

SAS – ПОСЛЕДОВАТЕЛЬНЫЙ ИНТЕРФЕЙС ДЛЯ ПОДКЛЮЧЕНИЯ УСТРОЙСТВ ХРАНЕНИЯ ДАННЫХ

Яковлев Д.С.

Научный руководитель – к.т.н., доцент Середкин В.Г.

Сибирский федеральный университет

В настоящее время большое внимание уделяется разработке подсистем хранения информации. В первую очередь идет увеличение объема хранимой информации на одном устройстве. Но кроме развития количественной характеристики накопителей, происходит и качественный рост, заключающийся в увеличении скорости обработки информации. Особенно это затрагивает системы уровня предприятий и высоконагруженных серверов, т. к. в них предъявляются очень жесткие требования к скорости работы системы. Но просто увеличить скорость работы самого устройства на данный момент уже недостаточно, поэтому на данный момент разработаны высокоскоростные интерфейсы, позволяющие производить обмен информации на скоростях, соответствующих возможностям устройств.

В данной работе рассматривается последовательный интерфейс SAS – замена высокопроизводительного параллельного интерфейса SCSI. С другой стороны, данный интерфейс разработан на основе физического интерфейса SATA и позволяет подключать устройства с интерфейсом SATA без дополнительных аппаратных или программных затрат. Данный интерфейс предназначен для работы в первую очередь в системах хранения данных предприятий.

Данный интерфейс включает три протокола:

- SSP (Serial SCSI Protocol) – работа с целевыми устройствами стандарта SAS. Это основной протокол, используемый в данном интерфейсе;
- STP (Serial ATA Tunneling Protocol) – работа с целевыми устройствами стандарта SATA. Данный протокол обеспечивает совместимость при работе с устройствами SATA, подключенными к системе на основе SAS;
- SMP (Serial Management Protocol) – работа с экспандерами (устройствами-расширителями), обеспечивающими связь устройства-инициатора с целевыми устройствами, подключенными в разветвленную конфигурацию.

Подключение устройства SAS обеспечивает трансивер Phy – приемник и передатчик последовательного интерфейса. Приемник и передатчик работают на одинаковых скоростях в полно дуплексном режиме (независимо друг от друга). Каждый ресивер в устройстве имеет собственный номер. Трансиверы устройства могут быть объединены в порты, имеющие свой уникальный адрес – глобальный уникальный 64-битный идентификатор.

Все SAS-устройства делятся на два класса:

1. Конечные устройства – к ним относятся как целевые устройства, так и хост-адаптеры;
2. Устройства-экспандеры – промежуточные устройства, служащие для объединения конечных устройств в иерархические структуры и увеличения количества устройств, которые могут быть подключены к хост-адаптеру.

Перед началом передачи данных между устройствами необходимо установить соединение. Для установления соединения инициатор посылает запрос, указывая адрес желаемого устройства. Но, хотя запрос идет к адресу порта, само соединение

устанавливается между трансиверами. Таким образом, если устройство имеет более одного трансивера, то возможна параллельная работа с этим устройством через каждый трансивер независимо от загрузки других трансиверов.

Архитектурная модель SAS состоит из следующих уровней:

- Физический уровень – определяет физические и электрические параметры приемников, передатчиков и кабелей;
- Интерфейс физического уровня – определяет способ кодирования данных и специальную «внеполосную» сигнализацию для служебных целей;
- Канальный уровень – обеспечивает идентификацию подключенных устройств и управления соединениями;
- Транспортный уровень – определяет структуры кадров и транспортные сервисы;
- Прикладной уровень – определяет процедуры выполнения команд SCSI в SSP и функции определения топологии и управления экспандерами в SMP.

Рассмотрим подробнее интерфейс физического уровня и канальный уровень. На уровне физического интерфейса определяется логическое кодирование 8B10B, «внеполосная» сигнализация и согласование скоростей. Все передачи организуются в виде двойных слов (32 бит). Внеполосная сигнализация используется для обнаружения факта подключения устройства, определения его типа (SAS, SATA) и согласования скорости интерфейса. На канальном уровне обеспечивается идентификация устройства, т.е. определение типа устройства (конечное, экспандер) и адреса порта, а так же производится согласование скоростей путем добавления в передаваемый поток специальных заполнителей, не несущих информативности.

Канальный уровень так же обеспечивает управление соединением. Для открытия соединения инициатор посылает запрос либо непосредственно целевому устройству, если оно подключено непосредственно к инициатору, либо экспандеру, через который это целевое устройство подключено. Целевое устройство либо подтверждает, либо отклоняет запрос на соединение, но если между ними находится экспандер, то как только к нему поступил запрос, он пересылает его далее, а обратно посылает специальный сигнал, информирующий инициатора о том, что запрос находится в процессе арбитража. Для закрытия соединения необходимо, чтобы приемник и передатчик переслали друг другу сообщения об окончании передачи, после чего оба устройства передают сигнал о закрытии соединения.

В случае, если к одному порту устройства приходит одновременно несколько запросов на соединение, то преимущество отдается запросу с наибольшим временем ожидания. Используя такой вариант доступа к разделяемому ресурсу, обеспечивается справедливость арбитража и ограниченная задержка обслуживания.

Для гарантированной доставки данных, передаваемых по интерфейсу SAS, для каждого кадра данных, полученных целевым устройством, генерируется квитанция о принятии, либо непринятии кадра. Инициатор, после отправки кадра, ожидает квитанции в течении 1 мс, если в течении этого времени пришла квитанция о принятии, то кадр считается успешно отправленным, иначе, если пришла квитанция о непринятии, либо квитанция не пришла в течении 1 мс, то кадр считается не переданным и влечет за собой закрытие соединения. Передача данных является не блокирующей, т.е. передатчик не ожидает приема квитанции об успешной доставке предыдущего кадра, а сразу же передает следующий.

Управление потоком передачи основано на кредитах. Для того, чтобы передать кадр данных, передатчик должен иметь ненулевой кредит. Кредит передатчику выдает приемник, при этом приемник должен иметь возможность принять полноразмерный кадр данных. Значение кредита может достигать 255.

Отдельным вопросом стоит работа с устройствами SATA, т.к. интерфейс SATA не предусматривает наличие соединения. В следствие этого возможны два варианта работы: если устройство SATA непосредственно подключено к хост-адаптеру – тогда обмен данными происходит полностью по интерфейсу SATA; если же устройство SATA подключено к хост-адаптеру через экспандер, то работа происходит по протоколу STP, который открывает соединение, затрачивая на это дополнительное время, но в дальнейшем работа происходит напрямую, используя команды интерфейса SATA.

Рассмотренный интерфейс SAS на данный момент является одним из самых перспективных стандартов. На сегодняшний день уже выпускаются устройства, поддерживающие вторую версию стандарта, позволяющие работу на скорости до 6 Гбит/с, что позволяет, при использовании устройств с четырьмя трансиверами, получить суммарную скорость до 2,4 Гбайт/с. Данная скорость все же ниже интерфейса Fiber Channel, но при этом интерфейс SAS значительно дешевле, что играет немаловажную роль при выборе используемого интерфейса на предприятии.

Данный интерфейс SAS используется в комплексе высокопроизводительных вычислений ИКИТ СФУ в подсистеме хранения информации, реализованном на серверах IBM DS 3400 в количестве 48 жестких дисков по 300 Гбайт (4 сервера по 12 дисков) и в серверах управления IBM X3650 общим объемом 1168 Гбайт. В данной системе для подключения жестких дисков с интерфейсом SAS не используются экспандеры, т.к. подключение организовано с помощью входящих в состав сервера RAID-контроллеров.

В заключении можно сказать, что интерфейс SAS имеет множество преимуществ, по сравнению своим предшественником – параллельным интерфейсом SCSI. Основными преимуществами являются:

- Значительный прирост в скорости передачи данных;
- Значительное удешевление аппаратной части;
- Более комфортное обслуживание дисковых массивов, за счет возможности их вынесения за пределы местонахождения сервера.