Amazon Alexa: How to talk with your Smart Home devices

Tsypuk Roman
- @tsypuk_r
- tsypuk.conf@gmail.com

Sofia, Bulgaria 2019
Staff Engineer @ Lohika Altran-Group

JVM-based languages: Java, Kotlin, Groovy

Spring ecosystem, Distributed Systems, AWS, Kubernetes, Data Processing and Streaming, Microservices, Security,

DevOps & Testing Practices, Infrastructure, IoT

Radio HAM, Software Defined Radio
Agenda

1. Evolution and stats for interaction
2. Amazon Alexa
3. Flash Brief Alexa skill
4. AWS Serverless Lambdas and Custom skill
5. Smart Home devices integration
6. Network and Messaging in depth
7. NodeMCU Wemo-compatible prototype
8. “Local Lambda” SW Gateway
Human-Machine Interaction Evolution

Source: University of Calgary, "History of Computer Interfaces" (Saul Greenberg)
Stats

Facts

September 2014
Baidu – 1 in 10 queries come through speech.

May 2016
Bing – 25% of searches performed on Windows 10 taskbar are voice searches per Microsoft reps.

2020
In five years time at least 50% of all searches are going to be either through images or speech.
Andrew Ng
Chief Scientist, Baidu (9/14)

June 2015
Siri – handles more than 1 billion requests per week through speech.

2015
Amazon Echo – fastest-selling speaker in 2015, @ for ~25% of USA speaker market, per 1010data.

May 2016
Android – 1 in 5 searches on mobile app in USA are voice searches & share is growing.

@KPCB
Voice Search Stat

- Google: 20% of all searches are voice
- 31% of smartphone users worldwide use voice tech at least once a week.
- 50% of all online searches will be voice-based by 2020.
- 30% of all web browsing sessions will be done without a screen by 2020.
- 55% of households are expected to own smart speaker devices by 2022.
- Playing music is the most popular use of smart speakers currently.
- Almost 20% of all voice search queries are triggered by just 25 keywords.
- Global voice commerce is expected to be worth $40B by 2022.

https://99firms.com/blog/voice-search-statistics/#gref
“Cloud-based voice service available on tens of millions of devices from Amazon and third-party device manufacturers. With Alexa, you build natural voice experiences interact models”

NLP, voice recognition, self educating AI
Alexa features
BMW, MINI & Ford Alexa integration

“By making this step of integrating Alexa into our models from mid-2018, BMW and MINI will form a more intrinsic part of our customers’ digital lifestyles,” remarked Dieter May, Senior Vice President Digital Services and Business Models at the BMW Group.

Today, 8.5 million BMW Group models worldwide are already connected. From mid-2018 Amazon Alexa is be available in all BMW and MINI vehicles in USA, UK and Germany.

https://www.youtube.com/watch?v=l-uHGOUpLlg
Skills

**Flash Briefing**
defines audio/video/text content added to the brief of news

“Alexa, tell me news”

**Custom**
everything is specified by developer

**Smart Home**
controls IoT devices like lights and thermostats

“Alexa, dim the light”

**Video**
video content is sent in response

“Alexa change to CNN”
Skills: Flash Brief

Diagram showing the integration of Amazon Alexa with Amazon Web Services (AWS) components:

1. XML data flow
2. JSON data flow
3. API Gateway
   - Amazon EC2
   - Lambda

Diagram illustrates the flow of data and services between these components.
Flash Brief’s endpoint requirements

- non-password-protected HTTPS endpoint
- either text (4500 characters) or audio content (256kbps mono or stereo MP3 10 minutes)
- JSON or RSS UTF-8 encoded
- available 24 hours a day/7 days a week
- items order from newest to oldest, based on the date. Alexa may ignore older items.
- each item in the feed should be unique and should not overlap with content in another feed item.
Demo

Flash Brief Skill

“JAVA NEWS”

Create Flash brief skill from the scratch

Alexa tells us latest Java world news

“Alexa give me latest news”
Skills: Custom skill

WebService
On-Premises or Cloud

1
 Lambda

2
Router for invocations

invocation
Alexa, ask “greeter” who is the winner?
Utterance-to-Intent mapping

Alexa, ask “greeter” who is the winner ?

utterance

utterances many-to-one

- find the winner
- who is lucky
- **who is the winner**
- get the winner
- winner

Lambda

winner

intent1

... intentN
“Greeter”

Kotlin-based lambda serverless.

Alexa will choose the winner from the audience.

“Alexa, ask greeter who is present?”

“How about our greeter winner?”
AWS Lambda
### AWS Lambda

<table>
<thead>
<tr>
<th>Provisioning</th>
<th>Need to provision, manage server</th>
<th>Serverless compute service</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Charge</strong></td>
<td><strong>On-demand</strong> charged hourly based on compute capacity, <strong>reserved</strong>, <strong>spot</strong></td>
<td>Charged for 100ms running and number of time is triggered, no idle time payment</td>
</tr>
<tr>
<td><strong>Scaling</strong></td>
<td>Need DNS, ELB, autoscaling, plus correct configuration and other techniques</td>
<td>Scales in parallel on each trigger</td>
</tr>
<tr>
<td><strong>Security</strong></td>
<td>Need manual security OS updates</td>
<td>No infrastructure to maintain</td>
</tr>
<tr>
<td><strong>Interaction</strong></td>
<td>Networking knowledge, routes to infrastructure</td>
<td>Triggered (API GW, IoT, ASK, Kinesis, S3, SNS, CodeCommit, CloudWatch, CloudFront)</td>
</tr>
<tr>
<td><strong>Runtime</strong></td>
<td>&lt;Any&gt;</td>
<td>C#, GO, Java, NodeJS, Python code</td>
</tr>
</tbody>
</table>
IoT Smart Home
Alexa in Wireshark

Frame 1: 149 bytes on wire (1192 bits), 149 bytes captured (1192 bits)

- User Datagram Protocol, Src Port: 50000 (50000), Dst Port: ssdp (1900)
- Simple Service Discovery Protocol
  - M-SEARCH * HTTP/1.1

[Expert Info (Chat/Sequence): M-SEARCH * HTTP/1.1]
Request Method: M-SEARCH
Request URI: *
Request Version: HTTP/1.1
HOST: 239.255.255.250:1900
MAN: "ssdp:discover"
MX: 15
ST: urn:schemas-upnp-remotecontrol:device:*

[Full request URI: http://239.255.255.250:1900/*]
[HTTP request 1/4]
[Next request in frame: 2]
Sequence diagram: Search devices

"Alexa, find my devices"

"Starting discovery"

"I have found 1 new device..."
"Alexa, turn meeting room on"

TCP: HTTP SOAP
POST ip:port/control/basicevent1

"OK"

TCP: HTTP SOAP
RESPONSE
Prototyping with ESP-8266

- Tensilica 32-bit RISC CPU Xtensa LX106
- Digital I/O Pins: 16  Analog Input Pins: 1
- UARTs: 1  SPIs: 1  I2Cs: 1
- Flash Memory: 4 MB
- SRAM: 16/32/64 KB
- Clock Speed: 80 Mhz
- Wi-Fi: IEEE 802.11 b/g/n
Demo

ESP-based light control

Room
Kitchen
Hall
End-to-End: How it works
Why not to create “Local Lambdas” for Alexa?
Demo Smart Gateway
1. Voice evolution
   https://voice-report.com/voice-is-here-and-it-is-here-right-now-heres-how-we-know/
   https://99firms.com/blog/voice-search-statistics/#gref

2. BMW announce of Alexa devices promo
   https://www.youtube.com/watch?v=mHCI6lM_Zl8

3. Flash Brief requirements
   https://developer.amazon.com/docs/flashbriefing/steps-to-create-a-flash-briefing-skill.html#create-a-flash-briefing-skill

4. Arduino WEMO switch

5. Alexa voice AI with filtering by AWS
   https://www.youtube.com/watch?v=2Bazibaz1F8
“THANK YOU”

Alexa, ask “greeter” say goodbye