



Exploring Compression Techniques for ROOT IO

Zhe Zhang
University of Nebraska-Lincoln
zhan0915@huskers.unl.edu



Agenda

- Introduction
- Motivation
- Approaches
- Results
- Conclusion



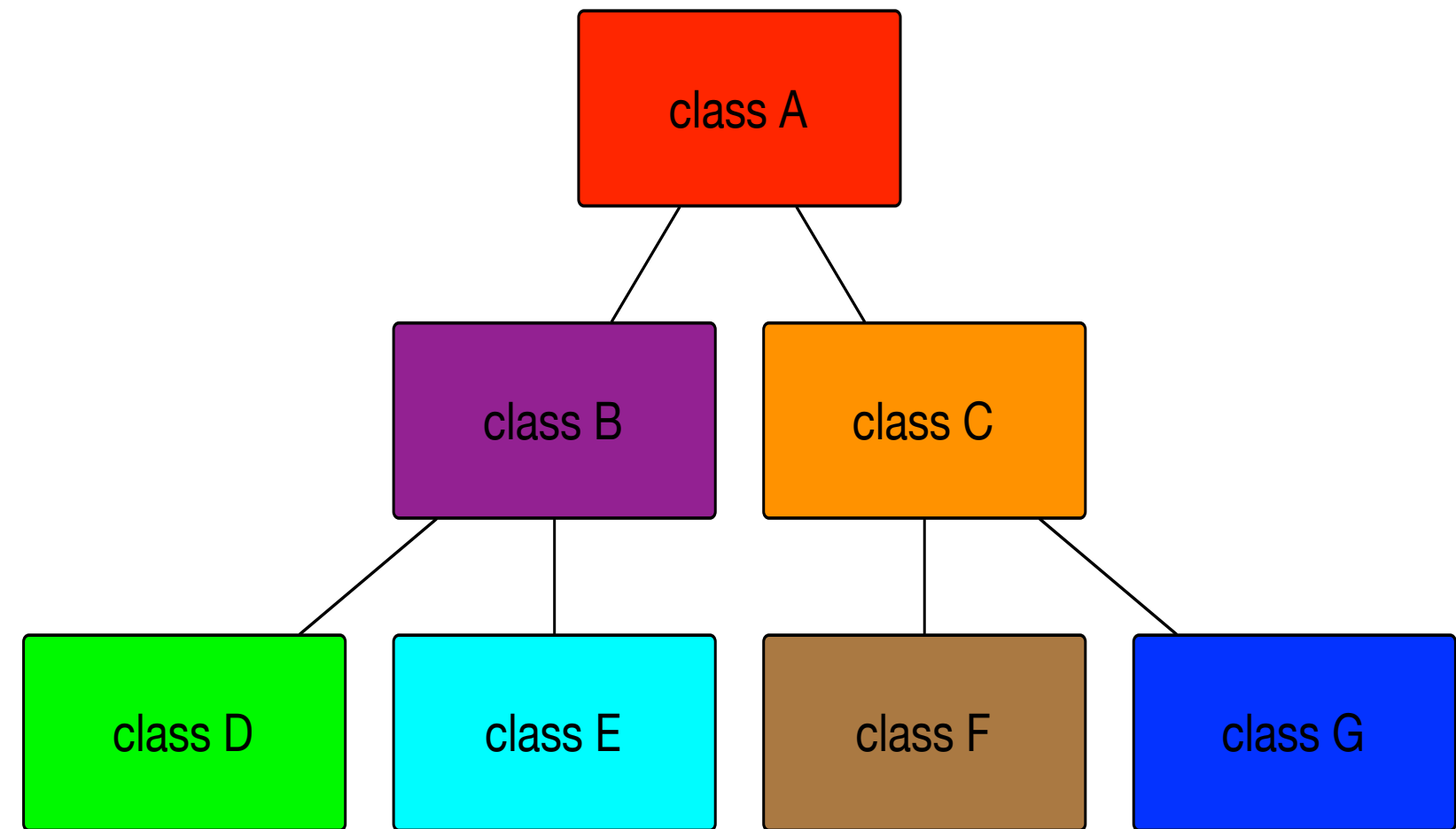
Introduction

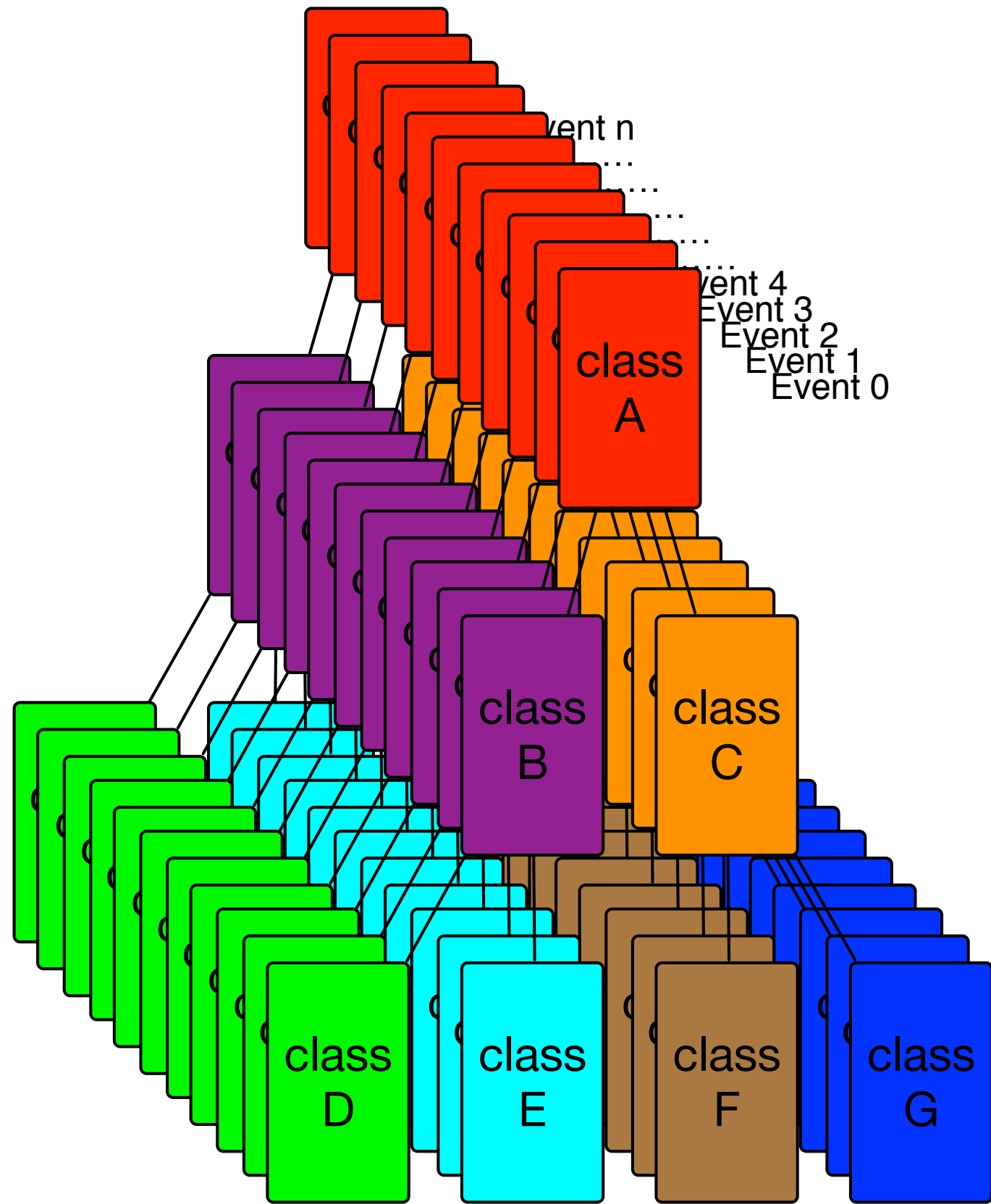
- ROOT uses tree to store large amount of objects.
- Trees allow sequential and random access to any entries.
- Trees have branches which behaves as column oriented database.



Event Structure:

```
class A {  
  class B {  
    class D;  
    class E;  
  };  
  class C {  
    class F;  
    class G;  
  };  
};
```





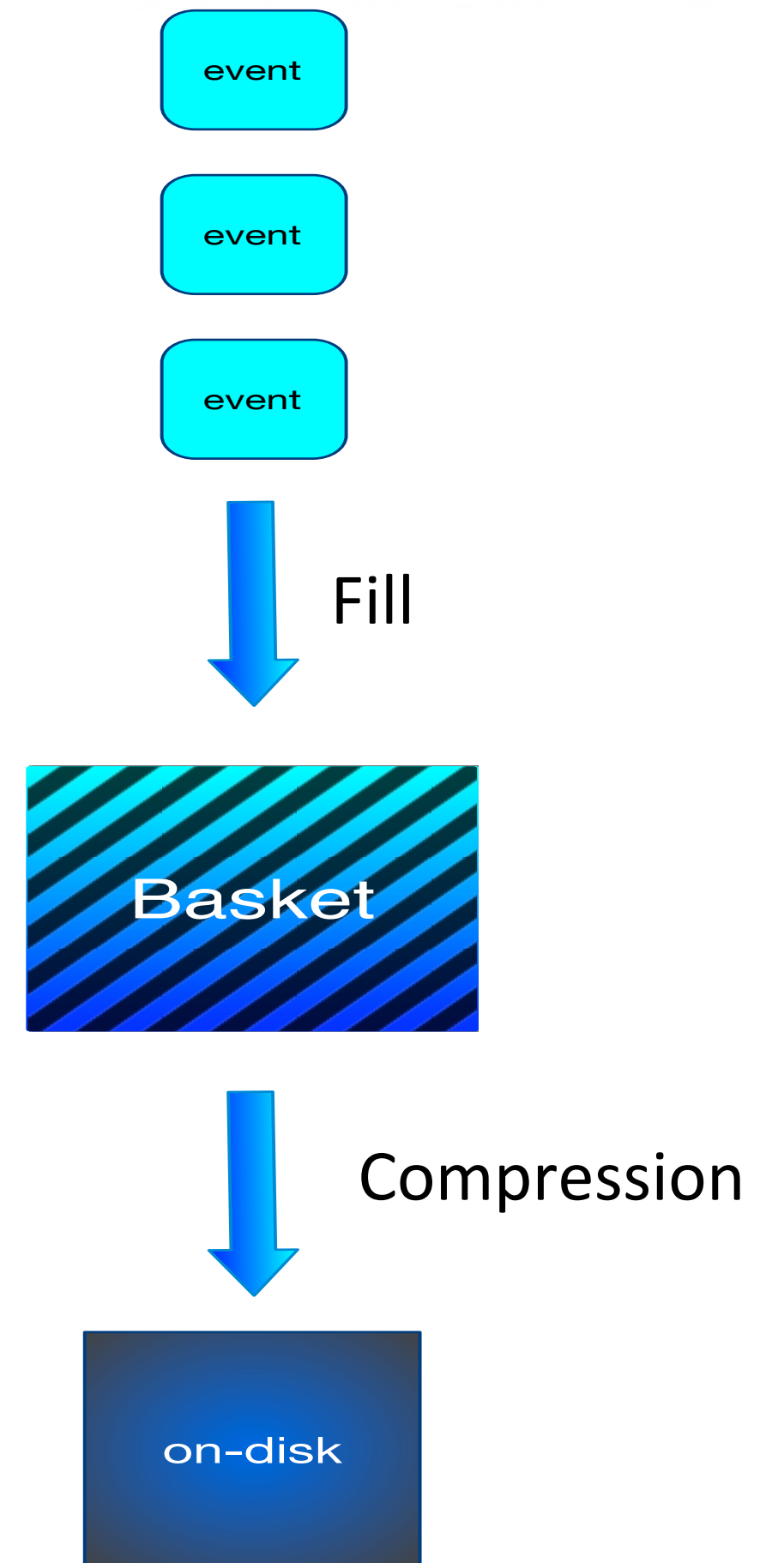
Trees are serialized by same branches:

- Easy comparison within the same branch.
- Redundant data are close together.





- Each branch has a basket.
- Each basket is filled with events.
- Each basket is compressed and flush to disk once it is full.





Motivation

- Why alternative compression algorithms ?
- Does ROOT compression compare well to naïve approach ?
- How random reads perform ?