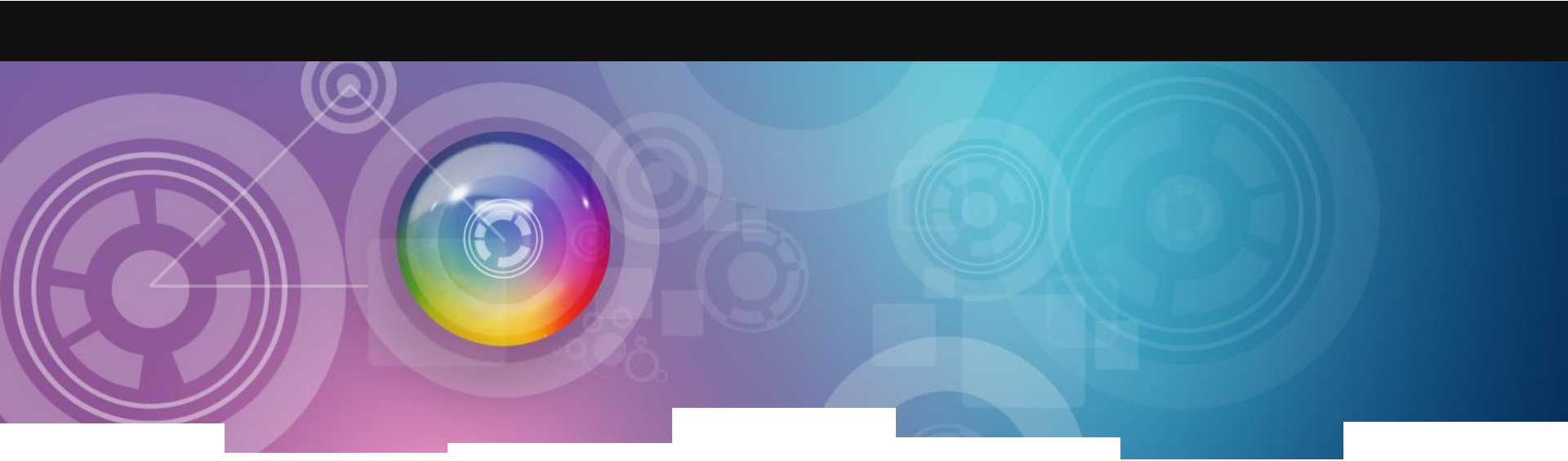


JVET-K0410 (Technicolor)

Comments on timing measurement variations for JVET experiments



Performed different assessments to show timing ratio variations:

- Anchor is VTM, Test is BMS (gives a timing ratio more than ~300%)
 - In practice, anchor is often close to Test, so the variation in timing ratio might be smaller.
- Anchor and Test are run in the same testing conditions
- Minimum of at least 4 runs is reported
- Study only decoder impact, as it is the most important for complexity assessment
- Vary some conditions:
 - some are difficult to impose (e.g. CPU generation),
 - some are easy to impose (e.g. compiler options)

Assessment 1 - Different runs in exactly same test configuration

Just re-run in the same conditions

→ ~0.5% variation

	VTM	BMS
Run 0	51.83	167.81
Run 1	51.87	167.35
Run 2	51.81	167.19
Run3	51.88	167.25

	description	Timing ratio variation
Assessment 1	Different runs in exactly same test configuration	0.5%
Assessment 2		
Assessment 3		
Assessment 4		
Assessment 5		
Assessment 6		
Assessment 7		
Assessment 8		
Assessment 9		
Assessment 10		

Assessment 2 - CPU generation

Change the CPU

→ ~24% variation

	VTM	BMS	ratio
SkyLake	53.8	170.2	316%
Broadwell	37.6	147.5	392%
Haswell	40.1	155.3	387%
IvyBridge	64.1	217.0	338%

	description	Timing ratio variation
Assessment 1	Different runs in exactly same test configuration	0.5%
Assessment 2	CPU generation	24%
Assessment 3		
Assessment 4		
Assessment 5		
Assessment 6		
Assessment 7		
Assessment 8		
Assessment 9		
Assessment 10		

Assessment 3 - Computer grid workload

Extreme load variation: 1 core busy or all cores busy

→ ~19% variation (using SSE42)

	VTM	BMS	Ratio
Minimum load	[37.5,37.6]	[147.5,147.9]	[393%,394%]
Maximum load	[52,53.1]	[175.7,179.3]	[331%,345%]

	description	Timing ratio variation
Assessment 1	Different runs in exactly same test configuration	0.5%
Assessment 2	CPU generation	24%
Assessment 3	Computer grid workload	19%
Assessment 4		
Assessment 5		
Assessment 6		
Assessment 7		
Assessment 8		
Assessment 9		
Assessment 10		

Assessment 4 - Computer grid turbo mode

Turbo ON and OFF. CPU power mode is the same, only turbo boost changed (intel pstate/no_turbo)

→ ~11% variation

	VTM	BMS	Ratio
Turbo off	[57,58.1]	[203.3,209.4]	[350%,367%]
Turbo on	[52,53.1]	[175.7,179.3]	[331%,345%]

	description	Timing ratio variation
Assessment 1	Different runs in exactly same test configuration	0.5%
Assessment 2	CPU generation	24%
Assessment 3	Computer grid workload	19%
Assessment 4	Turbo mode ON/OFF	11%
Assessment 5		
Assessment 6		
Assessment 7		
Assessment 8		
Assessment 9		
Assessment 10		

Assessment 5 - Compiler version

Change only the compiler version. Note that gcc 7.2 seems broken

→ ~21% variation

	VTM	BMS	ratio
gcc 4.8	65.4	267.6	409%
gcc 5.4	64.1	217.0	338%
gcc 6.3	62.9	213.5	339%
gcc 7.2	59.3	305.4	515%
gcc 8.1	58.8	240.5	409%

	description	Timing ratio variation
Assessment 1	Different runs in exactly same test configuration	0.5%
Assessment 2	CPU generation	24%
Assessment 3	Computer grid workload	19%
Assessment 4	Turbo mode ON/OFF	11%
Assessment 5	Compiler version	21%
Assessment 6		
Assessment 7		
Assessment 8		
Assessment 9		
Assessment 10		

Assessment 6 - Compiler settings in CMake files

Change the build environment: some missing options in the CMakeLists set the default optimization level, depending on the environment (OS, distribution etc.)

→ ~9% variation

	VTM	BMS	ratio
-O2 optimization	77.9	292.5	375%
-O3 optimization	65.4	267.6s	409%

	description	Timing ratio variation
Assessment 1	Different runs in exactly same test configuration	0.5%
Assessment 2	CPU generation	24%
Assessment 3	Computer grid workload	19%
Assessment 4	Turbo mode ON/OFF	11%
Assessment 5	Compiler version	21%
Assessment 6	Cmake settings/Compiler optimization	9%
Assessment 7		
Assessment 8		
Assessment 9		
Assessment 10		

Assessment 7 - SIMD level

Change the SIMD level on the command line

→ ~13~18% variation (depending on the turbo mode and load)

	VTM	BMS	Ratio
SCALAR	77.1	246.9	320%
SSE42	57	203.3	357%
AVX2	54.8	198.4	362%

	description	Timing ratio variation
Assessment 1	Different runs in exactly same test configuration	0.5%
Assessment 2	CPU generation	24%
Assessment 3	Computer grid workload	19%
Assessment 4	Turbo mode ON/OFF	11%
Assessment 5	Compiler version	21%
Assessment 6	Cmake settings/Compiler optimization	9%
Assessment 7	SIMD level	13~18%
Assessment 8		
Assessment 9		
Assessment 10		

Assessment 8 - Decoder options

Change the decoder options (even worse for small sequences)

→ ~27% variation

	VTM	BMS	ratio
MD5, output /dev/null	56.3	171.9	305%
MD5, no output	51.8	167.2	323%
No MD5, output /dev/null	44.3	160.0	362%
No MD5, no output	40.1	155.3	387%

	description	Timing ratio variation
Assessment 1	Different runs in exactly same test configuration	0.5%
Assessment 2	CPU generation	24%
Assessment 3	Computer grid workload	19%
Assessment 4	Turbo mode ON/OFF	11%
Assessment 5	Compiler version	21%
Assessment 6	Cmake settings/Compiler optimization	9%
Assessment 7	SIMD level	13~18%
Assessment 8	decoder software options	27%
Assessment 9		
Assessment 10		

Assessment 9 - Bitstream concatenation

Decode each chunk independently (cf JVET-B0036) or decode the whole bitstream

→ ~3% variation

	VTM	BMS	ratio
Log concatenation	67.6	225.1	333%
Bitstream concatenation	51.8	167.2	323%

	description	Timing ratio variation
Assessment 1	Different runs in exactly same test configuration	0.5%
Assessment 2	CPU generation	24%
Assessment 3	Computer grid workload	19%
Assessment 4	Turbo mode ON/OFF	11%
Assessment 5	Compiler version	21%
Assessment 6	Cmake settings/Compiler optimization	9%
Assessment 7	SIMD level	13~18%
Assessment 8	decoder software options	27%
Assessment 9	With/without Bitstream concatenation	3%
Assessment 10		

Assessment 10 - VTM software type

VTM software from BMS with JEM macro off, or BMS with JEM tools disabled by configuration file

→ ~1% variation

	VTM	BMS	Ratio
BMS macro off	65.2	217.0	333%
BMS macro on	64.1	217.0	338%

	description	Timing ratio variation
Assessment 1	Different runs in exactly same test configuration	0.5%
Assessment 2	CPU generation	24%
Assessment 3	Computer grid workload	19%
Assessment 4	Turbo mode ON/OFF	11%
Assessment 5	Compiler version	21%
Assessment 6	Cmake settings/Compiler optimization	9%
Assessment 7	SIMD level	13~18%
Assessment 8	decoder software options	27%
Assessment 9	With/without Bitstream concatenation	3%
Assessment 10	BMS macro ON/off	1%

Fix what can be fixed (e.g. cmake files, compiler, SIMD levels, decoding conditions) in the CTCs or in the CEs test conditions

Parameters	Decoder	Encoder
SIMD settings	SCALAR. Build with auto-vectorization off.	SSE42
Compiler	Inform in contribution the compiler used for testing.	free
Compiling option	-O3	-O3
Decoder options	No MD5, no output	
Bitstream concatenation	With Bitstream concatenation	
Grid turbo	Inform in the contribution	Inform in the contribution
BMS macro	Always ON (VTM activated by cfg)	Always ON (VTM activated by cfg)

Assume a ~10% inaccuracy on the timing ratio

Rely more on code analysis for complexity at decoder

Do not rely on SIMD optimized code to assess complexity