

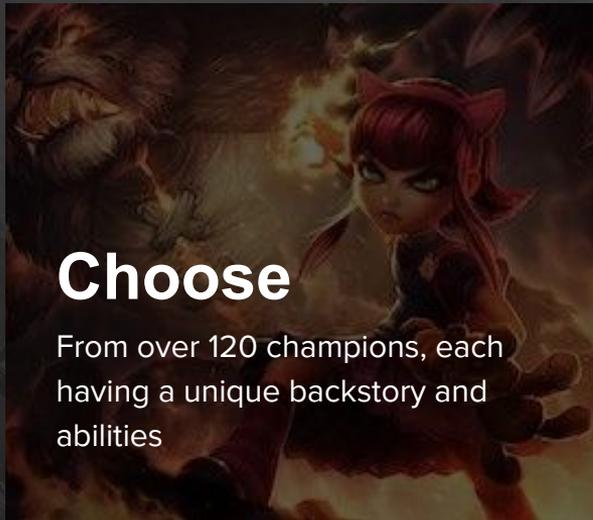
A person's silhouette is visible on the left side of the frame, looking towards the right. The background is a dimly lit room with a large, dark diamond-shaped graphic overlay in the center. The text "VIDEO GAMES AT SCALE" is written in white, bold, uppercase letters within the diamond. Below the text, there are two horizontal lines. At the bottom of the diamond, the text "Improving the gaming experience with Spark" is written, with "Spark" in red. The overall scene is dark and atmospheric, suggesting a gaming or esports environment.

VIDEO GAMES AT SCALE

Improving the
gaming experience
with **Spark**

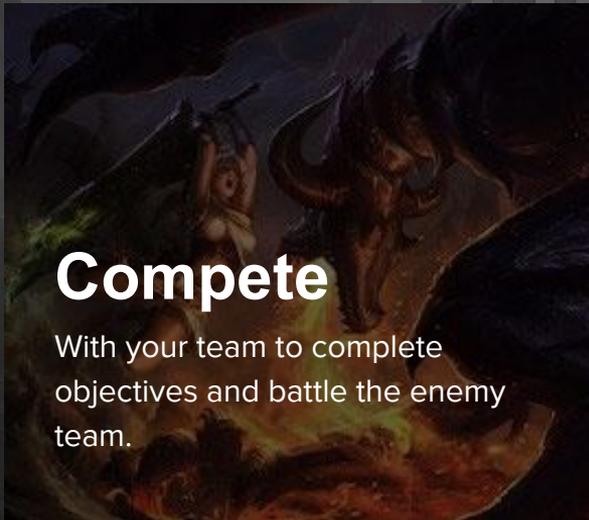
The background of the image is a detailed, stylized landscape. In the center, the 'League of Legends' logo is rendered in a golden, metallic font with a dark, shadowed interior. The logo is set against a backdrop of ancient, crumbling stone ruins with arches and pillars. To the left, a dense forest of dark green evergreen trees is visible, with a small figure standing near a stream. In the distance, jagged, snow-capped mountains rise against a sky filled with soft, golden light, suggesting a sunrise or sunset. The overall aesthetic is one of a vast, ancient, and somewhat mysterious world.

LEAGUE OF LEGENDS®



Choose

From over 120 champions, each having a unique backstory and abilities



Compete

With your team to complete objectives and battle the enemy team.



Win

Take down defenses and destroy the enemy nexus

OUR MISSION

WE ASPIRE

TO BE THE MOST

PLAYER

FOCUSED

GAME COMPANY IN THE

WORLD



AGENDA

DATA@RIOT

WHY SPARK?

USE CASES

FUTURE

Data

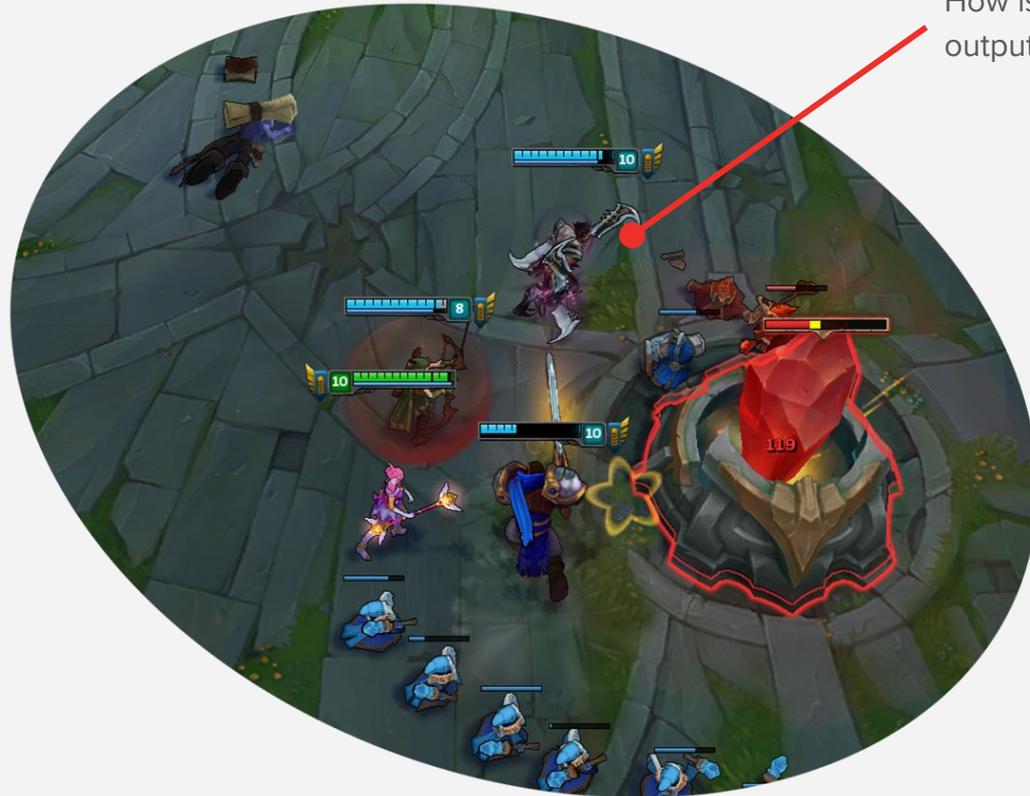
What can data tell us?

How do we interact with these data?



Game Balance

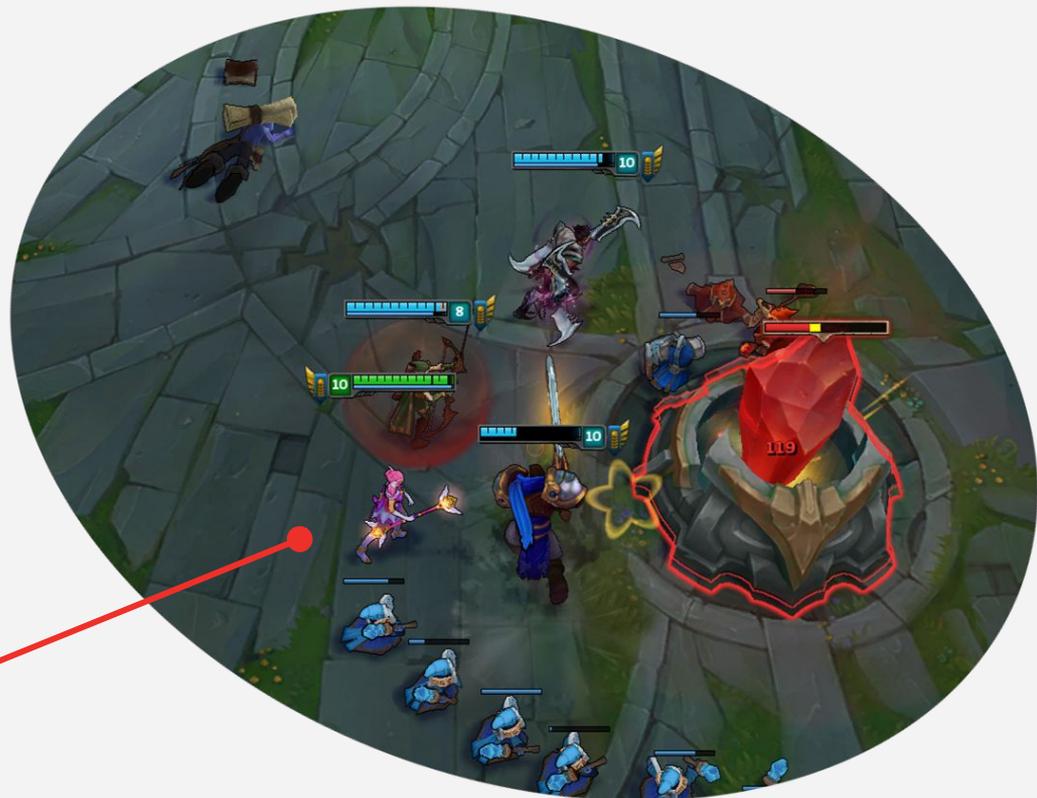
How is Draven's damage output since the patch?



Network Performance

Player noted lag during gameplay. Is his ISP having trouble?





Customization

Are there customizations similar to Star Guardian Lux this player may enjoy?



Players & Data

WORLDWIDE



67+ million

monthly active players



500+ billion

data points per day



26 petabytes

data collected since beta

Statistics released Jan 2014



THE DATA SCIENCE TOOLKIT



- Scheduled ETLs
- Ad Hoc Queries
- Data pulls



- Recent
- 10^6 events/sec
- Monitoring
- Fraud/Anomaly



- **Desktops**
- R/Python
- Tableau

THE DATA SCIENCE TOOLKIT



- Scheduled ETLs
- Ad Hoc Queries
- Data pulls



- Recent
- 10^6 events/sec
- Monitoring
- Fraud/Anomaly



- **Desktops**
- R/Python
- Tableau

Our data and ecosystem are scaling fine,
but our analytic tools were *not*.

AGENDA

DATA@RIOT

WHY SPARK?

USE CASES

FUTURE

Data Science At Scale

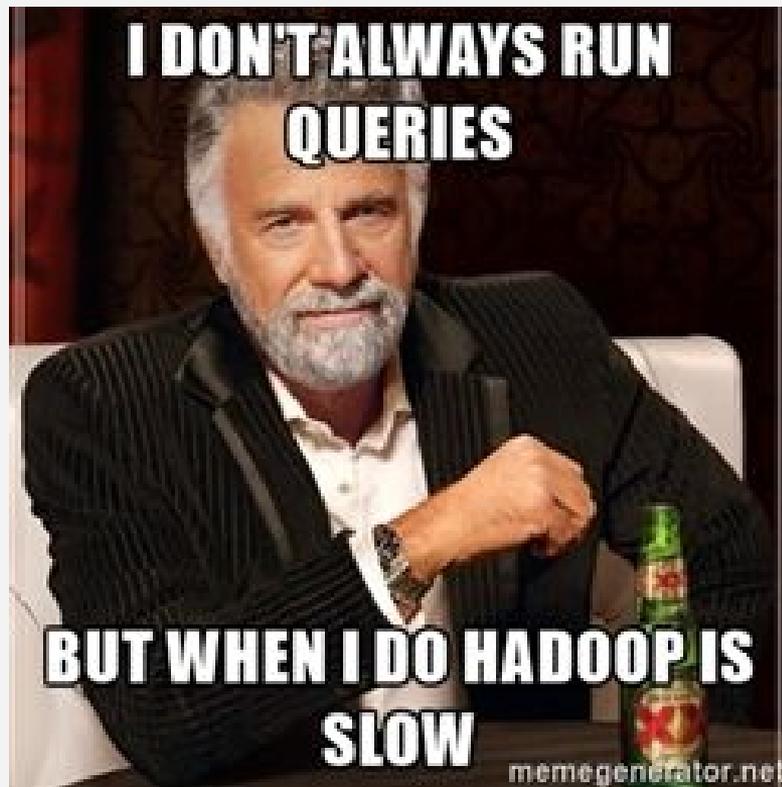
 databricks™



Why *Spark* ?



Spark SQL





Spark SQL

Spark Streaming

I SHOULD START



STREAMING



Spark SQL

Spark Streaming

Spark ML





Spark SQL

Spark Streaming

Spark ML

Empowerment



AGENDA

DATA@RIOT

WHY SPARK?

USE CASES

FUTURE

How did we apply this?

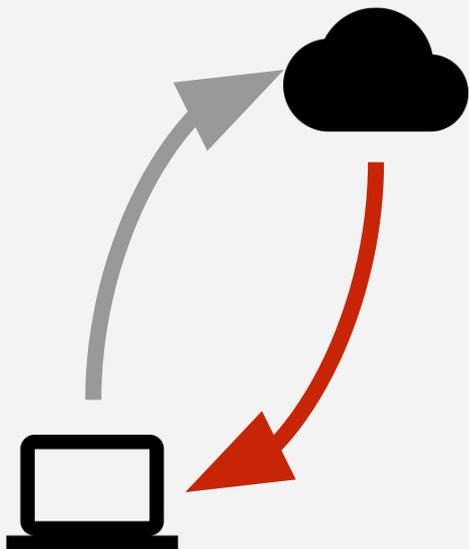


How to Play

(with data)



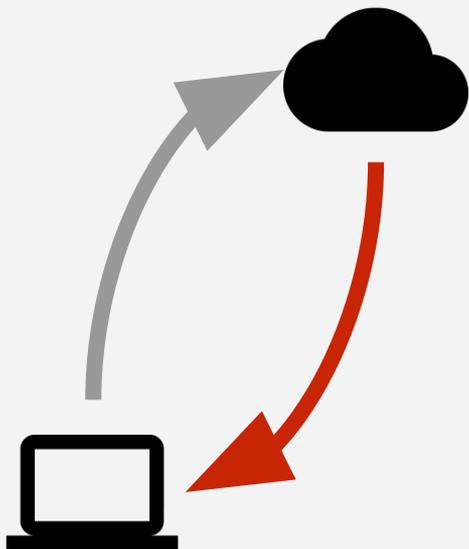
SparkSQL



Pain Points

FOR DATA EXPLORATION + REPORTING

Iterative query development



Pain Points

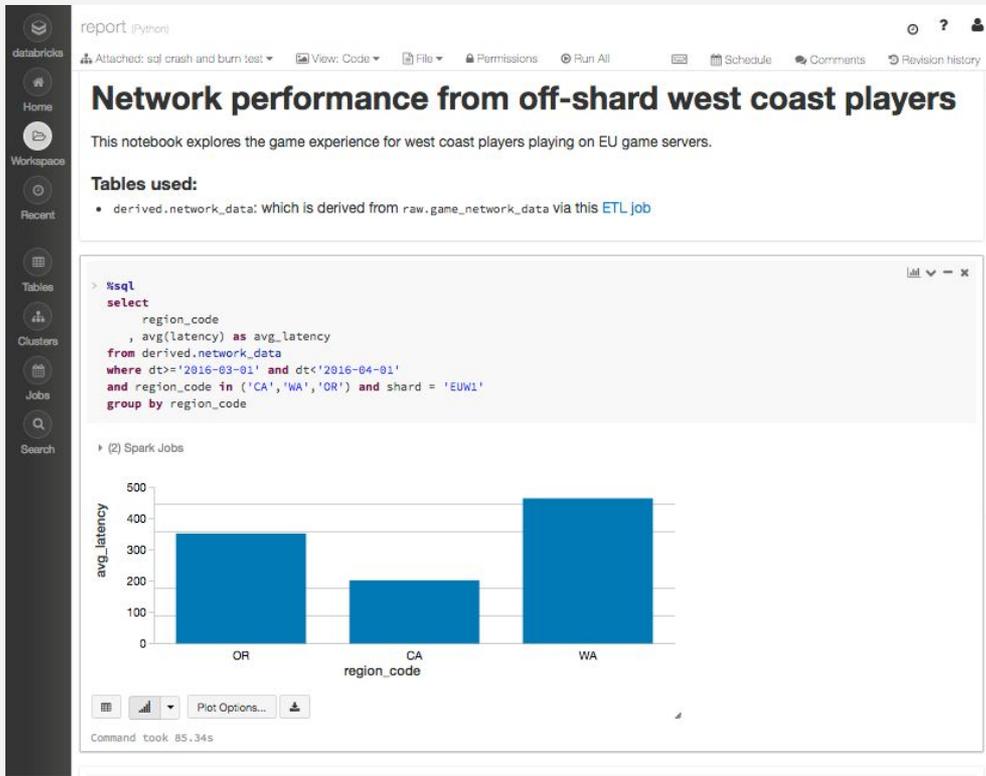
FOR DATA EXPLORATION + REPORTING

Iterative query development

Data pulls are slow

Workflow is inefficient.
Data is separated from tools.





Data, easier.



Integrated



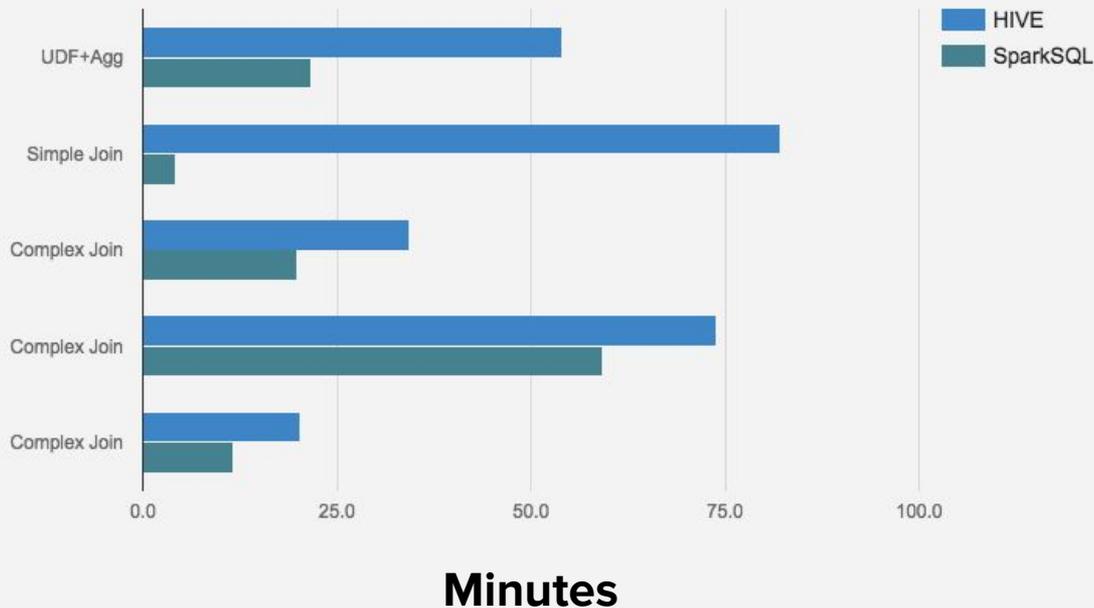
Interactive



Faster

Head-to-head tests between **EMR** and **SparkSQL** .

(equal cost clusters)





Managed/SparkSQL

TAKEAWAY

Data exploration efficiency increased

Performant SQL queries

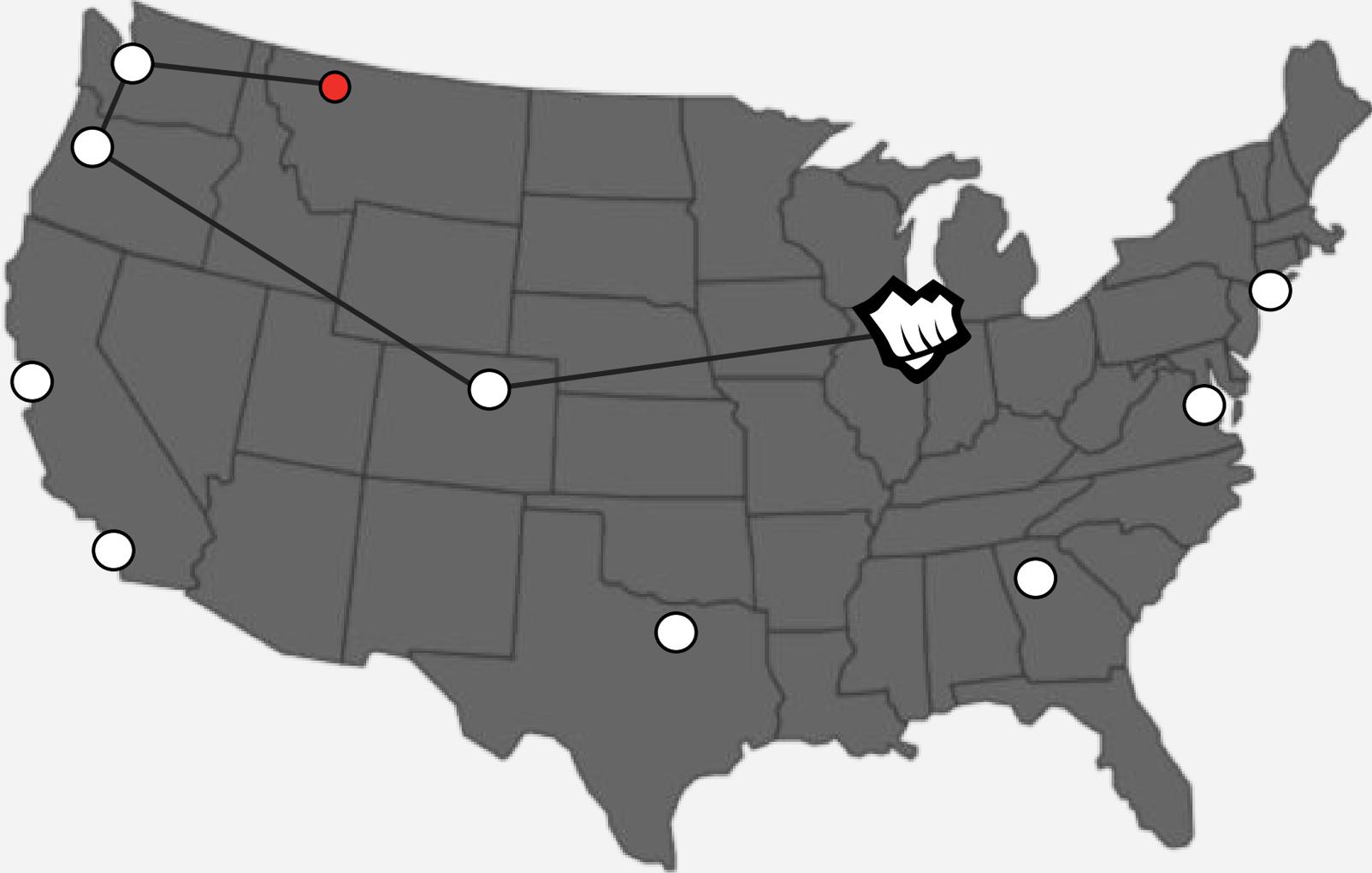
Self managed cluster deploys



Winning the war on lag

With Spark Streaming

**Gameplay highly dependent on
network connection.**

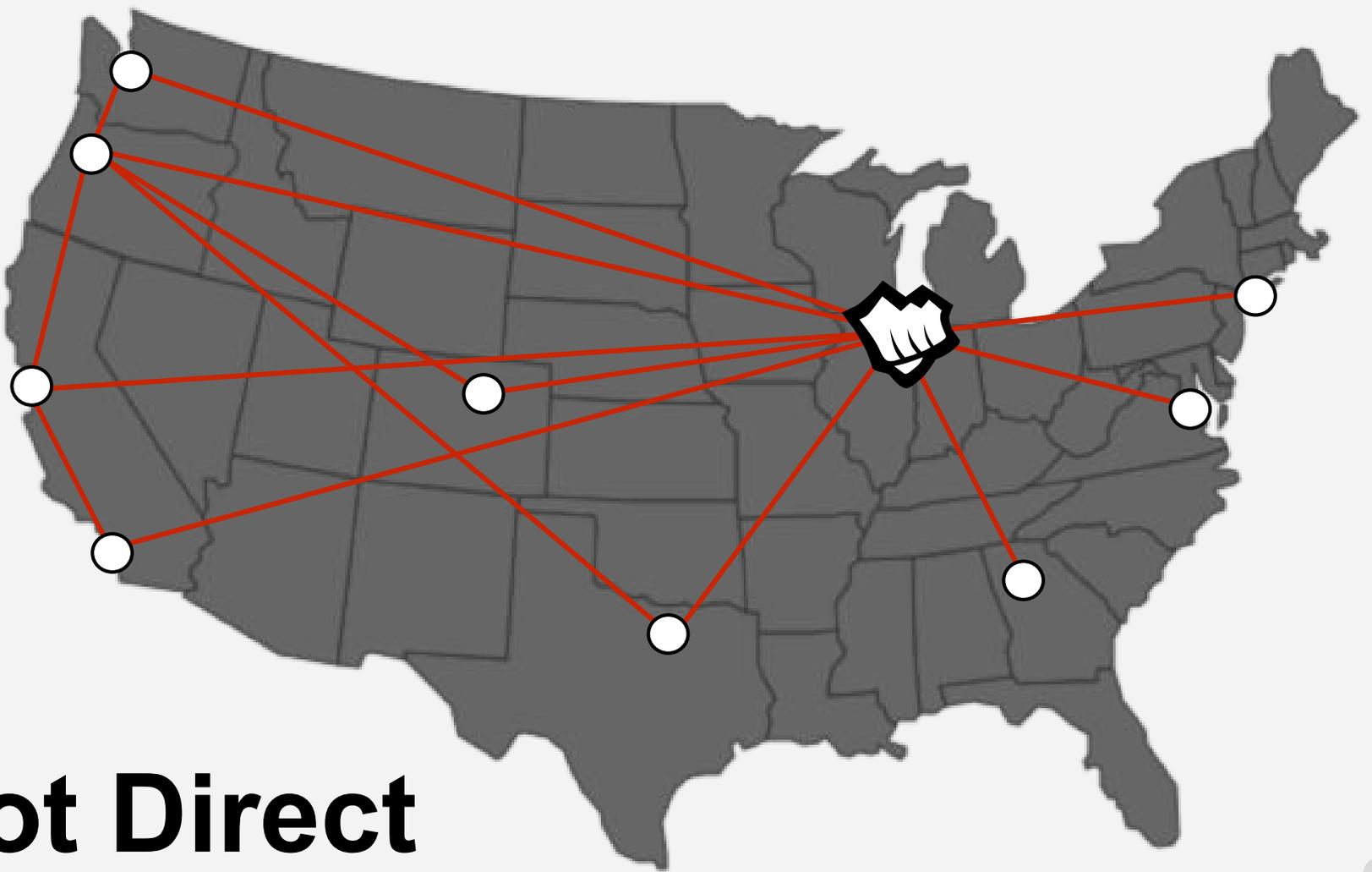


The existing network limits our players.

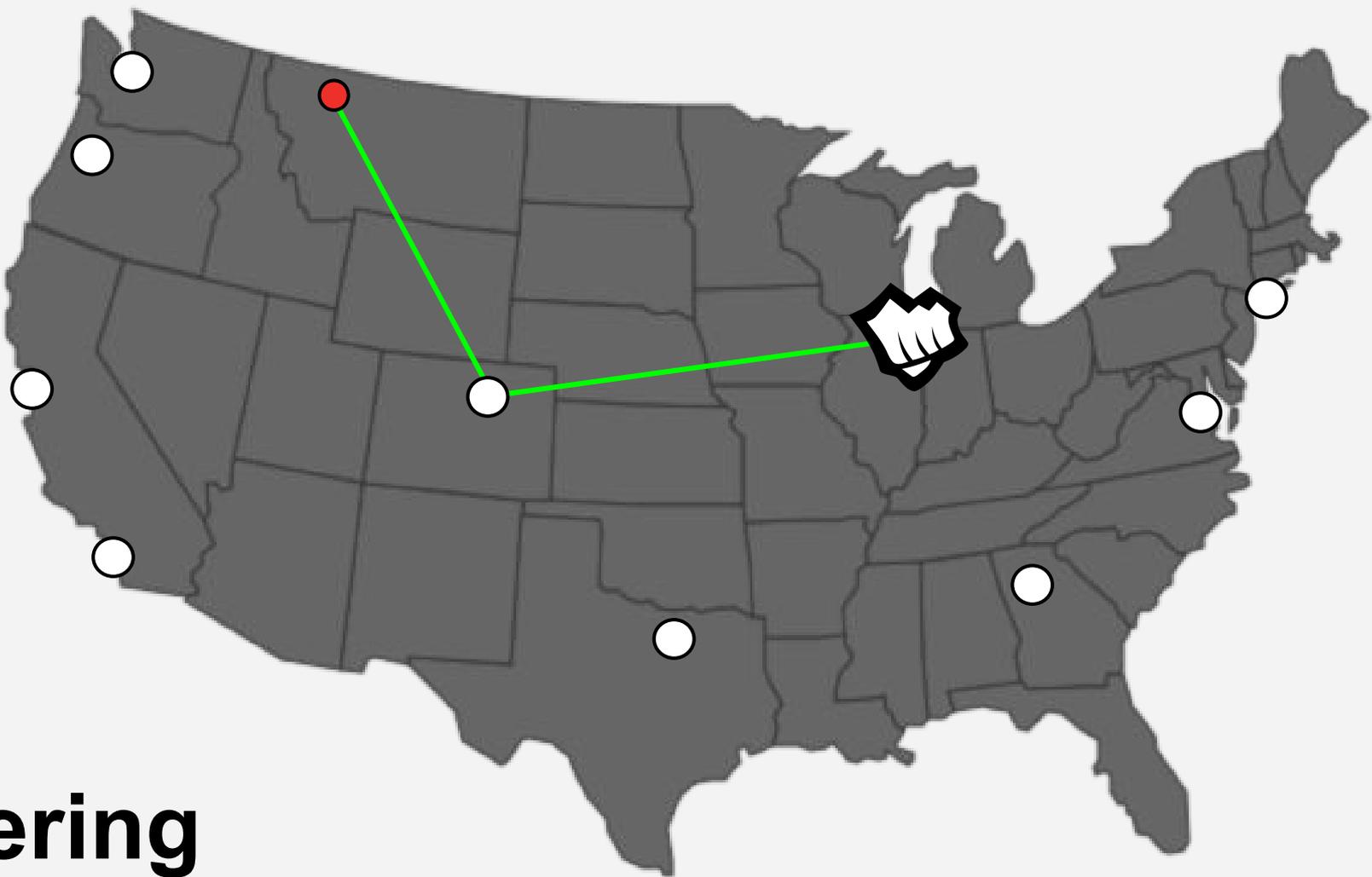


So we built our own!





Riot Direct



Peering

We have this awesome tool to fix problems.

But we need to find them!

Scope of Data

WORLDWIDE



17,000+

Unique ISPs



171,000+

City/ISP combinations

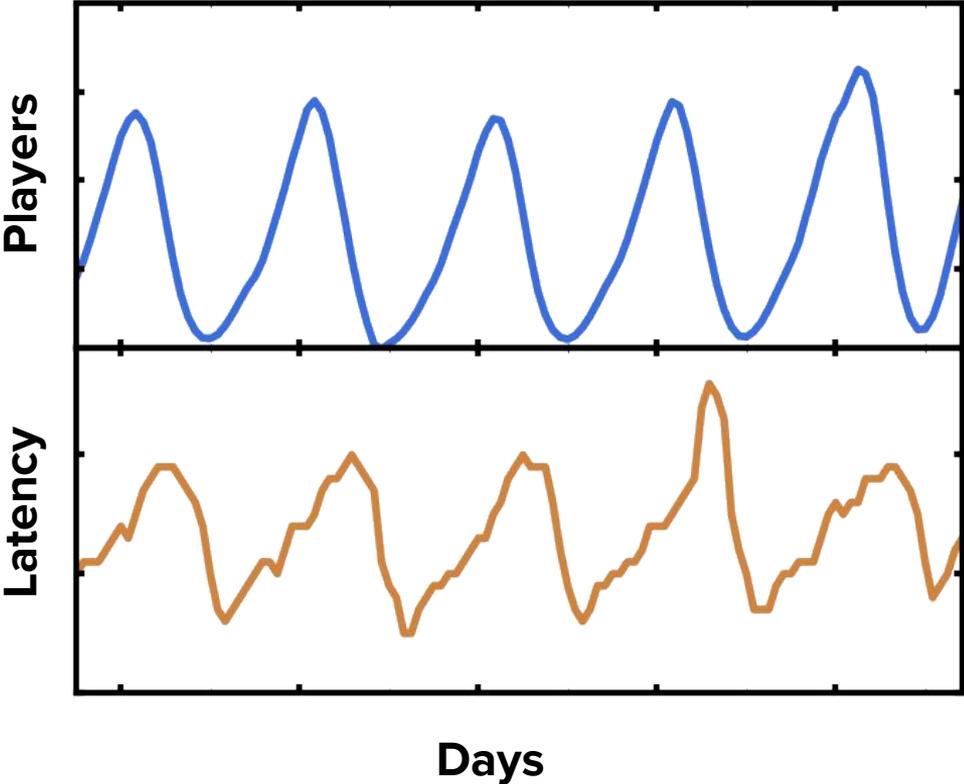


250,000+

Network stat messages per
second

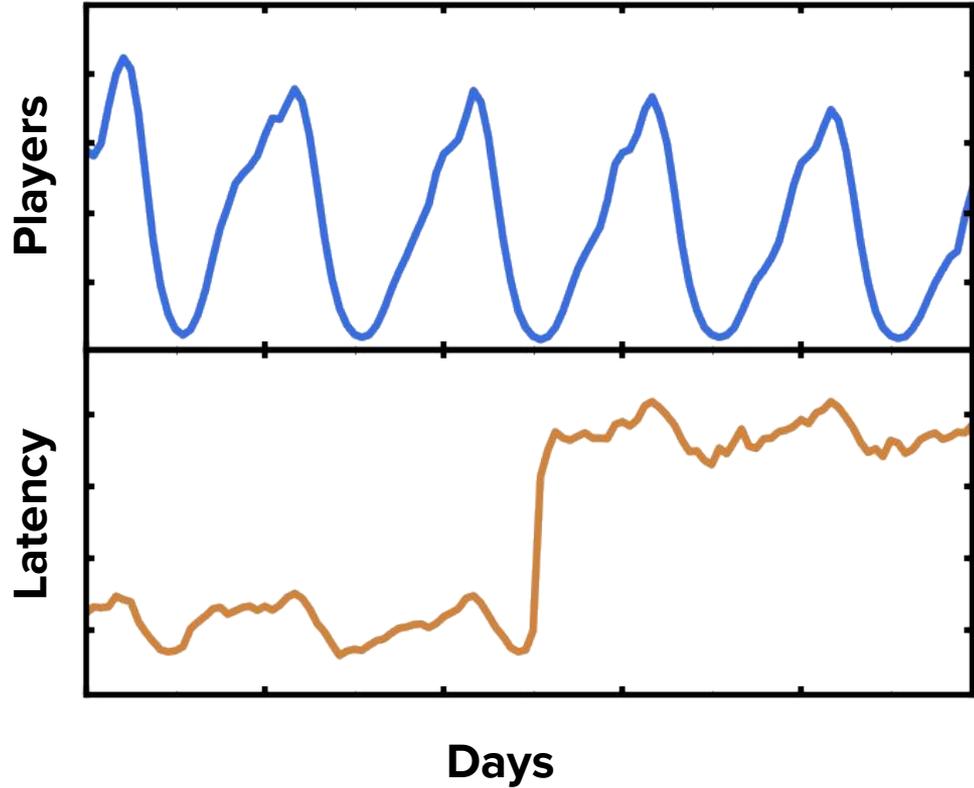


Model this



Model this

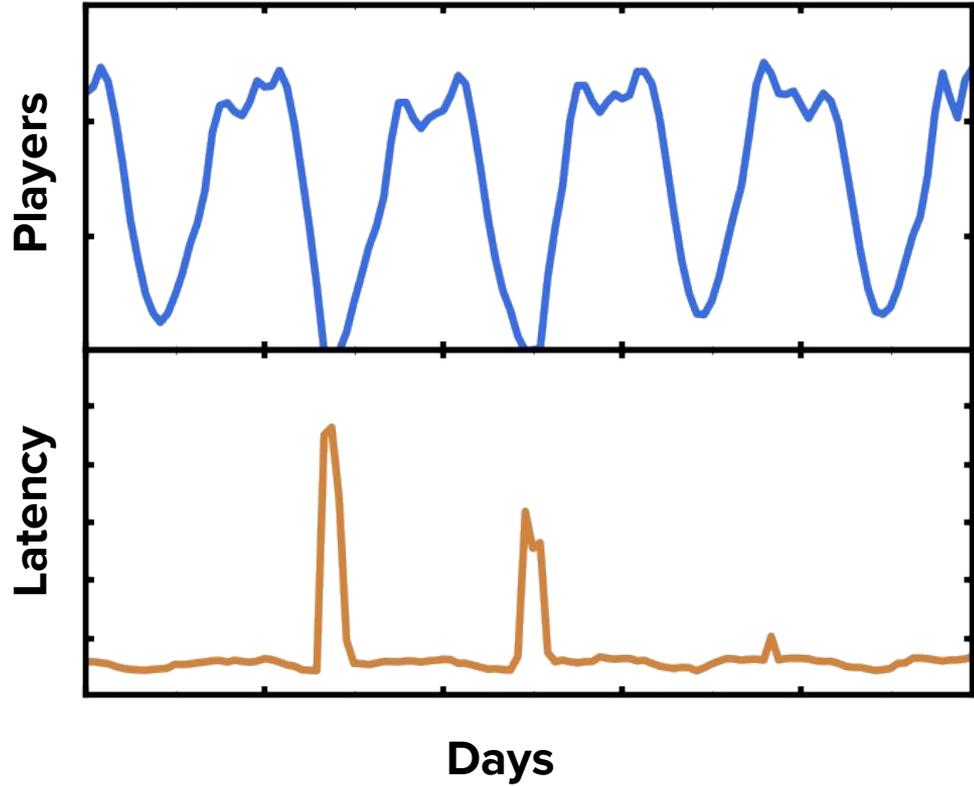
**...so we can
detect *this***

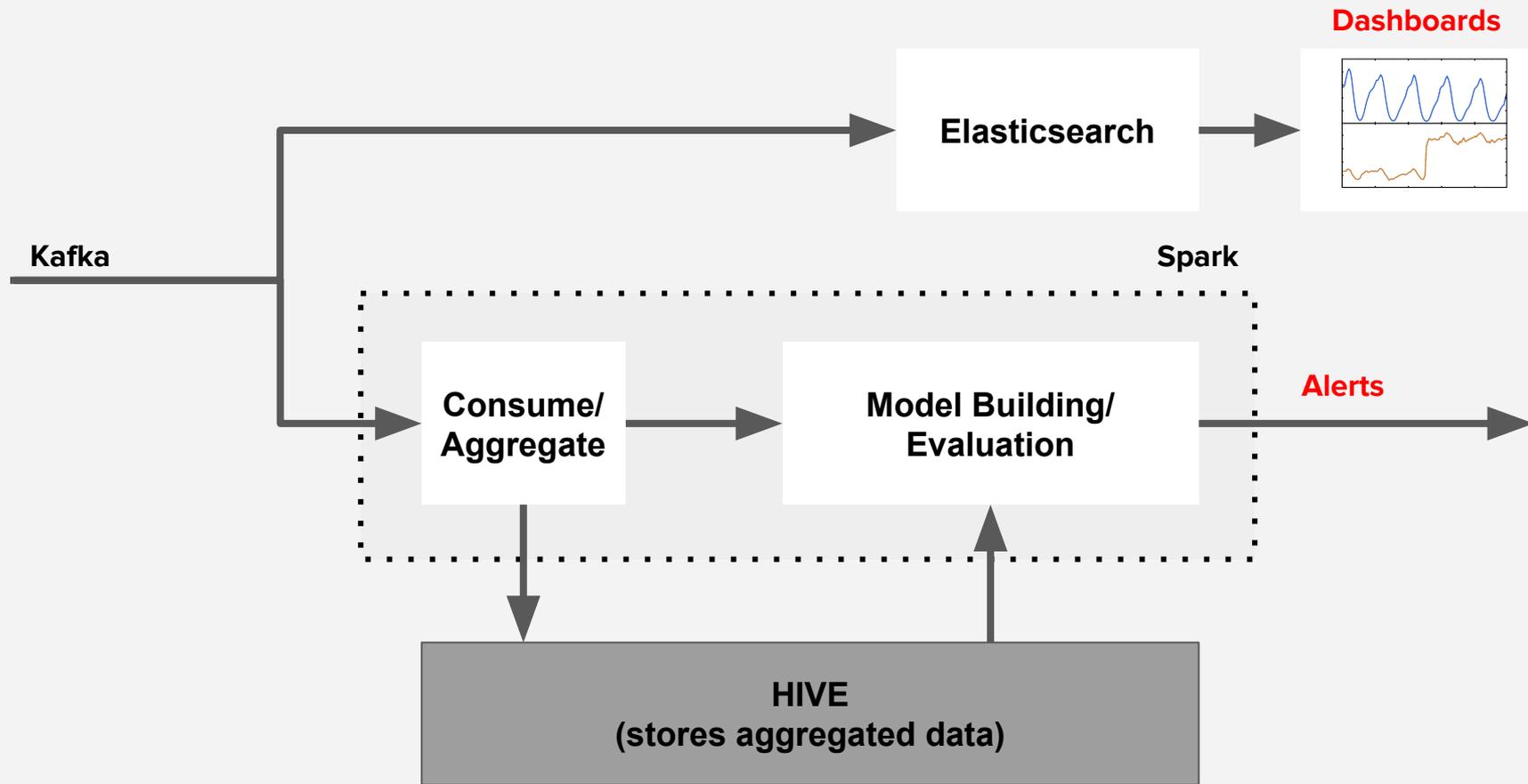


Model this

**...so we can
detect *this***

...or *this*.







Spark Streaming

TAKEAWAY

Works

Intuitive

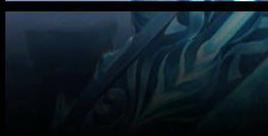
Will be a lot easier in 2.0!



Secret Store

The Recommendation System







LEAGUE of LEGENDS

CATEGORIES

Skins

Chroma Packs

Search

Show Owned

Sort By: Release Date ↓

OWNERSHIP

Champion Owned

TYPES

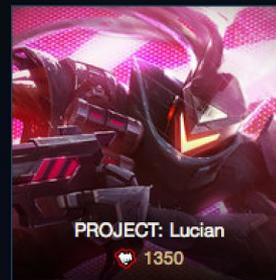
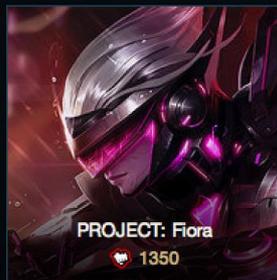
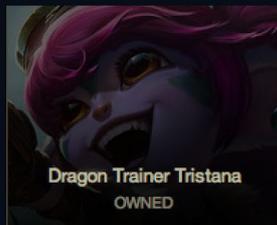
Limited Availability

Legendary

Ultimate

SALE STATUS

On Sale



FEATURED

CHAMPIONS

SKINS

GAMEPLAY

ACCESSORIES

BUNDLES

+ RP



CATEGORIES

◆ Skins

◆ Chroma Packs

Search

Show Owned

Sort By: Release Date ↓

OWNERSHIP

Champion Owned

TYPES

Limited Availability

Legendary

Ultimate

SALE STATUS

On Sale

Void Bringer Illaoi
1350

Dragon Trainer Tristana
OWNED

Spirit Fire Brand
1350

Demon Vi
1350

Cosmic Reaver Kassadin
1350

Shadowfire Kindred
1350

Ironsides Malphite
975

Marauder Alistar
750

Marauder Olaf
750

Warden Jax
750

Warden Karma
750

PROJECT: Fiora
1350

PROJECT: Leona
1350

PROJECT: Lucian
1350

PROJECT: Zed
1350



Why **personalized** recommendations?



Why **personalized** recommendations?



Why **personalized** recommendations?



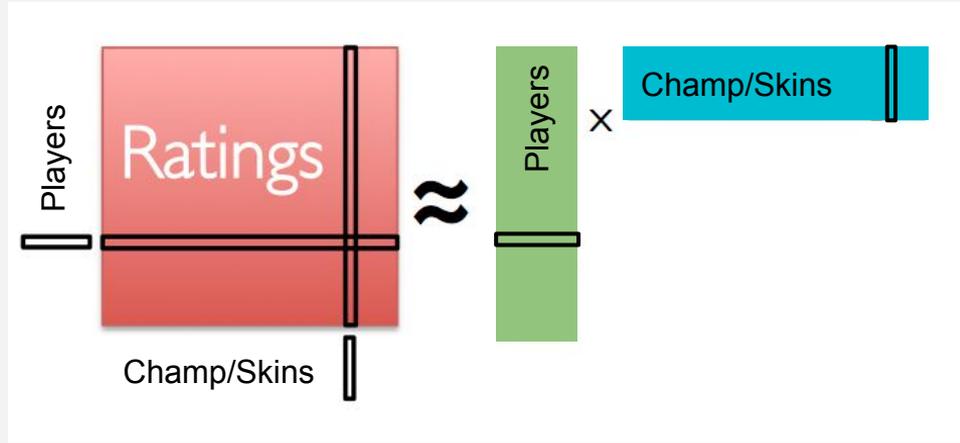




67+ million
monthly active players

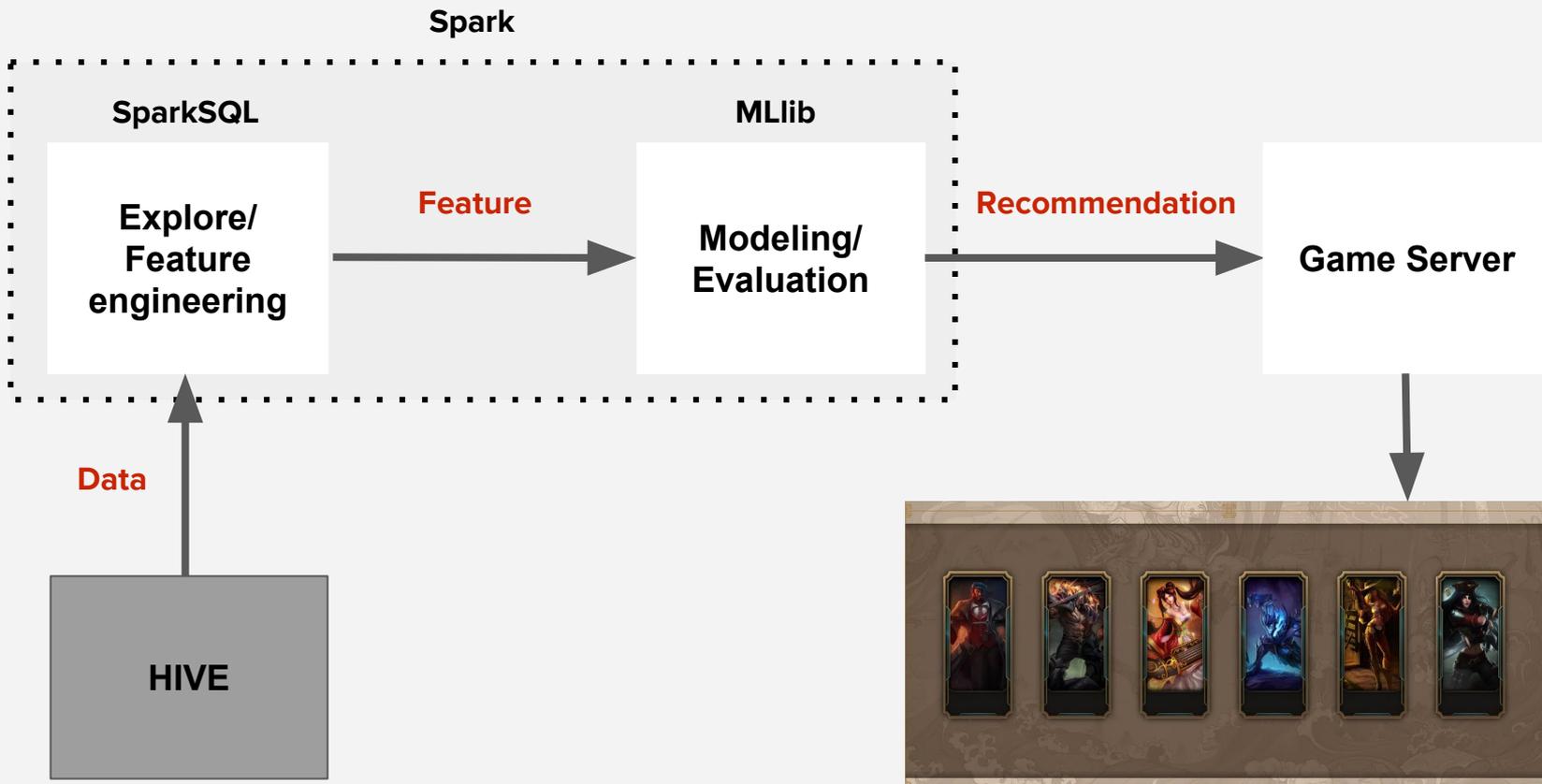


500+ billion
data points per day



~1000
champion/skin combo







Recommendation System

TAKEAWAY

Fast prototyping

Works on big matrix!

Easy automation

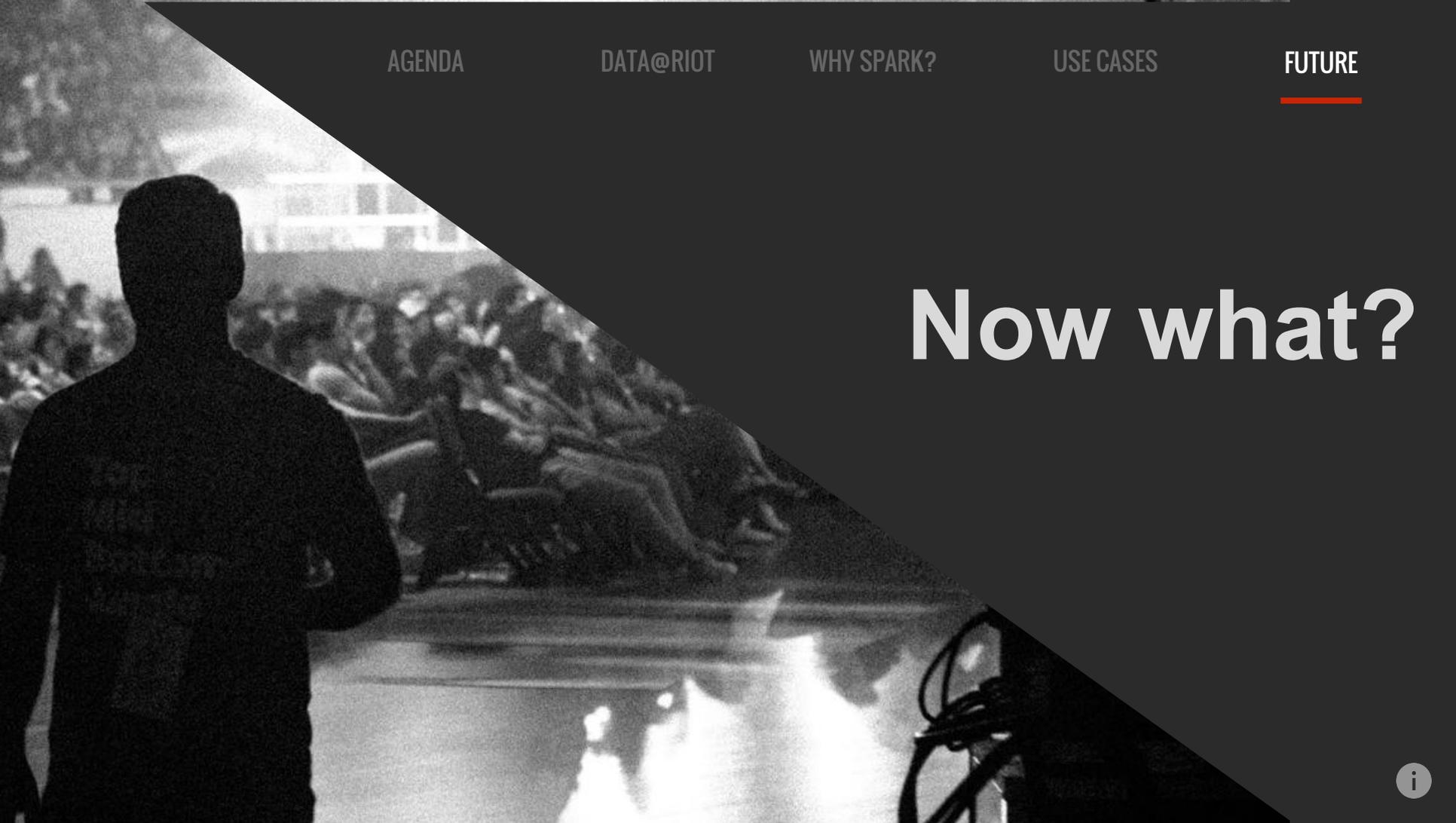
AGENDA

DATA@RIOT

WHY SPARK?

USE CASES

FUTURE



Now what?



**With Spark, we think that we're
just getting started.**

Wanna learn more?

(we're hiring!)



Colin Borys

CBORYS @ RIOTGAMES.COM



Xiaoyang Yang

XYANG @ RIOTGAMES.COM

