Alma Mater Studiorum · Università Di Bologna

Campus Di Cesena

SCUOLA DI INGEGNERIA E ARCHITETTURA
Corso di Laurea in Ingegneria Elettronica, Informatica e Telecomunicazioni

Tesi di Laurea in Sistemi Operativi

RUOLO DEL CLOUD NELL'AMMINISTRAZIONE DEI SISTEMI INFORMATICI MODERNI

Relatore:
Prof.
ALESSANDRO RICCI

Presentata da:
MARCO LORENZINI

Sessione III
Anno Accademico 2013/2014

Indice

Elenco delle figure

Elenco delle tabelle

Elenco delle abbreviazioni

AH	Authentication Header
AWS	Amazon Web Services
CEO	Chief Executive Officer
CEP	Complex Event Processing
DaaS	Data as a Service
DNS	Domain Name System
DPI	Deep Packet Inspection
DRS	Disaster Recovery System
ESP	Encapsulating Security Payload
HaaS	Hardware as a Service
IaaS	Infrastructure as a Service
IDS	Intrusion Detection System
IT	Information Technology
ITO	Information Technology Outsourcing
LSB	Least Significant Bit
MD5	Message Digest
MDM	Master Data Management
PaaS	Platform as a Service
RAID	Redundant Array of Independent Disks
SaaS	Software as a Service
SHA	Secure Hash Algorithm
SSO	Single Sign-On
STaaS	Storage as a Service
VPN	Virtual Private Network

Abstract

Per far fronte alla necessità di sviluppo incrementale ed evolutivo dell'IT, è necessario un continuo rinnovamento del sistema aziendale. Negli ultimi anni, il progresso informatico è stato caratterizzato dall'espansione delle tecnologie Cloud, in grado di semplificare la complessità amministrativa delle aziende, aumentare la scalabilità dei sistemi e ridurre i costi di produzione. Nel corso di questo elaborato analizzerò nel dettaglio come il Cloud sia capace di rivoluzionare, in primis dal punto di vista amministrativo, ma non solo, i sistemi informatici moderni, valutando i molteplici benefici da esso generati e gli inevitabili trade-off che ne derivano.

Keywords: Cloud Computing, administration.

Capitolo 1

Il Cloud Computing

Il Cloud Computing è un paradigma informatico che permette a un cliente di avvalersi di risorse hardware e software fornite da un provider, sfruttando un dispositivo connesso alla rete. Un modello Cloud è fondamentalmente basato su alcuni concetti primari [4]:

- Il provider investe in infrastrutture informatiche che saranno in seguito "affittate" da uno o più clienti, su richiesta, seguendo la classica architettura client-server.
- L'unico requisito tecnologico necessario all'utilizzo delle risorse distribuite è l'accesso alla rete Internet tramite un dispositivo informatico.
- La possibilità di "affittare" una risorsa informatica genera scenari ad elevata scalabilità, a seconda delle esigenze di ogni singolo cliente.

1.1 Classificazioni e servizi

Le tecnologie Cloud sono classificate secondo tre dimensioni:

- *Distribuzione*: una suddivisione dei servizi Cloud può essere effettuata in relazione ai privilegi di accesso forniti. In questo caso, possiamo definire i seguenti concetti [7]:

Figura 1.1: Cloud Computing

  - *Public Cloud*: rete posseduta e gestita da una compagnia che la utilizza per offrire ad altre aziende o a singoli individui un rapido accesso a risorse informatiche. Grazie alle Public Cloud, gli utenti non devono necessariamente acquistare risorse o un'apposita infrastruttura pur potendo usufruirne liberamente.
  - *Private Cloud*: rete posseduta e gestita da una compagnia che controlla le modalità in base alle quali le risorse virtualizzate e distribuite possono essere utilizzate da specifici gruppi di utenti.
  - *Hybrid Cloud*: rete che risulta dalla combinazione di una Private Cloud con una Public Cloud. Uno scenario comune prevede che l'azienda sfrutti al contempo la propria rete Cloud privata e un servizio pubblico che consenta di gestire il carico extra nei periodi di maggiore intensità operativa (*Cloud Bursting*). In questo modo, è possibile pagare risorse supplementari solo quando queste sono effettivamente necessarie. Questa strategia è risultata estremamente utile a colossi informatici quali eBay e Amazon, aziende di e-commerce che in alcuni periodi dell'anno (es: festività) registrano carichi di lavoro estremamente più elevati di quelli registrati in altri periodi dell'anno (Il caso Amazon a p. 9 per ulteriori approfondimenti).

- *Servizio*: a seconda del tipo di risorse informatiche che il provider mette a disposizione, si possono distinguere alcune principali tipologie di servizi Cloud:

  - *DaaS* (Data as a Service): si fornisce l'accesso a dati memorizzati su dispositivi remoti, al quale l'utente può accedere pubblicamente oppure attraverso autenticazione.
  - *SaaS* (Software as a Service): si fornisce l'accesso a risorse software localizzate su infrastrutture remote. Il cliente è quindi in grado di utilizzare un programma informatico senza tuttavia possederne

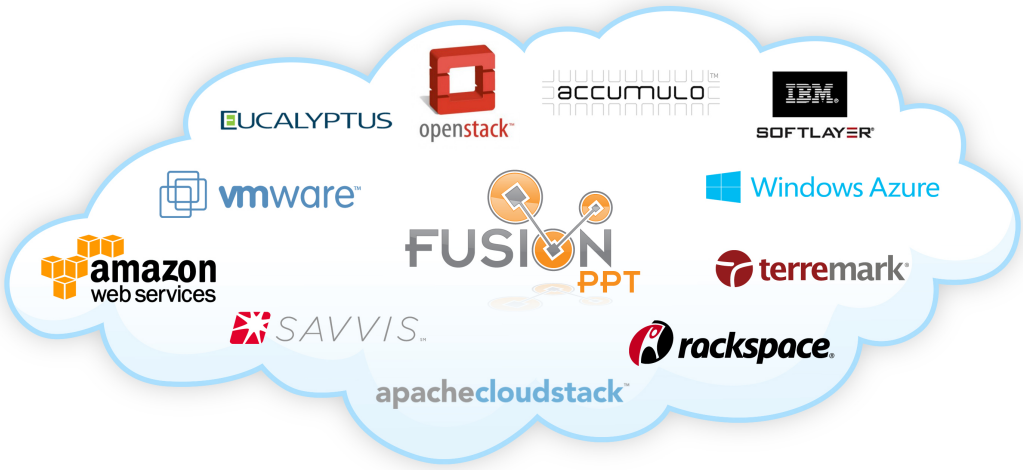

Figura 1.2: Principali provider Cloud a livello mondiale

una copia o una licenza d'uso completa: in questo caso è il software stesso ad essere concesso in "affitto" al cliente [2].

– *HaaS* (Hardware as a Service): si fornisce l'accesso a risorse hardware remote. L'utente ha dunque la possibilità di sfruttare la capacità computazionale delle risorse distribuite per l'elaborazione dei propri dati [3]. Una particolare sotto-categoria del servizio HaaS è il tipo *STaaS* (Storage as a Service), che permette all'utente di effettuare l'upload di dati su dispositivi remoti. Dal punto di vista tecnico, ciò significa che l'utente sfrutta un disco rigido di memoria, e per questo il servizio è riconducibile alla tipologia principale legata alla concessione di risorse hardware.

– *IaaS* (Infrastructure as a Service): rappresenta un'evoluzione del servizio HaaS, in cui le risorse fornite come servizio, anziché essere singole componenti hardware, sono centri di dati, ovvero edifici che contengono infrastrutture informatiche pronte all'uso.

Figura 1.3: Esempio di Private Cloud

  - *PaaS* (Platform as a Service): può essere considerato come implementazione del servizio IaaS, ed è il risultato della fusione di risorse hardware con quelle software. Tramite PaaS, infatti, il provider fornisce un centro di dati completo, in cui, oltre all'infrastruttura hardware, è presente anche una serie di software dedicati (es: sistemi operativi, applicazioni).

- *Carico di lavoro*: una terza ed ultima classificazione dei servizi Cloud avviene in merito alle tipologie di operazioni che l'azienda vuole effettuare. I tipici carichi di lavoro gestiti tramite Cloud sono [13]:

  - *Online Transaction Processing (OLTP)*: insieme di tecniche che permettono l'esecuzione delle funzioni fondamentali dell'azienda, facilitando la gestione delle applicazioni orientate alle transazioni. Relativamente ai sistemi informativi, ad esempio nel caso di inserimento di dati nel database e relativo recupero, gli aggiornamenti sono rapidi e caratterizzati da semplici query.

  - *Online Analytical Processing (OLAP)*: insieme di tecniche per l'analisi di grandi quantità di dati, allo scopo di facilitare la pianificazione e le strategie decisionali dell'azienda. Gli aggiornamenti sono periodici ma di lunga durata, con query complesse per l'interrogazione del database.

  - *Complex event processing (CEP)*: insieme di tecniche di analisi delle informazioni relative agli eventi, allo scopo di trarre conclusioni e ottimizzare le performance del sistema. Il CEP è affine ai metodi di Data Mining, l'insieme di pratiche che analizzano i sistemi e i database informativi allo scopo di estrarre modelli, schemi e tendenze, segmentando il set di informazioni e valutando la probabilità di eventi futuri sulla base di algoritmi matematici [14].

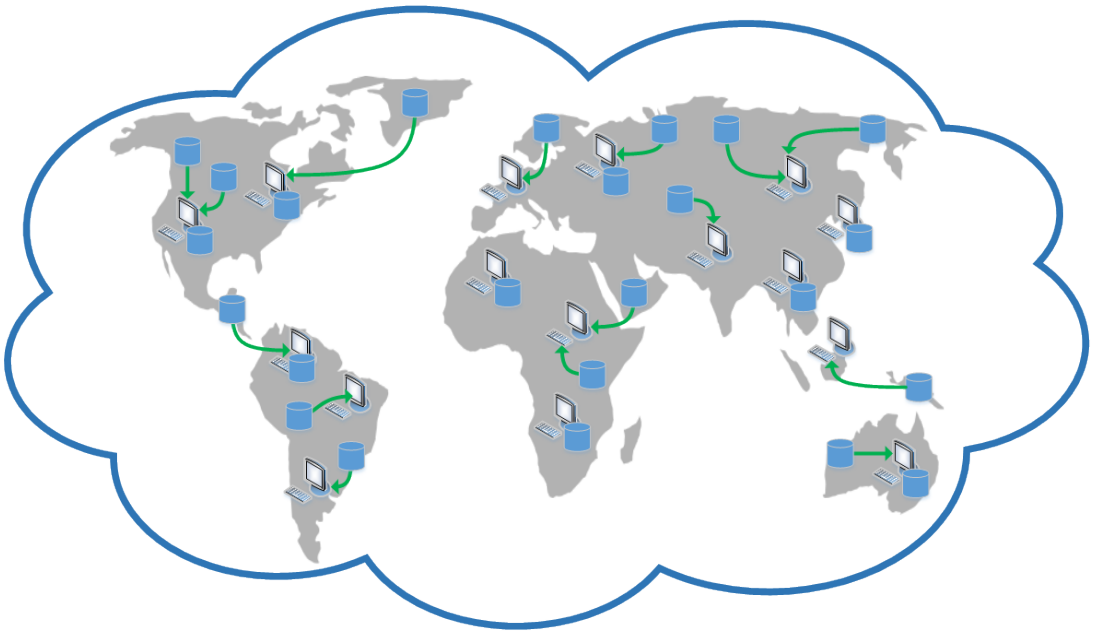

Figura 1.4: Scalabilità geografica del Cloud Computing

1.2 Vantaggi e svantaggi

Analizziamo ora pro e contro dell'utilizzo del Cloud Computing.

- Innanzitutto, i sistemi Cloud offrono una ottima scalabilità, ossia la capacità di un sistema di crescere o diminuire di scala in funzione delle necessità e delle disponibilità [6]. Relativamente ai sistemi distribuiti, la scalabilità è principalmente suddivisa in:
  - *scalabilità di carico*: aumentare o diminuire le prestazioni del sistema in funzione della potenza di calcolo necessaria alla sua esecuzione. Relativamente ai sistemi distribuiti, tale concetto è principalmente applicato al numero di processori disponibili, piuttosto che alla potenza dei singoli processori;
  - *scalabilità geografica*: mantenere inalterate la disponibilità e l'utilizzabilità del sistema indipendentemente dalla distanza geografica delle risorse e degli utenti;

- Il paradigma Cloud permette inoltre la realizzazione del concetto "pay-as-you-use" [7] , con conseguente riduzione dei costi fissi, dato che si sostituiscono gli investimenti iniziali relativi a grandi infrastrutture informatiche con costi operazionali proporzionali all'utilizzo delle risorse distribuite necessarie [4] [5].

- Si possono ottenere elevati livelli di tolleranza agli errori: grazie alla condivisione delle informazioni, se uno dei server utilizzati dal sistema subisce un guasto, altri server saranno in grado di continuare a garantire il funzionamento del sistema senza che questo venga compromesso [4].

- Rapida predisposizione del sistema, che può essere pienamente operativo in pochi minuti.

- Minor necessità di possedere conoscenze specializzate all'interno dell'azienda, con conseguente possibilità di concentrarsi sulle funzioni primarie.

Valutiamo ora gli aspetti negativi del Cloud Computing:

- per prima cosa, si ha l'inevitabile perdita di controllo dovuta all'esternalizzazione di una o più funzioni informatiche a terze parti. Il rischio più grande, in questo caso, è che il collegamento con il fornitore si interrompa per un guasto nelle comunicazioni, compromettendo l'operatività del sistema. Persino i migliori fornitori di servizi Cloud offrono una percentuale di *uptime* (operatività) variabile tra il 99% e il 99.9%, indicando che l'immunità assoluta da problemi tecnici non è raggiungibile. Alcune aziende scelgono di ridurre le possibilità di incorrere in uno scenario simile mantenendo parte del sistema informatico interno all'azienda (es: le funzioni primarie) ed affidando a terzi, tramite Cloud, le funzioni secondarie o di supporto.

- Il fatto che uno stesso provider gestisca il traffico di molteplici clienti implica numerosi rischi per la sicurezza delle informazioni. Anche

adoperando opportune misure di sicurezza, vi è una possibilità che i server del provider vengano infettati a causa della mancanza di attenzione da parte di un altro cliente, il quale, caricando dati e applicazioni mediante canali di comunicazione poco affidabili, mette a repentaglio la sicurezza delle informazioni di tutti gli altri utenti che si avvalgono dei servizi dello stesso provider.

1.3 Il caso Amazon

Amazon, primo sito di commercio elettronico, nacque nel 1995, quando Jeff Bezos (fondatore e attuale CEO) vide nella crescita esponenziale di internet un'eccellente opportunità di business [27]. L'idea si rivelò assolutamente vincente, tuttavia l'enorme espansione di Amazon mise in evidenza i limiti delle infrastrutture tecnologiche del momento, basate su sistemi monolitici e poco scalabili. Fu proprio la necessità di creare una struttura più performante e scalabile che portò al frazionamento del sistema in numerose sottoparti, allo scopo di rendere i sottoservizi il più possibile autonomi e minimizzare le interdipendenze (Descartes, nel suo *Discorso sul Metodo*, esponeva la necessità di "dividere ognuna delle difficoltà sotto esame nel maggior numero di parti possibile"). La globalità del sistema fu quindi scomposta in più sottosistemi, e successivamente si arrivò all'idea che consacrò Amazon quale principale colosso del Cloud Computing a livello mondiale: si pensò di non usare le infrastrutture create ad uso esclusivo interno, relativamente alle vendite online, bensì anche di condividere tali infrastrutture con chiunque avesse una temporanea necessità di aumentare in modo scalabile la propria potenza di calcolo, senza tuttavia potersi permettere un investimento considerevole per la creazione di un'infrastruttura privata. Furono quindi inizializzati gli Amazon Web Services (AWS), che ottennero un enorme successo grazie allo sfruttamento del paradigma "pay for use" (p. 8). Successivamente, Amazon consolidò il proprio valore nell'ambito dell'IT grazie all'implementazione dei Disaster Recovery Systems (p. 27), anch'essi disponibili in qualità di servizi

Tabella 1.1: Amazon AWS EC2 Pricing: General Purpose [31]

Modello	CPU	Memoria (GiB)	Capacità (GB)	Prezzo
t2.micro	1	1	/	$0.013/h
t2.small	1	2	/	$0.026/h
t2.medium	2	4	/	$0.052/h
m3.medium	1	3.75	1 x 4 SSD	$0.070/h
m3.large	2	7.5	1 x 32 SSD	$0.140/h
m3.xlarge	4	15	2 x 40 SSD	$0.280/h
m3.2xlarge	8	30	2 x 80 SSD	$0.560/h

Tabella 1.2: Amazon AWS EC2 Pricing: Compute Optimized [31]

Modello	CPU	Memoria (GiB)	Capacità (GB)	Prezzo
c4.large	2	3.75	/	$0.116/h
c4.2xlarge	8	15	/	$0.464/h
c4.8xlarge	36	60	/	$1.856/h
c3.xlarge	4	7.5	2 x 40 SSD	$0.210/h
c3.2xlarge	8	15	2 x 80 SSD	$0.420/h
c3.8xlarge	32	60	2 x 320 SSD	$1.680/h

distribuiti. Attualmente, Amazon è quotato in borsa con una capitalizzazione di circa 177 miliardi di dollari [28].

Il concetto del "pay-as-you-use", già introdotto a p. 8, pone alcuni interrogativi sui costi della tecnologia: quanto costa, esattamente, usufruire dei servizi distribuiti di un Cloud provider? Amazon offre diverse categorie di server on-demand, a partire dai "general purpose", fino ad arrivare a quelli specializzati in capacità computazionale, RAM o spazio di memorizzazione.
I prezzi variano a seconda della nazione in cui si trovano i server. Vi è quindi una lieve differenza se le risorse distribuite di cui si usufruisce si trovano in California oppure in Irlanda.

In alternativa al servizio on-demand, è possibile ottenere l'esclusività di risorse hardware o di intere infrastrutture informatiche sottoscrivendo un

Tabella 1.3: Amazon AWS EC2 Pricing: Memory Optimized [31]

Modello	CPU	Memoria (GiB)	Capacità (GB)	Prezzo
r3.large	2	15	1 x 32 SSD	$0.175/h
r3.xlarge	4	30.5	1 x 80 SSD	$0.350/h
r3.2xlarge	8	61	1 x 160 SSD	$0.700/h
r3.4xlarge	16	122	1 x 320 SSD	$1.400/h
r3.8xlarge	32	244	2 x 320 SSD	$2.800/h

Tabella 1.4: Amazon AWS EC2 Pricing: Storage Optimized [31]

Modello	CPU	Memoria (GiB)	Capacità (GB)	Prezzo
i2.xlarge	4	30.5	1 x 800 SSD	$0.853/h
i2.2xlarge	8	61	2 x 800 SSD	$1.705/h
i2.4xlarge	16	122	4 x 800 SSD	$3.410/h
i2.8xlarge	32	244	8 x 800 SSD	$6.820/h
hs1.8xlarge	16	117	24 x 2048	$4.600/h

contratto annuale o triennale. Questa modalità permette di risparmiare dal 30 al 40% rispetto ai consumi on-demand, sebbene non consenta gli estremi livelli di scalabilità dell'altra modalità.

Capitolo 2

Adattare le applicazioni all'ambiente Cloud

Adattare le applicazioni aziendali al nuovo sistema Cloud significa mantenere i benefici derivanti dagli investimenti passati e, allo stesso tempo, sostenere l'innovazione tecnologica [5]. Tale operazione richiede che le applicazioni da trasferire siano reimplementate secondo un nuovo paradigma informatico, e che sia effettuata un'accurata analisi dei rischi, con particolare riferimento ai costi di migrazione, reingegnerizzazione dell'applicazione, sicurezza e regolamentazione [8]. Sorgono quindi alcune domande: quali parti dell'applicazione è opportuno migrare, e quali invece possono essere lasciate nel sistema locale? In che modo effettuare l'adattamento delle applicazioni a questo ambiente distribuito? Sarà effettivamente una scelta efficace quella di trasferire l'intera applicazione sul Cloud, in termini di costi e benefici? Per rispondere a queste domande, è opportuno conoscere innanzitutto le modalità tramite le quali è possibile adattare un'applicazione ad un sistema Cloud [5]:

- Sostituire le risorse hardware attuali con quelle distribuite offerte da un provider Cloud, mantenendo la logica applicativa corrente.
- Trasferire alcune delle funzionalità applicative nel Cloud, eseguendo una parziale reimplementazione del software, e mantenendo parte della

logica applicativa corrente.

- Incapsulare l'applicazione in una macchina virtuale da eseguire sul Cloud. La virtualizzazione è un'astrazione delle risorse hardware che permette di realizzare un emulatore in grado di eseguire applicazioni e svolgere le medesime funzioni di un elaboratore fisico vero e proprio. Uno dei principali vantaggi della virtualizzazione consiste nella condivisione delle risorse, la quale permette il funzionamento di più sistemi in grado di sfruttare le medesime risorse.

- Trasformare completamente l'applicazione e renderla pienamente operativa nel Cloud: le funzionalità dell'applicazione vengono reimplementate come composizione di servizi in esecuzione sul Cloud. E' la soluzione più efficace, tuttavia la re-ingegnerizzazione completa dell'applicazione richiede un impegno considerevole.

Dato che fallire nel tentativo di trasferire il proprio sistema in ambiente Cloud implica possibili aumenti di costi e perdita di business, eliminando ogni potenziale beneficio del Cloud Computing, è opportuno seguire un programma che scandisca le fasi di migrazione e permetta di valutare attentamente ogni scelta. Tale programma può essere così articolato [8]:

- *Valutazione delle applicazioni e del carico di lavoro*: si individuano le applicazioni più adatte ad essere trasferite in ambiente distribuito. Spesso è opportuno iniziare con le applicazioni che hanno minore impatto sul sistema, e che pertanto comportano rischi minori, prima di passare alla migrazione di applicazioni del core business aziendale. Va tenuto in considerazione il ciclo di vita dell'applicazione: se un software sta per divenire obsoleto, è altamente sconsigliato il suo trasferimento su Cloud, ed è preferibile investire risorse in una nuova applicazione appositamente sviluppata per l'ambiente distribuito. Anche l'architettura dell'applicazione svolge un ruolo determinante nella scelta della migrazione: un'elevata modularità permette il trasferimento di determinate

parti del programma; alternativamente, il software deve essere trasferito globalmente sulla nuvola. Infine, non può mancare l'analisi relativa alla sicurezza: per ottenere i migliori risultati, le applicazioni trasferite sulla nuvola informatica devono essere opportunamente implementate affinché siano compatibili con i metodi utilizzati dal provider per garantire protezione al sistema. Ad esempio, se il provider supporta efficaci protocolli di comunicazione criptata, ma l'applicazione dell'utente non è in grado di sfruttarli oppure è danneggiata, e scambia informazioni utilizzando canali non sicuri, la sicurezza del sistema virtuale sarà comunque compromessa.

- *Realizzazione dello studio di fattibilità*: si esegue uno studio di fattibilità che analizzi accuratamente i costi del progetto, i livelli del servizio attesi e l'impatto sul business aziendale. Relativamente all'analisi dei costi, si prendono in considerazione:

  - i costi del *servizio fornito*, valutando possibili variabili dovute, ad esempio, a tasse extra richieste per la gestione dei picchi di lavoro;
  - i costi per la *gestione del servizio*, più che altro in termine di abilità necessarie (da qui, la necessità di valutare i costi necessari allo sviluppo di conoscenze approfondite del Cloud tramite il potenziamento delle risorse umane);
  - i costi di *reingegnerizzazione dell'applicazione*;
  - i costi di i*ntegrazione con il sistema locale*: l'applicazione trasferita su Cloud deve essere resa compatibile con le applicazioni che l'azienda sceglie di mantenere all'interno del sistema locale.

  Dopo aver valutato con attenzione tutte le variabili di costo, è necessario assicurarsi altresì che il livello di servizio offerto dal provider sia compatibile con quello attuale, per evitare una riduzione delle performance aziendali. Tale verifica avviene attraverso la stipulazione di appositi accordi sulla qualità del servizio, che permettono di valutare,

per ogni applicazione da trasferire su Cloud, la disponibilità del servizio, le performance, la sicurezza, la privacy e molti altri dettagli (*SLA* a p. 24).

Infine, si valuta l'impatto atteso dell'investimento sul business aziendale: lo spostamento dell'applicazione su Cloud ridurrà i tempi di consegna dei servizi agli utenti finali? Permetterà di garantire una disponibilità del sistema più elevata, aumentando così il grado di soddisfazione degli utenti?

- *Definizione dell'approccio tecnico*: sulla base di quanto indicato nella valutazione del carico di lavoro e nello studio di fattibilità, si specifica una soluzione tecnica scelta per la realizzazione del progetto, indicando nel dettaglio tempi, modi e realizzazione. In questo caso gli elementi chiave da valutare sono:
  - le *abilità tecniche* che l'azienda possiede in materia di migrazione verso i sistemi Cloud: a seconda del tipo di migrazione scelta, e del grado di responsabilità che avrà il provider del servizio, saranno necessarie più o meno competenze specifiche. La mancanza di determinate conoscenze all'interno dell'azienda può determinare la necessità di investire ulteriormente sulle risorse umane;
  - la *sicurezza* e il grado di protezione dei dati inviati al provider Cloud (crittografia a p. 31)
  - la *scalabilità*, uno dei principali vantaggi del Cloud Computing: le applicazioni devono essere ristrutturate appositamente da permettere di sfruttare l'estensibilità e la scalabilità del sistema, e questo può richiedere modifiche al codice sorgente applicativo. In particolare, la riprogrammazione di un'applicazione allo scopo di permettere l'utilizzo di processori o elaboratori multipli, in modo parallelo, può risultare dispendiosa in termini di tempo, specialmente se non si dispone di una documentazione accurata relativa al codice originale dell'applicazione stessa.

Solitamente, questa fase di analisi si suddivide in:

  - *Segmentazione del progetto*: suddivisione del progetto in moduli, ognuno dei quali è finalizzato al completamento di un sottosistema specifico.
  - *Riepilogo delle acquisizioni e realizzazioni previste*: sintetizza le acquisizioni previste (strumenti hardware e software, e più in generale risorse informatiche utili alla realizzazione del progetto).
  - *Piano del progetto*: esplicita la sequenza di fasi di sviluppo e le dipendenze tra le principali attività del progetto. Esistono numerose rappresentazioni grafiche che delineano il modello di un progetto. Fra queste, si ha la WBS (Work Breakdown Structure), che suddivide le attività in livelli, e consente un'analisi di tutti i processi coinvolti nel progetto, il diagramma di Gantt, una tabella in cui le righe indicano le attività previste dalla WBS, e le colonne sono utilizzate per indicare i tempi necessari alla loro realizzazione (utile per pianificare le tempistiche, verificare la fattibilità temporale del progetto e permettere a tutti gli interpreti del progetto di avere un quadro generale delle tempistiche stimate), e il diagramma di PERT, che modella anche l'interdipendenza tra le attività, e serve ad esprimere vincoli temporali e forme di sincronizzazione (es: l'attività B non può iniziare finché non sarà ultimata l'attività A).

- *Adottare un modello di integrazione flessibile*: tipicamente le applicazioni trasferite su Cloud presentano numerose connessioni con altre applicazioni del sistema. L'integrazione di un'applicazione con la globalità del sistema consiste principalmente in:

  - Integrazione di processo, quando un'applicazione ne invoca una seconda per eseguire una determinata operazione;

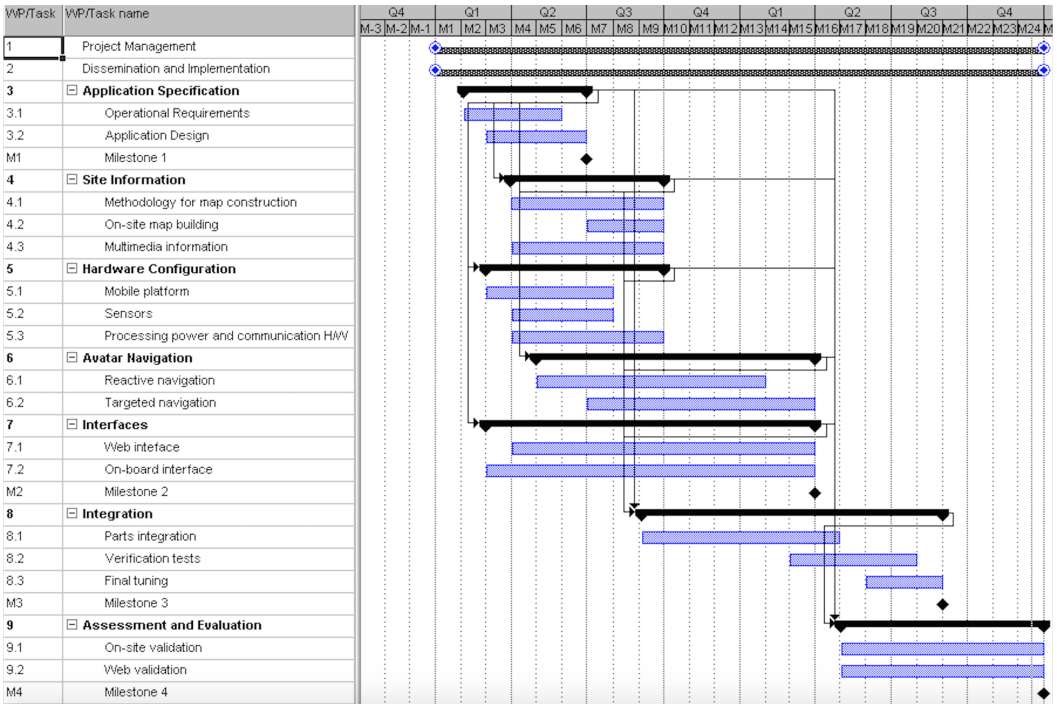

Figura 2.1: Diagramma di GANTT

– integrazione dei dati, quando le applicazioni condividono dati, o quando l'output di un'applicazione diventa l'input di una seconda applicazione.

E' pertanto fondamentale che l'applicazione sia flessibile, in grado di gestire differenti protocolli operativi o di comunicazione a seconda delle specifiche situazioni, e che sia costruita sfruttando regole standard, allo scopo di ottimizzare la sua integrazione con applicazioni future.

Infine, se l'azienda ha implementato un sistema MDM[1] per gestire i

[1]Master Data Management (MDM): Il MDM è l'insieme di discipline, tecnologie e soluzioni in grado di creare, e mantenere consistenti, aggiornati e completi, i dati di importanza critica per l'azienda e di fornirne una visione unica a utenti e processi sia interni sia esterni. Nello specifico, si parla di PIM (Product Information Managment) quando si gestiscono i dati critici di prodotto, mentre si dice CDI (Customer Data Integration) il sistema che gestisce i dati critici relativi a clienti, fornitori e utenti. La creazione di un'apposita architettura per la gestione dei MD è giustificata solo a fronte di quantità

dati critici, è probabile che l'applicazione trasportata su Cloud invochi frequentemente il sistema per ottenere dati di cui non dispone, quindi è importante che l'applicazione, sebbene abbia aderito alle specifiche Cloud del provider, sia ancora in grado di comunicare con un sistema esterno.

- *Specificare i requisiti di sicurezza e privacy*: sicurezza e privacy sono due fra i problemi principali relativi al Cloud Computing. A seconda delle caratteristiche dell'applicazione e dalla criticità delle informazioni gestite, il sistema (o sottosistema) può essere sbilanciato a favore di elevate performance, oppure di solida protezione dei dati (Security Management a p. 30).

- *Amministrare la migrazione*: dopo aver dettagliatamente definito i motivi alla base del progetto, e le modalità in cui si intende operare, si dà il via alla migrazione vera e propria, che può essere così articolata:

  - *Dislocazione dell'ambiente Cloud:* si installano i componenti necessari a costituire il sistema Cloud nel quale sarà eseguita l'applicazione a migrazione conclusa.
  - *Installazione e configurazione delle applicazioni*: si installano le applicazioni di supporto sui server Cloud.
  - *Rafforzamento del sistema*: si installano strumenti supplementari per la continuità del business aziendale e la sicurezza del sistema. Alcuni di questi servizi sono tipicamente forniti dal provider, ma anche in questo caso è opportuno che siano testati.
  - *Eseguire una migrazione di prova*: si effettua una prova di migrazione allo scopo di rivelare risultati e problemi inattesi.

elevate di informazioni, e maggiore è il valore di una informazione, maggiore la probabilità che sia considerato un MD. Le caratteristiche tipiche di un MD sono la riusabilità (dunque informazioni riutilizzate da più sistemi aziendali e in più progetti) e la centralità (dati alla base di molteplici applicazioni).

– *Esecuzione finale della migrazione*: si trasferisce l'applicazione nell'ambiente Cloud predisposto a supportarla.

Capitolo 3

Come incide il Cloud Computing sulla funzione IT?

Dato che, con l'utilizzo di un sistema Cloud, la funzione IT diventa sostanzialmente un'unità di gestione, piuttosto che di servizio operativo, è utile disporre di alcune figure professionali che amministrino l'integrazione dei sistemi esterni e gestiscano il progetto aziendale, anziché operatori e tecnici poco specializzati in funzioni gestionali, e in particolare: un Project Manager[1], un Data Administrator[2] e un Security Manager[3]. Nei seguenti capitoli analizzerò nel dettaglio queste figure professionali, evidenziando in che modo i rispettivi compiti vengono influenzati dal Cloud Computing.

3.1 Project Management

Il project management è l'applicazione di conoscenze, attitudini, tecniche e strumenti alle attività di un progetto al fine di conseguirne gli obiettivi [9]. L'obiettivo di un project manager è il raggiungimento degli obiettivi di

[1]Project Manager: responsabile della valutazione, pianificazione, realizzazione e controllo di un progetto.

[2]Data Administrator: responsabile della pianificazione, dell'organizzazione e del controllo delle risorse informative.

[3]Security Manager: responsabile dell'organizzazione e della gestione della sicurezza.

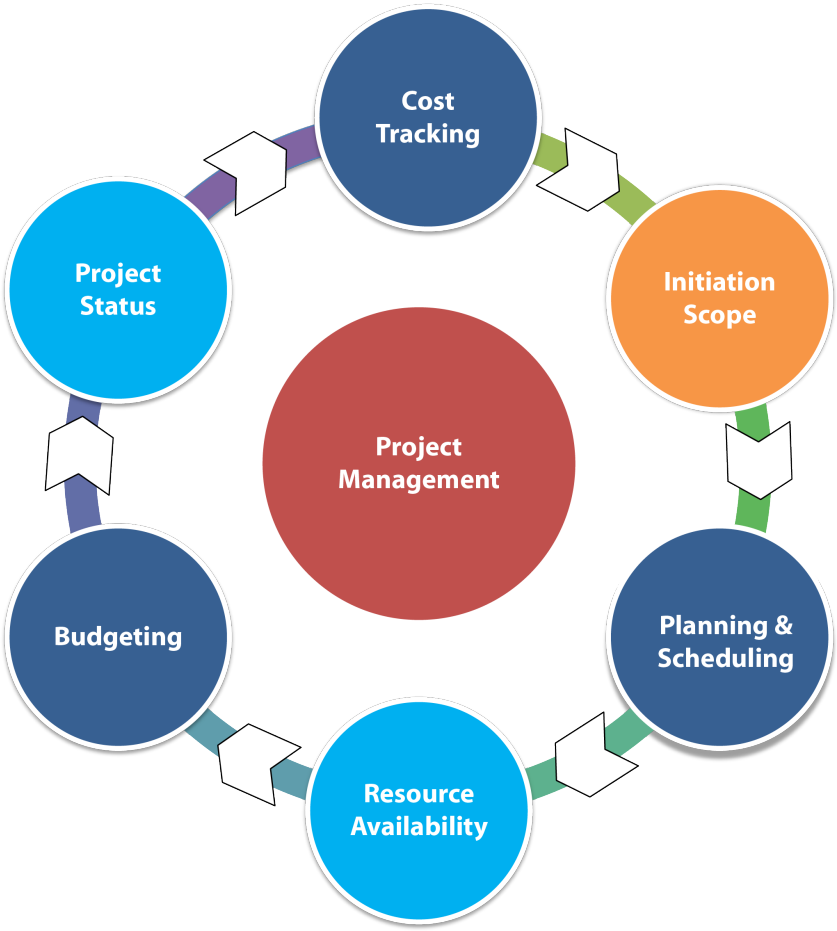

Figura 3.1: Project Management

progetto, assicurando il rispetto dei costi, dei tempi e della qualità concordati, e infine il raggiungimento della soddisfazione del cliente [10]. L'area di lavoro di un project manager si suddivide in:

- area operativa, legata alle metodologie e agli strumenti che permettono nel miglior modo il raggiungimento degli obiettivi;

- area relazionale: legata alle dinamiche di comunicazione che facilitano la collaborazione con gli altri reparti aziendali.

Il ruolo del project manager richiede elevate competenze gestionali ed amministrative, ma non necessariamente una consolidata conoscenza dei dettagli tecnici e operativi dei progetti.

Il ciclo di gestione di un progetto si articola in:

- Fase di iniziazione (*Initiation Scope*): si definiscono gli obiettivi del progetto, si stabiliscono i ruoli e si realizza lo studio di fattibilità.

- Pianificazione e programmazione (*Planning & Scheduling*): in fase di pianificazione, si determinano tutte le sottoattività del progetto necessarie al raggiungimento degli obiettivi prefissi, tenendo in considerazione anche vincoli di sincronizzazione. Lo scheduling è invece relativo alla definizione delle risorse necessarie allo svolgimento delle sottoattività sopracitate.

- Disponibilità delle risorse (*Resource Availability*): sulla base delle informazioni indicate in fase di scheduling, si verifica la disponibilità delle risorse necessarie.

- Controllo del bilancio (*Budgeting*): qualora una o più risorse non dovessero essere disponibili, si effettua una valutazione del budget a disposizione per il progetto, quindi si determina il modo più opportuno per reperire le risorse mancanti.

- Stato del progetto (*Project Status*): si definiscono template informativi che verranno utilizzati, a progetto avviato, per effettuare controlli periodici relativi allo stato di avanzamento dei compiti e all'insorgere di eventuali problemi.

- Stima finale dei costi (*Cost Tracking*): si identificano e quantificano i principali fattori di costo del progetto, effettuando una stima dei costi totali del progetto.

3.1.1 In che modo il Cloud Computing influisce sul compito di un Project Manager?

Uno degli aspetti di project management ad essere maggiormente influenzato dal Cloud Computing è il controllo degli standard di qualità. Affidando una o più funzioni informatiche ad un provider Cloud, la qualità del sistema aziendale non dipende esclusivamente dal settore IT interno, bensì è strettamente connessa alla capacità operativa delle terze parti alle quali ci si affida. Per questo motivo, l'utilizzo del paradigma Cloud, in linea con il concetto di esternalizzazione[4], genera la necessità di stipulare un contratto con i fornitori, nel quale si specificano i livelli di qualità che il provider deve garantire al cliente nell'erogazione del servizio. Tali accordi sono conosciuti come *SLA (Service Level Agreements).* Il livello di qualità del servizio è definito QoS (Quality of Service), *"the collective effect of service performance which determines the degree of satisfaction of a user of the service"* [15]. Il termine QoS è stato utilizzato per la prima volta nell'ambito delle reti informatiche, allo scopo di definire un set di caratteristiche relative alla performance della rete. Le principali caratteristiche relative al QoS sono: ritardo, percentuale di pacchetti persi, jitter[5], disponibilità del provider tramite la rete Internet, percentuale di errori nell'invio di dati e tempo di disservizio in seguito a un errore.

Alcuni provider Cloud offrono veri e propri servizi distribuiti di Project Management, semplificando enormemente il compito di un Project Manager attraverso applicazioni *Cloud-based* che permettono, con elevata facilità, di gestire efficacemente tutte le fasi del ciclo di vita del progetto ed eseguire accurate analisi dei dati, semplificando così la scelta delle decisioni. Un

[4]Esternalizzazione (o outsourcing): affidamento di una o più funzioni aziendali a terze parti, presupponendo una certa forma di stabilità nel rapporto di collaborazione tra l'impresa soggetto e le terze parti, con una prospettiva di medio-lungo termine. In questo caso trattasi di Information Technology Outsourcing (ITO), ossia esternalizzazione delle attività di sviluppo, produzione e/o manutenzione delle risorse informatiche.

[5]Jitter: variazione di una o più caratteristiche di un segnale (es: l'ampiezza).

MANAGED WEB HOSTING SERVICE LEVEL AGREEMENT (SLA)

THIS AGREEMENT is made this <<CurrentDay>> day of <<CurrentMonth>>, <<CurrentYear>> by and between <<Company>> ("Company") and <<CustCompany>> ("Customer").

The purpose of this Agreement (hereafter referred to as the "Agreement") is to set forth a detailed Service Level Agreement ("SLA") under which Company will provide a service to <<CustCompany>> in order to ensure the reliability and stability of all Web Hosting Services covered under this SLA.

Agreements

In consideration of the mutual covenants set forth in this Agreement, Customer and Company hereby agree as follows:

As a service, the standard Managed Hosting Service Level Agreement (SLA) with the Company is provided below.

1. Network Availability and Uptime.

<<Company>> guarantees that its Network and Connectivity shall be made available at all times. This 100% guarantee covers the availability of all Internet switches, peering, cabling, hubs, routers, DNS servers, load balancers, centralized servers, network appliances, backup and storage devices, management consoles, gateways and other equipment, now or in the future deemed as a requirement for connecting to the Internet and providing Company's services to Customer.

2. Infrastructure Availability and Uptime.

<<Company>> guarantees that its Infrastructure shall be made available at all times. This 100% guarantee covers the availability of all power requirements, components, HVAC, fire suppression, security systems, UPS/PDU, appliances, power cabling, phone systems and other infrastructure or equipment, now or in the future deemed as a requirement for maintaining the network infrastructure and providing Company's services to Customer. This infrastructure availability and uptime guarantee shall not extend to individual computer power supplies or computers or servers that are shut down due to excessive heat problems.

3. Uptime Guarantee and Customer Credits.

In the even that Customer suffers any "downtime" or lack of network or infrastructure availability, the Customer shall receive a credit on their account subject to the table below. All requests for credit must be made within ten (10) days from the occurrence of the downtime and must be made in writing via a support ticket. All credit requests must be verified by Company staff and credits may take up to thirty (30) days to show up on Customer's bill. Company reserves the right to revoke any credit for downtime issued that is later discovered to have been be caused or attributed to Customer activity or external forces not related to Company network or hardware.

Customer Initials _____

Figura 3.2: Facsimile di un Service Level Agreement

esempio è *Oracle Project Management* [24]. Tale suite migliora anche la comunicazione tra il responsabile del progetto e il suo team, permettendo ad ogni utente di accedere alle informazioni di scheduling del progetto tramite un'interfaccia per dispositivi mobili semplice ed efficace (Oracle PPM Mobile). Ogni dipendente può quindi controllare i lavori assegnati e, in modo attivo, registrare sul database i progressi effettuati e le notifiche da inviare al project manager, in modo che quest'ultimo possa sempre essere aggiornato relativamente allo stato di avanzamento dei lavori.

3.2 Data Administration

Le informazioni aziendali sono tipicamente immagazzinate nei database informativi, definiti per l'appunto come collezioni di dati. Nel caso in cui uno o più settori informatici siano gestiti da terze parti, è opportuno che all'interno dell'azienda si effettuino operazioni di amministrazione dei dati, con particolare riferimento a quelli integrati[6] da fonti esterne. Una delle principali azioni di data administration è la pulizia dei dati, un'operazione che ha lo scopo di eliminare eventuali inconsistenze dei dati, dovute, ad esempio, alla differenza di formato, all'incoerenza tra valori di campi correlati o ad indesiderate ridondanze.

3.2.1 In che modo il Cloud Computing influisce sul compito di un Data Administrator?

Relativamente al Cloud Computing, la possibilità che il formato dei dati integrati o gestiti esternamente non sia compatibile con quello utilizzato all'interno dell'azienda è assai elevata, dato che attualmente non esistono standard universalmente accettati dai venditori di servizi Cloud. Alcune organizzazioni, tuttavia, hanno recentemente iniziato a proporre alcuni stan-

[6]Integrazione (nell'ambito dei Sistemi Informativi): insieme di attività atte a rendere consistenti gli schemi sorgenti dai quali il sistema trae informazioni.

dard da utilizzare come linee guida per lo sviluppo dei servizi distribuiti, pertanto la missione relativa al miglioramento dell'interoperabilità lascia intravedere incoraggianti possibilità. Rimane, tuttavia, la necessità di gestire e controllare le informazioni elaborate dalle terze parti, affinché risultino coerenti con la globalità del sistema aziendale.
Passando agli aspetti positivi, il Cloud Computing facilita il compito del Data Administrator per quanto riguarda i piani di recupero e di ridondanza delle informazioni. Esistono appositi sistemi distribuiti, definiti *Disaster Recovery Systems* (DRS), che assicurano la salvaguardia delle informazioni in situazioni critiche, grazie alla realizzazione di protocolli di ridondanza automatici e procedure di recupero dei dati [11]. E' compito del DRS consentire un servizio continuo e sempre disponibile, anche in caso di disastri naturali quali incendi, allagamenti, o più semplicemente interruzione di corrente e cali di tensione [17]. I DRS si suddividono in:

- Storage Layer DRS: tipicamente basati su specifici dispositivi di memorizzazione; ogni volta che avviene una scrittura sul volume principale, questa viene eseguita anche su uno o più volumi di supporto. In questo caso, è comune l'utilizzo di un RAID, un insieme ridondante di dischi di memorizzazione che sfrutta principi di ridondanza dei dati e di parallelismo in accesso per garantire incrementi di prestazioni e tolleranza ai guasti [25]. Il RAID è una tecnica impiegata per i server che devono gestire grandi volumi di dati. Ovviamente, a seconda di ogni situazione specifica, è necessario valutare il trade-off tra sicurezza dovuta alla ridondanza e appesantimento del sistema in termini di spazio di memorizzazione.
  In alternativa, gli Storage Layer DRS possono essere basati su virtualizzazione, caso in cui è previsto che il sistema di memorizzazione sia interamente virtualizzato.

- Application Layer DRS: basati su tecniche che consentono il salvataggio di applicazioni e database tramite il backup delle operazioni SQL o altri strumenti di recupero software.

Per un'azienda, un disastro naturale, imprevedibile per natura, implica una brusca interruzione delle funzionalità operative, o di parte di esse, con conseguente perdita di entrate e riduzione del valore aziendale. Cisco [30] discute le modalità di attuazione di un buon piano di Disaster Recovery per le imprese.

Un altro vantaggio è garantito dalle Cloud API[7], interfacce il cui scopo è fornire ai clienti una comoda gestione dei servizi e la possibilità di realizzare applicazioni Cloud o effettuare manutenzione delle informazioni del sistema con relativa semplicità. Grazie a tali API, la necessità di realizzare complessi software ad hoc per la gestione del database e, più in generale, del sistema globale, è notevolmente ridotta, massimizzando così sia la semplicità di manutenzione, sia quella di utilizzo da parte dei clienti. I sistemi crescono continuamente nel tempo, e divengono più complessi, sia in termini di funzionalità, sia in termini di gestione dei dati: partire già dall'inizio con un sistema complesso significa partire con un netto svantaggio che può tradursi, prima o poi, in difficoltà gestionali. Kernighan e Plauger [32] scrivevano: "Debugging is twice as hard as writing the code in the first place. Therefore, if you write the code as cleverly as possible, you are, by definition, not smart enough to debug it." La stessa cosa vale per i sistemi distribuiti: ogni ora impiegata a semplificare il design, l'architettura e la struttura dei dati di un sistema Cloud si traduce in un notevole guadagno in termini di tempo e risorse una volta che il sistema sarà operativo [33].

Relativamente alla memorizzazione dei dati, il Cloud Computing è in grado di fornire enormi vantaggi, ma anche dare vita a problemi di consistenza dei dati. Si parte dal presupposto che il modo più semplice di immagazzinare dati in un dispositivo informatico sia quello di inserire tutte le informazioni in una singola macchina. Questo metodo, tuttavia, presenta numerosi problemi di efficienza e sicurezza: il singolo dispositivo esaurisce il suo spazio a disposizione molto rapidamente, ha una limitata potenza di calcolo, il che significa che non si possono eseguire elevate quantità di letture e scrittura

[7]Application Programming Interface (API): insieme di procedure che permettono o semplificano l'esecuzione di determinate funzioni all'interno di un programma [18].

simultanee, e, in caso di guasto del dispositivo, si perde il 100% delle informazioni memorizzate. Con il Cloud Computing, si scompongono i dati in più frammenti, e questi sono memorizzati in più dispositivi. In questo modo, la capacità di memorizzazione è unicamente limitata dal numero di macchine a disposizione; la capacità computazionale aumenta sensibilmente, semplificando i calcoli paralleli, e un singolo guasto ad una delle macchine non comporta la perdita della totalità dei dati memorizzati. Gli aspetti primari relativi alla gestione dei dati, in ambito di Cloud Computing, sono riassunti dal Principio *CAP*, e sono: Consistency, Availability, e Partition Tolerance. Il principio stabilisce che un sistema possa garantire, al più, due di queste caratteristiche, con inevitabile generazione di un trade-off.

- *Consistenza*: indica che tutti i nodi del sistema leggono gli stessi dati, sempre aggiornati. Se un sistema sta eseguendo un aggiornamento, tutti gli utenti, collegandosi alle molteplici repliche del database centrale, vedranno l'aggiornamento nello stesso istante. I sistemi che non garantiscono la consistenza totale dei dati possono garantire una consistenza parziale: ad esempio, possono assicurare che ogni aggiornamento raggiungerà la totalità delle repliche dei dati in un certo intervallo di tempo. Un esempio di sistema Cloud per il quale la consistenza ha un ruolo fondamentale è il sistema di online banking. Si pensi a due persone che coordinano le operazioni per prelevare l'intera somma di denaro dallo stesso account, allo stesso tempo, da due nazioni differenti: sicuramente, i due utenti si connetteranno a due repliche differenti del sistema centrale, per motivi di scalabilità geografica (p. 7), e se l'aggiornamento del primo prelievo non è sufficientemente rapido a raggiungere l'altra replica, anche il secondo utente potrebbe essere in grado di prelevare fondi dal conto anche se questi sono già stati prelevati.

- *Disponibilità*: garantisce che ogni richiesta riceva una risposta, sia in caso positivo, sia in caso negativo. In altre parole, indica che il sistema è sempre funzionante e i dati sono continuamente disponibili, senza interruzioni di servizio.

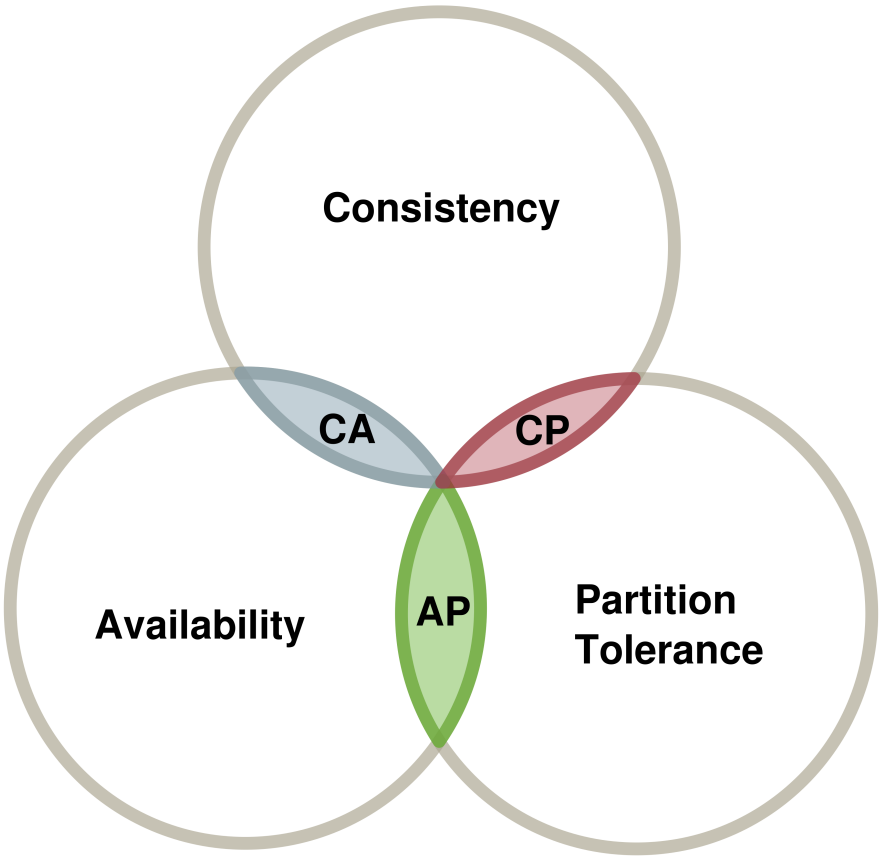

Figura 3.3: Principio CAP [37]

- *Tolleranza agli errori*: come già anticipato a p. 8, indica che il sistema è in grado di continuare la propria attività anche se alcune sue sottoparti si guastano o non sono disponibili.

MySQL [34] è un esempio di sistema CA, Voldemort [35] un esempio di AP, MongoDB [36] è invece CP.

3.3 Security Management

La sicurezza informatica consiste nell'insieme di attività atte a difendere le informazioni memorizzate da accessi non autorizzati. Il principio di sicurezza informatica si basa sui concetti di autenticazione (l'accesso ai dati deve essere

limitato ai soli utenti autorizzati) e integrità (i dati devono rimanere integri e corretti nel tempo, senza subire anomalie) [12].

3.3.1 In che modo il Cloud Computing influisce sul compito di un Security Manager?

Dato che il Cloud Computing fa della virtualizzazione uno dei suoi elementi chiave, è necessario che i protocolli di sicurezza siano implementati anche a livello virtualizzato, oltre che fisico. Pertanto, la sicurezza di un sistema Cloud è data, in parte, dalle medesime strategie di protezione di un sistema informatico tradizionale, e, in parte, dai metodi di sicurezza di ultima generazione nel campo della virtualizzazione e dei sistemi distribuiti.

A causa della notevole condivisione delle informazioni, il primo passo da effettuare è applicare una solida crittografia ai dati personali. Tale tecnica consiste nel processo di trasformazione di un testo semplice in un nuovo testo codificato che risulti incomprensibile, a meno che non si possieda la chiave di lettura necessaria a risolvere l'algoritmo e ricavare il messaggio originale. E' compito del Security Manager stabilire quali dati necessitino crittografia, e quali invece no, tenendo in considerazione il fatto che questa tecnica comporta un inevitabile appesantimento del sistema e un rallentamento nei tempi di lettura e di invio delle informazioni. Una variante di crittografia è la steganografia, tecnica che consente di nascondere informazioni all'interno di altre informazioni. Si può utilizzare, ad esempio, un immagine JPEG e criptare al suo interno un file di testo, ricavabile solo se si conosce la chiave di lettura. Tale tecnica è definita Steganografia LSB, e prevede che modificare un solo bit per ogni pixel dell'immagine non consenta di notare differenze rispetto alla versione originale; tali bit possono comporre il messaggio nascosto.

Altra area di sicurezza fondamentale nell'ambito del Cloud Computing è il controllo degli accessi. Per monitorare tutti gli accessi alla rete e agli host del sistema, è opportuno utilizzare appositi sistemi di rilevamento delle intrusioni (*Intrusion Detection Systems*) capaci di identificare accessi non autorizzati al sistema e compilare registri di revisione allo scopo di conservare informazioni

relative ai tentativi di accesso non autorizzati e alle eccezioni del sistema. Gli IDS si suddividono in [16]:

- IDS basati sulla *conoscenza*: sfruttando le conoscenze relative agli attacchi informatici, il sistema è in grado di riconoscere azioni malevole se rileva un particolare set di istruzioni sospette. Sono gli IDS più precisi, ma è fondamentale che gli aggiornamenti di sicurezza siano periodici e frequenti, per ampliare continuamente il database di conoscenze del sistema.

- IDS basati sul *comportamento*: sfruttando stime relative al carico di lavoro, e ipotesi riguardanti il comportamento tipico degli utenti, il sistema è in grado di riconoscere azioni malevole se le azioni di un particolare utente o di un sottosistema non corrispondono a quelle previste. Questo non significa che ogni singola azione non preventivata sia necessariamente malevola: infatti, questa tipologia di IDS genera un certo numero di falsi positivi, e richiede una costante manutenzione manuale per verificare quali, fra le azioni rilevate, siano effettivamente intrusioni.

Gli IDS sono inoltre:

- basati su *host*, quando monitorano gli eventi che si verificano all'interno di un singolo server, alla ricerca di malware e abusi di privilegi;

- basati su *rete*, quando monitorano una rete informatica alla ricerca di accessi non autorizzati e intercettazione del traffico.

Quali altri aspetti da tenere in considerazione nel passaggio a un sistema Cloud?

- Occorre ricordare che eventuali danni alla sicurezza del sistema possono generare, oltre alla ovvia perdita di business e di valore per l'azienda, anche azioni legali da parte dei clienti la cui privacy è stata compromessa [8].

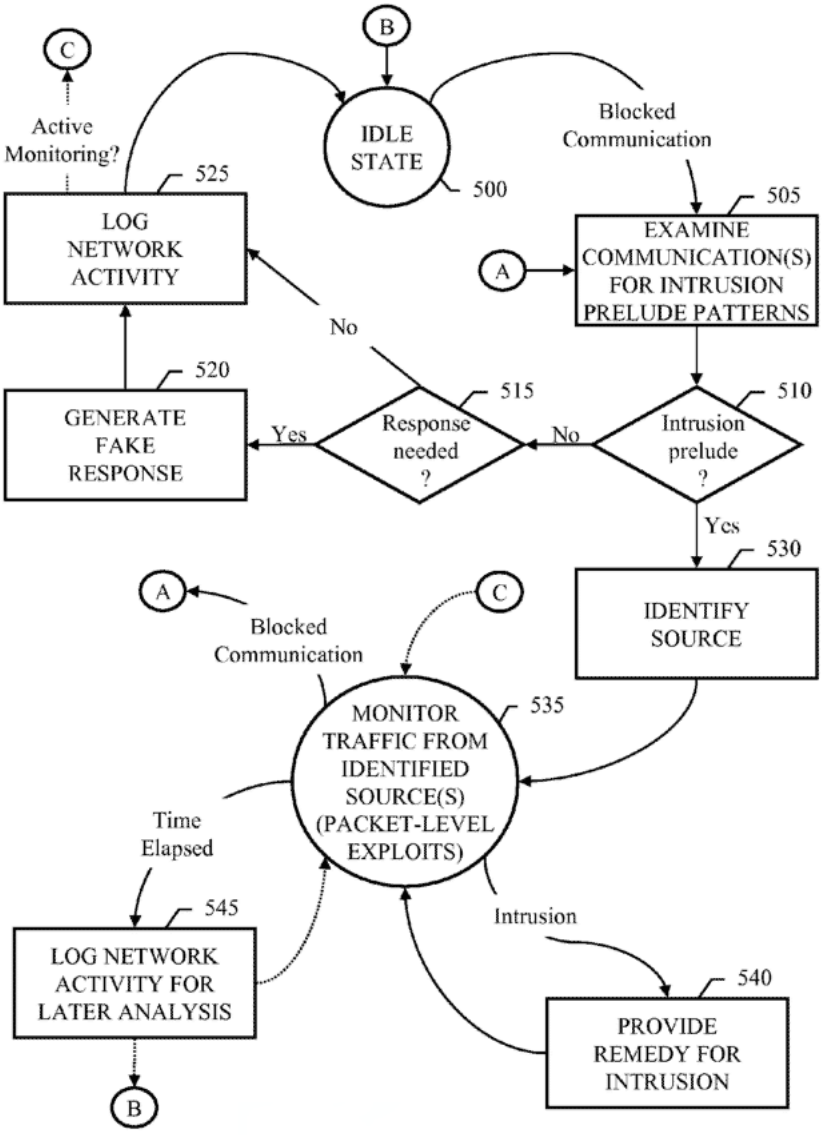

Figura 3.4: Schema comportamentale tipico di un IDS [19]

- E' di fondamentale importanza implementare un sistema *Single Sign-On* (SSO), tramite il quale le identità degli utenti sono gestite in modo univoco attraverso più servizi distribuiti: l'utente effettua l'accesso una sola volta, quindi è immediatamente connesso a tutti i servizi. Questo è estremamente utile dato che la mancanza di questo tipo di autenticazione può indurre gli utenti a sfruttare il salvataggio automatico delle password, via browser, allo scopo di minimizzare i tempi di accesso, e tale funzione favorisce gli hacker nei tentativi di intercettazione dei dati.

- Un caso recente di attacchi informatici è relativo alla diffusione del malware *Cryptowall*, in grado di criptare tutti i dati contenuti sull'hard disk del dispositivo bersaglio, trasformandoli in una serie di caratteri apparentemente non interpretabile. Il malware rende altresì impossibile il ripristino dei files a un punto precedente all'infezione. A questo punto, l'utente non potrà più leggere alcun documento presente sul suo computer, a meno di utilizzare la chiave di decriptazione che possiede l'hacker, il quale, tuttavia, esige un'ingente somma di denaro prima di comunicarla, una sorta di ricatto per la "restituzione" del computer manomesso. In questo caso, risulta estremamente utile l'utilizzo di un servizio Cloud di Disaster Recovery (p. 27), tramite il quale è possibile ripristinare interamente il sistema ad una data antecedente all'infezione del malware [29].

- Per quanto riguarda il furto di informazioni, sfortunatamente, le principali soluzioni sono in contrasto con la politica di ridondanza dei DRS: evitare backup di informazioni minimizza il rischio che i dati sensibili vengano copiati. La gestione delle chiavi crittografiche, inoltre, deve essere accuratamente valutata: idealmente, sarebbe opportuno creare nuovi chiavi crittografiche ad ogni singola comunicazione, in modo tale che il furto di una chiave comporti, nel peggiore dei casi, la decifrazione di un singolo messaggio. Certamente, questa soluzione comporta un ap-

pesantimento delle comunicazioni, ed è pertanto consigliabile un'analisi dei rischi che permetta di capire quali aree del sistema necessitino protocolli di sicurezza di questo tipo, e quali invece richiedano una minore attenzione.

- Con particolare riferimento alla navigazione sul web, è comune l'utilizzo di librerie di sicurezza che permettono di codificare e decodificare i messaggi inseriti nei form online. *AntiXSS* di Microsoft [38], ad esempio, fornisce numerose funzioni utili a difendere il sistema da attacchi XSS[8].

[8]XSS (Cross site scripting): tecniche che permettono all'hacker di rubare dati sensibili inseriti dagli utenti nei form di siti web dinamici che non adottano sufficienti misure di sicurezza.

Capitolo 4

Caso di studio: FlashStart

FlashStart è un filtro internet per la gestione di reti private, implementato da Collini Consulting [21], che permette filtraggio del traffico, blocco del materiale illegale e una gestione della rete in grado di ridurre sensibilmente i tempi improduttivi dell'azienda e, al contempo, monitorare le applicazioni dei clienti per garantire massimo supporto [20].

Il modello di base del filtro, definito FlashStart Cloud, presenta numerosi vantaggi per aziende, PA e scuole:

- disabilitazione di siti non idonei all'attività lavorativa (blacklist), e conseguente minimizzazione delle perdite di tempo durante l'orario di lavoro[1];
- blocco di siti infetti che possono causare problemi ai pc;
- tutela del dirigente responsabile, relativamente ai rischi legali legati all'utilizzo scorretto della rete;
- monitoraggio costante delle applicazioni e report in tempo reale dell'utilizzo della rete.

[1]Si effettuano aggiornamenti automatici periodici che consentono ai dispositivi dei clienti di ricevere protezione dalle minacce del web più recenti

Figura 4.1: Pannelllo di controllo Flashstart

Mediante un'appropriata configurazione dei DNS del router, il traffico di rete dei clienti, o parte di esso, viene indirizzato ai server del provider, i quali eseguono i controlli di sicurezza, ad esempio verificando se il sito al quale si tenta di accedere appartenga o meno a una blacklist predefinita: in caso positivo, il cliente viene reindirizzato su una pagina statica appositamente pensata per informare il cliente del blocco del traffico per motivi di sicurezza; altrimenti, viene inoltrata la richiesta originale utilizzando i DNS di default[2]. Sfruttando un pannello di controllo, il cliente è comunque in grado di gestire le blacklist a proprio piacimento, aggiungendo o rimuovendo siti in ogni momento. Quindi, il provider gestisce i servizi del cliente in modalità Cloud senza avere la necessità di intervenire direttamente sui dispositivi degli utenti.

Una versione avanzata del filtro FlashStart offerto da Collini Consulting, chiamata FlashStart Hybrid, offre sia il medesimo filtraggio della navigazione in modalità Cloud (con l'aggiunta del modulo Autentica Web, che imposta la gestione del filtraggio del traffico sulla base dell'autenticazione dell'uten-

[2]In alcuni casi, ad esempio quando il router utilizza un IP dinamico, non è possibile sfruttare la configurazione dei DNS, ed è pertanto necessario l'utilizzo di un dispositivo fisico, definito Push Device, che supporta il reindirizzamento del traffico.

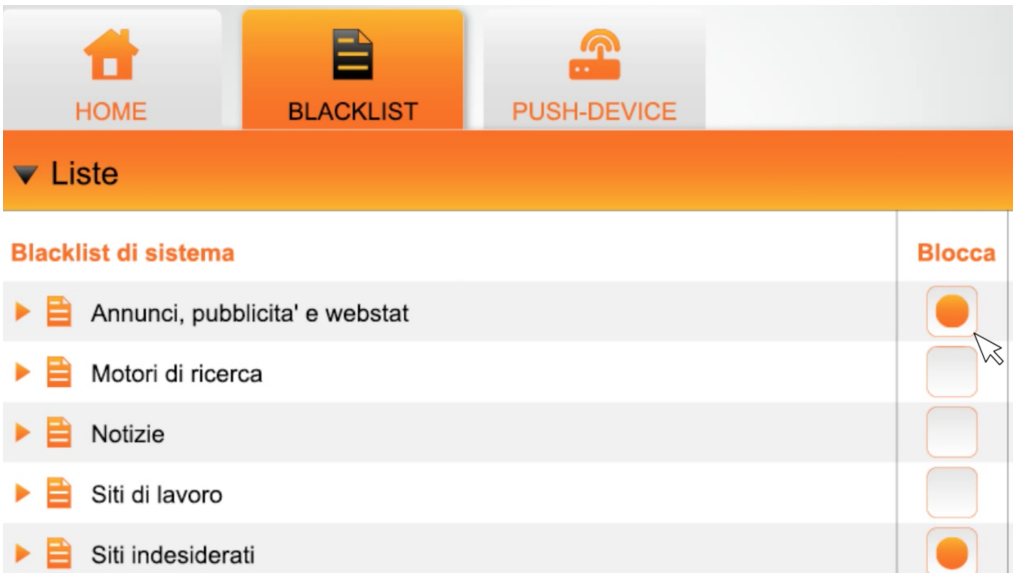

Figura 4.2: Implementazione delle blacklist su Flashstart

te), sia funzionalità aggiuntive gestite da un apposito dispositivo inserito all'interno della rete privata del cliente. Alcune delle principali funzionalità consistono in:

- *Firewall*: strumento di difesa perimetrale che controlla il traffico di rete in entrata e in uscita, gestendo l'apertura e la chiusura delle porte in ingresso ed in uscita, offre protezione da tentativi di intrusione dall'esterno e impedisce ad applicazioni interne alla rete di accedere ad Internet senza autorizzazione [22]. In linea con le tecnologie Firewall più recenti, vengono implementate tecniche DPI, filtraggi avanzati del traffico che controllano sia l'header dei pacchetti di dati, sia il contenuto del messaggio, allo scopo di individuare possibili divergenze dai protocolli che indichino attività fraudolente [26].

- *VPN*: gestione di reti private virtuali in modalità IPsec e con supporto NAT.

  - *IPsec*: è un'estensione del protocollo IP che implementa procedure sicurezza nell'ambito della comunicazione. I due protocolli

utilizzati per fornire autenticazione e garantire la confidenzialità della comunicazione sono AH (Authentication Header) e ESP (Encapsulating Security Payload) [23]. Il primo ha il compito di fornire un controllo di integrità dei dati, pacchetto per pacchetto, verificando l'autenticità del mittente tramite utilizzo delle chiavi condivise. Il secondo garantisce la confidenzialità della comunicazione, controllando, anziché l'header dei pacchetti, il payload, ossia il messaggio vero e proprio. In entrambi i casi, il controllo di integrità e autenticità viene eseguito tramite funzioni di hash (tipicamente con protocolli MD5 o SHA).

- *NAT*: tecnica che consiste nell'assegnamento di un indirizzo pubblico a un dispositivo informatico (o un gruppo di dispositivi) all'interno di una rete privata. Il principale scopo del NAT è limitare il numero di IP pubblici utilizzabili, per motivi di sicurezza.
- *Geolocalizzazione*: offre la possibilità di bloccare il traffico di navigazione e posta privata proveniente da una specifica area geografica. Questa funzione è particolarmente utile per prevenire attacchi informatici, tra i quali il sopracitato Cryptowall (p. 34).

Capitolo 5

Conclusioni

Lo scopo di questo lavoro è stato mostrare l'importanza del Cloud Computing all'interno della realtà informatica moderna, e in che modo tale tecnologia influenzi sia i piani strategici aziendali, con particolare focus sull'amministrazione dei sistemi, sia le applicazioni destinate all'utenza. Il passaggio all'ambiente Cloud ottimizza la scalabilità del sistema, permette la creazione di efficienti Disaster Recovery Systems, utili a massimizzare la conservazione dei dati, e semplifica la creazione di applicazioni distribuite grazie all'utilizzo delle API fornite dai provider.

Le agevolazioni relative alla gestione manageriale e all'amministrazione dei dati, che permettono di ottimizzare le risorse a disposizione e concentrarsi sul core aziendale, sono bilanciate da una maggiore necessità di implementare avanzate misure di sicurezza in grado di contrastare le tecniche di hacking all'avanguardia, sia da parte dei provider, che devono garantire la salvaguardia delle informazioni dei clienti, sia da parte di questi ultimi, che devono adattare le proprie applicazioni ai sistemi moderni. Si realizzano quindi Intrusion Detection Systems capaci di rilevare i tentativi di intrusioni nel perimetro di sicurezza del sistema, Firewall avanzati che controllano il traffico dei dati in entrata e in uscita, procedure crittografiche e consolidamento dei protocolli di comunicazione.

Con il mercato delle tecnologie Cloud in costante crescita, e il trend che

vede sempre più aziende abbandonare i sistemi classici per aderire alla nuvola informatica, tutto fa pensare che il Cloud Computing rappresenti non solo il presente dell'Information Technology, ma anche il suo futuro.

Riferimenti

[1] Mell, P., Grance, T., 2011: The NIST Definition of Cloud Computing, *National Institute of Standards and Technology, USA, Special Publication 800-145.*

[2] Sun, W., Zhang, K., Chen, S.K., Zhang, X., Liang, H., 2007: Software as a Service: An Integration Perspective, Service-Oriented Computing – ICSOC 2007, 5th International Conference, Vienna, Austria, 17-20 September 2007, p. 558.

[3] Magoules, F., 2009: Fundamentals of Grid Computing: Theory, Algorithms and Technologies, Chapman & Hall, pp. 131-132.

[4] Rajaraman, V., 2014: Cloud Computing, Resonance – Journal of Science Education, March 2014, pp. 242-258.

[5] Andrikopoulos, V., Binz, T., Leymann, F., Strauch, S., 2012: How to adapt applications for the Cloud environment: Challenges and solutions in migrating applications to the Cloud.

[6] Wikipedia, Scalability.
`http://en.wikipedia.org/wiki/Scalability`

[7] IBM, What is Cloud Computing?
`http://www.ibm.com/cloud-computing/in/en/what-is-cloud-computing.html`

[8] Cloud Standards Customer Council, Migrating Applications to Public Cloud Services: Roadmap for Success.
`http://www.cloudstandardscustomercouncil.org/Migrating-Apps-to-the-Cloud-Final.pdf`

[9] Project Management Institute, Guida Al Project Management Body of Knowledge, 3ª ediz., Project Management Institute, 2003, ISBN 1-930699-22-0.

[10] Wikipedia, Project Manager.
`http://it.wikipedia.org/wiki/Project_manager`

[11] Zheng, W., Fang, B., 2009: Structure-independent disaster recovery: Concept, architecture and implementations. *Sci China Ser F-Inf Sci, 2009, Volume 52, Issue 5*, pp. 813-823.

[12] Pfleeger, C.P., Pfleeger, S.L., 2008: Sicurezza in informatica, Seconda edizione italiana, Pearson, pp. 355-362.

[13] Kossmann, D., Kraska, T., 2010: Data management in the cloud: promises, state-of-the-art and open questions, Datenbank Spektrum, Volume 10, Issue 3, pp. 121-129.

[14] Oracle, Data Mining Concepts r11.1. `http://docs.oracle.com/cd/B2835901/datamine.111/b28129/`

[15] ITU-T Recommendations, Terms and definitions related to quality of service and network performance including dependability.

[16] Scarfone, K., Mell, P., 2007: NIST Guide to Intrusion Detection and Prevention Systems (IDPS).

[17] Wang, K., Su, Rui-dan, Li, Zeng-xin, Cai, Z., Zhou, Li-hua, 2006: Robust Disaster Recovery System Model, *Wuhan University Journal of Natural Sciences.*

[18] Microsoft, Calling Windows API.
`https://msdn.microsoft.com/en-us/library/172wfck9.aspx`

[19] Patent Application Pubblication, US 2010/0122317 A1, 2010, p6.

[20] Collini Consulting Sas, FlashStart.
`http://www.flashstart.it/`

[21] Collini Consulting Sas.
`http://www.colliniconsulting.it`

[22] Microsoft, What is a firewall?
`http://windows.microsoft.com/en-us/windows/what-is-firewall#1TC=windows-7`

[23] IPsec Howto, Cosa è IPsec?
`http://www.ipsec-howto.org/italian/x151.html`

[24] Oracle, Project Management Services.
`https://cloud.oracle.com/_downloads/Datasheet_ProjectExec_PPMCS/PPM_Cloud_Service_Datasheet.pdf`

[25] Vadala, D., 2002: Managing RAID on Linux, O' Reilly.

[26] Huang, K., Zhang, DF., 2010: An index-split Bloom filter for deep packet inspection, Science China Information Sciences, Vol. 54, No. 1, p.23.

[27] Otherplus Tech, La nascita degli Amazon Web Services e del cloud computing.
`https://otherplus.com/tech/la-nascita-degli-amazon-web-services-e-del-cloud-computing/`

[28] Yahoo! Finance, Amazon.com Inc. (AMZN) - NasdaqGS.
`https://it.finance.yahoo.com/q?s=AMZN`
Controllato il 15 febbraio 2015.

[29] Collini Consulting Sas, Malware Cryptowall: prima il danno e poi il ricatto!
http://blog.flashstart.it/2014/08/08/nuovo-malware-cryptowall-prima-il-danno-e-poi-il-ricatto/

[30] Cisco, Disaster Recovery: Best Practices.
http://www.cisco.com/en/US/technologies/collateral/tk869/tk769/white_paper_c11-453495.html

[31] Amazon, AWS EC2 Pricing.
http://aws.amazon.com/ec2/pricing/

[32] Kernighan, Plauger, 1978: The Elements of Programming Style.

[33] Limoncelli, T.A., Chalup, S.R., 2014: The Practice of Cloud System Administration: Designing and Operating Large Distributed Systems: 2, Addison Wesley.

[34] MySQL.
http://www.mysql.com

[35] Project Voldemort.
http://www.project-voldemort.com/voldemort/

[36] mongoDB.
http://www.mongodb.org

[37] Brewer, E.A.: Towards robust distributed systems, *Proceedings of the nineteenth annual ACM symposium on Principles of distributed computing*, New York, NY, USA.

[38] Microsoft, Anti-Cross Site Scripting Library.
https://msdn.microsoft.com/en-us/security/aa973814.aspx