Blockchain e distributed ledger: di cosa stiamo parlando?

Author : Fabrizio Baiardi

Date : 4 Luglio 2019

Blockchain sta diventando sempre più un termine in grado di scatenare amori folli e altrettanti odî sconfinati. Raramente una tecnologia informatica ha generato reazioni così forti. Ma innanzitutto dobbiamo capire quale sia l'**oggetto di amore o di odio**.

Quindi dobbiamo distinguere nettamente tra *distributed ledger*, che è una tecnologia, e *blockchain*, che è una particolare base di dati fondata sulla tecnologia. La *distributed ledger* è tecnologia che genera tutta una classe di basi di dati che possono resistere a malfunzionamenti dovuti a guasti o a comportamenti maliziosi.

Le **proprietà interessanti** delle basi di dati basate su *distributed ledger*, in breve, sono:

a) possono essere estese con nuovi dati ma, grazie a meccanismi crittografici, non permettono di modificare i dati già inseriti. Ad esempio, si inserisce in ogni elemento del *ledger* il codice *hash* di quelli precedenti. Questo equivale a inserire nell'elemento la versione estremamente compressa della base di dati precedente;

b) sono replicate tra più gestori, ognuno dei quali può proporre nuovi elementi da inserire. L'elemento da inserire viene scelto tra quelli proposti dai vari gestori;

c) la scelta dell'elemento da inserire si basa su un meccanismo di consenso che garantisce che tutti i gestori ben funzionanti scelgano lo stesso elemento e quindi modifichino nello stesso modo la base di dati.

Di conseguenza, la base di dati può solo crescere in un ciclo ripetuto all'infinito in cui vengono proposti degli elementi dai vari gestori, uno di essi viene scelto, tutti i gestori ben funzionanti lo inseriscono nella loro copia del *ledger*. I **gestori malfunzionanti** faranno a piacer loro, ma questo non invalida quanto fatto dagli altri. Il termine "malfunzionanti" va inteso sia in senso di *safety* che di *security*, quindi un gestore può non funzionare correttamente a causa di un guasto o di un attacco informatico di cui è stato bersaglio. Il *distributed ledger* fallisce sempre e comunque quando il numero di gestori malfunzionanti o maliziosi supera una soglia che dipende dalle specifiche soluzioni utilizzate.

La lettura della base di dati avviene interrogando un certo numero di copie - cioè di gestori - e poi scegliendo la risposta restituita dalla maggioranza dei gestori. Il numero dei gestori da interrogare dipende dalla soglia sul numero di gestori che possono essere malfunzionanti.

Se non interessa costruire un sistema in grado di tollerare guasti o attacchi informatici, sicuramente la tecnologia *distributed ledger* non interessa, perché esistono soluzioni più semplici ed efficienti per gestire una base di dati replicata. Queste soluzioni non possono però in molti casi tollerare guasti e attacchi. La **resilienza** rispetto ad attacchi e guasti è quindi discriminante ed è più influente delle prestazioni possibili. Trascurare questa resilienza vuol dire lasciare fuori dal dibattito mezzo problema e questo non permette di raggiungere conclusioni ben fondate.

Ovviamente anche qui vale il teorema "*No Free Lunch*" e quindi la resilienza ha necessariamente un **costo**, che aumenta con la soglia sul numero di gestori guasti che si intende tollerare. Un vantaggio dei *distributed ledger* su cui riflettere è che impediscono la strategia, purtroppo molto popolare, che rinvia la sicurezza e la resilienza "alla prossima versione". In un contesto in cui minacce ibride e *dottrina Gerasimov* non ci aiutano a capire se viviamo in pace o in conflitti non dichiarati, una soluzione informatica che abbia intrinseche proprietà di resilienza dovrebbe essere vista come un vantaggio, non come un problema. In altre parole, i *distributed ledger* non sono una tecnologia in cerca di un problema, perché tutto manca meno i problemi a cui applicare la tecnologia.

Prima di parlare di *bitcoin* dobbiamo ancora distinguere tra versioni *permissioned* e *not permissioned* dei *ledger*. Un *distributed ledger* è ***permissioned*** quando i gestori sono noti e decisi a priori. Ad esempio, alcune università stanno creando un *distributed ledger* con i diplomi che ognuna ha concesso. In questo caso, ogni università ha il ruolo di gestore e, ovviamente, organizzazioni pubbliche o private diverse dalle università non possono essere gestori in quanto non coinvolti nel processo di rilascio dei titoli di studio. Un *distributed ledger* è ***not permissioned*** quando chiunque può assumere il ruolo di gestore. In questo caso, il numero di gestori cambia dinamicamente e dipende da ragioni economiche o di altro tipo. Anche se il caso *not permissioned* è più attraente per anarchici e spiriti liberi, raramente organizzazioni con vincoli economici e sociali lo accetteranno: non è ad esempio molto credibile che uno Stato sovrano decida di appaltare il *ledger* con il risultato delle sue elezioni a un insieme di gestori sconosciuti e stabiliti in altri Paesi.

La *blockchain* di *bitcoin* è un *distributed ledger* di tipo *not permissioned*, in cui il meccanismo di consenso è basato sul lavoro e chiunque può diventare gestore. Il primo gestore che dimostra di aver eseguito un certo lavoro è quello che ha il diritto di scegliere il nuovo elemento da inserire nella *blockchain*. In *bitcoin* il lavoro è un calcolo complesso che richiede un tempo non banale. Se *n* gestori eseguono in parallelo un calcolo per ottenere il diritto di inserire un elemento è chiaro che *n-1* di questi calcoli saranno inutili e sarà sprecata l'energia utilizzata. Sicuramente questa soluzione non è ecologica ed è anche lenta.

Questo dimostra solamente che **la *blockchain* di *bitcoin* non rispetta i principi della buona ingegneria**. Come ricorda Steven Bellovin *"Bitcoin was deployed by enthusiasts who in essence let experimental code escape from a lab to the world, without thinking about the*

engineering issues — and now they're stuck with it. Perhaps another, better cryptocurrency can displace it, but it's always much harder to displace something that exists than to fill a vacuum”.

Ma la *blockchain* di Bitcoin è solo uno dei *distributed ledger* possibili e il suo algoritmo di consenso è solo uno degli algoritmi possibili. Altri algoritmi sono più “ecologici” e richiedono tempi di elaborazione minori. Un possibile esempio è un algoritmo di tipo *byzantine fault tolerance* che permette di tollerare comportamenti anomali in al più un terzo dei gestori. Altri algoritmi di consenso utilizzano una votazione in cui ogni gestore riceve i voti di un sottoinsieme di altri gestori. La votazione continua fino al raggiungimento di una maggioranza. Di nuovo, il numero di votazioni dipende dal numero di gestori guasti che si intende tollerare. Tutte queste soluzioni hanno prestazioni e consumi energetici diversi da quelle della *blockchain*. Quindi usare la *blockchain* come modello del consumo energetico della tecnologia dei *distributed ledger* è solo un esempio di *strawman fallacy*.

Un’altra **obiezione** che viene avanzata è l’impossibilità di modificare una informazione erronea se inserita in un *distributed ledger*. Il problema esiste, anche se ovviamente il fatto che l’informazione inserita sia erronea non dipende da dove venga inserita ma da chi la raccolga.

Comunque, dal punto di vista puramente informatico, il **problema di aggiornare informazioni** in una struttura dati che non può essere aggiornata è già stata affrontato. Consideriamo, ad esempio, un certificato di chiave pubblica: il fatto che debba essere aggiornato viene affrontato mediante strumenti come le liste di revoca o simili. Nel caso del *distributed ledger* possiamo inserire un nuovo elemento collegato a quello da aggiornare e prevedere che il *software* di accesso al *ledger* controlli se esistano modifiche in elementi successivi. Se un documento viene inserito in un *ledger* e poi si decide di modificarlo - si pensi a uno stato di famiglia o a un certificato di residenza - se ne inserirà una nuova versione in un elemento successivo. Possiamo anche inserire in una informazione nel *ledger* una data di scadenza, come per il latte.

Riassumendo, se il problema prevede la cooperazione tra più organizzazioni, se una organizzazione può essere fraudolenta o soggetta a guasti, se si vuole privilegiare la capacità di resistere ai guasti e agli attacchi rispetto alla velocità di inserzione di nuovi elementi, se è necessario dover ricostruire la storia delle operazioni eseguite allora la tecnologia dei *distributed ledger* diventa fondamentale e, spesso, senza alternative.

Non è casuale che **uno dei primi campi di utilizzo** sia quello delle *supply chain* per ricostruire la storia di un certo pezzo di ricambio o di un alimento. In una *supply chain* cooperano fornitori diversi, sono necessarie garanzie di sicurezza e tolleranza ai guasti per resistere a fornitori maliziosi o a furti di materiale. Un altro campo di applicazione estremamente interessante è la costruzione di inventari intelligenti di sistemi o infrastrutture informatiche complesse. La costruzione di questi inventari è spesso il primo passo di un’organizzazione per garantirsi il controllo del proprio sistema informativo, problema di non facile soluzione.

Ma vi sono interessanti applicazioni anche nella **sanità pubblica**, perchè l’esistenza di un *database* replicato che garantisca sicurezza e che sia resistente rispetto a guasti e attacchi è indubbiamente fondamentale, molto più importante del tempo necessario per inserire un *health-record*, che rimane comunque accettabile. Inoltre, l’uso di un singolo *database* che permetta a

ospedali diversi di accedere a informazioni sullo stesso paziente minimizza la probabilità di errori in diagnosi o cure. I problemi di compatibilità tra *database* diversi gestiti da più ospedali restano ancora molto seri. Ad esempio, pochi giorni fa un giornale USA ricordava questi problemi con il significativo titolo di "*Death by a thousand clicks*" dove le migliaia di *click* sono quelle di chi tenta di accedere, senza successo, alle informazioni.

Se la preoccupazione è invece quella della ***privacy***, ricordiamo che in questo caso deve essere utilizzato un *permissioned distributed ledger*, le cui informazioni possono essere cifrate e che può anche conservare informazioni su chi ha avuto accesso a cosa. Il fatto che l'informazione sugli accessi sia, per definizione, non modificabile dovrebbe essere visto come un vantaggio.

Infine, l'uso di un *ledger* nella **pubblica amministrazione** può semplificare molte pratiche e ridurre lo scambio di informazioni, nel rispetto delle norme sulla protezione dei dati personali. Infatti, una amministrazione non ha ragione di richiedere un documento che può agevolmente ritrovare nel proprio *ledger* o in un *ledger* di una diversa amministrazione. Inoltre, il *ledger* può fornire in maniera nativa informazioni sulla storia di una pratica, su chi ha elaborato dei documenti, li ha emessi e via dicendo. Queste informazioni sono fondamentali per analizzare e ricostruire i processi della P.A., verificarne l'efficacia e l'efficienza. Può essere vero che la digitalizzazione di molte P.A. sia ancora lontana da una situazione che permetta l'adozione di *distributed ledger* ma non è nemmeno detto che i processi di digitalizzazione della P.A. debbano necessariamente ripercorrere tutto lo sviluppo tecnologico.

È possibile, forse auspicabile, che si facciano dei salti, evitando tecnologie che in altri settori sono già state superate.

Articolo a cura di **Fabrizio Baiardi** e **Cosimo Comella**