

数据库第14次作业

19336035 陈梓乐

1. 利用内置包 DBMS_JOB 建立定时任务，每 5 分钟在特定表中插入一条记录，内容为 emp 表的行数以及时间戳。

创建记录表

```
create table emp_record (  
    time date,  
    record number(10)  
);
```

创建记录过程

```
create or replace procedure emprec is  
begin  
    insert into emp_record values(  
        sysdate,  
        (select count(*) from emp)  
    );  
    commit;  
end;  
/
```

创建定时任务

```
declare jobnum number := 1;  
begin  
    dbms_job.submit(  
        :jobnum,  
        'emprec;',  
        sysdate,  
        'sysdate + 1/24/12'  
    );  
    commit;  
end;  
/
```

终端 问题 输出 调试控制台

SQL = select job, next_date, next_sec, failures, broken from user_jobs;

JOB	NEXT_DATE	NEXT_SEC	FAILURES	B
23	20-5月 -22	15:35:16	0	N

SQL = select * from emp_record;

TIME	RECORD
20-5月 -22	14

SQL =

2. 批量创建 99 个用户 A01-A99，口令均为 tiger，在这 99 位用户下都建立 emp 表并且把 scott 的 emp 表内容复制过去。

基本想法

利用 PL/SQL 的循环重复下列语句：

终端 问题 输出 调试控制台

SQL = create user A01 identified by tiger;

用户已创建。

SQL = grant resource,connect,dba to A01;

授权成功。

SQL = create table A01.emp as (select * from scott.emp);

表已创建。

代码

```
declare
    usernum varchar2(4);
begin
    for i in 1..99 loop
        usernum := substr(to_char(i, '09'), 2, 2);
        execute immediate 'create user A'||usernum||' identified by tiger';
        execute immediate 'grant resource,connect,dba to A'||usernum;
        execute immediate 'create table A'||usernum||'.emp as (select * from scott.emp)';
    end loop;
end;
/
```

查看效果

终端 问题 输出 调试控制台

SQL = connect A01/tiger;

已连接。

SQL = select empno, ename, sal from emp where sal > 2500;

EMPNO	ENAME	SAL
7566	JONES	3210
7698	BLAKE	3230
7782	CLARK	2950
7788	SCOTT	3210
7839	KING	5500
7902	FORD	3210

已选择6行。

SQL =

3. 利用 UTL_FILE 包（而不是 sqlldr ）把“上证指数历史数据”中的股票交易信息导入到 Oracle 的指定表中

创建表

```
create table ss001 (
    day      char(10),
    open     number(12, 6),
    high     number(12, 6),
    low      number(12, 6),
    close    number(12, 6),
    adjclose number(12, 6),
    volume   number(18, 6)
);
```

创建目录变量

```
create or replace directory data as 'data';
```

导入数据

```
declare
    l_file utl_file.file_type;
    l_text varchar2(32767);
    i      number;
begin
    l_file := utl_file.fopen('DATA', '000001.SS.csv', 'R', 32767);
    for i in 1..6189 loop
        utl_file.get_line(l_file, l_text, 32767);
        if instr(l_text, 'null') = 0 then
            insert into ss001 values(
                regexp_substr(l_text, '^[^,]+', 1, 1),
                regexp_substr(l_text, '^[^,]+', 1, 2),
                regexp_substr(l_text, '^[^,]+', 1, 3),
                regexp_substr(l_text, '^[^,]+', 1, 4),
                regexp_substr(l_text, '^[^,]+', 1, 5),
                regexp_substr(l_text, '^[^,]+', 1, 6),
                regexp_substr(l_text, '^[^,]+', 1, 7)
            );
        end if;
    end loop;
    utl_file.fclose(l_file);
    commit;
end;
/
```

效果展示

终端 问题 输出 调试控制台

```
SQL = set linesize 1000;
SQL = select * from ss001 where rownum <= 10;
```

DAY	OPEN	HIGH	LOW	CLOSE	ADJCLOSE	VOLUME
1999-03-16	1164.79102	1166.96497	1155.93506	1166.93005	1166.93005	0
1999-03-17	1167.71899	1173.04102	1158.58801	1159.09595	1159.09595	0
1999-03-18	1155.271	1160.67896	1147.56897	1151.99097	1151.99097	0
1999-03-19	1150.60901	1172.92102	1150.33997	1172.92102	1172.92102	0
1999-03-22	1175.67505	1176.65002	1168.33606	1172.74304	1172.74304	0
1999-03-23	1172.97705	1174.47595	1163.28003	1170.94898	1170.94898	0
1999-03-24	1171.302	1177.34094	1167.104	1169.03406	1169.03406	0
1999-03-25	1167.86401	1167.875	1155.01294	1156.854	1156.854	0
1999-03-26	1155.34802	1162.41602	1154.93604	1161.12	1161.12	0
1999-03-29	1160.57605	1169.50305	1158.08704	1169.48901	1169.48901	0

已选择10行。

```
SQL = select count(*) from ss001;
```

COUNT(*)
6017

4. 利用 `UTL_HTTP` 包抓取 [ITPUB 网站](#) 的首页和所有次级页面，注意不抓外链，不重复抓取，把结果记录在一张表中，包含链接，网页全文，以及从网页中提取的一些信息，例如标题等（可能要用 Oracle 的正则表达式函数族 `regex_xxxx`）

修改控制访问列表

将ITPUB加入到控制访问列表中，并授予 scott 访问的权限。这只需要在 sys 用户下运行如下代码：

```

begin
    dbms_network_acl_admin.create_acl (          -- 创建访问控制文件（ACL）
        acl            => 'http.xml',          -- 文件名称
        description    => 'HTTP Access',       -- 描述
        principal      => 'SCOTT',             -- 授权或者取消授权账号，大小写敏感
        is_grant        => TRUE,               -- 授权还是取消授权
        privilege       => 'connect',          -- 授权或者取消授权的权限列表
        start_date      => null,              -- 起始日期
        end_date        => null,              -- 结束日期
    );

    dbms_network_acl_admin.add_privilege (      -- 添加访问权限列表项
        acl            => 'http.xml',          -- 刚才创建的acl名称
        principal      => 'SCOTT',             -- 授权或取消授权用户
        is_grant        => TRUE,               -- 与上同
        privilege       => 'resolve',          -- 权限列表
        start_date      => null,
        end_date        => null
    );

    dbms_network_acl_admin.assign_acl (        -- 该段命令意思是允许访问acl名为utl_http.xml下授权的用户，使用oracle网络访问包，
        acl            => 'http.xml',
        host            => '*',

        lower_port      => 1,                 -- 且建议使用ip地址或者使用域名，若用localhost，当oracle不是安装在本机上的情况
        upper_port      => 9999              -- 允许访问的起始端口号
        -- 允许访问的截止端口号
    );
    commit;
end;
/

```

爬取网站信息

由于信息较长，我们使用 `clob` 的数据类型并分块输出。创建一个函数用于爬取网站信息，在 `scott` 下代码如下：

```

create or replace function http_get(u varchar2)
return clob as
begin
    declare
        resp    utl_http.html_pieces;
        txt     clob := '';
        i       pls_integer;
    begin
        resp := utl_http.request_pieces(u, 100);
        for i in 1..resp.count loop
            txt := txt || resp(i);
        end loop;
        return txt;
    end;
end;
/

```

阶段性效果展示

将存放于 `clob` 中的数据取出并去除重复的空格和换行，效果如下：

```

create or replace procedure print_clob(c clob) is
    i int := 1;
begin
    loop
        exit when i > dbms_lob.getlength(c);
        dbms_output.put_line(
            translate(replace(
                dbms_lob.substr(c, 255, i), ' ', ''
            ), chr(10)||chr(11)||chr(13), ' '
        ));
        i := i + 255;
    end loop;
end;
/

set serveroutput on;
declare
    c clob;
begin
    c := http_get('itpub.net');
    print_clob(c);
end;
/

```

终端 问题 输出 调试控制台

SQL = @"src\14-6 print-clob.sql"

过程已创建。

```

<!DOCTYPE html> <html lang="en"> <head> <meta charset="UTF-8"> <title>ITPUB技术论坛_专业的IT技术社区</title> <link r
el="stylesheet"
href="/itpubcms/static/css/idangerous.swiper.css"> <link href="/static/css/style_index_list.css" rel="stylesheet" />
<link rel="stylesheet" href="/itpubcms/static/css/base.css?v1"> <script src="/itpubcms/static/js/jquery.min.js"></sc
ript> <script
src="/itpubcms/static/js/idangerous.swiper.min.js"></script> <script> if(/Android|webOS|iPhone|iPod|BlackB
erry/i
.test(navigator.userAgent)) { window.location.href = "http://m.itpub.net"; } </script> </head> <b
ody> <script
src="http://168.it168.com/a/h1089.js" type="text/javascript"></script> <div style="background:#fff"><div class="w120
0"><div id="ADV_2020
8" style="padding:0 30px"><script>LG('ADV_20208')</script></div></div></div> <div id="header"> <div class="w1200 cle
arfix"> <div

```

筛选有用信息

先建立一张表格，用来记录网站的有用信息，表格结构如下：

```

create table webinfo (
    website varchar2(500),
    label varchar2(10),
    sublabel varchar2(10),
    content varchar2(500)
);

```

为了方便，我们只捕获网站中标签为 `title` 的内容，执行过程如下：

```

create or replace procedure bs(u varchar2) is
begin
    declare
        c clob;
        v varchar2(32767);
        i int := 1;
    begin
        c := http_get(u);
        loop
            -- 截取信息片段用于处理
            exit when i > dbms_lob.getlength(c);
            v := replace(dbms_lob.substr(c, 32767, i), ' ', '');
            i := i + 20000;

            -- 利用正则表达式找到 <title> 标签的内容
            v := regexp_substr(v, '<title>.*?</title>', 1, 1);
            v := substr(v, 8, length(v) - 15);
            if v is not null then
                insert into webinfo values(
                    u,
                    'title',
                    null,
                    v
                );
            end if;
        end loop;
        commit;
    end;
end;
/

```

效果如下：

```

SQL = call bs('http://www.itpub.net/');
调用完成。

SQL = call bs('http://blog.itpub.net/69955379/viewspace-2895029/');
调用完成。

SQL = select substr(website, 1, 20) website, label, sublabel, substr(content, 1, 20) content from webinfo;

```

WEBSITE	LABEL	SUBLABEL	CONTENT
http://www.itpub.net	title		ITPUB技术论坛_专业的IT技术
http://blog.itpub.net	title		CentOS7中使用efibootmgr

```

SQL =

```

找出网站所有二级链接

使用找 `title` 的经验，找出所有标签为 `a` 的链接以及描述，并对这些链接爬取源代码分析出其 `title`。


```

set serveroutput on;
create or replace procedure get_link(u varchar2) is
begin
    declare
        c clob;
        v varchar2(32767);
        l varchar2(32767);
        t varchar2(32767);
        i int := 1;
        j int := 1;
    begin
        c := http_get(u);
        <<first_loop>>
        loop
            -- 截取信息片段用于处理
            exit first_loop when i > dbms_lob.getlength(c);
            v := replace(dbms_lob.substr(c, 32767, i), ' ', '');
            i := i + 20000;
            j := 1;

            -- 利用正则表达式找到 <a> 标签的内容
            <<second_loop>>
            loop
                t := regexp_substr(v, '<a.*?</a>', 1, j);
                j := j + 1;
                exit second_loop when t is null;

                -- 找出链接
                l := regexp_substr(t, 'href=".*?"', 1, 1);
                l := substr(l, 7, length(l) - 7);
                if l like '/' then
                    l := 'http://www.itpub.net' || l;
                end if;
                dbms_output.put_line(l);

                -- 找出描述
                t := regexp_substr(t, '>.*<', 1, 1);
                t := substr(t, 2, length(t) - 2);
                dbms_output.put_line(t);
            end loop;
        end loop;
        commit;
    end;
end;
/

```

效果如下：



```

终端 问题 输出 调试控制台

SQL = call get_link('itpub.net');
http://www.itpub.net/
<imgsrc="/itpubcms/static/images/logo.png"style="height:40px;margin-left:-13px;"alt="ITPUB">
https://z.itpub.net
技术栈
http://edu.itpub.net
学院
http://www.itpub.net/forum.php
社区
http://blog.itpub.net
博客
http://wenku.it168.com
文库
http://dtcc.it168.com
DTCC
http://sacc.it168.com
SACC

```

综合上述代码

```

create or replace procedure record_webinfo(u varchar2) is
begin
    declare
        c clob;
        v varchar2(32767);
        l varchar2(32767);
        t varchar2(32767);
        i int := 1;
        j int := 1;
    begin
        c := http_get(u);
        bs(u);
        commit;
        <<first_loop>>
        loop
            -- 截取信息片段用于处理
            exit first_loop when i > dbms_lob.getlength(c);
            v := replace(dbms_lob.substr(c, 32767, i), ' ', '');
            i := i + 20000;
            j := 1;

            -- 利用正则表达式找到 <a> 标签的内容
            <<second_loop>>
            loop
                t := regexp_substr(v, '<a.*?</a>', 1, j);
                j := j + 1;
                exit second_loop when t is null;

                -- 找出链接
                l := regexp_substr(t, 'href=".*?"', 1, 1);
                l := substr(l, 7, length(l) - 7);
                if l like '/' then
                    l := 'http://www.itpub.net' || l;
                end if;
                if l like 'http:%' then
                    -- 记录
                    insert into webinfo values (
                        u,
                        'a',
                        'href',
                        l
                    );

                    -- 找出描述
                    t := regexp_substr(t, '>.*<', 1, 1);
                    t := substr(t, 2, length(t) - 2);
                    insert into webinfo values (
                        u,
                        'a',
                        null,
                        t
                    );
                    commit;
                    if l like '%itpub.net%' then
                        bs(l);
                    end if;
                end if;
            end loop;
        end loop;
        commit;
    end;
end;
/

```


终端问题输出调试控制台

SQL = select substr(website, 1, 20) website, label, sublabel, substr(content, 1, 30) content from webinfo;

WEBSITE	LABEL	SUBLABEL	CONTENT
itpub.net	a		<imgsrc="/itpubcms/static/imag
http://www.itpub.net	title		ITPUB技术论坛_专业的IT技术社区
itpub.net	a	href	http://edu.itpub.net
itpub.net	a		学院
http://edu.itpub.net	title		ITPUB学院_IT社区
itpub.net	a	href	http://www.itpub.net/forum.php
itpub.net	a		社区
http://www.itpub.net	title		论坛-ITPUB论坛-专业的IT技术社区
itpub.net	a	href	http://blog.itpub.net
itpub.net	a		博客
http://blog.itpub.ne	title		ITPUB博客_IT社区
itpub.net	a	href	http://wenku.it168.com
itpub.net	a		文库
itpub.net	a	href	http://dtcc.it168.com
itpub.net	a		DTCC
itpub.net	title		ITPUB技术论坛_专业的IT技术社区
itpub.net	a	href	http://www.itpub.net/
itpub.net	a	href	http://sacc.it168.com
itpub.net	a		SACC
itpub.net	a	href	http://cto.itpub.net
itpub.net	a		CTO俱乐部
http://cto.itpub.net	title		ITPUBCTO? ? ±? 1? é? "-? ? ? ? ? -? ? ? ? ? ?
itpub.net	a	href	http://zt.itpub.net/topic/salo
itpub.net	a		直播
http://zt.itpub.net/	title		ITPUBa? ? IT168
itpub.net	a	href	http://www.itpub.net/videolist
itpub.net	a		视频专辑
http://www.itpub.net	title		视频–ITPUB
itpub.net	a	href	http://account.itpub.net/regis
itpub.net	a		注 册
http://account.itpub	title		注册
itpub.net	a	href	http://account.itpub.net/ucen
itpub.net	a		+user.username+
http://account.itpub	title		登录
itpub.net	a	href	http://www.itpub.net/
itpub.net	a		全部
http://www.itpub.net	title		ITPUB技术论坛_专业的IT技术社区
itpub.net	a	href	http://www.itpub.net/database

筛选

找出前30个二级链接

SQL = select min(content) from webinfo where website = 'itpub.net' and sublabel = 'href' and rownum <= 30 group by content;

MIN(CONTENT)

http://www.itpub.net/forum.php
http://zt.itpub.net/topic/salon/
http://blog.itpub.net/69992972/viewspace-2895357/
http://edu.itpub.net
http://sacc.it168.com
http://www.itpub.net/cloudcomputing
http://www.itpub.net/mobiledev
http://account.itpub.net/ucenter/user/index
http://www.itpub.net/itinfar
http://www.itpub.net/it
http://www.itpub.net/blockchain
http://dtcc.it168.com
http://www.itpub.net/bigdata
http://www.itpub.net/structruing
http://blog.itpub.net/69992972
http://cto.itpub.net
http://account.itpub.net/register
http://www.itpub.net/appdevelop
http://www.itpub.net/linux
http://blog.itpub.net
http://www.itpub.net/
http://www.itpub.net/database
http://www.itpub.net/ai
http://www.itpub.net/autooperation
http://www.itpub.net/embedded
http://www.itpub.net/datamanage
http://wenku.it168.com
http://www.itpub.net/videolist
http://www.itpub.net/digitaltransfor

已选择29行。

SQL =

5. 写一个触发器，使 emp 表只有在周一到周五 8:00-18:00 这个时间段才可以被修改

```
create or replace trigger emp_time
before insert or update or delete on emp
begin
    if (to_char(sysdate,'DY') in ('星期六','星期日')) or
       (to_char(sysdate,'HH24') not between '08' and '18')
    then
        if deleting then
            RAISE_APPLICATION_ERROR (-20502,'You may delete from
            EMP table only during business hours.');
        elsif inserting then
            RAISE_APPLICATION_ERROR (-20500,'You may insert into
            EMP table only during business hours.');
        elsif updating then
            RAISE_APPLICATION_ERROR (-20503,'You may update
            EMP only during business hours.');
        else
            RAISE_APPLICATION_ERROR (-20504,'You mayupdate
            EMP table only during normal hours.');
        end if;
    end if;
end;
/
```