



Real-time подходы

Vybornov Artyom, avybornov@bigdatateam.org

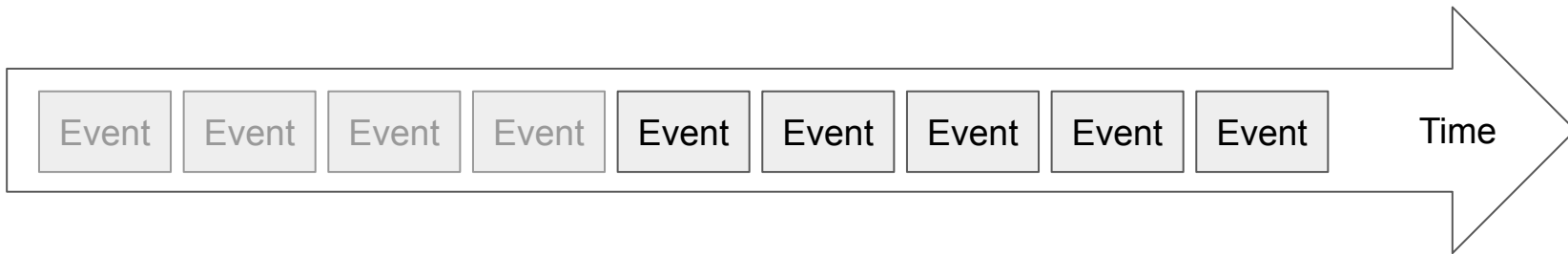
Big Data Instructor, <http://bigdatateam.org/>

Head of Data Platform, Rambler Group

<https://www.linkedin.com/in/artvybor/>

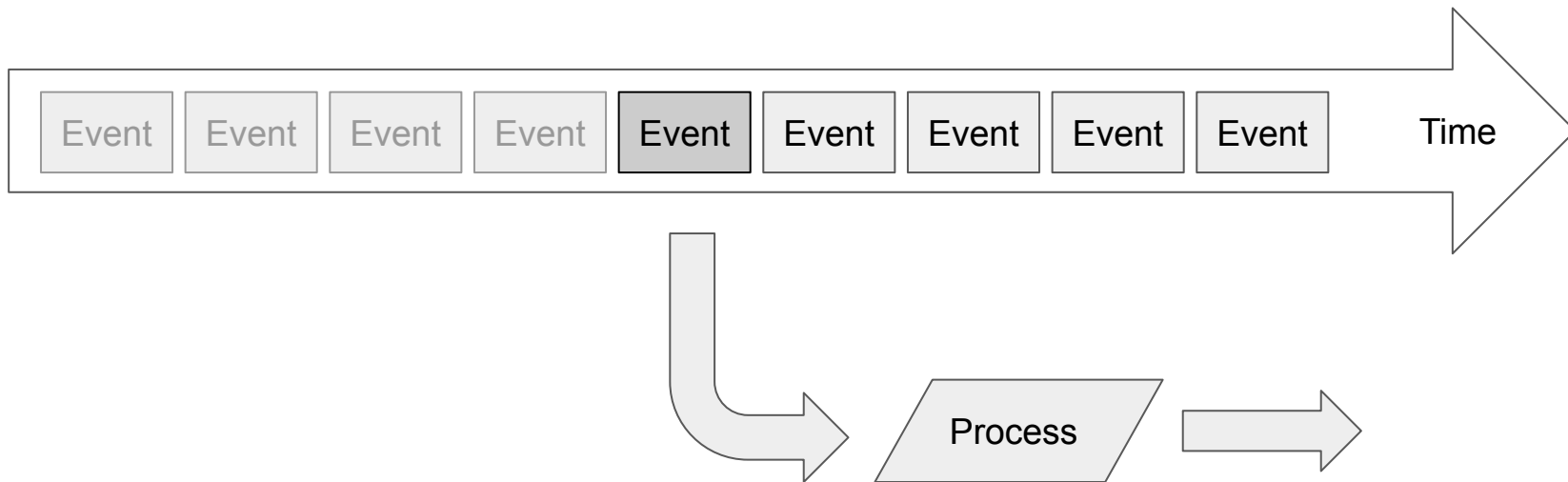


Пособытийный подход





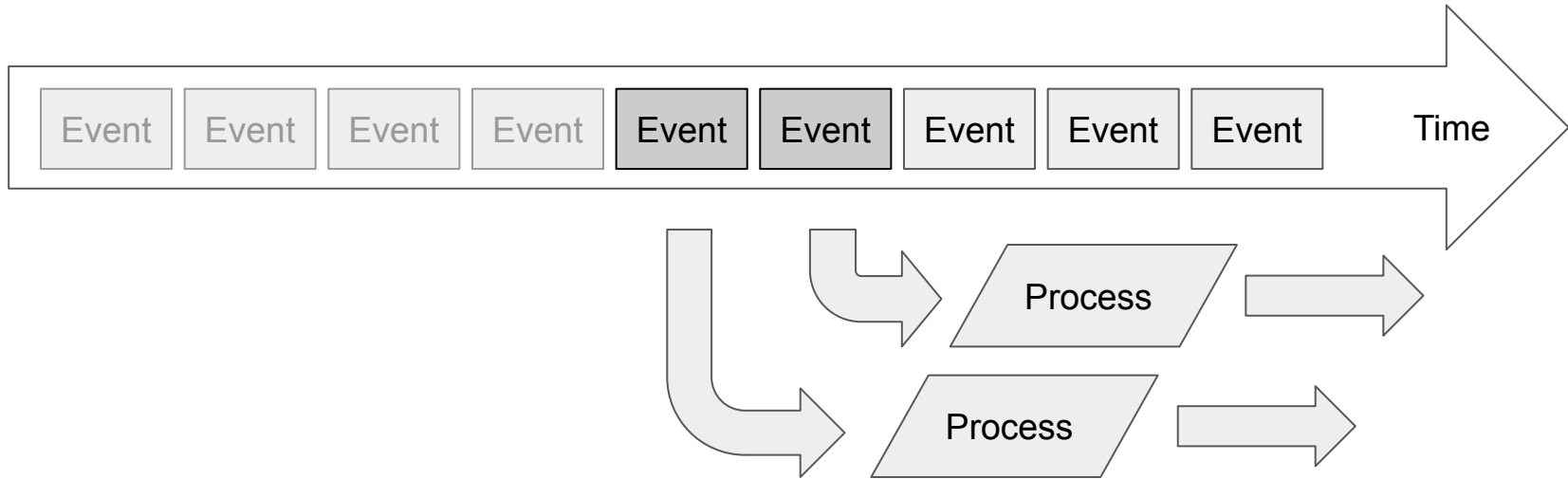
Пособытийный подход



Обработка одного события за раз



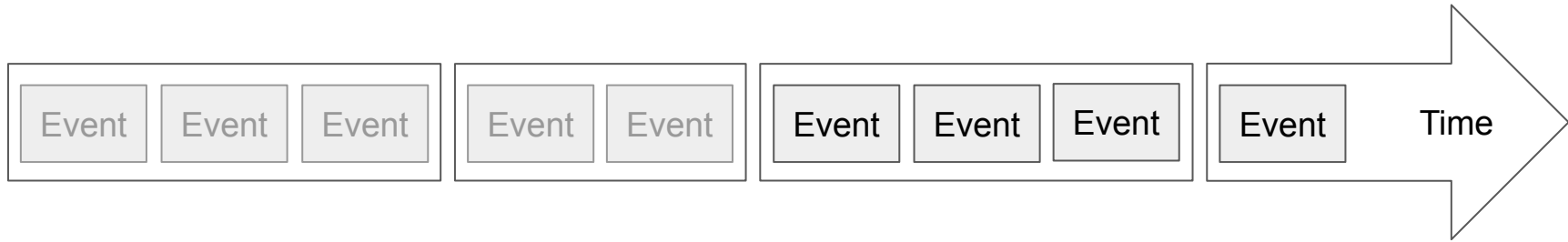
Пособытийный подход



- ▶ Обработка одного события за раз
- ▶ События обрабатываются независимо друг от друга
- ▶ Задержка ~10ms



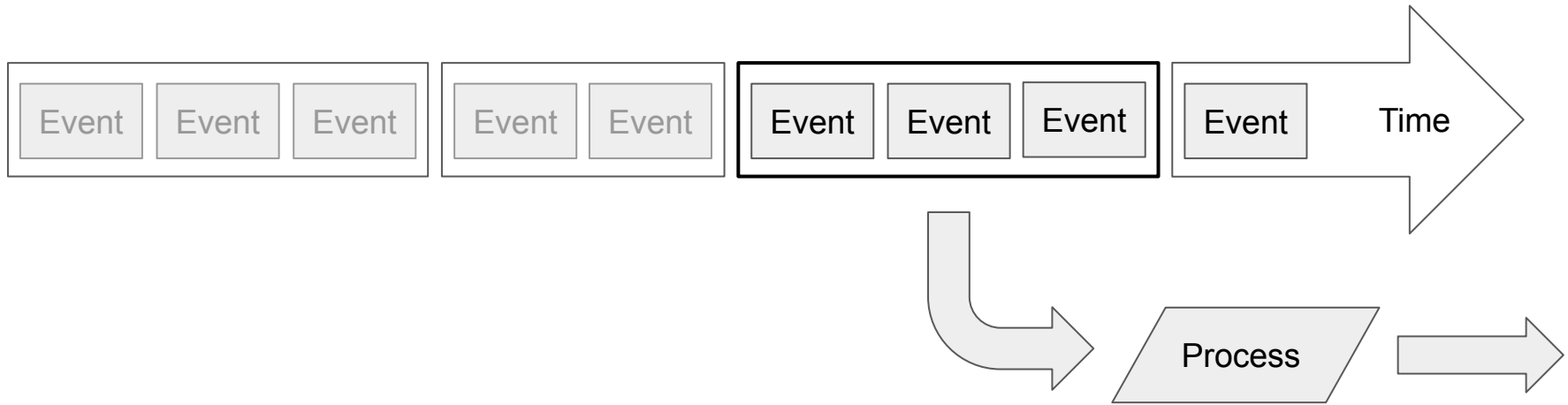
Микробатчевый подход



Поток событий нарезается на батчи (к примеру батч формируется из событий собранных за каждые 10 секунд)



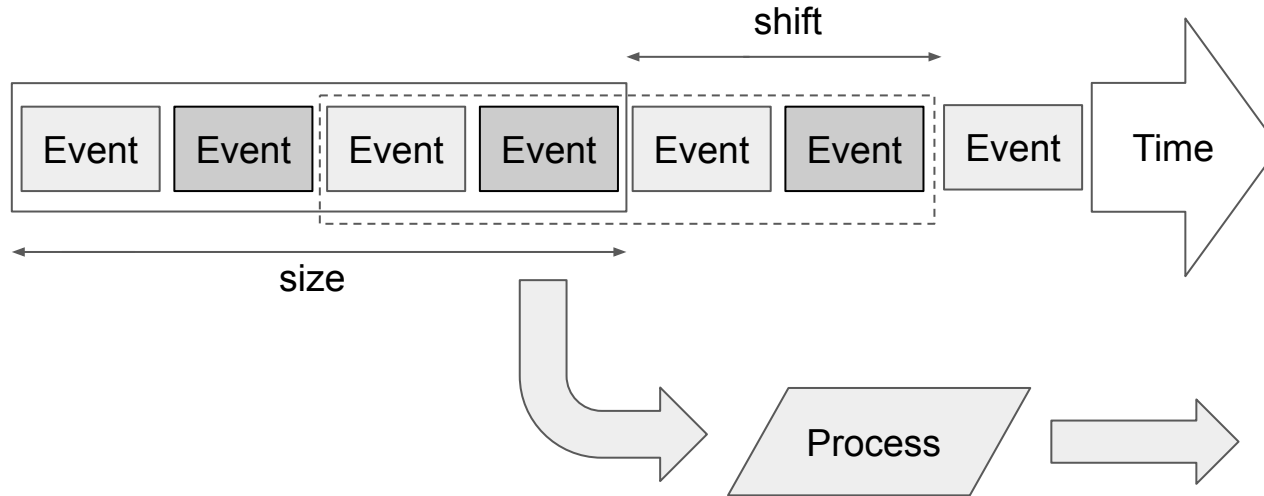
Микробатчевый подход



- ▶ Поток событий нарезается на батчи (к примеру батч формируется из событий собранных за каждые 10 секунд)
- ▶ Обычно батчи обрабатывают последовательно
- ▶ Задержка $\gg 1s$



Оконный подход



- ▶ Батч формируется с помощью скользящего окна
- ▶ Подвид микробатчевого подхода



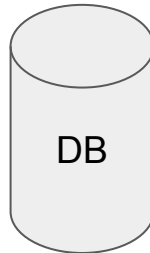
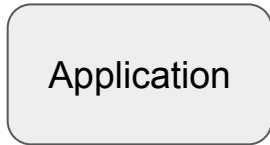
Пособытийный vs Микробатч

- ▶ Пособытийный подход позволяет достичь наименьшего лага
- ▶ Микробатч позволяет сэкономить ресурсы



Пособытийный vs Микробатч

- ▶ Пособытийный подход позволяет достичь наименьшего лага
- ▶ Микробатч позволяет сэкономить ресурсы





Пособытийный vs Микробатч

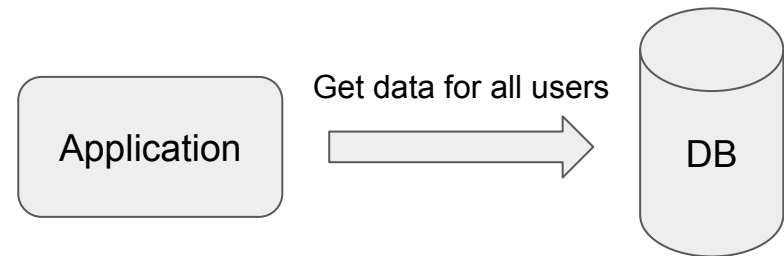
- ▶ Пособытийный подход позволяет достичь наименьшего лага
- ▶ Микробатч позволяет сэкономить ресурсы





Пособытийный vs Микробатч

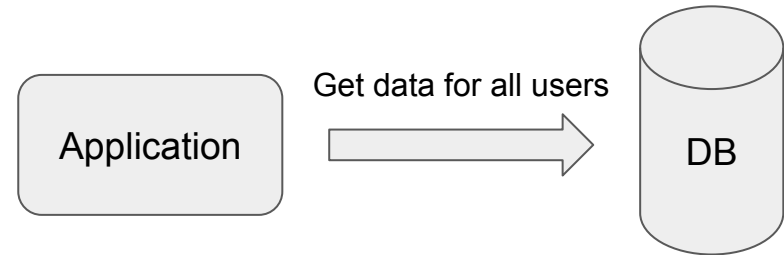
- ▶ Пособытийный подход позволяет достичь наименьшего лага
- ▶ Микробатч позволяет сэкономить ресурсы





Пособытийный vs Микробатч

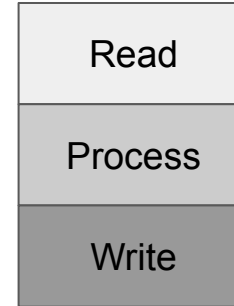
- ▶ Пособытийный подход позволяет достичь наименьшего лага
- ▶ Микробатч позволяет сэкономить ресурсы





Пособытийный vs Микробатч

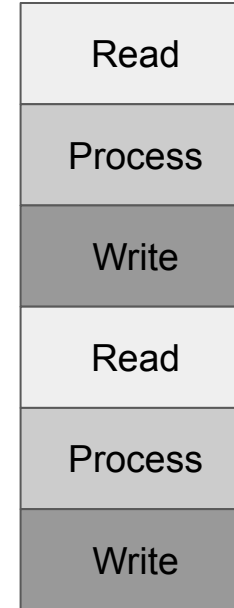
- ▶ Пособытийный подход позволяет достичь наименьшего лага
- ▶ Микробатч позволяет сэкономить ресурсы





Пособытийный vs Микробатч

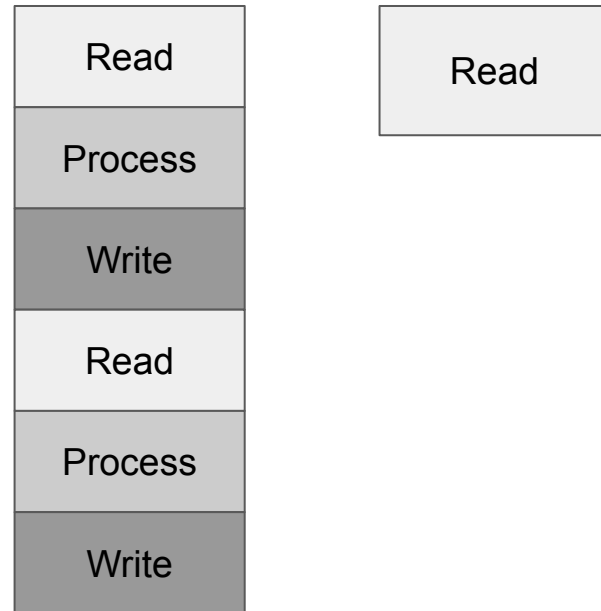
- ▶ Пособытийный подход позволяет достичь наименьшего лага
- ▶ Микробатч позволяет сэкономить ресурсы





Пособытийный vs Микробатч

- ▶ Пособытийный подход позволяет достичь наименьшего лага
- ▶ Микробатч позволяет сэкономить ресурсы





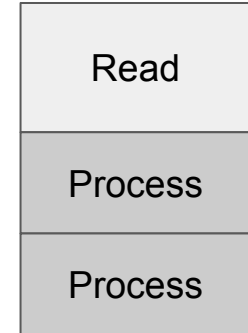
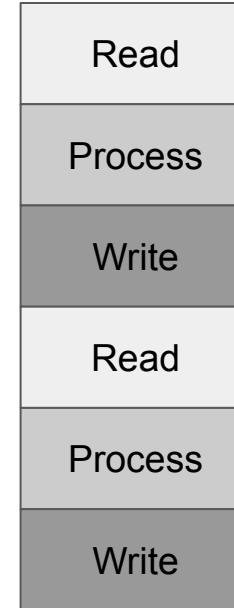
Пособытийный vs Микробатч



Пособытийный подход позволяет достичь наименьшего лага



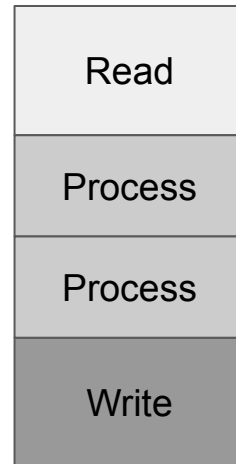
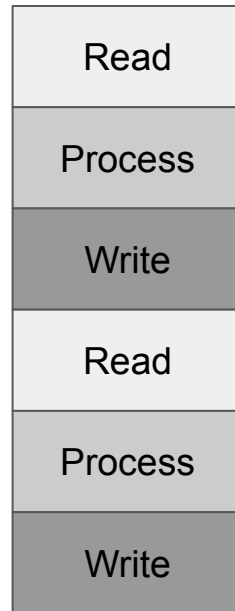
Микробатч позволяет сэкономить ресурсы





Пособытийный vs Микробатч

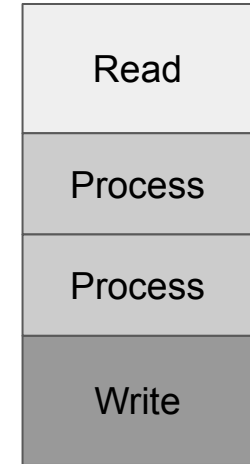
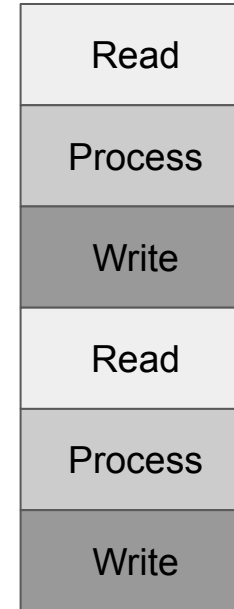
- ▶ Пособытийный подход позволяет достичь наименьшего лага
- ▶ Микробатч позволяет сэкономить ресурсы





Пособытийный vs Микробатч

- ▶ Пособытийный подход позволяет достичь наименьшего лага
- ▶ Микробатч позволяет сэкономить ресурсы
- ▶ Микробатч позволяет обработать больше данных на тех же ресурсах в сравнении с пособытийным





Пропускная способность VS Задержка

- ▶ В реальной жизни ресурсы всегда ограничены
- ▶ Большие данные требуют огромной пропускной способности => в большинстве случаев используем микробатч
- ▶ В каждом отдельном случае вы должны выбирать решение согласно задаче



Пропускная способность VS Задержка

