Paper Number AMERICA196

Performance Analysis of WHIPS Incremental Maintenance

Yue Zhuge and Hector Garcia-Molina
Computer Science Department
Stanford University
Stanford, CA 94305-2140, USA
{zhuge,hector}@cs.stanford.edu

Abstract

We consider materialized view maintenance, where views are integrated from multiple distributed data sources and stored in a data warehouse. Incremental maintenance incorporates new changes automatically and continuously into the views, and is a good solution for very large warehouses. However, the performance of incremental maintenance algorithms is not well understood, and commercial incremental maintenance systems are still not widely available. In this paper, we study the performance of WHIPS, a prototype system developed at Stanford for studying warehouse maintenance. We examine the efficiency of change propagation and study factors that determine performance. We also compare the relative performance of different warehouse maintenance algorithms.

1 Introduction

We consider materialized view maintenance, where views are stored in a data warehouse and integrated from multiple, often distributed and autonomous data sources. One of the major challenges is to correctly and efficiently update the views as the source data changes. There are basically two types of techniques: refreshing and incremental maintenance. With refreshing, a view is recomputed, either on demand or periodically. With incremental maintenance, new information is gathered and integrated into the view without throwing away all existing data.

Most current warehousing systems use refreshing because of its simplicity. However, incremental maintenance offers at least two important advantages that may make it more attractive than refresh. First, a complete refresh may be infeasible because of its duration. In particular, a warehouse may be unavailable to users during refresh, so refresh is often done at night or on a weekend. However, for large warehouses, the refresh time may exceed the available "update window" of a few hours. Furthermore, global corporations may need 24×7 operation and cannot have any downtime. Incremental maintenance performs a relatively small number of operations on each update, so it may be done concurrently with warehouse user transactions. Second, some applications, such as one that tracks stock market changes to perform trading, may need updates to be reflected quickly. Only incremental maintenance can provide this data freshness.

Incremental maintenance has its disadvantages, however. For example, update algorithms are more complex, and may have to deal with concurrent read operations. Also, incremental maintenance may lead to inconsistent data in a view, if conventional algorithms are directly applied (Section 6). Algorithms that guarantee consistency have been developed, but they are more complex to understand and implement.

To answer questions about incremental maintenance, we built a warehouse prototype called WHIPS (WareHousing Information Project at Stanford). This prototype uses incremental algorithms to automatically and consistently maintain materialized views. As far as we know, WHIPS is the first such working experimental prototype. The primary objective of WHIPS is *not* high performance. Rather, the goal was to build a modular system where different techniques could be tested, and where the *relative* performance of components could be measured easily. In this paper we describe the prototype and present experimental results that provide insights into the incremental operation of a warehouse. In particular, the questions we address include:

- *Costs of incremental view maintenance.* What are the major components in the propagation time of an update to the warehouse? How much data is transferred from the sources and between different warehouse modules? What are the potential system bottlenecks? What system parameters have the greatest impact on performance?
- *Relative performance of different view management algorithms.* As we will see in Sections 6.2 and 6.3, views can be incrementally maintained using different levels of correctness and complexity. Which algorithm works best for a given environment and update scenario? How much more expensive is it to use a more complex yet "more correct" view maintenance algorithm, as opposed to a simpler but "less correct" one? Some algorithms only work if warehouse tables contain all source keys. Do these algorithms perform better? If so, can we quantify the benefits?
- *Batching and sequencing of updates.* Updates can be batched together to reduce overhead. Delaying updates (or batches) so they are not in close proximity can also reduce some overheads. How does batch size and update separation impact performance? What is the impact on response time?

We believe that our experiences and results can provide some useful guidance to future warehouse designers and implementers. This paper starts with some preliminaries: related research (Section 2), the WHIPS architecture (Section 3), and our performance criteria and relevant system parameters (Section 4). In Sections 5 through 8 we focus on important performance factors: modularity, keys, consistency, source update scenarios and batching.

Due to space limitations, in this paper we only present selected results and highlights. We give more extensive and detailed experimental results in a technical report [ZGM98]. There, we also present an analytic model that roughly predicts WHIPS performance. This model helped verify the correctness of the system implementation and served as a guide for our performance experiments.

2 Related Work

Overviews of data warehousing research and technologies can be found in [LW95, SC97]. Basic incremental view maintenance algorithms are described in [BLT86, GMS93, GM95, GL95]. Reference [Han87] discusses performance tradeoffs for materialized views in a conventional DBMS. The algorithms implemented in WHIPS are developed in [ZGMHW95, ZGMW96]; they handle distributed autonomous data sources and continuously interleaving updates, and provide view consistency. In the workshop paper [WGL+96] we presented a short preliminary description of the WHIPS architecture and applications. Another incremental maintenance prototype that handles Select-Project views is the Squirrel Project [HZ96a]. A related preliminary performance study comparing the performance of virtual and materialized approaches is given in [HZ96b].

3 WHIPS

In this section we describe the WHIPS architecture, and we illustrate its run-time functionality by following the propagation of one source update. WHIPS handles relational Select-Project-Join(SPJ) views, as well as derived views defined by union or aggregate operators. However, this performance study focuses on base SPJ views only.

3.1 Architecture

Figure 1 depicts the part of the WHIPS architecture that is relevant to our performance study. The key WHIPS feature is modularity. As shown in the figure, the system is composed of distinct modules that communicate with each other. Each module is a separate process. We implemented each module as a CORBA object, using the ILU implementation of CORBA [CJS+94].

We believe that this modular design achieves several important goals. (1) Heterogeneity: WHIPS is able to handle various types of data sources, as may exist in a real data warehousing environment. Any database can be used as the data warehouse. There are also a variety of view maintenance algorithms that can be used. (2) Distribution and scalability: As the load grows, maintenance work can be distributed among different processes, across multiple machines, and performed concurrently. (3) Implementation flexibility: Each integration module can be implemented and optimized independently. However, as we will see in Section 5, this design also bought us performance penalties.

In general, WHIPS is divided into three layers: the data sources, the integration modules, and the data warehouse. Each data source is encapsulated by a source-specific monitor and wrapper. A *monitor* is responsible for detecting modifications to the source data and notifying the integrator. We have implemented trigger-based monitors for cooperative data sources and snapshot monitors for other sources [LGM96]. A *source wrapper* translates single source queries from the internal WHIPS representation into queries in the native source language. Thus, the wrapper hides the

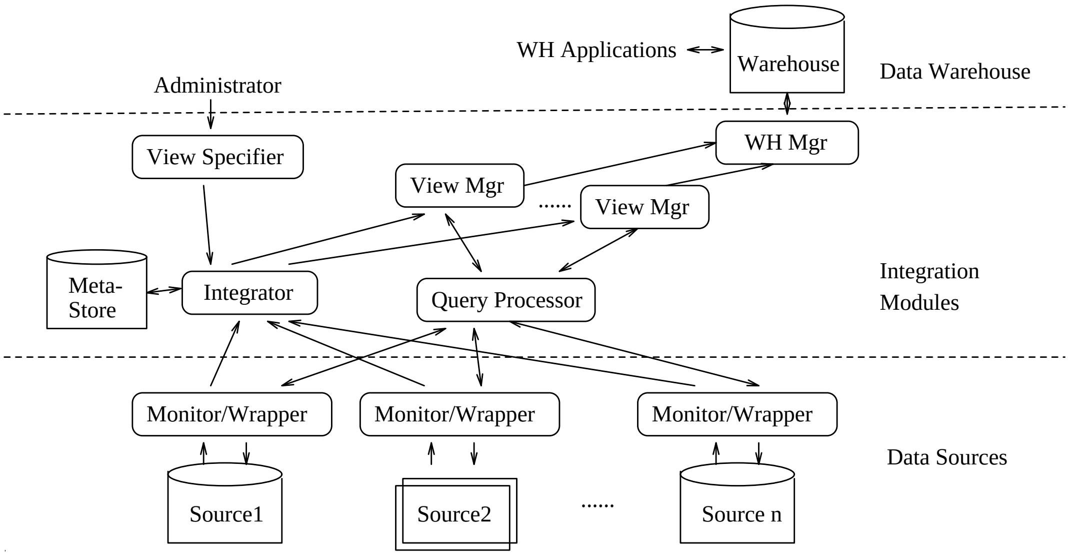

Figure 1: The WHIPS system architecture

source-specific querying details from all other modules. Data sources in WHIPS are autonomous; the only requirement is that they be capable (through their wrappers) of answering simple select queries. Currently, WHIPS wrappers handle three types of data sources: flat files, Sybase databases, and Mini SQL databases. New types of data sources can be added to the system as long as wrappers and monitors are written for them.

The *integrator* coordinates the different WHIPS components. At initialization, the integrator locates the warehouse manager and creates one or more query processors. When a new source becomes available, the integrator receives a notification from the source's monitor and wrapper. The integrator then stores the source information (such as the source type, schema, query interface) in the *meta-store*, a persistent store for system information. In the current prototype, the warehouse applications may define their views using a subset of SQL extended with constructs for specifying the desired view consistency (see Section 6). When a view is defined, the *view specifier* parses it and sends the corresponding view information to the integrator. The *integrator* then creates one *view manager* to maintain this view at the desired level of consistency. After initialization, the integrator coordinates view maintenance by routing source updates to the appropriate view managers.

Each *view manager* is responsible for maintaining a view, using one of the algorithms that we will discuss in Section 6. The view manager first initializes the corresponding view, then receives relevant updates and computes the corresponding view "deltas" (by interacting with the query processor if necessary), and finally sends those view deltas to the warehouse manager for execution. Views can be added and deleted at run time, and view maintenance for the different views can be performed concurrently, possibly using different maintenance algorithms.

A *query processor* receives global queries from the view managers and answers them by posing the appropriate single-source queries to the source wrappers. It then passes the composite global query answers back to the view managers. In WHIPS, a query processor performs distributed query processing using standard techniques such as sideways information passing and filtering of selection conditions to prune the queries it poses to the wrappers.

The *warehouse manager* receives deltas (lists of the specific inserts and deletes that must be performed to bring the materialized views up to date), and applies them to the views. All modifications received by the warehouse manager in a single message are applied to the warehouse in one transaction, as needed by the view consistency algorithms [ZGMW96]. Like the source wrapper, the warehouse manager also translates the view deltas to the specific syntax of the warehouse database, thus shielding all other WHIPS modules from the particulars of the warehouse and allowing any database to be used.

3.2 Propagating one source update

An incremental view maintenance algorithm computes and applies $\triangle V$, the changes to view V, when changes occur to the source data from which V is derived.

Example 1: Incremental view maintenance

Consider a company that handles magazine subscriptions and related issues. One department maintains subscription information in a relation $S = Subscription(Name, Magazine)$, and a second department maintains promotion information $P = Promotion(Magazine, Gift)$. A third department, in charge of shipping and handling, needs to find out what free gift to send to each subscriber. This information can be obtained by examining view $V = \Pi_{Name,Gift}(S \bowtie P)$. View V can be materialized in a data warehouse with the first two department databases as data sources. Figure 2 provides some sample data for S, P, and V. (The rest of the figure illustrates how WHIPS propagates updates; ignore for now.) For example, since Joe subscribes to Newsweek, he should receive the promotional CD gift.

For an update U_s on relation S, the corresponding view delta is $\triangle V = \Pi_{Name,Gift}(U_s \bowtie P)$. For an update U_p on relation P the corresponding delta is $\triangle V = \Pi_{Name,Gift}(S \bowtie U_p)$. Whether the view delta should be inserted into or deleted from the view is determined by whether the source update is an insertion or a deletion. □

The job of the WHIPS integration components is to compute and apply $\triangle V$. We now walk though a WHIPS update to illustrate the WHIPS components at work. In general, each propagation follows the message paths depicted in Figure 2. An asynchronous ILU procedure call is used to send each message. Note that in WHIPS, updates may also be propagated in batches (see Section 8) that follow the same route.

Consider view V of Example 1 and assume it is correctly initialized. Say the initial value of the base relations and the view is as shown in Figure 2. Let there be a source insertion $U_s = insert(S, [Tom, \underline{Time}])$, which indicates that Tom has subscribed to Time magazine. Update U_s is propagated as follows, where each point of time indicates the system starting or finishing an action.

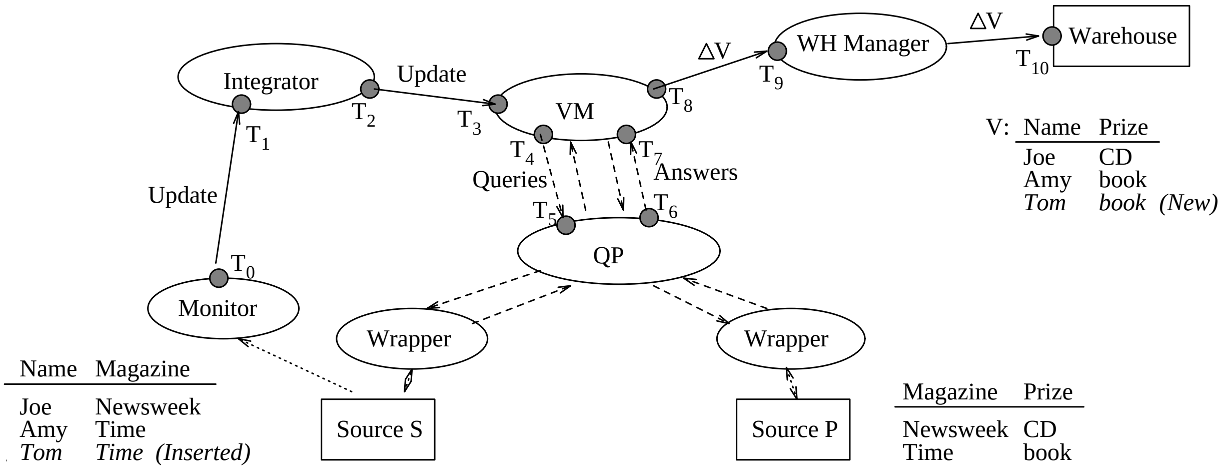

Figure 2: Propagating one source update

T_0: Update $U_s = insert(S, [Tom, \underline{Time}])$ is ready at the monitor of source S;

T_1: The integrator receives U_s;

T_2: The integrator forwards U_s to the view manager (VM) for view V;

T_3: This VM receives U_s;

T_4: The VM sends a query $Q_s = \Pi_{Name,Gift}(U_s \bowtie P)$ to the query processor (QP). We call this query an *update query* because it is triggered by a source update;

T_5: The QP receives query Q_s;

T_6: The QP finishes processing Q_s and sends the answer $A_s = [Tom, \mathit{book}]$ to the VM. Between T_5 and T_6, the QP sends a *wrapper query* to the source P wrapper asking for all gifts offered by Time magazine. The QP then composes answer A_s from the wrapper query answer;

T_7: The VM receives answer A_s;

T_8: The VM sends delta view $\triangle V$ to the warehouse manager. Note that $\triangle V$ contains the tuple $[Tom, \mathit{book}]$ obtained from A_s with an indication that it is to be inserted into the view;

T_9: The warehouse manager receives $\triangle V$;

T_{10}: $\triangle V$ appears in the view. That is, the transaction that applies $\triangle V$ to V commits, and the view contains the new tuple $[Tom, \mathit{book}]$.

Not all update propagations go through each of the above steps. For example, some view managers do not send queries to the QP when processing deletions. We let all view deltas, including empty ones, reach the warehouse manager, because this is needed to ensure multiple view consistency (which is not studied in this performance paper). Because of compensating queries (Section 6.3), there may be more than one pair of queries and answers between the VM and the QP (between T_4 and T_7). In this case, we assume T_4 is when the VM sends the first update query and T_7 is when the VM receives the last answer.

4 Performance Criteria and System Setup

In this section we first outline the major performance criteria, then describe the basic setup of our performance experiments. We also present some basic timing results.

4.1 Performance criteria

The main criteria we use is *update propagation time*(UPT). We define the UPT for a specific source update and view as the time elapsed from when the integrator receives the update (T_1 in Figure 2) to when the warehouse manager receives the corresponding view deltas (T_9). The time is measured using the Unix function `gettimeofday()`. Since in our experiments all the integration modules are located on the same machine, the UPT is just the time difference between T_9 and T_1. To refer to sub-components of the UPT, we use $t_{i,j}$ to represent the elapsed time between T_i and T_j in Figure 2. For example, $t_{5,6}$ is the time between the QP receiving a query Q and sending the corresponding answer, and $t_{6,7}$ is the time between the QP sending the answer and the VM receiving that answer. The full UPT of an update is $t_{1,9}$. A longer time to propagate an update to the warehouse indicates more stale views and an increased possibility of interference between consecutive update propagations.

We did not include in the UPT the time before the update reaches the integrator (before T_1) because $t_{0,1}$ is specific to the connection between the monitor and the integrator. Similarly, we did not include the time to apply $\triangle V$ to the warehouse (after T_9) because $t_{9,10}$ is specific to the database used as the warehouse. Neither of these times is determined by our integration techniques.

Another performance criterion is the number of messages passed between modules during the propagation of one or more updates. The most interesting numbers of messages are the number of queries sent by the view manager to the query processor and the number of source wrapper queries. These two numbers various for different views, for different types of view managers, and they affect UPT substantially. For some experiments in [ZGM98], we study the system throughput — the number of updates processed in one unit of time. In general, UPT and throughput are indications of the system efficiency, and the number and size of messages are indications of system load.

4.2 System

For all experiments presented in this paper, both the warehouse views and the source relations are stored in a Sybase database installed on a RS/6000 machine. The WHIPS integration modules run on a DEC Alpha 3000/300LX under OSF/1 3.0 with 125MH CPU, 64MB memory, and 128MB swap space. No other user processes were run on the system during the experiments. Source wrappers are on separate DEC Alpha machines with the same configuration. Any two machines are connected by 10BaseT Ethernet. As we have mentioned, most of the WHIPS system is implemented using C++ and ILU, a CORBA compliant object system developed by Xerox PARC. The entire system is about 19,000 lines of C++ code.

4.3 Views and base relations

View parameters affect the update propagation include the number of base relations, the sizes of base relations, selectivity (the fraction of tuples that satisfies the select condition(s)), project factor (the size of all projected attributes, which determines the tuple size of the view), and the join factor (the number of tuples in a relation that joins with one tuple in a neighboring relation). Given this diversity of factors, it is not easy to obtain a few "representative" scenarios. However, after some experimentation we chose the following "reasonable" views, representing views with different numbers of base relations and different result sizes.

$V_1 = \Pi_{A,B}(R), \quad V_2 = \Pi_{R.A,R.B,S.B,S.C}(R \bowtie S), \quad V_3 = \Pi_{R.A,R.B,S.B,S.C,T.C,T.D}(R \bowtie S \bowtie T),$
$V_4 = \Pi_{R.A,R.B,S.B,S.C,T.C,T.D,U.D,U.E}(R \bowtie S \bowtie T \bowtie U)$

In the view definitions, R, S, T, U are base relations. In the default setup, the cardinality of the base relations is set to 100 tuples. Each tuple in a relation is 100 bytes, and each projection attribute is 10 bytes. By default, we populate the tables such that one tuple of R joins with exactly one tuple of S, then joins with exactly one tuple of T. This is also true for new tuples; that is, if a new tuple is inserted into R, and this tuple is to be joined with S to compute $\triangle V$, we set the attribute value such that this new tuple will join with one tuple in S. This fixed join factor reduces variability and makes our results easier to interpret.

We use small relations mainly because there is an ILU upper bound on message size, which limits the sizes of the initial views, as well as of the query answers. To get around these limits we could have ported WHIPS to another CORBA implementation (e.g., Orbix has higher limits on message sizes). We could have also implemented a facility to break up large messages into "packets" to be sent in pipelined fashion. We did not consider this effort worthwhile for our evaluation because most of the performance results of interest are *not* affected by the base relation size. In particular, since we are studying the performance for *incremental* maintenance, base relation cardinality plays little role in computing $\triangle V$; the more important factors are the join factors between relations. Also, we are focusing on the performance of WHIPS integration components. Larger base relations may increase the time to evaluate a source wrapper query if the relation is not properly indexed. However, this is just one feature of the data source that the integration modules may take into consideration. Finally, we often want to compare the relative cost of different algorithms or design choices. For example, in Section 6.3 we are interested in comparing "complete consistency" strategies with "no consistency" ones. Absolute timings are not very relevant to this comparison.

4.4 Update scenarios

In our performance experiments we use an *update simulator* to generate source updates according to specified parameters. The simulator updates the source data and sends the updates to the integrator. Thus, the simulator behaves like a WHIPS monitor. We consider two types of update propagation scenarios: non-interleaving (default scenario), in which the updates are separated by sufficiently long intervals such that the propagation of consecutive updates do not interfere with each other, and interleaving, in which at least the propagation of one update U_k interferes with others. This interference may cause resource conflicts, (as well as additional compensation work, as described in Section 6). Thus, the UPT of U_k may be longer than in the non-interleaving scenario.

4.5 Confidence of timing results and some basic timing results

To measure propagation times we use the average time $\bar{T} = (t_1 + t_2 + \ldots + t_P)/P$ from P repetitive experiments. In other words, each data point in the graphs that follow is an average number obtained from P repetitive experiments. For all the experiments in this paper, we have chosen a large enough P such that with a degree of confidence 0.99 we guarantee that the expected value of T is within $\bar{T} \pm 0.1\bar{T}$. Details on our choice of P and the evaluation of confidence intervals is given in [ZGM98].

There are two basic timing results that are determined by the specific system configuration and are not affected by the view maintenance algorithms. The first one is the time to pass an asynchronous message between two modules. With the described configuration, a message containing 100 bytes of data takes about 0.004 seconds if the two modules are located on the same machine, and about 0.005 seconds if they are located on different machines. The second factor is the time to evaluate a source query, including the time to communicate with the source wrapper, execute the query, and get the results. When the result size is one tuple with 100 bytes of data, the time to execute a selection wrapper query with one selection condition takes about 0.028 seconds. This number remains roughly constant as the base relation ranges from 1 to 1000 tuples. We study the effect of message size on message time further in [ZGM98]

5 Cost of Modularity

WHIPS was designed as a modular system to make experimentation easier. Even though modularity is a requirement, it is important to study the cost of this modularity, to help us interpret the results that follow, and to help us understand the relative costs of the various update propagation stages.

Recall that we defined the UPT of an update as $t_{1,9} = \sum_{j=1}^{j=8} t_{j,j+1}$. Among those times, $t_{2,3}$, $t_{4,5}$, $t_{6,7}$ and $t_{8,9}$ are communication times between WHIPS integration modules. These times are mainly determined by the size of the message. For example, $t_{2,3}$ is the time that the integrator takes to forward a source update to the view manager. It is roughly the same for all single-tuple updates, and it increases with the batch size if source updates are combined. Similarly, $t_{6,7}$ is the time the query processor takes to send a query answer to the view manager, and it is mainly determined by the answer size. Let $t_{commu} = t_{2,3} + t_{4,5} + t_{6,7} + t_{8,9}$, the communication overhead entailed by WHIPS's modular design.

Let $t_{computation} = t_{1,2} + t_{3,4} + t_{7,8}$. During these periods the WHIPS modules process data, e.g., decide the type of an update or check the correctness of a query answer. Let $t_{query} = t_{5,6}$, the time to process a distributed query. The value t_{query} may involve both communication (with source wrappers) and computation (collecting results). The length of t_{query} is determined by the number of wrapper queries, the network speed, and the time it takes sources to execute queries.

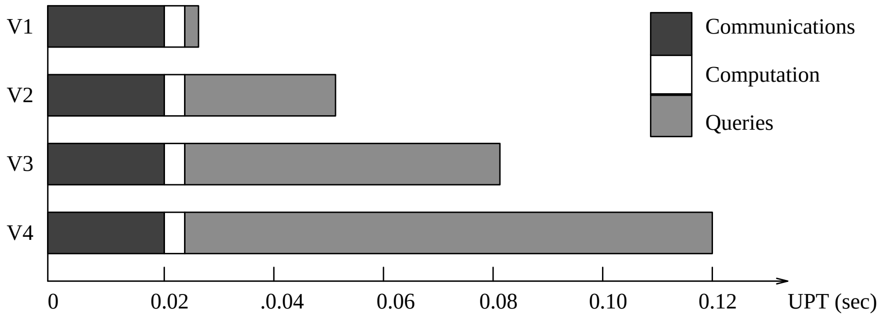

Figure 3: Update propagation time breakdown for different views

In Figure 3 we show the UPT breakdown for propagating a source insertion in the default setup. The results we present here are obtained using one type of view manager but results for other view managers are almost identical (see Section 6.3). In the non-interleaving scenario, the UPTs for the four views given in Section 4.3 are 0.025 seconds, 0.053 seconds, 0.083 seconds and 0.120 seconds. As expected, for different views, the t_{commu} times are almost the same, because messages with the same type and almost the same size are exchanged. The $t_{computation}$ times are also similar for all views, and they are shorter than the communication and query times. For V_1, which contains only one relation, no wrapper query is necessary to propagate source updates. Therefore, communication time dominates, and t_{commu} is about 85% of the total UPT. As the number of base relations in the view increases, the query time quickly increases and becomes the dominant factor. For V_4, t_{commu} is about 17% of the UPT and t_{query} is about 80%. While the absolute communication time remains essentially constant, its fraction of the entire process decreases.

From Figures 3, we see that query times can be very significant for some views, and the communication times are a significant overhead. Let us discuss these in turn.

Query times are high when the sources have to be contacted as part of the maintenance process. These times can be reduced by having faster sources, or by improving the query processor (e.g., through smart join ordering or condition filtering). Query times can be reduced even more by copying all source relations to the warehouse, avoiding going back to the sources entirely. (WHIPS can maintain copies of source data by defining views that are simply the same as source data.) In this case, queries are still performed, but entirely with warehouse data. The benefits of copying source data must be balanced with the costs: increased storage space and a larger number of updates that must be processed. The latter is because the copied source data must be maintained at the warehouse.

The communication costs can be reduced by using a more efficient CORBA implementation, or by coding the message passing directly. Another option is to combine all the integration modules, including the warehouse manager, into the same process, i.e., put all the integration modules close to the warehouse. With this approach, different integration modules may share a memory structure to exchange information. We believe that using separate modules is better for a testbed such as WHIPS, where ease of implementation is important, and that a combined approach would be more effective for a high performance system. This choice between tightly and lose integration of components is one faced by many system designers.

An option for reducing communication overhead while still maintaining modularity is to give each view manager its own query processor. Thus, each view manager and its instance of the query processor would share a process, and data exchanges between them would be efficient. This approach is especially beneficial when there are many queries from the view manager to the query processor, e.g., when updates frequently interleave and the view manager sends many compensating queries (Section 6.3, 7). The main disadvantage is that it limits our ability to do load balancing for query processing.

In a separate study we varied the size of the answer relations that the query processor composed ([ZGM98]) and we obtained similar results as those shown in Figure 3. In summary, if the efficiency of update propagation is critical, then a modular design based on CORBA is probably not be the best choice. If it is feasible to copy all source data to the warehouse, this can significantly reduce query times and should be considered. Of course, the types of views being maintained play an important role in these decisions, and the results shown in this section may help quantify the costs and benefits.

6 View Maintenance Algorithms

WHIPS implements six view maintenance algorithms, divided into two groups, as shown in Table 1. The detailed algorithms were presented in [ZGMHW95, ZGMW96]. *Key algorithms* assume that the projection list in the view definition contains all key attributes from the base relations. *Duplicate algorithms* can maintain all Select-Project-Join views, with or without key information. WHIPS has available three view maintenance algorithms for each group, with different consistency levels.

	Key Algorithms	Duplicate Algorithms
Inconsistent	KeyConventional	Conventional
Strongly Consistent	Strobe	ECA
Complete	CStrobe	CECA

Table 1: Consistency levels of view maintenance algorithms

In Section 6.1 we illustrate view consistency levels through examples, and show how key algorithms and duplicate algorithms achieve consistency. In Section 6.2 we compare the performance of the two groups of algorithms. Then in Section 6.3 we compare algorithms that guarantee different levels of view consistency. Finally we discuss the coding complexity of all algorithms implemented.

6.1 View maintenance algorithms and view consistency

Recall that sometimes more information from the sources is needed to compute $\triangle V$, and WHIPS sends source wrapper queries to obtain that information. If there are concurrent updates at data sources while these wrapper queries are evaluated, conventional view maintenance algorithms such as those described in [BLT86, GMS93] may yield incorrect results. The following is an example:

Example 2: Inconsistent view maintenance

Consider the view $V = \Pi_{Name,Gift}(S \bowtie P)$ defined in Example 1 and the two initial relations depicted in Figure 2. Initially, the view contains two tuples $[Joe, CD]$ and $[Amy, book]$. This view content is consistent with the base data. Now, say Joe discontinues his subscription to Newsweek, at the same time that Newsweek cancels its promotional CD for its subscribers. There are two source updates: $U_s = delete(S, [Joe, \underline{Newsweek}])$ and $U_p = delete(P, [\underline{Newsweek}, CD])$. Corresponding to U_s, query $Q_s = \Pi_{Name,Gift}([Joe, \underline{Newsweek}] \bowtie P)$ is evaluated at the source of P. Now that the tuple containing Newsweek is removed from P, the result of Q_s is empty. The query corresponding to U_p, $Q_p = \Pi_{Name,Gift}(S \bowtie U_p)$ also returns an empty answer because currently no one is subscribing to Newsweek. The view will not be changed if one simply applies both results as suggested by conventional view maintenance algorithms. The final view (containing tuple $[Joe, CD]$) is inconsistent with the final source data. $\square$

We have designed algorithms to avoid this type of inconsistency. As mentioned, some algorithms take advantage of base relation keys. For Example 2, if $Name$ is a key of $Subscription$, (i.e., each person subscribes to only one magazine), then for $U_s = delete(S, [Joe, \underline{Newsweek}])$, the view maintenance algorithm just needs to delete all tuples in the view with attribute $V.Name = Joe$. If $Gift$ is a key of $Promotion$ (i.e., each gift is offered by only one magazine), then for U_p, one may delete all tuples in the view with attribute $V.Gift = CD$. For views with duplicate semantics, *compensating queries* must be sent to correct the missing or redundant query results caused by update interleaving. In the above example, when U_p is processed, the maintenance algorithm may realize that update U_p will cause the result of Q_s to be incorrect. Then the query for U_p can be set to $Q_p = \Pi_{Name,Gift}(S \bowtie U_p) \cup \Pi_{Name,Gift}(U_s \bowtie U_p)$ ($\Pi_{Name,Gift}(U_s \bowtie U_p)$ is a compensating query), such that the result of Q_p will contain tuple $[Joe, CD]$. The final view is correct after this tuple is removed from the view.

A materialized view should reflect base data as it changes. In [ZGMHW95, ZGMW96] we formally defined different levels of materialized view consistency. Here we use a simple example to illustrate these levels informally. Let there be three source updates U_1, U_2, U_3, which occurred in that order at a data source. Let V be a related materialized view. A *complete* view maintenance algorithm will maintain V such that V reflects each of the U_is in the order they occurred. For the three updates in the example, the complete algorithm will update the view three times, each reflecting one update. A *strongly consistent* algorithm, on the other hand, may skip some source states but still guarantees eventual correctness. For example, it may update V to reflect U_1 and U_2 together, then reflect U_3. An *inconsistent* algorithm may maintain V such that its intermediate or final value is different from anything that could have been computed from source data. For example, it may finish all maintenance work and leave V in a state inconsistent with base data after U_3. As we saw in Example 1, a conventional (centralized) view maintenance algorithm is inconsistent if directly applied to the warehousing environment. We implemented the conventional algorithms to compare their performance with that of other algorithms.

6.2 Keeping keys in a view

Now we compare the performance of key and duplicate algorithms. We specifically compare key algorithm Strobe with duplicate algorithm ECA. In the non-interleaving scenario, the other two key algorithms behave the same as Strobe, and the other two duplicate algorithms behave the same as ECA (Section 6.3). Figure 4 shows the UPT of an update on base relation R.

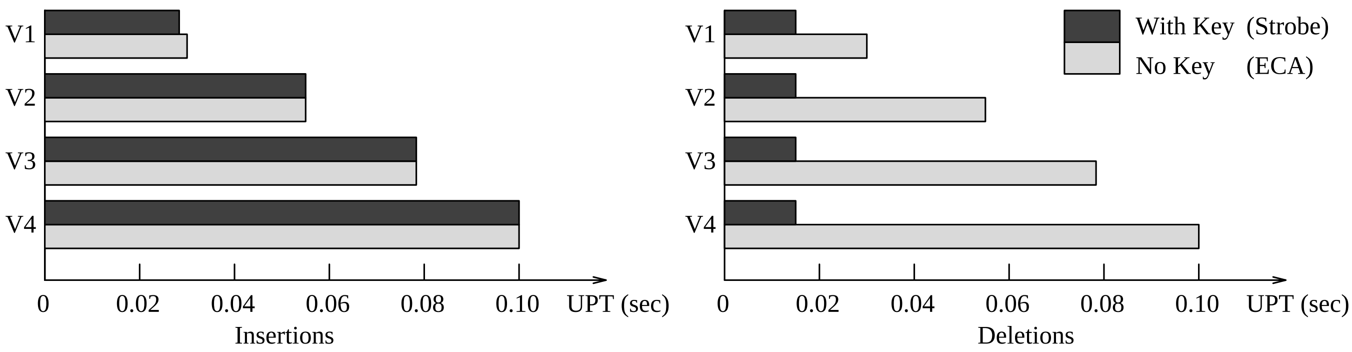

Figure 4: Update propagation time for algorithms with and without key information

As we may predict, for the same view, the UPTs for insertions are similar for both algorithms because they both need to send source queries and obtain answers. The UPT using Strobe is sometimes slightly shorter because Strobe uses a simpler data structure. For deletions, the UPTs are quite different. Duplicate algorithm ECA treats deletions exactly the same as insertions. Strobe, on the other hand, always forwards the deleted key attributes directly to the warehouse and lets the warehouse manager apply a delete action. Therefore for Strobe, the UPTs of all deletions are the same, regardless of the view definition, and they are shorter than those for insertions. (In the rest of this paper, we focus mainly on duplicate view managers under insertions. This is because the UPT for using other view managers can be inferred from these results.)

The benefit of using a key algorithm is the efficient processing of deletions. The higher percentage of deletions in the source update streams, or the longer it takes to propagate insertions (e.g., when the view contains many base relations), the more beneficial it is to use a key algorithm.

When source updates include deletions, maintaining a view using key algorithms also incurs fewer messages between the VM and the QP, and thus generates fewer wrapper queries and less work at the source. To propagate a total of K updates including K_1 insertions, there are K update queries sent from the VM to the QP for duplicate algorithms and K_1 for key algorithms.

We believe that for many applications, requiring base keys in view definitions is a reasonable assumption, since keys make it easier for the applications to interpret and handle the warehouse data. Furthermore, if a user-specified view does not contain sufficient key information, the warehouse can sometimes add the key attributes to the view and project them out when presenting the view content to the user. Nevertheless, it is not always appropriate to keep all base key attributes in a view, e.g., when base data use duplicate semantics, when the warehouse has no knowledge of base keys, or when the warehouse does not want to change user-defined views. In all these cases duplicate tuples or tuple counts need to be preserved in the views in order to incrementally maintain the view corresponding to source deletions [BLT86, GMS93, ZGMHW95].

6.3 View consistency

Now we study the cost and complexity of achieving different levels of view consistency. We performed experiments under both the non-interleaving scenario and a predictable interleaving scenario. The average UPT results are shown in Figure 5.

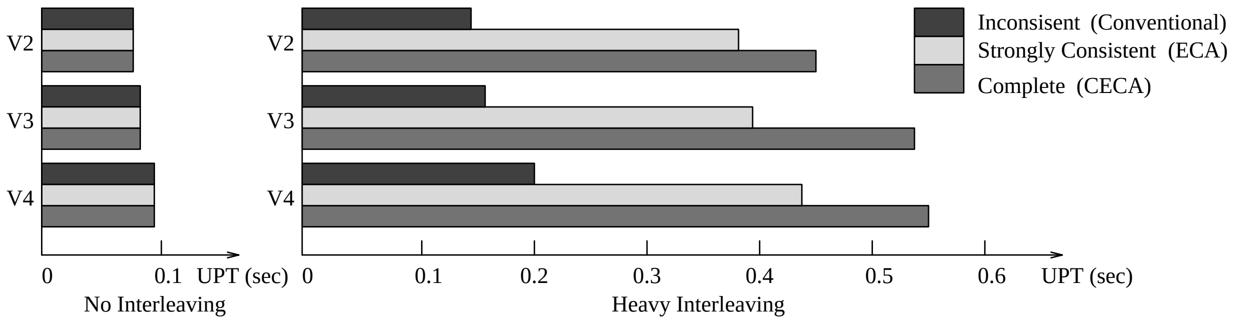

Figure 5: Update propagation time for views maintained to different consistency levels

When there is no source update interleaving, all three types of algorithms maintain materialized views correctly. The average UPTs for the three algorithms are almost the same. This is as expected since all view managers should send and receive the same messages with the same contents. Between T_7 and T_8, the strong and complete view managers execute more CPU operations due to their more complex logic, but the effect is insignificant. In this scenario, all three algorithms also reflect each single source update in the view.

In the interleaving scenario, two tuples are simultaneously inserted, one into base relation R and the other into S, and reported to the integrator one after the other. Let these two updates be U_r and U_s, U_r joins with U_s. Update U_s occurs before the propagation of U_r finishes, i.e., before the update query corresponding to U_r is evaluated. Thus U_s interferes with the propagation of U_r. We measure the average UPT of this group of two updates and repeat the experiment enough times to satisfy the specified confidence level. We stress that this interleaving scenario is a pessimistic one because every single update observes interference. In Section 7 we study a scenario with stochastic interference.

For the interleaving scenario, the Conventional algorithm may lead to incorrect view maintenance. The UPT for the Conventional algorithm is a lower bound since that algorithm logically processes the updates as if there were no conflicts among them. The UPT increase for the Conventional algorithm between the no-interleaving and interleaving scenarios is due to resource contention (see Section 7). The strong and complete algorithms will send compensating queries to correct potential inconsistencies. Specifically, a compensating query of the form $U_r \bowtie U_s \bowtie T$ is sent to the QP for the two updates U_r and U_s. Since evaluating compensating queries requires communication between the VM and QP, and possibly between the QP and source wrappers, the impact of compensating queries on UPT is substantial. For example, for V_3, the UPT using the ECA algorithm is about 2.6 times longer than the Conventional UPT, and that using CECA algorithm is about 3.0 times longer than the Conventional UPT. Complete algorithms take the longest time to propagate updates because updates must be processed one by one, while strongly consistent algorithms may combine updates. The increase of the average UPT for complete and strongly consistent view managers is the cost of maintaining data consistency.

Now we look at the number of messages exchanged. Let the total number of updates be K, where K is an even number. Let the number of messages exchanged by modules m_1 and m_2 be $msg(m_1, m_2)$. For the non-interleaving scenario, we have $msg(VM, QP) = K$, $msg(QP, WR) = 2K$ (WR represents all source wrappers) and $msg(VM, WH) = K$ for all three algorithms. For the interleaving scenario, every pair of updates triggers two update queries and one compensating query, thus $msg(VM, QP) = 3K/2$. Each update query invokes two source wrapper queries, and each compensating query invokes one wrapper query, so the total number of wrapper queries is $msg(QP, WR) = 2 * K + 1 * K/2 = 5K/2$. For the ECA view manager, every two updates are bundled together and invoke only one view-update message, so $msg(VM, WH) = K/2$. For the CECA view manager, each update invokes one view update message, so $msg(VM, WH) = K$. The number of messages is summarized in the following table.

	$msg(VM, QP)$	$msg(QP, WR)$	$msg(VM, WH)$
All three algorithms (non-interleaving) and Conventional (interleaving)	K	2K	K
ECA (interleaving)	3K/2	5K/2	K/2
CECA (interleaving)	3K/2	5K/2	K

As we can see, the increase in number of messages is not dramatic. The total message overhead goes from $4K$ in the non-interleaved case, to at most $5K$ with interleaving. Thus, the main costs of guaranteeing consistency when updates interleave are delays and the overhead of additional source queries.

6.4 Implementation complexity

The following table shows the number of lines of code for the view manager algorithms, excluding comments and common code (e.g., initialization). As we can see, key algorithms are slightly less complex to implement; for example, the implementation of Strobe view manager is about 70 lines of code shorter than the ECA view manager. Within the same group, the two complete view managers are the most complex, as expected, followed by the strongly consistent view managers, followed by the Conventional view managers.

View Managers	Lines of Code	View Managers	Lines of Code
vmCStrobe.cc	468	vmCECA.cc	479
vmStrobe.cc	311	vmECA.cc	381
vmKeyConv.cc	201	vmConv.cc	190

7 Source Update Scenarios

Interleaving of update propagations impacts the system in two ways:

1. Resource contention. A message for one update may be blocked if the module it tries to reach is busy processing another update.
2. Generation of extra work. Some of the algorithms, such as ECA or CECA, compensate for interleaving by sending compensating queries and processing the results of those queries. This creates additional work.

In Section 6.3 we focused on the delays caused by the extra work generated by compensating algorithms. In this section we present one experiment results where both contention and extra work affect system performance. In this experiment, we perform a series of tests with randomly-generated source updates. These tests provide a more realistic scenario and a better idea of how WHIPS may operate in practice. In [ZGM98], we present more detailed study on the effect of resource contention on UPT and the system throughput.

We concurrently generate 40 random insertions to base relation R and 40 random insertions into relation S. Each new tuple has about a 1/5 chance to join with an existing tuple in a neighboring relation. The update interval is a random number between 0 and I seconds. Figure 6 shows the performance results for view V_3, using CECA algorithm. Each data point represents one run where the 80 updates are processed.

Figure 6(a) shows the average UPT. As expected, when the update frequency is the highest (update interval I is 0), we have the longest average UPT due to update interference. When I is small, the UPT is quite variable, since the precise ways in which the updates interfere is very unpredictable. The average UPT decreases as the update interval grows, and once the time interval is greater than a certain threshold, the average UPT stays almost constant. Let this time interval be I_b. Intuitively, there is little interference between the propagations of consecutive updates if they are separated by a time interval larger than I_b.

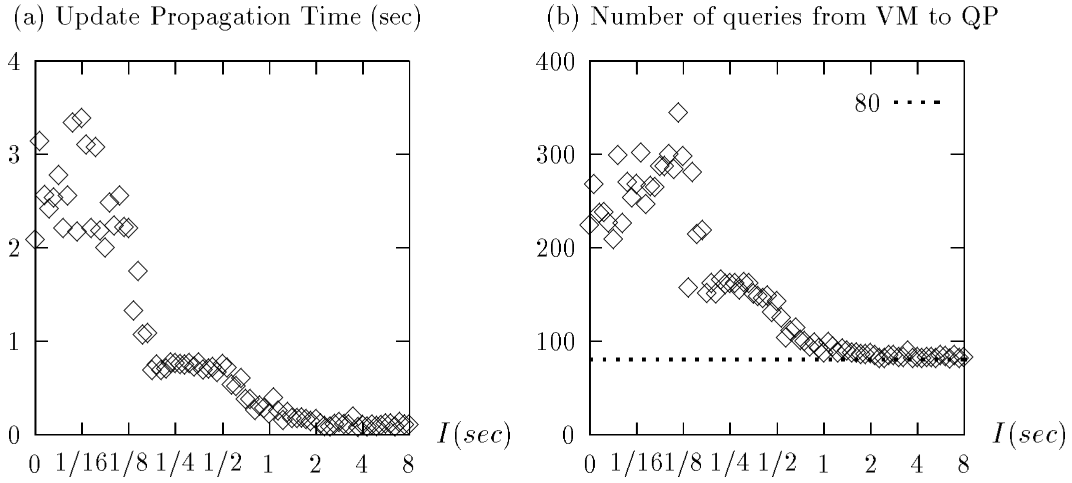

Figure 6: Performance results for propagations with random update intervals

In the current experiment, updates are delayed both because of resource contention and because of compensating queries. To illustrate this, Figure 6(b) shows the number of queries sent from the VM to the QP. Of all the queries, 80 are update queries, and the rest are compensating queries. We can see that the query number increases almost directly with the increase in average UPT. There are many compensating queries when the update frequency is high; the actual number depends on how updates interfere with each other.

We also noticed that there is no significant increase in the number of actual wrapper queries (See [ZGM98] for results). Because the chance of two updates U_r and U_s joining is small (about 1/200), most of the compensating queries of the form $U_r \bowtie U_s \bowtie T$ result in an empty answer and can be evaluated by the QP locally. Thus, we see a good optimization opportunity for the heavy interleaving scenario: the view manager may screen compensating queries before sending them to the QP. If a query will clearly yield an empty answer (as is the case with $U_r \bowtie U_s \bowtie T$ when U_r and U_s do not join), the view manager can skip it and reduce the compensation load.

8 Batching

Source updates can be batched and propagated to the warehouse together. In general, batching updates not only reduces the number of messages, but it also reduces the chance of interference between updates. However, batching may cause updates to wait before being sent, increasing their UPT. Batching can be done at various steps along the update propagation path, i.e., at the source monitor, the integrator, or the view manager. We assume that source monitors collect updates and send them together to the integrator, according to one of the following batching policies.

1. Batch every B updates, where the batch size $B \geq 1$. The $B = 1$ scenario has been discussed in previous sections, where each update is propagated individually. When $B > 1$, some updates may wait at the source monitors to be batched with later updates, and this waiting times would be considered part of the propagating time. As we will find out later in this section,

for certain update scenarios, there is an optimal batch size.

2. Batch all updates generated in certain time interval, e.g., in a minute or in an hour. This policy is useful if it is only possible to propagate updates during certain time windows, so as to not interfere with other activities (e.g., user queries). At the beginning of an update window, it is possible to break up a batch into smaller ones.
3. Batch when the "oldest" update has waited for a certain time. This method may ensure that a view does not become too "stale." It can also be used in combination with Policy 1 to make sure all updates are propagated to the warehouse after waiting for a certain time.

In this section, we will discuss only Policy 1 involving the optimal batch size. Once we have determined the optimal batch size for Case 1, its value may be used in Policies 2 and 3. For instance, in Policy 3 we may wait to collect the optimal number of updates, unless the first update is delayed too long.

There are two major goals for batching. The first is to reduce the workload. Batching reduces the number of messages, so it is most useful when messages are expensive. Batching may also reduce the number of operations done by the integration module. For example, to compute $\triangle V = \triangle R \bowtie S$ where there are n updates in $\triangle R$, batching the n tuples will enable the S source to perform one join instead of n joins. In general, the larger the batch size, the more the workload is reduced. Of course, considering view freshness requirements and message size limitations, there is a practical limit on the largest reasonable B.

The second major goal of batching is to speed update propagation by reducing the interference among update propagations. The following experiment shows how batching reduces the average UPT compared with individual updates. It also examines the optimal batch size under certain update scenarios. We study view $V_2 = R \bowtie S$ using algorithm CECA. A total of K updates on base relation R are grouped into batches of size B, and sent to the integrator with a fixed interval I between batches. Figure 7 shows the average UPT results for $K = 20$ and $B = 1, 2, 5, 10$. As an example, when $B = 2$, there are 10 batches of updates, each containing 2 tuples. These 10 batches are sent to the integrator separated by interval I.

Figure 7 shows that each batch size is optimal for some update frequency. In general, when the update frequency is high (I is small), updates bundled in larger batches are propagated faster. When the update frequency is low, smaller batches are better. When I is large enough such that the propagation of updates do not interleave, sending updates individually (B=1) is the fastest way to propagate them.

The average UPT shown in Figure 7 can be broken into two components: the actual update propagation time and the wait time. (The latter is the average time updates wait at the source before being batched and sent to the integrator.) For the actual update propagation time only, a larger batch size is always better, since there are fewer messages and less contention. For the wait time, the source waits until B updates are available before sending them together, giving the

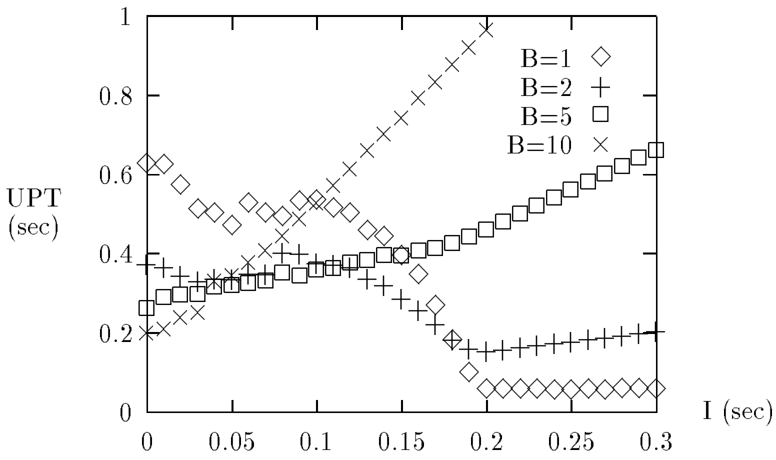

Figure 7: Average UPT for different batch sizes

average wait time for batch size B as $(B-1) * I/2$. The wait time increases linearly with the update interval as well as with B.

When we consider both the actual propagation time and wait time, the former dominates when I is short, and a larger batch size gives shorter UPT. One special case is when $I = 0$ (all updates are ready to be propagated). This is the case when updates are generated by the same source transaction or updates accumulate during non-propagation times. Batching them in the largest possible batch size is often the best. When I is long, wait time becomes the dominant factor and smaller batches may perform better. In reality, optimal batch size or batch window may be determined according to the expected frequency of updates.

9 Conclusion

In this paper we studied the performance of WHIPS, an incremental warehouse maintenance testbed. Performance depends on a variety of factors, including hardware parameters, the base relations and views defined, the update frequency and batch size, and the maintenance algorithm used. It is thus hard to make particular performance predictions. Nevertheless, we believe that our experiments shed some light on the major design tradeoffs faced in incremental maintenance. In particular,

- We evaluated the costs of a modular system design, by studying update propagation and quantifying the cost of using CORBA as an infrastructure (Section 5). We concluded that CORBA is quite expensive, but offers great design and implementation flexibility.
- We evaluated the delay caused by sending queries to data sources (Section 5). The results can help us decide whether to copy base data at the warehouse, and how beneficial it may be to optimize global query evaluation.
- We compared the relative performance of different view maintenance algorithms and studied the cost of maintaining view consistency (Section 6).

- We analyzed the behavior of WHIPS under a spectrum of source update scenarios, quantifying the impact of update interleaving (Section 7).
- We presented results that provide insights to system designers. For example, we discussed how to determine optimal batch size (Section 8), how to send updates (Section 7, 8), which query processor optimizations are crucial (Section 5, 6, 7), and what is a superior method for evaluating compensating queries (Section 6, 7).

References

[BLT86] J.A. Blakeley, P.-A. Larson, and F.W. Tompa. Efficiently updating materialized views. In *SIGMOD*, pages 61–71, June 1986.

[CJS+94] A. Courtney, W. Janssen, D. Severson, M. Spreitzer, and F. Wymore. Inter-language unification, release 1.5. Technical Report ISTL-CSA-94-01-01 (Xerox accession number P94-00058, Xerox PARC, May 1994.

[GL95] T. Griffin and L. Libkin. Incremental maintenance of views with duplicates. In *SIGMOD*, pages 328–339, May 1995.

[GM95] A. Gupta and I.S. Mumick. Maintenance of materialized views: Problems, techniques, and applications. *IEEE Data Engineering Bulletin, Special Issue on Materialized Views and Data Warehousing*, 18(2):3–18, June 1995.

[GMS93] A. Gupta, I. Mumick, and V. Subrahmanian. Maintaining views incrementally. In *SIGMOD*, pages 157–166, May 1993.

[Han87] E.N. Hanson. A performance analysis of view materialization strategies. In *SIGMOD*, pages 440–453, 1987.

[HZ96a] R. Hull and G. Zhou. A framework for supporting data integration using the materialized and virtual approaches. In *SIGMOD*, pages 481–492, June 1996.

[HZ96b] R. Hull and G. Zhou. Towards the study of performance trade-offs between materialized and virtual integrated views. In *Workshop on Materialized Views*, pages 91–102, June 1996.

[LGM96] W. Labio and H. Garcia-Molina. Efficient snapshot differential algorithms in data warehousing. In *VLDB*, pages 63–74, September 1996.

[LW95] D. Lomet and J. Widom, editors. *Special Issue on Materialized Views and Data Warehousing,* IEEE Data Engineering Bulletin 18(2), June 1995.

[SC97] U. Dayal S. Chaudhuri. An overview of data warehousing and olap technology. *SIGMOD Record*, 26(1):65–74, March 1997.

[WGL+96] J.L. Wiener, H. Gupta, W.J. Labio, Y. Zhuge, H. Garcia-Molina, and J. Widom. A system prototype for warehouse view maintenance. In *Workshop on Materialized Views*, pages 26–33, June 1996.

[ZGM98] Y. Zhuge and H. Garcia-Molina. Performance analysis of WHIPS. Technical report, Stanford University, May 1998. http://www-db.stanford.edu/pub/papers/zg-whipsperf.ps.

[ZGMHW95] Y. Zhuge, H. Garcia-Molina, J. Hammer, and J. Widom. View maintenance in a warehousing environment. In *SIGMOD*, pages 316–327, May 1995.

[ZGMW96] Y. Zhuge, H. Garcia-Molina, and J.L. Wiener. The Strobe algorithms for multi-source warehouse consistency. In *PDIS*, pages 146–157, December 1996.