JsQuery
the jsonb query language with GIN indexing support
October, 2014, Madrid, Spain

Alexander Korotkov, Intaro
Oleg Bartunov, Teodor Sigaev, SAI MSU
Oleg Bartunov, Teodor Sigaev

• Locale support
• Extendability (indexing)
  • GiST (KNN), GIN, SP-GiST
• Full Text Search (FTS)
• Jsonb, VODKA
• Extensions:
  • intarray
  • pg_trgm
  • ltree
  • hstore
  • plantuner

https://www.facebook.com/oleg.bartunov
obartunov@gmail.com, teodor@sigaev.ru
https://www.facebook.com/groups/postgresql/
Alexander Korotkov

- Indexed regexp search
- GIN compression & fast scan
- Fast GiST build
- Range types indexing
- Split for GiST
- Indexing for jsonb
- jsquery
- Generic WAL + create am (WIP)

aekorotkov@gmail.com
Agenda

• Introduction into jsonb
• Jsonb indexing
• Jquery - Jsonb Query Language
• Jsonb GIN opclasses with JsQuery support
• Future of Jsonb querying
Introduction to hstore

• Hstore benefits
  • In provides a flexible model for storing a semi-structured data in relational database
  • hstore has binary storage and rich set of operators and functions, indexes

• Hstore drawbacks
  • Too simple model!
    Hstore key-value model doesn't supports tree-like structures as json
    (introduced in 2006, 3 years after hstore)

• Json — popular and standartized (ECMA-404 The JSON Data Interchange Standard, JSON RFC-7159)

• Json — PostgreSQL 9.2, textual storage
Nested hstore

Title: One step forward true json data type. Nested hstore with array support.

We present a prototype of nested hstore data type with array support. We consider the new hstore as a step forward true json data type.

Recently, PostgreSQL got json data type, which basically is a string storage with validity checking for stored values and some related functions. To be a real data type, it has to have a binary representation, which could be a big project if started from scratch. Hstore is a popular data type, we developed years ago to facilitate working with semi-structured data in PostgreSQL. Our idea is to extend hstore to be nested (value can be hstore) data type and add support of arrays, so its binary representation can be shared with json.

We present a working prototype of a new hstore data type and discuss some design and implementation issues.
Nested hstore & jsonb

- Nested hstore at PGCon-2013, Ottawa, Canada (May 24) — thanks Engine Yard for support!

  One step forward true json data type. Nested hstore with arrays support

- Binary storage for nested data at PGCon Europe — 2013, Dublin, Ireland (Oct 29)

  Binary storage for nested data structures and application to hstore data type

- November, 2013 — binary storage was reworked, nested hstore and jsonb share the same storage. Andrew Dunstan joined the project.

- January, 2014 - binary storage moved to core
Nested hstore & jsonb

- Feb-Mar, 2014 - Peter Geoghegan joined the project, nested hstore was cancelled in favour to jsonb (Nested hstore patch for 9.3).
- Mar 23, 2014  Andrew Dunstan committed jsonb to 9.4 branch!

`pgsql: Introduce jsonb, a structured format for storing json.`

Introduce jsonb, a structured format for storing json.

The new format accepts exactly the same data as the json type. However, it is stored in a format that does not require reparsing the orgiginal text in order to process it, making it much more suitable for indexing and other operations. Insignificant whitespace is discarded, and the order of object keys is not preserved. Neither are duplicate object keys kept - the later value for a given key is the only one stored.
Jsonb vs Json

```
SELECT '{"c":0,  "a":2,"a":1}'::json, '{"c":0,   "a":2,"a":1}'::jsonb;
```

<table>
<thead>
<tr>
<th>json</th>
<th>jsonb</th>
</tr>
</thead>
<tbody>
<tr>
<td>{&quot;c&quot;:0,  &quot;a&quot;:2,&quot;a&quot;:1}</td>
<td>{&quot;a&quot;: 1, &quot;c&quot;: 0}</td>
</tr>
</tbody>
</table>

(1 row)

- json:    textual storage «as is»
- jsonb:   no whitespaces
- jsonb:   no duplicate keys, last key win
- jsonb:   keys are sorted
Jsonb vs Json

• Data
  • 1,252,973 Delicious bookmarks

• Server
  • MBA, 8 GB RAM, 256 GB SSD

• Test
  • Input performance - copy data to table
  • Access performance - get value by key
  • Search performance contains @> operator
Jsonb vs Json

Data
• 1,252,973 bookmarks from Delicious in json format (js)
• The same bookmarks in jsonb format (jb)
• The same bookmarks as text (tx)

=\# \texttt{\textbackslash dt+}

List of relations

<table>
<thead>
<tr>
<th>Schema</th>
<th>Name</th>
<th>Type</th>
<th>Owner</th>
<th>Size</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>public</td>
<td>jb</td>
<td>table</td>
<td>postgres</td>
<td>1374 MB</td>
<td>overhead is &lt; 4%</td>
</tr>
<tr>
<td>public</td>
<td>js</td>
<td>table</td>
<td>postgres</td>
<td>1322 MB</td>
<td></td>
</tr>
<tr>
<td>public</td>
<td>tx</td>
<td>table</td>
<td>postgres</td>
<td>1322 MB</td>
<td></td>
</tr>
</tbody>
</table>
Jsonb vs Json

- Input performance (parser)
  Copy data (1,252,973 rows) as text, json, jsonb

  copy tt from '/path/to/test.dump'

  Text: 34 s - as is
  Json: 37 s - json validation
  Jsonb: 43 s - json validation, binary storage
Jsonb vs Json (binary storage)

• Access performance — get value by key
  • Base: SELECT js FROM js;
  • Jsonb: SELECT j->>'updated' FROM jb;
  • Json: SELECT j->>'updated' FROM js;

  Base: 0.6 s
  Jsonb: 1 s 0.4
  Json: 9.6 s 9

Jsonb ~ 20X faster Json
EXPLAIN ANALYZE SELECT count(*) FROM js WHERE js #>>'{tags,0,term}' = 'NYC';

<table>
<thead>
<tr>
<th>QUERY PLAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate (cost=187812.38..187812.39 rows=1 width=0)</td>
</tr>
<tr>
<td>(actual time=10054.602..10054.602 rows=1 loops=1)</td>
</tr>
<tr>
<td>-&gt; Seq Scan on js (cost=0.00..187796.88 rows=6201 width=0)</td>
</tr>
<tr>
<td>(actual time=0.030..10054.426 rows=123 loops=1)</td>
</tr>
<tr>
<td>Filter: ((js #&gt;&gt; '{tags,0,term} '::text[]) = 'NYC '::text)</td>
</tr>
<tr>
<td>Rows Removed by Filter: 1252850</td>
</tr>
<tr>
<td>Planning time: 0.078 ms</td>
</tr>
<tr>
<td>Execution runtime: 10054.635 ms</td>
</tr>
<tr>
<td>(6 rows)</td>
</tr>
</tbody>
</table>

Json: no contains @> operator, search first array element
Jsonb vs Json (binary storage)

EXPLAIN ANALYZE SELECT count(*) FROM jb WHERE jb @> '{"tags": [{"term": "NYC"}]}':::jsonb;

<table>
<thead>
<tr>
<th>QUERY PLAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggregate (cost=191521.30..191521.31 rows=1 width=0)</td>
</tr>
<tr>
<td>(actual time=1263.201..1263.201 rows=1 loops=1)</td>
</tr>
<tr>
<td>-&gt; Seq Scan on jb (cost=0.00..191518.16 rows=1253 width=0)</td>
</tr>
<tr>
<td>(actual time=0.007..1263.065 rows=285 loops=1)</td>
</tr>
<tr>
<td>Filter: (jb @&gt; '{&quot;tags&quot;: [{&quot;term&quot;: &quot;NYC&quot;}]}':::jsonb)</td>
</tr>
<tr>
<td>Rows Removed by Filter: 1252688</td>
</tr>
<tr>
<td>Planning time: 0.065 ms</td>
</tr>
<tr>
<td>Execution runtime: 1263.225 ms</td>
</tr>
</tbody>
</table>

Execution runtime: 10054.635 ms

(6 rows)

Jsonb  ~ 10X faster Json
Jsonb vs Json (GIN: key && value)

CREATE INDEX gin_jb_idx ON jb USING gin(jb);

EXPLAIN ANALYZE SELECT count(*) FROM jb WHERE jb @> '{"tags": [{"term": "NYC"}]}':::jsonb;

---

Aggregate (cost=4772.72..4772.73 rows=1 width=0)
(actual time=8.486..8.486 rows=1 loops=1)
   -> Bitmap Heap Scan on jb (cost=73.71..4769.59 rows=1253 width=0)
      (actual time=8.049..8.462 rows=285 loops=1)
         Recheck Cond: (jb @> '{"tags": [{"term": "NYC"}]}':::jsonb)
         Heap Blocks: exact=285
      -> Bitmap Index Scan on gin_jb_idx (cost=0.00..73.40 rows=1253 width=0)
         (actual time=8.014..8.014 rows=285 loops=1)
            Index Cond: (jb @> '{"tags": [{"term": "NYC"}]}':::jsonb)
Planning time: 0.115 ms
Execution runtime: 8.515 ms
(8 rows)

Jsonb ~ 150X faster Json
Jsonb vs Json (GIN: hash path.value)

CREATE INDEX gin_jb_path_idx ON jb USING gin(jb jsonb_path_ops);

EXPLAIN ANALYZE SELECT count(*) FROM jb WHERE jb @> '{"tags": [{"term": "NYC"}]}::jsonb;

---

Aggregate (cost=4732.72..4732.73 rows=1 width=0)
(actual time=0.644..0.644 rows=1 loops=1)
  -> Bitmap Heap Scan on jb (cost=33.71..4729.59 rows=1253 width=0)
  (actual time=0.102..0.620 rows=285 loops=1)
    Recheck Cond: (jb @> '{"tags": [{"term": "NYC"}]}::jsonb)
    Heap Blocks: exact=285
  -> Bitmap Index Scan on gin_jb_path_idx
  (cost=0.00..33.40 rows=1253 width=0) (actual time=0.062..0.062 rows=285 loops=1)
    Index Cond: (jb @> '{"tags": [{"term": "NYC"}]}::jsonb)
Planning time: 0.056 ms
Execution runtime: 0.668 ms
(8 rows)

Jsonb ~ 1800X faster Json
MongoDB 2.6.0

• **Load data - ~13 min SLOW !**
  
  mongoimport --host localhost -c js --type json < delicious-rss-1250k
  
  2014-04-08T22:47:10.014+0400
  
  3700 1233/second
  
  ...
  
  2014-04-08T23:00:36.050+0400
  
  1252000 1547/second
  
  2014-04-08T23:00:36.565+0400 check 9 1252973
  
  2014-04-08T23:00:36.566+0400 imported 1252973 objects

• **Search - ~ 1s (seqscan) THE SAME**
  
  db.js.find({tags: {$elemMatch: {term: "NYC"}}}).count()
  
  285
  
  -- 980 ms

• **Search - ~ 1ms (indexscan) Jsonb 0.7ms**
  
  db.js.ensureIndex( {"tags.term" : 1} )
  
  db.js.find({tags: {$elemMatch: {term: "NYC"}}}).
### Summary: PostgreSQL 9.4 vs Mongo 2.6.0

- **Operator contains @>**
  - json : 10 s seqscan
  - jsonb : 8.5 ms GIN jsonb_ops
  - jsonb : 0.7 ms GIN jsonb_path_ops
  - mongo : 1.0 ms btree index

- **Index size**
  - jsonb_ops : - 636 Mb (no compression, 815Mb)
  - jsonb_path_ops : - 295 Mb
  - jsonb_path_ops (tags) : - 44 Mb USING gin((jb->'tags') jsonb_path_ops
  - mongo (tags) : - 387 Mb
  - mongo (tags.term) : - 100 Mb

- **Table size**
  - postgres : 1.3Gb
  - mongo : 1.8Gb

- **Input performance:**
  - Text : 34 s
  - Json : 37 s
  - Jsonb : 43 s
  - mongo : 13 m
Jsonb (Apr, 2014)

• Documentation
  • JSON Types, JSON Functions and Operators

• There are many functionality left in nested hstore
  • Can be an extension

• Need query language for jsonb
  • <,>,&& ... operators for values
    a.b.c.d && [1,2,10]
  • Structural queries on paths
    *.d && [1,2,10]
  • Indexes !
Currently, one can search jsonb data using:

- **Contains operators** - `jsonb @> jsonb`, `jsonb @@ jsonb` (GIN indexes)
  
  `jb @> '{"tags": ["term": "NYC"]}'::jsonb`

  Keys should be specified from root

- **Equivalence operator** — `jsonb = jsonb` (GIN indexes)

- **Exists operators** — `jsonb ? text`, `jsonb ?! text[]`, `jsonb ?& text[]` (GIN indexes)
  
  `jb WHERE jb ?| '{tags,links}'`

  Only root keys supported

- **Operators on jsonb parts** (functional indexes)
  
  `SELECT ('{"a": {"b":5}}'::jsonb -> 'a'->>'b')::int > 2;`

  `CREATE INDEX ....USING BTREE ( (jb->'a'->>'b')::int);`

  Very cumbersome, too many functional indexes
Jsonb querying an array: simple case

Find bookmarks with tag «NYC»:

```
SELECT * 
FROM js 
WHERE js @> '{"tags":[{"term":"NYC"}]}'
```
Jsonb querying an array: complex case

Find companies where CEO or CTO is called Neil.
One could write...

SELECT * FROM company
WHERE js @> '{"relationships":[{"person":
  {"first_name":"Neil"}],}' AND
  (js @> '{"relationships":[{"title":"CTO"}],}' OR
   js @> '{"relationships":[{"title":"CEO"}],}');
Jsonb querying an array: complex case

Each «@>> is processed independently.

```sql
SELECT * FROM company
WHERE js @> '{"relationships":[{"person":
    {
        "first_name":"Neil"
    }]}]}' AND
  (js @> '{"relationships":[{"title":"CTO"}]}' OR
   js @> '{"relationships":[{"title":"CEO"}]}'
);
```

Actually, this query searches for companies with some CEO or CTO and someone called Neil...
Jsonb querying an array: complex case

The correct version is so.

```sql
SELECT * FROM company
WHERE js @@ ' {
  "relationships": [
    { "title": "CEO",
      "person": { "first_name": "Neil" } ]
  }'
OR js @@ ' {
  "relationships": [
    { "title": "CTO",
      "person": { "first_name": "Neil" } ]
  }
'};
```

When constructing complex conditions over same array element, query length can grow exponentially.
Jsonb querying an array: another approach

Using subselect and jsonb_array_elements:

```
SELECT * FROM company
WHERE EXISTS (
    SELECT 1
    FROM jsonb_array_elements(js -> 'relationships') t
    WHERE t->>'title' IN ('CEO', 'CTO') AND
    t ->'person'->>'first_name' = 'Neil');
```
Jsonb querying an array: summary

Using «@>»
- Pro
  - Indexing support
- Cons
  - Checks only equality for scalars
  - Hard to explain complex logic

Using subselect and jsonb_array_elements
- Pro
  - Full power of SQL can be used to express condition over element
- Cons
  - No indexing support
  - Heavy syntax
Jsonb query

• Need Jsonb query language
  • Simple and effective way to search in arrays (and other iterative searches)
  • More comparison operators
  • Types support
  • Schema support (constraints on keys, values)
  • Indexes support

• Introduce Jsquery - textual data type and @@ match operator

  jsonb @@ jsquery
Jsonb query language (Jsquery)

```
Expr ::= path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT value_expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
       '(' expr ')' |
       path value_expr |
       path HINT value_expr |
       NOT expr |
       NOT HINT value_expr |
       path '(' expr ')' |
�
```
Jsonb query language (Jsquery)

• # - any element array

```sql
SELECT '{"a": {"b": [1,2,3]}}'::jsonb @@ 'a.b.# = 2';
```

• % - any key

```sql
SELECT '{"a": {"b": [1,2,3]}}'::jsonb @@ '%.b.# = 2';
```

• * - anything

```sql
SELECT '{"a": {"b": [1,2,3]}}'::jsonb @@ '*.# = 2';
```

• $ - current element

```sql
select '{"a": {"b": [1,2,3]}}'::jsonb @@ 'a.b.# ($ = 2 OR $ < 3)';
```

• Use "double quotes" for key!

```sql
select 'a1."12222" < 111':jsquery;
```
Jsonb query language (Jsquery)

• Scalar

```sql
select '{"a": {"b": [1,2,3]}}'::jsonb @@ 'a.b.# IN (1,2,5)';
```

• Test for key existence

```sql
select '{"a": {"b": [1,2,3]}}'::jsonb @@ 'a.b = *';
```

• Array overlap

```sql
select '{"a": {"b": [1,2,3]}}'::jsonb @@ 'a.b && [1,2,5]';
```

• Array contains

```sql
select '{"a": {"b": [1,2,3]}}'::jsonb @@ 'a.b @> [1,2]';
```

• Array contained

```sql
select '{"a": {"b": [1,2,3]}}'::jsonb @@ 'a.b <@ [1,2,3,4,5]';
```
Jsonb query language (Jsquery)

• Type checking

```sql
select '{"x": true}' @@ 'x IS boolean'::jsquery,
     '{"x": 0.1}' @@ 'x IS numeric'::jsquery;
?column? | ?column?
----------+----------
t        | t

select '{"a":{"a":1}}' @@ 'a IS object'::jsquery;
?column?
----------
t
select '{"a":['"xxx"]}' @@ 'a IS array'::jsquery, '[$ IS array'::jsquery;
?column? | ?column?
----------+----------
t        | t
```

<table>
<thead>
<tr>
<th>IS BOOLEAN</th>
<th>IS NUMERIC</th>
<th>IS ARRAY</th>
<th>IS OBJECT</th>
<th>IS STRING</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>t</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Jsonb query language (Jsquery)

• How many products are similar to "B000089778" and have product_sales_rank in range between 10000-20000?

• SQL
  SELECT count(*) FROM jr WHERE (jr->>'product_sales_rank')::int > 10000 and (jr->>'product_sales_rank')::int < 20000 and 
  ....boring stuff

• Jsquery
  SELECT count(*) FROM jr WHERE jr @@ 'similar_product_ids && ['B000089778'] AND product_sales_rank($ > 10000 AND $ < 20000)' 

• Mongodb
  db.reviews.find( { $and : [ { similar_product_ids: { $in : ['B000089778']}}, 
  {product_sales_rank:{$gt:10000, $lt:20000}} ] } ).count()
«#», «*», «%» usage rules

Each usage of «#», «*», «%» means separate element

• Find companies where CEO or CTO is called Neil.

```sql
SELECT count(*) FROM company WHERE js @@ 'relationships.#(title in ("CEO", "CTO") AND person.first_name = "Neil")'::jsquery;
```

<table>
<thead>
<tr>
<th>count</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
</tr>
</tbody>
</table>

• Find companies with some CEO or CTO and someone called Neil

```sql
SELECT count(*) FROM company WHERE js @@ 'relationships(#.title in ("CEO", "CTO") AND #.person.first_name = "Neil")'::jsquery;
```

<table>
<thead>
<tr>
<th>count</th>
</tr>
</thead>
<tbody>
<tr>
<td>69</td>
</tr>
</tbody>
</table>
Jsonb query language (Jsquery)

```
explain( analyze, buffers) select count(*) from jb where jb @> '{"tags":[{"term":"NYC"}]}'::jsonb;
```

**QUERY PLAN**

---

```
Aggregate  (cost=191514.16..191514.16 rows=1 width=0) (actual time=1039.422..1039.423 rows=1 loops=1)
  Buffers: shared hit=97841 read=78011
  ->  Seq Scan on jb  (cost=0.00..191514.16 rows=1253 width=0) (actual time=0.006..1039.310 rows=285 loops=1)
       Filter: (jb @> '{"tags":[{"term":"NYC"}]}'::jsonb)
       Rows Removed by Filter: 1252688
       Buffers: shared hit=97841 read=78011
Planning time: 0.074 ms
```

**Execution time: 1039.444 ms**

```
explain( analyze, costs off) select count(*) from jb where jb @@ 'tags.#.term = "NYC"';
```

**QUERY PLAN**

---

```
Aggregate (actual time=891.553..891.553 rows=1 loops=1)
  ->  Seq Scan on jb (actual time=0.010..891.553 rows=285 loops=1)
       Filter: (jb @@ 'tags.#.term = "NYC"':jsquery)
       Rows Removed by Filter: 1252688
```

**Execution time: 891.745 ms**
Jsquery (indexes)

- **GIN opclasses with jsquery support**
  - `jsonb_value_path_ops` — use Bloom filtering for key matching
    \[
    \{"a":\{"b":\{"c":10\}\}\} \rightarrow 10.( \text{bloom}(a) \text{ or } \text{bloom}(b) \text{ or } \text{bloom}(c) )
    \]
    - Good for key matching (wildcard support), not good for range query
  
  - `jsonb_path_value_ops` — hash path (like `jsonb_path_ops`)
    \[
    \{"a":\{"b":\{"c":10\}\}\} \rightarrow \text{hash}(a.b.c).10
    \]
    - No wildcard support, no problem with ranges

---

<table>
<thead>
<tr>
<th>Schema</th>
<th>Name</th>
<th>Type</th>
<th>Owner</th>
<th>Table</th>
<th>Size</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>public</td>
<td>jb</td>
<td>table</td>
<td>postgres</td>
<td></td>
<td>1374 MB</td>
<td></td>
</tr>
<tr>
<td>public</td>
<td>jb_value_path_idx</td>
<td>index</td>
<td>postgres</td>
<td>jb</td>
<td>306 MB</td>
<td></td>
</tr>
<tr>
<td>public</td>
<td>jb_gin_idx</td>
<td>index</td>
<td>postgres</td>
<td>jb</td>
<td>544 MB</td>
<td></td>
</tr>
<tr>
<td>public</td>
<td>jb_path_value_idx</td>
<td>index</td>
<td>postgres</td>
<td>jb</td>
<td>306 MB</td>
<td></td>
</tr>
<tr>
<td>public</td>
<td>jb_path_idx</td>
<td>index</td>
<td>postgres</td>
<td>jb</td>
<td>251 MB</td>
<td></td>
</tr>
</tbody>
</table>
explain( analyze, costs off) select count(*) from jb where jb @@ 'tags.#.term = "NYC"';

QUERY PLAN

-------------------------------------------------------------------------------------------------
Aggregate (actual time=0.609..0.609 rows=1 loops=1)
 -> Bitmap Heap Scan on jb (actual time=0.115..0.580 rows=285 loops=1)
   Recheck Cond: (jb @@ "tags".#."term" = "NYC"::jsquery)
   Heap Blocks: exact=285
   -> Bitmap Index Scan on jb_value_path_idx (actual time=0.073..0.073 rows=285 loops=1)
     Index Cond: (jb @@ "tags".#."term" = "NYC"::jsquery)

Execution time: 0.634 ms
(7 rows)
explain( analyze, costs off) select count(*) from jb where jb @@ '*'.term = "NYC";

QUERY PLAN

-------------------------------------------------------------------------------------------------
Aggregate (actual time=0.688..0.688 rows=1 loops=1)
  -> Bitmap Heap Scan on jb (actual time=0.145..0.660 rows=285 loops=1)
      Recheck Cond: (jb @@ '*'."term" = "NYC"::jsquery)
      Heap Blocks: exact=285
  -> Bitmap Index Scan on jb_value_path_idx (actual time=0.113..0.113 rows=285 loops=1)
      Index Cond: (jb @@ '*'."term" = "NYC"::jsquery)
Execution time: 0.716 ms
(7 rows)
Citus dataset

- 3023162 reviews from Citus 1998-2000 years
- 1573 MB

```json
{
  "customer_id": "AE22YDHSBFYIP",
  "product_category": "Business & Investing",
  "product_group": "Book",
  "product_id": "1551803542",
  "product_sales_rank": 11611,
  "product_subcategory": "General",
  "product_title": "Start and Run a Coffee Bar (Start & Run a)",
  "review_date": {
    "$date": 31363200000
  },
  "review_helpful_votes": 0,
  "review_rating": 5,
  "review_votes": 10,
  "similar_product_ids": [
    "0471136174",
    "0910627312",
    "047112138X",
    "0786883561",
    "0201570483"
  ]
}
```
explain (analyze, costs off) select count(*) from jr where
jr @@ 'similar_product_ids' && ['B000089778'];

QUERY PLAN

-----------------------------------------------------------------------------------
Aggregate (actual time=0.359..0.359 rows=1 loops=1)
  -> Bitmap Heap Scan on jr (actual time=0.084..0.337 rows=185 loops=1)
     Recheck Cond: (jr @@ "similar_product_ids" && ['B000089778']::jsquery)
     Heap Blocks: exact=107
        -> Bitmap Index Scan on jr_path_value_idx (actual time=0.057..0.057 rows=185 loops=1)
           Index Cond: (jr @@ "similar_product_ids" && ['B000089778']::jsquery)

Execution time: 0.394 ms
(7 rows)
Jsquery (indexes)

• No statistics, no planning :( 

```
explain (analyze, costs off) select count(*) from jr where
jr @@ 'similar_product_ids' && "B000089778"
AND product_sales_rank( $ > 10000 AND $ < 20000)
```

**QUERY PLAN**

```
--------------------------------------------------------------------------------------------------------------------------------------
Aggregate (actual time=126.149..126.149 rows=1 loops=1)
 -> Bitmap Heap Scan on jr (actual time=126.057 ..126.143 rows=45 loops=1)
   Recheck Cond: (jr @@ '("similar_product_ids" && ["B000089778"] &
"product_sales_rank"($ > 10000 & $ < 20000))':'jsquery)
   Heap Blocks: exact=45
 -> Bitmap Index Scan on jr_path_value_idx (actual time=126.029..126.029 rows=45 loops=1)
   Index Cond: (jr @@ '("similar_product_ids" && ["B000089778"] &
"product_sales_rank"($ > 10000 & $ < 20000))':'jsquery)
```

Execution time: 129.309 ms !!! No statistics
(7 rows)
MongoDB 2.6.0

```javascript
db.reviews.find( { $and : [ {similar_product_ids : { $in : ["B000089778"]}}, {product_sales_rank : { $gt : 10000, $lt : 20000 }} ] } ).explain()
{
    "n" : 45,
------------------------
    "millis" : 7,
    "indexBounds" : {
        "similar_product_ids" : [
            ["B000089778",
             "B000089778"
            ]
        ]
    },
}
```
Jsquery (indexes)

- If we rewrite query and use planner

```
explain (analyze,costs off) select count(*) from jr where
jr @@ 'similar_product_ids' && ['B000089778']
and (jr->>'product_sales_rank')::int>10000 and (jr->>'product_sales_rank')::int<20000;
```

```
Aggregate (actual time=0.479..0.479 rows=1 loops=1)
-> Bitmap Heap Scan on jr (actual time=0.079..0.472 rows=45 loops=1)
  Recheck Cond: (jr @@ "similar_product_ids" && ['B000089778']::jsquery)
  Filter: (((jr ->> 'product_sales_rank'::text))::integer > 10000) AND
      (((jr ->> 'product_sales_rank'::text))::integer < 20000))
  Rows Removed by Filter: 140
  Heap Blocks: exact=107
-> Bitmap Index Scan on jr_path_value_idx (actual time=0.041..0.041 rows=185 loops=1)
  Index Cond: (jr @@ "similar_product_ids" && ['B000089778']::jsquery)
```

Execution time: **0.506 ms**  Potentially, query could be faster Mongo!

(9 rows)
Jsquery (optimizer) — NEW!

- Jsquery now has built-in simple optimiser.

```sql
explain (analyze, costs off) select count(*) from jr where
jr @@ 'similar_product_ids && ["B000089778"]
AND product_sales_rank( $ > 10000 AND $ < 20000)'
```

Aggregate (actual time=0.422..0.422 rows=1 loops=1)
  -> Bitmap Heap Scan on jr (actual time=0.099..0.416 rows=45 loops=1)
    Recheck Cond: (jr @@ '("similar_product_ids" && ["B000089778"] AND
"product_sales_rank"($ > 10000 AND $ < 20000))':jsquery)
    Rows Removed by Index Recheck: 140
    Heap Blocks: exact=107
  -> Bitmap Index Scan on jr_path_value_idx (actual time=0.060..0.060 rows=185 loops=1)
    Index Cond: (jr @@ '("similar_product_ids" && ["B000089778"] AND
"product_sales_rank"($ > 10000 AND $ < 20000))':jsquery)

Execution time: **0.480 ms vs 7 ms MongoDB!**
Jsquery (optimizer) — NEW!

- Since GIN opclasses can't expose something special to explain output, jsquery optimiser has its own explain functions:

  - text `gin_debug_query_path_value(jsquery)` — explain for `jsonb_path_value_ops`
    
    ```
    # SELECT gin_debug_query_path_value('x = 1 AND (*.y = 1 OR y = 2)');
    
    gin_debug_query_path_value
    -----------------------------
    x = 1 , entry 0           +
    ```

  - text `gin_debug_query_value_path(jsquery)` — explain for `jsonb_value_path_ops`
    
    ```
    # SELECT gin_debug_query_value_path('x = 1 AND (*.y = 1 OR y = 2)');
    
    gin_debug_query_value_path
    --------------------------------
    AND                       +
    x = 1 , entry 0         +
    OR                      +
    *.y = 1 , entry 1     +
    y = 2 , entry 2       +
    ```
Jsquery now has built-in optimiser for simple queries.
Analyze query tree and push non-selective parts to recheck (like filter)

Selectivity classes:
1) Equality (x = c)
2) Range (c1 < x < c2)
3) Inequality (c > c1)
4) Is (x is type)
5) Any (x = *)
AND children can be put into recheck.

```sql
# SELECT gin_debug_query_path_value('x = 1 AND y > 0');
<table>
<thead>
<tr>
<th>gin_debug_query_path_value</th>
</tr>
</thead>
<tbody>
<tr>
<td>x = 1 , entry 0</td>
</tr>
</tbody>
</table>
```

While OR children can't. We can't handle false negatives.

```sql
# SELECT gin_debug_query_path_value('x = 1 OR y > 0');
<table>
<thead>
<tr>
<th>gin_debug_query_path_value</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR</td>
</tr>
<tr>
<td>x = 1 , entry 0</td>
</tr>
<tr>
<td>y &gt; 0 , entry 1</td>
</tr>
</tbody>
</table>
```
Can't do much with NOT, because hash is lossy. After NOT false positives turns into false negatives which we can't handle.

```sql
# SELECT gin_debug_query_path_value('x = 1 AND (NOT y = 0)');
gin_debug_query_path_value
-----------------------------
x = 1 , entry 0                +
```
Jsquery (optimizer) — NEW!

- Jsquery optimiser pushes non-selective operators to recheck

```sql
explain (analyze, costs off) select count(*) from jr where jr @@ 'similar_product_ids && ["B000089778"] AND product_sales_rank( $ > 10000 AND $ < 20000)'
```

```
Aggregate (actual time=0.422..0.422 rows=1 loops=1)
  -> Bitmap Heap Scan on jr (actual time=0.099..0.416 rows=45 loops=1)
    Recheck Cond: (jr @@ '("similar_product_ids" && ["B000089778"] AND "product_sales_rank"($ > 10000 AND $ < 20000))')::jsquery)
    Rows Removed by Index Recheck: 140
    Heap Blocks: exact=107
  -> Bitmap Index Scan on jr_path_value_idx (actual time=0.060..0.060 rows=185 loops=1)
    Index Cond: (jr @@ '("similar_product_ids" && ["B000089778"] AND "product_sales_rank"($ > 10000 AND $ < 20000))')::jsquery)
Execution time: 0.480 ms
```
Jsquery (HINTING) — NEW!

- Jsquery now has HINTING (if you don't like optimiser)!

```
explain (analyze, costs off) select count(*) from jr where jr @@ 'product_sales_rank > 10000'
```

```
Aggregate (actual time=2507.410..2507.410 rows=1 loops=1)
  -> Bitmap Heap Scan on jr (actual time=1118.814..2352.286 rows=2373140 loops=1)
    Recheck Cond: (jr @@ "product_sales_rank" > 10000::jsquery)
    Heap Blocks: exact=201209
  -> Bitmap Index Scan on jr_path_value_idx (actual time=1052.483..1052.48 rows=2373140 loops=1)
    Index Cond: (jr @@ "product_sales_rank" ::jsquery)
Execution time: 2524.951 ms
```

- Better not to use index — HINT /*--noindex */

```
explain (analyze, costs off) select count(*) from jr where jr @@ 'product_sales_rank /*-- noindex */ > 10000';
```

```
Aggregate (actual time=1376.262..1376.262 rows=1 loops=1)
  -> Seq Scan on jr (actual time=0.013..1222.123 rows=2373140 loops=1)
    Filter: (jr @@ "product_sales_rank" /*-- noindex */ > 10000::jsquery)
    Rows Removed by Filter: 650022
Execution time: 1376.284 ms
```
Jsquery (HINTING) — NEW!

- If you know that inequality is selective then use HINT /* --index */

```sql
# explain (analyze, costs off) select count(*) from jr where jr @@ 'product_sales_rank /*-- index*/ > 3000000 AND review_rating = 5'::jsquery;
```

**QUERY PLAN**

```
Aggregate (actual time=12.307..12.307 rows=1 loops=1)
  -> Bitmap Heap Scan on jr (actual time=11.259..12.244 rows=739 loops=1)
    Recheck Cond: (jr @@ '("product_sales_rank" /*-- index*/ > 3000000 AND "review_rating" = 5')::jsquery)
    Heap Blocks: exact=705
  -> Bitmap Index Scan on jr_path_value_idx (actual time=11.179..11.179 rows=739 loops=1)
    Index Cond: (jr @@ '("product_sales_rank" /*-- index*/ > 3000000 AND "review_rating" = 5')::jsquery)
```

**Execution time:** 12.359 ms vs 1709.901 ms (without hint)
(7 rows)
Contrib/jsquery

- Jquery index support is quite efficient (0.5 ms vs Mongo 7 ms!)
- Future direction
  - Make jjquery planner friendly
  - Need statistics for jsonb
- Availability
  - Jquery + opclasses are available as extensions
  - Grab it from https://github.com/akorotkov/jsquery (branch master), we need your feedback!
  - We will release it after PostgreSQL 9.4 release
  - Need real sample data and queries!
Stop following me, you fucking freaks!

PostgreSQL 9.4+
- Open-source
- Relational database
- Strong support of json
Better indexing ...

- GIN is a proven and effective index access method
- Need indexing for jsonb with operations on paths (no hash!) and values
  - B-tree in entry tree is not good - length limit, no prefix compression

<table>
<thead>
<tr>
<th>Schema</th>
<th>Name</th>
<th>Type</th>
<th>Owner</th>
<th>Table</th>
<th>Size</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>public</td>
<td>jb</td>
<td>table</td>
<td>postgres</td>
<td></td>
<td>1374 MB</td>
<td></td>
</tr>
<tr>
<td>public</td>
<td>jb_uniq_paths</td>
<td>table</td>
<td>postgres</td>
<td></td>
<td>912 MB</td>
<td></td>
</tr>
<tr>
<td>public</td>
<td>jb_uniq_paths_btree_idx</td>
<td>index</td>
<td>postgres</td>
<td>jb_uniq_paths</td>
<td>885 MB</td>
<td>text_pattern_ops</td>
</tr>
<tr>
<td>public</td>
<td>jb_uniq_paths_spgist_idx</td>
<td>index</td>
<td>postgres</td>
<td>jb_uniq_paths</td>
<td>598 MB</td>
<td>now much less!</td>
</tr>
</tbody>
</table>
Better indexing ...

- Provide interface to change hardcoded B-tree in Entry tree
  - Use spgist opclass for storing paths and values as is (strings hashed in values)
- We may go further - provide interface to change hardcoded B-tree in posting tree
  - GIS aware full text search!
- New index access method

```
CREATE INDEX ... USING VODKA
```
GIN History

• Introduced at PostgreSQL Anniversary Meeting in Toronto, Jul 7-8, 2006 by Oleg Bartunov and Teodor Sigaev

**Generalized Inverted Index**

• An inverted index is an index structure storing a set of (key, posting list) pairs, where 'posting list' is a set of documents in which the key occurs.

• Generalized means that the index does not know which operation it accelerates. It works with custom strategies, defined for specific data types. GIN is similar to GiST and differs from B-Tree indices, which have predefined, comparison-based operations.
GIN History

- Introduced at PostgreSQL Anniversary Meeting in Toronto, Jul 7-8, 2006 by Oleg Bartunov and Teodor Sigaev
- Supported by JFG Networks (France)
- «Gin stands for Generalized Inverted iNdex and should be considered as a genie, not a drink.»
- Alexander Korotkov, Heikki Linnakangas have joined GIN++ development in 2013
GIN History

• From GIN Readme, posted in -hackers, 2006-04-26

TODO

----

Nearest future:

  * Opclasses for all types (no programming, just many catalog changes).

Distant future:

  * Replace B-tree of entries to something like GiST (VODKA ! 2014)
  * Add multicolumn support
  * Optimize insert operations (background index insertion)
GIN index structure for jsonb

```json
{
    "product_group": "Book",
    "product_sales_rank": 15000
},
{
    "product_group": "Music",
    "product_sales_rank": 25000
}
```
Vodka index structure for jsonb

```json
{
    "product_group": "Book",
    "product_sales_rank": 15000
},
{
    "product_group": "Music",
    "product_sales_rank": 25000
}
```
CREATE INDEX … USING VODKA

• Delicious bookmarks, mostly text data

set maintenance_work_mem = '1GB';

<table>
<thead>
<tr>
<th>Schema</th>
<th>Name</th>
<th>Type</th>
<th>Owner</th>
<th>Table</th>
<th>Size</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>public</td>
<td>jb</td>
<td>table</td>
<td>postgres</td>
<td></td>
<td>1374 MB</td>
<td>1252973 rows</td>
</tr>
<tr>
<td>public</td>
<td>jb_value_path_idx</td>
<td>index</td>
<td>postgres</td>
<td>jb</td>
<td>306 MB</td>
<td>98769.096</td>
</tr>
<tr>
<td>public</td>
<td>jb_gin_idx</td>
<td>index</td>
<td>postgres</td>
<td>jb</td>
<td>544 MB</td>
<td>129860.859</td>
</tr>
<tr>
<td>public</td>
<td>jb_path_value_idx</td>
<td>index</td>
<td>postgres</td>
<td>jb</td>
<td>306 MB</td>
<td>100560.313</td>
</tr>
<tr>
<td>public</td>
<td>jb_path_idx</td>
<td>index</td>
<td>postgres</td>
<td>jb</td>
<td>251 MB</td>
<td>68880.320</td>
</tr>
<tr>
<td>public</td>
<td>jb_vodka_idx</td>
<td>index</td>
<td>postgres</td>
<td>jb</td>
<td>409 MB</td>
<td>185362.865</td>
</tr>
<tr>
<td>public</td>
<td>jb_vodka_idx5</td>
<td>index</td>
<td>postgres</td>
<td>jb</td>
<td>325 MB</td>
<td>174627.234 new spgist</td>
</tr>
</tbody>
</table>

(6 rows)
CREATE INDEX ... USING VODKA

```sql
select count(*) from jb where jb @@ 'tags.#.term = "NYC"';
```

Aggregate (actual time=0.423..0.423 rows=1 loops=1)
  -> Bitmap Heap Scan on jb (actual time=0.146..0.404 rows=285 loops=1)
    Recheck Cond: (jb @@ "tags".#."term" = "NYC"::jsquery)
    Heap Blocks: exact=285
  -> Bitmap Index Scan on jb_vodka_idx (actual time=0.108..0.108 rows=285 loops=1)
    Index Cond: (jb @@ "tags".#."term" = "NYC"::jsquery)

Execution time: 0.456 ms (0.634 ms, GIN jsonb_value_path_ops)

```sql
select count(*) from jb where jb @@ '*.term = "NYC"';
```

Aggregate (actual time=0.495..0.495 rows=1 loops=1)
  -> Bitmap Heap Scan on jb (actual time=0.245..0.474 rows=285 loops=1)
    Recheck Cond: (jb @@ "."."term" = "NYC"::jsquery)
    Heap Blocks: exact=285
  -> Bitmap Index Scan on jb_vodka_idx (actual time=0.214..0.214 rows=285 loops=1)
    Index Cond: (jb @@ "."."term" = "NYC"::jsquery)

Execution time: 0.526 ms (0.716 ms, GIN jsonb_path_value_ops)
CREATE INDEX ... USING VODKA

- CITUS data, text and numeric

set maintenance_work_mem = '1GB';

<table>
<thead>
<tr>
<th>Schema</th>
<th>Name</th>
<th>Type</th>
<th>Owner</th>
<th>Table</th>
<th>Size</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>public</td>
<td>jr</td>
<td>table</td>
<td>postgres</td>
<td></td>
<td>1573 MB</td>
<td>3023162 rows</td>
</tr>
<tr>
<td>public</td>
<td>jr_value_path_idx</td>
<td>index</td>
<td>postgres</td>
<td>jr</td>
<td>196 MB</td>
<td>79180.120</td>
</tr>
<tr>
<td>public</td>
<td>jr_gin_idx</td>
<td>index</td>
<td>postgres</td>
<td>jr</td>
<td>235 MB</td>
<td>111814.929</td>
</tr>
<tr>
<td>public</td>
<td>jr_path_value_idx</td>
<td>index</td>
<td>postgres</td>
<td>jr</td>
<td>196 MB</td>
<td>73369.713</td>
</tr>
<tr>
<td>public</td>
<td>jr_path_idx</td>
<td>index</td>
<td>postgres</td>
<td>jr</td>
<td>180 MB</td>
<td>48981.307</td>
</tr>
<tr>
<td>public</td>
<td>jr_vodka_idx3</td>
<td>index</td>
<td>postgres</td>
<td>jr</td>
<td>240 MB</td>
<td>155714.777</td>
</tr>
<tr>
<td>public</td>
<td>jr_vodka_idx4</td>
<td>index</td>
<td>postgres</td>
<td>jr</td>
<td>211 MB</td>
<td>169440.130 new spgist</td>
</tr>
</tbody>
</table>

(6 rows)
CREATE INDEX ... USING VODKA

explain (analyze, costs off) select count(*) from jr where jr @@ 'similar_product_ids && ["B000089778"]';

QUERY PLAN

-------------------------------------------------------------------------------------------
Aggregate (actual time=0.200..0.200 rows=1 loops=1)
  -> Bitmap Heap Scan on jr (actual time=0.090..0.183 rows=185 loops=1)
     Recheck Cond: (jr @@ "similar_product_ids" && ["B000089778"])::jsquery)
     Heap Blocks: exact=107
  -> Bitmap Index Scan on jr_vodka_idx (actual time=0.077 ..0.077 rows=185 loops=1)
     Index Cond: (jr @@ "similar_product_ids" && ["B000089778"])::jsquery)

Execution time: 0.237 ms (0.394 ms, GIN jsonb_path_value_idx)
(7 rows)
There are can be different flavors of Vodka
New VODKA concept

- Posting list/tree is just a way of effective duplicate storage
- Entry tree can consist of multiple levels of different access methods
- VODKA is a way to combine different access method in single index: VODKA CONNECTING INDEXES
JsQuery limitations

• Variables are always on the left size
  \[ x = 1 \quad \text{OK} \]
  \[ 1 = x \quad \text{Error!} \]

• No calculations in query
  \[ x + y = 0 \quad \text{Error!} \]

• No extra datatypes and search operators
  \[ \text{point}(x,y) <@ '((0,0),(1,1),(2,1),(1,0))'::\text{polygon} \]
JsQuery limitations

Users want jsquery to be as rich as SQL...
JsQuery limitations

Users want jsquery to be as rich as SQL ...

... But we will discourage them ;}
JsQuery language goals

• Provide rich enough query language for jsonb in 9.4.
• Indexing support for 'jsonb @@ jsquery':
  • Two GIN opclasses are in jsquery itself
  • VODKA opclasses was tested on jsquery

It's NOT intended to be solution for jsonb querying in long term!
What JsQuery is NOT?

It's **not** designed to be another **extendable, full weight:**
• Parser
• Executor
• Optimizer

It's NOT SQL inside SQL.
Jsonb querying an array: summary

Using «@>»

• Pro
  • Indexing support
• Cons
  • Checks only equality for scalars
  • Hard to explain complex logic

Using subselect and jsonb_array_elements

• Pro
  • SQL-rich
• Cons
  • No indexing support
  • Heavy syntax

JsQuery

• Pro
  • Indexing support
  • Rich enough for typical applications
• Cons
  • Not extendable

Still looking for a better solution!
Jsonb query: future

Users want jsonb query language to be as rich as SQL. How to satisfy them?..
Jsonb query: future

Users want jsonb query language to be as rich as SQL. How to satisfy them?

Bring all required features to SQL-level!
Functional equivalents:

- SELECT * FROM company WHERE EXISTS (SELECT 1 FROM jsonb_array_elements(js->'relationships') t WHERE t->>'title' IN ('CEO', 'CTO') AND t->'person'->>'first_name' = 'Neil');
- SELECT count(*) FROM company WHERE js @@ 'relationships(#.title in ("CEO", "CTO") AND #.person.first_name = "Neil")':jsquery;
- SELECT * FROM company WHERE ANYELEMENT OF js-> 'relationships' AS t ( t->>'title' IN ('CEO', 'CTO') AND t ->'person'->>'first_name' = 'Neil');
Jsonb query: ANYELEMENT

Possible implementation steps:
• Implement ANYELEMENT just as syntactic sugar and only for arrays.
• Support for various data types (extendable?)
• Handle ANYELEMENT as expression not subselect (problem with alias).
• Indexing support over ANYELEMENT expressions.
Another idea about ANYELEMENT

Functional equivalents:

• SELECT t
  FROM company,
    LATERAL (SELECT t FROM
        jsonb_array_elements(js->'relationships') t) el;

• SELECT t
  FROM company,
    ANYELEMENT OF js->'relationships' AS t;
Summary

• contrib/jsquery for 9.4
  • Jquery - Jsonb Query Language
  • Two GIN opclasses with jsquery support
  • Grab it from https://github.com/akorotkov/jsquery (branch master)

• Prototype of VODKA access method (supported by Heroku)
• New VODKA concept
• Idea of Jsonb querying in SQL
We invite to PGConf.RU in Moscow, February 2015!

WELCOME!
Thanks for support

WARGAMING.NET

LET'S BATTLE