



## Un'indagine dedicata ad Apache Spark rivela un aumento della crescita di utenti e di nuovi carichi di lavoro tra cui Exploratory Data Science e machine learning

*Il 57% degli intervistati menziona Cloudera come la piattaforma Spark preferita per i propri casi d'uso più importanti*

**Milano, 21 novembre 2016** – Al fine di comprendere meglio il ruolo crescente di [Apache Spark](#) nell'ambito dei big data, Taneja Group ha condotto un importante progetto di ricerca di mercato su un campione di circa 7.000 persone, composto da ruoli tecnici e manageriali direttamente coinvolti in progetti big data in tutto il mondo. L'indagine, che ha ricevuto un enorme riscontro, ha preso in esame le esperienze con l'ipotesi di adozione e di distribuzione di Spark, oltre alle percezioni attuali, dei maggiori fornitori e del futuro di Spark in sé. [Cloudera](#), fornitore globale della piattaforma di analisi e gestione dei dati più veloce, semplice e sicura basata su [Apache Hadoop™](#) e delle più recenti tecnologie open source, sponsor di questo progetto di ricerca, ha presentato i risultati dello studio.

Parte integrante della soluzione CDH, supportato da Cloudera Enterprise, Spark è lo standard aperto per l'elaborazione flessibile dei dati in-memory che consente analisi batch avanzate e in tempo reale sulla piattaforma Apache Hadoop.

“Apache Spark si è evoluto rapidamente in uno dei principali progetti open source nell'ambito dei big data”, ha dichiarato Mike Matchett, analista senior e consulente presso Taneja Group. “Abbiamo scoperto che, analizzando una vasta gamma di settori, dimensioni aziendali e livelli di maturità dei big data rappresentati differenti, oltre la metà degli intervistati sta già utilizzando attivamente Spark. Si sta rivelando prezioso: il 64% di coloro che attualmente utilizza Spark sta pianificando di aumentarne notevolmente l'utilizzo entro i prossimi 12 mesi. Con un numero crescente di carichi di lavoro che richiedono lo streaming dei dati in tempo reale per l'analisi, l'emergere di applicazioni di apprendimento automatico e di casi d'uso in ambito data science, la piattaforma Spark è destinata a imporsi con l'obiettivo di dare seguito a queste richieste”.

### **Spark: la leadership di Cloudera**

Cloudera è diventata il primo fornitore di Hadoop a distribuire e supportare Spark all'inizio del 2014, quando è rapidamente diventato il riferimento per un'elaborazione batch più veloce. Cloudera ha investito fin da subito al suo sviluppo. E oggi molti utenti di Cloudera hanno trasferito i carichi di lavoro di elaborazione dei dati da MapReduce a Spark all'interno dei loro sistemi di produzione, riducendo drasticamente i tempi di elaborazione dei dati. Secondo il sondaggio, inoltre, questa tendenza sta accelerando.

I clienti di Cloudera richiedono Spark su scala aziendale, supportato da esperti già coinvolti nel processo iniziale volto a rendere la piattaforma, di fatto, il motore di elaborazione dei dati per Hadoop. Cloudera, inoltre, continua a innovare attraverso il programma [One Platform Initiative](#), volto a migliorare le funzionalità di Spark rispetto a gestione, sicurezza, scalabilità, streaming e cloud. Attraverso questa iniziativa, Cloudera è impegnata ad aiutare l'ecosistema nell'adozione di Spark in qualità di motore di esecuzione dei dati di default per i carichi di lavoro analitici.

Cloudera, inoltre, collabora con i propri partner al fine di certificare nuove soluzioni basate su Spark e fornisce le risorse e il supporto necessari per portare queste soluzioni differenziate sul mercato più rapidamente, assicurando che i clienti possano risolvere nuovi e impegnativi casi d'uso.

### **I risultati della ricerca**

I principali risultati dello studio [Apache Spark Market Research Study](#) comprendono un elevato livello di crescita e di slancio nell'utilizzo di Spark al di là dei carichi di lavoro di elaborazione dati/ETL/progettazione e di un futuro passaggio alle distribuzioni in cloud. Altri risultati degni di nota sono:

- quasi la metà degli intervistati, il 54%, sta già utilizzando attivamente Spark. Di quelli che attualmente stanno utilizzando Spark, il 64% afferma che la piattaforma si sta dimostrando preziosa e che intendono ampliare l'uso di Spark entro i prossimi 12 mesi;
- anche l'adozione da parte di nuovi utenti Spark è in crescita con 4 persone su 10 esperti di progetti big data che prevedono di distribuire Spark a breve;
- il 57 % si affida a Spark, fornito da Cloudera, per i casi di utilizzo più importanti, oltre il doppio dei tre fornitori di Apache Hadoop più prossimi combinati insieme. I clienti che hanno scelto Cloudera rispetto ad altre soluzioni hanno evidenziato quali importanti elementi il suo modello in ambito sicurezza e governance conforme alle normative, stabilità e prestazioni, portabilità cloud e integrazione con una suite completa di servizi per l'elaborazione di dati, query, analisi e machine learning;
- a parte i carichi di lavoro previsti di elaborazione dei dati/progettazione/ ETL che costituiscono il 55 % dell'uso di Spark a oggi, le principali iniziative attive di Spark includono l'elaborazione di flusso in tempo reale, la scienza esplorativa dei dati e l'affermazione di Spark per l'apprendimento automatico, tutte aree in cui Cloudera continua a investire;
- le barriere per l'adozione e le problematiche rimangono però le stesse e sono in gran parte attribuite al gap di competenze nell'ambito dei big data e alla possibilità di ottenere una formazione adeguata in diversi formati (online, di persona, conferenza o fiera). Cloudera forma più professionisti Apache Spark rispetto a qualsiasi altro fornitore di Hadoop e li sostiene attraverso servizi professionali dedicati, consulenza e un ampio spettro di partner.

“In Cloudera siamo focalizzati sulla leadership aziendale e mettiamo a disposizione quella sicurezza fondamentale, governance dei dati e conformità di cui i nostri clienti hanno bisogno,” ha dichiarato Mike Olson, fondatore e chief strategy officer di Cloudera. “I risultati del sondaggio confermano l'importanza attribuita alla fornitura di soluzioni enterprise e alla capacità di supportare i futuri casi d'uso di Spark: si

tratta della ragione fondamentale per cui i clienti scelgono in grande maggioranza la soluzione Spark di Cloudera rispetto a quella di altri fornitori”.

L'indagine descrive, inoltre, il ruolo superiore del cloud pubblico e di Spark: “È interessante notare che, mentre le distribuzioni di Spark on-premise sono oggi dominanti, c'è un forte interesse nell'effettuare la transizione di molte di queste distribuzioni nel cloud”, ha dichiarato Matchett. “Nel complesso la distribuzione di Spark nel cloud privato e pubblico (IaaS o PaaS) dovrebbe aumentare in modo significativo dal 23% al 36% in futuro”.

Cloudera ha creato un'[infografica](#) con il dettaglio dei risultati del report.

### Per maggiori informazioni

- Scarica l'executive summary [The Apache Spark Market Research Study](#)
- Consulta l'[infografica](#) su Apache Spark
- Scarica [Spark 2.0 Beta Parcel](#)
- Prendi in esame Cloudera come [piattaforma di Data Engineering](#)

### A proposito di Taneja Group

Taneja Group è una società di analisti che fornisce consulenza e analisi per il settore tecnologico. Le nostre attività di ricerca e consulenza sono rivolte ai fornitori di tecnologia, utenti finali di soluzioni IT e venture capitalist. Gli analisti del gruppo Taneja studiano le tecnologie nelle seguenti aree: tutti gli aspetti di archiviazione, virtualizzazione dei server, ottimizzazione WAN, accelerazione di storage e applicazioni, eDiscovery e corporate governance. Storage cloud, Big Data e convergenza del data center sono inclusi nei segmenti citati. Il data center sta subendo una metamorfosi fondamentale e gli analisti del gruppo sono in prima linea per i clienti per consigliare le tecnologie determinanti e quando dovrebbero essere implementate per la massima efficacia. Per il settore eDiscovery analizziamo tutti gli aspetti del flusso dei contenitori, nonché i processi di business correlati, tra cui la governance, la compliance, la gestione dei record e della conservazione dei dati.

###

### A proposito di Cloudera

Cloudera offre la piattaforma per la gestione e l'analisi dei dati aziendali basata su Apache Hadoop e le più recenti tecnologie di open source. Le principali aziende di tutto il mondo si affidano a Cloudera per affrontare le principali problematiche di business con Cloudera Enterprise, la piattaforma dati più veloce, semplice e sicura disponibile ad oggi. I clienti di Cloudera sono in grado di acquisire, archiviare, elaborare e analizzare efficientemente una vasta quantità di dati per utilizzare analisi avanzate al fine di elaborare decisioni di business più velocemente, in modo più flessibile e con costi inferiori rispetto al passato. Per assicurare il successo dei propri clienti, Cloudera offre ampio supporto, sessioni di training e servizi professionali. Ulteriori informazioni sono disponibili su <http://cloudera.com>.

### Per saperne di più:

Web: [cloudera.com](http://cloudera.com)

Blog: [cloudera.com/blog](http://cloudera.com/blog) e [vision.cloudera.com](http://vision.cloudera.com)

Twitter: [twitter.com/clouderaITA](https://twitter.com/clouderaITA)

Facebook: [facebook.com/cloudera](https://facebook.com/cloudera)

LinkedIn Group: [linkedin.com/Cloudera Italia](https://linkedin.com/Cloudera Italia)

Cloudera Community: [community.cloudera.com](https://community.cloudera.com)

**Informazioni stampa**

Deborah Wiltshire

Cloudera

+1 (650) 644-3900 ext. 5907

[press@cloudera.com](mailto:press@cloudera.com)

Tania Acerbi, Francesco Petruzzi, Daniela Rimicci

Prima Pagina

0291339811

[cloudera@primapagina.it](mailto:cloudera@primapagina.it)