Kinesis Data StreamsBrock TubreINSTRUCTOR The Kinesis FamilyKINESIS DATA STREAMS KinesisThe Kinesis Family KinesisData StreamsKinesisData FirehoseKinesis Video StreamsKinesisData AnalyticsData ProducersKinesis Data StreamsKinesis Data StreamsKinesis Data StreamsKinesis Data StreamsKinesis Data StreamsKINESIS Data Streams 1-500 shards default limitData Producerers (Processing Tools) EC2Lambda KinesisData AnalyticsEMR Kinesis Data StreamsKINESIS DATA STREAMS Shard 2Shard 3 Shard N Kinesis Streams 1-500 shards limite predefinitoProduttori di dati BI Tools Storage and Analysis Data Consumers (Processing Tools) EC2Lambda KinesisData AnalyticsEMR S3 DynamoDB RedshiftKinesis Data StreamsKINESIS DATA STREAMS Shard 2Shard 3 Shard N Kinesis Streams 1-500 shards limite predefinitoProduttori di datiKinesis StreamsShard 1Partition Key: e55fab...Sequenza: 1 Data: <some fancy data>Partition Key: e55fab... Sequence: 2 Data: <some fancy data>Shard 2Partition Key: 45ab15.... Sequence: 1 Data: <some fancy fancy Data>Partition Key: 45ab15...Sequenza: 2 Data: <some fancy data>Kinesis Data StreamsKINESIS DATA STREAMSData ProductoriKinesis Data StreamsKINESIS DATA STREAMS Train Id = Partition KeyTrain Car = Sequence numberPassengers = Data •Each shard consists of a sequence of data records.These can be ingerited at 1000 records for second. •Default limit of 500 shards, but you can request increase to unlimited shards. •A data record is the unit of data captured.•sequence number•partition key•data blob (your payllow, up Kinesis Data StreamsKINESIS Data StreamsInteragire con Kinesis Data StreamsKINESIS Data StreamsKINESIS Data Streams Kinesis Data Streams Kinesis Client Library (KCL)Integrato direttamente con KPL per applicazioni consumer per consumare ed elaborare dati da Kinesis Data Stream. 2Kinesis API (AWS SDK) Utilizzato per operazioni API di basso livello per inviare registrazioni a un Kinesis Data Stream. 3 1Kinesis Producer Library (KPL) Facile da usare libreria che consente di scrivere a un Kin esis Data Stream. KPL vs APIKINESIS DATA STREAMSKinesis Producer Library (KPL) • Fornisce uno strato di astrazione specificamente per l'ingestione dei dati• Meccanismo automatico e configurabile di retry• Ulteriori ritardi di elaborazione possono verificarsi per maggiori efficienze di imballaggio e migliori prestazioni• Java wrapperKinesis API• Chiamate API di basso livello (PutRecords and GetRecords)• Creazioni di stream, reharding e messa e ottenere record vengono gestiti manualmente • Nessun ritardo nel trattamento• Qualsiasi AWS SDKKinesis Data StreamsKINESIS DATA STREams AMS Kinesis Data StreamsKINESIS DATA STREAMS Quando si devono utilizzare Kinesis Data Streams? • I bisogni di essere trattati dai consumatori.•Real time analytics.•Feed in altri servizi in tempo reale.•Alcune azioni devono verificarsi sui vostri dati.•La conservazione dei dati è facoltativa.•La conservazione dei dati è importante. Kinesis Data StreamsKinesis Data StreamsKINESIS DATA STREAMSKinesis Data Streams - Use CasesKINESIS DATA STREAMS •Processare e valutare i log immediatamente Esempio: Analizzare i log di sistema e applicazione continuamente e procedere all'interno del sec. Onds.•Real-time data analytics Esempio: Eseguire analisi in tempo reale su click stream data ed elaborarlo in pochi secondi.