



Russian Academy of Sciences Program Systems Institute

ServNet: service network of the SKIF family supercomputers

Objectives

Creation of a simple and reliable scheme realizing remote low-level management functions of the cluster computing nodes.

ServNet advantages

- scalability;
- simplicity and reliability;
- ease of assemblage and adjustment;
- technological effectiveness;
- low cost;

Constructive solutions

- single-sided board;
- one Atmel microcontroller;
- two Maxim integrated circuits (MAX232A and MAX487);
- one microchip with ferroelectric sequential RAM (FRAM) RAMTRON;
- small quantity of auxiliary components

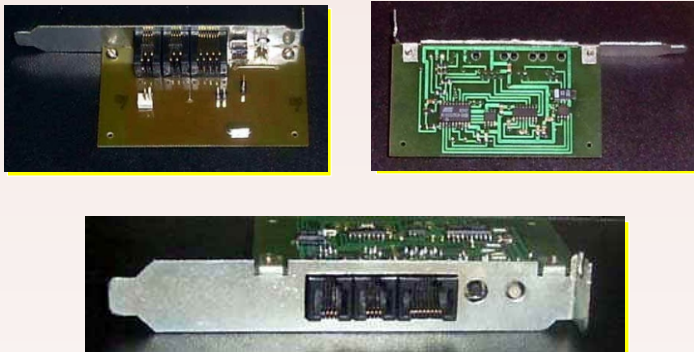


Fig. 1 Boards

Functionality

Abovementioned possibility of work with Linux serial console of the computing node and possibility of remote work in the BIOS Setup mode. Possibility to control the boot loader (LILO) of the computing node. LILO control can be configured onto the serial console and if several different operating systems are installed on the nodes it is possible to choose from the control station the type of operating system loaded on each node. So, a particular OS can be loaded (of those pre-installed on the node). on the whole cluster (or on part of the nodes) Possibility to change Linux kernel loading parameters on every node.

The possibility of "posthumous" viewing of a few (up to 4 Kbytes) last lines sent out onto the serial console. In case of the computing system hang with inoperability of

the cluster system network and the auxiliary network the service network adapter holds information about the state of the system before the hang in its nonvolatile memory. Thus, it is possible to recover from the control station the picture of the last "moments of life" of the computing system and understand the cause of the failure.

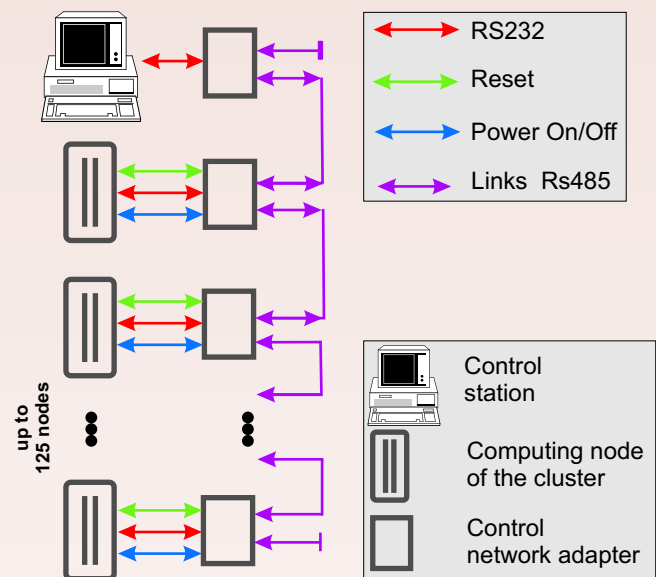


Fig. 2. Block diagram of the control network of the cluster

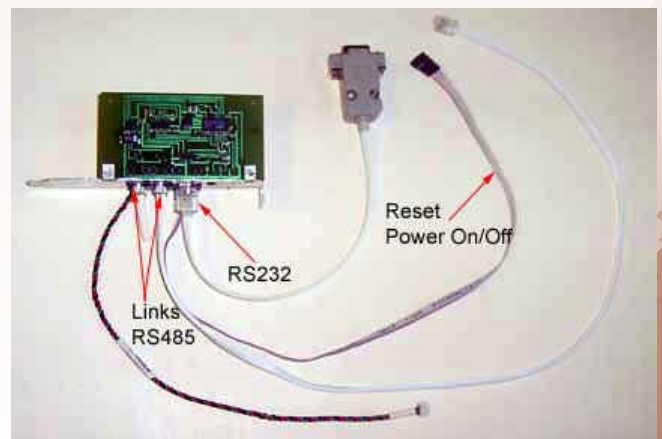


Fig. 3. Board cable connection diagram

ServNet: service network of the SKIF family supercomputers

Cluster management

- Control network adapter connected to the control station ("head adapter") functions as a master device on the bus. All the other adapters (service adapters) receive the right to transfer data to the bus only by the command of the head adapter. Before receiving command from the head adapter the adapter transmitters are in high-impedance state.
- Adapter is connected to the control station through the RS232 interface.
- Adapters are connected to the computing nodes through the RS232 interface as well. Besides, special adapter connectors are linked to the node allowing node power on/off and hardware reset operations.
- All service adapters are addressed. The procedure of assigning addresses to adapters is executed from the control station at a final stage of assembling the cluster.

Access to the serial console provides:

- Changing BIOS node parameters;
- LILO: choice of OS to be loaded, parameters of Linux kernel load;
- Any commands in the console mode;
- Monitoring of the OS critical messages;
- Posthumous reading (from the nonvolatile memory of the ServNET board) of a few last console messages.

Production development

- **Development** — PSI of RAS. **Production** — Computer Research Institute "NII EVM" (Minsk).
- **Startup production lot has been manufactured** (more than 200 items). ServNET is installed in:
 - T-Forge32 (16 nodes).
 - SKIF K-500 (64 nodes) and other clusters in Minsk.
 - Clusters in Pereslavl.
 - Clusters in RCC MSU (Research Computing Centre) (58 ServNET boards).
- **ServNET v.2: 2004, development of PSI of RAS and T-Platforms, production by T-Platforms** (Dimensions (66 x 33 mm) compared to the first version were reduced by half).

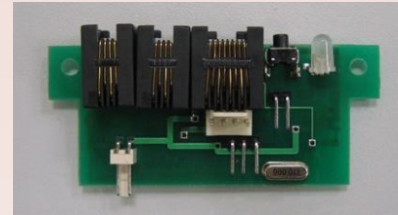


Fig. 4. ServNet board, ver.2, side "a"

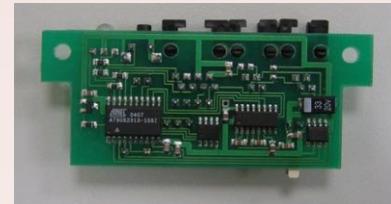


Fig. 5. ServNet board, ver.2, side "b"

Major features of the control network of the SKIF cluster

Interface for interaction between the adapter and the node:	Rs232
Data exchange rate with the node:	9600 bits/sec
Interface for interaction between the adapter and the control station network bus:	RS-485
Rate of RS-485 bus exchange	115200 bits/sec
Maximum possible rate of RS-485 bus exchange:	250000 bits/sec
Maximum possible length of the control station bus at the rate of bus exchange of 115200 bits/sec.:	1000 m
Maximum possible length of the control station bus at the rate of bus exchange of 57600 bits/sec:	1200 m
Maximum possible length of the control station bus at maximum rate of exchange	500 m
Maximum number of adapters on the control station bus:	127
Consumption current at the control station adapter:	not more than 150 mA