Netflix

Scritto da Febogamedev

Storia

Netflix è un servizio di streaming video on demand e live.

La storia di Netflix è particolarmente interessante perché non coincide semplicemente con la crescita di una piattaforma di streaming.
È piuttosto una successione di trasformazioni del prodotto, del modello di business e dell'infrastruttura tecnologica.
Netflix nasce infatti come servizio di noleggio di DVD, diventa progressivamente una piattaforma di streaming, costruisce una propria rete globale per distribuire video, sposta gran parte della propria infrastruttura nel cloud, entra direttamente nella produzione audiovisiva e, infine, amplia nuovamente il proprio modello introducendo pubblicità e contenuti live.
Questa evoluzione è importante perché permette di osservare come le scelte tecnologiche non siano indipendenti dal servizio offerto: ogni volta che Netflix cambia il proprio prodotto, deve modificare anche l'architettura necessaria per sostenerlo.

Netflix viene fondata nel 1997 da Reed Hastings e Marc Randolph.
L'idea iniziale consiste nell'utilizzare Internet per consentire agli utenti di scegliere i film da casa e riceverli fisicamente tramite posta.
Il servizio commerciale viene aperto nel 1998 come negozio online dedicato al noleggio e alla vendita di DVD (Netflix, 1998).
Il DVD era allora un formato relativamente recente e, proprio grazie alle sue dimensioni ridotte e alla maggiore resistenza rispetto alle videocassette, poteva essere spedito con relativa facilità.
In questa prima fase Netflix non è quindi una società di streaming. Internet viene utilizzato soprattutto come interfaccia per la selezione e la gestione del catalogo, mentre il contenuto audiovisivo continua a essere distribuito fisicamente.

La prima innovazione importante riguarda però già il modello di fruizione.
Nel 1999 Netflix passa progressivamente a un sistema basato sull'abbonamento: l'utente paga una quota periodica e può ricevere DVD senza il tradizionale modello di noleggio legato a ogni singolo titolo. Questa logica introduce uno degli elementi che rimarranno centrali anche nello streaming: l'utente non acquista il singolo contenuto, ma paga per accedere continuativamente a un catalogo (Netflix, Inc., 2003).

Già nei primi anni emerge inoltre un'altra caratteristica destinata a diventare fondamentale per Netflix: la personalizzazione.
Nel 2000 l'azienda introduce Cinematch, un sistema che utilizza le valutazioni fornite dagli utenti per costruire un profilo dei loro gusti e proporre film potenzialmente interessanti (Netflix, 2000).

Netflix inizia quindi a utilizzare i dati non soltanto per amministrare il catalogo, ma anche per decidere quali contenuti mostrare a ciascun utente.

Si forma così molto presto una struttura concettuale che rimarrà riconoscibile negli anni successivi:
Catalogo → comportamento dell'utente → dati → raccomandazioni → nuova scelta del contenuto.

Il passaggio decisivo avviene nel 2007.
Netflix introduce una funzione che permette agli abbonati di guardare immediatamente una selezione di film e serie televisive sul proprio computer.
Al momento del lancio il catalogo disponibile in streaming comprende circa un migliaio di titoli e il servizio viene inizialmente proposto come funzionalità aggiuntiva dell'abbonamento DVD (Netflix, 2007).
Questa scelta modifica radicalmente la natura tecnica di Netflix.

- Con i DVD il problema principale consisteva nella gestione di un'infrastruttura logistica: Catalogo → magazzino → spedizione → utente.
- Con lo streaming il percorso diventa invece: Catalogo digitale → infrastruttura informatica → rete Internet → player → utente.

La qualità del servizio non dipende più soltanto dalla disponibilità del titolo, ma anche dalla capacità di elaborarlo, archiviarlo e trasferirlo attraverso Internet con sufficiente velocità. Entrano quindi in gioco problemi come compressione video, bitrate, buffering, compatibilità dei dispositivi, capacità dei server e distribuzione geografica dei contenuti.

L'aumento della diffusione della banda larga rende progressivamente questo modello sempre più praticabile. Netflix comincia quindi a estendere la riproduzione oltre il personal computer, raggiungendo progressivamente televisori, console, dispositivi mobili e altri apparecchi connessi.

Il cambiamento non è soltanto tecnologico.
Lo streaming elimina l’attesa della spedizione fisica, pur mantenendo un tempo di avvio legato alla rete, al buffering e al dispositivo.
La distribuzione audiovisiva si trasforma così da un processo logistico asincrono a un servizio digitale che deve reagire quasi immediatamente alla richiesta dell'utente.

La rapida crescita dello streaming pone però un secondo problema: l'infrastruttura informatica deve essere in grado di crescere molto più rapidamente rispetto al passato.
Netflix individua un punto di svolta nel 2008, quando un grave problema al database impedisce per circa tre giorni la spedizione dei DVD (Izrailevsky et al., 2016).
L'azienda conclude che un'architettura basata su grandi sistemi centralizzati e su singoli punti critici non è sufficientemente resiliente per sostenere la crescita futura.
Nello stesso anno inizia quindi il percorso di migrazione verso Amazon Web Services (*AWS*).

Il cambiamento è concettualmente molto importante.

- L'infrastruttura tradizionale tendeva a crescere soprattutto attraverso il potenziamento dei singoli sistemi: Server più grandi → maggiore capacità.

- Netflix inizia invece a orientarsi verso sistemi distribuiti e scalabili orizzontalmente: più istanze → più servizi distribuiti → capacità elastica.

Il cloud consente di aggiungere rapidamente risorse computazionali e storage quando aumenta la domanda, evitando che la crescita internazionale richieda continuamente la costruzione di nuovi data center proprietari.

La migrazione dura diversi anni.
Prima del 2015 la maggior parte dei servizi rivolti agli utenti era già stata trasferita nel cloud; nel 2016 Netflix annuncia il completamento della migrazione dei sistemi legati al servizio di streaming e la dismissione degli ultimi componenti rimasti nei propri data center tradizionali (Izrailevsky et al., 2016).

AWS diventa fondamentale per il backend: autenticazione, account, database, raccomandazioni, elaborazione dei dati, transcodifica e numerosi altri servizi (Izrailevsky et al., 2016).
La distribuzione vera e propria dei file video seguirà invece una strada differente.

Più aumenta il numero degli utenti, infatti, più diventa difficile distribuire enormi quantità di video attraverso *CDN* generaliste.
Durante i primi anni dello streaming Netflix utilizza *CDN* commerciali esterne.
Con l'aumento del traffico, però, l'azienda comprende che la propria situazione presenta una caratteristica particolare: dispone di un catalogo finito di file audiovisivi e può conoscere con buona approssimazione quali contenuti saranno richiesti maggiormente nelle diverse aree.

Nel 2011 inizia quindi lo sviluppo di una propria Content Delivery Network, presentata pubblicamente nel 2012 con il nome Open Connect (Netflix, n.d.-e; Haspilaire, 2022).
La logica è molto diversa da quella di un unico grande data center centrale.
- Anziché far viaggiare continuamente ogni film attraverso grandi distanze: Origin remoto → Internet → ISP → utente,
- Netflix può collocare copie dei contenuti molto più vicino al pubblico: Open Connect Appliance → rete ISP → utente.

Gli Open Connect Appliances (*OCA*) sono server specializzati che conservano copie dei contenuti Netflix e costituiscono gli edge server della *CDN* Open Connect.
- Alcuni vengono installati direttamente nelle reti degli ISP. Netflix fornisce l'hardware senza addebitarne il costo all'operatore; l'ISP mette invece a disposizione spazio nel data center, alimentazione elettrica e connettività di rete (Netflix, n.d.-e).
- Altri vengono collocati nei siti Open Connect e nei punti di interconnessione con le reti degli operatori, da cui possono distribuire i contenuti agli utenti.

La peculiarità è che il sistema non deve necessariamente aspettare la richiesta di un utente per trasferire ogni contenuto. Netflix può analizzare la popolarità prevista dei titoli e distribuire preventivamente i file verso gli *OCA* delle diverse regioni.
Nel 2022 l'azienda dichiarava circa 18.000 server distribuiti in 6.000 località di 175 paesi (Haspilaire, 2022).

Inizia così a emergere la separazione che ancora oggi rende l'architettura Netflix particolarmente interessante:

- *AWS* → calcolo, dati e logica del servizio.
- Open Connect → distribuzione dei contenuti audiovisivi.

Questa distinzione tra control plane e media delivery diventerà fondamentale per comprendere l'architettura attuale.
Mentre modifica la propria infrastruttura, Netflix cambia anche il proprio ruolo nell'industria audiovisiva.
Inizialmente il catalogo è costituito principalmente da contenuti prodotti da altri soggetti e concessi a Netflix attraverso licenze. Progressivamente, però, la società comincia a investire direttamente in contenuti destinati alla propria piattaforma.

Nel 2012 debutta Lilyhammer, che Netflix considera l'inizio della propria storia di serie originali. La stagione viene resa disponibile interamente, invece di pubblicare un episodio alla settimana. Questo approccio contribuisce allo sviluppo del comportamento che diventerà comunemente conosciuto come binge watching (Sarandos, 2022).
Nel 2013 House of Cards consolida ulteriormente la trasformazione: Netflix non è più soltanto un'infrastruttura attraverso cui guardare contenuti prodotti da altri, ma diventa direttamente un soggetto centrale nella produzione e distribuzione audiovisiva. La serie diventa inoltre la prima serie originale online a ottenere importanti nomination ai Primetime Emmy (Netflix, 2018).

Questa trasformazione influenza anche il modello tecnologico.
Netflix non sta quindi più semplicemente trasferendo file video: sta costruendo una vera e propria pipeline industriale del contenuto digitale.

Parallelamente Netflix espande progressivamente il proprio mercato.
Dopo gli Stati Uniti, il servizio di streaming arriva prima in Canada e successivamente in America Latina, Europa, Australia, Nuova Zelanda e Giappone.
All'inizio del 2016 Netflix era già disponibile in circa 60 paesi.
Il 6 gennaio 2016 arriva un ulteriore salto di scala: Netflix annuncia contemporaneamente l'espansione in oltre 130 nuovi paesi, trasformando il servizio in una piattaforma sostanzialmente globale (Izrailevsky et al., 2016).
Non è casuale che questa espansione avvenga nello stesso periodo in cui viene completata la migrazione al cloud.
La crescita globale richiede infatti:

- Backend distribuito e scalabile → AWS.
- Distribuzione locale del video → Open Connect.
- Scoperta dei contenuti → recommendation system globale.

Nel 2016 Netflix descrive anche la creazione di un sistema globale di raccomandazione capace di identificare comunità di utenti con gusti simili indipendentemente dal paese di appartenenza (Gomez-Uribe, 2016).
L'idea è significativa: invece di considerare esclusivamente “gli utenti italiani” o “gli utenti statunitensi”, il sistema può individuare persone con comportamenti simili distribuite in nazioni differenti.
La globalizzazione riguarda quindi contemporaneamente rete, catalogo e dati.

Per molti anni Netflix mantiene un modello estremamente riconoscibile:
abbonamento mensile → accesso illimitato → assenza di pubblicità.

Anche questa caratteristica influenza l'esperienza del prodotto: Netflix può progettare la riproduzione senza dover inserire interruzioni pubblicitarie all'interno del flusso.

Nel 2021 l'azienda inizia però una nuova fase di diversificazione introducendo Netflix Games, inizialmente sui dispositivi mobili (Verdu, 2021).
I videogiochi vengono inclusi nell'abbonamento e rappresentano un primo tentativo di ampliare Netflix da servizio specificamente dedicato a film e serie a piattaforma più generale di intrattenimento digitale.
La trasformazione più rilevante per il modello di business arriva tuttavia nel 2022, quando Netflix introduce un piano di abbonamento supportato dalla pubblicità, inizialmente in dodici paesi, tra cui l'Italia (Peters, 2022).
Netflix affianca così allo SVOD senza pubblicità un abbonamento che combina sottoscrizione e ricavi pubblicitari. È un modello ibrido: il piano con annunci rimane a pagamento e non equivale a un servizio AVOD gratuito.

L'architettura deve ora gestire anche nuovi elementi:
Utente → piano → autorizzazioni → contenuto → eventuale pubblicità → playback.

La pubblicità non è quindi soltanto una modifica commerciale: introduce nuove esigenze relative a selezione degli annunci, misurazione delle impression, targeting consentito, inserimento degli spot e reporting verso gli inserzionisti.

Nel 2023 avviene un passaggio simbolico importante.
Dopo venticinque anni, Netflix spedisce il proprio ultimo DVD negli Stati Uniti e chiude definitivamente il servizio che aveva dato origine all'azienda (Sarandos, 2023).

Nello stesso periodo Netflix inizia però una trasformazione opposta sul fronte dello streaming.
Il *VOD* tradizionale presenta un vantaggio enorme dal punto di vista tecnologico: il contenuto esiste prima della richiesta dell'utente.
Un film può quindi essere analizzato, codificato, verificato e distribuito sui server Open Connect anche molte ore o giorni prima della visione.
Con il live streaming, invece, questo vantaggio scompare.

Nel marzo 2023, con lo speciale di Chris Rock Selective Outrage, Netflix realizza il proprio primo evento globale trasmesso direttamente in live streaming (Fedorov et al., 2025).
Nel novembre dello stesso anno Netflix trasmette anche il proprio primo evento sportivo live, la Netflix Cup (Netflix, 2023b).
Negli anni successivi l'offerta live aumenta ulteriormente.
L'accordo annunciato nel 2024 con WWE porta progressivamente programmi come Raw, SmackDown, NXT e i principali eventi WWE all'interno di Netflix in numerosi mercati; dall'aprile 2026 l'offerta WWE live viene estesa anche all'Italia (Netflix, 2026a).

Osservata nel suo insieme, l'evoluzione di Netflix può quindi essere rappresentata come una successione di trasformazioni:

```
1998
DVD via Internet
↓
1999
Subscription
↓
2000
Personalization
↓
2007
Streaming
↓
2008-2016
Cloud / AWS
↓
2011-2012
Open Connect
↓
2012-2013
Original Content
↓
2016
Global Platform
↓
2021
Games
↓
2022
Advertising
↓
2023+
Live Streaming
```

Modello di business

Il modello di business di Netflix nasce da un principio relativamente semplice: l’utente non acquista il singolo film o episodio, ma paga per accedere a un catalogo di contenuti per un determinato periodo di tempo.
Questa impostazione colloca Netflix storicamente nel modello *SVOD*, Subscription Video on Demand.

Con il tempo, tuttavia, il modello è diventato più articolato.

Netflix continua a ricavare la maggior parte dei propri introiti dagli abbonamenti, ma oggi affianca alla sottoscrizione anche pubblicità, partnership commerciali, utenti extra, prodotti ed esperienze collegate ai propri contenuti (Netflix, Inc., 2026a).
L'azienda stessa dichiara che il proprio obiettivo principale non è semplicemente aumentare il numero degli abbonati, ma massimizzare i ricavi nel lungo periodo, perché maggiori ricavi consentono di investire in contenuti, tecnologia e crescita del servizio (Netflix, Inc., n.d.).

Nel 2025 Netflix ha registrato circa 45,2 miliardi di dollari di ricavi, con un margine operativo del 29,5%. Nello stesso anno ha superato i 325 milioni di abbonamenti paganti (Netflix, Inc., 2026b).
La pubblicità ha generato oltre 1,5 miliardi di dollari, crescendo rapidamente ma rappresentando ancora una quota minoritaria rispetto agli abbonamenti (Netflix, Inc., 2026b).

Il nucleo del modello rimane l'abbonamento ricorrente.
Questo modello può essere rappresentato come:
Users → Subscription → Netflix

L'utente paga periodicamente Netflix e, in cambio, ottiene il diritto di accedere al catalogo disponibile nel proprio mercato.
Non acquista quindi la proprietà dei film o delle serie e non paga ogni singola riproduzione.
Il vantaggio economico per Netflix è la presenza di un flusso di ricavi ricorrente.
Ogni abbonamento genera ricavi fintanto che l'utente rimane all'interno del servizio.
Di conseguenza, non è sufficiente convincere una persona a iscriversi una volta.
È necessario convincerla a continuare a percepire il servizio come sufficientemente utile da mantenere l'abbonamento.
Il modello dipende quindi fortemente da due fenomeni:
Acquisition → Retention
Dove:

- Acquisition: Portare nuovi utenti all'interno del servizio.
- Retention: Convincere gli utenti esistenti a rimanere abbonati.

Netflix non propone necessariamente a tutti gli utenti lo stesso servizio allo stesso prezzo.
Utilizza invece una strategia di price segmentation: diversi livelli di prezzo corrispondono a differenti caratteristiche del servizio.

A seconda del mercato possono cambiare elementi come:

- Presenza o assenza della pubblicità.
- Numero di riproduzioni simultanee.
- Qualità massima del video.
- Numero di dispositivi utilizzabili per i download.
- Possibilità di aggiungere utenti extra.
- Altre funzionalità del servizio.

Netflix dichiara esplicitamente di utilizzare differenti fasce di prezzo e di sperimentare periodicamente combinazioni di piani e prezzi per comprendere la domanda degli utenti (Netflix, Inc., n.d.).
Il principio economico è quello della discriminazione di prezzo attraverso la differenziazione del prodotto.

Due persone possono attribuire un valore differente al servizio.

- Un utente potrebbe desiderare semplicemente Netflix al prezzo più basso possibile.
- Un altro potrebbe essere disposto a pagare di più perché desidera maggiore qualità, nessuna pubblicità e più dispositivi simultanei.

Netflix può quindi trasformare queste differenti disponibilità a pagare in differenti offerte.
In forma semplificata:

```
Differenti necessità
↓
Differenti piani
↓
Differenti prezzi
↓
Ottimizzazione dei ricavi
```

L'obiettivo non è necessariamente spingere tutti verso il piano più costoso, ma trovare una combinazione che massimizzi complessivamente: numero di utenti × ricavo medio × permanenza nel servizio.

Dal 2022 Netflix ha introdotto un secondo meccanismo di monetizzazione: la pubblicità.
Nel piano con pubblicità avvengono contemporaneamente due scambi economici:

- User → Subscription fee → Netflix
- Advertiser → Advertising spend → Netflix

Netflix offre quindi agli utenti un prezzo d'ingresso inferiore e, contemporaneamente, monetizza una parte della loro attenzione vendendo spazi pubblicitari agli inserzionisti.

Il modello diventa:

```
User
↓ subscription
Netflix
↑ advertising revenue
Advertiser
```

In cambio del denaro investito, l'inserzionista ottiene accesso agli spazi pubblicitari disponibili.
Quindi lo stesso utente genera valore sia pagando l'abbonamento sia generando opportunità pubblicitarie.

Nel maggio 2026 Netflix dichiarava più di 250 milioni di Monthly Active Viewers raggiunti dalla pubblicità a livello globale. La metrica moltiplica gli abbonati che hanno visto almeno un minuto di annunci nel mese per una stima delle persone nel nucleo domestico ricavata da ricerche interne: è quindi una stima di pubblico, non il numero di abbonamenti né un conteggio individuale verificato di tutti gli spettatori (Netflix, 2026b; Reinhard, 2025).

Il business pubblicitario è ormai sufficientemente importante da aver portato Netflix a sviluppare una propria infrastruttura, la Netflix Ads Suite, attraverso cui gli inserzionisti possono acquistare, targettizzare e misurare campagne (Netflix, 2026b).

I ricavi ottenuti dagli utenti e dagli inserzionisti devono però alimentare il principale elemento che rende possibile il modello: il catalogo.

Netflix deve continuamente acquisire o produrre contenuti che convincano gli utenti a: iscriversi → guardare → tornare → rimanere abbonati.

Il catalogo può essere costruito principalmente attraverso due modalità.

La prima consiste nell'acquisizione di licenze. Netflix paga un produttore, uno studio o un altro titolare dei diritti per poter distribuire un contenuto secondo determinate condizioni.

Una licenza può stabilire:
- Territori nei quali il contenuto può essere distribuito.
- Durata della disponibilità.
- Lingue disponibili.
- Tipologia di utilizzo.
- Eventuali periodi di esclusiva.
- Compatibilità con determinati modelli commerciali, incluso il piano con pubblicità.

Questo spiega perché il catalogo Netflix non sia necessariamente identico in tutti i paesi e perché alcuni titoli possano scomparire dopo un certo periodo.

La seconda modalità consiste nell'investimento in Netflix Originals, attraverso produzione diretta, commissione, co-produzione o acquisizione di diritti esclusivi (Netflix, Inc., n.d.).
In questo caso Netflix cerca di controllare maggiormente la relazione tra contenuto e piattaforma.
Dal punto di vista strategico, un contenuto esclusivo presenta un vantaggio evidente:
Se l'utente vuole vedere quel contenuto, deve utilizzare Netflix.
Un titolo non esclusivo può invece essere disponibile anche presso altri servizi.

Il contenuto diventa quindi contemporaneamente:
- Prodotto, perché è ciò che l'utente guarda.
- Strumento di acquisizione, perché può attirare nuovi abbonati.
- Strumento di retention, perché può convincerli a rimanere.
- Asset competitivo, perché può differenziare Netflix dai concorrenti.

Il ciclo economico fondamentale diventa quindi:

```
Revenue
↓
Content investment
↓
Attractive catalog
↓
```

```
Viewing
↓
Acquisition + Retention
↓
Revenue
```

Questa relazione è probabilmente il cuore del modello Netflix.
Netflix non guadagna direttamente di più ogni volta che un abbonato guarda un episodio, almeno nel modello SVOD tradizionale.
Questo produce una situazione interessante.
Una volta pagato l'abbonamento: guardare dieci film non genera necessariamente dieci volte il ricavo di guardarne uno.
Perché allora Netflix vuole che l'utente utilizzi molto il servizio?
Perché l'utilizzo può aumentare il valore percepito dell'abbonamento.

Se una persona paga Netflix ma non trova nulla da vedere per diverse settimane, aumenta la probabilità che si chieda:
"Perché sto continuando a pagare?"
Se invece trova continuamente contenuti interessanti, il costo mensile appare più giustificato.

L'engagement diventa quindi uno strumento soprattutto di retention (Netflix, Inc., n.d.).
Possiamo rappresentarlo così:

```
Relevant content
↓
Viewing
↓
Perceived value
↓
Retention
↓
Recurring revenue
```

Nel piano con pubblicità esiste anche una relazione più diretta:

```
Viewing time
↓
Advertising opportunities
↓
Ad impressions
↓
Advertising revenue
```

L'engagement assume quindi un doppio valore:

- *SVOD* → retention.
- Ad-supported plan → fidelizzazione + advertising inventory.

Questo aiuta anche a capire perché Netflix investa così tanto nella personalizzazione.
Raccomandazioni e personalizzazione come strumenti economici
Il recommendation system non serve soltanto a migliorare l'esperienza utente.
Ha anche una funzione economica.

Il catalogo Netflix contiene una quantità di contenuti troppo grande perché l'utente possa esplorarlo interamente. Se l'utente non riesce rapidamente a trovare qualcosa di interessante, la dimensione del catalogo perde parte del proprio valore.

Il sistema di raccomandazione prova quindi a trasformare:
Catalogo disponibile -> Catalogo percepito come rilevante per quel particolare utente.

Netflix può utilizzare informazioni come cronologia di visione, valutazioni, interazioni e caratteristiche del contesto per personalizzare ciò che viene presentato (Netflix, n.d.-c).

Il meccanismo può essere sintetizzato in questo modo:

```
Viewing
↓
Behavioral data
↓
Recommendation
↓
Relevant content
↓
More viewing
```

Da un punto di vista economico:
Personalization → Engagement → Fidelizzazione → Revenue

Naturalmente non significa che ogni raccomandazione produca direttamente denaro.
La relazione è indiretta: l'obiettivo è aumentare la probabilità che l'utente continui a trovare valore nel servizio.

Per molti anni la condivisione delle credenziali tra nuclei domestici è stata diffusa, anche se non va confusa con un'autorizzazione contrattuale generale.
Dal punto di vista economico, questo utilizzo poteva favorire la diffusione del servizio, ma comportava anche che più nuclei domestici utilizzassero un unico abbonamento: la prima relazione è un'interpretazione plausibile, non una causalità dimostrata qui.
Dal punto di vista dell'utente il valore dell'abbonamento aumentava, ma dal punto di vista di Netflix molte persone utilizzavano il servizio senza generare un ricavo aggiuntivo.
Con l'estensione delle iniziative di paid sharing nel 2023, Netflix ha rafforzato l'applicazione delle regole sul nucleo domestico. L'account è destinato alle persone dello stesso nucleo; le persone esterne devono creare un proprio account oppure, nei piani e mercati compatibili, essere aggiunte come utenti extra a pagamento (Netflix, 2023a; Netflix, n.d.-g).
Anche questa scelta può essere letta attraverso il modello di business.

- Prima: 1 subscription → multiple households

- Dopo: accesso per nucleo domestico → abbonamento proprio o utente extra, quando consentito.

L'obiettivo è quindi trasformare una parte dell'utilizzo precedentemente non monetizzato in:
- Nuovi abbonamenti.
- Ricavi aggiuntivi dall'account esistente.

È interessante notare che Netflix ha accompagnato questa transizione con la funzione di Profile Transfer: un utente che utilizzava il profilo appartenente a un altro account può trasferire cronologia, lista e raccomandazioni nel proprio nuovo account (Netflix, n.d.-f).

Questa funzionalità può anche ridurre gli ostacoli al passaggio da utente condiviso a cliente pagante: è un'inferenza basata sulla possibilità di conservare cronologia e preferenze, non una misura dell'effetto sulle conversioni.

Netflix non vende necessariamente tutti gli abbonamenti direttamente.
Il servizio può essere incluso in offerte di:
- Operatori telefonici.
- Provider Internet.
- Operatori televisivi.
- Produttori o piattaforme partner.
- Altri servizi in abbonamento.

In questi casi l'utente può pagare Netflix indirettamente attraverso un pacchetto più ampio.
Per esempio: User → Telecom operator → Netflix

Nel proprio bilancio Netflix specifica che, quando il prezzo pagato dall'utente viene determinato da un partner e non esiste un prezzo autonomo Netflix separato, viene contabilizzato come ricavo l'importo netto ricevuto dal partner (Netflix, Inc., 2026a).

Il bundle offre vantaggi a entrambe le parti.
L'operatore può rendere più interessante la propria offerta:
"Internet + mobile + Netflix."

Netflix può invece raggiungere utenti che magari non avrebbero sottoscritto autonomamente il servizio.
È quindi una strategia di distribuzione commerciale, non molto diversa concettualmente dalla distribuzione di un prodotto attraverso differenti canali di vendita.

Netflix sta inoltre cercando di aumentare il valore dell'abbonamento attraverso contenuti che vanno oltre film e serie tradizionali.
Ne fanno parte:
- Giochi.
- Live entertainment.
- Sport ed eventi.
- Video podcast.
- Altre forme di contenuto.

Non necessariamente ognuna di queste categorie deve generare direttamente un nuovo ricavo.
Un gioco incluso nell'abbonamento, per esempio, può essere economicamente utile anche se non viene venduto separatamente.
Può contribuire alla relazione:
More value → more engagement → more retention.

Lo stesso vale per alcuni eventi live.

- Un evento molto importante può funzionare come evento di acquisizione: una persona si abbona perché vuole vedere quell'evento.
- Può contemporaneamente funzionare come un evento di fidelizzazione: un abbonato decide di non cancellare perché sa che arriveranno nuovi eventi.

Quindi l'espansione del catalogo non deve essere letta soltanto come:
"Netflix vuole offrire più cose."

È più precisamente una strategia volta ad aumentare il numero di occasioni nelle quali Netflix può diventare rilevante per l'utente.

Una delle innovazioni più riconoscibili introdotte da Netflix non riguarda direttamente la tecnologia dello streaming, ma il modo in cui una serie viene pubblicata.
La televisione tradizionale era costruita principalmente intorno a una programmazione lineare: 1 episodio → attesa di una settimana → episodio successivo.

Lo spettatore doveva quindi adattarsi al calendario stabilito dall'emittente.
Con lo streaming questa necessità tecnica scompare. Il contenuto può essere disponibile in qualsiasi momento e l'utente può decidere autonomamente quando guardarlo.

Netflix sfrutta questa possibilità fin dalle prime produzioni originali. Nel 2013 pubblica l'intera stagione di House of Cards contemporaneamente, rendendo immediatamente disponibili tutti gli episodi. Netflix stessa ha successivamente indicato questo modello come una rottura rispetto alla programmazione televisiva tradizionale.
Si consolida così il modello del full-season drop, cioè la pubblicazione simultanea di un'intera stagione.

- Lo spettatore non deve più adattarsi al palinsesto.
- È la piattaforma ad adattarsi allo spettatore.

La scelta era coerente con uno dei principi fondamentali del prodotto Netflix: offrire all'utente il controllo sulla propria esperienza di visione.
Netflix sottolineava infatti come, eliminando il vincolo della programmazione settimanale, ogni spettatore potesse costruire autonomamente il proprio ritmo di consumo (Sarandos, 2022).
Esisteva inoltre una motivazione narrativa.
Nella televisione tradizionale il primo episodio, il pilot, doveva convincere immediatamente il pubblico a tornare la settimana successiva.
Netflix analizzò invece il comportamento degli utenti e osservò che, per numerose serie, il punto nel quale gli spettatori diventavano realmente coinvolti si trovava spesso dopo diversi episodi.

Nel 2015 Netflix definì hook point l'episodio dopo il quale almeno il 70% degli spettatori che lo avevano visto completava la prima stagione; nel campione analizzato non coincideva con il pilot. È una misura osservazionale di completamento: non dimostra che pubblicare tutta la stagione insieme causi maggiore fidelizzazione (Netflix, 2015).
La disponibilità immediata degli episodi può quindi consentire allo spettatore di proseguire senza l'attesa settimanale; il suo effetto su coinvolgimento e permanenza nell'abbonamento va verificato separatamente.
Questa struttura favorisce il binge watching, cioè la visione consecutiva di più episodi in un intervallo relativamente breve. Il comportamento era possibile anche con DVD e altri supporti: lo streaming lo rende più accessibile, ma non lo inventa (Sarandos, 2023).
Nel 2016 Netflix riportò una mediana di cinque giorni per completare la prima stagione delle serie studiate. Il dato riguardava soltanto gli account che avevano completato la stagione, non tutti gli utenti che l'avevano iniziata, e variava in funzione delle serie e dei generi (Netflix, 2016).

Il modello presenta però anche alcuni limiti dal punto di vista commerciale.
Se tutti gli episodi vengono pubblicati contemporaneamente, un utente può vedere tutta la stagione in pochissimo tempo.

Se quell'utente si era abbonato soprattutto per quella serie, dopo averla terminata deve trovare subito qualcos'altro di interessante, altrimenti può aumentare la probabilità che interrompa l'abbonamento al termine del mese.
La serie può inoltre concentrare gran parte dell'attenzione mediatica nelle settimane immediatamente successive al lancio.
Dal punto di vista economico questo può ridurre la durata durante la quale un singolo titolo contribuisce attivamente a mantenere l'attenzione degli utenti.

Un modello settimanale produce invece una dinamica differente:

```
Episode 1
↓
discussion
↓
Episode 2
↓
discussion
↓
Episode 3
↓
discussion
```

La serie rimane potenzialmente presente nella conversazione pubblica per un periodo molto più lungo.
Questo può contribuire a:
engagement → recurring visits → conversation → retention

Nei servizi supportati dalla pubblicità, il modello settimanale può inoltre distribuire le opportunità pubblicitarie lungo un periodo più esteso.

Non significa però che il modello settimanale sia necessariamente migliore. Impone nuovamente allo spettatore un ritmo stabilito dalla piattaforma e può interrompere la continuità narrativa.

Le piattaforme possono quindi adottare strategie intermedie. La tabella seguente confronta possibili effetti progettuali e commerciali: sono ipotesi e tendenze plausibili, non risultati causali dimostrati per ogni serie o pubblico.

Modello	**Descrizione**	**Punti positivi**	**Punti negativi**
Full-season drop	Tutti gli episodi della stagione vengono pubblicati nello stesso momento. Ep. 1–8 → same day	Massima libertà per l'utente. Favorisce il binge watching. Permette di seguire la storia senza interruzioni. Può generare un forte impatto iniziale e attirare rapidamente nuovi abbonati.	L'interesse può concentrarsi in pochi giorni. La conversazione online tende a esaurirsi più rapidamente. L'utente può completare subito la serie e avere meno motivi per tornare periodicamente sulla piattaforma. Maggiore rischio di spoiler.
Weekly release	Viene pubblicato un episodio alla volta, normalmente ogni settimana. Ep. 1 → Ep. 2 → Ep. 3 → Ep. 4	Mantiene l'attenzione sulla serie per molte settimane. Favorisce discussione, teorie e passaparola. Incentiva il ritorno periodico sulla piattaforma. Può contribuire alla fidelizzazione e distribuire nel tempo le opportunità pubblicitarie.	Riduce la libertà dello spettatore. L'attesa può interrompere la continuità narrativa. Alcuni utenti possono preferire aspettare la conclusione della stagione prima di iniziarla.
Batch release	Gli episodi vengono pubblicati in gruppi successivi. Ep. 1–4 → Ep. 5–7 → Ep. 8–10	Combina binge watching e ritorno periodico. Ogni blocco può generare un nuovo momento di attenzione. Prolunga la durata commerciale della stagione senza imporre	L'interruzione tra i blocchi può risultare artificiale. È necessario scegliere attentamente dove dividere la narrazione. Può creare confusione se il calendario non viene comunicato

		necessariamente un episodio alla settimana.	chiaramente.
Split season / Volumes	La stagione viene divisa in due o più grandi parti pubblicate a distanza di alcune settimane o mesi. Part 1 → wait → Part 2	Mantiene il binge watching all'interno di ciascun volume. Crea più eventi di lancio e prolunga l'attenzione mediatica. Può sostenere fidelizzazione e marketing per un periodo più lungo.	Una lunga pausa può spezzare il coinvolgimento. La divisione può sembrare motivata più commercialmente che narrativamente. Il pubblico può dimenticare dettagli della prima parte.
Multi-episode premiere + weekly	La serie debutta con due o più episodi contemporaneamente e successivamente passa a una pubblicazione settimanale. Ep. 1–3 → Ep. 4 → Ep. 5 → Ep. 6	Gli episodi iniziali permettono allo spettatore di entrare rapidamente nella storia e superare la fase introduttiva. Successivamente il ritmo settimanale prolunga engagement e conversazione. È un buon compromesso tra binge e appuntamento periodico.	Dopo il lancio iniziale il passaggio al ritmo settimanale può risultare frustrante. Richiede episodi iniziali sufficientemente forti da creare interesse. Rimane meno libero rispetto al full-season drop.

Il calendario di pubblicazione diventa quindi una vera variabile del modello di business.
Ogni strategia produce un equilibrio differente tra:

- User freedom
- Bingeability
- Engagement duration
- Marketing visibility
- Retention
- Advertising opportunities

È importante quindi non interpretare l'evoluzione di Netflix come un semplice ritorno alla televisione tradizionale.
La differenza fondamentale rimane che la piattaforma può scegliere il modello di pubblicazione più adatto a ogni prodotto.

- Una serie limitata può essere pubblicata interamente.
- Una grande produzione può essere divisa in due parti.
- Un reality può essere distribuito in batch settimanali.
- Una produzione legata a una programmazione televisiva o internazionale può seguire un calendario settimanale.

Lo streaming ha quindi trasformato il calendario da vincolo del modello distributivo e del palinsesto a variabile progettuale e commerciale.
Nella televisione lineare l'episodio settimanale era in larga parte conseguenza del palinsesto.
Nello streaming la domanda diventa invece: "Quale ritmo di pubblicazione produce l'esperienza e il comportamento più adatti a questo contenuto e agli obiettivi della piattaforma?"

Il modello Netflix può infine essere rappresentato come un ciclo.

```
Sottoscrizione + Pubblicità
↓
Ricavi
↓
Contenuto + Tecnologia + Infrastruttura
↓
Catalogo e User Experience
↓
Visualizzazione e Coinvolgimento
↓
Acquisizione + Fidelizzazione
↓
Sottoscrizione + Pubblicità
```

La piattaforma deve quindi mantenere un equilibrio continuo tra ricavi, costo dei contenuti, tecnologia, acquisizione degli utenti e redditività.

- Il contenuto attira e trattiene il pubblico.
- La tecnologia permette di distribuirlo efficientemente.
- La personalizzazione rende il catalogo più facilmente utilizzabile.
- Il coinvolgimento aumenta il valore percepito dell'abbonamento e, nel piano con pubblicità, produce anche spazi pubblicitari.
- I ricavi così ottenuti possono essere nuovamente investiti in contenuti e piattaforma.

In questa prospettiva Netflix non vende semplicemente film e serie. Vende accesso continuativo a un sistema di intrattenimento personalizzato, cercando di massimizzare il valore economico generato dalla relazione di lungo periodo con ciascun nucleo domestico.

Architettura

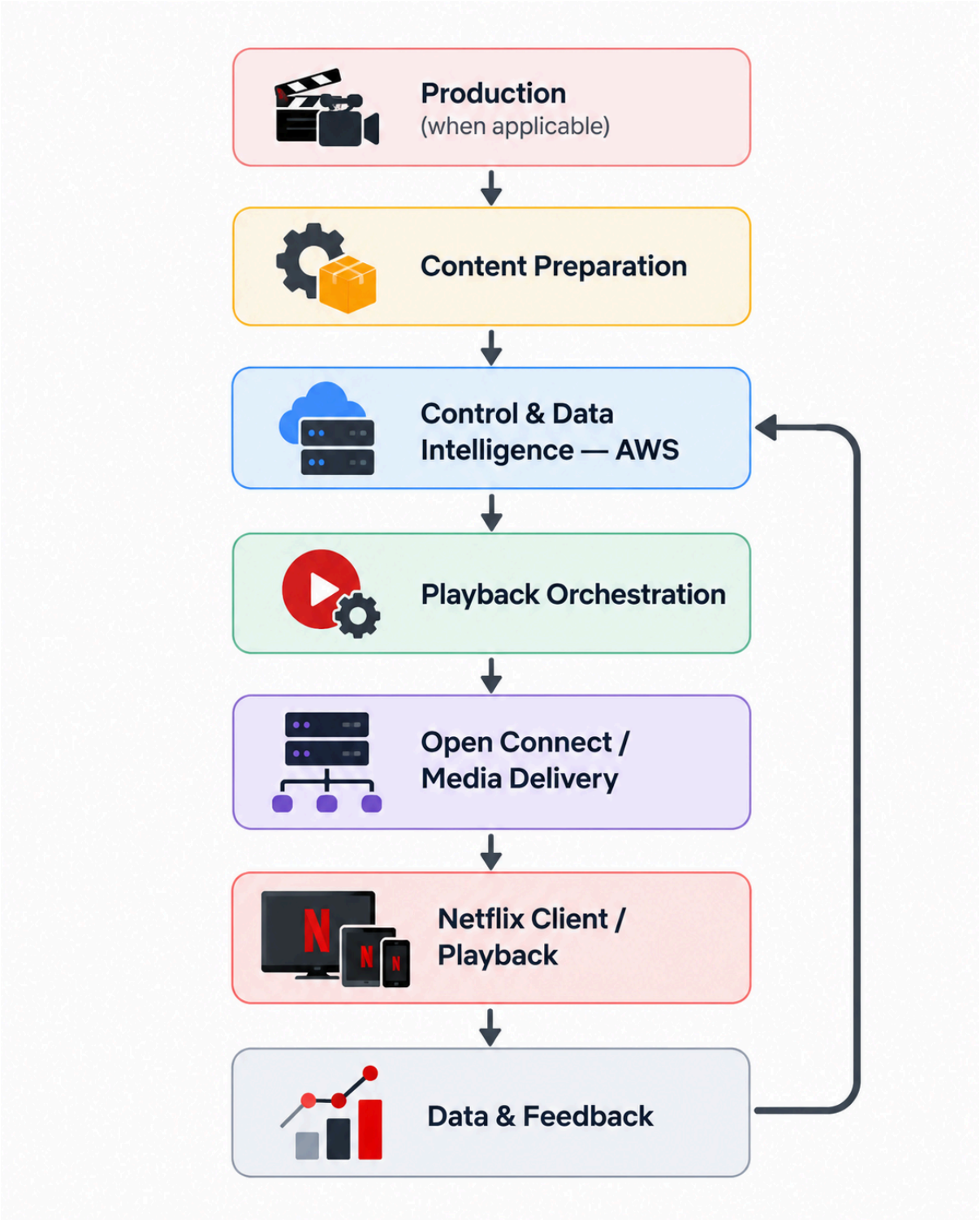
L'architettura di Netflix è progettata per risolvere un problema apparentemente semplice ma, su scala globale, estremamente complesso: permettere a milioni di utenti di scegliere un contenuto e iniziare a riprodurlo rapidamente, con una qualità adeguata al dispositivo e alla connessione disponibili, evitando per quanto possibile interruzioni e buffering.

Per raggiungere questo obiettivo Netflix non utilizza un unico grande sistema centralizzato. L'infrastruttura è invece suddivisa in componenti e fasi con responsabilità differenti, che collaborano durante l'intero ciclo di vita del contenuto.

- Production: quando il contenuto viene prodotto direttamente da Netflix o attraverso workflow produttivi gestiti nel suo ecosistema, la piattaforma può intervenire già durante la produzione e la post-produzione, gestendo materiali originali, metadati e asset attraverso strumenti dedicati fino alla realizzazione del master finale.
- Content Preparation: una volta disponibile il master, il contenuto viene preparato per lo streaming attraverso operazioni come encoding e transcoding, packaging, cifratura e archiviazione, producendo le differenti versioni necessarie ai dispositivi e alle condizioni di rete supportate.
- Control Plane: una parte dell'infrastruttura, eseguita prevalentemente nel cloud *AWS*, comprende i servizi backend che gestiscono account, profili, catalogo, autorizzazioni, licenze DRM, raccomandazioni, personalizzazione e configurazione della sessione di playback.
- Media Plane: la distribuzione effettiva dei contenuti audiovisivi viene affidata principalmente a Open Connect, la CDN proprietaria di Netflix, che utilizza server edge distribuiti geograficamente per avvicinare i contenuti agli utenti.
- Client Playback: il client riceve i contenuti dall'infrastruttura di distribuzione, seleziona dinamicamente la rappresentazione più appropriata, gestisce il buffer e il *DRM*, decifra e decodifica i flussi audio e video e ne effettua la riproduzione.
- Data & Feedback: durante l'intero processo vengono prodotti eventi e dati di telemetria che ritornano alla piattaforma e alimentano analytics, sistemi di Machine Learning, raccomandazioni, sperimentazione, monitoraggio e ottimizzazione dell'infrastruttura.

L'architettura può quindi essere interpretata come un grande ciclo:

```
Production (quando applicabile)
↓
Content Preparation
↓
Control & Data Intelligence — AWS
↓
Playback Orchestration
↓
Open Connect / Media Delivery
↓
Netflix Client / Playback
↓
Data & Feedback
↺
Control & Data Intelligence
```



Nel caso di Netflix il ciclo di vita di un contenuto può iniziare ancora prima della sua preparazione per lo streaming.
Quando un film o una serie viene prodotto direttamente da Netflix, o attraverso un workflow produttivo gestito all'interno del suo ecosistema, la piattaforma può infatti intervenire già durante le fasi di produzione e post-produzione.
In questa fase non esiste ancora necessariamente un master audiovisivo definitivo. Esistono invece numerosi materiali di lavorazione: riprese provenienti dalle camere, registrazioni

audio, clip, differenti versioni delle scene, materiali destinati agli effetti visivi e altri asset necessari alla produzione.

Netflix mette a disposizione strumenti dedicati alla gestione di questi materiali.
Un ruolo importante è svolto da **Content Hub**, utilizzato come ambiente per l'acquisizione, l'organizzazione e lo scambio degli asset di produzione.
La sezione dedicata al Footage Management, per esempio, permette ai team di produzione, post-produzione e tecnici di verificare i materiali acquisiti, distinguendo elementi come camera rolls, sound rolls, clip e altri tipi di upload, oltre a fornire informazioni tecniche e report relativi all'ingest (Netflix, n.d.-b).

Il processo può quindi iniziare con il trasferimento dei materiali originali prodotti sul set:

```
Riprese / Registrazioni audio
↓
Materiali originali di produzione
↓
Footage Ingest
↓
Content Hub
↓
Montaggio / VFX / Post-produzione
↓
Master finale
```

Content Hub viene utilizzato anche all'interno di workflow più specifici.
Nel caso degli effetti visivi, per esempio, materiali e relative informazioni possono essere caricati dai team di produzione o dagli studi VFX, sottoposti a controlli ed elaborazioni e successivamente resi disponibili ai sistemi utilizzati per la revisione delle diverse versioni.
Questa infrastruttura permette quindi di mantenere organizzato il materiale durante la lavorazione, facilitando gli scambi tra produzione, montaggio, effetti visivi e altri reparti fino alla realizzazione della versione definitiva del contenuto.

La fase di Production non deve tuttavia essere considerata obbligatoria per ogni titolo presente nel catalogo. Un contenuto già completato e acquisito da Netflix attraverso accordi di distribuzione può entrare nel processo molto più avanti, attraverso la consegna dei master e degli altri asset necessari. La presenza della fase Production è quindi particolarmente significativa quando Netflix partecipa direttamente al workflow produttivo del contenuto.

Quando Netflix riceve i master, essi non vengono semplicemente copiati sui server e distribuiti direttamente agli utenti: vengono analizzati, controllati e trasformati in numerose versioni differenti.
La necessità nasce dall'enorme varietà dei dispositivi sui quali viene utilizzato Netflix.

- Una Smart TV moderna può supportare risoluzioni 4K, HDR e codec recenti come AV1 o HEVC.
- Un dispositivo meno recente potrebbe invece essere compatibile soltanto con H.264/AVC e risoluzioni inferiori.

- Due utenti dotati dello stesso televisore potrebbero inoltre disporre di connessioni completamente differenti.

Netflix deve quindi preparare una famiglia di rappresentazioni dello stesso contenuto:

```
Master sorgente
↓
Analisi e controllo qualità
↓
Codifica / Transcodifica
↓
Codec multipli
↓
Risoluzioni e bitrate multipli
↓
Audio + Lingue + Sottotitoli
```

Un singolo film non corrisponde quindi a un unico file pronto per lo streaming.
Bisogna piuttosto immaginare:
Video→ molte rappresentazioni → selezione della rappresentazione appropriata durante la riproduzione.

La combinazione tra profili video, codec, risoluzioni, bitrate, tracce audio e sottotitoli produce numerosi asset. In un esempio pubblicato nel 2017, Netflix indicava circa 1.200 file per un episodio di The Crown: è un caso documentato, non un numero fisso per ogni episodio (Netflix Technology Blog, 2017).
Questa moltiplicazione non rappresenta uno spreco: permette alla piattaforma di consegnare a ciascun dispositivo le rappresentazioni più appropriate rispetto allo stato della connessione.

Netflix non applica inoltre la stessa strategia di compressione a tutti i contenuti.
Un film d'animazione caratterizzato da grandi superfici uniformi e movimenti relativamente semplici può essere più facile da comprimere rispetto a una scena notturna ricca di grana, pioggia, fumo e movimenti rapidi.
Utilizzare lo stesso bitrate in entrambi i casi sarebbe inefficiente.
Netflix ha quindi sviluppato il Per-Title Encoding, con il quale la configurazione di codifica può essere adattata alle caratteristiche dello specifico titolo (Aaron et al., 2015).
L'obiettivo non è ottenere il massimo bitrate possibile, ma raggiungere:
qualità percepita desiderata → con la minore quantità di dati necessaria.

Questo è particolarmente importante in una piattaforma globale, perché ridurre la quantità di dati necessaria mantenendo la stessa qualità percepita significa contemporaneamente ridurre:
banda necessaria → traffico *CDN* → rischio di buffering → spazio occupato nelle cache.

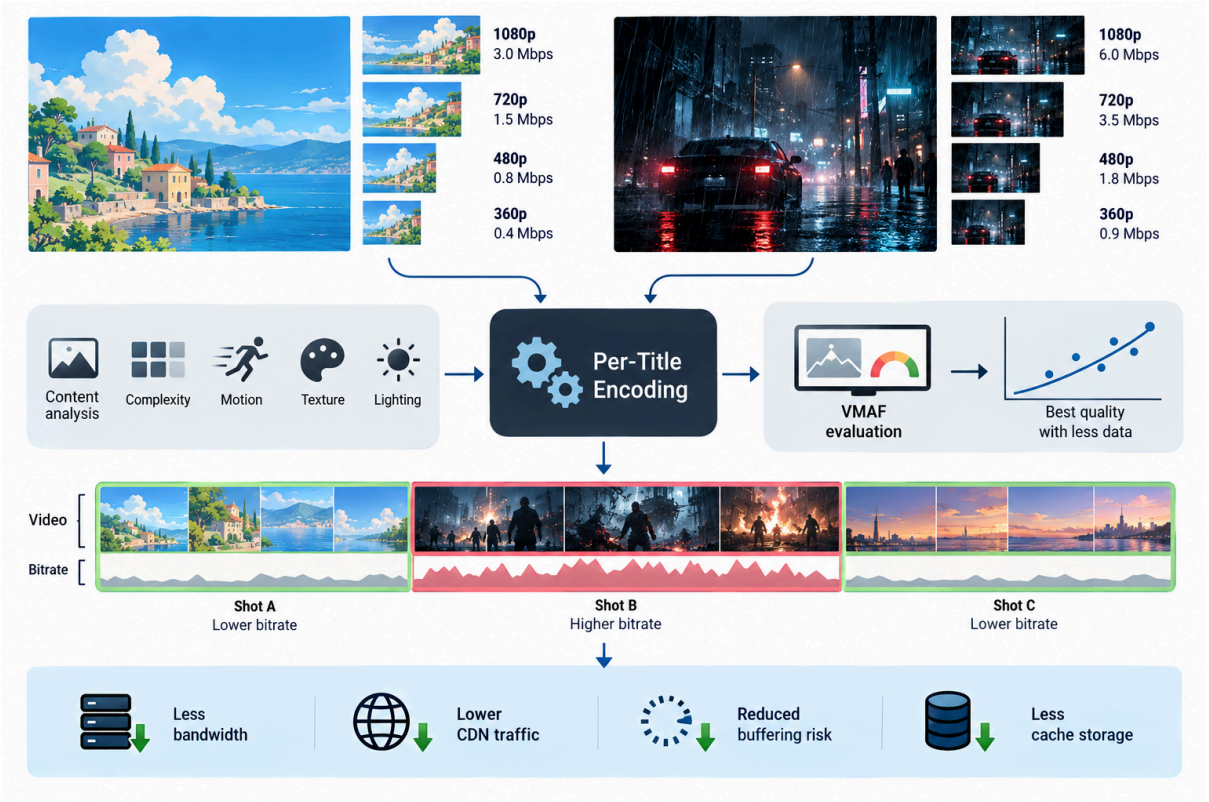
Per valutare queste scelte Netflix ha sviluppato VMAF (Video Multi-Method Assessment Fusion), una metrica che stima la qualità visiva percepita rispetto a un riferimento. Il

punteggio dipende dal modello e dalle condizioni di valutazione e non descrive, da solo, l'intera qualità dell'esperienza (Li et al., 2018).
Il bitrate da solo non è infatti sufficiente per stabilire la qualità di un video.
Un flusso da 5 Mbit/s non è necessariamente migliore di uno da 4 Mbit/s: il risultato dipende dal codec utilizzato, dalla risoluzione, dalla complessità delle immagini e dall'efficienza della compressione.
VMAF cerca quindi di rispondere a una domanda più significativa:
quanto è vicina, dal punto di vista percettivo, questa versione compressa alla qualità di riferimento?

Netflix utilizza queste valutazioni per confrontare differenti configurazioni e individuare quelle che producono il miglior rapporto tra qualità e quantità di dati.
Il principio è stato successivamente esteso fino ad analizzare il contenuto a livello di singolo shot.
Un film non presenta infatti la stessa complessità per tutta la propria durata.
Possiamo avere:

- Shot A — semplice → meno bit.
- Shot B — complesso → più bit.
- Shot C — semplice → meno bit.

Sistemi come il Dynamic Optimizer permettono quindi di distribuire più efficacemente i dati in funzione della complessità visiva delle diverse parti del contenuto (Manohara et al., 2018).



Netflix mantiene inoltre contemporaneamente differenti famiglie di codec. H.264/AVC rimane importante per garantire compatibilità con molti dispositivi, mentre codec come HEVC, VP9

e AV1 consentono una compressione più efficiente sui dispositivi compatibili (Guo et al., 2025; Manohara et al., 2018).
La rappresentazione realmente disponibile per una specifica sessione di riproduzione deriva dall'interazione di più elementi:
contenuto + dispositivo + codec supportati + risoluzione + *HDR* + piano + condizioni della rete.

Una volta preparate tutte queste rappresentazioni, entra in gioco la parte dell'architettura che deve decidere quali di esse siano appropriate per uno specifico utente.
Quando l'utente apre Netflix, il client non inizia immediatamente a scaricare un film.
La maggior parte delle prime operazioni riguarda informazioni e logica applicativa.
Il backend deve sapere chi è l'utente, quale profilo sta utilizzando, quale piano possiede, quali contenuti sono disponibili nella sua regione, quali titoli ha già guardato, quali autorizzazioni possiede e quali contenuti potrebbero risultare interessanti.
Netflix utilizza prevalentemente Amazon Web Services (*AWS*) per questa parte della propria infrastruttura.
Durante la migrazione al cloud l'azienda ha progressivamente sostituito una struttura più monolitica con numerosi servizi distribuiti dedicati a differenti funzioni (Izrailevsky et al., 2016).

In maniera semplificata:

```
Client Netflix
↓
Autenticazione / Account
↓
Profilo / Piano
↓
Catalogo / Disponibilità
↓
Raccomandazioni / Ricerca
↓
Servizi di playback
```

Questa parte può essere interpretata come il control plane.

Una delle funzioni più importanti di questo livello riguarda la personalizzazione.
Ogni interazione dell'utente può infatti generare segnali utili alla piattaforma: contenuti riprodotti, episodi completati o abbandonati, titoli aggiunti alla propria lista, valutazioni, ricerche, elementi mostrati nella homepage e interazioni con l'interfaccia.

Netflix non "sa" che un determinato utente apprezzerà necessariamente un video.
Il sistema cerca piuttosto di stimare quanto quel contenuto possa essere rilevante per quel particolare profilo.
Il problema può quindi essere espresso come:
User + Context + Content → Prediction

Una singola azione dice relativamente poco.

Guardare una volta un film di fantascienza non significa necessariamente essere appassionati del genere.
I sistemi di recommendation acquistano valore osservando numerosi segnali nel tempo e cercando regolarità nelle interazioni.
Un profilo potrebbe, per esempio, guardare frequentemente thriller, completare molte serie investigative, valutare positivamente alcuni titoli e interagire spesso con contenuti che presentano determinate caratteristiche.
Il sistema può utilizzare queste informazioni per stimare una maggiore affinità verso altri contenuti.
Si tratta di costruire rappresentazioni numeriche e modelli capaci di individuare relazioni statistiche tra utenti, contenuti e comportamenti.

Netflix può inoltre utilizzare informazioni aggregate relative a utenti con pattern simili.
Concettualmente:

```
All'utente A piacciono X, Y, Z
All'utente B piacciono X, Y, Z
↓
Modelli di preferenza simili
↓
All'utente A piace anche W
↓
W potrebbe essere rilevante per l'utente B
```

Questo richiama i principi del filtraggio collaborativo, anche se i sistemi reali combinano molte tecniche differenti (Netflix, n.d.-c).
Il filtraggio collaborativo (collaborative filtering, CF) utilizza regolarità nelle interazioni tra utenti e contenuti per formulare raccomandazioni; il confronto fra utenti simili è una possibile rappresentazione didattica di questo principio.

Il recommendation system non deve necessariamente confrontare nello stesso modo ogni elemento dell'intero catalogo ogni volta che viene costruita una pagina.
Una rappresentazione concettuale utile divide il processo in due fasi.

La prima è la candidate generation:
Quali contenuti potrebbero essere plausibilmente interessanti per questo profilo?
Il sistema riduce così l'enorme catalogo a un insieme di candidati.

Successivamente interviene il ranking:
In quale ordine dovrebbero essere presentati?
Per esempio, con valori puramente illustrativi e non ricavati da Netflix, un modello può assegnare ai candidati uno score di rilevanza stimata:

- Candidate A → 0,92.
- Candidate B → 0,81.
- Candidate C → 0,63.

Il sistema può quindi ordinare: A → B → C.
Questa graduatoria non rappresenta la qualità assoluta delle opere.

Significa soltanto che per questo profilo e in questo contesto, A viene considerato più rilevante di B e C.

Due utenti possono quindi accedere allo stesso catalogo ma visualizzarlo in maniera profondamente differente.
Netflix utilizza il machine learning anche per costruire la homepage.
Il sistema deve decidere quali righe mostrare, quali titoli inserire all'interno di ogni riga e in quale ordine presentare sia le righe sia i singoli contenuti (Netflix, n.d.-c).
La pipeline può quindi diventare:

```
Catalogo
↓
Righe candidate
↓
Ordinamento delle righe
↓
Ordinamento dei titoli
↓
Composizione della pagina
↓
Homepage personalizzata
```

La personalizzazione deve inoltre evitare di diventare eccessivamente ripetitiva.
Se il sistema deducesse che un utente è interessato alla fantascienza e mostrasse esclusivamente righe dedicate a quel genere, la previsione potrebbe essere statisticamente coerente ma l'esperienza risulterebbe povera.
Il recommendation system deve quindi bilanciare elementi come:
rilevanza + diversità + novità + familiarità.

Quando un'interfaccia mostra una percentuale di compatibilità, questa va interpretata come una previsione personalizzata; la disponibilità di tale indicatore può variare tra versioni e dispositivi.
Una percentuale elevata non significa: "il 95% degli utenti considera eccellente questo film".
Significa invece che Netflix stima una forte compatibilità tra quel contenuto e il particolare profilo.

La personalizzazione può inoltre reagire ai segnali recenti.
Un utente potrebbe normalmente guardare documentari ma, durante una specifica sessione, iniziare a esplorare commedie. Il sistema può quindi considerare non soltanto preferenze di lungo periodo, ma anche il contesto corrente della sessione (Netflix, n.d.-c).

Quando l'utente seleziona finalmente un contenuto, il control plane deve trasformare quella scelta in una vera sessione di streaming.
È in questo momento che il control plane e il media plane si incontrano.
Gli Open Connect Appliances, gli edge server della CDN Netflix, comunicano periodicamente ai servizi di controllo informazioni relative al proprio stato (Netflix, n.d.-e).
Tra queste informazioni possiamo avere:

- La salute del server.

- Le rotte BGP apprese dai router degli ISP: indicano quali prefissi di rete sono raggiungibili e aiutano il sistema a individuare gli OCA appropriati per servire i client.
- I file disponibili.

Quando il client invia una richiesta di riproduzione, i servizi di playback verificano innanzitutto che l'utente sia autorizzato a vedere quel contenuto.
Devono inoltre considerare informazioni come:

- Account.
- Piano.
- Licenza.
- Funzionalità del dispositivo.
- Condizioni della rete.

Una volta individuate le rappresentazioni necessarie, entra in gioco il **servizio di steering**.
Lo steering non distribuisce direttamente il film.
Il suo compito è decidere da quale *OCA* il client dovrebbe recuperarlo.
Per farlo utilizza le informazioni ricevute dagli appliance, come:

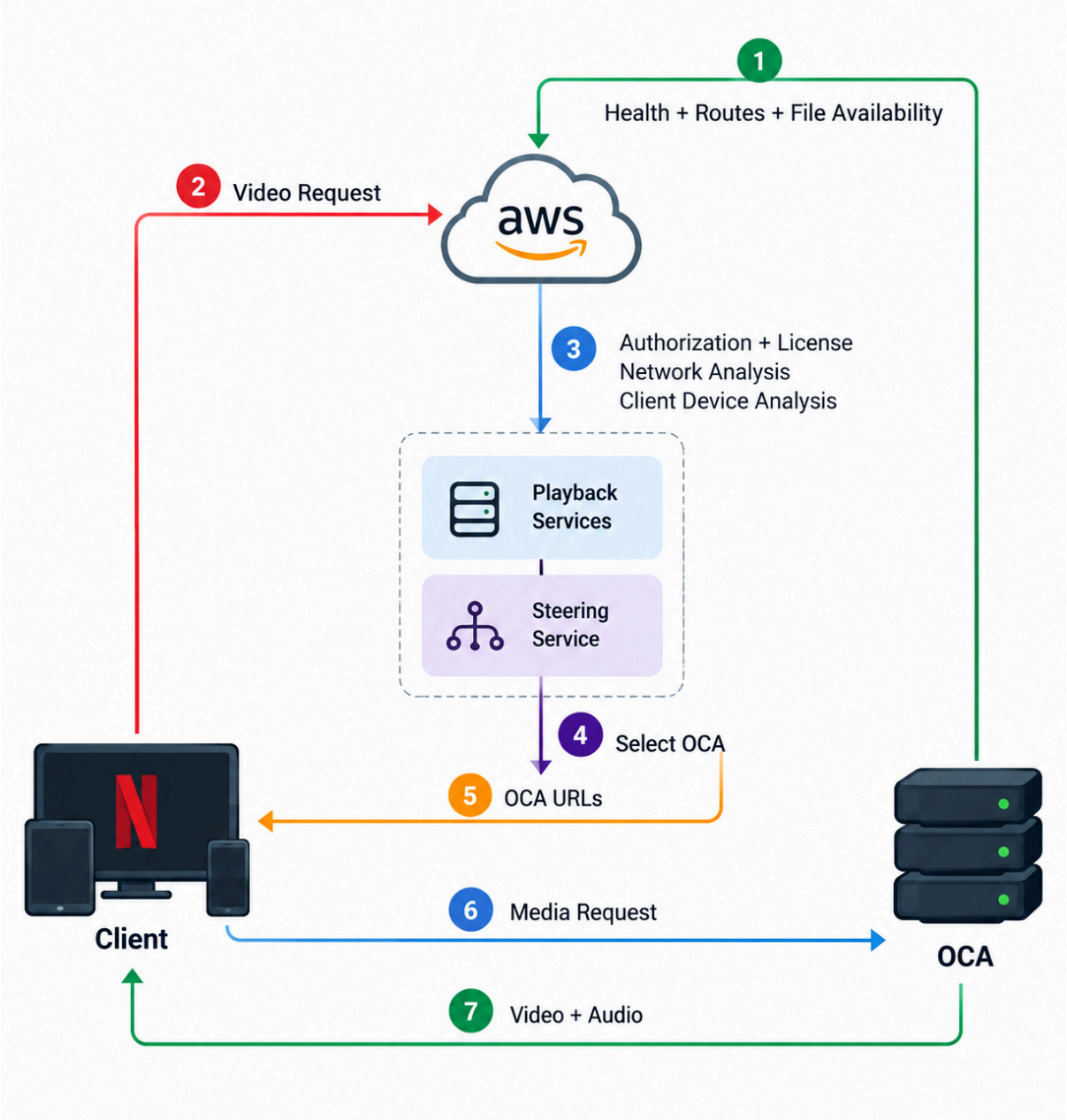
- Disponibilità del file.
- Salute *OCA*.
- Prossimità.
- Percorso di rete.
- Requisiti client.

Non è quindi corretto dire semplicemente che Netflix sceglie "il server geograficamente più vicino".
Un *OCA* molto vicino potrebbe non possedere quel contenuto; un altro potrebbe essere temporaneamente indisponibile; un terzo potrebbe offrire un percorso di rete migliore.

Il processo può essere rappresentato così (Netflix, n.d.-e):

```
1. OCA → AWS
Salute + Rotte + Disponibilità file
↓
2. Client → AWS
Richiesta video
↓
3. Playback Services
Autorizzazione + Licenza + Analisi network + Analisi del dispositivo del client
↓
4. Steering Service
Selezione OCA
↓
5. AWS → Client
OCA URLs
↓
6. Client → OCA
```

```
Richiesta Media
↓
7. OCA → Client
Video + Audio
```



Una caratteristica particolarmente interessante di Open Connect è il **proactive caching**.
Una *CDN* puramente reattiva può recuperare un contenuto dall'origine dopo che viene richiesto e conservarlo successivamente nella cache.
Netflix possiede però un vantaggio: gran parte del proprio catalogo è conosciuta in anticipo. Può quindi stimare quali contenuti saranno maggiormente richiesti nelle diverse aree e distribuirli preventivamente agli *OCA*.

Il processo può diventare:

```
Previsione della popolarità
↓
Decisione di collocazione del contenuto
↓
Riempimento nelle ore di bassa attività
↓
OCA regionale
↓
Richieste future degli utenti
```

Gli aggiornamenti possono essere effettuati durante periodi off-peak, cioè nelle fasce nelle quali la rete è meno utilizzata.
Con un modello di push fill, nuovi titoli, nuove codifiche e aggiornamenti possono essere trasferiti agli appliance prima dell'aumento della domanda (Netflix, n.d.-e).
Questo significa che, quando migliaia di utenti iniziano a guardare un nuovo episodio durante la serata, buona parte dei dati può essere già presente all'interno o molto vicino alle reti che dovranno distribuirlo.

Non bisogna però immaginare che ogni *OCA* contenga l'intero catalogo.
Lo storage è limitato.
Netflix deve quindi decidere continuamente: quali contenuti → su quali *OCA* → in quali regioni.
La previsione della popolarità diventa quindi anche un problema di ottimizzazione dello storage e del traffico di rete (Vora et al., 2017).

Open Connect non deve essere immaginata come un insieme di server completamente indipendenti che recuperano sempre i contenuti direttamente dall'origine.
Gli OCA possono recuperare aggiornamenti anche da altri nodi Open Connect durante il riempimento della rete. Nel VOD, però, la documentazione descrive soprattutto un posizionamento preventivo dei file, non una catena generalizzata di cache miss attivata da ogni richiesta di playback (Netflix, n.d.-e; Liu et al., 2025).
Durante la riproduzione VOD lo steering indirizza il client verso OCA che dispongono degli asset richiesti. Il recupero a richiesta attraverso una gerarchia di cache è invece particolarmente rilevante per gli asset live e per altre categorie di oggetti, come descritto nella sezione successiva (Netflix, n.d.-e; Liu et al., 2025).
Il trasferimento e il posizionamento dei contenuti nella rete possono essere rappresentati così; lo schema non implica un cache miss VOD per ogni richiesta:

```
Origin / Sorgente Netflix
↓
OCA di livello superiore
↓
OCA regionali o integrati negli ISP
↓
Client
```

Il percorso varia con il tipo di asset e con la fase operativa. Occorre distinguere il riempimento degli OCA dalla successiva distribuzione al client: nel VOD prevale il pre-positioning, mentre il live richiede il recupero progressivo dei segmenti appena prodotti (Liu et al., 2025).
Possiamo quindi distinguere, senza attribuire lo stesso comportamento a tutti gli asset:

- File VOD disponibile → lo steering indica un OCA idoneo, che lo distribuisce al client.
- *Peer / tier fill → trasferimento tra nodi Open Connect, con modalità legate al tipo di asset.*
- *Origin fill → recupero dalla sorgente, per esempio dei nuovi segmenti live.*

Il riuso dei contenuti distribuiti nella rete riduce il carico sui sistemi di origine; tempi e meccanismi di riempimento differiscono tra VOD, live e altri asset.

Aver portato il contenuto vicino all'utente non garantisce ancora una buona esperienza.
Le condizioni della rete possono cambiare anche durante la riproduzione.
Un utente potrebbe iniziare con una connessione molto veloce e subire pochi secondi dopo congestione, interferenze Wi-Fi o variazioni della banda disponibile.
Per questo Netflix utilizza l'*ABR*.
Lo stesso contenuto è disponibile a differenti livelli qualitativi e il client può cambiare rappresentazione durante la riproduzione.

Possiamo immaginare:

```
OCA
↓
High Quality
Medium Quality
Low Quality
↓
ABR Algorithm
↓
Buffer
↓
DRM / Decryption
↓
Decoder
↓
Audio + Video
```

L'algoritmo cerca continuamente un equilibrio.

- Una rappresentazione con bitrate elevato può offrire maggiore qualità, ma richiede più banda.
- Una rappresentazione con bitrate inferiore può ridurre la qualità visiva, ma aumenta la probabilità di mantenere continua la riproduzione.

L'obiettivo è quindi massimizzare la qualità senza esaurire il buffer.

Gli OCA mettono a disposizione i media tramite HTTP/HTTPS, mentre i dettagli specifici della distribuzione possono variare a seconda del dispositivo e dell'implementazione (Netflix, n.d.-e).

I media distribuiti da Netflix sono protetti da copyright e non devono poter essere utilizzati liberamente al di fuori delle condizioni previste dal servizio.
Netflix utilizza sistemi DRM per controllare la riproduzione dei contenuti protetti. Per esempio, su Android la documentazione richiede Widevine L1 per la riproduzione HD; questo requisito specifico non implica che tutti i dispositivi Netflix utilizzino lo stesso DRM (Netflix, n.d.-d).

Con l'introduzione del piano supportato dalla pubblicità, la sessione può contenere contenuti pubblicitari aggiuntivi.

Nel piano con pubblicità, Netflix non deve decidere soltanto quale contenuto distribuire e da quale *OCA* recuperarlo. Deve anche determinare se mostrare pubblicità, dove inserirla e quali annunci utilizzare.
Il modello di business entra quindi direttamente nella pipeline tecnica.
Il Control Plane conosce le condizioni di accesso associate all'utente e può distinguere, in generale, due situazioni:

- Ad-free Plan → Content Playback.
- Ad-supported Plan → Content + Ad Opportunities.

Questa distinzione rappresenta il funzionamento ordinario dei due piani, anche se alcuni contenuti, come determinati eventi live, possono prevedere eccezioni (Netflix, n.d.-a).
Netflix utilizza la propria infrastruttura pubblicitaria, denominata Netflix Ads Suite, per supportare operazioni come gestione delle campagne, targeting, distribuzione degli annunci, misurazione e reporting (Netflix, 2026b).
Non è necessario assumere che gli annunci siano conservati esattamente sugli stessi *OCA* utilizzati per film e serie.
È più corretto mantenerli concettualmente separati e farli convergere nel client:

- Open Connect → Content Delivery
- Advertising Infrastructure → Ad Delivery

Il percorso non termina quando il film viene riprodotto.
Durante la sessione Netflix raccoglie continuamente informazioni sul funzionamento della piattaforma.
Gli *OCA* forniscono dati relativi al proprio stato, alla disponibilità dei file e alla raggiungibilità della rete.
Il client può invece produrre telemetria relativa alla qualità dell'esperienza.
Tra le metriche più importanti possiamo avere:

- Startup delay: tempo necessario prima che inizi la riproduzione.
- Rebuffering: frequenza e durata delle interruzioni causate dall'esaurimento del buffer.
- Bitrate : quantità di dati per unità di tempo.
- Visual quality: stima della qualità visiva ricevuta, distinta dal solo bitrate.
- Playback Errors: errori verificatisi durante la sessione.

A questi dati si aggiungono gli user events utilizzati per comprendere l'interazione con catalogo e interfaccia e, nelle sessioni supportate dalla pubblicità, gli advertising events necessari a misurare l'erogazione delle campagne.

Si formano quindi almeno tre categorie di feedback:

- Playback Telemetry → come sta funzionando lo streaming?
- User Events → come sta utilizzando il servizio l'utente?
- Advertising Events → come è stata erogata la campagna?

Questi dati possono alimentare sistemi differenti:

- Analytics.
- Machine Learning.
- Recommendation.
- Experimentation.
- Advertising Measurement.
- Infrastructure Optimization.

Netflix può quindi adattarsi su due scale temporali.

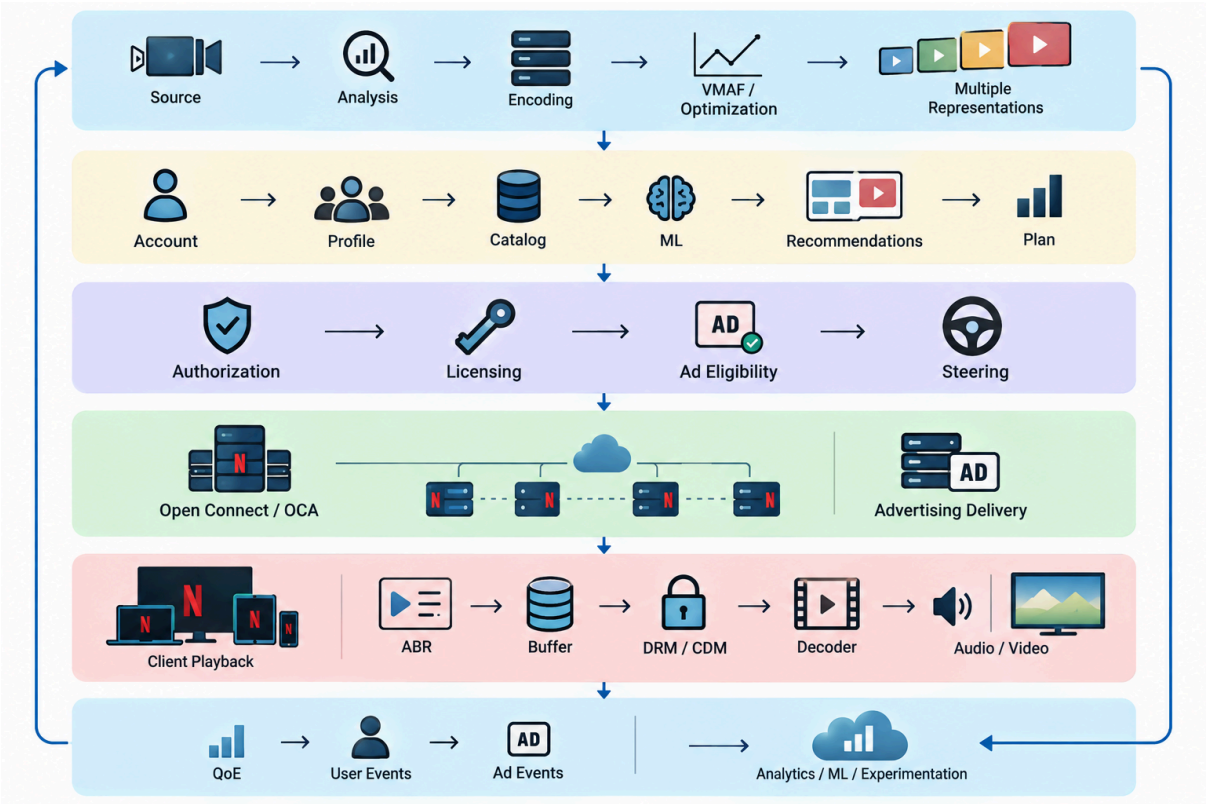
- Durante la singola sessione l'*ABR* reagisce rapidamente alle condizioni della rete.
- Nel medio e lungo periodo l'analisi di milioni di sessioni permette invece di migliorare modelli di raccomandazione, codifiche, algoritmi *ABR*, strategie di caching, interfaccia e infrastruttura.

L'intelligenza artificiale non è quindi un componente separato aggiunto alla fine della piattaforma.
È integrata nel modo in cui il catalogo viene filtrato, ordinato, presentato e continuamente adattato al profilo.

In conclusione l'intera architettura Netflix può essere ricondotta a sei grandi livelli funzionali.

```
CONTENT PREPARATION
Source → Analysis → Encoding → VMAF / Optimization → Multiple
Representations
↓
CONTROL & DATA INTELLIGENCE -- AWS
Account → Profile → Catalog → ML → Recommendations → Plan
↓
PLAYBACK ORCHESTRATION
Authorization → Licensing → Ad Eligibility → Steering
↓
MEDIA DELIVERY
Open Connect / OCA + Advertising Delivery
↓
CLIENT PLAYBACK
ABR → Buffer → DRM/CDM → Decoder → Audio / Video
↓
DATA & FEEDBACK
```

```
QoE + User Events + Ad Events → Analytics / ML / Experimentation
↺
```



Nota:
Lo schema rappresenta in forma semplificata l'architettura complessiva. Per quanto riguarda il *DRM*, nella pratica il contenuto multimediale cifrato e la licenza *DRM* non percorrono necessariamente la stessa sequenza: il primo viene distribuito attraverso l'infrastruttura di media delivery, mentre la licenza viene ottenuta separatamente attraverso i servizi di autorizzazione del backend. I due flussi convergono sul client, dove il componente *DRM* utilizza la licenza per consentire la decifratura del contenuto prima della decodifica e della riproduzione.

Questa struttura permette di comprendere perché Netflix sia molto più di un insieme di server che distribuiscono file video.

- Il contenuto viene prima preparato per poter essere distribuito in modo efficiente.
- Il control plane stabilisce chi è l'utente, che cosa può vedere e quale esperienza deve ricevere.
- I sistemi di machine learning contribuiscono a decidere quali contenuti mostrargli e come organizzare il catalogo.
- Quando l'utente preme Play, i servizi di playback verificano autorizzazioni e requisiti, mentre lo steering utilizza le informazioni ricevute dagli OCA per stabilire da dove recuperare i media.
- Open Connect provvede quindi alla distribuzione effettiva del contenuto.
- Il client sceglie continuamente quale rappresentazione utilizzare, mantiene il buffer, applica le autorizzazioni *DRM* e decodifica audio e video.

- Nel piano supportato dalla pubblicità, a questo percorso viene affiancata una seconda infrastruttura che seleziona e distribuisce gli annunci.
- Infine, gli eventi prodotti dalla riproduzione e dall'interazione ritornano alla piattaforma dati, dove vengono utilizzati per migliorare infrastruttura, algoritmi e personalizzazione.

Netflix può quindi essere interpretata non come una semplice pipeline lineare, ma come un sistema adattivo e ciclico, nel quale cloud computing, machine learning, encoding, *CDN*, networking, advertising e software client collaborano continuamente.
Premere Play rappresenta soltanto l'azione visibile all'utente.
Dietro quell'azione, l'infrastruttura deve simultaneamente decidere che cosa mostrare, se l'utente può vederlo, quale versione utilizzare, da quale server recuperarla, a quale qualità riprodurla, come proteggerla, se integrare pubblicità e quali dati raccogliere per migliorare la sessione successiva.

Architettura live

L'architettura descritta finora è particolarmente adatta al Video on Demand.

Nel *VOD* Netflix dispone del contenuto completo prima che l'utente ne richieda la riproduzione. Un film o un episodio può essere analizzato, codificato, suddiviso in differenti rappresentazioni e distribuito attraverso Open Connect ore o giorni prima che qualcuno prema Play.

Nel live streaming questa condizione cambia radicalmente.
Durante una trasmissione in diretta, infatti, il contenuto viene prodotto nello stesso momento in cui deve essere elaborato e distribuito. La piattaforma non può quindi preparare anticipatamente l'intero asset: deve trasformare continuamente il segnale proveniente dall'evento in segmenti che, pochi secondi dopo, possano essere riprodotti da milioni di dispositivi.

Netflix controlla questo processo lungo gran parte della catena, dalla produzione fino allo schermo dello spettatore. La pipeline generale può essere rappresentata così:

```
Live Production / Venue
↓
Contribution Feed
↓
Netflix Broadcast Operations Center
↓
AWS MediaConnect
↓
AWS MediaLive
↓
Netflix Custom Packager
↓
```

```
Netflix Live Origin
↓
Open Connect
↓
Netflix Client / Playback
```

Il primo elemento è la **Live Production**, cioè la produzione audiovisiva dell'evento. Può trattarsi, per esempio, di un incontro sportivo, uno spettacolo, un evento *WWE* o un'altra produzione trasmessa in diretta.

Le telecamere e le altre sorgenti non vengono però inviate direttamente agli utenti. Sul luogo dell'evento viene realizzato un programma completo, nel quale sono già state effettuate operazioni come regia, selezione delle camere, missaggio audio e altre attività proprie della produzione televisiva.
Questo programma viene quindi inviato attraverso un **Contribution Feed**.

Il termine contribution indica il trasporto del segnale professionale dalla sede nella quale viene prodotto l'evento verso l'infrastruttura che dovrà successivamente elaborarlo e distribuirlo.
Non si tratta ancora dello streaming destinato agli spettatori: è il collegamento che trasporta il segnale dalla produzione verso Netflix.

A seconda dell'evento possono essere utilizzati differenti sistemi di trasmissione, come collegamenti su fibra, reti IP o altre infrastrutture broadcast.
Netflix utilizza inoltre percorsi ridondanti, perché la perdita del contribution feed interromperebbe alla sorgente l'intera trasmissione.

Il segnale raggiunge quindi il **Netflix Broadcast Operations Center**, o strutture equivalenti dedicate alle operazioni broadcast.
Il Broadcast Operations Center costituisce il punto di incontro tra il mondo della produzione televisiva tradizionale e quello dell'infrastruttura streaming.

Qui il programma ricevuto può essere sottoposto a operazioni di signal ingest, controllo e ispezione.
Netflix descrive queste strutture come centri nei quali vengono gestiti anche elementi come sincronizzazione audio/video, closed captioning (un sistema che mostra sullo schermo il testo corrispondente ai dialoghi e a tutti i suoni rilevanti di un video), grafica, metadati e segnali utilizzati per la gestione pubblicitaria. Il loro compito è quindi assicurarsi che il feed che entrerà nella pipeline streaming sia corretto e utilizzabile (Fedorov et al., 2025).

Possiamo quindi distinguere:
- Contribution Feed → porta il programma dalla produzione a Netflix.
- Broadcast Operations Center → riceve, controlla e prepara operativamente il segnale prima dell'ingresso nella pipeline streaming.

Una volta verificato, il programma deve entrare nell'infrastruttura cloud.
Netflix utilizza AWS Elemental MediaConnect e AWS Elemental MediaLive per questa parte della pipeline. I due servizi svolgono però funzioni differenti (Deal et al., 2025).

AWS MediaConnect è principalmente un sistema di trasporto affidabile del video live.
Riceve il feed e permette di trasferirlo attraverso l'infrastruttura *AWS* garantendo caratteristiche particolarmente importanti per contenuti professionali, come affidabilità, monitoraggio, ridondanza e meccanismi per limitare gli effetti della perdita di pacchetti.

Il flusso viene quindi passato a AWS MediaLive.
MediaLive svolge invece la fase di encoding e transcoding in tempo reale.
Il programma ricevuto dal Broadcast Operations Center rappresenta infatti un singolo feed professionale, ma Netflix deve distribuirlo a dispositivi e connessioni estremamente differenti.

MediaLive genera quindi contemporaneamente diverse rappresentazioni del live:

```
Live Feed
↓
AWS MediaLive
↓
SD / HD / UHD
+ bitrate differenti
+ codec differenti quando necessari
```

Questo costituisce la bitrate ladder utilizzata successivamente dal sistema *ABR*.
Netflix utilizza pipeline cloud ridondanti, distribuite anche su regioni differenti (Deal et al., 2025).
La stessa trasmissione può quindi essere elaborata attraverso pipeline indipendenti, riducendo la probabilità che un singolo problema di encoding, rete o infrastruttura interrompa l'evento.

Dal 26 gennaio 2026 Netflix utilizza inoltre Variable Bitrate (VBR) per tutti i propri eventi live (Teixeira et al., 2026).
Netflix utilizzava da tempo strategie di Variable Bitrate nel *VOD*, dove il contenuto può essere analizzato e ottimizzato prima della distribuzione.
Nel live lo stesso principio deve invece essere applicato in tempo reale.
Il VBR è implementato attraverso la modalità QVBR di AWS MediaLive. Il bitrate può variare in funzione della complessità visiva della scena, entro i vincoli configurati per la codifica (Teixeira et al., 2026).
Il principio è:

- Scena semplice → meno dati.
- Scena complessa → più dati.

Una ripresa relativamente statica può essere compressa utilizzando pochi bit, mentre una sequenza sportiva con movimento rapido, pubblico, luci e molti dettagli può richiederne molti di più.

Questo migliora l’efficienza, ma introduce anche una conseguenza importante sull’infrastruttura di distribuzione.

Un server che durante una scena semplice sta trasferendo relativamente pochi dati può sembrare avere molta capacità disponibile. Se il sistema gli assegnasse troppe nuove sessioni e poco dopo il bitrate aumentasse improvvisamente, il server potrebbe trovarsi sovraccarico.
Per questo Netflix non considera soltanto il traffico istantaneo, ma riserva capacità considerando anche il bitrate nominale delle rappresentazioni. Il live dimostra quindi molto chiaramente come encoding, capacity management (la gestione della capacità disponibile dell’infrastruttura, cioè quante sessioni e quanti dati un sistema può sostenere senza degradare il servizio), traffic steering e CDN non possano essere progettati come sistemi indipendenti (Teixeira et al., 2026).

Le rappresentazioni prodotte da MediaLive devono successivamente essere trasformate in una struttura utilizzabile dal sistema di streaming.
Questa funzione viene svolta da un **Custom Packager** sviluppato da Netflix.

Il packager non ricodifica nuovamente il video. Riceve i flussi già codificati e li organizza nella forma richiesta dall’architettura Netflix, producendo i segmenti e le informazioni necessarie alla distribuzione e alla riproduzione.

Concettualmente:

```
Encoded Live Streams
↓
Netflix Custom Packager
↓
Streaming Segments + Playback Information
```

Netflix ha scelto di sviluppare un proprio packager per integrarlo maggiormente con i sistemi di playback, DRM e distribuzione già esistenti (Deal et al., 2025).
Il packager svolge inoltre un ruolo nella resilienza. Netflix effettua controlli leggeri sui segmenti generati e può associare informazioni che indicano eventuali problemi, come segmenti mancanti, troppo corti o temporalmente incoerenti. Queste informazioni possono successivamente essere utilizzate dal sistema per evitare di distribuire una versione difettosa quando ne è disponibile una corretta proveniente da una pipeline ridondante (Liu et al., 2025).

I segmenti prodotti vengono quindi inviati al **Netflix Live Origin**.
Il Live Origin è uno dei componenti più importanti che distinguono l’architettura live dall'architettura *VOD*.

- Nel VOD gli asset possono essere preparati e preposizionati con grande anticipo sugli OCA.
- Nel live, invece, ogni nuovo segmento viene creato pochi istanti prima di essere richiesto dagli spettatori.

Netflix ha quindi sviluppato un origin specifico per il live, collocato tra le pipeline cloud e Open Connect:

```
Custom Packager
↓
Netflix Live Origin
↓
Open Connect
```

Il packager pubblica continuamente i nuovi segmenti sul Live Origin.
Open Connect può quindi recuperarli e distribuirli agli edge server e, successivamente, ai client.
Netflix descrive il Live Origin come il punto di separazione tra la pipeline live eseguita nel cloud e la propria CDN (Liu et al., 2025).

Ma il Live Origin non svolge soltanto il ruolo di deposito temporaneo.
Poiché Netflix utilizza pipeline ridondanti, possono esistere più versioni candidate dello stesso segmento.
Il Live Origin conosce le differenti pipeline e regioni e può selezionare un segmento valido quando uno degli altri presenta un problema (Liu et al., 2025).

Il percorso diventa quindi:

```
Pipeline A → Segment A
Pipeline B → Segment B
↓
Netflix Live Origin
↓
Selezione del segmento valido
↓
Open Connect
```

Questa decisione viene presa lato server, evitando di trasferire al client tutta la complessità della gestione.

Dopo il Live Origin entra in gioco Open Connect, la *CDN* proprietaria di Netflix.

- Nel Video on Demand, Netflix conosce i contenuti con largo anticipo e può quindi utilizzare strategie di pre-positioning, distribuendo preventivamente sugli *OCA* i titoli che prevede saranno richiesti in una determinata area geografica.
- Nel live questo meccanismo non può essere utilizzato nello stesso modo, perché il prossimo segmento dello stream non esiste ancora. Ogni segmento viene generato dal packager pochi istanti prima di essere pubblicato sul Live Origin.

Open Connect deve quindi recuperare progressivamente questi nuovi segmenti mentre l'evento è in corso.
Questo non significa però che ogni OCA contatti direttamente il Live Origin.

Gli *OCA* fanno parte di una gerarchia di cache e possono recuperare un contenuto anche da altri nodi Open Connect.
Quando un nuovo segmento live diventa disponibile, alcuni nodi della rete possono recuperarlo dal Live Origin e conservarlo temporaneamente nella propria cache.
Altri *OCA* possono successivamente ottenere lo stesso segmento da questi nodi attraverso meccanismi di peer fill o tier fill, evitando che tutte le richieste raggiungano direttamente l'origine.
Il percorso può quindi essere rappresentato, in forma semplificata, così:

```
Netflix Live Origin
↓
OCA di livello superiore
↓
OCA regionali / ISP
↓
Client
```

Se il segmento è già presente nella cache dell'*OCA* che serve l'utente, viene distribuito immediatamente.
Se invece non è disponibile localmente, l'*OCA* può recuperarlo da un altro nodo Open Connect. Solo quando necessario la richiesta risale progressivamente verso i nodi che possono recuperare il segmento dal Live Origin.
Il comportamento può quindi essere sintetizzato così:

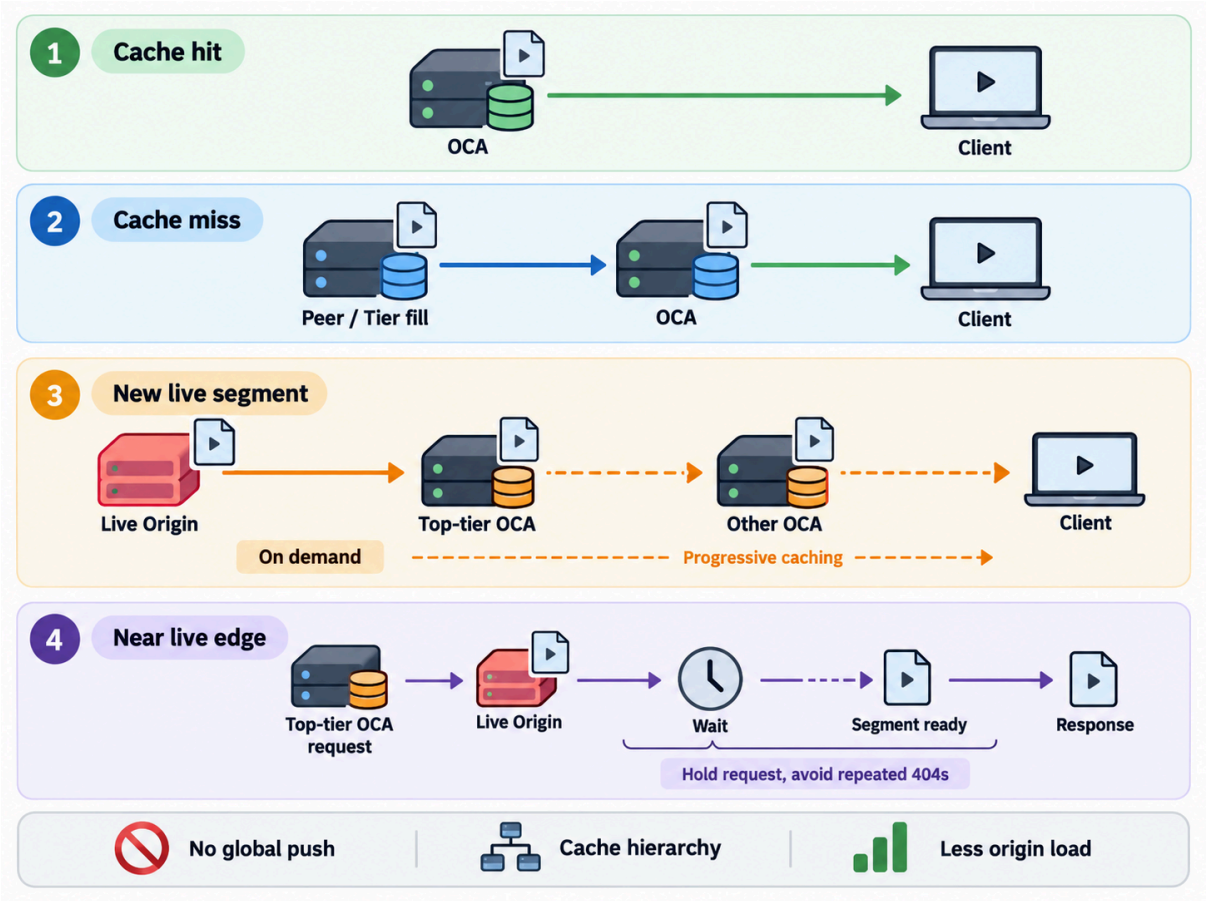
- Cache Hit → OCA → Client
- Cache Miss → Peer / Tier Fill → OCA → Client

Oppure, quando il segmento deve ancora entrare nella rete Open Connect:
Live Origin → OCA di livello superiore → altri OCA → Client

In questo modo un nuovo segmento entra progressivamente nella rete Open Connect in funzione delle richieste effettive degli spettatori e può essere riutilizzato dalle cache per servire molti client.
Invece, se un top-tier OCA richiede il prossimo segmento leggermente prima che sia stato pubblicato, il Live Origin può mantenere aperta la richiesta e rispondere appena il segmento arriva, invece di restituire immediatamente un errore 404.
Questo riduce le continue richieste ripetute vicino al live edge.

Nel live Open Connect passa quindi da un modello basato soprattutto sul pre-positioning anticipato a un modello nel quale i nuovi segmenti vengono recuperati, cachati e redistribuiti progressivamente mentre la trasmissione è in corso (Liu et al., 2025).

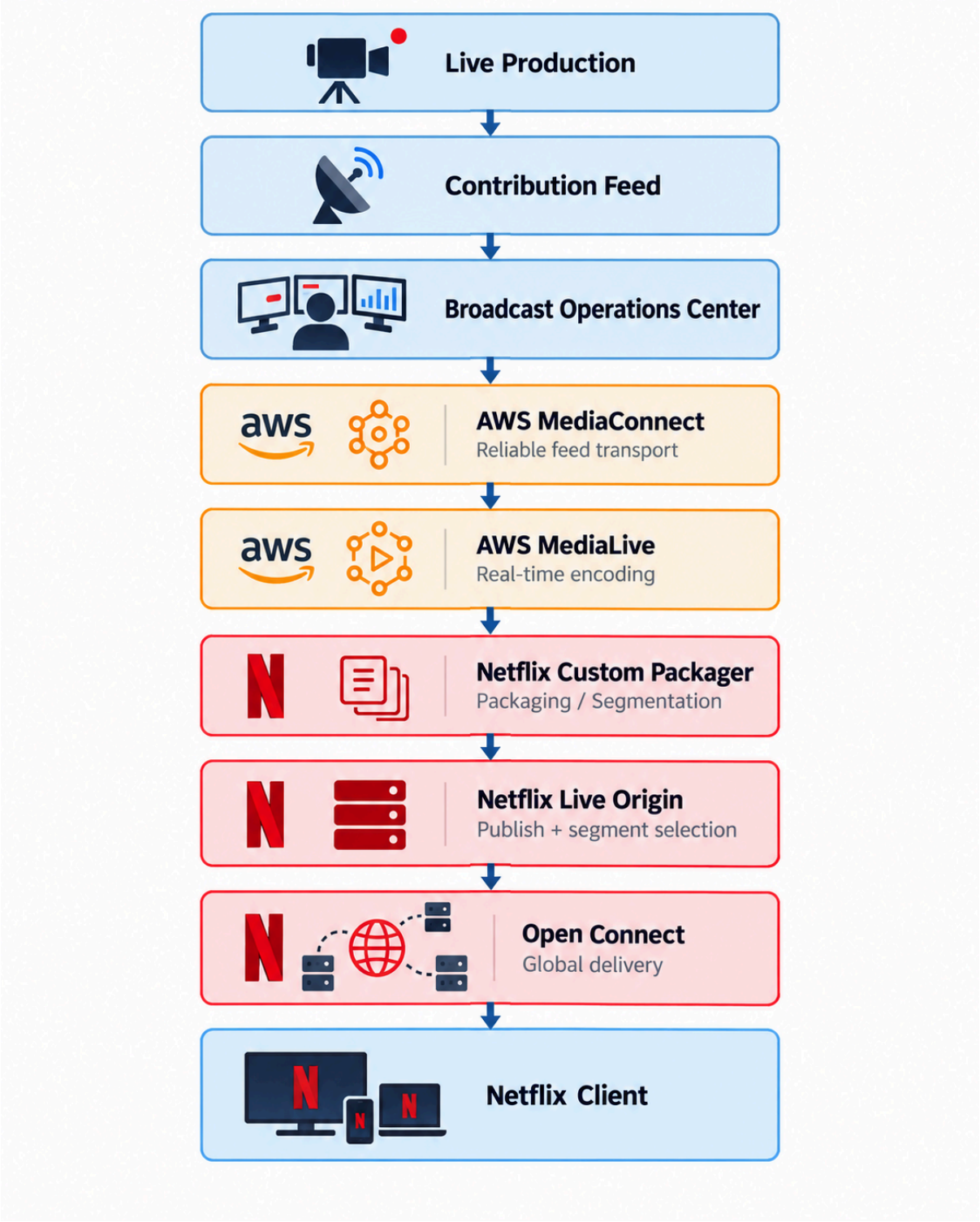
Questo permette alla *CDN* di continuare a svolgere la propria funzione fondamentale: evitare che milioni di richieste degli spettatori raggiungano direttamente il Live Origin e distribuire invece il traffico attraverso la rete di cache Open Connect.



La pipeline completa diventa quindi:

```
Live Production
↓
Contribution Feed
↓
Broadcast Operations Center
↓
AWS MediaConnect
Trasporto affidabile del feed
↓
AWS MediaLive
Encoding / Transcoding real-time
↓
Netflix Custom Packager
Packaging / Segmentation
↓
Netflix Live Origin
Pubblicazione + selezione dei segmenti
↓
Open Connect
Distribuzione globale
↓
```

```
Netflix Client
```



Durante la visione il player osserva le condizioni della rete e può cambiare rappresentazione per mantenere un compromesso tra qualità e stabilità tramite il principio dell'*ABR*.

All'inizio della riproduzione il client riceve dal cloud una configurazione di playback e un manifest che descrive, tra le altre cose, le rappresentazioni disponibili e i server utilizzabili.

Nell’architettura live descritta da Netflix nel 2025, i segmenti hanno una durata di circa 2 secondi e il manifest include un segment template: il client può ricavare gli URL dei segmenti dal tempo della trasmissione, senza richiedere continuamente un nuovo manifest (Fedorov et al., 2025).

Il problema fondamentale del live è quindi il tempo.
Per questo latenza, ridondanza, sincronizzazione, osservabilità e capacità di recuperare rapidamente da un errore diventano requisiti centrali.
Netflix raccoglie inoltre telemetria praticamente lungo l’intera pipeline, dall’ingest fino al playback. Nei grandi eventi i sistemi di monitoraggio hanno elaborato fino a circa 38 milioni di eventi al secondo, permettendo agli operatori di osservare rapidamente problemi relativi a encoder, CDN, playback e qualità dell’esperienza (Fedorov et al., 2025).

Anche la pubblicità introduce una difficoltà aggiuntiva.
Nel *VOD* le opportunità pubblicitarie possono essere associate a un contenuto già conosciuto.
Nel live, invece, i segnali relativi alle interruzioni possono essere prodotti durante la trasmissione e devono propagarsi attraverso la pipeline in tempo utile.

Il Broadcast Operations Center monitora anche i metadati relativi all’inserimento pubblicitario, mentre Netflix utilizza tecnologie di Dynamic Ad Insertion per alcuni eventi live, consentendo di selezionare dinamicamente gli annunci destinati alle differenti sessioni (Fedorov et al., 2025; Reinhard, 2025).
Due spettatori dello stesso evento possono quindi ricevere pubblicità differenti durante la stessa interruzione.

Il live non rappresenta quindi semplicemente una versione più veloce del VOD.
È una pipeline nella quale produzione, contribution, controllo broadcast, trasporto cloud, encoding, packaging, origin, CDN e client devono operare contemporaneamente, mantenendo una catena continua nella quale ogni nuovo segmento viene prodotto, elaborato e distribuito mentre l’evento sta ancora accadendo.
È proprio questa continuità in tempo reale a rendere il live una delle parti più complesse dell’architettura Netflix.

Materiale di supporto

Bibliografia

Aaron, A., Li, Z., Manohara, M., De Cock, J., & Ronca, D. (2015). Per-Title Encode Optimization. https://netflixtechblog.com/per-title-encode-optimization-7e99442b62a2

Deal, A., Concolato, C., Ribeiro, F., Ford, H., Afonso, M., Symborski, T., & Liu, X. (2025). Behind the Streams: Building a Reliable Cloud Live Streaming Pipeline for Netflix. Part 2. https://netflixtechblog.com/building-a-reliable-cloud-live-streaming-pipeline-for-netflix-8627c608c967

Fedorov, S., Pham, C., Ribeiro, F., Newton, C., & Wei, W. (2025). Behind the Streams: Three Years Of Live at Netflix. Part 1. https://netflixtechblog.com/behind-the-streams-live-at-netflix-part-1-d23f917c2f40

Gomez-Uribe, C. (2016). A Global Approach to Recommendations. https://about.netflix.com/en/news/a-global-approach-to-recommendations

Guo, L., Li, Z., Radford, S., & Watts, J. (2025). AV1 — Now Powering 30% of Netflix Streaming. https://netflixtechblog.com/av1-now-powering-30-of-netflix-streaming-02f592242d80

Haspilaire, G. (2022). Open Connect: Celebrating a Decade of Smooth and Efficient Streaming. https://about.netflix.com/en/news/open-connect-celebrating-a-decade-of-smooth-and-efficient-streaming

Izrailevsky, Y., Vlaovic, S., & Meshenberg, R. (2016). Completing the Netflix Cloud Migration. https://about.netflix.com/en/news/completing-the-netflix-cloud-migration

Li, Z., Bampis, C., Novak, J., Aaron, A., Swanson, K., Moorthy, A., & De Cock, J. (2018). VMAF: The Journey Continues. https://netflixtechblog.com/vmaf-the-journey-continues-44b51ee9ed12

Liu, X., Lynch, J., & Newton, C. (2025). Netflix Live Origin. https://netflixtechblog.com/netflix-live-origin-41f1b0ad5371

Manohara, M., Moorthy, A., De Cock, J., Katsavounidis, I., & Aaron, A. (2018). Optimized shot-based encodes: Now Streaming!. https://netflixtechblog.com/optimized-shot-based-encodes-now-streaming-4b9464204830

Netflix (1998). First Online DVD Rental Store Opens. https://about.netflix.com/en/news/first-online-dvd-rental-store-opens

Netflix (2000). Personalized Movie Recommendor Provides Netflix Visitors With Highly Accurate Film Recommendations Based on Their Individual Movie Taste History. https://about.netflix.com/en/news/personalized-movie-recommendor-provides-netflix-visitors-with-highly

Netflix (2007). Netflix Offers Subscribers the Option of Instantly Watching Movies on Their PCs. https://about.netflix.com/en/news/netflix-offers-subscribers-the-option-of-instantly-watching-movies-on-their

Netflix (2015). Do You Know When You Were Hooked? Netflix Does. https://about.netflix.com/en/news/do-you-know-when-you-were-hooked-netflix-does

Netflix (2016). Netflix & Binge: New Binge Scale Reveals TV Series We Devour and Those We Savor. https://about.netflix.com/en/news/netflix-binge-new-binge-scale-reveals-tv-series-we-devour-and-those-we-savor-1

Netflix (2018). Cody Fern Joins ‘House of Cards’ as Series Regular. https://about.netflix.com/en/news/cody-fern-joins-house-of-cards-as-series-regular

Netflix (2023a). An Update on Sharing. https://about.netflix.com/en/news/an-update-on-sharing

Netflix (2023b). The Netflix Cup Crowns Its Winners. https://about.netflix.com/en/news/netflix-cup-2023-winners

Netflix (2026a). All WWE Is Now Live Only on Netflix in Italy. https://about.netflix.com/en/news/all-wwe-is-now-live-only-on-netflix-in-italy

Netflix (2026b). Netflix Upfront 2026: Get Closer. https://about.netflix.com/en/news/netflix-upfront-2026-get-closer?pubDate=20260514

Netflix (n.d.-a). Ads on Netflix. Retrieved October 28, 2026, from https://help.netflix.com/en/go/ads

Netflix (n.d.-b). Footage Management page in Content Hub. Retrieved October 28, 2026, from https://partnerhelp.netflixstudios.com/hc/en-us/articles/49769502060947-Footage-Management-page-in-Content-Hub

Netflix (n.d.-c). How Netflix's Recommendations System Works. Retrieved October 28, 2026, from https://help.netflix.com/en/node/100639

Netflix (n.d.-d). How to use Netflix on your Android phone or tablet. Retrieved October 28, 2026, from https://help.netflix.com/en/node/23939

Netflix (n.d.-e). Open Connect Overview. Retrieved October 28, 2026, from https://openconnect.netflix.com/Open-Connect-Overview.pdf

Netflix (n.d.-f). Profile transfers. Retrieved October 28, 2026, from https://help.netflix.com/en/node/122698

Netflix (n.d.-g). Sharing your Netflix account. Retrieved October 28, 2026, from https://help.netflix.com/en/node/123277

Netflix Technology Blog (2017). Content Popularity for Open Connect. https://netflixtechblog.com/content-popularity-for-open-connect-b86d56f613b

Netflix, Inc. (2003). Annual Report 2002 (Form 10-K). https://www.sec.gov/Archives/edgar/data/1065280/000095016803001155/d10k.htm

Netflix, Inc. (2026a). Annual Report 2025 (Form 10-K). https://www.sec.gov/Archives/edgar/data/1065280/000106528026000034/nflx-20251231.htm

Netflix, Inc. (2026b). Q4 2025 Letter to Shareholders (January 20, 2026). https://s22.q4cdn.com/959853165/files/doc_financials/2025/q4/FINAL-Q4-25-Shareholder-Letter.pdf

Netflix, Inc. (n.d.). Top Investor Questions. Retrieved October 28, 2026, from https://ir.netflix.net/ir-overview/top-investor-questions/

Peters, G. (2022). Netflix Starting From $6.99 a Month. https://about.netflix.com/en/news/announcing-basic-with-ads-us

Reinhard, A. (2025). Netflix's Third Season of Ads and a Look Ahead at What's Next. https://about.netflix.com/en/news/netflix-third-season-of-ads-and-a-look-ahead-at-whats-next-uk

Sarandos, T. (2022). Celebrating 10 Years of Netflix Series With 'Lilyhammer'. https://about.netflix.com/en/news/celebrating-10-years-of-netflix-series-with-lilyhammer

Sarandos, T. (2023). Netflix DVD – The Final Season. https://about.netflix.com/en/news/netflix-dvd-the-final-season

Teixeira, R., Li, Z., Mahajan, R., & Wei, W. (2026). Smarter Live Streaming at Scale: Rolling Out VBR for All Netflix Live Events. https://netflixtechblog.com/smarter-live-streaming-at-scale-rolling-out-vbr-for-all-netflix-live-events-c8f833b238cc

Verdu, M. (2021). Let the Games Begin: A New Way to Experience Entertainment on Mobile. https://about.netflix.com/en/news/let-the-games-begin-a-new-way-to-experience-entertainment-on-mobile

Vora, M., Berglund, A., Sadafal, V., Pfitzner, D., & Livengood, E. (2017). Distributing Content to Open Connect. https://netflixtechblog.com/distributing-content-to-open-connect-3e3e391d4dc9

Approfondimenti

- Valentina Re, Che cos'è Netflix, Carocci, 2026 — Un saggio che ricostruisce la storia culturale, economica e tecnologica di Netflix, analizzando il rapporto tra l'evoluzione della piattaforma e le trasformazioni dell'industria audiovisiva. L'autrice esamina come il passaggio dalla distribuzione fisica dei DVD allo streaming abbia modificato non soltanto le modalità di accesso ai contenuti, ma anche la produzione cinematografica e televisiva. Il libro è particolarmente pertinente alle sezioni storiche ed economiche della dispensa, perché permette di collocare le scelte di Netflix in un contesto più ampio. Offre inoltre una prospettiva critica sul rapporto tra piattaforme digitali, produzione originale, distribuzione e nuove abitudini di consumo audiovisivo.

- Netflix, Open Connect, Documentazione ufficiale Netflix — Una presentazione dell'infrastruttura proprietaria utilizzata da Netflix per distribuire contenuti audiovisivi attraverso Internet. La documentazione approfondisce il rapporto tra Netflix e gli Internet Service Provider, illustrando il ruolo degli Open Connect Appliances e delle interconnessioni dirette nella distribuzione locale del traffico. È particolarmente pertinente alla sezione della dispensa dedicata alla separazione tra Control Plane e Media Plane, perché permette di comprendere come Netflix possa gestire servizi e logica applicativa nel cloud affidando contemporaneamente la distribuzione dei contenuti a una CDN specializzata. Offre inoltre una base per approfondire le differenze tra server di origine, edge server e dispositivi collocati nelle reti degli operatori.
- Netflix, In che modo Netflix trasmette video in streaming, Centro assistenza Netflix — Una spiegazione del percorso attraverso cui i contenuti audiovisivi vengono trasferiti dall'infrastruttura Netflix ai dispositivi degli spettatori. Il documento descrive il ruolo dei server, delle reti degli operatori Internet e delle connessioni domestiche nella distribuzione dei dati video, evidenziando come la congestione possa influenzare le prestazioni della riproduzione. È particolarmente pertinente alla descrizione della CDN Open Connect e del processo attraverso cui un client riceve i contenuti. Permette inoltre di comprendere perché la qualità dello streaming non dipenda esclusivamente dalla capacità dei server Netflix, ma dall'intero percorso di rete che collega la piattaforma al dispositivo dell'utente.
- Amazon Web Services, MediaConvert – Guida per l'utente, Documentazione AWS — Un manuale tecnico dedicato alla preparazione dei contenuti audiovisivi attraverso processi di elaborazione e transcodifica basati sul cloud. La documentazione descrive come un master video possa essere trasformato in differenti rappresentazioni, utilizzando codec, risoluzioni, bitrate e formati di distribuzione diversi. È particolarmente pertinente alla sezione della dispensa dedicata alla Content Preparation di Netflix, soprattutto per comprendere il rapporto tra encoding, transcoding, packaging e Adaptive Bitrate Streaming. Pur descrivendo un servizio AWS e non l'intera pipeline proprietaria di Netflix, fornisce un esempio tecnico concreto delle operazioni necessarie per preparare un catalogo audiovisivo alla distribuzione su dispositivi eterogenei.
- Netflix, Funzionamento del sistema di consigli di Netflix, Centro assistenza Netflix — Una spiegazione ufficiale dei principi utilizzati dalla piattaforma per personalizzare la scoperta dei contenuti. Netflix descrive come il sistema consideri informazioni relative alla cronologia di visione, alle valutazioni, ai comportamenti di utenti con preferenze simili e alle caratteristiche dei titoli disponibili. Il documento approfondisce inoltre la personalizzazione della homepage, compresa la selezione delle righe, dei contenuti presenti al loro interno e del relativo ordinamento. È particolarmente pertinente alla sezione della dispensa dedicata ai sistemi di raccomandazione, perché permette di comprendere come Machine Learning, dati comportamentali e ranking collaborino per trasformare un catalogo molto ampio in una selezione di contenuti potenzialmente rilevanti per ciascun utente.
- Netflix, Come ottenere la massima qualità video, Centro assistenza Netflix — Una guida tecnica sui fattori che determinano la qualità audiovisiva effettivamente disponibile durante la riproduzione. Il documento approfondisce il rapporto tra risoluzione, HDR, piano di abbonamento, caratteristiche del dispositivo, condizioni della rete e compatibilità delle connessioni video. È particolarmente pertinente alle

sezioni della dispensa dedicate all'Adaptive Bitrate Streaming, alla selezione delle rappresentazioni e al funzionamento del client. Permette inoltre di osservare concretamente perché la qualità video non dipenda soltanto dal bitrate disponibile, ma anche dalle capacità hardware, dalle autorizzazioni del piano e dai requisiti di protezione dei contenuti, inclusa la compatibilità HDCP.

- Netflix, Pubblicità su Netflix, Centro assistenza Netflix — Una documentazione ufficiale dedicata al funzionamento dell'esperienza supportata dalla pubblicità e alle differenze rispetto ai piani senza interruzioni pubblicitarie ordinarie. Netflix descrive la collocazione degli annunci, alcuni criteri utilizzati per la loro selezione, la compatibilità dei dispositivi e le particolarità della pubblicità durante gli eventi live. È particolarmente pertinente alle sezioni della dispensa dedicate alla trasformazione del modello economico da SVOD a un sistema ibrido e all'integrazione della pubblicità nella pipeline tecnica. Permette di comprendere perché il piano dell'utente influenzi le decisioni del Control Plane e come la gestione degli annunci introduca ulteriori requisiti relativi ad autorizzazioni, disponibilità dei contenuti, targeting e riproduzione.
- Microsoft, Panoramica di PlayReady, Microsoft Learn — Un approfondimento tecnico sui sistemi Digital Rights Management utilizzati per controllare l'accesso e l'utilizzo dei contenuti audiovisivi protetti. La documentazione introduce il funzionamento delle licenze, delle chiavi crittografiche, dei livelli di sicurezza e delle restrizioni relative alla riproduzione. È particolarmente pertinente alla sezione della dispensa dedicata al Client Playback, perché permette di comprendere come il contenuto cifrato distribuito attraverso una CDN possa essere riprodotto soltanto quando il dispositivo dispone delle autorizzazioni necessarie. Pur riguardando PlayReady e non tutti i DRM utilizzati da Netflix, il testo offre una base per comprendere l'architettura multi-DRM richiesta da un servizio disponibile su browser, smartphone, console e Smart TV differenti.
- Amazon Web Services, Come funziona AWS Elemental MediaLive, Documentazione AWS — Una documentazione tecnica dedicata all'elaborazione dei flussi audiovisivi live attraverso un'infrastruttura cloud. Il testo descrive il rapporto tra sistemi upstream, acquisizione dei segnali, transcodifica in tempo reale e sistemi downstream incaricati della successiva distribuzione. È particolarmente pertinente alla parte finale della dispensa, che analizza la differenza tra l'architettura Video on Demand e quella utilizzata da Netflix per gli eventi live. Permette di comprendere il ruolo di AWS MediaLive nella produzione di rappresentazioni video con caratteristiche differenti e perché una piattaforma live debba coordinare continuamente acquisizione, encoding, packaging, distribuzione e riproduzione mentre l'evento è ancora in corso.

Nota editoriale

Per la realizzazione di questi materiali ho utilizzato strumenti di intelligenza artificiale generativa come supporto alla scrittura, in particolare per migliorare la forma del testo, riorganizzare i contenuti, correggere la formulazione e velocizzare alcune attività redazionali.

Lavorando autonomamente alla produzione di questi materiali, cerco infatti di automatizzare tutte le attività che possono esserlo, così da poter dedicare più tempo alla ricerca, alla progettazione e allo sviluppo dei contenuti.

L'intelligenza artificiale non determina però i contenuti dell'opera: la scelta degli argomenti, la struttura, le idee, le interpretazioni, gli esempi e l'impostazione didattica sono elaborati da me. L'IA viene quindi utilizzata principalmente come strumento di supporto alla produzione e alla revisione formale, mentre la responsabilità autoriale e progettuale dei contenuti rimane mia.

Salvo diversa indicazione, questo materiale è distribuito con licenza Creative Commons Attribuzione – Non commerciale – Condividi allo stesso modo 4.0 Internazionale (CC BY-NC-SA 4.0).

È quindi possibile condividere, ridistribuire, modificare e creare opere derivate a partire dal materiale, a condizione di attribuirne correttamente la paternità, di non utilizzarlo per scopi commerciali e di distribuire eventuali versioni modificate o derivate con la stessa licenza.

