

2      **Development of CMS L1 Tracking Trigger Vertical  
3      Slice System Demonstration**

4      Ted Liu\*, Gregory Deptuch, Jim Hoff, Siddhartha Joshi, Sergo Jindariani, Jamieson Olsen, Luciano Ristori, Nhan  
5      Tran, Marcel Trimpl, Zijun Xu, Hang Yin  
6      Fermilab  
7      Kristian Hahn\*, Kevin Sung, Marco Trovato, Stanislava Sevova  
8      Northwestern University  
9      Ivan Furic\*, Souvik Das, Michele de Gruttola, Jacobo Konigsberg  
10     University of Florida

### Abstract

12     CMS plans to pursue a rich physics program in the HL-LHC era. In order to fulfill these aggressive  
13    goals in the environment of very high pile-up envisioned for HL-LHC, CMS must preserve its ability to  
14    identify in real time signatures of events originating from interesting physics processes. Use of tracker  
15    information in L1 provides a highly efficient handle for pile-up mitigation and will be one of the main  
16    goals of the CMS Phase 2 upgrades. In this proposal we request funding for an R&D program aimed at  
17    developing a tracking trigger solution for the CMS upgrade. This R&D program will allow US to lead  
18    the construction of a Vertical Slice Demonstration System, which will comprise a full tracking trigger  
19    path, running at full speed. This demonstraton system will be tested with simulated high-luminosity  
20    data to measure trigger latency and efficiency, to study overall system performance and to identify  
21    appropriate solutions to possible bottlenecks. Success of this project will provide the needed proof-  
22    of-existance of L1 silicon-based tracking trigger, and will allow the design of the upgraded tracker to  
23    be finalized. Funding is requested to support engineering efforts to develop critical components of the  
24    system: the ATCA hardware used for data dispatching, the associative memory (AM) chip for pattern  
25    recognition, and the pattern recognition mezzanine, including both the hardware and firmware.

---

\*) Denotes PI/Contact person for the corresponding institution. Overall project contact: Ted Liu thliu@fnal.gov

# **Contents**

<b>26</b>	<b>1 R&amp;D Overview</b>	<b>2</b>
<b>27</b>	<b>2 Technical Description and Deliverables</b>	<b>4</b>
<b>28</b>	2.1 Introduction . . . . .	4
<b>29</b>	2.2 System Overview . . . . .	5
<b>30</b>	2.3 Tracker-Trigger interface . . . . .	5
<b>31</b>	2.4 Track Trigger System Architecture . . . . .	6
<b>32</b>	2.4.1 Tracker geometry and Trigger Towers . . . . .	6
<b>33</b>	2.4.2 Formation of Trigger Towers . . . . .	6
<b>34</b>	2.4.3 System architecture . . . . .	6
<b>35</b>	2.4.4 Architecture Flexibility . . . . .	7
<b>36</b>	2.5 Vertical Slice Demonstrator System: Overview and Methodology . . . . .	9
<b>37</b>	2.6 Data Source Stage . . . . .	10
<b>38</b>	2.7 Data Input . . . . .	10
<b>39</b>	2.8 Pattern Recognition Board . . . . .	11
<b>40</b>	2.9 Pattern Recognition Mezzanine Card . . . . .	11
<b>41</b>	<b>3 Relation to Existing Efforts</b>	<b>12</b>
<b>42</b>	3.1 Introduction . . . . .	12
<b>43</b>	3.2 Existing effort on Data Formatting for Silicon-based tracking trigger . . . . .	12
<b>44</b>	3.3 Existing Effort on Associative Memory developments . . . . .	13
<b>45</b>	<b>4 Schedule, Milestones, and Resources</b>	<b>14</b>
<b>46</b>	4.1 Overall Schedule . . . . .	14
<b>47</b>	4.2 Milestones . . . . .	15
<b>48</b>	4.3 Facilities, Equipment, and Other Resources . . . . .	15
<b>49</b>	<b>5 Budget for FY14 and Preliminary Budget for FY15</b>	<b>16</b>

## 51 1 R&D Overview

52 As outlined recently by the European Strategy reports and the CERN management, continued LHC running through  
53 2020s is the top priority for the particle physics community. While the discovery of the Higgs boson is a major  
54 achievement, many questions about particle physics remain, in particular precise determination of the mechanism  
55 of electroweak symmetry breaking and understanding the nature of dark matter. In order to maximize the potential  
56 for the discovery, CMS must preserve or improve its ability to identify, in real time, events with signatures con-  
57 sistent with the Higgs boson and new particle decays. Identification of the signatures in real time at early trigger  
58 stage will suffer greatly from high pile-up conditions anticipated in the HL-LHC era.

59 At LHC, the only major detector not used at L1 trigger are tracking detectors. As we have learned from current  
60 HLT, tracking provides a highly effective handle on reducing pileup effects. Therefore, development of the L1  
61 tracking trigger system is of utmost importance for CMS for HL-LHC, in order to maintain physics acceptances  
62 for basic objects (leptons, photons, jets and MET). Without L1 tracking, most quantities traditionally used for  
63 triggering are washed out and become unusable due to the huge "background" created by the overlap of too many  
64 collisions in the same beam crossing, putting most of the anticipated physics program out of reach.

65 Consequently, the design of the Phase-II CMS Tracker must allow for an effective implementation of the tracking  
66 trigger. Since the construction of the Phase-II Tracker will take many years, its design must be finalized soon. A  
67 silicon-based L1 tracking trigger has never been realized at this scale and thus it is imperative that its feasibility  
68 be demonstrated before the design of the Phase-II Tracker can be finalized. Silicon-based Level-2 tracking trigger  
69 systems based on associative memory approach were successfully implemented in the past [1] [2] [3] and are  
70 being actively explored at present [4] [5]. Experience with these systems will serve as useful input to the design  
71 of the CMS L1 tracking trigger. However the higher occupancies anticipated at the HL-LHC and the low latencies  
72 required at L1 (about  $10 \mu\text{s}$  total and a few  $\mu\text{s}$  for track finding stage) present us with a formidable set of challenges  
73 that we need to attack with a well organized R&D campaign in oder to have good chances of success. As such,  
74 participation of CMS institutions with strong expertise in modern high-speed electronics and pattern recognition  
75 technologies is crucial for the success of this important R&D.

76 The short time available for the L1 tracking trigger demonstration, as the Phase-II tracker design has to be finalized  
77 in a few years, poses additional challenges as it does not allow sufficient time for starting R&D efforts from scratch.  
78 Fortunately, a number of USCMS institutions lead by Fermilab have established a strong generic R&D program  
79 in the area of silicon based tracking trigger. This R&D program, funded mostly by non-CMS sources over the  
80 past few years, has so far yielded excellent prototype results and put USCMS in the unique position to develop a  
81 working solution for the CMS L1 track trigger. The long-term goal of this R&D effort is to develop these critical  
82 technologies to the point where we can ultimately propose them as a viable solution to the problems of HL-LHC  
83 L1 track triggering. Given the progress made by this R&D program in the last few years, it is now time to take  
84 the next important step and establish a Vertical Slice Demonstration System. In fact, we have recently proposed a  
85 new architecture and system demonstration design [6] to the CMS Phase-II Tracker community and has been well  
86 received. This system will comprise a full tracking trigger path and will be used with simulated high-luminosity  
87 data to measure trigger latency and efficiency, to study overall system performance and to identify appropriate  
88 solutions to possible bottlenecks.

89 Processing each beam crossing implies finding and fitting thousands of tracks starting from a collection of Pt  
90 "stubs" (hit pairs). We need to process 40 million beam crossing per second with a maximum latency of order  
91 of a few microseconds. The total raw computation power needed to solve this problem is huge, several orders  
92 of magnitude larger than that has ever been done in the past. We obviously need to resort to massive parallelism  
93 and we choose to process in parallel different crossings coming at different times (time multiplexing) and different  
94 regions of the detector for the same crossing (regional multiplexing). For this purpose we divide the detector into  
95 48 angular regions (6 in eta times 8 in phi) we call "towers". We assign multiple processing engines to each tower  
96 so that data from that tower and from different crossings may be processed in parallel. In such a parallel system,  
97 a significant problem we need to address and solve is how to dispatch the right data to the right processors. Data  
98 from the same crossing, coming from different detector elements, must be assembled and delivered to the same  
99 processing unit for track reconstruction. Data from different crossings, coming from the same detector element,  
100 must be delivered to different processing units for optimal time multiplexing. The subdivision of the detector into  
101 geographical towers, does not lead to an exact corresponding subdivision of the track parameter space. Data coming  
102 from a given geographical tower may need to be delivered to multiple parameter space regions. This happens, in  
103 particular, when a stub comes from a detector element close to the border between geographical towers, due to  
104 the finite curvature of charged particles in the magnetic field and finite size of the beam luminous region along  
105 the beam axis. In additon to the complex data dispatching challenge, there is the obvious challenge of finding and

106 fitting hundreds of billions of tracks every second which requires extremely fast pattern recognition algorithms.

107 The Associative Memory [1] uses a massively parallel architecture to tackle the intrinsically complex combinatorics of track finding algorithms, avoid the typical power law dependance of the execution time on occupancy, and  
108 solve the pattern recognition in a time roughly proportional to the number of hits. This is of crucial importance  
109 to be able to deal with the large occupancy fluctuations typical of hadronic collisions. However, the design of an  
110 Associative Memory system capable of dealing with the much higher complexity of the HL-LHC collisions, and  
111 with the much shorter latency required by Level 1 triggering, poses significant, still unsolved, technical challenges.  
112 While we have a very aggressive R&D program at Fermilab to advance the state-of-the-art associative memory  
113 technology (the 3D VIPRAM [12] R&D is funded by DOE CDRD program [13]), we are open to possible new  
114 alternative approaches. Since the Associative Memory approach is so far the only proven solution to tracking trig-  
115 gers in a hadron collider environment, it is chosen as the baseline in what follows, but we will make sure that the  
116 architecture we are proposing lend itself to testing and comparing other possible alternatives.  
117

118 For the reasons above, the design of the overall architecture is focused on the need for efficient dispatching of the  
119 data for time and regional multiplexing and on the capability of providing a common flexible framework to test  
120 different possible solutions for track finding and fitting. Since the efficient data dispatching for time and regional  
121 multiplexing requires high bandwidth, low latency, and flexible real time communication among processing nodes,  
122 a full mesh backplane based hardware platform is a natural fit. A custom full mesh enabled ATCA board called  
123 Pulsar II has been designed at Fermilab with the goal of creating a scalable architecture abundant in flexible, non-  
124 blocking, high bandwidth board-to-board communication channels. The Pulsar II hardware will be the workhorse  
125 for the vertical slice demonstration. Most of the hardware design work for the Puslar II has been done [9], including  
126 prototyping work [10] [11], and the main work left for the vertical slice demonstration will be firmware for FY14  
127 and FY15. In addition, a pattern recogniton mezzanine card will need to be designed in FY14, and this will be the  
128 pattern recogniton engine and can host the new associative memory chips being developed at Fermilab.

129 The full-mesh ATCA architecture we are proposing for the CMS L1 tracking trigger demonstration permits high  
130 bandwidth inter-board communication. The full-mesh backplane is used to time-multiplex the high volume of in-  
131 coming data in such a way that I/O demands are manageable at the board and chip level. The resulting architecture  
132 is scalable, flexible and will enable us to provide an early technical demonstration using existing technology. The  
133 ATCA architecture will allow us to explore and compare various pattern recognition architectures and algorithms  
134 within the same hardware platform. Given that Advanced Mezzanine Card (AMC) specifications are designed to  
135 work with both ATCA and microTCA, the architecture naturally allows for the long-term integration of Tracker  
136 DAQ (AMC based) and tracking trigger activities.

137 The purpose of this R&D proposal is to build a vertical slice demonstration test bench to demonstrate the track  
138 finding and to identify possible bottlenecks and find solutions, using the technology that is available today. This  
139 "Vertical Slice" will process simulated data with HL-LHC occupancy at full speed, to allow us to study and  
140 improve the performance (such as latency, efficiency and fake rate). Based on past experience, we believe that the  
141 final technology choices for the final implementation of L1 track finding can be delayed until about four years from  
142 the start of commissioning. The goal of this R&D is to demonstrate that the track finding can be done so that the  
143 tracker design can be finalized in the near future. Once we can demonstrate a functional vertical slice with today's  
144 technology, Moore's law and the semiconductor industry will only work in our favour toward the HL-LHC era.

145 The architecture of the system is described first in the following section. It is scalable, flexible and, although  
146 not meant to be what we will actually implement in the final system, will enable us to provide an early technical  
147 demonstration of the feasibility of a L1 tracking trigger for CMS. We will then describe an affordable demo system  
148 which can be designed and built within 2-3 years from now using current state of the art technology. We will define  
149 the "Vertical Slice" that we plan to build and test as the deliverable of this R&D project. Toward the end, we will  
150 also describe the track finding approach we are pursuing, the advantages and challenges, the work to be done in  
151 the coming two years, and the funding request.

152 This USCMS R&D project, once funded, will help our community focus the attention on the real issues, compare  
153 different possible solutions to the fundamental pattern recognition and track fitting problems at HL-LHC era, and  
154 gain the necessary experience to move, in due time, toward the design of the final system.

155 **2 Technical Description and Deliverables**

156 **2.1 Introduction**

157 This section describes the architecture we propose for the system we want to build to demonstrate the feasibility  
158 of track reconstruction at L1 for CMS in the HL-LHC era. The proposed system is by no means meant to be final  
159 but serves the purpose of an existence proof while helping us identify bottlenecks and unanticipated possible road  
160 blocks.

161 Current estimates show that only a few microseconds will be available for track finding and fitting at Level 1. This  
162 includes data dispatching for trigger tower formation, pattern recognition, track fitting and any necessary further  
163 processing such as, for example, vertexing. Data dispatching is where the stubs from many thousands silicon  
164 modules must be organized and delivered to the appropriate eta-phi trigger towers. Due to the finite size of the  
165 beam's luminous region in z and the finite curvature of charged particles in the magnetic field, some stubs must be  
166 duplicated and sent to multiple towers in an intelligent way. Since all this must be done within a very short time (of  
167 the order of a micro-second), communication between processing elements in different towers requires very high  
168 bandwidth and very low latency. In addition, extremely fast and effective track fitting is also required. Extensive  
169 R&D and experimentation of innovative ideas is obviously needed in this area.

170 The architecture we propose is based on ATCA with full-mesh backplane. The large inter-board communication  
171 bandwidth provided by the full-mesh backplane is used to time multiplex the high volume of incoming data in such  
172 a way that the I/O bandwidth demands are manageable at the board and chip level, making it possible for an early  
173 technical demonstration with existing technology. The resulting architecture is scalable, flexible and open. For  
174 example, it allows different pattern recognition architectures and algorithms to be explored and compared within  
175 the same platform. Also, given that AMC specifications are designed to work with both ATCA and MicroTCA, this  
176 architecture allows a natural long term integration of TK-DAQ (AMC card based) and TK-TRIG (ATCA based).

177 Track reconstruction typically consists of two steps: pattern recognition followed by track fitting. Pattern recogni-  
178 tion involves choosing, among all the hits present in the detector, those hits that were potentially caused by the  
179 same particle. This stage produces a set of hits of interest. Track fitting involves extracting track parameters from  
180 the coordinates of the hits of interest. When time constraints are not so stringent, track reconstruction is imple-  
181 mented in software, often using processors running in the upper levels of a data acquisition system. However,  
182 software algorithms running on standard CPUs are typically not fast enough for low level triggers.

183 Hardware-based pattern recognition for fast silicon-based triggering on charged tracks was first developed for the  
184 CDF Silicon Vertex Trigger (SVT) at the Fermilab Tevatron in the 1990's. The method used there [1] was based  
185 on a massively parallel architecture - the Associative Memory - to efficiently identify patterns at high speed, and  
186 has provided an effective solution to fast track triggers in a hadron collider environment. It was successfully used  
187 in CDF in Run II at trigger Level 2 enabling a large number of physics results over more than 10 years. The  
188 same approach is now being implemented for Atlas (FTK), also at Level 2, albeit with a much improved hardware  
189 architecture implemented with modern technology.

190 The Associative Memory solves the combinatorial problem, inherent to this kind of pattern recognition algorithms,  
191 by employing a massively parallel architecture to compare each detector hit to a large number of pre-calculated  
192 geometrical patterns simultaneously. Then, the selected patterns are processed using fast FPGAs to perform track  
193 fitting. Since each pattern corresponds to a very narrow "road" through the detector, the usual helical fit is much  
194 simplified and fast by using a pre-calculated set of parameter values for the center of the road and applying correc-  
195 tions that are a linear function of the actual hit positions in each layer.

196 The Associative Memory (AM) uses a massively parallel architecture to tackle the intrinsically complex combina-  
197 torics of track finding algorithms, avoid the typical power law dependence of the execution time on occupancy, and  
198 solve the pattern recognition in a time roughly proportional to the number of hits. This is of crucial importance to  
199 be able to deal with the large occupancy fluctuations typical of hadronic collisions. The time it takes for the AM to  
200 solve the pattern recognition problem is virtually zero. As soon as all the hits have been stored in the AM, found  
201 tracks are ready to be output. The whole latency incurred is the time needed to load the hits plus the time to read  
202 the matched patterns. In this sense the AM is hard to beat for this particular task.

203 add here track fitting part.

204 For these reasons, although the system we propose is designed to allow the experimentation of different solutions  
205 to fast track finding and fitting, we consider the Associative Memory approach as the baseline against which other  
206 methods will be compared. However, the design of an Associative Memory system capable of dealing with the

207 much higher complexity of the HL-LHC collisions, with an input event rate several orders of magnitude larger  
208 than ever done before, and with the much shorter latency required by Level 1 triggering, poses significant, still  
209 unsolved, technical challenges. A significant progress is needed both in the pattern density and in the processing  
210 speed of AM chips. An aggressive R&D program is needed and is currently ongoing at Fermilab.

211 The proposed architecture and system demonstration concept has been well received within the tracker Phase 2  
212 upgrade community and work is in progress to better define the concept. Clearly, establishing international collabora-  
213 tions within CMS to work on this project is essential. Both INFN/Italy and Lyon/France have joined us to work  
214 on the vertical slice demonstration using associative memory approach, while others are interested in contributing  
215 or exploring possible new track finding algorithms on the same hardware platform. For example, INFN Italy has  
216 been working together with the Atlas FTK team exploring the possibility of using FTK associative memory chips  
217 for CMS L1 tracking trigger demonstration, to learn as much as we can using the FTK AMchips developed for  
218 Level 2 applications. INFN will work closely together with Fermilab to develop the pattern recognition mezzanine  
219 for Pulsar II board to host the FTK AMchips (another mezzanine will be designed specifically for the associative  
220 memory chips being developed at Fermilab). Experience with using the FTK AMchips will be useful to guide the  
221 associative memory chip design at Fermilab dedicated for CMS L1 application. Development of the simulations  
222 for the Associative Memory approach in CMS was started by the Pisa and Lyon groups with significant progress  
223 made towards making the machinery work in the standalone mode. Now, together with Padova, they are migrat-  
224 ing the tool into CMSSW. Fermilab/LPC have been working closely with them on the simulation efforts. Lyon  
225 group has been also investigating the possibility of using a new algorithm for track fitting stage, based on Hough's  
226 transforms, to replace the conventional SVT-style, linearized track fitting algorithm. This new algorithm will be  
227 combined with the associative memory stage.

228 One of the main activities in the coming year in USCMS (FY14) will consist of extensive simulation efforts, by  
229 physicists, to establish technical specifications based on Phase 2 physics goals. Due to the intrinsic massive parallel  
230 processing hardware nature of the AM operation, there is also a clear challenge in using software based simulation  
231 to emulate the hardware performance. Massive parallel software based processing technology, such as GPUs, will  
232 be explored to significantly speed up the simulation process, leveraging our experiences in GPU R&D over the past  
233 few years [7] [8]. At the same time, students and postdocs from all groups will be offered a unique opportunity to  
234 develop hardware experience by getting involved with the design, construction and commissioning of the vertical  
235 slice demonstration over the next few years. Some of the groups, for the longer term, are also interested in getting  
236 involved with the development of the algorithms for the post-track-finding stages and of the interfaces with the  
237 global trigger.

## 238 **2.2 System Overview**

239 Many unique challenges must be faced at the different stages of the processing chain: first, data need to be trans-  
240 ferred out of the tracker at the necessary speed, stubs from thousands of silicon modules must be formatted,  
241 organized into  $\eta - \phi$  trigger towers, duplicated and shared across tower boundaries as needed, then we need to  
242 perform pattern recognition and track fitting, and finally process all the tracks reconstructed by the previous stages  
243 to form an intelligent trigger decision. A coherent system design for a Level-1 track trigger will include all these  
244 aspects.

245 In this document we will make the working assumption that there will be a total of 15K detector modules/fibers,  
246 each fiber with 3.25 Gbps payload bandwidth capability. The detector will be partitioned into 48 trigger towers,  
247 6 in  $\eta$  and 8 in  $\phi$ . Each trigger tower will therefore handle 312 modules/fibers on average. The cabling of the  
248 modules will need to be optimized for trigger requirements. For simplicity, we will assume that the FEDs are  
249 upstream and receive the fibers from the modules and pass the relevant data to the track trigger system even though  
250 the architecture could allow the FED to reside in the same ATCA shelf as the track trigger data input boards on  
251 dedicated AMC ATCA carrier boards. The focus of this document is the Vertical Slice Demonstration System,  
252 not the DAQ readout, so FED details are not discussed here (they belong to TK-DAQ). The FED interface will  
253 need to be defined for demonstration purposes, even though the actual FEDs do not have to be involved in the  
254 demonstration.

## 255 **2.3 Tracker-Trigger interface**

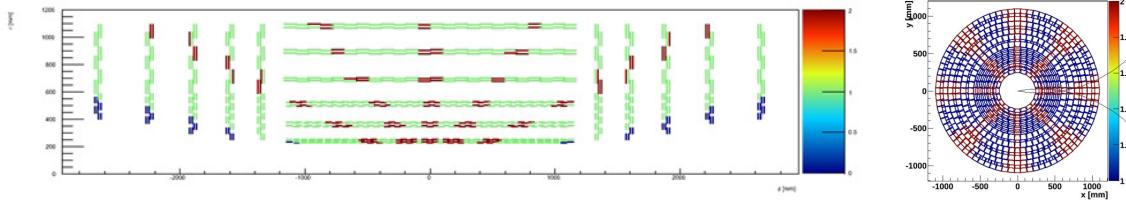
256 The found stubs are sent from the modules using a block synchronous data transfer scheme which tolerates random  
257 occupancy fluctuations while bonding latency. The current plan is to have the data from 8 consecutive beam cross-  
258 ings as one block. The front-end designers are still investigating different format variants for robustness against

259 rate fluctuations, ease of implementation, impact on power consumption, etc. While choosing the 8 crossings  
 260 scheme as our current working assumption, our strategy is to design the downstream components to be flexible  
 261 enough to handle different possible formats.

## 262 **2.4 Track Trigger System Architecture**

### 263 **2.4.1 Tracker geometry and Trigger Towers**

264 Detailed studies have been done for the BE tracker geometry with different trigger tower partitions, and the 6 (in  
 265  $\eta$ )  $\times$  8 (in  $\phi$ ) = 48 trigger tower partition has been chosen as the default baseline configuration (see Figures 1).



266 Figure 1: Six sectors in  $\eta$  (left). Note that the symmetry around  $\eta = 0$  will provide for easier cable grouping. Eight  
 267 sectors in phi (right).

268 Stubs from the 15K silicon modules must be delivered to the correct trigger towers. Detailed studies have been  
 269 performed on data sharing assuming the default 48 tower partition with a minimum  $p_T$  of 2 GeV and track origin  
 270 smearing in  $z \pm 7$  cm. Figure 2 shows the number of trigger towers that stubs from a given module must be  
 271 delivered to under these conditions. When a stub is in the middle of the trigger tower, it will have to be delivered  
 272 to only one tower (to the native trigger tower). When a stub is at the boundary in phi or eta (but not both), it will  
 273 have to be delivered to two towers. If a stub is at both the boundaries in eta and phi, it will have to be delivered to  
 274 four towers. Note that four towers is the maximum number of towers any stub must be delivered to.

### 275 **2.4.2 Formation of Trigger Towers**

276 The subdivision of the tracker into 48 trigger towers is shown in Figure 2 (right), where the colored lines indicate  
 277 all needed interconnections among the trigger towers. The unique feature of this arrangement is that any given  
 278 trigger tower only needs to be connected and share stubs with its immediate eight neighbors and detailed studies  
 279 show that this feature is more or less independent from the minimum  $p_T$  threshold and track origin smearing in  $z$   
 requirements. This inter-connection structure will be used as the basis of the proposed trigger system architecture.

### 280 **2.4.3 System architecture**

281 The tower processor platform must support large numbers of fiber transceivers, which are used for receiving input  
 282 links and sharing data between neighboring towers. A flexible, high bandwidth backplane is also required to  
 283 quickly transfer data between boards. The boards should be large enough to support pattern recognition engines  
 284 and fiber connections. Given these requirements, we conclude that a full mesh 14 slot ATCA shelf is a natural  
 285 fit for the tower processor. An ATCA shelf is typically an air-cooled 13U rack mounted chassis consisting of  
 286 14 slots. The first two slots are reserved for Ethernet switch blades. Switch blades may include a fast CPU and  
 287 are often used for controls and other system functions. The remaining 12 slots are used for processor or payload  
 288 blades. In a full mesh ATCA backplane each pair of slots is directly connected with a multi-lane bidirectional  
 289 serial channel capable of supporting sustained 40 Gbps data transfers. A modern "40G" full mesh ATCA shelf has  
 a total aggregate bandwidth of over 7 Tbps, not including external I/O.

290 For simplicity and illustration purpose, let's simply assume one ATCA shelf per trigger tower for the moment as a  
 291 starting point. Following this assumption, if we were building the L1 Tracking Trigger system today using existing  
 292 technology, we could propose a system comprised of 48 ATCA shelves with possibly an additional shelf acting as  
 293 a second stage processor, as shown in Figs. 3 and 4. Of course, the actual system will most likely be smaller. Note  
 294 that connections between tower processor shelves are limited to eight nearest neighbors, and this can be easily  
 295 achieved.

296 The generic processor blade concept is shown in Figure 5 (left). The front board measures 8U x 280mm and is  
 297 designed around a single FPGA. This FPGA connects directly to the full mesh backplane fabric, mezzanine cards,

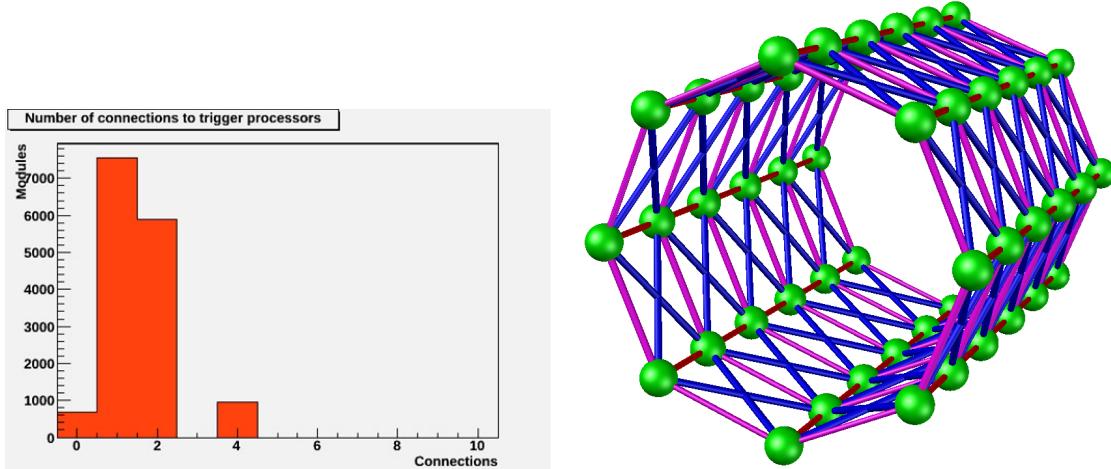


Figure 2: Left: Distribution of the number of trigger towers each module needs to be connected to. Entries at zero are from modules that do not participate in triggering. Right: Conceptual view of the proposed CMS phase II L1 tracking trigger towers. The formation is organized as 48 trigger towers ( $6 \eta \times 8 \phi$ ). Because the phase II tracker is being designed for tracking trigger purposes, it is possible to arrange the towers in such a way that data sharing only requires communication with immediate neighbor towers. Each node in this diagram represents a trigger tower processor engine. Within each processor engine crate the full mesh backplane is used for time multiplexing of the incoming data, while the data sharing between towers is handled with inter-crate fiber links.

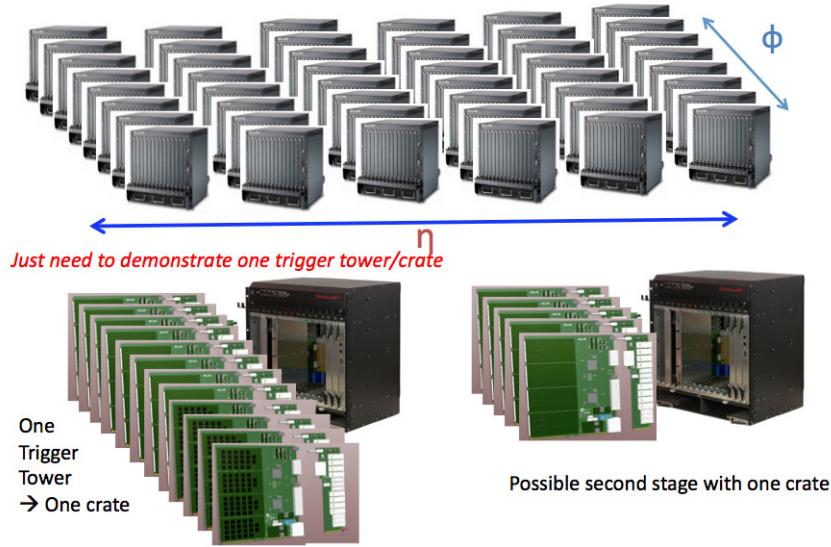


Figure 3: Possible system configuration with today's technology by simply assuming one ATCA shelf per trigger tower (will be smaller in the future)

298 and fiber transceivers located on a rear transition module (RTM). For the most part communication channels are  
 299 high speed serial point to point links and are directly supported by SERDES transceivers in the FPGA. The actual  
 300 design of Pulsar IIb is also shown in Figure 5 (right).

301 The fundamental processing element is a pattern recognition mezzanine (PRM) card shown on Fig. 6 which per-  
 302 forms both track finding and fitting. Time multiplexed data transfers into several parallel PRMs can reduce band-  
 303 width requirement to manageable level. PRM's using different approaches to track finding and fitting may be tested  
 304 and compared within the same overall high-level system architecture and data dispatching scheme.

#### 305 2.4.4 Architecture Flexibility

306 A major advantage of the full mesh backplane is that it effectively blurs the distinction between boards, thus  
 307 enabling system architects to experiment with different shelf configurations. In the following sections we briefly

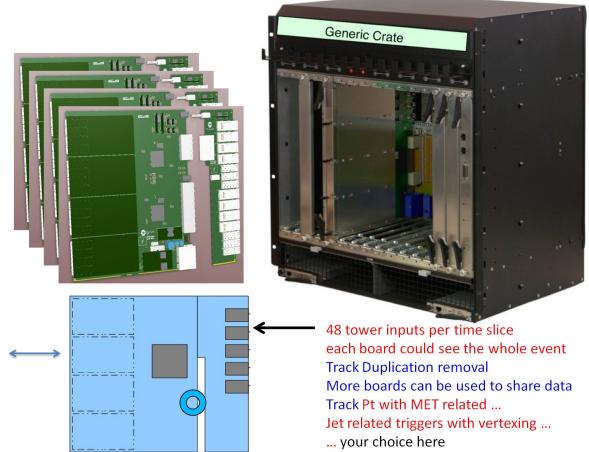


Figure 4: The second processing stage shelf.

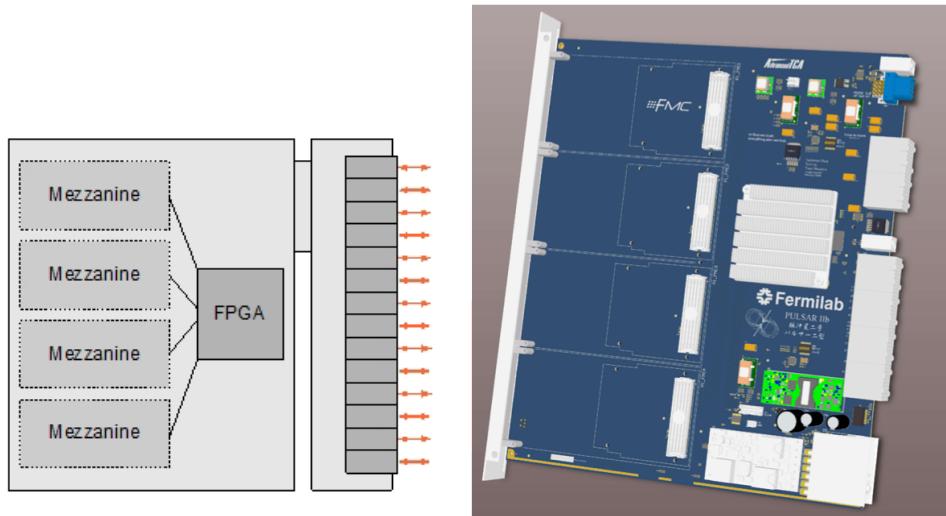


Figure 5: Generic processor blade concept (left), and the actual design of Pulsar IIb [9] at final layout stage (right)

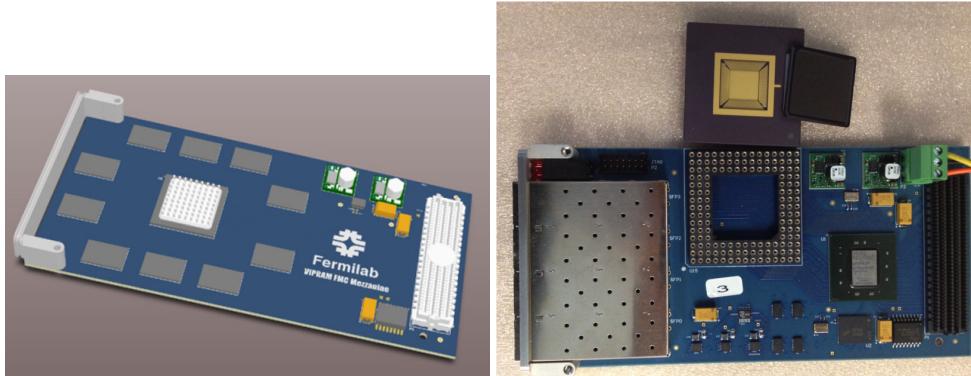


Figure 6: Left: Concept of a pattern recognition mezzanine design for testing different pattern recognition algorithms. Right: a test mezzanine prototype designed and built at Fermilab, features four SFP+ pluggable serial transceivers (for standalone data receiving), a Kintex 7 FPGA, configuration flash memory, DDR3 memory, power supplies, local oscillators, a test socket for associative memory chips developed at Fermilab, and FMC connectors.

308 illustrate two kinds of tower processor systems made possible by the flexibility of the full mesh architecture.

309 **N DIB and M PRM configuration ( $N + M \leq 12$ )** The most straightforward tower processor architecture  
 310 consists of N data input boards (DIB), which receive input links and perform zero suppression. A DIB may be

311 built using the generic ATCA processor blade (Figure 5) if the data is coming from FEDs or directly from the  
 312 detector modules. It is also possible to use a generic ATCA carrier board and several FED AMC mezzanines  
 313 directly if FED AMC card will include the DIB functionality (to pass the data for L1 track trigger to PRBs). After  
 314 zero suppression, the N DIBs transfer the event data to M number of pattern recognition boards (PRB), which  
 315 contain Mx4 pattern recognition mezzanine (PRM) cards. Data transfers from the DIBs to the PRMs are time  
 316 multiplexed, thereby the bandwidth requirements can be significantly relaxed.

317 Data entering the PRB can be time multiplexed again and transferred to the four PRMs to further reduce bandwidth  
 318 requirements and allow for longer processing times. The full mesh backplane fabric supports any variant of these  
 319 configuration (assuming that  $N + M \leq 12$ ), and different variant may have different demands on hardware.  
 320 Example variations are sketched on Fig.7 and bandwidth requirements for the worst case scenario (assuming 500  
 321 stubs per event per trigger tower) are summarized in Table 1. Note that current study show that on average,  
 322 we expect only about 100 to 200 stubs per trigger tower per beam crossing, here we assumed 500 stubs to be  
 323 conservative.

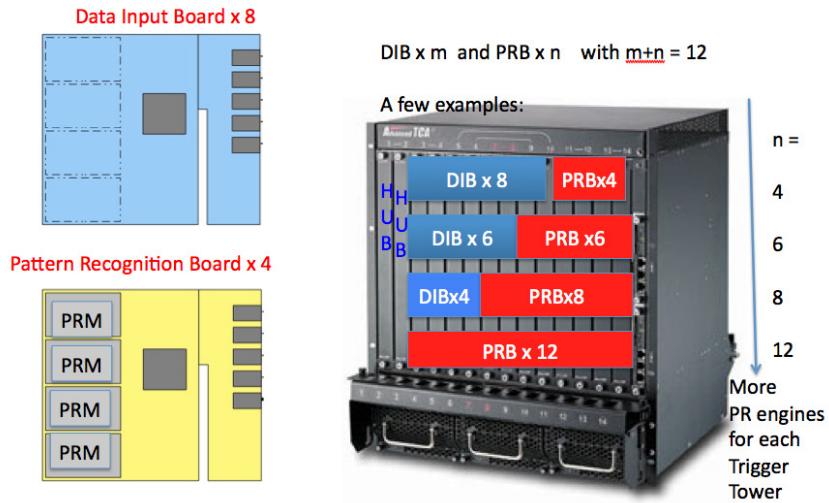


Figure 7: System flexibility: many configurations possible & being studied to select the right one for demonstration purpose

DIB/PRB/PRM Count	Fabric Channel BW (minimum)	PRM Input BW (minimum)
8/4/16	20 Gbps	40 Gbps
6/6/24	20 Gbps	27 Gbps
4/8/32	20 Gbps	20 Gbps

Table 1: Data sharing between towers occurs on the PRB board level. Each PRB connects to the corresponding PRB in the eight nearest tower processor shelves. The above numbers assume a worst case scenario of 500 32-bit stubs per trigger tower per event (every 25 ns). An example of special configuration with eight DIBs and four PRBs will be used as a simple example in Section 3.

324 **DIB/PRB combo configuration** In the limit of  $N=0$  and  $M=12$  from the "N DIB and M PRB" configuration, the  
 325 DIB and PRB functionalities can be combined into one blade design.

326 A tower shelf would then consist of 10 Processor blades, one Gateway blade (for data sharing), and one Collector  
 327 blade (for tracks found). These three different blade functionalities can be implemented in the same hardware.  
 328 Backplane transfers are described in a series of fully pipelined sequences shown in Figure 8.

329 This processor architecture uses every channel in the full mesh backplane. By using the full mesh fabric more  
 330 effectively we are able to decrease the channel bandwidth requirement from 20 Gbps down to 6 Gbps with no  
 331 significant latency increase.

## 332 2.5 Vertical Slice Demonstrator System: Overview and Methodology

333 The flexible architecture described above lends itself to an early technical demonstration of the system. The main  
 334 goal of the demonstration system is to identify possible problems in the architecture design and, hopefully, find

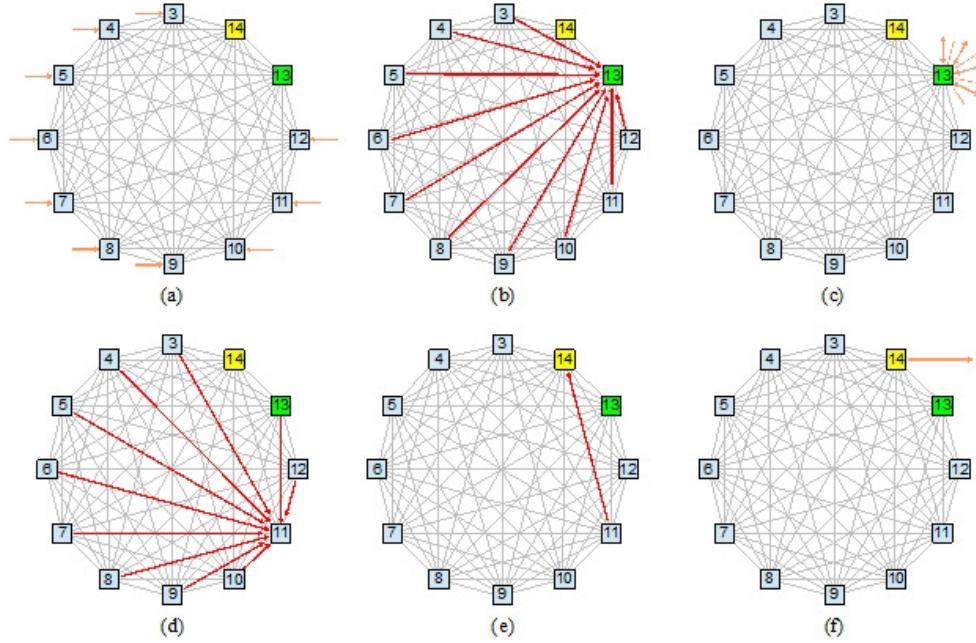


Figure 8: Backplane transfers sequences using the combined blade design. First, the input fibers are received on the Processor blades (a). Each Processor blade then transfers a portion of the input data to the Gateway blade (b), where it is exchanged with neighboring towers over fiber links (c). The Processor blades and Gateway blade transfer the event (including neighbor data) to the target Processor blade in a time multiplexed, round robin scheme (d). Results from the Processor blades are then transferred to the Collector blade (e) for any final formatting and processing before transmission downstream (f).

335 solutions. We would study, measure and optimize trigger latency and efficiencies at different stages of the system  
 336 using hardware prototypes being developed. This will involve extensive simulation work, to guide the hardware  
 337 implementation and to compare actual measurements with expectations. A possible Vertical Slice Demonstration  
 338 System is shown in Figure 9. Each stage is described in more detail in the following sections.

339 Although the architecture is flexible enough to allow for different configurations, for the sake of clarity and sim-  
 340 plicity, in what follows we will often use the specific configuration with eight Data Input Boards and four Pattern  
 341 Recognition Boards as an example. We will decide only at a later time which specific configuration we will actually  
 342 use for the demonstration system.

343 This demo system will be implemented in stages, at mezzanine level, board level, crate level and multi crate level.  
 344 These different stages would naturally proceed in sequence, from the bottom up. This way, we will have the  
 345 opportunity to learn along the way about the performance of the different components of the system before having  
 346 to decide exactly how the whole thing will be cabled up. Also, the extra crate, with three neighbor towers, will  
 347 need to come into play only at a very advanced stage, towards the end when the system dynamics need to be  
 348 demonstrated.

## 349 2.6 Data Source Stage

350 The Data Source mimics the data flow out of the upgraded Phase II outer-tracker running at the HL-LHC. It  
 351 will drive 300+ fibers (one/module) to the trigger tower under study exactly as the data were coming from the  
 352 real detector at high luminosity and full speed. Each fiber connection will transmit data at 3.25 Gbps payload  
 353 bandwidth, in the same way as the actual modules in the future real system. The data will be derived from  
 354 simulation, appropriately formatted, stored into on-board memories, and then played back at full speed. The  
 355 Pulsar IIb board can be used for Data Source stage.

## 356 2.7 Data Input

357 The Data Input Blade (DIB) is responsible for receiving data from the upstream detector electronics (or Data  
 358 Source output) and transferring them to the PRBs. Up to about 40 fiber links will be received by each DIB. These

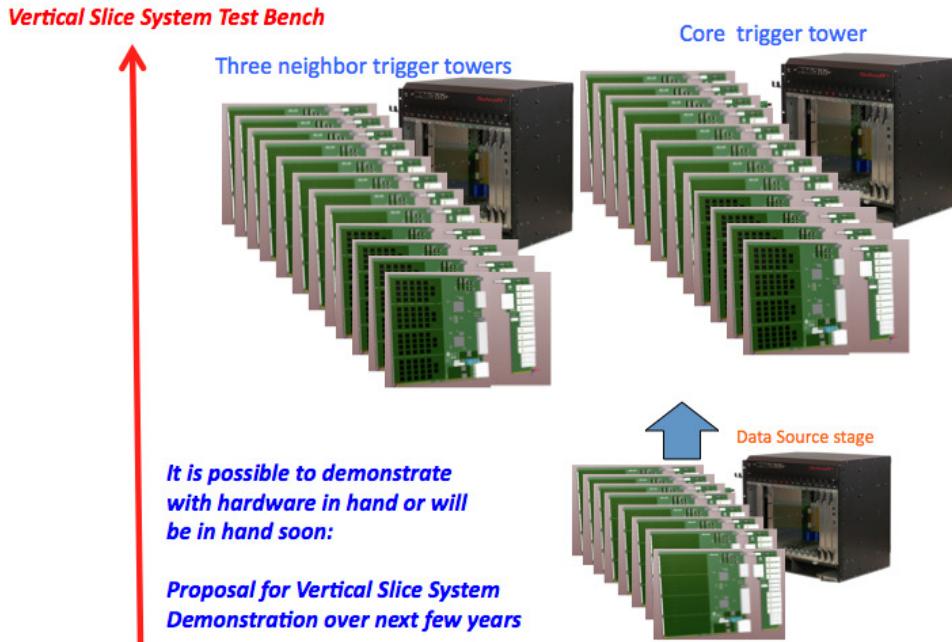


Figure 9: Vertical slice test bench principle.

359 input links may terminate on the RTM or mezzanine cards. Input fiber links are nominally 3.25 Gb/s payload  
 360 bandwidth. Again, the Pulsar IIb board can be used as DIB. The Data Input Board will perform zero suppression,  
 361 pack the stubs into a new format and send them to the PRBs. Current estimates indicate that a rate of about 200  
 362 stubs per event per trigger tower, which yields a data rate of roughly 256 Gb/s (200 stubs\*32 bits/25 ns) entering  
 363 each trigger tower on average.

364 As an example, in the configuration with eight DIBs and four PRBs, each DIB will be receiving an average of  
 365 about  $256/8 = 32$  Gb/s of stub data (after zero suppression). Each of the eight DIBs in the shelf sends data to four  
 366 PRBs in a round-robin, time multiplexed fashion. Since data is sent to four PRBs, these transfers can take place at  
 367 a quarter of the input rate, or  $32/4 = 8$  Gb/s assuming 200 stubs per trigger tower per event.

368 In Figure 9, the ATCA shelf devoted to the "core" trigger tower is shown equipped with 8 DIB boards and 4 PRB  
 369 boards while the shelf devoted to the "neighbor" towers is equipped with 12 PRB boards, 4 PRB boards for each  
 370 of the three neighbor towers. In general, each tower needs to share data with 8 neighbors but 3 are sufficient in the  
 371 demonstration system to test all possible data sharing cases (eta, phi and "diagonal"). Simulated data corresponding  
 372 to three neighbor towers are delivered from PRB boards in the "neighbor" shelves to the corresponding PRB boards  
 373 in the "core" shelf.

## 374 2.8 Pattern Recognition Board

375 Using again the special configuration above (8 DIB + 4 PRB), each PRB will be receiving 64 Gb/s stub data on  
 376 average. While receiving the data and sending them to the mezzanine cards, each PRB will exchange data with  
 377 the corresponding PRB, processing the same time slice, in the neighboring tower for data sharing in the overlap  
 378 regions. In this case, each PRB can use four 40Gb/s links (QSFP) for the connections in the eta and phi directions,  
 379 and four 10Gb/s links (SFP+) can be used for data sharing in the "diagonal" directions. The PRB FPGA drives  
 380 data received from the DIBs to the Pattern Recognition Mezzanine (PRM) boards. This can also be done in a 4x  
 381 time multiplexed fashion. The 4x time multiplexed transfers from the PRB FPGA to the PRM would require a  
 382 bandwidth of about 16Gb/s this way. Again, the Pulsar IIb board can be used as PRB.

## 383 2.9 Pattern Recognition Mezzanine Card

384 Each PRB supports four Pattern Recognition Mezzanine (PRM) boards. These boards are based on the FMC  
 385 standard and support high speed LVDS and SERDES connections to the PRB FPGA. In one possible incarnation,  
 386 each PRM will contain an FPGA, on board memory to act as Data Buffer, and an array of pattern recognition  
 387 devices. In our example configuration, we need to support 16 Gb/s between the PRB FPGA and the PRM. The

388 FPGA-PRAM (associative memory) channel bandwidth needs will be a fraction of the PRM input bandwidth,  
389 because only relevant stubs will be sent to the relevant pattern recognition chip covering the relevant regions of the  
390 trigger tower.

391 All track fitting algorithms can be implemented in FPGA on the PRMs, therefore they can be studied and compared  
392 directly using the same vertical slice demonstration setup. Generally speaking, PRM's using different approaches  
393 to track finding and fitting can be tested and compared within the same overall high-level system architecture and  
394 data dispatching scheme. The track fitting can occur on the PRM FPGA. The traditional CDF SVT/FTK-style  
395 track fitting stage [16] can be used to benchmark the performance of this stage.

## 396 **3 Relation to Existing Efforts**

### 397 **3.1 Introduction**

398 Fermilab has had a focused R&D program in developing hardware-based technology that advances the state-of-  
399 the-art for pattern recognition and track reconstruction for fast triggering. Specifically, we have been addressing the  
400 most important challenges for tracking trigger by developing new full-mesh ATCA based Data Formatting system  
401 as well as new Associative Memory technology using both conventional 2D and 3-dimensional (3D) fabrication  
402 technology. The 3D approach to associative memory implementation is an important component because adding a  
403 third dimension opens up the possibility for new architectures that could dramatically enhance pattern recognition  
404 capability.

405 The long-term goal of the R&D program is to develop the critical technologies to the point where we can ultimately  
406 propose it as a viable solution for track triggering for CMS in Phase 2 operations. The first step is to design and  
407 build prototypes for the ATCA based Data Formatting hardware as well as the associative memory chips. Recently  
408 we have finished the first round of prototype design for both projects, and we have successfully tested the first Data  
409 Formatting hardware (ATCA motherboard, nicknamed Pulsar-IIa, RTM as well as mezzanine cards) and they all  
410 work well, in fact better than expectations [10]. Soon we will test the first associative memory 2D prototype chips  
411 early 2014. All these work were supported by non-CMS funds from the past.

### 412 **3.2 Existing effort on Data Formatting for Silicon-based tracking trigger**

413 The overall design goal is to create a uniquely scalable architecture abundant in flexible, non-blocking, high band-  
414 width board to board communication channels while keeping the design as simple as possible. Expandability  
415 and scalability are achieved through three mechanisms. First, each board supports up to four mezzanine cards  
416 connected to the main FPGAs. Each mezzanine card may contain FPGAs, pattern recognition ASICs, fiber optic  
417 transceivers, or any other custom hardware. Our mezzanine cards use the FPGA Mezzanine Card (FMC) standard  
418 which has become popular with many third party vendors. Secondly, additional boards may be installed in the  
419 crate. Unlike a shared bus system, adding boards to the mesh network has a minimal impact on the system la-  
420 tency while dramatically increasing system processing power and I/O capability. Lastly, we have reserved several  
421 transceivers on rear transition modules (RTM) for dedicated serial links between boards in different crates.

422 Our first prototype, called the Pulsar IIa, is designed around a pair of FPGAs. These FPGAs feature multiple  
423 high speed serial transceivers which are directly connected to the ATCA full mesh backplane and to pluggable  
424 transceivers on the rear transition module (RTM). The Kintex FPGAs we have selected for Pulsar IIa have 16  
425 10Gb/s GTX serial transceivers so our first prototype boards offer a subset of the full backplane and RTM connec-  
426 tivity. Leveraging the experience we gained through designing, building and testing the Pulsar IIa board we are in  
427 the final stages of laying out the next generation board, the Pulsar IIb. The new board design replaces the two Kin-  
428 tex K325T devices with a single large Virtex-7 FPGA. The GTX transceiver count has increased up to 80 channels,  
429 providing a significant bandwidth increase to the RTM, Fabric and Mezzanine cards. The power regulator sections  
430 of the board have been redesigned to handle the estimated increased power required by the Virtex-7 FPGA. The  
431 Pulsar IIb design work is supported by non-CMS funds.

432 The Data Formatter design was originally motivated by Atlas FTK needs. The performance of the actual Data  
433 Formatter Pulsar II design far exceeds the original FTK requirements. This high performance scalable architecture  
434 may find applications beyond tracking triggers, and can serve as a starting point for future Level-1 silicon based  
435 tracking trigger R&D for CMS. The Pulsar IIb, as it is designed, can be used as the workhorse for the Vertical  
436 Slice Demonstration system for CMS L1 tracking trigger. The main work remaining for CMS tracking trigger

437 demonstration includes the firmware for Pulsar II, and a new mezzanine card design (hardware and firmware) to  
438 host the protoVIPRAM2D associative memory chips (described below).



Figure 10: Left: Pulsar IIa prototype board, together with its mezzanine cards and RTM. Right: Pulsar IIa crate level testing.

### 439 **3.3 Existing Effort on Associative Memory developments**

440 As mentioned earlier, the CDF SVT-style Associative Memory chip (we will call it PRAM, Pattern Recognition  
441 Associative Memory, to emphasize its purpose for HEP) is a departure beyond conventional CAMs. Like conventional  
442 CAMs, PRAMs store address patterns and look for matches between incoming hits and those addresses for a  
443 given detector layer. At this level, the match is expected to be either exact (Binary CAM) or partial (Ternary CAM)  
444 and an array of Match Flags is the typical output. A PRAM has an array of Match Flag Latches which capture and  
445 hold the results of the match until reset for the next event. As the hits from the various layers of the detector for  
446 the same event arrive, the PRAM is looking for matches from one candidate address to one or more stored address  
447 patterns. The PRAM organizes stored address patterns into roads, which are linked arrays of several stored ad-  
448 dress patterns from different detector layers. Each stored address in a road is from a different layer in the detector  
449 system. These linked addresses represent a path that a particle might traverse through the layers of the detector  
450 (hence the name "road"). The ultimate goal of the PRAM is to match real particle trajectories to those roads. Like  
451 a conventional CAM, a PRAM flags a match when a candidate address matches a stored pattern address for a given  
452 detector layer. However, before the PRAM does anything with that match, it must find matches in all (or majority  
453 of) the elements (layers) that constitute a road.

454 It should be emphasized that compared to commercially available CAMs, such as Network Search Engine, the  
455 PRAM has the unique ability to search for correlations among input words received on different clock cycles.  
456 This is essential for tracking trigger applications since the input words are the detector hits arriving from different  
457 layers at different times. They arrive at the chip without any specific timing correlation. Each pattern has to store  
458 each fired layer until the pattern is matched or the event is fully processed. Even in the case of a level-1 trigger  
459 application, which is largely synchronous, this feature will still be important. One unique feature of this approach  
460 is that the pattern recognition of the event is done as soon as the last hit arrives, which makes the approach a  
461 promising candidate for L1 track trigger. However, the requirements for L1 track trigger application will be very  
462 different from that for L2, and the system interface of the chip has to be fully redesigned and the performance has  
463 to be optimized.

464 The PRAM pattern density can be improved by optimizing the design in single-layer chips (2D), using custom  
465 cell designs with smaller feature size technology. There is an R&D effort by INFN using 65 nm technology to  
466 improve design for Atlas FTK application for L2 trigger (AMchip05 or 06). INFN has now in hands a 65 nm  
467 version prototype (Amchip04), developed for FTK purpose which is at Level 2 trigger stage. INFN AM05 has

468 been submitted recently, and the AM06 is expected to be submitted in the Spring 2014. The AMchip05 could be  
469 used for initial testing. Note that the FTK AMchips are not designed or optimized for L1 track trigger applications.

470 As mentioned earlier, there is an on-going R&D effort at Fermilab using both conventional 2D and the emerging 3D  
471 technology to design future generation of PRAM chip [12],[13] specifically for the needs of the L1 CMS tracking  
472 trigger needs (ProtoVIPRAM series). The first 2D prototype chip (ProtoVIPRAM01) is expected to arrive early  
473 2014. The Fermilab VIPRAM R&D project has two goals. The first is the increase in pattern density through the  
474 use of vertical integration and circuit and geometrical (layout) enhancements. This project will continue through  
475 FY14 for proof-of-principle of the 3D VIPRAM concept and is funded by DOE CDRD. The second is the increase  
476 in speed and the improvements in system interface, specifically with regard to Level 1 Tracking Trigger applications  
477 for CMS at HL-LHC, through the use of system, circuit and layout enhancements. The second goal is for CMS  
478 and it is this part of the work that we are requesting funding support from USCMS.

479 From the beginning, our design methodology has been to develop concepts and circuitry in 2D to confirm function-  
480 ality as economically as possible and then translate, where necessary, those ideas into 3D. The first step taken by  
481 the VIPRAM Project was the development of a 2D prototype (protoVIPRAM1) in which the associative memory  
482 building blocks were laid out with an eye toward future vertical integration. In fact, the associative memory build-  
483 ing blocks were laid out as if this was a 3D design. Room was left for as yet non-existent Through Silicon Vias  
484 and routing was performed to avoid these areas. The vertical integration approach taken thus far by the VIPRAM  
485 project reconfigures the pattern recognition algorithm into CAM cells and Control cells each of which ultimately  
486 will be integrated on different 3D Tiers. The readout circuitry is deliberately simplified to allow direct perfor-  
487 mance studies of the CAM and Control cells. protoVIPRAM1 was designed and fabricated in a 130nm Low Power  
488 CMOS process that has been used previously in High-Energy Physics 3D designs. The design has been thoroughly  
489 simulated at all levels and the prototype will be tested in early 2014 both for functionality and performance using  
490 a custom test setup.

491 The protoVIPRAM1 is the first step to develop the next generation AM chips for L1 applications. The next two  
492 steps will be done in parallel and will, in fact, feed off of one another. The protoVIPRAM3D takes the circuitry  
493 designed in protoVIPRAM1 and vertically integrates it. The Control cells are moved onto a Control Tier and  
494 the CAM cells become a CAM Tier. The basic idea here is that given properly designed sub-circuits, vertical  
495 integration is a solution to pattern density limitations. The protoVIPRAM2 for CMS, on the other hand, attempts  
496 to improve the data input and readout speed of the associative memory chips and bring the system-level interface  
497 to maturity using conventional 2D VLSI with an eye towards CMS Level 1 trigger applications. Several of the  
498 ideas created for the protoVIPRAM1, most notably the square layout of the CAM cells and the simplified readout  
499 architecture, will be used as stepping stones for increasing readout speed and flexibility.

## 500 **4 Schedule, Milestones, and Resources**

### 501 **4.1 Overall Schedule**

502 The schedule of the proposed program is driven by the current state of the generic R& D efforts and the CMS  
503 desire to perform the Vertical Slice Demonstration by 2016. In order to meet this aggressive deadline, we have  
504 to finalize design of the hardware in 2014 with year 2015 dedicated mainly to the firmware development, setting  
505 up the demonstration test-stand and performing the actual testing. This is only possible because we have already  
506 done so much tracking trigger R&D over the past few years (mostly with non-CMS funds).

507 Currently the prototype layout of the Pulsar-IIb is finalized and we expect to have the board available for testing  
508 in early Spring 2014. Detailed studies and testing require development of the associated software and firmware  
509 and based on past experience are expected to take at least 6 months. It is not inconceivable that depending on the  
510 results of the testing, another round of prototyping may be required (Pulsar IIc). Final revisions will be made to  
511 the board and the final version will be submitted for production in Fall 2014, followed by the production version  
512 testing in late 2014.

513 Similarly, the first prototype of the associative memory chip using conventional 2D technology is expected to be  
514 delivered early 2014. As outlined above, this prototype chip was intended for testing all the important design  
515 blocks of the core functionalities of associative memory. For this purpose the design was kept simple and for this  
516 version we intentionally did not include all features needed for L1 applications. Preparation for testing as well  
517 as testing itself is expected to take 3-6 months. Engineering work to adjust the 2D chip design to accommodate  
518 higher pattern density, faster speed and sparsified readout needed for L1 applications will start as soon as funding  
519 is available and will take into consideration results of the protoVipram testing. We expect protoVIPRAM2 to be

520 submitted for productionin Fall 2014 with delivery in late 2014 - early 2015, if funding is becoming available  
521 starting Jan. 2014.

522 Design of the Pattern Recognition Mezzanine card will proceed in parallel and in close communication with the  
523 protoVIPRAM2 development with submission taking place at approximately the same time. Since production of  
524 the card is expected to take less time than for the ASIC, the new mezzanine prototype is expected to be available  
525 in Fall 2014. Testing will take place in late 2014 prior to the arrival of protoVIPRAM2.

526 Early 2015 will be dedicated to the final tests of the hardware components and integration. Setup of the test stand  
527 will take place in Spring 2015. It is expected that most of the engineering effort will be spent on the firmware  
528 development for the Pulsar and PRM cards. Different versions of firmware will be needed for the Pulsar board  
529 to function as DIB and PRB. The Pulsar II boards can be used as data sourcing boards as well, in which case the  
530 firmware will be from TK-DAQ. Integration will take place in Spring-Summer 2015 followed by crate-level testing  
531 and measurements of the system performance parameters.

## 532 **4.2 Milestones**

- 533 • Pulsar-II/RTM design finalized: by FY2014Q3
- 534 • Pulsar-II/RTM design tested: by FY2014Q4
- 535 • Initial Pulsar-II firmware for DIB, PRB finished: FY2014Q4
- 536 • Pattern Recognition Mezzanine card design finalized: FY2014Q4
- 537 • Pattern Recognition Mezzanine card fully tested: FY2015Q1
- 538 • ProtoVIPRAM2 initial design dedicated for CMS L1 Track Trigger finished: FY2014Q4
- 539 • ProtoVIPRAM2 prototype tested: FY2015Q2
- 540 • Crate level test-stand setup: FY2015Q3
- 541 • Initial system level performance study: FY2015Q4

## 542 **4.3 Facilities, Equipment, and Other Resources**

543 The proposed R&D would be carried out as a collaborative effort among Fermilab, Northwestern, University of  
544 Florida, and Tezzaron Semiconductor [14]. Some of the physicists in this collaboration have been involved in the  
545 design, building, commissioning, operation and upgrade of the CDF SVT system, as well as the design work of  
546 the FTK system. It is worth mention that Luciano Ritorì, one of the original inventors of the SVT, has recently  
547 joined this project. A few years ago, Fermilab also collaborated closely with INFN Pisa and Frascati in Italy on  
548 the 2D development of AMchip04 [17] in 65 nm. Fermilab contributed to the new Majority Logic design as well  
549 as the pattern readout algorithm using Fisher Tree approach. The extensive experience in associative memory  
550 approach within the collaboration will be important for carrying out this R&D project. In addition, this proposal  
551 will leverage unique areas of engineering expertise at Fermilab.

552 The Fermilab ASIC Design Group is a leader in 3D ASIC design, and has expertise with the design of the mem-  
553 ory cells. The preliminary protoVIPRAM design work already done by the group would be a starting point for  
554 this project to design a dedicated associative memory device for CMS L1 tracking trigger demonstration using  
555 conventional technology (plan to use 130 nm to keep the cost low).

556 We believe that there is already enough critical mass of technical and scientific expertise to move forward with the  
557 design and construction of the vertical slice demonstration for CMS Level 1 tracking trigger, and we are actively  
558 looking for more collaborators to join the project, both from USCMS and outside.

## 559    5 Budget for FY14 and Preliminary Budget for FY15

560    The fully loaded funding request for FY2014 is \$637,429.53. This consists of \$497,320.86 of labor and \$138,442.00  
561    of M&S including travel to CERN. Following the feedback from the LOI review stage, we have reduced the re-  
562    quired amount by approximately \$120,000 by requesting one ATCA shelf worth of hardware in FY2014 instead of  
563    two. As outlined above, this is reasonable at the early stage because the demonstration system will be implemented  
564    in stages from bottom up, at mezzanine card level, board level, crate level and finally multi-crate level.

565    In FY2014 the requested funds are for:

566    1. The development of the Pulsar II hardware/firmware (engineer-III time, 1FTE): \$248,660.00 fully loaded

- 567       • Pulsar II hardware design work (minor), and firmware work (major).
- 568       • Pulsar II RTM (with minor modifications from existing RTM)
- 569       • Pulsar II mezzanine design and firmware (major)

570    2. The development of 2D protoVIPRAM2 (engineer-III time, 1FTE): \$248,660.00 fully loaded

- 571       • Dedicated chip design for CMS L1 track trigger
- 572       • Chip submission cost not included here (either from Tracker or in FY15)

573    3. Hardware needs: \$123,550.00 including all overheads

- 574       • One ATCA shelf
- 575       • Pulsar IIc/RTMs/Mezzanine cards

576    4. Travel (for two FNAL engineers to discuss technical details with collaborators). Four trips to CERN (one  
577    week per visit, 2 visits per engineer): \$14,892.00 including all overheads.

578    The anticipated request in FY2015 is expected to be roughly similar to FY2014. It is prorated 3% (to account for  
579    the inflation), totals \$656,552 and includes:

- 580       1. Firmware work and integration/testing
- 581       2. Second ATCA shelf to test crate-to-crate communication
- 582       3. Pulsar II/RTMs/Mezzanine cards for the second shelf
- 583       4. Travel (same number of trips as in 2014)

## 584 References

- 585 [1] M. Dell Orso, L. Ristori - "VLSI Structure For Track Finding" - Nucl.Instr. and Meth. A278 (1989), 436;  
586 L. Ristori and G. Punzi, Ann. Rev. Nucl. Part. Sci. 60, 595-614 (2010).
- 587 [2] J. Adelman *et al.* - "Real time secondary vertexing at CDF" - Nucl.Instr. and Meth. A569 (2006), 111
- 588 [3] J. Adelman *et al.* - "The Silicon Vertex Trigger upgrade at CDF" - Nucl.Instr. and Meth. A572 (2007), 361
- 589 [4] - "FTK: A Hardware Track Finder for the ATLAS Trigger" - Technical Proposal (2010)
- 590 [5] - "FastTracKer(FTK)TechnicalDesignReport" - CERN-LHCC-2013-007, ATLAS-TDR-021-2013, 2013
- 591 [6] T. Liu *et al.* - "System Architecture for CMS L1 Track Trigger and work plan for Vertical Slice System  
592 Demonstration" - Draft proposal for discussion v1.9, Nov. 16, 2013:  
593 [http://hep.uchicago.edu/~thliu/PulsarII/TT\\_1.9.pdf](http://hep.uchicago.edu/~thliu/PulsarII/TT_1.9.pdf)  
594 For more detailed presentation and discussion, please see most recent talks at the "L1 Track Finding meeting  
595 during the tracker Phase II days" in Nov 19th 2013:  
596 <https://indico.cern.ch/conferenceDisplay.py?confId=278418>
- 597 [7] W. Ketchum *et al.* - "Performance study of GPUs in real-time trigger applications for HEP experiments ",  
598 in Proceedings of the 2nd International Conference on Technology and Instrumentation in Particle Physics  
599 (TIPP 2011), vol. 37, 2012, p. 1965. [Online]. Available at:  
600 <http://www.sciencedirect.com/science/article/pii/S1875389212019207>
- 601 [8] A. Gianelle *et al.* - "Applications of Many-Core Technologies to On-line Event Reconstruction in High  
602 Energy Physics Experiments": <http://arxiv.org/abs/1312.0917>
- 603 [9] J. Olsen, T. Liu and Y. Okumura - "A Full Mesh ATCA-based General Purpose Data Processing Board: Pulsar II "  
604 - FERMILAB-CONF-13-526-CMS-PPD (Nov. 22, 2013), presented at TWEPP 2013 and submitted  
605 to JINST:  
606 [http://hep.uchicago.edu/~thliu/PulsarII/PulsarII\\_twepp\\_paper.pdf](http://hep.uchicago.edu/~thliu/PulsarII/PulsarII_twepp_paper.pdf)
- 607 [10] Y. Okumura, J. Olsen, T. Liu and H. Ying - "Prototype Performance studies of a Full Mesh ATCA-based Gen-  
608 eral Purpose Data Processing Board " - FERMILAB-CONF-13-527-CMS-PPD (Nov. 22, 2013), presented  
609 at IEEE NSS 2013:  
610 [http://hep.uchicago.edu/~thliu/PulsarII/PulsarII\\_NSS\\_paper.pdf](http://hep.uchicago.edu/~thliu/PulsarII/PulsarII_NSS_paper.pdf)
- 611 [11] J. Olsen, T. Liu and Y. Okumura - Pulsar IIa prototype work web page:  
612 <http://www-ppd.fnal.gov/EEDOffice-w/Projects/atca/>
- 613 [12] T. Liu *et al.* - "A New Concept of Vertically Integrated Pattern Recognition Associative Memory (VIPRAM)"  
614 - FERMILAB-CONF-11-709-E, Proceedings of TIPP 2011 conference, Volume 37, 2012, Pages 1973:  
615 <http://www.sciencedirect.com/science/article/pii/S1875389212019165>  
616 or  
617 <http://lss.fnal.gov/archive/2011/conf/fermilab-conf-11-709-e.pdf>
- 618 [13] T. Liu *et al.* - "Development of 3D Vertically Integrated Pattern Recognition Associative Memory (VIPRAM)"  
619 - FERMILAB-TM-2493-CMS-E-PPD-TD:  
620 <http://lss.fnal.gov/archive/test-tm/2000/fermilab-tm-2493-cms-e-ppd-td.pdf>
- 622 [14] "Tezzaron Semiconductor": <http://www.tezzaron.com>
- 623 [15] A. Annovi *et al.* - "A VLSI Processor for Fast Track Finding Based on Content Addressable Memories" -  
624 IEEE Transactions on Nuclear Science, vol. 53, no. 4 (2006), 1
- 625 [16] A. Annovi *et al.* - "The GigaFitter: A next generation track fitter to enhance online tracking performances at  
626 CDF" - Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE (2009), 1143
- 627 [17] A. Annovi *et al.* - Variable Resolution Associative Memory for High Energy Physics " - Submitted to the  
628 Advancements in Nuclear Instrumentation, Measurement Methods and their Applications (ANIMMA), Ghent  
629 Belgium, 6-9 June, 2011.