Compression

Data compression

- Compression is used to reduce the storage used by the tables.
- Teradata offers some Technique to compress the data:
  - Multi Value Compression(MVC)
  - Algorithmic Compression(ALC)
  - Block Level Compression(BLC)
- Primary Index column cannot be compressed.
- Volatile tables cannot be compressed.

MVC

- It can compress up to 255 distinct values including NULL.
- Can be added at table creation using `CREATE TABLE` or after table creation using `ALTER TABLE`.
- When compression is applied on a column, the values for this column are not stored with the row. Instead the values are stored in the Table header in each AMP and only presence bits are added to the row to indicate the value.
- **No overhead**

Example

```
CREATE SET TABLE employee
(
EmployeeNo integer,
FirstName CHAR(30),
LastName CHAR(30),
BirthDate DATE FORMAT 'YYYY-MM-DD-',
JoinedDate DATE FORMAT 'YYYY-MM-DD-',
Gender CHAR(1) ,
DepartmentNo CHAR(02) COMPRESS(1,2,3)
)
UNIQUE PRIMARY INDEX(EmployeeNo);
```

ALC

- Compression using algorithms instead of simple one-hot encoding.
- No limit of 255 values.
- ALC is limited to the `CHAR`, `VARCHAR` and `BYTE` data type while MVC can incorporate all the numeric as well as date data types.
- There is overhead involved to compress and decompress data in ALC, but in case of MVC there is no such kind of overhead.
- Recommended for large character columns which are not accessed often.

ALC - Example

```
CREATE TABLE Student
    (Roll_No       INTEGER,
    Student_Name VARCHAR(50),
    Student_Address CHAR(200) CHARACTER SET UNICODE
    COMPRESS USING TransUnicodeToUTF8
    DECOMPRESS USING TransUTF8ToUnicode)
UNIQUE PRIMARY INDEX(Roll_No);
```

BLC

- Compress the table by blocks of rows.
- To access a row, the entired block needs to be decompressed.
- Does not carry over across sessions.
- In case of MVC and ALC values need to be define in CREATE TABLE statement, but BLC is activated outside of table definition.
- More space reduction compared to MVC and ALC (up to 60%).
- Recommended mostly for *cold data* as there is overhead in compressing/decompressing.

BLC - Example

BLC can be applied in two ways:

- For an empty table query band can be used to apply BLC.

```
/* Turn on BLC */
SET QUERY_BAND = 'BLOCKCOMPRESSION=YES;' FOR SESSION;
/* Insert data into empty table */
INSERT INTO STUDENT_MVC AS SELECT * FROM STUDENT;
/* Turn off BLC */
SET QUERY_BAND = 'BLOCKCOMPRESSION=NO;' FOR SESSION;
```

- For a non-empty table, Ferret utility can be used to either compress all the data block in its or to decompress it.