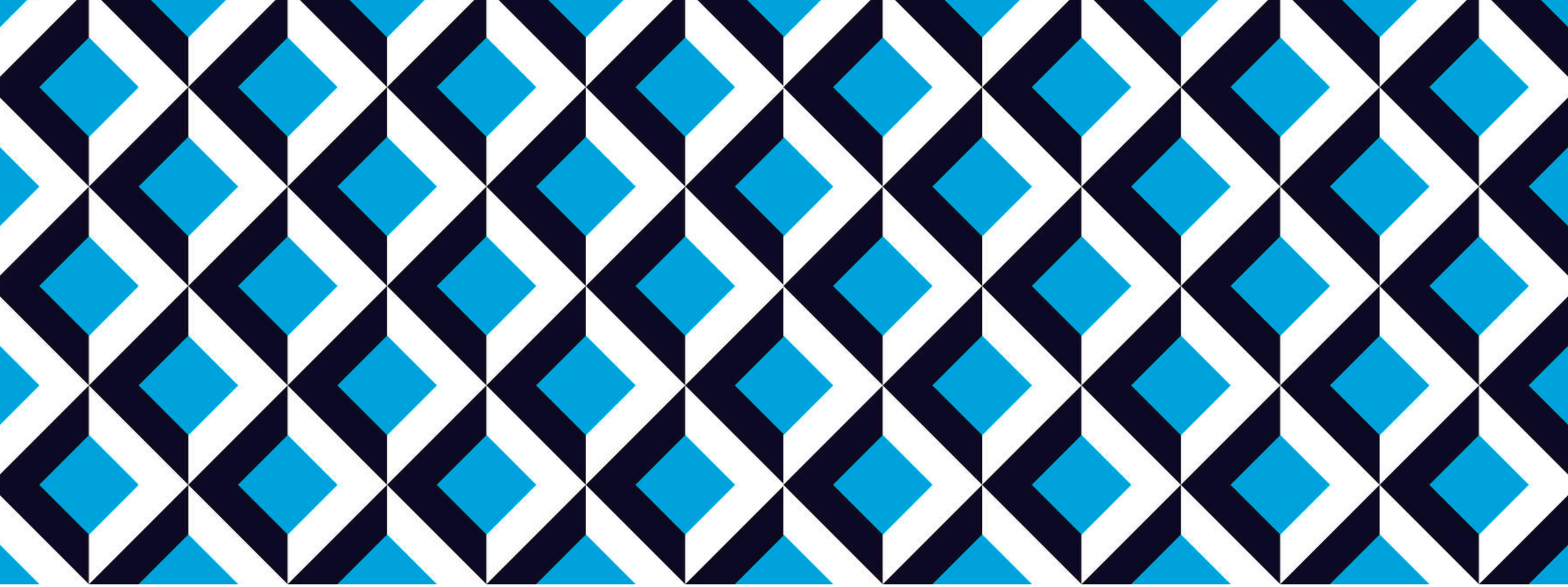


MARIO GRANDI

ALEX CERRI

# FUTURE DEVELOPMENTS SWIFT-HEP / FPGAS

UNIVERSITY OF SUSSEX

- **S**oft**W**are **I**n**F**rastructure and **T**echnology for **H**igh **E**nergy **P**hysics (SWIFT - HEP)
  - UK based project

- Goal is to study and develop new software and algorithm techniques that will be required for the next-generation experiments
  - Most HEP experiments are becoming more complex, with higher event rates and/or data volumes


- New software techniques are need to be developed in parallel to the new hardware that will be used for these future experiments
  - Existing codes used by many HEP experiments evolved from an era of entirely different hardware and software technologies
    - Have now become obsolete and inadequate for the task

# SUSSEX CONTRIBUTION TO SWIFT-HEP

- The deployment and performance comparison of tracking and vertex algorithms in **FPGA accelerators** for real-time and offline reconstruction using the expertise gained from ATLAS and CMS
  - The goal is to identify the techniques required to enable **algorithms** to be deployed cost effectively both **online** and **offline** in computing environments that employ **both accelerators and CPUs**.
- Use of **GPUs** to improve efficiency of **MC** simulations (Josh)

- The project will require to take existing algorithms and replicate them in FPGA accelerators making use of **High-Level-Synthesis** (HLS)
  - Compare algorithms performances between ATLAS and CMS, hough transform, NN…

- Document best practice for higher level language programming and experience with the vector-specific development environments and tools.

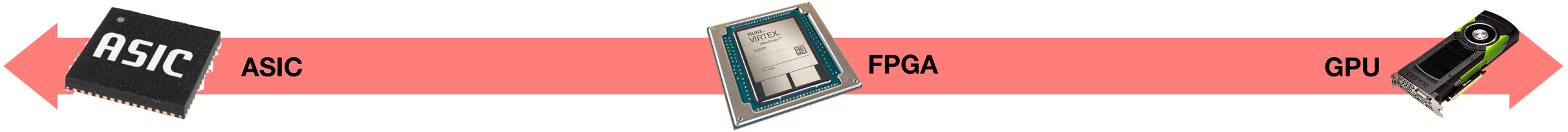
- The deployment and performance comparison of tracking and vertex algorithms in **FPGA accelerators** for real-time and offline reconstruction using the expertise gained from ATLAS and CMS
  - The goal is to identify the techniques required to enable **algorithms** to be deployed cost effectively both **online** and **offline** in computing environments that employ **both accelerators and CPUs**.
- Use of **GPUs** to improve efficiency of **MC** simulations (Josh)

- The project will require to take existing algorithms and replicate them in FPGA accelerators making use of **High-Level-Synthesis** (HLS)
  - Compare algorithms performances between ATLAS and CMS, hough transform, NN…

- Document best practice for higher level language programming and experience with the vector-specific development environments and tools.

LETS DISCUSS ACCELERATOR TECHNOLOGY FIRST

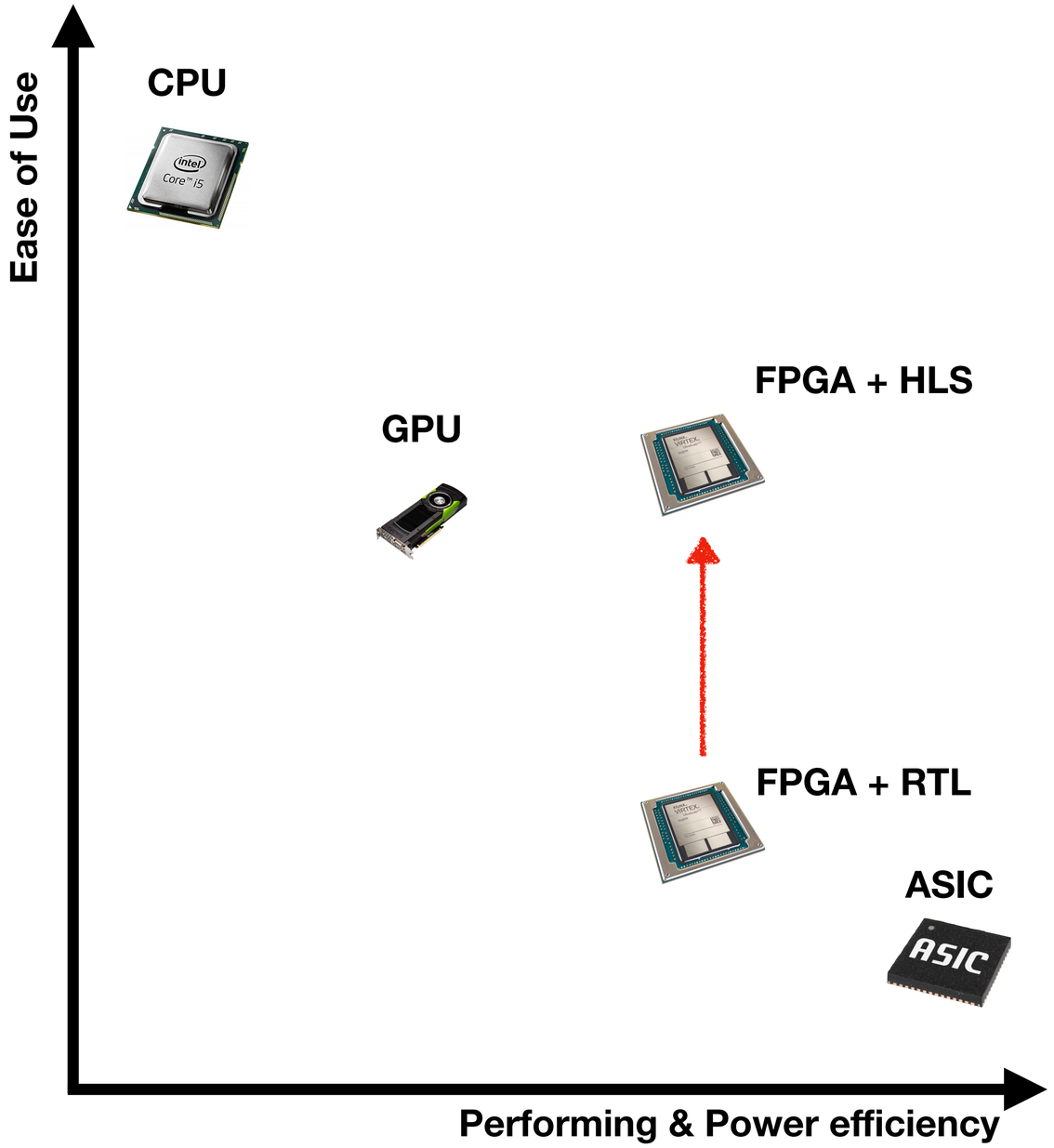
- High performing accelerators are essential towards any sort of high performance software, required by modern based digital electronics
- Most common accelerators are:
  - **Central Processing Unit (CPU)** - basic architecture implemented at hardware that allows for handle of wide range of tasks via software optimisation
    - Good performance in terms of clock speed but limited in number of concurrent tasks that can be running
  - **Graphics Processing Unit (GPU)** - basic architecture already implemented at hardware level with some software customisation possible
    - Good performance but with restricted level of possible tasks
  - **Field-Programmable Gate Array (FPGA)** - allow the designer to create a custom circuit implementation of an algorithm that is reconfigurable
    - Not as efficiency as ASICs and fairly pricy but allow for re-configuration
  - **Application-Specific Integrated Circuit (ASIC)** - custom circuit implementation of algorithm applied at hardware level
    - High performance and high cost, can only perform the task that is configured to do and cannot be re-configured



# MORE ON ACCELERATORS

- Hardware Definition Languages (HDL) are used to describe the hardware
  - **R**egister-**T**ransfer **L**evel (**RTL**)
    - **V**ery high speed integrated circuits **H**ardware **D**escription **L**anguages (**VHDL**)
  - **H**igh **L**evel **S**ynthesis *(**HLS**)*
    - Works at higher level of abstraction
      - Starts with the description of the algorithm in a high level language (C++) and gives the hardware designer better control over optimisation of design architecture
    - The tool will do the RTL implementation for hardware description on the FPGA

- **ASIC**s are the highest performing accelerators but are very expensive and hard to use
  - require high level of specialised training
- **FPGA**s have second highest performance
  - Improved ease of use when developed using HLS

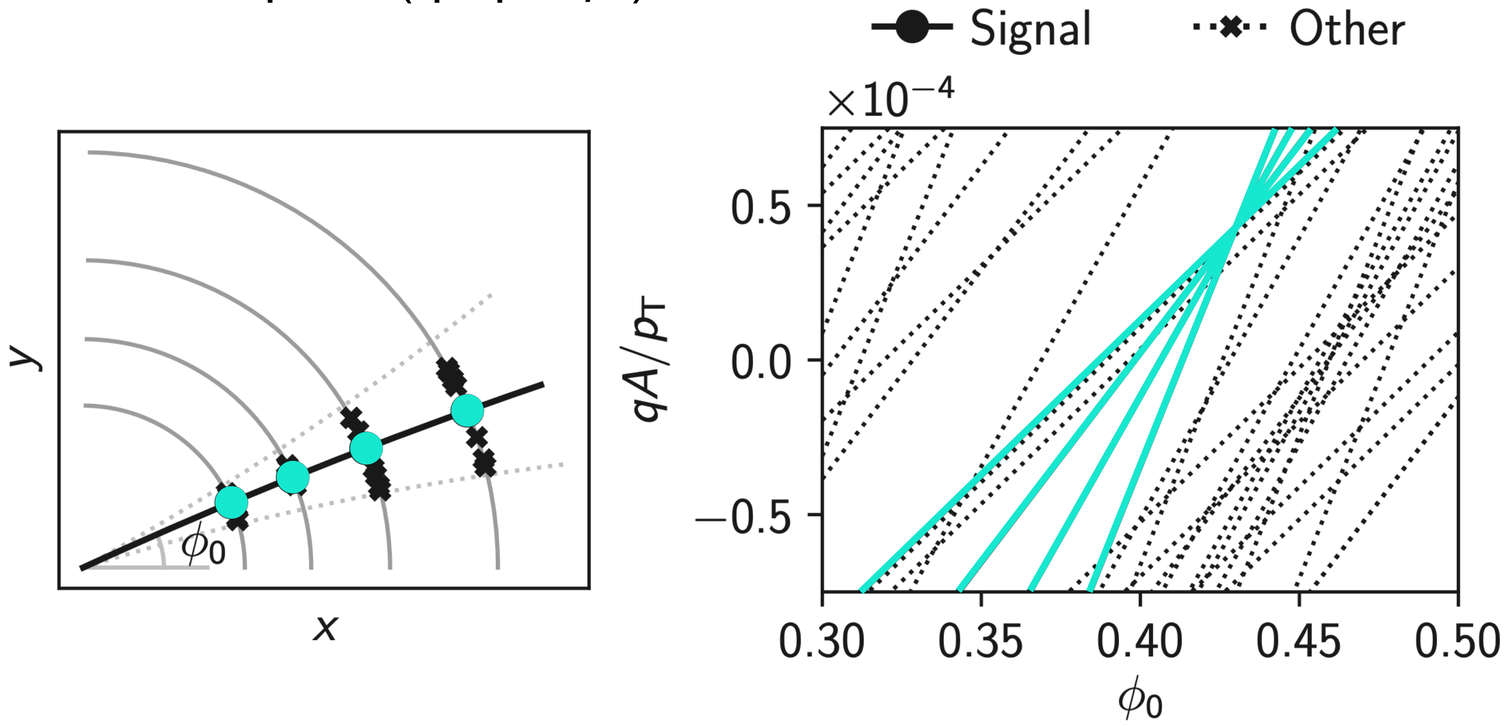


- The deployment and performance comparison of tracking and vertex algorithms in **FPGA accelerators** for real-time and offline reconstruction using the expertise gained from ATLAS and CMS
  - The goal is to identify the techniques required to enable **algorithms** to be deployed cost effectively both **online** and **offline** in computing environments that employ **both accelerators and CPUs**.
- Use of **GPUs** to improve efficiency of **MC** simulations (Josh)

- The project will require to take existing algorithms and replicate them in **FPGA** accelerators making use of **High-Level-Synthesis** (HLS)
  - Compare algorithms performances between ATLAS and CMS, hough transform, NN…

- Document best practice for higher level language programming and experience with the vector-specific development environments and tools.

- Will be focusing firstly on Hough transform track reconstruction algorithm
  - Used to project cluster positions in a detector tracker in an (x, y) transverse plane to a curve in the track parameter space ($qA/p_T$, $\phi_0$) called **accumulator**



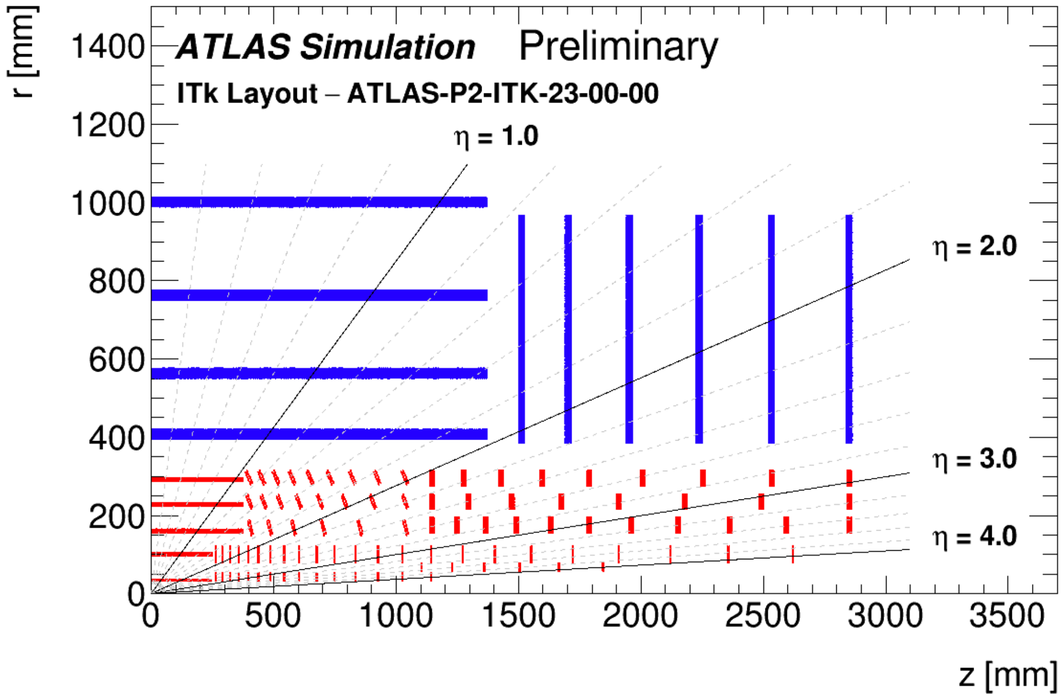

- Will be followed by Kalman filter, and Machine learning algorithm to reconstruct vertices from reconstructed tracks

- Work performed by Sussex master students on other AM based algorithms could have potentially interesting FPGA implementations
  - Use **online** pattern recognition with ML software to generate and populate pattern bank throughout data taking
    - Patterns would evolve as the running conditions change
      - Remove obstacles introduced by simulations, as otherwise running condition would have to be simulated ahead of time
    - Software would have to be extremely light-weight and process data very quickly to perform at the level required online
  - Perform **offline** pattern recognition on data to generate and populate pattern bank
    - Pattern recognition software would have more time and resources for higher accuracy/ performance
    - Limited by amount of data collected (previous years of data collection cannot be used due to severely different running conditions)

- Both ATLAS and CMS have began work towards the development of software tracking algorithms in FPGA for the planned upgrades to the HL-LHC environment

## ATLAS

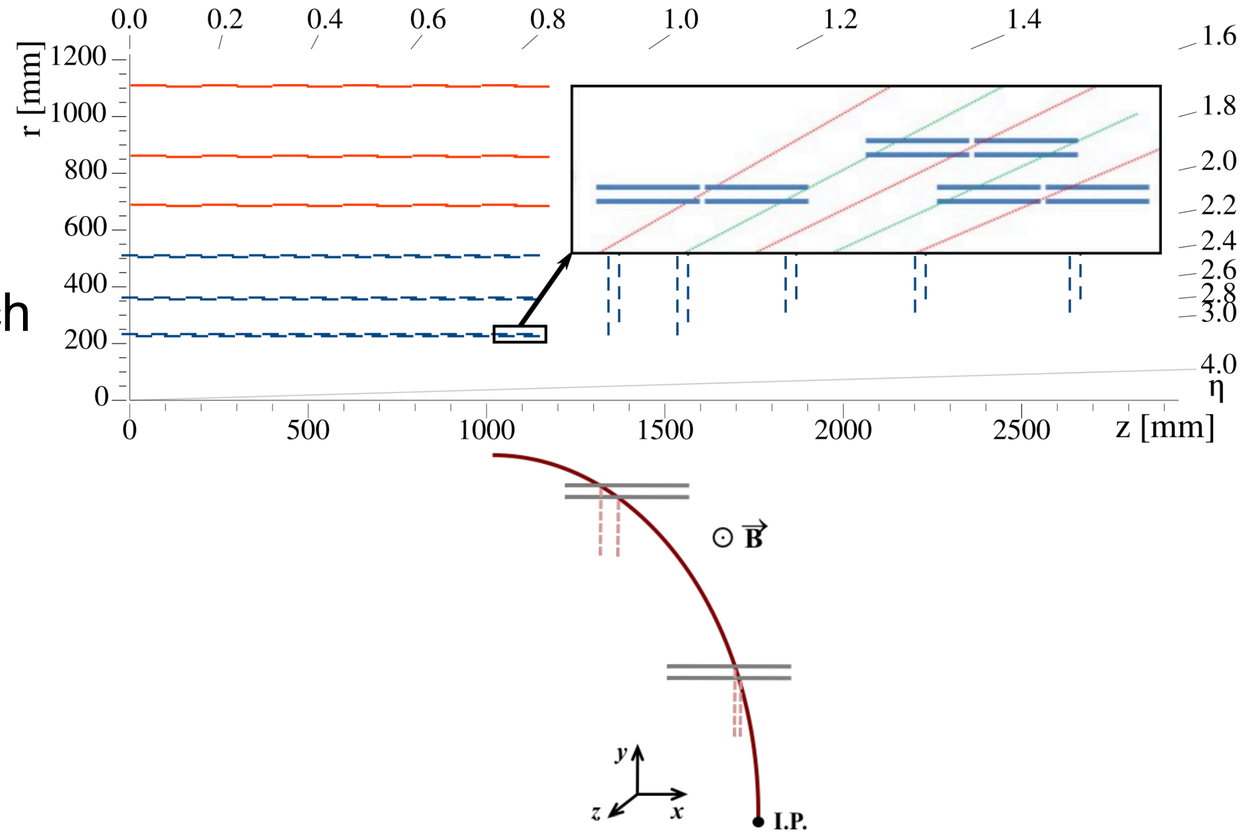
- New Inner Tracker (ITk) detector will provide better coverage for track fitting
- Online tracking is foreseen to be performed in software, possibly with the accelerator-based support
- There is ongoing work for several alternative solutions to the software-only tracking
  - Hough Transform track finding firmware developed on FPGAs using VHDL
  - Fast CPU based software trigger tracking
  - GPU based regional tracking




- A prototype for ASICs based tracking has been developed by ATLAS
  - Could have significant applications for pattern matching/recognitions in other fields
    - genetic sequencing, medical diagnostics etc…

## CMS

- Tracker will be replaced with a new tracker made up of modules composed of two layers which will provide improved tracking efficiency
  - Tracking will be performed on vector information rather than individual hits
    - This results in a significantly less complex pattern recognition problem compared to the one faced by ATLAS
- Three methods have been developed
  - An ASIC+FPGA approach
  - An all-FPGA hough transform approach
  - An all-FPGA road search algorithm approach
- The three methods have been merged to develop a single system design based on commercial FPGAs
  - Will enable joint exploration of tracking algorithms to optimise the final system performance



- **SWIFT-HEP** is a UK based project that involves several institutes around the UK
  - Goal to study and develop new **software algorithms and techniques** that will be essential for the **next-generation HEP experiments**

- Sussex involvement includes the deployment and performance comparison of **tracking and vertex** algorithms with **FPGA accelerators** for both online and offline reconstruction
  - Using resources and expertise gained from ATLAS and CMS
  - Significant overlap is found with the ATLAS activities in their study of accelerators for trigger level reconstruction

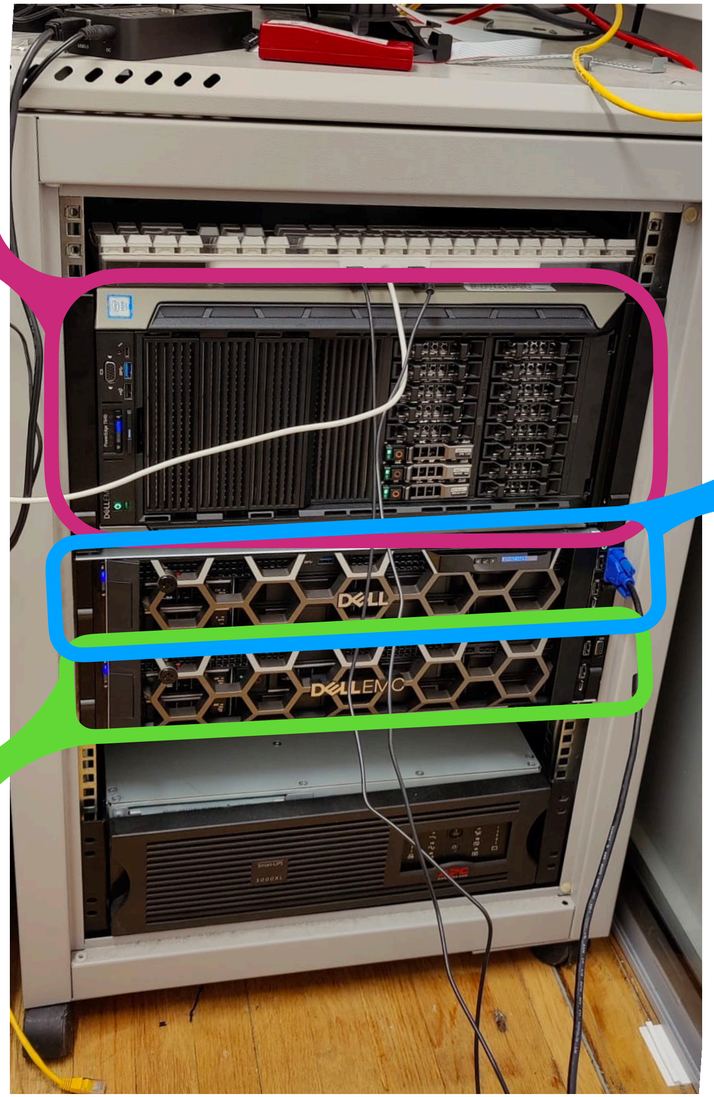
- Promising work has already been performed by both ATLAS and CMS
  - Provides a very good starting point to begin implementation of currently available algorithms and start to perform performance comparisons

- The technologies studied and developed in this project have a wide range of applications that possibly extend beyond HEP and into other fields of science

# THANK YOU

# MACHINERY AT SUSSEX

**T640 (Alveo)**
contains compatible slots for FPGA accelerators
1x U280 Alveo card (FPGA accelerator card)

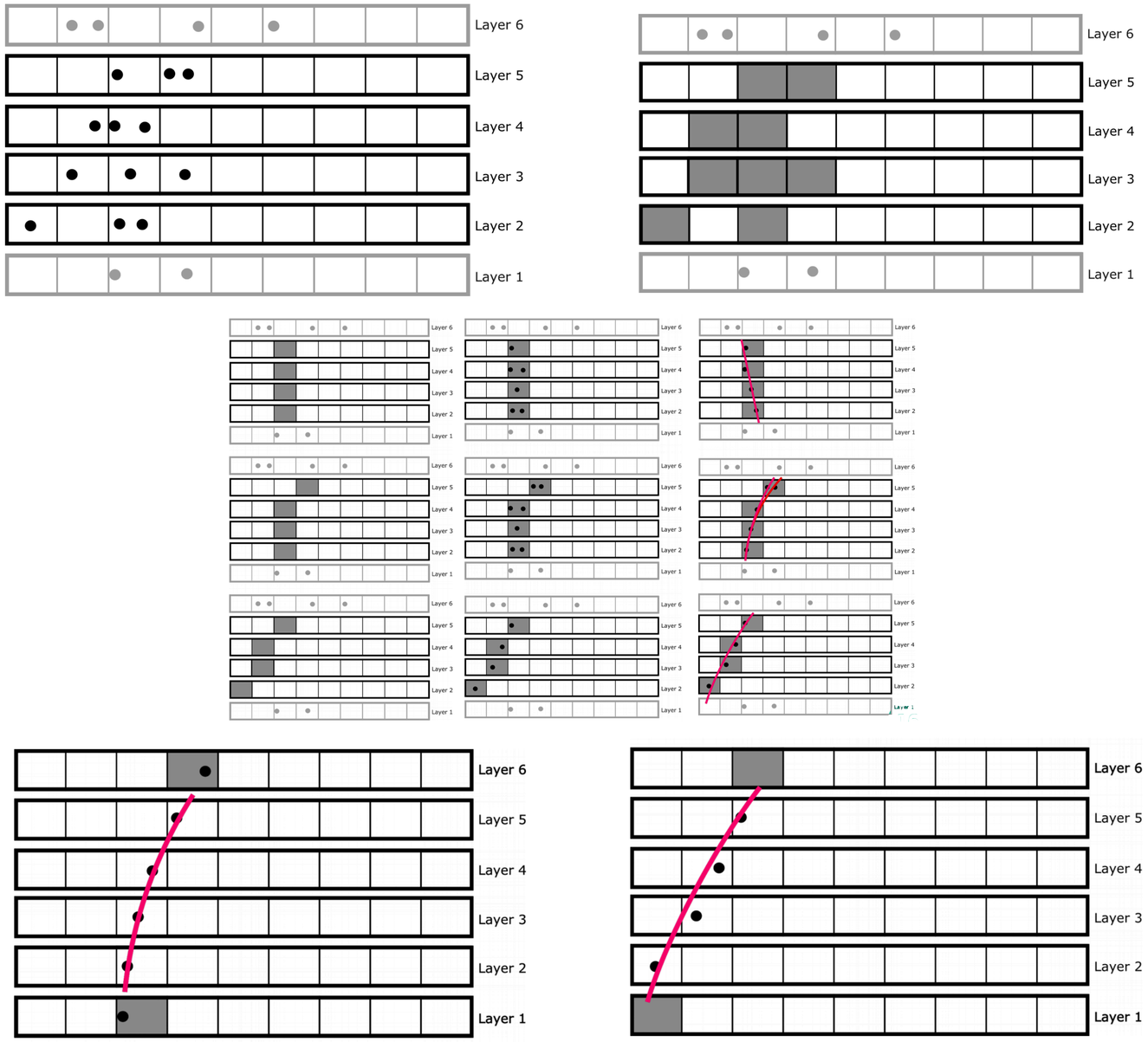



**R7920 (Muon)**
contains compatible slots for FPGA accelerators
No FPGA accelerator card

**R??? (Vivado)**
No compatible slots for FPGA accelerators available
Used as firmware synthesis “workhorse”

# CURRENT PATTERN MATCHING METHOD USED BY ATLAS WITH AM ASICS






- From a cluster of hits in the pixel strips define super-strips (coarser resolution strips) for the inner layers
- Find super-strip for each cluster
- Pattern match super-strips
- Get full resolution clusters within the pattern matched super-strips
- Fit all possible combinations of clusters and remove bad patterns via $\chi^2$ cuts
- Remove duplicates
- Finally extrapolate to super-strips into remaining layers, recalculate $\chi^2$ and remove bad patterns
- Remaining patterns will populate the pattern bank used to identify patterns online