Infrastruttura IPA

Documentazione tecnica

Sistema Pubblico di Connettività

Scopo del Documento:

Lo scopo del presente documento è descrivere l'infrastruttura del servizio di Indice della PA.

Sommario

0. GENERALITA' 3
0.1 APPLICABILITÀ 3
0.2 RIFERIMENTI 3
0.3 DEFINIZIONE ED ACRONIMI 3
1. INDICE DELLA PA 4
1.1 ORGANIZZAZIONE LOGICA E STRUTTURA DELL'INDICE DELLE PA 4
2. REALIZZAZIONE DEL DIRECTORY SERVER 5
3. INFRASTRUTTURA DEL SERVIZIO DI INDICE DELLA PA 6
3.1 FRONT-END WEB 7
3.2 BACK-END 7
3.3 BACK-END APPLICATIVO 8
3.4 STORAGE & BACKUP 9
3.5 RIEPILOGO DEI PRINCIPALI SOFTWARE IN USO E RELATIVE VERSIONI 10

0. GENERALITA'

0.1 Applicabilità

Il presente documento si applica all'infrastruttura "SICA" e specificamente ai sistemi preposti all'erogazione del servizio di Indice della PA

0.2 Riferimenti

Identificativo	Titolo/Descrizione
SPCoop-Schema_Interop_IndicePA	Schema per l'Interoperabilità dell'Indice delle Pubbliche Amministrazioni
CO-CA-SW-IPA-Descrizione data base Mysql	Descrizione del database Mysql

0.3 Definizione ed Acronimi

Definizione / Acronimo	Descrizione
SPC	Sistema Pubblico di Connettività
IGPEC	Indice dei Gestori di Posta Elettronica Certificata
IPA	Indice delle Pubbliche Amministrazioni
IANA	Internet Assigned Numbers Authority
DigitPA	Ente nazionale per la digitalizzazione della Pubblica Amministrazione

1. Indice della PA

Da un punto di vista funzionale, il servizio “Indice delle Pubbliche Amministrazioni e delle Aree Organizzative Omogenee” viene suddiviso nelle seguenti componenti:

- un servizio di directory LDAP;
- una interfaccia web per la consultazione e la ricerca dei dati;
- un servizio di gestione dell’Indice e di aggiornamento dei dati.

La prima componente realizza la base dati dell’Indice delle PA, accessibile tramite il protocollo LDAP versione 3, definito dalla RFC 1777; all’interno di questa base dati, vengono inseriti e gestiti tanto i dati relativi alla Struttura Organizzativa (Indice delle UO), quanto quelli riguardanti le Aree Organizzative Omogenee (Indice delle AOO) delle singole PA accreditate.

L’interfaccia web costituisce un front end per la ricerca guidata e la visualizzazione di informazioni contenute nell’indice; attraverso di essa, le informazioni presenti nella base dati possono essere accedute da un qualsiasi browser web attraverso il protocollo HTTP.

La terza componente, riguarda l’implementazione di tutte le funzionalità di gestione dei dati presenti nell’indice che, nel tempo, devono essere mantenuti coerenti con la struttura organizzativa delle AOO e degli uffici di tutte le PA accreditate.

1.1 Organizzazione logica e struttura dell’Indice delle PA

Da un punto di vista strutturale, l’indice delle PA risulta costituito da due indici tra di loro correlati:

1. l’indice delle Aree Organizzative Omogenee, che contiene e riporta i dati previsti dal d.P.C.M del 31 Ottobre 2000;
2. l’indice delle Unità Organizzative, che contiene i dati di struttura delle amministrazioni accreditate

Entrambi questi indici sono ospitati in una unica struttura di Directory Information Tree (DIT), le cui caratteristiche sono descritte in “SPCoop-Schema_Interop_IndicePA”.

2. Realizzazione del Directory Server

Lo schema di un directory server rappresenta la definizione:

1. delle classi di oggetti che esso può contenere (le objectclass);
2. del set di attributi che caratterizzano l'appartenenza degli oggetti ad una particolare classe.

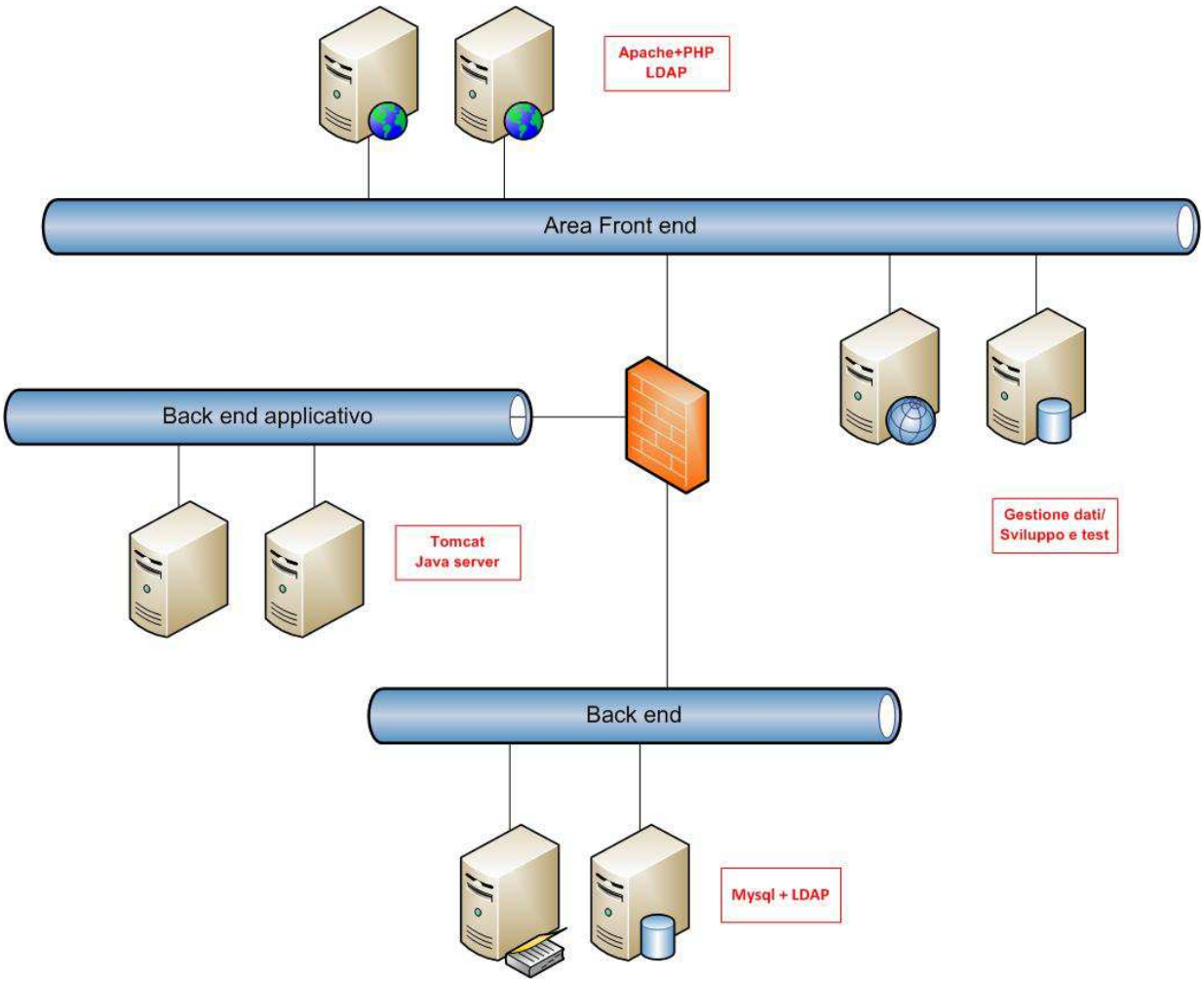
La descrizione completa delle objectclass e degli attributi definiti per il servizio IPA è contenuta nel documento "SPCoop-Schema_Interop_IndicePA".

3. Infrastruttura del servizio di Indice della PA

L'architettura implementata per il servizio prevede:

- Un directory server LDAP strutturato su diversi livelli; un livello di back-end per la validazione dei dati in ingresso e per l'autenticazione delle utenze di accesso al sito ed un livello di front-end per la pubblicazione dei dati dell'IPA attraverso il protocollo LDAP.
- Una batteria di front-end web realizzati con tecnologia open per la pubblicazione di un portale con funzionalità di ricerca e visualizzazione dei dati pubblicati.
- Un servizio di gestione dei dati realizzato con tecnologia open per la gestione dei dati di pertinenza delle pubbliche amministrazioni.

La seguente figura illustra l'infrastruttura di rete e le diverse componenti del servizio.



Nel seguito saranno descritte la realizzazione e le funzionalità delle varie componenti.

3.1 Front-end web

Il servizio di front-end web, realizzato mediante una batteria di server apache paritetici, assolve le seguenti funzioni:

- pubblicazione di un portale realizzato in PHP, versione 5, per la visualizzazione e la ricerca dei dati relativi alle pubbliche amministrazioni.
- reindirizzamento, attraverso il modulo mod_proxy di apache, verso i sistemi dell'area applicativa per le funzionalità di modifica dei dati .

Il bilanciamento sui diversi server di front-end avviene attraverso un appliance dedicato della Cisco.

Nell'area front-end sono presenti, inoltre:

- un terzo sistema web, non esposto al pubblico, che svolge il duplice ruolo di:
    1. pubblicazione delle pagine di amministrazione dei dati dedicate al Gestore dell'IPA;
    2. realizzazione di un server DNS master per l'editing delle zone pubbliche relative al servizio;
- un sistema web di pre-produzione per l'analisi di modifiche evolutive o correttive al codice prima della messa in produzione.

3.2 Back-end

Il servizio di back-end, come detto, realizza la funzionalità di database per i dati dell'Indice della PA.

Il servizio è realizzato mediante l'utilizzo di due database distinti e configurati in cluster active-standby:

- Un database relazionale, implementato mediante il software Mysql, versione 5.0.x, per il recepimento immediato delle modifiche alla base dati attraverso le funzionalità di gestione. Il database è descritto nel documento "CO-CA-SW-IPA-Descrizione data base Mysql".
- Un directory server LDAP, versione 2.3.43, per il consolidamento delle modifiche giornaliere, per l'autenticazione di referenti e utenze di accesso e per la pubblicazione dei dati.

In condizioni di funzionamento normale ogni server ha in esecuzione uno solo dei due servizi: in caso di malfunzionamento di un nodo, il sistema cluster provvede a spostare il servizio del nodo caduto sull'altro nodo.

Un batch notturno provvede ad esportare i dati presenti su Mysql in formato LDIF ed a caricarli sul directory LDAP per la loro pubblicazione definitiva.

Al proposito si riporta la tabella di riepilogo con la descrizione di tutte le procedure di movimentazione dati in esecuzione sul back-end.

Procedura	Schedulazione	Descrizione
`/opt/bin/check-logo-2.sh`	Ogni 5 minuti	Procedura di copia dei loghi dall'area di upload all'area pubblica.
`/opt/bin/check-attach.sh`	Ogni 5 minuti	Procedura di copia degli allegati dall'area di upload all'area riservata.
`/opt/bin/igpec_update/igpec.sh`	Ore 22:45	Procedura di allineamento dei dati di PEC presenti nel db Mysql rispetto al contenuto di IGPEC.
`/opt/batch/tRunJob_LDIF/tRunJob_LDIF_run.sh`	Ore 00:05	Procedura di export dei dati Mysql in formato LDIF
`/opt/bin/ldap-update.sh`	Ore 00:20	Procedura di caricamento dei dati LDIF esportati da Mysql su back-end e front-end LDAP.
`/opt/bin/IPABATCH/PECFATTEL/bin/ALLstart.sh`	Ore 04:00	Procedura in uso per il sistema di fatturazione elettronica.
`/opt/bin/LDAPACCESS/bin/LDAPACCESSstart.sh`	Ore 04:30	Procedura di caricamento sul db dei log degli accessi LDAP.
`/opt/bin/IPABATCH/IPASTAT/bin/IPASTATstart.sh`	Ore 05:00	Procedura di esecuzione di statistiche sulla PEC delle amministrazioni.
`/opt/bin/IPABATCH/OPENDATA/bin/OPENDATAstart.sh`	Ore 06:00	Procedura di export dei dati in formato Open Data per distribuzione sui front-end.

3.3 Back-end applicativo

Nel back-end applicativo è presente una batteria di java-server paritetici, realizzati mediante il software tomcat, versione 6.0.24, che espongono le funzioni di modifica e gestione dati.

Il bilanciamento tra i vari server è realizzato direttamente dai server di front-end in modalità active-standby per garantire il mantenimento delle sessioni di lavoro: ogni sistema apache, cioè, punta sempre allo stesso server tomcat e passa sull'altro solo in caso di indisponibilità del primo.

I pacchetti applicativi java (war) che eseguono le servlet relative alle varie funzioni vengono installati ed eseguiti su ciascuna macchina attraverso le funzionalità di deploy del container tomcat; i pacchetti installati sono i seguenti:

- IPA.war, per l'esecuzione delle servlet che riguardano la gestione dati da parte dei referenti delle amministrazioni accreditate;
- IPAACCRED.war, per l'esecuzione della servlet che si occupa di gestire le richieste di nuovi accreditamenti;
- IPAASS.war, per l'esecuzione della servlet che si occupa delle funzioni di assistenza ed amministrazione (gestione ticket, modifica dati, etc).

Ciascuno di questi pacchetti viene eseguito da tomcat in un diverso contesto web che serve anche a definire le regole di accesso.

L'accesso ai diversi context è regolato direttamente dal modulo mod_proxy di apache che, prima di indirizzare le richieste verso le diverse applicazioni verifica se le credenziali fornite garantiscono il livello di accesso richiesto e precisamente:

- Per il contesto /IPAACCRED non è richiesta autenticazione alcuna (è una funzione pubblica).
- Per l'accesso al contesto /IPA sono richieste le credenziali di accesso dei referenti delle amministrazioni accreditate all'IPA.
- Per l'accesso al contesto /IPAASS sono richieste le credenziali dell'utente amministratore.

3.4 Storage & backup

Il sottosistema dischi dell'infrastruttura di Indice PA è implementato mediante uno storage array HP in configurazione RAID5 su cui sono stati allocati i seguenti spazi:

- 350Gb circa per Mysql
- 90Gb circa per LDAP

I due volumi in questione sono gestiti direttamente dal cluster Redhat e contengono i dati dei due database ed i log e le elaborazioni delle procedure notturne.

Il backup dell'infrastruttura avviene con cadenza giornaliera e prevede i seguenti passi:

- Su ciascun back-end viene effettuato l'export, nella directory /var/backup, del contenuto delle due basi dati sui dischi locali del sistema che ospita il relativo servizio
- Su ciascun front-end vengono ruotati i log di accesso ed archiviati nella cartella /WEBIPA/logs/_ARCHIVE
- I file di export, i log ruotati, le configurazioni delle applicazioni ed i contenuti del sito web, vengono salvati su nastro attraverso l'agent di backup del prodotto HP Dataprotector. La policy di backup prevede un salvataggio incrementale con retention pari ad un mese per le configurazioni, sei mesi per i dati di ciascun database e due anni per i file di log.

3.5 Riepilogo dei principali software in uso e relative versioni

Software	Versione	Descrizione
Mysql	5.0.x	Database relazionale installato sui back-end
Openldap	2.3.43	Directory server ldap installato su back-end e front-end
Apache	2.2.3	Web server installato su front-end
PHP	5.1.x	Interprete PHP installato su front-end
mod_authz_ldap	0.26-9	Modulo per autenticazione LDAP su apache
Tomcat	6.0.24	Java container installato su back-end applicativo
Java	JRE 1.6.0_24	Interprete java installato su back-end applicativo