Objective of the present work

The primary objective of the present work is to demonstrate that current buffer management techniques for OODBMSs can be made more effective by simple synergistic modifications. The stand-alone configuration allows us to ignore network issues, thus enabling us to focus on techniques that minimize the effects of disk IO. However, the techniques that we will developed can be extended to work in the general peer-to-peer multi-node network model. We will choose the peer-to-peer network model since it allows multiple clients and one server to exist on a single node of the system. Page grained caching is chosen for its popularity and its superior performance when the system is well clustered.

We consider the following system entities for our model

- A database server serving one or more client processes using one or more CPUs.
- A number of concurrently running client processes on the same machine as the server, running applications requesting objects from the server.
- A disk only accessible to the server.
- A main memory page cache that the clients and server share.
- A disk queue, that rearranges disk page requests.
- A dynamic clustering thread which periodically changes the object to page mapping.
The evidence shall be presented with the help of empirical studies in which various existing buffer management techniques undergo synergistic transformations via simple and general transformation frameworks. In addition, an integrated buffer management cost model shall also be developed.

A more detailed breakdown of the proposed work is:

- To Propose an integrated cost model that encapsulates the interaction between static clustering, dynamic clustering, buffer replacement and pre-fetching. This cost model will perform two functions: to identify the various ways the buffer management techniques affect system performance; and to provide a starting point by which other researchers can approach the problem of developing synergistic techniques.

- To propose an opportunistic priority clustering framework (OPCF) which transforms static clustering algorithms into dynamic clustering algorithms. Opportunism refers to restricting re-organisation to in-memory pages only. Prioritisation refers to re-organising
the worst clustered pages first. Cost models shall be used to explain the performance advantages of this approach.

- To propose a Cache Conversant Clustering Framework (C3), that produces clustering algorithms that exploit the synergies between clustering and buffer replacement.

- To propose a Path and Cache Conversant Pre-fetching Framework (PCCP) which aims to produce a pre-fetching algorithms that exploit the synergy between pre-fetching and buffer replacement.

We will highlight the advantages of developing dynamic clustering algorithms that incorporate the synergies between dynamic clustering and static clustering algorithms. We will develop the opportunistic priority clustering framework (OPCF). OPCF is a general framework that creates dynamic clustering algorithms by placing opportunism and prioritization into existing static clustering algorithms.

We will highlight the performance gains resulting from exploiting the synergies between static clustering and buffer replacement. We will describe the cache Conversant clustering framework (C3). C3 produces static clustering algorithms which use knowledge of how buffer replacement algorithms behave to make clustering decisions. Like OPCF, C3 is simple and straightforward to apply and general in that it can be used to produce a whole family of different static clustering algorithms. Using the C3 framework, we will produced a new static clustering algorithm (C3-GP) which outperforms highly competitive existing algorithms in a variety of situations.

We will highlight the performance advantages of prefetching algorithms that incorporate synergies between prefetching and buffer replacement. We will describe the path and cache conscious prefetching framework (PCCP). Using PCCP, We will produce four new prefetching algorithms. We will compare PCCP against three highly competitive existing prefetching algorithms, PPM-1, PPM-2 and SP, in a variety of situations.
Future Work

Single Integrated Framework: We try to develop the synergistic frameworks, each exploit the synergies between two different buffer management techniques. However, we believe frameworks that exploit synergies between all four buffer management techniques would produce better results. Although we does not propose such a framework it nonetheless has setup the foundations for one to be developed. One possible approach is to produce a new framework by building OPCF on top of C3. Such a framework would be able to produce cache conscious dynamic clustering algorithms. We can then use PCCP to produce prefetching algorithms that use clustering information from the combined OPCF and C3 framework. This is only one of many possible fully synergistic frameworks (frameworks that exploit synergies from all four buffer management techniques) that can be produced.

General peer-to-peer and client/server models: We will try to develop techniques for a stand-alone single node of the peer-to-peer network model. However, the techniques can be extended to work for the general peer-to-peer and client/server network models. We will also discuss how each of the frameworks can be extended to work for the general peer-to-peer and client/server models.

Experiments on real OODBMS: The experiments will be conducted using a OODBMS simulator. Although the simulator provides good insights into algorithm performance through total IO and total IO stall time, it nonetheless is limited in its ability to simulate and measure all system components. Some system components that can be measured using real OODBMSs include buffer management computation time, multi-threading costs, locking costs and data structure space costs.

There are a number of areas the proposed work does not include. Following is a list of these as a way of defining the scope of the proposed study and to provide starting points for interested researchers to extend the work:

- The three frameworks OPCF, C3 and PCCP each address synergy for different pairs of the four buffer management areas (static clustering, dynamic clustering, buffer
replacement and prefetching). However, techniques that incorporate synergies of all four areas are beyond the scope of the study.

- This exploration of buffer management synergy is not aimed to be exhaustive. There may be numerous synergies between any pair of the buffer management areas. For example, there are numerous synergies between pre-fetching and buffer replacement: only non-memory resident pages are candidates for pre-fetching and buffer replacement algorithms play a large part in deciding which pages are more likely to be non-memory resident; early pre-fetching of pages may force the buffer replacement algorithm to evict in-memory pages earlier; buffer replacement algorithms may want to evict a wrongly pre-fetched page; etc. The proposed work does not attempt to exhaustively document all possible synergies between buffer management areas. Instead it aims to establish the fact that synergy does exist between these areas and that simple synergistic techniques can produce better performance.