

UNIVERSITA' DEGLI STUDI DI
NAPOLI FEDERICO II



Facoltà di Ingegneria

Corso di Studi in Ingegneria Informatica

Elaborato finale in **Reti di calcolatori**

***Piattaforme di monitoraggio per
sistemi “Cloud”***

Anno Accademico 2010-2011

Candidato:

Vincenzo Cinque

matr. N46 / 289

Ringraziamenti

Voglio ringraziare tutti quelli che mi vogliono bene, in particolar modo i miei genitori che mi hanno sempre sostenuto economicamente e moralmente in questo mio percorso; la mia fidanzata che con me ha vissuto sia i momenti di stress, come i giorni che precedono un esame, che quelli di gioia come quando le comunicavo un buon esito.

Infine i miei colleghi - amici di corso - con i quali ho condiviso sforzi ed esperienze durante questi anni.

Indice

Introduzione	4
Capitolo 1. Cloud computing	5
1.1 Caratteristiche principali	6
1.2 Modelli di servizio	7
1.3 Architettura e casi d'uso	8
1.4 Modelli di distribuzione	9
1.5 Vantaggi e perplessità	10
Capitolo 2. Monitoraggio del Cloud	12
2.1 Misurazioni e test	17
2.2 Architettura funzionale di un sistema di monitoraggio	20
2.3 Difficoltà e sfide nel monitoraggio del Cloud	22
Capitolo 3. Soluzioni commerciali e open-source per il Cloud computing	25
3.1 L'ultimo decennio del Computing	25
3.2 Cloud commerciali e le loro soluzioni di monitoraggio	26
3.2.1 Amazon Web Services	26
3.2.2 Microsoft Windows Azure	27
3.2.3 RackSpace	28
3.2.4 Google App Engine	29
3.2.5 Altre piattaforme commerciali	29
3.3 Cloud open-source e le loro soluzioni di monitoraggio	30
3.3.1 Eucalyptus	30
3.3.2 OpenNebula	31
3.3.3 OpenStack	32
3.3.4 Altre piattaforme open-source	33
3.4 Le soluzioni di CloudWatch e Nagios	35
3.5 Tabella comparativa tra le piattaforme di monitoraggio	36
Breve conclusione	37
Bibliografia	38

Introduzione

L'evoluzione tecnologica, compresa quella dei mezzi di comunicazione è inarrestabile. Attualmente, attraverso strumenti come internet, viene offerta una miriade di servizi e di opportunità, con accessi sempre più rapidi e semplici. Grazie alla diffusione di dispositivi mobili sempre più evoluti, come palmari, smartphone e tablet, è facile spostarsi rimanendo sempre immerso nella rete, quasi senza interrompere le attività informatiche quotidiane, che possono spaziare dal semplice svago, come giochi, chat e social network, a impegni lavorativi di ogni specie.

Questa tesi vuole esaminare l'infrastruttura del Cloud Computing che si sta proponendo prepotentemente, come nuovo modello per la fornitura di servizi sul web e, come per ogni nuova tecnologia che emerge, ne descriveremo le caratteristiche salienti, i vantaggi e le potenzialità ma anche le perplessità e gli ostacoli che il suo uso può suscitare.

Passeremo, poi, al vero tema centrale che è il monitoraggio di questi sistemi, valutando i suoi diversi aspetti e implementazioni, secondo le varie prospettive ed esigenze.

Termineremo valutando alcune tra le soluzioni offerte dai maggiori Cloud provider. E, su questo tema, guarderemo da vicino sia le piattaforme offerte nel panorama commerciale che quelle Open-source, soffermandoci e citando alcuni lavori di ricerca centrati su questo tema.

Per raggiungere questo risultato, ho svolto molte ricerche in rete e letto diversi articoli scientifici del settore.

Spero di aver trattato al meglio l'argomento, in maniera semplice e comprensibile ma allo stesso tempo con la giusta forma e completezza.

Capitolo 1

Cloud computing

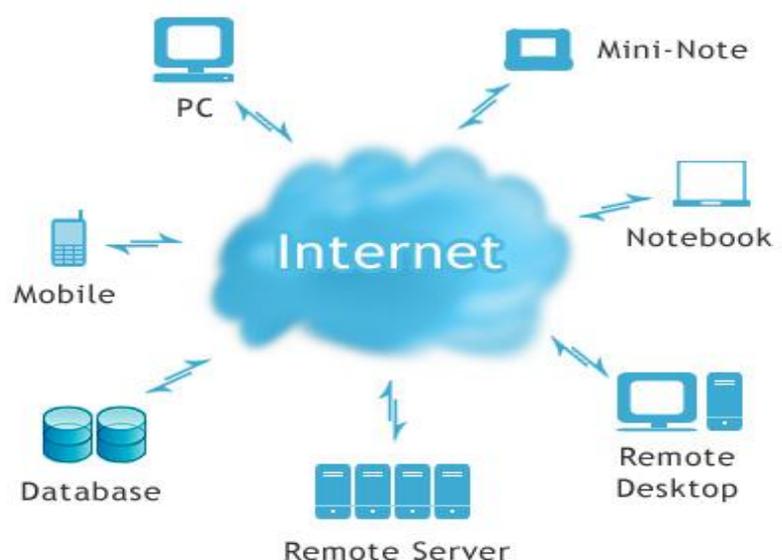
Nella crescente complessità della società dell'informazione vengono prodotti ed elaborati un numero sempre maggiore di dati. Oggi, i nostri elaborati vengono salvati sul computer di casa o dell'ufficio, ma la gestione e la condivisione di una quantità crescente d'informazioni richiede un sistema più flessibile, che ne garantisca il salvataggio, la gestione e la condivisione. La tecnologia Cloud consente di utilizzare qualsiasi tipo di documento senza aver bisogno di chiavette Usb, hard disk e archivi digitali.

Di volta in volta, si potranno acquistare canzoni, noleggiare film, sfogliare libri elettronici, come e quando si vuole, pagando quello che si consuma.

Anziché nel computer, i software sono installati direttamente sulla rete, in una sorta di «nuvola».

I dati che, fino ad oggi, venivano salvati sui pc saranno decentrati nei vari data center: giganteschi archivi digitali a cui l'utente può accedere attraverso il web. In concreto, invece di archiviare i nostri documenti e programmi sul computer di casa, li depositeremo su Internet e potremo accedervi con diversi dispositivi.

Ciò può essere visto come un'opportunità, non solo per semplici utenti del web, che collegati ad un Cloud provider possono svolgere le loro attività tramite un semplice internet browser ma, soprattutto, per imprese piccole e medie nonché sviluppatori di applicazioni. Infatti, il contenimento dei costi, uniti a performance sempre più elevate, sono ormai



concetti di cui tener conto per essere competitivi sul mercato. Attraverso il Cloud si riduce la complessità dell'infrastruttura IT grazie a piattaforme e servizi on-demand, così il contenimento dei costi è immediato grazie al concetto di "pay-per-use". L'idea di base è, dunque, quella di affidare a qualcun altro il compito di gestire l'infrastruttura informatica aziendale, abbattendo la spesa e i tempi di messa in opera e di gestione.

E' bene chiarire da subito, che il Cloud computing non è una nuova tecnologia ma un nuovo modello, se vogliamo, filosofia che rivoluzionerà sia l'uso del web che il mondo dell'IT.

Alla luce di quanto detto, sarà semplice comprendere la seguente definizione tratta dalla più grande enciclopedia online Wikipedia:

*“In informatica con il termine inglese **Cloud computing** si indica un insieme di tecnologie che permettono, tipicamente sotto forma di un servizio offerto da un provider al cliente, di memorizzare/archiviare e/o elaborare dati (tramite CPU o software) grazie all'utilizzo di risorse hardware/software distribuite e virtualizzate in Rete”.*

La virtualizzazione fa riferimento alla possibilità di sfruttare le risorse di una macchina fisica per creare una o più macchine virtuali (VM), in grado di simulare l'hardware fisico di un sistema di elaborazione. Questo consente di dividere tra più utenti, attraverso le varie istanze, l'uso della stessa risorsa computazionale, dando a ognuno l'impressione di esserne l'unico utilizzatore. Attraverso strumenti come i "virtual machine monitor" o "hypervisor" si realizza la virtualizzazione per il consolidamento di più server su di una macchina fisica. Così facendo, questa diviene la principale tecnologia di base per fornire tutti i tipi di servizi Cloud.

Il termine Cloud (nuvola in inglese) proviene dalla seguente considerazione:

Nei diagrammi di rete i collegamenti Internet sono genericamente rappresentati con il simbolo di una nuvola. Il computer A, passando dalla nuvola, chiede dati al computer B. La nuvola quindi rappresenta l'infrastruttura di rete che unisce i due punti, non meglio definita e mutevole. In ultima analisi, il medium attraverso cui i due enti comunicano. La novità è che, con questo nuovo sistema, il computer A chiederà i dati direttamente alla nuvola che, da semplice media di transito, diviene magazzino di dati e applicazioni.

I paragrafi di questo capitolo, sono stati realizzati a partire da informazioni tratte da articoli e forum trovati in rete e nella bibliografia finale sono riportati i link ^[1] ^[2] ^[3], per ulteriori approfondimenti.

1.1 Caratteristiche principali

- **On-demand self-service:** Il consumatore può, unilateralmente, disporre di capacità di calcolo come server time e network storage, se necessario, senza richiedere l'interazione umana con alcun fornitore del servizio.
- **Broad network access:** Le capacità sono disponibili in rete e sono accessibili tramite meccanismi standard che promuovono l'uso eterogeneo di thin o thick client (ad esempio, i telefoni cellulari, laptop e PDA).

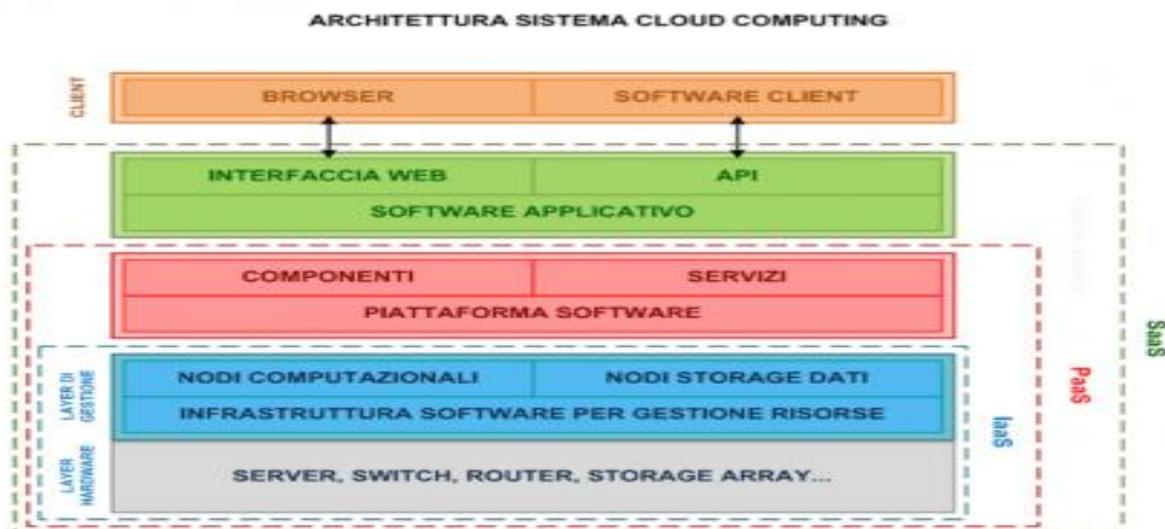
- **Resource pooling:** Le risorse di calcolo sono messe al servizio di tutti i consumatori utilizzando un modello multi-tenant (una singola istanza di un software gira su un server, servendo più clienti: i "tenants"), con diverse risorse fisiche e virtuali dinamicamente riassegnate in base alla domanda dei consumatori. In genere, l'utente non ha alcun controllo o conoscenza dell'esatta posizione delle risorse. Esempi di risorse includono la RAM, la CPU, lo spazio disco, larghezza di banda della rete, e le macchine virtuali.
- **Rapid elasticity:** Le risorse possono essere rapidamente ed elasticamente incrementate per adeguarsi alla potenza di calcolo richiesta e, altrettanto rapidamente liberate quando non più necessarie. Spesso per il consumatore, la capacità delle risorse disponibili sembra essere infinita e può esserne acquistata in qualsiasi quantità, in qualsiasi momento.
- **Measured Service:** I Sistemi Cloud controllano e ottimizzano automaticamente le risorse, misurandole adeguatamente (ad esempio, lo storage, il processing, la banda, e gli utenti attivi). L'utilizzo può essere monitorato, per offrire trasparenza al consumatore e adattabilità del servizio alle varie esigenze e situazioni.

1.2 Modelli di servizio

Si sono diffusi tre modelli di servizio che, solitamente, vengono forniti dalle aziende del settore:

- **IaaS** (Infrastructure as a Service)
- **PaaS** (Platform as a Service)
- **SaaS** (Software as a Service)

Essi si distinguono per il livello a cui il servizio fornisce all'utente finale l'accesso all'architettura Cloud. Nell'immagine che segue, è schematizzata l'architettura di un sistema di Cloud computing, con evidenziate le componenti presenti nei tre servizi nominati:



Dallo schema è possibile constatare come gli IaaS siano i servizi che permettono di avere l'accesso al livello più basso del Cloud. Un'azienda fornitrice di IaaS mette a disposizione un'infrastruttura hardware remota, fornendo la documentazione necessaria affinché la piattaforma software del cliente possa utilizzarne le risorse. Tipicamente, l'hosting provider crea un ambiente virtuale che ospita su un proprio server. In questa macchina virtuale vengono caricate e configurate tutte le risorse hardware e software richieste dal cliente, ma le spese per le licenze sono a suo carico. I clienti sono generalmente aziende che si affidano a questi servizi per costituire parte del proprio sistema informatico.

Sono identificabili come PaaS, invece, tutti i servizi che mettono a disposizione una piattaforma che può essere costituita da diversi ambienti, programmi, librerie, framework preinstallati, pronta per il deployment delle applicazioni.

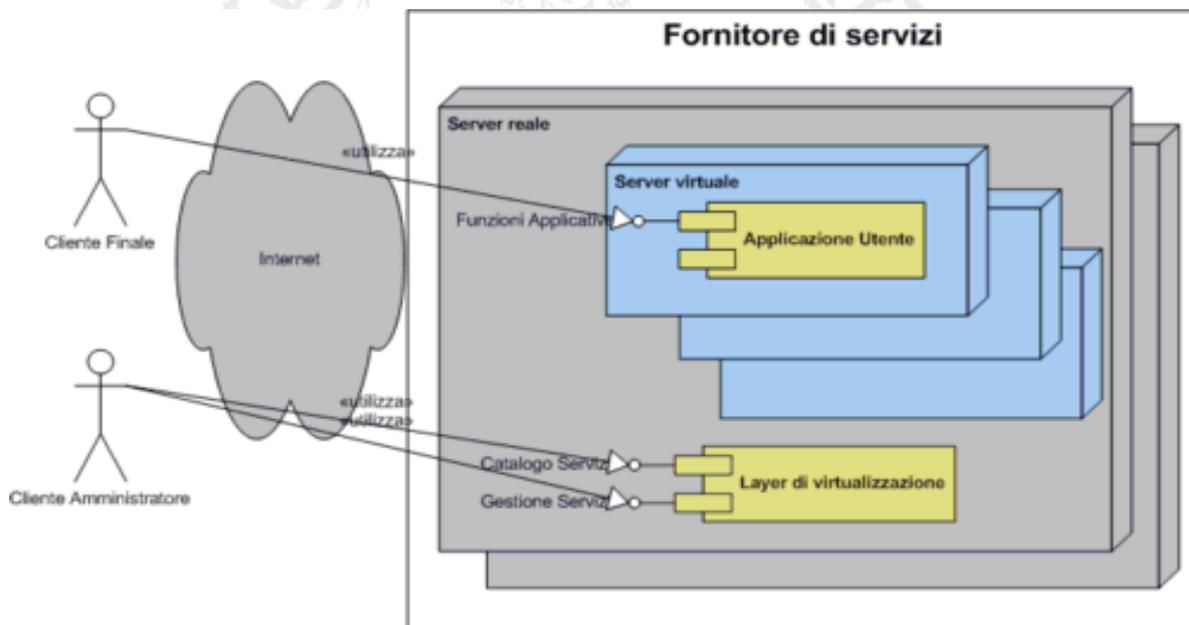
L'utente finale, infatti, non deve occuparsi della gestione software/hardware, ma può iniziare subito a lavorare, dopo aver caricato sull'infrastruttura del fornitore la propria applicazione.

I SaaS, infine, sono i servizi che mettono a disposizione dell'utente finale un software completo, al più configurabile, che sopperisce in toto a determinate richieste. Le web mail, conosciute ed utilizzate da tutti, sono un esempio di SaaS, perché permettono di poter gestire le proprie e-mail completamente dal web. Un altro chiaro esempio di SaaS, infine, è dato dai software da ufficio della suite Google Docs.

In definitiva, spostarsi nel Cloud significa abbracciare il concetto di BPO (Business Process Outsourcing) ovvero esternalizzazione dei processi di Business, con tutti i pregi e i difetti che ne conseguono.

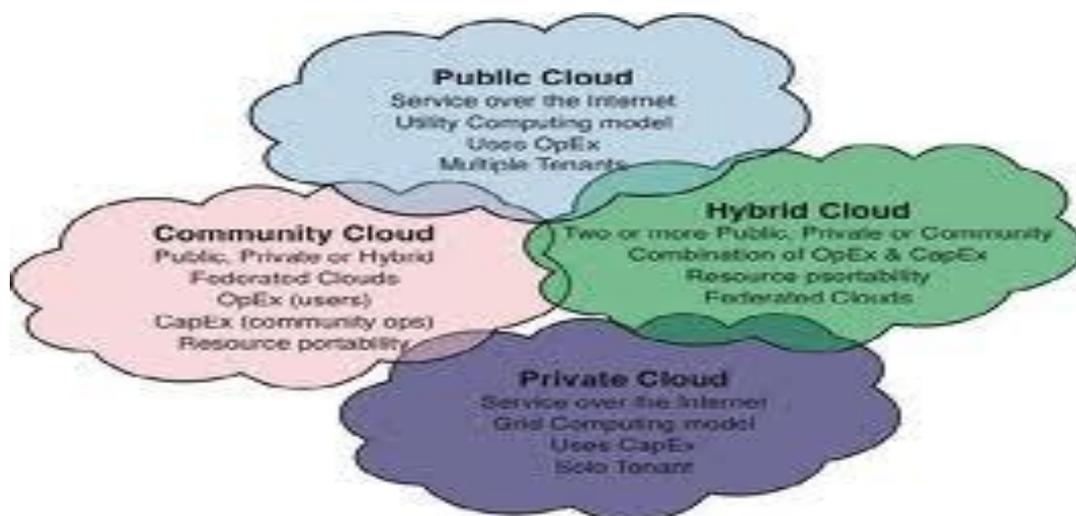
1.3 Architettura e casi d'uso

L'architettura del Cloud computing prevede uno o più server reali, generalmente in architettura ad alta affidabilità e fisicamente collocati presso il data center del fornitore.



Il fornitore di servizi (colui che offre servizi come server virtuali, storage, applicazioni complete ecc, generalmente secondo un modello "pay-per-use") espone delle interfacce per elencare e gestire i propri servizi. Il cliente amministratore utilizza tali interfacce per selezionare il servizio richiesto (ad esempio un server virtuale completo oppure solo storage) e per gestirlo (configurazione, attivazione, disattivazione). Il cliente finale (in molti casi coincide con l'utente amministratore) utilizza il servizio così configurato. Le caratteristiche fisiche dell'implementazione (server reale, localizzazione del data center) sono irrilevanti.

1.4 Modelli di distribuzione



- **Cloud pubblica:** La Cloud infrastructure è di proprietà di un'organizzazione che vende i servizi Cloud al pubblico o a un gruppo di grandi imprese.
- **Cloud privata:** La Cloud infrastructure è di proprietà o in leasing di una sola organizzazione ed è utilizzata esclusivamente dall'organizzazione. Le Cloud private possono essere paragonate ai tradizionali "data center" nei quali, però, sono usati degli accorgimenti tecnologici che permettono di ottimizzare l'utilizzo delle risorse disponibili e di potenziarle attraverso investimenti contenuti e attuati progressivamente nel tempo.
- **Cloud comunitaria:** La Cloud infrastructure è condivisa da diverse organizzazioni e sostiene una comunità specifica che ne condivide degli ambiti (ad esempio, la mission, i requisiti di sicurezza, la policy, e le considerazioni di conformità).
- **Cloud ibrida:** L'infrastruttura è una composizione di due o più Cloud (privata, comunitaria, o pubblica) in un'entità unica, tenute insieme da tecnologie proprietarie tali da permettere la portabilità dei dati e delle applicazioni.

1.5 Vantaggi e perplessità

In seguito sono riportati i principali vantaggi:

- Dal punto di vista economico:
 - **Abbassamento del TCO** (Total Cost of Ownership): risparmio sull'acquisto, installazione, manutenzione e dismissione di hardware e software.
 - **Maggiore flessibilità**: nel caso in cui servano maggiori o minori risorse è possibile effettuare un adeguamento contrattuale; cosa impossibile o limitata in caso di infrastruttura di proprietà.
 - **È possibile focalizzarsi sul proprio core business**: non è necessario assumere personale specializzato che si occupi della gestione dell'infrastruttura, essendo quest'ultima migrata nella Cloud.

- Dal punto di vista tecnico:
 - **Maggiore scalabilità**: nel caso in cui servano maggiori risorse (per picchi o quant'altro) il sistema di gestione e monitoring del Cloud ha la possibilità di allocare le risorse necessarie a fronte della nuova richiesta.
 - **Dati e applicazioni accessibili in mobilità**: per la definizione stessa di Cloud computing non potrebbe essere altrimenti.
 - **Maggiore sicurezza**: tutti i dati sono centralizzati e sottostanno alle policy di sicurezza centralizzate create; pertanto si annulla la possibilità di una fuga di dati dovuta alla sottrazione di materiale informatico nell'azienda.
 - **Possibilità di beneficiare di piani di Disaster Recovery**: se la società scelta per la fornitura dei servizi Cloud lo prevede, è possibile mettersi al sicuro da eventuali problemi che possono affliggere il data center dove sono raccolti i dati, disponendo di sedi secondarie in cui essi sono salvati ed elaborati.

Sono da menzionare, inoltre, i non meno importanti benefici apportati all'ambiente: utilizzando infrastrutture Cloud centralizzate è possibile ridurre il consumo energetico attraverso l'utilizzo di sistemi ad alta efficienza e/o costruendo datacenter in luoghi ad hoc (si pensi alle web farm dove si fa uso del free cooling, aiutate dal clima rigido).

Affidare la propria infrastruttura a servizi Cloud non ha solo pregi ma anche difetti che, attualmente, non permettono ad un'azienda di poter pensare di smantellare completamente la propria infrastruttura IT a favore dell'utilizzo del Cloud computing.

Si pensi a:

- **Difficoltà di migrazione dei dati:** Nel caso di un eventuale cambio del gestore dei servizi Cloud, non esistendo uno standard definito tra i gestori dei servizi un eventuale cambio di operatore risulta estremamente complesso. Tutto ciò risulterebbe estremamente dannoso in caso di fallimento del gestore dei servizi cui ci si è affidati.
- **Privacy:** non avendo più sotto controllo l'infrastruttura, dobbiamo fidarci di quanto riferitoci dal provider, contando sulla sua buona fede e sull'onestà dei suoi tecnici; questi avrebbero accesso ad un'enorme mole di dati (anche sensibili ed eventualmente coperti da segreto industriale)
- **La qualità e continuità del servizio:** L'infrastruttura di un provider che fornisce servizi Cloud è, naturalmente, molto complessa; più un sistema è complesso, più diventa difficile comprendere e gestire i rischi sia nelle fasi di implementazione che in quelle di gestione e utilizzo. Tutto ciò si traduce in un maggior rischio di insorgenza di problemi tecnici che portano a disservizi. Delegando a un servizio esterno la gestione dei dati e la loro elaborazione, l'utente si trova fortemente limitato, nel caso in cui i suddetti servizi non siano operativi (*out of service*). Essendo impossibile garantire ai clienti che l'infrastruttura sia bug-free, un provider serio dovrebbe necessariamente prevedere un piano di disaster recovery, che permetta di spostare tutto il carico di lavoro da un'infrastruttura ad un'altra (affrontando ulteriori ingenti investimenti).

Oltre a quanto descritto, bisogna pensare anche che il mantenimento di un'infrastruttura complessa e solida non è affatto poco costoso; ciò comporta che le percentuali di utilizzo dell'infrastruttura debbano essere alte per avere profitti.

Proprio per realizzare alcuni dei vantaggi citati, come la scalabilità e far fronte alle perplessità emerse, i Cloud provider sviluppano sistemi di monitoraggio grazie ai quali essi stessi e i clienti possono tenere sotto controllo le prestazioni, la qualità dei servizi e garantire maggiore sicurezza.

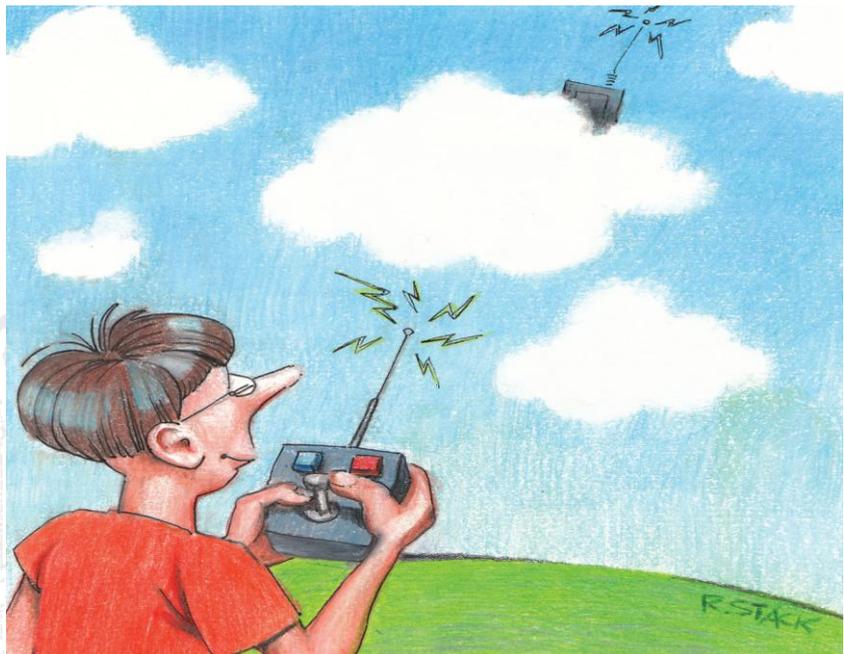
Capitolo 2

Monitoraggio del Cloud

Il monitoraggio è un'attività molto importante negli ambienti Cloud, poiché valutare la salute delle infrastrutture, risorse hardware/software, piattaforme e applicazioni, è un aspetto cruciale sia per i provider che per gli utilizzatori.

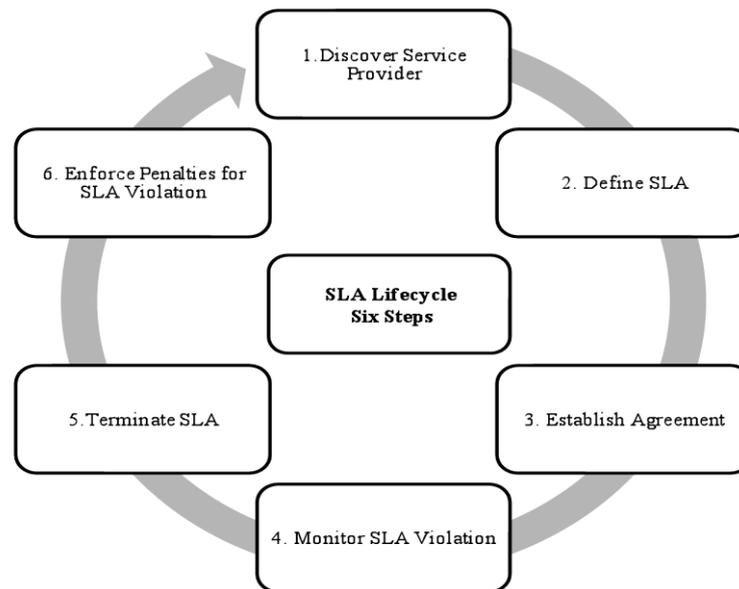
Il monitoraggio può essere visto sotto più aspetti: da un lato è indispensabile misurare e controllare sia i dispositivi fisici che quelli virtuali, al fine di garantire il buon funzionamento di tutto il sistema; dall'altro, le stesse misurazioni forniscono una serie di informazioni sulla qualità del servizio offerto.

Quindi è indispensabile, in questi anni in cui il Cloud computing si sta diffondendo, osservarne le prestazioni e valutarne i reali vantaggi, rispetto a soluzioni tradizionali. Per le aziende la possibilità di acquisire proprie strumentazioni per mettere su il sistema informatico, mentre per gli sviluppatori affidarsi a provider di hosting classici per le proprie applicazioni. Per valutare le prestazioni di un singolo sistema Cloud e confrontarle fra i vari provider c'è, quindi, bisogno di test di misurazione mirati, e di metriche di giudizio adeguate. Più avanti, vengono mostrati alcuni esempi tratti da articoli scientifici. Cerchiamo ora di rispondere ad alcune domande:



Chi ha l'esigenza di monitorare?

Tipicamente, clienti e provider redigono, tra loro, un contratto formale che prende il nome di SLA (Service Level Agreement). In questo, vengono stabiliti i dettagli del servizio fornito, in termini di parametri e obiettivi concordati tra le parti (ad es. latenza e disponibilità), le sanzioni e/o compensazioni in caso di violazione delle aspettative.



Ciclo di vita del SLA ^[4]

Così viene definita la qualità del servizio da offrire e stabilite le responsabilità tra le parti. Dalle seguenti osservazioni, si intuisce l'importanza del monitoraggio. Infatti:

- Il provider può prevenire violazioni del SLA e adattare il servizio ai cambiamenti del contesto. Per esempio, monitorando l'andamento delle risorse assegnate e dei carichi di lavoro sui server, può individuare le situazioni di disagio e agire tempestivamente con adeguate tecniche di load balancing; garantendo, così ai propri clienti, la continuità e la qualità del servizio fornito.
- Il cliente può tenere sotto controllo i propri dati e i trattamenti che vengono loro applicati. La possibilità di valutare la qualità del servizio offerto è particolarmente importante nel caso PaaS e IaaS, dove i clienti sono a loro volta fornitori di servizi web, attraverso proprie applicazioni residenti nel Cloud. Risulta, quindi di fondamentale importanza, per il loro business, avere uno strumento che fornisca, in tempo reale, informazioni sull'andamento delle proprie applicazioni, per agire prontamente in caso di malfunzionamenti o degradazione delle prestazioni e non creare disservizi ai rispettivi utenti.

Cosa si tiene sotto controllo?

Qualsiasi componente facente parte della struttura: l'hardware delle risorse, ma anche middleware e software, vanno monitorati e tenuti sotto controllo. In tal senso, il provider ha, ovviamente, le maggiori responsabilità poiché deve garantire il buon funzionamento e la qualità delle prestazioni, secondo quanto concordato ma è anche importante disporre di strumenti che consentano ai clienti di avere un resoconto sull'andamento dei servizi offerti.

Per questo il monitoraggio si divide in due livelli: ^[5]

High-level: Si propone di raccogliere, e mettere a disposizione dei clienti, informazioni che si focalizzano sulla piattaforma virtuale. Essi hanno grande interesse a valutarne lo stato di salute e quello delle proprie applicazioni che vi girano sopra. A tal proposito, molti, tra i Cloud provider, mettono a disposizione (direttamente o attraverso terzi) piattaforme di monitoraggio realizzate a questo scopo. Nel terzo capitolo, ci soffermeremo su alcune di queste.

Low-level: Si propone di raccogliere informazioni che riflettono lo stato dell'infrastruttura fisica, dall'hardware al livello di virtualizzazione delle risorse.

Questo monitoraggio ha grande importanza per i Cloud provider, dovendo garantire tutta una serie di automatismi (un esempio è il bilanciamento del carico), per tenere costante il livello di qualità delle prestazioni offerte, anche in situazioni impreviste.

Le informazioni raccolte in questo monitoraggio, vengono, generalmente, tenute nascoste al cliente e riguardano vari aspetti: dall'accessibilità e funzionalità delle risorse, alla sicurezza.

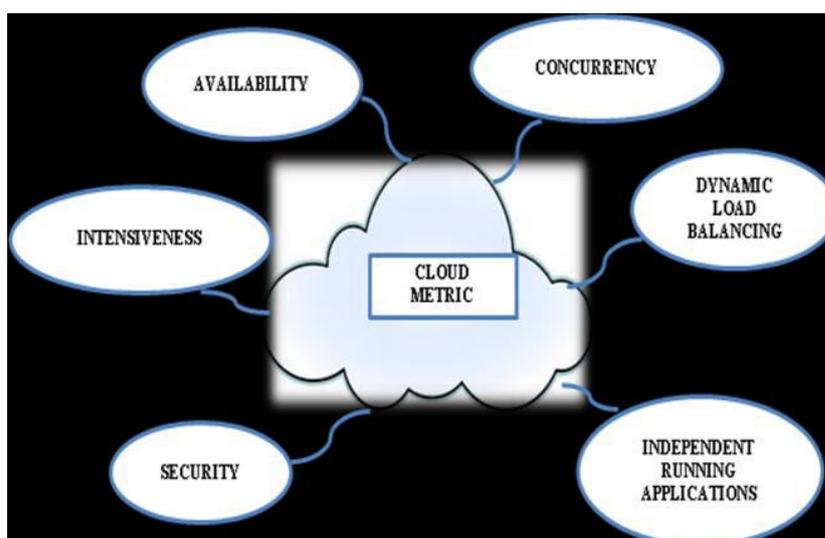
Il monitoraggio a questo livello si realizza attraverso: ^[6]

- Misurazioni delle risorse hardware con l'ausilio di utility dedicate, che consentono di avere resoconti sulle prestazioni dei server, in termini di impegno delle memorie e della CPU, dei valori di temperature e dei voltaggi, ma anche carichi di lavoro, in modo da sviluppare servizi efficienti, robusti e dinamici.
- Controllo delle strutture e dei locali con sistemi di video sorveglianza e di allarme che, con rigide politiche di accesso ai dati (come gestioni di privilegi, autenticazioni, ecc) garantiscono la sicurezza dei data center, in modo da evitare che persone non autorizzate possano fisicamente accedere a risorse e dati.
- Monitoraggio e protezione della rete di accesso ai data center, per ridurre, il più possibile, attacchi da parte di malintenzionati, attraverso robusti firewall, sistemi d'intrusion detection (IDSs) e intrusion prevention (IPSs) e, inoltre, evitando la divulgazione d'informazioni sensibili attraverso sistemi di criptaggio avanzati sui dati. I dispositivi di difesa sono le postazioni migliori per raccogliere informazioni su eventi che riguardano la sicurezza delle reti.
- Analisi e controllo di OS e middleware per la verifica di errori di configurazione e vulnerabilità del codice.

Che cosa deve avere un buon Cloud ?

Per valutare al meglio le prestazioni e confrontare i diversi servizi, c'è bisogno di avere delle metriche di giudizio ben definite a cui far riferimento anche se, allo stato attuale, non esistono metriche uniche e standardizzate.

Qui di seguito sono riportate in breve alcune metriche interessanti tratte dall'articolo: "CM-Measurement Facets for Cloud Performance" [7] che ritengo abbastanza significative per giudicare un buon servizio Cloud.



- **Availability (Disponibilità):** La natura on-demand, elastica e scalabile del Cloud, va considerata durante la realizzazione dei suoi servizi. Molti utenti potrebbero accedere alle stesse applicazioni contemporaneamente e da alcune statistiche è emerso che un utente non attende più di qualche secondo prima di abbandonare un sito web o applicazione che non risponde istantaneamente. E' quindi di essenziale necessità adempiere la richiesta senza sprechi di tempo, in modo da rendere subito disponibili le risorse. Per garantire che le varie istanze delle applicazioni, residenti nel Cloud, siano accedute con successo, si richiedono, per esempio, tecniche avanzate di virtualizzazione e di bilanciamento del carico sui server.

Alcuni criteri che consentono di valutare la disponibilità in funzione del numero di utenti concorrenti su una o su un gruppo di applicazioni sono i seguenti:

- Il tempo di risposta di cui un'applicazione necessita per rispondere alla richiesta dell'utente.
- La trasparenza dei tempi di risposta, poiché la raccolta d'informazioni per rispondere a una determinata richiesta può variare da richiesta a richiesta ed è bene che sia nota all'utente in modo che possa conoscere, in anticipo, il tempo di attesa.
- L'ammontare di tempo speso dagli utenti nel visitare una o più applicazioni.

- **Concurrency (Contemporaneità):** Fornire accessi concorrenti agli utenti è indispensabile per garantire a questi ultimi servizi ininterrotti e sempre disponibili. Uno dei criteri di concorrenza può essere l'ipermedialità delle applicazioni, in modo da garantire accessi rapidi, non lineari e guidati ai contenuti informativi.
- **Dinamic load balancing (Bilanciamento dinamico del carico):** La presenza di carichi elevati imprevedibili, può minare l'accessibilità delle applicazioni e delle risorse da parte degli utenti; quindi, è necessaria una buona gestione di questi eventi in maniera decifrabile e robusta. Il bilanciamento del carico consiste nel migliorare le prestazioni, distribuendo le richieste degli utenti su un diverso numero di server, riducendo, così, il carico singolo nei momenti di maggiore stress. Senza il load balancing l'infrastruttura computazionale non è scalabile.
- **Independent running applications (Possibile esecuzione di diverse applicazioni)**
Nel Cloud girano simultaneamente diverse applicazioni ma, per garantire la disponibilità dell'intero sistema, queste non devono interferire o intralciarsi fra loro.
L'accessibilità delle applicazioni distribuite nel Cloud è raggiunta attraverso server virtuali e schedulazioni delle risorse, in modo da dare all'utilizzatore la sensazione di avere un sistema completamente dedicato alle sue esigenze.
- **Security (Sicurezza):** I servizi Cloud si estendono sulla rete mondiale e sono, quindi, accessibili anche ad hacker e cyber-terroristi che, attraverso virus, worms e attacchi di ogni specie possono minare la sicurezza degli utenti e delle imprese che fanno uso del Cloud. È, quindi, necessario mettere in sicurezza i contenuti delle applicazioni, attraverso resistenti misure di sicurezza: come sistemi di crittografia avanzati, per garantire la riservatezza e integrità dei dati, e servizi come HTTPS (Hypertext Transfer Protocol Secure) e SSH (Secure Shell) per garantire trasferimenti sicuri.
Inoltre fornire servizi di disaster recovery in modo da poter rimediare ai danni subiti e non causare problematiche ai clienti.
- **Intensiveness (Intensità):** Le applicazioni distribuite nel Cloud devono servire le esigenze di diverse comunità di utenti, in modo rapido efficiente e semplice.
L'accessibilità a informazioni di rilevante importanza, come per esempio servizi forniti alla comunità da organi statali, devono essere reperibili in modo semplice; design accattivanti e interfacce grafiche usabili aiutano molto a questo riguardo. Si può, così, ridurre al minimo coloro che restano esclusi dal mondo del web, come anziani o disabili.

2.1 Misurazioni e test

Come detto in precedenza, il monitoraggio è fatto di attente e precise misurazioni che, tra l'altro, forniscono importanti informazioni ai fini valutativi, per questo, realizzare test di misura sui servizi offerti dal Cloud è, allo stato attuale, molto importante.

Questi test vengono realizzati dagli stessi provider, in prima persona oppure delegando terzi o anche gruppi di ricerca, per prendere coscienza della qualità di ciò che offrono.

Questi test si possono dividere in due categorie:

Computation-based tests : In questa categoria ricadono le misurazioni tese a valutare le performance computazionali dei server reali e virtuali su cui girano le applicazioni in esecuzione nel Cloud.

Attraverso adeguate strumentazioni hardware e software è possibile per esempio misurare: (queste sono alcune misure utilizzate e citate nell'articolo ^[8])

- **Server throughput** (#req/sec): Misura quantitativamente il numero massimo di richieste servite con successo al secondo, nel retrieval di documenti web.
- **CPU time per execution** (µs/exe): È un indicatore di performance che mostra il tempo della CPU medio ottenuto in microsecondi durante ogni esecuzione.
- **CPU utilization** (%): Per comprendere l'utilizzazione condivisa della CPU tra le VMs (virtual machines) eseguite su una singola macchina fisica, attraverso misurazioni dell'uso medio di CPU da parte di ogni VM.
- **Network I/O per second** (kByte/sec): L'ammontare di traffico nella rete I/O, in termini di KB/sec, trasferiti da e verso un web server remoto per il corrispondente carico di lavoro.
- **Memory pages exchange per second** (pages/sec): Misura del numero di pagine di memoria scambiate al secondo in un canale I/O. Indica quanto è efficiente un processo di I/O.
- **Memory pages exchange per execution** (pages/exe): Indicatore di performance della memoria, che mostra la media di pagine usufruite durante un'esecuzione.

Questi test sono generalmente sviluppati dagli stessi provider, che hanno piena autonomia nella scelta delle metriche e dei fattori che ritengono maggiormente rilevanti. Molto spesso, vengono eseguiti anche da compagnie terze autorizzate come per esempio Hyperic Inc che monitora le performance sia di EC2 che App Engine, in tempo reale e pubblica i risultati sul proprio website CloudStatus. ^[9]

Network-based tests : Misurano le performance della rete, nella gestione delle richieste da parte delle applicazioni residenti nel Cloud. Includono metriche come round-trip time (RTT), network throughput, dataloss, larghezza di banda, ritardi, latenza, e molti altri.

Molti gruppi di ricerca sviluppano questi test al fine di valutare, da un punto di vista di chi usa il Cloud, le performance e la convenienza rispetto ai tradizionali servizi di web-hosting.

A questo proposito, vorrei citare uno di questi test, sviluppato dal dipartimento di Computer Sciences dell'università del Texas in Austin, USA. I risultati sono stati pubblicati nell'articolo: "Network-based Measurements on Cloud Computing Services" ^[10], dal quale sono tratte le informazioni di questo paragrafo.

Tale lavoro mira a valutare l'efficienza dei sistemi Cloud, rispetto a quelli di web-hosting tradizionali, prendendo come riferimento il servizio offerto da Google App Engine (GAE), che vedremo più da vicino nel capitolo 3. ^[11]

Una volta applicato a un tale sistema e a uno tradizionale come Synthesite (SYN) un'applicazione web che si occupa d'immagine retrieval, si può utilizzare PlanetLab ^[12] per generare richieste parallele, simulando tre diversi livelli di carico di lavoro (light, medium, heavy).

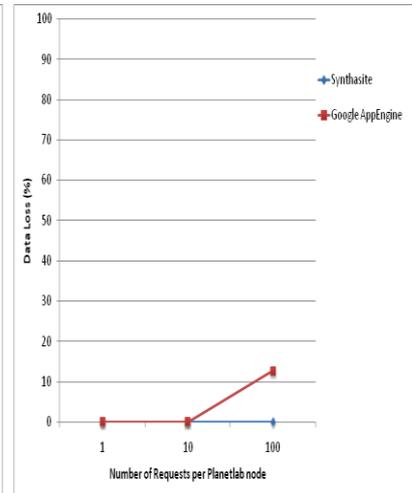
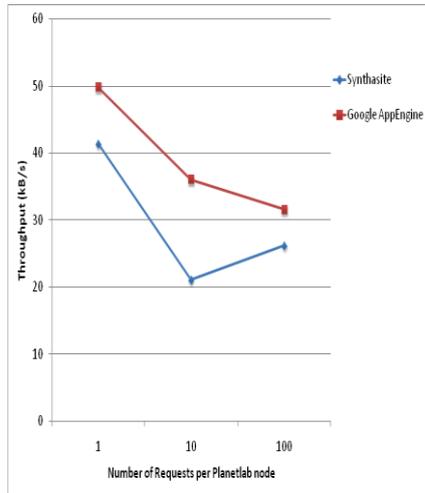
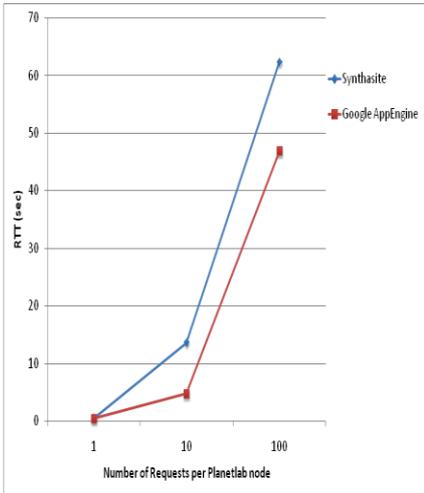
Utilizzando, poi, strumenti come httpperf ^[13] vengono ottenuti valori come:

- **Round-trip time (RTT)**: definito come il tempo impiegato da un pacchetto di dimensione trascurabile per viaggiare fino alla postazione ricevente e quello del suo ritorno alla sorgente.
- **Network Throughput (bit/s)**: È definito come il tasso medio di trasferimento dati, con successo (quantità di dati trasmessi in un'unità di tempo), attraverso una connessione di rete tra due host. È da distinguere dalla larghezza di banda (bandwidth), che è la capacità di un sistema di trasferire dati su una connessione, in un dato periodo. Dalla prospettiva del cliente, anche se i provider basano le proprie fatturazioni sulla banda e non sul throughput, il secondo è più importante poiché aiuta a definire la velocità percepita dall'utente nel ricevere i dati a fronte di una richiesta.
- **Packet/Data Loss**: Quantità di dati o pacchetti che, durante un trasferimento dati tra due host di una rete, non giungono a destinazione. Questa è una metrica importante, come valutazione quantitativa dei dati che un client riceve da un server.

Sono usati 100 nodi Planetlab programmati per inviare 1, 10, 100 richieste parallele e immagini di dimensioni small (12 KB), medium (350 KB) e large (1MB), per simulare le diverse condizioni di carico a cui i sistemi sono sottoposti.

Risultati:

- Synthasite - Google App Engine

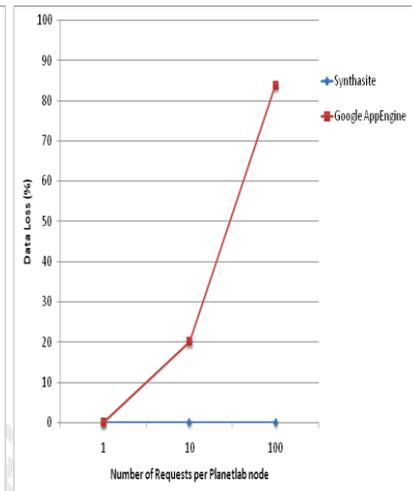
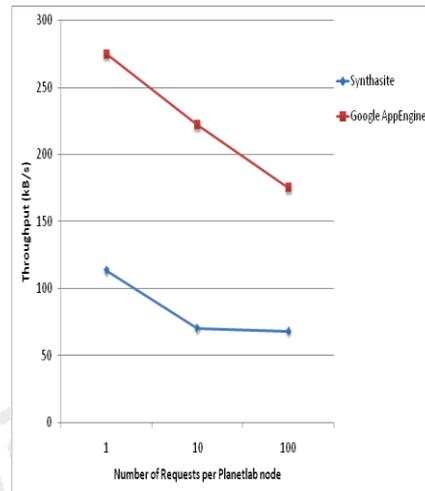
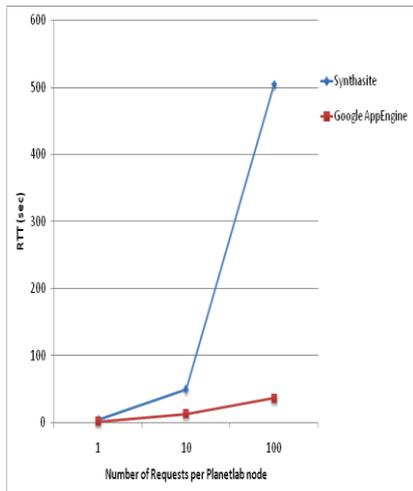


RTT

Throughput

DataLoss

(a) Small image (12 KB)

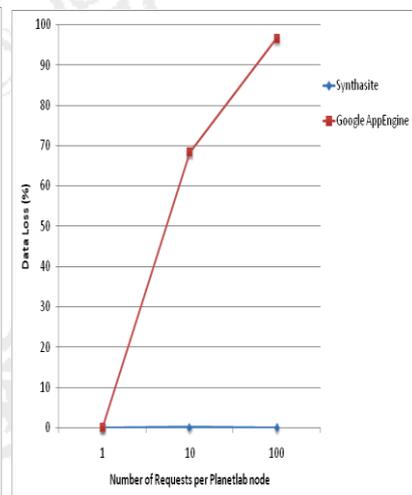
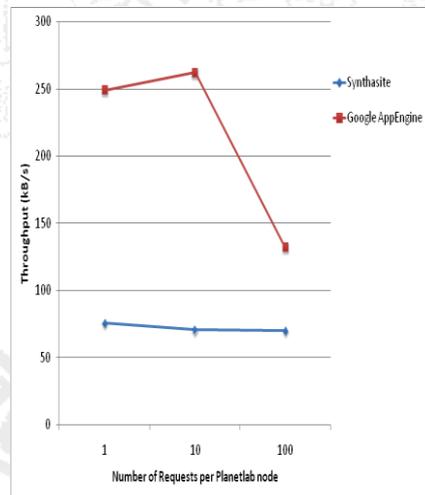
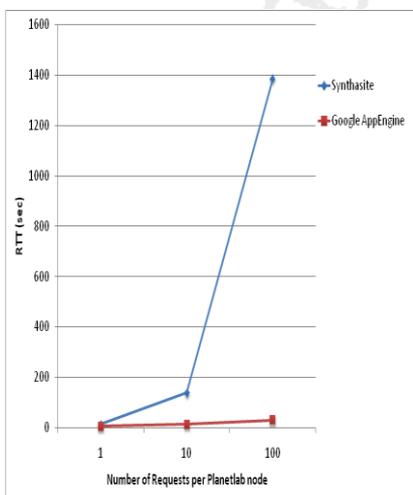


RTT

Throughput

DataLoss

(b) Medium image (350KB)



RTT

Throughput

DataLoss

(c) Large image (1MB)

Analisi dei risultati: Per quanto riguarda RTT, si nota che al crescere della quantità di byte aumenta maggiormente in SYN rispetto a GAE (migliori prestazioni si hanno con RTT bassi). Si ha throughput più elevato nel caso GAE che SYN, in tutti e tre le situazioni (small, medium, large), con valori più alti aumentando il numero di byte trasmessi. Come si può notare, nel caso GAE, il throughput diminuisce fortemente al crescere del numero delle richieste degradando maggiormente le prestazioni rispetto a SYN che le mantiene maggiormente costanti. Infine, si mostra, come nel caso GAE, che al crescere del carico aumenta fortemente il Data-Loss; questo porta, ovviamente, a un forte degrado delle prestazioni, rendendole molto scarse sotto carichi pesanti.

I ricercatori affermano di non dover considerare tali risultati pienamente attendibili, poiché strumenti come httpperf, utilizzato per ottenere le misurazioni, possono non essere stati configurati nella maniera più opportuna.

Ho, comunque, voluto citare questo studio, per mostrare come siano importanti i lavori di valutazione e, sia i provider che il mondo della ricerca devono impegnarsi in tal senso. Infine, come ogni nuovo modello emergente, dovrà essere speso tempo e impegno affinché le prestazioni siano ottimizzate fino a fornire elevati livelli di qualità.

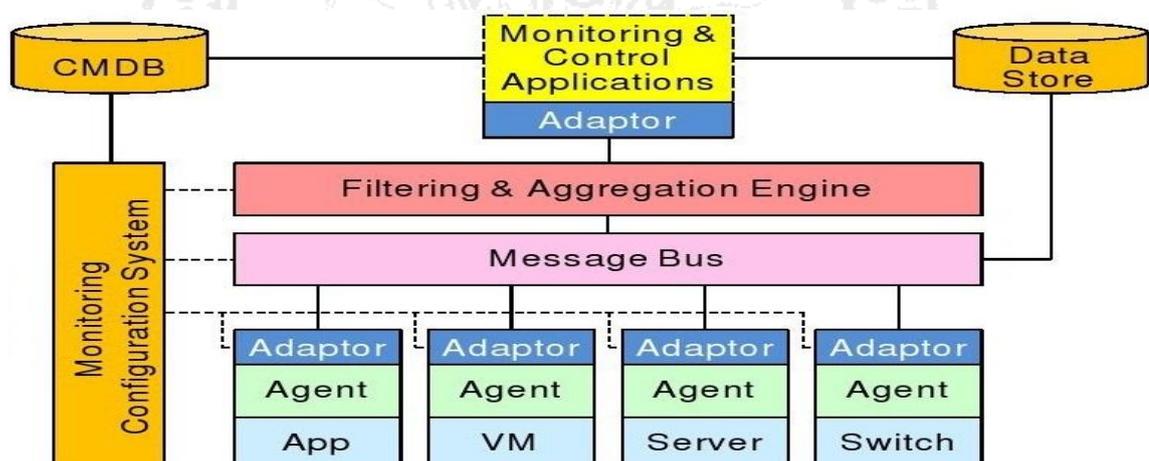
2.2 Architettura funzionale di un sistema di monitoraggio

Per fornire informazioni di monitoraggio adeguate e il più possibile efficaci allo scopo, c'è bisogno di sistemi di supervisione dotati di particolari requisiti strutturali e che si adattino, al meglio, in ogni situazione.

Alcuni di questi requisiti possono essere, la scalabilità, il dinamismo, il multi-tenancy, la semplicità.

Ho voluto riportare di seguito un'architettura tratta dall'articolo "Towards Holistic Multi-Tenant Monitoring for Virtual Data Centers" ^[14]

Tale struttura è un esempio di come può essere realizzato un sistema di monitoraggio. Non entrerò nel dettaglio delle varie tecnologie che possono essere adottate per realizzare i vari componenti, mi interessa descrivere le funzionalità generali di ognuno di essi, ritenendo che, anche se con altri nomi o simbologie, tali componenti sono presenti in qualsiasi piattaforma di monitoraggio presa in considerazione.



Il blocco principale di quest'architettura è il sistema di propagazione delle notifiche di eventi, chiamato "message bus" (può essere realizzato in differenti modi). A questo è collegato una struttura di filtraggio e aggregazione, allo scopo di far passare ai livelli superiori le notifiche, secondo i criteri decisi in fase di configurazione del sistema. Può essere considerato "evento" una qualsiasi situazione, di rilevante importanza per il monitoraggio, verificatasi su una risorsa, catturato e notificato dagli appositi agenti (dispositivi a cui è affidata la misurazione e la generazione delle notifiche). Ad esempio, un evento può essere la rilevazione, da parte dell'agente presente su una data risorsa, che i valori misurati sono esterni all'intervallo ritenuto funzionale in fase di configurazione. Ogni risorsa, monitorabile in un data center, ha collegati uno o più agenti. Questi, effettuando misure periodiche o su richiesta, sono le sorgenti delle informazioni del monitoraggio che, filtrate e aggregate, viaggiano attraverso il message bus e sono a disposizione delle applicazioni di controllo. Vengono salvate nell'apposito data store, per creare registri permanenti e consentono analisi successive.

Tali applicazioni possono essere realizzate in modo da implementare il comportamento che soddisfi al meglio le esigenze e gli scopi che il cliente del Cloud o il provider stesso necessitano (Es. report in tempo reale, statistiche e/o grafici per aiutare le analisi).

Per garantire l'interoperabilità dei vari blocchi, si pongono degli adattatori, che astraggano i dati scambiati dalle particolari tecnologie utilizzate.

Le informazioni sugli agenti, le loro capacità, le risorse assegnate e tutte le altre informazioni, relative alle configurazioni del sistema, sono salvate nel Configuration Management Database (CMDB). Questo può essere consultato dalle applicazioni, per conoscere quali agenti sono attivi, nonché gli utenti autorizzati a richiedere informazioni di monitoraggio e ricevere notifiche.

Il sistema deve, inoltre, garantire, attraverso autenticazione degli utenti e identificativi univoci sugli agenti, la sicurezza dei messaggi di notifica generati, in modo che giungano solo alle giuste destinazioni.

Di fondamentale importanza è il sistema di configurazione che, attraverso adeguate GUI (Graphical user interface) rende ad esempio, possibile, agli utenti del sistema di monitoraggio:

- conoscere gli agenti attivi sulle risorse;
- fissare le soglie di avvertenza per generare le notifiche;
- fissare la frequenza delle misurazioni, o richiederne una.

Inoltre è gestito il filtraggio e l'aggregazione delle notifiche, per esempio per valutare l'utilizzazione delle CPU dei vari server, accumulando tali informazioni per avere un resoconto sullo stress dell'intero server farm.

Quest'architettura è ovviamente scalabile e dinamica poiché è possibile, proprio grazie ad adeguati adattatori e agenti attivabili, incrementare o diminuire il numero di risorse da monitorare, in base alle varie esigenze dei diversi utenti. Il sistema, poi, segue il principio di semplicità suddividendo su più componenti le varie funzionalità ed è multi-tenancy poiché gestisce in maniera indipendente ed efficiente una vasta comunità di utenti interessati all'andamento del sistema Cloud attraverso un unico sistema di monitoraggio.

Di seguito, è riportato, con un esempio, uno dei possibili formati per i messaggi di notifica.

```
sourceType : http://example.org/NetworkLoad
sourceID : 192.168.1.20
data Type: float
tenants: customer86
timestamp: 2009-12-10T05:52:19.358
data: <value>0.21</value>
```

In cui:

- sourceType: campo che contiene un identificativo unico del tipo di informazione che è trasferita nel messaggio.
- sourceID: campo che identifica univocamente l'origine dell'informazione di monitoraggio.
- dataType: campo che indica il tipo di dato trasferito nella notifica.
- tenants: lista degli identificativi degli utenti a cui è destinata la notifica.
- timeStamp: istante in cui è stata creata la notifica.
- data: valore dell'informazione di monitoraggio trasportata

2.3 Difficoltà e sfide nel monitoraggio del Cloud

L'infrastruttura di un provider che fornisce servizi Cloud è, naturalmente, molto complessa. Tale complessità si traduce in maggiori difficoltà e sforzi per la gestione e il monitoraggio di tali ambienti.

La maggiore scalabilità e dimensione massima che le strutture Cloud vantano rispetto a quelle degli hosting tradizionali, comportano sistemi di monitoraggio più complessi, che siano, a loro volta, scalabili, robusti e veloci. Essi devono essere in grado di gestire e verificare un numero elevato di risorse e devono farlo in maniera efficiente. Questa efficienza viene realizzata attraverso tempi di misurazione rapidi e sistemi di allerta sempre pronti a notificare cali di prestazioni o altre problematiche, in modo da garantire interventi tempestivi come l'assegnazione di nuove risorse.

I Cloud provider dovranno, quindi, gestire strutture molto più grandi rispetto ai sistemi di hosting classici, per riuscire a fornire ad un numero elevato di clienti (tra cui le aziende per l'infrastrutture IT) i propri servizi nel rispetto del SLA. Per questo, le tecniche di monitoraggio vanno raffinate e adattate alle diverse situazioni presenti in ambienti di larga scala e fortemente dinamici come le strutture Cloud.

Il monitoraggio non ha più un ruolo marginale ma è tra le attività più importanti nella gestione della struttura poiché è indispensabile per implementare tutta una serie di automatismi che caratterizzano i servizi Cloud.

In questo paragrafo sono messe in evidenza alcune difficoltà (citate nell'introduzione dell'articolo ^[15]) che si presentano maggiormente accentuate nella realizzazione del monitoraggio di sistemi Cloud rispetto a quelli tradizionali. Come:

- **Gestione degli eventi:** Applicazioni, server, OS, dispositivi di rete possono generare un elevato ammontare di eventi, che il sistema deve catturare comprendere e gestire. La gestione in ambienti di larga scala non è semplice e il sistema di monitoraggio deve essere sempre in allerta.

- **Affidabilità del sistema:** La complessità del sistema di monitoraggio può portare a imprevedibili malfunzionamenti, che si possono ripercuotere sulla consistenza dei risultati. Metriche di consistenza per il monitoraggio in larga scala vanno implementate proprio per indicare la precisione dei risultati di monitoraggio e riconoscere dati inconsistenti prodotti dal fallimento del sistema.
- **Aggregazione dei dati:** Il numero elevato di misurazioni, al fine di ottenere un resoconto completo porta alla generazione di un ammontare di dati molto elevato; l'abilità di raggruppare questo volume di dati provenienti da diverse postazioni distribuite, è un aspetto cruciale per il monitoraggio dello stato globale del Cloud e fare questo in maniera efficiente ed efficace è una bella sfida.
- **Spreco di risorse:** Risultati inconsistenti nelle misurazioni possono allertare il sistema inutilmente, approvvigionando maggiori risorse non necessarie, causandone degli sprechi. Da notare anche che, sistemi di monitoraggio complessi per indagare a fondo lo stato di salute delle risorse hanno bisogno anch'essi di risorse computazionali dedicate alle esecuzioni dei propri task e di banda per le comunicazioni dei dati.

Ulteriori difficoltà si presentano a causa dell'uso intensivo delle tecniche di virtualizzazione, alla base della realizzazione dei servizi Cloud; come sulla precisione del time-stamping. Quella pratica di marcatura temporale, che i diversi componenti di un sistema eseguono sui propri eventi e prodotti (come i pacchetti inviati e ricevuti dai vari dispositivi), ed è molto importante in fase di misurazione e controllo.

Le VMs implementate su un nodo fisico richiedono meccanismi di scheduling e switching realizzati sul nodo stesso. I pacchetti appartenenti a una VM saranno quindi accodati fino a quando il sistema fisico non li gestirà, comportando un impreciso time-stamping di alcuni micro o millisecondi rispetto al reale tempo di arrivo del pacchetto.

In ambienti virtuali il controllo e le misurazioni, per realizzare analisi sulle prestazioni, rappresentano le principali sfide del monitoring del Cloud poiché introducono diverse problematiche.

Nell'articolo "TOWARDS PERFORMANCE PREDICTION FOR CLOUD COMPUTING ENVIRONMENTS BASED ON GOAL-ORIENTED MEASUREMENTS" (Verso la predizione delle prestazioni negli ambienti di elaborazione Cloud, basati sulle misure orientate all'obiettivo) ^[16] ne sono citate alcune e viene mostrato un approccio goal-oriented per ottenere automaticamente delle predizioni sulle performance e consentire analisi valutative in grandi ambienti virtuali come lo sono quelli Cloud.

- *Come incidono sulle prestazioni finali le performance delle piattaforme virtuali?*

Le prestazioni delle applicazioni, che girano su di una piattaforma virtuale, sono inevitabilmente dipendenti dalle performance di quest'ultima e, quindi, è importante esaminarne le proprietà rilevanti e come influenzano tali applicazioni.

Va tenuto conto dell'overhead che la virtualizzazione introduce per gestire le contese sulle risorse fisiche da parte delle diverse VM. L'hypervisor è il software che fa

questo, quindi gli algoritmi di schedulazione adottati influenzano le performance e mirare a realizzarne di più efficienti è una sfida importante.

- *Come progettare esperimenti per ottenere i valori rilevanti delle performance, indipendentemente dalla tecnologia?*

Al fine di realizzare previsioni sulle performance dei diversi ambienti Cloud gli esperimenti di misurazione vanno progettati in modo generico senza tener conto della specifica tecnica di virtualizzazione o hypervisor.

- *Come condurre accurate misurazioni in ambienti virtuali?*

Effettuare precise misurazioni in ambienti virtuali non è facile, se non si conoscono bene gli effetti sulle performance che la virtualizzazione introduce, si rischia di ottenere risultati inappropriati.

L'approccio considerato segue i passi mostrati in figura 1 :

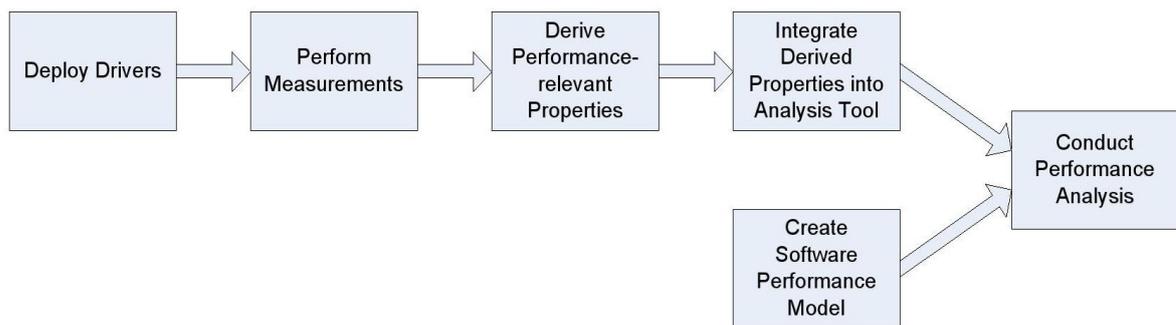


Figura 1: Flusso di lavoro della rilevazione automatica delle proprietà rilevanti per le prestazioni.

1. **Deploy drivers (Distribuire i driver):** Per accedere all'ambiente da monitorare vengono installati dei driver (nell'articolo non è specificata la loro implementazione).
2. **Perform measurements (Misurazione delle prestazioni):** I driver vanno distribuiti e calibrati in modo da eseguire diverse misurazioni che devono essere robuste rispetto al rumore e generiche per essere applicabili a piattaforme virtuali eterogenee. Possibili misure sono la virtualization overhead per alcune risorse richieste, network I/O throughput, latenze. Le misure sono avviate da un controller che ha anche il compito di raccolta dei risultati.
3. **Derive performance-relevant properties (Derivare le proprietà importanti per le prestazioni):** I risultati delle misurazioni sono analizzati anche con l'ausilio di metodi statistici per derivare proprietà interessanti.
4. **Integrate Derived Properties into Analysis Tool (Integrare le proprietà derivate negli strumenti di analisi):** Le proprietà ottenute sono integrate in appositi strumenti che implementano diversi modelli di analisi configurabili per una o gruppo di proprietà specifiche ottenute in precedenza.

Capitolo 3

Soluzioni commerciali e open-source per il Cloud computing

3.1 L'ultimo decennio del Computing ^[17]

L'avvento del Web e l'aumento della banda disponibile per i collegamenti portò alla nascita dei primi Application Service Provider (ASP), ovvero delle prime aziende che fornivano applicazioni disponibili tramite accesso remoto. Tra gli ASP troviamo Citrix, che nel 1995 presenta al pubblico *Citrix WinFrame*: un sistema operativo basato su WindowsNT 3.51 al quale era stata aggiunta la possibilità di essere utilizzato in multiutenza. Microsoft fiuta le potenzialità del prodotto di Citrix e decide di non concedere più in licenza i sorgenti di WinNT 4, scegliendo di dare vita al noto *Microsoft Terminal Services*. Purtroppo i tempi non erano ancora maturi e le soluzioni proposte non ebbero il successo sperato, a causa dei costi delle licenze.

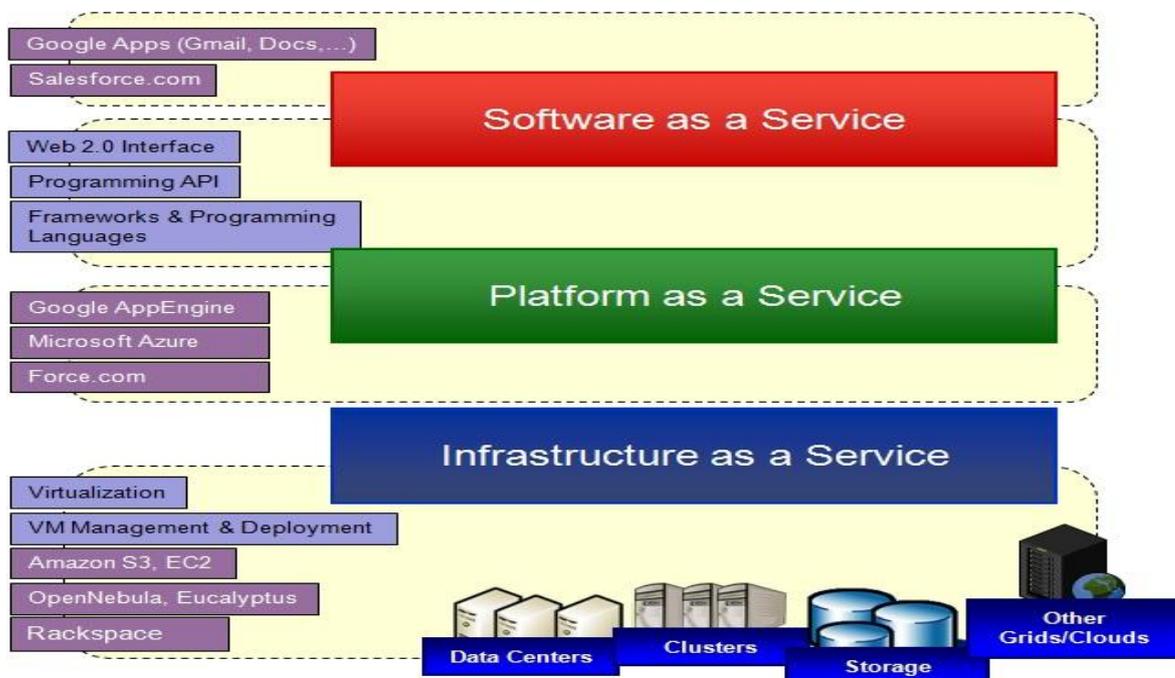
Accantonata l'idea di fare uso massivo dei servizi sviluppati dagli ASP, Connectix prima, VMWare poi, presentarono alle aziende i loro prodotti per la virtualizzazione. VMWare Workstation ed ESX Server fecero il loro esordio a cavallo degli anni 2000 e furono accolti positivamente dalle aziende, che videro nella virtualizzazione la possibilità concreta di ottimizzare l'infrastruttura IT ed abbassare il TCO. Proprio la virtualizzazione è uno dei pilastri su cui poggiano i servizi Cloud attualmente erogati dai provider mondiali.

Agli inizi del terzo millennio, in ambito prettamente software, nasce l'idea di una nuova architettura, in cui l'infrastruttura software viene vista come un insieme di servizi organizzati per concorrere allo stesso obiettivo; l'architettura in oggetto viene chiamata Service Oriented Architecture (SOA). Nel corso del passato decennio il modello SOA viene completamente definito ed il consorzio OASIS standardizza i linguaggi ed i protocolli che ne sono alla base (SOAP, WSDL, UDDI, WS-*, etc).

Il paradigma SOA è stato ed è tutt'ora adottato da molti vendor, per gli indubbi vantaggi che l'architettura porta con sé. La fusione tra le tecnologie di virtualizzazione ed il paradigma SOA ha permesso lo sviluppo di applicativi altamente portabili, ad architettura distribuita, oltre alla possibilità di interconnettere ed utilizzare facilmente software eterogenei ed eventualmente gestiti da altre aziende. Al giorno d'oggi, il Cloud computing è un paradigma

di uso delle risorse computazionali basato su un insieme di tecnologie mature e sviluppate nel corso dell'ultimo decennio.

L'attuale mercato del Cloud computing può essere riassunto dalla seguente figura: ^[4]



3.2 Cloud commerciali e loro soluzioni di monitoraggio

In questo paragrafo sono citati alcuni tra i maggiori Cloud provider commerciali presenti attualmente sul mercato con breve riferimento ai servizi offerti e alle piattaforme di monitoraggio utilizzate.

3.2.1 Amazon Web Services



Amazon ^[18] è stata una delle prime compagnie a lanciare al grande pubblico servizi di Cloud computing e oggi, attraverso una collezione di servizi web tra cui i più noti sono Simple Storage Service (S3) ed Elastic Compute Cloud (EC2) offerti sul mercato attraverso il sito web amazon.com, è divenuta leader del mercato.

Il primo, (S3), consente agli utenti registrati di avere a disposizione un "online storage web service", ovvero un servizio che consente loro di depositare sulle strutture di memorizzazione messe a disposizione da Amazon i propri dati e accedere a essi in download e upload in ogni momento attraverso interfaccia web.

Amazon fornisce tale servizio, dietro pagamento di una quota mensile e in funzione dei gigabyte usufruiti nella memorizzazione e della banda usata nei trasferimenti dei dati. Amazon ha dichiarato di mantenere memorizzati 762 bilioni di oggetti nel 2011, rispetto ai 102 bilioni del 2010, e ai 64 bilioni del 2009.

Il secondo servizio, EC2, è di maggiore interesse ai nostri scopi, perché, con esso, gli utenti possono noleggiare istanze di macchine virtuali, che prendono il nome di Amazon Machine Instances (AMIs), sulle quali creare la propria struttura informatica e far girare le proprie applicazioni. Un utente, può creare, lanciare e terminare le istanze in base alle proprie esigenze, pagando in base alle ore in cui queste sono attive; da ciò viene il termine Elastic.

[5] Per quanto riguarda il monitoraggio (così come fanno anche altri Cloud commerciali), il low-level monitoring system che Amazon utilizza per acquisire informazioni sui suoi cluster fisici è tenuto segreto. È invece fornito agli utenti un servizio che prende il nome di CloudWatch, una piattaforma che permette il monitoraggio dei servizi come quello di EC2, dove le informazioni ricavate e accessibili dai clienti del Cloud sono strettamente riferite alle loro piattaforme virtuali.

Il servizio CloudWatch raccoglie i valori di differenti tipi di misurazioni e li conserva per un periodo di due settimane, che rappresenta il periodo di scadenza di tutte le misure disponibili, realizzando un archivio storico di riferimento, dei grafici e delle statistiche per valutare l'andamento, nel tempo, di tali valori. CloudWatch è attualmente un meccanismo generale di misurazione, aggregazione e recupero di questi dati ed inoltre, ha anche una funzione di allarme, che scatta ogni qualvolta una misura acquisita supera o è al di sotto di una soglia pre-configurata.

È un servizio che Amazon addebita separatamente, con un singolo prezzo per ora, indipendentemente dalle risorse monitorate. Recentemente Amazon sta cambiando il basic monitoring plan che prevede la raccolta delle informazioni ogni cinque minuti, mirando a renderlo gratuito e lasciando a pagamento il detail monitoring plan che consente raccolte di dati ogni minuto.

3.2.2 Microsoft Windows Azure



Windows Azure [19] è la piattaforma Cloud offerta dalla Microsoft che consente di compilare rapidamente, distribuire e gestire applicazioni attraverso una rete globale di data center gestiti da Microsoft. È possibile compilare applicazioni utilizzando qualsiasi linguaggio, strumento o framework. E' classificata come PaaS ed è parte dei servizi Cloud offerti da Microsoft, insieme per esempio a SQL Azure (può essere vista come versione based-cloud di SQL Server) e Windows Azure AppFabric (insieme di servizi di supporto alle applicazioni).

[5] Per consentire il monitoraggio delle applicazioni sviluppate su Windows Azure, gli sviluppatori hanno a disposizione una libreria software che facilita la diagnostica e il controllo delle applicazioni. Questa libreria è integrata nel Azure SDK ed è dotata di:

- performance counters che sono user-defined e possono riguardare qualsiasi valore dell'applicazione Cloud che è quantificabile.
- strutture di registrazione (logs) con le quali consentire allo sviluppatore di registrare ciò che è di suo interesse da cui partire per diagnosticare eventuali problematiche.

Il prezzo da pagare non è per l'uso della libreria ma per usufruire di un servizio di storage per conservare in maniera permanente i file di log.

3.2.3 RackSpace



RackSpace^[20] è da anni una dei maggiori provider di web-hosting; con il lancio della piattaforma RackSpace Cloud (originariamente come Mosso LLC) fu tra le prime compagnie a fornire servizi Cloud ed è oggi uno dei maggiori provider IaaS.

I tre servizi offerti sono Cloud Files, Cloud Servers e Cloud Sites.

Il primo è un servizio di storage online (come fa Amazon con S3), il secondo fornisce agli utenti piattaforme virtuali (i Cloud Server), per l'installazione delle proprie applicazioni, (come fa Amazon con EC2), mentre il terzo è simile a un tradizionale servizio di web-hosting sviluppato a partire da un'elevata scalabilità dell'infrastruttura hardware.

Cloud Sites include l'utilizzo mensile, da parte del cliente, di 10.000 compute cycle, che è approssimativamente equivalente all'esecuzione su di un server di un moderno processore da 2.8 GHz nello stesso periodo.

[5] RackSpace consente ai propri clienti attraverso lo stesso Cloud Sites di monitorare alcuni parametri come il contatore del compute cycle e l'uso della banda.

Inoltre vengono forniti i così detti Cloud tools, strumenti che forniscono una completa soluzione di monitoraggio sia generale sia specifica sulle macchine virtuali, potendo anche configurare sistemi di allerta.

Recentemente, RackSpace ha acquisito CloudKick, una piattaforma di gestione multi-cloud che consente una vasta gamma di funzionalità di monitoraggio sulle macchine virtuali; include differenti metriche di monitoraggio di basso e/o alto livello come utilizzazione di CPU, RAM, disk, banda, latenza e altro.

Il monitoraggio delle metriche può essere esteso da custom plug-in, permettendo la gestione di tutto ciò che esso definisce.

I dati possono essere visualizzati in tempo reale, attraverso appositi tools di visualizzazione forniti, mentre funzionalità di allerta configurabili consentono al cliente di essere sempre al corrente di eventuali problematiche, attraverso invii di e-mail o SMS.

3.2.4 Google App Engine



La società Google è specializzata in SaaS e PaaS.

Come SaaS, viene offerto Google Apps: un insieme di applicazioni web con funzionalità simili ai tradizionali office suites, come Gmail, Google Groups, Google Calendar, Google Talk, Google Docs e Google Sites.

Come PaaS viene fornito Google App Engine ^[21], una piattaforma di sviluppo e hosting web applications gestite dai data center della Google. Sono messe a disposizione due sandbox collaudate per sviluppare le proprie applicazioni: una in Java e l'altra in Python e anche un suo framework in Java per creare applicazioni Web 2.0, Google Web Toolkit. Esiste infine una terza Sandbox implementata in via sperimentale che consente di programmare in Go. Go è un linguaggio di programmazione semplice, nuovo, estremamente efficiente e nato proprio per lavorare su ambienti cluster e Cloud.

Un discorso a parte merita il datastore. Per memorizzare i dati da gestire con l'applicazione, Google mette a disposizione GQL un suo sistema molto simile al SQL classico, ma con alcune limitazioni, per renderlo efficiente su un sistema estremamente distribuito come è il suo Cloud. GQL, ad esempio, permette di lavorare con una sola tabella alla volta e non consente, in alcun modo, le operazioni di join. In pratica tutte le funzionalità "relazionali" sono disabilitate, cosa che lo rende molto simile a NoSQL.

Per quanto riguarda il monitoraggio ci si avvale di uno strumento fornito dalla Hyperic ovvero CloudStatus ^[9], uno tra i primi servizi di monitoraggio indipendenti; utilizzato, tra l'altro, anche su piattaforme Amazon. Fornisce all'utente una visione delle performance dei servizi Cloud e metodologie per determinare le cause di eventuali cambiamenti nelle performance delle applicazioni residenti nel Cloud. È un servizio che fornisce report in tempo reale e trends settimanali, riportando tempi di risposta, latenze e altre metriche specifiche per valutare le performance dei servizi monitorati.

In parole povere, consente agli sviluppatori di avere sempre una finestra aperta sulle proprie applicazioni nel Cloud, in modo da valutarne l'andamento e tutelare il proprio business.

3.2.5 Altre piattaforme commerciali

Per terminare la nostra panoramica sui Cloud commerciali, possiamo citare GoGrid, che fornisce servizi d'infrastruttura ed è il maggiore concorrente di RackSpace; Wmware che presenta sia soluzioni PaaS che servizi per Cloud private in generale; Salesforce.com specializzato nel fornire soluzioni CRM (Customer relationship management) a diversi tipi di aziende; infine non può mancare l'Apple che con i-cloud permette ai suoi utenti di accedere a musica, foto e documenti direttamente dal telefono cellulare o dall'i-Pad, ovunque si trovino.

Così come accade solitamente, per ogni nuovo modello nascente, si genera un indotto di aziende che producono servizi e prodotti direttamente collegati. Si è vista, così, la nascita di una serie di piattaforme che rispondono a una sola esigenza, quella di monitorare le applicazioni su Cloud e le varie prestazioni connesse, in modo da raccogliere in un'unica schermata, tutti i dati essenziali al lavoro di tutti i giorni.

Oltre ai già citati CloudKick e CloudStatus vale la pena ricordare: ^[22]

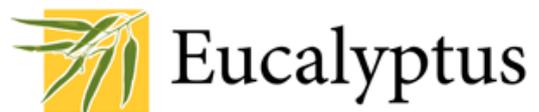
- Nimsoft, che con la sua soluzione Nimsoft Monitoring Solution (NMS), monitora le applicazioni presenti sia nei data center di Cloud private che pubbliche. Consente questo ai propri utenti, attraverso una singola "unified monitoring dashboard", permettendo piena visibilità sull'intera IT infrastructure e sui servizi forniti da Google Apps, RackSpace Cloud, Amazon web services, Salesforce.com e altri.
- Monitis che, attraverso i suoi agenti installati automaticamente sui server da monitorare, informa gli utenti del servizio sulle performance e genera notifiche quando le risorse sono ritenute scarse. È un servizio attualmente centrato principalmente sul Cloud di Amazon.

3.3 Cloud open-source e le loro soluzioni di monitoraggio

Il mercato Open-source è ancora agli albori, anche se negli ultimi anni si sono affermate alcune piattaforme Cloud affidabili e sono emersi nuovi progetti.

Per quanto riguarda l'offerta di piattaforme di monitoraggio, il mercato è ancora povero ma ci sono diversi gruppi di ricercatori che lavorano per proporre soluzioni adeguate; in seguito ne viene citato uno.

3.3.1 Eucalyptus



Eucalyptus^[23] (il cui nome è un acronimo che sta per "Elastic Utility Computing Architecture for Linking Your Programs To Useful System") è una piattaforma open source per la creazione di Cloud computing. Il sistema è compatibile con i servizi Amazon EC2 e S3 ma, essendo la piattaforma modulare, permette la compatibilità con altri tipi di servizi. Eucalyptus funziona con la maggior parte delle distribuzioni Linux attualmente disponibili (tra cui Ubuntu), ma anche con Microsoft Windows e può utilizzare una varietà di tecnologie di virtualizzazione compresi VMware, Xen e KVM hypervisor per attuare le Cloud abstractions.

Eucalyptus implementa IaaS (Infrastructure as a Service) Cloud privato e ibrido. La piattaforma fornisce una singola interfaccia che consente agli utenti di accedere a risorse di calcolo delle infrastrutture (macchine, reti e storage), disponibile in Cloud private.

Fornisce, inoltre, il proprio insieme di strumenti a linea di comando, chiamato Euca2ools, che può essere utilizzato internamente, per interagire con le installazioni Cloud di

Eucalyptus private, o esternamente, per interagire con le offerte di Cloud pubblico, tra cui Amazon EC2.

^[5] Fin dalla versione 2.0 Eucalyptus ha introdotto capacità di monitoraggio per i suoi componenti in esecuzione, per le macchine virtuali istanziate e per i servizi di storage.

Ciò viene realizzato attraverso l'integrazione di Eucalyptus in servizi di monitoraggio già esistenti, come Nagios (al quale dedicheremo qualche riga dopo), attraverso aggiornamenti delle loro configurazioni.

3.3.2 OpenNebula



OpenNebula^{[24][25]} è un toolkit di Cloud computing open source per la gestione delle infrastrutture di data center distribuiti e eterogenei. OpenNebula gestisce un'infrastruttura virtuale di elaborazione dati per costruire IaaS (Infrastrutture come servizio) pubbliche, private ed ibride. Il toolkit fornisce funzioni per l'integrazione, la gestione, la scalabilità, la sicurezza e la gestione delle risorse. Inoltre si focalizza sulla standardizzazione, sulla interoperabilità e sulla portabilità, fornendo agli utenti e agli amministratori del Cloud diverse interfacce (EC2 Query, OGF OCCI e vCloud) e hypervisor (Xen, KVM and VMware), oltre ad un'architettura flessibile che può far convivere diverse combinazioni di hardware e software nello stesso data center.

^[5] Le capacità di monitoraggio di OpenNebula sono focalizzate sull'interesse dei Cloud provider sulle risorse fisiche. Questa funzionalità si trova in un modulo chiamato Information Manager. Esso lavora utilizzando "sonde" per ricavare informazioni dai nodi del cluster. Le sonde sono costituite da script eseguiti sui nodi fisici che hanno come output coppie "attributo = valore" che vengono raccolti e centralizzati; hanno bisogno che i nodi fisici siano raggiungibili con SSH senza password.

Attualmente, queste sonde sono focalizzate solo nel raccogliere informazioni che sottolineano lo stato dei nodi fisici e non quello delle macchine virtuali che vi girano sopra.

Per avere anche queste informazioni la comunità di OpenNebula raccomanda di usare strumenti di service manager come entità separate da OpenNebula, soluzioni commerciali specializzate o rivolgersi a soluzioni di monitoraggio cluster open-source.

3.3.3 OpenStack



OpenStack ^[26] è un progetto open source per il Cloud Computing IaaS, nato nel 2010 da una collaborazione tra RackSpace e la Nasa, successivamente affiancate da più di 100 aziende tra cui Citrix Systems, Dell, AMD, Intel, Nec, Linux, Hp e Cisco. Si tratta di una piattaforma che sta guadagnando rapidamente consensi; basti pensare che nelle prossime versioni di Ubuntu, dovrebbe essere incluso proprio il supporto alle API di OpenStack per creare istanze virtuali basate su questa piattaforma.

OpenStack è composto da 3 progetti principali.

- *Compute*: è il progetto che si occupa della creazione e della gestione dell'intera piattaforma Cloud, creando reti di macchine virtuali scalabili e interconnesse in modo sicuro. Offre, inoltre, le API necessarie all'amministrazione dell'infrastruttura e le interfacce per la gestione delle istanze in esecuzione e della rete.
- *Object Store*: è un sistema per l'archiviazione a lungo termine di grandi quantità di dati statici (definiti oggetti), che possono essere recuperati, controllati e aggiornati. Si appoggia su cluster di server standard e ha un'architettura completamente distribuita, senza la presenza di nodi centrali o super nodi. Questo garantisce la scalabilità, la ridondanza e la permanenza dei dati.
- *Image Service* offre le funzionalità per la gestione completa, la registrazione e il recupero delle immagini delle istanze gestite dai vari nodi compute, permettendo la loro memorizzazione in un normale file system oppure utilizzando i servizi offerti da Object Store.

Il monitoraggio in tale ambiente può essere svolto attraverso Nagios ^[27], un sistema open source che si occupa di controllare le risorse degli host (carico della cpu, memoria utilizzata, numero di processi attivi, etc.), di verificare la disponibilità dei servizi offerti e di inviare notifiche nel caso di malfunzionamenti. Talvolta può effettuare azioni preventive per limitare i danni. Nagios è l'acronimo di Notice Any Glitches In Our System (notifica qualsiasi malfunzionamento nel nostro sistema).

Nagios presenta una struttura modulare composta da tre componenti, che ne consente grande flessibilità e adattabilità in vari ambienti:

- Nagios core è un demone, un programma sempre attivo eseguito in background e residente in memoria. È il fulcro dell'infrastruttura di monitoraggio e consente di effettuare diverse operazioni:
 - monitorare qualsiasi servizio senza essere a conoscenza dell'effettiva entità che si sta monitorando, grazie all'uso dei plug-in;

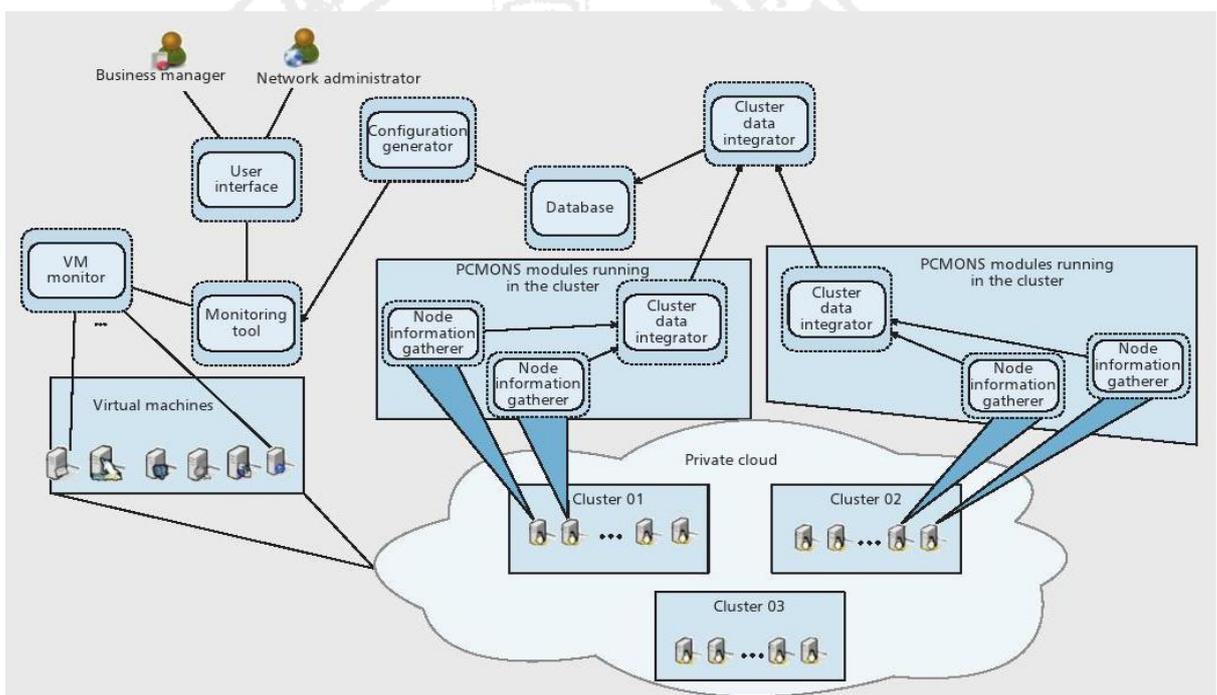
- definire event handler da eseguirsi all'occorrenza di determinati problemi, al fine di risolverli tempestivamente;
- visualizzare rapidamente lo stato corrente della rete, tramite una semplice interfaccia web che contiene le notifiche inviate, i file di log e la lista problemi verificatisi.
- Plugin: A differenza di tanti altri tool, Nagios non include alcun meccanismo concreto per il monitoraggio dei servizi. I veri artefici del monitoraggio sono i Nagios Plugin, script o eseguibili scritti da terzi in grado di monitorare host e servizi sulla rete.
- Estensioni modulari o Addon: Nagios presenta una struttura flessibile e facilmente ampliabile attraverso estensioni modulari o Addon. Questo permette di caricare solo le funzionalità necessarie, senza appesantire inutilmente il sistema, consentendo possibili estensioni future.

3.3.4 Altre soluzioni open-source

Oltre alle già citate proposte per il Cloud computing, nel mondo open-source ci sono anche altre soluzioni Cloud altrettanto importanti, come CloudStack, un software open source scritto in Java che permette di implementare diversi tipi di servizi Cloud, e Nimbus.

Come già detto in precedenza, le strumentazioni open-source per il monitoraggio sono ancora poche. Potrei citare, per esempio, Zenoss che fornisce una soluzione per gestire dispositivi fisici e virtuali, specializzato nei servizi di Amazon. Tanti gruppi di ricerca però sono attivi per proporre soluzioni valide. Tra questi, uno che ha pubblicato un articolo nel dicembre 2011 sull'IEEE Communications Magazine, propone un proprio monitoring framework per Cloud private chiamato PCMONS (Private Cloud MONitoring System) [28].

Sono previsti diversi moduli e la figura seguente mostra un possibile scenario:



- **Node Information Gatherer:** Questo modulo è responsabile di raccogliere informazioni locali sui nodi Cloud. La raccolta d'informazioni potrebbe variare secondo specifiche esigenze, anche se l'attuale versione raccoglie informazioni sulle VMs locali e invia questi al Cluster Data Integrator.
- **Cluster Data Integrator:** Come la maggior parte degli strumenti Cloud, organizza i nodi in clusters e c'è uno specifico agente che raccoglie e prepara i dati per i livelli successivi.
- **Monitoring Data Integrator:** Raccoglie e salva i dati riferiti al Cloud nel database e li fornisce al Configuration Generator.
- **VM Monitor:** Questo modulo inserisce script eseguiti sulle VMs monitorate, recuperando dati utili.
- **Configuration Generator:** Recupera informazioni dal database, per generare i file di configurazione necessari ai vari strumenti, come quelli di visualizzazione.
- **Monitoring Tool Server:** Questo modulo è preposto alla ricezione dei dati di monitoraggio delle differenti risorse e aggiornare il database. Attualmente sono stati presi alcuni accorgimenti per far passare le informazioni di monitoraggio direttamente all'interfaccia di Nagios.
- **User interface:** Molti monitoring tool hanno la propria interfaccia utente, consentendo un accesso diretto ai servizi forniti dal sistema, secondo le varie esigenze. L'interfaccia di Nagios è sufficiente, grazie ai file di configurazione creati dal monitoring tool server che consentono di controllare e visualizzare le informazioni di monitoraggio sulla sua interfaccia standard.

La prima release di PCMONS è compatibile con Eucalyptus per il livello infrastruttura e con Nagios per quello di view (livello che consente visualizzazione e analisi dei dati di monitoraggio). È stato scelto quest'ultimo per la popolarità, per la documentazione aperta e per la flessibilità, realizzata grazie all'estensibilità apportabile con lo sviluppo di nuovi plug-in. Inoltre, Eucalyptus fornisce alcuni semplici Nagios script per il monitoraggio di base dei suoi componenti. Questa breve descrizione è tratta dalla pubblicazione citata in precedenza. Da questa emerge che i futuri lavori, svolti del gruppo di ricerca, saranno incentrati a migliorare PCMONS così da supportare maggiori metriche di valutazione e adattarsi a diversi ambienti e strumenti Open-source.

3.4 Le soluzioni di CloudWatch e Nagios

Questo paragrafo, descrive come due, tra le principali problematiche evidenziate in questo testo, vengono affrontate da CloudWatch e Nagios:

- riuscire ad accedere ad un gran numero di risorse eterogenee per valutarne la salute e il buon funzionamento dei servizi.
- gestire gli eventi che si verificano durante il monitoraggio.

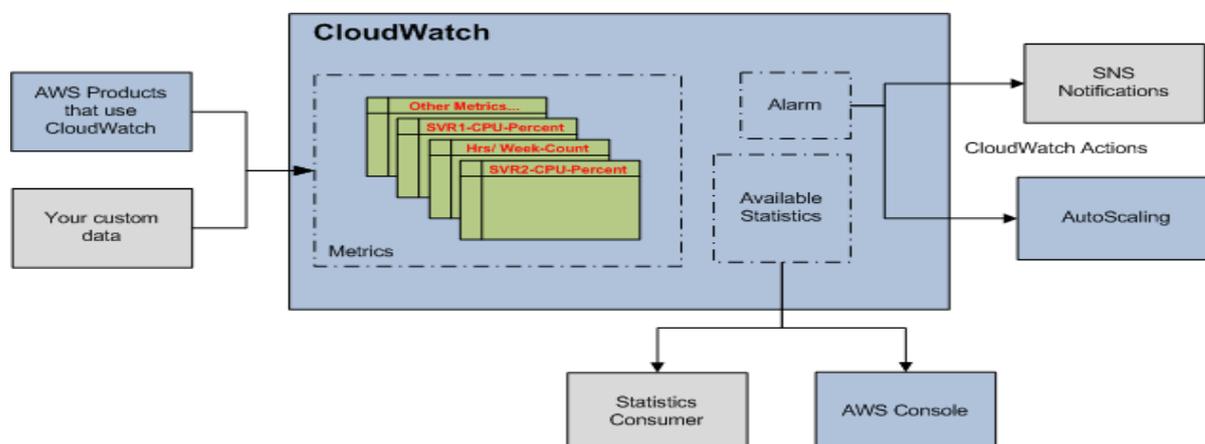
Per risolvere il primo, Amazon CloudWatch prevede molte metriche di controllo e la possibilità di definirne altre. Attraverso il pannello di controllo o tramite programmi personalizzati con l'uso delle API messe a disposizione, l'utente della piattaforma può gestirle a suo piacimento, attivando quelle che interessano al monitoraggio del particolare Amazon service.

Per EC2 le metriche disponibili sono: CPUUtilization, DiskReadOps, DiskWriteOps, DiskReadBytes, DiskWriteBytes, NetworkIn, NetworkOut, che valutano il consumo di risorse computazionali, disco e rete, da parte dell'istanza monitorata.

La gestione degli eventi viene realizzata, invece, con la creazione dei cosiddetti "alarms".

Basandosi sui dati, ottenuti dalle metriche scelte ed osservate per un certo tempo definito, gli alarms avviano una serie di azioni. Queste, comprendono la notifica degli eventi avvalendosi del servizio Amazon SNS e/o l'avvio di politiche di Autoscaling che consentono di adattare le risorse dell'istanza monitorata in base alle varie esigenze.

La figura seguente schematizza tutto ciò che è stato detto ^[29].



^[27] Per riuscire a monitorare le tante risorse e servizi di cui è composta un'infrastruttura Cloud, Nagios organizza la propria architettura in maniera flessibile, attraverso i plug-in. Il vantaggio di questo tipo di architettura consiste nella possibilità di monitorare qualsiasi cosa, senza essere limitati alle funzioni presenti nel tool. Gli sviluppatori possono così concentrarsi solo sul fulcro dell'infrastruttura di monitoraggio, lasciando a terzi la creazione di plug-in.

Nagios offre una svariata lista di controlli su servizi di base, quali memoria usata, carico della cpu, etc. Nel caso piuttosto frequente in cui i plug-in di base non dovessero soddisfare le esigenze del cliente, è possibile crearne di nuovi ad hoc. I plug-in ritornano dei valori che

descrivono lo stato di una risorsa o di un servizio. Quando si verifica un cambiamento di stato, l'event handler cattura questo evento. L'uso degli event handler rappresenta l'abilità di Nagios di risolvere preventivamente i problemi prima che essi siano notificati, o in ogni caso prima di un possibile intervento umano. Un uso comune degli event handler potrebbe essere il riavvio di un servizio fallito o il salvataggio dei file di log in un database. Fondamentalmente, esistono due tipi di event handler all'interno di Nagios: "global event handlers" e "specific event handlers".

Il primo tipo, definito direttamente da Nagios, ha capacità limitate di ripristino da guasti; il secondo, eseguito in caso di fallimento del primo, è definito dall'amministratore della risorsa stessa e permette un'azione correttiva definita ad hoc.

Entrambe le piattaforme, forniscono strumenti estendibili per un monitoraggio completo e una gestione immediata delle difficoltà, proponendo dove è possibile, valide soluzioni automatiche oltre alla semplice notifica.

3.5 Tabella comparativa tra le piattaforme di monitoraggio

La seguente tabella riassume e mette a confronto le diverse piattaforme di monitoraggio citate in precedenza, evidenziando a quali di queste appartengono le caratteristiche e funzionalità riportate.

	Cloud Watch	CloudKick	CloudStatus	Nimsoft	Monitis	Nagios	Zenoss
Multi-Cloud		✓	✓	✓		✓	
Multi-istanza*	✓	✓	✓	✓	✓	✓	✓
Estensibilità (utilizzo di plug-in o definizione di nuove metriche)	✓	✓				✓	
Report in Real-time	✓	✓	✓	✓	✓	✓	✓
Trend settimanali	✓		✓	✓			✓
Realizzazione di grafici & statistiche	✓	✓	✓	✓	✓	✓	✓
Funzionalità d'allerta	✓	✓	✓	✓	✓	✓	✓
Notifiche attraverso SMS & E-mail	✓	✓	✓			✓	✓

* Si intende per Multi-istanza, la possibilità di monitorare e visualizzare contemporaneamente l'andamento di più piattaforme virtuali, anche di diversi provider, se è multi-Cloud.

Breve conclusione

In questo lavoro sono stati considerati gli aspetti generali del Cloud computing, valutando con la giusta obiettività gli aspetti a favore e considerando con altrettanta obiettività alcune perplessità emerse. Si rileva l'esigenza dell'introduzione di un monitoraggio che vada a coprire i diversi aspetti d'interesse: dall'analisi della qualità delle prestazioni, alla sicurezza e privacy dei dati.

Da questo studio è emerso che il Cloud non è più solo una prospettiva, ma una vera realtà. Osservando il mercato si nota che è pieno di valide soluzioni, adottate sia nel settore pubblico, che in quello privato. Anche in Italia qualcosa si sta muovendo, in tal senso.

In questi anni, si vanno sviluppando iniziative di standardizzazione, come "Open Cloud Standards Incubator" e "The Open Cloud Computing Interface (OCCI) Working group", che mirano a definire le basi, per l'interoperabilità dei vari ambienti Cloud, e le interazioni con i consumatori; mentre qualche soluzione Open-source garantisce che questo nuovo modello di servizio non resti prerogativa di pochi grandi colossi commerciali.

Monitorare e testare i servizi offerti dai provider è un aspetto fondamentale per valutare i reali vantaggi di questo modello, rispetto all'hosting tradizionale e, com'è emerso da alcuni lavori di ricerca citati, difficoltà e ostacoli non mancano sia nella valutazione che nella realizzazione di buone prestazioni.

L'obiettivo di questo lavoro è di mostrare l'importanza del monitoraggio, come aspetto cruciale per la realizzazione efficiente ed efficace dei servizi Cloud, ma anche di come il controllo e il testing delle prestazioni può guidare gli utilizzatori a scelte consapevoli sulla qualità dei servizi offerti dai vari provider.

Bibliografia

- [1] http://www.moebiusonline.eu/fuorionda/cloud_computing.shtml/
- [2] <http://www.slideshare.net/ilcavi87/cloud-computing-andrea-cavicchini>
- [3] <http://www.hostingtalk.it/articoli/cloud-computing/5264/cloud-computing-cos-e-e-quali-sono-le-aspettative-attese-e-disattese-f?page=3/>
- [4] [www.masterprocurement.it/ckfinder/userfiles/files/Cardellini\(1\).ppt](http://www.masterprocurement.it/ckfinder/userfiles/files/Cardellini(1).ppt)
- [5] Eddy Caron, Frédéric Desprez, Luis Rodero-Merino, Adrian Muresan University of Lyon - CNRS - ENS Lyon - UCB Lyon 1 – INRIA LIP Laboratory, UMR 5668. Lyon, France, “Auto-Scaling, Load Balancing and Monitoring in Commercial and Open-Source Clouds”
- [6] Lori M. Kaufman, Bruce Potter, IEEE SECURITY E PRIVACY MARCH/APRIL 2011 “Monitoring Cloud computing, by Layer part1”
- [7] Dr. Gurdev Singh, Shanu Sood, Amit Sharma, Computer science & Engineering Eternal University Baru Sahib (India) , “CM- Measurement Facets for Cloud Performance”
- [8] Yiduo Mei, Ling Liu, Xing Pu, Sankaran Sivathanu, “Performance Measurements and Analysis of Network I/O Applications in Virtualized Cloud”
- [9] <http://www.hyperic.com/products/cloud-status-monitoring>
- [10] Vinod Venkataraman, Ankit Shah , Yin Zhang , “Network-based Measurements on Cloud Computing Services”
- [11] Google App Engine, <http://appengine.google.com/>
- [12] PlanetLab , <http://www.planet-lab.org/>

- [13] Httpperf tool, <http://www.hpl.hp.com/research/linux/httpperf/>
- [14] Peer Hasselmeyer, Nico d'Heureuse, NEC Laboratories Europe, NEC Europe, Ltd, 69115 Heidelberg, Germany, "Towards Holistic Multi-Tenant Monitoring for Virtual Data Centers"
- [15] Shicong Meng, Student Member, IEEE, Ling Liu, Senior Member, IEEE, and Ting Wang, Student Member, IEEE , " State Monitoring in Cloud Datacenters"
- [16] Michael Hauck, Jens Happe, Ralf H. Reussner Research Center for Information Technology, Karlsruhe, Germany "TOWARDS PERFORMANCE PREDICTION FOR CLOUD COMPUTING ENVIRONMENTS BASED ON GOAL-ORIENTED MEASUREMENTS"
- [17] <http://www.mycloudbeans.com/content/cloud-computing-rebrand-o-nuova-tecnologia-breve-storia-del-computing-dai-mainframe-al-cloud?page=0,1>
- [18] <http://aws.amazon.com>.
- [19] <http://www.windowsazure.com>
- [20] <http://www.rackspace.com/>
- [21] <http://www.azns.it/2011/06/17/cose-google-app-engine/>
- [22] <http://www.datacenterknowledge.com/archives/2010/02/04/cloud-monitoring-services-a-resource-guide/>
- [23] <http://www.infolinux.it/utility/eucalyptus-il-cloud-open-source/>
- [24] <http://openebula.org/documentation:archives:rel2.0:img>
- [25] <http://en.wikipedia.org/wiki/OpenNebula>
- [26] <http://docs.openstack.org/diablo/openstack-compute/admin/os-compute-adminguide-trunk.pdf>
- [27] <http://nagios.sourceforge.net/docs/nagioscore-3-en.pdf>
- [28] De Chaves, S.A.; Uriarte, R.B.; Westphall, C.B.; Post Graduation Program in Comput. Sci., Fed. Univ. of Santa Catarina, Florianopolis, Brazil; "Toward an Architecture for Monitoring Private Clouds"
- [29] <http://awsdocs.s3.amazonaws.com/AmazonCloudWatch/latest/acw-dg.pdf>