Go To CS848 Site

CS848 Paper Submission Site

Submit or Update A Review For Paper #9

Go To CS848 Submissions

It is currently Sunday 07th of November 2004 08:39:03 PM EST

Your review has been finalized!

Paper # 9 **(Download paper of type application/pdf, 1877532 bytes)**	
Title:	Join processing in database systems with large main memories
Abstract:	

You have already finalized your review for this paper.
You can no longer modify it, but you may view it.

If you made a mistake in your review and you want it "unfinalized",
you may send mail to the program chair asking them to unfinalize it

Send yourself this review by email

Attribute	Value
Are you finished with this review?	Finalize, I am done editing
Provide a short summary of the paper	This paper presents four algorithms for computing the join of two relations assuming that the smaller relation will fit almost entirely into main memory. The author begins with a description of the problem as well as some related work. The notation used throughout the paper is presented to the reader. The author presents the sort-merge join algorithm that makes use of root(card(S)) blocks of memory to perform the sort and merge operations (where S is the smaller relation). The cost of performing this join is also presented. The three hash join algorithms considered by the author (Simple, GRACE and hybrid) are all presented, as well as cost formulas for their evaluation. The author next presents a comparison of the algorithms assuming that no buffer overflows take place. The comparison is made using the cost formulas derived in the previous discussions. Some solutions to dealing with bucket overflow, such as repartitioning and reassigning buckets, are presented. The author next presents a method of performing memory management if there is not enough free memory to provide each join the previously defined minimum amount of memory. In this case, the easiest and most efficient means to provide memory is using hot pages and virtual memory. Hot pages are the number of pages guaranteed to be available to the join at any given time. A performance evaluation using the models developed earlier is shown. The author concludes with a discussion of tools such as Babb arrays and database filters that provide a more efficient join algorithm, as well as some concluding and summary remarks.

What is the strength of the paper? (1-3 sentences)	This paper presents a hybrid algorithm for performing hash joins. It also characterizes several methods of join processing using cost formulas.
What is the weakness of the paper? (1-3 sentences)	This paper makes some unreasonable assumptions early on. As well, performance modeling is performed using some unverified cost formulas instead of empirical data.
Your qualifications to review this paper	I know the material, but am not an expert
Writing Quality	Excellent
Relevance to query processing?	The paper is relevant to query processing
Experimental Methodology	**Poor**
Novelty of paper	This is a new contribution to an established area
Overall paper merit	**The paper is a novel or new contribution with average/weak methodology, or an incremental contribution that has good methodology. Someone in the area should read it**
In your opinion, will this paper be important over time?	Average
Provide additional detailed comments to the author	You have presented a novel join method. However, there are several issues that I take with your presentation: -A deeper discussion of the cost formulas and how they were derived would have been useful. In some cases, short single line descriptions are provided, but are not enough. -In your GRACE algorithm, assuming that each set Ri is of equal size assumes that the data is uniformly distributed. Is this really a valid assumption to make? -Your definitions of "large" in section 3 appear to be very inappropriate. A table that is large may contain several hundred thousand rows, not one thousand! -Is it really appropriate to compare your algorithms using the cost models you have defined? What if your models are incorrect? Would some empirical evidence not have been more appropriate to compare your algorithms and validate your models? Overall, you have provided a very well written and insightful paper. However, I believe that you should fix the above assumptions and errors and resubmit your paper at a later date.
Additional comments to PC (not seen by author)	**This paper presents a novel join processing method. However, there are major flaws with the performance analysis and some assumptions made throughout the paper. I believe this paper should be rejected until such deficiencies can be overcome.**

Goto Main Index   Close Window