

# Cisco TechClub

Lokální edukační on-line webináře každých 14 dní

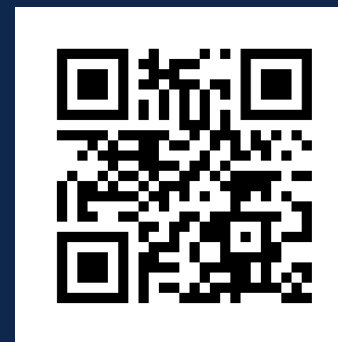
Cisco TechClub  
hlavní portál



Cisco TechClub  
registrační stránka



Cisco TechClub  
Webex Space



The bridge to possible



# AI in Collaboration

Jaroslav Martan  
jmartan@cisco.com  
21.1.2025

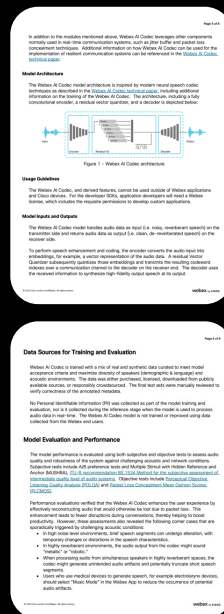
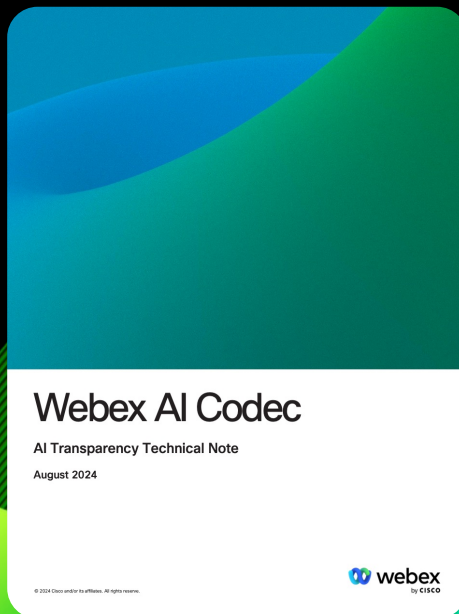
# Agenda

- Introduction
- Real-time Media Models
- AI Agent
- AI Assistants

# Introduction



# Preserving your trust with AI governance



Transparency

Fairness

Accountability

Reliability

Security

Privacy

Cisco  
**Responsible AI**  
Principles

Example of 22 Webex AI  
Transparency Technical  
Notes available from [Cisco](#)  
[Trust Portal](#)

# Incredible AI momentum

## Webex Suite

AI Codec  
Meeting Summaries  
Be-right-back Detection  
Call Summaries<sup>1</sup>  
Format, tone, translation  
Vidcast Highlights  
Ask AI Assistant – recordings  
Custom Dictionary

## Cisco Devices

Meeting Zones  
Frames & People Focus  
Cross-view  
Cinematic Meetings  
Virtual AI Lens  
Extended Speaker Tracking

## Webex Contact Center

Agent Wellness  
Suggested Responses<sup>1</sup>  
Auto CSAT  
Dropped Call Summaries  
Topic Analytics, Volume Drivers  
Native Voice Virtual Agent<sup>1</sup>

## Webex Connect

Code Generation with Evaluate Node  
Chat Summarization  
Multilingual Chatbots  
'Cold Start' Bot training data generation  
AI Outbound Video<sup>1</sup>

## Webex Control Hub

AI Assistant for Control Hub  
AI Assistant for Developers<sup>1</sup>  
AI Workspace Ranking<sup>1</sup>

<sup>1</sup> in beta

# Real-time Media Models

# Business Cases – Distance Zero Meetings



Eyes

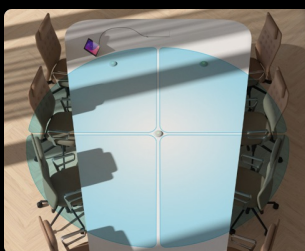


Knowing where people are in the room

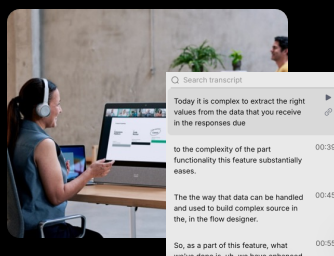


Knowing who is in the room

Ears



Knowing where the active speaker is sitting



Knowing who is speaking  
(Recording transcript)

Brain



Give the best view of the room



Framing is a blend of art and science



Focus only on those in the meeting

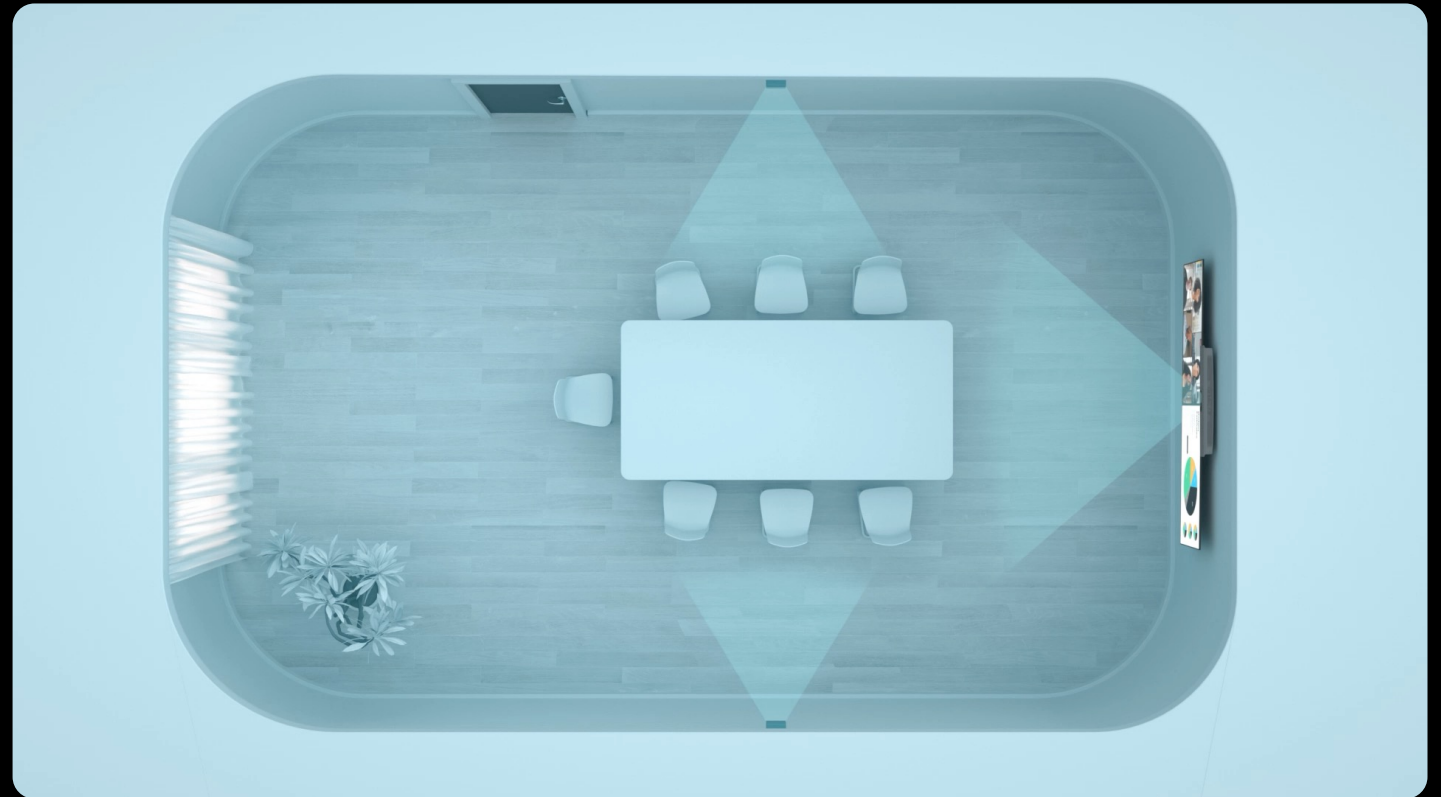
# Meeting Zones

Distraction-free meetings in open spaces framing and tracking only people in the meeting zone



# Cross View

Follow where the speaker is looking





# Be right back

Blur your camera and  
mute microphone  
automatically when you  
step away



# Revolutionizing video with **super resolution**

Generative AI video will rebuild packets in poor bandwidth conditions



Low resolution video captured at 270p



Super Resolution rescales 270p back into 1080p



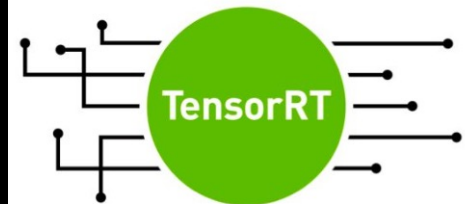
# Methodology

# Working with Deep Learning





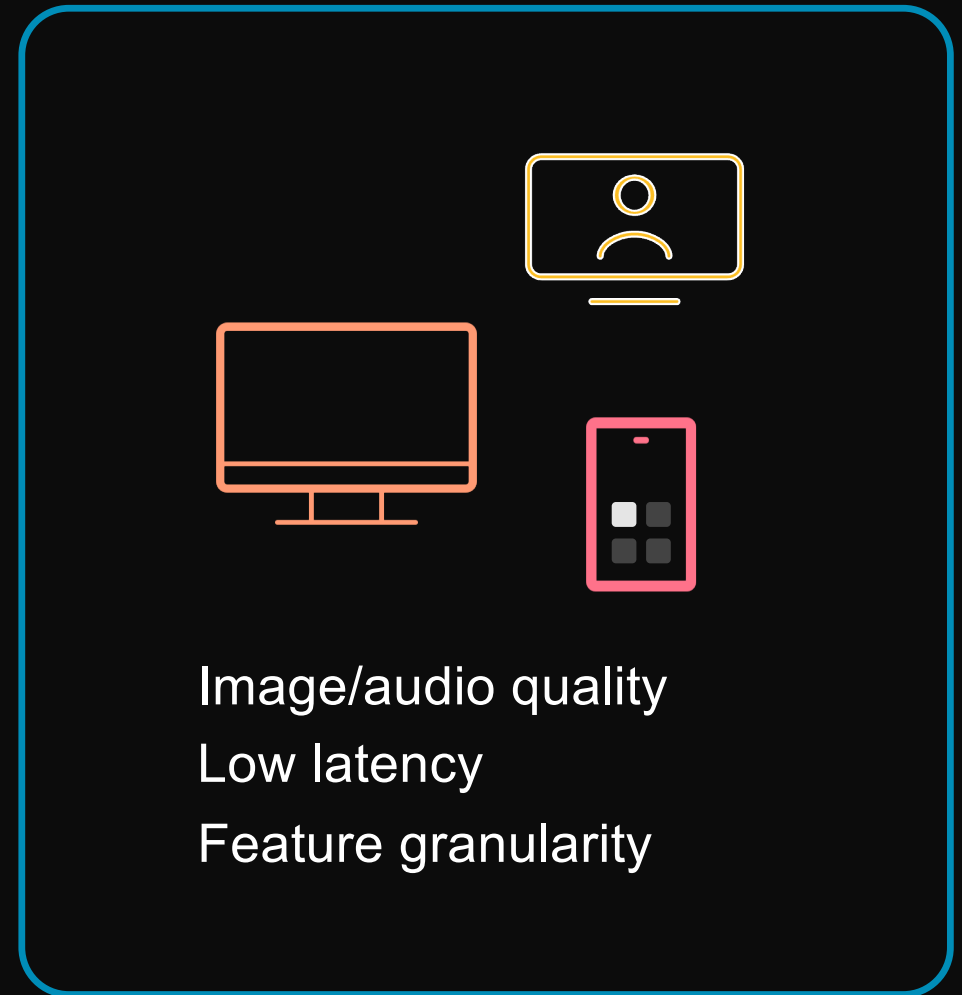
ONNX



# Cloud or Edge?



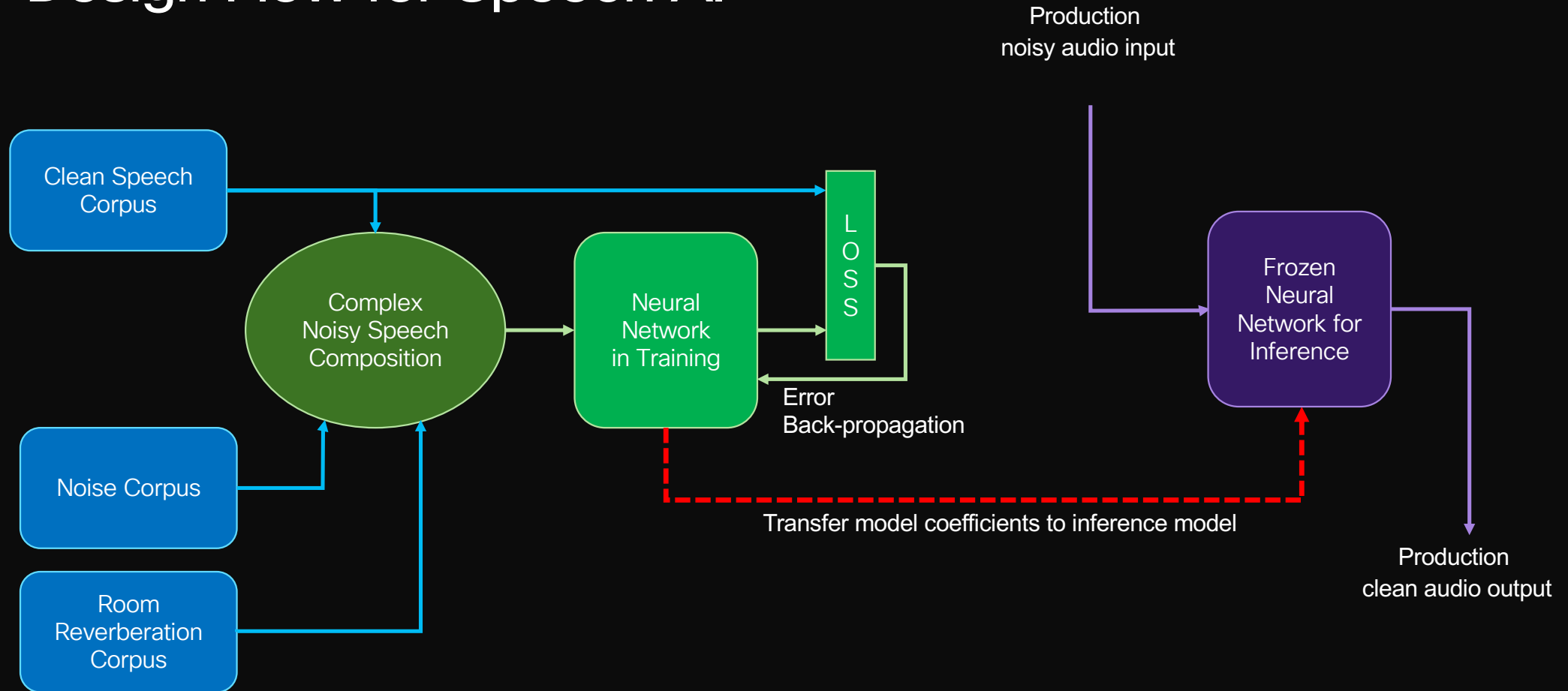
Performance(?)  
Scalability (LLMs)



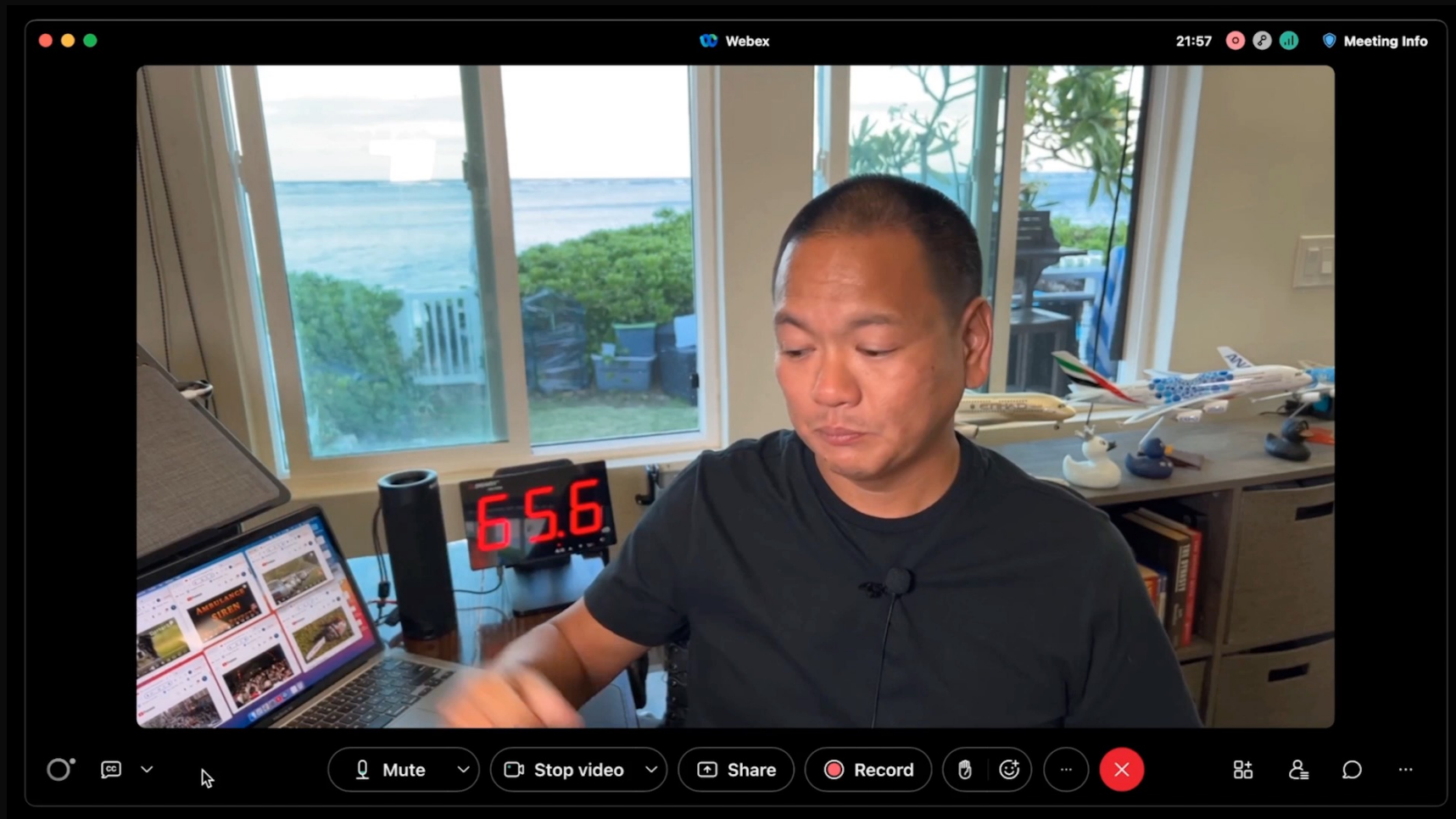
Image/audio quality  
Low latency  
Feature granularity

# Processing Audio

# Design Flow for Speech AI

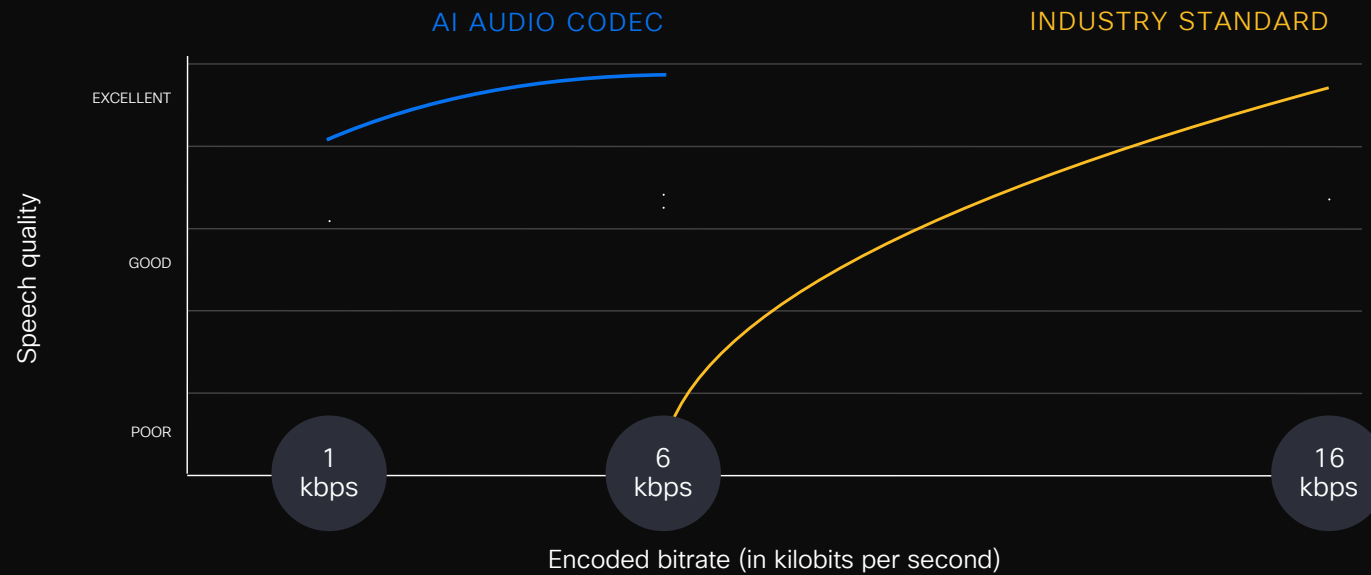


# Processed at Endpoint or in Cloud

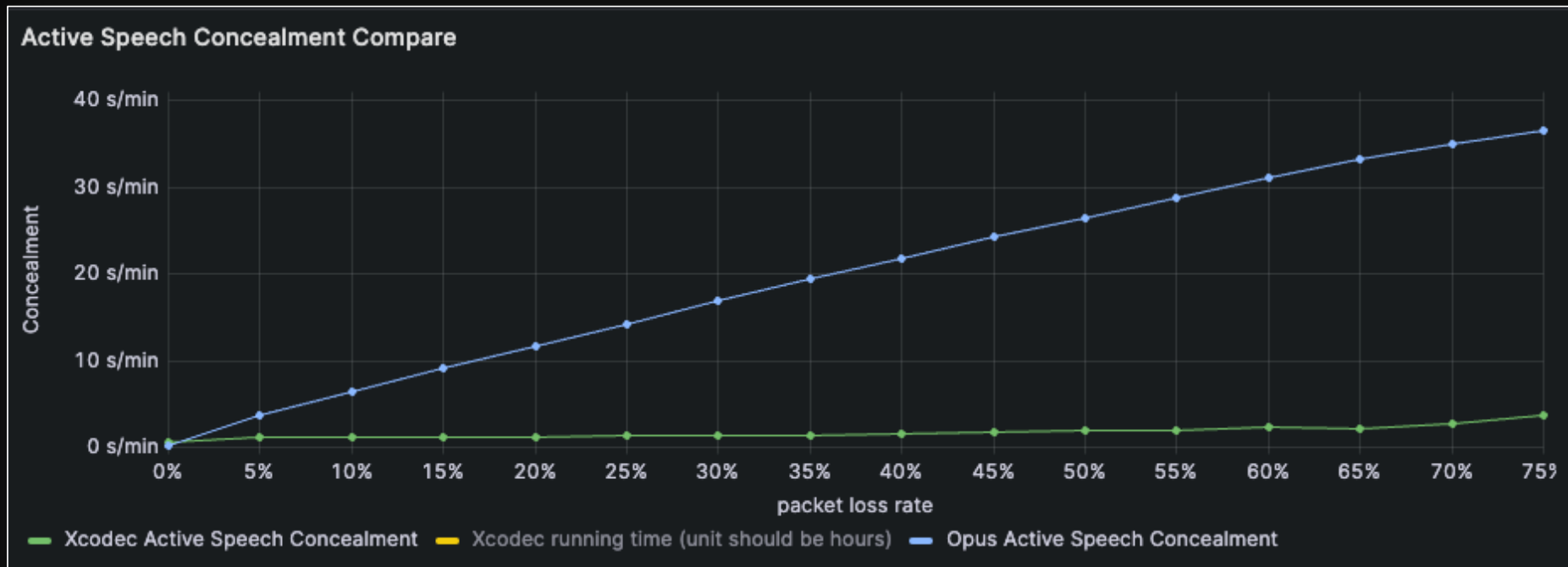


# INTRODUCING AI Audio Codec

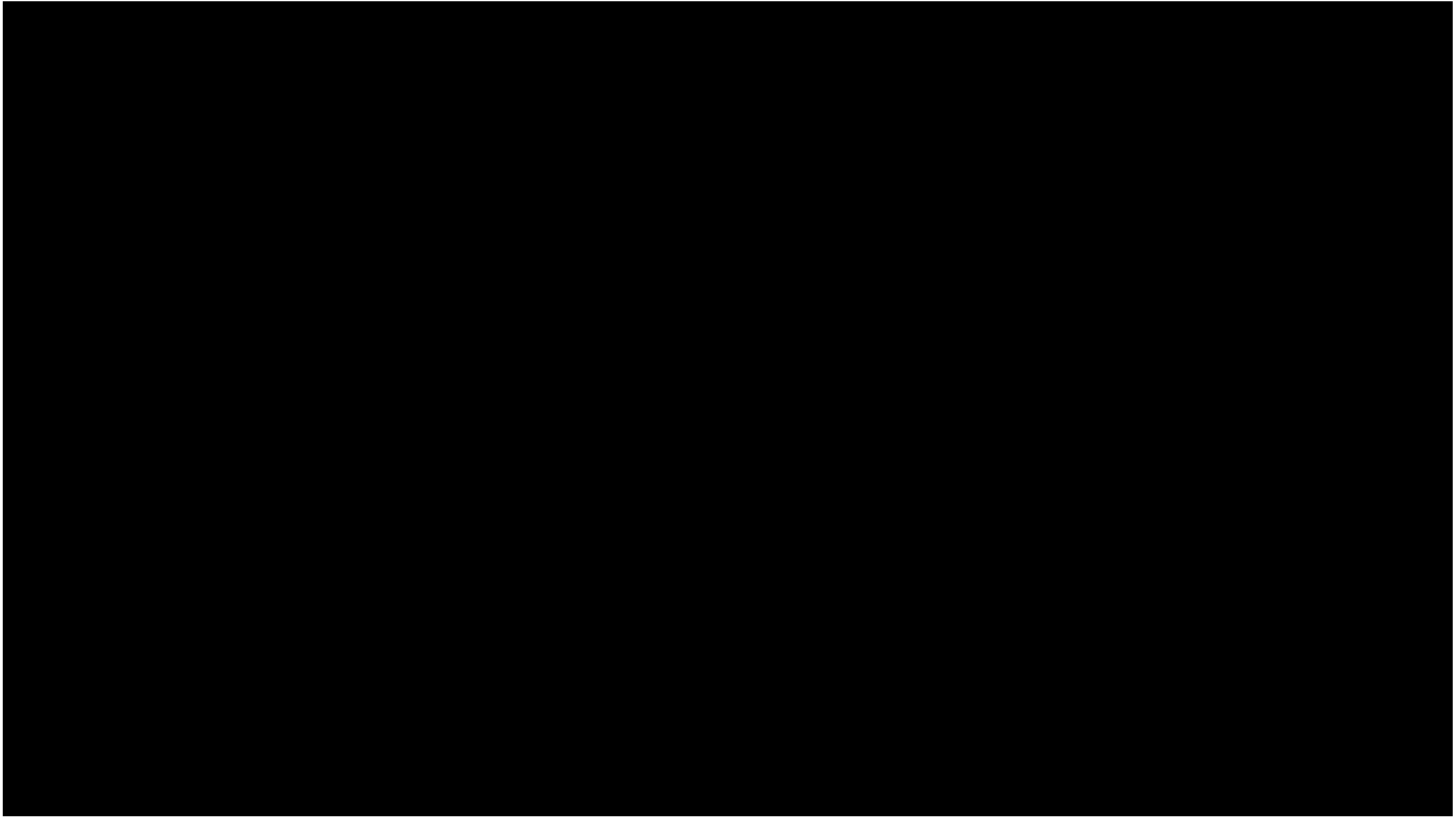
Optimal quality audio at a fraction of the bandwidth

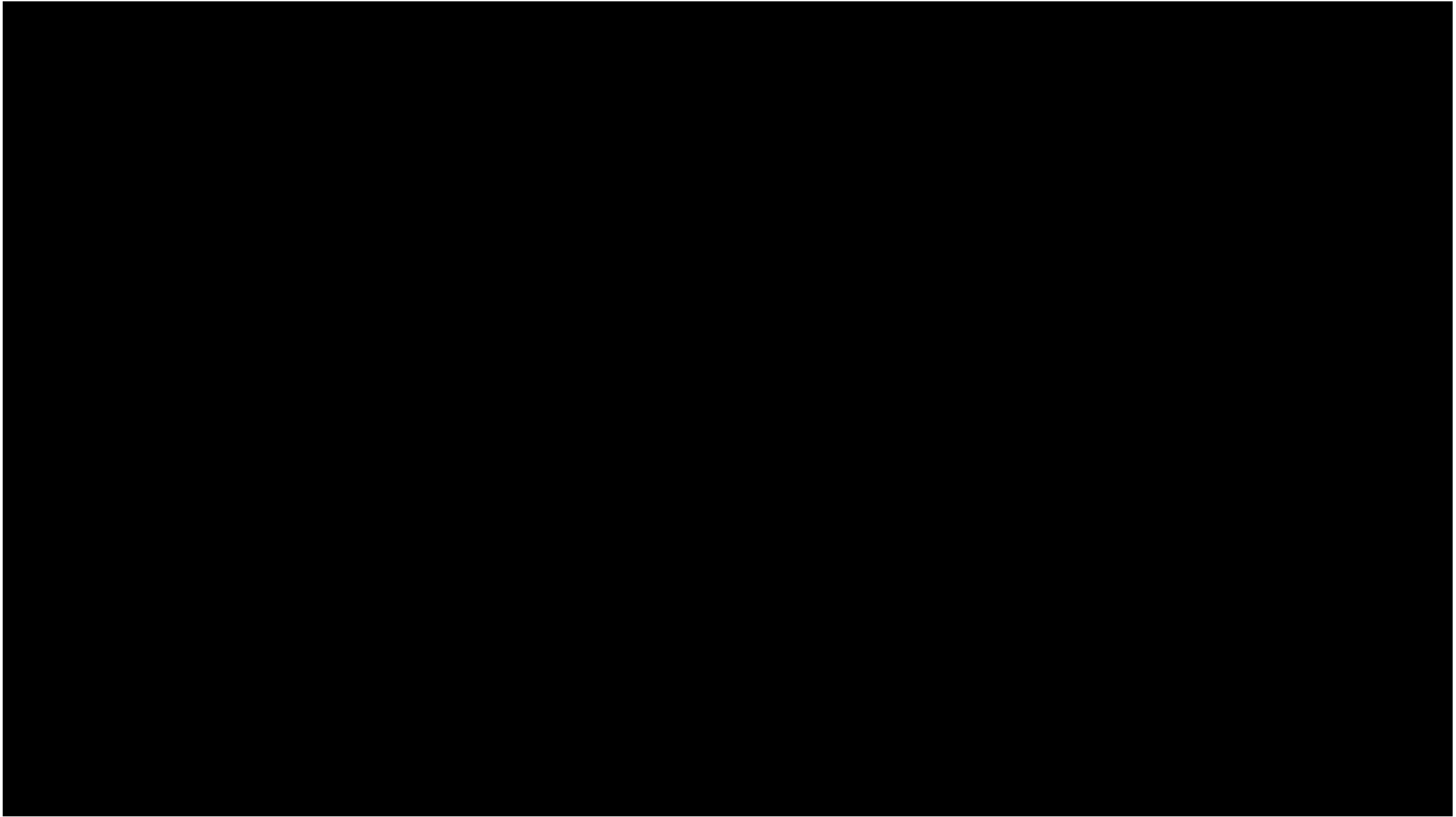




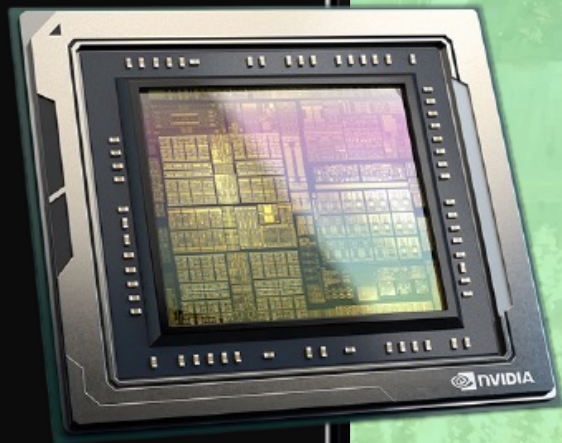


Even with high packet loss, AI Codec maintains speech concealment close to 0 sec, while Opus' speech concealment increases linearly with packet loss.



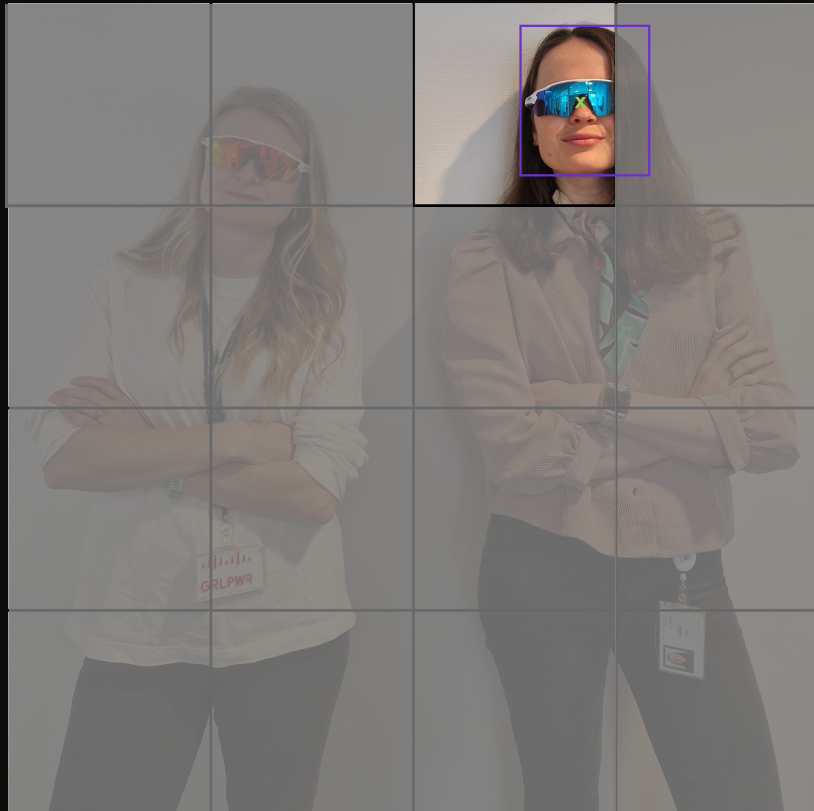


# Processing Video



webex

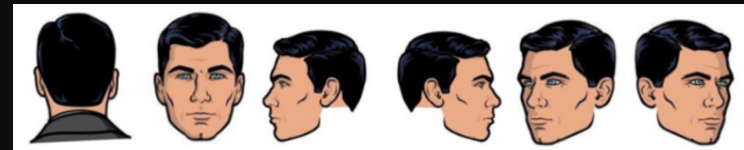
# Output of Head Detect model



The last 6 numbers are probabilities that the head belongs to one of 6 *head pose classes*

$$[t_x, t_y, t_w, t_h, P_{obj}, P_0, \dots, P_{C-1}]$$

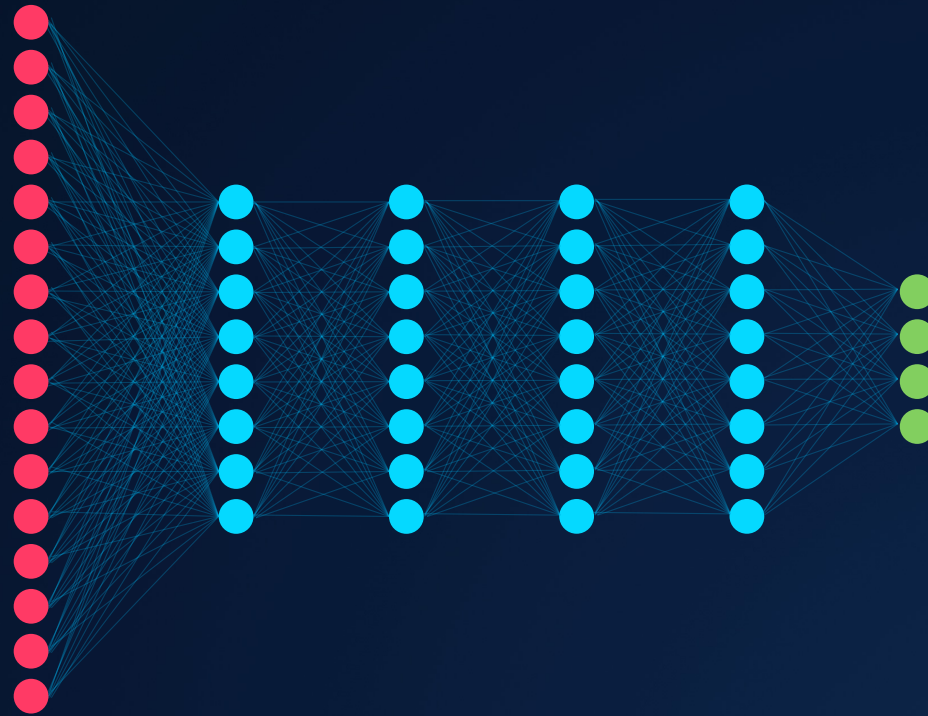
0 = head (back of the head)  
1 = frontal  
2 = full\_left  
3 = full\_right  
4 = half\_left  
5 = half\_right



# Artificial neural networks for machine vision



Source  
pixels

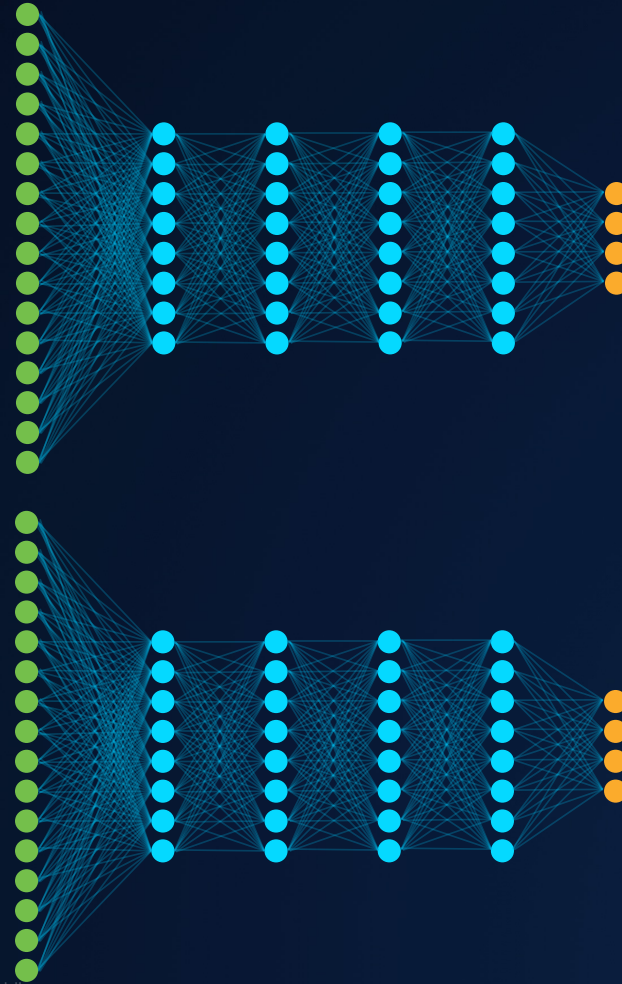


Background  
segmentation



# Optimizing the Architecture

Image  
Input



Background  
segmentation

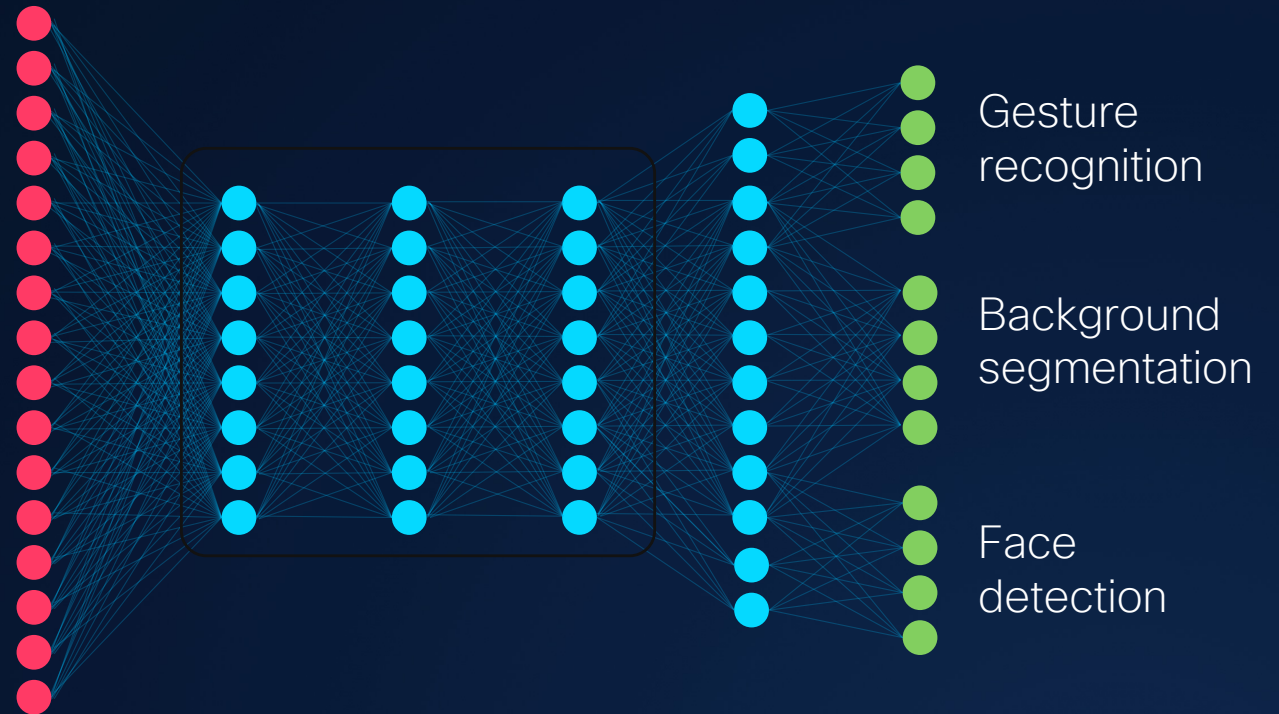
Face  
detection



# Artificial neural networks for machine vision



Source  
pixels



# From Lab to Product

# Software

- Portability
  - Windows, Mac, Linux
  - iOS, Android
  - Cisco-specific hardware (Linux appliance)
- Modularity
  - activate features based on the hardware capabilities and performance
- Stability
- Privacy & Security
  - Responsible AI

# Hardware

- future-proof
- 5-7 years lifespan
- FPGA -> GPU



Cisco AI Agent

# AI Agent

Autonomous interactions,  
natural-sounding language,  
and true self service

# AI Assistant

Suggested responses,  
automation of tasks, and  
greater agent efficiency

INTRODUCING

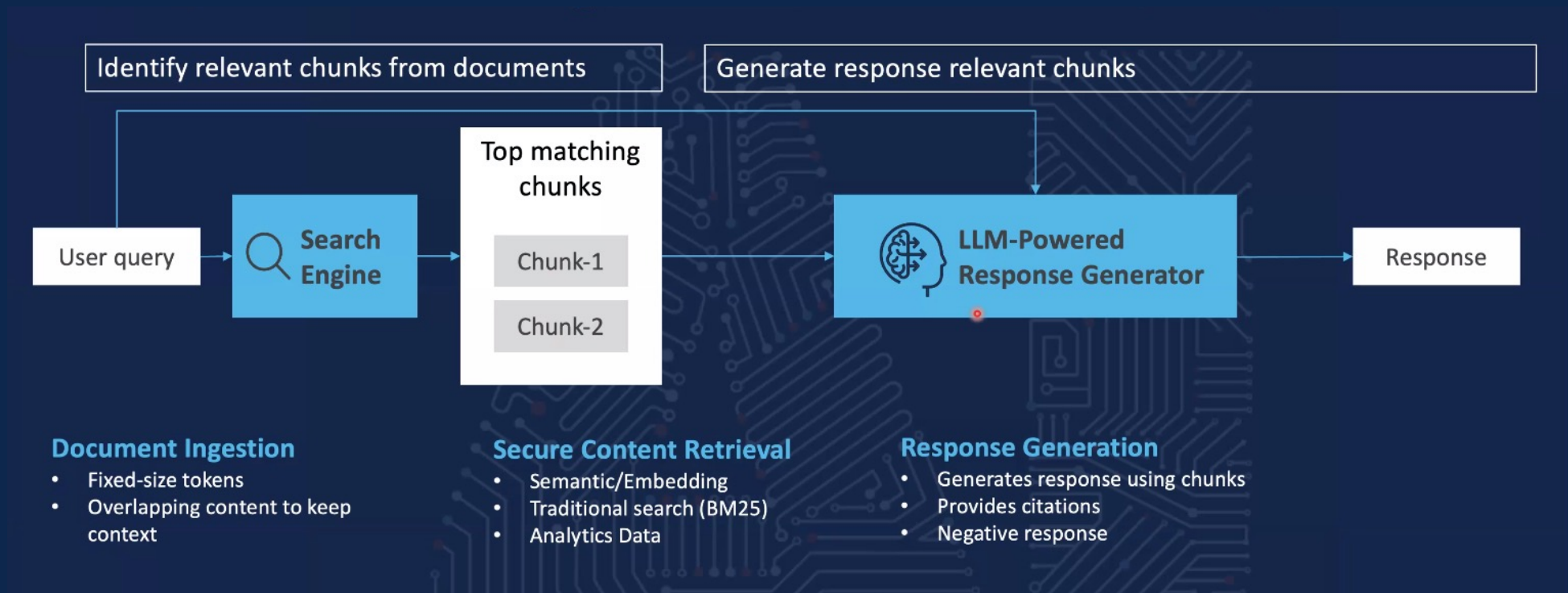
# Webex AI Agent

# Webex AI Agent





# Retrieval Augmented Generation (RAG)

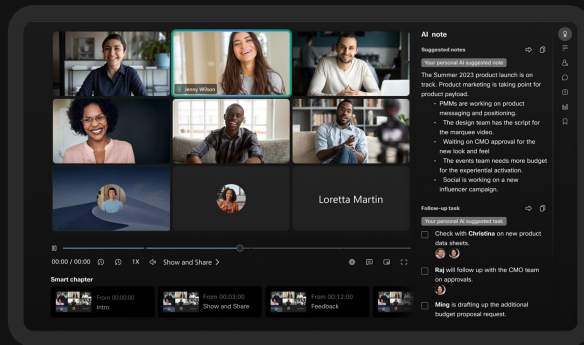


# Demo

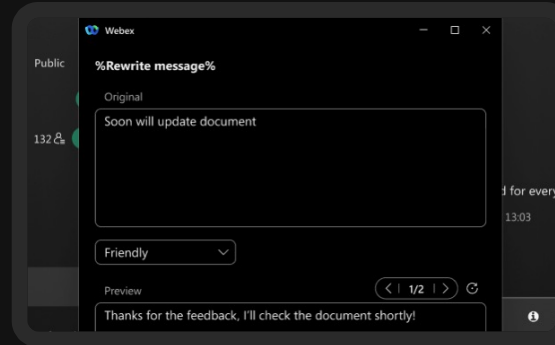


AI Assistants

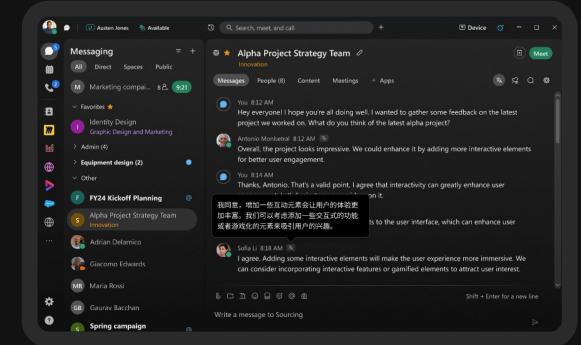
# AI Assistants in Webex



Catch up quickly with  
Meeting, Messaging, Vidcast,  
and Slido Summaries



Send the perfect message with  
Change the Tone



Communicate globally with  
Translation in 30+ languages

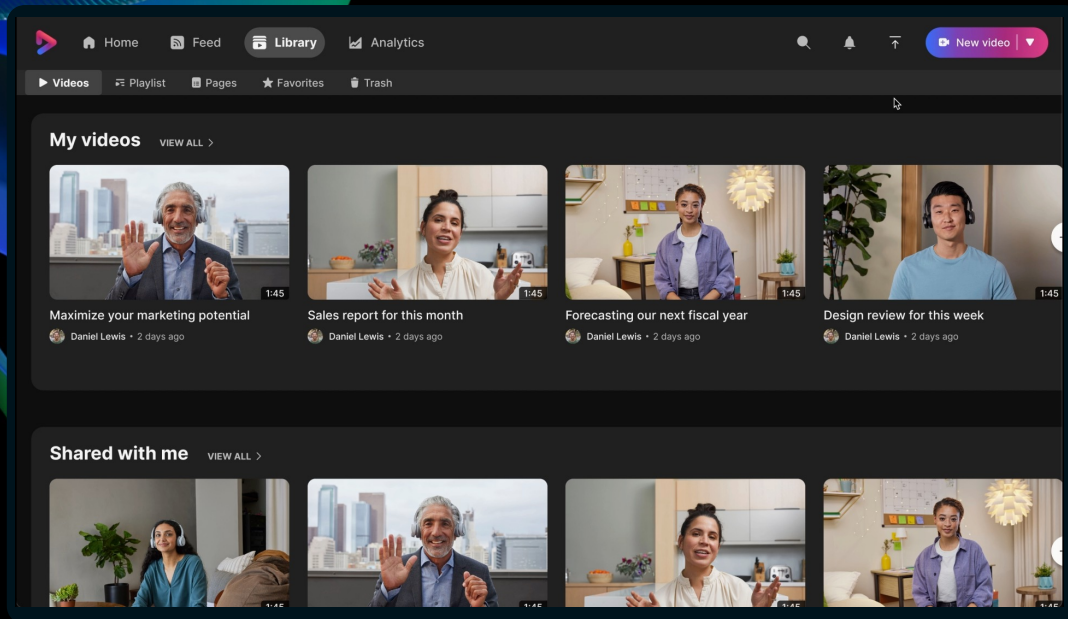
ANNOUNCING

# Elevate engagement with AI-generated Vidcast

No script needed—generate scripts from  
slide content, with full editing control

Translate into language of your choice

Choose your tone and style



ANNOUNCING

# Foster inclusivity with AI- generated Polls

Generates Slido polls based on  
presentation content

Increases participant engagement

Reduces workload and barriers to  
adoption for presenters

