

AoE (ATA over Ethernet)

Статус документа

В этом документе приведена спецификация облегченного протокола доступа к устройствам ATA через сеть Ethernet.

1. Введение

AoE используется для поддержки очень простого механизма удаленного вызова процедур (RPC¹) между клиентом и сервером в виде устройства ATA. Сервер воспринимает команды и генерирует отклики на базе кода команд в заголовке AoE.

В этом документе описывается формат заголовка AoE и набор команд для первой версии протокола. Все значения, для которых существенен порядок байтов, приводятся с использованием сетевого порядка байтов². Для зарезервированных полей при передаче любого сообщения должны устанавливаться нулевые значения.

Каждое сообщение содержит заголовок, за которым следует поле аргумента. Формат поля аргумента определяется кодом команды в заголовке.

1.1. Соединения

Протокол AoE не базируется на прямых соединениях³. Каждое сообщение, переданное серверу, следует рассматривать, как уникальное, доставка которого не гарантируется.

2. Формат заголовка AoE

0	1	2	3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1			
+-----+			
0	Ethernet Destination Addr		
+-----+			
4	Ethernet Destination Addr		Ethernet Source Addr
+-----+			
8	Ethernet Source Addr		
+-----+			
12	Ethernet Type (0x88A2)		Ver Flags Error
+-----+			
16	Major		Minor Command
+-----+			
20	Tag		
+-----+			
24	Arg		
+-----+			

2.1. Адреса Ethernet Source, Destination и Ethernet Type

Стандартный заголовок Ethernet MAC, используемый для кадров IEEE 802.3 Ethernet.

Протокол AoE использует зарегистрированное значение типа Ethernet - 0x88A2.

2.2. Ver (Версия)

Поле номера версии определяет формат заголовка AoE, а также набор кодов команд. Все будущие версии протокола должны содержать это поле в заголовке с сохранением позиции. В этом документе описывается протокол версии 1.

Коды команд для этой версии протокола описаны в параграфе 2.6. Command.

Отклик Query Config Information включает номер версии протокола, поддерживаемой сервером.

2.3. Флаги

```
+---+---+
|R|E|Z|Z|
+---+---+
```

Поле Flags содержит биты флагов, описанные ниже.

Бит R устанавливается для сообщений, являющихся откликами. Бит E устанавливается в сообщениях с откликами, если соответствующее отклику сообщение с командой привело к возникновению ошибки AoE. Флаги Z являются резервными и должны иметь значение 0.

¹Remote Procedure Call. Прим. перев.

²От старшего байта к младшему - т. е., первым указывается старший байт. Этот порядок еще называют big-endian. Прим. перев.

³Connectionless. Прим. перев.

2.4. Error

Если в поле Flags установлен бит E, это поле содержит один из определенных ниже кодов ошибок.

Error 1: Unrecognized command code — неизвестный код команды

Сервер не понял кода команды, заданного в поле Command.

Error 2: Bad argument parameter — некорректный параметр

В поле Arg имеется неприемлемое значение.

Error 3: Device unavailable — устройство недоступно

Сервер больше не способен воспринимать команды ATA.

Error 4: Config string present — присутствует строка конфигурации

Сервер не может установить строку конфигурации, поскольку она не пуста (см. параграф 3.2).

Error 5: Unsupported version — неподдерживаемая версия

Сервер не понимает номер версии протокола, указанный в поле Ver.

Error 6: Target is reserved - «цель» забронирована.

Команда не может быть выполнена, поскольку «цель» забронирована (см. параграф 3.4).

2.5. Major, Minor

Каждый сервер AoE обладает двумя адресами — старшим (major) и младшим (minor). До обработки поля Command в заголовке сервер должен сравнить свои адреса с соответствующими полями Major и Minor в заголовке. Сервер будет принимать командное сообщение для обработке при выполнении обоих приведенных ниже условий:

значение поля Major в заголовке совпадает со старшим адресом сервера или содержит только единицы (0xffff);

значение поля Minor в заголовке совпадает с младшим адресом сервера или содержит только единицы (0xff).

Если любое из условий не выполняется, сервер должен игнорировать команду.

Сервер должен включать свой старший и младший адрес в каждый отклик, даже в тех случаях, когда адрес в соответствующем запросе содержал только единицы.

2.6. Command

Это поле содержит код команды. В данной версии протокола определены следующие коды команд:

Command 0: Issue ATA Command — ввести команду ATA;

Command 1: Query Config Information — запросить конфигурационные данные;

Command 2: Mac Mask List — список масок MAC;

Command 3: Reserve / Release – обратить, освободить.

Коды 240-255 (0xf0-0xff) зарезервированы для конкретного использования производителями.

2.7. Tag

Поле Tag позволяет клиенту устанавливать корреляцию откликов с соответствующими командами. Это поле сервер копирует в отклик из сообщения с командой; при обработке команды сервером значение этого поля не используется.

2.8. Arg

Поле Arg содержит входные данные для заданной кодом команды.

3. Коды команд

3.1. Command 0, Issue ATA Command

	0								1								2								3															
	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
24	AFlags								Err/Feature								Sector Count								Cmd/Status															
28	lba0								lba1								lba2								lba3															
32	lba4								lba5								Reserved																							
36	Data																																							

Команда 0 используется для ввода команды ATA в подключенное к серверу устройство ATA.

Все данные, связанные с командой, должны помещаться в одно сообщение. Максимальный размер стандартного кадра Ethernet (1520 байтов) ограничивает возможность однократного считывания или записи двумя секторами.

Операции чтения/записи ATA должны быть защищены списком бронирования, как описано в параграфе 3.4. Если доступ к «цели» ограничен бронированием, должно возвращаться сообщение об ошибке AoE с кодом 6.

Формат поля AoE Arg показан на рисунке выше.

Поле AFlags показано на рисунке.

```

+-----+-----+
|Z|E|Z|D|Z|Z|A|W|
+-----+-----+

```

Флаг W должен устанавливаться в тех случаях, когда команда ATA требует записи в устройство. Бит A должен устанавливаться в тех случаях, когда запрос на запись будет выполняться в асинхронном режиме. Значение бита D определяется регистром ATA Device/Head и принимается во внимание только при установленном бите E. Установленный флаг E говорит о том, что команда относится к числу расширенных команд LBA48. Флаги Z являются резервными и должны иметь значение 0.

Флаг W вместе с полем Sector Count определяют направление переноса данных — на устройство или с устройства:

при записи данных на устройство бит W должен быть установлен, а поле Data должно содержать Sector Count * 512 байтов данных;

при считывании данных с устройства бит W должен быть сброшен (0), а поле Sector Count должно указывать количество секторов, считываемых с устройства; при успешном выполнении команды поле Data в отклике будет содержать считанные с устройства данные;

если переноса данных не выполняется, поле Sector Count должно иметь нулевое значение, а бит W игнорируется.

Если установлены оба флага A и W, сервер может кэшировать запросы на запись в своей памяти и отвечать на запрос незамедлительно, возвращая значение поля Arg без изменений. Сервер может ввести аппаратный запрос на запись в любое удобное время, если обеспечивается возврат кэшированных данных при запросе на считывание этого сектора. Однако при возникновении на сервере проблем с питанием кэшированные данные могут быть потеряны. При завершении исполнения кэшированного запроса на запись никакого сообщения не передается, если в процессе записи не возникло ошибок.

Если установлен флаг E, сервер будет переносить содержимое каждого поля в регистры устройства ATA, как показано ниже:

```

Device      <- (AFlags & 0x50 | 0xA0)
LBA Low     <- lba3
LBA Low     <- lba0
LBA Mid     <- lba4
LBA Mid     <- lba1
LBA High    <- lba5
LBA High    <- lba2
Sector Count <- 0
Sector Count <- Sector Count
Err/Feature <- Err/Feature
Cmd/Status  <- Cmd/Status

```

Биты E,D поля AFlags соответствуют битам L,D в регистрах устройства. По отношению к ним используется операция ИЛИ 0xA0 для установки утративших значение битов в 1 в соответствии со спецификацией ATA.

Если флаг E не установлен, сервер будет переносить содержимое каждого поля в регистры устройства ATA, как показано ниже:

```

Device      <- lba3
LBA Low     <- lba0
LBA Mid     <- lba1
LBA High    <- lba2
Sector Count <- Sector Count
Err/Feature <- Err/Feature
Cmd/Status  <- Cmd/Status

```

За исключением случаев асинхронной записи отклик не генерируется, пока не будет завершено выполнение команды ATA (успешно или с ошибкой). По завершении обработки команды ATA регистры устройства ATA копируются в поля сообщения, как показано ниже:

```

Err/Feature <- Err/Feature
Sector Count <- Sector Count
lba0        <- LBA Low
lba1        <- LBA Mid
lba2        <- LBA High
Cmd/Status  <- Cmd/Status

```

Порядок копирования регистров в поля сообщения сохраняется независимо от состояния бита E.

При установленном бите E дополнительно будут копироваться перечисленные ниже значения регистров устройства ATA в поля lba3, lba4 и lba5 после установки флага в регистре ATA Device Control.

```

lba3        <- LBA Low
lba4        <- LBA Mid
lba5        <- LBA High

```

Во всех остальных полях отклика сохраняются значения из соответствующих полей сообщения с командой.

3.2. Command 1, Query Config Information

```

      0           1           2           3
0  0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
24 |           Buffer Count           |           Firmware Version           |
+-----+-----+-----+-----+-----+-----+-----+-----+
28 | Sector Count | AoE | CCmd |           Config String Length           |
+-----+-----+-----+-----+-----+-----+-----+-----+
32 |           Config String           |
+-----+-----+-----+-----+-----+-----+

```

Command 1 позволяет получить от сервера конфигурационную информацию, а в некоторых случаях - поменять параметры конфигурации. Формат поля Arg в команде и отклике показан на рисунке.

Buffer Count

Максимальное число сообщений, которые могут сохраняться в очереди обработки сервера. При достижении порога новые сообщения будут отбрасываться.

Firmware Version

Номер версии ПО сервера (firmware).

Sector Count

Отличное от нуля значения счетчика задает число секторов, обрабатываемых по одной команде. Нулевое значение в целях совместимости со старыми устройствами трактуется, как 2.

AoE

Версия протокола AoE, поддерживаемая сервером.

CCmd

Субкоманда запроса/установки строки конфигурации (Config string).

Config String Length

Размер строки конфигурации.

Config String

Строка конфигурации размером до 1024 байтов, включительно.

В сообщениях с командами для полей Buffer Count, Firmware Version и AoE клиент должен устанавливать значение 0, а сервер должен игнорировать эти поля. Остальные поля могут использоваться для считывания или установки строки конфигурации.

Допустимые значения CCmd перечислены ниже.

CCmd 0: read config string - считать строку конфигурации

Считать конфигурационную строку сервера без какой-либо проверки или отклика.

CCmd 1: test config string - проверить строку конфигурации

Отклик выдается только при точном совпадении строки в аргументе со строкой конфигурации сервера.

CCmd 2: test config string prefix - проверить префикс строки конфигурации

Отклик выдается только при точном совпадении префикса в аргументе со строкой конфигурации сервера.

CCmd 3: set config string - установить строку конфигурации

Если строка конфигурации сервера пуста, она устанавливается в соответствии со строкой в аргументе и возвращается в отклике. Если строка конфигурации сервера не пуста, возвращается отклик с установленным битом E в поле E и кодом ошибки 4 в поле Error.

CCmd 4: force set config string - принудительно установить строку конфигурации

Строка конфигурации сервера устанавливается в соответствии со строкой аргумента и возвращается в отклике.

В откликах сервер должен указывать свои текущие значения для всех полей.

Когда сервер завершил загрузку и готов к обработке команд, ему следует передать в широковещательном режиме отклик Query Config Information с нулевым значением тега.

3.3. Command 2, Mac Mask List

Команда Mac Mask List используется для считывания списка масок MAC для цели AoE и управления этим списком. Цель AoE может использовать список масок для контроля доступа. Предполагается, что реализации «целей» будут вести себя, как описано ниже.

Цель AoE будет воспринимать полученный кадр с командой при выполнении любого из приведенных ниже условий:

1. список масок пуст;
2. адрес отправителя кадра имеется в списке масок.

Если ни одно из условий не выполняется, цель должна игнорировать команду. Такая проверка должна выполняться для всех команд AoE, включая команды Mac Mask List.

Формат поля Arg показан на рисунке.

	0									1									2									3												
	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
24	Reserved									MCmd									MError									Dir Count												
28	Directive 0																																							
32	Directive 0																																							

MCmd

Субкоманда Mac Mask List. MCmd может принимать значения:

MCmd 0: прочесть список;

MCmd 1: редактировать список.

MError

При возникновении ошибки в процессе обработки списка для поля MError устанавливается одно из значений:

MError 1: Unspecified Error - неизвестная ошибка;

MError 2: Bad DCmd directive - некорректная директива DCmd;

MError 3: Mask list full - список масок полон.

Dir Count

Счетчик числа директив, содержащихся в сообщении.

	0								1								2								3															
	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
0	Reserved								DCmd								Ethernet Address																							
4	Ethernet Address																																							

Directive

Директива для маски MAC. Каждая директива занимает 8 байтов и включает перечисленные ниже поля.

DCmd

Код директивы, который может принимать одно из значений:

- 0 - нет директивы;
- 1 - добавить MAC-адрес в список масок;
- 2 - удалить MAC-адрес из списка масок.

Ethernet Address

Адрес Ethernet, включаемый в список масок или удаляемый из списка.

В командах считывания списка масок должно быть установлено MCmd=0. Отклик будет возвращать MAC-адреса в директивах, следующих после заголовка. Число директив указывается в поле Dir Count.

В командах редактирования списка масок должно задаваться значение MCmd=1. Кадр с командой должен содержать число директив, указанное в поле Dir Count. Сервер, обрабатывающий такую команду, должен брать директивы в соответствии с порядком их размещения в кадре. При возникновении ошибки в процессе обработки директив сервер должен прервать обработку и вернуть отклик со значением поля Dir Count, указывающим на связанную с ошибкой директиву, и кодом ошибки в поле MError. Отклик сервера на ошибку должен содержать список масок из соответствующей команды.

Если при обработке списка не возникло ошибок, сервер должен вернуть в отклике список масок, включая значение Dir Count для счетчика директив. Попытка включения дубликата адреса Ethernet в список масок не считается ошибкой. Точно так же не относится к числу ошибок и попытка удаления из списка отсутствующего в нем адреса Ethernet. Сервер должен игнорировать такие попытки.

Здесь не предполагается считывания списка масок, содержащего более 255 элементов. Серверам, поддерживающим эту команду, рекомендуется ограничивать размер списка масок значением 255. Однако сервер не обязан поддерживать 255 масок в своем списке.

Следует отметить, что при добавлении клиентом в пустой список масок того или иного набора адресов, не включающего адрес самого клиента, этот клиент утратит дальнейшую возможность обмена данными с «целью». Рекомендуется при любом изменении списка адресов включать Ethernet-адрес клиента в качестве первой директивы.

3.4. Command 3, Reserve/Release

Команда Reserve/Release служит для управления блокировкой доступа к цели AoE по списку MAC-адресов. Забронированная цель AoE будет выполнять команды ATA на чтение/запись только в тех случаях, когда MAC-адрес отправителя из кадра с командой будет присутствовать в списке бронирования. Остальные команды (запрос конфигурации, идентификация ATA и т.п.) обрабатываются независимо от адреса отправителя. Когда цель недоступна в результате бронирования, возвращается сообщение об ошибке AoE с кодом 6: Target is reserved.

Формат поля Arg показан на рисунке.

	0								1								2								3															
	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
24	RCmd								NMacs								Ethernet Address 0																							
28	Ethernet Address 0																																							
32	Ethernet Address 1																																							
36	Ethernet Address 1								...																															

RCmd

Код команды бронирования/освобождения, который может принимать одно из следующих значений:

- 0 - прочесть список бронирования;
- 1 - установить список бронирования;
- 2 - принудительно выполнить команду бронирования/освобождения.

По команде RCmd 0 возвращается список бронирования.

RCmd 1 используется для бронирования или освобождения цели. Бронирование осуществляется путем создания списка с одним или множеством MAC-адресов. Освобождение выполняется путем сброса списка адресов (организации пустого списка) за счет установки NMacs=0. Команда RCmd 1 будет обрабатываться только в тех случаях, когда текущий список бронирования пуст или включает MAC-адрес отправителя команды. Если команда RCmd 1 не может быть выполнена в результате бронирования, должно возвращаться сообщение об ошибке AoE с кодом 6, как сказано выше.

Команда RCmd 2 используется для принудительного бронирования/освобождения цели, независимо от наличия MAC-адреса отправителя в списке бронирования. Эта команда используется прежде всего для изменения устаревших списков бронирования. Применять эту команду следует лишь в тех случаях, когда имеется достоверная информация о некорректности списка бронирования.

NMacs

Число MAC-адресов в сообщении.

Ethernet Address [0..n]

Один или множество 6-байтовых адресов Ethernet.

В откликах о выполнении или ошибке должен возвращаться текущий список бронирования для сведения клиента.

Адреса авторов

The Brantley Coile Company, Inc.

565 Research Drive

Athens, GA 30605

{sah,brantley}@borf.com

Перевод на русский язык

Николай Малых

nmalykh@protokols.ru

Благодарности

Большое спасибо Norman Wilson за критику и предложения.