



UNIVERSITÀ DI PISA

CORSO DI LAUREA IN INGEGNERIA INFORMATICA

# PROGETTO DEL CORSO DI BASI DI DATI

*Antonio Querci, Giosuè Ratti*

ANNO ACCADEMICO 2022-2023

# Indice

1– Introduzione	3
2– Progettazione concettuale	4
2.1 Glossario dei termini	4
2.2 Dizionario dei Dati	5
2.2.1 Tabella delle entità	5
2.2.2 Tabella delle relazioni	7
2.3 Descrizione aree	10
2.3.1 Area contenuti	10
2.3.2 Area formati	11
2.3.3 Area utenti	13
2.3.4 Area streaming	14
2.4 Tavola dei volumi	15
3– Ristrutturazione dello schema ER	17
3.1 Area utenti	17
3.2 Area formati	18
3.3 Area contenuti	20
3.4 Tabella delle entità	21
3.5 Tabella delle relazioni	23
3.6 Tavola dei volumi	24
4– Operazioni significative	26
4.1 Operazione 1	27
4.2 Operazione 2	28
4.3 Operazione 3	29
4.4 Operazione 4	33
4.5 Operazione 5	36
4.6 Operazione 6	37
4.7 Operazione 7	39
4.8 Operazione 8	40
5– Business rules e vincoli di integrità	42
5.1 Area contenuti	42
5.2 Area utenti	42
5.3 Area streaming	44
5.4 Area formati	44
5.5 Area geografica	44
6– Traduzione logica e normalizzazione	45
6.1 Area contenuti	45
6.1.1 Vincoli di integrità referenziale	45
6.1.2 Dipendenze funzionali	46
6.2 Area utenti	49
6.2.1 Vincoli di integrità referenziale	49

6.2.2 Dipendenze funzionali	50
6.3 Area formati	52
6.3.1 Vincoli di integrità referenziale	52
6.3.2 Dipendenze funzionali	52
6.4 Area geografica	53
6.4.1 Vincoli di integrità referenziale	53
6.4.2 Dipendenze funzionali	53
6.5 Area streaming	53
6.5.1 Vincoli di integrità referenziale	53
6.5.2 Dipendenze funzionali	53
6.6 Relazione generiche	54
6.6.1 Vincoli di integrità referenziale	54
6.6.2 Dipendenze funzionali	55
7– Operazione 2.4.4 (Raccomandazione di contenuti)	57
8– Caching	59
9– Operazione per il cambio di server	61
10– Area analytics	66
9.1 Classifiche	66
9.2 Bilanciamento del carico	67
9.3 Classifiche di gradimento dei piani di abbonamento	68

## Introduzione

La traccia del progetto richiede di progettare ed implementare una Base di Dati per un servizio di memorizzazione e streaming di film online. Il database inoltre memorizza informazioni circa gli utenti che interagiscono con il servizio e informazioni più di carattere tecnico come quelle legate ai formati e quella legati alla rete di server necessari per la distribuzione del servizio in rete.

# PROGETTAZIONE CONCETTUALE

## Glossario dei termini

Termine	Descrizione	Sinonimi	Collegamenti
Film	Il prodotto offerto dal sistema informativo di FilmSphere, portale per la memorizzazione e streaming di contenuti video online	Contenuto video	Utente, Formato, Regista, Attore, Rating
Rating	Valutazione complessiva del film, stabilita in base ai voti degli utenti e a fattori esterni	Valutazione	Film
Formato	Le varie specifiche con cui può essere disponibile il film.		Film, Area geografica
Utente	Il fruitore del servizio	Cliente	Film, Rating Connessione, Server
Piano di abbonamento	Le varie modalità in cui viene erogato il servizio		Area geografica, Utente
Server	Macchina connessa in rete che in questo caso invia un contenuto multimediale		Connessione, Utente, Area geografica
Area geografica	Una precisa zona del mondo		Formato, Piano di abbonamento, Connessione, Server e Formato
Connessione	La singola sessione in cui l'utente si collega al sito		Utente, Server, Area geografica

## Dizionario dei dati

### Tabella entità

Entità	Descrizione	Attributi	Identificatore
Film	Film disponibile sul sito	Titolo, durata, Anno di produzione, Paese di produzione, Descrizione, Genere, Lingua (1,N), Lingua sottotitoli (0,N), premio (0,N), Visualizzazioni, Rating Recensioni	Anno di produzione + titolo
Utente	Cliente del sito	Codice, Nome, Cognome, E-mail, Password	E-mail
Attore	Attore del film che può eventualmente avere vinto premi	Codice fiscale, Nome, Cognome. Popolarità	Codice fiscale
Regista	Regista del film	Codice fiscale, Nome, Cognome, Popolarità	Codice fiscale
Formato	Formati supportati dal sito di streaming. Ogni formato può avere una versione passata oltre a quella corrente	Codice, Larghezza rapporto, Altezza rapporto, Lunghezza risoluzione, Altezza risoluzione, Bitrate, Data rilascio, Qualità audio, Qualità video	Codice
Formato corrente	Il formato attualmente disponibile	Codice, Larghezza rapporto, Altezza rapporto,	Codice

		Lunghezza risoluzione, Altezza risoluzione, Bitrate, Data rilascio, Qualità audio, Qualità video	
Formato passato	Una delle versioni precedenti del formato	Codice, Larghezza rapporto, Altezza rapporto, Lunghezza risoluzione, Altezza risoluzione, Bitrate, Data rilascio, Qualità audio, Qualità video, Data fine validità	Codice
Area geografica	Alcuni piani di abbonamento e film possono avere limitazioni legati all'area geografica	Codice, IP iniziale, IP finale	Codice
Connessione	Ogni singola connessione, sia passata che in corso, dell'utente al sito di streaming	Inizio, indirizzo IP	Inizio + identificatore esterno utente
Connessione passata	Le connessioni passate fatte da un utente.	Inizio, Indirizzo IP, Fine, Film, Formato	Inizio + identificatore esterno utente
Connessione attuale	Connessione attualmente in corso di un utente che sta guardando un film	Dispositivo, Inizio, Indirizzo IP	Inizio + identificatore esterno utente
Fatturazione	Il singolo pagamento effettuato per acquistare un piano di abbonamento	Data, Scadenza, Fattura, Carta di credito	Fattura
Piano di abbonamento	Il tipo di abbonamento. Può avere limitazioni di età, geografica e alcune funzionalità esclusive	Pacchetto, Tariffa, Max ore, Max gigabyte, Caratterizzazioni (0,N), Limiti di età (0,N)	Pacchetto

Server	Il server tramite il quale avviene la connessione	Identificatore, Banda, Capacità, Off counter	Identificatore
Dispositivo	I vari dispositivi da cui può avvenire la connessione	Nome	Nome

### Tabella relazioni

Relazione	Descrizione	Entità coinvolte	Attributi
Direzione	Collega il regista con i film da lui diretti	Regista (0,N), Film (1,N)	/
Recitazione	Rappresenta se un attore ha recitato un film, il ruolo che aveva e se ha vinto dei premi per la sua recitazione	Attore (0,N), Film (1,N)	Principale, Premio
Visualizzazione attuale	Se un utente sta visualizzando un contenuto	Utente (0,1), Film (0, N)	/
Visualizzazione passata	Rappresenta il fatto che un utente ha visto un film e quante volte lo ha visto	Utente (0,N), Film (0,N),	Numero volte, Stelle
Disponibilità	Lo stesso film può essere disponibile in più formati	Film (0,N), Formato (0,N)	Dimensione file
Disponibilità passata	Un film può essere stato disponibile in una versione passata di un formato	Film (0,N), Formato passato (0,N)	Dimensione file
Pagamento	Collega un utente al pagamento che ha effettuato per un certo piano di	Utente (0,N), Fatturazione (1,1)	/

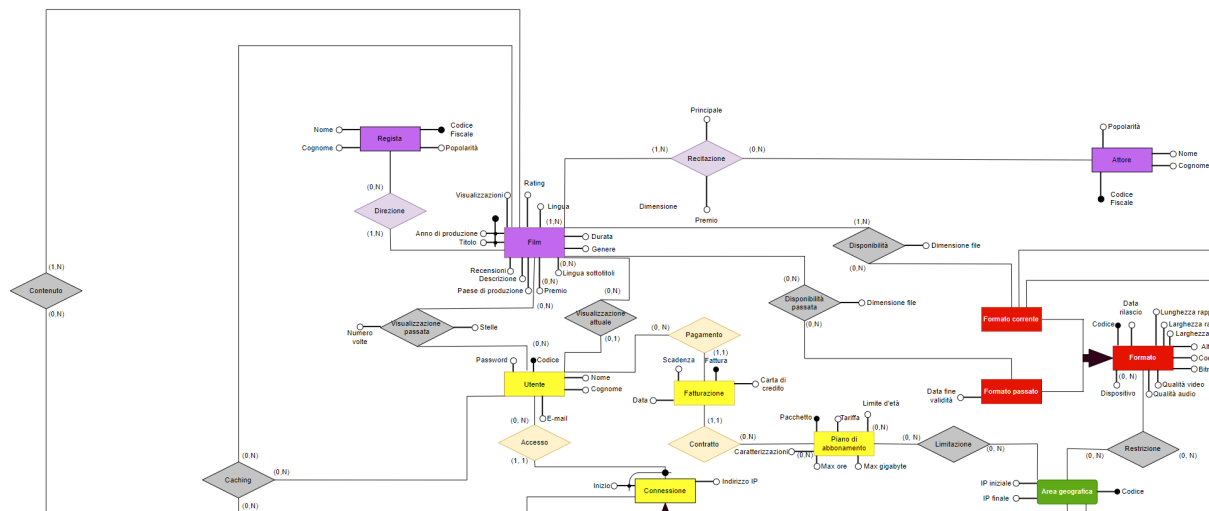


	abbonamento		
Contratto	Collega un certo pagamento al piano di abbonamento che è stato acquistato	Fatturazione (1,1), Piano di abbonamento (0,N)	/
Limitazione	Alcuni piano di abbonamenti possono avere limitazioni di tipo geografico	Piano di abbonamento (0,N), Area geografica (0,N)	/
Restrizione	Alcuni film potrebbero non essere disponibili in certe aree geografiche	Formato (0,N), Area geografica (0,N)	/
Accesso	Collega un utente ad ogni suo singolo connessione al sito	Utente (0,N), Connessione (1,1)	/
Locazione	Ogni server si trova in una precisa area geografica	Server (1,1), Area geografica (0,N)	/
Sessione	Ogni connessione avviene ad un preciso server, non necessariamente il più vicino fisicamente	Connessione attuale (1,1), Server (0,N)	/
Sessione Passata	Memorizza a quale server sono avvenute le connessioni passate	Connessione passata (1,1), Server (0,N)	/
Caching	Alcuni film, i preferiti dall'utente, possono essere disponibili in cache presso alcuni server	Utente (0,N), film (0,N), Server (0,N)	/
Contenuto	Un film è contenuto solo in alcuni server	Film (1,N), Server (0,N)	/

Distanza	La distanza tra un server e un'area geografica	Server (0,N), Area geografica (1,N)	Kilometri
Supporto	Indica da che tipo di dispositivo avviene la connessione	Dispositivo (0,N), Connessione (1,1)	/
Ideale	Ogni dispositivo ha un formato ideale per la riproduzione di contenuti	Formato corrente (0,N), Dispositivo (1,1)	/
Formato visualizzato	Collega ogni connessione con il formato del film che sta venendo visualizzato	Formato corrente (0,N), Connessione attuale (1,1)	/

## DESCRIZIONE AREE

### Area contenuti



Ogni film è identificato dal titolo e dall'anno di produzione. Un film può essere disponibile in più lingue e anche i sottotitoli possono essere disponibili in più lingue. Ogni film ha una durata, un paese di produzione, un genere e una descrizione. Le informazioni riguardanti gli attori che hanno recitato nel film sono memorizzate tramite l'entità *Attore*. La relazione *Recitazione* ha due attributi: Principale, che descrive il ruolo dell'attore nel film, e Premio, ad indicare se l'attore ha ottenuto qualche riconoscimento per la sua prestazione.

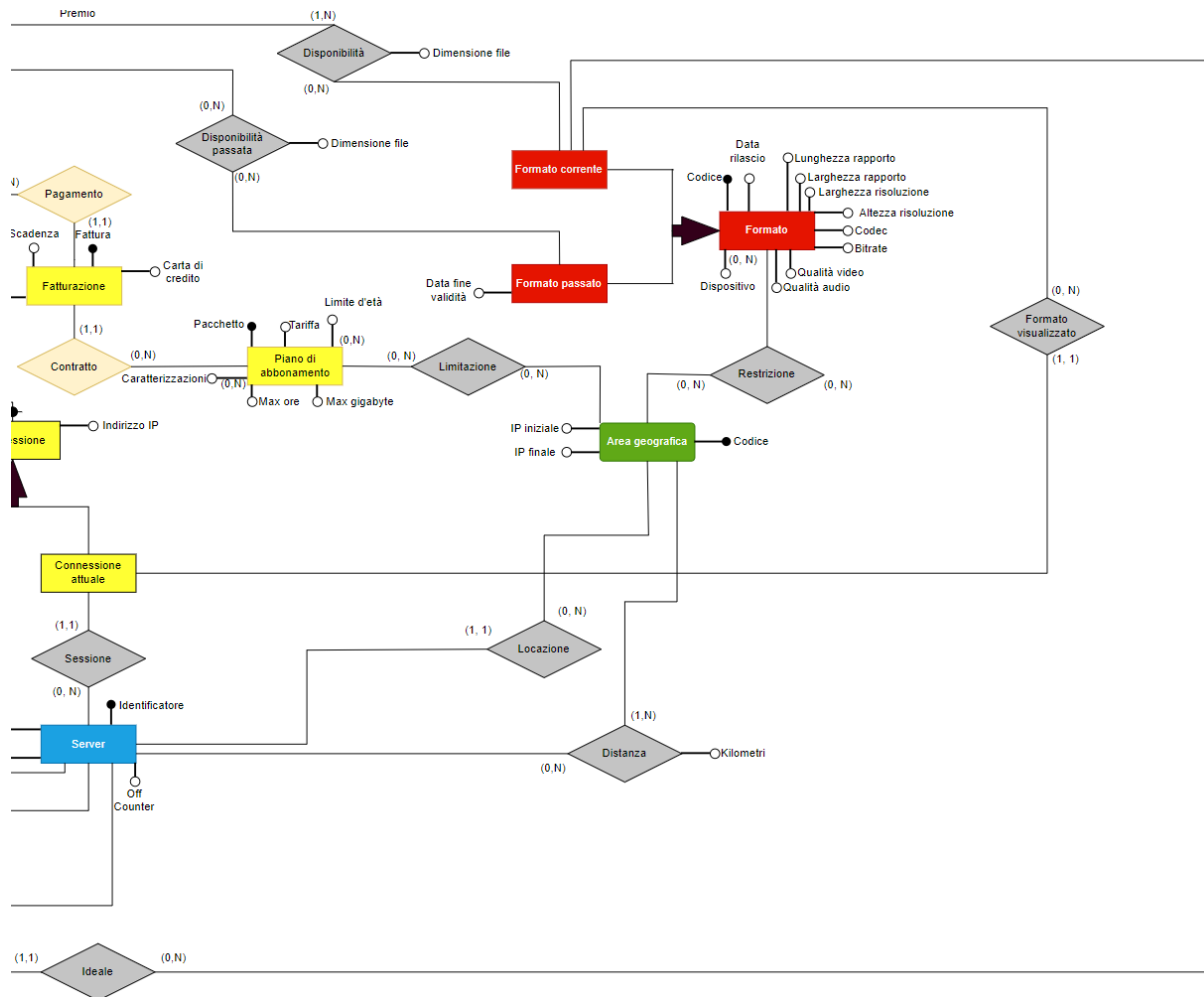
Le informazioni riguardanti i registi sono anch'esse memorizzate tramite un'entità *Regista*, che possiede l'attributo Popolarità, che servirà a calcolare il rating del film. Ogni film infatti, possiede un rating rappresentato dall'attributo *Rating*, calcolato come media aritmetica di vari fattori, tra cui per l'appunto popolarità dei registi. Infine ogni film può possedere anche un altro tipo di valutazione, ovvero quelle degli utenti, che però sono attribuite in base al gusto personale (si assegna tramite stelle) e non calcolate matematicamente.

Il sistema inoltre memorizza per ogni utente tutti i film visualizzati e il numero di volte che sono stati visti, un'informazione chiave per il meccanismo di raccomandazione contenuti.

Per ogni utente, il film preferito e un film raccomandato dal sistema sulla base dei contenuti che ha visualizzato sono disponibili in caching presso due server in modo da migliorare l'esperienza di visione.

Uno stesso film può essere visualizzato da più utenti in contemporanea, ma un utente può visualizzare un solo film alla volta.

## Area formati

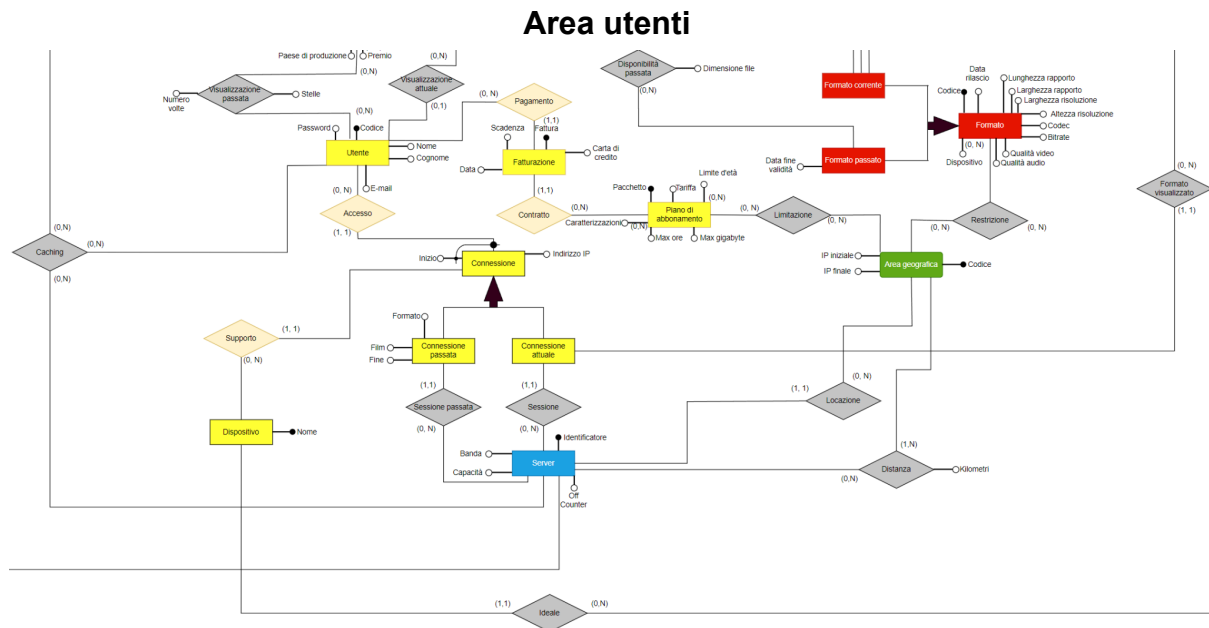


I film sono disponibili in vari formati, ognuno dei quali è identificato da un codice univoco. Ogni formato ha poi un insieme di caratteristiche tecniche. La relazione *Restrizione* descrive il fatto che alcuni formati potrebbero non essere disponibili in determinate aree geografiche.

La generalizzazione è necessaria per tenere traccia delle modifiche ai formati esistenti (storizzazione di entità), sebbene i film possono essere disponibili solamente nei formati correnti, motivo per cui la relazione *Disponibilità* si collega con *Formato corrente*.

Il formato con cui un utente connesso attualmente sta guardando un film nel presente è memorizzato nel database tramite la relazione *Formato visualizzato*.

Quando la connessione termina, questo dato sarà salvato nell'attributo Formato dell'entità *Connessione passata*.



Ogni utente che utilizza il servizio è identificato univocamente da un codice; in più, per poter accedere al servizio, un utente deve fornire anche alcune informazioni personali, come una sua e-mail (che identifica univocamente l'utente), il suo nome, il suo cognome e anche una password.

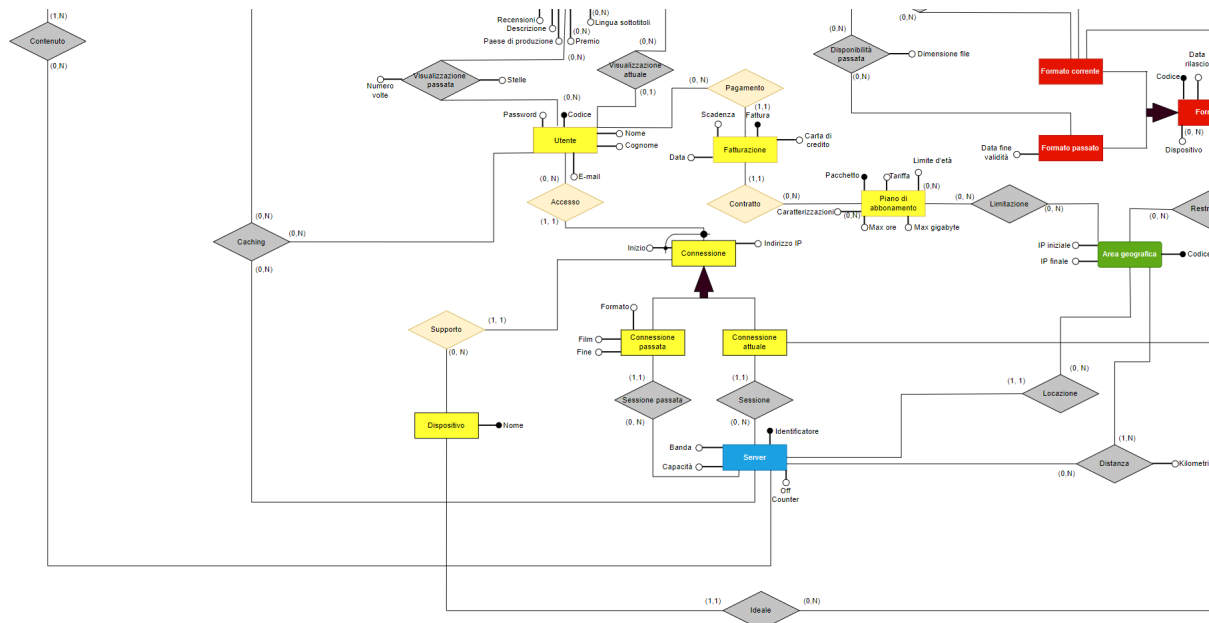
Per usufruire del servizio, è necessario che l'utente si abboni scegliendo un piano disponibile tra quelli forniti dall'entità Piano di abbonamento.

Ogni singola volta in cui l'utente rinnova l'abbonamento questo atto viene memorizzato nell'entità Fatturazione, la quale è identificata univocamente da un codice, indicato con l'attributo *Fattura*.

L'attributo "Scadenza" di Fatturazione serve a indicare, di norma, la data di scadenza, però potrebbe anche indicare il fatto che un abbonamento possa essere stato disdetto. In più, se un abbonamento non è stato disdetto, permette di risalire più veloce alla data in cui questo scadrà (evitandoci di dover passare ogni volta da Contratto e Piano di abbonamento). Poiché non necessariamente la scadenza è esattamente un mese dopo la data della fatturazione, l'attributo non è ridondante.

Un utente, al fine di sfruttare quanto offerto, deve collegarsi al servizio tramite un suo dispositivo. Ogni qual volta l'utente si collega al servizio, questa connessione viene memorizzata tramite l'entità connessione, la quale distingue tra le connessioni in corso e quelle passate. Nelle connessioni passate sono memorizzati anche il film visualizzato e il formato in cui è stato visualizzato, informazioni necessarie per il meccanismo di Classifiche nell'area Analytics 2.6.1.

## Area streaming



Il sito si serve di una rete di distribuzione di contenuti costituita da server, per diminuire la distanza fisica tra l'utente e il server con il contenuto richiesto in modo da velocizzare il servizio.

Ogni server è descritto da un identificatore, dall'area geografica in cui si trova, dalla banda e dalla capacità. In ogni momento il server si trova in uno stato determinato dal traffico di dati attualmente in corso sul server. È disponibile un sistema di caching che coinvolge film, utenti e server. Sostanzialmente, nei server fisicamente più vicini ad un utente, definiti *edge server*, possono essere disponibili i film preferiti dal cliente, al fine di velocizzare lo streaming. Ogni connessione avviene ad un server preciso, quello considerato più "conveniente" in base alla distanza fisica e al traffico attuale; durante la riproduzione questo server può cambiare, la nuova connessione avverrà al nuovo server più "conveniente". Inoltre, i server possono trovarsi momentaneamente in stato OFF, nel caso ci siano troppe persone connesse al server e non si può garantire un servizio di livello soddisfacente. Viene anche memorizzato il numero di volte in cui ogni server è in stato OFF, informazione necessaria per il meccanismo di bilanciamento del carico dell'area Analytics 2.6.2.

**Tavola dei volumi**

Concetto	Tipo	Volume
Film	E	200
Utenti	E	1000
Formato	E	60
Formato Passato	E	45
Formato Corrente	E	15
Area Geografica	E	30
Connessione	E	25.000 (25*Utenti)
Connessione Attuale	E	500
Connessione Passata	E	24.500 (25*Utenti - 500)
Server	E	40
Piano di abbonamento	E	5
Fatturazione	E	3000 (3*Utente)
Regista	E	50
Attore	E	500
Dispositivo	E	20
Visualizzazione Passata	R	30.000 (30*Utente)
Visualizzazione Attuale	R	500
Direzione	R	220
Recitazione	R	1200 (6*Film)
Disponibilità	R	1600 (8*Film)
Restrizione	R	120 (2*Formato)
Accesso	R	25.000 (25*Utente)
Sessione	R	500
Sessione Passata	R	24.500 (25*Utenti - 500)
Distanza	R	1200



Contenuto	R	20
Caching	R	4.000 (4*Utente)

### Osservazioni

In media:

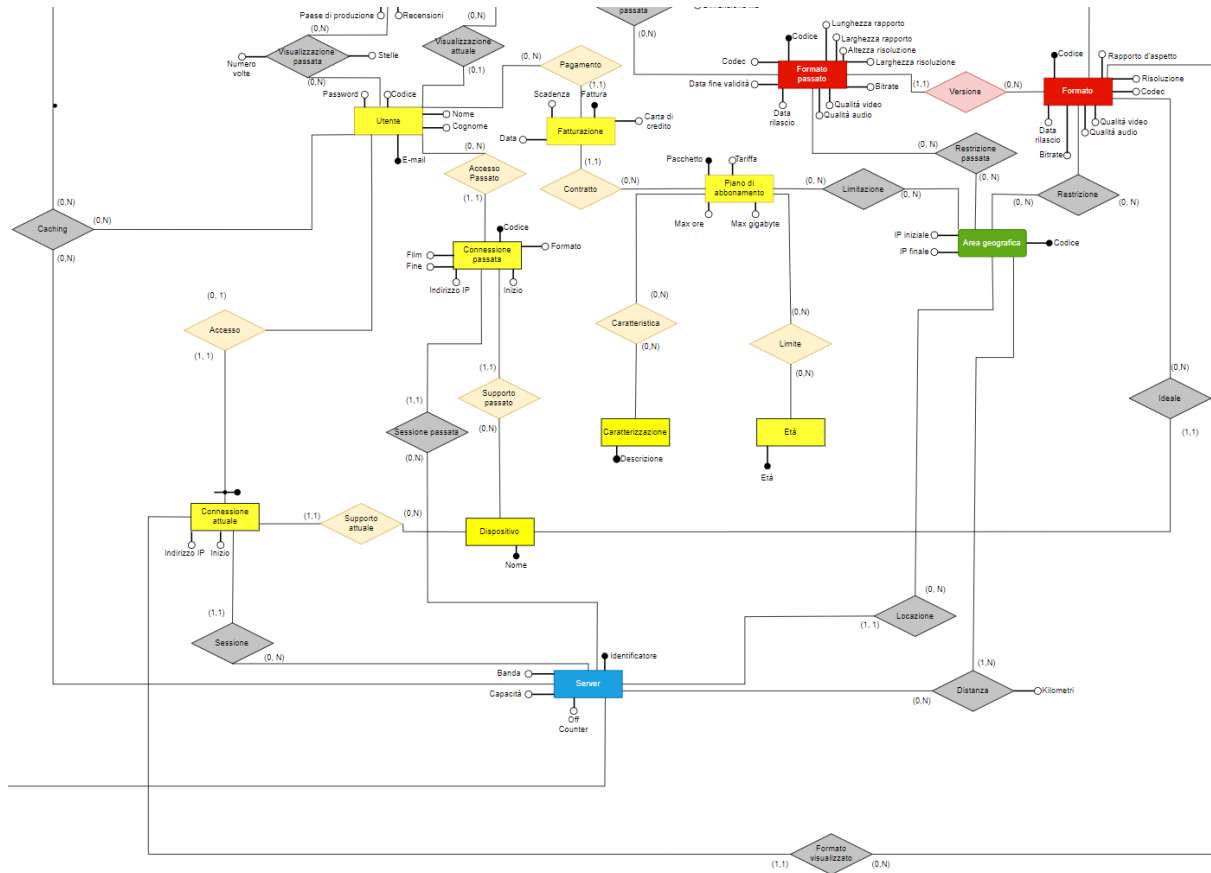
- Ci sono 25 connessioni (attuali + passate) per utente (analogo per gli accessi, attuali e passati).
- È connesso attualmente al servizio un utente su 2 (analogo per gli accessi attuali).
- Ci sono 3 fatturazioni per utente.
- Un attore ha recitato in 6 film.
- Un film è disponibile in 8 formati.
- Un formato non è disponibile in 2 aree geografiche.
- Ci sarebbero 1,1 registi per il film. In base all'operazione in cui utilizziamo la *relationship* Direzione, faremo le dovute considerazioni.

Esattamente:

- Ci sono 4 occorrenze della *relationship* "Caching" per ogni occorrenza dell'entità Utente. Infatti, per utente, ci sono 2 film in caching presenti in contemporanea su 2 server diversi.

## Ristrutturazione dello schema ER

## Area Utente

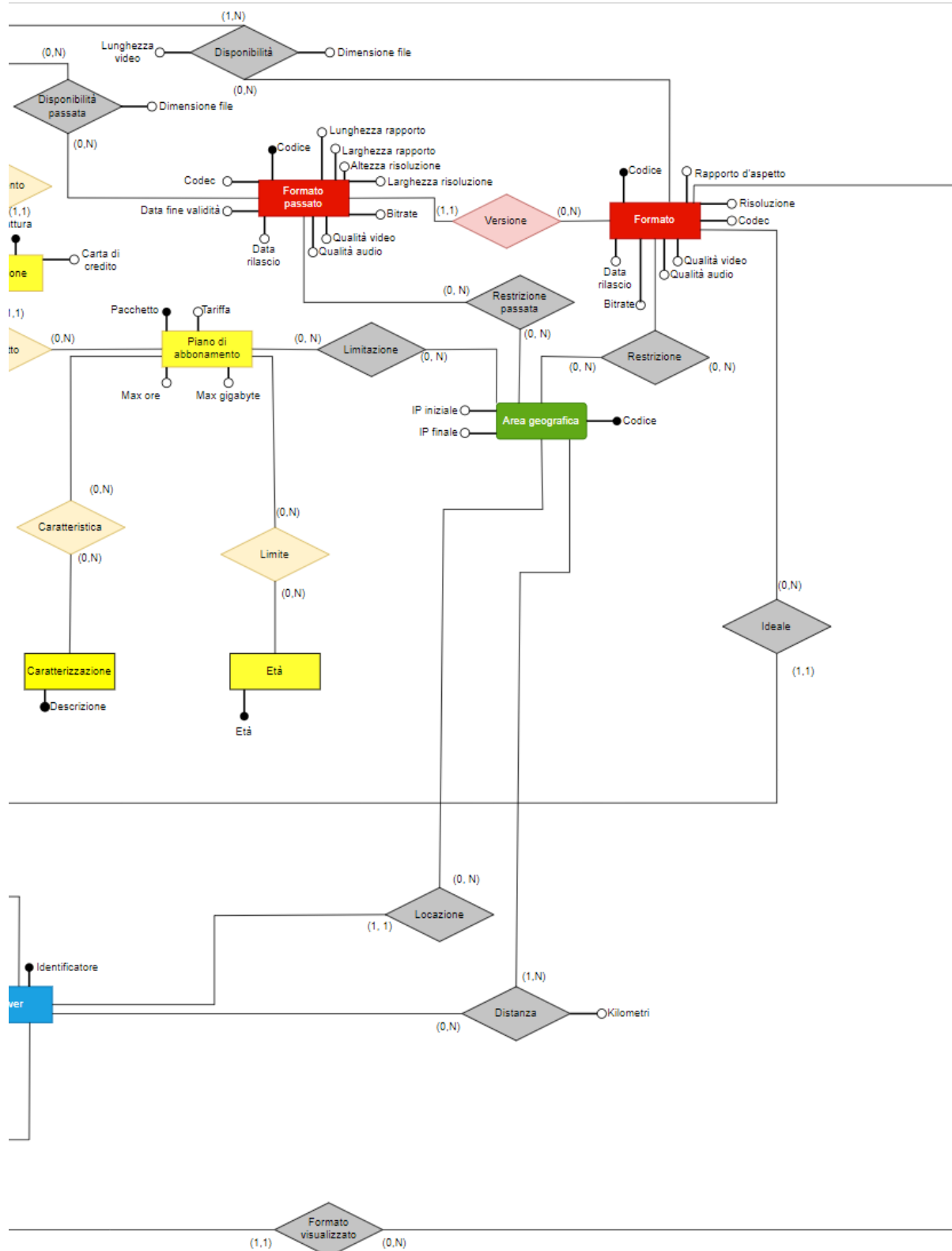


Nello schema sono presenti due generalizzazioni, quelle che coinvolgono l'entità connessione e dell'entità formato. Poiché gli accessi al padre e alle figlie non sono contestuali e inoltre la generalizzazione è totale, decidiamo di accorpare il padre alle figlie. La relationship *Accesso* verrà anch'essa sostituita da due relazioni, *Accesso passato* e *Accesso attuale*. Infine, anche la relazione *Supporto* viene sostituita da due relazioni, *Supporto attuale* e *Supporto passato*.

Inoltre, decidiamo di sostituire l'identificatore esterno con un codice univoco che funga da identificatore primario per quanto riguarda le connessioni passate, invece per le connessioni attuali, che possono essere massimo una per account contemporaneamente, scegliamo la chiave esterna *Utente*.

Reifichiamo i due attributi multivalore “Caratterizzazioni (0,N)” e “Limiti d’età (0,N)” in due entità, Caratterizzazione ed Età, che andiamo a collegare con due nuove relazioni, Caratteristica e Limite, alla relazione Piano di Abbonamento, alla quale appartenevano i due attributi multivalore in precedenza.

## Area formato

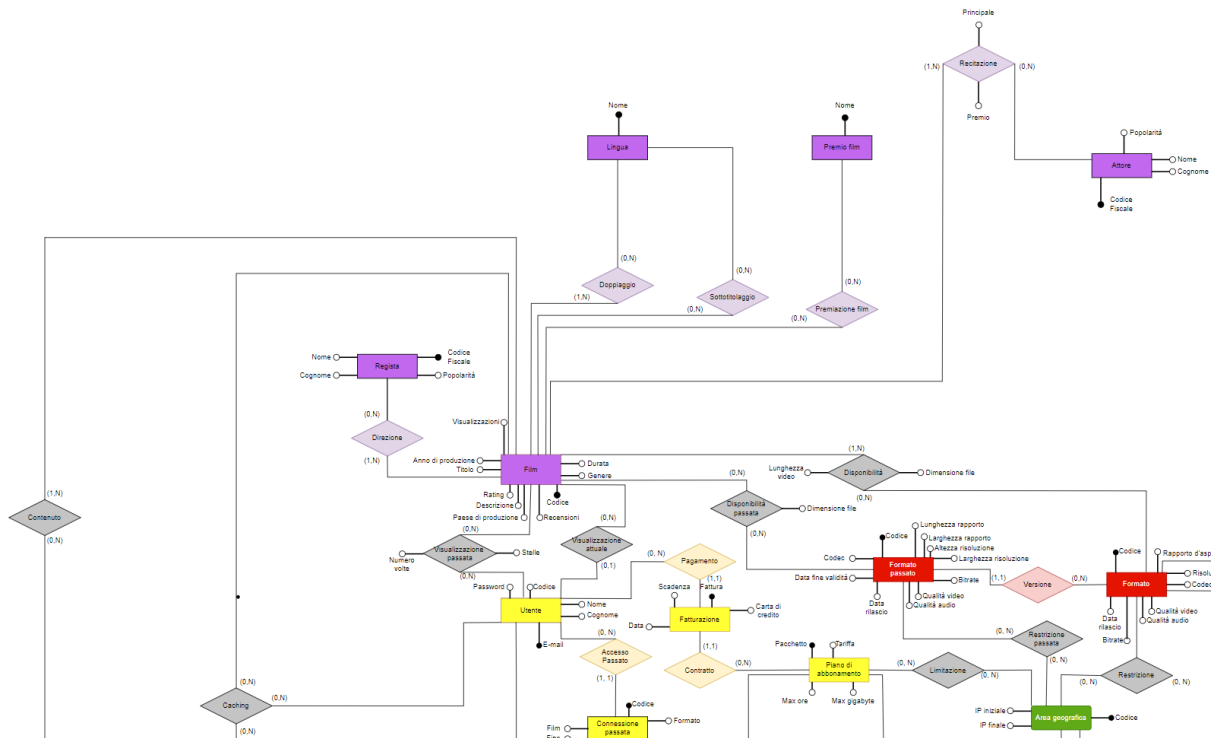


La seconda generalizzazione coinvolge l'entità genitore formato e le due entità figlie formato corrente e formato passato. Decidiamo di accorpare il genitore alle figlie e di aggiungere una relazione *Versione* per rappresentare la storicizzazione di un concetto senza perdere i dati delle versioni passate e anche il fatto che un formato può avere più di una versione passata con specifiche diverse.

Duplichiamo la relationship *Restrizione*, che inizialmente coinvolgeva l'entità *Formato*.

Inoltre l'attributo multivalore di *Formato* *Dispositivo* viene sostituito con una relazione e un'entità, *Ideale* e *Dispositivo*. *Dispositivo* viene anche collegato a connessione passata per memorizzare da che tipo di dispositivo sono avvenute le connessioni. Tutti i dispositivi sono “compatibili” con tutti i formati. La relazione *Ideale* serve solo per stabilire quale sia il formato consigliato per quel particolare dispositivo.

## Area contenuti



L'entità *Film* aveva tre attributi multivalore: *Lingua* (1,N), *Lingua sottotitoli* (0,N), *Premio* (0,N). Li sostituiamo con due nuove entità, *Lingua* (a cui si fa riferimento sia per il doppiaggio che per il sottotitolaggio), e *Premio film*, connesse a *Film* tramite altrettante *relationship*. Inoltre *Film* era identificato fino a questo momento tramite l'anno di produzione, il titolo e il regista (identificatore esterno). Decidiamo di introdurre un codice identificativo per ogni film che funga da identificatore primario.

**\*TABELLA DELLE ENTITÀ (dopo la ristrutturazione):**

Entità	Descrizione	Attributi	Identificatore
Film	Un film visualizzabile tramite il servizio.	Codice, Durata, Genere, <u>Rating</u> , Recensioni, Descrizione, Paese di produzione, Titolo, Anno di produzione, <u>Visualizzazioni</u>	Codice
Lingua	Una delle lingue in cui è disponibile il film	Nome	Nome
Premio film	Un premio vinto da un film.	Nome	Nome
Caratterizzazione	Una delle possibili funzionalità esclusive di un piano di abbonamento	Descrizione	Descrizione
Età	Alcuni piano di abbonamento possono avere limitazioni legate all'età	Età	Età
Formato	I formati con cui sono attualmente disponibili i film	Codice, Larghezza rapporto, Altezza rapporto, Lunghezza risoluzione, Altezza risoluzione, Bitrate, Qualità video, Qualità audio, Codec, Data rilascio	Codice
Formato passato	Le vecchie versioni di un formato	Codice, Larghezza rapporto, Altezza rapporto, Lunghezza risoluzione,	Codice

		Altezza risoluzione, Bitrate, Qualità video, Qualità audio, Data rilascio, Data fine validità, Codec	
--	--	--	--

**\*TABELLA DELLE RELAZIONI (dopo la ristrutturazione):**

Relazione	Descrizione	Entità coinvolte	Attributi
Doppiaggio	L'audio dei film è disponibili in varie lingue	Film (1,N), Lingua (0,N)	/
Sottotitolaggio	I sottotitoli possono essere disponibili in varie lingue	Film (0, N), Lingua sottotitoli (0,N)	/
Premiazioni film	Un film può avere vinti diversi premi	Film (0,N), Premio film (0,N)	/
Versione	Collega un formato attualmente disponibile con una delle sue versioni passate.	Formato (0,N), Formato passato (1,N)	/
Restrizione passata	Indica se una versione vecchia di un formato non fosse disponibile in alcune aree geografiche	Formato passato (0,N), Area geografica (0,N)	/
Caratteristica	Collega un pacchetto con una delle funzionalità specifiche del pacchetto	Caratterizzazione (0,N), (0,N)	/
Limite	Alcuni piani di abbonamento possono avere restrizioni legate all'età	Età (0,N), piano di abbonamento (0,N)	/
Accesso	Lega un utente alla connessione attualmente in corso	Utente (0,1), Connessione attuale(1,1)	/
Accesso passato	Lega un utente alle sue connessioni passate	Utente (0,N), Connessione passata (1,1)	/



Supporto attuale	Lega una connessione attuale al dispositivo utilizzato	Connessione attuale (1,1), Dispositivo (0,N)	/
Supporto passato	Lega una connessione attuale al dispositivo utilizzato	Connessione passata (1,1), Dispositivo (0,N)	/

**\*TAVOLA DEI VOLUMI (dopo la ristrutturazione)**

Concetto	Tipo	Volume
Lingua	E	20
Premio film	E	15
Caratterizzazione	E	10
Età	E	5
Formato Corrente	E	15
Formato Passato	E	45
Doppiaggio	R	3000 (15*Film)
Sottotitolaggio	R	4000 (20*Film)
Premiazione film	R	200
Versione	R	45
Restrizione	R	75
Restrizione passata	R	200
Caratteristica	R	30 (6 * Piano di abbonamento)
Limite	R	25 (5*Piano di abbonamento)
Accesso	R	500
Accesso passato	R	24.500

Disponibilità Passata	R	400
Supporto attuale	R	500
Supporto passato	R	24.500
Ideale	R	20
Formato visualizzato	R	500

\*Per evitare ripetizioni, abbiamo riportato in questa tavola solo le nuove entità/entità che hanno subito modifiche. Analogo per la tabella relazione e per la tavola dei volumi.

### **Osservazioni**

In media:

- Un film è doppiato in 10 lingue.
- Un film è sottotitolato in 15 lingue.

## Operazioni significative

- 1) Inserimento di un nuovo film
- 2) Inserimento di un nuovo utente (con relativo abbonamento).
- 3) Calcolo del rating + calcola media voti utenti + stampa informazioni film.
- 4) Inserimento di una nuova visualizzazione attuale
- 5) Calcolo di tutti i film di un certo genere con valutazione sopra una certa soglia data in input (sfrutta la ridondanza)
- 6) Rinnovo/Disdetta/Cambio abbonamento
- 7) Elenco degli utenti connessi e tutte le loro informazioni (area geografica, dispositivo, indirizzo IP)
- 8) Tutti i film con più di tot visualizzazioni (sfrutta la ridondanza Numero Visualizzazioni su film).

### TAVOLA DELLE OPERAZIONI

Operazione	Tipo	Frequenza
Operazione 1	B	1 volta ogni 3 giorni
Operazione 2	I	10 volte al giorno
Operazione 3	I	30 volte al giorno
Operazione 4	I	500 volte al giorno
Operazione 5	I	10 volte al giorno
Operazione 6	I	10 volte al giorno
Operazione 7	B	18 volte al giorno
Operazione 8	I	5 volte al giorno

## Analisi delle prestazioni delle operazioni

### OPERAZIONE 1

**Descrizione:** l'operazione consiste nell'inserire un nuovo film sul sito di streaming.

**Input:** Codice film, titolo, durata, anno di produzione, paese di produzione, descrizione, genere, media voto utenti (= 0), i formati in cui è disponibile il film (mediamente 8)

**Output:** nessuno

#### Tavola degli accessi

Concetto	Costrutto	Accessi	Tipo	Motivazione
Film	Entità	1	S	Avere un costante aggiornamento del catalogo di film disponibili è essenziale per aver un servizio soddisfacente.
Disponibilità attuale	Relazione	8	S	Per definire in quali formati sia disponibile il film.
Formato	Entità	8	L	Per ricavare risoluzione e rapporto d'aspetto

Si noti che questa operazione non riguarda in alcun modo le ridondanze.

La dimensione di un film (in *gigabyte*) in un certo formato è data da:

$\text{RisoluzioneVideo} * \text{RapportoDAspetto} * \text{Durata} / 10^7$

La qualità audio e gli altri parametri del formato sono trascurabili.

Durata è in minuti (scritto anche nella sezione vincoli).

## **OPERAZIONE 2**

**Descrizione:** l'operazione consente la registrazione di un nuovo utente al servizio, che si iscrive con uno specifico piano di abbonamento.

**Input:** Nome, cognome, e-mail, password, carta di credito, pacchetto

**Output:** nessuno

### **Tavola degli accessi**

Concetto	Costrutto	Accessi	Tipo	Motivazione
Utente	Entità	1	S	Crea un nuovo utente
Pagamento	Relazione	1	S	Crea l'atto del pagamento.
Fatturazione	Entità	1	S	Crea una nuova fatturazione
Contratto	Relazione	1	S	Stabilisce a quale piano di abbonamento la fatturazione si riferisca

Si noti che questa operazione non riguarda le ridondanze.

### OPERAZIONE 3

**Descrizione:** stampa tutti i dati riguardanti un film (tranne il suo codice, che è dato in input)

**Input:** Codice film

**Output:** Rating, valutazione media degli utenti, regista, attori, anno di produzione, durata, titolo, descrizione, paese di produzione, le lingue in cui è disponibile, le lingue dei sottotitoli in cui il film è disponibile, eventuali premi vinti dal film, i formati in cui è disponibile, gli attori che hanno recitato

#### **Tavola degli accessi (con ridondanza)**

Concetto	Costrutto	Accessi	Tipo	Motivazione
Film	E	1	L	Leggere le informazioni sul film
Direzione	R	1/2 (2 nel caso peggiore)*	L	Recuperare i registi del film
Regista	E	1/2 (2 nel caso peggiore)*	L	Ci serve il regista
Doppiaggio	R	10	L	Per ottenere le lingue in cui è doppiato un film.
Sottotitolaggio	R	15	L	Per ottenere le lingue dei sottotitoli
Premiazione film	R	1	L	Per sapere se un film è stato premiato
Disponibilità	R	8	L	Per sapere in quali formati è stato doppiato un film
Recitazione	R	6	L	Per sapere se un attore ha recitato nel film.
Attore	E	6	L	Gli attori che hanno recitato nel film

\*Di media, ci sarebbero 1,1 registi per film. Quindi, nei casi migliore e medio ci sarà un solo regista per film, mentre nel caso peggiore il film avrà 2 registi (quindi faremo sia 2 accessi a direzione che a Regista).

Visto che un accesso in scrittura corrisponde a due accessi in lettura, gli accessi in lettura totali sono 49 per richiesta: essendoci 30 richieste di questa operazione al giorno, ci sono 1470 accessi in lettura al giorno.

Il costo dell'attributo ridondante "Rating", in termini di memoria, è 4 byte (un *float*) a film. Visto che abbiamo 200 film, abbiamo 800 byte utilizzati, quindi meno di un 1 *kilobyte*.

Si noti che l'analisi delle ridondanze contava un solo accesso all'entità Regista, e non 2.

La formula per il calcolo del rating di un film è:

$AVG(VotoMedioUtenti + PopolaritaAttori + PopolaritaRegisti) + 0.1 \times PremiOscar$

In caso di utilizzo della ridondanza, c'è bisogno di tenerla aggiornata (visto che è possibile che la media del voto degli utenti cambi nel tempo). Un *event* si adatta perfettamente alla richiesta, essendo il voto medio un parametro abbastanza stabile nel tempo (e comunque, anche se fosse sempre perfettamente aggiornato non sarebbe un problema particolarmente significativo) e, quindi, soddisfacente anche se aggiornato con una frequenza ridotta:

Concetto	Costrutto	Accessi	Tipo	Motivazione
Film	E	200	L	Per accedere alle sue valutazioni degli utenti
Film	E	200	S	Per riscrivere l'attributo rating.
Recitazione	E	1200 (6*Film)	L	Per sapere se attori che hanno recitato in quel film hanno vinto l'Oscar e se sono attori principali
Attore	E	500	L	Per sapere la popolarità degli attori principali di quel film
Direzione	R	220	L	Per risalire al regista del film
Regista	E	50*	L	Per risalire alla popolarità del regista

Questa operazione viene eseguita ogni 2 settimane. Quindi, visto che gli accessi in scrittura contano doppio rispetto a quelli in lettura, abbiamo 12370 accessi in ~~più ogni 2 settimane (di media, circa 884 al giorno).

Perciò, avremmo 1470 accessi al giorno + 884 al giorno (in media), per un totale di 2354 accessi al giorno in media.



Senza ridondanza, ogni volta che l'operazione viene eseguita (quindi 30 volte al giorno) dovremmo fare gli accessi dell'*event* per il singolo film, quindi avremmo, di media,  $12370/200 = 62$  accessi per film, perciò 1860 accessi in più al giorno. In 14 giorni, questi diventerebbero 26040. Vista la poca memoria occupata, conviene tenere la ridondanza.

\*Supponiamo che ogni regista abbia diretto almeno un film, quindi accederemo a ogni occorrenza dell'entità Regista.

## **OPERAZIONE 4**

**Descrizione:** inserire una nuova visualizzazione (attuale) di un film da parte di un utente.

**Input:** l'email dell'utente e un film.

**Output:** stampa l'esito della richiesta

Quando un utente vuole vedere un film:

- 1) Prima di tutto, si controlla se il film dato in input è disponibile in caching per quel determinato utente. Se lo è, si controlla lo stato dei server in caching, se sono in uno stato valido, si crea l'occorrenza di relationship visualizzazione attuale e la connessione attuale (visto che i server in caching si sanno a priori").
- 2) Se non lo è: si crea comunque l'occorrenza di relationship di visualizzazione attuale, tramite la relazione contenuto si ricavano i server in cui il film è disponibile.
- 3) Si crea una nuova connessione con inizio, dispositivo e indirizzo IP dati in input.
- 4) Con l'indirizzo IP si risale all'area geografica da cui l'utente si connette; tramite la relazione distanza si trovano i server più "convenienti". Trovato il server migliore, si crea la connessione attuale e la relationship sessione tra server e connessione attuale.

## Tavola degli accessi

Concetto	Costrutto	Accessi	Tipo	Motivazione
Utente	E	1	L	Per ricavare le informazioni sul caching
Caching	R	4	L	Per verificare se il film è disponibile in caching per quell'utente e su quali server
Visualizzazione attuale	R	1	S	Per creare la visualizzazione attuale
Film	E	1	S	Per mantenere aggiornata la voce Numero Visualizzazioni
Accesso	R	1	S	Per far connettere l'utente al servizio
Connessione attuale	E	1	S	Per creare la connessione tra l'utente e il servizio.
Sessione	R	1	S	Per stabilire a quale server la connessione faccia riferimento.
Film	E	1	L	Per ricavare i server in cui il film è disponibile
Contenuto	R	20	L	Per trovare i server che contengono il film
Area geografica	E	1	L	Per trovare i server più vicini all'area geografica da cui si connette l'utente
Distanza	R	20	L	Per trovare la distanza minima tra i 20 server che hanno il film disponibile e l'area geografica da cui si connette l'utente. Si fa l'accesso diretto alle occorrenze di Distanza perché da

				Contenuto si ha la chiave di Server, mentre dall'indirizzo IP si risale all'area geografica dell'Utente
--	--	--	--	---

**N.B.** La parte **verde** viene eseguita solo se il film non è presente in *caching* per quell'utente.

**\*\***Spiegato nell'operazione del cambio di server (pagina 55). Visto che, a priori, non sappiamo stimare il numero di connessioni attuali per server, supponiamo di dover scansionare tutte le connessioni attuali, ma è solo un'approssimazione.

Se esiste un server *edge* (quindi che fornisce il *caching*) per quell'utente e per quello specifico film, non sono necessari gli accessi nella parte rossa della tabella.

L'inserimento di una nuova visualizzazione nell'attributo Visualizzazioni dell'entità Film sarà effettuata nell'operazione 8 (tramite un *event*). Quindi, questa operazione non riguarda le ridondanze.

L'accesso in scrittura a Film sarà implementato con un Trigger e servirà per aumentare l'efficienza dell'operazione 8.

Lo implementeremo con un *Trigger* - e non con un *Event* - perché:

- Usando un Trigger, dovremmo accedere solo a 500 occorrenze di entità di Film al giorno (ogni volta che c'è una visualizzazione, appunto). In una settimana, questi sarebbero 3500.
- Se utilizzassimo un *Event*, invece, supponendo un aggiornamento settimanale delle visualizzazioni totali di ogni film, dovremmo accedere a tutte le occorrenze di visualizzazione passata + a tutte quelle di visualizzazione attuale + a tutte quelle di film (a Film in scrittura), quindi fare  $30.000 + 500 + 2 \cdot 200 = 30.700$  accessi.  
Questo nel caso in cui, ogni volta, le visualizzazioni per ogni film si ricalcolino da 0.
- Si noti che effettuare un *incremental refresh* tramite una *materialized view* risulti non ottimale, visto che il *trigger di push* per aggiornare la *log table* della *materialized view* richiederebbe esattamente lo stesso numero di operazioni elementari del *trigger* applicato all'entità Film, più quelle necessarie aggiornare la base di dati con i dati della *log table*, rendendo l'*incremental refresh* automaticamente meno conveniente del *trigger* diretto.

## **OPERAZIONE 5**

**Descrizione:** per un certo genere restituisce i film che hanno un certo rating minimo.

**Input:** Valutazione minima (intero), Genere del Film.

**Output:** elenco dei film validi.

### **Tavola degli accessi (con ridondanza)**

Concetto	Costrutto	Accessi	Tipo	Motivazione
Film	E	200	L	Per trovare i film di quel genere e sapere la media voto utenti

Con la ridondanza, gli accessi totali sono 200 in lettura per operazione. Considerata la frequenza di 10 volte al giorno, gli accessi totali giornalieri sono 2000.

Come già detto durante la descrizione dell'operazione 3, la memoria necessaria per la ridondanza è di 800 byte.

### **Tavola degli accessi (senza ridondanza)**

In caso di ridondanza, bisogna eseguire gli accessi che fa l'*event* descritto all'operazione 3 (ovviamente, per un singolo film), che sono 62 per film. In totale, sarebbero, vista la frequenza di 10 volte al giorno,  $2000 + 620$  accessi al giorno. Visto che la ridondanza serve anche per velocizzare quest'operazione, si dimostra ancora una volta conveniente.

## OPERAZIONE 6

**Descrizione:** il cliente comunica se intende rinnovare, cambiare o disdire l'abbonamento.

**Input:** Codice utente, carta di credito, stringa 'Rinnovo' o 'Disdetta' o 'Cambio', piano di abbonamento (utile solo per il cambio).

**Output:** Stampa l'esito dell'operazione.

### **Tavola degli accessi (Rinnovo)**

Concetto	Costrutto	Accessi	Tipo	Motivazione
Utente	E	1	L	Per risalire al suo ultimo pagamento
Pagamento	R	3	L	Per risalire all'ultima fatturazione
Fatturazione	E	3	L	Per risalire alla data di scadenza dell'abbonamento.
Contratto	R	1	L	Per trovare il piano di abbonamento
Piano di abbonamento	E	1	L	Per trovare il pacchetto e la sua tariffa, che andrà nella fattura.
Contratto	R	1	S	Per inserire il nuovo contratto dell'utente
Pagamento	R	1	S	Registrare l'atto di pagamento dell'utente.
Fatturazione	E	1	S	Aggiungere una nuova fatturazione con relativi dati.

### **Tavola degli accessi (Cambio)**

Concetto	Costrutto	Accessi	Tipo	Motivazione
Pagamento	R	1	S	Per aggiungere il nuovo atto del

				pagamento
Fatturazione	E	1	S	Per aggiungere la nuova fatturazione
Contratto	R	1	S	Per aggiungere il nuovo contratto

#### Tavola degli accessi (Disdetta)

Concetto	Costrutto	Accessi	Tipo	Motivazione
Fatturazione	E	3	L	Per trovare la fatturazione più recente
Fatturazione	E	1	S	Per modificare la data di scadenza alla data odierna

Effetti dell'operazione:

- Se sceglie 'Rinnovo', si va a cercare l'ultima fatturazione per trovare che piano di abbonamento aveva per rinnovarlo a partire dalla sua scadenza.
- Se sceglie 'Cambio', il nuovo abbonamento sovrascriverà quello vecchio a partire dalla data corrente.
- Se sceglie 'Disdetta', la data di scadenza verrà impostata su quella odierna. Si osservi che la disdetta di un abbonamento non comporta la cancellazione di alcun dato riguardante l'utente. Visto che la durata di un abbonamento è una sua specifica caratteristica e visto che sappiamo la data della fatturazione, resterà comunque possibile risalire, all'occorrenza, alla data in cui tale abbonamento sarebbe scaduto.

Si noti che l'operazione non riguarda le ridondanze.

## OPERAZIONE 7

**Descrizione:** mostra tutte le connessioni attuali, comprendenti tutte le varie informazioni, con aggiunta di area geografica in cui è collocato il server e di area geografica dalla quale si connette l'utente.

**Input:** nessuno

**Output:** elenco degli utenti con area geografica da cui si connettono, il server a cui sono connessi e la posizione del server a cui sono connessi. Si stampano anche le varie informazioni della connessione attuale (indirizzo IP, inizio della connessione, dispositivo da cui l'utente è connesso).

### Tavola degli accessi

Concetto	Costrutto	Accessi	Tipo	Motivazione
Connessione attuale	E	500	L	Avere le informazioni sulle varie connessioni attuali
Accesso	R	500	L	Sapere quali utenti sono connessi attualmente
Sessione	R	500	L	Per risalire al server a cui è connesso l'utente
Server	E	40	L	Per poter accedere alla locazione del server*
Locazione	R	40	L	Per risalire all'area geografica in cui è collocato il server

\*Si suppone che ogni server abbia almeno un utente connesso, quindi devo guardare tutte le occorrenze di Server.

Si noti come l'utilizzo di una ridondanza per un'operazione come questa sia sempre scorretto, visto che si richiede una lista aggiornata sul momento ogni volta.

Questa operazione è per gli amministratori, e non per gli utenti casuali.



## OPERAZIONE 8

**Descrizione:** restituisce i film che hanno più di un certo numero di visualizzazioni.

**Input:** numero di visualizzazioni

**Output:** stampa tutti i film che hanno numero di visualizzazioni maggiore del numero dati in input

### Tavola degli accessi (con ridondanza)

Concetto	Costrutto	Accessi	Tipo	Motivazione
Film	E	200	L	Per trovare tutti i film in cui Numero Visualizzazioni è maggiore del numero dato in input e risalire alle informazioni sul film

### Tavola degli accessi (senza ridondanza)

In questo caso, oltre agli accessi a Film (che restano invariati), sarà necessario ricavare tutti gli attori che hanno recitato in ogni film.

Concetto	Costrutto	Accessi	Tipo	Motivazione
Film	E	200	L	Per risalire alle visualizzazioni
Visualizzazioni e passata	R	30000	L	Per vedere quante visualizzazioni passate ha il film
Visualizzazioni e attuale	R	500	L	Per vedere quante visualizzazioni attuali ha un film (per poterle sommare alle visualizzazioni passate)

La frequenza dell'operazione 8 è di 5 volte al giorno.

Quindi, con ridondanza, avendo un costo di 200 accessi in lettura per operazione, il suo costo giornaliero sarà di 1000 accessi in lettura. La ridondanza occupa 2 *byte* per film (Un film, di media, ha 225 visualizzazioni: 2 *byte* sono pienamente sufficienti), quindi 400 *byte* in totale, visto che abbiamo 200 film.

Senza ridondanza, oltre ai 200 accessi in lettura a film per operazione, dobbiamo risalire a tutte le occorrenze di visualizzazione passata (che sono in totale 30000, circa 30 per utente) e alle occorrenze di visualizzazione attuale, che sono 500. Il

costo totale giornaliero sarà quindi di  $30700 \cdot 5 = 153.300$  visite in letture, data la frequenza di 5 volte al giorno.

Vista la pochissima memoria utilizzata dalla ridondanza, conviene tenerla.

Per mantenere aggiornata la ridondanza non sono necessari ulteriori accessi né in lettura né in scrittura, in quanto questa è mantenuta aggiornata dall'operazione 4, che ogni volta che viene aggiunta una nuova visualizzazione aggiorna anche il contatore di visualizzazioni per il film che viene visualizzato, la cui convenienza è già stata discussa.

## Business rules

### Area contenuti

#### Vincoli di integrità

- 1) “Anno di produzione” deve essere minore o uguale dell’anno corrente.
- 2) L’attributo “Stelle” della *relationship* “Valutazione” deve assumere un valore intero tra 1 e 5 compresi.
- 3) L’attributo “Numero volte” della *relationship* “Visualizzazione Passata” deve essere un intero strettamente maggiore di zero.
- 4) Un “Codice Fiscale” è una stringa di 5 caratteri per i registi e massimo 20 per gli attori.
- 5) Un cambio di server durante la visione di un film non comporta un’aggiunta ulteriore a una visualizzazione passata.
- 6) Una volta terminata la visione di un film da parte di un determinato utente, si aggiunge un’unità all’attributo “Numero volte” della *relationship* Visualizzazione passata.
- 7) L’attributo Durata della relazione Film si misura in minuti.

#### Regole di derivazione

Il processo per ricavare i valori contenuti negli attributi ridondanti è stato spiegato per ogni ridondanza nelle operazioni in cui essa è stata impiegata.

### Area utente

#### Vincoli di integrità

- 1) Un’occorrenza di “Connessione Attuale” per un certo utente si crea solo dopo che l’utente è in una occorrenza di Visualizzazione Attuale. Gestito dall’operazione 4.

- 2) L'attributo "Password" di "Utente" deve essere una stringa di almeno 8 caratteri, di al massimo 30 caratteri e contenere almeno un carattere speciale ("\$", "%", "&", "@", "£", "-", ".", ":", "\_").
- 3) "Connessione Attuale" può cambiare il server a cui è legata tramite "Sessione".
- 4) In "Connessione Passata", l'attributo "Inizio" deve essere strettamente minore di "Fine".
- 5) "Età", attributo di "Età", è un intero che indica l'età che dà il limite superiore all'intervallo, il limite inferiore è standard e vale 0 anni.
- 6) In "Fatturazione", "Data" indica una data anteriore rispetto a "Scadenza".
- 7) La durata dell'abbonamento è uguale per tutti i pacchetti, ed è di un mese.
- 8) "Pacchetto" in "Piano Di Abbonamento" deve essere una stringa tra le seguenti: *Basic, Premium, Pro, Deluxe, Ultimate*.
- 9) "Carta di credito", attributo dell'entità "Fatturazione", deve essere un intero fatto in questo modo: "ABCDEFGHJKLMNOP", dove una maiuscola indica un numero naturale.
- 10) "Tariffa", attributo di "Piano Di Abbonamento" è un intero e deve essere misurato in euro.
- 11) L'attributo "Formato" dell'entità "Connessione Passata" deve essere lo stesso formato presente che era presente nella relazione "Formato Visualizzato" quando la connessione - ora passata - era "attuale".
- 12) L'attributo "Film" dell'entità "Connessione Passata" deve essere lo stesso film presente che era presente nella relazione "Visualizzazione Attuale" quando la connessione, ora passata, era "attuale".
- 13) Un utente deve essere riconnesso per guardare più film di fila, visto che le *relationship* "Connessione Attuale" e "Visualizzazione Attuale" sono collegate da vincoli d'integrità.
- 14) L'indirizzo IP permette di risalire all'area geografica da cui l'utente si è connesso, visto che, per ogni area geografica, abbiamo degli intervalli di indirizzi IP ammissibili.  
Si presume che l'indirizzo "IPIniziale" sia ammissibile come valore per le connessioni a quell'area geografica; "IPFinale" fa connettere gli utenti all'area geografica successiva.
- 15) Prima che un'occorrenza dell'entità "ConnessioneAttuale" venga cancellata, i suoi dati saranno inseriti in una nuova occorrenza dell'entità "ConnessionePassata".
- 16) La relazione "Formato Visualizzato" ha quel formato con cui l'utente sta visualizzando attualmente quel preciso film. Quindi, deve esistere una occorrenza della *relationship* Disponibilità contenente il codice del film e il formato con cui l'utente lo sta guardando.
- 17) L'attributo "Popolarità", sia nell'entità Regista che nell'entità attore può assumere valori compresi tra 1 e 10.

## **Area streaming**

### **Vincoli di integrità**

- 1) "Identificatore", attributo di Server, è una stringa di 4 caratteri.
- 2) La Banda di un server è compresa tra 0,2 e 0,9 Gbit/s
- 3) Capacità di trasmissione è un intero e l'unità di misura sono i *bit/s*.
- 4) Il meccanismo di *Caching* sarà spiegato nell'operazione richiesta al paragrafo 2.5.3.

## **Area formato**

### **Vincoli di integrità:**

- 1) In Formato passato, "Data Rilascio" deve essere strettamente minore di "Data Fine Validità".
- 2) In Formato Corrente e Formato Passato, "Codice" è una stringa in questa forma "NomeFormato\_Versione", dove "NomeFormato" è una stringa indicante il nome del formato e "Versione" è un naturale.
- 3) "Codec" è una stringa e può contenere sia numeri che lettere.
- 4) "Risoluzione" è una stringa nella forma "ABCDxEFGH" dove ABCD e EFGH sono numeri e l'unità di misura sono i pixel.
- 5) "Qualità audio" e "Qualità video" sono interi da 1 a 100, che costituiscono una scala oggettiva per misurare la qualità.
- 6) "Bitrate" è un intero e l'unità di misura sono i *bit/s*.
- 7) "Dimensione file", attributo della *relationship* "Disponibilità", è un *float* e si rappresenta in *gigabyte*.

## **Area geografica**

### **Vincoli di integrità referenziale:**

- 1) "Codice", attributo di "Area Geografica", è una stringa di massimo 4 caratteri che indica un acronimo dello stato a cui appartiene quell'area geografica.

## SCHEMA LOGICO E NORMALIZZAZIONE

**N.B.** Qui il termine “relazione” si riferisce al significato del modello relazionale.

**N.B.** I vincoli d'integrità referenziale qui riportati sono riferiti alla chiave primaria.

**N.B.** Qui si aggiungono quei vincoli derivanti dalla traduzione dello schema ER ristrutturato nel modello relazionale.

**N.B.** Gli **attributi ridondanti** sono indicati con un “\*” alla fine del loro nome, a destra.

**N.B.** Ogni vincolo d'integrità referenziale sarà implementato con un opportuno *trigger*.

### Area Contenuti

**Film**(Codice, Durata, Genere, AnnoDiProduzione, Titolo, Rating\*, Descrizione, PaeseDiProduzione, Visualizzazioni\*).

**Direzione**(Film, Regista)

**Regista**(CodiceFiscale, Nome, Cognome, Popolarità).

**Lingua**(Nome).

**Doppiaggio**(Film, Lingua)

**PremioFilm**(Nome)

**PremiazioneFilm**(Premio, Film)

**Sottotitolaggio**(Film, Lingua)

**Attore**(CodiceFiscale, Nome, Cognome)

**PremioAttore**(Nome)

**Recitazione**(Film, Attore)

**PremiazioneAttore**(Attore, Film, Premio)

Chiedere se mantenere Regista a Pistolesi in Film.

### **Vincoli d'integrità referenziale**

- Tra l'attributo Regista della relazione Film e la chiave (Codice fiscale) della relazione Regista.
- Tra l'attributo Film della relazione Doppiaggio e la chiave (Codice) della relazione Film.
- Tra l'attributo Lingua della relazione Doppiaggio e la chiave (Nome) della relazione Lingua.
- Tra l'attributo Premio della relazione PremiazioneFilm e la chiave (Nome) della relazione PremioFilm.
- Tra l'attributo Film della relazione PremiazioneFilm e la chiave (Codice) della relazione Film.

- Tra l'attributo Film della relazione Sottotitolaggio e la chiave (Codice) della relazione Film.
- Tra l'attributo Lingua della relazione Sottotitolaggio e la chiave (Nome) della relazione Lingua.
- Tra l'attributo Film della relazione Recitazione e la chiave (Codice) della relazione Film.
- Tra l'attributo Attore della relazione Recitazione la chiave (CodiceFiscale) della relazione Attore.
- Tra l'attributo Attore della relazione PremiazioneAttore e la chiave (CodiceFiscale) della relazione Attore.
- Tra l'attributo Film della relazione PremiazioneAttore e la chiave (Codice) della relazione Film.
- Tra l'attributo Premio della relazione PremiazioneAttore e la chiave (Nome) della relazione PremiazioneAttore.
- Tra l'attributo Regista della relazione Direzione e la chiave (CodiceFiscale) della relazione Regista.
- Tra l'attributo Film della relazione Direzione e la chiave (Codice) della relazione Film.

## Dipendenze funzionali

### Relazione Film

- Codice → Durata, Genere, Anno di produzione, Titolo, Rating\*, Descrizione, PaeseDiProduzione, AnnoDiProduzione, Visualizzazioni\*.
- Titolo, Anno di produzione, Regista → Codice, Durata, Rating\*, Descrizione, PaeseDiProduzione, AnnoDiProduzione, Visualizzazioni\*

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

### Relazione Regista

- CodiceFiscale → Nome, Cognome, Popolarità.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

### Relazione Direzione

- solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

#### Relazione Lingua

- solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

#### Relazione Doppiaggio

- solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

#### Relazione LinguaSottotitoli

- solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

#### Relazione Sottotitolaggio

- solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

#### Relazione PremioFilm

- solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

#### Relazione PremiazioneFilm

- solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

#### Relazione Attore



- CodiceFiscale → Nome, Cognome

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

#### Relazione Recitazione:

- Film, Attore → Principale

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

#### Relazione PremioAttore

- solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

#### Relazione PremiazioneAttore

- solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

**N.B.** Tutti i vincoli d'integrità sono stati implementati come **trigger**.

Sono stati implementati anche tutti i controlli necessari nelle operazioni richieste.

## **Area utenti**

**Utente**(Email, Codice, Nome, Cognome, Password).

**Fatturazione**(Fattura, Utente, Scadenza, Data, CartaDiCredito, PianoDiAbbonamento)

**PianoDiAbbonamento**(Pacchetto, Tariffa, MaxOre, MaxGigabyte)..

**Caratterizzazione**(Descrizione).

**Caratteristica**(PianoDiAbbonamento, Caratterizzazione)

**Eta**(Anni).

**Limite**(PianoDiAbbonamento, Età).

**ConnessionePassata**(Codice, IndirizzoIP, Dispositivo, Inizio, Fine, Utente, Film, Formato, Server)

**ConnessioneAttuale**(Codice, IndirizzoIP, Dispositivo, Inizio, Utente, Server, Formato)

**Dispositivo**(Nome, FormatoIdeale)

## **Vincoli di integrità referenziale**

- Tra l'attributo utente della relazione Fatturazione e la chiave (Codice) della relazione utente.
- Tra l'attributo PianoDiAbbonamento di Fatturazione e la chiave (Pacchetto) della relazione PianoDiAbbonamento.
- Tra l'attributo PianoDiAbbonamento della relazione Caratteristica e la chiave (Pacchetto) della relazione PianoDiAbbonamento.
- Tra l'attributo Caratterizzazione della relazione Caratteristica e la chiave (Descrizione) della relazione Caratterizzazione.
- Tra l'attributo Utente della relazione ConnessionePassata e la chiave (Codice) della relazione Utente.
- Tra l'attributo Utente della relazione ConnessioneAttuale e la chiave (Codice) della relazione Utente.
- Tra l'attributo Server della relazione ConnessioneAttuale e la chiave (Identificatore) della relazione Server\*\*.
- Tra l'attributo Server della relazione ConnessionePassata e la chiave (identificatore) della relazione Server\*\*.
- Tra l'attributo Formato della relazione ConnessioneAttuale e la chiave (Codice) della relazione Formato
- Tra l'attributo Formato della relazione ConnessionePassata e la chiave (Codice) della relazione Formato o la chiave (Codice) della relazione FormatoPassato.
- Tra l'attributo dispositivo della relazione ConnessioneAttuale e la chiave (Nome) della relazione Dispositivo.

- Tra l'attributo dispositivo della relazione ConnessionePassata e la chiave (Nome) della relazione Dispositivo.
- Tra l'attributo Film della relazione ConnessionePassata e la chiave (Codice) della relazione Film.
- Tra l'attributo FormatoIdeale della relazione Dispositivo e la chiave (Codice) di Formato.
- Tra l'attributo PianoDiAbbonamento della relazione Limite e la chiave (Pacchetto) della relazione PianoDiAbbonamento.
- Tra l'attributo Età della relazione Limite e la chiave (Età) dell'entità Età.

**\*\***La relazione Server è presente nella sezione "Area server" della traduzione nel modello logico. Nonostante più in basso vi sia un'apposita sezione per le relazioni generiche, abbiamo tradotto ora la *relationship* Sessione, visto che è (1,1) con ConnessioneAttuale.

## **Dipendenze funzionali**

### Relazione Utente

- Codice → Email, Nome, Cognome, Password
- Email → Codice, Nome, Cognome, Password

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

### Relazione Fatturazione:

- Fattura → Utente, Scadenza, Data, CartaDiCredito, PianoDiAbbonamento

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

**N.B.:** Scadenza non è univocamente determinata da Data, in quanto in caso di disdetta la Scadenza può subire variazioni.

### Relazione PianoDiAbbonamento

- Pacchetto → Tariffa, MaxOre, MaxGigabyte

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

### Relazione Caratterizzazione

- solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

#### Relazione Caratteristica

- solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

#### Relazione Età

- solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

#### Relazione Limite

- solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

#### Relazione ConnessionePassata

- Codice → IndirizzoIP, Dispositivo, Inizio, Fine, Utente, Film, Formato, Dispositivo, Server.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

#### Relazione ConnessioneAttuale

- Codice → IndirizzoIP, Dispositivo, Inizio, Utente, Server, Formato, Dispositivo

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

#### Relazione Ideale\*\*\*

- Nome → Formato

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

\*\*\*La relazione in realtà sarebbe tra quelle classificate "generiche", ma avendo cardinalità (1,1) con un'entità dell'area utente, decidiamo di inserire anch'essa nell'area utente.

### **Area formati**

**Formato**(Codice, LarghezzaRapporto, AltezzaRapporto, LunghezzaRisoluzione, AltezzaRisoluzione, Codec, Bitrate, QualitàVideo, QualitàAudio, DataRilascio).

**FormatoPassato**(Codice, LarghezzaRapporto, AltezzaRapporto, LunghezzaRisoluzione, AltezzaRisoluzione, Bitrate, Codec, QualitàVideo, QualitàAudio, DataRilascio, DataFineValidità, FormatoAttuale).

### **Vincoli di integrità referenziale**

- Tra l'attributo FormatoAttuale della relazione FormatoPassato e la chiave (Codice) della relazione Formato.

### **Dipendenze funzionali**

#### **Relazione Formato**

- Codice → RapportoDiAspetto, Risoluzione, Codec, Bitrate, QualitàVideo, QualitàAudio, DataRilascio.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

#### **Relazione FormatoPassato**

- Codice → RapportoDiAspetto, Risoluzione, Codec, Bitrate, QualitàVideo, QualitàAudio, DataRilascio, DataFineValidità, FormatoAttuale.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

## **Area geografica**

### **AreaGeografica(Codice)**

Vincoli di integrità referenziale:  
nessuno.

### **Dipendenze funzionali**

#### **Relazione AreaGeografica:**

- solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

## **Area streaming**

### **Server(Identificatore, Capacità, Banda, AreaGeografica, OffCounter)**

### **Vincoli di integrità referenziale**

- Tra l'attributo AreaGeografica e la chiave (Codice) della relazione AreaGeografica.

### **Dipendenze funzionali**

- Identificatore → Capacità, Banda, AreaGeografica, OffCounter

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

### **Relazioni “generiche”**

**Contenuto**(Server, Film).

**VisualizzazionePassata**(Utente, Film, NumeroVisualizzazioni, Stelle).

**VisualizzazioneAttuale**(Utente, Film).

**Caching**(Utente, Server, Film).

**Disponibilita**(Film, Formato, DimensioneFile).

**DisponibilitaPassata**(Film, FormatoPassato, DimensioneFile)

**Limitazione**(PianoDiAbbonamento, AreaGeografica).

**Restrizione**(Formato, AreaGeografica).

**RestrizionePassata**(Formato, AreaGeografica).

**Distanza**(Server, AreaGeografica)

### **Vincoli di integrità referenziale**

- Tra l'attributo Server di Contenuto e la chiave (identificatore) della relazione Server.
- Tra l'attributo Film di Contenuto e la chiave (codice) della relazione Film.
- Tra l'attributo Utente di VisualizzazionePassata e la chiave (codice) della relazione Utente.
- Tra l'attributo Film di VisualizzazionePassata e la chiave (codice) della relazione Film.
- Tra l'attributo Utente di VisualizzazioneAttuale e la chiave (codice) della relazione Utente.
- Tra l'attributo Film di VisualizzazioneAttuale e la chiave (codice) della relazione Film.
- Tra l'attributo Server di Caching e la chiave (identificatore) della relazione Server.
- Tra l'attributo Film di Caching e la chiave (codice) della relazione Film.
- Tra l'attributo Utente di Caching e la chiave (codice) della relazione Utente.
- Tra l'attributo Film di Disponibilita e la chiave (codice) della relazione Film.
- Tra l'attributo Formato di Disponibilita e la chiave (Codice) della relazione Formato.
- Tra l'attributo Film di Disponibilita passata e la chiave (Codice) della relazione Film.
- Tra l'attributo Formato di Disponibilita passata e la chiave (Codice) della relazione Formato passato.
- Tra l'attributo Formato di Restrizione e la chiave (Codice) della relazione Formato.

- Tra l'attributo AreaGeografica di Restrizione e la chiave (Codice) della relazione AreaGeografica.
- Tra l'attributo Formato di RestrizionePassata e la chiave (Codice) della relazione Formato.
- Tra l'attributo AreaGeografica di RestrizionePassata e la chiave (Codice) della relazione AreaGeografica.
- Tra l'attributo PianoDiAbbonamento della relazione Limitazione e la chiave (Pacchetto) della relazione PianoDiAbbonamento.
- Tra l'attributo AreaGeografica della relazione Limitazione e la chiave (Codice) della relazione AreaGeografica.
- Tra l'attributo Server di Distanza e la chiave (identificatore) della relazione Server.
- Tra l'attributo AreaGeografica di Distanza e la chiave (Codice) di AreaGeografica.

## **Dipendenze funzionali**

### Relazione Contenuto

- Solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

### Relazione Valutazione

- Utente, Film → Stelle

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

### Relazione VisualizzazionePassata

- Utente, Film → NumeroVisualizzazioni

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

### Relazione VisualizzazioneAttuale

- Solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.



### Relazione Caching

- Solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

### Relazione Disponibilit 

- File, Formato → DimensioneFile

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

### Relazione Limitazione

- Solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

### Relazione Restrizione

- Solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

### Relazione RestrizionePassata

- Solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

### Relazione Distanza

- Solo dipendenze funzionali banali.

La relazione è in BCNF, visto che una superchiave (chiave) è presente a sinistra di ogni dipendenza funzionale.

#### Operazione 2.4.4 (Raccomandazione di contenuti)

Per ogni utente, si guarda ogni occorrenza della *relationship* Visualizzazione Passata e la correlata occorrenza di Film (per sapere il genere del Film). Così facendo, sappiamo il numero di visualizzazioni di quell'utente per ogni genere da lui visto.

Questa funzione consiglierà a lui il film col rating più alto che non ha ancora visto del suo genere preferito (in caso di parità di rating, si prende il film più visualizzato).

Per quanto riguarda il formato da consigliarli, si ragiona in questo modo:

- Tramite l'entità Connessione Passata, si guarda quale sia il dispositivo più utilizzato dall'utente.
- Tramite la *relationship* Ideale, si guarda quale sia il formato consigliato per quel dispositivo
- A questo punto, si consiglia all'utente il formato ideale per il dispositivo da lui più utilizzato.

Si noti che saranno consigliati solo formati attuali.

Se i film non sono disponibili nel formato ideale per quel dispositivo, allora non restituisce un formato consigliato per quel film.

#### **Tavola degli accessi**

Concetto	Costrutto	Accessi	Tipo	Motivazione
Utente	E	1	L	Per risalire alle visualizzazioni e alle connessioni passate
Visualizzazione passata	R	30	L	Per risalire ai film e al numero di volte che li ha visti.
Film	E	30	L	Per risalire al genere del film e al voto medio del film che ha visto
Film	E	170	L	Per risalire ai film che non ha ancora visto.
Accesso Passato	R	25	L	Per risalire alle connessioni passate
Connessione passata	E	25	L	Per risalire ai dispositivi con cui si connette

Supporto	R	25	L	Per risalire ai dispositivi
Dispositivo	E	25	L	Per trovare il suo dispositivo “preferito”
Ideale	R	1	L	Per trovare il formato ideale per il suo dispositivo “preferito”
Disponibilità Attuale	R	3	L	Per verificare se il film trovato è disponibile nel formato ideale per il dispositivo trovato

**N.B.** Gli accessi passati per utente sarebbero 24.5, e non 25. Abbiamo deciso di approssimarli a 25.

Trovato il dispositivo da lui usato più frequentemente, basta vedere il formato ideale per quel solo tipo di dispositivo.

Si noti che questa operazione sfrutta la ridondanza. Se non la utilizzasse, avremmo dovuto ricalcolare il voto medio per ogni film del suo genere preferito.  
Le considerazioni sono analoghe a quelle fatte per l'operazione 3.

### Operazione 2.5.3 (Caching)

#### Passaggi

- 1) Si cerca il film che l'utente ha guardato di più. In caso di parità si sceglie quello col voto medio più alto.
- 2) Si cerca il server a cui l'utente si è connesso più volte\*.
- 3) Si inserisce quel film (se non ancora presente) in quel server.
- 4) Si usa l'operazione 2.4.4 per consigliare un film nuovo all'utente (quindi si inserisce quel film - appunto, se non ancora presente - nel server di *caching*): il film con voto medio degli utenti maggiore per genere preferito dall'utente stesso.
- 5) Si fa questo per due server a utente.

\*Si ipotizza che i due server più vicini all'utente siano quelli a cui si è connesso più spesso nel tempo.

#### Tavola degli accessi

Concetto	Costrutto	Accessi	Tipo	Motivazione
Utente	E	1.000	L	Per trovare il film che ha visto più volte e il server a cui si è connesso più di frequente
Visualizzazione e Passata	R	30.000	L	Per trovare i film visualizzati più spesso
Film	E	200 (i possibili pari merito)	L	Per trovare il voto medio dei pari merito.  Per ottenere il codice dei film non parimerito.  Per risalire al genere del film e al voto medio del film che ha visto.  Per risalire ai film che non ha ancora visto.
Accesso Passato	R	24.500	L	Per trovare i server a cui l'utente si è connesso più spesso.

Connessione Passata	E	24.500	L	Per trovare i server a cui si è connesso più spesso
Sessione Passata	R	24.500	L	Per trovare i server a cui si è connesso più spesso
Server	E	40	L	Per ottenere l'identificatore del server
Caching	E	4000	S	Per memorizzare in Caching le n-uple Film-Server-Utente

Vista la grande copiosità di operazioni, questa operazione sarà fatta una sola volta al mese tramite un *event*. Sarà fatta in un orario in cui il traffico è minimo (le 03:00 di mattina).

**N.B.** Si suppone che tutte le occorrenze di entità utilizzate in quest'operazione restino nel *buffer* per tutta la durata dell'operazione stessa. Dunque, visto che la modalità di accesso a ogni occorrenza di entità non cambia durante l'operazione (sono tutte letture - tranne per la *relationship Caching*, che è una scrittura), non dovremo ri-accedere alla stessa occorrenza di entità per tutta la durata dell'operazione (quindi ogni entità appare una sola volta nella tavola degli accessi).

## Operazione per il cambio di server durante la visione di un film

Lo stato di un server dipende dal:

- Numero di sessioni in quel server (quindi di utenti collegati)
- Dalla sua banda
- Dalla sua capacità.

In particolare, esso rispetta la formula:

$$Stato = Banda \times Capacità / 10.000$$

Dove:

- Con “Stato”, si indica il numero massimo di connessioni attuali a quel server per cui viene garantita una performance adeguata.
- Banda si misura in Gbit/s e è la velocità di download di un film da quel server.
- Capacità è il numero massimo di utenti che possono essere connessi a quel server.

La banda di un server, come scritto nella parte riguardante i vincoli d'integrità, è compresa tra 0,2 Gbit/s e 0,9 Gbit/s.

Quando un server raggiunge uno stato pari a  $Banda \times Capacità / 10.000$  e viene svolto un controllo dello stato dei server, esso sarà “svuotato” di metà utenti. Si noti che lo stato non indica che il server non possa più ricevere connessione per via del raggiungimento della massima capacità, ma perché, oltre quel numero di connessioni, esso calerebbe nettamente in prestazioni.

Quindi, quando un server raggiunge uno stato maggiore o uguale a  $Banda \times Capacità / 10.000$ , esso non è più in condizioni ottimali da un punto di vista prestazionale.

Gli utenti vengono nel server più vicino alla loro area geografica che contiene il film che stavano guardando (escluso questo server, a meno che il server in questione, che è stato parzialmente svuotato, non sia tornato in uno stato “buono”. In tal caso, l'utente non sarà spostato dal suo server.

Tutto questo accade senza il rischio che il server torni in uno stato negativo, dato che l'integrità dei server viene controllata per ogni inserimento).

Visto che la formula per il calcolo di uno stato dipende dalle proprietà fisiche del server, risulta impossibile stabilire una tavola degli accessi “universale” che sia soddisfacente da un punto di vista di numero di accessi; pertanto, mostreremo solo quali entità (e come) saranno coinvolte in questa operazione.

Si osservi che questa operazione non sfrutta le ridondanze: per calcolare lo stato di un server, basta contare le connessioni attuali che lo riguardano e confrontare quel numero con  $Banda \times Capacità / 10.000$ .

Questa operazione sarà implementata con un *event*, cosicché non sia possibile eccedere la capacità del server.

L'*event* scatterà ogni 15 minuti.

Quindi, per ogni server dovremo fare i seguenti accessi:

#### Tavola degli accessi

Concetto	Costrutto	Accessi	Tipo	Motivazione
Server	E	1	S	Per aggiornare il suo off-counter*
Sessione	E	Non stimabile senza errori trascurabili**	S	Per calcolare il nuovo stato del server e verificare se necessario effettuare un cambio di server per alcuni utenti.

\*L'off-counter viene gestito dall'operazione per il bilanciamento del carico (pagina 61).

\*\*Visto che ci sono server che sostengono un maggior numero di utenti mantenendo un'alta prestazione e server che non lo fanno, non ha senso parlare di "numero medio di connessioni attuali per server". In più, ci saranno più connessioni ai server prossimi alle aree geografiche in cui risiede la maggior parte degli utenti che utilizzano *FilmSphere*, pertanto risulta ancora meno sensato parlare di "numero medio di utenti per server". Volendo fare una stima, le connessioni attuali per server sono mediamente  $500/40 = 12,5$ , che tuttavia risulta probabilmente poco precisa, dal momento che, date anche le differenze di orario, il traffico non sarà mai distribuito equamente nel globo.

Nel caso ci sia una richiesta di connessione che manderebbe il server oltre la sua capacità, gli utenti saranno smistati in server decisi con i criteri discussi sopra. Pertanto, l'operazione avrà anche questa parte:

#### Tavola degli accessi

Concetto	Costrutto	Accessi	Tipo	Motivazione
Sessione	R	500	S	Per cancellare la sessione attuale
Sessione	R	$500+500^*$	L	Per sapere quante



				connessioni attuali ci sono per i potenziali nuovi server per quell'utente
Connessione Attuale	E	500	S	Per cancellare la connessione attuale
Accesso	R	500	S	Per cancellare l'accesso attuale
Supporto Attuale	R	500	S	Per cancellare il supporto attuale
Formato Visualizzato	R	500	S	Per cancellare il formato visualizzato
Sessione Passata	R	500	S	Per mettere tra quelle passate la sessione appena terminata
Connessione Passata	E	500	S	Per mettere tra quella passate la connessione appena terminata
Accesso passato	R	500	S	Per mettere tra quelli passati l'accesso appena terminato
Supporto Passato	R	500	S	Per riferire il dispositivo della precedente connessione attuale alla nuova connessione passata
Connessione Attuale	E	500	S	Per inserire la connessione attuale
Server	E	40	L	Per trovare l'identificatore del nuovo server a cui l'utente si conatterà
Accesso	R	500	S	Per legare la nuova connessione attuale al "nuovo" server
Supporto Attuale	R	500	S	Per legare la nuova connessione attuale al dispositivo con cui

				l'utente si connette
--	--	--	--	----------------------

\*Le sessioni totali iniziali sono 500. Tuttavia, nel caso peggiore, ogni volta che un utente viene disconnesso e riconnesso a un altro server, si crea una nuova occorrenza di sessione che non era tra quelle iniziale e che quindi va aggiunta agli accessi. Supponendo che tutti e 500 le sessioni in corso debbano essere cambiate, il 500-esimo utente vedrà 500 diverse occorrenze di sessione rispetto a quelle iniziale, perciò ci saranno 500 accessi ulteriori totali.

Questa è la parte di cui non siamo in grado di stimare gli accessi senza compiere errori non trascurabili (i risultati stimati non sarebbero validi):

- Se un server è piccolo, l'operazione non sarà molto costosa.
- Se un server è grande, il costo sarà influente.

La tabella sopra rappresenta il caso peggiore. Nel caso medio, si suppone che ci siano 250 accessi a ogni concetto (e 20 ai server), dunque, sempre considerando doppi gli accessi in scrittura rispetto a quelli in lettura, si avrebbero 3520 accessi in lettura ogni 15 minuti.

Gli accessi sono molti, ma utilizzare una ridondanza per questa operazione risulta inefficace perché servono le informazioni in *real time*. Mantenere aggiornata una ridondanza in ogni istante sarebbe estremamente costoso.

## Area analytics

### Classifiche

Generalmente, una persona si connette al server che ha più vicino. Le classifiche sono stilate in base alla geolocalizzazione dell'utente, e non in base alla locazione del server.

Tramite l'attributo Film dell'entità Connessione Passata è possibile sapere quali film sono stati guardati, quante volte per ogni utente. Questo attributo è ricavato dal vincolo d'integrità che sussiste tra "Visualizzazione Attuale" e "Connessione Attuale): appena una connessione attuale diventa "passata", il codice del film che l'utente stava visualizzando viene inserito nell'attributo Film dell'entità Connessione passata. Visto che ogni connessione è associata a un solo utente, è possibile sapere il piano di abbonamento dell'utente che ha guardato il film.

Quindi, abbiamo tutti i dati necessari per stilare una classifica in cui, per ogni area geografica (intesa, appunto, come geolocalizzazione dell'utente), a parità di modalità di abbonamento, sappiamo i film più visti.

Tramite la relazione "Formato visualizzato", sappiamo il formato con cui un utente sta visualizzando il film presente nell'occorrenza di relationship di Visualizzazione attuale.

Nel momento in cui una connessione attuale diventa una connessione passata, tramite la relazione Formato visualizzato si ricava l'attributo Formato di Connessione Passata e tramite la relazione Visualizzazione attuale si ricava l'attributo Film sempre di Connessione Passata.

Così ogni Connessione Passata è legata ad una coppia Film-Formato ed è possibile stilare una classifica sulla base di questi due fattori, che mostrerà le accoppiate Film-Formato più di successo in ogni area geografica e a parità di piano di abbonamento.

Si noti che la classifica è in base alle visualizzazioni che quel film ha avuto in quell'area geografica a parità di abbonamento. In più, è stato affiancato a ogni film il formato in cui, in quell'area geografica e a parità di abbonamento, è stato visualizzato più volte.

### **Bilanciamento del carico**

Ogni server ha un contatore che memorizza quante volte il server è stato in stato OFF in un certo lasso di tempo (questo contatore viene azzerato periodicamente da un *Event*). Tutti i server che hanno questo contatore oltre una certa soglia subiranno un bilanciamento del carico. Tramite le connessioni passate si ricava il film più visto presso quel server. Dopo, si ricava il server o nella stessa area geografica o in generale nell'area geografica più vicina che non ha questo film tramite le relazioni Contenuto e Distanza e si memorizza il film anche su questo server, solamente se questo server ha un contatore di stati OFF sotto una certa soglia (in caso di parimerito a livello di distanza, si metterà nel server col numero minore di off-counter), nel qual caso si cercherà il secondo server più vicino senza quel contenuto disponibile, e così via.

Affinché un server sia in sovraccarico, deve valere:

Off-Counter  $\geq 5$

(Se la distanza tra un server in sovraccarico e il primo disponibile supera 1000 km, il film non sarà inserito nel nuovo server).

L'*Event* è mensile.

## **Custom analytics: classifiche di gradimento dei Piani di Abbonamento**

La funzionalità che proponiamo stila mensilmente (ogni primo giorno del mese, per l'esattezza) 3 classifiche:

- Gli abbonamenti in ordine decrescente rispetto al numero di rinnovi.
- Gli abbonamenti in ordine decrescente rispetto al numero di disdette.
- Gli abbonamenti in ordine decrescente rispetto a quante volte un abbonamento sia stato scelto come abbonamento iniziale.

In più, stampa le informazioni riguardanti gli abbonamenti primi classificati.

Questa funzionalità è interessante, perché permette agli amministratori del servizio di analizzare quanto un abbonamento sia richiesto e, in caso di molte disdette, comprenderne il perché. Per esempio:

- Se un abbonamento ha molte disdette mensili, si possono analizzare vari fattori:
  - Il prezzo
  - I film offerti dall'abbonamento meno visti dagli utenti
  - Le sue limitazioni.

Ad esempio, se un abbonamento offre dei film che, in media, sono visti da pochissimi utenti e si osserva che gli utenti stessi, dopo aver guardato dei precisi contenuti (magari gli stessi - circa - per ogni utente), smettono di connettersi al servizio, significa che questo abbonamento dovrebbe cambiare una parte dei suoi contenuti: questa parte è proprio quella dei film non visti.

Invece, se troviamo un abbonamento che ha moltissimi rinnovi mensili, possiamo analizzare sempre i film visti dagli utenti e il numero di connessioni di uno stesso utente per capire cosa lo spinga a utilizzare così tanto il servizio. Avere queste classifiche ci aiuta, perché ci rende chiaro quali fattori dobbiamo analizzare affinché si voglia incrementare l'affluenza di utenti.