



Università degli Studi di Napoli "Parthenope"

DIPARTIMENTO DI SCIENZE E TECNOLOGIE

Corso di laurea in Corso di Laurea in Informatica

Tesi di Laurea Triennale

**PROGETTAZIONE E SVILUPPO DI UN
SISTEMA DISTRIBUITO PER L'ANALISI
DELLE TRANSAZIONI BITCOIN**

Relatore

prof. Alessio Ferone

Laureando

Antonio Riccardi
mat. 0124000411

Anno accademico 2018 - 2019

Ai miei genitori.

Ringraziamenti

TODO
Grazie a soreta.

Sommario

La seguente tesi ha come obiettivo la realizzazione di un'applicazione distribuita in grado di fare analisi di transazioni provenienti dalla blockchain di Bitcoin.

Il nocciolo della questione affrontata è quello di studiare la struttura di ogni singola piattaforma distribuita e di implementare componenti in grado di ricavare informazioni utili per l'analisi delle transazioni. In particolare, ci si soffermerà sulle principali tecnologie informatiche che permettono di costruire sistemi distribuiti, come Spark ed Hadoop, e sulle strutture dati che facilitano l'immagazzinamento delle informazioni come Neo4j.

Nella prima parte della tesi si procede con l'introduzione alla tecnologia alla base della moneta virtuale Bitcoin, facendo luce su cos'è la Blockchain e quali vantaggi/svantaggi presenta rispetto ad altri sistemi che gestiscono transazioni monetarie. A seguire, sono presi in esame le principali soluzioni di analisi di transazioni già esistenti.

Il secondo capitolo dell'elaborato mostra una panoramica del disegno dell'architettura del sistema realizzato, mettendo in luce i principali componenti del sistema. A tal fine è presente un'ampia e dettagliata descrizione di ogni singolo componente del sistema specificando i vantaggi ottenuti dal proprio utilizzo.

Il terzo capitolo descrive come è stata fatta l'implementazione del sistema, mostrando porzioni di codice e screen dell'applicazione che ci aiutano nella comprensione di come funziona il sistema.

Infine nell'ultimo capitolo vengono trattate le considerazioni finali del lavoro svolto soffermandosi anche sui possibili sviluppi.

Indice

Elenco delle figure	6
Elenco delle tabelle	7
Elenco dei listati	8
1 Bitcoin: Fonte di bigdata	9
1.1 Blockchain	9
1.2 Bitcoin	11
1.2.1 Come avviene una transazione	12
1.2.2 Come avviene una transazione dal punto di vista tecnico	13
1.3 Analisi dati su sistemi distribuiti	17
1.3.1 Caratteristiche di un sistema distribuito	18
1.3.2 Vantaggi e Svantaggi	18
1.4 Stato d'arte	19
2 Overview e progettazione di sistema	20
2.1 Scopo del progetto	20
2.2 Architettura del progetto	20
2.3 Sistema distribuito	22
2.3.1 Bitcoind	23
2.3.1.1 ZeroMQ	24
2.3.2 Apache Spark	25
2.3.2.1 Spark Streaming	27
2.3.2.2 GraphX	28
2.3.3 Hadoop HDFS	29
2.3.4 Neo4j	32
2.3.5 Zookeeper	36
2.3.5.1 Kafka	38
2.4 WebApp	39
2.4.1 Node.js	41
2.4.1.1 Express.js ed Handlebars	43
2.4.1.2 WebSocket	45
2.4.1.3 MaterialCSS	45
2.4.1.4 D3js	45

3	Implementazione	46
3.1	Diagramma delle classi	46
3.2	Codice	46
3.3	Github	46
4	Conclusioni e sviluppi futuri	47
	Riferimenti bibliografici	50

Elenco delle figure

1.1	Rete distribuita di nodi paritari.	10
1.2	Catena di blocchi nella blockchain.	10
1.3	Logo Bitcoin.	12
1.4	Come avviene una transazione Bitcoin.	13
1.5	Azioni del ricevente.	14
1.6	Prima parte. Azioni del pagante.	14
1.7	Seconda parte. Azioni del pagante.	15
1.8	Azione dei miner.	15
1.9	Merkle tree.	16
1.10	Differenza tra sistema distribuito e centralizzato.	17
2.1	Architettura completa.	21
2.2	Architettura in dettaglio del sistema distribuito.	23
2.3	Messaggio inviato con la modalità Publish-Subscribe.	25
2.4	Visualizzazione di un RDD partizionato.	26
2.5	Infrastruttura Spark.	27
2.6	Come vengono gestiti i dati in Spark Streaming.	28
2.7	Logo GraphX.	28
2.8	Visualizzazione invio dati ad HDFS.	30
2.9	Architettura HDFS.	32
2.10	Interfaccia web di Neo4j nella quale è eseguita una query.	34
2.11	Funzionamento di Zookeeper.	37
2.12	Come funziona Kafka	39
2.13	Architettura in dettaglio della webapp.	41
2.14	Gestione eventi con Node.js.	42
2.15	Visualizzazione della catena di funzioni per rispondere ad una richiesta	44

Elenco delle tabelle

Elenco dei listati

2.1	Query Cypher.	35
2.2	Esempio di HTML scritto con Handlebars.	44
2.3	Esempio di variabile passata dal controller al template engine.	45

Capitolo 1

Bitcoin: Fonte di bigdata

Nel lontano 2009 uno pseudonimo Satoshi Nakamoto pubblicò il manifesto *"Bitcoin: A Peer-to-Peer Electronic Cash System"*[21] sancendo la nascita di una nuova moneta digitale chiamata Bitcoin, che fa del suo punto di forza la libertà e la sicurezza. Nel corso degli anni il valore della moneta è accresciuto sino a raggiungere picchi di 19290 dollari nel 17 Dicembre 2018[6]. Di pari passo, anche il numero di transazioni giornaliere sulla rete bitcoin ha raggiunto cifre esorbitanti riuscendo a validare 5 transazioni al secondo[5], questa enorme mole di dati è possibile processarla solo con sistemi distribuiti che gestiscono i BigData. Non a caso l'elaborato di tesi, utilizza un sistema distribuito creato ad hoc per i bigdata come Spark ed Hadoop. I quali permettono l'elaborazione parallela e distribuita di grosse quantità di dati.

Bitcoin, oltre all'impatto economico, ha avuto un importante ruolo nel campo della ricerca, donandoci un sistema che gestisce le transazioni tra due parti in un network peer-to-peer, diversamente dai sistemi tradizionali, denominato Blockchain.

1.1 Blockchain

La blockchain come dice il nome è una catena di blocchi che implementano un database aperto e distribuito, atto a memorizzare le transazioni tra due parti in modo sicuro, verificabile e permanente. In altre parole, la blockchain rappresenta il libro contabile (o libro mastro), ossia il registro sul quale sono riportati tutti gli scambi tra le parti. Questo libro mastro è distribuito (Distributed Ledger) replicato e sincronizzato tra tutti i membri di una rete. In questo database vengono registrate le transazioni (come lo scambio di beni o informazioni) tra i partecipanti alla rete.

I dati non sono memorizzati su un solo computer ma su più macchine collegate tra loro attraverso una rete peer-to-peer [1.1] sottoforma di blocchi.

Ciascun nodo è autorizzato ad aggiornare e gestire il libro contabile distribuito in modo indipendente, ma sotto il controllo consensuale degli altri nodi. Infatti, gli aggiornamenti non sono più gestiti, come accadeva tradizionalmente, sotto il controllo rigoroso di un'autorità centrale, ma sono invece creati e caricati da ciascun nodo in modo appunto indipendente. In questo modo ogni partecipante è in grado di processare e controllare ogni

transazione ma, nello stesso tempo ogni singola transazione, essendo gestita in autonomia, deve essere verificata, crittografata e approvata dalla maggioranza dei partecipanti alla rete.



Figura 1.1: Rete distribuita di nodi paritari.

Ogni blocco contenuto nella blockchain archivia un insieme di transazioni validate correlate da un Marcatore Temporale (Timestamp) ed un hash (una stringa alfanumerica) che identifica il blocco in modo univoco e che permette il collegamento con il blocco precedente[1.2].

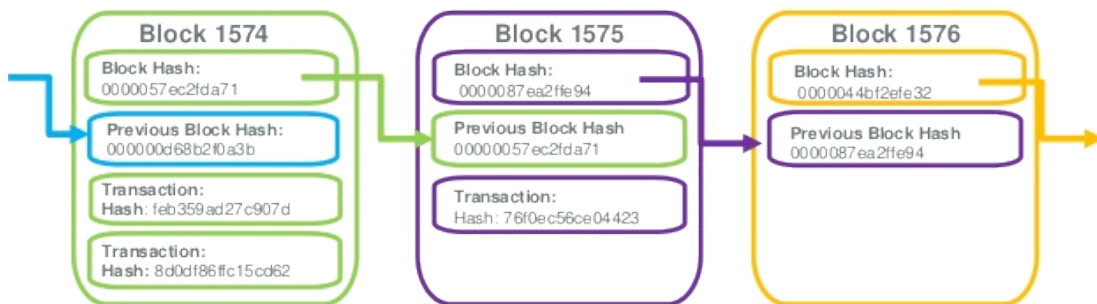


Figura 1.2: Catena di blocchi nella blockchain.

Perché un nuovo blocco di transazioni sia aggiunto alla Blockchain è necessario appunto che sia controllato, validato e crittografato. Solo con questo passaggio può poi diventare attivo ed essere aggiunto alla Blockchain. Per effettuare questo passaggio è necessario che ogni volta che viene composto un blocco venga risolto un complesso problema matematico che richiede un cospicuo impegno anche in termini di potenza e di capacità elaborativa. Questa operazione viene definita come “Mining” ed è svolta dai “Miner”. La risoluzione del problema è un processo irreversibile per cui il blocco valido aggiunto alla

catena diviene immutabile.

In conclusione, questa tecnologia può essere utilizzata in tutti gli ambiti in cui è necessaria una relazione tra più persone o gruppi. Ad esempio può garantire il corretto scambio di titoli e azioni o sostituire un atto notarile, perché ogni transazione viene sorvegliata da una rete di nodi che ne garantiscono la correttezza senza l'ausilio di intermediari.

Bitcoin sfrutta questa tecnologia per gestire le transazioni finanziarie in modo sicuro e del tutto autonomo.

1.2 Bitcoin

Bitcoin (codice: BTC o XBT) è una criptovaluta e un sistema di pagamento mondiale creato nel 2009 da un anonimo inventore, noto con lo pseudonimo di Satoshi Nakamoto, che sviluppò un'idea da lui stesso presentata su Internet a fine 2008. Per convenzione se il termine Bitcoin è utilizzato con l'iniziale maiuscola si riferisce alla tecnologia e alla rete, mentre se minuscola (bitcoin) si riferisce alla valuta in sé.[14]

Il progetto Bitcoin è nato con l'intento di risolvere i problemi di fiducia, trasparenza e responsabilità tra due parti nello scambio di denaro per beni e servizi su Internet, senza l'utilizzo di intermediari. Bitcoin, infatti, rappresenta la prima rete di pagamento che utilizza una tecnologia distribuita peer-to-peer per operare senza un'autorità centrale: la gestione delle transazioni e l'emissione di denaro sono effettuati collettivamente dalla rete.

La rete Bitcoin utilizzando la tecnologia blockchain, è una catena di blocchi concatenati tra di loro, infatti ogni blocco contiene il riferimento al hash del blocco precedente. Questo sistema è immutabile perché se si volesse modificare un blocco si dovrebbero modificare tutti i blocchi successivi ad esso, per fare ciò si dovrebbe creare una catena di blocchi che è più lunga di quella esistente. Ovviamente, la distribuzione e la natura delle tempistiche del processo rendono praticamente impossibile che ciò accada.

Inoltre, utilizzando un sistema distribuito, permette di tenere traccia di tutti i trasferimenti in modo da evitare il problema del double spending (spendere due volte). Tutti gli utenti sono a conoscenza di ciò che accade e pertanto non c'è bisogno di una autorità centrale che gestisca le transazioni.

In aggiunta, Bitcoin consente il possesso e il trasferimento anonimo delle monete, ed infatti i dati necessari per usufruire dei propri bitcoin sono salvati in uno o più personal computer sotto forma di digital wallet (portafoglio). I bitcoin sono trasferiti tramite Internet a chiunque disponga di un indirizzo bitcoin. La struttura peer-to-peer della rete e l'assenza di un ente centrale rende impossibile a qualsiasi autorità il blocco dei trasferimenti, il sequestro dei bitcoin senza il possesso delle relative chiavi o la svalutazione dovuta all'immissione di nuova moneta. Per questo motivo i bitcoin sono la moneta più utilizzata nel Deep Web.

Ricapitolando, le principali caratteristiche su cui si basa questo sistema di enorme successo sono le seguenti:

- **Nessuna autorità centrale:** non dipende da nessuna terza parte privata o ente governativo, e, pertanto, il valore dei BTC è liberamente contrattato sul mercato.

Si tratta quindi di un sistema decentralizzato;

- **Irreversibile e non falsificabile:** una volta che una transazione è stata effettuata ed è inclusa nella blockchain, non può più essere annullata, nemmeno dal mittente;
- **Anonimo:** chiunque può scaricare il software ed iniziare ad effettuare transazioni, senza registrarsi, senza comunicare dati personali e senza svelare la propria identità.
- **Sistema distribuito su rete Peer-to-peer (rete paritaria o paritetica):** qualsiasi nodo è in grado di avviare e completare una transazione in modo autonomo.
- **Inflazione determinata a priori:** L'emissione di nuovi BTC è determinata dall'algoritmo stesso del programma, e non può essere modificata.

Per capire meglio questi punti, verrà illustrata in modo dettagliato una transazione tra due utenti Alice e Bob che vogliono scambiarsi bitcoin.



Figura 1.3: Logo Bitcoin.

1.2.1 Come avviene una transazione

Un esempio pratico di scambio di moneta aiuterà nel capire come funziona il sistema. I nostri attori saranno Alice (pagante) e Bob (ricevente).

I passi da eseguire sono:

1. Bob comanda alla sua applicazione (wallet) su PC o smartphone di creare un indirizzo. Il software restituisce una sequenza alfanumerica da 26 a 35 caratteri. Questo è un esempio: 12gXGyXWkvyDAjVKZHyGGstVYyXJ6ZjgqV
2. Bob copia l'indirizzo e lo mostra al mittente tramite qualunque mezzo: via mail, messaggio, QR code, pubblicato su un sito internet, scritto su carta, dettato al telefono etc.
3. Alice inserisce nel suo software l'indirizzo di Bob e la quantità di Bitcoin da inviare, inoltre specifica l'ammontare della commissione da pagare al minatore (detto "miner") per convalidare la transazione. Ad oggi i più comuni software stabiliscono automaticamente una commissione fissa a 0,00001 bitcoin.
4. Bob vede immediatamente sul suo software la transazione avvenuta, ma prima di considerare il pagamento effettuato attende che la transazione sia inserita nella blockchain.
5. In un massimo di 10 minuti circa la transazione viene inserita in un blocco della blockchain da parte del miner, che ha convenienza a inserirla perché in questo modo ottiene la commissione pagata da Alice.

6. Bob può sentirsi sicuro che la transazione è confermata con l'aumentare di blocchi che vengono aggiunti alla blockchain conseguentemente a quello che contiene la transazione.
7. Bob avrà nel proprio portafoglio i bitcoin inviati da Alice per spenderli in una nuova transazione.



Figura 1.4: Come avviene una transazione Bitcoin.

1.2.2 Come avviene una transazione dal punto di vista tecnico

La stessa transazione affrontata con un occhio più tecnico, soffermando l'attenzione sulla parte di creazione delle chiavi e della conferma ed aggiunta del blocco.

I passi da eseguire sono:

1. Tramite il software che si occupa del wallet, Bob genera in modo del tutto casuale una chiave privata, che viene salvata sul suo computer.
2. La chiave privata viene convertita in una chiave pubblica tramite un procedimento matematico. Comunemente è il software di Bob a generare automaticamente la chiavi quando Bob chiede all'applicazione di creare un indirizzo da comunicare al mittente. In realtà Bob potrebbe utilizzare una chiave privata facilmente ricordabile, come "sum qui sum", e ricavare Public key e indirizzo da questa. Il procedimento matematico si basa su un algoritmo chiamato "Elliptic Curve Digital Signature Algorithm"[10] che utilizza la curva ellittica.
È teoricamente possibile ma statisticamente impraticabile scoprire la chiave privata partendo da quella pubblica: poiché il procedimento matematico applicato è unidirezionale, il processo inverso per indovinare la chiave privata richiederebbe una quantità di tentativi e una potenza di calcolo talmente enorme da essere al di là di ogni possibilità
3. La chiave pubblica viene a sua volta crittografata e accorciata tramite un hash. Possiamo chiamare la nuova chiave pubblica Public Key Hash. La chiave pubblica originaria invece è detta Full Public Key.
4. L'hash della chiave pubblica viene convertito in una riga di massimo 35 caratteri (per comodità pratica), che costituisce l'indirizzo del portafoglio di Bob. Per esempio: 12gXGyXWkvyDAjVKZHyGGstVYyXJ6ZjqV.
5. Bob spedisce l'indirizzo ad Alice.

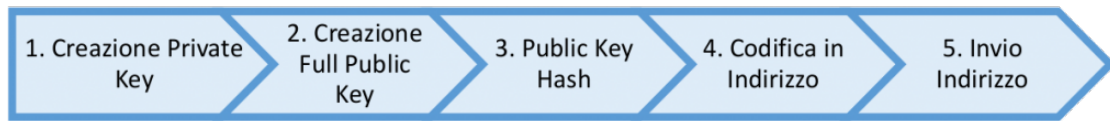


Figura 1.5: Azioni del ricevente.

6. Il software di Alice decodifica immediatamente l'indirizzo in una normale Public Key Hash
7. Alice crea la transazione. Si può pensare la transazione come un codice che contiene diverse informazioni, ciascuna rappresentabile come una stringa composta da molti caratteri:
 - l'input: uno o più output di una transazione precedente fatta nei confronti di Alice, da cui ella attinge i bitcoin che «spedisce» nel nuovo output
 - l'output: la quantità di bitcoin spediti. Possono esserci più output per ogni transazione, ciascuno identificato con un ID specifico
 - l'istruzione per la firma (la "signature script"): ovvero le istruzioni che Bob dovrà fornire per convalidare la transazione, dimostrando di essere il possessore del nuovo output. È proprio per la creazione dello script che il software di Alice ha bisogno del Public Key Hash fornito da Bob. Le informazioni necessarie per validare la firma sono due, entrambe già in possesso di Bob: la full Public Key e la Private Key, che dovranno combaciare col Public Key Hash specificato da Alice nello script.

Queste informazioni vengono processate insieme nella creazione di un unico hash chiamato txid (transaction identifier)



Figura 1.6: Prima parte. Azioni del pagante.

8. Tutti i bitcoin che Alice ha a disposizione su un particolare input vengono "spediti" nella transazione. Infatti nella transazione è coinvolta sempre l'intera quantità di bitcoin presenti nell'input anche se Bob ne ha richiesti molti meno. Se Alice dispone di un input di 100 bitcoin e ne trasferisce 20 a Bob, l'input è sempre trasferito nella sua interezza di 100 bitcoin. In questo caso avrà due output diversi, uno di 80 bitcoin (al lordo della commissione per il miner) che tornano al portafoglio di Alice (il change output), l'altro di 20 bitcoin che vanno all'indirizzo di Bob. L'unico caso di transazione che abbia un solo input e un solo output è quello in cui l'input corrisponde esattamente all'ammontare richiesto da chi riceve i bitcoin. Spesso le

transazioni hanno più output, e quindi i bitcoin trasferiti vanno ad indirizzi con diverse chiavi pubbliche e private.

9. Alice trasmette via internet al software di tutti gli altri nodi tutte le informazioni relative alla transazione. I nodi sono rappresentati da tutti coloro che hanno il software Bitcoin Core sui propri pc/dispositivi o numerosi altri software che permettono di "collegarsi" al network Bitcoin.



Figura 1.7: Seconda parte. Azioni del pagante.

10. I minatori inseriscono le transazioni ancora non confermate nella blockchain.
Per inserire le transazioni all'interno della blockchain il miner deve creare un nuovo blocco, processo che richiede una quantità di calcolo molto elevata e dunque una spesa in energia elettrica e strumenti. Un miner ha interesse a inserire quante più transazioni nel blocco che vuole creare poiché guadagnerà tutte le commissioni pagate su ciascuna transazione. Se una transazione non include alcuna commissione, il miner non ha alcun interesse economico nell'inserirla nel blocco.

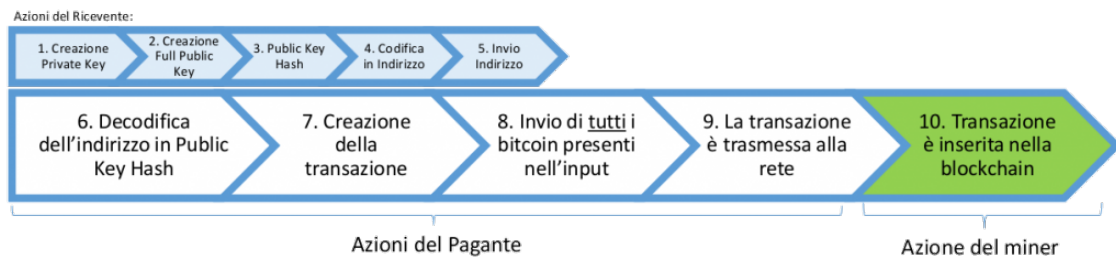


Figura 1.8: Azione dei miner.

Per inserire le transazioni nel blocco, i minatori partono dagli id delle transazioni (txid), ciascuno dei quali rappresenta l'hash di tutte le informazioni inerenti una singola transazione (passo 7).

I txid sono accoppiati due a due, creando un hash per ogni coppia di transazioni. Ogni hash viene poi accoppiato con un altro hash, creando un hash figlio dei due hash precedenti, e così via finché non si arriva a un unico hash. In caso di numeri dispari un hash viene processato con una sua copia identica. Questo procedimento può essere rappresentato come un albero, il cosiddetto Merkle tree [1.9], dove le foglie sono le transazioni txid, i rami (biforcute) gli hash intermedi e la radice l'hash finale, prodotto di tutti gli altri hash: la Merkle root. L'hash finale è come l'ultimo di una stirpe e porta con sé il "DNA" di tutti gli hash precedenti[11].

Grazie alla struttura del Merkle tree, non è necessario conoscere tutte le transazioni

includere in un blocco per verificare che una singola transazione ne faccia parte, è invece sufficiente seguire un particolare ramo che collega una foglia (una transazione) alla merkle root.

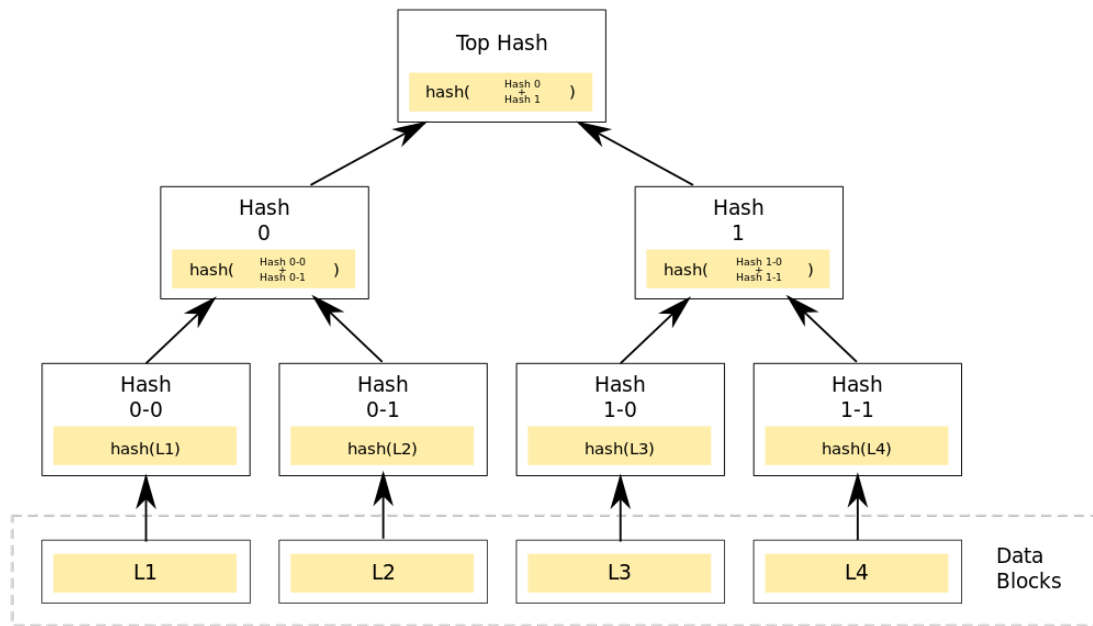


Figura 1.9: Merkle tree.

11. Bob ora vuole spendere i suoi nuovi bitcoin in una nuova transazione, il destinatario è Charlie. Bob segue la stessa procedura che ha seguito Alice, creando una transazione specificando output, signature script (per cui gli serve il public key hash di Charlie), timestamp e versione del software.

Bob però deve dimostrare di essere il possessore dei bitcoin che invia a Charlie, ovvero i bitcoin presenti nell'input della transazione. Tale input è anche l'output della transazione fra Alice e Bob. Quest'ultimo per dimostrare di possedere l'output di quella transazione deve porre la sua firma (signature). Bob inserisce quindi la sua Full Public Key, verificando che corrisponde al Public Key Hash dato in precedenza ad Alice, e la sua Private Key, che rappresenta la conferma che Bob solo è la persona che ha originato inizialmente quella Public Key. Infatti seppur sia teoricamente possibile, è per motivi statistici «infattibile» scoprire la Private Key partendo dalla Public Key. Il procedimento di firma è del tutto automatizzato dal software Bitcoin[26].

1.3 Analisi dati su sistemi distribuiti

Il numero di utenti che ogni giorno spende bitcoin è in costante crescita. La tecnologia Bitcoin quindi è costretta a lavorare con volumi di dati sempre crescente. Per questo motivo, le applicazioni che fanno analisi devono adattarsi scegliendo sistemi consoni a queste moli di dati: i Sistemi distribuiti.

Esistono più definizioni (più o meno equivalenti fra loro) di un sistema distribuito, fra cui:

- "Un sistema distribuito è una porzione di software che assicura che un insieme di calcolatori appaiano come un unico sistema coerente agli utenti del sistema stesso" (Maarten van Steen, 2016).[1]
- "Un sistema distribuito consiste di un'insieme di calcolatori autonomi, connessi fra loro tramite una rete e un middleware di distribuzione, che permette ai computer di coordinare le loro attività e di condividere le risorse del sistema, in modo che gli utenti percepiscano il sistema come un unico servizio integrato di calcolo" (Wolfgang Emmerich, 1997).[8]
- "Un sistema distribuito è un sistema in cui il fallimento di un computer che non sapevi neppure esistere può rendere il tuo computer inutilizzabile" (Leslie Lamport, 1987).[19]

Sintetizzando, un sistema distribuito è un insieme di processori indipendenti (con proprie risorse hardware/software) interconnessi da una rete di comunicazione, che cooperano per condividere alcune delle risorse ovunque distribuite ed eseguire algoritmi parallelamente. Questi sistemi riescono ad apparire all'utente come un singolo sistema, permettendo di avere una certa estrazione dall'hardware dei singoli nodi.

I sistemi distribuiti [1.10a] si contrappongono ai sistemi centralizzati, nella quale tutto il lavoro è eseguito da un solo calcolatore [1.10b]. Un esempio di sistema distribuito è la rete Internet stessa, che si estende a livello mondiale comprendendo risorse fisicamente molto distanti tra loro, in cui processi con funzioni diverse e connessi da reti di vario tipo si scambiano messaggi informativi basati su disparati protocolli di comunicazione.

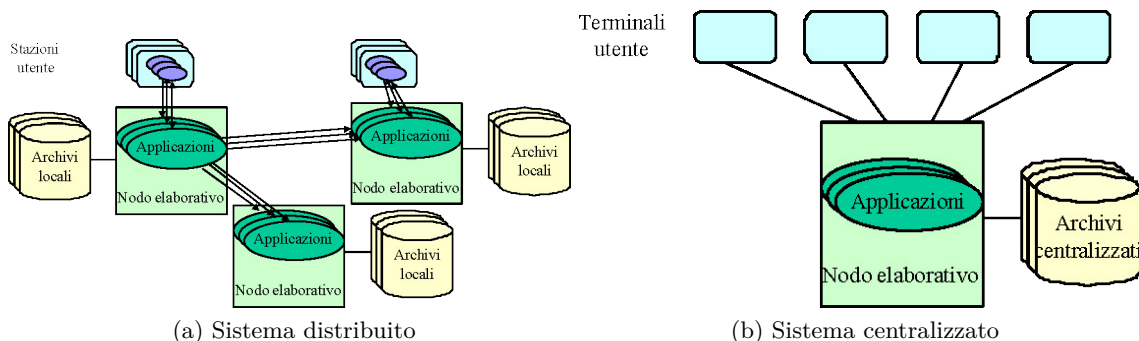


Figura 1.10: Differenza tra sistema distribuito e centralizzato.

L'applicazione creata è eseguita su un sistema distribuito questo perché abbiamo bisogno di tempi di risposta bassi ed alta affidabilità dei dati. Per questo motivo il codice viene partizionato in piccoli problemi ed eseguiti dai nodi del nostro cluster. Il processo di partizione dei dati viene automatizzato dal framework Spark [2.3.2] che si occupa della gestione, invio e recupero dei dati tra i vari nodi.

1.3.1 Caratteristiche di un sistema distribuito

Un sistema distribuito si definisce tale se rispetta alcune caratteristiche come:

- **Remoto:** le componenti di un sistema distribuito devono poter essere trattate allo stesso modo sia che siano in locale che in remoto.
- **Concorrenza:** è possibile eseguire contemporaneamente due o più istruzioni su macchine differenti.
- **Malfunzionamenti parziali:** Ogni componente del sistema può smettere di funzionare correttamente indipendentemente dalle altre componenti del sistema; questo non deve compromettere le funzionalità dell'intero sistema. I fallimenti che possono affliggere i processi possono essere di varia natura.
- **Eterogeneità:** un sistema distribuito è eterogeneo per tecnologia sia hardware che software. Si realizza in tutti i contesti come rete di comunicazione, protocollo di rete, linguaggi di programmazione, applicazioni, etc.
- **Autonomia:** un sistema distribuito non ha un singolo punto dal quale può essere controllato, coordinato e gestito. La collaborazione va ottenuta inviando messaggi tra le varie componenti del sistema e gestita tramite politiche di condivisione e di accesso che devono essere rigorosamente seguite.
- **Evoluzione:** un sistema distribuito può cambiare sostanzialmente durante la sua vita, sia perché cambia l'ambiente sia perché cambia la tecnologia utilizzata. L'obiettivo è quello di assecondare questi cambiamenti senza costi eccessivi.[18]

1.3.2 Vantaggi e Svantaggi

I sistemi distribuiti hanno dei pro e dei contro che sono da prendere in considerazione quando si vogliono adottare soluzioni di questo tipo.

I vantaggi nell'utilizzo dei sistemi sono:

- **Connettività e collaborazione:** possibilità di condividere risorse hardware e software (compresi dati e applicazioni)
- **Prestazioni e scalabilità:** la possibilità di aggiungere risorse fornisce la capacità di migliorare le prestazioni e sostenere un carico che aumenta (scalabilità orizzontale)
- **Tolleranza ai guasti:** grazie alla possibilità di replicare risorse e dati

- **Apertura:** l'uso di protocolli standard aperti favorisce l'interoperabilità di hardware e software di fornitori diversi
- **Economicità:** i sistemi distribuiti offrono spesso (ma non sempre) un miglior rapporto prezzo/qualità che i sistemi centralizzati basati su mainframe

Gli svantaggi invece sono:

- **Complessità:** i sistemi distribuiti sono più complessi di quelli centralizzati e quindi risultano più difficili da capire, inoltre, lo sviluppo delle applicazioni deve essere implementato ad hoc.
- **Sicurezza:** l'accessibilità in rete pone problematiche di sicurezza
- **Gestibilità:** è necessario uno sforzo maggiore per la gestione del sistema operativo e delle applicazioni

1.4 Stato d'arte

Capitolo 2

Overview e progettazione di sistema

In questo capitolo viene messo in luce lo scopo dell'elaborato realizzato con particolare attenzione alle scelte progettuali adottate e alle ricerche effettuate per implementare l'argomento trattato, onde presentare un quadro generale completo. Viene, inoltre, descritta l'architettura del software, sottolineando ed approfondendo le sue principali componenti.

2.1 Scopo del progetto

Lo scopo principale del lavoro di tesi è quello di creare un sistema distribuito che permetta la visualizzazione real time, la storicizzazione e l'analisi delle transazioni provenienti dalla blockchain di bitcoin. L'obiettivo, dunque, è quello di riuscire a creare un sistema che gestisca grandi quantità di dati in un ambiente distribuito, garantendo affidabilità e consistenza dei dati anche in caso di guasti.

Il sistema si rivolge ad un pubblico che vuole fare analisi delle transazioni processate dalla rete bitcoin. L'utente infatti, può controllare in real time le ultime transazioni elaborate, monitorare l'intera rete blockchain per risalire a tutta la catena di transazioni, oppure controllare gli hash (indirizzi) che hanno avuto maggior punteggio di PageRank.

Le ultime transazioni elaborate, hanno una serie di informazioni di dettaglio come il blocco di appartenenza, l'hash che ha generato quella transazione, il timestamp ed i destinatari dei bitcoin. Mentre per quanto riguarda il monitoraggio dell'intera blockchain, l'utente può navigare con l'utilizzo del proprio mouse, in un grafo orientato rappresentante la storia di tutte le transazioni. Inoltre, il fruitore può controllare i nodi con il maggior punteggio di PageRank e localizzarlo all'interno del grafo.

2.2 Architettura del progetto

Le principali scelte di progetto sono state prese coerentemente con lo stato dell'arte e si è tentato di non introdurre nuova complessità al panorama esistente, ricorrendo a tecnologie, protocolli e standard già esistenti ed affermati, senza definirne di nuovi. Quindi, l'architettura dovrà essere sufficientemente generale in modo da poter garantire

nuovi sviluppi ed evoluzioni future e da non comportare l'esclusione a priori di determinate soluzioni e tecnologie.

Il sistema, quindi, si divide logicamente in due moduli:

- **Il Sistema distribuito (back-end):** E' la parte non visibile all'utente. Si occupa del recupero dei dati dalla blockchain di bitcoin, della storicizzazione e della pubblicazione sui topic di kafka. Interamente scritto in Java, comprende Bitcoind, Spark, Hadoop, Neo4j, Zookeeper.
- **Webapp (front-end):** E' la parte visibile all'utente finale. Si occupa della rappresentazione grafica delle transazioni. Scritto in principalmente in Javascript, utilizza la potenza di NodeJS per creare l'interfaccia grafica. Integra nel proprio ecosistema Express Handlebars, WebSocket, MaterialCSS e D3js.

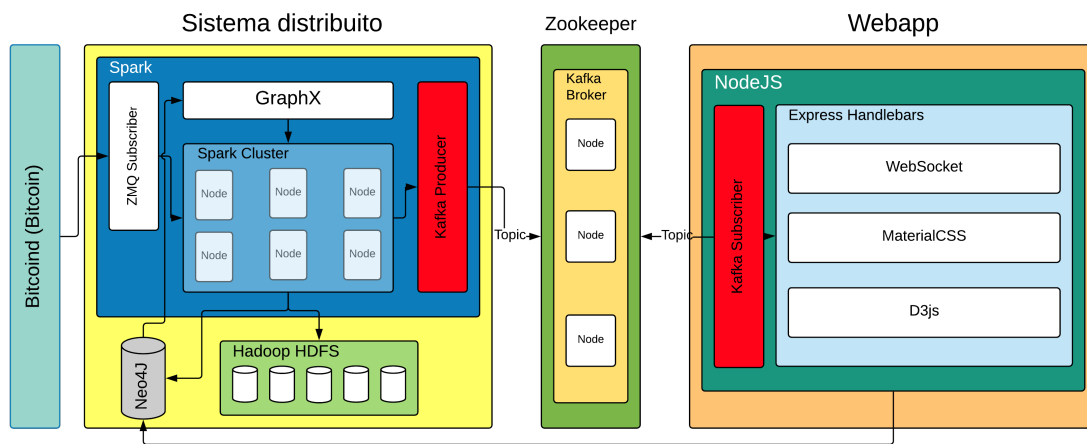


Figura 2.1: Architettura completa.

Per la costruzione di questa infrastruttura, la prima grande sfida, è stata quella di trovare un framework o un tool che permettesse di creare e programmare su di un sistema distribuito, senza complicarci la vita. Facendo ricerche sul web, la tecnologia che più si accostava meglio al mio problema è stata Apache Spark [2.3.2]. Questo strumento riesce a garantire a pieno i vincoli che ci siamo imposti, regalandoci la possibilità di creare un sistema distribuito su vari nodi impostando opportuni file di configurazione.

Risolto il problema infrastrutturale si è proceduto all'analisi dei singoli sottoproblemi. L'applicazione, per testare il carico di lavoro sul sistema, fa uso dei blocchi grezzi provenienti dal software nativo del progetto Bitcoin: Bitcoind [2.3.1]. Bitcoind è un demone che invia blocchi o transazioni (a seconda di come lo si imposta) su di una coda di tipo publisher-subscriber tramite protocollo ZeroMQ [2.3.1.1]. I dati, quindi, sono prelevati da Bitcoind grazie all'implementazione di un connettore creato ad hoc.

Ottenuti i blocchi dalla coda, è nata l'esigenza di conservare i dati ottenuti così da poterli processare ed analizzare. Fortunatamente, Spark offre una nativa collaborazione con il FileSystem distribuito Hadoop [2.3.3], permettendomi di tenerli salvati su una memoria

di massa distribuita.

Oltre ad Hadoop, i dati sono stati immagazzinati in Neo4j [2.3.4]. Un database NoSQL che permette il salvataggio dei dati sottoforma di grafo, così da poter gestire facilmente i collegamenti tra le varie transazioni.

L'ultimo step, è stato quello di fare analisi delle transazioni, trovando i nodi con il maggior PageRank [??]. Anche in questo caso Spark è venuto in contro grazie al modulo GraphX, contenuto nel framework, il quale contiene algoritmi (come il PageRank) già sviluppati per l'analisi sui grafi.

Una volta che il sistema distribuito è completo, non resta che mostrare i risultati ottenuti. Le scelte nel campo del front-end sono migliaia ma per semplicità ed una forte attitudine ai sistemi real-time si è preferito usare NodeJS [2.4.1]. NodeJS ha dei moduli che permettono l'accesso a Kafka, il tramite tra la parte di back-end e front-end. Quindi, con NodeJS è stato creato un sito web consentendo agli utenti dal proprio browser, di visualizzare lo stato delle transazioni, i valori del PageRank e le transazioni che arrivano in real time.

2.3 Sistema distribuito

Come si accennava in precedenza, per lo sviluppo dell'applicazione è stato necessario ricorrere ad un sistema distribuito. Questo sistema, perno di tutta la trattazione, è invisibile all'utente ma di fondamentale importanza. Infatti, è colui che si occupa del processamento dei dati provenienti da Bitcoin, garantendo che nessun guasto infici sul risultato.

Il sistema elabora su più macchine distribuite per ottimizzare i tempi di risposta. In particolare, utilizza un "cluster Spark" (insieme di macchine) per ridistribuire il carico di lavoro equamente sui diversi nodi. Così facendo, ha la possibilità di elaborare grandi quantità di dati in real-time. Inoltre, utilizza la tecnica della replicazione dei dati così che, in caso di rottura di uno dei nodi, si possano recuperare le informazioni perse.

Terminata l'elaborazione dei dati, il cluster invia i risultati ai sistemi dedicati al salvataggio e alla pubblicazione. Due, sono quelli dedicati alla storicizzazione permanente dei dati: Neo4j e Hadoop. Mentre per quanto riguarda la pubblicazione, i dati saranno inviati a Kafka (tramite un producer) che li renderà disponibili per essere fruiti. La scelta dei framework Hadoop e Kafka, è data dal fatto che sono entrambi capaci di lavorare in ambienti distribuiti. Diversa invece è la scelta di Neo4j, l'utilizzo di tale sistema, infatti, è dovuto al fatto che si adatta al meglio con l'idea di transazione, salvando i dati come grafi.

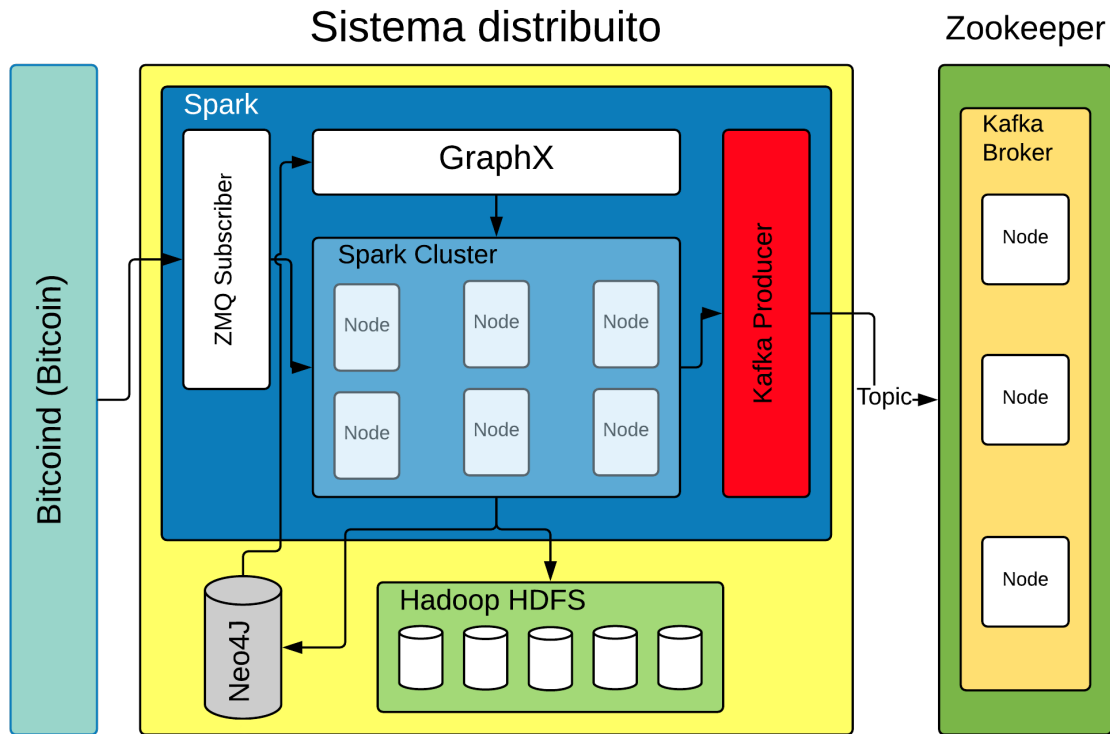


Figura 2.2: Architettura in dettaglio del sistema distribuito.

2.3.1 Bitcoin

Il primo componente del sistema distribuito ha il compito di fornire i dati da elaborare. Questa funzione è svolta dal demone Bitcoin.

Bitcoin, formalmente, è un software che implementa il protocollo Bitcoin per l'utilizzo delle remote procedure call (RPC). Esso è anche il secondo client Bitcoin nella storia del network [9]. Per sua natura, è eseguito come processo in background quindi l'utente per interagire con esso ha bisogno di farlo tramite una interfaccia da riga di comando chiamata *bitcoin-cli*. Il demone, inoltre, funge da nodo della rete Bitcoin, infatti si sincronizza con la blockchain, verifica le transazioni ed invia blocchi. Esiste una versione anche con interfaccia grafica del demone chiamata *Bitcoin-QT* o *Bitcoin Core*, ma per lo scopo dell'elaborato si è preferito utilizzare la versione lite per limitare l'utilizzo di risorse.

La versione demone, inoltre, ha il vantaggio di creare una coda ZeroMQ per la comunicazione con applicazioni esterne. Il sistema distribuito, utilizza questa funzione per recuperare i blocchi in formato grezzo (sequenza di byte) ogni qualvolta sono validati dalla blockchain.

2.3.1.1 ZeroMQ

ZeroMQ (anche conosciuto come ØMQ, 0MQ, o zmq) è una libreria di messagistica asincrona ad alte prestazioni, destinata all'uso in applicazioni distribuite o concorrenti. Fornisce code di messaggi, ma a differenza dei middleware orientati ai messaggi, il sistema ZeroMQ può essere eseguito senza un broker di messaggi dedicato [12]. La libreria, inoltre, funziona molto bene grazie al suo modello interno di threading, che può superare le tradizionali applicazioni TCP in termini di throughput utilizzando una tecnica di batching automatico dei messaggi [28]. Il suo modello I/O asincrono offre la possibilità di creare applicazioni multicore scalabili, costruite come attività asincrone di elaborazione dei messaggi. Inoltre, la comunità di sviluppatori, ha creato una quantità enorme di wrapper per la libreria, così da renderla funzionante sulla maggior parte dei sistemi operativi. All'interno del sistema distribuito, infatti, è stato utilizzato il wrapper *jzmq* che permette l'utilizzo della libreria tramite Java API.

ZMQ, può avere diverse modalità di invio di messaggi atomici:

- Request-reply: Connette un insieme di clienti ad un insieme di servizi. Questo è una Remote Procedure Call (RPC).
- Publish-subscribe: Connette un insieme di produttori ad un insieme di consumatori.
- Push-pull (pipeline): Connette i nodi in un pattern fan-out/fan-in che può avere più passaggi e cicli. Distribuisce in maniera parallela i messaggi.
- Exclusive pair: Collega due socket in maniera esclusiva.

Bitcoin utilizza una socket di tipo Publish-subscribe. Il publisher, in questo caso bitcoin, può spedire un messaggio a molti consumers attraverso un canale virtuale chiamato topic. Uno stesso topic può essere condiviso da più subscriber (client). Per ricevere i messaggi, i client devono "sottoscrivere" al topic creato da bitcoin. Ovviamente, questo implica che c'è relazione temporale tra i publisher e i subscribers, nel senso che, un client che si sottoscrive ad un topic può consumare solamente i messaggi pubblicati dopo la sua sottoscrizione. Qualsiasi messaggio spedito sul topic, viene consegnato a tutti i consumer sottoscritti, ciascuno dei quali riceve una copia identica di ciascun messaggio inviato. I messaggi quindi, vengono automaticamente inviati in broadcast ai consumer, senza che questi ne abbiano fatto esplicita richiesta.

Nel caso del sistema distribuito, il client che si sottoscrive al topic *pubrawblock* di bitcoin è il connettore-subscriber implementato all'interno di Spark.

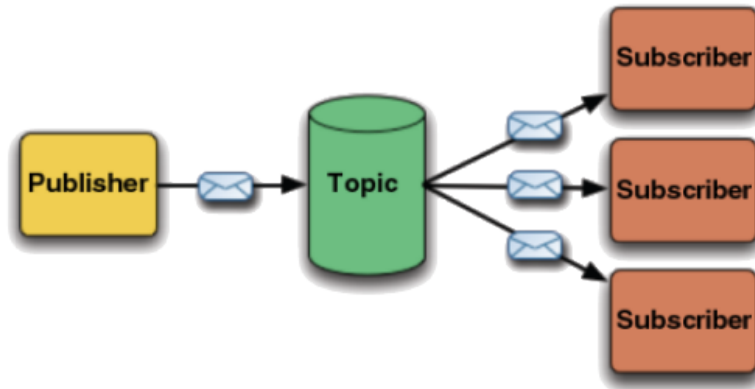


Figura 2.3: Messaggio inviato con la modalità Publish-Subscribe.

2.3.2 Apache Spark

Apache Spark è un framework per la computazione di grandi moli di dati su cluster, progettato e implementato nel 2010 da un gruppo di ricercatori dell'Università di Berkeley a San Francisco [25]. Viene scherzosamente descritto come "quello più veloce di Hadoop" perché nel confronto con il suo predecessore ha prestazioni 10-100 volte maggiori.

Questo progetto nasce dall'esigenza di migliorare le prestazioni dei sistemi distribuiti "MapReduce". Ad alto livello infatti, un'applicazione Spark è formata da un driver program, che contiene la funzione main scritta dall'utente, e di una serie di parallel operation definite nel programma che verranno eseguite sui vari nodi worker che compongono il cluster. Fin qui nulla di nuovo, è il modello del calcolo distribuito master/slave. La vera innovazione è stata introdotta nel modo di definire i dati da elaborare.

Infatti Spark mette a disposizione un'astrazione molto potente, il *resilient distributed dataset* (RDD), che rappresenta una collezione di dati "immutabili" a cui il programmatore si può riferire direttamente tramite l'oggetto associato. Un RDD rappresenta un set di dati che è suddiviso in partizioni (Una tabella chiave-valore suddivisa in tante sottotabelle o un file diviso in tanti segmenti).

L'RDD realizza la fault tolerance grazie all'informazione di lineage: quando una partizione di un RDD si perde a causa di un guasto o di un altro errore, l'RDD ha tutte le informazioni riguardo la storia di quella partizione, in termini di operazioni effettuate su di esso che gli consentono di ricostruirla. La creazione, avviene a partire dai dati su disco (presi da HDFS) o da altre fonti di dati. Una volta creato, un RDD può restare in memoria centrale oppure può essere materializzato su disco.

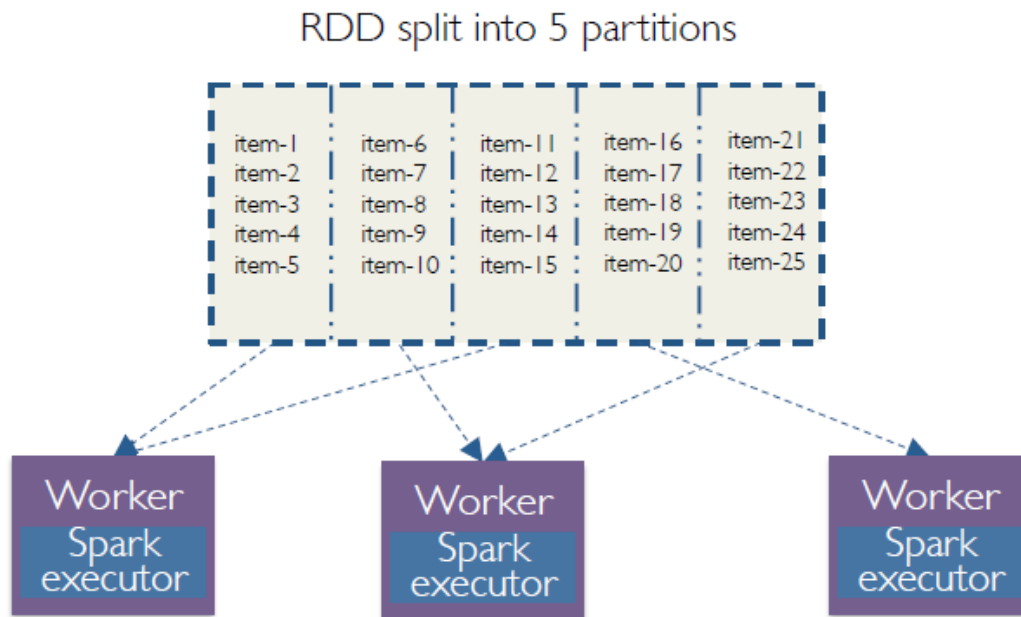


Figura 2.4: Visualizzazione di un RDD partizionato.

Spark nasce come un sistema per creare e gestire job di analisi basati su trasformazioni di RDD. Dato che gli RDD nascono e vivono in memoria, l'esecuzione di lavori iterativi, o che trasformano più volte un set di dati, sono immensamente più rapide di una sequenza di MapReduce; questo perché il disco non viene mai (o quasi mai) impiegato nell'elaborazione.

Il vantaggio principale dell'utilizzo di Spark è la sua estrema velocità nell'eseguire programmi di elaborazione dati. Il motivo di queste prestazioni risiedono in una miglior gestione della memoria; aspetto che lo diversifica da Hadoop. Prendendo come esempio una generica elaborazione fatta con il paradigma MapReduce, in Hadoop questa produce il seguente schema di lavoro, che può essere iterato più volte (semplificando):

- Load dei dati da disco locale verso i nodi worker del cluster;
- Esecuzione della funzione assegnata;
- Store dei dati su disco locale.

Le ripetute fasi di load/store rendono il sistema complessivamente lento. Spark, invece, cerca di mantenere in memoria i dati, esegue le operazioni di trasformazione e solo alla fine memorizzare i dati sul disco.

Come detto in precedenza, Spark non utilizza MapReduce come motore di esecuzione; invece, utilizza il proprio runtime distribuito (DAG) per l'esecuzione di jobs su un cluster. Quando viene invocata un'azione su un RDD, viene creato un "job". Un Directed Acyclic Graph o DAG è un grafo aciclico in cui ogni nodo è una partizione di RDD e ogni vertice è

una trasformazione. A differenza di MapReduce, il motore DAG di Spark può processare pipeline arbitrarie di operatori e tradurle in un unico “job” per l’utente.

Spark, infine, sta dimostrando di essere una buona piattaforma su cui costruire strumenti di analisi, infatti ha moduli per il Machine learning (MLlib), Elaborazione grafica (GraphX), Elaborazione di stream (Spark Streaming) ed SQL (Spark SQL) [25].

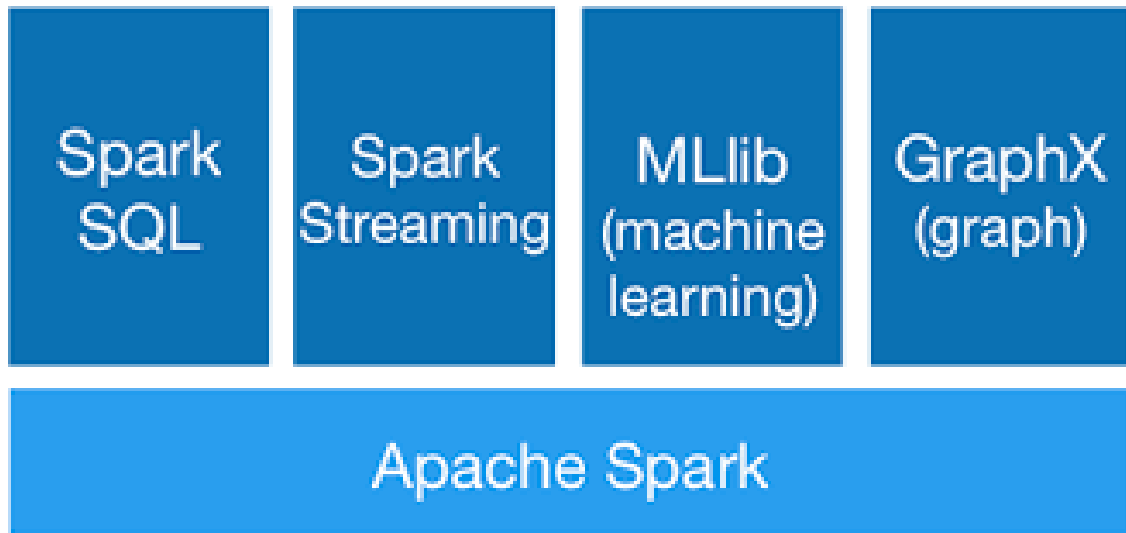


Figura 2.5: Infrastruttura Spark.

2.3.2.1 Spark Streaming

Nel sistema distribuito, poichè c’è l’esigenza di recuperare i dati in real time dalla coda di bitcoind, viene utilizzato Spark Streaming. Questo modulo è un’estensione dell’API Spark di base che consente l’elaborazione streaming, scalabile, ad alto throughput e con tolleranza agli errori dei flussi di dati in tempo reale. I dati, che possono provenire da diverse fonti, sono elaborati utilizzando algoritmi complessi espressi con funzioni di alto livello come *map*, *reduce*, *join* e *window*. I dati processati, infine possono essere inviati a filesystem (Hadoop) o database (Neo4j) per il salvataggio oppure ad altri moduli di Spark, dediti all’analisi, tipo machine learning (MLlib) o graph processing (GraphX).

Internamente, Spark Streaming riceve streams di dati di input e li divide in batch, che vengono quindi elaborati dal motore Spark per generare il flusso finale di risultati. Per consentire il facile utilizzo di questi dati, fornisce un’astrazione di alto livello chiamata stream discretizzato o DStream, che rappresenta un flusso continuo di dati. E’ possibile creare Dstreams da flussi di dati di input da sorgenti come Kafka, Flume e Kinesis o applicando operazioni di alto livello su altri Dstreams [24]. Internamente, un Dstream è rappresentato come una sequenza di RDD sulla quale possono essere effettuate le operazioni descritte in precedenza. Infatti, i Dstream possono essere trasformati usando le operazioni di trasformazione simili a quelle dei RDD, come *map* e *filter*.



Figura 2.6: Come vengono gestiti i dati in Spark Streaming.

2.3.2.2 GraphX

La diffusione dei grafi nei sistemi informatici ha portato a un grande lavoro di analisi su di essi. Facendone un utilizzo sempre più frequente, ci si trova a dover fare ricerche, interrogazioni e misure su questa struttura dati e di trovare un modo per memorizzare efficientemente questi oggetti. Anche Spark si è occupato del problema e ha reso disponibile uno strumento, basato sul framework principale, che ottimizza la gestione dei grafi e consente di applicare ad essi funzioni e metodi in modo molto intuitivo. Si tratta del progetto GraphX. Questo modulo è usato nel sistema distribuito, per fare analisi dei dati provenienti dall'elaborazione di Spark.

GraphX è un nuovo componetene di Apache Spark per grafi ed il calcolo parallelo su di essi. Spark ha introdotto l'RDD, un'astrazione comoda per memorizzare i dati e risparmiando al programmatore parecchio lavoro. GraphX estende il concetto di RDD introducendo il *Resilient Distributed Graph* (RDG). Inoltre, per aiutare nell'analisi, espone un insieme di operatori fondamentali (sottografo, joinVertices e aggregateMessages) come variante ottimizzata dell'API Pregel. In più, Graphx include una crescente collezione di algoritmi e costrutti per grafi per semplificare le attività di analisi [23]. Attualmente le sue API sono scritte in Scala, Java e Python.

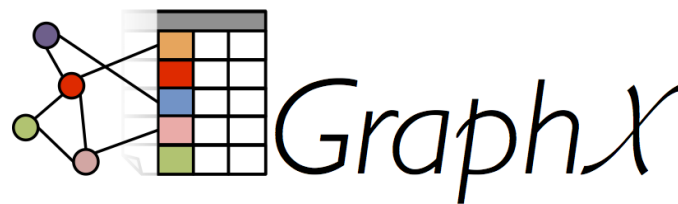


Figura 2.7: Logo GraphX.

GraphX, come detto in precedenza, gestisce i dati in memoria come se fossero grafi. Infatti, utilizza archi e vertici che hanno delle proprietà. Ogni vertice possiede un identificativo univoco a 64bit (VertexID), mentre, allo stesso modo, gli archi contengono gli identificativi di origine e partenza. Queste proprietà sono definite dal programmatore e tenuti in memoria come oggetti. Su di essi si possono eseguire metodi per l'analisi già inclusi nella libreria di GraphX come Connected components, Triangle Counting, Page-Rank, etc. Nell'elaborato di tesi, per trovare i nodi con maggior importanza all'interno

del grafo, è stato usato l'algoritmo *PageRank*.

Il PageRank è un algoritmo di analisi che assegna un peso numerico ad ogni elemento di un collegamento ipertestuale di un insieme di documenti, come ad esempio il World Wide Web, con lo scopo di quantificare la sua importanza relativa all'interno della serie. L'algoritmo di PageRank è stato brevettato (brevetto US 6285999) dalla Stanford University; è inoltre un termine ormai entrato di fatto nel lessico dei fruitori dei servizi offerti dai motori di ricerca. Il nome PageRank è un marchio di Google ed il suo nome si deve a Larry Page [27], uno dei due fondatori di quell'azienda [17].

L'algoritmo completo per il calcolo del PageRank fa ricorso all'uso della teoria dei processi Markov ed è classificato nella vera categoria degli algoritmi di Link Analysis Ranking. Dalla formula inizialmente sviluppata dai fondatori di Google, Sergey Brin e Larry Page, è possibile comprendere come il PageRank viene distribuito tra le pagine:

$$PR[A] = \frac{1-d}{N} + d \left(\sum_{K=1}^n \frac{PR[P_k]}{C[P_k]} \right)$$

Dove:

- **PR[A]** è il valore di PageRank della pagina A che vogliamo calcolare.
- **N** è il numero totale di pagine note.
- **n** è il numero di pagine che contengono almeno un link verso A. P_k rappresenta ognuna di tali pagine.
- **PR[P_k]** sono i valori di PageRank di ogni pagina P_k .
- **C[P_k]** sono il numero complessivo di link contenuti nella pagina che offre il link.
- **d (damping factor)** è un fattore deciso da Google e che nella documentazione originale assume valore 0,85. Può essere aggiustato da Google per decidere la percentuale di PageRank che deve transitare da una pagina all'altra e il valore di PageRank minimo attribuito ad ogni pagina in archivio.

Dalla formula si nota quindi che all'aumentare del numero di link complessivi dei siti che puntano ad A il PageRank aumenta.

2.3.3 Hadoop HDFS

Apache Hadoop è un framework open-source che supporta applicazioni distribuite con elevato accesso ai dati, è uno dei primi sistemi per l'analisi di Big Data, ed è stato considerato un modello da seguire per tutti i sistemi che sono stati creati successivamente. Apache Hadoop è stato ideato per la memorizzazione e la gestione di grandi quantità di dati in parallelo, su cluster di grandi dimensioni (costituiti da migliaia di nodi) assicurando un'elevata affidabilità, scalabilità e disponibilità (fault-tolerant). Tutti i moduli, infatti, sono progettati in modo tale che il software del framework gestisca automaticamente gli eventuali "down" dell'hardware, molto frequenti in un sistema parallelo.

Hadoop nacque per sopperire a un grave problema di scalabilità di Nutch, un crawler

open source basato sulla piattaforma Lucene di Apache. I programmatori Doug Cutting e Michael J. Cafarella hanno lavorato a una versione iniziale di Hadoop a partire dal 2004; proprio in quell'anno furono pubblicati documenti tecnici riguardanti il Google File System e Google MapReduce, documenti da cui Doug e Michael attinsero le competenze fondamentali per lo sviluppo di HDFS e di un nuovo e innovativo pattern per l'elaborazione distribuita di elevate moli di dati, MapReduce, che oggi rappresenta uno dei componenti fondamentali di Hadoop. Circa quattro anni più tardi, nel 2008, nacque la prima release come progetto Open Source indipendente di Apache. A oggi Hadoop è un insieme di progetti tutti facenti parte della stessa infrastruttura di calcolo distribuito. Hadoop offre librerie che permettono la suddivisione dei dati da elaborare direttamente sui nodi di calcolo e permette di ridurre al minimo i tempi di accesso, questo perché i dati sono immediatamente disponibili alle procedure senza pesanti trasferimenti in rete. Il framework garantisce inoltre un'elevata affidabilità: le anomalie e tutti gli eventuali problemi del sistema sono gestiti a livello applicativo anziché utilizzare sistemi hardware per garantire alta disponibilità. Un'altra caratteristica di Hadoop è la scalabilità che è realizzabile semplicemente aggiungendo nodi al cluster in esercizio. I principali vantaggi di Hadoop risiedono nelle sue caratteristiche di agilità e di flessibilità. L'architettura base del framework Hadoop, si compone dei seguenti elementi principali:

- **HDFS:** Il filesystem distribuito di Hadoop, nella quale vengono fisicamente salvati i file.
- **MapReduce:** Un framework per la creazione di applicazioni in grado di elaborare grandi quantità di dati in parallelo basandosi sul concetto di functional programming, ed è, quindi, il principale responsabile del processo di calcolo.
- **Yarn:** Si occupa della schedulazione dei task, ossia delle sotto-attività che compongono le procedure map e reduce eseguite nel processo di calcolo.

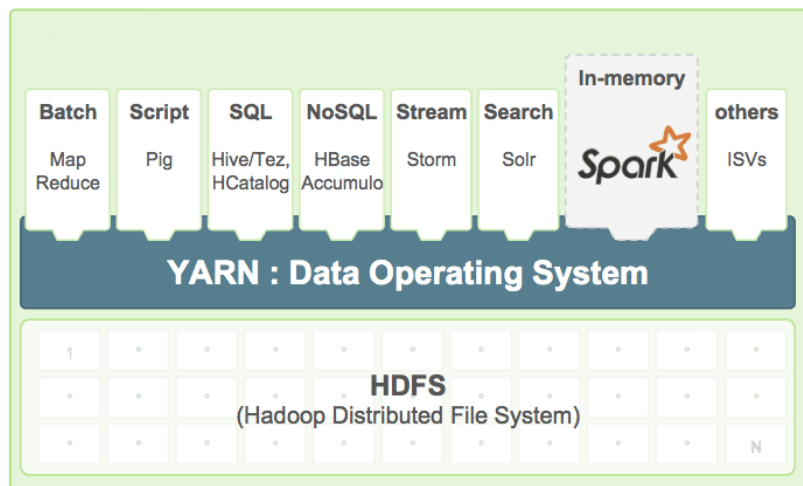


Figura 2.8: Visualizzazione invio dati ad HDFS.

All'interno del sistema distribuito, per il salvataggio veloce e distribuito dei dati su dispositivi fisici, viene utilizzato l'HDFS di Hadoop.

L'HDFS (Hadoop Distributed File System), come riportato nella documentazione ufficiale, è il file system distribuito di Hadoop, progettato appositamente per immagazzinare un'enorme quantità di dati (dell'ordine dei Gigabyte e Terabyte), in modo da ottimizzare le operazioni di archiviazione e accesso a un ristretto numero di file di grandi dimensioni, a differenza dei tradizionali file system che sono ottimizzati per gestire numerosi file di piccole dimensioni; infatti, quando la mole dei dati diventa non più gestibile da una singola macchina, si rende necessario partizionare gli stessi su un certo numero di macchine separate, interconnesse da una rete (cluster), rendendo il filesystem di fatto distribuito. HDFS presenta i file organizzati in una struttura gerarchica di cartelle. Dal punto di vista dell'architettura, un cluster è costituito dai seguenti tipi di nodi:

- **NameNode:** è l'applicazione che gira sul server principale. Gestisce il file system e in particolare il namespace, Inoltre, determina come i blocchi dati siano distribuiti sui nodi del cluster e la strategia di replicazione che garantisce l'affidabilità del sistema. Il NameNode monitora anche che i singoli nodi siano in esecuzione senza problemi e in caso contrario decide come riallocare i blocchi.
- **DataNode:** applicazione/i che girano su altri nodi del cluster, generalmente una per nodo, e gestiscono fisicamente lo storage di ciascun nodo. Queste applicazioni eseguono, logicamente, le operazioni di lettura e scrittura richieste dai client e gestiscono fisicamente la creazione, la cancellazione o la replica dei blocchi dati.
- **SecondaryNameNode:** si tratta di un servizio che aiuta il NameNode a essere più efficiente
- **BackupNode:** è il nodo di failover e consente di avere un nodo simile al SecondaryNameNode sempre sincronizzato con il NameNode.

In HDFS le richieste di lettura dati seguono una politica relativamente semplice [13] avvengono scegliendo i nodi più vicini al client che effettua la lettura e, ovviamente, in presenza di dati ridondati, risulta più semplice soddisfare questo requisito. Inoltre, occorre precisare che la creazione di un file non avviene direttamente attraverso il NameNode. Infatti, il client HDFS crea un file temporaneo in locale e solo quando tale file supera la dimensione di un blocco, è preso in carico dal NameNode. Quest'ultimo crea il file all'interno della gerarchia del file system, identifica un DataNode e i blocchi su cui posizionare i dati. Successivamente DataNode e blocchi sono comunicati al client HDFS che provvede a copiare i dati dalla cache locale alla sua destinazione finale.

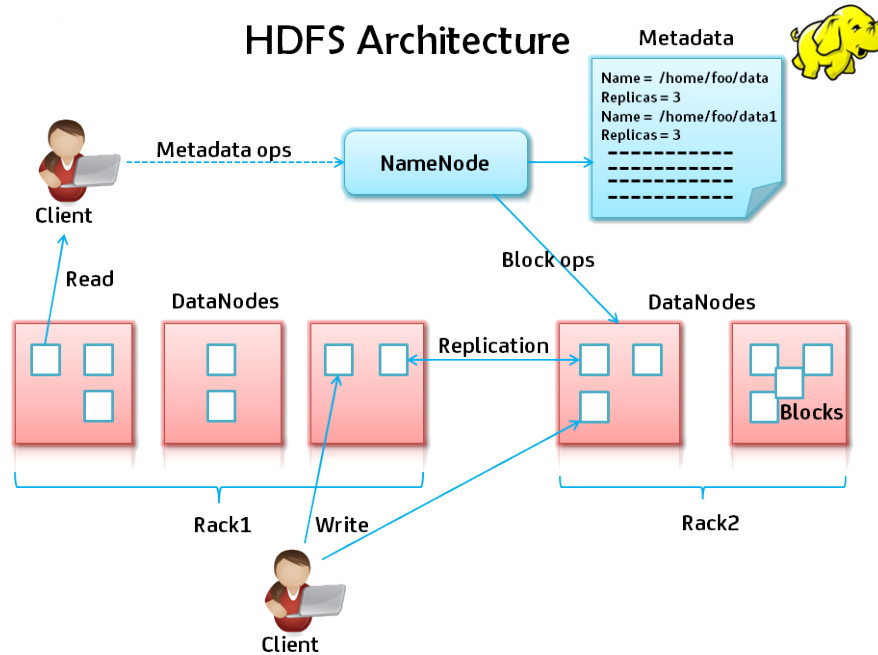


Figura 2.9: Architettura HDFS.

Quanto detto fino a ora, ci permette di concludere che quando vengono trattati file di grandi dimensioni, HDFS è molto efficiente. Invece con file di piccole dimensioni, dove per piccole dimensioni s'intendono dimensioni inferiori al blocco, il trattamento dei file è molto inefficiente, questo perché i file utilizzano spazio all'interno del namespace, cioè l'elenco dei file mantenuti dal NameNode, che ha un limite dato dalla memoria del server che ospita il NameNode stesso.

2.3.4 Neo4j

Negli ultimi anni è incredibilmente aumentata la popolarità delle tecnologie di immagazzinamento di informazioni conosciute con il nome di *NoSQL*, acronimo che sta per Not only SQL [7]. I *NoSQL Database Management System (DBMS)* sono sistemi software che consentono di immagazzinare e organizzare i dati senza fare affidamento sul modello relazionale, solitamente impiegato da database tradizionali.

I NoSQL DBMS sono inoltre contraddistinti dal fatto che non utilizzano un sistema transazionale ACID e spesso sono schema-less [7], ovvero non possiedono uno schema fisso a cui devono attenersi, evitando spesso così le operazioni di join.

Le strutture dei dati memorizzati nei sistemi NoSQL (ad esempio documenti, grafi o coppie chiave/valore) rendono alcune operazioni più veloci rispetto a generici database relazionali. Si utilizzano infatti diversi tipi di database NoSQL in base al problema da risolvere. Le tipologie di database NoSQL più utilizzate sono:

- **Key/Value Store:** rappresentano una tipologia di database NoSql che si basa sul concetto di *associative array*, implementati attraverso HashMap. In un array associativo si hanno un insieme di record che possono essere identificati sulla base

di una chiave univoca. La tipologia di memorizzazione adottata dai key/value stores garantisce tempi di esecuzione costanti per tutte le operazioni applicabili sui dati: *add*, *remove*, *modify* e *find*.

- **Document-oriented database:** questo tipo di database permette di memorizzare in maniera efficiente dati semistruutturati. Essi possono essere considerati una specializzazione dei database key/value, nei quali vengono permesse strutture innestate. Sono spesso usati i formati JSON o XML per la memorizzazione dei dati. MongoDB è il più diffuso database orientato ai documenti ed è uno dei DBMS NoSQL più utilizzati.
- **Column-oriented database:** i DBMS colonnari, a differenza dei tradizionali RDBMS, che memorizzano i dati riga per riga, sfruttano la memorizzazione dei dati per colonna. Per ogni colonna si memorizzano coppie chiave/valore, dove la chiave è l'identificativo di riga ed il valore è il valore associato a quella colonna per la specifica riga. Questa rappresentazione permette di risparmiare una notevole quantità di spazio in caso di sparsità dei dati. I DBMS di tipo colonnare più diffusi sono HBase e Cassandra.
- **Graph database:** un database a grafo utilizza un insieme di nodi con un insieme di archi che li connettono per memorizzare le informazioni, in cui le “relazioni” vengono rappresentate come grafi. Le strutture a grafo si prestano molto bene per la rappresentazione di determinati dati semistruutturati e altamente interconnessi come, ad esempio, i dati dei social network e del Web. I database a grafi più diffusi sono GraphDB, OrientDB e Neo4j.

La struttura dati che più si avvicina all'idea di transazione è quella del grafo, con nodi che rappresentano i mittenti/destinatari dei bitcoin e gli archi che identificano l'ammontare della transazione. Per questo motivo, nel sistema distribuito, si è preferito fare utilizzo del database NoSQL Neo4j per il salvataggio. Inoltre, le performance dei Graph DBMS tendono ad essere ottimali quando i dati da archiviare sono altamente connessi e la mole del dataset è estremamente grande. Questo perché al contrario degli RDBMS (Relational Database Management System), la loro natura gli consente di evitare le onerose operazioni di join semplicemente attraversando le relazioni che connettono i nodi.

Neo4j, dunque, è un software per basi di dati a grafo open source sviluppato interamente in Java. È un database totalmente transazionale, che viene integrato nelle applicazioni permettendone il funzionamento stand alone e memorizza tutti i dati in una cartella. È stato sviluppato dalla Neo Technology, una startup di Malmö, Svezia e della San Francisco Bay Area [15]. Il database può essere usato sia in modalità embedded che server. Nella modalità embedded si incorpora il database nell'applicazione (con maven o includendo i file JAR) e questo viene eseguito all'interno della JVM, quindi nello stesso processo ma accettando vari thread concorrenti. Nella modalità server invece il database è un processo a sé stante a cui si accede tramite REST facendo delle query e ricevendo i dati in remoto. È robusto, scalabile e ad alte prestazioni. È dotato di:

- Transazioni ACID.

- High Availability.
- Può memorizzare miliardi di nodi e relazioni.
- Alta velocità di interrogazione tramite attraversamenti.
- Linguaggio di interrogazione dichiarativo e grafico chiamato Cypher
- DBMS schema-less, ciò sta a significare che i suoi dati non devono attenersi ad alcuna struttura di riferimento prefissata.

La index-free adjancency è alla base delle sue alte prestazioni di attraversamento, d'interrogazione e di scrittura, ed è uno degli aspetti chiave della sua architettura. L'index-free adjancency è una lista (o tabella), ove ogni suo elemento è composto da un nodo del grafo e dai puntatori ai nodi connessi ad esso.

Neo4j, inoltre, salva i dati dentro di una serie di store file, contenuti all'interno di un'unica cartella. Ognuno di questi file contiene al suo interno le informazioni relative ad una singola parte del grafo (e.g. nodi, relazioni, proprietà). Questa separazione della struttura del grafo facilita il suo attraversamento. In Neo4j le unità fondamentali che compongono un grafo, dunque, sono i nodi e le relazioni. I nodi vengono solitamente impegnati per rappresentare le entità, ma a seconda della sfera delle relazioni possono essere utilizzati per scopi differenti. A parte proprietà e relazioni, i nodi possono anche essere etichettati con zero o più Label. Le relazioni tra i nodi, invece, sono una parte chiave dei database a grafo. Ci permettono di trovare le informazioni connesse. Come per i nodi, le relazioni possono avere le proprietà, hanno sempre una direzione e possiedono sempre un nodo di partenza e uno di arrivo. Infine, una label è un *named graph construct* , viene usata per raggruppare i nodi in sottoinsiemi; tutti i nodi etichettati con la stessa label fanno parte dello stesso insieme.

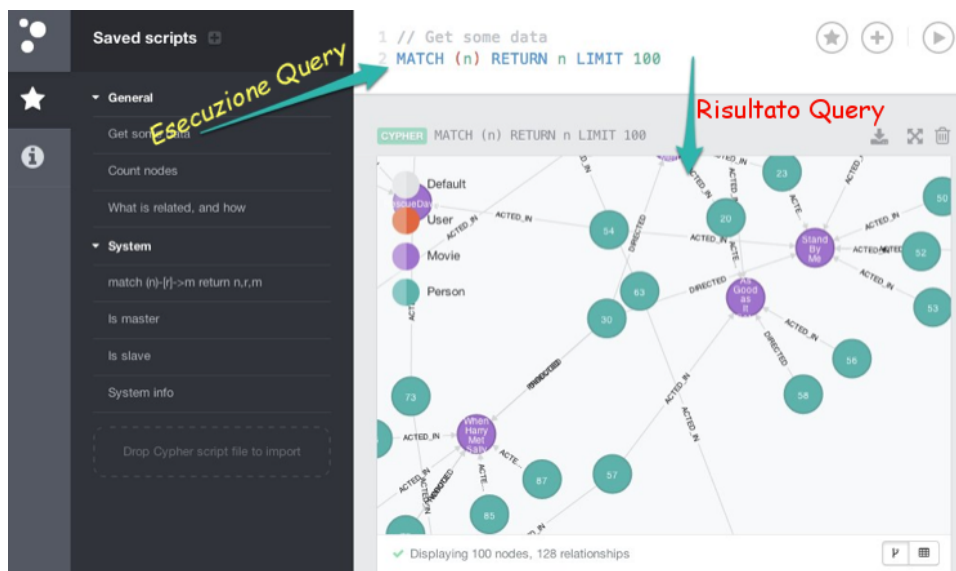


Figura 2.10: Interfaccia web di Neo4j nella quale è eseguita una query.

Tra i vantaggi che possiamo ottenere nell'utilizzo di Neo4j troviamo:

- **Distribuito:** Qualora si voglia garantire che il proprio sistema sia in grado di fornire un servizio di erogazione dei dati continuo, senza failure point e in grado di bilanciare e gestire un'enorme mole di richieste, Neo4j High Available (HA) è la risposta a questa esigenza. Neo4j HA, infatti, è stato progettato per rendere semplici, le transazioni da una singola macchina ad una macchina multipla, senza dover cambiare la tipologia delle istanze che andranno a comporre il cluster.
- **REST API:** Il server è munito di una ricca REST API che permettono ai client di spedire richieste in formato JSON per mezzo del protocollo HTTP. Le risposte vengono restituite all'interno di documenti JSON arricchiti con Hypermedia Links che mettono in risalto ulteriori caratteristiche del dataset. Sono molteplici le funzionalità messe a disposizione dalla REST API, ma il suo più grande vantaggio è quello di potervi accendere per mezzo di una semplice applicazione browser, come Firefox, Chrome o Internet Explorer.
- **Platform Independence:** Dato che le informazioni contenute nel server vengono accedute per mezzo di documenti JSON spediti attraverso l'HTTP, le applicazioni client possono essere costruite su qualsiasi tipo di piattaforma, basta possedere delle librerie client HTTP.
- **Scaling Independence:** Quando Neo4j viene eseguito in modalità server possiamo aumentare o diminuire il numero di componenti del cluster indipendente dal tipo di applicazione.
- **Per-request transactional:** Ogni richiesta da parte del client viene eseguita come una singola transazione, atomicamente separata dalle altre. Tuttavia, la REST API fornisce un supporto per l'esecuzione di operazioni in batch (ovvero l'esecuzione "accorpata" delle operazioni).

Neo4j, infine utilizza un sistema di query dichiarativo e grafico che prende il nome di Cypher. Esso è un linguaggio grafico, ovvero si basa sulla riproduzione grafica del sotto-grafo che si vuole estrarre. Con le query è possibile creare, modificare, eliminare e interrogare i dati del database. Il sotto-grafo riprodotto nelle query viene chiamato pattern, e per produrlo non servono strumenti particolari, ma basta seguire delle semplici regole che permettono di disegnarlo impiegando i caratteri ASCII (i caratteri presenti sulla tastiera).

Cypher permette, quindi, agli utenti (o ad una applicazione che agisce per conto dell'utente) di interrogare il database cercando i dati che corrispondono ad una specifica struttura. In termini da profano, chiediamo al database di *"cercare tutti quegli oggetti simili o che assomigliano"* ad un certo pattern.

```
1 Match (a: NODE_LABEL) -[r: RELATION_LABEL] -> (b: NODE_LABEL)
2 RETURN a, b, r;
```

Listato 2.1: Query Cypher.

Questo è un esempio di query scritta col linguaggio Cypher. Questo pattern descrive un path (percorso), che connette (a) a (b) ritornando sia i nodi che le relazioni che soddisfano la query.

2.3.5 Zookeeper

ZooKeeper è un servizio, centralizzato e open source, di coordinamento affidabile per applicazioni distribuite [4]. I servizi di coordinamento sono notoriamente difficili da ottenere; sono, infatti, particolarmente inclini a errori derivanti, ad esempio, da condizioni di stallo. ZooKeeper, invece, fornisce servizi operativi per cluster di grandi dimensioni. Esempi di questi servizi includono: informazioni di configurazione distribuita, un servizio di sincronizzazione e un registro di denominazione per i sistemi distribuiti. Le applicazioni sfruttano questi servizi per coordinare l'elaborazione distribuita tra cluster di grandi dimensioni. ZooKeeper inoltre, mira a semplificare la natura di questi diversi servizi in un'interfaccia molto semplice creando un servizio di coordinamento centralizzato. Il servizio stesso è distribuito e altamente affidabile. Questo perché, questo tipo di servizi risulta essere di difficile implementazione per le applicazioni distribuite.

Lo spazio dei nomi è costituito da registri di dati, chiamati *znodes*, in linguaggio ZooKeeper, che sono simili ai file e alle directory. A differenza di un tipico file system, progettato per l'archiviazione, i dati di ZooKeeper vengono conservati in memoria, il che significa che ZooKeeper può ottenere alti throughput e bassi numeri di latenza. Il database replicato di ZooKeeper comprende un albero di *znodi*, entità che rappresentano approssimativamente i nodi del file system (simili a directory). Ogni *znode* può essere arricchito da una matrice di byte, che memorizza i dati. Inoltre, ogni *znode* può avere altri *znodi* sotto di esso, praticamente formando un sistema di directory interno.

Zookeeper è considerato un servizio robusto, dal momento che i dati persistenti sono distribuiti tra più nodi (questo insieme di nodi è chiamato "ensemble") e un client si connette a uno di essi (cioè un "server" specifico), migrando se un nodo fallisce; Finché funziona la stragrande maggioranza dei nodi, l'insieme dei nodi ZooKeeper è vivo. In particolare, un nodo master viene scelto dinamicamente per consenso all'interno dell'ensemble; se il nodo principale non funziona, il ruolo del master passa a un altro nodo. È interessante notare che il nome di uno *znode* può essere sequenziale, il che significa che il nome che il client fornisce quando si crea lo *znode* è solo un prefisso: il nome completo è dato anche da un numero sequenziale scelto dall'ensemble. Ciò è utile, ad esempio, per scopi di sincronizzazione: se più client desiderano ottenere un blocco su una risorsa, possono contemporaneamente creare uno *znode* sequenziale su una posizione: chi ottiene il numero più basso ha diritto al blocco.

Inoltre, uno *znode* può essere "effimero" (temporaneo): questo significa che viene distrutto non appena il client che lo ha creato si disconnette. Ciò è utile soprattutto per sapere quando un client fallisce, il che può essere rilevante quando il client stesso ha delle responsabilità che dovrebbero essere prese da un nuovo client. Tutte le richieste di scrittura dei client vengono inoltrate a un singolo server, chiamato leader: in questo modo è possibile garantire che le scritture siano mantenute in ordine, vale a dire che le scritture siano lineari. Il resto dei server ZooKeeper, chiamati follower, riceve proposte di messaggi dal leader e concorda la consegna dei messaggi. Il livello di messaggistica si occupa di sostituire i

leader in caso di fallimento e di sincronizzare i follower con il leader. ZooKeeper utilizza un protocollo di messaggistica atomica personalizzato. Poiché il livello di messaggistica è atomico, ZooKeeper può garantire che le repliche locali non divergano mai. Quando il leader riceve una richiesta di scrittura, calcola lo stato del sistema quando deve essere applicata la scrittura e lo trasforma in una transazione che cattura questo nuovo stato. Ogni volta che un client scrive sull'ensemble, la maggior parte dei nodi mantiene l'informazione: questi nodi includono il server per il client e ovviamente il leader. Ciò significa che ogni scrittura rende il server aggiornato con il leader. Significa anche, tuttavia, che non è possibile avere scritture simultanee. La garanzia delle scritture lineari è la ragione del fatto che ZooKeeper non offra prestazioni ottimali per i carichi di lavoro dominanti dalla scrittura. In particolare, non dovrebbe essere usato per l'interscambio di grandi dati, come i media. Fintanto che la comunicazione coinvolga dati condivisi, ZooKeeper è utile. Quando i dati possono essere scritti contemporaneamente, ZooKeeper interviene, perché impone un rigoroso ordine di operazioni anche se non strettamente necessario dal punto di vista di chi scrive. Per la lettura, invece, ZooKeeper eccelle, questo perché le letture sono simultanee poiché sono servite dallo specifico server al quale il client si connette.

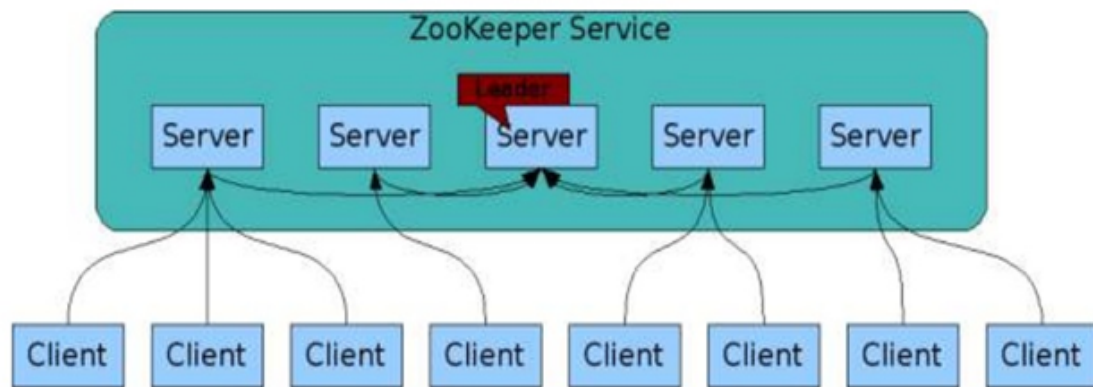


Figura 2.11: Funzionamento di Zookeeper.

ZooKeeper, in sintesi, è servizio molto veloce e molto semplice. La velocità di zookeeper è data da carichi di lavoro in cui le letture dei dati sono più comuni delle scritture. Il rapporto di lettura / scrittura ideale è di circa 10:1. Inoltre, Zookeeper mantiene uno spazio dei nomi gerarchico standard, simile a file e directory, questo lo rende un sistema sostanzialmente semplice. ZooKeeper viene replicato su un insieme di host (chiamato “ensemble”) e i server sono a conoscenza l’uno dell’altro. Finché sarà disponibile una massa critica di server, sarà disponibile anche il servizio ZooKeeper. Non c’è un singolo punto di errore. Per questo Zookeeper è detto: affidabile.

Dal momento che il suo obiettivo, però, è quello di essere una base per la costruzione di servizi più complicati, come la sincronizzazione, fornisce un insieme di garanzie tra cui:

- *Consistenza sequenziale*: gli aggiornamenti da un client verranno applicati nell’ordine in cui sono stati inviati.

- *Atomicità*: gli aggiornamenti hanno esito positivo o negativo. Nessun risultato parziale.
- *Immagine del sistema singolo*: un client vedrà la stessa vista del servizio indipendentemente dal server a cui si connette.
- *Affidabilità*: una volta applicato un aggiornamento, esso persisterà da quel momento fino a quando un client non sovrascriverà l'aggiornamento.
- *Tempestività*: la visualizzazione dei client del sistema è garantita per essere aggiornata entro un certo limite di tempo.

2.3.5.1 Kafka

Apache Kafka è una piattaforma di streaming open source distribuita che può essere utilizzata per compilare applicazioni e pipeline di dati in streaming in tempo reale. Kafka offre anche una funzionalità di broker di messaggi simile a una coda di messaggi, dove è possibile pubblicare e sottoscrivere flussi dei dati denominati [20]. Apache Kafka permette la gestione di centinaia di megabyte di traffico in lettura e scrittura al secondo da parte migliaia di Client. È stato sviluppato per la prima volta su LinkedIn e viene spesso utilizzato in combinazione con Apache Spark Streaming per l'elaborazione dei flussi in tempo reale.

Kafka ha un'architettura che differisce significativamente da altri sistemi di messaggistica. Ogni nodo prende il nome di broker. Kafka offre un'API Producer per la pubblicazione di record in un topic Kafka e un'API Consumer che viene utilizzata quando si sottoscrive un topic. Un topic è un nome di categoria o feed a cui i record sono pubblicati. I Topic in Kafka sono sempre multi-abbonato; cioè, un argomento può avere zero, uno o più consumatori che si abbonano ai dati scritti su di esso [2]. Kafka archivia i record in topic. Un topic è un insieme di messaggi della stessa categoria; per ogni topic il cluster Kafka mantiene un registro partizionato. Ogni partizione è una sequenza ordinata ed immutabile di messaggi che vengono aggiunti in continuazione all'interno del registro. I record, che come abbiamo visto sono archiviati in topic, vengono prodotti da un producer e usati dai consumer. I Producers o produttori pubblicano i messaggi (i dati) all'interno di un topic. Il Producer è responsabile di scegliere in quale partizione del registro del topic inserire un proprio messaggio. Ogni Producer sceglie il proprio algoritmo di assegnamento (per esempio un semplice round robin). I Consumers o consumatori leggono (consumano) i dati presenti all'interno del topic. Nel caso del sistema distribuito il produttore è integrato nell'ecosistema Spark, mentre il consumatore è una libreria creata ad-hoc presente all'interno della Webapp, in particolare nell'ecosistema NodeJS.

Ogni partizione è una sequenza ordinata e immutabile di record che viene continuamente aggiunta a un registro di commit strutturato. Ai record nelle partizioni viene assegnato un numero ID sequenziale chiamato offset che identifica in modo univoco ogni record all'interno della partizione [2].

In Kafka le partizioni sono replicate tra i nodi per garantire protezione in caso di interruzioni dei nodi (broker). La replicazione di una partizione viene gestita automaticamente in modo che sia assegnata a diversi brokers. Kafka elegge per ogni Broker una partizione

“Leader” (indicato con (L)) e tutte le scritture e le letture dovranno passare alla partizione “Leader” scelta. Il traffico dei producer viene indirizzato al leader di ogni nodo, usando lo stato gestito da ZooKeeper.

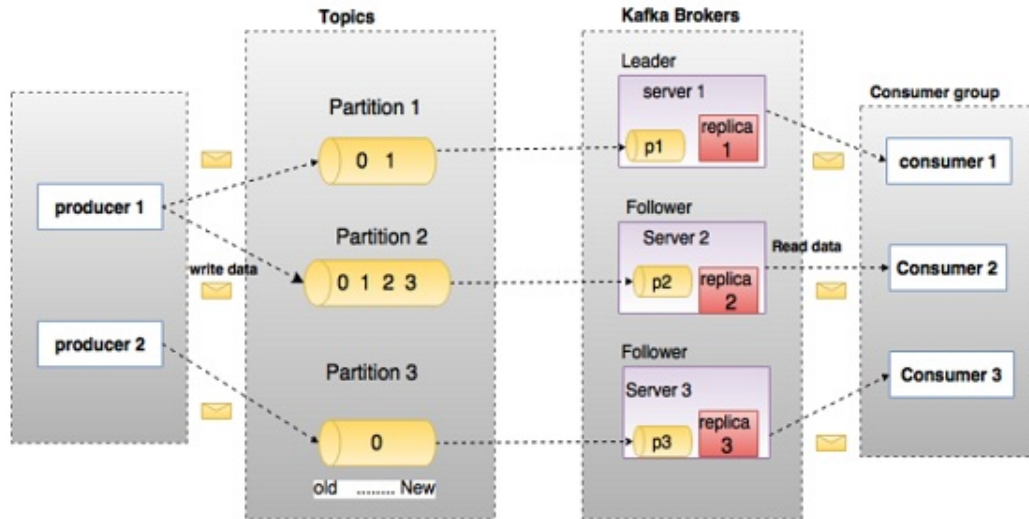


Figura 2.12: Come funziona Kafka

Infine, la messaggistica in Kafka prevede solo due di tipi di modelli:

- *Queuing (Coda)*: un pool di Consumers può leggere dal Server e ciascuno può leggere i dati solamente durante il suo turno;
- *Publish - subscribe*: il messaggio viene trasmesso a tutti i Consumers.

Kafka offre una sola implementazione dell'entità Consumer che generalizza entrambi i modelli, il consumer group. Un Consumer etichetta se stesso con il nome del gruppo a cui decide di far parte e ciascun messaggio pubblicato in un topic, seguito dal gruppo, viene consegnato a ciascun Consumer presente. Se tutti i Consumer hanno lo stesso consumer group, il sistema si reduce ad una semplice coda con priorità first in - first out (FIFO). Se tutti i consumer hanno un consumer group diverso, il sistema automaticamente diventa di tipo publish-subscribe.

2.4 WebApp

Una applicazione web è una applicazione client/server per un ambiente stateless, cioè senza memoria, che utilizza le tecnologie internet. Con il termine Webapp si descrive un'applicazione accessibile via web per mezzo di una network, come ad esempio una Intranet o attraverso la rete Internet. Pertanto, saper programmare per il web significa conoscere i diversi meccanismi e strumenti per conservare o passare i dati, detti parametri, tra le diverse pagine dell'applicazione web. In pratica una Web-application, è un

programma che non necessita di essere installato nel computer in quanto esso si rende disponibile su un server in rete e può essere fatto funzionare attraverso un normale Web browser (es. Google Chrome, Mozilla Firefox, Opera, ecc.). I client, dopo aver instaurato una connessione con il server, invia la richiesta per un servizio; il server dopo aver elaborato i dati necessari rende disponibile al client il servizio richiesto. A differenza dei siti web statici (HTML), l'applicazione web viene realizzata con una o più tecnologie (Javascript, Ajax, Servlet, Database ecc.) che permettono la creazione di un sito dinamico, cioè di un sito nel quale il contenuto delle pagine varia durante l'interazione.

Le applicazioni Web si pongono come valida alternativa alle tradizionali applicazioni Client/Server per vari motivi:

- *Facilità di distribuzione e aggiornamento*: un'applicazione Web risiede interamente sul server, per cui la sua pubblicazione coincide con la distribuzione e l'aggiornamento ed è automaticamente disponibile a tutti gli utenti.
- *Accesso multi-piattaforma*: l'accesso all'applicazione è indipendente dall'hardware e dal sistema operativo utilizzato dagli utenti;
- *Riduzione del costo di gestione*: l'uso di Internet come infrastruttura per un'applicazione Web riduce notevolmente sia i costi di connettività che i costi di gestione dei client;
- *Scalabilità*: un'applicazione Web ben progettata può crescere insieme alle esigenze dell'azienda senza particolari problemi;

Per poter mostrare l'elaborazione di Spark nel sistema distribuito, si è preferito utilizzare una webapp per i motivi sopra citati. Questa applicazione web fa uso di tecnologie di ultima generazione quali NodeJs ed HTML5. Questo la rende fruibile a tutti gli utenti provvisti di Pc, Tablet o Smartphone con un browser installato. Nel seguente capitolo dunque vengono messi in luce i principali componenti che compongono la webapp, mostrando i punti forti di ogni tecnologia.

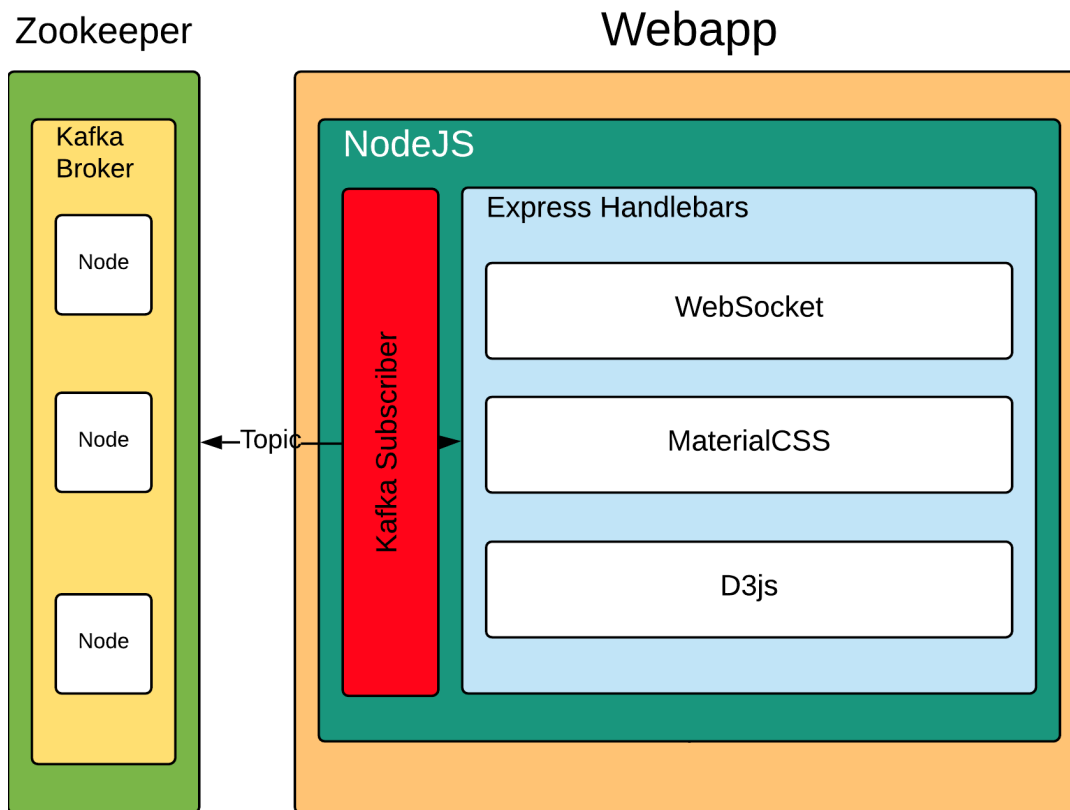


Figura 2.13: Architettura in dettaglio della webapp.

2.4.1 Node.js

Node.js (anche conosciuto come Node o Nodejs) è un framework/piattaforma molto potente basato sul motore *JavaScript V8* di Google Chrome, per creare facilmente applicazioni di Web veloci e scalabili. Pubblicata da Ryan Dahl nel 2009, viene utilizzato per sviluppare applicazioni Web intensive di I/O come siti di streaming video, applicazioni a pagina singola e altre applicazioni Web. Node.js è un ambiente open source, multi-piattaforma per lo sviluppo di applicazioni lato server completamente gratuito, utilizzato da migliaia di sviluppatori in tutto il mondo. Le applicazioni Node.js sono scritte in Javascript e possono essere eseguite all'interno del runtime di Node.js su OSX, Microsoft Windows e Linux.

Utilizza un modello I/O non bloccante e basato sugli eventi che lo rende leggero ed efficiente, perfetto per applicazioni in tempo reale ad alta intensità di dati che funzionano su dispositivi distribuiti [22]. Il modello di networking su cui si basa Node.js non è quello dei processi concorrenti, ma I/O event-driven: ciò vuol dire che Node richiede al sistema operativo di ricevere notifiche al verificarsi di determinati eventi, e rimane quindi in

sleep fino alla notifica stessa. Solo in tale momento torna attivo per eseguire le istruzioni previste nella funzione di *callback*, così chiamata perché da eseguire una volta ricevuta la notifica, che il risultato dell'elaborazione è disponibile. Tale modello di networking, implementato anche nella libreria Event machine per Ruby e nel framework Twisted per Python, è ritenuto più efficiente nelle situazioni critiche in cui si verifica un elevato traffico di rete [16]. Di fronte alle esigenze di migliorare le performance dei software di rete Ryan Dahl ha creato una piattaforma che esegue le operazioni di I/O particolarmente lente (comunicazioni di rete o accesso al disco) in modo asincrono, rendendo la programmazione su Node JS diversa da qualsiasi esperienza con altri linguaggi. In definitiva, l'obiettivo di Node.js è quello di fornire un modo veloce per realizzare applicazioni web scalabili in termini di gestione delle connessioni da parte dei client verso il web server.

Nel momento in cui una operazione di I/O considerata lenta (di solito lo è se riguarda la rete o il disco fisso) viene eseguita da un programma in Node JS, V8 si occupa di trasferire la chiamata su un thread non bloccante fra quelli che ha a disposizione nella sua *thread-pool base*. In questo modo, il thread principale con il codice può continuare la sua esecuzione senza context switch. Nel momento in cui una operazione collegata ai thread non bloccanti è terminata il kernel segnala che questo thread può tornare in coda di esecuzione. A questo punto però V8 si occuperà di intercettare il messaggio, mettere nella propria coda di esecuzione la funzione di callback specificata con l'operazione di I/O terminata e di rimettere il thread non bloccante a disposizione per altre operazioni di I/O. Così facendo, virtualmente il thread che esegue codice non si ferma mai, avvicinando le funzioni di callback delle varie operazioni terminate [2.14].

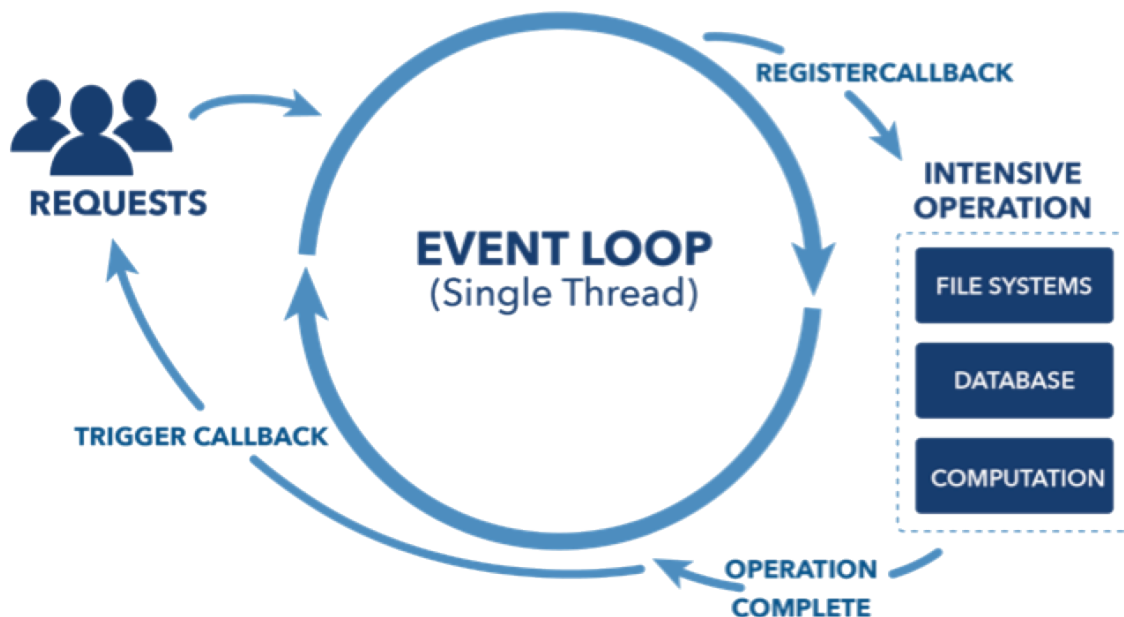


Figura 2.14: Gestione eventi con Node.js.

Bisogna però sempre tenere a mente, quando si progetta di scrivere un programma con Node.js, che questa architettura ha un effetto collaterale molto pesante; le operazioni

che occupano per lungo tempo il thread in cui viene eseguito il codice (operazioni di calcolo onerose) bloccano l'interno software. Per questo motivo Node.js è assolutamente sconsigliato in caso di operazioni di calcolo complesse e nella fase di progettazione di programmi che usano questa piattaforma è sempre necessario utilizzare questa caratteristica come criterio fondamentale per la scrittura del codice. Di seguito sono alcune delle caratteristiche importanti che rendono Node.js la prima scelta di architetti del software:

- **Asynchronous and Event Driven:** Tutte le API della libreria Node.js sono asincrone, ovvero non bloccanti. Significa essenzialmente che un server basato su Node non aspetta mai che un'API restituisca i dati. Il server passa all'API successiva dopo averlo chiamato ed un meccanismo di notifica di eventi consente al server di ottenere una risposta dalla precedente chiamata.
- **Molto veloce:** Essendo costruito sul motore Javascript V8 di Google Chrome, Node.js è molto veloce nell'esecuzione del codice.
- **Thread singolo ma altamente scalabile:** Node utilizza un singolo thread con loop di eventi. Il meccanismo degli eventi aiuta il server a rispondere in modo non bloccante e rende il server altamente scalabile rispetto ai server tradizionali che creano thread limitati per gestire le richieste. Quindi, Node utilizza un singolo programma con thread e lo stesso programma può fornire il servizio a un numero molto più grande di richieste rispetto ai server come Apache HTTP Server.
- **Nessun Buffering:** Le applicazioni create con node non bufferizzano mai alcun dato. Queste applicazioni generano semplicemente i dati in blocchi.
- **Licenza:** Node è rilasciato sotto licenza MIT.

Esiste un mondo attorno a Javascript composto da librerie che ne estendono le funzionalità. Stesso discorso vale per Node.js, in quanto è attiva una comunità di sviluppo che ha realizzato in questi anni molte librerie per realizzare particolari tipi di supporto (database, Network, . . .) ed un sistema di installazione di questi moduli che si occupa anche di eventuali dipendenze: *npm*. Nei prossimi capitoli saranno analizzati alcuni progetti legati a Node.js usati nella creazione della webapp.

2.4.1.1 Express.js ed Handlebars

Express.js (anche chiamato semplicemente Express) è un framework basato su Node.js che offre un insieme robusto di utilità per realizzare agilmente un'architettura *MVC (Model-View-Controller)* sul lato server di applicazioni web single-page, multi-page ed ibride.

Basato sul modulo di Node chiamato *connect*, risulta essere un ottimo "connettore" o *middleware* tra le diverse librerie che possono popolare una webapp, tra cui WebSocket, Passport, Mustache.js, Handlebars, etc.

Le funzionalità offerte sono il *routing*, la possibilità di gestire le configurazioni dell'applicazione ed un motore di templating.

Al centro di questa libreria c'è il concetto di flusso di funzioni, o come piace dire a certe persone, un set di *livelli di funzione*. Infatti, per creare un'applicazione Express c'è bisogno di creare una sequenza di funzioni (livelli) che il framework può navigare.

Quando una di queste decide di entrare in gioco, può completare il processo e fermare la sequenza. Dopodiché Express riprende a scorrere la sequenza fino alla fine.

Queste funzioni hanno delle callback associate che sono eseguite nel momento in cui un client effettua una richiesta. Questo processo prende il nome di Routing.

Il routing, dunque, è una funzionalità di Express.js, che determina la risposta ad una richiesta client ad un endpoint particolare; il quale può essere un URI (o percorso) o un metodo di richiesta HTTP specifico. In pratica, con Express, si possono gestire le richieste e le risposte *in the middle*, cioè fra server e client. Quando il server riceve una richiesta HTTP la racchiude all'interno di un oggetto *ServerRequest*. Questo oggetto, insieme all'oggetto *ServerResponse*, viene passato al primo middleware che ne può modificare il contenuto, o aggiungere proprietà. Una volta terminata la modifica, il middleware richiamerà il successivo nell'eventuale catena (di funzioni) presente.



Figura 2.15: Visualizzazione della catena di funzioni per rispondere ad una richiesta

Se l'ultima funzione della catena deve restituire un HTML come risultato, Express chiama in causa il *template engine*. Questo motore, infatti si occupa di fare il tramite tra i dati presenti nel layer model e quello di presentazione. In particolare elabora la richiesta producendo, in fase di run-time, un HTML dinamico partendo da una struttura definita (template). Esistono diversi motori che possono essere utilizzati con Express ognuno con le proprie particolarità e specifiche.

Per questo elaborato la scelta del templating engine è caduta su *Handlebars.js*. La libreria di templating Handlebars consente di creare un'interfaccia utente ricca includendo HTML statico e contenuto dinamico, che possono essere specificati nelle doppie parentesi graffe. Handlebars.js è molto popolare, semplice da usare e con una grande community. È basato sul linguaggio dei modelli di Mustache, ma lo migliora in molti aspetti. Con Handlebars, si può separare la generazione di HTML dal resto del JavaScript e scrivere codice più pulito. Inoltre, aggiunge costrutti (if e cicli for) che permettono di creare dinamicamente l'HTML. Infine, introduce un sistema di *partial* che permette allo sviluppatore di inserire nelle proprie pagine, porzioni di HTML provenienti da file esterni.

Express, quindi, per costruire la pagina di risposta da inviare al client, chiama Handlebars che preleva i file con estensione *.handlebars* li "compila" e genera il risultato finale.

```
1 <div class="entry">
2   <h1>{{title}}</h1>
3   <h2>By {{author.name}}</h2>
4   <div class="body">
5     {{body}}
6   </div>
7 </div>
```

Listato 2.2: Esempio di HTML scritto con Handlebars.

Il listato 2.2 è un esempio di HTML scritto con la sintassi di Handlebars. I valori *title*, *author.name* e *body* saranno sostituiti, in fase di run-time, con i valori provenienti dal controller che passerà un oggetto all'engine templating. vedi listato 2.3.

```
1 var context = {
2   title: "My First Blog Post!",
3   author: {
4     id: 47,
5     name: "Yehuda Katz"
6   },
7   body: "My first post. Wheeeee!"
8 };
```

Listato 2.3: Esempio di variabile passata dal controller al template engine.

2.4.1.2 WebSocket

Come detto in precedenza, l'applicazione web dell'elaborato di tesi, prevede una interfaccia grafica per la visualizzazione delle ultime transazioni provenienti da Bitcoin. Per poter implementare questa funzionalità, c'è bisogno di utilizzare tecniche che permettano l'invio di dati tra client e server in modo asincrono. Tra le novità portate da HTML5, i WebSocket rappresentano quella di maggior importanza dal punto di vista dell'interazione tra client e server. WebSocket è una tecnologia per effettuare comunicazioni bidirezionali in tempo reale. Essi prevedono un canale di comunicazione sempre attivo, a bassa latenza, tra client e server, utilizzabile da entrambi sia in scrittura che in lettura. Tale canale è costituito da una connessione TCP persistente, garantito da un handshaking client-key iniziale ed un modello di sicurezza originbased. Per la protezione dei dati trasmessi contro lo sniffing sono applicate apposite maschere. Ecco le caratteristiche principali dei WebSocket:

2.4.1.3 MaterialCSS

2.4.1.4 D3js

Capitolo 3

Implementazione

3.1 Diagramma delle classi

3.2 Codice

3.3 Github

Capitolo 4

Conclusioni e sviluppi futuri

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut

porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Bibliografia

- [1] Maarten van Steen e Andrew S. Tanenbaum. «A brief introduction to distributed systems». In: (2008). URL: <https://www.cs.vu.nl/~ast/Publications/Papers/computing-2016.pdf>.
- [2] Apache. *Apache Kafka*. URL: <http://kafka.apache.org>.
- [3] Apache. «Apache Software Foundation. Official Website Apache Spark». In: URL <http://spark.apache.org> ().
- [4] Apache. *Apache Zookeeper*. URL: <http://zookeeper.apache.org>.
- [5] Blockchain.com. *Chart bitcoin transactions*. URL: <https://www.blockchain.com/it/charts/n-transactions?timespan=all> (visitato il 2018).
- [6] Blockchain.com. *Chart value of Bitcoin*. URL: <https://www.blockchain.com/charts/market-price?timespan=all> (visitato il 2018).
- [7] www.nosql database.org. *NoSQL Database*. URL: <http://www.nosql-database.org/> (visitato il 2018).
- [8] Wolfgang Emmerich. «Distributed System Principled». In: (1997). URL: <http://www0.cs.ucl.ac.uk/staff/ucacwxe/lectures/ds98-99/dsee3.pdf>.
- [9] en.bitcoinwiki.org. *Bitcoind*. URL: <https://en.bitcoinwiki.org/wiki/Bitcoind> (visitato il 2018).
- [10] en.bitcoinwiki.org. *ECDSA*. URL: https://en.bitcoinwiki.org/wiki/Elliptic_Curve_Digital_Signature_Algorithm (visitato il 2018).
- [11] en.wikipedia.org. *Merkle Tree*. URL: https://en.wikipedia.org/wiki/Merkle_tree (visitato il 2018).
- [12] en.wikipedia.org. *ZeroMQ*. URL: <https://en.wikipedia.org/wiki/ZeroMQ> (visitato il 2018).
- [13] Alex Holmes. *Hadoop in practice*. Manning Publications Co., 2012.
- [14] it.wikipedia.org. *Bitcoin*. URL: <https://it.wikipedia.org/wiki/Bitcoin> (visitato il 2018).
- [15] it.wikipedia.org. *Neo4j*. URL: <https://it.wikipedia.org/wiki/Neo4j> (visitato il 2018).
- [16] it.wikipedia.org. *Node.js*. URL: <https://it.wikipedia.org/wiki/Node.js>.

- [17] it.wikipedia.org. *PageRank*. URL: <https://it.wikipedia.org/wiki/PageRank> (visitato il 2018).
- [18] it.wikipedia.org. *Sistema Distribuito*. URL: https://it.wikipedia.org/wiki/Sistema_distribuito#Caratteristiche (visitato il 2018).
- [19] Leslie Lamport. «Distribution». In: (1997). URL: <https://www.microsoft.com/en-us/research/uploads/prod/2016/12/Distribution.pdf>.
- [20] Microsoft. *Introduzione ad Apache Kafka*. URL: <http://docs.microsoft.com/it-it/azure/hdinsight/kafka/apache-kafka-introduction>.
- [21] Satoshi Nakamoto. «Bitcoin: A Peer-to-Peer Electronic Cash System». In: (2009). URL: <https://bitcoin.org/bitcoin.pdf>.
- [22] nodejs.org. *Node.js*. URL: <https://nodejs.org/en>.
- [23] spark.apache.org. *GraphX*. URL: <https://spark.apache.org/docs/latest/graphx-programming-guide.html> (visitato il 2018).
- [24] spark.apache.org. *Spark Streaming*. URL: <https://spark.apache.org/docs/latest/streaming-programming-guide.html> (visitato il 2018).
- [25] Tom White. *Hadoop: The definitive guide*. "O'Reilly Media, Inc.", 2012.
- [26] www.albertodeluigi.com. *Come avviene una transazione bitcoin*. URL: <http://www.albertodeluigi.com/index/bitcoin/transazione-bitcoin#panel-108-0-0-1> (visitato il 2018).
- [27] www.google.com. *Google Press Center: Fun Facts*. URL: <https://web.archive.org/web/20090424093934/http://www.google.com/press/funfacts.html> (visitato il 2018).
- [28] zeromq.org. *How come ØMQ has higher throughput than TCP although it's built on top of TCP?* URL: <http://zeromq.org/area:faq#toc6> (visitato il 2018).