

РЫНОК 100GE-УСТРОЙСТВ: от анонсов к реальности

С.Голощупов

О важности и значимости технологии 100 Gigabit Ethernet (100GE) сказано и написано немало. Однако как обстоит дело с ее реальным, коммерческим воплощением? Казалось бы, за два-три года, прошедших с первых анонсов оборудования, рынок должен быть наполнен решениями всех уровней – аппаратурой, унифицированными модулями, элементной базой. Но пока этого не происходит, освоение технологии, по сути, только началось.

Не подлежит сомнению, что в ближайшие годы ожидается значительный рост глобального интернет-трафика. Поэтому понятен интерес к высокоскоростным транспортным протоколам, в числе которых 100GE выглядит одним из наиболее перспективных. Однако практическая реализация этой технологии оказалась значительно сложнее, чем это представлялось вначале, и выпуск серийных устройств на рынок происходит гораздо медленнее прогнозируемого.

Сегодня на рынке в основном представлено дорогое оборудование операторского класса. Причем на рынке в свободном доступе практически нет специализированной элементной базы для 100GE, особенно в части сетевых процессоров. Поэтому каждая компания-производитель оборудования 100GE разрабатывает собственный уникальный чипсет. Задача непростая – ведь необходимо обеспечить высочайшую производительность системы, причем при неравномерной загрузке интерфейса. В частности, при перестановках пакетов внутри одного IP-потока технически сложно распараллелить поддержку одного полнодуплексного 100GE-интерфейса между несколькими сетевыми процессорами. Тем не менее, задача эта решается, пусть и не столь быстро, как ожидалось.

ОБОРУДОВАНИЕ 100GE

Примечательно, что в 2008–2009 годах сразу несколько ведущих производителей объявили о поддержке технологии 100GE. Однако реальные поставки оборудования начались значительно позже.

Так, корпорация Cisco еще в 2008 году объявила об успешных испытаниях оборудования 100GE по существующей оптической инфраструктуре компании

Comcast между городами Филадельфия и Маклин в США. Использовались маршрутизаторы Cisco CRS-1. Однако эта демонстрация не воспроизводила полностью полнодуплексный 100GE-канал, поскольку маршрутизатор CRS-1 поддерживает скорость до 40 Гбит/с на слот.

Первой платформой Cisco, способной обеспечить работу 100GE-интерфейсов, стал маршрутизатор CRS-3 со скоростью 140 Гбит/с на слот. По этой причине первые настоящие испытания 100GE-устройств Cisco состоялись лишь в 2010 году, а о появлении первых коммерческих клиентов (компании AT&T и Comcast) объявили в апреле 2011 года. В июле 2011 года Cisco проводила демонстрации 100GE-интерфейсов на маршрутизаторах с ядром ASR9000, но дата начала поставок не называлась.

Одним из пионеров освоения технологии 100GE выступила и компания Juniper. Еще в июне 2009 года она объявила о поддержке 100GE в своей платформе маршрутизаторов T1600. В 2010 году компания продемонстрировала 100GE-интерфейсы между платформами T1600 и MX3D. В марте 2011 года Juniper анонсировала поставку решений 100GE крупному североамериканскому оператору Verizon, на месяц опередив Cisco. В тот же период Juniper поставляла оборудование и менее крупным клиентам (например, JANET UK) и к середине 2011 года уже располагала существенной базой потребителей 100GE-решений. К концу 2011 года Juniper подготовила сразу два новых магистральных продукта с поддержкой 100GE: обновленных маршрутизаторов T-серии T4000 (240 Гбит/с на слот) и совершенно нового MPLS-коммутатора PTX (480 Гбит/с на слот).

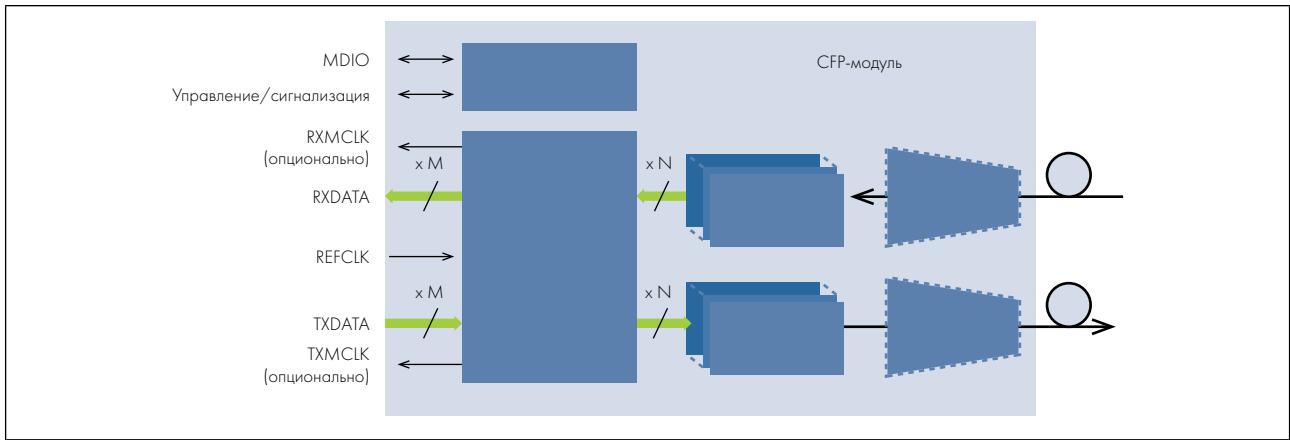


Рис.1. Структура модуля CFP MSA

Интересно, что рынок 100GE-решений для маршрутизаторов в целом повторил ситуацию десятилетней давности с продуктами 10GE. Тогда пионером поставок выступила компания Juniper, на несколько месяцев опередившая своего крупнейшего соперника Cisco.

Чуть позже другая заявила о поддержке 100GE фирма Brocade. Осенью 2010 года она объявила о поддержке 100GE в своей платформе MLXe. Тем не менее, уже летом 2011 года компания анонсировала первый коммерческий запуск технологии 100GE на площадке обмена трафиком AMX-IX в Амстердаме. Для платформы MLXe Brocade выпускает двухпортовую линейную карту в форм-факторе CFP MSA.

Компания Alcatel-Lucent впервые анонсировала 100GE-интерфейсы для маршрутизаторов 7450 ESS/7750 SR в июне 2009 года. Через год, летом 2010, прошли тесты. Однако весной 2011 года 100GE-решения Alcatel-Lucent все еще находились лишь в стадии демонстрационных образцов. Для этого семейства высокоскоростных маршрутизаторов компания разра-

ботала собственный сетевой процессор FP2 с полнуплексной производительностью 50 Гбит/с. Для получения 100GE-потока используют два FP2 без балансировки нагрузки по потокам между ними. В перспективе Alcatel-Lucent планирует перевести свою высокоскоростную платформу 7750 на анонсированный в мае 2011 года процессор FP3 (400 Гбит/с).

Китайская компания Huawei представила свою разработку 100GE-интерфейса для маршрутизатора еще в октябре 2008 года. Следующим шагом фирмы стал анонс законченной системы для передачи 100GE в сентябре 2009 года. Система включала в себя оптический транспорт OSN6800/8800 и одну линейную 100GE-карту в форм-факторе CFP MSA для маршрутизаторов NE5000e на основе процессора собственной разработки Solar 2.0 PFE2A.

В октябре 2010 года Huawei получила контракт на строительство IP/MPLS-сети от российской компании "Мегафон". Что любопытно, в посвященном этому событию пресс-релизе Huawei отчиталась лишь о поставке 40GE-систем NE5000e, с возможностью дальнейшего масштабирования до 100GE на слот. В апреле 2011 года компания выпустила новый анонс линейной карты LPU-200 (2x100GE) для NE5000e на том же чипе Solar 2.0.

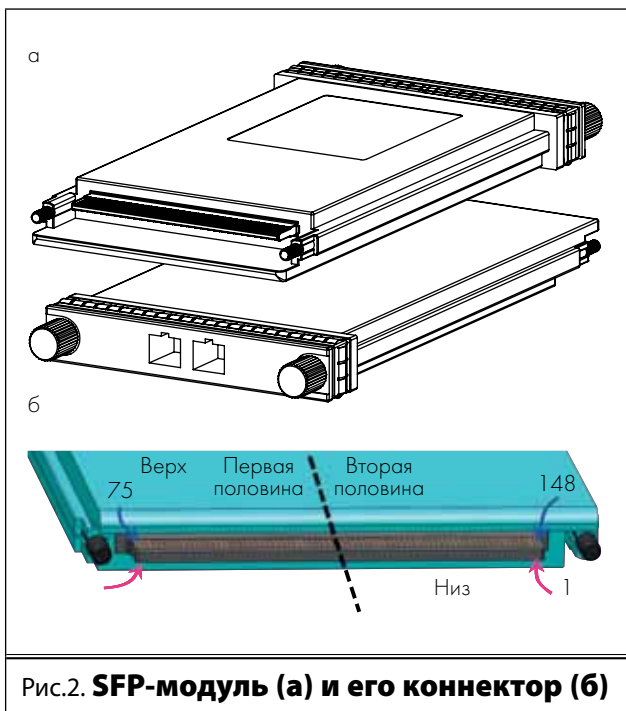


Рис.2. SFP-модуль (а) и его коннектор (б)

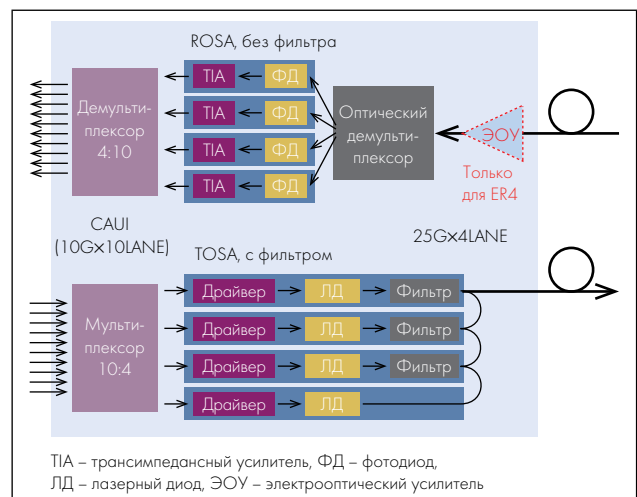


Рис.3. Структура модулей FIM37100 / FIM37200 компании Fujitsu

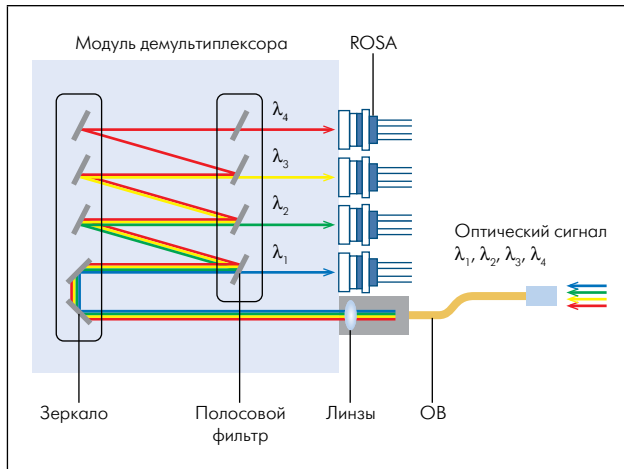


Рис.4. Схема оптического демультиплектора ROSA

В нем приводились данные по поставкам 40GE-версии чипа Solar 1.0 (120 тыс. комплектов), но не было информации по поставкам Solar 2.0.

В августе 2011 года Huawei рапортовала о коммерческой установке 100GE DWDM-систем в сетях операторов KPN и China Telecom, но не указала ни одного

покупателя 100GE-решений на базе NE5000e. А 12 октября Huawei объявила об успешном тестировании решений 100GE на участке действующей сети оператора "Ростелеком" Москва – Самара. Тесты проводились при полной загрузке оптических каналов 100GE-трафиком. По заявлению участников, все пункты программы испытаний были выполнены успешно и в полном объеме. По результатам тестирования оборудование признано готовым к коммерческому применению и может быть рекомендовано для строительства современных высокоскоростных магистральных ВОЛС большой протяженности.

Таким образом, оборудование 100GE уже стало промышленной реальностью. Но пока оно чрезвычайно дорого, поскольку доступно очень немногим производителям. Ситуация изменится, когда на рынке в свободном доступе появится элементная база (прежде всего – трансиверы и сетевые процессоры) и унифицированные модули. Первый шаг в этом направлении – появление форм-фактора CFP MSA.

ФОРМ-ФАКТОР CFP MSA

Возросшие требования к размерам и охлаждению оптической части для 100GE создали дополнительные трудности в разработке линейных карт стандарта IEEE P802.3ba. В условиях дефицита на рынке

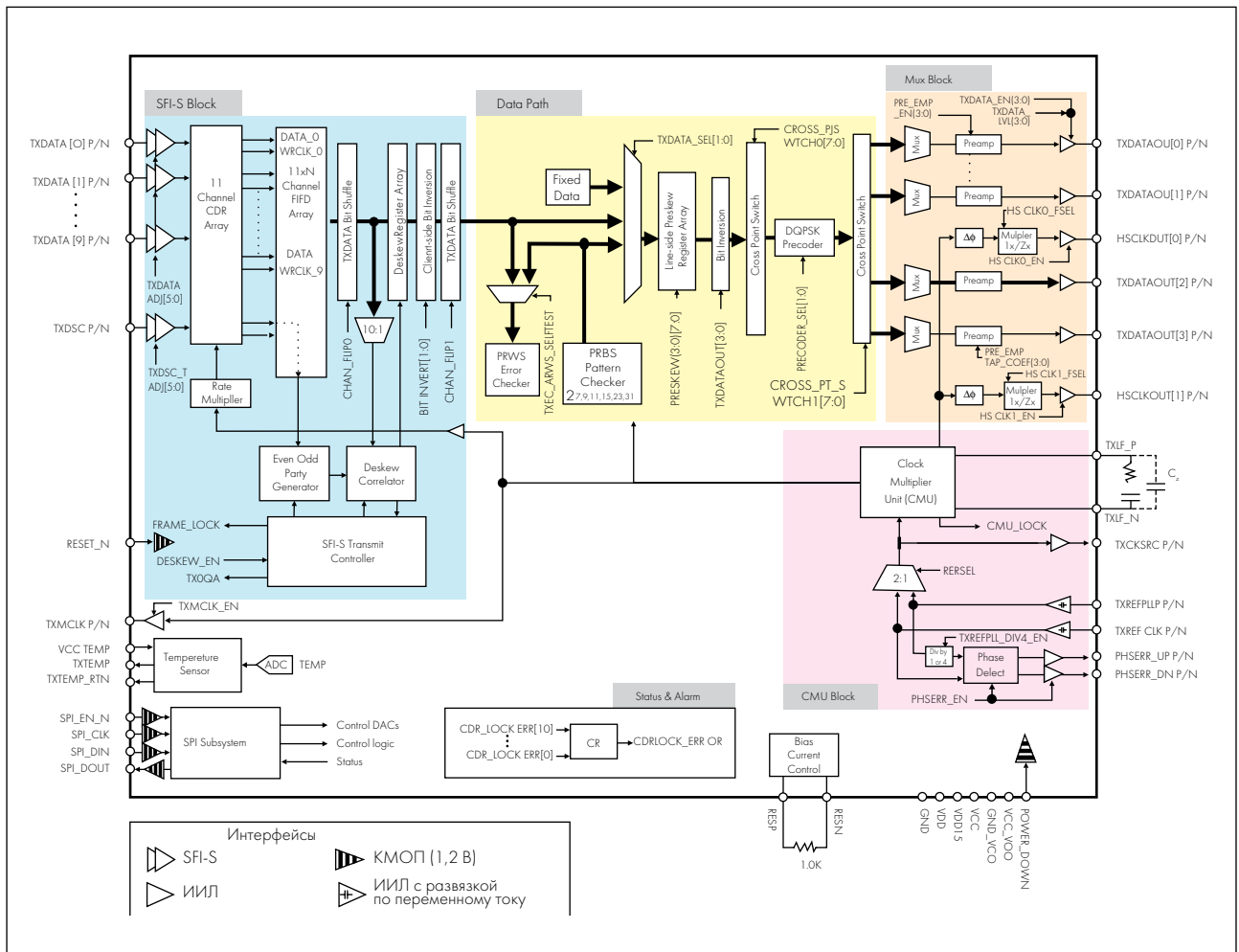


Рис.5. Мультиплексор SMI10021 100GE 10:4 компании Semtech

100GE-трансиверов поставщики 100GE-оборудования вынуждены вести совместные разработки оптоэлектроники, чтобы уложиться в жесткие линейные и энергетические ограничения современных сетевых устройств. Один из примеров такого сотрудничества – разработка стандарта встраиваемых модулей форм-фактора CFP MSA (C Form-factor Pluggable, Multi-Source Agreement) фирмами Avago, Finisar, Fujitsu, Opnext и Sumitomo Electric. Применение модулей форм-фактора CFP MSA [1] позволяет ускорить проектирование устройств, экономя время и ресурсы по сравнению с собственными разработками на дискретных компонентах. Кроме того, использование в разработке слота для модуля форм-фактора CFP MSA делает архитектуру устройства более унифицированной, легко адаптируемой под разнообразные типы оптических интерфейсов.

Модуль форм-фактора CFP MSA позволяет подключить от трех интерфейсов XLAUI* (40 Gigabit Attachment Unit Interface) к одному интерфейсу OTL4.10 (Optical channel Transport Lane level 4). Это возможно за счет использования стандартной скорости одного потока – 10 Гбит/с (точнее – 10,3125 Гбит/с), а элементная база для обработки таких потоков уже давно разработана. Модуль для физических интерфейсов 100GBASE-LR4 и 100 GBASE-ER4 преобразовывает десять потоков по 10 Гбит/с от CAUI* в четыре потока по 25 Гбит/с (рис.1).

Оптический передатчик для интерфейсов 100GBASE-LR4 и 100 GBASE-ER4 выполнен на базе лазера с электроабсорбционным модулятором с распределенной обратной связью (EA-DFB – Electro-Absorption Distributed FeedBack). Но допускается также применение и обычных DFB-лазеров. В качестве оптического приемника трансивера обычно выступает PIN-фотодиод.

Одна из особенностей модулей CFP MSA – встроенный разъем особой конструкции, допускающий горячую замену модуля без отключения питания устройства. Самый мощный класс модулей CFP MSA рассеивает мощность около 32 Вт, что дает ток в 10 А при напряжении питания в 3,3 В. Такие электрические характеристики подключаемых 100GE-модулей, особенно для физических интерфейсов 100GBASE-LR4

(дальность 10 км) и 100GBASE-ER4 (дальность 40 км), создают дополнительную проблему – большой ток при их отключении от устройства в режиме горячей замены. Большой ток в момент подключения/отключения способен повредить устройство, если не предпринять специальных мер.

Для уменьшения электрической дуги, возникающей при отключении модуля CFP в момент его горячей замены, электрический разъем имеет разную длину контактов. Два самых коротких называются "контакт отсутствия модуля" и "контакт малой мощности". С помощью контакта малой мощности модуль может определить момент своего извлечения из устройства и запустить процедуру быстрого выключения питания. Тем самым удается существенно снизить ток потребления в момент разъединения контактов питания в разъеме, что позволяет избежать электрической дуги. Контакты отсутствия модуля и малой мощности требуют другой конструкции разъемов по сравнению, например, с устройствами 10GE. Так, если раньше в качестве разъема служил край печатной платы (слот), то теперь в модуле CFP MSA предусмотрен специальный встроенный разъем с четко определенными механическими размерами контактов (рис.2). Такая конструкция обеспечивает более качественное электрическое соединение, что улучшает производительность всей системы за счет уменьшения помех в разъеме.

В целом, конструкция получилась удачная, и на данный момент большая часть линейных карт 100GE выполнена в стандарте CFP MSA. Его поддерживают практически все производители оборудования. Причем как со стороны шасси, так и собственно линейных карт 100GE. Уже сегодня можно купить модули форм-фактора CFP MSA компаний FCOM, Fujitsu, Reflex Photonics, XENA Networks.

ЭЛЕМЕНТНАЯ БАЗА – ПЕРВЫЕ ЛАСТОЧКИ?

Рассмотрим устройство модулей форм-фактора CFP MSA на примере трансиверов FIM37100/FIM37200 компании Fujitsu (рис.3). Мультиплексор преобразует 10 потоков по 10 Гбит/с от интерфейса CAUI в четыре потока по 25 Гбит/с, которые далее поступают в оптический мультиплексор TOSA (Transmitter Optical Sub-Assembly) и передаются в оптическое волокно. В обратном направлении все происходит аналогично, сигнал из волокна обрабатывается сначала оптическим мульт-

* XLAUI и CAUI – это описанный в стандарте IEEE P802.3ав физический внутренний интерфейс [2].

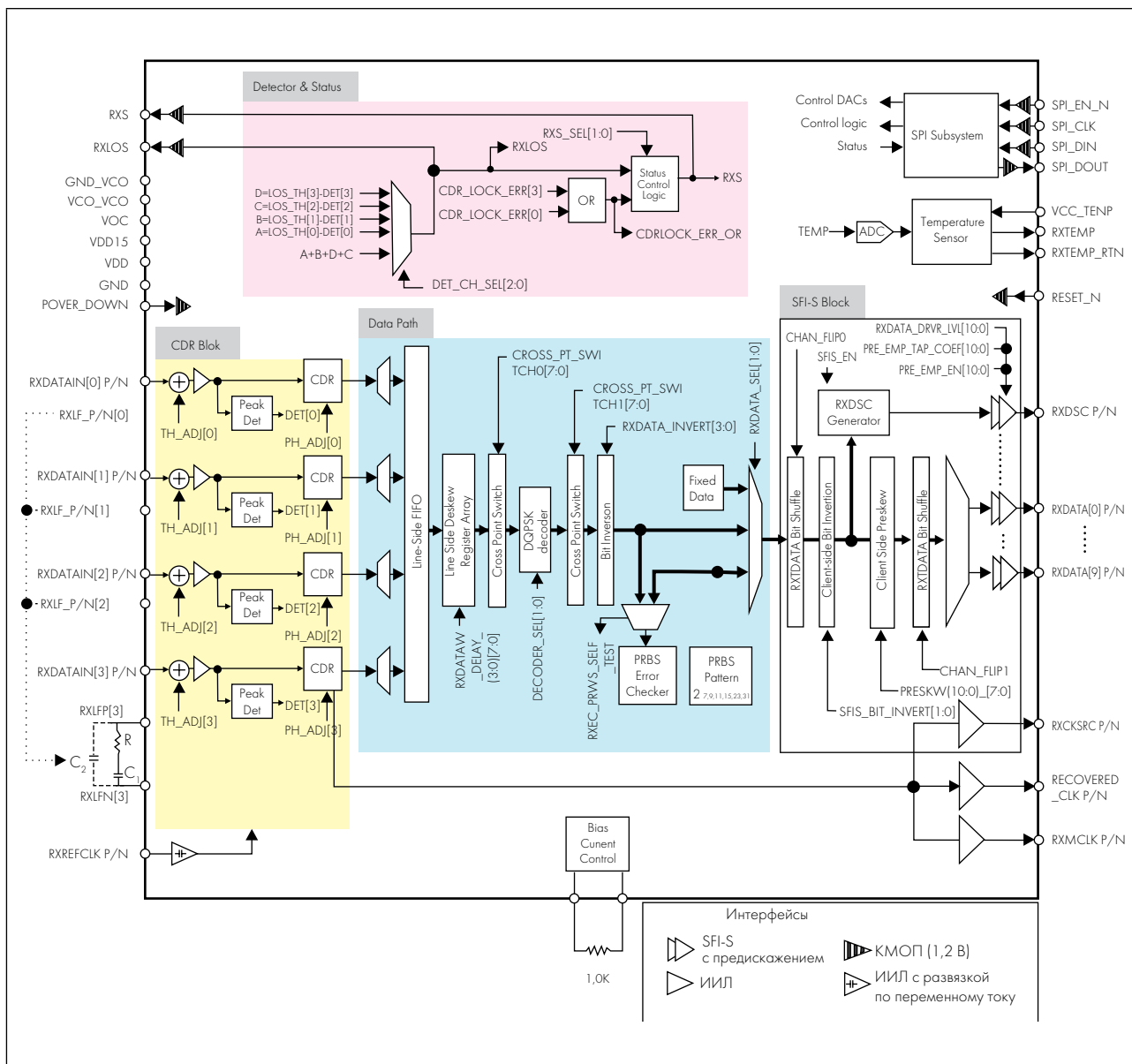


Рис.6. Демультимплексор SMI10031 100GE 4:10 компании Semtech

типлексором, затем поступает в оптический приемник ROSA (Receiver Optical Sub-Assembly) и в мультиплексор 4:10.

В оптическом приемнике ROSA входной поток поступает на зигзагообразную структуру зеркал (рис.4). Зеркала пропускают световой поток заданной длины волны, выполняя роль оптических фильтров. В оптическом мультиплексоре TOSA происходит обратный процесс, поток от нескольких лазеров через оптические фильтры смешивается в общий выходной поток.

Компания Fujitsu для своих модулей использует компоненты собственной разработки. Но на рынке появляется элементная база и сторонних производителей. Например, компания Semtech анонсировала в марте 2011 года чипсет для мультиплексоров SMI10021 (10:4,

10 потоков по 10 Гбит/с в четыре потока по 25 Гбит/с) (рис.5) и демультимплексоров SMI10031 (4:10, 4x25 Гбит/с в 10x10 Гбит/с, рис.6).

В заключение отметим, что процесс воплощения 100GE в аппаратуре происходит гораздо медленнее, чем предполагалось, тем не менее, динамика явно положительная. Постепенно рынок наполняется компонентами различного уровня – модулями, чипсетам, оптическими сборками, что в свою очередь подстегивает процесс разработки и выпуска на рынок новых устройств стандарта IEEE P802.3ba.

ЛИТЕРАТУРА

1. CFP MSA Hardware Specification Revision 1.4 7 June 2010 /Ed. M.Traverso. – www.cfp-msa.org
2. Голощапов С., Шахнович И. 100GB Ethernet: основные принципы. – Первая миля, 2011, №3, с.22– 29.



РОССИЙСКАЯ
НЕДЕЛЯ
ЭЛЕКТРОНИКИ

ОТ КОМПОНЕНТОВ ДО МОДУЛЕЙ
ЭЛЕКТРОННОЙ АППАРАТУРЫ!



МОСКВА, ЭКСПОЦЕНТР, 1-3 НОЯБРЯ 2011
www.RussianElectronicsWeek.ru

MOBILE & WIRELESS

МЕЖДУНАРОДНАЯ ВЫСТАВКА

БЕСПРОВОДНЫЕ И МОБИЛЬНЫЕ ТЕХНОЛОГИИ

INTERNATIONAL EXHIBITION

MOBILE & WIRELESS TECHNOLOGIES



Москва, Экспоцентр, 1-3 ноября 2011 г.

Организатор:
ЗАО "ЧипЭКСПО"
(495)221-50-15, info@chipexpo.ru
<http://chipexpo.chipexpo.ru/wm.html>