

Hadoop MapReduce 기틀 재설계

김동원(발표자), 박성우
포항공대 프로그래밍 언어 연구실
2014-01-14

빅데이터 처리 프로그램



빅데이터, 경영을 바꾸다

한유근 · 채승범 지음

- 25 million, number of tweets sent every second
- 170 companies, top ranked by worldwide revenue
- 270 items, products customers purchase per second
- 24 products, data per day generated by Google
- 20 billion, video uploads to YouTube
- 1.3 exabytes, data sent and received by mobile internet user
- 700 billion, total minutes spent on Facebook each month

데이터는 답을 알고 있다!
빅데이터 시대의 새로운 기회를 찾아서

삼성경제연구소



빅데이터 Big Data 세상을 바꾸다

박사 연구 주제 :

하둡 기틀(architecture) 재설계를 통한 **처리** 시간 단축

입력



처리



결과



• 기대 효과

- 경영을 **더 빨리** 바꿀 수 있다
- 미래를 **더 빨리** 예측할 수 있다
- 세상을 **더 빨리** 바꿀 수 있다



센터장님의 예상 질문



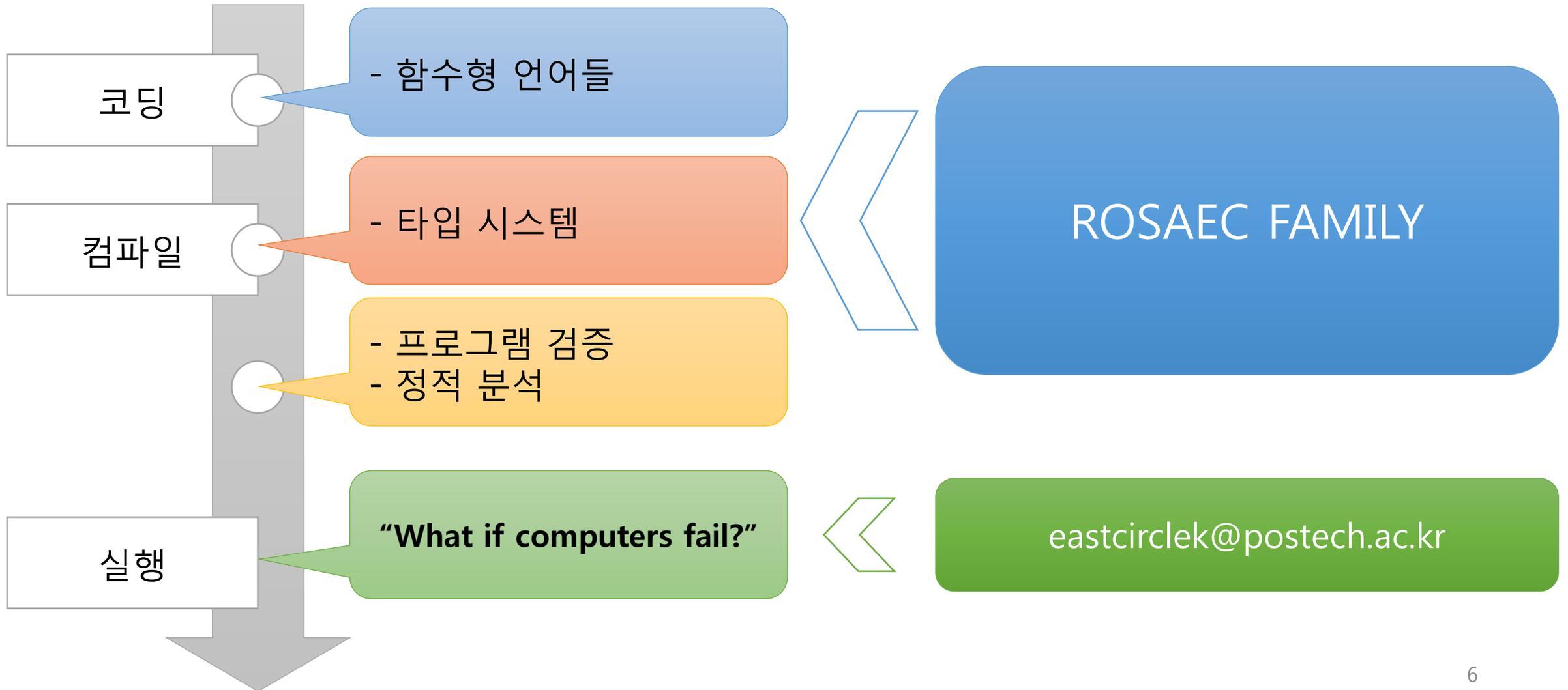
도대체 **빅데이터 처리 프로그램**하고
소프트웨어 무결점연구센터와
무슨 관계가 있는 거죠?

질문에 대한 고찰

- 소프트웨어 **무결점** 연구 센터
 - Research on **Software Analysis** for **Error-free Computing**
- 빅데이터 처리 프로그램에 요구되는 것 (하둡 & 새기틀 하둡 모두)
 - Software Analysis? No!
 - **실행도중 결함**을 감내할 수 있어야 함 (Runtime fault-tolerance)



프로그램의 생애 **결점** 방지를 위한 노력들



센터장님의 예상 질문에 대한 저의 대답



도대체 **빅데이터 처리 프로그램**하고
소프트웨어 **무결점**연구센터와
무슨 관계가 있는 거죠?

실행도중 결함감내도
소프트웨어의 **무결점**을 추구하는
기술이기 때문입니다!

Software analysis를 보완할 수 있는 기술이
라고 볼 수도 있습니다!



목차

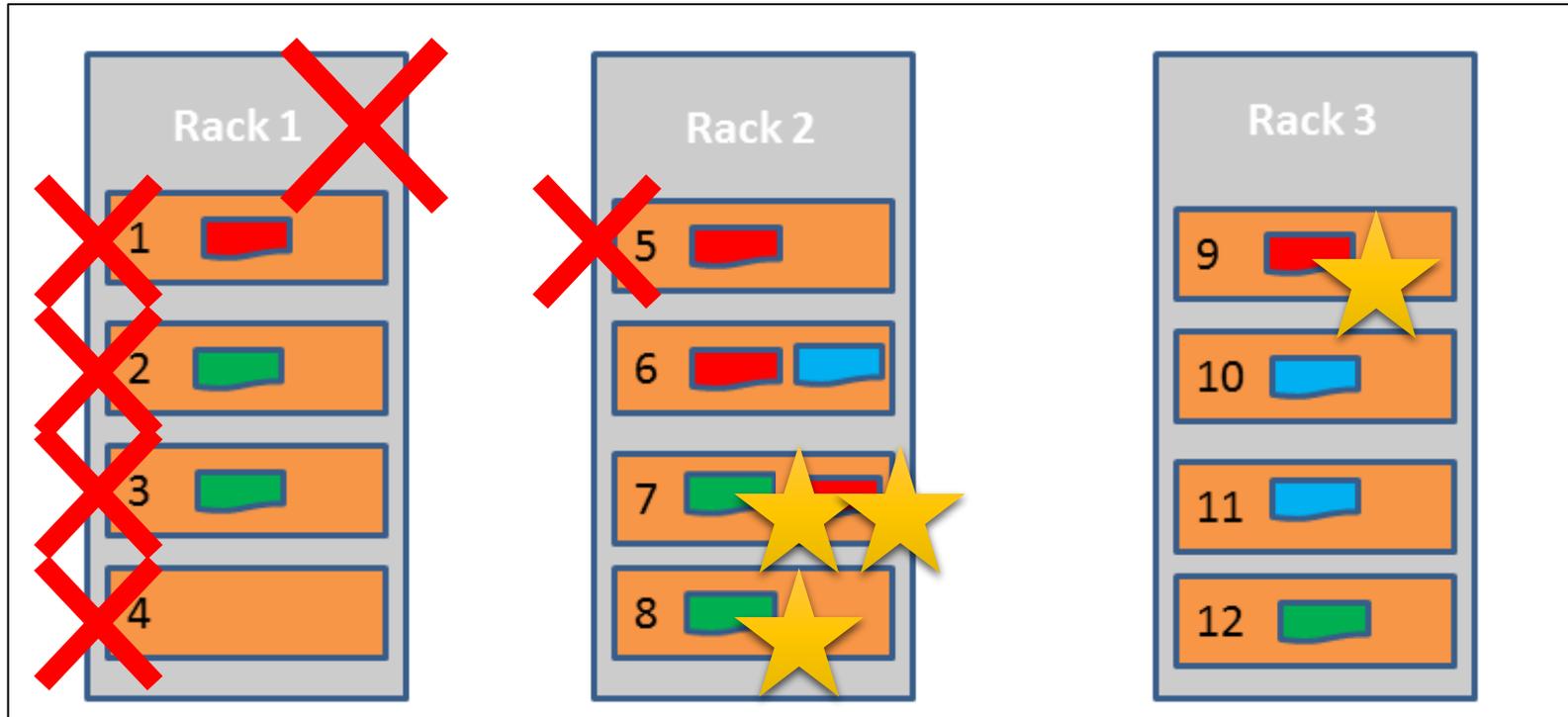
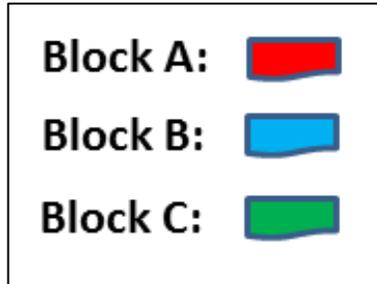
- 빅데이터 처리 프로그램 연구 @ ROSAEC
 - 기여 방안 : 실행도중 결함감내 연구
- **Hadoop**
 - 실행도중 결함감내와 성능
- Hadoop 기틀 재설계

하둡의 실행도중 결함감내를 위한 노력

(1) 입력 중복 저장

(2) 일꾼 프로세스들

(3) 데이터 도중 저장



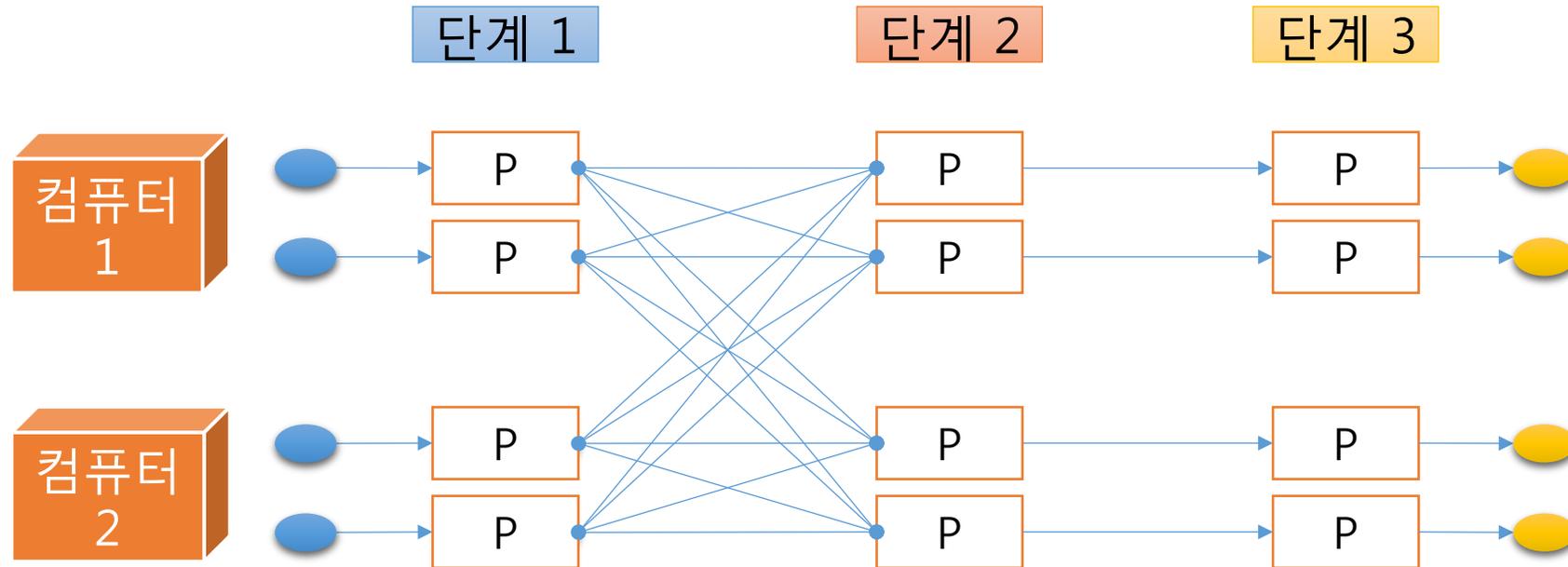
하둡의 실행도중 결함감내를 위한 노력

(1) 입력 중복 저장

(2) 일꾼 프로세스들

(3) 데이터 도중 저장

- 일꾼들이 프로세스로 실행 됨



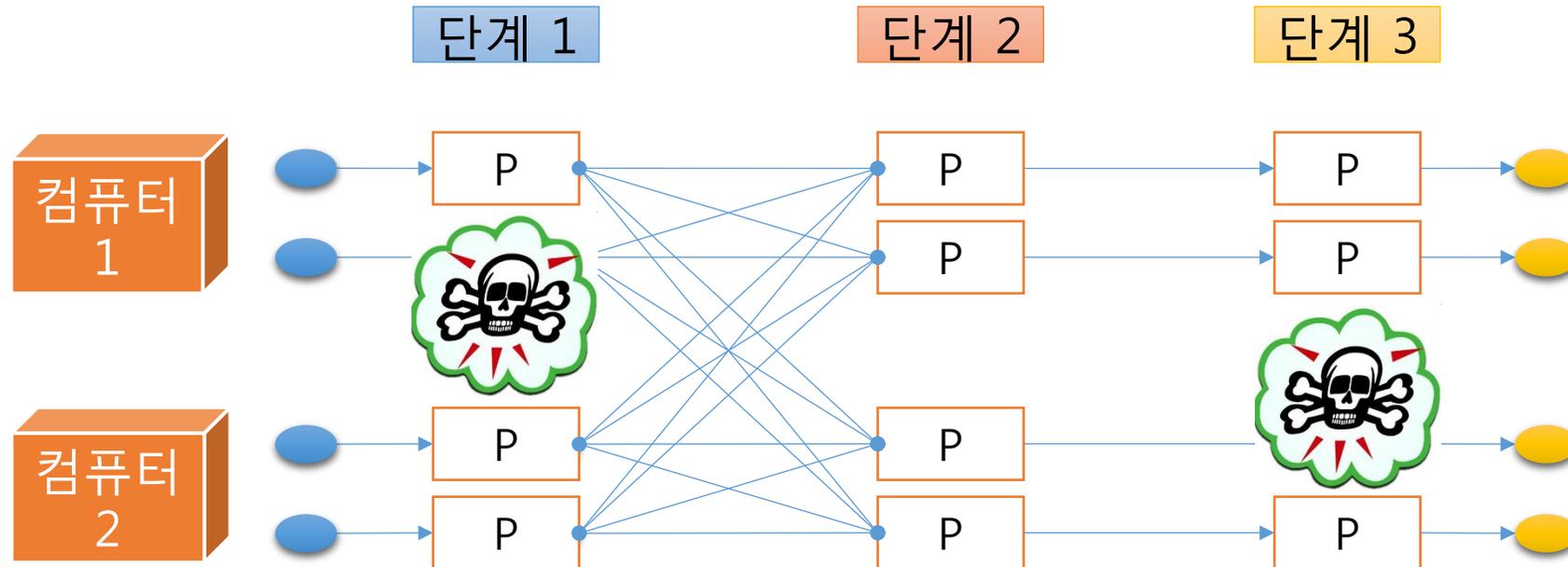
하둡의 실행도중 결함감내를 위한 노력

(1) 입력 중복 저장

(2) 일꾼 프로세스들

(3) 데이터 도중 저장

- 한 일꾼 프로세스의 죽음은 다른 일꾼들에게 피해를 안 끼침



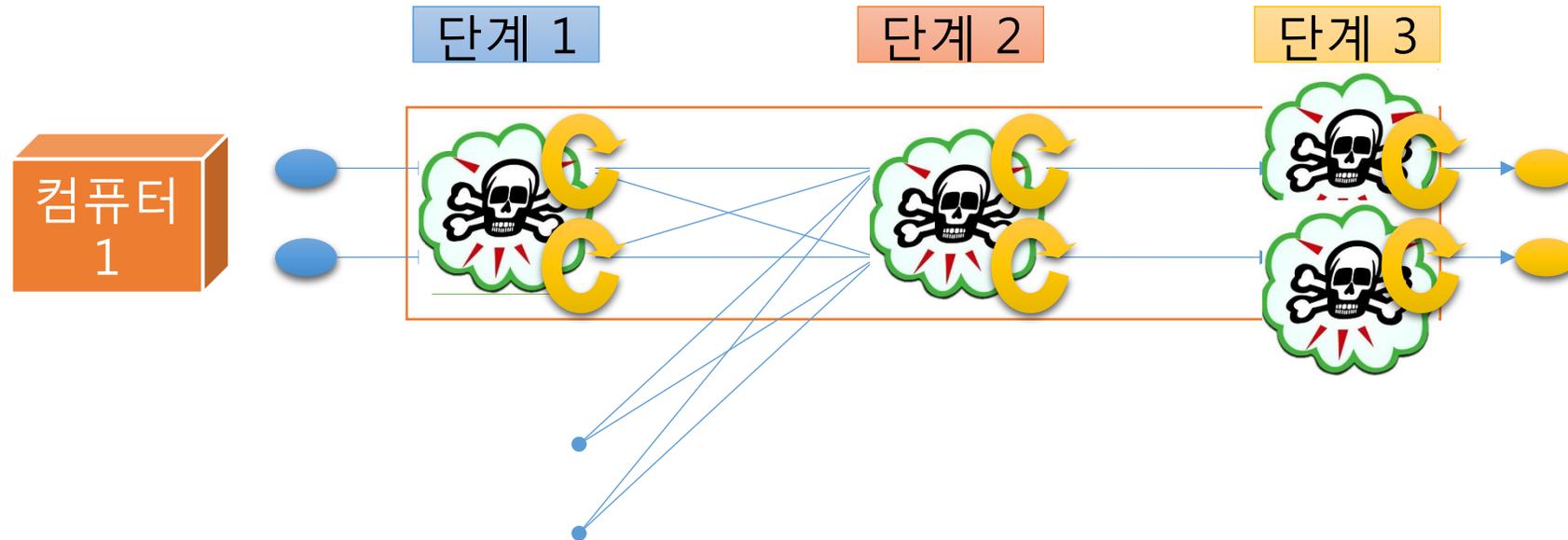
하둡의 실행도중 결함감내를 위한 노력

(1) 입력 중복 저장

(2) 일꾼 프로세스들

(3) 데이터 도중 저장

- 만약 일꾼들이 한 Process내 Thread들로 실행된다면?



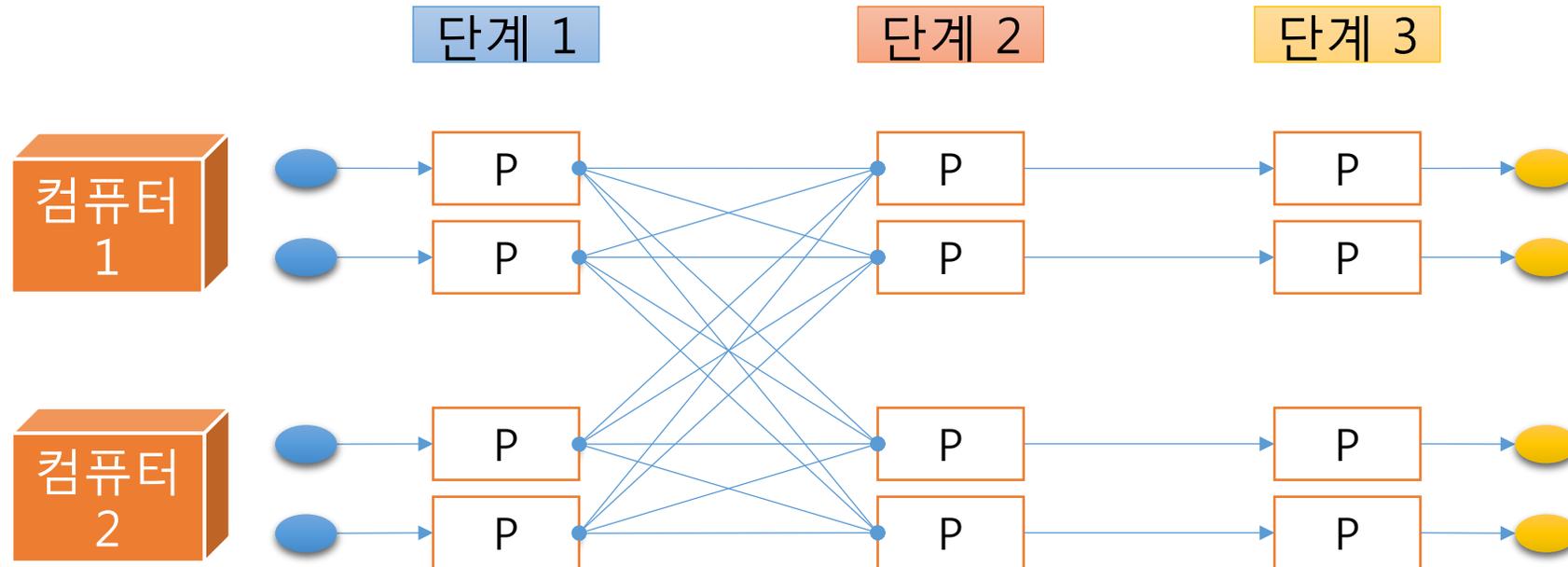
하둡의 실행도중 결함감내를 위한 노력

(1) 입력 중복 저장

(2) 일꾼 프로세스들

(3) 데이터 도중 저장

- 일꾼들이 데이터를 어떻게 주고 받을까요? 공유 메모리? 파이프? 소켓 통신?



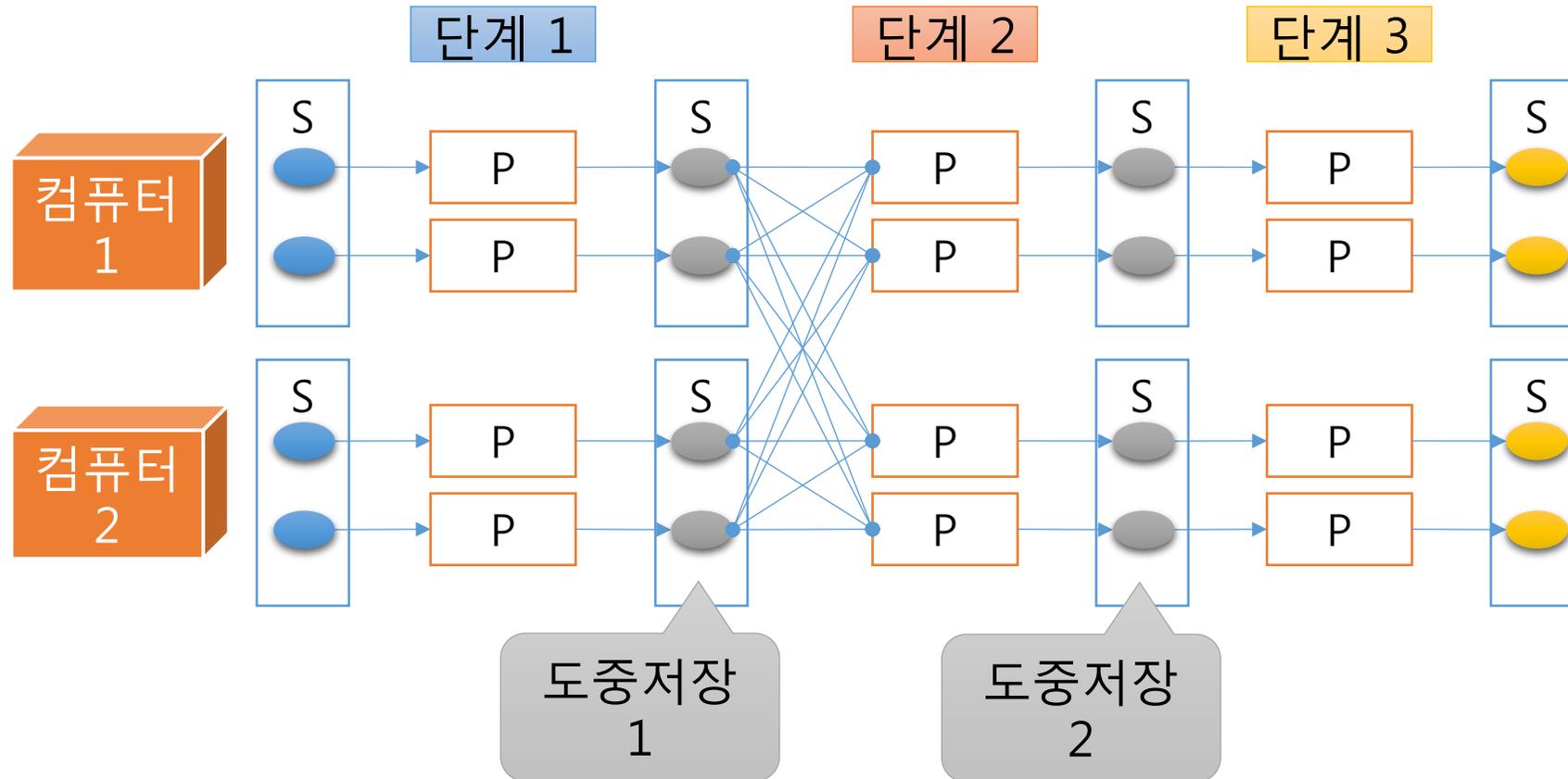
하둡의 실행도중 결함감내를 위한 노력

(1) 입력 중복 저장

(2) 일꾼 프로세스들

(3) 데이터 도중 저장

- 직전 단계 일꾼이 'S'에 저장한 결과를 읽어서 처리 ('S' : HDD나 SSD 같은 저장 장치)



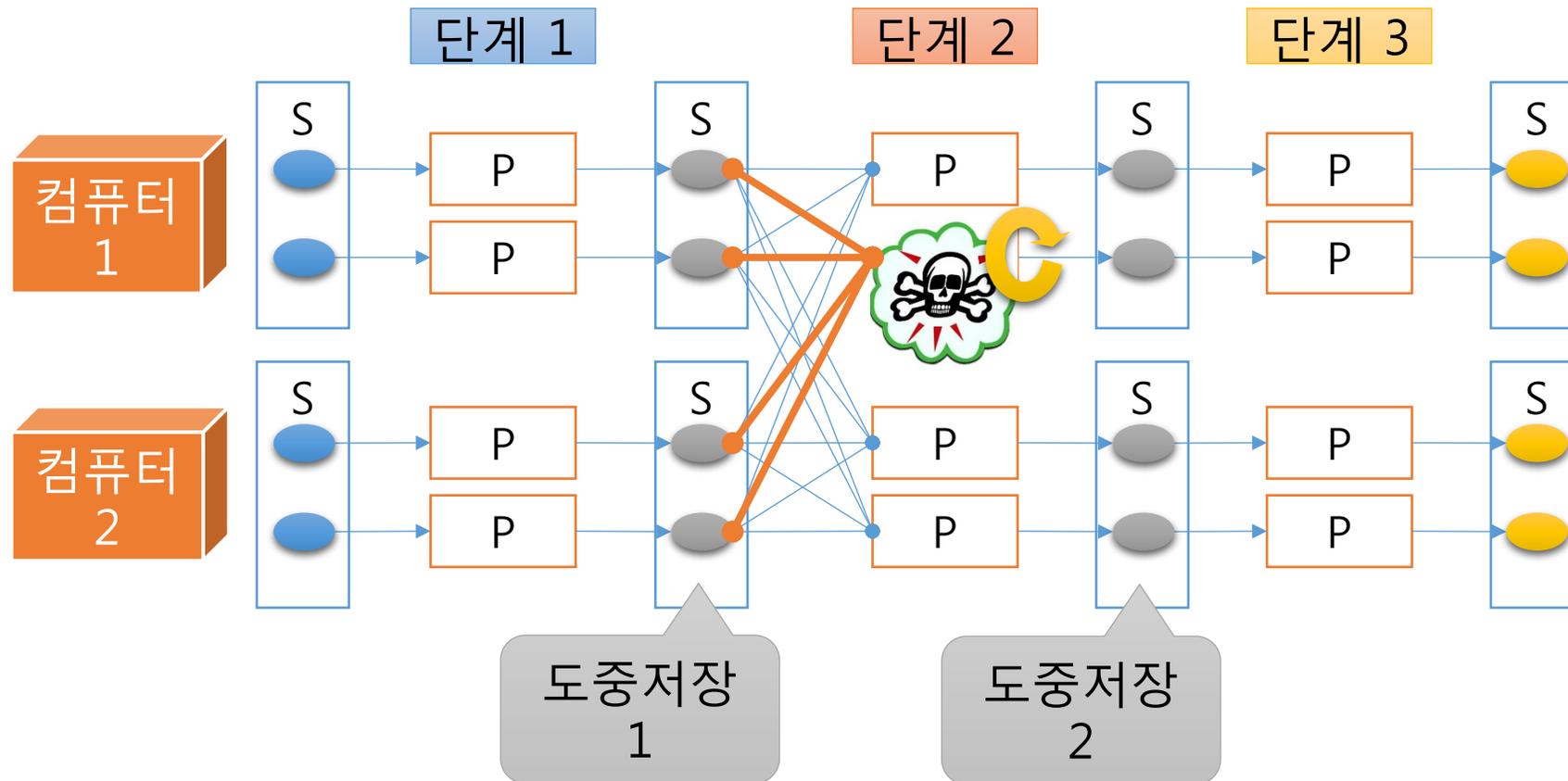
하둡의 실행도중 결함감내를 위한 노력

(1) 입력 중복 저장

(2) 일꾼 프로세스들

(3) 데이터 도중 저장

- 중간에 일꾼 프로세스가 죽으면?



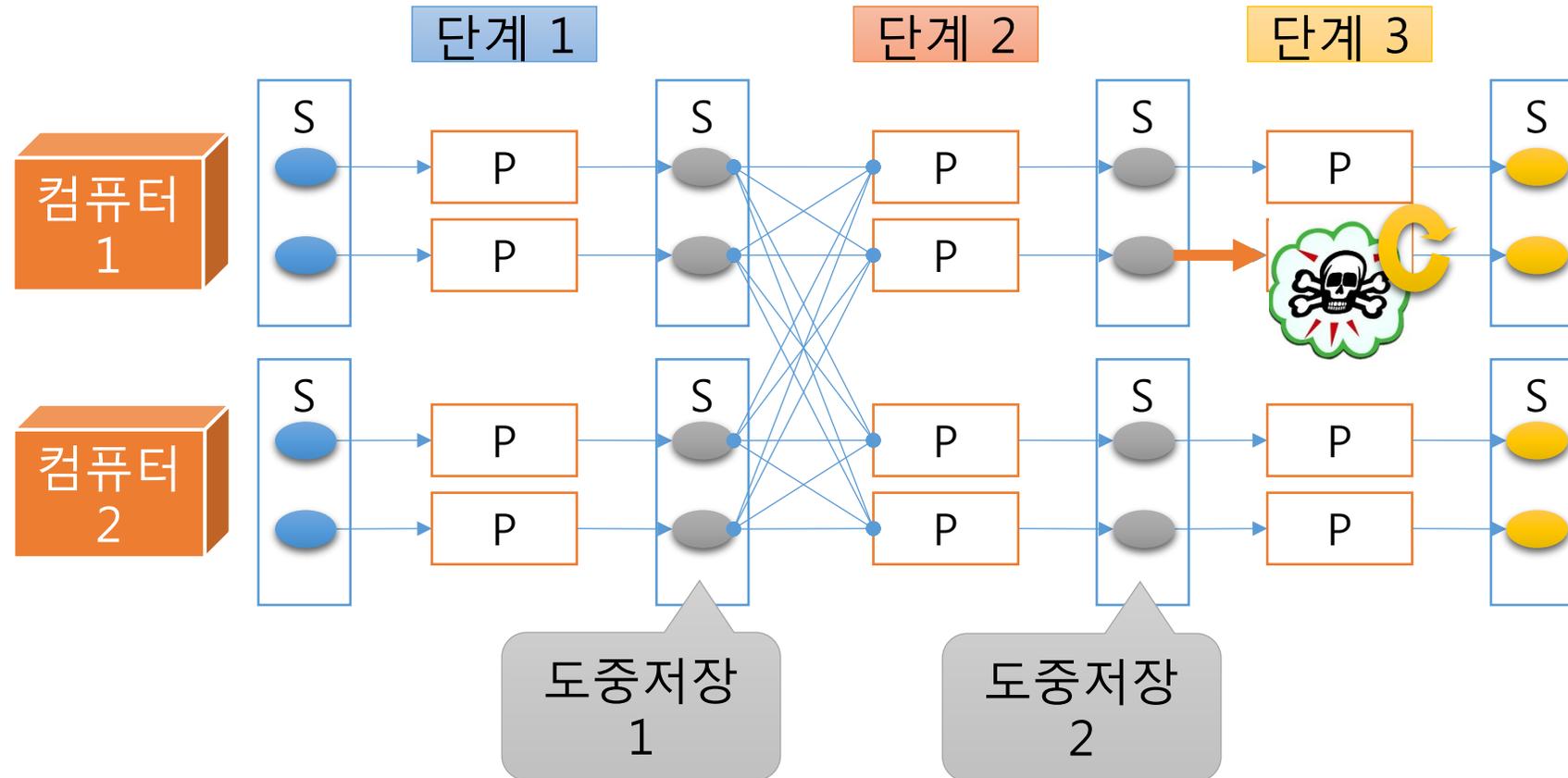
하둡의 **실행도중 결함감내**를 위한 노력

(1) 입력 중복 저장

(2) 일꾼 프로세스들

(3) 데이터 도중 저장

- 중간에 일꾼 프로세스가 죽으면?



하둡의 실행도중 결함감내와 성능

결함감내 잡았다!

- 입력 중복 저장
- 일꾼 프로세스들
- 데이터 도중 저장



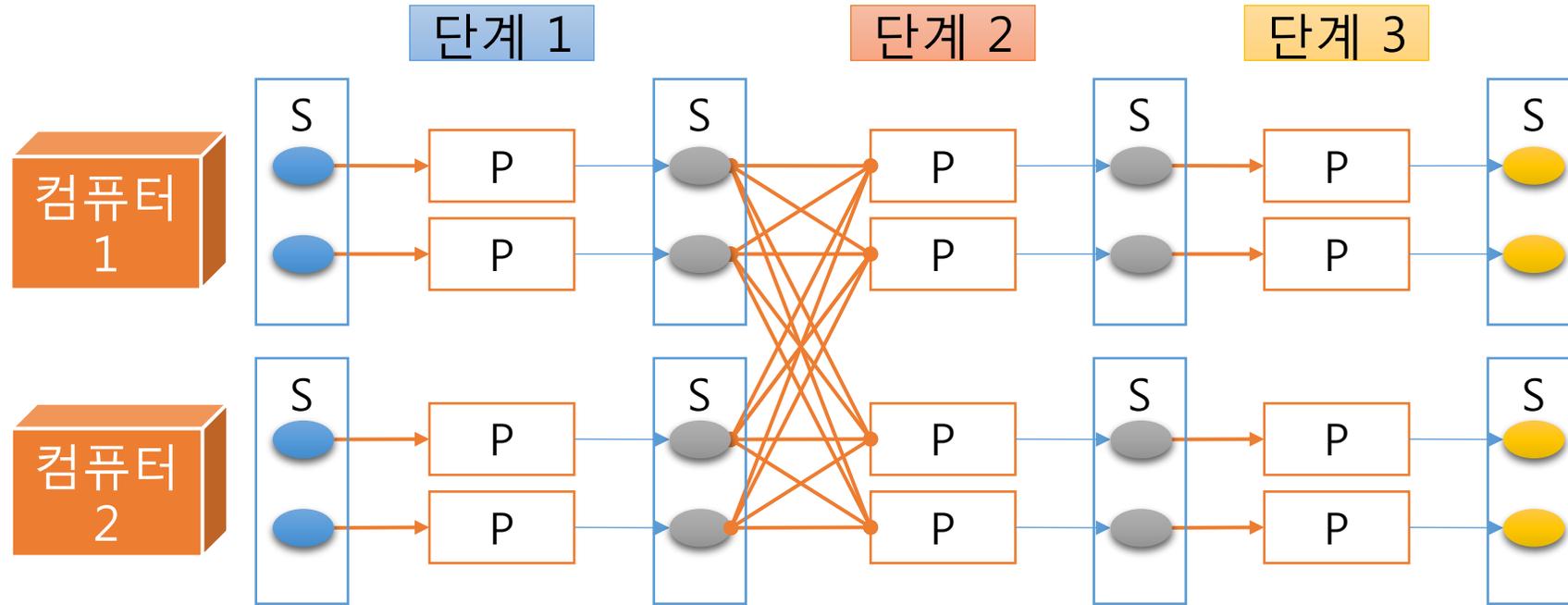
성능 농침
ㄷㄷ

하둡의 실행도중 결함감내와 성능

결함감내
잡았다!



성능 농침
TTT



- 성능이 도망간 이유?
 - 일꾼들이
 - 자기 입력만 **단순&무식하게** 열심히 읽어서
 - **Disk random seek**이 많이 발생

목차

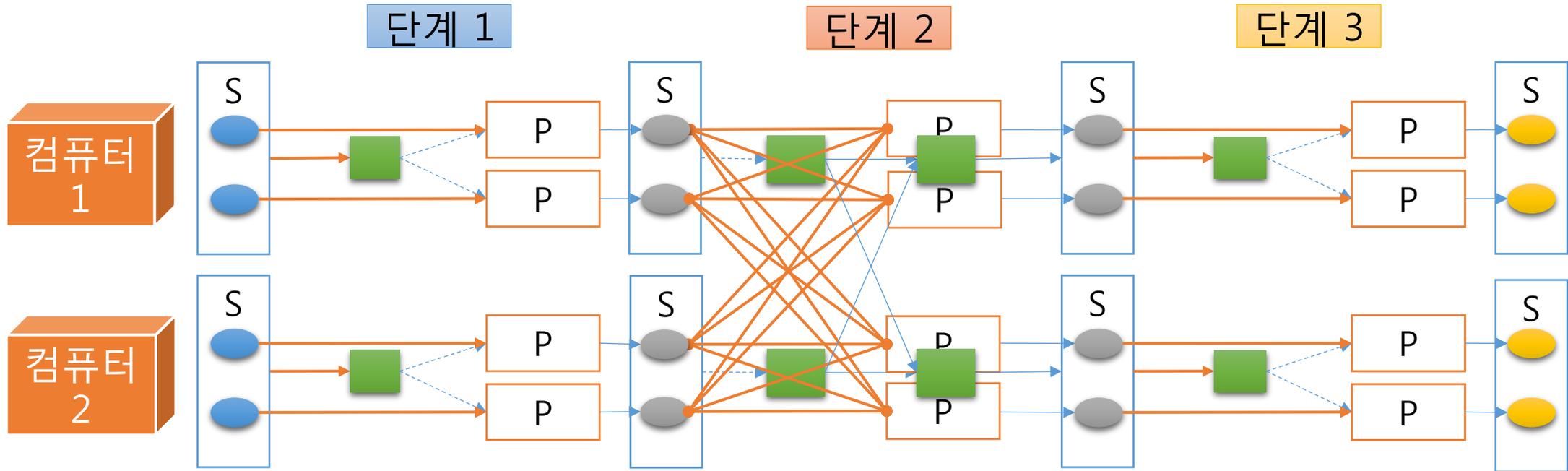
- 빅데이터 처리 프로그램 연구 @ ROSAEC
 - 기여 방안 : 실행도중 결함감내 연구
- Hadoop
 - 실행도중 결함감내는 잡았지만 성능은 놓침
- Hadoop 기틀 재설계
 - 핵심 아이디어 - 화살표 줄이기
 - 성능

기틀 재설계



두 마리 토끼를 다 잡기

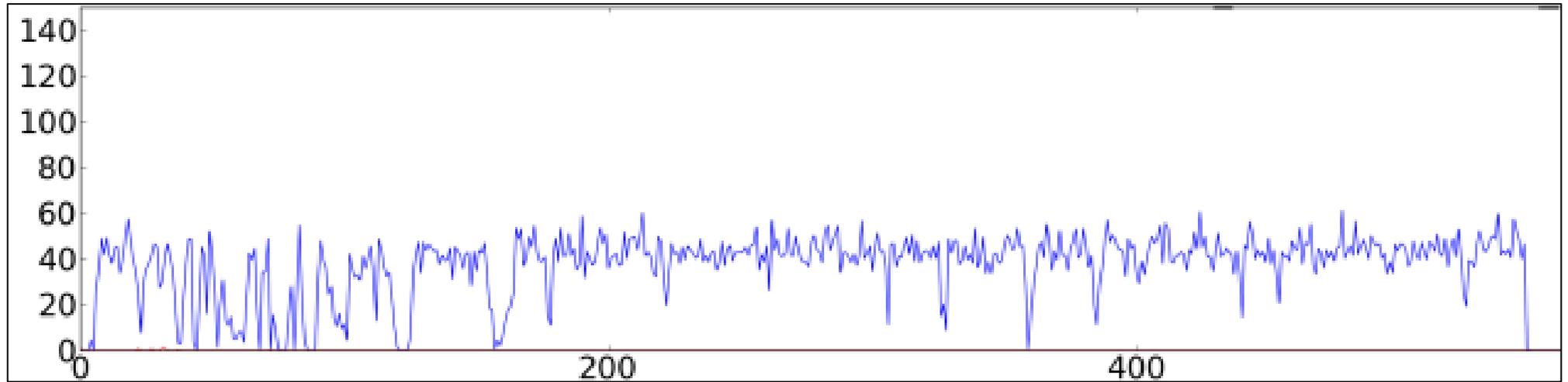
기틀 재설계 - 빨간 화살표 줄이기



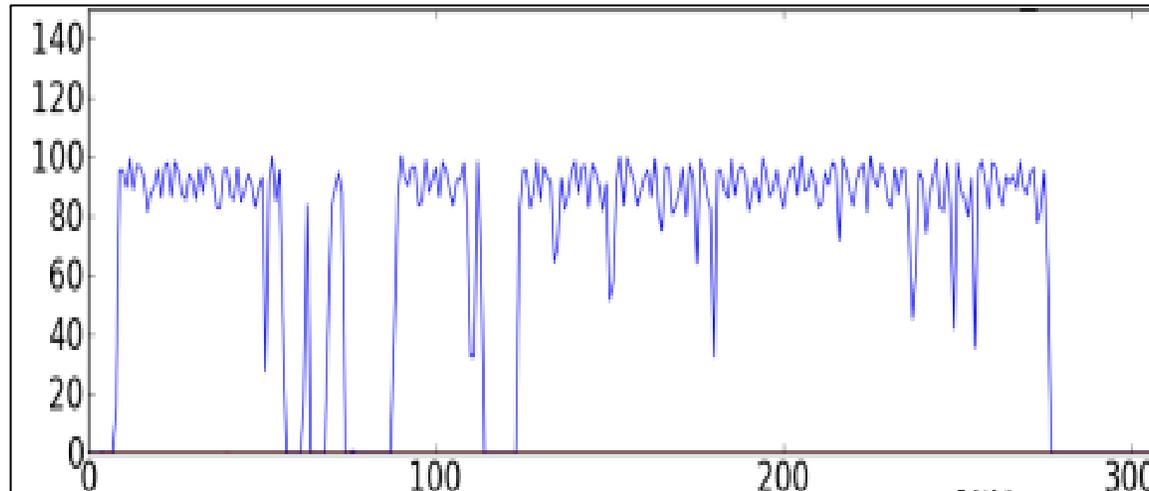
- [성능] 화살표 수의 감소 (하둠 24개 → 새 기틀 4개)
- [실행도중 결함감내] 여전히 각 단계의 결과는 디스크에 기록함

단계 1 디스크 읽기 속도

하둠

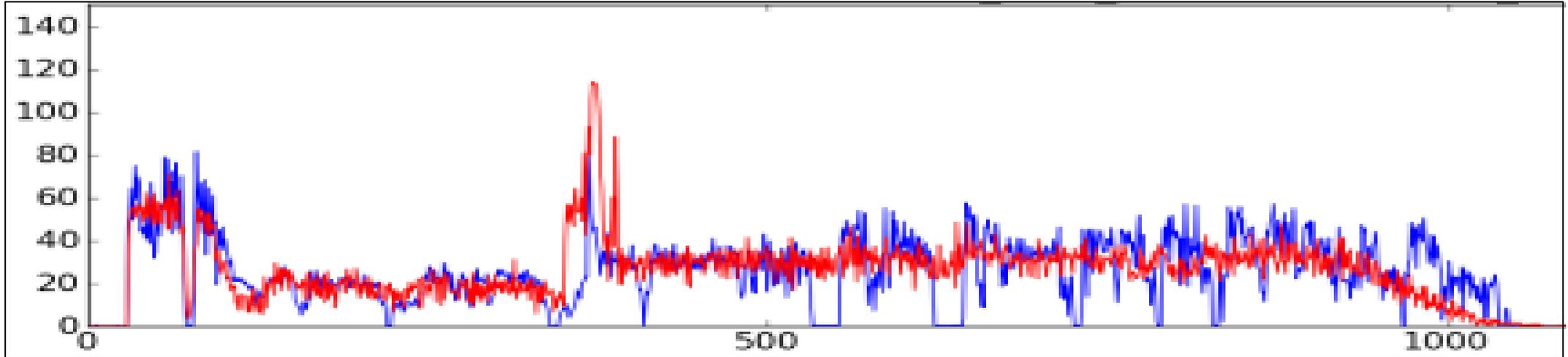


새 기틀

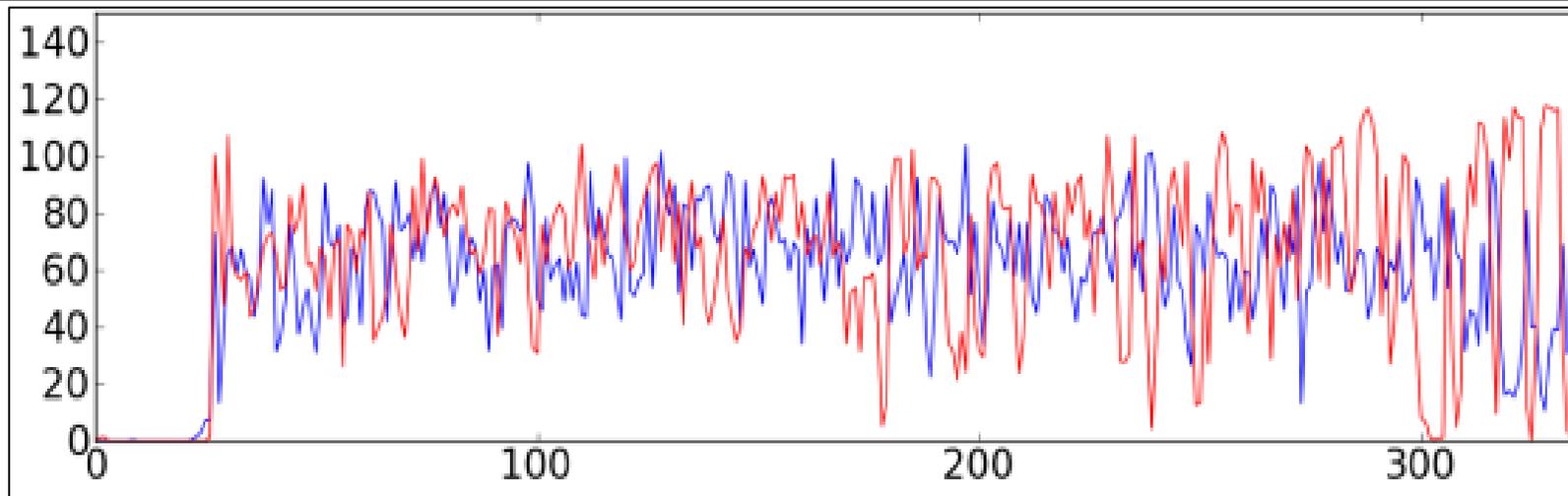


단계 2 네트워크 송수신 속도

하둠

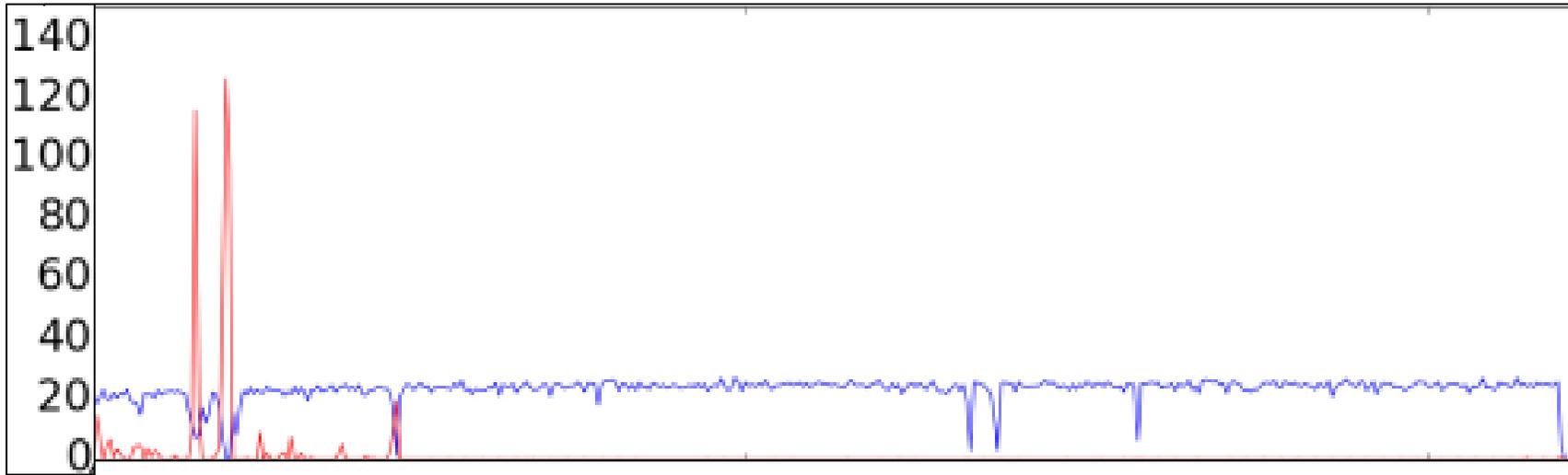


새 기틀

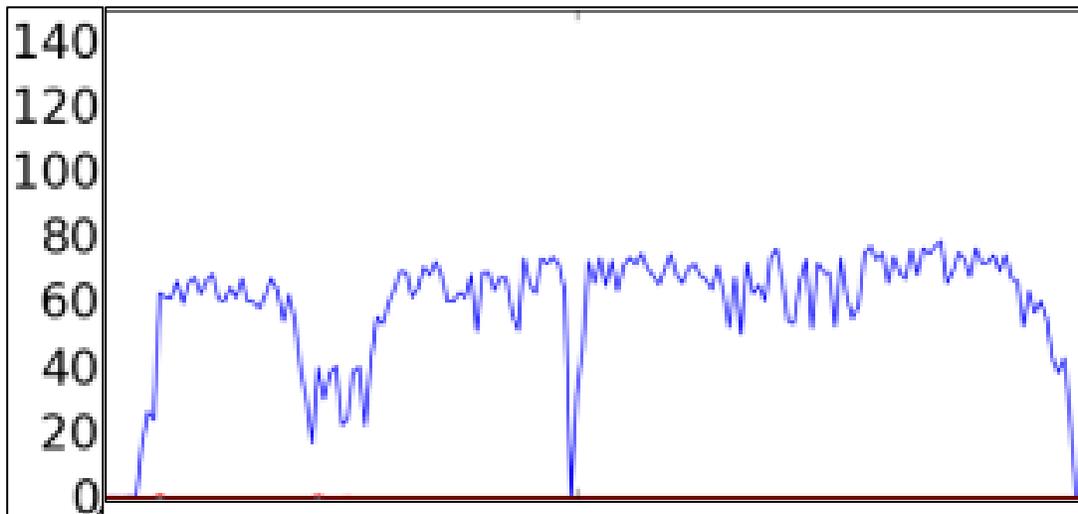


단계 3 디스크 읽기 속도

하둡



새 기틀



정리

- Hadoop
 - **실행도중 결함감내**는 잡았지만 **성능**은 놓침
- Hadoop 기틀 재설계
 - 디스크의 효율적인 동작을 유도
 - **실행도중 결함감내**와 **성능**을 둘 다 잡음
 - 추가적으로
 - **실행도중 결함감내** 끄고 키기 → disk write 감소
 - 몇몇 중요한 파라미터 값을 쉽게 고를 수 있음



END