Prof. Dr. U. Brinkschulte
Institut für Mikrorechner und Automation
Universität Karlsruhe

Database Services

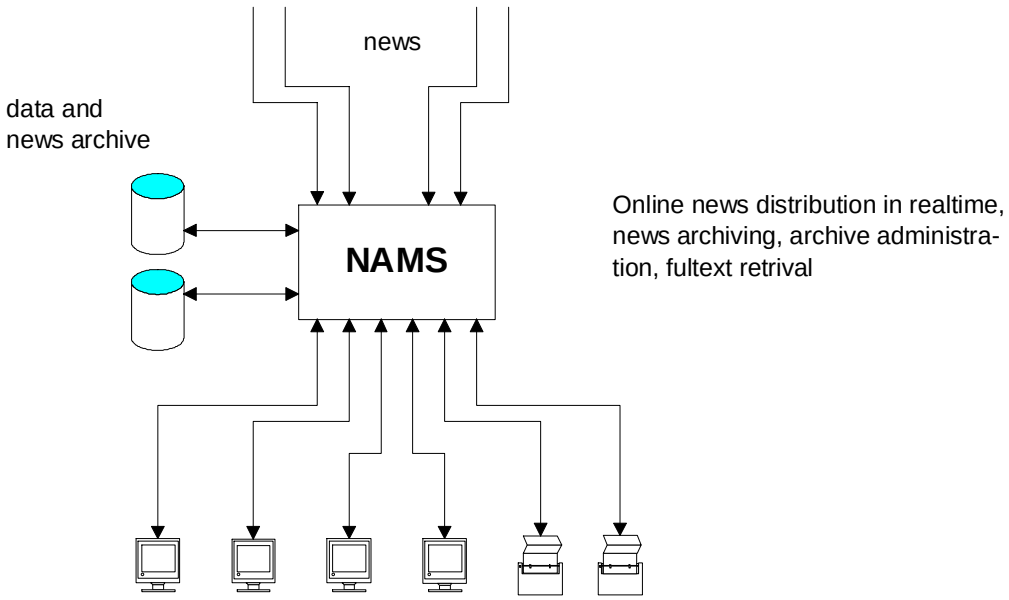
1. Database Requirements in Modern Systems

Database services are an essential part of modern information and automation systems. To be used in such systems, database services must meet some base requirements, such as:

- a powerful, but simple interface
- comfortable database operations
- data security
- network capabilities, cooperative and distributed operations
- realtime features
- a compact outfit
- platform independency (portability)

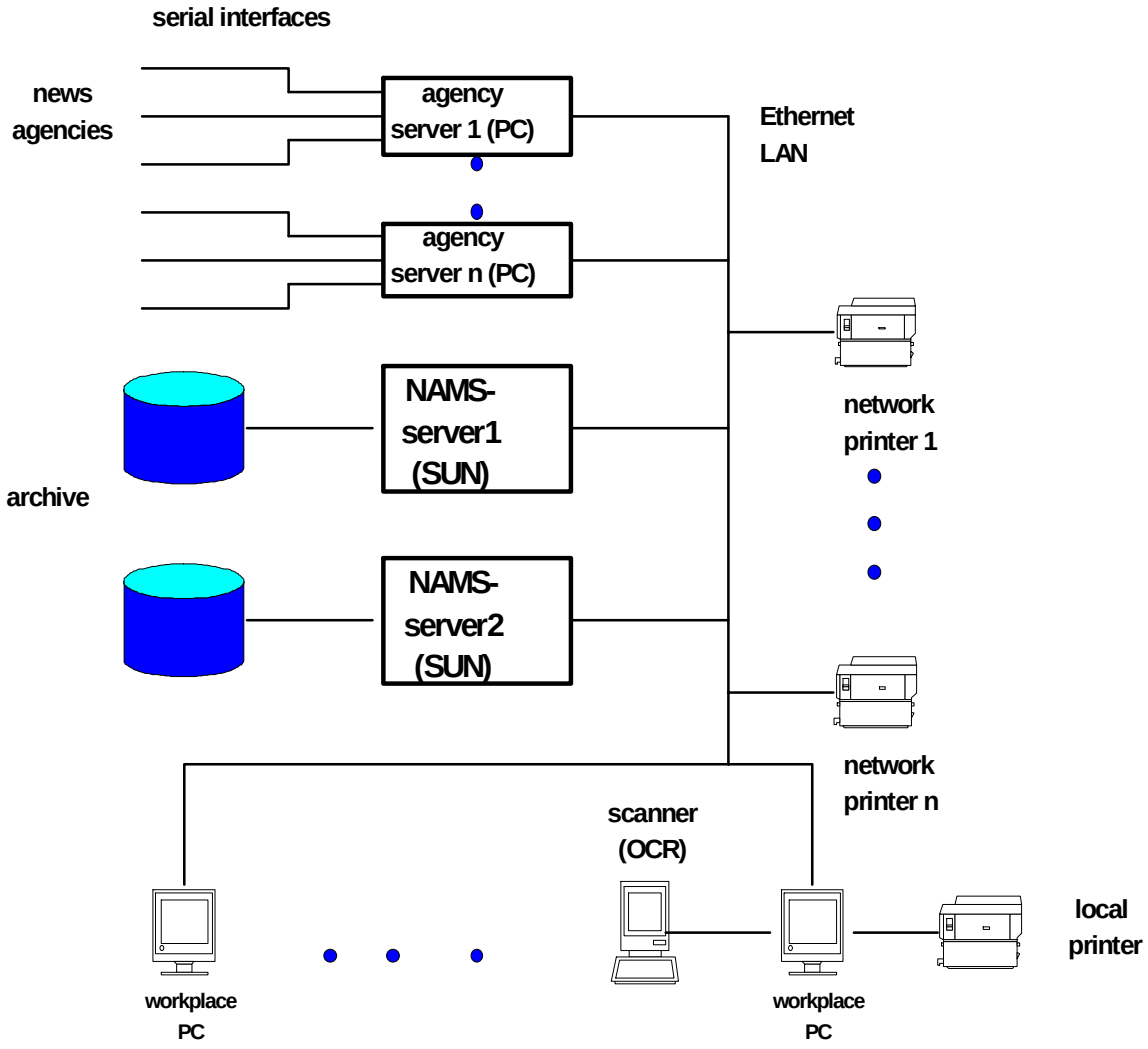
2. An Application Example

An example for a modern information and automation system is a news and archive management system:



The main task of this system is the online distribution of incoming news from news agencies in realtime. This means, users can select various distribution critereas, for example ‘all news from ressort politics with the keywords Clinton or Jelzin’. News, which meet these criterias are distributed to the user’s workplace. Another task of the system is the news and data archiving. By this a user can retrieve older news and other archive datas, e.g. biographies of important persons, using fultext retrieval methods.

The following picture shows the architecture of the news and archive management system ConText /1/. This system was developed by the GAI Informationssysteme GmbH in Germany and is used in the german governement and in several big german firms and newspapers.



ConText is a heterogenous distributed system that consists of several parts:

- Agency Servers

  The task of the agency servers is the reading and temporarily storing of agency news. This prevents news loss in case of server damage or system maintenance. Another task is the preanalysis and convertion of news (agency formats, keymaps, etc.) and the news transport to the NAMS-servers.

- Workplace PC´s

  The workplace PC's provide a graphical user interface for the interaction with ConText.

- NAMS Servers

  The NAMS (News and Archive Management System) servers are the kernel parts of the system. They execute the realtime news distribution to the workplace PC's and network printers according to the various distribution criterias and the the archiving and retrieval of news and other documents. Furthermore, they perform common database task for system administration.

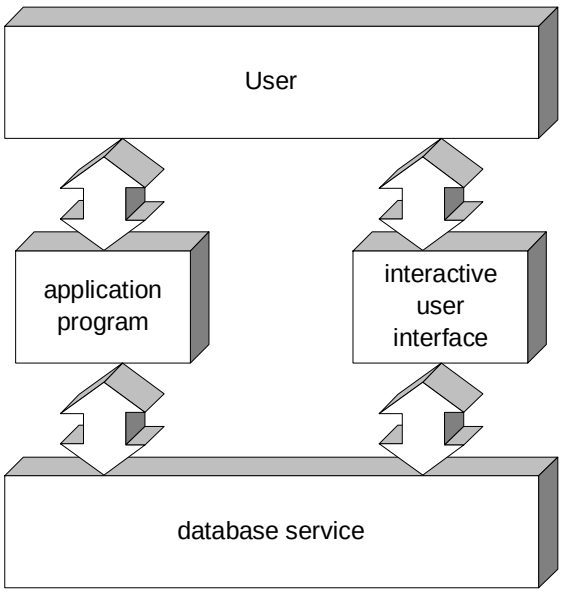
It is obvious, that this system provides a lot of tasks for database services, for example the storage and administration of user datas (names, rights, etc.), screen masks for the workplace PC´s, distribution criterias and other common administration data (statistical information about documents, users, resources etc.).
The requirements for these database services mainly are a distributed and cooperative database management in a heterogenous network, a fast data access for realtime news distribution, a 24 hour online availability and the protection against data loss.

3. Conception and Architecture of the MERLIN Database Service

The MERLIN database service was designed to meet the requirements of modern information and automation systems /3/. Its main properties are:

- an easy to use but mighty interface to the application programm
- cooperative and distributed network database management
- a very fast system operation
- 24 hour online availablity
- configurable data security
- a small amount of system resources needed
- portability

The MERLIN database service was used for example for all database tasks in ConText /2/. The following picture shows the conception of MERLIN.

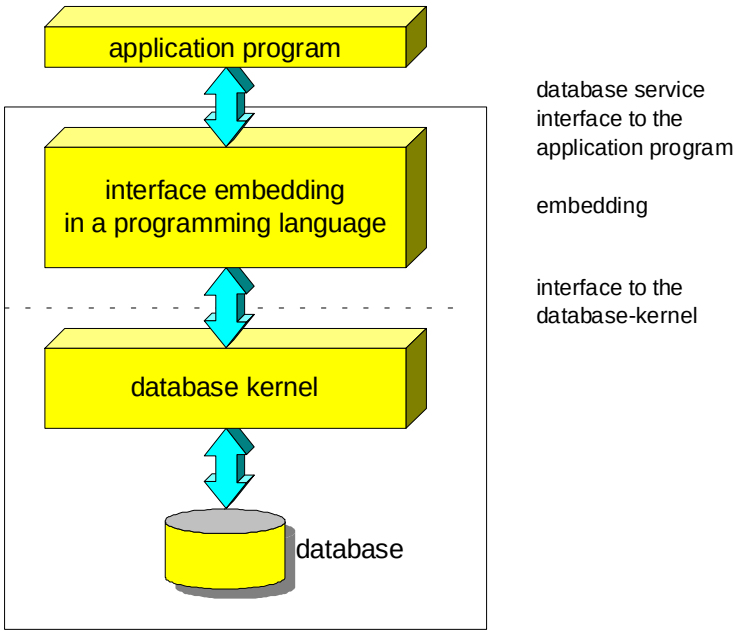


Database management is available to the application program as a service. This means a strict seperation between the application programm and the database management, leading to:

- reduction of programming costs
- improved abilities of modification and maintainement
- platform independency

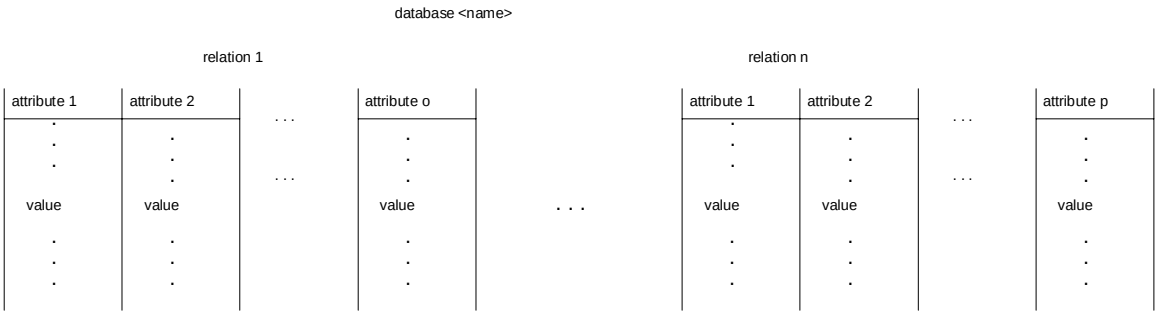
Additionaly MERLIN provides an interactive user interface for editing data structures and datas.

The architecture of MERLIN is divided in two parts, the application program interface embedding in a programming language and the database kernel.



3.1 Application Program Interface

First let´s have a closer look to the application program interface and its embedding .
For data storage, MERLIN uses relational data structures. This allows a simple modelling of even very complex data objects. Therefore, a MERLIN database looks as follows:



These data structures are embedded in the programming language C/C++. This provides best platform independency, because the language is available on nearly every computer system. The embedding of the relational data structures is done by a language extention. This means, new language elements for the definition of databases, relations, tupels, cursors etc. have been addes to C/C++:

```
database {relation name, ...} database name;

relation tupelname withkey {attribute name,...} relation name;

tupel    {attribute type : attribute name; ...} tupel name;

cursor   relation name cursor name;
```

For data access and manipultion, MERLIN offers a record oriented operational interface to the application program. This has prooved to be better to handle in a programming language as a set oriented interface.
The main database operations are:

Database Management
CREATE: database name x user x password x database
INIT: database name x user x password x database
EXIT: database
SAVE: database

Cursor Management
INITCURSOR: database x relation x cursor
EXITCURSOR: cursor

Access Operations
GET: cursor x mode x tupel x found
SEARCH: cursor x mode x condition x tupel x found
EXIST: cursor x mode x condition x found

Modification Operations
INSERT: database x relation x tupel
UPDATE: cursor x tupel
DELETE: cursor

These operations are also embedded by procedures and language extention for parameter declaration, for example:

```
db_create (database,database name,user,password);

db_open   (database,database name,user,password);

db_search (cursor,mode,condition,tupel);

db_insert (database,relation,tupel);
    ....
```

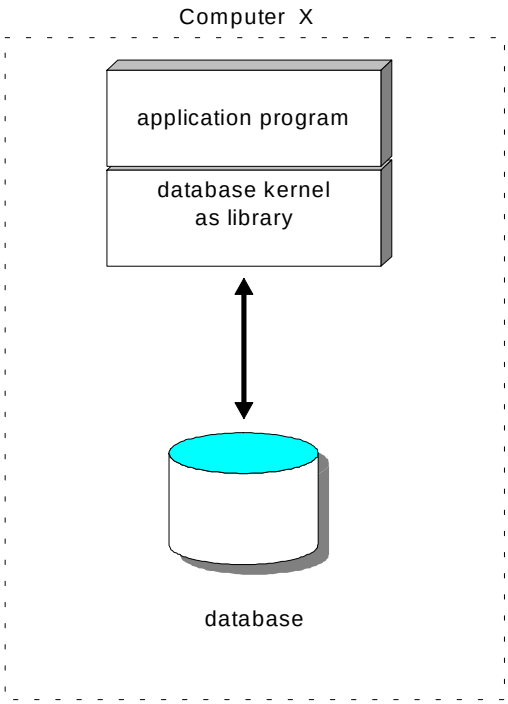
The language extention itself is realised by a precompiler. The principle is, that the new language elements are compiled to standard C/C++ and database kernel calls. The code created by the precompiler can now be compiled with a standard C/C++ compiler.

This embedding technique has several advantages: First, the language extention improves the programming comfort and the program testability compared with standard language elements and procedural embedding. Furthermore the precompiler concepts keeps up platform independency, because after precompiling any standard compiler can be used.
It must be said, that there is also a disadvantage: this technique means a little increase of programming inflexibility, because many tests are done at precompile time. For this reason, database structure must be known at precompile time. This is no restriction to most programs. However, MERLIN provides also a procedural interface without language extention, which allows to write programs without knowing the database structure. But higher flexibility is than bought with less comfort and higher error rates.

3.2 Connection to the Database Kernel

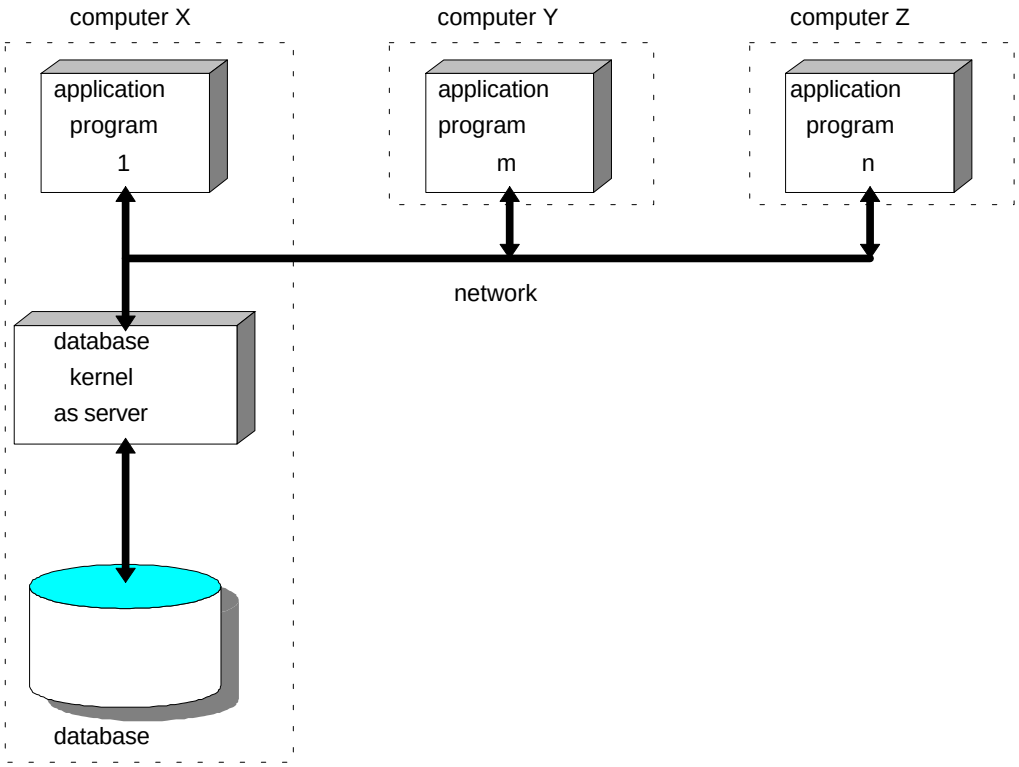
MERLIN allows three variants to connect an application program to the database kernel.

Variant 1 is the exclusive standalone connection. The database kernel is linked as a library to the application program. This is the fasted connection, but only one application program can access the database at a time.



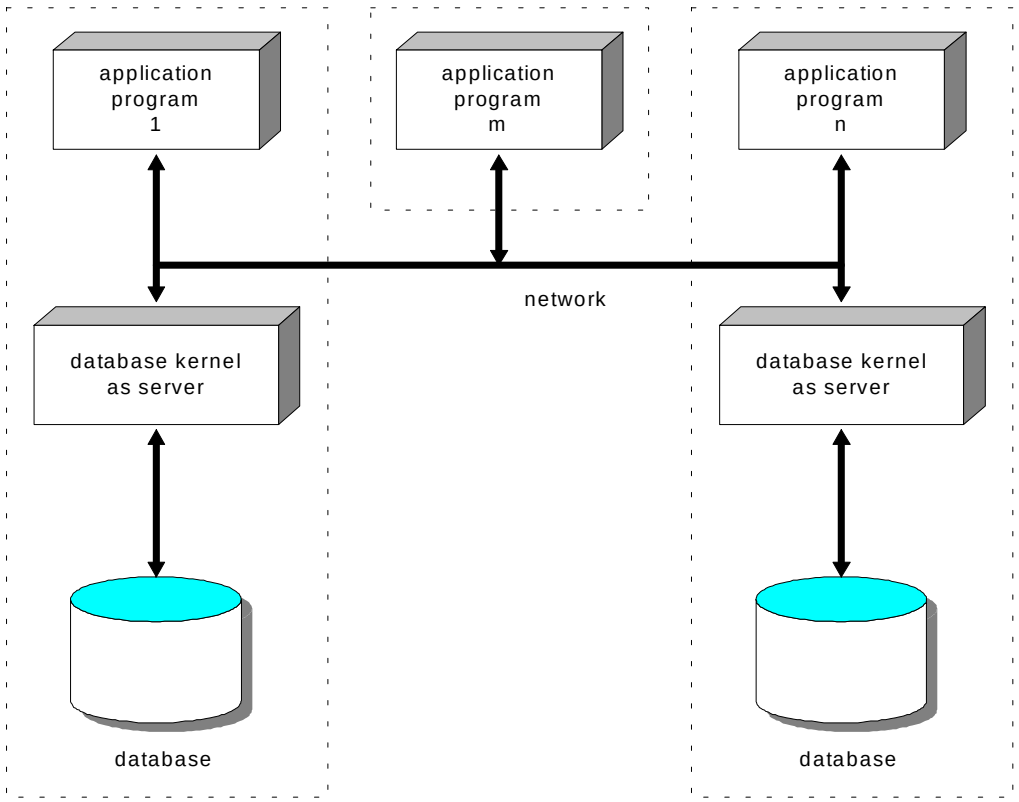
variant 1: exclusive standalone connection

Variant 2 is the cooperative networkwide connection. Here, the database kernel is a server. Application programs are clients spreaded over the network. More than one client can access the database at a time.



variant 2: cooperation networkwide connection

Variant 3 is the same as variant 2, but with multiple servers. Therefore, the databases can be distributed over the network. Each client can access more than one server at a time.



variant 3: cooperation networkwide connection with multiple servers

The point is, that the interface is the same for all variants. This means, application programs don´t need to know, which variant actually is used. The connection method is hidden to the application program and can be altered without altering the application program.

3.3 Database Kernel

The MERLIN database kernel is designed for realtime applications. His main properties are
- high performance
- 24 hour online operation
- configurable data security

This was reached by the use of new methods and scientific knowledge, for example
- all data is stored in main memory as far as possible.
- new storage replacement strategies, specially tailored for B-trees, are used.
- a new very efficient free storage management method is used.
- a configurable protocol mechanism for best compromise between security and speed has been implemented.

More detailed information about the MERLIN database kernel can be found in /4/.

4. Experiences

Implementations of MERLIN exist at the moment on PC under the operating systems Msdos, Windows, Windows NT, Linux and Sco-Unix, on SUN under Solaris 1 and 2 and on VAX under Vms.

First industrial applications are for example :
- the database management in a news and archive management system (as shown)
- the database management in a process management system
- the database management in a document management system

In addition there are some scientific applications and trial use at the university of Karlsruhe.

The first application experiences are very positive. They have shown
- a wide application spectrum of MERLIN
- a significant simplification of networkwide database management
- a significant reduction of programming costs
- an improved maintainability and modificability of application programs
- very good MERLIN realtime features

It seems, that the MERLIN database service is a useful tool for database management in modern information and automation systems.

References

/1/ GAI GmbH: ConText Systembeschreibung
GAI Informationssysteme GmbH, Friedrichshafen

/2/ U. Brinkschulte: Nachrichten und Archivsysteme, Innovation eines kleinere mittelständischen Unternehmens auf Basis der Informatik Systemtechnik
In: Conference Proceedings, 2. Beckman Kolloqium, Wismar Juni 1994

/3/ U. Brinkschulte: Architektur eines Datenhaltungsdienstes
In: Conference Proceedings, 39. Wissenschaftliches Kolloqium, TU Ilmenau, September 1994

/4/ U. Brinkschulte: MERLIN - Ein Prozeßdatenhaltungssystem für Echtzeitanwendungen
In: Conference Proceedings, Echtzeit 93, Karlsruhe, 1993
Herausgeber: Prof. Dr. H. Rzehak, L. Drebinger GmbH, München