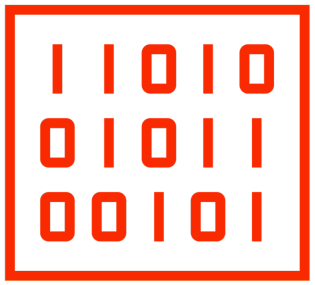
Exploring AVRO

Bogdan Sucaciu
SOFTWARE ENGINEER

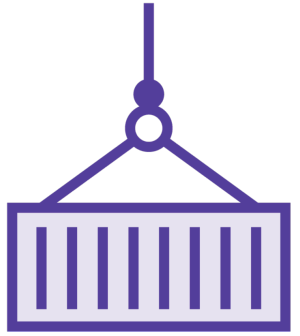
@BSucaciu bsucaciu.com

AVRO Features

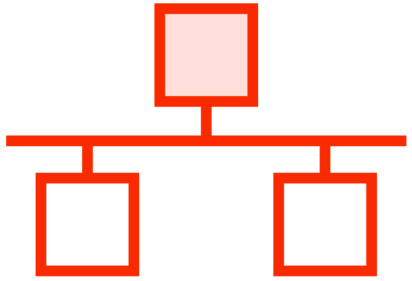
Rich Data Structures



Compact Binary Data



Container File



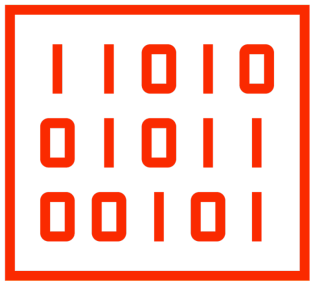
RPC

Integration

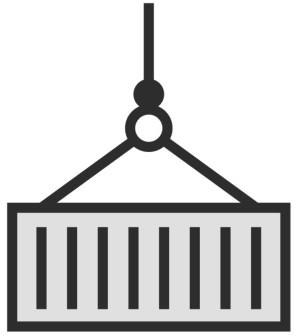
Schemas

AVRO Features

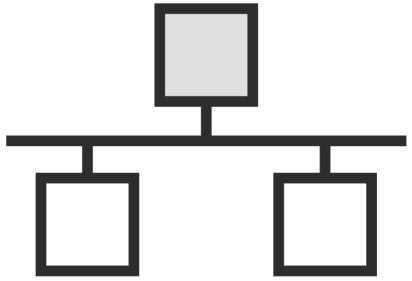
Rich Data Structures



Compact Binary Data



Container File



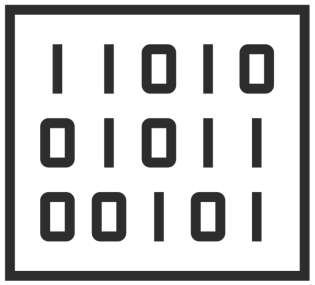
RPC

Integration

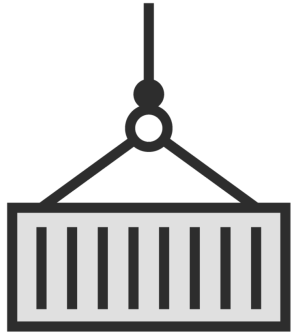
Schemas

AVRO Features

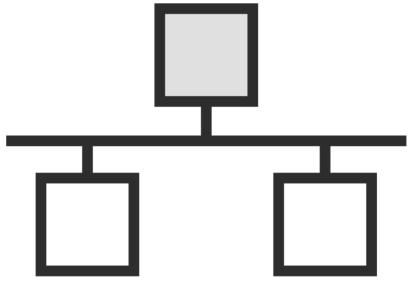
Rich Data Structures



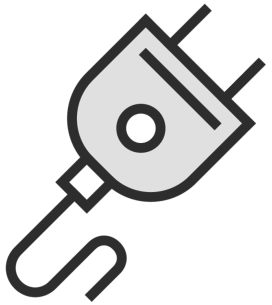
Compact Binary Data



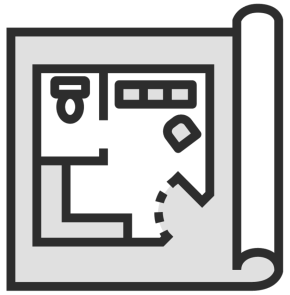
Container File



RPC

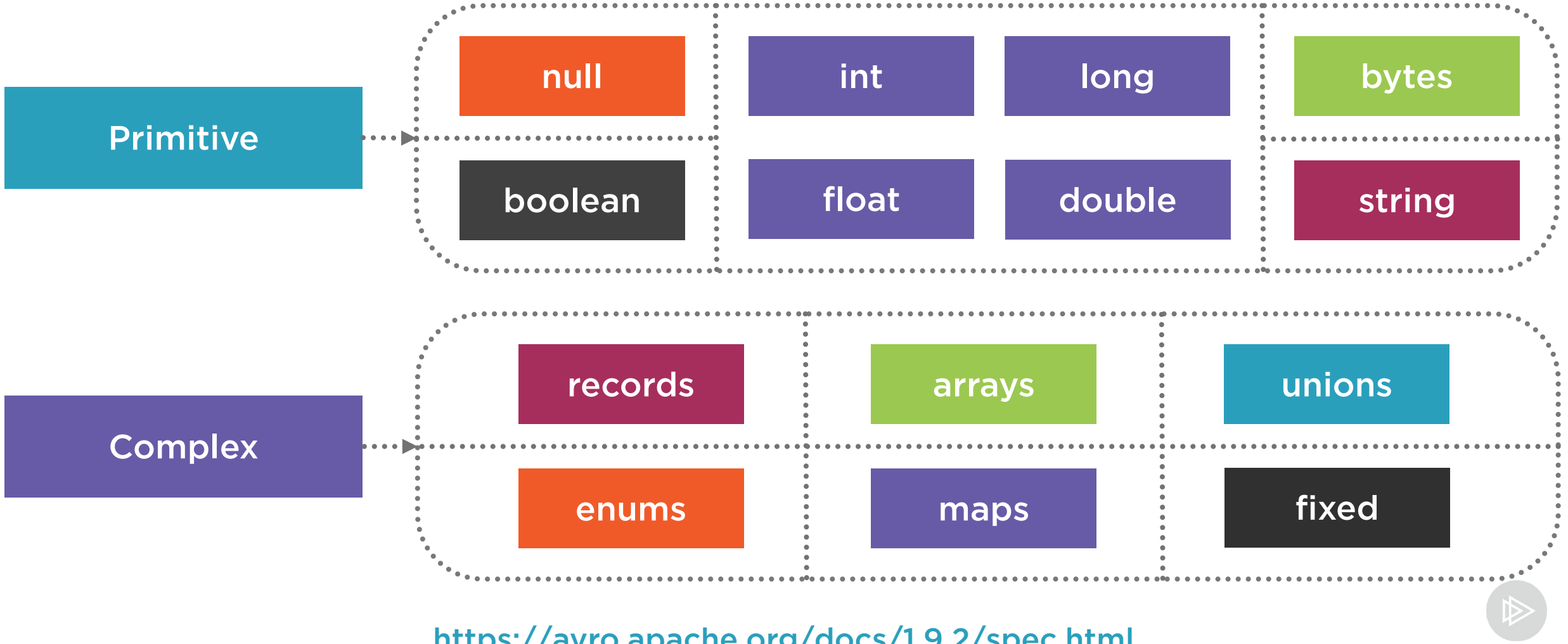


Integration



Schemas

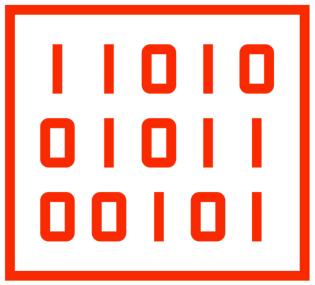
Rich Data Structures



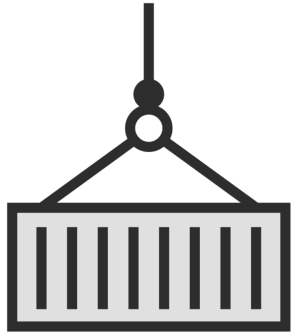
https://avro.apache.org/docs/1.9.2/spec.html

AVRO Features

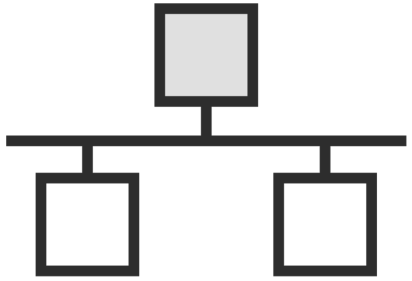
Rich Data Structures



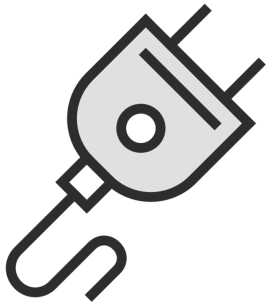
Compact Binary Data



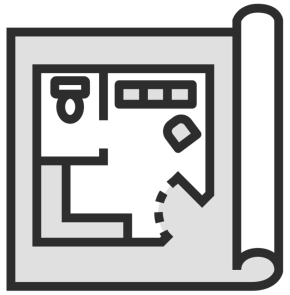
Container File



RPC

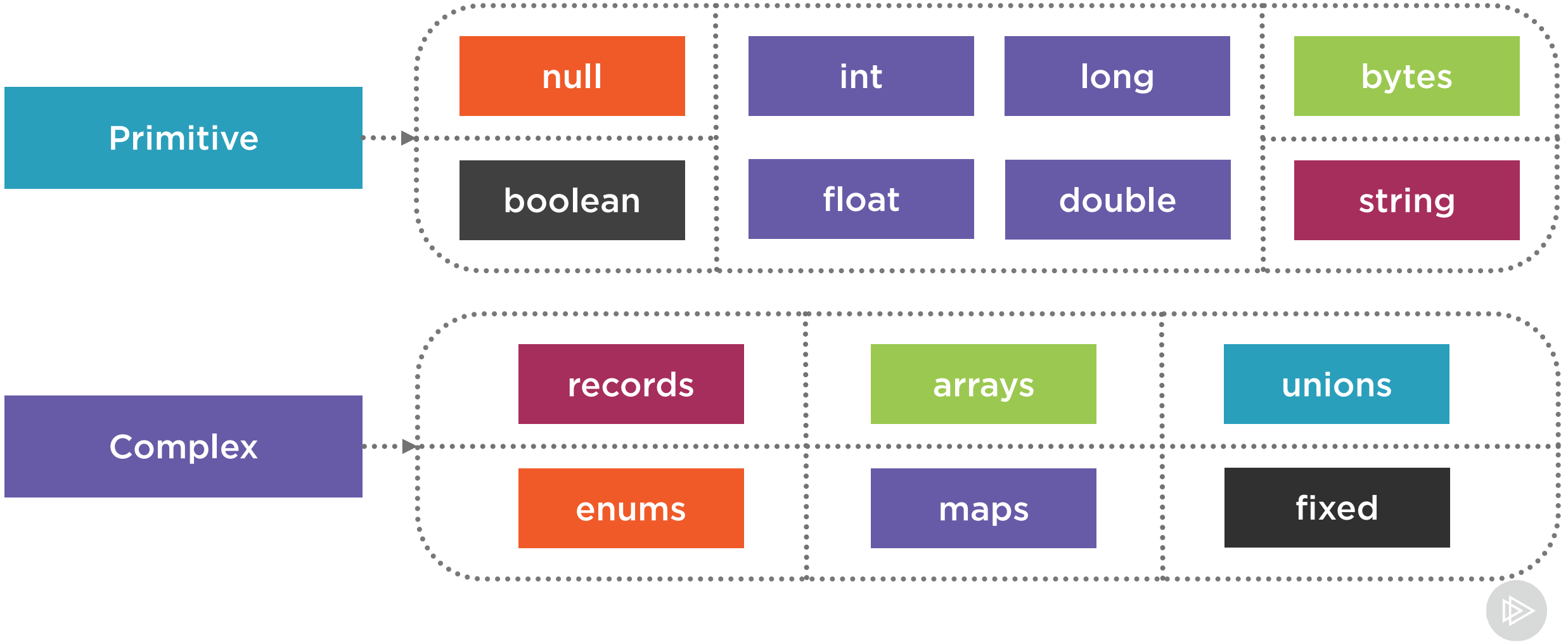


Integration

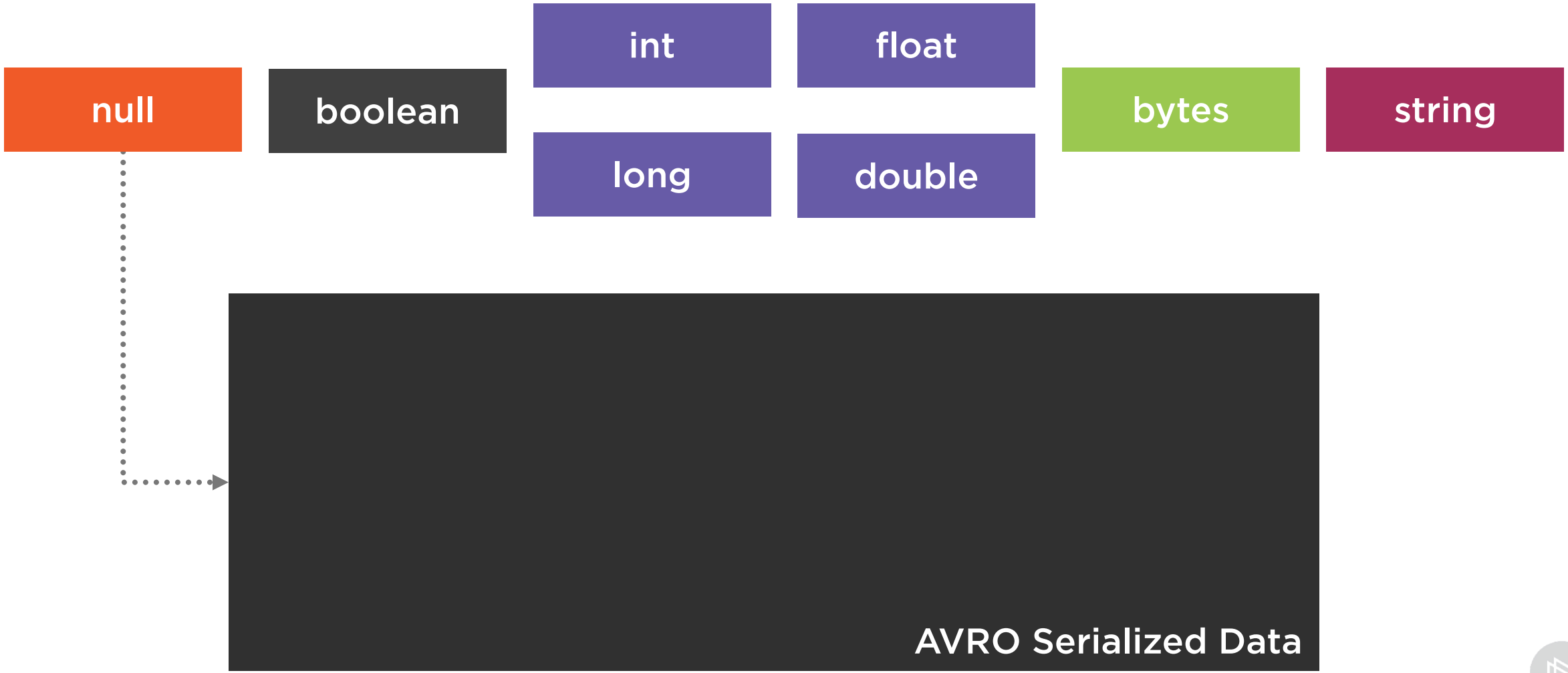


Schemas

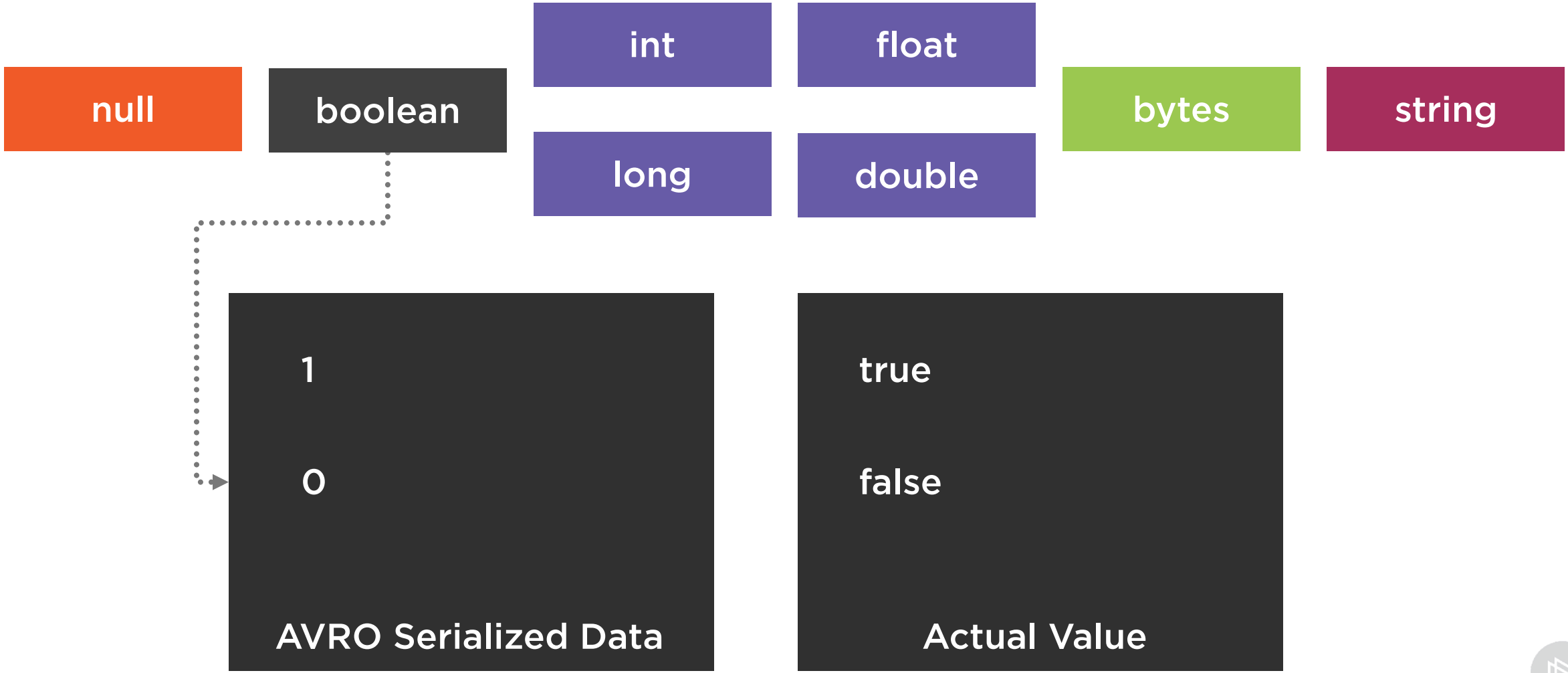
Compact Binary Data



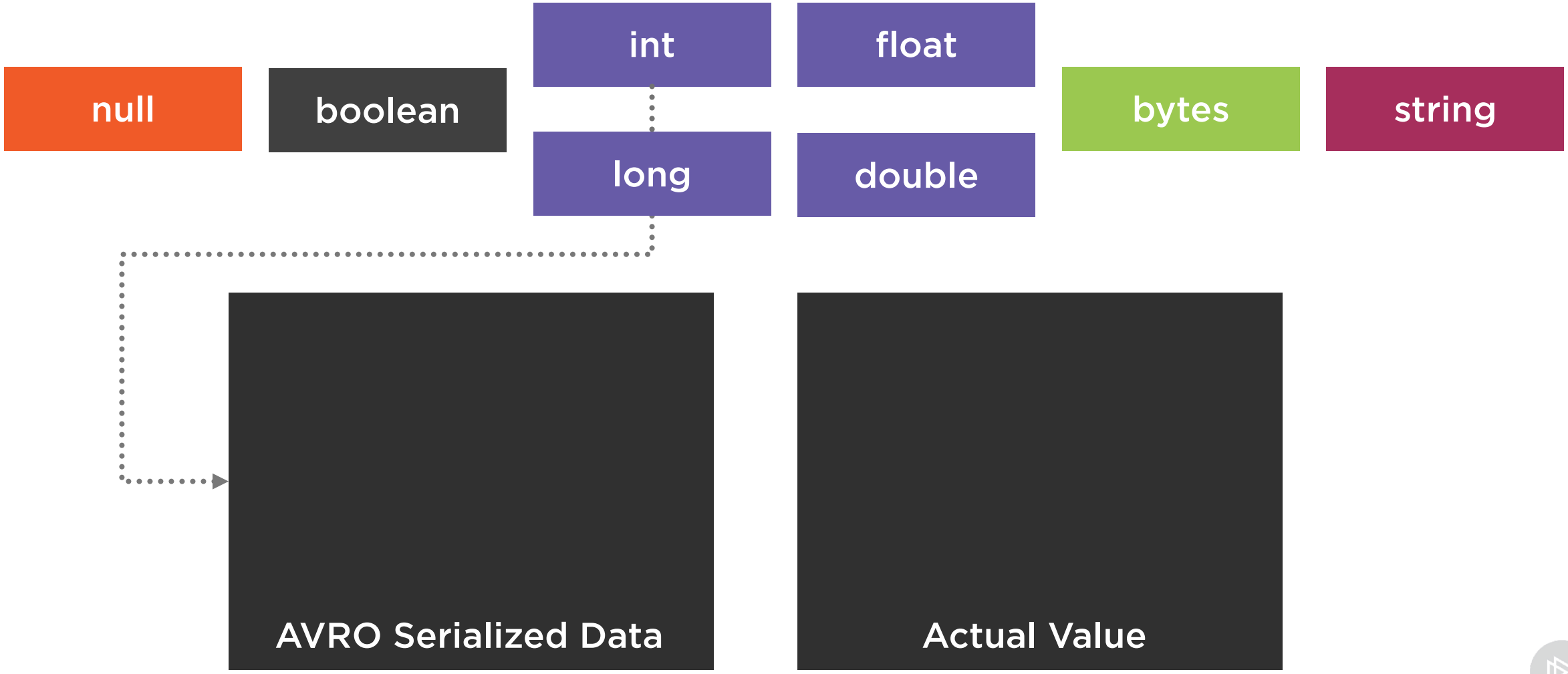
Compact Binary Data

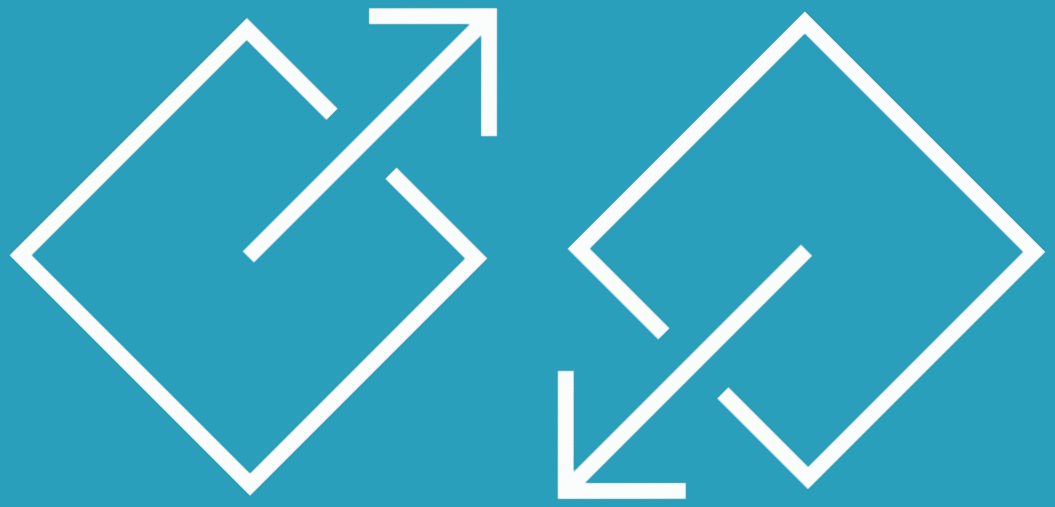


Compact Binary Data



Compact Binary Data

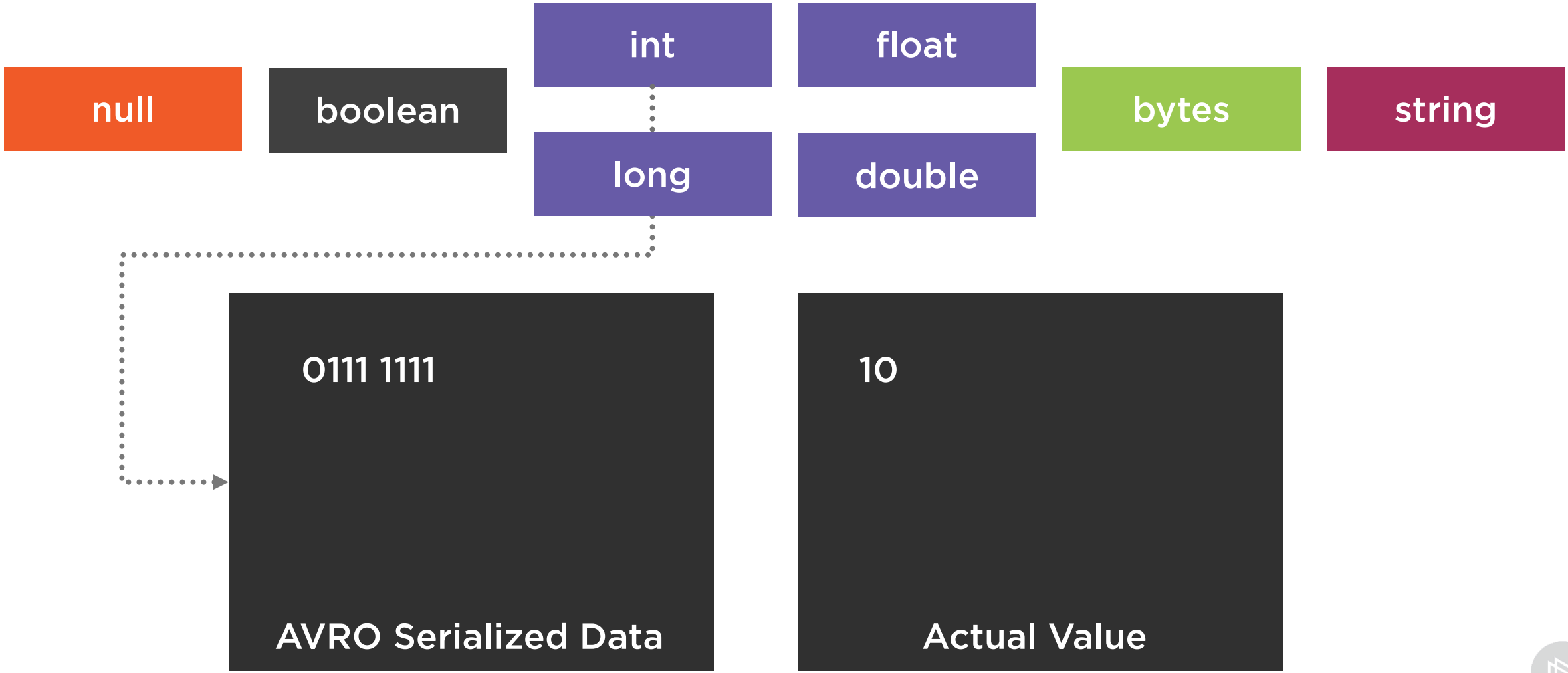




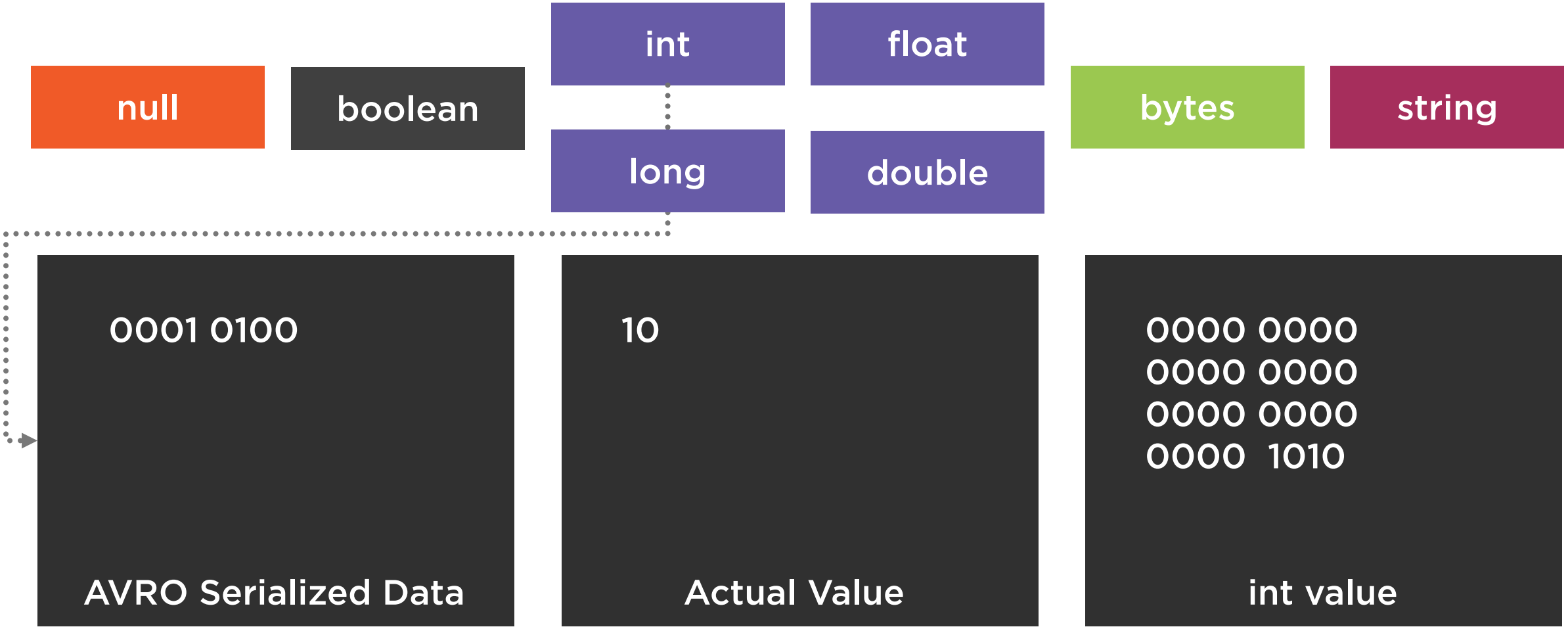
Variable-length zig-zag encoding

Value	Hex
0	00
-1	01
1	02
-2	03
2	04
...	...

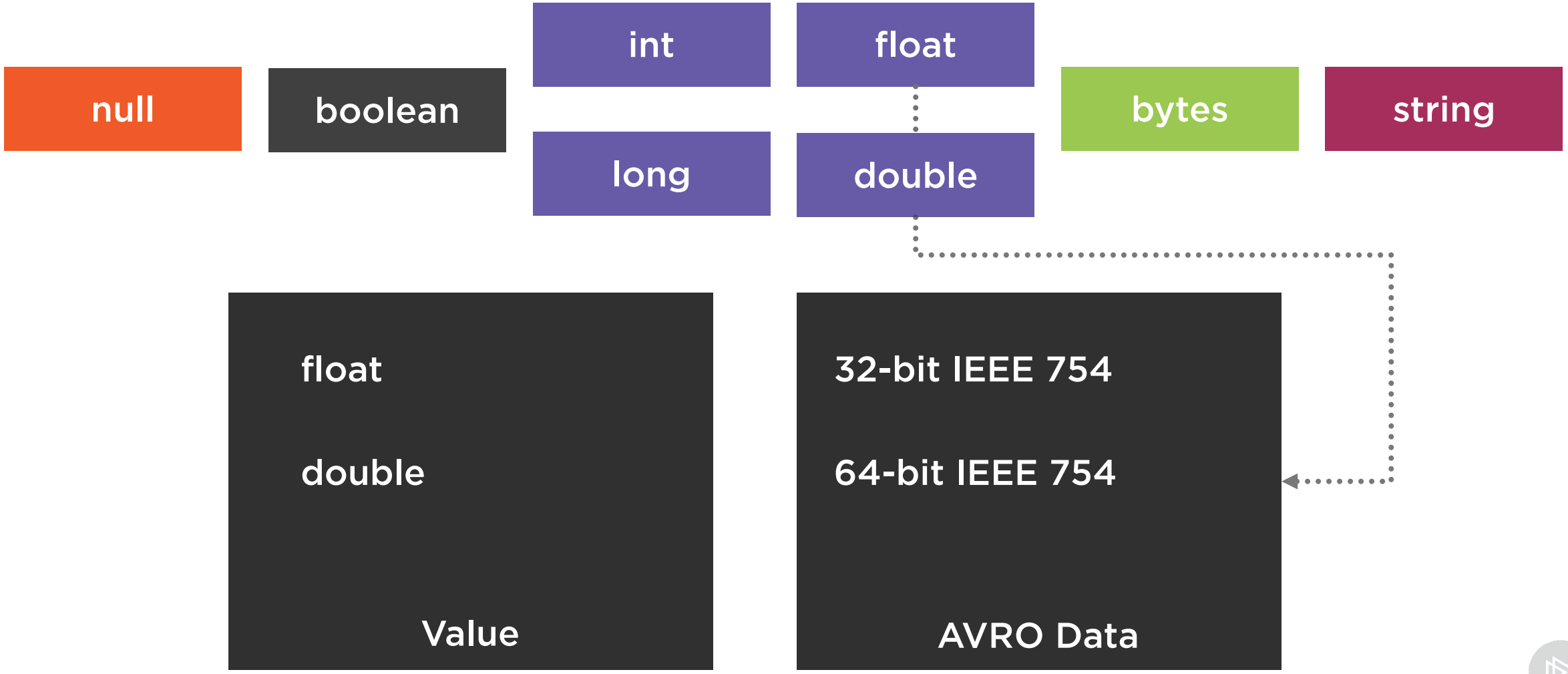
Compact Binary Data



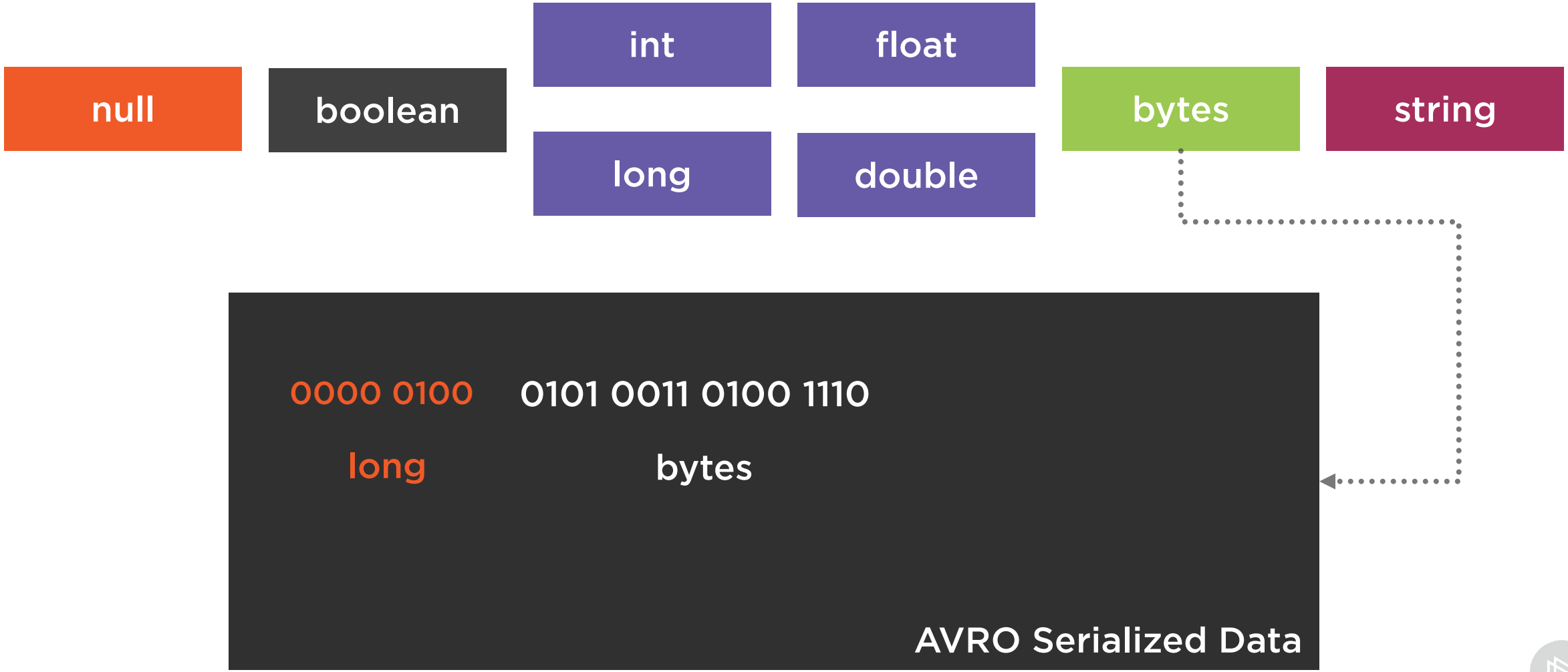
Compact Binary Data



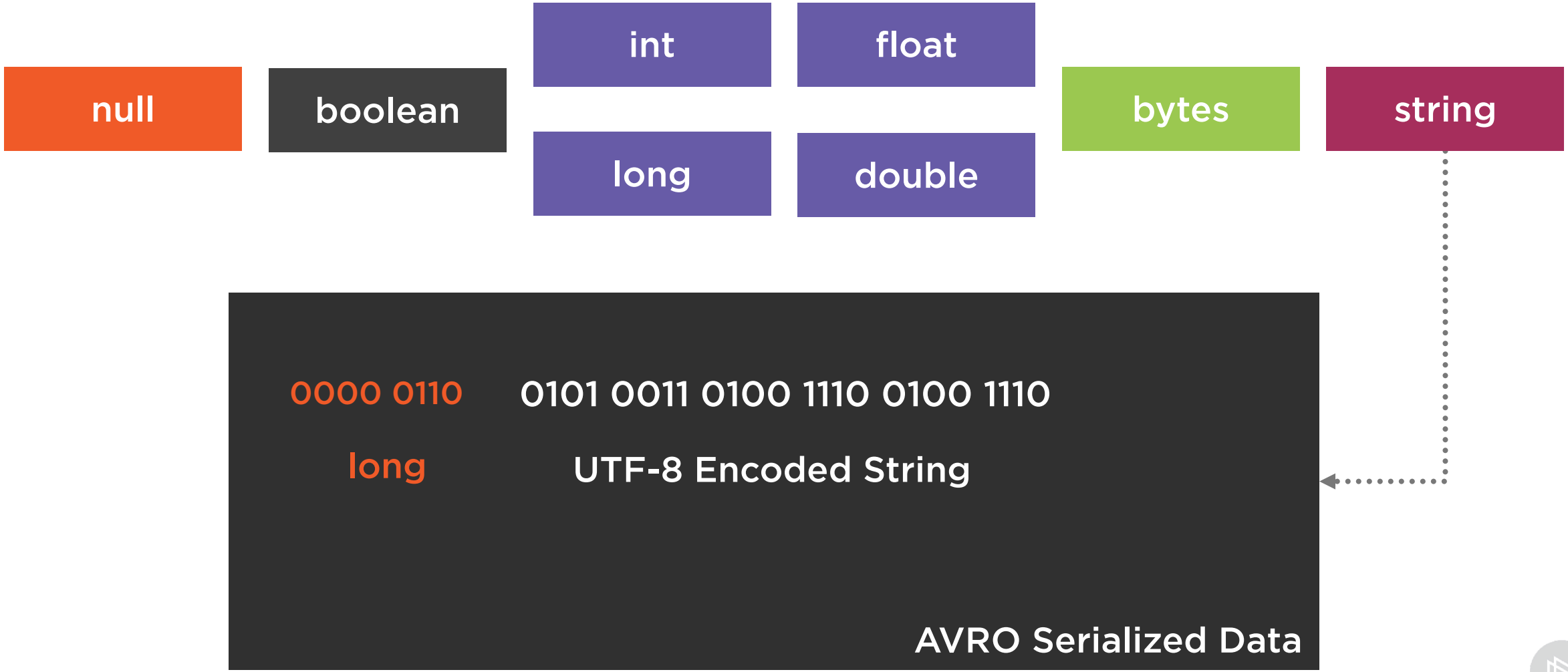
Compact Binary Data



Compact Binary Data

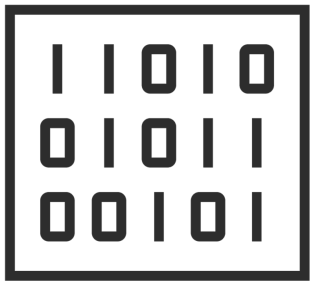


Compact Binary Data

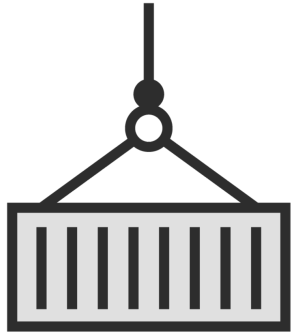


AVRO Features

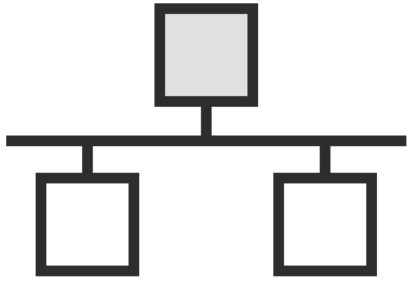
Rich Data Structures



Compact Binary Data

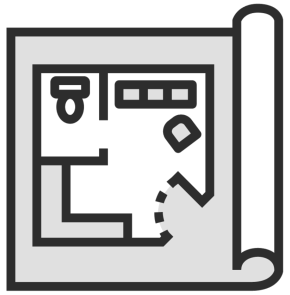


Container File



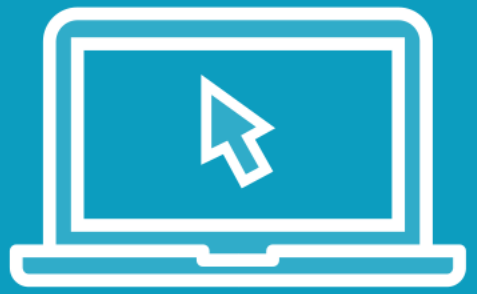
RPC

Integration



Schemas

Demo



Java integration

Set up Schema Repository

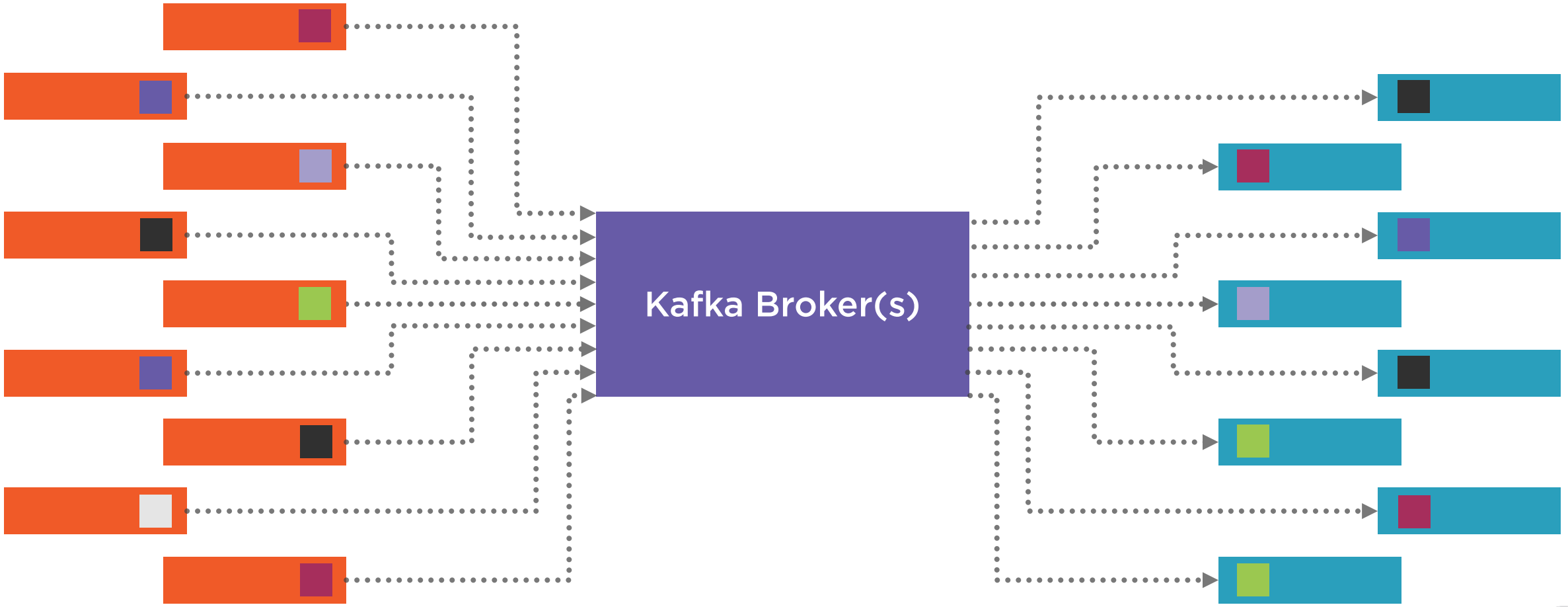
Kafka-powered System

Microservices

Data Pipelines

Kafka-powered System

Kafka-powered System



Centralized Schema Repository

Schemas

AVRO Schema

Schema.avsc

```
{
  "type": "string"
}
```

AVRO Schema

Schema.avsc

```
{
  "name": ”Schema",
  "namespace": "com.pluralsight",
  …
}
```

AVRO Schema

Schema.avsc

```
{
  "name": ”Schema",
  "namespace": "com.pluralsight",
  "type": ”record”,
  ...
}
```

AVRO Schema

Schema.avsc

```
{
  "name": ”Schema",
  "namespace": "com.pluralsight",
  "type": "record",
  "fields": [
  ...
  ]
}
```

AVRO Schema

Schema.avsc

```
{
  "name": ”Schema",
  "namespace": "com.pluralsight",
  "type": "record",
  "fields": [
    {
      "name": ”age",
      "type": "int"
    },
    ...
  ]
}
```

AVRO Schema

Schema.avsc

```
{
  "name": ”Schema",
  "namespace": "com.pluralsight",
  "type": "record",
  "fields": [
    {
      "name": ”address",
      "type": "record",
      "fields": [
        ...
      ]
    },
    ...
  ]
}
```

AVRO Schema

Schema.avsc

```
{
  "name": ”Schema",
  "namespace": "com.pluralsight",
  "type": "record",
  "fields": [
    {
      "name": "daysOfTheWeek",
      "type": "enum",
      "symbols": ["MON", ”TUE", ”WEN", ”THU", ”FRI", ”SAT", ”SUN"]
    },
    ...
  ]
}
```

AVRO Schema

Schema.avsc

```
{
  "name": ”Schema",
  "namespace": "com.pluralsight",
  "type": "record",
  "fields": [
    {
      "name": ”middleName",
      "type": ["null", "string"]
    },
    ...
  ]
}
```

AVRO Schema

Schema.avsc

```
{
  "name": ”Schema",
  "namespace": "com.pluralsight",
  "type": "record",
  "fields": [
    {
      "name": "md5",
      "type": "fixed",
      "size": 16
    },
    ...
  ]
}
```

Demo

AVRO schemas

Produce / Consume using AVRO

Summary

AVRO features

Building schemas