

Optimize MERGE Conversion Script

This is the full test script we used to convert a non-optimized MERGE plan into an optimized plan.

```
cat make_merge_optimized01.sql
\echo This test will UPDATE REC2UPDATE records and INSERT REC2INSERT records x 2 fields,
represented in the SOURCE table
\echo into the TARGET table which includes dozens of fields.
\echo
\echo Updating only two fields does not meet the optimized MERGE INTO query plan conditions,
\echo Therefore, we create a temporary table with all field values, in preparation for UPDATE
\echo and INSERT clauses to include every column in the target table.

-- Assign Variables
-- Records to INSERT in this test
\set REC2INSERT 1000000
-- Records to UPDATE in this test
\set REC2UPDATE 1000000
-- Find and set max user_id for the INSERT test below
\set MaxUserID '' `vsq1 -t -c 'SELECT MAX(user_id) FROM MYSCHEMA.TARGET;` ''

-- Cleanup before creating objects
DROP TABLE IF EXISTS MYSCHEMA.SOURCE CASCADE;
DROP TABLE IF EXISTS MYSCHEMA.TMPSOURCE CASCADE;
DROP TABLE IF EXISTS MYSCHEMA.TMPISOURCE CASCADE;

-- Create objects
CREATE TABLE MYSCHEMA.SOURCE(user_id int, user_first_name varchar(64), user_last_name varchar(64));
CREATE PROJECTION SOURCE_b0 (
    user_id ENCODING COMMONDELTA_COMP,
    user_first_name ENCODING AUTO,
    user_last_name ENCODING AUTO
)
AS SELECT user_id, user_first_name, user_last_name FROM MYSCHEMA.SOURCE
ORDER BY user_id, user_first_name, user_last_name
SEGMENTED BY HASH (user_id) ALL NODES OFFSET 0;

CREATE PROJECTION SOURCE_b1 (
    user_id ENCODING COMMONDELTA_COMP,
    user_first_name ENCODING AUTO,
    user_last_name ENCODING AUTO
)
AS SELECT user_id, user_first_name, user_last_name FROM MYSCHEMA.SOURCE
ORDER BY user_id, user_first_name, user_last_name
SEGMENTED BY HASH (user_id) ALL NODES OFFSET 1;
SELECT refresh();

-- Prepare records for INSERT
INSERT INTO MYSCHEMA.SOURCE
SELECT user_id + :MaxUserID::INT,
CONCAT(left(user_first_name,15),'FI'),
CONCAT(left(user_last_name,15),'LI')
FROM MYSCHEMA.MASTER
WHERE length(user_first_name) > 2 LIMIT :REC2INSERT;

-- Prepare records for UPDATE
INSERT INTO MYSCHEMA.SOURCE
```

```

SELECT user_id,
CONCAT(left(user_first_name,15),'FU'),
CONCAT(left(user_last_name,15),'LU')
FROM MYSCHEMA.MASTER
WHERE length(user_first_name) > 2 LIMIT :REC2UPDATE;
SELECT ANALYZE_HISTOGRAM('VDB.MYSCHEMA.SOURCE',100);

----- TEST STARTS FROM HERE -----
\timing
-- Create MYSCHEMA.TMPSOURCE table for UPDATE records only
CREATE TABLE MYSCHEMA.TMPSOURCE(
  user_id int,
  user_first_name varchar(64),
  user_last_name varchar(64),
  f04 int, f05 int, f06 int, f07 int, f08 varchar(30),
  f09 varchar(64), f10 varchar(255), f11 varchar(8), f12 varchar(256),
  f13 varchar(30), f14 timestamp, f15 int, f16 numeric(20,2),
  f17 numeric(20,2), f18 timestamp, f19 int, f20 numeric(20,2),
  f21 timestamp, f22 timestamp, f23 int, f24 int,
  f25 timestamp, f26 timestamp, f27 date
);

INSERT INTO MYSCHEMA.TMPSOURCE SELECT
  s.user_id,
  s.user_first_name,
  s.user_last_name,
  t.f04, t.f05, t.f06, t.f07, t.f08, t.f09, t.f10,
  t.f11, t.f12, t.f13,t.f14, t.f15, t.f16, t.f17, t.f18,
  t.f19, t.f20, t.f21, t.f22, t.f23, t.f24, t.f25, t.f26, t.f27
WHERE s.user_id = t.user_id;

-- Create MYSCHEMA.TMPISOURCE table for INSERT records only
CREATE TABLE MYSCHEMA.TMPISOURCE(
  user_id int,
  user_first_name varchar(64),
  user_last_name varchar(64),
  f04 int, f05 int, f06 int, f07 int, f08 varchar(30),
  f09 varchar(64), f10 varchar(255), f11 varchar(8), f12 varchar(256),
  f13 varchar(30), f14 timestamp, f15 int, f16 numeric(20,2),
  f17 numeric(20,2), f18 timestamp, f19 int, f20 numeric(20,2),
  f21 timestamp, f22 timestamp, f23 int, f24 int, f25 timestamp,
  f26 timestamp, f27 date
);

INSERT INTO MYSCHEMA.TMPISOURCE SELECT
  s.user_id, s.user_first_name, s.user_last_name
FROM MYSCHEMA.SOURCE s
WHERE s.user_id NOT IN (SELECT user_id from MYSCHEMA.TMPSOURCE);
SELECT COUNT(*) FROM MYSCHEMA.MASTER;
SELECT COUNT(*) FROM MYSCHEMA.TARGET;
SELECT COUNT(*) FROM MYSCHEMA.SOURCE;
SELECT COUNT(*) FROM MYSCHEMA.TMPSOURCE;
SELECT COUNT(*) FROM MYSCHEMA.TMPISOURCE;

\echo Optimized execution plan for a MERGE statement:
\o | grep --color 'Semi'
EXPLAIN MERGE INTO MYSCHEMA.TARGET TGT
  USING MYSCHEMA.TMPSOURCE SRC ON SRC.user_id=TGT.user_id
WHEN MATCHED THEN UPDATE SET
  user_id=SRC.user_id,

```

```

    user_first_name=SRC.user_first_name,
    user_last_name=SRC.user_last_name,
    f04=SRC.f04, f05=SRC.f05, f06=SRC.f06, f07=SRC.f07, f08=SRC.f08,
    f09=SRC.f09, f10=SRC.f10, f11=SRC.f11, f12=SRC.f12, f13=SRC.f13,
    f14=SRC.f14, f15=SRC.f15, f16=SRC.f16, f17=SRC.f17, f18=SRC.f18,
    f19=SRC.f19, f20=SRC.f20, f21=SRC.f21, f22=SRC.f22, f23=SRC.f23,
    f24=SRC.f24, f25=SRC.f25, f26=SRC.f26, f27=SRC.f27
WHEN NOT MATCHED THEN INSERT VALUES(
    SRC.user_id,
    SRC.user_first_name,
    SRC.user_last_name,
    SRC.f04, SRC.f05, SRC.f06, SRC.f07, SRC.f08, SRC.f09, SRC.f10,
    SRC.f11, SRC.f12, SRC.f13, SRC.f14, SRC.f15, SRC.f16,
    SRC.f17, SRC.f18, SRC.f19, SRC.f20, SRC.f21, SRC.f22,
    SRC.f23, SRC.f24, SRC.f25, SRC.f26, SRC.f27
);
\o

\echo Non-optimized execution plan for a MERGE statement,
\echo Since "user_last_name=SRC.user_last_name," is ommited from the UPDATE:
\o | grep --color 'RightOuter'
EXPLAIN MERGE INTO MYSCHEMA.TARGET TGT USING MYSCHEMA.TMPSOURCE SRC ON
    SRC.user_id=TGT.user_id
WHEN MATCHED THEN UPDATE SET
    user_id=SRC.user_id,
    user_first_name=SRC.user_first_name,
    f04=SRC.f04, f05=SRC.f05, f06=SRC.f06, f07=SRC.f07, f08=SRC.f08,
    f09=SRC.f09, f10=SRC.f10, f11=SRC.f11, f12=SRC.f12, f13=SRC.f13,
    f14=SRC.f14, f15=SRC.f15, f16=SRC.f16, f17=SRC.f17, f18=SRC.f18,
    f19=SRC.f19, f20=SRC.f20, f21=SRC.f21, f22=SRC.f22, f23=SRC.f23, f24=SRC.f24,
    f25=SRC.f25, f26=SRC.f26, f27=SRC.f27
WHEN NOT MATCHED THEN INSERT VALUES(
    SRC.user_id,
    SRC.user_first_name,
    SRC.user_last_name,
    SRC.f04, SRC.f05, SRC.f06, SRC.f07, SRC.f08, SRC.f09, SRC.f10,
    SRC.f11, SRC.f12, SRC.f13, SRC.f14, SRC.f15, SRC.f16,
    SRC.f17, SRC.f18, SRC.f19, SRC.f20, SRC.f21, SRC.f22, SRC.f23, SRC.f24,
    SRC.f25, SRC.f26, SRC.f27
);
\o

-- Start UPDATE -- Optimized Merge
MERGE INTO MYSCHEMA.TARGET TGT USING MYSCHEMA.TMPSOURCE SRC
    ON SRC.user_id=TGT.user_id
WHEN MATCHED THEN UPDATE SET
    user_id=SRC.user_id,
    user_first_name=SRC.user_first_name,
    user_last_name=SRC.user_last_name,
    f04=SRC.f04, f05=SRC.f05, f06=SRC.f06, f07=SRC.f07, f08=SRC.f08,
    f09=SRC.f09, f10=SRC.f10, f11=SRC.f11, f12=SRC.f12, f13=SRC.f13, f14=SRC.f14,
    f15=SRC.f15, f16=SRC.f16, f17=SRC.f17, f18=SRC.f18, f19=SRC.f19, f20=SRC.f20,
    f21=SRC.f21, f22=SRC.f22, f23=SRC.f23, f24=SRC.f24,
    f25=SRC.f25, f26=SRC.f26, f27=SRC.f27
WHEN NOT MATCHED THEN INSERT VALUES(
    SRC.user_id,
    SRC.user_first_name,
    SRC.user_last_name,
    SRC.f04, SRC.f05, SRC.f06, SRC.f07, SRC.f08, SRC.f09, SRC.f10,
    SRC.f11, SRC.f12, SRC.f13, SRC.f14, SRC.f15, SRC.f16,

```

```

SRC.f17, SRC.f18, SRC.f19, SRC.f20, SRC.f21, SRC.f22, SRC.f23, SRC.f24,
SRC.f25, SRC.f26, SRC.f27
);

-- Start INSERT -- Optimized Merge
MERGE INTO MYSCHEMA.TARGET TGT USING MYSCHEMA.TMPISOURCE SRC
  ON SRC.user_id=TGT.user_id
WHEN MATCHED THEN UPDATE SET
  user_id=SRC.user_id,
  user_first_name=SRC.user_first_name,
  user_last_name=SRC.user_last_name,
  f04=SRC.f04, f05=SRC.f05, f06=SRC.f06, f07=SRC.f07, f08=SRC.f08, f09=SRC.f09,
  f10=SRC.f10, f11=SRC.f11, f12=SRC.f12, f13=SRC.f13, f14=SRC.f14, f15=SRC.f15,
  f16=SRC.f16, f17=SRC.f17, f18=SRC.f18, f19=SRC.f19, f20=SRC.f20,
  f21=SRC.f21, f22=SRC.f22, f23=SRC.f23, f24=SRC.f24, f25=SRC.f25, f26=SRC.f26, f27=SRC.f27
WHEN NOT MATCHED THEN INSERT VALUES(
  SRC.user_id,
  SRC.user_first_name,
  SRC.user_last_name,
  SRC.f04, SRC.f05, SRC.f06, SRC.f07, SRC.f08, SRC.f09, SRC.f10,
  SRC.f11, SRC.f12, SRC.f13, SRC.f14, SRC.f15, SRC.f16, SRC.f17, SRC.f18,
  SRC.f19, SRC.f20, SRC.f21, SRC.f22, SRC.f23, SRC.f24, SRC.f25,
  SRC.f26, SRC.f27
);

-- Start INSERT and UPDATE -- Un-Optimized Merge
MERGE INTO MYSCHEMA.TARGET TGT USING MYSCHEMA.SOURCE SRC
  ON SRC.user_id=TGT.user_id
WHEN MATCHED THEN UPDATE SET
  user_id=SRC.user_id,
  user_first_name=SRC.user_first_name,
  user_last_name=SRC.user_last_name
WHEN NOT MATCHED THEN INSERT VALUES(
  SRC.user_id,
  SRC.user_first_name,
  SRC.user_last_name
);

\echo ##### Test END #####

```