

# DTrace Probes in 8.4 Lightning Talk

**May 22, 2009**

**Robert Lor**  
**[robert.lor@sun.com](mailto:robert.lor@sun.com)**

# Probes in 8.2

- > transaction-start
- > transaction-commit
- > transaction-abort
- > lock-acquire
- > lwlock-release
- > lwlock-wait-start/done
- > lwlock-condacquire
- > lwlock-condacquire-fail
- > lock-wait-start/done

# Probes in 8.3

- > transaction-start
- > transaction-commit
- > transaction-abort
- > lock-acquire
- > lwlock-release
- > lwlock-wait-start/done
- > lwlock-condacquire
- > lwlock-condacquire-fail
- > lock-wait-start/done

# Probes in 8.4

- 
- > transaction-start
  - > transaction-commit
  - > transaction-abort
  - > query-start/done
  - > query-parse-start/done
  - > query-rewrite-start/done
  - > query-plan-start/done
  - > query-execute-start/done
  - > statement-status
  - > checkpoint-start/done
  - > clog-checkpoint-start/done
  - > subtrans-checkpoint-start/done
  - > multixact-checkpoint-start/done
  - > buffer-checkpoint-start/done
  - > twophase-checkpoint-start/done
  - > buffer-sync-start/done
  - > buffer-sync-written
  - > buffer-flush-start/done
  - > buffer-read-start/done
  - > buffer-write-dirty-start/done
  - > deadlock-found
  - > sort-start/done
  - > lwlock-acquire
  - > lwlock-release
  - > lwlock-wait-start/done
  - > lwlock-condacquire
  - > lwlock-condacquire-fail
  - > lock-wait-start/done
  - > ...

# What can the probes do?

- Time spent in query processing
- Transaction processing
- Access to a relation
- Monitor buffer usage (read, write, hit, miss)
- Checkpoint statistics
- Determine type and time spent for sort
- Monitor lock contention

# Can I use on my OS?

- Now
  - > OpenSolaris / Solaris
  - > OS X Leopard
- Future
  - > FreeBSD
  - > Linux (Systemtap / DTrace???)

# PostgreSQL DTrace Toolkit

PgFoundry: DTrace Probes: Project Filelist

http://pgfoundry.org/frs/?group\_id=10001


pgFoundry This project's trackers Search Advanced search Log In New Account

Home My Page Project Tree Project Openings **DTrace Probes**

Summary Forums Tracker Lists Tasks Docs News SCM **Files**

[Files](#) | [Admin](#)

Below is a list of all files of the project. The release you have chosen is **highlighted**. Before downloading, you may want to read Release Notes and ChangeLog (accessible by clicking on release version).

Package	Release & Notes	Filename	Size	Date		
				D/L	Arch	Type
PostgreSQL DTrace Toolkit 						
<b>DTrace Toolkit 2009.03.29</b>			2009-03-30 04:07			
<a href="#">PostgreSQL-DTrace-Toolkit-2009.03.29.tar.gz</a>			3.84 MB	71	i386	.gz
<b>Project totals</b>		<b>1</b>	<b>3.84 MB</b>	<b>71</b>		

Find:  Next Previous  Highlight all  Match case

Done

# DTrace Script

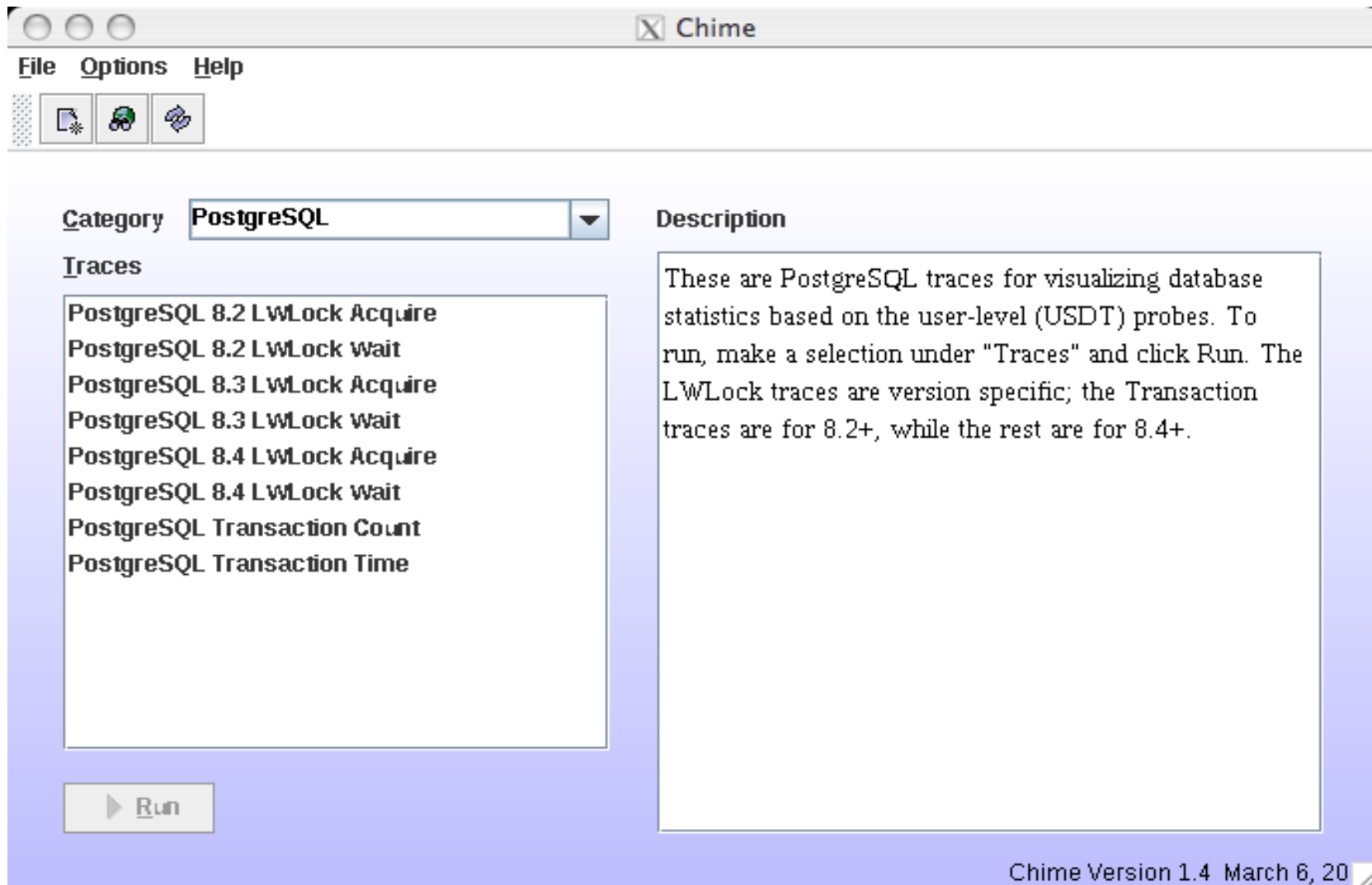
*What is this cryptic thing?*

```
#!/usr/sbin/dtrace -qs  
postgresql*:::query-start  
{  
    self->query = copyinstr(arg0);  
    self->ts = timestamp;  
}
```

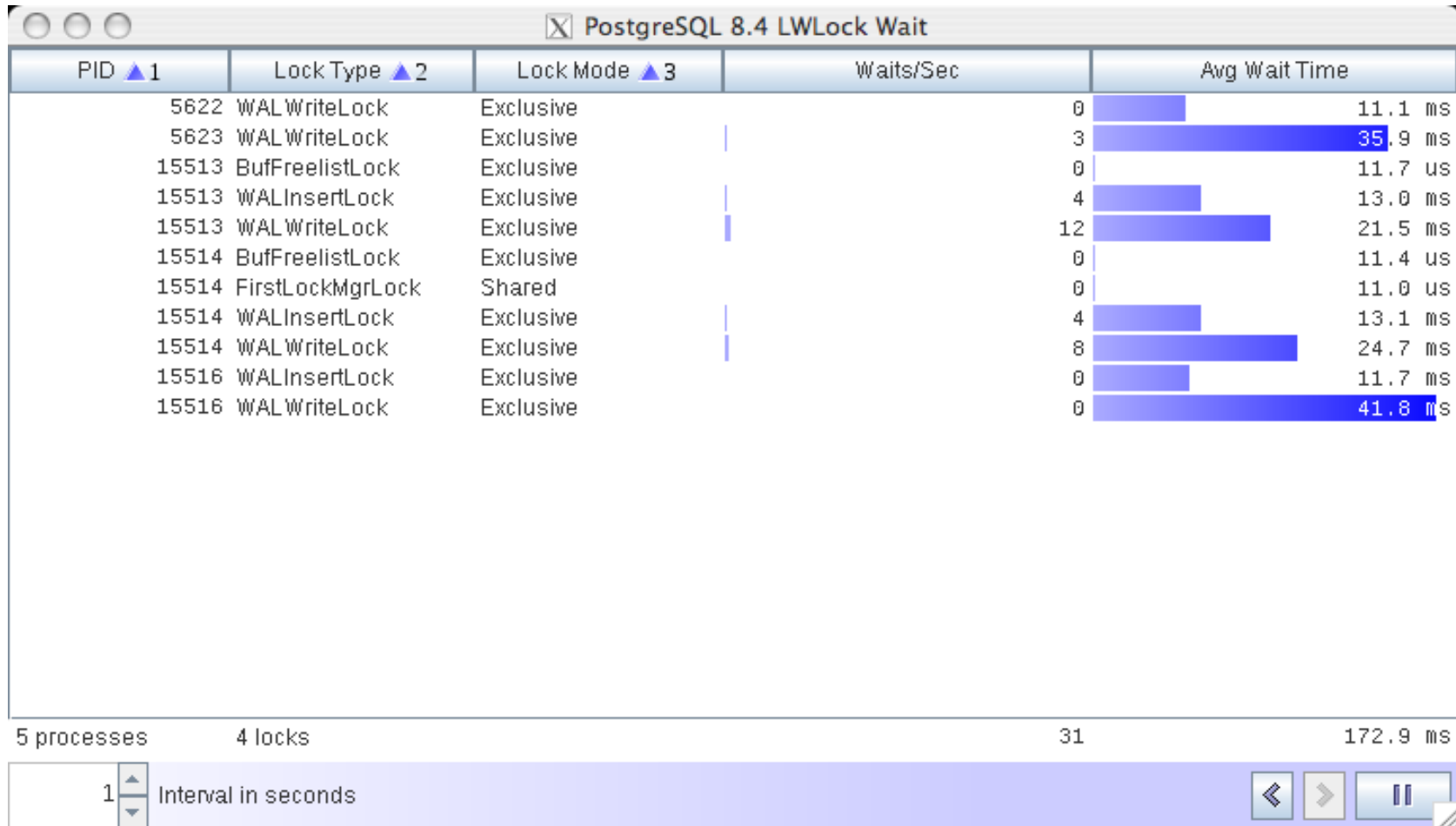




# GUI Tool



# LWLock Stats





# Demo

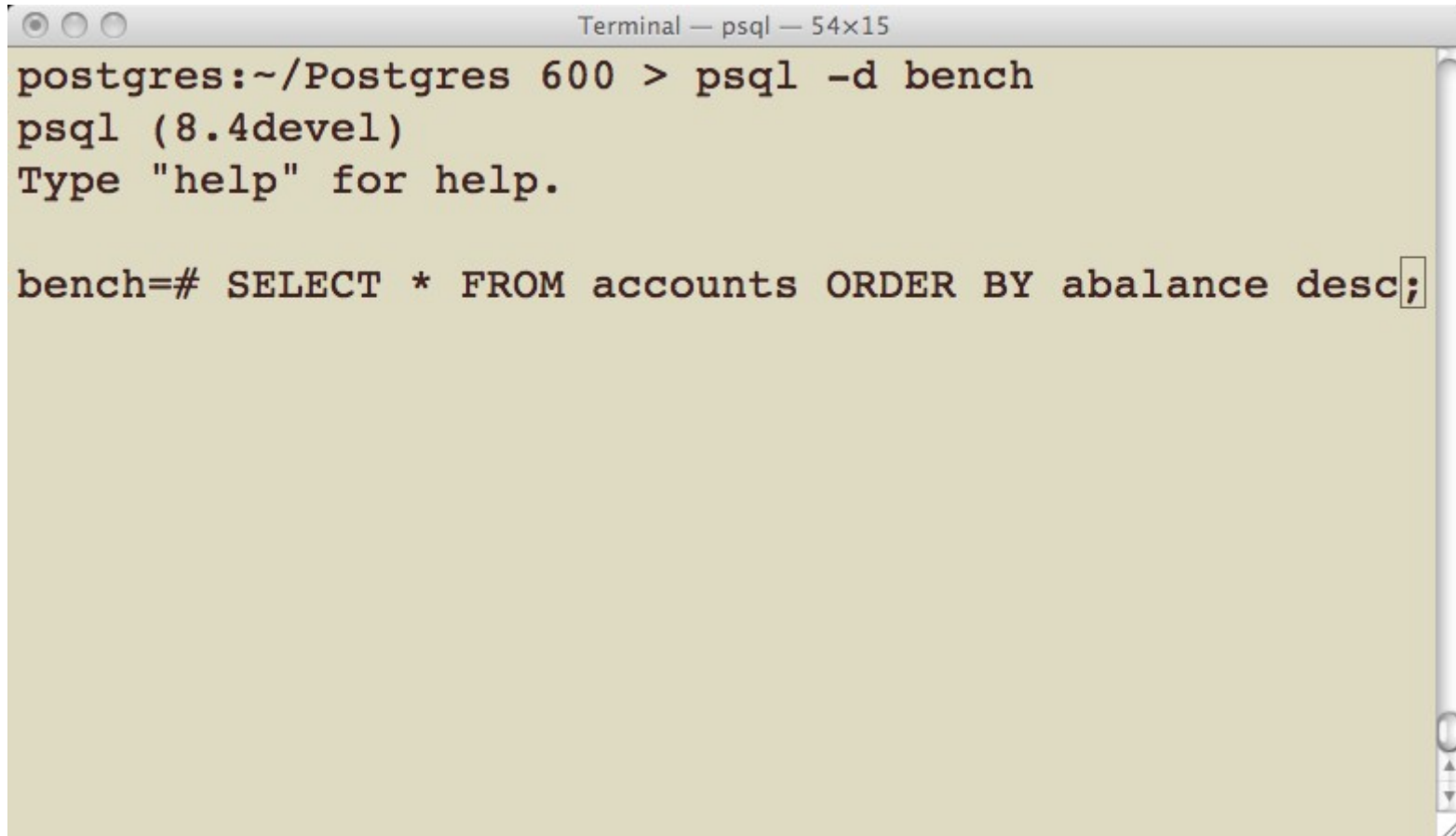
# Slow Query #1

# Identify Slow Query

```
Terminal — bash — 62x15
bash-3.2# ./query_time.d
Tracing... Hit Ctrl-C to end.
^C
  TIME (ns) QUERY
=====
3240393069 SELECT * FROM accounts ORDER BY abalance desc;

bash-3.2#
```

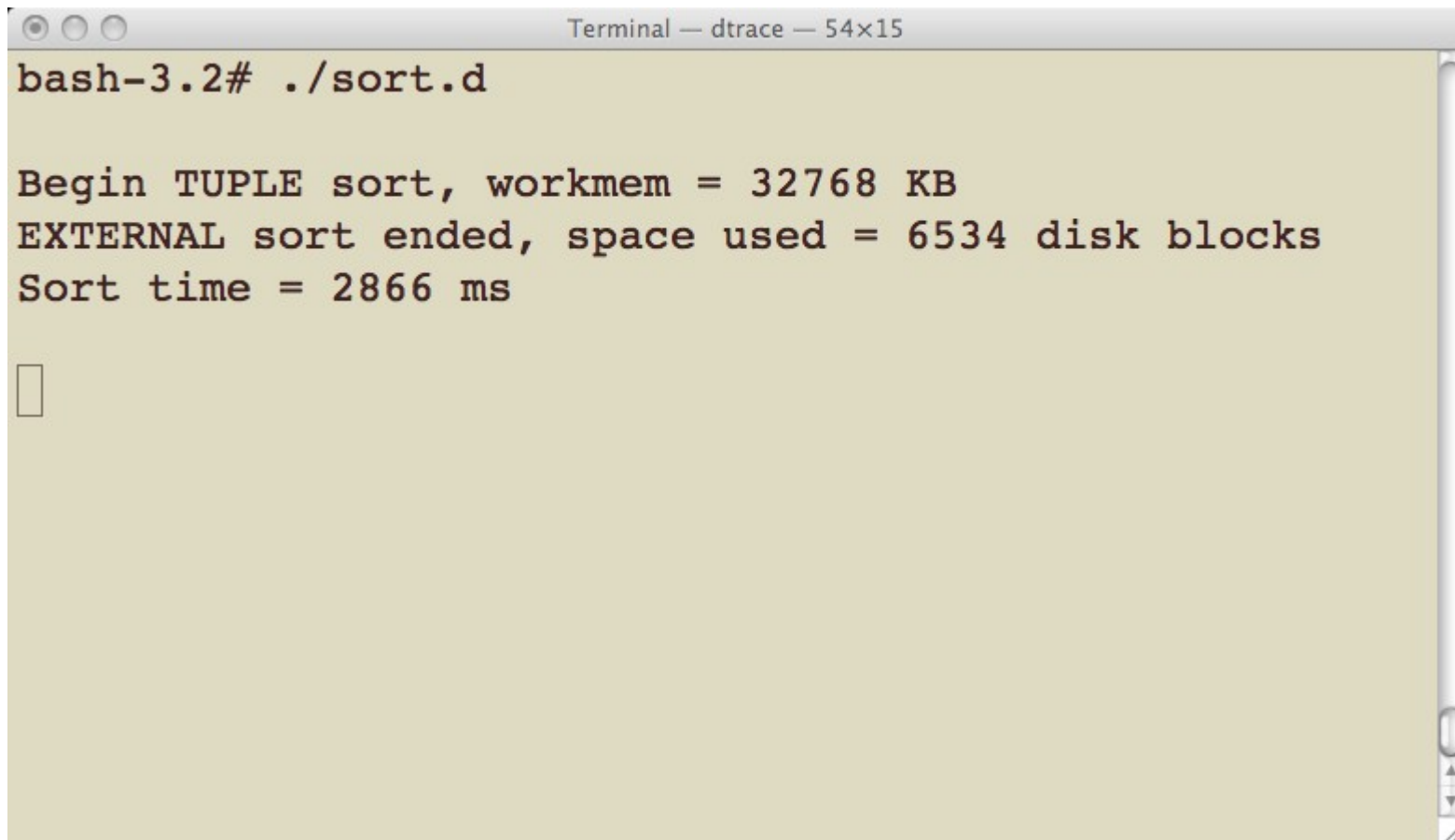
# Drill Down

A terminal window titled "Terminal — psql — 54x15" with a light beige background. The window shows the following text: "postgres:~/Postgres 600 > psql -d bench", "psql (8.4devel)", "Type \"help\" for help.", and "bench=# SELECT \* FROM accounts ORDER BY abalance desc;". The cursor is positioned at the end of the last line. The window has standard macOS window controls (red, yellow, green buttons) in the top-left corner and a vertical scrollbar on the right side.

```
postgres:~/Postgres 600 > psql -d bench
psql (8.4devel)
Type "help" for help.

bench=# SELECT * FROM accounts ORDER BY abalance desc;
```

# Dtrace Output



```
Terminal — dtrace — 54x15
bash-3.2# ./sort.d

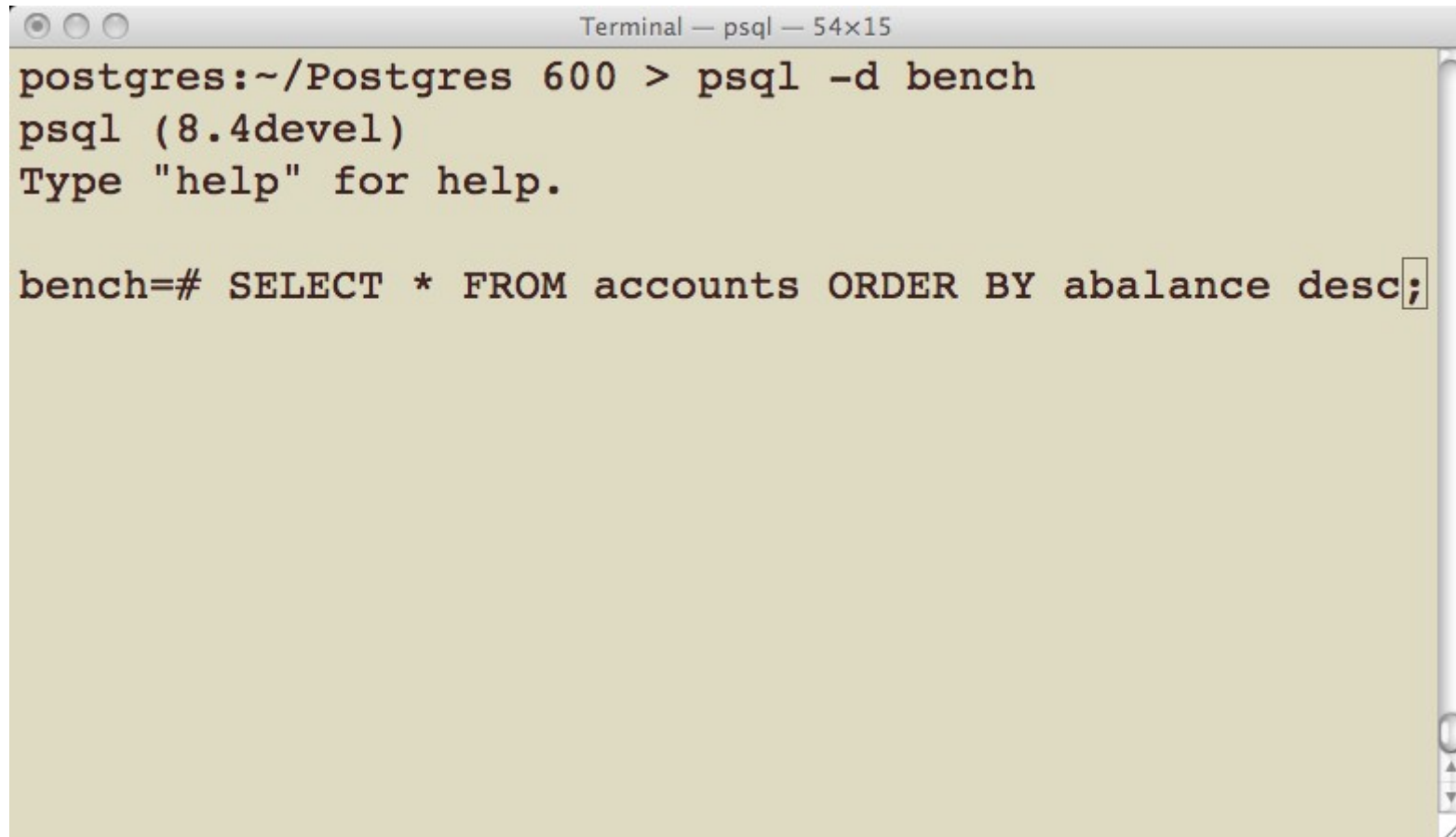
Begin TUPLE sort, workmem = 32768 KB
EXTERNAL sort ended, space used = 6534 disk blocks
Sort time = 2866 ms

□
```



**Increase work\_mem**

# Rerun Query

A terminal window titled "Terminal — psql — 54x15" with three window control buttons in the top-left corner. The terminal text shows a user at a postgres prompt running 'psql -d bench', which opens a psql prompt. The user then enters a SQL query to select all records from the 'accounts' table, ordered by 'abalance' in descending order.

```
postgres:~/Postgres 600 > psql -d bench
psql (8.4devel)
Type "help" for help.

bench=# SELECT * FROM accounts ORDER BY abalance desc;
```

# Dtrace Output



```
Terminal — dtrace — 54x15
bash-3.2# ./sort.d

Begin TUPLE sort, workmem = 98304 KB
INTERNAL sort ended, space used = 74599 KB
Sort time = 1536 ms

□
```

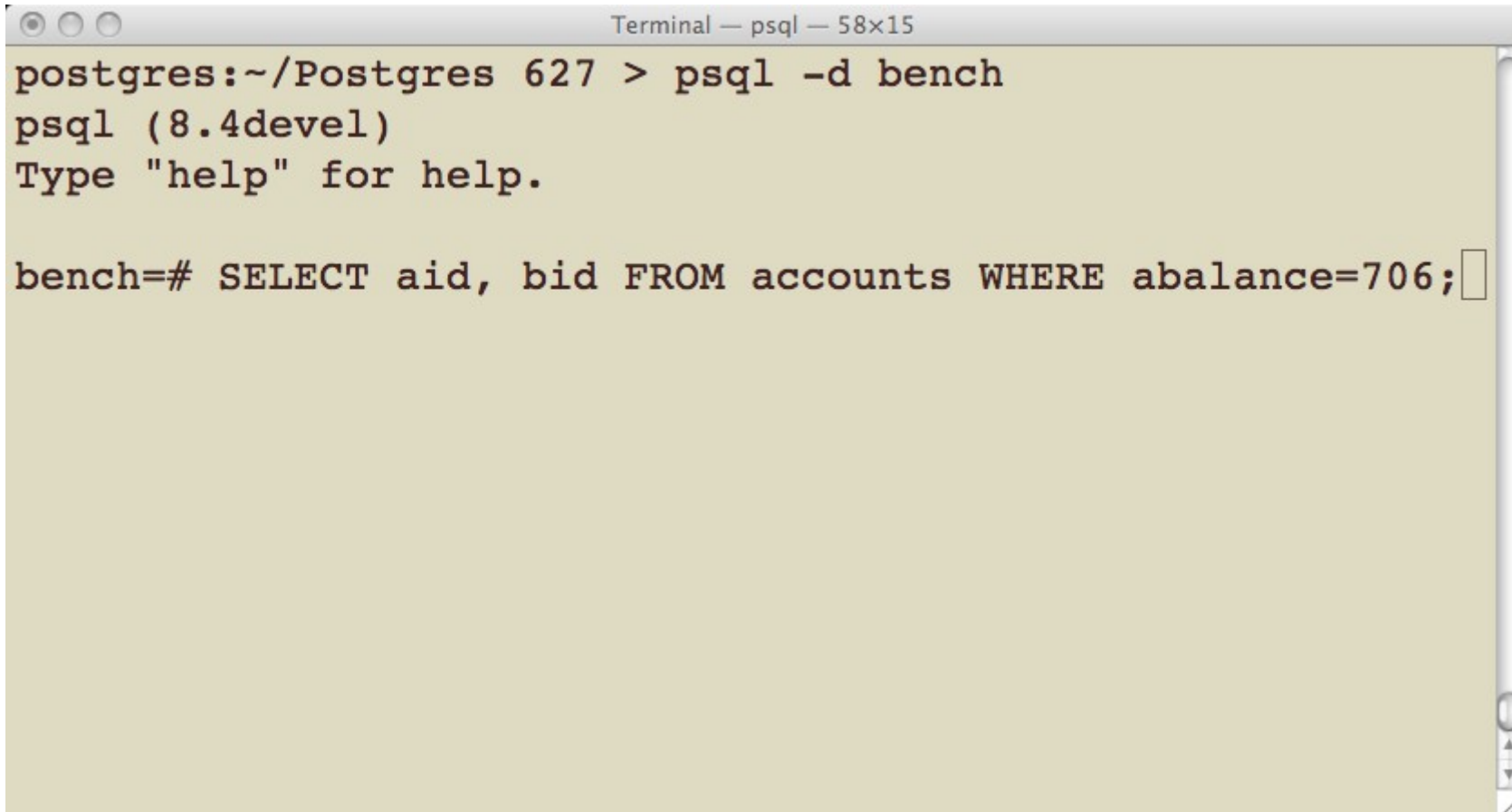
# Slow Query #2

# Identify Slow Query

```
Terminal — bash — 78x14
bash-3.2# ./query_time.d
Tracing... Hit Ctrl-C to end.
^C
TIME (ns) QUERY
=====
 97689 SELECT abalance FROM accounts WHERE aid = 100274;
153096 SELECT abalance FROM accounts WHERE aid = 100160;
309538 SELECT abalance FROM accounts WHERE aid = 100129;
613759 UPDATE tellers SET tbalance = tbalance + 997 WHERE tid = 35;
642923 UPDATE tellers SET tbalance = tbalance + -1002 WHERE tid = 23;
1888027 UPDATE tellers SET tbalance = tbalance + -100 WHERE tid = 16;
205590467 SELECT aid, bid FROM accounts WHERE abalance=706;

bash-3.2#
```

# Drill Down

A terminal window titled "Terminal — psql — 58x15" with a light beige background. The prompt is "postgres:~/Postgres 627 >". The user has entered "psql -d bench", and the prompt has changed to "psql (8.4devel)". Below that, the instruction "Type 'help' for help." is shown. The user has entered the SQL query "bench=# SELECT aid, bid FROM accounts WHERE abalance=706;" and a cursor is visible at the end of the line. The terminal has standard window controls (three circles) at the top left and a scroll bar on the right side.

```
postgres:~/Postgres 627 > psql -d bench
psql (8.4devel)
Type "help" for help.

bench=# SELECT aid, bid FROM accounts WHERE abalance=706;
```

# Dtrace output

```

Terminal — vim — 66x24
bash-3.2# ./query_trace.d

===== Buffer Read Counts =====
Tablespace      Database      Table      Count
      1663      16384      2610         1
      1663      16384      2678         1
      1663      16384      1249         4
      1663      16384      2659         8
      1663      16384      16397       8239

===== Buffer Write Request Counts =====
Tablespace      Database      Table      Count

===== Additional Statistics =====
Backend PID      : 40203
SQL Statement    : SELECT aid, bid FROM accounts WHERE abalance=706;
Execution time   : 0.357 sec (357409894 ns)

Total buffer cache hits      :      329
Total buffer cache misses    :      7937
Average read time from cache :    12637 (ns)
Average read time from disk  :    23348 (ns)

```

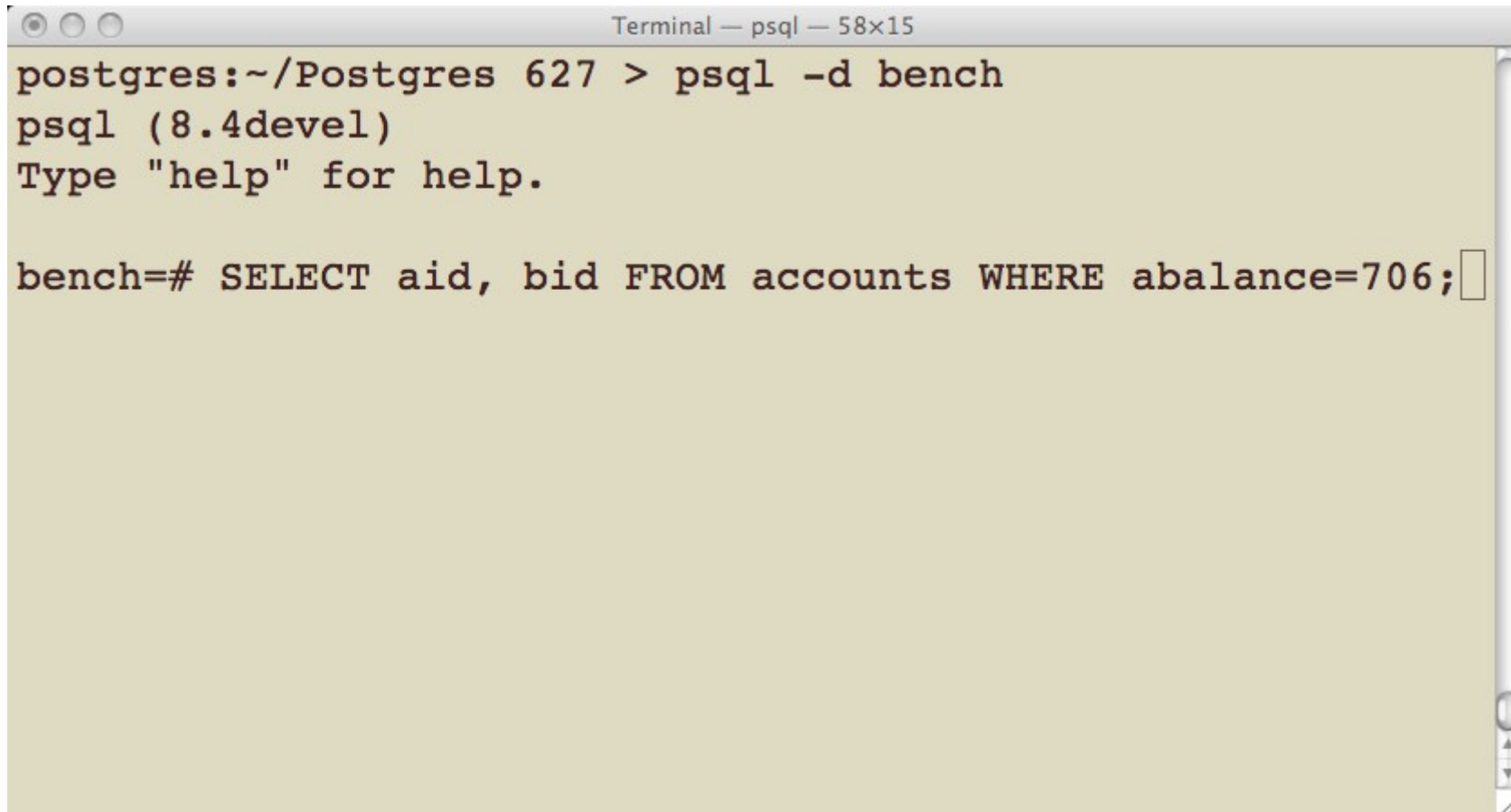
# Check for Index

```
Terminal — psql — 58x15
bench=# \d accounts;
          Table "public.accounts"
  Column  |      Type      | Modifiers
-----+-----+-----
 aid     | integer        | not null
 bid     | integer
 abalance | integer
 filler  | character(84)
Indexes:
    "accounts_pkey" PRIMARY KEY, btree (aid)

bench=# CREATE INDEX accounts_balance ON accounts(abalance
);
CREATE INDEX
bench=#
```



# Rerun Query

A terminal window titled "Terminal — psql — 58x15" with a light beige background. The prompt is "postgres:~/Postgres 627 >". The user has entered "psql -d bench", and the prompt has changed to "psql (8.4devel)". Below that, it says "Type 'help' for help." The user has entered "bench=# SELECT aid, bid FROM accounts WHERE abalance=706;" and the cursor is at the end of the line.

```
postgres:~/Postgres 627 > psql -d bench
psql (8.4devel)
Type "help" for help.

bench=# SELECT aid, bid FROM accounts WHERE abalance=706;
```

# Dtrace Output

```

Terminal — vim — 66x24
bash-3.2# ./query_trace.d

===== Buffer Read Counts =====
Tablespace      Database      Table         Count
    1663         16384        1255           1
    1663         16384        2678           1
    1663         16384        2617           2
    1663         16384        1259           5
    1663         16384        1249           6
    1663         16384        2659          12
    1663         16384        16397          29

===== Buffer Write Request Counts =====
Tablespace      Database      Table         Count

===== Additional Statistics =====
Backend PID      : 40203
SQL Statement    : SELECT aid, bid FROM accounts WHERE abalance=706;
Execution time   : 0.005 sec (5818620 ns)

Total buffer cache hits      :          52
Total buffer cache misses    :          31
Average read time from cache :      37253 (ns)
  
```



# For More Information

- PostgreSQL 8.4 Probes  
<http://developer.postgresql.org/pgdocs/postgres/dynamic-trace.html>
- PostgreSQL DTrace Toolkit  
<http://pgfoundry.org/projects/dtrace/>
- DTrace toolkit contains a lot of very useful scripts  
<http://www.opensolaris.org/os/community/dtrace/dtracetoolkit/>
- Solaris DTrace doc  
<http://wikis.sun.com/display/DTrace/Documentation>
- Open Solaris DTrace community page  
<http://www.opensolaris.org/os/community/dtrace/>