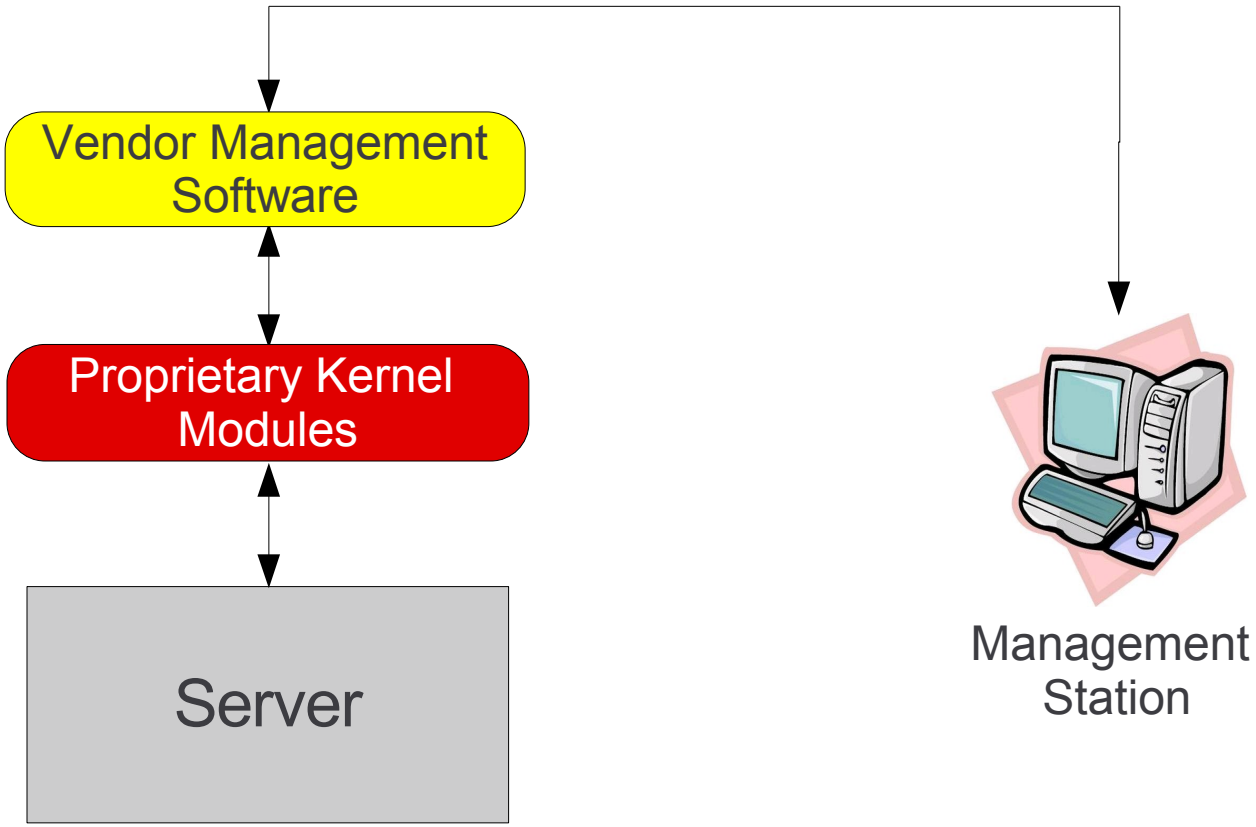
Server Management and Monitoring with IPMI

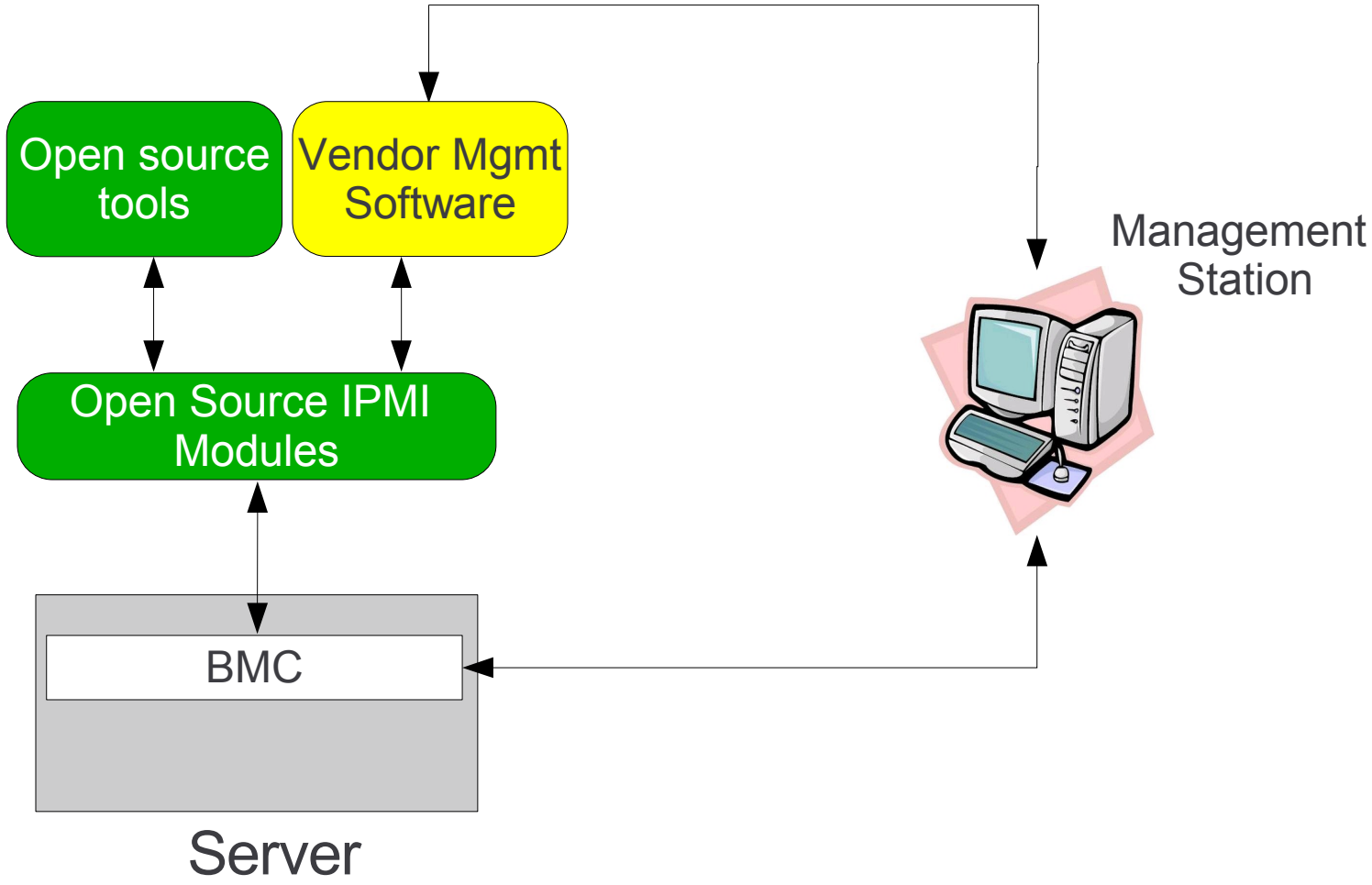
Amit Bhutani
Software Engineer
Linux Engineering
Dell, Inc.
Amit_Bhutani@Dell.com

Server Management – In the Past



Proprietary management solution, varies among vendors

Current Server Management with IPMI



IPMI - Intelligent Platform Management Interface

- Open-standard hardware manageability interface specification developed by Dell/Intel/IBM/HP
- Provide common interface for accessing system components:
  - Environmental sensors (temperature, voltage)
  - Chassis power control and identification
  - System Event Logs
  - Watchdog timer (to send alerts when OS crashes)
- Asynchronous, message based interface
- Cross platform support, standard IPMI utilities will manage different vendor hardware

Intelligent Platform Management Interface (IPMI)

- IPMI v1.5
  - supports remote connection via NIC and serial over LAN
  - IPMI 1.5 used Remote Control Management Protocol (RCMP) - Basic authentication only
- **IPMI v2.0**
  - Backwards compatible with v1.5
  - Enhanced security authentication and encryption for network packets
    - Incorporates authentication based on SHA-1 (Secure Hash Algorithm-1) and supports AES (Advanced Encryption Standard)
    - Uses RCMP+ (lanplus) - more robust authentication + encryption

BMC – Baseboard Management Controller

- The BMC is the heart of an IPMI based system
- Has its own firmware, power, NIC connection
- Physical NIC port may be shared with host system, BMC has separate MAC address, configured with static or DHCP IP address
- Circuitry monitors all sensors in system, generates events to SEL
- Replaces previous sensor monitoring methods such as smbus and i2c
- BMC supports three system interface methods; KCS (keyboard controller style), SMIC (system management interface controller), BT (block transfer), determine method from SMBIOS table
- BMC is active even when system is powered off
- Doesn't require an agent on host system
- Allows issuing commands pre-boot environment

Limitations of IPMI

- Only monitors “IPMI devices”
  - Does not monitor RAID Status
  - Does not monitor network status
- Does not return SMBIOS asset information
- Provides text-only console redirection
- No built-in Webserver
- No Virtual media support
- Third party HW/SW solutions continue to fill these gaps
  - Remote Assistant Card (assumes function of BMC if installed)

IPMI Definitions

SDR – System Data Repository

- Stored in NVRAM on the BMC
- Contains data about all devices that may be in the system
  - Sensors: SDR contains sensor address, name, type (Temperature, Voltage, Fan, etc), Units (RPM, volts, etc), granularity, valid thresholds, and entity
    - Sensors support up to 6 thresholds: Upper/Lower Non-Recoverable, Critical, Non-Critical
    - Entity describes which object is being monitored: CPU, Power Supply, etc
  - Field Replaceable Unit (FRU) Devices

SEL – System Event Log

- Stored in NVRAM on the BMC
  - Logs all sensor threshold violations
  - Memory ECC errors
  - Power ON/OFF requests
  - Watchdog timer
- Each SEL Entry Contains
  - Date/time of event
  - Device that caused event
  - Information on event (current sensor reading, failed DIMM ID)

SOL – Serial Over LAN

- Feature in IPMI 1.5, allows redirecting serial port over the network
  - Ability to view pre-OS boot screens; BIOS setup
- Requires SOLProxy service installed on management station
- Doesn't require separate terminal server hardware

RHEL/SLES and IPMI

- RHEL3 U7, RHEL4 U2, SLES9 SP3 and SLES10 support IPMI based systems out of the box
  - ipmi_msghandler – main module, handles message synchronization
  - ipmi_si – System interface for KCS, SMIC, and BT system interfaces
  - ipmi_devintf – IOCTL interface for userspace programs
  - ipmi_watchdog – Watchdog timer interface, supports Linux watchdog IOCTLs
- Generates SEL entry at kernel panic
- OpenIPMI provides startup service and libs
- ipmitool provides CLI to BMC using drivers

OpenIPMI Drivers and Libraries

- OpenIPMI project designed to provide an open source IPMI library on Linux
- Project home: http://openipmi.sourceforge.net
- Abstracts methods of system interface access from user
- Fully asynchronous, event driven interface
- Composed of Linux kernel drivers and userspace libraries
  - Now included with all major Linux distros

ipmitool

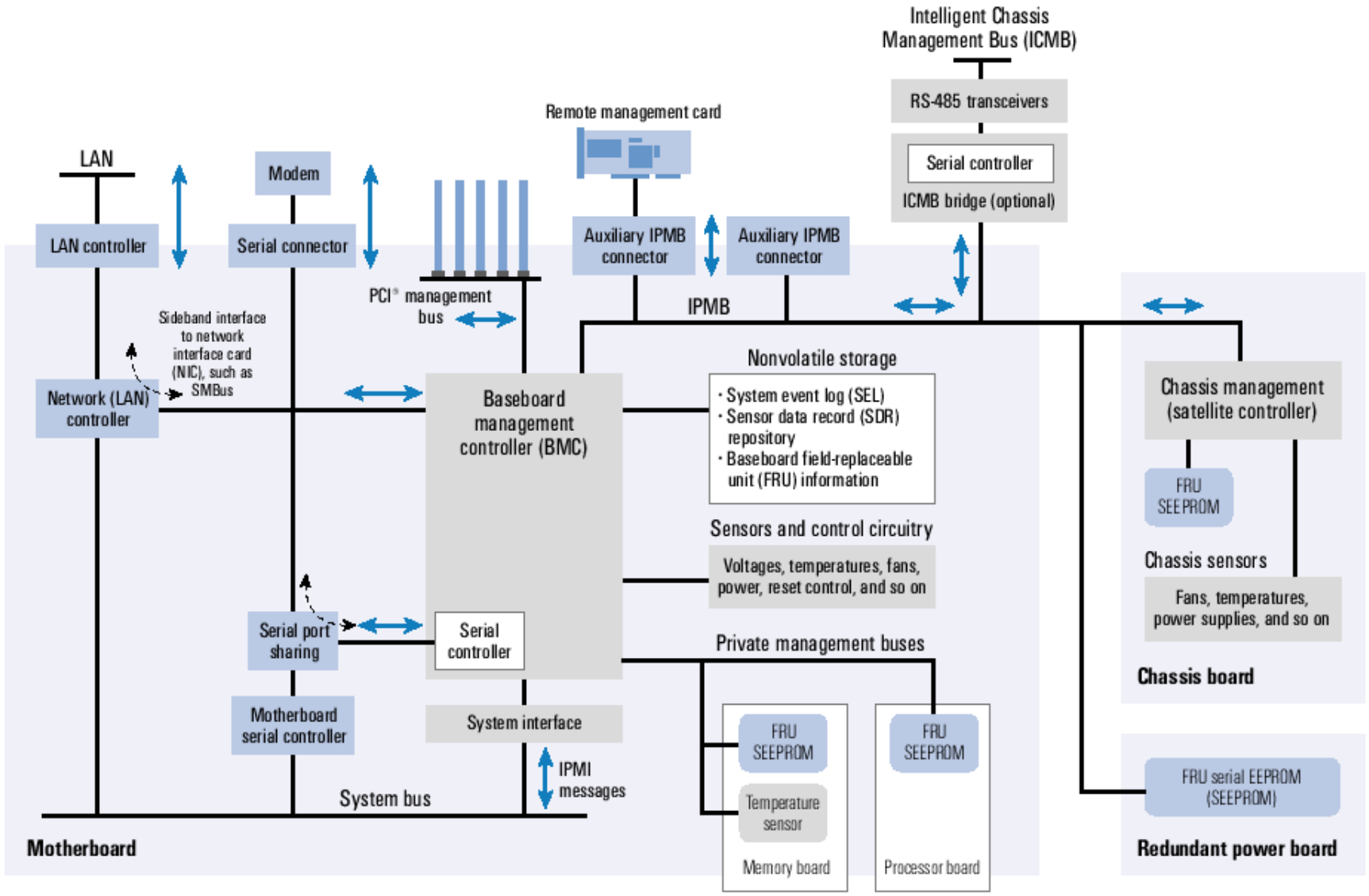
- Linux CLI tool
  - Can be used on local or remote system
- Can query sensors, SDR, SEL, power states, MC info
- Can set power states (on, off, cycle, reset)
- Can configure MC, NIC, SOL, MC users
- Can send events to SEL
- Can set boot order

Additional Information

Helpful Information

- IPMI Specifications, Reference Drivers, Test Tools
  - http://www.intel.com/design/servers/ipmi
- IPMITool/OpenIPMI Projects
  - http://ipmitool.sourceforge.net
  - http://openipmi.sourceforge.net
  - http://ipmiutil.sourceforge.net (formerly the panicsel project)
- Dell OpenIPMI drivers (RHEL/SLES packages)
  - http://linux.dell.com/files/openipmi
- Dell PowerSolutions Articles
  - Managing and Monitoring High-Performance Computing Clusters with IPMI
    - http://www.dell.com/downloads/global/power/ps4q04-20040138-Fang.pdf
  - Managing Dell PowerEdge servers using IPMITool
    - http://www.dell.com/downloads/global/power/ps4q04-20040204-Murphy.pdf

IPMI System Layout



IPMI Message Format

- IPMI commands are grouped into several different functional classes
- All IPMI commands use a common message format
  - 3 byte command header: netfn cmd lun
  - n bytes of data
    - netfn is command class (Application, Storage, Sensor, Chassis, OEM)
    - cmd is operation to perform, depends on netfn
      - Sensor:read sensor
      - Sensor:set threshold
      - Chassis:identify
      - Chassis:power off
    - lun is logical unit (usually 0)
    - data is command-specific data (sensor id, sensor thresholds)
- IPMI Specification allows for future expansion

ipmitool

- Command-line utility for issuing common IPMI requests
- Allows remote operation
- Usage: ipmitool [-v] [-I intf] [-H host] [-k key] [-U user] [-P password] [-E] command...
  - -v : Verbosity, can be specified multiple times -vv
  - -I intf : IPMI interface to use
    - open – OpenIPMI driver (default)
    - lan – LAN connection (remote connection, requires -H/-U/-P arguments)
    - lanplus – LANplus connection (IPMI 2.0) Requires H/U/P arguments supplied
  - -H host : Hostname of remote system (-I lan only)
  - -k key : KG Key (System password) (-I lanplus only)
  - -U user : Username on remote system (-I lan only)
  - -P pass : Password for user on remote system (-I lan only)
  - -E : Read password from IPMI_PASSWORD environment variable
  - If -E and -P are not specified on a remote connection, the utility prompts for a password

Ipmitool – User administration

- BMC Supports multiple users, username/password required for remote connections
- Users are local to BMC, not synchronized with OS
- Adding a user to the BM
  - ipmitool <opts> user set name 2 <username>
  - Ipmitool <opts> user set password 2 <password>
  - ipmitool <opts> user enable 2

Ipmitool – Configuring NIC

- Lan configuration is best done locally
- Display current LAN configuration
  - ipmitool <opts> lan print 1
- Configure static IP Address
  - ipmitool <opts> lan set 1 ipsrc static
  - ipmitool <opts> lan set 1 ipaddr x.x.x.x
  - ipmitool <opts> lan set 1 netmask x.x.x.x
- Configure DHCP IP Address
  - ipmitool <opts> lan set 1 ipsrc dhcp

Ipmitool – Sensor commands

- Displaying all objects in SDR
  - ipmitool <opts> sdr list
  - Ipmitool <opts> sdr dump <filename> (Dump SDR contents to a file)
- Displaying all sensors in the system
  - ipmitool <opts> sensor list
- Displaying an individual sensor
  - ipmitool <opts> sensor get “Temp”
- Changing sensor threshold
  - ipmitool <opts> sensor thresh “Temp” ucr 100
  - Thresholds are: unr, ucr, unc, lnc, lcr, lnr

Ipmitool – Chassis commands

- Chassis Identify
  - ipmitool <opts> chassis identify (defaults to 15 seconds)
  - ipmitool <opts> chassis identify off
- Controlling System Power
  - ipmitool <opts> chassis power status
  - ipmitool <opts> chassis power off
  - ipmitool <opts> chassis power on
  - ipmitool <opts> chassis power cycle
  - ipmitool <opts> chassis power soft (Performs safe OS shutdown)
- Changing System Boot Order
  - ipmitool <opts> chassis bootdev pxe
  - ipmitool <opts> chassis bootdev harddisk
  - ipmitool <opts> chassis bootdev cdrom

Ipmitool – SEL Commands

- Retrieving information about SEL
  - ipmitool <opts> sel info
  - Displays date/time of last event, last log clear time, # of entries
- Displaying SEL
  - ipmitool <opts> sel list
- Clearing SEL
  - ipmitool <opts> sel clear

Ipmitool – Configuring Serial-over-LAN

- Remote System Steps:
  - Setup BIOS to peform console redirection
    - Serial Port : BMC NIC
    - Console Redirection : Serial Port 1
  - Modify grub.conf to support serial port
    - serial –unit=0 –speed=19200 –word=8 –parity=no –stop=1
    - kernel .... console=ttyS0,19200 console=tty0
  - Modify /etc/inittab to spawn serial console
    - s0:345:respawn:/sbin/agetty ttyS0 19200 vt100
- Local System Steps
  - Set baud rate for remote system serial -> LAN
    - ipmitool <opts> isol setup 19200
  - Install osabmcutil RPM on management station
    - rpm -ivh osabmcutil-1.0.1_A00.i386.rpm