CHAPTER 3. AN OVERVIEW OF SOVIET HIGH-PERFORMANCE COMPUTING

3.1 Introduction

Soviet high-performance computing has a long and complex history. To set the stage for subsequent discussion, this chapter provides an overview of this sector from the inception of digital computing in the USSR through the present, highlighting the role of the Institute of Precision Mechanics and Computer Technology (ITMVT) in Moscow. We present the history and chronology of projects which reached at least the prototype stage, or for which prototypes are currently being contructed. This chapter raises a number of issues of technological development and HPC within the Soviet context which will be discussed later, and lays the groundwork for a comparison of a variety of different lines of HPC development.

3.2 HPC Efforts at ITMVT

3.2.1 Early Uniprocessors

The Father of Soviet computing, S. A. Lebedev, began work on the first Soviet digital, stored-memory computer in 1946 at the Institute of Electro-Technology in Kiev. Called the MESM (*malaya elektronnaya schetnaya mashina*-Small Electronic Calculating Machine), it was built under difficult post-War conditions outside of Kiev between 1947 and 1950, when the first programs were run on it. The MESM was formally completed and put into use in 1951 [Liso88, 20-21; Crow93]. Figure 3-1 presents a timeline of early uniprocessors built under Lebedev's direction.

In 1950, Lebedev was invited to come to Moscow to head a new laboratory which had been established for the development of electronic digital computers at the Institute

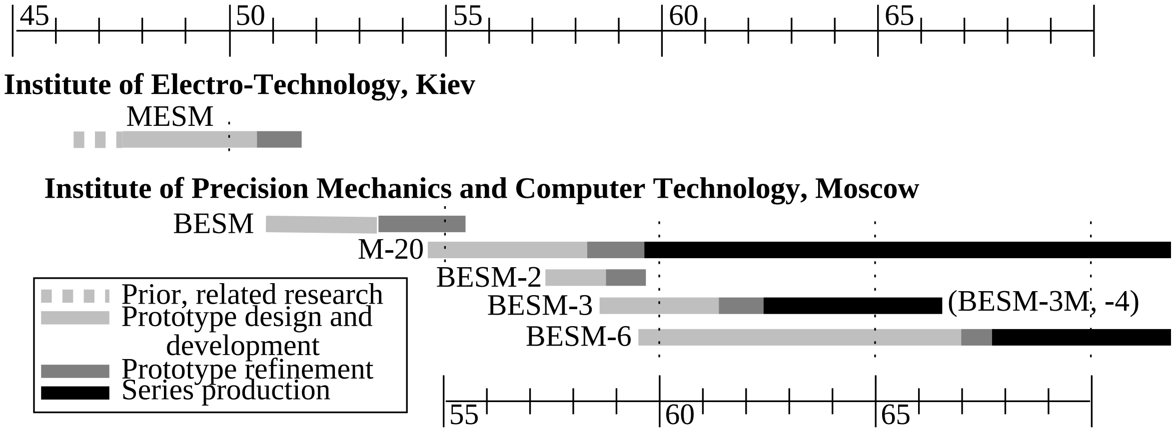

Figure 3-1 Early Soviet High-performance Uniprocessors

of Precision Mechanics and Computer Technology [Bard87; Bard88]. As the head of this laboratory and director of ITMVT (1953-1973) Lebedev established ITMVT as the premier developer of high-performance computing systems in the Soviet Union. ITMVT has maintained that role to this day. Lebedev cultivated a strong tradition of building fast machines with original architectures which were manufactured by domestic industry and used in real-world applications. He built complete systems, and pioneered advances in all aspects of computing, including components, power supplies, construction technologies, peripheral devices, software, and, of course, new architectures. Although software was a necessary part of any system, Lebedev and others felt that the mission of ITMVT was to build the fastest machines possible, even at the expense of software compatibility. The volume of software written for the early machines was small, and he felt that the users would always be ready to re-write existing code. Software compatibility did not become a major design goal until the 1970s, with the El’brus computers described below.

Lebedev began work on the BESM (*bystrodeystvuyushaya elektronnaya schetnaya mashina*-High-Speed Electronic Calculating Machine) in 1950. A three-address, floating-point machine with a 39-bit word length, the BESM was built using vacuum tubes and mercury delay lines for operational memory. The machine had a main memory capacity of 1024 words and a processing rate of approximately 1,000 operations per second. It was completed in 1952, and accepted by a State commission in 1953 [Crow93].

During these years, Lebedev’s group competed with others for equipment and components. In 1951, the USSR Council of Ministers decided to fund two competing projects with the goal of creating a world-class electronic digital computer. One was Lebedev’s BESM project; the other, called the Strela, was built in the Ministry of Machine and Instrument Construction in Moscow. The latter had closer ties to industry (and the supply of components) and considerable political influence. This project made it difficult for Lebedev to acquire the quantities of electro-static cathode ray tubes he needed. Although Lebedev did get the CRTs necessary after considerable complaining to the Academy of Sciences and government officials, the experience further convinced him of the need to establish close relationships with industry. Such ties were achieved during the early 1960s when ITMVT and a number of applied research institutes of the Academy of Sciences were transferred to industrial ministries that had equipment and production facilities the Academy lacked [Lakh90, 39-40]. Lebedev promoted close contacts and active involvement by the factories from the earliest stages of research and development as a means of significantly shortening the development cycle [Bard87].

In 1954, Lebedev started a new project to build a computer using germanium semiconductor diodes. In 1955, he began working in close cooperation with a special-design bureau at the Moscow Calculating-Analytic Machines (SAM) Plant in Moscow for the development and series production of such a machine, called the M-20. Some problematic circuits led to strained relations between Lebedev and his SAM counterparts, and Lebedev re-focused his efforts on an an industrial version of the BESM, called the BESM-2. Both the M-20 and the BESM-2 were completed in 1958. The latter had an operational memory of 2048 words and an average processing rate of 7-8,000 operations per second. The M-20 had a 45-bit word length and a nominal performance rate of 20,000 floating-point operations per second.

Lebedev subsequently developed the transistor-based BESM-3 and two slightly upgraded production versions, the BESM-3M and BESM-4. This three-address machine was based more directly on the M-20 than the BESM-2, having a 45-bit word length and using the M-20 instruction set. It had 4-8K words of memory and ran at 20,000 operations per second, like the M-20 [Grub77; Targ80]. The BESM-4 was completed in 1961 and was in series production from 1962-1966 [Gryz88, 93].

In 1959, Lebedev began preliminary work on a machine based entirely on a semiconductor component base. The first mock-up of the BESM-6, Lebedev's most enduring work, was completed in 1964. The BESM-6 was a single-address machine with 48-bit words and a performance of one million operations per second. A full prototype underwent factory testing in 1966, and the machine passed state testing in 1967[1] [Laut91, 11].

The BESM-6 had a profound impact on the Soviet scientific computing community. Between 1966 and approximately 1980 over 350 units were manufactured at the Moscow SAM Plant [Supe91b, 14]. For approximately two decades the machine was the computational workhorse of Soviet science, probably logging more hours of use per unit manufactured than any other serially produced computer in the world. It was beaten in the HPC race only by the 3 MIPS CDC 6600 (1964). Thousands of programmers worked on the

---

[1] This chapter presents a number of figures indicating machine chronologies. We consider the prototype development period to last from the time a decision was made to build a physical machine to successful completion of ‘‘state testing.’’ For state testing, a government commission consisting of representatives of principal users, prominent members of industry and academia, and government policy makers–usually under the chair of Academician A. A. Dorodnitsyn–examined the machine, ran tests, and verified that the prototype conformed to the technical statement of work established at the outset of development. Passing state testing signaled successful completion of the development phase, although did not rule out further modifications to the system.

machine which became the first Soviet computer with a full and rich set of systems and applications software. The machine was very reliable, by Soviet standards, and easily maintained. Over the years, existing systems were upgraded in the field by adding semiconductor memory and enhanced peripherals. The BESM-6 incorporated many "firsts" for Soviet computing. These included virtual memory, pipeline processing, hardware memory protection, scratch-pad memory, and many others [Laut91, 8-10]. Many of the lead engineers of prominent HPC projects at ITMVT and elsewhere in later years were part of the BESM-6 development team.

3.2.2 ITMVT Computers of the Late 1960s and 1970s

3.2.2.1 BESM-10 and AS-6

Following the completion of the BESM-6, its development team divided to work on two different projects. The chief engineer of the BESM-6 project, V. A. Mel'nikov, worked on a system called the BESM-10, which was to be a 64-bit successor built using integrated circuits. A. A. Sokolov began work around 1969 on a system called the AS-6. Little has been written about this machine; A. A. Sokolov has a reputation for caring little for writing and publishing. The AS-6 was a multi-machine system designed for reliable real-time control of objects. Running at 5 MIPS, its strengths were in inter-computer interaction, such as high data transmission throughput and direct memory access by I/O subsystems. Some distributed operating system and network protocol ideas were pioneered in it. The AS-6 was built in relatively few years thanks to the experience of the development team, the use of the same (i.e. available) component and construction base as the BESM-6, and the urgency of finishing the machine in time for use in the mission control system for the Apollo-Soyuz space project in 1975. The system for that mission included a number of tightly-coupled AS-6 and BESM-6 computers. This machine was not compatible with the BESM-6. Sokolov felt that requirements for software compatibility ‘‘tied ones hands’’ and that achieving high performance was more important. Between 1975 and approximately 1980 on the order of 15 such systems were manufactured.

3.2.2.2 El’brus

Simultaneously, V. S. Burtsev and others, including B. A. Babayan, began designing a general-purpose multiprocessor dubbed the El’brus. Burtsev had previously worked on special-purpose multiprocessors for the military at ITMVT, and was proposing a large multiprocessor system which offered not only increased performance, real-time capability, high reliability, and increased main and peripheral storage, but also ease of programming and inter-generational software compatibility. The emphasis on software was something of a departure from ITMVT tradition, but would serve this line of development well in the future.

At the time the designs for the BESM-10, AS-6, and El’brus were emerging, Soviet leadership was becoming increasing concerned about the ‘‘computer gap’’ with the West in high-performance computing. Decisions had been made in the late 1960s to develop a Soviet line of IBM System/360 compatible machines to address deficiencies in general-purpose mainframes for economic applications. The introduction of Control Data Corporation’s CDC 7600 in 1969 highlighted a serious lack in high-performance systems for scientific applications. The CDC 7600 was over an order of magnitude faster than the BESM-6. While the latter was in full series production, no machines were ready to succeed it. There were few high-performance alternatives to the AS-6, BESM-10, or El’brus.[2] Discussions of which systems to support coincided with the need to select a

---

[2] The M-10, a vector-oriented parallel system designed for image processing is described in section 7.6. It undoubtedly profited from policy makers’ sense of urgency. It too had to be designed and developed, however. Supporting development of this machine would not resolve the issue of which systems to develop at ITMVT.

successor to the aging S. A. Lebedev who would soon (in 1973) step down as director of ITMVT. The AS-6, incorporating the same technology as the BESM-6, could easily be built at the Moscow SAM Plant. With the large-scale efforts underway to tool up for the manufacture of ES mainframes, Minradioprom did not have the production resources to support both the BESM-10 and the El'brus lines, however. Ultimately, this issue was settled less on technological grounds than political ones. Burtsev was able to line up more high-level support than Mel'nikov in the Military-Industrial Commission (VPK) and Minradioprom itself. The El'brus was supported, and Burtsev was selected to succeed Lebedev in 1973.

The El'brus computers are discussed in depth in chapter 4. These 64-bit machines were designed for the most demanding military and civilian applications. They had to offer not only high performance, but high reliability. The principal customers also wanted to be able to develop real-time software quickly and effectively. El'brus designers reviewed a number of Western projects and felt that these requirements were best met in the Burroughs 700 Series. Many of the features of these machines were worked into the El'brus: a stack-based multiprocessor architecture, modular construction, multiprocessing, dynamic resource allocations, hardware tags, hardware oriented towards efficient compilation and execution of programs written in high-level languages, software compatibility between system generations. The El'brus architecture is not a copy of the Burroughs systems, however, as we shall see.

With the El'brus computers, ITMVT continued its tradition of driving development in the Soviet computer industry. High requirements forced the development of new generations of integrated circuits, power supplies, cabling, design tools, printed-circuit boards, connectors, etc. The El'brus computers pushed the boundaries of each of these technologies simultaneously. The development times suffered as a result. The technical and administrative challenges were staggering. Technically, delays resulted from having to develop, debug, and incorporate so many new technologies simultaneously. Administratively, each relationship with a factory had to be negotiated through a long chain of Ministry and Party officials. The factories themselves were generally disinclined to upset current production schedules with the introduction of new technologies and exercised a non-trivial *de facto* influence over production. As we can see from figure 3-2, the El'brus-1 and -2 took over 10 and 15 years, respectively, to reach series production. Similarly long development periods characterize other industry driving projects.

During the latter half of the 1970s, several events worked concurrently to shape ITMVT's future developments. First, the BESM-6 had been in production for a decade. Hundreds of systems had been installed, and a large and active user community had grown up around it. Something had to be done to support this community in the future. Second, the first prototypes of the El'brus-1 were under construction. Although less powerful than the El'brus-2 which embodied the original aspirations of the designers, the El'brus-1 gave users hope that a new generation of systems would be available shortly. Third, the AS-6 had been completed. This system would not be in production long, however, because many potential customers, believing that El'brus computers would soon be available, preferred to wait for the new machines. ITMVT in fact promoted this view. Fourth, Cray Research Inc. introduced the Cray-1 in 1976. Its clear demonstration of the viability of the vector-pipelined approach made a strong impression on computer engineers in both the USA and the Soviet Union. Its peak performance of 160 Mflops further underscored the lag in high-performance computing between the two countries. Fifth, during the late 1970s and early 1980s the political climate between the Soviet Union and the West worsened considerably. Computers were becoming more and more important in defense systems, and support for technologies of strategic importance in general grew. In

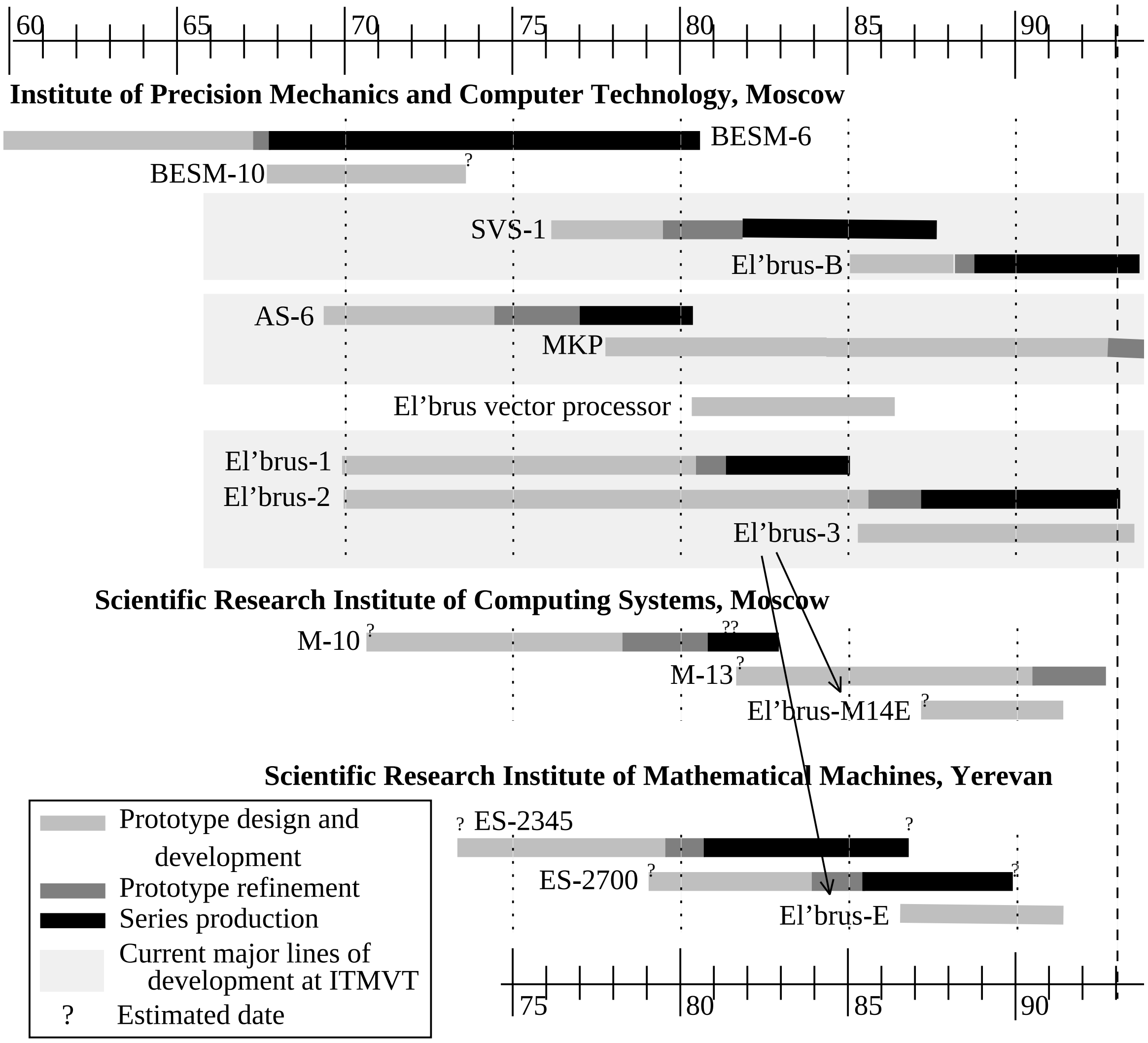

Figure 3-2 High-performance Computers at ITMVT and Related Institutes

the following sections we shall see how these factors shaped the direction of R&D at ITMVT.

3.2.2.3 SVS-1

To provide a migration path for BESM-6 users, engineers at ITMVT built a BESM-6 compatible processor called the SVS-1 which, using some of the same components and technologies as the El'brus computers, could be incorporated into an El'brus configuration as a specialized processor. Like the BESM-6, it was a 48-bit machine. Its architecture was largely the same as the BESM-6's; small changes to improve performance were made to the adder, multiplier, and communications unit [Tyap82]. It was compatible with the BESM-6, although the BESM-6's operating system had to be modified to be able to function in the context of an El'brus configuration [Bala82].

One or two such processors could be incorporated into an El'brus configuration [Mvke80]. Configurations in which these were the only CPUs were called El'brus-1K2 and El'brus-1K4. SVS-1 based systems offered BESM-6 users improved performance and significantly better peripheral storage systems. The SVS-1 processor had an average performance of 2.5-3.5 MIPS. Because of the improved peripheral storage capability provided by the El'brus configurations, it could run 10-12 times faster than a BESM-6 on applications with large volumes of data [Mvke80].

Initiated around 1975 or 1976, the SVS-1 passed state testing in 1979 [Katk78; Tsan79; Mvke80; Supe91b, 13]. Between 60 and 100 units were manufactured between 1979 and 1987 [Supe91b, 13].

3.2.3 ITMVT Computers of the 1980s

3.2.3.1 El'brus Vector Processor

The Cray's vector-pipelined architecture inspired the development of a specialized vector-pipelined processor which could be incorporated into an El'brus configuration. Built using the El'brus-2 component base and construction, the processor employed a rather traditional vector-pipelined architecture with multiple, pipelined vector functional units and register files. In contrast to the Cray-1, which had three floating-point functional units (add, multiply, reciprocal), the El'brus-2 vector processor had two add, two multiply, and one division floating-point units.[3] The El'brus vector processor also lacked scalar processing capabilities; this was to be provided by the native El'brus-2 processors. Like the Cray, the El'brus vector processor provided chaining between functional units to allow the output of one to be used directly as the input of another. The El'brus vector processor used a switch to provide chaining between any two functional units. Designed for a 40 nsec clock period, the vector processor would have had a theoretical peak performance of 125 Mflops [Burt85; Burt89].[4]

The vector processor was never completed. During the early 1980s conflicts arose between Burtsev and others within ITMVT and without. The details of the conflicts are not public, but they undoubtedly involved dissatisfaction with the progress of the El'brus program, differences over the future direction of El'brus development, and personal issues. Burtsev was removed as director in 1984 and replaced by G. G. Ryabov, formerly the head of ITMVT's computer-aided design division. Ryabov did not support the project, although he allowed it to continue for a couple of years with dwindling support. In 1986 the project was terminated and the division was disbanded [Mvs89, 44].

One of the basic issues about the best path of HPC development which divided Burtsev and others like Ryabov was the use of special-purpose vs. general-purpose processors. Burtsev felt that high-performance could best be achieved through the use of a general-purpose system with attached specialized processors. "[T]he solution to the problem of achieving maximum speed is best achieved in a multiprocessor complex with extensive use of specialized vector processors, alongside universal processors (oriented towards scalar computations)" [Burt89, 8]. Others, like Ryabov, held that any reasonable

---

[3] The El'brus vector processor also had a logic unit. Besides the floating-point units, the Cray-1 had three vector units (add, logic, shift), four scalar (add, logic, shift, POP/LZ) and two address (add, multiply). Since the Cray-1 could produce two floating-point results per 12.5 nsec clock period, the theoretical peak performance was 160 Mflops.

[4] Five results per clock period.

application would consist of a mix of scalar and vector operations and that general-purpose processors were the best means of achieving high performance on a broad set of applications. When Ryabov became director, the latter view prevailed. Work on the vector processor ended, while work on the Modular Pipeline Processor was accelerated. The use of specialized processors in the El’brus-3, a successor to the El’brus-2, also was rejected.

3.2.3.2 Modular Pipeline Processor

Following his work on the AS-6, A. A. Sokolov began working on a new computer called the Modular Pipeline Processor, discussed in section 7.2. On the one hand, this machine was influenced by Sokolov’s previous work on the BESM-6 and AS-6 and by the vector-pipeline ideas of the Cray systems. The MKP implements pipelining ideas differently from the classical Cray-like architecture in that it integrates the vector and scalar operations more tightly into a single, general-purpose processor. On the other hand, the MKP was designed for a user community which could not necessarily afford systems as expensive as the El’brus. This factor made economy of hardware, ease of maintenance, and the ability to adapt the configuration to a variety of user needs major design and construction goals. The MKP is “modular” in the sense that a variable number of processors (theoretically with a variety of functional unit sets) can be combined in configurations with an assortment of other systems and peripheral devices.

Sokolov had worked on the MKP design since the late 1970s. Shortly after Ryabov became director, the project began to receive strong support. At this time, the technical characteristics of the next generation of ECL gate arrays were becoming known and the MKP was designed to achieve one Gflops in a dual configuration with a clock period of 10 nsec. The first full prototypes of this system were completed in 1990.

3.2.3.3 El’brus-B

In 1985, engineers began working on upgrading the SVS-1. In 1986, they changed plans and decided to develop a machine which, while compatible with the BESM-6 and SVS-1, had a new architecture and a new component base. The El’brus-B (also called the El’brus-KB) has three operating modes. The first mode provides full compatibility with the BESM-6 and SVS-1 processors. It uses a 48-bit word and 15-bit effective address and address registers. In the second mode, the word-length remains 48-bits, but the addressing is expanded to 27 bits. The third mode uses 64-bit operands and 27-bit effective addressing [Chai88, 25].

The El’brus-B is constructed using the same ECL chips used in the El’brus-2. By 1988 when the machine was first constructed, problems in the manufacture of these components had largely been ironed out; the El’brus-B is considered rather reliable. It runs at approximately five Mflops and has 64 Mbytes of main memory [Supe91b, 14; Usdi91, 30].

The El’brus-B passed state testing in 1988. On the order of 70 units have been manufactured at the Moscow SAM Plant since then [Supe91b, 13].

3.2.3.4 Expansion of the El’brus Family

As the El’brus-2 prototype neared completion, Babayan and others began work on the design of the next generation system. The new machines preserved many of the design principles of the El’brus-2: coarse-grain processors, modular construction, shared memory, multiple functional units, hardware tags, hardware support for high-level language constructs. One of the earlier El’brus design decisions which was to have a particularly profound influence on the new generation of systems was the use of a high-level programming language for all software, without exception. The lack of an assembler language made it possible to alter the underlying architecture significantly while maintaining

	Mico-El’brus	El’brus-E	El’brus-M14E	El’brus-3
Number of processors	1-10	2-8	2-8	1-16
Clock period (nsec)	125	40	25	10
Peak performance (processor)	8 MIPS	50 Mflops	160 Mflops	700 Mflops
Peak performance (configuration)	80 MIPS	400 Mflops	1280 Mflops	11 Gflops
Technology	CMOS	ECL	ECL	ECL
Architecture	RISC-like	VLIW	VLIW	VLIW
R&D facility	ITMVT	YERNIIMM	NIIVK	ITMVT

ITMVT	Institute of Precision Mechanics and Computer Technology, Moscow
NIIVK	Scientific Research Institute of Computing Systems, Moscow
YERNIIMM	Scientific Research Institute of Mathematical Machines, Yerevan

Table 3-1 Target Characteristics of Recent El’brus Computers
Source: [Baba89, 879]

software compatibility. While the El’brus-2 has a stack-based architecture with ‘‘as much control as possible’’ given to the hardware, the El’brus-3 takes a very-long-instruction-word approach in which decisions about scheduling are made not by the hardware, but by the compiler. Existing code, including systems software, would have to be re-compiled, but not altered.

This feature made it possible during the early 1980s to consider expanding the El’brus line to embrace a broad spectrum of systems, from microprocessor to high-end supercomputer. Inspired by similar efforts by DEC and IBM, El’brus engineers worked on a series of upwardly compatible systems whose target characteristics are shown in table 3-1.

The El’brus-3 is discussed in chapter 4. The two mid-range El’brus computers were developed at the Yerevan Scientific Research Institute of Mathematical Machines and the

Scientific Research Institute of Computing Systems, Moscow. Both these projects were carried out in close cooperation with ITMVT. The VLIW machines are all compatible at the level of the architecture, but each implementation was carried out independently. At the time of this writing, construction of an El'brus-3 prototype processor is nearing completion. The other projects have been terminated.

3.3 The Proliferation of Soviet HPC Efforts (1978-1985)

During the late 1970s and early 1980s the number of HPC development projects within industry and academia increased significantly. Strong push and pull forces drove this trend. Except for the AS-6, no new high-performance systems had entered series production since the BESM-6 in 1966/7. Military users had a growing need for general-purpose and specialized high-performance computing systems for the new generation of weapons. For national security reasons, they often refused to rely on Western systems for critical applications. Pushing the development of new systems were many researchers within the Academy of Sciences, the Ministry of Higher Education, and industry who were dissatisfied with the policy of copying Western work which characterized the ES mainframe and SM minicomputer lines. By the end of the 1970s several largely theoretical or paper designs had progressed to the point where building a prototype was a possibility. Prominent academicians and members of the Academy of Sciences such as G. I. Marchuk, V. M. Glushkov, V. A. Mel'nikov and others wanted to build new machines with their own, often radical, architectural ideas.

3.3.1 Industrial Projects

3.3.1.1 Attached Array Processors

During the late 1970s and early 1980s the Soviet, Bulgarian, and East German computer industries began manufacturing attached array processors which could be used in ES mainframe configurations for computationally intensive tasks. In the West, a number of generations of attached array processors were developed during the late 1960s and early 1970s, the most successful of which was the AP-120B by Floating-Point Systems, Inc., introduced in 1976 [Hock88, 30-32]. Traditionally the primary applications have been image processing and the analysis of seismological applications, and the attached array processors initially were hardware systems designed for the execution of fast Fourier transform and other algorithms. Soviet, Bulgarian, and East German industry followed this trend and introduced a number of models described in section 7.12, beginning in 1979.

The attached array processors were not truly ‘‘high performance,’’ since the most widely used models (ES-2335, ES-2345, MAMO-1M, ES-2706) offered only 5-60 MIPS performance on 32- or 38-bit data. In the absence of other computers such as the El’brus, however, they offered significant performance improvements over the general-purpose mainframes available to the scientific community. It would be incorrect to say that such systems were developed because of delays in the El’brus program, but it is likely that because the latter were not available, the attached array processors proliferated to a greater degree than would have been the case otherwise. Hundreds of attached array processors were manufactured in the Soviet Union and Eastern Bloc countries, primarily in Bulgaria. By far, most of the latter were sent to the Soviet Union [Ivan88; Prat90]. Most of the AAP were used in geophysics applications [Tcha92, 33; Niko82; Prat90].

3.3.1.2 The PS- series

One of the only high-performance lines developed outside of ITMVT to reach series production was the PS-2000 (*perestraivayemaya struktura*–reconfigurable architecture) designed by researchers at the Institute of Control Problems (IPU) in Moscow and built at the Scientific-Research Institute of Control Computing Systems (NIIUVM) in Severodonetsk, Ukraine. Both institutes were part of the Ministry of Instrument-Building, Means of Automation, and Control Systems (Minpribor). Although not the most powerful, the PS-2000 was among the most successful Soviet HPC systems in terms of development cycle length and number of units installed.

The PS-2000 and its successors, the PS-2100 and PS-2300 are discussed at length in chapter 5. The PS-2000 achieved high performance through high levels of parallelism. It had a a single-instruction multiple-data (SIMD) architecture in which an instruction is executed simultaneously on multiple pieces of data by multiple processing elements. Part of the inspiration for such an architecture came from the ILLIAC IV computer developed at the University of Illinois during the late 1960s. With the ability to modify the communications links between processing elements and dynamically turn individual processing elements on or off, the system could be used successfully in a variety of applications with high degrees of data parallelism. The PS-2000 was developed primarily for use within the Ministry of Geology and Ministry of the Oil and Gas Industry for seismic exploration. With up to 64 processing elements and a theoretical peak performance of 200 MIPS on 24-bit data, the machine offered 1-2 orders of magnitude more computing power for applications which did not require great precision (>24 bits) than was generally available to these users.

This system had a remarkably short R&D cycle, as shown in figure 3-3. The first prototype was built within four years from the time financing was secured and a decision to

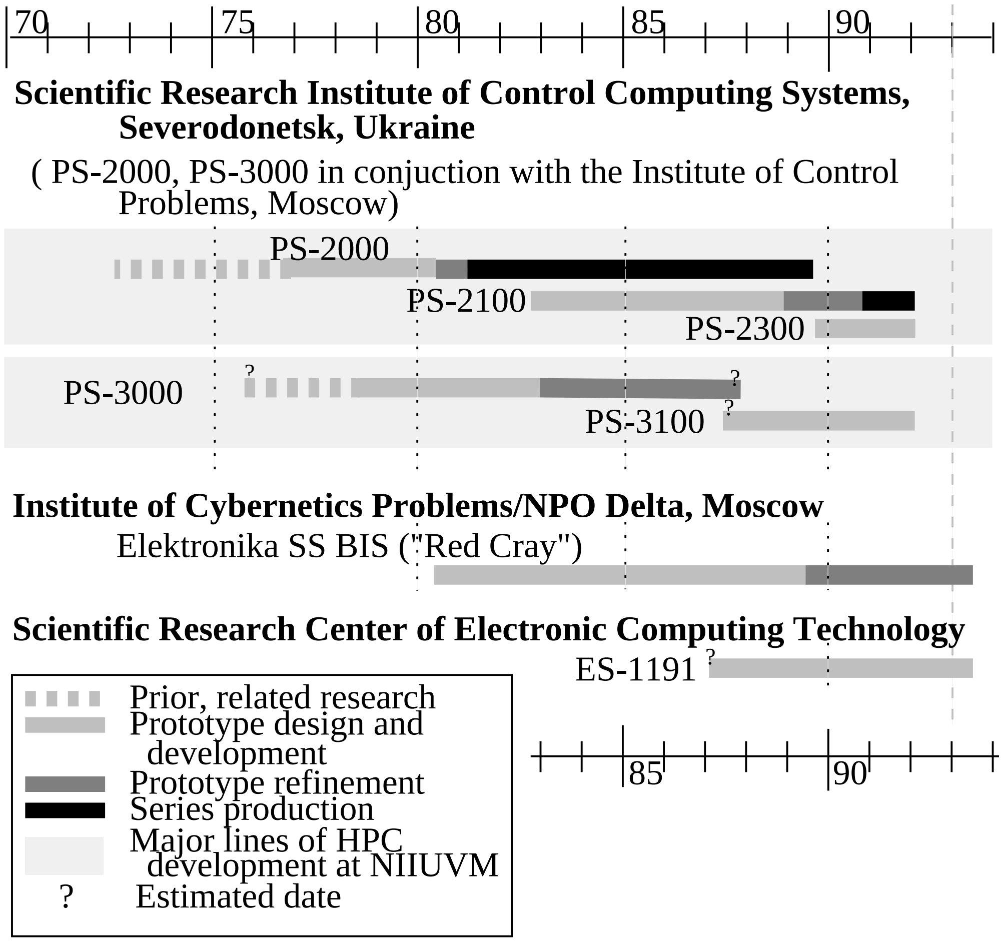

Figure 3-3 PS- and Other Industrial High-performance Computers

build the machine was made (1976). The machine entered series production only a year later and approximately 200 units were manufactured before production ended in 1989.

Several factors contributed to this success. The developers at NIIUVM had long been involved in developing systems for industry, and built the PS-2000 with a pragmatic eye for what could realistically be accomplished with the technologies at hand. Having many standard, identical pieces, the PS-2000 was also easy to manufacture, and relationships with the nearby factory were well established, thanks to administrative links in the Impul's NPO and prior development of control systems. The PS-2000 was not an industry driver. To speed development, engineers almost exclusively used components already in series production. The project also profited from a demanding timetable and considerable high-level support. While the PS-2000 was under development, the Soviet leadership began a massive campaign to develop the nation’s oil and gas resources. The PS-2000 was as well suited as any to address the computational needs of such a campaign, especially since during the late 1970s and 1980s American Presidents Carter and Reagan imposed technology embargoes, making it difficult or impossible for the oil and gas ministries to obtain and maintain Western computers [Mche81].

Thanks to stable target applications, requirements, and steady support for R&D from Minpribor and the Ministry of Geology, the PS-2100 is a very smooth and natural evolution of the PS-2000. The former incorporates up to ten times as many (640) processing elements arranged in modules of 64 PEs each, increased word-length (32- vs. 24-bit), more main and peripheral memory, etc. The theoretical peak performance of a full configuration is 1.5 GIPS. Primarily because of the need to develop a new generation of gate arrays for the processors, the PS-2100 took at least 50% longer to develop than the PS-2000. Approximately forty base modules (64 PEs each) were built before production ended in 1991.

The PS-3000 (described in section 7.5.1) was a parallel project also conducted jointly at IPU and NIIUVM. It was designed for use as the top level of complex hierarchical data processing and real-time control systems. Although it also has the PS- designation, the 32-bit PS-3000 has a very different architecture than the PS-2000. It consisted of two or four scalar processors, with each scalar processor associated with a vector processor. The machine incorporated two key features: a dynamic allocation of computing resources in which control units and processor fields were decoupled from each other (as opposed to the static coupling of traditional SIMD designs), and pipelining of processors. Because of the limitations of the available hardware, however the true pipeline processing was not implemented. The vector processors consisted of a number of processing elements which in many respects mimic the operation of a true vector-pipelined processor. The peak performance of a full configuration was 20 MIPS.

Although it used an available component base, the PS-3000 had a much longer development time than the PS-2000. Possible contributing factors were the need to rework the preliminary design in order to implement it using the available component base, a desire on the part of systems developers to work on a new generation system rather than push the system through into production, and low demand for the machine. Only about 10 systems were manufactured.

3.3.1.3 Elektronika SSBIS

The Elektronika SSBIS was the most direct Soviet response to the introduction of the Cray-1 in 1976. In a sector in which original design was the rule, the Elektronika SSBIS is an exception. Work on a machine patterned after the Cray-1 began in 1980 by a team led by V. A. Mel’nikov, the chief-engineer of the BESM-6. He had left ITMVT in 1978 to work at the Del’ta Scientific Production Association in the Ministry of the Electronics Industry. That he, a disciple of the Lebedev school, should agree to implement a non-indigenous design speaks strongly of the emphasis policy makers at this time placed on initiating an effort to follow in Cray’s footsteps [Sher92b, 1]. Nevertheless, as described in section 7.3, engineers did ultimately incorporate a number of distinctly non-Cray elements.

While the El’brus computers had some high-level characteristics in common with the Burroughs 700 Series computers, the underlying architecture has significant differences. The Elektronika SSBIS, on the other hand, was designed to be nearly compatible with the Cray at the level of assembler. The two machines are not binary compatible because designers lacked detailed descriptions of the Cray construction and desired to improve the implementation of individual instructions. The Elektronika SSBIS also differs in the number of functional units, the number of instruction buffers, and other implementation specifics.

The Elektronika SSBIS exhibits the characteristically long development time of industry-driving Soviet large-scale, advanced computing projects. The machine took nearly ten years to build; the first prototype was completed in 1989. The machine was based not only on a new technological base, but also a new infrastructure. This was the first high-performance computer developed within the Ministry of the Electronics Industry, and although components were a primary product of the ministry, considerable time and effort were needed to establish relationships with the scores of additional factories and enterprises needed to develop and manufacture everything from CAD systems to printed circuit-boards, cables, etc.

3.3.2 Academic Projects

As in the West, Soviet researchers have for decades viewed parallelism as a promising path of achieving high computational performance. Given perennial problems with the performance and reliability of the component base and subsystems, it was widely held that parallel systems, in principle, offered a means by which performance comparable with Western machines could be achieved using the slower, less reliable components of Soviet industry. Parallel systems were also attractive to the research community because of the rich research opportunities. The most significant Soviet parallel systems developed by the academic research community, shown in figure 3-4, cover a very broad spectrum of architectural approaches.

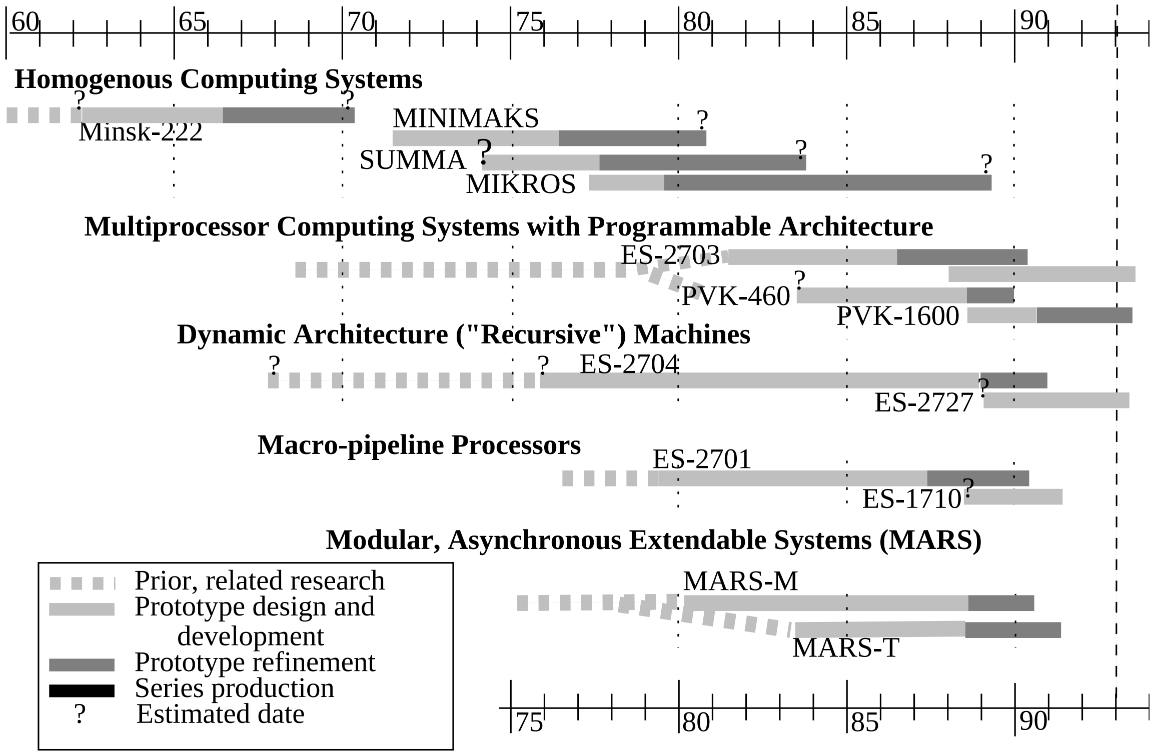

Figure 3-4 Academic Parallel Computing Systems

3.3.2.1 Homogeneous Computing Systems

Without question, the father of parallel computing in the Soviet Union was E. V. Yevreinov, who began research on parallel, multi-machine systems in 1960 at the Institute of Mathematics of the Siberian Department of the USSR Academy of Sciences. During the early 1960s he began working on combining individual computers into a single system to improve performance and reliability. He published his first work on this subject in 1962 [Yevr62].

Yevreinov pioneered the development of so-called homogeneous computing systems (OVS). This label covers a category of approaches to coarse-grain parallelism rather than a specific architecture. Tightly coupled parallel systems and geographically distributed systems can all be homogeneous computing systems [Yevr81]. All OVS share three characteristics [Dimi82, 84]:

- parallel computation;
- a programmable, or reconfigurable structure.
- a homogeneous (single-type) of computing element;

The OVS are, fundamentally, a collection of computing modules which are linked together via a regular, tunable interconnect network. The computing module can be as simple as a processor plus memory, or as complex as a complete general-purpose computer with processor, memory, communications channels, external memory, I/O devices, etc. As a rule, the computing modules are all of a single type. This leads to a shorter development cycle, ease of manufacture, and ease of maintenance. It also facilitates the inclusion of additional computing modules and the development of systems software [Dimi82, 89-90]. A so-called elementary machine consists of a computational portion, the computational module, and a system unit which managed the interconnection and interaction with other computational modules.

All of the OVS developed by Yevreinov or his colleagues at other institutes were manufactured on the basis of computers already in series production.

The ability to program the interconnect system allows the OVS to be tuned to match the structure of a particular task. Bus, pipeline, matrix, and hierarchical configurations could be established by software. This capability also allowed the system to be reconfigured in the event of a hardware or software failure [Dimi82, 90].

Yevreinov was one of the first to explore issues of distributed processing, interconnect protocols, distributed operating systems, and software development for distributed-memory hardware platforms. Many of the prominent Soviet parallel computing projects carried out during the 1970s and 1980s trace their roots, directly or indirectly, back to

Yevreinov’s work. A conference on homogeneous computing systems held in 1966 reportedly drew many of the individuals who would later lead parallel projects in programmable architecture systems, recursive architectures, reconfigurable systems, and others.

Yevreinov’s first running system, called the Minsk-222, was completed in 1966 at the Institute of Mathematics in Novosibirsk [Dimi82, 81]. It consisted of two Minsk-22 and one Minsk-2 computers. The former was a 37-bit machine with 4 Kwords of memory running at 5000-6000 operations per second which entered production in 1964. The Minsk-222 was used designed for controlling scientific experiments [Dimi82, 81; Apok74, 214]. Another version of this system was completed by 1972. It consisted of 1-18 so-called elementary machines (EM) physically linked in a ring. Each elementary machine had a processing rate of up to 4000 operations/sec and 32K words of associated memory. The EMs are linked by seven control lines which, under the control of a lead-EM, govern the interaction of all EMs [Dimi82, 175-193].

Three other systems, some with several versions, were developed at the Institute of Mathematics during the 1970s and early 1980s: the MINIMAKS, SUMMA, and MIKROS.

Work on the MINIMAKS was started in 1971 and a prototype was completed in 1976 [Golo80]. Research using this machine was conducted into the 1980s, however [Yush81]. The design allowed for 2-64 elementary machines. The machine could accommodate any computer of the Aggregate System of Computer Technology (ASVT) minicomputers developed at the Impul’s Scientific Production Association. This family of control computers was originally based on Hewlett-Packard minicomputers, but quickly evolved into an indigenous line. The M-6000, used in the MINIMAKS prototype entered series production in 1972, and had a performance of 200 KOPS. The MINIMAKS had two types of interconnect systems: one for data which linked each elementary machine with its four nearest neighbors, and a control line which linked each elementary machine with two neighbors. The interconnect system employed the communications channels of the ASVT minicomputers. Tuning MINIMAKS structure took place through altering tuning registers which indicated the ‘‘active links’’ for a given elementary machine [Dimi82, 193-217].

Built during the late 1970s by V. G. Khoroshevskiy, a colleague of Yevreinov’s, the SUMMA consisted of 1-10 elementary machines, based on the Elektronika-100I minicomputer which had a processing rate of 30 KOPS [Kash79; Sedu79; Golo80; Khor82; Khom87]. Each elementary module was linked with three neighbors via half-duplex communications lines. Unlike MINIMAKS, it used only one type of communication line for both control and data transmission. The greater stability and reduced need for real-time response of the target SUMMA applications made this possible. Information was passed via a packet-switched approach using either address or name identification for routing [Dimi82, 217-229]. The first version of the SUMMA was completed in 1977.

During the late 1970s and early 1980s, the number of organizations directly involved with Yevreinov and others in OVS development expanded. This was in large part due to the fact that Yevreinov worked at a number of different institutes during this period. The MIKROS (*MIKRoprotsessornaya Odnorodnaya Sistema*–Microprocessor Homogeneous System) was developed by researchers at the Institute of Mathematics in Novosibirsk, the Novosibirsk Electro-Technical Institute, and the All-Union State Design-Construction Institute (VGPTI) of the State Statistical Agency in Moscow. It was based on the Elektronika-60 computer, a minicomputer software compatible with the Western LSI-11 minicomputer. In contrast to the OVS mentioned above, MIKROS had a much more flexible structure. While elementary machines in the MINIMAKS and SUMMA were connected to a fixed number of neighbors, elementary machines in MIKROS could be connected with 1-12 others. The MIKROS could also be constructed as a geographically distributed system through adapters to telecommunications channels [Dimi82, 238-245]. This approach was used to build systems in other affiliated institutes, such as the Smolensk Branch of the VGPTI [Zuyk82]. The first MIKROS was completed in 1979, but work on it and others continued through the late 1980s. In 1989 a system with 24 elementary machines was being used as a platform for parallel systems software development [Mvs89, 6].

Other OVS were constructed at other organizations. Researchers at the Moscow Electro-Technical Institute for Communications built a universal communications computing system called the ‘‘Kollektiv,’’ designed for use in the control of communications systems [Dimi82, 232-235; Mamz84; Kudr84; Mamz89]. Also based on the Elektronika-60, this system used a ring-shaped trunck system channel containing 16 data and 13 control busses, making it possible to vary the number of computer modules (with full connectivity) from 2-16 [Yush81].[5] Based on the SUMMA architecture, the ME-80 was developed using Intel 8080 microprocessors at the Institute of High-Energy Physics in Dubna [Bald82]. The Statistika-1.0 OVS was developed at the All-Union Design-Engineering Institute of the State Statistical Association for use in an integrated system for information processing related to accounting, information services, planning, and management.[6] It also is based on the Elektronika-60. Its structure is basically hierarchical; Statistika-1.0 systems could be joined together in a hierarchical integrated system [Dimi82, 235-238].

Researchers in three other major lines of development–multiprocessor systems with programmable architecture, dynamic architecture (‘‘recursive’’) machines, and macro-

---

[5] As of late 1991, Yevreinov was reportedly employed here.
[6] By 1982, Yevreinov was working at this insitute [Ilyu82].

pipelined computers–view the homogeneous computing systems philosophy as an inspiration to their own projects. Each of these systems, to one degree or another, consists of a collection of identical processing elements with distributed memory and control. Although the PS-2000 is not considered an OVS because of its centralized control, the head of the laboratory (and later director) at the Institute of Control Problems where the PS-2000 was built, I. V. Prangishvili, had worked on OVS ideas during the late 1960s and early 1970s [Yevr74].

3.3.2.2 Multiprocessor Computing Systems with Programmable Architecture

The ‘‘multiprocessor computing systems with programmable architecture’’ developed at the Scientific Research Institute of Multiprocessor Computing Systems (NIIMVS) in Taganrog, Russia, under the direction of A. V. Kalyayev, were strongly inspired by the reconfigurable nature of Yevreinov’s homogeneous computing systems. This work is discussed in section 7.9. During the 1960s Kalyayev worked on building digital systems which modeled the functions of analog integrators. During the 1970s, he began working on more general-purpose systems based on so-called ‘‘macro-processors,’’ ‘‘macro-switches,’’ and ‘‘macro-memory’’ units. Underlying this work was the notion that the performance of a system is best when there is a good match between an algorithm’s control and data flows and the underlying hardware which executes it. The macro-switches provide a reconfigurable set of links which can be ‘‘tuned’’ dynamically to provide an interconnect system reflecting the nature of the problem being solved.

Kalyayev felt that a system should be user-oriented, i.e., should enable a user to operate the system in a language close to that used by practitioners in the applicataion domain. The MVS PA computers were designed primarily for mathematicians and designed so that the user would access a set of ‘‘macro-operations’’ (running on ‘‘macro-processors’’) which consisted of high-level mathematical operations (fast Fourier transform, matrix multiplication, integration, etc.). During the 1980s, researchers at Kalyayev's institute built a number of systems incorporating these principles and designed and manufactured a number of processor, switching, and memory components to support their architectures.

3.3.2.3 Dynamic Architecture Machines

V. A. Torgashev and others worked on building dynamic (‘‘recursive’’) architecture systems, first at the Leningrad Institute of Aviation Instrument-Building, and later at the Leningrad Institute of Informatics and Automation. These ideas were first widely publicized at the International Federation of Information Processing IFIP ’74 conference in Stockholm [Glus74]. These computers fall into a category of non-von Neumann architectures called data- or demand-driven. In data-driven (data flow) machines, the availability of operands triggers the execution of an operation. In demand-driven (reduction) machines, a request for a result prompts the execution (possibly recursively) of the operation which will generate it. In principle, such systems can ‘‘seek out’’ parallelism inherent in a program but not recognized by a programmer, improving performance. Torgashev's systems are discussed in section 7.7. In the dynamic architecture systems, a program is represented as a set of automata which represent data, operations, relations, references, or physical machine resources. Automata can be combined into more complex automata. A problem, represented by a dynamic automata network, its executed by applying transformations to the network until a final state is achieved which represents the solution. During the 1980s Torgashev developed a prototype system of which several units were built.

3.3.2.4 Macro-pipelined Machines

Researchers at the Institute of Cybernetics in Kiev worked on so-called ‘‘macro-pipeline’’ computers described in section 7.8. This approach incorporates a number of coarse-grain processors which can operate in a systolic, pipelined fashion in which the data for an algorithm propagate across the processors, with each processor performing a significant amount of computation before sending the results to a neighboring processor. The programming language associated with such a machine, MAYaK, was designed to allow traditional, sequential programs to be converted to parallel systems. By packaging existing code with MAYaK instructions specifying inter-module communication, the program could be broken into pieces which could run in parallel and communicate with each other on multiple processors. A number of prototype systems were manufactured during the 1980s.

3.3.2.5 The ES-270x Systems

The basic research for the three systems just mentioned was conducted for years before full prototype systems were constructed. Carried out in academic institutes with little or no manufacturing capability, such systems could not be constructed without significant help from industry. Around 1980, individuals at Scientific Research Center for Electronic Computer Technology (NITsEVT) began an effort to provide research teams with ES mainframe technology and development resources to build prototype implementations of their ideas. Each of the projects had concentrated on demonstrating new ideas for constructing a computational engine. Researchers had designed the systems not as complete, industrial systems, but primarily to demonstrate new possibilities for carrying out parallel computation. NITsEVT agreed to provide both the technology to construct processor prototypes, as well as that which was lacking to enable the processors to perform useful work. As a result, the above systems were built on the basis of ES mainframe technology and were linked by standard ES channels to a mainframe host which provided needed I/O, systems management, and user interaction facilities.

Although the projects were independent of each other, they each received the designation ES-270x, indicating that they were a processor attached to an ES mainframe system: ES-2700 (attached-array processor), ES-2701 (macro-pipelined system), ES-2702 (symbolic processor at the Institute of Applied Mathematics, Moscow), ES-2703 (multiprocessor computing system with programmable architecture), ES-2704 (dynamic architecture machine), ES-2705 (analog parallel system, Riga Polytechnical Institute). Another system fitting into this naming scheme, the ES-2706, was a Bulgarian attached array processor and not directly supported by NITsEVT. NITsEVT coordinated efforts to push proposals through policy-making and funding organizations. During the early 1980s a series of resolutions of the USSR Council of Ministers provided the necessary support for the various development stages of the ES-270x machines. Such efforts were in keeping with some of the stated goals of the 11th Five Year Plan [Uprs81; Myas82].

NITsEVT committed to supporting these projects through the prototype stage. Production of the systems depended on finding production facilities willing to manufacture the systems, and gathering enough political support to overcome any resistance by factories to introducing new products. Both depended on the existence of customers who would purchase the machines once they were manufactured; having adequate financing for R&D was not sufficient. Factories were also interested in the ease with which machines could be manufactured using the technology available.

Efforts to arrange for series productions of the machines were ultimately unsuccessful. Although some (single digits) of potential customers were found for some of the machines, the market was not large enough to make it worth factories' while to manufacture them. From the users' perspective, there were fundamental problems with each of the systems. Lacking their own I/O facilities and systems management capabilities, the systems had been built out of necessity as attachments to standard ES mainframe hosts.

These configurations were generally not able to supply data to the processors at a rate that would support the claimed performance levels. Plans were made for giving the processors direct access to main and peripheral storage of the host computers, but the prototypes which were constructed did not have such features. Second, with the exception of the ES-2701, the systems required significant re-coding in non-standard languages. Third, although manufactured with NITsEVT support, the projects varied in the degree to which they conformed to accepted construction practices at the factories. The greater the number of different kinds of circuits and modules, the greater the factory’s difficulty in manufacturing the system. Reportedly, the ES-2704 had a relatively small number of different types of modules, while the others were more complex, and therefore less desirable from the perspective of a factory.

3.3.2.6 MARS

The Modular, Asynchronous, Extendable Systems (MARS) were developed at the Computing Center of the Siberian Department of the USSR Academy of Sciences. The conception for these systems were developed by G. I. Marchuk and V. Ye. Kotov and published in 1978. Formerly the director of the Computing Center, Marchuk at that time was President of the Siberian Department. Later, as the chairman of the State Committee of Science and Technology, he was able to provide considerable high-level support for the implementation of the MARS ideas.

In their analysis of computing trends during the 1970s, Kotov and Marchuk identified as key the broadening of the sphere of application of computer technology, the compound or systemic nature of the problems to be solved, the transition from traditional von Neumann machines to computing systems with a variety of configurations, capabilities, and purposes, the miniaturization of the component base, and the trend towards higher-level man-machine interfaces.

The key architectural principles of the MARS are parallelism, decentralized information processing and data flow, asynchronous interaction of devices and processes, a hierarchical structure, specialized systems components and hardware implementation of complex data processing functions, modularity, reconfigurability, and self-identification of data and processes.

Like the academic systems described above, the implementation of the MARS conceptions required the assistance of industry. While the ES-270x systems were supported by NITsEVT, MARS was supported by ITMVT, thanks in part to the close relationship between Marchuk and V. S. Burtsev, ITMVT director from 1973-1984. Marchuk was also able to secure funding through the GKNT.

Two systems embodying MARS ideas to varying degrees were developed between 1980 and 1990. The first of these, called the Mini-MARS (later, the MARS-M) was a shared-memory heterogeneous multiprocessor incorporating a variety of advanced architectural ideas including decoupled architecture, very-long-instruction-word (VLIW), and virtual heterogeneous multiprocessors. The second, called MARS-T, was a multiprocessor with computational elements patterned after the transputer and the stack-based Lilith processor developed by Niklaus Wirth (author of Pascal and Modula-2).

Between 1985 and 1988, these projects and others were supported through the so-called START project. In 1981 the Japanese had launched the Fifth Generation Project, which placed artificial intelligence at the core of a program oriented towards the development of a new generation of computers. Soviet researchers and policy makers began discussing the need to launch a Soviet response. Several prominent AI researchers joined with Kotov to push an AI/hardware/software program through policy making channels, and found it useful to cast their work as a response to the Japanese efforts. Thanks in large part to the help of G. I. Marchuk, these efforts were successful and the three-year

START program was created. The START program, and the MARS computers more generally are the subject of chapter 6.

Prototypes of the MARS-M and MARS-T were built by the conclusion of the START project in 1988. In spite of the high levels of government support, the high-profile nature of the project, and favorable organizational conditions, the projects still suffered from the administrative gap between industry and the Academy of Sciences, and the nature of supplies and operations at industrial factories through the mid-1980s. During the late 1980s, support for MARS research and demand for such computers dwindled. Work on hardware development ended by 1991.

3.4 1985-present

Soviet HPC developments during the latter half of the 1980s took place within the context of a growing gap in computing power and availability between the East and the West; established HPC projects initiated during the 1970s and early 1980s; and major, disruptive efforts to reform Soviet society and the economy. Technological advance in the West during the early 1980s was relentless, with firms like Cray and Control Data Corporation introducing new models every 2-4 years with double or triple the performance of their predecessors.

3.4.1 Policy-making Developments

From the perspective of high-performance computing, key policy-making events were the creation in 1983 of the Department of Computer Technology, Informatics, and Automation (OIVTA) of the USSR Academy of Sciences, and the Comprehensive Program of Scientific-Technical Progress of Member Countries of the Council of Mutual Economic Assistance to the Year 2000 (Program to the Year 2000), adopted in December, 1985 [Sama85; Prav851219]. The latter called for broad cooperative efforts in five basic categories: ‘‘electronification’’ of the national economy, comprehensive automation of manufacturing, nuclear energy, bio-technology, and development of new technology for the production and processing of new materials. The goals of ‘‘electronification’’ included the development of supercomputers having speeds of more than one billion operations per second, large numbers of personal computers and software, the development of a unified systems for transmitting digital information, the development of new satellite communications and television systems, and the development of new electronic devices and fast, reliable integrated circuits [Prav851219]. The GKNT was given over-all responsibility for organizing and coordinating under this program [Byal90, 19].

The Academy of Sciences had lost much influence in establishing computing policy during the late 1960s and early 1970s when the ES and SM programs were implemented. With the creation of the OIVTA the Academy sought to re-assert its role as a consultative body in the field of computing and consolidate and coordinate informatics research within the Academy by providing an official forum in which issues related to computer technology could be discussed and policy formulated [Mikh84].

The OIVTA played a leading role in crafting the Program to the Year 2000 [Guse83; Yasm85]. Several individuals prominent in Soviet HPC were members of the OIVTA: G. G. Ryabov, B. A. Babayan, V. S. Mikhalevich (director of the Institute of Cybernetics in Kiev), V. P. Ivannikov, V. Ye. Kotov, and A. V. Kalyayev [Alek84; Izv841224]. The Institute of Cybernetics Problems was created as part of the OIVTA and became home to the Elektronika SSBIS project. One of the main goals of the new department was the creation of supercomputers and, thanks to OIVTA efforts, the Program to the Year 2000 included plans for the ‘‘development of new supercomputers having speed of more than one billion operations per second...’’ and a machine with a performance of 10 billion operations per second by 1990 [Prav851219; Veli85; Sama85; Veli87; Marc87].

Several Soviet high-performance computing projects received financing under the umbrella of this program from a variety of sources such as the State Committee on Science and Technology, the Academy of Sciences, and individual Ministries and enterprises. Among the projects targeting the one billion-plus operations per second goal were the MKP and El’brus-3 of ITMVT, the ES-1191 of NITsEVT, the PVK-1600 and others at NIIMVS, ES-2727 successor to the ES-2704, and the ES-1710 successor to the ES-2701 [Przh89, 36-37]. The MARS (START) program, with its emphasis on new computer architecture and artificial intelligence also benefited from the Program to the Year 2000’s emphasis [Koto87].

The Program to the Year 2000 was designed to be an international effort, and some lip-service was paid to cooperation with other countries in HPC R&D. For example, [Ivan87] reports participation by Hungarian and Polish specialists in the creation of high-level artificial intelligence languages and power supplies for the El’brus-3 and MKP. In practice, however, whatever international contact there may have been in these and other high-performance projects, they were at best minimal and did not have a significant impact on the technology. Soviet high-performance projects remained almost exclusively Soviet.

3.4.2 Developments in High-Performance Computing Systems

Although it provided a vehicle for channeling development funds to the high-performance computing sector, the Program to the Year 2000 did not qualitatively change the landscape of Soviet high-performance computing. The technological developments in high-performance computing during the latter half of the 1980s are best characterized as a logical extention of lines of development begun during the 1970s and early 1980s in which basic design philosophies were preserved.

The ES-1191 (figure 3-3) and the loosely-coupled array processor configurations were the only ‘‘new entries’’ in the Soviet HPC sector during the latter half of the 1980s. Other projects had either been initiated previously, or were evolutionary extensions to previous projects. A number of projects, including the ES-1068.17 and ‘‘Sibir’’’ systems described in section 7.12.6, consisted of a mainframe host with a number of attached array processors. The ES-1191 represents the first direct efforts of the Scientific-Research Center of Electronic Computing Technology (NITsEVT) to develop a high-performance computing system. NITsEVT had assisted with the ES-270x projects, but until the mid-1980s had concentrated its own development efforts on building high-end mainframe systems.

The ES-1191 was perhaps inspired by IBM’s efforts to enter the supercomputer market through the introduction in 1985 of its 3090 VF, a 3090 mainframe with one or more integrated vector processors. The ES-1191 also integrates vector processors into a system based on a general-purpose mainframe, but the design differs considerably from the IBM machines. Seeking not only to break the 1 Gflops performance barrier, but also provide high average processing rates, preserve the systems software, and clearly separate the computationally intensive tasks from the routine system management tasks, engineers designed a machine with a computational subsystem including both scalar and vector resources. The scalar processors treat the vector processors as a shared resource. This machine is described further in section 7.4. The ES-1191, originally scheduled for completion in 1989 experienced delays and reductions in financing. Nominal work continues on the system, but the prospects for completion are very poor.

In spite of the inability to secure production facilities, R&D continued into the early 1990s for successors to the original ES-270x systems. Support for the MKP, El’brus-3, and Elektronika-SSBIS remained constant, if not necessarily adequate. The latter had production facilities available, but the relationships with the myriad of upstream factories became more tenuous as economic relationships were transformed from a centralized command mode to horizontal links based on mutual benefit. The system of state orders *(goszakaz)* which replaced the centralized Plan as an agent of centralized control of production helped maintain these links, but the state orders grew increasing unable to compensate for basic production shortfalls, discontinuities in the supply chains, and growing local autonomy. The deteriorating financial state of nearly all Soviet organizations, their *khozraschet*-related tendency to manage their money more tightly, and the highly uncertain government budget allocations devastated the market for HPC.

As work progressed during the late 1980s, the Soviet economy deteriorated, the infrastructure necessary for building new machines began to fracture, and the market evaporated. In 1989 the Eastern Bloc dissolved and the Program to the Year 2000 became irrelevant. The system of Five Year Plans ended in 1990 and large-scale ‘‘goal-directed programs’’ for the advancement of one branch of the economy were largely phased out as tools of economic management. Trade relationships with the Eastern European countries were severely disrupted with the conversion to trade based on world prices at the beginning of 1991 [Izv910720; Newm91; Seme92]. In particular, the supply of attached array processors from Bulgaria and the former East Germany ended as prices started being computed in hard-currency terms, and Soviet users lacked, or were unwilling to pay, the necessary amount of money. Financial support for high-performance computing became more fragmented, dependent on the agendas and abilities of individual government funding organizations and HPC customers.

3.4.3 Soviet Computing Associations

In response to the growing crisis, developers, manufacturers, and users of high-performance computers formed two associations to promote their own interests. Leaders at ITMVT, the Scientific-Research Computing Center at Moscow State University, and other organizations spearheaded the creation of the ‘‘Supercomputer Association of Users, Developers, and Manufacturers of High-Performance Systems,’’ officially organized in January, 1991. This association, described at greater length in section 4.8.3, was created to draw together the various players in the HPC sector to encourage closer cooperation, and serve as a mouthpiece and lobbying force for the HPC sector to help secure government support. The Association held its first conference on the ‘‘Problems of the Development of High-Performance Computing Systems’’ in October, 1991 to coordinate efforts, and discuss issues related to the basic survival of the industry. Although the association has been so strapped for funds that its activities in 1992 were minimal, its efforts in 1991 did play a role in securing basic funding for all prominent on-going HPC projects during 1992.

In 1990, the Soviet Transputer Association was formed [Usdi91, 30].[7] The name of the association is almost a misnomer, since it includes not only organizations and individuals using Western transputers, but also Soviet organizations working on distributed memory multiprocessor systems. Most of the parallel projects discussed above which trace their inspiration back to Yevreinov’s homogeneous computing systems (OVS) fall into this category: Torgashev’s dynamic architecture machines, the macro-pipeline systems at the Institute of Cybernetics, Kiev, the multiprocessor computing systems with programmable architecture in Taganrog, Khoroshevskiy’s work on homogeneous computing systems in Novosibirsk, and others. Each of these groups is a member of the Russian Transputer Association, and each has begun working to port their systems software to a transputer-based hardware platform [Niim90, 7-8]. Individuals at the Scientific Re-

---

[7] Following the breakup of the Soviet Union, the association was renamed the Russian Transputer Association.

search Institute ‘‘Kvant’’ in Zelenograd built a three-chip transputer-like processor [Niik92; Korn92]. Researchers at the Institute of Applied Mathematics and elsewhere use transputer platforms to model other computer architectures, develop parallel algorithms, and perform complex data processing tasks [Zabr91].

In spite of the broad differences in computer architectures, the multiprocessor, distributed memory approaches taken by these groups are not altogether incompatible with the basic ideas of transputer construction. In fact, the transputers simply provide a hardware platform of basic processors together with integrated communications facilities. With the appropriate systems software, the basic multiprocessor hardware can be used to model a variety of other architectures. While the resulting system might not be as efficient as a customized hardware system, transputers provide basic facilities that free researchers who want to conduct research in parallel systems, algorithms, and applications from having to expend much time and resources on building the hardware itself. In the Soviet context, as trade (legal and illegal) with the West in computer components grows and the prospects for enlisting computer factories in the development of new hardware systems were worsening, transputers can provide an alternative to many Soviet scientists.

The Russian Transputer Association reportedly has over two hundred members. It was formed on the initiative of V. K. Levin, the director of NII Kvant, in part in response to a government program to promote transputer-related research within the Soviet Union [Mvs89, 58]. The association does not finance research, but works to facilitate contacts with the Western transputer community and the flow of information among individuals involved in Soviet transputer-related research. Several international conferences organized by the RTA have been held in Russia and RTA members have traveled abroad to transputer conferences. The association also provided expert advise to policy makers who had to evaluate and fund research proposals in computing.