

# What is going on?

PostgreSQL Conference Fall 2007  
October 20, 2007

Mark Wong  
markwkm@gmail.com

# Overview

Use commonly available system tools and PostgreSQL's system catalog tables to examine how the system is being utilized.

# At first glance, general systems statistics

What tools do you like to use?

# Suggestions for some tools

- top
- iostat
- ps

# top

\$ top

top - 13:21:07 up 6 days, 16:03, 6 users, load average: 1.36, 1.19, 2.38

Tasks: 140 total, 3 running, 137 sleeping, 0 stopped, 0 zombie

Cpu0 : 14.0%us, 5.0%sy, 0.0%ni, 51.0%id, 29.7%wa, 0.0%hi, 0.3%si, 0.0%st

Cpu1 : 11.0%us, 5.0%sy, 0.0%ni, 43.5%id, 40.2%wa, 0.0%hi, 0.3%si, 0.0%st

Cpu2 : 14.7%us, 4.3%sy, 0.0%ni, 52.0%id, 28.7%wa, 0.0%hi, 0.3%si, 0.0%st

Cpu3 : 12.0%us, 5.0%sy, 0.0%ni, 43.1%id, 39.8%wa, 0.0%hi, 0.0%si, 0.0%st

Mem: 4084736k total, 4069248k used, 15488k free, 97984k buffers

Swap: 8388480k total, 23808k used, 8364672k free, 3739776k cached

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
10994	markwkm	15	0	41536	33m	30m	R	47	0.8	0:13.79	postgres
11021	markwkm	16	0	43904	36m	30m	R	21	0.9	0:06.46	postgres
11141	markwkm	16	0	39744	31m	29m	D	1	0.8	0:00.21	postgres
30748	markwkm	15	0	5632	3968	2624	R	1	0.1	0:29.16	top
374	root	10	-5	0	0	0	D	0	0.0	0:25.57	kwindfarm

# What I look for

```
$ top
top - 13:21:07 up 6 days, 16:03, 6 users, load average: 1.36, 1.19, 2.38
Tasks: 140 total, 3 running, 137 sleeping, 0 stopped, 0 zombie
Cpu0  : 14.0%us, 5.0%sy, 0.0%ni, 51.0%id, 29.7%wa, 0.0%hi, 0.3%si, 0.0%st
Cpu1  : 11.0%us, 5.0%sy, 0.0%ni, 43.5%id, 40.2%wa, 0.0%hi, 0.3%si, 0.0%st
Cpu2  : 14.7%us, 4.3%sy, 0.0%ni, 52.0%id, 28.7%wa, 0.0%hi, 0.3%si, 0.0%st
Cpu3  : 12.0%us, 5.0%sy, 0.0%ni, 43.1%id, 39.8%wa, 0.0%hi, 0.0%si, 0.0%st
Mem:   4084736k total, 4069248k used, 15488k free, 97984k buffers
Swap: 8388480k total, 23808k used, 8364672k free, 3739776k cached
```

PID	USER	PR	NI	VIRT	RES	SHR	S	%CPU	%MEM	TIME+	COMMAND
10994	markwkm	15	0	41536	33m	30m	R	47	0.8	0:13.79	postgres
11021	markwkm	16	0	43904	36m	30m	R	21	0.9	0:06.46	postgres
11141	markwkm	16	0	39744	31m	29m	D	1	0.8	0:00.21	postgres
30748	markwkm	15	0	5632	3968	2624	R	1	0.1	0:29.16	top
374	root	10	-5	0	0	0	D	0	0.0	0:25.57	kwindfarm

# ps

```
$ ps -aef | head -n 1; ps -aef | grep postgres | grep -v grep
```

UID	PID	PPID	C	STIME	TTY	TIME	CMD
markwkm	10594	1	0	13:20	pts/7	00:00:00	/home/markwkm/local/pgsql-8.2.3/bin/postgres
markwkm	10597	10594	0	13:20	?	00:00:00	postgres: writer process
markwkm	10598	10594	0	13:20	?	00:00:00	postgres: stats collector process
markwkm	10994	10594	51	13:20	?	00:00:14	postgres: markwkm dbt3 [local] SELECT
markwkm	11021	10594	24	13:20	?	00:00:07	postgres: markwkm dbt3 [local] SELECT
markwkm	11141	10594	0	13:20	?	00:00:00	postgres: markwkm dbt3 [local] INSERT

# Found 10994 again

```
$ ps -aef | head -n 1; ps -aef | grep postgres | grep -v grep
```

UID	PID	PPID	C	STIME	TTY	TIME	CMD
markwkm	10594	1	0	13:20	pts/7	00:00:00	/home/markwkm/local/pgsql-8.2.3/bin/postgres
markwkm	10597	10594	0	13:20	?	00:00:00	postgres: writer process
markwkm	10598	10594	0	13:20	?	00:00:00	postgres: stats collector process
markwkm	10994	10594	51	13:20	?	00:00:14	postgres: markwkm dbt3 [local] <b>SELECT</b>
markwkm	11021	10594	24	13:20	?	00:00:07	postgres: markwkm dbt3 [local] SELECT
markwkm	11141	10594	0	13:20	?	00:00:00	postgres: markwkm dbt3 [local] INSERT



# iostat

```
$ iostat -x 60 # 60 second sample length
```

```
Linux 2.6.22-gentoo-r6 (fungcho)      09/14/07
```

Device:	rrqm/s	wrqm/s	r/s	w/s	rsec/s	wsec/s	avgrq-sz	avgqu-sz	await	svctm	%util
sda	31.79	3.45	115.86	3.07	16154.79	55.94	136.30	1.82	15.28	7.84	93.23
...											
dm-82	0.00	0.00	147.48	5.67	16152.24	45.33	105.77	2.16	14.14	6.08	93.18

Device:	rrqm/s	wrqm/s	r/s	w/s	rsec/s	wsec/s	avgrq-sz	avgqu-sz	await	svctm	%util
sda	8.63	512.63	26.00	30.73	3823.56	14307.15	319.63	24.81	387.21	7.44	42.18
...											
dm-82	0.00	0.00	34.36	546.64	3822.83	15421.70	33.12	399.34	616.36	0.72	42.08

# Columns I look at

```
$ iostat -x 60 # 60 second sample length
```

```
Linux 2.6.22-gentoo-r6 (fungcho)          09/14/07
```

Device:	rrqm/s	wrqm/s	r/s	w/s	rsec/s	wsec/s	avgrq-sz	avgqu-sz	await	svctm	%util
sda	31.79	3.45	115.86	3.07	16154.79	55.94	136.30	1.82	15.28	7.84	93.23
...											
dm-82	0.00	0.00	147.48	5.67	16152.24	45.33	105.77	2.16	14.14	6.08	93.18
Device:	rrqm/s	wrqm/s	r/s	w/s	rsec/s	wsec/s	avgrq-sz	avgqu-sz	await	svctm	%util
sda	8.63	512.63	26.00	30.73	3823.56	14307.15	319.63	24.81	387.21	7.44	42.18
...											
dm-82	0.00	0.00	34.36	546.64	3822.83	15421.70	33.12	399.34	616.36	0.72	42.08

# Diving a little deeper, looking inside the database

Look into the system catalog tables to see  
what those processes are doing.

# What is the database doing?

```
$ psql -c "  
SELECT datname, procpid, current_query  
FROM pg_stat_activity  
ORDER BY procpid;  
"
```

datname	procpid	current_query
dbt3	10994	select sum(l_extendedprice * l_discount) as revenue from
dbt3	11021	select s_name, count(*) as numwait from supplier, lineite
dbt3	11141	insert into lineitem (select * from tmp_lineitem2);
dbt3	11155	: SELECT datname, procpid, current_query : FROM pg_stat_activity : ORDER BY procpid; :

(4 rows)

# Point out PID 10994

```
$ psql -c "  
SELECT datname, procpid, current_query  
FROM pg_stat_activity  
ORDER BY procpid;  
"
```

```
datname | procpid | current_query  
-----+-----+-----  
dbt3    | 10994   | select sum(l_extendedprice * l_discount) as revenue from  
dbt3    | 11021   | select s_name, count(*) as numwait from supplier, lineite  
dbt3    | 11141   | insert into lineitem (select * from tmp_lineitem2);  
dbt3    | 11155   |  
: SELECT datname, procpid, current_query  
: FROM pg_stat_activity  
: ORDER BY procpid;  
:  
(4 rows)
```

# What are the effects these queries have on the database

- What tables are being accessed?
- What indexes are being used?
- Any locks waiting to be acquired?

# Some table statistics

```
$ psql -c "  
SELECT relname, seq_scan, seq_tup_read, n_tup_ins, n_tup_upd, n_tup_del  
FROM pg_stat_user_tables  
ORDER BY relname;  
"
```

relname	seq_scan	seq_tup_read	n_tup_ins	n_tup_upd	n_tup_del
customer	64	4800000	0	0	0
lineitem	46	144519980	17968	0	0
nation	1181765	29542875	0	0	0
orders	32	24084000	4501	0	0
part	63	6200000	0	0	0
partsupp	9	4000000	0	0	0
region	10349	51665	0	0	0
supplier	74	290000	0	0	0
time_statistics	105	9224	105	105	0

(9 rows)

# Some index statistics

```
$ psql -c "  
SELECT relname, indexrelname, idx_scan, idx_tup_read, idx_tup_fetch  
FROM pg_stat_user_indexes  
ORDER BY relname, indexrelname;  
"
```

relname	indexrelname	idx_scan	idx_tup_read	idx_tup_fetch
customer	i_c_nationkey	0	0	0
customer	pk_customer	4	600000	600000
lineitem	i_l_commitdate	0	0	0
lineitem	i_l_orderkey	1388565	26191210	26188763
lineitem	i_l_orderkey_quantity	0	0	0
lineitem	i_l_partkey	32033	981915	187059
lineitem	i_l_receiptdate	4	3648680	0
lineitem	i_l_shipdate	20	13120485	0
lineitem	i_l_suppkey	1552	934560	0
lineitem	i_l_suppkey_partkey	193876	1459950	1459669
lineitem	pk_lineitem	0	0	0
nation	i_n_regionkey	0	0	0
nation	pk_nation	19984	19984	19984
orders	i_o_custkey	76377	50962	50962
orders	i_o_orderdate	16	3208273	0
...				



# Waiting for locks?

```
$ psql -c "  
SELECT pid, mode, current_query  
FROM pg_locks, pg_stat_activity  
WHERE granted = false  
      AND locktype = 'transactionid'  
      AND pid = procpid  
ORDER BY pid, granted;  
"
```

pid	mode	current_query
2047	ShareLock	update time_statistics set int_time = 3 where task_name =

(1 row)

# Correlating statistics back to i/o statistics

# Where is this index?

```
$ psql -c '\d i_l_commitdate'
```

```
Index "public.i_l_commitdate"  
  Column | Type  
-----+-----  
  l_commitdate | date  
btree, for table "public.lineitem"  
Tablespace: "dbt3_i_l_commitdate"
```

```
$ psql -c '\db dbt3_i_l_commitdate'
```

```
                List of tablespaces  
      Name      | Owner | Location  
-----+-----+-----  
  dbt3_i_l_commitdate | markwkm | /mnt/dbt3/i_l_commitdate/ts  
(1 row)
```

# What device is this tablespace on?

```
$ psql -c '\db dbt3_i_l_commitdate'
```

```
                List of tablespaces
      Name          | Owner  | Location
-----+-----+-----
 dbt3_i_l_commitdate | markwkm | /mnt/dbt3/i_l_commitdate/ts
(1 row)
```

```
$ df -h /mnt/dbt3/i_l_commitdate/ts
```

```
Filesystem          Size  Used Avail Use% Mounted on
/dev/mapper/vg-dbt3 38G   7.8G   28G   22% /mnt/dbt3
```

# Really, what device is the tablespace on?

```
$ df -h /mnt/dbt3/i_l_commitdate/ts
```

Filesystem	Size	Used	Avail	Use%	Mounted on
/dev/mapper/vg-dbt3	38G	7.8G	28G	22%	/mnt/dbt3

```
$ vgdisplay -v vg
```

```
--- Volume group ---
VG Name                vg
...
--- Physical volumes ---
PV Name                /dev/sda5
PV UUID                GP7m9E-1CNr-6ilF-RfNB-76wV-AFS4-Dm8nxE
PV Status              allocatable
Total PE / Free PE    57314 / 0
```

# Is iostat more helpful?

```
$ iostat -x 60 # 60 second sample length
```

```
Linux 2.6.22-gentoo-r6 (fungcho)          09/14/07
```

Device:	rrqm/s	wrqm/s	r/s	w/s	rsec/s	wsec/s	avgrq-sz	avgqu-sz	await	svctm	%util
sda	31.79	3.45	115.86	3.07	16154.79	55.94	136.30	1.82	15.28	7.84	93.23
...											
dm-82	0.00	0.00	147.48	5.67	16152.24	45.33	105.77	2.16	14.14	6.08	93.18

  

Device:	rrqm/s	wrqm/s	r/s	w/s	rsec/s	wsec/s	avgrq-sz	avgqu-sz	await	svctm	%util
sda	8.63	512.63	26.00	30.73	3823.56	14307.15	319.63	24.81	387.21	7.44	42.18
...											
dm-82	0.00	0.00	34.36	546.64	3822.83	15421.70	33.12	399.34	616.36	0.72	42.08

# Remember

- Make sure the following parameters are enabled:
  - stats\_start\_collector
  - stats\_command\_string
  - stats\_block\_level
  - stats\_row\_level