

Progetto “Centro di elaborazione”

Lab. di Basi di Dati - A.A. 2014-2015

Prof. P. Pierini

Premessa

La specifica informale del problema data nei paragrafi seguenti è, come in ogni caso reale, incompleta e, in alcuni punti, ambigua o contraddittoria. Lo studente dovrà quindi raffinare e disambiguare le specifiche simulando l'interazione con il committente. In alcuni casi allo studente sarà richiesto di valutare diverse possibili alternative, per poi sceglierne una in maniera motivata. Le motivazioni di tutte le scelte interpretative, progettuali e implementative andranno sempre chiaramente documentate nel progetto e verranno discusse in sede di revisione.

Problema

Il progetto prevede la realizzazione della base di dati per la gestione di un centro di elaborazione dati.

Il sistema è strutturato come un insieme di apparati di vario genere (server, data storage, switch, router, ups, impianti di raffreddamento) allocati all'interno di rack. Ogni apparato è caratterizzato da un codice di “inventory” univoco, da una serie di dati tecnici quali il produttore, il modello, il tipo di alimentazione e informazioni sull'equipaggiamento hardware.

E' necessario determinare, per ciascun apparato, in quale rack è allocato, in quale posizione all'interno del rack e quante posizioni sono complessivamente occupate (si consideri che ogni rack è composto di 42U “rack unit”).

I locali del centro di elaborazione sono organizzati con 10 corridoi paralleli e numerati ordinatamente, con due file di 18 rack ciascuna ai lati destro e sinistro di ciascun corridoio, pertanto ogni rack è localizzabile da: il numero del corridoio dove è situato, il lato e la posizione nella fila.

Gli apparati sono connessi tra loro tramite cavi di alimentazione o di trasmissione, in rame o in fibra ottica, con specifiche caratteristiche standard. In particolare per la trasmissione si usano fibre ottiche multimodali con connettori SC o cavi in rame con connettori RJ45; per l'alimentazione si usano cavi H05VV-F con connettori C13. Per cui di ogni cavo bisogna conoscerne la tipologia, la tecnologia e lo standard con il quale è realizzato. Ciascuno cavo è identificato da una targhetta alfanumerica e connette una coppia di apparati. Bisogna anche conoscere in quali connettori, degli apparati connessi, ciascun cavo è inserito. Si consideri che le alimentazioni sono ridondate (i.e. duplicate) su ogni apparato così come le connessioni di trasmissione. Switch e router avranno un numero più elevato di connettori di trasmissione, così come gli UPS avranno un numero maggiore di connettori di alimentazione. Da notare che la rete di trasmissione è strutturata ad albero e prevede che soltanto gli switch ed i router possano essere interconnessi tra loro; le altre tipologie di apparato sono connesse soltanto verso gli switch.

Su ogni apparato sono configurati uno o più servizi (ad esempio su un server può essere configurato un servizio “http”, su un altro un servizio email, etc..). Si vogliono archiviare i file di backup delle configurazioni di ciascun servizio per ogni apparato. Per ogni configurazione bisogna conoscere la posizione del file di backup nel data storage e, ovviamente, il servizio e l'apparato relativo.

Per ogni apparato si vogliono memorizzare anche le informazioni diagnostiche che vengono prodotte periodicamente con frequenza oraria. Ogni tipologia di apparato produce diagnostica specializzata in funzione dei servizi che vi sono configurati. In particolare:

- i server, gli switch ed i router producono informazioni sul volume di traffico scambiato espresso in byte/secondo,
- i data storage indicano il volume complessivo dei dati memorizzati espressi in byte,
- gli UPS indicano i chilowatt assorbiti dagli apparati che alimenta,
- gli impianti di raffreddamento ed i rack riportano le temperature delle aree calde e delle aree fredde dei corridoi.

Per ciascuna di queste grandezze, su ogni apparato, sono preconfigurati dei valori di soglia che, se superati, evidenziano delle criticità di funzionamento.

Infine, per i 40 operatori che lavorano nel centro, è necessario memorizzare la matricola, le informazioni anagrafiche, i codici di sicurezza per l'accesso ai sistemi relativamente a tre possibili attività:

- installazione e riconfigurazione fisica degli apparati,
- configurazione dei servizi,
- accesso ai dati diagnostici

Ogni operatore può essere autorizzato a compiere una o più attività e, per ogni attività, dovrà possedere un codice di sicurezza specifico.

Le operazioni previste sul data base, ripetute giornalmente, sono le seguenti:

1. Per ogni entità notevole individuata, fornire gli statement di inserimento/modifica/eliminazione delle informazioni corrispondenti.
2. Realizzare il censimento degli apparati elencando per ciascuno di essi, oltre ai dati tecnici, anche l'allocazione del rack che lo contiene e la posizione all'interno del rack.
3. Dato un router si vogliono determinare gli switch direttamente collegati.
4. Dato un router e un apparato ad esso connesso, si vuole determinare attraverso quali switch è realizzata la connessione (da notare che i percorsi sono duplicati e che gli switch coinvolti avranno in comune con l'apparato gli elementi di localizzazione ovvero corridoio o rack).
5. Si vuole determinare quali apparati sono alimentati da un dato UPS (fare attenzione che gli UPS avranno anche collegamenti di trasmissione).
6. Dato un apparato si vogliono conoscere gli apparati ad esso direttamente collegati, riportando la tipologia e gli identificativi dei connettori e dei cavi utilizzati.
7. Dato un servizio si vuole sapere su quali apparati è erogato.
8. dato un servizio su uno specifico apparato, si vuole recuperare la configurazione di backup.
9. Elencare giornalmente i superamenti dei valori di soglia sui dati statistici, indicando l'apparato e l'ora in cui il superamento è avvenuto .
10. Calcolare la crescita del volume di traffico su ciascun router rispetto al giorno precedente.
11. Su base giornaliera, calcolare quale UPS, in quale ora, ha erogato il maggior picco di potenza.
12. Elencare quali operatori sono abilitati all'installazione di apparati.
13. Ritrovare i dati di accesso di un dato operatore.
14. Dato un apparato ricostruire la cronologia di tutti gli interventi, indicando il tipo di attività e l'operatore che l'ha eseguita

E' possibile inserire operazioni e procedure di gestione addizionali che si ritengano utili.

Individuare le regole di vincolo e di derivazione che possono essere applicate ai contenuti di questa base di dati. L'individuazione dei vincoli e la loro implementazione costituiscono un requisito fondamentale per lo sviluppo di un progetto realistico, e ne verrà tenuto conto durante la valutazione finale.

Alcune delle funzionalità richieste potrebbero non essere realizzabili con singole query, ma richiedere l'uso di strumenti più avanzati messi a disposizione dal DBMS, come le procedure e/o i trigger.

Progetto

Lo sviluppo del progetto comprende i seguenti punti:

1. Progettazione concettuale consistente in:
 - a. Definizione formale, analisi e strutturazione dei requisiti.
 - b. Modellazione dello schema concettuale utilizzando il modello ER.
 - c. Formalizzazione delle regole di vincolo e di derivazione.
2. Progettazione logica consistente in:
 - a. Ristrutturazione ed ottimizzazione dello schema ER.
 - b. Analisi dei volumi e degli accessi.
 - c. Traduzione nel modello relazionale.
3. Implementazione del progetto.
 - a. Sviluppo del codice SQL
 - b. Implementazione di un'interfaccia grafica web-oriented (opzionale)

Tutte le fasi del progetto dovranno essere corredate da adeguata documentazione che illustri quanto viene realizzato e le scelte intraprese. In particolare, dovranno essere necessariamente inclusi nella documentazione gli schemi ER risultanti dai passi (1) e (2), debitamente commentati, il modello relazionale della base di dati ottenuto al passo (2c), e tutto il codice SQL usato nell'implementazione, comprendente gli statement necessari alla creazione della base di dati, alcuni statement che mostrino esempi di inserimento dati nel DB, e il codice di tutte le query/operazioni richieste.

Per l'implementazione del progetto si può usare un qualunque DBMS preventivamente concordato con il docente, le seguenti sono alcune scelte possibili:

- Oracle [Windows e Unix, commerciale]
- MS SQL Server [Windows, commerciale]
- DB2 [Windows e Unix, commerciale]
- PostgreSQL [Linux, gratuito]
- Interbase [Windows e Linux, la versione 6 è gratuita]
- Firebird [Windows e Linux, gratuito, basato sulla versione 6 di Interbase]
- MySQL [Windows e Linux, gratuito]

Nota: MSAccess è fortemente sconsigliato.

Il DB realizzato può essere dotato di un'interfaccia grafica scritta con un linguaggio di programmazione a scelta (ad. Es. Java, PHP). La valutazione NON terrà conto delle caratteristiche estetiche della grafica di interfaccia bensì dei criteri di interazione tra DBMS e linguaggio di programmazione.

Informazioni

Questa specifica è disponibile in formato PDF sulla pagina web del corso di Laboratorio di Basi di Dati, all'indirizzo <http://www.di.univaq.it/pierluigi.pierini/>.

Ulteriori informazioni e chiarimenti sulle specifiche possono essere richiesti direttamente via email all'indirizzo Pierluigi.Pierini@intecs.it.

Si ricorda che i progetti vanno svolti singolarmente o in piccoli gruppi (fino a 3 persone).

Tutti coloro che intendono sostenere l'esame di Laboratorio di Basi di Dati dovranno inviare l'elaborato del progetto, in formato PDF, via email all'indirizzo Pierluigi.Pierini@intecs.it, rispettando le scadenze di volta in volta indicate per ogni appello. In fase di discussione del progetto, si potrà presentare l'implementazione di una interfaccia grafica.