

Progetto di

Basi di Dati

A.A. 2022-2023

Università di Pisa

Dipartimento di **Ingegneria dell'Informazione**
Corso di Laurea Triennale **Ingegneria Informatica**



Studenti:

Federico Nardi
Lorenzo Melani

Docenti:

prof. Tonellotto
prof. Pistolesi

Sommario

1. Introduzione.....	4
2. Progettazione Concettuale	5
2.1 Dizionario delle Entità	5
2.2 Dizionario delle Relazioni.....	7
2.3 Regole di Derivazione.....	9
2.4 Descrizione E-R	10
2.4.1 Formato.....	10
2.4.2 IP range e geolocalizzazione.....	10
2.4.3 Restrizioni.....	11
3. Ristrutturazione.....	12
3.1 Eliminazione delle Generalizzazioni	12
3.2 Analisi delle Ridondanze.....	12
4. Tavole dei Volumi	13
4.1 Area Contenuti.....	13
4.2 Area Film.....	14
4.3 Area Utente	15
4.4 Area Streaming	16
5. Analisi delle Operazioni sui Dati	17
5.1 Nuova Connessione.....	17
5.2 Registrazione Utente.....	18
5.3 Fine Erogazione.....	19
5.4 Emissione Nuova Fattura.....	21
5.5 Pagamento Fattura	22
5.6 Stampa Premi Attore.....	23
5.7 Stampa Premi Regista.....	24
5.8 Stampa Premi Film	25
6. Progettazione Logica	26
6.1 Traduzione in Schema Logico	26
6.2 Implementazione	27
6.3 Vincoli.....	30
6.4 Analisi delle Dipendenze Funzionali e Normalizzazione.....	31
7. Area Analytics.....	33

7.1 Classifiche	33
7.2 Bilanciamento del Carico	35
8. Implementazione Fisica	37
8.1 Note su Vincoli	37
8.2 Popolamento	39
8.3 Trigger	40
8.3.1 Controllo Password	40
8.3.2 Controllo Numero Carta di Credito	40
8.3.3 Numero Dispositivi Connessi	40
8.3.4 Conteggio Stelle dei Critici	40
8.3.5 Conteggio Stelle degli Utenti	41
8.3.6 Calcolo Bitrate Totale di un File	41
8.3.7 Calcolo Banda Disponibile	41

1. Introduzione

La traccia del Progetto dell'A.A. 2022-2023 propone di realizzare una Base di Dati chiamata FilmSphere per l'implementazione di un servizio di memorizzazione e streaming di contenuti video online.

Il portale, come delineato dalla traccia, memorizza anche informazioni sensibili degli utenti, come dati di fatturazione, abbonamenti e connessioni ed informazioni tecniche, come ad esempio la dislocazione dei contenuti nei server che compongono la CDN.

Ci sarà anche modo di gestire il lato della sicurezza dei dati e dell'analisi dei dati sensibili riguardanti soprattutto l'area dell'Utente e tutte le fatture emesse.

Ad ogni componente dello schema ER che riguarda una delle diverse Aree della base di dati viene univocamente associato un colore:



Figura 1: Legenda ripresa dal nostro schema E-R

Più avanti avremo modo di approfondire come abbiamo progettato e successivamente gestito la nostra Base di Dati coi rispettivi capitoli della Documentazione.

2. Progettazione Concettuale

2.1 Dizionario delle Entità

NOME	ATTRIBUTI	PRIMARY KEY	DESCRIZIONE
Utente	Codice, Email, Nome, Cognome, Password	Codice	Persona che usufruisce del servizio offerto dalla piattaforma.
Dispositivo	IDDispositivo, Modello, RisoluzioneSchermo	IDDispositivo	Strumento attraverso il quale l'utente si connette alla piattaforma.
Connessione	Oralizio, OraFine, IP	Oralizio, Dispositivo (FK)	Istanza di accesso dell'utente alla piattaforma.
Erogazione	Oralizio, OraFine, IDErogazione	IDErogazione	Distribuzione di un contenuto richiesto da un utente connesso.
Server	BandaDisponibile, MaxCapacita, IDServer, Banda, Latitudine, Longitudine	IDServer	Centro di immagazzinamento e distribuzione dei contenuti.
Piano di Abbonamento	Nome, MaxGiga, MaxOre, MaxQualita, FunzioneAggiuntiva, Costo	Nome	Categoria di servizio di cui l'utente può usufruire.
Fattura	IDFattura, Scadenza, DataEmissione, Importo	IDFattura	Documento di richiesta di pagamento indirizzato all'utente.
Carta di Credito	NumeroCarta, CVV, DataScadenza	NumeroCarta	Mezzo di pagamento del servizio.
File	IDFile, Dimensione	IDFile	Contenuto multimediale offerto dalla piattaforma.
Formato	Codice, Bitrate	Codice	Codifica dei dati relativi a un file.
Audio (entità figlia)	Frequenza		Codifica dei dati audio.

Video (entità figlia)	Risoluzione, Rapporto		Codifica dei dati video.
Paese	Nome, IPMin, IPMax, Latitudine, Longitudine	Nome	Nazione che ha prodotto contenuti e/o verso cui vengono distribuiti.
Film	ID, Titolo, Descrizione, AnnoP, VotoTotale, Durata, Genere	ID	Produzione cinematografica disponibile sulla piattaforma.
Critico	IDCritico, Nome, Cognome	IDCritico	Esperto cinematografico che valuta i film.
Lingua	NomeLingua	NomeLingua	Linguaggio in cui un film è doppiato o sottotitolato.
Attore	IDAttore, Nome, Cognome	IDAttore	Artista che recita nelle produzioni cinematografiche.
Regista	IDRegista, Nome, Cognome	IDRegista	Artista che si occupa della direzione della produzione.
Premio	IDPremio, Nome	IDPremio	Riconoscimento verso un film, un attore o un regista.

In fase di ristrutturazione è stata eliminata la generalizzazione dell'entità Formato accorpendo il padre alle entità figlie (Formato Audio e Formato Video).

Sempre in fase di ristrutturazione sono state aggiunte, per maggiore efficienza nella manipolazione dei dati, i seguenti attributi:

- *BitrateTotale* in File
- *StelleUtenti* in Film
- *StelleCritici* in Film

2.2 Dizionario delle Relazioni

NOME	DESCRIZIONE	PARTECIPAZIONE	ATTRIBUTI
Utilizzo	Associazione di un utente a un dispositivo per la connessione.	Utente (1,N), Dispositivo (1,1)	-
Associazione	Individua da quale dispositivo viene svolta una connessione.	Dispositivo (1,N), Connessione (1,1)	-
Gestione	Presa in carico di un'erogazione da parte di una connessione.	Connessione (0,N), Erogazione (1,1)	-
Presenza	Associazione tra ciascun server e i file che vi sono memorizzati.	Server (1,N), File (1,N)	-
Richiesta	Corrispondenza tra l'erogazione di un contenuto e il server che la sta eseguendo.	Server (1,N), Erogazione (1,1)	-
Accesso	Corrispondenza tra l'erogazione di un contenuto e il file interessato.	File (0,N), Erogazione (1,1)	-
Sottoscrizione	Associazione tra l'utente e il piano di abbonamento scelto.	Utente (1,1), Piano di Abbonamento (1,N)	-
Disponibilità	Corrispondenza tra un piano di abbonamento e i contenuti visualizzabili.	File (1,N), Piano di Abbonamento (1,N)	-
Intestazione	Associazione tra una fattura e l'utente intestatario.	Fattura (1,1), Utente (1,N)	-
Proprietà	Associazione tra una carta di credito e l'utente che la utilizza per pagare.	Utente (1,N), Carta di Credito (1,N)	-

Pagamento	Corrispondenza tra una fattura e la carta con cui è stata pagata.	Fattura (0,1), Carta di Credito (0,N)	-
Codifica	Associazione tra un file e il formato dei suoi dati.	File (1,1), Formato (1,N)	DataRilascio
Restrizione Formato	Associazione tra ciascun formato e gli eventuali paesi in cui esso non è disponibile.	Formato (0,N), Paese (0,N)	-
Restrizione Geografica	Associazione tra un file e gli eventuali paesi in cui esse non è disponibile.	File (0,N), Paese (0,N)	-
Produzione	Corrispondenza tra un film e il paese in cui è stato prodotto.	Film (1,1), Paese (0,N)	-
Rappresentazione	Corrispondenza tra un file e il film che codifica.	File (1,1), Film (1,N)	-
Recensione Utente	Valutazione di un utente relativa a un film	Utente (0,N), Film (0,N)	ValutazioneU
Critica	Valutazione di film da parte di un critico.	Critico (1,N), Film (1,N)	ValutazioneC
Sottotitolaggio	Associazione tra un film e le lingue in cui sono disponibili i sottotitoli.	Film (0,N), Lingua (0,N)	-
Doppiaggio	Associazione tra un film e le lingue in cui è stato doppiato.	Film (0,N), Lingua (0,N)	-
Recitazione	Associazione tra un attore e i film in cui ha recitato.	Film (0,N), Attore (1,N)	-
Regia	Associazione tra un regista e i film che ha diretto.	Film (1,N), Regista (1,N)	-
Premiazione Attore	Associazione tra un attore e i premi che ha vinto.	Attore (0,N), Premio (1,N)	-
Premiazione Regista	Associazione tra un regista e i premi che ha vinto.	Regista (0,N), Premio (1,N)	-
Premiazione Film	Associazione tra un film e i premi che ha vinto.	Film (0,N), Premio (1,N)	-

In fase di ristrutturazione, in seguito all'eliminazione della generalizzazione dell'entità Formato, la relazione Restrizione Formato è stata divisa in Restrizione Audio e Restrizione Video; lo stesso vale per la relazione Codifica che è stata divisa in Codifica Audio e Codifica Video.

2.3 Regole di Derivazione

In fase di ristrutturazione sono stati aggiunti gli attributi sopra elencati. Essi, insieme ad altri già presenti nell'ER non ristrutturato, sono valori derivabili da altre tabelle. Vediamo come:

- BitrateTotale in File è calcolato come la somma dei campi Bitrate di Formato Audio e Formato Video relativi a un certo file;
- StelleUtenti in Film è calcolato come la somma delle votazioni degli utenti rispetto a un certo film;
- StelleCritici in Film è calcolato come la somma delle votazioni dei critici rispetto a un certo film;
- VotoTotale in Film è calcolato come la media pesata delle valutazioni degli utenti e dei critici rispetto a un certo film;
- BandaDisponibile in Server è calcolato come la differenza tra la banda massima e i bitrate totali dei file erogati dal server ad un certo istante.

2.4 Descrizione E-R

In questo paragrafo viene spiegata la struttura generale del diagramma ER e si prova a chiarire punti che potrebbero risultare ambigui.

Il diagramma E-R proposto si divide in **quattro aree**:

- **Area Clienti**: racchiude tutto ciò che riguarda l'utente e i suoi dati, compresi i pagamenti;
- **Area Contenuti**: sono le entità e relazioni che riguardano i file e le loro specifiche tecniche;
- **Area Streaming**: contiene ciò che riguarda l'erogazione dei contenuti dai server ai dispositivi;
- **Area Film**: presenta tutte le informazioni relative ai film, premi, recensioni, attori e registi.

2.4.1 Formato

Nell'ER non ristrutturato è presente una generalizzazione esclusiva e totale per quanto riguarda l'entità Formato. E' stata preferita questa soluzione rispetto a quella che avrebbe previsto due entità separate in quanto i formati audio e quelli video condividono due attributi fondamentali per il database: il codice e il bitrate. I formati video sono caratterizzati dalla risoluzione, che ne indica la qualità, e dal rapporto; ai formati audio è associata invece la frequenza. File codificati in formati di qualità maggiore avranno ovviamente un bitrate e una dimensione maggiori, con le dovute conseguenze sulla loro erogazione da parte del server (maggior occupazione della banda) e sull'occupazione di memoria nei Point of Presence.

2.4.2 IP range e geolocalizzazione

L'entità Paese presenta gli attributi IPMin e IPMax: abbiamo supposto che a ogni paese sia assegnato staticamente un range di indirizzi IP. A partire dall'indirizzo IP relativo a una connessione è possibile geolocalizzare da dove un utente si sta connettendo e scegliere a quale server indirizzare le richieste sulla base della distanza calcolata con le coordinate relative a un paese e quelle dei server. Si tratta chiaramente di un'approssimazione poiché, avendo condensato un paese in una singola coppia di coordinate, un utente che si connette potrebbe trovarsi più vicino a un server diverso rispetto a

quello che gli viene assegnato ma considerando una granularità a livello nazionale si tratta comunque di una buona approssimazione.

2.4.3 Restrizioni

Le restrizioni previste dal nostro database riguardano interi contenuti o soltanto certi tipi di formato: una restrizione rispetto a un contenuto in un paese potrebbe essere dovuta a motivi di censura, mentre la restrizione di un certo tipo di formato potrebbe essere causata da motivi tecnologici. Non sono stati posti vincoli di partecipazione obbligatoria alle relazioni di restrizione in quanto un contenuto può o non può essere sottoposto a restrizioni (lo stesso vale per un formato) e un paese può o non può applicare restrizioni.

3. Ristrutturazione

In fase di ristrutturazione sono state eseguite due azioni principali: eliminazione di generalizzazioni e analisi e inserimento di ridondanze.

3.1 Eliminazione delle Generalizzazioni

Il diagramma ER non ristrutturato presenta una sola generalizzazione: quella dell'entità formato. Si tratta di una generalizzazione esclusiva, in quanto se un formato è di tipo audio non può essere di tipo video e viceversa, e completa, in quanto tutti i tipi di formato sono racchiusi in queste due categorie. Proprio guidati dall'esclusività, abbiamo scelto di non accorpare le entità figlie in un'unica entità, che a nostro avviso sarebbe risultata troppo generale, ma di mantenerle separate distribuendo su ciascuna delle due gli attributi che si trovavano sull'entità padre Formato.

3.2 Analisi delle Ridondanze

A causa della scelta fatta sulla generalizzazione di Formato abbiamo scelto di inserire l'attributo ridondante BitrateTotale in File: questa ridondanza ci permette di non dover risalire alle tabelle Formato Audio e Formato Video ogni volta in cui abbiamo bisogno del bitrate di un certo file. Considerato il fatto che questo dato è necessario ogni qualvolta si voglia iniziare una nuova erogazione o chiuderne una in corso per aggiornare la banda disponibile di un certo server abbiamo ritenuto necessario l'inserimento del nuovo attributo. Il costo di questa ridondanza è uno scan della tabella Formato Audio e uno scan della tabella Formato Video al momento dell'inserimento di un nuovo file nel database. Il risparmio è non dover fare queste ricerche ogni inizio o fine erogazione.

Anche l'attributo BandaDisponibile in Server è una ridondanza in quanto calcolabile a partire dalla Banda del server e dai bitrate dei file che esso sta erogando in un certo istante. Il costo di questa ridondanza è la necessità di essere aggiornata ogni inizio e fine erogazione, il guadagno consiste nell'avere continuamente a disposizione un indicatore di carico del server oltre a un considerevole alleggerimento dell'operazione di ribilanciamento del carico che, altrimenti, necessiterebbe di calcolare sempre il grado di occupazione della banda scorrendo la tabella Erogazione che ha un volume molto superiore a quella Server.

Lo stesso discorso vale per gli attributi StelleUtenti e StelleCritici che facilitano il calcolo del rating di un certo film.

4. Tavole dei Volumi

4.1 Area Contenuti

NOME	TIPO	CALCOLO VOLUME	MOTIVAZIONE
File	E ¹	2.000 ²	1 Film ha 2 Formati
FormatoAudio	E	6	Supposizione iniziale
FormatoVideo	E	8	Supposizione iniziale
Codifica Video	R ³	2.000	Relazione 1:1 con File
CodificaAudio	R	2.000	Relazione 1:1 con File
RestrizioneGeografica	R	3.000	Il 30% dei file non è disponibile nel 10% dei paesi
RestrizioneAudio	R	3.000	Il 30% dei file non è disponibile nel 10% dei paesi
RestrizioneVideo	R	3.000	Il 30% dei file non è disponibile nel 10% dei paesi
Paese	E	50	Supposizione iniziale

¹ E: indica Entità

² . : indica la separazione delle migliaia

³ R: indica Relazione

4.2 Area Film

NOME	TIPO	CALCOLO VOLUME	MOTIVAZIONE
Film	E	1.000	Supposizione iniziale
Produzione	E	1.000	Relazione (1,1) con Film
Lingua	E	30	Supposizione iniziale
Sottotitolaggio	R	3.000	1 Film è sottotitolato 3 lingue
Doppiaggio	R	5.000	1 Film è doppiato in 5 lingue
Attore	E	600	Il 60% dei Film, tenendo traccia degli Attori Principali
Recitazione	R	1.200	1 Attore ha recitato 2 Film in Media
Regista	E	100	Il 10% dei Film
Regia	R	1.100	1 Regista ha prodotto 11 Film in Media
Premio	E	50	Supposizione iniziale (40% Attori, 20% Registi e 40% Film)
PremiazioneAttore	R	600	Il 5% degli Attori ha vinto 1 Premio
PremiazioneRegista	R	50	Il 5% dei Registi ha vinto 1 Premio
PremiazioneFilm	R	1.000	Il 5% dei Film ha vinto 1 Premio
Rappresentazione	R	2.000	Relazione (1,1) con File
Critico	E	50	Supposizione Iniziale
Critica	R	1.500	1 Critico ha recensito 30 Film in Media
RecensioneUtente	R	75.000	1 Utente ha recensito 1.5 Film in Media

4.3 Area Utente

NOME	TIPO	CALCOLO VOLUME	MOTIVAZIONE
Utente	E	50.000	Supposizione iniziale
Sottoscrizione	R	50.000	Relazione (1,1) con Utente
PianodiAbbonamento	E	5	Supposizione iniziale
Disponibilita	R	8.000	In media ogni piano di abbonamento può usufruire dell'80% dei file
Intestazione	R	150.000	Relazione (1,1) con Fattura
Fattura	E	150.000	1 Utente ha mediamente 3 Fatture intestate
Pagamento	R	105.000	Il 70% delle Fatture sono state pagate
Proprieta	R	55.000	1 Utente può avere più Carte e ciascuna Carta può essere Intestata a più Utenti
CartadiCredito	E	60.000	Ogni Utente ha 1,2 Carte di Credito in Media

4.4 Area Streaming

NOME	TIPO	CALCOLO VOLUME	MOTIVAZIONE
Dispositivo	E	100.000	Ogni Utente utilizza 2 Dispositivi in Media
Utilizzo	R	100.000	Relazione (1,1) con Dispositivo
Connessione	E	1.000.000	Ogni Dispositivo si è connesso 10 volte in Media
Erogazione	E	1.500.000	Ogni Connessione gestisce 1.5 Erogazioni in Media
Gestione	R	1.500.000	Relazione (1,1) con Erogazione
Accesso	R	1.500.000	Relazione (1,1) con Erogazione
Richiesta	R	1.500.000	Relazione (1,1) con Erogazione
Presenza	R	8.000	Ogni Server contiene 50% dei File in Media
Server	E	8	Supposizione Iniziale
Associazione	R	1.000.000	Relazione (1,1) con Connessione

5. Analisi delle Operazioni sui Dati

5.1 Nuova Connessione

Input: IDDispositivo, IP, Inizio_Connessione, Fine_Connessione

Output: Booleano di controllo

Frequenza stimata: 30.000 nuove connessioni giornaliere

Descrizione: L'operazione permette di effettuare una nuova connessione al database a partire dall'indirizzo IP, un dispositivo e l'orario.

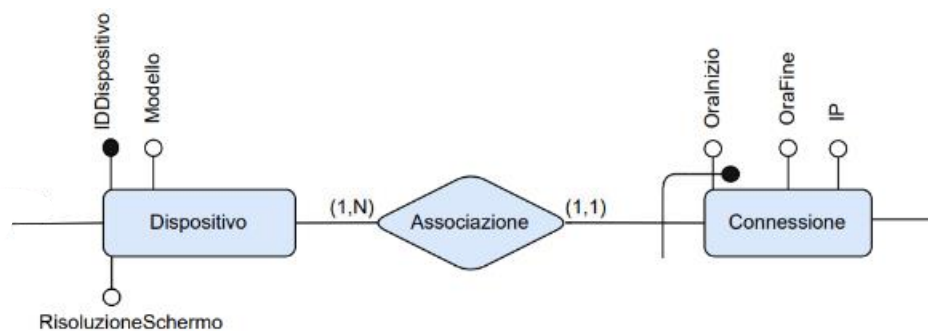


Tavola degli Accessi:

NOME	TIPO	CALCOLO DEGLI ACCESSI (GIORNALIERI)	MOTIVAZIONE
Connessione	E	30.000	Ipotizziamo 30.000 Connessioni giornaliere. L'operazione richiede 1 singolo accesso in scrittura

5.2 Registrazione Utente

Input: Codice, Nome, Cognome, Password, Email, Abbonamento

Output: Booleano di controllo

Frequenza stimata: 300 nuovi utenti giornalieri

Descrizione : Si consente l'inserimento di un nuovo utente nel database a partire dai dati personali che lo caratterizzano.

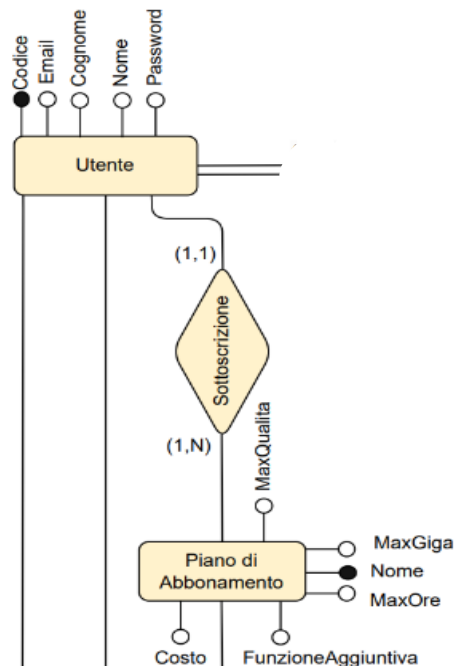


Tavola degli Accessi:

NOME	TIPO	CALCOLO DEGLI ACCESSI (GIORNALIERI)	MOTIVAZIONE
Utente	E	300	Ipotizziamo 300 Utenti abbonati giornalieri. Per ogni nuovo Abbonato è necessario 1 accesso in scrittura nella tabella Utente.

5.3 Fine Erogazione

Input: idErogazione

Output: Booleano di controllo

Frequenza stimata: 45.000 terminazioni di erogazione al giorno (abbiamo precedentemente ipotizzato 30.000 connessioni giornaliere, consideriamo una media di 1.5 erogazioni per connessione)

Descrizione: Operazione per porre fine all'erogazione di un contenuto. Prima di eseguire l'operazione aggiornando l'orario di fine erogazione si controlla che quest'ultimo valga NULL, cioè che si stia provando a chiudere un'erogazione in corso e non una già chiusa.

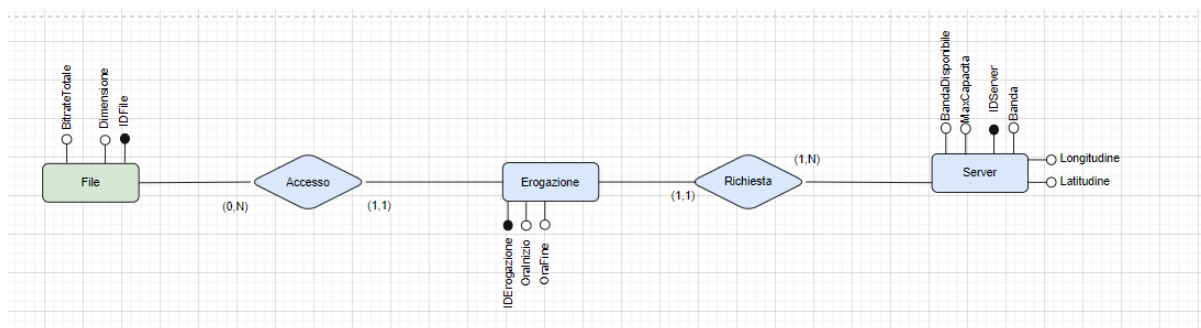


Tavola degli Accessi:

NOME	TIPO	CALCOLO DEGLI ACCESSI (GIORNALIERI)	MOTIVAZIONE
Erogazione	E	45.000	Ipotizziamo 30.000 Connessioni giornaliere che gestiscono mediamente 1.5 Erogazioni ciascuna. Ogni Erogazione deve essere terminata.
File	E	45.000	Accesso puntuale attraverso la chiave che si trova in Erogazione.
Server	E	45.000	Accesso puntuale attraverso la chiave che si trova in Erogazione.

Accesso	R	45.000	Ad ogni record di Erogazione corrisponde esattamente 1 File.
Richiesta	R	45.000	Ad ogni record di Erogazione corrisponde esattamente 1 Server.

5.4 Emissione Nuova Fattura

Input: IDFattura, Intestatario, Importo

Output: Booleano di controllo

Frequenza stimata: 1.700 fatture emesse al giorno (ipotizziamo una cadenza mensile di pagamento e una distribuzione omogenea dei 50.000 utenti sui 30 giorni)

Descrizione: Operazione che permette al gestore della piattaforma di emettere una nuova fattura a carico di un intestatario registrato.

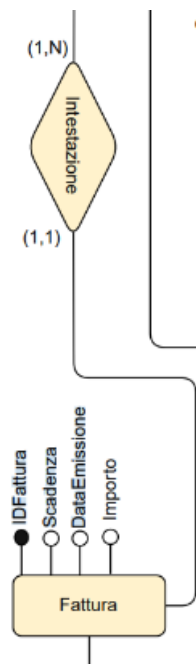


Tavola degli Accessi:

NOME	TIPO	CALCOLO DEGLI ACCESSI (GIORNALIERI)	MOTIVAZIONE
Fattura	E	1.700	Ipotizziamo una distribuzione giornaliera omogenea delle Fatture dei 50.000 Utenti sui 30 giorni mensili

5.5 Pagamento Fattura

Input: IDFattura, CartaPagamento

Output: Booleano di controllo

Frequenza stimata: 1.200 fatture pagate giornalmente (il 70% di quelle emesse)

Descrizione: Si registra l'avvenuto pagamento di una fattura, ammesso che non risulti già pagata.

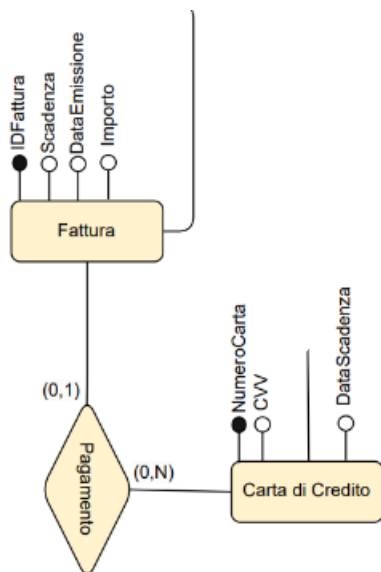


Tavola degli Accessi:

NOME	TIPO	CALCOLO DEGLI ACCESSI (GIORNALIERI)	MOTIVAZIONE
Fattura	E	1.200	Ipotizziamo il pagamento del 70% delle Fatture emesse giornalmente.

5.6 Stampa Premi Attore

Input: Identificatore dell'attore

Output: stampa dei premiattori richiesti

Frequenza stimata: 15.000 (metà degli utenti che si connettono giornalmente richiedono la lista dei premi)

Descrizione (con disegno e poche righe): Si permette, su richiesta dell'utente, di stampare i premi vinti da un attore.

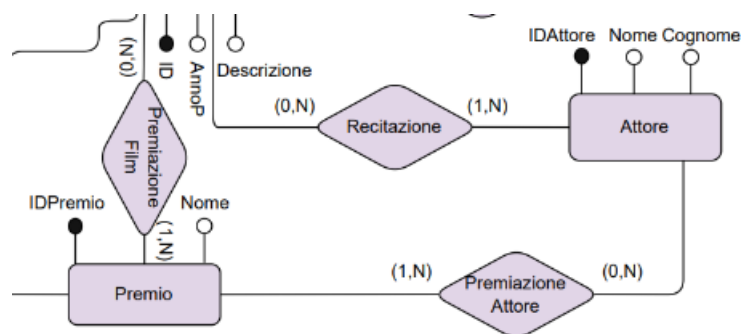


Tavola degli Accessi:

NOME	TIPO	CALCOLO DEGLI ACCESSI (GIORNALIERI)	MOTIVAZIONE
Attore	E	15.000	Ipotizziamo 1 Accesso puntuale per trovare il nome dell'Attore.
Premio	E	15.000	Ipotizziamo 1 Accesso puntuale per trovare il nome del Premio.
PremiazioneAttore	R	9.000.000	Ipotizziamo che metà degli Utenti che si connettono ogni giorno vogliono usufruire di questa Operazione. PremiazioneAttore è composta di 600 record ed è necessario un fullscan ad ogni chiamata.

5.7 Stampa Premi Regista

Input: Identificatore del regista

Output: stampa dei premiregista richiesti

Frequenza stimata: 15.000 (metà degli utenti che si connettono giornalmente richiedono la lista dei premi)

Descrizione: Si permette, su richiesta dell'utente, di stampare i premi vinti da un regista.

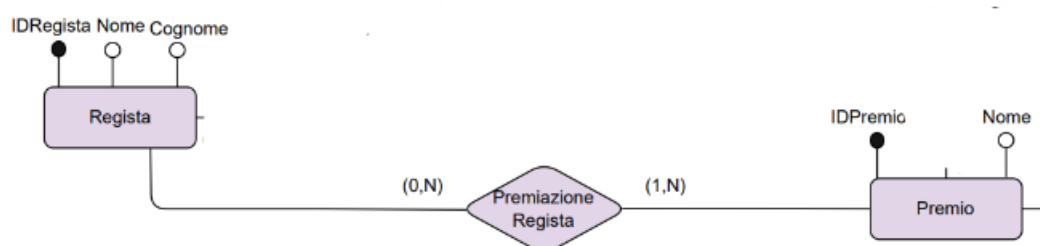


Tavola degli Accessi:

NOME	TIPO	CALCOLO DEGLI ACCESSI (GIORNALIERI)	MOTIVAZIONE
Regista	E	15.000	Ipotizziamo 1 Accesso puntuale per trovare il nome del Regista.
Premio	E	15.000	Ipotizziamo 1 Accesso puntuale per trovare il nome del Premio.
PremiazioneRegista	R	750.000	Ipotizziamo che metà degli Utenti che si connettono ogni giorno vogliono usufruire di questa Operazione. PremiazioneRegista è composta di 600 record ed è necessario un fullscan ad ogni chiamata.

5.8 Stampa Premi Film

Input: Identificatore del film

Output: stampa dei premi dei film richiesti

Frequenza stimata: 15.000 (metà degli utenti che si connettono giornalmente richiedono la lista dei premi)

Descrizione: Si permette, su richiesta dell'utente, di stampare i premi vinti da un film.

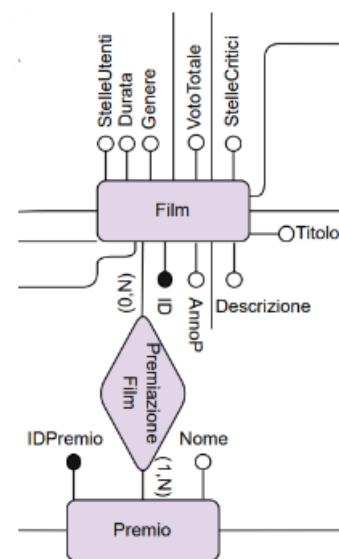


Tavola degli Accessi:

NOME	TIPO	CALCOLO DEGLI ACCESSI (GIORNALIERI)	MOTIVAZIONE
Film	E	15.000	Ipotizziamo 1 Accesso puntuale per trovare il titolo del Film.
Premio	E	15.000	Ipotizziamo 1 Accesso puntuale per trovare il nome del Premio.
PremiazioneFilm	R	15.000.000	Ipotizziamo che metà degli Utenti che si connettono ogni giorno vogliono usufruire di questa Operazione. Premiazione Film è composta di 1.000 record ed è necessario un fullscan ad ogni chiamata.

6. Progettazione Logica

6.1 Traduzione in Schema Logico

La traduzione dello schema concettuale in schema logico produce le seguenti tabelle:

- PianoDiAbbonamento(Nome, MaxQualita, MaxGiga, MaxOre, Costo, FunzioneAggiuntiva)
- Utente(Codice, Nome, Cognome, Password, Email, Abbonamento)
- Dispositivo(IDDispositivo, Modello, RisoluzioneSchermo, Utente)
- Connessione(OraInizio, Dispositivo, IP, OraFine)
- Server(IDServer, MaxCapacita, BandaDisponibile, Banda, Latitudine, Longitudine)
- Paese(Nome, IPMin, IPMax, Latitudine, Longitudine)
- Film(ID, AnnoP, Descrizione, Titolo, StelleCritici, StelleUtenti, Durata, Genere, VotoTotale, Paese)
- File(IDFile, Dimensione, BitrateTotale, Film, FormatoAudio, FormatoVideo, DataRilascioA, DataRilascioV)
- Erogazione(IDErogazione, OraInizio, OraFine, Server, File, Dispositivo)
- Presenza(Server, File)
- RestrizioneGeografica(File, Paese)
- FormatoVideo(Codice, Risoluzione, Bitrate, Rapporto)
- FormatoAudio(Codice, Frequenza, Bitrate)
- RestrizioneVideo(Paese, FormatoVideo)
- RestrizioneAudio(Paese, FormatoAudio)
- Lingua(NomeLingua)
- Sottotitolaggio(Film, Lingua)
- Doppiaggio(Film, Lingua)
- Attore(IDAttore, Nome, Cognome)
- Regista(IDRegista, Nome, Cognome)
- Regia(Film, Regista)
- Recitazione(Attore, Film)
- Premio(IDPremio, Nome)
- PremiazioneRegista(Regista, Premio)
- PremiazioneFilm(Film, Premio)

- PremiazioneAttore(Attore, Premio)
- Critico(IDCritico, Nome, Cognome)
- Critica(Critico, Film, ValutazioneC)
- RecensioneUtente(Utente, Film, ValutazioneU)
- Disponibilita(File, PianoDiAbbonamento)
- CartadiCredito(NumeroCarta, CVV, DataScadenza)
- Proprieta(Utente, CartadiCredito)
- Fattura(IDFattura, DataEmissione, Scadenza, Intestatario, Importo, CartaPagamento)

6.2 Implementazione

TABELLA	ATTRIBUTO	CODIFICA	NOTE
PianoDiAbbonamento	Nome MaxQualita MaxGiga MaxOre Costo FunzioneAggiuntiva	VARCHAR(100) INT UNSIGNED INT UNSIGNED INT UNSIGNED DOUBLE VARCHAR(100)	NULLABLE NULLABLE NULLABLE NULLABLE
Utente	Codice Nome Cognome Password Email Abbonamento	INT UNSIGNED VARCHAR(50) VARCHAR(50) VARCHAR(20) VARCHAR(100) VARCHAR(100)	
Dispositivo	IDDispositivo Modello RisoluzioneSchermo Utente	INT UNSIGNED VARCHAR(100) VARCHAR(100) INT UNSIGNED	
Connessione	Oralnizio Dispositivo IP OraFine	DATETIME INT UNSIGNED BIGINT DATETIME	NULLABLE
Server	IDServer MaxCapacita BandaDisponibile Banda Latitudine	INT UNSIGNED INT UNSIGNED INT UNSIGNED INT UNSIGNED DOUBLE	

Server	Longitudine	DOUBLE	
Paese	Nome IPMin IPMax Latitudine Longitudine	VARCHAR(100) BIGINT BIGINT DOUBLE DOUBLE	
Film	ID AnnoP Descrizione Titolo StelleCritici StelleUtenti Durata Genere VotoTotale Paese	INT UNSIGNED INT UNSIGNED VARCHAR(200) VARCHAR(50) INT UNSIGNED INT UNSIGNED INT UNSIGNED VARCHAR(100) DOUBLE UNSIGNED VARCHAR(100)	NULLABLE NULLABLE
File	IDFile Dimensione BitrateTotale Film FormatoAudio FormatoVideo DataRilascioA DataRilascioV	INT UNSIGNED INT UNSIGNED INT UNSIGNED INT UNSIGNED INT UNSIGNED INT UNSIGNED DATE TIME DATE TIME	NULLABLE
Erogazione	IDErogazione Oralnizio OraFine Server File Dispositivo	INT UNSIGNED DATE TIME DATE TIME INT UNSIGNED INT UNSIGNED INT UNSIGNED	NULLABLE
Presenza	Server File	INT UNSIGNED INT UNSIGNED	
RestrizioneGeografica	File Paese	INT UNSIGNED INT UNSIGNED	
FormatoVideo	Codice Risoluzione Bitrate Rapporto	INT UNSIGNED INT UNSIGNED INT UNSIGNED VARCHAR(10)	
FormatoAudio	Codice	INT UNSIGNED	

	Frequenza Bitrate	INT UNSIGNED INT UNSIGNED	
RestrizioneAudio	Paese FormatoAudio	VARCHAR(100) INT UNSIGNED	
RestrizioneVideo	Paese FormatoVideo	VARCHAR(100) INT UNSIGNED	
Lingua	NomeLingua	VARCHAR(100)	
Sottotitolaggio	Film Lingua	INT UNSIGNED VARCHAR(100)	
Doppiaggio	Film Lingua	INT UNSIGNED VARCHAR(100)	
Attore	IDAttore Nome Cognome	INT UNSIGNED VARCHAR(100) VARCHAR(100)	
Regista	IDRegista Nome Cognome	INT UNSIGNED VARCHAR(100) VARCHAR(100)	
Regia	Film Regista	INT UNSIGNED INT UNSIGNED	
Recitazione	Attore Film	INT UNSIGNED INT UNSIGNED	
Premio	IDPremio Nome	INT UNSIGNED VARCHAR(100)	
PremiazioneRegista	Regista Premio	INT UNSIGNED INT UNSIGNED	
PremiazioneFilm	Film Premio	INT UNSIGNED INT UNSIGNED	
PremiazioneAttore	Attore Premio	INT UNSIGNED INT UNSIGNED	
Critico	IDCritico Nome Cognome	INT UNSIGNED VARCHAR(100) VARCHAR(100)	
Critica	Critico Film ValutazioneC	INT UNSIGNED INT UNSIGNED INT UNSIGNED	
RecensioneUtente	Utente Film ValutazioneU	INT UNSIGNED INT UNSIGNED INT UNSIGNED	
Disponibilita	File PianoDiAbbonamento	INT UNSIGNED VARCHAR(100)	

CartadiCredito	NumeroCarta CVV DataScadenza	BIGINT INT UNSIGNED DATE	
Proprieta	Utente CartadiCredito	INT UNSIGNED BIGINT	
Fattura	IDFattura DataEmissione Scadenza Intestatario Importo CartaPagamento	INT UNSIGNED DATE DATE INT UNSIGNED DOUBLE BIGINT	NULLABLE

6.3 Vincoli

Nella progettazione della nostra Base di Dati abbiamo avuto modo di tenere traccia dei **vincoli** con cui ci siamo scontrati.

Partendo dai **Vincoli di Integrità Referenziale** abbiamo posizionato la nostra attenzione particolarmente sull'uso delle *Foreign Key*, dove nei capitoli successivi mostreremo la loro implementazione su codice.

Per quanto riguarda i **Vincoli di Tupla**, invece, tolto le Primary Key ed alcuni attributi che non possono essere NOT NULL, ce ne sono degli altri che possono essere messe a NULL, tra questi:

1. Piano di Abbonamento:
 - a. *MaxQualita*
 - b. *MaxGiga*
 - c. *MaxOre*
 - d. *Costo*
 - e. *FunzioneAggiuntiva*
2. Connessione:
 - a. *IP*
3. Film:
 - a. *StelleCritici*
 - b. *StelleUtenti*
4. File:
 - a. *BitrateTotale*

5. Erogazione:

a. *OraFine*

6. Fattura:

a. *CartaPagamento*

Abbiamo inserito anche dei **Vincoli di Dominio** sul dominio di alcuni attributi, tra questi:

1. Utente (*Password*): abbiamo impostato un controllo della password dell'Utente che deve essere di caratteri alfanumerici compresi tra 6 e 20;
2. CartadiCredito(*NumeroCarta*): abbiamo impostato un controllo del numero della Carta di Credito che deve essere cifre uguale a 16.

Infine, abbiamo aggiunto un ulteriore **vincolo intrarelazionale** che possiamo definire di "policy", permette ad un Utente di connettersi contemporaneamente alla piattaforma con un numero massimo di 2 dispositivi. Il controllo viene effettuato nella tabella *Connessione* quando si prova ad inserire una nuova connessione.

6.4 Analisi delle Dipendenze Funzionali e Normalizzazione

Con l'introduzione di identificatori ad hoc per quasi ogni entità si ha che per molte tabelle l'unica dipendenza funzionale non banale sia quella dalla chiave primaria verso gli altri attributi della n-upla: prendiamo per esempio Attore dove *IDAttore* implica *Nome* e *Cognome*.

Ci sono altri casi in cui, per la stessa relazione, esistono più dipendenze funzionali: vediamo il caso della relazione Paese. In Paese *IPMin* implica gli altri attributi, *IPMax* lo stesso e pure la coppia *Latitudine-Longitudine* determina il resto della tupla. È vero però che questi attributi risultano essere chiavi per la relazione Paese in quanto ogni Paese ha un proprio range di IP e una propria coppia di coordinate univoche. Per quanto riguarda la relazione Server ci troviamo nella stessa situazione: latitudine e longitudine implicano gli altri attributi, ma trattandosi di valori univoci per ciascuna tupla si ha sempre una relazione in forma normale.

C'è un caso particolare in cui abbiamo scelto di non portare una relazione in forma normale: la relazione Fattura. In questa relazione abbiamo l'identificatore che determina il resto della n-upla e lo stesso fa la coppia di

valori *Intestatario-DataEmissione*. Tuttavia vi è una dipendenza funzionale tra *DataEmissione* e *Scadenza* in quanto la scadenza sarà sempre due settimane dopo all'emissione: in questo caso l'implicante non è una chiave quindi non siamo in forma normale. Abbiamo deciso di mantenere questo caso speciale per diverse ragioni: esso riguarda un'area del database che non è soggetta a frequenti aggiornamenti o modifiche quindi il rischio di incongruenze è minimo, inoltre abbiamo fornito all'utilizzatore del database la possibilità di gestire l'emissione di una nuova fattura con un'apposita procedura che maneggia in modo consistente le date relative all'operazione. Abbiamo valutato che fosse utile non eliminare l'attributo *Scadenza* per lasciare le porte aperte all'implementazione semplificata di operazioni, come la verifica di eventuali debiti insoluti da parte di un utente, quindi la relazione non si trova in forma normale.

7. Area Analytics

7.1 Classifiche

Frequenza stimata: 3.000 richieste giornaliere

Descrizione: Vengono mostrate le classifiche in base al tipo di abbonamento e formati di visualizzazione. Il criterio di classificazione è il semplice numero di visualizzazioni di un film.

Tavola degli Accessi:

NOME	TIPO	CALCOLO DEGLI ACCESSI (GIORNALIERI)	MOTIVAZIONE
Erogazione	E	4.500.000.000	Ipotizziamo che il 10% degli Utenti che si connettono giornalmente vogliono usufruire di questa operazione (3.000 * 1.500.000)
File	E	6.000.000	Viene eseguito un full scan della tabella File
Connessione	E	3.000.000.000	Leggiamo la tabella per risalire a Dispositivo
Dispositivo	E	300.000.000	Leggiamo la tabella per risalire ad Utente
Utente	E	150.000.000	Leggiamo la tabella per conoscere il tipo di Abbonamento
Film	E	3.000.000	Leggiamo la tabella per conoscere il titolo dei Film
Accesso	R	4.500.000.000	Leggiamo la tabella per risalire al File relativo ad 1 Erogazione
Gestione	R	4.500.000.000	Leggiamo la tabella per risalire alla Connessione relativa ad 1 Erogazione

Associazione	R	3.000.000.000	Leggiamo la tabella per risalire al Dispositivo associato ad 1 Connessione
Utilizzo	R	300.000.000	Leggiamo la tabella per risalire all'Utente associato ad 1 Dispositivo
Rappresentazione	R	6.000.000	Leggiamo la tabella per risalire al Film rappresentato da 1 File

7.2 Bilanciamento del Carico

Frequenza stimata: 48 volte al giorno (una ogni mezz'ora)

Descrizione: Il bilanciamento del carico permette di spostare le erogazioni da un server a un altro in base al carico di lavoro di ciascuno di essi. Un server è considerato sovraccarico quando il rapporto tra la banda disponibile e la banda totale scende sotto al 20%. Per ogni server che soddisfa questa condizione si ricercano le erogazioni a suo carico ancora in corso e si spostano sul server con meno carico fra quelli disponibili finché non si scende sotto la soglia critica di occupazione della banda.

Tavola degli Accessi:

NOME	TIPO	CALCOLO DEGLI ACCESSI (GIORNALIERI)	MOTIVAZIONE
Erogazione	E	72.000.480	Ipotizziamo 1.500.000 di accessi in Lettura per trovare le Erogazioni gestite da un Server al momento dell'esecuzione dell'operazione e 10 in Scrittura per aggiornare le Erogazioni da spostare in un altro Server. Il tutto viene eseguito ogni 30 minuti.
Presenza	R	480.240	Ogni Server compare mediamente 1.000 volte nella tabella Presenza, quindi devo eseguire 1.000 accessi in lettura per ogni volta che sposto un'erogazione verso un altro Server. Supponiamo di dover spostare 10 erogazioni da un server sovraccarico e che in 5 casi il server destinazione

Presenza	R		non possenga il contenuto: sono necessari 5 accessi in scrittura.
Server	E	1.748	Ogni volta che viene eseguito il bilanciamento si fanno 2 scan completi per trovare il Server con Max Banda Disponibile e controllare se un Server è Sovraccarico. Supponiamo di trovare un Server sovraccarico da cui eliminare 10 Erogazioni (8+8+10+10 x48) Ipotesi che il Server non sarà più sovraccarico.
Accesso	R	72.000.480	Si fanno gli stessi accessi fatti in Erogazione.
File	E	480	Si accede solo per controllare il Bitrate del File una volta deciso quale spostare. Sono circa 10 accessi per ogni 30 minuti.
Richiesta	R	72.000.480	Si fanno gli stessi accessi fatti in Erogazione

8. Implementazione Fisica

8.1 Note su Vincoli

Come avevamo già accennato nel capitolo dei vincoli, abbiamo avuto modo di gestire i vincoli di integrità referenziale con l'utilizzo delle *Foreign Key*. Qui di seguito vi mostreremo quali sono le tabelle coinvolte usando il codice MySQL commentato:

- **TABELLA Utente:** FOREIGN KEY (`Abbonamento`) REFERENCES `PianoDiAbbonamento` (`Nome`)
- **TABELLA Dispositivo:** FOREIGN KEY (`Utente`) REFERENCES `Utente` (`Codice`)
- **TABELLA Connessione:** FOREIGN KEY (`Dispositivo`) REFERENCES `Dispositivo` (`IDDispositivo`)
- **TABELLA Film:** FOREIGN KEY (`Paese`) REFERENCES `Paese` (`Nome`)
- **TABELLA File:** FOREIGN KEY (`Film`) REFERENCES `Film` (`ID`)
- **TABELLA Erogazione:** FOREIGN KEY (`Server`) REFERENCES `Server` (`IDServer`),
FOREIGN KEY (`File`) REFERENCES `File` (`IDFile`),
FOREIGN KEY (`Dispositivo`) REFERENCES `Connessione` (`Dispositivo`)
- **TABELLA Presenza:** FOREIGN KEY (`Server`) REFERENCES `Server` (`IDServer`),
FOREIGN KEY (`File`) REFERENCES `File` (`IDFile`)
- **TABELLA RestrizioneGeografica:** FOREIGN KEY (`File`) REFERENCES `File` (`IDFile`),
FOREIGN KEY (`Paese`) REFERENCES `Paese` (`Nome`)
- **TABELLA RestrizioneVideo:** FOREIGN KEY (`Paese`) REFERENCES `Paese` (`Nome`),
FOREIGN KEY (`FormatoVideo`) REFERENCES `FormatoVideo` (`Codice`)
- **TABELLA RestrizioneAudio:** FOREIGN KEY (`Paese`) REFERENCES `Paese` (`Nome`),
FOREIGN KEY (`FormatoAudio`) REFERENCES `FormatoAudio` (`Codice`)
- **TABELLA Sottotitolaggio:** FOREIGN KEY (`Film`) REFERENCES `Film` (`ID`),
FOREIGN KEY (`Lingua`) REFERENCES `Lingua` (`NomeLingua`)
- **TABELLA Doppiaggio:** FOREIGN KEY (`Film`) REFERENCES `Film` (`ID`),
FOREIGN KEY (`Lingua`) REFERENCES `Lingua` (`NomeLingua`)

- **TABELLA Regia:** FOREIGN KEY (`Film`) REFERENCES `Film` (`ID`),
FOREIGN KEY (`Regista`) REFERENCES `Regista` (`IDRegista`)
- **TABELLA Recitazione:** FOREIGN KEY (`Film`) REFERENCES `Film` (`ID`),
FOREIGN KEY (`Attore`) REFERENCES `Attore` (`IDAttore`))
- **TABELLA PremiazioneRegista:** FOREIGN KEY (`Regista`) REFERENCES
`Regista` (`IDRegista`),
FOREIGN KEY (`Premio`) REFERENCES `Premio` (`IDPremio`)
- **TABELLA PremiazioneFilm:** FOREIGN KEY (`Film`) REFERENCES `Film`
(`ID`),
FOREIGN KEY (`Premio`) REFERENCES `Premio` (`IDPremio`)
- **TABELLA PremiazioneAttore:** FOREIGN KEY (`Attore`) REFERENCES
`Attore` (`IDAttore`),
FOREIGN KEY (`Premio`) REFERENCES `Premio` (`IDPremio`))
- **TABELLA Critica:** FOREIGN KEY (`Critico`) REFERENCES `Critico`
(`IDCritico`),
FOREIGN KEY (`Film`) REFERENCES `Film` (`ID`)
- **TABELLA RecensioneUtente:** FOREIGN KEY (`Utente`) REFERENCES
`Utente` (`Codice`),
FOREIGN KEY (`Film`) REFERENCES `Film` (`ID`)
- **TABELLA Disponibilita:** FOREIGN KEY (`File`) REFERENCES `File` (`IDFile`),
FOREIGN KEY (`PianoDiAbbonamento`) REFERENCES
`PianoDiAbbonamento` (`Nome`)
- **TABELLA Proprieta:** FOREIGN KEY (`Utente`) REFERENCES `Utente`
(`Codice`),
FOREIGN KEY (`CartadiCredito`) REFERENCES `CartadiCredito`
(`NumeroCarta`)
- **TABELLA Fattura:** FOREIGN KEY (`Intestatario`) REFERENCES `Utente`
(`Codice`),
FOREIGN KEY (`CartaPagamento`) REFERENCES `CartaDiCredito`
(`NumeroCarta`)

8.2 Popolamento

Il popolamento è stato fatto a scopo esemplificativo senza la pretesa di rispettare i volumi riportati nelle apposite tavole.

Abbiamo cercato di mantenere una coerenza coi dati inseriti rispettando i vincoli posti in fase di progettazione concettuale e logica.

Naturalmente i dati inseriti non fanno riferimento a persone o cose realmente esistenti, tranne i titoli dei Film, ma sono stati generati grazie all'utilizzo dell'AI ChatGPT 3.5 di OpenAI.

8.3 Trigger

Abbiamo inserito dei trigger di controllo di dominio sui dati di rilevante importanza quali la Password e la Carta di Credito.

Un ulteriore trigger controlla il numero dei dispositivi connessi contemporaneamente relativi ad un Utente.

Infine, vi sono i trigger per la gestione delle ridondanze.

8.3.1 Controllo Password

Tipo: BEFORE INSERT

Tabella: Utente

Descrizione: Verifica che la password sia composta da un numero di caratteri alfanumerici compresi tra 6 e 20.

8.3.2 Controllo Numero Carta di Credito

Tipo: BEFORE INSERT

Tabella: CartadiCredito

Descrizione: Verifica che il numero della Carta di Credito sia esattamente 16 cifre.

8.3.3 Numero Dispositivi Connessi

Tipo: BEFORE INSERT

Tabella: Connessione

Descrizione: Prima di inserire una nuova connessione verifichiamo che l'utente che la richiede non abbia già più di 2 Connessioni attive.

8.3.4 Conteggio Stelle dei Critici

Tipo: AFTER INSERT

Tabella: Critica

Descrizione: Dopo l'inserimento di una nuova critica si aggiorna il valore delle stelle dei Critici assegnati ad un certo Film.

8.3.5 Conteggio Stelle degli Utenti

Tipo: AFTER INSERT

Tabella: RecensioneUtente

Descrizione: Dopo l'inserimento di una nuova recensione si aggiorna il valore delle stelle degli Utenti ad un certo Film.

8.3.6 Calcolo Bitrate Totale di un File

Tipo: BEFORE INSERT

Tabella: File

Descrizione: Prima di inserire un record nella tabella File si imposta il valore del Bitrate Totale del record che si vuole inserire pari alla somma dei Bitrate dei Formati Audio e Video.

8.3.7 Calcolo Banda Disponibile

Tipo: AFTER INSERT

Tabella: Erogazione

Descrizione: Quando inizia una nuova Erogazione si aggiorna la Banda Disponibile del Server che eroga il contenuto sottraendovi il Bitrate del contenuto erogato.