# Calorimeter Trigger Layer 2 Design and Status

Greg Iles, Andrew Rose – Imperial College London

Mark Baber, Aaron Bundock, Simon Fayer, Sarah Greenwood, Geoff Hall, Sarah Ives, Robyn Lucas, Jad Marrouche, Alex Tapper – Imperial College London

John Jones – Iceberg Technology

Robin Aggleton, Jim Brooke, Sudeshna Dasgupta, Mark Grimes, Christopher Lucas, Dave Newbold, Sudarshan Paramesvaran, Tom Williams – University of Bristol

Tim Durkin, Kristian Harder, Sam Harper, David Petyt, Dave Sankey, Claire Shepherd-Themistocleous, Alessandro Thea – Rutherford Appleton Laboratory

Florian Beaudette, Philippe Busson, Luca Mastrolorenzo, Benoit Renaud, Thierry Romanteau, Jean-Baptiste Sauvan, Alex Zabi – LLR

#### Overview

- Introduction
- Status of the TMT concept
  - Results of the September TMT Integration Test
- Status of the Layer-2 hardware
  - The MP7 processor card
  - Status of standalone testing
  - MP7 R1
  - Planning and Production
- Status of the Layer-2 firmware
- Status of the Layer-2 software
- Overall Project Status
- Conclusions

# Introduction

## Conventional trigger



#### **Conventional Trigger**

- Data is processed in regions
- Boundaries between regions must be handled by sharing or duplicating inputs
- Volume of data reduced at each stage by selecting and discarding candidates
- When volume of data has been sufficiently reduced it can be passed to the global trigger

#### Time-multiplexed trigger



#### Time-Multiplexed Trigger

- Data from an event is buffered and retransmitted to the first processing node over *N* bunch crossings
- Data from the next event is buffered and retransmitted to the second processing node, again, over *N* bunch crossings
- Process is repeated in round-robin fashion across ≥N processing nodes
- Because both algorithm latency and data volume are constant, dataflow is fully deterministic and no complex scheduling mechanism is required

#### Reminder of TMT concept

- The Time-multiplexed architecture allows all data to arrive in geometric order:
  - Towers at given  $\phi$  always occupy the same bits on the same optical links
  - Towers arrive in order of increasing |η|
- This converts a 2D geometric problem to a 1D problem
- This allows all algorithms to be fully pipelined +:
  - The processing is localised
  - Fan-outs reduced
  - Routing delays minimised
  - Register duplication eliminated
  - Also only need to develop one FPGA design

<sup>†</sup>That is, pipelined at the full data rate, not at the bunch-crossing rate

Imperial College London

# Status of the TMT concept: TMT integration test

## September TMT test objectives

https://twiki.cern.ch/twiki/pub/CMS/L1Calo2013/TMTtestdemonstration\_final.pdf

#### 7. Test criteria:

A Pass of the TMT test requires the following objectives to be met:

- Reliable transmission of data @ 10Gbps asynch between PP and MP (36 links with full error checking) "reliable" to mean running for a series of at least 6 hour runs with zero alignment errors
- Successful alignment of all links
- Implementation of an algorithm and successful transmission of data through it
- Verification of latency and how it compares to TDR value -in particular the SerDes link

#### TMT layout

#### MP7's used here as PP's



#### TMT test setup @ 904



#### 720 Gbps of data going through the MP board The MP is processing data from the entire calorimeter

#### Results

#### 7. Test criteria:

A Pass of the TMT test requires the following objectives to be met:

- Reliable transmission of data @ 10Gbps asynch between PP and MP (36 links with full error checking) "reliable" to mean running for a series of at least 6 hour runs with zero alignment errors
- Successful alignment of all links

The stability of the MP7 links was extensively tested several times overnight for 8 hours at a time, 72 links operating, monitoring the CRC counters and the alignment flags.

#### NO CRC errors NO alignment errors

### September TMT test objectives

Key aim of the September integration test:

• Implementation of an algorithm and successful transmission of data through it

Must necessarily be representative of final complexity!

Calorimeter trigger algorithms can be broadly divided into three categories:

- Compact objects e/γ/τ
- Extended objects Jets
- Global objects Ring sums

#### September TMT test objectives

- Compact objects e/γ/τ
  - Use the 2×2 Wisconsin clustering algorithm
  - Baseline algorithm currently in CMSSW
- Extended objects Jets
  - Use an 8×8 tower circular jet
  - Equivalent to cone jet (r=0.35)
  - Other algorithms available
- Global objects Ring sums
  - Use full granularity equivalent of what is done in current trigger

## TMT algorithms - Clusters



Many algorithms available, each with comprehensive test suite Subset of algorithms were used in TMT integration tests

Algorithm	Written	Full VHDL Test-bench	Used in integration test
2X2 SUM	$\checkmark$	$\checkmark$	$\checkmark$
Cluster overlap filter	$\checkmark$	$\checkmark$	×
Cluster classifier (e, $\gamma, \tau$ )	$\checkmark$	$\checkmark$	×
Cluster isolation	partially	×	×
Cluster sort in φ	$\checkmark$	$\checkmark$	$\checkmark$
Cluster sort in η	$\checkmark$	$\checkmark$	<b>×</b> †

## TMT algorithms - Jets

Many algorithms available, each with comprehensive test suite Subset of algorithms were used in TMT integration tests

Algorithm	Written	Full VHDL Test-bench	Used in integration test
2x1, 4x1, 6x1, 8x1 strip formation	$\checkmark$	$\checkmark$	$\checkmark$
2-, 3-, 4-, 5-, 6-, 7-, 8-strip wide sums	$\checkmark$	$\checkmark$	$\checkmark$
• 4x4, 6x6, 8x8 square jets	$\checkmark$	$\checkmark$	×
• 4x4, 6x6 circular jets	$\checkmark$	$\checkmark$	×
8x8 circular jets	$\checkmark$	$\checkmark$	$\checkmark$
Jet overlap filter	×	×	×
Jet sort in φ	$\checkmark$	$\checkmark$	$\checkmark$
Jet sort in η	$\checkmark$	$\checkmark$	<b>×</b> †

Image: starting sta

#### TMT algorithms - Ring sums

Many algorithms available, each with comprehensive test suite Subset of algorithms were used in TMT integration tests

Algorithm	Written	Full VHDL Test-bench	Used in integration test
Multiply by trigonometric coefficients	$\checkmark$	$\checkmark$	$\checkmark$
3x1, 9x1, 18x1, 36x1, 72x1 strip formation	$\checkmark$	$\checkmark$	$\checkmark$
• Scalar sum (ET)	$\checkmark$	$\checkmark$	$\checkmark$
• Vector sum (MET)	$\checkmark$	$\checkmark$	$\checkmark$
Count towers over threshold (PU estimator)	$\checkmark$	$\checkmark$	$\checkmark$
Accumulate in η	×	×	×

Frame 0043 : Frame 0044 :	505050bc 505050bc 505050b 505050bc 505050bc 505050b	c 505050bc c 505050bc	505050bc 505050bc	505050b 505050b	c 50505 c 50505	50bc 50505 50bc 50505	0bc 50505 0bc 50505	0bc 0bc											
Frame 0045 : Frame 0046 :	505050bc 505050bc 505050b 505050bc 505050bc 505050b	c 505050bc c 505050bc	505050bc 505050bc	505050b 505050b	c 50505 c 50505	50bc 50509 50bc 509	60bc 50505	0bc 50505	0bc 50505	0bc 50505	0bc 50505	0bc 50505	0bc			_			
Frame 0047 : Frame 0048 :	505050bc 505050bc 505050b 505050bc 505050bc 505050b	c 505050bc c 505050bc	505050bc 505050bc	505050b 505050b	c 50505 c 50505	50bc 50! 50bc 50!	Algo	orith	m		Clo	ck c	ycles	5	B>	<			
Frame 0049 : Frame 0050 :	505050bc 505050bc 505050b 505050bc 505050bc 505050b	c 505050bc c 505050bc	505050bc 505050bc	505050b	c 50505 c 50505	0bc 50!							,						
Frame 0051 :	505050bc 505050bc 505050b	c 505050bc	505050bc	505050b	c 50505	0bc 50!	Clust	ters			08-	71 =	27		1.				
Frame 0052 : Frame 0053 :	505050bc 505050bc 505050b	c 505050bc	505050bc	505050b	c 50505	50bc 50!					90	/	~/		4.	ך כ	ominat	od by s	ort
Frame 0054 : Frame 0055 :	505050bc 505050bc 505050b 505050bc 505050bc 505050b	c 505050bc c 505050bc	505050bc 505050bc	505050b	c 50505 c 50505	50bc 50! 50bc 50!											ommat	eu by s	
Frame 0056 :	505050bc 505050bc 505050b	c 505050bc	505050bc	505050b	c 50505	0bc 50!	Jets				104	+-71 =	= 33		5.5	5			
Frame 0058 :	505050bc 505050bc 505050b	c 505050bc	505050bc	505050b	c 50505	0bc 50!													
Frame 0059 : Frame 0060 :	505050bc 505050bc 505050b 505050bc 505050bc 505050b	c 505050bc	505050bc	5 Frame	0066 :	000000	Rina	sum	าร		82-	71 =	11		1.8	3			9000
Frame 0061 :	505050bc 505050bc 505050b	c 505050bc	505050bc	5 Frame	0067 : 0068 :	000006		5011				/-				-			3006
Frame 0063 :	505050bc 505050bc 505050b	c 505050bc	505050bc	5 Frame	0069 :	00000045	00000039	00000043	00000046	00000045	00000039	00000043	00000046	00000000	00000000	00000000	00000000	00000000	00000000
Frame 0064 :	505050bc 505050bc 505050b	c 505050bc	505050bc	5 Frame	0070 :	00000045	00000039	00000043	00000046	00000045	00000039	00000043	00000046	000000000	00000000	00000000	000000000	000000000	00000000
Frame 0065 : Frame 0066 :	505050bc 505050bc 505050b	c 505050bc	505050bc	5 Frame	0072 :	00000045	00000039	00000043	00000046	00000045	00000039	00000043	00000046	0000000000	000000000	000000000	000000000	000000000	00000000
Frame 0067 :	505050bc 050 <mark>50bc 505050b</mark>	505050bc	505050bc	5 Frame	0073 :	00000045	00000039	00000043	00000046	00000045	00000039	00000043	00000046	00000000	90000000	00000000	00000000	00000000	00000000
Frame 0068 : Frame 0069 :	505050bc 050531kt 505150b 505050bc 05050bc 505050b	505050bc	505050bc	5 Frame	0074 :	00000045	00000039	00000043	00000046	00000045	00000039	00000043	000000046	000000000 0	300000000	000000000	000000000	000000000	000000000
Erame 0070 :	00040001 01040001 0204000	03040001	04040001	0 Frame	0076 :	00000045	00000039	00000043	00000046	00000045	00000039	00000043	00000046	00000			000000000000	000000000	00000000
Frame 0071 :	7c4a2f4f f43b8fdd 2645b8f	0 343aa950	2df84254	3 Frame	0077 :	000000045	00000039	00000043	00000046	00000045	00000039	00000043	00000046	0000000000	000000000	000000000	000000000000000000000000000000000000000	000000000	000000000
Frame 0072 :	d3alddc5 40cd6e3f 06971a6	4 ae8b2fa3	a17304d2	2 Frame	0079 :	00000045	00000039	00000043	00000046	00000045	00000039	00000043	00000046	00000 00 0	00000000	00000000	0000000000	000000000	00000000
Frame 0074 :	307932aa 33220c87 fec1f13	5 5dd438a0	118bb6b3	7 Frame	0080 :	00000045	00000039	00000043	00000046	00000045	00000039	00000043	00000046	000000000	00000000	00000000	00000000000	000000000	00000000
Frame 0075 : Frame 0076 :	227b8b08 2262cbae 7abb7f5	/ b69/4ac2 3 f2dd48d2	201826c9 805aa127	aFrame	0082 +	00000045	00000039	00000043	00000040	00000045	00000039	00000043	00000040	000004904 0	9465b9be	04637606	0000003d	000000000	000000000
Frame 0077 :	9ec95bcb 3ede43df 8c8ed42	9 ef6dc227	2cb7ab7d	2 Frame	0083 :	00000045	00000039	00000043	00000046	00000045	00000039	00000043	00000046	0000466e 0	9460bb58	045b6b8a	0000003c	000000000	00000000
Frame 0078 :	885414c6 91b5f975 d4c902a	5 140c9526	fcc97ac3	2 Frame ₀ Frame	0084 :	00000045	00000039	00000043	00000046	00000045	00000039	00000043	00000046	00004931 0 0000495c 0	94633080 9463666c	045/3326 045ba69c	0000003T	000000000	000000000
Frame 0080 :	e5303359 64087ff3 219cd11	3 17745efc	1624bb30	g Frame	0086 :	00000045	00000039	00000043	00000046	00000045	00000039	00000043	00000046	000041c0 0	945bd73b	04656aca	0000003b	000000000	00000000
Frame 0081 :	1e87823c b0bdb5d1 5e2926e	9 21803c79	e5fe9ce0	e Frame	0087 :	00000045	00000039	00000043	00000046	00000045	00000039	00000043	00000046	00004834 (	9465fa1c	04634a7d	0000003e	000000000	000000000
Frame 0082 :	4e36d6d4 305809b9 9ddd387 0be8da3f 40eb24b7 a6ba228	a e0401837 a bae915f9	72f76dbd	7 Frame	0089 :	00000045	00000039	00000043	00000046	00000045	00000039	00000043	00000046	0000492e (	945dc521	0463fc76	000000040	000000000	00000000
Frame 0084 :	e8aa0e3f bf19ca79 f946e79	9 02a5bc3a	bad96950	b Frame	0090 :	00000045	00000039	00000043	00000046	00000045	00000039	00000043	00000046	00004cfa 0	94673219	04638151	0000003f	000000000	00000000
Frame 0085 : Frame 0086 :	c0c883e5 71eaa9bd 40d9371 a2303ee5 071ed8aa fbbecae	7 9a8ea8da 7 bfa9cdf4	2dc2a619 f344d103	3 Frame 5 Frame	0091 :	00000045	00000039	00000043	00000046	000000045	00000039	00000043	00000046	0000482e 0	946405a0 945c5079	04615C20 045edfa6	0000003c	000000000	000000000
Frame 0087 :	7c0910ac a467c3b1 7727e13	a e9cb4bf3	39637066	d Frame	0093 :	00000045	00000020	00000043	00000046	00000045	00000039	00000043	00000046	000044e1 0	9466c5a5	0462fcf7	00000040	000000000	00000000
Frame 0088 :	6e231211 693fef2f 30e5783	b 87fd6cfe	b00c12d0	⊖ Frame ⊖ Frame	0094 :	0000004	00000039	00000043	000000046	00000045	00000039	00000043	00000046	000048ab 0	94643786 94667daf	045cb2ff 045d3c39	0000003e 00000040	000000000	000000000000000000000000000000000000000
Frame 00000 :	2a887511 c97a9129 1522ac4	4 618288e4	87698c7d	2 Frame	0096 :	0000004	00000039	00000043	000000046	00000045	00000039	00000043	00000046	00004b41 (	94697c31	04581123	0000003e	000000000	00000000
Frame 0091 :	5cd72620 26db2c2a 5c110a6	b f5cc5e8a	dc71e4fe	8 Frame	0097 :	0000004	000000000	00036690	00036556	00000045	00000039	00000043	00000046	00004deb (	945c6ba3	045a7ff0	00000040	000000000	00000000
Frame 0092 : Frame 0093 :	e6cfff78 62335a48 02faf40	a /129ce50 7 3d40dede	244abb48 e922a226	5 Frame	0099 :	00031720	00031610	00030985	00021328	00000045	00000033	00000043	00000046	00003e17 (	045dcff9	0455ea24	00000035	000000000	00000000
Frame 0094 :	1f0a8deb 80f4fc90 56380de	9 bf6227be	e5b08855	2 Frame	0100 :	00034728	00034527	00030ea9	00030018	0000004	00000039	00000043	00000046	00004f69 (	946069da	0465b4dd	00000040	000000000	00000000
Frame 0095 : Frame 0096 :	46c02a3a 5d5/a00b be130a3 8f26a963 bd854c51 0fa7648	d 92b2t826 4 1fcfac36	a8/dc42b 361f6111	1 Frame	0102 :	0003411c	000331b8	00033398 00032a45	00032e12	00000004	00000039	0000043	000000046	0000437a (	94645130 945b7d18	04598007 0457a018	0000003b	000000000	000000000
Frame 0097 :	2e8fef79 a427e37c 54a0444	4 0386f4db	88cdaed0	a Frame	0103 :	0002fea9	0002fd8a	0002ed3d	0002d238	00000045		00000015	00000046	00004671 0	94444447	045d7c64	0000003e	000000000	00000000
				Frame Frame	0104 :	00032346	00030e2b	0002df0f	0002bd1c	00111a43 001173c2	001e5995 001e2/43	001e1d91 001e2096	001ddd10 001dfac1	00004038 0 00004bcf 0	945aet20 945aac4b	045dba5a	000000039	000000000	000000000
				Frame	0106 :	000356aa	00032344	00031d31	0003150a	001dd042	001dafc3	001d4616	001d0598	00003d0f 0	945b11d7	0455dc6d	00000033	000000000	00000000
				Frame	0107 :	0003710f 00034d92	00037010 00030f0a	000324b0 0002c309	0003208a 0002bc30	001d8e43 001ccf95	001cdb16 001c9a42	001cc114 001c7122	001c950e 001c3998	000046d4 0	9465251e 945238cc	04616c8d 045c9362	00000031	000000000	000000000000000000000000000000000000000
	Algorithms	sare		Frame	0109 :	00031298	0002ee2a	0002c61e	0002c28a	001d27a5	001cd123	001c3797	001c3514	000044f1 (	9456bc76	04643d43	0000003d	000000000	00000000
				Frame	0110 :	00034c98	00031517	0002f11e	0002dbaf	001cf938	001ce4c1	001ce308	001c740e	000046ea (	946494ac	0461d36d	0000003c	000000000	00000000
	running	at		Frame	0112 :	0003beld	00033c19	00032288	0003141a	00203409	001e648a	001e440c	001de1c1	000000000	000000000	000000000	000000000	000000000	00000000
	ronning	at		Frame	0113 :	00038093	000364a3	00035e29	00034b05	001fb989	001f0b0a	001ea90b	001dda88	00000000 0	00000000	00000000	00000000	000000000	00000000
				Frame	0114 : 0115 :	000358b5	00032728	00030418	00030129	001f2689	001ed58b	001ea18c	001e250b	000000000 (	300000000	000000000	000000000	000000000	000000000
	250IVIH	Z		Frame	0116 :	0002ee8d	0002e935	0002e0c4	0002dbc3	001e6f8c	001e098b	001de98e	001d560d	00000000	90000000	00000000	00000000	000000000	00000000
				Frame Frame	0117 : 0118 :	00032f45	00032318	000320aa	0002fb10	001df196	001d9b17	001d6994	001c831a	000000000 6	300000000	000000000	000000000	000000000	000000000000000000000000000000000000000
				Frame	0119 :	00032498	00030aa1	0002ba06	0002b4aa	001e090c	001db617	001da415	001d7423	00000000	00000000	00000000	00000000	00000000	00000000
13/11	/2013			Frame	0120 :	0002d6bb	0002b32f	0002b1b1	6002b02a	001f3696	001eb40d	001e9c18	001e8817	00000000	900000000	00000000	00000000	000000000	00000000

#### Results (1)

Data injected

into PP

- Implementation of an algorithm and successful transmission of data through it
- Random data passed through an emulator was used in the testing of the algorithms

Compared emulated results (solid line) with those from the MP7 (markers)

Time-

multiplexed

C++ emulator and hardware match precisely



#### Results (2)

Data injected

into PP

- Implementation of an algorithm and successful transmission of data through it
- Random data passed through an emulator was used in the testing of the algorithms

Compared emulated results (solid line) with those from the MP7 (markers)

Time-

multiplexed

C++ emulator and hardware match precisely



#### Results (3)

- Implementation of an algorithm and successful transmission of data through it
- Random data passed through an emulator was used in the testing of the algorithms



#### Results

• Implementation of an algorithm and successful transmission of data through it

TMT integration test successfully demonstrated simultaneous operation of:

- Compact algorithms ( $e/\gamma/\tau$ )
- Extended algorithms (jets)
- Global algorithms (ring sums)

These algorithms are representative of the complexity of algorithms expected in the final system

#### Results – Latency Measurement

• Verification of latency and how it compares to TDR value -in particular the SerDes link

Source of Latency	BX (TDR)	BX –measured in Sept 2013		
L1 processing + TM	10	7		
L1/L2 SerDes (Tx+Rx) @ 10Gbps	5	5		
L1/L2 SerDes Align Data	1	1		
L1/L2 cable (20m)	4	4		
L2 Processing	8	5.5 (clustering, jets, ring sums)		
L2/GT SerDes (Tx+Rx)	5.5	5 (identical link to L1/L2 above)		
L2/GT SerDes Align Data	1	1 (identical link to L1/L2 above)		
L2/GT cable	0.5	0.5		
De-multiplex	6	7		
TOTAL	41	36		

13/11/2013

#### Lessons learned from test

- Floor-planning
  - Huge impact on algorithm design
  - Structure the algorithm to map optimally onto the FPGA
  - Reduces risk that after many hours of routing 6 million nets just 1 or 2 fail to meet timing - exceedingly annoying
  - Significant timing improvement
  - Only viable if signals remain relatively local
- Full pipelining of algorithms is essential even relatively innocuous looking fan-outs in chips this large have the potential to kill off the entire design
- uHAL and IPbus are working well, and have made debugging and multi-user crate access both possible and easy
- AMC13 has worked extremely well for us during the test

#### Integration test summary

- Have demonstrated successful TMT setup @ CERN
- Pipeline tested from injecting data into PP and capturing at the MP
- Many algorithms were loaded and results were in perfect agreement with expectations
- The TMT test was very successful can be measured by the objectives setout before the test began:

#### All test criteria were met

#### Integration test setup: Next Steps

- Plan is to continue implementing and testing algorithms over the next few weeks

   have a testbench, now performance and resource-usage can be measured
- Improve structuring of algorithms
- Integrate with other CMS components, i.e. DAQ via AMC13, IPMI and other  $\mu\text{TCA}$  infrastructure
- Make the MP7s and infrastructure firmware available to others to start working with and gaining experience
- Start putting into place 'polished' online software

# Hardware:

#### The Master-Processor, Virtex-7 (MP7)

- uTCA form factor
- ~1.5Tb/s signal processor
- Latest-generation Xilinx Virtex-7 FPGA technology
- GbE, AMC<sub>13</sub>/TTC/TTS, PCIe, SAS, SATA, SRIO
- Advanced boot-loader & diagnostics (full system test at start-up)
- On-board firmware repository
- Pin-compatible FPGAs allow costperformance balance
- 2×144Mbit 550MHz QDR RAM (optional)
- USB-OTG serial-debug console



XC7VX690T, 72 TX + 72 RX @ 10Gb/s

## History

- Summer 2011 started considering successor to the Mini-T5
- November 2011 schematics started (before Virtex-7 or MiniPODs were available!)
- June 2012 first manufacturing run
- Autumn 2012 to Summer 2013 extensive testing, several further manufacturing runs
- June 2013 submit a first revision of the board prior to production runs
- August 2013 first revision boards now in hand and under test





#### **Current Hardware in Hand**

- 2 revision-0 cards with 48 links (XC7VX485T) engineering silicon
- **3** revision-0 cards with 48 links (XC7VX485T) production silicon
- 3 revision-0 cards with 72 links (XC7VX69oT) engineering silicon
- 1 revision-1 cards with 48 links (XC7VX485T) production silicon
- **5** revision-1 cards with 72 links (XC7VX69oT) production silicon

14 cards in-hand<sup>+</sup>

1 at Imperial College, UK
1 at RAL, UK
1 at Berkeley, USA
5 at CERN B904
3 at CERN Meyrin under final validation
2 at CERN Meyrin under test
1 at CERN Meyrin in display case

# Hardware: Standalone Testing

# Working at the cutting edge

As noted previously, the board was designed before engineering silicon was even available

A lot of the work to understand the board involved direct contact with the Xilinx silicon engineers. This has included alerting them to errors in the datasheets. It turns out that exactly following the datasheet is no substitute for hands-on experience.

A lot of things changed between the original FPGA specification documents and the production silicon

## Working at the cutting edge (ii)

- In particular, power consumption and thermal dissipation were far higher than Xilinx original estimate
- These are now both well understood and controlled but it took time to understand these issues



Anodized production heatsink

#### Working at the cutting edge (iii)





Imperial College London

#### Working at the cutting edge (iv)



## **Optical link performance**

Optical links worked perfectly out-of-the-box

- MiniPODs are using default pre-emphasis (minimal)
- TX differential swing is minimal
- No pre-emphasis/de-emphasis in MGTs





	GTH_X0Y0				
Y MGT Settings					
- MGT Alias	GTH0_210				
- Tile Location	GTH_X0Y0				
- MGT Link Status	10.0 Gbps				
- PLL Status	QPLL LOCKED				
– Loopback Mode	None 💌				
– Channel Reset	Reset				
- TX/RX Reset	TX Re RX Re				
- TX Polarity Invert					
- TX Error Inject	Inject				
- TX Diff Output Swing	250 mV (0 💌				
- TX Pre-Cursor	0.00 dB (0 💌				
- TX Post-Cursor	0.00 dB (0 💌				
- RX Polarity Invert					

13/11/2013

Imperial College London

#### Link validation

Links are validated between boards rather than in loopback – <u>a more rigorous test</u>



72 (a) 10Gb/s  $\rightarrow$  62.2Pb/day  $\rightarrow$  ~1Eb every 2 weeks
### 72 links validated simultaneously

- Eye-diagrams measured inside • XC7VX690T FPGA
- All 72 links running PRBS7 at 10Gb/s, LPM





#### Link validation

XC7VX485T – GTX transceivers



- Various line rates validated
- Eye-diagrams for all line-rates meet industry standards
- MGT configuration <u>done at runtime, not build-time</u>, giving significantly greater flexibility and control

#### Link validation

- LHC-synchronous board-board communication has been demonstrated at 6.4Gb/s
- Both boards received the TTC clock over the backplane from the AMC-13
- <u>The MP7 has been demonstrated to work both SYNCHRONOUSLY and</u> <u>ASYNCHRONOUSLY, whichever way the CMS trigger decides to go</u>

## Clocking

- The clock architecture and clocking resources of the Virtex-7 are very different from previous generations
- Several months of work has gone into understanding the configurationoptions, the performance and the limitations of each of these clocking resources <u>so that there are no nasty surprises in future</u>
- This was one area with many "undocumented" features which required direct dialog with Xilinx
- We are now confident that this is thoroughly understood

#### Other tests

All the auxiliary subsystems on the MP7 have also been validated:

- Mezzanine I/O
- QDR-II+ RAM
- Board communication (Ethernet)
- USB 2.0 OTG, IPMI, SD card repository (microcontroller)
- SPI PROM (protected bootloader)
- Thermal dissipation, humidity & temperature monitoring
- Power consumption, phasing & monitoring



I/O fan-out mezzanine for debugging 30 differential pairs to/from the FPGA 3v3, 2v5 and 1v8 supplies Dedicated I2C lines to the microcontroller

# Hardware: MP7 Revision 1

### List of changes on R1 card

- Swap LTM4628 power module for LTM4620 (higher current rating, footprint compatible)
- Swap LTM4606 power module for LTM4601 (higher current rating, not low EMI part, different footprints)
- Add potential divider on one P/I/V/T monitor
- Invert polarity of enable line on an oscillator
- Change the value of resistors to make tri-colour LEDs glow white, not purple
- Add four extra bulk decoupling caps

### Testing of the R1 cards

- 1 board already in use in 904
- 3 boards undergoing final validation (links already tested)
- 2 boards partially tested

Note that both R1 cards have been assembled with production silicon:

These are the FINAL FPGAs on the FINAL board design

# Hardware: Planning and Production

#### **Production and Planning**

A production specification and manufacturing contract document has been drawn up based on a CERN template

Been in discussion with CERN purchasing



### MP7 users





#### Documentation

Memorandum of Understanding has been drafted specifying what will be provided and supported when a user purchases an MP7 card, in terms of

- Hardware
- Firmware
- Software

Detailed documentation on the specification and design of the board exists

Testing, servicing and maintenance documents still to be done



### MP7 timeshare

The MP7 cards in CERN building 904 (those used in TMT integration test) are currently available on a "timeshare" basis for people who plan to use the MP7 to gain experience.



# Firmware



#### Implementation Status



#### Floorplan of FPGA: Integration Test



Imperial College London



#### Floorplan of FPGA

Towers

Sorting



#### Resource usage of FPGA

Resource	Fracti Entire	Fraction of	
	Used for infrastructure	Used for algorithm + infrastructure	used
Registers	9%	21%	20%
LUTs	19%	36%	35%
DSPs	0%	8%	13%
BRAM	12%	12%	0%

All Algorithms and Infrastructure: 7 hour build-time

Much better than the >24 hours reported elsewhere

#### Some Firmware Notes

- All MP7 firmware is in the official CMS CACTUS repository
  - <u>http://svnweb.cern.ch/reps/cactus/trunk/boards/mp7/base\_fw</u>
  - Around 5ok lines of HDL requires serious code management now
  - The source is available for re-use by other projects
  - Scripted flow allows bitfile to be built 'out of the box'
- Collaboration on firmware has been a positive experience
  - Re-use of HCAL TTC blocks, debugging & development of IPbus/UDP
- Algorithm development looking ahead
  - Developing a UK / LLR proposal for *fully modular* algorithm firmware
    - Automatically ensures firmware-emulator bit-level consistency
    - Uniform interface to algorithm sub-blocks, with flexible dataflow
    - Data insertion / capture at any internal connection between blocks
  - Reminder: algorithm development and tuning is a shared effort across several institutes



# Software

#### Software

- The MP7 has been built around the official CMS IPbus release
- The control software for the board is fully based on the official CMS uHAL framework, <u>as required by the Level-1 Trigger Online Software specification</u> currently under peer-review
- A set of core software modules exist to control key firmware components of the MP7 and to control peripheral components.
- These components are fully reusable and are currently being added to the official CMS trigger-upgrade SVN (CACTUS) repository
- This software has been being used, tested and developed for over a year now!
  - Used in the July Integration Test
  - Used in the September Integration Test

#### Software details

- MP7 firmware and software components tested and validated:
  - X-points, Clocking, TTC interface, Transceivers, Spy buffers
- Additional µTCA control and configuration software developed for the July Integration test
- A basic graphical frontend providing access to multiple
  - MP7 boards via IPbus & IPMI
- AMC13 frontend based on the AMC13Tool: TTC clock config, L1A and BGos monitoring
- IPMI interface
  - Based on the the IPMITool package
  - Power management, board discovery, sensor monitoring



### System-level integration

General observation:

The level-1 trigger cannot afford the effort required to repeat legacy online system development

Calo layer-2 group held a kick-off meeting on how to tackle this, which was followed up with a meeting with representatives of other subsystems and then presented to the L1-trigger community as a whole

#### The Basic Idea

Make maximum use of the flexibility offered by  $\mu$ TCA, IPbus and uHAL!



13/11/2013

Imperial College London



#### Software conclusions

We are making maximum use of the advanced position of the Layer-2 hardware, firmware and low-level software to start attacking the questions of system-level integration which will affect everyone

We are in open dialogue with all level-1 subsystems on this as this will affect everyone

# **Overall Project Status**

#### CMS Trigger Upgrade Schedule

#### Relevant to Layer 2 -

Imperial Co

Title	#	2013      2014        O1      O2      O3      O4      O1      O2      O3      O4
🖞 CMS Phase 1 Trigger Upgrade Project	0	
Calorimeter Trigger	1	
oSLB/oRM	2	
Design and Prototyping	3	
Procurement	4	
Production and Testing	5	
Integration Testing at CERN	6	
Remove cables and install Fibers at P5	7	
Installation at P5	8	· · · · · · · · · · · · · · · · · · ·
Commissioning at P5	9	· · · · · · · · · · · · · · · · · · ·
System commissioned and ready for data taking	10	
Calorimeter Trigger	11	
Hardware Development and Testing	12	······································
Stage 1 oRSC Card	13	
Prototype Design, Maunfacture, Testing	14	
Pre- production Prototype Design, Manufacture and Testing	15	
Production Readiness Review	16	
Procurement	17	
Production and Testing	18	
Layer 1 processor	19	
Prototype Design, Manufacture and Testing	20 (	
Pre- production Prototype Design, Manufacture and Testing	21	
Production Readiness Review	22	. <b>↓</b> ♦–
Procurement	23	
Production and Testing	24	
Layer 2 processor	25	,
Prototype Design, Manufacture and Testing	26	
Pre- production Prototype Design, Manufacture and Testing	27	
Production Readiness Review	28	;   <mark>+</mark> 0
Procurement	29	
Production and Testing	30	
Hardware production completed	31	
Software and Firmware Development	32	
Infrastructure Software and Firmware Development	33 (	
Infrastructure Software and Firmware complete	34	
Algorithm software and Firmware development	35	
Initial algorithms complete	41	◆ · · · · · · · · · · · · · · · · · · ·
Testing	42	
Testing in lab	43	
Testing at CERN	44	
TMT Processing Demonstration completed	45	i L•∲
Integration with CMS infrastruture complete	46	i i i i i i i i i i i i i i i i i i i
Integration and Commissioning	47	
Integration testing at CERN	48	
Installation at P5	49	
Installation at P5 completed	50	
Commissioning with HCAL and parallel operation with LHC beam	51	
System commissioned and ready for data taking	52	

#### CMS Trigger Upgrade Schedule

		Status
Layer 2 processor	25	-
Prototype Design, Manufacture and Testing	26	Completed on schedule 🗸
Pre- production Prototype Design, Manufacture and Testing	27	Completed on schedule 🗸
Production Readiness Review	28	Pending
Procurement	29	Pending (Contract prepared)
Production and Testing	30	Pending
Hardware production completed	31	Pending
Software and Firmware Development	32	
Infrastructure Software and Firmware Development	33	Completed except DAQ
Infrastructure Software and Firmware complete	34	Completed except DAQ
Algorithm software and Firmware development	35	<u> </u>
Initial algorithms complete	41	Completed and demonstrated ahead of schedule $\checkmark$
Testing	42	-
Testing in lab	43	Testing complete 🗸
Testing at CERN	44	Testing complete 🗸
TMT Processing Demonstration completed	45	Completed on schedule (using MP7s as Pre-processors) $\checkmark$
Integration with CMS infrastruture complete	46	On-going
Integration and Commissioning	47	
Integration testing at CERN	48	Partially complete (waiting on other systems)
Installation at P5	49	Pending
Installation at P5 completed	50	Pending
Commissioning with HCAL and parallel operation with LHC beam	51	Pending
System commissioned and ready for data taking	52	Pending

#### Personnel

#### <u>Management</u>

- Hall, Tapper Imperial
- Newbold, Paramesvaran Bristol

#### Physics Studies

- Baber, Fayer, Ives, Marrouche, (Rose, Tapper) – Imperial
- Aggleton, Brooke, Grimes Bristol
- Harper, Petyt, Shepherd-Themistocleous – RAL
- Beaudette, Mastrolorenzo, Sauvan, Zabi – LLR

#### <u>Hardware</u>

- Bundock, Greenwood, Iles, Rose Imperial
- Jones Iceberg Technology
- Durkin RAL

#### <u>Firmware</u>

- Bundock, Iles, Rose Imperial
- Jones Iceberg Technology
- Dasgupta, (Newbold), Williams Bristol
- Harder, Sankey RAL
- Busson, Renaud, Romanteau LLR

#### Online Software

- Bundock, Rose Imperial
- Brooke, Lucas, Paramesvaran, Williams
  Bristol
- Harder, Sankey, Thea RAL
- + new students joining the project
- new groups have expressed interest in joining physics/algorithm effort

# Conclusions

### Conclusions: TMT

- The September TMT Integration Test was a success
- Many lessons were learned which reaffirmed the motivations for choosing the TMT architecture over a conventional architecture

#### Conclusions: Hardware

- The MP7 card has now been in-hand for over a year
- Every part of the card has been thoroughly validated and understood, in particular the optical links and the Virtex-7 which are the most important and most complex parts of the boards
- A first revision has been received, which has only minor changes from the original design, and are either in use or in the final stages of validation

#### Conclusions: Software and Firmware

- The firmware for the MP7 is in an advanced state and available in the CMS trigger-upgrade SVN (CACTUS) repository
- The low-level software for the MP7 meets the official Level-1 online software requirements and is in the CMS trigger-upgrade SVN (CACTUS) repository
- All firmware is available in the CMS trigger-upgrade SVN repository
- All software and firmware has been designed to be reused by others, should they wish to do so
- Discussions on the system-level software have been initialized by the layer-2 groups with the aim of minimizing software effort required by the trigger community as a whole

#### The take-home message

- a) The TMT principle has been proven to be very successful.
- b) The MP7 is an exceptionally advanced board and is very well understood.
- c) The firmware for the MP7 is in an advanced state.
- d) The low-level software for the MP7 is in an advanced state.
- e) We are using the advanced position of the rest of the project to start attacking the system-level software.
# Spares

#### MP7 board – Bottom half



#### Website





Background Specs

Gallery

The MP7

#### Breaking News

On 18th March 2013, Avago proudly announced the release of MiniPODs capable of operating at 14Gbps. This will, in theory, extend the bandwidth of the MP7 from the already impressive 0.75+0.75Tbps, up to a phenomenal 0.94+0.94Tbps.

We eagerly await the results of testing the MP7 with these new optics!

ee here for the Avago press release

The Imperial Master Processor, Virtex-7 (MP7) is a high-performance all-optical, data-stream processor designed to operate in the challenging conditions of the CMS trigger system at the Large Hadron Collider (LHC). Utilising the high performance Xilinx Virtex-7 FPGA and stateof-the-art fibre optics technologies, the MP7 has the capability to input and output data at a rate of 3/4 Tbit per second, equivalent to the mean global traffic of the entire Internet in 2001. These features are crucial in the operation of the trigger at the LHC where a latency budget of 3.2µs is afforded to readout and process the large volume of data from the detector subsystems, equivalent to processing data at a rate of up to 10 Tbits per second. The MP7 is the baseline trigger processor board for the CMS calorimeter trigger upgrade, whose capabilities are expected to achieve an improvement in the physics performance of the CMS detector under the more challenging full-energy and high-luminosity conditions that will be experienced in the



The Imperial Master Processor, Virtex-7.

#### Features

- A powerful processing capability provided by a high-performance Xilinx Virtex-7 FPGA.
- A large total optical bandwidth input and output of up to 740 Gbps in each direction.
- LVDS I/O with speeds of up to 50 Gbps.
- 288 Mbit fast QDR II+ SRAM, giving memory access of up to 550MHz DDR (1100MHz) per chip.

#### www.hep.ph.ic.ac.uk/mp7

Imperial College London



#### Clusters











## Power supply validation

- We were very concerned about the possibility of power supply changes introducing noise
- Several test cards were made to test alternative power supply designs
- Noise was measured both electrically and by its effect on the error-rate of the 10Gb/s optical links
- No increase in bit error-rate was observed and we were happy to proceed
- These tests have left one board unsuitable for normal, non-benchtop, use. We have enough boards in-hand, however, that we could "burn" a board for the sake of progress





### Automated Link alignment firmware

- Fixed, low Latency GTX/GTH Serdes Link
- Asynchronous or Synchronous with LHC clock
- Rx buffer bypassed for low latency & asynchronous clock capability
- CRC32 integrity checks
- Locks to TTC timing
- Spy buffer
- Only 256 orbits to align the entire system
- Has been in continuous use in hardware for many months



#### Screenshots



#### An Upgrade Processor



### Software Goals

Can we rationalize the board control software?

- MP7, CTP7, MFT7 have a very similar structure
- One software class to control them all?

Can we use the same software model for production and testing?

- Standardize the concept of 'test' (what goes in, what goes out)
- Introduce flexible system-wise tests: from {system A, buffer X} to {system B, buffer Y}
- Are concurrent modes (test/production/development) possible?

