Immagine che contiene Carattere, cerchio, Elementi grafici, logo

Descrizione generata automaticamente

Simone Calò – Nicola Guerrieri

**Documentazione del Progetto di Basi di Dati**

Anno Accademico 2022-23

**SOMMARIO**

[GLOSSARIO DEI TERMINI 3](#_Toc140142076)

[DESCRIZIONE DIAGRAMMA-ER 4](#_Toc140142077)

* [AREA CONTENUTI 4](#_Toc140142078)
* [AREA FORMATI 5](#_Toc140142079)
* [AREA CLIENTI 6](#_Toc140142080)
* [AREA STREAMING 7](#_Toc140142081)

[RISTRUTTURAZIONE DIAGRAMMA-ER 8](#_Toc140142082)

[TAVOLA DEI VOLUMI 12](#_Toc140142083)

[OPERAZIONI SUI DATI 15](#_Toc140142084)

* [OPERAZIONE 1 15](#_Toc140142085)
* [OPERAZIONE 2 17](#_Toc140142086)
* [OPERAZIONE 3 18](#_Toc140142087)
* [OPERAZIONE 4 19](#_Toc140142088)
* [OPERAZIONE 5 20](#_Toc140142089)
* [OPERAZIONE 6 21](#_Toc140142090)
* [OPERAZIONE 7 23](#_Toc140142091)
* [OPERAZIONE 8 24](#_Toc140142092)

[PROGETTAZIONE LOGICA 25](#_Toc140142093)

* [AREA CONTENUTI 25](#_Toc140142094)
* [AREA FORMATI 25](#_Toc140142095)
* [AREA CLIENTI 26](#_Toc140142096)
* [AREA STREAMING 26](#_Toc140142097)
* [VINCOLI DI INTEGRITA’ GENERICI 27](#_Toc140142098)

[ANALISI DIPENDENZE FUNZIONALI E NORMALIZZAZIONE 29](#_Toc140142099)

* [AREA CONTENUTI 29](#_Toc140142100)
* [AREA FORMATI 30](#_Toc140142101)
* [AREA CLIENTI 30](#_Toc140142102)
* [AREA STREAMING 31](#_Toc140142103)

[FUNZIONALITÀ BACKEND 32](#_Toc140142104)

[DATA ANALYTICS 34](#_Toc140142105)

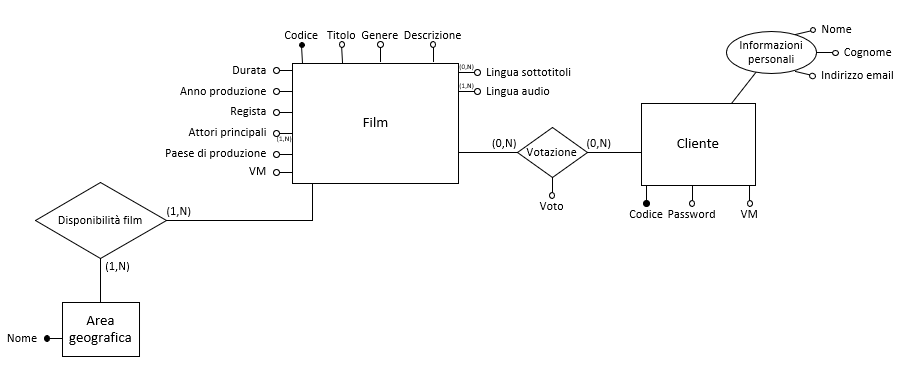
* [CLASSIFICHE 34](#_Toc140142106)
* [BILANCIAMENTO DEL CARICO 34](#_Toc140142107)

GLOSSARIO DEI TERMINI

|  |  |  |  |
| --- | --- | --- | --- |
| **Termine** | **Descrizione** | **Sinonimi** | **Collegamenti** |
| Film | Film disponibile su FilmSphere | Video online | Rating, Formato, Area geografica |
| Rating | Rating di un film | Valutazione | Film, Cliente |
| Formato | Formato di un film |  | Film, Area geografica |
| Area geografica | Area geografica di interesse per FilmSphere | Restrizione, suddivisione geografica | Film, Formato, Connessione, Server |
| Cliente | Persona iscritta a FilmSphere | Utente | Film, Rating, Visualizzazione |
| Connessione | Connessione a FilmSphere da parte di un certo cliente |  | Cliente, Area geografica, Visualizzazione |
| Visualizzazione | Visualizzazione di un film da parte di un certo cliente | Visione | Film, Cliente, Connessione |
| Piano di abbonamento | Piano di abbonamento con certe funzionalità | Modalità di abbonamento | Abbonamento |
| Abbonamento | Abbonamento di un certo cliente |  | Cliente, Piano di abbonamento, Fattura |
| Fattura | Documento relativo al pagamento di un certo abbonamento |  | Cliente, Abbonamento |
| Server | Server della CDN in una certa area geografica |  | Film, Cliente, Area geografica, Visualizzazione |
| Cache | Memorizzazione di un certo film in un certo server |  | Film, Server, Visualizzazione |

DESCRIZIONE DIAGRAMMA-ER

* AREA CONTENUTI

****

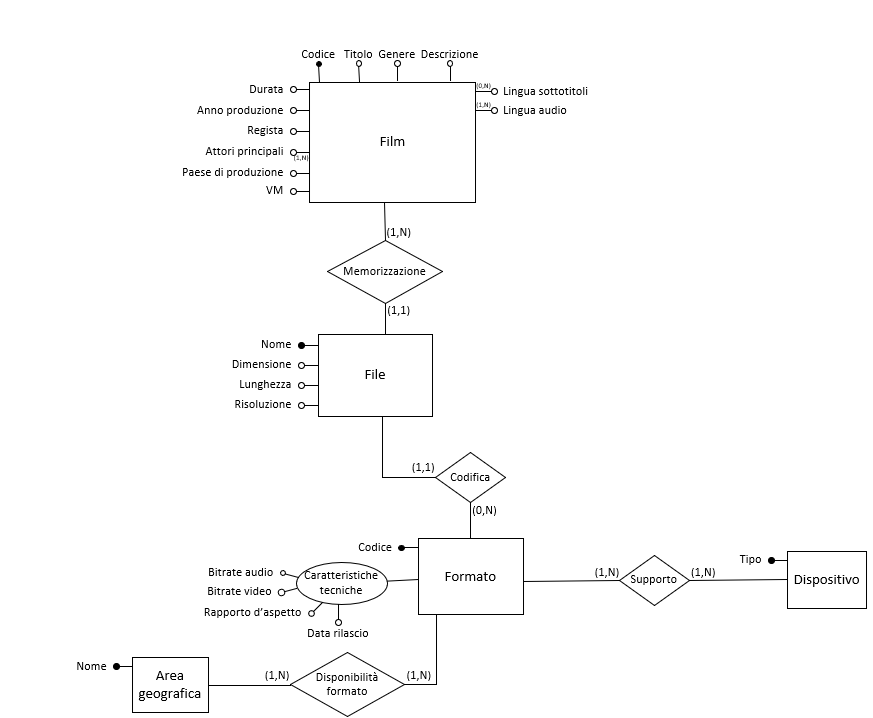
Questa porzione del diagramma E-R descrive l’area contenuti di FilmSphere.

Ogni *Film* è dotato di un attributo *VM* che vale:

* 0 se il *Film* non è vietato ai minori;
* 1 se il *Film* è vietato ai minori.

Ogni *Film* è disponibile in almeno un’*Area geografica*. Ogni *Cliente* può effettuare una *Votazione* di un *Film*, esprimendo un *Voto* intero compreso tra 1 e 10.

* AREA FORMATI

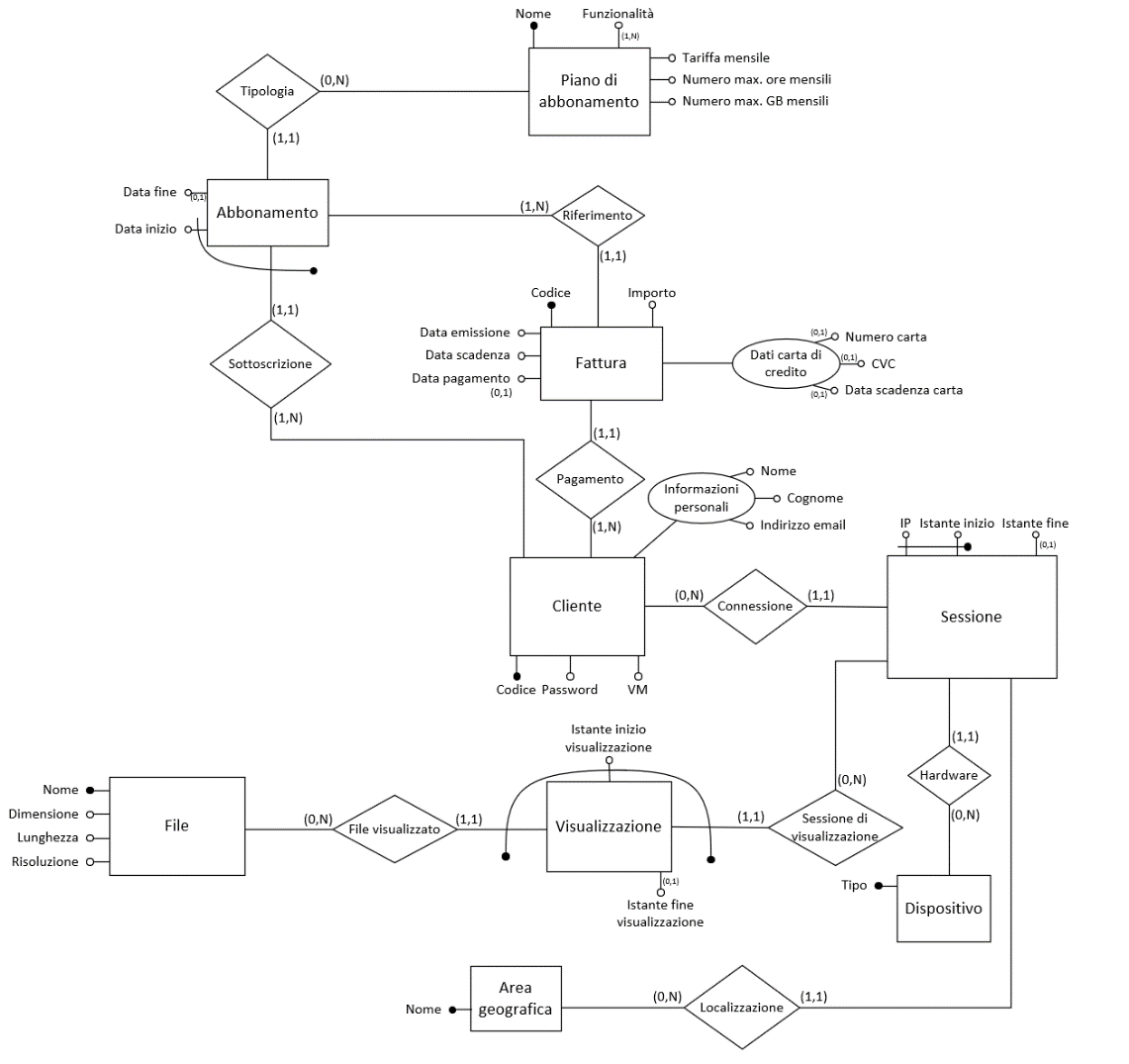


Questa porzione del diagramma E-R descrive l’area formati di FilmSphere.

Ogni *Film* è memorizzato in almeno un *File*, ognuno dotato di un *Nome*, una *Dimensione*, una *Lunghezza* e una *Risoluzione*.

Ogni *File* è codificato in uno e un solo *Formato*. Ciascun *Formato* è supportato da almeno un *Dispositivo*, ed è disponibile in almeno un’*Area geografica*.

* AREA CLIENTI



Questa porzione del diagramma E-R descrive l’area clienti di FilmSphere.

Ogni *Cliente* è dotato di un attributo *VM* che vale:

* 0 se il *Cliente* non può visualizzare *Film* vietati ai minori;
* 1 se il *Cliente* può visualizzare *Film* vietati ai minori.

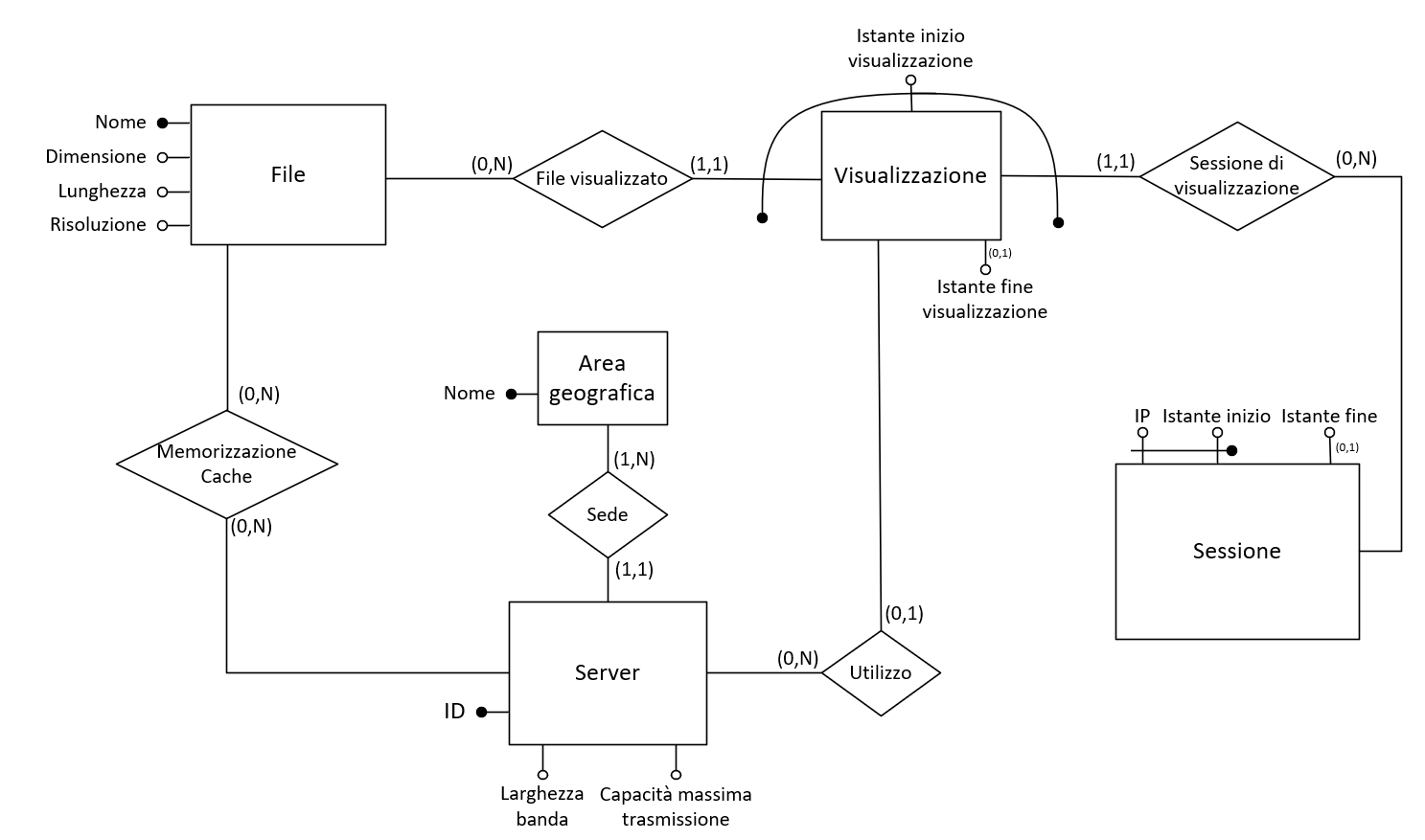
Ogni volta che un *Cliente* si connette al servizio FilmSphere, dà origine ad una *Sessione.* Ogni *Sessione* è identificata dall’*IP* con cui è stata effettuata la connessione e l’*Istante inizio.* Il valore dell’attributo *Istante fine* permette di distinguere una *Sessione* in corso da una terminata.Ciascuna *Sessione* è poi caratterizzata dal *Dispositivo* usato per la connessione e dall*’Area geografica* di provenienza.

Durante una *Sessione* si possono effettuare delle visualizzazioni. Una *Visualizzazione* è identificata dal *File* visualizzato, la *Sessione* in cui tale File è stato visualizzato e dall’*Istante inizio visualizzazione*. Il valore dell’attributo *Istante fine visualizzazione* permette di distinguere una *Visualizzazione* in corso da una terminata.

Ogni *Cliente* può sottoscrivere uno o più abbonamenti, di una certa *Tipologia*. Un *Abbonamento* è identificato dal *Cliente* che l’ha sottoscritto e dalla *Data inizio*. Il valore dell’attributo *Data fine* permette di distinguere un *Abbonamento* in corso da uno terminato. Si suppone che, in fase di iscrizione, ogni *Cliente* debba necessariamente sottoscrivere un abbonamento.

Ogni Cliente effettua poi uno o più pagamenti di fatture relative ad un certo *Abbonamento*. Una *Fattura* è identificata da un *Codice*, e il valore dell’attributo *Data pagamento* permette di distinguere le fatture pagate da quelle non ancora pagate. Si suppone che le fatture siano mensili e che, non appena si sottoscrive un *Abbonamento*, venga subito emessa una *Fattura* relativa a tale *Abbonamento*.

* AREA STREAMING

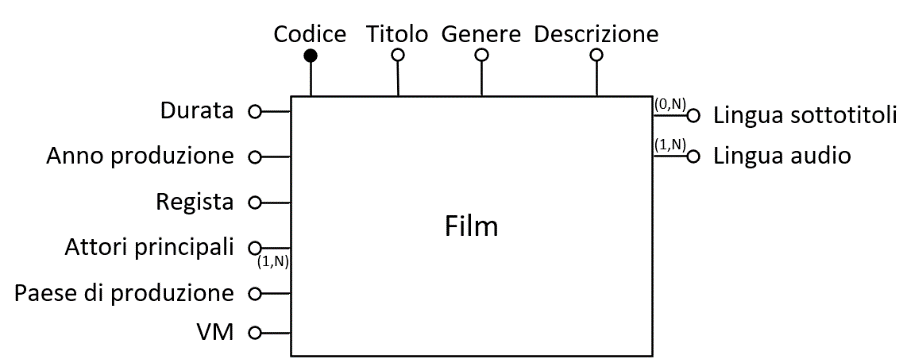
****

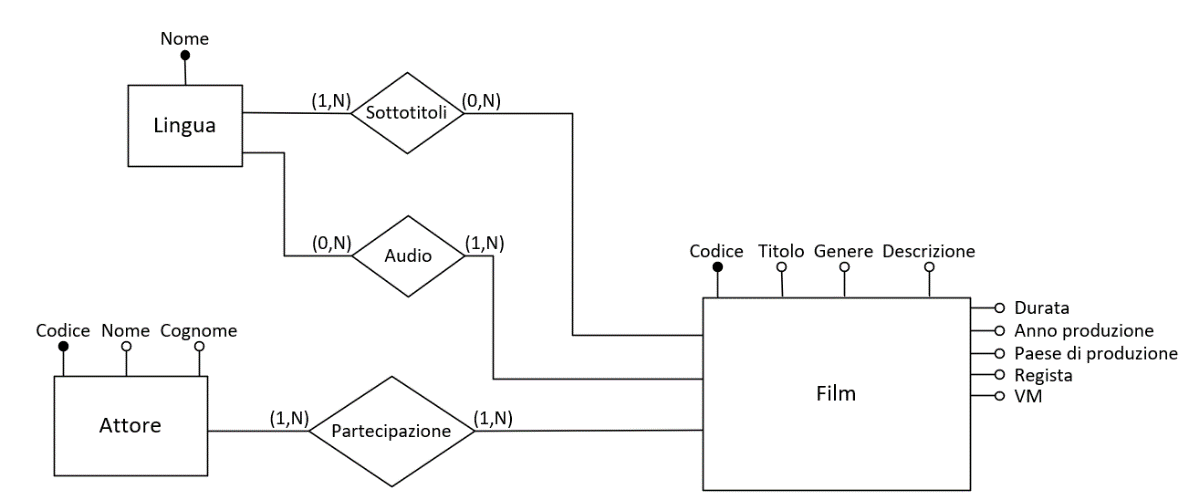
Questa porzione del diagramma E-R descrive l’area streaming di FilmSphere.

Ogni *Server* ha *Sede* in una e una sola *Area geografica* e può presentare in ogni momento dei *File* memorizzati in cache. Ogni *Visualizzazione* può sfruttare in ogni momento un *Server* che presenta il *File* visualizzato in cache.

*Utilizzo* e *Memorizzazione Cache* riguardano dunque rispettivamente utilizzi correnti e memorizzazioni correnti.

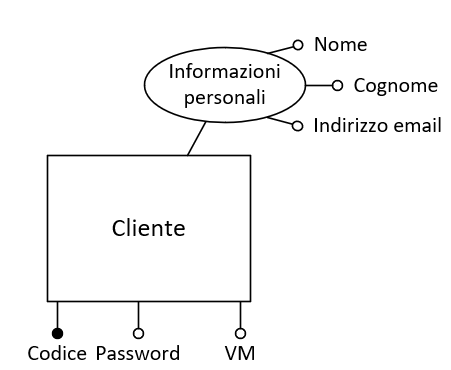
RISTRUTTURAZIONE DIAGRAMMA-ER

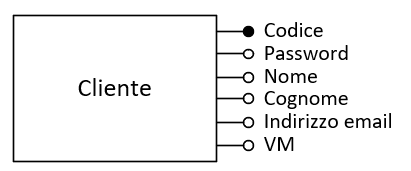
****

****

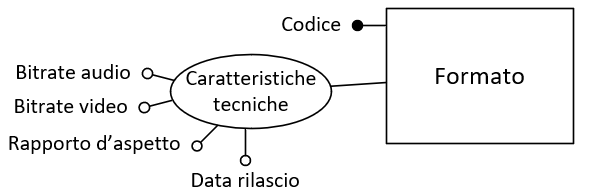
L’attributo multivalore *Attori principali* è stato trasformato in un’entità *Attore,* legata all’entità *Film* tramite la relationship *Partecipazione*.

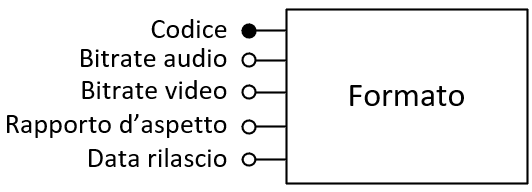
Gli attributi *Lingua sottotitoli* e *Lingua audio* sono stati trasformati in un’entità *Lingua*, legata all’entità *Film* tramite le relationship *Sottotitoli* e *Audio*.



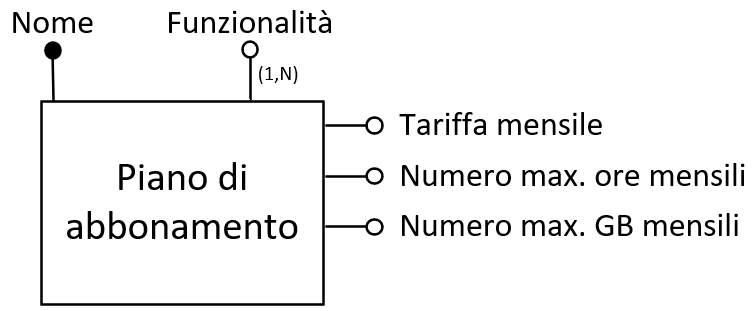


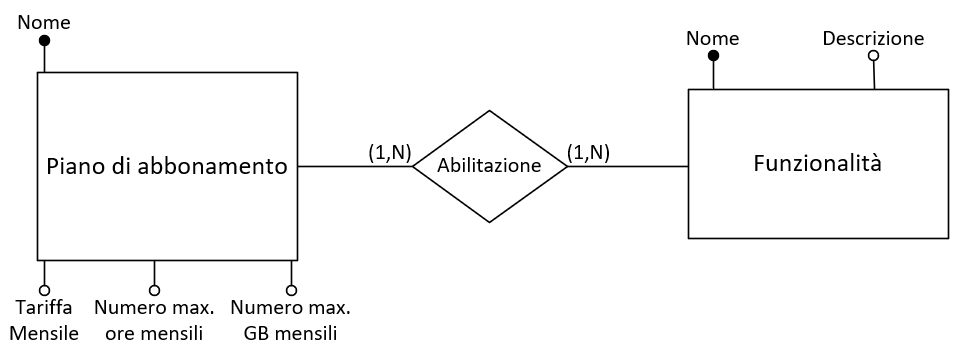
L’attributo composto *Informazioni personali* è stato scomposto nei singoli attributi *Nome*, *Cognome* e *Indirizzo email*



****

L’attributo composto *Caratteristiche tecniche* è stato scomposto nei singoli attributi *Bitrate audio*, *Bitrate video, Rapporto d’aspetto* e *Data rilascio.*

**



L’attributo multivalore *Funzionalità* è stato trasformato in un’entità *Funzionalità*, legata all’entità *Piano di abbonamento* tramite la relationship *Abilitazione*.

TAVOLA DEI VOLUMI

Nell’ipotesi che il servizio *FilmSphere* sia stato lanciato da 30 giorni, la tavola dei volumi è la seguente:

|  |  |  |  |
| --- | --- | --- | --- |
| **CONCETTO** | **TIPO** | **VOLUME** | **MOTIVAZIONE** |
| **AREA CONTENUTI** | | | |
| Film | E | 30 | Ipotesi |
| Lingua | E | 20 | Ipotesi |
| Attore | E | 80 | Ipotesi |
| Cliente | E | 200 | Ipotesi |
| Area geografica | E | 7 | Ipotesi |
| Sottotitoli | R | 120 (30 x 4) | Ipotesi: ogni film è sottotitolato in media in 4 lingue |
| Audio | R | 90 (30 x 3) | Ipotesi: ogni film è doppiato in media in 3 lingue |
| Partecipazione | R | 160 (80 x 2) | Ipotesi: ogni attore partecipa in media a 2 film |
| Votazione | R | 1000 (200 x 5) | Ipotesi: ogni cliente recensisce in media 5 film |
| Disponibilità film | R | 150 (30 x 5) | Ipotesi: ogni film è disponibile in media in 5 aree geografiche |
| **AREA FORMATI** | | | |
| Formato | E | 10 | Ipotesi |
| File | E | 60 (30 x 2) | Ipotesi: ogni film è memorizzato in media in due file |
| Dispositivo | E | 5 | Ipotesi |
| Memorizzazione | R | 60 | Da File, relationship cui partecipa con cardinalità (1,1) |
| Codifica | R | 60 | Da File, relationship cui partecipa con cardinalità (1,1) |
| Supporto | R | 25 (5 x 5) | Ipotesi: un dispositivo supporta in media 5 formati |
| Disponibilità formato | R | 63 (7 x 9) | Ipotesi: ogni area geografica supporta in media 9 formati |

|  |  |  |  |
| --- | --- | --- | --- |
| **AREA CLIENTI** | | | |
| Sessione | E | 6000 (200 x 1 x 30) | Ipotesi: ogni cliente si connette a FilmSphere in media una volta al giorno |
| Visualizzazione | E | 7200 (6000 x 1,2) | Ipotesi: durante ogni sessione sono visualizzati in media 1,2 file |
| Piano di abbonamento | E | 5 | Da specifiche |
| Abbonamento | E | 220 (200 x 1,1) | Ipotesi: ogni cliente ha sottoscritto in media 1,1 abbonamenti |
| Funzionalità | E | 8 | Ipotesi |
| Fattura | E | 220 | Ipotesi: le fatture sono mensili e la prima viene emessa il giorno della sottoscrizione di un abbonamento, e avendo considerato come periodo di tempo i primi 30 giorni di FilmSphere, le fatture sono tante quanti gli abbonamenti |
| Riferimento | R | 220 | Da Fattura, relationship cui partecipa con cardinalità (1,1) |
| Connessione | R | 6000 | Da Sessione, relationship cui partecipa con cardinalità (1,1) |
| Sessione di visualizzazione | R | 7200 | Da Visualizzazione, relationship cui partecipa con cardinalità (1,1) |
| File visualizzato | R | 7200 | Da Visualizzazione, relationship cui partecipa con cardinalità (1,1) |
| Hardware | R | 6000 | Da Sessione, relationship cui partecipa con cardinalità (1,1) |
| Localizzazione | R | 6000 | Da Sessione, relationship cui partecipa con cardinalità (1,1) |
| Pagamento | R | 220 | Da Fattura, relationship cui partecipa con cardinalità (1,1) |

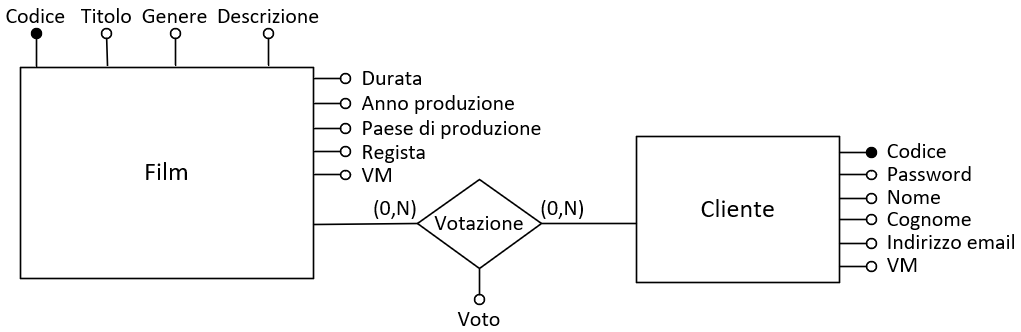
|  |  |  |  |
| --- | --- | --- | --- |
| Sottoscrizione | R | 220 | Da Abbonamento, relationship cui partecipa con cardinalità (1,1) |
| Tipologia | R | 220 | Da Abbonamento, relationship cui partecipa con cardinalità (1,1) |
| Abilitazione | R | 25 (5 x 5) | Ipotesi: ogni piano di abbonamento abilita in media 5 funzionalità |
| **AREA STREAMING** | | | |
| Server | E | 35 (7 x 5) | Ipotesi: per ogni area geografica ci sono in media 5 server |
| Utilizzo | R | 10 | Ipotesi: ad ogni istante ci sono in media 10 visualizzazioni in corso (e quindi 10 utilizzi di server) |
| Sede | R | 35 | Da Server, relationship cui partecipa con cardinalità (1,1) |
| Memorizzazione cache | R | 280 (35 x 8) | Ipotesi: ad ogni istante ogni server ha memorizzati in media 8 file |

OPERAZIONI SUI DATI

1. Vedere il voto complessivo di un film;
2. Calcolare il numero di minuti di film visualizzati da un cliente;
3. Inserire una nuova visualizzazione;
4. Calcolare il numero di abbonamenti in corso relativi ad un piano di abbonamento;
5. Trovare il codice, la data di emissione e la data di scadenza delle fatture non pagate di un cliente, con indicato nome e cognome;
6. Calcolare da quanti giorni un cliente è iscritto;
7. Trovare informazioni sui file in cui è memorizzato un film, e informazioni sui rispettivi formati;
8. Trovare i file memorizzati in un server, con codice del formato e codice del film relativi.

OPERAZIONE 1

* **Descrizione**: l’operazione permette di vedere la media dei voti assegnati ad un certo film;
* **Input**: codice di un film;
* **Output**: media dei voti assegnati al film;
* **Frequenza**: 1000 volte al giorno (si ipotizza che sia costituito dal numero di volte in cui il cliente stesso controlla questa informazione, ad esempio nella scelta del film da visualizzare, più il numero di volte in cui viene controllato lato back-end per statistiche).

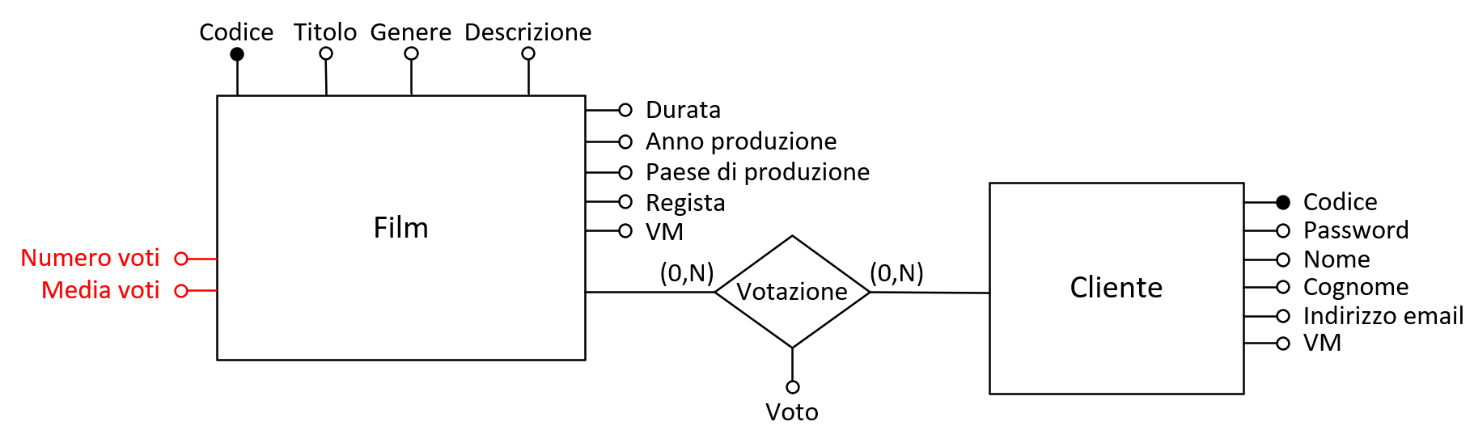


|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CONCETTO** | **COSTRUTTO** | **ACCESSI** | **TIPO** | **DESCRIZIONE** |
| Votazione | R | 34 (le votazioni sono 1000, i film sono 30, dunque un film è stato votato in media 34 volte) | L | Lettura dell’attributo *Voto* delle votazioni relative al film interessato |

Numero di operazioni elementari: 34.

Il costo giornaliero dell’operazione è 1000 x 34 = 34000.

Vediamo il caso in cui sono presenti gli attributi ridondanti *Numero voti* e *Media voti* in Film:



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CONCETTO** | **COSTRUTTO** | **ACCESSI** | **TIPO** | **DESCRIZIONE** |
| Film | E | 1 | L | Lettura dell’attributo *Media voti* del film interessato |

Numero di operazioni elementari: 1.

Il costo giornaliero dell’operazione è 1000 x 1 = 1000.

Vediamo il costo per mantenere aggiornate le ridondanze. Supponendo che un film venga votato in media 1,2 volte al giorno (ogni film ha 34 votazioni in media, e il periodo di tempo considerato è 30 giorni), e i film sono 30, la ridondanza va aggiornata in media 36 volte al giorno.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CONCETTO** | **COSTRUTTO** | **ACCESSI** | **TIPO** | **DESCRIZIONE** |
| Votazione | R | 1 | L | Lettura dell’attributo *Voto* della votazione appena inserita relativa al film interessato |
| Film | E | 1 | L | Lettura degli attributi *Numero voti* e *Media voti* relativi al film interessato |
| Film | E | 1 | S | Aggiornamento degli attributi *Numero voti* e *Media voti* relativi al film interessato |

Numero di operazioni elementari: 4.

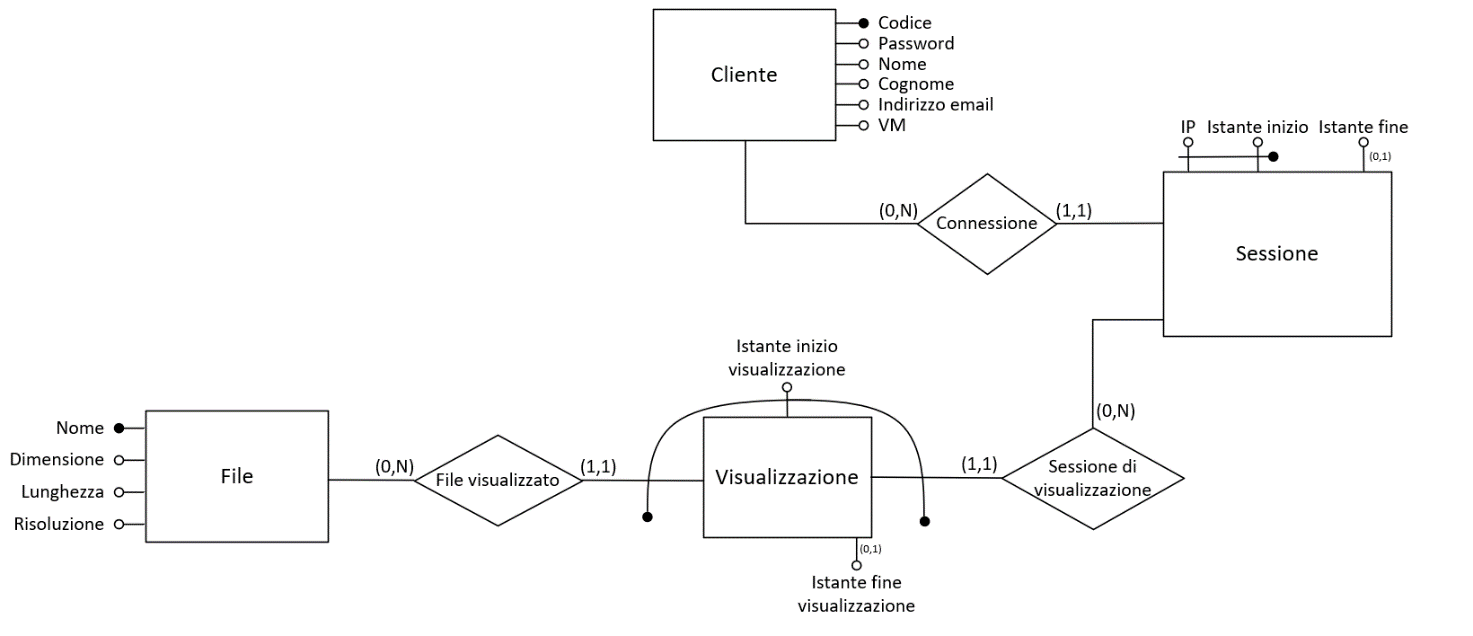
Il costo giornaliero per mantenere aggiornata la ridondanza è 36 x 4 = 144.

Il costo totale giornaliero è 1000 + 144 = 1144.

1144 < 33000, dunque si decide di inserire gli attributi ridondanti *Numero voti* e *Media voti* in Film.

OPERAZIONE 2

* **Descrizione**: l’operazione permette di conoscere il numero di minuti di film che un certo cliente ha visualizzato. Questo viene calcolato come somma del tempo trascorso tra l’istante di inizio visualizzazione e l’istante di fine visualizzazione, per ogni visualizzazione effettuata dal cliente;
* **Input**: codice di un cliente;
* **Output**: numero di minuti di film che il cliente ha visualizzato;
* **Frequenza**: 600 volte al giorno (si ipotizza che sia costituito dal numero di volte in cui il cliente stesso controlla questa informazione, più il numero di volte in cui viene controllato lato back-end per statistiche).



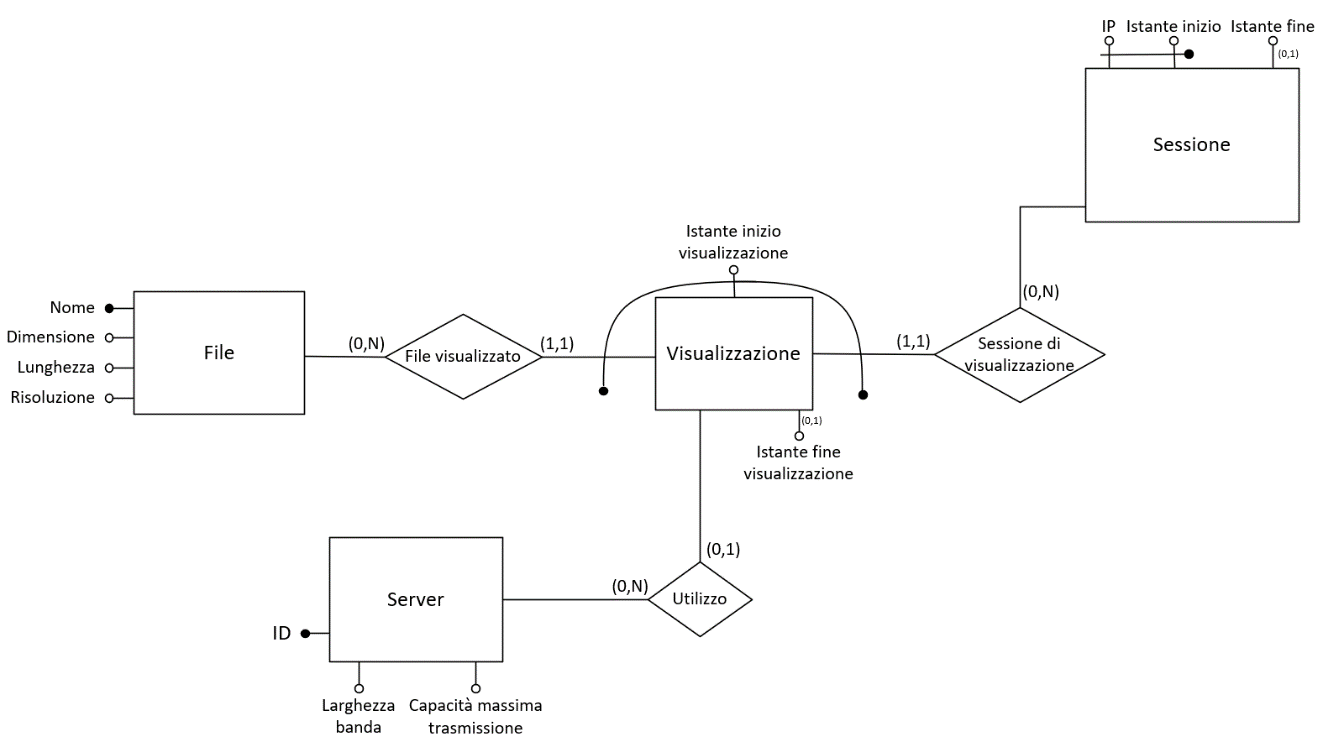
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CONCETTO** | **COSTRUTTO** | **ACCESSI** | **TIPO** | **DESCRIZIONE** |
| Connessione | R | 30 (ogni cliente si connette al servizio in media una volta al giorno, e si sta considerando come periodo i primi 30 giorni) | L | Recupero di tutte le sessioni effettuate dal cliente interessato |
| Sessione di visualizzazione | R | 36 (ogni cliente guarda in media 1,2 file per sessione) | L | Recupero di tutte le visualizzazioni effettuate dal cliente interessato |
| Visualizzazione | E | 36 | L | Lettura dell’attributo *Istante fine visualizzazione* delle visualizzazioni interessate |

Numero di operazioni elementari: 102.

Il costo giornaliero dell’operazione è 600 x 102 = 61200.

OPERAZIONE 3

* **Descrizione**: l’operazione permette di inserire una nuova visualizzazione di un certo file, effettuata in una certa sessione, in un certo istante;
* **Input**: IP e istante di inizio di una sessione, nome di un file, istante di inizio e di fine della visualizzazione;
* **Output**: N/A;
* **Frequenza**: 240 volte al giorno (ci sono 200 sessioni al giorno, e durante una sessione sono visualizzati in media 1,2 file).



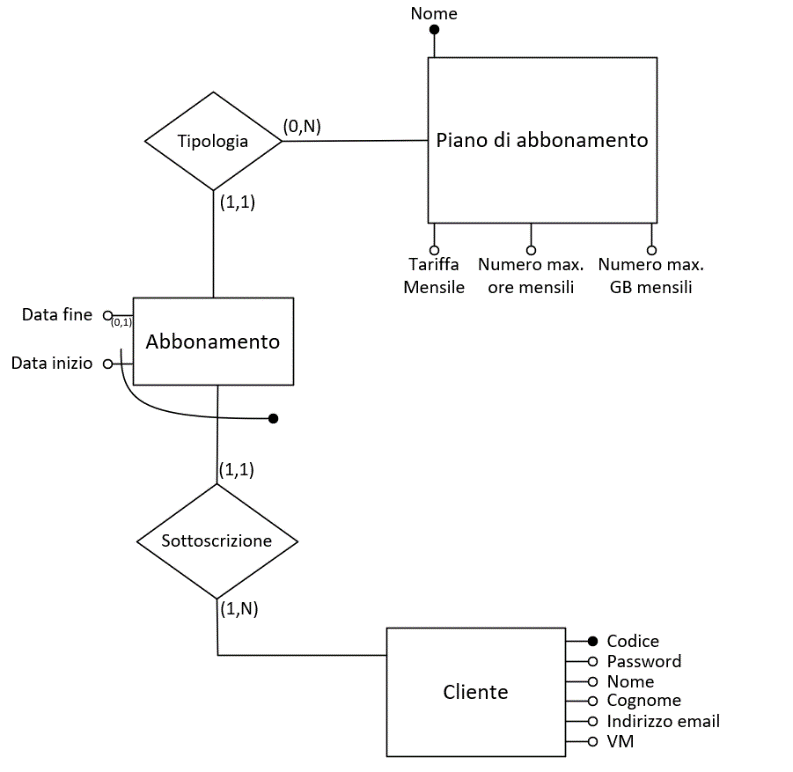
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CONCETTO** | **COSTRUTTO** | **ACCESSI** | **TIPO** | **DESCRIZIONE** |
| Visualizzazione | E | 1 | S | Inserimento di una nuova visualizzazione di un certo file in una certa sessione |
| File visualizzato | R | 1 | S | Si crea la relazione con File |
| Sessione di visualizzazione | R | 1 | S | Si crea la relazione con Sessione |
| Utilizzo | R | 1 | S | Si crea la relazione con Server |

Numero di operazioni elementari: 8.

Il costo giornaliero dell’operazione è 240 x 8 = 1920.

OPERAZIONE 4

* **Descrizione**: l’operazione permette di calcolare il numero di abbonamenti in corso relativi a un certo piano di abbonamento;
* **Input**: nome di un piano di abbonamento;
* **Output**: numero di abbonamenti in corso relativi al piano di abbonamento;
* **Frequenza**: 20 volte al giorno (si ipotizza che sia costituito dal numero di volte in cui viene controllato lato back-end per statistiche).



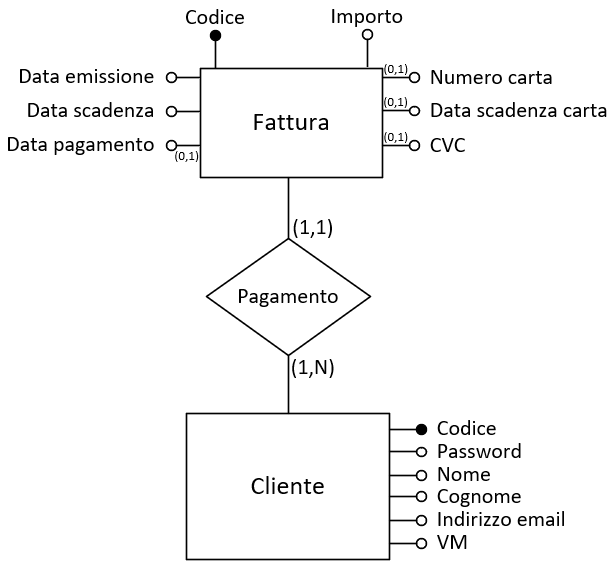
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CONCETTO** | **COSTRUTTO** | **ACCESSI** | **TIPO** | **DESCRIZIONE** |
| Tipologia | R | 44  (gli abbonamenti sono 220, i piani di abbonamento sono 5, dunque ci sono in media 44 abbonamenti per piano) | L | Recupero di tutti gli abbonamenti che riguardano il piano di abbonamento interessato |
| Abbonamento | E | 44 | L | Lettura dell’attributo *Data Fine* degli abbonamenti interessati, per trovare gli abbonamenti in corso |

Numero di operazioni elementari: 88.

Il costo giornaliero dell’operazione è 20 x 88 = 1760.

OPERAZIONE 5

* **Descrizione**: l’operazione permette di trovare il codice, la data di emissione e la data di scadenza delle fatture non pagate di un certo cliente, indicando il suo nome e cognome;
* **Input**: codice di un cliente;
* **Output**: fatture non pagate del cliente;
* **Frequenza**: 20 volte al giorno (si ipotizza che sia costituito dal numero di volte in cui il cliente stesso controlla questa informazione, più il numero di volte in cui viene controllato dagli impiegati del servizio FilmSphere).



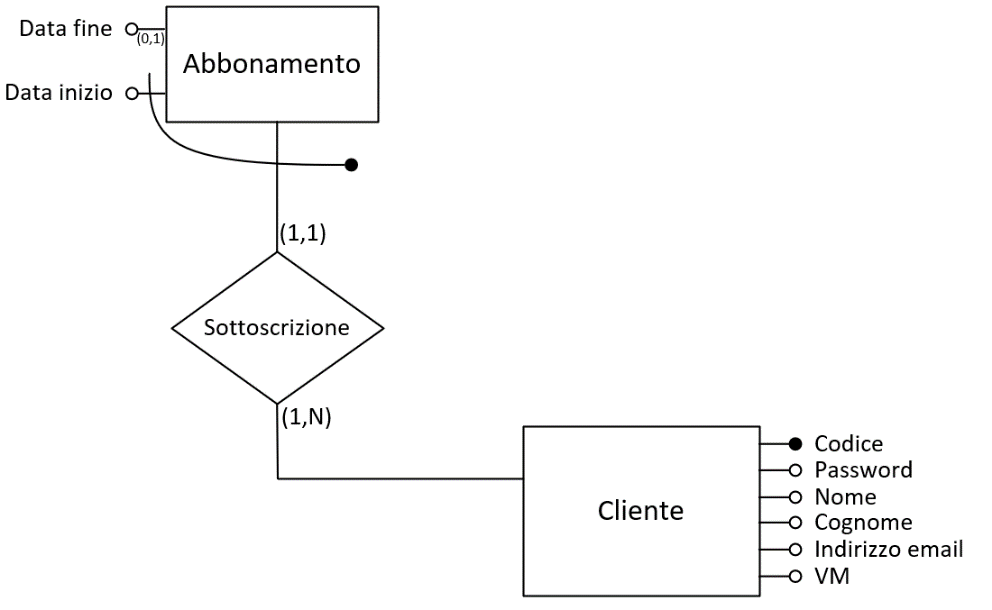
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CONCETTO** | **COSTRUTTO** | **ACCESSI** | **TIPO** | **DESCRIZIONE** |
| Cliente | E | 1 | L | Lettura degli attributi *Nome* e *Cognome* del cliente interessato |
| Pagamento | R | 1,1 (le fatture sono 220, i clienti sono 200, dunque ci sono in media 1,1 fatture per cliente) | L | Recupero delle fatture del cliente interessato |
| Fattura | E | 1,1 | L | Lettura dell’attributo *Data Pagamento* delle fatture interessate, per trovare quelle non pagate, e degli attributi *Data emissione* e *Data scadenza* |

Numero di operazioni elementari: 3,2.

Il costo giornaliero dell’operazione è 20 x 3,2 = 64.

OPERAZIONE 6

* **Descrizione**: l’operazione permette di calcolare da quanti giorni un cliente è iscritto. Visto che, in fase di iscrizione, è obbligatorio sottoscrivere un abbonamento, questo corrisponde a trovare il numero di giorni passati dalla data d’inizio dell’abbonamento più vecchio che il cliente ha sottoscritto;
* **Input**: codice di un cliente;
* **Output**: numero di giorni passati da quanto il cliente si è iscritto;
* **Frequenza**: 600 volte al giorno (si ipotizza che sia costituito dal numero di volte che il cliente stesso controlla questa informazione, più le volte che viene controllato lato back-end per statistiche).

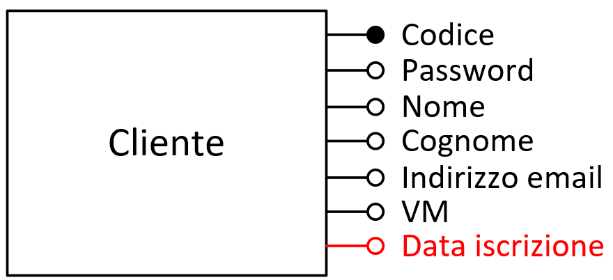


|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CONCETTO** | **COSTRUTTO** | **ACCESSI** | **TIPO** | **DESCRIZIONE** |
| Sottoscrizione | R | 1,1 (ogni cliente ha sottoscritto in media 1,1 abbonamenti) | L | Recupero degli abbonamenti effettuati dal cliente interessato, con lettura dell’attributo *Data Inizio* |

Numero di operazioni elementari: 1,1.

Il costo giornaliero dell’operazione è 600 x 1,1 = 660.

Vediamo il caso in cui è presente l’attributo ridondante *Data iscrizione* in Cliente:



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CONCETTO** | **COSTRUTTO** | **ACCESSI** | **TIPO** | **DESCRIZIONE** |
| Cliente | E | 1 | L | Lettura dell’attributo *Data iscrizione* del cliente interessato |

Numero di operazioni elementari: 1.

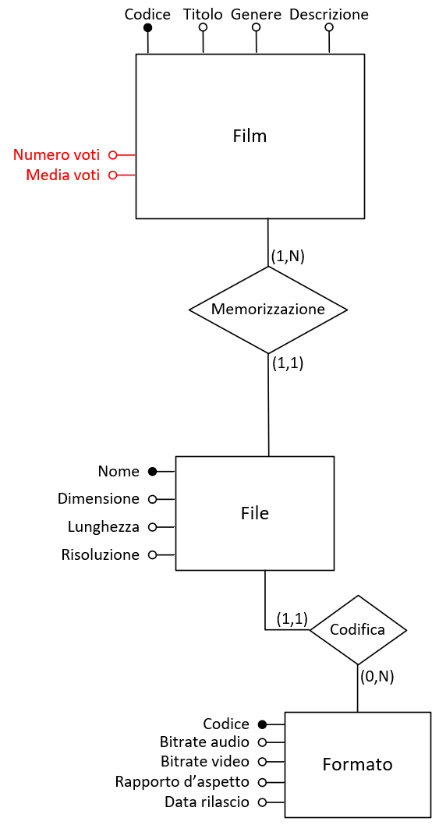
Il costo giornaliero dell’operazione è 600 x 1 = 600.

La ridondanza non va aggiornata, dunque il costo totale giornaliero è 600.

600 < 660. Considerato inoltre che, col passare del tempo, può aumentare il numero di abbonamenti sottoscritti da uno stesso cliente, si decide di inserire l’attributo ridondante *Data iscrizione* in Cliente.

OPERAZIONE 7

* **Descrizione**: l’operazione permette di trovare informazioni sui file in cui è memorizzato un certo film, e informazioni sui rispettivi formati;
* **Input**: codice di un film;
* **Output**: informazioni sui file in cui è memorizzato il film, e informazioni sui rispettivi formati;
* **Frequenza**: 240 volte al giorno (si ipotizza che queste informazioni vengano controllate prima della visualizzazione di ogni film, e il numero è leggermente maggiore del numero di visualizzazioni al giorno perché possono esserci clienti che, una volta visto che un film non è nel formato desiderato, decidono di non visualizzarlo).



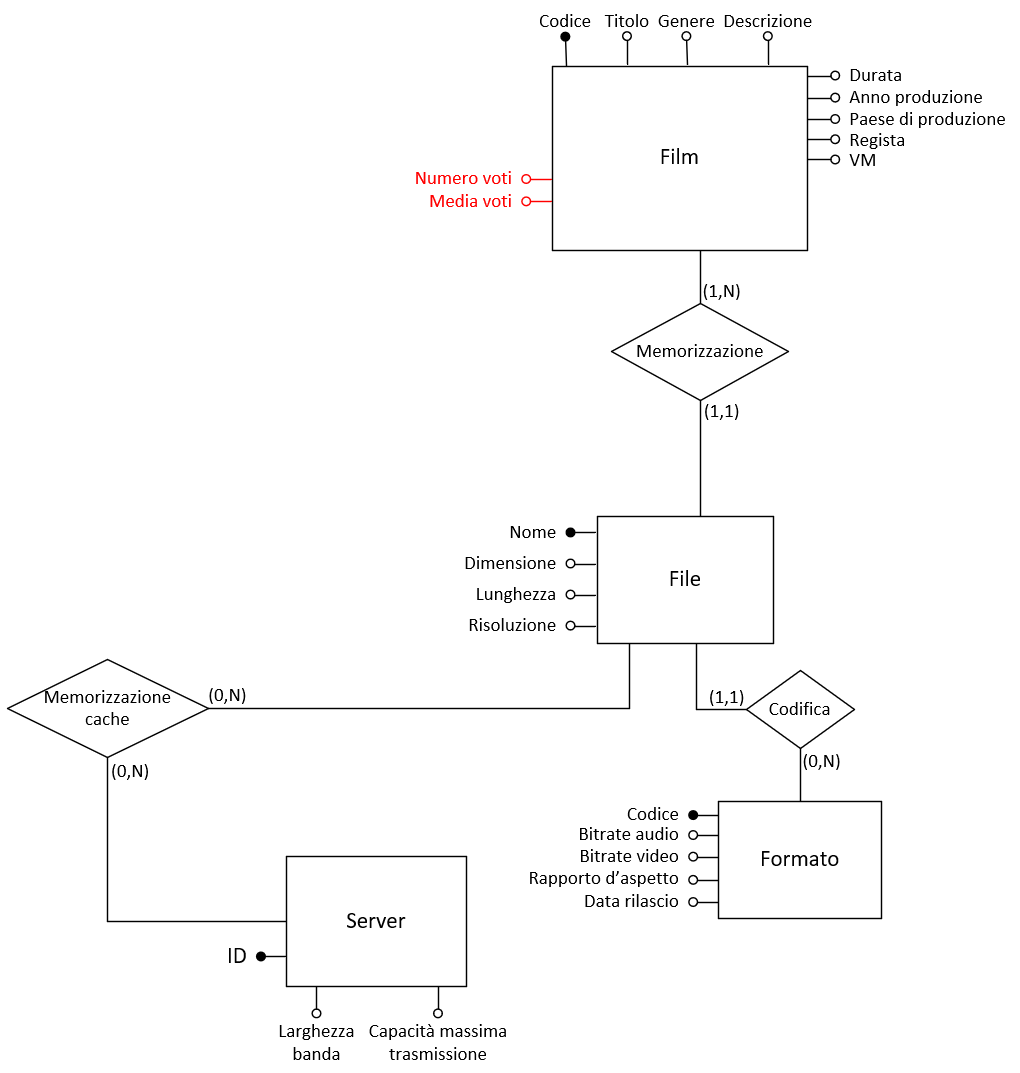
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CONCETTO** | **COSTRUTTO** | **ACCESSI** | **TIPO** | **DESCRIZIONE** |
| Memorizzazione | R | 2 (ogni film è memorizzato in media in due file) | L | Recupero dei file in cui è memorizzato il film interessato |
| File | E | 2 | L | Lettura degli altri attributi dei file interessati |
| Codifica | R | 2 | L | Recupero dei formati in cui sono codificati i file interessati |
| Formato | E | 2 | L | Lettura degli altri attributi dei formati interessati |

Numero di operazioni elementari: 8

Il costo giornaliero dell’operazione è 240 x 8 = 1920.

OPERAZIONE 8

* **Descrizione**: l’operazione permette di trovare informazioni sui file memorizzati in un certo server, insieme al codice del formato e il codice del film relativi;
* **Input**: ID di un server;
* **Output**: informazioni sui file memorizzati nel server, insieme al codice del formato e il codice del film relativo;
* **Frequenza**: 1680 volte al giorno (si ipotizza che venga controllato ogni mezz’ora, per ciascun server).



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CONCETTO** | **COSTRUTTO** | **ACCESSI** | **TIPO** | **DESCRIZIONE** |
| Memorizzazione cache | R | 8 (ogni server memorizza in media 8 file) | L | Recupero dei file memorizzati nel server interessato |
| Codifica | R | 8 | L | Recupero dei formati in cui sono codificati i file interessati |
| File | E | 8 | L | Lettura degli altri attributi dei file interessati |
| Memorizzazione | R | 8 | L | Recupero dei film interessati |

Numero di operazioni elementari: 32.

Il costo giornaliero dell’operazione è 1680 x 32 = 53760.

PROGETTAZIONE LOGICA

AREA CONTENUTI

Film(**Codice,** Titolo, Genere, Descrizione, Durata, AnnoProduzione, PaeseProduzione, Regista, VM, NumeroVoti, MediaVoti)

Lingua(**Nome**)

Attore(**Codice,** Nome, Cognome)

Cliente(**Codice,** Password, Nome, Cognome, IndirizzoEmail, VM, DataIscrizione)

AreaGeografica(**Nome**)

Sottotitoli(**Film, Lingua**)

Audio(**Film, Lingua**)

Partecipazione(**Film, Attore**)

Votazione(**Cliente, Film,** Voto)

DisponibilitaFilm(**Film, AreaGeografica**)

Con vincoli di integrità referenziale tra:

* L’attributo *Film* della relazione Sottotitoli e la chiave *Codice* della relazione Film;
* L’attributo *Lingua* della relazione Sottotitoli e la chiave *Nome* della relazione Lingua;
* L’attributo *Film* della relazione Audio e la chiave *Codice* della relazione Film;
* L’attributo *Lingua* della relazione Audio e la chiave *Nome* della relazione Lingua;
* L’attributo *Film* della relazione Partecipazione e la chiave *Codice* della relazione Film;
* L’attributo *Attore* della relazione Partecipazione e la chiave *Codice* della relazione Attore;
* L’attributo *Cliente* della relazione Votazione e la chiave *Codice* della relazione Cliente;
* L’attributo *Film* della relazione Votazione e la chiave *Codice* della relazione Film;
* L’attributo *Film* della relazione DisponibilitaFilm e la chiave *Codice* della relazione Film;
* L’attributo *AreaGeografica* della relazione DisponibilitaFilm e la chiave *Nome* della relazione AreaGeografica.

AREA FORMATI

Formato(**Codice**, BitrateAudio, BitrateVideo, RapportoAspetto, DataRilascio)

File(**Nome,** Film, Formato, Dimensione, Lunghezza, Risoluzione)

Dispositivo(**Tipo**)

Supporto(**Dispositivo, Formato**)

DisponibilitaFormato(**Formato, AreaGeografica**)

Con vincoli di integrità referenziale tra:

* L’attributo *Film* della relazione File e la chiave *Codice* della relazione Film;
* L’attributo *Formato* della relazione File e la chiave Codice della relazione Formato;
* L’attributo *Dispositivo* della relazione Supporto e la chiave *Tipo* della relazione Dispositivo;
* L’attributo *Formato* della relazione Supporto e la chiave *Codice* della relazione Formato;
* L’attributo *Formato* della relazione DisponibilitaFormato e la chiave *Codice* della relazione Formato;
* L’attributo *AreaGeografica* della relazione DisponibilitaFormato e la chiave *Nome* della relazione AreaGeografica.

AREA CLIENTI

Sessione(**IP, IstanteInizio,** Cliente, Dispositivo, AreaGeografica, IstanteFine)

Visualizzazione(**File, IstanteInizioVisualizzazione, IPSessione, IstanteInizioSessione,** IstanteFineVisualizzazione)

PianoAbbonamento(**Nome,** TariffaMensile, NumeroMaxOreMensili, NumeroMaxGBMensili)

Abbonamento(**DataInizio, Cliente,** PianoAbbonamento, DataFine)

Funzionalita(**Nome,** Descrizione)

Fattura(**Codice,** Importo, DataInizioAbbonamento, Cliente, DataEmissione, DataScadenza, DataPagamento, NumeroCarta, DataScadenzaCarta, CVC)

Abilitazione(**PianoAbbonamento, Funzionalita**)

Con vincoli di integrità referenziale tra:

* L’attributo *File* della relazione Visualizzazione e la chiave *Nome* della relazione File;
* Gli attributi *IPSessione, IstanteInizioSessione* della relazione Visualizzazione e la chiave *IP, IstanteInizio* della relazione Sessione;
* L’attributo *Cliente* della relazione Abbonamento e la chiave *Codice* della relazione Cliente;
* L’attributo *Cliente* della relazione Sessione e la chiave *Codice* della relazione Cliente;
* L’attributo *Dispositivo* della relazione Sessione e la chiave *Tipo* della relazione Dispositivo;
* L’attributo *AreaGeografica* della relazione Sessione e la chiave *Nome* della relazione AreaGeografica;
* L’attributo *Cliente* della relazione Fattura e la chiave *Codice* della relazione Cliente;
* Gli attributi *DataInizioAbbonamento, Cliente,* della relazione Fattura e la chiave *DataInizio, Cliente* della relazione Abbonamento.
* L’attributo *PianoAbbonamento* della relazione Abbonamento e la chiave *Nome* della relazione PianoAbbonamento;
* L’attributo *PianoAbbonamento* della relazione Abilitazione e la chiave *Nome* della relazione PianoAbbonamento;
* L’attributo *Funzionalita* della relazione Abilitazione e la chiave *Nome* della relazione Funzionalita.

AREA STREAMING

Server(**ID,** AreaGeografica, LarghezzaBanda, CapacitaMaxTrasmissione)

Utilizzo(**IstanteInizioVisualizzazione, File, IPSessione, IstanteInizioSessione,** Server)

Non è stata fusa in Visualizzazione per non introdurre troppi valori NULL

MemorizzazioneCache(**Server, File**)

Con vincoli di integrità referenziale tra:

* L’attributo *Server* della relazione Utilizzo e la chiave *ID* della relazione Server;
* Gli attributi *File, IstanteInizioVisualizzazione, IPSessione, IstanteInizioSessione* della relazione Utilizzo e la chiave *File*, *IstanteInizioVisualizzazione*, *IPSessione*, *IstanteInizioSessione* della relazione Visualizzazione;
* L’attributo *AreaGeografica* della relazione Server e la chiave *Nome* della relazione AreaGeografica;
* L’attributo *Server* della relazione MemorizzazioneCache e la chiave *ID* della relazione Server;
* L’attributo *File* della relazione MemorizzazioneCache e la chiave *Nome* della relazione File.

VINCOLI DI INTEGRITA’ GENERICI

In grassetto quelli implementati.

1. **L’attributo *Durata* della relazione Film deve assumere un valore positivo;**
2. L’attributo *Durata* della relazione Filmdeve essere espresso in minuti;
3. **L’attributo *AnnoProduzione* della relazione Film deve assumere un valore positivo e non maggiore dell’anno corrente;**
4. **L’attributo *VM* della relazione Film ha dominio {0, 1};**
5. **L’attributo *DataIscrizione* della relazione Cliente non può essere successivo alla data corrente;**
6. **Non possono esistere valori duplicati per l’attributo *IndirizzoEmail* della relazione Cliente;**
7. **L’attributo *VM* della relazione Film ha dominio {0, 1};**
8. **L’attributo *Voto* della relazione Votazione ha dominio {1, 2, 3, 4, 5, 6, 7, 8, 9, 10};**
9. **Gli attributi *BitrateAudio*, *BitrateVideo* della relazione Formato devono assumere un valore positivo;**
10. Gli attributi *BitrateAudio* e *BitrateVideo* devono essere espressi in kbps;
11. L’attributo RapportoAspetto della relazione Formato deve assumere un valore nella forma “a : b”, dove a e b sono due numeri naturali;
12. **L’attributo *DataRilascio* della relazione Formato non può essere successivo alla data corrente;**
13. **Gli attributi *Dimensione, Lunghezza* della relazione File devono assumere un valore positivo;**
14. L’attributo *Dimensione* della relazione *File* deve essere espresso in MB;
15. L’attributo *Lunghezza* della relazione *File* deve essere espresso in minuti;
16. L’attributo *Risoluzione* della relazione File deve assumere un valore nella forma “a x b”, dove a e b sono due numeri naturali;
17. L’attributo *IP* della relazione Sessione deve assumere un valore nella forma “a.b.c.d”, dove a, b, c, d sono quattro numeri interi in [0, 255];
18. **Gli attributi *IstanteInizio, IstanteFine* della relazione Sessione non possono essere successivi all’istante corrente, e *IstanteFine* deve essere successivo a *IstanteInizio*. La sessione, inoltre, è valida solo se il cliente ha un abbonamento in corso nell’istante di inizio della sessione;**
19. **Gli attributi *IstanteInizioVisualizzazione, IstanteFineVisualizzazione* della relazione Visualizzazione non possono essere successivi all’istante corrente, e *IstanteFineVisualizzazione* deve essere successivo a *IstanteInizioVisualizzazione*. *IstanteInizioVisualizzazione* e *IstanteFineVisualizzazione* inoltre, devono essere compresi tra l’istante di inizio e l’istante di fine della sessionein cui avviene tale visualizzazione;**
20. **L’attributo *File* della relazione Visualizzazione deve essere relativo ad un film disponibile nell’area geografica della sessione in cui avviene tale visualizzazione, deve avere un formato disponibile nell’area geografica, deve essere supportato dal dispositivo usato nella sessione e, se il film è vietato ai minori, il cliente deve essere abilitato a guardarlo;**
21. **Gli attributi *TariffaMensile, NumeroMaxOreMensili, NumeroMaxGBMensili* della relazione PianoAbbonamento devono assumere un valore positivo;**
22. **Gli attributi *DataInizio, DataFine* della relazione Abbonamento non possono essere successivi alla data corrente, e *DataFine* deve essere successiva a *DataInizio*. *DataInizio*, inoltre, non deve essere precedente a *DataIscrizione* del cliente cui l'abbonamento è relativo;**
23. Ogni cliente può usufruire unicamente dei servizi che il suo attuale abbonamento gli concede;
24. **L’attributo *Importo* della relazione Fattura deve assumere un valore positivo;**
25. **Gli attributi *DataEmissione, DataPagamento* della relazione Fattura non possono essere successivi alla data corrente, e *DataPagamento* e *DataScadenza* non possono essere precedenti a *DataEmissione*. *DataEmissione*, inoltre, non deve essere precedente a *DataInizioAbbonamento*;**
26. L’attributo *NumeroCarta* della relazione Fattura deve essere costituito da 16 cifre;
27. L’attributo *DataScadenzaCarta* della relazione Fattura deve assumere un valore nella forma “mm/aa”, dove mm è il numero di un mese, e aa sono le ultime due cifre di un anno. Inoltre, la carta non deve essere scaduta quando si paga una fattura;
28. L’attributo *CVC* della relazione Fattura deve essere costituito da 3 cifre;
29. **Gli attributi *LarghezzaBanda, CapacitaMaxTrasmissione* della relazione Server devono assumere un valore positivo, e *CapacitaMaxTrasmissione* deve essere minore o uguale a *LarghezzaBanda*;**
30. Gli attributi *LarghezzaBanda* e *CapacitaMaxTrasmissione* della relazione Server devono essere espressi in Mbps.

ANALISI DIPENDENZE FUNZIONALI E NORMALIZZAZIONE

AREA CONTENUTI

Film(**Codice,** Titolo, Genere, Descrizione, Durata, AnnoProduzione, PaeseProduzione, Regista, VM, NumeroVoti, MediaVoti)

Codice -> Titolo, Genere, Descrizione, Durata, AnnoProduzione, PaeseProduzione, Regista, VM, NumeroVoti, MediaVoti

E’ già in BCNF.

Lingua(**Nome**)

E’ già in BCNF.

Attore(**Codice,** Nome, Cognome)

Codice -> Nome, Cognome

E’ già in BCNF.

Cliente(**Codice,** Password, Nome, Cognome, IndirizzoEmail, VM, DataIscrizione)

Codice -> Password, Nome, Cognome, IndirizzoEmail, VM, DataIscrizione

IndirizzoEmail -> Codice, Password, Nome, Cognome, VM, DataIscrizione (per il vincolo 6)

E’ già in BCNF (la parte sinistra di entrambe le dipendenze funzionali costituisce una superchiave).

AreaGeografica(**Nome**)

E’ già in BCNF.

Sottotitoli(**Film, Lingua**)

E’ già in BCNF.

Audio(**Film, Lingua**)

E’ già in BCNF.

Partecipazione(**Film, Attore**)

E’ già in BCNF.

Votazione(**Cliente, Film,** Voto)

Cliente, Film -> Voto

E’ già in BCNF.

DisponibilitaFilm(**Film, AreaGeografica**)

E’ già in BCNF.

AREA FORMATI

Formato(**Codice**, BitrateAudio, BitrateVideo, RapportoAspetto, DataRilascio)

Codice -> BitrateAudio, BitrateVideo, RapportoAspetto, DataRilascio

E’ già in BCNF.

File(**Nome,** Film, Formato, Dimensione, Lunghezza, Risoluzione)

Nome -> Film, Formato, Dimensione, Lunghezza, Risoluzione

E’ già in BCNF (Film non determina nulla, in quanto dato uno stesso film possono esserci file di formato diverso, dimensione diversa, lunghezza diversa e risoluzione diversa)

Dispositivo(**Tipo**)

E’ già in BCNF.

Supporto(**Dispositivo, Formato**)

E’ già in BCNF.

DisponibilitaFormato(**Formato, AreaGeografica**)

E’ già in BCNF.

AREA CLIENTI

Sessione(**IP, IstanteInizio,** Cliente, Dispositivo, AreaGeografica, IstanteFine)

IP, IstanteInizio -> Cliente, Dispositivo, AreaGeografica, IstanteFine

E’ già in BCNF (IP non determina AreaGeografica, perché in un istante potrebbe essere relativo ad una certa area geografica; in un altro istante potrebbe essere relativo ad un’altra)

Visualizzazione(**File, IstanteInizioVisualizzazione, IPSessione, IstanteInizioSessione,** IstanteFineVisualizzazione)

File, IstanteInizioVisualizzazione, IPSessione, IstanteInizioSessione ->IstanteFineVisualizzazione

E’ già in BCNF.

PianoAbbonamento(**Nome,** TariffaMensile, NumeroMaxOreMensili, NumeroMaxGBMensili)

Nome ->TariffaMensile, NumeroMaxOreMensili, NumeroMaxGBMensili

E’ già in BCNF.

Abbonamento(**DataInizio, Cliente,** PianoAbbonamento, DataFine)

DataInizio, Cliente ->PianoAbbonamento, DataFine

E’ già in BCNF.

Funzionalita(**Nome,** Descrizione)

Nome -> Descrizione

E’ già in BCNF.

Fattura(**Codice,** Importo, DataInizioAbbonamento, Cliente, DataEmissione, DataScadenza, DataPagamento, NumeroCarta, DataScadenzaCarta, CVC)

Codice ->Importo, DataInizioAbbonamento, Cliente, DataEmissione, DataScadenza, DataPagamento, NumeroCarta, DataScadenzaCarta, CVC

E’ già in BCNF (NumeroCarta non determina DataScadenzaCarta e CVC, perché rinnovando la carta si potrebbe mantenere lo stesso NumeroCarta, ma con DataScadenzaCarta e CVC diversi)

Abilitazione(**PianoAbbonamento, Funzionalita**)

E’ già in BCNF.

AREA STREAMING

Server(**ID,** AreaGeografica, LarghezzaBanda, CapacitaMaxTrasmissione)

ID ->AreaGeografica, LarghezzaBanda, CapacitaMaxTrasmissione

E’ già in BCNF.

Utilizzo(**IstanteInizioVisualizzazione, File, IPSessione, IstanteInizioSessione,** Server)

IstanteInizioVisualizzazione, File, IPSessione, IstanteInizioSessione -> Server

E’ già in BCNF.

MemorizzazioneCache(**Server, File**)

E’ già in BCNF.

FUNZIONALITÀ BACKEND

Quando viene avviata una nuova sessione da parte di un cliente, viene fatto caching dei contenuti di cui tale cliente potrebbe richiedere la visualizzazione in opportuni server del CDN.

Si ipotizza che, ad ogni nuova visualizzazione, il server coinvolto riservi ad essa 10 Mbps. Definendo allora la *Capacità di trasmissione libera* di un server come la differenza tra la sua *Capacità massima di trasmissione* e la *Capacità di trasmissione* riservata alle visualizzazioni in corso (10 \* il numero di visualizzazioni che coinvolgono il server), entrambe espresse in Mbps, la metrica per determinare quanto un server è opportuno per una certa sessione è la seguente:

Con *k* una costante che vale:

* 1 se il server si trova nella stessa area geografica della sessione;
* 0,5 se il server si trova in un’area geografica differente.

Più la metrica è alta, più il server è conveniente per la sessione, e la costante *k* permette di dare maggiore priorità ai server più vicini, in modo da offrire un servizio più efficiente possibile. Un server con la   
metrica ≤ 0 è considerato sovraccarico.

Per la scelta dei contenuti, si è ipotizzato che il cliente che ha avviato la sessione vedrà molto probabilmente altri film del suo genere preferito. Allora, una volta determinati i 5 server con la metrica più alta e non sovraccarichi, questi vengono riempiti con al massimo 5 file relativi ai film con la media voti più alta del genere preferito del cliente, disponibili nell’area geografica della sessione, in un formato supportato dal dispositivo della sessione e disponibile nell’area geografica della sessione, escludendo i film vietati ai minori se il cliente non è abilitato a visualizzarli.

Questa funzionalità è stata implementata tramite una stored procedure, che viene chiamata da un trigger dopo ogni inserimento di una *Sessione* in corso.

Quando un cliente richiede di visualizzare un file, viene scelto il server del CDN più conveniente che ha il contenuto richiesto, usando come misura della convenienza la stessa metrica definita alla pagina precedente.

Una volta individuato il server più conveniente che ha memorizzato il file richiesto, se non è sovraccarico, verrà utilizzato per tale visualizzazione.

Questa funzionalità è stata implementata tramite una stored procedure, che viene chiamata da un trigger dopo ogni inserimento di una *Visualizzazione* in corso.

Periodicamente, per ogni visualizzazione in corso, viene rideterminato qual è il server del CDN più conveniente che ha il contenuto richiesto, usando come misura della convenienza la stessa metrica definita alla pagina precedente.

Per ogni visualizzazione in corso, se il server più conveniente che ha memorizzato il file richiesto è diverso da quello attualmente coinvolto e non è sovraccarico, tale server verrà utilizzato per la visualizzazione al posto del precedente.

Questa funzionalità è stata implementata tramite un event che va in esecuzione ogni 30 minuti e, per ogni visualizzazione in corso, chiama la stored procedure della funzionalità precedente.

Periodicamente, i server del CDN vengono svuotati dei file non coinvolti in alcuna visualizzazione in corso.

Questa funzionalità è stata implementata tramite un event che va in esecuzione ogni 6 ore.

DATA ANALYTICS

CLASSIFICHE

**CLASSIFICA PER AREA GEOGRAFICA**

Questa funzionalità stampa a video una classifica dei contenuti più visti e dei loro formati, a seconda dell’area geografica.

Per ogni area geografica, vengono considerati solo i film disponibili nell’area geografica, e solo i formati disponibili nell’area geografica. Nella classifica compaiono anche contenuti che non sono ancora stati visualizzati.

Questa funzionalità è stata implementata tramite una stored procedure.

**CLASSIFICA PER ABBONAMENTO**

Questa funzionalità stampa a video una classifica dei contenuti più visti e dei loro formati, a seconda dell’abbonamento dei clienti al momento della visualizzazione.

Nella classifica compaiono anche contenuti che non sono ancora stati visualizzati, e piani di abbonamento con i quali nessun cliente ha mai effettuato una visualizzazione.

Questa funzionalità è stata implementata tramite una stored procedure.

BILANCIAMENTO DEL CARICO

Questa funzionalità stima i potenziali sovraccarichi di un server e suggerisce spostamenti di contenuti da un server per non sovraccaricarlo, indicando un server di destinazione.

La metrica per stimare quanto un server è potenzialmente sovraccarico è la seguente:

Più la metrica è alta, meno il server è sovraccarico. Se un server ha la metrica < 0,2, allora è potenzialmente sovraccarico, e sarà interessato nei suggerimenti di spostamenti di suoi contenuti.

Inizialmente, per ciascun server, viene considerato il numero di visualizzazioni in corso in cui ciascun file memorizzato è attualmente coinvolto.

A questo punto, per ogni file memorizzato, viene individuato il server con la metrica più alta (e cioè, il server meno sovraccarico al momento), e se questo è diverso dal server che ha attualmente memorizzato il file, il server che ha attualmente memorizzato il file è potenzialmente sovraccarico e spostando il file nel server con la metrica più alta (e, con esso, le visualizzazioni in corso coinvolte) questo non diventa a sua volta potenzialmente sovraccarico, allora la funzionalità suggerirà lo spostamento tra i due server.

La funzionalità sfrutta due tabelle:

* *Suggerimento*, in cui vengono aggiunti via via i suggerimenti di spostamenti di contenuti;
* *ServerEUtilizziTeorici,* che, in base ai suggerimenti che vengono via via inseriti in *Suggerimento*, tiene conto di come si aggiorna la metrica e la capacità di trasmissione libera di ogni server.

Alla fine, la funzionalità stamperà a video la tabella *Suggerimento*.

Questa funzionalità è stata implementata tramite una stored procedure.