Advanced DataTools Webcast

from the IBM Informix Champions

Informix Update Statistics Best Practices for Informix DBAs by Lester Knutsen

Thursday, November 21, 2019 at 2:00pm EDT

Lester Knutsen



Lester Knutsen is President of Advanced DataTools Corporation, and has been building large data warehouse and business systems using Informix Database software since 1983. Lester focuses on large database performance tuning, training, and consulting. Lester is a member of the IBM Gold Consultant program and was presented with one of the Inaugural IBM Information Champion awards by IBM. Lester was one of the founders of the International Informix Users Group and the Washington Area Informix User Group.

> lester@advancedatatools.com www.advancedatatools.com 703-256-0267 x102

Webcast Guidelines

- The Webcast is being recorded. The Webcast replay and slides will 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

Running Update Statistics to Optimize Queries

Informix Update Statistics Webcast Agenda

- Why is Update Statistics important?
- What does Update Statistic do?
- What are Data Distributions?
- Recommendations for running Update Statistics
- How to improve performance of running Update Statistics?
- How do you monitor Update Statistics?
- Using Art Kagel's DoStats
- What is Informix Automated Update Statistics (AUS) that is built into the database server?
- How to manage AUS with InformixHQ?

Why is Update Statistics important?

- Update Statistics collects metrics to enable the server to run your SQL faster
- One of the most important tasks for a DBA
- Informix has Automated Update Statistics built into the server but you still need to monitor it

What does Update Statistic do?

- Updates the system catalogs
- Collects data distribution
- Collects information for the Query Optimizer
- Re-optimizes (compiles) stored procedures

Update Statistics Levels

- Low gathers row counts for a table
- Medium scans a *sampling* of data to collect distributions

 High – scans *every row* to collect data distributions and sorts every column in the list

Update Statistics Syntax

Syntax

>>-UPDATE STATISTICS----->

	LOW			
>+-	+-+ Table and Column Scope	+		++-><
	I	'-DROP DISTRIB	UTIONS++-' +-FORCE-+	⊦
1	I		'-ONLY-' '-AUTO'	' I
1	'-+-MEDIUM-+ Table and Column Sco	ope +	'	I
I	'-HIGH'	I	(1)	Ι
I		'- Resolut	ion Clause '	I
I	(2)			I
۰ <u>-</u>	Routine Statistics			'

Table and Column Scope

Update Statistics Low

- Default for Syntax
 - Update Statistics
- Fastest but least data gathered
- Updates:
 - Systables
 - Syscolumns
 - Sysindexes
- Does NOT update Sysdistrib

Statistics Collected During Update Statistics Low

- Systables
 - Number of rows
 - Number of pages to store the data
- Syscolumns
 - Second largest value for a column
 - Second smallest value for a column
- Sysindexes
 - Number of unique values for the lead key
- Sysindexes
 - How highly clustered the values are for the lead key

Update Statistics Medium

- Scans sampling of rows in the table to gather data distributions
- Not a full table scan
- Faster than Update Statistics High
- Default:
 - Creates 50 bins of distribution data
 - Resolution is 2.5%
 - Default Medium with confidence_level .95 and percent 2.5% will sample 2,963 records

Update Statistics Medium

- Creates or updates rows in Sysdistrib with a sampling of data distributions
- Syntax:
 - Update Statistics medium for table table_name [(column_names)] distributions only;
- Distributions only option is faster and does not redo the Update Statistics low
- Do Update Statistics low before doing medium or high

Update Statistics Resolution Clause

- Use the Resolution clause to adjust the size of the distribution bin, designate whether or not to avoid calculating data on indexes, and with the MEDIUM mode to adjust the confidence level.
 - Confidence_level A factor of the number of samples (estimated fraction of the time that sampling in MEDIUM mode should produce the same results as the exact HIGH mode). Default level is 0.95. Must be within the range from 0.80 (minimum) to 0.99 (maximum).
 - Percent Percentage of sample in each bin of distribution. Default is 2.5 for MEDIUM and 0.5 for HIGH.
 - Example
 - 100,000 rows in the table
 - Resolution of 2%
 - Each bin will represent 2,000 rows
- Default Medium with confidence_level = .95 and percent = 2.5 will sample 2,963 records.

Update Statistics High

- Scans all rows in the table to gather data distributions
- Resolution clause default percentage of data distributed to every bin is 0.5
- Default:
 - Creates 200 bins of distribution data
 - Samples all records
 - Sorts all columns in the list
- Scan and sort of large tables takes time

Update Statistics High

- Creates or updates rows in Sysdistrib with data distributions from all rows
- Distributions only option is faster and does not redo the Update Statistics low
- Do Update Statistics low before doing medium or high

Update Statistics High

- Syntax:
 - Update Statistics high for table table_name [(column_names)] distributions only;
- Every column in the list will cause a sort of the data to gather distributions
- The more columns in the list, the more sorts will be performed and the longer Update Statistics will take to complete
- Limit columns to key columns

Example Script

- -- ## Module: @(#)updatestats_zip.sql 2.0 Date: 01/01/2015
- -- ## Author: Lester Knutsen Email: lester@advancedatatools.com
- -- ## Advanced DataTools Corporation

update statistics;

update statistics low for table state; update statistics low for table zip; update statistics low for table benchmark;

update statistics medium for table state distributions only; update statistics medium for table zip distributions only; update statistics medium for table benchmark distributions only;

update statistics high for table state (state) distributions only; update statistics high for table zip (zip, state) distributions only; update statistics high for table benchmark (id, zip, state) distributions only;

What are Data Distributions?

- Example a table has a status column with four values
 - -A = Active = 49% of the data -Scan
 - I = Inactive = 49% of the data Scan
 - -P = Pending = 1% of the data Use Index
 - E = Error = 1% of the data Use Index
- Data Distributions will help the Query Optimizer decide when to use an Index

Data Distributions

- Help the SQL Optimizer pick between columns and indexes
- Not needed for every column in a table
- Select Key Columns
 - columns that are listed first in an index
 - key columns used in joins
 - key columns used as filters

Data Distributions

- Created with Update Statistics MEDIUM or HIGH
- Distributions are a mapping of the data in the column into a carefully chosen set of column values.
- The data in the column is examined and divided into bins, which represent a percentage of data. For example, if a bin might hold 2 percent of the data; 50 bins would hold all the data.
- RESOLUTION is the percent of data that is held in each bin.
 - MEDIUM = 2.5% (50 bins)
 - HIGH = 0.5% (200 bins)
- The Optimizer uses distributions for columns referenced in a WHERE clause to estimate the effect of a WHERE clause on the data.
- You must have DBA privileges, or be the owner of the table, in order to create HIGH or MEDIUM distributions.

Process for Collecting Distributions

- Develop scan plan based on available resources
- Scan table with Isolation Dirty Read
 - High = All rows
 - Medium = Sample of rows
- Sort each column
- Collect distributions into bins
- Begin transaction
 - Delete old columns distributions
 - Insert new columns distributions
- Commit transaction

Use Dbschema to See Distributions

- dbschema –d database –hd table
 dbschema –d stores_demo –hd orders
- Displays histograms of the distribution for columns of a specified table, or "all" for all tables.

Example High Data Distribution

Distribution for informix.zip.zip

Constructed on 2015-10-15 19:00:43.00000

High Mode, 0.500000 Resolution

-- DISTRIBUTION ---

	(401)
1:	(209,	209,	1074)
2:	(209,	209,	1614)
3:	(209,	209,	2158)
4:	(209,	209,	2754)
5:	(209,	209,	3285)
6:	(209,	209,	4038)
7:	(209,	209,	4469)
8:	(209,	209,	4912)
9:	(209,	209,	5474)
10:	(209,	209,	6076)
11:	(209,	209,	6510)
12:	(209,	209,	7088)
13:	(209,	209,	7757)
14:	(209,	209,	8223)
15:	(209,	209,	8863)

Bin (# of rows, # unique, highest value)

Example Medium Data Distributions

Distribution for informix.zip.city

Constructed on 2015-10-15 10:39:29.00000

Medium Mode, 2.500000 Resolution, 0.950000 Confidence

-- DISTRIBUTION ---

	(ABBEVILLE
1:	(1046,	121,	ANNAPOLIS
2:	(1046,	96,	BALTIMORE
3:	(1046,	123,	BIRMINGHAM
4:	(1046,	127,	BRONSTON
5:	(1046,	123,	CANNON BALL
6:	(1046,	121,	CHARLOTTESVILLE
7:	(1046,	113,	COCOA
8:	(1046,	119,	CREAL SPRINGS
9:	(1046,	131,	DENNISON
10:	(1046,	117,	EAST LANSING
11:	(1046,	140,	EUCHA
12:	(1046,	121,	FOREMAN
13:	(1046,	115,	GARRISON
14:	(1046,	140,	GREAT FALLS
15:	(1046,	117,	HARTSVILLE

Bin (# of rows, # unique, highest value)

Example Data Distribution Overflows

00000

(OVER	FLOW -					
1:	(537,			NEV	۷	YORK
Dist	ribu	tion f	for in	nfo	rmi>	(.)	zip.state
Const	truc	ted or	n 2015	5-10	9–15	5 :	19:00:43
High	Mod	e, 0.5	500000) R	eso]	Lu	tion
[DIST	RIBUTI	ON				
	(AK)	
1:	(37,		4,	WY)	
(OVER	FLOW -					
1:	(264,			AK)	
2:	(822,			AL)	
3:	(709,			AR)	
4 :	(442.			A7		

When one value exceeds 25% of a Bin it will get placed in the Overflow Bin (# values, value)

Update Statistics to Drop Distributions

- Clear out old data Distributions
- Important after upgrades or major data changes
- Syntax:

 Update statistics LOW for table tablename [(columns)] drop distribution

Update Statistics for Stored Procedures

- Stored Procedures are compiled and optimized when created
- Update Statistics for Stored Procedure recompiles and reoptimizes them to account for data changes
- Syntax:
 - Update statistics for stored_procedure_name

Update Statistics for Stored Procedures

- Update Statistics for function
- Update Statistics for procedure
- Update Statistics for routine

Update Statistics for Stored Procedures

- The current PDQPRIORITY setting used to run Update Statistics will be compiled into the procedure
 - Do NOT Run Update Statistics for Stored Procedures with PDQPRIORITY Set

Recommendations for running Update Statistics

- Execute UPDATE STATISTICS (LOW) for the database on a regular basis.
- Execute UPDATE STATISTICS MEDIUM DISTRIBUTIONS ONLY for selected tables, or for the entire database if time is available.
 - The DISTRIBUTIONS ONLY keyword prevents re-updating the index data and speeds up the process.
- Execute UPDATE STATISTICS HIGH DISTRIBUTIONS ONLY
 for:
 - all columns that are listed first in an index
 - key columns used in joins
 - key columns used in SQL where clauses as filters
- The goal is to balance the time required to execute update statistics vs improving query performance.

Improving Performance for Update Statistics

- Turn on PDQ when running Update Statistics for tables only
- DO NOT use PDQ when updating statistics for Stored Procedures
- When running High or Medium, increase the memory Update Statistics has to work with
 - Enable parallel sorting (i.e. PDQPRIORITY, PSORT_NPROCS)
 - Enable parallel temp dbspace (PSORT_DBTEMP, DBSPACETEMP)

Update Statistics after Upgrades

- Best practice is to run Update Statistics after an Informix server upgrade to:
 - drop old distributions
 - create new distributions for new version
- Update Statistics will also convert indexes to the format of the new database server version
- Migration to a newer database server is the only context in which the UPDATE STATISTICS statement causes table indexes to be implicitly dropped and re-created.

How do you monitor Update Statistics? Demo

- Example Scripts Demo
 - 01_aus_mon_tab_profile.sql
 - 02_aus_cmd_comp.sql
 - 03_aus_mk_todo.sql
 - 04_aus_ph_task.sql
 - 05_aus_db_info.sql
 - 06_mk_updatestatslow.sql
 - 07-aus_last_run.sql
 - 08_aus_db_info.sql

Review the Evaluator Task

- -- ## Module: @(#)01_aus_mon_tab_profile.sql 2.0 Date: 11/01/2019
- -- ## Author: Lester Knutsen Email: lester@advancedatatools.com
- -- ## Advanced DataTools Corporation
- -- ## Update Statistics Webcast Sample Scripts

- -- Select AUS Evaluator Stats from mon_table_profile
- -- Number of tables and Indexes

select count(*) Number_of_tables from mon_table_profile;

-- Usage count in sysptntab

select count(*) Number_of_ucount from mon_table_profile where ucount > 0;

-- Open count in sysptntab

select count(*) Number_of_ocount from mon_table_profile where ocount > 0;

-- See all Data

select * from mon_table_profile order by ucount desc;

See the AUS Commands

-- ## Module: @(#)02_aus_cmd_comp.sql.sql 2.0 Date: 11/01/2019

- -- ## Author: Lester Knutsen Email: lester@advancedatatools.com
- -- ## Advanced DataTools Corporation
- -- ## Update Statistics Webcast Sample Scripts

-- Get Count of AUS Commands by State

select aus_cmd_state, count(*) from aus_command

group by aus_cmd_state;

-- All AUS Commands

select * from aus_command order by aus_cmd_priority desc;

- -- AUS Commands to Do
- -- select * from aus_cmd_list;
- -- AUS Commands completed
- -- select * from aus_cmd_comp;
Unload the AUS Commands

-- ## Module: @(#)03_aus_unload.sql.sql 2.0 Date: 11/01/2019

- -- ## Author: Lester Knutsen Email: lester@advancedatatools.com
- -- ## Advanced DataTools Corporation
- -- ## Update Statistics Webcast Sample Scripts

-- AUS Commands unload

unload to current_low_update_statics.sql delimiter ';'

select aus_cmd_exe from aus_cmd_list where aus_cmd_type in ("I", "L");

unload to current_med_update_statics.sql delimiter ';'

select aus_cmd_exe from aus_cmd_list where aus_cmd_type in ("m", "M");

unload to current_high_update_statics.sql delimiter ';' select aus_cmd_exe from aus_cmd_list where aus_cmd_type in ("h", "H");

See the AUS Tasks

-- ## Module: @(#)04_aus_ph_task.sql 2.0 Date: 11/01/2019

- -- ## Author: Lester Knutsen Email: lester@advancedatatools.com
- -- ## Advanced DataTools Corporation
- -- ## Update Statistics Webcast Sample Scripts

database sysadmin;

-- View the status of the AUS Tasks in the Scheduler select * from ph_task where tk_name in ("Auto Update Statistics Evaluation",

"Auto Update Statistics Refresh")

See AUS Info by Databases

- -- ## Module: @(#)05_aus_db_info.sql.sql 2.0 Date: 11/01/2019
- -- ## Author: Lester Knutsen Email: lester@advancedatatools.com
- -- ## Advanced DataTools Corporation
- -- ## Update Statistics Webcast Sample Scripts

database sysadmin;

-- View the AUS Status by Database select * from aus_cmd_info

Script to Generate Update Statistics Low Commands for a Database

- -- ## Module: @(#)genupdatestatslow.sql 2.0 Date: 01/01/2015
- -- ## Author: Lester Knutsen Email: lester@advancedatatools.com
- -- ## Advanced DataTools Corporation

unload to updstats_low.sql delimiter ";" select "update statistics low for table " II trim (tabname) from systables where tabid > 99 and tabtype = "T";

When was AUS last run?

-- ## Module: @(#)aus_last_run.sql 2.0 Date: 01/01/2018

-- ## Author: Lester Knutsen Email: lester@advancedatatools.com

-- ## Advanced DataTools Corporation

-- ## Show when Update Status (AUS) was last run by table for a server

-- unload to "aus_last_run.uld"

select

substr(b.name,1,20) as db, substr(c.tabname,1,20) as table, aus_cmd_type as level, aus_cmd_time as when

from

sysadmin:aus_command a, sysmaster:sysdatabases b, sysmaster:systabnames c where a.aus_cmd_dbs_partnum = b.partnum and a.aus_cmd_partnum = c.partnum and b.name not matches "sys*" and c.tabname not matches "sys*" order by aus_cmd_id;

Update Statistics Last Run by Table and Column

-- ## Module: @(#)updstats_info.sql 2.0 Date: 01/01/2018

-- ## Author: Lester Knutsen Email: lester@advancedatatools.com

- -- ## Advanced DataTools Corporation
- -- ## Show when update status for last run for a table or columns in the current database
- -- Select a database before running this script

select t.tabname as tabname,

t.ustlowts as low_update,

c.colname as column,

d.constructed as upddate,

d.mode as mode,

d.constr_time as updtime,

d.ustbuildduration as updduration,

max(d.seqno) as maxseqno

from systables t, outer (sysdistrib d, syscolumns c)

where t.tabid > 99

```
and t.tabtype = "T"
```

and t.tabid = d.tabid

and d.tabid = c.tabid

```
and d.colno = c.colno
```

group by 1,2,3,4,5,6,7

order by 1, 3

Using Art Kagel's Dostats

Synopsis:

Dostats automatically generates optimal UPDATE STATISTICS statements for a table(s) or database(s) per the 7.2 release notes, the latest Performance Guide manual, and John Miller III's paper on recent improvements to way that UPDATE STATISTICS works internally and how to take advantage of those improvements. Options control what databases and tables are affected, whether commands are executed or output to a script, whether and how stored procedures are handled, the level of verbosity, tweaking the granularity of the statistical distributions captured, specifying criteria for selecting tables to update, and much more.

Art's latest release of utilities - utils2_ak.gz http://www.askdbmgt.com/my-utilities.html

Using Art Kagel's Dostats

Usage: ./dostats [-h host] -d database [-t table] [-p] [-f filename] [-g] [-N] [-i table_1 [-i table_2 [...]]] [-x [table_a|dbase_a:] [-x...]] [-i 0[file1] [-i 0[file2]]] [-x 0[file1] [-x 0[file2]]] [-G] [-i '!where clause'|subquery] [-x '!where clause'] [-S] [-v<-1|0|1>] [-e|-E] [-s] [-V] [-F] [-T] [-I] [-C] [-M] [-L] [-P procPDQ] [-Q tablePDQ] [-w Nsecs] [-X] [-R HighRes] [-r MedRes] [-c MedConf] [-Z MinSamp|MinPct] [-m] [-b[-B BrowseThreshHold]] [-a [-A AgingLimit]] [-n errlimit] [--procedure=procname] [--proc-local] [--execute-local] [--schedule [--frequency [d|w[0-6]]m[0-6]]]

> [--first-time <datetime>] --distributions-high <list file> [--one-time]] [--aus-thresholds] [--clean-distributions] [--time-display] [--display-local]

[--drop-distributions] [--force_run] [--include-owner]

[--small-tables-high [--small-tables-threshold=<val>]]

- [--isolation <c|d|l>] [--force-run] [--auto-run]
- -h Host defaults to \$INFORMIXSERVER. (The program uses multiple connections to the host. This requires a non-shared memory connection.)
- -d Database may contain valid "matches" clause wildcards. The string 'ALL' is equivalent to '*' and will cause ./dostats to execute against all databases except sysmaster.
- -t Table may contain valid "matches" clause wildcards. Table defaults to '*', all user tables (No system catalog tables.) The special name "none" can be used with the -i option to specify a limited list of tables with dissimilar names.
- -G Get physical row count using COUNT(*). On servers with a very large number of tables this option can improve runtime as sysptnhdr has no indices and must be scanned.
- -b Browse mode. Select only tables which have had their row count change by a percentage specified by the -B, browse threshhold, option since statistics were last updated. Overrides inclusions. This option requires a non-shared memory connection and so is not compatible with -S.

Using Art Kagel's Dostats

 Requires the Client SDK and LD_LIBRARY_PATH set

export LD_LIBRARY_PATH=\$INFORMIXDIR/lib:\$INFORMIXDIR/lib/esql:\$INFORMIXDIR/lib/tools

- Requires TCP connection
- Art's recommendation:

dostats -d <database> -Q <PDQ level> -E --display-local

Using Art Kagel's Dostats Example: dostats –d database

informix@tiger1:~/AKutilities train1tcp > ./dostats -d benchmark3 Warning! Optimized servers produce statistics fastest with PSORT_NPROCS set. Warning! Optimized servers produce statistics fastest with PSORT_DBTEMP set. Warning! Optimized servers produce statistics fastest with DBUPSPACE set. Working on database: benchmark3. Working on table warehouse: w_id(LOW)...SUCCESSFUL w_id(HIGH)...SUCCESSFUL w_ytd, w_tax, w_name, w_street_1, w_street_2, w_city, w_state, w_zip(MEDIUM)...SUCCES Table warehouse completed Working on table district: d_w_id, d_id(LOW)...SUCCESSFUL d w id(HIGH)...SUCCESSFUL d_id, d_ytd, d_tax, d_next_o_id, d_name, d_street_1, d_street_2, d_city, d_state, d_z Table district completed Working on table customer: c_w_id, c_d_id, c_id(LOW)...SUCCESSFUL c_w_id, c_d_id, c_last, c_first(LOW)...SUCCESSFUL c_w_id, c_id, c_last(HIGH)...SUCCESSFUL c_d_id, c_discount, c_credit, c_first, c_credit_lim, c_balance, c_ytd_payment, c_paym

Automated Update Statistics (AUS)

- AUS is a set of tasks that run in the Sysadmin database on Informix Server 11.50 and greater
- AUS uses a combination of scheduler sensors, tasks, thresholds, and tables to evaluate and run Update Statistics
- Evaluate daily identify what tables need Update Statistics run based on data change thresholds
- Execute weekly default is to run the Update Statistics commands that need to run on Saturday and Sunday

AUS Works with Logged Databases

- AUS updates the statistics for tables that are in logged databases
- AUS will work with any database locale
- Sysadmin database is Logged and can only connect and work in other Logged Databases
- IBM Knowledge Center AUS web page

https://www.ibm.com/support/knowledgecenter/en/SSGU8G_14.1.0/com.ibm.perf.doc/ids_prf_742.htm

AUS Evaluation Task

- Scheduler daily runs mon_table_profile to identify how many rows have changed in each table
- Determines which tables need updates based on the expiration policies
- Inserts Update Statistics commands into the aus_command table to be run

AUS Updates Statistics Refresh Task

- Runs the statements in the aus_command table
- Default schedule is Saturday and Sunday between 1:00 AM and 5:00 AM
- Updates aus_command table with results aus_cmd_state
 - P = Pending
 - I = In progress
 - E = Error
 - C = Complete without errors

AUS Tables and Views

- Table mon_table_profile Tracks inserts, updates, and deletes by table from sysmaster:sysactptnhdr
- Table aus_command Update Statistics commands and status
- Table ph_task Scheduler for AUS Tasks
- View aus_cmd_comp Completed commands
- View aus_cmd_list Update Statistics commands
- Table aus_cmd_info Info on AUS by database

AUS Rules in ph_threshold

Threshold Name	Default	Description
AUS_AGE	30 days	Statistics are updated for a table after this amount of time regardless of how much data has changed
AUS_AUTO_RULES	1 - enables	 Rules: All tables are updated in LOW mode. All the leading index keys are updated in HIGH mode. All non-leading index keys are updated in MEDIUM mode. The minimum resolution for MEDIUM mode is 2.0. The minimum confidence for MEDIUM mode is 0.95. The minimum resolution for HIGH mode is 0.5. If the UPDATE STATISTICS statement was run manually, AUS will continue with the same level, resolution, confidence, or sampling size options.
AUS_PDQ	10	PDQ priority for UPDATE STATISTICS statements run by the AUS maintenance system
AUS_SMALL_TABLES	100 Rows	Statistics or distributions are updated every time for a table that has fewer than this number of rows

AUS Prioritizing Databases

- Default all databases are prioritized as medium
- Assign a higher priority to ensure your critical databases are done first in AUS
- Insert a record into ph_threshold to set the priority to High

INSERT INTO ph_threshold

(id, name, task_name, value, value_type, description) VALUES

(0, "AUS_DATABASE_HIGH", "Auto Update Statistics Evaluation", "database_name", "STRING", "Rank high priority to get done first")

Changing the AUS Schedule

- Update the task in the scheduler to change the schedule, the days AUS runs, and how long it runs
- Update stop time to 8:00am:
 - Update ph_task SET tk_stop_time = "08:00:00" WHERE tk_name = "Auto Update Statistics Refresh";

Change schedule to run every day

UPDATE ph_task SET tk_monday = "T", tk_tuesday = "T", tk_wednesday = "T", tk_thursday = "T", tk_friday = "T" WHERE tk_name = "Auto Update Statistics Refresh"; tk_name is case sensitive

tk name is case sensitive

AUS Onconfig Paramaters

- AUTO_STAT_MODE Enables (1) or disables (0) Update Statistics automatic mode where statistics of table, fragment, or index are rebuilt only if existing statistics are considered stale, based on STATCHANGE
- STATCHANGE Percent of change to trigger a rebuild of statistics for table, fragment, or index; default is 10%
- USTLOW_SAMPLE Enables (1) or disables (0) the use of sampling during Update Statistics Low operations for large indexes

Update Statistics Force and Auto Keywords

- FORCE keyword refreshes the statistics for all tables and columns and overrides automatic mode
- AUTO keyword causes the database server to run automatic mode for tables and fragments whose statistics are missing or stale, based on the STATCHANGE value

Manage AUS with InformixHQ

- Example InformixHQ Screens
 - Auto Update Statistics Overview
 - Auto Update Statistics Configuration
 - Auto Update Statistics Alerts
 - Auto Update Statistics Commands
 - Task Scheduler Auto Update Statistics
 Evaluation
 - Task Scheduler Auto Update Statistics
 Refresh

Manage AUS with InformixHQ

Informix					⊕ ~ <u>●</u> 1
٩	Root Group > Group-Train1 > tiger1-train1				
tiger1-train1 \checkmark	tiger1-train1 Server Data Agent Data			Vie	w last: 4 hours 🗘 🚺 🔺
Setup	Status		Incidents		
Dashboards	Server: Online Database spaces:	0 spaces < 5% free	O incidents have occurred yet.		
Monitoring	Server type: Standard Auto Ledits Statistics				
Alerting	Online log: Q 4 errors, Auto Opulate Statistics. 0 warnings	O OINIOWI	Storage Performance		
Permissions	Availability		Last Checkpoint		Cache Hit Rates
Incidents	▲ Server is not part of a high availability cluster.		0.007 seconds		
Configuration			2 minutes ago	100	Cached Write %
Logs >	Threads			50	
Performance >	Tatel Theode Deads Theode Molifes Theore	de Weiting en Muteu		0	6:44:00 pm 6:49:00 pr
Replication >	Total Inreads Ready Inreads Waiting Inread	25 Waiting on Mutex			Latest data timestamp: 06:48:19 P
Schema Manager	41 0 8	0	Foreground Writes (per seco	nd)	Sequential Scans
Server Administration ~			1.0		50332
Auto Update Statistics	Host		0.5		
Privileges	Operating System Memory		-0.5		
Task Scheduler	Informix Memory		-1.0 6:39:00 pm 6:44:00 pm Latest data time:	6:49:00 pm stamp: 06:48:19 PM	
Storage >	OS Memory				
SQL Tracing	0 B 2 GB 4 GB 6 GB 8 GB 10 GB	3 12 GB 14 GB 16 GB	Sessions		
System Reports			Total Sessions	Largest Session Memory	Average Session Memory
System Resources >	Server Info VPs: 14 total 4 cpu Version: IBM Informix Dynamic Server Version 14.10.FC2		11	1.58 MB	548.36 KB
	Uptime: 20 days 7 hours 13 minutes 3 seconds Host: tiger1 OS: Linux x86_64 3.10.0-1062.4.1.el7.x86_64				

AU

Auto Update Statistics Overview

Auto Update Statistics

Overview Configuration Alerts Commands	
tatistics Summary Last evaluation completed at 2019-11-18 18:40:03	Evaluate Now Clean Up
Tables Missing Statistics	
0	
Large Tables Needing Statistics Refreshed	
25	
Small Tables Needing Statistics Refreshed	
514	
Tables With Refreshed Statistics	

0

Auto Update Statistics by Database Auto Update Statistics Refresh will run at 2019-11-23 01:11:00

Last Time Checked $\hat{\downarrow}$	Database 🝦	Tables Missing Statistics $\hat{\downarrow}$	Large Tables Needing Statistics Refreshed $\frac{1}{2}$	Small Tables Needing Statistics Refreshed $\frac{^{\scriptscriptstyle A}}{^{\scriptscriptstyle V}}$	Tables With Refreshed Statistics $\stackrel{\scriptscriptstyle +}{\scriptscriptstyle \vee}$
2019-11-18 01:00:00	sysmaster	0	0	211	0
2019-11-18 01:00:00	sysutils	0	1	40	0
2019-11-18 01:00:00	sysuser	0	1	46	0
2019-11-18 01:00:00	sysadmin	0	18	61	0
2019-11-18 01:00:01	benchmark1	0	1	38	0
2019-11-18 01:00:01	benchmark2	0	1	37	0
2019-11-18 01:00:01	benchmark3	0	2	38	0
2019-11-18 01:00:01	stores_demo	0	1	43	0

Auto Update Statistics Configuration

Auto Update Statistics	;										
Overview Configuration	Alerts Commands										
Auto Update Statistics Schedule											
Name 🗘	Start Time 🗘	Stop Time 🍦	Run Frequency 🗘	М ‡	τ 🔅	w ‡	Τ 🗘	F ÷	S +	S ÷	Enable 🔆
Auto Update Statistics Evaluation	01:00:00	01:10:00	1 00:00:00	~	×	×	×	*	×	×	~
Auto Update Statistics Refresh	01:11:00	05:00:00	1 00:00:00	8	8	8	8	8	*	×	*
Auto Update Statistics Configuration	on										Configure Database Priority
Name	Description					Value					
AUS_AGE	The statistics are rebuilt after this many days	L				30 days	s. "#*				
AUS_CHANGE	The statistics are rebuilt after this percentage	e of data has changed.				10 % _	n ^a				
AUS_AUTO_RULES	Ensures a base set of guidelines are followed	when building statistics.				On 💉	•				
AUS_SMALL_TABLES	Tables containing less than this number of ro	ws will always have their statist	tics rebuilt.			100 rov	vs 🔎				
AUS_PDQ	Update statistics executes with this PDQ prio	rity.				10 prior	rity 🦯				
Thread	The Auto Update Statistics Refresh task runs	with this number of threads.				1 thread	ds 🎤				

Auto Update Statistics Alerts

Auto Update Statistics

Overview Configuration

Alerts Commands

Auto Update Statistics Alerts

Time 🖕	Туре 🗦	Color ≑	Message 🍦
2019-11-18 01:00:01	0	•	Found 44 table(s) in database stores_demo which need statistics updated.
2019-11-18 01:00:01	0	•	Found 40 table(s) in database benchmark3 which need statistics updated.
2019-11-18 01:00:01	0	•	Found 38 table(s) in database benchmark2 which need statistics updated.
2019-11-18 01:00:01	0	•	Found 39 table(s) in database benchmark1 which need statistics updated.
2019-11-18 01:00:01	0	•	Found 79 table(s) in database sysadmin which need statistics updated.
2019-11-18 01:00:00	0	•	Found 47 table(s) in database sysuser which need statistics updated.
2019-11-18 01:00:00	0	•	Found 41 table(s) in database sysutils which need statistics updated.
2019-11-18 01:00:00	0	•	Found 186 table(s) in database sysmaster which need statistics updated.
2019-11-17 01:00:01	0	•	Found 44 table(s) in database stores_demo which need statistics updated.
2019-11-17 01:00:01	0	•	Found 40 table(s) in database benchmark3 which need statistics updated.
2019-11-17 01:00:01	0	•	Found 38 table(s) in database benchmark2 which need statistics updated.
2019-11-17 01:00:01	0	•	Found 39 table(s) in database benchmark1 which need statistics updated.
2019-11-17 01:00:01	0	•	Found 78 table(s) in database sysadmin which need statistics updated.
2010-11-17 01-00-00	A	•	Found 47 tabla(e) in database eveneer which need statistics undated

Auto Update Statistics Commands

Auto Update Statistics

Overview Configuration Alerts Commands Commands Pending Update Statistics Commands Q Search Command 🔅 UPDATE STATISTICS LOW FOR TABLE sysadmin:informix.systables UPDATE STATISTICS HIGH FOR TABLE sysadmin:informix.systables(tabname.tabid) RESOLUTION 0.50 DISTRIBUTIONS ONLY UPDATE STATISTICS MEDIUM FOR TABLE sysadmin:informix.systables(owner) RESOLUTION 2.00 0.95 DISTRIBUTIONS ONLY UPDATE STATISTICS LOW FOR TABLE sysadmin:informix.syscolumns UPDATE STATISTICS HIGH FOR TABLE sysadmin:informix.syscolumns(tabid,extended_id) RESOLUTION 0.50 DISTRIBUTIONS ONLY UPDATE STATISTICS MEDIUM FOR TABLE sysadmin:informix.syscolumns(colno) RESOLUTION 2.00 0.95 DISTRIBUTIONS ONLY UPDATE STATISTICS LOW FOR TABLE sysadmin:informix.sysindices UPDATE STATISTICS HIGH FOR TABLE sysadmin:informix.sysindices(idxname,tabid) RESOLUTION 0.50 DISTRIBUTIONS ONLY UPDATE STATISTICS MEDIUM FOR TABLE sysadmin:informix.sysindices(owner) RESOLUTION 2.00 0.95 DISTRIBUTIONS ONLY UPDATE STATISTICS LOW FOR TABLE sysadmin:informix.systabauth UPDATE STATISTICS HIGH FOR TABLE sysadmin:informix.systabauth(tabid) RESOLUTION 0.50 DISTRIBUTIONS ONLY UPDATE STATISTICS MEDIUM FOR TABLE sysadmin:informix.systabauth(grantor,grantee) RESOLUTION 2.00 0.95 DISTRIBUTIONS ONLY UPDATE STATISTICS LOW FOR TABLE sysadmin:informix.sysviews UPDATE STATISTICS HIGH FOR TABLE sysadmin:informix.sysviews(tabid) RESOLUTION 0.50 DISTRIBUTIONS ONLY UPDATE STATISTICS MEDIUM FOR TABLE sysadmin:informix.sysviews(seqno) RESOLUTION 2.00 0.95 DISTRIBUTIONS ONLY UPDATE STATISTICS LOW FOR TABLE sysadmin:informix.sysdepend

Task - Auto Update Statistics Evaluation

AUS

sk Setup						×
Task Name & Group						
Task Name			Select Gro	oup		
Auto Update Statistics Evaluation	1		PERFORM	ANCE		\$
Description						
To Evaluate which columns and	tables should have the	statistics an	d distribution:	s refreshed.		4
Specify schedule for Task						
Start Time (24 hrs.)			Stop Time	(24 hrs.) Never		
01 🔅 : 00 🔅 : 00 🔅			01 🤤	10 🔅 : 00 🤤		
Frequency Never 01 0 : 00 0 : 00 Days	00					
Monday Tuesday Wed	nesday Thursday	Friday	Saturday	Sunday		
Specify the command for Tas Command	k					
aus_evaluate_stats						~
Enable Task View Parameters					Cancel	Edit
						_

Task - Auto Update Statistics Refresh

AUS

sk Setup		×
Task Name & Group		
Task Name	Select Group	
Auto Update Statistics Refresh	PERFORMANCE	\$
Description		
Refreshes the statistics and distributions which were recommend	ded by the evaluator.	
Specify schedule for Task		
Start Time (24 hrs.)	Stop Time (24 hrs.) Never	
01 (1) : 11 (1) : 00 (1)	05 (\$) : 00 (\$) : 00 (\$)	
Frequency Never 01 0 0 00 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0		
Days		
Monday Tuesday Wednesday Thursday Friday	Saturday Sunday	
specify the command for Task		
Command		
aus_refresh_stats		
Enable Task		
View Parameters	Са	ncel Edit

Update Statistics Summary

- Correct Statistics is a key factor in the performance of your server
- Without Statistics the SQL Optimizer may pick an incorrect execution plan
- Auto Update Statistics is very effective and improved over previous versions
- A DBA is still responsible for monitoring

Questions?



Send follow-up questions to Lester@advancedatatools.com



Free Informix Webcasts

from the IBM Informix Champions

<u>Next Year – Informix Tutorials Webcast Series:</u> <u>a step by step approach to using Informix Database Servers</u>

- Getting Started with Informix by Lester Knutsen on January 30th, 2020 at 2:00 pm EDT This Webcast is a step-by-step guide to installing and getting up and running with a basic Informix Server. We will introduce using InformixHQ, the IBM Informix Knowledge Center, and how to find resources to get started with Informix.
- Configuring a New Informix Server by Lester Knutsen on February 27th, 2020 at 2:00 pm EDT This Webcast will be an introduction to the Informix ONCONFIG file and configuring memory, CPUs, network, and disk for a more extensive Informix Server.
- Managing Informix Disk Space Dates TBD
- Managing Informix Logs Dates TBD
- Informix Backup, Recovery, and High Availability Dates TBD
- Connecting Users to Informix Servers Dates TBD
- Creating Databases and Tables in Informix Dates TBD
- Basic Informix Server Monitoring Dates TBD

Registration and more information: https://advancedatatools.com/Informix/NextWebcast.html

Informix Training Updated for Informix 14.10

Attend classes online on the web or in person at our training center in Virginia. All you need is a web browser to connect to our WebEx training system, and an SSH client (like Putty) to connect to our training lab for hands-on exercises. Each student uses an 8-core Linux server, with 16GB RAM, SSD drives with Informix 14, and several large databases for benchmark exercises.

Informix Training in 2020

- May 18-21, 2020 Informix for Database Administrators
- July 6-9, 2020 Advanced Informix Performance Tuning
- October 5-8, 2020 Informix for Database Administrators

More information and registration at:

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

Informix 14 Training



Each student in class will have a server running Informix 14.10 with:

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

Class size is limited to 8 students.

Attend online or in person!

Coming Soon - New Website

Please test and send up feedback https://live-advancedatatools.pantheonsite.io





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, 3 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


Thank You Advanced DataTools Corporation



For more information:

Lester@advancedatatools.com http://www.advancedatatools.com

Advanced DataTools