Intel® Switch Module IXM5414E

Flow Control Settings Application Note

Revision 1.0

April 2005

Contents

Revision History

Date	Revision	Description
April 2005	1.0	Initial release

1.0 Executive Summary

Intel's IXM5414E switch module is one of up to four switch modules that can be installed in Intel's Server Blade Chassis Enterprise (SBCE) blade server chassis. This high-performance IXM5414E switch module is ideally suited for networking environments that require superior microprocessor performance, efficient memory management, flexibility and reliable data storage.

Each IXM5414E switch module has a total of twenty ports consisting of:

- Four external 1000BASE-T ports for making 10/100/1000 Mbps connections to backbone switches, end stations, or servers
- Fourteen internal full-duplex gigabit ports; each connected to one blade server
- Two internal full-duplex 100 Mbps ports connected to the management modules.

The IXM5414E switch module implements the IEEE 802.3x full-duplex flow control mechanism for the fourteen internal gigabit ports and the four external ports. In the default switch module configuration, the flow control settings of all the ports are “Off”.

This paper describes the flow control mechanism and the implications of setting the mechanism “On” or “Off” in the switch module.

2.0 What is Flow Control

The IEEE 802.3x committee has standardized a method of flow control using a PAUSE function in the full-duplex Ethernet environment. At the data link layer, this type of flow control is applicable only on devices that are connected through a point-to-point link and not end-to-end for the hosts, unless they are directly connected. This means flow control mechanism may be set up between:

- two stations on a point-to-point link
- an end station and a switch
- two switches

If the receiving station becomes congested due to the source sending a large number of frames, it can send back a frame called a “pause frame” to the source, instructing that station to stop sending packets for a specific period of time. The sending station waits for the requested time before sending more data. The receiving station can also send a frame back to the source with a time-to-wait of zero, instructing the source to begin sending data again. This helps to minimize frame loss due to buffer overflow at the receiving station.

This flow-control mechanism is developed to match the sending and receiving devices' throughputs. For example, a server may transmit to a client at a rate of 3000 pps. The client, however, may not be able to accept packets at that rate because of CPU interrupts, excessive network broadcasts, or multitasking within the system. In such cases, the client sends out a pause frame and requests the server to delay transmission for a certain period of time. This gap allows it to complete processing frames already received in its buffer and be prepared to receive more frames from the server.

Flow control is thus a mechanism to deal with temporary overload conditions by reducing inbound traffic when a buffer overflow occurs at the receiver side. If the sustained (i.e. - a steady state) traffic level exceeds the level the device is designed to handle, then the flow control mechanism may not be the ideal solution. In other words, the PAUSE based flow control mechanism cannot cure a sustained overload of network traffic.

3.0 Enabling Flow Control in IXM5414E

The IXM5414E 4X release has the following default configurations:

Table 1. IXM5414E 4X default configurations

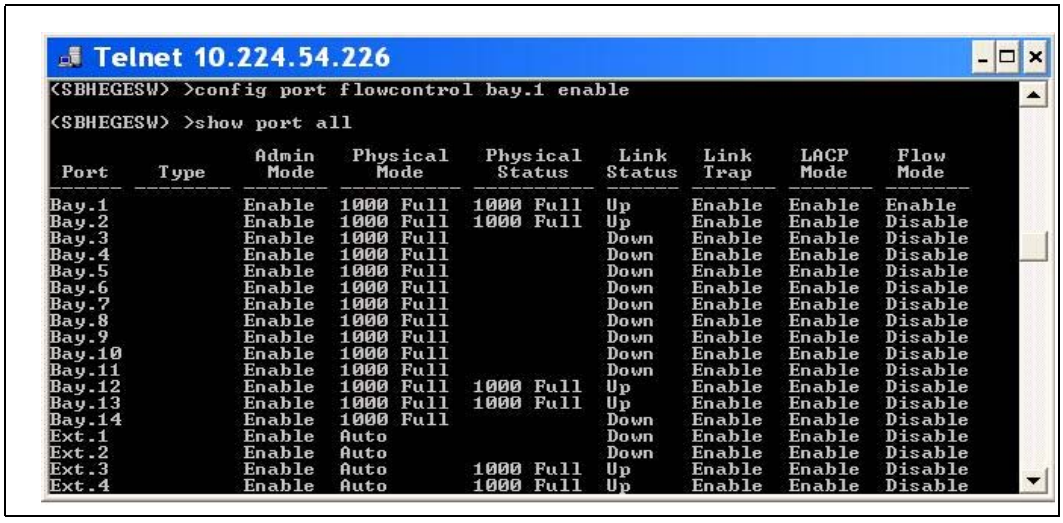
Switch Module Ports	Default Flow Control Settings	Notes
Bay Ports	Off	The flow control can be turned on or off, for each port using CLI commands or Web-based GUI.
External Ports	Off	The flow control can be turned on or off, for port using CLI commands or Web-based GUI.

To enable flow control using CLI

1. Telnet to the management IP address of the switch.
2. After logging in, execute the following command:
   `config port flowcontrol <port/listofports/all> <enable/disable>`
3. Execute the command to check the status of the settings on all ports:
   `show port <port/listofports/all>`

Figure 1 depicts the execution of the preceding on a switch module.

Figure 1. Enabling flow control using CLI



To enable flow control using web-based GUI

1. Log into the Switch management GUI.
2. From the **System** drop down menu on the left panel, select **Port**, then **Configuration**.
3. From the drop down box, select the port for which you need to change the settings.
4. From the **Flow Control Mode** drop down box, enable flow control by selecting the **Enable** option.
5. Choose the **Apply** button.
   The Summary option under the **Port** menu reflects the present configurations' status for all the ports.

Figure 2 depicts a sample configuration screen.

Figure 2. Sample configuration screen

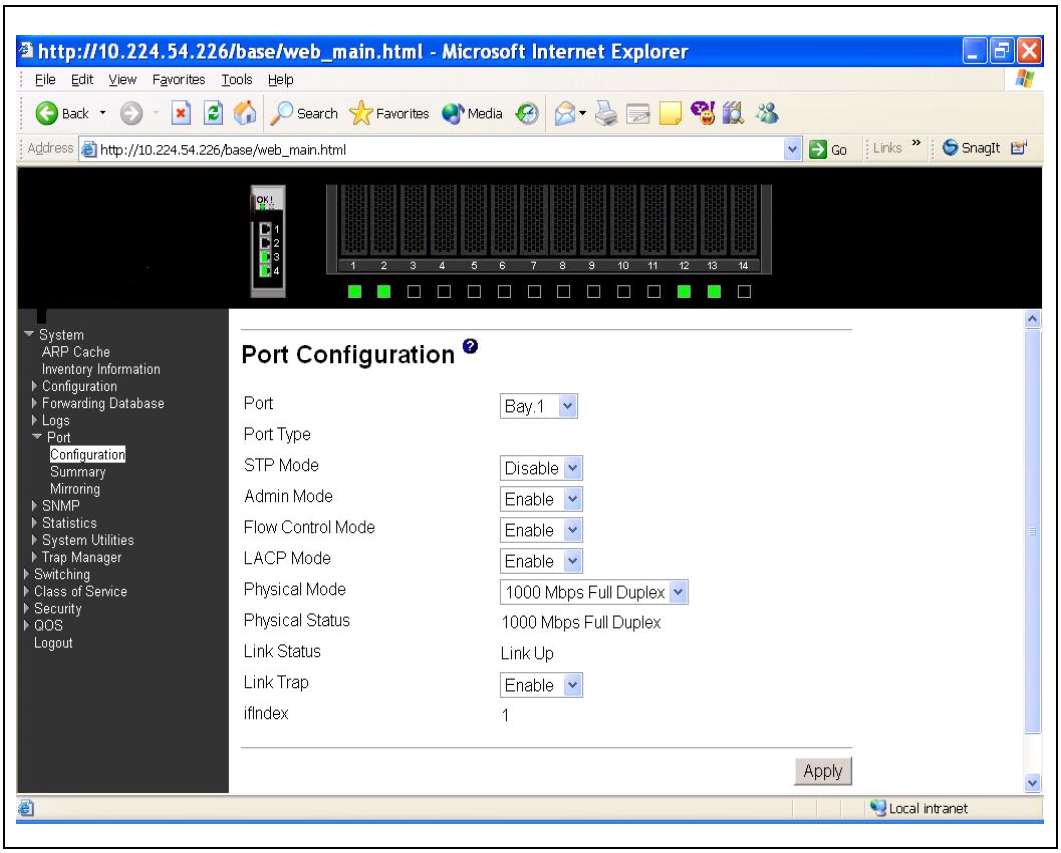
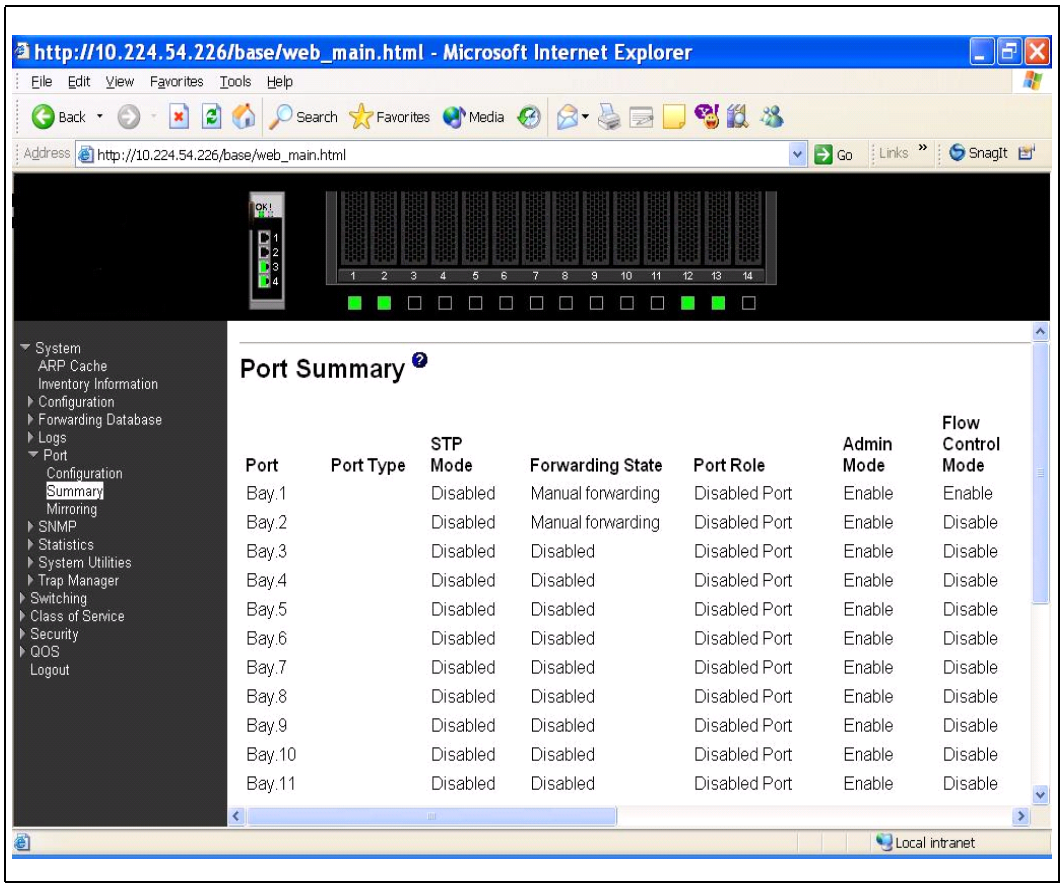


Figure 3 depicts a Port Summary screen.

Figure 3. Port Summary screen



Note: When required, the same commands with relevant parameters or GUI options may be used to disable the flow control settings for each port

4.0 When to Enable Flow Control

In a modular server environment, flow control is generally necessary when frame or data loss is observed during communication between a blade server operating at a higher speed (1Gbps) and an external client operating at a much lower speed (10/100 Mbps). In this case, the flow control needs to be enabled on the switch module for two ports - one external port where the client is connected and the bay port where the blade server connects.

Flow control may not be necessary when the traffic is one-to-one between server blades in the same chassis. But, flow control may need to be turned on, when there is many-to-one traffic at line rate (1 Gbps). For example, if the server blades at bay port 1, 2 and 3 are pumping line rate traffic to server blade at bay port 4, then the flow control should be enabled on all four ports (bay ports 1-4) of the switch module.

Flow control may also be needed when traffic goes from one server blade in one chassis to another one in a different chassis, both chassis being connected through the external ports.

5.0 Tips for Configuring Flow Control

Consider the following tips while configuring the flow control settings:

1. Due to a certain limitation of the IXE5416 ASIC used in the IXM5414E switch module, when flow control settings are enabled on the external ports of the switch connected directly to the external ports on another IXM5414E switch and a steady, excessive traffic flow is maintained between the two switch modules, then occasionally, both IXM5414E switch modules may stop switching traffic and become unresponsive to ping, telnet or http requests.

   The way to recover from such a “hang” condition is to reset power to the switch modules; either manually or through their respective management module Web-based GUIs.

   An IXM5414E connected to any other switch, that does not use the IXE5416 ASIC, will not experience such a “hang” condition, when flow control settings are enabled.

2. In an IXM5414E switch module, flow control is effective only when traffic remains localized within two groups of ports - one group consisting of bay ports 1 through 8 and the other group consisting of bay ports 9 through 14 and the external ports 1 through 4.

6.0 Summary

The flow control mechanism is a useful feature that ensures that network devices operating at different speeds can communicate without any data loss. The flow control mechanism may also be used to ensure fairness of switching bandwidth on the backplane when different operating systems are used on the server blades in the chassis. In the default configuration of the IXM5414E switch module, the flow control settings are turned off for all ports and can be turned on either using the switch module CLI or the GUI interfaces.