
Rule DAS622: The number of index buffers should be increased (STRNO value)

Finding: The number of index buffers was not greater than the number of strings specified in the Access Control Block (ACB) STRNO value. The number of index buffers should be increased for optimal performance of the VSAM data sets listed. This finding applies only if SMF Type 42 (Data Set Statistics)¹ and SMF Type 64 (VSAM Statistics) records are available in a MXG performance data base.

Impact: This can have a LOW IMPACT, MEDIUM IMPACT, or HIGH IMPACT on the performance of applications referencing the VSAM data. The level of impact depends on the number of direct I/O operations.

Discussion: A VSAM file structure consists of one or more *Control Intervals (CIs)* and one or more *Control Areas (CAs)*.

- A **Control Interval** is a continuous area of direct access storage that VSAM uses to store logical records. When a logical record is read from direct access storage, the entire Control Interval containing the record is read into a VSAM buffer in virtual storage. The desired logical record is then transferred from the VSAM buffer to a user-defined buffer or work area. While logical records within a Control Interval may vary in length, all Control Intervals in a specific VSAM data set are of the same length.
- A **Control Area** contains one or more Control Intervals. The Control Intervals are grouped together into fixed-length contiguous areas of direct access storage. A VSAM data set is composed of one or more Control Areas.

Each VSAM data set is defined as a cluster of one or more components.

- The *data component* is the part of a VSAM data set, alternate index, or catalog that contains the data records. The minimum size of a Control Area for a data component is one track, and the maximum size is one cylinder of DASD storage.
- The *index component* is a collection of logically sequenced keys. A key is a value taken from a fixed defined field in each logical record in the VSAM data set. The key identifies the record's position in the data set. Using the index, VSAM is able to randomly retrieve a record from the data component when a request is made for a record with a certain key.

¹%LET TYPE42DS = Y; must be specified in USOURCE(GENGUIDE) must be specified in USOURCE(GENGUIDE) or USOURCE(DASGUIDE) to advise CPEXpert that TYPE42DS is available.

The size of the Control Area for an index component is one track of DASD storage.

Key-sequenced data sets (KSDS) and variable-length relative records data sets (VRRDS) contain both a data component and an index component. Additionally, each alternate index contains a data component and an index component. Entry-sequenced data sets (ESDS), linear data sets (LDS), and fixed-length relative record data sets (RRDS) contain only a data component.

The index component consists of two parts: *sequence set* and *index set*.

- The sequence set is the lowest level of index control intervals and directly points to the data Control Interval in the data Control Area. There is one Control Interval in the sequence set for each *data* Control Area. This *index* Control Interval contains pointers and high key information for each *data* Control Interval. This index Control Interval also contains horizontal pointers from one sequence set Control Interval to the next higher keyed sequence set Control Interval.
- The records in all levels of the index above the sequence set are called the index set. If there is more than one sequence set Control Interval, VSAM automatically builds another index level. An entry in an index set record consists of the highest possible key in an index record in the next lower level, and a pointer to the beginning of that index record.

I/O buffers are used by VSAM to read and write control intervals from DASD to virtual storage. A minimum of one buffer is required for an index Control Interval. Having only one index I/O buffer does not hinder performance when the VSAM data set is accessed sequentially, because VSAM gets to the next Control Interval by using the horizontal pointers in sequence set records rather than the vertical pointers in the index set.

A string is a request to a VSAM data set requiring data set positioning. If different *concurrent* accesses to the same data set are necessary, multiple strings are used. If multiple strings are used, each string requires exclusive control of an index I/O buffer. Therefore, the value specified for the STRNO parameter (in the ACB or GENCB macro, or AMP parameter) is the minimum number of index I/O buffers required when requests that require concurrent positioning are issued.

If the number of I/O buffers provided for index records is greater than the number of requests that require concurrent positioning (specified by the STRNO parameter), one buffer is used for the highest-level index record. Any additional buffers are used, as required, for other index-set index records. With direct access, the minimum number of index buffers should be one more than the value of the STRNO if VSAM is to keep the

highest-level index record resident. Keeping the highest-level index record resident can significantly improve performance, at a modest increase in virtual storage used for index buffers

For optimum performance, the number of index buffers should be at least as large as the number of high-level index set Control Intervals, **plus one per string to contain the entire high-level index set and one sequence set control interval per string in virtual storage.**

After applying the screening criteria specified for VSAM data sets, and extracting SMF Type 64 information for those VSAM data sets, CPEXpert examines SMF Type 42 (Data Set Statistics) information for the selected VSAM data sets. CPEXpert uses the TYPE42DS information to compute the percent of accesses to the VSAM data set that were direct. CPEXpert uses the TYPE42DS information to compute the percent of direct accesses to the VSAM data set, using the following algorithm:

$$\text{Percent direct accesses} = \frac{S42AMDRB}{S42AMSRB + S42AMDRB}$$

where: S42AMSRB = Blocks read using sequential access

S42AMDRB = Blocks read using direct access

CPEXpert produces Rule DAS622 under the following conditions:

- The TYPE42DS S42DSBUF variable showed that NSR was used for KSDS or VRRDS VSAM data sets, **and**
- The percent of direct accesses for the index component was greater than **DIRINDEX** guidance variable in USOURCE(DASGUIDE), **and**
- The STRNO specification in the ACB (the MXG ACBSTRNO) was greater than one (indicating that concurrent accesses had been specified for direct processing), **and**
- The number of buffers (the MXG BUFDRNO variable) assigned to the index component was less than the ACBSTRNO value, plus 1.

The default value for the DIRINDEX guidance variable is 25%, so CPEXpert will produce Rule DAS622 for NSR VSAM data sets when more than 25% of the accesses were direct for the index component and the number of buffers assigned to the index component was less than the STRNO value, plus 1.

The following example illustrates the output from Rule DAS622:

RULE DAS622: THE NUMBER OF INDEX BUFFERS SHOULD BE INCREASED (STRNO VALUE)

VOLSER: MVS902. Non-Shared resources (NSR) was specified as the buffering technique for the below VSAM data sets, and most of the access were direct processing. However, relatively few index buffers were assigned to the data sets. You should consider increasing the number of index buffers to 1 more than the number of strings specified in the STRNO in the ACB. The I/O RATE is for the time the data set was open.

SMF TIME STAMP	JOB NAME	VSAM DATA SET	I/O RATE	-ACCESS TYPE (PCT)-		BUFFERS	
				SEQUENTIAL	DIRECT	STRINGS	ASSIGNED
10:10,19SEP2002	NDMMON..	SDPDPA.PK.MVSP.RT.NDMTCF.INDEX.....	39.2	0.0	100.0	5	5
10:10,19SEP2002	PCGORDER	SDPDPA.PK.MVSP.RT.NDMDRD.INDEX.....	12.1	0.0	100.0	3	3
10:10,19SEP2002	NDMMON..	SDPDPA.PK.MVSP.RT.NDMTCF.INDEX.....	242.8	0.0	100.0	5	5
10:14,19SEP2002	PGDOS06R	SDPDPA.PK.MVSP.CS.NDMDRD.INDEX.....	12.9	0.0	100.0	3	3
10:14,19SEP2002	PGDOS06R	SDPDPA.PK.MVSP.CS.NDMDRD.INDEX.....	16.5	0.0	100.0	3	3

While not shown in the above example, CPEXpert also shows the OPEN time for the VSAM data set. This normally is the duration of the current OPEN. If %LET VSAMSMRY=Y; was specified in USOURCE(DASGUIDE), the OPEN time represents the sum of the times the VSAM data was OPEN for all TYPE64 records in the performance data base.

Suggestion: If CPEXpert produces Rule DAS622, you should consider the following alternatives:

- **Increase the number of buffers.** You should consider increasing the number of index buffers for the VSAM data sets identified by Rule DAS622 to be one greater than the number of concurrent accesses (as specified by the STRNO parameter).
- Alternatively, you can specify System Managed Buffering (SMB)² for the VSAM data set. VSAM can use system-managed buffering to determine the number of buffers and the type of buffer management to use for VSAM data sets. VSAM also determines the number of buffers to locate in HiperSpace for use in direct optimization. To indicate that VSAM is to use SMB, specify either of the following options:
 - Specify the ACCBIAS subparameter of the JCL DD statement AMP parameter and specify Sequential Optimized (SO) for the record access bias.
 - Specify Record Access Bias in the data class and an application processing option in the ACB.

²Please review *DFSMS: Using Data Sets* (Section 2.5.4.2.3: Processing Guidelines and Restrictions) before implementing system managed buffering.

For system-managed buffering (SMB), the data set must use both of the following options:

- System Management Subsystem (SMS) storage
- Extended format (DSNTYPE=ext in the data class)
- If the above actions are not appropriate, you can change the **DIRINDEX** guidance variable in USOURCE(DASGUIDE). Section 3 describes how to change the DIRINDEX guidance variable if you feel that Rule DAS622 is produced prematurely.
- Alternatively, you can exclude the reported VSAM data sets from analysis. Section 3 describes how to exclude VSAM data sets from analysis. However, you should be aware that no analysis of potential VSAM problems will be performed on data sets that are excluded from analysis.

Reference: *DFSMS: Using Data Sets* (SC26-7339 for OS/390; SC26-7410 for z/OS)
Section 2.5.4.4.2: Index Buffers for Direct Access
Section 2.5.4.2: Tuning for System-Managed Buffering

IBM Redbook: *VSAM Demystified* (SG24-6105)
Section 2.6.9 (Buffering options)