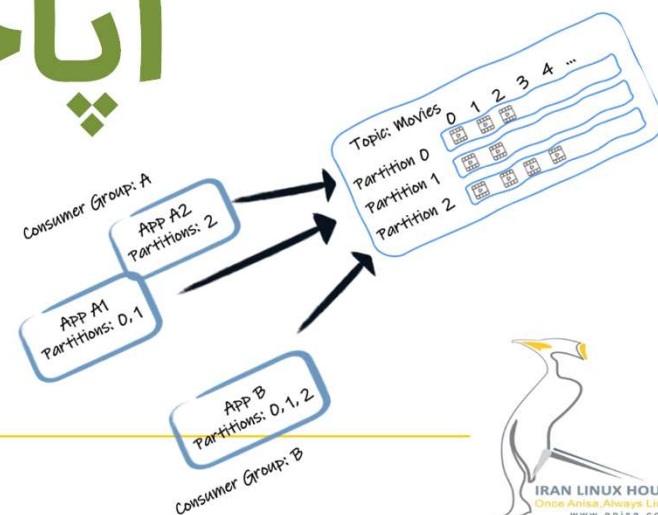




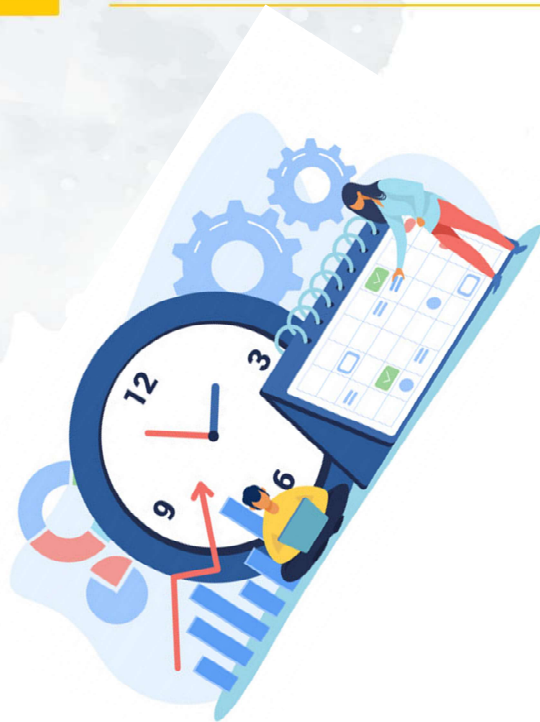
# آشنایی با

# آپاچی کافکا



# آنچه خواهیم دید

- مروری بر سیستم‌های پیام‌رسانی
- تاریخچه و فلسفه ایجاد کافکا
- مفاهیم پایه کافکا
- خطوط پردازش داده
- سایر پیام‌رسان‌های توزیع‌شده



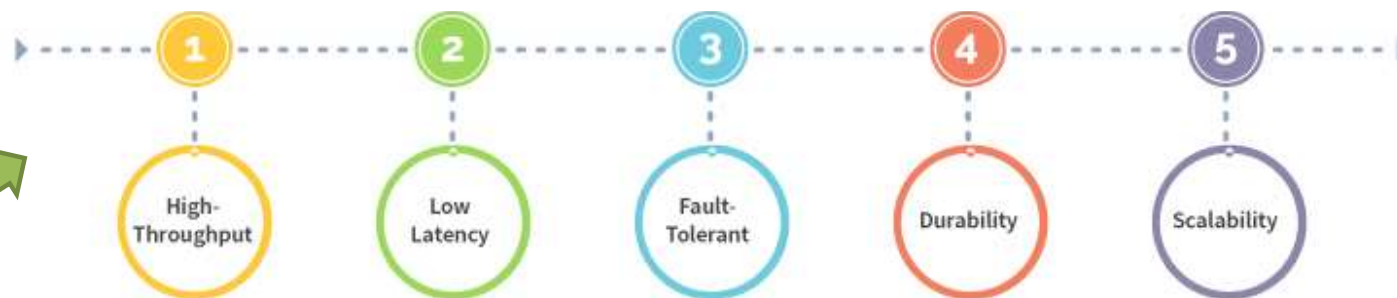
# کافکا چیست ؟

## APACHE KAFKA

*More than 80% of all Fortune 100 companies trust, and use Kafka.*

Apache Kafka is an open-source distributed event streaming platform used by thousands of companies for high-performance data pipelines, streaming analytics, data integration, and mission-critical applications.

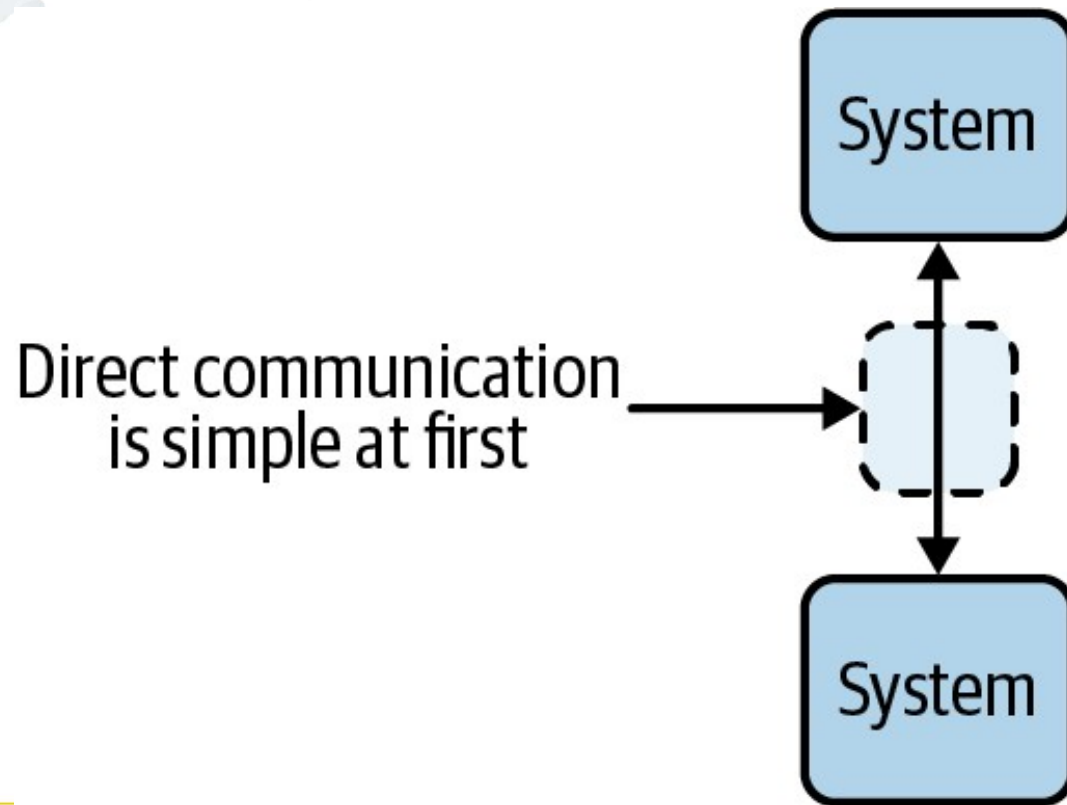
در اکثر معماری‌های کلان‌داده، کافکا نقش محوری و واسط بین سرویس‌های مختلف را ایفا می‌کند



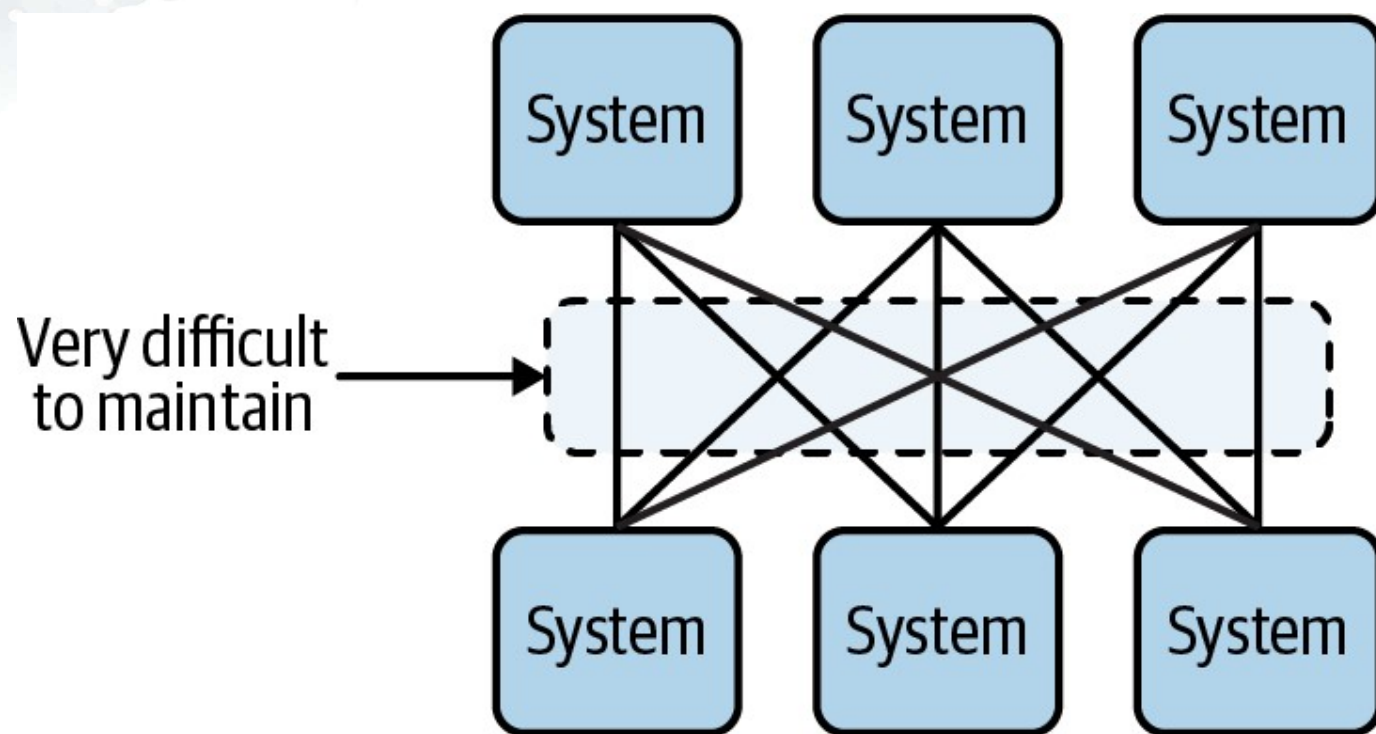
مزایای اصلی

LI  
C

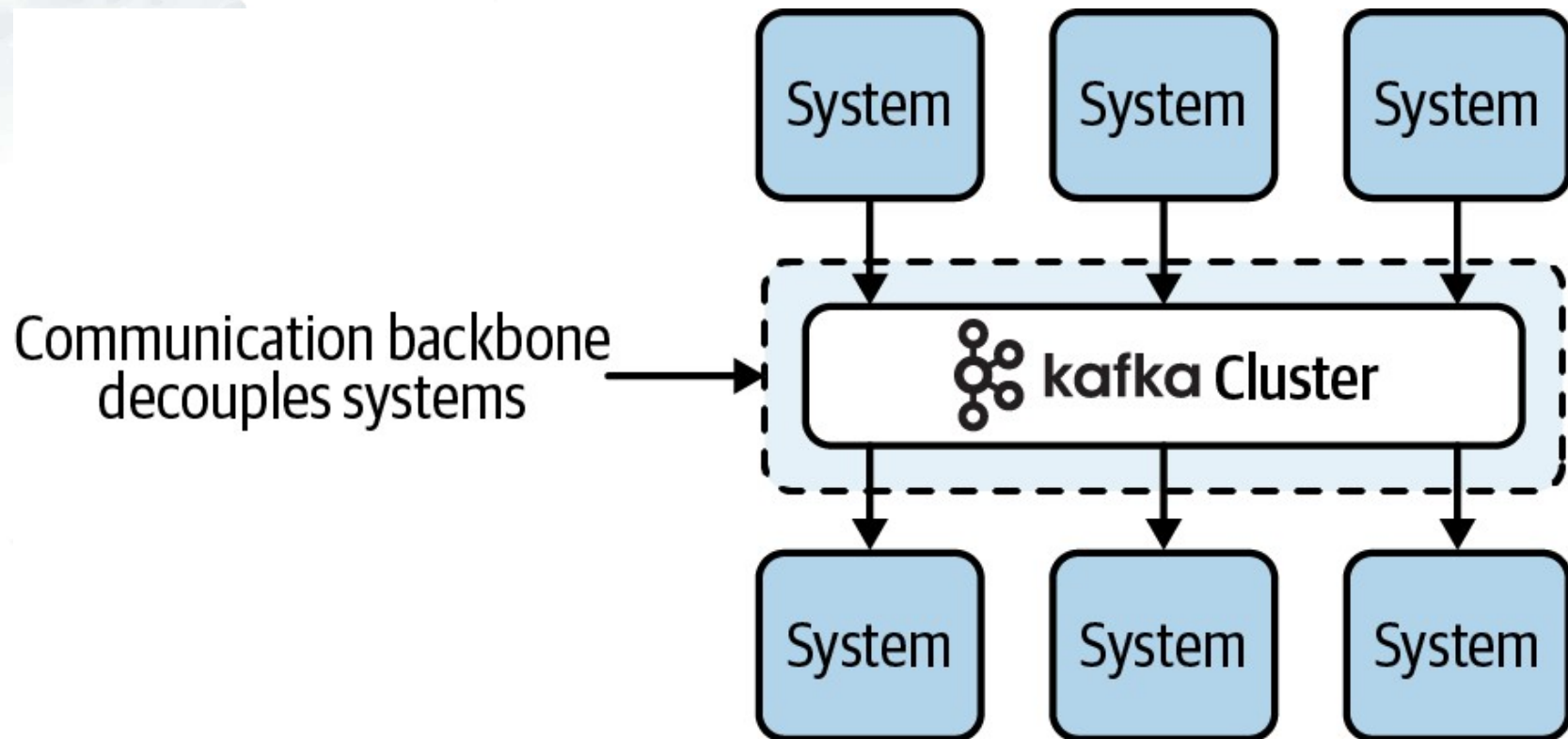
# قصه از کجا شروع شد



# قصه از کجا شروع شد

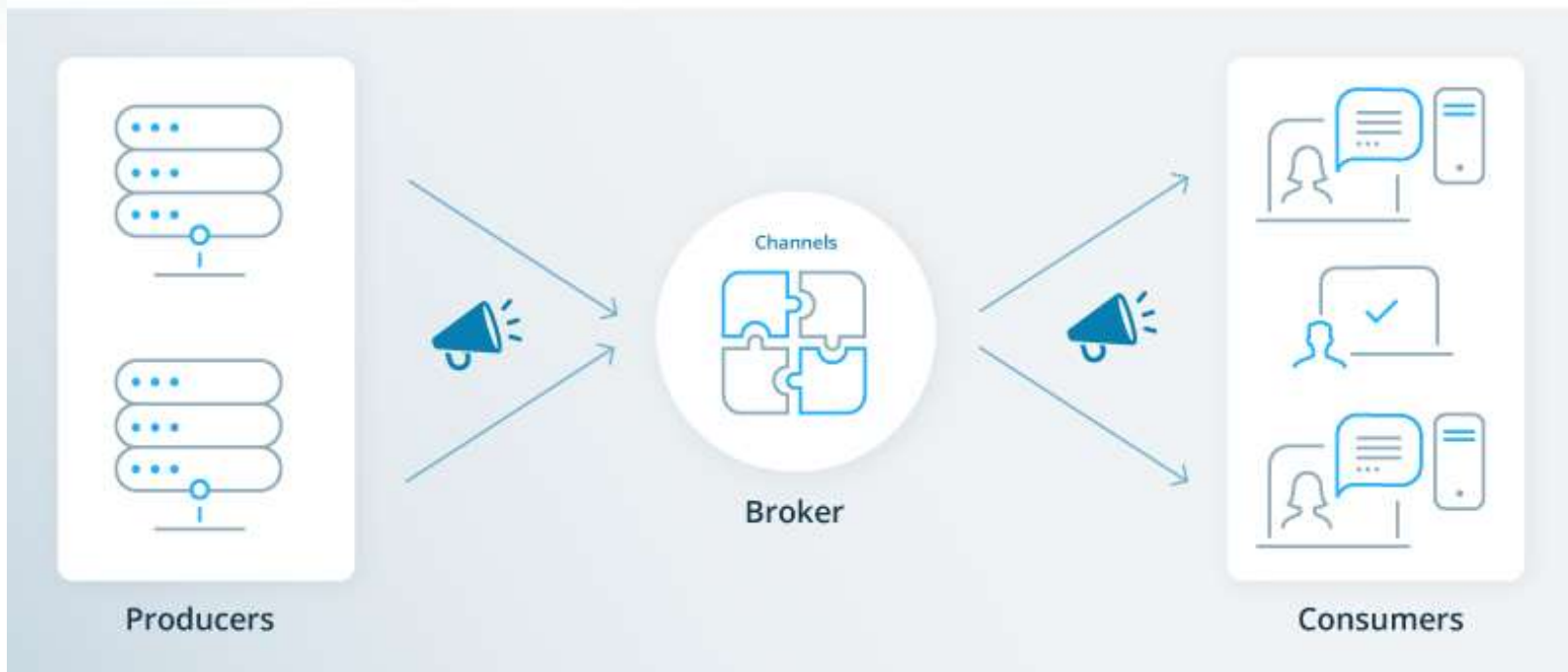


# کافکا وارد می شود



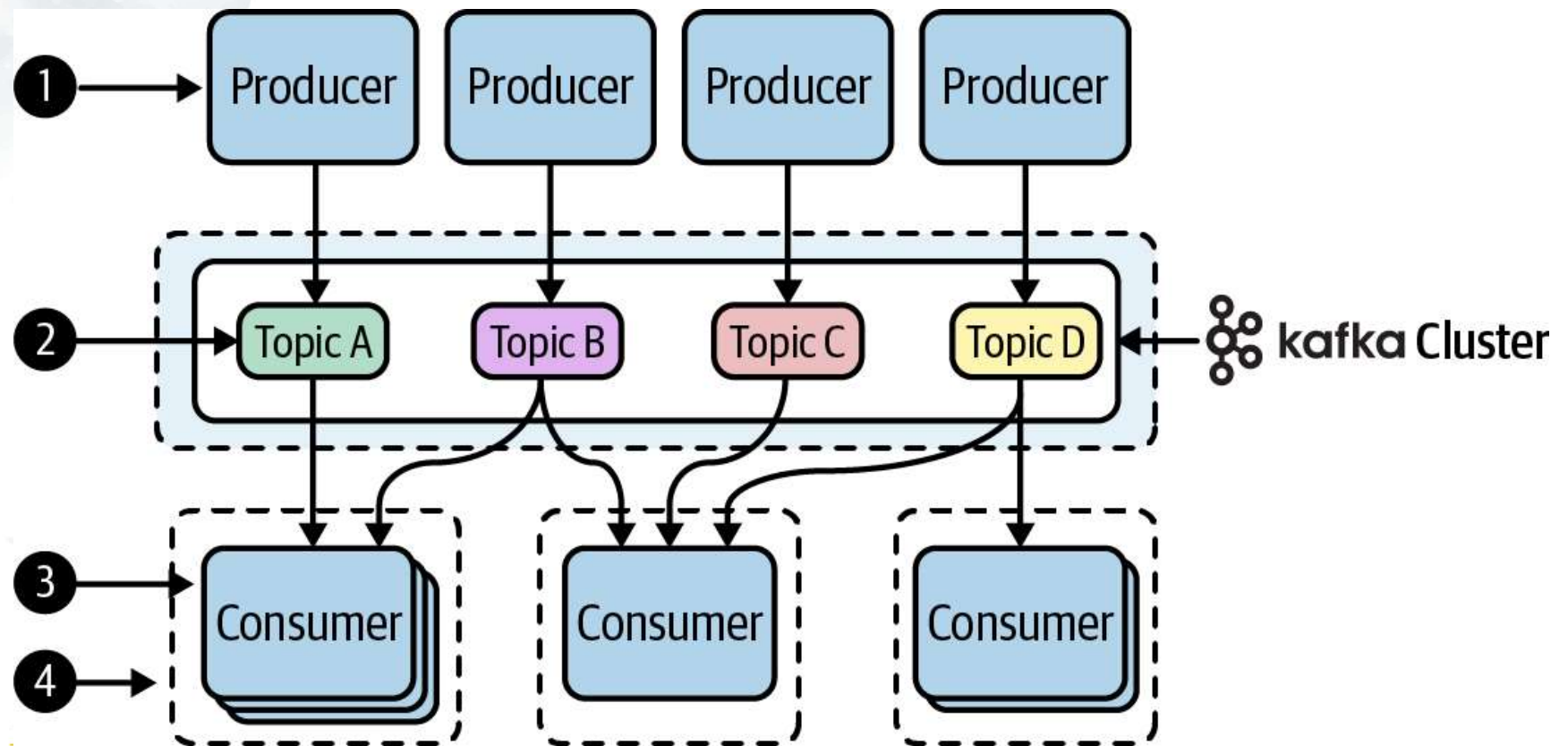
# یک نیازمندی جدید: معماری رویدادگرا

## Event-Driven Architecture





# معماری یک سامانه مبتنی بر کافکا

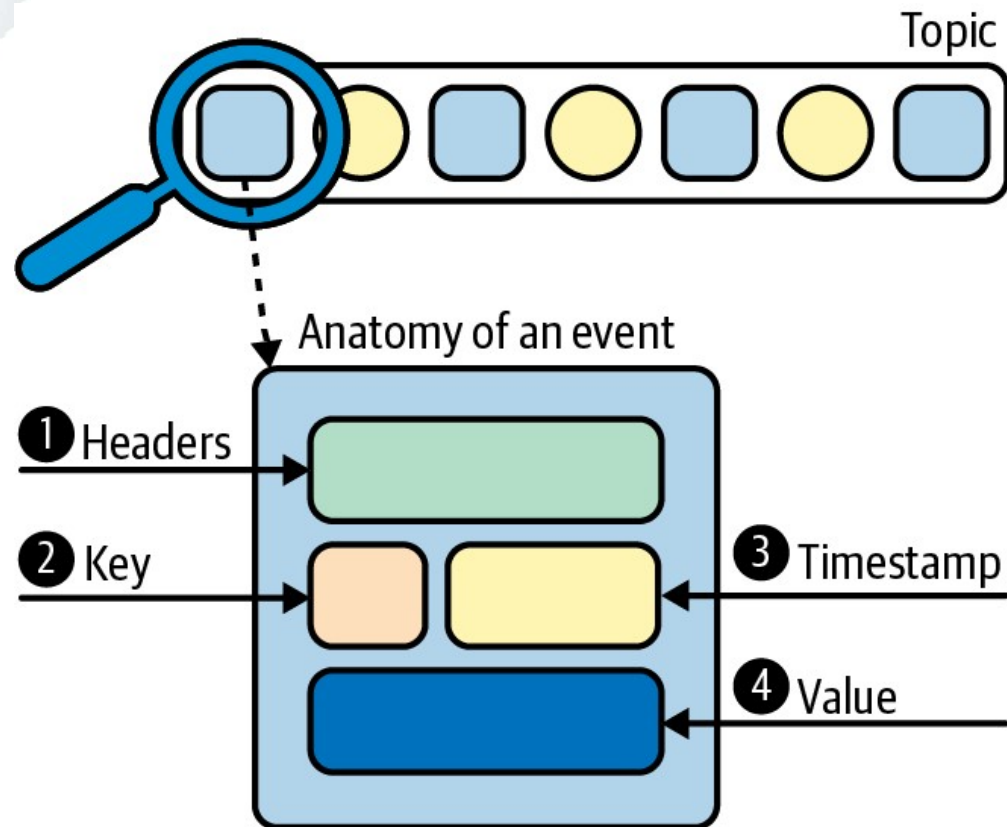




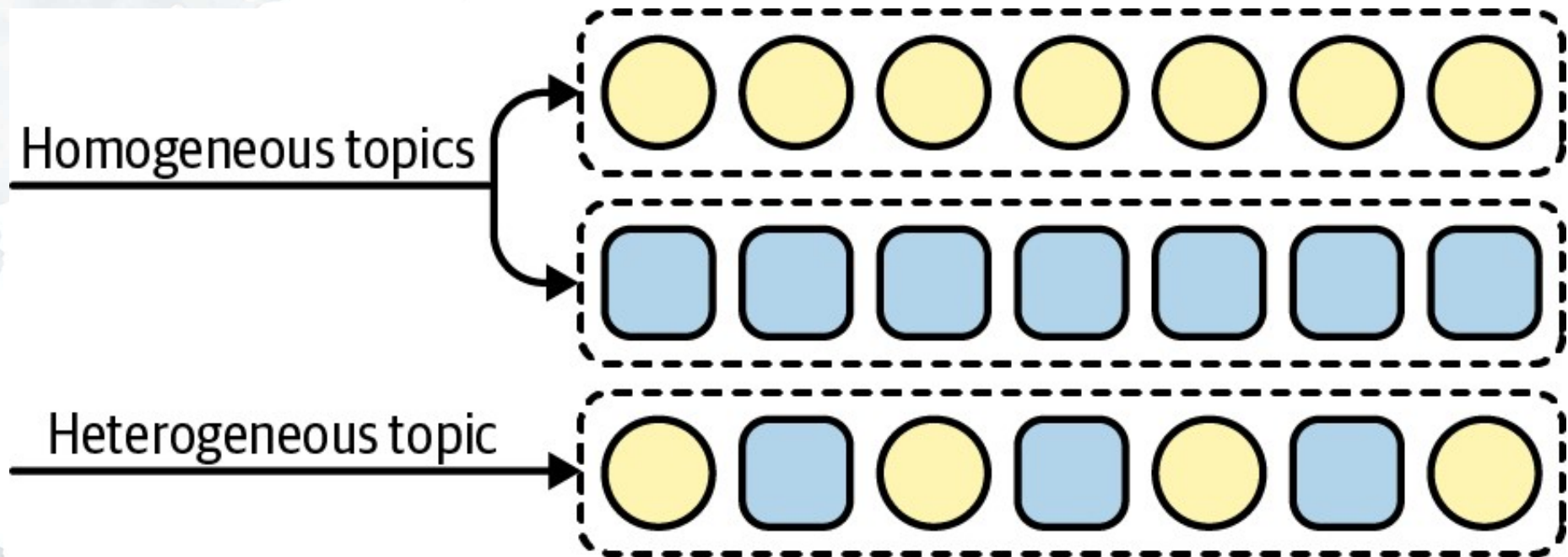
# مفاهیم پایه کافکا

- **پیام / رکورد** : کوچکترین واحد ذخیره و دریافت داده‌ها
- **بروکر** : به هر نود کافکا در شبکه که قابلیت دریافت یا ارسال پیام را دارد، بروکر می‌گوییم.
- **تولیدکننده** : هر منبعی که تولیدکننده پیام است و آنرا به کافکا ارسال می‌کند.
- **مصرف‌کننده** : هر منبعی که مصرف‌کننده پیام‌ها در کافکاست.
- **گروه مصرف‌کننده** : مجموعه‌ای از مصرف‌کنندگان با یک شناسه مشترک گروه.
- **تاپیک** : برای گروه‌بندی و مدیریت پیام‌ها استفاده می‌شود. برای ارسال پیام یا دریافت پیام از کافکا، ابتدا باید تاپیک آنرا مشخص کرد.
- **پارتیشن** : هر تاپیک، به بخش‌هایی تقسیم می‌شود که خواندن و ذخیره پیام‌ها در هر پارتیشن انجام می‌شود.
- **زوکپیر** : مدیریت کلاستر کافکا و ذخیره آفست‌های هر پارتیشن به ازای هر گروه مصرف‌کننده بر عهده زوکپیر است. (در نسخه‌های جدید در حال حذف است)

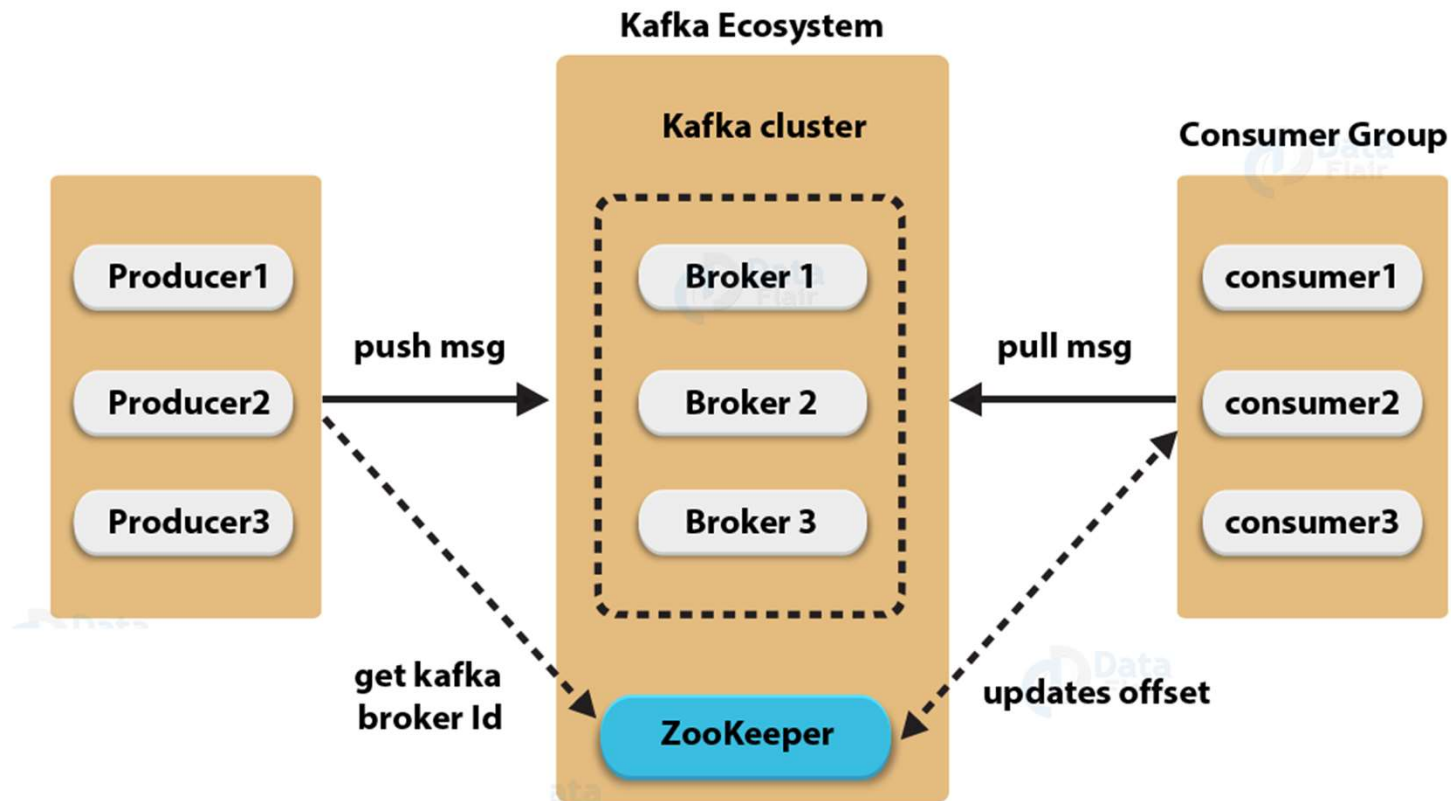
# ساختار پیام در کافکا



# انواع پیام‌ها : همگن/ناهمگن



# نقش زوکیپر در کلاستر کافکا



# نقش آفست در دریافت و ارسال پیام

گروه مصرف‌کنندگان

**Partition 0**



**Partition 1**



**Partition 2**



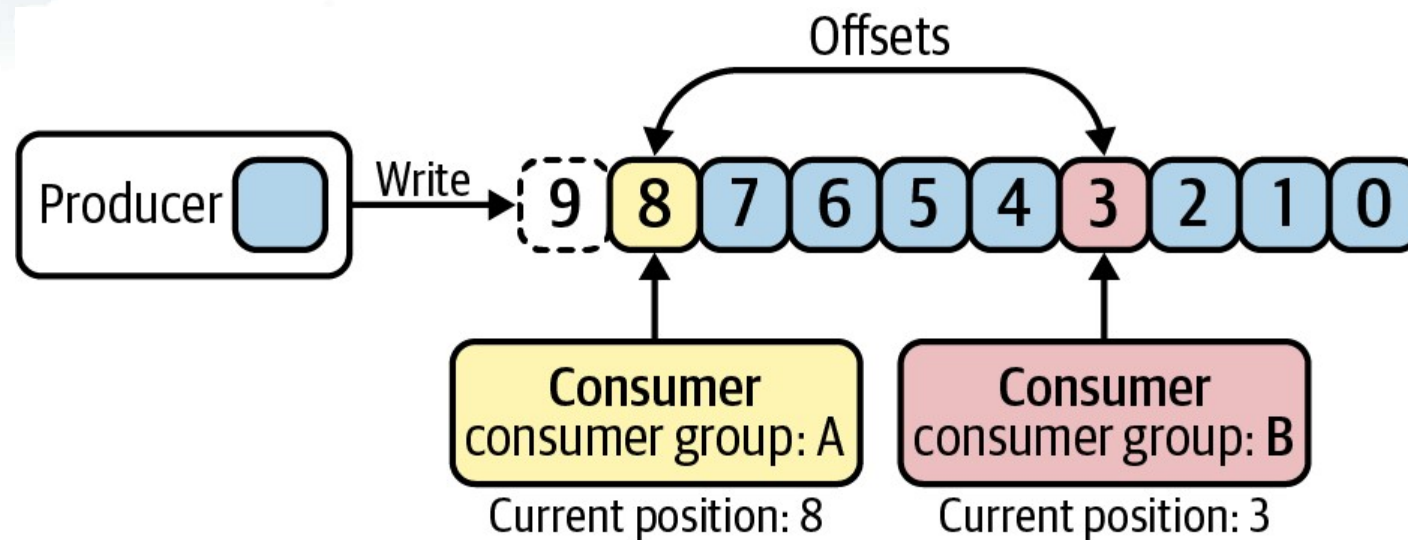
**Writes**

تولید کننده

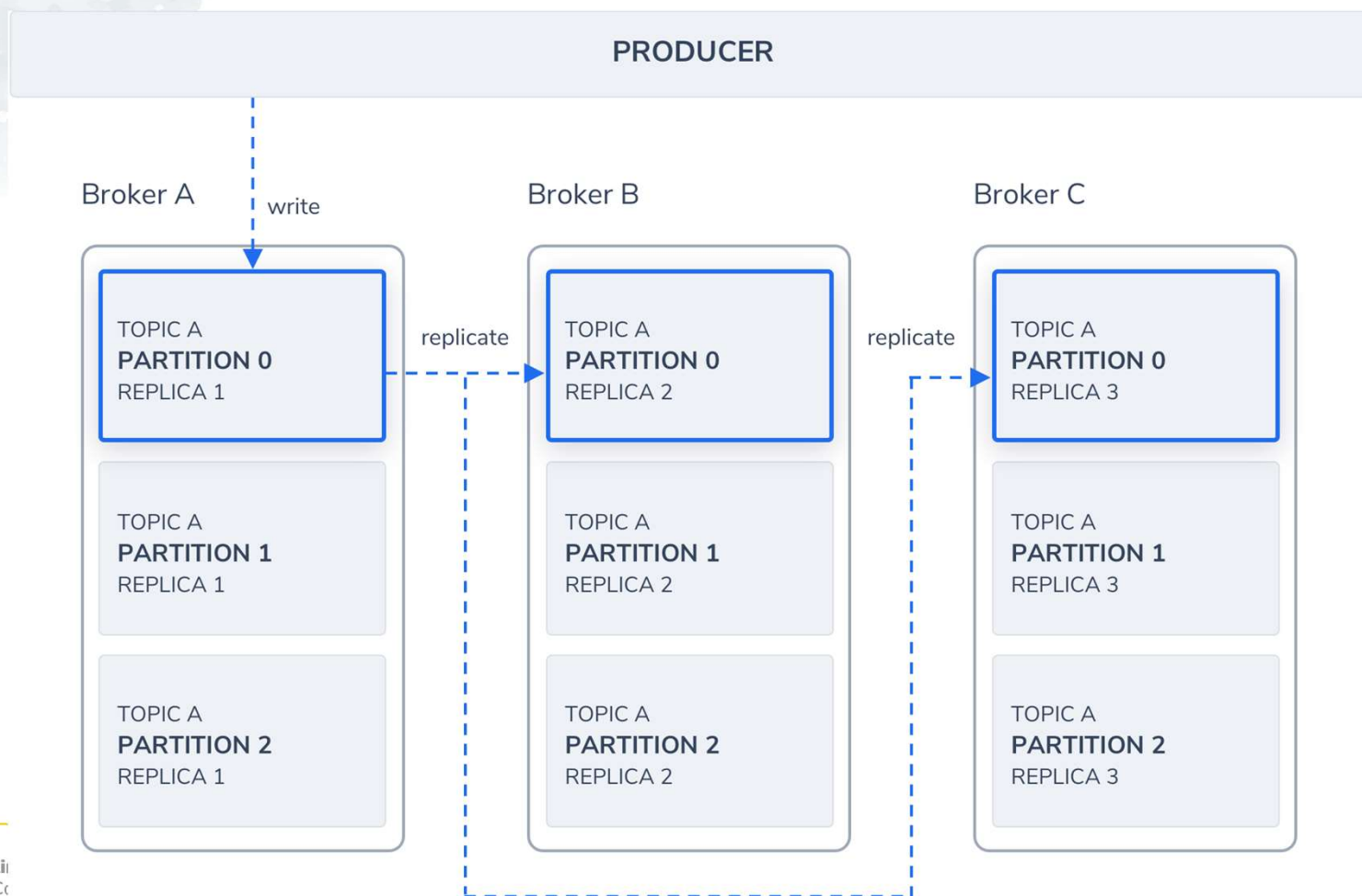
Old

New

# مصرف کنندگان گوناگون بر روی یک تاپیک



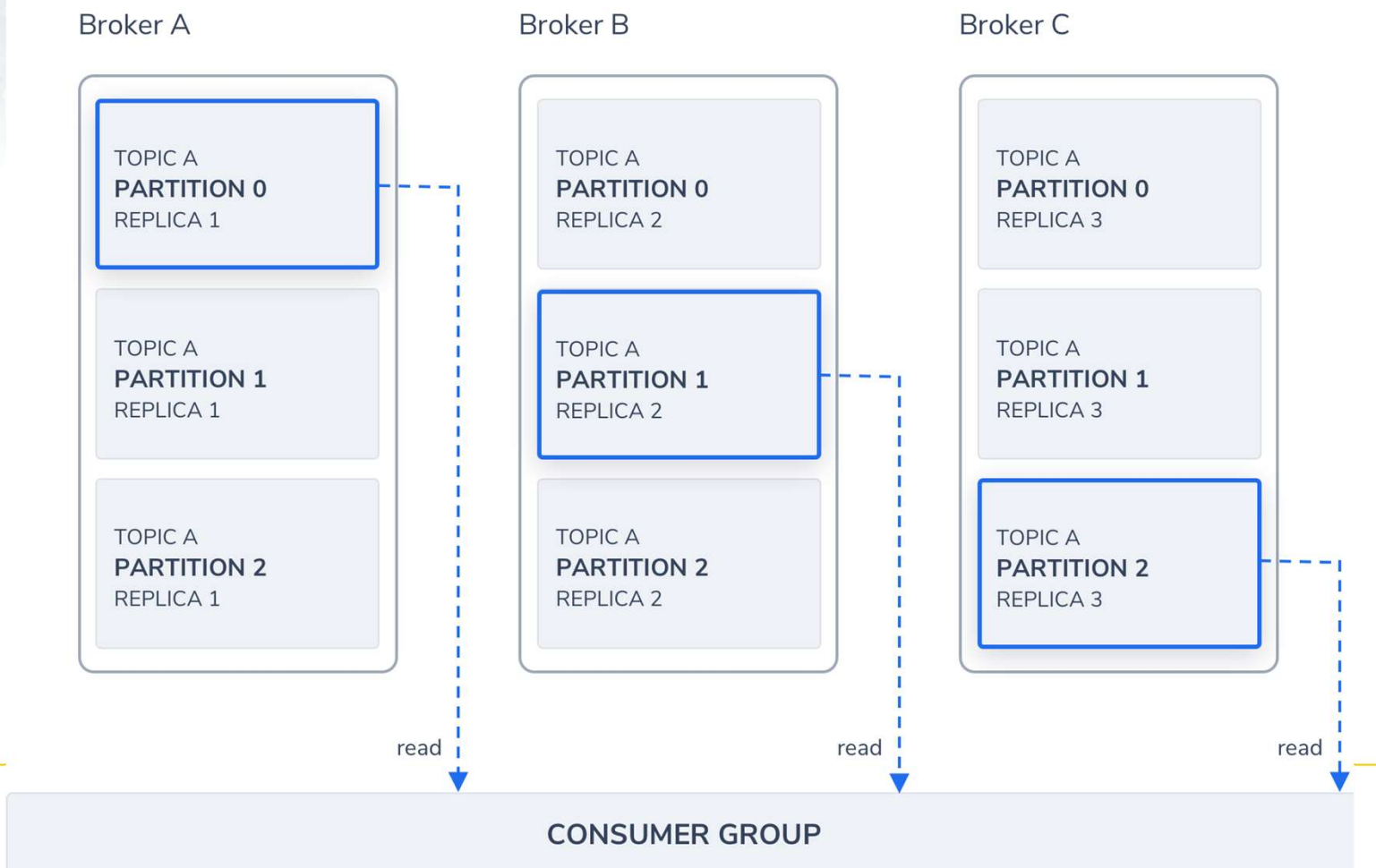
# پارتیشن / رپلیکا



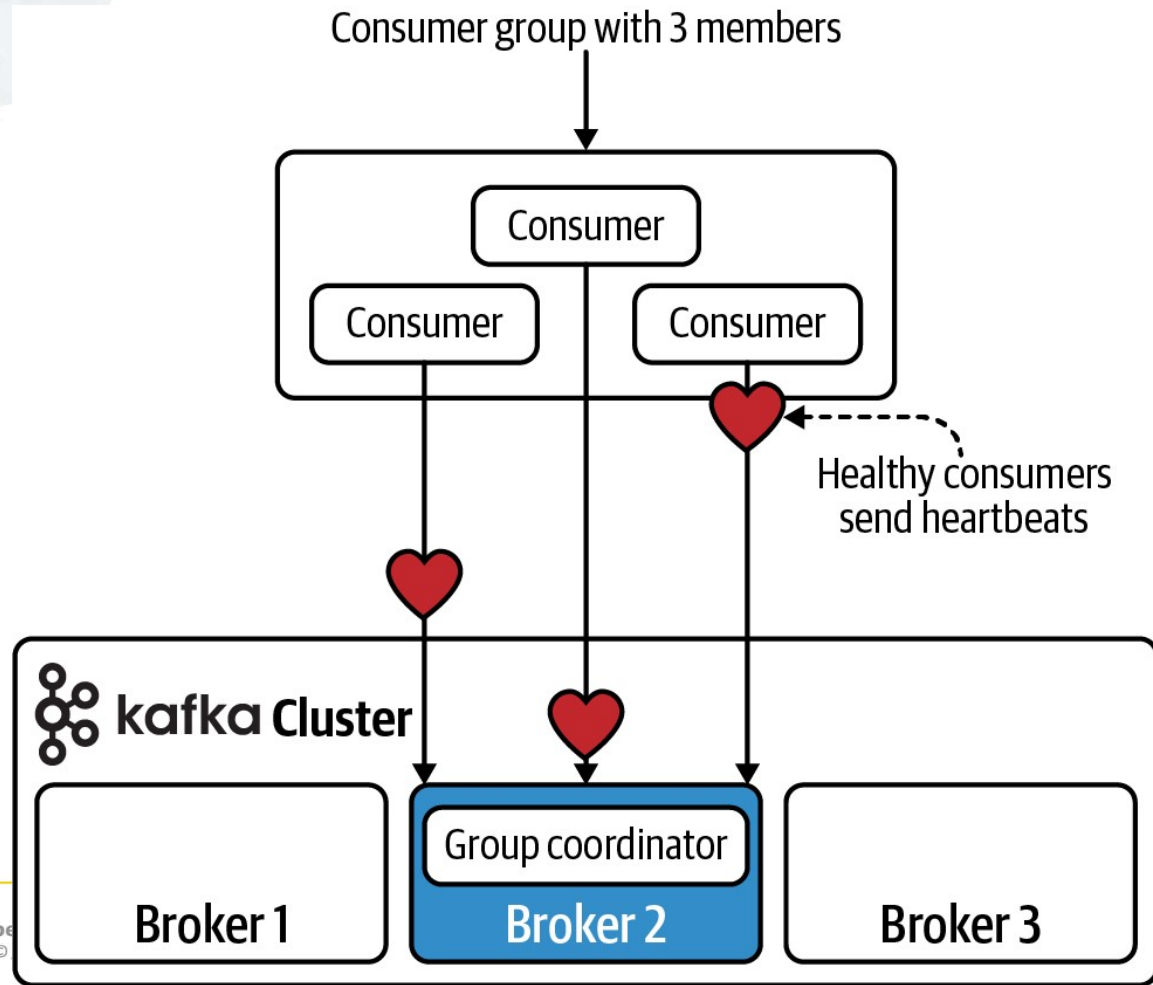
ضریب تکرار ۳



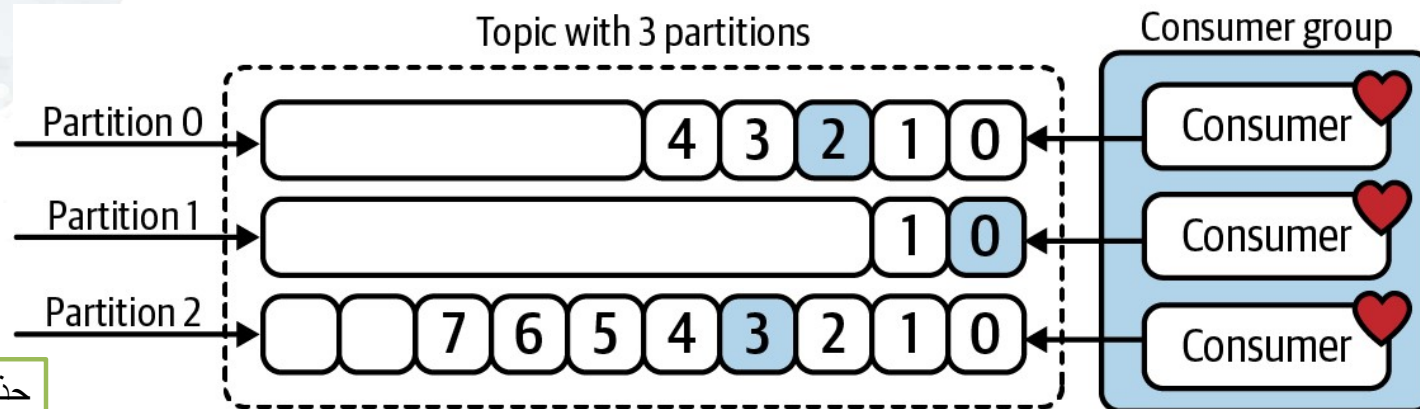
# رپلیکا و افزایش سرعت خواندن



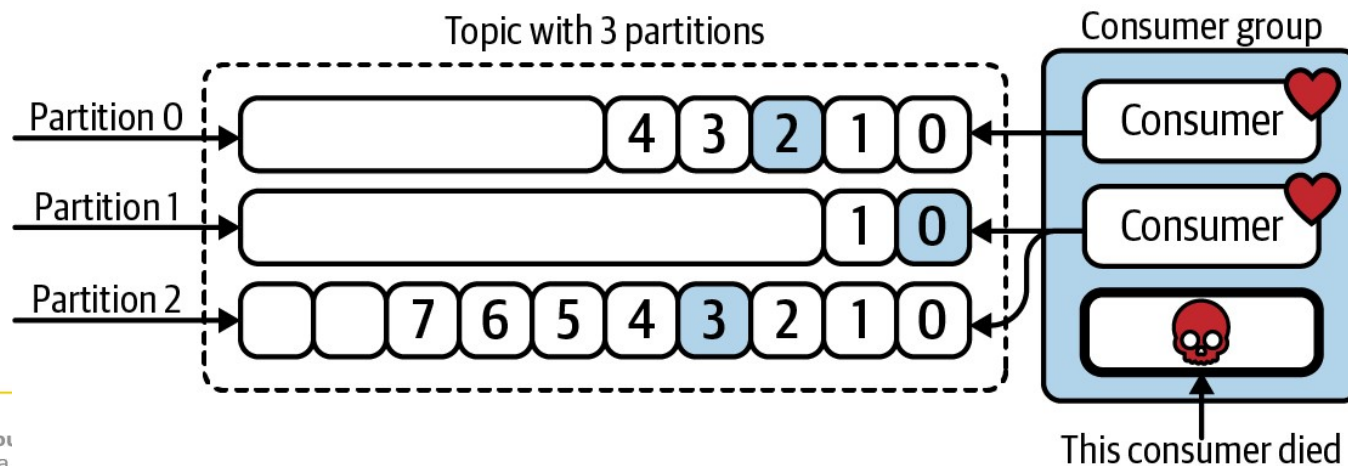
# فرآیند مصرف و دریافت پیام



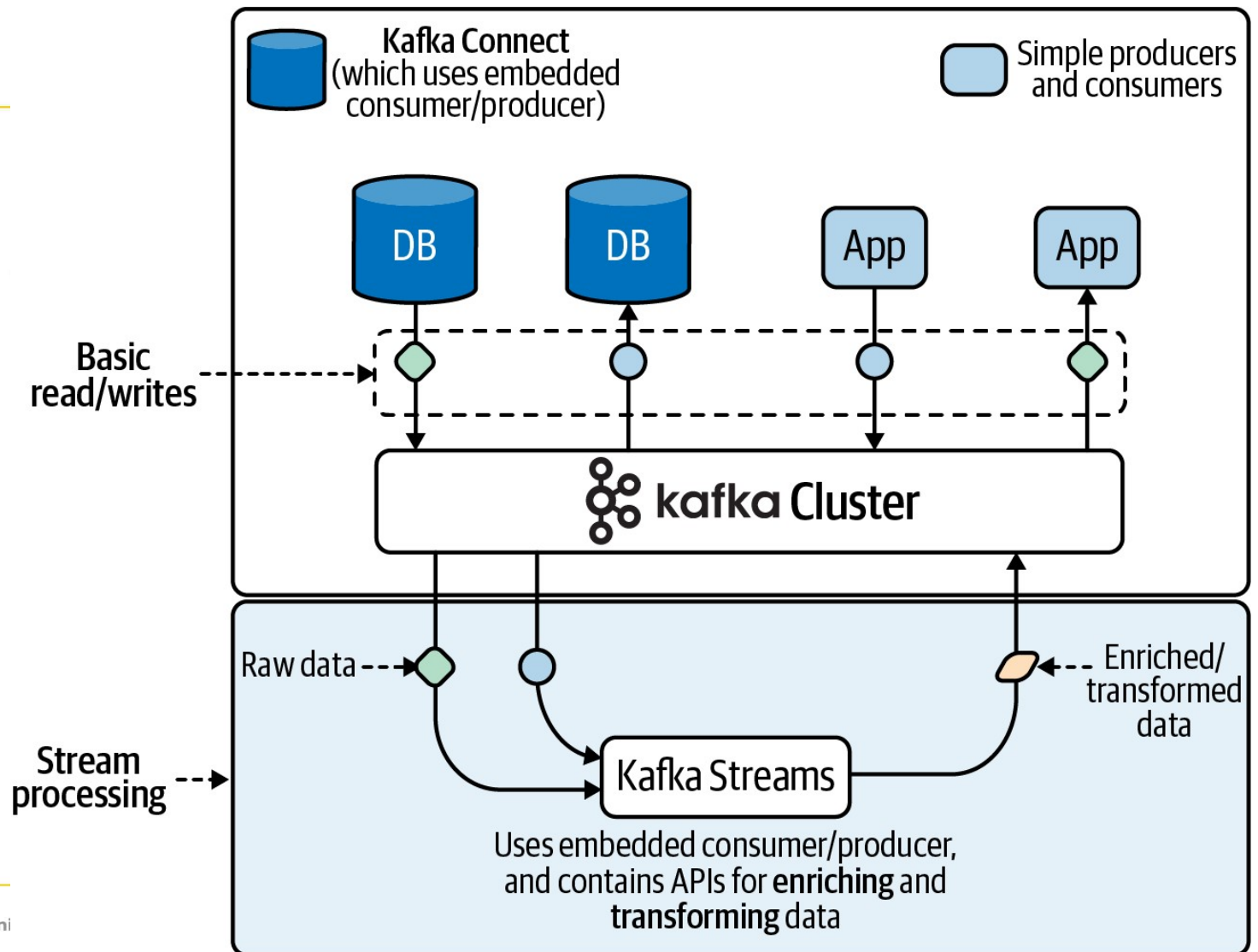
# فرآیند مصرف و دریافت پیام



حذف یک مصرف کننده



# اکوسیستم نوین کافکا



# سایر بازیگران این حوزه

	APACHE KAFKA	APACHE PULSAR	RABBITMQ (AMQP)
GENERAL			
License	Apache v2	Apache v2	Mozilla Public
Components	Kafka + Zookeeper (ZK is being removed)	Pulsar + Zookeeper + BookKeeper + RocksDB	RabbitMQ
Message consumption model	Pull	Push	Push
Storage architecture	Log	Index	Index

