

# Единицы измерения количества информации

Материал из Википедии — свободной энциклопедии

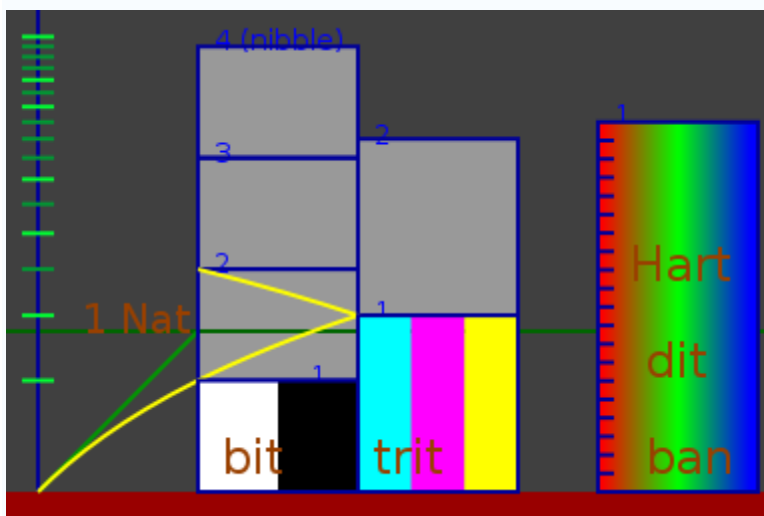
**Единицы измерения информации** служат для измерения объёма **информации** — величины, исчисляемой логарифмически.<sup>[1]</sup> Это означает, что когда несколько объектов рассматриваются как один, количество возможных состояний **перемножается**, а количество информации — **складывается**. Не важно, идёт речь о **случайных величинах** в математике, **регистрах** цифровой памяти в технике или в **квантовых системах** в физике.

Чаще всего измерение информации касается объёма **компьютерной памяти** и объёма данных, передаваемых по **цифровым каналам связи**.

## Содержание

- [1 Первичные единицы](#)
- [2 Единицы, производные от бита](#)
  - [2.1 Байт](#)
  - [2.2 Килобайт](#)
  - [2.3 Мегабайт](#)
  - [2.4 Гигабайт](#)
- [3 Что такое «байт»?](#)
- [4 Чему равно «кило»?](#)
- [5 Примечания](#)
- [6 См. также](#)
- [7 Ссылки](#)

## Первичные единицы



Сравнение разных единиц измерения информации. Дискретные величины представлены прямоугольниками, единица «нат» — горизонтальным уровнем. Риски слева — логарифмы натуральных чисел.

Объёмы информации можно представлять как [логарифм](#)<sup>[2]</sup> количества состояний.

Наименьшее целое число, логарифм которого положителен — [2](#). Соответствующая ему единица — [бит](#) — является основой исчисления информации в цифровой технике.

Единица, соответствующая числу 3 ([трит](#)) равна  $\log_2 3 \approx 1.585$  бита, числу 10 (*хартли*) —  $\log_2 10 \approx 3.322$  бита.

Такая единица как [нат](#) (nat, е-бит), соответствующая [натуральному логарифму](#) применяется в вычислительной технике в инженерных и научных расчётах. [Основание натуральных логарифмов](#) не является [целым числом](#).

## Единицы, производные от бита

Целые количества бит отвечают количеству состояний, равному степеням двойки.

Особое название имеет 4 бита — [ниббл](#) (полу[байт](#), [тетрада](#), четыре двоичных разряда), которые вмещают в себя количество информации, содержащейся в одной [шестнадцатеричной](#) цифре.

### [Байт](#)

<a href="#">Измерения в байтах</a>						
<a href="#">Десятичная приставка</a>			<a href="#">Двоичная приставка</a>			
<a href="#">Название</a>	<a href="#">Символ</a>	<a href="#">Степень</a>	<a href="#">Название</a>	<a href="#">Символ</a>	<a href="#">Степень</a>	
				<a href="#">МЭК</a>	ГОСТ	
<a href="#">байт</a>	В	$10^0$	<a href="#">байт</a>	В	байт	$2^0$
<a href="#">килобайт</a>	кВ	$10^3$	<a href="#">кибибайт</a>	КиВ	Кбайт	$2^{10}$
<a href="#">мегабайт</a>	МВ	$10^6$	<a href="#">мебибайт</a>	МиВ	Мбайт	$2^{20}$
<a href="#">гигабайт</a>	ГВ	$10^9$	<a href="#">гибибайт</a>	ГиВ	Гбайт	$2^{30}$
<a href="#">терабайт</a>	ТВ	$10^{12}$	<a href="#">тебибайт</a>	ТиВ	Тбайт	$2^{40}$
<a href="#">петабайт</a>	ПВ	$10^{15}$	<a href="#">пебибайт</a>	ПиВ	Пбайт	$2^{50}$
<a href="#">эксабайт</a>	ЕВ	$10^{18}$	<a href="#">эксбибайт</a>	ЕиВ	Эбайт	$2^{60}$
<a href="#">зеттабайт</a>	ЗВ	$10^{21}$	<a href="#">зебибайт</a>	ЗиВ	Збайт	$2^{70}$
<a href="#">йоттабайт</a>	УВ	$10^{24}$	<a href="#">йобибайт</a>	УиВ	Йбайт	$2^{80}$

Следующей по порядку популярной единицей информации является **8** бит, или *байт* (о терминологических тонкостях [написано ниже](#)). Именно к байту (а не к биту) непосредственно приводятся все большие объёмы информации, исчисляемые в компьютерных технологиях.

Такие величины как [машинное слово](#) и т. п., составляющие несколько байт, в качестве **единиц измерения** почти никогда не используются.

## Килобайт

Для измерения больших количеств байтов служат единицы «килобайт» = **1000** байт и «Кбайт»<sup>[31]</sup> (*кибибайт*, kibibyte) = **1024** байт (о путанице десятичных и двоичных единиц и терминов [см. ниже](#)). Такой порядок величин имеют, например:

- [Сектор диска](#) обычно равен 512 байтам то есть половине Кбайт, хотя для некоторых устройств может быть равен одному или двум Кбайт.
- Классический размер «блока» в [файловых системах UNIX](#) равен одному Кбайт (1024 байт).
- «Страница памяти» в процессорах [x86](#) (начиная с модели [Intel 80386](#)) имеет размер 4096 байт, то есть 4 Кбайт.

Объём информации, получаемой при считывании [дискеты](#) «3,5" высокой плотности» равен 1440 Кбайт (ровно); другие форматы также исчисляются целым числом Кбайт.

## Мегабайт

Единицы «мегабайт» = 1000 килобайт = **1000000** байт и «Мбайт»<sup>[31]</sup> (*мебибайт*, mebibyte) = 1024 Кбайт = 1 048 576 байт применяются для измерения объёмов носителей информации.

Объём адресного пространства процессора [Intel 8086](#) был равен 1 Мбайт.

[Оперативную память](#) и ёмкость [CD-ROM](#) меряют двоичными единицами (мебибайтами, хотя их так обычно не называют), но для объёма НЖМД десятичные мегабайты были более популярны.

Современные жёсткие диски имеют объёмы, выражаемые в этих единицах минимум шестизначными числами, поэтому для них применяются гигабайты.

## Гигабайт

Единицы «гигабайт» = 1000 мегабайт = **1000000000** байт и «Гбайт»<sup>[31]</sup> (*гибибайт*, gibibyte) = 1024 Мбайт =  $2^{30}$  байт измеряют объём больших носителей информации, например [жёстких дисков](#). Разница между двоичной и десятичной единицами уже превышает 7 %.

Размер 32-битного адресного пространства равен 4 Гбайт  $\approx$  4,295 гигабайт. Такой же порядок имеют размер [DVD-ROM](#) и современных носителей на [флеш-памяти](#). Размеры жёстких дисков уже достигают сотен и тысяч гигабайт.

Для исчисления ещё больших объёмов информации имеются единицы [терабайт](#)—*тебибайт* ( $10^{12}$  и  $2^{40}$  соответственно), *петабайт*—*небибайт* ( $10^{15}$  и  $2^{50}$  соответственно) и т. д.

## Что такое «байт»?

В принципе, байт определяется для **конкретного компьютера** как минимальный шаг адресации памяти, который на старых машинах не обязательно был равен 8 битам (а память не обязательно состоит из битов — см., например: троичный компьютер). В современной традиции, байт часто считают **равным восьми битам**.

В таких обозначениях как **Кбайт** (русское) или **KB** (английское) под байт (B) подразумевается именно 8 бит, хотя сам термин «байт» не вполне корректен с точки зрения теории.

Во французском языке используются обозначения **о**, **Ко**, **Мо** и т. д. (от слова octet) дабы подчеркнуть, что речь идёт именно о 8 битах.

## Чему равно «кило»?

Долгое время разнице между множителями 1000 и 1024 старались не придавать большого значения. Во избежание недоразумений следует чётко понимать различие между:

- двоичными кратными единицами, обозначаемыми согласно ГОСТ 8.417-2002 как «Кбайт», «Мбайт», «Гбайт» и т. д. (два в степенях кратных десяти);
- единицами *килобайт*, *мегабайт*, *гигабайт* и т. д., понимаемыми как научные термины (десять в степенях кратных трём).

Последние **по определению** равны соответственно  $10^3$ ,  $10^6$ ,  $10^9$  байт.

В качестве терминов для «Кбайт», «Мбайт», «Гбайт» и т. д. МЭК предлагает «кибибайт», «мебибайт», «гибибайт» и т. д., однако эти термины критикуются за непроизносимость и не встречаются в устной речи.

В различных областях информатики предпочтения в употреблении десятичных и двоичных единиц тоже различны. Причём, хотя со времени стандартизации терминологии и обозначений прошло уже несколько лет, далеко не везде стремятся прояснить точное значение используемых единиц. В английском языке для «киби»=1024 иногда используют прописную букву **К**, дабы подчеркнуть отличие от обозначаемой строчной буквой приставки СИ *кило*. Однако, такое обозначение не опираются на авторитетный стандарт, в отличие от российского ГОСТа касательно «Кбайт».

## Примечания

1. ↑ С точки зрения физики, величина информации (как и близкая к ней по смыслу энтропия) безразмерна. На практике, как и при измерении безразмерных же углов, пользуются различными практически удобными единицами.
2. ↑ «логарифм» на answers.com (англ.) Не важно, какой именно логарифм выбрать, поскольку численные величины логарифмов по разным основаниям пропорциональны. Таким образом, вопрос выбора единицы измерения информации фактически равнозначен выбору основания для логарифма количества состояний. Следует также заметить, что информация случайной величины точно равна логарифму количества состояний лишь при равномерном распределении. Во всех прочих случаях количество информации будет меньше.
3. ↑ <sup>1 2 3</sup> ГОСТ 8.417-2002 «Единицы величин»

## См. также

- [Двоичные приставки](#)
- [Машинное слово](#)
- [Теория информации](#)
- [Трайт](#)

## Ссылки

- [В 2006 году в мире создан 161 миллиард гигабайтов информации](#) «Компьюлента», 9 марта 2007