

Data Engineering

데이터엔지니어링

[Lambda Expression과 Stream]



Lambda Expression과 Parallel Stream의 필요성

학습내용

- 1 Lambda Expression의 필요성 이해
- 2 Parallel Stream의 필요성 이해

학습목표

- 실 세계 데이터 분석 사례를 통해 Lambda Expression의 필요성을 설명할 수 있다.
- 실 세계 데이터 분석 사례를 통해 Parallel Stream의 필요성을 설명할 수 있다.

Lambda Expression의 필요성

```
if (!occurrence.containsKey(from)) {  
    occurrence.put(from, 1);  
} else {  
    occurrence.put(from, occurrence.get(from) + 1);  
}
```

compute Method 활용

```
occurrence.compute(from, new BiFunction<Integer, Integer, Integer>() {  
    @Override  
    public Integer apply(Integer key, Integer value) {  
        if (value == null)  
            return 1;  
        else {  
            return value+1;  
        }  
    }  
});
```

간결하게 구현

Lambda Expression의 필요성

```
occurrence.compute(from, new BiFunction<Integer, Integer, Integer>() {  
    @Override  
    public Integer apply(Integer key, Integer value) {  
        if (value == null)  
            return 1;  
        else {  
            return value+1;  
        }  
    }  
});
```

Lambda Expression으로 구현한
compute Method 활용

```
occurrence.compute(from, (key,value) -> {  
    if (value == null)  
        return 1;  
    else {  
        return value + 1;  
    }  
});
```

함축적이고 간결한
논리 구현 가능

Lambda Expression

Java 언어에서의 함수형 프로그래밍 지원
(버전 8 이후 지원)

Method 정의가 한 개 있는 Interface
(Functional Interface)의 구현을 간결하게 표현

매개변수

()



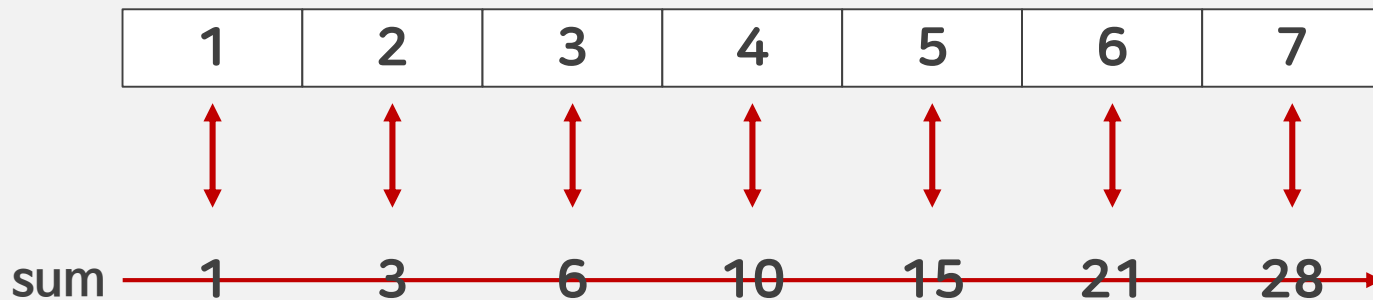
{ Functional Interface 구현 }

Parallel Stream의 필요성

List<Integer>의 모든 수의 합

멀티 스레드를 통한 성능 향상이 가능할까?

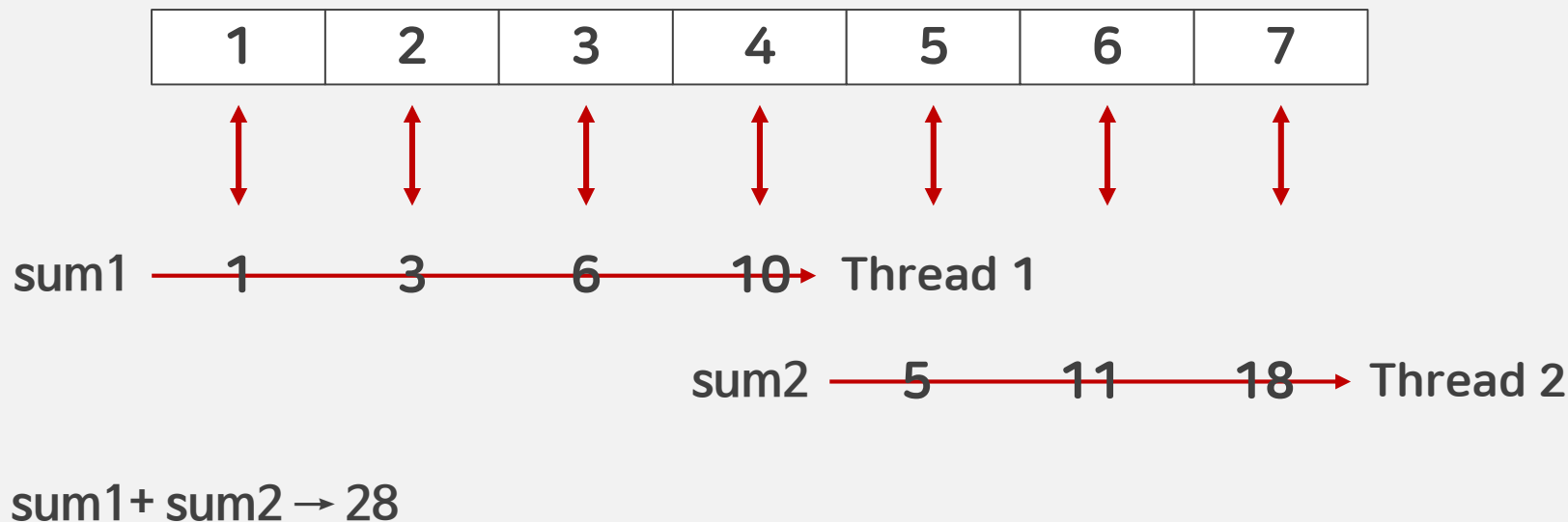
Single Thread



Parallel Stream의 필요성

List<Integer>의 모든 수의 합

멀티 스레드를 통한 성능 향상이 가능할까?



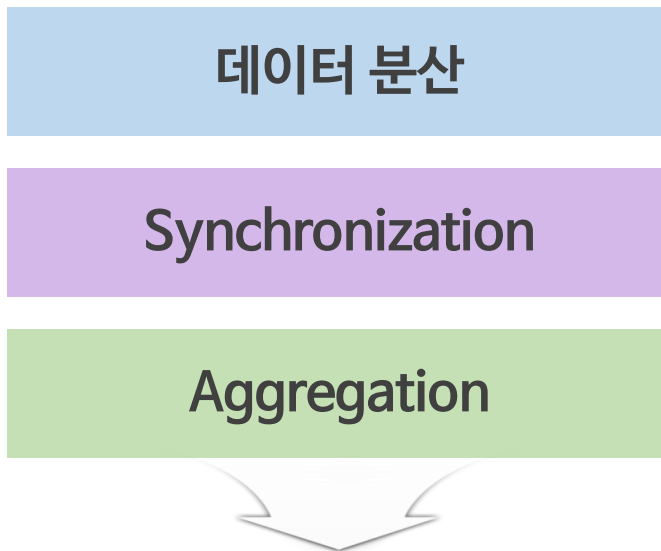
Parallel Stream의 필요성

데이터 분산

Synchronization

Aggregation

직접 처리



Parallel Stream의 필요성

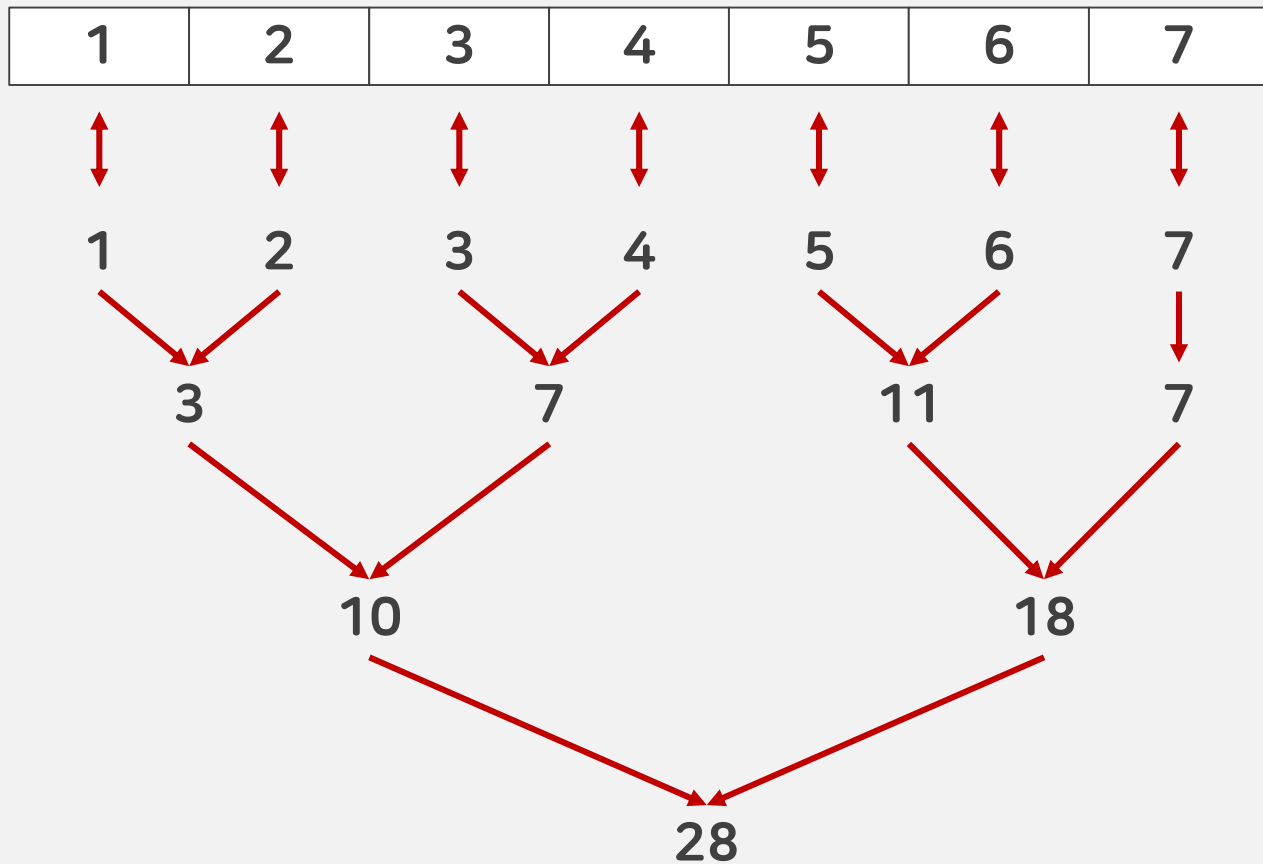
자바 버전 8 이후 지원하는
Java Collection의 Parallel Stream 활용

Parallel Stream 사용

```
Stream.of(1,2,3,4,5,6,7).parallel()  
.mapToInt(Integer::intValue).sum();
```

List<Integer>의 모든 수의 합

멀티 스레드를 통한 성능 향상이 가능할까?



Parallel Stream의 필요성



Lambda Expression 사용 없이는
코드가 매우 커짐

Lambda Expression 사용은 필수임

Remind

Lambda Expression 필요성 이해

Parallel Stream 필요성 이해