



Summary

Today's work focused on preparing the Cloud9 environments for future GPU-enabled TTS deployment and establishing the Lambda → Cloud9 API bridge for integration. The FishSpeech model was installed and set up for text-to-speech inference, with plans to replace its web UI dependency with a callable Python function. The TTS environment is now connected to Lambda for end-to-end testing, though GPU installation and full call flow integration remain pending.

오늘은 GPU 기반 TTS 배포를 위한 Cloud9 환경 준비와 Lambda → Cloud9 API 브리지 구축에 집중했습니다. FishSpeech 모델을 설치 및 구성하여 텍스트를 음성으로 변환할 수 있는 환경을 마련했으며, 현재는 Web UI 의존성을 제거하고 직접 호출 가능한 Python 함수 형태로 대체할 계획입니다. TTS 환경은 Lambda 와 연동되어 종단 간 테스트가 가능한 상태이며, GPU 설치와 전체 콜 플로우 통합은 아직 진행 중입니다.

English

1. Summary of Accomplishments

1. Cloud9 Environment Setup for GPU-Ready TTS

- Reviewed Cloud9 environment types and instance settings to prepare for GPU-based deployment in the future.
- Created a dedicated TTS environment for integration with the STT API.
- Installed core dependencies for FishSpeech, including required libraries.

2. FishSpeech Model Setup

- Obtained and authenticated access to the `fishaudio/openaudio-s1-mini` model on HuggingFace.
- Verified model inference readiness for speech generation (pending GPU acceleration).

3. Lambda → Cloud9 API Bridge

- Developed and tested a Python API on Cloud9 to receive and process text input.
- Configured Cloud9 security group to allow inbound API requests from Lambda.
- Verified successful Lambda call to Cloud9 API returning processed output.

4. Refactoring for Automation

- Began development to remove reliance on FishSpeech's web UI.
- Planned extraction of core inference logic into a callable Python function.

2. Ongoing Tasks

- Install GPU drivers, CUDA, and PyTorch with GPU support in the TTS Cloud9 environment (requires GPU-backed instance).
 - Finalize Python inference code to convert text into speech without web UI.
 - Integrate TTS and STT APIs for seamless processing.
 - Extend current Lambda integration to accept Amazon Connect call flow input and return generated TTS output.
 - Perform full end-to-end testing with real call scenarios.
 - Finalize deployment documentation for reproducibility.
-

한국어

1. 주요 작업 요약

1. GPU 기반 TTS 준비를 위한 Cloud9 환경 구성

- 향후 GPU 기반 배포를 위해 Cloud9 환경 유형과 인스턴스 설정 검토.
- STT API 와 연동 가능한 전용 TTS 환경 생성.
- FishSpeech 실행에 필요한 주요 라이브러리 설치.

2. FishSpeech 모델 설정

- HuggingFace `fishaudio/openaudio-s1-mini` 모델 접근 권한 획득 및 인증 완료.
- 음성 생성을 위한 모델 추론 준비 완료(단, GPU 가속은 대기 중).

3. Lambda → Cloud9 API 브리지

- Cloud9에서 텍스트 입력을 수신·처리하는 Python API 개발 및 테스트.
- Cloud9 보안 그룹 설정을 수정하여 Lambda에서 API 요청 허용.
- Lambda 출력을 통한 Cloud9 API 응답 반환 성공 검증.

4. 자동화를 위한 리팩토링

- FishSpeech Web UI 의존성 제거를 위한 개발 착수.
 - 핵심 추론 로직을 직접 호출 가능한 Python 함수로 추출할 계획 수립.
-

2. 다음 작업 예정

- GPU 인스턴스 전환 후 GPU 드라이버, CUDA, GPU 지원 PyTorch 설치.
 - Web UI 없이 텍스트를 음성으로 변환하는 Python 추론 코드 완성.
 - TTS ↔ STT API 통합으로 원활한 처리 구현.
 - Amazon Connect 콜 플로우 입력을 수신하고 생성된 TTS 를 반환하는 Lambda 통합 확장.
 - 실제 콜 시나리오를 통한 종단 간 전체 테스트 진행.
 - 재현 가능한 배포를 위한 문서화 최종 완성.
-

Remarks / 참고 사항

- Lambda 와 Cloud9 간의 API 브리지는 정상적으로 동작하여 향후 전체 파이프라인 통합 시 핵심 역할을 수행할 예정입니다.
- GPU 가속 및 전체 콜 플로우 테스트가 완료되면, 실시간 음성 생성 및 재생 기능을 안정적으로 제공할 수 있을 것으로 예상됩니다.