Alma Mater Studiorum · Università di Bologna

facoltà di scienze matematiche, fisiche e naturali
Corso di Laurea Triennale in Informatica

L'Open-Source in ambito aziendale
Un caso di studio

Relatore:
Prof. Fabio Panzieri

Correlatore:
Ing. Davide Bolcioni

Presentata da:
Diego Luis Neto

Sessione III
Anno Accademico 2009/2010

La grandezza dell'uomo si misura in base a quel che cerca
e all'insistenza con cui egli resta alla ricerca.
M.H.

Introduzione

All'interno di questo documento verrà descritto un caso reale di utilizzo esclusivo di prodotti Open-Source all'interno di una realtà aziendale di medie dimensioni, focalizzandosi sull'aspetto infrastrutturale, ovvero sull'organizzazione della rete, delle macchine e dei servizi offerti internamente ed esternamente alla società.

L'azienda in questione è Servabit Srl, una società nata nel 2008 che vede al suo interno, non solo personale strettamente tecnico, ma figure professionali estremamente diverse e variegate come economisti, consulenti, analisti, sviluppatori software e sistemisti, che ha scelto di basare il suo lavoro quotidiano e il proprio business esclusivamente sul software libero, ritenendolo "vincente", sia dal punto di vista tecnico che strategico, rispetto al più diffuso e celebrato modello basato su software proprietario.

Prima di dedicarsi alla descrizione strettamente tecnica del lavoro svolto, è però opportuno soffermarsi a riflettere sui motivi che hanno spinto Servabit a scegliere l'Open-Source per il proprio lavoro e per quello dei propri clienti. Innanzitutto, l'assoluta convinzione della maggiore qualitá dei prodotti Open-Source utilizzati, rispetto ai corrispettivi proprietari presenti sul mercato, in quanto frutto del lavoro coordinato di grandi comunità di sviluppo, guidate nella maggior parte dei casi, non da interessi strettamente economici ma dalla semplice volontá di voler creare il miglior prodotto possibile. A questo segue uno dei motivi più ovvi ma non trascurabile, ovvero l'assenza

di costi di licenza per i software installati su ciascuna macchina partendo dal sistema operativo, *Ubuntu Linux* scelto per la sua facilità di gestione e di utilizzo da parte degli utenti, fino ad arrivare a tutte le applicazioni utilizzate da un utente canonico nel proprio lavoro quotidiano (gestione documenti, navigazione internet, gestione email, ecc.). Infine la chiave strategica delle scelte fatte, legata alla consapevolezza che in questo periodo di grandi cambiamenti a livello economico, le aziende abbiano bisogno di strumenti flessibili in grado di cambiare assieme a loro adattandosi rapidamente alle loro specifiche esigenze, adattamento impossibile attraverso l'utilizzo di prodotti proprietari rigidi per definizione.

Proprio a causa di questi cambiamenti, l'infrastruttura Servabit è stata nel corso dei suoi tre anni di vita in continua evoluzione, passando da un sistema "originale" di gestione di un'azienda con soli quattro collaboratori, localizzati in maniera stabile in un unico ufficio, all'attuale stato dell'arte estremamente più complesso, ovvero un sistema che gestisce in maniera centralizzata, circa venti collaboratori, tre diverse sedi operative nel centro di Bologna e diversi collaboratori che svolgono il loro lavoro lontani da queste. Oggetto di questo documento sarà quindi la descrizione dell'attuale stato dell'arte consolidato dell'infrastruttura informatica.

L'infrastruttura in questione é stata progettata e realizzata interamente da un team di collaboratori interni sotto la guida esperta dell'Ing. Davide Bolcioni, stimato professionista all'interno del mondo Open-Source nonchè correlatore di questo documento.

Capisaldi dell'architettura

Come detto in precedenza l'intera infrastruttura, dalla sua nascita all'attuale stato dell'arte, si è evoluta divenendo via via più complessa, restando però sempre fedele ai principi che hanno ispirato la sua progettazione e realizzazione. Il più importante di questi è l'impiego massiccio della virtualizzazione di macchine e di segmenti di rete per la gestione dei diversi servizi di cui un'azienda ha bisogno. Attraverso questa tecnica è stato possibile infatti, nei limiti delle risorse disponibili sulla macchina fisica su cui le macchine virtuali sono in esecuzione, effettuare un'opera di consolidamento dei server, evitando la proliferazione di macchine fisiche sotto-utilizzate tramite la virtualizzazione dei sistemi che necessitano di una macchina dedicata. Tutto ciò con un notevole risparmio economico, sia a livello energetico che sui costi di acquisto dell'hardware, ed effettuando tutti i test in assoluta sicurezza e tranquillità, in virtù delle proprietà intrinseche ai sistemi virtuali. Un sistema del genere risulta estremamente flessibile e funzionale ma purtroppo non è esente dal rischio di guasti e avarie...in altre parole un guasto sulla macchina fisica che "ospita" un numero n di macchine virtuali e che tiene quindi fisicamente in vita i servizi da queste erogati, rischia di farli crollare contemporaneamente, lasciando di fatto l'azienda immobilizzata fino alla risoluzione del problema. Un'eventualità del genere sarebbe di certo assai meno nefasta su un sistema basato su macchine fisiche ognuna relativa ad un servizio, in quanto una rottura andrebbe ad interrompere solo il servizio erogato dalla macchina danneggiata. Per scongiurare questa eventualità ci si é dunque affidati ad un cluster di due macchine fisiche (una attiva ed una passiva pronta a sostituire la prima in caso di necessità) governato da istanze gemelle del software *Heartbeat*(componente del progetto Linux High Avaibility), che è in grado di stabilire se un nodo ha cessato di erogare il servizio tramite un continuo "pinging" ed effettuare il failover sul nodo passivo, e *DRBD*, un sistema di ridondanza tra i due nodi del cluster. La combinazione di questi strumenti permette dunque, in caso di danneggiamento del nodo attivo del cluster, di passare al nodo passivo ripristinando rapidamente su questo, tutte

le macchine virtuali in funzione e dunque i relativi servizi. Il sistema appena decritto è stato chiamato *albireo* (come la famosa stella doppia) e rappresenta il cuore di tutta l'infrastruttura Servabit.

Al momento della sua creazione *albireo* era fisicamente sito nell'allora unica sede di Servabit in via S.Stefano a Bologna. Attualmente invece, il tutto è stato migrato nella nuova sede di via Fondazza da dove il cluster continua ad erogare tutti i servizi (che nel frattempo sono aumentati esponenzialmente), all'interno di questa e alle altre due sedi di Servabit attraverso tunnel VPN, che utilizzano la rete internet per lo scambio cifrato dei dati. In questo modo i collaboratori Servabit hanno la possibilità di accedere a tutti servizi (autenticazione centralizzata, posta, condivisione documenti, spazio disco personale, accesso a repository *svn*, *trac* e molti altri che saranno descritti nel corpo del documento) come se si trovassero sempre tutti nello stesso unico grande ufficio, con evidenti vantaggi in termini di efficienza aziendale.

Struttura del documento

Questo documento sarà composto da tre parti:

1. Nella prima parte sarà descritto lo stato dell'arte dell'architettura sviluppata analizzando la struttura e la configurazione delle reti interne a ciascuna delle tre sedi Servabit. Particolare attenzione sarà riservata alla sede sita in via Fondazza, dove come già detto, risiede fisicamente il cluster *albireo* e dalla quale vengono quindi erogati i servizi principali. In questa sezione saranno quindi descritti oltre ai diversi segmenti di rete virtuale anche le diverse macchine virtualizzate "all'interno" di *albireo*.

2. Nella seconda parte saranno trattate le implementazioni relative ai servizi fruibili attraverso l'infrastruttura Servabit. Saranno dunque descritti gli strumenti selezionati e le motivazioni che hanno indirizzato la scelta su questi anzichè su altri. Trattandosi comunque di servizi, anche se in senso più lato, saranno trattati in questo capitolo anche i software *Heartbeat*, *DRBD* (già introdotti precedentemente), *KVM* e *Bridge-Utils* utilizzati per la virtualizzazione rispettivamente di macchine e segmenti di rete.
   Infine verranno descritti gli strumenti messi a disposizione degli utenti su tutti i client "standard Servabit". Come anticipato, all'interno di Servabit parte del personale non ha competenze tecniche nel campo dell'informatica, e prima di arrivare all'interno dell'azienda non aveva mai avuto nessun contatto con il software libero a cominciare proprio dal sistema operativo *Gnu/Linux*. Per questo motivo è stata necessaria particolare attenzione nella scelta dei software da rendere disponibili agli utenti (dalla versione di Linux al Desktop environment, dal client di posta al file-manager) e nella creazione di una suite standard di programmi e configurazioni predefinite, con l'obbiettivo di rendere il passaggio al nuovo ambiente di lavoro il meno traumatico possibile.

3. Nell'ultimo capitolo del documento saranno invece riportate le consid-

erazioni finali sul lavoro svolto, i risultati dei primi tre anni di lavoro di un'azienda interamente Open-Source e i prossimi interventi in programma per migliorare l'infrastruttura Servabit aumentando il numero di servizi offerti, il grado di automazione e la sicurezza.

Indice

Elenco delle figure

Capitolo 1

Architettura sviluppata

Come già detto in fase di introduzione, Servabit dispone di un'infrastruttura informatica non ordinaria per una qualsiasi società start-up di medie dimensioni, il cui attuale stato dell'arte rappresenta il risultato di tre anni di evoluzione dettata dalle problematiche tecniche affrontate e dai grandi cambiamenti avvenuti all'interno dell'azienda a livello di dimensioni, di organizzazione e di tipologia di business. Attualmente esistono tre sedi, localizzate in tre punti diversi del centro di Bologna, che comunicano tra di loro attraverso canali virtuali.

Come rappresentato in figura 1.1 le tre sedi sono site in via S.Stefano, via Fondazza e via Guerrazzi, e riflettono fisicamente le diverse aree commerciali presidiate da Servabit, profondamente diverse ma strettamente correlate all'interno del sistema d'offerta aziendale:

Via S.Stefano sede legale e amministrativa dell'azienda.

Via Fondazza base operativa dei sistemisti che garantiscono il corretto funzionamento di tutti i servizi aziendali.

Via Guerrazzi centro di analisi e sviluppo software.

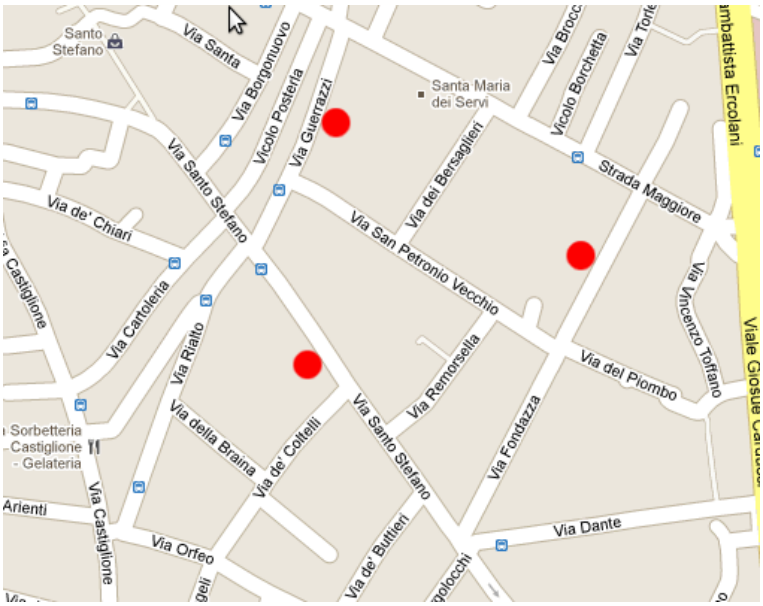

Figura 1.1: Localizzazione sedi aziendali

Come già accennato la comunicazione tra i diversi stabili avviene attraverso due tunnel VPN (Virtual Private Network) che utilizzano la rete internet per lo scambio dei dati in modo cifrato. La velocità di trasferimento dati di ogni segmento VPN, ha le caratteristiche specifiche di banda definita dal tipo di contratto che è stato definito con il provider. Si noti inoltre che, per come l'architettura è stata concepita e strutturata, le sedi di via Guerrazzi e via S.Stefano comunicano tra di loro sempre in maniera indiretta, inoltrando quindi le richieste prima alla sede di via Fondazza, nodo centrale dell'intero sistema, dove particolari regole di routing si occupano di redirigere il traffico a destinazione.

Nelle sezioni successive sarà dunque descritta nel dettaglio la struttura interna delle tre sedi Servabit, concentrandosi sulla configurazione delle reti e dei servizi. Questi ultimi, verranno descritti in maniera generica con dei rapidi accenni agli strumenti utilizzati per la loro implementazione, strumenti di cui si parlerà dettagliatamente nel capitolo successivo.

1.1 Architettura rete Fondazza

Come ormai noto, gran parte dei servizi attualmente erogati all'interno dell'azienda sono localizzati all'interno della sede di via Fondazza.
In questa, é attivo un contratto di allacciamento alla rete esterna Fastweb Small Business in fibra ottica (FTTH: Fiber To The Home) con le seguenti caratteristiche:

- Banda Download: 100 Mbit/s
- Banda Upload: 100 Mbit/s
- Banda minima garantita: 10 Mbit/s
- Doppia linea telefonica

In figura 1.2, subito successiva, è rappresentata schematicamente l'intera organizzazione delle sotto-reti interne e delle macchine (fisiche e virtuali) allacciate a queste. Per motivi grafici e non essendo particolarmente rilevanti, all'interno di questa è stata descritta in maniera semplificata la parte hardware di collegamento del cluster alla rete esterna, realizzata (partendo dall'HAG) attraverso un router Cisco direttamente collegato all'HAG fornito dal provider, uno switch esterno al quale sono connessi i due nodi del cluster *albireo*, *albireo* stesso e un altro switch al quale sono collegate tutte le macchine fisiche della sede. Per quanto riguarda invece la simbologia utilizzata, tralasciando tutto ciò che è stato rappresentato utilizzando immagini stilizzate dei dispositivi fisici, le linee nere in grassetto rappresentano i diversi segmenti di rete, i rettangoli azzurri rappresentano le diverse macchine virtuali mentre le linee azzurre tratteggiate mostrano i due tunnel vpn che collegano il cluster *albireo* con i due mini-pc *gula* e *koala* (rettangoli color arancio) situati nelle sedi di via S.Stefano e via Guerrazzi.

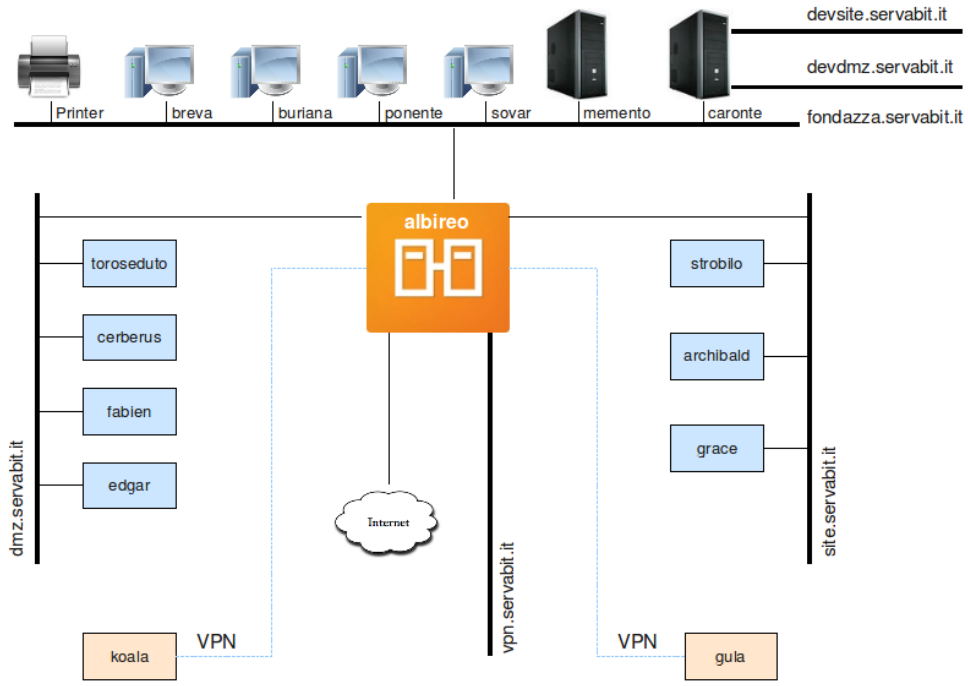

Figura 1.2: Architettura di rete sede via Fondazza

Cluster albireo

Il cluster *albireo*, è composto da due Server DELL PowerEdge 860 con 8GB di RAM e 1600GB di spazio su disco divisi equamente sui due nodi *primary* e *secondary*. Su questo è installata la versione 8.04 di *Ubuntu Server*, ultima release LTS ai tempi della messa in opera del sistema. *Albireo* svolge una funzione insostituibile all'interno dell'infrastruttura informatica Servabit, fornendo fisicamente la maggior parte dei servizi essenziali alla vita dell'azienda. Proprio a causa della sua estrema importanza è stato progettato, assieme al resto dell'infrastruttura, per lavorare senza alcuna interruzione di erogazione e senza il rischio di perdere i propri dati o il proprio lavoro. Le soluzioni adottate nella realizzazione di *albireo* hanno cercato dunque di essere le più complete ed efficaci per le maggiori problematiche che incorrono

nell'erogazione dei servizi a 360 gradi. Queste garantiscono continutità ai servizi erogati internamente ed esternamente, permettendo accessibilità sicura e recuperabilità dei dati eventualmente corrotti. Tutto questo è garantito essenzialmente dalla ridondanza hardware dei due server che forniscono i servizi e dalla virtualizzazione di reti e macchine. Tramite specifici protocolli di comunicazione implementati dai software *Heartbeat* e *DRBD* (che saranno descritti nel capitolo successivo nella sezione relativa ai servizi di alta affidabilità), mentre il server primario svolge il suo lavoro, il secondario controlla costantemente la sua vitalità, pronto a sostituire il suo clone non appena questo incorrerà in un problema. I servizi stessi sono forniti attraverso macchine virtualizzate tramite il software *KVM*, in esecuzione sui server e raggiungibili tramite reti virtuali gestite mediante *Bridge-Utils*.

Servizi diretti

Come detto il cluster *albireo* fornisce tutti i servizi vitali per l'azienda. Al suo interno infatti, sono presenti sette macchine virtuali, ognuna dedicata ad una tipologia di servizio, posizionate su due segmenti di rete virtualizzati. I due segmenti in questione sono il *site* ed il *dmz*, distinti l'uno dall'altro in quanto i server virtuali presenti su questi forniscono servizi relativamente all'interno e all'esterno dell'azienda, necessitando quindi di politiche di sicurezza completamente diverse.
Albireo però non ha solo il ruolo di grande e solido contenitore di macchine virtuali sulle quali "scaricare" l'erogazione dei servizi; al contrario svolge compiti fondamentali per il funzionamento e la sicurezza dell'intera infrastruttura. Oltre agli strumenti relativi ad alta affidabilità e virtualizzazione, al suo interno sono dunque attivi:

- Il firewall *UFW (Uncomplicated FireWall)*, fornito con la distribuzione standard di *Ubuntu Server 8.04* installata su *albireo*, ma spogliato della sua interfaccia semplificata che è risultata incompatibile con la versione installata del software *Heartbeat*. In altre parole, il "motore"

del firewall, *iptable*, lavora all'interno del cluster implementando prevalentemente due differenti politiche di sicurezza, una per ciascuno dei segmenti di rete (*site* e *dmz*) introdotti precedentemente.

- Il DHCP relay , che si occupa di inoltrare le richieste DHCP ad un server che non si trova nella stessa sotto-rete. Nel nostro caso la macchina che fornisce il servizio di DHCP è *strobilo*.

- Il servizio di VPN implementato attraverso *OpenVPN*, fondamentale per le connessioni della rete di via Fondazza con il resto del mondo aziendale e viceversa. Questo infatti svolge una duplice funzione:

  - Consente a tutti i collaboratori Servabit, previa autorizzazione attraverso certificati personali, di accedere al sistema e ai relativi servizi aziendali da ovunque si trovino attraverso il segmento VPN *vpn.servabit.it*.
  - Gestisce le connessioni con i due gateway *koala* e *gula*, implementate attraverso due interfacce dedicate tun97 e tun99, localizzati rispettivamente nelle sedi di via Guerrazzi e via S.Stefano.

Architettura di rete

Come è possibile vedere in figura 1.2 a pag 4, la rete della sede sita in via Fondazza è composta da tre segmenti di rete principali:

- *dmz.servabit.it*
- *site.servabit.it*
- *fondazza.servabit.it*

Segmento dmz

Il segmento di rete *dmz.servabit.it*, già introdotto precedentemente, è uno dei due segmenti di rete virtualizzata gestiti da *albireo* ed è relativo a tutti i servizi che necessitano di essere resi disponibili all'esterno dell'azienda. Su questo sono installate quattro macchine virtuali con sistema *Ubuntu Linux 8.04* (ad eccezione della macchina *fabien* su cui si stà testando la nuova LTS *Ubuntu Server 10.04* per una successiva migrazione collettiva) che forniscono ciascuna una differente tipologia di servizio:

cerberus Funge da DNS primario per tutto il segmento di rete sul quale si trova e si occupa dell'autenticazione centralizzata attraverso il software *OpenLDAP* che permette a ciascun collaboratore, di effettuare l'accesso a tutte le macchine e a tutti i servizi aziendali utilizzando lo stesso username e la stessa password.

toroseduto Gestisce tutti i servizi di posta sia in ingresso che in uscita e fornisce una webmail accessibile da tutti i collaboratori. Su questa macchina é inoltre attivo il sito aziendale.

fabien ed edgar Due macchine raggiungibili dall'esterno su cui é installato *OpenERP*, un applicativo gestionale Open-Source basato su database *PostgreSQL* sul quale Servabit ha incentrato gran parte del proprio business. Queste due macchine vengono utilizzate dunque per il test interno degli applicativi sviluppati sulla piattaforma e per le demo presso i potenziali clienti.

Segmento site

Il segmento *site.servabit.it* è il secondo segmento di rete virtuale gestito da *albireo* e si contrappone al segmento *dmz*, precedentemente descritto, in quanto su di esso sono posizionati i server virtuali che erogano i servizi accessibili strettamente dall'interno di Servabit. Anche in questo caso le tre macchine presenti sulla rete montano *Ubuntu Linux 8.04*, la LTS più recente ai tempi della loro installazione. Le tre macchine in questione sono:

strobilo Fornisce servizi di DHCP, DNS primario e condivisone dei file all'interno dell'azienda attraverso *Samba*.

archibald Oltre a svolgere le funzioni di DNS secondario ospita i software *SVN* e *Trac*, strumenti fondamentali per soddisfare le esigenze di sicurezza, coordinamento e condivisione del team di sviluppo software.

grace Fornisce servizi di monitoring delle risorse e delle reti attraverso una serie di strumenti che saranno descritti nel capitolo successivo. Al suo interno vengono inoltre gestiti i certificati di autenticazione, generati e gestiti internamente tramite *OpenSSL*.

Segmento fondazza

Infine l'ultimo segmento di rete relativo alla rete di via Fondazza, ovvero il segmento *fondazza.servabit.it*. Su questo, come visibile in figura 1.2 sono collegate fisicamente le workstation dei diversi collaboratori (attualmente quattro) che svolgono il loro lavoro presso la sede di via Fondazza. Oltre a queste sono presenti su questo segmento altri due server fisici che svolgono funzioni non trascurabili:

memento Una macchina che svolge la funzione di backup-server attraverso il software *BackupPc*.

caronte Un server DELL PowerEdge T3000, equipaggiato con processore quad Core Xeon da 2.83GHz, 250GB di spazio su disco e 4GB di RAM, che gestisce due ulteriori segmenti di rete virtuale, *devsite.servabit.it* (interna) e *devdmz.servabit.it* (raggiungibile dall'esterno). Su questi i

membri del team di sviluppo hanno facoltà di creare (e distruggere) macchine virtuali per effettuare ogni tipo di test ritengano opportuno per i loro applicativi. Caratteristica importante di questo server è il sistema operativo scelto: *Debian Squeeze*, anche questo installato per testare i comportamenti di macchine server con sistema operativo diverso da *Ubuntu Sever 8.04*.

1.2 Architettura rete Guerrazzi

La sede sita in via Guerrazzi ospita la parte di analisi e sviluppo software di Servabit e, come già detto in precedenza, utilizza un tunnel VPN con la sede di via Fondazza per usufruire dei servizi erogati da *albireo*, *caronte* e *memento*. La sede è provvista di un allacciamento alla rete esterna fornito dal provider EhiWeb, con cui è stato sottoscritto un abbonamento ADSL con le seguenti caratteristiche:

- Banda download max 7 Mbit/s - garantiti 256 Kbit/s
- Banda upload max 600 Kbit/s - garantiti 256 Kbit/s
- Una linea telefonica

In figura 1.3 , subito in basso, è rappresentata la struttura generale della rete di via Guerrazzi, estremamente più semplice di quella già descritta per la sede di Fondazza.

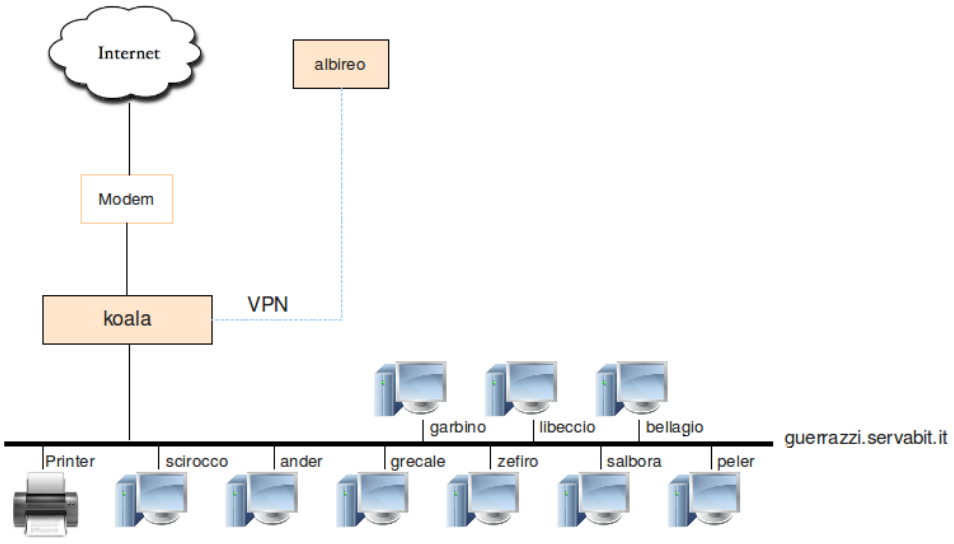

Figura 1.3: Architettura di rete sede via Guerrazzi

Come è possibile vedere la rete è strutturata su un unico segmento *guerrazzi.servabit.it*. Su questo sono posizionate dunque le nove workstation dei collaboratori e il mini-pc *koala* che, all'interno della rete svolge le funzioni fondamentali per la gestione di questa.

Mini-PC koala

Koala, introdotto precedentemente, è un mini-pc equipaggiato con un processore Intel Atom da 1,6 GHz, 80 GB di spazio disco e 2GB di RAM su cui è installato, come nella maggior parte delle macchine server Servabit, *Ubuntu Server 8.04*.
Koala è dotato di due interfacce di rete:

1. la prima collegata al modem *Fritz* fornito dal provider la cui configurazione è stata modificata per gestire solo ed esclusivamente la rete ADSL e instradare all'esterno il traffico generato internamente.

2. la seconda invece è collegata direttamente ad uno switch Cisco per permettere la gestione delle workstation dei collaboratori che svolgono il loro lavoro all'interno della sede.

Servizi

Come accennato in precedenza il mini-pc *koala* eroga all'interno della sede tutti i servizi indispensabili per la gestione della rete:

VPN Si occupa della gestione del tunnel VPN che mette in comunicazione la sede di via Guerrazzi con quella principale di via Fondazza. Questa è configurata in modalità point-to-point tra *koala* e *albireo*, che utilizzano una chiave condivisa per le loro comunicazioni.

DNS Fornisce il servizio di DNS per le macchine presenti sul segmento di rete *guerrazzi.servabit.it* fornendo a questi tutte le informazioni necessarie per tutti i sotto-domini Servabit. Nel dettaglio, il DNS di *koala* utilizza una configurazione "ibrida" agendo sia come master che

come slave in funzione del sotto-dominio interessato. Premettendo che i termini master e slave sono da intendersi con i seguenti significati:

- slave - indica che la sotto-rete specificata viene gestita all'esterno e quindi il DNS non è autoritativo per le risposte che fornisce, in quanto i dati vengono aggiornati ad intervalli regolari facendo una copia delle informazioni necessarie;
- master - indica che la sotto-rete specificata è gestita all'interno e quindi il DNS è autoritativo per le riposte che fornisce e mette a disposizione i dati ad altri DNS che ne facciano richiesta;

koala svolge dunque la funzione di master per il sotto-dominio *guerrazzi.servabit.it*. I sotto-domini per cui invece è definito come slave sono:

- *sstefano.servabit.it*
- *devsite.servabit.it*
- *devdmz.servabit.it*
- *dmz.servabit.it*
- *servabit.it*
- *site.servabit.it*
- *fondazza.servabit.it*

Avendo scelto di utilizzare un servizio di DHCP per l'assegnazione degli indirizzi ai client di rete, è necessario che le informazioni per il sottodominio *guerrazzi.servabit.it* vengano aggiornate dinamicamente. Per la configurazione scelta, essendo il server DHCP attivo sulla stessa macchina, soltanto localhost ha la possibilità di aggiornare le zone interessate, aggiornamento che avviene in modalità sicura tramite l'utilizzo di una chiave condivisa.

DHCP Come anticipato precedentemente, su koala è presente un servizio DHCP per l'assegnazione dinamica degli indirizzi ai diversi dispositivi connessi alla rete

LDAP Ultimo servizio erogato da *koala* è quello di autenticazione centralizzata tramite server LDAP, per permettere ai collaboratori di accedere alle diverse macchine sempre con le stesse credenziali. In realtà però questo servizio non è fornito direttamente dal server *koala* bensì dalla macchina virtuale *cerberus* situata in via Fondazza all'interno di *albireo*. Su *koala* è invece presente una replica delle configurazioni LDAP presenti su *cerberus* aggiornata ad intervalli regolari attraverso direttiva *syncrepl*, allo scopo di non lasciare i collaboratori impossibilitati ad accedere alle workstation nel caso di un grave disservizio della macchina *cerberus* o del cluster *albireo* in generale.

1.3 Architettura rete S.Stefano

Resta da descrivere infine la rete della sede legale e amministrativa di Servabit sita in via S.Stefano. Questa è allacciata alla rete esterna tramite un collegamento in fibra ottica (FTTH - Fiber To The Home) fornito dal provider Fastweb. Il tipo di contratto stipulato è lo stesso della sede in via Fondazza e prevede banda simmetrica in download e upload di 10Mbit/s con 1 Mbit/s garantito.
In figura 1.4, subito successiva, è rappresentata la struttura generale della rete di via S.Stefano.

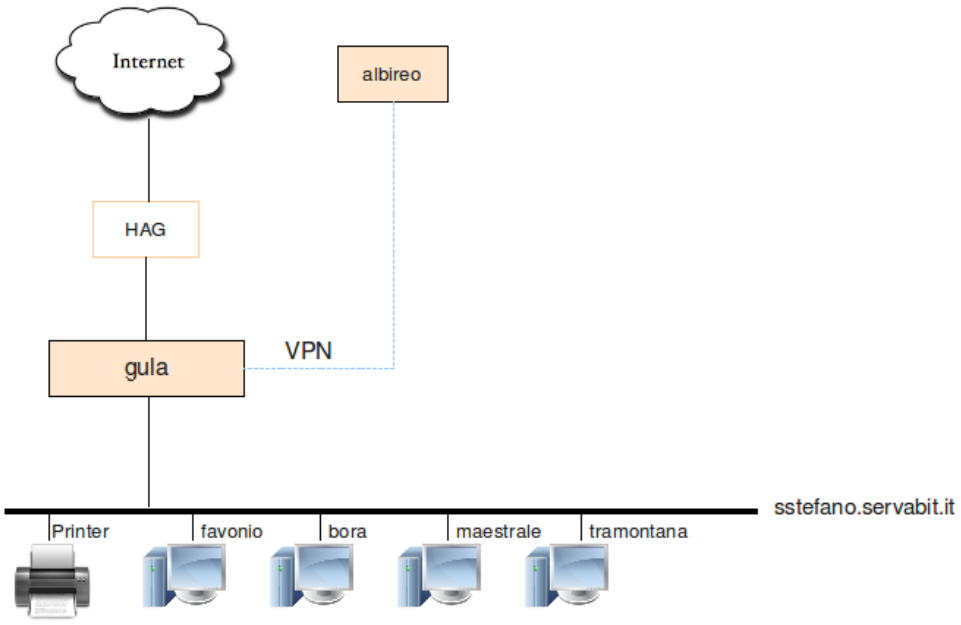

Figura 1.4: Architettura di rete sede via S.Stefano

Come è possibile vedere la rete è strutturata su un unico segmento chiamato *sstefano.servabit.it*. Su questo sono posizionate dunque le quattro workstation dei collaboratori e il mini-pc *gula* che, all'interno della rete svolge le funzioni fondamentali per la gestione di questa.

Mini-PC gula

Gula, è un mini-pc, "gemello" di *koala* (descritto nella sezione rete Guerrazzi), anche lui quindi equipaggiato con un processore Intel Atom da 1,6 GHz, 80 GB di spazio disco e 2GB di RAM sul quale è installato, come sulla maggior parte delle macchine server Servabit, *Ubuntu Server 8.04*.
Anche *gula* è dotato di due interfacce di rete:

- la prima collegata all' HAG fornito dal provider;
- la seconda invece è collegata direttamente ad uno switch Cisco per permettere la gestione delle workstation dei collaboratori che svolgono il loro lavoro all'interno della sede.

Servizi

Come accennato in precedenza il mini-pc *gula* fornisce tutti i servizi fondamentali per la gestione della rete:

VPN Si occupa della gestione del tunnel VPN che mette in comunicazione la sede di via S.Stefano con quella principale di via Fondazza. Questa è configurata in modalità point-to-point tra *gula* e *albireo*, che utilizzano una chiave condivisa per le loro comunicazioni.

DNS Fornisce il servizio di DNS per le macchine presenti sul segmento di rete *sstefano.servabit.it* fornendo a questi tutte le informazioni necessarie per tutti i sottodomini Servabit. Nel dettaglio, il DNS di *gula* utilizza una configurazione "ibrida" agendo sia come master che come slave in funzione del sotto-dominio interessato. Premettendo che i termini master e slave sono da intendersi con i seguenti significati:

- slave - indica che la sotto-rete specificata viene gestita all'esterno e quindi il DNS non è autoritativo per le risposte che fornisce, in quanto i dati vengono aggiornati ad intervalli regolari facendo una copia delle informazioni necessarie;

- master - indica che la sotto-rete specificata è gestita all'interno e quindi il DNS è autoritativo per le riposte che fornisce e mette a disposizione i dati ad altri DNS qualora ne facciano richiesta;

gula svolge il ruolo di master per il sotto-dominio *sstefano.servabit.it*. I sottodomini per cui è invece definito come slave sono:

- *guerrazzi.servabit.it*
- *devsite.servabit.it*
- *devdmz.servabit.it*
- *dmz.servabit.it*
- *servabit.it*
- *site.servabit.it*
- *fondazza.servabit.it*

Avendo scelto di utilizzare un servizio di DHCP per l'assegnazione degli indirizzi ai client di rete, è necessario che le informazioni per il sotto-dominio *sstefano.servabit.it* vengano aggiornate dinamicamente. Per la configurazione scelta, essendo il server DHCP attivo sulla stessa macchina, soltanto localhost ha facoltà di aggiornare dinamicamente le zone interessate, aggiornamento che avviene in modalità sicura tramite l'utilizzo di una chiave condivisa.

DHCP Come anticipato precedentemente, su *gula* è presente un servizio DHCP per l'assegnazione dinamica degli indirizzi ai diversi dispositivi connessi alla rete

LDAP Ultimo servizio erogato da *gula* è quello di autenticazione centralizzata tramite server LDAP, che permette ai collaboratori di accedere alle diverse macchine sempre con le stesse credenziali. In realtà, così come

succede per la sede di via Guerrazzi, questo servizio non è governato direttamente dal server *gula* bensì dalla macchina virtuale *cerberus* situata in via Fondazza all'interno di *albireo*. Su *gula* è invece presente una replica delle configurazioni LDAP presenti su *cerberus* aggiornata ad intervalli regolari tramite direttiva syncrepl, allo scopo di non lasciare i collaboratori impossibilitati ad accedere alle workstation nel caso di un grave disservizio della macchina *cerberus* o dell'intero cluster *albireo*.

Capitolo 2

Implementazione servizi

Dopo aver fornito una descrizione dell'attuale stato dell'arte dell'infrastruttura informatica Servabit, in questa sezione verranno affrontate le tematiche relative alle diverse soluzioni adottate per realizzare quanto precedentemente descritto. Saranno quindi trattate le problematiche affrontate in relazione agli strumenti impiegati per risolverle, analizzando i motivi alla base delle scelte fatte, le diverse configurazioni e i risultati raggiunti. I diversi servizi, dove per servizio ci si riferisce a qualsiasi cosa l'intero sistema metta a disposizione degli utenti, saranno dunque descritti partendo da quelli di livello più basso, come possono essere alta affidabilità e virtualizzazione delle macchine, fino ad arrivare a quelli di livello più alto, come la configurazione delle diverse workstation.

2.1 Alta affidabilità

Il progetto Servabit nasce con l'obbiettivo di realizzare, utilizzare e fornire un servizio informatico eccellente per qualità ed affidabilità, utilizzando esclusivamente software libero. Questo servizio dovrà essere disponibile e amministrabile per le reti interne all'azienda, nonchè fruibile da terze parti esterne. Qualità e affidabilità possono essere facilmente tradotte con continuità di servizio e, per raggiungere questo obbiettivo, Servabit ha preso in considerazione ogni possibile problematica che possa insorgere durante la sua erogazione.

Risolto il banale problema di un possibile blackout elettrico con un gruppo di continuità, ci si è concentrati sui problemi di soluzione meno intuitiva. È possibile infatti che, durante l'erogazione dei servizi, il sistema del Server possa andare in crash creando un blackout di durata relativa all'efficacia e all'efficienza delle soluzioni adottate. Il crash stesso, che può verificarsi con un riavvio spontaneo del sistema, nella maggior parte dei casi tende a manifestarsi con il blocco totale delle risorse. In tal caso un'auto-diagnosi del problema a livello software risulterebbe impossibile poichè un sistema bloccato non potrà mai accorgersi di esserlo, avendo bisogno, per farlo, delle stesse risorse inutilizzabili. Non essendo in grado di accorgersi del proprio stato di blocco, il server resterà dunque in stallo fino al verificarsi di un evento esterno che lo sbloccherà. Si è rivelata dunque essenziale la presenza di dispositivi specifici, in grado di determinare il verificarsi di un crash di sistema. A tal proposito, è opportuno notare che molti server dispongono di microcontroller in grado di effettuare semplici diagnosi di attività e di trovare altrettanto semplici soluzioni (nella maggior parte dei casi viene effettuato un riavvio del sistema), ai problemi riscontrati. Questa soluzione risulta però molto limitata e stride con gli obbiettivi di continuità di servizio del progetto, in quanto comporta la perdita completa del lavoro eseguito dal server fino al momento dell'incidente. Risulta quindi piú efficace una struttura esterna al sistema, ben più complessa, che riesca a diagnosticare accuratamente il

problema e che trovi la soluzione migliore per un ripristino fedele del sistema in stallo. Come anticipato nella descrizione dell'architettura realizzata Servabit, alla luce di queste riflessioni, ha deciso di affidare l'erogazione dei suoi servizi fondamentali a macchine virtuali ospitate all'interno di un cluster per alta affidabilità (High-availability cluster), basato esclusivamente su software libero.

Clustering

Per meglio comprendere il funzionamento del cluster *albireo* e degli strumenti utilizzati al suo interno, è forse opportuno fermarsi a descrivere cosa si intende per clustering, le potenzialità offerte da questa tecnica e i diversi ambiti in cui questa può essere utilizzata.

Con il termine cluster si intende una particolare configurazione hardware e software progettata con l'obiettivo di fornire all'utente un insieme di risorse computazionali estremamente potenti ed affidabili. Il cluster risponde all'esigenza di utenti ed amministratori di sistema di assemblare insieme più macchine per garantire prestazioni, capacità, disponibilità e scalabilità ad un buon rapporto prezzo/prestazioni. Si può pertanto parlare propriamente di cluster quando un insieme di computer completi ed interconnessi presenta le seguenti proprietà:

- Le varie macchine appaiono all'utente come una singola risorsa computazionale.
- Le varie componenti sono risorse dedicate al funzionamento dell'insieme.

In base al tipo di bisogni per le quali può essere utile affidarsi ad un cluster anzichè ad un unico server possiamo dunque distinguere:

- Cluster per il bilanciamento dei carichi
- Cluster a scopi di computazione

- Cluster per alta affidabilità

Come intuibile, viste le esigenze di Servabit, all'interno dell'azienda è stata realizzata questa terza tipologia di cluster per garantire un servizio stabile e continuativo. Questa, basata sulla ridondanza di più nodi hardware, è stata realizzata con configurazione attivo/passivo , in cui, come illustrato nel capitolo 1, mentre un nodo attivo si occupa di erogare i servizi, l'altro passivo resta in attesa pronto a sostituirlo in caso in cui il nodo attivo dovesse smettere di svolgere il suo lavoro in maniera corretta. Questo processo si sostituzione, conosciuto con il nome di failover, è possibile grazie alla parte software di gestione del cluster, essenzialmente composta da *Heartbeat* e *DRBD*.

Heartbeat

Heartbeat è uno strumento software inserito nel progetto Linux-HA che consente di monitorare il funzionamento di uno o più nodi di un cluster HA (High-Availability) e stabilire i comportamenti da adottare in caso di malfunzionamento di un nodo. Attraverso questo demone, i due nodi del cluster *albireo* si interrogano vicendevolmente per verificare la reciproca vitalità attraverso l'invio di pacchetti *Heartbeat* tramite un segmento ed un'interfaccia di rete dedicati. Al momento in cui il nodo attivo dovesse smettere di rispondere ai segnali inviati dal nodo passivo questo si occuperà dunque di svolgere l'operazione di failover, con le modalità descritte in figura 2.1.

All'interno di questa è possibile vedere la situazione del cluster prima (sopra) e dopo (sotto) l'operazione di failover dovuta ad un qualche problema sul nodo attivo. I due nodi del cluster sono rappresentati attraverso i due rettangoli di color arancio.

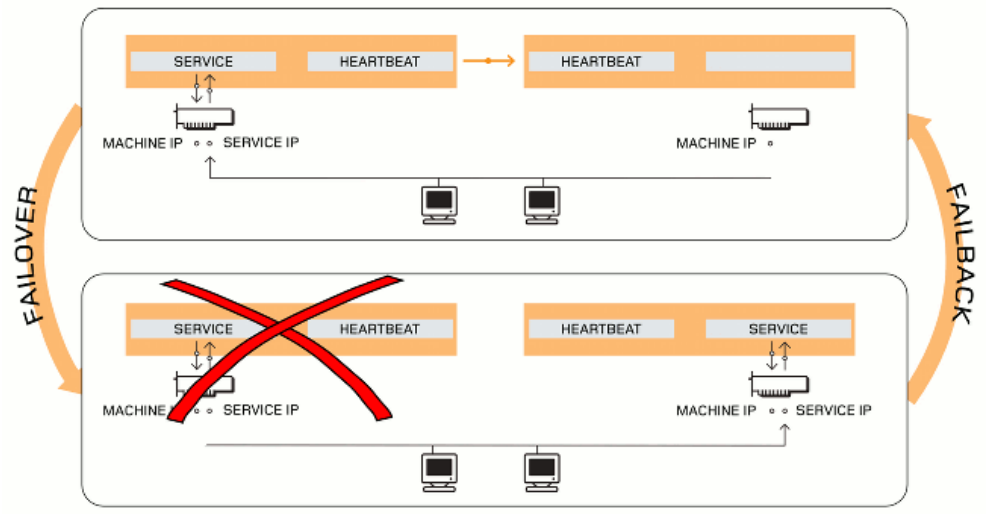

Figura 2.1: Gestione del failover attraverso Heartbeat

DRBD

DRBD è un sistema di storage distribuito per sistemi GNU/Linux composto da un modulo del kernel, diverse applicazioni per la sua gestione e alcuni script di shell. All'interno di un cluster-HA (High-availability) come quello di Servabit, si occupa di replicare tutto ciò che accade nel nodo attivo su quello passivo in maniera analoga a quanto accade nei dischi in configurazione RAID 1. *DRBD*, acronimo di Distributed Replicated Block Device, come suggerito dal nome, lavora al di sopra dei dispositivi a blocchi (hard disk o volumi logici LVM), copiando interamente ciascun blocco di dati da un nodo attivo ad uno passivo al momento della sua scrittura. All'interno del cluster questa operazione di mirroring viene eseguita in maniera sincrona, ovvero attraverso scritture strettamente accoppiate. Ciò implica che l'invio della segnalazione di fine scrittura sia inviata al file-system del nodo attivo, solo quando questa è stata effettivamente ultimata su entrambi i nodi del cluster. Al contrario della modalità asincrona, consigliata per un mirroring su lunghe distanze o che privilegi le prestazioni all'affidabilità, risulta essere la tipologia di replica ideale per i cluster-HA, in quanto garantisce transazioni corrette anche in caso di crash completo del nodo attivo.

In figura 2.2, subito seguente, è schematizzato il processo di mirroring tra i due nodi del cluster rappresentati dai due riquadri color arancio. All'interno di questi sono presenti tutte le classiche componenti del kernel *Linux*. Il normale flusso di dati tra queste è descritto attraverso le frecce nere, mentre con le frecce color arancio è rappresentato il flusso di dati governato da *DBRD*, dal nodo attivo, sul quale sono dunque attivi i servizi, a quello passivo.

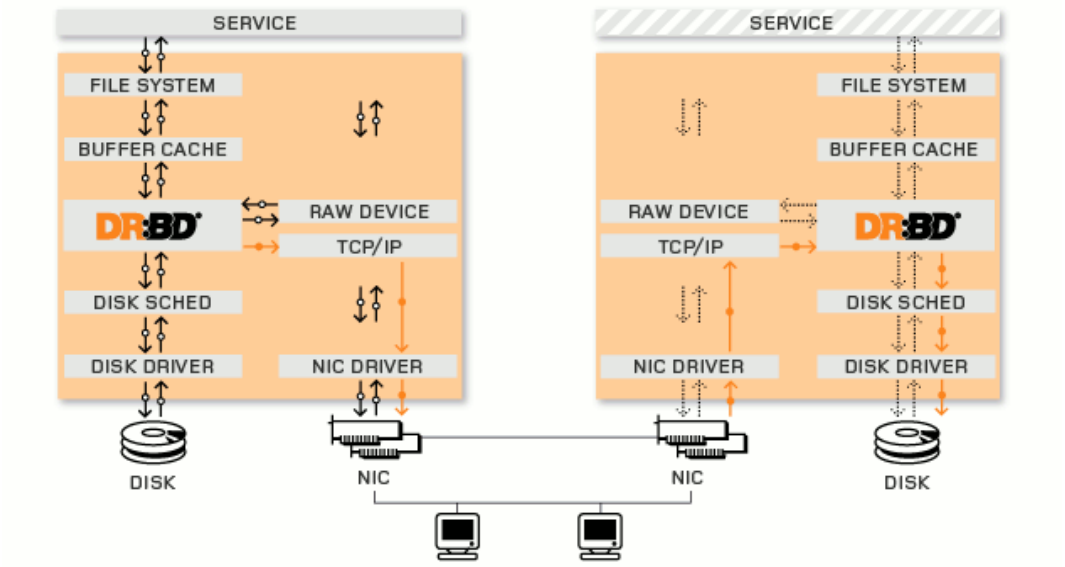

Figura 2.2: Processo di mirroring tramite DRBD

In caso di caduta di un nodo *DRBD*, si occuperà automaticamente di risincronizzare il nodo temporaneamente non disponibile alla versione più recente dei dati, tutto ció in background senza interferire con il servizio in corso. In caso invece di crash totale di entrambi i nodi, al riavvio delle macchine, *DRBD* provvederà a ripristinare la sincronizzazione tra i due nodi sulla base di quale dei due è rimasto inattivo per un lasso di tempo maggiore. In caso di crash della rete di replicazione, il software sarà in grado di ristabilire la connessione effettuando successivamente la risincronizzazione tra i nodi. *DRBD* offre inoltre svariate opzioni di supporto all'utente per il recupero automatico e manuale in caso di split-brain, situazione in cui, a causa della caduta di tutti i collegamenti di rete tra i nodi del cluster, entrambi i nodi diventano attivi con conseguenze potenzialmente molto dannose.

Non bisogna dimenticare infine che il cluster *albireo* è un server concepito per ospitare macchine virtuali, che in caso di caduta del nodo attivo dovranno tornare perfettamente attive dopo l'operazione di failover. Poichè una macchina virtuale altro non è che un file modificato costantemente, in caso

di failure del nodo attivo del cluster, queste dovranno essere recuperate dal nodo secondario al loro ultimo stato di aggiornamento. Per assicurare dunque il corretto ripristino delle macchine virtuali, queste dovrebbero trovarsi salvate in una locazione sicura e quindi distinta da una macchina che rischia improvvisamente di trovarsi in uno stato inutilizzabile. Normalmente, all'interno di architetture commerciali largamente diffuse all'interno delle aziende, questo problema viene risolto attraverso l'impiego di uno storage condiviso che risulta dunque fisicamente separato dal cluster. Attraverso l'impiego di *DRBD* invece, questo non è necessario in quanto la replicazione, eseguita in modo continuo e costante sui due nodi, fornirà al nuovo nodo attivo le copie delle macchine virtuali perfettamente sincronizzate e aggiornate al loro stato appena precedente al crash. Queste potranno quindi essere tranquillamente avviate ripristinando in tempi relativamente brevi i servizi da esse erogati.

2.2 Backup

Dopo quanto detto in precedenza sulle soluzioni per alta affidabilità adottate in Servabit, verrà ora trattata la tematica del backup dei dati. Il fatto di avere un sistema estremamente robusto, non garantisce infatti una protezione universale contro tutti i problemi in cui può incorrere un sistema aziendale. Vista l'importanza delle informazioni contenute all'interno delle diverse macchine, si rende quindi indispensabile un sistema in grado di salvare periodicamente i dati contenuti all'interno di queste e di permetterne un pratico e immediato ripristino. Proprio per questa sua importanza il sistema di backup, nel corso della vita di tutto il sistema aziendale si è evoluto notevolmente, passando da un sistema estremamente semplice e spartano ad uno più strutturato, complesso ed efficace.

La nascita del primo sistema di backup dei dati, coincide con la messa in opera del cluster ed era implementato attraverso uno script eseguito quotidianamente all'interno di *albireo*. Lo script in questione si occupava dunque di effettuare il "mount" di un disco USB esterno, copiare i dati da salvare al suo interno ed effettuare il successivo "umount". Con questo sistema venivano salvati all'interno di due dischi esterni, alternati manualmente con cadenza settimanale, tutti i file immagine delle macchine virtuali in formato qcow2(formato che rende le immagini consistenti e più facili da amministrare o da esplorare), i loro rispettivi monitor, il contenuto del file di configurazione del nodo ospitante (/usr/local , /etc) e tutti gli script di amministrazione del sistema. Questo sistema, nonostante risultasse perfettamente funzionante, presentava però alcuni inconvenienti fondamentali:

- Malgrado la ridondanza, i due nodi di *albireo* sono due macchine locali distinte ed ognuna è programmata per tentare l'esecuzione del backup. Essendo il disco di backup collegato ad un solo nodo, questo veniva effettivamente effettuato solo da una delle due macchine.

- Per motivi di unicità di ogni macchina locale, erano conservati all'in-

terno del backup esclusivamente i file di configurazione del nodo che lo aveva eseguito, non essendo le cartelle /usr/local ed /etc oggetto di duplicazione per *DRBD*. Se fosse stato dunque necessario ripristinare la cartella /etc del nodo *primary* possedendo solo quella di *secondary*, si sarebbe rivelato necessario modificare il contenuto di svariati file per eseguire un ripristino corretto.

- Il disco esterno doveva sempre essere collegato al nodo attivo del cluster, in caso dunque di switch dei nodi era necessario cambiare manualmente il collegamento fisico del disco.
- Nessun controllo sullo spazio disco libero sui dischi esterni, la cui gestione era dunque affidata alla consapevolezza degli amministratori di sistema
- Nessuna notifica in caso di fallimento del backup.

Alla luce di questi problemi, e diventando l'intero sistema infrastrutturale sempre più complesso si è dunque cercato un prodotto più evoluto per sostituire la procedura descritta precedentemente sia dal punto di vista hardware che software. Per quanto riguarda l'hardware, i dischi USB, sono stati sostituiti con un backup-server fisico dotato di 2 dischi in configurazione RAID 1, battezzato *memento* e posizionato sul segmento di rete *fondazza.servabit.it*. Per la parte software si è invece deciso di sostituire lo script sviluppato internamente con un prodotto dedicato strettamente alle procedure di backup. Il primo software testato, selezionato tra i tanti esistenti è stato *Bacula*, un applicativo Open-Source che sembrava poter rispondere alle esigenze di Servabit. Questo è stato installato all'interno di *memento* e configurato per effettuare un backup di tutte le configurazioni delle macchine virtuali presenti nel sistema, con cadenza settimanale per il full-backup e giornaliera per il differential-backup. Ci si è presto accorti però che lo strumento scelto risultava instabile e con seri problemi di salvataggio dei dati, che molto spesso venivano corrotti o salvati con nomi errati. Tutto ciò, unito alla poca praticità dello strumento in fase di navigazione e ripristino dei file salvati,

ha portato il team tecnico ad abbandonare questo strumento per andare alla ricerca di qualcosa di più solido.
Dopo una più attenta ricerca all'interno del panorama dei software open per la gestione dei backup è stato selezionato l'applicativo *BackupPC*, attualmente installato sul server *memento* per sostituire *Bacula*. *BackupPC*, oltre a funzionare perfettamente, presenta diverse caratteristiche che lo rendono un software di alta qualità particolarmente adatto alle esigenze aziendali:

- Gestione intelligente dei salvataggi. I file identici tra i diversi backup e le diverse macchine sono salvati un'unica volta attraverso l'impiego di Hard-Link. In questo modo si ottiene un sensibile risparmio di spazio sui dischi e un minor numero di scritture effettuate su questi.

- Un efficace sistema di compressione in grado di ridurre lo spazio utilizzato fino al 40%, senza appesantire eccessivamente la CPU in quanto effettuata solo per i file non presenti in backup precedenti.

- Interfaccia web pratica ma allo stesso tempo potente. Attraverso questa è possibile configurare tutte le componenti del software, gestire dunque le macchine e i file da sottoporre a procedura di backup, navigare attraverso i file già salvati e all'occorrenza ripristinarli in maniera estremamente immediata.

- Non necessita di nessun software client-side installato in quanto sfrutta protocolli o comandi standard in base al sistema ospitante.

All'interno di Servabit, l'applicativo è accessibile all'indirizzo *memento.fondazza.servabit.it/backuppc* ed stato configurato per collegarsi via SSH con privilegi di root alle diverse macchine, effettuando un backup di tipo incrementale dei file selezionati tramite direttiva *rsync*. Attualmente sono salvate giornalmente le configurazioni dei due nodi del cluster *albireo*, di tutte le macchine server presenti sui segmenti *site*, *dmz*, *fondazza*, *devsite* e *devdmz*, e quelle dei due mini-pc *koala* e *gula* che gestiscono le reti rispettivamente

di via Guerrazzi e via S.Stefano. Unica eccezione a questa regola è la directory */srv* della macchina *strobilo* sul segmento *site*, che contenendo file in condivisione tra tutti i collaboratori modificati molto frequentemente, viene sottoposta a backup due volte al giorno, in modo da minimizzare la quantità di informazioni perse nell'eventualità di un grave crash di sistema.

2.3 Virtualizzazione

Dopo aver parlato di clustering ed alta affidabilità verrà ora descritto l'altro aspetto peculiare dell'infrastruttura Servabit, ovvero l'erogazione dei servizi attraverso macchine virtuali gestite del cluster *albireo*. In termini generali, uno dei vantaggi dell'utilizzo di sistemi virtuali è proprio quello di poter disporre di sistemi di test per i più svariati usi, ad esempio uno sviluppatore potrà testare i propri software sui vari sistemi senza avere una pletora di installazioni diverse, mentre un sistemista potrà testare diverse configurazioni in tutta tranquillità senza dover cercare server per creare gli ambienti di test. In ambienti aziendali invece, la virtualizzazione permette di ridurre i costi e consolidare la propria infrastruttura aziendale, accorpando su un sistema fisico più macchine virtuali, in modo da ridurre sia il costo dell'hardware che i consumi energetici. Alla luce di queste riflessioni e di quanto detto fin'ora sulle attività di Servabit, appare subito chiaro quanto questa soluzione sia adeguata ai bisogni espressi in termini di servizi e di risorse disponibili per un'azienda di modeste dimensioni.

Nel campo della virtualizzazione il software commerciale leader del settore è *VMware*, uno strumento proprietario prodotto dalla compagnia omonima ed estremamente diffuso all'interno delle aziende che fanno uso di questa tecnologia. Essendo però Servabit, orientata esclusivamente al software libero per tutte le ragioni descritte precedentemente, si è cercata una valida alternativa Open-Source in grado di fornire soluzioni di egual, se non migliore qualità. Questa è stata individuata in *KVM*, uno strumento Open-Source per la creazione e la gestione di macchine virtuali, scelto rispetto a *XEN*, suo maggiore concorrente open, per le sue migliori prestazioni e per lo sviluppo coordinato con quello del kernel *Linux*. A questo è stato inoltre affiancato un'altro tool già introdotto precedente, ovvero *Bridge-Utils*, che assieme a *KVM* permette cosi la creazione di un ambiente virtuale completo.

KVM

KVM, acronimo per Kernel-based Virtual Machine è basato, come suggerito dal nome, su un modulo integrato nel kernel *Linux* (dalla release 2.6.20) che permette di sfruttare le estensioni di virtualizzazione dei processori di nuova generazione. Di per sè *KVM* non effettua nessun tipo di emulazione, si limita quindi ad attivare le estensioni della CPU e a mettere a disposizione dello userspace un dispositivo in grado di riservare le risorse hardware da fornire ai sistemi virtuali. Un programma userspace (in particolare le *libvirt* o *QEMU*), si occuperà dunque di richiedere al modulo di riservare determinate risorse e di svolgere le operazioni necessarie per l'esecuzione della macchina virtuale. *KVM* si occupa inoltre di aggiungere una terza modalità all'esecuzione del kernel, la *Guest Mode*, una modalità particolare in cui il kernel del sistema host consente ai kernel dei sistemi virtualizzati di eseguire operazioni privilegiate come l'accesso diretto all'hardware virtuale. Ciò che rende *KVM* un ottimo prodotto è proprio questa sua semplicità; il cuore del sistema di virtualizzazione, l'hypervisor, svolge infatti poche fondamentali operazioni, mentre tutta la parte di emulazione dell'hardware viene demandata a componenti esterne e specializzate (come le *libvirt* o *Qemu*). Questa semplicità dell'hypervisor è però strettamente legata alla dipendenza da hardware specializzato, punto debole del sistema *KVM*. Questo infatti riesce ad essere così semplice e rapido proprio grazie all'utilizzo delle estensioni speciali inserite nei processori x86 di ultima generazione: AMD-V per i processori AMD e VT-x per i processori Intel, senza le quali *KVM* non potrebbe funzionare.

Bridge-Utils

Bridge-Utils è una utility che permette la creazione e la gestione dei bridge su macchine *Linux*, permettendo così il collegamento tra diverse reti ethernet. In altre parole consente di configurare diversi segmenti di rete sul quale posizionare successivamente le macchine virtuali create, attraverso l'associazione delle rispettive interfacce TAP al bridge relativo alla sotto-rete desiderata. In questo modo viene superato il limite posto dai tradizionali sistemi

di virtualizzazione, che rendono disponibili le macchine create solo ed esclusivamente sulla stessa sottorete IP della macchina ospitante. È possibile creare dunque, diverse sotto-reti distinte su cui applicare differenti politiche di sicurezza, sulla base dei servizi erogati dalle macchine allacciate a queste, esattamente come realizzato all'interno del sistema Servabit per i segmenti *site* e *dmz*.

In un sistema che presenta una configurazione di rete cosi articolata ovviamente sono necessari dei meccanismi per fare in modo che le richieste vengano inoltrate alla macchina virtuale corretta. All'interno del sistema Servabit attualmente questo avviene attraverso delle regole di NAT che definiscono la corrispondenza tra le diverse porte, sulle quali possono arrivare le richieste, e gli indirizzi delle macchine virtuali dedicate. Una volta effettuata questa conversione il traffico viene instradato sul bridge corretto attraverso specifiche regole di routing arrivando poi a destinazione all'indirizzo specificato.

2.4 Monitoring

Come ormai chiaro, l'intera infrastruttura Servabit presenta una complessità non indifferente; complessità la cui gestione necessita di molte risorse per verificare costantemente che la rete, le macchine e i servizi funzionino regolarmente. L'obbiettivo che ci si è posti è stato dunque quello di selezionare e configurare una serie di strumenti che soddisfino queste esigenze di natura sistemistica, scaricando gli amministratori di sistema dalla necessità di predisporre piani specifici per la verifica del buon funzionamento dell'infrastruttura nel suo complesso. Il risultato atteso è stato quindi quello di ottenere un tipo di amministrazione "ondemand", dove il sistemista dovrà attivarsi immediatamente solo all'insorgere di problemi.
Considerate le esperienze pregresse nell'utilizzo dell'infrastruttura e date le competenze in essere, sono state individuate alcune tipologie di monitoraggio che è necessario attivare per raggiungere gli obbiettivi prefissati:

Monitoraggio continuativo Lo scopo di questa tipologia di monitoraggio è quello di raccogliere in modo continuativo informazioni sul funzionamento dell'infrastruttura, in modo da renderle immediatamente disponibili ogni qualvolta sia necessario, che possano fornire un utile supporto sia in fase di analisi dello stato generale dell'infrastruttura, sia per i processi decisionali concernenti gli interventi previsti per il miglioramento dei servizi. Per questo scopo è dunque necessario avere uno storico sufficientemente profondo dei dati rilevati.

Monitoraggio puntuale Questa tipologia di monitoraggio si pone invece l'obbiettivo di avere un'informazione il più possibile puntuale e tempestiva sui problemi rilevati. Idealmente dovrebbe essere dunque il singolo server a notificare lo stato delle cose.

Monitoraggio servizi La caratteristica fondamentale di questo tipo di monitoraggio è invece quella di essere tagliata su misura in base ai servizi erogati. Devono essere tenuti quindi in considerazione i parametri

specifici dei singoli servizi, in modo da valutarne sia la disponibilità che le risorse utilizzate. In questo frangente è dunque indispensabile avere un meccanismo che sia in grado di reagire in modo autonomo e indipendente nel caso in cui il servizio monitorato cessi di essere disponibile.

Strumenti di monitoraggio

Preso atto che strumenti in grado di coprire tutte le categorie di monitoraggio precedentemente descritte esistono, ma rischiano di essere di complessa configurazione o di non completa aderenza o adattamento al mutare delle necessità, si è scelto di muoversi verso una soluzione che preveda l'utilizzo di strumenti diversi e specifici, ovviamente Open-Source.

Munin

Munin è lo strumento scelto per il monitoraggio di tipo continuativo. Dotato di una pratica interfaccia web e raggiungibile all'indirizzo *munin.site.servabit.it* da qualsiasi macchina aziendale, fornisce dei precisi grafici dello stato delle risorse monitorate, mantenendone traccia per un anno. *Munin* sfrutta un'architettura a plugin attivandone, già in fase di installazione, un buon numero per il monitoraggio del sistema e mettendone a disposizione molti altri per il monitoraggio del qualsivoglia servizio. Attualmente, attraverso *Munin*, sono sottoposti a monitoraggio continuo, oltre ovviamente ai due nodi del cluster *albireo*, i segmenti di rete *fondazza.servabit.it*, *guerrazzi,servabit.it, sstefano.servabit.it*, *devsite.servabit.it*, *devdmz.servabit.it* assieme ovviamente a tutte le macchine server su questi posizionate.

Logcheck

Uno strumento che all'interno del sistema di monitoraggio si occupa di segnalare, ad intervalli regolari e configurabili, tutto ciò che accade all'interno del sistema su cui è installato. Bisogna però tenere in considerazione che

il software è ignorante rispetto a tutto ciò che lo circonda, ragion per cui necessita di una specifica configurazione volta a indicare al sistema cosa va ignorato e cosa invece notificato. La notifica dei problemi riscontrati avviene tramite email sulla mailing-list riservata ai sistemisti aziendali, che in questo modo potranno assicurare un intervento tempestivo per la risoluzione del problema.

Monit

Un software modulare che si occupa, non solo di individuare e segnalare eventuali problemi ma anche di intervenire automaticamente per cercare di risolverli. Attraverso una configurazione, che può risultare anche estremamente complessa, *Monit* è in grado di monitorare sia servizi locali che servizi remoti e, al verificarsi di condizioni che l'amministratore del sistema può definire secondo necessità, agire conseguentemente segnalando il problema, riattivando il servizio o addirittura eseguendo un comando ad-hoc.

Fail2ban

Fail2ban è un altro strumento evento-reazione per il controllo degli accessi da remoto montato localmente su ogni macchina. Questo, attraverso la verifica dei log specificati in fase di configurazione, individua tentativi di accesso da parte di utenti non autorizzati (in base al numero di richieste errate in un determinato intervallo di tempo) aggiornando successivamente le regole del firewall *iptable* per respingere automaticamente tutte le richieste arrivate dagli indirizzi degli attaccanti.

2.5 Accesso centralizzato

Al giorno d'oggi, ogni volta che si parla di servizi rivolti agli utenti, ci si imbatte inevitabilmente nel problema dell'autenticazione, necessaria per garantire la privacy dell'utente e la personalizzazione del servizio stesso sulla base delle sue specifiche preferenze. Questa caratteristica, implementata nella maggioranza dei casi attraverso uno username ed una password personali, necessita anche nel più semplice di casi, di un database (o una qualche struttura che funga come tale) dove conservare le credenziali di accesso dei singoli utenti. Il problema dell'autenticazione cresce in maniera esponenziale nel momento in cui ci si trova a dover gestire gli accessi e i privilegi di tanti utenti a svariati servizi, che possono essere molto diversi tra di loro ed erogati da sorgenti diverse. All'interno di una realtà aziendale, non è dunque ipotizzabile una gestione delle credenziali di accesso ad una serie di servizi in maniera isolata su ciascuno di questi, gestione che dopo un certo numero di modifiche, aggiunte e cancellazioni finirebbe per generare un inevitabile disallineamento sui diversi sistemi interessati. L'unica soluzione a questo problema è centralizzare tutte le informazioni relative all'autenticazione degli utenti in un'unica banca dati e fare in modo che tutti i diversi servizi si rivolgano direttamente a questa per la gestione degli accessi. In questo modo si evita la ridondanza delle informazioni semplificando in maniera drastica la gestione dei diversi utenti. Per avere un'idea concreta basti pensare al processo di inserimento di un nuovo utente all'interno di un'infrastruttura che fornisce un numero n di servizi; con un sistema che gestisce gli accessi localmente ad ogni singolo servizio, sarebbero necessarie n operazioni e l'utente in questione potrebbe avere potenzialmente n username abbinati a n password distinte. Con un sistema di gestione delle credenziali centralizzato invece, è necessaria una sola operazione per permettere al nuovo utente di accedere a tutti gli n servizi con lo stesso username e stessa password. Stesso principio vale ovviamente per tutti i generi di attività di gestione delle credenziali, non solo inserimento dunque ma anche modifica e cancellazione di queste.
Alla luce di queste considerazioni, e trovandosi di fronte ad una forte crescita

aziendale sia per numero di utenti che di servizi offerti, si è dunque cercato lo strumento in grado di fungere da banca dati unica e centrale, all'interno della quale raccogliere tutte le informazioni riguardanti le credenziali di autenticazione dei diversi collaboratori. Lo strumento individuato è stato il protocollo LDAP (Lightweight Directory Access Protocol) nella sua implementazione Open-Source, ovvero *OpenLDAP*.

Configurazione accesso centralizzato aziendale

Come detto precedentemente, all'interno dell'infrastruttura informatica Servabit viene utilizzato *OpenLDAP* per la gestione delle credenziali di autenticazione di ciascun collaboratore, credenziali che saranno dunque valide per l'accesso a tutti i servizi aziendali. Prima di tutto è però fondamentale soffermarsi a riflettere sulle sui motivi della scelta effettuata riguardo allo strumento *OpenLDAP*. Questa infatti non è l'unica implementazione esistente per il suddetto protocollo, al contrario sono presenti un gran numero di prodotti commerciali, come *SunONE Directory Server* (ex-Sun Microsystem), *Novell eDirectory* e *Microsoft Active Directory*, che implementano un gran numero di funzionalità aggiuntive rispetto al semplice LDAP. La scelta nonostante la concorrenza di aziende cosi grandi e blasonate è però ricaduta, come sempre all'interno di Servabit, sul software Open per diverse ragioni:

- Il codice di *OpenLDAP* è libero e, in quanto tale, è disponibile per il download all'indirizzo *http://www.openldap.org/* rilasciato sotto licenza *OpenLDAP Public License*. Il codice sorgente disponibile può inoltre costituire un'ottima integrazione alla documentazione già esistente.

- *OpenLDAP 2* è completamente compatibile con le specifiche della versione 3 del protocollo LDAP

- *OpenLDAP* è disponibile per una grande varietà di piattaforme che include *Linux*, *Solaris*, *MacOS* e *Microsoft Windows* (nelle sue varie incarnazioni).

- Il progetto *OpenLDAP* rappresenta inoltre l'evoluzione del progetto originale relativo al server LDAP dell'univerisità del Michigan, rispetto al quale tutte le implementazioni commerciali risultano estremamente involute.

All'interno dell'infrastruttura informatica il server LDAP, che viene dunque interrogato tutte le volte che viene richiesta autenticazione per l'accesso ad un servizio, è attivo all'interno della macchina *cerberus* sul segmento di rete *dmz.servabit.it*.
I servizi attualmente coperti da autenticazione centralizzata sono:

Login workstation e server Attraverso questa configurazione ogni collaboratore potrà effettuare l'accesso su tutte le macchine Servabit (fisiche e virtuali) sempre con le stesse credenziali e senza che l'utente sia stato manualmente inserito all'interno di queste. Tutto ciò è reso possibile da due strumenti che, se propriamente configurati, consentono di estendere il meccanismo di autenticazione di Unix fondamentalmente basato sui file */etc/passwd*, */etc/shadow*, */etc/group*:

NSS ovvero *Name Service Switch*, consente, all'interno di ambienti *Unix-like*, di configurare il sistema in modo che il name-service del sistema operativo faccia riferimento a sorgenti diverse da quelle predefinite per il reperimento delle informazioni necessarie. Nel nostro caso specifico si è sfruttata dunque questa funzionalità per "istruire" i diversi sistemi a cercare le informazioni sia all'interno dei file locali che all'interno di LDAP. In questo modo oltre a consentire l'autenticazione centralizzata su tutte le macchine si lascia aperta la possibilità di utilizzare utenti creati localmente che possono risultare molto utili per la gestione di utenze "occasionali" o di gravi disservizi dell'intero sistema infrastrutturale.

PAM acronimo inglese per *Pluggable Authentication Modules*, è un meccanismo per integrare più schemi di autenticazione a basso

livello in un'unica API ad alto livello, permettendo a programmi che necessitino di una forma di autenticazione, di essere scritti indipendentemente dallo schema di autenticazione sottostante utilizzato.

Posta e *SVN* Come per l'accesso alle diverse macchine, anche per i servizi relativi alla posta elettronica ed a *Subversion* (*SVN*), è presente un meccanismo di autenticazione basato sulle informazioni salvate sul server LDAP. Per questi servizi però non sono necessarie configurazioni particolari in quanto il loro sistema di autenticazione utilizza le credenziali di accesso del sistema ospitante, ragion per cui è stato necessario esclusivamente impostare nella maniera corretta gli strumenti *NSS* e *PAM* descritti precedentemente.

OpenERP* e *Trac Per questi applicativi, presentati brevemente nel capitolo 1, il discorso è invece completamente diverso in quanto dispongono di meccanismi di autenticazione autonomi. Fortunatamente entrambi gli strumenti scelti dispongono di moduli e plugin dedicati che, ovviamente dopo una corretta configurazione, consentono l'interfacciamento automatico con un database LDAP per il recupero delle credenziali di accesso dei diversi utenti.

In base a quanto detto fin'ora ci si rende subito conto di quanto il sistema di autenticazione centralizzata sia fondamentale all'interno dell'infrastruttura Servabit. Basti pensare che nel caso in cui dovesse esserci un qualsiasi problema che impedisca la comunicazione tra le diverse macchine e *cerberus*, ci si troverebbe nell'impossibilità di poter compiere anche la più elementare delle operazioni come ad esempio utilizzare un computer aziendale. Inutile dire che all'interno di una realtà lavorativa, un'eventualità del genere non può non essere tenuta in considerazione. Per prevenire dunque situazioni di blocco completo delle autenticazioni e conseguente impossibilità di utilizzo dei servizi sono stati effettuati diversi interventi per garantire la disponibilità costante delle informazioni necessarie per l'accesso di tutti gli utenti ai servizi

aziendali. Il primo di questi, estremamente banale ma altrettanto efficace, è stata l'impostazone di un sistema di caching delle credenziali all'interno delle singole macchine, implementato attraverso la libreria *libpam-ccreds*, che consente di memorizzare localmente le credenziali di accesso al sistema al primo accesso di un utente. I diversi sistemi dunque, al momento dell'accesso interrogheranno normalmente il server LDAP e, solo nel caso in cui questo risultasse non raggiungibile per un qualsiasi motivo, andrebbero ad utilizzare per l'accesso le credenziali salvate localmente relative all'utente in questione. Con questo sistema inoltre, tutte le macchine client Servabit diventano indipendenti dal sistema LDAP, permettendone l'utilizzo anche in assenza di connessioni di rete.
Non bisogna dimenticare inoltre che l'azienda dispone di tre sedi che utilizzano i servizi erogati da *albireo*, tra cui figura anche LDAP. Per quanto riguarda la sede di via Fondazza essendo il cluster ospitato al suo interno non esistono problemi di raggiungibilità della macchina *cerberus* e quindi del sistema LDAP. Per quanto riguarda invece le sedi di via S.Stefano e via Fondazza il problema esiste, amplificato dal fatto che queste sono connesse attraverso tunnel VPN, una tecnologia che si è rivelata estremamente delicata. Per questo motivo si è scelto di esporre la macchina virtuale *cerberus* direttamente all'esterno posizionandola sul segmento *dmz*, in modo tale da permettere l'aggiornamento periodico dei server LDAP replica presenti su *gula* e su *koala* (vedi capitolo 1.2 e 1.3), attraverso la rete internet esterna e non attraverso i canali VPN. Essendo inoltre la macchina *cerberus* accessibile dall'esterno all'indirizzo *cerberus.dmz.servabit.it*, i diversi collaboratori che lavorano lontani da Bologna, se provvisti di rete, possono effettuare l'accesso sulle loro macchine utilizzando le credenziali di accesso aggiornate, leggendole direttamente dal server LDAP centrale e non attraverso le loro credenziali locali.

2.6 Condivisione file

Ciò che differenzia le modalità di lavoro all'interno di un'azienda strutturata da quelle tipiche del lavoro free-lance, è la necessità di coordinamento e condivisione delle informazioni tra tutti i collaboratori aziendali. Per rispondere a questa esigenza all'interno di una realtà estremamente sfaccettata come quella di Servabit, si è reso necessario progettare un sistema che permetta la condivisione dei sorgenti software, della documentazione tecnica interna, dei documenti formali relativi al business aziendale e di tutte le tipologie di file la cui consultazione potrebbe risultare utile. Per quanto riguarda il codice, il problema è stato risolto in maniera semplice ed elementare attraverso l'impiego di *Subversion*, uno strumento estremamente diffuso nel campo della programmazione, che permette agli sviluppatori membri dello stesso team, di lavorare in maniera coordinata sui diversi progetti di sviluppo software. Stesso discorso vale per la documentazione tecnica gestita con *Trac*, che oltre essere particolarmente indicato per la raccolta di guide e how-to grazie ad un wiki interno estremamente curato, fornisce una pratica interfaccia grafica (web) su eventuali repository *SVN* relative ai progetti di sviluppo in essere. Situazione completamente diversa è invece quella della gestione di tutti gli altri tipi di file come *.pdf*, *.odt*, *.ppt* estremamente importanti e diffusi all'interno dell'azienda. Per la gestione "comune" di questi si rende dunque necessario un vero e proprio "file-system condiviso" organizzato in directory sulle quali, essendo non tutte le informazioni di pubblico dominio, i diversi collaboratori abbiano privilegi differenti sulla base del ruolo ricoperto all'interno dell'azienda. Per risolvere questa problematica, dopo aver definito una locazione fisica dove depositare i file, si è passati alla fase di ricerca dello strumento giusto per gestire la fruibilità delle informazioni da qualsiasi workstation collegata alla rete aziendale. Il primo strumento impiegato è stato NFS (Network File System) abbandonato però dopo un breve periodo di test perchè ritenuto troppo legato al mondo *Unix* e quindi non compatibile con l'idea di esportazione e diffusione del "modello Servabit", che intende restare interamente Open-Source, mantenendo però la compatibilità con gli

altri sistemi eventualmente presenti all'interno di un'azienda. La scelta è dunque ricaduta su *Samba*, un software libero che permette di condividere risorse sulla rete anche tra client che utilizzano sistemi operativi differenti da *Linux*, in quanto re-implementazione del protocollo *SMB/CIFS*. Oltre a permettere l'interroperabilità tra sistemi *Linux* , *Mac* e *Windows*, *Samba* fornisce un servizio di file-locking coordinato con *OpenOffice*(Suite per la produzione individuale utilizzata e proposta da Servabit), fondamentale alla luce delle attività svolte all'interno dello spazio disco condiviso aziendale. Questo infatti non funge solo da raccoglitore di file ma da vera e propria area di lavoro condivisa, all'interno della quale tutti i collaboratori possono dunque redigere i diversi documenti in maniera coordinata, senza il rischio di compromettere il lavoro altrui.

Configurazione condivisione file aziendale

La condivisione file Servabit è fisicamente posizionata nella directory */srv* della macchina virtuale *strobilo*, situata sul segmento *site* e quindi accessibile solo dall'interno della rete aziendale. Al suo interno esistono diverse directory condivise, soggette a politiche restrittive regolate sulla base dei gruppi di appartenenza dei diversi collaboratori. I permessi sui diversi file e directory sono dunque definiti da nove campi suddivisi in tre triple che indicano rispettivamente i permessi (lettura, scrittura, esecuzione) per il proprietario del file, per gli appartenenti al gruppo e per gli atri utenti.
Attualmente all'interno della condivisione servabit sono disponibili le seguenti aree, modellate sulla base dell'organizzazione aziendale e sugli specifici bisogni dei collaboratori:

Servabit Area di lavoro condivisa tra tutti i collaboratori appartenenti al gruppo *servabit*.

Commerciale Directory in cui sono raccolti tutti i documenti relativi al comitato commerciale aziendale. Tutti i collaboratori, appartenenti al

gruppo *servabit*, possono leggere all'interno di questa ma non scrivere. I permessi di scrittura sono riservati al gruppo *commerciale*.

Direzione Directory condivisa solo tra i membri del consiglio di direzione aziendale (gruppo *direzione*).

Societa Directory riservata ai soci (gruppo *societa*).

Collaboratore X Directory personale di ogni collaboratore. Ognuna di queste può essere letta e scritta solo dal legittimo proprietario. Ciascun collaboratore ha dunque a disposizione un luogo "sicuro" dove salvare i suoi file personali, accessibile da qualsiasi workstation Servabit (o da ovunque tramite VPN), eliminando allo stesso tempo il rischio di perdita dei propri dati in quanto soggetti a backup periodico.

Per mantenere coerente l'organizzazione dei permessi sui file condivisi, ogni file o directory creato attraverso *Samba*, avrà dunque come proprietario il suo creatore e come gruppo lo stesso gruppo della directory all'interno della quale il file è stato creato. In altre parole se l'utente *mrossi* crea un file all'interno della condivisione *direzione*, questo avrà come proprietario *mrossi* e come gruppo *direzione*, in modo che anche tutti gli altri utenti inseriti in questo gruppo possano leggere e modificare il file in questione.

2.7 Servizi di posta elettronica

Uno dei servizi indispensabili all'interno di un'azienda è il servizio di posta elettronica, in quanto principale mezzo di comunicazione tra i diversi collaboratori. Proprio per questa sua importanza, nella progettazione del sistema di posta elettronica aziendale si è cercato di trovare le soluzioni migliori per rendere le caselle email, sicure, pratiche e sempre raggiungibili.
Per ottenere questo obbiettivo, è stata selezionata e installata (sulla macchina *toroseduto* esposta all'esterno) una serie di applicativi, rigorosamente Open-Source, che oltre ai classici sistemi di invio e recezione della posta, fornisce anche un efficace organizzazione dei messaggi, la possibilità di gestire diverse mailing-list dedicate ai diversi reparti aziendali(tecnici, sviluppatori, ecc.), una buona protezione contro lo spam e una pratica webmail.

Strumenti di gestione posta elettronica

Postfix

Postfix è un demone di posta SMTP (categorizzato comunemente come *MTA*, o *Mail Transfer Agent*) che attualmente è presente nella quasi totalità delle distribuzioni *Linux* nonchè *MTA* di default su tutte le versioni di *Ubuntu*.
Installato su *toroseduto*, si occupa della gestione del servizio SMTP, ovvero dell'invio e della ricezione di tutte le email e del posizionamento di queste nelle giuste mailbox relative ai diversi collaboratori. Questo posizionamento avviene attraverso un sistema di aliasing impostato per mappare ogni indirizzo specificato come destinatario di una messaggio ad una o più mailbox. Con questo sistema, oltre a permettere una gestione flessibile di diverse sotto-directory per i messaggi personali, si riescono a gestire le diverse mailing-list aziendali, impostando esclusivamente alcune semplici regole ai aliasing per indicare la corrispondenza di un indirizzo (destinatario) alle diverse mailbox dei collaboratori inseriti nella mailing-list in questione.

Procmail

Procmail è un agente di consegna dei messaggi (*MDA - Mail Delivery Agent*) descrivibile anche come filtro dei messaggi, un programma dunque per elaborare i messaggi in entrata su una macchina, largamente impiegato nei sistemi *Unix/Linux*.
All'interno del sistema Servabit questo strumento, interrogato da *Postfix*, si occupa di filtrare i messaggi in arrivo secondo determinate regole, anche estremamente complesse, permettendo quindi un'azione estremamente accurata.
Proprio in virtù di questa flessibilità, *Procmail* è utilizzato per la gestione dello spam. Questo viene individuato dalla utility *Bogofilter*, che si occupa di "marcare" i messaggi come spam, sulla base di quante volte determinate parole segnalate compaiono all'interno di un messaggio. Una volta individuati, *Procmail* si occuperà dunque di spostare automaticamente tutti i messaggi marcati "SPAM = YES" nella mailbox creata appositamente per lo spam, impedendogli di intasare le caselle di posta dei destinatari.

Dovecot

Dovecot non è altro che lo strumento che si occupa della gestione del servizio IMAPS (IMAP + cifratura SSL), ovvero che consente ai collaboratori di leggere la propria posta personale. Open-Source anch'esso, oltre che per la sua semplicità di installazione e manutenzione è stato scelto tra i diversi prodotti disponibili per il suo basso impatto sulle risorse del sistema.

SquirrelMail

SquirrelMail, fornisce la webmail aziendale. Consente dunque, a tutti i collaboratori Servabit, di consultare la loro posta personale ovunque ci sia una connessione ad internet accedendo all'indirizzo *www.servabit.it/webmail*.
Estremamente diffuso, *SquirrelMail* è stato scelto tra la moltitudine di applicativi che forniscono servizi di webmail per diversi motivi:

- Supporto a protocollo STMP e IMAP.
- Generazione della pagine in puro HTML 4.0 e quindi estremamente compatibili con tutti i tipi di browser e visualizzabili senza necessità di supporto *Javascript*.
- Presenza delle sole features essenziali per un client di posta: Manipolazione delle directory, Rubrica personale, forte supporto dei tipi MIME.
- Facile da installare e configurare.
- Ottima esperienza fatta all'interno della facoltà di informatica di Bologna (presso la quale è in uso), dalla quale proviene gran parte del personale tecnico Servabit.

2.8 Configurazione client

Dopo aver parlato del motore dell'infrastruttura informatica Servabit, è giunto ora il momento occuparsi di tutto quello che c'è sopra, ovvero dell'impostazione delle workstation che permettono ai collaboratori di svolgere il loro lavoro quotidiano usufruendo di tutti i servizi aziendali in maniera semplice e pratica.
Il mondo dei client, e in modo particolare il mondo dei client aziendali, è letteralmente dominato dal software proprietario. Generalmente, infatti, su un comune PC aziendale troviamo Microsoft *Windows XP* (e relativo antivirus) e tutta la selva di software proprietari correlati per svolgere le normali azioni quotidiane. Ci riferiamo dunque a prodotti come Microsoft *Internet Explorer*, Microsoft *Outlook*, Microsoft *Office* e via discorrendo su questo filone di applicazioni che per diversi motivi, sono ormai solidamente radicati non solo nelle case e nelle aziende di tutto il mondo, ma anche nell'immaginario e nelle abitudini di chiunque sappia anche solo cosa è un computer.
Da questo punto di vista appare subito chiara la scommessa fatta da tutto il team Servabit che ha scelto, anche in questo frangente, di utilizzare solo software libero spinto non solo dall'assenza dei costi di licenza, che all'interno di un'azienda start-up possono diventare particolarmente onerosi, ma soprattutto dalla forte convinzione di poter ottenere e fornire un servizio migliore attraverso l'impiego degli strumenti ritenuti di maggiore qualità e flessibilità. Il software Open, dinamico per definizione, consente infatti all'azienda di rimanere costantemente alla frontiera tecnologica, non ingessandola all'interno di prodotti obsoleti (Windows XP, sistema operativo attualmente più diffuso all'interno delle aziende, risale ormai al lontano 2001) il cui sviluppo è decisamente lento, la sicurezza opinabile e il cui l'aggiornamento a versioni successive può risultare estremamente oneroso in termini economici.

Software Suite Standard Servabit

Dopo un periodo di selezione ed evoluzione delle scelta fatte è stata definita una suite di applicativi e configurazioni standard, installata su tutte le macchine aziendali, che comprende tutti gli strumenti per soddisfare i bisogni di tutte le tipologie di utenti presenti all'interno dell'azienda. Non bisogna dimenticare infatti che all'interno di Servabit , è presente una parte tecnica, già avvezza all'utilizzo di strumenti alternativi, ma anche una folta componente di personale non tecnico o abituato a lavorare esclusivamente con un determinato tipo di applicativi. Proprio per questa ragione il lavoro di selezione del software è stato svolto con particolare attenzione, cercando gli strumenti adatti a rendere il passaggio ai nuovi sistemi il meno traumatico possibile e la fruizione di servizi semplice e immediata, arrivando a sviluppare dei piccoli software ad hoc per i casi in cui questa esigenza di praticità non potesse essere soddisfatta con i prodotti già disponibili. Proprio alla ricerca di questa praticità si è intervenuto anche in fase di installazione e configurazione dei diversi client, implementando dei semplici programmi che automaticamente dopo l'installazione di una macchina vergine, si occupano di installare e configurare automaticamente tutti i software della *suite standard Servabit* per l'accesso ai diversi servizi aziendali di tutti gli utenti. Attraverso questo intervento sono stati abbattuti i tempi di ripristino o messa in opera effettiva di una nuova macchina, eliminando allo stesso tempo l'elevato rischio di errore umano in fase di configurazione.

Sistema Operativo

Il sistema operativo scelto per le workstation dei collaboratori è stato *Ubuntu Linux* che è oggettivamente il sistema operativo Open-Source più user-friendly, che sta crescendo con maggiore velocità rispetto ai suoi concorrenti e che offre il miglior riconoscimento automatico dell'hardware delle diverse macchine, vero tasto dolente del mondo Open-Source, anche in questo caso più per motivi politico-economici che tecnici. La versione di *Ubuntu* installata inizialmente è stata la 8.04 LTS, scelta con l'idea di non essere sos-

tituita almeno fino alla fine del suo ciclo di supporto e ritenuta abbastanza stabile da poter essere utilizzata da un utente canonico. Per motivi di compatibilità con l'hardware di nuova generazione, non supportato automaticamente dalla versione 8.04 ci si è trovati a dover effettuare obbligatoriamente, il passaggio alla versione 9.04 ed ora, essendo terminato il supporto di questa si stà procedendo alla migrazione di tutti i client sulla nuova LTS *Ubuntu* 10.04.

Desktop environment

Per quanto riguarda l'ambiente grafico da mettere a disposizione degli utenti, la scelta è ricaduta inizialmente su *KDE* V3.XX, scelto oggettivamente oltre che per la sua lunga tradizione e l'ampio numero di applicazione messe a disposizione, per la sua innegabile somiglianza grafica con Microsoft *Windows*, elemento non trascurabile nello scenario descritto precedentemente. Con il passaggio alla versione di *Ubuntu* 9.04 su tutti i client, il conseguente passaggio alla versione 4 di KDE all'epoca ancora profondamente instabile, e il progressivo abbandono delle applicazioni KDE native (*Konqueror*, *Kmail*, ecc.) si è deciso di passare all'ambiente *Gnome* decisamente più leggero e stabile del precedente.

Browser Internet

Per navigare il web e per lo sviluppo delle sue web application, Servabit ha scelto di sposare Mozilla *Firefox*, diffusissimo ormai anche in ambiente *Windows*, e apprezzatissimo anche per la vasta disponibilità di plugin che aiutano gli utenti e gli sviluppatori nel loro lavoro quotidiano.

Client email

Come client di posta all'interno dello standard Servabit è stato inserito il software *Thunderbird*, anche questo molto diffuso e apprezzato per la sua stabilità, semplicità e ricchezza di funzionalità. All'interno dello standard,

Thunderbird è impostato automaticamente oltre che per inviare e ricevere mail, anche per sincronizzare automaticamente la rubrica dei collaboratori aziendali con le informazioni personali ottenute dal server LDAP.

Produttività individuale

In questo ambito Servabit utilizza *OpenOffice*, sostanzialmente l'unico software Open-Source che possa contrastare Microsoft *Office*, strumento per cui è inoltre offerta una buona contabilità dei formati.

Gestione dischi condivisi

Per la gestione di dischi condivisi, di cui si è già trattata l'importanza sono state provate diverse soluzioni per il relativo montaggio e la consultazione dei contenuti. La prima soluzione attraverso mounting manuale dei dischi è risultata decisamente poco user-friendly. Lo strumento *SMB4K* è stato utilizzato per un periodo di tempo relativamente lungo ma oltre a generare confusione negli utenti presentava problemi nella gestione dei permessi sui diversi file. Alla fine si è deciso di implementare un semplice script, che con l'ausilio di funzionalità grafiche offerte attraverso *zenity*, permette agli utenti di effettuare il montaggio e lo smontaggio dei dischi desiderati con un semplice click, navigando poi i contenuti attraverso *Nautilus* (file manager predefinito di *Ubuntu*) proprio come se i file fossero fisicamente presenti sulla macchina.

Conclusioni

All'interno del documento è stata descritta l'impostazione dell'infrastruttura informatica di Servabit Srl, esempio concreto di come l'Open-Source non sia ne una mera imitazione di prodotti proprietari estremamente costosi ne tanto meno un qualcosa di estremamente esclusivo e riservato ad una ristretta cerchia di utenti dotati di grandi conoscenze tecniche. Al contrario si è cercato di mostrare come, attraverso l'utilizzo esclusivo di software libero, sia possibile creare un'infrastruttura informatica completa, che sia non solo di alta qualità, ma anche robusta, low-cost (rispetto ad una soluzione equivalente facente uso di tecnologie differenti) e che si adatti ad aziende di tutti i generi non necessariamente inserite nel panorama IT.

Ovviamente nei tre anni di vita del sistema, prima di arrivare all'attuale configurazione che risulta solida e quindi esportabile e proponibile sotto forma di "prodotto", ci si è imbattuti in una serie di problemi legati non solo all'evoluzione dell'azienda e alla conseguente revisione delle scelte fatte, ma anche a situazioni abbastanza tipiche del modo Open-Source, legate dunque all'hardware, agli strumenti scelti e al loro stato di sviluppo. Ci si riferisce dunque a tutte le situazioni in cui determinate scelte a livello software sono risultate, all'atto pratico, obbligate dall'inconpatibilità (o all'eccessivo investimento da fare per una configurazione "manuale") dei prodotti in uso nel sistema con le nuove componenti hardware, raramente supportate dai sistemi più datati e ancor più raramente provviste di driver specifici per i sistemi open. Altre volte invece, in sede di selezione software, ci si è dovu-

ti confrontare con una folta selva di applicativi Open-Source, nella quale è estremamente facile affidarsi ad uno strumento non adeguato alle proprie esigenze, o non sufficientemente solido da poter essere utilizzato all'interno di un'azienda. In altre situazioni invece, lo sviluppo estremamente rapido degli applicativi, grande pregio del software open dal punto di vista tecnologico, é invece risultato un elemento confusione per gli utenti non tecnici, utenti che avendo già fatto un grande sforzo mentale per cambiare completamente metodologia di lavoro, hanno avuto problemi di adattamento ai nuovi sistemi in veloce evoluzione e cambiamento. Nonostante queste difficoltà si è riusciti a trovare, il giusto compromesso tra avanguardia tecnologica e stabilità del sistema arrivando a ottenere un'infrastruttura di livello estremamente alto per un'azienda di dimensioni cosi modeste.

Sviluppi futuri

Nonostante il buon livello di solidità e di maturità raggiunto, il sistema infrastrutturale Servabit è ancora, e forse vista la natura delle soluzioni adottate resterà sempre , in fase evolutiva. Come accennato nel corso della trattazione infatti, attualmente ci sono già dei progetti di aggiornamento in atto o previsti per il lungo periodo, volti a migliorare il funzionamento complessivo del sistema Servabit:

Upgrade SO client Attualmente è in programma l'aggiornamento del sistema operativo di tutti i client dall'attuale *Ubuntu 9.04*, giunta ormai al termine del suo ciclo di supporto, alla nuova LTS *Ubuntu 10.04*. Questo aggiornamento richiede particolare attenzione in quanto prima di procedere con la fase di installazione su tutte le workstation aziendali, è necessario testare meticolosamente la compatibilità con tutti gli altri strumenti trattati nel corso del documento, in quanto questi per funzionare con il nuovo sistema installato, potrebbero necessitare di configurazioni differenti dalla versione precedente.

Autenticazione LDAP su Samba Contestualmente all'aggiornamento dei client sarà inoltre messa in opera l'autenticazione centralizzata tramite LDAP anche sul servizio di condivisione file attraverso *Samba*, attualmente unico servizio a possedere un sistema di autenticazione autonomo.

Upgrade SO Server Anche per i server, attualmente equipaggiati con *Ubuntu* 8.04 è in programma l'aggiornamento del sistema operativo ad *Ubuntu* 10.04, attualmente in test sulla macchina virtuale *fabien*.

Aggiornamento architettura *albireo* Uno dei progetti più ambiziosi in programma per un futuro non estremamente remoto, riguarda la rivisitazione integrale dell'architettura del cluster-HA *albireo*. Questa prevede il passaggio dall'attuale configurazione di tipo attivo/passivo ad una di tipo attivo/attivo, dove dunque non ci sarà più spreco di risorse causato dalla presenza di una macchina server praticamente inutilizzata. Con questa configurazione sarà possibile demandare la gestione delle diverse macchine virtuali ad uno o all'altro nodo, mantenendo la logica di failover descritta precedentemente, ovvero di un nodo pronto a prendersi carico delle funzioni dell'altro se questo dovesse incorrere in qualche problema. In tutto ciò verrà quindi mantenuto il software *DRBD* come sistema di ridondanza tra i due nodi in quanto già pronto per il supporto a cluster con configurazione attivo/attivo, mentre verrà abbandonato *Heartbeat* per passare a *Pacemaker*, uno strumento di nuova generazione (già predisposto per lavorare con *DRBD*), per la gestione delle risorse del cluster. Questo, oltre a svolgere tutte le funzioni del suo predecessore permette di superare molte delle limitazioni imposte da *Heartbeat*, prime fra tutte, l'incapacità di individuare casi di failures a livello di servizio e di gestione di un cluster di sole due macchine.

Sempre riguardo *albireo*, nei piani di ancor più lungo periodo, in ottica di evitare la presenza di risorse perennemente inutilizzate, si è iniziato a

riflettere sull'impostazione di un sistema di assegnazione dinamica delle risorse fisiche della macchina ospitante alle diverse macchine virtuali. Trattandosi attualmente solo di idee e progetti è però ancora prematuro entrare in merito ad una trattazione più tecnica delle modalità con cui questa sarà realizzata.

Sostituzione ADSL Guerrazzi La linea ADSL in uso presso la sede di via Guerrazzi, che comunica con *albireo* attraverso un tunnel VPN, nonostante i diversi provvedimenti adottati per mantenere la comunicazione il più leggera possibile, si è rivelata particolarmente lenta, in particolar modo per i servizi di condivisione file e accesso alle repository *SVN* che comportano un notevole traffico dati sulla linea. Per risolvere questo inconveniente è stata presa in considerazione l'eventualità di installazione di una linea in fibra ottica che garantirebbe maggiore velocità. Attualmente questo provvedimento, proposto dal team di sviluppo software che svolge il suo lavoro presso la suddetta sede, è in fase di discussione da parte del consiglio direttivo aziendale.

Bibliografia

[1] Florian Haas, Philipp Reisner, Lars Ellenberg, *The DRBD User's Guide*

[2] Florian Haas, *The Linux-HA User's Guide*

[3] *Introduzione ai cluster*, http://www.bitportal.it/

[4] http://www.linux-kvm.org

[5] Francesco Pedrini, *Introduzione alla virtualizzazione con KVM*, http://www.miamammausalinux.org

[6] Gerald Carter, *LDAP System Administration*, O'Reilly Media, 2003

[7] Gerald Carter, Jay Ts, Robert Eckstein, *Using Samba, Third Edition*, O'Reilly Media, 2007

[8] Giorgio Richero, Alessandro Dotti Contra, Michele Pellegrini, Matteo Atti, Matteo Davì, *Documentazione interna Servabit*

[9] *BackupPC Documentation*, http://backuppc.sourceforge.net

[10] *SquirrelMail Documentation*, http://squirrelmail.org

Ringraziamenti

Sembra sia giunto il momento dei ringraziamenti e, al contrario di quanto si dice, risulta tutto abbastanza semplice...forse perché tutte queste persone sono state, sono e saranno per me veramente importanti. Senza di loro con ogni probabilità questa pagina non esisterebbe, niente tesi, niente laurea...niente di niente.

Prima di tutto, e come se fosse possibile farlo in 4 righe, devo dire grazie ancora una volta a Mamma, Cesare e nonna Maria, per avermi permesso di crescere in un ambiente sano, di studiare, inseguire i miei desideri, le mie ambizioni e i miei sogni, spaccandosi letteralmente le ossa giorno dopo giorno e rinunciando a buona parte della loro vita in funzione della mia.
Un ringraziamento speciale e un bacio grande va poi ad Ambra, la mia ragazza, non solo per aver riletto interminabili volte la bozza di questa tesi, ma per essermi stata accanto in questi mesi densi di preoccupazioni e problemi, sopportando nervosismo, paranoie, annessi e connessi con infinita dolcezza e pazienza.
La famiglia tutta e in particolare Gianluca, dimostratosi molto più fraterno di quanto dica il debole legame di sangue...perchè anche solo la domanda: "Ma quando ti laurei?" vuol dire che in fondo ci credi.
Non si può non menzionare Servabit; il suo fondatore Luigi per avermi permesso di crescere e formarmi professionalmente sopportando, più o meno tranquillamente, tutti quelli che possono essere gli atteggiamenti di un ragazzino di vent'anni catapultato in una realtà lavorativa; tutti i colleghi, di

ieri e di oggi e in modo particolare Giorgio e Michele, che hanno contribuito enormemente alla stesura di questa tesi, colmando le mie svariate lacune in merito.
Infine gli amici, compagni di mille stronzate e a mio parere vero tesoro dell'intera esperienza universitaria...un grosso abbraccio dunque ad Andrea, Edo, Francesco, Giorgio, Marcello, Marco e Stefano, con la speranza che rimangano sempre le persone splendide che ho conosciuto.

Grazie a tutti...di cuore.