How Netflix directs 1/3rd of Internet Traffic

QCon
San Francisco
Nov 16, 2015

Haley Tucker
Mohit Vora
This is Washington. There's always a leak. All 13 episodes will launch February 27.
CONGRATULATIONS
EMMY® AWARD WINNER

NETFLIX ORIGINAL
| ORANGE is the new BLACK |

Praise Norma! It’s another Emmy win for Uzo Aduba.
How do we build a streaming “tape”?
Determine the preferred experience

**TITLE**

**COUNTRY**

**DEVICE**

**CUSTOMER**

**NETWORK**

**CONNECTIONS**

- Broadband - wired or wifi
- Cellular - Edge, 3G, LTE, ...

- Dansk
- Deutsch
- English
- Español
- Français
- Nederlands
- Norsk bokmål
- Português
- Suesi
- Svenska
That’s exactly what I want
...now where can I get it?
Point the device to appropriate locations
Uh-oh, the content is encrypted!
And...Action!
SESSION EVENTS

SESSION (START, STOP, PAUSE, RESUME, KEEPALIVE)
PLAYBACK LIFECYCLE

- PLAY
- LICENSE
- SESSION (START, STOP, PAUSE, RESUME, KEEPALIVE)
- GENERATE PLAYBACK MANIFEST
- PLAYBACK MANIFEST
What is a Content Delivery Network?
Open
Connect
A NETFLIX ORIGINAL
PREDICTABLE VIEWING PATTERNS
FILLING WHEN YOU SLEEP
READ XOR WRITE
ADVENTURE... EXCITEMENT...

A JEDI CRAVES NOT THESE THINGS
Rodents of unusual size? I don't think they exist.
Content Delivery Mechanisms
OPEN CONNECT

NETFLIX DEVICE

DEVICE CONTROL PLANE

STRAIN

CDN CONTROL PLANE

DON'T KEEP SECRETS

Network Proximity
Content Positioning
Load Distribution
Network Proximity
By Specification?
By Specification?

Doesn’t scale
Border Gateway Protocol

BGP ROUTE
175.231.128.0/24
(+ proximity attributes)
BGP ROUTE
175.231.128.0/24
(+ proximity attributes)
Content Positioning
LOCALIZE TRAFFIC
HOW DO WE DETERMINE WHAT CONTENT WILL BE POPULAR TOMORROW?
Evolving Member Tastes
MINIMIZE FILL CHURN
USE HISTORICAL DATA
\[\text{bytesStreamed}/\text{bytesStored}\]
IS ONE DAY OF HISTORY ENOUGH?
**TAKEAWAY** Weigh Recent Data Higher

\[
p(n) = p_0 \alpha^n + p_1(1-\alpha) + p_2(\alpha^2) + \ldots + p_n \alpha^n
\]

\[
p(0) = p_0 \alpha^0 + p_1(1-\alpha) + p_2(\alpha^2) + \ldots + p_n \alpha^n
\]

**EXPONENTIALLY WEIGHTED MOVING AVERAGE**
HOW SHOULD CONTENT BE ALLOCATED?
HOW SHOULD CONTENT BE ALLOCATED?

MILLIONS
OF FILES

THOUSANDS
OF SERVERS
TAKEAWAY

Consistent Hashing

ALLOCATE MULTIPLE REPLICA
RESILIENT TO CLUSTER CHANGES
REPEATABLE
Load Distribution
CONTENT WITH CONFLICTING CONSTRAINTS
TAKEAWAY Tier Infrastructure

WITHIN CLUSTERS

ON EACH SERVER
HOW DO WE BALANCE LOAD?

ACROSS EQUIDISTANT CLUSTERS

ACROSS SERVERS WITHIN CLUSTERS
HOW DO WE BALANCE LOAD?

USING CONTENT DISTRIBUTION
AND WHEN WE HAVE EQUALLY ATTRACTIVE LOCATIONS TO SERVE FROM –

FLIP A COIN
HOW DO WE LOAD SERVERS OPTIMALLY?
... AMIDST EVER CHANGING INTERNET WEATHER
AND DAILY TRAFFIC EBBS AND FLOWS
WE INTRODUCE A FEEDBACK LOOP
TAKEAWAY | PID CONTROLLER
TAKEAWAY | PID CONTROLLER

<table>
<thead>
<tr>
<th>Process Variable</th>
<th>Current RPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set Point</td>
<td>Desired RPM</td>
</tr>
<tr>
<td>Control Variable</td>
<td>Input Voltage</td>
</tr>
</tbody>
</table>

DC MOTOR
## TAKEAWAY

**PID CONTROLLER**

<table>
<thead>
<tr>
<th>Process Variable</th>
<th>Current RPM</th>
<th>System Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set Point</td>
<td>Desired RPM</td>
<td>System Metrics Max</td>
</tr>
<tr>
<td>Control Variable</td>
<td>Input Voltage</td>
<td>Controlled Traffic</td>
</tr>
</tbody>
</table>

**LOADING SERVERS**
0.0 < CONTROL VAR < 1.0
TRAFFIC SHIFTS TO NEXT HOP LOCATION
Steering
Got URLs for f1, f2, ..., fn?

Yes, here's the URLs

Got URLs for f1, f2, ..., fn?
Architecture Evolution
5 CHALLENGES
How did we evolve from here...
CHALLENGE

High dimensionality

TITLE

CUSTOMER

DEVICE

COUNTRY

NETWORK

CONNECTIONS

Broadband - wired or wifi
Cellular - Edge, 3G, LTE, ...

Dansk  Norsk bokmål
Deutsch  Português
English  Suomi
Español  Svenska
Français  日本語
Nederlands
How can we quickly alter the playback experience in a targeted manner?
USE CASE Stream Filtering

ALL STREAMS FOR CONTENT

RULES

ENGINE

BEST STREAMS FOR SESSION
EXAMPLE RULES

<item identity="1.0.76">
  <description>Filter out video bitrates > 1050 for my device.</description>
  <dimension type="ESN">NFCDCCH-MC-12345</dimension>
  <dimension type="DL_PROFILE_TYPE">VIDEO</dimension>
  <dimension type="DL_BITRATE" operation="GT">1050</dimension>
</item>

<item identity="1.0.75">
  <description>Filter out English audio description for my device.</description>
  <dimension type="ESN">NFCDCCH-MC-12345</dimension>
  <dimension type="DL_PROFILE_TYPE">AUDIO</dimension>
  <dimension type="DL_AUDIO_TYPE">Assistive</dimension>
  <dimension type="DL_AUDIO_LANGUAGE">en</dimension>
</item>
UPDATING RULES

CONFIGURATION MANAGEMENT UI

PUBLISH

TOPIC

RULES

SUBSCRIBE

ENGINE
TAKEAWAY Dynamic Business Rules
CHALLENGE | Pinpoint what is broken
3:00 AM: Pager goes off
METRICS AND ALERTING
OK...error code 105 is elevated. But why?
Detailed Domain Insights

API

MANIFEST

INSIGHTS

RULES

SESSION

LICENSE

STEERING

DRM

TAKEAWAY
Large amount of state
How can we enable faster UIs and low-end devices?
We introduced a server-side caching tier
Watch out for resiliency issues!!
TAKEAWAY: Reduce client state
Managing device protocols
Can we allow devices to define their own protocols?
TAKEAWAY

Client-driven protocols
Enabling high-velocity innovation
How can we expose new data with the least amount of churn?
This works from API:
* `stream.getBitrate()`
* `stream.getDynamicData().get("FRAME_RATE")`

Works both ways!
This works from CLIENT SCRIPT!

- `stream.getDynamicData().get("BIT_RATE")`
- `stream.getDynamicData().get("FRAME_RATE")`
TAKEAWAY | Data pass-thru

- CLIENT SCRIPTS
- SERVICE LAYER
- API
- MANIFEST
- INSIGHTS
- RULES
- SESSION
- LICENSE
- DRAM
- STEERING
- CACHE
TAKEAWAYS

- BGP based proximity
- Tiered Infrastructure
- PID Controller
- EWMA for historical data
- Consistent Hashing

- Dynamic business rules
- Detailed domain insights
- Reduce client state
- Client-driven protocols
- Data pass-thru
TAKEAWAYS

- BGP based proximity
- Tiered Infrastructure
- PID Controller
- EWMA for historical data
- Consistent Hashing

- Dynamic business rules
- Detailed domain insights
- Reduce client state
- Client-driven protocols
- Data pass-thru

Haley Tucker  @hwilson1204  Mohit Vora  @mohitvora
Image Attributions

- Background image from https://www.flickr.com/photos/centralasian/4099515384, Image was cropped and red lines and dots were drawn on top, https://creativecommons.org/licenses/by/2.0/.
- Image from https://www.flickr.com/photos/28705377@N04/4142872268, No modifications made, https://creativecommons.org/licenses/by/2.0/.
- Image of cassette is from https://www.flickr.com/photos/comedynose/6939206771, Image was cropped, https://creativecommons.org/licenses/by/2.0/.
- Image of speaker is from https://www.flickr.com/photos/av_hire_london/5578975575, No changes made, https://creativecommons.org/licenses/by/2.0/.
- Image of television is from https://www.flickr.com/photos/jvcamerica/3660897684/, No changes made, https://creativecommons.org/licenses/by/2.0/.
- Image of text is from https://www.flickr.com/photos/dno1967b/5754743006, No changes made, https://creativecommons.org/licenses/by/2.0/.
- Background image from https://www.flickr.com/photos/mcgraths/866572532, Image was cropped, https://creativecommons.org/licenses/by/2.0/.
- Image from https://www.flickr.com/photos/thatguyfromcchs08/2300190277, Image is dimmed, https://creativecommons.org/licenses/by/2.0/.
- Image from https://www.flickr.com/photos/mknowles/3134373590, Image was cropped, https://creativecommons.org/licenses/by-sa/2.0/.