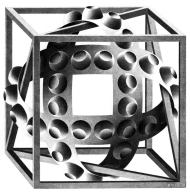
LA NUOVA ARCHITETTURA
SECOLI

LINEE GUIDA



FATA INFORMATICA
Roberto Reale
Marta Serpietri

X Kal. Sext. MMXIII

SOMMARIO

Alla luce delle nuove sfide cui il portale web SECOLI è chiamato a far fronte, si esaminano criticità e requisiti, tracciando i lineamenti di una proposta progettuale.

1 PREMESSE

Il portale web SECOLI è uno strumento d'avanguardia che l'Amministrazione offre ai cittadini italiani residenti all'estero. *il portale SECOLI*

In linea con i contenuti dell'Agenda Digitale italiana ed europea, il portale eroga per via telematica una vasta gamma di servizi, in precedenza fruibili soltanto recandosi di persona in Consolato. Anche in quei casi in cui la presenza allo sportello sia necessaria, attraverso le pagine del portale è possibile fissare un appuntamento e predisporre in anticipo i dati di interesse (documenti da produrre, percezioni da corrispondere, etc.).

Il portale si presenta come canale privilegiato tra la sede consolare e la comunità residente all'estero: di importanza primaria, dunque, in termini di visibilità e prestigio dell'Amministrazione. *canale privilegiato con la comunità residente all'estero*

Da un punto di vista tecnico, individuiamo nella "missione" del portale SECOLI tre aree critiche:

- comunità mondiale di utenti;
- suddivisione in un *front end* (esposto su rete Internet, rivolto al connazionale) e in un *back end* (esposto su RIPA, accessibile esclusivamente dagli operatori consolari);
- interazione complessa con altri servizi e piattaforme MAE, in particolare con il *Sistema Integrato Funzioni Consolari* (SIFC).

Ad esse corrispondono i seguenti requisiti di progetto:

- continuità del servizio e disponibilità del portale 24 ore su 24, 7 giorni su 7. L'estensione territoriale e sociale dell'utenza non consente l'individuazione di fasce orarie privilegiate, ma, viceversa, un disservizio che interessi le ore notturne a Roma comporterebbe un costo elevato nelle Americhe o in Asia non soltanto in termini operativi, ma anche e soprattutto di immagine; *continuità del servizio*
- resistenza ai guasti hardware e software (*fault tolerance*) e predisposizione ad un accesso intenso con potenziali picchi; *resistenza ai guasti*
- sicurezza dei sistemi e dei dati; *sicurezza*
- capacità di crescita e di adattamento alle condizioni di carico (*scalability*). Il progetto, dal canto suo, deve contemplare sin dall'inizio spazi di crescita, in forma tale che sia possibile, in ogni momento, intervenire ad incrementare la potenza di calcolo (mediante aggiunta di nuovi nodi fisici o virtuali) senza *scalabilità*

ledere la normale operatività del portale e senza necessità di ridisegnare l'architettura;

- interconnessione ben definita con altre piattaforme; *integrazione*
- semplicità di gestione.

Come ottenere il rispetto di questi requisiti-cardine?

- Ripartizione chiara in moduli indipendenti;
- ridondanza e bilanciamento di ogni unità funzionale; *ridondanza*
- punto unico d'accesso, in cui concentrare le misure di prevenzione degli attacchi;
- adozione di tecnologie e protocolli standard e aperti (art. 68 del *Codice dell'Amministrazione Digitale*);
- virtualizzazione.

Dal punto di vista della gestione, le condizioni che rendono un portale come SECOLI adeguatamente fruibile dall'utente equivalgono alle più generali premesse su cui riposa la fornitura affidabile, efficace e ininterrotta di un servizio.

Quest'ultima dipende dai seguenti parametri:

- un buon progetto, che tenga da conto appropriati standard e buone prassi, consolidate nel tempo e confermate dall'esperienza; *best practices*
- la stesura di procedure che coprano tutti gli aspetti operativi, dalla normale manutenzione alle condizioni di emergenza e al *disaster recovery*, e di protocolli che assicurino la continuità operativa. Si noti che parte integrante di questo aspetto è la puntuale e frequente verifica di tutte le procedure di ripristino, incluse naturalmente quelle che coinvolgono le politiche di backup e restore dei dati;
- un approccio olistico alla sicurezza, parte integrante del quale è l'adozione di formati aperti di interscambio dei dati (si tengano anche presenti le linee guida normative);
- una gestione che faccia sue le direttive stabilite in sede di progetto nonché durante i successivi momenti di messa a punto e di aggiornamento;
- una condivisione libera e piena della conoscenza all'interno del gruppo sistemi, dell'intero progetto SECOLI e dell'Amministrazione.

Questi i criteri che informano le presenti linee guida progettuali. Lo spirito che ad essi è sotteso dovrà tuttavia costituire la guida privilegiata anche nelle fasi successive, e in ogni momento della vita del progetto. Fermo restando, naturalmente, il riconoscimento pieno di una perfettibilità del disegno iniziale e della necessità che al confronto continuo con la pratica si accompagni un'attività mai interrotta di revisione.

2 PROSPETTIVE

2.1 CLUSTER E RIDONDANZA

ridondanza a livello applicativo

OVE POSSIBILE, sarà da preferirsi una ridondanza realizzata a livello applicativo o di servizio anziché di nodo. I vantaggi di quest'approccio sono infatti non indifferenti:

- flessibilità, granularità e controllo maggiori: dal che, a sua volta, segue la possibilità di implementare efficacemente strategie proattive di intervento a fronte di crash;
- misure di *server consolidation*, ossia di ottimizzazione delle risorse in modo da ridurre il numero complessivo di server impiegati.

La ridondanza è assicurata mediante:

- la coesistenza di istanze multiple dello stesso applicativo, non necessariamente corrispondenti a nodi virtuali o fisici distinti;
- tecnologie native, specifiche dei singoli applicativi, che garantiscano ove possibile la cooperazione tra le diverse istanze;
- impiego di un *modulo* o *componente di smistamento*, al quale si chiede di instradare le richieste utente verso una tra le molteplici istanze disponibili. Tale componente provvederà ad assegnare le richieste in base a criteri di bilanciamento del carico, e potrà implementare, ove richiesto, restrizioni di accesso, oltre naturalmente ad assicurare l'alta disponibilità dell'insieme.

modulo di smistamento

Oracle RAC

IN PARTICOLARE, si farà riferimento alla tecnologia nativa *Oracle RAC* per implementare la ridondanza e l'alta disponibilità nell'ambito del database.

2.2 HUB DATI

ex "Replichetta"

L'ATTUALE database ausiliario "Replichetta", rispondendo alla sua vocazione originaria, s'avvia ad assumere nel nuovo quadro strutturale il ruolo di collettore ed accentratore di ogni flusso dati tra SECOLI ed altri servizi MAE, ad eccezione di quelli gestiti dall'*Enterprise Service Bus*. Si estende e generalizza, dunque, la primitiva funzione di semplice contenitore dei dati SIFC.

Un passo in questa direzione è stato già compiuto, peraltro, con l'allacciamento al database EmbAddress.

integrazione con EmbAddress

Opportune misure potranno poi predisporsi per assicurare una tracciabilità dei dati di transito.

2.3 SISTEMA GEOGRAFICO CONSOLARE

Il Sistema Geografico Consolare, o SCE, è un elemento dell'architettura SECOLI vitale tanto per la normale operatività dell'applicativo, quanto per il dialogo con il SIFC. Con la designazione SCE s'intende, *in primis*, il database contenente l'estensione territoriale delle circoscrizioni consolari, a cui l'applicativo SECOLI fa riferimento per:

database SCE

- determinare la sedi di competenza in vari contesti;
- trattare la codifica numerica del luogo di nascita, di residenza, di matrimonio, e così via, indispensabile per l'interoperabilità con il SIFC.

Il database SCE viene fornito dal gruppo Anagrafe, nel corso di aggiornamenti periodici che coinvolgono simultaneamente il portale SECOLI e i server SIFC decentralizzati delle sedi estere, sotto forma di una collezione di file binari *SQLite*; in un formato, dunque, eterogeneo rispetto a quello preferenziale (Oracle), e tale da richiedere l'intervento di un componente software specializzato: il *gateway SCE*.

gateway SCE

2.4 DATABASE DI STAND-BY

L'applicazione dei princìpi progettuali che ispirano questo documento, se condotta ad un livello ulteriore, si evolve in modo naturale nel concetto di *database di stand-by*. Quest'ultimo è una copia funzionalmente identica del database di esercizio (a cui spetterà la qualifica di *primario*), quiescente in condizioni ordinarie, ma pronta a subentrare in ogni momento, e a tutti gli effetti, al primario al verificarsi di eventi traumatici oppure, più semplicemente, durante attività di manutenzione che richiedano un fermo prolungato.

L'alta disponibilità, la protezione dei dati e le capacità di *disaster recovery* risultato estremamente potenziate dalla presenza di un database replicato; non solo, ma quest'ultimo è l'unica misura progettuale che possa mantenere entro limiti accettabili e controllati il *Mean time to recover* (MTTR), ovvero il tempo necessario per il ripristino di una disponibilità operativa anche a fronte di gravi emergenze.

In termini operativi, la realizzazione di un database replicato deve rispondere ai seguenti requisiti:

- assicurare un aggiornamento costante e fluido della replica durante la normale operatività del sito principale;
- disporre di procedure di scambio rapido (*switch*) tra il database primario e quello di stand-by.

Data Guard

Considerazioni legate al costo delle licenze dovranno intervenire, in sede di valutazione, a guidare nella scelta tra un meccanismo di replica nativo, che faccia uso cioè degli strumenti propri del prodotto Oracle (la tecnologia *Data Guard* è l'unica in grado di mantenere una copia transazionalmente coerente del database primario), e soluzioni per così dire "artigianali", che dovranno tuttavia essere concepite, se mai, in modo da ridurre al minimo il rischio di inconsistenze.

Per quanto riguarda le risorse hardware necessarie, esse potranno essere agevolmente ricavate dalla dismissione dell'attuale infrastruttura; ovviamente, va previsto, per il sito di stand-by, un carico di lavoro conseguente ad una attività assai limitata nel tempo.

2.5 MONITORAGGIO

monitoraggio proattivo

L'architettura proposta non potrà dirsi completa senza la scelta di uno strumento di monitoraggio che assista lo staff sistemistico nella pronta percezione dei problemi, nella diagnosi e ricerca dell'eziologia e, nella misura del possibile, anche in un intervento attivo. Ad un apparato del genere si chiederà, in altre parole, di dotarsi di una "intelligenza" che gli consenta di incrementare la proattività del sistemista, ossia la capacità di prevenire ed anticipare i problemi e, più in generale, i bisogni futuri.

Individuiamo i seguenti requisiti come preferenziali:

- possibilità di attingere ad un vasto patrimonio di moduli aggiuntivi (*plugin*);
- strumenti di amministrazione che consentano uno sguardo olistico sullo stato dell'intera infrastruttura;
- supporto per il monitoraggio di apparati di rete tramite protocollo SNMP;

monitoraggio di rete

- capacità di operare in una configurazione distribuita.

Per sua natura intrinseca, lo strumento di monitoraggio prescelto dovrà attestarsi al di fuori dell'infrastruttura oggetto della sua supervisione, di modo che un fallimento di questa non ne pregiudichi le capacità di notifica e di intervento.

3 AMBIENTI

3.1 ESERCIZIO

resilienza a carichi elevati

Va detto che non è possibile, ad oggi, fornire una stima certa del carico a cui sarà sottoposto il portale SECOLI. Va però anche osservato che l'utenza potenziale coincide con l'intera comunità italiana all'estero, e che picchi elevati di accesso potranno senz'altro verificarsi.

È importante, pertanto, dotarsi di un'infrastruttura che sia facilmente scalabile, ossia tale da permettere l'accrescimento delle sue risorse (aggiunta di ulteriori nodi), se necessario, con il minor sforzo possibile.

struttura dell'ambiente

Si propone, pertanto, la seguente articolazione iniziale:

- due nodi oppure quattro nodi Oracle Database in configurazione RAC, secondo la disponibilità delle licenze, da ripartirsi idealmente su due macchine fisiche distinte;
- almeno quattro nodi per application server di front end, ripartiti su almeno due macchine fisiche distinte;
- *idem* per l'application server di back end;
- web server Apache o Nginx di front end, in cluster;
- web server Apache o Nginx di back end, in cluster;
- almeno quattro nodi per il gateway SCE, con due nodi di smistamento in cluster;
- Enterprise Service Bus;
- Hub dati.

3.2 TEST E COLLAUDO

importanza di una buona corrispondenza con l'esercizio

L'ambiente di collaudo dovrà essere per quanto possibile vicino a quello di esercizio: non in termini di dati, ovviamente, quanto piuttosto di struttura.

Nella pratica corrente si commette sovente l'errore di rinunciare ad un'aderenza piena (il più delle volte per questioni legate ai costi di messa in opera e gestione); noi, tuttavia, insistiamo sulla necessità di

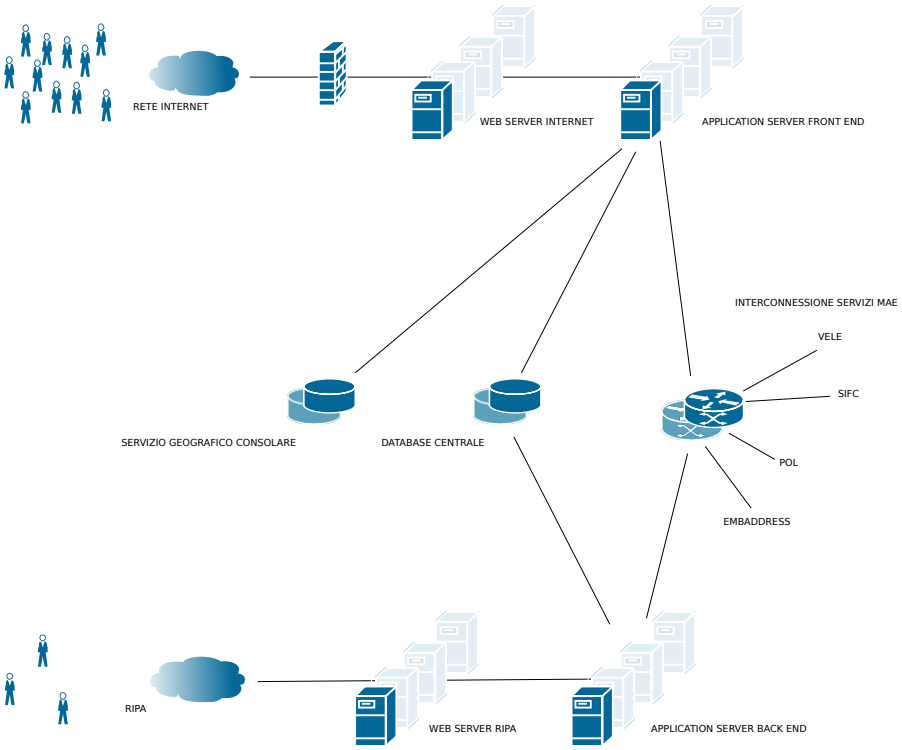


Figura 1: Ambiente di esercizio SECOLI

mantenere anche nel collaudo in livello almeno minimo di ridondanza, al fine di evitare che problematiche legate intrinsecamente ad un ambiente distribuito non si evidenzino che in esercizio.

RAC di collaudo

Per quanto riguarda la base dati, sarà preferibile ricorrere ancora ad una configurazione di tipo RAC (ove lo permettano le licenze). La complessità aggiuntiva derivante dalla necessità di gestire un secondo ambiente RAC, accanto a quello di esercizio, verrà ampiamente controbilanciata dal vantaggio di un passaggio più fluido e sicuro dal collaudo alla produzione.

struttura dell'ambiente

Si propone la struttura che segue:

- due nodi Oracle database in RAC;
- due nodi per l'application server di front end;
- due nodi per l'application server di back end;
- web server di front end (singola istanza);
- web server di back end (singola istanza);
- due nodi per il gateway SCE;

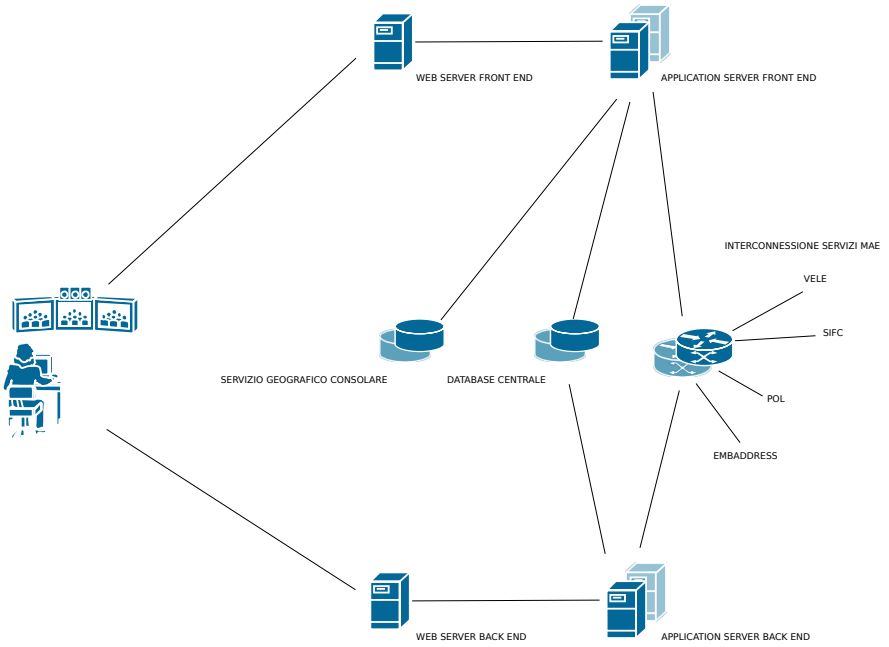


Figura 2: Ambiente di collaudo SECOLI

- singola istanza ESB;
- Hub dati.

I due nodi per il gateway SCE, l'istanza ESB e l'Hub dati potranno essere condivisi con l'ambiente corsi e con quello di sviluppo.

3.3 CORSI

specificità dell'ambiente corsi

L'AMBIENTE CORSI possiede una sua individualità, situata a metà strada sul continuum che corre dal collaudo all'esercizio. Esso offre un servizio effettivo ad una comunità di utenti distinti dal gruppo di sviluppo (gli operatori consolari chiamati ad esercitarsi), sicché non viene meno, pur scemandone il rigore, l'imperativo di evitare indisponibilità prolungate. Va tenuto presente quanto segue:

- i dati trattati nell'ambiente corsi sono fittizi, ragion per cui non sarà necessario mettere in atto tutto il complesso delle misure di backup tipiche dell'ambiente di esercizio;
- stante il carico di lavoro relativamente modesto, si potrà rinunciare ad implementare un'infrastruttura ridondata;

infrastruttura non ridondata

- per quel che riguarda l'applicativo vero e proprio, la versione deployata in ambiente corsi dovrà coincidere esattamente con

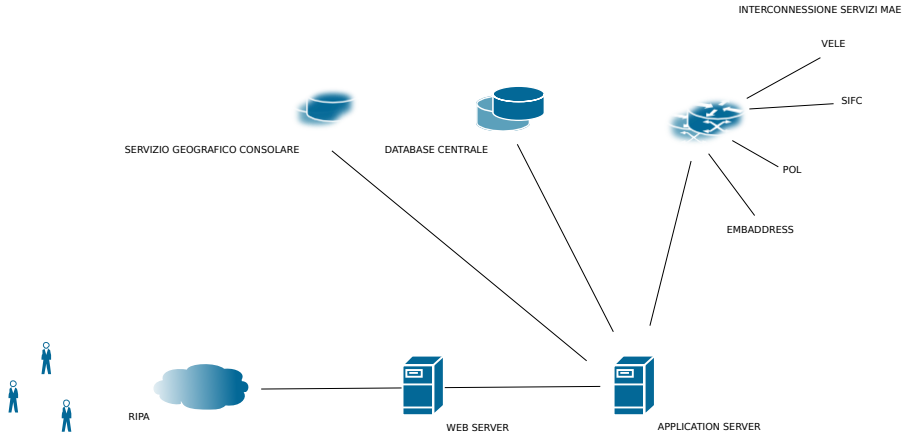


Figura 3: Ambiente corsi

quella in esercizio, salvo speciali esigenze. Un medesimo automatismo di deploy potrà eventualmente farsi carico di entrambi gli ambienti.

Si potrà costruire l'ambiente corsi come segue:

- database (a singola istanza);
- application server di front end e web server di front end, ad istanze singole su un nodo condiviso;
- application server di front end e web server di back end, ad istanze singole su un nodo condiviso;
- il gateway SCE, l'istanza ESB e l'Hub dati potranno essere condivisi con l'ambiente di collaudo.

3.4 SVILUPPO APPLICATIVO
SVILUPPO SISTEMI

L'ambiente di sviluppo è, tra tutti, quello che avanza pretese più modeste. Si predisporrà un pool di macchine destinate ad accogliere l'applicativo (web server, application server) e una macchina ospitante il database.

macchine per lo sviluppo sistemistico

Sarà poi indispensabile tenere da parte altre macchine da utilizzare per lo sviluppo sistemistico (ottimizzazione su database, sperimentazione di ambienti, redazione di script) nonché per installazioni SIFC.

3.5 GESTIONE

strumenti di gestione

Più che di un ambiente vero e proprio, si tratta di una collezione eterogena di macchine deputate a compiti specifici; tra queste (ma l'elenco non sarà esaustivo) menzioniamo:

- macchine per il backup: catalogo RMAN per il backup Oracle, server di backup Amanda;
- collettore dei log;
- repository del sistema di controllo versione Subversion;
- file server;
- gestore di progetto Redmine.

NOTA SUL DOCUMENTO

A questo documento diede LaTeX veste tipografica. Il numero di revisione è 1477.