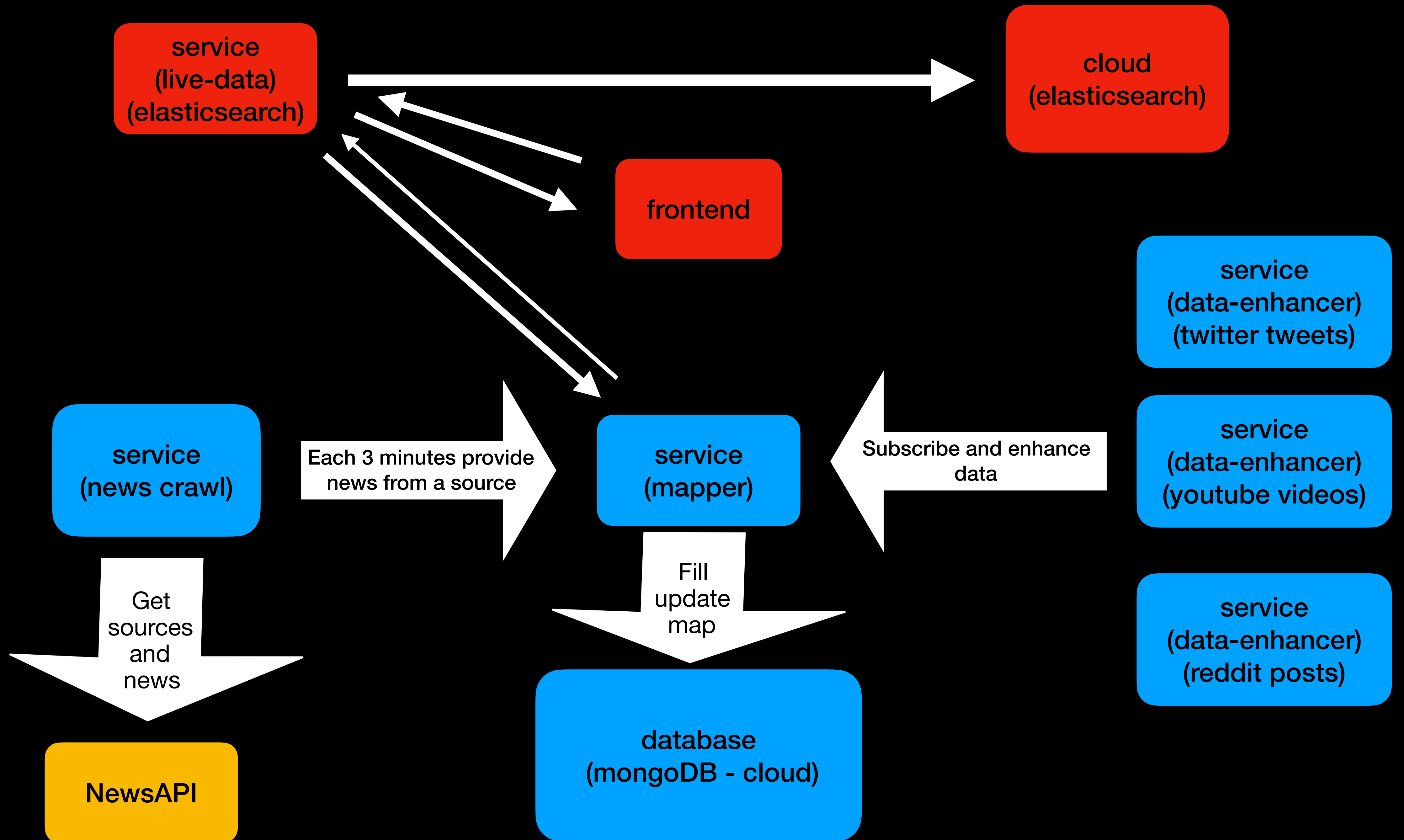


Daily News



Data Sample

Source

```
{
  "id": "abc-news",
  "name": "ABC News",
  "description": "Your trusted source for breaking news, analysis,
exclusive interviews, headlines, and videos at ABCNews.com.",
  "url": "https://abcnews.go.com",
  "category": "general",
  "language": "en",
  "country": "us"
}
```

Data Sample

News with source

```
{
  "source": {
    "id": "abc-news",
    "name": "ABC News",
    "description": "Your trusted source for breaking news, analysis, exclusive interviews, headlines, and videos at ABCNews.com.",
    "url": "https://abcnews.go.com",
    "category": "general",
    "language": "en",
    "country": "us"
  },
  "author": "COLLIN BINKLEY AP Education Writer",
  "title": "Nation's public schools need major repairs, report finds",
  "description": "",
  "url": "https://abcnews.go.com/Politics/wireStory/nations-public-schools-major-repairs-report-finds-71070921",
  "urlToImage": "null",
  "publishedAt": "2020-06-04T17:30:29Z",
  "content": "The majority of U.S. public school districts need to make major building repairs in at least half of their schools, yet many districts are unable to pay for updates or have postponed them to improve ... [+4431 chars]"
}
```

Data Sample

Enhanced News with YouTube data-enhancer service

```
{
  "source": {
    "id": "abc-news", "name": "ABC News", "description": "Your trusted source for breaking news, analysis, exclusive interviews, headlines, and videos at ABCNews.com.",
    "url": "https://abcnews.go.com", "category": "general", "language": "en", "country": "us"
  },
  "author": "COLLIN BINKLEY AP Education Writer",
  "title": "Nation's public schools need major repairs, report finds",
  "description": "",
  "url": "https://abcnews.go.com/Politics/wireStory/nations-public-schools-major-repairs-report-finds-71070921",
  "urlToImage": "null",
  "publishedAt": "2020-06-04T17:30:29Z",
  "content": "The majority of U.S. public school districts need to make major building repairs in at least half of their schools, yet many districts are unable to pay for updates or have postponed them to improve ... [+4431 chars]",
  "videos": [
    {
      "youtubeId": { "kind": "youtube#video", "videoId": "od3s3lZWbWM" },
      "publishedAt": "2019-01-23T20:59:56Z",
      "channelId": "UCn4sPeUomNGIr26bElVdDYg",
      "title": "How Black High School Students Are Hurt by Modern-Day Segregation | NowThis",
      "description": "Students at Block High School in Jonesville, Louisiana, are still facing segregation and inequality every day at school, more than 50 years after Martin Luther ...",
      "thumbnails": {
        "default": { "url": "https://i.ytimg.com/vi/od3s3lZWbWM/default.jpg", "width": 120, "height": 90 },
        "medium": { "url": "https://i.ytimg.com/vi/od3s3lZWbWM/mqdefault.jpg", "width": 320, "height": 180 },
        "high": { "url": "https://i.ytimg.com/vi/od3s3lZWbWM/hqdefault.jpg", "width": 480, "height": 360 }
      },
      "channelTitle": "NowThis News",
      "publishTime": "2019-01-23T20:59:56Z"
    }
  ]
}
```

Data Crawl

NewsAPI

- 500 API request limit per day
- 128 news source
- For each 3 min period;
 - Headlines for a source; max 10 news; avg. ~ 4 news
- For 24 hour period; ~4 different request for a source
 - $128 * 4 \text{ (request per day for a source)} * 4 \text{ (avg news)} \Rightarrow 2048 \text{ news per day}$

Data Crawl

- Youtube videos
- Reddit Posts
- Tweets

Elasticsearch

Used Features

- Bulk API
- Highlighting
- CRU(D) Operations for an index
- Fields
 - Dates, full-text, videos [, tweets, images], source, language