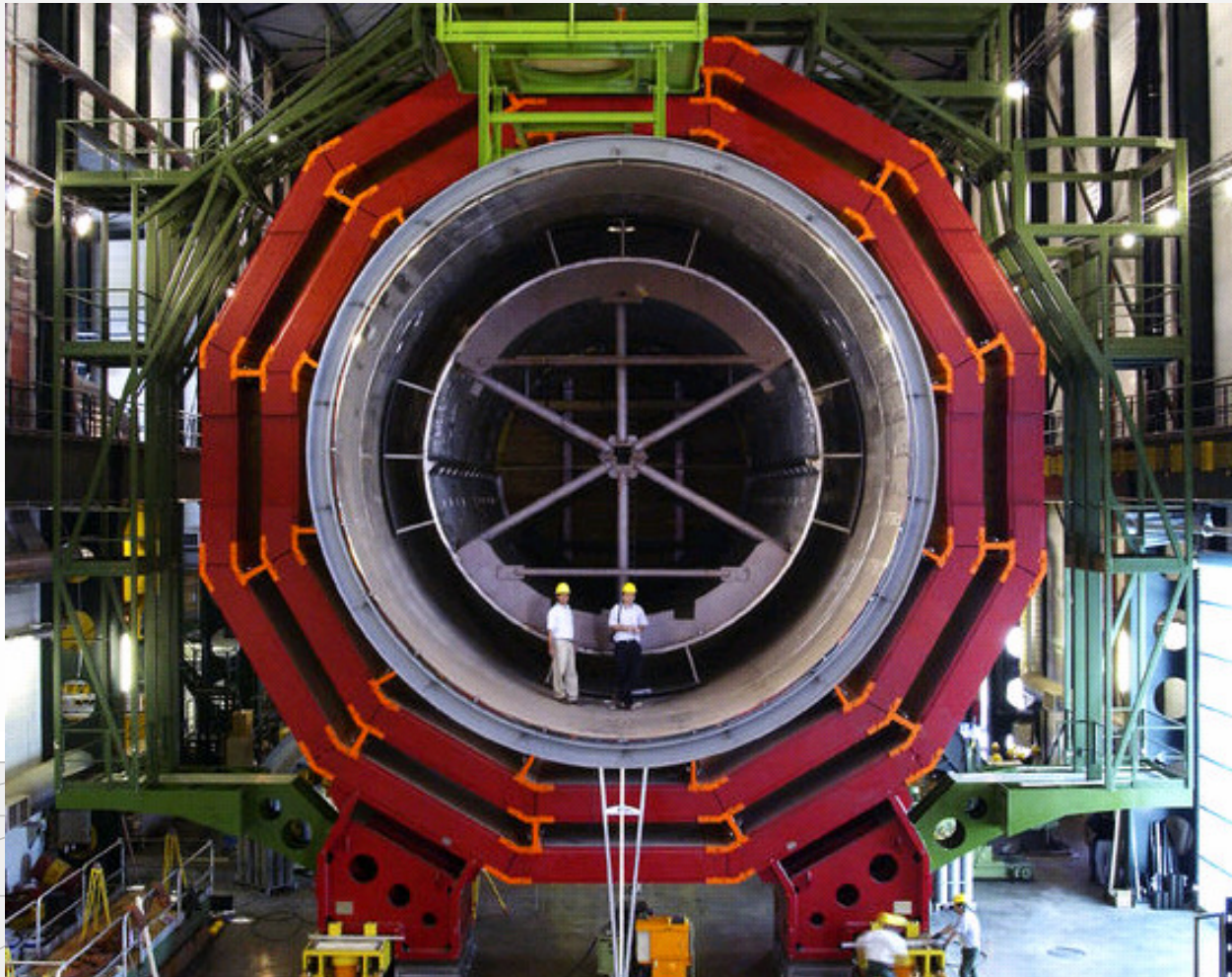


Введение в системы сбора данных

ССД больших экспериментов

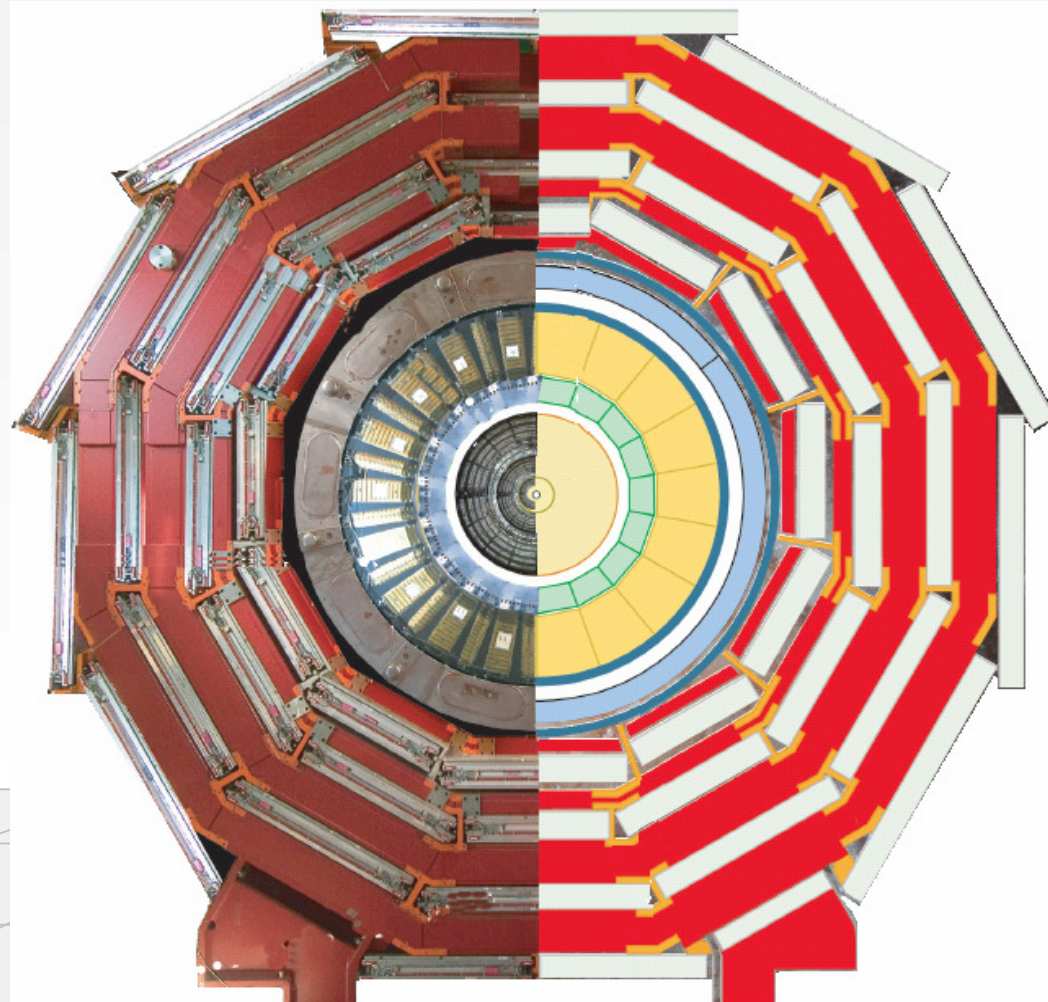
О.Соловьянов на основе
материалов N.Neufeld

Большие эксперименты



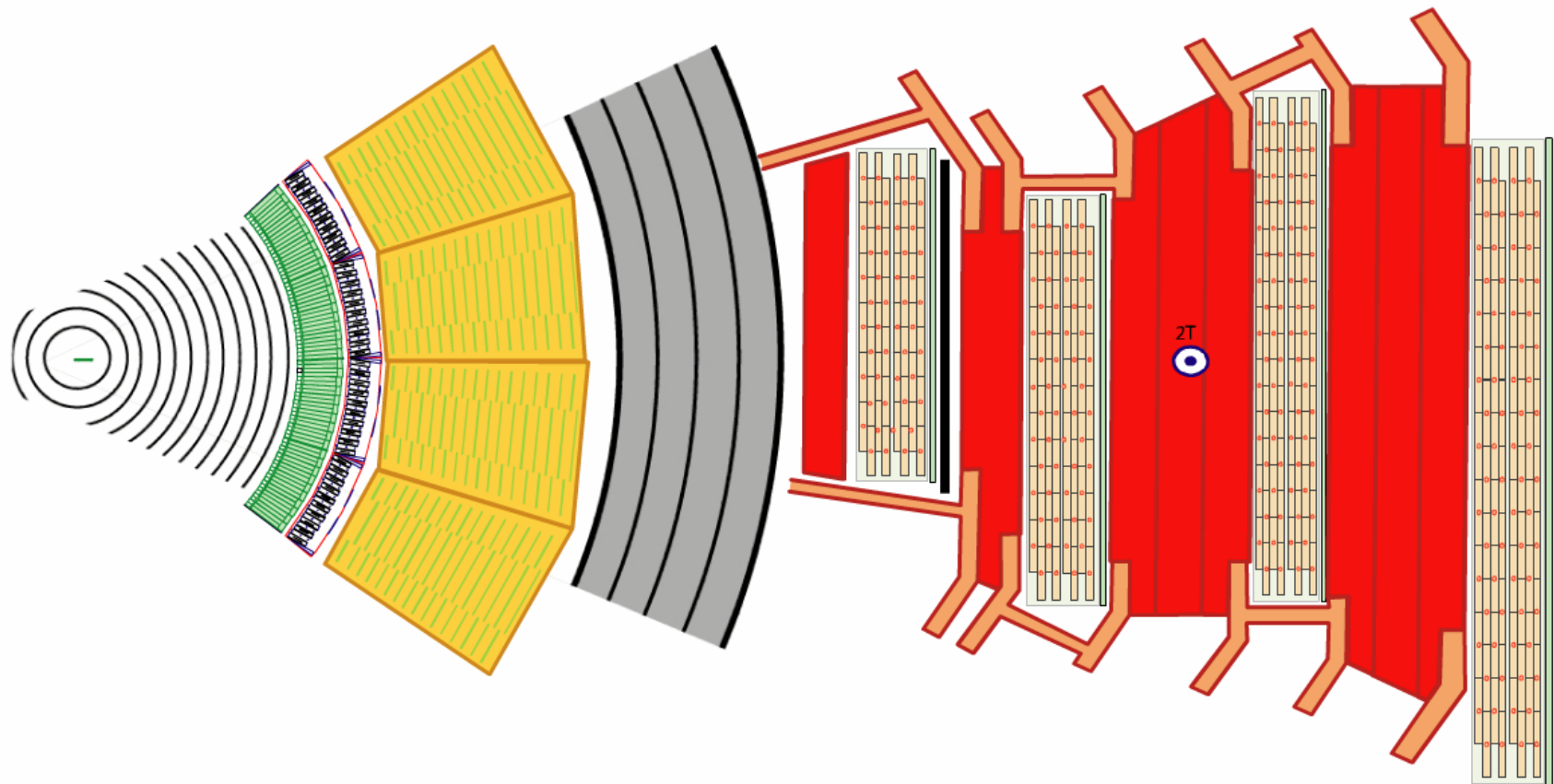
Детектор CMS (Compact Muon Solenoid)

Большие эксперименты



Детектор CMS

Большие эксперименты



Детектор CMS

Большие эксперименты

- 15 миллионов каналов
- 40 MHz
- $= \sim 15 * 1,000,000 * 40 * 1,000,000$ bytes
- $= \sim 600$ TB/sec



Проектирование ССД для большого эксперимента



- Что значит «большого»?
 - Число каналов: для экспериментов на LHC $O(10^7)$ каналов детектора
 - оцифрованный канал может быть от 1 до 16 бит
 - Частота: для экспериментов на LHC все происходит с частотой 40.08 MHz, частотой пересечения пучков
 - это соответствует 24.9500998 ns ну или 25 ns для краткости записи
- Эксперименты в физике высоких энергий обычно состоят из множества разных под-детекторов:
 - трекинг
 - калориметрия
 - идентификация частиц
 - детекторы мюонов

Основные вопросы

- Можем ли мы или хотим спасти все данные?
- Как мы отбираем нужные данные?
- Нужен ли постоянное чтение данных, т.е. это эксперимент на коллайдере?
- Или есть свободные промежутки времени вместе с периодом прихода событий – обычно в экспериментах на фиксированной мишени
- Как сделать так чтобы данные от множества каналов разных под-детекторов соответствовали одному событию (столкновению пучков)?

Что нужно для чтения детектора?

- Механизм отбора событий (“триггер”)
- Чтение (электронное) сенсоров и датчиков детекторов (“регистрирующая электроника”)
- Система синхронизации (“часы”)
- Система для сбора отобранных данных (“ССД”)
- Система конфигурации, управления и наблюдения за детектором и ССД
- Время, деньги, физики, кофе... и конечно студенты 😊

Большие ССД



Система сбора данных

- Данные события оцифрованы, предварительно обработаны и помечены уникальным, увеличивающимся номером
- Но данные события находятся во множестве разных источников (электронных блоках)
- Для следующей ступени отбора, или для того чтобы просто сохранить эти данные (например на диске): необходима ССД

ССД на основе сетей

- В больших экспериментах обычно нужно прочитать большое количество детекторов, которые очень часто далеки друг от друга → *шины этого сделать не могут!*
- Сетевые технологии решают задачу масштабирования
 - В сети устройства равны («партнеры»)
 - В сети устройства общаются напрямую друг с другом
 - Нет необходимости в арбитрации
 - Гарантированная пропускная способность
 - Данные и управление используют те же линии
 - Меньше проводников (например всего два в традиционном Ethernet)
 - На сигнальном уровне сети обычно используют параллельные медные проводники. Сетевые технологии могут быть так же оптическими, беспроводными, и обычно являются последовательными линиями.

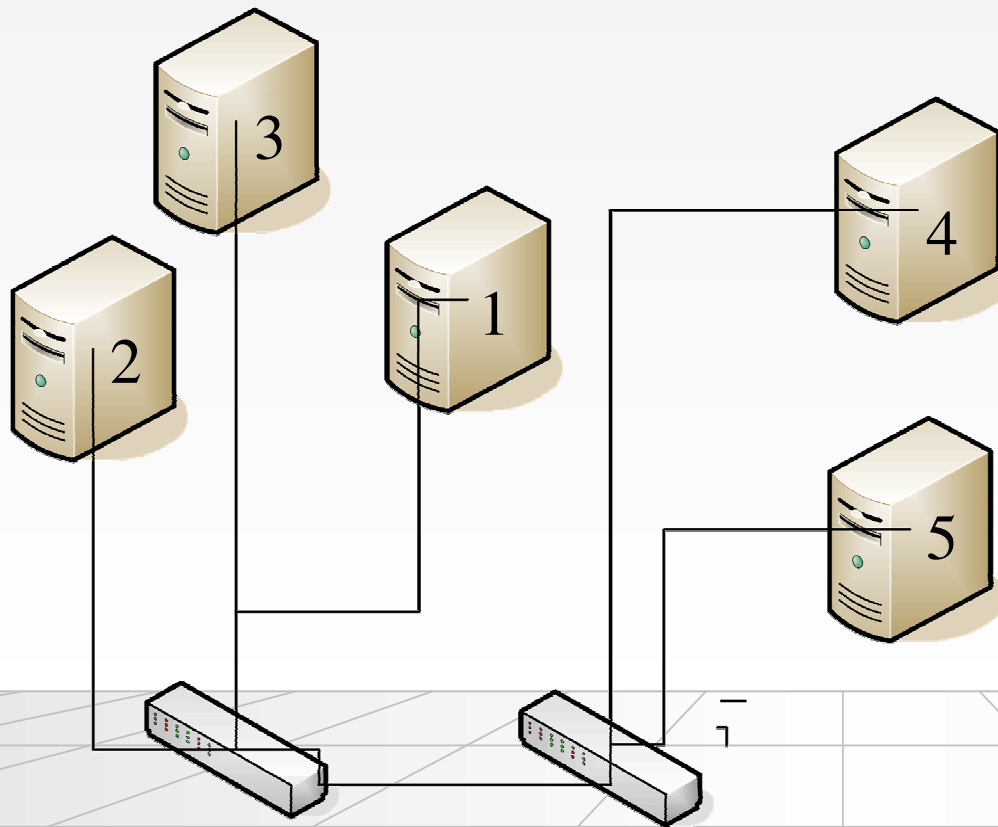
Общение между устройствами

- В сети устройство наделено **сетевым адресом**
- Устройства общаются посредством передачи сообщений (пакетов) друг-другу
- Некоторые устанавливают **соединения** подобно телефонной сети, некоторые просто посылают сообщения
- Современные сети являются **коммутируемыми со связью точка-точка**
 - Коммутация соединений, коммутация пакетов

Коммутируемые сети

- В коммутируемой сети каждый узел подсоединен к одному или нескольким **коммутаторам**
- Коммутаторы могут быть подсоединены к другим коммутаторам
- Путь от одного узла к другому проходит через один или несколько коммутаторов (их число иногда называют числом участков ("**hops**"))

Коммутируемая сеть



- Пока 2 посылает данные 1, то 4 и 3 могут посылать данные 5 на полной скорости
- 2 может распределить пропускную способность между собой и 1 и 4 по необходимости

Коммутаторы

- Производительность коммутаторов является ключевой в производительности всей сети
- Коммутаторы должны надежно и как можно быстро передавать пакеты данных между узлами сети
- Перед ними стоят две проблемы
 - Найти правильный путь для пакета
 - Скопление пакетов (два или более пакета пришли одному узлу в одно и тоже время)

Ethernet

- Дешевый
- Ненадежный – но на практике частота ошибок передачи низка
- Доступен с разными скоростями в разных передающих средах
- Обычно используется протокол TCP/IP
- Наиболее широко используемая технология для построения локальных сетей

IP пакеты в Ethernet

Ethernet Header



IP Header

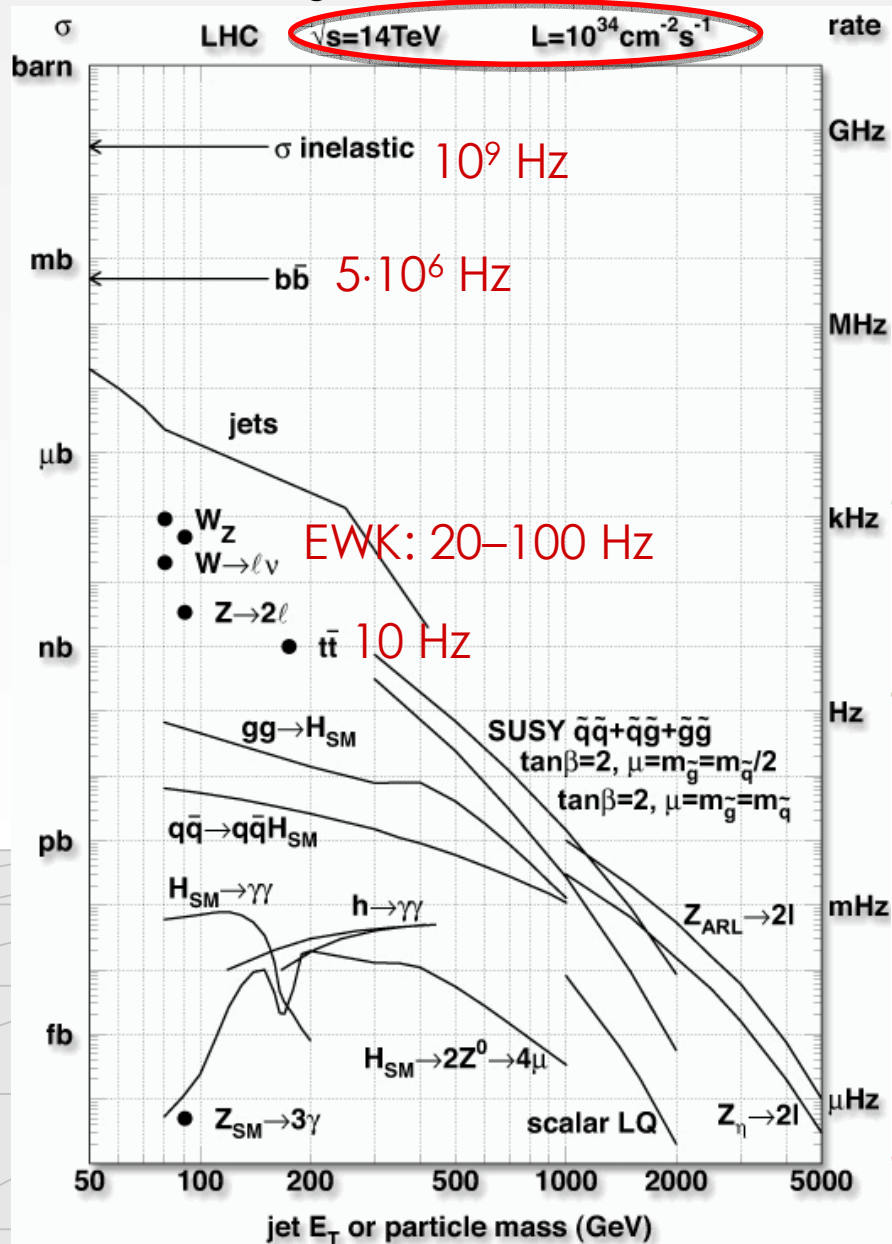
UDP Header

Data

0 ... 32 bits

Триггер

Нужно ли читать все подряд?

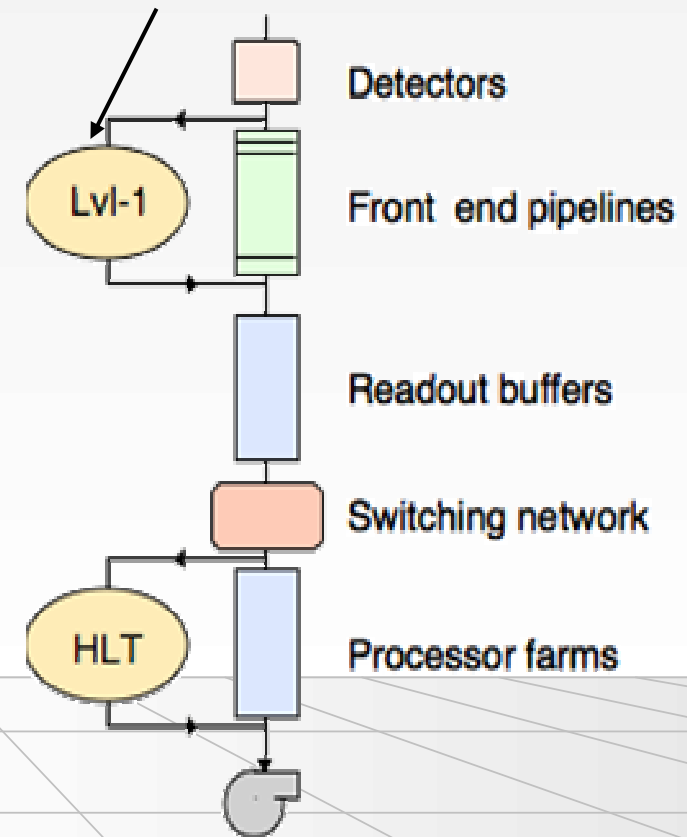


- Типичное столкновение “скучное”
 - Но нам необходимо некоторое количество таких “скучных” данных для проверки, калибровки, а также для некоторой важной “низкоэнергетичной” физики
- “Интересная” физика случается примерно на 6–8 порядков реже
- “Новая” физика, включая новые частицы/открытия, примерно на 9 порядков реже
 - 100 GeV Higgs - 0.1 Hz
 - 600 GeV Higgs - 0.01 Hz
- Нам *всего лишь* ☺ надо эффективно отделить эти события от фона прежде чем прочитать и записать

Триггер на LHC

- Никакая (доступная) ССД не в состоянии прочитать $O(10^7)$ каналов с частотой 40 MHz \rightarrow 400 TBit/s – даже если подразумевать двоичные данные (да/нет)!
- Хуже того: большинство из этих миллионов событий в секунду абсолютно неинтересны: одно событие с рождением бозона Хиггса предполагается каждые 0.02 секунды
- Триггер первого уровня (*Level-1, L1*) должен как-то отобрать наиболее интересные события и сказать нам с какими из них дальше работать

Черная магия ☺



Внутри черного ящика: как работает триггер первого уровня?

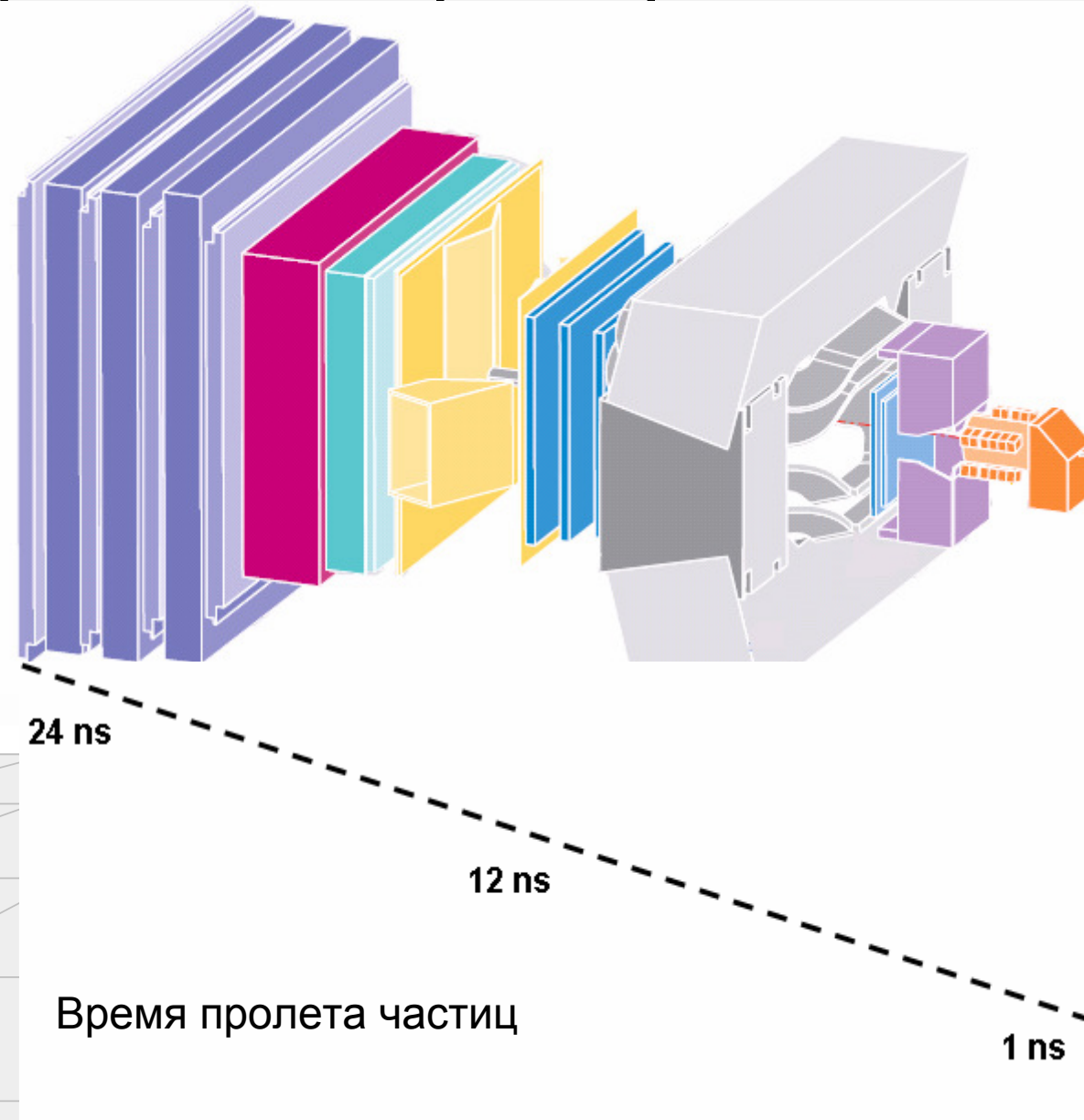


- Миллионы каналов →: надо работать с как можно более “локальной” информацией
 - Малое число межсоединений
- Быстрый: поиск “простых” признаков
 - Сохранить хорошие, выбросить плохие
 - Надежный, может быть сделан в «железе» (быстрый)
- Принцип построения:
 - Быстрый: чтобы не выйти за размер буферов
 - Новое событие каждые 25 наносекунд: должен принять решение за несколько микросекунд (μs): задержка триггера

Проблемы триггера L1 на LHC

- N (каналов) $\sim O(10^7)$; ≈ 20 взаимодействий каждые 25 ns
 - Требуется большое число соединений
- Надо синхронизовать элементы детекторов лучше чем 25 ns
- В некоторых случаях: длина сигнала или время пролета > 25 ns
 - Собрать информацию о нескольких пересечениях
 - Необходимо определить номер пересечения...
- Прямо сейчас (нельзя вернуться назад и перечитать давние события)
 - Необходимо наблюдать за эффективностью и качеством отбора – нужен хороший контроль за текущими установками и характеристиками

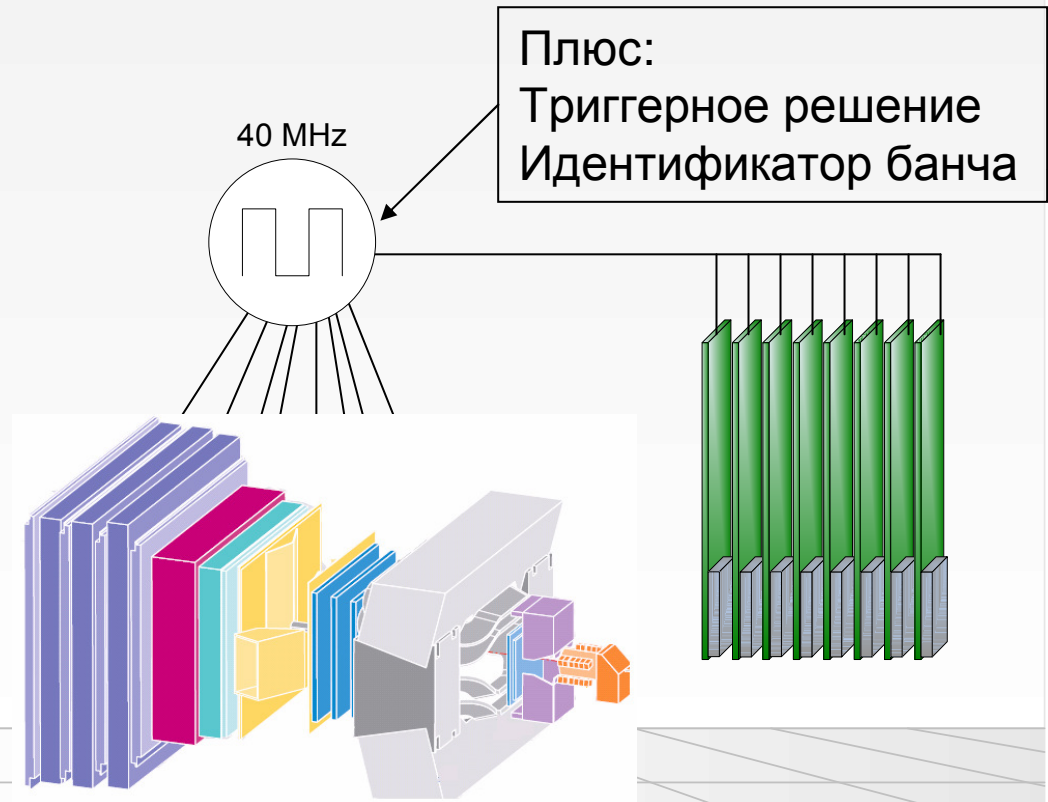
Проблемы триггера L1 на LHC



Распределение тактовой частоты и синхронизация

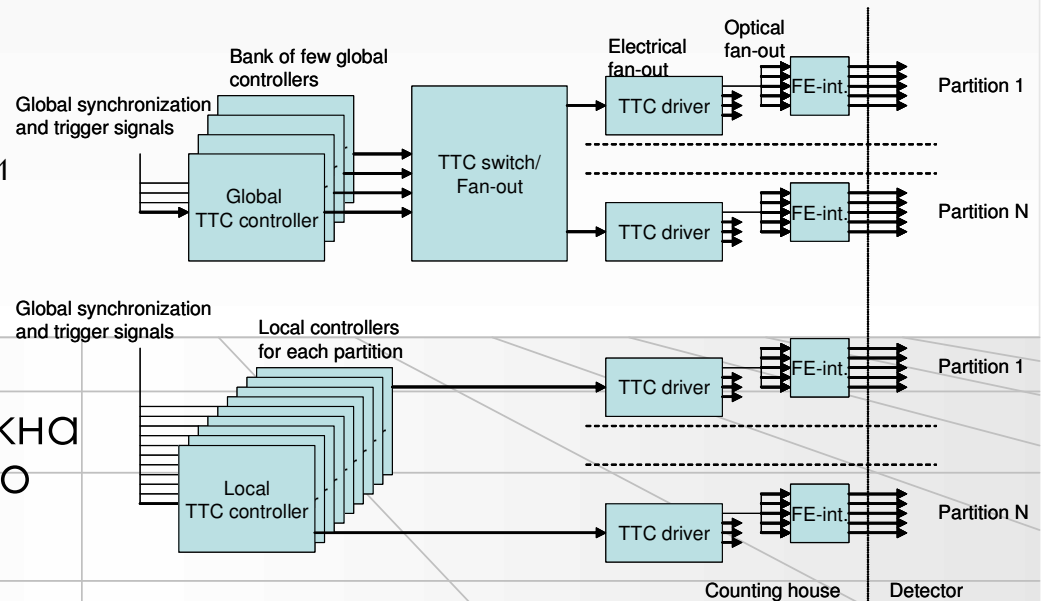
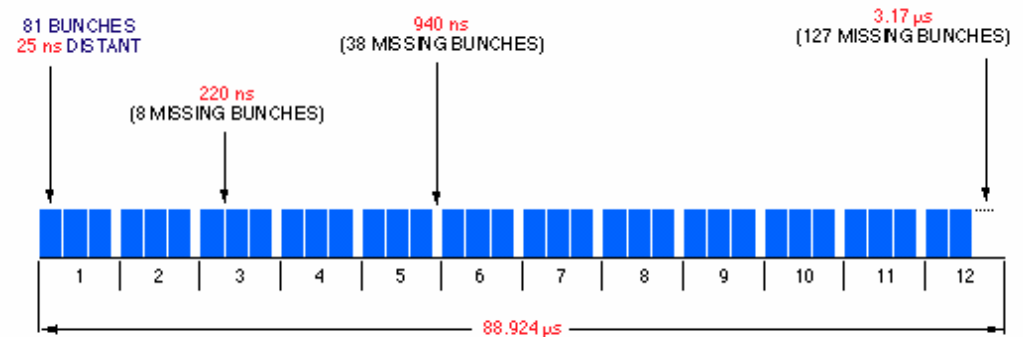


- **Событие** это «моментальная фотография» значений всех каналов детектора, в которых есть значимый отклик
- Общий тактовый сигнал должен быть передан всем элементам детектора
 - Поскольку скорость света (сигнала) постоянна, детекторы велики а электроника быстра, **элементы детектора** должны быть точно **установлены по времени**
- Общей системой для LHC экспериментов является **TTC (Trigger Timing and Control)** основанная на радиационно-стойкой оптоэлектронике



Временные привязки и контроль

- Тактовая частота с малым дрожанием (jitter)
- Синхронный сброс
- Синхронизация со структурой пучка (банчей) ускорителя
- Калибровка
- Триггер (с типом события)
- Временная привязка всех частей детектора
 - Программируемые задержки
- Разветвление – в одном направлении
 - На весь эксперимент
 - На под-детектор
- Должна быть высоко надежна и точна, потом очень трудно скорректировать нежелательные эффекты



Знай своего врага: pp соударения при 14 TeV и $10^{34} \text{ cm}^{-2}\text{s}^{-1}$

- $\sigma(\text{pp}) = 70 \text{ mb}$ -
→ $> 7 \times 10^8 / \text{s}$ (!)
- В ATLAS и CMS
~20 событий
наложатся
друг на друга
(pileup)

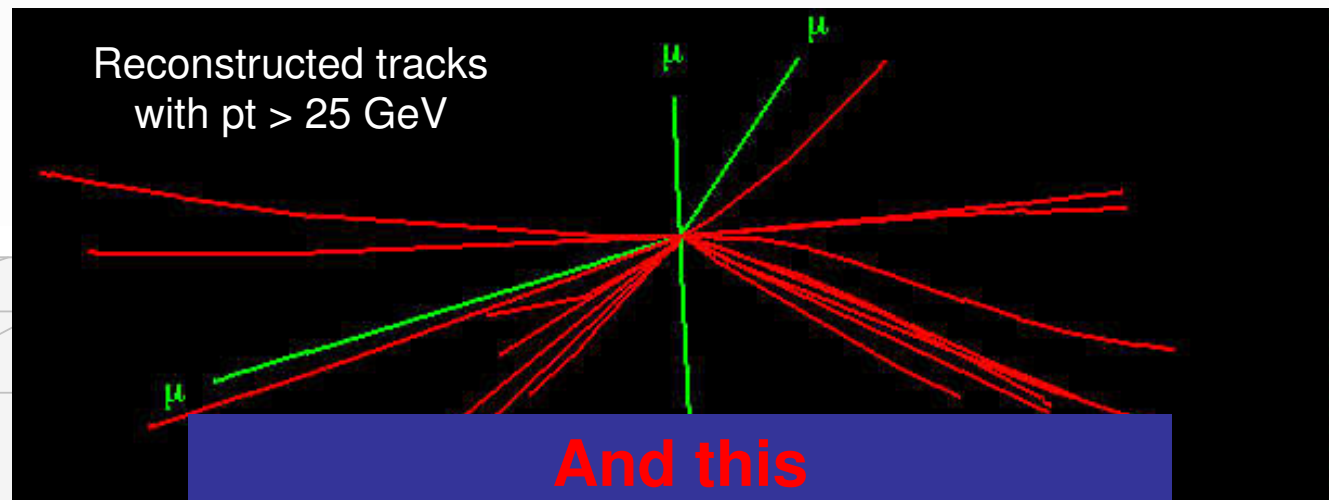
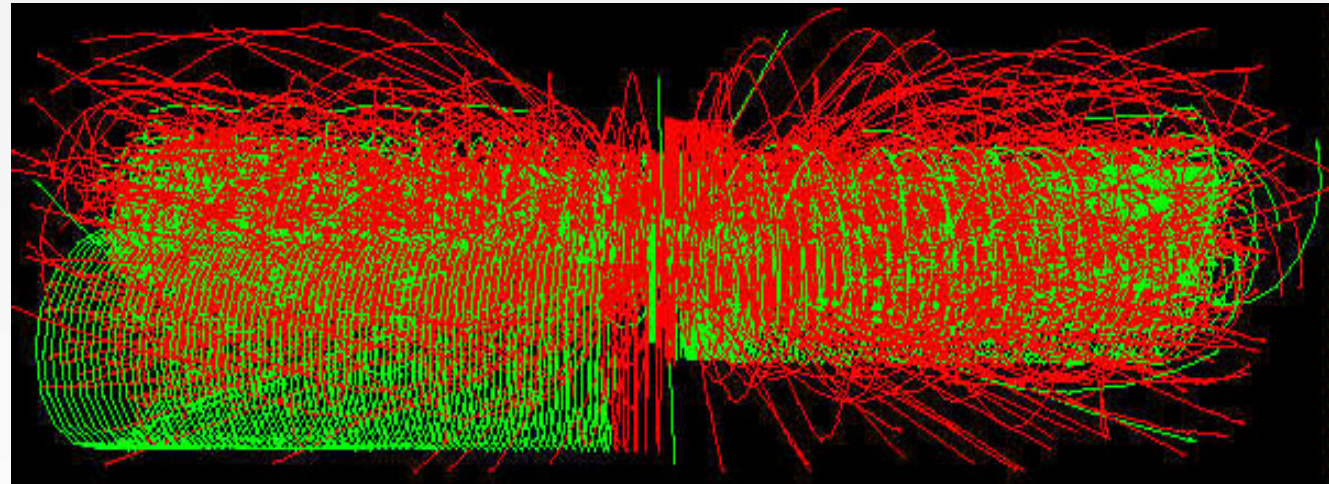
• $\text{H} \rightarrow \text{ZZ}$

$\text{Z} \rightarrow \mu\mu$

$\text{H} \rightarrow 4 \text{ muons}$:

Наиболее чистая
("золотая")

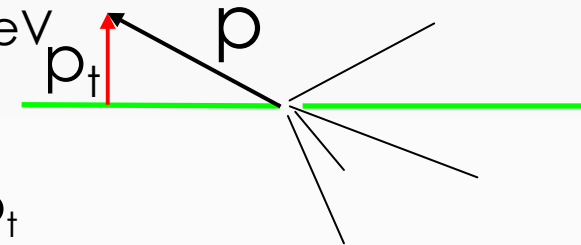
сигнатура



**And this
(not the H though...)
repeats every 25 ns...**

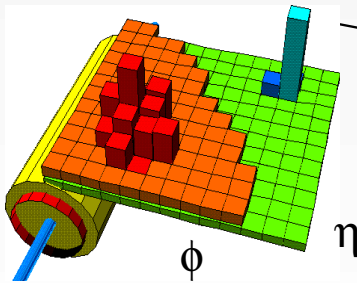
Триггер на поперечный импульс

- pp соударения производят в основном адроны с поперечным импульсом " p_{\perp} " ~ 1 GeV
- Интересная физика (старая и новая) состоит из частиц (лептонов и адронов) с большими p_{\perp} :
 - $W \rightarrow e\nu$: $M(W) = 80 \text{ GeV}/c^2$; $p_{\perp}(e) \sim 30\text{-}40 \text{ GeV}$
 - $H(120 \text{ GeV}) \rightarrow \gamma\gamma$: $p_{\perp}(\gamma) \sim 50\text{-}60 \text{ GeV}$
 - $B \rightarrow \mu D^{*+} \nu$ $p_{\perp}(\mu) \sim 1.4 \text{ GeV}$
- Стоит отбирать частицы с большими p_{\perp}
 - Требуется определения типа частиц; возможно для электронов, мюонов и "струй"; за пределами этого требует сложных алгоритмов
- Вывод: в триггере первого уровня (L1) нужно уметь выделять электроны, мюоны или струи с большим поперечным импульсом



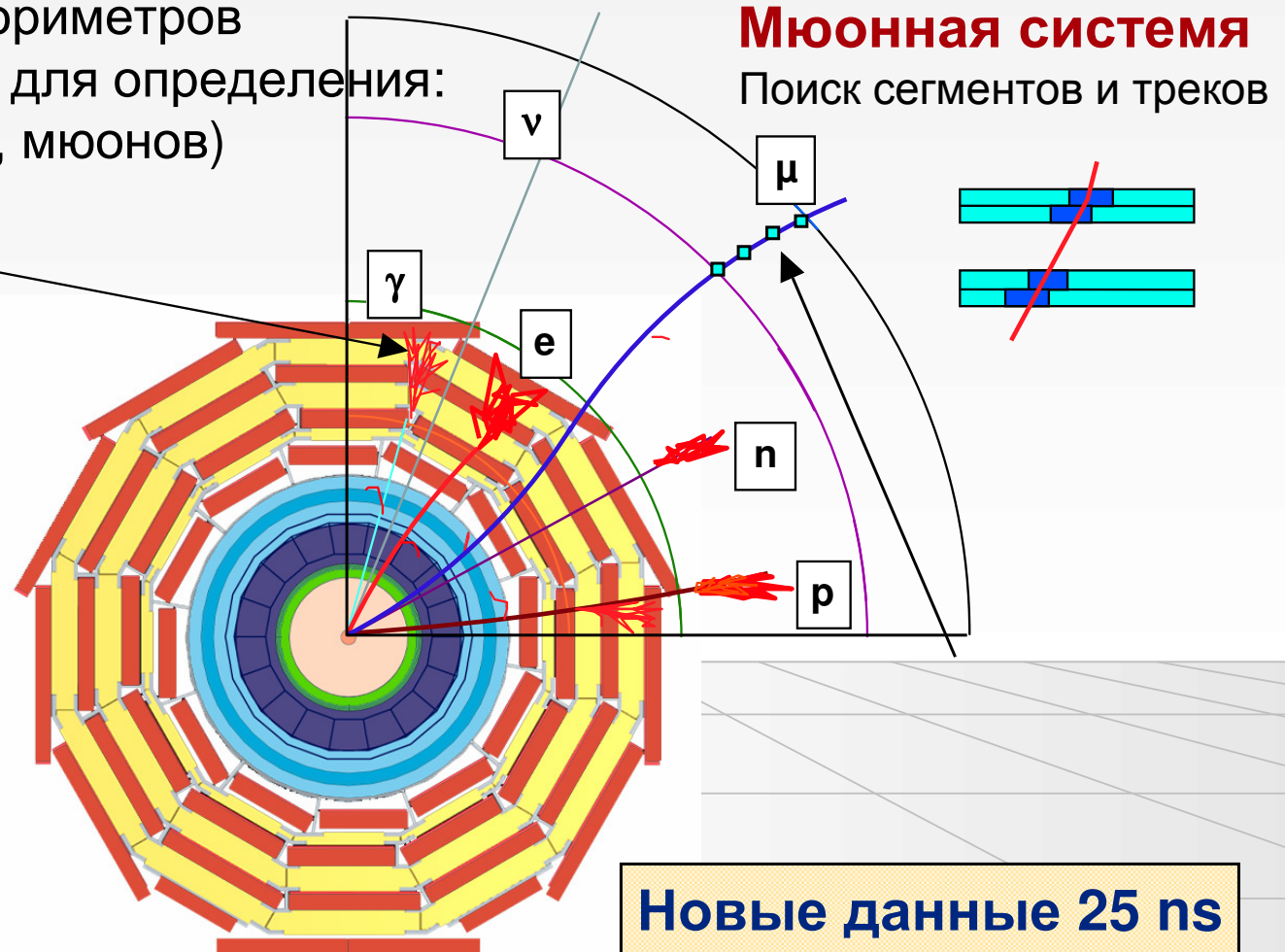
Как избавиться от фона: поперечный импульс p_t

Использование калориметров
и мюонной системы для определения:
Частиц (электронов, мюонов)
с большим p_t



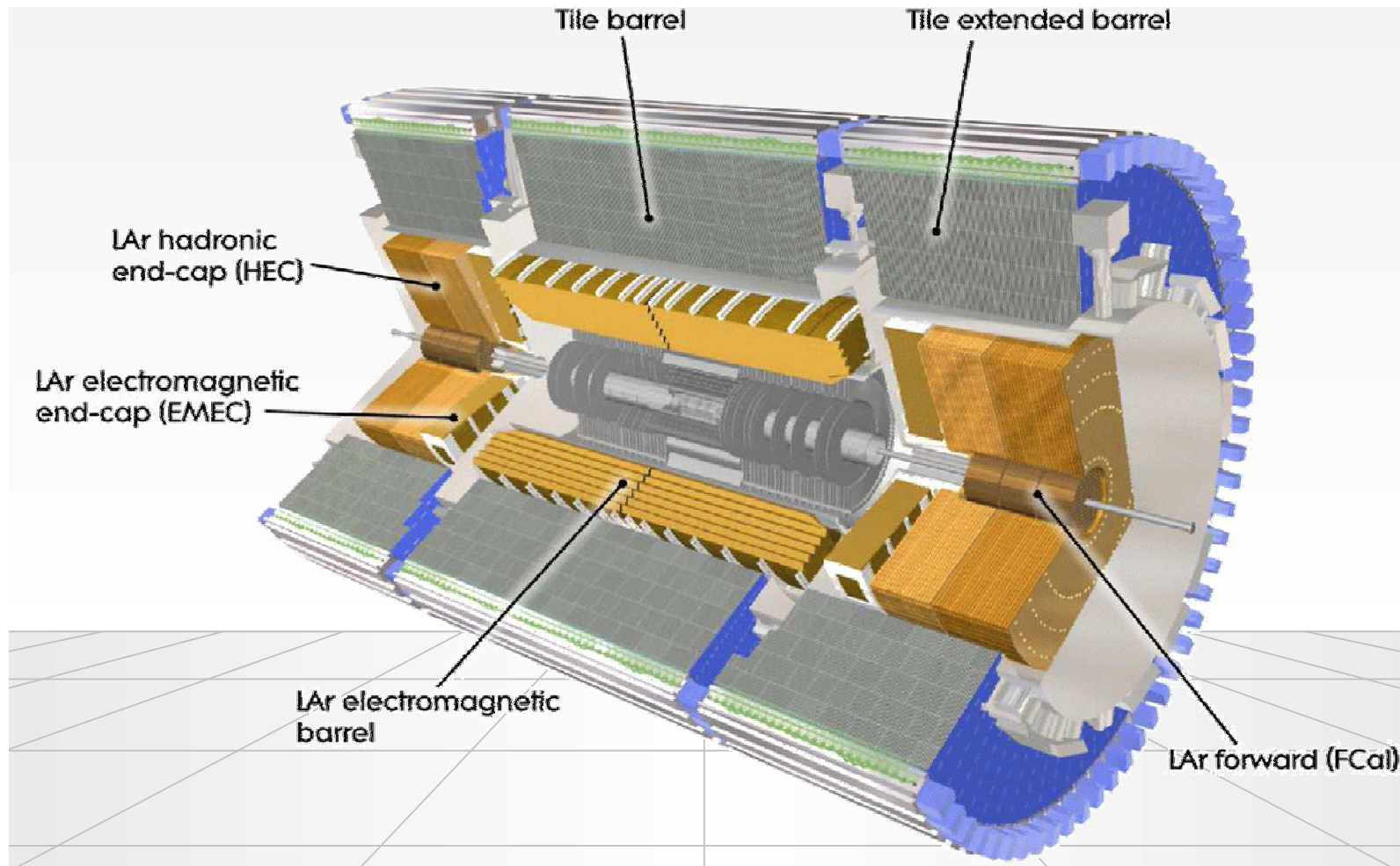
Калориметры

Поиск кластеров
Определение энергии



Новые данные 25 ns
Время решения $\sim \mu\text{s}$

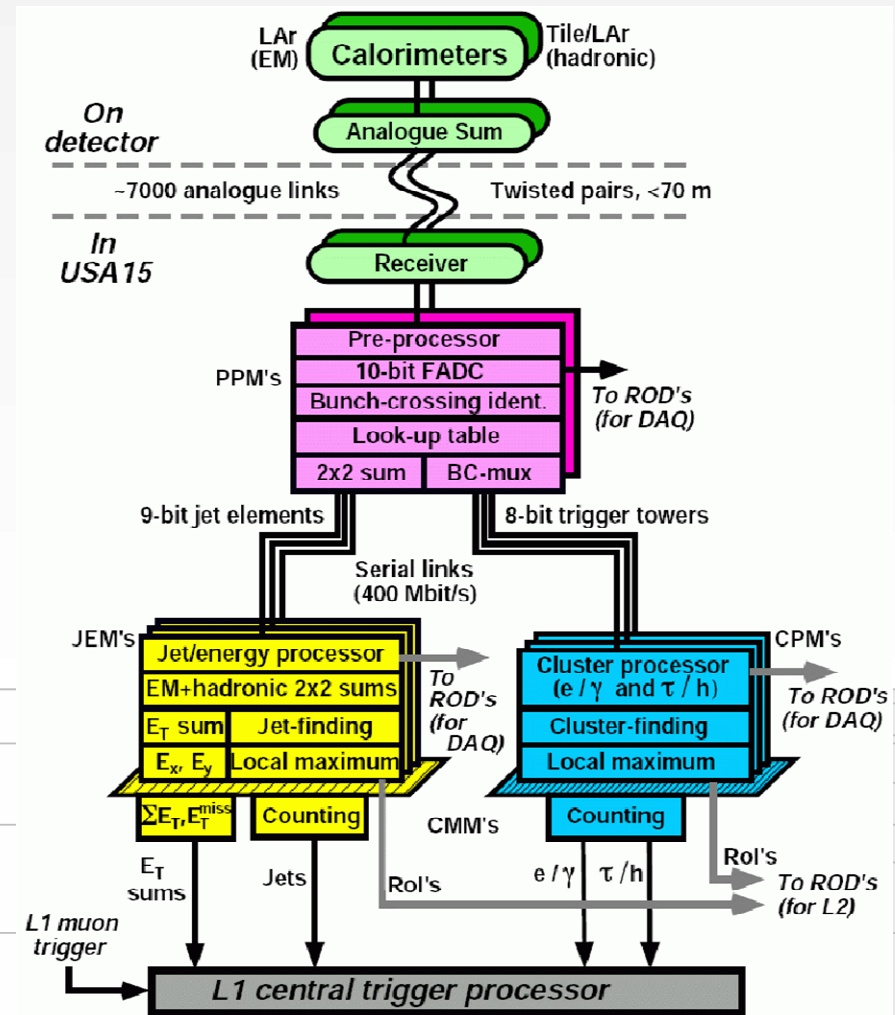
Калориметры ATLAS



Калориметрический триггер ATLAS L1

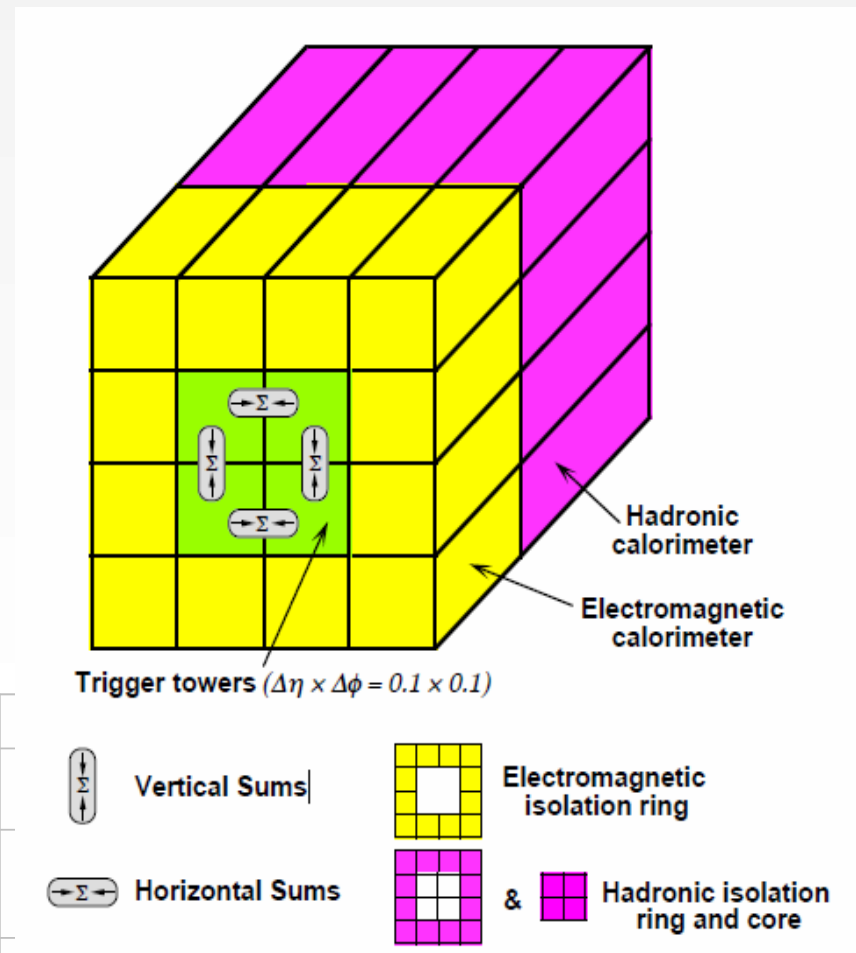


- Аналоговые «башни» 0.1×0.1 ($\delta\eta \times \delta\phi$)
- Оцифровка, идентификация
- таблица (LUT) $\rightarrow E_T$
- Копирование данных струи/суммарная
- Струйный (JEP) и Кластерный (CP) процессоры
- Выдача решения через $1.5 \mu s$ после столкновения
- Передача триггерной информации на следующий уровень



Алгоритм ATLAS L1-Calor

- Суммирование энергий ячеек в квадрате 4x4
 1. 4 э.м. кластера
 2. адронное ядро
 3. 4 адронных кластера (сумма п. 1 и 2)
 4. э.м. и адр. изоляционные кольца
 5. Пороги на энергию кластеров и на отсутствие энергии в изоляционном кольце



Архитектура ATLAS L1-Calorimeter

Буферизация сигналов

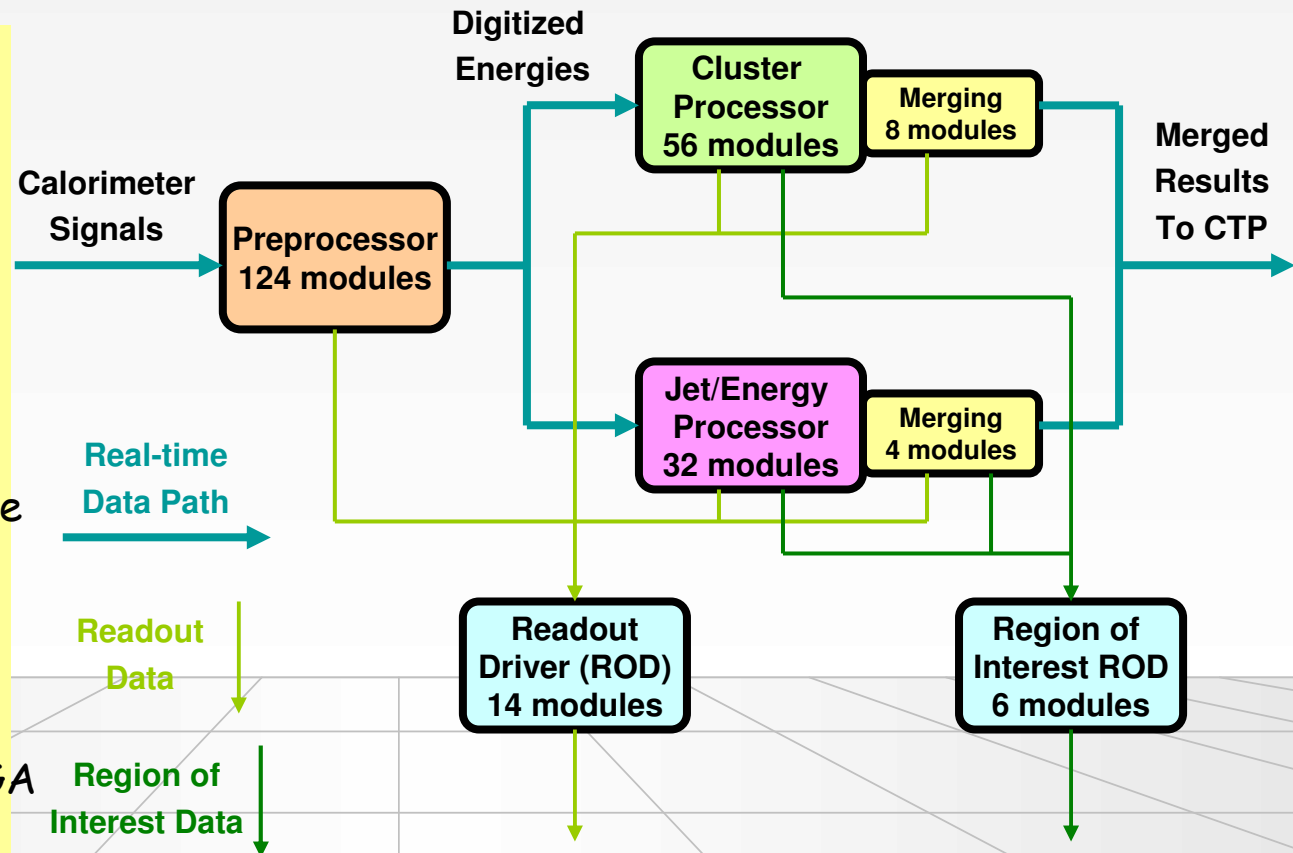
Задержка $< 1 \mu s$

Много стадий обработки

Параллельное исполнение

Многофункциональные модули

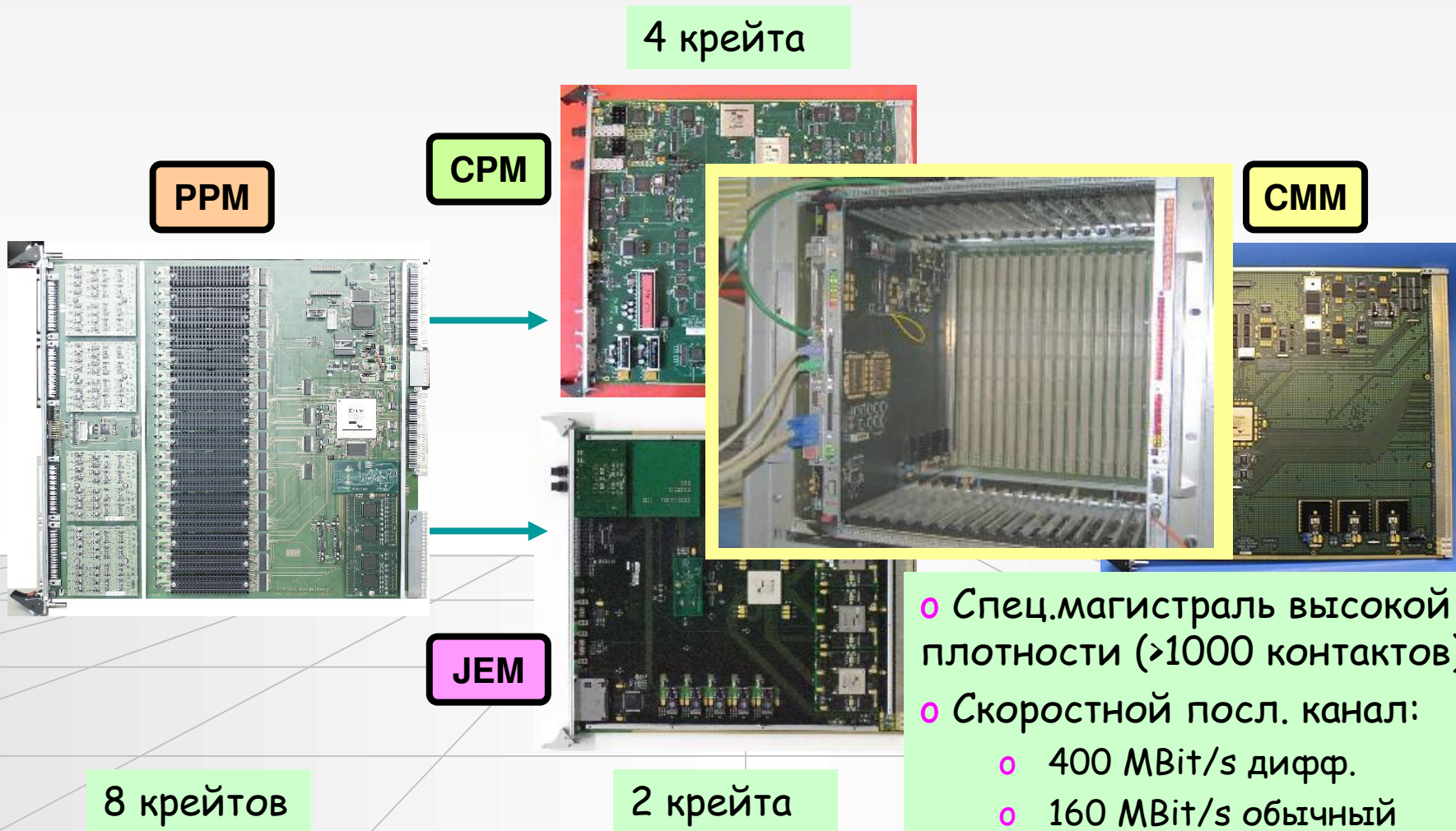
Широкое применение FPGA



5 основных типов
9U модулей

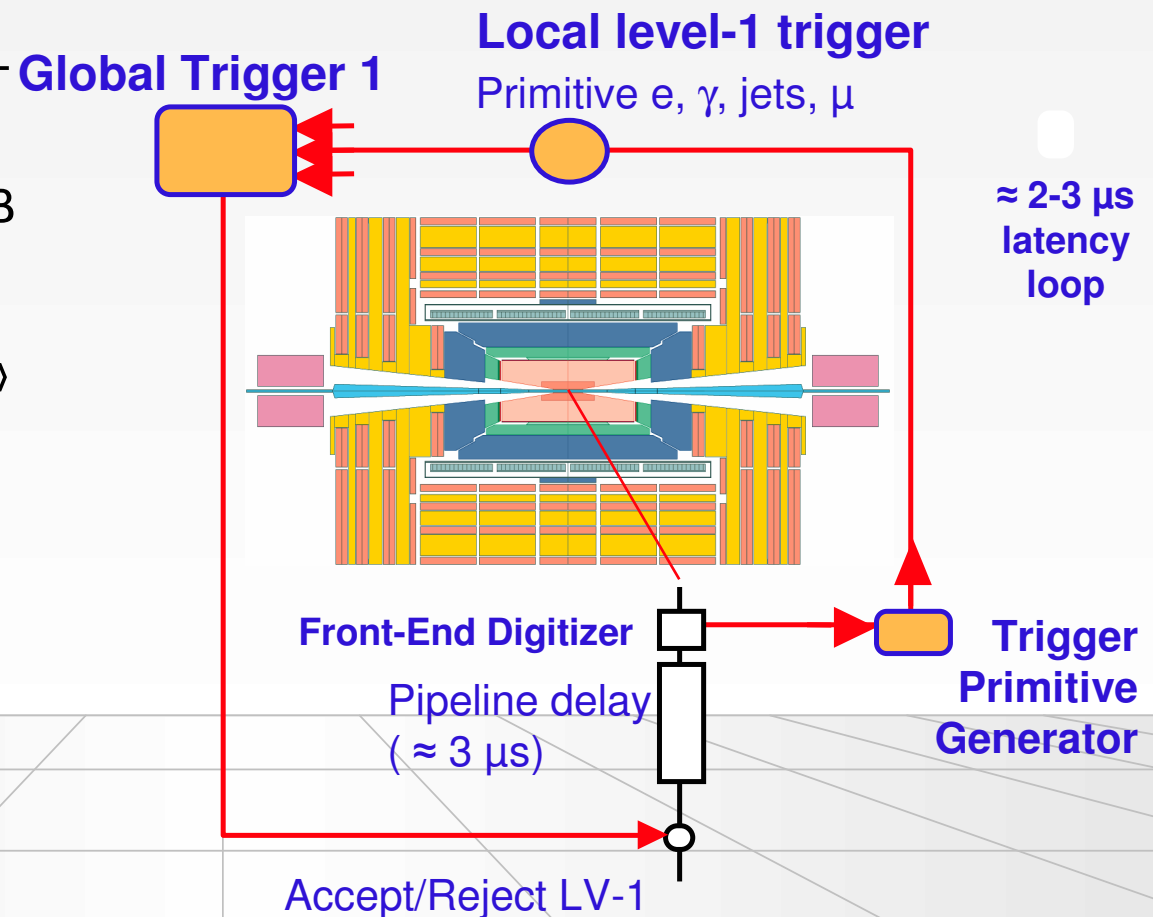


Модули ATLAS L1-Calor

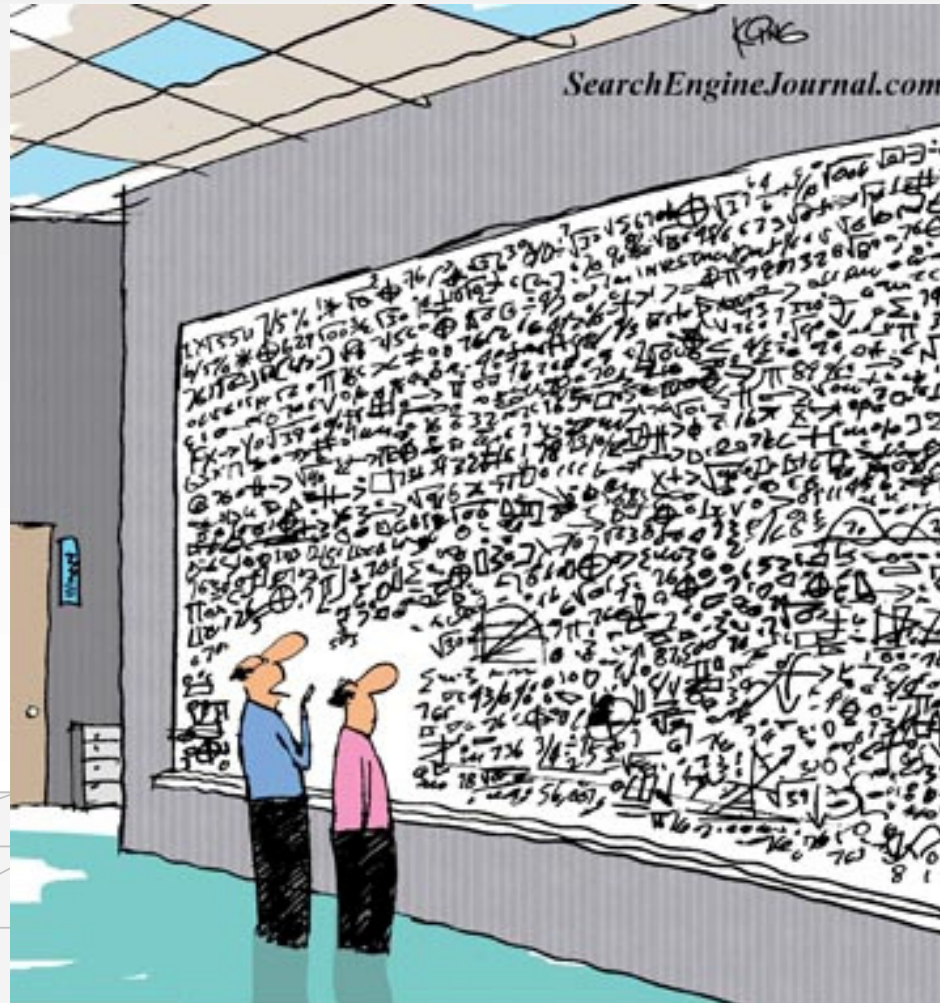


Распространение триггера L1

- Допустим что наш черный ящик выдает да/нет на каждое пересечение пучков
 - Триггер не для философов, тут нет понятия «может быть»
- Это решение должно быть доведено до всей регистрирующей электроники чтобы она послала, либо выкинула данные, полученные в этом столкновении

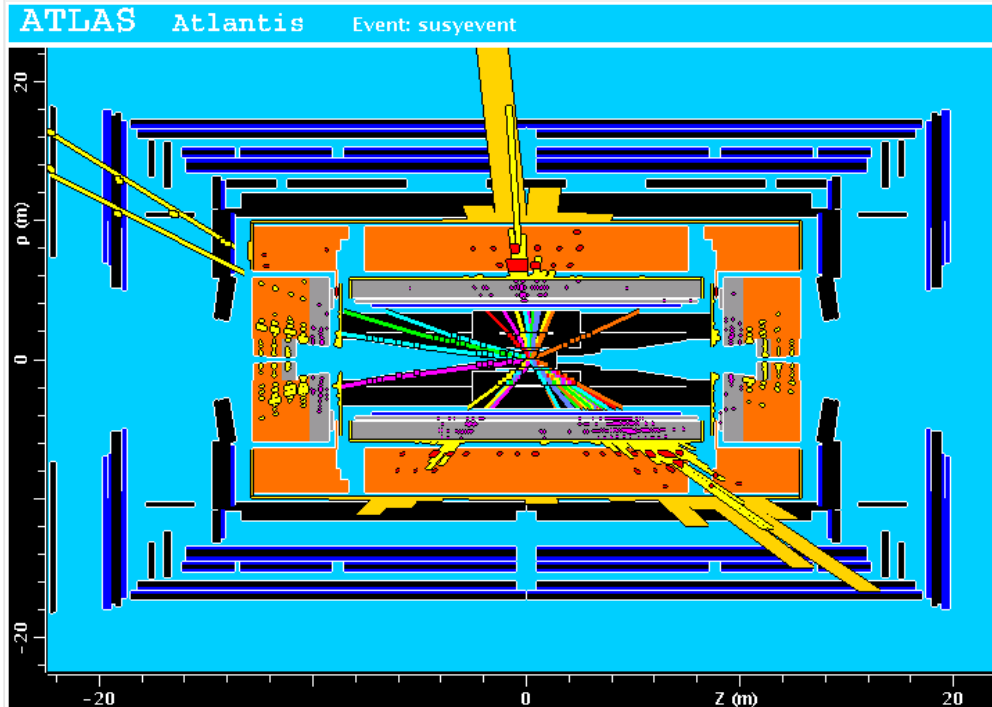


Триггер высокого уровня (HLT)

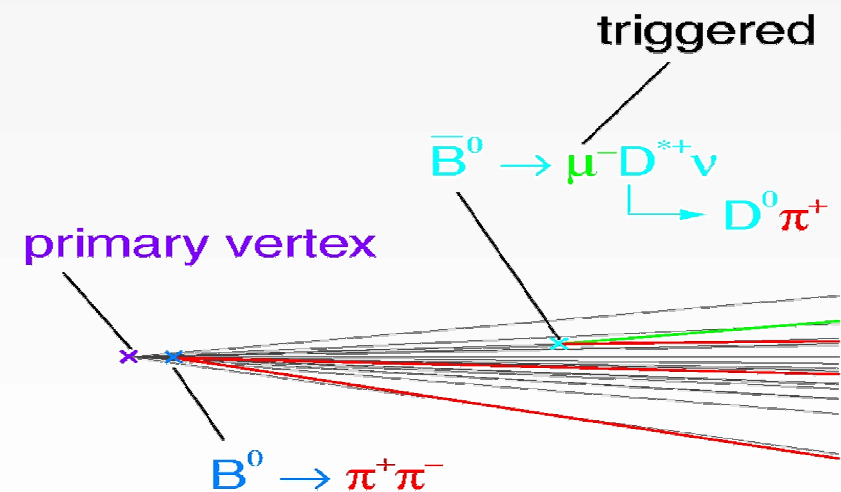


А это, в общих чертах, то что мы делаем в триггере высокого уровня

Триггер высокого уровня

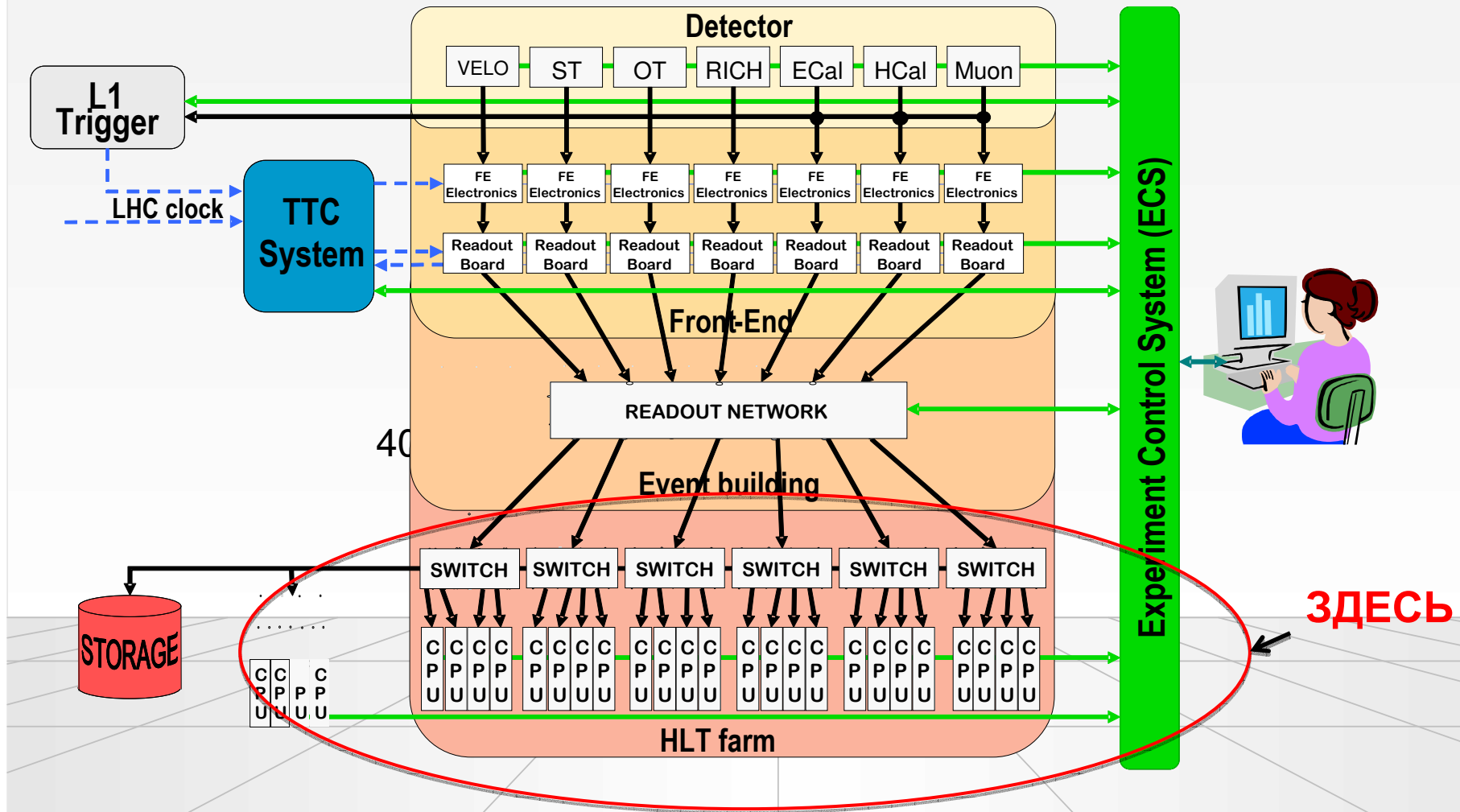


Сложная структура события требует информации от всего детектора



Методы и алгоритмы такие же как в системе реконструкции данных (offline)

Триггер высокого уровня

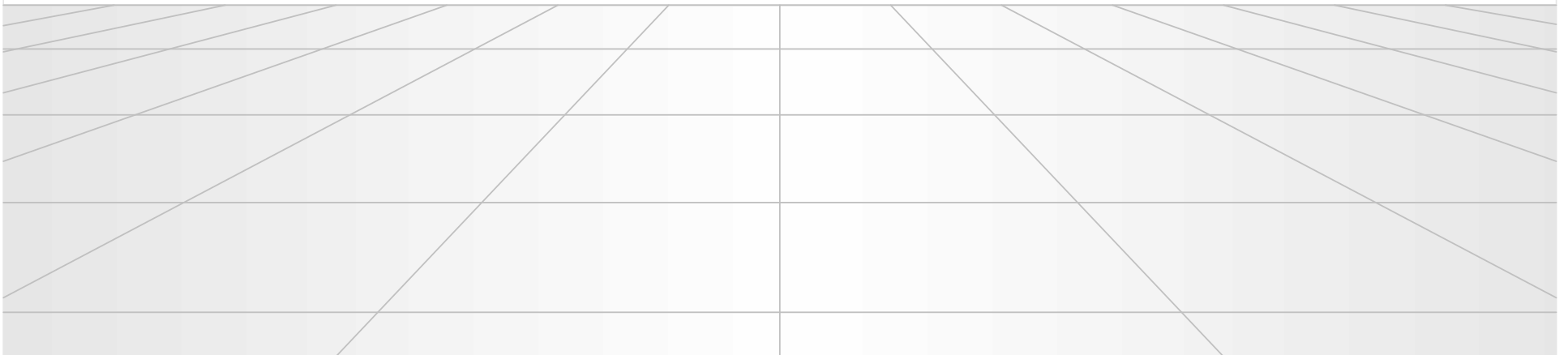


После триггера L1: что дальше?

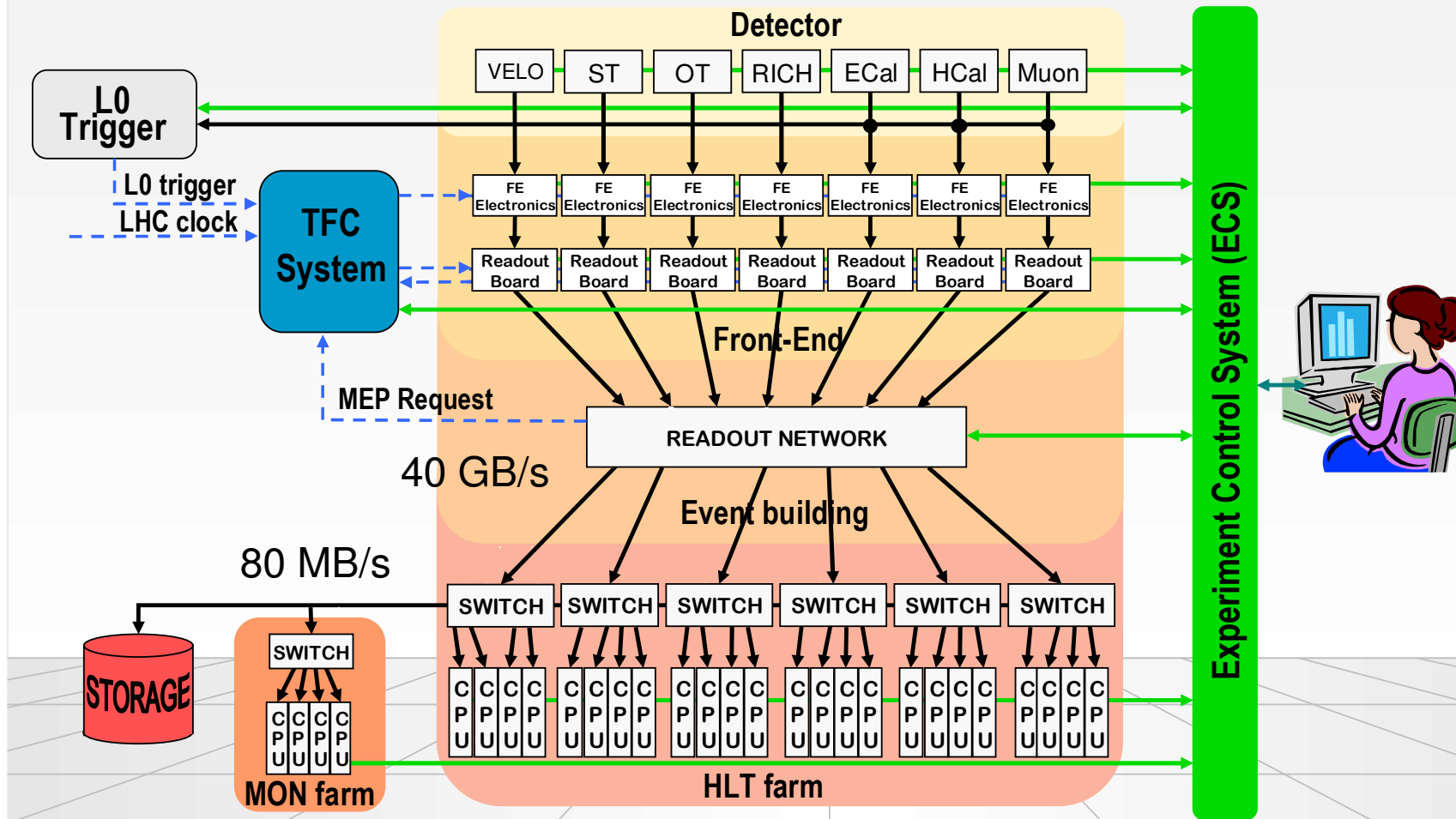
- Условия L1
 - ATLAS and CMS : частота ~ 75 to 100 kHz, размер события ~ 1 -2 MB
 - LHCb: частота 1 MHz, размер события 40 kB / ALICE: $O(\text{kHz})$ и $O(\text{GB})$
- В идеале
 - Запустить полнофункциональный физический анализ
 - Это занимает $O(s)$. Понадобится очень большая и дорогая вычислительная система **100 MCHF?**
- В реальности:
 - Смотрим **только на часть детектора, по указке триггера 1^{го} уровня**
 - LHCb: < 10 ms/event
 - ATLAS: < 40 ms/event
- **→ Уменьшение частоты примерно в 30 раз, а теперь можно запускать полномасштабный анализ данных**

Построение событий

(сбор данных для HLT)



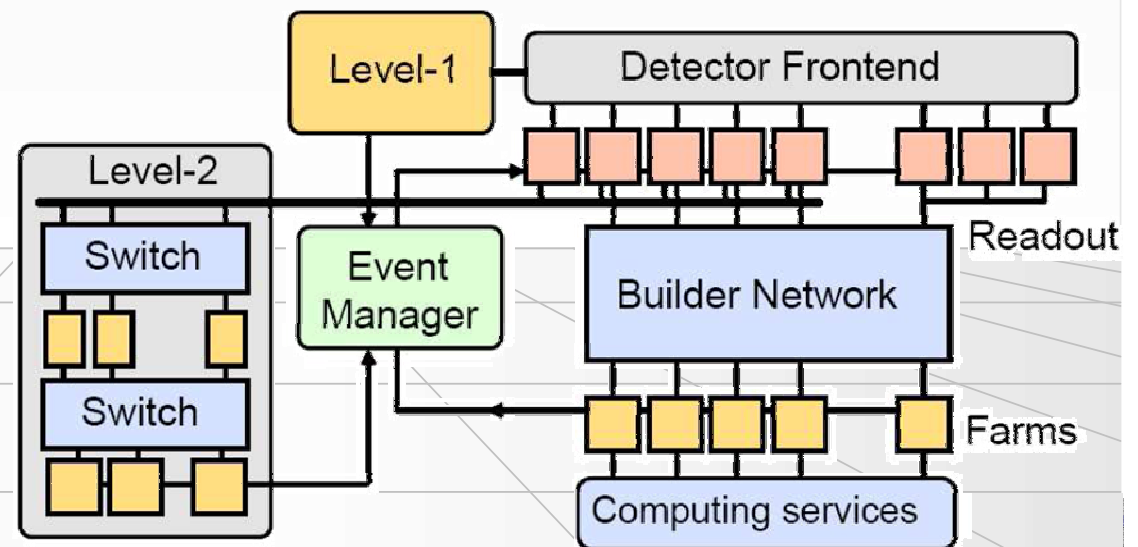
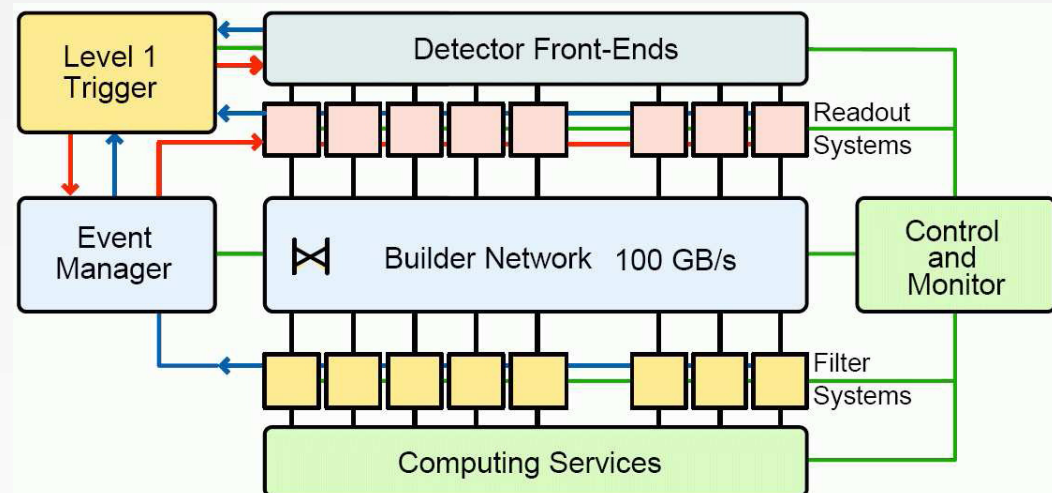
LHCb DAQ



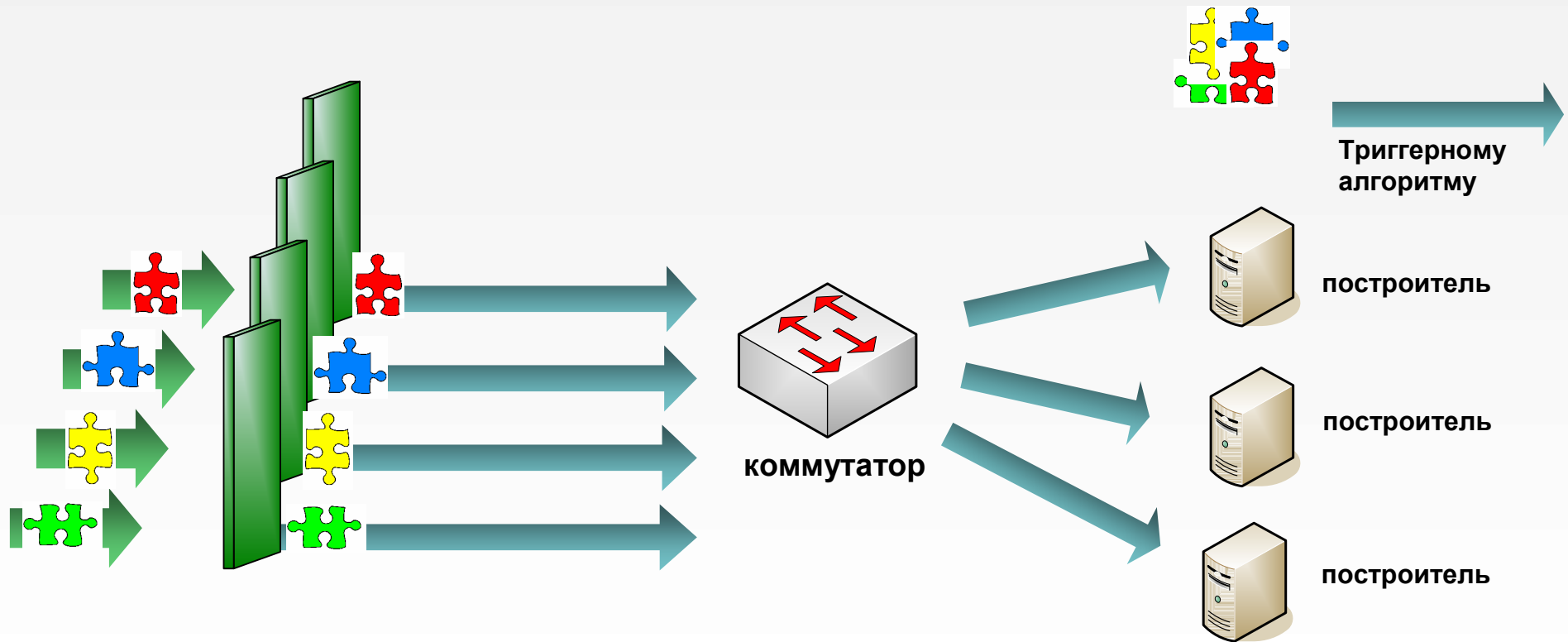
Средний размер события 40 kB
 Средняя частота фермы 1 MHz
 Средняя частота записи 2 kHz

Две философии

- Послать все, спросить потом (ALICE, CMS, LHCb)
- Послать часть, уточнить и послать все в случае интересного события (ATLAS)

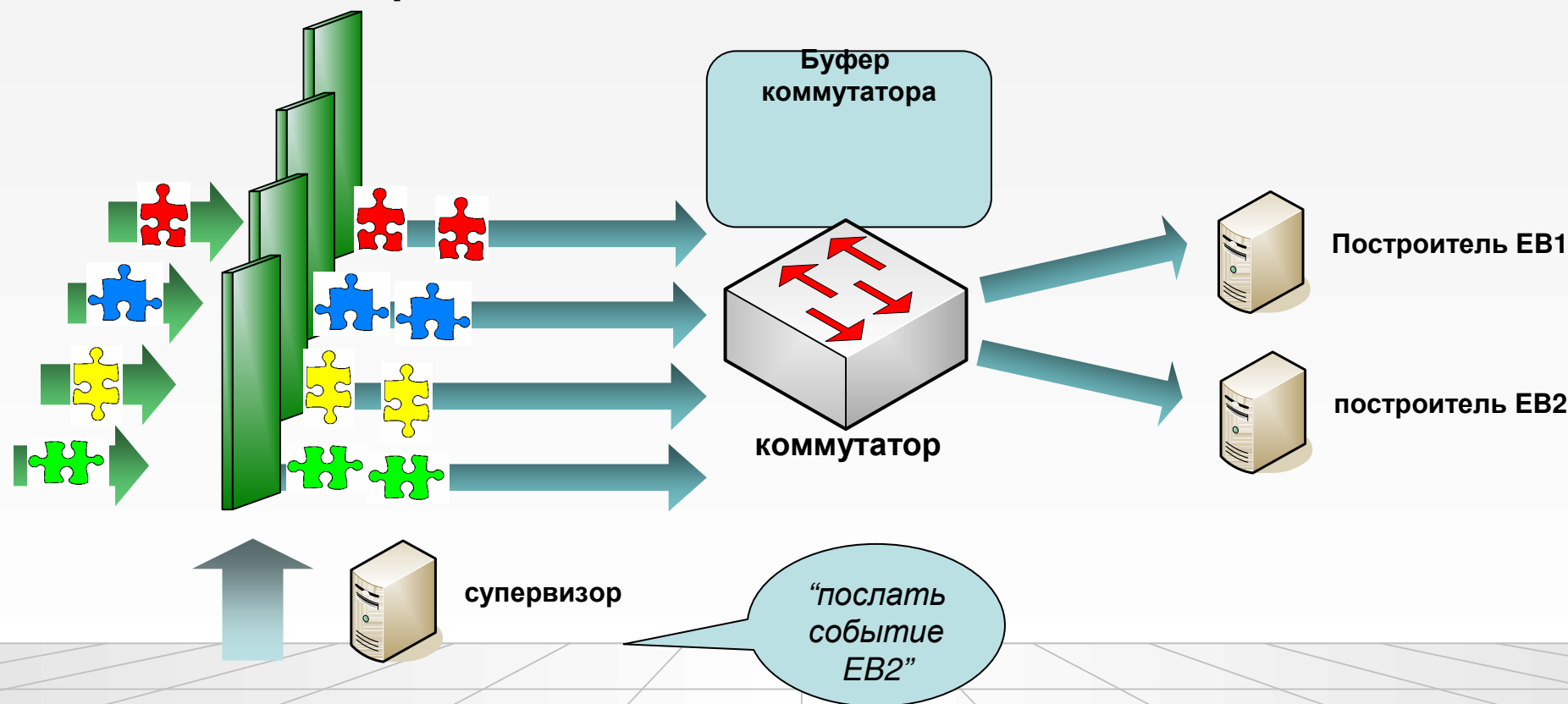


Построение событий



- 1 фрагменты события получены от регистрирующей электроники
- 2 фрагменты события вычитываются по сети
- 3 построитель события собирает фрагменты в единое целое
- 4 Полные события обрабатываются триггерными алгоритмами

Построение события: проталкиванием

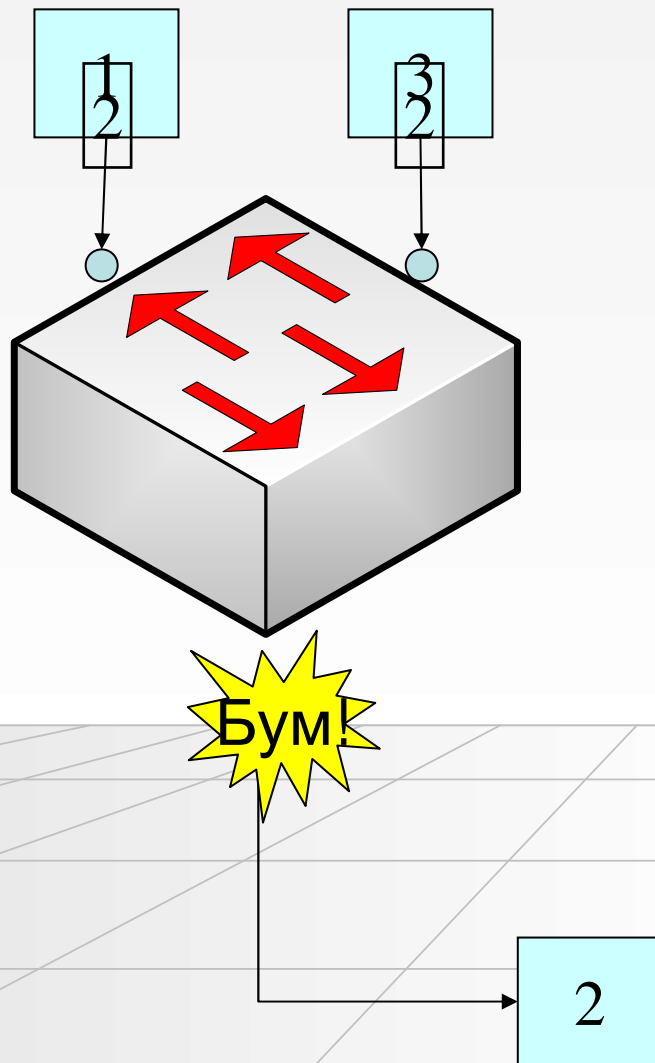


1 Супервизор чтения
говорит куда
посылать
фрагменты

2 Платы чтения без
буферизации,
должен
буферизовать
коммутатор

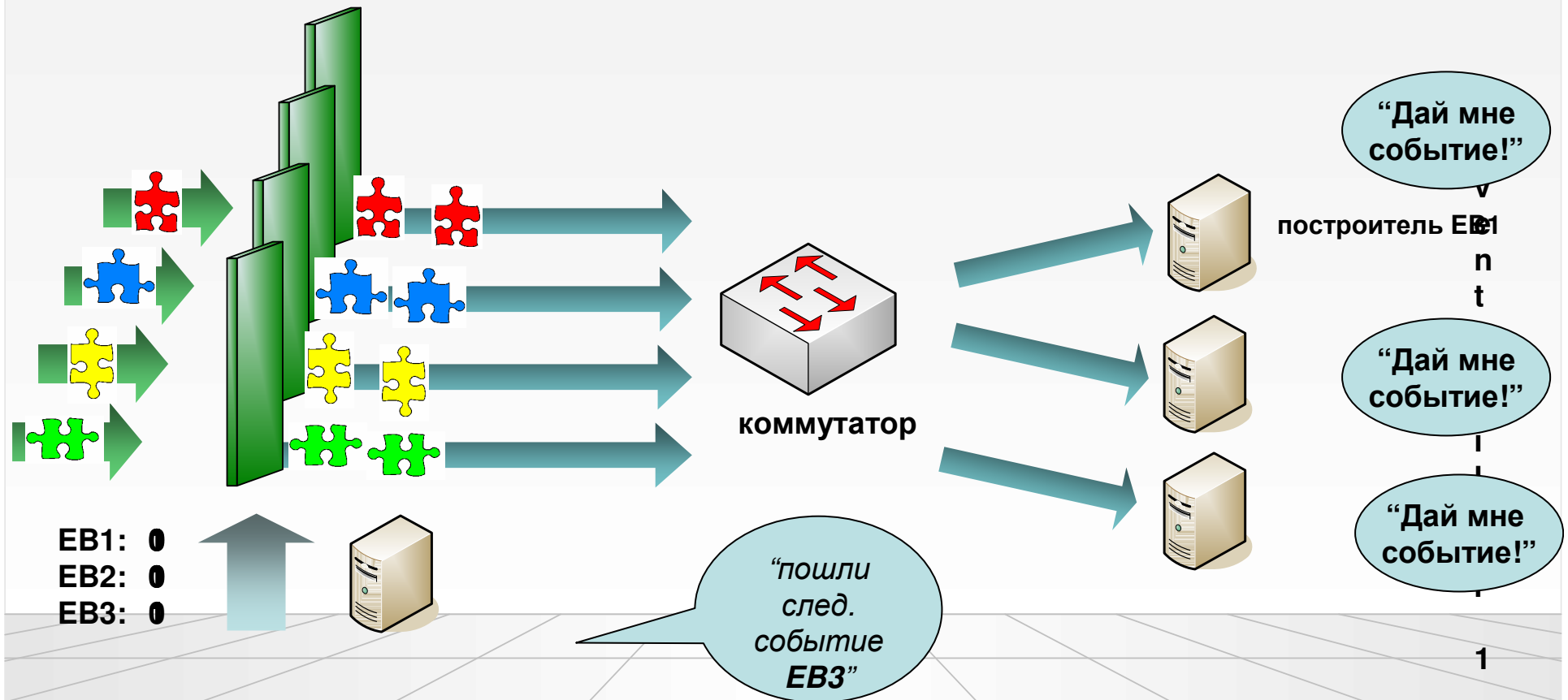
3 Нет обратной связи
от построителей
событий

Перегрузка



- «Бум!» означает неконтролируемую потерю пакетов
- В сети Ethernet это обычное явление
- Протоколы высокого уровня должны справляться с потерей пакетов **отсутствие буферизации**
- Проблема возникает из-за **синхронизованных** источников посылающих данные одному и тому же приемнику **в тоже самое время**

Построение события: вытаскиванием



1 Построители сообщают супервизору о своей загрузке

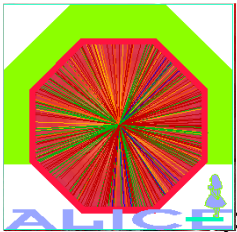
2 Супервизор посылает данные менее загруженным

3 Система сбора опирается на обратную связь от построителей

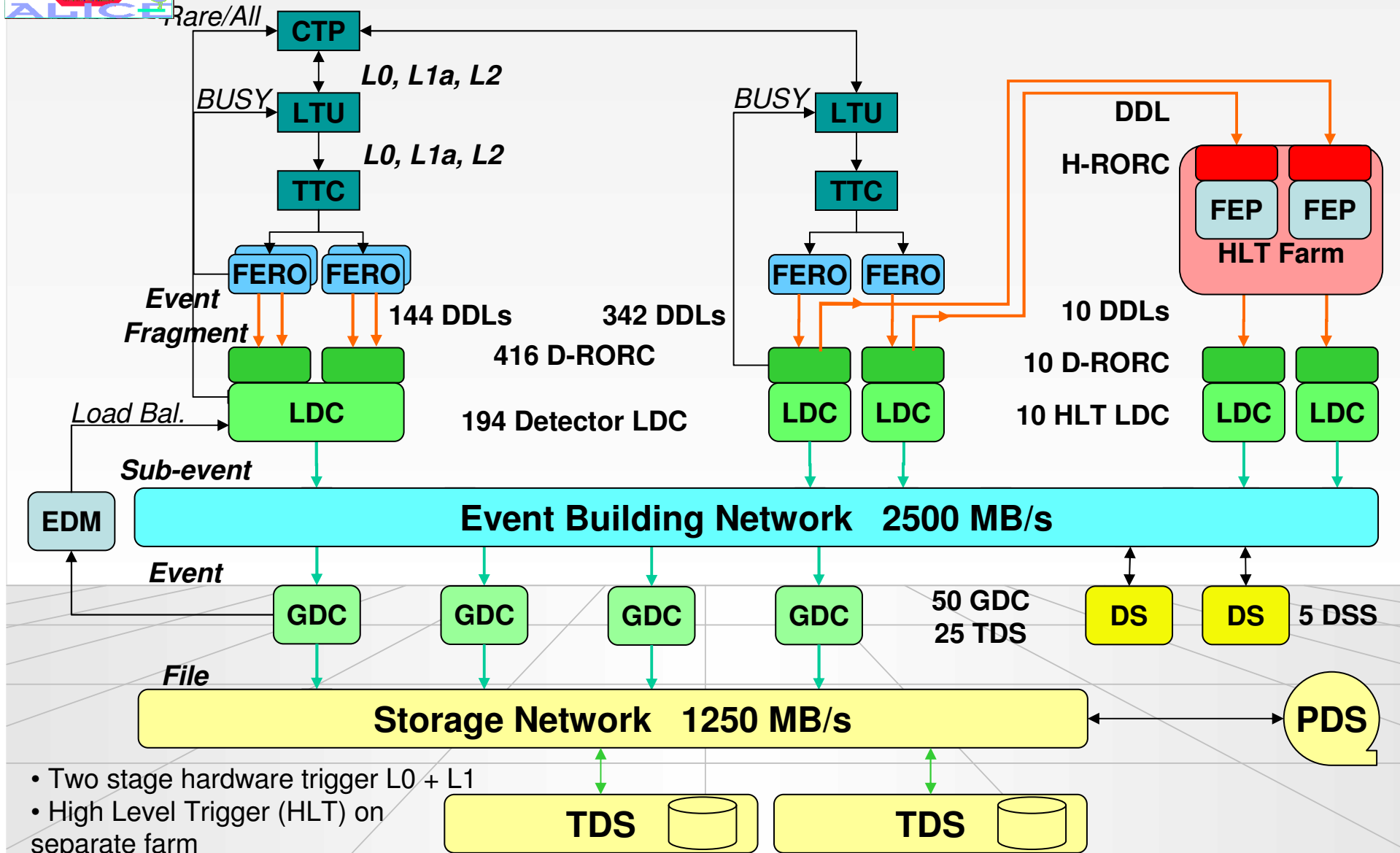
AACL

ALICE, ATLAS, CMS, LHCb

DAQs в 4-х слайдах

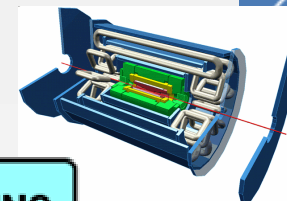


ALICE DAQ



ATLAS DAQ

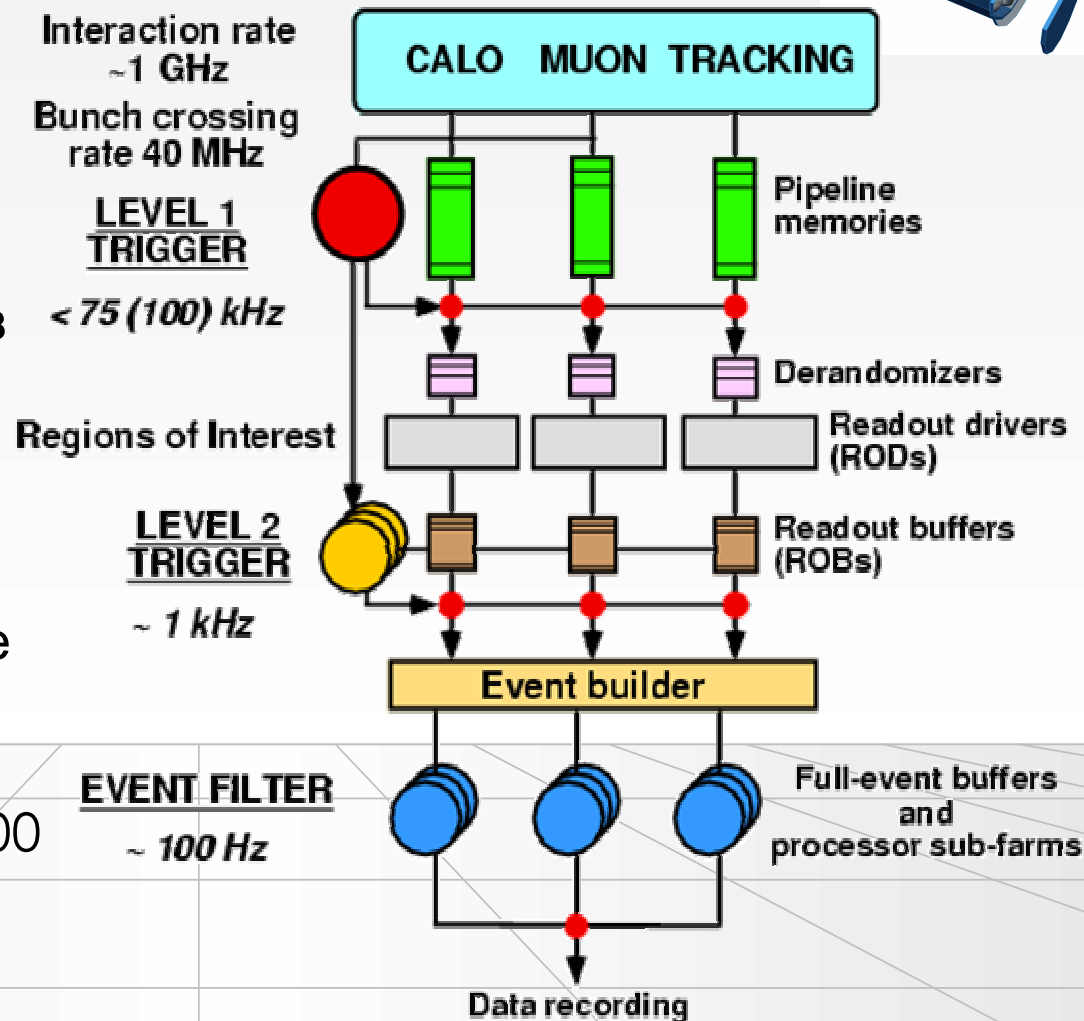
ATLAS



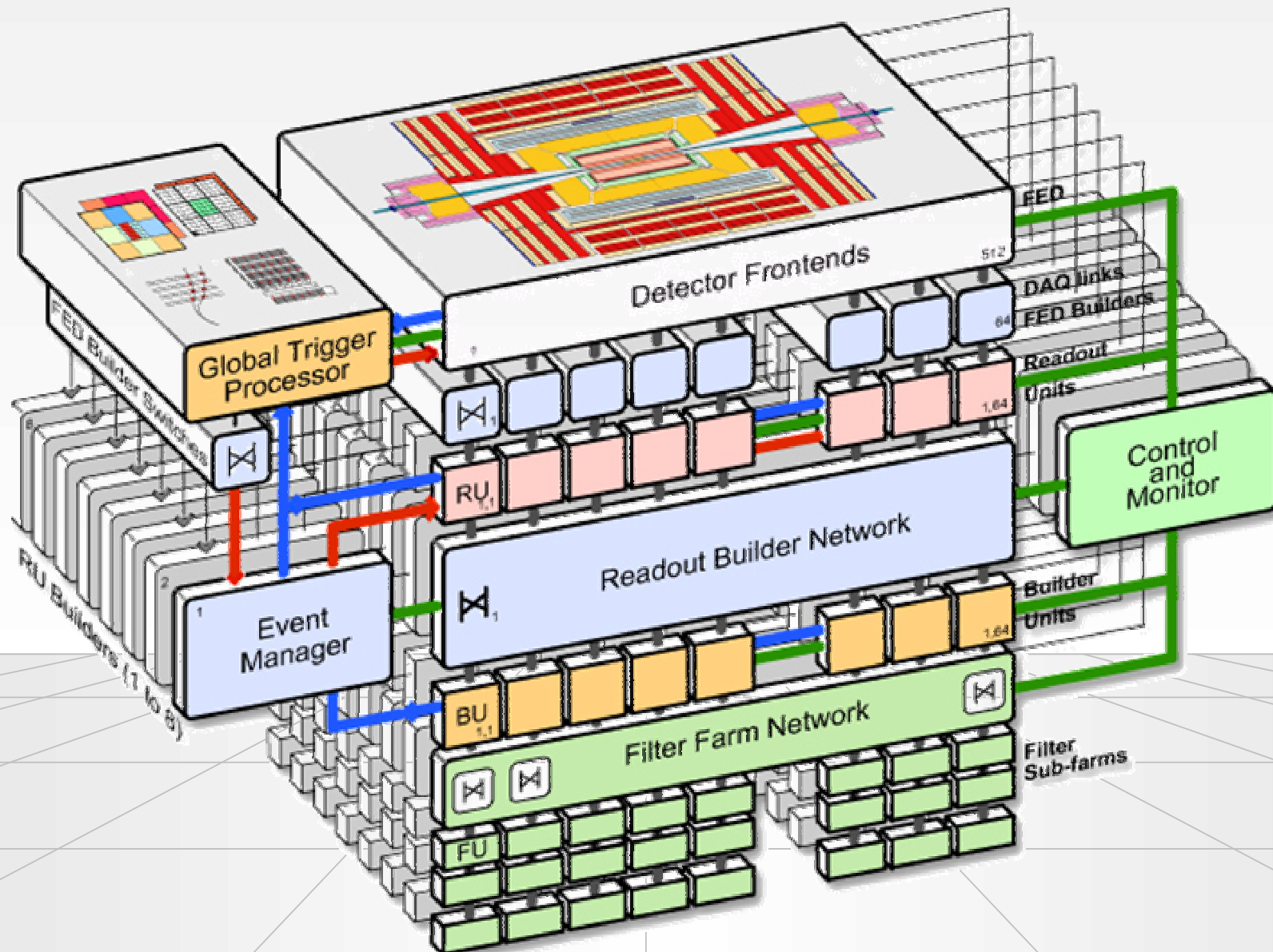
- L1 отбирает события с частотой 100 kHz и обозначает *интересные места (RoI)*

- L2 вытаскивает данные из RoI и обрабатывает их на ферме L2, L2 выдает решения с частотой ~1 kHz

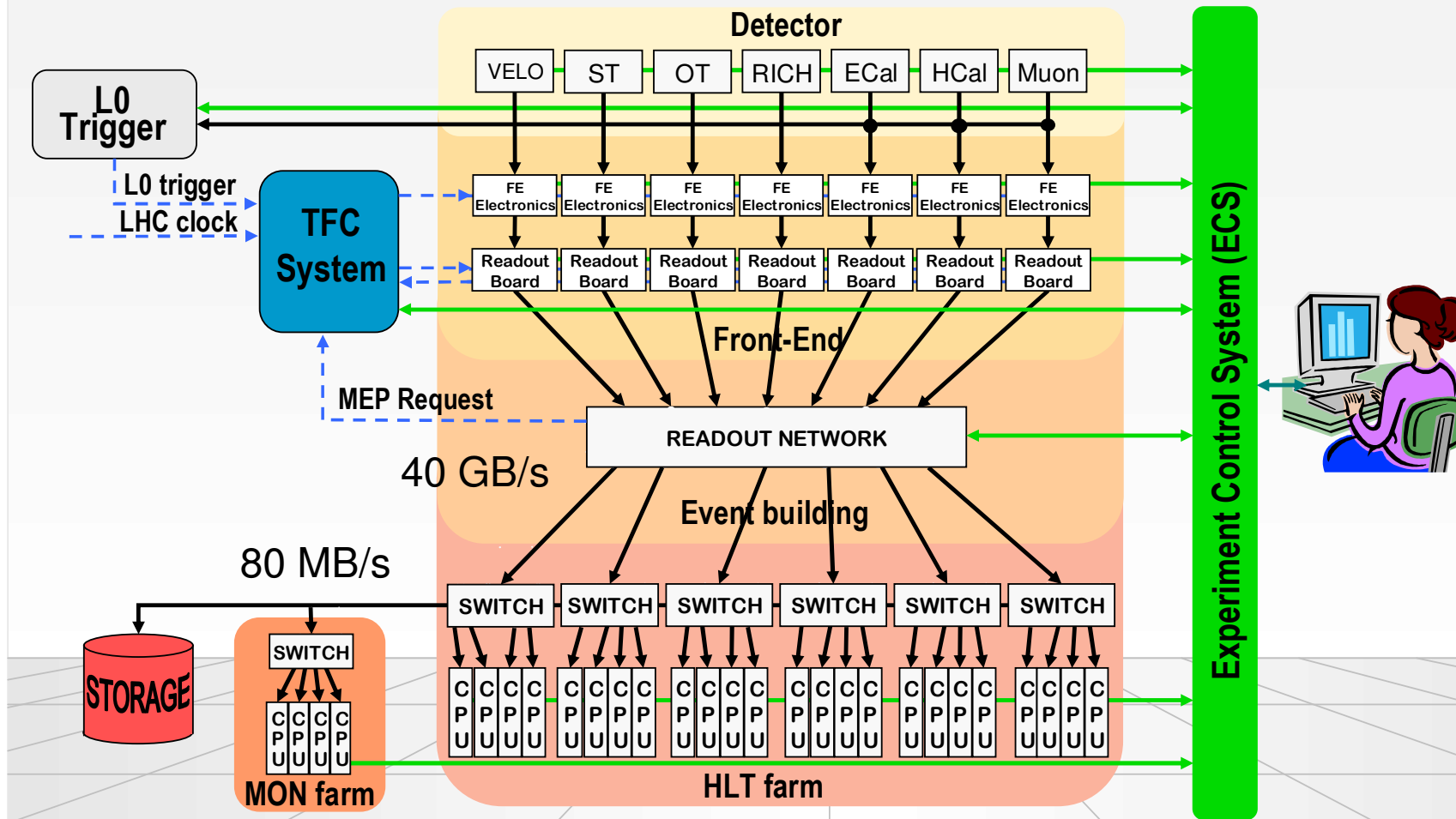
- Event Filter читает данные всего детектора обрабатывает события на ферме и дает на выходе 100 Hz



CMS DAQ

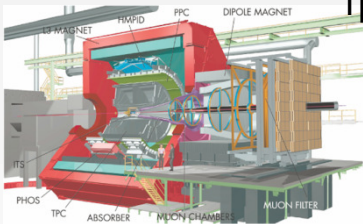
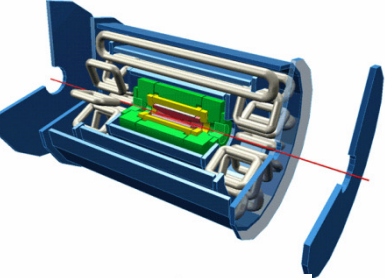
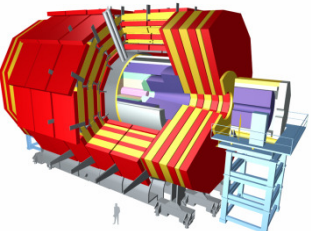
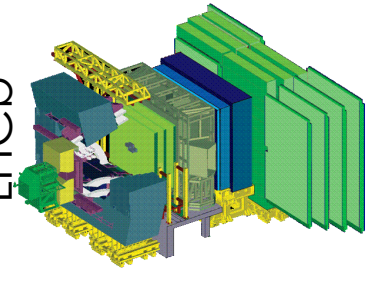


LHCb DAQ



Средний размер события 40 kB
 Средняя частота фермы 1 MHz
 Средняя частота записи 2 kHz

Параметры ССД

	Уровни триггера	L0,1,2 частота (Hz)	Событие размер (Byte)	Чтение полоса.(GB/c)	HLT MB/s (собр./с)
ALICE		Pb-Pb 500	5×10^7	25	1250 (10^2)
		p-p 10^3	2×10^6		200 (10^2)
ATLAS		LV-1 10^5 LV-2 3×10^3	1.5×10^6	4.5	300 (2×10^2)
CMS		LV-1 10^5	10^6	100	~ 1000 (10^2)
LHCb		LV-0 10^6	3.5×10^4	35	70 (2×10^3)

На ленту...а затем в GRID

