



Making your engine hum with  
ALICE, and other  
BTSCANNER tips

Mark Jamison  
IBM

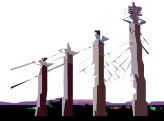
Session D07  
Day, April 0, 2008 • 00:00 a.m. – 00:00 a.m.

2008 IIUG Inform*i*x Conference



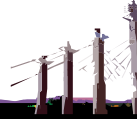
## Agenda

- Learn about how BTSCANNER works
- Learn about ALICE
- Learn to Monitor Index Cleaning
- Basic troubleshooting



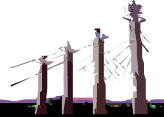
## How IDS used to Clean Indices

- A single thread, named btclean, handled all of the cleaning.
- When an index item was removed, a delete flag was set in the item's key.
- A request was created and inserted onto the btclean's request queue.
- The btclean thread then serially processed each non-duplicate cleaning request.



## How IDS Cleans with BTSCANNER

- At least one thread, named btscanner[0-x], handles cleaning.
- When an index item was removed, a delete flag was set in the item's key.
- Increase hit count for Index, or table if attached index, Partition.
- Scan list of Partitions with hit counts, add all partitions to cleaning list that have a hit count greater than the designated threshold
- Clean using the appropriate method, and compress where possible.
- For all Partitions that have been cleaned, reset hit counter to 0.

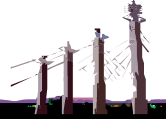


BTSCANNER was introduced in the following versions  
7.31.UD8, 9.21.UC7, and 9.30.UC6 and greater.

In addition to being more efficient to the BTCLEANER, by Default, it also allows for Index cleaning to be handled by multiple thread, instead of being single threaded.

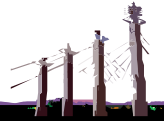
## The Hit Counter and Hot List

- Hit counter is the number of time an index page that contains deleted items has been scanned.
- The Hot List is the list of indices that need to be cleaned based on their hit counter having exceeded the defined threshold.



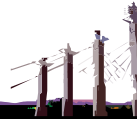
## A Note on Thread Priority

- BTSCANNER has 2 priority Modes
  - High
  - Low
- Priority was removed in 10.00.FC7 and 11.10.UC1



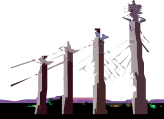
## What cleaning methods are there?

- Leaf Scan
- Range Scan
- ALICE Scan



## Leaf Scan Mode

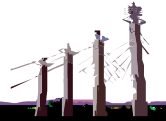
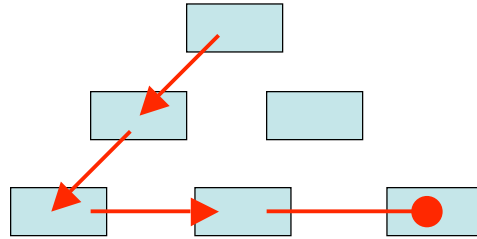
- Leaf Scan Mode:
  - Start at the root node of a partition.
  - Traverse down the left side of the tree until the leaf level is reached.
  - Scan across entire index looking for the delete flag.
  - Cleans each page with a delete flag, and attempts to compress pages where possible.





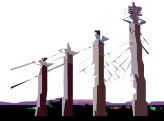
## Leaf Scan Mode

- Example of Leaf Scan.



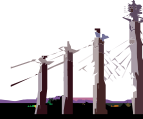
## Leaf Scan Mode

- Positives:
  - Extremely efficient for small indices.
  - Can be used for detached or attached indices.
- Negatives:
  - I/O intensive.



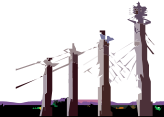
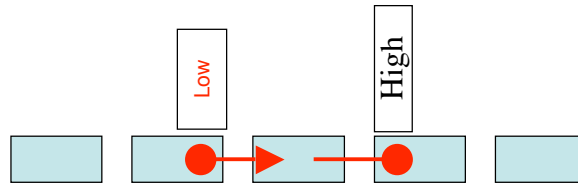
## Range Scan

- Range Scan Mode
  - Start the scan by reading a block of pages starting with the lowest logical page for which a request for cleaning has been made.
  - Examine the block for pages which have deleted items.
  - Request next block of pages to be read asynchronously.
  - Read the list of pages to be cleaned into the buffer pool, clean and, if possible, compress.
  - Stop the scan when the highest logical page for which a request has been made is encountered



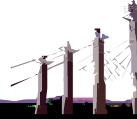
## Range Scan Mode

- Example of a Range Scan



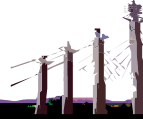
## Range Scan Mode

- Positives:
  - Much more efficient from an I/O perspective.
- Negatives:
  - Does not work on table that has multiple attached indices.
  - Range scan is dependent on index layout, and not on the data contained in index.



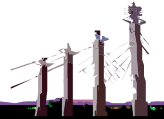
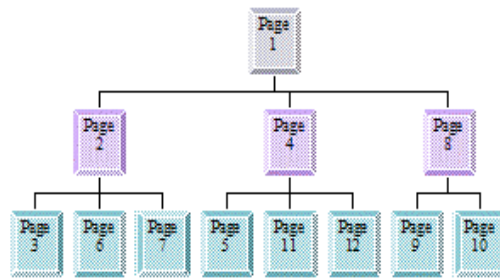
## ALICE Scan Mode

- (A)utonomic (L)inear (I)ndex (C)l(E)aning Mode
  - When an item is deleted, the index region's bitmap is updated to indicate that cleaning needs to occur in that region.
  - ALICE only reads the pages from the regions whose bitmap entry indicates that cleaning needs to occur.
  - Once the region is cleaned the related bitmap entry is reset.



## ALICE Scan Mode

- ALICE takes a Normal Index like the following:

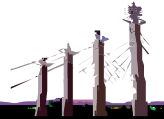


## ALICE Scan Mode

- And views that Index as a linear object like the following:

Linear Representation of the Index

P 1	P 2	P 3	P 4	P 5	P 6	P 7	P 8	P 9	P 10	P 11	P 12
--------	--------	--------	--------	--------	--------	--------	--------	--------	---------	---------	---------



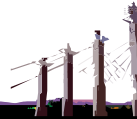


## Alice Scan Mode

- Divide the index into regions like the following:

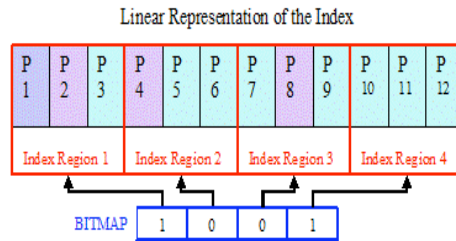
Index Regions

P 1	P 2	P 3	P 4	P 5	P 6	P 7	P 8	P 9	P 10	P 11	P 12
Index Region 1			Index Region 2			Index Region 3			Index Region 4		



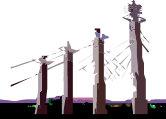
## Alice Scan Mode

- Finally it represents each region as part of a bitmap like the following:



## ALICE Scan Mode

- Positives:
  - Only reads the regions of an index that need cleaning.
  - Is promote able.
- Negatives:
  - Bitmaps require space and present a size issue to memory strapped instances.



Note that the size is not

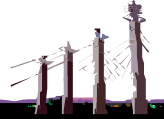
## Configuring BTSCANNER

- BTSCANNER  
num=#,priority=#,threshold=#,rangesize=#,alice=#
  - Num = Number of btree scanner threads to start at IDS initialization
  - Priority = what priority should the btree scanner threads be
    - Note, eliminated in 10.00.UC7 and greater
  - Threshold = number the hit counter must exceed before it will be placed onto the hot list.
  - Rangesize = the number of pages required in the index before a range scan will be done.
  - Alice = the ALICE mode to set as default at the instance level



## Configuring BTSCANNER

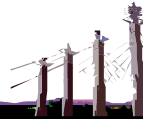
- onmode -C start # = start # of btscanner threads
- onmode -C stop # = stop # of btscanner threads.
- onmode -C threshold # = Set the threshold.
- onmode -C rangesize # = set the rangesize.
- onmode -C duration # = duration to expire Hot List
- onmode -C alice # = set the alice mode.



Also available


## Monitoring BTSCANNER

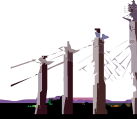
- onstat -C
- onstat -C hot
- onstat -C part
- onstat -C clean
- onstat -C range
- onstat -C alic
- onstat -C map



## Monitoring BTSCANNER

The **onstat -C** output has the following fields:

<b>Field</b>	<b>Description</b>
BTS info	
Id	BTSCANNER ID
Prio	Current priority of BTSCANNER
Partnum	The partition number for the index this thread is currently working on.
Cmd	Current command this thread is processing.
 Table Name	Name of the table or index.

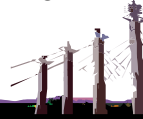


## Monitoring BTSCANNER

The **onstat -C hot** output has the following fields:

<b>Field</b>	<b>Description</b>
Partnum	The partition number for an index.
Key	Index Key,
Hits	The current value of the Hit counter.

“\*” This indicates that this partition has been cleaned during this hot list duration.





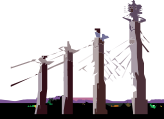
## Monitoring BTSCANNER

The **onstat -C part** output has the following fields:

<b>Field</b>	<b>Description</b>
Partnum	The partition number for an index
Key	Index Key.
Positions	Number of times index has been read.
Compress	Number of pages which have been compressed.
Split	Number of splits that have occurred.

“C” Indicates partition is busy being cleaned.

“N” Index partition no longer eligible for cleaning.



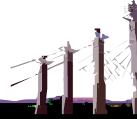
## Monitoring BTSCANNER

The **onstat -C clean** output has the following fields:

<b>Field</b>	<b>Description</b>
Partnum	The partition number for an index
Key	Index Key
Dirty Hits	Number of times a dirty page has been scanned
Clean Time	Total time spent, in seconds
Pg Examined	Number of pages examined by btscanner thread.
Items Del	Number of items removed from this index.
Pages/Sec	Number of pages examined per second.

“C” Indicates partition is busy being cleaned.

“N” index partition is no longer eligible for cleaning.



## Monitoring BTSCANNER

The **onstat -C range** output has the following fields:

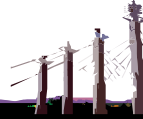
<b>Field</b>	<b>Description</b>
Partnum	The partition number
Key	Index Key
Low	Low boundary for range scan.
High	High boundary for index scan.
Size	Size of index in pages.
Saving	Percentage of time saved versus a full scan.

“C” Indicates partition is busy being cleaned.

“N” Index partition is no longer eligible for cleaning.



[www.iiug.org](http://www.iiug.org)



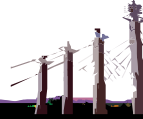
## Monitoring BTSCANNER

The **onstat -C alice** output has the following fields:

<b>Field</b>	<b>Description</b>
Partnum	The partition number for an index.
Mode	The alice mode for the current partition.
BM_Sz	The size allocated for the bitmap.
Used_Pg	The size of the index in pages (used)
Examined	Number of pages examined during cleaning.
Dirty_Pg	Number of dirty pages
# I/O	Number of pages read
Found	Number of dirty pages found in reads.
Eff	How efficient was the bitmap.



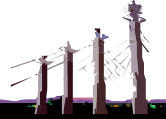
www.iiug.org



## Monitoring BTSCANNER

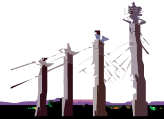
The **onstat -C map** output has the following fields:

<b>Field</b>	<b>Description</b>
Partnum	The partition number for an index
Key	Index Key,
Map	Alice bitmap



## BTSCANNER Tips

- Avoid leaf scanning, if possible.
- Use onstat -C part to identify unused indices, and problem indices.
- In versions prior to 10.00.F7 watch closely to run away BTSCANNER threads.



Session D07  
BTSCANNER

Mark Jamison  
IBM  
[mjamison@us.ibm.com](mailto:mjamison@us.ibm.com)

