

### Авторские права

© Postgres Professional, 2017–2024

Авторы: Егор Рогов, Павел Лузанов, Илья Баштанов, Алексей Береснев Фото: Олег Бартунов (монастырь Пху и пик Бхрикути, Непал)

# Использование материалов курса

Некоммерческое использование материалов курса (презентации, демонстрации) разрешается без ограничений. Коммерческое использование возможно только с письменного разрешения компании Postgres Professional. Запрещается внесение изменений в материалы курса.

# Обратная связь

Отзывы, замечания и предложения направляйте по адресу: edu@postgrespro.ru

#### Отказ от ответственности

Компания Postgres Professional не несет никакой ответственности за любые повреждения и убытки, включая потерю дохода, нанесенные прямым или непрямым, специальным или случайным использованием материалов курса. Компания Postgres Professional не предоставляет каких-либо гарантий на материалы курса. Материалы курса предоставляются на основе принципа «как есть» и компания Postgres Professional не обязана предоставлять сопровождение, поддержку, обновления, расширения и изменения.

# Темы занятия



Основные понятия

Установка из исходных кодов

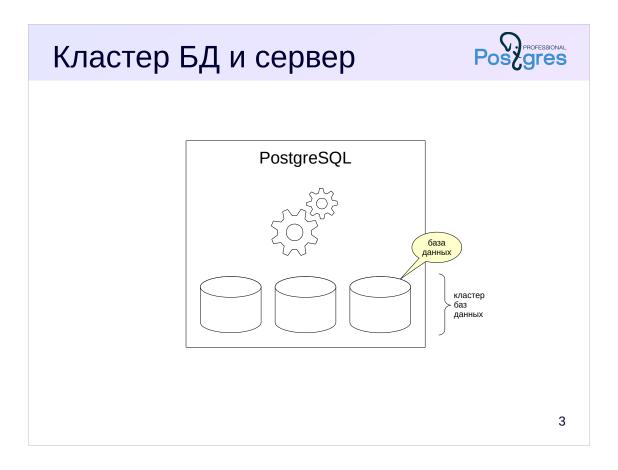
Управление сервером

Пакетная установка

Управление сервером в Ubuntu

Облачные решения

2



Прежде чем говорить об установке, рассмотрим некоторые основные понятия.

PostgreSQL — программа, которая относится к классу систем управления базами данных.

Когда эта программа выполняется, мы называем ее *сервером* PostgreSQL или *экземпляром сервера*. Пока сервер представляется для нас «черным ящиком», но постепенно мы познакомимся с тем, как он устроен.

Данные, которыми управляет PostgreSQL, хранятся в базах данных. Один экземпляр PostgreSQL одновременно работает с несколькими базами данных. Этот набор баз данных называется кластером баз данных. Подробнее мы будем говорить о базах данных в теме «Организация данных. Базы данных и схемы».

Итак: кластер баз данных — это данные в файлах; сервер или экземпляр сервера — программа, управляющая кластером баз данных.

# Исходные коды



### Установка из исходных кодов

стабильная версия сервера можно использовать нестандартные параметры или собрать на специфичной архитектуре

## Установка из репозитория git

текущая версия сервера необходимо в первую очередь для разработчиков ядра; требует более широкого набора инструментов

4

Полезно иметь представление о том, как происходит установка из исходных кодов, чтобы в случае необходимости можно было собрать PostgreSQL с нестандартными параметрами или на специфичной архитектуре.

По адресу <a href="https://www.postgresql.org/ftp/source">https://www.postgresql.org/ftp/source</a> доступен исходный код в двух вариантах: gzip и bzip2. В виртуальной машине необходимый файл уже находится в домашнем каталоге пользователя student.

Исходные коды можно взять и непосредственно из git-репозитория проекта: git://git.postgresql.org/git/postgresql.git (проект имеет зеркала, в том числе на github).

В этом случае можно собрать не только стабильную версию, но и версию на любой конкретный коммит, в том числе наиболее актуальную.

При установке из репозитория git требуется более широкий набор инструментов. Например, лексический и синтаксический анализатор сделан с помощью Flex и Bison, и в git хранятся именно исходные коды для этих инструментов. При сборке из них генерируются файлы на Си, которые затем компилируются. При этом в архиве с исходными кодами находятся уже сгенерированные файлы Си.

# Необходимое ПО



#### Обязательно

tar, gzip/bzip2, GNU make, компилятор Си (С89)

### Используются, но можно отключить

библиотеки GNU Readline, zlib, lz4, zstd

### Дополнительно

языки программирования Perl, Python, Tcl для использования PL/Perl, PL/Python, PL/Tcl

Kerberos, OpenSSL, OpenLDAP, PAM для аутентификации и шифрования

библиотека ICU для кроссплатформенной поддержки UNICODE

Отдельные инструменты при сборке из репозитория git

5

Сборка и установка традиционно основывается на GNU Make и Autoconf:

https://postgrespro.ru/docs/postgresgl/16/install-make

В этом курсе рассматривается традиционный способ сборки.

Начиная с версии 16 можно также использовать систему сборки программного обеспечения Meson:

https://postgrespro.ru/docs/postgresql/16/install-meson

Для сборки требуется ряд программ и утилит.

Библиотека readline дает возможность редактировать командную строку, пользоваться историей команд и автодополнением. В серверной инсталляции может быть и не нужна, если не предполагается запускать psql в консоли.

Библиотеки zlib, lz4 и zstd используются для сжатия.

#### Распаковка архива

```
Архив с исходными кодами находится в домашнем каталоге пользователя student. Распаковываем архив:
```

#### Создание конфигурации

Настройку сборки выполним в каталоге с распакованным исходным кодом, используя скрипт configure.

```
student$ cd /home/student/postgresql-16.0
```

Внимание! Если сборка производится повторно, например с другими параметрами конфигурации, то предварительно нужно убрать файлы, оставшиеся от предыдущей сборки:

```
student$ make distclean
```

В команде configure можно указать различные параметры конфигурации. Например:

- --prefix каталог установки, по умолчанию /usr/local/pgsql;
- --enable-debug для включения отладочной информации.

Полный список параметров есть в документации.

Также принимаются во внимание переменные окружения. Например, СС и CFLAGS настраивают компилятор С.

Скрипт configure по умолчанию подготавливает сборку и установку исполняемых файлов, библиотек и прочих файлов PostgreSQL в каталог /usr/local/pgsql. Обычно каталог установки доступен для записи только суперпользователю ОС.

В целях демонстрации ключом --prefix укажем каталог установки ПО сервера в подкаталог pgsql16 домашнего каталога пользователя student. Кроме того, будем использовать порт 5555 вместо порта 5432 по умолчанию:

```
student$ ./configure --prefix=/home/student/pgsql16 --with-pgport=5555
```

Вывод этой и следующих команд занимает очень много места, мы не будем его показывать.

#### Сборка и установка PostgreSQL

Возможные варианты:

- make сборка только сервера;
- make world сборка сервера, всех расширений и документации.

Собираем сервер. В зависимости от характеристик компьютера команда make может работать достаточно долго (минуты и даже десятки минут).

```
student$ make
```

Теперь устанавливаем сервер.

```
student$ make install
```

Для сборки расширений нужно повторить make и make install в подкаталоге contrib.

```
student$ cd contrib
```

#### student\$ make

student\$ make install

Если бы мы собирали сервер с расширениями и документацией (make world), то можно было бы установить все вместе командой make install-world.

# Создание кластера



#### Утилита

initdb

#### Особенности

новый кластер баз данных не может принадлежать суперпользователю ОС

PGDATA — удобная переменная для каталога кластера файлы конфигурации создаются в каталоге PGDATA



🛕 желательно включить расчет контрольных сумм для страниц с данными

7

После установки сервера необходимо создать кластер баз данных. Для этого предназначена утилита initdb.

В целях безопасности каталог, в котором инициализируется кластер, не может принадлежать суперпользователю ОС. Владельцем кластера обычно делают пользователя postgres.

Владелец кластера может определить переменную окружения PGDATA, указывающую на каталог кластера. Эту переменную используют некоторые утилиты сервера, когда им нужно узнать расположение кластера. К таким утилитам относится initdb, а также основная утилита управления сервером pg ctl, о которой мы будем говорить дальше.

В процессе инициализации кластера утилита initdb создает конфигурационные файлы в этом же каталоге PGDATA. Подробнее о файлах конфигурации мы расскажем в одной из следующих тем.

Утилита initdb имеет множество ключей, влияющих на ее работу. Один из важных — -k, или --data-checksums — включает подсчет контрольных сумм на страницах данных. Проверка контрольной суммы будет выполняться при обращении к любой странице данных в кластере. Это несколько снижает производительность, но позволяет оперативно обнаружить повреждения в данных.

https://postgrespro.ru/docs/postgresgl/16/app-initdb

#### Создание кластера

Теперь необходимо создать каталог для данных. Для промышленного использования владельцем каталога обычно назначают отдельного пользователя ОС. В нашем случае, пусть это будет каталог пользователя student внутри /home/student/pgsql16

student\$ mkdir /home/student/pgsql16/data

Этот каталог часто называют PGDATA по имени переменной окружения, которую удобно использовать при работе с утилитами сервера.

student\$ export PGDATA=/home/student/pgsql16/data

Все утилиты сервера установлены в подкаталоге bin, который стоит добавить в переменную РАТН:

student\$ export PATH=/home/student/pgsql16/bin:\$PATH

Для инициализации кластера баз данных используется утилита initdb.

- В ключе -U можно указать имя суперпользователя СУБД. Если его не задать, то используется имя пользователя ОС, в нашем случае student. Обычно таким пользователем является postgres, так что зададим это имя явно.
- Ключ -k включает подсчет контрольных сумм страниц, что позволяет своевременно обнаруживать повреждения данных.
- Если переменная PGDATA не установлена, то в ключе -D следует указать каталог для данных. В нашем случае этот ключ можно было бы не указывать.

student\$ initdb -U postgres -k -D /home/student/pgsql16/data





Утилита

pg\_ctl

### Основные задачи

запуск, останов и получение статуса сервера обновление параметров конфигурации переключение на реплику

9

К основным операциям управления сервером относятся запуск и останов сервера, получение текущего статуса сервера, обновление конфигурации и некоторые другие.

Для выполнения этих действий предназначена утилита pg\_ctl, идущая в составе PostgreSQL.

Запускать pg\_ctl следует от имени владельца кластера баз данных.

Более подробная информация об управлении сервером для администраторов баз данных:

https://postgrespro.ru/docs/postgresql/16/app-pg-ctl

https://postgrespro.ru/docs/postgresql/16/runtime

#### Управление сервером

Теперь все готово к запуску сервера.

- В ключе -1 укажем файл журнала сообщений сервера.
- Ключ -D опускаем, поскольку задана переменная PGDATA.

```
student$ pg_ctl start -l /home/student/logfile
```

Для проверки работы можно вызвать утилиту psql, которая подключится к серверу и вернет текущее время:

```
now
2024-07-08 14:49:56.128765+03
(1 row)
```

Для останова сервера используется следующая команда:

```
student$ pg_ctl stop
```

# Установка пакета



Готовые пакеты — предпочтительный способ

Linux (ubuntu<sup>®</sup>, Debian, Red Hat, SUSE и другие)

входит в дистрибутив ОС репозиторий (yum, apt) или пакеты RPM, DEB

FreeBSD, OpenBSD

пакеты из Ports and Packages Collection

Mac OS X

Windows

11

Предпочтительным вариантом является использование готовых пакетов, так как в этом случае получается понятная, поддерживаемая и легко обновляемая установка.

Пакеты существуют для большинства широко распространенных систем. В каждой из них могут быть свои особенности, с которыми стоит ознакомиться перед установкой.

В пакетные репозитории некоторых систем уже входит PostgreSQL, но, обычно, не самой последней версии. Актуальная версия может быть взята из репозитория PostgreSQL Global Development Group (PGDG):

https://www.postgresql.org/download/

Мы рассмотрим особенности пакета PostgreSQL для Ubuntu.





#### Утилита

pg\_createcluster \mathbb{m}initdb

#### Особенности

инициализация выполняется при установке пакета и создает кластер «main»

🛕 подсчет контрольных сумм по умолчанию не включен расположение файлов конфигурации и журнала сервера автоматический запуск сервера при старте ОС

12

Для инициализации кластера в Ubuntu создана обертка pg createcluster над утилитой initdb.

Справку по использованию pg createcluster можно получить командой: \$ man pg createcluster

Утилита pg createcluster автоматически запускается при установке пакета и создает кластер баз данных с именем «main».

Стоит обратить внимание, что инициализация кластера проходит с отключенным подсчетом контрольных сумм в страницах данных.

Исполняемые файлы, файлы конфигурации и журнал сервера размещены в соответствии с правилами, принятыми в Ubuntu.

Кроме того, настраивается автоматический запуск и останов сервера PostgreSQL при запуске и останове операционной системы.

Для удаления кластера используется утилита pg dropcluster.

Утилиты pg createcluster и pg dropcluster специфичны для Ubuntu. В других системах нужно явно инициализировать кластер с помощью initdb и задействовать подходящие средства операционной системы.

#### Создание кластера в Ubuntu

```
В виртуальной машине курса PostgreSQL установлен из пакета командой:
student$ sudo apt install -y postgresql-16
Каталог установки PostgreSQL:
student$ sudo ls -l /usr/lib/postgresql/16
total 8
drwxr-xr-x 2 root root 4096 июн 21 16:28 bin
drwxr-xr-x 4 root root 4096 июн 21 16:28 lib
Владелец ПО сервера — пользователь root.
Утилита pg config показывает, с какими параметрами был собран сервер:
student$ sudo /usr/lib/postgresql/16/bin/pg_config --configure
 '--build=x86 64-linux-anu' '--prefix=/usr' '--includedir=${prefix}/include'
'--mandir=${prefix}/share/man' '--infodir=${prefix}/share/info' '--sysconfdir=/etc'
'--localstatedir=/var' '--disable-option-checking' '--disable-silent-rules'
'--libdir=${prefix}/lib/x86_64-linux-gnu' '--runstatedir=/run'
'--disable-maintainer-mode' --disable-dependency-tracking' '--with-tcl' '--with-perl'
'--with-python' '--with-pam' '--with-openssl' '--with-libxml' '--with-libxslt'
'--mandir=/usr/share/postgresql/16/man' '--docdir=/usr/share/doc/postgresql-doc-16'
'--sysconfdir=/etc/postgresql-common' '--datarootdir=/usr/share/'
'--datadir=/usr/share/postgresgl/16' '--bindir=/usr/lib/postgresgl/16/bin'
'--libdir=/usr/lib/x86 64-linux-gnu/' '--libexecdir=/usr/lib/postgresql/'
'--includedir=/usr/include/postgresgl/' '--with-extra-version= (Ubuntu
16.3-1.pgdg22.04+1)' '--enable-nls' '--enable-thread-safety' '--enable-debug'
'--enable-dtrace' '--disable-rpath' '--with-uuid=e2fs' '--with-gnu-ld' '--with-gssapi'
'--with-ldap' '--with-pgport=5432' '--with-system-tzdata=/usr/share/zoneinfo' 'AWK=mawk'
'MKDIR P=/bin/mkdir -p' 'PROVE=/usr/bin/prove' 'PYTHON=/usr/bin/python3' 'TAR=/bin/tar
'XSLTPROC=xsltproc --nonet' 'CFLAGS=-g -02 -flto=auto -ffat-lto-objects -flto=auto
-ffat-lto-objects -fstack-protector-strong -Wformat -Werror=format-security
-fno-omit-frame-pointer' 'LDFLAGS=-Wl,-Bsymbolic-functions -flto=auto -ffat-lto-objects
-flto=auto -Wl,-z,relro -Wl,-z,now' '--enable-tap-tests' '--with-icu' '--with-llvm
'LLVM CONFIG=/usr/bin/llvm-config-15' 'CLANG=/usr/bin/clang-15' '--with-lz4'
'--wiTh-zstd' '--with-systemd' '--with-selinux' 'build alias=x86 64-linux-qnu'
'CPPFLAGS=-Wdate-time -D FORTIFY SOURCE=2' 'CXXFLAGS=-g -02 -flto=auto -ffat-lto-objects
-flto=auto -ffat-lto-objects -fstack-protector-strong -Wformat -Werror=format-security'
Кластер баз данных с именем main уже инициализирован и находится в каталоге /var/lib/postgresql/16/main.
Владельцем каталога является пользователь postgres. Вот содержимое каталога:
student$ sudo ls -l /var/lib/postgresql/16/main
total 92
drwx----- 1 postgres postgres 4096 июн 21 17:32 base
drwx----- 1 postgres postgres 4096 июл 8 14:49 global
drwx----- 2 postgres postgres 4096 июн 21 17:31 pg commit ts
drwx----- 2 postgres postgres 4096 июн 21 17:31 pg_dynshmem
drwx----- 4 postgres postgres 4096 июн 21 17:32 pg logical
drwx----- 4 postgres postgres 4096 июн 21 17:31 pg_multixact
drwx----- 2 postgres postgres 4096 июн 21 17:31 pg notify
drwx----- 2 postgres postgres 4096 июн 21 17:31 pg_replslot
drwx----- 2 postgres postgres 4096 июн 21 17:31 pg serial
drwx----- 2 postgres postgres 4096 июн 21 17:31 pg snapshots
drwx----- 1 postgres postgres 4096 июл 8 14:49 pg stat
drwx----- 2 postgres postgres 4096 июн 21 17:31 pg stat tmp
drwx----- 2 postgres postgres 4096 июн 21 17:31 pg subtrans
drwx----- 2 postgres postgres 4096 июн 21 17:31 pg_tblspc
```

-rw----- 1 postgres postgres 88 июн 21 17:31 postgresql.auto.conf -rw----- 1 postgres postgres 130 июл 8 14:49 postmaster.opts -rw----- 1 postgres postgres 109 июл 8 14:49 postmaster.pid





#### Утилита



### Основные задачи

запуск, останов и получение статуса сервера перечитывание параметров конфигурации переключение на реплику

14

В пакетном дистрибутиве для Ubuntu доступ к утилите pg\_ctl осуществляется не напрямую, а через специальную обертку pg\_ctlcluster. Справку по использованию pg\_ctlcluster можно получить командой:

\$ man pg\_ctlcluster

В других системах утилита pg\_ctl может использоваться непосредственно.

#### Управление сервером в Ubuntu

14:49:59 MSK

При установке из пакета в настройки запуска ОС добавляется запуск PostgreSQL. Поэтому после загрузки операционной системы отдельно стартовать PostgreSQL не нужно.

Можно явным образом управлять сервером с помощью утилиты pg\_ctlcluster, которая запускается либо от имени пользователя ОС postgres, либо через sudo.

Утилита pg\_ctlcluster позволяет управлять несколькими серверами разных версий. В качестве параметров ей передаются:

```
• 16 — номер версии сервера;
   • таіп — название сервера;
   • команда
Остановить сервер:
student$ sudo pg_ctlcluster 16 main stop
student$ pg_lsclusters 16 main
Ver Cluster Port Status Owner
                                  Data directory
16 main 5432 down postgres /var/lib/postgresql/16/main
/var/log/postgresql/postgresql-16-main.log
Запустить сервер:
student$ sudo pg_ctlcluster 16 main start
Перезапустить:
student$ sudo pg ctlcluster 16 main restart
Получить текущий статус сервера:
student$ sudo pg_ctlcluster 16 main status
pg_ctl: server is running (PID: 58483)
/usr/lib/postgresql/16/bin/postgres "-D" "/var/lib/postgresql/16/main" "-c"
"config file=/etc/postgresql/16/main/postgresql.conf"
Перечитать файлы конфигурации:
student$ sudo pg_ctlcluster 16 main reload
Журнал сообщений сервера при пакетной установке находится здесь:
student$ ls -l /var/log/postgresql/postgresql-16-main.log
-rw-r---- 1 postgres adm 44288 июл 8 14:50 /var/log/postgresql/postgresql-16-main.log
Заглянем в конец журнала:
student$ tail -n 10 /var/log/postgresql/postgresql-16-main.log
2024-07-08 14:49:59.847 MSK [58431] LOG: shutting down
2024-07-08 14:49:59.857 MSK [58431] LOG: checkpoint starting: shutdown immediate
2024-07-08 14:49:59.931 MSK [58431] LOG: checkpoint complete: wrote 3 buffers (0.0%); 0
WAL file(s) added, 0 removed, 0 recycled; write=0.021 s, sync=0.011 s, total=0.084 s;
sync files=2, longest=0.006 s, average=0.006 s; distance=0 kB, estimate=0 kB;
lsn=0/1914DE0, redo lsn=0/1914DE0
2024-07-08 14:49:59.944 MSK [58430] LOG: database system is shut down
2024-07-08 14:50:00.164 MSK [58483] LOG: starting PostgreSQL 16.3 (Ubuntu
16.3-1.pgdg22.04+1) on x86_64-pc-linux-gnu, compiled by gcc (Ubuntu
11.4.0-lubuntu1~22.04) 11.4.0, 64-bit
2024-07-08 14:50:00.164 MSK [58483] LOG: listening on IPv4 address "127.0.0.1", port 5432 2024-07-08 14:50:00.176 MSK [58483] LOG: listening on Unix socket
"/var/run/postgresql/.s.PGSQL.5432"
2024-07-08 14:50:00.206 MSK [58486] LOG: database system was shut down at 2024-07-08
```

2024-07-08 14:50:00.228 MSK [58483] LOG: database system is ready to accept connections 2024-07-08 14:50:02.651 MSK [58483] LOG: received SIGHUP, reloading configuration files

# Облачные решения



# PostgreSQL в виртуальной среде

накладные расходы на виртуализацию администратор имеет полный доступ к экземпляру

### PostgreSQL as a Service

предлагается многими облачными провайдерами основное администрирование берет на себя провайдер доступ только к ограниченным инструментам для управления, резервного копирования, мониторинга

16

Средства виртуализации могут использоваться и для баз данных, в том числе PostgreSQL. Надо понимать, что за удобства этой технологии приходится расплачиваться накладными расходами. Для высоконагруженных систем любые дополнительные промежуточные слои между СУБД и железом нежелательны.

Кроме того, многие ведущие зарубежные и российские облачные провайдеры предлагают использовать PostgreSQL как сервис (Database as a Service, managed database).

В этом случае провайдер берет на себя большинство задач администрирования. Полного доступа к экземпляру в этом случае нет, а возможности управления ограничиваются предоставленными инструментами для таких задач, как мониторинг производительности или резервное копирование и восстановление.

В рамках курса мы не будем рассматривать особенности отдельных облачных решений. Для этого стоит изучить документацию, предоставляемую провайдерами услуг.

# Итоги



Два варианта установки — пакеты или исходный код Доступны облачные решения

Сервер выполняется под выделенным пользователем ОС

Перед использованием сервера необходимо инициализировать кластер баз данных

Специальные команды для управления сервером

17

# Практика



Включение расчета контрольных сумм в кластере.

- 1. Остановите сервер.
- 2. Проверьте, рассчитываются ли контрольные суммы в кластере.
- 3. Включите расчет контрольных сумм.
- 4. Запустите сервер.

18

2, 3. Используйте утилиту pg\_checksums. Для ее запуска укажите полный путь: /usr/lib/postgresql/16/bin/pg\_checksums <a href="https://postgrespro.ru/docs/postgresql/16/app-pgchecksums">https://postgrespro.ru/docs/postgresql/16/app-pgchecksums</a>

#### 1. Останов сервера

```
student$ sudo pg_ctlcluster 16 main stop
```

#### 2. Проверка

```
Чтобы узнать, включен ли расчет контрольных сумм страниц, запустим утилиту pg_checksums с ключом --check: student$ sudo /usr/lib/postgresql/16/bin/pg_checksums --check -D /var/lib/postgresql/16/main pg_checksums: error: data checksums are not enabled in cluster
```

#### 3. Включение расчета контрольных сумм

Выполняем pg\_checksums c ключом --enable:

```
student$ sudo /usr/lib/postgresql/16/bin/pg_checksums --enable -D /var/lib/postgresql/16/main
```

Checksum operation completed

Files scanned: 1244 Blocks scanned: 3741 Files written: 1026 Blocks written: 3741

pg\_checksums: syncing data directory
pg\_checksums: updating control file
Checksums enabled in cluster

Расчет контрольных сумм включен.

#### 4. Запуск сервера

student\$ sudo pg\_ctlcluster 16 main start

# Практика+



1. Установите PostgreSQL из исходных кодов так, как это показано в демонстрации.

Создайте кластер баз данных, запустите сервер.

Убедитесь, что сервер работает.

Остановите сервер.

19

### 1. Сборка PostgreSQL из исходных кодов

Все необходимые команды были показаны в демонстрации, но они не выполнялись. Повторите эти действия своими руками в виртуальной машине курса.

Имейте в виду, что компиляция исходных кодов может занимать продолжительное время (от нескольких минут до десятков минут в зависимости от мощности компьютера).