

Designing PCI-Express Systems using PI2PCIE Signal Switches

By Paul Li

Introduction

The PCI-Express architecture is rapidly becoming the mainstream of system designs. This application note will introduce Pericom PCI-Express signal switches for use in PCI-Express system designs.

PCI-Express Packet Switch versus Signal Switch

There are two different types of PCI-Express switches: the packet switch as defined in the PCI-Express base specification, and the signal switch that Pericom developed for PCI-Express applications. Well recognized in the electronics industry, Pericom is the technology and market leader in high-speed differential switches using CMOS technology.

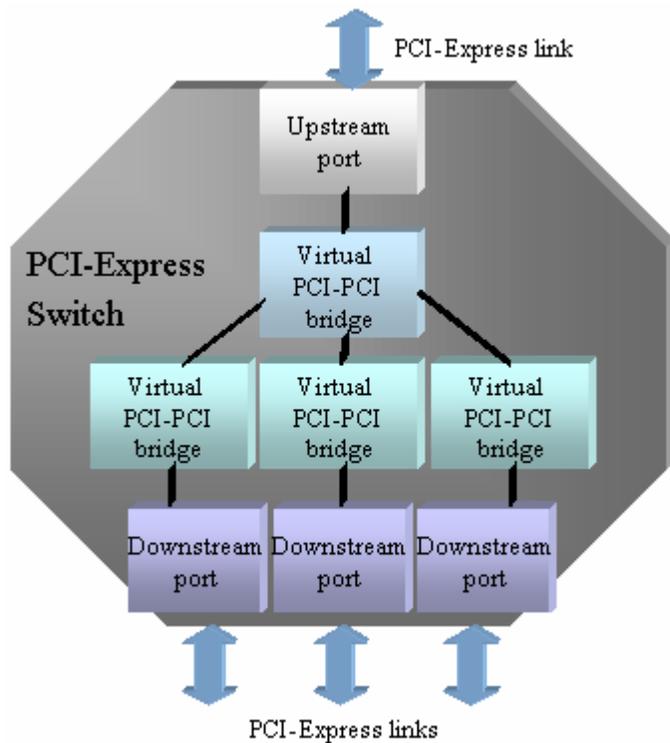


Figure 1: Block diagram of PCI-Express packet switch

Per the PCI-Express base specification revision 1.0a, a packet switch is defined as a logical assembly of multiple virtual PCI-to-PCI bridge devices as illustrated in Figure 1.

All packet switches are governed by the following rules.

- Switches appear to configuration software as two or more logical PCI-to-PCI bridges.
- A switch forwards transactions using PCI bridge mechanisms; e.g., address based routing.
- Except as noted in the spec, a switch must forward all types of transaction layer packets between any set of ports.

This definition tells us that in the PCI-Express architecture, a packet switch virtually performs the mechanisms similar as what the PCI-bridge does in the PCI architecture. This explains why the PCI-Express is downward compatible with all software developed for the PCI environment, because internally, the packet switch is compliant to all the rules defined for the PCI-bridge. Figure 2 is an example of a packet switch application in the PCI-Express base specification.

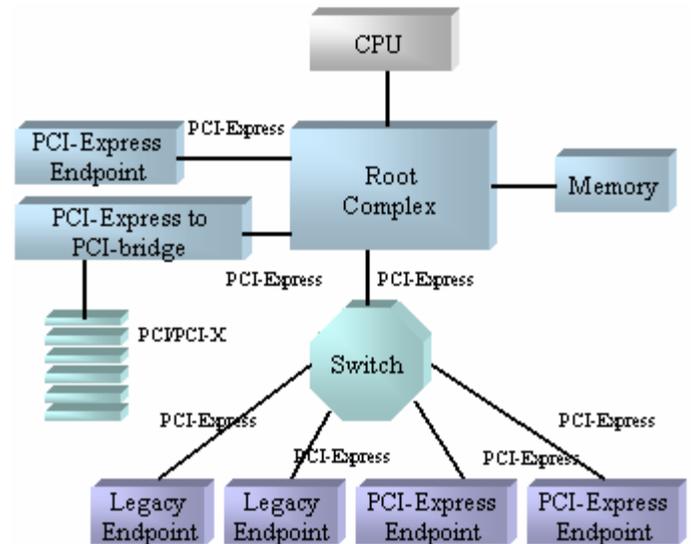


Figure 2: Application of the PCI-Express packet switch

In real PCI-Express applications, there are often conditions where a root complex port needs to connect to two endpoint chipsets or connectors (one at a time) as depicted in Figure 5.

All of the connection selection is done during the power-off of the system. There is no need to switch the connection while the system is running. Both packet switch, and signal switch can be used for these kinds of applications. However, using a packet switch for these applications is overkill since a packet switch is more expensive and comes with a longer design cycle.

The Pericom PCI-Express Signal Switch, as shown in Figure 3, is ideally designed at a far lower cost and is benefited with drastically shorter design cycles.

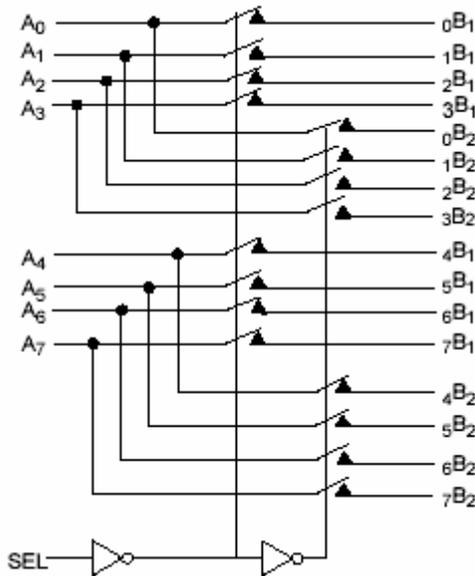


Figure 3: Block diagram of PI2PCIExxx signal switch

Why is a switch needed for the connection between the root complex and endpoints?

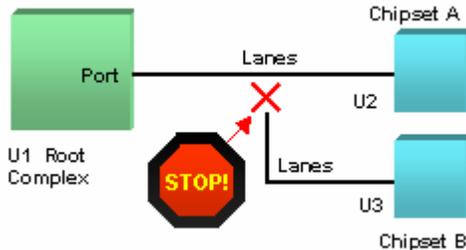


Figure 4: Data transmission of PCI-Express is point-to-point

The PCI-Express signal transmission is a point-to-point architecture, which does not allow the sub-trace as shown in Figure 4. If the root complex was connected to both the endpoint chipsets A and B via traces as the multi-drop architecture of the PCI bus, then the sub-traces will cause the reflections on the signal rising and falling edges as the “knees”. These “knees” will degrade the signal integrity and will cause errors bits.

For example, since the velocity of the signal on the FR4 PC board is about 5” per 1ns, the “knees” caused by a 1” sub-trace is 200ps. This 200ps “knees” will add to the rising and falling edges of the signal, totaled at 400ps, which will corrupt the PCI-Express signal bit with 400ps duration.

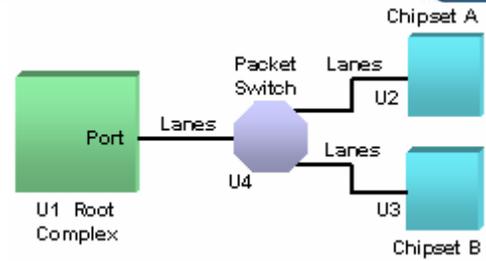


Figure 5: Packet switch

Using packet and signal switches to isolate the two traces for chipset A and B in Figure 5 will prevent the reflections. The 50-ohm resistors at the input of the packet switch will terminate the signals from the root complex and then will re-start the transmission line at its output with a re-drive. This will ensure that the connection between the root complex and the chipsets A and B are point-to-point architectures. But the packet switch is overkill for this application since it is more expensive and has a longer design cycle.

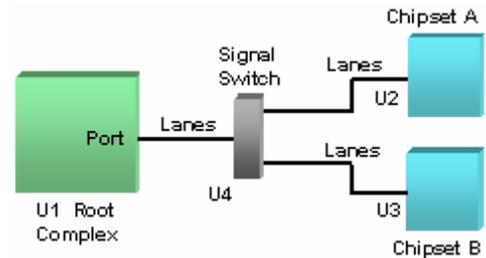


Figure 6: Signal Switch

Pericom Signal Switches are ideally suited for these applications at much lower cost and a faster design cycle. In Figure 6, when the trace from chipset A (or B) is connected to the root complex, the switch will internally disconnect the trace from chipset B (or A). Therefore there is no reflection from the disconnected trace to the other trace.

Typical PCI-Express applications using PI2PCIExxx Signal Switches

The two x2 ports in the root complex U1 in Figure 7 can be re-configured as one x4 port to J1, which is an x4 PCI-Express connector but can also be used as a x2 connector. When the switch U2 connects between the root complex U1 and J1, J1 becomes an x4 connector and J2 is not in use. When the switch connects to J2, both J1 and J2 become the x2 connectors. This application will provide the flexibility of re-configuration. This approach can further derive to re-configure many other combinations of ports, lanes and connectors. You name it.

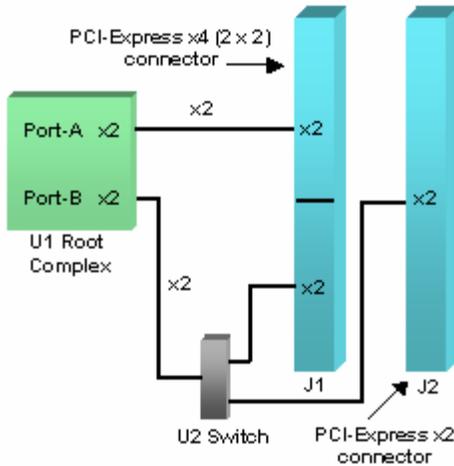


Figure 7: Reconfiguration of two x2 PCI-Express connectors

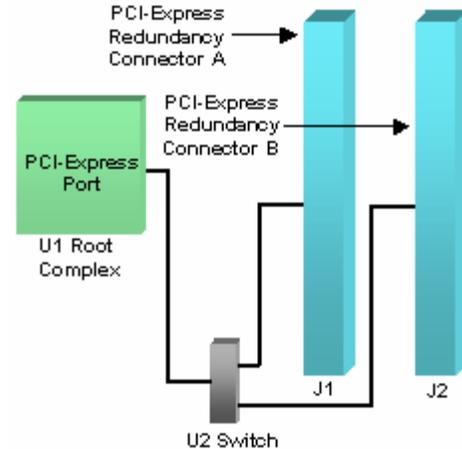


Figure 9: Signal Switch between the two redundancy connectors for the server application

The root complex in Figure 8 has a x16 graphic port. J1 is a x16 graphic connector and J2 is an x8 connector. When the switch connects to J1, J1 becomes a x16 graphic connector while J2 is not in use. When the switch connects to J2, both J1 and J2 become x8 connectors (for instance, the application using two x8 graphic cards specially designed for high-speed and high-resolution 3D gaming).

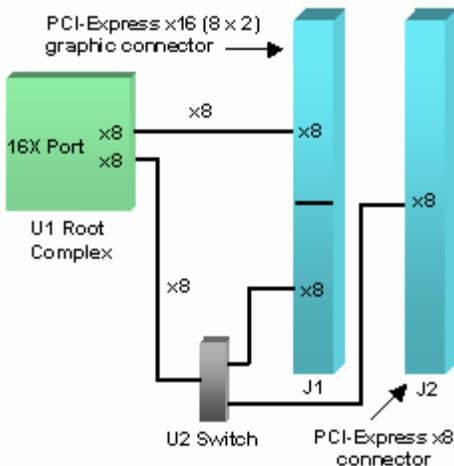


Figure 8: Reconfiguration between the two graphic cards options

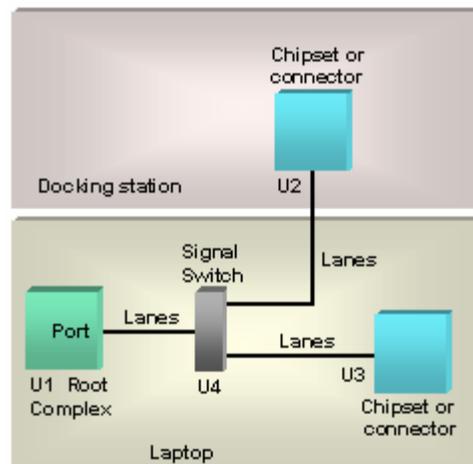


Figure 10: Signal Switch between laptop and docking station

In figure 9, the switch U2 will switch to the redundancy connector B when the card in the connector A is failed.

In Figure 10, the switch U4 will connect to the endpoint chipset U3 when the laptop is working at stand-long mode. U4 will switch to the chipset U2 in the docking station when the laptop is docked.

The topology of the switch layout

An ideal differential switch would be a wire with 100-ohm differential impedance, zero capacitance, and zero resistance. Thus, when this ideal switch is inserted in differential traces at 100-ohm impedance, there will be no insertion loss from the switch's C-on & R-on, nor return loss from the impedance mismatching.

But due to technology limitations, the resistance and capacitance from the switch could not be completely negated. Besides, there is no such process that a switch can be made with 100-ohm impedance. Thus, when a switch is inserted in a trace, it will cause the insertion loss. It will also cause return loss, which will generate the reflections that degrade the signal integrity.

Based on theoretical analysis and system application experience, we have developed application techniques that will optimize insertion loss and return loss caused by inserting the switch into the trace.

1. Locate the switch as close as possible to the input/output pins of the driver and receiver (in the root complex or the endpoints), within 1/4" to 1/2" if feasible, as illustrated in the block A in Figure 11.

When the switch is located very close to the output pins of the driver, the reflections caused by the short trace is minimized and the driver will directly absorb the C-on and R-on without trace delay. Thus, the switch is considered merged into the driver instead of an inserted object in the transmission line. This topology will ensure the minimal reflections from the switch.

When the switch is populated very close to the input pins of the receiver, the reflections caused by the short trace is minimized while the reflections from the switch to the long trace to the driver will go toward the driver and not be seen by the receiver. Thus, the switch at receiver will only cause attenuation but not reflection.

Besides, the R-on from the switch will cause an RC effect that will attenuate the signal. When the switch is at the end of the trace where the capacitance is the lowest, the RC effect from R-on will be minimal.

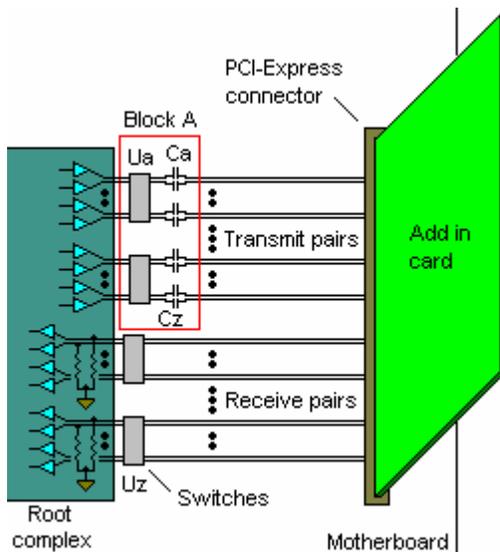


Figure 11: Recommended location of the switch

2. When the routing space allows, separating the transmission pairs and the reception pairs by using two separated switch groups, as in Figure 11, will help reduce the interference. When the signal pairs converge into the narrow area of the switch pins without enough space between the pairs for isolation, it could cause extra interference. The interlaced transmission pair and reception pair tend to have more edge phase difference than the adjacent pairs in the same transmission/reception groups. Thus, in the separated group's topology, the noise from the adjacent pairs will impact on the edges that are more immune thereby reducing jitter.

3. The allowable input signal range of the PI2PCIExxx Signal Switches are from -0.5V to 1.8V, at 1.8V Vcc. For most root complex chipsets, for instance the Intel chipsets for the 915 and 925 motherboards, the Vcc is 1.5V and the output signal voltage is from 0.6V to 1.35V, and covered by PI2PCIExxx Signal Switches. For signal levels at 1.8V or below, directly locate the switch to the root complex as in Figure 11. But in case the signal swing range is above 1.8V, then locate the AC coupling capacitors between the switch and the root complex as in the block B in Figure 12. This will offset the signal down to 0V and the signal applied to the switch will ride on the 0V offset, as +/- 0.4V.

In normal conditions, the root complex output signal level should not exceed its Vcc level.

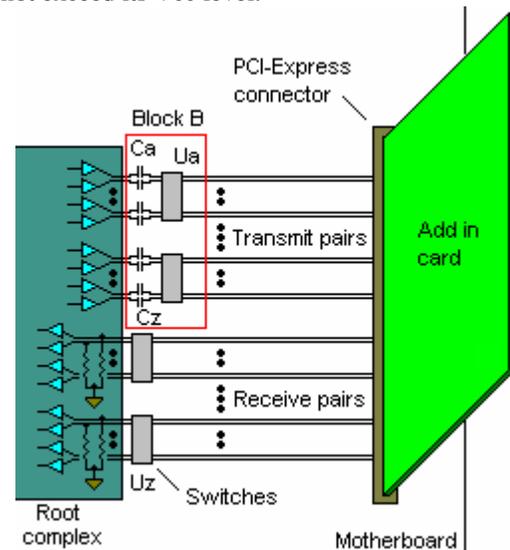


Figure 12: Recommended location of the switch for signal voltage level higher than 1.8V

4. For detailed PCI-Express PCB layout guidelines, refer to the PCI-SIG presentation of "PCI Express™ Technical Training Day PCI Express™ Technical Training Day". Visit the web site www.pcisig.com, register, log-in, and click on http://www.pcisig.com/events/express_trainingday, then click on "PCI Express Technical Training Day presentation".

Pericom demo boards

Pericom has two types of demonstration boards; both are available by request by visiting www.pericom.com/techsupport. The system test board in Figure 13 is designed to test the performance of PI2PCIE Signal Switches in a real system condition, using an Intel motherboard (915 or 925) and a x16 PCI-Express graphic card, both available in the PC market. The trace on the test board is 11", plus the 3" and 2.5" on the motherboard and graphic card, for a total of 16.5", which is the test against the nominal maximum 16" trace specified in the above PCI-Express document.

There are eight PI2PCIE switches on the system test board, each for two lanes, switching between the two connectors either plugged with the x16 graphic card connecting to a monitor. The

header “sw1” on the top-left corner of the system test board is for the connector selection. When a jumper connects the left and the middle pins of sw1, the upper connector is selected. When the jumper connects the right and the middle pins of sw1, the lower connector is selected.

It is not recommended to perform the Rx eye compliance test using the system test board since the switches on the system test board could not be close to the driver nor the receiver preferred and will generate reflections as discussed above. Therefore we have made a dedicated Rx eye compliance test board as shown in Figure 14.

For the Rx eye compliance test, it is recommended to use active differential probes and scopes recommended on the PCI-SIG web site at:

http://www.pcisig.com/specifications/pciexpress/compliance/compliance_library/.

Connect the two differential probes between TP1 and TP2 and the two ground-via near by TP1 and TP2: the two positive pins to TP1 and TP2 and the two negative pins to the ground.

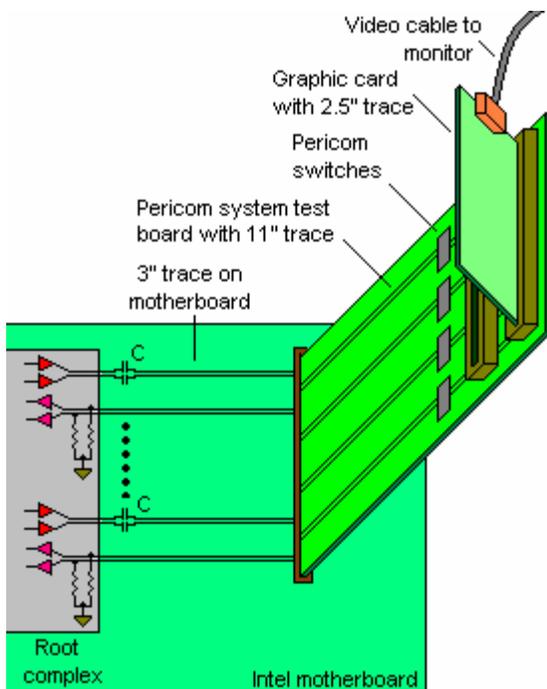


Figure 13: System test board

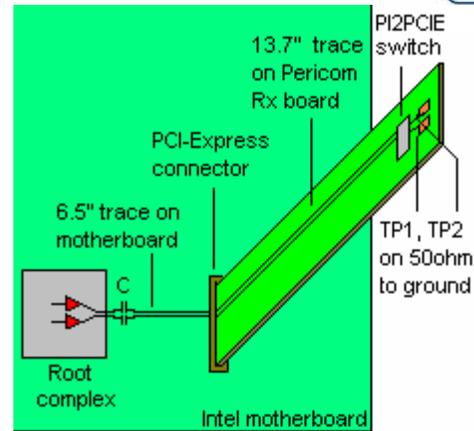


Figure 14: Rx eye compliance test board

Refer to the detailed document on the PCI-SIG web site above on how to use a particular name-brand scope, capturing the waveform data, and save in “xxxx.csv” files. Then, run the “xxxx.csv” files in the “SigTest 2.0 Beta” compliance test software. To download the free “SigTest 2.0 Beta” software, visit the web page www.pcisig.com, register and log-in, click on:

http://www.pcisig.com/specifications/pciexpress/compliance/compliance_library/, then click on the “PCI Express electrical test software” in the middle of the page.

The waveforms in Figures 15 and 16 were captured using an Agilent 54855A 6 GHz, 20 GSa/s differential scope with two Agilent 1134A 7 GHz probes (channel 1 and channel 3), as specified in the PCI-SIG compliance test document “PCI Express Signal Quality Test Methodology Users Guide for Agilent infiniium 54855A” at:

http://www.pcisig.com/specifications/pciexpress/compliance/compliance_library/PCIE_SQ_TEST_r0_7_INF54855A.pdf

Table 1, Figures 15 and 16 illustrate the compliance test result of the “SigTest 2.0 Beta” software, using the “xxx.csv” files captured on TP1 and TP2 on the Rx test board with a total 19.2” trace (Figure 14) as described above.

Overall Result:	Pass
Data Rate: 2.489056 GB/s	Pass
Median to Peak Jitter: 39.793455 ps	Pass
Peak to Peak Jitter: 78.888076 ps	Pass
Eye Violations: 0 points:	Pass

Table 1: Rx eye compliance test result of the 19.2” Rx test board

How to choose the template files in "SigTest 2.0 Beta" for compliance test

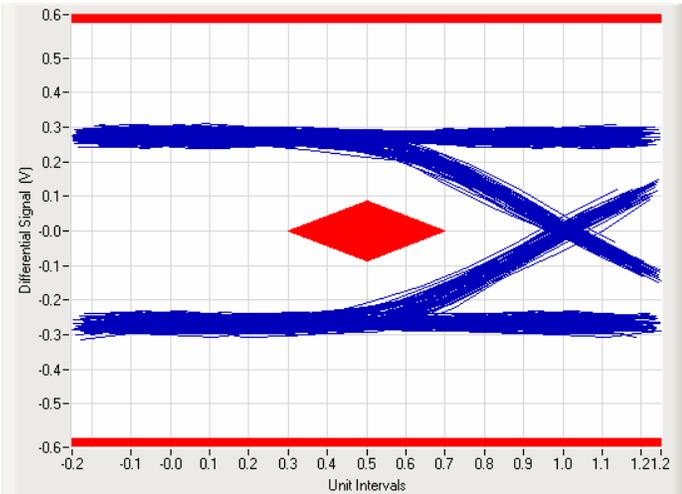


Figure 15: Worst non-transition signal eye of the Rx test board in Figure 14

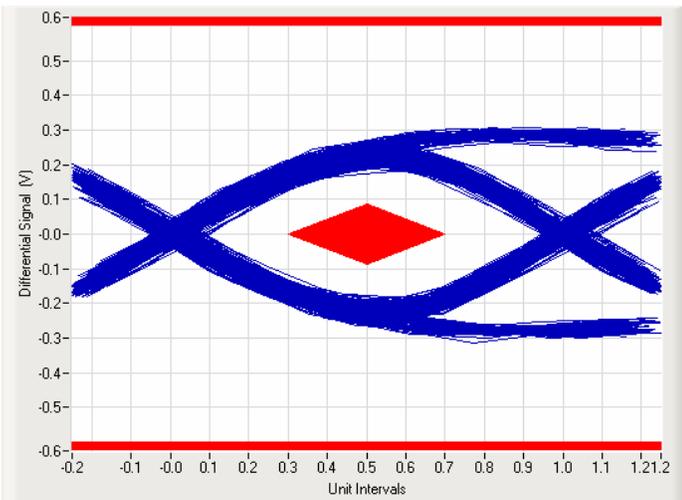


Figure 16: Worst transition signal eye measured on Rx test board in figure 14

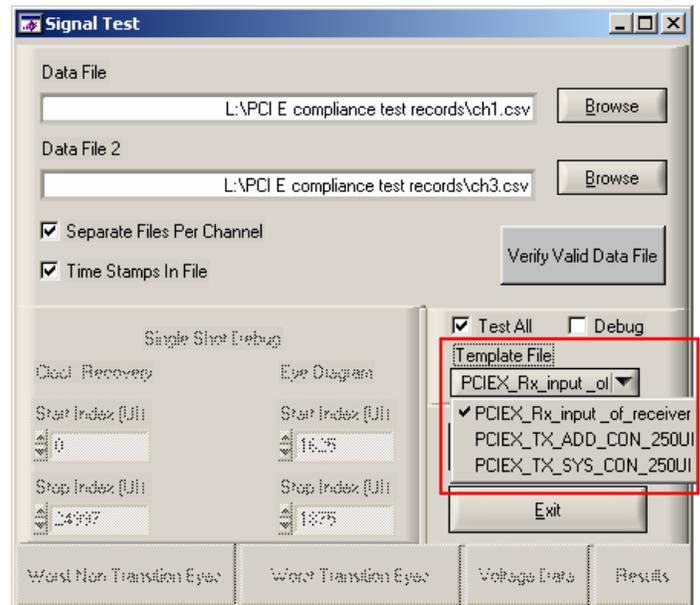


Figure 17: Pop-up window of "SigTest 2.0 Beta" for compliance test

The template file "PCIEX_TX_ADD_CON_250UI" in the red area in Figure 17 is for the add-in card transmitter test in Figure 18, in which the signal on the edge-finger of the add-in card should be able to further drive the maximum 13" trace on the motherboard. Therefore, it uses the transmitter spec of 514mV/237ps for the transition bit with emphasis and uses the spec of 360mV/237ps for the non-transition bit with de-emphasis, as indicated in Tables 4-6 on page 44 of the "PCI-Express card electromechanical specification revision 1.0a".

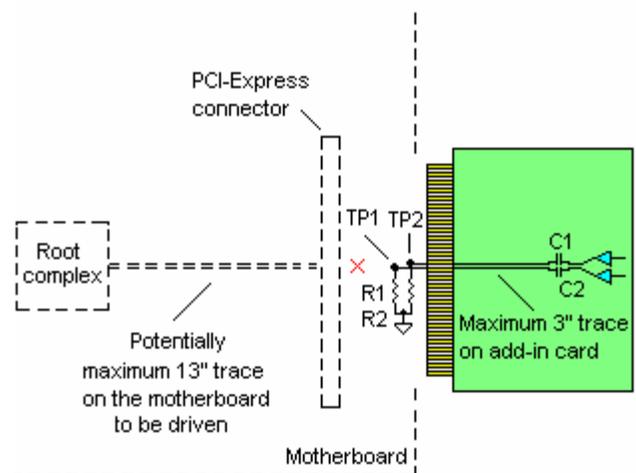


Figure 18: Add-in card transmitter test condition for the template file "PCIEX_TX_ADD_CON_250UI"

Likewise, the template file "PCIEX_TX_SYS_CON_250UI" is for the system board test as in Figure 19, using the spec in the Tables 4-8 on page 46 of the "PCI-Express card electromechanical specification revision 1.0a". Since the signal at the motherboard connector (Figure 19) should be capable of further driving the maximum 3" trace on the add-in card, its specs of 274mV/183ps and 253mV/183ps are higher than the minimum 175mV/0.4UI Rx eye spec specified in the "PCI-Express base specification revision 1.0a".

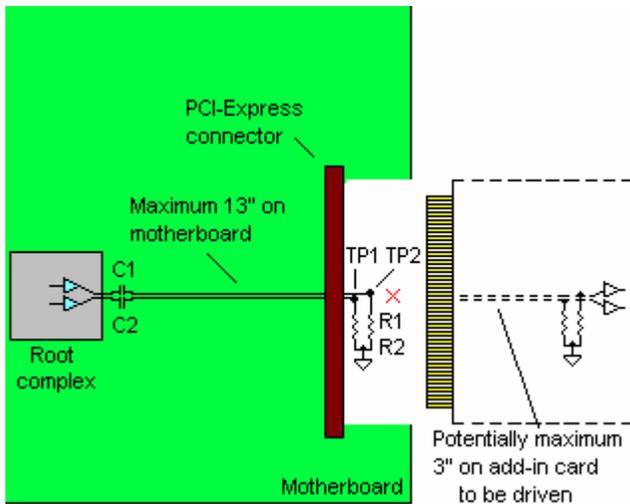


Figure 19: System test condition for the template file "PCIEX_TX_SYS_CON_250UI"

The template file "PCIEX_Rx_input_of_receiver" in Figure 17 created for the Rx test in Figure 14 was not provided in the default template files in "SigTest 2.0 beta". The template file "PCIEX_Rx_input_of_receiver" complies with the minimum 175mV/0.4UI Rx eye spec specified in the Figures 4-26 on page 219 of the "PCI-Express base specification revision 1.0a".

How to create the template files for the Rx compliance test in the "SigTest 2.0 Beta"

The procedure below shows how to create a template file to be used in the "SigTest 2.0 beta", as the template file "PCIEX_Rx_input_of_receiver" in Figure 17 for the Pericom Rx demo board or for any other tests.

1. Go to the directory where the SigTest 2.0 Beta is installed, it is normally in the C:\program files\SigTest 2.0 Beta\templates. If it cannot be found in the C drive, try your other local drives to find where it was installed.

2. In the subdirectory C:\program files\SigTest 2.0 Beta\templates, copy an existing template, e.g. the "PCIEX_TX_ADD_CON_250UI.dat", to a new template file, and rename it to "PCIEX_Rx_input_of_receiver.dat"

3. Open this new file in Microsoft Notepad or equivalent.

4. Towards the end of the file looking for the following statements:

```
; The eye template values
[template]
Upper=.600
Lower=-.600
Point1=".204 0"
Point2=".500 .180"
Point3=".796 0"
Point4=".500 -.180"
```

```
; The Transition eye template values
[TransitionTemplate]
Upper=.600
Lower=-.600
Point1=".204 0"
Point2=".500 .257"
Point3=".796 0"
Point4=".500 -.257"
MedianPeakJitterLimit=81.5E-12
```

The "Point1...Point4" above are the (x, y) co-ordinates that define the template (the red diamond in the compliance test result) of the "PCIEX_TX_ADD_CON_250UI". The ".180" and the "-.180" define the +/-360mV template for the non-transition eye and the ".257" and the "-.257" define the +/-514mV template for the transition eye, as indicated in the table 4-6 in page 44 of the "PCI-Express card electromechanical specification revision 1.0a".

5. In the following, modify the "Point1...Point4" complying the spec of 175mV/0.4UI, as the minimum receiver eye specified in the Figure 4-26 in page 219 of the "PCI-Express base specification revision 1.0a", then save in the "PCIEX_Rx_input_of_receiver.dat" template file.

```
; The eye template values
[template]
Upper=.600
Lower=-.600
Point1=".300 0"
Point2=".500 .0875"
Point3=".700 0"
Point4=".500 -.0875"
```

```
; The Transition eye template values
[TransitionTemplate]
Upper=.600
Lower=-.600
Point1=".300 0"
Point2=".500 .0875"
Point3=".700 0"
Point4=".500 -.0875"
MedianPeakJitterLimit=108.5E-12
```

6. The next time when the “SigTest 2.0 Beta” is used, the new template “PCIEX_Rx_input_of_receiver” is available in the pop-up window as in the Figure 17, in addition to the “PCIEX_TX_ADD_CON_250UI” and the “PCIEX_TX_SYS_CON_250UI”, for the Pericom Rx demo board compliance test.