Russian Academy of Sciences
Program Systems Institute

ServNet-3: Service net of series 3 supercomputers of SKIF family

The aim of the development is to establish a simple and reliable system, implementing functions of a remote low-level management of computing nodes and other subsystems of Series 3 supercomputers of the SKIF family.

Functionality of ServNet-3 service net

Series 3 supercomputers of the SKIF family have been built using blade computing modules and chassis in which they are installed.

ServNet-3 service net monitors all blade computing nodes and fan modules of all the chassis and supports:

- the following operations with any (randomly chosen) computing node of the supercomputer:
  - switching on/off;
  - status poll switched on, switched off and so forth;
  - serial console session discussed in detail below;
  - voltage measurements in the power supply and temperature system (at the point where the ServNet T60 card is mounted);
- the following operations with any (randomly chosen) fan of any chassis in the supercomputer:
  - fan rotation speed measurement;
  - fan rotation speed management.

Further development of monitoring and control functions of the ServNet-3 service net is provided for.

ServNet-3 service net architecture

The ServNet-3 service net has a two-level architecture:

- **ServNet bus**. Inside the mounting cabinet the monitoring and control boards of the ServNet-3 net are interconnected with the ServNet bus that incorporates RS485 and power supply (12 volt). Data transfer rate is 625 kbit/s. The ServNet bus can cover up to 11 chassis (up to 127 devices connected to the bus).
- **TCP/IP net, Ethernet 10/100BaseTX**. The ServNet-3 service net segments assembled in the framework of the supercomputer mounting cabinets are interconnected into a TCP/IP net with the help of Ethernet 10/100BaseTX. Access to the same net is provided for control computers.

This allows scaling the ServNet-3 net for supercomputers of any size.

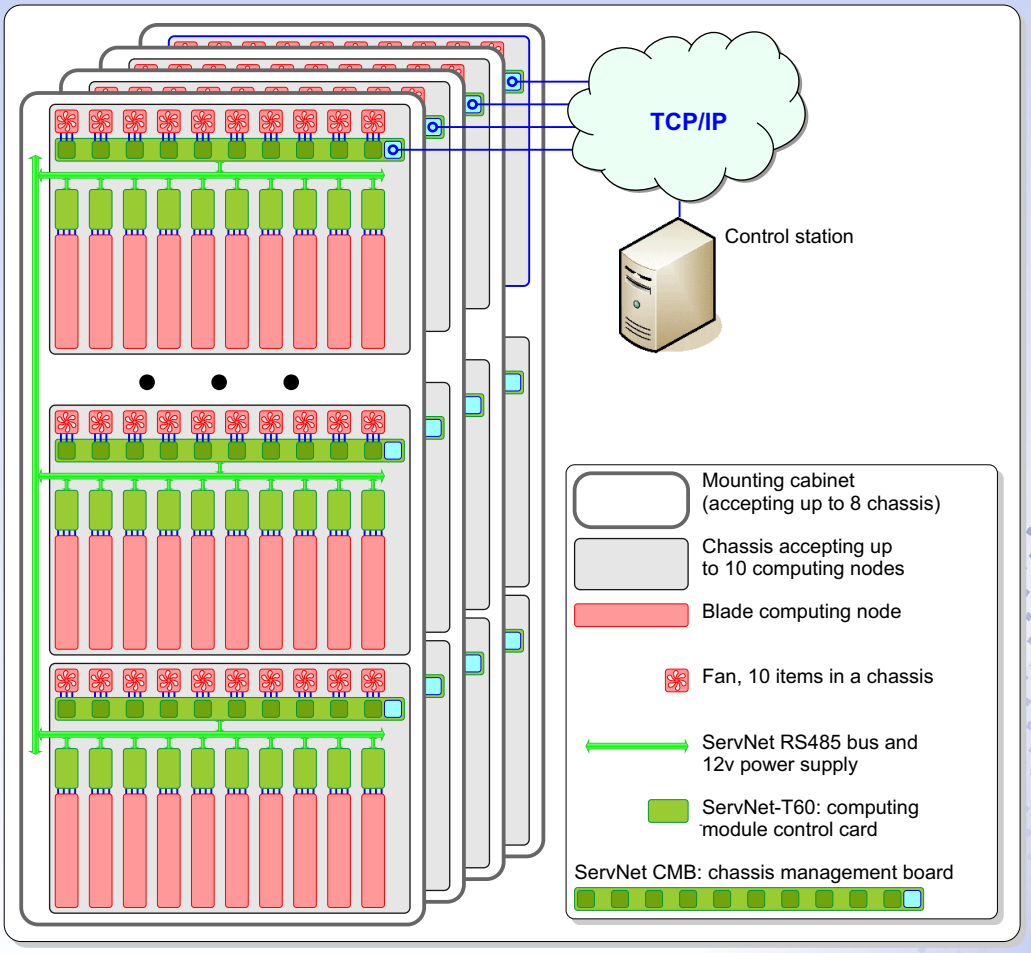

Fig. 1. Block diagram of the service net for Series 3 supercomputers of the SKIF family.

Russian Academy of Sciences
Program Systems Institute

ServNet-T60 card: a blade computing node control card

Every computing module of Series 3 supercomputers of the SKIF family contains a ServNet-T60 card connected to the motherboard of the node and to the ServNet bus of the mounting cabinet. The ServNet-T60 card has:

- **a ServNet bus interface** for receiving and monitoring data, control commands and data transferred in the framework of sessions of work with the serial console;
- **Reset signal** to reboot the motherboard;
- **On/off signal** to switch power of the motherboard on and off;
- **Info signals** for receiving information about the state of the motherboard (if it is switched on/off and so on);
- **RS232 interface** to realize access to the motherboard in the serial console mode.

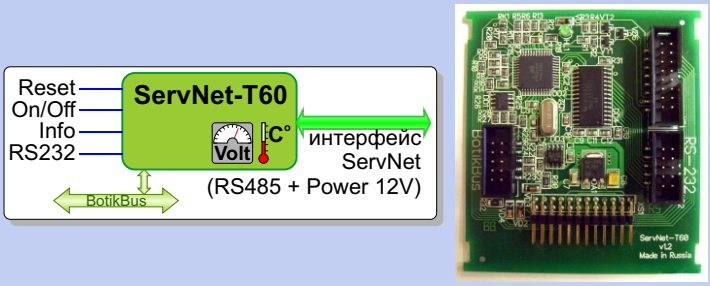

Fig. 2. ServNet T60: Blade computing node control card, 80×70 mm

There is a nonvolatile memory on the ServNet-T60 card in which the last lines (up to 32 Kbytes) brought out onto the serial console are stored. Owing to this the possibility of 'posthumous' viewing those lines is realized.

Further, local temperature sensors are implemented on the ServNet-T60 card, as well as supply voltage sensing devices and the BotikBus interface that is provided for further developing monitoring functions of the ServNet service net. Development of smaller modules that can be connected to the BotikBus interface is planned. They allow to read information from the control bus of the motherboard or to connect to the system various auxiliary sensors (to read temperature, air flow speeds, unauthorized access and so forth).

ServNet CMB: chassis fans control card

The ServNet CMB card enables measuring rotation speed of each fan in the chassis and regulate it. This is important since if temperature conditions permit reducing speed at least by 30%, then electro-consumption of the fans falls by half.

The ServNet CMB (Chassis Management Board) includes:

- the ServNet bus interface for transceiving monitoring data and control commands;
- ten fan control units (on the base of the AVR Tiny24 microcontroller) to which 10 fans that are in the chassis are connected;
- one EtherBox32 module that realizes Ethernet 10/100BaseTX interface. In the mounting cabinet the ServNet CMB of one of the chassis is connected to the TCP/IP net uniting all the ServNet-3 net segments.

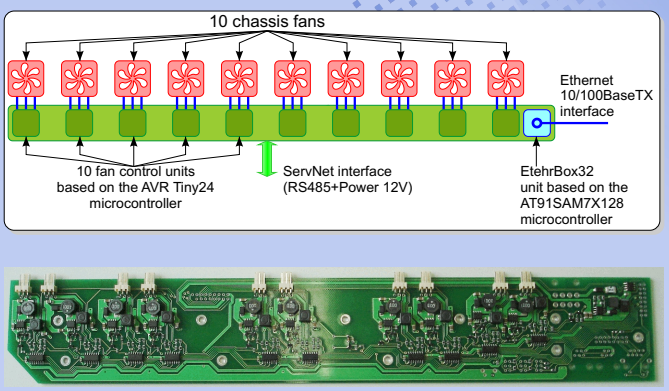

Fig. 3. ServNet CMB: chassis fan management board, 372×60 mm.

Support for sessions of the serial console in ServNet-3

Support for sessions of the serial console with an arbitrary supercomputer computing node implements:

- remote work with a node in the BIOS Setup mode;
- boot loader control (LILO) of the computing node;
- the possibility of "posthumous" viewing of a few (up to 32 Kbytes) last lines brought out onto the serial console.

The last two possibilities are considered in more detail.

Boot loader control (LILO) of the computing node. If there have been installed a few various operating systems on the node and LILO control can be configured onto the serial console, then from the control station the type of operating system loaded on every node can be chosen, and thus in the entire cluster (or on some of the nodes) this or that OS (of those pre-installed on the node) can be loaded. Boot-up parameters of the Linux core can also be changed in every node.

The possibility of "posthumous" viewing of a few (up to 32 Kbytes) of the last lines brought out onto the serial console. When there is an error or failure of a computing node or of the entire supercomputer (even in cases when neither the system network of supercomputers nor the auxiliary one works) the last lines brought out onto the serial console remain in the nonvolatile memory of the ServNet-T60 card. As a rule, they hold information about the state of the system before the error and from the control station it is possible to recover the picture of the latest developments and understand the cause of the error.

Pereslavl-Zalessky Yaroslavl Region Russia, 152020
Tel/Fax: +7 (48535) 98064 E-mail: abram@botik.ru Web-site: http://www.botik.ru/PSI

Pereslavl-Zalessky, 2008