관 인 생 략

출 원 번 호 통 지 서

출 원 일 자	2013.12.30
특 기 사 항	심사청구(무) 공개신청(무)
출 원 번 호	10-2013-0167898 (접수번호 1-1-2013-1208820-10)
출원인 명칭	주식회사 케이티(2-1998-005456-3) 외 1명
대리인 성명	특허법인 명문(9-2004-100021-1)
발명자 성명	정현호 김봉기 리샤드 알리예프 서동원 이시형 이희조 입누 무바로크 황영헌
발명의 명칭	네트워크 트래픽과 서버 부하에 따른 복제서버 자원의 동적 관리 방법, 그 장치 및 시스템

특 허 청 장

<< 안내 >>

1. 귀하의 출원은 위와 같이 정상적으로 접수되었으며, 이후의 심사 진행상황은 출원번호를 통해 확인하실 수 있습니다.

2. 출원에 따른 수수료는 접수일로부터 다음날까지 동봉된 납입영수증에 성명, 납부자번호 등을 기재하여 가까운 우체국 또는 은행에 납부하여야 합니다.

※ 납부자번호 : 0131(기관코드) + 접수번호

3. 귀하의 주소, 연락처 등의 변경사항이 있을 경우, 즉시 [출원인코드 정보변경(경정), 정정신고서]를 제출하여야 출원 이후의 각종 통지서를 정상적으로 받을 수 있습니다.

※ 특허로(patent.go.kr) 접속 > 민원서식다운로드 > 특허법 시행규칙 별지 제5호 서식

4. 특허(실용신안등록)출원은 명세서 또는 도면의 보정이 필요한 경우, 등록결정 이전 또는 의견서 제출기간 이내에 출원서에 최초로 첨부된 명세서 또는 도면에 기재된 사항의 범위 안에서 보정할 수 있습니다.

5. 외국으로 출원하고자 하는 경우 PCT 제도(특허·실용신안)나 마드리드 제도(상표)를 이용할 수 있습니다. 국내출원일을 외국에서 인정받고자 하는 경우에는 국내출원일로부터 일정한 기간 내에 외국에 출원하여야 우선권을 인정받을 수 있습니다.

※ 제도 안내 : http://www.kipo.go.kr-특허마당-PCT/마드리드

※ 우선권 인정기간 : 특허·실용신안은 12개월, 상표·디자인은 6개월 이내

※ 미국특허상표청의 선출원을 기초로 우리나라에 우선권주장출원 시, 선출원이 미공개상태이면, 우선일로부터 16개월 이내에 미국특허상표청에 [전자적교환허가서(PTO/SB/39)를 제출하거나 우리나라에 우선권 증명서류를 제출하여야 합니다.

6. 본 출원사실을 외부에 표시하고자 하는 경우에는 아래와 같이 하여야 하며, 이를 위반할 경우 관련법령에 따라 처벌을 받을 수 있습니다.

※ 특허출원 10-2010-0000000, 상표등록출원 40-2010-0000000

7. 기타 심사 절차에 관한 사항은 동봉된 안내서를 참조하시기 바랍니다.

【명세서】

【발명의 명칭】

네트워크 트래픽과 서버 부하에 따른 복제서버 자원의 동적 관리 방법, 그 장치 및 시스템 {Dynamic Management Methods of Replica Server Resorces, Apparatus Thereof, And Sytem Comprising The Same}

【기술분야】

<0001> 본 발명은 분산 서비스 거부 공격 등의 네트워크 트래픽 및 이로 인한 서버 부하를 관리하기 위한 방법에 관한 것으로서, 보다 상세하게는 네트워크 트래픽 및 서버 부하에 따라 클라우드 기반의 복제서버를 동적으로 관리하는 방법에 관한 것이다.

【발명의 배경이 되는 기술】

<0002> 분산 서비스 거부 공격은 여러 대의 공격자를 분산 배치하여 동시에 동작하게 함으로써 특정 사이트를 공격하는 해킹 방식의 하나이다. 이 방식은 서비스 공격을 위한 도구들을 여러 대의 컴퓨터에 심어놓고 공격 목표인 사이트의 컴퓨터 시스템이 처리할 수 없을 정도로 엄청난 분량의 패킷을 동시에 범람시킴으로써 네트워크의 성능을 저하시키거나 시스템을 마비시키게 된다.

<0003> 기존 DDoS 공격에 대한 방어 기법은 DDoS 공격의 일정한 규칙을 이용하여 트래픽을 차단하여 폐기하는 데 주력하였다. 그러나 최근의 DDoS 공격 방법은 정상 트래픽 패턴과 유사하므로 이와 같은 규칙을 적용하더라도 많은 양의 악성 트래픽이 여전히 공격 대상 서버에 도달하게 된다. 또한 이러한 규칙 기반의 대응 방법을

이용하는 경우 플래시 크라우드(Flash crowds)와 같은 정상적인 트래픽 집중 현상이 발생할 때도 악성 트래픽으로 오인하는 경우가 발생한다.

<0004> 한편, 한국등록특허 제900491호는 오리진 서버를 포함하는 복수의 서버로 구성하고, 오리진 서버의 트래픽 상태가 분산 서비스 거부 공격 상태로 판단되는 경우, 오리진 서버의 IP 주소를 복수의 서버로 IP 주소로 변경하여 오리진 서버의 부하를 경감하는 방식의 대응 방법을 제시하고 있다.

<0005> 그러나, 이 방식의 경우 공격의 강도에 따라 코어망에 유입되는 트래픽을 최소화 하기 위한 서버 자원의 동적 관리에 관한 고려가 없다.

【발명의 내용】

【해결하고자 하는 과제】

<0006> 상기 종래 기술의 문제점을 해결하기 위하여 본 발명은, 복제서버의 자원을 동적으로 관리하는 방법을 제공하는 것을 목적으로 한다.

<0007> 또한, 본 발명은 상기 복제서버 자원을 동적으로 관리하기 위한 복제서버 관리장치 및 시스템을 제공하는 것을 목적으로 한다.

【과제의 해결 수단】

<0008> 상기 기술적 과제를 달성하기 위하여 본 발명은, 메인서버의 성능 메트릭을 측정하는 단계; 상기 메인서버의 성능 메트릭과 사전 설정된 제1 기준을 대비하여 상기 메인서버로의 경로 상에 복제서버의 개시 여부를 결정하는 단계; 및 상기 메인서버의 성능 메트릭과 상기 사전 설정된 제1 기준보다 낮은 제2 기준을 대비하여 상기 메인서버로의 경로 상의 복제서버의 중지 여부를 결정하는 단계를 포함하는

복제서버의 자원 관리 방법을 제공한다.

<0009> 본 발명에서 상기 성능 메트릭은 메인서버의 응답시간 정보 또는 CPU 부하를 포함할 수 있다.

<0010> 이 때, 상기 성능 메트릭은 소정 기간 동안 측정된 값이고, 상기 사전 설정된 제1 기준은 제1 임계값 및 상기 제1 임계값 보다 낮은 제2 임계값을 포함하고, 상기 소정 기간 측정된 성능 메트릭과 상기 제1 임계값 또는 상기 제2 임계값과의 대비에 따라 복제서버의 개시 여부가 결정될 수 있다.

<0011> 또, 상기 성능 메트릭의 상기 사전 설정된 제2 기준은 제3 임계값 및 상기 제3 임계값 보다 낮은 제4 임계값을 포함하고, 상기 소정 기간 측정된 성능 메트릭과 상기 제3 임계값 또는 제4 임계값을 대비하여 복제서버의 중지 여부가 결정될 수 있다.

<0012> 본 발명은 상기 메인서버의 성능 메트릭의 변화에 따라 상기 복제서버를 스케일링하는 단계를 더 포함할 수 있다.

<0013> 또한, 본 발명에서 상기 측정 단계는, 네트워크에 산재된 복수의 센서로부터의 성능 메트릭을 수신하는 단계를 포함할 수 있다. 이 때, 상기 성능 메트릭은 상기 메인서버로의 응답시간 정보를 포함하는 것이 바람직하다. 또한, 상기 성능 메트릭은 상기 센서로부터 상기 메인서버로의 네트워크 상의 경로 정보를 포함할 수 있다. 또한, 상기 성능 메트릭이 상기 제1 기준을 초과하는 경우 상기 성능 메트릭에 포함된 경로 정보로부터 추출된 링크들 중 상기 메인서버로부터 가장 먼 링크에 연결된 복제서버를 우선적으로 개시하는 것이 바람직하다.

<0014> 상기 다른 기술적 과제를 달성하기 위하여 본 발명은 네트워크에 산재된 복수의 센서로부터 성능 메트릭 정보를 수신하는 상태 모니터링부; 상기 성능 메트릭 정보와 사전 설정된 기준을 대비하여 DDoS 공격 여부를 판단하는 공격 판단부; DDoS 공격이 있는 것으로 판단되는 경우 상기 네트워크상에 복제서버의 위치를 결정하는 복제서버 위치결정부; 및 상기 복제서버를 개시 및 중지하며 상기 복제서버의 자원을 관리하는 복제서버 자원관리부를 포함하는 복제서버 관리장치를 제공한다.

<0015> 본 발명에서 상기 성능 메트릭 정보는 상기 복수의 센서와 상기 메인서버 간의 링크 정보를 포함하고, 상기 복제서버 위치결정부는 상기 성능 메트릭 정보가 사전 설정된 기준을 초과하는 경우 해당 센서와 메인서버간의 링크 정보로부터 정체 상태인 링크 중 네트워크 경로 상의 거리가 가장 먼 링크를 식별한다.

<0016> 본 발명에서 상기 복제서버 자원 관리부는 DDoS 공격 강도에 따라 상기 복제서버의 자원을 증감시키는 것이 바람직하다.

<0017> 상기 또 다른 기술적 과제를 달성하기 위하여 본 발명은 네트워크에 산재된 복수의 복제서버 팜에 인접하여 메인서버의 성능 메트릭을 측정하는 측정 센서; 및 상기 측정 센서로부터의 성능 메트릭 데이터로부터 상기 네트워크의 트래픽을 측정하고, 측정된 네트워크 트래픽에 따라 복제서버의 자원을 관리하는 복제서버 관리장치를 포함하는 복제서버 관리 시스템을 제공한다.

【발명의 효과】

<0018> 본 발명에 따르면, DDoS 등의 공격자의 공격 강도에 따라 복제서버를 개시 및 중지하고 확장 및 축소하여 클라우드 기반 복제서버를 효율적으로 운용할 수 있게 된다.

<0019> 또한, 본 발명에 따르면, 메인서버의 인스톨이나 변경 없이도 복제서버의 자원을 효율적으로 관리할 수 있게 된다.

<0020> 또한, 본 발명에 따르면 메인서버로 이르는 네트워크 경로 중 메인서버에서 가장 먼 정체 링크에 속하는 복제서버를 개시함으로써, DDoS 공격 등을 효율적으로 차단하는 한편, 선의의 사용자에게는 메인서버에서 제공하는 서비스를 지장없이 제공할 수 있게 된다.

【도면의 간단한 설명】

<0021> 도 1은 본 발명의 바람직한 실시예에 따른 공격 트래픽 분산 방법을 구현하기 위한 전체 시스템 구성을 모식적으로 도시한 도면이다.

도 2는 본 발명의 구현예로서 복제서버와 관련한 자원관리 절차를 개략적으로 도시한 도면이다.

도 3은 본 발명의 바람직한 실시예에 따른 복제서버 관리장치의 구성을 개략적으로 도시한 도면이다.

도 4는 본 발명의 바람직한 실시예에 따른 복제서버의 동작의 일례를 설명하기 위한 시스템 구성을 모식적으로 도시한 도면이다.

도 5의 (a) 및 (b)는 본 발명의 바람직한 실시예에 따라 복제서버의 위치를 결정하는 방법을 모식적으로 도시한 도면이다.

도 6은 시간의 경과에 따른 성능 메트릭의 변화를 예시적으로 나타내는 그래프이다.

도 7의 (a) 및 (b)는 성능 메트릭에 기초하여 복제서버를 동적으로 개시 및 중지하기 위한 실시예를 예시적으로 설명하기 위한 도면이다.

도 8의 (a) 및 (b)는 성능 메트릭에 기초하여 복제서버를 동적으로 개시 및 중지하기 위한 다른 실시예를 예시적으로 설명하기 위한 도면이다.

도 9의 (a) 및 (b)는 성능 메트릭에 기초하여 복제서버를 동적으로 개시 및 중지하기 위한 또 다른 실시예를 예시적으로 설명하기 위한 도면이다.

도 10의 (a) 및 (b)는 성능 메트릭에 기초하여 복제서버를 동적으로 개시 및 중지하기 위한 또 다른 변형례를 예시적으로 설명하기 위한 도면이다.

【발명을 실시하기 위한 구체적인 내용】

<0022> 이하, 첨부한 도면을 참조하여, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 본 발명을 상술한다. 그러나 본 발명은 아래에서 예시한 것과 다른 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다. 또, 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

<0023> 본 발명을 설명함에 있어서, 관련된 공지 구성 또는 기능에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우에는 그 상세한 설명은 생략한다.

<0024> 도 1은 본 발명의 바람직한 실시예에 따른 공격 트래픽 분산 방법을 구현하기 위한 전체 시스템 구성을 모식적으로 도시한 도면이다.

<0025> 도시된 바와 같이, 메인 서버(S_M)와 다수의 복제서버(R_{11}, R_{12}, R_{21}, R_{23})가 인터넷 서비스 공급자(Internet Service Provider)에 의해 제공되는 네트워크(Network; 10)에 분산되어 있다.

<0026> 본 발명에서 메인 서버(S_M)는 예컨대 'www.naver.com'과 같은 도메인 네임을 가지고 웹 페이지 상의 서비스를 구현하는 웹 서버일 수 있다.

<0027> 상기 복제서버(R_{11}, R_{12}, R_{21}, R_{23})는 상기 메인 서버가 보유한 원본 콘텐츠의 전부 또는 일부의 사본을 구비하여 메인 서버(R_{11}, R_{12}, R_{21}, R_{23})의 기능을 제공할 수 있는 서버를 말한다.

<0028> 복제서버(R_{11}, R_{12}, R_{21}, R_{23})는 다양한 유형으로 구현될 수 있다. 먼저, 복제서버는 메인 서버의 전체 콘텐츠를 복제서버에 복사하는 형태로 구성될 수 있다. 복제 소요 시간이 오래 걸리고, 저장 장치의 자원이 많이 요구되지만, 사용자에게 가장 안정적인 서비스를 제공할 수 있다.

<0029> 이와 달리, 사용자의 요청이 빈번한 특정 콘텐츠를 복제서버에 복사하는 형태로 복제서버가 구성될 수 있다. 이와 같은 관심기반 복제서버는 사용자 요청이 빈번한 콘텐츠 인지 여부를 콘텐츠에 대한 사용자의 요청 횟수에 기반하여 판단할 수 있다. 관심 기반 복제서버는 상대적으로 적은 자원이 요구되지만, 서비스 제공

자는 어떤 콘텐츠에 사용자들이 관심을 두는지 모니터링 해야 하고, 이에 따라 복제서버의 콘텐츠를 갱신해야 한다.

<0030> 또한, 멀티미티어 파일, 문서 파일, 사용자 파일 등으로 콘텐츠의 타입을 나누어 콘텐츠를 저장하는 콘텐츠 타입 기반 복제서버로 구성될 수도 있다. 즉, 하나의 복제서버는 하나 이상의 콘텐츠 타입을 담당하게 된다. 이 때, 콘텐츠 타입이란, 콘텐츠의 파일 형식이 될 수 있고, 기설정된 콘텐츠의 분류가 될 수도 있다.

<0031> 본 발명에서 복제서버(R_{11}, R_{12}, R_{21}, R_{23})는 클라우드 컴퓨팅 기술로 구현될 수 있다. 클라우드 컴퓨팅은 서로 다른 물리적인 위치에 존재하는 컴퓨터들의 자원을 가상화 기술로 통합해 제공하는 기술로, 복제서버의 자원을 효율적으로 사용할 수 있게 하며, 가상 공간에 있는 서버의 자원을 이용하여 복제서버를 구축할 수 있도록 한다. 물리적으로, 상기 복제서버는 메인 서버 관리자가 관리하는 데이터 센터 또는 상기 관리자와 약정 관계에 있는 다른 사업자의 데이터 센터 내에 위치할 수 있다.

<0032> 본 발명에서 상기 복제서버가 반드시 클라우드 컴퓨팅 기술로 구현되어야 하는 것은 아니고 별도의 내부 또는 외부의 리소스로 구성되는 것도 가능하다.

<0033> 본 발명에 따르면, 네트워크(10)에 연결된 사용자(U)와 공격자(A)를 포함하는 네트워크 트래픽은 적절히 분리된다.

<0034> 정상 동작 모드에서 상기 사용자(U)에 의해 전송되는 정상 패킷(normal packet)은 네트워크(10) 상의 경로를 거쳐 메인 서버(S_M)로 전송된다.

<0035> 악의의 사용자(A)의 DDoS 공격에 의한 패킷(Attack Packets) 범람시 메인 서버로 향하는 네트워크 경로 상의 적절한 위치에 복제서버가 활성화된다. 상기 복제서버는 활성화 될 위치에 관계된 데이터 센터의 관리자로 전송되는 서버 용량, 개수 및 활성화 시점 등을 포함하는 개시 메시지(invoke message)에 의해 활성화될 수 있다. 상기 복제서버는 상기 메인 서버와 동기화된다. 상기 복제서버와 메인 서버의 동기화를 위하여 별도의 네트워크 예컨대 콘텐츠 전달 네트워크(CDN)가 구성될 수도 있을 것이다.

<0036> DDoS 공격시 범람 패킷(attack flood)은 상기 메인 서버로의 네트워크 경로 상에 위치한 활성화 된 복제서버(R_{11}, R_{12})로 리디렉션 된다.

<0037> 본 발명에서 상기 복제서버의 활성화 위치(activate location) 및 요구되는 리소스를 결정하기 위하여 상기 네트워크(10)에는 복제서버 관리 장치(100)가 구비된다.

<0038> 상기 복제서버 관리장치(100)는 네트워크(10)의 트래픽 및 상기 메인 서버(S_M)의 상태를 모니터링하고, 이에 기초하여 복제서버와 관련한 자원(Resoruce)을 관리한다. 상기 복제서버 관리장치(100)는 복제서버의 개시(invoction) 여부를 결정하기 위하여 별도의 센서를 운용할 수 있다. 또한, 상기 복제서버 관리장치(100)는 별도의 IDS(Intrusion Detection System)로부터 DDoS 공격 여부에 대한 정보를 수신하여 복제서버의 개시 여부 결정에 참조할 수 있다. 본 발명에서 복제서버의 동작에 관해서는 후술한다.

<0039> 도 2는 본 발명의 구현예로서 복제서버와 관련한 자원관리 절차를 개략적으로 도시한 도면이다.

<0040> 도 2를 참조하면, DDoS 등의 공격에 의한 패킷 범람이 검출되고 공격 상황으로 판단되면, 해당 공격을 방어하기 위한 적절한 복제서버 위치가 활성화 된다. 네트워크 상태, 메인 서버 상태 및 복제서버의 상태가 지속적으로 모니터링되고 해당 위치에서 복제서버가 개시(invoke)되거나 중지(revoke)되며 복제서버는 동적으로 확장 또는 축소(scaling) 된다. 이어서, 해당 위치와 관련한 DDoS 공격 상태가 해소되면, 복제서버 위치는 비활성화 된다.

<0041> 도 3은 본 발명의 바람직한 실시예에 따른 복제서버 관리장치(100)의 구성을 개략적으로 도시한 도면이다.

<0042> 도 3을 참조하면, 복제서버 관리장치(100)는 상태 모니터링부(120), DDoS 공격판단부(140), 복제서버 위치결정부(160) 및 복제서버 자원관리부(180)를 포함한다.

<0043> 상기 상태 모니터링부(140)는 다양한 상태정보를 모니터링한다. 예컨대, CPU 로드 등 메인서버 및 복제서버의 상태가 모니터링 될 수 있다. 또한, 네트워크 내외에 위치한 센서로부터의 트래픽 상태가 모니터링 될 수 있다. 또한, IDS와 같은 외부 장치에서 취득된 DDoS 공격 정보가 모니터링 될 수 있다. 이 때, OpenNMS, Nagios, PA Server Monitor 등과 같은 서버 모니터링 툴이 사용될 수 있다.

<0044> DDoS 공격 판단부(140)는 모니터링된 정보를 기초로 DDoS 공격 여부와 복제서버의 개시 여부를 판단한다. 이 때, 응답 시간(Response Time; RT), CPU 로

드(load), 네트워크 I/O(input/output), 메모리 사용량이나, 단위 시간당 패킷, 단위시간당 접속, 단위 시간당 리퀘스트(request)와 같은 통계 데이터를 지표(indicator)로 사용될 수 있다. 전술한 지표 중 최소한 하나 이상을 반영한 성능 메트릭(performance metric)으로부터 DDoS 공격 여부와 복제서버의 동작 여부를 판단한다. 예컨대, 상기 성능 메트릭은 예컨대 응답시간과 같은 수치이고, 상기 수치가 사전 설정된 임계값을 초과하는 경우 DDoS 공격이 발생하였음을 판단할 수 있다.

<0045> 물론 당업자라면 상기 성능 메트릭이 하나 이상의 지표가 통합되어 운용될 수 있음을 알 수 있을 것이다. 또한, 상기 성능 메트릭의 하나 이상의 지표에는 가중치가 부여될 수도 있을 것이다.

<0046> 상기 DDoS 공격 판단부(140)에 의해 복제서버의 개시(invoke)가 결정되는 경우, 상기 복제서버 위치결정부(160)는 활성화 될 복제서버의 위치를 결정한다. 복제서버의 위치 결정 방법에 대해서는 후술한다.

<0047> 상기 복제서버 위치결정부(160)가 활성화 될 복제서버의 위치를 결정하면, 복제서버 자원관리부(180)는 해당 위치에서의 복제서버의 개수 및 용량을 포함하여 요구되는 복제서버 자원을 결정한다. 이 때, 상기 복제서버 자원관리부(180)는 공격의 강도를 고려하여 복제서버 자원을 결정할 수 있다.

<0048> 상기 복제서버 자원관리부(180)는 복제서버의 개시(invoke)를 의미하는 메시지를 상기 복제서버 또는 상기 복제서버의 관리자에게 전송하여 복제서버를 활성화할 수 있다. 상기 개시 메시지(invoke msg.)에는 복제서버 자원의 크기를 특정하는

메시지가 포함될 수 있다.

<0049> 또한, 복제서버 자원관리부(180)는 상기 복제서버 자원을 동적으로 관리한다. 복제서버의 개시 이후 공격의 강도에 맞추어 적절한 복제서버 자원이 추가되거나 삭제될 수 있다. 상기 복제서버 자원관리부(180)는 개시 메시지(invoke msg.) 및 중지 메시지(revoke msg.)를 통해 복제서버 자원을 동적으로 관리한다.

<0050> 이하에서는 도 4 및 도 5를 참조하여 본 발명에 따른 복제서버의 개시 및 중지 방법을 설명한다.

<0051> 도 4는 본 발명의 바람직한 실시예에 따른 복제서버의 동작의 일례를 설명하기 위한 시스템 구성을 모식적으로 도시한 도면이다.

<0052> 도시된 같이, 네트워크 상에는 복수의 복제서버 팜(replica farm; RF)이 산재되어 있다. 상기 복제서버 팜(RF_1, RF_2, RF_3, RF_4)은 각각 최소한 하나 이상의 복제서버들(R1, R2, R3)이 구비되어 있다. 상기 복제서버 팜(RF_1, RF_2, RF_3, RF_4)은 예컨대 ISP 사업자가 운영하는 데이터 센터 또는 데이터 저장소일 수 있다.

<0053> 네트워크 상의 적절한 위치에 센서 에이전트(S)가 구비된다. 도시된 바와 같이 상기 센서 에이전트들(S_{I1}, S_{I2}, S_{I3}, S_{I4})은 상기 복제서버 팜(RF_1, RF_2, RF_3, RF_4) 내부에 구비될 수 있으며, 일부 센서(S_{01})는 사용자와 동일한 환경에 있도록 네트워크 외부에 구비될 수 있다. 상기 센서 에이전트들(S_{I1}, S_{I2}, S_{I3}, S_{I4}, S_{01})은 메인서버(S_M)로의 성능 메트릭을 측정한다. 상기 센서에 의해 측정되는 성능 메트릭은 예

컨대 임의의 요청에 대한 메인서버의 응답시간(RT)이 될 수 있다.

<0054> 응답시간은 상기 센서 에이전트들(S_{I1}, S_{I2}, S_{I3}, S_{I4}, S_{01})이 메인서버로 TCP SYN 신호를 전송하거나 다른 간단한 HTTP 요청을 전송하고, 상기 메인서버로부터 응답을 수신함으로써 측정될 수 있다. 상기 센서들은 메인서버의 응답시간에 대응하는 정보를 포함하는 센싱 데이터를 복제서버 관리장치(100)로 전송한다. 상기 센싱 데이터는 요청과 응답이 송수신된 네트워크 경로의 링크 정보를 포함한다. 성능 메트릭으로 응답시간을 이용하는 방식은 단순한 정보 요청에 의해 구현될 수 있으며, 메인서버에 대한 인스톨이나 변경 문제를 유발하지 않는 장점을 갖는다.

<0055> 복제서버 관리장치(100)는 상기 센서들에 의해 전송된 정보에서 응답시간 정보 및 정체 링크 정보를 추출한다. 예컨대, 정상 동작시를 기준으로 한 평균 응답시간을 현저히 초과하는 응답시간이 발생한 해당 센서와 메인서버의 경로 상에 정체가 발생하였음을 의미한다. 상기 복제서버 관리장치(100)는 응답시간이 사전 설정된 값을 초과한 경우 센서로부터 전송된 경로 정보로부터 정체가 발생한 링크를 추출하고, 해당 링크에 관련된 복제서버를 개시할 수 있다.

<0056> 또한, 상기 센서는 상기 복제서버 관리장치(100)로 응답시간 정보를 주기적으로 전송한다. 복제서버가 개시된 상태에서 전송된 응답시간 정보로부터 추출된 응답시간이 사전 설정된 값 미만으로 하락한 경우 상기 복제서버 관리장치(100)는 해당 링크의 정체 상태가 해소되었음을 판단하고 해당 링크에 관련된 복제서버를 중지할 수 있다. 복제서버의 중지를 위한 사전 설정된 값은 메인서버가 핸들링 가

능한 트래픽을 기준으로 설정될 수 있다.

<0057> 도 5의 (a) 및 (b)는 본 발명의 바람직한 실시예에 따라 복제서버의 위치를 결정하는 방법을 모식적으로 도시한 도면이다.

<0058> 도 5의 (a)에 도시된 바와 같이, 네트워크에 산재된 각 복제서버 팜(RF)의 센서들(S)은 주기적으로 TCP SYN 신호를 메인서버로 전송하고 그 응답으로서 ACK 신호를 수신한다. 상기 센서들(S)이 수신한 데이터는 응답시간 정보로서 메인서브를 경유하여 또는 직접 복제서버 관리장치(100)로 전송된다.

<0059> 복제서버 관리장치(100)는 전송된 정보로부터 정체된 링크를 판별한다. 예컨대, 악의의 공격자(A)에 의한 공격 트래픽(실선)이 발생하면, 그 경로 상의 복제서버 팜(RF_1, RF_2)의 센서들(S_{I1}, S_{I2})로부터 전송된 데이터에는 링크 정체 상태가 포함될 수 있다. 따라서, 상기 복제서버 관리장치(100)는 상기 정체 링크에 연결된 복제서버 팜(RF_1, RF_2)의 복제서버를 활성화 할 수 있을 것이다.

<0060> 도 5의 (b)는 상기 복제서버 관리장치(100)는 정체된 링크에 연관된 복제서버 중 메인서버로부터 가장 먼 링크에 연관된 복제서버 팜(RF_1)의 복제서버를 우선적으로 활성화 하는 것을 보여주고 있다. 이와 같이 정체 링크들 중 메인서버로의 경로상의 가장 먼 링크에 연관된 복제서버를 활성화 함으로써 공격자들이 메인서버로 접속하는 것을 차단하고 해당 링크를 이용하지 않는 다수의 사용자들은 메인서버의 서비스에 장애 없이 접속할 수 있게 된다. 메인서버로의 경로에서 가장 먼 링크는 센서 데이터에 포함된 경로 정보 중 홉 거리(hop distance) 또는 TTL(Time to

Live) 정보를 이용하여 추출될 수 있다.

<0061> 상술한 실시예는 하나의 복제서버 팜을 활성화 하는 것을 기술한 것이지만, 공격이 여러 지역의 공격자로부터 동시 다발적으로 이루어지는 상황에서는 서로 다른 링크에 연관된 복수의 복제서버 팜이 활성화될 수 있음은 당업자라면 누구나 알 수 있을 것이다.

<0062> 또한, 본 발명에서 복제서버는 공격의 강도에 따라 동적으로 증감(scaling)된다.

<0063> 복제서버의 개시에도 불구하고 응답시간의 감소가 유발되지 않고 응답시간이 여전히 사전 설정된 임계값 이상을 유지하는 경우 복제서버의 크기는 증가하여야 하고, 반대로 응답시간이 감소하여 소정 기준 이하로 되는 경우 복제서버의 크기는 감축되어야 한다. 또한, 응답시간 값이 소정 기준값 상하로 변동하는 경우 복제서버의 크기는 증가 또는 감소되어야 한다.

<0064> 본 발명에서 복제서버의 스케일링은 사전 설정된 용량 단위로 수행될 수 있다. 예컨대 복제서버는 예컨대 메인서버의 용량과 동일한 크기 단위로 개시 및 중지될 수 있고 또 이를 증감 단위로 하여 스케일링 될 수 있다. 또한, 본 발명에서 복제서버의 스케일링은 복제서버의 개시 및 중지 명령과 동일한 명령에 의하여 수행될 수 있다.

<0065> 이하 본 발명의 구현예로써 성능 메트릭에 기초하여 복제서버의 개시, 중지 및 스케일링을 실행하는 방법을 예시적으로 설명한다.

<0066> 본 발명은 복제서버의 개시나 중지 결정을 위한 성능 메트릭의 판단 기준으

로서 각각 하나의 임계값과 임계값들 사이에 존재하는 완충값을 사용한다. 완충값은 DDoS 공격에 의한 트래픽의 요동(fluctuation)시 복제서버의 점진적인 스케일링을 가능하게 하며, 메인서버 및 복제서버에 의한 서비스의 지속성을 보장한다.

<0067> 도 6은 시간의 경과에 따른 성능 메트릭의 변화를 예시적으로 나타내고 있다.

<0068> 전술한 바와 같이, 상기 성능 메트릭(Performance Metric; 이하 'PM'이라 한다)으로는 응답시간(RT), CPU 부하 등의 성능 지표가 사용될 수 있다. 바람직하게는 상기 성능 메트릭은 네트워크 내부나 외부의 센서(S)로부터 전송되는 응답시간 정보이다.

<0069> 도시된 바와 같이, PM에 대응하여 4개의 사전 설정된 값 즉 개시 임계값(threshold invoke; Θ_{inv}), 개시 완충값(bumper invoke; β_{inv}), 중지 완충값(bumper revoke; β_{rev}) 및 중지 임계값(Θ_{rev})이 표시되어 있다. 개시 완충값(β_{inv}) 이상의 영역은 개시 영역(invoke area)이라 하고, 중지 완충값(β_{rev}) 미만의 영역은 중지 영역(revoke area), 그 사이의 영역은 안정 영역(stable area)이라 부른다.

<0070> 도시된 PM은 중지 영역을 거쳐 점진적으로 증가하여 개시 임계값(Θ_{inv})을 지나 상승하고 있다. 이와 같이 평균 PM이 개시 임계값(Θ_{inv}) 이상이고, 현재의 PM 또한 개시 임계값(Θ_{inv}) 이상인 경우 초기 DDoS 공격으로 판단하고 복제서버가 개

시된다.

<0071> 복제서버가 개시된 후 PM의 요동(fluctuation)이 존재하는 경우에는 주어진 모니터링 시간($T_{monitor}$) 동안의 PM의 평균값이 개시 완충값(bumper invoke; β_{inv})을 초과할 때에 복제서버가 개시되는 것이 바람직하다. 또한, 주어진 시간 동안 성능 메트릭의 평균값이 중지 완충값(bumper revoke; β_{rev}) 미만일 때에는 복제서버가 중지되는 것이 바람직하다. 이하에서는 평균 PM 및 최종 PM 값을 기준으로 복제서버의 개시 및 중지를 판단하는 다양한 예를 설명한다.

<0072> 도 7의 (a) 및 (b)는 PM에 기초하여 복제서버를 동적으로 개시 및 중지하기 위한 제1 실시예를 예시적으로 설명하기 위한 도면이다.

<0073> 우선 도 7의 (a)를 참조하면, 소정의 모니터링 개시점(M_{u_start})과 모니터링 중지점(M_{u_stop}) 사이의 주어진 모니터링 기간($T_{monitor}$) 동안 PM이 모니터링 되고 그 기간의 평균 PM값이 개시 완충값(β_{inv})을 초과하고, 모니터링 기간 중 마지막으로 얻어진 PM값 즉 M_{u_stop}에서의 PM값이 개시 완충값(β_{inv}) 이상일 때 복제서버가 개시된다.

<0074> 도 7의 (b)를 참조하면, 모니터링 기간($T_{monitor}$) 동안의 평균 PM이 중지 완충값(β_{inv}) 미만이고, 최종 PM값이 중지 완충값(β_{inv}) 미만일 때 복제서버는 중지된다.

<0075> 도 8의 (a) 및 (b)는 PM에 기초하여 복제서버를 동적으로 개시 및 중지하기

위한 제2 실시예를 예시적으로 설명하기 위한 도면이다.

<0076> 도 8의 (a)와 같이, 모니터링 기간에서의 최종 PM이 안정 영역(stable area)에 존재하더라도, 평균 PM이 개시 임계값(Θ_{inv})을 초과하는 경우에는 복제서버가 개시된다. 또한, (b)와 같이 모니터링 기간에서의 최종 PM이 안정 영역에 존재하더라도, 평균 PM이 중지 임계값(Θ_{rev}) 미만일 때에는 복제서버는 중지된다.

<0077> 도 9의 (a) 및 (b)는 성능 메트릭에 기초하여 복제서버를 동적으로 개시 및 중지하기 위한 제3 실시예를 예시적으로 설명하기 위한 도면이다.

<0078> 도 9의 (a)와 같이, 최종 PM이 개시 임계값(Θ_{inv})을 초과하고, 평균 PM이 중지 완충값(β_{inv}) 미만일 경우에는 안정 상태로 판단하여 복제서버는 개시되지 않는다.

<0079> 이와 유사하게, 도 9의 (b)와 같이, 최종 PM이 개시 임계값(Θ_{rev})을 초과하고, 평균 PM이 중지 완충값(β_{rev}) 이상일 경우에는 안정 상태로 판단하여 복제서버는 중지되지 않는다.

<0080> 도 10의 (a) 및 (b)는 성능 메트릭에 기초하여 복제서버를 동적으로 개시 및 중지하기 위한 제4 실시예를 예시적으로 설명하기 위한 도면이다.

<0081> 도 10의 (a)와 같이, 측정된 최종 PM 값이 개시 임계값(threshold invoke; Θ_{inv}) 이상이고, 그 기간 동안의 평균 PM이 개시 임계값(threshold invoke; Θ_{inv}) 이상인 경우 복제서버가 개시된다. 이 때, 복제서버의 용량은 단위 복제서버의 크

기의 2배가 된다.

<0082> 이와 유사하게, 도 10의 (b)와 같이, 측정된 최종 PM 값이 개시 임계값(threshold invoke; Θ_{rev}) 미만이고, 그 기간 동안의 평균 PM이 개시 임계값(threshold invoke; Θ_{rev}) 미만인 경우 복제서버가 중지된다. 이 때, 중지되는 복제서버의 용량은 단위 복제서버의 크기의 2배가 된다.

<0083> 상술한 복제서버의 개시 및 중지에 의하여 메인서버로의 트래픽은 복제서버로 분산된다. 트래픽의 분산은 DNS 라운드 로빈(DNS Round Robin)이나 스위치 기반 부하 분산 방법 등의 통상의 기법에 의해 수행될 수 있다. DNS 라운드 로빈은 예를 들어 www.example.com에 대한 서비스를 1.1.1.1이라는 IP를 소유한 서버가 담당하고 있었다면, 과도한 트래픽이 집중될 시에는 1.1.1.2, 1.1.1.3 등의 복제서버의 IP를 해당 도메인의 담당 서버로 등록하여 사용자의 트래픽이 복제서버로 분산될 수 있도록 하는 기법을 말한다. 한편, 스위치 기반 부하 분산 방법은 특정 IP 영역을 근원지 IP로 가지는 패킷 혹은 일정 확률로 선택된 패킷을 지정된 대상으로 전달하는 네트워크 스위치 기능을 이용하여 복제서버로 트래픽을 분산하는 기법을 말한다.

【부호의 설명】

<0084>

A	공격자
U	사용자

R	복제서버
RF	복제서버 팜
S	센서
10	네트워크
100	복제서버 자원관리장치
120	상태 모니터링부
140	DDoS 공격 판단부
160	복제서버 위치결정부
180	복제서버 자원관리부

【특허청구범위】

【청구항 1】

메인서버의 성능 메트릭을 측정하는 단계;

상기 메인서버의 성능 메트릭과 사전 설정된 제1 기준을 대비하여 상기 메인서버로의 경로 상에 복제서버의 개시 여부를 결정하는 단계; 및

상기 메인서버의 성능 메트릭과 상기 사전 설정된 제1 기준보다 낮은 제2 기준을 대비하여 상기 메인서버로의 경로 상의 복제서버의 중지 여부를 결정하는 단계를 포함하는 복제서버의 자원 관리 방법.

【청구항 2】

제1항에 있어서,

상기 성능 메트릭은 메인서버의 응답시간 정보 또는 CPU 부하를 포함하는 것을 특징으로 하는 복제서버의 자원 관리 방법.

【청구항 3】

제1항에 있어서,

상기 성능 메트릭은 소정 기간 동안 측정된 값이고,

상기 사전 설정된 제1 기준은 제1 임계값 및 상기 제1 임계값 보다 낮은 제2 임계값을 포함하고,

상기 소정 기간 측정된 성능 메트릭과 상기 제1 임계값 또는 상기 제2 임계값과의 대비에 따라 복제서버의 개시 여부를 결정하는 것을 특징으로 하는 복제서버의 자원 관리 방법.

【청구항 4】

제1항에 있어서,

상기 성능 메트릭은 소정 기간 동안 측정된 값이고,

상기 사전 설정된 제2 기준은 제3 임계값 및 상기 제3 임계값 보다 낮은 제4 임계값을 포함하고,

상기 소정 기간 측정된 성능 메트릭과 상기 제3 임계값 또는 제4 임계값을 대비하여 복제서버의 중지 여부를 판단하는 것을 특징으로 하는 복제서버의 자원 관리 방법.

【청구항 5】

제1항에 있어서,

상기 메인서버의 성능 메트릭의 변화에 따라 상기 복제서버를 스케일링 하는 단계를 더 포함하는 것을 특징으로 하는 복제서버의 자원 관리 방법.

【청구항 6】

제1항에 있어서,

상기 측정 단계는,

네트워크에 산재된 복수의 센서로부터의 성능 메트릭을 수신하는 단계를 포함하는 것을 특징으로 하는 복제서버의 자원 관리 방법.

【청구항 7】

제6항에 있어서,

상기 성능 메트릭은 상기 메인서버로의 응답시간 정보를 포함하는 것을 특징

으로 하는 복제서버의 자원 관리 방법.

【청구항 8】

제6항에 있어서,

상기 성능 메트릭은 상기 센서로부터 상기 메인서버로의 네트워크 상의 경로 정보를 포함하는 것을 특징으로 하는 복제서버의 자원 관리 방법.

【청구항 9】

제8항에 있어서,

상기 복제서버 개시 여부의 결정 단계는,

상기 성능 메트릭이 상기 제1 기준을 초과하는 경우 상기 성능 메트릭에 포함된 경로 정보로부터 추출된 링크들 중 상기 메인서버로부터 가장 먼 링크에 연결된 복제서버를 우선적으로 개시하는 것을 특징으로 하는 복제서버의 관리 방법.

【청구항 10】

네트워크에 산재된 복수의 센서로부터 성능 메트릭 정보를 수신하는 상태 모니터링부;

상기 성능 메트릭 정보와 사전 설정된 기준을 대비하여 DDoS 공격 여부를 판단하는 공격 판단부;

DDoS 공격이 있는 것으로 판단되는 경우 상기 네트워크상에 복제서버의 위치를 결정하는 복제서버 위치결정부; 및

상기 복제서버를 개시 및 중지하며 상기 복제서버의 자원을 관리하는 복제서버 자원관리부를 포함하는 복제서버 관리장치.

【청구항 11】

제10항에 있어서,

상기 성능 메트릭 정보는 상기 복수의 센서와 상기 메인서버 간의 링크 정보를 포함하고,

상기 복제서버 위치결정부는 상기 성능 메트릭 정보가 사전 설정된 기준을 초과하는 경우 해당 센서와 메인서버간의 링크 정보로부터 정체 상태인 링크 중 네트워크 경로 상의 거리가 가장 먼 링크를 식별하는 것을 특징으로 하는 복제서버 관리장치.

【청구항 12】

제10항에 있어서,

상기 복제서버 자원 관리부는 DDoS 공격 강도에 따라 상기 복제서버의 자원을 증감시키는 것을 특징으로 하는 복제서버 관리장치.

【청구항 13】

네트워크에 산재된 복수의 복제서버 팜에 인접하여 메인서버의 성능 메트릭을 측정하는 측정 센서; 및

상기 측정 센서로부터의 성능 메트릭 데이터로부터 상기 네트워크의 트래픽을 측정하고, 측정된 네트워크 트래픽에 따라 복제서버의 자원을 관리하는 복제서버 관리장치를 포함하는 복제서버 관리 시스템.

【요약서】

【요약】

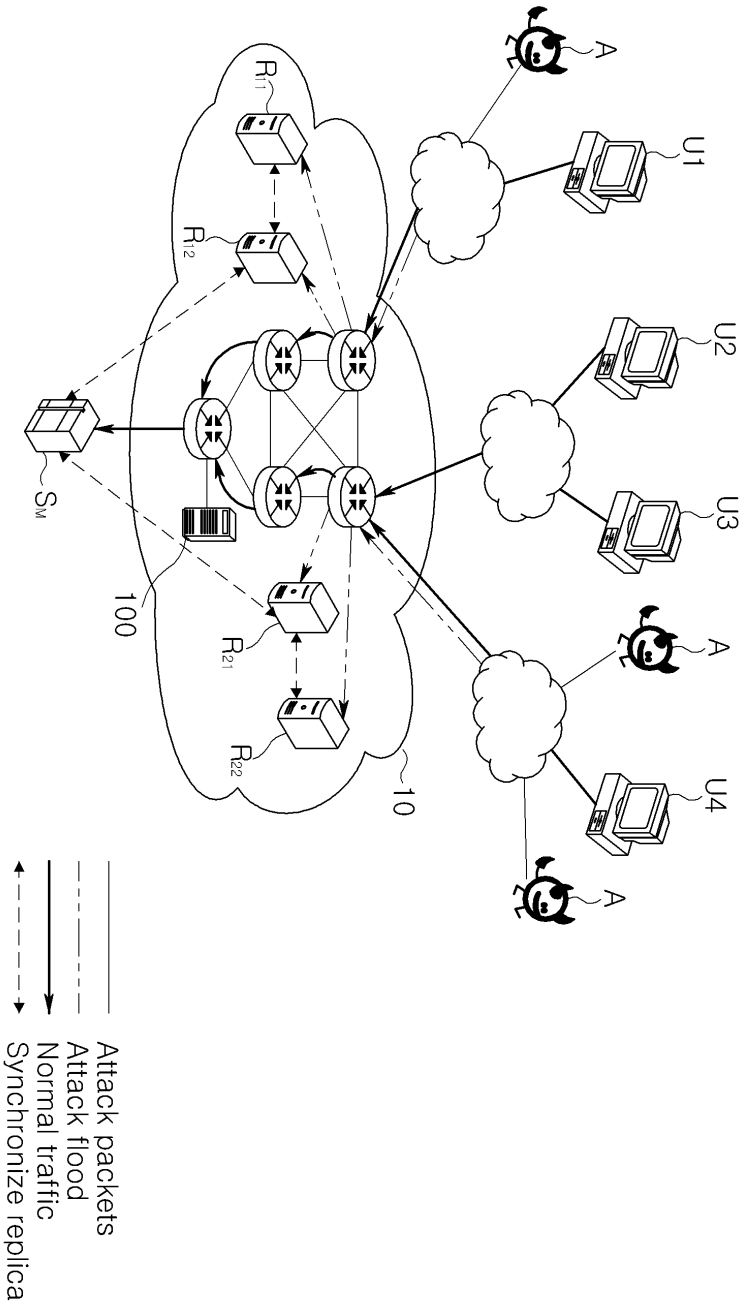
본 발명은 네트워크 트래픽 및 서버 부하에 따라 클라우드 기반의 복제서버를 동적으로 관리하는 방법에 관한 것이다. 본 발명은, 메인서버의 성능 메트릭을 측정하는 단계; 상기 메인서버의 성능 메트릭과 사전 설정된 제1 기준을 대비하여 상기 메인서버로의 경로 상에 복제서버의 개시 여부를 결정하는 단계; 및 상기 메인서버의 성능 메트릭과 상기 사전 설정된 제1 기준보다 낮은 제2 기준을 대비하여 상기 메인서버로의 경로 상의 복제서버의 중지 여부를 결정하는 단계를 포함하는 복제서버의 자원 관리 방법을 제공한다. 본 발명에 따르면, DDoS 등 공격자의 공격 강도에 따라 복제서버를 개시 및 중지하고 확장 및 축소하여 클라우드 기반 복제서버를 효율적으로 운용할 수 있게 된다.

【대표도】

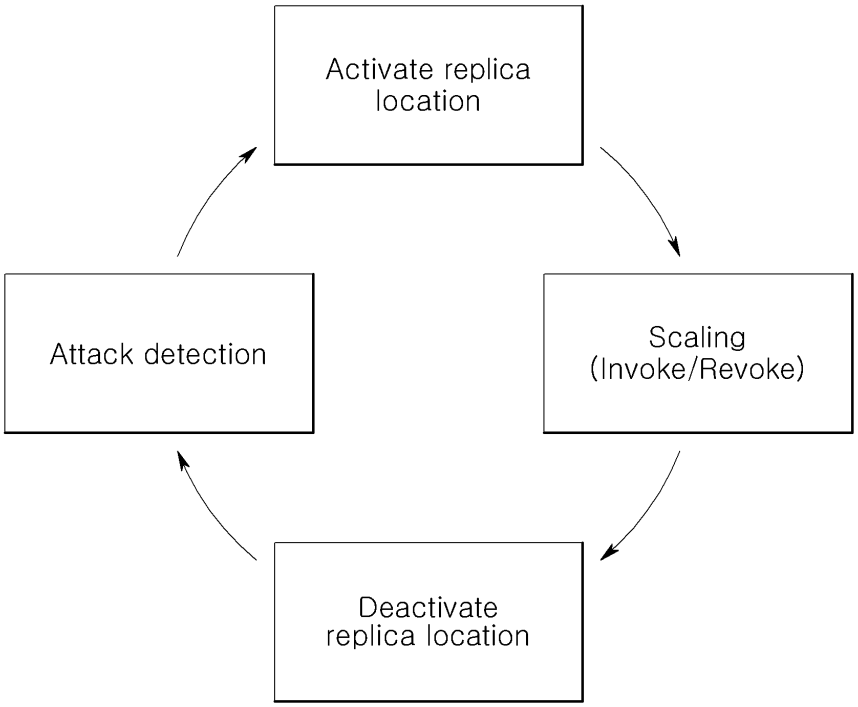
도 2

【도면】

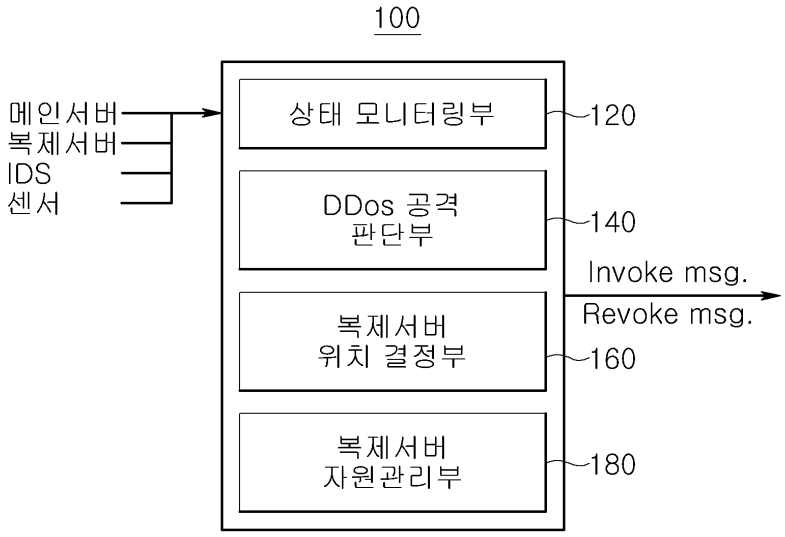
【도 1】


A
U1
U2
U3
A
U4
A
R11
R12
R21
R22
SM
100
10
Attack packets
Attack flood
Normal traffic
Synchronize replica

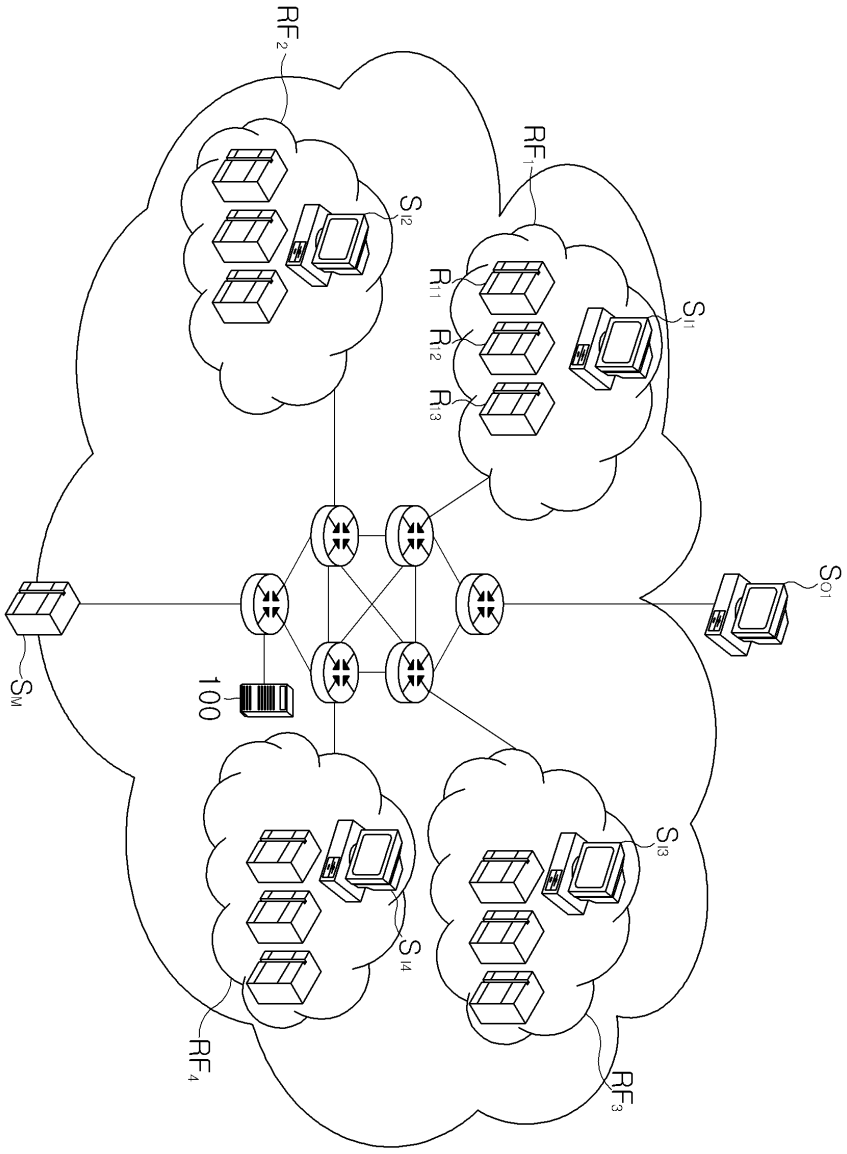
【도 2】


Activate replica location
Scaling (Invoke/Revoke)
Deactivate replica location
Attack detection

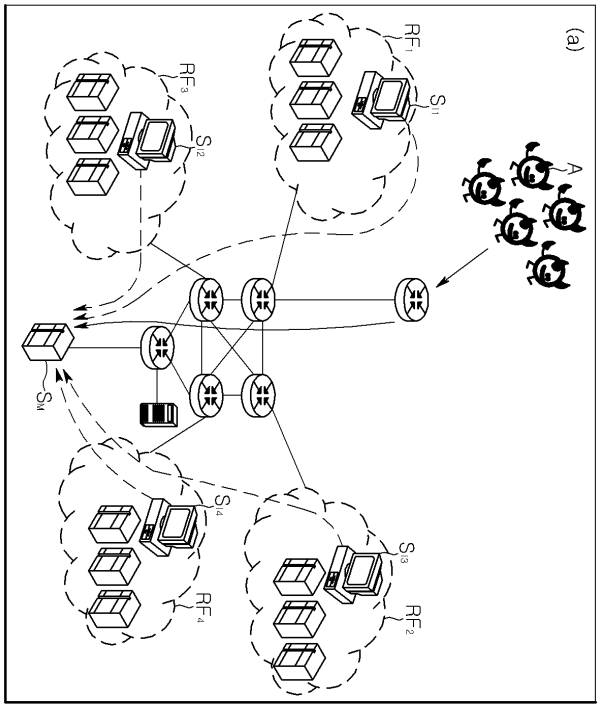
【도 3】

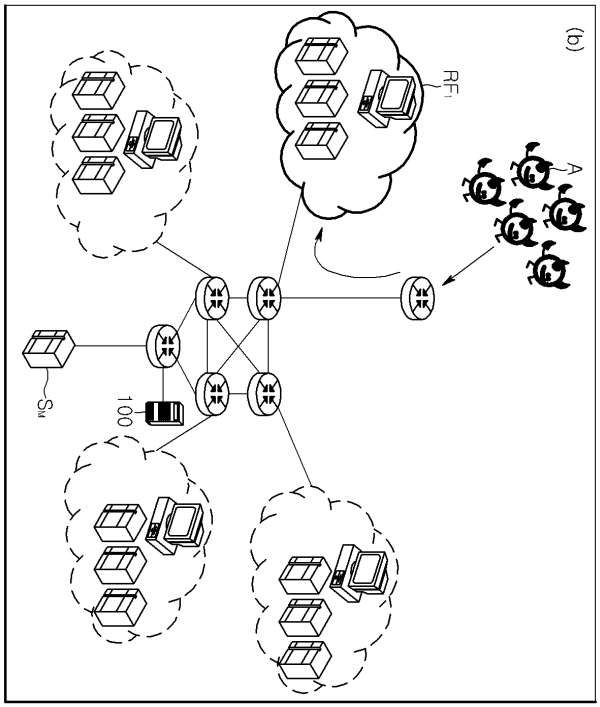

100
메인서버
복제서버
IDS
센서
상태 모니터링부
120
DDos 공격 판단부
140
Invoke msg.
Revoke msg.
복제서버 위치 결정부
160
복제서버 자원관리부
180

【도 4】

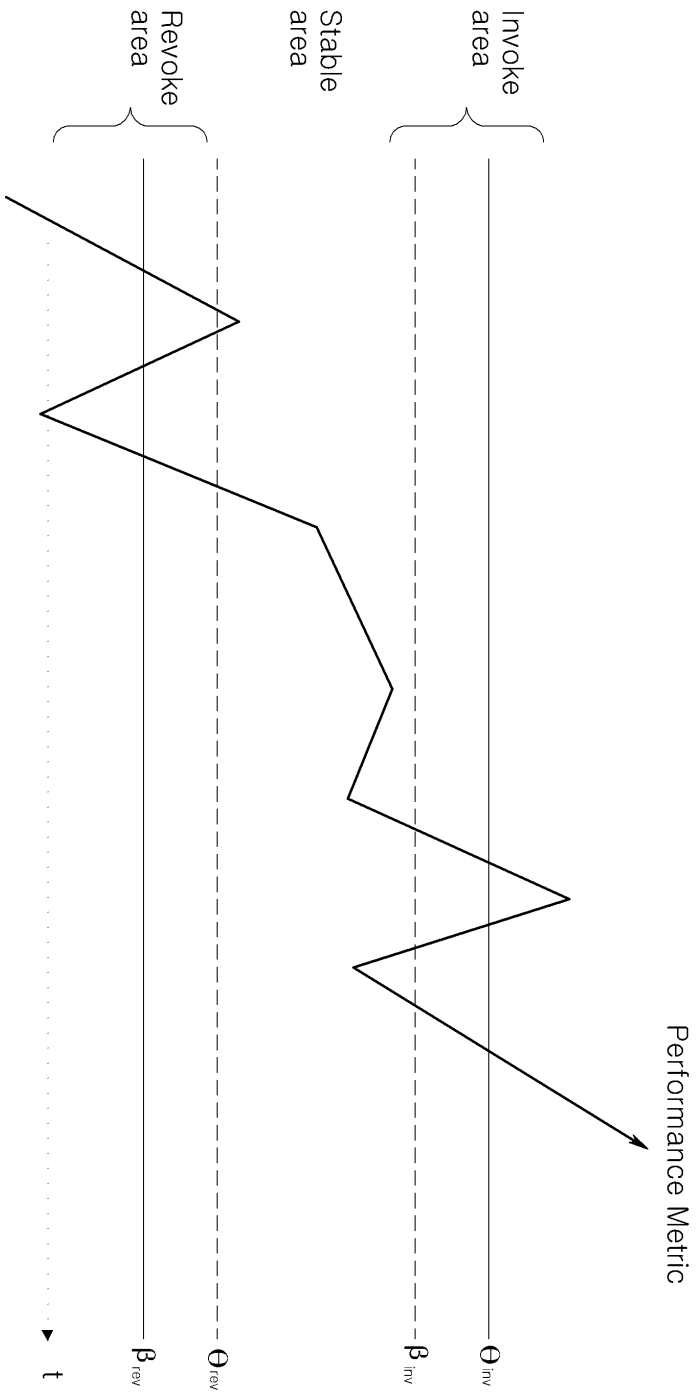


【도 5】

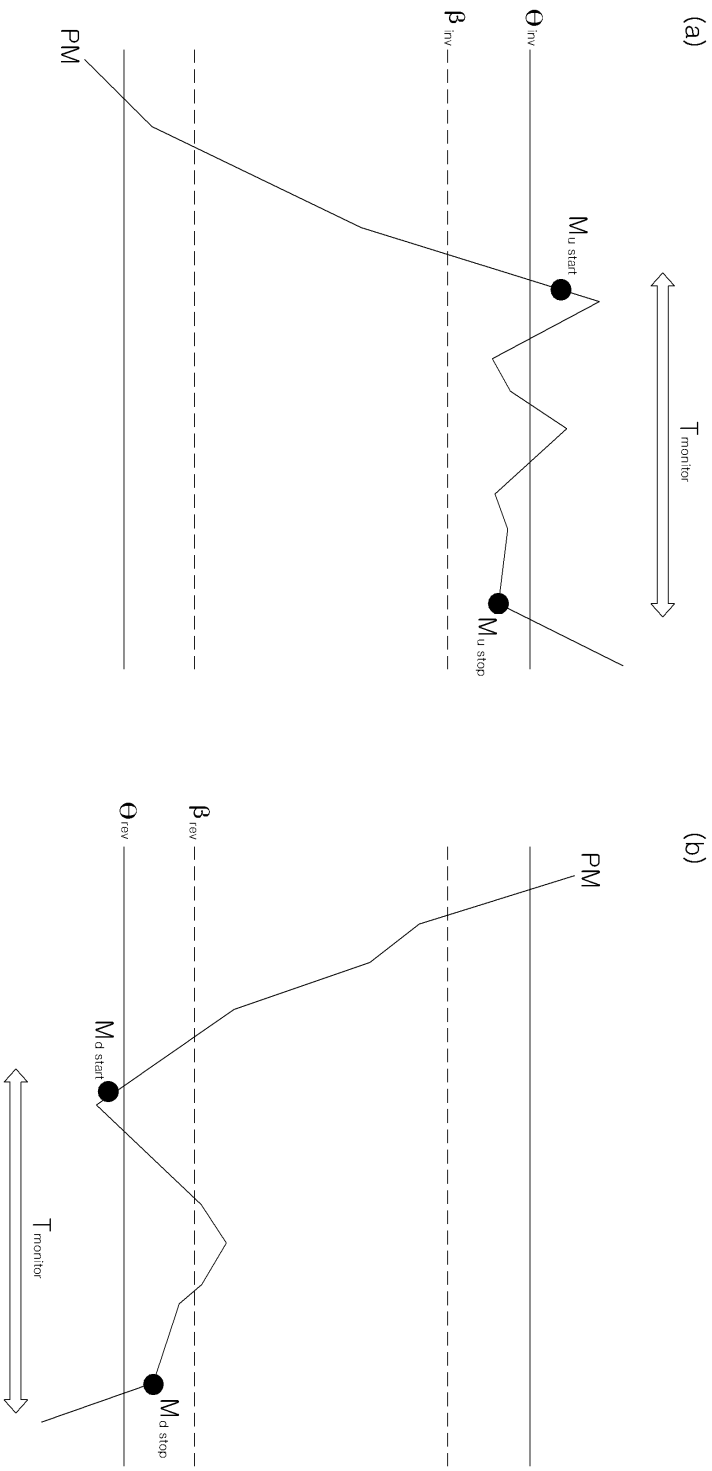




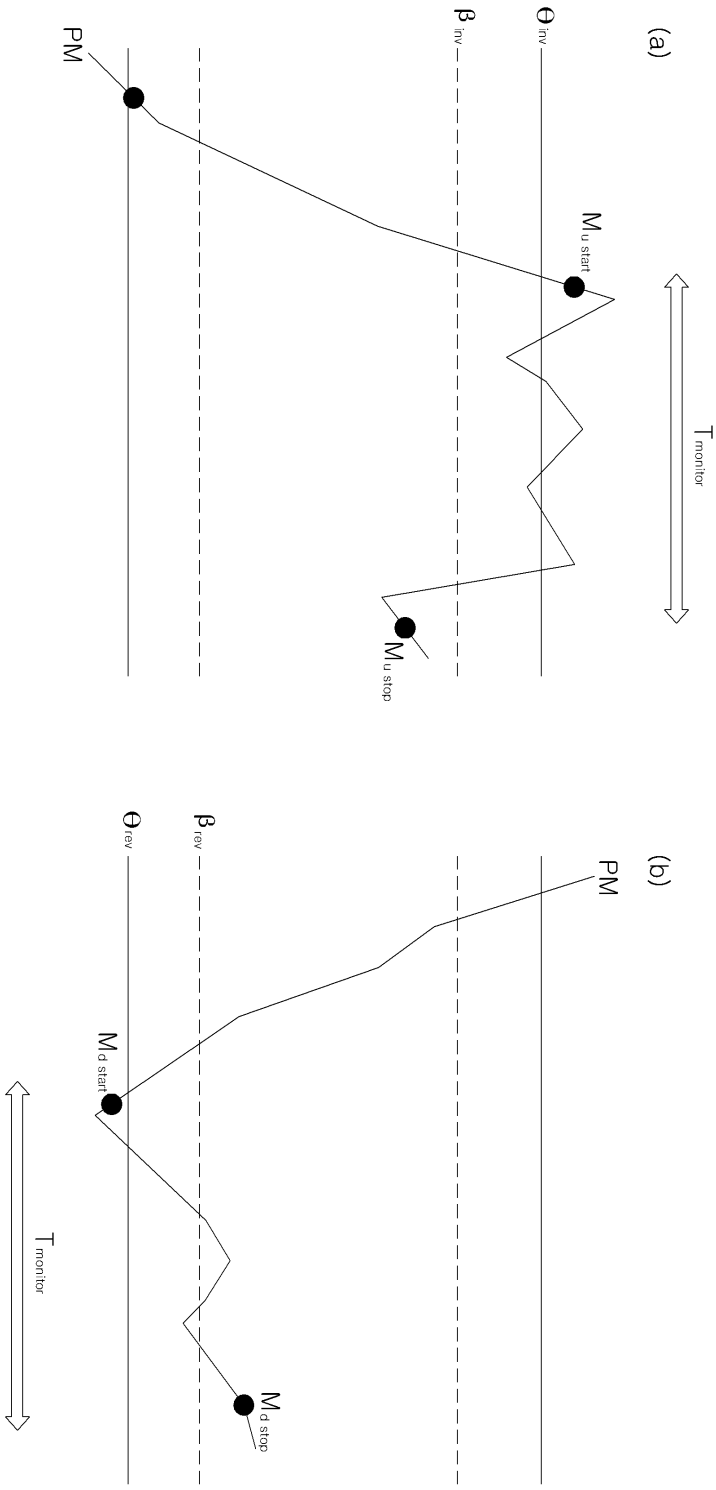
【도 6】


Invoke area
Stable area
Revoke area
Performance Metric
Θ_{inv}
β_{inv}
Θ_{rev}
β_{rev}
t

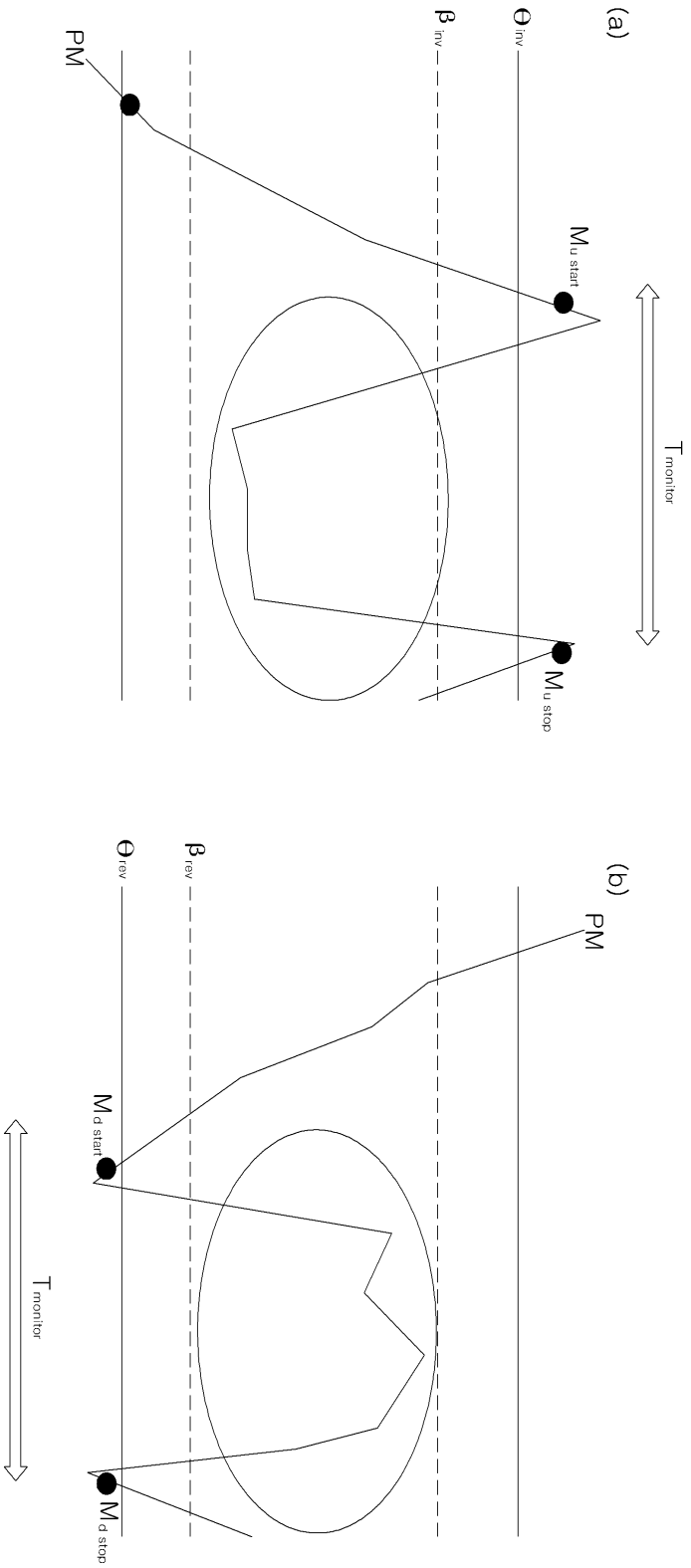
【도 7】


(a)
PM
θ inv
β inv
M u start
M u stop
T monitor
(b)
PM
θ rev
β rev
M d start
M d stop
T monitor

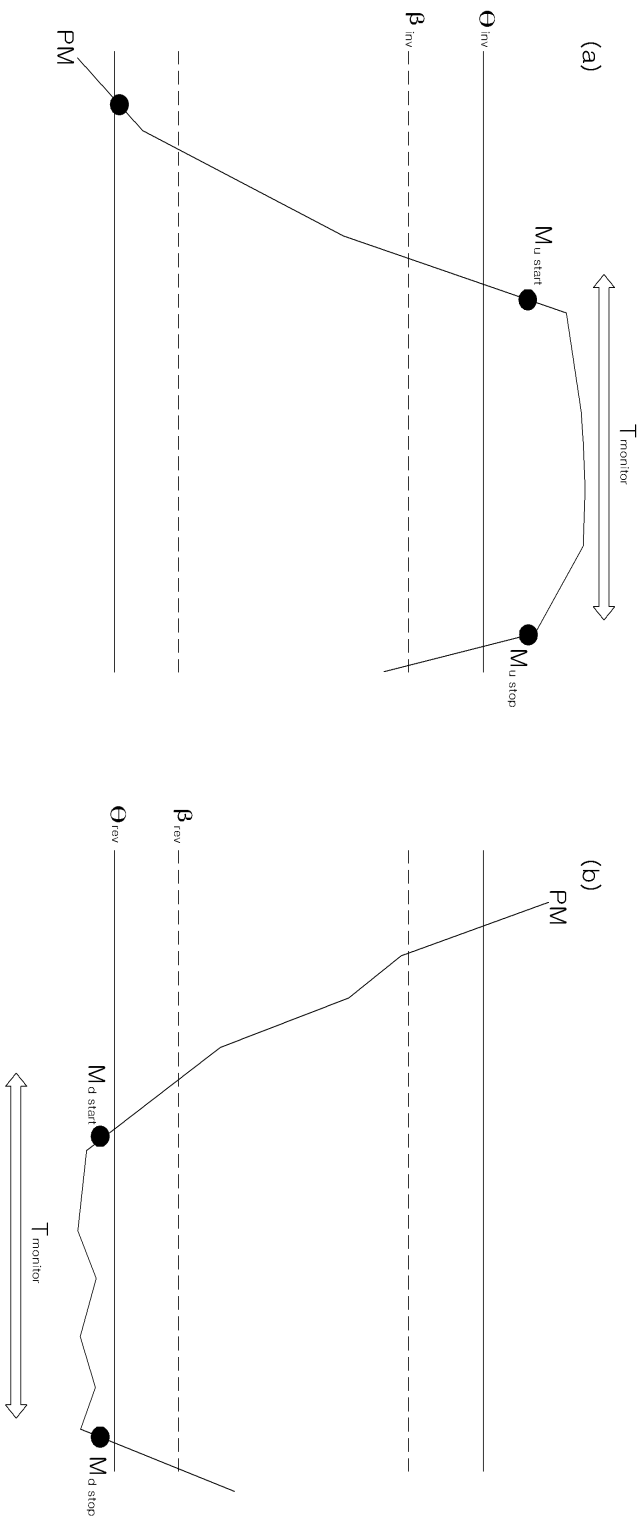
【도 8】


(a)
PM
β inv
Θ inv
M u start
M u stop
T monitor
(b)
PM
β rev
Θ rev
M d start
M d stop
T monitor

【도 9】


(a)
PM
β inv
Θ inv
M u start
M u stop
T monitor
(b)
PM
Θ rev
β rev
M d start
M d stop
T monitor

【도 10】


(a)
PM
β inv
Θ inv
M u start
T monitor
M u stop
(b)
PM
β rev
Θ rev
M d start
T monitor
M d stop