

Using CoSort for Change Data Capture (CDC) Reports

With [CoSort's](#) Sort Control Language ([SortCL](#)) program, you can identify and [report](#) on additions, deletions, and modifications to data in files. This changed data capture (or [delta reporting](#)) example demonstrates SortCL's join functionality. Based on customer data that changes on a daily basis, a full outer join is used to produce a report that shows differences between previous and current sets of data.

The input files “*previous*” and “*current*” have both been presorted on their DEBT_ID fields. The final report displays the following information in the last column of each record:

- Insert (**Ins**), where a current record exists for a given DEBT_ID value where a previous record did not (that is, a non-matched record).
- Delete (**Del**), where a previous record exists for a given DEBT_ID value where a current record does not (that is, a non-matched record).
- Update (**Upd**), where the Checksum value of the previous and current matched record is changed (requires a matching record).

No change (), where the Checksum value of the previous and current record are the same (requires a matching record)

A subset of the presorted file *previous* is as follows:

ZUNTG341374451	271180
ZUNTG341374452	270626
ZUNTG341374453	298035
ZUNTG341374454	269036
ZUNTG341374455	281366
ZUNTG341374456	272174

A subset of the presorted file *current* is as follows:

ZUNTE341371586	8346534510	305232	23/11/2001	52.97	10/12/2001	52.97
ZUNTE341371587	8344660410	287910	23/11/2001	93.54	10/12/2001	93.54
ZUNTE341371588	8344660510	298545	23/11/2001	142.71	10/12/2001	142.71
ZUNTE341371589	8344660610	307105	23/11/2001	5.37	10/12/2001	5.37
ZUNTE341371590	8344660712	299498	23/11/2001	314.92	10/12/2001	314.92
ZUNTE341371591	8344660815	282840	23/11/2001	61.38	10/12/2001	61.38
ZUNTE341371592	8343919615	283904	23/11/2001	59.86	10/12/2001	59.86

The following SortCL script uses a full outer join and a conditional field, `delta_flag`, that is appended as the last column in the output:

```

/INFILE=previous      # ordered list, pre-sorted over DEBT_ID
/FIELD=(DEBT_ID,POS=1,SIZE=13)
/FIELD=(Checksum,POS=17,SIZE=6)
/CONDITION=(prev_exist,TEST=(previous.Checksum NE "      "))  # check for non-empty
/CONDITION=(prev_nonexist,TEST=(previous.Checksum EQ "      "))

/INFILE=current       # ordered list, pre-sorted over DEBT_ID
/FIELD=(DEBT_ID,POS=1,SIZE=13)
/FIELD=(CONSUMER,POS=17,SIZE=10)
/FIELD=(Checksum,POS=33,SIZE=6)
/FIELD=(transdate,POS=41,SIZE=10)
/FIELD=(amount,POS=52,SIZE=9,NUMERIC)
/FIELD=(date_due,POS=63,SIZE=10)
/FIELD=(paid,POS=74,SIZE=9,NUMERIC)
/CONDITION=(curr_exist,TEST=(current.Checksum NE "      "))  # check for non-empty
/CONDITION=(curr_nonexist,TEST=(current.Checksum EQ "      "))
/CONDITION=(update,TEST=(previous.Checksum NE current.Checksum))
/CONDITION=(no update,TEST=(previous.Checksum EQ current.Checksum))

```

```
/JOIN FULL OUTER previous current WHERE previous.DEBT_ID EQ current.DEBT_ID
```

```
/OUTFILE=delta.out
```

```
  /HEADREC="Prev: DebtID      Cksum      Curr: DebtID      Consumer      Cksum      TransDate      Amount      DateDue
Paid  Delta"
  /FIELD=(previous.DEBT_ID, POS=1, SIZE=13)
  /FIELD=(previous.Checksum, POS=17, SIZE=6)
  /FIELD=(current.DEBT_ID, POS=29, SIZE=13)
  /FIELD=(current.CONSUMER, POS=44, SIZE=10)
  /FIELD=(current.Checksum, POS=56, SIZE=6)
  /FIELD=(transdate, POS=64, SIZE=10)
  /FIELD=(amount, POS=76, SIZE=9, NUMERIC)
  /FIELD=(date_due, POS=87, SIZE=10)
  /FIELD=(paid, POS=99, SIZE=9, NUMERIC)
  /DATA="      "
  /FIELD=(delta_flag, POS=110, IF update AND prev_exist \
AND curr_exist THEN "Upd" \
ELSE IF prev_exist AND curr_nonexist THEN "Del" \
ELSE IF prev_nonexist AND curr_exist THEN "Ins" \
ELSE "") # conditional field logic for delta flag column
```

A subset of the output file, *delta.out*, is as follows:

Prev: DebtID	Cksum	Curr: DebtID	Consumer	Cksum	TransDate	Amount	DateDue	Paid	Delta
ZUNTG34137445	291385	ZUNTG34137445	8342769910	269036	15/01/2001	0.00	01/02/2001	0.00	Upd
ZUNTG34137445	291385	ZUNTG34137445	8342770110	281366	15/01/2001	12.83	01/02/2001	12.83	Upd
ZUNTG34137445	291385	ZUNTG34137445	8342770510	272174	15/01/2001	9.77	01/02/2001	9.77	Upd
		BNDZ21163338	2100115413	322734	10/06/2002	150.00	15/10/2002	150.00	Ins
		BNDZ21163349	2100572747	321929	14/09/2002	150.00	04/11/2002	150.00	Ins
		BNDZ21163370	2101808433	321963	12/09/2002	150.00	06/11/2002	150.00	Ins
		CR1Z4015759	4000785910	340645	16/09/2002	-100.00		-100.00	Ins
		CR1Z4015788	4002164110	457459	17/09/2002	-100.00	25/11/2002	-100.00	Ins
		CR1Z4015797	4000620810	624339	17/09/2002	-250.00	25/10/2002	-250.00	Ins
		CSHZ21385085	2101050610	268992	08/06/2001	-215.08	09/07/2001	-215.08	Ins
		CSHZ21385090	2101006401	268703	08/06/2001	-136.49	02/07/2001	-136.49	Ins
ZUNTG34137446	298257								Del
ZUNTG34137446	303976								Del
ZUNTG34137446	323950								Del
ZUNTG34137447	297622								Del
ZUNTG34137447	443942								Del
ZUNTG34137448	320098								Del
ZUNTE34769171	276502	ZUNTE34769171	8345997810	276502	09/01/2001	10.88	26/01/2001	10.88	
ZUNTE34769172	292611	ZUNTE34769172	8345997911	292611	09/01/2001	202.17	26/01/2001	202.17	
ZUNTE34769173	293219	ZUNTE34769173	8346709112	293219	09/01/2001	142.80	26/01/2001	142.80	
ZUNTE34769174	314550	ZUNTE34769174	8345998013	314550	09/01/2001	128.03	26/01/2001	128.03	

As you can see, the full outer join produces both left outer join results (the **Del** case) and right outer join results (the **Ins** case). The full outer join also returns the matches, where the **Upd** and "" (empty) cases are determined -- by comparing the Checksum values of the matching records (shown in columns 2 and 5 of the output).

If you have any questions about how this, or another delta reporting, application can work in your environment, please contact support@cosort.com or call 1-800-333-SORT, main menu option 3.