View Maintenance in a Warehousing Environment*

Yue Zhuge, Hector Garcia-Molina, Joachim Hammer, Jennifer Widom

Computer Science Department
Stanford University
Stanford, CA 94305-2140, USA
{zhuge,hector,joachim,widom}@cs.stanford.edu

Abstract

A warehouse is a repository of integrated information drawn from remote data sources. Since a warehouse effectively implements materialized views, we must maintain the views as the data sources are updated. This view maintenance problem differs from the traditional one in that the view definition and the base data are now decoupled. We show that this decoupling can result in anomalies if traditional algorithms are applied. We introduce a new algorithm, ECA (for "Eager Compensating Algorithm"), that eliminates the anomalies. ECA is based on previous incremental view maintenance algorithms, but extra "compensating" queries are used to eliminate anomalies. We also introduce two streamlined versions of ECA for special cases of views and updates, and we present an initial performance study that compares ECA to a view recomputation algorithm in terms of messages transmitted, data transferred, and I/O costs.

1 Introduction

Warehousing is an emerging technique for retrieval and integration of data from distributed, autonomous, possibly heterogeneous, information sources. A *data warehouse* is a repository of integrated information, available for queries and analysis (e.g., decision support, or data mining) [IK93]. As relevant information becomes available from a source, or when relevant information is modified, the information is extracted from the source, translated into a common model (e.g., the relational model), and integrated with existing data at the warehouse. Queries can be answered and analysis can be performed quickly and efficiently since the integrated information is directly available at the warehouse, with differences already resolved.

1.1 The Problem

One can think of a data warehouse as defining and storing integrated *materialized views* over the data from multiple, autonomous information sources. An important issue is the prompt and correct propagation of updates at the sources to the views at the warehouse. Numerous methods have been developed for materialized view maintenance in conventional database systems; these methods are discussed in Section 2.

*This work was supported by ARPA Contract F33615-93-1-1339, by the Anderson Faculty Scholar Fund, by the Center for Integrated Systems at Stanford University, and by equipment grants from Digital Equipment Corporation and IBM Corporation.

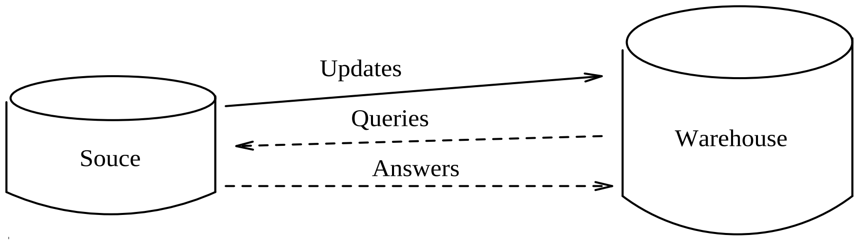

Figure 1.1: Processing of updates in a single source model

Unfortunately, existing materialized view maintenance algorithms fail in a warehousing environment. Existing approaches assume that each source understands view management and knows the relevant view definitions. Thus, when an update occurs at a source, the source knows exactly what data is needed for updating the view.

However, in a warehousing environment, the sources can be legacy or unsophisticated systems that do not understand views. Sources can inform the warehouse when an update occurs, e.g., a new employee has been hired, or a patient has paid her bill. However, they cannot determine what additional data may or may not be necessary for incorporating the update into the warehouse views. When the simple update information arrives at the warehouse, we may discover that some additional source data is necessary to update the views. Thus, the warehouse may have to issue queries to some of the sources, as illustrated in Figure 1.1. The queries are evaluated at the sources later than the corresponding updates, so the source states may have changed. This decoupling between the base data on the one hand (at the sources), and the view definition and view maintenance machinery on the other (at the warehouse), can lead the warehouse to compute incorrect views.

We illustrate the problems with three examples. For these examples, and for the rest of this paper, we use the relational model for data and relational algebra for views. Although we are using relational algebra, we assume that duplicates are retained in the materialized views. Duplicate retention (or at least a replication count) is essential if deletions are to be handled incrementally [BLT86, GMS93]. Note that the type of solution we propose here can be extended to other data models and view specification languages.

Also, in the examples and in this paper we focus on a single source, and a single view over several base relations. Our methods extend to multiple views directly. Handling a view that spans several sources requires the same type of solution, but introduces additional issues; see Section 7 for a brief discussion.

Example 1: Correct View Maintenance

Suppose our source contains two base relations r_1 and r_2 as follows:

r_1 :

W	X
1	2

r_2 :

X	Y
2	4

Suppose the view at the warehouse is defined by the relational algebra expression:

$$V = \Pi_W(r_1 \bowtie r_2)$$

Initially the materialized view at the warehouse, MV, contains the single tuple [1]. Now suppose tuple [2,3] is inserted into r_2 at the source. For notation, we use $insert(r, t)$ to denote the insertion of tuple t into relation r (similarly for $delete(r,t)$), and we use $([t_1],[t_2],\ldots,[t_n])$ to denote a relation with tuples $t_1,t_2,\ldots,t_n$. The following events occur:

1. Update $U_1 = insert(r_2,[2,3])$ occurs at the source. Since the source does not know the details or contents of the view managed by the warehouse, it simply sends a notification to the warehouse that update U_1 occurred.
2. The warehouse receives U_1. Applying an incremental view maintenance algorithm, it sends query $Q_1 = \Pi_W(r_1 \bowtie [2,3])$ to the source. (That is, the warehouse asks the source which r_1 tuples match with the new [2,3] tuple in r_2.)
3. The source receives Q_1 and evaluates it using the current base relations. It returns the answer relation $A_1 = ([1])$ to the warehouse.
4. The warehouse receives answer A_1 and adds $([1])$ to the materialized view, obtaining ([1],[1]).

In this example the final view at the warehouse is correct, i.e., it is equivalent to what one would obtain using a conventional view maintenance algorithm directly at the source.[1] The next two examples show how the final view can be incorrect.

Example 2: A View Maintenance Anomaly

Assume we have the same relations as in Example 1, but initially r_2 is empty:

r_1:

W	X
1	2

r_2:

X	Y

Consider the same view definition as in Example 1: $V = \Pi_W(r_1 \bowtie r_2)$, and now suppose there are two consecutive updates: $U_1 = insert(r_2,[2,3])$ and $U_2 = insert(r_1,[4,2])$. The following events occur. Note that initially $MV = \emptyset$.

1. The source executes $U_1 = insert(r_2,[2,3])$ and sends U_1 to the warehouse.
2. The warehouse receives U_1 and sends $Q_1 = \Pi_W(r_1 \bowtie [2,3])$ to the source.
3. The source executes $U_2 = insert(r_1,[4,2])$ and sends U_2 to the warehouse.
4. The warehouse receives U_2 and sends $Q_2 = \Pi_W([4,2] \bowtie r_2)$ to the source.
5. The source receives Q_1 and evaluates it on the current base relations: $r_1 = ([1,2],[4,2])$ and $r_2 = ([2,3])$. The resulting answer relation is $A_1 = ([1],[4])$, which is sent to the warehouse.
6. The warehouse receives A_1 and updates the view to $MV \cup A_1 = ([1],[4])$.
7. The source receives Q_2 and evaluates it on the current base relations r_1 and r_2. The resulting answer relation is $A_2 = ([4])$, which is sent to the warehouse.
8. The warehouse receives answer A_2 and updates the view to $MV \cup A_2 = ([1],[4],[4])$.

[1] As stated earlier, for incremental handling of deletions we need to keep both [1] tuples in the view. For instance, if [2,4] is later deleted from r_2, one (but not both) of the [1] tuples should be deleted from the view.

If the view had been maintained using a conventional algorithm directly at the source, then it would be ([1]) after U_1 and ([1],[4]) after U_2. However, the warehouse first changes the view to ([1],[4]) (in step 6) and to ([1],[4],[4]) (in step 8), which is an incorrect final state. The problem is that the query Q_1 issued in step 4 is evaluated using a state of the base relations that differs from the state at the time of the update (U_1) that caused Q_1 to be issued. □

We call the behavior exhibited by this example a *distributed incremental view maintenance anomaly*, or *anomaly* for short. Anomalies occur when the warehouse attempts to update a view while base data at the source is changing. Anomalies arise in a warehousing environment because of the decoupling between the information sources, which are updating the base data, and the warehouse, which is performing the view update.

Example 3: Deletion Anomaly

Our third example shows that deletions can also cause anomalies. Consider source relations:

r_1 :

W	X
1	2

r_2 :

X	Y
2	3

Suppose that the view definition is $V = \Pi_{W,Y}(r_1 \bowtie r_2)$. The following events occur. Note that initially $MV = ([1,3])$.

1. The source executes $U_1 = \mathit{delete}(r_1, [1,2])$ and notifies the warehouse.
2. The warehouse receives U_1 and emits $Q_1 = \Pi_{W,Y}([1,2] \bowtie r_2)$.
3. The source executes $U_2 = \mathit{delete}(r_2, [2,3])$ and notifies the warehouse.
4. The warehouse receives U_2 and emits $Q_2 = \Pi_{W,Y}(r_1 \bowtie [2,3])$.
5. The source receives Q_1. The answer it returns is $A_1 = \emptyset$ since both relations are now empty.
6. The warehouse receives A_1 and replaces the view by $MV - A_1 = ([1,3])$. (Difference is used here since the update to the base relation was a deletion [BLT86].)
7. Similarly, the source evaluates Q_2, returns $A_2 = \emptyset$, and the warehouse replaces the view by $MV - A_2 = ([1,3])$.

This final view is incorrect: since r_1 and r_2 are empty, the view should be empty too. □

1.2 Possible Solutions

There are a number of mechanisms for avoiding anomalies. As argued above, we are interested only in mechanisms where the source, which may be a legacy or unsophisticated system, does not perform any "view management." The source will only notify the warehouse of relevant updates, and answer queries asked by the warehouse. We also are not interested in, for example, solutions where the source must lock data while the warehouse updates its views, or in solutions where the source must maintain timestamps for its data in order to detect "stale" queries from the warehouse. In the following potential solutions, view maintenance is autonomous from source updating:

- *Recompute the view* (RV). The warehouse can either recompute the full view whenever an update occurs at the source, or it can recompute the view periodically. In Example 2, if the warehouse sends a query to the source to recompute the view after it receives U_2, then the source will compute the answer relation $A = ([1],[4])$ (assuming no further updates) and the warehouse will correctly set $MV = ([1],[4])$. Recomputing views is usually time and resource consuming, particularly in a distributed environment where a large amount of data might need to be transferred from the source to the warehouse. In Section 6 we compare the performance of our proposed solution to this one.

- *Store at the warehouse copies of all relations involved in views* (SC). In Example 2, suppose that the warehouse keeps up-to-date copies of r_1 and r_2. When U_1 arrives, Q_1 can be evaluated "locally," and no anomaly arises. The disadvantages of this approach are: (1) the warehouse needs to store copies of all base relations used in its views, and (2) copied relations at the warehouse need to be updated whenever an update occurs at the source.

- *The Eager Compensating Algorithm* (ECA). The solution we advocate avoids the overhead of recomputing the view or of storing copies of base relations. The basic idea is to add to queries sent to the source *compensating queries* to offset the effect of concurrent updates. For instance, in Example 2, consider the receipt of $U_2 = insert(r_1, [4,2])$ in step 4. If we assume that messages are delivered in order, and that the source handles requests atomically, then when the warehouse receives U_2 it can deduce that its previous query Q_1 will be evaluated in an "incorrect" state—Q_1 will see the [4,2] tuple of the second insert. (Otherwise, the warehouse would have received the answer to Q_1 before it received the notification of U_2.) To compensate, the warehouse sends query:

  $$Q_2 = \Pi_{W,Y}([4,2] \bowtie r_2) - \Pi_{W,Y}([4,2] \bowtie [2,3])$$

  The first part of Q_2 is as before; the second part compensates for the extra tuple that Q_1 will see. We call this an "eager" algorithm because the warehouse is compensating (in step 4) even before the answer for Q_1 has arrived (in step 6). In Section 5.2 we briefly discuss a "lazy" version of this approach. Continuing with the example, we see that indeed the answer received in step 6, $A_1 = ([1],[4])$, contains the extra tuple [4]. But, because of the compensation, the A_2 answer received in step 8 is empty, and the final view is correct. In Section 5.2 we present the Eager Compensating Algorithm in detail, showing how the compensating queries are determined for an arbitrary view, and how query answers are integrated into the view.

We also consider two improvements to the basic ECA algorithm:

- *The ECA-Key Algorithm* (ECA^K). If a view includes a key from every base relation involved in the view, then we can streamline the ECA algorithm in two ways: (1) Deletions can be handled at the warehouse without issuing a query to the source. (2) Insertions require queries to the source, but compensating queries are unnecessary. To illustrate point (1), consider Example 3, and suppose W and Y are keys for r_1 and r_2. When the warehouse receives $U_1 = delete(r_1, [1,2])$ (step 1), it can immediately determine that all tuples of the form [1,x] are deleted from the view (where x denotes any value)—no query needs to be sent to the source. Similarly, $U_2 = delete[r_2, [2,3]]$ causes all [x,3] tuples to be deleted from the view, without querying the source. The final empty view is correct. Section 5.4 provides an example illustrating point (2), and a description of ECA^K. Note that ECA^K does have the disadvantage that it can only be used for a subset of all possible views—those that contain keys for all base relations.

- *The ECA-Local Algorithm* (ECA^L). In ECA^K, the warehouse processes deletes locally, without sending any queries to the source, but inserts still require queries to be sent to the source. Generalizing this idea, for a given view definition and a given update, it is possible to determine whether or not the update can be handled locally; see, e.g., [GB94, BLT86]. ECA^L combines local handling of updates with the compensation approach for maintenance of arbitrary views.

1.3 Outline of Paper

In the next section we briefly review related research. Then, in Section 3, we provide a formal definition of correctness for view maintenance in a warehousing environment. As we will see, there are a variety of "levels" of correctness that may be of interest to different applications. In Sections 4 and 5 we present our new algorithms, along with a number of examples. In Section 6 we compare the performance of our ECA algorithm to the view recomputation approach. In Section 7 we conclude and discuss future directions of our work. Additional details—additional examples, proofs, analyses, etc.—that are too lengthy and intricate to be included in the body of the paper are presented in Appendix A to Appendix D.

2 Related Research

Many incremental view maintenance algorithms have been developed. Most of them are designed for a traditional, centralized database environment, where it is assumed that view maintenance is performed by a system that has full control and knowledge of the view definition, the base relations, the updated tuples, and the view [HD92, QW91, SI84]. These algorithms differ somewhat in the view definitions they handle. For example, [BLT86] considers select-project-join (SPJ) views only, while algorithms in [GMS93] handle views defined by any SQL or Datalog expression. Some algorithms depend on key information to deal with duplicate tuples [CW91], while others use a counting approach [GMS93].

A series of papers by Segev et al. studies materialized views in distributed systems [SF90, SF91, SP89a, SP89b]. All algorithms in these papers are based on timestamping the updated tuples, and the algorithms assume there is only one base table. Other incremental algorithms, such as the "snapshot" procedure in [LHM$^+$86], also assume timestamping and a single base table. In contrast, our algorithms have no restrictions on the number of base tables, and they require no additional information. Note that although we describe our algorithms for SPJ views, our approach can be applied to adapt any existing centralized view maintenance algorithm to the warehousing environment.

In both centralized and distributed systems, there are three general approaches to the timing of view maintenance: *immediate* update [BLT86], which updates the view after each base relation is updated, *deferred* update [RK86], which updates the view only when a query is issued against the view, and *periodic* update [LHM$^+$86], which updates the view at periodic intervals. Performance studies on these strategies have determined that the efficiency of an approach depends heavily on the structure of the base relations and on update patterns [Han87]. We assume immediate update in this paper, but we observe that with little or no modification our algorithms can be applied to deferred and periodic update as well.

3 Correctness

Our first task is to define what correctness means in an environment where activity at the source is decoupled from the view at the warehouse. We start by defining the notion of *events*. In our context, an event corresponds to a sequence of operations performed at the same site. There are two types of events at the source:

1. *S_up*: the source executes an update U, then sends an update notification to the warehouse.
2. *S_qu*: the source evaluates a query Q using its current base relations, then sends the answer relation A back to the warehouse.

There are two types of events at the warehouse:

1. *W_up*: the warehouse receives an update U, generates a query Q, and sends Q to the source for evaluation.
2. *W_ans*: the warehouse receives the answer relation A for a query Q and updates the view based on A.

We will assume that events are atomic. That is, we assume there is a local concurrency control mechanism (or only a single user) at each site to ensure that conflicting operations do not overlap. With some extensions to our algorithms, this assumption could be relaxed. We also assume that, within an event, actions always follow the order described above. For example, within an event *S_up*, the source always executes the update operation first, then sends the update notification to the warehouse.

We use e to denote an arbitrary event, se to denote a source event, and we to denote a warehouse event. Event types are subscripted to indicate a specific event, e.g., S_up_i, or W_up_j. For each event, relevant information about the event is denoted using a functional notation: For an event e, $query(e)$, $update(e)$, and $answer(e)$ denote respectively the query, update, and answer associated with event e (when relevant). If event e is caused ("triggered") by another event, then $trigger(e)$ denotes the event that triggered e. For example, $trigger(W\_up_j)$ is an event of type *S_up*. The state of the data after an event e is denoted by $state(e)$.

It is useful to immediately rule out algorithms that are trivially incorrect; for example, where the source does not propagate updates to the warehouse, or refuses to execute queries. These two examples are captured formally by the following rules:

- $\forall U = update(S\_up_i)$: $\exists\ W\_up_j$ such that $update(W\_up_j) = U$.
- $\forall Q = query(W\_up_i)$: $\exists\ S\_qu_j$ such that $query(S\_qu_j) = Q$.

There are a number of other obvious rules such as these that we omit for brevity.

Finally, we define the binary event operator "$<$" to mean "occurs before". We assume that messages are delivered in order and are processed in order. In particular, let e_1, e_2, e_3, e_4 be four events. If $trigger(e_2) = e_1$, $trigger(e_4) = e_3$, and e_1 and e_3 occurred at the same site, then $e_1 < e_3$ iff $e_2 < e_4$. We also use the $<$ relation to order states. That is, we say that $s_i < s_j$ if $state(e_i) = s_i$, $state(e_j) = s_j$, and $e_i < e_j$.

3.1 Levels of Correctness

During the execution of a view maintenance algorithm, the system will process a sequence of updates $U_1, U_2, \ldots, U_n$. In doing so, the source executes events $se_1, se_2, \ldots se_p$ with corresponding resulting states $ss_1, ss_2, \ldots, ss_p$. At the warehouse the triggered events are $we_1, we_2, \ldots we_q$ with corresponding resulting states $ws_1, ws_2, \ldots, ws_q$. When we define the notion of *convergence* (below), we consider executions that are finite, i.e., that have a last update U_n and last events se_p and we_k. However, in general executions may be finite or infinite.

At the warehouse, the state of materialized view V after event we_i is given by $V[ws_i]$, where V is the view definition. Similarly, at the source, $Q[ss_i]$ is the result of evaluating query expression Q on the relations existing after event se_i. If we apply the view definition V to a source state, $V[ss_i]$, we get the state of the view had it been evaluated at the source after event se_i.

An algorithm for warehouse view maintenance may exhibit the following properties:

- *Convergence*: For all finite executions, $V[ws_q] = V[ss_p]$. That is, after the last update and after all activity has ceased, the view is consistent with the relations at the source.
- *Weak Consistency:* For all executions and for all ws_i, there exists an ss_j such that $V[ws_i] = V[ss_j]$. That is, every state of the view corresponds to some valid state of the relations at the source.
- *Consistency:* For all executions and for every pair $ws_i < ws_j$, there exist $ss_k \leq ss_l$ such that $V[ws_i] = V[ss_k]$ and $V[ws_j] = V[ss_l]$. That is, every state of the view corresponds to a valid source state, and in a corresponding order.
- *Strong consistency:* Consistency and convergence.
- *Completeness:* Strong consistency, and for each ss_i there exists a ws_j such that $V[ws_j] = V[ss_i]$. That is, there is a complete order-preserving mapping between the states of the view and the states of the source.

Although completeness is a nice property since it states that the view "tracks" the base data exactly, we believe it is too strong a requirement and unnecessary in most practical warehousing scenarios. In some cases, convergence may be sufficient, i.e., knowing that "eventually" the warehouse will have a valid state, even if it passes through intermediate states that are invalid. In other cases, consistency (weak or not) may be required, i.e., knowing that every warehouse state is valid with respect to some source state. Examples 2 and 3 showed that a straightforward incremental view maintenance algorithm is not even weakly consistent in the environments we consider. We will show that our Eager Compensating Algorithm is strongly consistent, and hence a satisfactory approach for most warehousing scenarios.

4 Views and Queries

Before presenting our algorithms, we must define the warehouse views we handle and the types of queries generated. In this paper, we consider views defined as:

$$V = \Pi_{proj}(\sigma_{cond}(r_1 \times r_2 \times \ldots \times r_n))$$

where *proj* is a set of attribute names, *cond* is a boolean expression, and $r_1, \ldots, r_n$ are base relations. Note that any relational algebra expression constructed with select, project, and join operations can be transformed into an equivalent expression of this form. For simplicity, we assume that $r_1, \ldots, r_n$ are distinct relations. Our algorithms can be extended to allow multiple occurrences of the same relation (e.g., by handling updates to such relations once for each appearance of the relation).

4.1 Signs

Our warehouse algorithms will handle two types of updates: insertions and deletions. (Modifications must be treated as deletions followed by insertions, although extensions to our approach could permit modifications to be treated directly.) For convenience, we adopt an approach similar to [BLT86] and use *signs* on tuples: $+$ to denote an inserted or an existing tuple, and $-$ to denote a deleted tuple. Tuple signs are propagated through relational operations, as we will illustrate.

Consider an update $U_1 = delete(r_1, [1,2])$, which causes the warehouse to issue a query $Q_1 = \Pi_W([1,2] \bowtie r_2)$. Using signs, we instead issue the query $Q_1 = \Pi_W(-[1,2] \bowtie r_2)$, where "$-$" attached to tuple [1,2] represents that this is a deleted tuple. Suppose that at the source there is an r_2 tuple [2,3] (which by default has a $+$ sign). The result of Q_1, which we call A_1, will contain the tuple $-[1]$, i.e., the minus sign carries through. Relation A_1 is then returned to the warehouse, where it is combined with the existing view by an operation (explained below): $MV \leftarrow MV + A_1$. Because of the minus sign, the [1] tuple in A_1 is removed from the view. Note that if the original update U_1 had been an insert, the tuple [1,2] would have a plus sign, and tuple [1] would instead have been added to the view. Using tuple signs allows us to handle inserts and deletes uniformly and compactly in our algorithms.

More formally, existing tuples and inserted tuples always have a plus sign, while deleted tuples always have a minus sign. When a relational algebra expression operates on signed tuples, the sign of the result tuples is given by the following tables, where t, t_1 and t_2 are signed tuples:

t	$\sigma_{cond}(t)$	$\Pi_{proj}(t)$
$+$	$+$	$+$
$-$	$-$	$-$

t_1	t_2	$t_1 \times t_2$
$+$	$+$	$+$
$+$	$-$	$-$
$-$	$-$	$+$
$-$	$+$	$-$

In addition, we define two binary operators, also called $+$ and $-$, which operate on relations with signed tuples. For a relation r, let $pos(r)$ denote the tuples with a plus sign and let $neg(r)$ denote the tuples with a minus sign. Then:

$$r_1 + r_2 = (pos(r_1) \cup pos(r_2)) - (neg(r_1) \cup neg(r_2))$$
$$r_1 - r_2 = r_1 + (-r_2)$$

Operators $+$ and $-$ are commutative and associative, and they generalize to relational expressions and to single tuples in the obvious way. The cross product $\times$ is distributive over $+$ and $-$.

Note that the use of signed tuples is a notational convenience only—it is not necessary for sources to handle signed tuples in order to participate in our algorithms.

4.2 Query Expressions

In maintaining a view over relations $r_1, \ldots, r_n$, our algorithms will generate queries that contain a collection of terms. Each term is of the form:

$$T = \Pi_{proj}(\sigma_{cond}(\tilde{r}_1 \times \tilde{r}_2 \times \ldots \times \tilde{r}_n)) \quad (4.1)$$

where each $\tilde{r}_i$ is either a relation r_i or an updated tuple t_i of r_i. A query is formed by a sum of terms:

$$Q = \sum_i T_i. \tag{4.2}$$

As an example, the following relational algebra expression is query we might generate:

$$Q = \Pi_{proj}(\sigma_{cond}(r_1 \times [2,3] \times r_3)) + \Pi_{proj}(\sigma_{cond}(-[1,2] \times [2,3] \times r_3))$$

In our algorithms we often derive queries from earlier queries or from view definitions. For example, say we are given a view definition $V = \Pi_{proj}(\sigma_{cond}(r_1 \times r_2))$, and we receive a deletion $U = delete(r_2, [3,4])$. Then the query we want to send to the source is V with r_2 substituted by $-[3,4]$, i.e., $Q = \Pi_{proj}(\sigma_{cond}(r_1 \times -[3,4]))$. We use $V\langle U \rangle$ to denote view expression V with the updated tuple of U substituted for U's relation.

More formally, consider any query (or view definition) of the form $Q = \sum_i T_i$ (recall Equation 4.2). Let U be an update involving relation r_k, and let $tuple(U)$ be the updated tuple. Then $Q\langle U \rangle = \sum_i T_i\langle U \rangle$, where for each T_i (recall Equation 4.1):

$$T_i\langle U \rangle = \begin{cases} \emptyset & \textit{if } \tilde{r}_k \textit{is an updated tuple} \\ \Pi_{proj}(\sigma_{cond}(\tilde{r}_1 \times \ldots \times \tilde{r}_{k-1} \times tuple(U) \times \tilde{r}_{k+1} \times \ldots \times \tilde{r}_n)) & \textit{if } \tilde{r}_k \textit{is relation } r_k \end{cases}$$

We also recursively define $Q\langle U_1, U_2, \ldots U_k \rangle = (Q\langle U_1, U_2, \ldots U_{k-1} \rangle)\langle U_k \rangle$. That is, $Q\langle U_1, U_2 \ldots, U_k \rangle$ is the query in which all updated relations have been replaced by the corresponding updated tuples. Note that, by definition, if any two or more of the updates $U_1, U_2, \ldots U_k$ occur on the same relation, then $Q\langle U_1, U_2, \ldots U_k \rangle = \emptyset$.

5 The ECA Algorithm

In this section we present the details of our *Eager Compensating Algorithm* (ECA), introduced in Section 1 as a solution to the anomaly problem. ECA is an incremental view maintenance algorithm based on the centralized view maintenance algorithm described in [BLT86]. ECA anticipates the anomalies that arise due to the decoupling between base relation updates and view modification, and ECA compensates for the anomalies as needed to ensure correct view maintenance. Before we present ECA and its extensions, we first review the original incremental view maintenance algorithm.

5.1 The Basic Algorithm

The view maintenance algorithm described in [BLT86] applies incremental changes to a view each time changes are made to relevant base relations. We adapt this algorithm to the warehousing environment and use our event-based notation:

Algorithm 5.1 (Incremental view maintenance algorithm)

At the source:

- S_up_i: execute U_i;
  send U_i to the warehouse;
  trigger event W_up_i at the warehouse

- S_qu_i: receive query Q_i;
  let $A_i = Q_i[ss_i]$;   (ss_i is the current source state)
  send A_i to the warehouse;
  trigger event W_ans_i at the warehouse

At the warehouse:

- W_up_i: receive update U_i;
  let $Q_i = V\langle U_i \rangle$;
  send Q_i to the source;
  trigger event S_qu_i at the source

- W_ans_i: receive A_i;
  update view: $MV = MV + A_i$

(end algorithm)

Notice that each update at the source triggers an S_up event, which then triggers a W_up event at the warehouse, triggering an S_qu event back at the source, and finally a W_ans event at the warehouse (recall Figure 1.1). As shown in Examples 2 and 3, this algorithm may lead to anomalies. Consequently, using our definitions from Section 3, this basic algorithm is neither convergent nor weakly consistent in a warehousing environment.

5.2 The Eager Compensating Algorithm

We start by defining the set of "pending" queries at the warehouse:

Definition: Consider the processing of an event we at the warehouse. Let the *unanswered query set* for we, $UQS(we)$, be the set of queries that were sent by the warehouse before we occurred, but whose answers were not yet received by we. We shorten $UQS(we)$ to UQS when we refers to the event being processed. □

When the warehouse receives an update U_i and UQS is not empty, then U_i may cause queries Q_j in UQS to be evaluated incorrectly. The incorrectness arises because the queries Q_j are assumed to be computed before U_i, but are actually computed after U_i. Thus, all queries in UQS will "see" a source state that already reflects update U_i. ECA takes this behavior into account: When ECA issues its query in response to update U_i, ECA incorporates one "compensating query" for each query in UQS. The compensating queries offset the effects of U_i on the results of queries in UQS.

An important and subtle consequence of using compensating queries is that the results of queries should be applied to the view only after the answer to the last compensating query has been received. If instead we updated the view after the receipt of each answer, then the view might temporarily assume an invalid state. (That is, in the terminology of Section 3.1, the algorithm would be convergent but not consistent.) To avoid invalid view states, ECA collects all intermediate answers in a temporary relation called `COLLECT`, and only updates the view when the last answer has been received (i.e. when $UQS = \emptyset$).

Algorithm 5.2 (ECA)

`COLLECT` $= \emptyset$.
The source events behave exactly as in Algorithm 5.1. At the warehouse:

- W_up_i: receive U_i;
  let $Q_i = V\langle U_i\rangle - \sum_{Q_j \in UQS} Q_j\langle U_i\rangle$
  send Q_i to the source;
  trigger event S_qu_i at the source
- W_ans_i: receive A_i;
  let `COLLECT` = `COLLECT` + A_i;
  if $UQS = \emptyset$
  then { $MV \leftarrow MV +$ `COLLECT`; `COLLECT` $\leftarrow \emptyset$ }
  else do nothing

(end algorithm)

5.3 Example

The following example illustrates ECA handling three insertions to three different base relations. A number of additional ECA examples are given in Appendix A.

Example 4: ECA

Consider source relations:

r_1:

W	X
1	2

r_2:

X	Y

r_3:

X	Y

Let the view definition be $V = \Pi_W(r_1 \bowtie r_2 \bowtie r_3)$, and suppose the following events occur. Initially, the materialized view at the warehouse is empty, and `COLLECT` is initialized to empty. For brevity, we omit the source events, only listing events occurring at the warehouse. We assume that the three updates occur at the source before any queries are answered.

1. Warehouse receives $U_1 = insert(r_1, [4,2])$

   Warehouse sends $Q_1 = V\langle U_1\rangle = \Pi_W([4,2] \bowtie r_2 \bowtie r_3)$

2. Warehouse receives $U_2 = insert(r_3, [5,3])$

   $UQS = \{Q_1\}$

   Warehouse sends $Q_2 = V\langle U_2\rangle - Q_1\langle U_2\rangle$
   $= \Pi_W(r_1 \bowtie r_2 \bowtie [5,3]) - \Pi_W([4,2] \bowtie r_2 \bowtie [5,3])$

3. Warehouse receives $U_3 = insert(r_2, [2,5])$

   $UQS = \{Q_1, Q_2\}$

   Warehouse sends $Q_3 = V\langle U_3\rangle - Q_1\langle U_3\rangle - Q_2\langle U_3\rangle$
   $= \Pi_W(r_1 \bowtie [2,5] \bowtie r_3) - \Pi_W([4,2] \bowtie [2,5] \bowtie r_3)$
   $- \Pi_W((r_1 - [4,2]) \bowtie [2,5] \bowtie [5,3])$

4. Warehouse receives $A_1 = [4]$

   `COLLECT` $= ([4])$, $UQS = \{Q_2, Q_3\}$

5. Warehouse receives $A_2 = [1]$
   COLLECT = ([1],[4]), $UQS = \{Q_3\}$
6. Warehouse receives $A_3 = \emptyset$
   COLLECT = ([1],[4]), $UQS = \emptyset$
   Warehouse updates view: $MV = \emptyset +$ COLLECT = ([1], [4]), which is correct. □

A complete proof showing that the algorithm is strongly consistent is given in Appendix B. Notice, however, that ECA is not complete. Recalling Section 3.1, completeness requires that every source state be reflected in some view state. Clearly some source states may be "missed" by the ECA algorithm while it collects query answers. We can modify ECA to obtain a complete algorithm that we call the *Lazy Compensating Algorithm* (LCA). For each source update, LCA waits until it has received all query answers (including compensation) for the update, then applies the changes for that update to the view. A detailed description of LCA is beyond the scope of this paper. LCA is less efficient than ECA, and we believe that strong consistency is sufficient for most environments and completeness is generally not needed. Hence, we expect that ECA will be more useful than LCA in practice.

5.4 The ECA-Key Algorithm

We can "streamline" ECA in the case where key attributes for each base relation are present in the view. (In fact, it could be advisable to design warehouse view definitions with this property in mind, for more efficient maintenance.) The *ECA-Key Algorithm* (ECA^K) proceeds as follows:

1. COLLECT is initialized with the current materialized view (not the empty set). Instead of storing modifications to MV, COLLECT is a "working copy" of MV.
2. When a delete is received at the warehouse, no query is sent to the source. Instead, the delete is applied to COLLECT immediately, using a special *key-delete* operation defined below.
3. When an insert is received at the warehouse, a query is sent to the source. However, no compensating queries are added. Thus, when an insert U is received, the query sent to the source is simply $V\langle U\rangle$.
4. As answers are received, they are accumulated in the COLLECT set, as in the original ECA. However, duplicate tuples are not added to the COLLECT set. When the view contains keys for all base relations, there can be no duplicates in the view. Thus, if a duplicate occurs, it is due to an anomaly and can be ignored.
5. If UQS is empty, the materialized view is updated, replacing it with the tuples in COLLECT.

Example 5: $\mathbf{ECA}^K$

Consider the following source relations, where W and Y are key attributes.

r_1:

W	X
1	2

r_2:

X	Y
2	3

Suppose that the warehouse view definition is $V = \Pi_{W,Y}(r_1 \bowtie r_2)$, and the following events occur. Initially, the materialized view at the warehouse is $MV = ([1,3])$, and COLLECT $= MV$.

1. $U_1 = insert(r_2, [2,4])$

   Warehouse sends $Q_1 = V\langle U_1 \rangle = \Pi_{W,Y}(r_1 \bowtie [2,4])$

2. $U_2 = insert(r_1, [3,2])$

   Warehouse sends $Q_2 = V\langle U_2 \rangle = \Pi_{W,Y}([3,2] \bowtie r_2)$

3. $U_3 = delete(r_1, [1,2])$

   Operation *key-delete*(V, r_1, [1,2]) (see below) deletes tuples of the form [1,x], obtaining `COLLECT` = `COLLECT` − ([1,3]) = $\emptyset$, $UQS = \{Q_1, Q_2\}$

4. Warehouse receives $A_1 = ([3,4])$

   `COLLECT` = `COLLECT` + ([3,4]) = ([3,4]), $UQS = \{Q_2\}$

5. Warehouse receives $A_2 = ([3,3],[3,4])$

   First, A_2 is added to `COLLECT`: duplicate tuple [3,4] is not added, so `COLLECT` = ([3,3],[3,4]). Next, since $UQS = \emptyset$, MV is set to `COLLECT`, so MV = ([3,3],[3,4]). Note that `COLLECT` is not reset to empty.

The special operation *key-delete*(MV, r_1, [1,2]) deletes from MV all tuples whose attribute corresponding to r_1's key (i.e., attribute W) is equal to the key value in tuple [1,2] (i.e., 1).

Observe that if we had installed `COLLECT` into MV in steps 3 or 4, without waiting for $UQS = \emptyset$, then the view would have temporarily assumed an invalid state (resulting in convergence but not consistency).

It is the presence of keys in the view definition that makes it possible to perform deletes at the warehouse without issuing queries to the source, and that eliminates the need for compensating queries in the case of inserts. Consider deletions first. Since each view tuple contains key values for all base relations, when a base relation tuple t is deleted, we can use the key values in t to identify which tuples in the view were derived using t. These are the view tuples that must be deleted.

Now consider insertions. Since insertions cause queries to be sent to the source, anomalies can still arise. However, all such anomalies result in either duplicate view tuples (which we can detect and ignore), or missing tuples that would have been deleted anyway. As illustration, suppose Q_1 of Example 5 had been executed when U_1 occurred; it would have evaluated to ([1,4]). Instead, because of the delay, we have A_1 = ([3,4]). A_1 is incorrect in two ways: (1) It contains an extra tuple [3,4] produced because Q_1 was evaluated after insert U_2. (2) It is missing tuple [1,4] because Q_1 was evaluated after delete U_3. However, both of these problems are resolved at the warehouse: (1) The extra tuple [3,4] is identified as a duplicate when A_2 is received. (2) The missing tuple [1,4] would have been deleted by U_3. Details and a sketch for the ECA^K correctness proof appear in Appendix C.

5.5 The ECA-Local Algorithm

The original ECA uses compensating queries to avoid the anomalies that may occur when queries are sent to the source. ECA^K relies on key properties to avoid compensating queries; furthermore, ECA^K introduces the concept of *local updates* (deletions, in the case of ECA^K), which can be handled at the warehouse without sending queries to the source. The last algorithm we discuss, the *ECA-Local Algorithm* (ECA^L) combines the compensating queries of ECA with the local updates of ECA^K to produce a streamlined algorithm that applies to general views.

In ECA^L, each update is handled either locally or non-locally. A number of papers, e.g., [BLT86, GB94, TB88], describe conditions when, for a particular view definition and a particular base relation update, the view can be updated without further access to base relations (i.e., the view is *autonomously computable*, using the terminology of [BLT86]). These results can be used to identify which updates ECA^L will process locally. For updates that cannot be processed locally, ECA is used (assuming the key condition for ECA^K does not hold), with compensating queries whenever necessary.

Unfortunately, maintaining the correct order of execution among local and non-local processing in ECA^L is not straightforward. Intuitively, we can process a local update as soon as all queries for previous updates have been answered and applied to the view. However, consider three updates, U_1, U_2, and U_3, where U_2 is the only local update. Suppose we process U_2 as soon as the query Q_1 for U_1 is answered. Since ECA^L uses compensating queries, the "true" view update corresponding to U_1 may include not only the answer for Q_1, but also compensations appearing in queries for later updates (such as U_3). To correctly handle this scenario, ECA^L must buffer updates and, in some cases, "split" query results (separating compensation from original), in order to process local updates on a correct version of the view. The details of ECA^L, and determining whether the additional overhead is worthwhile, is left as future work.

5.6 Properties of the ECA family

ECA, and its extensions ECA^K and ECA^L, have the following desirable properties:

1. They are incremental, meaning that they update the warehouse based on updates to the source, rather than recomputing the complete warehouse view from scratch.

2. They do not place any additional burden on the sources (e.g., timestamps, locks, etc.).

3. When the update frequency is low, i.e. when the answer to a warehouse query comes back before the next update occurs at the source, then the ECA algorithms behave exactly like the original incremental view maintenance algorithm. (Compensating queries are used only when the answer to a query has not been received before the next update occurs at the source.)

6 Performance Evaluation

In Section 1.2 we outlined several strategies for view maintenance in a warehousing environment: recomputing the view (RV), storing copies of all base relations (SC), our Eager Compensating Algorithm (ECA), and our extensions ECA-Key (ECA^K) and ECA-Local (ECA^L). All of these approaches provide strong consistency (recall Section 3.1). Thus, a natural issue to explore is the relative performance of these strategies. In this paper we evaluate only the basic algorithms, RV and ECA. From the performance perspective, ECA^K is simply an enhancement to ECA that eliminates querying the source in certain cases. Since there is very little additional overhead in ECA^K, ECA^K should certainly be used when it is possible to arrange for warehouse views to include all base relation keys. Storing copies of base relations (SC) can be seen as an enhancement to any of our algorithms, requiring an "orthogonal" performance comparison (based on warehouse storage costs, etc.) that is beyond the scope of this paper. As discussed in Section 5.5, ECA^L requires complex processing at the warehouse, the measurement of which falls outside the scope of our performance evaluation. We plan to address performance issues for SC and ECA^L, along with a more extensive evaluation of ECA, as future work.

When we intuitively compare RV and ECA, it seems that ECA should certainly outperform RV, since ECA is an incremental update algorithm while RV recomputes the view from scratch. However, ECA may need to send many more queries to the source than RV. In addition, ECA's queries grow in complexity as compensations upon compensations are added. Hence, we seek to determine when it is more effective to recompute the entire view, rather than maintaining it incrementally with the associated extra activity.

To answer this question, we provide an initial performance analysis of RV and ECA. We note that this is not a comprehensive analysis, as there are a number of parameters we have not fully studied (ranging from the number of relations, to the exact sequence and timing of the updates, to the way queries are optimized at the source). Rather, we have selected a "representative" scenario that serves to illustrate the performance tradeoffs.

For the analysis, we focus on three separate cost factors: M, the number of messages sent between the source and warehouse, B, the number of bytes transferred from the source to the warehouse, and IO, the number of I/O's performed at the source. In RV and ECA, identical update notification messages are sent to the warehouse, so these costs are not included in our calculations. Throughout this section, we use the variables listed in Table 1, shown with their default values. (The default values are starting points for our evaluation; we vary values as part of our analysis.) The RV algorithm is described informally in Section 1.2; a formal specification is provided in Appendix D.

Name	Meaning	Default Value
M	Number of messages sent	N/A
B	Total number of bytes transferred	N/A
IO	Number of I/O's	N/A
C	Cardinality of a relation	100
S	Size of projected attributes	4 bytes
σ	Selection factor	1/2
J	Join factor	4
K	# of tuples per physical block	20
k	# of updates at the source	N/A
$s-1$	# of updates skipped before recomputing view	$s \leq k$

Table 1: List of variables.

6.1 Performance Based on Number of Messages

Assume there is a sequence of k updates. For RV, assume the warehouse sends a query message to the source asking it to recompute the view after every s updates, $s \leq k$. Counting both the query and answer messages, the total number of messages is $M_{RV} = \lceil \frac{k}{s} \rceil \times 2$. Thus, RV generates at least 2 messages (if the view is only recomputed once; $s = k$) and at most $2k$ messages (if the view is recomputed after every update; $s = 1$). For ECA, if there are k updates, we always have k queries and k answers, so there are $2k$ messages.[2]

Thus, in the situation least favorable for ECA, ECA sends $2k$ messages while RV sends 2 messages. Of course, the price of the RV approach in this case is that the state of the warehouse view lags well behind the state of the base relations. In the most favorable situation for ECA, ECA and RV both send $2k$ messages.

6.2 Performance Based on Data Transferred

To analyze the number of bytes transferred (and later on the number of I/O's), we introduce a sample scenario consisting of a particular view and a particular sequence of update operations. As mentioned earlier, we have chosen to focus on a sample scenario to illustrate the performance tradeoffs while keeping the number of parameters manageable.

Example 6: Example warehouse scenario

Base relation schema: $r_1(W, X)$, $r_2(X, Y)$, $r_3(Y, Z)$
View definition: $V = \Pi_{W,Z}(\sigma_{cond}(r_1 \bowtie r_2 \bowtie r_3))$
Updates: U_1 = insert[r_1, t_1], U_2 = insert[r_2, t_2], U_3 = insert[r_3, t_3] $\square$

The condition *cond* involves a comparison between attributes W and Z (e.g., $W > Z$). (This condition has an impact on the derivation of the the number of I/O's performed.) Later we extend this example to a sequence of k updates.

We make the following assumptions in our analysis:

1. The cardinality (number of tuples) of each relation is some constant C.
2. The size of the combined W, Z attributes is S bytes.
3. The join factor $J(r_i, a)$ is the expected number of tuples in r_i that have a particular value for attribute a. We assume that the join factor is a constant J in all cases. So, for example, if we join a 20-tuple relation with a second relation, we expect to get $20J$ tuples.
4. The selectivity for the condition *cond* is given by σ, $0 \leq \sigma \leq 1$.
5. We assume that C, J and our other parameters do not change as updates occur. This closely approximates their behavior in practice when the updates are single-tuple inserts and deletes (so the size and selectivity do not change significantly), or when C and J are so large that the effect of updates is insignificant.

Not surprisingly, for RV the fewest bytes are transferred (B_{RVBest}) when the view is recomputed only once, after U_3 has occurred. The worst case ($B_{RVWorst}$) is when the view is recomputed after each update. For ECA, the best case ($B_{ECABest}$) is when no compensating queries are needed, i.e., the updates are sufficiently spaced so that each query is processed before the next

[2]Because ECA uses signed queries (recall Section 4.1), and some sources—such as an SQL server—may need to handle the positive and negative parts of such queries separately, we may need to send a pair of queries for some updates instead of a single query. We assume the pair of queries is "packaged" as one message, and the pair of answers also is returned in a single message.

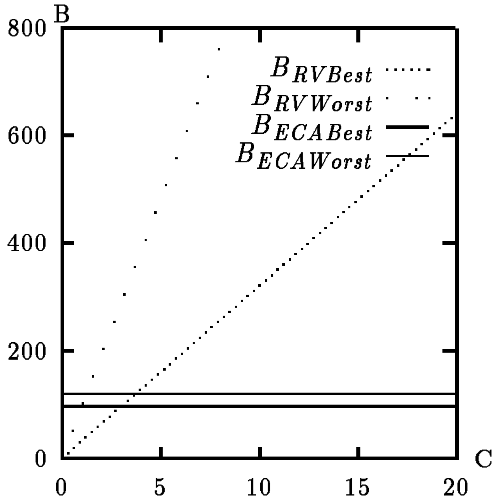

Figure 6.2: B versus C

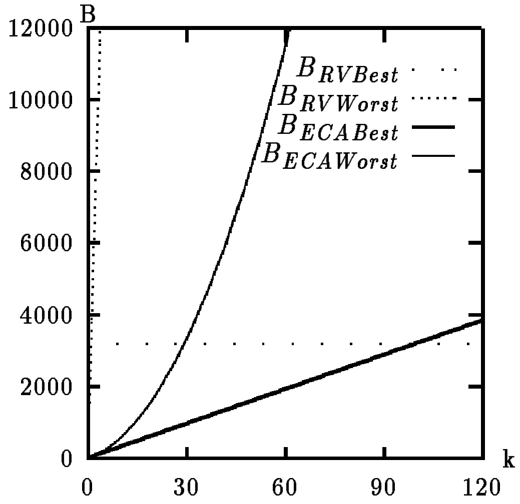

Figure 6.3: B versus k

update occurs at the source. Note that in this case, ECA performs as efficiently as the original incremental algorithm (Algorithm 5.1). The worst case for ECA ($B_{ECAWorst}$) is when all updates occur before the first query arrives at the source. Intuitively, the difference between the best and worst cases of ECA represents the "compensation cost."

The calculations for analyzing our algorithms are rather complex and therefore omitted here; the complete derivations can be found in Appendix D. In particular, the expressions derived in Appendix D for the number of bytes transferred are:

$$\begin{aligned} B_{RVBest} &= S\sigma C J^2 \\ B_{RVWorst} &= 3S\sigma C J^2 \\ B_{ECABest} &= 3S\sigma J^2 \\ B_{ECAWorst} &= 3S\sigma J(J+1) \end{aligned}$$

Figure 6.2 shows the number B of bytes transferred as a function of the cardinality C. (In all of our graphs, parameters have the default values of Table 1 unless otherwise indicated.) Best and worst cases are shown for both algorithms. Thus, for each algorithm, actual performance will be somewhere in between the best and worst case curves, depending on the timing of update arrivals (for ECA) and the frequency of view recomputation (for RV).

From Figure 6.2 we see that in spite of the compensating queries, ECA is much more efficient than RV (in terms of data transferred), unless the relations involved are extremely small (less than approximately 5 tuples each). This result continues to hold over wide ranges of the join selectivity J, except if J is very small (see equations above).

One of the reasons ECA appears to perform so well is that we are considering only three updates, so the amount of "compensation work" is limited. In Appendix D we extend our analysis to an arbitrary number k of updates and obtain the following equations:

$$\begin{aligned} B_{RVBest} &= S\sigma C J^2 \\ B_{RVWorst} &= kS\sigma C J^2 \\ B_{ECABest} &= kS\sigma J^2 \\ B_{ECAWorst} &= kS\sigma J^2 + k(k-1)S\sigma J/3 \end{aligned}$$

Figure 6.3 shows the number of bytes transferred as a function of k when $C = 100$. As expected, there is a crossover point beyond which recomputing the view once (RV's best case) is superior to even the best case for ECA. For our example, this crossover is at 100 updates. In the ECA worst case, when all updates occur at the source before any of the warehouse queries arrive, each warehouse query must compensate every preceding update. This behavior results in ECA transmitting additional data that is quadratic on the number of updates. Hence, in the situation least favorable for ECA, RV outperforms ECA when 30 or more updates are involved. Bear in mind that this situation occurs only if all updates precede all queries. If updates and queries are interleaved at the source, then performance will be somewhere between the ECA best and worst cases, and the crossover point will be somewhere between 30 and 100 updates.

Also notice that Figure 6.3 is for relatively small relations ($C = 100$); for larger cardinalities the crossover points will be at larger number of updates. Finally, note that the RV best case we have been comparing against assumes the view is recomputed once, no matter how many updates occur. If RV recomputes the view more frequently (such as once per some number of updates), then its cost will be substantially higher. In particular, $B_{RVWorst}$ is very expensive and always substantially worst than $B_{ECAWorst}$.

6.3 Performance Based on I/O

We continue using Example 6 and now estimate the number of I/O's performed at the source. We consider two extreme scenarios. In the first scenario, indexing is used and ample memory is available, while in the second scenario memory is very limited and there are no indexes. Studying these extremes lets us discover the conditions that are most favorable for the algorithms we consider.

For our analysis we do not consider optimization across query terms or the effects of caching: If a query consists of several terms (recall Definition 4.2), each one is evaluated independently. Furthermore, if in evaluating a term we need a particular page that was accessed in evaluating the previous term, we still assume that the page has to be read from disk (as opposed to a cache). Since only ECA uses multi-term queries, our results for ECA will be pessimistic. (Our no-caching assumption could potentially hurt RV as well, but only in the case where recomputations are frequent enough that pages from previous evaluations are still in memory.)

For the first scenario (call it Scenario 1) we assume that main memory is large enough to hold the parts of all three relations that participate in the join. We also assume that there are clustering indexes on X for r_1 and r_2, a clustering index on Y for r_3, and a non-clustering index on Y for r_2. These indexes reside in memory and their access generates no I/O.

In Appendix D we derive the following equations for the number of I/O's in Scenario 1:

$$
\begin{aligned}
IO_{RVBest} &= 3I \\
IO_{RVWorst} &= 9I \\
IO_{ECABest} &= 3min(I, J) + 3 \\
IO_{ECAWorst} &= 3min(I, J) + 6
\end{aligned}
$$

where $I = \lceil C/K \rceil$ is the number of I/O's needed to read an entire base relation. From these equations we can see that in the situation least favorable for ECA, ECA is at most a constant factor of 6 I/O's more expensive than RV. However, if $J < I$, then ECA can outperform RV arbitrarily (bounded by $I - J$). For large relations, we would expect J to be substantially smaller than I, in which case even ECA's worst case will have substantially fewer I/O's than RV.

For the second I/O scenario (Scenario 2) we assume that there are no indexes, and that there are only three free memory blocks to be used for nested-loop join processing. The relevant equations, derived in Appendix D, are:

$$\begin{aligned}
IO_{RVBest} &= I^3 \\
IO_{RVWorst} &= 3I^3 \\
IO_{ECABest} &= 3II' \\
IO_{ECAWorst} &= 3I(I'+1)
\end{aligned}$$

where $I = \lceil C/K \rceil$ and $I' = \lceil C/(2K) \rceil$. It is apparent from these equations that, as expected, the I/O costs for this scenario are much higher than for Scenario 1. However, the difference between ECA and RV is even more dramatic now. Unless the relations are very small (occupying less than 3 blocks), even in the situation least favorable for ECA, ECA outperforms RV by a factor of I.

As with our evaluation of the amount of data transferred, the performance of ECA benefits from the fact that we have been considering only three updates. Hence, we extend both scenarios to the case of k updates. The equations, again derived in Appendix D, follow. For brevity, we consider here only the (likely) case in which $J < I$:

Scenario 1:

$$\begin{aligned}
IO_{RVBest} &= 3I \\
IO_{RVWorst} &= 3kI \\
IO_{ECABest} &= k(J+1) \\
IO_{ECAWorst} &= k(J+1) + k(k-1)/3
\end{aligned}$$

Scenario 2:

$$\begin{aligned}
IO_{RVBest} &= I^3 \\
IO_{RVWorst} &= kI^3 \\
IO_{ECABest} &= kII' \\
IO_{ECAWorst} &= kII' + Ik(k-1)/3
\end{aligned}$$

Figures 6.4 and 6.5 show the results when $J = 4$ and $C = 100$ (and hence $I = 5$). The shape of these curves is similar to those in Figure 6.3, and thus our conclusions for I/O costs are similar to those for data transmission. The main difference is that the crossover points occur with smaller update sequences: $k = 3$ for Scenario 1, and $5 < k < 8$ for Scenario 2, as opposed to a crossover between $k = 30$ and $k = 100$ when data transfer is the metric. Intuitively, this means that ECA is not as effective at reducing I/O costs as it is at reducing data transfer. However, ECA can still reduce I/O costs over RV significantly, especially if relations are larger than the 100 tuples considered for these figures. Also, we expect that the I/O performance of ECA would improve if we incorporated multiple term optimization or caching into the analysis.

As a final note, we remind the reader that our results are for a particular three-relation view. In spite of this, we believe that our results are indicative of the performance issues faced in choosing between RV and ECA. Our results indicate that when the view involves more relations, ECA should still generally outperform RV.

7 Conclusion

Data warehousing is an emerging (and already very popular) technique used in many applications for retrieval and integration of data from autonomous information sources. However, warehousing typically is implemented in an ad-hoc way. We have shown that standard algorithms for maintaining a materialized view at a warehouse can lead to anomalies and inconsistent modifications to

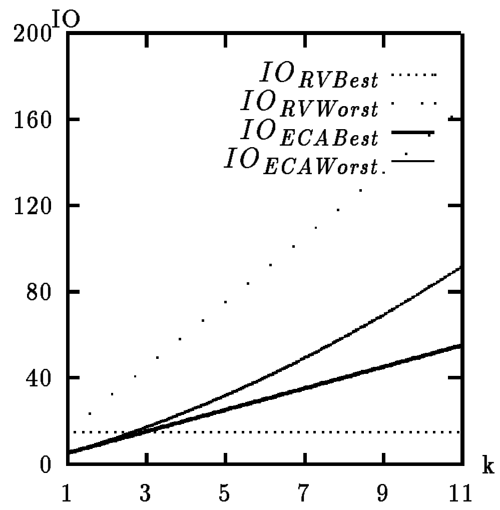

Figure 6.4: IO versus k, Scenario 1

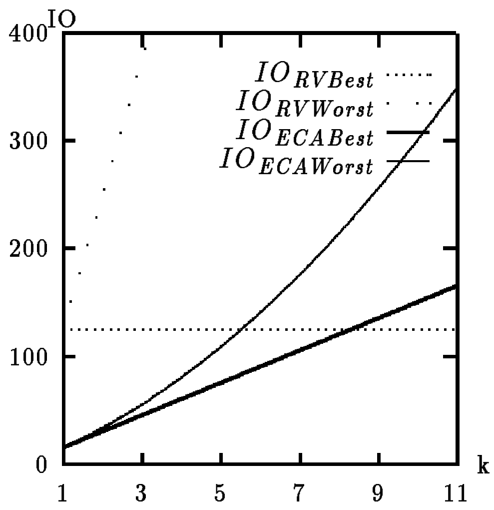

Figure 6.5: IO versus k, Scenario 2

the view. The anomalies are due to the fact that view maintenance at the warehouse is decoupled from updates at the data sources, and we cannot expect the data sources to perform sophisticated functions in support of view management. Consequently, previously proposed view maintenance algorithms cannot be used in this environment.

We have presented a new algorithm, and outlined two extensions, for correctly maintaining materialized views in a warehousing environment. Our Eager Compensating Algorithm, ECA, and its streamlined versions, ECA^K and ECA^L, are all strongly consistent, meaning that the warehouse data always corresponds to a meaningful state of the source data. An initial performance study analyzing three different cost factors (messages, data traffic, and I/O) suggests that, except for very small relations, ECA is consistently more efficient than periodically recomputing the warehouse view from scratch.

Although in this paper we have addressed a restricted warehousing environment with only one source and one view, ECA can readily be adapted to more general scenarios. For example, in a warehouse consisting of multiple views where each view is over data from a single source, ECA is simply applied to each view separately.

In the future we plan to address the following additional issues.

- We will consider how ECA (and its extensions) can be adapted to views over multiple sources. Many aspects of the anomaly problem remain the same. However, additional issues are raised because warehouse queries (both regular queries and compensating queries) must be fragmented for execution at multiple sources. While fragmenting itself does not pose a novel problem (at least in the straightforward relational case), coordinating the query results and the necessary compensations for anomaly-causing updates may require some intricate algorithms. We expect that the protocols reported in [GW94] for evaluating constraints in loosely-coupled databases may prove to be useful for extending ECA to the multi-site case.

- We will consider how ECA can be extended to handle a set of updates at once, rather than one update at a time. Since we expect that in practice many source updates will be "batched," this extension should result in a very useful performance enhancement.

- We will modify the algorithms to handle views defined by more complex relational algebra expressions (e.g., using union and/or difference) as well as other relational query languages (e.g., SQL or Datalog).

- We will explore how the algorithms can be adapted to other data models (e.g., an object-based data model).

References

[BLT86] J.A. Blakeley, P.-A. Larson, and F.W. Tompa. Efficiently updating materialized views. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pages 61–71, Washington, D.C., June 1986.

[CW91] S. Ceri and J. Widom. Deriving production rules for incremental view maintenance. In *Proceedings of the Seventeenth International Conference on Very Large Data Bases*, pages 577–589, Barcelona, Spain, September 1991.

[GB94] Ashish Gupta and J. A. Blakeley. Updating materialized views using the view contents and the update. In *unpublished document*, 1994.

[GMS93] A. Gupta, I. Mumick, and V. Subrahmanian. Maintaining views incrementally. In *Preceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, pages 157–166, Washington, D.C., May 1993.

[GW94] P. Grefen and J. Widom. Integrity constraint checking in federated databases. Submitted for publication, 1994.

[Han87] E.N. Hanson. A performance analysis of view materialization strategies. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pages 440–453, 1987.

[HD92] J.V. Harrison and S.W. Dietrich. Maintenance of materialized views in a deductive database: An update propagation approach. In *Proceedings of the 1992 JICLSP Workshop on Deductive Databases*, pages 56–65, 1992.

[IK93] W.H. Inmon and C. Kelley. *Rdb/VMS: Developing the Data Warehouse.* QED Publishing Group, Boston, London, Toronto, 1993.

[LHM+86] B. Lindsay, L.M. Haas, C. Mohan, H. Pirahesh, and P. Wilms. A snapshot differential refresh algorithm. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, Washington, D.C., May 1986.

[QW91] X. Qian and G. Wiederhold. Incremental recomputation of active relational expressions. *IEEE Transactions on Knowledge and Data Engineering*, 3(3):337–341, September 1991.

[RK86] N. Roussopoulos and H. Kang. Preliminary design of ADMS+-: A workstation-mainframe integrated architecture for database management systems. In *Proceedings of the Twelfth International Conference on Very Large Data Bases*, pages 355–364, Kyoto, Japan, August 1986.

[SF90] A. Segev and W. Fang. Currency-based updates to distributed materialized views. In *Proceedings of the Sixth International Conference on Data Engineering*, pages 512–520, Los Alamitos, February 1990.

[SF91] A. Segev and W. Fang. Optimal update policies for distributed materialized views. *Management Science*, 37(7):851–70, July 1991.

[SI84] O. Shmueli and A. Itai. Maintenance of views. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, pages 240–255, Boston, Massachusetts, May 1984.

[SP89a] A. Segev and J. Park. Maintaining materialized views in distributed databases. In *Proceedings of the Fifth International Conference on Data Engineering*, pages 262–70, Los Angeles, February 1989.

[SP89b] A. Segev and J. Park. Updating distributed materialized views. *IEEE Transactions on Knowledge and Data Engineering*, 1(2):173–184, June 1989.

[TB88] F.WM. Tompa and J.A. Blakeley. Maintaining materialized views without accessing base data. *Information Systems*, 13(4):393–406, 1988.

[ZGMHW94] Y. Zhuge, H. Garcia-Molina, J. Hammer, and J. Widom. View maintenance in a warehousing environment. Technical report, Stanford University, October 1994. Available via anonymous ftp from host `db.stanford.edu` as `pub/zhuge/1994/anomaly-full.ps`.

A Examples of ECA Applications

The following three scenarios are additional examples of ECA applications in a warehousing environment.

Example 7: Insertion

This example has the same base relations and source updates as in Example 4, but the event order is different:

r_1 :

W	X
1	2

r_2 :

X	Y

r_3 :

X	Y

The view is defined to be $V = \Pi_W(r_1 \bowtie r_2 \bowtie r_3)$ as in Example 4, and is empty originally.

1. Warehouse receives $U_1 = insert(r_1, [4, 2])$;
   Warehouse sends $Q_1 = V\langle U_1 \rangle = \Pi_W([4, 2] \bowtie r_2 \bowtie r_3)$.
2. Warehouse receives $U_2 = insert(r_3, [5, 3])$; Now $UQS = \{Q_1\}$
   Warehouse sends $Q_2 = V\langle U_2 \rangle - Q_1\langle U_2 \rangle = \Pi_W(r_1 \bowtie r_2 \bowtie [5, 3]) - \Pi_W([4, 2] \bowtie r_2 \bowtie [5, 3])$
3. Warehouse receives $A_1 = \emptyset$, COLLECT $= \emptyset$, $UQS = \{Q_2\}$
4. Warehouse receives U_3= $insert(r_2, [2, 5])$; $UQS = \{Q_2\}$. Warehouse sends
   $$\begin{aligned} Q_3 &= V\langle U_3 \rangle - Q_2\langle U_3 \rangle \\ &= \Pi_W(r_1 \bowtie [2, 5] \bowtie r_3) - \Pi_W(r_1 \bowtie [2, 5] \bowtie [5, 3]) + \Pi_W([4, 2] \bowtie [2, 5] \bowtie [5, 3]) \end{aligned}$$
5. Warehouse receives $A_2 = [1]$, COLLECT $= ([1])$, $UQS = \{Q_3\}$.
6. Warehouse receives $A_3 = [4]$, COLLECT $=$ ([1],[4]), $UQS = \emptyset$. The view is updated to be: $MV = \emptyset +$ COLLECT $= ([1], [4])$,

□

Example 8: Deletion

Base relations:

r_1

W	X
1	2
4	2

r_2

X	Y
2	3

The view is defined to be $V = \Pi_W(r_1 \bowtie r_2)$. It contains two tuples ([1], [4]) before any updates.

The event list at the warehouse is:

1. Warehouse receives $U_1 = delete(r_1, [4, 2])$
   Warehouse sends $Q_1 = V\langle U_1 \rangle = \Pi_W((-[4, 2]) \bowtie r_2)$

2. Warehouse receives $U_2 = delete(r_2, [2,3])$; $UQS = \{Q_1\}$.

   Warehouse sends

$$
\begin{aligned}
Q_2 &= V\langle U_2\rangle - Q_1\langle U_2\rangle \\
&= \Pi_W(r_1 \bowtie (-[2,3])) - \Pi_W((-[4,2]) \bowtie (-[2,3])) \\
&= -\Pi_W(r_1 \bowtie [2,3]) - \Pi_W([4,2] \bowtie [2,3])
\end{aligned}
$$

3. Warehouse receives $A_1 = \emptyset$, `COLLECT` $= \emptyset$, $UQS = \{Q_2\}$

4. Warehouse receives $A_2 = (-[4], -[1])$, `COLLECT` = (-[4], -[1]), $UQS = \emptyset$ So the view is updated to be $MV = ([1],[4]) +$ `COLLECT` $= \emptyset$, which is correct. □

Example 9: Insertion and Deletion

Base relations:

r_1

W	X
1	2
4	2

r_2

X	Y
-	-

The view is defined to be $V = \Pi_W(r_1 \bowtie r_2)$ as in Example 8. $MV = \emptyset$ before any updates. The event list at the warehouse is:

1. Warehouse receives $U_1 = delete(r_1, [4,2])$

   Warehouse sends $Q_1 = V\langle U_1\rangle = \Pi_W((-[4,2]) \bowtie r_2)$

2. Warehouse receives $U_2 = insert(r_2, [2,3])$; $UQS = \{Q_1\}$.

   Warehouse sends

$$
\begin{aligned}
Q_2 &= V\langle U_2\rangle - Q_1\langle U_2\rangle \\
&= \Pi_W(r_1 \bowtie [2,3] - \Pi_W((-[4,2]) \bowtie ([2,3])) \\
&= \Pi_W(r_1 \bowtie [2,3] + \Pi_W([4,2] \bowtie ([2,3]))
\end{aligned}
$$

3. Warehouse receives $A_1 = -[4]$, `COLLECT` = (-[4]), $UQS = \{Q_2\}$.

4. Warehouse receives $A_2 = [1] + [4]$, `COLLECT` = ([1]), $UQS = \emptyset$.

   Now the view is updated and the result is $MV = \emptyset +$ `COLLECT` $= [1]$, which is correct. □

B Correctness of the Eager Compensating Algorithm

In this section we provide a correctness proof for the Eager Compensating Algorithm (ECA) as presented in Section 5. Theorem B.1 shows that ECA is strongly consistent by proving that it is both consistent and convergent. Before getting to the theorem, we need to establish five lemmas.

First, we claim that not all events will affect outcome of a query evaluation. For example, if a source event se has type S_qu, then the base relations are not modified. Consequently, the result of evaluating another query before or after se are the same.

Formally, let se_i be an arbitrary source event; the source state before se_i is ss_{i-1} and after is ss_i. Let we_i be a warehouse event and the warehouse state before we_i is ws_{i-1} and after is ws_i.

Lemma B.1 *Independent events using ECA.*

se_i *has type* $S\_qu_i \Longrightarrow V[ss_{i-1}] = V[ss_i],\ Q[ss_{i-1}] = Q[ss_i]$

we_i *has type* $W\_up_i \Longrightarrow V[ws_{i-1}] = V[ws_i]$

we_i *has type* W_ans_i *with* $UQS \neq \emptyset \Longrightarrow V[ws_{i-1}] = V[ws_i]$

Proof: Obvious from the definition of events (Section 3) and events described in ECA. □

With ECA, the triggering relationship among source and warehouse events is as shown in Figure B.6.

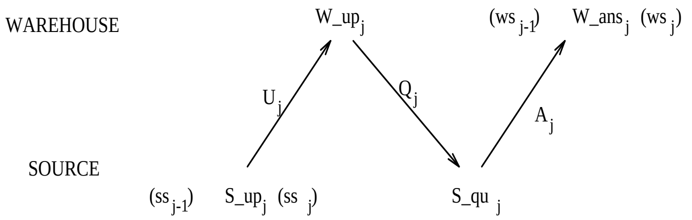

Figure B.6: Triggering relations among events.

Since there is a one-to-one correspondence between updates and the four events they trigger, in the rest of the appendix we will label events by the id of the update that generated them. Specifically, let $U_1, U_2, \ldots$ be the updates that occurred during the execution of ECA algorithm. For each U_j, the events it triggers are S_up_j, W_up_j, S_qu_j and W_ans_j. Also, let Q_j be the query involved in the processing of U_j and A_j be the answer relation of Q_j. According to Lemma B.1, at the source only events of type S_up affect query evaluation, so we can assume the source state only change after each S_up event, that is, we can label the source state before S_up_j as ss_{j-1} and label it after S_up_j as ss_j. Similarly, assume warehouse states only change after each type W_ans events.

In the next lemma, we show how a type S_up event affects the evaluation of a query. Assume Q is an arbitrary query and is targeted for execution before a source update U_j(at state ss_{j-1}). However it is actually executed after U_j(at state ss_j).

Lemma B.2 $Q[ss_{j-1}] = Q[ss_j] - Q\langle U_j\rangle[ss_j]$ *for any query* Q.

Proof:

Consider a term T in Q as defined by Formula 4.1: $T = \Pi_X(\sigma_C(\tilde{r}_1 \times \tilde{r}_2 \times \ldots \times \tilde{r}_n))$. Assume U_j is an update on relation r_m. If r_m is not used in T, then $T\langle U_j\rangle = \emptyset$, and $T[ss_j] = T[ss_{j-1}]$.

Therefore we have $T[ss_j] = T[ss_{j-1}] + T\langle U_j\rangle[ss_j]$ holds. If r_m is used in defining T, then for r_l other than r_m, $\tilde{r}_l[ss_j] = \tilde{r}_l[ss_{j-1}]$. And $\tilde{r}_m[ss_j] = \tilde{r}_m[ss_{j-1}] + tuple(U_j)$. Then

$$\begin{aligned}
T[ss_j] &= \Pi_X(\sigma_C(\tilde{r}_1[ss_j] \times \tilde{r}_2[ss_j] \times \ldots \times \tilde{r}_m[ss_j] \ldots \times \tilde{r}_n[ss_j])) \\
&= \Pi_X(\sigma_C(\tilde{r}_1[ss_j] \times \tilde{r}_2[ss_j] \times \ldots \times (\tilde{r}_m[ss_{j-1}] + tuple(U_j)) \ldots \times \tilde{r}_n[ss_j])) \\
&= \Pi_X(\sigma_C(\tilde{r}_1[ss_{j-1}] \times \tilde{r}_2[ss_{j-1}] \times \ldots \times \tilde{r}_m[ss_{j-1}] \ldots \times \tilde{r}_n[ss_{j-1}])) + \\
&\quad\ \Pi_X(\sigma_C(\tilde{r}_1[ss_j] \times \tilde{r}_2[ss_j] \times \ldots \times tuple(U_j) \ldots \times \tilde{r}_n[ss_j])) \\
&= T[ss_{j-1}] + T\langle U_j\rangle[ss_j].
\end{aligned}$$

Since $Q = \sum_i T_i$ defined by Formula 4.2 is a sum of terms, we have $Q[ss_j] = Q[ss_{j-1}] + Q\langle U_j\rangle[ss_j]$. Thus $Q[ss_{j-1}] = Q[ss_j] - Q\langle U_j\rangle[ss_j]$. $\square$

In ECA, if $UQS(W\_ans_j) = \emptyset$, then the warehouse updates the materialized view MV. The following lemma states that when the warehouse updates MV, the last answer it gets(A_j) is the intended one for Q_j .

Lemma B.3 $UQS(W\_ans_j) = \emptyset \implies A_j = Q_j[ss_j]$.

Proof:

Assume there exists an S_up_m, such that $S\_up_j < S\_up_m < S\_qu_j$. Then from the triggering rule described in Section 3 we know that $W\_up_j < W\_up_m < W\_ans_j$, and $W\_ans_j < W\_ans_m$. Therefore, $Q_m \in UQS(W\_ans_j)$ and this implies that $UQS(W\_ans_j) \neq \emptyset$. This is a contradiction, therefore all source events between S_up_j and S_qu_j(if there are any) must be of type S_qu. By Lemma B.1 and our labeling convention, the source state right before S_qu_j is still ss_j, and A_j is evaluated at this state, $A_j = Q_j[ss_j]$. $\square$

Definition: An event of type W_ans updates the materialized view at the warehouse when its UQS is empty. Define W_ans_i be the *Previous-Updating-Event* of a warehouse event we if $W\_ans_i < we$, $UQS(W\_ans_i) = \emptyset$, and $\not\exists W\_ans_m$ such that $W\_ans_i < W\_ans_m < we$ and $UQS(W\_ans_m) = \emptyset$.

The next lemma shows that for two consecutive updates on the warehouse view, if the first one updates the view into a consistent state, then so does the second one.

Lemma B.4 *Assume* $UQS(W\_ans_j) = \emptyset$ *and* W_ans_i *is the Previous-Updating-Event of* W_ans_j. *Then* $V[ws_i] = V[ss_i] \implies V[ws_j] = V[ss_j]$.

Proof: We know that the updates occurring between U_i and U_j (if any) are U_{i+1}, U_{i+2}, ..., U_{j-1}. Before updating MV at event W_ans_j, the $COLLECT$ set is $COLLECT = \sum_{l=i+1}^{j} A_l$ since the last time $COLLECT$ was set to empty was at event W_ans_i. Furthermore we have $V[ws_j] = V[ws_i] + COLLECT$.

Now we can reduce the proof that $V[ws_j] = V[ss_j]$ assuming $V[ws_i] = V[ss_i]$ to the proof of the following equation with $COLLECT = \sum_{l=i+1}^{j} A_l$:

$$V[ss_i] + COLLECT = V[ss_j] \tag{B.3}$$

The proof of Equation B.3 is done by induction on the number of updates between U_i and U_j, which is $m = j - i - 1 \geq 0$.

1. Base case: $m = 0$.

   There is no update between U_i and U_j so U_i and U_j are consecutive updates. Since $UQS(W\_ans_j) = \emptyset$, we have $Q_j = V\langle U_j\rangle - \emptyset = V\langle U_j\rangle$.

$$\begin{aligned} V[ss_i] + A_j &= V[ss_i] + Q_j[ss_j] && (Lemma\ B.3) \\ &= V[ss_i] + V\langle U_j\rangle[ss_j] && (ECA) \\ &= V[ss_j] && (Lemma\ B.2) \end{aligned}$$

   Thus, Equation B.3 holds for the base case.

2. Induction step:

   Induction hypothesis: Assume Equation B.3 holds for all $m < c$.

   When $m = c$, we have c updates between U_i and U_j. Let us refer this scenario with c updates as Scenario S. In Scenario S, some of the queries in $Q_{i+1} \ldots Q_{j-1}$ are evaluated before U_j occurs at the source, some after.

   Let

$$Q_{before} = \sum_{\substack{l=i+1 \\ Q_l \notin UQS(W\_up_j)}}^{j-1} Q_l \quad \text{and} \quad Q_{after} = \sum_{\substack{l=i+1 \\ Q_l \in UQS(W\_up_j)}}^{j-1} Q_l$$

   Correspondingly, let

$$A_{before} = \sum_{\substack{l=i+1 \\ Q_l \notin UQS(W\_up_j)}}^{j-1} A_l \quad \text{and} \quad A_{after} = \sum_{\substack{l=i+1 \\ Q_l \in UQS(W\_up_j)}}^{j-1} A_l$$

   So

$$COLLECT = A_{before} + A_{after} + A_j \tag{B.4}$$

   Let Scenario S' be exactly the same as Scenario S except U_j and all four corresponding events do not occur. In Scenario S' we label all events and states with a prime. Thus the updates are $U_1', U_2', \ldots$, the queries are $Q_1', Q_2', \ldots$, and so on. Many of the S' events and states correspond to the ones in S. In particular, $V' = V$(we have the same view), and $\forall l, i \leq l \leq j-1$, $U_l' = U_l$, $Q_l' = Q_l$ and $ss_l' = ss_l$.

   Since U_j is absent in Scenario S', $UQS(W\_ans_{j-1}') = \emptyset$. There are $c - 1 < c$ updates between U_i and U_{j-1}. By induction hypothesis we have

$$V'[ss_i'] + COLLECT' = V'[ss_{j-1}'] \tag{B.5}$$

   In Scenario S', $COLLECT'$ is the sum of answers A_i' though A_{j-1}'. Let us divide the answers into two sets:

   A_{before}' = answers to the queries that constitute Q_{before}

   A_{after}' = answers to the queries that constitute Q_{after}

   Thus, $COLLECT' = A_{before}' + A_{after}'$, Equation B.5 can be written as

$$V[ss_i] + A_{before}' + A_{after}' = V[ss_{j-1}] \tag{B.6}$$

A_{before} are those answers computed before U_j in Scenario S, they are computed at the same state in Scenario S', thus $A_{before} = A'_{before}$.

A_{after} contains those answers that were computed after U_j occur in Scenario S. Since $UQS(W\_ans_j) = \emptyset$, by similar arguments to those in Lemma B.3, we know that all answers in A_{after} are computed at state $ss_j = state(S\_up_j)$:

$$A_{after} = Q_{after}[ss_j] \tag{B.7}$$

In Scenario S', answers in A'_{after} must be computed after U_{j-1} occurs, otherwise they will not be evaluated after U_j in Scenario S. So $A'_{after} = Q_{after}[ss_{j-1}]$, and Equation B.6 becomes:

$$V[ss_i] + A_{before} + Q_{after}[ss_{j-1}] = V[ss_{j-1}] \tag{B.8}$$

Furthermore, we know that $Q_j = V\langle U_j\rangle - Q_{after}\langle U_j\rangle$ from Algorithm ECA. This gives us $Q_j[ss_j] = V\langle U_j\rangle[ss_j] - Q_{after}\langle U_j\rangle[ss_j]$. From Lemma B.2 we have that $Q_{after}[ss_{j-1}] = Q_{after}[ss_j] - Q_{after}\langle U_j\rangle[ss_j]$. Subtracting the latter from the former equation we obtain

$$Q_{after}[ss_j] + Q_j[ss_j] = Q_{after}[ss_{j-1}] + V\langle U_j\rangle[ss_j] \tag{B.9}$$

The intuitive meaning of Equation B.9 is that if the queries in Q_{after} and Q_j are evaluated in state ss_j as in Scenario S, the result will be the same as evaluating those queries in Q_{after} in state ss_{j-1} as in Scenario S', and adding the term $V\langle U_j\rangle[ss_j]$ that reflects the effect of U_j on the view.

Combining all of our results we can show that

$$\begin{aligned}
& V[ss_i] + COLLECT \\
= \; & V[ss_i] + A_{before} + A_{after} + A_j && (\,B.4) \\
= \; & V[ss_i] + A_{before} + Q_{after}[ss_j] + A_j && (\,B.7) \\
= \; & V[ss_i] + A_{before} + Q_{after}[ss_j] + Q_j[ss_j] && (Lemma\ B.3) \\
= \; & V[ss_i] + A_{before} + Q_{after}[ss_{j-1}] + V\langle U_j\rangle[ss_j] && (\,B.9) \\
= \; & V[ss_{j-1}] + V\langle U_j\rangle[ss_j] && (\,B.8) \\
= \; & V[ss_j] && (Lemma\ B.2)
\end{aligned}$$

This shows that Equation B.3 holds in Scenario S, and completes the induction proof. As stated earlier, Equation B.3 is sufficient to prove Lemma B.4. □

Lemma B.5 *Assume* $V[ws_0] = V[ss_0]$, *then* $UQS(W\_ans_j) = \emptyset \Longrightarrow V[ws_j] = V[ss_j]$ *for any* j.

<u>Proof:</u> The proof is done by induction on j.

1. Base case: $j = 1$.

   Assume there is an 'empty update' U_0 which occurs before any updates and its four events (which are also empty) are processed before any other events. After S_up_0 the source state is ss_0 and after W_ans_0 the warehouse state is ws_0. Since there's no pending queries when W_ans_0 is processed, W_ans_0 is the *Previous-Updating-Event* of W_ans_1.

Since $V[ws_0] = V[ss_0]$ and $UQS(W\_ans_1) = \emptyset$, by Lemma B.4 we have $V[ws_1] = V[ss_1]$.

2. Induction step: Induction hypothesis: the lemma holds for any $j < c$ ($c > 1$).

   When $j = c$, let W_ans_i be the *Previous-Updating-Event* of W_ans_j. So $i < c$ and by induction hypothesis we have $V[ws_i] = V[ss_i]$. Then from Lemma B.4 we can prove $V[ws_j] = V[ss_j]$.

□

Theorem B.1 *Algorithm ECA is strongly consistent for a finite sequence of updates $U_1, U_2, \ldots, U_k$.*

Proof:

Assume $V[ws_0] = V[ss_0]$, the warehouse is in a consistent state with the source before any events occur.

`Consistent:`

The definition of consistency is given in Section 3. For every pair of warehouse states $ws_i < ws_j$ with corresponding events $we_i \leq we_j$, let $W\_ans_{i'}$ be the Previous-Updating-Event of we_i and $W\_ans_{j'}$ be the Previous-Updating-Event of we_j. $W\_ans_{i'} \leq W\_ans_{j'}$ obviously. From Lemma B.1, we know $V[ws_i] = V[ws_{i'}]$ and $V[ws_j] = V[ws_{j'}]$.

Since the UQS of both $W\_ans_{i'}$ and $W\_ans_{j'}$ are empty, Lemma B.5 gives that $V[ws_{i'}] = V[ss_{i'}]$ and $V[ws_{j'}] = V[ss_{j'}]$. Also, $W\_ans_{i'} \leq W\_ans_{j'}$ implies $ss_{i'} \leq ss_{j'}$.

So there exist $ss_n = ss_{i'} \leq ss_l = ss_{j'}$ such that $V[ws_i] = V[ss_n]$ and $V[ws_j] = V[ss_l]$. By definition, ECA is consistent.

`Convergent:`

$UQS(W\_ans_k) = \emptyset$ since U_k is the last update. Lemma B.5 gives $V[ws_k] = V[ss_k]$. At the source, the last update is in S_up_k and after that, all events(if any) have type S_qu. So ss_k is the final source state. At the warehouse, W_ans_k is the last event so ws_k is the final warehouse state. Thus ECA is convergent.

Since ECA is consistent and convergent, it is strongly consistent. □

C Sketch of Proof of Correctness for ECA-Key algorithm

Consider a warehouse event we_f where $UQS(we_f) = \emptyset$. The warehouse state right after this event, $state(we_f) = ws_f$. Let F be the set of all updates that have been seen at the warehouse by we_f. Let $U_f \in F$ be the last update received before or by we_f.

Let se_f be the *S_up* source event that processed U_f, and let ss_f be the source state immediately after se_f. In the rest of this sketch we discuss how to show that $V[ss_f] = V[ws_f]$. Given this, it is straightforward to see that ECA^K is strongly consistent, since the materialized view only changes at the warehouse when $UQS = \emptyset$.

We can show that $V[ss_f] = V[ws_f]$ by contradiction and considering the various cases. That is, assume that $V[ss_f] \neq V[ws_f]$. Then there must be a tuple t that is either missing from $V[ws_f]$ or is an extra tuple in $V[ws_f]$. (In what follows, we do not have to worry about t appearing more than once either at the source or the warehouse due to our key condition.)

Case I: $t \in V[ss_f]$ and $t \notin V[ws_f]$.

Subcase I(a): There is (at least one) insert at the source (before se_f) that generates t. Let U_i be the last of these events. Insert U_i adds a tuple at the source that contains one of the keys involved in t. (Recall that t has a key value for each of the relations involved in the view definition.)

Note there can be no deletes involving a t key after U_i is processed at the source (at least not before or at se_f). If there were, they would remove t from $V[se_f]$, a contradiction.

Under ECA^K, U_i gets propagated to the warehouse, a query Q is generated, answered at the source, and an answer returned to the warehouse. At the warehouse, both the processing of U_i and the receipt of the Q answer must occur before or at we_f. (If U_i were received after we_f, then $U_i \notin F$ and U_i must have been processed after se_f. If the Q answer arrived after we_f, then $UQS(we_f)$ would not have been empty.) When Q is processed at the source, it sees all the key values involved in t, so the answer to Q contains t. Thus, t is inserted into the materialized view. Because there are no deletes affecting t at the source between the processing of U_i and se_f, there can be no such deletes at the warehouse, and t remains until we_f, a contradiction.

Subcase I(b): There are no inserts at the source that generate t. This means that t must have been in $V[ss_0]$, where ss_0 is the initial source state. Since the initial materialized view is correct, t must also be in $V[ws_0]$.

We can now proceed similarly to Case I(a), arguing that there are no deletes affecting t at the source (before se_f) and hence no such deletes at the warehouse. Thus, t remains at the warehouse, a contradiction.

Case II: $t \notin V[ss_f]$ and $t \in V[ws_f]$.

Subcase II(a): There is at least one delete at the source that involves a t key value. The proof is analogous to Case I(a), with one extra case to consider. We consider the last such source delete U_d, where se_d is the processing of U_d at the source and we_d is the processing of U_d at the warehouse. Clearly, we_d deletes t from the materialized view. Since $t \in V[ws_f]$, it must have been reinserted. Notice that there are no inserts that could generate t at the source after se_d, so there cannot be any corresponding inserts that warehouse. However, it could be the case that an insert U_j occurring at the source *before* U_d, could generate an answer that is processed at the warehouse *after* we_d, adding t. The query generated for U_j must have been processed at the source after se_d (else the answer would have arrived before we_d). Since the query is executed at the source after U_d is processed, it does not see one of the key values of t (i.e., the one deleted by U_d). Thus, there is no way the answer could contain t, and t is not added to the materialized view, a contradiction.

Subcase II(b): There are no source deletes that involve a key value of t. Therefore t must not be in the initial $V[ss_0]$, and since the initial materialized view is correct, t is not in $V[ws_0]$. Since somehow t appears in $V[ws_f]$, it must have been inserted by some answer received at the warehouse. Say se_k is the source event that generates the answer. At that point $V[state(se_k)]$ must contain t. Since there are no deletes affecting t at the source, then t remains, a contradiction.

D Performance Evaluation

D.1 Recomputing the View

The following is a description of the "recompute view" strategy (RV), it updates MV after every s source updates.

Algorithm D.1 (Algorithm RV)

Events at the source are as in Algorithm 5.1. At the warehouse (initially `Count` $= 0$):

- W_up_i: receive U_i;
  let `Count` = `Count`+1;
  if `Count` = s
  then {let `Count` = 0; let $Q_i = V$;
  send Q_i to the source;
  trigger event S_qu_i at the source
  }
- W_ans_i: receive A_i;
  let $MV \leftarrow A_i$

(end algorithm)

D.2 Performance Based on Number of Bytes Transferred

1. RV best and worst case computation:

   Let $card(r_i)$ be the cardinality of relation r_i. The total bytes transferred is equal to the size of the resulting view:

   $$\begin{aligned} B_{RVBest} &= (W, Z\text{bytes}) \times (\text{ number of tuples in the resulting view relation}) \\ &= S \times \sigma \times (card(r_1) \times J(r_2, X) \times J(r_3, Y)) \\ &= S\sigma C J^2 \end{aligned}$$

   For the worst case, the view is recomputed after each update, i.e., three times:

   $$B_{RVWorst} = 3 \times B_{RVBest} = 3S\sigma CJ^2$$

2. ECA best case computation:

   The warehouse events occur in this order: $U_1, Q_1, A_1, U_2, Q_2, A_2, U_3, Q_3, A_3$

   Using ECA, we know

$Q_1 = V\langle U_1\rangle = \Pi_{W,Z}(\sigma_{cond}(t_1 \bowtie_X r_2 \bowtie_Y r_3))$
$Q_2 = V\langle U_2\rangle = \Pi_{W,Z}(\sigma_{cond}(r_1 \bowtie_X t_2 \bowtie_Y r_3))$
$Q_3 = V\langle U_3\rangle = \Pi_{W,Z}(\sigma_{cond}(r_1 \bowtie_X r_2 \bowtie_Y t_3))$

The size of A_1 can be evaluated as: $B = S \times \sigma \times 1 \times J(r_2, X) \times J(r_3, Y) = S \times \sigma \times J^2$

The size of A_2 and A_3 can be estimated in a similar fashion.

$B_{ECABest} = 3S\sigma J^2$

3. ECA worst case computation:

   The event order is: $U_1, Q_4, U_2, Q_5, U_3, Q_6, A_4, A_5, A_6$

   The queries are:

   $Q_4 = V\langle U_1\rangle = Q_1$

   $Q_5 = V\langle U_2\rangle - Q_4\langle U_2\rangle = V\langle U_2\rangle - \Pi_{W,Z}(\sigma_{cond}(t_1 \bowtie_X t_2 \bowtie_Y r_3))$

   $Q_6 = V\langle U_3\rangle - Q_4\langle U_3\rangle - Q_5\langle U_3\rangle = V\langle U_3\rangle - \Pi_{W,Z}(\sigma_{cond}(t_1 \bowtie_X r_2 \bowtie_Y t_3)) - \Pi_{W,Z}(\sigma_{cond}(r_1 \bowtie_X t_2 \bowtie_Y t_3)) + \Pi_{W,Z}(\sigma_{cond}(t_1 \bowtie_X t_2 \bowtie_Y t_3))$

   Notice that the cost for the first terms of Q_4, Q_5, Q_6 is equal to $B_{ECABest}$. For the second term in Q_5, $\Pi_{W,Z}(\sigma_{cond}(t_1 \bowtie_X t_2 \bowtie_Y r_3))$, $B = S \times \sigma \times J(r_3, Y)$, i.e., $B = S\sigma J$. The cost of the second and third terms in Q_6 is the same. The cost of the last term in Q_6 is 0; it does not have to be sent to the source for evaluation.

   For the total cost, the costs for these terms are added. (Recall that "$-$" is not a set difference operator; it is a union of signed tuples.) Thus $B_{ECAWorst} = B_{ECABest} + 3S\sigma J$, or $B_{ECAWorst} = 3S\sigma J(J + 1)$.

4. For the k update case, we assume that the first three updates are as before, and then we get $k - 3$ additional ones.

   For RV it is easy to see that B_{RVBest} does not change, and $B_{RVWorst} = kB_{RVBest} = kS\sigma CJ^2$.

   For the best ECA case, it is easy to see that our computation above generalizes to $B_{ECABest} = kS\sigma J^2$.

   For the ECA worst case, consider a fourth update U_4 on relation r_1. Assuming all queries are executed after all updates (worst case), ECA issues queries Q_4, Q_5, Q_6 as before. The last query is $Q_7 = V\langle U_4\rangle Q_1\langle U_4\rangle Q_2\langle U_4\rangle Q_3\langle U_4\rangle$. That is, $Q_7 = V\langle U_4\rangle - V\langle U_1, U_4\rangle - V\langle U_2, U_4\rangle + V\langle U_1, U_2, U_4\rangle - V\langle U_3, U_4\rangle + V\langle U_1, U_3, U_4\rangle + V\langle U_2, U_3, U_4\rangle - V\langle U_1, U_2, U_3, U_4\rangle$. Since U_1 and U_4 update the same relation, all terms of the form $V\langle U_1, U_4, ...\rangle$ are empty. Similarly, all terms with three or more updates are empty (no compensating query needs to be sent since all data needed is already at the warehouse). This leaves the first term with data transfer cost $S\sigma J^2$ and two compensating queries with costs $S\sigma J$. Notice that the cost would be the same if U_4 had modified any of the two other relations.

   So, in general, we see that the cost of processing update U_j is $S\sigma J^2$ (for the Q_j term $V\langle U_j\rangle$) plus a term $S\sigma J$ (for $V\langle U_j, U_p\rangle$) for each previous update U_p such that $rel(U_j) \neq rel(U_p)$. If we assume that updates access any of the three relations with equal probability, then the cost for U_j is $S\sigma J^2 + (2(j-1)/3)S\sigma J$. Adding the terms from 1 to k we get $B_{ECAWorst} = kS\sigma J^2 + k(k-1)S\sigma J/3$.

D.3 Performance Based on I/O

D.3.1 Scenario 1

1. RV best and worst case computation:

   Assuming that J is not too small, the best strategy is for the source to read into memory all three relations.

   $$IO_{RVBest} = \lceil \frac{card(r_1)}{K} \rceil + \lceil \frac{card(r_2)}{K} \rceil + \lceil \frac{card(r_3)}{K} \rceil = 3\lceil \frac{C}{K} \rceil = 3I$$

   where $I = \lceil \frac{C}{K} \rceil$ is the number of IOs needed to read an entire base relation.

   In the worst case, RV computes the view three times, so $IO_{RVWorst} = 3 \times IO_{RVBest} = 9I$.

2. For the ECA best case, we have the same three queries Q_1, Q_2, Q_3 as in Section D.2. Let us first assume that $J < I$ so it is best to use the indexes for evaluating all the queries.

   For Q_1, we use the X value in t_1 to fetch the matching r_2 tuples (using the X index on r_2). Since the X index is clustered, all the r_1 tuples are contiguous. This gives the first term below. Then, for each of the $J(r_2, X)$ tuples found, we get the matching r_3 tuples, giving us the second term. Thus,

   $$IO_1 = \lceil \frac{J(r_2, X)}{K} \rceil + (J(r_2, X) \times \lceil \frac{J(r_3, Y)}{K} \rceil).$$

   Assuming that $J \leq K$ (which holds for the value ranges we are interested in here), this reduces to

   $$IO_1 = 1 + J.$$

   As discussed in the main body of the paper, the select condition in Q_1 is evaluated after the join since it involves W and Z attributes. This implies that the condition does not save us any I/Os. Note that we did not consider caching effects: whenever we probe a relation, we go to disk to read the block. Hence, the results for ECA are pessimistic, even for the evaluation of a single term.

   For Q_2, we use the t_2 tuple values to probe the r_1 and r_3 indexes. Thus,

   $$IO_2 = \lceil \frac{J(r_1, X)}{K} \rceil + \lceil \frac{J(r_3, Y)}{K} \rceil = 2.$$

   For Q_3, we have to use an un-clustered index to probe r_2, so for each of the $J(r_2, Y)$ tuples we need an I/O. This gives us the first term below. Then we use the X index to probe r_1 giving us the second term.

   $$IO_3 = J(r_2, Y) + (J(r_2, Y) \times \lceil \frac{J(r_1, X)}{K})\rceil = J + J.$$

   Therefore $IO_{ECABest} = IO_1 + IO_2 + IO_3 = 3J + 3$.

   If $J \geq I$, Q_3 can be best evaluated by reading in both r_1 and r_2, for a total cost of $2I$. For Q_1, it is also better to read in r_3 fully (cost is I) rather than piecemeal (cost J), so the cost of Q_1 is $1 + I$. Query Q_2 is still done best with indexes. Thus, the total cost when $J \geq I$ is $3I + 3$.

In general, and assuming that the query optimizer can select the best plan, the cost of evaluating the three queries will be $IO_{ECABest} = 3min(J, I) + 3$.

3. ECA worst case computation:

   For $V\langle U_1\rangle, V\langle U_2\rangle$ and $V\langle U_3\rangle$, the number of I/Os is the same as obtained for the best case.

   For term the $\Pi_{W,Z}(\sigma_{cond}(t_1 \bowtie_X t_2 \bowtie_Y r_3))$ in Q_5, $IO = \lceil J(r_3, Y)/K \rceil = 1$

   Similarly, for the extra terms in Q_6 we have $IO = \lceil J(r_2, X)/K \rceil + \lceil J(r_1, X)/K \rceil = 2$

   Thus, $IO_{ECAWorst} = IO_{ECABest} + 3 = 3min(J, I) + 6$

D.3.2 Scenario 2

We assume there are no indices on any attributes. We also assume that the memory can only hold three blocks, which are used for a nested join algorithm.

1. RV computation:

   To compute $r_1 \bowtie r_2 \bowtie r_3$, the source can load one r_1 and one r_2 block into memory and loop through r_3. Then it repeats this for all possible pairs of r_1, r_2 blocks. Thus,

   $IO_{RVBest} = I^3$

   $IO_{RVWorst} = 3 \times IO_{RVBest} = 3I^3$.

2. $IO_{ECABest}$ computation:

   Using Q_1 as an example, the source can load two blocks of r_2 into memory and loop through the r_3 records (number of I/Os is I). This is repeated $\lceil C/(2K) \rceil$ times, so $IO_1 = I \times I'$, where $I' = \lceil C/(2K) \rceil$ is the number of double-block buffers needed to hold a relation.

   The cost for the other queries is the same, so altogether, $IO_{ECABest} = 3II'$.

3. $IO_{ECAWorst}$ computation:

   For each of the three extra terms in the compensating queries we need to read a relation, so $IO_{ECAWorst} = 3II' + 3I = 3I(I' + 1)$

D.3.3 *k* Update Case

We extend the Scenario 1 and 2 results for the case of k updates. The analysis is similar to that at the end of Section D.2. For simplicity, assume that $J < I$.

For Scenario 1:

1. $IO_{RVBest} = 3I$ since the view is still recomputed once.
2. $IO_{RVWorst} = 3kI$.
3. $IO_{ECABest} = k \times (3J + 3)/3 = k(J + 1)$ because we assumed updates have equal probability on any of the three relations, and the total IO for three updates on three relations are $3J + 3$ as we computed before (and $J < I$).
4. $IO_{ECAWorst} = k(J + 1) + k(k - 1)/3$

   The cost of evaluating $V\langle U_j\rangle$ is $J + 1$, and the cost of evaluating $V\langle U_j, U_p\rangle$ is 1, so the total is $\sum_{j=1} k((J + 1) + 2(j - 1)/3) = k(J + 1) + k(k - 1)/3$.

For Scenario 2:

1. $IO_{RVBest} = I^3$ since the view is still recomputed once.

2. $IO_{RVWorst} = kI^3$.

3. $IO_{ECABest} = kII'$

4. $IO_{ECAWorst} = kII' + Ik(k-1)/3$.

   To compute the last expression, we note that for update U_j, the query Q_j contains $V\langle U_j \rangle$ plus (on average) $2(j-1)/3$ terms of the form $V\langle U_j, U_p\rangle$ where $rel(U_j) \neq rel(U_p)$. The cost of evaluating Q_j is thus $II' + 2(j-1)I/3$. Adding from Q_1 to Q_k we obtain the expression above.