

Ltree: extending syntax

Dmitry Belyavskiy

PgConf.ru

Moscow, Russia

February 4-6, 2019

Who am I?

- Dmitry Belyavskiy, Moscow, Russia
- Main job: Cryptocom LTD (<http://www.cryptocom.ru/>, in Russian)
 - Senior Developer
- Participation in Open Source projects:
 - GOST engine
 - XMLSec
 - Some Perl modules (Net::DRI, some OpenSSL wrappers)
 - PostgreSQL Itree contrib — subject of this presentation

What is ltree?

<https://www.postgresql.org/docs/11/ltree.html>

Type Ltree — representation of hierarchical structures

```
Top.Science.Astronomy
```

```
Top.Science.Astronomy.Astrophysics
```

```
Top.Science.Astronomy.Cosmology
```

```
Top.Collections.Pictures.Astronomy
```

```
Top.Collections.Pictures.Astronomy.Stars
```

```
Top.Collections.Pictures.Astronomy.Galaxies
```

Types lquery/ltxquery — requests

```
SELECT path FROM test WHERE path ~ '*.Astronomy.*';
```

Why extending syntax?

Limited set of symbols

Real-life example

65 — «Экономика. Экономические науки»

65.2/4 — «Сектора и отрасли экономики»

65.32 — «Экономика сельского хозяйства»

65.321 — «Экономика сельскохозяйственных предприятий»

65.321.8 — «Сельскохозяйственные кооперативы».

Brave New syntax

Problem: Limited set of characters

Solution: Escaping

Usability: Backslash or quotes

Examples: “a b c”, a\ b\ c, “a\ b\ c”

- Internal form: a b c
- External form: “a b c”

Special symbols **MUST** be escaped

Some internal details

State machine:

```
#define LTPRS_WAITNAME 0  
#define LTPRS_WAITDELIM 1  
#define LTPRS_WAITESCAPED 2  
#define LTPRS_WAITDELIMSTRICT 3
```

No limited lists of symbols

memcpy => char-by-char copying

Asserts: bytes copied == bytes expected

Implementing, testing, debugging

- Use USE_PGXS=1, Luke!
- make USE_PGXS=1 all install
- make USE_PGXS=1 check
 - Many, many tests

If you do not see what's wrong...

- `SELECT pid FROM pg_stat_activity()`
- `gdb --attach`

Result:

<https://commitfest.postgresql.org/22/1977/>

Questions?

Drop 'em all

beldmit@gmail.com

beldmit@cryptocom.ru