

UNIVERSITÀ DELLA CALABRIA
Facoltà di Scienze Matematiche Fisiche e Naturali
Corso di laurea in Informatica

TESI DI LAUREA

TITOLO

Cloud Computing con Software Libero

Relatori

Prof. Salvatore Di Gregorio

Dott. Vincenzo Bruno

Ing. Alessandro Tarasio

Candidato

Vincenzo Pirrone

Anno Accademico 2009-2010

L'OPERA (COME SOTTO DEFINITA) È MESSA A DISPOSIZIONE SULLA BASE DEI TERMINI DELLA PRESENTE LICENZA "CREATIVE COMMONS PUBLIC LICENCE" ("CCPL" O "LICENZA"). L'OPERA È PROTETTA DAL DIRITTO D'AUTORE E/O DALLE ALTRE LEGGI APPLICABILI. OGNI UTILIZZAZIONE DELL'OPERA CHE NON SIA AUTORIZZATA AI SENSI DELLA PRESENTE LICENZA O DEL DIRITTO D'AUTORE È PROIBITA.

CON IL SEMPLICE ESERCIZIO SULL'OPERA DI UNO QUALUNQUE DEI DIRITTI QUI DI SEGUITO ELENCATI, TU ACCETTI E TI OBBLIGHI A RISPETTARE INTEGRALMENTE I TERMINI DELLA PRESENTE LICENZA AI SENSI DEL PUNTO 8.f. IL LICENZIANTE CONCEDE A TE I DIRITTI QUI DI SEGUITO ELENCATI A CONDIZIONE CHE TU ACCETTI DI RISPETTARE I TERMINI E LE CONDIZIONI DI CUI ALLA PRESENTE LICENZA.

1. Definizioni. Ai fini e per gli effetti della presente licenza, si intende per

- a. "Collezione di Opere", un'opera, come un numero di un periodico, un'antologia o un'enciclopedia, nella quale l'Opera nella sua interezza e forma originale, unitamente ad altri contributi costituenti loro stessi opere distinte ed autonome, sono raccolti in un'unità collettiva. Un'opera che costituisce Collezione di Opere non verrà considerata Opera Derivata (come sotto definita) ai fini della presente Licenza;
- b. "Opera Derivata", un'opera basata sull'Opera ovvero sull'Opera insieme con altre opere preesistenti, come una traduzione, un arrangiamento musicale, un adattamento teatrale, narrativo, cinematografico, una registrazione di suoni, una riproduzione d'arte, un digesto, una sintesi, o ogni altra forma in cui l'Opera possa essere riproposta, trasformata o adattata. Nel caso in cui un'Opera tra quelle qui descritte costituisca già Collezione di Opere, essa non sarà considerata Opera Derivata ai fini della presente Licenza. Al fine di evitare dubbi è inteso che, quando l'Opera sia una composizione musicale o registrazione di suoni, la sincronizzazione dell'Opera in relazione con un'immagine in movimento ("synching") sarà considerata Opera Derivata ai fini di questa Licenza;
- c. "Licenziante", l'individuo o l'ente che offre l'Opera secondo i termini e le condizioni della presente Licenza;
- d. "Autore Originario", il soggetto che ha creato l'Opera;
- e. "Opera", l'opera dell'ingegno suscettibile di protezione in forza delle leggi sul diritto d'autore, la cui utilizzazione è offerta nel rispetto dei termini della presente Licenza;
- f. "Tu"/"Te", l'individuo o l'ente che esercita i diritti derivanti dalla presente Licenza e che non abbia precedentemente violato i termini della presente Licenza relativi all'Opera, o che, nonostante una precedente violazione degli stessi, abbia ricevuto espressa autorizzazione dal Licenziante all'esercizio dei diritti derivanti dalla presente Licenza.

2. Libere utilizzazioni. La presente Licenza non intende in alcun modo ridurre, limitare o restringere alcun diritto di libera utilizzazione o l'operare della regola dell'esaurimento del diritto o altre limitazioni dei diritti esclusivi sull'Opera derivanti dalla legge sul diritto d'autore o da altre leggi applicabili.

3. Concessione della Licenza. Nel rispetto dei termini e delle condizioni contenute nella presente Licenza, il Licenziante concede a Te una licenza per tutto il mondo, gratuita, non esclusiva e perpetua (per la durata del diritto d'autore applicabile) che autorizza ad esercitare i diritti sull'Opera qui di seguito elencati:

- a. riproduzione dell'Opera, incorporazione dell'Opera in una o più Collezioni di Opere e riproduzione dell'Opera come incorporata nelle Collezioni di Opere;
- b. creazione e riproduzione di un'Opera Derivata;
- c. distribuzione di copie dell'Opera o di supporti fonografici su cui l'Opera è registrata, comunicazione al pubblico, rappresentazione, esecuzione, recitazione o esposizione in pubblico, ivi inclusa la trasmissione audio digitale dell'Opera, e ciò anche quando l'Opera sia incorporata in Collezioni di Opere;

- d. distribuzione di copie dell'Opera o di supporti fonografici su cui l'Opera Derivata è registrata, comunicazione al pubblico, rappresentazione, esecuzione, recitazione o esposizione in pubblico, ivi inclusa la trasmissione audio digitale di Opere Derivate;
- e. Al fine di evitare dubbi è inteso che, se l'Opera sia di tipo musicale:
 - i. Compensi per la comunicazione al pubblico o la rappresentazione o esecuzione di opere incluse in repertori. Il Licenziante rinuncia al diritto esclusivo di riscuotere compensi, personalmente o per il tramite di un ente di gestione collettiva (ad es. SIAE), per la comunicazione al pubblico o la rappresentazione o esecuzione, anche in forma digitale (ad es. tramite webcast) dell'Opera.
 - ii. Compensi per versioni cover. Il Licenziante rinuncia al diritto esclusivo di riscuotere compensi, personalmente o per il tramite di un ente di gestione collettiva (ad es. SIAE), per ogni disco che Tu crei e distribuisce a partire dall'Opera (versione cover).
- f. Compensi per la comunicazione al pubblico dell'Opera mediante fonogrammi. Al fine di evitare dubbi, è inteso che se l'Opera è una registrazione di suoni, il Licenziante rinuncia al diritto esclusivo di riscuotere compensi, personalmente o per il tramite di un ente di gestione collettiva (ad es. IMAIE), per la comunicazione al pubblico dell'Opera, anche in forma digitale.
- g. Altri compensi previsti dalla legge italiana. Al fine di evitare dubbi, è inteso che il Licenziante rinuncia al diritto esclusivo di riscuotere i compensi a lui attribuiti dalla legge italiana sul diritto d'autore (ad es. per l'inserimento dell'Opera in un'antologia ad uso scolastico ex art. 70 l. 633/1941). Al Licenziante spettano in ogni caso i compensi irrinunciabili a lui attribuiti dalla medesima legge (ad es. l'equo compenso spettante all'autore di opere musicali, cinematografiche, audiovisive o di sequenze di immagini in movimento nel caso di noleggio ai sensi dell'art. 18-bis l. 633/1941).

I diritti sopra descritti potranno essere esercitati con ogni mezzo di comunicazione e in tutti i formati. Tra i diritti di cui sopra si intende compreso il diritto di apportare all'Opera le modifiche che si rendessero tecnicamente necessarie per l'esercizio di detti diritti tramite altri mezzi di comunicazione o su altri formati. Tutti i diritti non espressamente concessi dal Licenziante rimangono riservati.

4. Restrizioni. La Licenza concessa in conformità al precedente punto 3 è espressamente assoggettata a, e limitata da, le seguenti restrizioni:

- a. Tu puoi distribuire, comunicare al pubblico, rappresentare, eseguire, recitare o esporre in pubblico l'Opera, anche in forma digitale, solo assicurando che i termini di cui alla presente Licenza siano rispettati e, insieme ad ogni copia dell'Opera (o supporto fonografico su cui è registrata l'Opera) che distribuisce, comunichi al pubblico o rappresenti, esegui, reciti o esponi in pubblico, anche in forma digitale, devi includere una copia della presente Licenza o il suo Uniform Resource Identifier. Non puoi proporre o imporre alcuna condizione relativa all'Opera che alteri o restringa i termini della presente Licenza o l'esercizio da parte del beneficiario dei diritti qui concessi. Non puoi concedere l'Opera in sublicenza. Devi mantenere intatte tutte le informative che si riferiscono alla presente Licenza ed all'esclusione delle garanzie. Non puoi distribuire, comunicare al pubblico, rappresentare, eseguire, recitare o esporre in pubblico l'Opera, neanche in forma digitale, usando misure tecnologiche miranti a controllare l'accesso all'Opera ovvero l'uso dell'Opera, in maniera incompatibile con i termini della presente Licenza. Quanto sopra si applica all'Opera anche quando questa faccia parte di una Collezione di Opere, anche se ciò non comporta che la Collezione di Opere di per sé ed indipendentemente dall'Opera stessa debba essere soggetta ai termini ed alle condizioni della presente Licenza. Qualora Tu crei una Collezione di Opere, su richiesta di qualsiasi Licenziante, devi rimuovere dalla Collezione di Opere stessa, ove materialmente possibile, ogni riferimento in accordo con quanto previsto dalla clausola 4.b, come da richiesta. Qualora tu crei un'Opera Derivata, su richiesta di qualsiasi Licenziante devi rimuovere dall'Opera Derivata stessa, nella misura in cui ciò sia possibile, ogni riferimento in accordo con quanto previsto dalla clausola 4.b, come da richiesta.
- b. Qualora Tu distribuisca, comunichi al pubblico, rappresenti, esegua, reciti o esponga in pubblico, anche in forma digitale, l'Opera o qualsiasi Opera Derivata o Collezione di Opere, devi mantenere intatte tutte le informative sul diritto d'autore sull'Opera. Devi riconoscere una menzione adeguata rispetto al mezzo di comunicazione o supporto che utilizzi: (i) all'Autore Originale (citando il suo nome o lo pseudonimo, se del caso), ove fornito; e/o (ii) alle terze parti designate, se l'Autore Originale e/o il Licenziante hanno designato una o più terze parti (ad esempio, una istituzione

finanziatrice, un ente editoriale) per l'attribuzione nell'informativa sul diritto d'autore del Licenziante o nei termini di servizio o con altri mezzi ragionevoli; il titolo dell'Opera, ove fornito; nella misura in cui sia ragionevolmente possibile, l'Uniform Resource Identifier, che il Licenziante specifichi dover essere associato con l'Opera, salvo che tale URI non faccia alcun riferimento alla informazione di protezione di diritto d'autore o non dia informazioni sulla licenza dell'Opera; inoltre, in caso di Opera Derivata, devi menzionare l'uso dell'Opera nell'Opera Derivata (ad esempio, "traduzione francese dell'Opera dell'Autore Originario", o "sceneggiatura basata sull'Opera originaria dell'Autore Originario"). Tale menzione deve essere realizzata in qualsiasi maniera ragionevole possibile; in ogni caso, in ipotesi di Opera Derivata o Collezione di Opere, tale menzione deve quantomeno essere posta nel medesimo punto dove viene indicato il nome di altri autori di rilevanza paragonabile e con lo stesso risalto concesso alla menzione di altri autori di rilevanza paragonabile.

5. Dichiarazioni, Garanzie ed Esonero da responsabilità

SALVO CHE SIA ESPRESSAMENTE CONVENUTO ALTRIMENTI PER ISCRITTO FRA LE PARTI, IL LICENZIANTE OFFRE L'OPERA IN LICENZA "COSÌ COM'È" E NON FORNISCE ALCUNA DICHIARAZIONE O GARANZIA DI QUALSIASI TIPO CON RIGUARDO ALL'OPERA, SIA ESSA ESPRESSA OD IMPLICITA, DI FONTE LEGALE O DI ALTRO TIPO, ESSENDO QUINDI ESCLUSE, FRA LE ALTRE, LE GARANZIE RELATIVE AL TITOLO, ALLA COMMERCIALIZZABILITÀ, ALL'IDONEITÀ PER UN FINE SPECIFICO E ALLA NON VIOLAZIONE DI DIRITTI DI TERZI O ALLA MANCANZA DI DIFETTI LATENTI O DI ALTRO TIPO, ALL'ESATTEZZA OD ALLA PRESENZA DI ERRORI, SIANO ESSI ACCERTABILI O MENO. ALCUNE GIURISDIZIONI NON CONSENTONO L'ESCLUSIONE DI GARANZIE IMPLICITE E QUINDI TALE ESCLUSIONE PUÒ NON APPLICARSI A TE.

6. Limitazione di Responsabilità. SALVI I LIMITI STABILITI DALLA LEGGE APPLICABILE, IL LICENZIANTE NON SARÀ IN ALCUN CASO RESPONSABILE NEI TUOI CONFRONTI A QUALUNQUE TITOLO PER ALCUN TIPO DI DANNO, SIA ESSO SPECIALE, INCIDENTALE, CONSEGUENZIALE, PUNITIVO OD ESEMPLARE, DERIVANTE DALLA PRESENTE LICENZA O DALL'USO DELL'OPERA, ANCHE NEL CASO IN CUI IL LICENZIANTE SIA STATO EDOTTO SULLA POSSIBILITÀ DI TALI DANNI. NESSUNA CLAUSOLA DI QUESTA LICENZA ESCLUDE O LIMITA LA RESPONSABILITÀ NEL CASO IN CUI QUESTA DIPENDA DA DOLO O COLPA GRAVE.

7. Risoluzione

- a. La presente Licenza si intenderà risolta di diritto e i diritti con essa concessi cesseranno automaticamente, senza necessità di alcuna comunicazione in tal senso da parte del Licenziante, in caso di qualsivoglia inadempimento dei termini della presente Licenza da parte Tua, ed in particolare delle disposizioni di cui ai punti 4.a e 4.b, essendo la presente Licenza condizionata risolutivamente al verificarsi di tali inadempimenti. In ogni caso, la risoluzione della presente Licenza non pregiudicherà i diritti acquistati da individui o enti che abbiano acquistato da Te Opere Derivate o Collezioni di Opere, ai sensi della presente Licenza, a condizione che tali individui o enti continuino a rispettare integralmente le licenze di cui sono parte. Le sezioni 1, 2, 5, 6, 7 e 8 rimangono valide in presenza di qualsiasi risoluzione della presente Licenza.
- b. Sempre che vengano rispettati i termini e le condizioni di cui sopra, la presente Licenza è perpetua (e concessa per tutta la durata del diritto d'autore sull'Opera applicabile). Nonostante ciò, il Licenziante si riserva il diritto di rilasciare l'Opera sulla base dei termini di una differente licenza o di cessare la distribuzione dell'Opera in qualsiasi momento; fermo restando che, in ogni caso, tali decisioni non comporteranno recesso dalla presente Licenza (o da qualsiasi altra licenza che sia stata concessa, o che sia richiesto che venga concessa, ai termini della presente Licenza), e la presente Licenza continuerà ad avere piena efficacia, salvo che vi sia risoluzione come sopra indicato.

8. Varie

- a. Ogni volta che Tu distribuischi, o rappresenti, esegui o reciti pubblicamente in forma digitale l'Opera o una Collezione di Opere, il Licenziante offre al destinatario una licenza per l'Opera nei medesimi termini e condizioni che a Te sono stati concessi dalla presente Licenza.
- b. Ogni volta che Tu distribuischi, o rappresenti, esegui o reciti pubblicamente in forma digitale un'Opera Derivata, il Licenziante offre al destinatario una licenza per l'Opera originale nei medesimi termini e condizioni che a Te sono stati concessi dalla presente Licenza.
- c. L'invalidità o l'inefficacia, secondo la legge applicabile, di una o più fra le disposizioni della presente Licenza, non comporterà l'invalidità o l'inefficacia dei restanti termini e, senza bisogno di ulteriori

azioni delle parti, le disposizioni invalide o inefficaci saranno da intendersi rettificata nei limiti della misura che sia indispensabile per renderle valide ed efficaci.

- d. In nessun caso i termini e le disposizioni di cui alla presente Licenza possono essere considerati rinunciati, né alcuna violazione può essere considerata consentita, salvo che tale rinuncia o consenso risultino per iscritto da una dichiarazione firmata dalla parte contro cui operi tale rinuncia o consenso.
- e. La presente Licenza costituisce l'intero accordo tra le parti relativamente all'Opera qui data in licenza. Non esistono altre intese, accordi o dichiarazioni relative all'Opera che non siano quelle qui specificate. Il Licenziante non sarà vincolato ad alcuna altra disposizione addizionale che possa apparire in alcuna comunicazione da Te proveniente. La presente Licenza non può essere modificata senza il mutuo consenso scritto del Licenziante e Tuo.
- f. Clausola iCommons. Questa Licenza trova applicazione nel caso in cui l'Opera sia utilizzata in Italia. Ove questo sia il caso, si applica anche il diritto d'autore italiano. Negli altri casi le parti si obbligano a rispettare i termini dell'attuale Licenza Creative Commons generica che corrisponde a questa Licenza Creative Commons iCommons.

Indice generale

INTRODUZIONE.....	3
1 Cos'è il Cloud Computing?.....	5
1.1 Caratteristiche chiave.....	7
1.2 Service Models.....	7
1.2.1 Software as a Service (SaaS).....	8
1.2.2 Platform as a Service (PaaS).....	9
1.2.3 Infrastructure as a Service (IaaS).....	9
1.3 Deployment Models.....	10
1.3.1 Public Cloud.....	10
1.3.2 Private Cloud.....	13
1.4 Cloud e Software Libero.....	14
1.5 Cloud e Virtualizzazione.....	16
1.6 Il Cloud Computing per la gente comune.....	16
1.6.1 Uso improprio del termine.....	17
1.7 Il Cloud Computing per le aziende.....	18
2 Piattaforme di Cloud Computing.....	19
2.1 Google Docs.....	19
2.2 Google App Engine.....	20
2.2.1 Java Runtime Environment.....	21
2.2.2 Datastore.....	21
2.2.3 Sviluppo.....	22
2.3 Amazon Web Services.....	22
2.3.1 Amazon Elastic Compute Cloud (EC2).....	23
2.3.2 Multiple Locations.....	23
2.3.3 Elastic IP Addresses.....	24
2.3.4 Security Groups.....	25
2.3.5 Amazon Elastic Block Store (EBS).....	26
2.3.6 Amazon Simple Storage Service (S3).....	26
2.3.7 Amazon CloudWatch.....	27
2.3.8 Elastic Load Balancing.....	27
2.3.9 Amazon Virtual Private Cloud (VPC).....	27
2.3.10 High Performance Computing (HPC) Clusters.....	28
2.3.11 Lavorare con gli AWS.....	28
2.4 OpenNebula.....	29
2.5 EyeOS.....	30
3 Caratteristiche di Eucalyptus.....	31
3.1 Architettura.....	31
3.1.1 Node Controller.....	33
3.1.2 Cluster Controller.....	33
3.1.3 Cloud Controller.....	34
3.2 Gestione degli utenti.....	35
3.3 Gestione EMI.....	35
3.4 Network Management.....	36
3.5 Controllo delle istanze.....	38
3.6 Euca2ools.....	39
3.7 Eucalyptus Community Cloud.....	39
3.8 Installazione.....	40
3.8.1 Ubuntu Enterprise Cloud (UEC).....	40
3.9 Eucalyptus Enterprise Edition.....	40
4 Installazione e sperimentazione di Eucalyptus.....	42
4.1 I Cluster del Grid del Dipartimento di Fisica.....	42

4.2 Cloud e HPC.....	43
4.3 Installazione di Eucalyptus.....	43
4.3.1 Setup di un nodo.....	44
4.3.2 Setup del front-end.....	45
4.4 Utilizzo di Eucalyptus.....	45
4.5 Avviare le macchine virtuali.....	49
4.6 Test e benchmark.....	51
4.7 Analisi dei risultati.....	52

INTRODUZIONE

Il lavoro di tesi si propone di analizzare una tecnologia molto in voga negli ultimi tempi, tra studiosi di informatica e non. Tuttavia sul significato di Cloud Computing e su cosa sia in realtà c'è davvero molta confusione. Il termine “nuvola” è un chiaro richiamo a internet, l'unica cosa in cui tutti sono d'accordo infatti è che per poter “fare” Cloud Computing è necessaria una connessione alle rete. Computing sulla nuvola in effetti rende abbastanza l'idea, qualcosa di simile ad utilizzare un computer attraverso il network, senza sapere cosa accade nella nube. A livello tecnico si tratta di usufruire di risorse di calcolo in remoto, siano esse CPU, memoria, disco, banda, sistemi operativi, librerie, ambienti o applicazioni; non solo, ma utilizzarle quando richiesto e nella misura necessaria. Sembra quasi una sorta di magia poter avere quello che si vuole quando si vuole (in ambito informatico), in realtà l'idea che sta alla base è molto semplice: poter ordinare e incrementare rapidamente le risorse necessarie per il proprio business senza dover acquisire hardware e software, senza dover configurare l'ambiente di lavoro, cercando di limitare i costi; rendere semplice l'utilizzo di servizi computazionali così come acqua, luce o gas.; come fossero servizi pubblici, riprendendo da questo modello anche la forma di pagamento a consumo.

Il Cloud Computing può quindi essere definito come un modello di fruizione delle risorse tecnologiche, le quali vengono portate al consumatore sotto forma di servizio; un nuovo modello capace di cambiare la concezione della rete, e in parte lo ha già fatto.

Il Cloud Computing ha avuto un enorme boom negli ultimi anni perché è ciò di cui la rete ha bisogno, offrire servizi nel modo più flessibile possibile. Data la forte crescita registrata nel campo sempre più aziende hanno cominciato a stanziare grandi investimenti, primi tra tutti i giganti Amazon, Yahoo, Google, IBM e Microsoft, che oggi offrono le principali e più svariate soluzioni Cloud.

Ma cosa c'è dentro la nuvola? Come vengono realizzati i servizi? Innanzi tutto, arrivare a ciò è stato possibile grazie allo sviluppo raggiunto oggi dal networking, dalle tecnologie web a da ciò che potrebbe essere definito il motore del modello Cloud, la virtualizzazione. Ma l'aspetto a cui si rivolge principalmente lo studio è il software. Nulla è infatti possibile senza un apparato software adeguato. Purtroppo le soluzioni software adottate dai grandi attori del settore sono proprietarie e non permettono il libero studio. Perciò questo lavoro riguarda il Cloud Computing

con software libero, per avere la libertà non solo di utilizzare liberamente il software, ma eventualmente di studiarlo nel suo interno e modificarlo secondo le proprie esigenze, come la definizione di software libero prevede.

Il lavoro di tesi è stato svolto in Uniclust, spin-off dell'Università della Calabria, il cui ambito principale è l'High Performance Computing. Le attività si sono svolte in collaborazione con il Dipartimento di Fisica dell'Università della Calabria¹ ed il Centro di Ricerca Hacklab Cosenza². Nella prima parte viene definito il concetto di Cloud Computing, analizzato in tutte le sue forme e studiate motivazioni, utilizzi, vantaggi e svantaggi. Lo studio prosegue con uno sguardo a servizi e piattaforme Cloud esistenti, sia libere che proprietarie, per dare un'idea di quello che il panorama offre. Di questi software ne è stato selezionato uno (libero ovviamente), per la sperimentazione, Eucalyptus, un sistema che permette facilmente il setup di macchine virtuali su misura. La piattaforma è stata installata su due macchine a disposizione dell'Uniclust, arrivando a toccare con mano il cuore della nuvola. Fine ultimo della tesi è testare le risorse usufruibili tramite Eucalyptus. La piattaforma Cloud è stata utilizzata per creare in forma virtuale un ambiente per il calcolo parallelo. Una soluzione Cloud è adatta a questo tipo di esigenze? Lo svantaggio della virtualizzazione è una certa perdita di prestazioni, lo studio effettuato consiste nel misurare la perdita e valutare se il Cloud Computing può essere utilizzato per High Performance Computing.

1 <http://fis.unical.it>

2 <http://hacklab.cosenzainrete.it>

1 Cos'è il Cloud Computing?

Inizialmente il termine cloud veniva utilizzato in riferimento alle reti di telecomunicazione e alla rete internet in quanto queste erano rappresentate come nuvole negli schemi tecnologici, a indicare aree in cui le informazioni venivano spostate ed elaborate; tuttavia, ciò avveniva senza che l'utente sapesse esattamente quel che stava realmente accadendo. Questa è la caratteristica principale del cloud computing: il cliente richiede e riceve informazioni o altre risorse senza sapere dove risiedono o secondo quale meccanismo il servizio nel cloud soddisfa la richiesta.[1]

Una definizione formale e ampiamente adottata è stata elaborata dal National Institute of Standards and Technology:

Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.[2]

Ovvero, un modello per abilitare un accesso conveniente e su richiesta a un insieme condiviso di risorse computazionali configurabili (ad esempio reti, server, memoria di massa, applicazioni e servizi), che possono essere rapidamente procurate e rilasciate con un minimo sforzo di gestione o di interazione con il fornitore del servizio.

Tipicamente le risorse vengono rilasciate da un fornitore, un Cloud Provider, in forma pubblica e pay-per-use. Le risorse possono essere rilasciate a diversi livelli a seconda delle esigenze del fruitore, a partire dall'applicazione che il cliente può utilizzare all'istante, al server su cui invece si ha completo controllo e capacità di personalizzazione. La definizione fa pensare molto a risorse "virtuali"; in effetti è proprio quello che sono, la tecnologia chiave del Cloud Computing è infatti la virtualizzazione.

Ma cosa comporta tutto ciò? Quali sono i possibili utilizzi? E i vantaggi? Innanzi tutto scompare quella fase in cui l'utente acquisisce i propri mezzi, in un certo senso l'informatica del Cloud è l'informatica senza computer, tutto sta nella nuvola. Ora l'utente non ha più bisogno di utilizzare il proprio PC, può usufruire delle sue applicazioni da qualunque postazione. La piccola azienda che ha bisogno di una piattaforma on-line non deve porsi più il problema di investire sui propri server o cercare un servizio con le caratteristiche adeguate, generalmente una soluzione Cloud fornisce tutto il necessario ad un costo ridotto; non solo, mentre spesso i servizi standard

risultano inadeguati, le risorse insufficienti o i costi di gestione troppo alti, la promessa del Cloud è di avere sempre e comunque il tipo di risorsa adatto, con tempi di cambiamento strettissimi. Ciò che prima richiedeva settimane (acquisire nuove macchine, acquisire personale, configurare il software) ora si è ridotto a pochi minuti e quelli che erano i costi più vari (hardware, tecnici, elettricità, raffreddamento, consulenze, outsourcing) ora sono tradotti in una sola formula, “si paga quello che si usa”.

Dal punto di vista del provider i benefici non sono minori, dotarsi di un'infrastruttura Cloud vuol dire riuscire a soddisfare varie tipologie di esigenze. Con gli stessi server a disposizione il modello Cloud permette di tenere pronti diversi tipi di ambiente ed essere sempre preparati a soddisfare le richieste dei clienti.

Molti esperti attribuiscono l'idea del Cloud Computing a John McCarthy che già negli anni '60 aveva affermato che *“computation may someday be organized as a public utility”*[3], la prima forma di servizi Cloud è invece stata realizzata da Salesforce.com che nel '99 aveva reso possibile utilizzare applicazioni tramite web. Ma il passo decisivo è stato realizzato da Amazon, il gigante dell'e-commerce nel 2002 ha reso possibile l'utilizzo di risorse cloud-based tramite gli Amazon Web Services e nel 2006 ha lanciato EC2, servizio che permette il renting di macchine virtuali on-demand. Il 2006 è stato l'anno decisivo, dopo Amazon anche Google e IBM sono entrati nel settore e di recente anche la Microsoft.[4]

Di recente (settembre 2010), anche Telecom Italia ha annunciato il suo ingresso nel Cloud, impostandosi come punto di riferimento nazionale del settore. Con Nuvola Italiana, questo il nome dato alla piattaforma, Oltre a fornire connettività l'azienda mira a vendere servizi supportati dalla banda, ad aziende e Pubblica Amministrazione.[5]

Il Cloud Computing può essere visto come una sorta di rivoluzione, un radicale cambiamento delle modalità di approvvigionamento delle risorse. Ecco perchè il settore negli ultimi anni ha avuto una crescita incredibile, dovuto soprattutto ai grossi investimenti dei grandi nel settore. Oggi si possono contare numerose conferenze sul tema, la più importante è la Cloud Computing Expo, la prima tenuta nel 2007 a New York con 450 delegati, la prossima si terrà a novembre 2010 e conta già più di 5.000 delegati e oltre 1.000 sponsor³; da citare sono anche l'IEEE Cloud che si svolge a marzo a Miami⁴; CloudSlam⁵, una conferenza telematica; conferenze importanti si sono tenute anche in Italia a Milano^{6,7}.

3 <http://cloudcomputingexpo.com/>

4 <http://www.thecloudcomputing.org/2010/>

5 <http://cloudslam10.com/>

6 <http://www.bci-italia.com/confexpo.asp?id=st20102>

7 <http://www.cvent.com/EVENTS/Info/Summary.aspx?e=6c3918fd-9062-4864-8335-f9f25e442014>

1.1 Caratteristiche chiave

La definizione di Cloud Computing è rimasta ancora un po' vaga, cosa è Cloud e cosa no? Si potrebbe arrivare a pensare che qualunque cosa al di fuori di un PC sia Cloud. Sempre stando alla definizione del NIST il modello Cloud ha cinque principali caratteristiche:

Self-service on-demand

L'utilizzatore del servizio può usufruire in ogni momento delle risorse computazionali richieste e senza alcuna necessità di interazione umana.

Ampio accesso via rete

Le risorse sono disponibili in rete e accedute attraverso meccanismi standard che promuovono lo sviluppo e l'utilizzo di client differenti, nonché l'accesso tramite diversi tipi di dispositivi (laptop, palmari, smartphone).

Pooling delle risorse

Tutte le risorse nel cloud sono organizzate e gestite come un pool comune condiviso, con risorse fisiche e virtuali differenti assegnate e riassegnate dinamicamente agli utenti che ne fanno richiesta. Generalmente l'utente non ha controllo e conoscenza sulla locazione delle risorse, al più ha la capacità di specificare la locazione ad un livello di astrazione più alto (es. regione, stato o data center).

Rapida elasticità

La “quantità” delle risorse utilizzate può essere modificata rapidamente e in ogni momento in base alle necessità dell'utente o del sistema, spesso in modo automatico.

Controllo delle risorse

Tutte le risorse sono controllate e monitorate ad uso sia del provider che del consumer.[2]

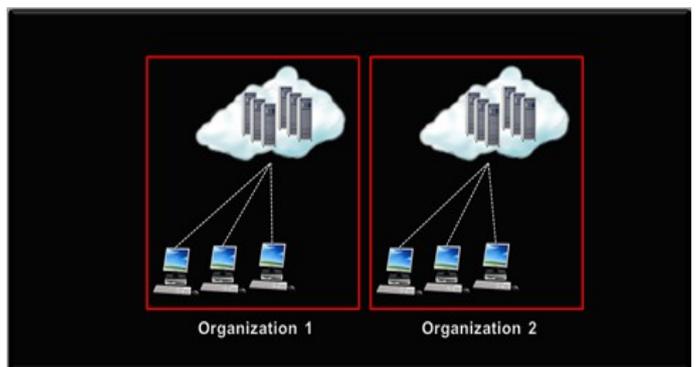


Figura 1: Private Cloud

1.2 Service Models

Si è parlato fin'ora di risorse e di servizi, senza definirli con precisione. Il modello Cloud può essere classificato proprio in base al tipo di risorsa offerta come servizio. Una risorsa può essere un'applicazione, un'intera piattaforma o un'infrastruttura. Un service model (modello di servizio) è identificato in base alla tipologia della risorsa. Ogni modello ha quindi un livello di astrazione differente e si rivolge ad una diversa tipologia di utenti.

1.2.1 Software as a Service (SaaS)

Il servizio offerto all'utente consiste nella capacità di utilizzare applicazioni eseguite in un'infrastruttura cloud. L'accesso all'applicazione viene effettuato tramite programmi client, spesso tramite browser web. L'utente non controlla né amministra l'infrastruttura sottostante (rete, server, sistema operativo, storage) né gestisce direttamente l'applicazione, può tuttavia personalizzarla in alcuni aspetti.[2]

Le caratteristiche principali di questa tipologia sono:

- accesso e amministrazione via rete del software;
- attività controllate da qualsiasi postazione piuttosto che in quella specifica del cliente;
- facilità il lavoro di gruppo, più utenti possono condividere e lavorare sugli stessi file;
- una sola istanza del software è utilizzata da più utenti;
- aggiornamento centralizzato, caratteristica che elimina la necessità per ogni utente di scaricare patch e aggiornamenti;
- tendenza a focalizzarsi sui vantaggi competitivi del software piuttosto che sull'infrastruttura.[3]

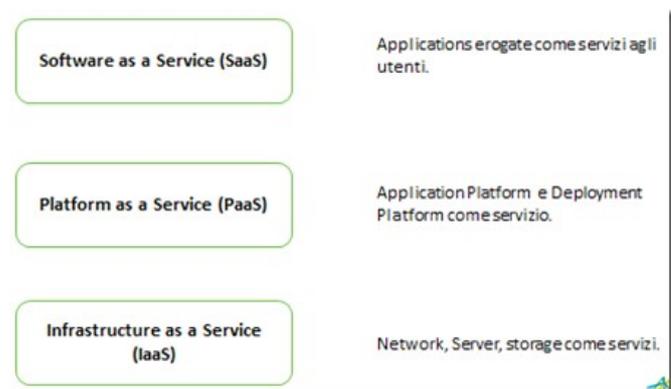
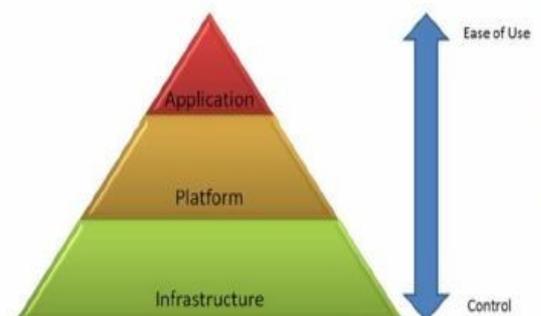


Figura 2: Modelli di servizio

Figura 3: Livelli



SaaS è la tipologia a livello più alto e si rivolge direttamente agli utenti finali dell'applicazione. Un valido esempio è Google Docs, una suite di ufficio utilizzabile completamente via browser.

1.2.2 Platform as a Service (PaaS)

Il servizio offerto all'utente consiste nella possibilità di sviluppare applicazioni e farne il deploy sull'infrastruttura Cloud del fornitore del servizio. Le applicazioni vengono sviluppate utilizzando gli strumenti forniti dal provider (la piattaforma appunto). L'utente non ha il controllo sull'infrastruttura,

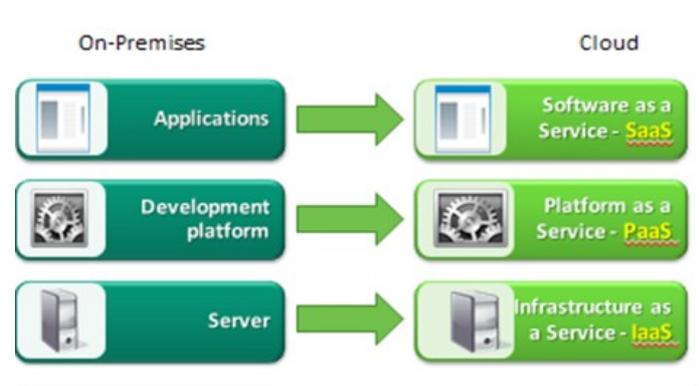


Figura 4: Modelli di servizio

quindi non è in grado di gestire rete, sistema e storage, ma a differenza del SaaS ha il controllo sull'applicazione e sull'ambiente di hosting.[2]

Le caratteristiche principali di questa tipologia sono:

- unico ambiente di sviluppo per la programmazione, testing, deploy, hosting e mantenimento delle applicazioni;
- strumenti per lo sviluppo di interfacce utente web-based;
- possibilità di utilizzo in multiutenza, grazie a gestione della concorrenza, scalabilità e sicurezza;
- integrazione con web services e database.[3]

Rispetto a SaaS il servizio è di un livello più basso, adatto agli sviluppatori. Google ad esempio mette a disposizione App Engine, piattaforma che permette lo sviluppo e il deploy di applicazioni web sulla propria infrastruttura Cloud.

1.2.3 Infrastructure as a Service (IaaS)

Il servizio offerto all'utente consiste nella possibilità di usufruire di capacità di calcolo, network e storage e di altre risorse informatiche di basso livello. L'utente è in grado di usufruire di sistema operativo e software arbitrario, non ha il controllo sull'infrastruttura ma può gestire direttamente rete, sistema, storage e applicazioni.[2] Questa tipologia è rivolta soprattutto ai sistemisti.

Tra i possibili utilizzi di un IaaS vi sono:

Testing

possono essere facilmente creati ambienti di testing e staging su misura, utilizzati per il tempo richiesto e rilasciati a lavoro compiuto. Scompaiono i tempi di attesa per l'acquisizione di hardware e la configurazione degli ambienti, quando necessario è sufficiente creare un nuovo ambiente e cancellarlo una volta terminati i test.

Deployment di applicazioni web

le applicazioni web possono risiedere in infrastrutture cloud, con la possibilità di incrementare le risorse erogate in base al traffico.

Storage

fornire storage remoto.

Calcolo ad alte prestazioni

quando necessario creare un ambiente personalizzato per l'elaborazione di grosse quantità di dati, cluster on-demand da utilizzare per determinati periodi di tempo.[6]

Amazon EC2 permette ad esempio il renting di server virtuali utilizzabili dagli utenti per creare la propria infrastruttura. IaaS è anche il modello di servizio dell'offerta Cloud Telecom Italia⁸.

1.3 Deployment Models

I Deployment Models (modelli di rilascio) indicano le modalità in cui vengono forniti i servizi, e in che modo essi arrivano ai fruitori. Il modello più comune è il Public Cloud, ovvero i servizi vengono offerti al pubblico da un provider. Gli altri casi possibili sono il Private Cloud, realizzato quando un'azienda decide di dotarsi di una propria infrastruttura e i servizi vengono utilizzati solo all'interno di essa; Community Cloud, realizzato quando più aziende con finalità simili utilizzano una stessa infrastruttura; infine Hybrid Cloud, realizzato utilizzando un'infrastruttura ibrida ottenuta combinando due o più modelli (public, private o community) che rimangono entità distinte ma sono utilizzati assieme grazie a tecnologie standard che abilitano la portabilità di dati e applicazioni (Ad esempio un'azienda che possiede un'infrastruttura cloud privata può combinarla con Public Cloud per garantire maggiore potenza di calcolo e scalabilità quando necessario).[2][7]

8 www.nuvolaitaliana.it/

1.3.1 Public Cloud

È il modello più diffuso.

Il servizio viene distribuito da provider tramite internet in forma pubblica. I servizi Public Cloud possono essere offerti gratuitamente o in forma pay-per-use.

Questa tipologia è una forma di Utility Computing, ovvero la possibilità di usufruire di risorse computazionali in modo simile alle risorse pubbliche (elettricità, acqua, gas, ecc)[3], ovvero in modo trasparente (per utilizzare l'energia elettrica non è necessario conoscerne la rete, similmente per utilizzare le risorse computazionali non è necessario conoscere l'infrastruttura), a consumo, e con minima o nessuna competenza tecnica (per utilizzare l'energia elettrica non è necessario essere degli elettricisti, similmente per utilizzare le risorse computazionali non sono necessarie le competenze tecniche per la configurazione dell'hardware).[3]

Il modello Public è particolarmente adatto per aziende e privati che necessitano di risorse per periodi di tempo limitati, o di potenza di calcolo scalabile. Utilizzando risorse Cloud si hanno abbattimento dei costi di acquisizione hardware, di reperimento personale e significativa riduzione dei tempi.

A livello realizzativo erogare servizi Public richiede enormi investimenti e soltanto le grosse aziende del campo informatico come Amazon, Microsoft e Google sono entrate nel campo. Il Public Cloud è implementato su migliaia di server distribuiti in centinaia di data center ubicati in decine di locazioni sparse in tutto il mondo.[6]

1.3.1.1 Critiche

Il modello Public è molto diffuso ma anche molto criticato, le principali critiche riguardano:

Riservatezza

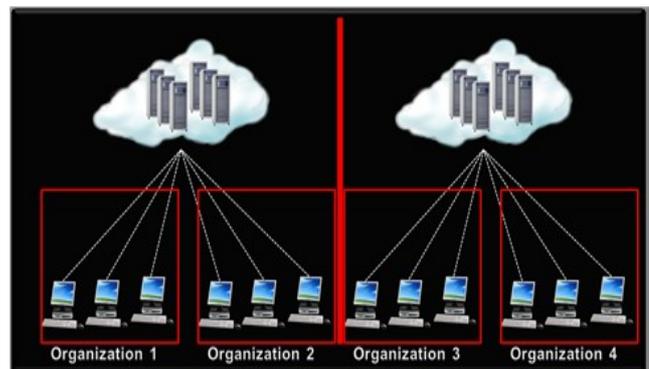


Figura 5: Community Cloud

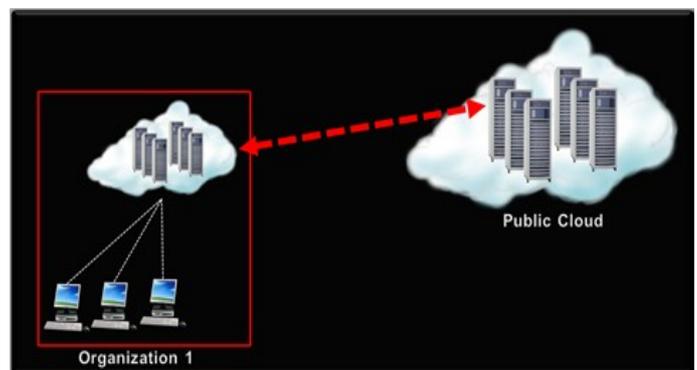


Figura 6: Hybrid Cloud

I dati personali vengono affidati a terzi, essi risultano in possesso dell'azienda che quindi se avesse un comportamento malevolo potrebbe accedere ai dati degli utenti al fine di eseguire indagini di mercato o di profilazione dell'utente.

Continuità del servizio

Delegando a un servizio esterno la gestione dei dati e la loro elaborazione l'utente si trova fortemente limitato nel caso i suddetti servizi non siano operativi. Un eventuale malfunzionamento inoltre colpirebbe un numero molto elevato di persone contemporaneamente dato che questi sono servizi condivisi.

Dipendenza da un singolo fornitore

Essendo i dati in possesso di un determinato provider potrebbe essere complicato o addirittura impossibile migrare ad un altro fornitore, a causa dei formati chiusi utilizzati o dell'impossibilità di fare backup.

Cambiamento improvviso del servizio

Nel caso in cui l'utente utilizzi un'applicazione sul proprio pc, può decidere se aggiornare il software o meno, o se installare un'applicazione con le medesime funzionalità. Per i servizi Cloud ciò non è valido, l'utente potrebbe trovarsi improvvisamente ad utilizzare un servizio o un'interfaccia diversa.

Cambiamento dei termini di servizio

Può capitare che un servizio gratuito diventi a pagamento, o che venga effettuato un aumento dei prezzi, in questo caso l'utente potrebbe trovarsi in condizioni spiacevoli, specie se ha assoluto bisogno del servizio.

Chiusura del codice

nella maggior parte dei casi ci si trova ad utilizzare codice proprietario e formati chiusi.[3]

1.3.1.2 Cloud e sicurezza

La sicurezza informatica si basa sulla fiducia, fiducia verso il produttore dell'hardware, verso il produttore del software, verso l'ISP. Con il paradigma cloud bisogna fidarsi anche del fornitore del servizio. Tuttavia c'è una certa differenza, se si ha accesso fisico alla macchina, si ha maggiore controllo sulla sicurezza, configurando firewall e controllando direttamente il traffico, la fiducia verso i produttori non è quindi del tutto necessaria. Se la risorsa invece è fornita da qualcun altro dalla nuvola ciò non è affatto vero, bisogna fidarsi completamente del fornitore del servizio.[8]

All'utente devono essere quindi garantiti:

Protezione dei dati

I dati devono essere conservati in modo sicuro, nessun altro deve avere la possibilità di accedervi, non devono esserci perdite di dati e il trasferimento da una locazione ad un'altra dev'essere altrettanto sicuro.

Controllo dell'identità

Devono essere forniti mezzi di autenticazione sicuri.

Sicurezza fisica

L'accesso alle macchine fisiche dev'essere ristretto e controllato, nonché l'accesso ai dati degli utenti.

Disponibilità

Dev'essere garantito l'accesso ai dati e alle applicazioni in ogni momento.

Privacy

I dati sensibili (come numeri di carte di credito) devono essere mascherati e l'accesso ad essi dev'essere riservato solo agli autorizzati, l'identità digitale e le credenziali nonché le attività degli utenti devono essere altrettanto protette.

Continuità del servizio e recupero dei dati

Il servizio e l'accesso ai dati devono essere garantiti anche a seguito di problemi tecnici.

Recupero dei dati in caso di termine di servizio

Qualora il servizio venisse terminato dev'essere garantito il recupero dei dati anche dopo tale evento.

mantenimento dei log

I log devono essere mantenuti e devono rimanere accessibili anche dopo lunghi periodi, sia per il controllo da parte degli utenti che per eventuali investigazioni giudiziarie.[3][9]

1.3.2 Private Cloud

È il caso in cui un'azienda si munisce di un'infrastruttura Cloud per erogare servizi all'interno, utile alle organizzazioni che necessitano o preferiscono avere maggiore controllo sui propri dati e modalità di servizio personalizzate.

Inoltre un modello Private Cloud permette all'azienda di centralizzare il controllo sui propri server, implementando i servizi in un unico data center. Ciò richiede un numero di personale ridotto, minore dispendio di energia e attrezzatura di raffreddamento.

Il Private Cloud diventa vantaggioso per un'azienda quando questa dispone di un elevato numero di server (da qualche decina in su) che erogano diversi servizi, magari su macchine

ridondate per garantire efficienza e continuità del servizio. Ma una infrastruttura di questo tipo genera molti sprechi se c'è disparità di utilizzo dei vari servizi o discontinuità di utilizzo nel tempo. La virtualizzazione di molte macchine su un numero più limitato di server permette di ottimizzare le risorse, riducendo gli sprechi. Il paradigma del Cloud Computing consente di spostare dinamicamente le risorse laddove è necessario, con una “elasticità” che sui grandi numeri paga in termini di efficienza, riduzione dei costi ed anche dei consumi energetici.

1.4 Cloud e Software Libero

Il concetto di software libero discende naturalmente da quello di libertà di scambio di idee e di informazioni. Negli ambienti scientifici, quest'ultimo principio è tenuto in alta considerazione per la fecondità che ha dimostrato; ad esso infatti è generalmente attribuita molta parte dell'eccezionale ed imprevedibile crescita del sapere negli ultimi tre secoli.

La libertà di scambio di idee non è tuttavia una questione puramente pratica: essa è anche alla base dei concetti di libertà di pensiero e di espressione. Analogamente alle idee, il software è immateriale, e può essere riprodotto e trasmesso facilmente. In modo simile a quanto avviene per le idee, parte essenziale del processo che sostiene la crescita e l'evoluzione del software è la sua libera diffusione.[10]

Il concetto di Software Libero è stato formalizzato per la prima volta negli anni '80 da Richard Stallman, allora ricercatore al MIT e fondatore della Free Software Foundation, per difendere il software dal crescente mercato dei programmi proprietari. Secondo la sua definizione un software (per essere considerato libero) deve garantire quattro libertà a chi lo utilizza⁹:

Libertà 0 – libertà di utilizzare il software per qualsiasi scopo;

Libertà 1 – libertà di studiarne il funzionamento e riadattarlo alle proprie esigenze; poter accedere e modificare il codice sorgente è una precondizione per questo;

Libertà 2 – Libertà di ridistribuire copie in modo da aiutare il prossimo;

Libertà 3 – Libertà di migliorare il software e distribuirne i miglioramenti, in modo che tutta la comunità ne tragga beneficio.

Il software è tutelato da licenze di copyright, le quali appoggiandosi alle leggi sul diritto d'autore specificano diritti e doveri sui programmi. Mentre il software proprietario è distribuito

9 <http://www.gnu.org/philosophy/free-sw.it.html>

con licenze che ne limitano l'uso, la distribuzione e la condivisione; il software libero è rilasciato con licenze che ne tutelano le quattro libertà. Prima e più importante fra queste licenze è la General Public License (GPL)¹⁰, la quale oltre a garantire la libertà del software, lo protegge da abusi, infatti non è possibile modificare un software rilasciato sotto GPL e ridistribuirlo con una licenza diversa.

Nel campo Cloud la maggior parte dei software utilizzati dai grandi provider e le stesse applicazioni usate come SaaS sono proprietarie. Stallman ha infatti definito il Cloud Computing come una trappola per costringere gli utenti ad utilizzare sempre di più sistemi chiusi e proprietari che in futuro avranno costi sempre maggiori.[11] Il primo software pensato per il Cloud Computing ad essere rilasciato sotto GPL è stato Eucalyptus nel 2008. Da allora tuttavia nel campo stanno emergendo sempre più software open e organizzazioni che lo promuovono, come Free Cloud Alliance¹¹, un'organizzazione no-profit che propone una serie di soluzioni cloud open. Anche la NASA si è immersa nel campo Cloud sviluppando una piattaforma open, Nebula¹², basata proprio su Eucalyptus. Di questo e in generale di open cloud si è parlato anche in una conferenza tenuta a giugno 2010 a Roma[12]. Al panorama Free Cloud non poteva inoltre mancare una delle principali aziende portavoce del software libero, Red Hat, che ha recentemente reso disponibile la propria soluzione¹³, che fa della portabilità e dell'interoperabilità la sua caratteristica principale, nel pieno rispetto dell'idea di un Cloud Open.

Azienda come Sun Microsystem, IBM e la stessa Red Hat hanno sottoscritto l'Open Cloud Manifesto¹⁴, un documento pubblico su principi e intenzioni dei Cloud provider con la convinzione che “la nuvola dovrebbe essere aperta”; manifesto che comunque non è stato sottoscritto da Microsoft e Amazon¹⁵.

Il Cloud fa sorgere tuttavia problemi sulle licenze. La GPL ad esempio nel campo Cloud ha una grossa limitazione, essa impone che il sorgente dev'essere distribuito insieme al software, ma le applicazioni cloud non vengono “distribuite”. Pertanto chi offre la possibilità di eseguire programmi via web può usufruire di codice GPL senza essere costretto a pubblicare le modifiche. Per impedire tale abuso la Free Software Foundation ha rilasciato la Affero General Public License (AGPL). I software rilasciati tramite AGPL devono mettere a disposizione il sorgente non solo per chi installa il programma, ma anche per chi lo esegue via web.[13]

10 <http://www.gnu.org/licenses/gpl.html>

11 <http://www.freecloudalliance.org/>

12 <http://nebula.nasa.gov>

13 <http://www.europe.redhat.com/solutions/cloud/>

14 <http://opencloudmanifesto.org>

15 http://news.cnet.com/8301-13860_3-10206077-56.html

1.5 Cloud e Virtualizzazione

Per virtualizzazione, o più precisamente per hardware virtualization, si intende la simulazione a livello software di un calcolatore, una macchina virtuale, la quale può eseguire a sua volta software arbitrario. Il software che rende possibile la simulazione viene chiamato virtual machine monitor o hypervisor. Esistono due principali metodi di virtualizzazione:

- Emulazione: l'hypervisor simula completamente l'hardware della macchina virtuale. Il software può essere eseguito su di essa come fosse una macchina reale;
- Paravirtualizzazione: la macchina virtuale non simula l'hardware ma offre speciali API che richiedono modifiche nel sistema operativo virtualizzato.

La virtualizzazione permette un rapido setup di sistemi operativi e ambienti senza preoccuparsi dell'hardware; una macchina virtuale (in quanto software) può essere facilmente “portata” da un computer ad un altro; più macchine virtuali possono essere eseguite su un singolo calcolatore dedicandone ognuna ad un compito specifico; il “crash” di una macchina virtuale non compromette il sistema ospitante e una macchina virtuale crashata può essere sostituita in brevissimo tempo, a differenza di una macchina reale. Visti i possibili benefici e la flessibilità offerta appare chiaro come la virtualizzazione sia la tecnologia cardine del Cloud Computing, tuttavia le applicazioni eseguite su ambiente virtuale non hanno diretto accesso alla macchina reale ma devono bensì “passare” per uno strato software, un ambiente virtualizzato è perciò in genere poco performante. Il calo di performance è in genere contenuto se si tratta di paravirtualizzazione in quanto più che simulare l'hardware l'hypervisor tende in questo caso a controllare l'accesso alle risorse reali, in questo modo il software ospitato ha quasi diretto accesso alla CPU, il prezzo da pagare è la necessaria modifica del codice del sistema operativo virtualizzato, operazione non sempre possibile (se il sistema in questione è proprietario ad esempio)[3]. I principali hypervisor sono KVM, VMware e Xen, quest'ultimo fa della paravirtualizzazione il suo punto forte ed è il software utilizzato da Amazon EC2.

1.6 Il Cloud Computing per la gente comune

Il termine Cloud Computing è ormai conosciuto anche dai non esperti di informatica. Il punto di vista della gente comune può essere espresso dalla definizione che ne ha dato nel 2006 il

direttore dell'azienda che è quasi diventata sinonimo di internet, così si pronunciò il CEO di Google Eric Schmidt:

it starts with the premise that the data services and architecture should be on servers. We call it cloud computing – they should be in a ‘cloud’ somewhere. And that if you have the right kind of browser or the right kind of access, it doesn't matter whether you have a PC or a Mac or a mobile phone or a BlackBerry or what have you – or new devices still to be developed – you can get access to the cloud...[14]

Lo scenario è quindi quello di un utente il quale, avendo un device (non importa quale: un PC, un palmare, uno smartphone, oppure chissà quale futuro diabolico apparecchio), un browser ed una connessione ad Internet può accedere alla nuvola giusta che gli fornisce i servizi e/o i dati che gli sono necessari. Spesso poi questi servizi saranno ‘composti’ a piacimento dall’utente nel contesto delle sue necessità. In questo modo egli può creare uno strumento ‘personalizzato’ fatto di un mosaico di funzionalità derivanti dalla somma di singoli ‘servizi web’.[15]

Il successo del Cloud Computing, e in particolare del modello SaaS, è in sintesi dovuto la fatto di portare quelle che erano le attività svolte al PC nella rete, “in the cloud”; e agevolare la diffusione dei dispositivi ultra mobile, i quali non hanno più necessità di eseguire applicazioni e memorizzare dati, devono solo garantire l'accesso alla nuvola. In ogni luogo posso scrivere i miei documenti, leggere le mie email, guardare i miei appuntamenti, chattare con i miei contatti; tutto quello di cui ho bisogno è un browser web, sempre Schmidt afferma che “il browser è il computer”¹⁶ (su questa idea Google ha sviluppato Chrome OS, il sistema operativo pensato per il Cloud).

1.6.1 Uso improprio del termine

La diffusione del termine ha però portato ad un suo abuso, la possibilità di accedere via browser ad applicazioni e dati ha creato non poca confusione: tra le attività che svolgo in rete quali possono essere definite cloud e quali no? Il termine è infatti utilizzato in senso più generico per indicare una forma di elaborazione dei dati via rete, e la nuvola è intesa come un qualcosa di incognito e incontrollabile, che può costituire un notevole rischio per la privacy. Stando a questa definizione la grande maggioranza dei servizi web 2.0 possono essere definiti cloud, quali i social network, il portale YouTube e affini. A livello tecnico invece un social network ad esempio non costituisce una forma di Cloud Computing in quanto si tratta di un'applicazione web eseguita su server reali e che quindi non presenta le caratteristiche chiave del paradigma (ad esempio

¹⁶ http://news.cnet.com/8301-17939_109-10250196-2.html

applicazioni web di questo tipo non sono scalabili, i server erogano sempre la stessa potenza indipendentemente dalle necessità del software).

1.7 Il Cloud Computing per le aziende

L'opportunità che offre il Cloud Computing, la capacità di usufruire di tutte le risorse desiderate in qualsiasi momento è un po' il sogno di ogni informatico. Il vantaggio in termini economici e in termini temporali che ciò può portare ad un'azienda è lampante ed è costituita principalmente da tre fattori:

- l'illusione di poter usufruire di risorse infinite on-demand eliminando la necessità di pianificarne il reperimento;
- l'eliminazione della necessità di reperimento hardware (o più in generale del servizio adatto) consentendo alle piccole compagnie di cominciare con ridotte quantità di risorse per poi incrementarle in base alle necessità;
- la possibilità di pagare solo per le risorse utilizzate.[16]

Il risparmio conseguente è stimato tra il 25 e il 60%¹⁷.

17 http://www.bci-italia.com/readnews_2010.asp?id=2010/18002

2 Piattaforme di Cloud Computing

Per piattaforma di Cloud Computing si intende l'insieme di software e tecnologie che abilitano il delivering di risorse nella forma Cloud (on-demand, scalabili, virtualizzate). Per comprendere meglio cosa vuol dire utilizzare servizi di Cloud Computing è necessario dare uno sguardo a questi framework, sia alle tecnologie e offerte dei public providers, utilizzati su larga scala; che ai free software che ne consentono libero studio, ricerca, sperimentazione e implementazioni private; cercando di avere una visione globale dei vari service models e delivery models.

Tra le più popolari piattaforme vi sono quelle di Google e, naturalmente, la pietra miliare del Cloud Computing, gli Amazon Web Services.

Come software libero è necessario citare prima di tutti Eucalyptus, il primo e più utilizzato. Ma progetti importanti sono anche OpenNebula ed EyeOS.

2.1 Google Docs



- Modello di servizio: SaaS
- Modello di rilascio: Public Cloud
- Sito web: docs.google.com

È quasi sinonimo di Cloud Computing, con Google Docs quelli che erano forse i programmi più utilizzati in ambito desktop, vale a dire i software di ufficio, sono migrati nella nuvola eliminando il problema di dover installare la suite sui PC e di inviare i documenti nelle maniere tradizionali. Google Docs permette di lavorare ai documenti in qualsiasi posto, tramite browser, ma soprattutto facilita la condivisione e il lavoro di gruppo. Per accedere ai propri documenti è sufficiente un Google Account, una volta effettuato l'accesso è possibile caricare documenti o crearne di nuovi, visualizzare e modificare i documenti esistenti, condividerli con gli account dei collaboratori, in sola lettura o in lettura scrittura. L'editing dei documenti è concesso a più utenti simultaneamente, e per l'import/export sono supportati i principali formati (es. odt, pdf, doc), più

in generale è possibile caricare qualunque file e distribuirli in cartelle utilizzando Docs come un filesystem remoto.

Google fornisce API in Java e Python che consentono di eseguire la maggior parte delle azioni (upload/download di file, cambio permessi, ecc.).

La conservazione dei dati non in locale pone però seri problemi di privacy per le aziende e per i singoli, sia per l'utilizzo delle informazioni a scopo di schedature che potrebbe essere fatto da chi gestisce il servizio, sia per il maggiore rischio di attacchi e manipolazioni da parte di soggetti esterni, che si verifica quando i dati risiedono su server sempre connessi a Internet.[3]

2.2 Google App Engine



- Modello di servizio: PaaS
- Modello di rilascio: Public Cloud
- Sito web: appengine.google.com

Google App Engine è un ottimo esempio di PaaS: un ambiente che permette lo sviluppo e il deploy di un'applicazione web sull'infrastruttura Cloud di Google. Sono garantite tutte le caratteristiche chiave del paradigma Cloud, in particolare la scalabilità: l'allocazione delle risorse che l'applicazione utilizzerà varia dinamicamente in base alle necessità del software.

I linguaggi di programmazione supportati sono Java e Python, per entrambi è possibile scaricare l'apposito ambiente di sviluppo.

Il servizio di base, con determinati limiti di risorse (500 MB di storage e 5 milioni di visualizzazioni al mese per pagina), è gratuito. È possibile richiedere le risorse necessarie con il dovuto pagamento, nella forma pay-for-use.

L'applicazione viene eseguita in una sandbox, ovvero un ambiente di esecuzione virtualizzato che mantiene un alto livello di sicurezza e limita l'accesso al sistema operativo sottostante. Ciò permette ad App Engine di distribuire le richieste a diversi server, e avviare e arrestare le macchine in base al traffico. L'ambiente impone però alcune limitazioni:

- l'applicazione può accedere a computer in rete solo attraverso URL o email. L'accesso dall'esterno è consentito solo tramite HTTP (o HTTPS) nelle porte standard.

- L'applicazione non può scrivere direttamente sul filesystem, è necessario l'uso del datastore fornito dalla piattaforma.
- Il codice può essere eseguito solo in risposta ad una web request, la quale deve dare risposta entro 30 secondi. Non è possibile creare sotto processi o eseguire codice dopo la risposta.

In sintesi le caratteristiche principali della piattaforma sono:

- Supporto per le più comuni tecnologie web.
- Storage persistente, con supporto per query, ordinamento e transazioni.
- Scaling e load balancing automatici.
- API per l'autenticazione e l'invio di email tramite Google Accounts.
- L'SDK simula in locale la piattaforma.
- Si possono creare task da mettere in coda (queued tasks) per eseguire codice all'infuori di una web request.
- Scheduling di operazioni in momenti specifici e ad intervalli regolari (scheduled tasks).

2.2.1 Java Runtime Environment

Tramite l'ambiente fornito è possibile utilizzare i normali strumenti di Java per il Web Development e le API standard (Servlet, JSP, ecc.). Le restrizioni della sandbox sono implementate nella JVM, ad esempio se l'applicazione tenta di aprire un socket o scrivere su file verrà lanciata un'eccezione.

L'applicazione accede alla maggior parte dei servizi di App Engine tramite le API di Java. Per l'accesso al database l'SDK include apposite implementazioni di Java Data Objects (JDO) e Java Persistence API. Per l'invio di email è possibile utilizzare JavaMail API

2.2.2 Datastore

Il datastore fornito da App Engine non è un database relazionale. I dati sono organizzati in oggetti (entità), ogni entità ha un tipo e un insieme di attributi. Le query possono cercare i dati in base al tipo di entità e ai valori degli attributi. Le interfacce Java e Python si adattano perfettamente a tale implementazione.

2.2.3 Sviluppo

Gli SDK forniti includono un web server capace di emulare tutti i servizi di App Engine in locale sul proprio computer. Ogni SDK include anche tutte le API e librerie disponibili. Il web server emula anche il sandbox environment, nonché un tool per eseguire l'upload dell'applicazione. Una volta creato codice, file statici e file di configurazione è sufficiente caricare il software tramite l'apposito strumento, il quale richiederà le credenziali di accesso del Google Account dell'utente.

Quando viene rilasciata una nuova versione di un software già presente in App Engine è possibile registrarla come nuova release, gli utenti continueranno ad utilizzare la vecchia versione finché lo sviluppatore non decide di sostituirla con la nuova. In questo modo la nuova release può essere testata mentre la versione precedente è ancora in esecuzione.

Per l'amministrazione del software è anche possibile utilizzare una console di amministrazione web-based, la quale può essere utilizzata per eseguire l'upload di nuove applicazioni, configurare domini, cambiare versione di un'applicazione, esaminarne accessi, datastore e log di errori. Inoltre è possibile invitare altre persone a partecipare allo sviluppo, consentendone l'accesso alla console e la possibilità di caricare nuove versioni.

2.3 Amazon Web Services



- Modello di servizio: IaaS
- Modello di rilascio: Public Cloud
- Sito web: <http://aws.amazon.com>

Gli Amazon Web Services (AWS) sono una collezione di servizi web interoperanti offerti dal noto provider che assieme offrono un servizio Cloud IaaS.

AWS è di gran lunga la piattaforma Cloud più utilizzata e popolare, tanto da essere un vero e proprio esempio nel settore, a cui molti altri software e piattaforme si sono rifatti.

Con AWS si può avere a disposizione, in sintesi, potenza di elaborazione, memoria di massa per archiviazione, servizi di database, attraverso una infrastruttura IT “elastica”, variabile in base alle esigenze di business, e in forma pay-per-use.[17]

2.3.1 Amazon Elastic Compute Cloud (EC2)

Amazon EC2 è il servizio fondamentale degli AWS. Si basa su un semplice concetto fondamentale: richiedere e utilizzare macchine virtuali (VM) su cui si ha pieno controllo.

Per utilizzare EC2 è sufficiente scegliere l'immagine delle macchine virtuali da avviare. L'immagine di una VM nel gergo di EC2 è denominata AMI (Amazon Machine Image). L'utente ha a disposizione una serie di AMI preconfigurate (le cosiddette Public AMI) in grado di soddisfare la gran parte delle necessità, I sistemi operativi disponibili sono GNU/Linux, OpenSolaris e Windows Server 2003. Eventualmente è possibile personalizzare le AMI o crearne di nuove.

La potenza di calcolo delle macchine è completamente personalizzabile, e quantificata in "EC2 Compute Units". Una EC2 Compute Unit equivale a una CPU virtuale operante alla frequenza di 1.0-1.2 GHz, le tariffe si basano sul tempo di utilizzo (in ore) e sulle Compute Units utilizzate, in questo modo utilizzare una Compute Unit per 100 ore è equivalente a utilizzarne 100 per un'ora.

Una VM avviata è denominata istanza, le istanze si possono avviare e terminare in pochi minuti in base alle necessità garantendo massima scalabilità, da qui il termine "elastic".

All'avvio di un'istanza ad essa vengono assegnati un indirizzo IP privato, utilizzabile esclusivamente per l'accesso da altre istanze, e un indirizzo IP pubblico per l'accesso dall'esterno. Insieme agli indirizzi vengono assegnati anche i rispettivi nomi DNS.

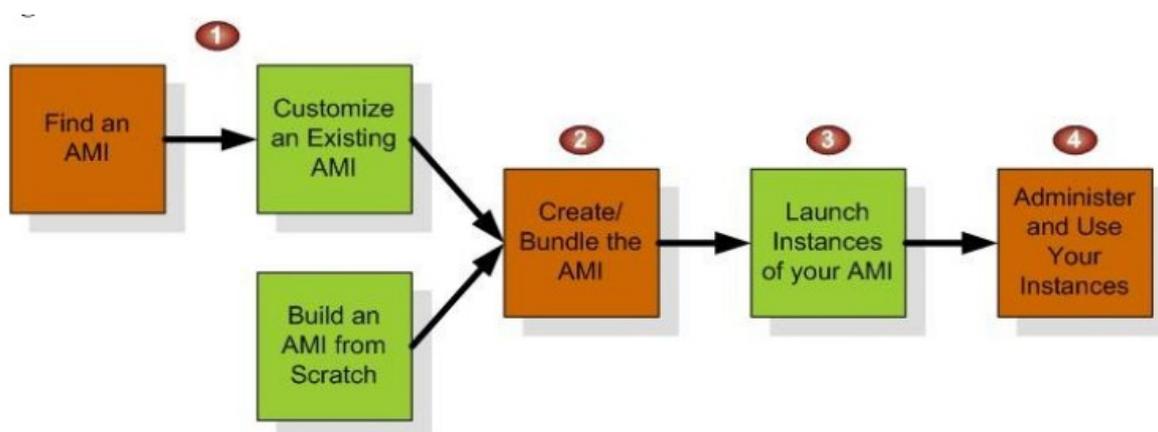


Figura 7: Utilizzo di EC2

2.3.2 Multiple Locations

Amazon EC2 prevede la possibilità di collocare le varie istanze in più sedi. Le locazioni di Amazon EC2 sono composte da "regioni" e da "availability zone". Queste ultime sono luoghi

ben distinti progettati per essere potenzialmente immuni da possibili guasti, sono poco costose, hanno basso tempo di latenza di connettività di rete e sono sempre collegate con le rimanenti availability zone presenti nella stessa Regione.

Con il lancio di istanze in availability zone distinte, si possono proteggere le applicazioni dai potenziali guasti avvenuti in una singola locazione. Le regioni consistono in una o più availability zone (per esempio la regione che si occupa dell'Europa è formata solo da due availability zone); sono distribuite sul territorio e sono in zone geografiche distinte o addirittura in stati differenti. EC2 è attualmente disponibile in due sole regioni: uno negli Stati Uniti e uno in Europa.

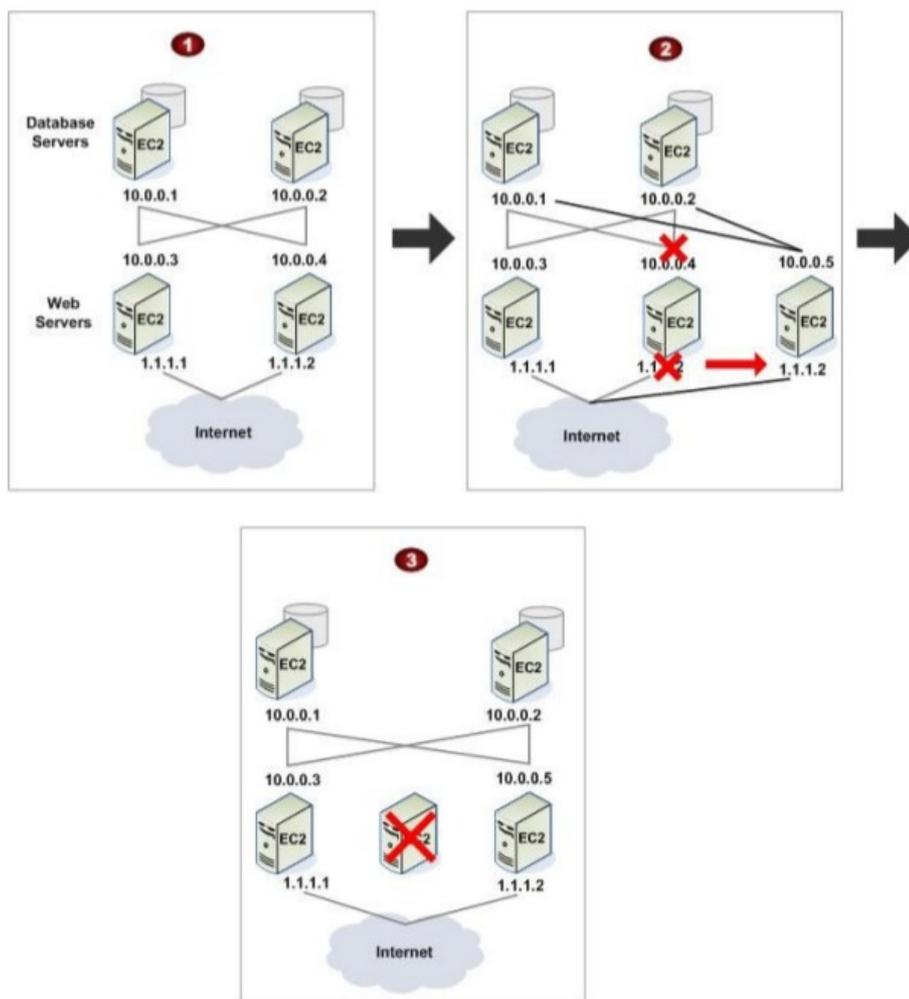


Figura 8: Elastic IP Addressing

2.3.3 Elastic IP Addresses

Gli Elastic Ip sono una funzionalità aggiuntiva di EC2, sono progettati per fornire indirizzi IP dinamici alle istanze. Ogni indirizzo IP elastico è associato con l'account dell'utente e non alla singola istanza, e l'utente continua a controllare tale indirizzo fino a quando non si scelga esplicitamente di rilasciarlo. A differenza dei tradizionali indirizzi IP statici, gli indirizzi IP elastici consentono di proteggere le istanze dai guasti delle availability zone attraverso una rimappatura sistematica degli indirizzi IP dell'utente e di tutte le istanze del suo account. Piuttosto che attendere un supporto tecnico (dati programmi etc) che riconfiguri o sostituisca l'host, o in attesa della trasmissione del DNS a tutti i suoi client, Amazon EC2 consente di risolvere rapidamente i problemi con le proprie istanze o con particolari software mediante l'invocazione della rimappatura dei propri indirizzi IP elastici in modo da ottenerne dei nuovi da sostituire a quelli temporaneamente non funzionanti.

2.3.4 Security Groups

EC2 permette di assegnare le istanze a gruppi definiti da'utente per facilitare la configurazione dei firewall. Il firewall viene impostato sul gruppo, le istanze ereditano quindi le impostazioni del Security Group di cui fanno parte.

2.3.5 Amazon Elastic Block Store (EBS)

EC2 non offre direttamente lo storage, o meglio, le VM sono provviste di disco virtuale e i dati vengono preservati anche dopo un reboot, ma quando un'istanza viene terminata il disco virtuale viene eliminato. Per utilizzare lo storage con EC2 occorre usufruire di EBS.

EBS fornisce volumi, ad esempio, di memorizzazione "off-instance" che persistono indipendentemente dalla vita di un'istanza. Gli Amazon EBS sono volumi di stoccaggio facilmente reperibili, altamente affidabili che possono essere facilmente connessi ad istanze di Amazon EC2 e sono rilevati come dispositivi standard a blocchi. I volumi EBS offrono notevoli miglioramenti in termini di durevolezza, indipendentemente da dove le istanze di Amazon EC2 siano memorizzate; infatti, i volumi di Amazon EBS sono automaticamente replicati tramite "backend" (in una unica availability zone).

Per quelli che vogliono ancora più affidabilità nello stoccaggio, Amazon EBS prevede la possibilità di creare "point-in-time", "istantanee" dei volumi selezionati che vengono poi immagazzinate in Amazon S3, e automaticamente replicate su più zone disponibili (multi

availability zone). Queste istantanee possono essere utilizzate come backup o come punto di partenza per i nuovi volumi Amazon EBS.

I volumi possono essere utilizzati anche come partizioni di avvio per le istanze di EC2, ciò permette di preservare i dati della partizione di boot e creare AMI personalizzate.

2.3.6 Amazon Simple Storage Service (S3)

Amazon S3 è un servizio web di storage, pensato per offrire piena scalabilità, alta disponibilità e bassa latenza. S3 memorizza oggetti arbitrari con dimensione massima di 5 GB più 2 KB di metadati. Essi sono organizzati in quelli che vengono definiti bucket, ogni bucket è identificato da una chiave univoca definita dall'utente. Oltre che tramite le apposite API e interfaccia web gli oggetti possono essere scaricati anche tramite BitTorrent.

In generale le operazioni che è possibile fare con S3 sono poche e semplici:

- Creare bucket;
- Memorizzare oggetti in un bucket;
- Scaricare oggetti da un bucket;
- Modificare i permessi di accesso ad un bucket.

S3 può essere utilizzato anche per memorizzare immagini di macchine virtuali da usare con EC2. Un'immagine può essere creata da qualsiasi disco (reale o virtuale) in possesso dell'utente, caricata in un bucket, e registrata in modo da essere utilizzabile su EC2.

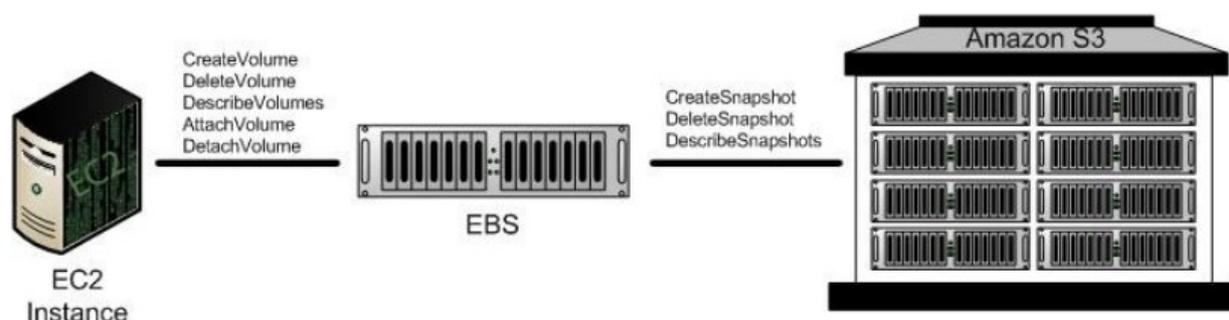


Figura 9: EBS API Overview

2.3.7 Amazon CloudWatch

Amazon CloudWatch è il web service che provvede al monitoring delle risorse degli AWS utilizzate. Oltre al monitoring il servizio abilita l'Auto Scaling, ovvero la possibilità di cambiare automaticamente il numero di istanze in esecuzione.

2.3.8 Elastic Load Balancing

L'Elastic Load Balancing è una funzionalità di EC2 che permette di distribuire il traffico tra le istanze in esecuzione, redirezionandolo alle istanze più scariche. Elastic Load Balancing può essere abilitato in una o anche più Availability Zone e l'attività di un Balancer può essere monitorata da CloudWatch.

2.3.9 Amazon Virtual Private Cloud (VPC)

VPC è un servizio che consente di collegare una infrastruttura IT esistente ad una nuvola di risorse e servizi AWS. Ciò avviene tramite una VPN (Virtual Private Network) e consente di estendere le caratteristiche dell'infrastruttura IT esistente (es. sicurezza, firewall, intrusion detection, ecc.) alla nuvola di risorse AWS, che pertanto viene isolata dal rimanente mondo AWS.

2.3.10 High Performance Computing (HPC) Clusters

Per gli utenti che necessitano di prestazioni elevate per calcolo ad alte prestazioni Amazon ha di recente predisposto istanze speciali da 33.5 Compute Units, che per di più hanno parziale accesso all'hardware sottostante permettendo una certa ottimizzazione prestazionale delle applicazioni. Tali istanze sono inoltre collocabili in cluster virtuali, placement group nel gergo di Amazon. Le istanze facenti parte del gruppo godono di una connessione a bassa latenza (10 Gbps).

2.3.11 Lavorare con gli AWS

Per permettere a tutti facile accesso ai servizi è possibile utilizzabile AWS Management Console, una comoda interfaccia web che permette l'accesso e l'utilizzo di tutte le funzionalità della piattaforma. Per avviare un'istanza di EC2 ad esempio, l'utente non

deve fare altro che seguire una procedura guidata che comprende la selezione della AMI da una lista dettagliata, la generazione di una coppia di chiavi per l'accesso SSH alla macchina virtuale,

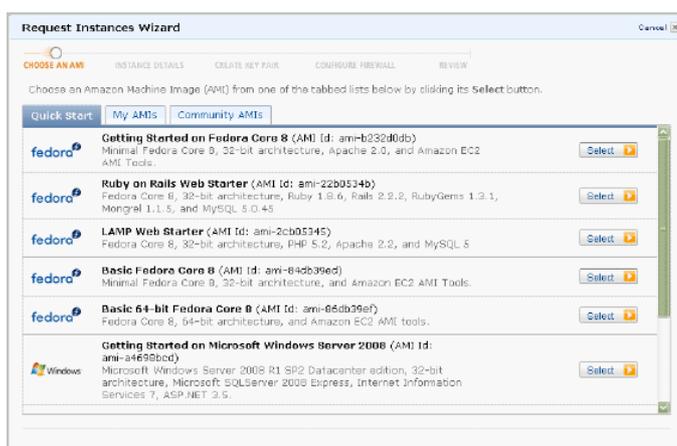


Figura 10: AWS Console - Selezione AMI

e le impostazioni delle policy del Security Group. Le istanze in esecuzione e le relative informazioni (tipo, stato, nome DNS) possono essere controllate sempre mediante la Console. In alternativa alla Console è possibile utilizzare i tool da riga di comando forniti sempre da Amazon.

A livello implementativo, l'accesso agli AWS viene effettuato principalmente tramite HTTP,

	Instance	AMI ID	Root Device Type	Type	Status	Lifecycle	Public DNS
<input type="checkbox"/>	 i-b58ed1dd	ami-b232d0db	ebs	m1.small	 running	normal	ec2-174-129-126

Figura 11: AWS Console - Istanza in esecuzione

utilizzando i protocolli standard REST¹⁸ o SOAP¹⁹, permettendo agli sviluppatori di realizzare i propri strumenti senza troppe complicazioni con qualsiasi linguaggio, ma Amazon fornisce anche librerie per Java, Ruby, PHP e .NET, nonché rispettivi SDK.

La comunità ha inoltre rilasciato librerie open basate sulle interfacce SOAP e REST, quali JetS3t e typica²⁰ (in Java), e Boto²¹ (in Python).

2.4 OpenNebula



- Modello di servizio: IaaS
- Modello di rilascio: Private/Hybrid Cloud
- Sito web: <http://www.opennebula.org/>

OpenNebula è un toolkit open-source per l'implementazione di servizi cloud pubblici, privati e ibridi di tipo IaaS.

Può essere installato su un cluster, ogni nodo dev'essere munito di hypervisor, e una macchina funge da front-end. Le macchine virtuali sono identificate da immagini contenenti il proprio

18 Representational state transfer - http://it.wikipedia.org/wiki/Representational_State_Transfer

19 Simple Object Access Protocol - <http://it.wikipedia.org/wiki/SOAP>

20 code.google.com/p/typica/

21 code.google.com/p/boto/

disco virtuale. Le immagini devono essere accessibili ad ogni nodo, per questo è possibile configurare un repository tramite NFS, SFTP o HTTP.

Per interagire con OpenNebula è possibile utilizzare le API native in XML-RPC o direttamente la libreria di virtualizzazione libvirt e qualsiasi tool legato ad essa (es: virsh, virt-manager). Inoltre OpenNebula implementa un sottoinsieme delle EC2 Query API.

Può interagire con Amazon EC2 per fornire Hybrid Cloud.

2.5 EyeOS



- Modello di servizio: PaaS
- Modello di rilascio Public/Private Cloud
- Sito web: eyeos.org

EyeOS – the cloud computing operative system - è un ambiente desktop web, classificabile quindi come PaaS. Scritto principalmente in PHP, XML e JavaScript è rilasciato su licenza AGPL.

L'utente può utilizzare eyeOS dal sito ufficiale, è sufficiente registrarsi a ecedervi tramite browser. Proprio attraverso il browser di presenta un desktop virtuale dal quale è possibile eseguire l'upload dei propri file ma soprattutto, come in tutti gli ambienti del genere, lanciare applicazioni, manovrare finestre, gestire file, ecc.

Il sistema è fornito di una serie di applicazioni di default quali suite di ufficio (videoscrittura e foglio di calcolo), calendario, diario, email, e naturalmente un file manager. Ulteriori applicazioni possono essere installate grazie ad un package manager.

EyeOS è concepito per il lavorare online, individualmente ma soprattutto in collaborazione, è possibile cercare e aggiungere contatti di altri utenti della piattaforma e creare gruppi di lavoro, condividere file, note e attività.

Gli sviluppatori possono creare ulteriori applicazioni grazie ad un set di librerie e funzioni: l'eyeOS toolkit. Inoltre è possibile creare un proprio repository basato su Portage²² per la distribuzione del software.

EyeOS può essere installato su qualsiasi sistema che supporti Apache e PHP5 e fornire un'ottima alternativa di Private Cloud.

²² Sistema di gestione dei pacchetti di Gentoo Linux

3 Caratteristiche di Eucalyptus

Poiché Eucalyptus è il software libero scelto per effettuare le nostre sperimentazioni, ad esso è dedicato per intero questo capitolo.

Eucalyptus (Elastic Utility Computing Architecture for Linking Your Programs To Useful Systems) è un'infrastruttura software open source per l'implementazione di Private o Hybrid Cloud di tipo IaaS utilizzando (uno o più) cluster di calcolatori, nato come progetto di ricerca del Computer Science department della University of California a Santa Barbara. Nel 2009 fu fondata la Eucalyptus Systems Inc. per supportarne la commercializzazione²³. In sostanza il sistema permette di avviare, controllare, accedere e terminare macchine virtuali; e supporta gli hypervisor Xen e KVM.

Eucalyptus implementa le API degli Amazon Web Services (sia SOAP che REST), quindi è possibile non solo interagire con la piattaforma con gli stessi tool utilizzabili con AWS, ma anche migrare le risorse da Amazon a Eucalyptus e viceversa. Eucalyptus riprende anche alcune delle terminologie di Amazon, quali la definizione di istanza e di immagine, quest'ultima a differenza delle AMI di EC2 viene definita EMI (Eucalyptus Machine Image).

Il sistema è utilizzato dalla Canonical (viene infatti installato con Ubuntu Enterprise Cloud) e dalla NASA.



Figura 12: il logo di Eucalyptus

3.1 Architettura

L'architettura di Eucalyptus è flessibile e modulare, con un design gerarchico. Il sistema è costituito da cinque componenti, implementati come Web Service:

Node Controller

Controlla l'esecuzione, l'ispezione e la terminazione di istanze di macchine virtuali su un singolo host.

²³ <http://open.eucalyptus.com/about/story>

Cluster Controller

Raccoglie informazioni e schedula l'esecuzione delle macchine virtuali sui Node Controller a esso collegati, controlla le funzioni di networking delle istanze.

Walrus

Un servizio di storage che implementa l'interfaccia di Amazon S3, permettendo lo stoccaggio dei dati degli utenti e l'archiviazione delle immagini delle macchine virtuali.

Storage Controller

Servizio che implementa l'interfaccia di Amazon EBS, fornendo il meccanismo per lo storage persistente delle macchine virtuali.

Cloud Controller

L'interfaccia mediante cui utenti e amministratori accedono ai servizi. È in grado di raccogliere informazioni dai nodi sulle risorse hardware e software, e si occupa dello scheduling ad alto livello, il tutto inoltrando richieste ai cluster controller.

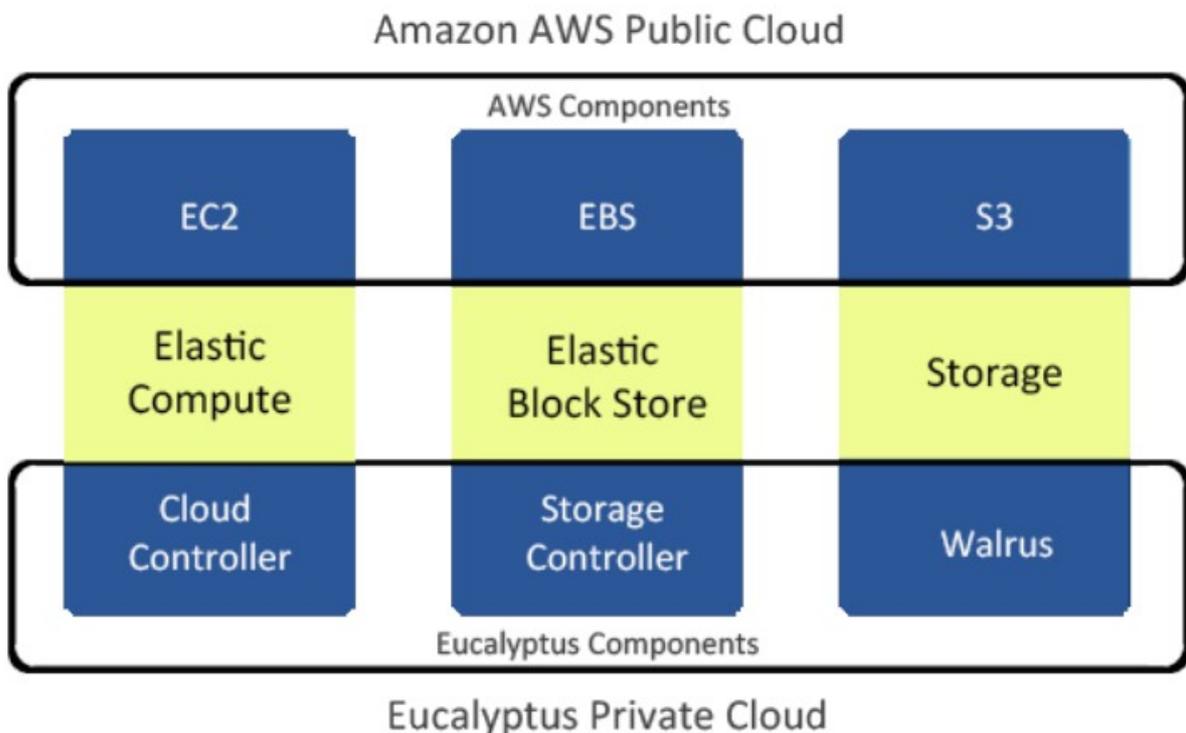


Figura 13: Implementazione di Eucalyptus dei servizi AWS

3.1.1 Node Controller

Il Node Controller (NC) viene installato su ogni nodo designato all'esecuzione delle macchine virtuali (VM), sul nodo deve quindi essere installato un hypervisor²⁴. Il NC tiene traccia delle risorse hardware e software della macchina su cui è installato, riceve richieste dal Cluster Controller e comunica con l'hypervisor.

Il nodo invia informazioni sulle risorse disponibili (CPU, memoria, disco) e sullo stato delle istanze a seguito delle rispettive richieste *describeResources* e *describeInstances* da parte del Cluster Controller. Il Cluster Controller ha il controllo sulle istanze mediante le richieste *runInstance* e *terminateInstance*. Ricevute tali richieste, il NC una volta controllate le autorizzazioni (solo il proprietario di un'istanza è autorizzato a terminarla) e confermato la disponibilità delle risorse le inoltra all'hypervisor.

I file necessari (kernel, ramdisk e filesystem dell'immagine) vengono prelevati dal repository quando è necessario avviare la corrispettiva EMI e mantenuti nella cache locale per eventuali utilizzi futuri.

3.1.2 Cluster Controller

Il Cluster Controller (CC) dev'essere installato sul front-end di un singolo cluster, o più precisamente su una macchina connessa a tutti i nodi del cluster (sul quale vengono installati i NC) e al Cloud Controller. Il CC ha tre funzioni principali: inoltrare richieste di avvio di macchine virtuali a specifici NC, controllare la copertura della rete virtuale delle istanze, e raccogliere informazioni sui NC. Ricevuta una richiesta di avvio di un set di VM, il CC invia la richiesta *describeResource* a tutti i nodi, e successivamente tramite la *runInstance* richiede al primo nodo che ha le risorse disponibili l'avvio di un'istanza.

Le funzioni di un cluster in Eucalyptus sono molto simili a quelle delle Availability Zone in AWS.

²⁴ Virtual machine monitor. Il software che si occupa della gestione delle macchine virtuali

3.1.3 Cloud Controller

Fornire l'accesso ai servizi e alle risorse virtuali sottostanti di Eucalyptus è compito del Cloud Controller (CLC). Il CLC è un insieme di Web Services suddivisi in tre categorie in base alla loro funzione:

Resource Services

Provvedono all'allocazione delle risorse, permettono all'utente di manipolare le proprietà delle macchine virtuali e della rete, e controllano componenti di sistema e risorse virtuali.

Data Services

Si occupano dei dati di sistema e forniscono un user environment per impostare le proprietà delle richieste di risorse.

Interface Services

Ovvero l'interfacciamento con l'utente, si occupano dell'autenticazione e della decodifica dei protocolli, e forniscono gli strumenti per il controllo del sistema.

I Resource services processano le richieste delle macchine virtuali e interagiscono con i CC per l'allocazione e la de-allocazione di risorse fisiche. Mediante la comunicazione con i CC vengono calcolate le risorse disponibili, utilizzate poi per verificare la fattibilità delle richieste degli utenti.

I Data Services gestiscono creazione, modifica, interrogazione e memorizzazione dei dati degli utenti. Gli utenti possono interrogare questi servizi per avere informazioni relative alle risorse disponibili (immagini e cluster) e manipolare i parametri applicabili alle macchine virtuali. Resource Services e Data Services interagiscono qualora devono essere effettuati cambiamenti a tali parametri (come regole di firewall).

In aggiunta alle interfacce SOAP e REST, gli Interface Services offrono un'interfaccia web per utenti e amministratori. Attualmente l'interfaccia web è in fase di sviluppo, gli utenti possono controllare i propri dati, la lista delle EMI, e scaricare le credenziali di accesso da usare con

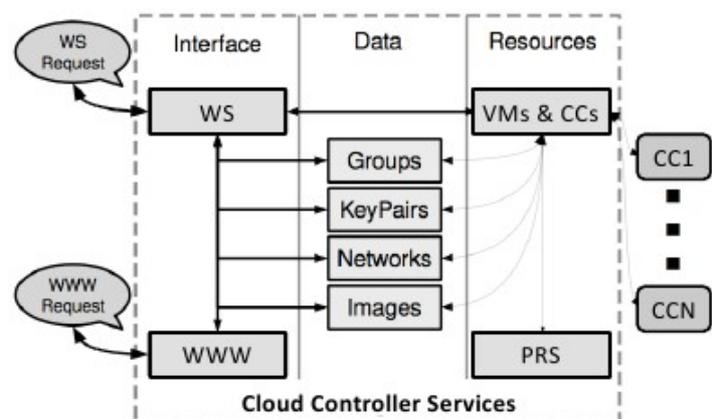


Figura 14: i servizi del Cloud Controller

i tool amministrativi; mentre gli amministratori possono configurare cluster, DNS, Walrus e dimensione delle risorse delle VM, nonché gestire le utenze.

3.2 Gestione degli utenti

Eucalyptus prevede due tipi di utenti, utenti ordinari o amministratori. Ciascun utente ha la possibilità di caricare i propri file nel Walrus, caricare le proprie EMI, avviare VM dalle proprie EMI o dalle immagini pubbliche e controllare le proprie istanze. L'utente può accedere all'interfaccia web tramite username e password e da lì ottenere le credenziali (standard x509) da utilizzare per l'accesso alle interfacce REST e SOAP.

3.3 Gestione EMI

Gli utenti possono caricare le immagini delle VM nel Walrus, che ne funge da repository. Una EMI (Eucalyptus Machine Image) è composta da un disco virtuale contenente il sistema da avviare e il relativo software, l'immagine del kernel, ed eventualmente l'immagine ramdisk, che il sistema utilizzerà, l'hypervisor infatti funge anche da bootloader per le istanze. È possibile specificare per ogni disco kernel e ramdisk da utilizzare, sia al

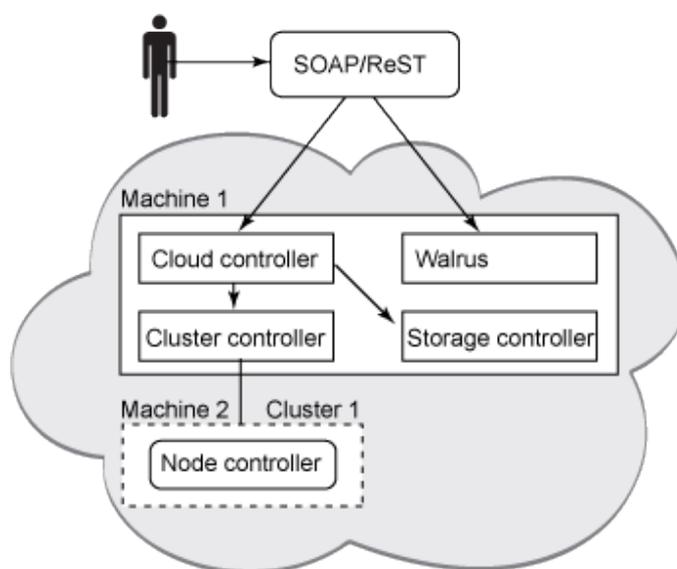


Figura 15: topologia di un'installazione mono-cluster

caricamento del file del disco che all'avvio dell'istanza; qualora non fossero specificati verranno usate le immagini di default definite dall'amministratore. Per essere utilizzati i file devono essere "impacchettati" e criptati utilizzando i certificati dell'utente, caricati e registrati. Ad ogni file

(disco virtuale, kernel e ramdisk) alla registrazione vengono assegnati id univoci, da utilizzare per l'avvio delle istanze.

3.4 Network Management

Uno degli aspetti chiave di un IaaS come Eucalyptus è la configurazione della rete delle VM. È infatti necessario garantire l'interconnettività tra le istanze, il loro accesso alla rete pubblica e facilitare quanto più possibile aspetti comuni delle VM quali firewall e indirizzi.

Per rendere l'amministrazione il più flessibile possibile Eucalyptus prevede quattro modalità di configurazione della rete, ognuna con i propri parametri, benefici e restrizioni:

SYSTEM Mode

La modalità più semplice, ma anche quella che permette meno opzioni. Le VM vengono agganciate direttamente nella rete fisica tramite un bridge²⁵ configurato sui nodi, l'indirizzo viene assegnato alle istanze tramite dhcp. Il server dhcp può essere installato su qualunque macchina attaccata alla rete e le VM possono essere viste dall'esterno come macchine reali.

STATIC Mode

Simile a SYSTEM, le VM vengono agganciate alla rete fisica ma gli indirizzi IP che ricevono devono essere pre-impostati dall'amministratore, specificando coppie MAC/IP. All'avvio di un'istanza essa riceve la prima coppia disponibile. Se gli indirizzi statici sono stati tutti assegnati non è più possibile avviare altre istanze. Per utilizzare questa modalità è necessario installare un server dhcp sul CC.

MANAGED Mode

La modalità che offre più potenzialità ma è anche la più ardua da configurare. Le macchine virtuali vengono collegate in una VLAN inaccessibile dall'esterno e completamente controllata da Eucalyptus. Per l'accesso alla rete pubblica le VM sono provviste di una seconda interfaccia, l'amministratore specifica gli indirizzi IP pubblici assegnabili e ogni istanza all'avvio ne riceve uno, tramite un server dhcp configurato sul CC. In modo simile agli Elastic IP Addresses di EC2 l'utente può allocare ulteriori indirizzi pubblici (e deallocarli quando non più necessari) e assegnarli alle istanze dinamicamente. Se gli indirizzi pubblici vengono esauriti non è più possibile avviare altre VM. Le istanze

25 Per bridgin si intende collegare una o più interfacce di rete virtuali ad un'unica interfaccia reale

possono inoltre essere collocate in gruppi con le stesse funzionalità dei Security Groups di Amazon, a livello implementativo gruppi diversi appartengono a VLAN differenti. Affinché tale modalità sia funzionante occorre che la rete fisica sia VLAN clean, ovvero che il firewall non blocchi i pacchetti contrassegnati come appartenenti ad una VLAN.

MANAGED-NOVLAN Mode

Simile a MANAGED Mode con l'unica differenza che le VM non sono collegate ad una VLAN ma ad una rete reale, che rimane comunque separata dalla rete fisica. Il CC funge da router tra rete fisica e rete delle VM, le interfacce virtuali delle istanze sono messe in bridge con l'interfaccia reale del nodo. La divisione in gruppi in questa modalità è effettuata dividendo la rete in subnet.

3.5 Controllo delle istanze

Per avviare le proprie istanze occorre selezionare una EMI tra quelle disponibili ed eseguire una sorta di richiesta. La richiesta identifica il cluster (o Availability Zone) che si occuperà di ospitare le istanze, il numero di istanze necessarie, il tipo di istanze, la coppia di chiavi da utilizzare per l'accesso SSH, il Security Group a cui assegnarle, l'id della EMI selezionata, immagini di kernel e ramdisk (ERI ed EKI) da utilizzare. Qualora uno o più dei parametri non fossero specificati verranno usati i valori di default.

Una richiesta può essere inviata ad un solo cluster. Ricevuta la richiesta il Cluster Controller analizza le risorse dei nodi ad esso collegati e se possibile avvia le istanze, altrimenti comunica l'indisponibilità all'utente.

Il tipo di istanza indica il livello prestazionale della macchina. Eucalyptus definisce cinque tipologie, `m1.small`, `c1.medium`, `m1.large`, `m1.xlarge` e `c1.xlarge`. I nomi sono già abbastanza indicativi (le small sono le meno performanti, mentre le xlarge hanno maggiore potenza di calcolo), le caratteristiche di ciascuna tipologia sono definite dall'amministratore, ma gli utenti possono interrogare il sistema per averne i dettagli.

L'accesso SSH alle VM viene effettuato utilizzando per l'autenticazione una coppia di chiavi RSA. Pertanto se si intende utilizzare SSH, prima di avviare un'istanza occorre definire il keypair, all'avvio della VM la chiave pubblica specificata verrà copiata nel relativo filesystem, permettendo il login utilizzando la chiave privata.

Un'istanza ha un id univoco assegnato all'avvio ed uno stato, l'utente può interrogare il sistema per avere informazioni sulle proprie istanze (id, stato, indirizzi, gruppo di appartenenza, chiave utilizzata). Appena avviata, lo stato di un'istanza è impostato su pending, vale a dire in preparazione, quando l'istanza sarà pronta il suo stato passerà a running, a questo punto esse avranno un indirizzo IP e sarà possibile accedervi. Tramite l'id delle istanze invece è possibile riavviare le VM, terminarle, allocare gli indirizzi IP dinamici (Elastic IP Addresses) e controllare l'output della console.

3.6 Euca2ools

Gli Euca2ools sono tool da riga di comando utilizzabili per interagire con Eucalyptus. Sono implementati in Python utilizzando la libreria Boto, sul modello dei tool forniti da Amazon, mantenendo la compatibilità con AWS.

Le caratteristiche includono:

- query sulle Availability Zone (cluster per Eucalyptus);
- gestione chiavi SSH (aggiungere, listare, cancellare);
- gestione macchine virtuali (avviare, listare, riavviare, terminare, leggere l'output della console);
- gestione Security Groups;
- gestione volumi e snapshot;
- gestione immagini;
- gestione Elastic IP Addresses.

3.7 Eucalyptus Community Cloud

Eucalyptus può essere sperimentato grazie al contributo di HP, CoreSite, InfoRelay e DasherTechnology, i quali insieme ad Eucalyptus System forniscono un ambiente di test. Il servizio è gratuito e può essere utilizzato da chiunque, sottostando ad alcuni limiti quali:

- un'istanza non può essere tenuta in esecuzione per più di sei ore;
- ciascun utente può eseguire al massimo quattro istanze;
- l'allocazione di un elastic IP non può durare più di sei ore;
- un bucket viene preservato nel Walrus al massimo per tre settimane;
- la dimensione massima di un bucket è di 5GB;
- un volume EBS viene preservato al massimo per tre settimane;
- la dimensione massima di un volume EBS è di 5GB.

La versione di Eucalyptus utilizzata è la 2.0, l'hypervisor KVM e la rete è configurata in MANAGED Mode.

3.8 Installazione

Eucalyptus è stato pensato per essere facile da installare e poco intrusivo, non occorre infatti apportare modifiche al sistema operativo o installare software particolari, i vari componenti di Eucalyptus e gli Euca2ools sono installabili tramite pacchetti binari su Debian, CentOS, openSuse e Ubuntu; e naturalmente sono disponibili i sorgenti per l'installazione su una generica distribuzione GNU/Linux. I requisiti minimi per l'installazione della piattaforma si aggirano intorno ad 1GHz di CPU e 1GB di memoria.

3.8.1 Ubuntu Enterprise Cloud (UEC)

Canonical ha recentemente rilanciato la versione Cloud della sua distribuzione, la quale utilizza Eucalyptus come componente fondamentale. Utilizzando UEC si ha la possibilità di installare, tramite CD, sistema operativo ed Eucalyptus, nel modo più semplice possibile.

3.9 Eucalyptus Enterprise Edition

Eucalyptus è disponibile anche nella versione Enterprise, con l'aggiunta di diverse funzionalità quali supporto per VMware e VM basate su Windows.

Eucalyptus Feature Comparison	Open Source	Enterprise Edition
Amazon AWS Interface Compatibility	X	X
Flexible Clustering and Availability Zones	X	X
Network Management, Security Groups, Traffic Isolation	X	X
Cloud Semantics and Self-Service Capability	X	X
Bucket-Based Storage Abstraction (S3-Compatible)	X	X
Block-Based Storage Abstraction (EBS-Compatible)	X	X
Xen and KVM Hypervisor Support	X	X
VMware Hypervisor Support		X
Virtual-to-Virtual Image Conversion for VMware		X
Microsoft Windows Guest Support		X
Direct SAN Integration		X
Quota Management and Accounting		X
User and Group Access Management		X
High-Performance MySQL Database Backend		X

Confronto delle caratteristiche della versione Open e della versione Enterprise di Eucalyptus²⁶

²⁶ <http://www.eucalyptus.com/products/eee>

4 Installazione e sperimentazione di Eucalyptus

La sperimentazione effettuata per questo lavoro di tesi si colloca nell'ambito dell'HPC e dei computer cluster, al fine di valutare in modo quantitativo l'impatto della virtualizzazione sulle performances delle macchine. Il calo inevitabile delle prestazioni deve essere quantificato al fine di poter valutare costi e benefici dell'utilizzo del paradigma di Cloud nell'ambito dei sistemi di calcolo scientifico. In un ambiente di calcolo con molte macchine e molti utenti si hanno solitamente necessità differenti, in termini di ambienti, librerie, ottimizzazioni di sistema. Il Cloud apporta in questo caso vantaggi indiscutibili, come ad esempio la possibilità di avviare all'occasione un cluster con macchine su cui è installato un ambiente specifico per un determinato calcolo, configurato appositamente ed ottimizzato per esso.

La finalità della parte sperimentale di questo lavoro di tesi è quella di installare il sistema cloud Eucalyptus su alcune macchine di un cluster, lanciare dentro di esse un ambiente di calcolo virtualizzato e confrontarne le prestazioni con quelle dello stesso ambiente di calcolo installato direttamente sulle macchine fisiche.

La sperimentazione e' stata effettuata sotto la supervisione di Uniclust S.r.l., lo Spin Off dell'Università della Calabria che opera nel settore HPC, su macchine del Dipartimento di Fisica.

4.1 I Cluster del Grid del Dipartimento di Fisica

Il Dipartimento di Fisica dispone, tra gli altri, di tre cluster in griglia, denominati Titania, Ambiente e Darwin.

Titania è composto da otto nodi, ognuno con due processori AMD Opteron dual-core a 64 bit con 2.0 GHz di clock. Il cluster può fornire potenza di calcolo pari a 32 CPU. I nodi hanno 3 GB di memoria e 80 GB di disco SATA ciascuno.

Ambiente è composto sempre da otto nodi, ognuno con doppio processore Intel Quad Xeon della serie 5400 a 64 bit da 2.5 GHz di clock, fornendo fino a 64 CPU per le operazioni di calcolo. Ogni nodo dispone di 8 GB di memoria.

Darwin è composto da quattro nodi ognuno con due processori Intel Quad Xeon della serie 5500 a 64 bit da 2.4 GHz di clock, fornendo 32 CPU. Ogni nodo dispone di 12 GB di memoria.

Tutte le macchine utilizzano come sistema operativo CentOS 4.5. Per la sperimentazioni sono stati utilizzati due dei nodi di Darwin.

4.2 Cloud e HPC

Sebbene le promesse di riduzione dei costi e massima flessibilità sembrano piuttosto allettanti, fin'ora il campo in cui il Cloud Computing ha avuto scarso successo è quello del calcolo ad alte prestazioni[18][19]. Il modello Cloud, pensato e realizzato per essere il più possibile general-purpose, ben poco si adatta ad un campo in cui sono richieste soluzioni il più possibile particolareggiate. Per ottenere prestazioni elevate è necessario avere diretto controllo e conoscenza dell'hardware, la virtualizzazione è una semplificazione dell'hardware, una maschera, uno strato aggiuntivo che aggiunge flessibilità ma ad un caro prezzo, la perdita di prestazioni può arrivare a essere notevole. Studi della Denison University[20] hanno stimato un divario prestazionale che va da un contenutissimo 3% ad un pesante 73%.

D'altro canto stanno emergendo soluzioni Cloud create appositamente per HPC di recente Amazon ha comunicato la sua offerta speciale[21], le cluster compute instances, in grado di offrire prestazioni molto elevate.

Resta da capire bene quale sia in realtà il divario e se esso sia colmabile, se è solo questione di tempo o se il vantaggio economico offerto dal modello cloud scompare quando si deve sopperire alle sue mancanze.

4.3 Installazione di Eucalyptus

Per l'installazione di Eucalyptus si è scelto di utilizzare sui due nodi Darwin un sistema Debian squeeze. D'ora in poi i due nodi saranno riferiti come test-cloud-1 e test-cloud-2. I due nodi sono connessi alla rete locale, per semplificare la configurazione può essere utile aggiungere le informazioni sui rispettivi indirizzi IP

```
test-cloud-1    192.167.201.218
test-cloud-2    192.167.201.219
```

al file

```
/etc/hosts
```

test-cloud-1 è stato configurato sia come front-end di Eucalyptus che come nodo, vi sono quindi stati installati tutti i componenti di Eucalyptus: Cloud Controller, Walrus, Cluster Controller, Storage Controller e Node Controller; test-cloud-2 funge solo da nodo, quindi su di esso è stato configurato solo il Node Controller.

L'ultima versione di Eucalyptus è la 2.0, per debian sono disponibili pacchetti precompilati, nel formato standard della distribuzione (.deb), per consentire una rapida e semplice installazione. L'installazione è divisa in diversi pacchetti (un pacchetto per ogni componente) in modo da garantire massima flessibilità. I pacchetti sono reperibili da un repository esterno, per poterli scaricare è sufficiente aggiungere la seguente riga

```
deb http://eucalyptussoftware.com/downloads/repo/eucalyptus/2.0.0/debian/
squeeze main
```

al file

```
/etc/apt/sources.list
```

ed eseguire da riga di comando

```
# apt-get update
```

inoltre bisogna sincronizzare gli orologi di sistema delle varie macchine, usando ad esempio ntp, e assicurarsi che il firewall permetta la comunicazione tra i componenti di Eucalyptus (nel front-end devono essere accessibili le porte 8443, 8773, 8774 e 9001; nei nodi la 8775).

4.3.1 Setup di un nodo

Su entrambe le macchine utilizzate dev'essere installato un Node Controller, entrambe devono essere configurate quindi come nodo di Eucalyptus. Bisogna perciò innanzi tutto installare un hypervisor, la scelta è ricaduta su quello più semplice da configurare, kvm; e configurare un bridge in modo da poter usare la modalità SYSTEM. Per installare i pacchetti necessari

```
# apt-get install qemu-kvm bridge-utils
```

per configurare il bridge è sufficiente editare il file

```
/etc/network/interfaces
```

aggiungendo le seguenti righe

```
auto br0
iface br0 inet dhcp
    bridge_ports eth0
```

in questo modo sull'interfaccia connessa alla rete locale (eth0) viene impostato un bridge e l'indirizzo della macchina assegnato tramite dhcp. Per far si che le impostazioni abbiano effetto

```
# /etc/init.d/networking restart
```

a questo punto è possibile installare il Node Controller

```
# apt-get install eucalyptus-nc
```

4.3.2 Setup del front-end

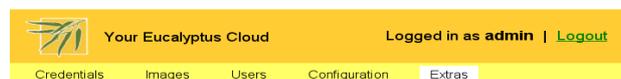
Test-cloud-1 funge da front-end, devono quindi essere installati i componenti necessari al controllo del sistema cloud

```
# apt-get install eucalyptus-cloud eucalyptus-walrus eucalyptus-sc
eucalyptus-cc
```

una volta installati, i vari componenti devono essere registrati

```
# euca_conf --register-walrus test-cloud-1
# euca_conf --register-cluster cluster1 test-cloud-1
# euca_conf --register-sc cluster1 test-cloud-1
# euca_conf --register-nodes test-cloud-1 test-cloud-2
```

dove **cluster1** è il nome assegnato al cluster.



Eucalyptus-certified Images

Name	Description
euca-centos-5.3-i386.tar.gz	CentOS 5.3 i386
euca-centos-5.3-x86_64.tar.gz	CentOS 5.3 x86_64
euca-debian-5.0-i386.tar.gz	Debian 5.0 (lenny) i386
euca-debian-5.0-x86_64.tar.gz	Debian 5.0 (lenny) x86_64
euca-fedora-10-x86_64.tar.gz	Fedora 10 x86_64
euca-fedora-11-i386.tar.gz	Fedora 11 i386
euca-ubuntu-9.04-i386.tar.gz	Ubuntu 9.04 (jaunty) i386
euca-ubuntu-9.04-x86_64.tar.gz	Ubuntu 9.04 (jaunty) x86_64

Eucalyptus-compatible Tools

Name	Description
euca2ools	Eucalyptus Client Tools
other_clients	Eucalyptus Ecosystem Page

4.4 Utilizzo di Eucalyptus

A installazione completata è possibile accedere all'interfaccia web del front end. Qui è possibile effettuare ulteriori configurazioni, ai fini della

Figura 16: La lista delle immagini certificate da Eucalyptus

sperimentazione occorre modificare le risorse allocate dalle VM. Tali dettagli sono editabili nella scheda **Configuration**, in fondo alla pagina, alla voce **VM Types**. Essendo interessati ad avere la massima prestazione possibile dalle istanze, le risorse utilizzate dalla **c1.xlarge** sono state impostate in modo da essere prossime a quelle della macchina reale, cioè 8 CPU e 11GB di memoria. Con questa configurazione Eucalyptus consentirà l'esecuzione di due sole istanze **c1.xlarge**, una per ciascun calcolatore.

Sempre da interfaccia web occorre scaricare le credenziali dell'utente, necessarie per utilizzare un qualsiasi client per Eucalyptus. Le credenziali possono essere trovate nella scheda **Credentials**.

La scheda **Extras** contiene link ad alcuni strumenti utilizzabili per interagire col sistema ma soprattutto da qui è possibile scaricare le immagini di macchine virtuali certificate dalla comunità di Eucalyptus, evitando così di configurarle manualmente. Sono disponibili le immagini delle distribuzioni più comuni: Ubuntu, Debian, OpenSuse e, quella utilizzata per la sperimentazione, CentOS.

Your Eucalyptus Cloud Logged in as admin | Logout

Credentials Images Users Configuration Extras

Account of user 'utente1' was updated

Users

Username	Email	Name	Status	Actions
utente1	test@test.com	Utente di prova	active	Edit Disable Delete
admin	n/a	admin	active & admin	Edit

Add user

Figura 18: Pagina di gestione utenze di Eucalyptus

Your Eucalyptus Cloud Logged in as admin | Logout

Credentials Images Users Configuration Extras

Cloud configuration:

Cloud Host: 192.167.201.219
 Default kernel: Default ramdisk:
 Save Configuration Loaded configuration from server

DNS configuration:

Domain name: localhost
 Nameserver: nshost.localhost IP: 127.0.0.1
 Save Configuration Loaded configuration from server

Walrus Configuration:

Walrus host: 192.167.201.219 Deregister
 Maximum bucket size (MB): 5120
 Space reserved for snapshots (GB): 50
 Maximum buckets per user: 5
 Buckets Path: /var/lib/eucalyptus/bukkits
 Space reserved for unbundling images (MB): 30720
 Register Walrus Save Walrus configuration Walrus configuration up to date

Clusters:

Name: cluster1 Deregister Cluster
 Cluster Controller
 Host: 192.167.201.219
 Dynamic public IP address assignment
 Reserve for assignment: 10 public IP addresses
 Maximum of: 5 public IP addresses per user
 Use VLAN tags: 10 through 4095
 Storage Controller
 Max volume size: 10
 Disk space reserved for volumes: 50
 Storage interface: eth0
 Volumes path: /var/lib/eucalyptus/volumes
 Zero-fill volumes
 Register cluster Save cluster configuration Clusters up to date

VM Types:

Name	CPUs	Memory (MB)	Disk (GB)
m1.small	1	1400	5
c1.medium	2	2800	5
m1.large	4	5600	10
m1.xlarge	7	9800	20
c1.xlarge	8	11200	20

Save VM Types

Figura 17: Pagina di configurazione di Eucalyptus

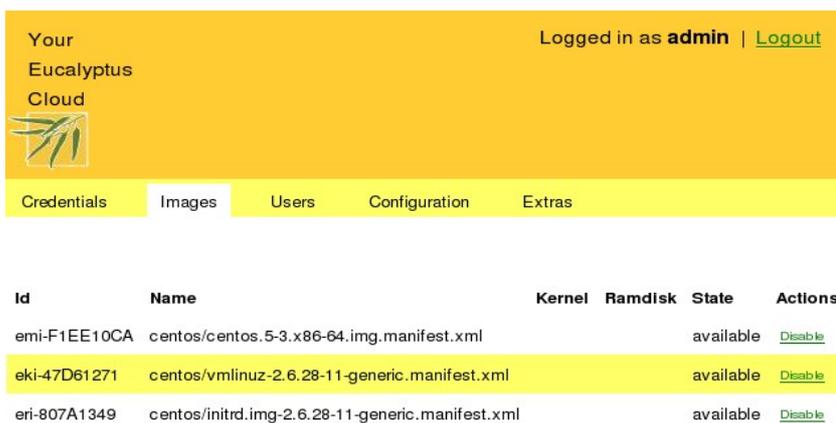
Una volta scaricata l'immagine della CentOS essa dev'essere caricata nel Walrus, utilizzando un qualsiasi client, nel nostro caso Euca2ools²⁷.

Gli Euca2ools utilizzano variabili d'ambiente per identificare l'indirizzo del front-end, le chiavi di accesso e i certificati, non occorre tuttavia impostarle manualmente in quanto sono definite nel file **euarc**, presente nella cartella delle credenziali. Prima di utilizzare i tool è sufficiente spostarsi nella cartella delle credenziali ed eseguire

```
$ source euarc
```

Per rendere disponibile la EMI di CentOS occorre ora impacchettare, caricare e registrare kernel, ramdisk e disco dell'immagine, tutti e tre i file sono presenti nell'archivio scaricato.

Comando:
<code>\$ euca-bundle-image -i vmlinuz-2.6.28-11-server</code>
Output:
Checking image Tarring image Encrypting image Splitting image... Part: vmlinuz-2.6.28-11-server.part.0 Generating manifest /tmp/vmlinuz-2.6.28-11-server.manifest.xml
L'immagine, in questo caso il kernel, viene compressa, criptata e divisa in più parti, in questo caso una sola, più un file descrittivo, il manifest.



Id	Name	Kernel	Ramdisk	State	Actions
emi-F1EE10CA	centos/centos.5-3.x86-64.img.manifest.xml			available	Disable
eki-47D61271	centos/vmlinuz-2.6.28-11-generic.manifest.xml			available	Disable
eri-807A1349	centos/initrd.img-2.6.28-11-generic.manifest.xml			available	Disable

Figura 19: La lista delle EMI disponibili

²⁷ http://open.eucalyptus.com/wiki/Euca2oolsGuide_v1.3

Comando:

```
$ euca-upload-bundle -b centos -m  
/tmp/vmlinuz-2.6.28-11-server.manifest.xml
```

-b indica il nome del bucket su cui sarà caricato il file, se il bucket non esiste viene creato.
-m indica il file manifest del bundle da caricare, ottenuto dal comando precedente

Output:

```
Checking bucket: centos/  
Creating bucket: centos/  
Uploading manifest file  
Uploading part: vmlinuz-2.6.28-11-server.part.0  
Uploaded image as centos/vmlinuz-2.6.28-11-server.manifest.xml
```

Comando:

```
$ euca-register centos/vmlinuz-2.6.28-11-server.manifest.xml
```

Output:

```
IMAGE      eki-36AD1232
```

Il comando abilita l'utilizzo del file, l'input è formato da bucket/file, in output viene restituito l'identificativo del file.

Occorre ripetere la stessa procedura con l'immagine del ramdisk e del disco virtuale, dopodiché l'EMI sarà a disposizione

Comando:

```
$ euca-describe-images
```

Output:

```
IMAGE      eri-6EA31316  centos/initrd.img-2.6.28-11-server.manifest.xml  admin      available public      x86_64      ramdisk

IMAGE      eki-36AD1232  centos/vmlinuz-2.6.28-11-server.manifest.xml  admin      available public      x86_64      kernel

IMAGE      emi-C9EB1037  centos/centos.5-3.x86.img.manifest.xml  admin      available public      x86_64      machine      eki-36AD1232      eri-6EA31316
```

Per ogni immagine vengono indicati l'id, la collocazione del file manifest, l'utente proprietario (admin) , lo stato (available public), l'architettura (x86_64), tipo di immagine (kernel/ramdisk/machine).

Nel caso della EMI, vengono specificati gli id delle immagini di kernel e ramdisk utilizzate di default.

4.5 Avviare le macchine virtuali

Registrata correttamente la EMI è possibile avviare la corrispondente istanza, prima però occorre generare una coppia di chiavi per permettere successivamente l'accesso SSH alla VM. Per fare ciò

```
$ euca-add-keypair mykey > mykey.private
```

la chiave pubblica verrà registrata da Eucalyptus per essere copiata nel disco delle istanze una volta avviate, il comando restituisce in output la chiave privata, che è bene memorizzare in un file e settarne opportunamente i permessi

```
$ chmod 600 mykey.private
```

The screenshot shows the Eucalyptus Cloud user interface. At the top, it says "Your Eucalyptus Cloud" and "Logged in as admin | Logout". Below this is a navigation menu with "Credentials" selected. The main content area is titled "User account Information" and displays the following details:

- Login: admin
- Name: admin
- Email: n/a

Below the details, there is a note: "Feel free to change the account information (except the login) and the password whenever you want. The cryptographic credentials for the Web services associated with this account, shown below, will not be affected by these changes." There are two buttons: "Edit Account Information" and "Change Password".

The next section is "Credentials ZIP-file". It contains a note: "Click the button to download a ZIP file with your Eucalyptus credentials. Use the public/private key pair included therein with tools that require X.509 certificates, such as Amazon's EC2 command-line tools." There is a "Download Credentials" button.

The next section is "Query interface credentials". It contains a note: "Use this pair of strings with tools - such as [euca2ools](#) - that utilize the 'query interface' in which requests and parameters are encoded in the URL." It displays "Query ID:" and "Secret Key:" with a "Show keys" button.

The final section is "Cloud Registration". It contains a note: "Your Eucalyptus cloud can be registered with [Rightscale](#), providing you and your users with an easy-to-use interface to both EC2 and your Eucalyptus cloud. The following two parameters will be necessary for registration:" It displays "Cloud URL: https://192.167.201.219:8443/register" and "Cloud ID: c3d04a47-cbaf-49ca-81db-5e657a5ecf44". There is a "Register with RIGHTSACLE" button.

Figura 20: Pagine delle credenziali dell'utente admin

È possibile ottenere in ogni momento le informazioni sulle chiavi registrate con

```
$ euca-describe-keypairs
```

Ora è possibile avviare l'istanza utilizzando l'id della EMI registrata in precedenza, specificandone chiave, tipologia e numero

Comando:				
\$ euca-run-instances	-k mykey	-t c1.xlarge	-n 2	emi-C9EB1037
Nome della chiave				
tipo di istanza				
numero di istanze da avviare				
id della EMI				
Output:				
RESERVATION	r-422408C2	admin	default	
INSTANCE	i-365F0765	emi-C9EB1037	0.0.0.0	0.0.0.0
pending	mykey 0	c1.xlarge		
	2010-09-16T09:58:30.444Z	cluster1	eki-36AD1232	
	eri-6EA31316			
INSTANCE	i-3F7A0739	emi-C9EB1037	0.0.0.0	0.0.0.0
pending	mykey 1	c1.xlarge		
	2010-09-16T09:58:30.447Z	cluster1	eki-36AD1232	
	eri-6EA31316			
La RESERVATION indica la richiesta effettuata dall'utente (admin) e il gruppo di cui fanno parte le relative istanze (default). Per ogni istanza vengono indicati id univoco (i-365F0765/ i-3F7A0739), id corrispettiva EMI (emi-C9EB1037), indirizzo IP pubblico e privato (non ancora assegnati), chiave utilizzata (mykey), tipo di istanza (c1.xlarge), data di attivazione, cluster su cui sono state eseguite (cluster1) e id di kernel e ramdisk utilizzati.				

Nella procedura di avvio il Cluster Controller decide su quali nodi devono essere avviate le istanze, in questo caso si può essere certi che ogni nodo avvierà un'istanza, si può comunque controllare su quale nodo verranno eseguite tramite un file di log nel front-end

```
# grep 'on resource' /var/log/eucalyptus/eucalyptus.log
```

Una volta contattati i nodi per l'avvio delle VM essi devono scaricare dal Walrus i file necessari, e istruire l'hypervisor. La modalità con cui Eucalyptus gestisce il networking di default è impostata su SYSTEM, quindi al boot le VM richiederanno un indirizzo al server dhcp

presente in rete. È possibile controllare l'intera procedura di avvio tramite i log del Node Controller

```
# tail -f /var/log/eucalyptus/nc.log
```

Le istanze dopo pochi minuti sono in esecuzione e il loro stato passa da pending a running, una volta avviata è possibile ottenere l'output della console dell'istanza tramite l'id, ad esempio

```
$ euca-get-console-output i-365F0765
```

naturalmente in ogni momento è possibile controllare stato e indirizzi delle istanze.

Comando:			
\$ euca-describe-instances			
Output:			
RESERVATION	r-422408C2	admin	default
INSTANCE	i-365F0765	emi-C9EB1037	192.167.201.50
192.167.201.50	running	mykey 0	c1.xlarge
2010-09-16T09:58:30.444Z		cluster1	eki-36AD1232
eri-6EA31316			
INSTANCE	i-3F7A0739	emi-C9EB1037	192.167.201.51
192.167.201.51	running	mykey 1	c1.xlarge
2010-09-16T09:58:30.447Z		cluster1	eki-36AD1232
eri-6EA31316			
L'output è lo stesso del comando precedente, ora si può notare che lo stato è passato a running e che le istanze hanno ricevuto un indirizzo IP			

Per accedere in SSH ad un'istanza

```
$ ssh -i mykey.private 192.167.201.50
```

4.6 Test e benchmark

Il test prestazionale è stato effettuato utilizzando HPL²⁸, un'implementazione portabile per sistemi a memoria distribuita del High-Performance Linpack Benchmark.

I Linpack Benchmark sono utilizzati per misurare la potenza di calcolo di una macchina su

²⁸ <http://www.netlib.org/benchmark/hpl/>

numeri a virgola mobile. I test si basano sulla velocità con cui il computer risolve un sistema di equazioni lineari, i risultati vengono riportati in miliardi di operazioni a virgola mobile per secondo, in linguaggio tecnico GFLOPs.[3]

HPL distribuisce il carico di lavoro tra i vari nodi del cluster utilizzando MPI²⁹, è il benchmark utilizzato per stilare la classifica dei top500, i 500 supercomputer più potenti al mondo, e nel campo del calcolo ad alte prestazioni è considerato uno standard.

La versione di Linpack utilizzata è la 2.0, sul sistema operativo CentOS 5.3, il compilatore GNU GCC 4.1.2, e la libreria software per algebra lineare ATLAS³⁰ 3.7.10. L'ambiente software è stato riprodotto fedelmente nel cluster virtuale realizzato tramite Eucalyptus e in un cluster reale con la stessa capacità di calcolo.

4.7 Analisi dei risultati

Il software utilizzato risolve un sistema lineare $Ax=b$ di ordine N . Sono stati effettuati cinque test variando l'ordine della matrice A e confrontate le prestazioni dei due cluster.

N	Cluster reale		Cluster virtuale		Rapporto
	Time	GFlops	Time	GFlops	
40000	1089.30.00	3.917e+01	1884.41.00	2.264e+01	57.65%
32000	556.60	3.925e+01	948.50	2.303e+01	58.68%
24000	228.79	4.029e+01	446.71	2.063e+01	51.20%
12000	35.34	3.260e+01	69.86	1.649e+01	50.58%
4000	3.23	1.320e+01	7.57	5.638e+00	42.71%

i risultati dei benchmark a confronto, catalogati in base alla taglia della matrice in input. Il tempo è espresso in secondi.

Come si evince dalla tabelle i risultati rispettano le aspettative, le prestazioni del cluster virtuale sono in media il 52% inferiori delle prestazioni del cluster reale, con un calo del 48% e un rapporto migliore con i dataset più grandi.

29 http://en.wikipedia.org/wiki/Message_Passing_Interface

30 http://en.wikipedia.org/wiki/Automatically_Tuned_Linear_Algebra_Software

Tra i fattori che possono incidere sulle prestazioni vi sono l'hypervisor utilizzato (nella presente sperimentazione è stato utilizzato KVM, l'uso di Xen o VMware ad esempio potrebbe portare a risultati molto differenti); il numero di core e la frequenza di operazioni di I/O.

La sperimentazione della Denison University[20] ha infatti dimostrato che il calo aumenta inversamente al numero di CPU: esso si mantiene a 22% utilizzando 12 core e passa al 73% utilizzando un solo processore; mentre applicazioni con limitate operazioni di I/O possono contenere il calo al 3%.

Un'altra analisi interessante emerge da uno studio della Vrije Universiteit (Amsterdam) e della Aachen University (Germania)[22], in cui vengono analizzate le prestazioni di cluster realizzati mediante Amazon EC2(sebbene allora non fossero ancora disponibili le cluster compute instances) sempre utilizzando HPL, affrontando anche l'impatto economico della soluzione Cloud in base ai GFlops ottenuti dal benchmark. I risultati di tale esperimento dimostrano che le prestazioni calano all'aumentare del numero dei nodi (aumentando la taglia dell'input proporzionalmente ad essi) così come i GFlops per dollaro ottenuti, e il costo di un GFlop aumenta esponenzialmente in base all'input.

Dai risultati della sperimentazione risulta un forte calo delle prestazioni dovuto alla virtualizzazione, che difficilmente si compensa con gli altri vantaggi delle soluzioni Cloud per il calcolo parallelo. Tuttavia la tecnologia è giovane e molti sforzi stanno facendo i produttori di Hardware e Software per elevare l'efficienza dei sistemi virtualizzati e ciò fa ben sperare per il prossimo futuro.

CONCLUSIONI

Il Cloud Computing è di sicuro una grande novità e rappresenta di sicuro una svolta nelle Information Technologies, e ciò è stato dimostrato dalla rapida crescita che ha avuto.

Tra le opportunità offerte vi sono la possibilità di usare i propri software ovunque; utilizzare i programmi senza più preoccuparsi dell'infrastruttura; riduzione di costi e tempi.

Con Google App Engine è possibile sviluppare un'applicazione web in brevissimo tempo senza pensare a dove e come dev'essere eseguita e senza tener conto di picchi di visite, crescita improvvisa o periodi di inattività.

Con EC2 in pochi minuti si può usufruire di tutte le macchine necessarie per gli usi più disparati.

Sebbene il Cloud sia governato più che altro dai potenti che non si preoccupano della libertà dei comuni mortali le soluzioni libere non mancano e anzi stanno prendendo piede. Eucalyptus è un'ottima piattaforma open per le aziende e organizzazioni che vogliono una nuvola all'interno delle loro mura. Basato sulla stessa idea di EC2 con cui ancora non può competere rappresenta comunque una buona alternativa per quelli che vogliono testare il modello o per quelli che hanno intenzione di abbandonare Amazon per far fronte alle proprie spese.

Tutto l'interesse che il Cloud Computing sta suscitando può interessare anche il mondo dell'HPC ed a tal fine è necessario effettuare studi specifici per valutarne i costi ed i benefici, per poter convincere il mercato degli eventuali vantaggi di questa tecnologia.

Ai fini della tesi è stato realizzato un cluster Eucalyptus perfettamente funzionante e si spera che ciò possa tornare utile anche in ambito universitario, in future applicazioni in sinergia con lo Spin Off Uniclust.

Per ora il modello Cloud ha almeno un limite, la scarsità di prestazioni, che in campi in cui la potenza di calcolo è un fattore cruciale, come l'HPC, rappresenta un forte deterrente. Probabilmente il problema è solo economico, o forse è il Cloud Computing a non essere ancora abbastanza maturo e sviluppi futuri potranno portare ad abbracciare questo ed altri settori.

Resta da dire che la sperimentazione è stata fatta con la virtualizzazione data da KVM. Altri hypervisor come Xen o Vmware, così come ulteriori ottimizzazioni e tuning, potrebbero dare risultati migliori e rendere l'HPC Cloud finalmente appetibile dal mercato.

Riferimenti

1. Cloud privato come sinonimo di business: riduzione dei costi e aumento della flessibilità - EMC <http://italy.emc.com>
2. The NIST Definition of Cloud Computing - Peter Mell, Tim Grance
<http://http://csrc.nist.gov/groups/SNS/cloud-computing>
3. Wikipedia - <http://wikipedia.org>
4. A history of cloud computing - -
<http://www.computerweekly.com/Articles/2009/06/10/235429/A-history-of-cloud-computing.htm>
5. Telecom e la Nuvola Italiana - Mauro Vecchio -
<http://punto-informatico.it/2994006/PI/News/telecom-nuvola-italiana.aspx>
6. Cloud services for your virtual infrastructure
<http://www.ibm.com/developerworks/opensource/library/os-cloud-virtual1>
7. The Tenets of the Cloud
<http://www.janakiramm.net/blog/the-tenets-of-the-cloud>
8. Be Careful When You Come to Put Your Trust in the Clouds - Bruce Schneier
<http://www.schneier.com/essay-274.html>
9. Seven cloud-computing security risks - Jon Brodtkin

<http://www.infoworld.com/d/security-central/gartner-seven-cloud-computing-security-risks-853>
10. Cos'è il software libero?
<http://softwarelibero.it/documentazione/softwarelibero.shtml>
11. Cloud computing is a trap, warns GNU founder Richard Stallman
<http://www.guardian.co.uk/technology/2008/sep/29/cloud.computing.richard.stallman>
12. Open Source Cloud Computing: Notes from a Conference - Roberto Galoppini
<http://robertogaloppini.net/2010/06/24/open-source-cloud-computing-notes-from-a-conference/>
13. AGPL: Open Source Licensing in a Networked Age
<http://redmonk.com/sogady/2009/04/15/open-source-licensing-in-a-networked-age/>
14. Google CEO's new paradigm: 'cloud computing and advertising go hand-in-hand'
<http://www.zdnet.com/blog/micro-markets/google-ceos-new-paradigm-cloud-computing-and-advertising-go-hand-in-hand/369>

15. Cloud Computing: l'ultimo trend di Internet
<http://www.7thfloor.it/2007/10/04/cloud-computing-lultimo-trend-di-internet/>
16. Above the Clouds: A Berkeley View of Cloud Computing - Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, Andy Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, and Matei Zaharia - 2009
<http://radlab.cs.berkeley.edu/>
17. Il cloud computing di Amazon Web Services
<http://segnalazionit.org/2009/11/aws-amazon-web-services/>
18. Will Public Clouds Ever Be Suitable for HPC? - Nicole Hemsoth
<http://www.hpcinthecloud.com/features/Will-Public-Clouds-Ever-Be-Suitable-for-HPC-97269804.html?viewAll=y>
19. High Performance Cloud Computing Still an Oxymoron - Michael Feldman
http://www.hpcwire.com/specialfeatures/isc09/blogs/High_Performance_Cloud_Computing_Still_an_Oxymoron-49108931.html
- 20: The Performance cost of virtual machines in Computer Clusters - Neal Barcelo, Nick Legg
<http://www.denison.edu/academics/departments/mathcs/>
21. Amazon Adds HPC Capability to EC2 - Nicole Hemsoth
<http://www.hpcinthecloud.com/features/Amazon-Adds-HPC-Capability-to-EC2-98349769.html>
22. Can Cloud Computing Reach The TOP500? - Jeffrey Napper, Paolo Bientinesi
<http://userweb.cs.utexas.edu>