Fonti dati per LDAP

Università di Modena e Reggio nell'Emilia

11 ottobre 2011

Introduzione

Queste slide introducono:

- i componenti del correlatore;
- censimento delle fonti dati;
- la struttura di un record ldap;
- descrizione del processo di inserimento di un nuovo record ldap;
- elenco degli eventi che permettono l'aggiornamento di un record.

I componenti del correlatore – introduzione

A partire da un'idea di Giova Faglioni, si è scelto di affidare la gestione delle identità a un sistema di messaggistica asincrona, che offre i vantaggi:

- disaccoppiamento degli elementi;
- funzionamento in tempo reale.

Il sistema precendente funzionava con la logica dello scambio di file.

I componenti del correlatore

Il correlatore è composto da tre elementi legati a un sistema di messaggistica asincrona (Apache ActiveMQ e Apache Camel): l'output di un compomente è un messaggio che viene consumato dal componente successivo.
I componenti sono:

- raccoglitore;
- correlatore (propriamente detto);
- delta.

Raccoglitore

Il raccoglitore: fa le query sql con chiave il codice fiscale e crea un file yaml (un formato a tag tipo xml ma più amichevole per gli umani) con tutti i dati di un certo utente;

Un file yaml

```
cf: MLVFNC69H12B819Z
contacts:
  -
    MAIL:
      - francesco.malvezzi@unimore.it
    UNIMOREMAILPERSONALE:
      - francesco.mlvz@gmail.com
    UNIMORETELEPHONENUMBER:
      - "39 059 2058018 "
    USERNAME:
      - malvezzi
usernames:
  -
    DATACAMBIOPASSWORD:
      - "1312372929"
    TYPE:
      - "Main Userid"
    UIDNUMBER:
      - "41312"
    USERNAME:
      - malvezzi
    USERPASSWORD:
      - "{SSHA}DYweocwnlCG55BTWA02va8+zqavoNSpw"
```

Il correlatore

Il correlatore è un programma che applica le logiche di composizione dei dati e crea degli ldif (sono i record ldap in formato testo):

```
dn: uid=malvezzi,ou=people,dc=unimore,dc=it
sambantpassword: EC9BAD0CB2B883A7843AE306B9A8A13A
uidnumber: 41312
userpassword: {SSHA}DYweocwnlCG55BTWA02va8+zqavoNSpw
givenname: Francesco
sn: MALVEZZI
unimorecodicefiscale: MLVFNC69H12B819Z
unimoredipmatricola: 020954
unimorediptessera: 170954
mail: francesco.malvezzi@unimore.it
unimoretelephonenumber: 39 059 2058018
OU: people
OU: Dipendenti
[...]
```

Delta

Il Delta verifica se ci sono differenze tra gli ldif ed i dati in ldap e produce i file di diff (sono i comandi ldap), che inoltre applica su ldap2-master.

```
# uid=171403,ou=people,dc=unimore,dc=it
dn: uid=171403,ou=people,dc=unimore,dc=it
changetype: modify
add: unimoremail
unimoremail: 171403@studenti.unimore.it
-
add: unimoremailprincipale
unimoremailprincipale: 171403@studenti.unimore.it
-
add: mail
mail: 171403@studenti.unimore.it
```

Le fonti dati del raccoglitore

- usernames: username, dati posix, password e scadenza (`CORRELA_USERNAME` su `cesia`);
- csa: contiene i dati dalle Risorse Umane (`CORRELA_CSA` su `cesia`);
- esterni: cioé il personale registrato con identity (`CORRELA_ESTERNI` su `cesia`);
- contacts: numeri di telefono (`CORRELA_CONTACTS` su `orasia`);
- mail: gli indirizzi di posta che provengono direttamente dal cluster di posta (query diretta al cluster della posta);
- entitlements: le autorizzazioni che shibboleth usa per alcune applicazioni (query diretta a `cesia`);

continua . . .

Le fonti dati del raccoglitore – 2

- account: gli account registrati con identity (`CORRELA_ACCOUNT` su `cesia`);
- registered: trova solo i registrati e i pre-immatricolati(`CORRELA_REGISTERED` su `cesia`);
- exalumni: trova solo gli studenti che hanno conseguito il titolo da più di tre anni (`CORRELA_EXALUMNI` su `cesia`);
- esse3: trova solo gli studenti con una posizione attiva (`CORRELA_ESSE3` su `cesia`).

Struttura di un record ldap

Ogni record ldap si compone di una parte generale di anagrafica della persona (nome, cognome, password, mail ecc.), più una o più parti relative agli incarichi.
Gli attributi della parte degli incarichi sono prefissati da un indice tra graffe:

```
unimoredipafferenzadidattica: {1}Non assegnato
unimoredipattivita: {1}In servizio
unimoredipcodiceattivita: {1}0001
unimoredipcodiceruolo: {1}ND
unimoredipcodicesettore: {1}000000000000
unimoredipruolo: {1}Personale tecnico amministrativo
```

Per gli studenti al posto degli incarichi c'è una sezione per ogni posizione attiva.

Logica di aggregazione

La logica di aggregazione dei dati è di formare per ogni codice fiscale:

- max un record da studente (al momento la username è numerica): se uno studente è iscritto a più corsi ha più ruoli. I dati delle stored procedure `registered` e `exalumni` sono scartati se uno studente ha una posizione attiva;
- max un record da dipendente (al momento la username è non numerica): qui si sommano i ruoli da dipendente e quelli da esterno;
- quanti account si vuole.

Inserimento di un nuovo utente

- Per i dipendenti e gli account il correlatore si aspetta di avere dei dati completi, cioé nessun utente entra in ldap se non ha almeno un incarico e una username (la username si sceglie con la scelta dell'username). Se non li ha, crea un ldif di cancellazione.
- Per gli studenti il correlatore si aspetta che possano arrivare dati incompleti della username. In questo caso crea un messaggio con i dati necessari e lo invia all'applicazione `registra_username`, che dopo la creazione della username manda un messaggio al raccoglitore.

Eventi che scatenano l'aggiornamento di un utente

I dati di un utente sono aggiornati quando il suo codice fiscale passa per il correlatore, cosa che accade con:

- `epi_replica`: la tabella di esse3 che contiene i dati degli utenti che hanno subito qualche cambiamento viene monitorata con un polling ogni 10 secondi;
- modifica dei dati di un utente (scelta della username, cambio mail o telefono, cambio password) – le applicazioni mandano un messaggio immediatamente;
- interventi amministrativi: (`check_all` su identity);
- eventi periodici: una specie di censimento fa passare tutti i dipendenti, tutti gli esterni e una parte degli stduenti attivi per il correlatore tutte le notti.

Conflitti tra le fonti dati

Se il correlatore dispone sia di un record studente che di un record dipendente/esterno, per tutti crea la parte generale (nome e cognome, indirizzo ecc.) a partire del record studente.

Riferimenti

- Enterprise Integration Patterns: `http://www.eaipatterns.com/`, ISBN-13: 978-0321200686