



Локализация



Назначение и возможности

Настройка сервера

Настройка клиента

Правила сортировки

Сообщения сервера и утилит

Работа с датами, числами, денежными единицами

Поддержка национальных языков и культурных традиций

Основана на возможностях локализации ОС

Поддерживаются основные кодировки русского языка,
включая UNICODE

Перевод сообщений сервера и утилит

Проверка локалей в ОС

Инициализация кластера

`initdb --locale=локаль`

формат задания *локали*: Язык_Территория[.НаборСимволов]

пример: ru_RU.UTF-8

можно изменить отдельные категории локали

Параметры сервера

<i>название</i>	<i>контекст</i>	<i>описание</i>
lc_ctype	internal	классификация символов, upper/lower
lc_collate	internal	правила сортировки символов
lc_messages	superuser	язык сообщений
lc_monetary	user	формат денежных единиц
lc_numeric	user	формат чисел
lc_time	user	формат даты и времени

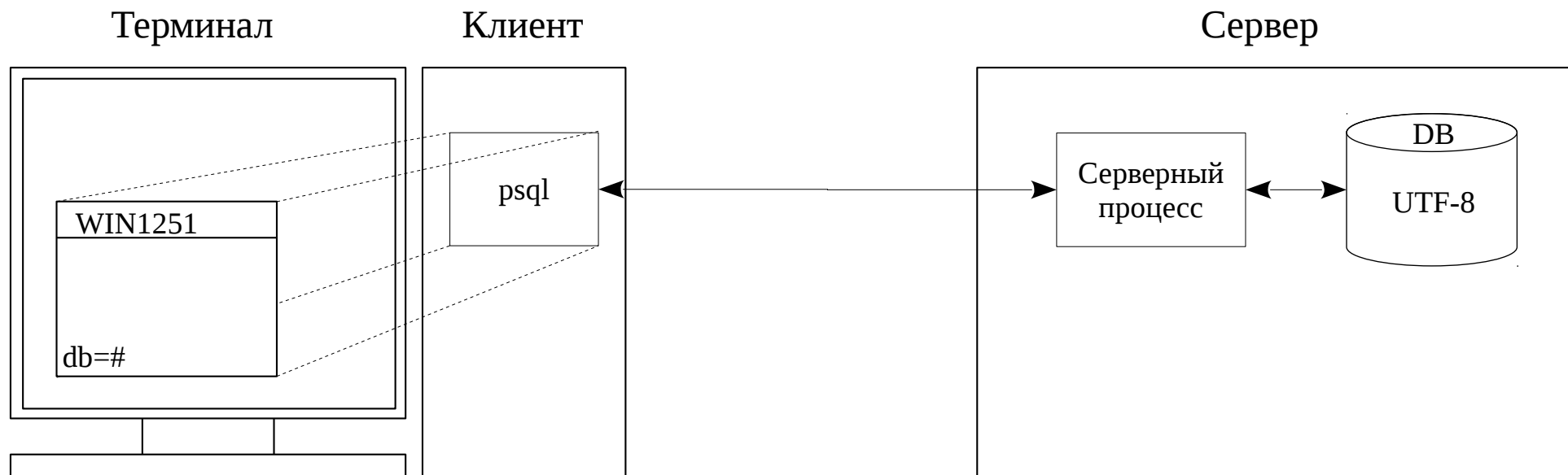
Опции локализации при создании БД

ENCODING
LC_CTYPE
LC_COLLATE

Информация

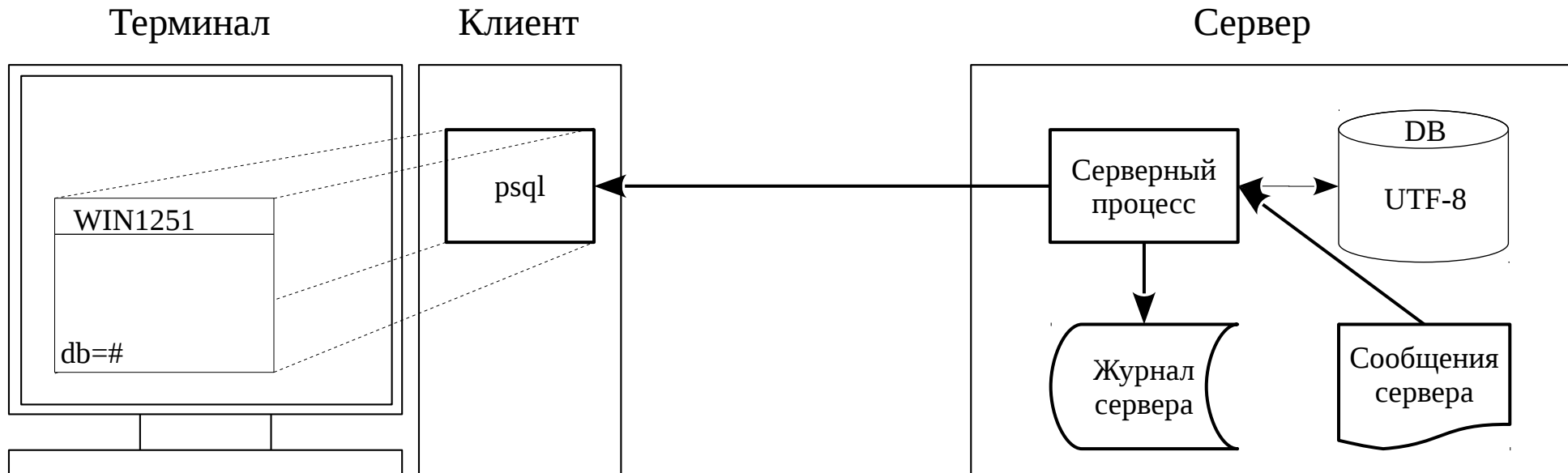
`pg_database(encoding, datcollate, datatype)`
`\list`

Настройка клиента



- Проверка настройки сервера
- Проверка локалей клиентской ОС
- Настройка сеанса ОС
- Настройка приложения
- Параметр `client_encoding`

Сообщения сервера

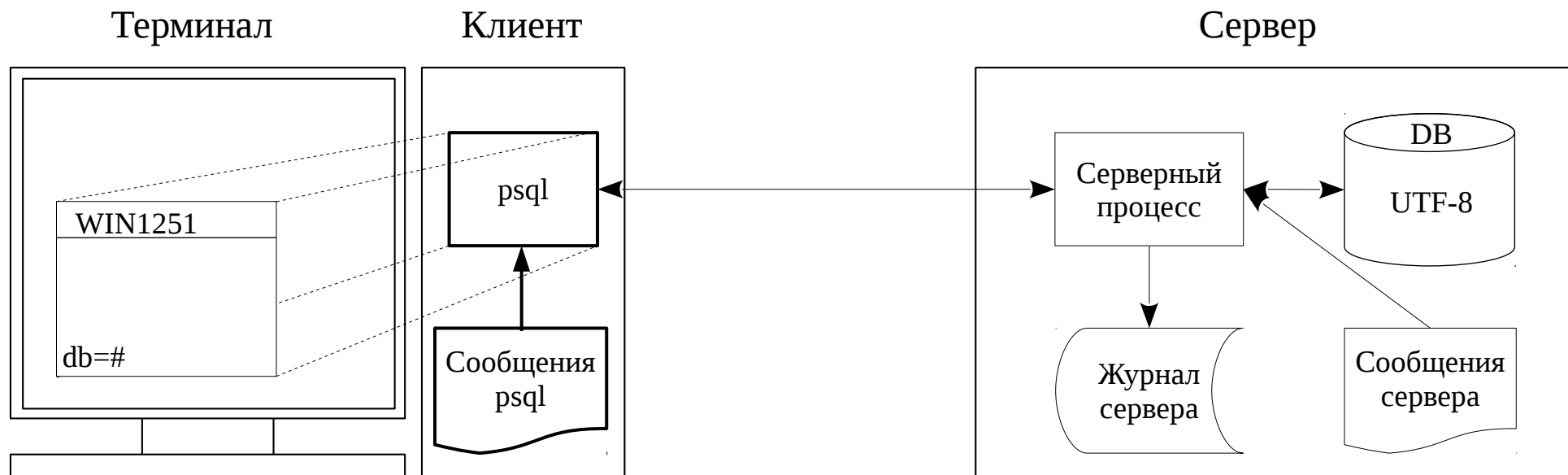


Проверка поддержки NLS на сервере

Параметр `lc_messages`

Сообщения отправляются не только клиенту, но и в журнал сервера

Сообщения утилит



Проверка поддержки NLS на клиенте
Переменная среды LC_MESSAGES

Параметры `lc_ctype`, `lc_collation`

классификация символов и правила сортировки
нельзя изменить после создания БД

Collation — объект БД

переопределение для отдельных столбцов и выражений
возможны разные значения для `lc_ctype` и `lc_collation`
правило сортировки «C»

Информация

`pg_collation`

`\dO`

Даты и время

`lc_time` — локаль для форматирования даты и времени

`DateStyle` — формат вывода и ввода значений даты и времени

`IntervalStyle` — формат вывода для значений интервалов

Числа

`lc_numeric` — локаль для форматирования чисел

Денежные единицы

`lc_monetary` — локаль для форматирования денежных сумм

Демонстрация



PostgreSQL использует возможности локализации ОС

Локаль задается при инициализации кластера или при создании БД

Клиент и сервер могут работать в разных кодировках.

Преобразование данных выполняется автоматически

LC_STYPE и LC_COLLATE задаются при создании БД

и не могут быть изменены. Их можно изменить для отдельных столбцов, индексов, выражений

Сообщения сервера и утилит переведены на несколько языков, включая русский

Для работы с датами, числами и денежными единицами есть соответствующие категории локали и параметры конфигурации

1. Кодировки базы данных

Создайте базу данных db22 с кодировкой KOI8R.

Сделайте копию базы db22 утилитой pg_dump с ключом -C (включить в копию команду create database).

Проверьте полученную копию. Можно ли загрузить ее на сервер, где не установлена локаль с кодировкой KOI8R?

2. Получите номер дня недели для функции now()

Условие: понедельник — первый день недели.



Авторские права

Курс «Администрирование PostgreSQL 9.4. Расширенный курс» разработан в компании Postgres Professional (2015 год).

Авторы: Егор Рогов, Павел Лузанов

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 косвенным, 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.

Назначение и возможности

Настройка сервера

Настройка клиента

Правила сортировки

Сообщения сервера и утилит

Работа с датами, числами, денежными единицами

Поддержка национальных языков и культурных традиций

Основана на возможностях локализации ОС

Поддерживаются основные кодировки русского языка, включая UNICODE

Перевод сообщений сервера и утилит

Под локализацией понимается возможность использования национальных языков и культурных традиций разных стран, например формат вывода даты, чисел и т. д.

PostgreSQL использует возможности локализации, которые предоставляет операционная система сервера.

Для русского языка поддерживаются все основные кодировки символов, включая UTF8, WIN1251, KOI8R, ISO_8859_5.

Кроме того, PostgreSQL предлагает инструменты для поддержки национальных языков при выводе сообщений сервера и утилит.

Подробнее о локализации в PostgreSQL:

<http://www.postgresql.org/docs/9.5/static/charset.html>

Проверка локалей в ОС

Инициализация кластера

```
initdb --locale=локаль
```

формат задания *локали*: Язык_Территория[.НаборСимволов]

пример: ru_RU.UTF-8

можно изменить отдельные категории локали

Начальная настройка локализации сервера БД.

Прежде всего нужно убедиться, что в операционной системе сервера установлены необходимые локали. Для Unix-систем проверка делается командой: `locale -a`.

Дальше можно пойти двумя путями:

- Перед запуском `initdb` выставить значения локалей в переменных ОС (`LC_*`). Тогда `initdb` будет использовать именно эти значения. Пакетные дистрибутивы могут предполагать, что локаль в ОС выставлена верно.
- При самостоятельной инициализации кластера использовать ключи локализации `initdb`: `encoding`, `locale`, `lc-messages`, `lc-*`

Подробнее об утилите `initdb`:

<http://www.postgresql.org/docs/9.5/static/app-initdb.html>

Параметры сервера

<i>название</i>	<i>контекст</i>	<i>описание</i>
lc_ctype	internal	классификация символов, upper/lower
lc_collate	internal	правила сортировки символов
lc_messages	superuser	язык сообщений
lc_monetary	user	формат денежных единиц
lc_numeric	user	формат чисел
lc_time	user	формат даты и времени

PostgreSQL поддерживает 6 категорий локали. Для каждой из них есть соответствующий параметр конфигурации.

lc_ctype и lc_collate определяются в момент создания БД и в дальнейшем их изменить нельзя.

Опции локализации при создании БД

```
ENCODING  
LC_STYPE  
LC_COLLATE
```

Информация

```
pg_database(encoding, datcollate, datatype)  
\list
```

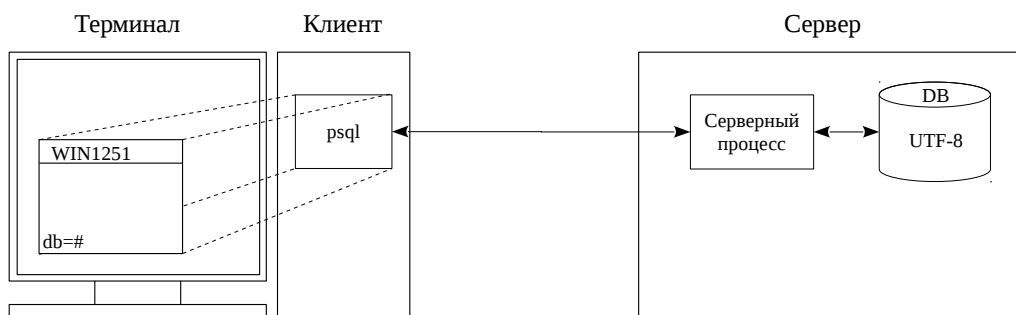
При создании базы данных можно указать параметры локализации, отличные от тех, что были выбраны при инициализации кластера.

Для этого у команды CREATE DATABASE есть опции ENCODING, LC_STYPE и LC_COLLATE.

Соответствующие локали должны быть установлены в ОС.

Важно отметить, что базу данных в таком случае можно создать только из шаблона template0.

Настройка клиента



Проверка настройки сервера
Проверка локалей клиентской ОС
Настройка сеанса ОС
Настройка приложения
Параметр `client_encoding`

7

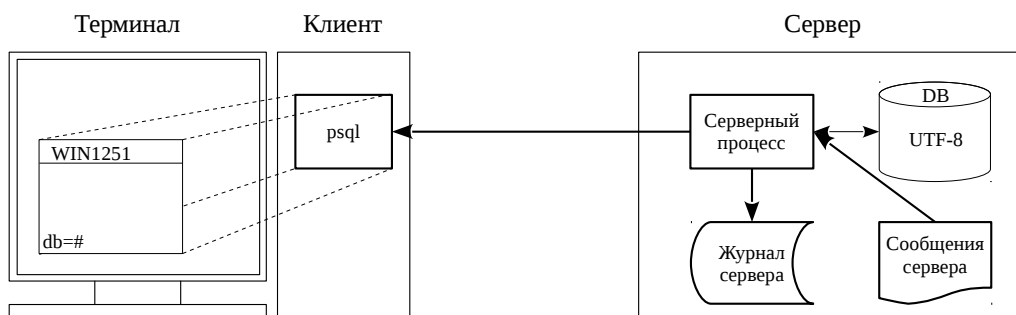
Для настройки локализации клиентского приложения нужно сделать следующее.

- Проверить, что настройки сервера корректны. Как минимум, что используется правильная кодировка БД (параметр `server_encoding`).
- Проверить, что в клиентской ОС установлены нужные локали.
- Настроить сеанс пользователя в ОС (переменные `LC_*`).
- Настроить кодировку приложения. Например, в Windows (`cmd.exe`) для кодировки `WIN1251` нужно:
 - Установить шрифты `Lucida Console`.
 - Выполнить команду `chcp 1251`.
- После этого подключиться к БД и проверить параметр `client_encoding`. При необходимости, установить его значение в значение кодировки приложения. Значение параметра можно задать и в переменной среды `PGCLIENTENCODING`.

Важно, что клиент и сервер могут работать в разных кодировках. Для большинства кодировок в PostgreSQL преднастроены процедуры автоматического преобразования.

Процедуры преобразования кодировок находятся в таблице системного каталога `pg_conversion`. Возможно создание дополнительных пользовательских процедур (`CREATE CONVERSION`).

Сообщения сервера



Проверка поддержки NLS на сервере

Параметр `lc_messages`

Сообщения отправляются не только клиенту, но и в журнал сервера

8

Сообщения сервера и утилит PostgreSQL переведены на несколько языков. В том числе и на русский.

Для того, чтобы сообщения сервера выводились на русском языке, нужно убедиться, что сервер PostgreSQL был собран с поддержкой NLS. Это можно проверить командой:

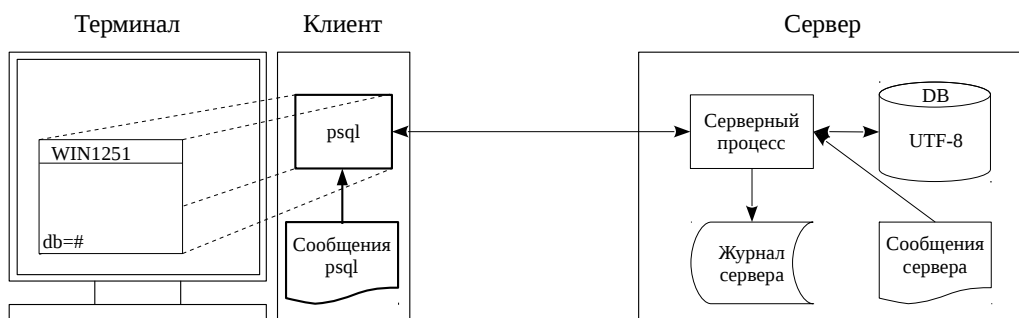
```
pg_config --configure
```

Параметр `lc_messages` управляет языком сообщений сервера.

Важно учитывать тот факт, что сообщения сервера отправляются не только клиенту, но и записываются в журнал сервера. При выборе языка, отличного от английского, нужно быть уверенным, что инструменты работы с журналом сервера понимают другие языки. Например, `pgbadger` требует установки `lc_messages` в «C».

Поддержка NLS – перевод сообщений сервера для переводчиков и разработчиков:

<http://www.postgresql.org/docs/9.5/static/nls.html>



Проверка поддержки NLS на клиенте
Переменная среды LC_MESSAGES

Утилиты PostgreSQL (psql, pg_dump и пр.) также поддерживают NLS.

Для того, чтобы сообщения утилит выводились на русском языке, нужно убедиться, что PostgreSQL на клиенте был собран с поддержкой NLS и установить переменную среды LC_MESSAGES.

Большинство ОС (включая Windows) используют следующий порядок просмотра переменных среды: LANGUAGE, LC_ALL, LC_MESSAGES, LANG.

Параметры `lc_ctype`, `lc_collation`

классификация символов и правила сортировки
нельзя изменить после создания БД

Collation — объект БД

переопределение для отдельных столбцов и выражений
возможны разные значения для `lc_ctype` и `lc_collation`
правило сортировки «C»

Информация

`pg_collation`
`\dO`

Параметры сервера `lc_ctype` и `lc_collate` определяют классификацию символов и правила сортировки. Эти параметры нельзя изменить после создания БД.

Но для отдельных столбцов и выражений есть способ использовать другие значения. Для этого используется специальный объект базы данных – `collation` (правила сортировки).

`Collation` представляет собой комбинацию значений `lc_ctype` и `lc_collate`. Начальный список правил сортировки формируется при инициализации кластера. Возможные значения считываются из локалей ОС.

В дальнейшем, при добавлении локалей в ОС можно создавать дополнительные правила сортировки.

Специальное правило сортировки «C» создается для всех кодировок сервера. Это правило говорит о том, что буквами будут считаться только латинские символы от A до Z, все остальные знаки будут сортироваться в соответствии со своими кодами в данной кодировке.

Посмотреть имеющиеся правила сортировки можно в таблице системного каталога `pg_collation`.

Дополнительная информация о работе с правилами сортировок:
<http://www.postgresql.org/docs/9.5/interactive/collation.html>

Даты и время

lc_time — локаль для форматирования даты и времени

DateStyle — формат вывода и ввода значений даты и времени

IntervalStyle — формат вывода для значений интервалов

Числа

lc_numeric — локаль для форматирования чисел

Денежные единицы

lc_monetary — локаль для форматирования денежных сумм

Для работы с датами, числами и денежными единицами используются различные параметры конфигурации.

Подробнее об этом в документации:

- функции форматирования данных:

<http://www.postgresql.org/docs/9.5/static/functions-formatting.html>

- параметры конфигурации:

<http://www.postgresql.org/docs/9.5/interactive/runtime-config-client.html#RUNTIME-CONFIG-CLIENT-FORMAT>

- поддержка даты и времени:

<http://www.postgresql.org/docs/9.5/static/datetime-appendix.html>



PostgreSQL использует возможности локализации ОС

Локаль задается при инициализации кластера или при создании БД

Клиент и сервер могут работать в разных кодировках.

Преобразование данных выполняется автоматически

LC_STYPE и LC_COLLATE задаются при создании БД

и не могут быть изменены. Их можно изменить для отдельных столбцов, индексов, выражений

Сообщения сервера и утилит переведены на несколько языков, включая русский

Для работы с датами, числами и денежными единицами есть соответствующие категории локали и параметры конфигурации

1. Кодировки базы данных

Создайте базу данных db22 с кодировкой KOI8R.

Сделайте копию базы db22 утилитой `pg_dump` с ключом `-C` (включить в копию команду `create database`).

Проверьте полученную копию. Можно ли загрузить ее на сервер, где не установлена локаль с кодировкой KOI8R?

2. Получите номер дня недели для функции `now()`

Условие: понедельник — первый день недели.

1. Для создания БД в кодировке KOI8R

Убедитесь, что в ОС установлена нужная локаль

В команде `CREATE DATABASE` используйте шаблон `template0` и опции `ENCODING`, `LC_STYPE`, `LC_COLLATE`

2. Для получения номера дня недели используйте функцию `to_char`.

Допустимые форматные маски даты:

<http://www.postgresql.org/docs/current/interactive/functions-formatting.html#FUNCTIONS-FORMATTING-DATETIME-TABLE>