

COMET Phase-I 実験に向けた ニューラルネットワークを実装したFPGAによる オンライン事象選別の研究

宮滝 雅己^A

m-miyataki@epp.phys.sci.osaka-u.ac.jp

藤井祐樹^B、中沢遊^C、吉田学立^D、上野一樹^{A,C}

MyeongJae Lee^E、青木正治^A

阪大理^A、Monash Univ.^B、KEK素核研^C、阪大RCNPD^D、成均館大^E

2022/10/17 計測システム研究会@J-PARC

目次

- COMET Phase-I 実験 @ J-PARC
- 研究動機
- トリガーアルゴリズム
- Neural Network実機試験
- まとめ

ミュオン電子転換過程

ミュオンがニュートリノを放出せずに電子へ転換する荷電レプトンフレーバー非保存過程

信号事象

- $\mu^- + N(A, Z) \rightarrow e^- + N(A, Z)$
- 電子は単色エネルギー
- Alの場合、105 MeV/c

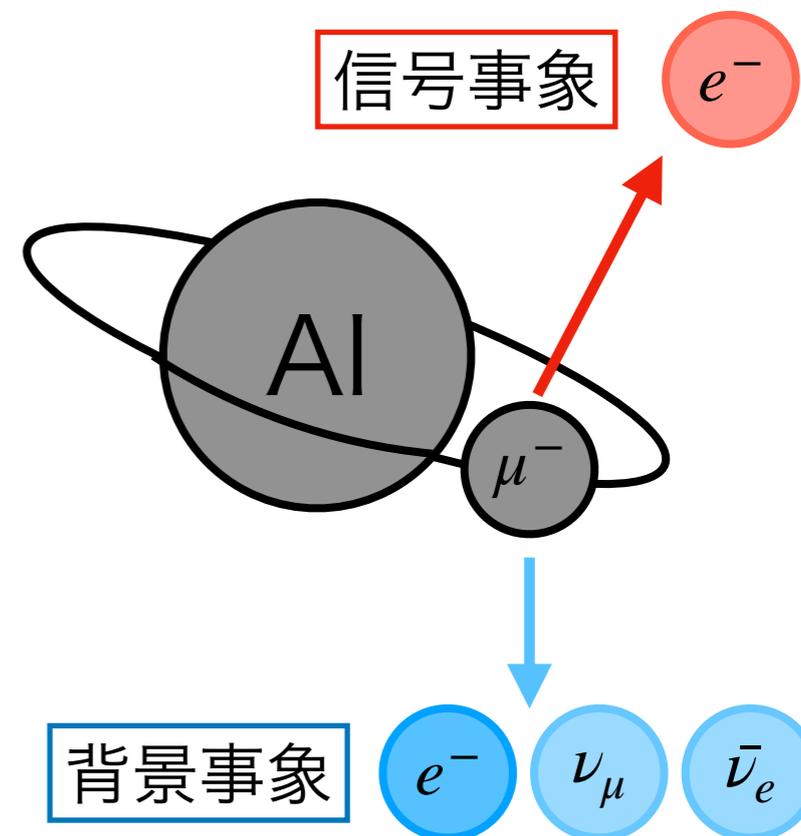
背景事象

- $\mu^- \rightarrow e^- + \nu_\mu + \bar{\nu}_e$
- 3体崩壊であり、電子のエネルギーは連続的に分布

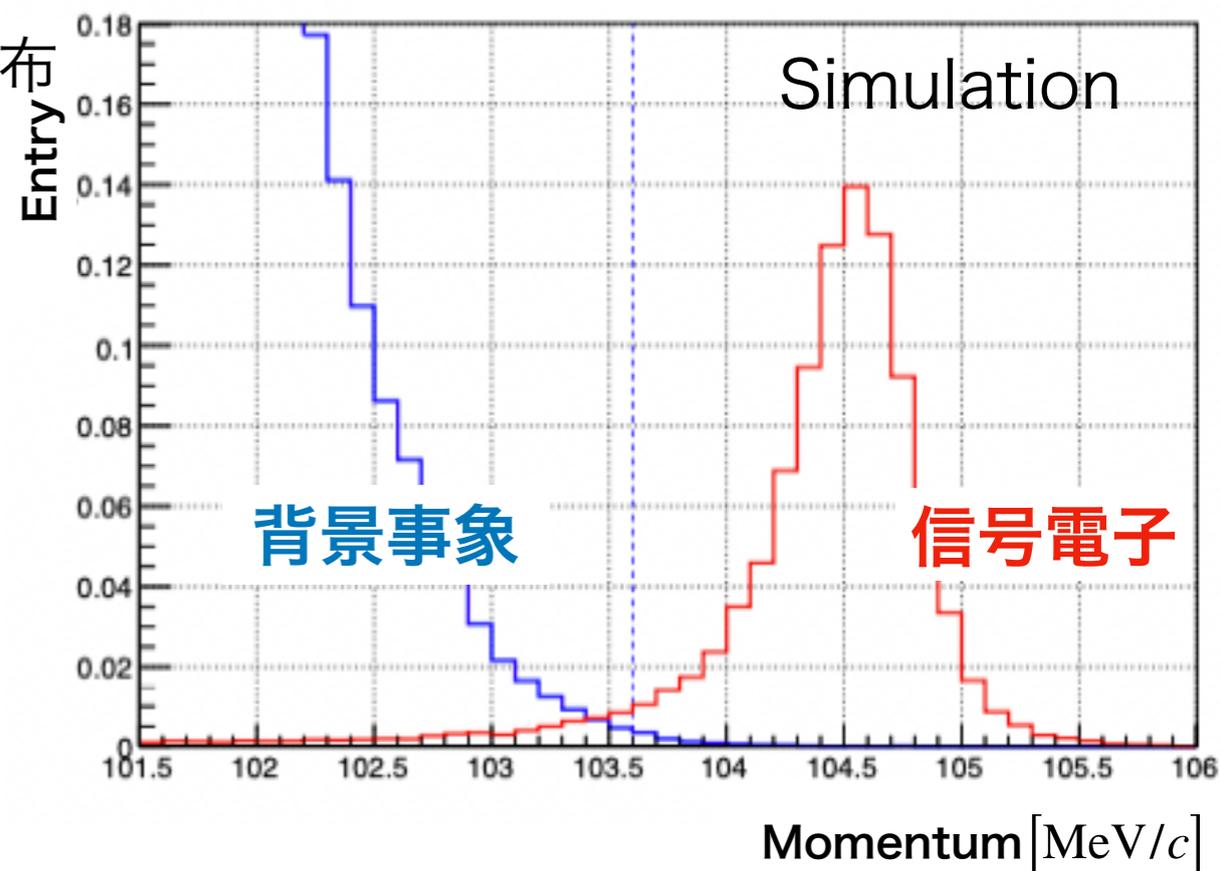
信号事象の分岐比

- 標準模型 + ニュートリノ振動 $O(10^{-54})$
- 標準模型を超えた物理 $O(10^{-15} \sim 10^{-17})$

COMET実験はミュオン電子転換過程を探索



$$BR = 3 \times 10^{-15}$$



COMET Phase-I 実験 @ J-PARC

茨城県東海村

目的： **Al原子中でのミュオン電子転換過程探索**

- 150日間の測定で感度 $\sim 7 \times 10^{-15}$ @ 90% C.L.

<-> 現在の上限値 SINDRUM II*

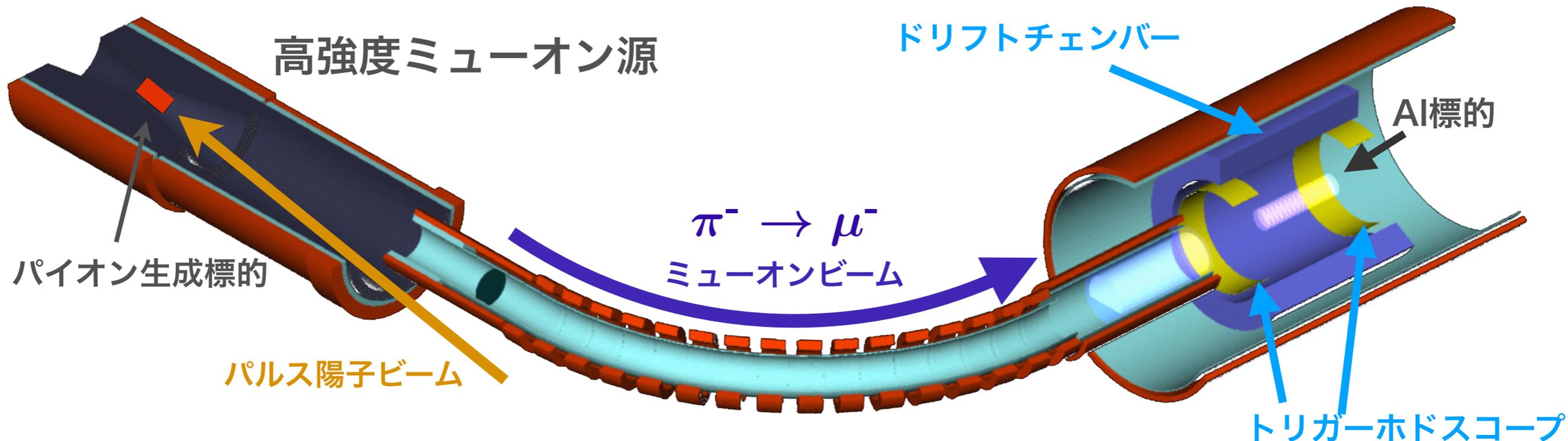
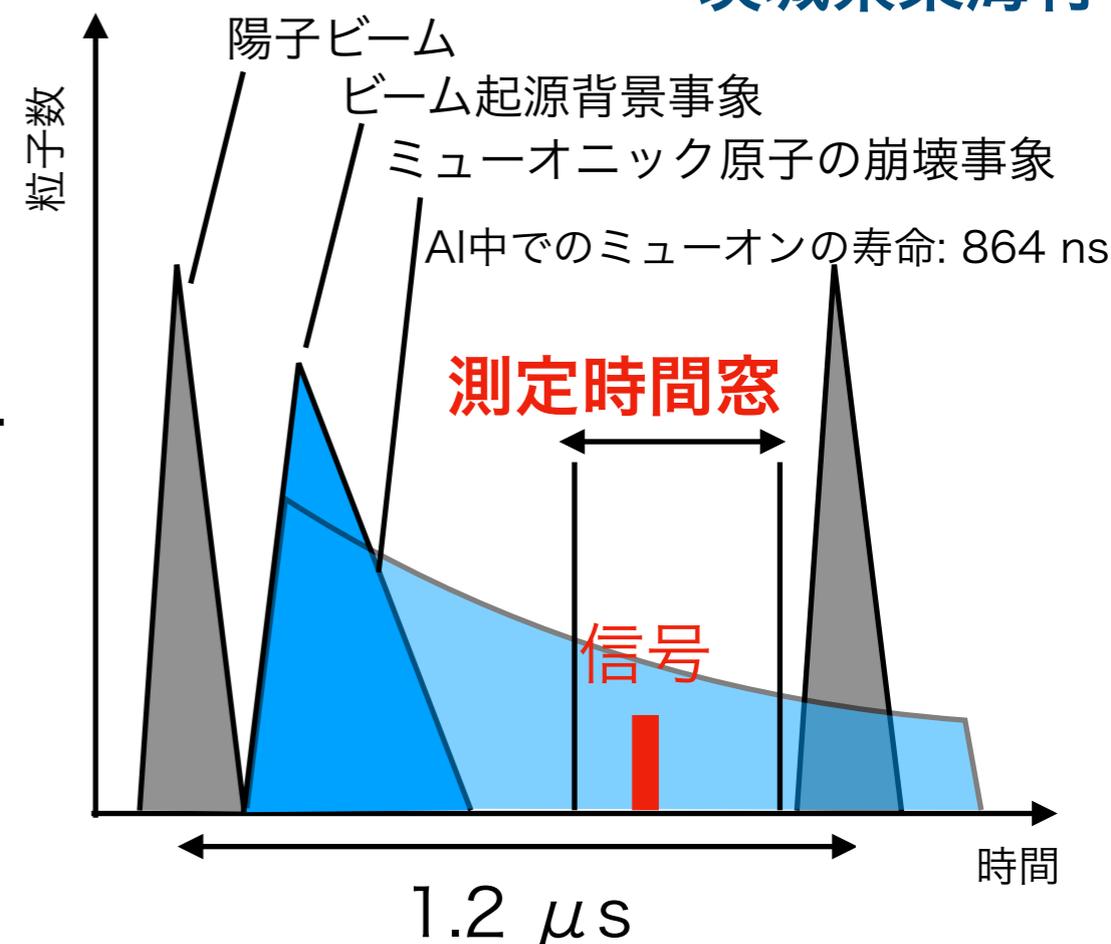
$$B(\mu^- + \text{Au} \rightarrow e^- + \text{Au}) < 7 \times 10^{-13} \text{ @ 90\% C.L.}$$

ビーム構造

：ビーム背景事象を排除するためパルス陽子ビーム

検出器：円筒型検出器

- ドリフトチェンバー + トリガーホドスコープ

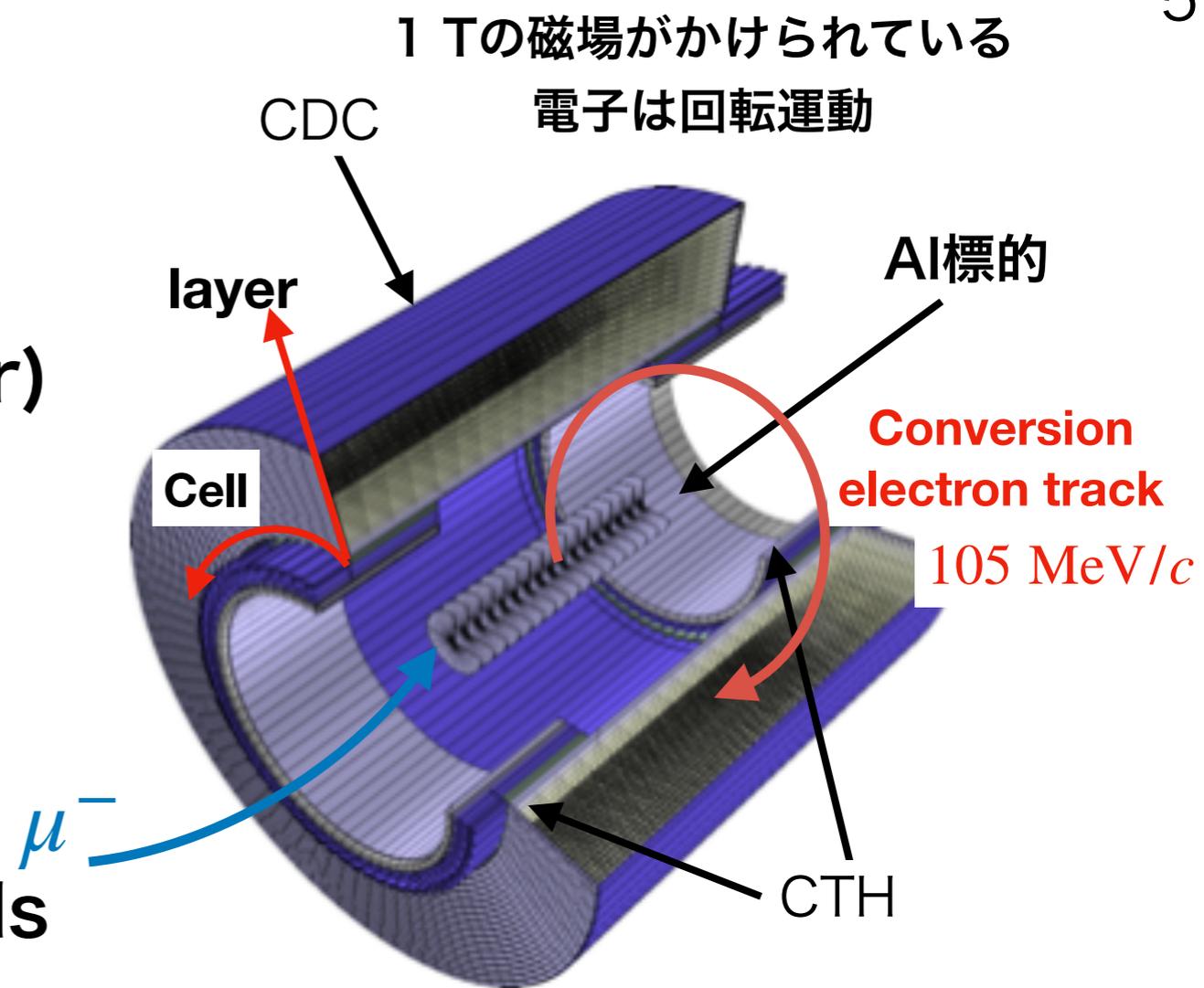


円筒型検出器

◆ ドリフトチェンバー

CDC (Cylindrical Drift Chamber)

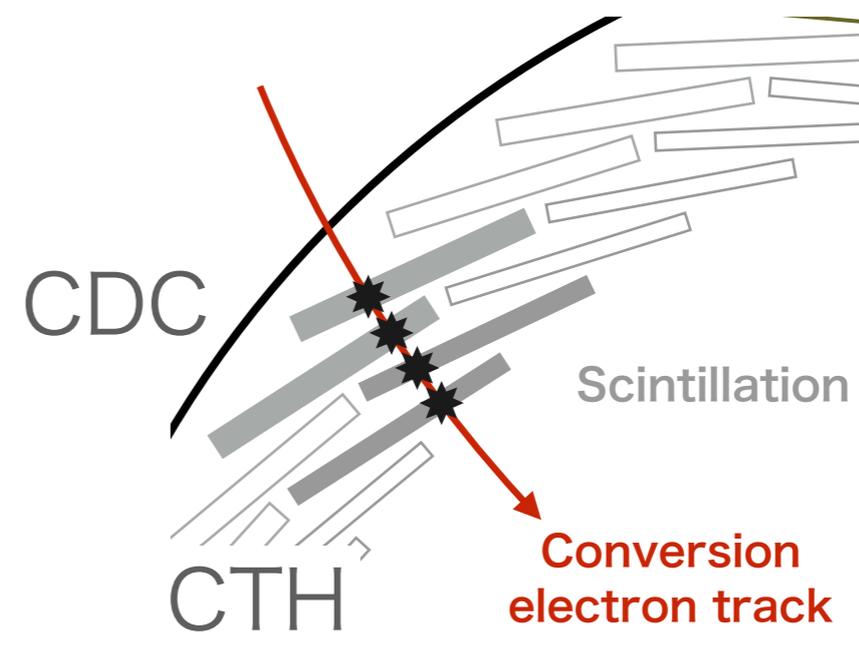
- 外径 ~1.7 m、内径 ~1 m、長さ ~1.5 m
- 粒子の運動量を測定
 - 4986 sense wires
 - **20 stereo layers x ~250 cells**



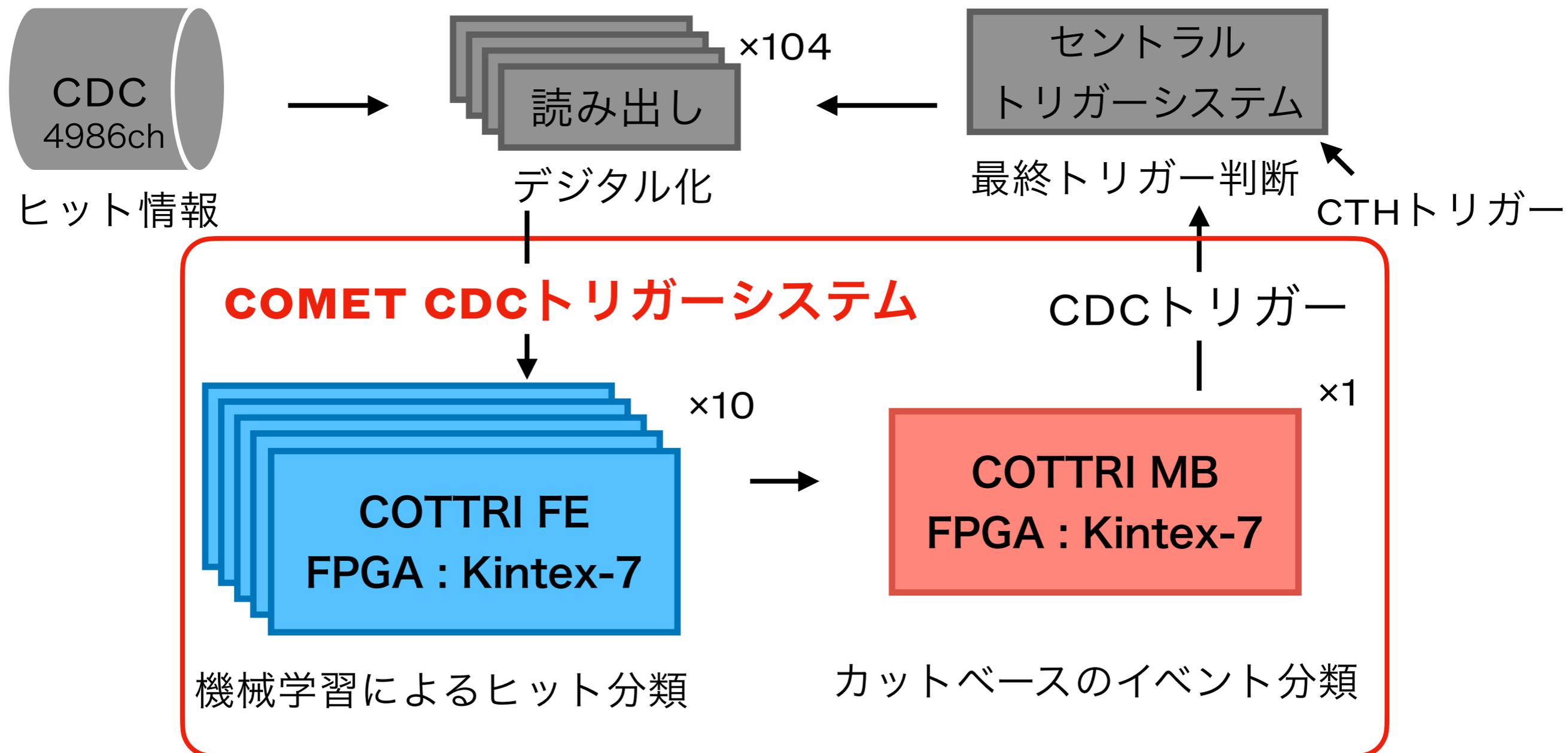
◆ トリガーホドスコープ

CTH (Cylindrical Trigger Hodoscope)

- 電子の時間情報を測定
- 4重コインシデンスにより偶発的なフェイクトリガー事象を除去



FPGAベースオンライントリガーシステム 6



- **トリガー要求性能を達成**

<https://doi.org/10.1109/tns.2021.3084624>

- 96 %の信号電子保持効率でトリガーレート13 kHz

(CTHのトリガーレートは91 kHz、DAQの要求は26 kHz)

- Latency : 3.2 μ s (要求は7.5 μ s)

トリガー効率を上げ測定時間窓を広げ、
シグナルアクセプタンスを向上させたい

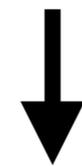
	現在	目標
測定時間窓(ns)	[700,1170]	[500,1170]
シグナルアクセプタンス(%)	4.2	7.0

x 1.7

現在のトリガーシステムは500 ns開始ではトリガーレートが13 kHzを超える

現在のトリガーアルゴリズム

= 機械学習によるヒット分類 + カットベースのイベント分類



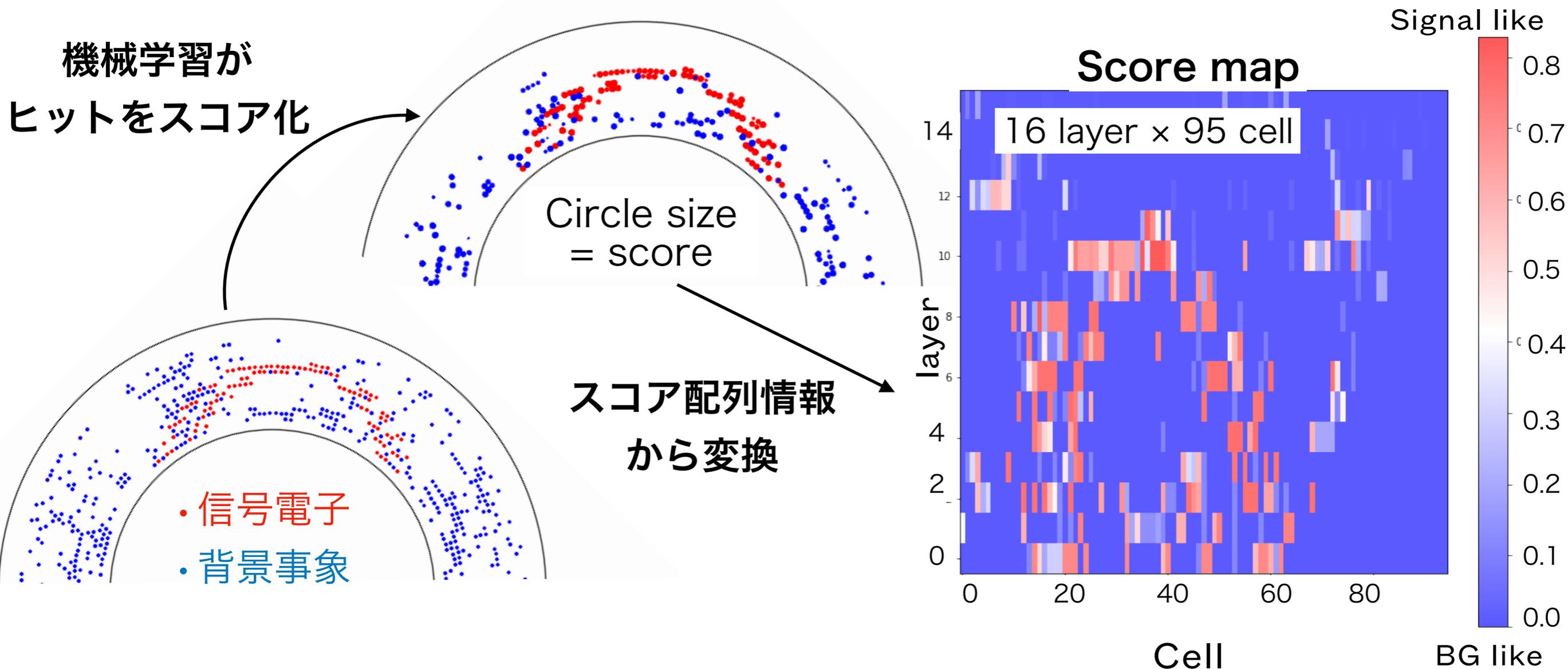
開発アルゴリズム

= 機械学習によるヒット分類 + **Neural Networkのイベント分類**

新たなトリガーアルゴリズムの開発

開発アルゴリズム

= 機械学習によるヒット分類 + Neural Networkのパターン認識によるイベント分類



- Score mapをNeural Networkの入力に用いることで、

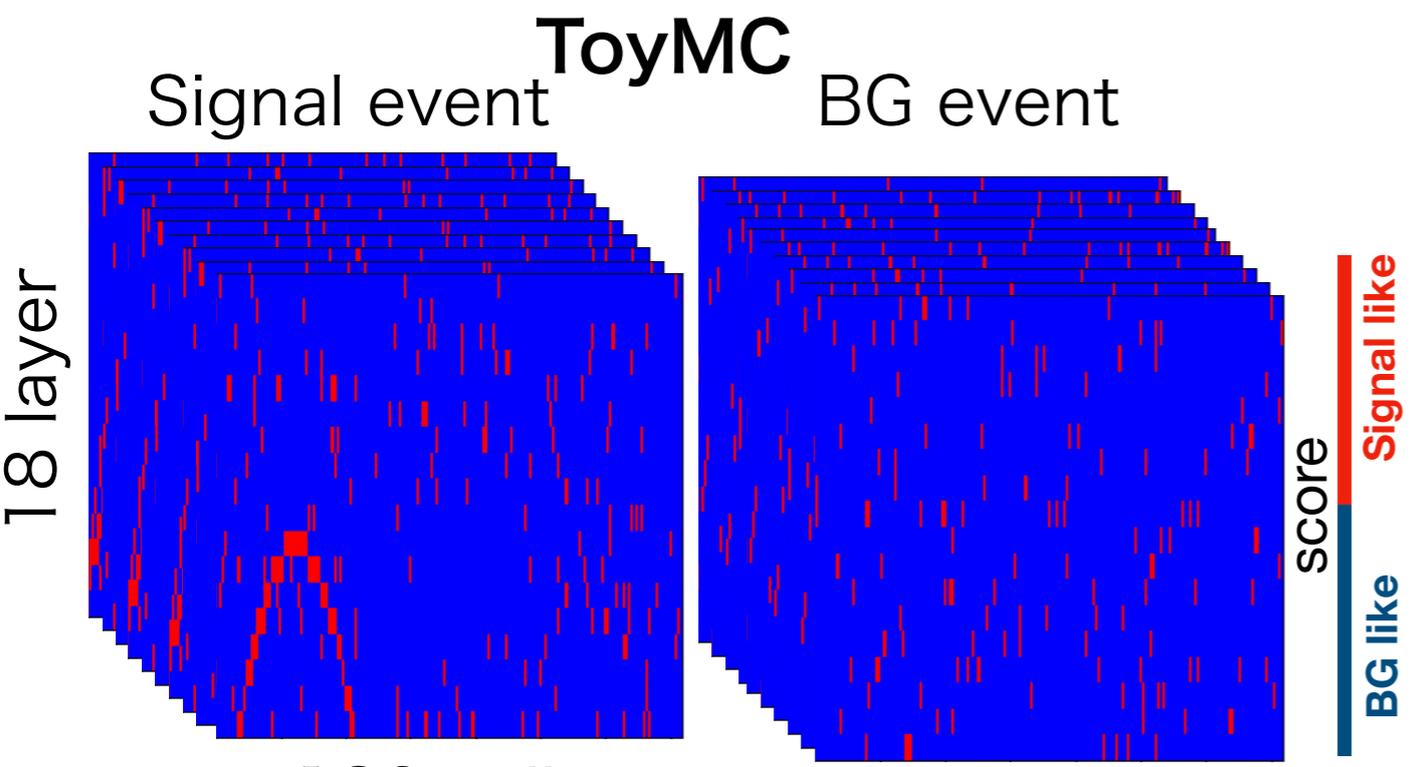
イベント分類に信号電子が描く軌跡のパターン認識を導入

問題設定とハードウェア制約を簡略化し、まず以下の3つを確認したい

- **ソフトウェア、FPGAファームウェア両サイドの開発スキーム**
- **実機FPGA上のNeural Networkの性能**
- **実機FPGA上でNeural Networkの推論にかかるlatency**

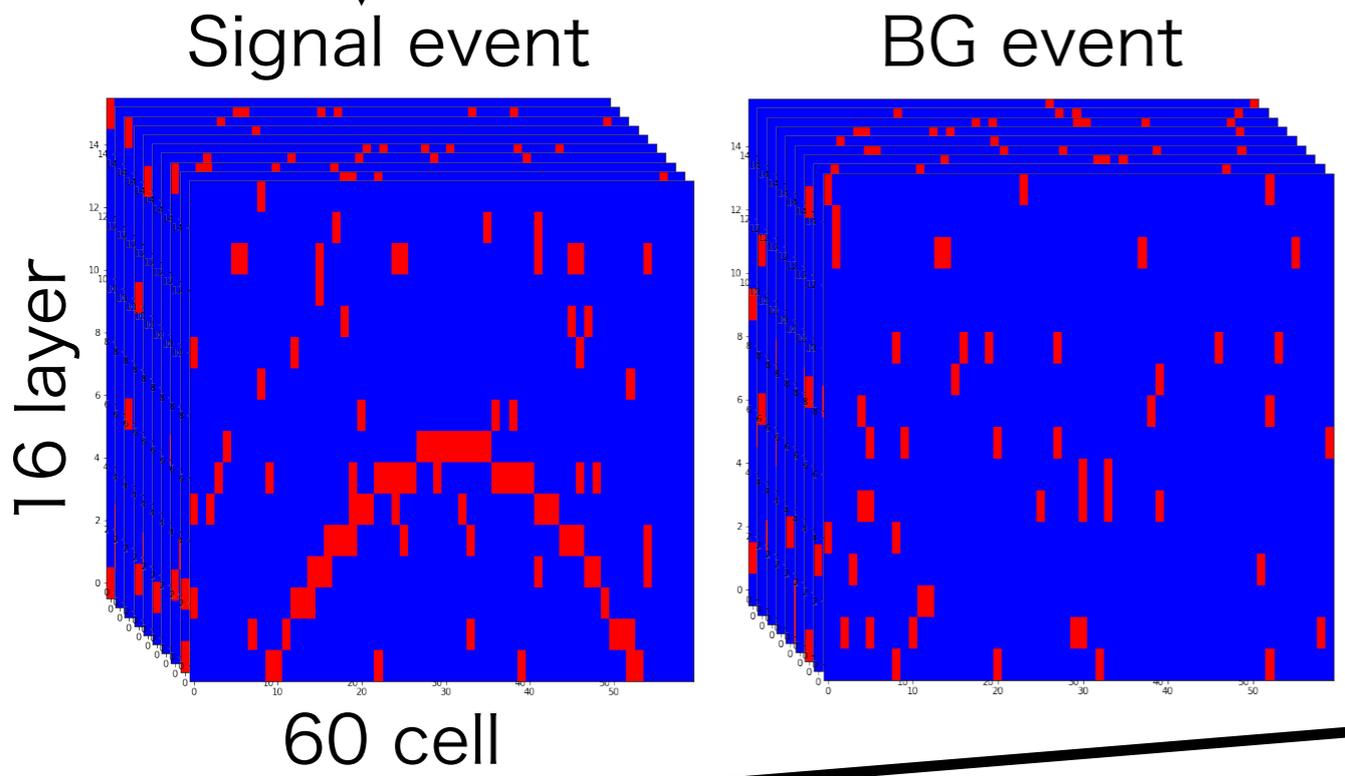
簡略化した設定でソフトウェア、FPGAファームウェアの開発を行った

データセットの用意

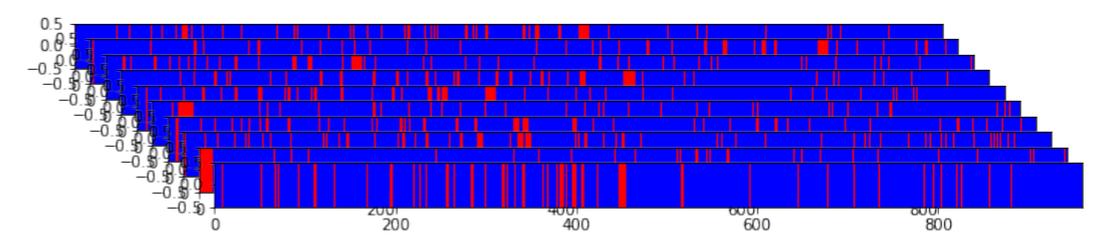


180 cell

score総和が大きい領域を抽出



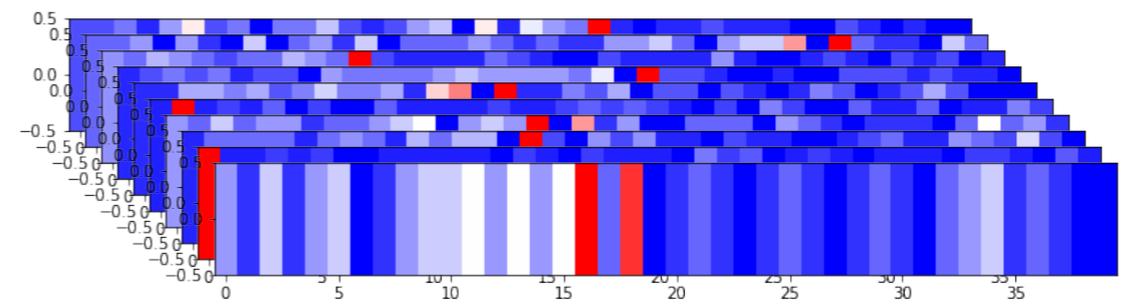
一次元化



1 bit x 960
(BGも同様に一次元化)

24個ずつSum
(何個足すかは仮決め)

NNデータセット



6 bit x 40
(BGも同様に圧縮)
25000 eventずつ用意

モデル構築 + モデルのhls変換

全結合型5層の量子化MLP(QMLP)を構築

 TensorFlow

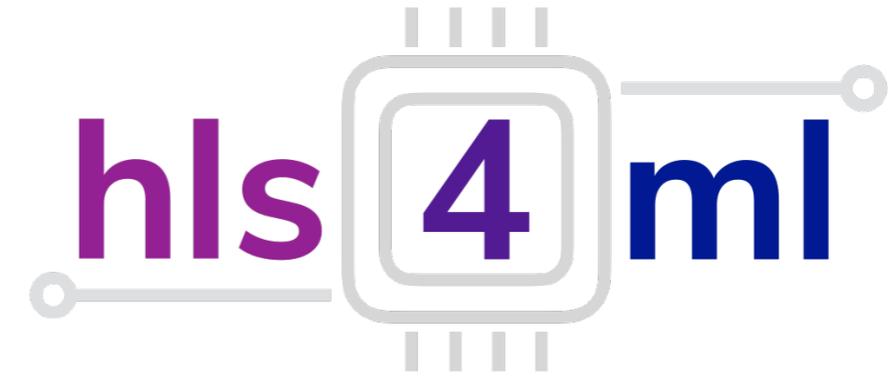
 Keras

QKeras*

<https://doi.org/10.48550/arXiv.2006.10159>

*Kerasの拡張ライブラリでヘテロな量子深層Neural networkが作成可能

訓練済みモデルをファームウェア
変換可能なコード(RTL)に変換



© Copyright 2021, Fast Machine Learning Lab.

<https://dx.doi.org/10.1088/1748-0221/13/07/P07027>

hls4mlは目標FPGAに対し実装に必要なlatencyとリソース使用率をレポートしてくれる

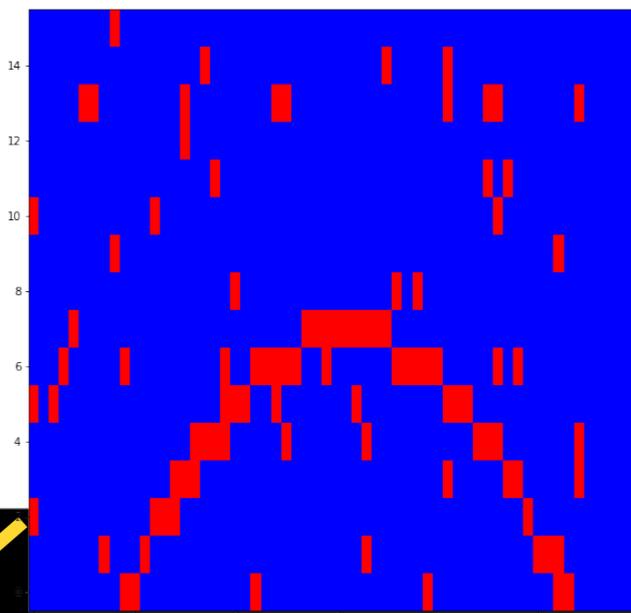
FPGA : AMD Xilinx Kintex-7 xck355t-ffg901-1

	Usage (%)			
Latency @200 MHz	BRAM	DSP	FF	LUT
130 ns 🍌	~0 🍌	~0 🍌	5 🍌	32 🍌

hls4mlで作成したQMLP moduleをsimulationで確認後、実機で試験を行った

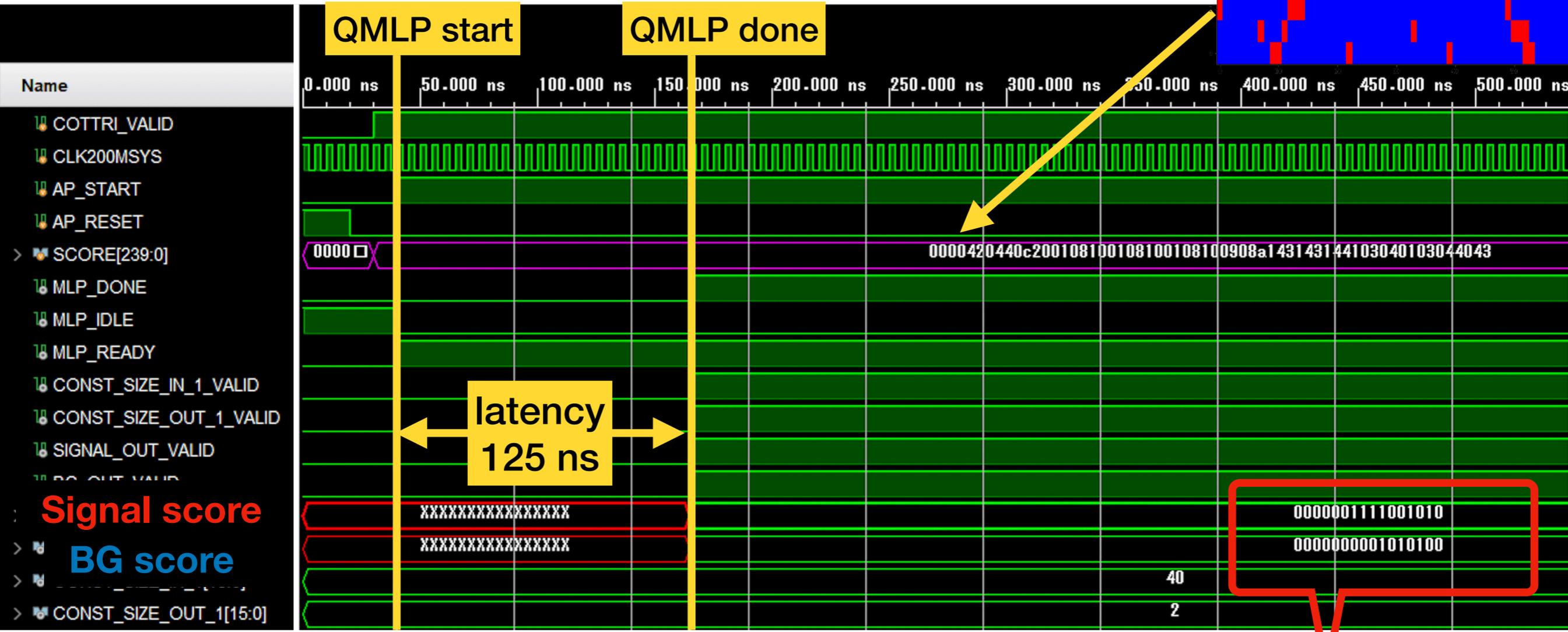
QMLP module simulation

Input event



Xilinx vivado simulationによりQMLP moduleのlatencyと出力を確認

信号イベント情報を入力した時のwaveform

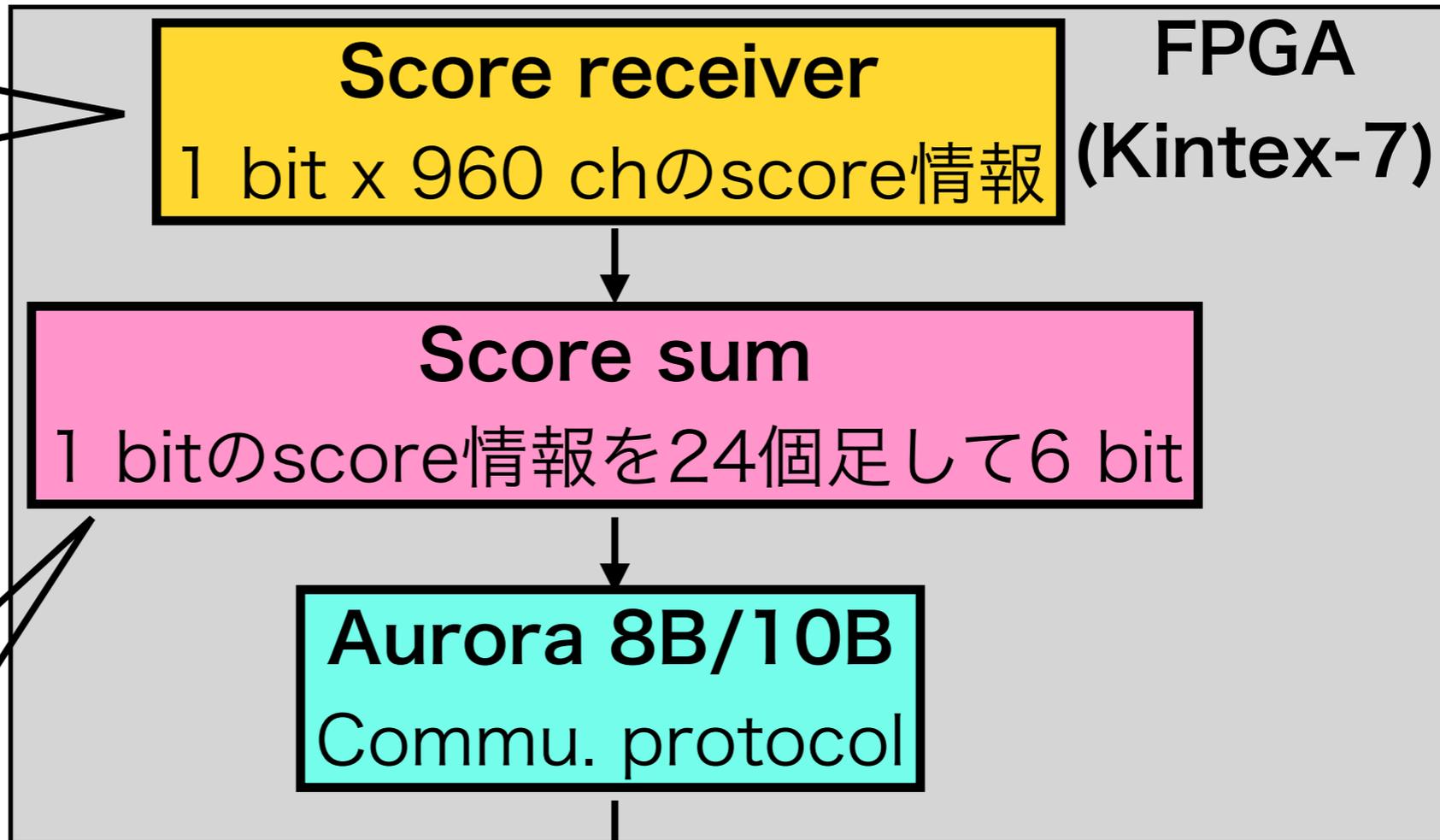
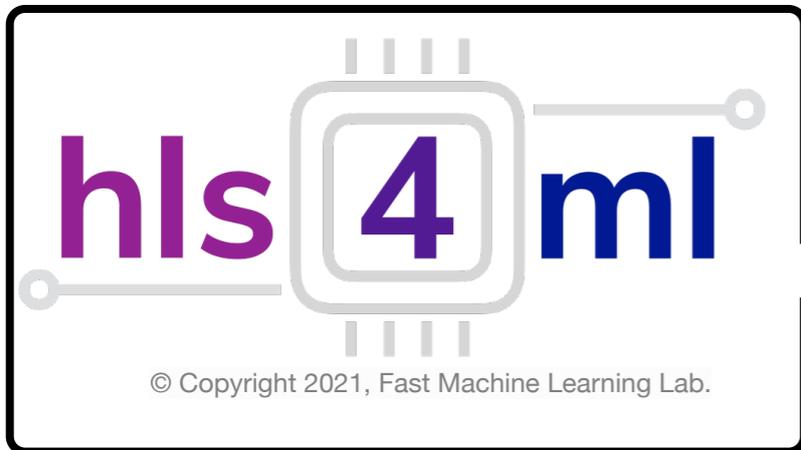


latency, score出力予想通り 😎

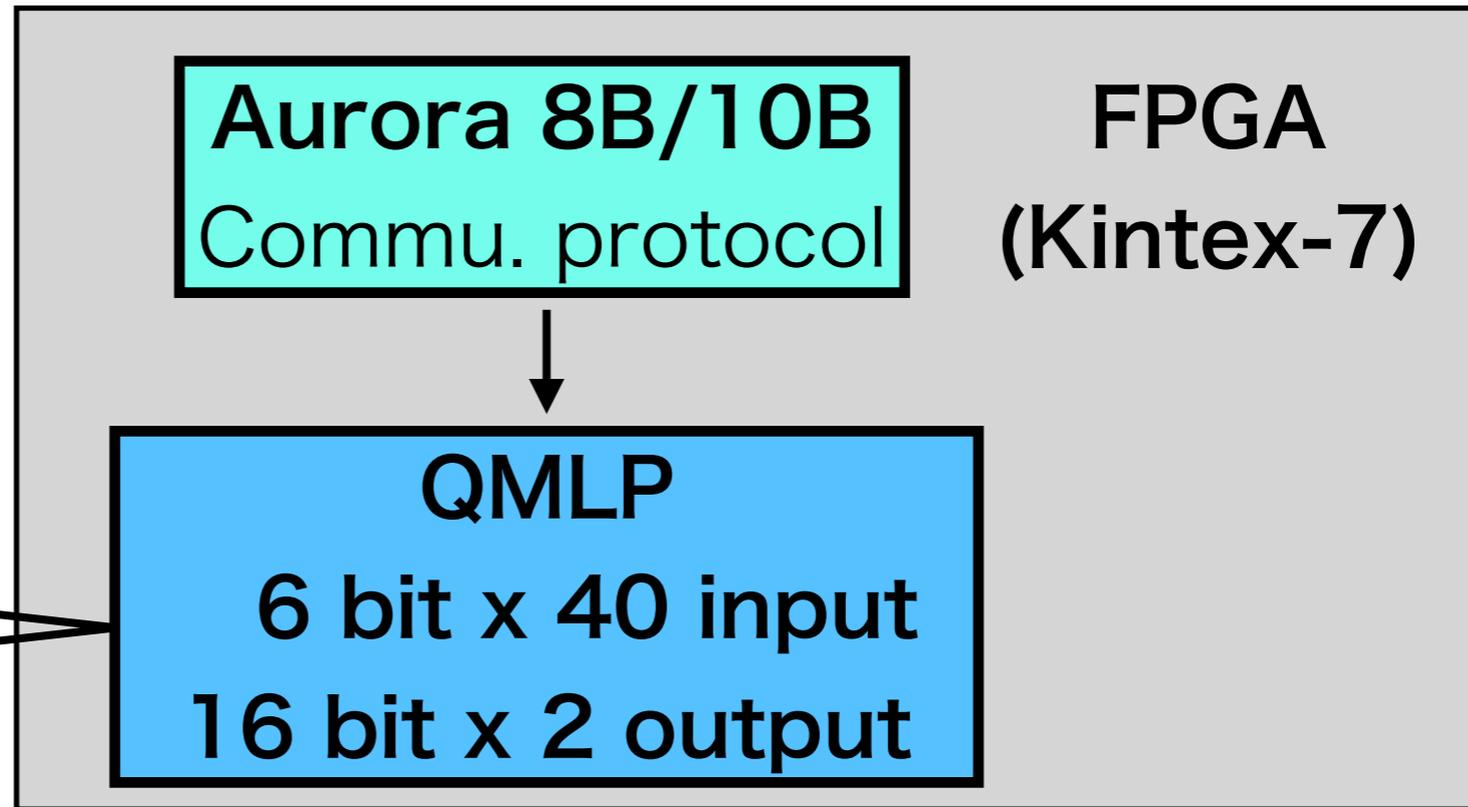
BG score < Signal score
0.08 0.95

分類させたいイベントのscore情報をSiTCPでアドレスに書き込み

FE-MB間のデータ転送は10MHz。
960 bit のscoreを10MHzで転送できないので、6 bit x 40 に圧縮



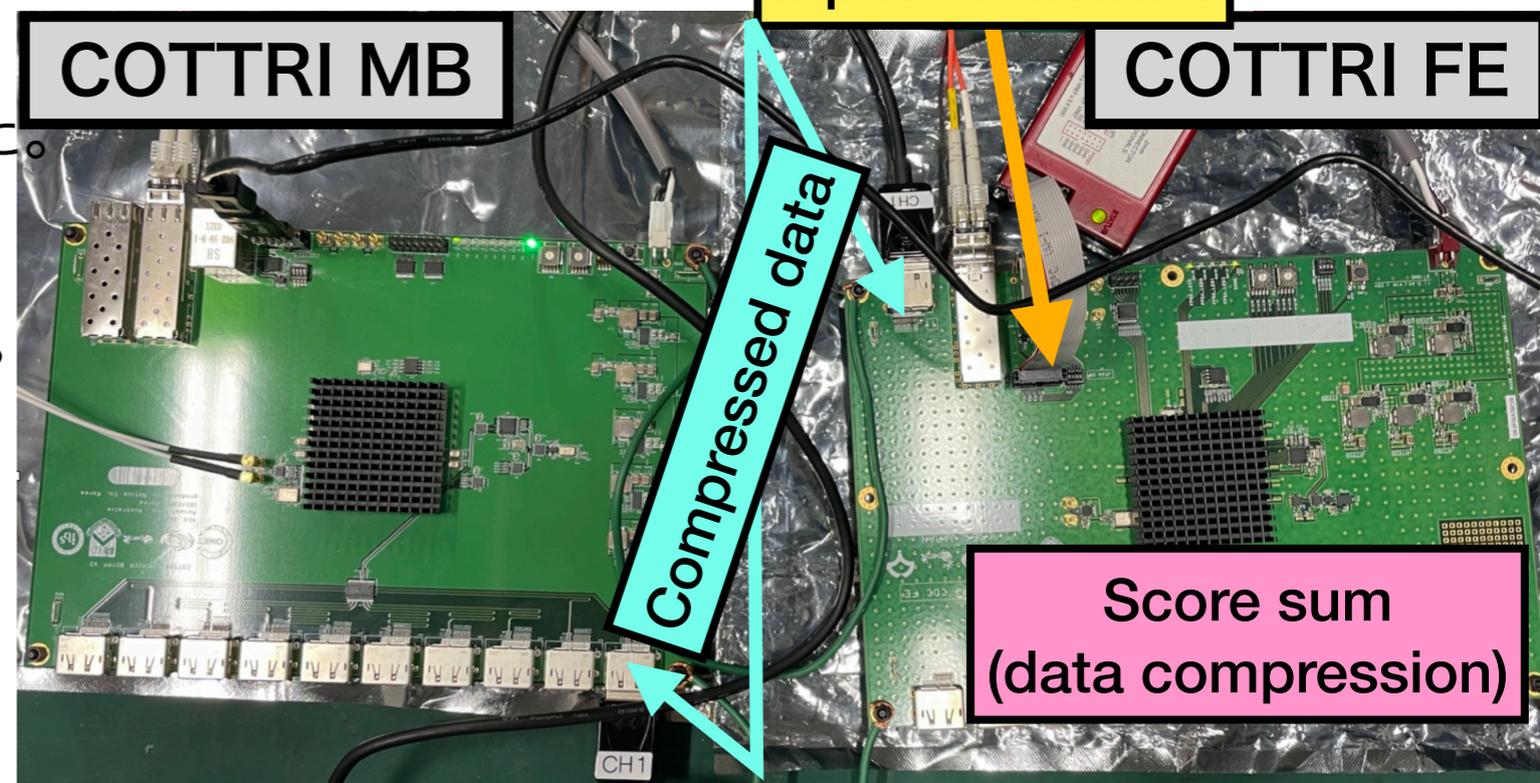
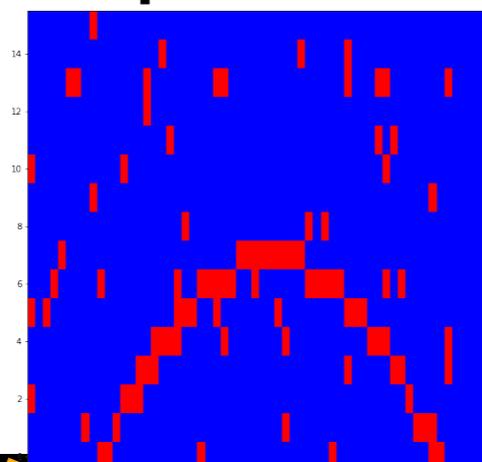
COTTRI MB



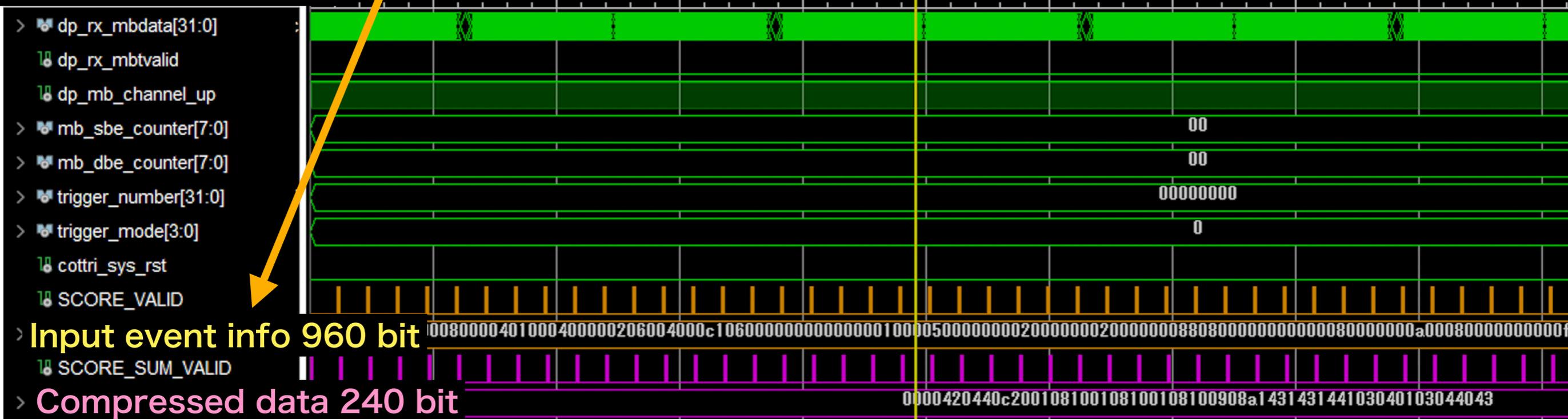
実機試験

- COTTRI FE,MB実機を用いて試験を行った。
- 入力情報はFEにUDP通信で書き込んだ。
- 出力はVivado ILA debug coreで確認した。

Input event



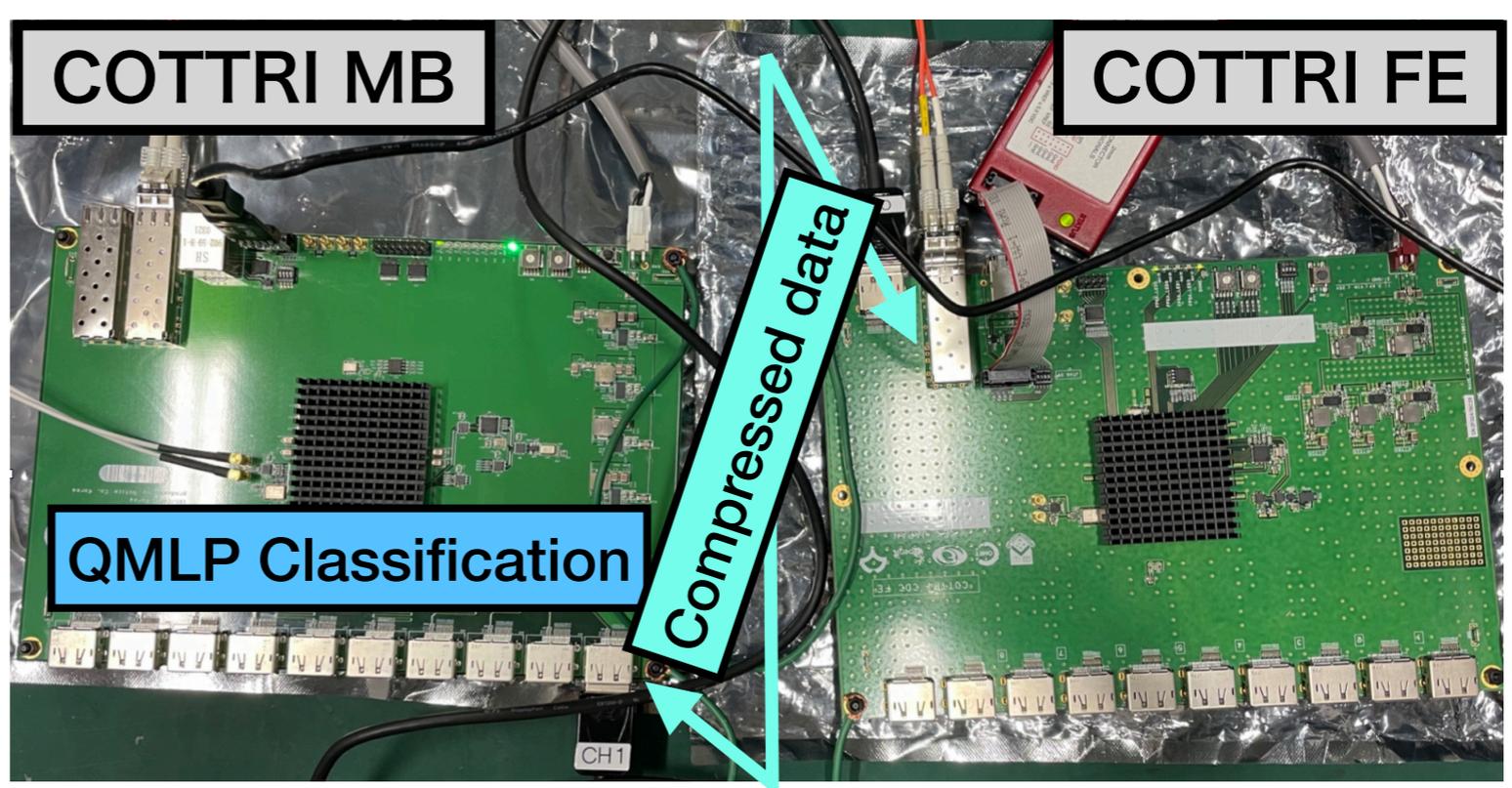
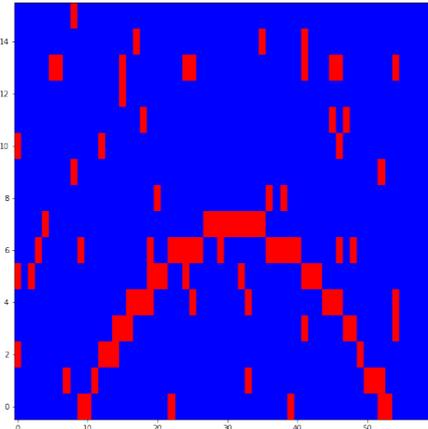
ILA : COTTRI FE



期待通りFWが動いていることを確認🕶️

実機試験

Input event



ILA : COTTRI MB

> NUM_OF_VALID_FE[7:0]

┆ COTTRI_VALID

> DpRxDataOutf239:01

Compressed data 240 bit
from COTTRI FE



┆ MLP_IDLE

┆ MLP_READY

┆ MLP_DONE

┆ SIGNAL_OUT_VALID

┆ BG_OUT_VALID

┆ CONST_SIZE_IN_1_VALID

┆ CONST_SIZE_OUT_1_VALID

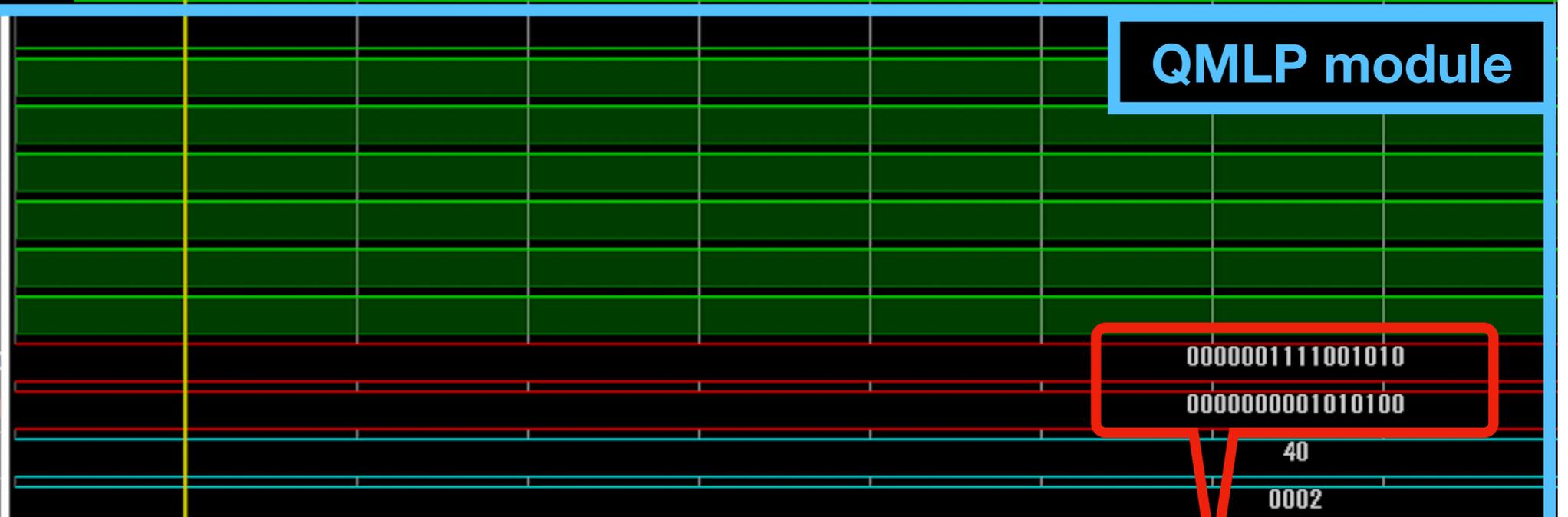
Signal score

BG score

> CONST_SIZE_IN_1[15:0]

> CONST_SIZE_OUT_1[15:0]

QMLP module



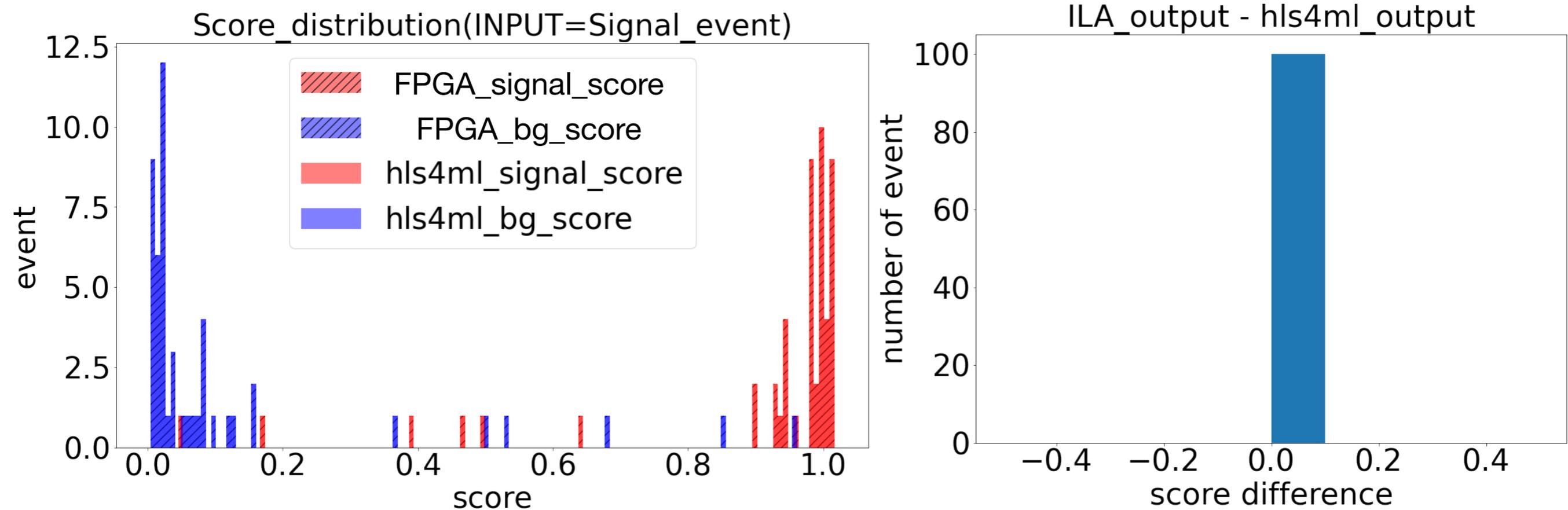
Score出力予想通り 😎

BG score < Signal score

0.08

0.95

QMLPのFPGA実機出力をSignal eventとBG eventの50 eventずつ確認



**hls4ml (software)と、FPGA実機(hardware)の
predictionが完全に一致**

まとめ

- COMET実験は標準理論で強く抑制されている μ -e転換過程探索実験
- COMET Phase-I 開始に向けて鋭意準備中
 - 目標感度 $\sim 7 \times 10^{-15}$ @90 C.L. (AI) ->現在の制限を100倍更新
- シグナルアクセプタンス向上のためにイベント分類Neural Networkをオンライントリガーシステムに実装するスタディが進行中
- 問題設定を簡略化することで以下を確認
 - Software、FPGA firmware両サイドの開発スキームを開発
 - データセットの用意、モデル構築、hls4mlによるRTL作成、VivadoによるSimulation, Neural Network module試験用ファームウェアの開発、FPGA実機出力とSoftware出力の比較
 - FPGA上でのNeural Networkの性能を評価
 - hls4ml(software)のpredictionとFPGA実機(hardware)のpredictionが一致することを確認

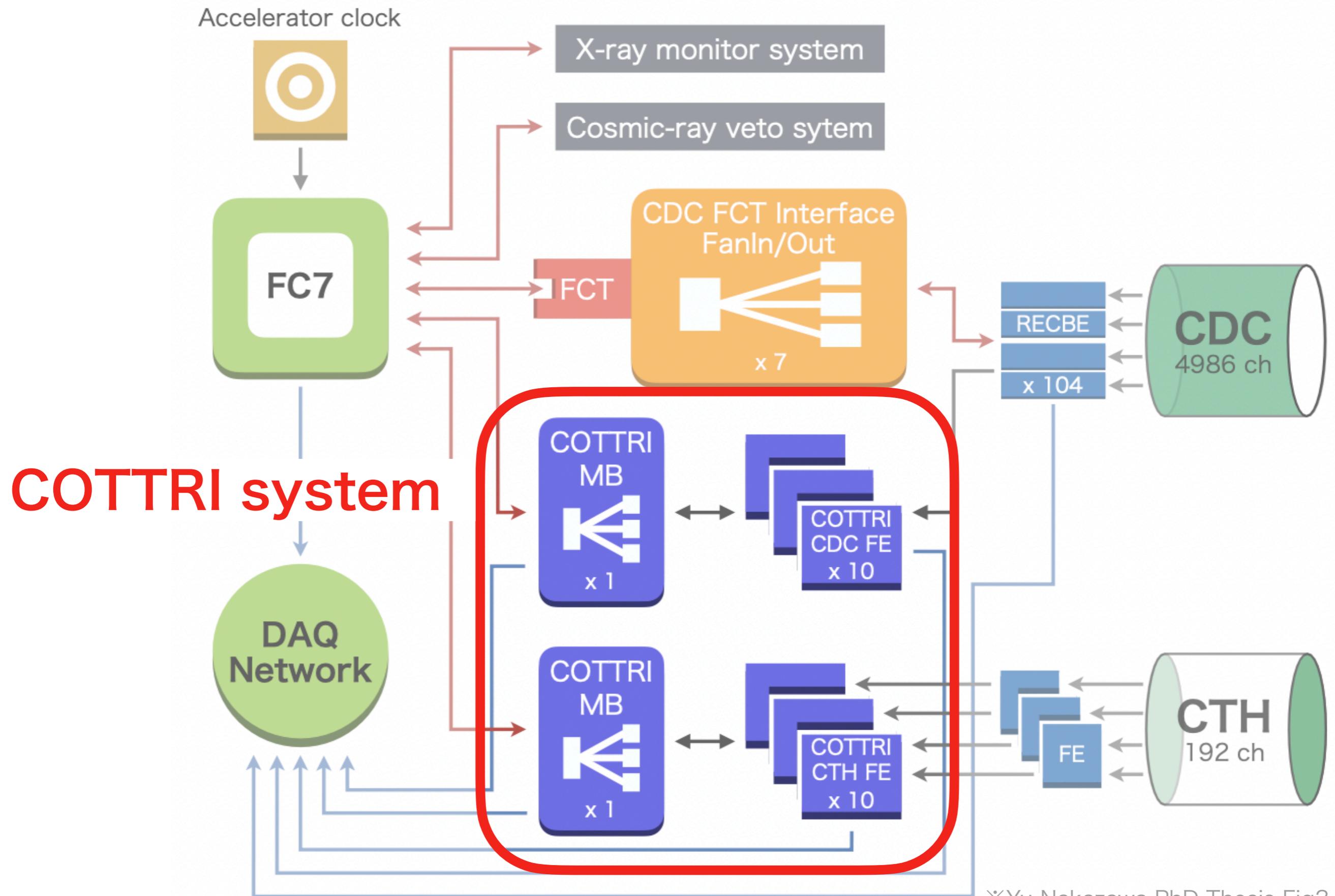
👉 実機でのlatency 測定する

問題設定、ハードウェア設定を現実のセットアップにスケールアップしていく

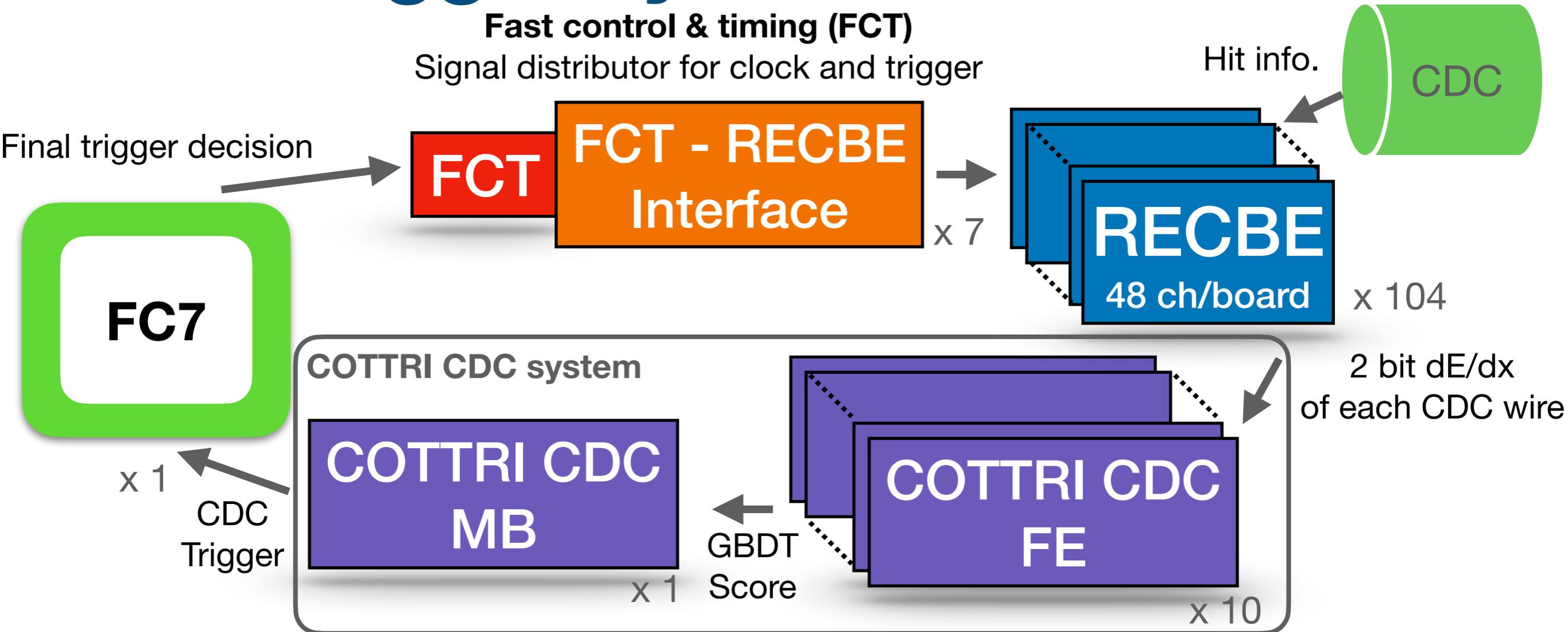
MBのリソースはまだ余裕があるのでよりdeepなNeural networkやCNN,GNN等を試す

Backup

Online trigger system



Online trigger system

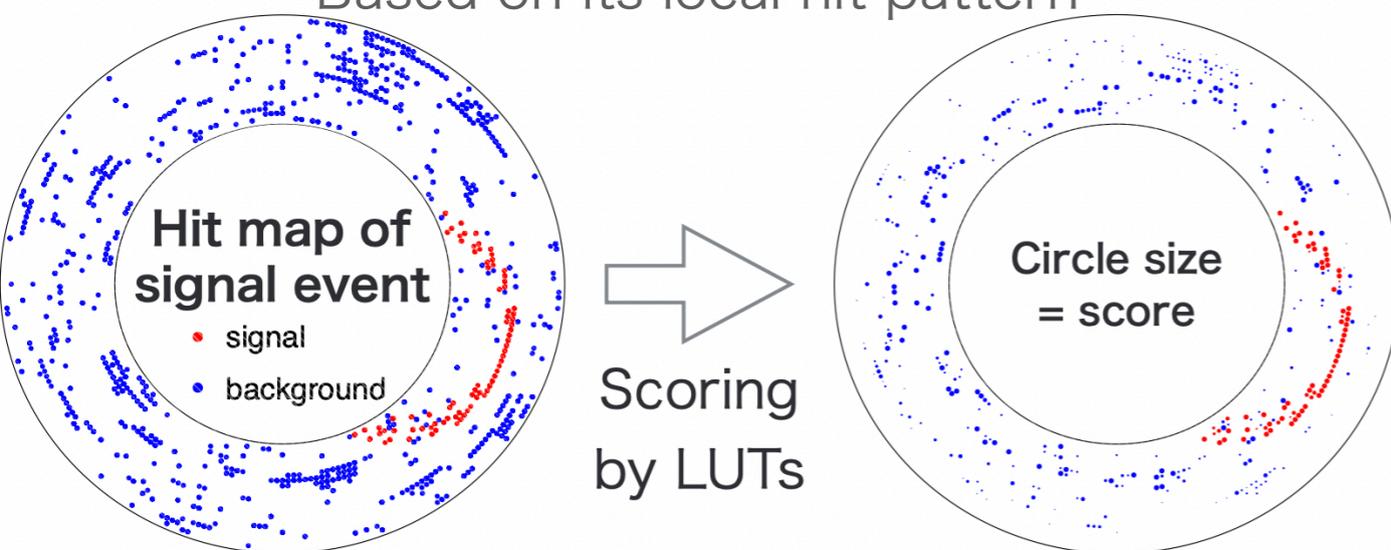


- RECBEs generate the 2 bit dE/dx information and send it @10 MHz
- COTTRI CDC system
 - FE : hit classification based on local/neighboring features.
 - Convert 2 bit data to GBDT scores in 400 ns integration time window
 - MB : event classification with the global feature.
 - Sumsup the GBDT scores and makes the CDC trigger decision @10 MHz

Trigger algorithm

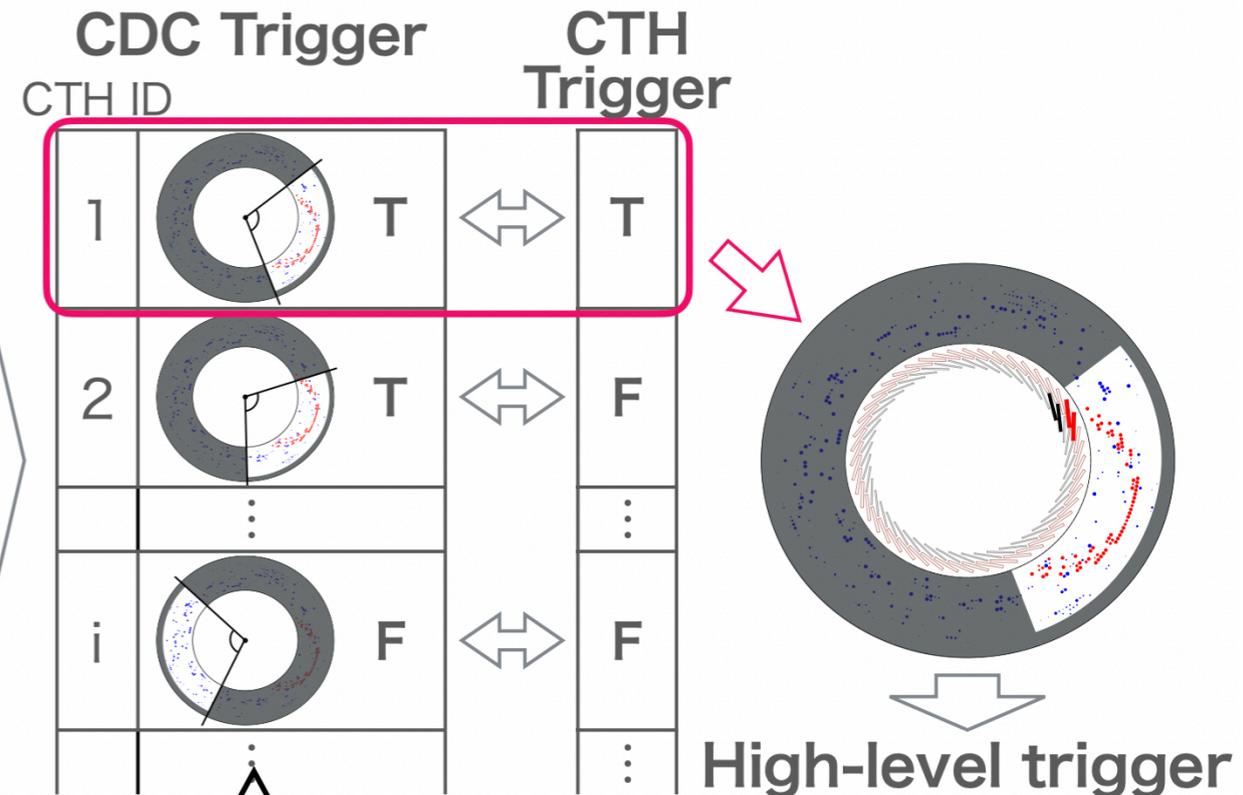
Hit classification

GBDT gives each wire hit a score
Based on its local hit pattern



Yu nakazawa PhD thesis Fig3.5

Event classification



The current CDC trigger algorithm

1. Set the CDC active section for each segment of CTH
2. Within each active section, count hits that exceed the score threshold
3. CDC trigger is issued when the count exceeds the threshold.

New CDC trigger algorithm under development

1. Set the CDC active section for each segment of CTH
- 2'. Execute Neural Network inference w/ score information of each active section as inputs
- 3'. CDC Trigger is issued based on Neural Network classification.

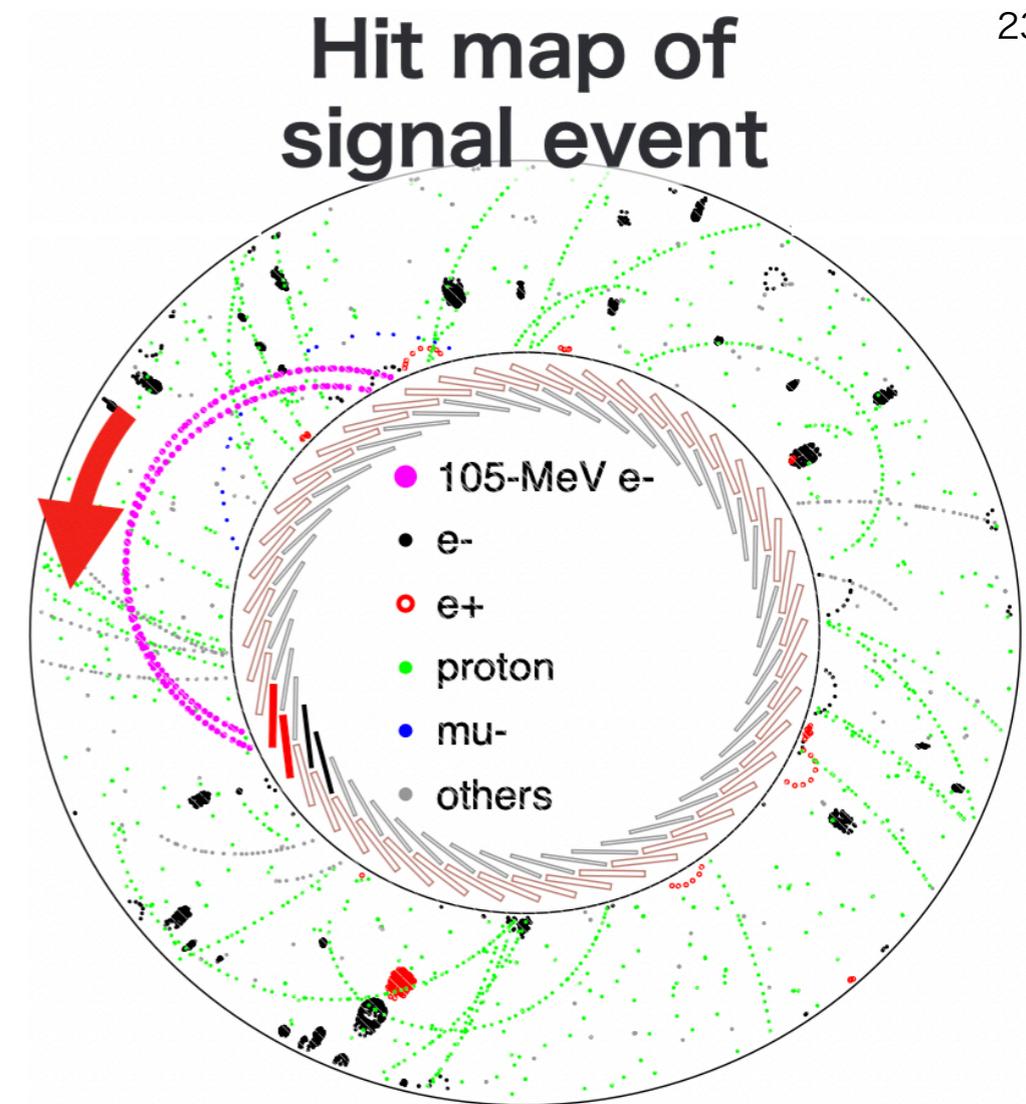
Signal and BG hits

Signal-hit characteristics

- **Contained helical tracks**
- **Single hit in the same wire**
- **MIP-level energy loss**

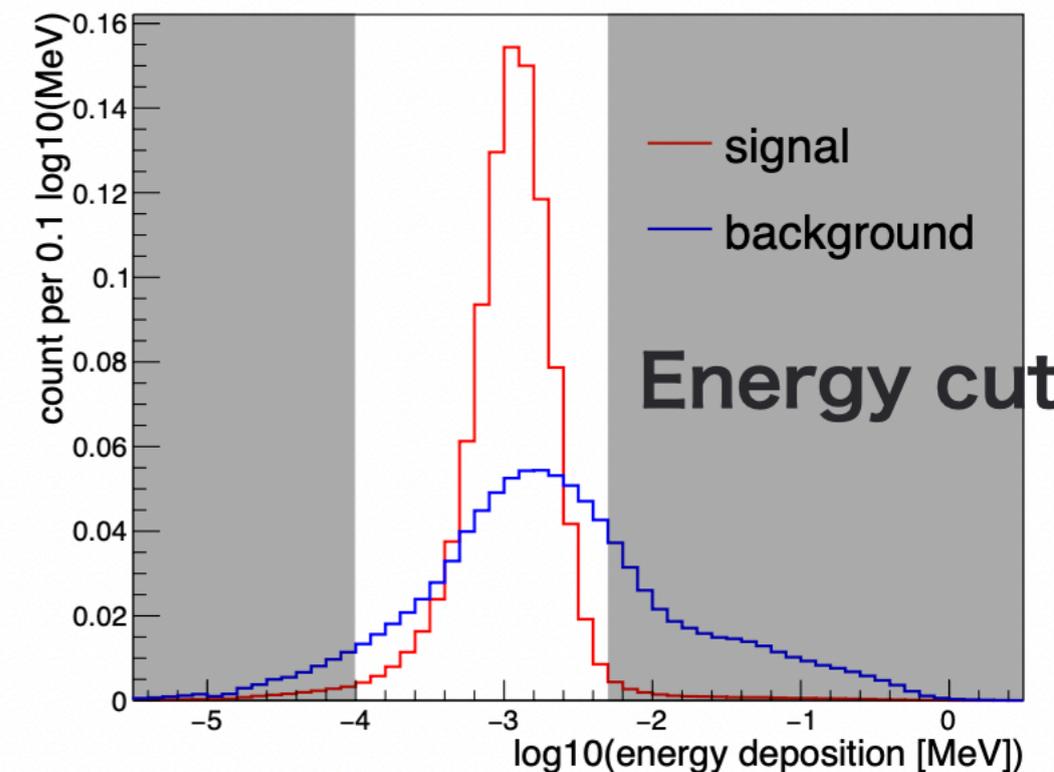
Background-hit characteristics

- Low energy electrons
 - Interaction of gamma rays at the CDC walls
 - Helical trajectory contained in the same cell
 - **Multi hits in the same wire**
- Protons (from muon nuclear capture)
 - Momentum higher than 100MeV/c
 - **Large energy loss**
 - ~40 protons/beam-pulse



✧ Yu Nakazawa PhD Thesis Fig3.2

Energy deposition

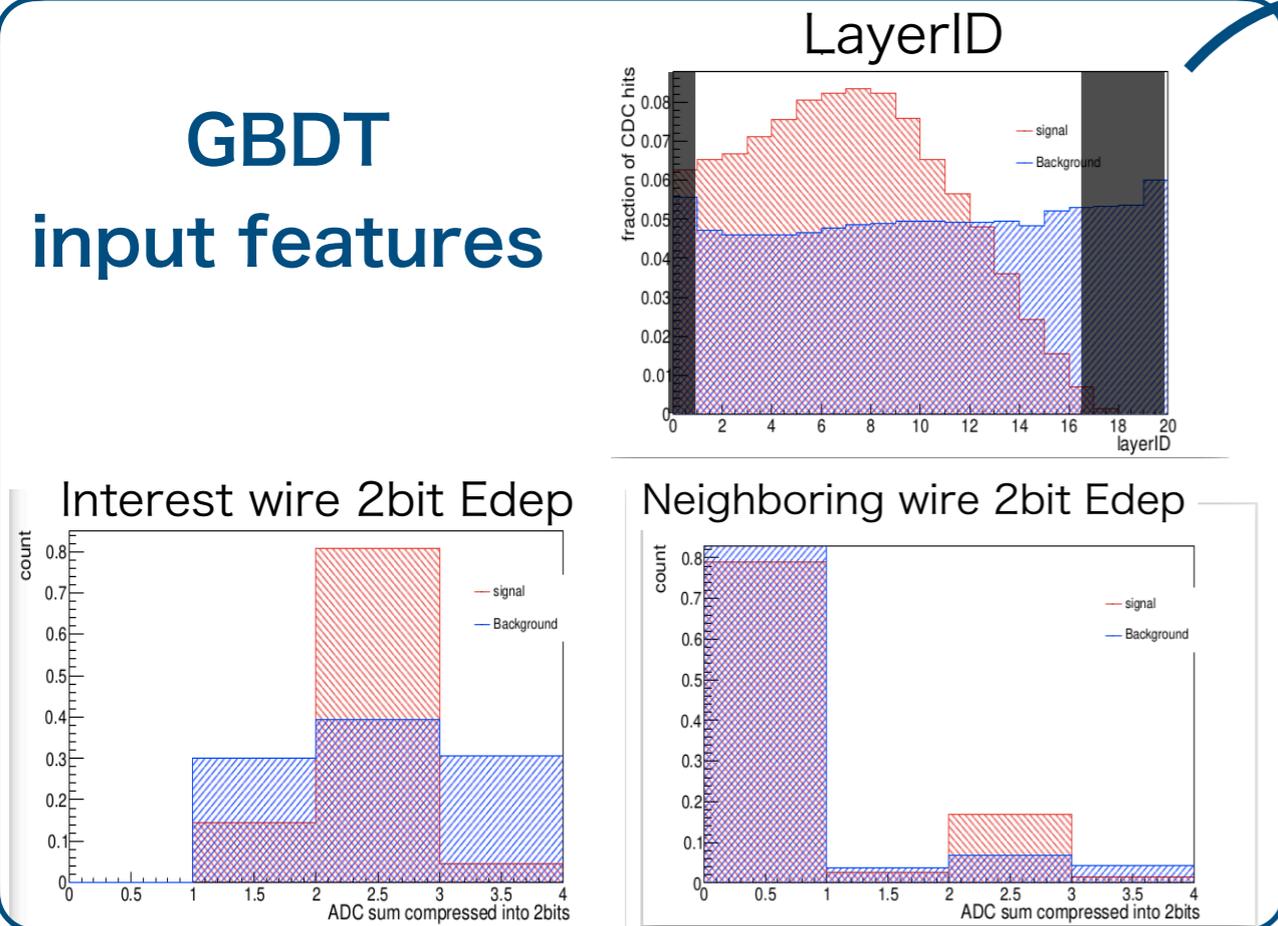


Hit classification

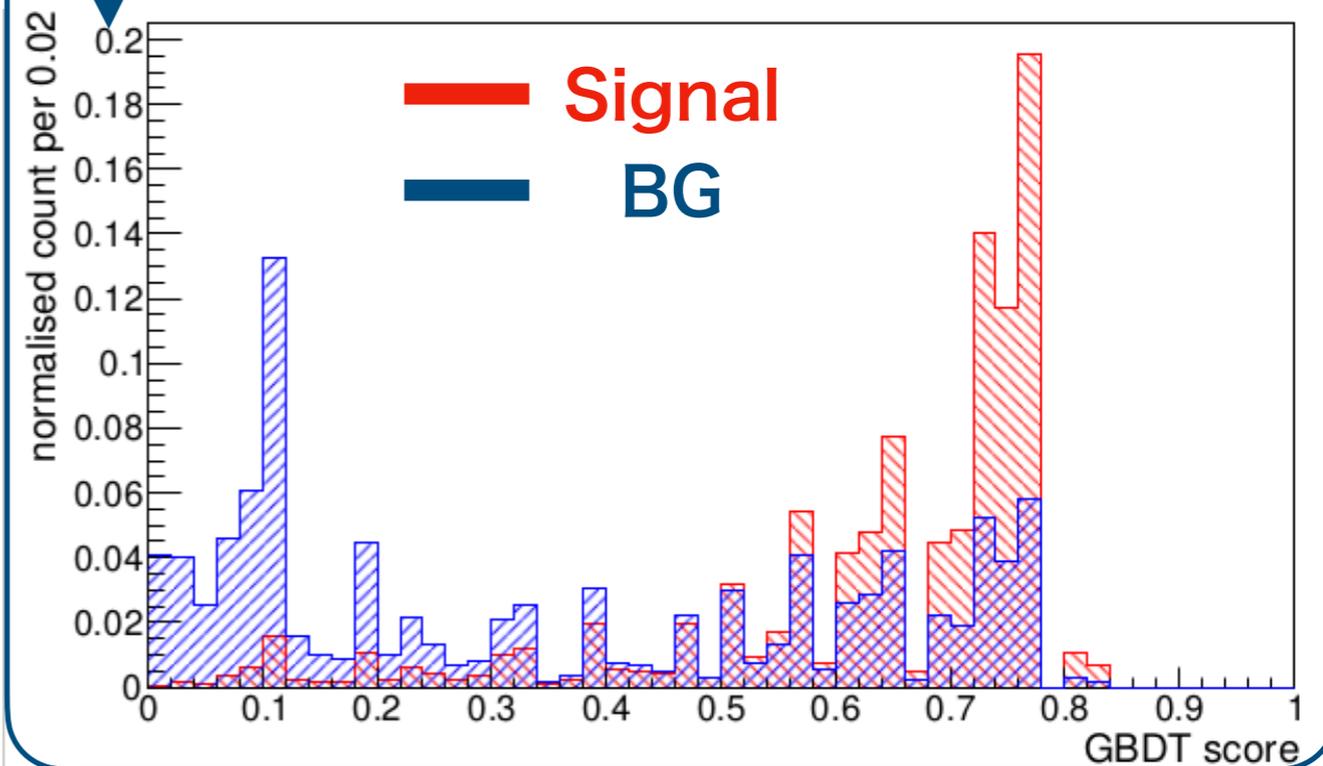
*GBDT = Gradient Boosted Decision Tree

- Machine learning algorithm (GBDT*) to score hit information for each wire based on energy loss and local patterns

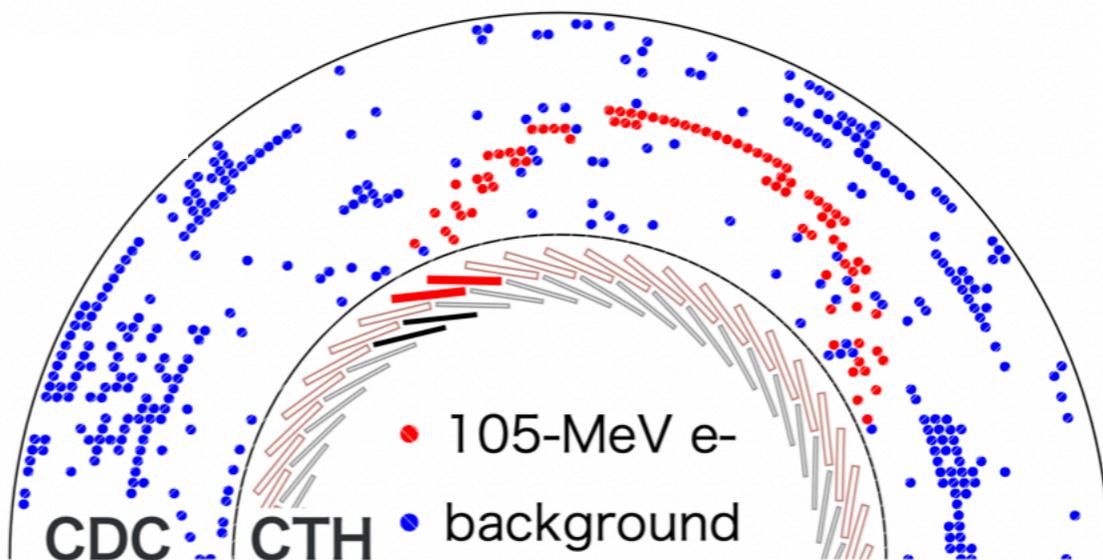
GBDT input features



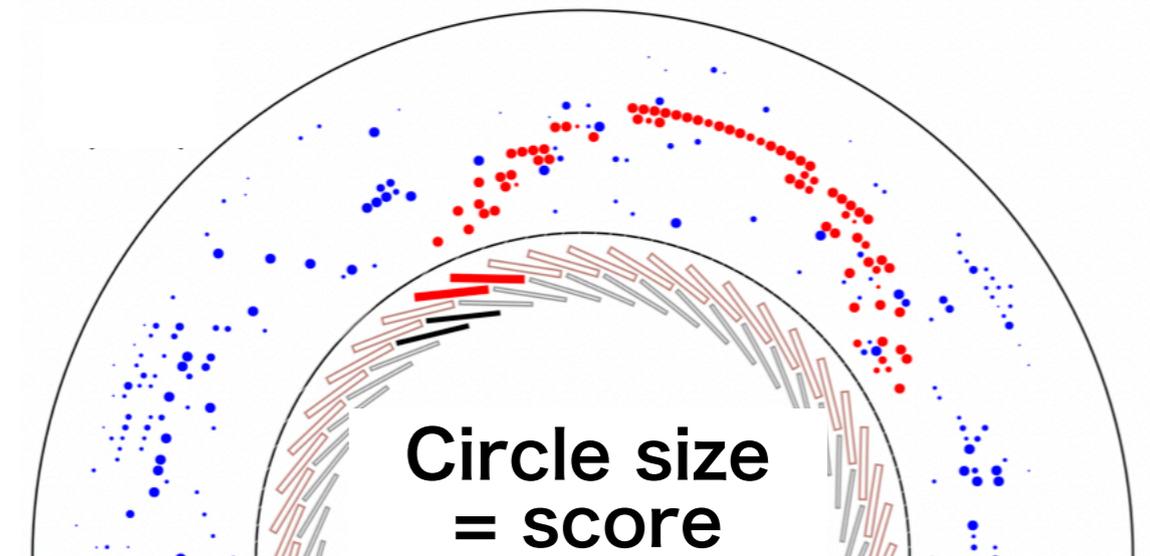
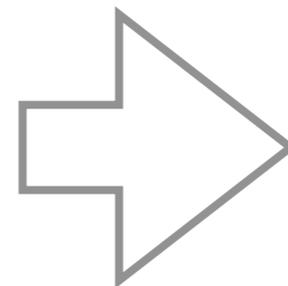
GBDT output score distribution



Hit map of signal event

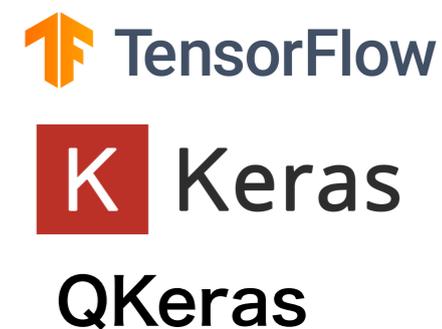
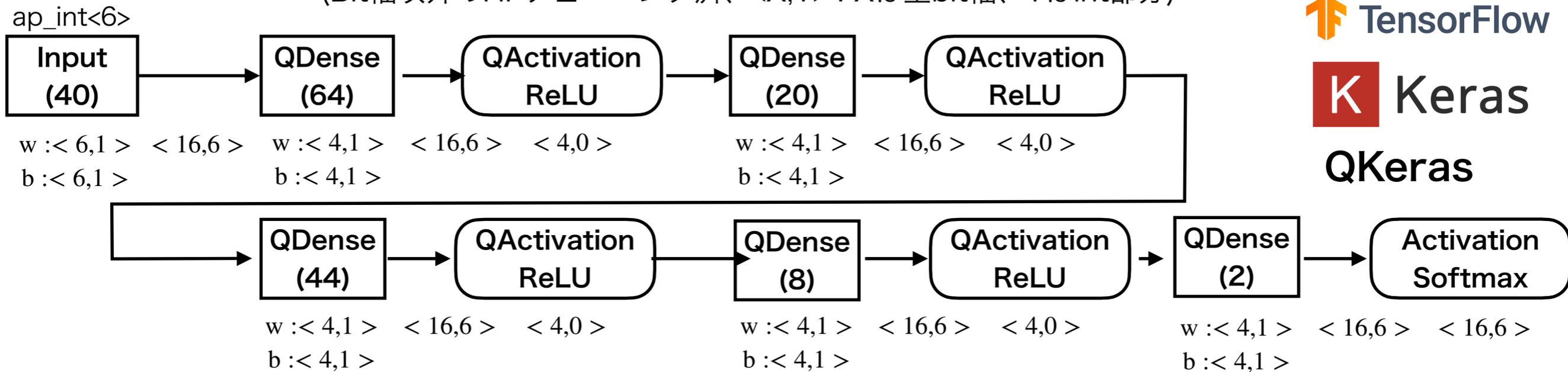


Scoring By GBDT



	CDC config	BG hit 占有率	ワイヤーヒット score情報	Active section
現実	20 layer x ~250 cell	~20 %	6 bit	~1500 ch
本スタディ	18 layer x 180 cell	5%	1 bit	960 ch

(Bit幅以外のHPチューニング済、 $\langle X, Y \rangle$: Xは全bit幅、Yはint部分)

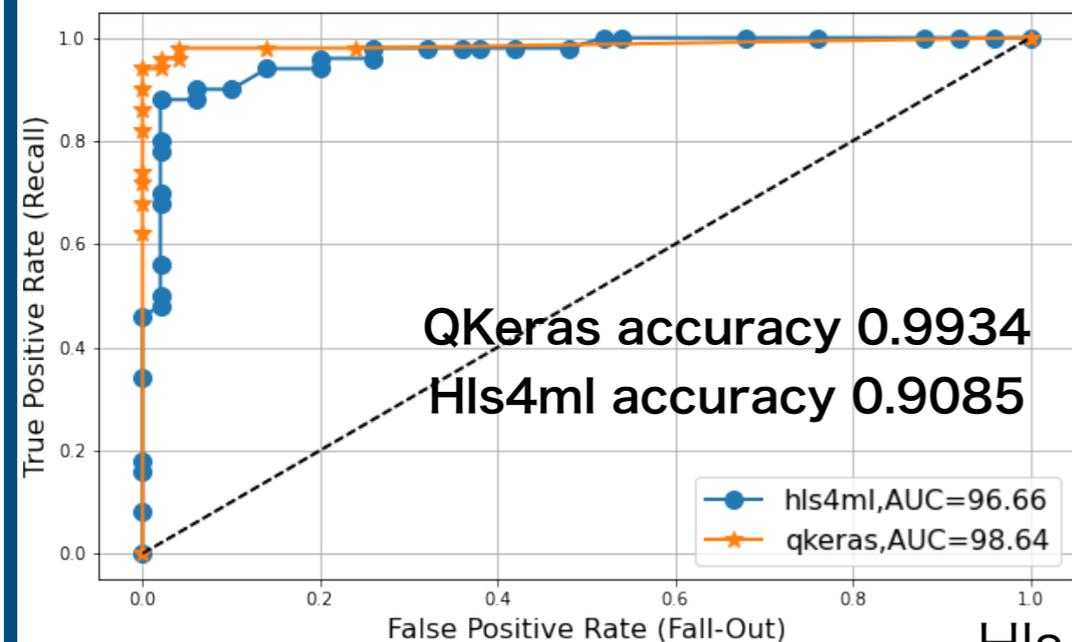


© Copyright 2021, Fast Machine Learning Lab.

モデルをハードウェア記述言語によるプログラミングなしに
ファームウェア変換可能なコード(RTL)に変換

<https://dx.doi.org/10.1088/1748-0221/13/07/P07027>

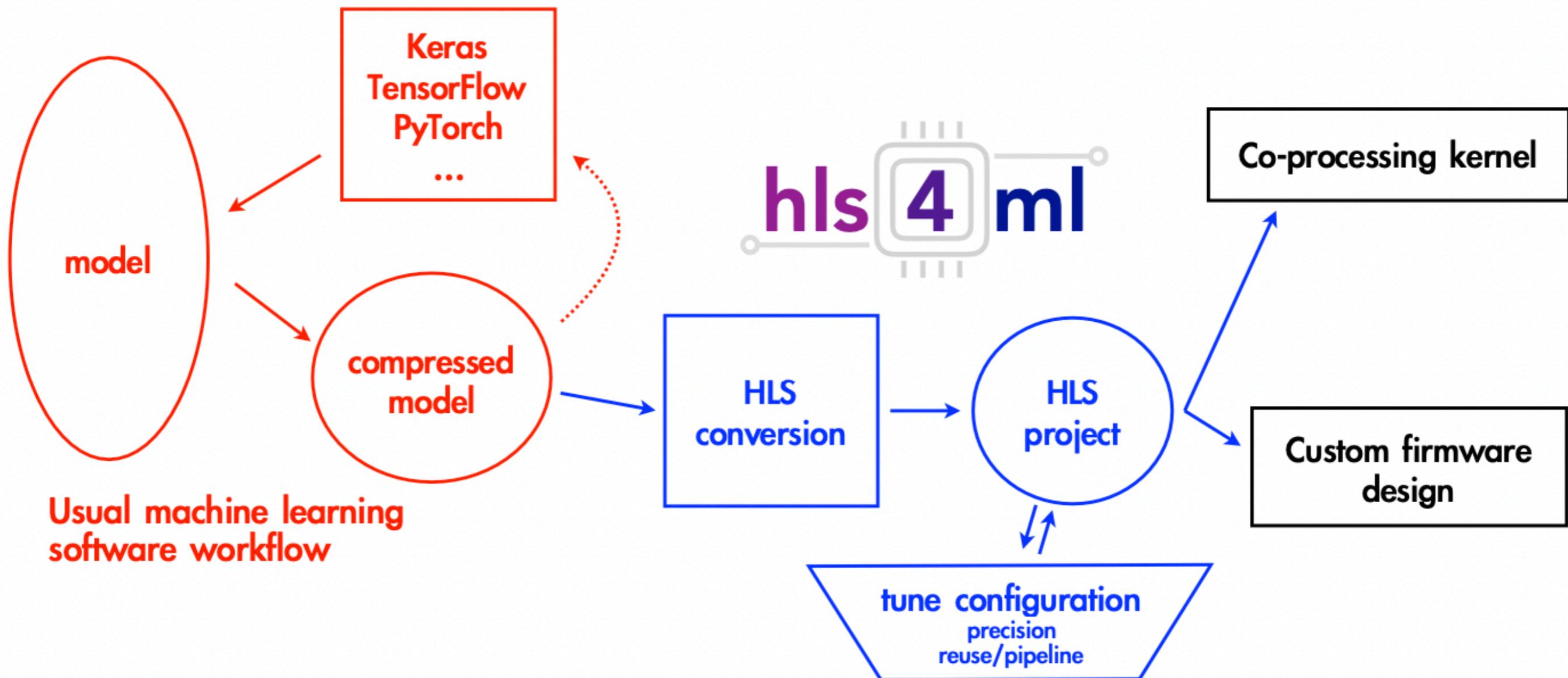
FPGA : AMD Xilinx Kintex-7 xck355t-ffg901-1



Latency @200MHz	Usage (%)			
	BRAM	DSP	FF	LUT
130 ns	~0	~0	5	32

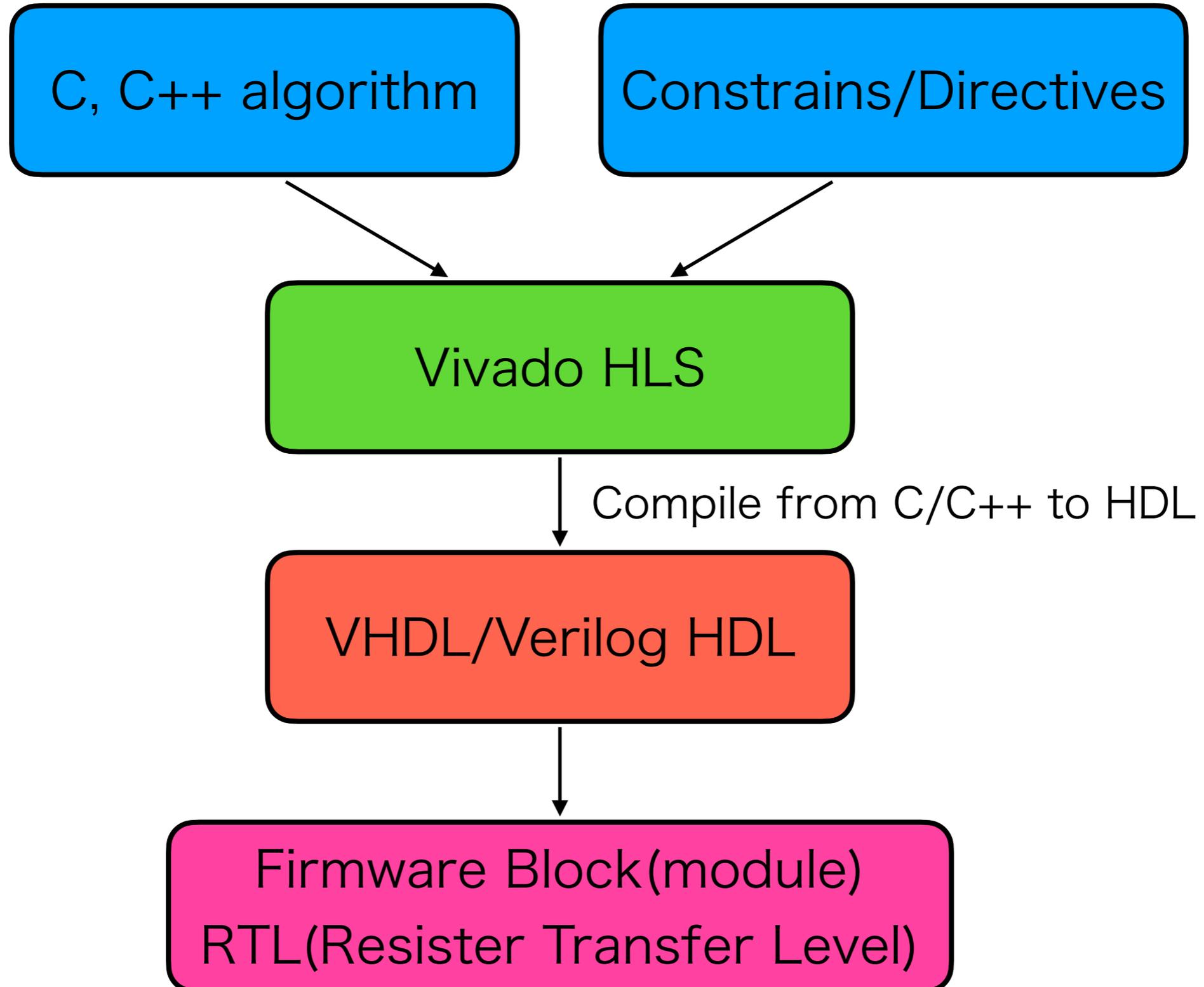
Hls4mlが生成したC++ファイルをVivado_hlsで高位合成した後、vivadoでこのQMLPのipを生成し、COTTRI MBのFWに組み込んだ

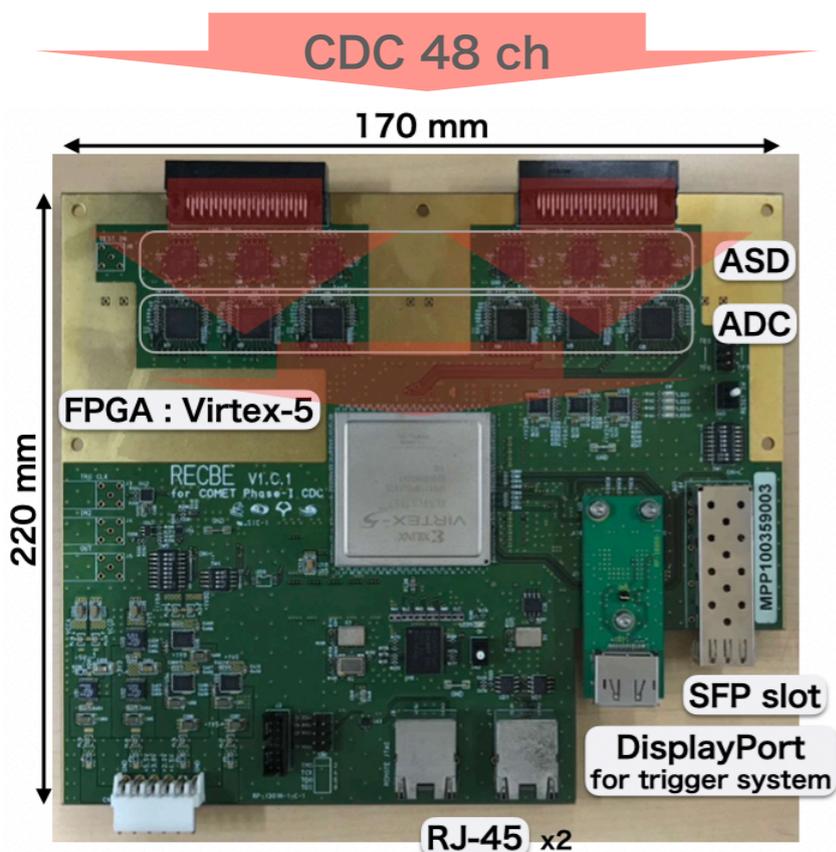
High level synthesis for machine learning 27



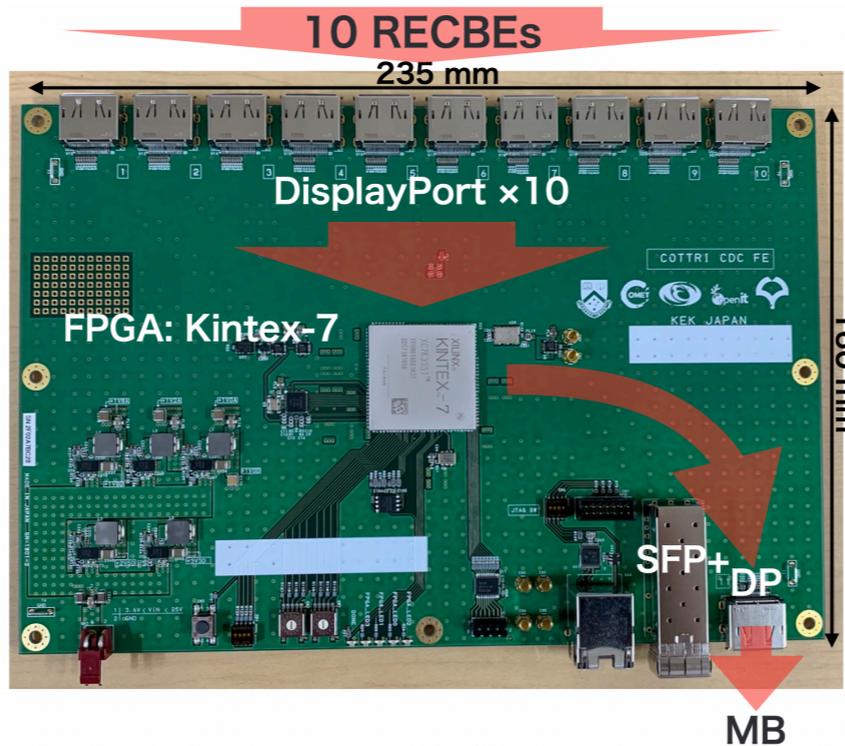
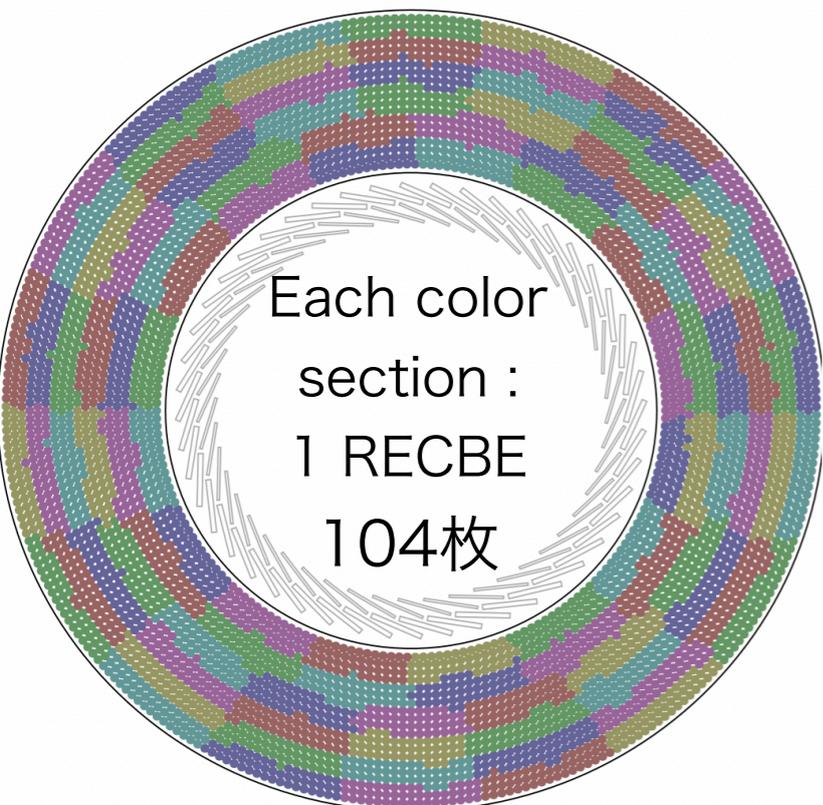
Fast inference of deep neural networks in FPGAs for particle physics Fig1.

FPGA programming Flow

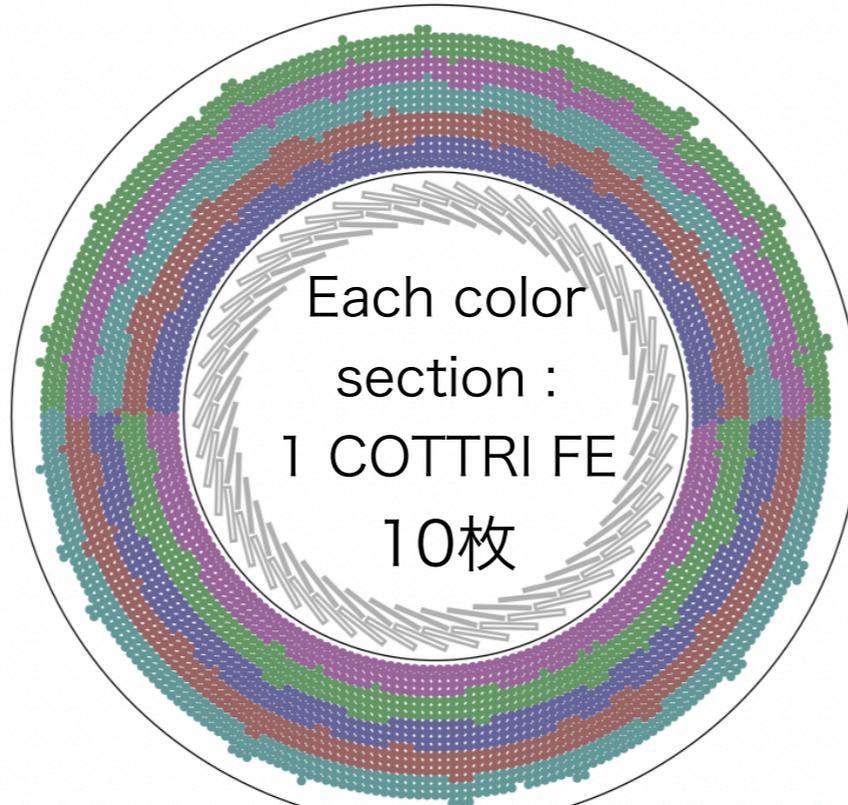




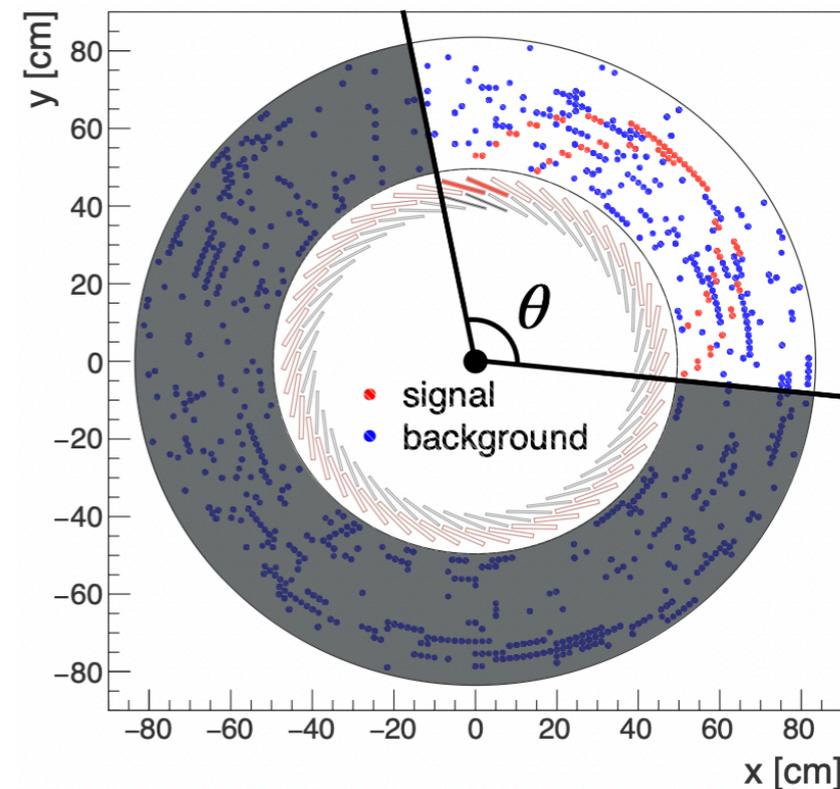
RECBE Configuration



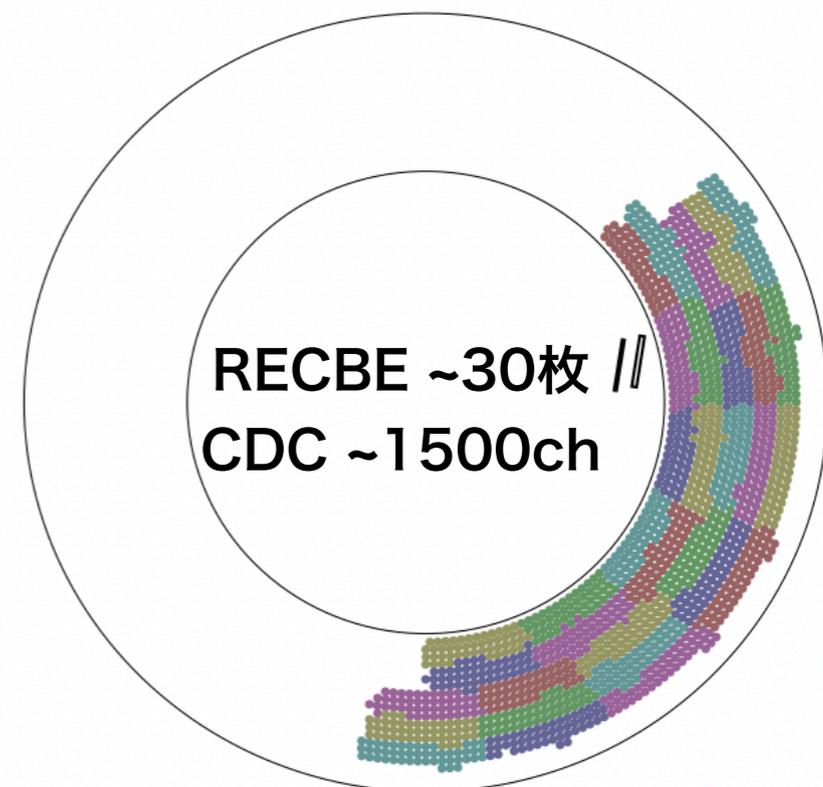
COTTRI FE Configuration



信号電子が軌跡を残すのは CDCの約1/3領域



Active sectionの例

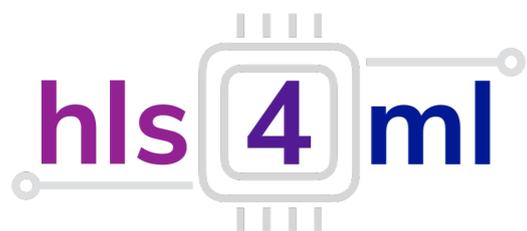


Firmware

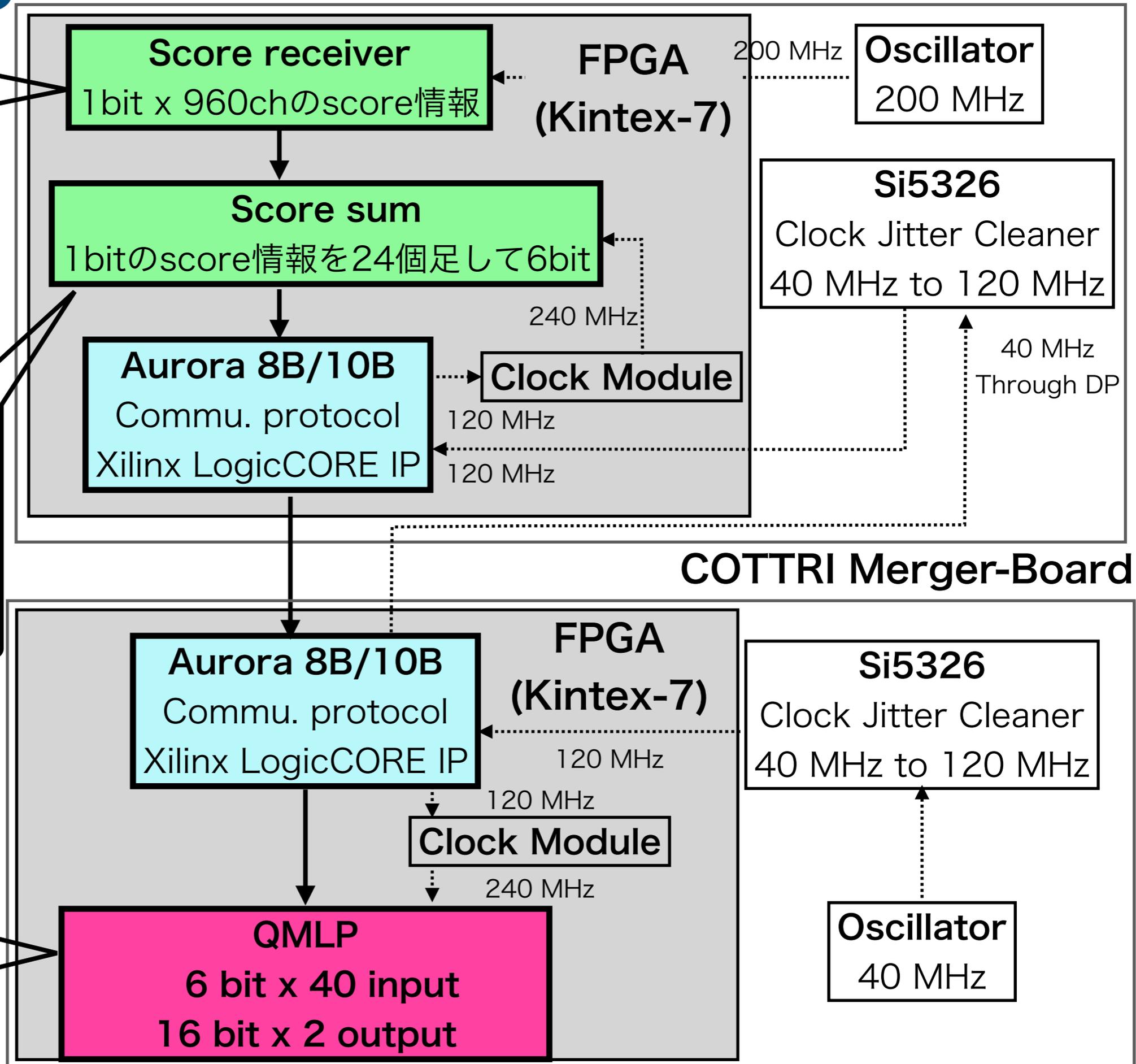
COTTRI Front-End³⁰

Active section
(CDC wire 960 ch)
の1bit score情報を
SiTCPでアドレスに書
き込み

FE-MB間のデータ転送は
10MHz。
960 bit のscoreを
10MHzで転送できないの
で、6bit x 40 に圧縮



© Copyright 2021, Fast Machine Learning Lab.



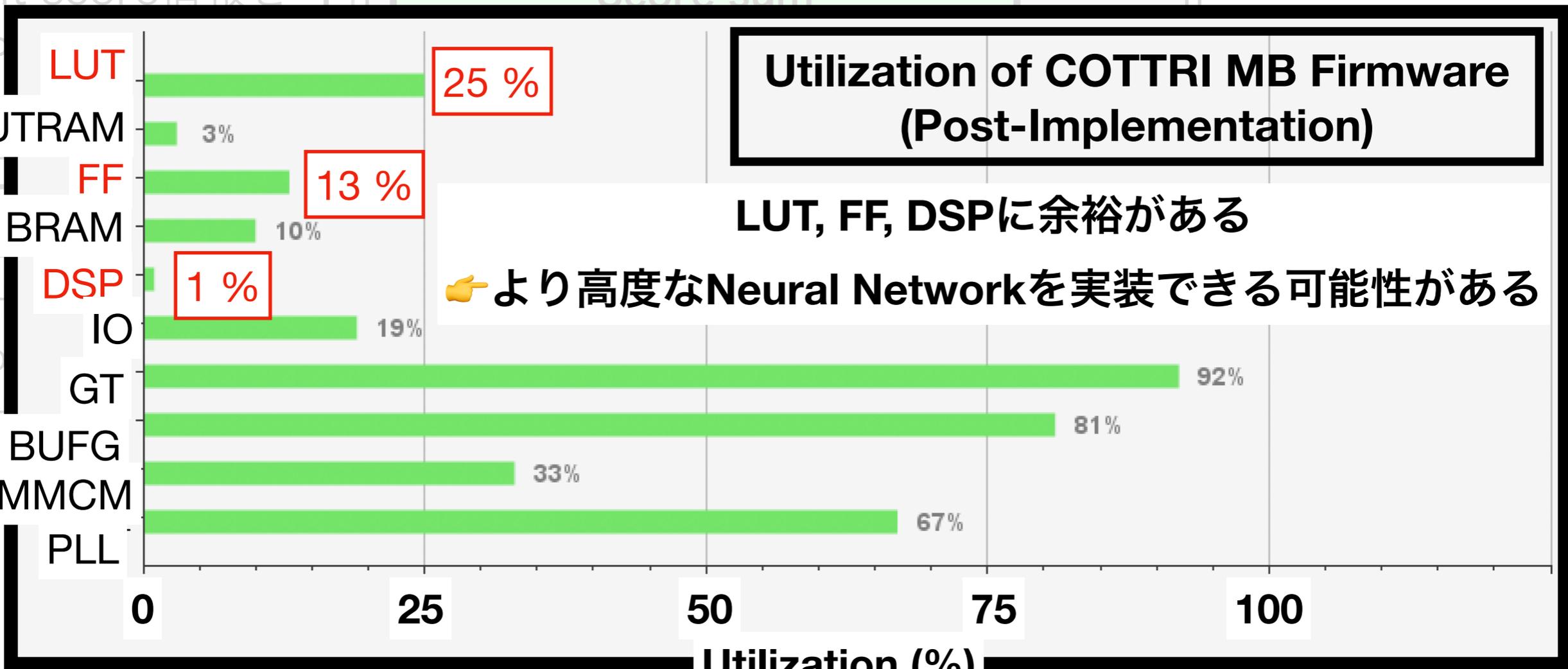
Active section
(CDC wire 960 ch)
の1bit score情報を
SiTCP



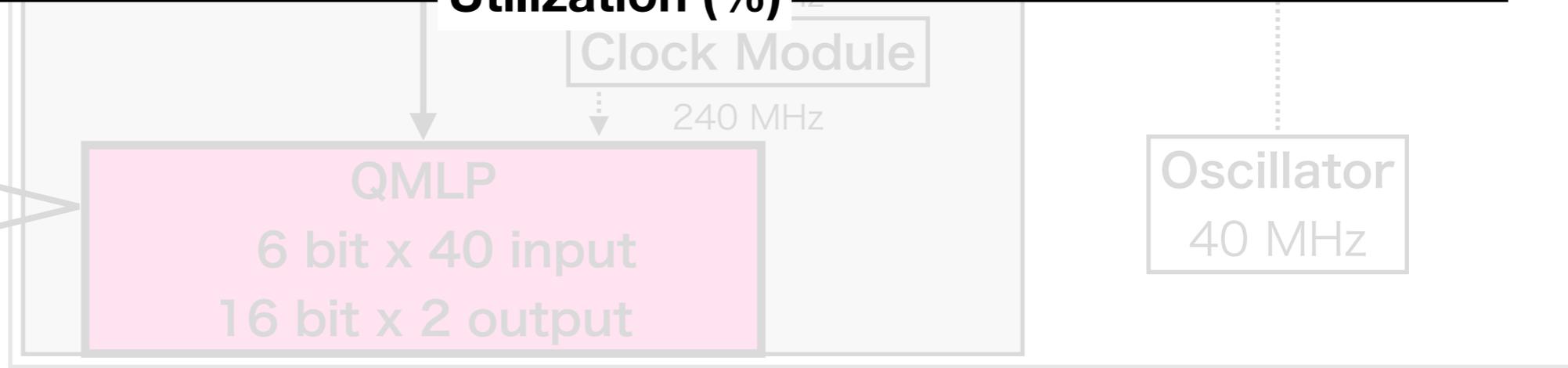
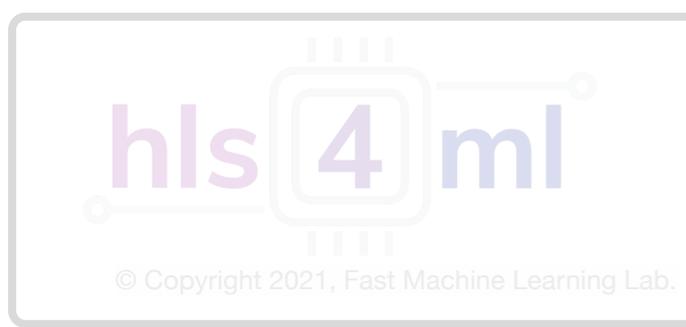
Utilization of COTTRI MB Firmware (Post-Implementation)

LUT, FF, DSPに余裕がある

👉 より高度なNeural Networkを実装できる可能性がある



FE-ML
10MHz
960 b
10MHz
で、6



- Maximum data transfer = 2.4 Gbps/lane x 2lane x 0.8 = 3.84 Gbps
- Data format : 1 header & 10 data packets

1 frame @ 10 MHz

Bit	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
Header	0	Parity bits			1	Sent number												Board ID														
Score	0	Parity bits			0	RECBE 9																										
	0	Parity bits			0	RECBE 8																										
	0	Parity bits			0	RECBE 7																										
	0	Parity bits			0	RECBE 6																										
	0	Parity bits			0	RECBE 5																										
	0	Parity bits			0	RECBE 4																										
	0	Parity bits			0	RECBE 3																										
	0	Parity bits			0	RECBE 2																										
	0	Parity bits			0	RECBE 1																										
	0	Parity bits			0	RECBE 0																										

COTTRI FE to COTTRI MB data format

For the preliminary study

1 frame @ 10 MHz

Bit	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
Header	0	Parity bits			1	Sent number											Board ID															
Score	0	Parity bits			0	Input39			Input38			Input37			Input36																	
	0	Parity bits			0	Input35			Input34			Input33			Input32																	
	0	Parity bits			0	Input31			Input30			Input29			Input28																	
	0	Parity bits			0	Input27			Input26			Input25			Input24																	
	0	Parity bits			0	Input23			Input22			Input21			Input20																	
	0	Parity bits			0	Input19			Input18			Input17			Input16																	
	0	Parity bits			0	Input15			Input14			Input13			Input12																	
	0	Parity bits			0	Input11			Input10			Input9			Input8																	
	0	Parity bits			0	Input7			Input6			Input5			Input4																	
	0	Parity bits			0	Input3			Input2			Input1			Input 0																	

$$precision = \frac{TP}{TP + FP}$$

陽性予測の正解率

$$recall = \frac{TP}{TP + FN}$$

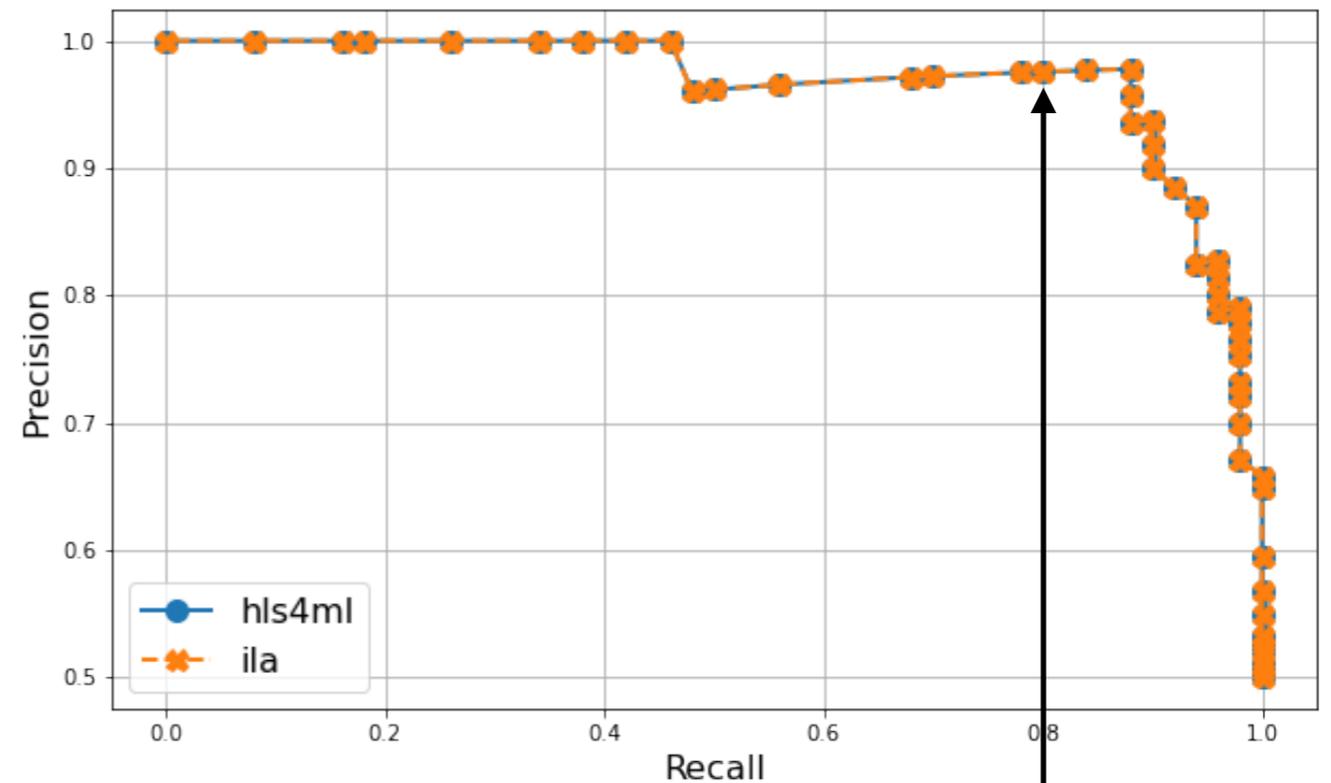
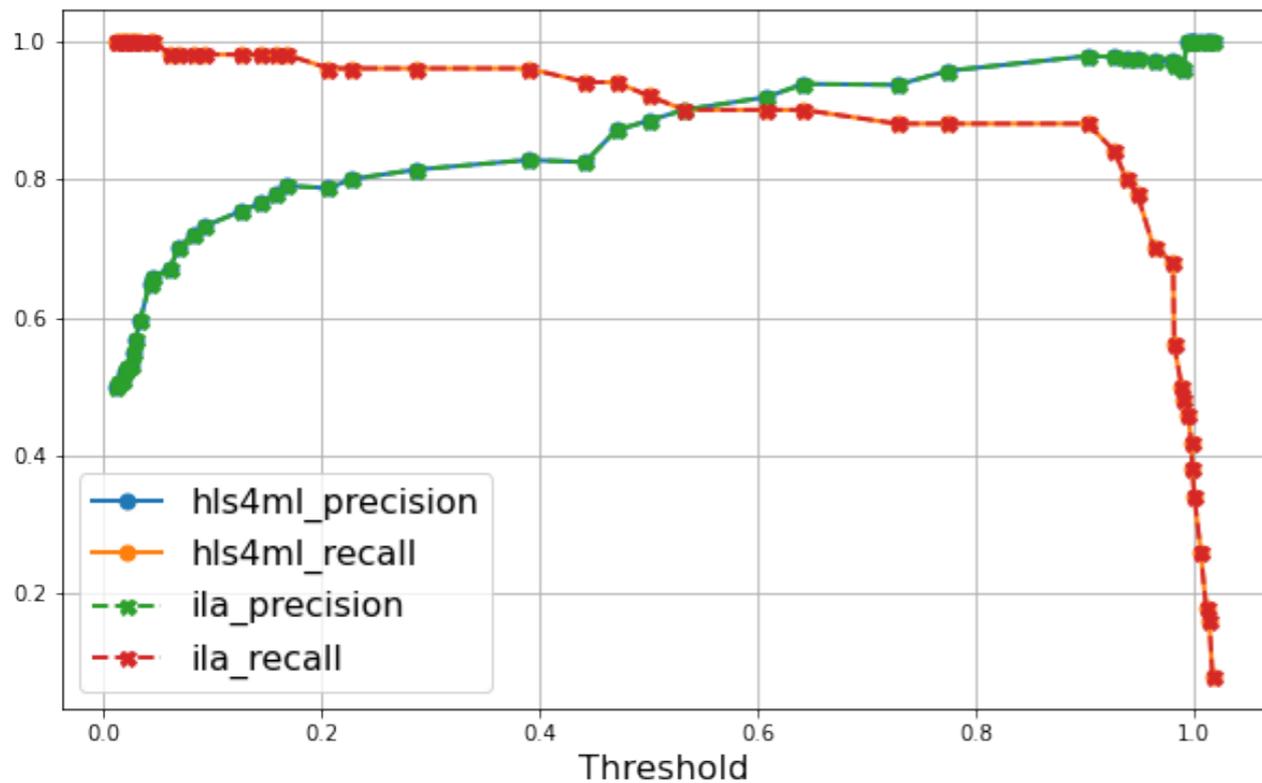
正しく分類した陽性の割合

TP : 真陽性の数 Signal eventをsignal eventであると分類した数

TN : 新陰性の数 BG eventをBG eventであると分類した数

FP : 偽陽性の数 BG eventをsignal eventであると分類した数

FN : 偽陰性の数 Signal eventをBG eventであると分類した数



分類器がSignal eventであると予測したとき、その予測が正しいのは98%
 全てのSignal eventの80%を検出

$$TPR = \frac{TP}{TP + FN}$$

正しく分類した陽性の割合

$$FPR = \frac{FP}{TN + FP}$$

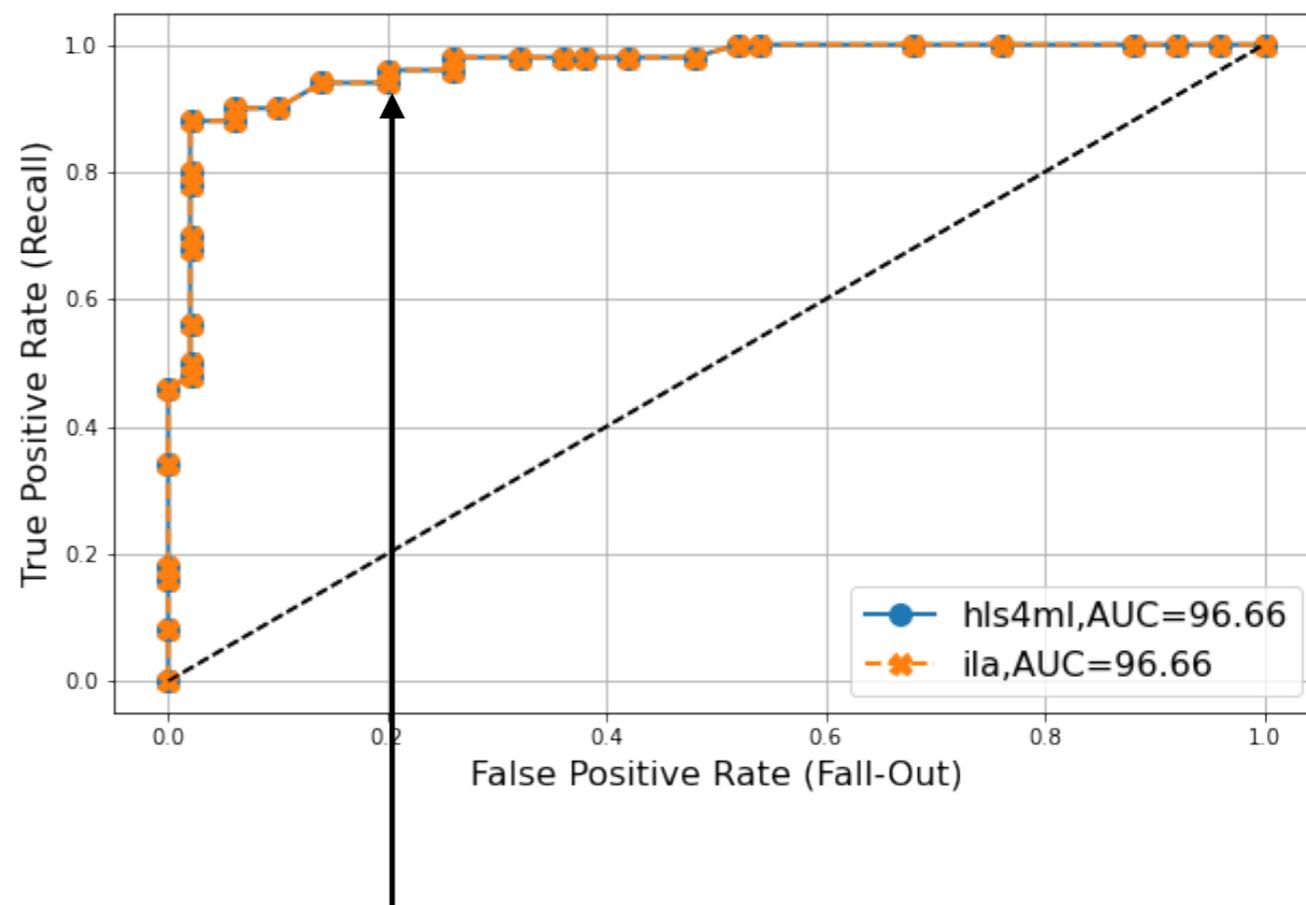
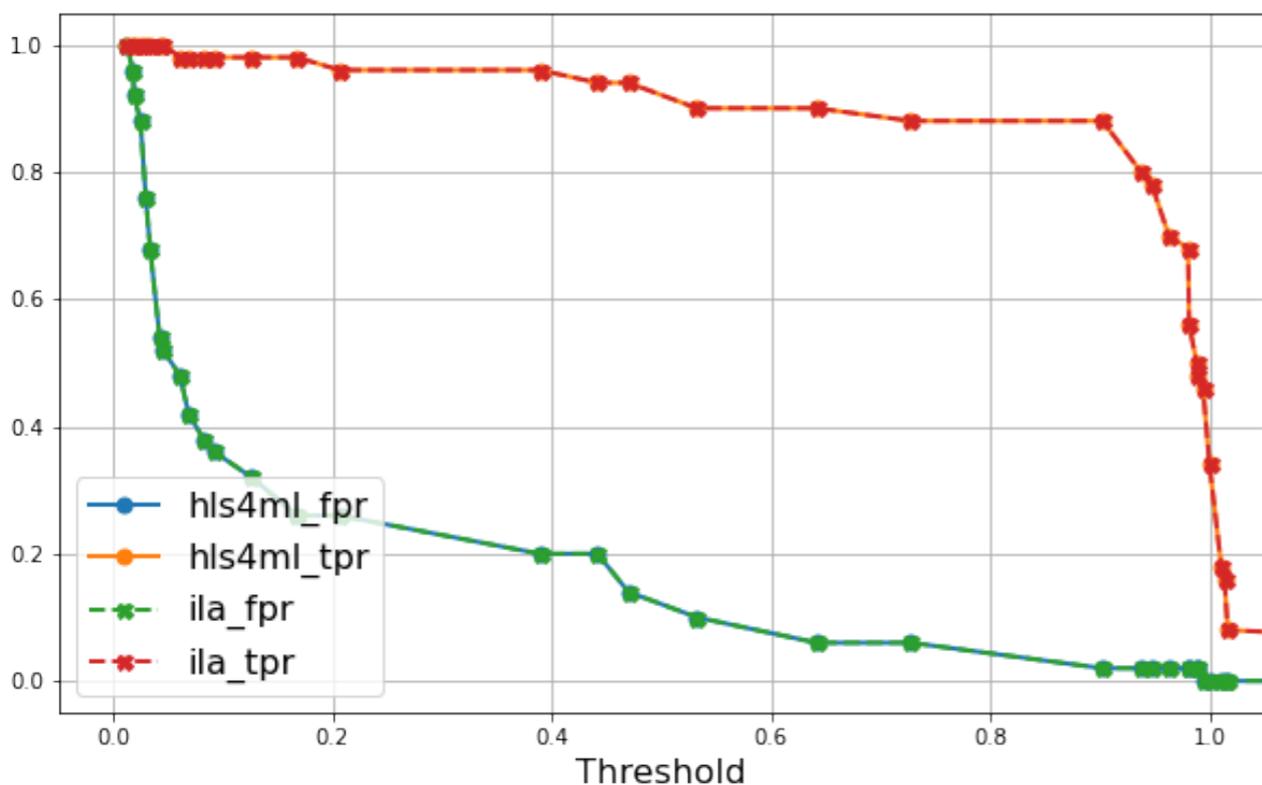
誤って陽性と分類された陰性の割合

TP : 真陽性の数 Signal eventをsignal eventであると分類した数

TN : 新陰性の数 BG eventをBG eventであると分類した数

FP : 偽陽性の数 BG eventをsignal eventであると分類した数

FN : 偽陰性の数 Signal eventをBG eventであると分類した数



分類器がSignal eventであると予測したとき、その予測が正しいのは98%

全てのBG eventの20%を誤ってSignal eventであると分類