Advanced DataTools Loaders

Webcast on Aug 23, 2018
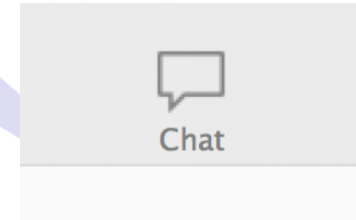## *2:00 PM EDT*

Advanced DataTools

Jack Parker

Jack is an Independent Consultant and Senior Database Consultant and Architect with Advanced DataTools. He has been working with Informix Databases since 1986. Jack is an experienced Unix, Database/Data Warehouse team leader. Skilled in large scale relational database design, development, tuning, operation and support. Expertise as an Agile practitioner. He has been an IBM Gold Consultant and IBM Information Champion.

jparker@advancedatatools.com
www.advancedatatools.com

Advanced DataTools
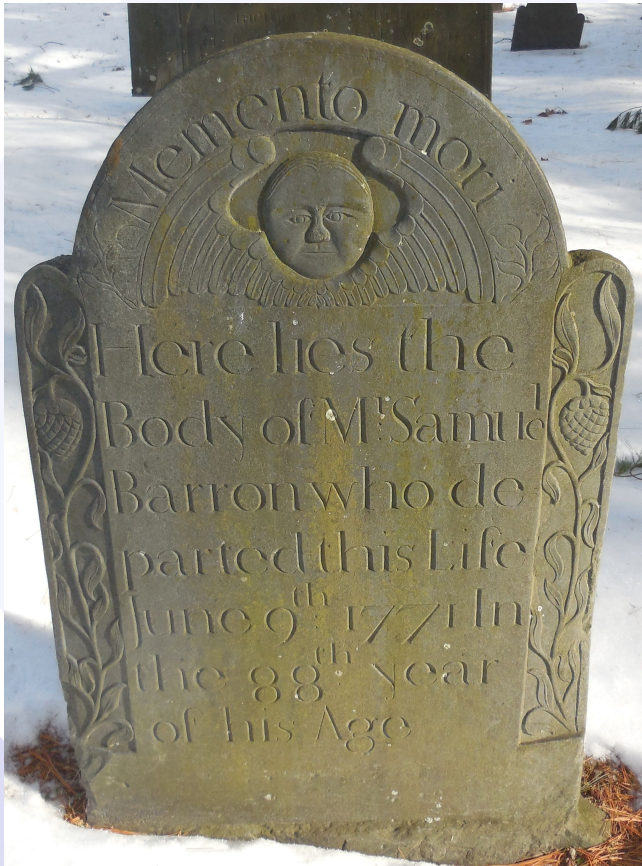
Webcast Guidelines

- The Webcast is being recorded. The Webcast replay and slides may be available in a few days.

- Please Mute your Line. Background sounds will distract everyone.

- Use the Chat Button in the upper right to ask questions.

Chat

Advanced DataTools

Load FAQ

- Posted at Advanced DataTools
- Very little change to it except with respect to external tables and HPL

Advanced DataTools

High Performance Loader / onpload / onpladm

Finicky
&
Confusing

But it does
do EBCDIC

Advanced DataTools

External Tables

- Ported from XPS to 11.x

- Much better control than HPL, especially for scripting.

Three basic loaders

- load (unload)
- dbload (seriously!?!)
- External tables
- TSL (Time series)

- Amongst our weaponry are such varied and diverse elements as…

Advanced DataTools

Considerations

- 2GB Limit (32 bit systems)

- Space

- Logging

- Violations tables

- Light Appends, Light Scans

- Partitioning

Advanced DataTools

Logging

- An insert is a logged operation
- Not into a RAW table
- Extent allocations are still logged

Advanced DataTools

Load / Unload

- Unload to file select * from table

- Pipe delimiters

  - Change with Environment variable DBDELIMITER

- Special characters escaped with "\"

  - Change with Environment variable DEFAULTESCCHAR

Advanced DataTools

Example (1 & 2)

unload to "data/months"

select * from months

unload to "data/tabcols"

select * from systables a, syscolumns b
where a.tabid=b.tabid

Load

- Load from file insert into table;
  - Very effective for small loads
  - Simple
  - Load from <filename> delimiter 'l' insert into table (col1, col3, col2)
- Example 3

Example

- load from data/months2 insert into months
- load from data/months2 insert into months2(month, mnthnbr)

Advanced DataTools

Microsoft

- Csv file
  - "This, is a test",

- Csv2pipe.awk
  - (awk –f csv2pipe.awk tst.csv)

Advanced DataTools

Oracle

- "Standard" Date format
  - 10-FEB-2018
- export DBTIME="%d-%b-%y"
- export USE_DTENV=1

Advanced DataTools

Oracle to Informix Date

```
create function o2idt(strg char(9)) returning date

DEFINE dt date;
DEFINE mm, dd, yy smallint;
DEFINE dts char(10);

LET dd=substr(strg,1,2);
LET yy=substr(strg,8,2)+2000;
select nbr INTO mm from months where month=substr(strg,4,3);

LET dts=mm II "/" II dd II "/" II yy;
LET dt=date(dts);

RETURN dt;

END FUNCTION;
```

Advanced DataTools

…Continued

- Insert into table select o2idt(date_col) from source;

Advanced DataTools

dbload

- Two parts,
  - Command file
  - Execute from command line
    - Examples 4 & 5
  - Dbload commits every 1000 rows (configurable)

Advanced DataTools

Partitioned table

- If the table is partitioned, you can load each partition in parallel

- Example 6

External tables

Create external table table_ext sameas table

using (datafiles

("DISK:/Users/jackparker/desktop/ADT/LOADERS/data/500mb"

),

format "delimited",

rejectfile
"/Users/jackparker/desktop/ADT/LOADERS/data/500mb.rej",

maxerrors 100,

express

);

Advanced DataTools

Just like a regular table

- Select count(*) from tab_ext;

- Select first 10 * from tab_ext;

- Insert into destination select * from tab_ext;

- Insert into tab_ext select * from source;

Advanced DataTools

Caution

- There are no statistics for the external table.

- The table has 2 billion rows in it.

- A join to such a table would be "interesting".

Advanced DataTools

Example

- 500MB load
- 7 (normal)
- 8 (partitioned)

Advanced DataTools

Loading from External Tables

1. Destination is raw table
   (means no replication)
2. No indexes
3. No triggers
4. Insert into …
5. Alter table <table> type (standard)
6. Take a backup (even if it is fake)
7. Create indexes and whatever else.
   Update statistics
   THIS IS NOT SOMETHING YOU SHOULD DO TO A TABLE ON A REPEATING BASIS

Advanced DataTools

Unload to compressed file

- mknod –p filename
- Create external table tab_ext sameas table
  - using (
    - datafiles("PIPE:filename"),
    - format Informix, express)
- cat filename l gzip - > file.gz &
- Insert into tab_ext select * from table;

Advanced DataTools

Load from compressed file

- Mknod –p filename
- Create external table tab_ext sameas table
  - using (
    - datafiles('PIPE:filename"),
    - format Informix, express)
- Gunzip -c file.gz > filename &
- Insert into table select * from tab_ext;

Advanced DataTools

26

Compression

- I do not recommend doing that.
- Too many moving parts

Example

- 9 (unload)

- 10 (load)

Advanced DataTools

Time Series

- More of an environment for you to write your own loader.

- Car Manual

- You aren't loading a table, just the time series within it. There may be multiple time series in a single table
  – you need to load each of them separately

Advanced DataTools

Example

EXECUTE PROCEDURE ifx_allow_newline ('t');

EXECUTE FUNCTION TSL_Init ('ts_hours','hours',2);

EXECUTE FUNCTION TSL_Put ('ts_hoursIhours',

'FILE:/Users/jackparker/desktop/ADT/LOADERS/data/ts.dat');

begin;

EXECUTE FUNCTION TSL_FlushAll('ts_hoursIhours',**261**);

commit;

EXECUTE FUNCTION TSL_SessionClose('ts_hoursIhours');

Advanced DataTools

261

- 256 + 5

- Reduced logging

- Duplicate elements replace existing elements.

Advanced DataTools

Serviceability
(aka Error Handling)

- LOAD – SOL

- Dbload – log file

- External tables – reject file

- Tsl_put – onstat –m (!)

Advanced DataTools

Use Case

- duplicate data
  - Create a table
  - start a violations table for it
  - Load it
  - All of the duplicates fall into the violations table.

Advanced DataTools

Example

start violations table for tenktwo;

set indexes jp2 filtering without error;

Load table

(in this example, no data gets loaded into the table)

Advanced DataTools

More Considerations

- PDQPRIORITY
- LIGHT SCANS
  - DIRTY READ
- Be Creative

Advanced DataTools

Questions?

Send follow-up questions to
info@advancedatatools.com

Advanced DataTools

Next Webcast

- ➢ **Capturing Informix SQL Metrics with SQLTrace**
  - ➢ **by Lester Knutsen - September 27,, 2018 at 2:00pm EST**
  - ➢ How do you capture the cost of SQL statements and activity on your database server. We will take a look at SQL Trace, and the Sysmaster tables that enable you to save and analyze the costs of your SQL.

- ➢ **Informix External Tables**
  - ➢ **by Mike Walker - November 15, 2018 at 2:00pm EST**
  - ➢ An overview of using Informix External Tables for unloading and loading data and the benefits that they offer.

- ➢ **Dostats and Informix Update Statistics**
  - ➢ **by Art Kagel December 6, 2018 at 2:00pm EST**
  - ➢ Art Kagel developed and maintains the popular open source utility dostats to enable you to automate and keep up to date catalog statistics and data distributiond that the Informix Optimizer needs in order to select the best method of running your SQL statements. He will describe how it works, how to compile it on several platforms, and why you should upgrade to the latest version. He will also discuss the Informix Automated Update Statistics (AUS) and how that compares to dostats.

  - ➢ Please register for each webcast here at:
  http://advancedatatools.com/Informix/NextWebcast.html

Advanced DataTools

Informix Training 2018

- **Advanced Informix Performance Tuning**

  - February 5-8, 2018 – **Completed and Filled up**

- **Informix for Database Administrators**

  - May 21-24, 2018 - **Completed and Filled up**

  - September 17-20, 2018 – **Registration in Progress**

- All courses can be taken online on the web from your desk or at our training center in Virginia.

- We guarantee to *NEVER* cancel a course and will teach a course as long as one student is registered!

- Please register early as the last two courses have filled up and we have not been able to accommodate everyone.

http://advancedatatools.com/Training/InformixTraining.html

Advanced DataTools

Eight New Training Servers

Each Student in class will have a server running Informix 12.10 with:

- 8 CPU Cores
- 16 GB RAM
- 1 SSD Disk
- 1- 4 disks

Advanced DataTools

Advanced DataTools

Informix Support and Training from the Informix Champions!

Advanced DataTools is an Advanced Level IBM Informix Data Management Partner, and has been an authorized Informix partner since 1993. We have a long-term relationship with IBM, we have priority access to high-level support staff, technical information, and Beta programs. Our team has been working with Informix since its inception, and includes 8 Senior Informix Database Consultants, 4 IBM Champions, 2 IIUG Director's Award winners, and an IBM Gold Consultant. We have Informix specialists Lester Knutsen and Art Kagel available to support your Informix performance tuning and monitoring requirements!

- *Informix Remote DBA Support Monitoring*
- *Informix Performance Tuning*
- *Informix Training*
- *Informix Consulting*
- *Informix Development*

Free Informix Performance Tuning Webcast replays at:
http://advancedatatools.com/Informix/Webcasts.html
Email: info@advancedatatools.com
Web: http://www.advancedatatools.com

40

Advanced DataTools

Thank You

Advanced DataTools Corporation

For more information:

info@advancedatatools.com

http://www.advancedatatools.com

Advanced DataTools