

## Разъяснение вопроса о регистре символов в DNS

### Domain Name System (DNS) Case Insensitivity Clarification

#### Статус документа

В этом документе описан предлагаемый стандарт протокола для сообщества Internet; документ служит приглашением к дискуссии в целях развития протокола. Информацию о текущем состоянии стандартизации протокола можно найти в документе "Internet Official Protocol Standards" (STD 1). Данный документ может распространяться свободно.

#### Авторские права

Copyright (C) The Internet Society (2006).

#### Аннотация

Система доменных имён DNS<sup>1</sup> не принимает во внимание регистр символов. В этом документе даны разъяснения этого вопроса и приведена чёткая спецификация правил. Эти разъяснения обновляют RFC 1034, 1035 и 2181.

## Оглавление

1. Введение.....	1
2. Независимость от регистра меток DNS.....	1
2.1. Представление «необычных» октетов в метках DNS.....	2
2.2. Примеры меток с Escape-символами.....	2
3. Просмотр имён, типы меток и CLASS.....	2
3.1. Исходные типы меток DNS.....	2
3.2. Независимость от регистра меток расширенного типа.....	2
3.3. Независимость от регистра значений CLASS.....	2
4. Регистр символов на вводе и выводе.....	3
4.1. Сохранение регистра на выходе DNS.....	3
4.2. Сохранение регистра на вводе DNS.....	3
5. Доменные имена на других языках.....	3
6. Вопросы безопасности.....	3
7. Благодарности.....	4
Нормативные документы.....	4
Дополнительная литература.....	4

## 1. Введение

Система доменных имён (DNS) представляет собой глобальную иерархическую систему распределённых баз данных для адресов Internet, почтовых проху-серверов и другой информации. Каждый узел в дереве DNS имеет имя, которое может включать метки [STD13, RFC1591, RFC2606], трактуемые независимо от регистра символов в них. В этом документе разъясняются вопросы независимости от регистра и вносятся обновления в RFC 1034, RFC 1035 [STD13] и [RFC2181].

Ключевые слова **необходимо** (MUST), **недопустимо** (MUST NOT), **требуется** (REQUIRED), **нужно** (SHALL), **не следует** (SHALL NOT), **следует** (SHOULD), **не нужно** (SHOULD NOT), **рекомендуется** (RECOMMENDED), **возможно** (MAY), **необязательно** (OPTIONAL) в данном документе интерпретируются в соответствии с [RFC2119].

## 2. Независимость от регистра меток DNS

Спецификация DNS была разработана в «эпоху [ASCII]». Предполагалось, что имена DNS будут выглядеть подобно именам хостов, правой половине почтовых адресов Internet (справа от символа @) или выражаться числами, как в пространстве имен in-addr.arpa. Например,

```
foo.example.net.  
aol.com.  
www.gnu.ai.mit.edu.
```

или

```
69.2.0.192.in-addr.arpa.
```

Чувствительные к регистру варианты [RFC3092] приведённых выше имён DNS будут иметь вид

```
Foo.Example.net.  
AOL.COM.  
WWW.gnu.AI.mit.EDU.
```

или

<sup>1</sup>Domain Name System.

Однако отдельные октеты имен DNS включают символы, не входящие в набор ASCII, коды которых являются 8-битовыми и могут принимать любые значения. Тем не менее, многие приложения интерпретируют такие символы, как ASCII.

## 2.1. Представление «необычных» октетов в метках DNS

В первичных файлах [STD13] и другом контексте ASCII, чтение и запись в котором рассчитаны на участие человека, требуется использование специального escape-символа для точки (0x2E, «.») и всех символов, которые выходят за пределы диапазона 0x21 (!) - 0x7E (~). Иными словами, специальное представление требуется для символа с кодом 0x2E и всех символов с кодами из диапазонов 0x00 - 0x20 и 0x7F - 0xFF.

Одним из соглашений для октетов, которые не соответствуют печатным символам ASCII, является использование символа обратной дробной черты (\), за которым следует октет целого числа без знака в форме трех десятичных цифр.

Такое же соглашение может использоваться для печатных символов ASCII, включая и сам символ \, который будет представляться в форме \092 или \\, а также символа точки (.), который будет представляться в форме \046 или \. Желательно избегать использования символа \ для кватирования сразу после кода непечатного символа ASCII во избежание сложностей при реализации.

Последовательности из символа \, за которым следует только одна или две десятичных цифры, не определены. Символ обратной дробной черты, за которым следует 4 десятичных цифры, порождает два октета, из которых первый образуется на основе значений трех первых цифр, трактуемых, как десятичное значение, а второй октет будет кодом символа для четвертой десятичной цифры.

## 2.2. Примеры меток с Escape-символами

Первый из приведенных ниже примеров включает метку с символами пробела и точкой, а во втором дана 5-октетная метка, в которой все биты второго октета равны 0, а все биты четвертого - 1.

```
Donald\032E\.\032Eastlake\0323rd.example.
```

и

```
a\000\\\255z.example.
```

## 3. Просмотр имён, типы меток и CLASS

В соответствии с изначальным устройством DNS сравнение при поиске имен в запросах DNS следует выполнять без учета регистра символов [STD13]. Т. е., поиск строки, символы которой представляются октетами из диапазона 0x41 — 0x5A (прописные буквы набора ASCII) **должен** давать такой же результат, как поиск для строки, содержащей символы с кодами из диапазона 0x61 - 0x7A (строчные буквы набора ASCII). Поиск строки с символом нижнего регистра ASCII **должен** давать такой же результат, как при поиске строки с соответствующим символом верхнего регистра ASCII<sup>1</sup>.

Одним из способов реализации этого правила является вычитание значения 0x20 из всех октетов с кодами из диапазона 0x61 - 0x7A перед операцией сравнения. Такую операцию обычно называют приведением регистра (case folding), но реализация такого приведения не требовалась спецификацией. Отметим, что независимость от регистра в DNS **не** соответствует приведению регистра, заданному в стандартах [ISO-8859-1] и [ISO-8859-2]. Например, октеты 0xDD (\221) и 0xFD (\253) **не** соответствуют друг другу, хотя в другом контексте, они могут интерпретироваться, как буква «У» в верхнем и нижнем регистре.

### 3.1. Исходные типы меток DNS

Метки DNS при кодировании для передачи, имеют связанный с ними тип. Исходная спецификация DNS [STD13] включала лишь два типа - метки ASCII размером от 0<sup>2</sup> до 63 октетов и косвенные (или сжатые) метки, которые состоят из указателя смещения по отношению к имени, хранящемуся где-то в сообщении DNS представленном для передачи. Метки ASCII следуют соглашению о преобразовании регистра, описанному здесь, и, как отмечено выше, могут на практике содержать произвольные значения байтов. Косвенные метки, по сути, заменяются именами, на которые они указывают, и после этого трактуются в соответствии с описанными здесь правилами независимости от регистра.

### 3.2. Независимость от регистра меток расширенного типа

В [RFC2671] было введено расширение DNS в части поддержки дополнительных типов меток<sup>3</sup>.

Соглашение о независимости от регистра меток ASCII применимо только к ASCII-меткам, т. е., меткам типа 0x0 независимо от того, указаны они явно или косвенно.

### 3.3. Независимость от регистра значений CLASS

Как описано в [STD13] и [RFC2929], DNS использует для также классы (CLASS). Единственным повсеместно применяемым классом является IN (Internet).

Обработка регистра меток DNS не зависит от класса. В исходной спецификации DNS предполагается, что рекурсивные преобразователи DNS будут способны обслуживать новые классы, которые не были определены на момент разработки спецификации. Это требует односторонней обработки независимости меток от регистра. Если потребуются, например, создание класса с чувствительными к регистру метками ASCII, для таких меток нужно будет выделить новый тип.

<sup>1</sup>Термины «верхний регистр» и «нижний регистр» были введены после появления перемещаемых кареток. Изначально эти термины относились к двум наборам шрифтов для строчных и прописных наборных литер, хранящихся в разных местах. До появления перемещаемых кареток использовались термины «прописная буква» и «строчная буква».

<sup>2</sup>Метка ASCII нулевого размера зарезервирована для использования в качестве имени корня дерева доменных имен.

<sup>3</sup>Единственным типом, определенным в соответствии с этим расширением является тип BINARY [RFC2673], имеющий статус экспериментального [RFC3363].

## 4. Регистр символов на вводе и выводе

Хотя сравнение меток ASCII не зависит от регистра символов, в [STD13] сказано, что регистр символов меток **должен** сохраняться при выводе и, по возможности, при вводе. Однако это требование не столь жестко, как может показаться, поскольку сохранение регистра при выводе **не** требуется, когда выходная информация оптимизируется за счет использования косвенных меток, как описано ниже.

### 4.1. Сохранение регистра на выходе DNS

[STD13] рассматривает пространство имен DNS, как дерево узлов. Вывод ASCII образуется путем принятия метки узла, чье имя выдается на выход, типографического представления строки ASCII с перемещением вверх по всему дереву и добавлением каждой метки в начало строки (между метками добавляется точка). При кодировании на уровне сигналов происходят те же действия с кодированием каждой метки, но без вставки точек. В этих операциях не выполняется преобразования регистра символов, т. е. регистр «сохраняется». Однако для оптимизации вывода могут применяться косвенные метки, указывающие на имена в ответе DNS. При определении имени, на которое указывает такая метка (например, QNAME), выполняется независимое от регистра сравнение, описанное выше. Таким образом, эта оптимизация может легко изменить регистр символов на выходе. Этот тип оптимизации обычно называют «сжатием имени».

### 4.2. Сохранение регистра на вводе DNS

Исходные данные DNS берутся из первичного файла ASCII, как определено в [STD13], или получаются в результате переноса зоны. Динамическое обновление DNS и инкрементальный перенос зон [RFC1995] были позднее добавлены [RFC2136, RFC3007] в качестве дополнительных источников данных DNS. При создании дерева имен DNS любым из этих способов преобразования регистра символов не происходит. Таким образом, все метки ASCII сохраняют регистр символов, если эти метки созданы для узлов. Однако если имя метки является входной информацией для узла, который уже существует в данных DNS, ситуация усложняется. Реализация может сохранить регистр символов первоначально загруженной метки, чтобы позволить при новом вводе изменить регистр, или поддерживать отдельные копии с сохранением исходного регистра.

Например, если загружаются данные с именем владельца foo.bar.example [RFC3092], а позднее вводятся данные с именем владельца хуз.BAR.example, имя метки узла bar.example (т. е., bar) может быть заменено на BAR в сохраняемых в DNS данных. Таким образом, при последующем поиске данных, сохраненных для хуз.bar.example, во всех возвращаемых данных может указываться метка хуз.BAR.example или хуз.bar.example, а также при возврате более одной записи RR могут указываться оба варианта вперемешку. Последний вариант маловероятен, поскольку оптимизация размера отклика за счет использования косвенных меток приводит к использованию одного варианта «хвоста» метки (bar.example или BAR.example) для всех возвращаемых RR. Отметим, что это не оказывает никакого влияния на число и полноту возвращаемого набора RR и может воздействовать лишь на регистр символов в именах.

Такие же соображения применимы и для случаев ввода множества записей, в которых имена владельцев отличаются лишь регистром символов. Например, если одна запись A является первой записью, сохраненной с именем владельца хуз.BAR.example, а вторая запись A сохраняется с именем владельца XYZ.BAR.example, вторая запись **может** быть сохранена с первым (строчные буквы) именем, **может** переписать первую с использованием прописных (заглавных) букв, а **могут** быть сохранены оба варианта в данных DNS. В любом случае при поиске данных будут возвращаться все RR, независимо от регистра символов в запросе.

Отметим, что порядок вставки в базу данных сервера имен узлов дерева DNS из первичного файла (Master File) не определен, поэтому несогласованное использование регистра символов в Master File ведет к непредсказуемому регистру символов на выходе.

## 5. Доменные имена на других языках

Схема была адаптирована для доменных имен на других языках (internationalized domain name и internationalized label), как описано в [RFC3490, RFC3454, RFC3491, RFC3492]. Это делает доступной большую часть символов [UNICODE] с помощью отдельного преобразования из доменных имен на национальных языках в имена DNS и обратно. Чувствительность к регистру символов в именах доменов на национальных языках зависит от используемого преобразования и обрабатывается в соответствии с [RFC3454] и [RFC3491]. Это не относится к протоколам DNS, стандартизованным в STD 13.

## 6. Вопросы безопасности

Эквивалентность некоторых типов меток DNS, различающихся регистром символов, как разъяснено в этом документе, может приводить к некоторым проблемам безопасности. Например, пользователей может смущать тот факт, что имена, различающиеся регистром символов, на самом деле окажутся тождественными.

Кроме того, доменные имена могут использоваться за пределами контекста DNS. Они могут применяться в качестве чувствительных к регистру символов индексов в некоторых базах данных или файловых системах. Возможна также интерпретация имен, как двоичных данных, в некоторых системах защиты целостности или аутентификации. Эти проблемы обычно решаются за счет использования стандартизованной или «канонической» формы меток DNS типа ASCII (т. е., приведения значений всех октетов символов ASCII в метках типа ASCII к верхнему или нижнему регистру). Пример канонической формы доменных имен (и канонического порядка их сортировки) приведен в разделе 6 [RFC4034]. См. также документ [RFC3597].

И, наконец, в DNS могут храниться имена, отличные от DNS, с ошибочным предположением о сохранении регистра символов в них. Например, хотя и достаточно редко, в системах с регистрочувствительной локальной (слева от @) частью адресов электронной почты попытка сохранить две записи RP<sup>1</sup> [RFC1183], различающиеся только регистром символов, может приводить к неожиданным результатам, способным оказать влияние на безопасность. Это обусловлено тем, что весь адрес электронной почты, включая регистрочувствительную локальную часть, преобразуется в имя DNS с возможным изменением регистра некоторых символов, как было описано выше.

<sup>1</sup>Responsible Person — ответственное лицо

## 7. Благодарности

Благодарим Rob Austein, Olafur Gudmundsson, Daniel J. Anderson, Alan Barrett, Marc Blanchet, Dana, Andreas Gustafsson, Andrew Main, Thomas Narten и Scott Seligman за их вклад в подготовку документа.

## Нормативные документы

- [ASCII] ANSI, "USA Standard Code for Information Interchange", X3.4, American National Standards Institute: New York, 1968.
- [RFC1995] Ohta, M., "Incremental Zone Transfer in DNS", [RFC 1995](#), August 1996.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, [RFC 2119](#), March 1997.
- [RFC2136] Vixie, P., Thomson, S., Rekhter, Y., and J. Bound, "Dynamic Updates in the Domain Name System (DNS UPDATE)", [RFC 2136](#), April 1997.
- [RFC2181] Elz, R. and R. Bush, "Clarifications to the DNS Specification", [RFC 2181](#), July 1997.
- [RFC3007] Wellington, B., "Secure Domain Name System (DNS) Dynamic Update", [RFC 3007](#), November 2000.
- [RFC3597] Gustafsson, A., "Handling of Unknown DNS Resource Record (RR) Types", RFC 3597, September 2003.
- [RFC4034] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, "Resource Records for the DNS Security Extensions", [RFC 4034](#), March 2005.
- [STD13] Mockapetris, P., "Domain names - concepts and facilities", STD 13, [RFC 1034](#), November 1987.  
Mockapetris, P., "Domain names - implementation and specification", STD 13, [RFC 1035](#), November 1987.

## Дополнительная литература

- [ISO-8859-1] International Standards Organization, Standard for Character Encodings, Latin-1.
- [ISO-8859-2] International Standards Organization, Standard for Character Encodings, Latin-2.
- [RFC1183] Everhart, C., Mamakos, L., Ullmann, R., and P. Mockapetris, "New DNS RR Definitions", [RFC 1183](#), October 1990.
- [RFC1591] Postel, J., "Domain Name System Structure and Delegation", RFC 1591, March 1994.
- [RFC2606] Eastlake 3rd, D. and A. Panitz, "Reserved Top Level DNS Names", BCP 32, RFC 2606, June 1999.
- [RFC2929] Eastlake 3rd, D., Brunner-Williams, E., and B. Manning, "Domain Name System (DNS) IANA Considerations", BCP 42, RFC 2929, September 2000.
- [RFC2671] Vixie, P., "Extension Mechanisms for DNS (EDNS0)", [RFC 2671](#), August 1999.
- [RFC2673] Crawford, M., "Binary Labels in the Domain Name System", [RFC 2673](#), August 1999.
- [RFC3092] Eastlake 3rd, D., Manros, C., and E. Raymond, "Etymology of 'Foo'", RFC 3092, 1 April 2001.
- [RFC3363] Bush, R., Durand, A., Fink, B., Gudmundsson, O., and T. Hain, "Representing Internet Protocol version 6 (IPv6) Addresses in the Domain Name System (DNS)", RFC 3363, August 2002.
- [RFC3454] Hoffman, P. and M. Blanchet, "Preparation of Internationalized Strings ("stringprep")", RFC 3454, December 2002.
- [RFC3490] Faltstrom, P., Hoffman, P., and A. Costello, "Internationalizing Domain Names in Applications (IDNA)", RFC 3490, March 2003.
- [RFC3491] Hoffman, P. and M. Blanchet, "Nameprep: A Stringprep Profile for Internationalized Domain Names (IDN)", RFC 3491, March 2003.
- [RFC3492] Costello, A., "Punycode: A Bootstring encoding of Unicode for Internationalized Domain Names in Applications (IDNA)", RFC 3492, March 2003.
- [UNICODE] The Unicode Consortium, "The Unicode Standard", <<http://www.unicode.org/unicode/standard/standard.html>>.

### Адрес автора

**Donald E. Eastlake 3rd**  
Motorola Laboratories  
155 Beaver Street  
Milford, MA 01757 USA  
Phone: +1 508-786-7554 (w)  
E-Mail: [Donald.Eastlake@motorola.com](mailto:Donald.Eastlake@motorola.com)

### Перевод на русский язык

**Николай Малых**

[nmalykh@protokols.ru](mailto:nmalykh@protokols.ru)

### Полное заявление авторских прав

Copyright (C) The Internet Society (2006).

К этому документу применимы права, лицензии и ограничения, указанные в BCP 78, и, за исключением указанного там, авторы сохраняют свои права.

Этот документ и содержащаяся в нем информация представлены "как есть" и автор, организация, которую он/она представляет или которая выступает спонсором (если таковой имеется), Internet Society и IETF отказываются от каких-либо гарантий (явных или подразумеваемых), включая (но не ограничиваясь) любые гарантии того, что использование представленной здесь информации не будет нарушать чьих-либо прав, и любые предполагаемые гарантии коммерческого использования или применимости для тех или иных задач.

### **Интеллектуальная собственность**

IETF не принимает какой-либо позиции в отношении действительности или объема каких-либо прав интеллектуальной собственности (Intellectual Property Rights или IPR) или иных прав, которые, как может быть заявлено, относятся к реализации или использованию описанной в этом документе технологии, или степени, в которой любая лицензия, по которой права могут или не могут быть доступны, не заявляется также применение каких-либо усилий для определения таких прав. Сведения о процедурах IETF в отношении прав в документах RFC можно найти в BCP 78 и BCP 79.

Копии раскрытия IPR, предоставленные секретариату IETF, и любые гарантии доступности лицензий, а также результаты попыток получить общую лицензию или право на использование таких прав собственности разработчиками или пользователями этой спецификации, можно получить из сетевого репозитория IETF IPR по ссылке <http://www.ietf.org/ipr>.

IETF предлагает любой заинтересованной стороне обратить внимание на авторские права, патенты или использование патентов, а также иные права собственности, которые могут потребоваться для реализации этого стандарта. Информацию следует направлять в IETF по адресу [ietf-ipr@ietf.org](mailto:ietf-ipr@ietf.org).

### **Подтверждение**

Финансирование функций RFC Editor обеспечено IETF Administrative Support Activity (IASA).