

PostgreSQL in Practice

Selena Deckelmann
Information Systems Manager
Chris King Precision Components
www.chrinking.com
pugs.postgresql.org/pdx

Why we chose PostgreSQL

Supported by our software vendor

Open source

Great user community

But...

Didn't know about commercial support options

Probably couldn't afford them anyway

Needed to learn more

pdxpug

OSCON 2006

Perl Mongers

Started a users group!

pugs.postgresql.org/pdx/

What I do

Information Systems Manager

Sysadmin, database admin, programmer

Support servers, desktops, embedded systems, CNC, timeclocks (yuck!), phones, router, switches, firewall, lots of custom applications

Database experience

MySQL

Filemaker Pro

Omnis (from Raining Data)

PostgreSQL

(I'm not a DBA)

Compiling/installing PostgreSQL

On Mac OS X -- this is pretty much it:

```
# ./configure --with-openssl --without-  
readline --with-pam --with-krb5
```

```
# make
```

```
# make install
```

```
# adduser postgres
```

What you just got

PAM authentication support

Kerberos support

OpenSSL support

Packaged installs

Fink

<http://developer.apple.com/internet/opensource/postgres.html>

Most Linux distributions have a binary (pre-compiled) package

Configuration

No one-size fits all solution

Good guideline:

[http://www.powerpostgresql.com/Downloads/
annotated_conf_80.html](http://www.powerpostgresql.com/Downloads/annotated_conf_80.html)

Configuration

Can control:

amount of memory per connection

amount of memory per query

size of disk cache for queries

`shared_buffers`, `work_mem`,

`maintenance_work_mem`,

`effective_cache_size`

Other neat things

ODBC (e.g. get data into Excel)

Client/sysadmin applications (pgAdmin, phpPgAdmin, psql)

Stored Procedures

PL/PgSQL, PL/Java, PL/Python and
(not sure why) PL/Tcl

An example

Enterprise Resource Planning

Work orders, Sales Orders, Invoices,
General Ledger, Purchasing, Inventory

Hundreds of tables

about 100 GB of data so far

Uses a lot of PL/PgSQL

Problem

Need to figure out how many hours we scheduled a “work center”

Workcenter: Bar Stock Saw

Operation: cutting 12' barstock down to 4' lengths

Takes 30 minutes to set up

Takes 10 minutes per cut

Performance Issue

Takes 5-10 minutes to display results

Planning uses query frequently

Many trips to coffee machine

So - wrote a test query that mimicked what the GUI was doing...

```

CREATE OR REPLACE FUNCTION woloadbyworkcenter2 (
    wrkcnt_ids integer[],
    cal_ids integer[]
) RETURNS SETOF numeric AS $$

DECLARE
    totaltime    numeric;
BEGIN
    FOR wloop IN 1..array_upper(wrkcnt_ids, 1) LOOP
        IF wrkcnt_ids[wloop] IS NULL THEN
            CONTINUE;
        END IF;

        FOR cloop IN 1..array_upper(cal_ids, 1) LOOP
            IF cal_ids[cloop] IS NULL THEN
                CONTINUE;
            END IF;

            SELECT INTO totaltime SUM(
                CASE WHEN (wooper_sucomplete) THEN 0
                     ELSE (wooper_sutime - wooper_suconsumed)
                END +
                CASE WHEN (wooper_rncomplete) THEN 0
                     ELSE (wooper_rntime - wooper_rnconsumed)
                END )
            FROM wooper
            WHERE ( (wooper_wrkcnt_id=wrkcnt_ids[wloop])
                AND (wooper_scheduled::DATE BETWEEN
                    findPeriodStart(cal_ids[cloop]) AND findPeriodEnd(cal_ids[cloop])) );

            RETURN NEXT totaltime;

        END LOOP;

    END LOOP;

END;
$$ LANGUAGE plpgsql;

```


Then...

EXPLAIN:

```
# EXPLAIN ANALYZE SELECT * from woloadbyworkcenter2(  
'{83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100}',  
'{135,136,137,138,139,140,141,142}');
```

QUERY PLAN

Function Scan on woloadbyworkcenter2

(cost=0.00..12.50 rows=1000 width=32)

(actual time=**21067.087**..21067.109 rows=144 loops=1)

Total runtime: 21067.171 ms

(2 rows)

OUCH!

Statistics tables

PostgreSQL profiling!

Turn stats gathering on in config

Then look at: `pg_stat_all_tables`

To sum up...

Calendar tables hit hard - 1000s of
Sequence Scans

Sequence Scans walk through tables
or indexes sequentially from beginning
to end - very slow!

Dug around in findPeriodStart &
findPeriodEnd functions..

EXCERPT

```
<snip>  
-- If weeks... (gotta be a better way)  
</snip>
```

Plus another 50 lines and six other loops...

Take this function out!

Reduce looping on the performance problem

Some logic was in the GUI

Focused on the main `SELECT` statement

Results:

```
# EXPLAIN ANALYZE SELECT * FROM woloadbyworkcenter2(  
'{83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100}',  
'{135,136,137,138,139,140,141,142}');
```

QUERY PLAN

```
-----  
Function Scan on woloadbyworkcenter2  
  (cost=0.00..12.50 rows=1000 width=32)  
    (actual time=477.910..477.931 rows=144  
      loops=1)  
Total runtime: 478.003 ms  
(2 rows)
```

(it was 21067.087 ms before!)

Resources

PostgreSQL, Douglas & Douglas

David Wheeler on PL/PgSQL

http://pugs.postgresql.org/pdx/archives/2006_07.html

Me on installing PostgreSQL

http://pugs.postgresql.org/pdx/archives/2006_08.html

Josh Berkus, Elein Mustain, Gabrielle Roth, OpenMFG

Thanks!

Selena Deckelmann

selenamarie@gmail.com

<http://pugs.postgresql.org/pdx/>