DECOMPRESS: Decompress an Adabas File

The DECOMPRESS function decompresses data either

- from output unloaded by the ADAULD UNLOAD utility function; or

- directly from a single compressed Adabas file when the file number is specified with the INFILE parameter.

When decompressing data directly from the INFILE file, DECOMPRESS first performs an ADAULD UNLOAD/MODE=SHORT function. This can save time over separate ADAULD and ADACMP DECOMPRESS operations.

```
ADACMP  DECOMPRESS  [CODE = cipher-code ]
                    [FORMAT = output-record-format-definition ]
                    [HEADER = { YES | NO } ]
                    [INFILE = file-number ]
                       [ETID = owner-id ]
                       [LPB = prefetch-buffer-size ]
                       [PASSWORD = ' password ' ]
                       [SORTSEQ = { descriptor [,NU] | ISN } ]
                       [UTYPE = { EXF | EXU } ]
                    [ISN]
                    [LOBVALUES = { YES | NO } ]
                    [MAXLOGRECLEN = buffer-size ]
                    [NOUSERABEND]
                    [NUMREC = number-of-records ]
                    [TRUNCATE]
                    [TZ = ' timezone-name ' [,DST] ]
                    [UACODE = userdata-alpha-key ]
                    [UARC = {architecture-key | 2 } ]
                    [UWCODE = userdata-wide-key ]
```

This chapter describes the syntax and parameters of the ADACMP DECOMPRESS function.

- Optional Parameters and Subparameters

- Decompressing Multiclient Files

- ADACMP DECOMPRESS Examples

Optional Parameters and Subparameters

CODE: Cipher Code

> If the file to be decompressed is ciphered, the cipher code that was used when the file was compressed must be specified with this parameter. See the *Adabas Security* documentation for additional information on the use of ciphering.

DST: Daylight Savings Indicator

The DST parameter can be specified to indicate that date-time data includes a daylight savings time indicator. If a time zone uses daylight savings time, you must be sure to store and retrieve the daylight savings indicator with your date-time data or there will be no way to distinguish date-time values in the hour before the time is switched back to standard time. The two-byte daylight savings indicator directly follows the date-time value in uncompressed input and specifies the hexadecimal value of the daylight saving time offset from standard time in seconds.

The DST parameter requires that the TZ parameter be set in the same ADACMP run. However, it should *not* be specified when the FORMAT parameter is specified in the run.

You must specify a DST parameter for files containing date-time fields defined with option TZ, when the time zone includes a daylight savings time indicator. If the DST parameter is not specified, date-time data is stored without a daylight savings time indicator. The default is store the date-time data without a daylight savings time indicator.

ETID: Multiclient File Owner ID

ETID specifies an owner ID for a multiclient file specified by INFILE. ADACMP DECOMPRESS selectively decompresses only those records in the multiclient file assigned to the owner ID specified by ETID. The ETID value must be the same as that assigned to the records when they were loaded into the multiclient file.

FORMAT: Output Record Format Definition

FORMAT allows decompression to a format other than that specified by the FDT. It can be used to change the FDT of an existing file and, in particular, the structure of a periodic (PE) group.

The FORMAT parameter syntax is the same as the format buffer syntax used for read commands except that text cannot be inserted (text is not compressible/decompressible); see the *Adabas Command Reference Guide* documentation for more information.

Note:
The FORMAT parameter does not check whether all related data fields have been processed during decompression.

For example, if a multiple-value (MU) field defined as:

```
01,AA,8,A,MU
```

has five occurrences, and the ADACMP DECOMPRESS FORMAT parameter specifies:

```
AA1-4
```

then only the first four AA field values are decompressed; no indication is given regarding the fifth field value. This also applies to PE field occurrences and length overrides.

HEADER

This optional parameter indicates whether or not the ADACMP decompression logic should produce the ADACMP segmented record headers (ADAH and ADAC) as part of the decompressed output. Valid values are YES or NO; the default is NO.

HEADER=NO is the format accepted and produced by ADACMP in Adabas versions prior to Adabas 8. When it is specified, the decompressed output records produced by ADACMP will contain only the data for the fields of the file being processed. Each data record must fit into one physical record of the sequential input data set (less than 32 KB). If the data exceeds this size, the records in error are written to the DDFEHL error data set.

HEADER=YES can only be specified if you are running Adabas 8. If HEADER=YES is specified, each output decompressed record produced by ADACMP begins with either an ADAH or ADAC header, relating the physical record with a logical record. Each logical record can be larger than 32 KB. The header in each physical record defines the position of the data following it within the logical record. For complete information about segmented records in ADACMP, read *Segmented Record Considerations*. DSECTs for the ADAH and ADAC headers can be found in members ADAH and ADAC of the distributed Adabas SRCE data set.

INFILE: Number of File to Be Decompressed

The INFILE parameter allows you to decompress a file without first unloading it with the ADAULD utility. If the INFILE parameter is not specified, the input is read from a sequential (DD/EBAND) file. With the ETID parameter, INFILE permits selectively decompressing records from a multiclient file. When decompressing multiclient files, refer to the section *Decompressing Multiclient Files*.

ISN: Include ISN in Decompressed Output

The ISN of each record is to be included with each decompressed record output. If this parameter is omitted, the ISN will not be included with each record.

If ISN is specified with HEADER=YES, the ISN immediately follows the ADAH header as part of the logical record. The DSECT for the ADAH header can be found in member ADAH of the distributed Adabas SRCE data set.

LPB: Prefetch Buffer Size

LPB specifies the size, in bytes, of the internal prefetch buffer for the ADACMP DECOMPRESS INFILE function. The maximum value is 32,760 bytes. The default is calculated by Adabas, depending on the ADARUN LU value in effect for the nucleus.

LOBVALUES: LB Field Size Indicator

The LOBVALUES parameter should only be specified for files containing large object (LB) fields.

This optional parameter indicates whether long LB field values (larger than 253 bytes) or short LB field values (up to 253 bytes) are expected in the ADACMP DECOMPRESS output data. Valid values for this parameter are "YES" and "NO"; the default is "NO".

If "NO" is specified for this parameter, the uncompressed output data may contain only LB fields up to 253 bytes long and references to LB field values stored in a *LOB file*. In this case, you cannot specify an LB field in the ADACMP DECOMPRESS FORMAT parameter. During processing,

ADACMP DECOMPRESS reads only the base file records as input; if the base file contains references to LB field values in a LOB file, ADACMP DECOMPRESS does not read them, but only reproduces the references in the output.

If "YES" is specified for this parameter, the uncompressed output data will contain the LB field values present for the record. In this case, the INFILE parameter must also be specified to identify the file number of the *base file* of a *LOB file group* whose data is to be decompressed. During processing, as ADACMP DECOMPRESS reads and decompresses the records from the base file, it reads any referenced LB field values from the *LOB file*.

Note:
An ADACMP DECOMPRESS function with LOBVALUES=NO followed by an ADACMP COMPRESS function with LOBVALUES=NO can be used to modify the FDT of the *base file* in the *LOB file group*.

MAXLOGRECLEN: Buffer Size

This optional parameter can be used to specify the size, in bytes, of a buffer used by ADACMP to assemble logical records that span one or more physical records of uncompressed output data. This buffer is allocated only if HEADER=YES is also specified. Otherwise, the setting of MAXLOGRECLEN is ignored. The default value of MAXLOGRECLEN is 1,048,576 bytes (1 MB).

If the value specified by MAXLOGRECLEN is appended by the letter "K", it is multiplied by 1024. The minimum value is 32768 bytes.

NOUSERABEND: Termination without Abend

When an error is encountered while the function is running, the utility prints an error message and terminates with user abend 34 (with a dump) or user abend 35 (without a dump). When NOUSERABEND is specified, it must be specified as the first parameter (before all other parameters) for the utility function.

If NOUSERABEND is specified, the utility will *not* abend after printing the error message. Instead, the message "utility TERMINATED DUE TO ERROR CONDITION" is displayed and the utility terminates with condition code 20.

NUMREC: Number of Records to Be Processed

NUMREC specifies the number of input records to be processed. If this parameter is omitted, all input records contained on the input data set are processed.

Use of NUMREC is recommended for the initial ADACMP execution if a large number of records are contained on the input data set. This avoids unneeded processing of all records when a field definition error or invalid input data causes a large number of rejected records. NUMREC is also useful for creating small files for test purposes.

PASSWORD: Password for INFILE

The PASSWORD parameter must specify the correct password if the file is to be decompressed directly from a password-protected Adabas file.

SORTSEQ: Processing Sequence for INFILE File

SORTSEQ determines the sequence in which the file is processed. If this parameter is omitted, the records are processed in physical sequence. SORTSEQ can be specified only when INFILE is also specified.

If a descriptor is specified, the file is processed in the logical sequence of the descriptor values. *Do not* use a hyperdescriptor, a phonetic descriptor, a multiple-value descriptor field, or a descriptor contained in a periodic group.

If the descriptor name refers to a field defined with the null suppression (NU) option, you must specify ",NU" after the descriptor name. In this case, records of the descriptor that contain null values are not decompressed. If NU is not specified in this case (the default), ADACMP rejects NU descriptors.

Note:
Even when the descriptor field is not null-suppressed, the record is *not* represented in the inverted list if the descriptor field or a field following it has never been initialized (held a value). Therefore, the record will be dropped when the utility is executed.

If ISN is specified, the file is processed in ascending ISN sequence. For the Adabas checkpoint or security file, only SORTSEQ=ISN is allowed.

TRUNCATE: Truncate Excess Alphanumeric Characters

The TRUNCATE parameter enables truncation of compressed alphanumeric data during decompression. When TRUNCATE is specified and ADACMP DECOMPRESS operation finds an alphanumeric field containing more characters than the FDT description allows for the field, the extra characters are truncated. If TRUNCATE is not specified, alphanumeric records with extra characters are written to the DDFEHL data set. Non-alphanumeric fields cannot be truncated.

TZ: Time Zone

The TZ parameter can be used to specify the local time zone. As records are decompressed and read from the file, the date-time data is converted from UTC time (Coordinated Universal Time, also known as Greenwich Mean Time) to the corresponding time for the specified time zone. Date-time field data is always stored in UTC time in an Adabas file. Valid time zone names are listed in the TZINFO member of the Adabas time zone library and are specified in single quotes. The following is an example of a valid TZ specification:

```
TZ='America/New_York'
```

Note:
Time zone names are case-sensitive.

This example sets the time zone to Eastern US. When decompressing input, local time is assumed to be Eastern US. When retrieving the stored UTC time data, it is converted to Eastern US time.

Adabas uses the time zone data taken from the tz database, which is also called the *zoneinfo* or *Olson* database. The specific list of time zone names that Adabas supports in any given release can be found in the TZINFO member of the Adabas time zone library (ADA*vrs*.TZ00). For more information about the TZINFO member of the time zone library, read *Supported Time Zones*.

Note:
Review important information about the daylight savings time (DST) parameter. If a time zone uses daylight savings time, you must be sure to retrieve the daylight savings indicator with your date-time data or there will be no way to distinguish date-time values in the hour before the time is switched back to standard time.

UACODE: Encoding Protocol for Output Alphanumeric Fields

UACODE defines the encoding of the sequential output of alphanumeric fields. This parameter allows you to override the user encoding for alphanumeric fields passed in the header of the compressed sequential input.

UARC: Architecture for Output Uncompressed User Data

The UARC parameter specifies the architecture of the sequential output of the uncompressed user data. This parameter allows you to override the user encoding passed in the header of the compressed sequential input.

The 'userdata-architecture-key' is an integer which is the sum of the following numbers:

byte order	b=0	high-order byte first
	b=1	low-order byte first
encoding family	e=0	ASCII encoding family
	e=2	EBCDIC encoding family (default)
floating-point format	f=0	IBM370 floating-point format
	f=4	VAX floating-point format
	f=8	IEEE floating-point format

The default is ARC = b + e + f = 2; that is, high-order byte first; EBCDIC encoding family; and IBM370 floating-point format (b=0; e=2; f=0).

User data from an Intel386 PC provides the example: b=1; e=0; f=8; or ARC=9.

UTYPE: User Type

The user type to be in effect when unloading the file specified by INFILE. Allowed values are

EXF	no access/update allowed for other users of the file.
EXU	access only is allowed for other users of the file. EXU is the default.

UWCODE: Encoding Protocol for Output Wide-Character Fields

UWCODE defines the encoding of the sequential output of wide-character fields. This parameter allows you to override the user encoding for wide-character fields passed in the header of the compressed sequential input.

Decompressing Multiclient Files

ADACMP decompresses Adabas data to a sequential user file. The DECOMPRESS function can decompress records selectively if the INFILE parameter specifies a multiclient file and a valid ETID value is specified.

The DECOMPRESS function skips the owner ID, if present. The output of a DECOMPRESS operation on a multiclient file contains neither owner ID nor any ETID information.

If the INFILE parameter specifies a multiclient file for the DECOMPRESS function, you can use the ETID parameter to limit decompression to records for a specific user only. ADACMP then reads and decompresses records only for the specified user. If the ETID parameter is not specified when decompressing a multiclient file, all records in the file are decompressed.

Example:

Only records owned by USER1 from file 20 are decompressed to a sequential output file:

```
 ADACMP   DECOMPRESS   INFILE=20,ETID=USER1
```

ADACMP DECOMPRESS Examples

Example 1:

The DECOMPRESS function is to be executed. The input data set to be used is the output of a previous execution of the ADAULD utility:

```
 ADACMP DECOMPRESS
```

Example 2:

Adabas file 23 is to be decompressed. The ISN of each record is to be included in the decompressed output:

```
 ADACMP DECOMPRESS INFILE=23,ISN
```