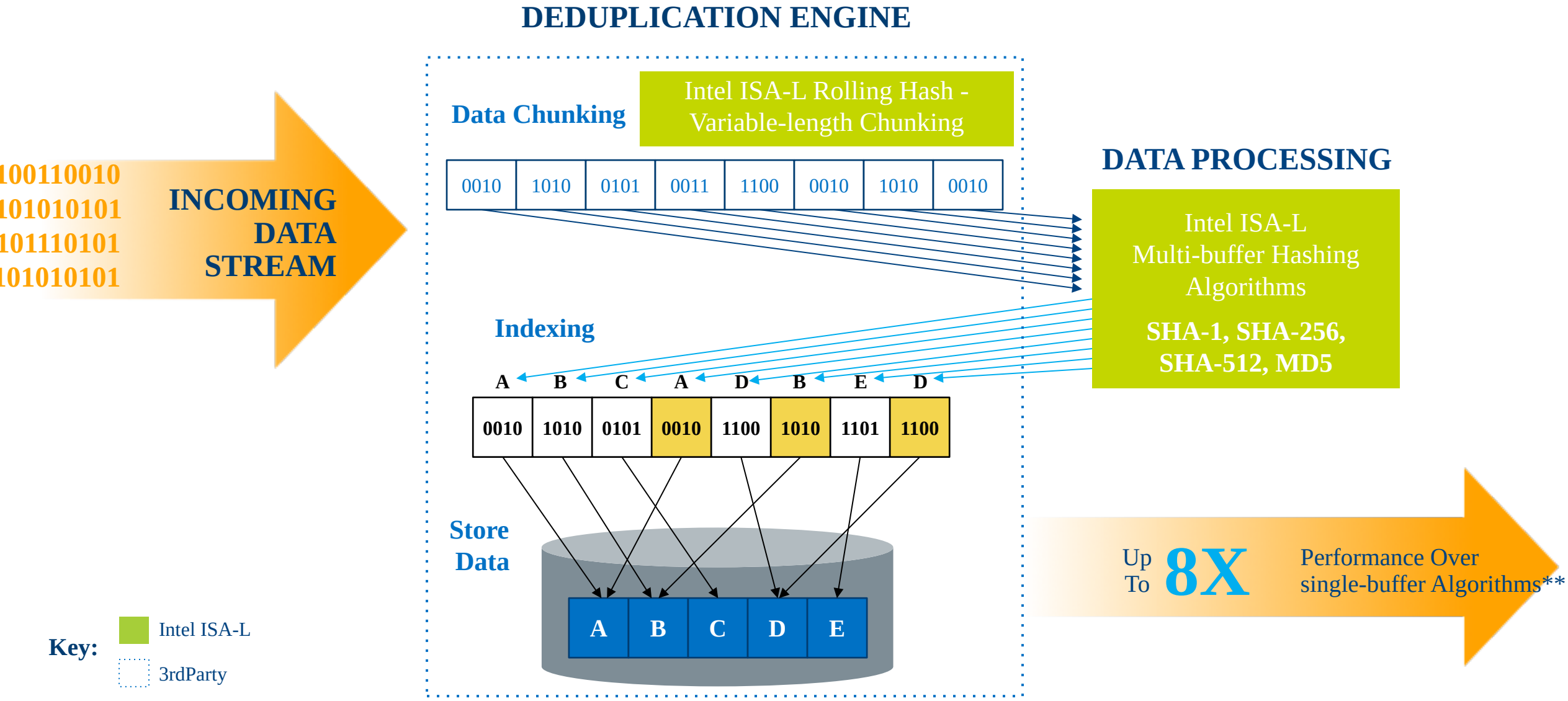
- Synchronous interface
- SHA1 != SHA1

Use Cases

- Data integrity
- Encryption
- Deduplication



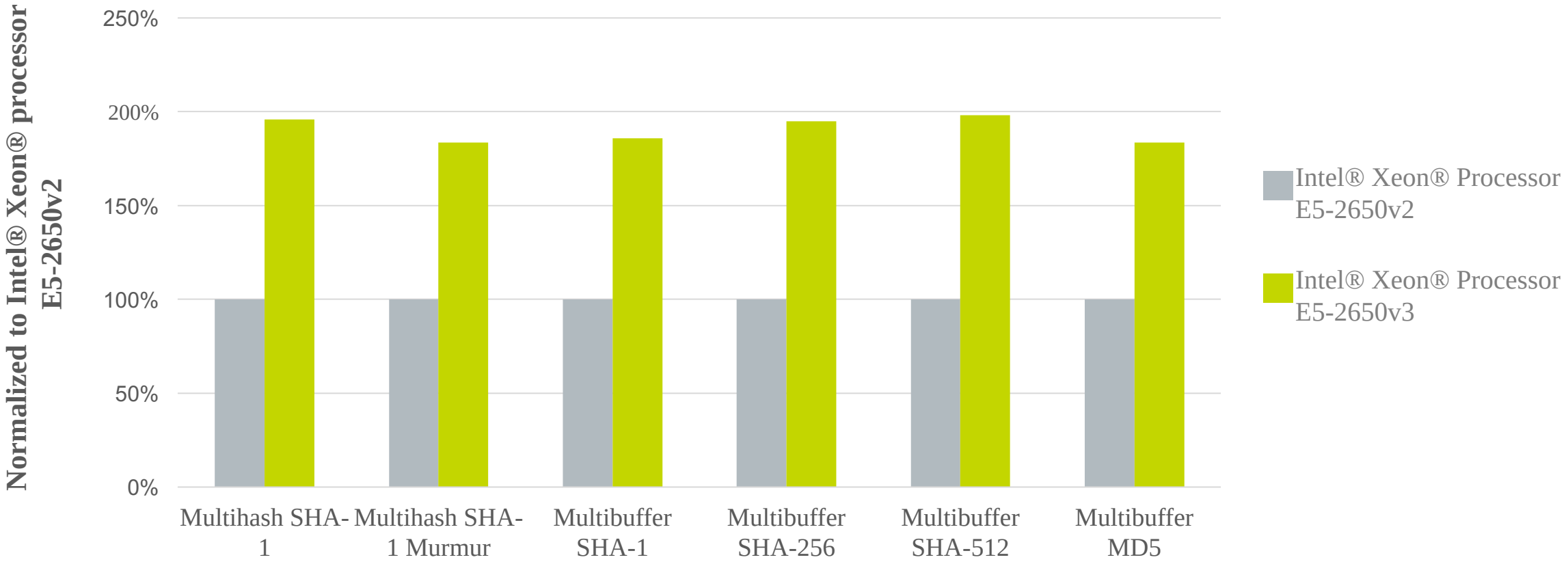
Hashing Usage: Data Deduplication Optimizations



Performance on the **Intel® Xeon® Processor**

Generational Cycle/Byte Comparison

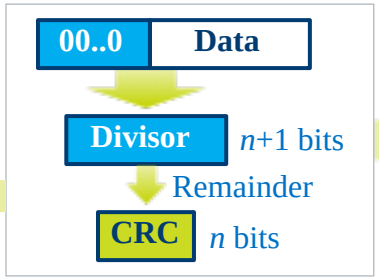
(higher is better)



E5-2560v2 Configuration: Rose City CRB, 2x Intel® Xeon® E5-2650v2, 4x 8GB DDR3 1600 MHz ECC RDIMM
E5-2650v3 Configuration: Aztec City CRB, 2x Intel® Xeon® E5-2650v2, 4x 8GB DDR4 2133 MHz ECC RDIMM
BIOS configuration: Hyperthreading: disabled; Turbo Boost: disabled; Speed Step: disabled; P- and C-states: disabled.

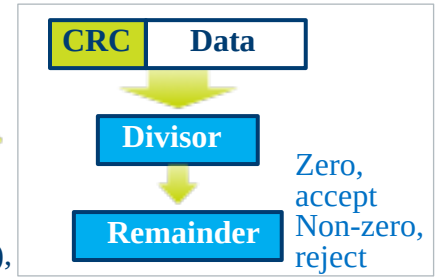
Intel® ISA-L **Functions: Erasure Coding**

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



Sender

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



Receiver

DATA PROTECTION

DATA INTEGRITY

PERFORMANCE OPTIMIZING

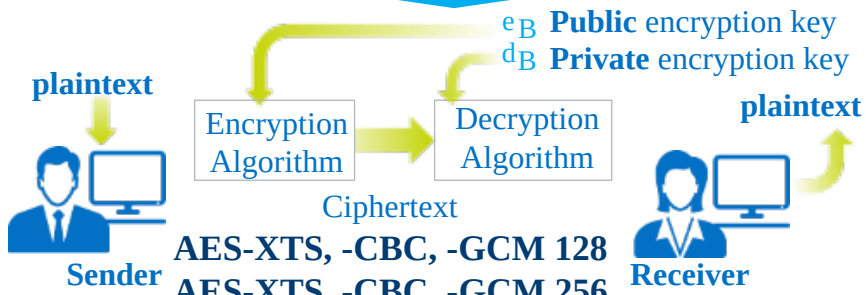
CRYPTOGRAPHIC HASHING

Multi-Buffer: SHA-1, SHA-256, SHA-512, MD5

COMPRESSION "DEFLATE"

IGZIP: Fast DEFLATE (zlib)

ENCRYPTION



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

ISA-L: Erasure Codes

Who is using Erasure Codes?

- "All the clouds" – distributed storage frameworks
- Hadoop HDFS, Ceph, Swift, hyperscalers...

Why are they using Erasure Codes?

- Irresistible economics:
  - (at least) as much redundancy as triple replication with half the raw data footprint
- Half the storage media costs = big capex and opex savings

Why wasn't everyone using them before?

- Previously RS-EC was computationally prohibitive
- E5-2600v4, ISA-L can generate ~5GB/s of EC!

ISA-L General Reed-Solomon Erasure Codes

General: Any Reed-Solomon block erasure code in GF(2^8) (m,k)

Performance does not depend on elements in encoding/decoding matrix

Optimal in the sense of highest theoretical recovery potential (MDS)

Flexible: Can be symmetric or not

Fast: High speed makes irrelevant any R-S replacements

$$\begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \end{vmatrix} \cdot \begin{vmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \end{vmatrix} = \begin{vmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \\ p_1 \\ p_2 \\ p_3 \end{vmatrix}$$

$$\mathbf{G \cdot d = c}$$

ISA-L: Get

- Open source available on GitHub
  https://github.com/01org/isa-l
  https://github.com/01org/isa-1_crypto
- Standard distros
  - FreeBSD ports - (http://www.freshports.org/devel/isa-l/)
  - Clear Linux
  - Debian – (sid/stretch/stable-backports)
  - Ubuntu - (yakkety/zesty)

backup

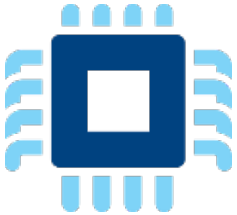
Performance Metrics

Intel® ISA-L Performance **Overview**

Functional Library Comparisons

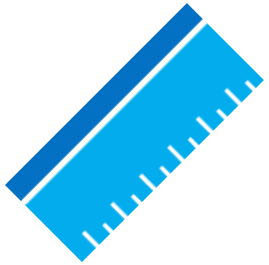
(performance vs. other libraries available)

- ISA-L 2.17
- OpenSSL 1.0.2g
- zlib 1.2.8



CPU Gen over Gen Performance

- Intel® Xeon® processor generation over generation performance metrics



Units of Measurement

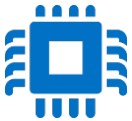
- Cycles/Byte
- Throughput (MB/s, GB/s)
- Calgary Corpus Weighted Ave
- Compression Ratio

Intel® ISA-L Performance **Overview**

Platform configuration details

Intel® Xeon® Processor E5-2600v4

- E5-2650v4, 12C, 2.2 GHz, M0
- Aztec City CRB
- 4x8 GB DDR4 2400 MT/s ECC RDIMM



BIOS Configuration

- P-States: Disabled
- Turbo: Disabled
- Speed Step: Disabled
- C-States: Disabled
- Power Performance Tuning: Disabled
- ENERGY_PERF_BIAS_CFG: PERF
- Isochronous: Disabled
- Memory Power Savings: Disabled

Cold Cache Tests

- Pick large data set by default (larger than last-level cache)
- Ensures memory fetch/put included

Turbo Off for Repeatability

Loop to Reduce Timer Latencies and Transients

- Start timer
- Iterate over data set
- Stop timer
- Report total bytes processed/time

Cycle/Byte Performance on the **Intel® Xeon® Processor E5-2600v4 Product Family** (cache cold cycle/byte)

ISA-L Function	Intel® Xeon® Processor E5-2650v4 @ 2.1 GHz 1 Socket			
	ISA-L		OpenSSL 1.0.2g	
	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	4.16	529 MB/s	-	-
Rolling Hash 64 bit	2.67	823 MB/s	-	-
Multihash SHA-1	1.09	2.0 GB/s	-	-
Multihash SHA-1 Murmur	1.36	1.6 GB/s	-	-
Multibuffer SHA-1	1.14	1.9 GB/s	4.22	521 MB/s
Multibuffer SHA-256	2.62	840 MB/s	12.44	177 MB/s
Multibuffer SHA-512	3.26	676 MB/s	7.95	277 MB/s
Multibuffer MD5	0.61	3.5 GB/s	4.96	443 MB/s
AES-XTS 128	0.72	3.0 GB/s	0.86	2.5 GB/s
AES-XTS 256	0.93	2.3 GB/s	1.15	1.9 GB/s
AES-CBC 128 Decode	0.65	3.3 GB/s	0.81	2.7 GB/s
AES-CBC 192 Decode	0.76	2.8 GB/s	0.93	2.3 GB/s
AES-CBC 256 Decode	0.89	2.4 GB/s	1.06	2.0 GB/s
AES-GCM 128	0.80	2.7 GB/s	1.97	1.1 GB/s
AES-GCM 256	1.05	2.1 GB/s	2.26	973 MB/s

Up to **5X** bandwidth boost

Up to **8X** bandwidth boost

Cycle/Byte Performance on the Intel® Xeon® Processor E5-2600v4 Product Family (cache cold cycle/byte)

ISA-L Function	Intel® Xeon® Processor E5-2650v4 @ 2.1 GHz 1 Socket			
	ISA-L		OpenSSL 1.0.2g	
	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)	0.11	19.0 GB/s	-	-
XOR Gen (16+1)	0.10	21.5 GB/s	-	-
Reed Solomon EC (10+4)	0.41	5.3 GB/s	-	-
CRC T10	0.18	12.0 GB/s		
CRC IEEE (802.3)	0.18	12.0 GB/s		
CRC32 iSCSI	0.18	11.7 GB/s		
CRC64 Normal	0.18	12.0 GB/s		
CRC64 Reflective	0.18	12.0 GB/s		
Compress - Stateless	7.86 CC WT AVE ratio 40.52 6.75 Silesia WT AVE ratio 41.35	280 MB/s 325 MB/s		
Decompress "Inflate"	6.07 CC WT AVE 5.20 Silesia WT AVE	362 MB/s 422 MB/s		

	Cycle/Byte Performance (lower is better)	Single CoreThroughput (higher is better)
zlib 1.2.8 - Deflate		
	50.89 CC WT AVE ratio 39.24% 48.59 Silesia WT AVE ratio 38.33%	43 MB/s 45 MB/s
zlib 1.2.8 - Inflate		
	12.48 CC WT AVE 12.04 Silesia WT AVE	176 MB/s 182 MB/s