```
[oracle@localhost ~]$ lsb_release -a
               :core-3.1-ia32:core-3.1-noarch:graphics-3.1-ia32:graphics-3.1-noarch
Distributor ID: EnterpriseEnterpriseServer
Description:
               Enterprise Linux Enterprise Linux Server release 5.5 (Carthage)
Release:
                   5.5
Codename:
                  Carthage
   SQL> select * from v$version where rownum<2;
BANNER
Oracle Database 11g Enterprise Edition Release 11.2.0.1.0 - Production
SQL> create table t(id number, name varchar2(10));
    表已创建。
已用时间: 00:00:00.15
SQL> set timing off;
SQL> show user;
USER 为 "HR"
SQL> insert into t values(1, 'a');
已创建 1 行。
SQL> insert into t values(2, 'b');
已创建 1 行。
SQL> commit;
提交完成。
   SQL> update t set name='c' where id=1;
已更新 1 行。
SQL> select * from t;
           ID NAME
             1 c
             2 b
```

注意没提交。

重新打开一个 session:

下面通过 undo 的一致性读分析 undo:

SQL> select * from t;

ID NAME

1 a

2 b

此时还是读取到修改之前的数据,这里的 a 是重 undo 里读取的,下面 dump 分析这个过程: SQL> select t.*, rowid from t;

ID NAME	ROWID
1 c	AAASunAAEAAABuuAAA
2 b	AAASunAAEAAABuuAAB

我们利用 oracle 提供的包,可以获得第一条数据所在的数据文件号及块号:

SQL> show user;

USER 为 "SYS"

SQL> select dbms_rowid.rowid_relative_fno(rowid) fno, dbms_rowid.rowid_block_number(rowid) bno from hr.t;

FNO	BNO
4	7086
4	7086

此时我们可以 dump 4号文件的第7086块:

SQL> alter system dump datafile 4 block 7086;

系统已更改。

下面是摘自部分转储文件:

*** 2015-04-20 15:26:34.151

Block header dump: 0x01001bae

Object id on Block? Y

seg/obj: 0x12ba7 csc: 0x00.6658e2 itc: 2 flg: E typ: 1 - DATA

brn: 0 bdba: 0x1001ba8 ver: 0x01 opc: 0

inc: 0 exflg: 0

Itl Xid Uba Flag Lck Scn

/Fsc

```
0x00c009ac.02f0.24 C--- 0 scn 0x0000.0066582c
0x01
      0x0003.017.0000098a
                         0x00c02e17.0258.28 ---- (表示事务锁定)
0x02
      0x0008.00d.00000a84
                                                                 1(锁定一条数
据)
     fsc 0x0000.00000000
bdba:
0x01001bae
data_block_dump, data header at 0xa18264
tsiz: 0x1f98
hsiz: 0x16
pbl: 0x00a18264
      76543210
flag=----
ntab=1
nrow=2
frre=-1
fsbo=0x16
fseo=0x1f88
avsp=0x1f70
tosp=0x1f70
0xe:pti[0] nrow=2 offs=0
0x12:pri[0] offs=0x1f90
0x14:pri[1] offs=0x1f88
block_row_dump:
tab 0, row 0, @0x1f90
tl: 8 fb: --H-FL-- lb (锁标记): 0x2 cc: 2
                                                 第一行第二列已经被上锁了,被一个事
务锁定。 lb:lock byte
col 0: [2] c1 02
                                                =>这列是从 4 号文件 7086 块中读取的
col 1: [1] 63
                                                   63 代表是 c,实际已经被改了
tab 0, row 1, @0x1f88
tl: 8 fb: --H-FL-- lb: 0x0 cc: 2
col 0: [2] c1 03
col 1: [1] 62
end\_of\_block\_dump
End dump data blocks tsn: 4 file#: 4 minblk 7086 maxblk 7086
   第二个 session 我们查出的第一行第二列却是 a, 也是从 4 号文件 7086 个块上读取得, 但是发现第二
列已经被上锁,不能读取。
```

我们通过 uba 提供的地址,可以获取修改之前的数据 a 存放的位置:

uba: 0x00c02e17.0258.28 先把这个十六进制数转换成十进制数,oracle 提供的十进制包可以获取文件号及块号: SQL> select to_number('00c02e17','XXXXXXXXXXXXXXXXXXXXX)) from dual; TO_NUMBER('00C02E17', 'XXXXXXXXXXXXXXXXXXXX') 12594711 SQL> select dbms_utility.data_block_address_file(12594711) from dual; DBMS_UTILITY.DATA_BLOCK_ADDRESS_FILE(12594711) SQL> select dbms_utility.data_block_address_block(12594711) from dual; DBMS_UTILITY. DATA_BLOCK_ADDRESS_BLOCK (12594711) 11799 在第二个 session 中, oracle 提示读取第一行第二列需要到 3 号文件上第 11799 号块上读取: SQL> alter system dump datafile 3 block 11799; 系统已更改。 下面宅在部分转储文件: uba: 0x00c02e17.0258.25 ctl max scn: 0x0000.00665307 prv tx scn: 0x0000.00665325 txn start scn: scn: 0x0000.006658a2 logon user: 91 prev brb: 12594708 prev bcl: 0 KDO undo record: KTB Redo op: 0x03 ver: 0x01 compat bit: 4 (post-11) padding: 1 op: Z Array Update of 1 rows: tabn: 0 slot: 0(0x0) flag: 0x2c lock: 0 ckix: 191

xtype: XAxtype KDO_KDOM2 flags: 0x000000080 bdba: 0x01001bae hdba: 0x01001baa itli: 2 ispac: 0 maxfr: 4858 vect = 3 col 1: [1] 61

ncol: 2 nnew: 1 size: 0

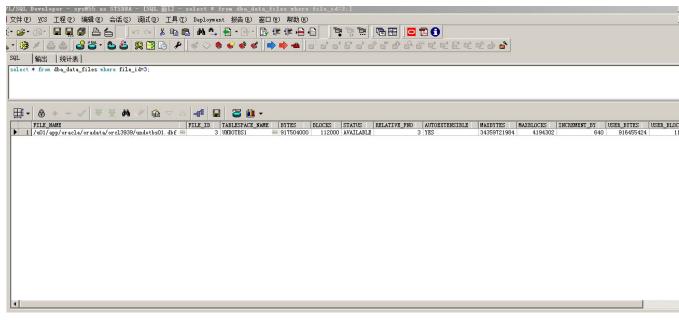
KDO Op code: 21 row dependencies Disabled

End dump data blocks tsn: 2 file#: 3 minblk 11799 maxblk 11799

所以通过 undo 读取了 61 (代表 a)

我们查看数据文件3是什么文件类型:

select * from dba_data_files where file_id=3;

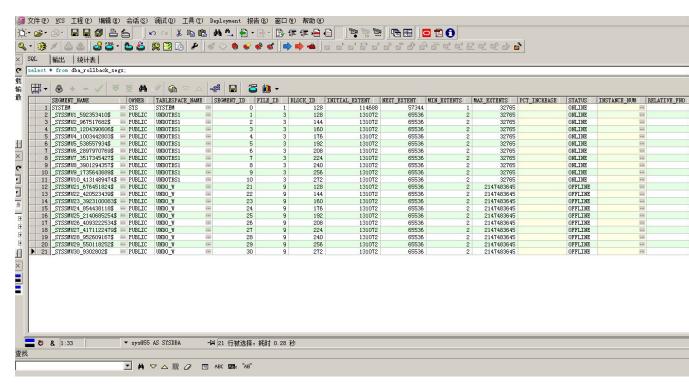


是 undo 数据文件。

SQL> show parameter undo;

NAME	TYPE	VALUE
undo_management	string	AUTO
undo_retention	integer	900
undo tablespace	string	UNDOTBS 1

默认的是 undotbs1 表空间,可以创建很多 undo 表空间,但是一个实例只用一个 undo 表空间。 undo 表空间是有 undo segments 组成,查看有多少 undo 段: select * from dba_rollback_segs;

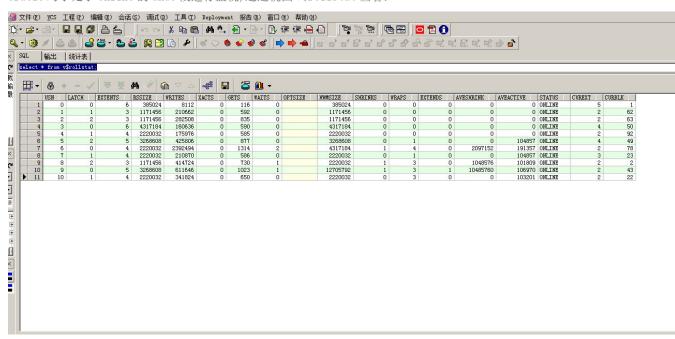


注意上面的 owner 列,如果是 public,则该实例创建的 undo 段可以被数据库其他实例使用,但是 sys 表示的是私有 undo 段,只可以被该 undo 段创建者使用。

注意到 undo 段的状态了没,现在默认的 undo 表空间是 undotbs1, 所以该 undo 段都是在线, undo_w 表空间的 undo 段都是离线。

我们可以通过修改参数 undo_tablespace 设置默认 undo 表空间。

oracle 对于处于 online 的 undo 段进行监视, 通过视图 v\$rollstat 查看:



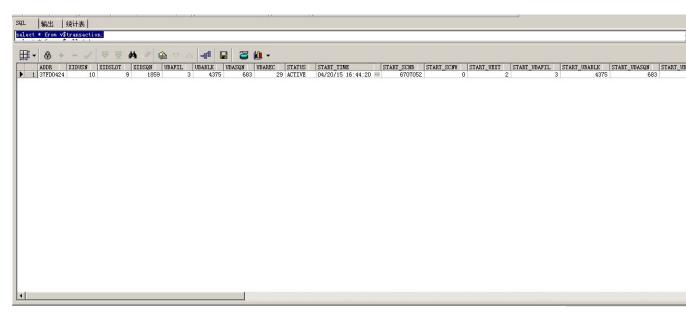
上面总共有11条, usn 是 undo 段编号 一个事务使用一个 undo 段

下面执行一个事物:

SQL> update t set name='c' where id=1;

已更新 1 行。

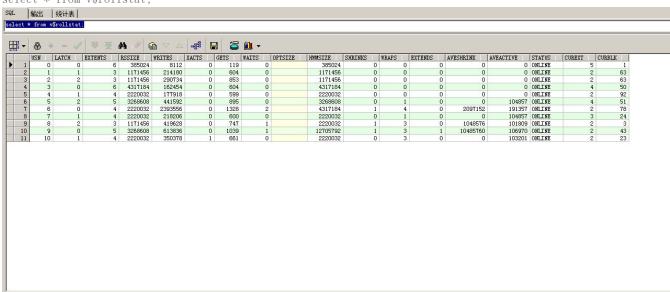
select * from v\$transaction:



注意 XIDUSN 列表示的是 undo 段编号,此时该事务使用的是 10号 undo 段

查看 10 号 undo 段:

select * from v\$rollstat;



XACTS 列表示的是该 10 号 undo 段上具有活动的事务数量此时修改默认 undo 表空间:

SQL> alter system set undo_tablespace=undo_w;

系统已更改。

SQL> show parameter undo;

NAME TYPE VALUE

undo_management string AUTO

undo_retention integer 900

undo_tablespace string UNDO_W

此时查看 undo 段状态:

select * from dba_rollback_segs;

ect * f	from dba_rollback_se	egs;	1000000											
	EGMENT_NAME	OWNER	TABLESPACE_NAME	SEGMENT_ID	D FILE_ID	BLOCK_ID	INITIAL_EXTENT	NEXT_EXTENT	MIN_EXTENTS		PCT_INCREASE STATUS	INSTANCE_NU		
1 SY	YSTEM	SYS	SYSTEM	***	0 '	1 128		57344	. 1	32765				
2 _5	SYSSMU1_592353410\$	- PUBLIC			1 7	3 128	131072	2 65536	3 2	2 32765	OFFLINE	<u>L</u>		
	SYSSMU2_967517682\$	PUBLIC			2 3	3 144								
4 5	SYSSMU3_1204390606\$	- PUBLIC			3 3	3 160						L		
5 5	SYSSMU4_1003442803\$	- PUBLIC	UNDOTBS1		4	3 176	131072	2 65536	3 2	2 32765	OFFLINE	L		
6 5	SYSSMU5_538557934\$	- PUBLIC			5 ′	3 192		2 65536	3 2	2 32765	OFFLINE	L		
7 _S	SYSSMU6_2897970769\$	- PUBLIC			6 ′	3 208	131072	2 65536	3 2	2 32765	OFFLINE	£		
8 _5	SYSSMU7_3517345427\$	- PUBLIC			7	3 224	131072	2 65536	2	2 32765	OFFLINE	L		
9 _5	SYSSMU8_3901294357\$	- PUBLIC	UNDOTBS1		8 ′	3 240	131072	2 65536	3 2	2 32765	OFFLINE	£		
10 5	SYSSMU9_1735643689\$	- PUBLIC			9 3	3 256	131072	2 65536	3 2	2 32765	OFFLINE	£		
11 _5	SYSSMU10_4131489474\$	\$ - PUBLIC			10 3	3 272	131072	2 65536	5 2	2 32765	ONLINE	AT .		
12 5	SYSSMU21_676451824\$	- PUBLIC				9 128	131072	2 65536	5 2	2 32765	ONLINE			
	SYSSMU22_420523439\$					9 144								
	SYSSMU23_3923100083\$				23 9	9 160								
15 _S	SYSSMU24_854438118\$	- PUBLIC			24 9	9 176								
	SYSSMU25_2140685254\$					9 192								
17 _5	SYSSMU26_4093222534\$	\$ PUBLIC			26 9	9 208								
18 S	SYSSMU27_4171122479\$					9 224								
19 S	SYSSMU28_952609167\$					9 240								
20 _S	SYSSMU29_550118252\$					9 256		2 65536						
21 S	SYSSMU30_9302802\$	- PUBLIC	UNIDO_W		30 9	9 272	131072	2 65536	3 2	2 32765	ONLINE			

undotbs1 里的 undo 段除了 10 号,其他的都处于 offline,因为它任被使用,事务结束后,自动变为 offline 通过查看视图:

select * from v\$rollstat;

			v\$rollsta		70 70	1-0														
						1-0														
	mt					. 41		اما		40-										
Ħ	# ▼	⊕ .	dr 4		V V	M //		-66		<u> </u>										
		USN	LATCH	EXT	TENTS	RSSIZE	WRITES	XACTS	GETS	WAITS	OPTSIZE	HWMSIZE	SHRINKS	WRAPS	EXTENDS	AVESHRINK	AVEACTIVE	STATUS	CUREXT	CUR
>	1	0	-	0	6	385024	18274	-	190			385024) (-	0		ONLINE	-	5
	2	10)	1	4	2220032	350378		1 664	(2220032) 3	0		103201	PENDING O	F	2
	3	21	1	0	2	122880	708	1	12	()	122880	0) 0	0	C	0	ONLINE		0
	4	22		1	2	122880	2428		16	()	122880) 0	0		0	ONLINE		0
	5	23		2	2				12			122880) 0		C		ONLINE		0
	6	24		0	2	122880	1522	1	14	()	122880) 0	0		0	ONLINE		0
	7	25	5	1	2	122880	1406	- 1	16	(3	122880) 0	0		0	ONLINE		0
	8	26		2	2			- 1	14	()	122880) 0	0	C		ONLINE		0
	9	27		0	2	122880	1216	- 1	14	()	122880) 0	0	0	0	ONLINE		0
	10	28		1	2				18			122880) 0				ONLINE		0
	11	29		2	2				14	()	122880) 0	0	C		ONLINE		0
	12	30	1	n	2	122880	1920		14	1	1	122880) (0			ONLINE		n '

发现 10号 undo 段状态是 pending offline (pending 在等待…期间)

会疑问,每个回滚段段上到底可以被几个事务使用呢?

 $\ensuremath{\mathsf{SQL}}\xspace^{>}\xspace$ show parameter roll

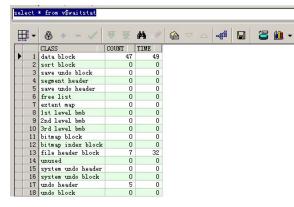
NAME		TYPE	VALUE
fast_start_parallel_rollback	string	LOW	
rollback_segments	Si	tring	
transactions_per_rollback_segment	integer	5	

可以被 5 个事务使用,但是这是在 u n d o 表空间没有自动管理之前,自从 u n d o 表空间自动管理后,该 parameter 不起作用。

一个 undo 段只能被一个事务使用,若 undo 被事务用完后,则 oracle background process smon 自 动创建 undo 段.

如果一个回滚段被多个事务使用的话, undo 段头会有等待, 影响并发性, 我们可以通过视图 V\$WAITSTAT 查看等待事件:

select * from v\$waitstat;



可以通过执行多个事务,模拟 smon 自动创建 undo 段,自行模拟试验。(smon 创建的 undo 段不会因为事务结束而回收)

下面看一下一个很重要的参数: undo_retention 单位是秒 (表示的是事务提交以后,放在 undo 里的数据保留的时间)

SQL> show parameter undo;

NAME	TYPE	VALUE
undo_management	string	AUTO
undo_retention	integer	900
undo_tablespace	string	UNDO_W

SQL> show parameter roll

很有名的 oracle 一个错误: ORA-01555 (快照太旧)

实际情况下查询发生在修改之前, 比较少。

出现这个错的可能情况:

undo 表空间太小

查询数据的时间过长(sql 查询性能差)

undo rentention 太小

我们通过视图 dba_tablespaces, 引出一个参数:

SELECT TABLESPACE_NAME, RETENTION FROM DBA_TABLESPACES;



retention 这列, undo 表空间缺省的是 NOGUARANTEE, 但是我们可以修改这个参数,强制保留。 SQL> alter tablespace undotbs1 retention GUARANTEE;

表空间已更改。

一定会保留 900 秒。

当让可以改 undo rentention

SQL> alter system set undo_retention=1200;

系统已更改。

上面列出的出现 ORA-01555 的三种情况,这三者之间有关系,比如增加了 undo_rentention, 从而需要的 undo 表空间要更多。

oracle 里提供了 undo advisor (顾问), 在 Undo Advisor 的帮助下设置 undo 表空间的尺寸,来平衡这几者关系,进入到 0EM:

[oracle@localhost ~]\$ emctl start dbconsole

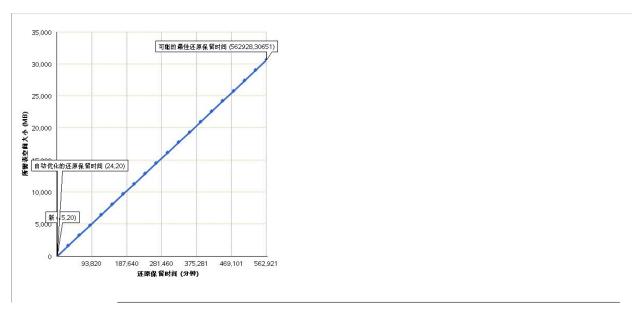
Oracle Enterprise Manager 11g Database Control Release 11.2.0.1.0

Copyright (c) 1996, 2009 Oracle Corporation. All rights reserved.

https://diy_os:1158/em/console/aboutApplication

Starting Oracle Enterprise Manager 11g Database Control

进入 em 后, 在自动管理界面, 提供了这一张图:



根据 undo_rentention 的大小,来确定所需的 undo 表空间的大小,可以更改分析时段,会提供建议图形界面。

图形是如何作出的呢?

有一个公式:

UndoSpace = [UR * (UPS * DBS)] + (DBS*24)

UR:undo_rentention in seconds

UPS:number of undo data blocks generated per second

DBS:overhead varies based on extent and file size(db_block_size)实际就是块大小: show

parameter db_block;

我们可以根据 oracle 的一个视图: v\$undostat;

1	* from v\$1				-											
zerec.	* from va	undostat;														
	1.0			_	21 2		A 1 22 1	-								
III		- 4	☆ ☆ ₩	. 0		VA	e	5 11 •								
	BEGIN TIM	ME .	END TIME			UNDOTSN	UNDOBLK	TXNCOUN	MAXQUERYLEN	MAXQUERYID	MAXCONCURRENCY	UNXPSTEALCHT	UNXPBLKRELCHT	UNXPBLKREUCHT	EXPSTEALCHT	EXPBLKRELCHT
•	1 2015/4/20	18:35:16	· 2015/4/2	0 18:	42:22	•	2	62	469 5	26 Orc4km05kgzb9	4		0) 0	0
	2 2015/4/20	18:25:16	2015/4/2	0 18:	35:16	•	2	31	163 2	22 Orc4km05kgzb9	4		0	() 0	0
	3 2015/4/20	18:15:16	2015/4/2	0 18:	25:16	-	2	347 1	099 8	22 Orc4km05kgzb9	4) 0) 0	0
	4 2015/4/20	18:05:16	2015/4/2	0 18:	15:16	•	2	12	248 13	16 Orc4km05kgzb9	3		0	() (0
	5 2015/4/20	17:55:16	2015/4/2	0 18:	05:16	•	2	96	146 6	59 Orc4kmO5kgzb9	4		0) 0	0
	6 2015/4/20	17:45:16	2015/4/2	0 17:	55:16	•	2	7	78 3	48 Orc4km05kgzb9	2) 0	() 0	0
	7 2015/4/20	17:35:16	2015/4/2	0 17:	45:16	-	11	9	57 9	50 Orc4km05kgzb9	2) 0	() 0	0
	8 2015/4/20	17:25:16	2015/4/2	0 17:	35:16	•	11	6	92 2	97 Orc4km05kgzb9	3		0	() (0
	9 2015/4/20	17:15:16	2015/4/2	0 17:	25:16	•	11	3	32 8	76 Orc4km05kgzb9	2		0) 0	0
1	0 2015/4/20	17:05:16	2015/4/2	0 17:	15:16	•	11	11	164 2	60 Orc4km05kgzb9	2		0) 0	0
1	1 2015/4/20	16:55:16	2015/4/2	0 17:	05:16	•	11	101	149 8	42 Orc4km05kgzb9	4) 0	() 0	0
1	2 2015/4/20	16:45:16	2015/4/2	0 16:	55:16	-	11	18	85 2	13 Ore4km05kgzb9	2) 0	() (0
1	3 2015/4/20	16:35:16	2015/4/2	0 16:	45:16	•	2	8	101 8	00 Orc4km05kgzb9	2		0) 0	0
1	4 2015/4/20	16:25:16	2015/4/2	0 16:	35:16	•	2	7	59 1	76 Orc4km05kgzb9	1) 0) 0	0
1	5 2015/4/20	16:15:16	2015/4/2	0 16:	25:16	•	2	12	66 7	65 Orc4km05kgzb9	3		0	() 0	0
1	6 2015/4/20	16:05:16	2015/4/2	0 16:	15:16	T	2	5	77 4	36 Ore4km05kgzb9	3) 0	() 0	0
1	7 2015/4/20	15:55:16	2015/4/2	0 16:	05:16	•	2	84	223 10	20 Orc4km05kgzb9	4) 0) 0	0
1	8 2015/4/20	15:45:16	2015/4/2	0 15:	55:16	•	2	8	59 3	90 Orc4km05kgzb9	5		0) 0	0
1	9 2015/4/20	15:35:16	2015/4/2	0 15:	45:16	•	2	9	81 9	B5 Orc4kmO5kgzb9	3		0	() 0	0
2	0 2015/4/20	15:25:16	2015/4/2	0 15:	35:16	₹	2	7	56 3	63 Ore4km05kgzb9	2) 0	() 0	0
2	1 2015/4/20	15:15:16	2015/4/2	0 15:	25:16	•	2	8	156 9	27 Orc4km05kgzb9	2) 0) 0	0
2	2 2015/4/20	15:05:16	· 2015/4/2	0 15:	15:16	•	2	7	65 3	14 Ore4km05kgzb9	2		0) 0	0
2	3 2015/4/20	14:55:16	2015/4/2	0 15:	05:16	•	2	69	138 9	02 Orc4km05kgzb9	4		0	() 0	0
2	4 2015/4/20	14:45:16	2015/4/2	0 14:	55:16	Ψ.	2	4	70 2	61 Orc4km05kgzb9	2		0	() 0	0
2	5 2015/4/20	14:35:16	2015/4/2	0 14:	45:16	•	2	15	56 8	61 Ore4km05kgzb9	4) 0) 0	0
2	6 2015/4/20	14:25:16	2015/4/2	0 14:	35:16	•	2	6	86 2	50 Orc4km05kgzb9	2		0) (0
2	7 2015/4/20	14:15:16	2015/4/2	0 14:	25:16	•	2	7		11 Orc4km05kgzb9			0) 0	0
2	8 2015/4/20	14:05:16	2015/4/2	0 14:	15:16	•	2	13		04 Orc4km05kgzb9			0	() 0	0
2	9 2015/4/20	13:55:16	2015/4/2	0 14:	05:16	•	2	109	191	0	4		0) 0	0
			· 2015/4/2				2	6		68 Orc4kmO5kgzb9	2) 0) 0	0
3	1 2015/4/20	13-35-16	9015/4/9	0 13	45-16	•	2	0	115 0	RS OvadlanOSleavki	3		0	r	1 0	n

每十分钟计算 undo 的数量。可以根据这个视图可以画出上述图形。