|  |  |  |
| --- | --- | --- |
| **Характеристики** | **Apache Spark** | **Apache Flink** |
| Поддержка лямбда-архитектуры | + | + |
| Наличие встроенных средств для  выполнения графовых операций  машинного обучения  SQL-аналитики | GraphX  MLlib  Spark SQL | Gelly  FlinkML  Table-API |
| Гарантия строго однократной доставки сообщений (exactly once) | + | + |
| Надежность  масштабируемость  отказоустойчивость | + | + |
| Время задержки (latency) | может составлять несколько секунд | в режиме реального времени с задержкой порядка 1 миллисекунды |
| Режим работы с данными | реализует микропакетный подход (micro-batch) | обеспечивает полноценную работу как в поточном, так и в пакетном режимах с помощью разных API – для потоков DataStream API и DataSet API для пакетов |
| Зрелость, поддержка сообщества и реализация в коммерческих продуктах | более распространен среди Big Data решений и сообщества специалистов | пока характерны некоторые проблемы на уровне реализации, например, высокое потребление CPU, чувствительность к сетевым проблемам |

