



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2019년10월16일
(11) 등록번호 10-2033112
(24) 등록일자 2019년10월10일

(51) 국제특허분류(Int. Cl.)
G06F 13/14 (2006.01) G06F 13/38 (2006.01)
(21) 출원번호 10-2013-0083093
(22) 출원일자 2013년07월15일
심사청구일자 2018년04월11일
(65) 공개번호 10-2015-0008746
(43) 공개일자 2015년01월23일
(56) 선행기술조사문헌
US20120144230 A1
US20130346665 A1

(73) 특허권자
한국전자통신연구원
대전광역시 유성구 가정로 218 (가정동)
(72) 발명자
최용석
대전 유성구 가정로 63, 110동 902호 (신성동, 럭키하나아파트)
(74) 대리인
팬코리아특허법인

기술이전 희망 : 기술양도, 실시권허여, 기술지도

전체 청구항 수 : 총 17 항

심사관 : 김세영

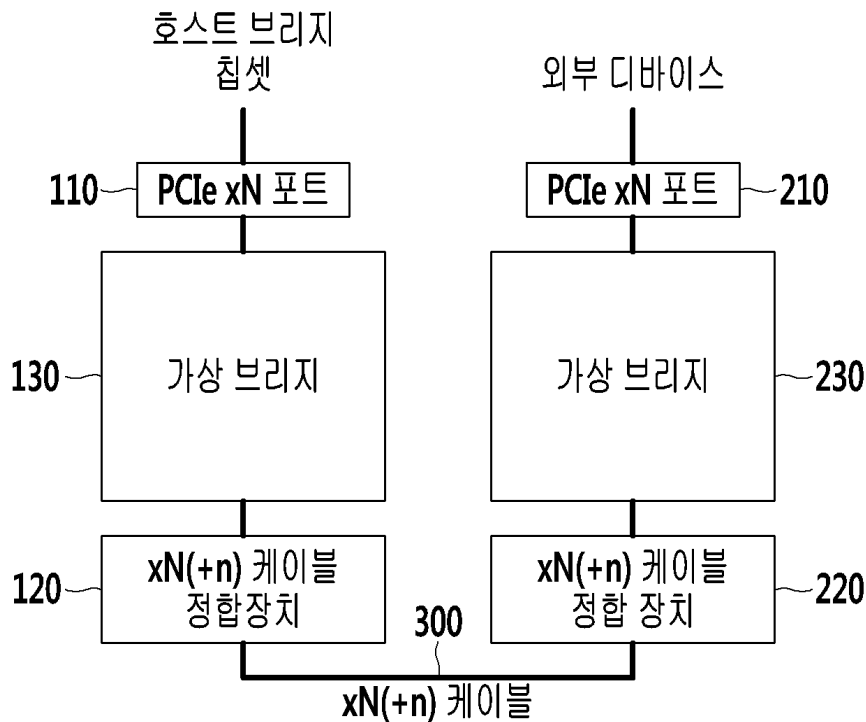
(54) 발명의 명칭 PCI 익스프레스 스위치 장치 및 그의 접속 제어 방법

(57) 요약

PCI 익스프레스 스위치 장치 및 그의 접속 제어 방법이 개시된다.

이 장치에서, 제1 가상 브리지는 제1 PCI 익스프레스(Peripheral Component Interconnect Express) 포트를 통해 컴퓨터 시스템에 접속되어 외부 디바이스와의 PCI 방식에 따른 데이터 송수신을 수행하고, 제2 가상 브리지는 제

(뒷면에 계속)
대표도 - 도5



1 가상 브리지 및 제2 PCI 익스프레스 포트를 통해 상기 외부 디바이스에 접속되며, 상기 제1 가상 브리지와 협력하여 상기 외부 디바이스가 상기 컴퓨터 시스템과 PCI 방식에 따른 데이터 송수신이 가능하도록 한다. 제1 케이블 정합 장치는 상기 제1 가상 브리지에 접속된다. 또한, 제2 케이블 정합 장치는 상기 제2 가상 브리지에 접속되며, 상기 제1 케이블 정합 장치와 PCI 케이블을 통해 접속된다. 또한, 상기 PCI 케이블은 상기 제1 가상 브리지와 상기 제2 가상 브리지가 데이터 송수신을 수행하는 최대 레인의 개수보다 많은 개수의 레인에 해당하는 케이블을 구비한다.

이 발명을 지원한 국가연구개발사업

과제고유번호	10038764
부처명	지식경제부
연구관리전문기관	한국산업기술평가관리원
연구사업명	산업원천기술개발사업(ETRI지원사업)
연구과제명	실리콘 나노포토닉스 기반 차세대 컴퓨터 칩기술
기 여 율	1/1
주관기관	한국전자통신연구원
연구기간	2011.03.01~2016.02.29

명세서

청구범위

청구항 1

제1 PCI 익스프레스(Peripheral Component Interconnect Express) 포트를 통해 컴퓨터 시스템에 접속되어 외부 디바이스와의 PCI 방식에 따른 데이터 송수신을 수행하는 제1 가상 브리지;

제2 PCI 익스프레스 포트를 통해 상기 외부 디바이스에 접속되며, 상기 제1 가상 브리지와 협력하여 상기 외부 디바이스가 상기 컴퓨터 시스템과 PCI 방식에 따른 데이터 송수신이 가능하도록 하는 제2 가상 브리지;

상기 제1 가상 브리지에 접속되는 제1 케이블 정합 장치; 및

상기 제2 가상 브리지에 접속되며, 상기 제1 케이블 정합 장치와 PCI 케이블을 통해 접속되는 제2 케이블 정합 장치를 포함하고,

상기 PCI 케이블은 상기 제1 가상 브리지와 상기 제2 가상 브리지가 데이터 송수신을 수행하는 최대 레인의 개수보다 많은 개수의 레인에 해당하는 케이블을 구비하는 것을 특징으로 하는 PCI 익스프레스 스위치 장치.

청구항 2

제1항에 있어서,

상기 최대 레인의 개수가 N (여기서, N 은 1 이상의 자연수임)이고, 상기 PCI 케이블이 구비하는 케이블의 개수가 $N+n$ (여기서, n 은 최대 N 인 자연수임)인 경우, 상기 제1 가상 브리지는 상기 제2 가상 브리지와 훈련 시퀀스(Training Sequence)를 송신하여 $N+n$ 케이블 중에서 최대 N 개의 레인을 결정하는 것을 특징으로 하는 PCI 익스프레스 스위치 장치.

청구항 3

제2항에 있어서,

상기 제1 가상 브리지 및 제2 가상 브리지는 PCI 익스프레스 계층 구조로 이루어지는 것을 특징으로 하는 PCI 익스프레스 스위치 장치.

청구항 4

제3항에 있어서,

상기 제1 가상 브리지는,

상기 컴퓨터 시스템으로 물리적으로 패킷을 전송하기 위해 상기 제1 PCI 익스프레스 포트에 접속되는 제1 업스트림 레인 물리 계층;

상기 외부 디바이스로 물리적으로 패킷을 전송하기 위해 상기 제1 케이블 정합 장치에 접속되는 제1 다운스트림 레인 물리 계층;

상기 제1 업스트림 레인 물리 계층과 상기 제1 다운스트림 레인 물리 계층을 통해 송수신되는 패킷의 조립 및 분해를 수행하는 제1 트랜잭션 계층;

상기 제1 트랜잭션 계층과 상기 제1 업스트림 레인 물리 계층 사이의 데이터 전달 구성요소로 작용하며, 상기 제1 트랜잭션 계층과 상기 제1 업스트림 레인 물리 계층 사이에서 패킷을 교환하기 위한 신뢰할 수 있는 메커니즘에 링크를 제공하는 제1 데이터 링크 계층; 및

상기 제1 트랜잭션 계층과 상기 제1 다운스트림 레인 물리 계층 사이의 데이터 전달 구성요소로 작용하며, 상기 제1 트랜잭션 계층과 상기 제1 다운스트림 레인 물리 계층 사이에서 패킷을 교환하기 위한 신뢰할 수 있는 메커니즘에 링크를 제공하는 제1 케이블 전송 데이터 링크 계층

을 포함하는 PCI 익스프레스 스위치 장치.

청구항 5

제4항에 있어서,

상기 제2 가상 브리지는,

상기 외부 디바이스로 물리적으로 패킷을 전송하기 위해 상기 제2 PCI 익스프레스 포트에 접속되는 제2 다운스트림 레인 물리 계층;

상기 컴퓨터 시스템으로 물리적으로 패킷을 전송하기 위해 상기 제2 케이블 정합 장치에 접속되는 제2 업스트림 레인 물리 계층;

상기 제2 다운스트림 레인 물리 계층과 상기 제2 업스트림 레인 물리 계층을 통해 송수신되는 패킷의 조립 및 분해를 수행하는 제2 트랜잭션 계층;

상기 제2 트랜잭션 계층과 상기 제2 업스트림 레인 물리 계층 사이의 데이터 전달 구성요소로 작용하며, 상기 제2 트랜잭션 계층과 상기 제2 업스트림 레인 물리 계층 사이에서 패킷을 교환하기 위한 신뢰할 수 있는 메커니즘에 링크를 제공하는 제2 데이터 링크 계층; 및

상기 제2 트랜잭션 계층과 상기 제2 다운스트림 레인 물리 계층 사이의 데이터 전달 구성요소로 작용하며, 상기 제2 트랜잭션 계층과 상기 제2 다운스트림 레인 물리 계층 사이에서 패킷을 교환하기 위한 신뢰할 수 있는 메커니즘에 링크를 제공하는 제2 케이블 전송 데이터 링크 계층

을 포함하는 PCI 익스프레스 스위치 장치.

청구항 6

제5항에 있어서,

상기 제1 가상 브리지와 상기 제2 가상 브리지는 N+n개의 상기 PCI 케이블을 통해 상호 혼련 시퀀스를 송신하여 각각 송신 가능한 레인을 확인하는 것을 특징으로 하는 PCI 익스프레스 스위치 장치.

청구항 7

제6항에 있어서,

상기 제1 가상 브리지는 상기 제2 가상 브리지는 상기 N개의 레인 중에서 장애가 발생된 레인이 있는 경우 추가 n개의 레인 중에서 사용 가능한 레인으로 대체하여 상기 N개의 레인을 구성하는 것을 특징으로 하는 PCI 익스프레스 스위치 장치.

청구항 8

제6항에 있어서,

상기 제1 다운스트림 레인 물리 계층과 상기 제2 업스트림 레인 물리 계층 각각은 N+n개의 케이블에 대응되는 레인별로 LTSSM(Link Training Sequence State Machine)을 구비하고, 상기 레인별 LTSSM에 의해 수신되는 혼련 시퀀스의 상태와 상태 천이를 상기 레인별 LTSSM이 공유하도록 하기 위해 상기 레인별 LTSSM에 공통으로 연결되는 버스를 포함하는 것을 특징으로 하는 PCI 익스프레스 스위치 장치.

청구항 9

컴퓨터 시스템과 외부 디바이스를 PCI 익스프레스(Peripheral Component Interconnect Express) 방식에 의해 접속하는 PCI 익스프레스 스위치 장치가 외부 디바이스의 접속을 제어하는 방법에 있어서,

상기 PCI 익스프레스 스위치 장치는 PCI 케이블을 통해 상기 외부 디바이스를 상기 컴퓨터 시스템에 접속하고, 상기 PCI 케이블은 상기 외부 디바이스와 데이터 송수신을 수행하는 최대 레인의 개수보다 많은 개수의 레인에 해당하는 케이블을 구비하며,

상기 PCI 익스프레스 스위치 장치는 상기 컴퓨터 시스템에 접속되는 제1 가상 브리지와 상기 외부 디바이스에 접속되는 제2 가상 브리지를 포함하고,

상기 제1 가상 브리지와 상기 제2 가상 브리지가 상호간에 혼련 시퀀스를 송신하여 장애가 발생된 레인을 확인

하는 단계;

장애가 발생된 레인이 확인되는 경우, 상기 장애가 발생된 레인을 제외하고 나머지 레인으로 링크 폭을 결정하는 단계; 및

결정되는 링크 폭을 통해 상기 컴퓨터 시스템과 상기 외부 디바이스와의 PCI 익스프레스 방식에 따른 데이터 전송을 제어하는 단계

를 포함하는 접속 제어 방법.

청구항 10

제9항에 있어서,

상기 최대 레인의 개수가 N (여기서, N은 1 이상의 자연수임)이고, 상기 PCI 케이블이 구비하는 케이블의 개수가 N+n (여기서, n은 최대 N인 자연수임)인 경우, 상기 장애가 발생된 레인을 확인하는 단계에서, 상기 제1 가상 브리지와 상기 제2 가상 브리지는 N+n개의 케이블을 통해 혼련 시퀀스를 송신하여 장애가 발생된 레인을 확인하는 것을 특징으로 하는 접속 제어 방법.

청구항 11

제10항에 있어서,

상기 장애가 발생된 레인을 확인하는 단계는,

상기 제1 가상 브리지에 의해 형성되는 다운스트림 레인이 설정된 링크 값과 PAD의 레인 값을 가지는 TS1을 상기 N+n개의 레인을 통해 상기 제2 가상 브리지에 의해 형성되는 업스트림 레인으로 송신하는 단계;

상기 업스트림 레인이 모두 PAD로 설정된 링크 및 레인 값을 가지는 TS1을 상기 N+n개의 레인을 통해 상기 다운스트림 레인으로 송신하는 단계;

상기 업스트림 레인은 상기 다운스트림 레인으로부터 수신한 링크 값과 PAD의 레인 값을 가지는 TS1을 TS1을 수신한 레인에 송신하되, 상기 업스트림 레인으로 송신하는 단계에서 TS1을 수신하지 못하여 장애가 발생된 레인으로 판단되는 레인으로는 모두 PAD의 링크 및 레인 값을 가지는 TS1을 송신하는 단계; 및

상기 다운스트림 레인은 상기 다운스트림 레인으로 송신하는 단계에서 TS1을 수신하지 못하여 장애가 발생된 레인으로 판단되는 레인을 확인하는 단계

를 포함하는 접속 제어 방법.

청구항 12

제11항에 있어서,

상기 링크 폭을 결정하는 단계는,

상기 다운스트림 레인이 상기 N+n개의 레인 중에서 장애가 발생된 레인을 제외하고 일련의 레인 값을 가지는 TS1을 송신하는 단계;

상기 업스트림 레인이 상기 N+n개의 레인 중에서 장애가 발생된 레인을 제외하고 일련의 레인 값을 가지는 TS1을 송신하는 단계;

상기 다운스트림 레인이 상기 업스트림 레인으로부터 송신되는 레인 값에 기초하여 상기 다운스트림 레인에서 사용 가능한 레인을 확인하는 단계; 및

상기 업스트림 레인은 상기 다운스트림 레인으로부터 송신되는 레인 값에 기초하여 상기 업스트림 레인에서 사용 가능한 레인을 확인하는 단계

를 포함하는 접속 제어 방법.

청구항 13

제12항에 있어서,

상기 데이터 전송을 제어하는 단계는,

상기 다운스트림 레인이 상기 다운스트림 레인에서 사용 가능한 레인에 대해 일련의 레인 값을 가지는 TS1을 송신하는 단계;

상기 업스트림 레인이 상기 업스트림 레인에서 사용 가능한 레인에 대해 일련의 레인 값을 가지는 TS1을 송신하는 단계;

상기 다운스트림 레인은 자신이 송신한 값과 일치하는 레인 값을 가지는 TS1을 상기 업스트림 레인으로부터 수신하는 경우 설정된 레인 값을 가지는 TS2를 송신하는 단계; 및

상기 다운스트림 레인과 상기 업스트림 레인이 TS2를 통해 설정된 레인 값을 갖는 레인을 통해서 상기 컴퓨터 시스템과 상기 외부 디바이스 사이의 데이터 전송을 제어하는 단계

를 포함하는 접속 제어 방법.

청구항 14

제13항에 있어서,

상기 다운스트림 레인과 상기 업스트림 레인은 각각 레인별로 LTSSM(Link Training Sequence State Machine)을 구비하고,

상기 LTSSM은 Configuration.Linkwidth.Start 상태, Configuration.Linkwidth.Accept 상태, Configuration.Lanenum.Wait 상태, Configuration.Lanenum.Accept 상태 및 Configuration.Complete 상태를 포함하며,

상기 다운스트림 레인과 상기 업스트림 레인 상호간에 수신되는 TS1에 의해 상태의 천이가 발생하는 것을 특징으로 하는 접속 제어 방법.

청구항 15

제14항에 있어서,

상기 다운스트림 레인 및 상기 업스트림 레인은 상기 Configuration.Linkwidth.Start 상태 및 Configuration.Linkwidth.Accept 상태의 천이를 통해 장애가 발생된 레인을 확인하는 것을 특징으로 하는 접속 제어 방법.

청구항 16

제14항에 있어서,

상기 다운스트림 레인 및 상기 업스트림 레인은 상기 Configuration.Lanenum.Wait 상태 및 Configuration.Lanenum.Accept 상태의 천이를 통해 링크 폭에 해당하는 레인을 결정하는 것을 특징으로 하는 접속 제어 방법.

청구항 17

제16항에 있어서,

상기 다운스트림 레인은 Configuration.Complete 상태에서 상기 링크 폭에 해당하는 레인 정보를 TS2에 포함시켜서 상기 업스트림 레인으로 전송하여 상기 다운스트림 레인과 상기 업스트림 레인 사이의 링크 폭 협상을 완료하는 것을 특징으로 하는 접속 제어 방법.

발명의 설명

기술 분야

본 발명은 PCI 익스프레스(Peripheral Component Interconnect Express, 이하 "PCI 익스프레스" 또는 "PCIe"라고 함) 스위치 장치 및 그의 접속 제어 방법에 관한 것이다.

[0001]

배경 기술

- [0002] PCI 익스프레스 기술은 컴퓨터 시스템 내부에서 CPU와 주변기기간의 입출력(I/O) 인터페이스를 하기 위한 기술이다. 이러한 PCI 익스프레스 기술은, 초기에는 칩-투-칩 또는 슬롯을 이용한 인터페이스에 적용되었으나, 최근에는 PCI 익스프레스 인터페이스를 외부로 확장하기 위하여, PCI 익스프레스 케이블 또는 기타 구리 및 광케이블을 이용하여 확장하는 방법이 도입되었다. 이로 인해, 컴퓨터 시스템 내부뿐만 아니라 시스템 외부의 확장 인터페이스로도 사용 가능하게 되었다. 특히, 최근에는 PCI 익스프레스 신호를 썬더볼트 케이블이라는 새로운 구리 케이블을 이용하여 비디오 신호와 함께 전송하는 방법이 도입되었다.
- [0003] PCI 익스프레스 신호를 케이블로 전송하는 종래 기술에서는 대개 PCI 익스프레스 스위치 등을 이용하여 PCI 익스프레스 규격으로 정의된 신호를 그대로 사용하고 있다. 그런데, PCI 익스프레스의 다중레인 접속 제어 방법의 특성상, 다중 케이블을 사용하여 신호를 송수신하거나 다중 신호를 송수신할 수 있도록 여러 개의 경로를 가지는 케이블을 사용할 경우에, 각 레인별 다중 신호 송수신 경로 중 하나라도 문제가 발생하면, 송수신 속도가 급격히 감소하거나, 심지어는 송수신이 전혀 불가능하게 될 수도 있다. 특히, 전송 경로가 짧은 경우에는 문제가 발생하였을 때 문제를 발생시킨 케이블이나 디바이스 교체 등으로 문제를 해결할 수 있으나, 전송 경로가 100m ~ 10km 등으로 원거리로 전송할 때에는 디바이스 교체는 가능하나, 케이블 자체의 교체에는 많은 어려움이 뒤따르므로, 서비스를 부득이하게 중단하게 될 경우도 발생한다.
- [0004] 그러나, 기존의 PCI 익스프레스 스위치 등을 이용하는 방법에서는 PCI 익스프레스 신호 자체를 케이블에서 송수신하는 방법을 채택하고 있으므로, 일부 전송경로에 장애가 발생할 경우 이에 대한 대비책이 전혀 없는 실정이다. 이로 인해 해결책이 필요하다.

발명의 내용

해결하려는 과제

- [0005] 본 발명이 이루고자 하는 기술적 과제는 PCI 익스프레스 상호 운용성을 위하여 PCI 익스프레스 규격을 만족하면서도 PCI 익스프레스 다중 신호를 외부로 확장함에 있어서 다중 신호 확장 케이블 중 일부에 문제가 발생하는 경우에도 지속적인 서비스를 제공할 수 있도록 하는 PCI 익스프레스 스위치 장치 및 그의 접속 제어 방법을 제공하는 것이다.

과제의 해결 수단

- [0006] 본 발명의 하나의 특징에 따른 PCI 익스프레스 스위치 장치는,
- [0007] 제1 PCI 익스프레스(Peripheral Component Interconnect Express) 포트를 통해 컴퓨터 시스템에 접속되어 외부 디바이스와의 PCI 방식에 따른 데이터 송수신을 수행하는 제1 가상 브리지; 제2 PCI 익스프레스 포트를 통해 상기 외부 디바이스에 접속되며, 상기 제1 가상 브리지와 협력하여 상기 외부 디바이스가 상기 컴퓨터 시스템과 PCI 방식에 따른 데이터 송수신이 가능하도록 하는 제2 가상 브리지; 상기 제1 가상 브리지에 접속되는 제1 케이블 정합 장치; 및 상기 제2 가상 브리지에 접속되며, 상기 제1 케이블 정합 장치와 PCI 케이블을 통해 접속되는 제2 케이블 정합 장치를 포함하고, 상기 PCI 케이블은 상기 제1 가상 브리지와 상기 제2 가상 브리지가 데이터 송수신을 수행하는 최대 레인의 개수보다 많은 개수의 레인에 해당하는 케이블을 구비하는 것을 특징으로 한다.
- [0008] 여기서, 상기 최대 레인의 개수가 N (여기서, N은 1 이상의 자연수임)이고, 상기 PCI 케이블이 구비하는 케이블의 개수가 N+n (여기서, n은 최대 N인 자연수임)인 경우, 상기 제1 가상 브리지는 상기 제2 가상 브리지와 훈련 시퀀스(Training Sequence)를 송신하여 N+n 케이블 중에서 최대 N개의 레인을 결정하는 것을 특징으로 한다.
- [0009] 또한, 상기 제1 가상 브리지 및 제2 가상 브리지는 PCI 익스프레스 계층 구조로 이루어지는 것을 특징으로 한다.
- [0010] 또한, 상기 제1 가상 브리지는, 상기 컴퓨터 시스템으로 물리적으로 패킷을 전송하기 위해 상기 제1 PCI 익스프레스 포트에 접속되는 제1 업스트림 레인 물리 계층부; 상기 외부 디바이스로 물리적으로 패킷을 전송하기 위해 상기 제1 케이블 정합 장치에 접속되는 제1 다운스트림 레인 물리 계층부; 상기 업스트림 레인 물리 계층부와 상기 다운스트림 레인 물리 계층부를 통해 송수신되는 패킷의 조립 및 분해를 수행하는 제1 트랜잭션 계층; 상기 트랜잭션 계층부와 상기 업스트림 레인 물리 계층부 사이의 데이터 전달 구성요소로 작용하며, 상기 트랜잭

선 계층부와 상기 업스트림 레인 물리 계층부 사이에서 패킷을 교환하기 위한 신뢰할 수 있는 메커니즘에 링크를 제공하는 제1 데이터 링크 계층부; 및 상기 트랜잭션 계층부와 상기 다운스트림 레인 물리 계층부 사이의 데이터 전달 구성요소로 작용하며, 상기 트랜잭션 계층부와 상기 다운스트림 레인 물리 계층부 사이에서 패킷을 교환하기 위한 신뢰할 수 있는 메커니즘에 링크를 제공하는 제1 케이블 전송 데이터 링크 계층부를 포함한다.

[0011] 또한, 상기 제2 가상 브리지는, 상기 외부 디바이스로 물리적으로 패킷을 전송하기 위해 상기 제2 PCI 익스프레스 포트에 접속되는 제2 다운스트림 레인 물리 계층부; 상기 컴퓨터 시스템으로 물리적으로 패킷을 전송하기 위해 상기 제2 케이블 정합 장치에 접속되는 제2 업스트림 레인 물리 계층부; 상기 다운스트림 레인 물리 계층부와 상기 다운스트림 레인 물리 계층부를 통해 송수신되는 패킷의 조립 및 분해를 수행하는 제2 트랜잭션 계층; 상기 제2 트랜잭션 계층부와 상기 제2 업스트림 레인 물리 계층부 사이의 데이터 전달 구성요소로 작용하며, 상기 제2 트랜잭션 계층부와 상기 제2 업스트림 레인 물리 계층부 사이에서 패킷을 교환하기 위한 신뢰할 수 있는 메커니즘에 링크를 제공하는 제2 데이터 링크 계층부; 및 상기 제2 트랜잭션 계층부와 상기 제2 다운스트림 레인 물리 계층부 사이의 데이터 전달 구성요소로 작용하며, 상기 제2 트랜잭션 계층부와 상기 제2 다운스트림 레인 물리 계층부 사이에서 패킷을 교환하기 위한 신뢰할 수 있는 메커니즘에 링크를 제공하는 제2 케이블 전송 데이터 링크 계층부를 포함한다.

[0012] 또한, 상기 제1 가상 브리지와 상기 제2 가상 브리지는 N+n개의 상기 PCI 케이블을 통해 상호 혼련 시퀀스를 송신하여 각각 송신 가능한 레인을 확인하는 것을 특징으로 한다.

[0013] 또한, 상기 제1 가상 브리지는 상기 제2 가상 브리지는 상기 N개의 레인 중에서 장애가 발생된 레인이 있는 경우 추가 n개의 레인 중에서 사용 가능한 레인으로 대체하여 상기 N개의 레인을 구성하는 것을 특징으로 한다.

[0014] 또한, 상기 제1 다운스트림 레인 물리 계층부와 상기 제2 업스트림 레인 물리 계층부 각각은 N+n개의 케이블에 대응되는 레인별로 LTSSM(Link Training Sequence State Machine)을 구비하고, 상기 레인별 LTSSM에 의해 수신되는 혼련 시퀀스의 상태와 상태 천이를 상기 레인별 LTSSM이 공유하도록 하기 위해 상기 레인별 LTSSM에 공통으로 연결되는 버스를 포함하는 것을 특징으로 한다

[0015] 본 발명의 다른 특징에 따른 접속 제어 방법은,

[0016] 컴퓨터 시스템과 외부 디바이스를 PCI 익스프레스 방식에 의해 접속하는 PCI 익스프레스 스위치 장치가 외부 디바이스의 접속을 제어하는 방법으로서, 상기 PCI 익스프레스 스위치 장치는 PCI 케이블을 통해 상기 외부 디바이스를 상기 컴퓨터 시스템에 접속하고, 상기 PCI 케이블은 상기 외부 디바이스와 데이터 송수신을 수행하는 최대 레인의 개수보다 많은 개수의 레인에 해당하는 케이블을 구비하며, 상기 PCI 익스프레스 스위치 장치는 상기 컴퓨터 시스템에 접속되는 제1 가상 브리지와 상기 외부 디바이스에 접속되는 제2 가상 브리지를 포함하고, 상기 제1 가상 브리지와 상기 제2 가상 브리지가 상호간에 혼련 시퀀스를 송신하여 장애가 발생된 레인을 확인하는 단계; 장애가 발생된 레인이 확인되는 경우, 상기 장애가 발생된 레인을 제외하고 나머지 레인으로 링크 폭을 결정하는 단계; 및 결정되는 링크 폭을 통해 상기 컴퓨터 시스템과 상기 외부 디바이스와의 PCI 익스프레스 방식에 따른 데이터 전송을 제어하는 단계를 포함한다.

[0017] 여기서, 상기 최대 레인의 개수가 N (여기서, N은 1 이상의 자연수임)이고, 상기 PCI 케이블이 구비하는 케이블의 개수가 N+n (여기서, n은 최대 N인 자연수임)인 경우, 상기 장애가 발생된 레인을 확인하는 단계에서, 상기 제1 가상 브리지와 상기 제2 가상 브리지는 N+n개의 케이블을 통해 혼련 시퀀스를 송신하여 장애가 발생된 레인을 확인하는 것을 특징으로 한다.

[0018] 또한, 상기 장애가 발생된 레인을 확인하는 단계는, 상기 제1 가상 브리지에 의해 형성되는 다운스트림 레인이 설정된 링크 값과 PAD의 레인 값을 가지는 TS1을 상기 N+n개의 레인을 통해 상기 제2 가상 브리지에 의해 형성되는 업스트림 레인으로 송신하는 단계; 상기 업스트림 레인이 모두 PAD로 설정된 링크 및 레인 값을 가지는 TS1을 상기 N+n개의 레인을 통해 상기 다운스트림 레인으로 송신하는 단계; 상기 업스트림 레인은 상기 다운스트림 레인으로부터 수신한 링크 값과 PAD의 레인 값을 가지는 TS1을 수신한 레인에 송신하되, 상기 업스트림 레인으로 송신하는 단계에서 TS1을 수신하지 못하여 장애가 발생된 레인으로 판단되는 레인으로는 모두 PAD의 링크 및 레인 값을 가지는 TS1을 송신하는 단계; 및 상기 다운스트림 레인은 상기 다운스트림 레인으로 송신하는 단계에서 TS1을 수신하지 못하여 장애가 발생된 레인으로 판단되는 레인을 확인하는 단계를 포함한다.

[0019] 또한, 상기 링크 폭을 결정하는 단계는, 상기 다운스트림 레인이 상기 N+n개의 레인 중에서 장애가 발생된 레인을 제외하고 일련의 레인 값을 가지는 TS1을 송신하는 단계; 상기 업스트림 레인이 상기 N+n개의 레인 중에서 장애가 발생된 레인을 제외하고 일련의 레인 값을 가지는 TS1을 송신하는 단계; 상기 다운스트림 레인이 상기

업스트림 레인으로부터 송신되는 레인 값에 기초하여 상기 다운스트림 레인에서 사용 가능한 레인을 확인하는 단계; 및 상기 업스트림 레인은 상기 다운스트림 레인으로부터 송신되는 레인 값에 기초하여 상기 업스트림 레인에서 사용 가능한 레인을 확인하는 단계를 포함한다.

[0020] 또한, 상기 데이터 전송을 제어하는 단계는, 상기 다운스트림 레인이 상기 다운스트림 레인에서 사용 가능한 레인에 대해 일련의 레인 값을 가지는 TS1을 송신하는 단계; 상기 업스트림 레인이 상기 업스트림 레인에서 사용 가능한 레인에 대해 일련의 레인 값을 가지는 TS1을 송신하는 단계; 상기 다운스트림 레인은 자신이 송신한 값과 일치하는 레인 값을 가지는 TS1을 상기 업스트림 레인으로부터 수신하는 경우 설정된 레인 값을 가지는 TS2를 송신하는 단계; 및 상기 다운스트림 레인과 상기 업스트림 레인이 TS2를 통해 설정된 레인 값을 갖는 레인을 통해서 상기 컴퓨터 시스템과 상기 외부 디바이스 사이의 데이터 전송을 제어하는 단계를 포함한다.

[0021] 또한, 상기 다운스트림 레인과 상기 업스트림 레인은 각각 레인별로 LTSSM(Link Training Sequence State Machine)을 구비하고, 상기 LTSSM은 Configuration.Linkwidth.Start 상태, Configuration.Linkwidth.Accept 상태, Configuration.Lanenum.Wait 상태, Configuration.Lanenum.Accept 상태 및 Configuration.Complete 상태를 포함하며, 상기 다운스트림 레인과 상기 업스트림 레인 상호간에 수신되는 TS1에 의해 상태의 천이가 발생하는 것을 특징으로 한다.

[0022] 또한, 상기 다운스트림 레인 및 상기 업스트림 레인은 상기 Configuration.Linkwidth.Start 상태 및 Configuration.Linkwidth.Accept 상태의 천이를 통해 장애가 발생한 레인을 확인하는 것을 특징으로 한다.

[0023] 또한, 상기 다운스트림 레인 및 상기 업스트림 레인은 상기 Configuration.Lanenum.Wait 상태 및 Configuration.Lanenum.Accept 상태의 천이를 통해 링크 폭에 해당하는 레인을 결정하는 것을 특징으로 한다.

[0024] 또한, 상기 다운스트림 레인은 Configuration.Complete 상태에서 상기 링크 폭에 해당하는 레인 정보를 TS2에 포함시켜서 상기 업스트림 레인으로 전송하여 상기 다운스트림 레인과 상기 업스트림 레인 사이의 링크 폭 협상을 완료하는 것을 특징으로 한다.

발명의 효과

[0025] 본 발명에 따르면, 레인이 추가되더라도 레인의 순서와 상관없이 가용한 레인의 개수를 모두 이용하여 링크 폭을 결정할 수 있으므로, 케이블과 같이 임의의 부분에 문제가 발생하더라도, 레인을 최대한 이용하여 대역폭을 보장할 수 있다.

도면의 간단한 설명

[0026] 도 1은 PCI 익스프레스 스위치 칩셋을 이용하여 케이블 연장을 달성하는 일반적인 방법에 따른 구조를 도시한 도면이다.

도 2는 도 1에 도시된 PCI 익스프레스 스위치 칩셋의 내부 구조를 도시한 도면이다.

도 3은 일반적인 링크 트레이닝 시퀀스 상태 머신의 동작을 설명하는 도면이다.

도 4는 일반적인 PCI 익스프레스 LTSSM에서 업스트림 레인의 순서와 다운스트림 레인의 순서가 같은 경우의 구성 상태 천이를 도시한 도면이다.

도 5는 본 발명의 실시예에 따른 PCI 익스프레스 스위치 장치의 구성을 도시한 도면이다.

도 6은 도 5에 도시된 가상 브리지의 구체적인 구성을 도시한 도면이다.

도 7은 도 6에 도시된 가상 브리지의 구체적인 구성을 도시한 도면이다.

도 8은 본 발명의 실시예에 따른 PCI 익스프레스 스위치 장치에서 x8 레인을 적용한 예를 도시한 도면이다.

도 9는 본 발명의 실시예에 따른 PCI 익스프레스 스위치 장치에서 링크 폭을 결정하기 위한 협상 과정을 도시한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0027] 아래에서는 첨부한 도면을 참고로 하여 본 발명의 실시예에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 그리고 도면에서 본 발명을 명확하게 설명하기 위

해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

- [0028] 명세서 전체에서, 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다. 또한, 명세서에 기재된 "...부", "...기", "모듈" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.
- [0029] 먼저, 도 1을 참조하여 PCI 익스프레스 스위치 칩셋을 이용하여 케이블 연장을 달성하는 방법에 대해 설명한다.
- [0030] 도 1은 PCI 익스프레스 스위치 칩셋을 이용하여 케이블 연장을 달성하는 일반적인 방법에 따른 구조를 도시한 도면이다.
- [0031] 도 1을 참조하면, PCI 익스프레스 신호를 외부로 확장하기 위해서 컴퓨터 시스템(10)의 내부 및 외부에 각각 PCI 익스프레스 스위치 칩셋(11, 20)을 사용하고, 이들 PCI 익스프레스 스위치 칩셋(11)과 PCI 익스프레스 스위치 칩셋(20)간을 xN 포트 to xN 케이블 정합장치(30, 40)를 이용하여 xN 케이블(50)로 접속한다. 여기서 N은 PCI 익스프레스 인터페이스의 레인의 개수로 1, 2, 4, 8, 12, 16, 32의 값을 가질 수 있다.
- [0032] PCI 익스프레스 스위치 칩셋(11, 20)에 존재하는 PCI 익스프레스 xN 포트(12, 13, 21, 22)는 PCI 익스프레스 규격과 호환성을 가져야 하므로, 첫 번째 레인 혹은 마지막 레인은 반드시 동작 가능하여야 하고, 정상 동작인 레인 역시 연속성을 가져야 한다.
- [0033] 예를 들어, 레인이 총 8개인 x8 레인 각각을 1레인, 2레인, 3레인, 4레인, 5레인, 6레인, 7레인, 8레인이라고 한다면, x8 동작을 위해서는 1~8레인이 모두 동작 가능하여야 하고, x4 동작을 위해서는 1~4레인이 모두 동작 가능하거나 5~8레인이 모두 동작 가능하여야 한다. 또한, x2 동작을 위해서는 1~2레인이 모두 동작 가능하거나 7~8레인이 모두 동작 가능하여야 한다. 만약, 2레인이 동작할 수 없다면, 동작 가능한 레인인 5~8레인을 이용하여 x4의 동작이 가능하며, 7레인이 동작할 수 없다면, 동작 가능한 레인인 1~4레인을 이용하여 x4의 동작이 가능하다. 다른 예로 3레인이나 6레인이 동작할 수 없다면, 1~2레인을 이용하거나 7~8레인을 이용하여 x2의 동작이 가능하다. 더 심한 경우, 2레인과 7레인이 동작하지 않는다면 동작하는 레인이 1, 3, 4, 5, 6, 8레인의 총 6개임에도 불구하고, 1~2레인과 7~8레인이 모두 동작 가능하지 않으므로, x2의 동작도 불가능하여 1레인 혹은 8레인만을 사용하는 x1의 동작만이 가능하다.
- [0034] 종래의 PCI 익스프레스 스위치 칩셋(11, 20)은 케이블 연장을 고려하여 만들어져 있지 않으므로, PCI 익스프레스 스위치 칩셋(11, 20)을 이용하여 케이블 연장을 달성하고자 한다면 호환성 유지를 위하여 위의 제약조건을 그대로 가지고 있어야 하며, 이는 변경될 수 없다.
- [0035] 도 2는 도 1에 도시된 PCI 익스프레스 스위치 칩셋(11, 20)의 내부 구조를 도시한 도면이다. 여기서, PCI 익스프레스 스위치 칩셋(11, 20)은 그 구조가 동일하므로, 하나의 PCI 익스프레스 스위치 칩셋(20)에 대해서만 설명한다.
- [0036] 도 2를 참조하면, PCI 익스프레스 스위치 칩셋(20)은 두 개의 PCI 익스프레스 xN 포트(21, 22), 두 개의 가상 브리지(23, 24) 및 내부 버스(25)를 포함한다.
- [0037] 가상 브리지(23)와 가상 브리지(24)를 연결하는 내부 버스(25)는 PCI 익스프레스 규격에서 정의되어 있지 않으며, 구현자에 따라 구현 방법을 임의로 결정할 수 있다.
- [0038] 따라서, 본 발명의 실시예에서는 이러한 내부 버스(25)를 이용하여 케이블 확장을 구현한다.
- [0039] 도 3은 일반적인 링크 트레이닝 시퀀스 상태 머신(Link Training Sequence State Machine, 이하 "LTSSM"이라 함)의 동작을 설명하는 도면이다.
- [0040] LTSSM은 검출 상태, 폴링 상태, 구성 상태, L0 상태 및 복원 상태를 포함하지만, 도 3에서는, PCI 익스프레스 규격에서 링크(Link)의 상태를 초기 설정하고 관리하는 LTSSM 중, 특히 레인(Lane)을 설정하는 단계를 나타내는 구성 상태를 나타내며, 링크가 정상적으로 동작하는 L0 상태까지의 진행 방법을 나타낸 것이다.
- [0041] 실제 상태 천이(status transition)가 발생하는 조건은 레인이 업스트림 레인(Upstream Lane)이나 다운스트림 레인(Downstream Lane)이냐에 따라 다르게 되는데, 이는 링크의 설정의 주체가 다운스트림 레인이기 때문이다.
- [0042] x8의 레인 폭을 가지는 PCI 익스프레스 링크 예에서 업스트림 레인에서 설정하는 레인과 다운스트림 레인에서

설정하는 라인 순서가 일치하는 경우 PCI 익스프레스 규격에서 정의된 링크 설정을 위한 상태 천이 순서는 다음의 도 4와 같다.

- [0043] 도 4는 일반적인 PCI 익스프레스 LTSSM에서 업스트림 레인의 순서와 다운스트림 레인의 순서가 같은 경우의 구성 상태 천이를 도시한 도면이다.
- [0044] 도 4를 참조하면, Configuration.Linkwidth.Start의 상태(61)의 다운스트림 레인은 설정된 링크(Link) 값과 패드(PAD)의 라인 값을 가지는 훈련 시퀀스(Training Sequence, 이하 "TS"라 함)1을 수신기를 감지한 모든 라인에 송신한다(S1).
- [0045] 업스트림 레인은 모두 PAD로 설정된 링크 및 라인 값을 가지는 TS1을 수신기를 감지한 모든 라인에 송신한다(S2).
- [0046] PAD의 링크 값을 가지는 연속 2개의 TS1을 아무 라인에서건 수신한 다운스트림 레인은 그 후에, Configuration.Linkwidth.Accept 상태(62)로 천이한다.
- [0047] Configuration.Linkwidth.Start 상태(71)의 업스트림 레인은 설정된 링크 번호와 PAD의 라인 번호를 가지는 TS1을 아무 라인에서건 2개 연속 수신한 후, Configuration.Linkwidth.Accept 상태(72)로 천이한다.
- [0048] 그 후, 업스트림 레인은 수신한 링크 값과 PAD 라인 값을 가지는 TS1을 TS1을 수신한 라인에 송신하며, TS1을 수신하지 못한 라인에는 모두 PAD의 링크 및 라인 값을 가지는 TS1을 송신한다(S3).
- [0049] Configuration.Linkwidth.Accept 상태(62)의 다운스트림 레인은 자신이 송신한 링크 값과, PAD의 라인 값을 가지는 TS1을 아무 라인에서건 수신한 후, Configuration.Lanenum.Wait의 상태(63)로 천이한다.
- [0050] 그 후, 다운스트림 레인은 이전의 링크 값과 원하는 라인 값을 가지는 TS1을 두 번째 라인에서 수신하지 못한 경우에는 첫 번째 라인에만 송신한다(S4). 그러나, 세 번째나 네 번째 라인에서 상기한 TS1을 수신하지 못한 경우, x2 설정이 가능한 경우에는 첫 번째와 두 번째 라인에 송신하고, 그렇지 않으면 첫 번째 라인에만 송신한다(S4). 또한, 다섯 번째 내지 여덟 번째 라인 중에서 상기한 TS1을 수신하지 못한 경우, x4 설정이 가능한 경우에는 첫 번째 내지 네 번째 라인에 송신하고, x2 설정이 가능한 경우에는 첫 번째와 두 번째 라인에 송신하며, x1 설정만이 가능한 경우에는 첫 번째 라인에만 송신한다(S4).
- [0051] 그러나, 모든 라인에서 상기한 TS1을 수신한 경우에는 모든 라인에 송신하며, 첫 번째 라인에서 상기한 TS1을 수신하지 못한 경우에는 동작을 중지하게 된다. 이 때 라인 값은 0에서 시작하여 연속되는 값을 가지게 되며, 나머지 라인에는 링크와 라인 값이 모두 PAD인 TS1을 송신한다(S4).
- [0052] Configuration.Linkwidth.Accept 상태(72)의 업스트림 레인은 설정된 링크 값과, PAD가 아닌 라인 값을 가지는 2개 연속의 TS1을 아무 라인에서건 수신한 후, Configuration.Lanenum.Wait 상태(73)로 천이한다.
- [0053] 그 후, 업스트림 레인은 수신한 TS1과 동일한 링크 및 라인 값을 가지는 TS1을 두 번째 라인에서 수신하지 못한 경우에는 첫 번째 라인에만 송신한다(S5). 그러나, 세 번째나 네 번째 Lane에서 상기한 TS1을 수신하지 못한 경우, x2 설정이 가능한 경우에는 첫 번째와 두 번째 라인에 송신하고, 그렇지 않으면 첫 번째 라인에만 송신한다(S5). 또한, 다섯 번째 내지 여덟 번째 라인 중에서 상기한 TS1을 수신하지 못한 경우, x4 설정이 가능한 경우에는 첫 번째 내지 네 번째 라인에 송신하고, x2 설정이 가능한 경우에는 첫 번째와 두 번째 라인에 송신하며, x1 설정만이 가능한 경우에는 첫 번째 라인에만 송신한다(S5). 그러나, 모든 라인에서 상기한 TS1을 수신한 경우에는 모든 라인에 송신하며, 첫 번째 라인에서 상기한 TS1을 수신하지 못한 경우에는 동작을 중지한다. 이 때 라인 값은 0에서 시작하여 연속되는 값을 가지게 되며, 나머지 라인에는 링크와 라인 값이 모두 PAD인 TS1을 송신한다(S5).
- [0054] Configuration.Lanenum.Wait 상태(63)의 다운스트림 레인은 처음 Configuration.Lanenum.Wait 상태(63)에 진입 시의 라인 값과 일치하지 않는 라인 값을 가지는 TS1을 아무 라인에서건 2개 연속 수신하거나, 또는 라인의 링크 값 중 적어도 하나가 아닌 경우나 자신이 송신한 것과 일치하는 링크 및 라인 값을 가지는 TS1을 수신하는 경우에, Configuration.Lanenum.Accept 상태(64)로 천이한다.
- [0055] 그리고, 다운스트림 레인은 자신이 송신한 것과 일치하는 링크 값과 라인 값을 가지는 TS1을 설정된 폭의 라인에서 2개 연속 수신하였으므로, Configuration.Complete 상태(65)로 천이하여 링크 및 라인 값이 설정된 TS2를 송신한다(S6).
- [0056] 한편, Configuration.Lanenum.Wait 상태(73)의 업스트림 레인은 TS2를 설정된 폭의 라인에서 2개 연속 수신한

후에, Configuration.Lanenum.Accept(74)로 천이하고, TS2를 설정된 폭의 레인에서 2개 연속 수신하였으므로, 이어서 Configuration.Complete 상태(75)로 진입한다.

- [0057] 위의 방식에서 총 8개의 레인이 존재하지만, 만약 두 번째 레인에 장애가 발생하는 경우 7개의 레인이 사용 가능함에도 불구하고 1개의 레인 폭으로만 설정이 가능하다. 또한, 세 번째나 네 번째 레인에 장애가 발생하는 경우에는, 6~7개의 레인이 사용 가능함에도 불구하고, 최대 2개의 레인 폭으로만 설정이 가능하다는 문제점이 존재한다.
- [0058] 한편, 업스트림 레인의 순서와 다운스트림 레인이 서로 역순인 경우에는 업스트림 레인과 다운스트림 레인 둘 다 또는 어느 한쪽에서 역순의 레인에 대한 처리를 할 수 있어야 한다. x8의 폭을 예를 들면 첫 번째나 마지막 레인 둘 중 하나는 반드시 동작이 가능하여야 하고, 두 번째나 마지막에서 두 번째 레인에서 장애가 발생할 경우에는 x1의 동작만이 가능하다. 또한, 세 번째나 네 번째 혹은 마지막에서 세 번째나 네 번째 레인에서 장애가 발생할 경우에는 x2의 동작만이 가능하며, 다섯 번째에서 여덟 번째 레인이나 마지막에서 다섯 번째에서 여덟 번째 레인에서 장애가 발생할 경우에는 x4의 동작만이 가능하다는 문제점을 가지고 있다.
- [0059] 이하 상기한 문제점을 해결하기 위한 본 발명의 PCI 익스프레스 스위치 장치에 대해 설명한다.
- [0060] 도 5는 본 발명의 실시예에 따른 PCI 익스프레스 스위치 장치의 구성을 도시한 도면이다.
- [0061] 도 5에 도시된 바와 같이, 본 발명의 실시예에 따른 PCI 익스프레스 스위치 장치는 PCI 익스프레스(PCIe) xN 포트(110, 210), xN(+n) 케이블 정합 장치(120, 220) 및 가상 브리지(130, 230)를 포함한다. 여기서, PCI 익스프레스 xN 포트(110)와 가상 브리지(130)는 컴퓨터 시스템(100)의 내부에 위치하는 것이 바람직하다.
- [0062] PCI 익스프레스 xN 포트(110)는 컴퓨터 시스템(100)을 제어하는 도 1에 도시된 바와 같은 호스트 브리지 칩셋(도시되지 않음)에 접속된다.
- [0063] PCI 익스프레스 xN 포트(210)는 PCI 익스프레스 방식을 통해 컴퓨터 시스템(100)에 연결되는 외부 PCI 익스프레스 디바이스(도시되지 않음)에 접속된다.
- [0064] xN(+n) 케이블 정합 장치(120)는 가상 브리지(130)에 접속되며 xN(+n) 케이블(300)을 통해 외부 디바이스측의 xN(+n) 케이블 정합 장치(220)에 연결된다. 이 때, xN(+n) 케이블(300)은 총 N+n개의 케이블로 구성됨을 의미하며, 컴퓨터 시스템(100)과 외부의 디바이스가 최대 N(여기서 N은 1, 2, 4, 8, 12, 16, 32 등의 값을 가짐)개의 레인을 사용하여 PCI 익스프레스 방식에 따라 데이터 전송이 가능한 경우 N개의 케이블 외에 추가로 n개의 케이블을 더 구비하고 있다. 여기서, n의 최대값은 N이다. 즉, xN(+n) 케이블(300)은 최소 N+1개의 케이블부터 최대 2N개의 케이블로 구성될 수 있다. 여기서, 추가의 n개의 케이블은 기본적인 N개의 케이블로 구성되는 N개의 레인에서 장애가 발생하는 경우 장애가 발생된 케이블을 대신하여 데이터 송수신이 가능하도록 하는 케이블이다.
- [0065] 따라서, xN(+n) 케이블 정합 장치(120)는 N+n개의 케이블로 구성되는 xN(+n) 케이블(300)을 통해 가상 브리지(130)와 xN(+n) 케이블 정합 장치(220) 사이에 데이터 전달을 수행한다.
- [0066] 가상 브리지(130)는 PCI 익스프레스 xN 포트(110)를 통해 호스트 브리지 칩셋에 연결되고, xN(+n) 케이블 정합 장치(120)를 통해서 외부 디바이스측의 xN(+n) 케이블 정합 장치(220)에 연결되어, 호스트 브리지 칩셋과 외부 디바이스 사이에 PCI 익스프레스 방식에 따른 데이터 송수신을 수행한다. 이 때, 가상 브리지(130)는 최대 N개의 레인을 통해서 데이터 송수신을 수행하는데, xN(+n) 케이블 정합 장치(120, 220) 사이에 연결되는 xN(+n) 케이블(300)의 N+n개의 케이블에서 데이터 송수신에 사용되는 최대 N개의 케이블을 선택하여 데이터 송수신에 사용한다. 즉, N개의 케이블 내에서 장애가 발생된 케이블이 있는 경우에는 추가의 n개의 케이블에서 장애가 없는 케이블을 사용하여 최대 N개의 레인을 구성할 수 있다. 이를 위해, 가상 브리지(130)는 N+n개의 케이블을 통해서 TS(Training Sequence) 데이터를 송수신하여 장애가 발생된 케이블을 확인하고 N+n개의 케이블 중에서 장애가 발생된 케이블을 제외한 케이블을 사용하여 데이터 송수신을 위한 레인을 구성한다. 이 때 구성되는 레인은 최대 N개가 된다.
- [0067] 이와 같이, 본 발명의 실시예에 따른 가상 브리지(130)는 N개의 케이블 외에 n개의 케이블을 추가로 구성하는 케이블 확장을 구현함으로써 PCI 익스프레스 규격을 벗어나지 않으면서도 레인에서 장애가 발생한 경우에도 지속적인 동작이 가능하다.
- [0068] 예를 들어, 최악의 경우 N개의 모든 케이블에 장애가 발생할 경우라도 추가의 n(여기서, n은 N이라고 가정함)개

의 케이블을 이용하여 데이터 송수신이 가능하도록 할 수 있다.

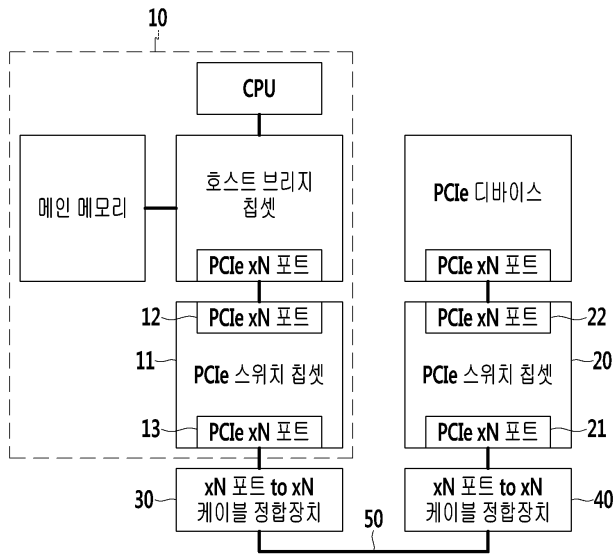
- [0069] xN(+n) 케이블 정합 장치(220)는 가상 브리지(230)에 접속되며 xN(+n) 케이블(300)을 통해 컴퓨터 시스템(100)측의 xN(+n) 케이블 정합 장치(120)에 연결된다. 이러한 xN(+n) 케이블 정합 장치(220)는 상기한 xN(+n) 케이블 정합 장치(120)와 그 구성 및 기능이 동일하다.
- [0070] 가상 브리지(230)는 PCI 익스프레스 xN 포트(210)를 통해 외부 디바이스에 연결되고, xN(+n) 케이블 정합 장치(220)를 통해서 컴퓨터 시스템(100)측의 xN(+n) 케이블 정합 장치(120)에 연결되어, 외부 디바이스와 호스트 브리지 칩셋 사이에 PCI 익스프레스 방식에 따른 데이터 송수신을 수행한다. 이러한 가상 브리지(230)도 상기한 가상 브리지(130)와 그 구성 및 기능이 동일하다. 즉, 가상 브리지(230)는 가상 브리지(130)와 N+n개의 케이블을 통한 TS 데이터 송수신 수행하여 가상 브리지(130)가 장애가 발생된 케이블을 확인하도록 하고, 가상 브리지(130)로부터 N+n개의 케이블 중에서 장애가 발생된 케이블을 제외한 유효한 케이블 정보를 수신하여 이들로 구성되는 레인을 통해 데이터 송수신을 수행한다. 이 때 구성되는 레인은 최대 N개가 된다.
- [0071] 도 6은 도 5에 도시된 가상 브리지(130)의 구체적인 구성을 도시한 도면이다.
- [0072] 도 6에 도시된 바와 같이, 가상 브리지(130)는 PCI 익스프레스 업스트림 레인 물리 계층(이하 "업스트림 레인 물리 계층"이라 함)(131), xN(+n) 케이블 전송 다운스트림 레인 물리 계층(이하 "다운스트림 레인 물리 계층"이라 함)(133), PCI 익스프레스 트랜잭션 계층(이하 "트랜잭션 계층"이라 함)(135), PCI 익스프레스 데이터 링크 계층(이하 "데이터 링크 계층"이라 함)(137) 및 케이블 전송 데이터 링크 계층(139)를 포함한다.
- [0073] 업스트림 레인 물리 계층(131)은 물리적으로 패킷을 컴퓨터 시스템(100)내 호스트 브리지 칩셋으로 전송하기 위해 PCI 익스프레스 xN 포트(110)에 접속된다.
- [0074] 다운스트림 레인 물리 계층(133)은 물리적으로 패킷을 외부 디바이스로 전송하기 위해 xN(+n) 케이블 정합 장치(120)에 접속된다.
- [0075] 트랜잭션 계층(135)은 업스트림 레인 물리 계층(131)과 다운스트림 레인 물리 계층(133)을 통해 송수신되는 패킷, 즉 트랜잭션 계층 패킷의 조립 및 분해를 수행한다. 이러한 트랜잭션 계층(135)에 대해서는 이미 잘 알려져 있으므로 구체적인 설명을 생략한다.
- [0076] 데이터 링크 계층(137)은 트랜잭션 계층(135)과 업스트림 레인 물리 계층(131) 사이의 중간 단계로 작용하며, 두 계층(131, 135) 사이에서 패킷을 교환하기 위한 신뢰할 수 있는 메커니즘에 링크를 제공한다.
- [0077] 케이블 전송 데이터 링크 계층(139)은 트랜잭션 계층(135)과 다운스트림 레인 물리 계층(133) 사이의 중간 단계로 작용하며, 두 계층(133, 135) 사이에서 패킷을 교환하기 위한 신뢰할 수 있는 메커니즘에 링크를 제공한다.
- [0078] 도 7은 도 6에 도시된 가상 브리지(230)의 구체적인 구성을 도시한 도면이다.
- [0079] 도 7에 도시된 바와 같이, 가상 브리지(230)는 PCI 익스프레스 다운스트림 레인 물리 계층(이하 "다운스트림 레인 물리 계층"이라 함)(231), xN(+n) 케이블 전송 업스트림 레인 물리 계층(이하 "업스트림 레인 물리 계층"이라 함)(233), PCI 익스프레스 트랜잭션 계층(이하 "트랜잭션 계층"이라 함)(235), PCI 익스프레스 데이터 링크 계층(이하 "데이터 링크 계층"이라 함)(237) 및 케이블 전송 데이터 링크 계층(239)를 포함한다.
- [0080] 다운스트림 레인 물리 계층(231)은 물리적으로 패킷을 외부 디바이스로 전송하기 위해 PCI 익스프레스 xN 포트(210)에 접속된다.
- [0081] 업스트림 레인 물리 계층(233)은 물리적으로 패킷을 컴퓨터 시스템(100)으로 전송하기 위해 xN(+n) 케이블 정합 장치(220)에 접속된다.
- [0082] 트랜잭션 계층(235)은 다운스트림 레인 물리 계층(231)과 업스트림 레인 물리 계층(233)을 통해 송수신되는 패킷, 즉 트랜잭션 계층 패킷의 조립 및 분해를 수행한다. 이러한 트랜잭션 계층(235)에 대해서는 이미 잘 알려져 있으므로 구체적인 설명을 생략한다.
- [0083] 데이터 링크 계층(237)은 트랜잭션 계층(235)과 다운스트림 레인 물리 계층(231) 사이의 중간 단계로 작용하며, 두 계층(231, 235) 사이에서 패킷을 교환하기 위한 신뢰할 수 있는 메커니즘에 링크를 제공한다.
- [0084] 케이블 전송 데이터 링크 계층(239)은 트랜잭션 계층(235)과 업스트림 레인 물리 계층(233) 사이의 중간 단계로 작용하며, 두 계층(233, 235) 사이에서 패킷을 교환하기 위한 신뢰할 수 있는 메커니즘에 링크를 제공한다.

- [0085] 한편, 가상 브리지(130)의 다운스트림 레인 물리 계층(133)과 가상 브리지(230)의 업스트림 레인 물리 계층(233)은 xN(n) 케이블(300)의 N+n개의 케이블을 통해서 TS 데이터를 송수신하여 장애가 발생된 케이블을 확인하고, N+n개의 케이블 중에서 장애가 발생된 케이블을 제외한 케이블을 사용하여 데이터 송수신을 위한 최대 N개의 레인을 구성한다.
- [0086] 도 8은 본 발명의 실시예에 따른 PCI 익스프레스 스위치 장치에서 x8 레인을 적용한 예를 도시한 도면이다.
- [0087] 도 8에서는 본 발명의 실시예에 따른 PCI 익스프레스 스위치 장치에서 x8 레인을 적용하고, n=2여서 2개의 여분의 케이블 인터페이스를 추가한 형태를 나타내고 있다. 또한, 레인 1과 레인 7에 장애가 발생하였을 경우를 나타내고 있다.
- [0088] 도 8에서는 LTSSM이 각 레인별로 존재한다. 이것은 PCI 익스프레스 아키텍처상 물리 계층에서 송수신하는 패킷이 레인별로 정의되어 있으며, 링크 폭의 결정도 레인별로 이루어지기 때문이다. 따라서, 링크 폭이 결정되기 전까지 레인별로 LTSSM이 동작하도록 하여야 레인에서 장애가 발생하였을 때 이를 격리하고 나머지 레인으로만 링크 폭을 결정할 수 있도록 하기 위함이다.
- [0089] LTSSM의 상태 천이는 모든 레인의 상태를 근거로 동시에 이루어져야 하는 경우가 있으므로, 이를 위해 각 레인별 TS의 수신 상태와 상태 천이가 발생하였는지를 외부의 버스로 출력하고, 또한 그 버스 내용 모두를 각 레인에 입력하여 레인별로 동기화가 이루어지도록 하고 있다.
- [0090] 도 8에서, x8의 레인 폭을 가지는 PCI 익스프레스 링크를 예로 들어 다운스트림 레인에서는 두 번째 레인(Lane 1)의 수신에 장애가 발생하였고, 업스트림 레인에서는 마지막 레인(Lane 7)에서의 수신에 장애가 발생하였을 경우를 나타낸다.
- [0091] 이하, 도 8에 도시된 예를 참조하여 본 발명의 실시예에 따른 PCI 익스프레스 스위치 장치에서 링크 폭을 결정하기 위한 상태 천이 순서에 대해 설명한다.
- [0092] 도 9는 본 발명의 실시예에 따른 PCI 익스프레스 스위치 장치에서 링크 폭을 결정하기 위한 협상 과정을 도시한 도면이다.
- [0093] 도 9를 참조하면, Configuration.Linkwidth.Start의 상태(410)의 다운스트림 레인은 설정된 링크 값과 PAD의 레인 값을 가지는 TS1을 수신기를 감지한 모든 레인에 송신한다(S100).
- [0094] 업스트림 레인은 모두 PAD로 설정된 링크 및 레인 값을 가지는 TS1을 수신기를 감지한 모든 레인에 송신한다(S110).
- [0095] 다운스트림 레인은 PAD의 링크 값을 가지는 연속 2개의 TS1을 레인 1을 제외한 레인에서 수신한 후에 Configuration.Linkwidth.Accept 상태(420)로 천이한다.
- [0096] 한편, Configuration.Linkwidth.Start 상태(510)의 업스트림 레인은 설정된 링크 넘버와 PAD의 레인 넘버를 가지는 TS1을 레인 7을 제외한 모든 레인에서 2개 연속 수신한 후 Configuration.Linkwidth.Accept 상태(520)로 천이한다.
- [0097] 그 후, 업스트림 레인은 수신한 링크 값과 PAD의 레인 값을 가지는 TS1을 TS1을 수신한 레인에 송신하며, 레인 7에는 모두 PAD의 링크 및 레인 값을 가지는 TS1을 송신한다(S120).
- [0098] Configuration.Linkwidth.Accept 상태(420)의 다운스트림 레인은 자신이 송신한 링크 값과 PAD의 레인 값을 가지는 TS1을 레인 1을 제외한 모든 레인에서 수신한 후 Configuration.Lanenum.Wait(430)로 천이한다.
- [0099] 그 후 다운스트림 레인은 레인 0에는 0, 레인 2에는 1, 레인 3에는 2, 레인 4에는 3, 레인 5에는 4, 레인 6에는 5, 레인 7에는 6, 레인 8에는 7, 레인 9에는 8의 레인 값을 가지는 TS1 명령 집합(Ordered Set)을 송신하며, 레인 1에는 PAD의 레인 값을 가지는 TS1 명령 집합을 송신한다(S130).
- [0100] Configuration.Linkwidth.Accept 상태(520)의 업스트림 레인은 설정된 링크 값과 PAD가 아닌 레인 값을 가지는 2개 연속의 TS1을 아무 레인에서건 수신한 후, Configuration.Lanenum.Wait 상태(530)로 천이한다.
- [0101] 그 후, 업스트림 레인은 수신한 TS1과 동일한 링크 값을 가지며, 레인 0에는 0, 레인 1에는 1, 레인 2에는 2, 레인 3에는 3, 레인 4에는 4, 레인 5에는 5, 레인 6에는 6, 레인 8에는 7, 레인 9에는 8의 레인 값을 가지는 TS1을 송신하며, 레인 7에는 링크와 레인 값이 모두 PAD인 TS1을 송신한다(S140).

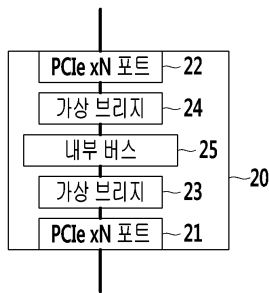
- [0102] Configuration.Lanenum.Wait 상태(430)의 다운스트림 레인은 처음 Configuration.Lanenum.Wait 상태(430) 진입시의 레인 값과 일치하지 않는 레인 값을 가지는 TS1을 레인 1, 2, 3, 4, 5, 6, 7에서 수신하였으므로, Configuration.Lanenum.Accept 상태(440)로 진입한다.
- [0103] 그리고, 다운스트림 레인은 자신이 송신 가능한 레인이 0, 2, 3, 4, 5, 6, 7, 8, 9이어서 공통적으로 사용 가능한 레인이 0, 2, 3, 4, 5, 6, 8, 9의 총 8개임을 확인한 후에 Configuration.Lanenum.Wait 상태(430)로 진입한다.
- [0104] 그 후, 다운스트림 레인은 설정한 링크 값을 가지며, 레인 0에는 0, 레인 2에는 1, 레인 3에는 2, 레인 4에는 3, 레인 5에는 4, 레인 6에는 5, 레인 8에는 6, 레인 9에는 7의 레인 설정값을 갖는 TS1을 송신한다(S150).
- [0105] 한편, Configuration.Lanenum.Wait 상태(530)의 업스트림 레인에서는 처음 Configuration.Lanenum.Wait 상태(530) 진입시의 레인 값과 일치하지 않는 레인 값을 가지는 TS1을 레인 1, 2, 3, 4, 5, 6, 8, 9에서 수신하였으므로, 자신이 송신 가능한 레인이 0, 1, 2, 3, 4, 5, 6, 8, 9이어서 공통적으로 사용 가능한 레인이 0, 2, 3, 4, 5, 6, 8, 9의 총 8개임을 확인한다.
- [0106] 그 후, 업스트림 레인은 Configuration.Lanenum.Accept 상태(540)로 진입하여서 수신한 링크 값을 가지며, 레인 0에는 0, 레인 2에는 1, 레인 3에는 2, 레인 4에는 3, 레인 5에는 4, 레인 6에는 5, 레인 8에는 6, 레인 9에는 7의 레인 설정값을 갖는 TS1을 송신한다.
- [0107] Configuration.Lanenum.Wait 상태(430)의 다운스트림 레인은 자신이 송신한 값과 일치하는 링크 및 레인 값을 가지는 TS1을 수신하였으므로, Configuration.Lanenum.Accept 상태(440)로 진입한 후에, Configuration.Complete 상태(450)로 진입하여 설정된 링크 값과 레인 값을 가지는 TS2를 송신한다.
- [0108] 한편, Configuration.Lanenum.Accept 상태(540)의 업스트림 레인은 TS2를 8개의 레인에서 2개 연속 수신한 후에 Configuration.Complete 상태(550)로 진입한다.
- [0109] 이상과 같이, 본 발명의 실시예에서는 레인이 추가되더라도 레인의 순서와 상관없이 가용한 레인의 개수를 모두 이용하여 링크 폭을 결정할 수 있으므로, 케이블과 같이 임의의 부분에 문제가 발생하더라도, 레인을 최대한 이용하여 대역폭을 보장할 수 있다.
- [0110] 또한, n개의 케이블을 추가 설치하더라도 LTSSM이 레인별로 분리되어 있으므로 설계상 추가가 용이하고, 추가된 케이블을 이용하여 N+n-k개(여기서 k는 장애가 발생한 케이블의 개수)의 레인을 이용하여 송수신을 할 수 있다.
- [0111] 따라서, n = N일 경우에는 최악의 경우 가용한 케이블이 절반일 지라도 N개의 레인 설정이 가능하므로, 대역폭에 전혀 지장을 받지 않도록 할 수 있다.
- [0112] 이상에서 본 발명의 실시예에 대하여 상세하게 설명하였지만 본 발명의 권리범위는 이에 한정되는 것은 아니고 다음의 청구범위에서 정의하고 있는 본 발명의 기본 개념을 이용한 당업자의 여러 변형 및 개량 형태 또한 본 발명의 권리범위에 속하는 것이다.

도면

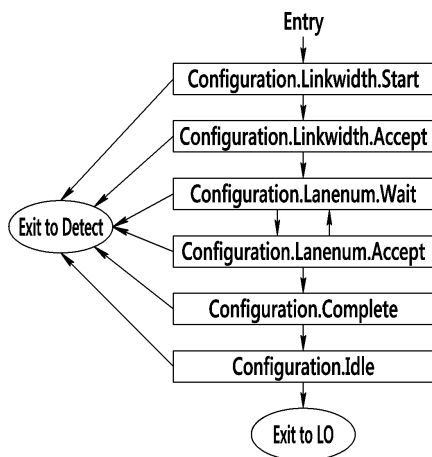
도면1



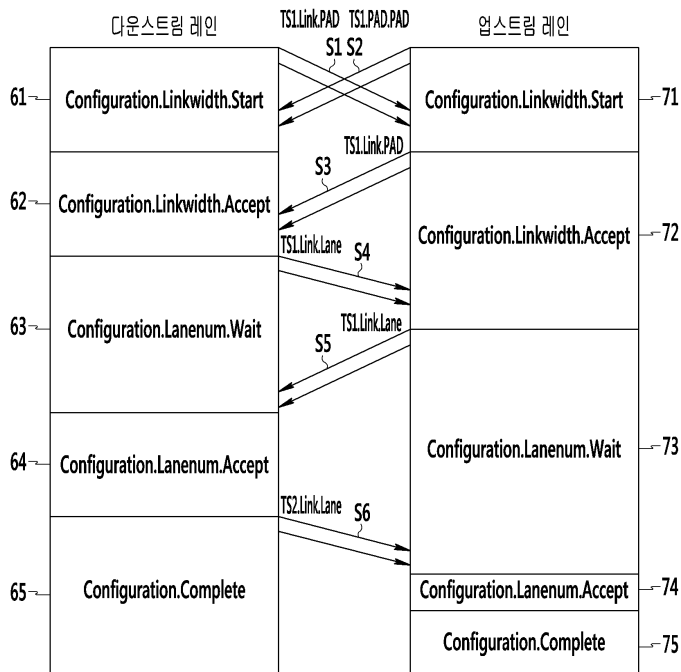
도면2



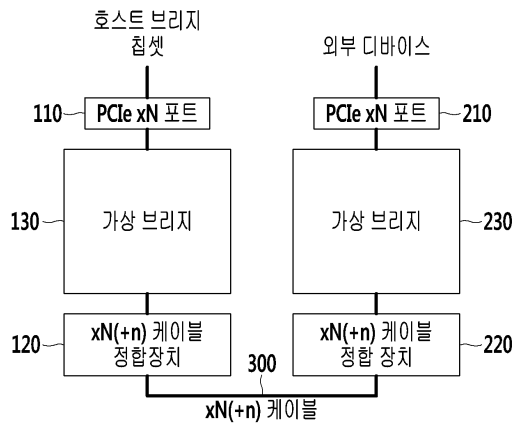
도면3



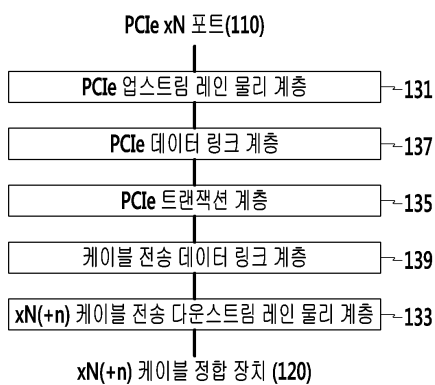
도면4



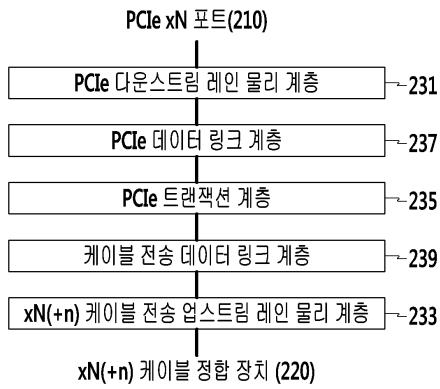
도면5



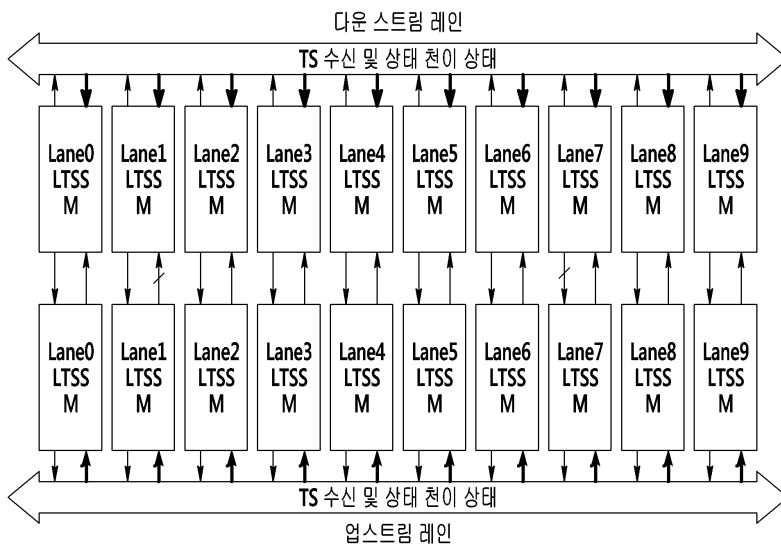
도면6



도면7



도면8



도면9

