HEWLETT PACKARD

DataMesh™ — scope and objectives

John Wilkes

19 July 1989

HPL–DSD–89–37rev1

Objectives

The DataMesh project will do research into *large*, *fast*, *highly functional storage servers*.

- *large* means in the range of 0.1–10 Terabytes of stored data;
- *fast* means that getting information from the DataMesh will be faster than getting it off local secondary storage devices;
- *highly functional* means that a DataMesh storage server will support:
  - tolerance of processor, storage device, and network failures
  - structured data (not just flat files), and many different data types simultaneously
  - 10:1 scalability in node count with incremental addition of processor / storage elements
  - a storage hierarchy of multiple device types
  - offloading data-intensive work from clients (e.g. image compression, tuple filtering)
  - client-specified data properties (e.g. access times, availability requirements)
  - automatic internal workload balancing.
- *storage servers* will be the preferred way to manage shared storage in a cooperative computing environment (CCE): their data will be uniformly and easily accessible throughout the network, and they will simplify centralized functions such as administration and backup.

Basic premises

VLSI CPUs are getting exponentially faster, and RAM is getting larger and cheaper at a similar rate. In the target timeframe (circa 1995), low-cost single-chip processors will provide 15–20 MIPS, and 16–64 Mbytes of RAM will fit into less than a dozen chips.

Mass storage devices are getting smaller, denser and cheaper, but neither their access times (seek plus rotational delay) nor their data transfer rates are keeping pace with the changes in CPUs and RAM.

Effective parallelism, caching, and data synthesis techniques will be necessary to improve the apparent speed of mass storage devices if their cost and storage density trends are to be retained.

The DataMesh premise is that pairing a powerful VLSI processor with each storage device will be a cost-effective way to build large, fast, highly functional storage servers.

Both client interfaces to storage servers, and the networks used to connect to them, will be standards-dominated, and optimized for local-area access.

Most clients of future storage servers will be system-level software, rather than individual applications.

The small physical scale of a DataMesh (one will fit into a computer room) will enable its internal interconnect to be optimized for its needs, and faster than the one used to connect the DataMesh to its clients.

Primary investigation topics

The key technical challenge for the DataMesh project will be to develop techniques to take full advantage of the parallelism available at the storage element and processor level, and combining these with ways to provide the functionality needed of a future storage service.

The primary areas of research in the DataMesh project are the following, in decreasing order of importance:

- high performance through aggressive use of parallelism, over a range of DataMesh sizes
- increasing the availability of data
- client-specified abstract data properties, and mapping them onto implementation choices
- support for a dynamically-extensible set of data types.

Secondary investigation topics

The following are also possible topics of interest, but of lesser importance to us than the ones listed above:

- designing an optimal storage hierarchy
- use of fast non-volatile storage
- extensive performance measurement and evaluation support
- file systems for write-once optical devices
- commercial data processing and supercomputers.

The DataMesh project will proceed by a sequence of experiments to develop and refine techniques, and by the construction of a sequence of DataMesh prototypes.