

```

1 Lab7. API를 통한 Kafka 실습
2
3 1. API 서버 역할의 EC2 인스턴스 생성
4   1)인스턴스 목록에서 [인스턴스 시작] 버튼 클릭
5   2)[인스턴스 시작] 페이지에서
6     -[이름] : {계정}-kafka-api-ec2
7     -[애플리케이션 및 OS 이미지] > [Quick Start] > Ubuntu
8     -Ubuntu Server 22.04 LTS (HVM), SSD Volume Type, 64비트(x86)
9     -[인스턴스 유형] : t2.micro
10    -[키 페어(로그인)] > [새 키 페어 생성]
11    -[키 페어 생성] 창에서,
12      --[키 페어 이름] : {계정}-kafka-api-ec2-key
13      --나머지는 기본값 그대로 사용
14      --[키 페어 생성] 버튼 클릭
15      --적당한 위치에 pem 파일 다운로드
16    -[네트워크 설정] > [편집] 버튼 클릭
17    -[VPC] : {계정}-datalake-vpc
18    -[서브넷] : {계정}-datalake-subnet-2a
19    -[퍼블릭 IP 자동 할당] : 활성화
20    -[방화벽(보안 그룹)] : 보안 그룹 생성
21    -[보안 그룹 이름] : {계정}-kafka-api-sg
22    -[설명] : Security group for Kafka API Server
23    -[스토리지 구성] : 30 GiB, gp2
24    -[인스턴스 시작] 버튼 클릭
25
26
27 2. API Server 사용 준비
28   1)방금 생성한 API EC2 서버의 [상태 검사]가 "2/2개 검사 통과"가 되면 Tabby를 통해 SSH 연결한다.
29     -Tabby 설정창에서 [New profile] 버튼 클릭
30     -[Name] : Kafka-api-server
31     -[Group] : kafka
32     -[Host] : {계정}-kafka-api-ec2의 퍼블릭 IPv4 DNS 값
33     -[Username] : ubuntu
34     -[Authentication method] : Key
35     -[Private keys] > [Add a private key] 버튼 클릭하여 {계정}-kafka-api-ec2-key.pem 파일 연결
36     -[Save] 버튼 클릭
37     -[Profiles] 목록에서 "kafka-api-server" 연결
38
39   2)API 서버에 필요한 패키지 설치하기
40     $ sudo apt update
41     $ sudo apt install python3-pip
42     $ pip3 install flask
43     $ pip3 install requests bs4
44     $ pip3 install confluent-kafka
45
46   3)app.py 파일 생성
47     $ nano app.py
48
49   4)API Server의 보안 그룹에 port 5000 추가
50     -{계정}-kafka-api-ec2 인스턴스의 상세페이지에서 [보안] 탭 클릭
51     -[보안 그룹]에 있는 {계정}-kafka-api-sg 링크 클릭
52     -{계정}-kafka-api-sg 페이지에서 [인바운드 규칙] 탭 선택
53     -[인바운드 규칙 편집] 버튼 클릭
54     -[인바운드 규칙 편집] 페이지에서 [규칙 추가] 버튼 클릭
55     -[유형] : 사용자 지정 TCP
56     -[포트 범위] : 5000
57     -[소스] : Anywhere IPv4(0.0.0.0/0)
58     -[규칙 저장] 버튼 클릭
59
60   5)API Server 접근하기
61     -app.py 실행하기
62       $ flask run --host=0.0.0.0
63
64     -웹 브라우저에서 확인
65       http://{계정}-kafka-api-ec2의 퍼블릭 IPv4 주소:5000
66       Hello, World
67
68       http://{계정}-kafka-api-ec2의 퍼블릭 IPv4 주소:5000/api
69       {"link": "http://www.yes24.com/Product/Goods/102418815", "title": "Let's Get IT 파이썬 프로그래밍", "author": "안지혜",
70        "publisher": "길벗", "pub_date": "2021년 07월", "price": "19,800원"}
71
72 3. kafka producer 설정하기
73   1)사전 설치
74     $ pip3 install requests confluent-kafka
75     $ pip3 install urllib3==1.26.15
76
77   2)producer2.py 파일 생성
78     $ pwd
79     /home/ec2-user/kafka-client/producer
80
81     $ nano producer2.py
82
83   3)새 topic 생성하기

```

```

84 -'{계정}-topic3' 생성
85 $ pwd
86 /home/ec2-user/kafka_2.12-2.8.1/bin
87
88 $ ./kafka-topics.sh --create --bootstrap-server
b-2.henrymskcluster.4ka9rz.c4.kafka.ap-northeast-2.amazonaws.com:9092,b-1.henrymskcluster.4ka9rz.c4.kafka.ap-nort
heast-2.amazonaws.com:9092 --replication-factor 2 --partitions 3 --topic henry-topic3
Created topic henry-topic3.
89
90
91 -topic 확인
92 $ ./kafka-topics.sh --list --bootstrap-server
b-2.henrymskcluster.4ka9rz.c4.kafka.ap-northeast-2.amazonaws.com:9092,b-1.henrymskcluster.4ka9rz.c4.kafka.ap-nort
heast-2.amazonaws.com:9092
93 __amazon_msk_canary
94 __consumer_offsets
95 henry-topic
96 henry-topic2
97 henry-topic3
98
99 4)producer2.py 테스트
100 -먼저 API Server를 실행한다.
101 -API Server를 연결해서
102 $ flask run --host=0.0.0.0
103
104 -Producer EC2 인스턴스에서 producer2.py 실행
105 $ pwd
106 /home/ec2-user/kafka-client/producer
107
108 $ python3 producer2.py
109 Message delivered to henry-topic3 [0]
110 Message delivered to henry-topic3 [1]
111 Message delivered to henry-topic3 [2]
112 Message delivered to henry-topic3 [1]
113 Message delivered to henry-topic3 [0]
114 Message delivered to henry-topic3 [0]
115 Message delivered to henry-topic3 [1]
116
117
118 4. kafka consumer 설정하기
119 1)사전 설치
120 $ pip3 install requests confluent-kafka
121
122 2)consumer2.py 파일 생성
123 $ pwd
124 /home/ec2-user/kafka-client/consumer
125
126 $ nano consumer2.py
127
128 3)consumer2.py 테스트
129 -먼저 API Server를 실행한다.
130 -API Server를 연결해서
131 $ flask run --host=0.0.0.0
132
133 -그다음, Producer EC2 인스턴스에서 producer2.py 실행
134 $ pwd
135 /home/ec2-user/kafka-client/producer
136
137 $ python3 producer2.py
138 Message delivered to henry-topic3 [0]
139 Message delivered to henry-topic3 [1]
140 Message delivered to henry-topic3 [2]
141 Message delivered to henry-topic3 [1]
142 Message delivered to henry-topic3 [0]
143 Message delivered to henry-topic3 [0]
144 Message delivered to henry-topic3 [1]
145
146 -마지막으로 Consumer EC2 인스턴스에서 consumer2.py 실행
147 $ pwd
148 /home/ec2-user/kafka-client/consumer
149
150 $ python3 consumer2.py
151 Received message: {'link': 'http://www.yes24.com/Product/Goods/96262613', 'title': '모두의 인공지능 with 파이썬', 'author':
'이영호', 'publisher': '길벗', 'pub_date': '2020년 12월', 'price': '19,800원'}
152 Received message: {'link': 'http://www.yes24.com/Product/Goods/97831032', 'title': '게임으로 배우는 파이썬', 'author': '다나카
겐이치로', 'publisher': '영진닷컴', 'pub_date': '2021년 03월', 'price': '9,900원'}
153 Received message: {'link': 'http://www.yes24.com/Product/Goods/107464875', 'title': '웹 크롤링 & 데이터 분석 with 파이썬',
'author': '장철원', 'publisher': '인사이트(insight)', 'pub_date': '2022년 02월', 'price': '16,000원'}
154 Received message: {'link': 'http://www.yes24.com/Product/Goods/97032765', 'title': '선형대수와 통계학으로 배우는 머신러닝 with
파이썬', 'author': '장철원', 'publisher': '비제이퍼블릭(BJ퍼블릭)', 'pub_date': '2021년 01월', 'price': '33,750원'}
155 Received message: {'link': 'http://www.yes24.com/Product/Goods/108947478', 'title': 'Do it! 쉽게 배우는 파이썬 데이터 분석',
'author': '김영우', 'publisher': '이지스퍼블리싱', 'pub_date': '2022년 05월', 'price': '20,700원'}
156 Received message: {'link': 'http://www.yes24.com/Product/Goods/116864757', 'title': '혼자 공부하는 데이터 분석 with 파이썬',
'author': '박해선', 'publisher': '한빛미디어', 'pub_date': '2023년 01월', 'price': '18,720원'}
157 Received message: {'link': 'http://www.yes24.com/Product/Goods/75193791', 'title': 'Do it! 점프 투 파이썬 -전면 개정판',

```

158 'author': '박응용', 'publisher': '이지스퍼블리싱', 'pub\_date': '2019년 06월', 'price': '13,000원'}  
Received message: {'link': 'http://www.yes24.com/Product/Goods/106349915', 'title': '파이썬으로 만드는 나만의 게임', 'author':  
159 '이수안', 'publisher': '비제이퍼블릭(BJ퍼블릭)', 'pub\_date': '2022년 01월', 'price': '22,500원'}  
Received message: {'link': 'http://www.yes24.com/Product/Goods/97032765', 'title': '선형대수와 통계학으로 배우는 머신러닝 with  
파이썬', 'author': '장철원', 'publisher': '비제이퍼블릭(BJ퍼블릭)', 'pub\_date': '2021년 01월', 'price': '33,750원'}  
160 Received message: {'link': 'http://www.yes24.com/Product/Goods/74258258', 'title': '파이썬 머신러닝 판다스 데이터 분석',  
'author': '오승환', 'publisher': '정보문화사', 'pub\_date': '2019년 06월', 'price': '22,500원'}  
161 Received message: {'link': 'http://www.yes24.com/Product/Goods/88922648', 'title': '머신러닝을 위한 파이썬 한 조각', 'author':  
'박성호', 'publisher': '비제이퍼블릭(BJ퍼블릭)', 'pub\_date': '2020년 02월', 'price': '19,200원'}  
162 Received message: {'link': 'http://www.yes24.com/Product/Goods/67025877', 'title': '윤성우의 열혈 파이썬 기초편', 'author':  
'윤성우', 'publisher': '오렌지미디어', 'pub\_date': '2018년 11월', 'price': '14,400원'}  
163 Received message: {'link': 'http://www.yes24.com/Product/Goods/80855594', 'title': '파이썬 웹 프로그래밍, 실전편', 'author':  
'김석훈', 'publisher': '한빛미디어', 'pub\_date': '2019년 11월', 'price': '28,800원'}  
164 Received message: {'link': 'http://www.yes24.com/Product/Goods/106349915', 'title': '파이썬으로 만드는 나만의 게임', 'author':  
'이수안', 'publisher': '비제이퍼블릭(BJ퍼블릭)', 'pub\_date': '2022년 01월', 'price': '22,500원'}  
165 Received message: {'link': 'http://www.yes24.com/Product/Goods/116864757', 'title': '혼자 공부하는 데이터 분석 with 파이썬',  
'author': '박해선', 'publisher': '한빛미디어', 'pub\_date': '2023년 01월', 'price': '18,720원'}  
166 Received message: {'link': 'http://www.yes24.com/Product/Goods/107490270', 'title': '만들면서 배우는 파이썬과 40개의 작품들',  
'author': '장문철', 'publisher': '앤써북', 'pub\_date': '2022년 02월', 'price': '16,920원'}