SCALABLE DATA ARCHITECTURES

from thousands to billions of events

@philip_pfo

NETFLIX

Please read the notes associated with each slide for the full context of the presentation

NETFLIX

Who am I?

Philip Fisher-Ogden

- Director of Engineering @ Netflix
- Playback Services (making "click play" work)
- 6 years @ Netflix, from 10 servers to 10,000s

Story

Netflix streaming – 2007 to present

Device Growth

2007
1 device

XBOX 360

2008
10s of devices

PS3

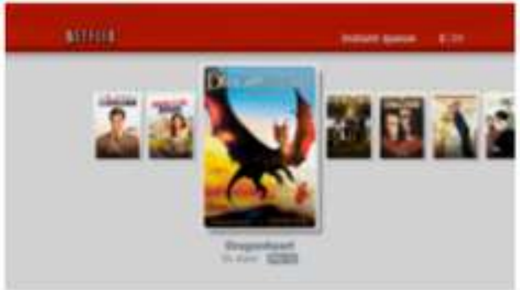
2009
10s of devices

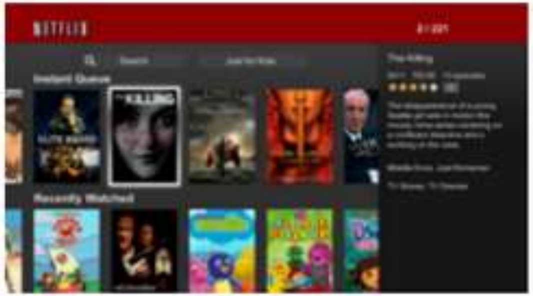
iPhone
iPad
Wii

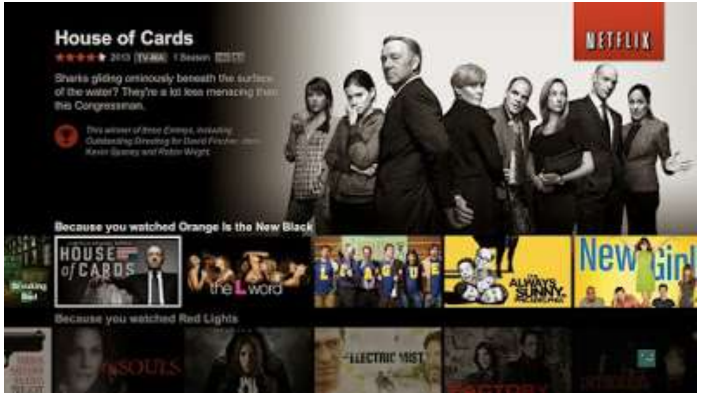
2010
100s of devices

2011+
1000+ devices

Experience Evolution






House of Cards

2013 TV-MA 1 Season

Sharks gliding ominously beneath the surface of the water? They're a lot less menacing than this Congressman.

Because you watched Orange Is the New Black

Because you watched Red Lights

NETFLIX

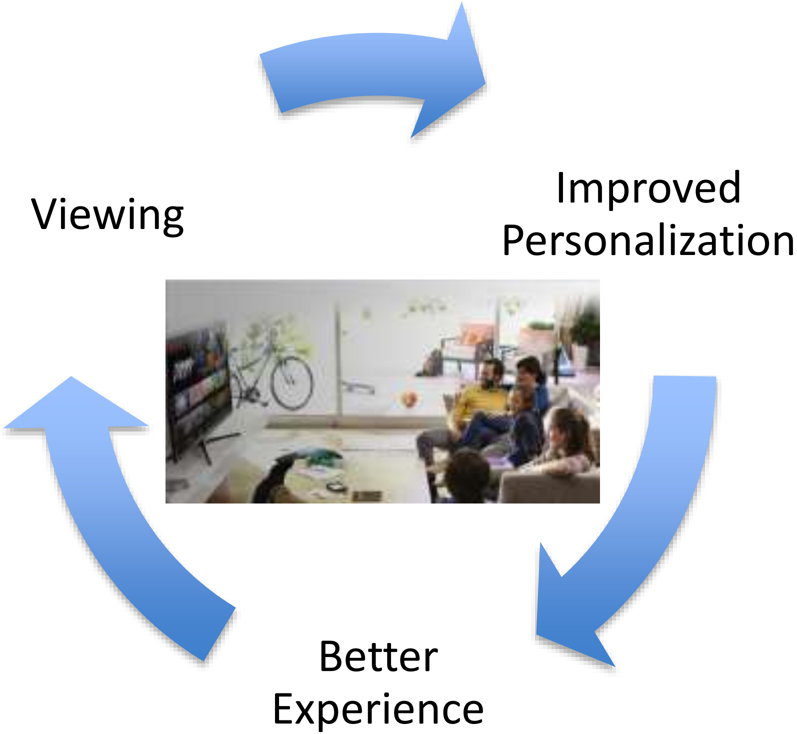
Subscribers & Viewing

53M global subscribers

50 countries

>2 billion hours viewed per month

Virtuous Cycle



Viewing

Improved Personalization

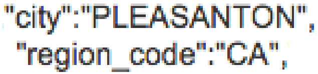
Better Experience

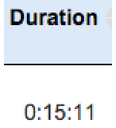
NETFLIX

Viewing Data

Who, What, When, Where, How Long

10/13/14


"city":"PLEASANTON",
"region_code":"CA",

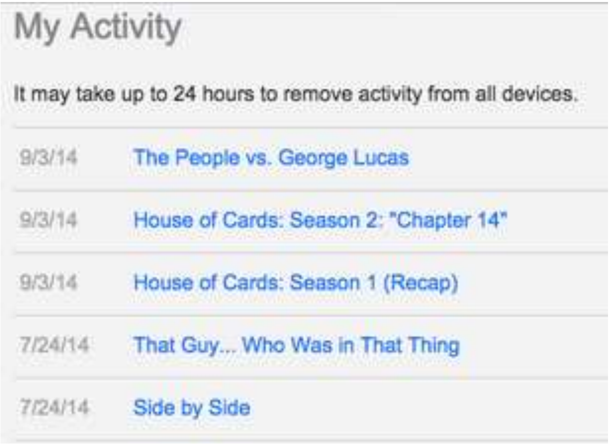


Duration
0:15:11

Latest Position

Real time data use cases

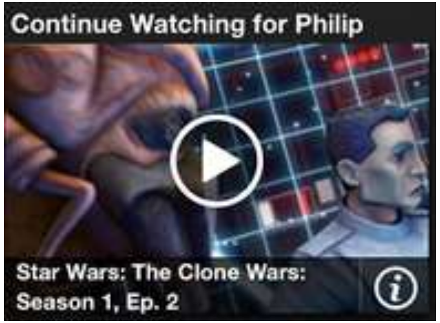
What have I watched?



My Activity

It may take up to 24 hours to remove activity from all devices.

9/3/14	The People vs. George Lucas
9/3/14	House of Cards: Season 2: "Chapter 14"
9/3/14	House of Cards: Season 1 (Recap)
7/24/14	That Guy... Who Was in That Thing
7/24/14	Side by Side

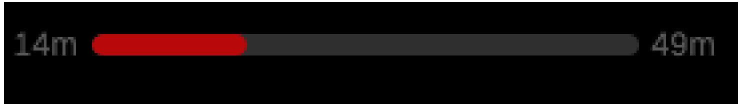


Continue Watching for Philip

Star Wars: The Clone Wars:
Season 1, Ep. 2

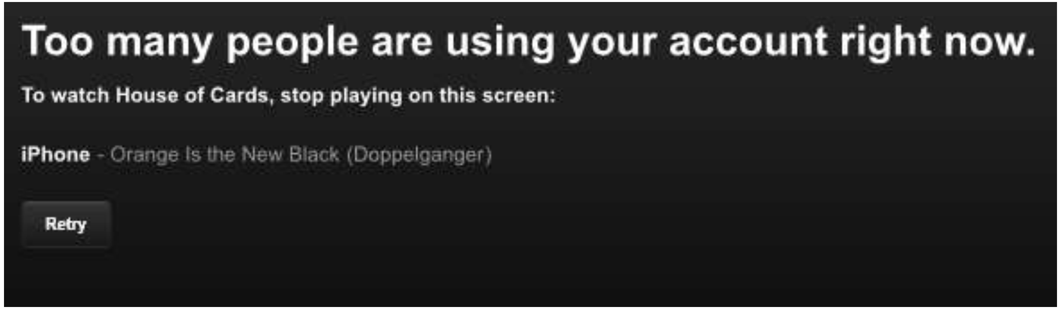
Real time data use cases

Where was I at?



Real time data use cases

What else am I watching?



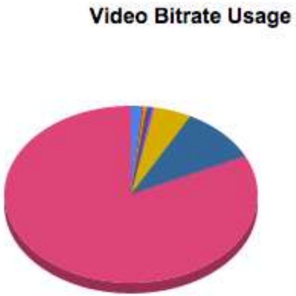
Too many people are using your account right now.

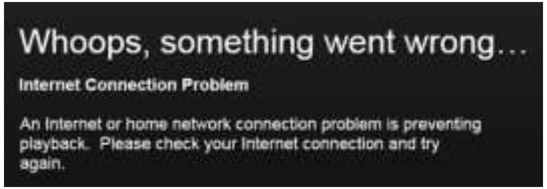
To watch House of Cards, stop playing on this screen:

iPhone - Orange Is the New Black (Doppelganger)

NETFLIX

Session Analytics

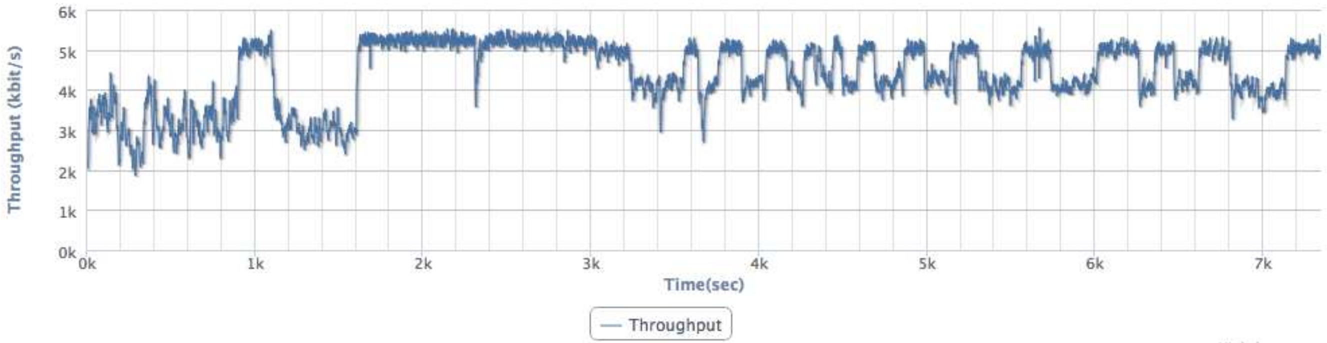


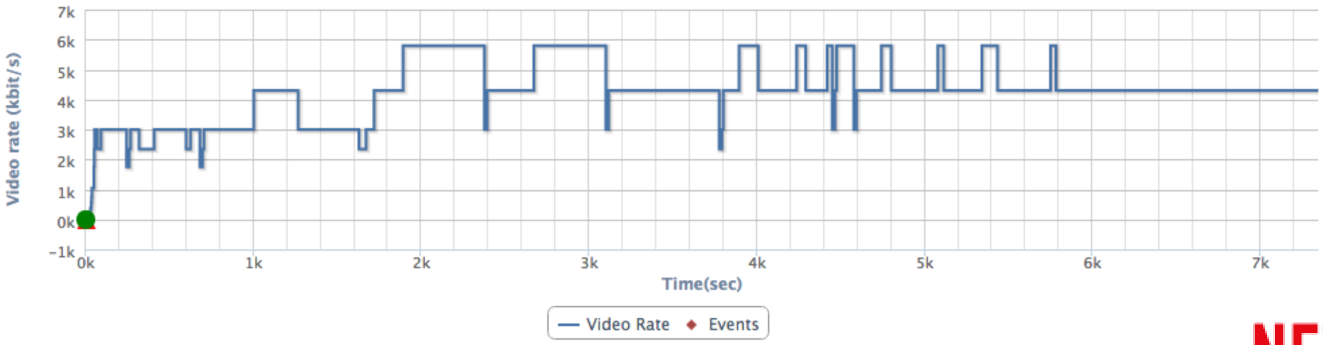

Whoops, something went wrong...

Internet Connection Problem

An Internet or home network connection problem is preventing playback. Please check your Internet connection and try again.

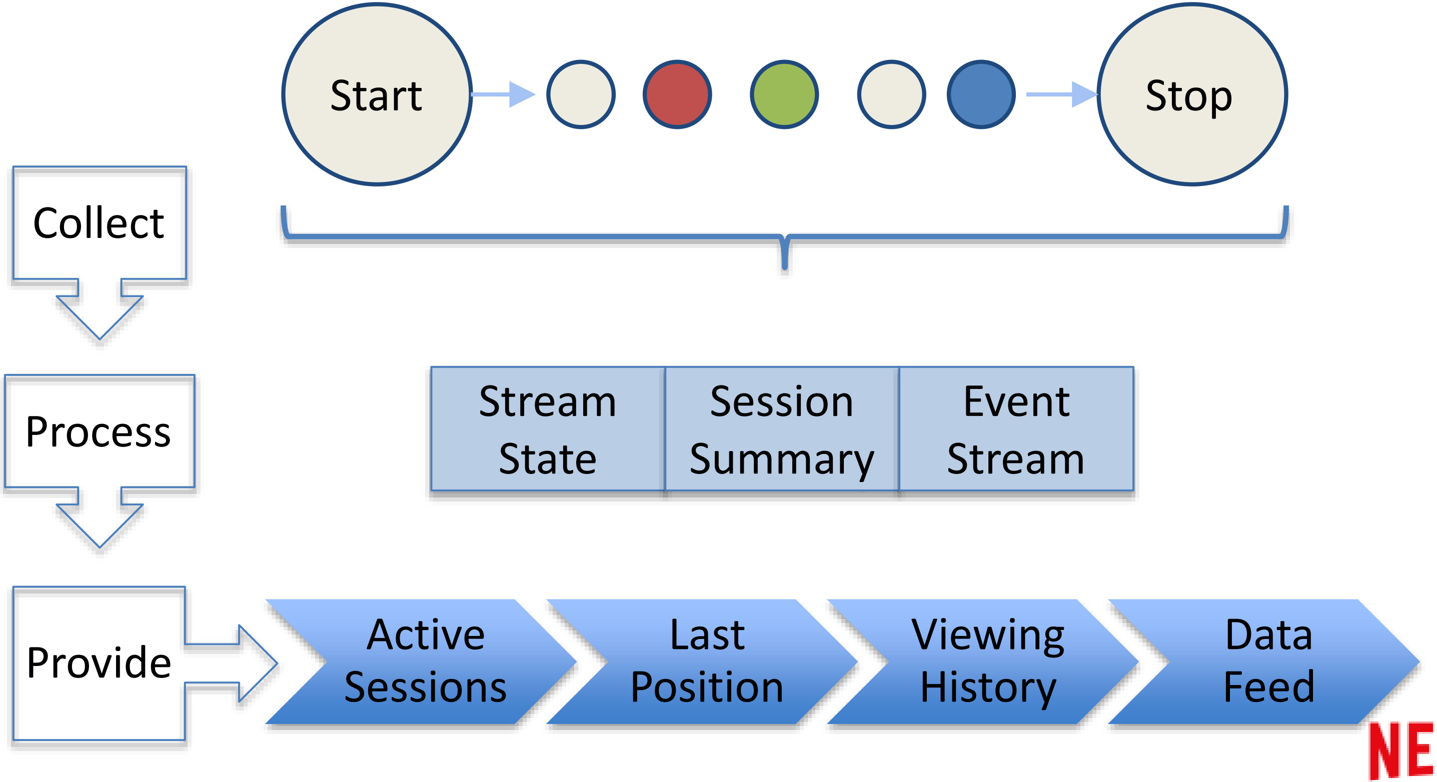
Session Analytics





NETFLIX

Generic Architecture

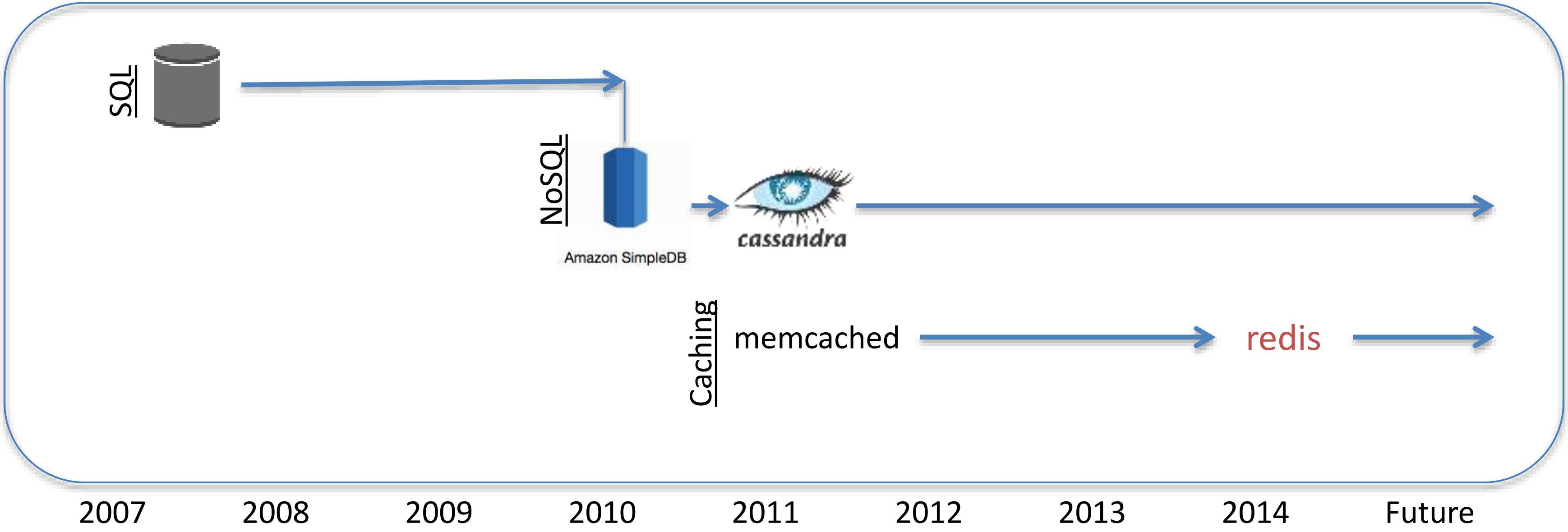


NETFLIX

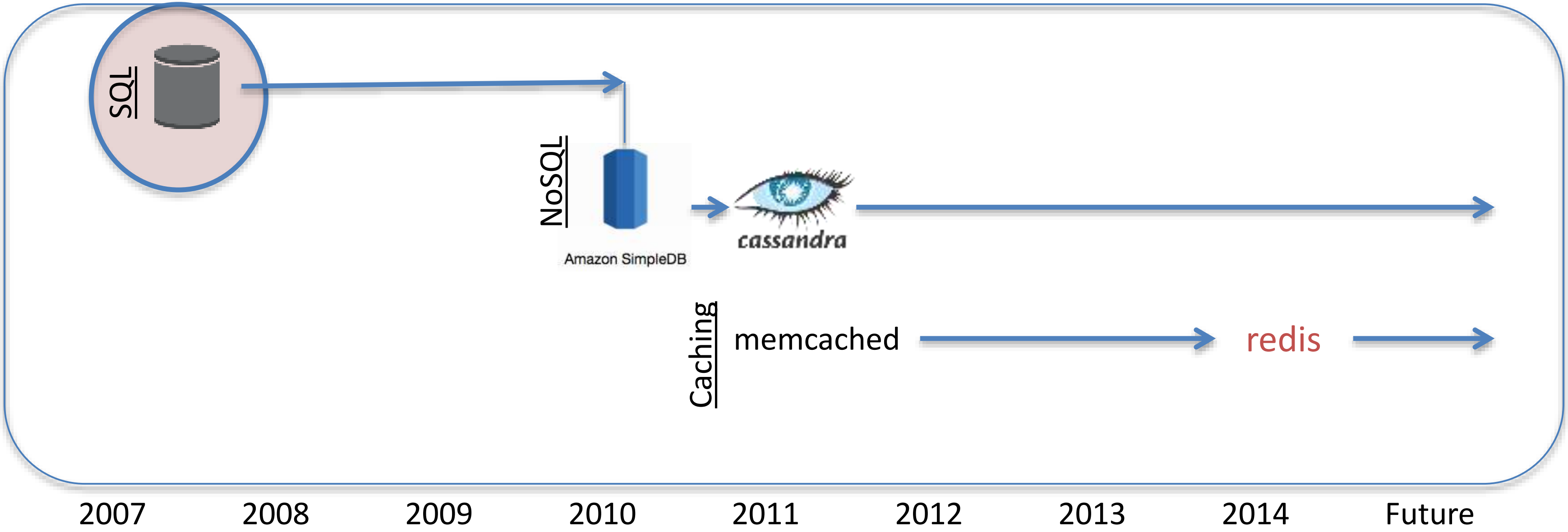
Architecture Evolution

- Different generations
- Pain points & learnings
- Re-architecture motivations

Real Time Data

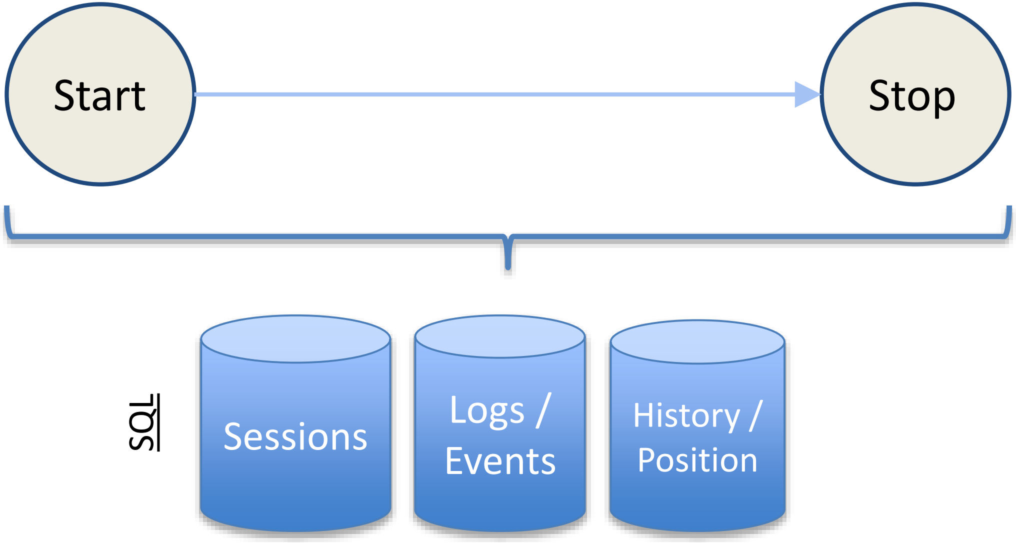


Real Time Data – gen 1



NETFLIX

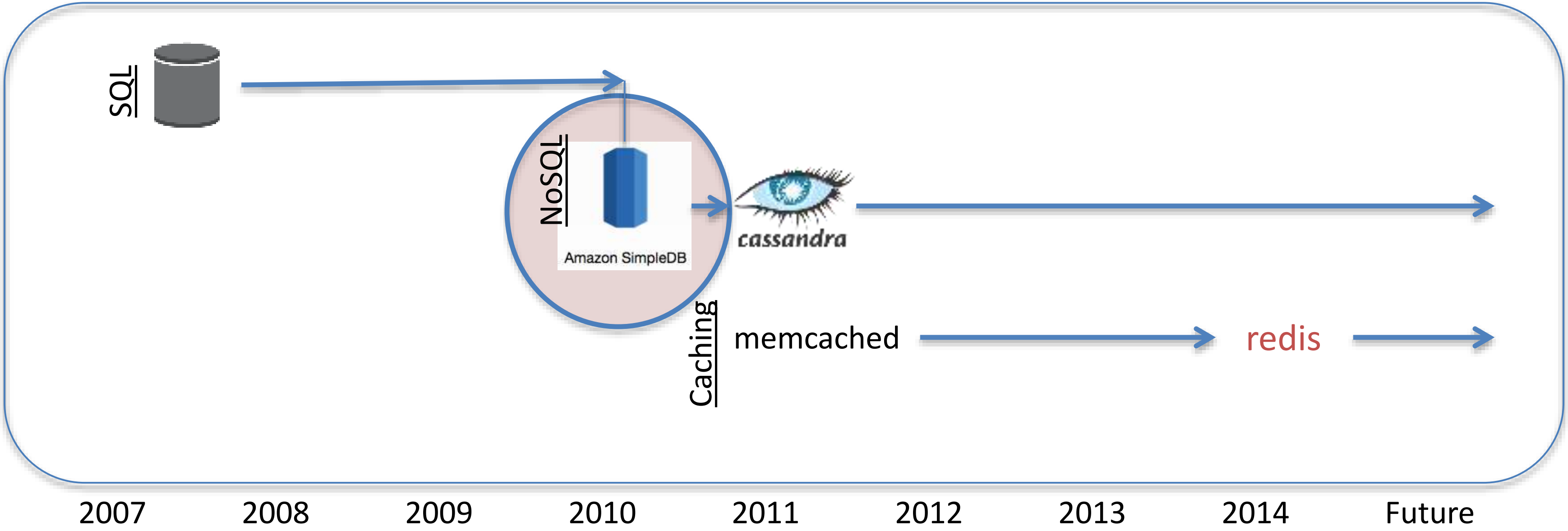
Real Time Data – gen 1



Real Time Data – gen 1 pain points

- Scalability
  - DB scaled up not out
- Event Data Analytics
  - ad hoc
- Fixed schema

Real Time Data – gen 2

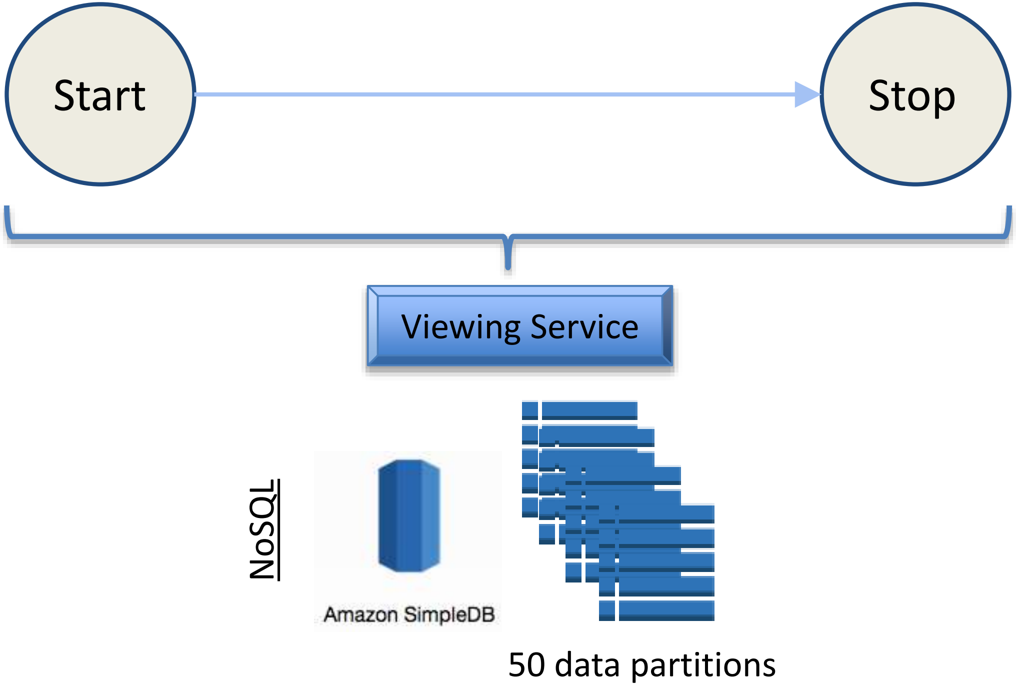


NETFLIX

Real Time Data – gen 2 motivations

- Scalability
  - Scale out not up
- Flexible schema
  - Key/value attributes
- Service oriented

Real Time Data – gen 2

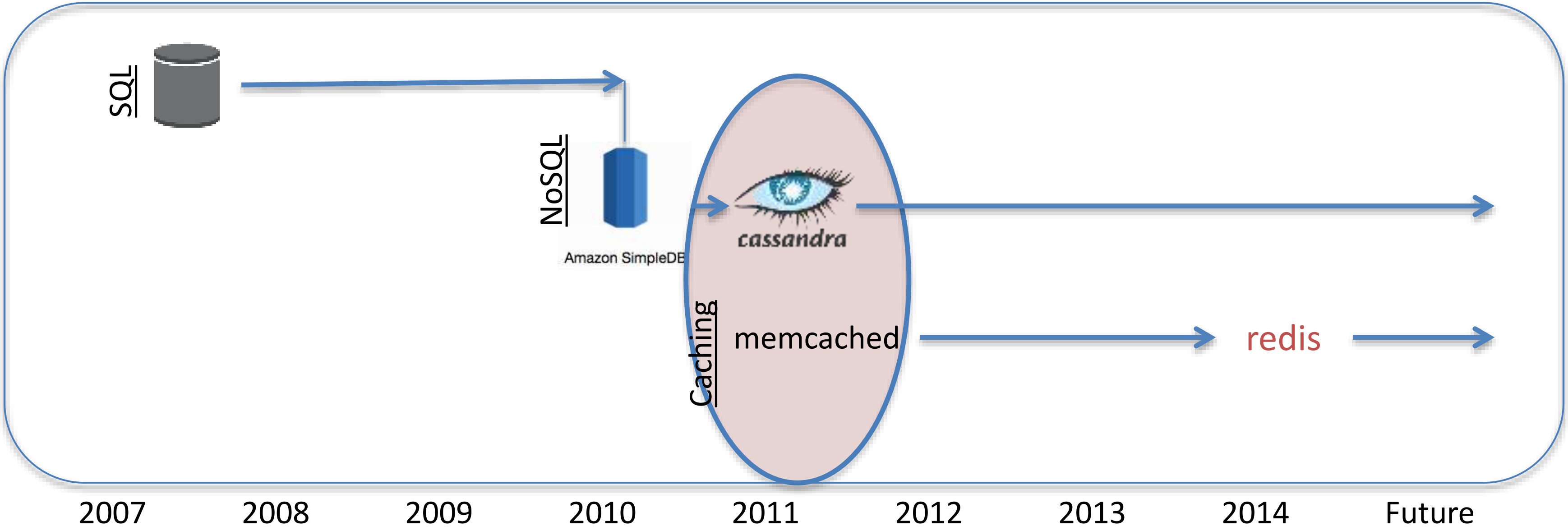


NETFLIX

Real Time Data – gen 2 pain points

- Scale out
  - Resharding was painful
- Performance
  - Hot spots
- Disaster Recovery
  - SimpleDB had no backups

Real Time Data – gen 3



Real Time Data – gen 3 landscape

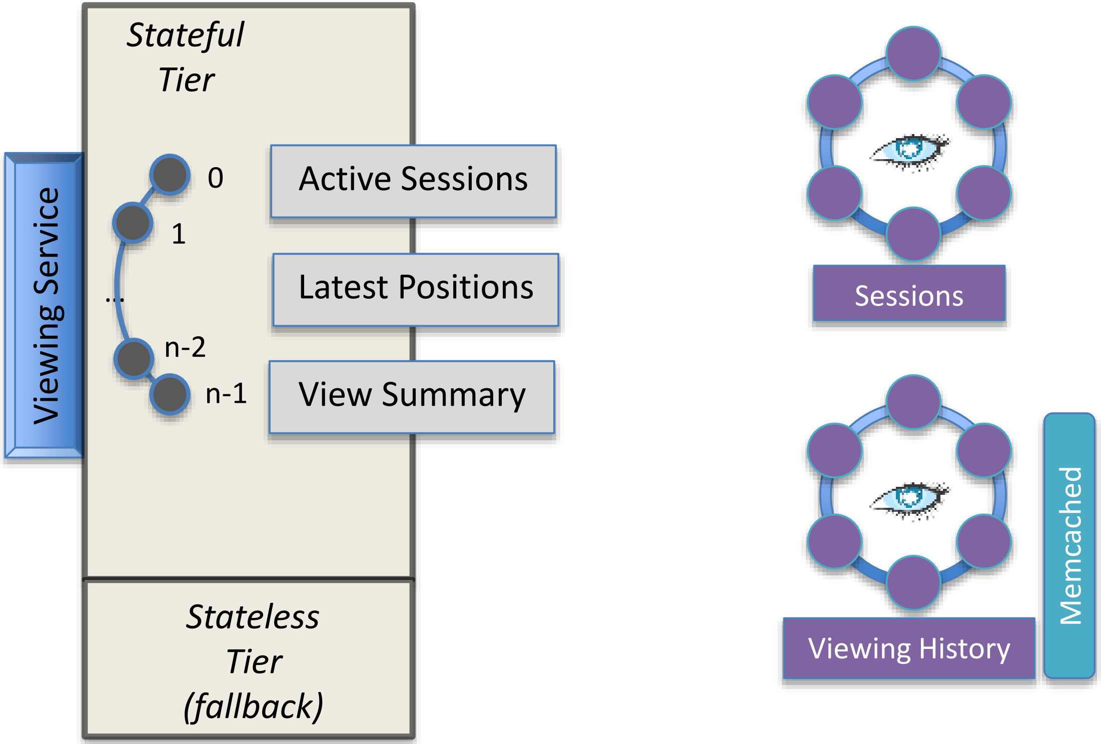
- Cassandra 0.6
- Before SSDs in AWS
- Netflix in 1 AWS region

Real Time Data – gen 3 motivations

- Order of magnitude increase in requests

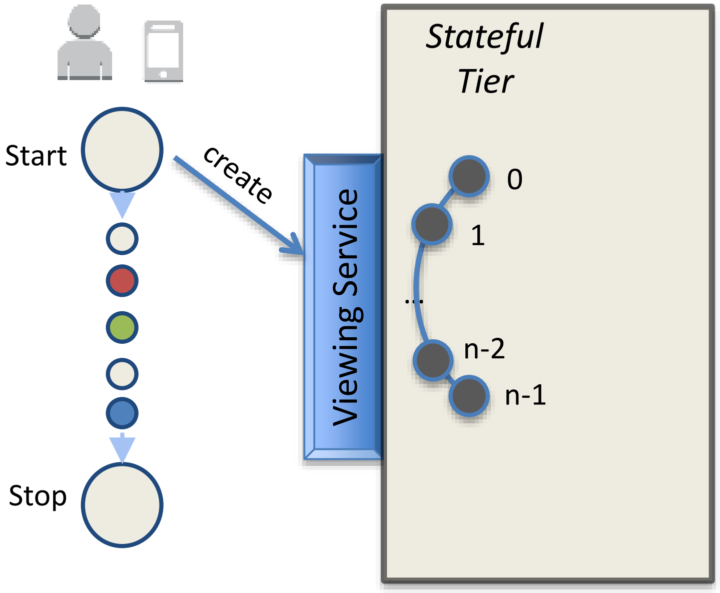
- Scalability
  - Actually scale out rather than up

Real Time Data – gen 3



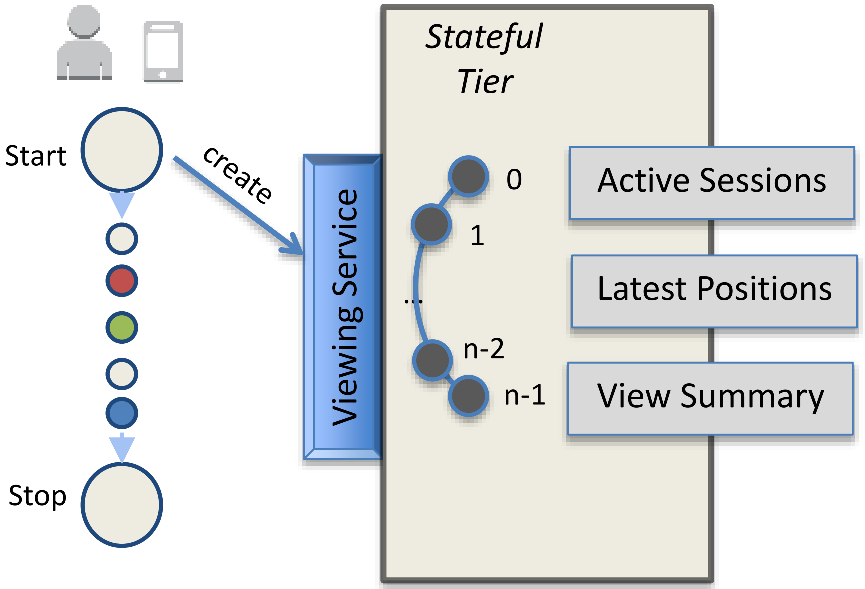
NETFLIX

Real Time Data – gen 3 writes



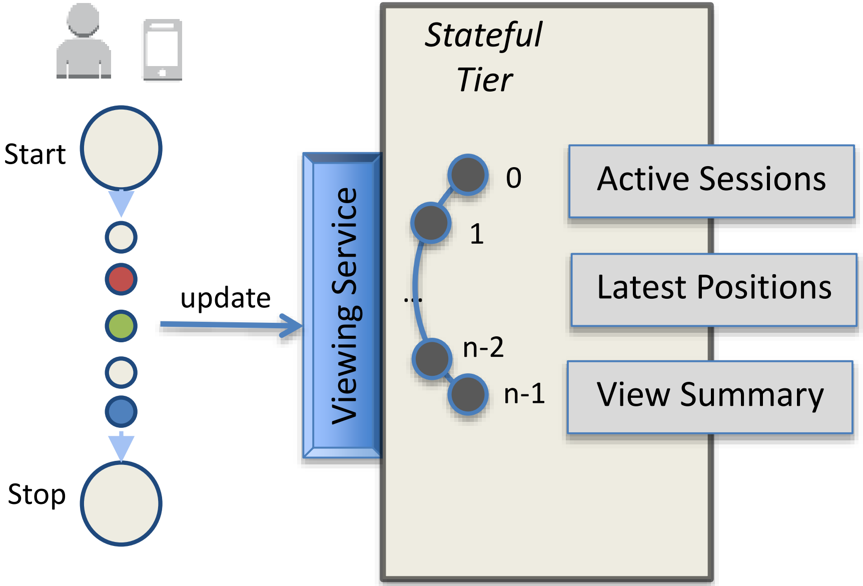
NETFLIX

Real Time Data – gen 3 writes

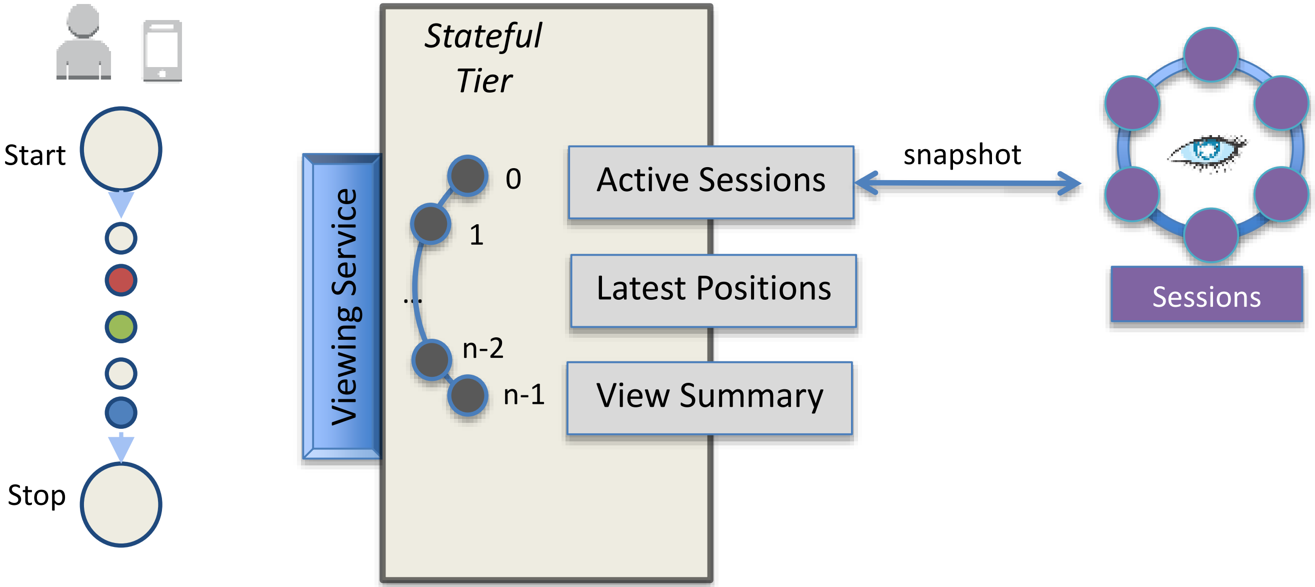


NETFLIX

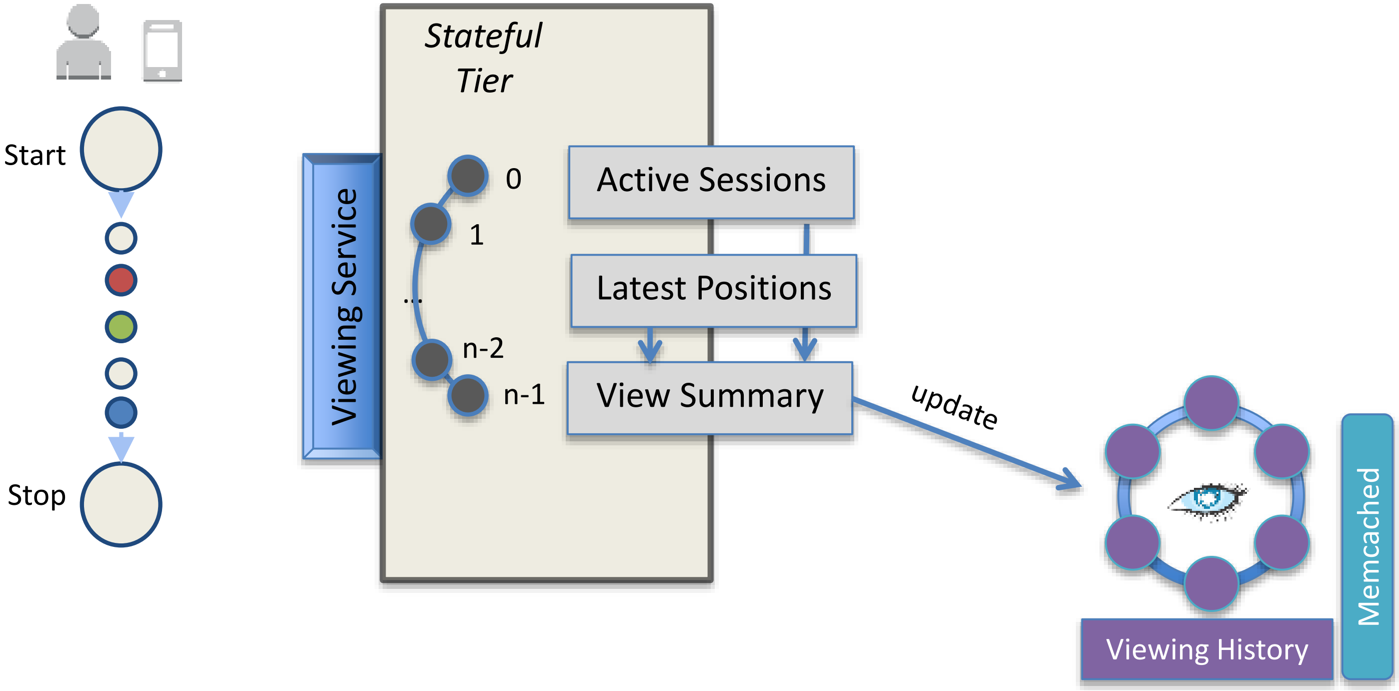
Real Time Data – gen 3 writes



Real Time Data – gen 3 writes

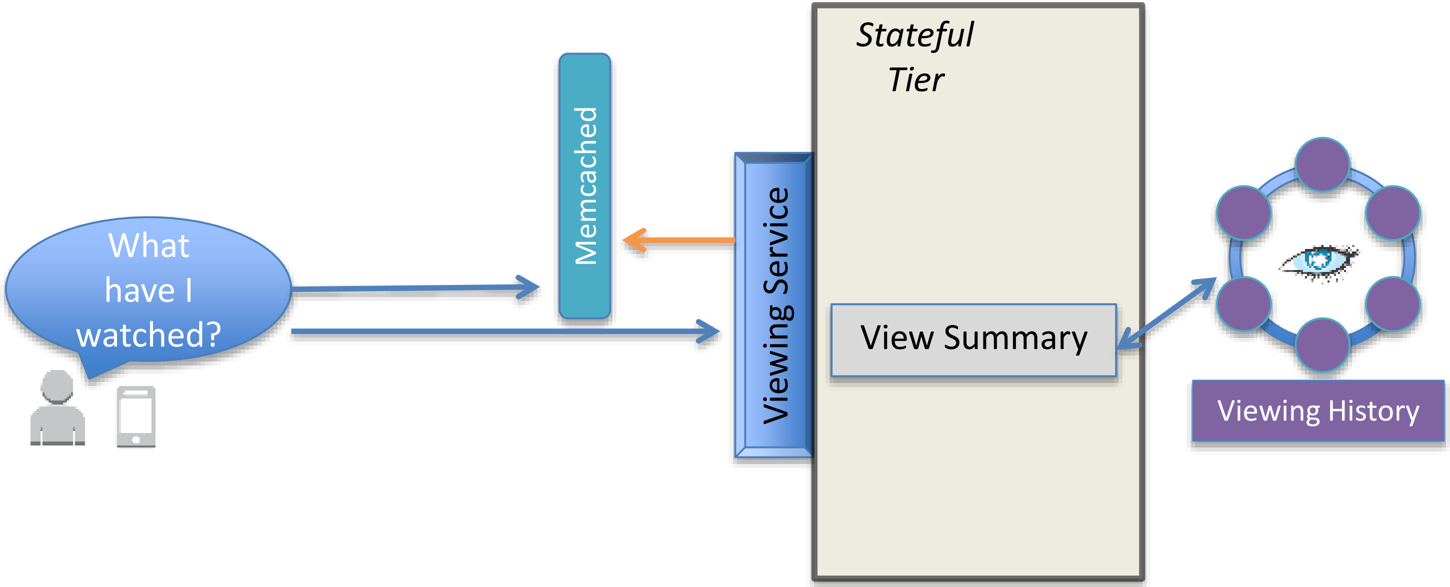


Real Time Data – gen 3 writes



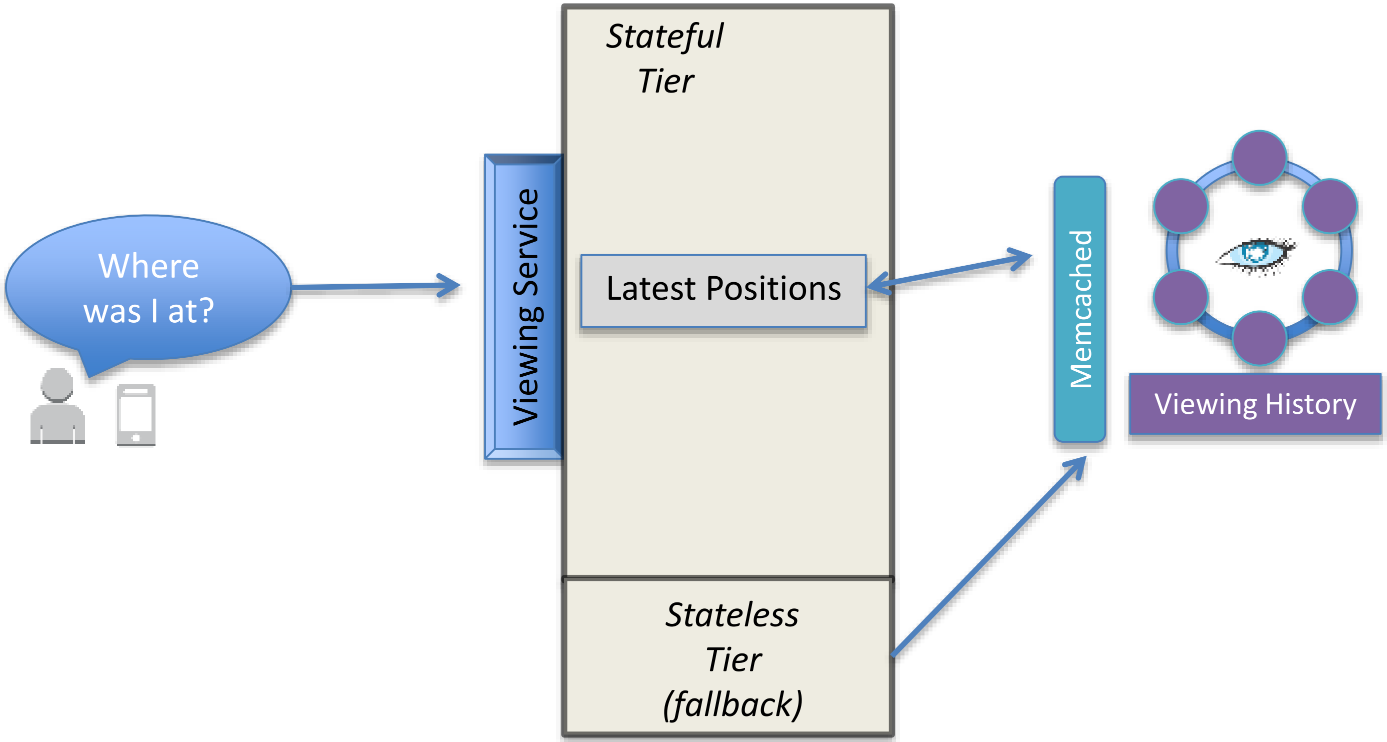
NETFLIX

Real Time Data – gen 3 reads



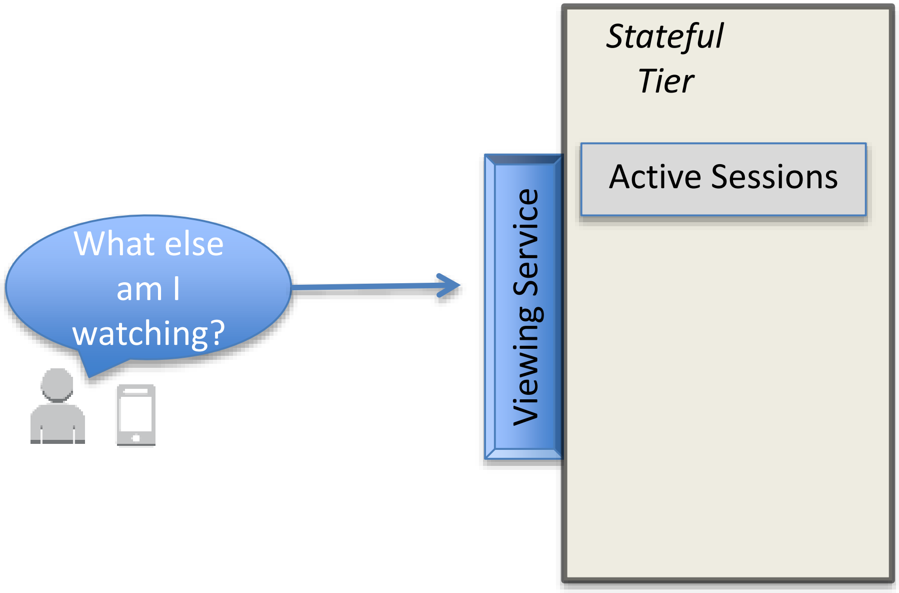
NETFLIX

Real Time Data – gen 3 reads



NETFLIX

Real Time Data – gen 3 reads



NETFLIX

gen 3 - Requests Scale

Operation	Scale
Create (start streaming)	1,000s per second
Update (heartbeat, close)	100,000s per second
Append (session events/logs)	10,000s per second
Read viewing history	10,000s per second
Read latest position	100,000s per second

gen 3 – Cluster Scale

Cluster	Scale
Cassandra Viewing History	~100 hi1.4xl nodes ~48 TB total space used
Viewing Service Stateful Tier	~1700 r3.2xl nodes 50GB heap memory per node
Memcached	~450 r3.2xl/xl nodes ~8TB memory used

Real Time Data – gen 3 pain points

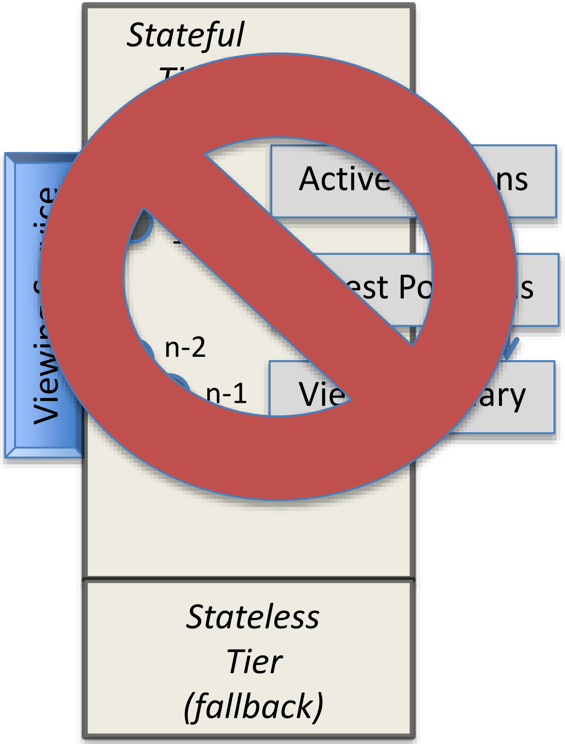
- Stateful tier
  - Hot spots
  - Multi-region complexity
- Monolithic service
- read-modify-write poorly suited for memcached

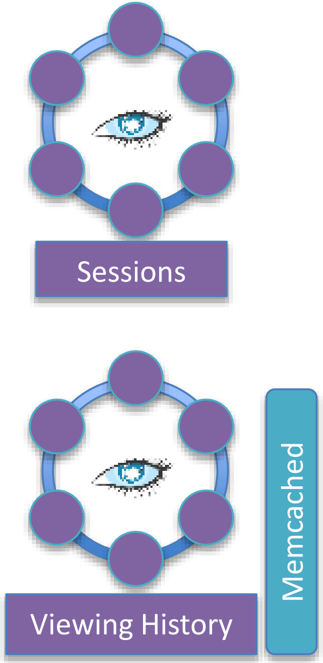
Real Time Data – gen 3 learnings

- Distributed stateful systems are hard
  - Go stateless, use C*/memcached/redis…
- Decompose into microservices

NETFLIX

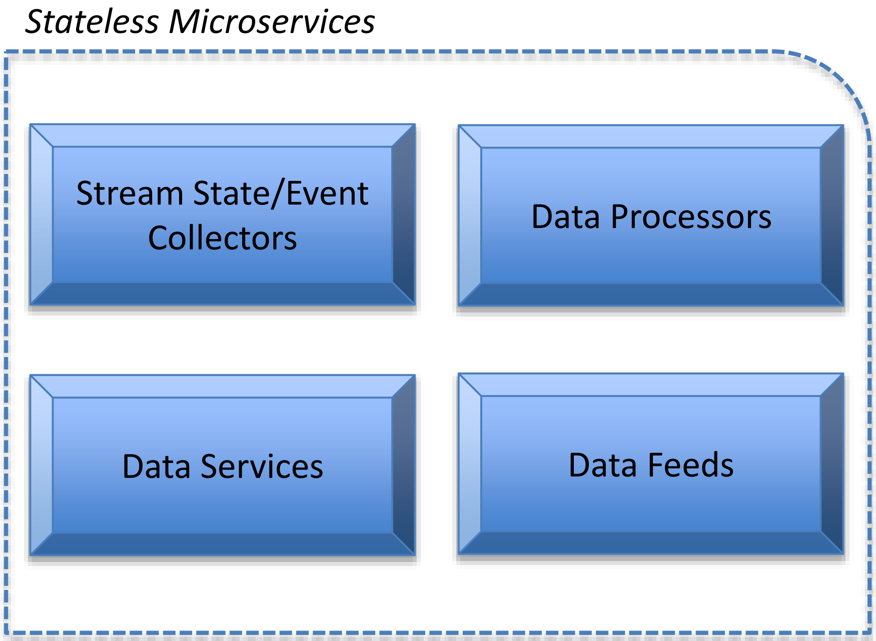
Real Time Data – gen 4





NETFLIX

Real Time Data – gen 4



Stateless Microservices

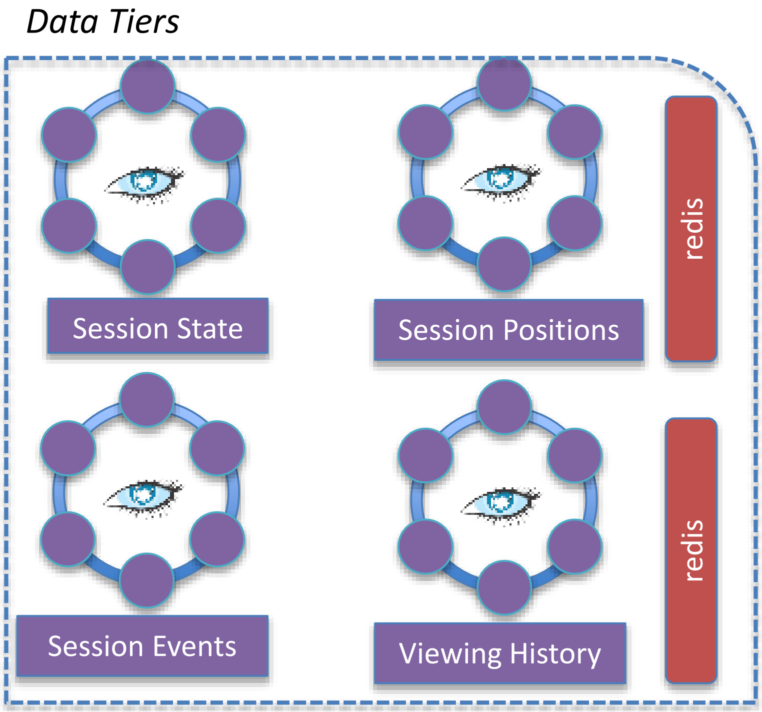
Stream State/Event Collectors

Data Processors

Data Services

Data Feeds

Real Time Data – gen 4

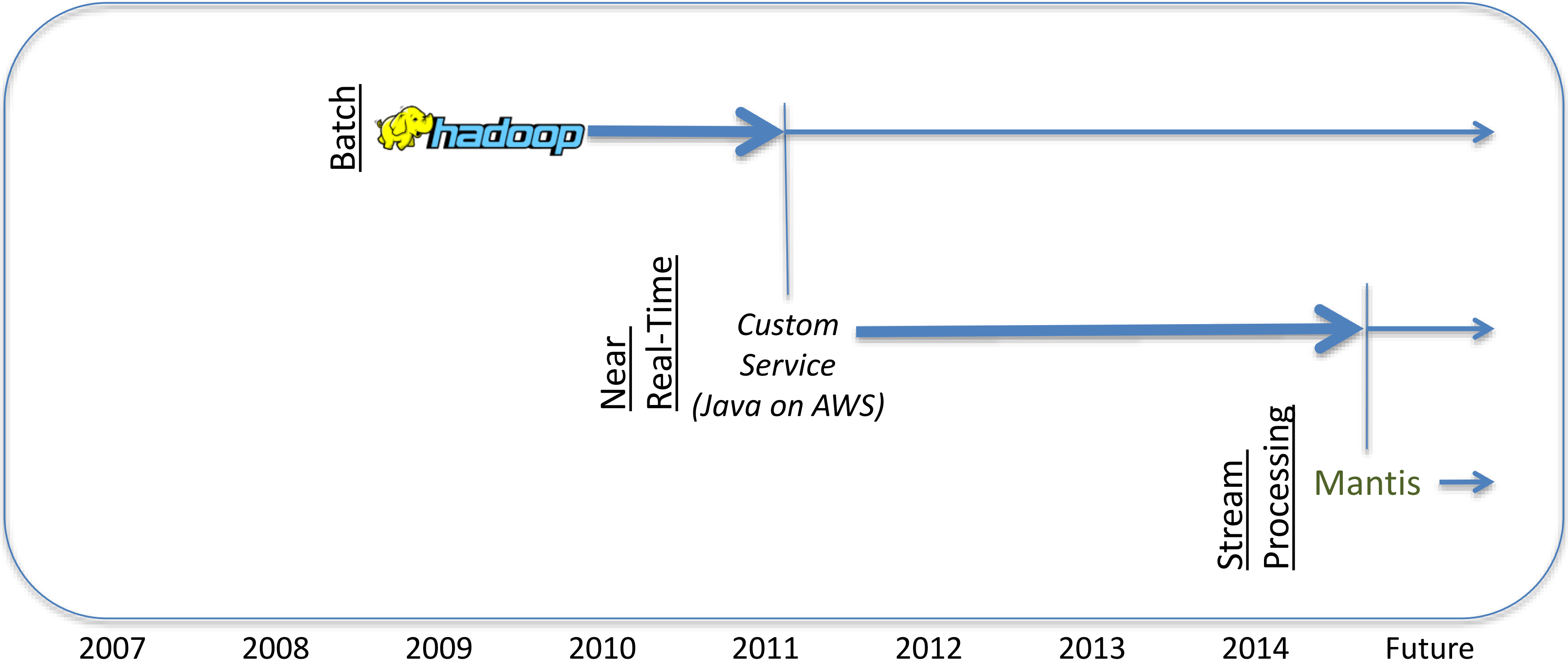


NETFLIX

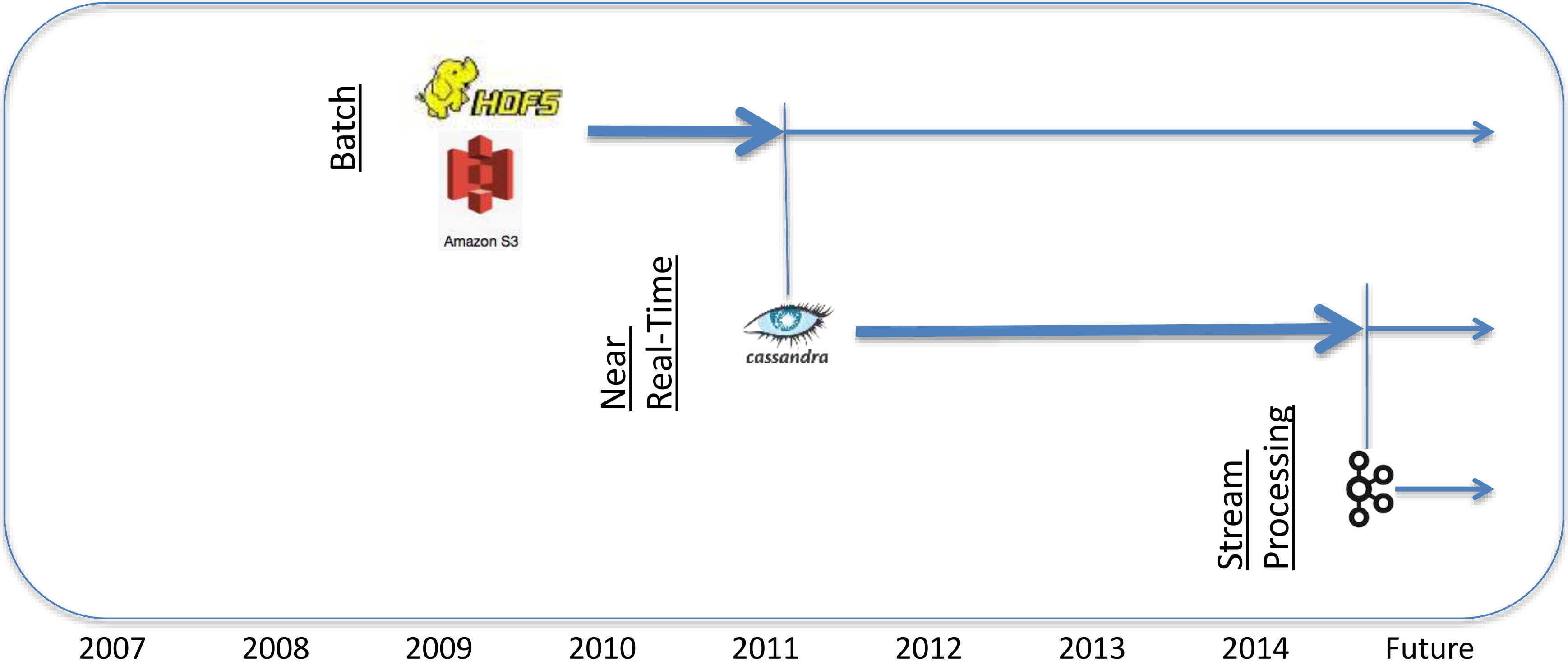
Session Analytics

- Summarize detailed event data
- Non-real time, but near real time
- Some shared logic with real time

Session Analytics - Processing

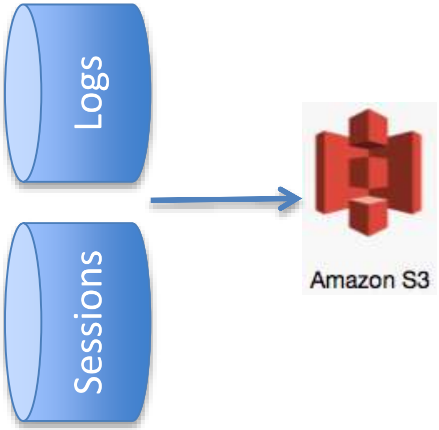


Session Analytics - Storage



Session Analytics – gen 1

- Storage



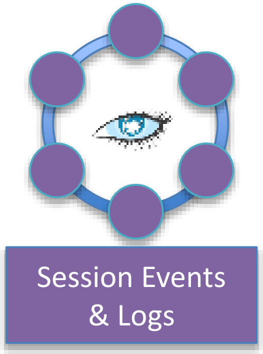
- Processing

Session Analytics – gen 1 pain points

- MapReduce good for batch
  - Not for near real time
- Complexity
  - Code in 2 systems / frameworks
  - Operational burden of 2 systems

Session Analytics – gen 2

- Storage



Session Events & Logs

- Processing

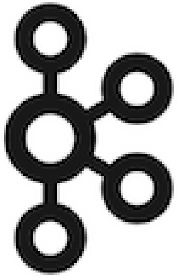
Java

Session Analytics – gen 2 learnings

- Reduced complexity
  - shared code and ops
- Batch still available
- New bottleneck
  - harder to extend logic

Session Analytics – gen 3 (*)

- Storage



- Processing

Storm

Mantis

Samza

Spark Streaming

Stream Processing Frameworks

Takeaways

- Polyglot Persistence
  - One size fits all *doesn't* fit all
- Strong opinions, loosely held
  - Design for long term, but be open to redesigns

Thanks!

@philip_pfo

NETFLIX