

PROGETTO DI RICERCA

SAMMClouds: Secure and Adaptive Management of Multi-Clouds

Finanziamento: Fondi di Ateneo per la Ricerca (FAR) 2014

Budget di progetto: 28000 euro

Durata del progetto: 18 mesi

Data di inizio del progetto: 1 Giugno 2015

Responsabile scientifico: Claudia Canali

Partecipanti al progetto:

Riccardo Lancellotti, Mirco Marchetti, Fabio Pierazzi, Luca Ferretti

Descrizione

Il Cloud Computing è ormai considerato la tecnologia abilitante per affrontare le sfide della nuova generazione di applicazioni, quali servizi Web 3.0, analisi di Big Data e Internet of Things. La capacità delle infrastrutture di cloud computing di superare nei prossimi anni le sfide legate alla crescente richiesta di efficienza, scalabilità e sicurezza rappresenterà un fattore determinante per il futuro sviluppo della società digitale.

Il progetto SAMMClouds (Secure and Adaptive Management of Multi-Clouds) mira ad analizzare e risolvere alcune delle problematiche connesse all'offerta di servizi integrati di tipo Software as a Service (SaaS) e Infrastructure as a Service (IaaS). Lo scenario di riferimento è un sistema multi-cloud in cui i fornitori di servizi (SaaS) ospitano le loro applicazioni presso molteplici infrastrutture cloud (IaaS). I fornitori SaaS competono fra loro per l'utilizzo delle infrastrutture, mentre i fornitori IaaS hanno l'esigenza di gestire efficacemente la propria infrastruttura. In questo scenario, il monitoraggio delle risorse, la loro gestione a run-time e le garanzie di sicurezza dei dati rappresentano aspetti critici per l'erogazione dei servizi.

Nello specifico, l'attività di ricerca si focalizzerà sullo studio e sulla proposta di soluzioni innovative riguardo a tre problematiche principali:

- Monitoraggio e gestione di sistemi IaaS cloud e multi-cloud
- Gestione delle risorse in sistemi SaaS multi-cloud
- Protezione e sicurezza dei dati in sistemi cloud

Per quanto riguarda il monitoraggio e la gestione di sistemi IaaS, i datacenter attuali si basano su soluzioni che usano euristiche e tecniche di riduzione ed aggregazione dei dati monitorati basati su best practices. Con l'aumentare delle dimensioni e della complessità dei datacenter, si rende necessario adottare tecniche di nuova generazione per garantire la scalabilità dei sistemi. A tal fine, si stanno studiando soluzioni in grado di identificare automaticamente gruppi di macchine virtuali (VM) che mostrano comportamenti simili attraverso tecniche di clustering. Tali similarità possono essere sfruttate per migliorare la scalabilità del sistema di monitoraggio e l'efficienza dell'allocazione delle risorse. Gli obiettivi in questo ambito riguardano sia lo sviluppo di un approccio adattativo al problema del clustering, in grado di lavorare quasi in tempo reale anche in presenza di variazioni dinamiche del carico di lavoro dei data center, sia la proposta di nuove famiglie di algoritmi per la gestione delle risorse, in grado di sfruttare la similarità tra VM per ottenere soluzioni scalabili in termini dei problemi di bin-packing che caratterizzano l'allocazione delle VM in un data center. La ricerca in tale ambito sarà condotta partendo da tracce di sistemi cloud reali e da dati ottenuti da piattaforme cloud implementate su centri di calcolo interni all'università. Le tracce che descrivono il comportamento dell'infrastruttura saranno usate come base sia per la validazione di metodologie di clustering adattative, sia per la ricerca di nuovi algoritmi di bin-packing che utilizzano i risultati del clustering. In quest'ultima fase, saranno usate tecniche di ottimizzazione basate su tecnologie che rappresentano lo stato dell'arte in questo ambito, come il linguaggio AMPL per la modellazione dei problemi di programmazione lineare, intera e mista.

Rispetto ai sistemi SaaS, è da sottolineare come lo scenario multi-cloud in cui diversi fornitori di servizi si appoggiano a più infrastrutture di tipo IaaS sia stato studiato solo parzialmente in letteratura, e si ritiene possa essere di grande interesse per l'evoluzione dei sistemi cloud di domani. Gli obiettivi di ricerca in tale ambito riguardano sia le strategie per suddividere il carico tra i fornitori IaaS in scenari con risorse limitate, sia lo studio di interazioni tra più fornitori SaaS che competono per le stesse risorse. Nell'ambito di tale ricerca si applicheranno strumenti come la teoria dei giochi, per identificare soluzioni di equilibrio di Nash generalizzato nell'ambito delle interazioni tra i fornitori cloud. Inoltre, si intende sviluppare un meccanismo di simulazione di infrastrutture Cloud e multi-Cloud per la valutazione prestazionale delle soluzioni ottenute.

L'approccio della ricerca prevede che il miglioramento delle prestazioni sia strettamente correlato con le garanzie di sicurezza. I dati affidati al cloud costituiscono un asset economicamente importante per i proprietari delle informazioni; una struttura cloud ad alte prestazioni che non offra elevati standard di sicurezza non è pertanto accettabile. Tuttavia, lo stato dell'arte offre livelli di sicurezza piuttosto bassi in quanto soluzioni in grado di offrire una garanzia di sicurezza più elevata sono raramente applicabili a causa del loro

alto costo computazionale. Il progetto si pone come obiettivo lo studio di soluzioni innovative per lo storage sicuro dei dati, che da un lato consentano al proprietario dei dati di manipolare le informazioni con garanzie di privacy elevate (anche rispetto a comportamenti malevoli del cloud provider), e dall'altro consentano comunque di offrire un livello competitivo di prestazioni e di costi del servizio. Risultati preliminari promettenti sono stati ottenuti usando tecniche basate su crittografia parzialmente omomorfa. Tuttavia, rimane un ampio spazio di ricerca sia per migliorare il livello di sicurezza, rendendolo adattivo al tipo di operazioni richieste sui dati, sia per integrare i vincoli che derivano dall'adozione delle nuove soluzioni per la sicurezza nei modelli di gestione dell'infrastruttura. In particolare, si considera estremamente promettente l'estensione dei modelli economici per scenari multi-cloud per quantificare il costo della sicurezza in scenari data-intensive (come big-data analysis o raccolta di informazioni da Internet of Things), in modo da poter studiare il trade-off tra sicurezza e costi aggiuntivi. Strumenti quali i modelli matematici e il simulatore di sistemi cloud sviluppati nell'ambito del progetto saranno affiancati in questa ricerca da soluzioni prototipali che dimostrino la fattibilità delle soluzioni proposte in contesti cloud reali.

Tra i principali risultati attesi del progetto ci sono:

- metodologie per il monitoraggio scalabile di sistemi cloud IaaS, adattive rispetto alle tipologie di carico di lavoro cui l'infrastruttura è sottoposta
- tecniche e algoritmi per l'allocazione di VM all'interno di un singolo data center e tra data center differenti con caratteristiche di scalabilità significativamente superiori rispetto allo stato dell'arte
- modelli matematici per la descrizione delle interazioni tra fornitori cloud sia di tipo IaaS che SaaS
- tecniche e modelli matematici avanzati per la suddivisione del carico tra fornitori SaaS in grado di garantire prestazioni elevate e stabili sia in termini di qualità del servizio sia in termini di costi
- strumenti per la valutazione prestazionale (mediante approcci simulativi, analitici e prototipali) di piattaforme cloud e multi-cloud
- metodologie e algoritmi per garantire la sicurezza dei dati conservati su piattaforme cloud
- modelli per rappresentare i trade-off tra prestazioni e sicurezza in sistemi cloud orientati allo storage di dati, possibilmente distribuiti su più fornitori