



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2017-0002415
(43) 공개일자 2017년01월06일

- (51) 국제특허분류(Int. Cl.)
G06F 9/50 (2006.01) G06F 17/30 (2006.01)
G06F 21/62 (2013.01) H04L 29/08 (2006.01)
- (52) CPC특허분류
G06F 9/5066 (2013.01)
G06F 17/30592 (2013.01)
- (21) 출원번호 10-2016-7030931
- (22) 출원일자(국제) 2015년05월05일
심사청구일자 없음
- (85) 번역문제출일자 2016년11월04일
- (86) 국제출원번호 PCT/US2015/029132
- (87) 국제공개번호 WO 2015/171539
국제공개일자 2015년11월12일
- (30) 우선권주장
14/271,548 2014년05월07일 미국(US)

- (71) 출원인
마이크로소프트 테크놀로지 라이선싱, 엘엘씨
미국 워싱턴주 (우편번호 : 98052) 레드몬드 원
마이크로소프트 웨이
- (72) 발명자
야할롬 자르
미국 워싱턴주 98052-6399 레드몬드 원 마이크로
소프트 웨이 엘씨에이 - 인터넷셔널 페이턴츠
(8/1172) 마이크로소프트 테크놀로지 라이선싱,
엘엘씨
바 나다브
미국 워싱턴주 98052-6399 레드몬드 원 마이크로
소프트 웨이 엘씨에이 - 인터넷셔널 페이턴츠
(8/1172) 마이크로소프트 테크놀로지 라이선싱,
엘엘씨
- (74) 대리인
제일특허법인

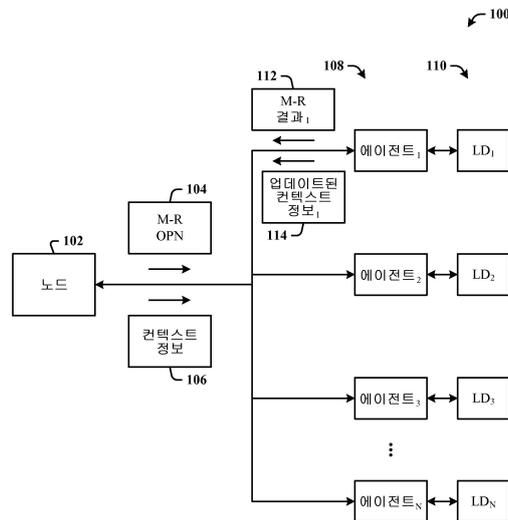
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 순회 맵-리듀스 아키텍처

(57) 요약

데이터 저장소와 장치 사이를 스킵할 수 있는 완전 컨텍스트 정보를 갖고 "순회하는" 맵-리듀스 동작에 관한 것이다. "순회" 측면은 맵-리듀스 동작 요청이 특정 에이전트로 통신되어 에이전트의 로컬 데이터에 대해 동작할 수 있음을 의미한다. 순회 맵-리듀스 동작이 프라이버시를 보호하고 사용자 개인 데이터의 유출을 피한다. 상기 순회 맵-리듀스 동작은 장기간 동안 실행될 수 있고 항상 연결되어 있는 것은 아님(오프라인 상태인) 데이터 저장소에서 작업할 수 있다. 아키텍처가 컨텍스트 프리 제어기 및 데이터 저장소(장치) 내에 위치하는 온-프레미스(장치 상에 있는) 에이전트의 세트를 이용한다.

대표도 - 도1



(52) CPC특허분류

G06F 21/6218 (2013.01)

H04L 67/10 (2013.01)

명세서

청구범위

청구항 1

시스템으로서,

맵-리듀스 동작(a map-reduce operation) 및 컨텍스트 정보(context information)를 에이전트로 전송함으로써 맵-리듀스 세션을 실시하도록 구성된 노드 - 상기 에이전트 각각은 연관된 로컬 데이터에 대해 맵-리듀스 동작을 실행시켜, 맵-리듀스 결과 및 대응하는 컨텍스트 정보의 업데이트를 획득하고, 상기 노드는 상기 맵-리듀스 세션을 기초로 상기 에이전트로부터 맵-리듀스 결과를 수신하며, 상기 에이전트로부터 업데이트된 컨텍스트 정보를 수신함 - 와,

상기 노드와 연관된 메모리에서 컴퓨터 실행가능 명령어를 실행하도록 구성된 적어도 하나의 마이크로프로세서를 포함하는

시스템.

청구항 2

제1항에 있어서,

상기 노드는 지정된 에이전트들에게 병렬로 상기 맵-리듀스 동작을 전송함으로써, 상기 맵-리듀스 세션을 병렬로 완료하고, 상기 맵-리듀스 동작을 완료한 에이전트의 대응하는 맵-리듀스 결과 및 업데이트된 컨텍스트 정보를 수신하는

시스템.

청구항 3

제1항에 있어서,

상기 노드는 상기 맵-리듀스 세션 내 다른 에이전트에 액세스하기 전에 상기 맵-리듀스 동작을 에이전트로 전송하고 맵-리듀스 결과 및 컨텍스트 업데이트 정보를 수신함으로써, 상기 맵-리듀스 세션을 순차적으로 완료하는

시스템.

청구항 4

제1항에 있어서,

에이전트의 상기 맵-리듀스 결과는 특정 에이전트 및 사용자 장치로부터 획득된 때 식별 불가능한 데이터를 포함하는

시스템.

청구항 5

제1항에 있어서,

하나의 에이전트는 도달 가능한 에이전트로부터 획득된 새로운 진행 중인(on-going) 동작 컨텍스트 및 새로운 맵-리듀스 결과가 상기 노드로 전달되기 전에, 진행 중인 동작 컨텍스트 및 맵-리듀스 결과를 도달 가능한 다른

에이전트로 전달하는
시스템.

청구항 6

하나 이상의 에이전트의 로컬 데이터에 대해 맵-리듀스 동작을 수행하기 위해 맵-리듀스 동작 요청을 노드로부터 대응하는 하나 이상의 에이전트로 전송하는 단계와,

상기 맵-리듀스 동작 요청을 기초로 상기 하나 이상의 에이전트로부터 맵-리듀스 결과 및 업데이트된 컨텍스트 정보를 상기 노드에서 수신하는 단계와,

상기 노드로부터 상기 맵-리듀스 결과 및 업데이트된 컨텍스트 정보를 출력하는 단계를 포함하는
방법.

청구항 7

제6항에 있어서,

특정 에이전트가 상기 맵-리듀스 동작을 완료했음을 식별하기 위해 컨텍스트를 업데이트하는 단계를 더 포함하
는

방법.

청구항 8

제6항에 있어서,

이전에 오프라인 상태였던 온라인 상태 에이전트로 상기 맵-리듀스 동작 요청을 리디렉트(redirect)하는 단계를
더 포함하는

방법.

청구항 9

제6항에 있어서,

지정 에이전트들의 리스트의 지정 에이전트들을 통해 순차적으로 상기 맵-리듀스 동작 요청을 전송하는 단계를
더 포함하는

방법.

청구항 10

제6항에 있어서,

하나의 에이전트로부터의 맵-리듀스 결과 및 업데이트된 컨텍스트 정보를, 또 다른 에이전트에 의해, 상기 노드
에서 점진적으로 누적하는 단계를 더 포함하는

방법.

발명의 설명

배경 기술

[0001] 맵-리듀스 알고리즘(map-reduce algorithm)을 실행하기 위해서는, 모든 데이터가, 보통 온라인 데이터 클러스터(데이터 저장소)에서, 쉽게 이용 가능해야 하며, 데이터 수집 및 맵-리듀스 정의를 조정하는 메인 제어기가 필요하다. 이는 분석되는 데이터가 공유 위치에 존재해야 함을 의미하는데, 이는 데이터를 프라이버시 침해에 노출시킬 수 있다. 분산되어 있고 사적일 수 있으며 지속적으로 이용 가능하지 않은 또는 프라이버시 규정으로 인해 이용 가능하지 않은 데이터에 맵-리듀스 동작을 실행시키는 어떠한 방법도 존재하지 않는다. 덧붙여, 장치들(또는 데이터 저장소)에 걸쳐 장기간의 맵-리듀스 동작의 실행을 가능하게 하는 컨텍스트 프리 제어기(context free controller)도 존재하지 않는다.

발명의 내용

[0002] 개요

[0003] 이하에서 본 명세서에 기재된 일부 신규한 실시예의 기본적인 이해를 제공하기 위해 단순화된 요약이 제공된다. 이 개요는 포괄적인 개괄이 아니며, 이는 본 발명의 핵심/필수 요소를 식별하거나 이의 범위를 명확화하기 위한 것이 아니다. 차후 제공되는 상세한 설명의 전제로서 단순화된 형태로 일부 개념들을 제공하기 위한 것이다.

[0004] 개시된 아키텍처는 데이터 저장소와 장치 간에 스킵할 수 있는 풀 컨텍스트(full context)를 갖는 "순회하는" 맵-리듀스 동작이다. "순회" 측면은 맵-리듀스 동작 요청, 컨텍스트 및 결과가 특정 에이전트로 그리고 이 에이전트를 통해 통신되어 에이전트의 로컬 데이터를 조작할 수 있음을 의미한다. 상기 순회 맵-리듀스 동작은 프라이버시를 보호하고 사용자 개인 데이터의 노출을 피한다. 순회 맵-리듀스 동작은 장기간 동안 실행될 수 있고 항상 연결되어 있는 것은 아닌(오프라인) 데이터 저장소 상에서 동작할 수 있다. 아키텍처는 "컨텍스트 프리(context free)" 온라인 제어기 및 데이터 저장소(장치) 내에 있는 온-프레미스(장치 상에 있는) 에이전트의 세트를 이용한다. 제어기는 맵-리듀스 동작이 에이전트 별로 이동할 때 컨텍스트가 자동으로 변하기 때문에 컨텍스트 프리이다.

[0005] 일반적인 동작 기재에서, 맵-리듀스 동작(요청)이 일부 소비자 또는 그 밖의 다른 서비스 또는 프로그램으로부터 제어기로 제출된다. 상기 맵-리듀스 동작은 맵 및 리듀스 동작 정의를, 맵-리듀스 동작에 참여하는 에이전트를 가리키는 에이전트 속성의 세트와 함께 포함한다. 제어기는 에이전트 중 하나 이상과 통신하여 맵-리듀스 동작을 제출할 수 있다. 에이전트는 데이터 프라이버시를 보존하면서 로컬 데이터에 대해 맵-리듀스 동작을 실행시킨다. 맵-리듀스 동작의 현재 실행 중인 에이전트와 관련된 컨텍스트 정보는 업데이트되어 로컬 데이터에 대한 맵-리듀스 동작을 실행한 결과를 반영할 수 있고 컨텍스트 내 에이전트 엔트리가 업데이트되어 에이전트가 동작을 완료했음을 반영할 수 있다. 에이전트가 완료될 때, 업데이트된 맵-리듀스 동작 컨텍스트 및 결과를 제어기 또는 또 다른 에이전트로 전송한다. 그 후 제어기가 순회 맵-리듀스 동작을 새 에이전트로 타깃재설정하거나 에이전트가 순회 맵-리듀스 동작을 실행 및 전달하고 프로세스가 반복되어 맵-리듀스 세션을 완료할 수 있다.

[0006] 상기의 그리고 관련된 목적을 이루기 위해, 특정 예시적 양태가 본 명세서에 다음의 설명 및 도면을 참조하여 기재된다. 이들 양태는 본 명세서의 원리가 실시될 수 있고 모든 양태 및 균등물이 본 발명의 범위 내에 있는 다양한 방식을 나타내는 것이다. 그 밖의 다른 이점 및 신규한 특징이 도면을 참조하여 고려할 때 다음의 상세한 설명으로부터 자명해질 것이다.

도면의 간단한 설명

- [0007] 도 1은 개시된 아키텍처에 따르는 시스템을 도시한다.
- 도 2는 개시된 아키텍처에 따르는 맵-리듀스 시스템의 순차적 구현 흐름도이다.
- 도 3은 개시된 아키텍처에 따르는 맵-리듀스 시스템의 피어 구현 흐름도이다.
- 도 4는 개시된 아키텍처에 따르는 맵-리듀스 시스템의 병렬 구현 흐름도이다.
- 도 5는 개시된 아키텍처에 따르는 맵-리듀스 시스템의 위치-기반 피어-투-피어 구현 흐름을 도시한다.
- 도 6은 개시된 아키텍처에 따르는 맵-리듀스 시스템의 조합 구현 흐름도이다.
- 도 7은 맵-리듀스 동작 세션을 이루기 위해 제어기 및 에이전트에 의해 통신될 수 있는 데이터세트를 도시한다.
- 도 8은 개시된 아키텍처에 따르는 방법을 도시한다.

도 9는 개시된 아키텍처에 따르는 대안적 방법을 도시한다.

도 10은 순회 맵-리듀스 아키텍처를 실행하는 컴퓨팅 시스템의 블록도이다.

발명을 실시하기 위한 구체적인 내용

- [0008] 맵-리듀스 처리는 일반적으로, 총괄하여 클러스터라고 지칭되는 많은 컴퓨터(노드)를 이용해 거대한 데이터셋들 간에 병렬로 문제를 처리하기 위한 프레임워크인 것으로 이해된다. 맵-리듀스는 저장 자산 상의 또는 그 근처의 데이터를 처리함으로써 데이터의 로컬성(locality)을 이용하여, 데이터 전송 비용을 낮출 수 있다.
- [0009] 개시된 아키텍처는 동작에 대해 선택된 장치를 기초로 데이터 저장소와 장치 사이를 스킵(skip)할 수 있는 폴 컨텍스트를 갖는 "순회(traveling)" 맵-리듀스 동작이다. "순회" 양태는 맵-리듀스 동작 요청이 특정 에이전트로 통신되고 이를 통해 에이전트의 로컬 데이터를 조작할 수 있음을 의미한다. 순회 맵-리듀스 동작은 프라이버시를 보호하고 사용자의 사적인 데이터의 노출을 피한다. 순회 맵-리듀스 동작은 장기간 동안 실행될 수 있으며 항상 연결되어 있는 것은 아닌(오프라인) 데이터 저장소 상에서 동작할 수 있다. 아키텍처가 컨텍스트 프리 온라인 제어기(context free online controller) 및 데이터 저장소(장치) 내에 있는 온-프레미스(on-premises)(장치 상에 있는) 에이전트의 세트를 채용한다.
- [0010] 일반적인 동작을 기재하자면, 맵-리듀스 동작(요청)이 제어기로 제출되어 일부 소비자 또는 그 밖의 다른 서비스 또는 프로그램을 위한 결과를 획득할 수 있다. 맵-리듀스 동작은 에이전트가 맵-리듀스 동작에 참여함을 가리키는 에이전트 속성의 세트와 함께, 맵 및 리듀스 동작 정의를 포함한다. 제어기는 에이전트와 통신하여 맵-리듀스 동작을 제출할 수 있다. 에이전트는 데이터 프라이버시를 보존하면서 로컬 데이터에 대해 맵-리듀스 동작을 실행시킨다. 맵-리듀스 동작의 현재 동작 에이전트에 대한 컨텍스트 정보가 업데이트되어 로컬 데이터에 대해 맵-리듀스 동작을 실행시킨 결과를 반영할 수 있고, 컨텍스트 내 에이전트 엔트리가 업데이트되어 상기 에이전트가 동작을 완료했음을 반영할 수 있다. 에이전트가 완료할 때, 에이전트는 업데이트된 맵-리듀스 동작 컨텍스트 및 결과를 제어기로 전송한다. 그 후 제어기는 순회 맵-리듀스 동작의 타깃을 새 에이전트로 재설정하고, 프로세스가 반복된다. 그러나 본 명세서에 기재되는 바와 같이, 에이전트는 제어기를 우회하여 맵-리듀스 동작을 또 다른 에이전트에게 직접 전달할 수 있다.
- [0011] 다음의 예시가 본 발명의 아키텍처의 현실적 이익을 보여준다. 제1 예시에서, (사용자 위치 정보를 보호하면서 순회 맵-리듀스 동작을 이용해) 도시의 가장 유행하는 장소를 찾고자 한다. 컨텍스트 프리 제어기가 가령 온라인 클라우드 서비스로서 실행 중이고, 맵-리듀스 에이전트(가령, 맵-리듀스 에이전트로서 동작하는 장치 측 서비스)를 실행하는 모바일 폰의 세트가 존재한다고 가정할 수 있다. 맵-리듀스 동작이 장치상 저장된(on-device stored) 데이터 위치에 대해 실행되고 장치상 위치에 따라 도시 타일별 카운터의 리스트를 생성하도록 정의된다. 타일 카운터 리스트가 맵-리듀스 동작 컨텍스트(map-reduce operation context)에 저장된다. 따라서 타일 카운터 리스트가 (필요에 따라 제어기를 중개자(mediator)로서 이용해) 하나의 모바일 장치에서 또 다른 모바일 장치로 순회한다. 사용자 위치 데이터는 장치 자체를 결코 떠나지 않고, 장치들 간에 대강의 위치 정보만 공유된다.
- [0012] 제2 예시에서, 순회 맵-리듀스 동작을 이용해 사용자 데이터를 보호하면서 가장 유행하는 음악 그룹을 찾고자 한다. 컨텍스트 프리 제어기가 가령 온라인 클라우드 서비스로서 실행되고, 맵-리듀스 에이전트(가령, 맵-리듀스 에이전트로서 동작하는 장치 측 서비스)를 실행시키는 모바일 폰의 세트가 존재한다고 가정할 수 있다. 맵-리듀스 동작이 장치상 저장된 음반 데이터에 대해 실행되고 장치상 음반에 따라 음악 그룹별 카운터의 리스트를 생성하도록 정의된다. 카운터 리스트는 맵-리듀스 동작의 컨텍스트에 저장되고 따라서 컨텍스트 정보가 다음 모바일 폰 에이전트로 전송되기 때문에, 카운터 리스트는 (필요에 따라 제어기를 중개자로서 이용하여) 하나의 모바일 장치에서 또 다른 모바일 장치로 순회한다. 사용자 데이터는 결코 장치 자체를 떠나지 않고, 특정 사용자 또는 사용자 장치와 다시 관련지어지는 데이터 없이, 대강의 음악 그룹 정보만 장치들 간에 공유된다.
- [0013] 제3 예시에서, 샌 프란시스코("베이 에어리어(Bay Area)")의 십대 여성에 의해 이뤄진 평균 통화 횟수를 찾고자 한다. 컨텍스트 프리 제어기가 예를 들어 온라인 클라우드 서비스로서 실행 중이고, 맵-리듀스 에이전트(가령, 맵-리듀스 에이전트로서 동작하는 장치 측 서비스)를 실행시키는 모바일 폰의 세트가 존재한다고 가정할 수 있다. 맵-리듀스 동작이 특정 속성을 갖는 에이전트를 트리거하도록 정의된다. 이 경우, 속성은 장소("베이 에어리어"), 성별("여성") 및 연령("12-17세 그룹")이다.
- [0014] 맵-리듀스 동작이 장치상 통화 로그 기록에 대해 실행되고 24시간 동안의 평균 통화 횟수를 생성하도록 설정된다. 평균 통화 횟수가 맵-리듀스 동작 컨텍스트에 저장되며 다음 에이전트로 전달된다. 따라서 맵-리듀스 컨텍

스트 및 결과가 (필요에 따라 제어기를 중개자로서 이용해) 하나의 모바일 장치에서 또 다른 모바일 장치로 "순회"한다. 사용자 전화기의 통화 데이터는 결코 장치 자체를 떠나지 않고, 특정 사용자 또는 사용자 장치와 다시 관련지어지는 일 없이 대강의 통계 정보만 장치들 간에 공유된다.

- [0015] 이하에서 도면을 참조하는데, 전체에서 유사한 도면 부호가 유사한 요소를 지칭하도록 사용된다. 이하의 기재에서, 설명 목적으로, 복수의 특정 상세사항이 완전한 이해를 제공하기 위해 제공된다. 그러나 이들 특정 세부사항 없이 새로운 실시예가 실시될 수 있음이 자명할 수 있다. 또 다른 경우, 설명을 용이하게 하기 위해 잘 알려진 구조 및 장치가 블록도 형태로 도시된다. 본 발명의 사상 및 범위 내에 있는 모든 변형예, 균등예, 및 대안예를 포함한다.
- [0016] 도 1은 개시된 아키텍처에 따르는 시스템(100)을 도시한다. 시스템(100)은 맵-리듀스 동작(M-R OPN)(104) 및 컨텍스트 정보(106)를 에이전트(에이전트_{1-N})(108)에게 전송함으로써 맵-리듀스 세션을 실시하도록 구성된 노드(102)를 포함할 수 있다. 상기 에이전트(108)는 로컬 데이터(LD_{1-N})의 연관된 로컬 데이터(LD)에 맵-리듀스 동작(104)을 실행시켜 맵-리듀스 결과(가령, M-R 결과₁(112)) 및 대응하는 컨텍스트 정보(가령, 업데이트된 컨텍스트 정보₁(114))의 업데이트를 획득할 수 있다. 노드(102)는 맵-리듀스 세션(의 일부)을 기초로 에이전트(108)로부터 맵-리듀스 결과를 수신하고 에이전트(108)로부터 업데이트된 컨텍스트 정보를 수신한다.
- [0017] 로컬 데이터가 서로 다른 프로그램, 예컨대, 일정관리 프로그램, 음성 프로그램, 이미지 처리 프로그램, 텍스트 생성 및 편집 프로그램 등과 연관되어 생성 및 저장된 데이터를 포함할 수 있다. 따라서 로컬 데이터는 텍스트, 이미지, 오디오, 비디오 및 임의의 조합을 포함할 수 있다. 로컬 데이터는 로컬 장치 상의 하나 이상의 위치, 가령, 단일 하드 드라이브 또는 복수의 하드 드라이브, 외부 드라이브 등에 저장될 수 있다.
- [0018] 사용자 장치와 상이한 위치에 저장될 수 있는 데이터에 대해, 맵-리듀스 동작이 원격 데이터까지의 경로(가령, 하이퍼링크)를 "따라가", 원격 데이터까지, 또는 대안적으로 온-프레미스 데이터를 처리할 수 있는 것이 본 발명의 아키텍처의 범위 내에 있다. 예를 들어, 충분한 로컬 저장소가 없지만 데이터를 생성하는 장치에서, 데이터는 정규 데이터 동작의 일부로서 원격 데이터 저장소로 업로드될 수 있다. 따라서 원격 데이터 저장소가 맵-리듀스 에이전트를 호스팅하는 경우, 원격 데이터 저장소는 에이전트를 실행하여 원격 데이터 저장소에 "로컬"인 데이터를 조작하고 결과 및 업데이트된 컨텍스트 정보를 장치로, 그리고 장치로부터 노드(102)로 반환할 수 있다.
- [0019] 노드(102)는 맵-리듀스 동작(104)을 에이전트(108) 중 지정된 (가령, 전체, 하나 또는 일부의) 에이전트(가령, 에이전트₁, 에이전트₃ 등)에게 병렬로 전송함으로써 병렬 방식으로 맵-리듀스 세션을 실행(완료)할 수 있고 대응하는 맵-리듀스 결과(가령, M-R 결과₁(112)) 및 맵-리듀스 동작을 완료한 에이전트(가령, 지정된 에이전트)의 업데이트된 컨텍스트 정보(가령, 업데이트된 컨텍스트 정보₁(114))를 수신한다.
- [0020] 노드(102)는 또한 맵-리듀스 동작을 한 에이전트(가령, 에이전트₁)에게 전송하고 또 다른 에이전트(가령, 에이전트₂)를 액세스하기 전에 상기 에이전트(가령, 에이전트₁)의 컨텍스트 업데이트 정보를 수신함으로써, 맵-리듀스 세션을 순차적으로 실행할 수 있다.
- [0021] 노드(102)는 모든 지정된 에이전트(108)에 대해 맵-리듀스 세션을 핸들링(관리)하는 제어기 노드일 수 있다. 이를 대체하여 또는 이와 조합하여, 에이전트(가령, 에이전트₁)는 제어기로서 동작할 수 있고 타 지정 노드(가령, 에이전트₂, 에이전트₃, 등)에 대한 맵-리듀스 세션을 핸들링할 수 있다. 노드(102)는 온라인 클라우드 서비스로서 실행될 수 있다. 각각의 에이전트(108)는 장치측 서비스로서 동작하는 맵-리듀스 프로그램이다.
- [0022] 에이전트(가령, 에이전트₁)의 맵-리듀스 결과(가령, M-R 결과₁(112))는 특정 에이전트로부터 도출(derived) 및 획득되며 특정 사용자 및 사용자 장치와 관련된 것으로 식별될 수 없는 데이터를 포함한다. 따라서 임의의 특정 사용자 장치로부터의 결과의 일부로서 포함되는 데이터의 프라이버시가 맵-리듀스 에이전트 동작의 일부로 유지된다. 결과의 최소 임계치가 에이전트(108)로부터 수신될 때, 노드(102)는 지정된 에이전트(108) 중 하나, 일부, 또는 전부로부터 맵-리듀스 결과 및 업데이트된 컨텍스트 정보를 소비자(도시되지 않음)에게로 출력한다. 소비자는 예를 들어 또 다른 네트워크 서비스일 수 있다.
- [0023] 하나의 구현예에서, 하나의 에이전트는, 또 다른 도달 가능한 에이전트로부터 획득된 새로운 진행 중인 동작 컨텍스트 및 새로운 맵-리듀스 결과가 노드(102)에게 전달되기 전에, 진행 중인 동작 컨텍스트 정보 및 맵-리듀스

결과를 상기 도달 가능한 에이전트에게 전달한다.

- [0024] 이하에서 다양한 맵-리듀스 구현예가 제공된다. 예를 들어, 구현예의 비제한적 예를 들면, 순차적, 병렬, 순차적과 병렬 모두, 에이전트 피어-투-피어(peer-to-peer), 위치 기반 실행 등이 있다.
- [0025] 도 2는 개시된 아키텍처에 따르는 맵-리듀스 시스템(200)의 순차 구현예 흐름을 도시한다. 상기 시스템(200)은 "컨텍스트 프리" 온라인 제어기(202) 및 온-프레미스(장치 상에 있는) 에이전트의 세트를 포함할 수 있고, 이때, 에이전트가 데이터 저장소(장치) 내에 위치하거나 데이터 저장소와 연관되어 위치한다.
- [0026] 먼저, ①에서, "순회" (분산 실행) 맵-리듀스 동작이 (노드(203)와 유사한) 제어기(202)로 제출된다. 맵-리듀스 동작은 맵 및 리듀스 동작 정의를, 이 맵-리듀스 세션에 참여하도록 지정된 에이전트의 에이전트 속성의 세트와 함께 포함한다.
- [0027] ②에서, 제어기(202)는 제1 "온-프레미스" 에이전트(204)(에이전트가 장치 상에 있거나 로컬 데이터를 갖고 있을 때 "온-프레미스"라고 함)와 통신하여 (정의를 포함하는) 맵-리듀스 동작을 지정 에이전트에게 제출할 수 있다.
- [0028] 제1 에이전트(204)가 리스트 상의 첫 번째 지정된 에이전트였지만, 제1 에이전트(204)가 제어기(202)에게 오프라인인 경우가 있을 수 있다. 이러한 경우, 제어기(202)는 지정 에이전트 리스트 상의 다음 에이전트를 접촉하게 될 수 있다. 이 프로세스는 온라인 에이전트가 발견될 때까지 계속될 수 있다. 덧붙여, 온라인 에이전트가 완전히 처리되면, 제어기(202) 및/또는 마지막 온라인 상태 에이전트가 맵-리듀스 동작을, 놓쳤거나 오프라인 상태였지만 현재는 온라인 상태가 된 다음 에이전트로 라우팅하여, "재시도"를 일부 지정 한계에 도달할 때까지 쓸 수 있다(가령, 에이전트가 도달될 수 없다고 간주되기 전에 최대 5번 재시도할 수 있다).
- [0029] 제어기(202)가 또한 컨텍스트 정보의 초기화된 세트를 제1 에이전트(204)에게도 전송하는 경우가 있을 수 있지만, 이러한 정보가 맵-리듀스 동작(요청)과 함께 수신되지 않은 경우 제1 에이전트(204)가 컨텍스트 정보를 자동으로 생성할 수 있기 때문에, 이는 필수가 아니다.
- [0030] ③에서, 제1 에이전트(204)가 자신의 로컬 데이터에 대해 맵-리듀스 동작을 실행하여, 맵-리듀스 결과를 획득할 수 있으며, 그 후 동작 컨텍스트 정보를 업데이트한다. 맵-리듀스 컨텍스트가 업데이트되어 제1 에이전트(204)의 로컬 데이터에 대한 맵-리듀스 동작을 실행한 결과를 반영할 수 있고, 컨텍스트 정보에서 "실행중인" 에이전트 엔트리가 업데이트되어 제1 에이전트(204)가 맵-리듀스 동작을 완료했음을 반영할 수 있다.
- [0031] 맵-리듀스 동작은 식별될 수 없는 소스(가령, 사용자 신원 또는 사용자 장치 신원)로부터 왔기 때문에 데이터 프라이버시를 보존하거나, 데이터 내에 있거나 상기 데이터와 연관될 수 있는 소스 신원 정보의 노출을 방지하는 방식으로 로컬 데이터를 처리한다.
- [0032] ④에서, 이 구현예에서 컨텍스트 정보에 포함되는 것으로서, 제1 에이전트(204)는 제어기(202)를 중개자(또는 프록시)로서 이용해, 진행 중인 동작 및 컨텍스트를 다음 에이전트, 가령, 제2 에이전트(206)로 전달할 수 있다. 제어기(202)는 업데이트 컨텍스트 정보를 기초로 분산된 맵-리듀스 동작의 타깃을 새로운 에이전트(가령, 제2 에이전트(206))로 재설정한다.
- [0033] ⑤에서, 제어기(202)는 제1 에이전트(204)로부터의 업데이트된 컨텍스트 정보에 따라 제2 "온-프레미스" 에이전트(206)와 통신하여, (정의를 포함하는) 맵-리듀스 동작을 리스트 상의 다음(가령, 온라인 상태의) 에이전트, 즉, 제2 에이전트(206)로 제출할 수 있다.
- [0034] ⑥에서, 제2 에이전트(206)는 자신의 로컬 데이터에 대해 맵-리듀스 동작을 실행하여 맵-리듀스 결과를 획득하고, 그 후, 동작 컨텍스트 정보를 업데이트할 수 있다. 맵-리듀스 컨텍스트 정보가 업데이트되어 제2 에이전트(206)의 로컬 데이터에 대해 맵-리듀스 동작을 실행시킨 결과를 반영할 수 있고, 컨텍스트 정보 내 에이전트 엔트리가 업데이트되어 제2 에이전트(206)가 맵-리듀스 동작을 완료했음을 반영할 수 있다.
- [0035] ⑦에서, 이 구현예의 경우 컨텍스트 정보에 포함되기 때문에, 제2 에이전트(206)가 이 특정 맵-리듀스 세션에서 마지막 에이전트인 경우, 제2 에이전트(206)는 맵-리듀스 동작, 연관된 컨텍스트 정보, 및 맵-리듀스 결과를 제어기(202)로 다시 전송한다.
- [0036] ⑧에서, 제어기(202)는 맵-리듀스 결과를 요청한 개체에 출력한다. 에이전트(204 및 206)로부터 수신된 때 결과의 최소 임계치에 이르는 경우에만 제어기(202)가 결과를 출력하는 경우가 있을 수 있다. 예를 들어, 최소 임계치가 도달 가능한(온라인) 에이전트 중 응답한 퍼센티지(가령, 80%)로 결정될 수 있다. 임계치가 수집되는 결과

의 특정 유형(가령, 날씨 상태) 및 결과가 희망되는 타이밍(가령, 현재 또는 다음 1시간 내에)에 따라 달라지는 경우가 있을 수 있다. 또한 요청되는 데이터의 유형, 가령, 이미지만 있는 데이터인지, 또는 비디오만 있는 데이터인지를 기초로 할 수 있다.

- [0037] 도 3은 개시된 아키텍처에 따르는 맵-리듀스 시스템(300)의 피어 구현에 흐름을 도시한다. 따라서 제어기(202)는 최종 에이전트가 완료될 때까지 맵-리듀스 동작의 피어 통신(peer communication) 동안 중개 기능부로서 우회된다. 시스템(300)은 온라인 제어기(202) 및 온-프레미스 에이전트의 세트, 즉, 제1 에이전트(204), 제2 에이전트(206), 제3 에이전트(302) 및 제4 에이전트(304)를 포함하며, 여기서 에이전트는 로컬 데이터 내에 위치하거나 로컬 데이터(가령, 장치 드라이브 저장부)와 연관되어 위치한다.
- [0038] 먼저, ①에서, "순회" (분산 실행) 맵-리듀스 동작이 (노드(102)와 유사하게) 제어기(202)로 제출된다. 맵-리듀스 동작은 맵 및 리듀스 동작 정의를 (이 맵-리듀스 세션에 참여하도록 지정된) 에이전트의 에이전트 속성의 세트와 함께 포함한다.
- [0039] ②에서, 제어기(202)는 제1 "온-프레미스" 에이전트(204)와 통신하여 (정의를 포함하는) 맵-리듀스 동작을 지정된 에이전트(가령, 제1 에이전트(204))로 제출할 수 있다.
- [0040] 제1 에이전트(204)가 리스트 상의 첫 번째 지정된 에이전트였지만, 상기 제1 에이전트(204)가 제어기(202)에게 오프라인인 경우가 있을 수 있다. 이러한 경우, 제어기(202)는 지정 에이전트 리스트 상의 다음 에이전트, 가령, 제2 에이전트(206)를 접촉할 수 있다. 이 프로세스는 온라인 상태 에이전트가 발견될 때까지 계속될 수 있다. 덧붙여, 온라인 상태 에이전트가 처리 완료되면, 제어기(202) 및 마지막 온라인 상태 에이전트(가령, 제4 에이전트(304))가 맵-리듀스 동작을, 놓쳤거나 오프라인 상태였지만 현재는 온라인 상태가 된 다음 에이전트로 라우팅하여, "재시도"를 일부 지정 한계에 도달할 때까지 쓸 수 있다(가령, 에이전트가 도달될 수 없다고 간주되기 전에 최대 5번 재시도할 수 있다).
- [0041] 제어기(202)가 또한 컨텍스트 정보의 초기화된 세트를 제1 에이전트(204)에게도 전송하는 경우가 있을 수 있지만, 컨텍스트 정보가 맵-리듀스 동작(요청)과 함께 수신되지 않은 경우 제1 에이전트(204)가 자동으로 이러한 정보를 생성할 수 있기 때문에, 이는 필수가 아니다.
- [0042] ③에서, 제1 에이전트(204)가 자신의 로컬 데이터에 대해 맵-리듀스 동작을 실행하여, 맵-리듀스 결과를 획득할 수 있고, 그 후 동작 컨텍스트 정보를 업데이트한다. 맵-리듀스 상황이 업데이트되어 제1 에이전트(204)의 로컬 데이터에 대해 맵-리듀스 동작을 실행한 결과를 반영할 수 있고, 컨텍스트 정보 내 "실행 중인" 에이전트 엔트리가 업데이트되어, 제1 에이전트(204)가 맵-리듀스 동작을 완료했음을 반영할 수 있다.
- [0043] 맵-리듀스 동작은 식별될 수 없는 소스(가령, 사용자 신원 또는 사용자 장치 신원)로부터 오기 때문에 데이터 프라이버시를 보존하고, 데이터에 존재할 수 있고 데이터와 연관될 수 있는 소스 신원 정보의 노출을 막는 방식으로 로컬 데이터를 처리한다.
- [0044] ④에서, 이 구현예의 경우 컨텍스트 정보에 포함되기 때문에, 제1 에이전트(204)는 진행 중인 동작(요청) 및 컨텍스트를 (제어기(202)를 우회하여) 제2 에이전트(206)로 직접 전달한다.
- [0045] ⑤에서, 제2 에이전트(206)가 자신의 로컬 데이터에 대해 맵-리듀스 동작을 실행하여 맵-리듀스 결과를 획득할 수 있고, 그 후, 동작 컨텍스트 정보를 업데이트한다. 맵-리듀스 컨텍스트 정보는 업데이트되어 제2 에이전트(206)의 로컬 데이터에 대한 맵-리듀스 동작을 실행시킨 결과를 반영할 수 있고, 컨텍스트 정보 내 에이전트 엔트리가 업데이트되어 제2 에이전트(206)가 맵-리듀스 동작을 완료했음을 반영할 수 있다.
- [0046] ⑥에서, 이 구현예의 경우 컨텍스트 정보에 포함되기 때문에, 제2 에이전트(206)는 진행 중인 동작(요청) 및 컨텍스트를 (제어기(202)를 우회하여) 제3 에이전트(302)로 직접 전달한다.
- [0047] ⑦에서, 제3 에이전트(302)는 자신의 로컬 데이터에 대해 맵-리듀스 동작을 실행하여 맵-리듀스 결과를 획득할 수 있고, 그 후 동작 컨텍스트 정보를 업데이트한다. 맵-리듀스 컨텍스트 정보가 업데이트되어 제3 에이전트(302)의 로컬 데이터에 대해 맵-리듀스 동작을 실행한 결과를 반영할 수 있고, 컨텍스트 정보 내 에이전트 엔트리가 업데이트되어 제3 에이전트(302)가 맵-리듀스 동작을 완료했음을 반영할 수 있다. 이 구현예의 경우 컨텍스트 정보에 포함되기 때문에, 제3 에이전트(302)는 진행 중인 동작(요청) 및 컨텍스트를 (제어기(202)를 우회하여) 제4 에이전트(304)로 직접 전달한다.
- [0048] ⑧에서, 제4 에이전트(304)는 자신의 로컬 데이터에 대해 맵-리듀스 동작을 실행하여 맵-리듀스 결과를 획득할 수 있고, 그 후 동작 컨텍스트 정보를 업데이트한다. 맵-리듀스 컨텍스트 정보가 업데이트되어 제4 에이전트

(304)의 로컬 데이터에 대해 맵-리듀스 동작을 실행한 결과를 반영할 수 있고, 컨텍스트 정보 내 에이전트 엔트리가 업데이트되어 제4 에이전트(304)가 맵-리듀스 동작을 완료했음을 반영할 수 있다.

- [0049] 이 구현예의 경우 컨텍스트 정보에 포함되기 때문에, 제4 에이전트(304)는 이 맵-리듀스 세션에서 온라인 상태이며 응답하는 마지막 에이전트이다. 따라서 ㉑에서, 제4 에이전트(304)는 맵-리듀스 동작, 연관된 컨텍스트 정보, 및 맵-리듀스 결과를 제어기(202)로 다시 전송한다.
- [0050] ㉑에서, 제어기(202)가 맵-리듀스 결과를 요청 개체에게 출력한다. 에이전트(204, 206, 302 및 304)로부터 수신된 때 결과의 최소 임계치에 이르는 경우에만 제어기(202)가 결과를 출력하는 경우가 있을 수 있다. 예를 들어, 최소 임계치가 도달 가능한 (온라인 상태의) 에이전트 중 응답한 정족수(가령, 단순히 과반수) 또는 퍼센티지(가령, 80퍼센트)로 결정될 수 있다. 수집되는 결과의 특정 유형(가령, 날씨 조건) 및 결과가 희망되는 타이밍(가령, 지금 또는 다음 1시간 내에)에 따라 임계치가 달라지는 경우가 있을 수 있다. 또한 요청되는 데이터의 유형, 가령, 이미지만 있는 데이터인지 또는 비디오만 있는 데이터인지를 기초로 할 수 있다.
- [0051] 도 4는 본 발명의 아키텍처에 따르는 맵-리듀스 시스템(400)의 병렬 구현 흐름도를 도시한다. 시스템(400)은 제어기(202) 및 에이전트가 데이터 저장소(장치) 내에 있거나 이와 연관되어 있는 온-프레미스(장치상에 있는) 에이전트의 세트를 포함할 수 있다.
- [0052] 먼저, ①에서, 분산 실행 맵-리듀스 동작이 (노드(102)와 유사하게) 제어기(202)로 제출된다. 맵-리듀스 동작은 이 맵-리듀스 세션에 참여하도록 지정된 에이전트의 에이전트 속성의 세트와 함께 맵 및 리듀스 동작 정의를 포함한다.
- [0053] ②, ③ 및 ④에서, 제어기(202)는 각각의 에이전트(204, 206 및 302)와 병렬로 통신하여 (정의를 포함하는) 맵-리듀스 동작을 지정 에이전트로 제출할 수 있다.
- [0054] 각각의 에이전트가 독립적으로 동작하여 자신의 로컬 데이터에 대해 맵-리듀스 동작을 실행하여 맵-리듀스 결과를 획득하고 그 후 대응하는 동작 컨텍스트 정보를 업데이트할 수 있다. 상기 맵-리듀스 컨텍스트가 업데이트되어, 에이전트(204, 206 및 302)의 로컬 데이터에 대한 맵-리듀스 동작을 실행한 결과를 반영하고, 컨텍스트 정보 내 "실행 중인" 에이전트 엔트리가 업데이트되어, 대응하는 에이전트(204, 206, 및 302)가 맵-리듀스 동작을 완료했음을 반영할 수 있다.
- [0055] 맵-리듀스 동작은 식별 될 수 없는 소스(가령, 사용자 신원 또는 사용자 장치 신원)로부터 오기 때문에 데이터 프라이버시를 보존하고, 데이터에 존재할 수 있고 데이터와 연관될 수 있는 소스 신원 정보의 노출을 막는 방식으로 로컬 데이터를 처리한다.
- [0056] ⑤, ⑥ 및 ⑦에서, 각각의 대응하는 에이전트(204, 206, 및 302)가 컨텍스트 정보와 연관된 맵-리듀스 동작을 전송하고, 제어기(202)로 맵-리듀스 결과를 다시 전송한다. 그 후 제어기(202)는 에이전트(204, 206, 및 302)의 결과 및 컨텍스트 정보를 결과 및 컨텍스트 정보의 최종 세트로 처리하여, 충분한 에이전트가 원하는 정보로 응답했음을 보장할 수 있다.
- [0057] ⑧에서, 제어기(202)가 맵-리듀스 결과를 요청 개체에게 출력한다. 앞서 언급한 바와 같이, 제어기(202)는 에이전트(204 및 206)로부터 수신된 때 결과가 최소 임계치에 도달할 때에만 출력한다. 예를 들어, 최소 임계치는 도달될 수 있는(온라인 상태의) 에이전트 중 응답한 퍼센티지(가령, 80퍼센트)로 결정될 수 있다. 임계치가 수집되는 결과의 특정 유형(가령, 날씨 조건) 및 결과가 희망되는 타이밍(가령, 현재 또는 다음 1시간 내)에 따라 달라지는 경우가 있을 수 있다. 이는, 요청된 데이터의 유형, 가령, 이미지만 있는 데이터인지, 또는 비디오만 있는 데이터인지를 기초로 할 수 있다.
- [0058] 도 5는 개시된 아키텍처에 따르는 맵-리듀스 시스템(500)의 위치 기반 피어-투-피어 구현 흐름을 도시한다. 이 구현예에서, 맵-리듀스 동작 세션이 특정 지리적 영역(502)과 연관되거나 영역(502)과 일정 레벨의 관련도를 가진다고 결정된 에이전트와 관련된다.
- [0059] 먼저, (1)에서, 분산 실행 맵-리듀스 동작이 제어기(202)로 제출된다. 맵-리듀스 동작은 이 맵-리듀스 세션에 참여하도록 지정된 에이전트의 에이전트 속성의 세트와 함께, 맵 및 리듀스 동작 정의를 포함한다. 제1 에이전트(204)가 세션에 대해 지정되는 동안, 정보의 소스는 제1 에이전트(204)가 영역(502)과 더는 연관되지 않고 일부 지정 시간 동안 연관되지 않기("만료") 때문에, 실질적으로 실시간 이익을 위해, 제1 에이전트(204)가 원하는 정보에 대해 더는 관련성이 없음을 가리킬 수 있다. 따라서 제어기(202)는 제1 에이전트(204)를 무효화하고 세션 프로세싱으로부터 해입할 수 있다.

- [0060] 제어기(202)(또는 제어기(202)로 인터페이스하는 그 밖의 다른 일부 적합한 구성요소)에 의한 이 사전-세션 참여 결정이, 가령, 지리적 정보에 대한, 지정된 에이전트(204, 206, 302, 및 304)로의 짧은 통신에 의해 이뤄질 수 있다. 어느 경우라도, 제2 에이전트(206), 제3 에이전트(302), 및 제4 에이전트(304)가 영역(502)과 밀접하게 연관된 것으로 결정되고 맵-리듀스 세션 동안 처리될 것이다.
- [0061] 따라서 ②에서, 제어기(202)는 맵-리듀스 동작(요청)을 제출함으로써, 에이전트, 가령, 제1 에이전트(204)와 통신하여, 세션을 개시할 수 있다. ③에서, 제2 에이전트(206)는 자신의 로컬 데이터에 대해 맵-리듀스 동작을 실행하여 맵-리듀스 결과를 획득할 수 있고, 그 후 동작 컨텍스트 정보를 업데이트한다. 맵-리듀스 컨텍스트는 업데이트되어 제2 에이전트(206)의 로컬 데이터에 대한 맵-리듀스 동작을 실행한 결과를 반영할 수 있고, 컨텍스트 정보 내 "실행 중인" 에이전트 엔트리가 업데이트되어 제2 에이전트(206)가 맵-리듀스 동작을 완료했음을 반영할 수 있다. 이 구현예의 경우 컨텍스트 정보에 포함되기 때문에, 제2 에이전트(206)는 진행 중인 동작(요청) 및 컨텍스트를 (제어기(202)를 우회하여) 제3 에이전트(302)에게 직접 전달한다.
- [0062] ④에서, 제3 에이전트(302)는 자신의 로컬 데이터에 대해 맵-리듀스 동작을 실행하여 맵-리듀스 결과를 획득하고, 그 후 동작 컨텍스트 정보를 업데이트한다. 맵-리듀스 컨텍스트 정보가 업데이트되어 제3 에이전트(302)의 로컬 데이터에 대해 맵-리듀스 동작을 실행한 결과를 반영할 수 있고, 컨텍스트 정보 내 에이전트 엔트리가 업데이트되어 제3 에이전트(302)가 맵-리듀스 동작을 완료했음을 반영할 수 있다. 이 구현예의 경우 컨텍스트 정보에서 지시되기 때문에, 제3 에이전트(302)는 진행 중인 동작(요청) 및 컨텍스트를 (제어기(202)를 우회하여) 제4 에이전트(304)로 직접 전달한다.
- [0063] ⑤에서, 제4 에이전트(304)는 자신의 로컬 데이터에 대해 맵-리듀스 동작을 실행하여 맵-리듀스 결과를 획득하고, 그 후 동작 컨텍스트 정보를 업데이트한다. 맵-리듀스 컨텍스트 정보가 업데이트되어 제4 에이전트(304)의 로컬 데이터에 대해 맵-리듀스 동작을 실행한 결과를 반영할 수 있고, 컨텍스트 정보 내 에이전트 엔트리가 업데이트되어 제4 에이전트(304)가 맵-리듀스 동작을 완료했음을 반영할 수 있다. 이 구현예의 경우 컨텍스트 정보에 포함되기 때문에, 제4 에이전트(304)는 결과 및 업데이트된 컨텍스트를 제어기(202)로 직접 전달한다. ⑥에서, 제어기(202)가 맵-리듀스 결과를 요청 개체에게 출력한다.
- [0064] 도 6은 개시된 아키텍처에 따르는 맵-리듀스 시스템(600)의 조합 구현 흐름을 도시한다. 여기서, 조합 구현예는 직렬, 병렬, 및 피어-투-피어 맵-리듀스 처리를 가능하게 한다. 이 예시에서, 제어기(202)는 제1 에이전트(204)에 순차로 맵-리듀스 동작을 개시하고, 이는 제2 에이전트(206)로 맵-리듀스 동작을 동시에 전송하는 것과 병렬로 이뤄질 수 있다.
- [0065] 그 후 제2 에이전트(206)는 제3 에이전트(302) 및 제4 에이전트(304)와의 맵-리듀스 세션을 피어 방식으로 계속 하도록 동작한다. 최종적으로, 제4 에이전트가 제1 에이전트(204)처럼 결과 및 업데이트된 컨텍스트를 제어기(202)로 반환한다.
- [0066] 상기 구현예 모두에서, 각각의 에이전트가 일부 또는 모든 에이전트 장치 및/또는 상이한 에이전트 장치 상에서 현재 세션의 일부가 아닌 것을 일부 이해할 수 있는(뷰어(viewer) 친화적인) 방식으로 본 최종 결과를 제공할 수 있는 것이 바람직하다.
- [0067] 덧붙여, 개시된 아키텍처에서, 특정 구성요소가 재배열, 조합, 생략될 수 있고 추가 구성요소가 포함될 수 있음이 이해될 것이다.
- [0068] 도시되지는 않았지만, 프라이버시 구성요소가 사용자 및 장치 정보의 보안된 핸들링의 추가 레이어에 대해 채용될 수 있다. 프라이버시 구성요소에 의해, 사용자가 로컬 데이터 액세스를 옵트-인(opt-in) 및 옵트-아웃(opt-out)할 수 있다.
- [0069] 도 7은 맵-리듀스 동작 세션을 이루기 위해 제어기 및 에이전트에 의해 통신될 수 있는 데이터세트(700)를 도시한다. 데이터세트(700)는 컨텍스트 정보(106), 맵-리듀스 동작(요청)(104), 및 에이전트 결과(702)를 포함할 수 있다. 컨텍스트 정보(106)는 맵-리듀스 동작에 대해 요청될 특정 에이전트를 가리키는 지정된 에이전트 리스트(704)를 더 포함할 수 있다. 리스트(704)는 특정 에이전트에 대한 맵-리듀스 동작의 에이전트 속성, 가령, 에이전트 식별자(가령, 에이전트1) 및 상태(상태1), 가령, "미완료" 또는 "완료"를 나타내는 값을 정의하는 것을 가능하게 한다.
- [0070] 또한 리스트(704)는 에이전트가 처리될 수 있는 순서, 가령, 등급화된 우선순위 또는 탑-다운 우선순위를 가리킬 수 있다. 제1 에이전트가 동작을 완료할 때, 제1 에이전트는 리스트(704)를 참조하여, 제2 에이전트가 맵-리

듀스 처리를 위한 다음 순서임을 알 수 있다. 그러나 제2 에이전트가 도달 가능하지 않은 경우(제1 에이전트에게 "오프라인"인 경우), 제1 에이전트는 다음 에이전트, 즉, 제3 에이전트로 이동한다. 이와 유사한 동작이 제어기(202)에 의해 수행되어, 컨텍스트 정보(106), 맵-리듀스 동작(요청)(104), 및 결과(702)를 리스트(704) 상의 다음 번 지시되는 에이전트에게 전달할 수 있다.

- [0071] 에이전트 결과(702)가 제어기(202)에 의해 최종적으로 처리될 대응하는 이전 에이전트의 이전 결과들의 집성, 또는 가령 각각의 에이전트가 완료될 후 획득된 결과들의 중간 결과 컴필레이션일 수 있다.
- [0072] 개시된 아키텍처의 신규한 양태를 수행하는 예시적 방법을 나타내는 흐름도의 세트가 본 명세서에 포함된다. 설명을 간단히 하기 위해, 흐름도의 형태로 본 명세서에 나타난 하나 이상의 방법이 일련의 동작들로 도시되고 기재되지만, 방법은 동작들의 순서에 한정되지 않으며, 일부 동작은 상이한 순서 또는 본 명세서에 도시되고 기재된 다른 동작과 동시에 발생할 수 있음이 이해되어야 한다. 예를 들어, 해당 분야의 통상의 기술자라면 방법이 일련의 상태 또는 이벤트로서, 가령, 상태 다이어그램으로 나타낼 수 있음을 알 것이다. 덧붙여, 방법에 도시된 동작 모두가 신규한 구현예를 위해 필요한 것은 아니다.
- [0073] 도 8은 개시된 아키텍처에 따르는 방법을 도시한다. (800)에서, 맵-리듀스 동작 요청이 하나의 노드에서 대응하는 하나 이상의 에이전트로 전송되어, 하나 이상의 에이전트의 로컬 데이터에 대해 맵-리듀스 동작을 수행할 수 있다. (802)에서, 맵-리듀스 결과 및 업데이트된 컨텍스트 정보가 맵-리듀스 동작 요청을 기초로 하나 이상의 에이전트로부터 노드에서 수신된다. (804)에서, 맵-리듀스 결과 및 업데이트된 컨텍스트 정보가 노드로부터 출력된다.
- [0074] 방법은 에이전트 상의 맵-리듀스 동작의 일부로서 로컬 데이터의 프라이버시를 보존하는 단계를 더 포함할 수 있다. 방법은 컨텍스트 정보를 업데이트하여 특정 에이전트가 맵-리듀스 동작을 완료했음을 식별하는 단계를 더 포함할 수 있다. 상기 방법은 맵-리듀스 동작 요청을 이전에 오프라인 상태였던 온라인 에이전트로 리디렉트하는 단계를 더 포함할 수 있다.
- [0075] 상기 방법은 지정된 에이전트에 병렬로 맵-리듀스 동작 요청을 전송하는 단계를 더 포함할 수 있다. 상기 방법은 지정된 에이전트의 리스트 중 지정된 에이전트를 통해 맵-리듀스 동작 요청을 순차적으로 전송하는 단계를 더 포함할 수 있다. 상기 방법은 하나의 에이전트로부터의 맵-리듀스 결과 및 업데이트된 컨텍스트 정보를 다른 에이전트에 의해 노드에서 점진적으로 누적하는 단계를 더 포함할 수 있다.
- [0076] 도 9는 개시된 아키텍처에 따르는 대안적 방법을 도시한다. 상기 방법은 마이크로프로세서에 의해 실행될 때 마이크로프로세서로 하여금 다음의 동작을 수행하게 하는 컴퓨터 실행 명령어를 포함하는 컴퓨터 판독형 저장 매체로 구현될 수 있다.
- [0077] (900)에서, 맵-리듀스 세션의 맵-리듀스 동작 요청이 하나의 노드로부터 지정된 에이전트로 전송되어, 에이전트의 로컬 데이터에 대한 맵-리듀스 동작을 수행할 수 있다. (902)에서, 통계치, 데이터, 및 컨텍스트 정보가 임의의 특정 에이전트로부터 도출된 식별 불가능 정보로서 수신된다. (904)에서, 식별 불가능 통계치, 데이터, 및 컨텍스트 정보가 지정된 에이전트들 간에 전달된다. (906)에서, 식별 불가능한 통계치, 데이터, 및 컨텍스트 정보가 노드로부터 출력된다.
- [0078] 컴퓨터 판독형 저장 매체는 지정된 에이전트에 병렬로 맵-리듀스 동작 요청을 전송하는 단계를 더 포함한다. 컴퓨터 판독형 저장 매체는 맵-리듀스 동작 요청을 지정된 제1 에이전트로 전송하는 단계, 상기 지정된 제1 에이전트로부터 식별 불가능한 통계치, 데이터, 및 컨텍스트 정보를 수신하는 단계, 지정된 제1 에이전트의 식별 불가능한 통계치, 데이터, 및 컨텍스트 정보를 지정된 제2 에이전트로 순차로 전달하는 단계를 포함할 수 있다. 컴퓨터 판독형 저장 매체로서, 여기서 노드가 맵 동작 정의 및 리듀스 동작 정의 및 맵-리듀스 세션에 참여하는 에이전트의 에이전트 속성의 세트를 수신한다.
- [0079] 본 명세서에서 사용될 때, "구성요소" 및 "시스템"이라는 용어는 컴퓨터 관련 개체, 즉 하드웨어, 소프트웨어 및 유형의 하드웨어의 조합, 소프트웨어, 또는 실행 소프트웨어를 지칭하는 것으로 의도된다. 예를 들어, 구성요소의 비제한적 예를 들면, 유형의 구성요소, 가령, 마이크로프로세서, 칩 메모리, 대용량 저장 장치(가령, 광학 드라이브, 솔리드 스테이트 드라이브, 및/또는 자기 저장 매체 드라이브), 및 컴퓨터, 및 소프트웨어 구성요소, 가령, 마이크로프로세서 상에서 실행되는 프로세스, 객체, 실행파일, (휘발성 또는 비휘발성 저장 매체에 저장된) 데이터 구조, 모듈, 실행 스레드 및/또는 프로그램일 수 있다.
- [0080] 예를 들어, 서버 상에서 실행되는 애플리케이션과 서버 모두 구성요소일 수 있다. 하나 이상의 구성요소가 프로세스 및/또는 실행 스레드 내에 위치할 수 있고, 구성요소는 하나의 컴퓨터 상에 로컬화되거나 및/또는 둘 이상

의 컴퓨터 간에 분산될 수 있다. "예시적"이라는 단어는 본 명세서에서 예시, 사례, 또는 일례로서 역할 함을 의미한다. 본 명세서에 "예시"로 기재된 임의의 양태 또는 설계가 반드시 그 밖의 다른 양태 또는 설계에 비해 바람직하거나 유리한 것으로 해석되는 것은 아니다.

- [0081] 도 10을 참조하면, 순회 맵-리듀스 아키텍처를 실행하는 컴퓨팅 시스템(1000)의 블록도가 도시된다. 그러나 개시된 방법 및/또는 시스템의 일부 또는 모든 측면이 시스템-온-칩으로 구현될 수 있으며, 여기서 아날로그, 디지털, 혼합 신호, 및 그 밖의 다른 기능이 단일 칩 기관 상에 조립된다.
- [0082] 다양한 측면에 대한 추가 맥락을 제공하기 위해, 도 10 및 이하의 설명은 다양한 측면이 제공되는 적합한 컴퓨팅 시스템(1000)의 간략하고, 일반적인 기재를 제공하려는 의도이다. 상기의 기재가 하나 이상의 컴퓨터 상에서 실행될 수 있는 컴퓨터-실행 명령어의 일반적인 맥락에서 제공되었지만, 해당 분야의 통상의 기술자라면 신규한 실시예가 또한 그 밖의 다른 프로그램 모듈과 조합되어 및/또는 하드웨어 및 소프트웨어의 조합으로서 구현될 수 있음을 알 것이다.
- [0083] 다양한 형태를 구현하기 위한 컴퓨팅 시스템(1000)은 마이크로프로세싱 유닛(1004)(마이크로프로세서 및 프로세서라고 지칭됨), 컴퓨터 판독형 저장 매체, 가령, 시스템 메모리(1006)(컴퓨터 판독형 저장 매체는 또한 자기 디스크, 광학 디스크, 솔리드 스테이트 드라이브, 외부 메모리 시스템 및 플래시 메모리 드라이브를 포함한다), 및 시스템 버스(1008)를 갖는 컴퓨터(1002)를 포함한다. 마이크로프로세싱 유닛(1004)은 상용화된 다양한 마이크로프로세서, 가령, 싱글-프로세서, 멀티-프로세서, 처리 및/또는 저장 회로의 싱글-코어 유닛 및 멀티-코어 유닛 중 임의의 것일 수 있다. 덧붙여, 해당 분야의 통상의 기술자라면 신규한 시스템 및 방법이 그 밖의 다른 컴퓨터 시스템 구성, 가령, 미니컴퓨터, 메인프레임 컴퓨터 및 개인 컴퓨터(가령, 데스크톱, 랩톱, 태블릿 PC 등), 핸드-헬드 컴퓨팅 장치, 마이크로프로세서-기반 또는 프로그래머블 소비자 전자장치 등에 의해 실시될 수 있으며, 이들 각각은 하나 이상의 연관된 장치에 동작 가능하게 연결될 수 있다.
- [0084] 컴퓨터(1002)는 휴대용 및/또는 모바일 컴퓨팅 시스템, 가령, 무선 통신 장치, 셀룰러 전화기, 및 그 밖의 다른 모바일 가능 장치를 위한 클라우드 컴퓨팅 서비스를 지원할 때 데이터센터 및/또는 컴퓨팅 자원(하드웨어 및/또는 소프트웨어)에서 사용되는 복수의 컴퓨터 중 하나일 수 있다. 클라우드 컴퓨팅 서비스의 비제한적 예를 들면, 서비스형 인프라구조, 서비스형 플랫폼, 서비스형 소프트웨어, 서비스형 저장장치, 서비스형 데스크톱, 서비스형 데이터, 서비스형 보안, 서비스형 API(application program interface)가 있다.
- [0085] 시스템 메모리(1006)는 컴퓨터 판독형 저장(물리 저장장치) 매체, 가령, 휘발성(VOL) 메모리(1010)(가령, 랜덤 액세스 메모리(RAM)), 및 비휘발성 메모리(NON-VOL)(1012)(가령, ROM, EPROM, EEPROM 등)를 포함할 수 있다. 기본 입/출력 시스템(BIOS)이 비휘발성 메모리(1012)에 저장될 수 있고, 가령 시동 동안, 컴퓨터(1002) 내 구성요소들 간 데이터 및 신호의 통신을 가능하게 하는 기본 루틴을 포함한다. 휘발성 메모리(1010)는 데이터를 캐싱하기 위한 고속 RAM, 가령, 정적 RAM을 더 포함할 수 있다.
- [0086] 시스템 버스(1008)는 시스템 구성요소를 위한 인터페이스, 가령, 시스템 메모리(1006)와 마이크로프로세싱 유닛(1004) 간 인터페이스를 제공한다. 상기 시스템 버스(1008)는 상용화된 다양한 버스 아키텍처 중 임의의 것을 이용해 (메모리 제어를 갖거나 갖지 않는) 메모리 버스 및 주변장치 버스(가령, PCI, PCIE, AGP, LPC 등)로 더 인터커넥트될 수 있는 복수 유형의 버스 구조 중 임의의 것일 수 있다.
- [0087] 컴퓨터(1002)는 저장 서브시스템(1014)을 시스템 버스(1008) 및 그 밖의 다른 바람직한 컴퓨터 구성요소 및 회로로 인터페이스하기 위한 머신 판독형 저장 서브시스템(1014) 및 저장 인터페이스(1016)를 더 포함한다. 저장 서브시스템(1014)(물리 저장 매체)은 예를 들어 하드 디스크 드라이브(HDD), 자기 플로피 디스크 드라이브(FDD), 솔리드 스테이트 드라이브(SSD), 플래시 드라이브 및/또는 광학 디스크 저장 드라이브(가령, CD-ROM 드라이브 DVD 드라이브) 중 하나 이상을 포함할 수 있다. 저장 인터페이스(1016)는 인터페이스 기술, 가령, EIDE, ATA, SATA, 및 IEEE 1394를 포함할 수 있다.
- [0088] 하나 이상의 프로그램 및 데이터가 메모리 서브시스템(1006), 머신 판독 및이동식 메모리 서브시스템(1018)(플래시 드라이브 폼 팩터 기법) 및/또는 저장 서브시스템(1014)(가령, 광학, 자기, 솔리드 스테이트), 가령, 운영 체제(1020), 하나 이상의 애플리케이션 프로그램(1022), 그 밖의 다른 프로그램 모듈(1024), 및 프로그램 데이터(1026)에 저장될 수 있다.
- [0089] 운영 체제(1020), 하나 이상의 애플리케이션 프로그램(1022), 그 밖의 다른 프로그램 모듈(1024), 및/또는 프로그램 데이터(1026)가 도 1의 시스템(100)의 아이템 및 구성요소, 시스템(200, 300, 400, 500, 및 600)의 구현 흐름의 아이템 및 구성요소, 도 5의 데이터세트(700)의 아이템, 및 도 8 및 9의 흐름도에 의해 표현되는 방법을

포함할 수 있다.

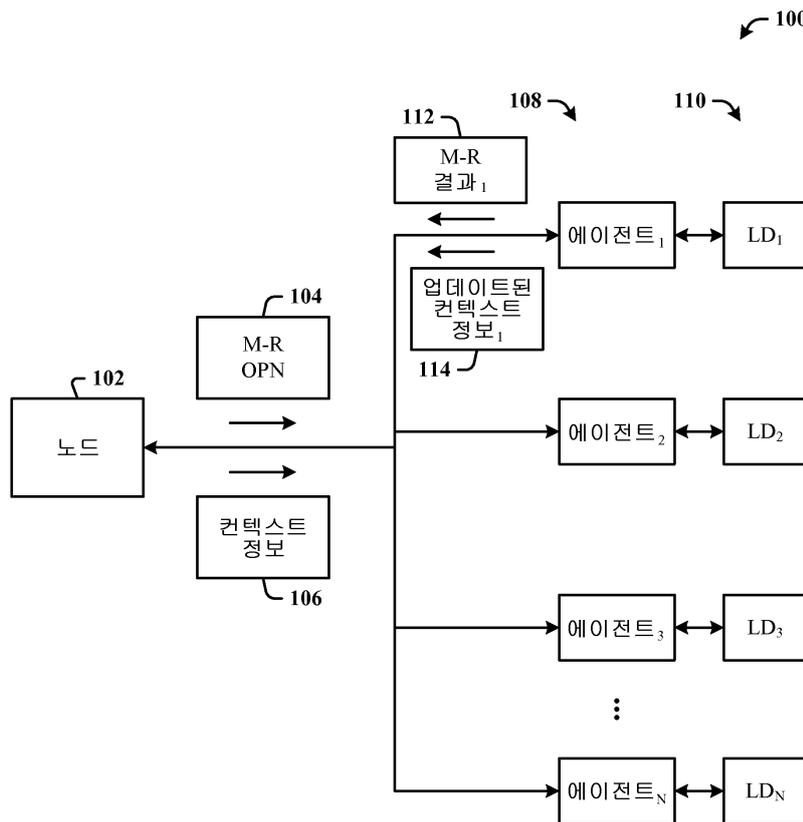
- [0090] 일반적으로, 프로그램은 특정 작업을 구현하거나, 기능하거나, 특정 추상화 데이터 형을 구현하는 루틴, 방법, 데이터 구조, 그 밖의 다른 소프트웨어 구성요소 등을 포함한다. 운영 체제(1020), 애플리케이션(1022), 모듈(1024) 및/또는 데이터(1026)가 메모리, 가령, 휘발성 메모리(1010) 및/또는 비휘발성 메모리에 캐싱될 수 있다. 개시된 아키텍처가 상용화된 다양한 운영 체제, 또는 운영 체제의 조합으로 (가령, 가상 머신으로서) 구현될 수 있음을 알아야 한다.
- [0091] 저장 서브시스템(1014) 및 메모리 서브시스템(1006 및 1018)이 데이터, 데이터 구조, 컴퓨터 실행 명령어 등의 휘발성 및 비휘발성 저장을 위한 컴퓨터 관독형 매체로서 역할한다. 이러한 명령어는, 컴퓨터 또는 그 밖의 다른 머신에 의해 실행될 때, 컴퓨터 또는 그 밖의 다른 머신이 방법 중 하나 이상의 동작을 수행할 수 있게 한다. 컴퓨터 실행 명령어는, 예를 들어, 범용 컴퓨터, 특수 컴퓨터, 또는 특수 마이크로프로세서 장치가 특정 기능 또는 기능 그룹을 수행하게 하는 명령어 및 데이터를 포함한다. 컴퓨터 실행 명령어는, 예를 들어, 바이너리, 중간 포맷 명령어, 가령, 어셈블리 언어, 또는 소스 코드일 수 있다. 동작을 수행하기 위한 명령어는 하나의 매체 상에 저장되거나, 복수의 매체에 걸쳐 저장되어, 모든 명령어가 동일한 매체 상에 있는지 여부와 무관하게, 명령어가 하나 이상의 컴퓨터 관독형 저장 매체 상에 집합적으로 나타나게 할 수 있다.
- [0092] 컴퓨터 관독형 저장 매체는 전파되는 신호 자체는 배제하고 컴퓨터(1002)에 의해 액세스될 수 있으며, 이동식 및/또는 비이동식인 휘발성 및 비휘발성 내부 및/또는 외부 매체를 포함할 수 있다. 컴퓨터(1002)의 경우, 다양한 유형의 매체가 임의의 적합한 디지털 포맷으로 된 데이터의 저장을 수용한다. 해당 분야의 통상의 기술자라면, 그 밖의 다른 유형의 컴퓨터 관독형 매체, 가령, 집 드라이브, 솔리드 스테이트 드라이브, 자기 테이프, 플래시 메모리 카드, 플래시 드라이브, 카트리지 등이 개시된 아키텍처의 신규한 방법(동작)을 수행하기 위한 컴퓨터 실행형 명령어를 저장하기 위해 사용될 수 있다.
- [0093] 사용자는 외부 사용자 입력 장치(1028), 가령, 키보드 및 마우스를 이용해, 그리고 음성 인식에 의해 가능해지는 음성 명령어에 의해, 컴퓨터(1002), 프로그램, 및 데이터와 대화할 수 있다. 그 밖의 다른 외부 사용자 입력 장치(1028)는 마이크로폰, IR(적외선) 원격 제어, 조이스틱, 게임 패드, 카메라 인식 시스템, 스타일러스 펜, 터치 스크린, 제스처 시스템(가령, 안구 운동, 바디 포즈, 가령, 손, 손가락, 팔, 두부 등과 관련된 것) 등을 포함할 수 있다. 사용자는 온보드 사용자 입력 장치(1030), 가령, 터치패드, 마이크로폰, 키보드 등을 이용해 컴퓨터(1002), 프로그램 및 데이터와 대화할 수 있으며, 이때 컴퓨터(1002)는 예컨대 휴대용 컴퓨터이다.
- [0094] 이들 및 그 밖의 다른 입력 장치가 시스템 버스(1008)를 통해 입/출력(I/O) 장치 인터페이스(1032)를 통해 마이크로프로세싱 유닛(1004)에 연결되지만 그 밖의 다른 인터페이스, 가령, 병렬 포트, IEEE 1394 직렬 포트, 게임 포트, USB 포트, IR 인터페이스, 단거리 무선(가령, 블루투스) 및 그 밖의 다른 개인 영역 네트워크(PAN) 기법 등에 의해 연결될 수 있다. I/O 장치 인터페이스(1032)는 또한 출력 주변장치(1034), 가령, 프린터, 오디오 장치, 카메라 장치 등, 가령, 사운드 카드 및/또는 온보드 처리 능력의 사용을 가능하게 한다.
- [0095] 하나 이상의 그래픽 인터페이스(1036)(또한 그래픽 처리 유닛(GPU)라고도 지칭됨)가 컴퓨터(1002)와 외부 디스플레이(1038)(가령, LCD, 플라즈마) 및/또는 온보드 디스플레이(1040)(가령, 휴대용 컴퓨터) 간 그래픽 및 비디오 신호를 제공한다. 그래픽 인터페이스(1036)는 또한 컴퓨터 시스템 기판의 일부로서 제작될 수 있다.
- [0096] 컴퓨터(1002)는 유선/무선 통신 서브시스템(1042)을 통한 하나 이상의 네트워크 및/또는 그 밖의 다른 컴퓨터로의 논리 연결을 이용해 네트워크 연결된 환경(가령, IP 기반)에서 동작할 수 있다. 그 밖의 다른 컴퓨터가 워크스테이션, 서버, 라우터, 개인 컴퓨터, 마이크로프로세서-기반 엔터테인먼트 기구, 피어 장치 또는 그 밖의 다른 공통 네트워크 노드를 포함할 수 있고, 일반적으로 컴퓨터(1002)에 대해 기재된 요소 중 다수 또는 전부를 포함한다. 논리 연결이 로컬 영역 네트워크(LAN), 광역 네트워크(WAN), 핫스팟 등으로의 유선/무선 연결을 포함할 수 있다. LAN 및 WAN 네트워킹 환경은 사무실 및 회사에서 흔하며 기업별 컴퓨터 네트워크, 가령, 인트라넷을 가능하게 하고 이들 모두는 전역 통신 네트워크, 가령, 인터넷으로 연결될 수 있다.
- [0097] 네트워킹 환경에서 사용될 때, 컴퓨터(1002)는 유선/무선 통신 서브시스템(1042)(가령, 네트워크 인터페이스 어댑터, 온보드 트랜시버 서브시스템 등)을 통해 네트워크로 연결되어 유선/무선 네트워크, 유선/무선 프린터, 유선/무선 입력 장치(1044) 등과 통신할 수 있다. 컴퓨터(1002)는 네트워크를 통해 통신을 구축하기 위한 모뎀 또는 그 밖의 다른 수단을 포함할 수 있다. 네트워크 연결된 환경에서, 분산 시스템과 연관될 때 컴퓨터(1002)와 관련된 프로그램 및 데이터가 원격 메모리/저장 장치에 저장될 수 있다. 도시된 네트워크 연결이 예시이며, 컴퓨터들 간 통신 링크를 확립하는 그 밖의 다른 수단이 사용될 수 있음이 자명할 것이다.

[0098] 컴퓨터(1002)는 무선 기법, 가령, IEEE 802.xx 표준 군을 이용해 유선/무선 장치 또는 개체, 가령, 무선 통신(가령, IEEE 802.11 오버-더-에어(over-the-air) 변조 기법)에서 동작 가능하게 배치된 무선 장치, 예를 들어, 프린터, 스캐너, 데스크톱 및/또는 휴대용 컴퓨터, 개인 디지털 보조기(PDA), 통신 위성, 무선 검출 가능한 태그와 연관된 임의의 장비 또는 위치(가령, 키오스크, 뉴스 가판대, 화장실) 및 전화기와 통신하도록 기능할 수 있다. 이는 핫스팟을 위한 (무선 컴퓨터 네트워킹 장치의 상호운용성을 증명하는 데 사용되는) 적어도 Wi-Fi™, WiMax, 및 Bluetooth™ 무선 기법을 포함한다. 따라서 통신은 적어도 2개의 장치들 간 종래의 네트워크, 또는 단순히 애드 혹(ad hoc) 통신을 갖는 지정된 구조일 수 있다. Wi-Fi 네트워크는 IEEE 802.11x(a, b, g 등)을 이용해, 보안되고, 신뢰할만하며, 고속의 무선 연결성을 제공할 수 있다. Wi-Fi 네트워크는 컴퓨터를 서로, 인터넷에, 그리고 (IEEE 802.3-관련 기법 및 기능을 이용하는) 유선 네트워크에 연결하도록 사용될 수 있다.

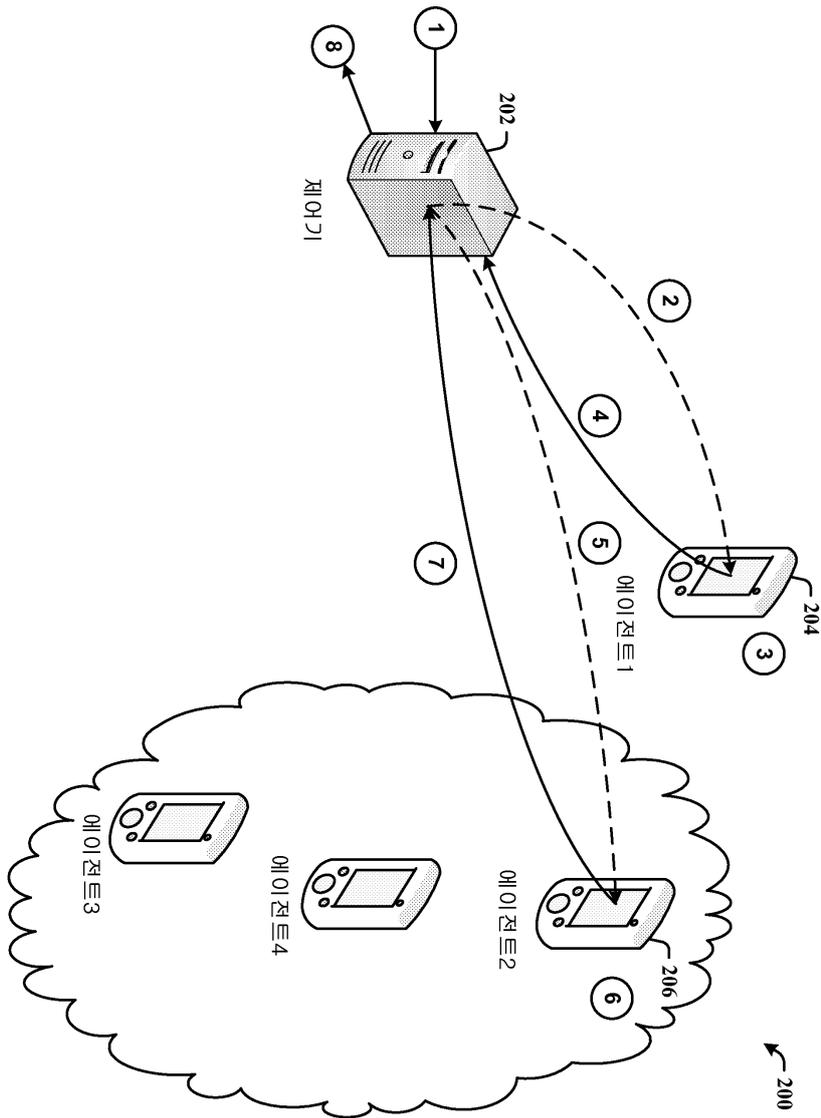
[0099] 앞서 본 발명의 아키텍처의 예시가 기재되었다. 물론, 구성요소 및/또는 방법의 상상할 수 있는 모든 조합을 기재하는 것이 불가능하며, 해당 분야의 통상의 기술자라면 많은 추가적인 조합 및 순열이 가능함을 알 수 있을 것이다. 따라서 신규한 아키텍처가 첨부된 청구항의 사상 및 범위 내에 모든 이러한 변경, 수정 및 변형을 포함하는 것으로 의도된다. 또한, 용어 "~를 포함하다(include)"가 상세한 설명 또는 청구범위에 사용될 때, 이러한 용어는 용어 "~를 포함하는(comprising)"가 청구범위에서 연결어로서 사용될 때 해석되는 것과 유사한 방식으로 포함된 것으로 의도되었다.

도면

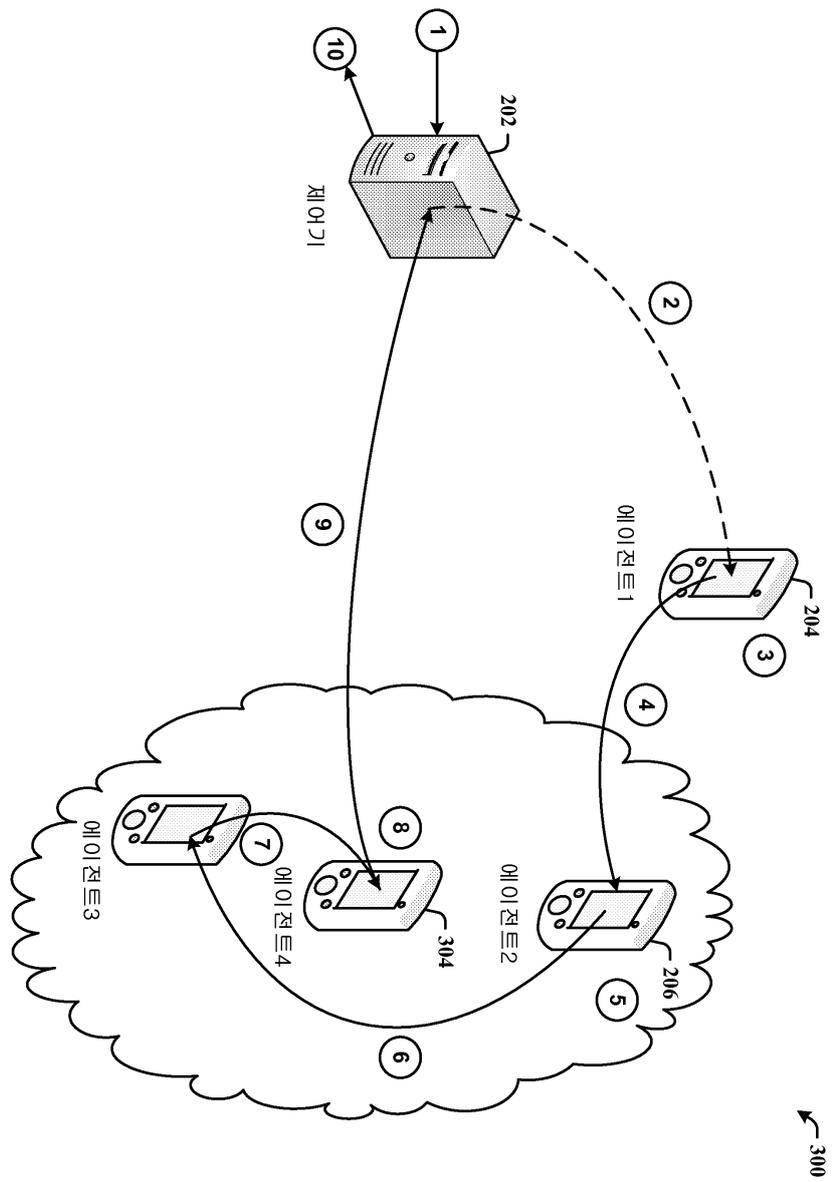
도면1



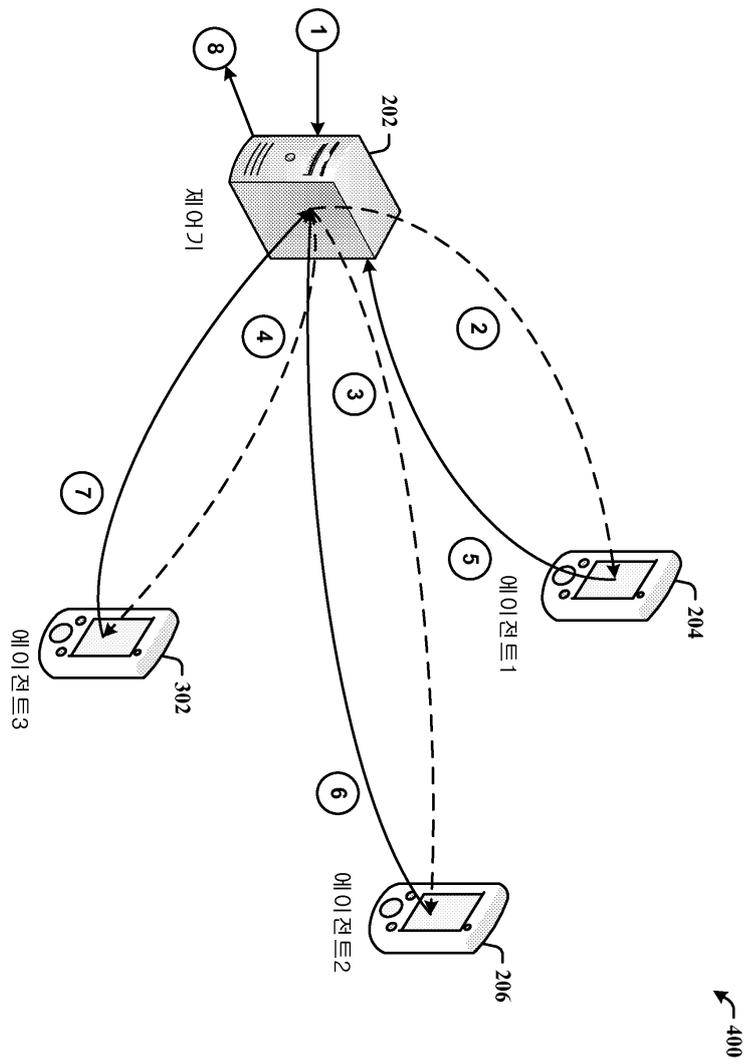
도면2



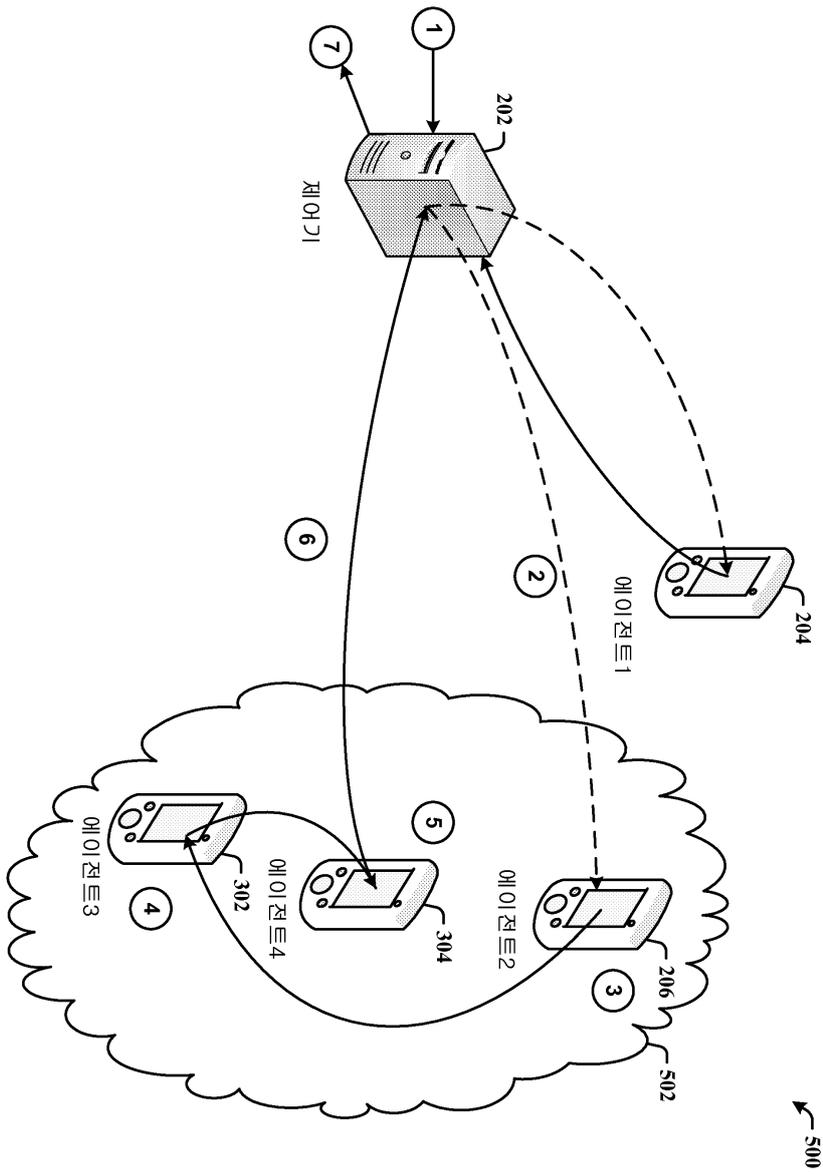
도면3



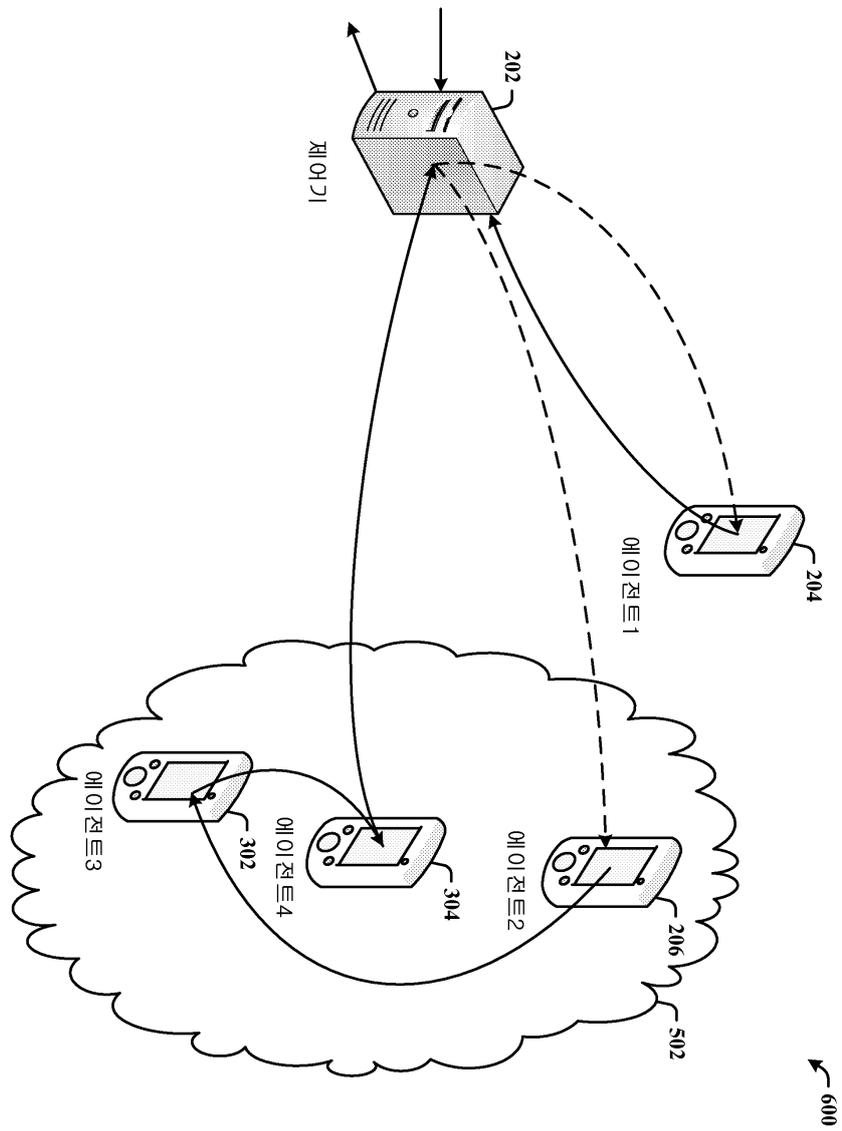
도면4



도면5

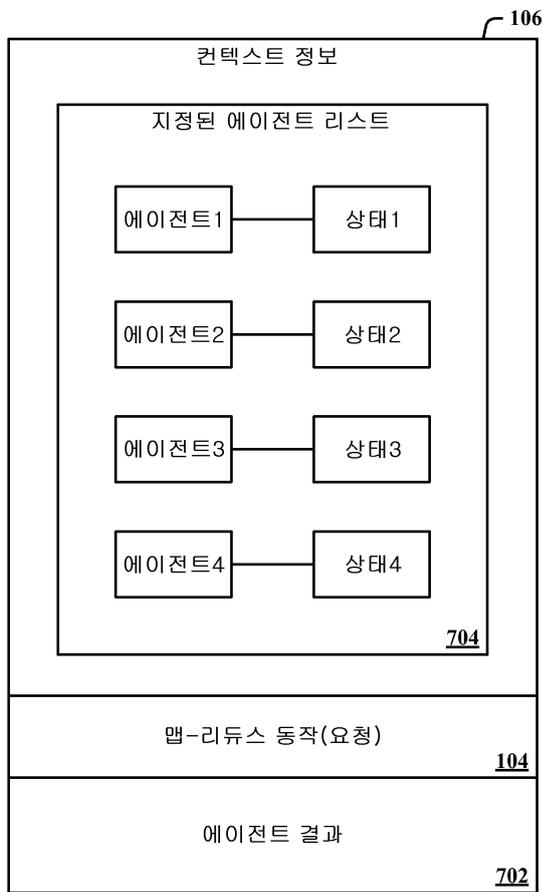


도면6

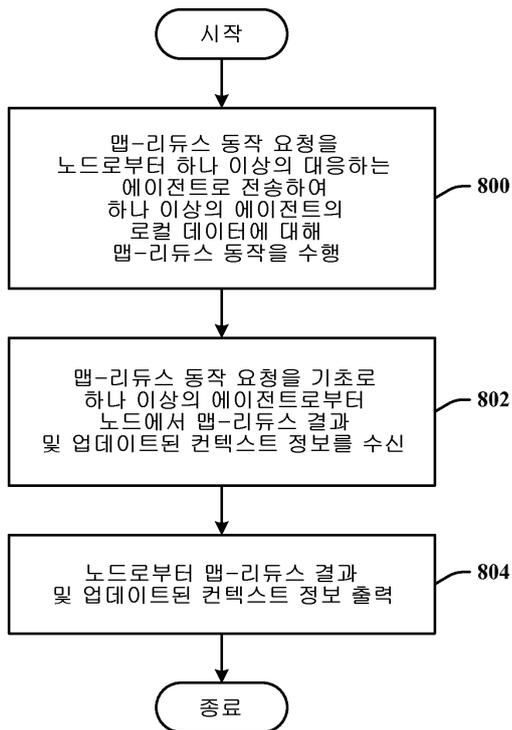


도면7

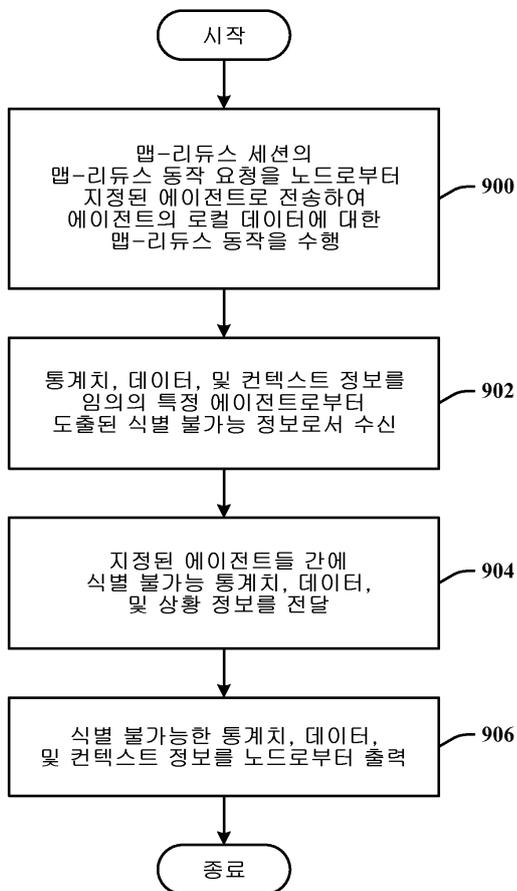
700



도면8



도면9



도면10

