

ARCHITETTURE SCALABILI PER MEMORIZZAZIONE, ANALISI, CONDIVISIONE E PUBBLICAZIONE DI GROSSE MOLI DI DATI

1. PREMESSA

Con l'incremento esponenziale della quantità di informazioni prodotte nel mondo, la capacità di analizzarle in maniera relativamente semplice e veloce sta diventando un elemento fondamentale. Enormi moli di informazioni, di tipologie e formati differenti, sono le caratteristiche principali dei cosiddetti "big data", un binomio che recentemente è entrato nell'uso comune sebbene le sue prime introduzioni risalgano addirittura all'inizio del nuovo millennio. Quando i computer hanno iniziato a diffondersi come strumenti di business, oltre che di ricerca e di analisi, la prima sfida è stata quella di riuscire ad archiviare e gestire le informazioni aziendali e quelle legate ai vari ambiti della ricerca scientifica. Le limitazioni tecnologiche di allora hanno portato alla creazione dei database così come li conosciamo, composti da righe e colonne, che i computer erano in grado di gestire senza particolari requisiti prestazionali. Il passaggio successivo è stato quello di iniziare ad utilizzare concretamente questi dati, dal momento che le informazioni hanno rappresentato da sempre una vera e propria ricchezza per le aziende e la possibilità di organizzarle all'interno dei database ha permesso di sfruttarne le potenzialità, grazie al supporto dei computer nelle attività essenziali di analisi e automazione. Oggi la nuova sfida riguarda invece i grandi volumi di dati, la pura quantità (e velocità di crescita) delle informazioni che vengono prodotte quotidianamente e che risultano problematiche da gestire anche per il database più potente al mondo.

Dal momento che la produzione dell'informazione avviene tramite strumenti e canali che sono i più variegati (dagli smartphone ai social network – con tecniche di crowdsourcing – fino ad arrivare a tutto ciò che può rientrare nell'Internet delle cose, come la sensoristica, la domotica, etc.), altrettanto diversificato è il tipo di informazioni prodotte. Siamo lontani dall'informazione fatta solo di parole e numeri, in quest'epoca dove anche immagini e video sono sempre più importanti nella nostra vita. Tutto è informazione, tutto è correlabile ed è sempre meno trascurabile, anche in ambiti disciplinari quali l'archeologia.

2. GLI STRUMENTI DI GESTIONE E DI ANALISI PER I BIG DATA

In passato, la gestione dei big data è stata un semplice problema tecnologico. Oggi, non è più così poiché esistono framework scalabili come Hadoop che nascono proprio per la gestione di grandi moli di dati oltre a database *ad hoc*

– raccolti sotto le definizioni di NOSQL e di NewSQL – che nascono con l’obiettivo di superare i limiti strutturali e intrinseci dei database relazionali (RDBMS). Settori quali il data warehousing e la business intelligence stanno convergendo sempre più verso le nuove metodologie di analisi dei dati che prendono il nome di “big data analytics”. Esattamente come per il data warehousing e la business intelligence, queste metodologie permettono di gestire grandi quantità di dati e di estrarre informazioni preziose per il business o per la ricerca scientifica quali, ad esempio, l’analisi dei trend dei consumatori o le correlazioni che si possono instaurare tra dati apparentemente non relazionabili. La necessità di tecniche di analisi dei big data, dunque, attraversa tutti i segmenti industriali e scientifici: classiche analisi e segmentazione della clientela, gestione delle reti di trasporto e di distribuzione dell’energia, iniziative di sicurezza civile e militare, gestione di Digital Libraries e del Digital Cultural Heritage e molto altro ancora.

Nel caso in cui i dati rientrino nella categoria dei big data, non è immediato effettuare il passaggio dalle metodologie tradizionali di analisi ai processi di tipo “big data analytics”: si parla, infatti, di un cambio di paradigma legato alle modalità di analisi e, di conseguenza, le competenze richieste al personale IT dedicato a questa funzionalità sono radicalmente diverse da quelle che sono spesso possedute da chi si occupa di analisi in ambito tradizionale. I tool stessi sono diversi, sia per la memorizzazione dei dati, sia per l’analisi: raramente si parla di database relazionali – principalmente per via dei loro limiti di scalabilità – e sempre più frequentemente si parla di structured storage, DBMS di tipo NOSQL/NewSQL, object store e file system distribuiti. Questi requisiti tecnologici, in ogni caso, non devono essere visti come problematiche o barriere all’ingresso, piuttosto come nuove opportunità: riqualificando il proprio personale IT – creando anche nuove figure professionali – e adottando tali strumenti e tecniche, ci si rende immediatamente conto dei notevoli vantaggi ottenibili in termini di tempistiche e di risultati complessivi delle analisi.

I database relazionali, come già detto, non rendono semplice e immediata la gestione di grandi moli di dati, specie se devono essere distribuiti su cluster composti da diversi nodi di calcolo. Basti pensare ai principali social network quali Facebook, Digg o Twitter e agli enormi sforzi tecnologici che, in passato, hanno dovuto compiere per gestire tutti gli utenti e i dati da loro prodotti. Tali sforzi, infatti, crescono in base alla quantità di informazioni da memorizzare e sono incentrati sulla gestione di basi dati replicate e distribuite, alle quali vengono richieste caratteristiche di sicurezza, affidabilità, tolleranza ai guasti e performance. In queste situazioni, la scalabilità orizzontale – cioè l’incremento del numero di macchine per aumentare la potenza a disposizione, sia essa di calcolo, sia di memorizzazione – è sempre stata un problema per i database relazionali in quanto essi richiedono l’utilizzo di tecniche *ad hoc* (sharding, replicazione, etc.) spesso complicate o costose sia in termini prestazionali, sia in termini economici. Proprio per risolvere questo problema, Facebook, ad esempio, ha deciso di sviluppare

internamente un proprio database che permettesse di ottenere facilmente la scalabilità, mantenendo solamente alcune delle caratteristiche tipiche dei RDBMS. In seguito a questo episodio, molti altri social network hanno deciso di migrare verso database di questo tipo per la gestione dei loro dati.

3. LE BASI DI DATI DI TIPO NOSQL E NEWSQL

In questo settore applicativo sono stati sviluppati molti prodotti, ognuno con caratteristiche differenti: alcuni sono ottimizzati per escludere i “single point of failure” tramite svariate tecniche di replica dei dati (fra rack differenti, fra data center differenti, etc.), altri puntano ad ottenere prestazioni elevate, nonostante la struttura distribuita – ad esempio, basandosi su strutture dati allocate interamente in memoria, come accade per i cosiddetti “in-memory database” – e altri ancora prediligono la semplicità di utilizzo e di gestione. Oltre ai social network citati in precedenza, molte altre aziende hanno avuto necessità analoghe e hanno fatto ricorso a prodotti specializzati, spesso abbandonando gli RDBMS di tipo general purpose che da anni adottavano nei loro ambienti di produzione. Il fattore che accomuna questa moltitudine di database, evidenziato anche dall'appartenenza ad una famiglia chiamata NOSQL – o, come più correttamente si dovrebbe definirla, structured storage – è proprio il fatto che essi non sfruttano il classico linguaggio di interrogazione SQL e che utilizzano una rappresentazione dei dati spesso slegata dal consueto concetto di tabella.

Il movimento NOSQL, al contrario di quello che il nome sembrerebbe indicare, non nasce in opposizione alla tradizione dettata dal linguaggio SQL, bensì come sua valida alternativa per la risoluzione di problematiche differenti, proprio legate alla continua crescita dei dati e all'eterogeneità tipica dei big data. Infatti, NOSQL è l'acronimo di Not Only SQL ed è un movimento che sostiene l'utilizzo dei database di questo tipo laddove i tradizionali RDBMS risultino carenti o troppo complessi, seguendo la filosofia del “right tool for the job”. I campi di applicazione ideali per le basi dati NOSQL sono quelli dove è necessario gestire ambienti distribuiti con dati replicati, specie se si è disposti a “sacrificare” alcune delle caratteristiche tipiche dei RDBMS – ad esempio quelle che vanno sotto l'acronimo ACID (Atomicity, Consistency, Isolation, Durability) – verso un insieme di caratteristiche differenti, definite dalla sigla BASE (Basically Available, Soft state, Eventually consistent). Tali connotati tecnologici permettono di ottenere, pur a discapito di una diversa gestione della consistenza dei dati e della loro disponibilità, un notevole incremento della scalabilità, secondo il “teorema CAP” per il quale non si può avere Consistency, Availability e Partition tolerance al tempo stesso.

Inoltre, uno degli obiettivi dei sistemi informativi distribuiti odierni è di poter garantire una disponibilità elevata, persino quando alcuni nodi non sono raggiungibili a causa di guasti improvvisi; è importante, quindi, che il servizio

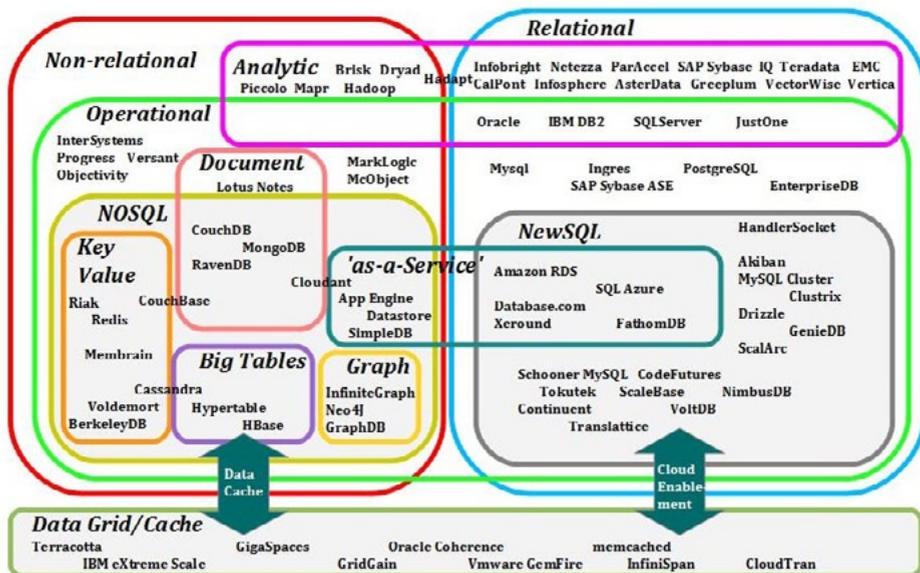


Fig. 1 – Una suddivisione schematica dei principali sistemi di basi di dati.

possa continuare a funzionare, anche in presenza di una o più failures che determinino un partizionamento della rete, e che ogni richiesta ricevuta sia comunque evasa anche se il sistema presenta al suo interno dei nodi non operativi. Le basi di dati NOSQL, fra le altre cose, cercano di risolvere almeno parzialmente la suddetta problematica e il fatto che molte di esse siano free ed open source permette di avere un'elevata adattabilità alle problematiche specifiche che si vogliono risolvere, oltre ad offrire un TCO (Total Cost of Ownership) nettamente inferiore rispetto all'utilizzo di soluzioni legate ai tradizionali RDBMS e alla possibilità di utilizzare del commodity hardware come nodi della propria infrastruttura di memorizzazione. A tutto questo si aggiunge una fervente community di sviluppatori che offrono costantemente nuove caratteristiche e contribuiscono al supporto dei vari prodotti. Questi database sembrano sposarsi perfettamente con la crescente metodologia del cloud computing, che richiede l'utilizzo di ambienti fortemente distribuiti, scalabili e dinamici, all'interno dei quali il numero di macchine fisiche o virtuali è in continua variazione. Inoltre, molti di questi prodotti sono in grado di gestire più o meno facilmente i big data e permettono l'integrazione con tool di big data analytics.

Accanto alla categoria dei database NOSQL, di recente è stato adottato anche il termine NewSQL (Fig. 1). I database appartenenti a questa famiglia tendono ad integrare i pregi del mondo NOSQL con quelli del mondo dei

RDBMS. Normalmente, essi nascono per funzionare in ambienti distribuiti e supportare la gestione di grandi moli di dati, pur mantenendo gli standard dominanti fra i RDBMS, cioè le caratteristiche ACID delle transazioni e la possibilità di utilizzare il linguaggio SQL.

4. LA SPERIMENTAZIONE EFFETTUATA

Utilizzando strumenti appartenenti a queste famiglie, nelle nostre sperimentazioni tecnologiche abbiamo cercato di definire un'infrastruttura che potesse memorizzare, analizzare e condividere ingenti moli di dati per potenziali applicazioni di gestione di basi di dati scientifiche e di contenuto archeologico/culturale. Il cuore dell'infrastruttura è dato da Hadoop, un framework che supporta applicazioni che vengono distribuite con un elevato livello di accesso ai dati e che lavorano con migliaia di nodi e petabyte di dati; Hadoop si ispira alla tecnologia "Map/Reduce" di Google e al Google File System ed è un progetto realizzato e utilizzato da una comunità globale di utenti e di contributori.

Intorno a questo cuore tecnologico, ruota una serie di altri strumenti come Fuse, HBase, HAProxy, Keepalived, R, Apache Web Server e il modulo WebDAV. Fuse, acronimo di "File system in USERSpace", è un progetto open source rilasciato sotto la licenza GPL e LGPL e volto alla realizzazione di un modulo per il kernel Linux che permetta agli utenti con permessi non amministrativi di creare un proprio file system senza la necessità di scrivere codice a livello del kernel. A differenza dei file system tradizionali, che si preoccupano principalmente di organizzare e memorizzare i dati su disco, i file system caricati nello userspace (o file system virtuali) non memorizzano realmente i dati per conto proprio: essi agiscono, infatti, da tramite fra l'utente ed il file system reale sottostante alla base di dati. Fuse è un sistema molto potente poiché, virtualmente, ogni risorsa disponibile ad essere implementata sfruttando Fuse può divenire un file system virtuale.

Apache HTTP Server, o più comunemente Apache, è il nome del web server modulare più diffuso, in grado di operare su sistemi operativi Unix/Linux e Microsoft. Apache è un software open source che realizza le funzioni di trasporto delle informazioni, di internetwork e di collegamento tra risorse di sistema, con il vantaggio di offrire anche funzioni di controllo per la sicurezza come quelle che compie normalmente un proxy. Questo applicativo, grazie all'installazione di un piccolo modulo, supporta anche il protocollo WebDAV – Web-based Distributed Authoring and Versioning – ovvero un set di istruzioni del protocollo HTTP che permettono all'utente di gestire in modo collaborativo dei file su di un server remoto, con lo scopo di rendere il World Wide Web un mezzo di lettura e scrittura, seguendo le linee guida dell'idea originale di Tim Berners-Lee. Molti sistemi operativi moderni, inclusi quelli per smartphone,

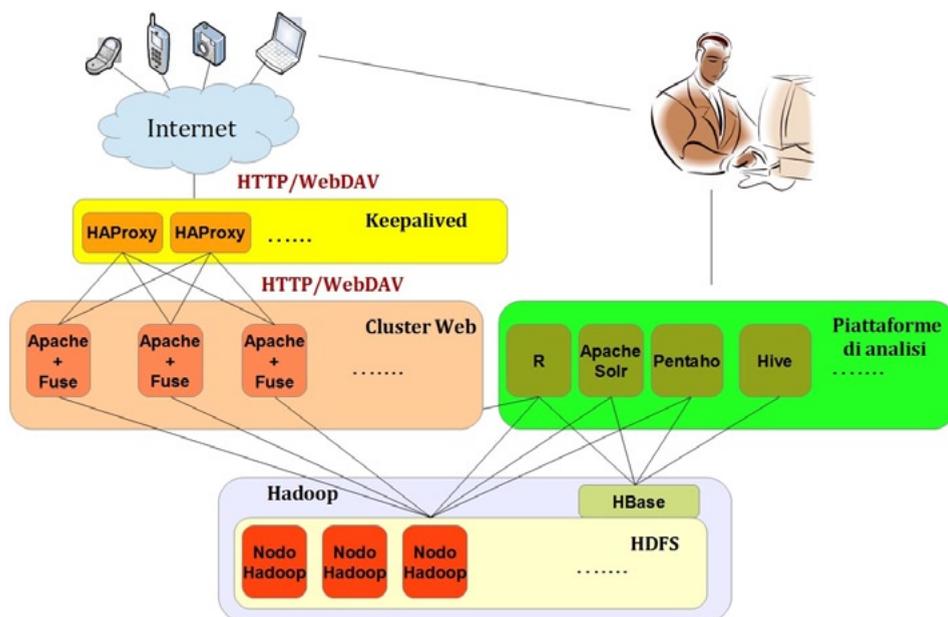


Fig. 2 – Architettura finale.

tablet e PDA, hanno un supporto integrato per il protocollo WebDAV e, con il giusto client ed una rete sufficientemente veloce, la gestione dei file su server WebDAV risulta facile quasi quanto la gestione dei file in locale.

HBase è una base di dati NOSQL distribuita, di tipo open source, modellata sull'idea di BigTable di Google e scritta in Java. Per usare HBase è necessario adottare un cluster Hadoop sul quale sia configurato HDFS – Hadoop Distributed File System – come sistema di storage. HAProxy è una soluzione open source per ottenere l'alta affidabilità e fare il load balancing di un servizio basato su protocollo TCP e replicato su più macchine. È particolarmente adatto per i server web che devono gestire migliaia di connessioni e, come tale, calza a pennello con la nostra architettura. Keepalived è un altro software open source che ci permette di costruire un'architettura replicata active-passive dove al fallimento di una macchina, un'altra ne prende immediatamente le veci, in modo da non interrompere il servizio e offrire continuità operativa. R è un ambiente di sviluppo specifico per l'analisi statistica dei dati; è un software libero, distribuito con la licenza GNU GPL ed è disponibile per diversi sistemi operativi. Il linguaggio su cui si basa R offre molte funzionalità per l'analisi dei dati, sia in campo statistico, sia in moltissimi altri campi – anche grazie ad una notevole estensibilità basata su pacchetti aggiuntivi – quali la gestione ed elaborazione dei dati e la loro rappresentazione grafica.

Mettendo insieme tutti questi applicativi open source (Fig. 2), si può ottenere un web-based file storage collaborativo, che può essere utilizzato con una moltitudine di dispositivi (PC, smartphone, tablet, etc.). Tale sistema integra già alcune metodologie adatte ad effettuare operazioni di big data analytics e al contempo si presta in maniera molto semplice ad interagire con sistemi di analisi avanzata già esistenti (ad esempio Solr) o creati *ad hoc*, grazie alla semplice integrazione di Hadoop con molti linguaggi di programmazione, Java *in primis*.

5. LA REALIZZAZIONE PRESSO IL LABORATORIO DI CSP

Un prototipo a scopo dimostrativo della suddetta architettura è stato ricreato all'interno alla struttura tecnologica dell'azienda CSP – Innovazione nelle ICT. Tale organismo di ricerca senza scopo di lucro si occupa di trasferimento tecnologico e di sviluppo sperimentale in ambito ICT, mettendo a disposizione del territorio, delle imprese e della pubblica amministrazione i risultati concreti delle proprie attività. Analizzando gli ultimi anni e limitandoci agli aspetti legati alla diffusione e allo sviluppo di servizi innovativi nel settore dei beni culturali, CSP ha contribuito con progetti quali:

- Astronomia in rete, nel cui ambito alcuni importanti osservatori astronomici piemontesi hanno reso pubbliche, tramite Internet, le loro immagini e i loro dati, visualizzandoli anche in tempo reale e permettendo di effettuare il puntamento di telescopi da remoto;
- Vivivalsesia, un portale di servizi per il turismo dove vengono fornite delle personal guide su dispositivi mobili, per scoprire le bellezze locali, oltre a realizzare dei servizi di video-valorizzazione per promuovere le risorse turistiche di un territorio ricco di tesori artistici come il Sacro Monte di Varallo, riconosciuto patrimonio dell'umanità dall'UNESCO.
- Orto botanico di Torino, un laboratorio di raccolta di dati legati ai cambiamenti climatici e alle risposte dei vegetali, svolto per conto del Dipartimento di Biologia vegetale dell'Università degli Studi di Torino ma destinato, inoltre, alla pubblicazione sul web.
- RoeroLab, un progetto mirato alla conservazione e gestione di beni storico/culturali dove vengono valorizzate e riutilizzate risorse esistenti, favorendo il reperimento di fondi anche tramite soluzioni di crowdfunding.
- Museo Nazionale del Cinema, un prototipo multicanale, con differenti accessi tecnologici, che prevede tre ambiti di sperimentazione: un canale web, un canale PDA per ampliare l'accesso ai contenuti grazie all'uso di dispositivi palmari e un canale di networking, dedicato alla sperimentazione della banda larga.
- Fondazione Torino musei, un'esperienza di progettazione e realizzazione dell'infrastruttura di rete wi-fi per Palazzo Madama a Torino, patrimonio mondiale dell'umanità UNESCO, nonché sede del Museo Civico d'Arte Antica.

- Centro conservazione e restauro di Venaria Reale, in cui è stato offerto il supporto nella progettazione di una banca dati condivisa sui materiali pittorici, parte del più ampio Progetto CE.R.MA. (Centro per la ricerca sui materiali pittorici).
- Multitouch e TouchEmeroteca, due applicazioni per superfici interattive di varie dimensioni, dedicate a rispondere a molteplici esigenze che spaziano dai servizi bibliotecari come la lettura di quotidiani su dispositivi dotati di touch screen, agli strumenti didattici per i disabili, dal settore turistico alla sicurezza, ampliando le possibilità di interazione uomo-macchina.

La sperimentazione illustrata in questa sede, invece, è stata realizzata direttamente all'interno del laboratorio CSP, dove è stata implementata un'istanza di Ubuntu Server, abilitando la virtualizzazione con KVM e installando i vari software descritti in precedenza, opportunamente configurati all'interno di alcune virtual machine, al fine di offrire un'idea di massima delle funzionalità applicative della nostra architettura. Una particolare attenzione è stata prestata alle caratteristiche di fault tolerance e di scalabilità "a caldo", oltre alle capacità di memorizzazione e di analisi di dati eterogenei provenienti da differenti tipi di device, permettendo di ottenere un sistema resistente e affidabile al tempo stesso. Per il limite contingente dell'hardware impiegato – un solo server fisico – e la gestione dell'intera infrastruttura operata in modalità virtuale, non ci è stato possibile addentrarci in analisi particolarmente avanzate in termini di moli di dati esaminate, tuttavia i primi test ci hanno offerto risultati molto interessanti che aprono la strada a nuove sperimentazioni in questo ambito applicativo, cruciale per la business intelligence e per la gestione dei dati scientifici. È inoltre da segnalare che l'architettura in questione funziona perfettamente su commodity hardware e non comporta dei requisiti tecnici particolari.

RODOLFO BORASO, DIEGO GUENZI
CSP – Innovazione nelle ICT

ABSTRACT

In this article we discuss a possible architecture to store, analyze and share large quantities of data (also known as Big Data). This data can be structured or not and can be generated from different sources: mobile devices, servers or, in general, sensors. This infrastructure can be integrated with analysis software like R, Solr and many others; it is developed with the cloud computing and its principles in mind so it can scale out easily and is high available. It is based on the Hadoop framework and, in particular, on the HDFS / HBase store mechanisms (a distributed file system and a NOSQL database). Those repositories are exported with Fuse and are published with a typical Apache HTTP server with the addition of WebDav module. Fault tolerance and load balancing is achieved with HAProxy and Keepalived. All those software are free and open source (FOSS), can be installed on commodity hardware and are constantly updated by a large community of users.