

AWS  
re:Invent

AIM222-R

# Monetizing text-to-speech AI

## **Robin Dautricourt**

Principal Product Manager  
Amazon Web Services AI/ML

## **Ron Jaworski**

Co-founder, CEO  
Trinity Audio

# Amazon Polly

- AWS machine learning (ML) service that converts text into lifelike speech
- Voices & languages:
  - 60+ voices across 29 languages
  - Neural and standard voices
  - Newscaster & conversational speaking styles
- Low-latency responses allow developers to build real-time systems
- Developers can store, replay, and distribute generated speech



Amazon Polly

# Standard TTS versus neural TTS

Text

Sentence to synthesize.

Phonetic transcription

'sɛntəns tə 'sɪnθə ,saɪz.

**Standard TTS**

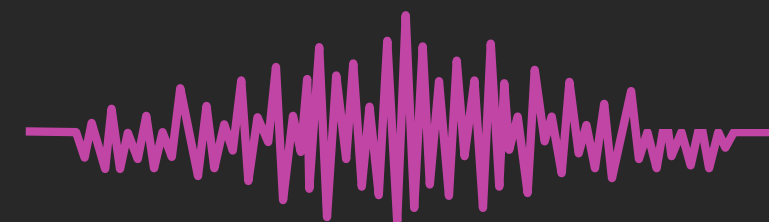
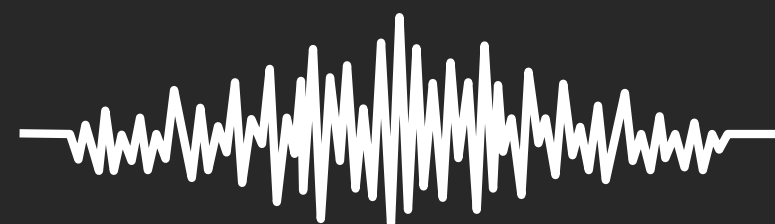
**Neural TTS**

'sɛnt

təns tə

'sɪnθ

ə ,saɪz.

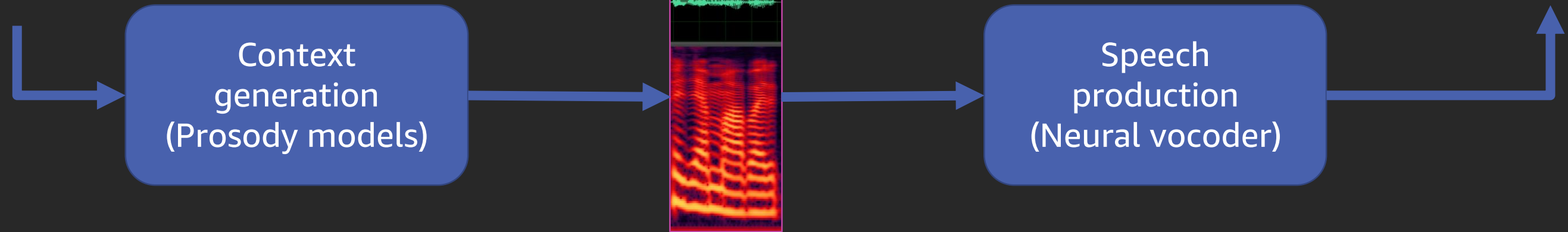


# How does neural TTS work?

## Phonemes & features

## Mel spectrograms

## Speech

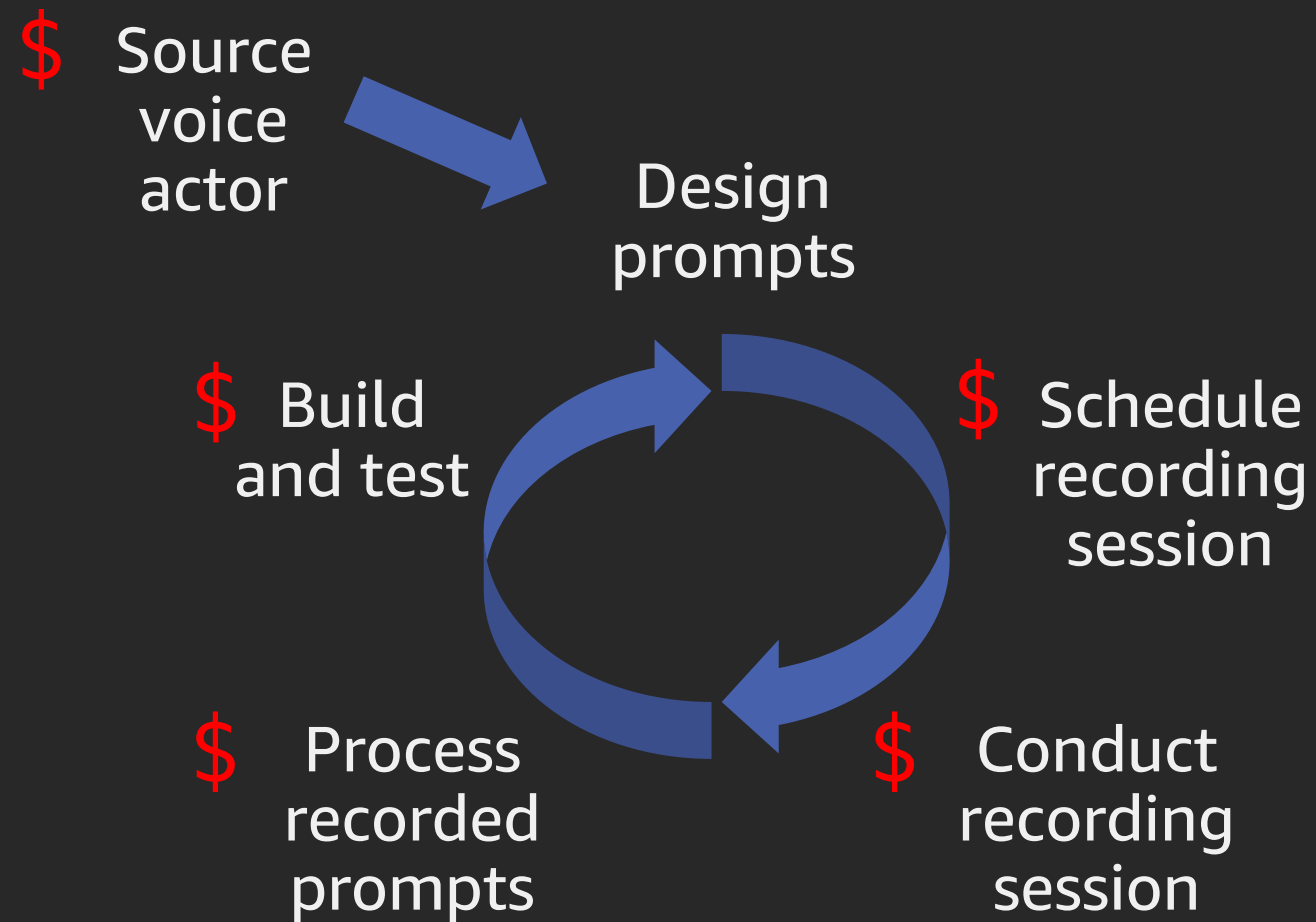


- Focused on intonation
- Speaker style-specific

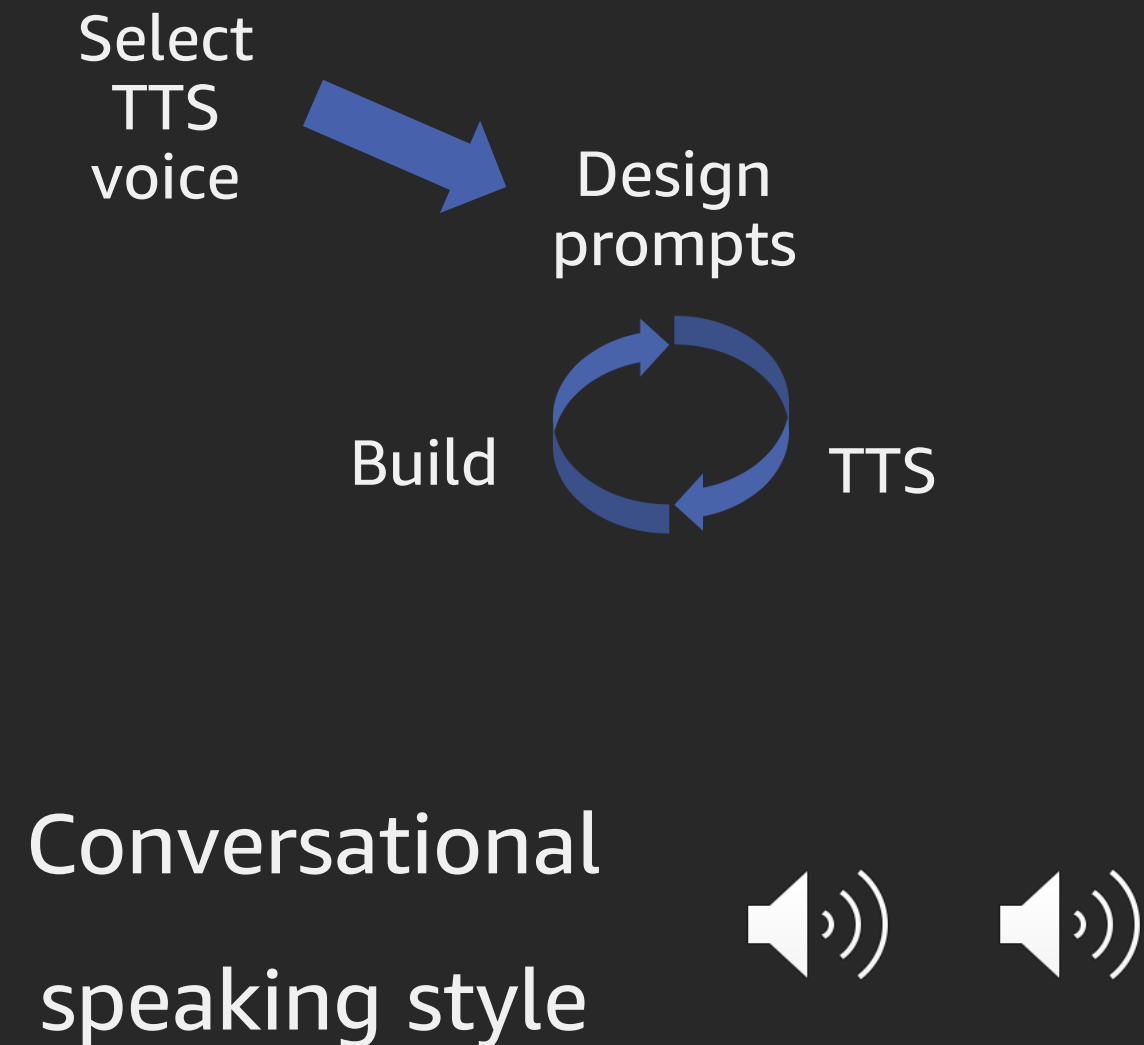
- Focused on Hi-Fi speech signal
- Independent of speaker or style

# Cost & time savings with text-to-speech

## Speech recordings



## Amazon Polly



# New monetization opportunities via speech



New modality  
(text and speech)

New channel  
(podcast, video)



Newscaster  
speaking style



# A few words about Trinity Audio and how we started

Trinity Audio is a company by Somoto (a publicly traded company in the TLV stock exchange). We are Adtech veterans with a deep understanding of the publisher landscape. We were all a part of Trinity's sister company TIM Media—a video monetization solution for publishers, which was acquired by Somoto in 2016.

In 2017, we set forth to develop a set of products that brings value to all 3 pillars of the ecosystem: publishers, users, and advertisers, with our unique content and technology solution.

The result: Trinity Audio.



General Demo x +

demo.trinityaudio.ai/general-demo/#/

**Customize your demo**

Select units to display

- Trinity Player
- Trinity Pulse

Select additional features

- FAB
- Multiple articles

Select article language

ENGLISH DEUTSCH FRANÇAIS

ESPAÑOL 汉语 ITALIANO

Select voice over accent (Eng only)

ENGLISH(US) ENGLISH(UK)

Select ad placements

- Preroll
- iPreroll
- Time-based midroll
- Postroll

Select additional ads

- Companion ad

Show more

APPLY

MOBILE DESKTOP

Artificial Intelligence Is Too Important to Leave to Google and Facebook Alone

Let's develop a public research consortium to take on useful projects that have no commercial prospects.

First Community

CLOSE MENU

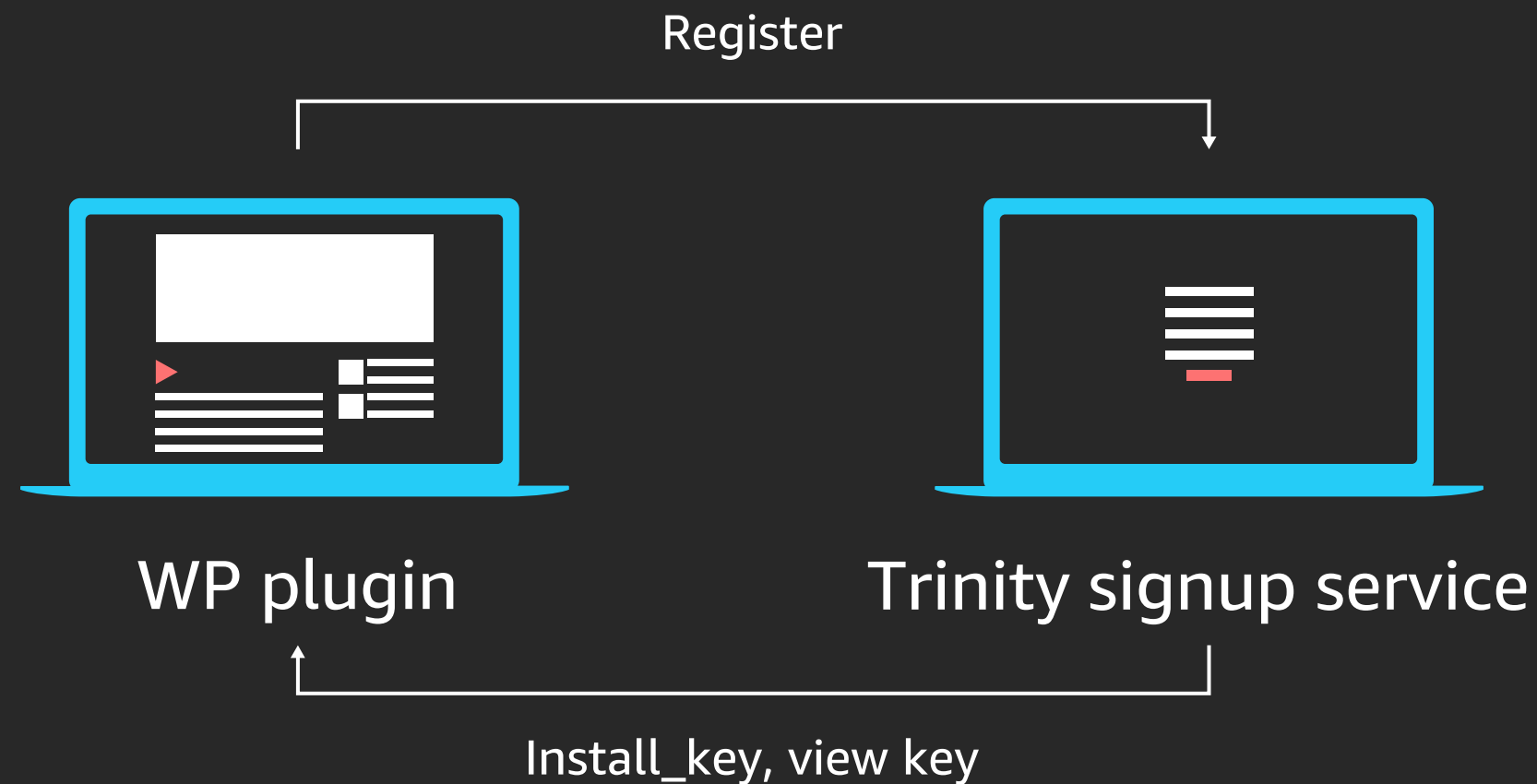
# WordPress plugin credit

- Credit system for reduction of risk
- Ads generated = more credits
- Credit execution = no more audio creation

# WordPress plugin architecture

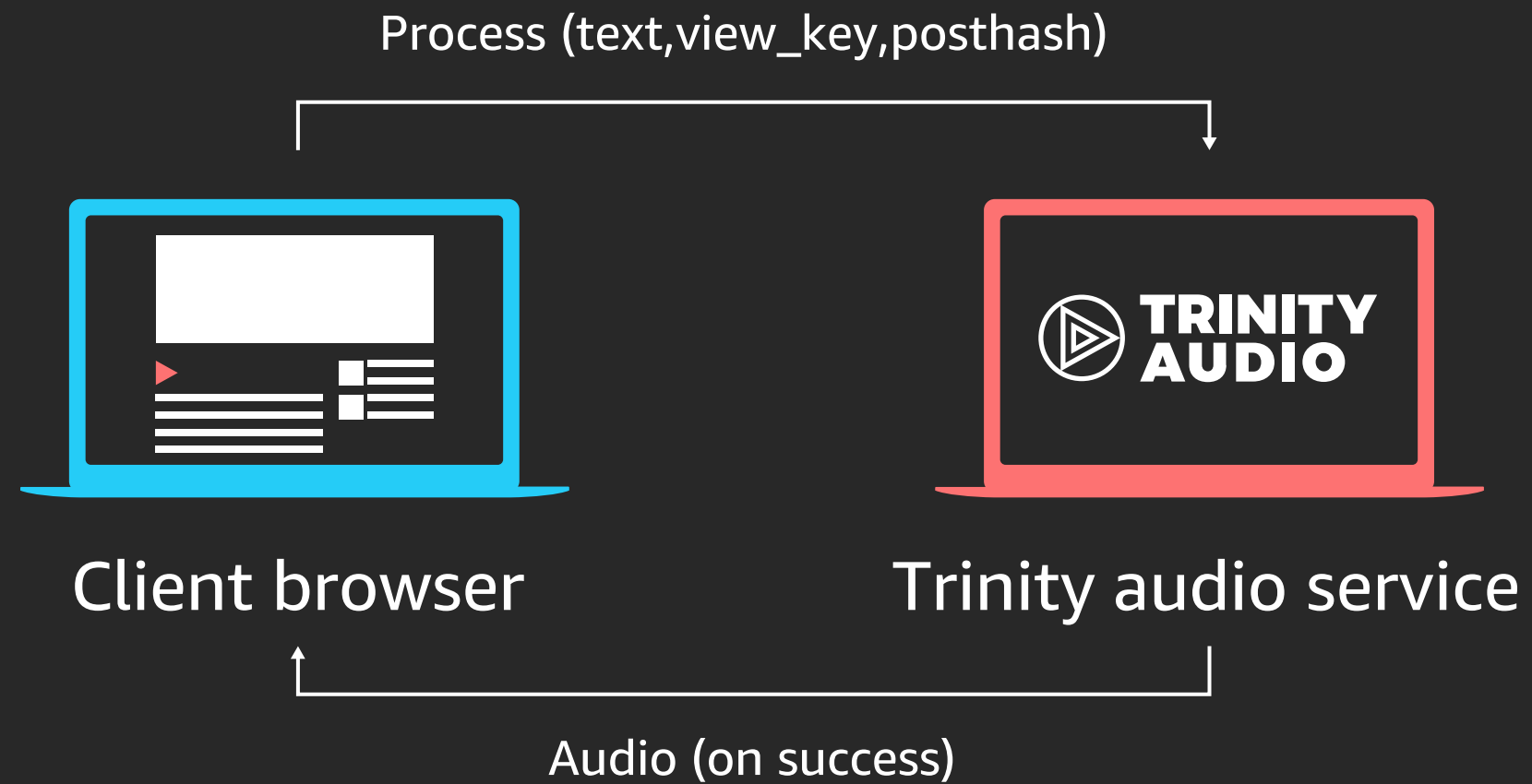
Goal:

To have an identifier to a plugin installation—which we can use for tracking usage and preventing abuse

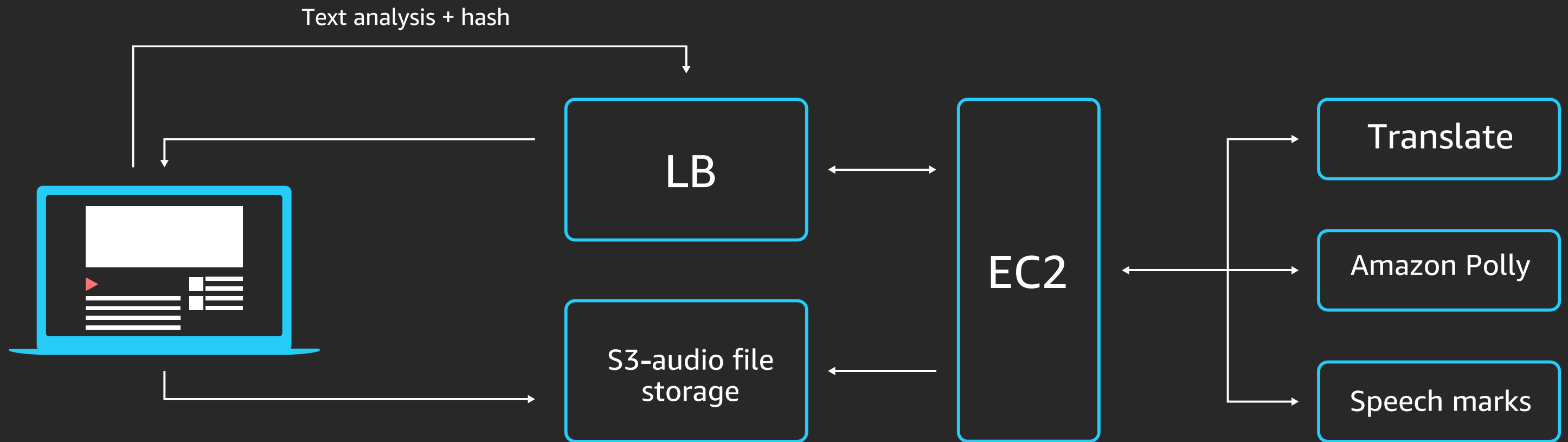


# WordPress plugin architecture

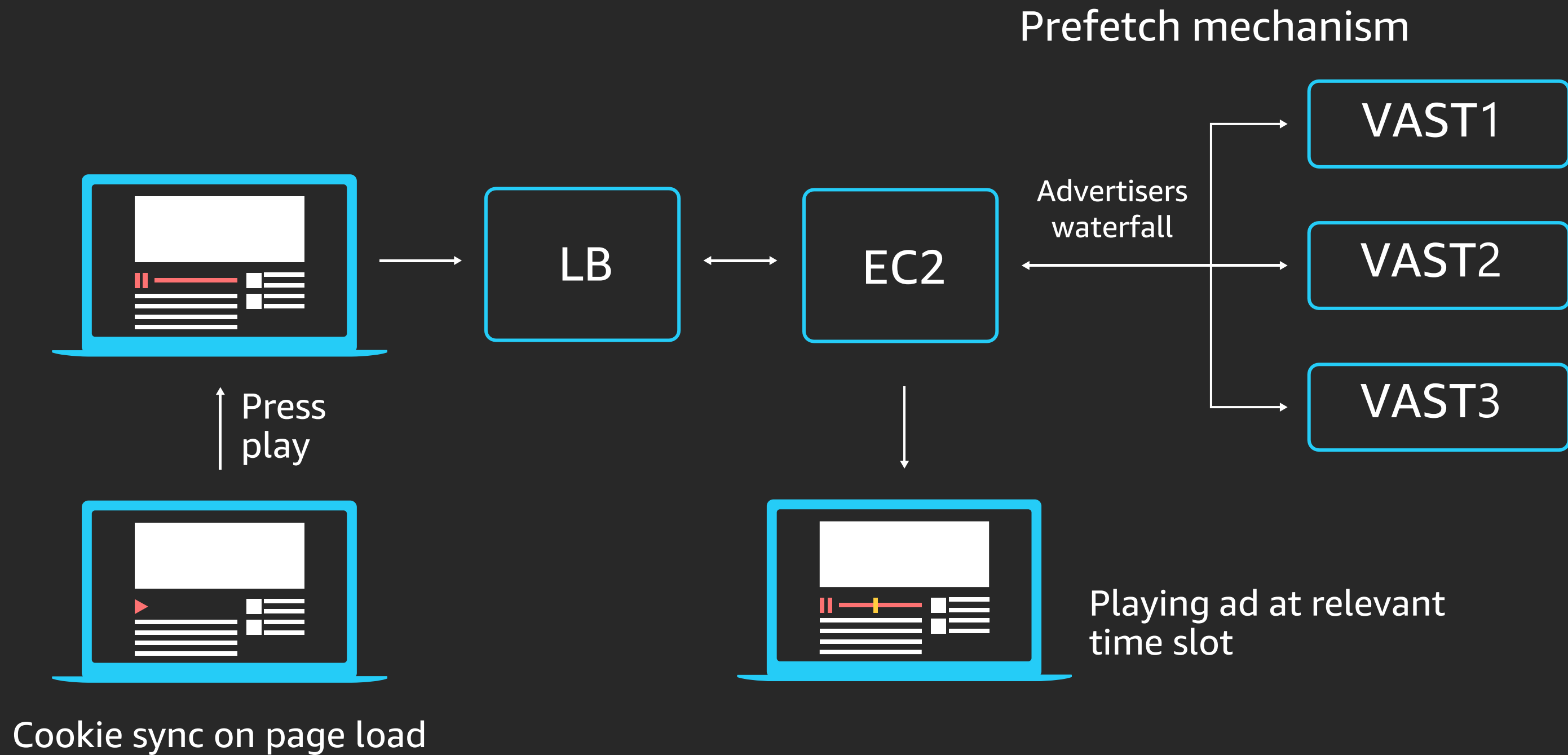
Player "play" clicked



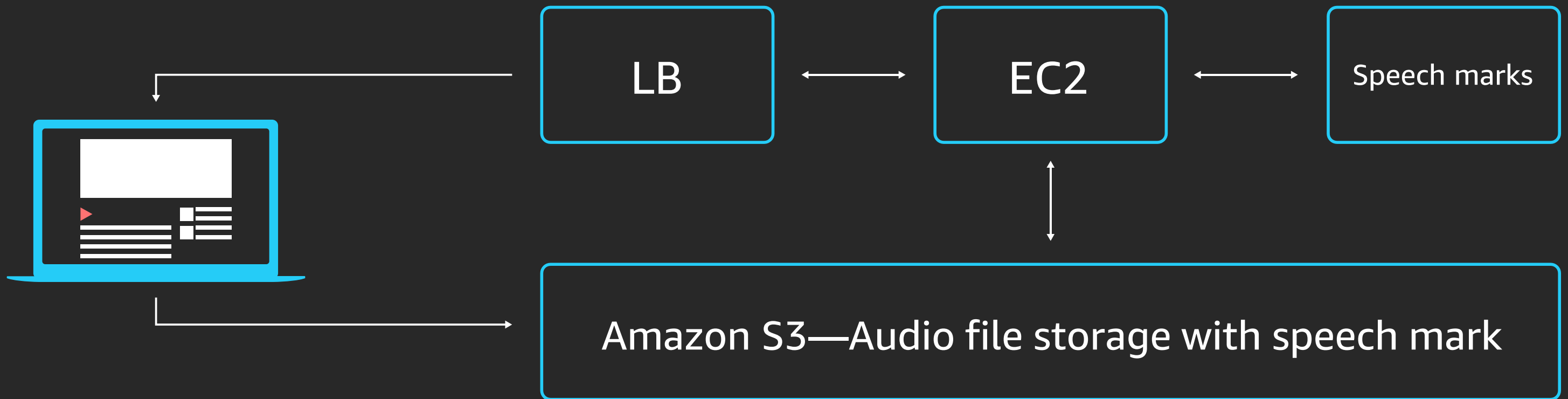
# Architecture of work with Amazon Polly



# Architecture of our ad server



# Using speech marks to identify the perfect spot to place an ad



Speech marks can be: end line, end paragraph, etc.

# Some stats

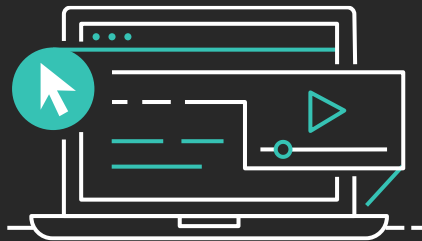
- Over 100 publishers
- Over 90% LTR (main advertiser KPI)
- Over 3M daily player loads
- Average CPM between \$3 and \$5
- Fill rate over 70%
- CTR ~3%

# Learn ML with AWS Training and Certification

The same training that our own developers use, now available on demand



Role-based ML learning paths for developers, data scientists, data platform engineers, and business decision makers



70+ free digital ML courses from AWS experts let you learn from real-world challenges tackled at AWS



Validate expertise with the  
**AWS Certified Machine Learning - Specialty** exam

Visit <https://aws.training/machinelearning>

# Thank you!

**Robin Dautricourt**

rdautric@amazon.com

**Ron Jaworski**

ron@trinityaudio.ai

# Questions?