



## **Cloudera lancia Kudu, un nuovo sistema di storage Hadoop per una rapida analisi dei dati**

*Il nuovo progetto Open Source per abilitare le applicazioni analitiche in real-time all'interno di Hadoop disponibile come Beta pubblica*

**Milano, 14 ottobre 2015** – [Cloudera](#), leader nella gestione dei dati analitici enterprise basata su piattaforma [Apache Hadoop™](#), annuncia il rilascio della beta pubblica di Kudu, un nuovo rivoluzionario sistema di storage basato su colonne (columnar store) per Hadoop che abilita una potente combinazione di rapide analitiche per i fast data. Complementare alle esistenti opzioni di storage Hadoop, HDFS e Apache HBase, Kudu è il primo engine di storage nativo Hadoop che supporta sia accesso casuale a bassa latenza e analitiche con throughput elevato, semplificando notevolmente le architetture Hadoop per casi d'uso real-time sempre più comuni. Una versione Beta pubblica di Kudu è già disponibile in base alla licenza open source di Apache e in futuro verrà trasferita nell'incubatore Apache Software Foundation.

Fino ad oggi, gli sviluppatori erano costretti a scegliere tra analitiche rapide con HDFS o aggiornamenti efficienti con HBase. In particolare, con l'aumento dei dati di streaming, si è verificata una crescente richiesta di combinare le due funzioni per creare applicazioni analitiche real-time sui dati da modificare – portando così gli sviluppatori a creare architetture complesse con le opzioni di storage disponibili. Kudu è complementare alle funzionalità di HDFS e HBase, fornendo inserimenti e aggiornamenti rapidi e scansioni colonnari efficienti. Questa potente combinazione abilita i carichi di lavoro analitici in real-time con un singolo livello di storage, eliminando l'esigenza di architetture complesse.

“Abbiamo migliorato Hadoop fin da subito”, spiega Charles Zedlewski, vice president, products, Cloudera. “La nostra è una missione ambiziosa: stimolare costantemente l'innovazione all'interno della comunità per introdurre la nuova generazione di analitiche supportate da Hadoop, in modo che le aziende possano adattarsi alle tecnologie più recenti. Cloudera ha già trasformato quanto è possibile con Hadoop – permettendo le attività di discovery interattiva e analitiche dei dati con Impala e l'elaborazione flessibile dei dati e lo streaming degli stessi con Apache Spark. Kudu prosegue in tal senso rivoluzionando l'architettura di storage di Hadoop

per supportare meglio lo sviluppo di applicazioni analitiche real-time e funge da passaggio cruciale verso il consolidamento di Hadoop quale piattaforma principale per le analitiche moderne”.

L’architettura di Kudu ottimizza l’esperienza dello sviluppatore nella creazione di applicazioni analitiche, supportando casi d’uso comune che includono l’analisi delle serie temporali, le analitiche dei dati macchina e la reportistica online. Inoltre, Kudu è progettato per sfruttare i mutevoli trend in ambito hardware ed elaborazione in-memory. Offre prestazioni eccellenti della CPU, sfrutta RAM e Flash, e migliora l’efficienza in ambito I/O.. Infine, poiché è un componente aperto nativo all’interno di Hadoop, Kudu è integrato e offre prestazioni più rapide in fase di query per i framework analitici più potenti. Gli utenti fanno già affidamento su molti di questi, tra cui Impala e Spark, per le applicazioni analitiche end-to-end su una singola piattaforma.

Kudu è stato progettato congiuntamente da Cloudera e Intel in anticipo rispetto al mutevole panorama hardware. Intel ha contribuito attivamente a Kudu per fare in modo che sfruttasse appieno le tecnologie di memoria e processore di Intel, attuali e future. Kudu è stato progettato per utilizzare nuove innovazioni di persistent memory (pmem) sviluppate attraverso il progetto pmem di Intel.

“Man mano che le analitiche di Hadoop evolvono, è fondamentale che vengano progettate tenendo ben presente l’hardware di nuova generazione”, afferma Vin Sharma, Director of Strategy & Products for Big Data Analytics di Intel. “Kudu è una tappa fondamentale critica per Hadoop, poiché supporta la crescente necessità di applicazioni real-time semplificate. Intel ha collaborato con Cloudera e la comunità per assicurare che Kudu sia ottimizzato per offrire prestazioni di rapida analisi oggi, ma anche che sia strutturato per utilizzare i miglioramenti delle piattaforme Intel che saranno rilasciati in futuro, come i moduli Intel DIMM con memoria 3D XPoint”.

Poiché si tratta di un progetto open source, Kudu ha attirato un grande coinvolgimento da parte della comunità. Xiaomi, uno dei principali sviluppatori di smartphone al mondo, è stato uno dei primi utenti beta di Kudu contribuendo attivamente al progetto. Altre aziende, tra cui AtScale, Splice Machine e Zoomdata, hanno contribuito allo sviluppo di Kudu.

“Xiaomi è un utente e sostenitore di lunga data dell’ecosistema Hadoop, e lo ha utilizzato per alimentare un’ampia gamma di casi d’uso nel proprio ambito di business” ha affermato Baoqiu Cui, Chief Architect di Xiaomi. “Il nostro gruppo dedicato all’infrastruttura ha lavorato con Cloudera per sviluppare Kudu, sfruttando la sua capacità unica di supportare le scansioni colonnari, inserimenti e aggiornamenti rapidi per continuare ad ampliare la portata del nostro ecosistema

Hadoop. L'utilizzo di Kudu, insieme a strumenti SQL interattivi come Impala, ci ha consentito di creare una piattaforma di analisi dei dati di nuova generazione per analisi e reporting online in real-time. Siamo entusiasti di continuare a lavorare con la comunità per dare un ulteriore impulso a Kudu e alle funzionalità di Hadoop nel suo complesso".

"Kudu abilita effettivamente la nuova generazione di architetture analitiche, soprattutto per la Business Intelligence (BI). Con il suo supporto per elevati volumi di lettura e scrittura sia casuali che sequenziali, è il sistema di storage ideale per architetture di BI a bassa latenza scale-out del tipo richiesto dai clienti di AtScale. Poiché le aziende desiderano democratizzare l'accesso ai dati e consentire a Hadoop di eseguire carichi di lavoro analitici veloci e su larga scala, Kudu ricoprirà un ruolo critico", afferma Josh Klahr, VP of Product Management di AtScale. "Forte sostenitore dell'ecosistema open source Apache Hadoop, AtScale è parte di questo impegno della comunità ed è lieta di poter contribuire al suo ulteriore sviluppo a favore dei propri clienti".

"Siamo entusiasti di far parte della comunità Kudu," ha affermato John Leach, co-founder and CTO, Splice Machine. "In Splice Machine abbiamo sviluppato un RDBMS conforme ACID che opera su Hadoop e stiamo andando oltre i limiti in termini di esecuzione di carichi di lavori misti su Hadoop. Di conseguenza, apprezziamo e supportiamo l'innovazione nell'architettura storage di Hadoop. Kudu è incredibilmente promettente in termini di capacità di gestire aggiornamenti in real-time unitamente ad analitiche prolungate. Rafforza l'ecosistema Hadoop fornendo un engine di storage scalabile alternativo e complementare a quelli esistenti".

"Kudu offre un'architettura storage semplificata per casi d'uso che sono piuttosto comuni tra gli utenti di Zoomdata", afferma Justin Langseth, CEO di Zoomdata. "Quale componente nativo di Hadoop, l'integrazione di Kudu con Impala e Spark semplifica l'apertura di questi dati utilizzando la soluzione di rapida analisi visiva di Zoomdata. Abbiamo collaborato fianco a fianco di Cloudera e della comunità per contribuire a sviluppare Kudu a soddisfare le esigenze dei nostri utenti - supportando la combinazione ottimizzata di applicazioni analitiche e in real-time - e siamo lieti di proseguire con quest'impegno con il rilascio della versione beta pubblica".

Per far sì che le aziende continuino a trarre beneficio da approfondimenti data-driven, l'architettura di Hadoop deve operare alla stessa velocità, sempre maggiore, con cui i dati vengono creati e modificati. Con Kudu, la comunità Hadoop dà inizio alla prossima generazione di applicazioni Hadoop con storage per rapide analisi dei fast data.

"Nell'era dei dati generati da computer, è sempre maggiore l'esigenza di analizzare i dati in tempo reale. Ciò è vero per un'ampia gamma di casi d'uso d'analisi, dal monitoraggio e business intelligence alla modellazione e segnalazione predittiva", ha affermato Curt Monash, presidente di [Monash Research](#). "Kudu, Spark e il resto dello stack Hadoop rappresentano un approccio promettente per soddisfare tali esigenze".

## **Risorse**

Per contribuire - [<http://getkudu.io/contributing.html>]

Per scaricare la Beta pubblica o provare la VM - [[www.cloudera.com/downloads](http://www.cloudera.com/downloads)]

Whitepaper tecnico - [[getkudu.io/kudu.pdf](http://getkudu.io/kudu.pdf)]

## **Per saperne di più:**

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

Blog: <http://www.cloudera.com/blog/> e <http://vision.cloudera.com/>

Twitter: <http://twitter.com/clouderaita>

Facebook: <http://www.facebook.com/cloudera>

Cloudera Community: <http://cloudera.com/community>

###

## **A proposito di Cloudera**

Cloudera rivoluziona la gestione dei dati aziendali offrendo la prima piattaforma unificata per i big data, un hub di dati aziendale basato su Apache Hadoop. Cloudera offre alle aziende uno spazio per archiviare, consultare, elaborare, proteggere e analizzare tutti i dati, consentendo loro di estendere il valore degli investimenti esistenti abilitando al contempo nuove modalità essenziali per ricavare valore dai propri dati. La piattaforma open source per i big data di Cloudera è la più diffusa in tutto il mondo, e Cloudera è tra i sostenitori più prolifici dell'ecosistema open source Hadoop. Con un ruolo fondamentale nella formazione di professionisti Hadoop, Cloudera ha formato oltre 40.000 persone in tutto il mondo. Oltre 1.800 partner e un esperto team di servizi professionali offrono supporto per assicurare un maggiore time to value. Le principali aziende di ogni settore industriale e del settore pubblico utilizzano le soluzioni Cloudera.

## **Informazioni stampa**

Deborah Wiltshire

Sara Argentina, Elena Bottioli

Cloudera  
+1 (650) 644-3900 ext. 5907  
press@cloudera.com

Prima Pagina  
0291339811  
cloudera@primapagina.it

###

*This information is not a commitment, promise or legal obligation to deliver any material, code or functionality. Cloudera does not guarantee that the beta software will be made generally available or that any individual feature in the beta version will be made generally available. Cloudera may make the beta software generally available, or not, in its sole discretion and without obligation to make any communication of any kind with regard to such availability.*