

PROJECT.

**ELK STACK**

(ELASTICSEARCH + KIBANA + LOGSTACH)

**+**

**APACHE NIFI**

01

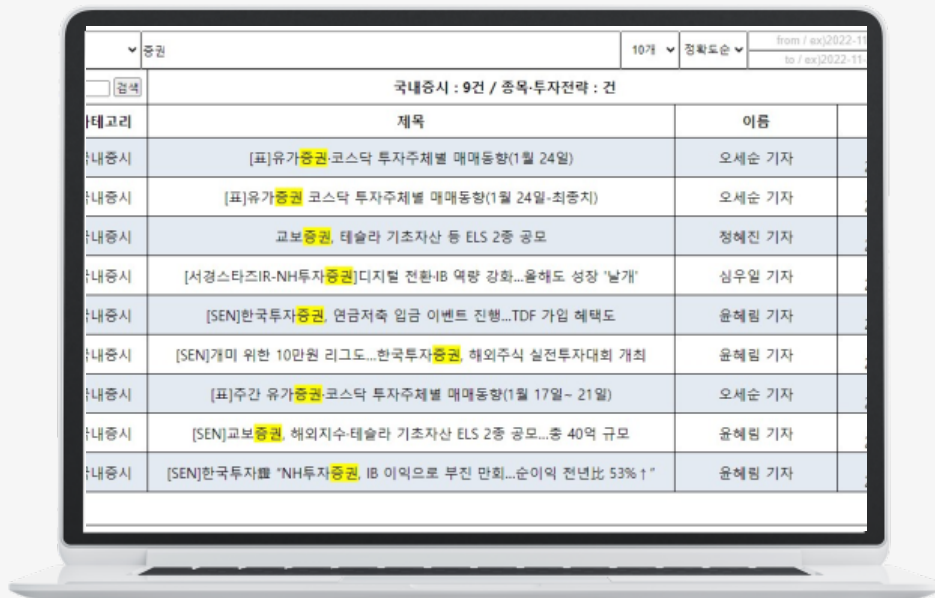
---

## Portal Search Engine

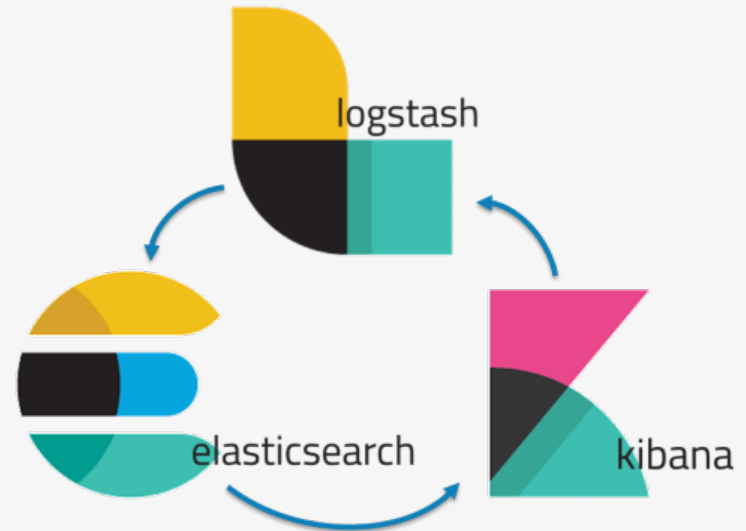
Elastic Search는 Apache Lucene 기반으로 개발한 오픈소스 실시간 분산 검색을 수행하며, 고성능의 확장가능한 Full-Text 검색 및 분석 엔진입니다.

## Data Visualizing

Kibana는 ELK의 관리, 적재된 데이터의 모니터링 및 대시보드를 활용한 데이터 시각화 기능을 제공하고 있습니다.



중권	10개	정확도순	from / ex(2022-11-11) to / ex(2022-11-11)
검색	국내중시 : 9건 / 중권·투자전략 : 건		
태고리	제목	이름	
국내중시	[표]유가중권 코스닥 투자주제별 매매동향(1월 24일)	오세순 기자	
국내중시	[표]유가중권 코스닥 투자주제별 매매동향(1월 24일-최종지)	오세순 기자	
국내중시	교보중권, 테슬라 기초자산 등 ELS 2종 공모	정혜진 기자	
국내중시	[서경스타즈IR-NH투자중권] 디지털 전환·IB 역량 강화...올해도 성장 '남개'	심우일 기자	
국내중시	[SEN]한국투자중권 연금저축 임금 이벤트 진행...TDF 가입 혜택도	윤혜림 기자	
국내중시	[SEN]개미 위한 10만원 리그드...한국투자중권 해외주식 실전투자대회 개최	윤혜림 기자	
국내중시	[표]주간 유가중권 코스닥 투자주제별 매매동향(1월 17일 - 21일)	오세순 기자	
국내중시	[SEN]교보중권, 해외지수·테슬라 기초자산 ELS 2종 공모...총 40억 규모	윤혜림 기자	
국내중시	[SEN]한국투자중권 "NH투자중권 IB 이익으로 부진 만회...순이익 전년比 53% ↑"	윤혜림 기자	





## Front + Back-End 주효명

### 개발분야

Log File 수집& 변형  
ElasticSearch 검색 & 분석  
Kibana 시각화 & 매니지

### 기여도

기획  100  
개발  100

02

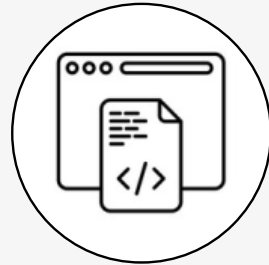
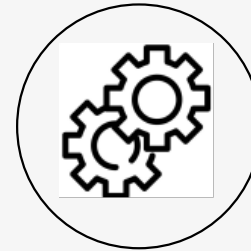
---

# SERVICE FLOW

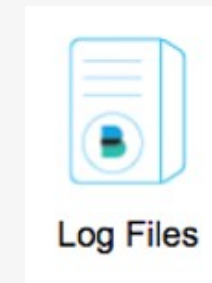
및 개발도구

# 개발도구

language / tool / database

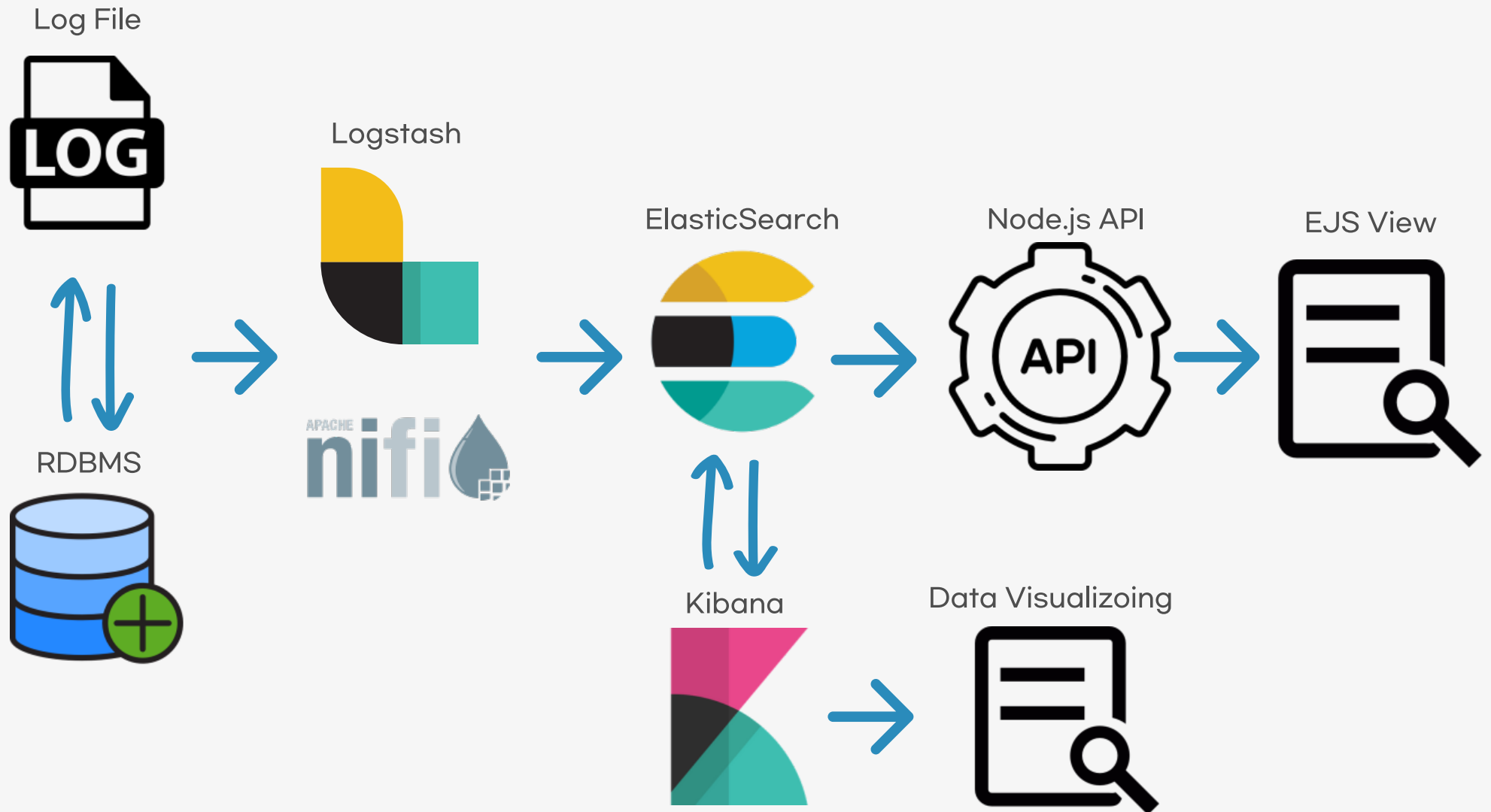


**EJS**



# Service Flow

데이터 축적 / API 구현 + 시각화 / 화면 출력



# PORTAL SEARCH ENGINE

검색 엔진

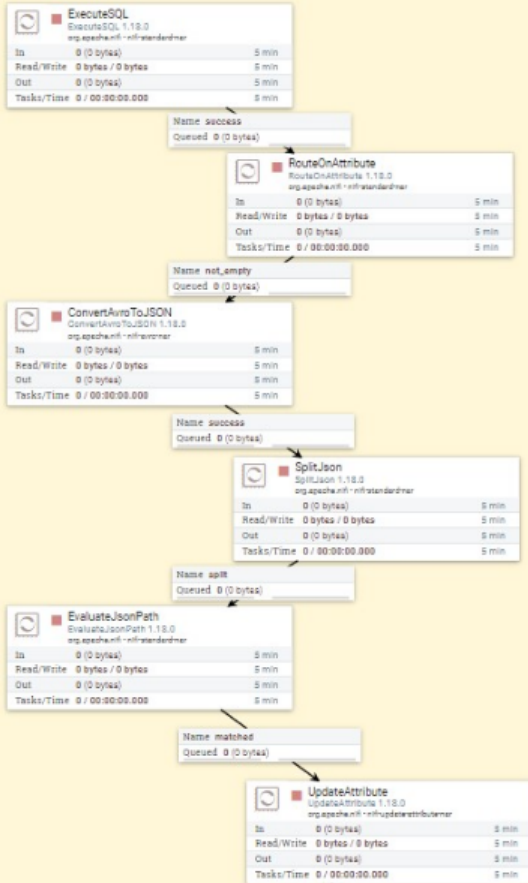
03

---

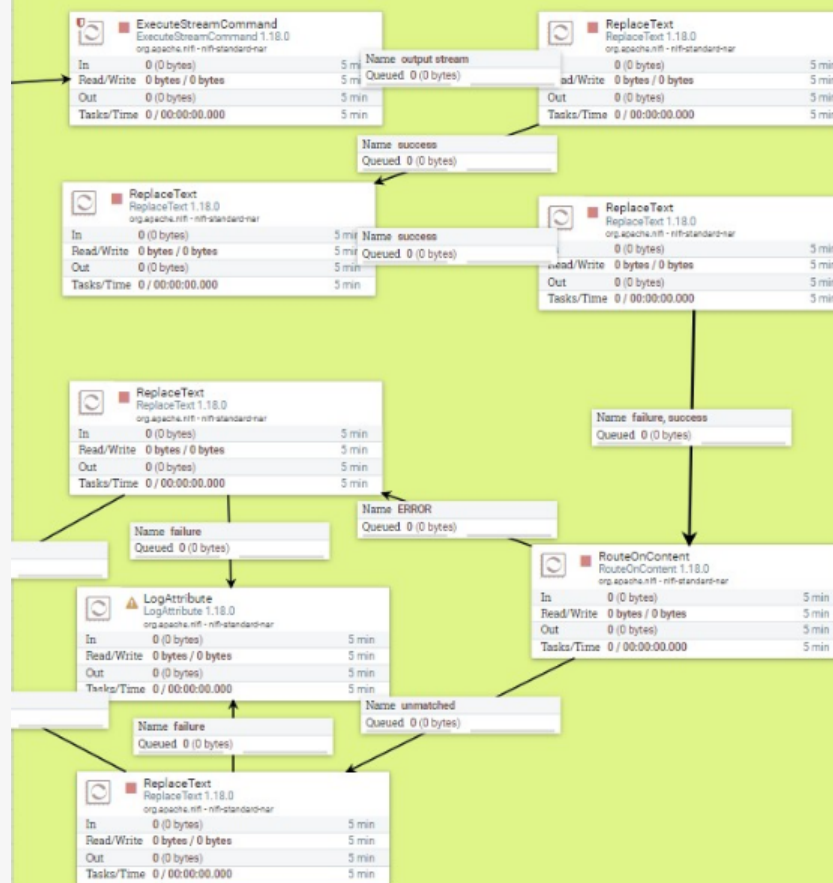
# Service Flow

Nifi -> Elastic Search 데이터 적재

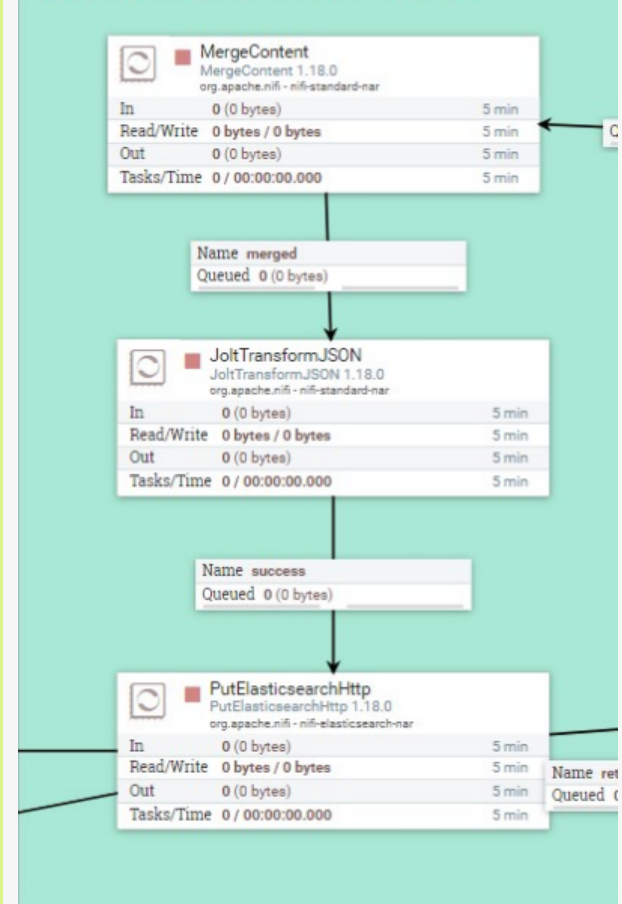
## RDBMS 데이터 추출 및 가공



## ERR 및 빈 데이터 체크



## 데이터 병합 및 Elastic Search 적재



Nifi를 활용하여 SQL데이터 추출 후 데이터를 가공하여 Elastic Search에 적재

# Service Flow

Elastic 데이터 가공

```
PUT _index_template/firma1
{
  "index_patterns": [
    "firma*"
  ],
  "priority": 1,
  "template": {
    "settings": {
      "index": {
        "max_ngram_diff": 20,
        "number_of_shards": "3",
        "number_of_replicas": "0",
        "refresh_interval": "3s"
      },
      "analysis": {
        "analyzer": {
          "firma_analyzer": {
            "type": "custom",
            "tokenizer": "my_nori_tokenizer",
            "filter": [
              "lowercase",
              "snowball",
              "unique",
              "nori_readingform",
              "my_ngram"
            ]
          },
          "email_analyzer": {
            "tokenizer": "email_tokenizer"
          },
          "firma_eng_analyzer": {
            "tokenizer": "standard",
            "filter": [
              "lowercase",
              "stop",
              "my_edge_ngram"
            ]
          }
        },
        "tokenizer": {
          "my_nori_tokenizer": {
            "type": "nori_tokenizer",
            "decompound_mode": "mixed"
          },
          "email_tokenizer": {
            "type": "uax_url_email",
            "max_token_length": 5
          }
        },
        "filter": {
          "my_ngram": {
            "type": "ngram",
            "min_gram": 2,
            "max_gram": 5,
            "token_chars": [
              "letter"
            ]
          },
          "my_edge_ngram": {
            "type": "edge_ngram",
            "min_gram": 1,
            "max_gram": 20
          }
        }
      }
    }
  }
}
```

```
"mappings": {
  "properties": {
    "BIZNO": {
      "type": "long"
    },
    "CHACNT_DEAL_BANK_CD": {
      "type": "long"
    },
    "CHACNT_DEAL_BANK_NM": {
      "type": "text",
      "term_vector": "with_positions_offsets",
      "fields": {
        "keyword": {
          "type": "keyword",
          "ignore_above": 256
        },
        "korean": {
          "type": "text",
          "analyzer": "kobrick",
          "search_analyzer": "kobrick_search"
        },
        "nori": {
          "type": "text",
          "analyzer": "firma_analyzer"
        },
        "eng": {
          "type": "text",
          "analyzer": "firma_eng_analyzer",
          "search_analyzer": "standard"
        }
      }
    },
    "CMP_ENG_NM": {
      "type": "text",
      "term_vector": "with_positions_offsets",
      "fields": {
        "keyword": {
          "type": "keyword",
          "ignore_above": 256
        },
        "nori": {
          "type": "text",
          "analyzer": "firma_analyzer"
        },
        "eng": {
          "type": "text",
          "analyzer": "firma_eng_analyzer",
          "search_analyzer": "standard"
        }
      }
    },
    "CMP_GHRL_NM": {
      "type": "text",
      "term_vector": "with_positions_offsets",
      "fields": {
        "keyword": {
          "type": "keyword",
          "ignore_above": 256
        },
        "nori": {
          "type": "text",
          "analyzer": "firma_analyzer"
        }
      }
    }
  }
}
```

```
GET firma1/_search
{
  "query": {
    "bool": {
      "should": [
        {
          "multi_match": {
            "query": "jswoo@dongsungprinting.com",
            "fields": [
              "CMP_NM.keyword",
              "CMP_NM.nori",
              "CMP_ENG_NM.nori",
              "CMP_ENG_NM.eng",
              "GRP_NM",
              "MNBS_LOANN_INSTT_NM.nori",
              "MNBS_LOANN_INSTT_NM.korean",
              "CHACNT_DEAL_BANK_NM.nori",
              "CHACNT_DEAL_BANK_NM.korean",
              "HMPG_ADDR",
              "EMAIL_ADDR",
              "LOCPLC_BASE_ADDR"
            ]
          }
        }
      ]
    }
  },
  "highlight": {
    "pre_tags": [
      "<strong>"
    ],
    "post_tags": [
      "</strong>"
    ],
    "fragment_size": 10,
    "fields": {
      "CMP_NM.keyword": {},
      "CMP_NM.nori": {},
      "CMP_ENG_NM.nori": {},
      "CMP_ENG_NM.eng": {},
      "GRP_NM": {},
      "MNBS_LOANN_INSTT_NM.nori": {
        "MNBS_LOANN_INSTT_NM.korean": {}
      },
      "CHACNT_DEAL_BANK_NM.nori": {
        "CHACNT_DEAL_BANK_NM.korean": {}
      },
      "HMPG_ADDR": {},
      "EMAIL_ADDR": {}
    }
  }
}
```

데이터 형식에 맞게 Re-Mapping 및 Analyzer 적용

하이라이트 검색 Query 작성



# Service Flow

Elastick 연동 API 작성

```
const express = require("express");
const APPROOT = require("app-root-path");
const config = require(`${APPROOT}/config/config`);
const proxyServer = require(`${APPROOT}/routes/index`);
const ejs = require("ejs");
const bodyParser = require("body-parser");

const app = express();
app.use(express.json());

//post
// app.unsubscribe(express.urlencoded({ extended: true }));
app.use(bodyParser.urlencoded({ extended: false }));
app.use(bodyParser.json());

proxyServer(app);

// Server Port
const port = config.API_SERVICE_PORT;
app.set(port);

// ejs
app.set("view engine", "ejs");
app.use(express.static("views"));
// app.use(express.static(__dirname + '/'));

app.get("/", function (req, res) {
  res.render("idx_first");
});

// Create Http Server
// const server = http.createServer(app);
// server.listen(port);

const handleListening = () => {
  console.log("Server listening on port http://localhost:${port}");
  app.listen(port, handleListening);
};

module.exports = app;
```

```
const express = require("express");
const router = express.Router();
const searchController = require("../search/index");

router.post("/search_post", (req, res) => {
  searchController.search(req, res);
});

router.post("/search_post_sp", (req, res) => {
  searchController.search(req, res);
});

module.exports = router;
```

```
const search = (data) => {
  let field = data.category;
  let search = data.search;
  let o_search = data.o_search;
  let o_category = data.o_category;
  let s_search = data.s_search;
  let scope = data.scope;
  if (!field && !search) {
    if (field === 'title') {
      field = field + '.korean';
    }
    const result = {
      index: 'practice1',
      body: {
        from: 0,
        size: scope,
        _source: {
          excludes: 'content',
        },
        aggs: {
          category: {
            terms: {
              field: 'category.keyword',
            },
          },
        },
        query: {
          bool: {
            must: {
              match: {
                [field]: search,
              },
            },
          },
        },
        highlight: {
          number_of_fragments: 1,
          fragment_size: 30,
          pre_tags: ['<mark>'],
          post_tags: ['</mark>'],
          fields: {
            [field]: {},
          },
        },
      },
    };
  }
};
```

Elastick에 연동되는 API작성

# Service Flow

화면 구현

title ▼		투자	10개 ▼	정확도순 ▼	from) to)2	제출
<div>상세검색</div> <div>검색</div>		국내증시 : 14건 / 종목·투자전략 : 건				
번호	카테고리	제목	이름	날짜		
1	국내증시	[SEN]DB금융투자 "해외선물옵션 투자하세요"...제3회 해선왕 선발대회 개최	윤혜림 기자	2022-01-25T01:00:22.000Z		
2	국내증시	[SEN]한화투자증권 "삼성바이오로직스, 장기성장 청사진 제시...투자 매력도 ↑"	윤혜림 기자	2022-01-24T23:31:20.000Z		
3	국내증시	[SEN]개미 위한 10만원 리그도...한국투자증권, 해외주식 실전투자대회 개최	윤혜림 기자	2022-01-24T00:57:34.000Z		
4	국내증시	[SEN]한국투자증권 "NH투자증권, IB 이익으로 부진 만회...순이익 전년比 53% ↑"	윤혜림 기자	2022-01-23T23:45:30.000Z		
5	국내증시	메리츠증권, 미·중·일·홍콩 투자 가능한 CFD 출시	정혜진 기자	2022-01-24T09:07:42.000Z		
6	국내증시	[표]유가증권·코스닥 투자주체별 매매동향(1월 24일)	오세순 기자	2022-01-24T06:41:40.000Z		
7	국내증시	[표]유가증권 코스닥 투자주체별 매매동향(1월 24일·최종치)	오세순 기자	2022-01-24T09:20:59.000Z		
8	국내증시	[SEN]하나금융투자, 불확실한 시장 대응...목표전환형 '모으기협합랩' 출시	윤혜림 기자	2022-01-24T00:25:35.000Z		
9	국내증시	[서경스타즈IR·NH투자증권]디지털 전환·IB 역량 강화...올해도 성장 '날개'	심우일 기자	2022-01-23T09:18:42.000Z		
10	국내증시	[SEN]한국투자증권, 연금저축 입금 이벤트 진행...TDF 가입 혜택도	윤혜림 기자	2022-01-25T02:04:14.000Z		
11	국내증시	[SEN투자전략]美 증시 하락세...국내 증시 실적주 반등 예상	최민정 기자	2022-01-23T23:45:17.000Z		
12	국내증시	[표]주간 유가증권·코스닥 투자주체별 매매동향(1월 17일~ 21일)	오세순 기자	2022-01-23T07:27:46.000Z		
13	국내증시	거래소 "쪼개기 상장·스톡옵션 매각 관련 투자자 보호책 마련할 것"	정혜진 기자	2022-01-25T03:51:07.000Z		
14	국내증시	[SEN]한화투자증권 "에스엘, 평균판매가 상승·물량 회복 기대...가치 재평가 기대"	윤혜림 기자	2022-01-23T23:27:07.000Z		
14						

EJS를 활용한 검색엔진 화면 구현

04

---

# KIBANA DATA VISUALIZING

데이터 시각화

# Service Flow

Elastic 연동 API 작성

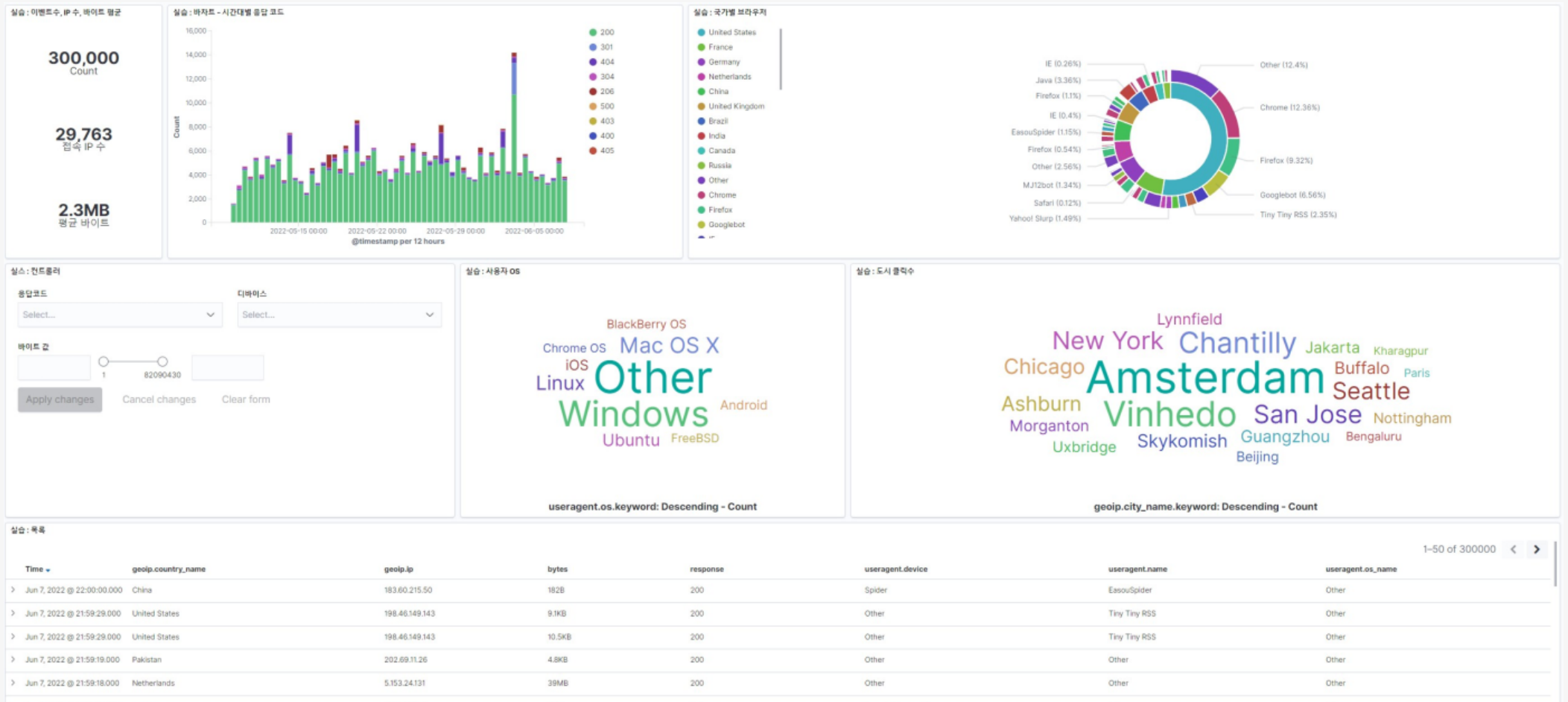
```
input {
  file {
    path => "D:/Elastic/weblog-sample.log"
    start_position => "beginning"
  }
}
filter {
  grok {
    match => { "message" => "%{COMBINEDAPACHELOG}" }
  }
  date {
    match => [ "timestamp", "dd/MMM/yyyy:HH:mm:ss Z" ]
  }
  mutate {
    convert => {
      "bytes" => "integer"
    }
  }
  mutate {
    remove_field => [ "timestamp", "host", "@version", "agent" ]
  }
}
```

```
output {
  elasticsearch {
    hosts => ["localhost:9200"]
    index => "apachelog-%{+yyyy.MM.dd}"
  }
  stdout {
    codec => "dots"
  }
}
```

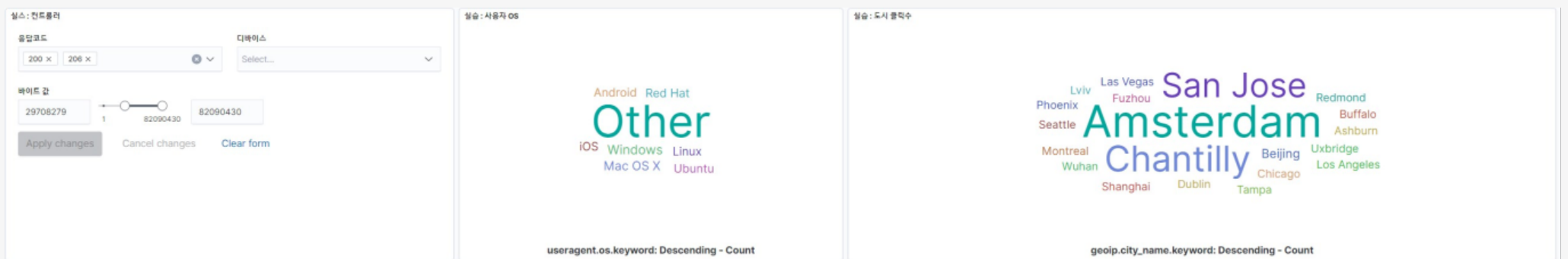
Logstash를 이용한  
데이터 Mapping & Filtering  
및  
Elastic Data 적체

# Service Flow

Kibana를 이용한 Elastic 적제 데이터 시각화



컨트롤러를 이용한 데이터 필터링





**THANK YOU**