WP -33

ZIB - A CASE STUDY OF SCIENTIFIC SOFTWARE DEVELOPMENT ORGANIZATION

by

Dr. V.B. Kaujalgi*

December 1992

INDIAN INSTITUTE OF MANAGEMENT
BANGALORE

**Professor, Indian Institute of Management, Bangalore*

P R E F A C E

This report describes the research as part of the Fellowship Programme of the "Centre for International Management Education". This is UNDP sponsored project (IND/87/011/C/99) at the Institute.

The study was carried out mainly in Germany with various Academic and Research Organizations from 16 September to 12 October 1991.

The Fellow thanks the Institute and various experts in Germany who helped in the present study.

Dr. V.B. KAUJALGI

I.I.M. Bangalore

ZIB - A CASE STUDY OF SCIENTIFIC SOFTWARE DEVELOPMENT ORGANIZATION

1. Introduction :

Organizations are being set up to develop software. Generally such organizations are concerned with the development of software for a specific customer or general packages in some of the functional areas of Management.

Scientific software is not very common in the majority of these organizations. Still the importance of such software cannot be reduced. Even though it takes more effort to develop such software, its impact on the specific applications is great. Marketing consideration may force only a few organizations in such activities, as the customers are few and specialized.

Generally, government of a country plays important role in setting up such organizations. ZIB is such an organization set up by the German Government.

Important objectives of ZIB are :

* Develop relevant scientific software
* Distribute such software to R & D and academic organizations in Germany
* Provide and Advice super computing facility to Academic and R & D organizations
* Create facility for access of Public Domain software world-wide through Electronic Mail

* Provide computing facility to graduate students, researchers and invitied visitors in the organization
* Hold seminars and workshops on scientific software development
* Access to current Library in relevant topics of scientific computing

2. **ZIB - Organization Structure :**

Eventhough, President and Vice-President are responsible for ZIB operations, they are also involved in the groups which are responsible for various types of development of scientific software.

As shown in Fig 1, the President is directly responsible for 3 groups and 2 groups indirectly. Similarly the Vice-President is directly associated with 3 groups and indirectly involved in other 2 groups. Both the functionaries are also responsible for one group each, which are their main interests.

The important groups as shown in Fig. 1 are :

* Numerik (Numeric Analysis)
* Numerische Software (Numeric Software)
* Symbolik (Symbolic Computation)
* Paralleles Rechnen (parallel computations)
* Kombinatorishe (combinatorics)

3. ZIB Computing Facility - Overview :

Any scientific software development organization like ZIB should have state-of-art computing facility. Computing power rather than input/output facilities are important in such activities. Fig. 2 shows the computing facilities at ZIB.

The main compouting power is available in CRAY X-MP/24 (2 cpus and 32 MB memory 8 GB disc space) and CRAY Y-MP2E/164 (1 cpu, 512 MB memory and 32 GB disc space). A robot controlled backup facility (STK4400 ACS, with 750 GB from about 8000 cartridges) forms important new facility. The OS of CRAY systems are is UNICOS (6.1.4).

CRAY X-MP has FORTRAN 77, C, ANSI C and PASCAL. Main libraries are from NAG and IMSL. For graphics there is New GKS. CRAY Y-MP has similar software facility.

The users have access to SUN3 and SUN4 workstations which have been networked. OS is the latest SUN OS (4.1.1).

Networking is through X.25 protocol using WAN and bridges. It is networked with Technical Univ, Berlin, Free Univ and other academic, R&D organizations in Berlin.

There are about 30 supercomputers in Germany from CRAY, NEC, etc as shown in Fig 3. They are accesseable through ERAN - electronic mail of Germany.

4. Scientific Software Projects - A Spectrum of Applications

Eventhough, R & D organization need not worry about advertising its products, ZIB has reasonable number of publications, brochures and annual reports for the users.

Total number of pre-prints upto December 1990 are 42 in which 21 pre-prints are for the year 1990. Similarly, there are 30 Technical Reports out of which 13 were published in the year 1990. There are 12 publications in scientific journals in 1990 by the scientists of ZIB.

R & D activities have to be at the leading edge of the topics. Hence, seminars by visitors is very important. There were 23 visitors in 1990 whose seminars covered wide spectrum of applications in scientific field.

Availability of software developed at ZIB to users is not standardized. This situation is typical in any R & D organization. Personal contacts play important roles. Some of the software are available in the Electronic Mail facility of ZIB.

REDUCE is a general purpose package for symbolic computation. The present version is 3.4. It is priced for SUN3 at $500, IBM RS6000 at $600 and CRAY at $1000. Support to users is not well defined. It is generally felt that more effort is required for such distribution activities and more programs should be brought in the form of packages at some

realistic price so that commercial organizations can also make use of R & D effort of ZIB. The procedure for determining price of a package is still not evolved. Most of the software efforts are dependent upon a few key scientists in that particular group.

5. eLib - Use of Electronic Mail in Scientific Computing

With the increase in information explosion in all the scientific fields, era of isolated computing is outdated. ZIB has made very good progress in Electronic Mail for the scientific community.

eLib is the electronic mail implementation by ZIB. It is connected to INTERNET, BITNET, uucp, X.400, RFC 822 etc. [1]. The main menu of eLib is shown in Table 1.

Important libraries of software available in eLib are listed in Table 2.

(a) GAMS : Library of software based on GAMS classifications scheme developed by NIST (National Institute of Standards and Technology) [2]. There are 20 major classifications.

(b) NetLib : It has software in 94 major headings [3]. The programs in this library are considered to be vey high quality.

---

[1] Interested readers may refer to :
"The Matrix" by J.S. Quarterman, Digital Press, 1990

[2] Contact : Ronals Boisvert of NIST, USA
Report No. NISTIR 90-4237

[3] Reference for NetLib : "Distribution of Math Software via Electronic Mail", Dongarra, E. Grosse, Comm ACM (1987) 30, 403-407.

(c) RedLib : Reduce Oriented library of Computation Algebra. It has 21 major classifications. Formal problems in mathematics, science and engineering are considered. A new set of operators in symbolic evaluation are added in this library.

(d) eLib : Contains CodeLib, the scientific software of ZIB and popular software from other important networks. It provides access to all the reports.
It is also dialogue server for the other libraries in (a), (b), (c) as described above.

According to ZIB, further research and effort are required to interface eLib, INTERNET and uucp completely. Incidently ERNET of India is uucp network. Large programs may be difficult to transfer in one attempt, particularly to uucp users in India because of the restrictions at the gateway.

Future attempts will be concerned with the transfer of files from eLib to uucp user at IIM, Bangalore.

6. **Supercomputing - Typical Benchmarks on CRAY**

Supercomputing facility is an important pre-requisite for scientific computing. A brief description of CRAY X-MP and CRAY Y-MP are in Section 3.

Benchmarking of supercomputers is gaining more importance to measure the computing power. Eventhough, there are no standard benchmarks available, typical results of a set of benchmark results from the programs in eLib of ZIB are shown in Table 3.

The benchmark programs were from David H. Bailey, "NAS (Numeric Aerodynamic Simulation Kernel Benchmark Program)" [4]. It consists of following seven programs:

1. MXMTST : Floating-Point Matrix Multiply Test
2. FFTTST : 2D FFT Test Program (Complex Radix 2 FFTS)
3. CHOTST : Cholskey Decomposition/Substitution Test Programme
4. BTRTST : Vectorized BLOCK Tri-Diagonal Solver in the J Direction for K = constant planes.
5. GMTTST : Compute Solid-Related Arrays (Gauss Elimination of matrix of Wall Influence Coefficients)
6. EMITST : Emit new Vartices to satisfy boundary conditions for forces, pressures, etc.
7. VPETST : Routine to invert 3 pentadiagonals simultaneously

Table 3 shows the MFLOPS (Millions of Floating Points per Second) for programms without vectorisation, after compiler vectorization (cv) and after FORTRAN Pre-Processor (fpp). These programs were run on both CRAY X-MP and CRAY Y-MP. MFLOPS for parallelization by fmp for CRAY X-MP were also recorded. MFLOPS were obtained, dividing number of Floating Point Operations by seconds. These results test hardware system software and FORTRAN Compiler characteristics for super computing.

---

[4] Programs developed by NAS at NASA Ames Research Centre by D.H. Bailey, June 24, 1988. .pa

Important observations are :

(a) CRAY Y-MP-E/164 is about 2.5 times faster than CRAY CRAY X-MP/24.

(b) Sometimes programs take longer time after fpp or fmp.

7. **Points for Discussion/Study :**

This case study is particularly interesting as scientific software development is becoming more important and complex compared to the commercial software.

Following points are worth discussion/study:

(a) Scientific software and commercial software - main differences in development in India and Germany.

(b) Costing of scientific software - methods in which R & D component is involved.

(c) Strategy for proper identification of scientific software for development.

(d) Organizational structure for scientific software development.

(e) Marketing/Distribution activities for scientific software packages

(f) Strategy for distribution/maintenance/enhancement for scientific software.

(g) Comparison of network with ERAN, bitnet, edunet, internet. Compare the networks in India like ERNET CSIR-NET, INDO-NET, etc. with eLib.

(h) Comment on the utility of elib and netlib software. How recent are these public domain software?

8. Concluding Remarks :

ZIB funding from German Government has similar advantages/disadvantages as R & D organizations in India. Both the countries may get benefitted by learning from the other.

Costing of software is an emerging subject in software engineering. This is particularly true for scientific software in view of higher level of skill of manpower and specialized resources. ZIB experience may be important for similar organizations in India.

Network and Electronic Mail are key facilities in scientific computing. India has to go a long way in this direction. There are important lessons to learn from the experience of ZIB.

Table 3 : Benchmark of CRAY supercomputers
(in MFLOPS : Millions of Floating Point Operations per Second)

PROGRAM	CRAY X-MP/24			CRAY Y-MP 2E/164		
	CV=ON fpp=OFF fmp=OFF	CV=ON fpp=ON fmp=OFF	CV=ON fmp=ON fmp=ON	CV=OFF fpp=OFF	CV=ON fpp=OFF	CV=ON fpp=ON
MXMTST	173.98	174.93	176.28	26.15	278.64	278.74
FFTTST	53.94	52.86	41.67	27.51	76.35	76.39
CHOTST	55.65	55.92	45.96	8.18	91.19	91.23
BTRTST	96.61	95.65	75.55	22.36	152.80	151.71
GMTTST	74.42	83.81	86.62	10.00	134.35	251.35
EMITST	115.78	116.79	114.92	19.81	180.59	182.20
VPETST	41.21	42.04	34.07	23.65	51.80	51.87

fpp - FORTRAN Pre-processor, CV - Compiler Vectorization,
fmp - FORTRAN Multi Tasking Processor

(Note: CRAY X-MP is duel processors and CRAY Y-MP is a single processor systems. Hence fmp is possible only on CRAY X-MP.)

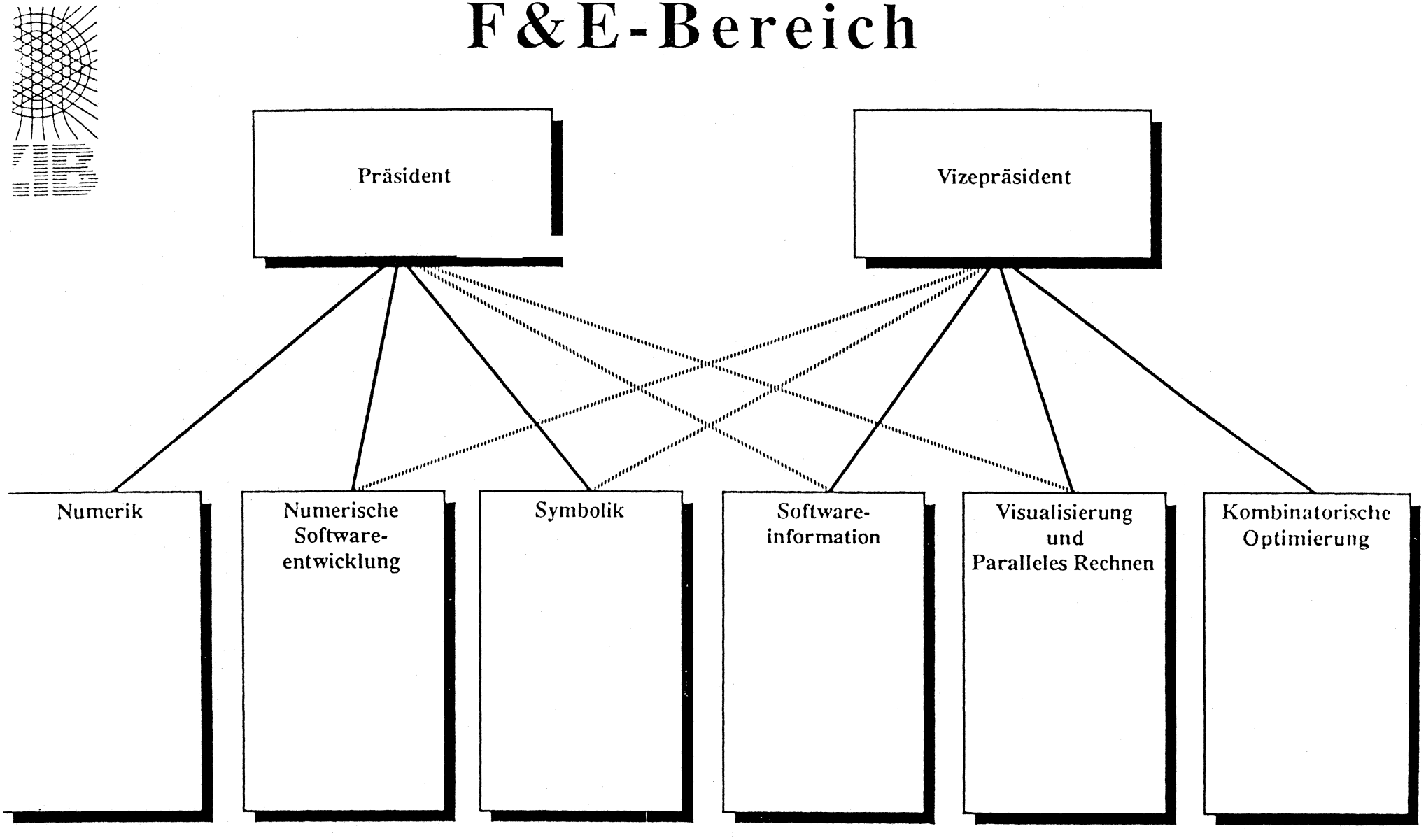

FIG 1 : ORGANIZATION STRUCTURE OF ZIB

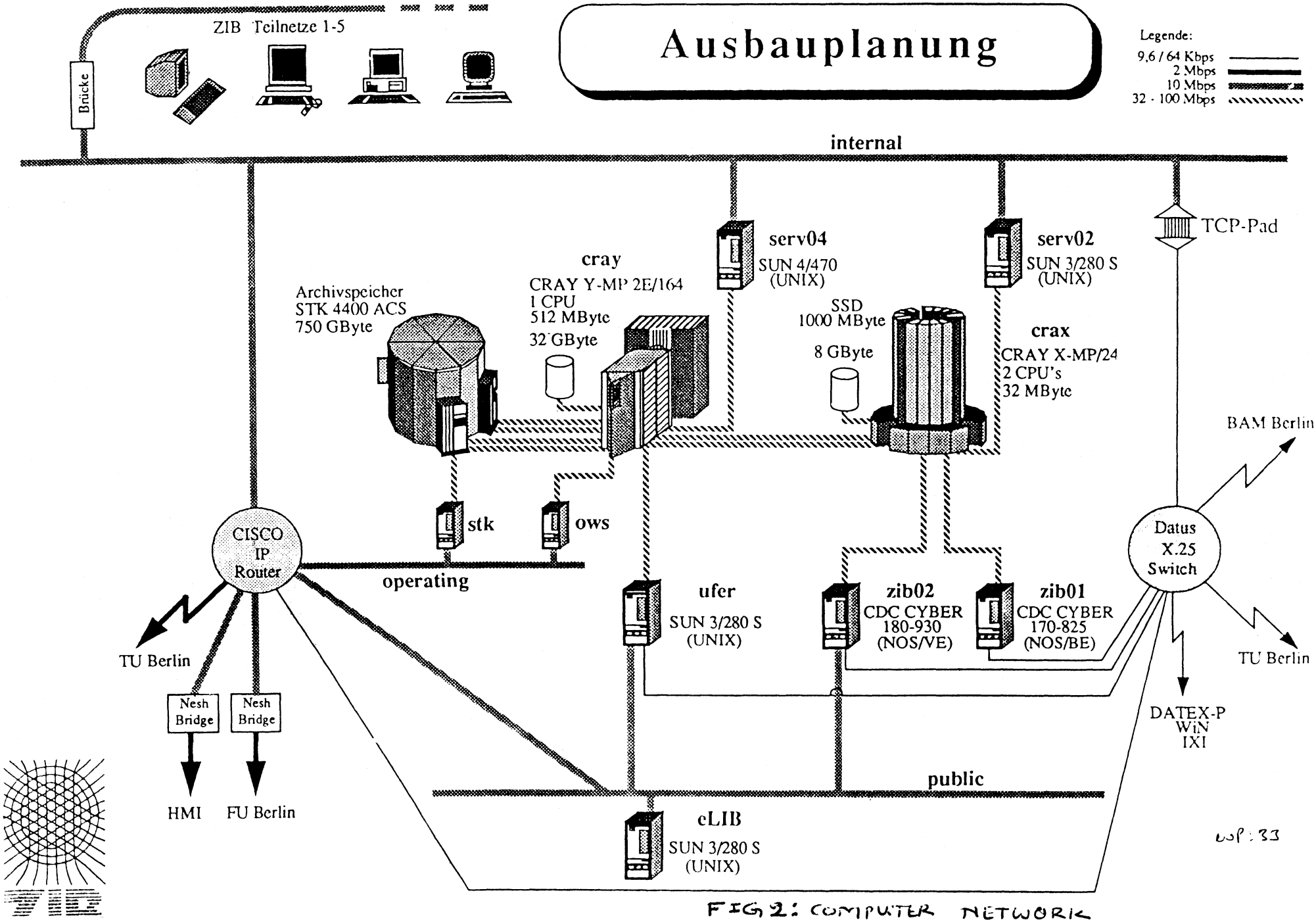

FIG 2: COMPUTER NETWORK

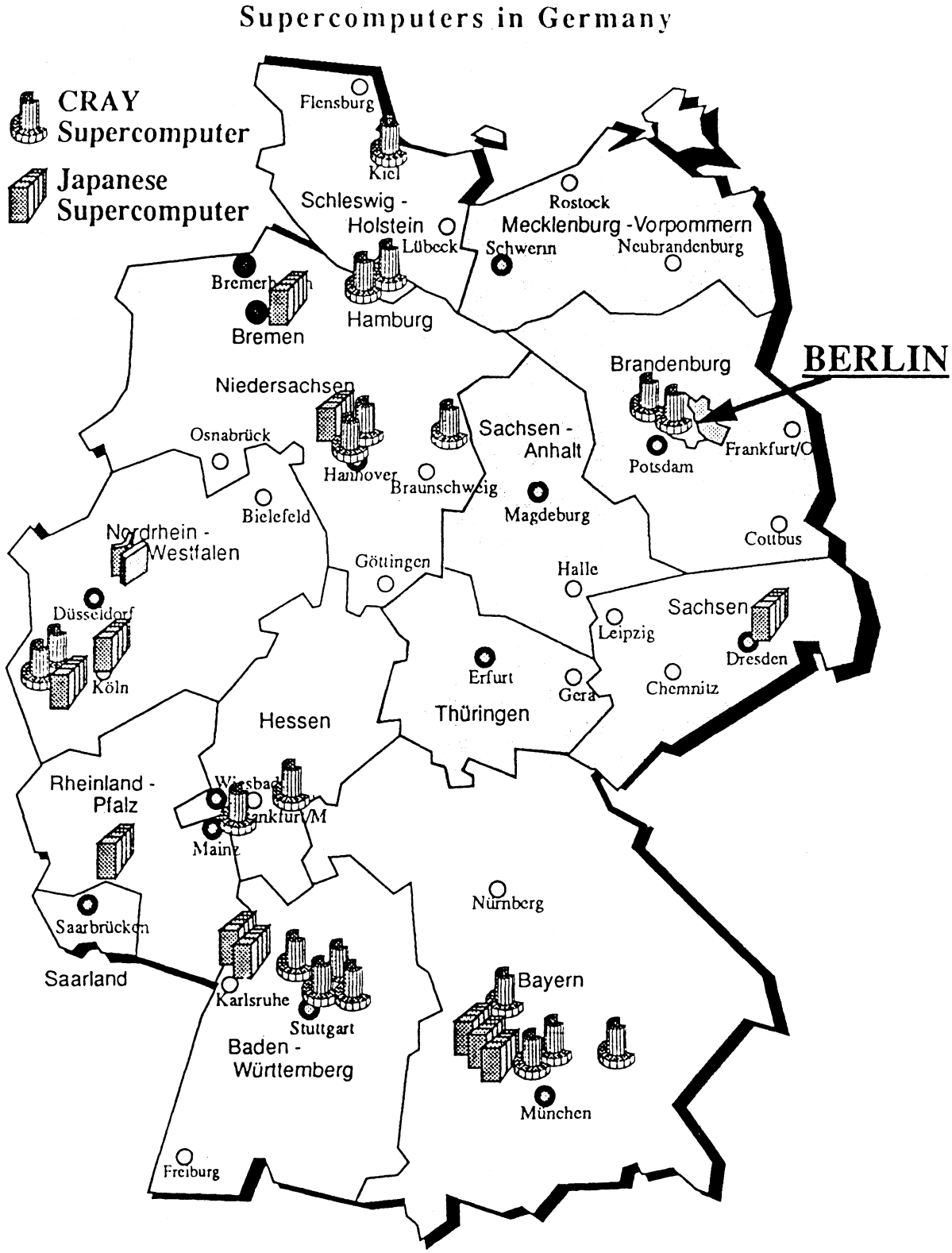

FIG 3 : SUPERCOMPUTER NETWORK FACILITY IN GERMANY