Data Processing Inside PostgreSQL

BRUCE MOMJIAN



There are indisputable advantages of doing data processing in the database rather than in each application. This presentation explores the ability to push data processing into the database using SQL, functions, triggers, and the object-relational features of POSTGRESQL.

https://momjian.us/presentations

Creative Commons Attribution License



Last updated: June 2024

Pre-SQL Data Access

No one wants to return to this era:

- Complex cross-table access
- Single index
- No optimizer
- Simple WHERE processing
- No aggregation



SQL Data Access

You probably take these for granted:

- Easy cross-table access, with optimizer assistance
- Complex WHERE processing
- Transaction Control
- Concurrency
- Portable language (SQL)

Post-Ingres

Welcome to the next generation of data storage.



https://www.flickr.com/photos/mtshaw/

Contents

- 1. SQL
- 2. Functions and Triggers
- 3. Customizing Database Features

1. SQL

Make full use of the SQL tools available.



2. Functions and Triggers

Put your programs in the database.



3. Customizing Database Features

Change the database features.



1. SQL



https://www.flickr.com/photos/ajc1/

Table Constraints

Table creation requires concentration.



Unique Test in an Application

```
BEGIN;
LOCK tab;
SELECT ... WHERE col = key;
if not found
    INSERT (or UPDATE)
COMMIT;
```

UNIQUE Constraint

```
CREATE TABLE tab
(
    col ... UNIQUE
);
CREATE TABLE customer (id INTEGER UNIQUE);
```

Preventing NULLs

```
if (col != NULL)
    INSERT/UPDATE;
```

NOT NULL Constraint

```
CREATE TABLE tab
(
    col ... NOT NULL
);
CREATE TABLE customer (name TEXT NOT NULL);
```

Primary Key Constraint

- UNIQUE
- NOT NULL CREATE TABLE customer (id INTEGER PRIMARY KEY);

Ensuring Table Linkage: Foreign → Primary

```
BEGIN;
SELECT *
FROM primary
WHERE key = col
FOR UPDATE;
if found
    INSERT (or UPDATE) INTO foreign;
COMMIT;
```

$Primary \rightarrow Foreign \\$

```
BEGIN;
SELECT *
FROM foreign WHERE col = key FOR UPDATE;
if found ? UPDATE/DELETE primary;
COMMIT;
```

Ensuring Table Linkage, Example

```
CREATE TABLE statename (
                        CHAR(2) PRIMARY KEY,
                   code
                         VARCHAR (30)
                   name
);
CREATE TABLE customer
                         customer id INTEGER,
                                     VARCHAR(30).
                         name
                         telephone
                                     VARCHAR(20),
                         street
                                     VARCHAR(40).
                                     VARCHAR(25),
                         city
                                     CHAR(2) REFERENCES statename,
                         state
                                     CHAR(10).
                         zipcode
                                     VARCHAR(20)
                         country
);
```

Larger Example

```
CREATE TABLE customer
                       customer id INTEGER PRIMARY KEY.
                                   VARCHAR(30).
                       name
                       telephone
                                   VARCHAR(20),
                       street
                                    VARCHAR(40),
                       city
                                    VARCHAR(25).
                                   CHAR(2),
                       state
                       zipcode
                                   CHAR(10),
                                   VARCHAR(20)
                       country
);
CREATE TABLE employee
                       employee id INTEGER PRIMARY KEY,
                                   VARCHAR(30),
                       name
                                   DATE
                       hire date
);
```

Larger Example

```
CREATE TABLE part (
                  part id
                              INTEGER PRIMARY KEY,
                              VARCHAR(30).
                   name
                              NUMERIC(8,2),
                   cost
                  weight
                              FLOAT
);
CREATE TABLE salesorder (
                        order id
                                      INTEGER,
                        customer id
                                     INTEGER REFERENCES customer,
                        employee id
                                     INTEGER REFERENCES employee,
                        part id
                                      INTEGER REFERENCES part,
                        order date
                                      DATE.
                        ship date
                                      DATE
);
```

Prevent Change to Primary

```
BEGIN;
SELECT ...
FROM foreign
WHERE col = key
FOR UPDATE;
if found
    ABORT;
UPDATE/DELETE primary;
COMMIT;
```

REFERENCES Constraint, NO ACTION/RESTRICT (default)

Cascade Change to Primary

```
BEGIN;
SELECT ...
FROM foreign
WHERE col = key
FOR UPDATE;
if found
          UPDATE/DELETE foreign;
UPDATE/DELETE primary;
COMMIT;
```

REFERENCES Constraint, CASCADE

```
CREATE TABLE foreign (
col ... REFERENCES primary (col)
ON UPDATE CASCADE
ON DELETE CASCADE
);
```

Set Foreign to NULL on Change to Primary

```
BEGIN;
SELECT ...
FROM foreign
WHERE col = key
FOR UPDATE;
if found
          UPDATE foreign SET col = NULL;
UPDATE/DELETE primary;
COMMIT;
```

REFERENCES Constraint, SET NULL

```
CREATE TABLE foreign
(
    col ... REFERENCES primary (col)
    ON UPDATE SET NULL
    ON DELETE SET NULL
);
```

Set Foreign to DEFAULT on Change to Primary

```
BEGIN;
SELECT ...
FROM foreign
WHERE col = key
FOR UPDATE;
if found
         UPDATE foreign SET col = DEFAULT;
UPDATE/DELETE primary;
COMMIT;
```

REFERENCES Constraint, SET DEFAULT

```
CREATE TABLE foreign
(
    col ... REFERENCES primary (col)
        ON UPDATE SET DEFAULT
        ON DELETE SET DEFAULT
);

CREATE TABLE order (cust_id INTEGER REFERENCES customer (id));
```

Controlling Data

```
if col > 0 ...
  (col = 2 OR col = 7) ...
length(col) < 10 ...
INSERT/UPDATE tab;</pre>
```

CHECK Constraint

```
CREATE TABLE tab
(
    col ... CHECK (col > 0 ...
);

CREATE TABLE customer (age INTEGER CHECK (age >= 0));
```

Check Constraint Example

```
CREATE TABLE friend2 (
            firstname VARCHAR(15).
            lastname VARCHAR(20).
            city VARCHAR(15).
            state CHAR(2)
                                  CHECK (length(trim(state)) = 2),
                                  CHECK (age >= 0).
            age INTEGER
            gender CHAR(1)
                                  CHECK (gender IN ('M', 'F')),
            last met DATE
                                  CHECK (last met BETWEEN '1950-01-01'
                                         AND CURRENT DATE).
            CHECK (upper(trim(firstname)) != 'ED' OR
                   upper(trim(lastname)) != 'RIVERS')
);
INSERT INTO friend2
VALUES ('Ed', 'Rivers', 'Wibbleville', 'J', -35, 'S', '1931-09-23');
ERROR:
       ExecAppend: rejected due to CHECK constraint friend2 last met
```

Default Column Values

```
if col not specified
  col = DEFAULT;
INSERT/UPDATE tab;
```

DEFAULT Constraint

```
CREATE TABLE tab
(
    quantity ... DEFAULT 1
);
CREATE TABLE customer (created timestamp DEFAULT CURRENT_TIMESTAMP);
```

Auto-Numbering Column

```
CREATE TABLE counter (curr INTEGER);
INSERT INTO counter VALUES (1);
...
BEGIN;
val = SELECT curr FROM counter FOR UPDATE;
UPDATE counter SET curr = curr + 1;
COMMIT;
INSERT INTO tab VALUES (... val ...);
```

SERIAL/Sequence

```
CREATE TABLE tab
    col SERIAL
);
CREATE TABLE tab
    col INTEGER DEFAULT nextval('tab col seq')
);
CREATE TABLE customer (id SERIAL):
CREATE SEQUENCE customer id seq;
CREATE TABLE customer (id INTEGER DEFAULT nextval('customer id seq'));
```

Constraint Macros, DOMAIN

```
CREATE DOMAIN phone AS CHAR(12)

CHECK (VALUE \sim '\sim[0-9]{3}-[0-9]{3}-[0-9]{4}$');

CREATE TABLE company ( ... phnum phone, ...);
```

Using SELECT's Features



ANSI Outer Joins - LEFT OUTER

```
SFLFCT *
FROM tab1, tab2
WHFRF tab1.col = tab2.col
UNION
SELECT *
FROM tab1
WHERE col NOT IN
    SELECT tab2.col
    FROM tab2
);
SFLFCT *
FROM tab1 LEFT JOIN tab2 ON tab1.col = tab2.col;
```

ANSI Outer Joins - RIGHT OUTER

```
SFLFCT *
FROM tab1, tab2
WHFRF tab1.col = tab2.col
UNION
SELECT *
FROM tab2
WHERE col NOT IN
    SELECT tab1.col
    FROM tab1
);
SFLFCT *
FROM tab1 RIGHT JOIN tab2 ON tab1.col = tab2.col;
```

ANSI Outer Joins - FULL OUTER

```
SELECT *
FROM tab1, tab2
WHERE tab1.col = tab2.col
UNION
SELECT *
FROM tab1
WHERE col NOT IN
    SELECT tab2.col
    FROM tab2
UNION
SELECT *
FROM tab2
WHERE col NOT IN
    SELECT tab1.col
    FROM tab1
);
SELECT *
FROM tab1 FULL JOIN tab2 ON tab1.col = tab2.col;
```

ANSI Outer Join Example

```
SELECT *
FROM customer LEFT JOIN order ON customer.id = order.cust_id;
```

Aggregates SUM()

```
total = 0
FOREACH val IN set
    total = total + val;
END FOREACH
SELECT SUM(val) FROM tab;
```

Aggregates MAX()

```
max = MIN_VAL;
FOREACH val IN set
   if (val > max)
        max = val;
END FOREACH

SELECT MAX(val) FROM tab;

SELECT MAX(cost) FROM part;
```

Aggregates GROUP BY SUM()

```
gsort(set)
save = '';
total = 0;
FOREACH val IN set
    if val != save and save != ''
        print save, total;
        save = val;
        total = 0:
    total = total + amt;
END FOREACH
if save != ''
    print save, total;
```

Aggregates, GROUP BY MAX()

```
save = '';
max = MIN VAL;
FOREACH val IN set
    if val != save and save != ''
        print save, max;
        save = val:
        max = MIN VAL;
    if (amt > max)
        max = amt:
END FOREACH
if save != ''
    print save, max;
SELECT val, MAX(amt) FROM tab GROUP BY val;
```

GROUP BY Examples

```
SELECT part, COUNT(*)
FROM order
ORDER BY part;

SELECT cust_id, SUM(due)
FROM order
GROUP BY cust_id
ORDER BY 2 DESC;
```

Merging SELECTs, UNION

```
SELECT * INTO TEMP out FROM ...
INSERT INTO TEMP out SELECT ...
INSERT INTO TEMP out SELECT ...
SELECT DISTINCT ...
SELECT *
UNION
SELECT *
UNION
SELECT *:
```

Joining SELECTs, INTERSECT

```
SELECT * INTO TEMP out;

DELETE FROM out WHERE out.* NOT IN (SELECT ...);

DELETE FROM out WHERE out.* NOT IN (SELECT ...);

SELECT *

INTERSECT

SELECT *

INTERSECT

SELECT *;
```

Subtracting SELECTs, EXCEPT

```
SELECT * INTO TEMP out;

DELETE FROM out WHERE out.* IN (SELECT ...);

DELETE FROM out WHERE out.* IN (SELECT ...);

SELECT *

EXCEPT

SELECT *

EXCEPT

SELECT *;
```

Controlling Rows Returned, LIMIT/OFFSET

```
DECLARE limdemo CURSOR FOR SELECT ...
FOR i = 1 to 5
    FFTCH IN limdemo
FND FOR
SELECT *
LIMIT 5;
DECLARE limdemo CURSOR FOR SELECT ...
MOVE 20 IN limdemo
FOR i = 1 to 5
    FETCH IN limdemo:
END FOR
SELECT *
OFFSET 20 LIMIT 5;
```

Controlling Rows Returned, LIMIT/OFFSET Example

SELECT order_id, balance FROM order ORDER BY balance DESC LIMIT 10;

Locking SELECT Rows, FOR UPDATE

```
BEGIN:
LOCK tab:
SELECT * FROM CUSTOMER WHERE id = 4452;
UPDATE customer SET balance = 0 WHERE id = 4452:
COMMIT:
BEGIN;
SFLECT *
FROM customer
WHERE id = 4452
FOR UPDATE:
. . .
UPDATE customer
SET balance = 0
WHERE id = 4452;
COMMIT:
```

Temporary Tables

```
CREATE TABLE tab (...);
. . .
DROP TABLE tab;
CREATE TEMP TABLE tab (...);
SELECT *
INTO TEMPORARY hold
FROM tab1, tab2, tab3
WHERE ...
```

Automatically Modify SELECT, VIEW - One Column

```
SELECT col4
FROM tab;

CREATE VIEW view1 AS
SELECT col4
FROM tab;

SELECT * FROM view1;
```

Automatically Modify SELECT, VIEW - One Row

```
SFLFCT *
FROM tab
WHERE col = 'ISDN';
CREATE VIEW view2 AS
SELECT *
FROM tab
WHERE col = 'ISDN';
SELECT * FROM view2;
```

Automatically Modify SELECT, VIEW - One Field

```
SFLFCT co14
FROM tab
WHERE col = 'ISDN';
CREATE VIEW view3 AS
SELECT col4
FROM tab
WHERE col = 'ISDN';
SELECT * FROM view3;
```

Automatically Modify, INSERT/UPDATE/DELETE Rules

```
INSERT INTO tab1 VALUES (...);
INSERT INTO tab2 VALUES (...);

CREATE RULE insert_tab1 AS ON INSERT TO tab1 DO
INSERT INTO tab2 VALUES (...);
INSERT INTO tab1 VALUES (...);
```

Automatically Modify INSERT/UPDATE/DELETE Rules Example

```
CREATE TABLE service request
    customer id INTEGER,
    description text,
   cre user text DEFAULT CURRENT USER,
   cre timestamp timestamp DEFAULT CURRENT TIMESTAMP
);
CREATE TABLE service request log
    customer id INTEGER,
    description text,
    mod type char(1),
   mod user text DEFAULT CURRENT USER,
  mod timestamp timestamp DEFAULT CURRENT TIMESTAMP
);
```

Rules Example - Rule Definition

```
CREATE RULE service request update AS -- UPDATE rule
ON UPDATE TO service request
D0
    INSERT INTO service request log (customer id, description, mod type)
    VALUES (old.customer id, old.description, 'U');
CREATE RULE service request delete AS -- DELETE rule
ON DELETE TO service request
D0
    INSERT INTO service request log (customer id, description, mod type)
    VALUES (old.customer id, old.description, 'D');
```

Multi-User Consistency

- Atomic Changes
- Atomic Visibility
- Atomic Consistency
- Reliability

Notification, LISTEN/NOTIFY

```
signal()/kill()
LISTEN myevent;
NOTIFY myevent;
```

Application Walk-through



2. Functions and Triggers

Placing Code Into the Database: Server-Side Functions



Single-Parameter Built-In Functions/Operator

Two-Parameter Built-in Function/Operator

```
SELECT date mi('2003-05-20'::date, '2001-10-13'::date);
date mi
584
SELECT '2003-05-20'::date - '2001-10-13'::date:
?column?
584
psql \dfS
2ob/ [psq
```

Custom Server-Side Functions

- Create function
- Call function, manually or automatically

Compute Sales Tax

```
total = cost * 1.06;
INSERT ... VALUES ( ... total ... );
INSERT ... VALUES ( ... cost * 1.06, ... );
CREATE FUNCTION total(float)
RFTURNS float
AS 'SELECT $1 * 1.06;'
LANGUAGE 'sal':
INSERT ... VALUES ( ... total(cost) ... )
```

Convert Fahrenheit to Centigrade

```
cent = (faren - 32.0) * 5.0 / 9.0
INSERT ... VALUES ( ... cent ... )
INSERT ... VALUES ( ... (faren - 32.0) * 5.0 / 9.0, ... )
CREATE FUNCTION ftoc(float)
RFTURNS float
AS 'SELECT ($1 - 32.0) * 5.0 / 9.0;'
LANGUAGE 'sal':
INSERT ... VALUES ( ... ftoc(faren) ... )
```

Compute Shipping Cost

```
if cost < 2
    shipping = 3.00
else if cost < 4
    shipping = 5.00
else shipping = 6.00

INSERT ... VALUES ( ... cost + shipping ... );</pre>
```

Shipping Cost Function

```
CREATE FUNCTION shipping(numeric)
RETURNS numeric
AS 'SELECT CASE
        WHEN $1 < 2 THEN CAST(3.00 AS numeric(8,2))
        WHEN 1 \ge 2 AND 1 < 4 THEN CAST(5.00 AS numeric(8,2))
        WHEN $1 \ge 4 THEN CAST(6.00 AS numeric(8.2))
END: '
LANGUAGE 'sql';
INSERT ... VALUES ( ... cost + shipping(cost) ... );
```

String Processing — PL/pgSQL

```
CREATE FUNCTION spread(text)
RETURNS text
AS $$
DECLARE
   str text:
   ret text:
   i integer;
    len integer;
REGIN
   str := upper($1);
   ret := ''; -- start with zero lenath
   i := 1:
   len := length(str);
   WHILE i <= len LOOP
       ret := ret || substr(str, i, 1) || ' ';
      i := i + 1:
   END LOOP:
   RETURN ret;
END:
$$
LANGUAGE 'plpgsgl':
SELECT spread('Major Financial Report'):
                   spread
MAJOR ETNANCTAL REPORT
```

State Name Lookup SQL Language Function

```
SFLFCT name
FROM statename
WHERE code = 'AL':
CREATE FUNCTION getstatename(text)
RETURNS text
AS 'SELECT name
    FROM statename
   WHERE code = $1:'
LANGUAGE 'sql';
SELECT getstatename('AL');
```

State Name Lookup From String PL/pgSQL Language Function

```
CREATE FUNCTION getstatecode(text)
RETURNS text
AS $$
DECLARE
   state str statename.name%TYPE;
   statename rec record;
   i integer:
   len integer:
   matches record:
   search str text;
BEGIN
    state str := initcap($1); -- capitalization match column
    len := length(trim($1));
    i := 2:
    SELECT INTO statename rec * -- first try for an exact match
    FROM statename
    WHERE name = state str;
    IF FOUND
    THEN RETURN statename rec.code;
    END IF;
```

State Name Lookup From String PL/pgSQL Language Function (Cont.)

```
WHILE i <= len LOOP -- test 2,4,6,... chars for match
        search str = trim(substr(state str, 1, i)) || '%';
        SELECT INTO matches COUNT(*)
        FROM statename
        WHERE name LIKE search str;
        IF matches.count = 0 -- no matches, failure
        THEN RETURN NULL:
        END IF:
        IF matches.count = 1 -- exactly one match, return it
        THFN
           SELECT INTO statename rec *
           FROM statename
           WHERE name LIKE search str;
            IF FOUND
            THEN RETURN statename rec.code;
           END IF:
        END IF:
        i := i + 2: -- > 1 match. try 2 more chars
    END LOOP:
    RETURN '':
LANGUAGE 'plpgsql';
```

END; \$\$

State Name Maintenance

```
CREATE FUNCTION change_statename(char(2), char(30))
RETURNS boolean
AS $$
DECLARE
    state_code ALIAS FOR $1;
    state_name ALIAS FOR $2;
    statename rec RECORD;
```

State Name Maintenance (Cont.)

```
BEGIN
    IF length(state code) = 0 -- no state code, failure
    THEN RETURN 'f';
   ELSE
        IF length(state name) != 0 -- is INSERT or UPDATE?
        THEN
           SELECT INTO statename rec *
            FROM statename
           WHERE code = state code;
            IF NOT FOUND -- is state not in table?
            THEN INSERT INTO statename
                   VALUES (state code, state name);
                  UPDATE statename
           ELSE
                    SET name = state name
                    WHERE code = state code;
            END IF;
           RETURN 't';
```

State Name Maintenance (Cont.)

```
ELSE -- is DELETE
            SELECT INTO statename rec *
            FROM statename
            WHERE code = state code;
            TF FOUND
            THEN DELETE FROM statename
                    WHERE code = state code;
                    RETURN 't':
            ELSE RETURN 'f':
            END IF:
        END IF:
    END IF:
END;
$$
LANGUAGE 'plpgsql';
SELECT change statename('AL','Alabama');
SELECT change_statename('AL','Bermuda');
SELECT change statename('AL','');
SELECT change statename('AL',''); -- row was already deleted
```

SELECT Inside FROM

```
SFLFCT *
FROM (SELECT * FROM tab) AS tab;
SFLFCT *
FROM ( SELECT 1.2.3.4.5 UNION
       SELECT 6.7.8.9.10 UNION
      SELECT 11,12,13,14,15) AS tab15;
col| col| col| col| col
---+---+---+----
1 | 2 | 3 | 4 | 5
6 | 7 | 8 | 9 | 10
11 | 12 | 13 | 14 | 15
```

Function Returning Multiple Values

```
CREATE FUNCTION func5() RETURNS
TABLE (x1 INT, x2 INT, x3 INT, x4 INT, x5 INT) AS
'SELECT 1,2,3,4,5;'
LANGUAGE SQL;

SELECT * FROM func5();
x1 | x2 | x3 | x4 | x5
---+---+----+----
1 | 2 | 3 | 4 | 5
```

Function Returning a Table Result

```
CREATE FUNCTION func15() RETURNS
TABLE (x1 INT, x2 INT, x3 INT, x4 INT, x5 INT) AS
   SELECT 1,2,3,4,5 UNION
   SELECT 6,7,8,9,10 UNION
   SELECT 11.12.13.14.15: '
LANGUAGE SOL:
SELECT * FROM func15() ORDER BY x1;
x1 | x2 | x3 | x4 | x5
----+----
 1 | 2 | 3 | 4 | 5
 6 | 7 | 8 | 9 | 10
11 | 12 | 13 | 14 | 15
```

Automatic Function Calls Trigger

- BEFORE/AFTER ROW
- INSERT/UPDATE/DELETE
- OLD/NEW

Trigger on Statename

```
CREATE FUNCTION trigger insert update statename()
RETURNS trigger
AS $$
BEGIN
    IF new.code !~ '^[A-Za-z][A-Za-z]$'
    THEN RAISE EXCEPTION 'State code must be two alphabetic characters.';
    END IF;
    IF new.name !~ '^[A-Za-z ]*$'
    THEN RAISE EXCEPTION 'State name must be only alphabetic characters.':
    END IF:
    IF length(trim(new.name)) < 3
    THEN RAISE EXCEPTION 'State name must longer than two characters.':
    END IF:
    new.code = upper(new.code); -- uppercase statename.code
    new.name = initcap(new.name): -- capitalize statename.name
    RETURN new;
END:
$$
LANGUAGE 'plpgsql';
```

Install Trigger On Statename

```
CREATE TRIGGER trigger_statename
BEFORE INSERT OR UPDATE
ON statename
FOR EACH ROW
EXECUTE PROCEDURE trigger_insert_update_statename();
INSERT INTO statename VALUES ('a', 'alabama');
INSERT INTO statename VALUES ('al', 'alabama2');
INSERT INTO statename VALUES ('al', 'alabama2');
INSERT INTO statename VALUES ('al', 'alabama');
```

Function Languages

- SQL
- PL/pgSQL
- PL/Perl
- PL/Python
- PL/Java
- PL/TCL
- PL/sh
- (

Function Examples

- /contrib/earthdistance
- /contrib/fuzzystringmatch
- /contrib/pgcrypto



3. Customizing Database Features

Adding New
Data and Indexing
Features



Creation

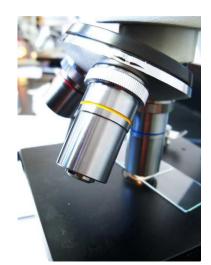
- CREATE FUNCTIONS in C
- CREATE TYPE
- CREATE OPERATOR
- CREATE OPERATOR CLASS (index type)

Create New Data Type With Operator and Index Support

- Write input/output functions
- Register input/output functions with CREATE FUNCTION
- Register type with CREATE TYPE
- Write comparison functions
- Register comparison functions with CREATE FUNCTION
- Register comparison functions with CREATE OPERATOR
- Register operator class for indexes with CREATE OPERATOR CLASS

Create New Data Type Examples

- /contrib/chkpass
- /contrib/isn
- /contrib/cube
- /contrib/ltree
- /src/backend/utils/adt



Conclusion





https://momjian.us/presentations