



UNIVERSITÀ DI PISA

Progetto Basi di Dati A.A. 2022/23

Michele Meazzini - Elia Tonci

Indice

Indice.....	1
1 Progettazione concettuale.....	3
1.1 Entità.....	3
1.1.1 Area Clienti.....	3
1.1.2 Area Contenuti.....	4
1.1.3 Area Server.....	5
1.1.4 Area Formati.....	5
1.2 Relazioni.....	6
1.2.1 Area Clienti.....	6
1.2.2 Area Contenuti.....	7
1.2.3 Area Server.....	7
1.2.4 Area Formati.....	7
1.3 Supposizioni.....	8
2 Ristrutturazione.....	10
2.1 Eliminazione generalizzazioni.....	10
2.1.1 Generalizzazione sulla tabella Artista.....	10
2.1.2 Generalizzazione sulla tabella Formato.....	10
2.2 Eliminazione degli attributi multivalore.....	10
2.3 Analisi ridondanze.....	10
3 Tavole dei volumi.....	12
3.1 Area Clienti.....	12
3.2 Area Contenuti.....	13
3.3 Area Server.....	13
3.4 Area Formati.....	14
4 Operazioni sul database.....	15
4.1 Operazione 1 - Registrazione utente.....	15
4.2 Operazione 2 - Nuova fattura.....	15
4.3 Operazione 3a - Sottoscrizione abbonamento.....	16
4.4 Operazione 3b - Scadenza abbonamento.....	17
4.5 Operazione 4 - Inserimento artista.....	18
4.6 Operazione 5 - Creazione film.....	18
4.7 Operazione 6a - Inserimento contenuto.....	19
4.8 Operazione 6b - Modifica offerta vecchi contenuti.....	20
4.9 Operazione 7 - Inserimento recensione.....	21
4.10 Operazione 8a - Visualizzazione contenuto.....	22
4.11 Operazione 8b - Fine visualizzazione.....	23
4.12 Operazione 8c - Sposta utenti su altri server.....	24
4.13 Operazione 9a - Download di contenuti.....	26
4.14 Operazione 9b - Elimina contenuti scaricati.....	26
4.15 Operazione 10 - Caching di contenuti.....	27
5 Area analytics.....	29
5.1 Classifiche.....	29

5.2 Bilanciamento del carico.....	29
6 Progettazione logica.....	31
6.1 Schema logico.....	31
6.1.1 Area Clienti.....	31
6.1.2 Area Contenuti.....	32
6.1.3 Area Server.....	32
6.1.4 Area Formati.....	32
6.2 Vincoli di integrità referenziale.....	33
6.3 Vincoli di tupla.....	34
6.4 Vincoli di dominio.....	34
6.5 Vincoli inter-relazionali.....	34
6.6 Dipendenze funzionali e Normalizzazione.....	35
6.6.1 Area Clienti.....	35
6.6.2 Area Contenuti.....	36
6.6.3 Area Server.....	36
6.6.4 Area Formati.....	37
7 Implementazione.....	38
7.1 Trigger e Events.....	38
7.2 Popolamento.....	39

1 Progettazione concettuale

1.1 Entità

1.1.1 Area Clienti

Entità	Attributi	Identificatori	Descrizione
Utente	ID, Nome, Cognome, Password, Email, Critico, ContenutiScaricati	ID	Utente registrato alla piattaforma, che può non aver sottoscritto un abbonamento
Connessione	ID_Utente, TimestampInizio, TimestampFine, IndirizzoIP, Dispositivo	ID_Utente, TimestampInizio	Una nuova connessione avviene quando un utente si connette alla piattaforma per guardare un contenuto
Abbonamento	Nome, TariffaMensile, ContenutiScaricabili, Caratteristiche	Nome	Abbonamento fornito dalla piattaforma
Pagamento*	ID, Importo, Data, DataPagamento, NumeroCarta, ScadenzaCarta, CVCCarta	ID	Pagamento di un utente salvato sulla piattaforma
CarteDiCredito	NumeroCarta, ScadenzaCarta, CVCCarta	NumeroCarta	Carta di credito utilizzata da un utente per pagare
Fattura	ID, Importo, Data, DataPagamento	ID	Documento contenente i dati riguardanti il pagamento per utilizzare la piattaforma

*In fase di ristrutturazione la tabella Pagamento è stata trasformata in una relazione e sono state aggiunte le entità CartaDiCredito e Fattura.

1.1.2 Area Contenuti

Entità	Attributi	Identificatori	Descrizione
Contenuto	ID, DataRilascio, Dimensione, BitrateTotale	ID	Contenuto visualizzabile dall'utente, è il film con la specifica combinazione di formato audio e video
Film	ID, Descrizione, Durata, Genere, Titolo, DataPubblicazione, PaeseDescrizione, LingueAudio, LingueSottotitoli, PunteggioRecensioni, PunteggioArtisti	ID	Film presente nella piattaforma
Artista	ID, Nome, Cognome, Ruolo, Popolarità	ID	Generalizzazione di Attore e Regista
Attore**	ID, Nome, Cognome, Ruolo, Popolarità, PremiAttore	ID	Specializzazione di Artista. Attore principale che recita nel film
Regista**	ID, Nome, Cognome, Ruolo, Popolarità, PremiRegista	ID	Specializzazione di Artista. Regista che produce il film

**In fase di ristrutturazione le tabelle Attore e Regista sono state accorpate in Artista.

1.1.3 Area Server

Entità	Attributi	Identificatori	Descrizione
Server	ID, Stato, AreaGeografica, CapacitàMax, LarghezzaBanda	ID	Server contenente determinati film, a lo scopo di far connettere più persone insieme
AreaGeografica	IndirizzoIP, NomePaese	IndirizzoIP	Nazione da dove avviene il collegamento al server

1.1.4 Area Formati

Entità	Attributi	Identificatori	Descrizione
Formato***	ID, Nome, Bitrate, Qualità	ID	Generalizzazione di FormatoVideo e FormatoAudio
FormatoVideo	ID, Nome, Bitrate, Rateo, Risoluzione, Qualità	ID	Specializzazione di Formato, Formato video in cui verrà visualizzato il contenuto
FormatoAudio	ID, Nome, Bitrate, Qualità	ID	Specializzazione di Formato. Formato audio con cui verrà visualizzato il contenuto

***In fase di ristrutturazione la tabella Formato è stata eliminata mantenendo le entità figlie FormatoVideo e FormatoAudio

1.2 Relazioni

1.2.1 Area Clienti

Relazione	Cardinalità	Attributi	Descrizione
Metodo di pagamento	Utente(o, N) CartaDiCredito(1, 1)	-	Registrazione del metodo di pagamento di ogni utente
Pagamento	CartaDiCredito(o, N) Fattura(1, 1)	-	Pagamento di una fattura con una carta di credito
Accesso	Connessione(1, 1) Utente(o, N)	-	Accesso alla piattaforma da parte di un utente
Iscrizione	Abbonamento(o, N) Utente(1, 1)	Scadenza	Iscrizione ad un abbonamento da parte di un utente
Offerta	Abbonamento(1, N) Contenuto(1, N)	Visualizzabile	Offerta di un contenuto fornita da un abbonamento
Visualizzazione	Utente(o, N) Contenuto(o, N)	TimestampInizio	Visualizzazione di un contenuto da parte di un utente
Recensione	Utente(o, N) Film(o, N)	PunteggioUtente	Recensione di un film da parte di un utente
Download	Utente(o, N) Contenuto(o, N)	-	Download di un contenuto da parte di un utente

1.2.2 Area Contenuti

Relazione	Cardinalità	Attributi	Descrizione
Disponibilità	Contenuto(1, 1) Film(o, N)	-	Disponibilità di un film in determinati formati
Produzione	Film(1, N) Artista(1, N)	-	Produzione di un film da parte di un artista

1.2.3 Area Server

Relazione	Cardinalità	Attributi	Descrizione
Possesso	Server(o, N) Contenuto(o, N)	-	Possesso di un contenuto da parte di un server
Collegamento	Server(o, N) Connessione(1, 1)	-	Collegamento ad un server da parte di un utente connesso
Copertura	Server(1, N) AreaGeografica(o, N)	Ping	Copertura di un'area geografica da parte di un server
Restrizione	AreaGeografica(o, N) Contenuto(o, N)	-	Contenuto proibito in un'area geografica

1.2.4 Area Formati

Relazione	Cardinalità	Attributi	Descrizione
Codifica****	Formato(1, N) Contenuto(1, N)	-	Codifica di un contenuto in un formato video e audio
CodificaVideo	FormatoVideo(o, N) Contenuto(1, 1)	-	Codifica di un contenuto in un formato video
CodificaAudio	FormatoAudio(o, N) Contenuto(1, 1)	-	Codifica di un contenuto in un formato audio

****In fase di ristrutturazione la relazione Codifica è stata eliminata e sono state create al suo posto le relazioni CodificaAudio e CodificaVideo

1.3 Supposizioni

- I premi considerati per i registi sono: Premio Oscar, Palma d'Oro a Cannes e Premi al festival di Venezia (leone d'oro o leone d'argento).
- I premi considerati per gli attori sono: Premio Oscar, Premio NAFTA e il Golden Globe.
- I premi ottenibili da un artista valgono 2, e non possono superare il valore totale di 30, inoltre ogni presenza dell'artista in un film vale 2 e non può superare il valore 20, questi due valori sommati formano l'attributo Popolarità (da 0 a 50).
- PunteggioRecensioni è calcolato come la media ponderata dei voti dati dagli utenti (x_1) e dai critici (x_{10}), mentre PunteggioArtisti è la media delle popolarità degli artisti che compaiono nella produzione del film.
- Lo Stato di un server è calcolato come la somma dei bitrate dei contenuti che quel server sta trasmettendo, esso è quindi un valore numerico che, se prossimo al valore di CapacitàMax indica che il server è pieno.
- Nella produzione di un film vengono considerati solo il regista e un unico attore principale così da facilitare l'operazione di inserimento di un nuovo film.
- I formati video considerati sono:
 - AVCHD: file di massima qualità con dimensioni ridotte
 - AVI: file di alta qualità e di grandi dimensioni
 - MOV: file con qualità di medio livello
 - MP4: file di minima qualità
- I formati audio considerati sono:
 - WAV: file di massima qualità senza perdita di dati
 - AAC: file di alta qualità con perdita di dati ma buona compressione
 - MP3: file di minima qualità con perdita di dati

- I vari formati si presentano in 6 combinazioni audio/video. Di seguito l'elenco delle combinazioni, da quella con qualità peggiore a quella con qualità migliore (tra parentesi gli abbonamenti per i quali sono disponibili):
 - MP3/MP4 (Basic, Premium, Pro, Deluxe, Ultimate)
 - MP3/MOV (Premium, Pro, Deluxe, Ultimate)
 - AAC/MOV (Premium, Pro, Deluxe, Ultimate)
 - AAC/AVI (Pro, Deluxe, Ultimate)
 - WAV/AVI (Pro, Deluxe, Ultimate)
 - WAV/AVCHD (Ultimate)
- Gli abbonamenti Pro, Deluxe e Ultimate comprendono la possibilità di scaricare rispettivamente 5, 10 e 15 contenuti (qualsiasi) da rivedere in un secondo momento senza collegamento internet.

2 Ristrutturazione

2.1 Eliminazione generalizzazioni

2.1.1 Generalizzazione sulla tabella Artista

La generalizzazione sulla tabella Artista è totale, specializzata in Attore e Regista, per distinguere il ruolo che un artista ha nella produzione di un film.

Gli unici attributi propri di Attore e Regista sono legati ai premi ricevuti.

La generalizzazione è stata eliminata accorpendo le figlie nel padre aggiungendo l'attributo Ruolo che assume i valori "attore" o "regista", quest'ultimo avrà l'attributo premiAttore a NULL e viceversa per gli attori.

Questa scelta è stata fatta dato che l'entità padre ha molti più attributi rispetto alle figlie, in questo modo si vanno a generare pochi valori nulli.

2.1.2 Generalizzazione sulla tabella Formato

La generalizzazione sulla tabella Formato è totale, specializzata in Formato audio e Formato video, tra le entità figlie solo Formato video ha attributi specifici.

La generalizzazione è stata eliminata accorpendo l'entità padre nelle figlie, sono così state create le entità FormatoAudio e FormatoVideo.

Entrambe le entità hanno relazioni solo con l'entità Contenuto tramite Codifica audio e Codifica video, entrambe di cardinalità (0, N) dalla parte delle entità Formato e (1, 1) dalla parte di Contenuto. Si può osservare che le cardinalità sono cambiate rispetto all'ER non ristrutturato dove avevamo (1, N) rispetto ad entrambe le entità.

È stato deciso di accorpare l'entità padre nelle figlie così da non andare a creare troppi attributi con valore nullo e consentire comunque una gestione efficiente dei dati.

2.2 Eliminazione degli attributi multivalore

L'unico attributo multivalore presente nel diagramma ER è Dati carta, dell'entità Pagamento, esso è stato eliminato mantenendo i tre attributi di cui era composto come attributi distinti.

2.3 Analisi ridondanze

L'attributo Visualizzabile della tabella Offerta non è strettamente necessario e rappresenta una ridondanza, calcolabile con la data di rilascio del contenuto, ma è stato inserito per semplificare l'algoritmo che aggiorna l'offerta degli abbonamenti che non comprendono la visualizzazione delle ultime uscite (Paragrafo 4.8).

Nell'E-R non ristrutturato l'entità Contenuto non comprende l'attributo BitrateContenuto, questo attributo rappresenta una ridondanza nell'E-R ristrutturato, ma è stato inserito per aumentare le prestazioni. Il bitrate totale del

Contenuto veniva infatti ricalcolato ogniqualvolta volessimo aggiornare o verificare l'attributo Stato nella tabella Server, per un totale di 3 volte nelle operazioni individuate, in ciascuna delle quali occorre 2 query per il calcolo del bitrate del formato video e del formato audio e un'operazione di somma. Dato che il ricalcolo del bitrate totale veniva fatto in operazioni molto frequenti è stato quindi deciso di calcolarlo una volta per tutte direttamente nella creazione del contenuto stesso (Paragrafo 4.7), operazione che avviene sicuramente con frequenza molto minore rispetto alla visualizzazione del contenuto.

L'attributo Stato nella tabella Server rappresenta una ridondanza, esso infatti è calcolabile sommando i bitrate necessari alla visualizzazione dei contenuti che quel server sta trasmettendo. Si è rivelata necessaria però la sua aggiunta nel database in quanto fondamentale, nel momento in cui un utente vuole visualizzare un contenuto, per verificare quali server siano disponibili alla fruizione.

3 Tavole dei volumi

3.1 Area Clienti

Concetto	Tipo	Volume	Note
Utente	E	1.000.000	Ipotesi iniziale
MetodoDiPagamento	R	2.000.000	Cardinalità (1, 1) con la tabella Carta di Credito
Carta di Credito	E	2.000.000	Si considerano in media due carte per ogni utente
Pagamento	R	12.000.000	Cardinalità (1, 1) con la tabella Fattura
Fattura	E	12.000.000	Si considera una fattura al mese per ogni utente per un anno
Accesso	R	180.000.000	Cardinalità (1, 1) con la tabella Connessione
Connessione	E	180.000.000	Si considerano in media una connessione ogni due giorni per ogni utente
Iscrizione	R	1.000.000	Cardinalità (1, 1) con la tabella Utente
Abbonamento	E	5	Ipotesi iniziale
Offerta	R	20.000	Deriva dalla somma delle 1.000 offerte per l'abbonamento basic, 3.000 per premium, 5.000 per pro e deluxe e 6.000 per ultimate
Visualizzazione	R	180.000.000	Ad ogni connessione corrisponde una visualizzazione
Recensione	R	50.000	Si ipotizzano in media 50 recensioni per ogni film
Download	R	5.000.000	Si considerano in media 5 download per ogni utente

3.2 Area Contenuti

Concetto	Tipo	Volume	Note
Contenuto	E	6.000	Si considerano in media 6 tipi di contenuto per ogni Film
Disponibilità	R	6.000	Cardinalità (1, 1) con la tabella Contenuto
Film	E	1.000	Ipotesi iniziale
Produzione	R	Max 2.000	(Al massimo possono esserci due artisti per ogni film)
Artista	E	Max 2.000	Al massimo possono esserci due artisti per ogni film

3.3 Area Server

Concetto	Tipo	Volume	Note
Server	E	100	Si considera 1 server per ogni Area Geografica (gli stati più grossi avranno bisogno di più server)
Possesso	R	500.000	Deriva dal prodotto $100 * 5.000$, si considera che ogni server possieda in media 5.000 contenuti
Collegamento	R	180.000.000	Cardinalità (1, 1) con la tabella Connessione
Area Geografica	E	100	Ipotesi iniziale
Copertura	R	400	Si considera che ogni server raggiunge in media 4 Aree Geografiche
Restrizione	R	62.500	Deriva dal prodotto di $25 * 2.500$ (Si considera che la metà dei contenuti non è visualizzabile da 1/4 delle Aree geografiche)

3.4 Area Formati

Concetto	Tipo	Volume	Note
Formato Video	E	4	Ipotesi iniziale
Codifica Video	R	6.000	Cardinalità (1,1) con la tabella Contenuto
Formato Audio	E	3	Ipotesi iniziale
Codifica Audio	R	6.000	Cardinalità (1,1) con la tabella Contenuto

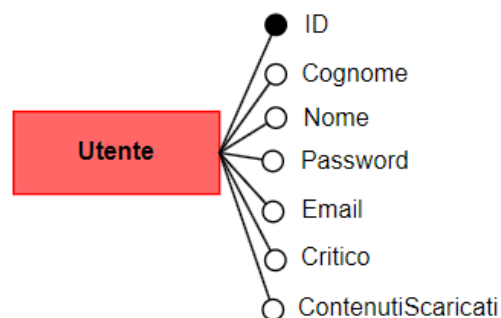
4 Operazioni sul database

4.1 Operazione 1 - Registrazione utente

L'operazione modifica la tabella Utente, inserendo una nuova riga con i dati passati come parametro.

INPUT: _id INT, _nome VARCHAR(50), _cognome VARCHAR(50), _passw VARCHAR(16), _email VARCHAR(100), _critico BOOL

OUTPUT: modifica tabella Utente



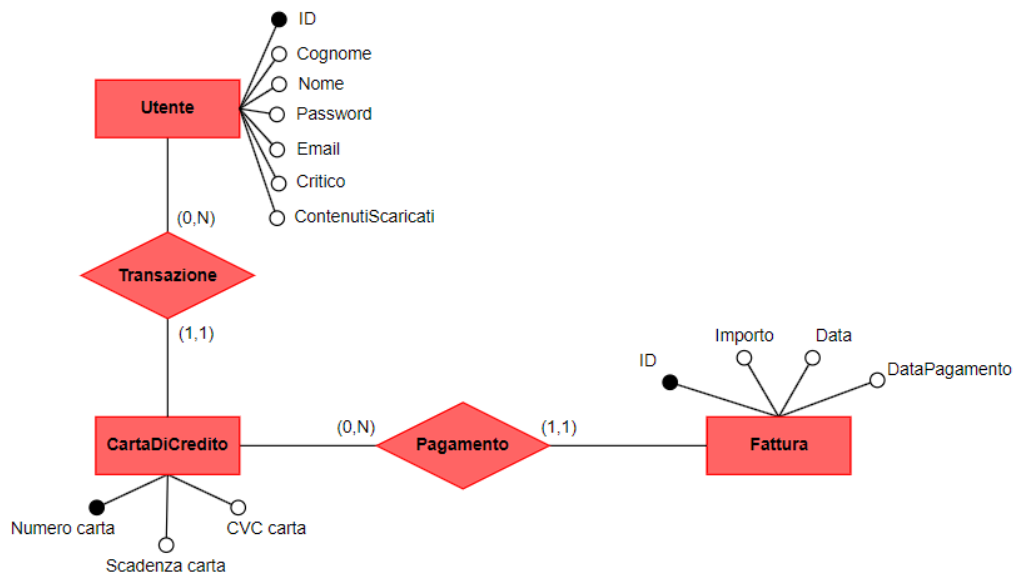
Concetto	Tipo	Accessi	L/S	Note
Utente	E	1.000.000	L	Cerco l'id più alto
Utente	E	1	S	Inserimento nuova riga

4.2 Operazione 2 - Nuova fattura

L'operazione verifica se la Carta di credito passata come parametro è presente nella tabella omonima, in caso contrario la inserisce. Dopodiché tramite una variabile di appoggio, calcola l'importo della fattura in base all'abbonamento richiesto come parametro, infine inserisce una nuova riga nella tabella Fattura.

INPUT: _idutente INT, _numerocarta VARCHAR(20), _scadenzacarta DATE, _cvc INT, _abbonamento VARCHAR(10)

OUTPUT: modifica tabelle CartaDiCredito e Fattura



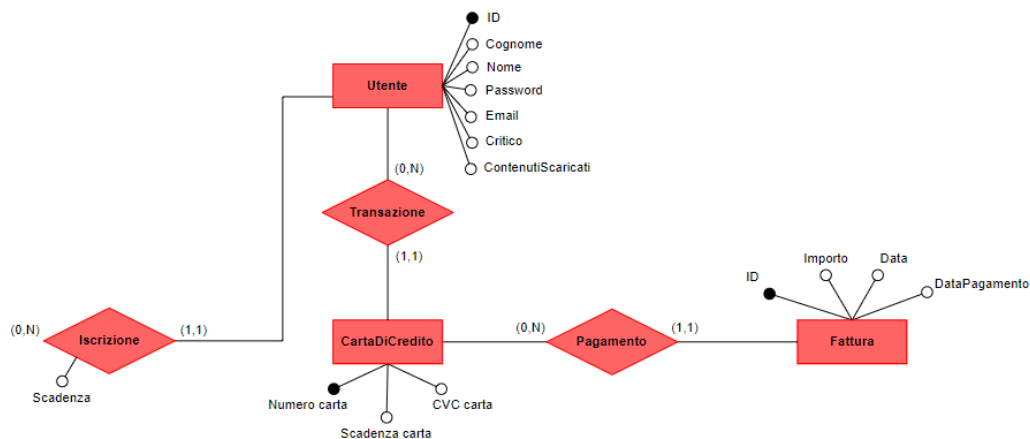
Concetto	Tipo	Accessi	L/S	Note
Fattura	E	12.000.000	L	Calcolo l'ID dell'ultima fattura
CartaDiCredito	E	1	L	Verifico se la carta è già registrata tramite la chiave
CartaDiCredito	E	1	S	Se non è registrata la inserisco in una nuova riga
Fattura	E	1	S	Inserisco la fattura in una nuova riga

4.3 Operazione 3a - Sottoscrizione abbonamento

L'operazione ricerca l'ultima fattura legata alla carta di credito passata come parametro, dopodichè controlla se la carta di credito è ancora in corso di validità, se non lo è restituisce un errore.

INPUT: _idutente INT, _numerocarta VARCHAR(20)

OUTPUT: modifica tabelle Fattura e Utente

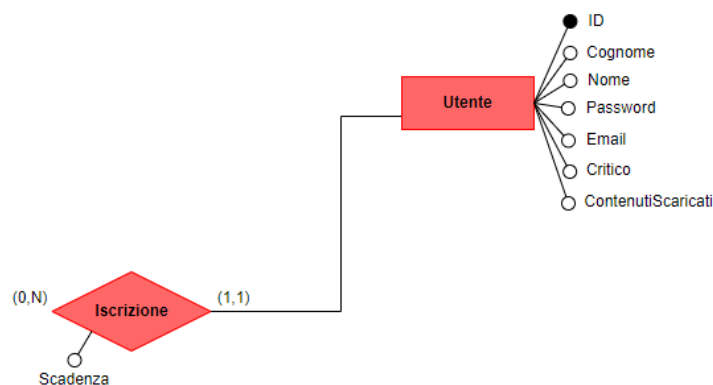


Concetto	Tipo	Accessi	L/S	Note
Fattura	E	6	L	In media ci sono 6 fatture per ogni carta di credito (12.000.000 / 2.000.000)
Fattura	E	6	L	Cerco l'ultima fattura di una specifica carta di credito
CartaDiCredito	E	1	L	Ricerca in CartaDiCredito tramite chiave
Fattura	E	1	L	Ricerco in Fattura tramite chiave
Fattura	E	1	S	Aggiornamento della data di pagamento
Utente	E	1	L	Ricerco in Utente tramite chiave
Utente	E	1	S	Aggiornamento dell'abbonamento e della data di scadenza

4.4 Operazione 3b - Scadenza abbonamento

L'operazione utilizza un Event quotidiano per controllare che l'abbonamento di un utente non sia scaduto, in tal caso annulla l'abbonamento.

OUTPUT: modifica tabella Utente



Concetto	Tipo	Accessi	L/S	Note
Utente	E	1.000.000	L	Si suppone che il 60% degli utenti iscritti possieda un abbonamento
Utente	E	20.000	S	Supponendo che le iscrizioni siano distribuite equamente nel corso del mese (30 giorni),

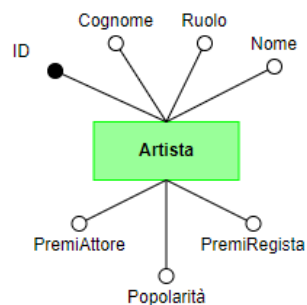
				ogni giorno scadranno 20.000 abbonamenti
--	--	--	--	--

4.5 Operazione 4 - Inserimento artista

L'operazione modifica la tabella Artista inserendo una nuova riga, con i dati passati per parametro, che può essere destinata ad un attore o ad un regista.

INPUT: _nome VARCHAR(255), _cognome VARCHAR(255), _premi INT, _ruolo VARCHAR(10)

OUTPUT: modifica tabella Artista



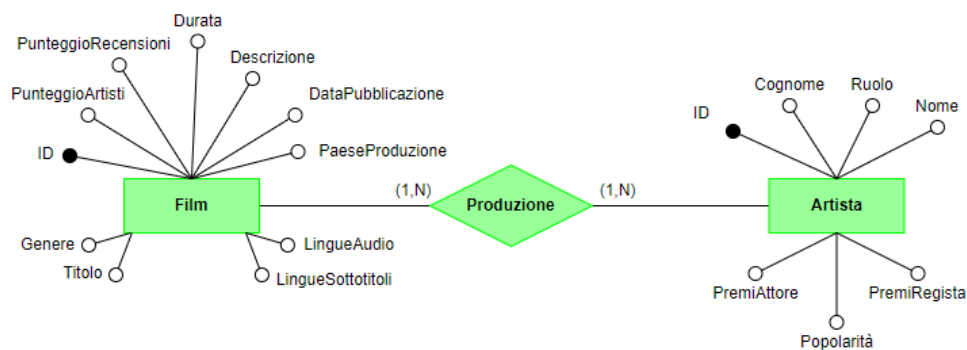
Concetto	Tipo	Accessi	L/S	Note
Artista	E	2.000	L	Ricerca dell'ID più alto
Artista	E	1	S	Inserimento di una nuova riga

4.6 Operazione 5 - Creazione film

L'operazione controlla che un artista sia già presente nel database, se non lo è da errore. In seguito aggiorna la popolarità dell'artista e inserisce una nuova riga in Film, per il nuovo film utilizzando i dati passati per parametro, e una nuova riga in Produzione.

INPUT: _titolo VARCHAR(255), _durata INT, _descrizione VARCHAR(255), _datapubb DATE, _paeseprod VARCHAR(50), _lingueaudio VARCHAR(100), _linguesott VARCHAR(100), _genere VARCHAR(50), _idatt INT, _idreg INT

OUTPUT: modifica tabelle Artista, Film e Produzione



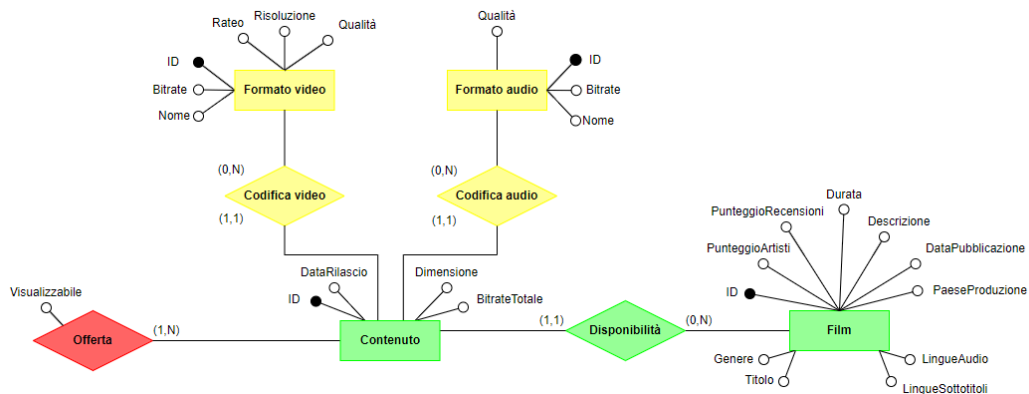
Concetto	Tipo	Accessi	L/S	Note
Film	E	1.000	L	Ricerca dell'id più alto
Artista	E	1	L	Cerco se l'attore esiste nella tabella Artista
Artista	E	1	L	Cerco se il regista esiste nella tabella Artista
Artista	E	2	L	Ricerca Artista tramite chiave
Artista	E	2	S	Aggiornamento della popolarità degli artisti
Artista	E	1	L	Calcolo popolarità attore
Artista	E	1	L	Calcolo popolarità regista
Film	E	1	S	Inserimento nuova riga
Produzione	R	2	S	Inserimento nuove righe

4.7 Operazione 6a - Inserimento contenuto

L'operazione calcola il bitrate totale del contenuto e aggiunge una nuova riga in Contenuto, utilizzando i dati passati per parametro, e una in Offerta. Il contenuto appena aggiunto fa parte delle "nuove uscite", non è quindi visualizzabile da utenti con abbonamento che non comprende questa caratteristica.

INPUT: _idfilm INT, _idformatoaudio INT, _idformatovideo INT

OUTPUT: modifica tabelle Contenuto e Offerta



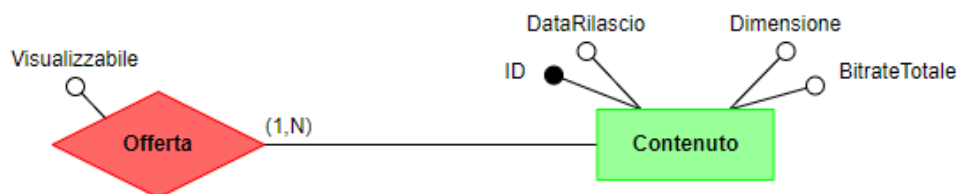
Concetto	Tipo	Accessi	L/S	Note
Film	E	1	L	Calcolo della durata del film
FormatoAudio	E	1	L	Calcolo del bitrate del formato audio
FormatoVideo	E	1	L	Calcolo del bitrate del formato video
Contenuto	E	6.000	L	Ricerca dell'id più alto
Contenuto	E	1	S	Inserimento nuova riga
Offerta	R	5	S	Inserimento nuove righe

4.8 Operazione 6b - Modifica offerta vecchi contenuti

L'operazione utilizza un Event quotidiano per aggiornare la tabella Offerta aggiungendo la possibilità, a chi ha un abbonamento basic, premium o pro, di vedere i contenuti che non sono più "nuove uscite".

INPUT: nessuno

OUTPUT: modifica tabella Offerta



Concetto	Tipo	Accessi	L/S	Note
Contenuto	E	6.000	L	Cerco i contenuti usciti da un mese esatto

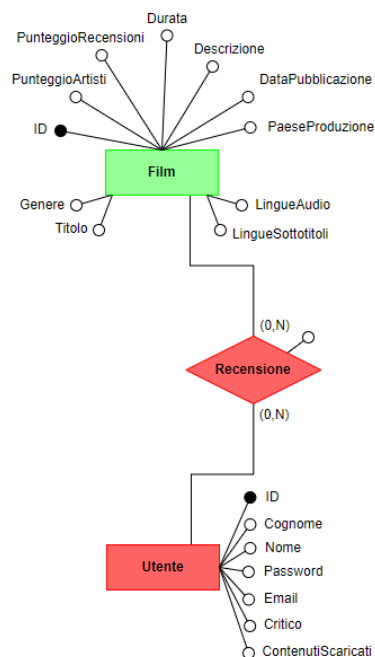
Offerta	R	9.000	L	Cerco le righe da aggiornare
Offerta	R	5	S	Supponendo che al mese escano 10 film, quindi circa 60 contenuti (in media 2 al giorno) e dato che le righe interessate sono quelle degli abbonamenti “basic”, “premium” e “pro”, calcoliamo una stima di aggiornamenti giornalieri pari a 5 righe (1 per basic, 2 per premium e pro)

4.9 Operazione 7 - Inserimento recensione

L'operazione controlla che l'utente, passato per parametro, non abbia già recensito il film, passato per parametro, in tal caso aggiorna la vecchia recensione. In seguito aggiorna il punteggio del film, considerando se l'utente è un critico o no.

INPUT: _idutente INT, _idfilm INT, _punteggio INT

OUTPUT: modifica tabelle Recensione e Film



Concetto	Tip o	Accessi	L/S	Note
Recensione	R	1	L	Controllo se l'utente ha già recensito quel film
Recensione	R	1	L	Cerco la recensione da eliminare

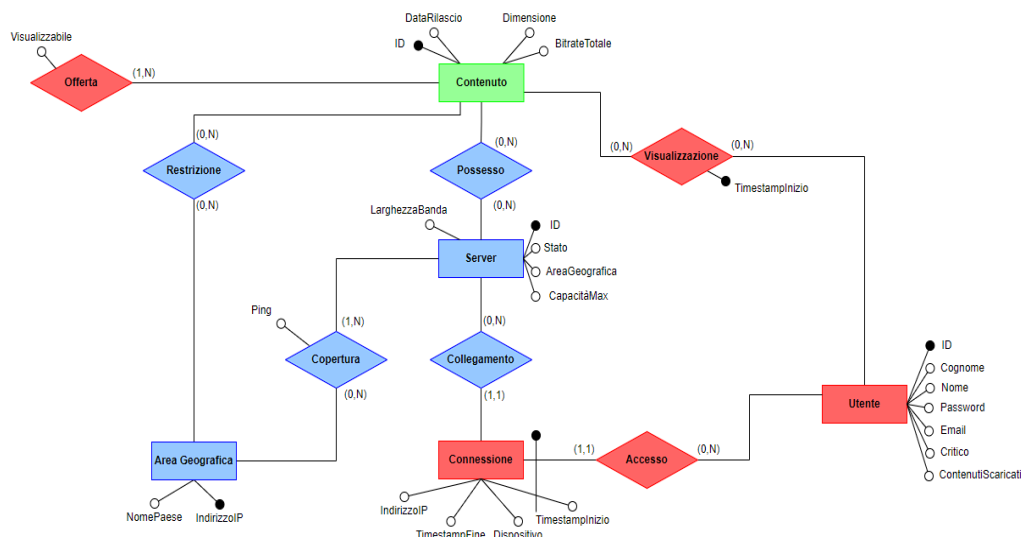
Recensione	R	1	S	Elimino la vecchia recensione (se presente)
Recensione	R	1	S	Inserisco la nuova recensione
Recensione	R	50.000	L	Si suppone che in media ogni film abbia 50 recensioni
Utente	E	1.000.000	L	Si suppone che il 20% degli utenti siano critici
Recensione	R	50.000	L	Si suppone che in media ogni film abbia 50 recensioni
Utente	E	1.000.000	L	Si suppone che l'80% degli utenti siano utenti comuni
Film	E	1	L	Ricerca del film tramite chiave
Film	E	1	S	Aggiornamento del punteggio del film

4.10 Operazione 8a - Visualizzazione contenuto

L'operazione controlla che l'utente, passato per parametro, possa visualizzare il contenuto, e inserisce una nuova riga in connessione e visualizzazione, dopo aver scelto il miglior server disponibile a cui collegarsi, aggiornando il suo stato.

INPUT: _idutente INT, _idcontenuto INT, _indirizzoip INT, _dispositivo VARCHAR(50)

OUTPUT: modifica tabelle Connessione, Visualizzazione e Server



Concetto	Tipo	Accessi	L/S	Note
Restrizione	R	1	L	Controllo se il contenuto è

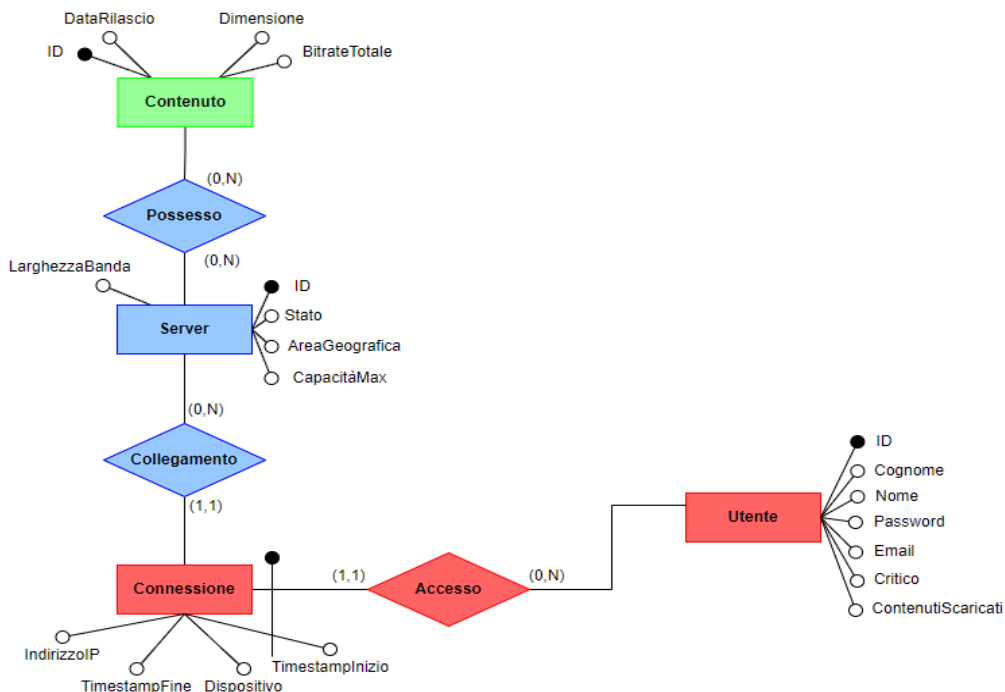
				visualizzabile in una data area geografica
Utente	E	1	L	Calcolo l'abbonamento dell'utente
Offerta	R	1	L	Controllo se il contenuto è compreso nell'abbonamento calcolato
Contenuto	E	1	L	Calcolo il bitrate necessario a visualizzare il contenuto
Server	E	100	L	Cerco i server non pieni
Possesso	R	500.000	L	Supponiamo che 30 server abbiano lo stesso contenuto (5 per continente, 6 continenti)
Copertura	R	400	L	Sui 30 server che possiedono il contenuto, in media 5 si trovano nel continente giusto
Connessione	E	1	S	Inserimento nuova riga
Visualizzazione	R	1	S	Inserimento nuova riga
Server	E	1	L	Ricerco il server destinato all'erogazione del contenuto
Server	E	1	S	Aggiorno lo stato del server target

4.11 Operazione 8b - Fine visualizzazione

L'operazione trova l'ultima connessione effettuata dall'utente, cioè quella ancora in corso, passato per parametro, e aggiorna il timestamp di fine connessione di quell'utente. In seguito aggiorna lo stato del server al quale era connesso.

INPUT: _idutente INT, _idcontenuto INT, _indirizzoip INT

OUTPUT: modifica tabelle Connessione e Server

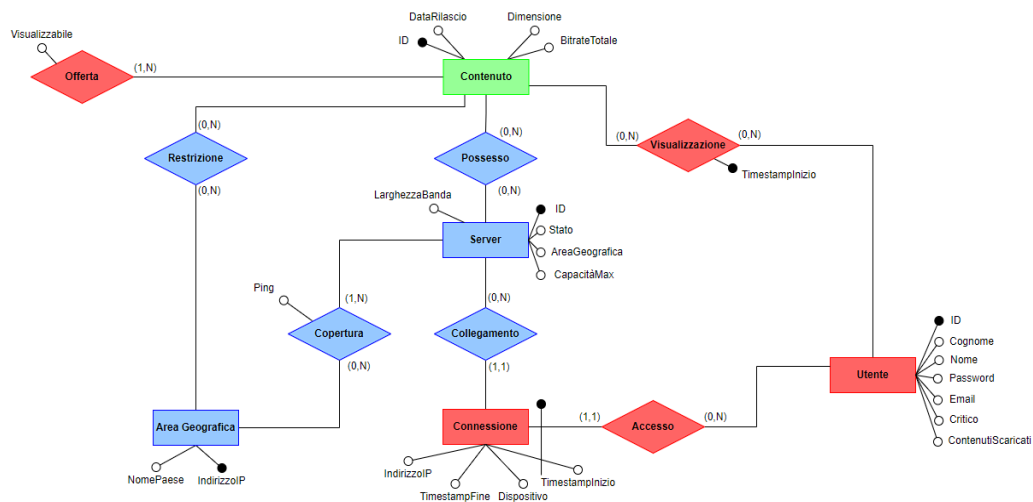


Concetto	Tipo	Accessi	L/S	Note
Connessione	E	180	L	Ricerco l'ultima connessione dell'utente utilizzando parte della chiave (ID_Utente)
Connessione	E	1	L	Cerco il server con cui è avvenuta l'ultima connessione
Connessione	E	1	L	Ricerca in Connessione tramite chiave
Connessione	E	1	S	Aggiorno il timestamp di fine connessione
Contenuto	E	1	L	Calcolo il bitrate del contenuto target tramite chiave
Server	E	1	L	Ricerca in Server tramite chiave
Server	E	1	S	Aggiorno lo stato del server a cui l'utente era connesso

4.12 Operazione 8c - Sposta utenti su altri server

L'operazione sposta un utente dal server al quale è collegato, ad un altro più conveniente, se esiste, quando si libera spazio in esso.

OUTPUT: modifica tabelle Server, Connessione e Visualizzazione



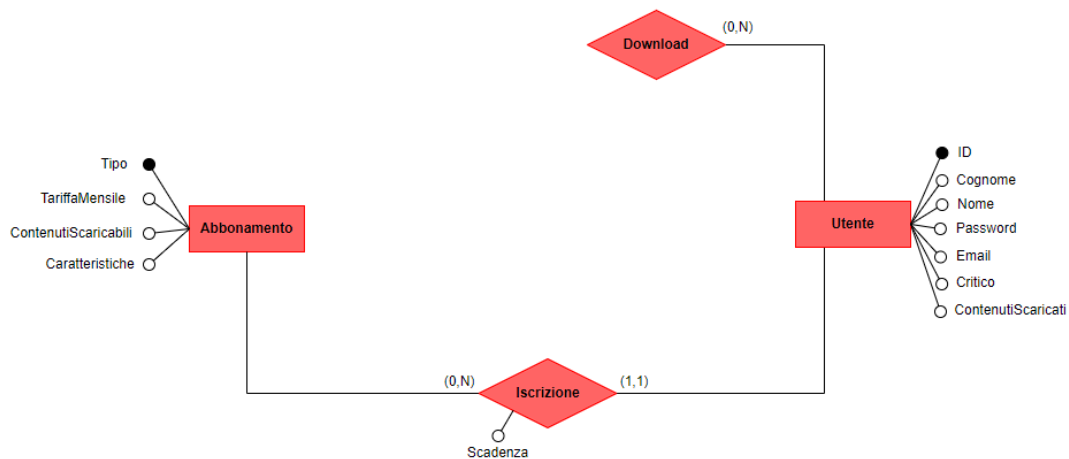
Concetto	Tipo	Accessi	L/S	Note
Connessione	E	180.000.000	L	Natural join con Visualizzazione (su TimestampInizio e ID_Utente)
Visualizzazione	R	180.000.000	L	Natural join con Connessione (su TimestampInizio e ID_Utente)
Copertura	R	300	L	Inner join con la tabella ottenuta sopra sulla chiave di Copertura. I valori di Ping nella tabella sono 20, 40, 60, 80, perciò ¾ delle righe hanno Ping > 20
Contenuto	E	1	L	Ricerca del bitrate del contenuto target
Server	E	1	L	Accesso in Server tramite chiave
Server	E	1	S	Aggiorno lo stato di trasmissione del server
Connessione	E	1	L	Accesso in Connessione tramite chiave
Connessione	E	1	S	Aggiorno il TimestampFine della connessione
Chiamata alla funzione Visualizzazione contenuto, vedere paragrafo 4.10				

4.13 Operazione 9a - Download di contenuti

L'operazione controlla che l'utente, passato per parametro, possa scaricare il contenuto, passato per parametro. Se può, inserisce una nuova riga in Download.

INPUT: _idutente INT, _idcontenuto INT

OUTPUT: modifica tabella Download



Concetto	Tipo	Accessi	L/S	Note
Download	R	1	L	Controllo che l'utente non abbia già scaricato quel contenuto
Utente	E	1	L	Controllo l'abbonamento dell'utente
Abbonamento	E	1	L	Controllo se l'abbonamento comprende la possibilità di scaricare contenuti offline
Download	R	5.000.000	L	Conto quanti download ha già fatto l'utente, in media 5 download per utente secondo le supposizioni nella tavola dei volumi
Download	R	1	S	Inserimento nuova riga

4.14 Operazione 9b - Elimina contenuti scaricati

L'operazione elimina una riga da download con i dati corrispondenti a quelli passati per parametro.

INPUT: _idutente INT, _idcontenuto INT

OUTPUT: modifica tabella Download

Download

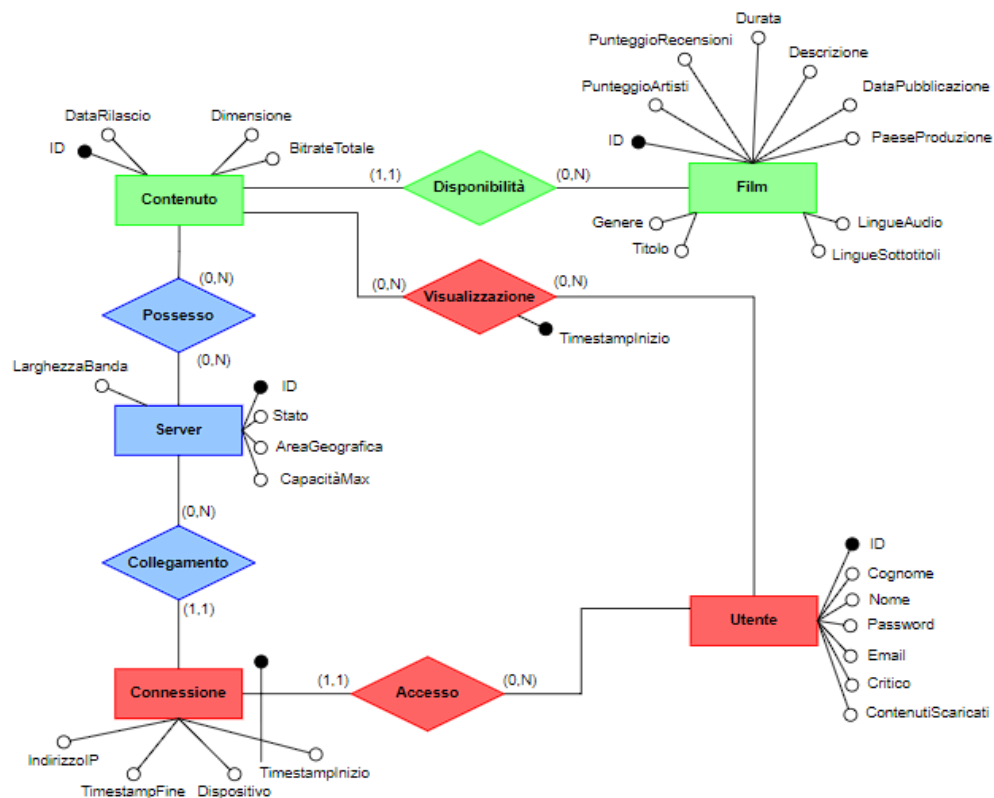
Concetto	Tipo	Accessi	L/S	Note
Download	R	1	L	Ricerca in Download tramite chiave
Download	R	1	S	Elimino una riga

4.15 Operazione 10 - Caching di contenuti

L'operazione calcola l'ultimo server al quale l'utente, passato per parametro, si è connesso, in seguito calcola il suo genere preferito. Da questo dato sceglie il contenuto che l'utente vorrà vedere con più probabilità, e lo sposta nel server più opportuno.

INPUT: _idutente INT

OUTPUT: modifica tabella Possesso



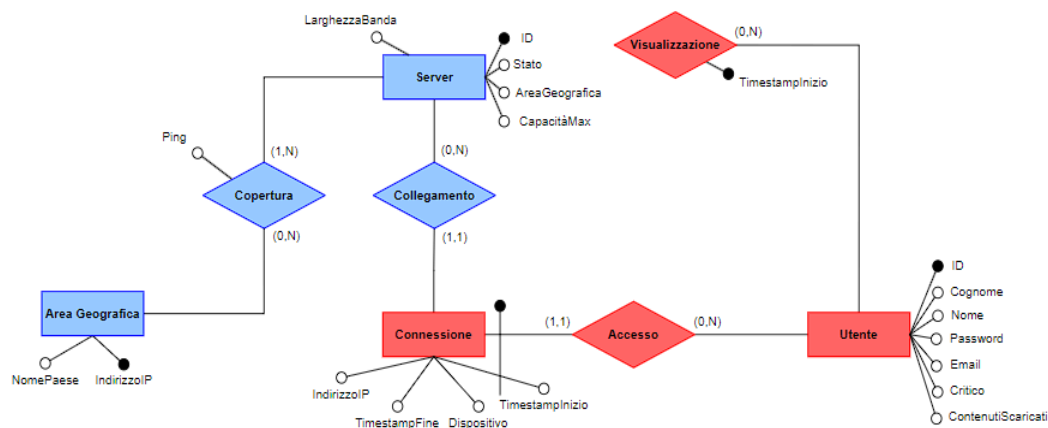
Concetto	Tipo	Accessi	L/S	Note
Connessione	E	180	L	Ricerco le connessioni dell'utente in esame
Visualizzazione	R	180.000.000	L	Inner join con Contenuto sulla chiave di Contenuto
Contenuto	E	6.000	L	Inner join con Visualizzazione sulla chiave di Contenuto
Film	E	1.000	L	Inner join con la tabella sopra, sulla chiave di Film
Utente	E	1	L	Ricerco l'abbonamento dell'utente tramite chiave
Visualizzazione	R	180.000.000	L	Inner join con Contenuto sulla chiave di Contenuto
Contenuto	E	6.000	L	Inner join con Visualizzazione sulla chiave di Contenuto
Film	E	1.000	L	Inner join con la tabella sopra, sulla chiave di Film, supponendo che ci siano 10 generi distinti di Film
Film	E	1.000	L	Cerco gli ID dei film con un genere specifico, che non siano già stati visti dall'utente
Contenuto	E	6.000	L	Supponendo che i contenuti riguardanti lo stesso film abbiano sempre id consecutivi e siano sempre 6, cerco l'id del primo contenuto
Possesso	R	1	S	Inserisco un nuovo contenuto in un server target

5 Area analytics

5.1 Classifiche

Vengono stilate delle classifiche dei contenuti più visti e dei loro formati, a seconda della zona geografica, e della modalità di abbonamento.

OUTPUT: due tabelle contenenti le classifiche



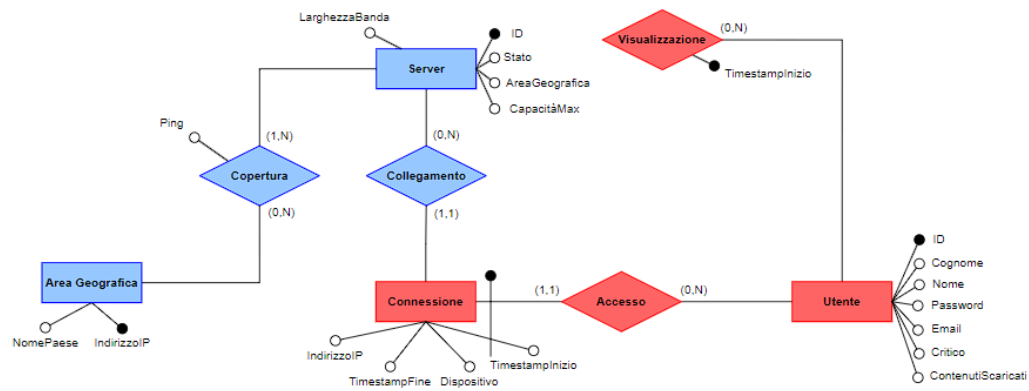
Concetto	Tipo	Accessi	L/S	Note
Connessione	E	180.000.000	L	Natural join con Visualizzazione
Visualizzazione	R	180.000.000	L	Natural join con Connessione
Area Geografica	E	100	L	Inner join Connessione
Visualizzazione	R	180.000.000	L	Inner join con Utente
Utente	E	1.000.000	L	Inner join con Visualizzazione

5.2 Bilanciamento del carico

Si stimano i potenziali sovraccarichi di un server, facendo un' analisi della CDN e dell'uso fatto dai clienti dei film memorizzati in essa. Vengono suggeriti degli spostamenti di contenuti da un server per non sovraccaricarlo indicando un server destinatario. Il parametro N indica quanti contenuti devono essere spostati per ogni paese, MaxPing fa da limite superiore al ping che i server destinatari devono avere in relazione al paese in esame.

INPUT: N INT, MaxPing INT

OUTPUT: una tabella contenente i suggerimenti



Concetto	Tipo	Accessi	L/S	Note
Connessione	E	180.000.000	L	Natural join con Visualizzazione
Visualizzazione	R	180.000.000	L	Natural join con Connessione
Area Geografica	E	100	L	Inner join Connessione
Copertura	R	400	L	Inner join con tabella derivata

6 Progettazione logica

6.1 Schema logico

6.1.1 Area Clienti

Concetti ER	Traduzione
Utente, Iscrizione	Utente (<u>ID</u> , Nome, Cognome, Passw, Email, Critico, NomeAbbonamento, Scadenza*)
CartaDiCredito, MetodoDiPagamento	CartaDiCredito (<u>NumeroCarta</u> , ID_Utente, ScadenzaCarta, CVC)
Fattura, Pagamento	Fattura (<u>ID</u> , Data, DataPagamento, NumeroCarta, Importo)
Connessione, Accesso, Collegamento	Connessione (<u>ID_Utente</u> , <u>TimestampInizio</u> , ID_Server, IndirizzoIP, Dispositivo, TimestampFine)
Abbonamento	Abbonamento (<u>Nome</u> , TariffaMensile, ContenutiScaricabili, Caratteristiche)
Offerta	Offerta (<u>NomeAbbonamento</u> , <u>ID_Contenuto</u> , Visualizzabile)
Visualizzazione	Visualizzazione (<u>ID_Utente</u> , <u>ID_Contenuto</u> , <u>TimestampInizio</u>)
Recensione	Recensione (<u>ID_Utente</u> , <u>ID_Film</u> , PunteggioUtente)
Download	Download (<u>ID_Utente</u> , <u>ID_Contenuto</u>)

*L'attributo Scadenza è stato incorporato in Utente

6.1.2 Area Contenuti

Concetti ER	Traduzione
Contenuto, Disponibilità, CodificaVideo, CodificaAudio	Contenuto(<u>ID</u> , ID_Film, DataRilascio, Dimensione, BitrateTotale)
Film	Film(<u>ID</u> , Descrizione, Durata, Genere, Titolo, DataPubblicazione, PaeseDescrizione, LingueAudio, LingueSottotitoli, PunteggioRecensioni, PunteggioArtisti)
Produzione	Produzione(<u>ID_Film</u> , <u>ID_Artista</u>)
Artista	Artista(<u>ID</u> , Nome, Cognome, Ruolo, Popolarità, PremiAttore, PremiRegista)

6.1.3 Area Server

Concetti ER	Traduzione
Server	Server(<u>ID</u> , Stato, AreaGeografica, CapacitàMax, LarghezzaBanda)
Possesso	Possesso(<u>ID_Contenuto</u> , <u>ID_Server</u>)
AreaGeografica	AreaGeografica(<u>IndirizzoIP</u> , NomePaese)
Copertura	Copertura(<u>ID_Server</u> , <u>IndirizzoIP</u> , Ping)
Restrizione	Restrizione(<u>ID_Contenuto</u> , <u>IndirizzoIP</u>)

6.1.4 Area Formati

Concetti ER	Traduzione
FormatoVideo	FormatoVideo(<u>ID</u> , Nome, Bitrate, Rateo, Risoluzione, Qualità)
FormatoAudio	FormatoAudio(<u>ID</u> , Nome, Bitrate, Qualità)

6.2 Vincoli di integrità referenziale

Connessione:

- ID_Utente in ID della tabella Utente
- ID_Server in ID della tabella Server
- IndirizzoIP in IndirizzoIP della tabella AreaGeografica

Contenuto:

- ID_Film in ID della tabella Film
- ID_FormatoAudio in ID della tabella FormatoAudio
- ID_FormatoVideo in ID della tabella FormatoVideo

Copertura:

- IndirizzoIP in IndirizzoIP della tabella AreaGeografica
- ID_Server in ID della tabella Server

Download:

- ID_Utente in ID della tabella Utente
- ID_Contenuto in ID della tabella Contenuto

Fattura:

- NumeroCarta in NumeroCarta della tabella CartadiCredito

Offerta:

- ID_Contenuto in ID della tabella Contenuto
- NomeAbbonamento in Nome della tabella Abbonamento

Possesso:

- ID_Server in ID della tabella Server
- ID_Contenuto in ID della tabella Contenuto

Produzione:

- ID_Film in ID della tabella Film
- ID_Artista in ID della tabella Artista

Recensione:

- ID_Film in ID della tabella Film
- ID_Utente in ID della tabella Utente

Restrizione:

- IP_AreaGeografica in IndirizzoIP della tabella AreaGeografica
- ID_Contenuto in ID della tabella Contenuto

Visualizzazione:

- ID_Utente in ID della tabella Utente
- ID_Contenuto in ID della tabella Contenuto

6.3 Vincoli di tupla

Ogni attributo che fa parte di una chiave di una tabella non deve mai assumere il valore NULL.

Nella tabella Connessione TimestampFine non deve mai precedere TimestampInizio.

Nella tabella Fattura DataPagamento non deve precedere Data.

Nella tabella Artista, le tuple con valore “attore” in ruolo deve avere valore NULL in premiRegista e le tuple con valore “regista” in ruolo deve avere valore NULL in premiAttore

6.4 Vincoli di dominio

Ogni attributo di tipo ID, così come ogni altro attributo numerico non deve avere valore negativo.

L'attributo CVC in CartadiCredito deve avere un valore compreso tra 100 e 999.

L'attributo DataPubblicazione di Film deve precedere CURRENT_DATE, e ogni Timestamp non nullo deve precedere CURRENT_TIMESTAMP.

L'attributo PunteggioUtente di Recensione deve avere un valore compreso tra 0 e 50.

L'attributo Popolarità di Artista è calcolato considerando i premi vinti dall'artista e il numero di film nella piattaforma in cui compare, ed ha un valore compreso tra 0 e 50.

Gli attributi PunteggioRecensioni e PunteggioArtisti hanno entrambi un valore compreso tra 0 e 50.

L'attributo Passw di Utente deve avere lunghezza pari o maggiore a 8.

L'attributo Email di Utente deve contenere la @ e il . per essere valida.

6.5 Vincoli inter-relazionali

Nella tabella Contenuto, il valore dell'attributo BitrateTotale deve essere uguale alla somma dei bitrate dei formati associati a quel contenuto, inoltre DataRilascio dev'essere successiva alla DataPubblicazione del film.

TimestampInizio della tabella Visualizzazione deve essere uguale al TimestampInizio della Connessione corrispondente.

DataPagamento della tabella Fattura deve essere precedente alla data di scadenza della carta con cui si sta pagando.

6.6 Dipendenze funzionali e Normalizzazione

6.6.1 Area Clienti

Utente (ID, Nome, Cognome, Passw, Email, Critico, NomeAbbonamento, Scadenza)

- ID → Nome, Cognome, Passw, Email, Critico, NomeAbbonamento, Scadenza

Si suppone che possano esistere due o più utenti con la stessa Email, altrimenti esisterebbe la DF: Email → Nome, Cognome, Passw, Critico, NomeAbbonamento, Scadenza.

→ La tabella è in forma BCNF

Pagamento (ID, DataFattura, DataPagamento, Importo, NumeroCarta, ScadenzaCarta, CVC)

In questa tabella è presente la dipendenza funzionale:

NumeroCarta → ScadenzaCarta, CVC, pertanto non è rispettata la BCNF (in quanto NumeroCarta non fa parte della chiave). Nella ristrutturazione la tabella è stata decomposta in CartaDiCredito e Fattura:

CartaDiCredito (NumeroCarta, ID_Utente, ScadenzaCarta, CVC)

- NumeroCarta → ID_Utente, ScadenzaCarta, CVC

→ La tabella è in forma BCNF

Fattura (ID, Data, DataPagamento, NumeroCarta, Importo)

- ID → Data, DataPagamento, NumeroCarta, Importo

→ La tabella è in forma BCNF

Connessione (ID_Utente, TimestampInizio, ID_Server, IndirizzoIP, Dispositivo, TimestampFine)

- ID_Utente, TimestampInizio → ID_Server, IndirizzoIP, Dispositivo, TimestampFine

→ La tabella è in forma BCNF

Abbonamento (Nome, TariffaMensile, ContenutiScaricabili, Caratteristiche)

- Nome → TariffaMensile, ContenutiScaricabili, Caratteristiche

→ La tabella è in forma BCNF

Offerta (NomeAbbonamento, ID_Contenuto, Visualizzabile)

- NomeAbbonamento → ID_Contenuto, Visualizzabile

→ La tabella è in forma BCNF

Visualizzazione (ID Utente, ID Contenuto, TimestampInizio)

→ La tabella è in forma BCNF

Recensione (ID Utente, ID Film, PunteggioUtente)

- ID Utente, ID Film → PunteggioUtente

→ La tabella è in forma BCNF

Download (ID Utente, ID Contenuto)

→ La tabella è in forma BCNF

6.6.2 Area Contenuti

Contenuto (ID, ID_Film, DataRilascio, Dimensione, BitrateTotale)

- ID → ID_Film, DataRilascio, Dimensione, BitrateTotale

→ La tabella è in forma BCNF

Film (ID, Descrizione, Durata, Genere, Titolo, DataPubblicazione, PaeseDescrizione, LingueAudio, LingueSottotitoli, PunteggioRecensioni, PunteggioArtisti)

- ID → Descrizione, Durata, Genere, Titolo, DataPubblicazione, PaeseDescrizione, LingueAudio, LingueSottotitoli, PunteggioRecensioni, PunteggioArtisti

Si considera che potrebbero esistere film con la stessa descrizione dato che questa è breve e generica, altrimenti esisterebbe la dipendenza funzionale:

Descrizione → ID, Durata, Genere, Titolo, DataPubblicazione, PaeseDescrizione, LingueAudio, LingueSottotitoli, PunteggioRecensioni, PunteggioArtisti

→ La tabella è in forma BCNF

Produzione (ID Film, ID Artista)

→ La tabella è in forma BCNF

Artista (ID, Nome, Cognome, Ruolo, Popolarità, PremiAttore, PremiRegista)

→ La tabella è in forma BCNF

6.6.3 Area Server

Server (ID, Stato, AreaGeografica, CapacitàMax, LarghezzaBanda, Ping)

- ID → Stato, AreaGeografica, CapacitàMax, LarghezzaBanda, Ping

→ La tabella è in forma BCNF

Possesso (ID Contenuto, ID Server)

→ La tabella è in forma BCNF

Area Geografica (IndirizzoIP, NomePaese)

- IndirizzoIP → NomePaese

→ La tabella è in forma BCNF

Copertura (ID_Server, IndirizzoIP, Ping)

- ID_Server, Indirizzo_IP → Ping

→ La tabella è in forma BCNF

Restrizione (ID_Contenuto, IndirizzoIP)

→ La tabella è in forma BCNF

6.6.4 Area Formati

Formato Video (ID, Nome, Bitrate, Rateo, Risoluzione, Qualità)

- ID → Nome, Bitrate, Rateo, Risoluzione, Qualità

→ La tabella è in forma BCNF

Formato Audio (ID, Nome, Bitrate, Qualità)

- ID → Nome, Bitrate, Qualità

→ La tabella è in forma BCNF

7 Implementazione

Il database è stato implementato utilizzando il software MySQL, per ogni tabella è stato scelto il motore di memorizzazione InnoDB, in quanto garantisce accesso mediante BTREE di default sugli indici generati dalle chiavi primarie, questo ottimizza la performance dei JOIN. Inoltre ciò permette di accedere velocemente alle tabelle sia per chiave primaria che per chiave esterna.

7.1 Trigger e Events

Tutti i vincoli, compresi quelli di integrità referenziale sono stati implementati tramite Trigger o Event.

Di seguito una lista in ordine di tabella alfabetico con una breve descrizione per ognuno.

Tabella	Implementato con	Descrizione
Abbonamento	Trigger before insert	Controlla che i dati rispettino i vincoli di Abbonamento
Area Geografica	Trigger before insert	Controlla che i dati rispettino i vincoli di Area Geografica
Artista	Trigger before insert	Controlla che i dati rispettino i vincoli di Artista
ArtistaUpdate	Trigger after update	Controlla che la popolarità di un artista non superi il valore 50
Carta di Credito	Trigger before insert	Controlla che i dati rispettino i vincoli di Carta di Credito
Connessione	Trigger before insert	Controlla che i dati rispettino i vincoli di Connessione
Contenuto	Trigger before insert	Controlla che i dati rispettino i vincoli di Contenuto
Copertura	Trigger before insert	Controlla che i dati rispettino i vincoli di Copertura
Download	Trigger before insert	Controlla che i dati rispettino i vincoli di

		Download
Fattura	Trigger before insert	Controlla che i dati rispettino i vincoli di Fattura
Film	Trigger before insert	Controlla che i dati rispettino i vincoli di Film
FormatoAudio	Trigger before insert	Controlla che i dati rispettino i vincoli di FormatoAudio
FormatoVideo	Trigger before insert	Controlla che i dati rispettino i vincoli di FormatoVideo
Offerta	Trigger before insert	Controlla che i dati rispettino i vincoli di Offerta
Possesso	Trigger before insert	Controlla che i dati rispettino i vincoli di Possesso
Produzione	Trigger before insert	Controlla che i dati rispettino i vincoli di Produzione
Recensione	Trigger before insert	Controlla che i dati rispettino i vincoli di Recensione
Restrizione	Trigger before insert	Controlla che i dati rispettino i vincoli di Restrizione
Server	Trigger before insert	Controlla che i dati rispettino i vincoli di Server
Utente	Trigger before insert	Controlla che i dati rispettino i vincoli di Utente
Visualizzazione	Trigger before insert	Controlla che i dati rispettino i vincoli di Visualizzazione
Offerta	Event	Vedi paragrafo 4.8
Utente	Event	Vedi paragrafo 4.4

7.2 Popolamento

Il popolamento è stato portato avanti per la maggior parte grazie all'utilizzo dell'intelligenza artificiale ChatGPT, con la quale sono stati creati elenchi di informazioni aleatorie, come quelle presenti nelle tabelle Utente, CartaDiCredito,

Connessione, Visualizzazione ed elenchi di informazioni reali reperite online, come quelle nelle tabelle Artista e Film.

Inoltre ChatGPT è stato utilizzato anche per popolare tabelle composte da associazioni numeriche e testuali che rispettano dei pattern nei valori, come le tabelle Offerta, Contenuto, Possesso e Recensione.

Per quanto riguarda tabelle contenenti dettagli tecnici, come le tabelle Server, FormatoAudio e FormatoVideo, sono state svolte ricerche per selezionare dati realistici e opportuni ai fini del progetto. La tabella Produzione è stata popolata manualmente ricercando regista e attore principale per ogni film.

Alle tabelle costruite con ChatGPT sono state apportate modifiche coerenti con i vincoli precedentemente descritti, in particolare sono state rielaborate le seguenti tabelle:

- Connessione e Visualizzazione, in modo da far corrispondere ogni connessione ad una visualizzazione (modifica su TimestampInizio).
- Fattura e CartaDiCredito, in modo da avere valori corrispondenti di NumeroCarta.
- Fattura e Utente, così da associare ad ogni utente l'abbonamento corretto per l'importo pagato in fattura.