



VoiceBase's reduced its software distribution time by 88%

Introduction

San Francisco, California-based [VoiceBase](#) provides speech-to-text for audio and video transcription, as well as tools for knowledge extraction, speech analytics, and predictive analytics to analyze spoken information. Its customers use these tools to monitor agents, analyze customer requests, coach sales reps, identify upsell opportunities in call centers and many more applications.

Challenge: Distributing 50GB to 400 servers, on a regular basis

The most demanding aspect of transcription is accuracy. VoiceBase's speech models have to be trained to recognize accents, include custom vocabularies for professions like medicine and legal, as well as handle multiple languages. At the outset, speech model files were around 500MB but with the constantly improving software, today they are closer to 10GB each.

For its speech tools, VoiceBase leverages barebone servers collocated across several sites. Each collocation contains a few racks, holding a switch, a router and as much hardware as it can hold. Since these are used for pure compute power, there is no redundancy, except that each rack is connected via a different ISP.

When speech models need to be distributed to the voice servers, a whopping 50GB of files have be disseminated to over 400 production servers as fast as possible. The main bottleneck was found to be within the data center itself, where 1GB cables feed each rack.

With new speech models being released every two-to-four-weeks, the distribution day used to be dreaded by the VoiceBase DevOps team. The dissemination process would take around eight-hours. Inside the data center the team was running a combination of homemade scripts based on open source tools like rsync and NFS. The distribution would routinely get bogged down because of expired passwords, changed paths, bottlenecks downloading from the servers, and most problematic; distributing from one file server to 30-40 machines within a rack.



Solution: Peer-to-peer file distribution

Christian Bryndum, VoiceBase's Director of Operations, knew he had to move to a peer-to-peer (P2P) based solution. Investing in very expensive routers for each rack would not speed up the distribution over the entire path and any TCP/IP-based option would result in the same roadblocks.

VoiceBase looked into several P2P distribution solutions, including Twitter's open-sourced Murder, which is based on the BitTorrent protocol. While investigating Murder, VoiceBase encountered Resilio Connect, which also is based on the very efficient BitTorrent protocol.

VoiceBase put Resilio Connect through meticulous testing on both its QA and Dev environments. Once Resilio Connect passed the testing phase, the team started working on deploying it into production. Since the dissemination process spans many locations, racks and servers, every step had to be automated, leaving no human intervention in the distribution process.

Results

The positive impact of deploying Resilio Connect was felt immediately with the new, highly-reliable dissemination process only taking 12-percent of the original download time!

"Resilio Connect enables us to reliably distribute our code, specifically new language models in a fraction of time. These copy jobs now take an hour, down from eight," says Bryndum, "Best of all, once Resilio Connect was installed, it just works: We never need to manually intervene in any way."

Industry

Tech