(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(51) Int. Cl.[7]
G06F 17/00

(11) 공개번호 특2000-0076638
(43) 공개일자 2000년12월26일

(21) 출원번호 10-2000-0006191
(22) 출원일자 2000년02월10일

(30) 우선권주장 9/247893 1999년02월10일 미국(US)
(71) 출원인 루센트 테크놀러지스 인크
미합중국 뉴저지 머레이 힐 마운틴 애비뉴 600 (우편번호 : 07974-0636)
(72) 발명자 에드워드토마스제이
미국,98112,워싱턴,시애틀,마드로나플레이스이스트131
율렌데이비드
미국,98052,워싱턴,레드몬드,NE185애비뉴,6464
(74) 대리인 이병호

심사청구 : 없음

(54) 서비스 목표 달성에 기초한, 작업의 경합 클래스에 서버자원의 동적인 배분 방법 및 시스템

요약

작업 처리 설비에서 비충족된 서비스 필요성에 기초하여 어떤 작업유형을 수행하는 서버 풀에 배분되는 다수의 서버의 조정을 위한 설비. 서버는 사람 및 기계식인 서비스 에전트를 포함한다. 서버 배분기 및 대응하는 서버 배분 방법은 서버 풀에서 서버 이용성을 조건부로 조정하는 작업 분배기 혹은 자동 호 분배기(ACD)에서 각각 채용될 수 있다. 서버 배분기는 작업유형에 따른 각 서버의 임계값에 대해 작업유형에 따른 복합 선호값을 비교한다. 복합 선호값이 서버의 임계값 이상인 것으로 서버 배분기가 판정하였을 때, 서버 배분기는 그 작업유형에 대해 서버 풀 내에 서버가 포함될 수 있음을 나타낸다. 각각의 서버는 상이한 작업유형마다 선호값 및 임계값을 갖는다. 선호값의 크기는 작업유형에 따른 친화도를 나타낸다. 서버의 임계값은 그 작업유형을 갖는 작업의 수행에 대한 저항도를 나타낸다. 통상, 서버는 해당 작업유형에 대해 서버의 임계값을 서버의 복합 선호값이 초과하는 것으로 서버 배분기가 판정하지 않았다면, 선호값이 임계값 미만인 작업유형의 작업을 받아들이지 않는다. 서버 배분기 및 대응하는 서버 배분 방법은 사용자가 선택할 수 있는 다수의 입력으로부터 복합 신호값을 계산하고 사용자가 선택할 수 있는 다수의 함수를 이용한다.

대표도

도1

색인어

서버 배분 시스템

명세서

도면의 간단한 설명

도 1은 본 발명의 실시예에 따라, 서버 배분기(102)가 실행할 수 있는 샘플 범용 컴퓨터 시스템(100)의 블록도.

도 2는 본 발명의 실시예에 따라, 작업 처리 설비에서 수행되는 작업유형을 나타내는 여러 가지 작업유형(202)에 대한 여러 서버의 선호를 도시한 서버 선호표(200).

도 3은 본 발명의 실시예에 따라, 작업유형에 따른 복합 선호값에 결합된 작업 처리 설비에서 수행되는 작업유형에 따른 서버 선호 및 임계를 예시한 표(300)를 도시한 도면.

도 4는 본 발명의 실시예에 따라, 서버 배분기(102) 또는 이의 대응하는 방법을 사용하여 복합 선호값의 생성을 도시한 흐름도.

도 5a 및 도 5b는 본 발명의 실시예에 따라 서버 급파 과정을 도시한 흐름도.

도 6a-6d는 본 발명의 실시예에 따라, 서버 배분기(102)에 의해 시간에 걸쳐 여러 서버 풀(603-605)에 서버의 배분 변화를 도시한 도면.

도 7은 본 발명의 대안 실시예에 따라, 동적 선호값 열(701)을 포함하는 테이블(700)을 도시한 도면.

도 8은 본 발명의 실시예에 따라, 작업 처리 설비에서 수행되는 작업유형을 나타내는 여러 가지 작업유형(202)에 대한 여러 서버의 선호를 예시한 서버 선호표(800).

도 9a는 본 발명의 대안 실시예에 따라, 임계값 결정기(911) 및 복합 선호 평가기(910)를 포함하는 또 다른 범용 컴퓨터 시스템(100)의 블록도.

도 9b는 본 발명의 실시예에 따라, 복합 선호 평가기(910)에 의해 수행되는 동작의 일부를 도시한 흐름도.

도 10은 본 발명의 실시예에 따라, 복합 선호값에 결합된 또 다른 작업 처리 설비에서 수행되는 작업유형에 따른 서버 선호 및 임계를 예시한 표(1000)를 도시한 도면.

도 11은 본 발명의 실시예에 따라, 분배 서버 배분 시스템(1101)의 도면.

〈도면의 주요 부분에 대한 부호의 설명〉

100 : 컴퓨터 시스템

102 : 서버 배분기

110: CPU

120 : 입력/출력장치

130 : 메모리

201 : 서버 ID

202 : 작업유형

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 서비스 제공자 분배 분야에 관한 것이다.

많은 서비스 조직은 어떤 목표를 달성하기 위해서 이들의 서버를 동적으로 배분할 필요가 있다. 이러한 배분은 통상 수동으로 수행된다. 서버는 사람 및 기계식인 서비스 에전트를 포함한다. 전달된 서비스에 대한 외부 수행 측정은 이용비용 및 노동비용과 같은 내부 비용측정을 좌우한다. 이러한 외부 측정은 흔히 어떤 트랜잭션들을 원하는 목표를 달성하는 것 혹은 달성하지 못하는 것으로 분류하고, 목표를 충족하는 트랜잭션 비율을 결정하는 것으로 구성된다. 이러한 비율을 서비스 수준이라고 한다. 서비스 수준은 일정 시간에 대해서 혹은 몇 개의 트랜잭션에 대해 측정된다.

서비스 수준의 예로서는, 더 이상의 활동없이 해결된 고객문제 백분율, 약속된 시간 내에 탑승자에 도착하는 급파된 택시의 백분율, 다른 서버에 보내거나 의뢰를 요구함이 없이 적합한 대리인이 처리한 전화 호의 비율, 지연없이 서버에 접속될 수 있는 전화 호의 비율, 24시간 내에 응답되는 전자우편 요청의 비율, 평일에 특정 버스노선 상의 도시 버스의 정기적인 출발 백분률, 고객 불만을 초래하지 않게 처리된 트랜잭션 비율, 완전히 적격의 서버에 의해 처리된 우선 고객 호의 비율, 스페인어에 유창한 서버에 의해 처리된 스페인 고객 백분율, 서버에 접속하기 전에 고객이 포기하지 않은 전화 호의 백분율, 전화국 교환기에서 차단하지 않은 고객 조회 전화 호의 백분율, 디스플레이를 기다리는 동안 중지되지 않은 자급 WWW(World Wide Web) 페이지와의 고객 세션의 백분율, 전화를 받고 있는 동안 즉시 완료될 수 있는 전화를 통한 고객 요청의 백분율, 요청시간부터 반시간 이내에 처리되는 대부 신청 백분율, 및 8초 내에 응답되며 적격의 서버에 의해 적합하게 처리되는 우선 전화 호의 백분율이다.

이러한 정황에서 서비스 수준에 대한 서비스 조직의 목표는 서비스 수준의 특정하게 원하는 값이다. 달성된 서비스 수준이 그 목표에 대해 원하는 서비스 수준만큼 높은 경우 목표가 달성되었다고 한다. 반대로, 실현된 서비스 수준이 원하는 서비스 수준 미만이면 목표는 달성되지 않았다고 한다. 예를 들면, 우선 고객으로부터의 전화 호를 적어도 85%로 12초 내에 응답하는 목표는 당일에 우선 고객으로부터 전화 호 중 87%를 12초 내에 응답하였다면 달성될 것이며, 반대로, 이러한 호의 단지 84%만이 12초 내에 응답되었다면, 목표는 달성되지 않는 것이다. 이러한 구조에서, 목표는 얻어지든지 얻어지지 않든지 어느 하나이다. 더욱이, 목표보다 훨씬 높은 서비스 수준의 달성으로부터 별도의 이점이 전혀 생기지 않는다.

서비스 수준 목표는 자원의 예외적인 배분을 포함하여 임시 작용을 위한 기준에서 일반적으로 사용되는 것이다. 이러한 정황에서 다른 관련된 수행 목표는 평균 대기시간에 대한 최대값, 현재 대기상태, 작업 시프트에서 트랜잭션 수율의 최소수, 및 질 서비스를 표시하는 여러 가지 결정가능한 인자를 포함하는 측정을 포함할 수 있다.

서비스 유형에 할당되는 서버 자원의 수는 그 유형의 서비스에 대해 달성되는 서비스 수준에 영향을 미친다. 그러한 경우일 때, 운영은 서비스 수준 목표를 달성하기 위해서 주 작업에 서버를 재배분할 수 있다. 그러나, 이러한 재배분은 다른 작업에 대한 서비스 수준이 나빠지기 때문에, 일반적으로 기회비용을 발생한다. 이러한 기회비용을 기업의 운영규칙에 의해 제정될 수도 있는 적합한 우선계층에 근거하여 정당화할 수 있다.

예를 들면, 콜 센터 내의 서버가 대부 서비스와 판매 서비스 트랜잭션을 처리할 수 있다고 가정한다.

보다 많은 서버가 판매활동에 배분되었을 때, 판매 서비스 트랜잭션은 응답 지연에 대해 보다 높은 서비스 수준을 경험한다. 즉, 각각의 판매 호에 응답하는데 필요한 시간양이 감소된다. 한편, 대부 서비스 호는 신속하게 응답되지 않아, 대부 서비스 트랜잭션에 대한 서비스 수준이 감소한다. 서비스 조직은 이에 대해서, 기존의 고객이 대부 회사를 바꿀 일은 거의 없기 때문에 대부 서비스는 비교적 덜 중요하며, 현재 회사는 신규 고객의 호에 즉시 응답하지 않는다면 그들의 업무를 경쟁사에 쉽게 넘길 수 있을 이들 신규 고객을 잡을 필요가 있다고 말함으로써 합리화할 수 있다. 서비스 조직은 판매에서 목표에 도달할 수 없는 대가를 치르지 않고 대부 서비스의 목표를 만족시키고자 한다. 판매 목적이 위태로운 상태에 있지 않을 때, 그러나 대부 서비스는 이의 목표를 총족시킬 수 없을 때, 서비스 조직은 보다 많은 자원을 대부 서비스에 배분하기를 원한다. 서비스 조직은 이들 두 목표를 만족시키기를 원하지만, 판매 목표는 대부 서비스보다 더 중요하여 선매권이 있다. 즉, 운영이 하나의 목표만을 만족할 수 있다면 그것은 판매 목표일 것이다.

한 세트의 서비스 목표 내에, 단기간의 "기회 윈도우"를 갖는 작업에 관계된 목표만이 아니라 장기간의 기회 윈도를 갖는 작업에 대한 목표가 있을 수 있다. 단기간 기회 윈도우 작업의 예는 전화 호이며, 이것은 수십초 내에 응답되지 않으면 참을성 한계로 콜러가 단념할 수 있다. 장기간 기회 윈도우 작업의 예는 완료하는데에 있어 다음 빌링 사이클까지 갖는 빌링 정산에 관한 고객으로부터 서신이다. 따라서, 운영은 단기간 기회 윈도우 작업에 관련된 서비스 목적을 달성함에 있어 매우 짧은 시간 프레임을 갖는 한편, 장기간 기회 윈도우 작업에 관련된 서비스 목적을 달성하는 기회의 윈도우는 훨씬 넓다. 그러므로, 단기간 기회 윈도우 작업에 대한 서비스 목표를 달성하기 위해서 장기간 기회 윈도우 작업을 수행하고 있는 보다 많은 자원을 단기간 기회 윈도우 작업에 일시적으로 할당하는 생산적인 운영전략이 있을 수 있다.

많은 서비스 조직은 원하는 결과를 달성하기 위해서 그들의 서버를 동적으로 배분할 필요가 있다. 수동 개입은 통상 이러한 재배분을 달성한다. 대부분의 서버는 이들이 작업유형에 대해 필요한 기능 프로파일을 만족할지라도 정지상태에 있고 어떤 형태의 작업에 사용하지 못한다. 그러나, 조건이 보장된다면 이들은 작업유형에 배분될 수 있을 것이다. 서버가 그 유형의 작업에 통상 이용되지 않는 경우에 그 서버를 그 작업에 대한 백업 서버라고 한다.

보다 많은 서버 자원을 활동에 배분하고자하는 것은 서버 자원이 수행할 수 있는 대안 활동 및 이러한 대안 활동에 대한 요구를 조건으로 한다. 이들 대안 활동 각각은 잠재적으로 여러 가지 서비스 수준에 관련되며, 그 각각은 목표 및 달성 수준을 갖는다. 결국, 자원 재배분은 서버 자원 각각에 관련된 모든 대안 작업유형에 따른 서비스 측정에 의존할 수 있다. 이러한 잠재적으로 복잡한 배분기능을 수동으로 수행하는 것은 현저하게 부분 최적화 결과를 초래할 수 있다. 흔히, 수동 배분은 너무 늦어 재분배된 서버가 목표를 달성하는데 충분히 빨리 이들의 선호된 작업으로 복귀하지 않았을 때 더 많은 문제에 이르게 된다.

대부분의 자동 호 분배기(ACD)는 "호 오버플로"라고 하는 특징을 갖고 있다. 호 오버플로는 선택된 상태를 조건으로 하는 호 대기에 대해 서버 그룹을 사용할 수 있게 한다. 그러나, ACD는 일반적으로 적합한 동적 상태 나름대로 호 수신으로부터 서버를 중지상태에 있게 하는 설비가 없다. 일반적으로, ACD는 서버에 호를 얻게 하는 처리의 명백한 제어를 필요로 하나, 호에 대해 서버를 얻는 처리의 명백한 제어가 거의 제공되지 않는다. 그러나, 서버의 작업유형에 따른 관점은 작업의 총체적인 완성을 최대화하기 위해서 대안이 되는 작업간 선택에 비판적이다.

각 개개의 서버의 기능이 서버를 배분할 때 고려되는, 기능에 기초한 라우팅의 출연으로 상황을 복잡하게 한다. 기능에 기초한 라우팅은 기능을 구별하는 것을 무시하는 모노리식 풀로 자원의 분할을 간단하게 하는 것을 감당할 수 없다. 이러한 이유로, 기존의 ACD 및 워크플로 자동 시스템은 이러한 필요성을 충족시킬 수 없다.

작업 분배 시스템은 전술한 바와 같이, 서버의 재배분을 달성하기 위해서 서버 "기능"을 사용자가 조작하게 한다. ACD 및 다른 자동 작업 분배기는 단지 기능 요구에 대해서만 그리고 분배 맵 상의 어떤 류의 대기점에 대해서만 "서비스 수준"을 보고할 수 있다. 그러나, 때때로 "기능"은 실제로는 서버의 능력의 속성 대신 작업유형이다. 이들 종래의 해결책은 서비스 조직에 중요한 서비스 수준의 감시 및 제어에 대해 잠재적으로 심각한 한계를 이룬다.

이러한 환경에서, 서비스 조직은 서비스 목표 및 이들 목표에 대해 표명된 우선도 면에서 최상의 작업 항목이 서버에 확실하게 주어지는 계층 내에서 작업 활동에 우선적인 처리를 제공하기를 원한다. 작업 처리 설비에서 작업유형에 속하는 결정할 수 있는 서비스 목표의 달성 혹은 비달성에 기초하여 작업유형에 대해 사용할 수 있는 서버의 풀을 동적으로 확장하는 자동화된 시스템은 현저한 유용성을 가질 것이다.

발명이 이루고자하는 기술적 과제

본 발명은 작업 처리 설비에서 비충족된 혹은 위태롭게된 서비스 필요성에 기초하여 특정 작업유형에 사용할 수 있는 다수의 서버를 조정하는 설비를 제공한다. 서버 배분기 및 대응하는 서버 배분방법은 서버 풀에 서버 배분을 조건부로 조정하는 작업 분배기에서 각각 채용될 수 있다. 작업이 인커밍 전화 호인 공통의 작업 분배기의 예는 자동 호 분배기("ACD")이다.

서버 배분기는 본 발명의 실시예에 따라 작업유형에 따른 각 서버의 임계값에 대해 작업유형에 따른 복합 선호값을 비교한다. 복합 선호값이 서버의 임계값 이상인 것으로 서버 배분기가 판정하였을 때, 서버 배분기는 그 작업유형에 대해 서버 풀 내에 서버가 포함될 수 있음을 나타낸다. 각각의 서버는 상이한 작업유형마다 선호값 및 임계값을 갖는다. 선호값의 크기는 관련된 작업유형에 따른 친화도를 나타낸다. 서버의 임계값은 그 작업유형의 작업 수행에 대한 저항도를 나타낸다. 통상, 서버는 본 발명의 실시예에 따라 해당 작업유형에 대해 서버의 임계값을 서버의 복합 선호값이 초과하는 것을 서버 배분기

가 판정하지 않았다면, 선호값이 임계값 미만인 작업유형의 작업을 받아들이지 않는다.

서버 배분기 및 대응하는 서버 배분 방법은 본 발명의 실시예에 따라 사용자가 선택할 수 있는 다수의 입력으로부터 복합 선호값을 계산하고 사용자가 선택할 수 있는 다수의 함수를 이용한다. 복합 선호값은 서비스 수준, 큐(queue) 상태, 알람 상태, 원하는 서비스 수준, 응답 지연, 호 방기율, 운영자 개입과 같은 다양한 입력으로부터 계산된 작업유형에 따른 동적인 선호값을 포함할 수 있다. 본 발명의 일 실시예에서, 복합 선호값 함수의 입력은 그 작업유형에 따른 서버의 선호도 및 그 작업유형의 동적 선호값을 포함한다. 따라서, 본 실시예에서의 복합 선호값은 여러 가지 종류의 작업에 대한 서버의 특정한 선호 및 목표를 포함하는 폭넓은 고찰에 기초하여 여러 가지 종류의 작업을 처리하는 작업 처리 설비의 현재의 일반적인 선호를 고찰한다.

서버 분배기는 본 발명의 실시예에 따라, 입력으로서 동적 선호값 및 작업유형에 따른 서버의 선호값을 수신하는 복합 선호값에 대한 무수한 함수 중 어느 것을 이용할 수 있을 것이다. 간단한 복합 선호값 함수는 작업유형에 따른 서버의 선호값과 작업유형의 동적 선호값을 합산하는 것을 포함한다.

서버 배분기 및 이의 대응하는 서버 배분 방법은 충족되지 않은 목표를 갖는 작업유형에 따른 서버 풀의 확정을 달성하기 위해서 작업의 동적인 방편 요구를 적용한다. 서버 배분기 및 서버 배분방법은 작업을 진척시키시는 동작규칙의 달성에 적합한 백업 서버를 작업 처리 설비가 이용하게 한다. 서버 배분기 및 서버 배분 방법은 작업 처리 설비 및 서버 조직에 무모한 제약없이 백업 서버를 이용할 수 있게 한다.

본 발명의 실시예는 다음 도면에 관련하여 이하 설명한다. 도면에서 유사 구성요소 및 단계는 동일한 참조부호를 갖는다.

발명의 구성 및 작용

본 발명은 서비스 목표의 달성을 최대화하기 위해서 작업 처리 설비에서의 비충족된 서비스 요구에 기초하여 특정 작업유형에 사용할 수 있는 서버 풀(pool)을 조정하기 위한 설비를 제공한다. 서버는 사람 및 기계식인 서비스 에전트를 포함할 수 있다. 서버 분배기 및 대응하는 서버 분배방법 각각은 서버 풀에 서버 분배를 조건부로 조정하는 작업 분배기에서 채용될 수 있다. 서버 풀은 한 세트의 서버를 나타내는데, 각각의 서버는 이를 어떤 작업유형 수행에 사용할 수 있게 하는데 충분한 선호도(preference)를 갖는다. 서버는 서버 풀에 전혀 속하지 않을 수도 있고 하나 이상의 서버 풀에 속할 수도 있다. 작업이 인커밍 전화 호인 공통의 작업 분배기의 예는 자동 호 분배기("ACD")이다.

각각의 서버는 상이한 작업유형들에 대한 선호값과 임계값을 갖는다. 선호값의 크기는 관련된 작업유형에 따른 친화도를 나타낸다. 서버의 임계값은 그 작업유형의 작업 수행에 대한 저항도를 나타낸다. 통상, 서버는 선호값이 임계값 미만인 작업유형의 작업을 받아들이지 않는다. 그러나, 작업유형에 따른 복합 선호값이 그 작업유형에 대해 서버의 대응하는 임계값보다 낮지 않은 것으로 판정하였을 때, 서버 분배기는 그 서버가 그 작업유형에 따른 서버 풀에 포함될 수 있음을 나타낸다.

서버는 일 실시예에 따라, 모든 작업유형에 대해 하나의 임계값만을 가질 수도 있다. 또한, 임계값 및 서버 선호값은 다른 자동화된 제어 시스템 및 감독 운영자로부터의 수동입력을 동적 서비스 및 이용특성에 통합하는 시스템에 동적으로 기초할 수 있다.

서버 분배기는 사용자가 선택할 수 있는 다수의 함수를 사용하여 사용자가 선택할 수 있는 다수의 입력으로부터 복합 선호값을 계산할 수 있다. 복합 선호값은 서비스 수준, 큐(queue) 상태, 알람 상태, 원하는 서비스 수준, 응답 지연, 호 방기율, 운영자 개입과 같은 다양한 입력으로부터 계산된 작업유형에 따른 동적인 선호값을 포함할 수 있다. 일 실시예에서, 복합 선호값 함수의 입력은 그 작업에 대한 서버의 선호도 및 그 작업유형의 속성인 동적 선호값을 포함한다. 동적 선호값은 통상 서버 선호도에 관계없이 다른 유형의 작업에 비교해 특정 작업유형을 서버가 수행하게 하는 작업 처리 설비의 선호도를 나타낸다.

서버 분배기는 입력으로서 동적 선호값 및 작업유형에 따른 서버의 선호값을 수신하는 복합 선호값에 대한 무수한 함수 중 어느 것을 이용할 수 있을 것이다. 이들 함수는 입력으로서 이를테면 작업 도착율, 목표달성, 및 기타 서버 분배기를 지원하는데 필요한 관련 정보를 수신할 수도 있다. 간단한 복합 선호값 함수는 작업유형에 따른 서버의 선호값과 동적 선호값을 합하는 것을 포함한다. 서버가 기본적으로 부적합한(선호값이 없는) 작업유형은 통상(적어도 바람직한 실시예에서) 동적 선호값에 관계없이 복합 선호값을 전혀 수신하지 않는다.

많은 유형의 속성의 무리는 기업의 작업 운영이 어떻게 조직되기를 원하는가에 따라 작업유형을 결정할 수 있다. 작업유형은 필요한 태스크, 제품, 통신매체, 지리, 정부관할, 인구통계, 언어, 숙련도, 서비스 유형, 고객 프로파일, 판매처리단, 및 규제 클래스의 조합에 의해 결정될 수 있을 것이다. 다음의 예를 고찰한다. 작업 분배 운영은 인바운드 전화 호, 전자 메일, WWW 히트, 미국 우편 서비스 메일, 기타 메일과 같이, 매체 유형으로 방편을 구별하고자 할 수 있다. 예상 고객, 신규고객, 기존 고객, 탈퇴 고객, 및 탐탐치 않은 고객과 같은, 고객 관계의 유형에 기초하한 상이한 서비스 프로파일을 제공하고자 할 수도 있다. 특정 판매 태스크를 다른 주(state)의 고객을 위한 임의의 서버가 처리할 수 있어도 클라이언트가 거주하는 주에서 면허가 있는 서버를 이 태스크가 요구하는 것과 같은, 규제 제약에 기초하여 상이한 처리를 요할 수도 있을 것이다. 셀룰라 전화회사에의 조회(inquiry)는 어떤 공익사업 위원회에 의한 규정에 따른 상이한 처리 제약을 만족시키기 위해서 지리적으로 분류될 수도 있다. 전망 및 고객은 시장분할 전략을 용이하게 하기 위해서 여러 가지 시장분할에 따라 분류될 수도 있을 것이다. 나이, 수입, 교육, 직업, 및 가족 규모와 같은 인구통계 요소를, 다양한 제품라인에 대한 다양한 목표 시장에서 고객들을 확인하는데 사용할 수 있을 것이다. 작업분배는 주소변경, 청구서 조회, 대부 요청, 권유 판매, 비권유 판매, 제품라인에 의한 판매, 기술지원, 불만, 문제 확인, 및 문제해결과 같은, 태스크 유형에 의한 분류를 필요로 할 수도 있을 것이다. 통상, 이러한 작업 및 고객 속성의 조합은 매우 복합한 한 세트의 태스크 유형을 초래한다. 본 발명은 소규모의 정적인 풀에 서버 집단을 무모하게 분

할하는 비능률없이 이러한 복잡성을 지원한다.

사용자가 선택할 수 있는 복합 선호값 함수는 본 발명의 바람직한 실시예에 따라, 동적 선호값 및 서버의 선호값에 관하여 단조롭게 증가해야 한다. 즉, 서버의 선호값이 증가하거나 감소한다면, 다른 모든 것들이 동일한 상태로 복합 선호값이 증가 혹은 감소한다. 마찬가지로, 동적 선호값이 증가하거나 감소한다면, 복합 선호값은 모든 다른 것들은 동일한 상태로 각각 증가하거나 감소한다. 물론, 숙련된 자이면 복합 선호값 함수는 서버 분배기 및 작업 처리 시스템의 필요에 따라 부가적인 입력 혹은 소수의 입력이라도 이를 이용할 수 있음을 알 것이며, 이들 입력으로부터 복합 선호값 함수를 생성하는 다양한 적합한 방법을 또한 알 것이다. 일부 실시예는 다른 원하는 동작(operating) 특성을 달성하기 위해서 복합 선호값 함수의 선택을 동적으로 조작하는 최우선의 제어법칙을 이용할 수도 있다. 다른 원하는 동작(operating) 특성의 예는 관련된 목표를 초과하여 서비스 수준 수행의 동등한 분배, 서버의 동등한 이용, 서비스 수준 달성에서 변동의 제동, 서비스 무위(idleness)에서 변동의 제동, 및 모드에 특정한 행동을 포함한다. 최우선의 제어법칙은 본 발명의 이들 실시예에서 사용자가 선택할 수 있는 복합 선호값의 방법에 더하여 또 다른 층의 제어를 구성할 수 있다.

복합 선호값에서 동적 성분을 계산하는 방법 및 시스템은 여기 참고로 포함시킨, 모자익스사에 양도된, 1999년 2월에 출원된 미국 특허 출원 번호.____, "Hierarchical Service Level Remediation for competing Classes Based Upon Achievement of Service Level Goals"에 개시되어 있다.

서버 분배기 및 이의 대응하는 서버 분배방법은 비충족된 혹은 최저로 달성된 목표를 갖는 작업 유형을 위한 서버 풀 확장을 달성하기 위해서 작업의 동적인 방편 요구를 적용한다. 마찬가지로, 서버 분배기는 사용할 수 있는 서버의 불균형 공급을 갖는 것으로 간주된 작업유형에 따른 서버 풀의 단축을 달성할 수 있다. 서버 분배기 및 서버 분배방법에 의해서, 서비스 동작은 운영규칙에 따라 작업을 진척시키고 공평하게 작업을 분해하는 운영규칙의 달성에 적합한 서버를 이용할 수 있다. 이것은 실현된 결과에 기초하여 서버를 재분배하기 위한 수동 개입을 제거하는데 도움을 줄 수 있다. 전술한 바와 같이, 수동 파견은 흔히 너무 늦어 재분배된 서버가 목표를 달성하는데 충분히 빨리 이들의 선호된 작업으로 복귀하지 않았을 때 더 많은 문제에 이르게 된다. 서버 분배기 및 서버 분배방법은 작업 처리 설비 및 서버 조직에 대한 무모한 제약없이 작업유형을 위한 백업 서버를 자동으로 공유하여 이용할 수 있게 한다. 서버 분배기 및 서버 분배방법은 서비스 목표를 충족시킴으로써 고객 서비스를 개선할 수 있으며 수동개입에 의해 달성되는 서버 재분배 필요성을 감소시킴으로써 감독에 대한 짐을 줄일 수 있다.

서버 분배기는 서비스 수준 목표에 주체인 작업유형에 관련된 서비스 요구를 완화시키는데 특히 도움이 된다. 전술한 바와 같이, 서비스 수준 목표는 기업의 운영규칙에 관계될 수 있다. 이들 작업유형은 여기에서 서비스 클래스라고 부르는 서비스 요구 클래스를 갖는다. 서비스 클래스는 단순히 목표에 적용하는 트랜잭션 세트를 정하는 서비스 수준 목표의 부분이다. 원하는 서비스 수준 및 서비스를 양호 및 불량으로 분류하는 기준은 서비스 클래스의 부분이 아니다. 예를 들면, 운영이 프리미엄 프로스펙트로부터 전자메일 조회의 80%를 6시간 내에 처리하는 목표를 갖는다고 가정한다. "프리미엄 프로스펙트로부터 전자메일 조회"는 목표에 속하는 서비스 클래스이다. 서버가 6시간 내에 조회를 처리하는지 여부는 서비스 클래스 내의 일원에 영향을 미치지 않고, 80% 목표값은 서비스 클래스의 정의의 부분이 아니다.

작업유형과 서비스 클래스간 관계는 빈번히 복잡하여 확실히 사소한 것은 아니다. 서비스 클래스는 반드시 상호간에 배타적일 필요도 없고 공동으로 포함될 필요도 없다. 서비스 클래스는 서버 함수에 관계되지 않은 속성에 의해 정의될 수도 있고, 혹은 서버 함수를 정의하는 동일한 속성에 의해 정의될 수도 있다. 부가적으로, 한 목표 이상이 서비스 클래스에 속할 수 있다. 서버 선호도에 관하여, 본 발명의 실시예에 따라 작업유형은 빈번히 서비스 클래스보다는 더 많이 관계된 특성이다.

작업유형의 동적 선호값은 작업 분배기에서 서버 배분기에 입력되는 제어로서 함수하는 벡터를 구성한다. 동적 선호값은 작업에 서버 배분에 대한 결정을 좌우할 수 있다. 작업 분배기는 작업유형의 선호값의 크기에 따라 작업유형에 서버 분배를 한다. 작업유형에 관련된 동적 선호값을 이용하는 머신의 동작없이, 작업 분배기 혹은 ACD는 여전히 함수할 수 있으나, 서버 풀을 자동으로 확장 혹은 단축시키는 능력이 없어지게 된다. 서부 배분기를 사용하는 머신에 의해서, 작업 분배기는 여러 가지 서비스 클래스에 관련된 목표를 달성할 수 있게 하는 동적 서버 풀을 누리고 있다.

예를 들면, 작업 분배기가 통상적으로 제 1 서비스 클래스의 트랜잭션에 배분될 수 있는 것으로 선정된 모든 서버를 소진하였을 수 있을 상황에서, 서버 배분기가 있을 때, 제 1 서비스 클래스에 관련된 서비스 조직의 목표가 일단 우선도 면에서 다른 목표에 비해 충분히 위태롭게 되면, 작업 분배기는 배분될 수 있는 것으로 선정된 부가적인 서버를 가질 수도 있다. 본 발명의 실시예에 따라, 서버 배분기는 활용할 수 있는 서버의 풀을 명료하게 확장해도, 작업 항목을 서버에 분배하는 제어는 결국 작업 분배기의 것으로 된다. 본 발명의 또 다른 실시예에 따라, 서버 배분기는 각 서버에 대한 작업항목을 선택하는 과정에서 활용할 수 있는 서버의 풀을 은연중에 확장한다. 동적 선호값은 최소한으로 작업 분배에 영향을 미칠 뿐이다. 예를 들면, 작업유형에 따른 높은 동적 선호값은 전적으로 부적합한 작업유형에 대해 서버를 풀에 두지는 않을 것이다. 어떠한 목표도 위태롭게 되어 있지 않을 때, 서버 배분기는 모든 서비스 클래스에 대해 제로의 동적 선호값을 할당할 수 있다. 이 경우, 작업 분배기는 동적 상태에 어떠한 보상도 없다면 서버에 필적한다.

목표를 위태롭게 한 작업유형에 대해 서버풀의 확장이 서버 배분기 및 이의 서버 배분방법에 있어서 전형적인 모드일 수 있으나, 자원이 서비스 수준 목표를 초과하는 것에 제한을 둘 필요가 있다. 모든 목표가 충족되고 있는 상황에서, 그리고 일부 서비스 수준이 이들의 목표를 다른 목표보다 훨씬 초과하고 있는 상황에서, 운영규칙은 본 발명의 실시예에 따라 우수한 서비스를 고르게 분배시키기 위해서 서버를 재배분할 수 있다. 예를 들면, 동적 선호도는 양이 아닌 음으로 될 수 있어, 이에 의해 복합 선호도가 서버의 임계치보다 낮아지게 한다. 복합 선호도가 임계치 이하로 될 때, 서버는 서비스 수준 목표가 충족되지 않게 되는 시간까지 이 작업유형에 대해 서버 풀로부터 일시적으로 제거될 것이다. 이에 따라,

모든 서비스 목표가 충족되고 있는 경우, 시스템은 작업유형에 대해 큰 일관성을 제공하기 위해서 임의의 한 작업유형에 대해 자원을 확장하고 자원을 감축함으로써 작업간에 균형을 이룰 수 있다.

도 1은 본 발명의 실시예에 따라 서버 배분기(102)가 수행할 수 있는 표본 범용 컴퓨터 시스템(100)의 블록도이다. 컴퓨터 시스템(100)은 본 발명의 실시예에 따라 작업 처리 설비 내에 포함될 수 있다. 본 발명의 또 다른 실시예에서, 작업 처리 설비는 그 자체가 컴퓨터 시스템(100) 내에 포함될 수도 있다. 컴퓨터 시스템(100)은 중앙처리장치(CPU)(110), 입력/출력장치(120), 및 컴퓨터 메모리(메모리)(130)을 포함한다. 입력/출력장치 중에서 하드 디스크 드라이브와 같은 저장장치(121) 및 컴퓨터가 읽을 수 있는 매체 드라이브(122)가 있고, 이것은 CD-ROM과 같은 컴퓨터가 읽을 수 있는 매체 상에 제공되는 서버 배분기(102)를 포함하는 소프트웨어 제품을 설치하는데 사용될 수 있다. 입력/출력장치(120)는 또한 서버 및 트랜잭션이 수신되는 트랜잭션 소스에의 접속부(123)를 포함한다. 예를 들면, 접속부(123)는 컴퓨터 시스템을 하나 이상의 컴퓨터 전화 인터페이스, WWW 서버, 혹은 WWW 브라우저를 접속하는데 사용될 수 있다. 이 접속부는 서버를 검사(examine)하고, 응용물 및 트랜잭션을 서버가 입수할 수 있게 하며, 서버에 의한 트랜잭션의 처리를 모니터하기 위해 서버 배분기(102)가 사용할 수 있다.

서버 배분기(102)는 전술한 바와 같이 구성된 컴퓨터 시스템 상에 구현되는 것이 바람직하나, 이 분야에 숙련된 자들은 다른 구성의 컴퓨터 시스템 상에 구현될 수도 있음을 알 것이다. 예를 들면, 서버 배분기(102)는 전술한 바와는 다른 구성요소를 갖는 컴퓨터 시스템에 구현될 수 있다. 서버 배분기(102)는 콜 센터(call center)의 것들과 같은 전용 컴퓨터 시스템에 구현될 수도 있다. 서버 배분기(102)는 예를 들면 전용 IC(ASIC)를 사용하여 전용 하드웨어로 소프트웨어없이 구현될 수도 있다. 콜 센터의 자동 호 분배기(ACD) 혹은 작업 흐름 자동 시스템에 대한 이의 기능성의 일부로서 서버 배분기(102)를 포함할 수도 있다.

서버 배분기(102)의 상세를 완전하게 기술하기 위해서, 특정한 예에 관련한 동작을 이하 논한다. 이 예는 서버 배분기(102)의 동작을 논하기 위한 손쉬운 원리를 제공하도록 한 것이다. 그러므로 서버 배분기(102)의 전형적인 적용의 모든 면을 반드시 대표하는 것은 아니다. 사실, 이 분야에 숙련된 자들은 서버 배분기(102)는 이 예에 묘사된 시나리오보다는 폭넓고 다양한 시나리오에 적용될 수 있음을 알 것이다.

도 2는 본 발명의 실시예에 따라, 작업 처리 설비에서 수행되는 작업의 유형을 나타내는 여러 가지 작업유형(202)에 대한 여러 서버의 선호도를 예시한 서버 선호도표(200)이다. 시스템 내의 각각의 서버는 서버 ID(201)를 갖는다. 각각의 서버는 이론적으로는 서비스 조직에 의해 식별되는 각각의 가능한 작업유형(202)에 관련된 작업을 수행할 수 있다.

각각의 서버는 특정한 작업유형(202)의 작업을 수행하는 선호값(203)를 갖는다. 선호값(203)의 크기는 관련된 작업유형(202)에 대한 친화도를 나타낸다. 선호값(203)의 크기는 관련된 작업유형(202)에 관하여 서버 함수의 측정을 나타낸다. 각각의 서버는 특정 작업유형(202)에 관련된 태스크를 수행하는 서버의 저항도를 나타내는 임계값을 갖는다. 임계값(204)은 특정 작업유형(202)에 관련된 태스크를 수행하는 것에 관하여 서버의 무능의 측정을 또한 나타낸다. 따라서, 서버 선호표(200)는 각각의 서버 ID(201) 및 각각의 작업유형(202)에 대응하는 엔트리(205-209)를 갖는다. 예를 들면, 엔트리(205)는 "프로스펙트부터의 호"라는 작업유형에 관하여 서버 ID(100)를 갖는 서버의 선호도를 나타낸다. 특정한 작업유형에 대해서, 서버 ID(100)는 서비스 클래스 "8"에 대한 선호값과 서비스 클래스 "9."에 대한 임계값을 갖는다. 통상, 서버는 선호값이 임계값 미만인 것에 대해서는 작용하지 않는다. 따라서, 서버 ID(100)는 "프로스펙트로부터의 호"에 관련된 태스크를 수행하지 않을 것이다. 한편, 엔트리(205)에 나타낸 바와 같이, 서버 ID(100)는 통상 "호 클레임"라는 작업유형에 따른 선호값(203)이 "4"이고 임계값(204)이 "3"이므로 이 작업유형에 관련된 태스크를 수행할 것이다.

서버가 기본적으로 부적합한 작업유형은 선호값이 없거나 "0"의 선호값을 갖는다. 예를 들면, 서버 ID(101)는 엔트리(209)에 보인 바와 같은 "우선 고객으로부터의 호"에 관련된 작업을 수행하는데는 적합하지 않다. 따라서, 작업 분배기는 우선 고객으로부터의 호에 관련된 작업을 수행하도록 서버 ID를 배분하지 않을 것이다.

도 3은 본 발명의 실시예에 따라, 복합 선호값에 결합된 작업 처리 설비에서 수행되는 작업유형에 따른 서버 선호도 및 임계값을 예시한 표(300)를 보인 도면이다. 복합 선호값은 특정한 작업유형에 따른 개개의 선호도만이 아니라 작업유형에 관한 작업 처리 설비의 선호도이다. 도 2 및 도 3를 비교하면, 표(300)의 열(301-304)의 내용은 표(200)에서 열(201-204)의 내용에 각각 대응함을 알 수 있다. 표(300)는 부가적인 열로서, 작업유형(302)를 포함하는 작업을 수행하는 각각의 서버에 대한 작업 처리 시스템의 전체 선호도를 나타내는 복합 선호값 열(305)를 포함한다. 예를 들면, 서버 ID(100)의 "프로스펙트로부터의 호"에 대한 복합 선호값은 "10"이고, "호 클레임"에 대한 서버 ID(100)의 복합 선호값은 "8"이다.

서버 배분기(102)는 작업유형에 따른 동적 선호를 포함할 수 있는 복합 선호값을 계산하기 위해서 다양한 함수를 이용할 수 있다. 작업유형에 따른 동적 선호값을 생성하기 위한 적어도 한 대표적인 과정은 여기 참고로 포함시킨, 모자익스사에 양도된, 1998년 9월 8일에 출원된 미국특허출원 제09/149,877호, "Dynamically Assigning Priorities for the Allocation of Server Resources to Competing Classes of Works Based Upon Achievement of Service Level Goals"에 개시되어 있다.

도 4는 본 발명에 따라, 서버 배분기(102) 혹은 이의 대응하는 방법을 사용하여 복합 선호값의 생성을 예시한 흐름도이다.

서버 배분기(102)는 선택된 작업항목의 작업유형에 따른 서버의 선호값을 찾는다(단계 401). 본 발명의 일부 실시예에서, 서버 배분기(102)는 선택된 작업항목에 대한 작업유형을 판정한다. 본 발명의 다른 실시예에서, 선택된 작업항목에 대한 작업유형의 판정은 서버 배분기(102)가 간단히 작업항목에 대해 판정된 작업유형을 찾을 때, 복합 선호값의 생성 전에 수행될 수 있다. 서버 배분기(102)는 도 2에 도시

한 서버 선호표(200)와 같은 서버 선호표를 조사할 수 있다. 사용할 수 있는 서버가 도 2의 서버 선호표(200)에 나타낸 서버 ID(100)이고 선택된 작업유형이 "프로스펙트로부터의 호"라고 가정한다. 따라서, 서버 배분기(102)는 "8"의 선호값을 갖는 "프로스펙트로부터의 호"에 대한 서버 ID(100)의 선호값을 확인할 것이다.

서버 배분기(102)는 작업항목에 대한 동적 선호값을 찾는다(단계 402). 전술한 바와 같이, 동적 선호값은 작업유형의 속성일 수도 있을 것이다. 작업유형에 따른 동적 선호값을 판정함에 있어 많은 함수가 이용될 수 있다. 예를 들면, 동적 선호값은 달성된 서비스 수준, 큐 상태, 원하는 서비스 수준, 알람 상태, 응답 지연, 호 방기율, 혹은 운영자 개입에 기초할 수 있다. 동적 선호값이 원하는 서비스 수준의 달성/비달성 및 작업유형에 따른 우선도의 조합에 기초한다고 가정한다. 따라서, 원하는 서비스 수준이 현재 달성되고 있지 않은 높은 우선도의 작업유형은 서비스 수준이 달성되고 있는 작업유형 혹은 보다 낮은 우선도의 다른 작업유형보다 높은 동적 우선값을 가질 것이다. "프로스펙트로부터의 호"에 대한 동적 우선값이 "4"라고 가정한다.

서버 배분기(102)는 작업 처리 설비 내에서 현재 선택된 복합 선호값 함수를 찾는다(단계 403). 전술한 바와 같이, 서버 배분기(102)는 이를 테면 복합 선호값 함수와 같은, 여러 가지 파라미터를 사용자가 선택할 수 있게 한다. 때때로, 설비관리는 예를 들면, 하나의 복합 선호값 함수가 다른 함수보다 선호될 수 있는 것을 결정할 수도 있다.

복합 선호값 함수는 작업항목에 대한 동적 선호값으로(단계 402) 작업항목의 작업유형에 따른 서버의 선호값을 더하는 것을 포함할 수도 있다(단계 401). 또 다른 복합 선호값 함수는 작업항목에 대한 동적 선호값으로(단계 402) 작업항목의 작업유형에 따른 서버의 선호값을 곱하는 것을 포함할 수 있다(단계 401). 또 다른 복합 선호값 함수는 또 다른 파라미터에 따라 가중치가 부여된 작업항목에 대한 동적 선호값으로(단계 402) 또 다른 파라미터에 따라 가중치가 부여된 작업항목의 작업유형에 따른 서버의 선호값을 더하는 것을 포함할 수 있다(단계 401). 복합 선호값 함수의 다른 실시예에서는 다양한 스케일링 율을 사용할 수도 있다.

선택된 복합 선호값 함수에 관계없이, 서버 배분기(102)는 복합 선호값을 계산하기 위해서 복합 선호값 함수를 이용한다(단계 404). 현재 선택된 복합 선호값 함수가 동적 선호값으로 서버의 선호값을 더하는 것을 요구한다고 가정한다. "프로스펙트로부터의 호"에 대한 동적 선호값("4")으로 "프로스펙트로부터의 호"에 대한 서버 ID(100)의 선호값("8")을 더하면 "12"가 된다.

전술한 바와 같이, 작업 처리 설비는 서버의 선호값이 동일한 작업유형에 따른 서버 임계값 미만일 때는 서버에 작업을 제공하지 않는다. "프로스펙트로부터의 호"에 대한 서버 ID(100)의 선호값("8")은 도 2에 도시한 바와 같이, 서버 ID(100)의 대응하는 임계값("9")보다 작다. 작업 처리 설비는 "프로스펙트로부터의 호"에 관련된 작업이 할당될 수 있는 서버의 풀 내에 서버 ID(100)이 있는 것으로 간주하지 않을 것이다. 그러나, 위에 보인바와 같이, 서버 ID(100)의 복합 선호값은 12이고, 이것은 서버 ID(100)의 임계값인 9를 초과한다. 즉, 서버 ID(100)가 "프로스펙트로부터의 호"를 처리할 작업 처리 설비의 필요성은 이러한 태스크 수행에 서버 ID(100)의 저항도를 능가한다. 따라서, 작업 처리 설비는 프로스펙트로부터의 호를 수신할 수도 있는 서버의 풀 내에 서버 ID(100)이 있는 것으로 간주할 수 있다. 물론, "프로스펙트로부터의 호" 작업유형에 대해 복합 선호값을 재계산하여 서버 ID(100)의 임계값 미만의 복합 선호값으로 될 때, 서버 ID(100)는 "프로스펙트로부터의 호" 서버 풀로부터 제거될 수 있다. 더욱이, 복합 선호값(바람직한 실시예에 따라)은 서버 풀에 부적합한 서버의 할당으로는 결코 되지 않을 것이다. 서버가 기본적으로 부적합한("0"의 선호값을 갖는) 작업유형은 "0"의 복합 선호값을 수신할 수 있다.

복합 선호값 함수는 본 발명의 실시예에 따라 서비스 클래스 선호값 및 작업유형에 따른 서버의 선호값에 대한 수 많은 함수 중 임의의 것일 수 있다. 복합 선호 함수는 여러 동작모드 및 정규 서비스 목표에 무관계한 다른 동작상태를 조건으로 하여 동적으로 선택될 수 있다. 이러한 모드는 서버집단의 효율, 공평성 고려, 일 증진 요구, 혹은 특히 예외적인 상태를 포함하는 2차 목표에 관계될 수도 있다. 동작상태에 대한 간단한 예는 하루 중 시간이며 또 다른 예는 업무가 다른 업무규칙 하에 운영되는 주중 일이다.

도 5a 및 도 5b는 본 발명의 실시예에 따라, 서버 급파 과정을 예시한 흐름도이다. 서버 배분기(102)는 검사간격을 설정한다(단계 501). 검사간격은 임의의 회수의 시간 기간 혹은 조건을 포함할 수 있다. 적어도 한 실시예에서, 검사간격은 작업처리 설비 내의 다른 함수성에 의해 설정될 수 있다.

서버 배분기(102)는 종료상태가 일어났는지 여부를 판정한다(단계 503). 종료상태는 작업 처리 설비의 정지를 포함하여, 많은 인자에 의해 일어날 수 있다. 종료상태가 일어났다면(단계 503), 서버 배분기(102)는 이의 동작을 종료한다. 아니면, 서버 배분기(102)는 종료상태가 일어날 때까지 어떤 규칙적인 간격으로 동작을 계속한다.

종료상태가 일어나지 않았으면(단계 503), 서버 배분기(102)는 검사간격이 종료되었는지 여부를 판정한다(단계 505). 검사간격이 종료되지 않았으면(단계 505), 서버 배분기(102)는 대기한 후(단계507) 종료상태가 일어났는지 (단계 503) 및/또는 검사간격이 종료되었는지 여부(단계505)를 재판정한다. 검사간격의 길이는 사용자가 선택할 수 있으며, 검사간격은 본 발명의 적어도 한 실시예에 따라 근본적으로 연속하게 되도록 자주 할 수 있다.

검사간격이 종료되었다면(단계505), 서버 배분기(102)는 모든 작업유형이 검사되었는지 여부를 판정한다(단계509). 모든 작업유형이 검사되었다면(단계 509), 서버 배분기(102)는 복귀하여 종료상태가 일어났는지 여부를 판정한 후(단계503) 그렇지 않으면 또 다른 검사간격이 종료되었는지 판정하도록 진행한다(단계 505).

물론, 숙련된 자는 서버 배분기(102)가 통상적으로는 항시 모든 작업유형 및 모든 서버를 분석하지 않는다라는 것을 알 것이다. 실제로, "검사"는 서버를 사용할 수 있게 되었을 때 시작할 수 있다. 검사는

해당 서버만에 대해서 그리고 이미 대기하고 있던 작업항목에 의해 나타낸 작업유형에 대해서만 관계할 것이다. 더욱이, 검사는 작업항목이 작업처리 설비 내에 도착할 때 시작할 수도 있다. 검사간격은 새로이 도착한 작업항목에 대해서 그리고 이미 작업을 기다리고 있는 서버에 관해서 관계할 것이다. 선호값의 증가 및 임계값의 감소는 검사를 촉발할 수도 있을 것이다. 서버는 서버가 현재 중단시킬 수 없는 작업항목에 의해 사용중에 있을 때는 새로운 서버 풀에 포함될 필요는 없다. 물론, 어떠한 작업도 현재 대기하고 있지 않는 작업유형에 대해 서버 풀 내의 서버 일원임을 판정할 필요성도 거의 일어나지 않는다. 숙련된 자가 아는 바와 같이, 이들 상태는 본 발명의 실시예에 따라, 서버 배분기(102)에 관련된 처리 요구를 감소시킬 것이다.

모든 작업유형이 검사되지 않았다면(단계 509), 서버 배분기(102)는 검사될 다음 작업유형을 선택한다(단계 511). 서버 배분기(102)는 모든 서버가 검사되었는지 여부를 판정한다(단계 513). 서버 배분기(102)가 모든 서버를 검사하였다면(단계513), 서버 배분기(102)는 모든 작업유형이 검사되었는지 여부를 판정하도록 되돌아간다(단계 509).

모든 서버가 검사되지 않았다면(단계 513), 서버 배분기(102)는 다음 서버를 선택한다(단계 515). 서버 배분기(102)는 선택된 서버의 선호값이 서버의 임계값 미만인지 여부를 판정한다(단계 517). 선택된 서버의 선호값이 서버의 임계값 미만이 아니면(단계 517), 서버 배분기(102)는 선택된 작업유형에 대해 서버를 사용할 수 있게 한다(단계 523). 선택된 작업유형에 대해 서버를 사용할 수 있게 함에 따라, 작업처리 시스템 내의 작업 할당 설비에 의해 이용되는 서버 풀을 수정이 수반된다. 어떤 경우엔 작업유형에 대해 서버 풀에 서버가 미리 할당되어 있을 수 있고, 부가적인 처리가 필요하지 않게 될 것이다. 물론, 전술한 바와 같이, 모든 서버 및 모든 작업항목이 항시 검사될 필요는 없는 것이다.

선택된 서버의 선호값이 선택된 서버유형에 대해 서버의 임계값 미만이면(단계 517), 서버 분배기(102)는 작업유형 및 서버에 관계된 복합 선호값을 갱신한다(단계 519). 도 4는 복합 선호값을 갱신하는 바람직한 과정을 제공한다. 작업유형 및 서버에 관련된 복합 선호값을 갱신하는 것은 서버 배분기(102)가 어느 복합 선호값함수를 이용할 것인가를 판정하는 것을 판정한다. 전술한 바와 같이, 사용자는 본 발명의 실시예에 따라, 서버 배분기(102)에 의해 사용되는 복합 선호값 함수를 선택할 수도 있다.

서버 배분기(102)는 복합 선호값이 서버의 임계값 이상인지 여부를 판정하기 위해서 선택된 복합 선호값 함수를 적용한다(단계 521). 복합 선호값이 서버의 임계값 이상이면, 서버 배분기(102)는 이 작업유형에 대해 서버를 사용할 수 있게 한다(단계 523). 서버 배분기(102)는 작업유형에 대헤 서버 풀에 서버 할당을 달성하는데 필요한 임의의 단계를 실제로 착수함이 없이, 간단히, 작업유형에 대해 서버 풀 내에 서버 포함될 것을 작업 처리 설비에게 지시할 수 있다. 복합 선호값이 서버의 임계값보다 크지 않거나 같다면, 서버 배분기(102)는 이 작업유형에 대해 서버를 선택하지 않고 다른 서버를 검사하는 것으로 돌아간다(단계 513).

서버 급파과정을 위한 도 5a 및 도 5B에 도시한 흐름도는 집중 및 분배 서버 급파 과정에 똑같이 적용할 수 있다. 따라서, 서버 급파 과정은 집중 실시예 및 분배 실시예에 존재할 수 있다. 예를 들면, 동적 선호의 평가는 하나 이상의 프로세서로 동작하는 하나의 처리에 의해 결정될 수 있는 한편 복합 선호는 각각이 서버 전용인 프로세서들에 의해 결정될 수 있을 것이며, 이 경우 이들 전용 프로세서들은 어떤 통신수단을 통해 한 세트의 동적 선호를 얻는다. 서버 급파 처리에 대한 실제 처리는 특정한 구현에 대해 실제적인 임의의 구성으로 실현될 수 있다.

도 6a-6d는 본 발명의 실시예에 따라, 서버 배분기(102)에 기인하여 시간에 걸쳐, 여러 서버 풀(603-605)에 서버의 배분의 변화를 도시한 것이다. 예시 목적으로, 서버 풀(603)에 관련된 작업유형(601)은 서버 풀(605)에 관련된 작업유형보다 높은 평균 복합 선호값을 갖는 서버 풀(604)에 관련된 작업유형보다 높은 평균 복합 선호값을 갖는 것으로 가정한다.

작업 처리 설비에서 비충족 필요성 때문에, 서버 배분기(102)는 목표가 충족되고 있지 않은 작업유형에 대해 보다 많은 서버를 사용할 수 있게 한다. 서버 배분의 변경은 도 6a에 도시한 차트에서 도 6b에 도시한 차트로 서버 풀 내의 다수의 가용 서버(602)를 변경한다. 도 6a에서, "프로스펙트로부터의 호"의 작업유형은 42개의 가용 서버의 서버 풀(603)을 갖는다. "프로스펙트로부터의 호"에 관련하여 달성되지 않은 목표 때문에, 서버 배분기(102)는 도 6b에 도시한 바와 같이, 서버 풀(603) 내의 서버의 수를 48개의 서버로 더욱 더 갱신한다. 물론, 서버는 하나의 서버 풀 이상에 할당될 수도 있다.

마찬가지로, "호 클레임" 서버 풀(604) 내의 서버 수는 도 6a 및 도 6b에 나타낸 시간 간격 도안 감소하였다. "호 클레임" 서버 풀(604) 내의 서버 수의 감소는 서버 선호값 변경 혹은 동적 선호값의 변경에 기인할 수 있다. "우선 고객으로부터 호" 작업유형을 위한 서버 풀(605)은 도 6a 및 도 6b에 나타낸 시간 간격 동안 일정하다. 물론, 도 6b에 도시한 "우선고객으로부터의 호" 서버 풀(605) 내의 서버들은 도 6a의 "우선고객으로부터의 호" 서버 풀(605)에 대해 도시한 서버들과 동일할 필요는 없다.

작업 처리 설비에서 비충족 필요성 때문에, 서버 배분기(102)는 도 6b와 도 6c를 비교하여 나타낸 바와 같이, "프로스펙트로부터의 호"에 대해 보다 많은 서버를 사용할 수 있게 하게 되는 동적 선호값을 계속 이용한다. 마찬가지로, "호 클레임" 서버 풀(604) 및 "우선고객으로부터의 호" 서버 풀(605)은 도 6b 및 도 6c에 나타낸 시간 간격 동안 일괄적으로 감소한다.

서버 배분기(102)는 본 발명의 실시예에 따라, 서버 풀에 대해 최대 및/또는 최소 수의 서버를 유지하도록 설계될 수 있다. 예를 들면, "우선고객으로부터의 호" 서버 풀(605)은 최소 서버 수 "10"을 가질 수 있다. 따라서, 서버 배분기(102)는 서버풀(605)이 예비 서버로도 알려진 최소 서버 수에 일단 도달하면 "우선고객으로부터의 호"로부터 서버 풀(603-604)로 서버의 재배분을 그만둔다.

서버 배분기(102)는 가용 서버의 불균형 공급을 갖는 것으로 간주된 작업유형에 따른 서버 풀을 감축할 수 있다. 목표를 위태롭게 한 작업유형에 따른 서버 풀의 확장이 서버 배분기(102) 및 이의 서버 배분 방법에 있어서 전형적인 모드이지만, 서버 배분기(102)는 본 발명의 실시예에 따라 서버 수준 목표를 초과하는 것으로부터 서버 자원을 한정하는 것이라 할 수도 있다. 모든 목표가 충족되고 있는 상황에서,

그리고 일부 서비스 수준이 다른 서비스 수준이 그들 각각의 목표를 초과하고 있는 것보다 훨씬 목표를 초과하고 있는 상황에서, 서버 배분기의 동작규칙은 서버 자원을 보다 균일하게 분배하도록 서버를 재배분한다.

예를 들면, 도 6c에 도시한 "프로스펙트로부터의 호" 서버 풀(603)이 관련된 서비스 목표를 달성하는데 필요한 것보다 많은 서버를 포함한다고 가정한다. "프로스펙트로부터의 호" 서버 풀(603)에 관련된 목표가 원하는 수준 이상으로 달성되고 있기 때문에, 서버의 수를 감소시켜 다른 서비스 목표가 달성될 수 있도록 한다. 따라서, "프로스펙트로부터의 호" 서버 풀(603) 내의 서버 수는 도 6c에서 도 6d로 감소한다. 서버 풀(604-605)에 할당된 서버 수는 대응하여 도 6c에서 도 6d로 증가한다.

서버 풀(603) 내의 서버 수를 감소시키기 위해서, 서버 배분기(102)는 양이 아닌 음으로 되어, 이에 의해 복합 선호값이 하나 이상의 서버의 임계치보다 낮아지게 된다. 복합 선호값이 서버의 임계치 이하로 될 때(이 경우 서버의 선호값도 임계값 보다 낮아짐), 서버는 서비스 수준 목표가 충족되지 않게 될 때까지 이 작업유형에 대해 서버 풀로부터 일시적으로 제거될 것이다. 이에 따라, 모든 서비스 목표가 충족되고 있는 경우, 서버 배분기(102)는 어떤 세트의 작업유형에 대해 큰 일관성을 제공하기 위해서 임의의 한 작업유형에 대해 자원(예를 들면, 서버)을 확장하고 자원(예를 들면 서버)을 감축함으로써 작업유형간에 균형을 이룰 수 있다. 복합 선호값을 구동하는 함수(들)이 복합 선호값을 서버의 임계값 밑으로 되게 할 때 동적 선호값은 제로에 있게 하거나, 증가하거나 고려되지 않을 수 있음을 당업자에겐자명할 것이다.

도 7은 본 발명의 대안 실시예에 따라, 동적 선호값 열(701)만이 아니라 작업 처리 설비에서 수행되는 작업유형에 따른 서버 선호 및 임계값 및 복합 선호값을 포함하는 표(700)을 도시한 도면이다. 도 3의 표(300과 표(700)을 비교하면, 표(300)의 열(301-305)는 표(700)의 열(301-305)에 대응함을 알 수 있다. 표(700)은 여러 가지 작업유형의 수행을 위한 서비스 조직의 현재 선호값을 나타내는 동적 선호값 열(701)을 포함한다.

전술한 바와 같이, 복합 선호값 열(305) 내에 기록되는 복합 선호값을 계산에 다수의 방법이 사용될 수 있다. 표(700)에서, 복합 선호값 열(305) 내의 복합 선호값은 작업유형(701)에 대한 동적 선호값으로 작업유형 열(303)에 대한 서버 선호을 함께 더함으로써 결정되었다. 예를 들면, 서버 ID(100)의 "프로스펙트로부터의 호"에 대한 복합 선호값은 이 작업유형에 따른 동적 선호값 "2"로 이 작업유형에 따른 서버 ID(100)의 선호값 "8"을 더하여 계산된 "10"이다. 모든 다른 면에서, 본 발명은 전술한 방법으로 동작한다.

작업유형에 따른 임계값은 특정한 서버에 대한 관계를 유지하는 것으로 앞에서 기술하였다. 예를 들면, 도 2에 도시한 바와 같이, "호 클레임" 작업유형에 따른 서버 ID(100)의 임계값은 "3"이고 "호 클레임" 작업유형에 따른 서버 ID(100)의 임계값은 "7"이다. 마찬가지로, "호 클레임" 작업유형에 따른 서버 ID(100)의 임계값은 "프로스펙트로부터의 호" 작업유형에 따른 서버 ID(100)의 임계값과는 다르다.

본 발명의 대안 실시예에서, 작업유형에 따른 임계값은 특정한 서버에 관계없이 설정될 수 있다. 도 8은 본 발명의 실시예에 따라 작업 처리 설비에서 수행되는 작업유형을 나타내는 여러 가지 작업유형(202)에 대한 여러 서버의 선호값을 도시한 서버 선호표(800)이다. 서버 선호표(800)은 서버 선호표(800)가 도 2에 도시한 작업유형 열(204)에 서버 임계를 포함하지 않는 것을 제외하곤 도 2에 도시한 서버 선호표(200)와 유사하다.

서버 선호표(800)은 본 발명의 대안 실시예에 따라 여러 가지 작업유형에 따른 적합한 임계값을 포함하는 임계값 표(806)에 관련할 수 있다. 이 실시예에서, 서버 배분기(102)는 특정 서버에 작업을 할당할 것이지 여부를 판정할 때 임계값 표(806) 내의 적합한 값을 참조할 것이다. 예를 들면, 임계값 표(806)은 일 실시예에 따라 모든 작업유형에 대해 단지 하나의 임계값을 가질 수 있다. 임계값 표(806)는 대신에 각각의 서버마다 모든 작업유형에 따른 하나의 임계값을 가질 수도 있다. 대안으로, 임계값 및 서버 선호값은 다른 자동화된 제어 시스템 및 감독 운영자로부터의 수동입력을 동적 서비스 및 이용특성에 통합하는 시스템에 동적으로 기초할 수 있다. 더욱이, 임계값은 작업유형마다 다르게 결정될 수도 있다. 물론, 서버 선호표(800) 및 임계값 표(806)는 공통의 메모리 요소에 저장될 수 있고 본 발명의 실시예에 따라 상호 관계된 데이터 구조를 가질 수 있다.

물론, 서버 선호의 결정을 위해 작업유형에 기여하는 요소는 본 발명의 실시예에 따라 동적 선호값을 결정하는 속성과는 전적으로 무관하다. 예를 들면, 도 8에 도시한 바와 같이, 서버 ID(100)에 대해 서버 선호를 얻는데 필요한 작업유형은 콜러(caller)의 위치 혹은 서버 ID(100)의 위치를 기술하는 지역 속성과 같은 다른 속성과는 무관하다. 그러나, 선호의 동적성분(서버와는 무관한)은 지역속성과 같은 다른 속성에 의해 부분적으로 결정될 수도 있다. 더욱이, 도 8에 도시한 작업유형은 예를 들면 선호의 동적 성분을 결정할 때 어느 정도 반드시 이용될 필요는 없다.

도 9a는 본 발명의 대안 실시예에 따라 임계값 결정기(911) 및 복합 선호 평가기(910)를 포함하는 또 다른 범용 컴퓨터 시스템(100)에 대한 블록도이다. 임계값 결정기(911)는 도 8에 도시한 임계값 표(806)에 대한 임계값을 계산할 수 있다. 복합 선호 평가기(910)은 현재 선택된 사용자 선택가능 복합 선호값 함수가 본 발명의 실시예에 따라 작업 처리 설비에 관하여 기업의 목적을 달성하는 가장 적합한 복합 선호값 함수인지 여부를 판정할 수 있다. 모든 다른 면에서, 컴퓨터 시스템(100)은 도 1에 도시한 컴퓨터 시스템(100)에 대해 앞에서 기술한 방식으로 동작한다. 물론, 복합 선호 평가기(910) 및 임계값 결정기는 서버 배분기(102)와 같은, 다른 기능성 내에 포함될 수 있다.

복합 선호 평가기(910)은 관련된 목표를 초과하여 서비스 수준 수행의 동등한 분배, 서버의 동등한 이용, 서비스 수준 달성에서 변동의 제동, 서비스 무위에서 변동의 제동, 및 모드에 특정한 행동과 같은, 다른 원하는 동작특성을 달성하기 위해서, 복합 선호값 함수의 선택을 동적으로 조잘하는 최우선의 제어 법칙을 이용할 수 있다. 복합 선호 평가기(910)은 이 실시예에서 앞에서 기술한 복합 선호값 함수에 더하여 또 다른 층의 제어를 구성하더록 고찰될 수 있다.

복합 선호 평가기(910)는 여러 동작모드 및 정규 서비스 목표에 무관계한 다른 동작상태를 조건으로 하여 동적으로 복합 선호함수를 선택할 수 있다. 이러한 모드는 서버집단의 효율, 공평성 고려, 일 증진 요구, 혹은 특히 예외적인 상태를 포함하는 2차 목표에 관계될 수도 있다.

도 9b는 본 발명의 실시예에 따라 복합 선호 평가기(910)에 의해 수행되는 d일부 동작을 도시한 흐름도이다.

복합 선호 평가기(910)는 기업 및 이의 관련된 작업 처리 상태를 평가하여 작업 처리 설비에 대한 기업의 목적의 전체 달성 수준을 결정한다(단계 901). 복합 선호 평가기(910)는 현재 선택된 사용자 선택가능 복합 선호값 함수가 기업 및 이의 관련된 작업 처리 상태의 평가에 비추어 적합한 복합 선호값 함수인지 여부를 판정한다(단계 902).

복합 선호 평가기(910)가 현재 선택된 사용자 선택가능 복합 선호값 함수가 기업 및 이의 관련된 작업 처리 상태의 평가에 비추어 적합한 복합 선호값 함수인 것으로 판정하면(단계 902), 복합 선호 평가기(910)는 대기상태에 들어간 후(단계 903) 한번 다시 기업 평가를 수행한다(단계 901).

복합 선호 평가기(910)가 현재 선택된 사용자 선택가능 복합 선호값 함수가 기업 및 이의 관련된 작업 처리 상태의 평가에 비추어 적합한 복합 선호값 함수가 아닌 것으로 판정하면(단계 902), 복합 선호 평가기(910)는 본 기업 평가에 기초하여 다른 사용자 선택가능 복합 선호값 함수를 선택한다(단계 904). 복합 선호 평가기(910)는 대기상태에 들어간 후(단계 903) 한번 다시 기업 평가를 수행한다(단계 901).

대기상태(단계 903)는 많은 수단 및 다양한 이유로 종료할 수 있을 것이다. 예를 들면, 대기상태는 소정의 시간 기간일 수 있을 것이다. 부가적으로, 복합 선호값 함수 혹은 다른 이벤트 이용을 조건으로 혹은 이에 동기하여 촉발될 수도 있을 것이다. 전술한 바와 같이, 많은 유형의 속성의 무리는 본 발명의 실시예에 따라 작업유형을 결정할 수 있다. 작업유형 필요한 태스크, 제품, 통신매체, 지리, 정부관할, 인구통계, 언어, 숙련도, 서비스 유형, 고객 프로파일, 판매처리단, 및 규제 클래스의 조합에 의해 결정될 수 있다.

도 10은 본 발명의 실시예에 따라, 복합 선호값에 결합된 작업 처리 설비에서 수행되는 작업유형에 따른 서버 선호 및 임계를 예시한 표(1000)이다. 표(1000)는 작업유형 열(302) 내의 작업유형이 단일 속성 이상을 포함할 수 있다는 것을 제외하곤 도 3에 도시한 표(300)와 유사하다. 예를 들면, 행(1001)은 "서부지역에서 프로스펙트로부터의 호"인 복수의 속성 작업유형을 포함한다. 표(300)에서 대응하는 행(806)은 "프로스펙트로부터의 호"의 단일 속성 작업유형을 포함한다. 작업유형에서 복수의 속성을 할당하는 것은 전체 시스템이 복잡도 및 작업유형 엔트리 총 수를 증가시킨다. 표(1000)가 서버 ID(100)에 대한 속성만을 목록하고 있으나, 다른 서버에 대해 유사한 정보를 넣을 수 있다.

표(1000)를 사용하는 다음의 예를 고찰한다. 셀룰라 전화회사에의 조회는 서버 선호의 결정을 위해 작업유형으로서 질의를 분류하는 것 외에 상이한 처리 제약을 만족시키기 위해서 지리적으로 분류될 수도 있다. 작업 처리 설비에 넣어지는 작업항목은 제 1 속성(예를 들면, "프로스펙트 호" 혹은 "호 클레임") 범주로 분류될 수 있을 뿐만 아니라 제 2 속성(예를 들면, 지역 설명)에 따라 분류될 수도 있다. 작업항목은 시장분할 전략을 용이하게 하기 위해서 여러 가지 시장분할에 따라 분류될 수도 있을 것이다. 나이, 수입, 교육, 직업, 및 가족 규모와 같은 인구통계 요소를, 다양한 제품라인에 대한 다양한 목표 시장에서 고객들을 확인하는데 사용될 수도 있을 것이다. 통상, 작업 및 고객속성의 조합은 매우 복잡한 세트의 작업유형을 야기한다. 본 발명은 소규모의 정적인 풀에 서버 집단을 무모하게 분할하는 비능률없이 이러한 복잡성을 지원한다.

표(1000)는 작업유형이 빈번하게 수행되지 않아도 일부 작업유형이 높은 우선도를 가질 필요성을 나타내고 있다. 예를 들면, 작업 처리 설비는 행(1005)에 도시한 "공익사업 집행 호"와 같은 여러 규제제약하에서 운영할 수도 있다. "공익사업 집행 호" 작업유형은 기업에 있어서는 거의 없는 확률의 호이지만, 그럼에도 불구하고 어떤 공무로 집행된 최소/최대 수준의 서비스를 받아야 하는 호를 포함할 수 있다. 따라서, 표(1000)는 "공익사업 집행 호"가 공적으로 집행되는 서비스 수준에 가깝게 도달하였음을(혹은 그 이하로 떨어졌음을) 나타내는, 높은 복합 선호값을 행(1005)에 예시하고 있다. 일단 서버 분배기(102)가 "공익사업 집행 호"에 대해 요구된 서비스 수준을 다시 달성하였으면, 복합 기준 값은 이의 높은 값으로 떨어진다. 규제제약 외에도, 표(1000)는 그 경우가 아니면 기업에게는 제한적으로 중요한, 이를 테면 여러 전기통신 서비스 노조와의 합의에 의해 집행되는 서비스 수준의 달성과 같은, 다른 중요한, 높은 값의 목표를 포함할 수 있다.

전술한 바와 같이, 서버 급파 과정은 집중 및 분배될 수 있다. 도 11은 본 발명의 실시예에 따라, 분배 서버 배분 시스템(1101)을 도시한 것이다. 분배 서버 배분 시스템(1101)은 컴퓨터 시스템(100a-100f)를 포함한다. 컴퓨터 시스템(100a-100f)는 서버 분배기(102a-102f)를 포함하며, 다른 것은 도 1에 도시한 컴퓨터 시스템(100)과 유사하다.

서버 분배기(102a-102f)는 서버 분배기(102)에 대해 기술된 방식으로 동작할 수 있다. 서버 분배기(102a-102f)는 단일 작업 처리 설비 내에 배치될 수 있으며 서로간에 작업항목(혹은 개개의 작업항목)을 공유하는 복수의 작업 처리 설비에 걸쳐 펼쳐져 있을 수 있다.

서버 분배기(102a-102f) 중 하나가 이의 서버 분배 태스크를 달성할 수 없으면, 서버 분배기는 망(1100)을 사용하여 다른 서버 분배기와 연락한다. 서버 분배기(102a-10f)는 망(1100)을 사용하여 서로간에 상태정보를 통신할 수도 있다.

일부 서버 배분기(102a-102f)는 본 발명의 실시예에 따라, 다른 서버 배분기(102a-102f)가 복합 선호값 함수를 결정하는 한편 동적 선호값 결정을 수행한다. 이 실시예에서, 동적 선호값 결정을 제공하는 특별화된 서버 배분기(들)과의 통신을 제외하곤 서버 배분기들간에 통신으로 없으며 이들은 독립적으로 행동한다.

발명의 효과

이 발명을 바람직한 실시예에 관련하여 도시하고 기술하였으나, 본 발명의 범위에서 벗어남이 없이 형태에서의 여러 가지 변경 혹은 수정이 행해질 수 있음을 당업자는 알 것이다. 예를 들면, 서버 배분기에 의해 이용되는 선호값은 모든 유형의 자원을 할당하는데 사용될 수 있다. 더욱이, 인접하는 정수가 아닌 선호 수준은 비수치적인 서비스 수준일 수도 있어 작업 처리 설비에 의해 할당될 수도 있다. 예를 들면, 특정한 작업유형에 관한 양호 혹은 불량 서비스를 만족하는 기준은 트랜잭션 속성의 부울린 함수나 문자 숫자식의 표시자, 예를 들면 A1〉C3와 같은 어떤 형태를 사용함으로써 정수나 소수의 사용을 필할 수 있다. 더욱이, 당업자는 임의의 소정의 작업 처리 설비에서 서버의 선호값, 임계값, 및 동적 선호값에 대해 사용하기 위해 값의 결정은 공지된 실험적인 과정을 쉽게 이용할 수 있음을 알 것이다.

여기서는 서버를 서버 풀에 속하는 것으로 논하였으나, 당업자는 특정한 서버 풀의 일원으로서 실제로는 식별할 필요가 전혀 없다는 것을 알 것이다. 즉, 서버는 한 유형 혹은 다른 유형의 작업만을 받을 수 있으므로, 풀이 물리적으로 존재하지 않아도 그 작업유형에 대해 서버 풀에 속하는 것으로서 간주될 수 있다.

본 발명은 작업유형 및 서비스 클래스에 관하여 작업항목이 기술되는 시스템에 또한 적용할 수 있다. 서비스 클래스는 작업항목의 특정한 면을 기술할 수 있으며 작업유형은 서버 관점에서 작업항목을 기술한다. 따라서, 본 발명의 실시예에서, 서비스 클래스는 동적 선호값 및 궁극적으로는 복합 선호값에 영향을 미칠 수 있다.

여기에서 다룬 예는 콜 센터에서의 콜에 관계하나, 본 발명은 사람 및/또는 기계식인 서버에 의한 콜 센터에서의 콜의 처리로 제한되지 않는다. 본 발명은 임의의 형태의 작업유형 및 임의의 형태의 서비스 클래스의 서비스에 적용할 수 있다. 예를 들면, 본 발명은 전자 메일 메시지 및 비디오 스트리밍 태스크의 처리에 똑같이 적용할 수 있다.

본 발명에 대해 이들 및 다른 변경을 상기 상세한 설명에 비추어 행할 수 있다. 일반적으로, 다음의 청구범위에서, 사용된 용어는 본 명세서에 개시된 특정 실시예로 본 발명을 한정하는 뜻으로 파악해서는 안되며, 청구범위는 모든 계층적인 개선 및 본 발명에 따라 동작하는 시스템을 포함하는 것으로 파악해야 한다. 따라서, 본 발명은 개시된 바에 의해 제한되는 것이 아니라 다음 청구범위에 의해 그 범위가 결정되어야 하는 것이다.

(57) 청구의 범위

청구항 1

작업 처리 설비에서 작업유형에 따라, 작업유형의 작업을 처리하는 서버들의 지시를 보유하도록 구성된 서버 풀에 서버를 배분할 것인지 판정하는 방법에 있어서,

작업유형에 따른 서버의 양의 표시자를 나타내는 서버의 제 1 값이, 그 작업유형에 따른 서버의 음의 표시자를 나타내는 서버의 제 2 값보다 작은지 판정하는 단계와,

상기 작업유형에 따른 상기 서버의 제 1 값이 상기 작업유형에 따른 상기 서버의 제 2 값보다 작으면 상기 작업유형에 따른 복합 선호값을 결정하는 단계와,

상기 작업유형에 따라 결정된 복합 선호값이 작업유형에 따른 상기 서버의 제 2 값보다 큰지 판정하는 단계와,

상기 작업유형에 따른 복합 선호값이 상기 작업유형에 따른 서버의 제 2 값보다 크면 상기 서버가 상기 작업유형에 대해 서버풀에 배분에 적합하다는 지시를 작업 처리 설비에 보내는 단계를 포함하는 작업유형에 따른 서버 풀에 서버의 배분방법.

청구항 2

제 1항에 있어서, 상기 복합 선호값을 판정하는 것은,

상기 작업유형의 작업을 서버들이 수행하게 하는 상기 작업 처리 설비의 필요성의 측정을 반영하는, 작업유형에 따른 동적 선호값을 결정하는 단계와,

복합 선호값을 결정하도록 구성되는, 사용자가 선택할 수 있는 복합 선호값 함수를 액세스하는 단계와,

상기 작업유형에 따라 상기 서버의 결정된 선호값과 상기 작업유형에 따라 상기 결정된 동적 선호값을, 상기 액세스된 사용자가 선택할 수 있으며 작업유형에 따른 상기 복합 선호값을 생성하는 복합 선호값 함수에 제공하는 단계를 포함하는 작업유형에 따른 서버 풀에 서버의 배분방법.

청구항 3

제 1 항에 있어서, 상기 작업유형에 따른 복합 선호값이 작업유형에 따른 상기 서버의 제 2 값보다 크지 않다면, 상기 서버의 제 1 값이 상기 서버의 제 2 값보다 작은 경우 작업유형의 서버 풀에 서버를 배분하지 않는 단계를 더 포함하는 작업유형에 따른 서버 풀에 서버의 배분방법.

청구항 4

제 1항에 있어서, 상기 작업유형에 따른 복합 선호값을 재결정하는 단계와,

상기 작업유형에 따라 상기 재결정된 복합 선호값이 상기 서버의 제 2 값보다 큰지 판정하는 단계와,

상기 작업유형에 따라 상기 재결정된 복합 선호값이 상기 서버의 제 2 값보다 작다면 상기 서버를 작업유형에 따른 서버 풀로부터 제거한다는 지시를 상기 작업 처리 설비에 보내는 단계를 더 포함하는 작업유형 대해 서버 풀에 서버의 배분방법.

청구항 5

제 1 항에 있어서, 상기 작업 처리 설비는 복수의 작업유형에 따라 복수의 서버 풀을 가지며,

복합 선호값은 작업유형에 따른 상기 서버의 제 1 값이 그 작업유형에 따른 상기 서버의 제 2 값보다 작다면 상기 복수의 작업유형의 각각의 작업유형에 따라 결정되는 작업유형에 따른 서버 풀에 서버의 배분방법.

청구항 6

작업 처리 설비에서 작업유형에 따라서, 작업유형의 작업을 처리하는 서버들의 지시를 보유하도록 구성된 서버 풀에 서버를 배분할 것인지 판정하는 시스템에 있어서,

작업유형에 따른 서버의 양의 표시자를 나타내는 서버의 제 1 값이, 그 작업유형에 따른 서버의 음의 표시자를 나타내는 서버의 제 2 값보다 작은지 판정하도록 구성된 제 1 비교기와,

상기 작업유형에 따른 상기 서버의 제 1 값이 상기 작업유형에 따른 상기 서버의 제 2 값보다 작으면 상기 작업유형에 따른 복합 선호값을 결정하도록 구성된 제 2 비교기와,

상기 작업유형에 따라 결정된 복합 선호값이 작업유형에 따른 상기 서버의 제 2 값보다 큰지 판정하도록 구성된 제3 비교기와,

상기 작업유형에 따른 복합 선호값이 상기 작업유형에 따른 서버의 제 2 값보다 크면 상기 서버가 상기 작업유형에 대해 서버 풀에 배분에 적합하다는 지시를 작업 처리 설비에 보내도록 구성된 결과 표시기를 포함하는 작업유형에 따른 서버 풀에 서버의 배분 시스템.

청구항 7

제 6 항에 있어서, 상기 제 2 비교기는,

상기 작업유형의 작업을 서버들이 수행하게 하는 상기 작업 처리 설비의 필요성의 측정을 반영하는, 작업유형에 따른 동적 선호값을 결정하도록 구성된 결정기와,

상기 작업유형에 따라 상기 서버의 결정된 선호값과 상기 작업유형에 따라 상기 결정된 동적 선호값을, 상기 액세스된 사용자가 선택할 수 있으며 작업유형에 따른 상기 복합 선호값을 생성하도록 구성된 복합 선호값 함수에 제공하도록 구성된 프로세서를 포함하는 작업유형에 따른 서버 풀에 서버의 배분 시스템.

청구항 8

제 6 항에 있어서, 상기 결과 표시기는 상기 작업유형에 따른 복합 선호값이 작업유형에 따른 상기 서버의 임계값보다 크지 않다면, 상기 서버가 작업유형의 서버 풀에 배분되지 않음을 지시하도록 구성된 작업유형 대해 서버 풀에 서버의 배분 시스템.

청구항 9

제 6 항에 있어서,

상기 제 2 비교기는 상기 작업유형에 따른 복합 선호값을 재결정하도록 구성되며, 상기 시스템은,

상기 작업유형에 대해 상기 재결정된 복합 선호값이 상기 서버의 임계값보다 큰지 판정하도록 구성된 제4 비교기와,

상기 작업유형에 대해 상기 재결정된 복합 선호값이 상기 서버의 제 2 값보다 작다면 상기 서버를 작업유형에 따른 서버 풀로부터 제거한다는 또 다른 지시를 보내도록 구성된 제 2 결과 표시기를 더 포함하는 작업유형에 따른 서버 풀에 서버의 배분 시스템.

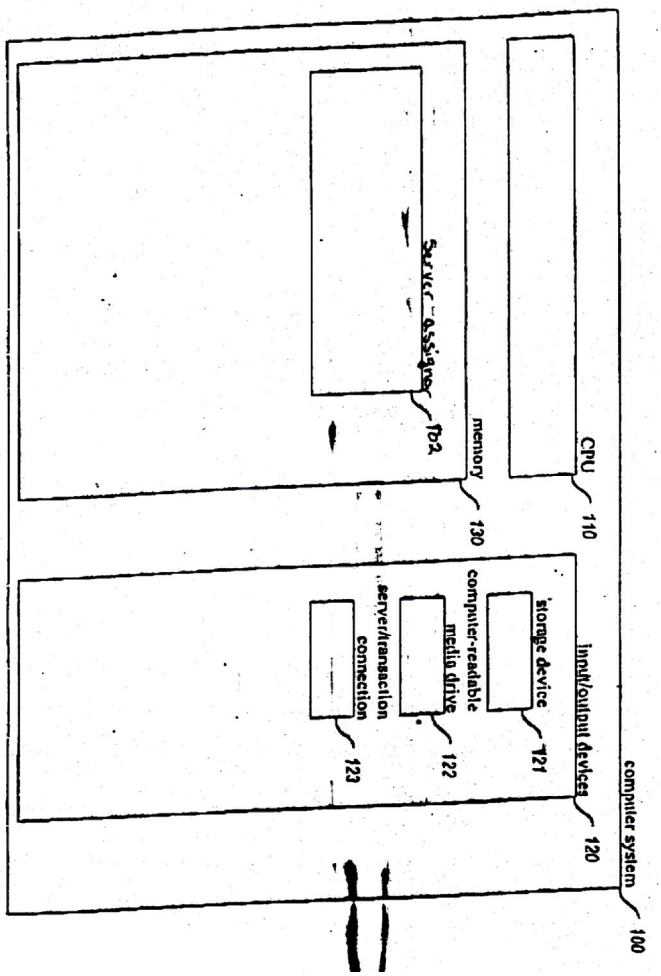
청구항 10

제 6 항에 있어서, 상기 작업 처리 설비는 복수의 작업유형에 따른 복수의 서버 풀을 가지며,

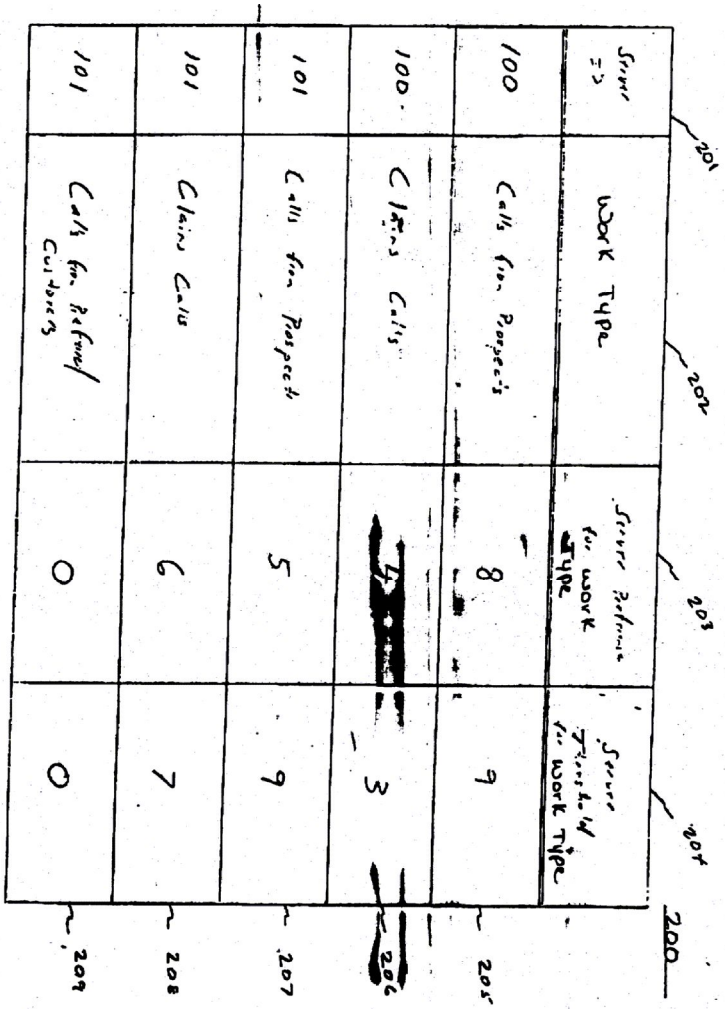
상기 제 2 비교기는 작업유형에 따른 상기 서버의 선호값이 그 작업유형에 따른 서버의 임계값보다 작다면, 복수의 작업 유형들의 각 작업유형에 따른 복합 선호 값을 결정하도록 구성되는 작업유형에 따른 서버 풀에 서버의 배분 시스템.

도면

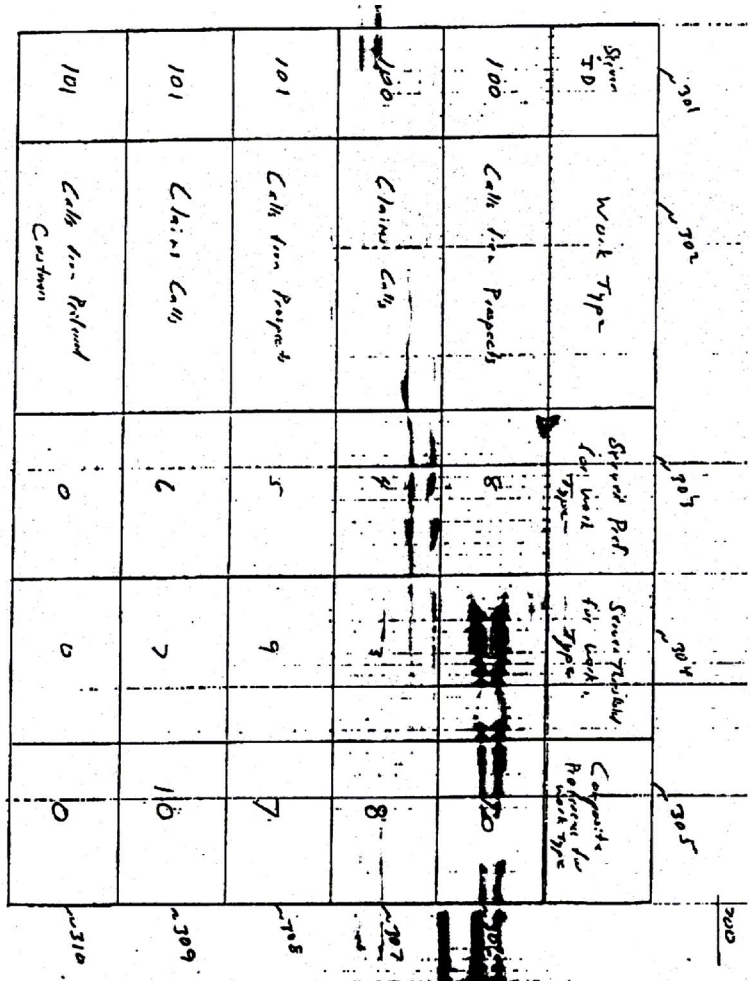
도면1


computer system
100
input/output devices
120
storage device
121
computer-readable
media drive
122
server/transaction
connection
123
CPU
110
memory
130

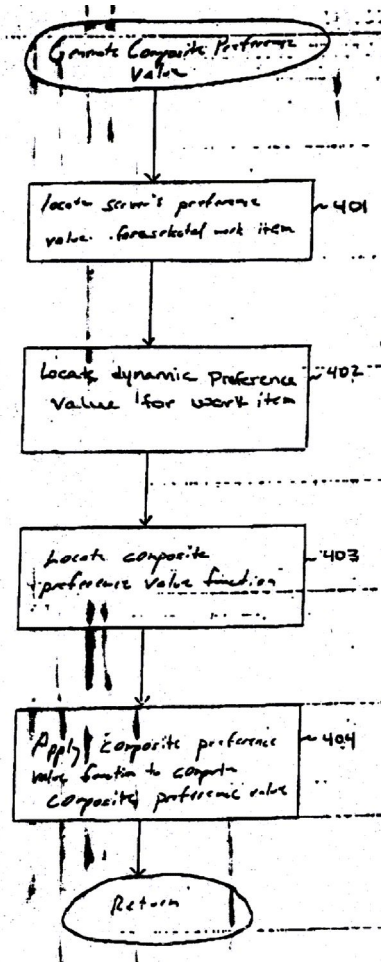
도면2


200
201
202
203
204
Server ID
Work Type
Server Performance for work Type
Server Throughput for work Type
100
Calls from Prospects
8
9
205
100
Closing Calls
4
3
206
101
Calls from Prospects
5
9
207
101
Claims Calls
6
7
208
101
Calls for Refund/ Customers
0
0
209

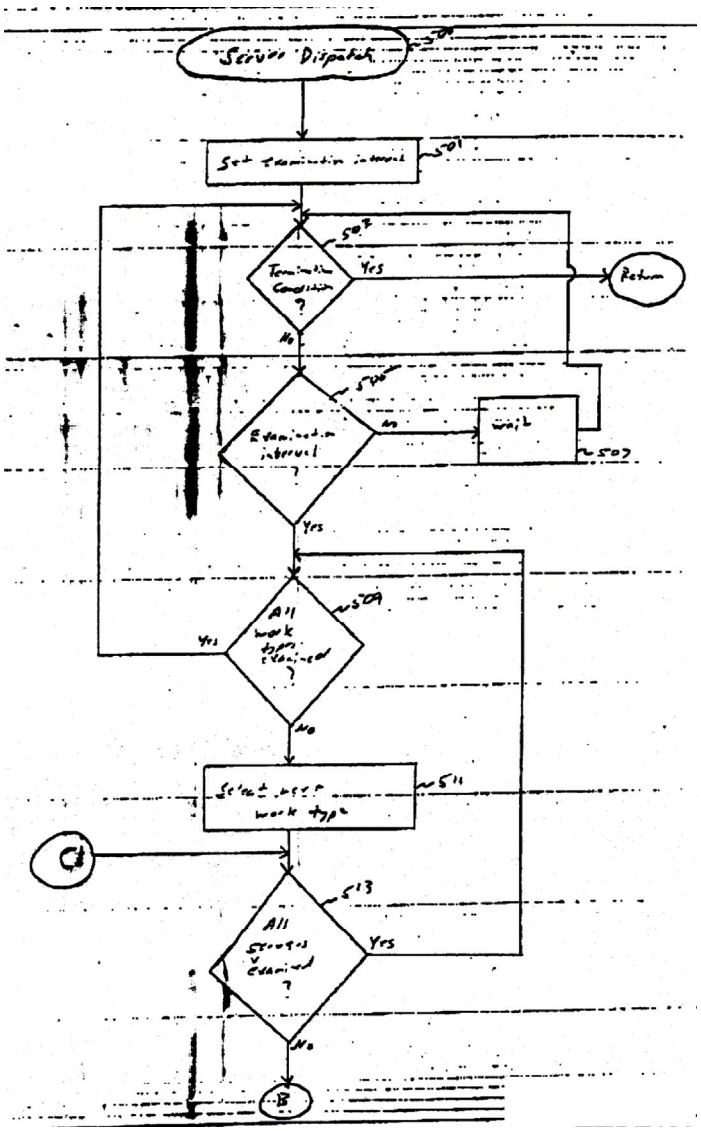
도면3


Server ID
Work Type
Server Prof. for Work Type
Server Reliability for Work Type
Composite Preference for Work Type
100
Calls from Prospects
8
10
101
Calls from Prospects
5
9
7
101
Claims Calls
6
7
10
101
Calls from Preferred Customers
0
0
0
301
302
303
304
305
306
307
308
309
310
200

도면4


Generate Composite Preference Value
Locate screen's preference value for selected work item
401
Locate dynamic Preference Value for work item
402
Locate composite preference value function
403
Apply composite preference value function to compute composite preference value
404
Return

도면5a


Server Dispatch
Set examination interval
Termination Condition ?
Yes
Return
No
Examination interval ?
No
wait
Yes
All work types examined ?
Yes
No
Select next work type
All servers examined ?
Yes
No
B

도면5b

