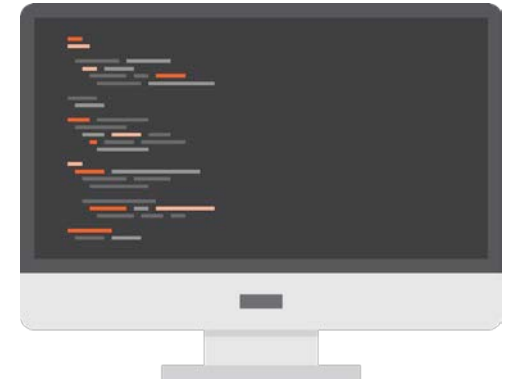
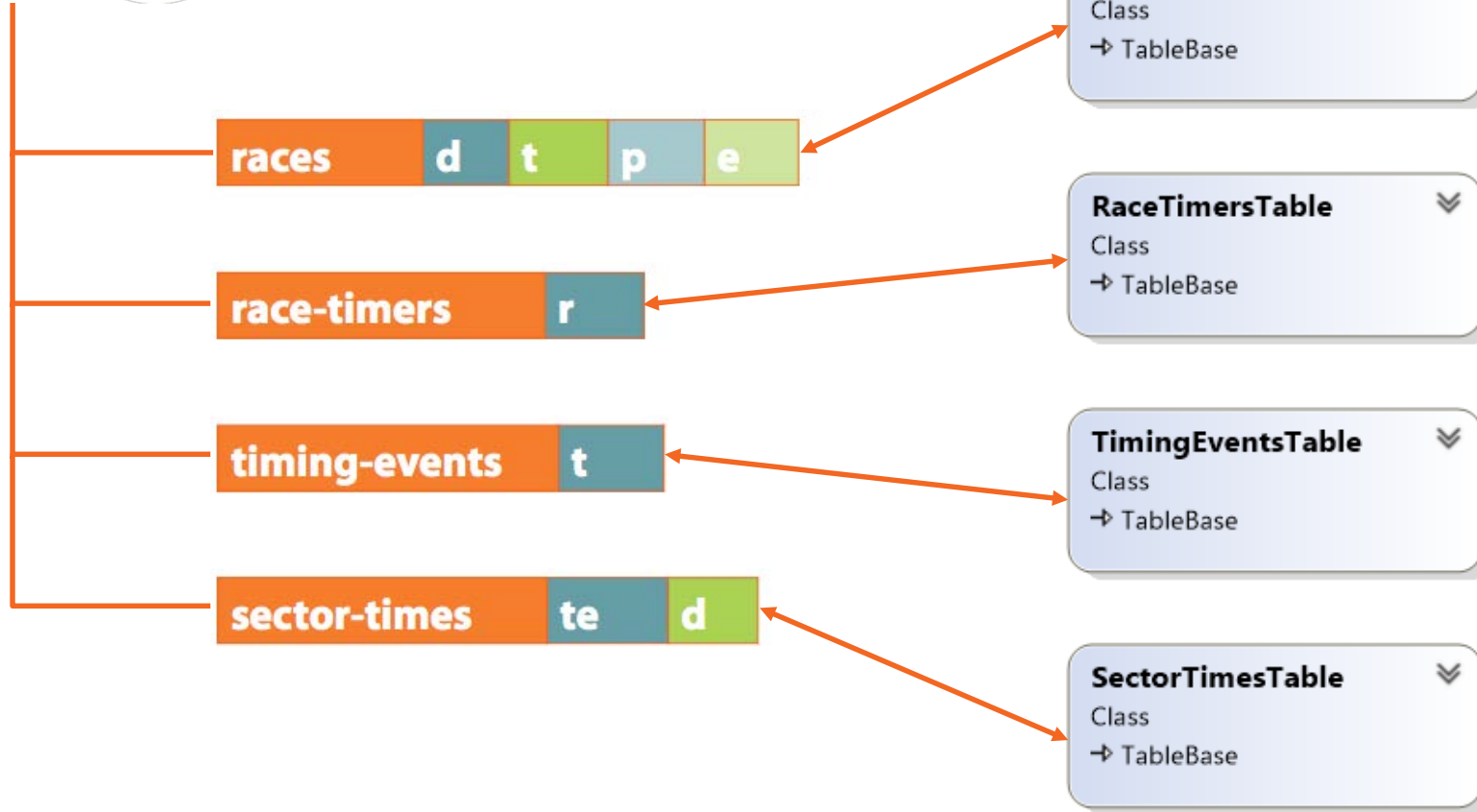


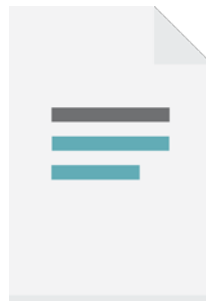
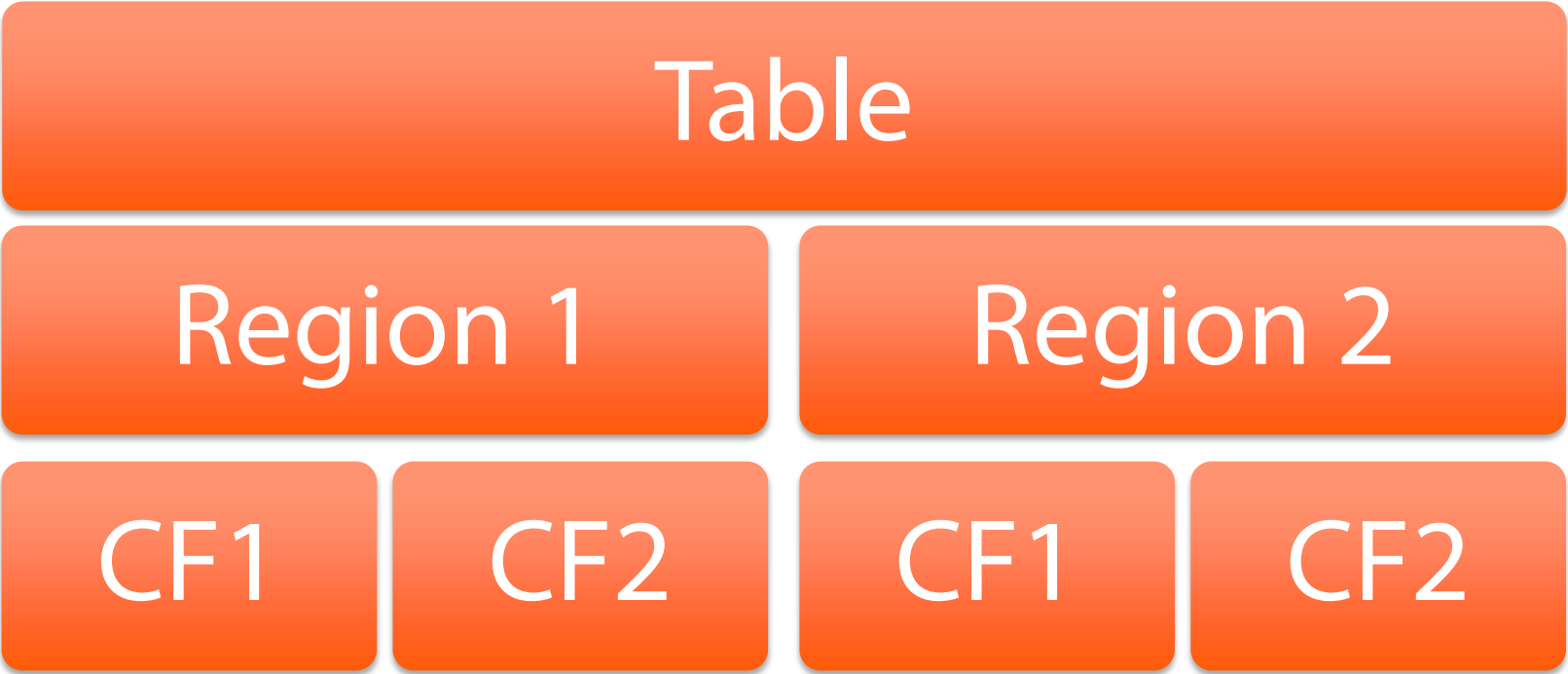
HBase Deep Dive



Elton Stoneman

@EltonStoneman | blog.sixeyed.com



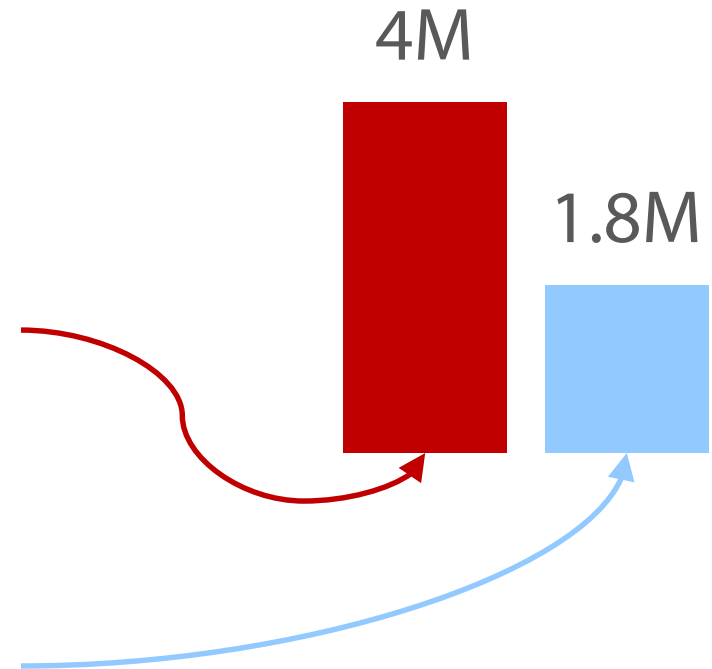




Node	Regions	Req/sec
workernode0	3	630
workernode1	3	1,205
workernode2	3	685
workernode3	3	403



APACHE
HBASE



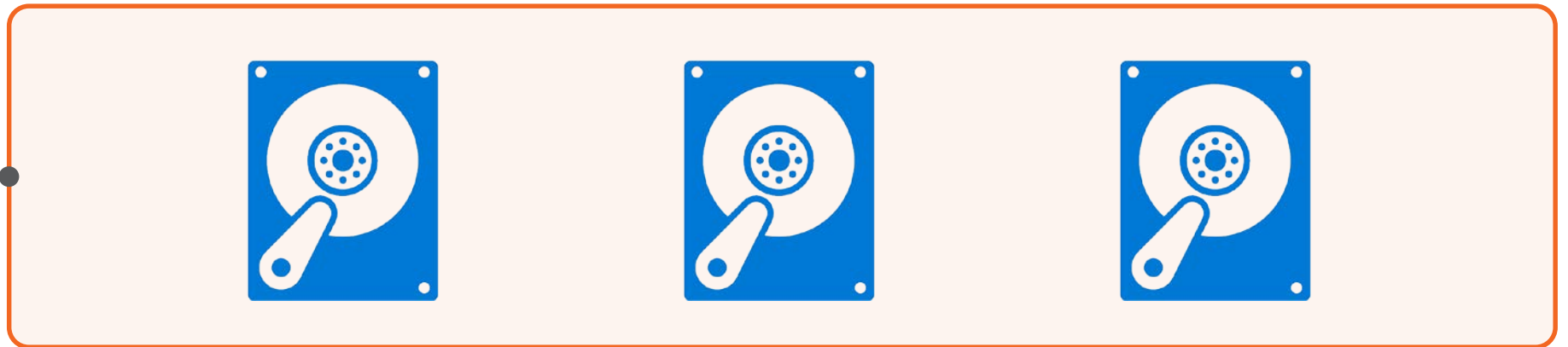
HFiles



HDFS



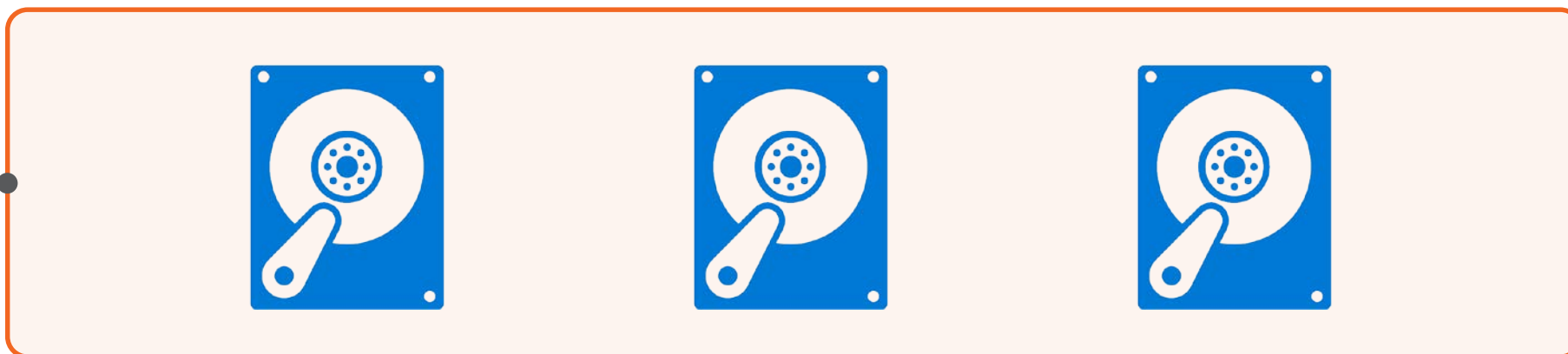
Disk



HFiles



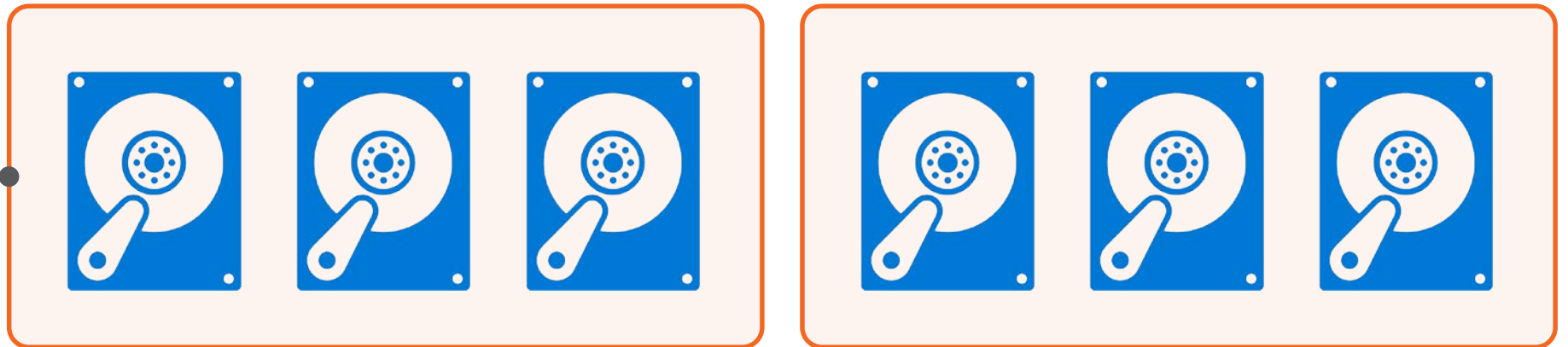
DC



HFiles



DC x2



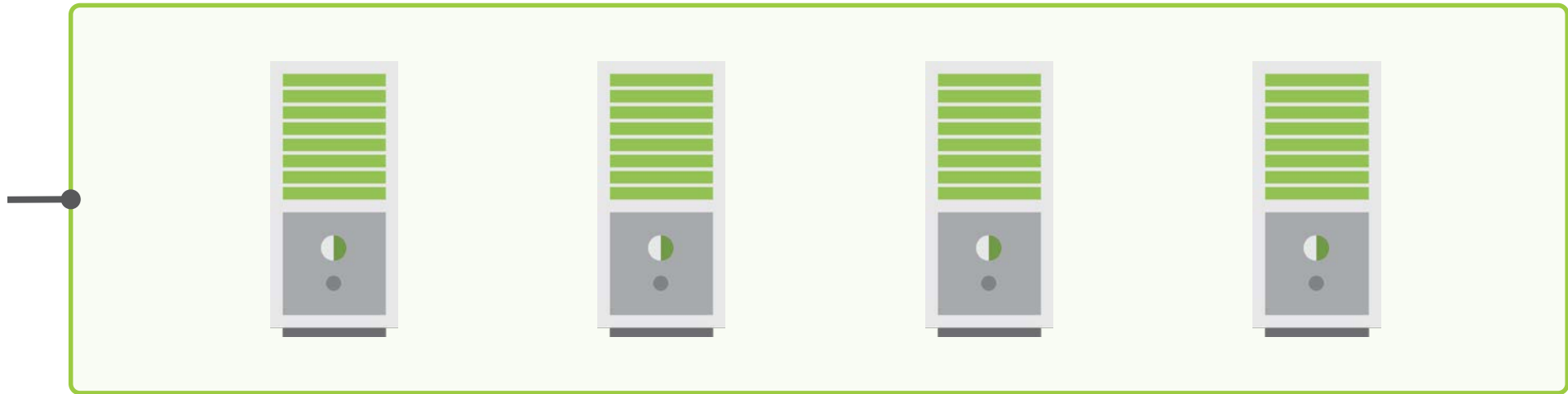
Region
Servers



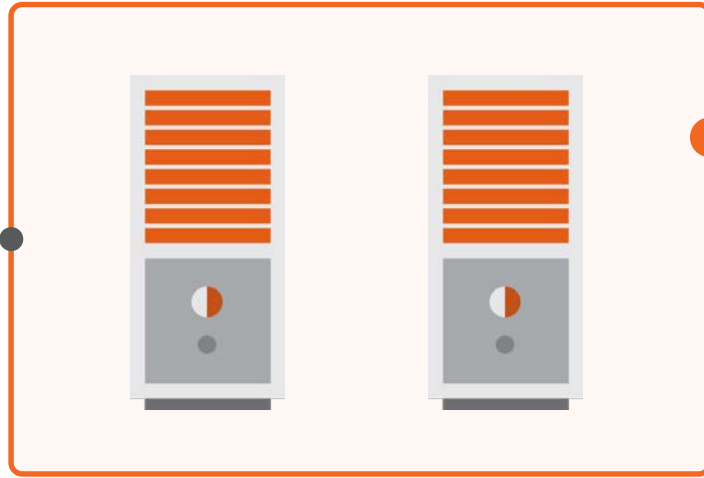
HFiles



Region
Servers

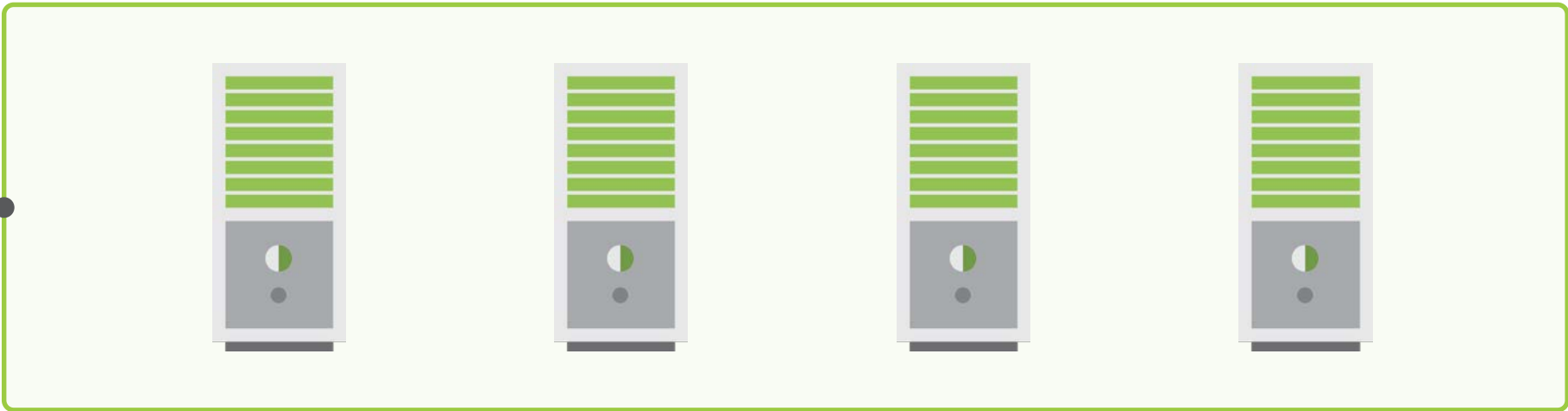


Master
Server(s)

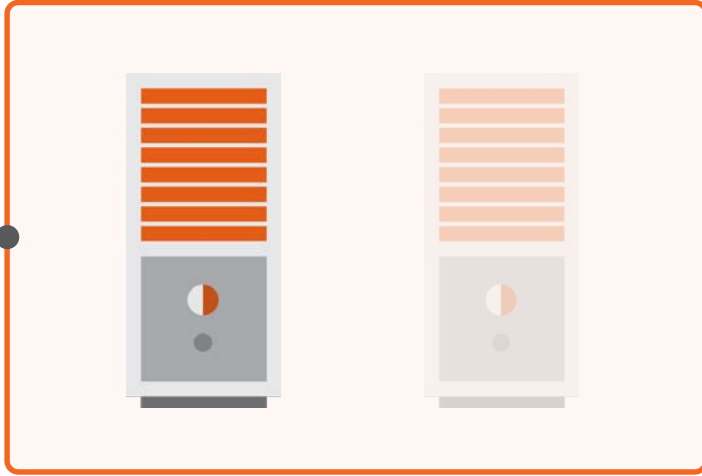


Schema changes
Cluster **management**
Data administration

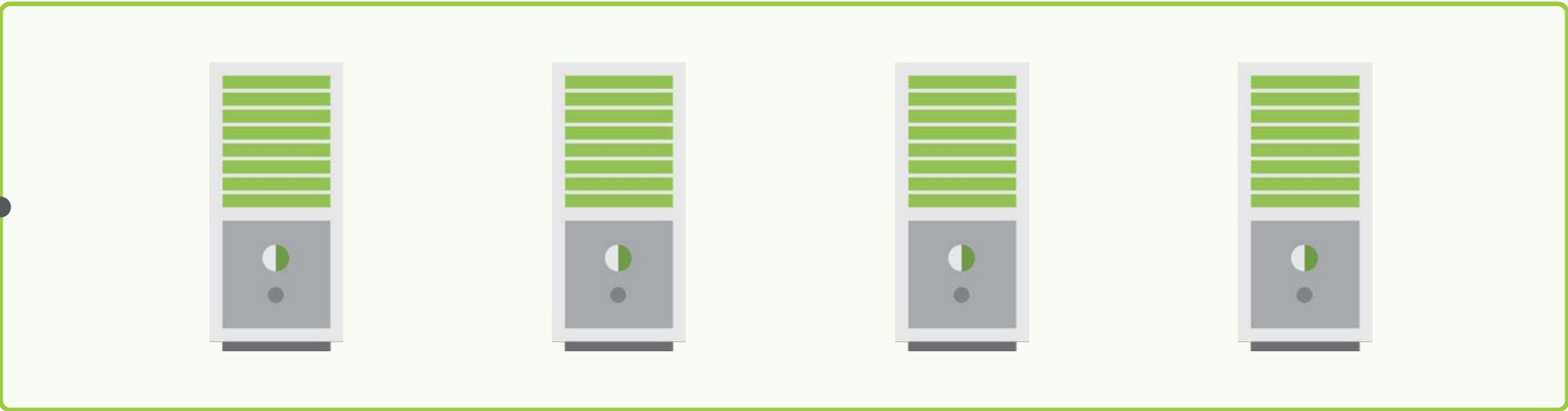
Region
Servers

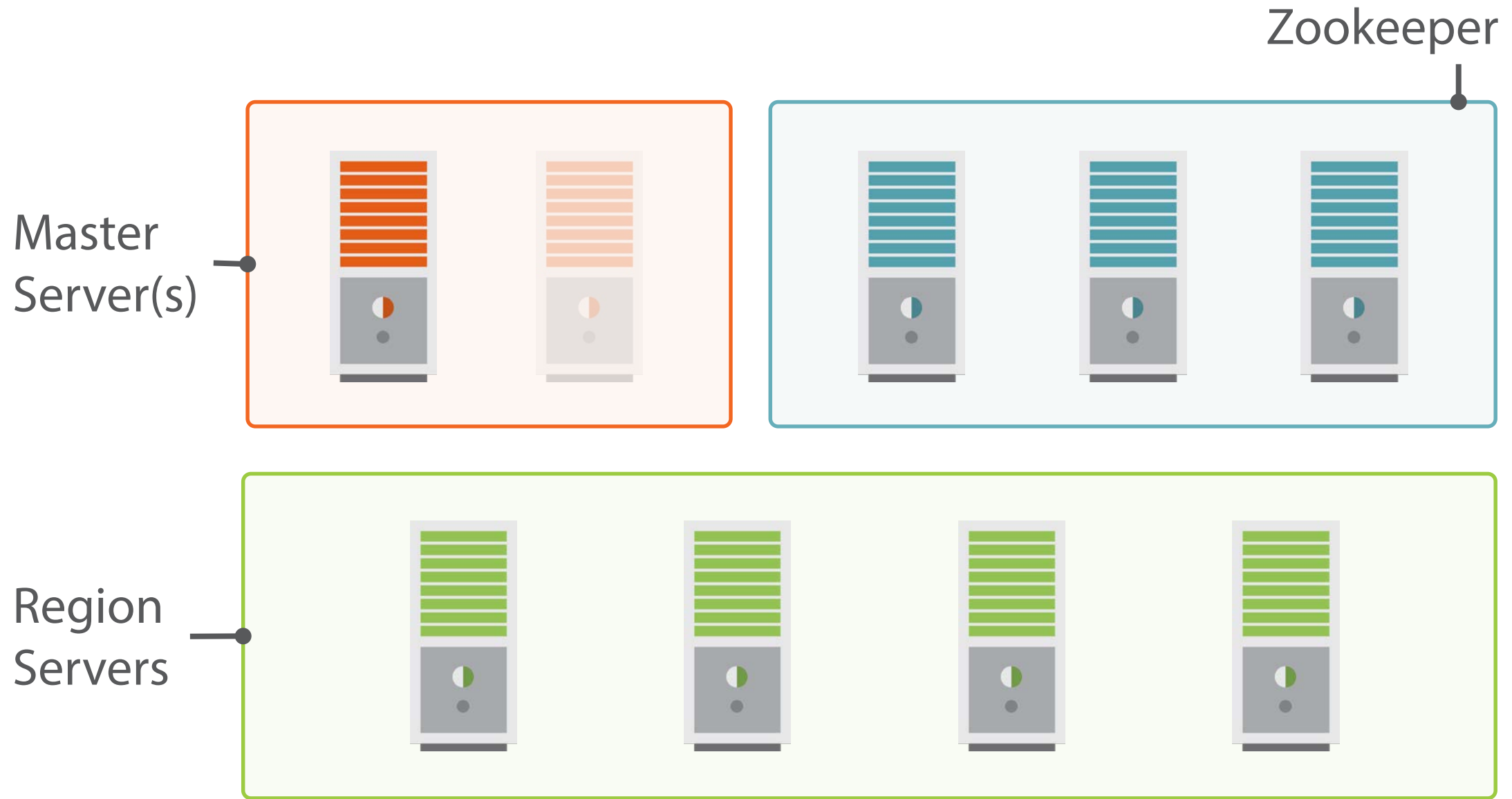


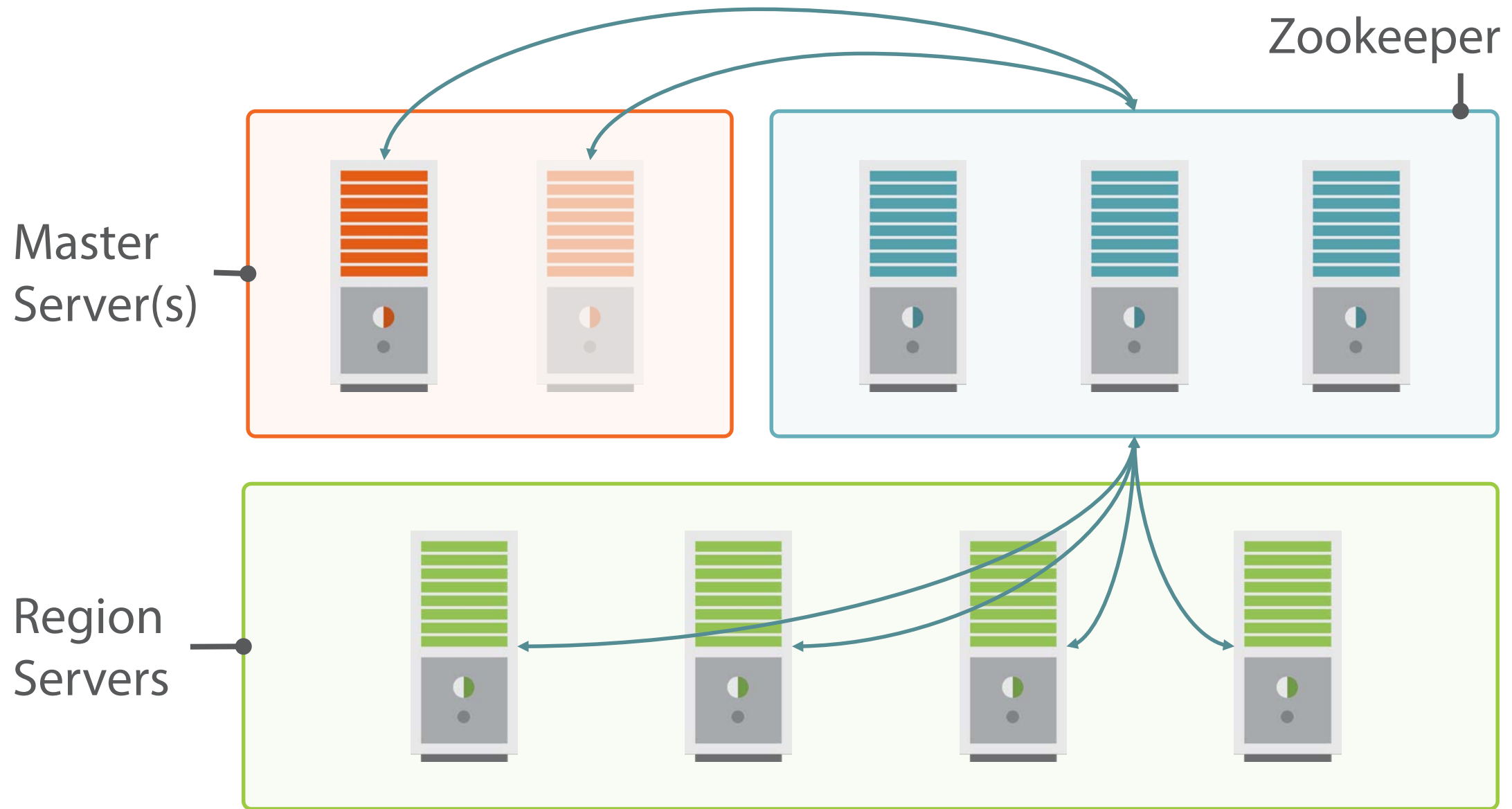
Master
Server(s)

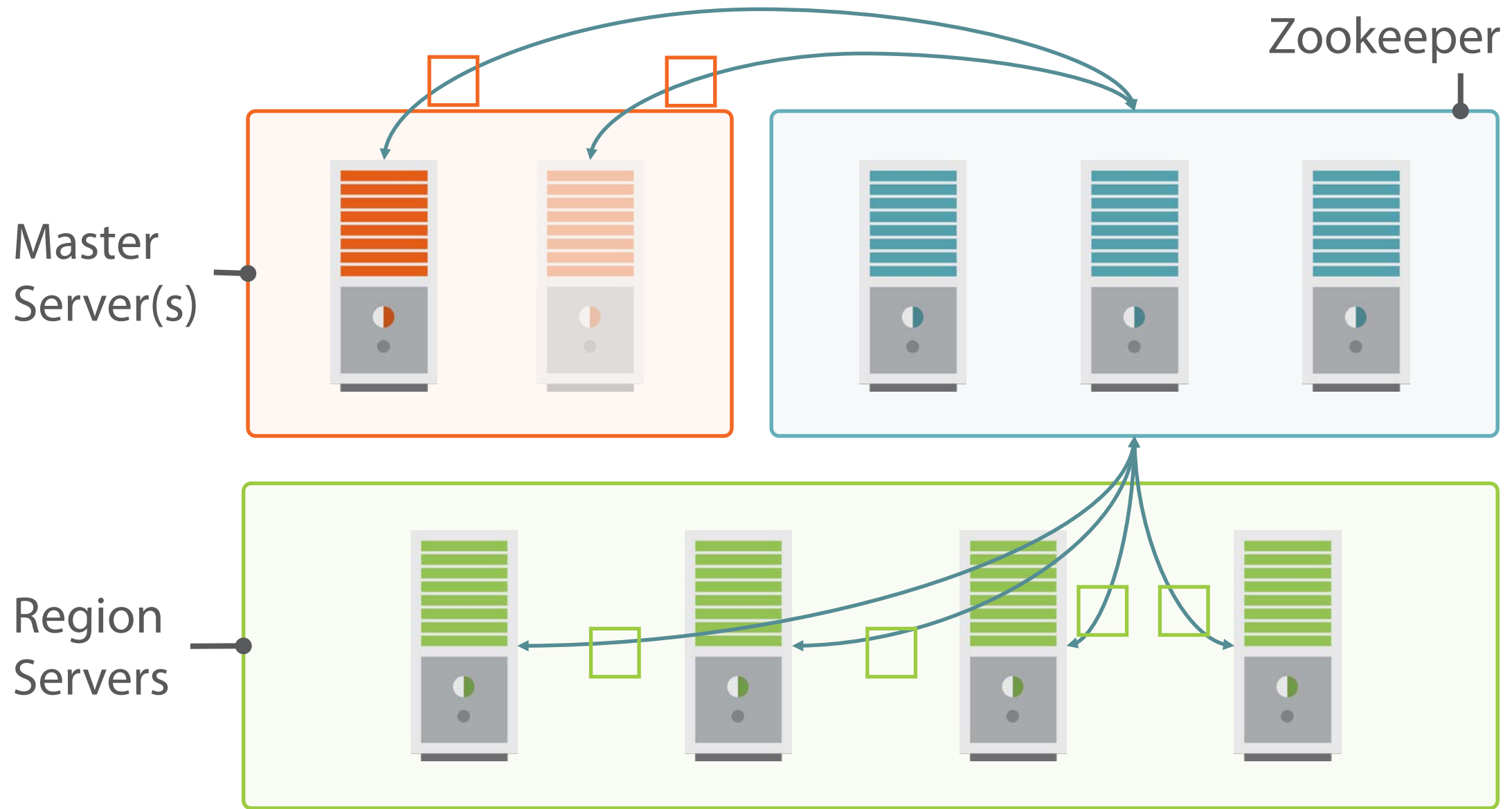


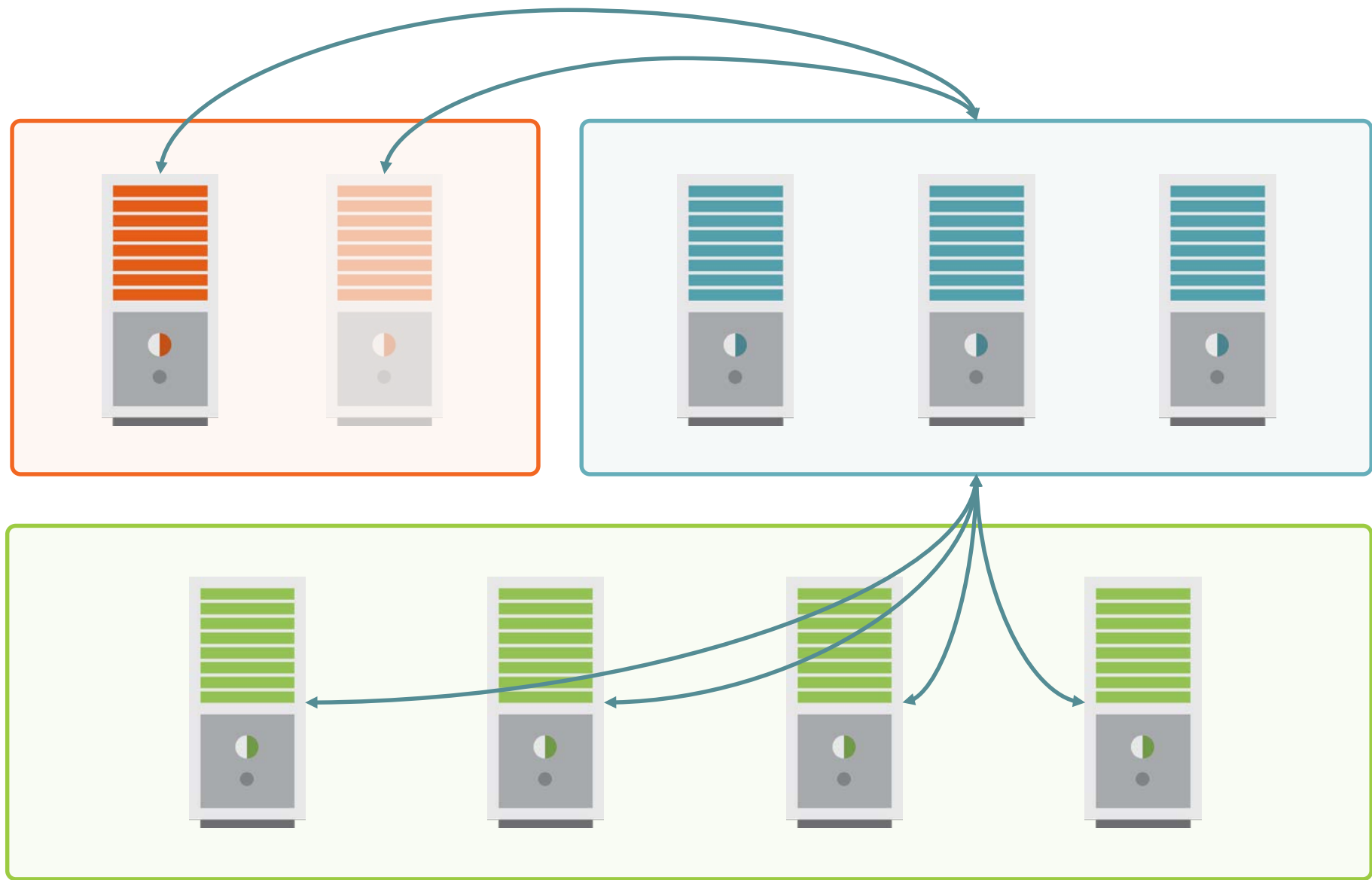
Region
Servers

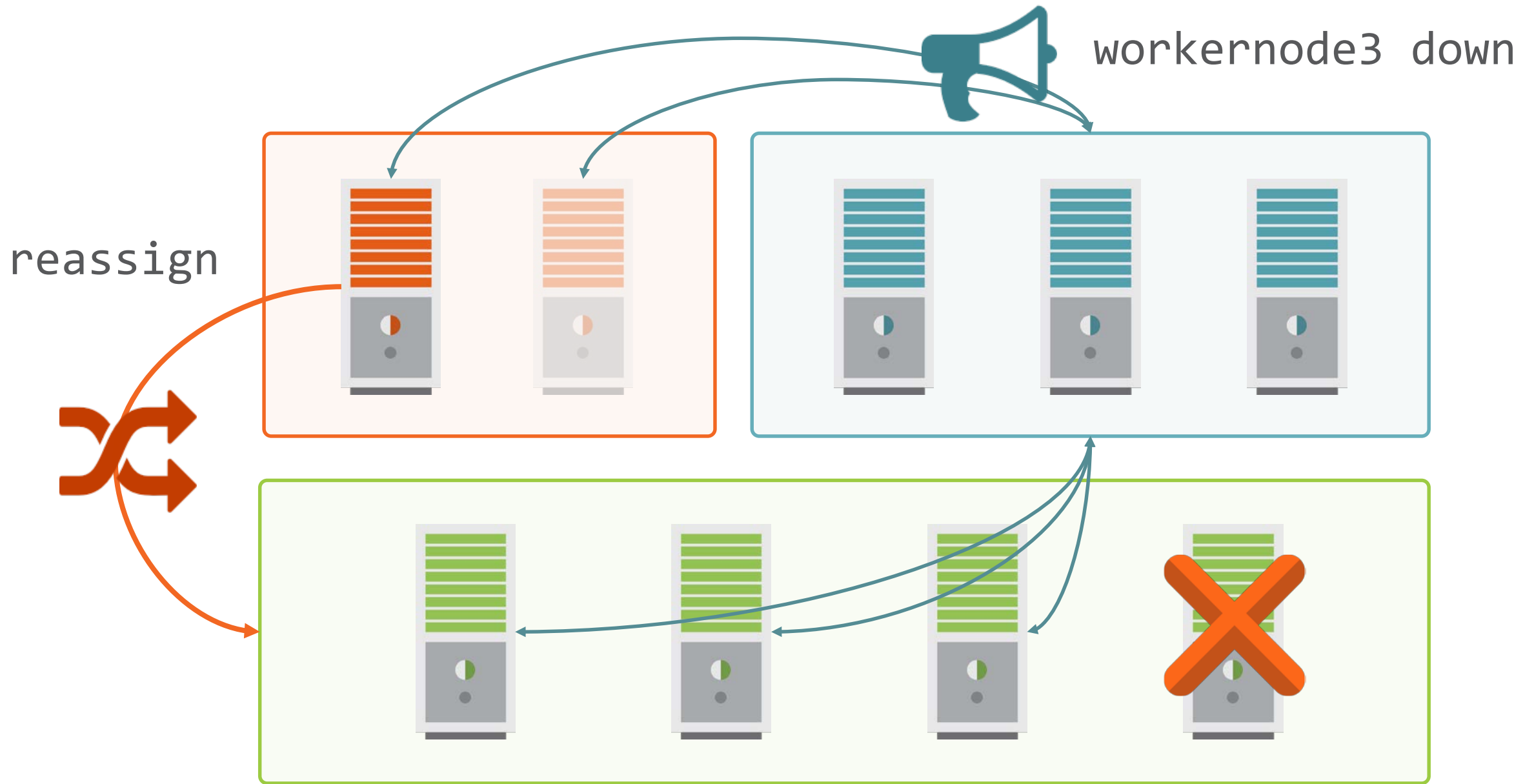


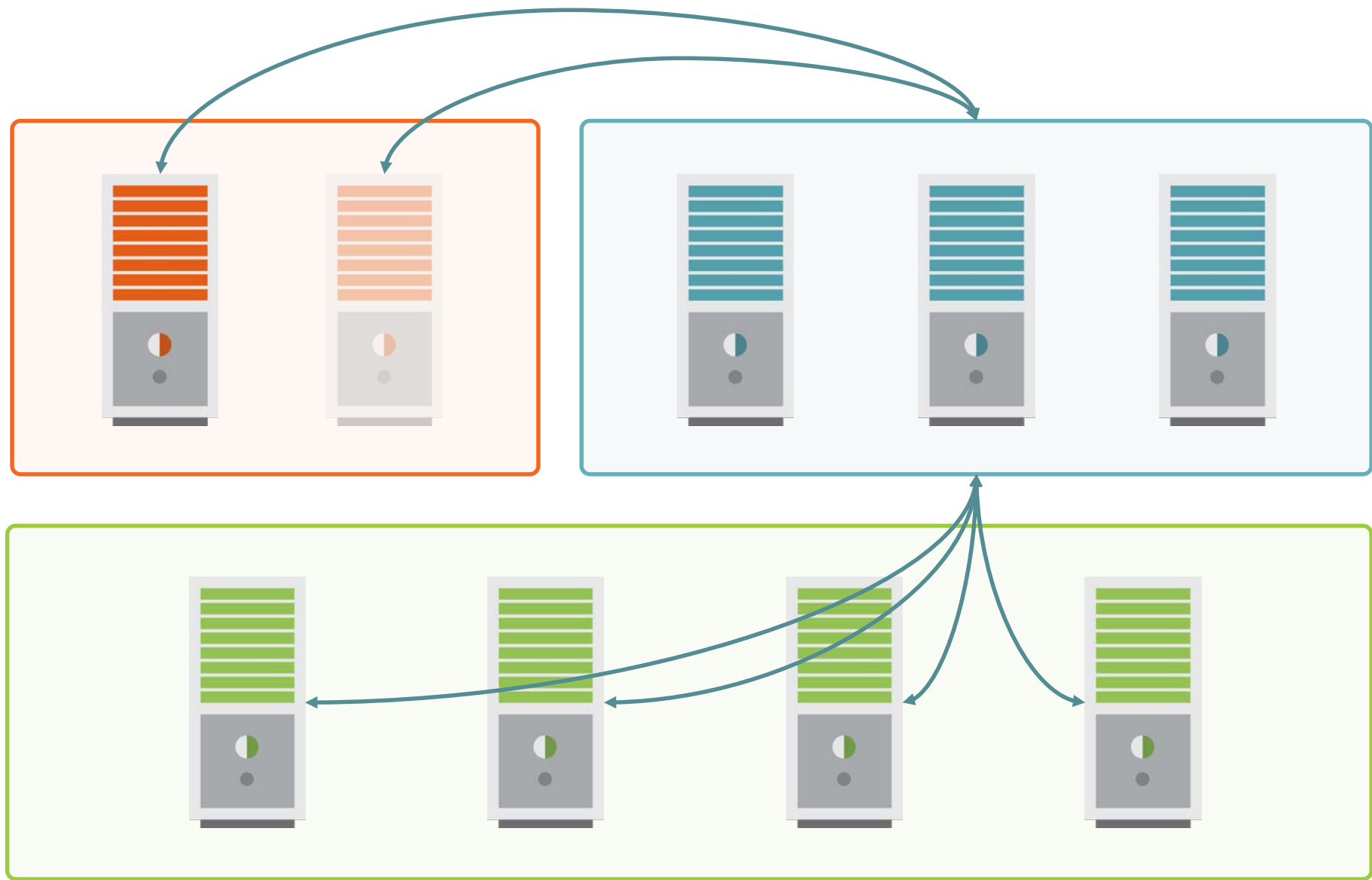






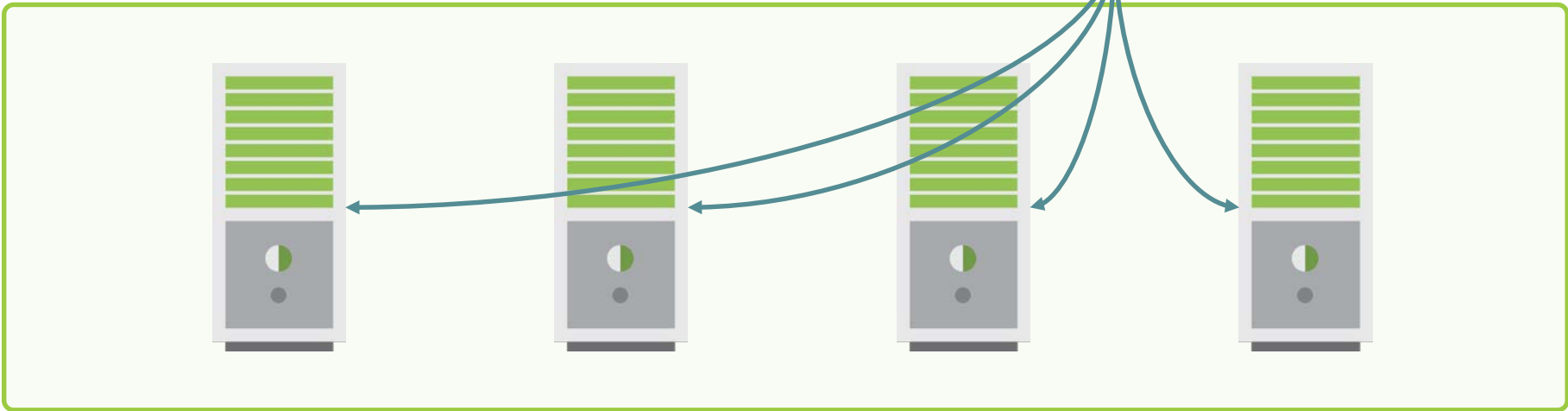
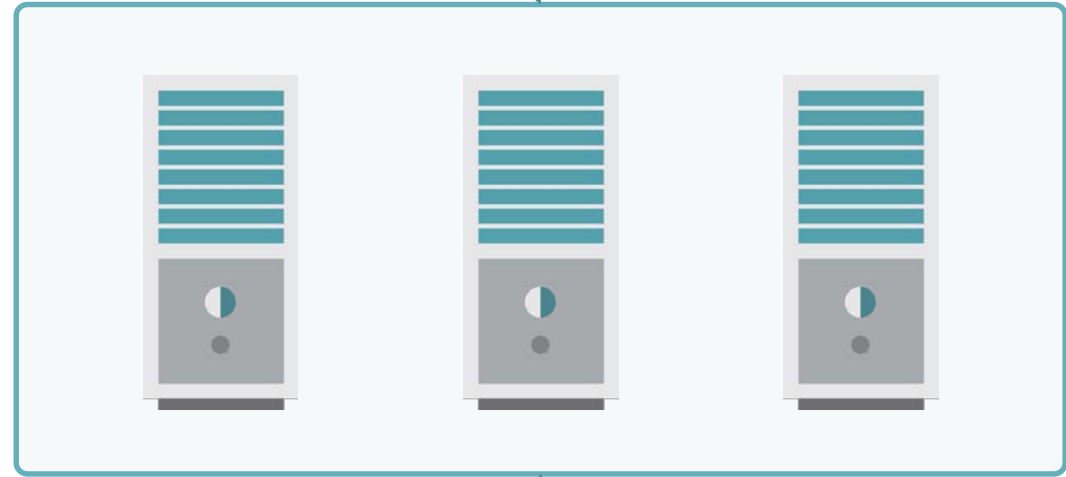






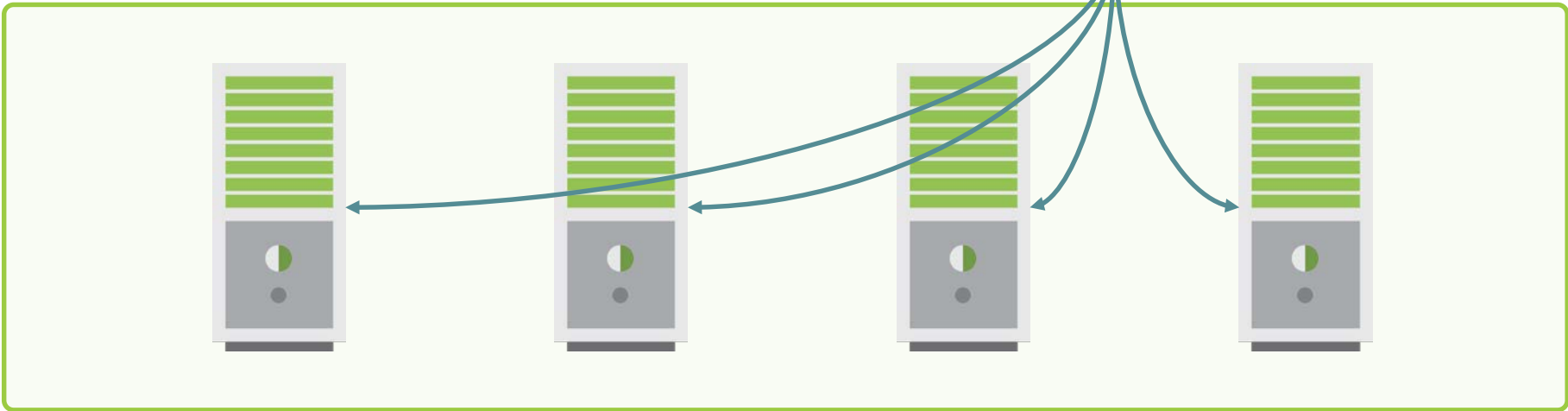
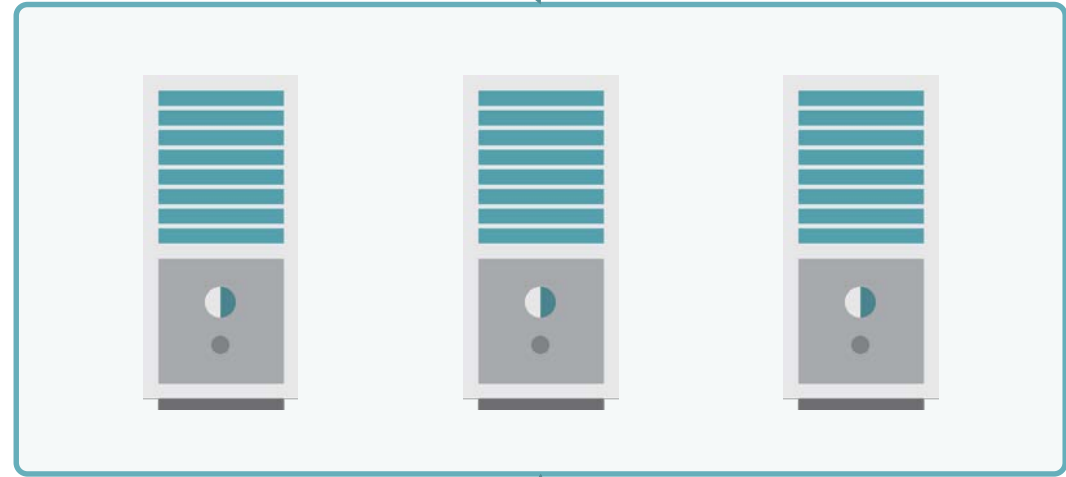


headnode0 down

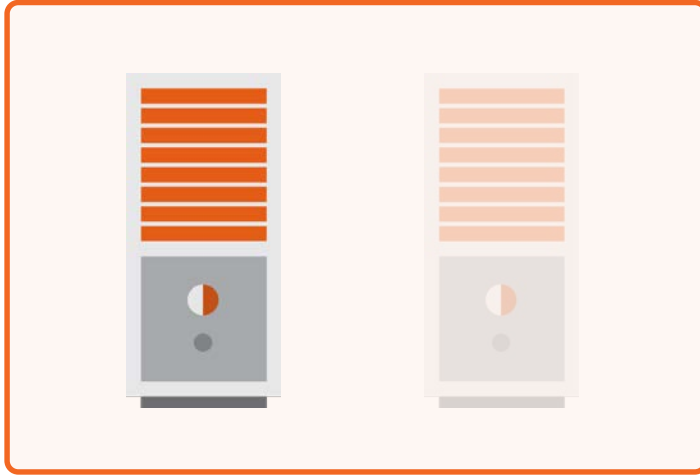




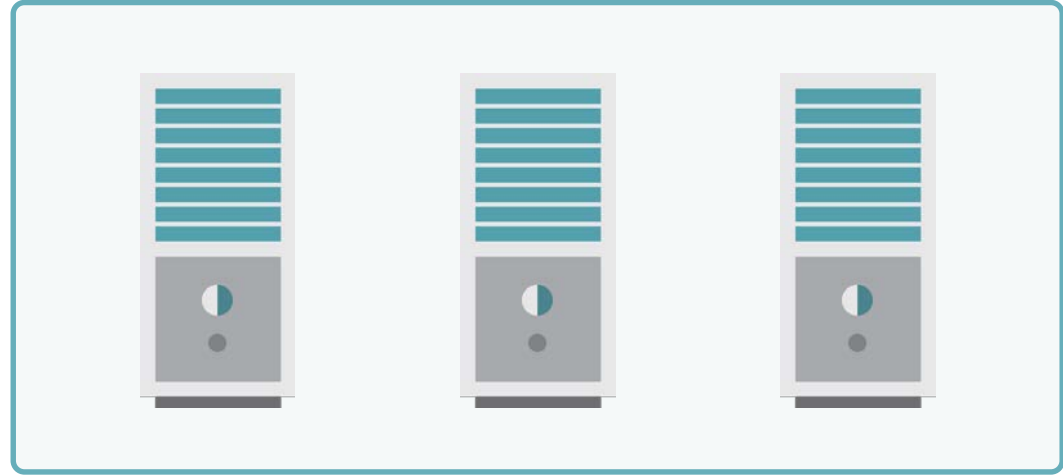
headnode0 down



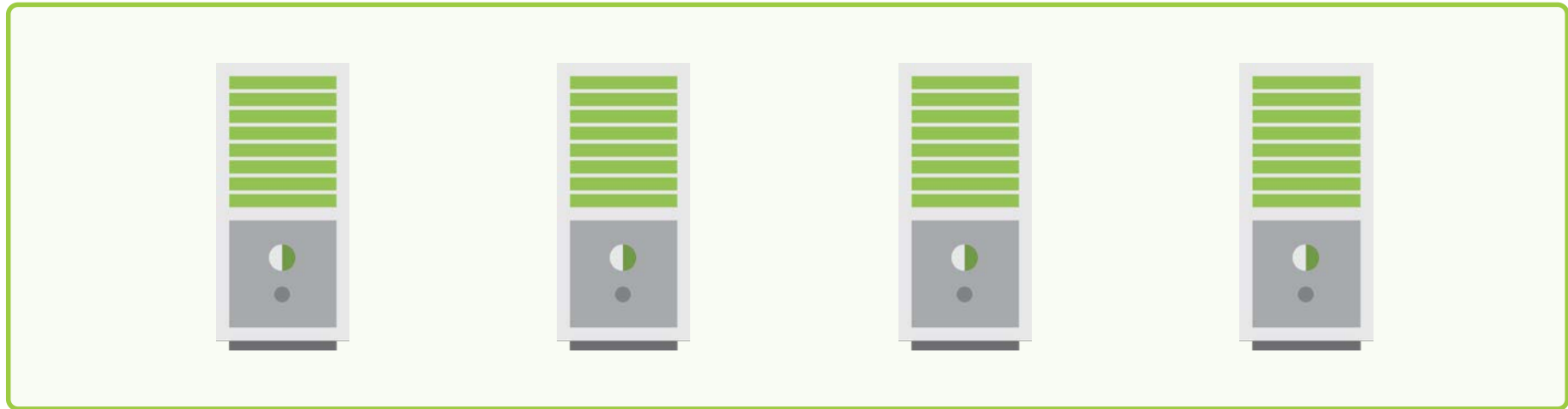
2x Master



3x Zookeeper



4x Region





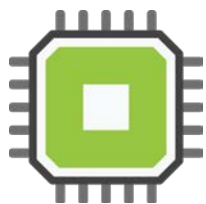
x4



x2



x3



4 cores

4 cores

4 cores



14 GB

7 GB

7 GB

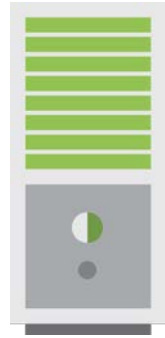


200 GB **SSD**

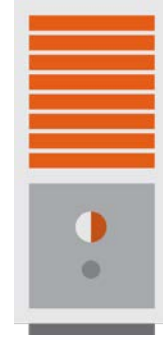
285 GB

285 GB

27
cores



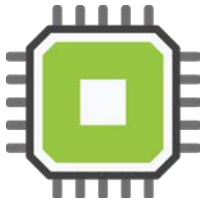
x4



x2



x3



4 cores

4 cores

1 core



7 GB

7 GB

3.5 GB



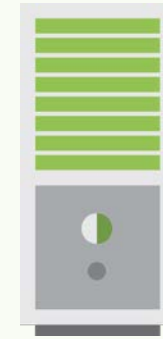
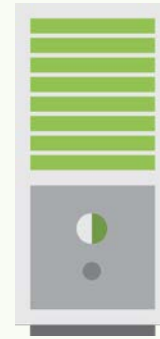
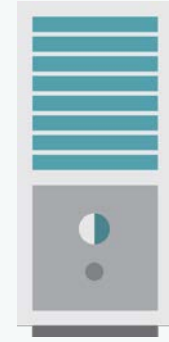
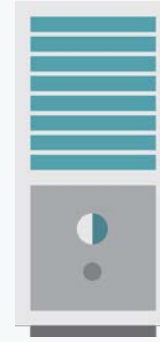
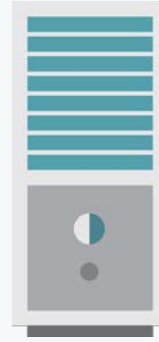
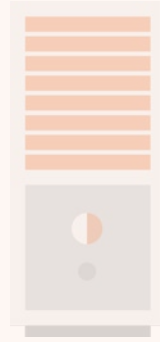
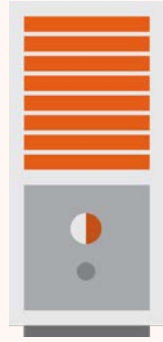
285 GB

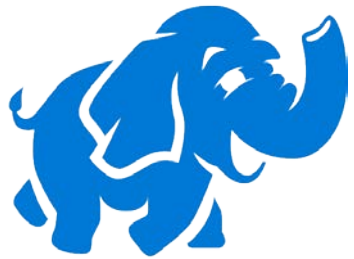
285 GB

50 GB **SSD**

\$1K /month

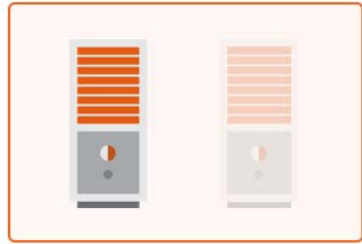
\$1.50 /hour





cluster-01

cluster-02



sector-times	te	d
--------------	----	---

RowKey	te	d
0000 a6545da436		
0123 a6545da436		
0fff a6545da436		
1000 a6545da436		
1123 a6545da436		
1fff a6545da436		

Region 1

Start key = 0000 |
End key = 0fff |

Region 2

Start key = 1000 |
End key = 1fff |

sector-times	te	d
--------------	----	---

RowKey	te	d
0000 a6545da436		
0123 a6545da436		
07ff a6545da436		
0800 a6545da436		
0a23 a6545da436		
0fff a6545da436		

Region 1
Start key = 0000 |
End key = 0fff |

sector-times	te	d
--------------	----	---

RowKey	te	d
0000 a6545da436		
0123 a6545da436		
07ff a6545da436		
0800 a6545da436		
0a23 a6545da436		
0fff a6545da436		

Region 1.1

Start key = 0000 |
End key = 07ff |

Region 1.2

Start key = 0800 |
End key = 0fff |

sector-times te d

RowKey	te	d
0000 ...		
0123 ...		
07ff ...		
0800 ...		
0a23 ...		
0fff ...		

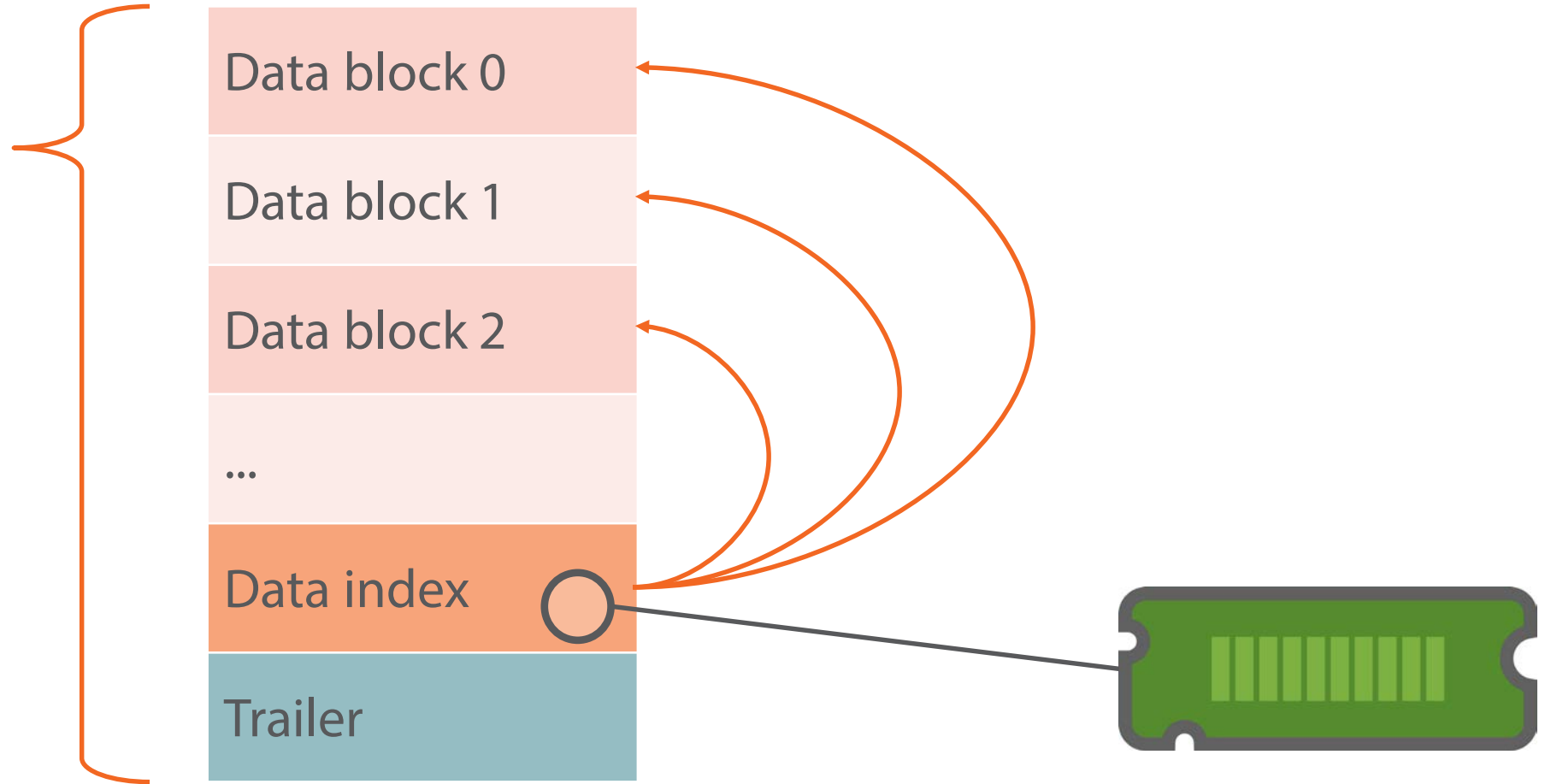
Region 1.1

Start key = 0000 |
End key = 07ff |

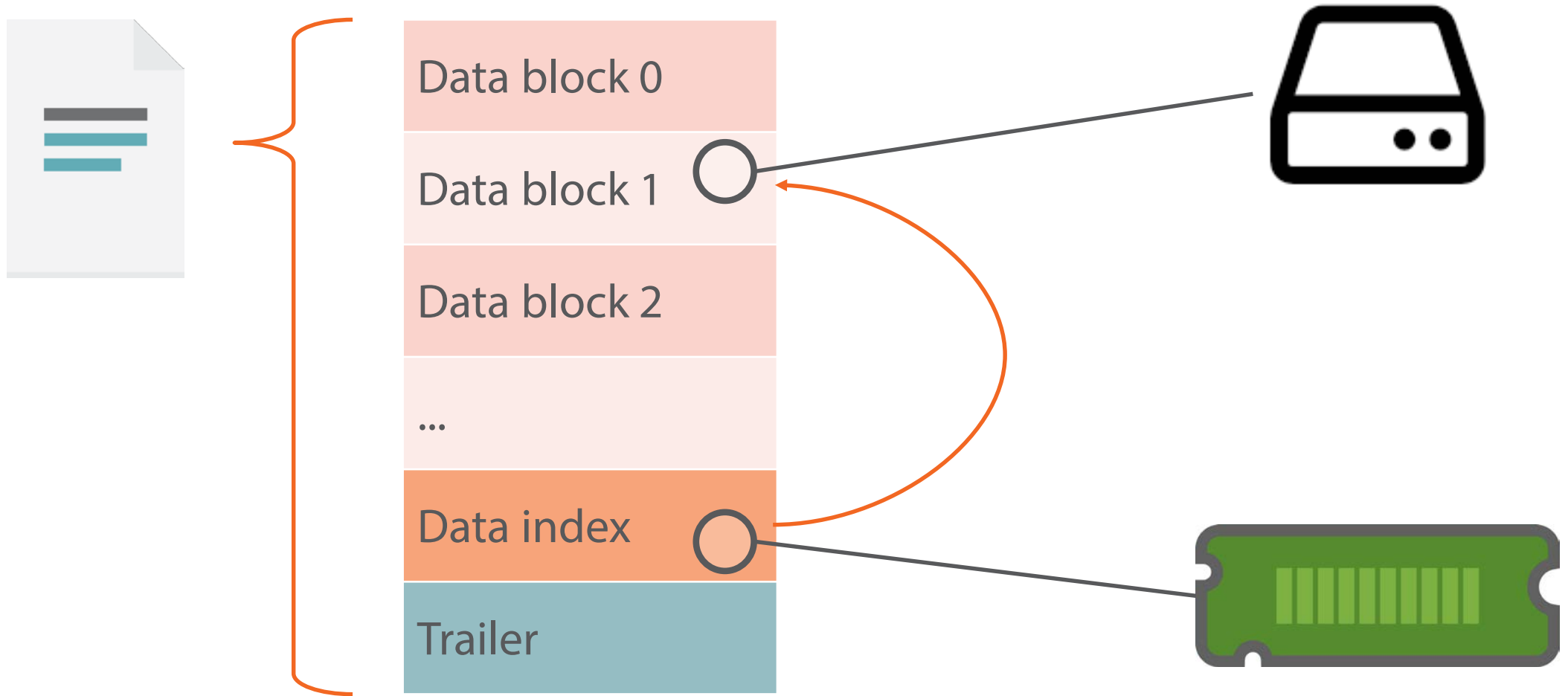
Region 1.2

Start key = 0800 |
End key = 0fff |

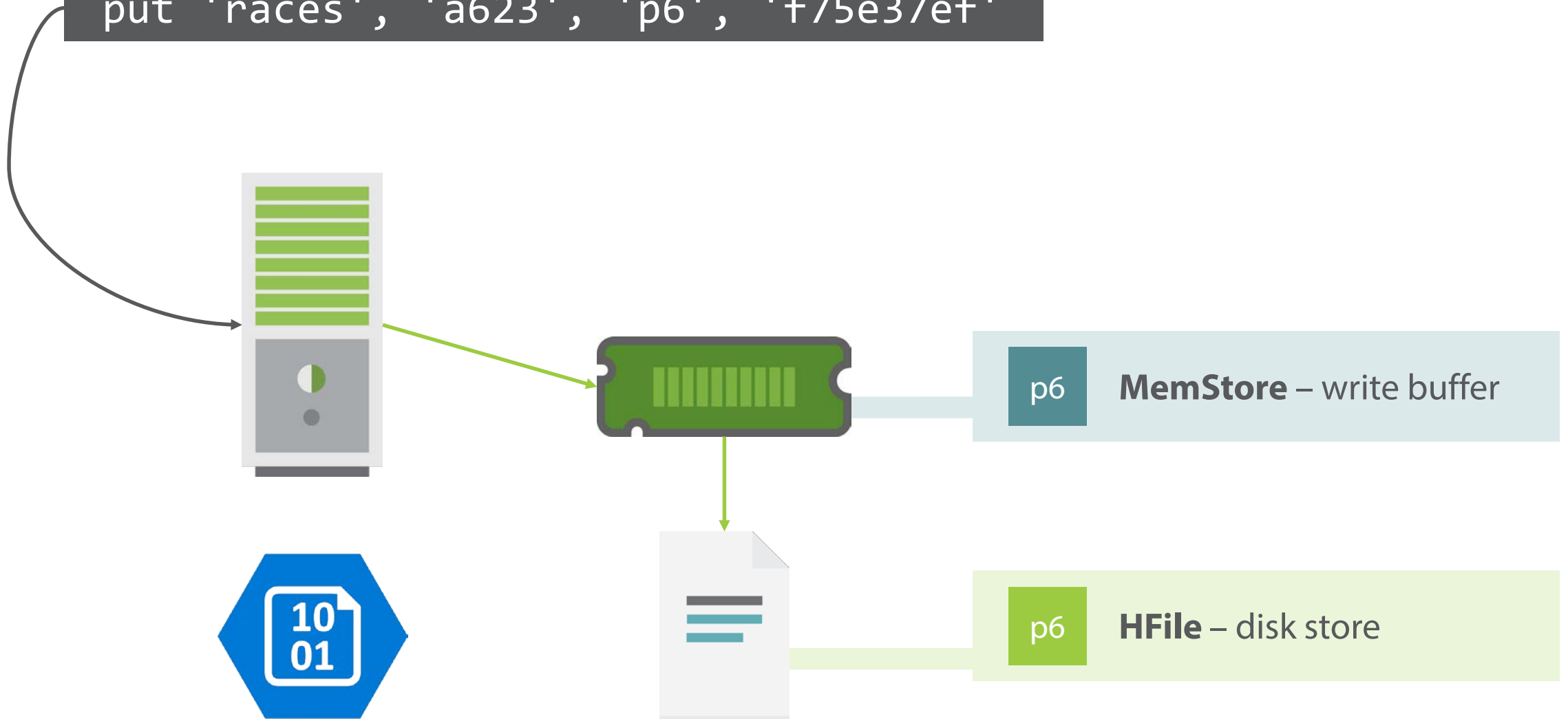
/sector-times/region-1.1/te/hfile-1



/sector-times/region-1.1/te/hfile-1

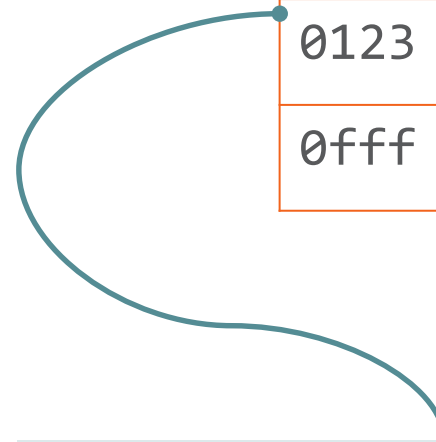


```
put 'races', 'a623', 'p6', 'f75e37ef'
```





RowKey	p
0000	
0123	
0fff	



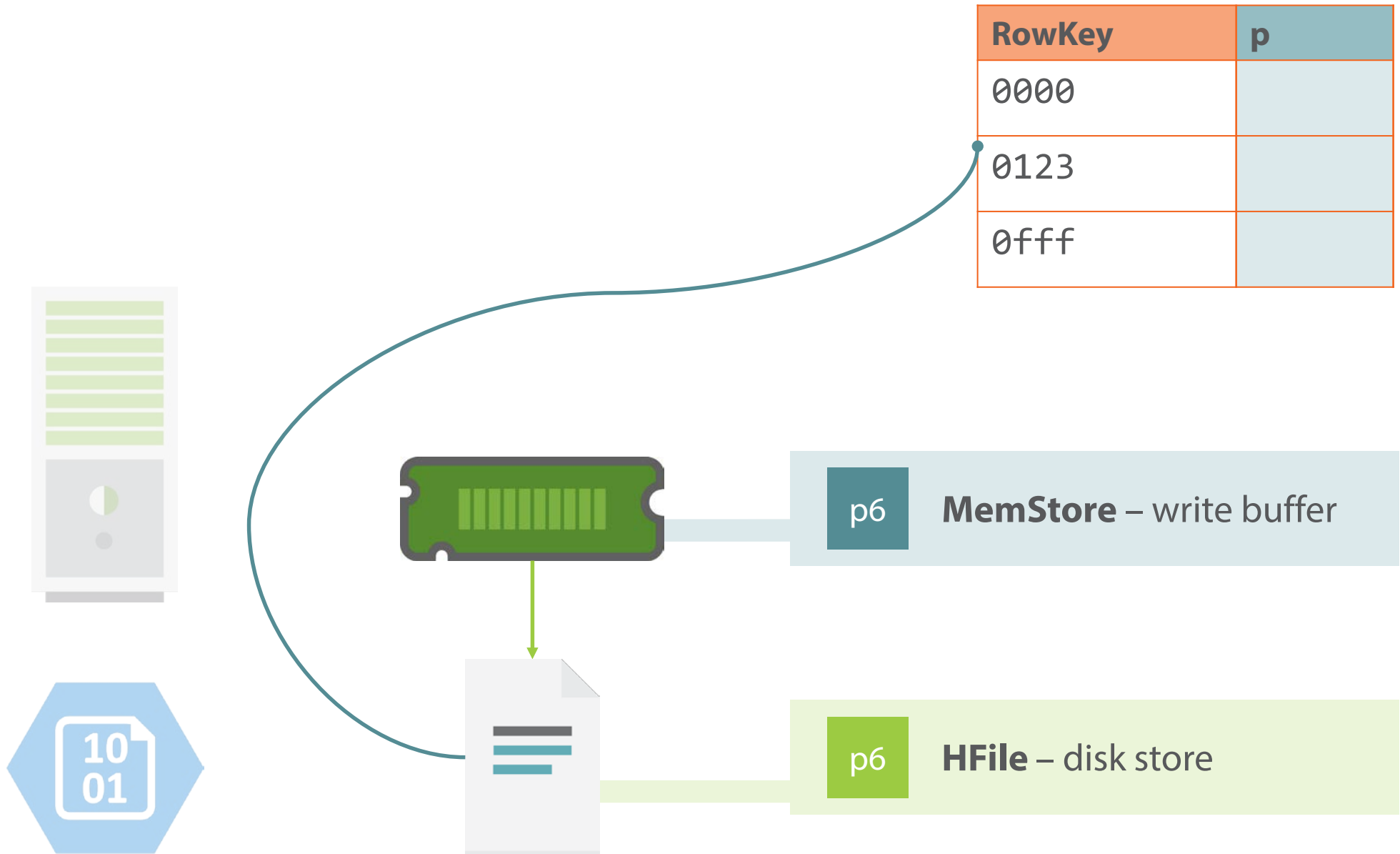
p6

MemStore – write buffer



p6

HFile – disk store

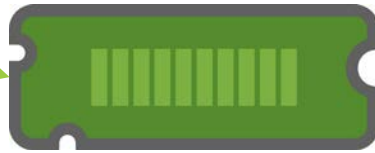


```
get 'races', 'a623'
```



p1

BlockCache - read cache



p9

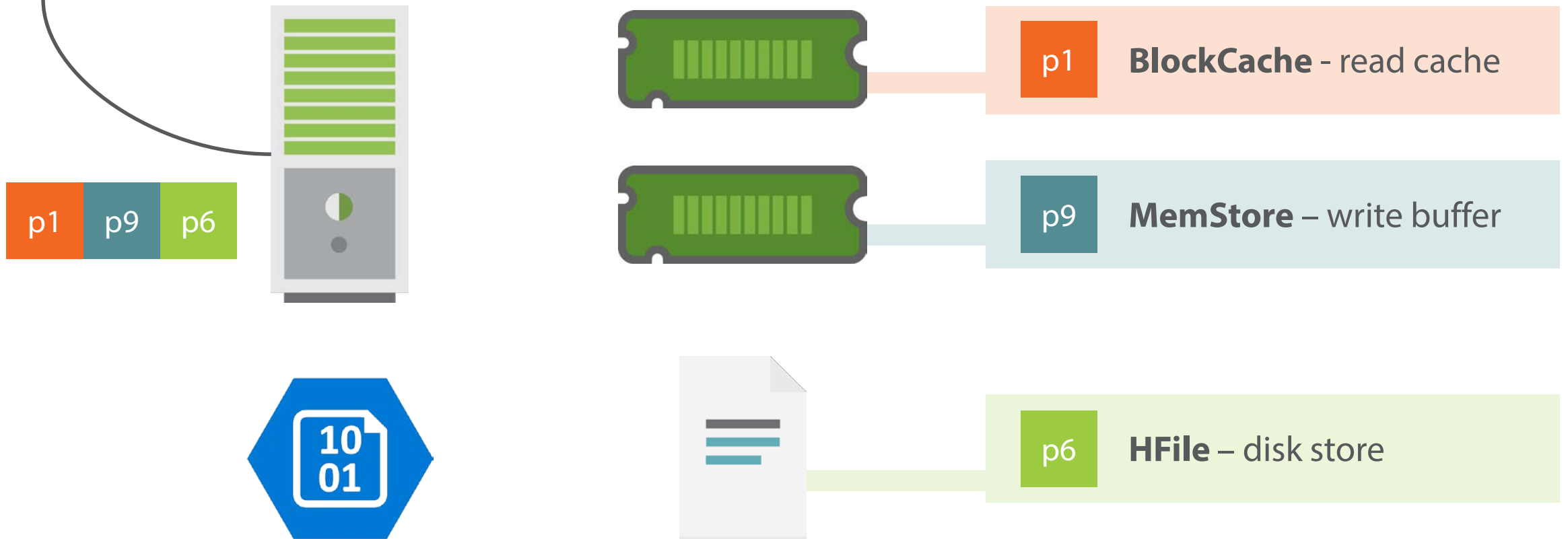
MemStore – write buffer

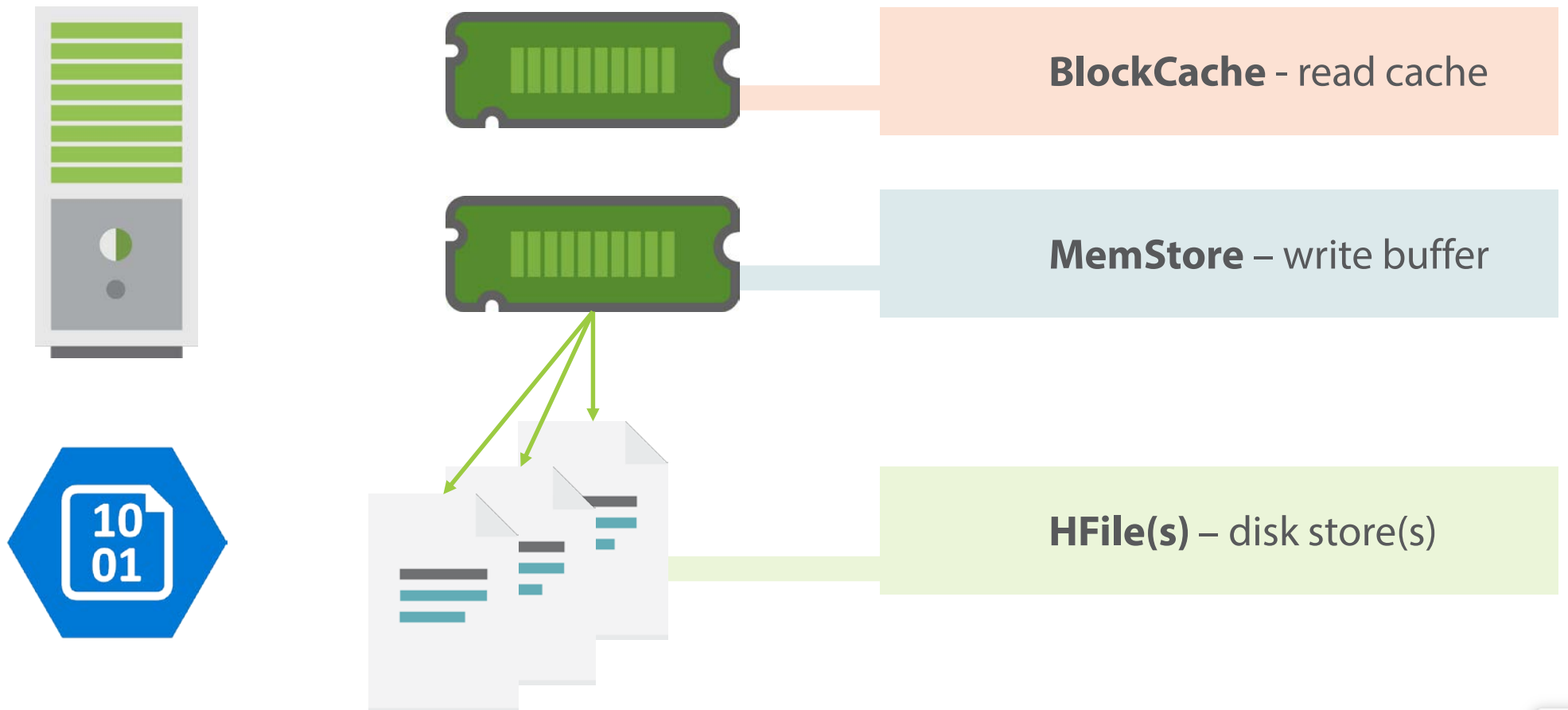


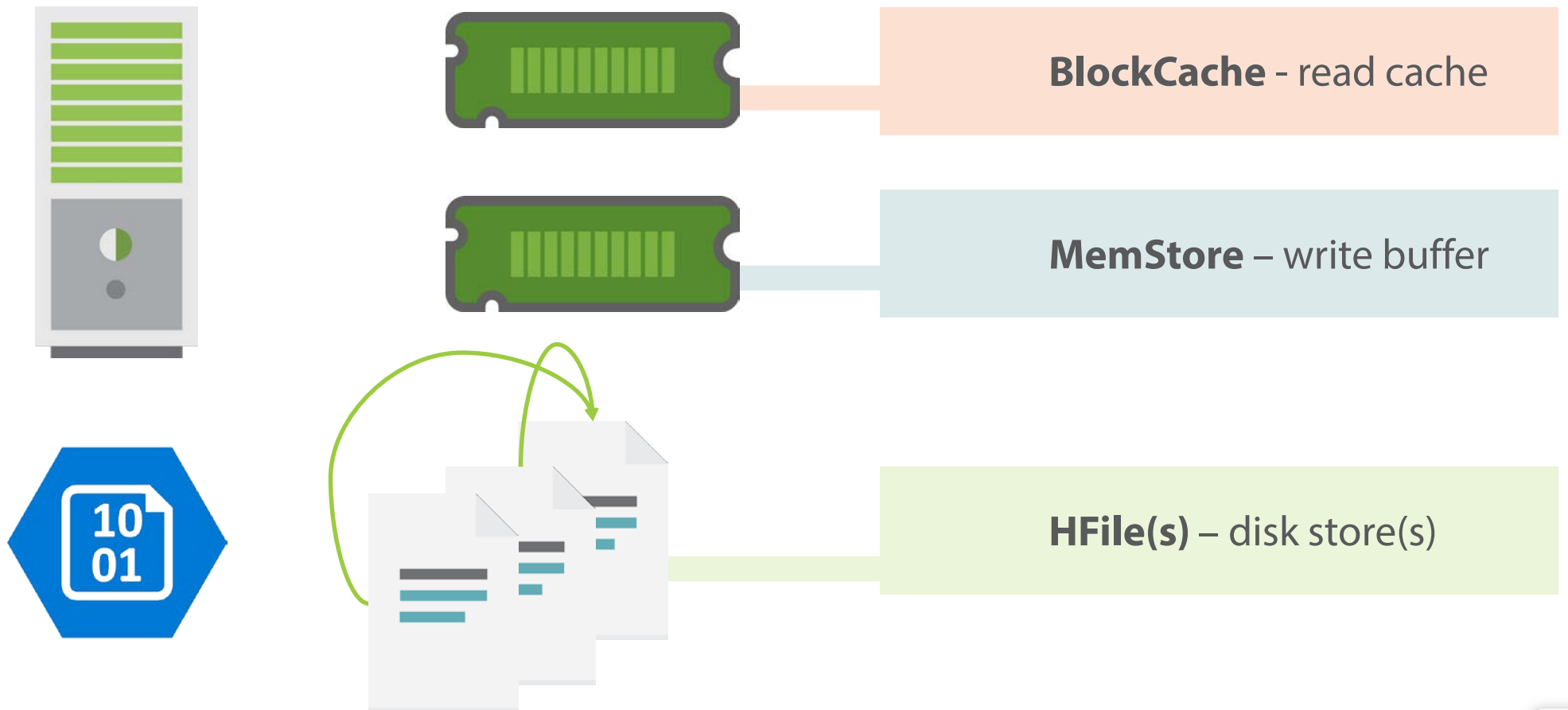
p6

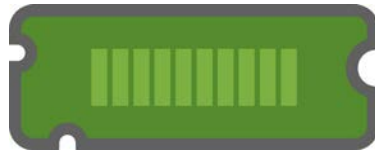
HFile – disk store

get 'races', 'a623'

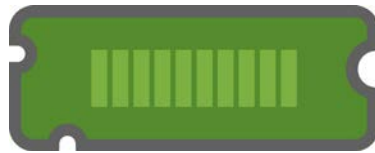








BlockCache - read cache

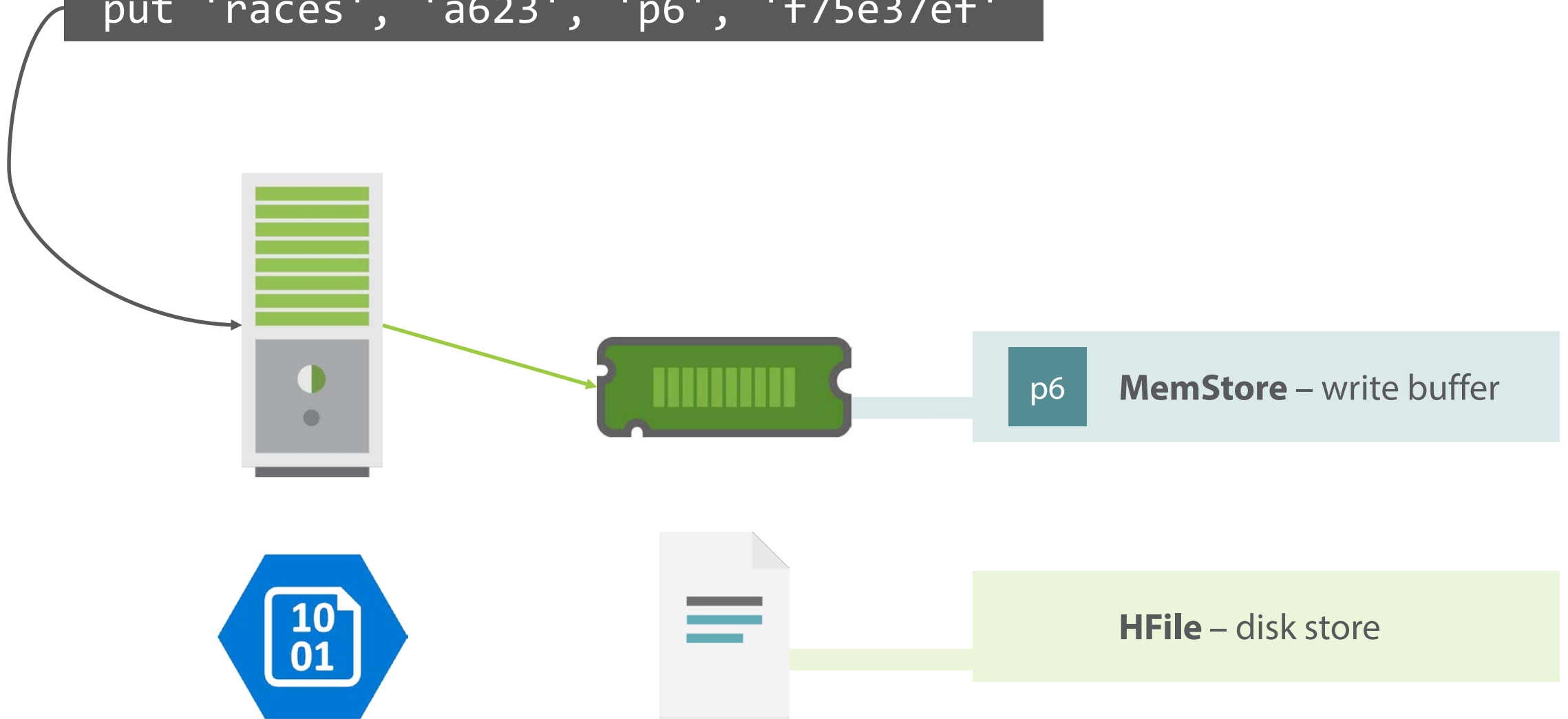


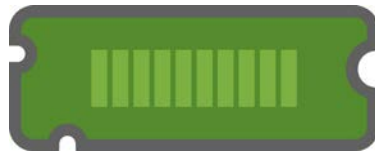
MemStore – write buffer



HFile(s) – disk store(s)

```
put 'races', 'a623', 'p6', 'f75e37ef'
```





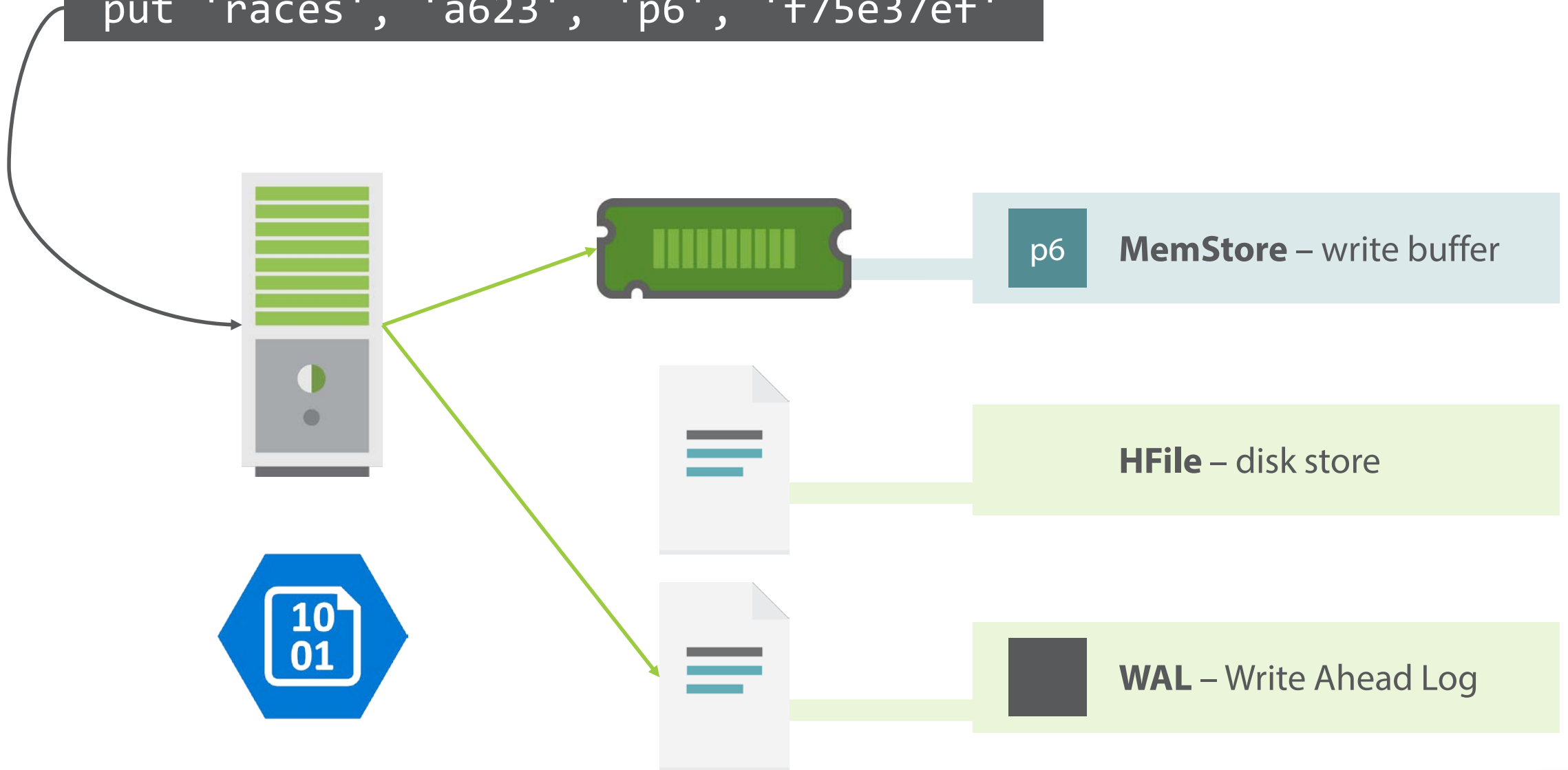
p6

MemStore – write buffer



HFile – disk store

```
put 'races', 'a623', 'p6', 'f75e37ef'
```





MemStore – write buffer

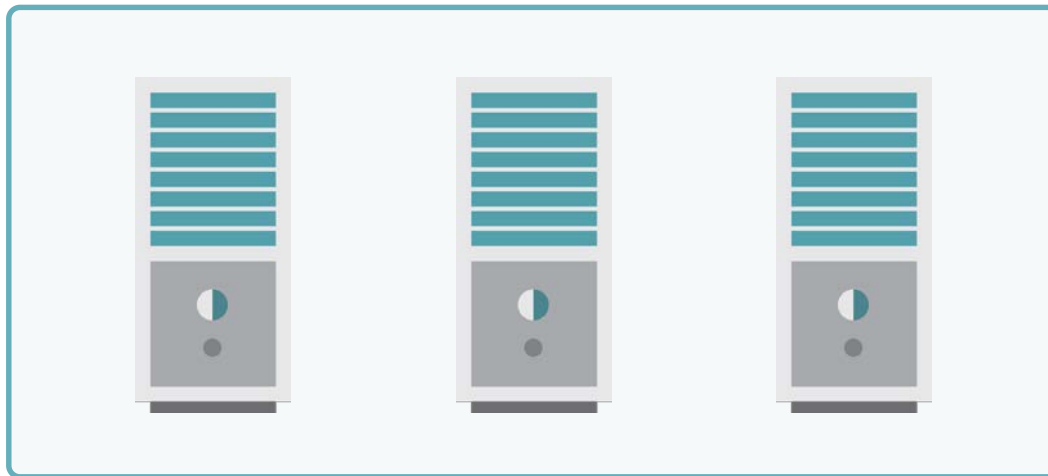
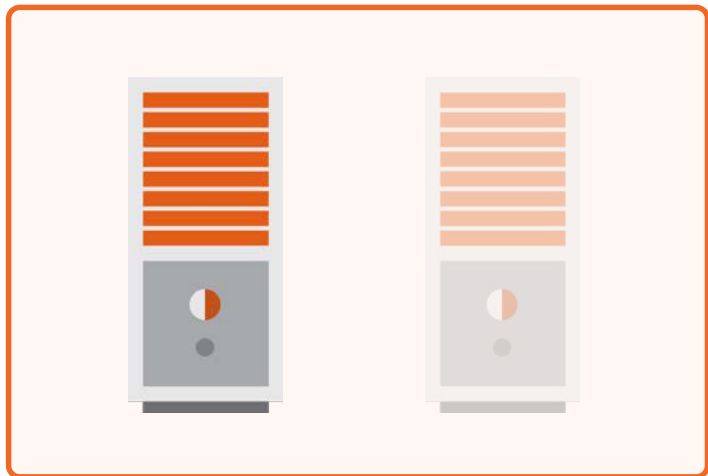


p6

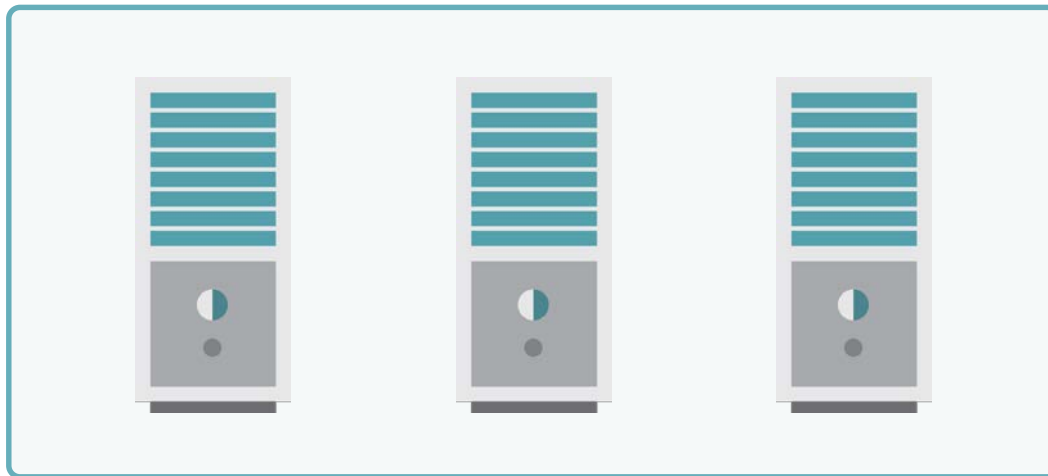
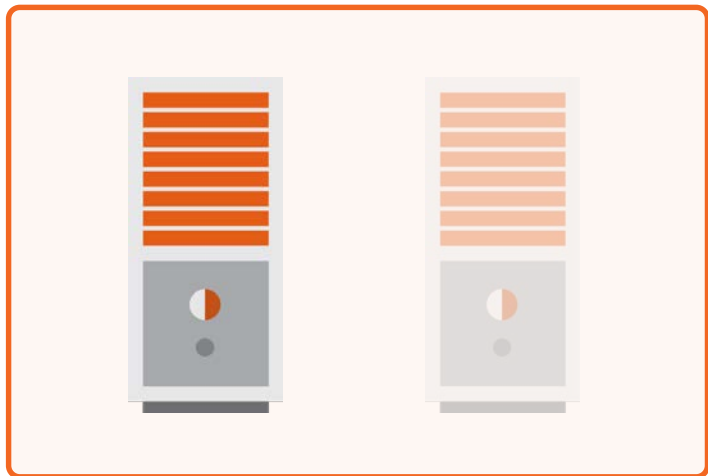
HFile – disk store



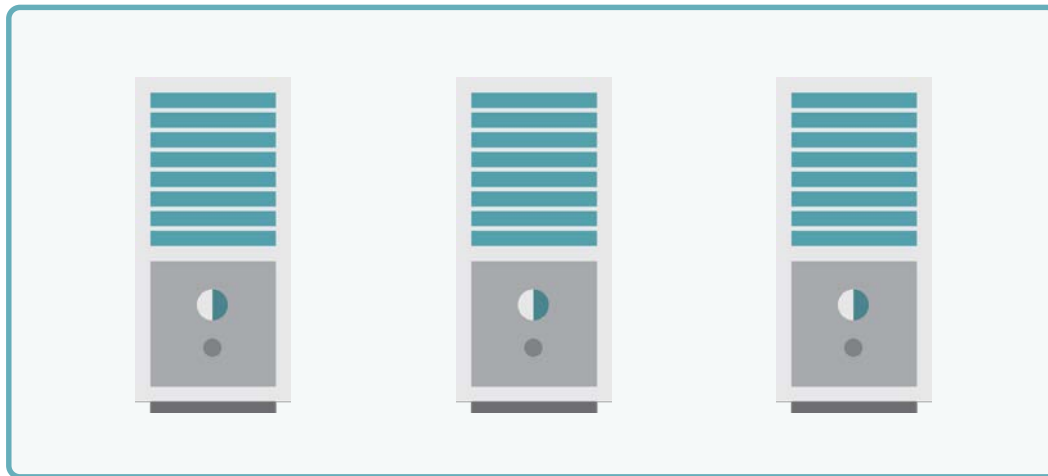
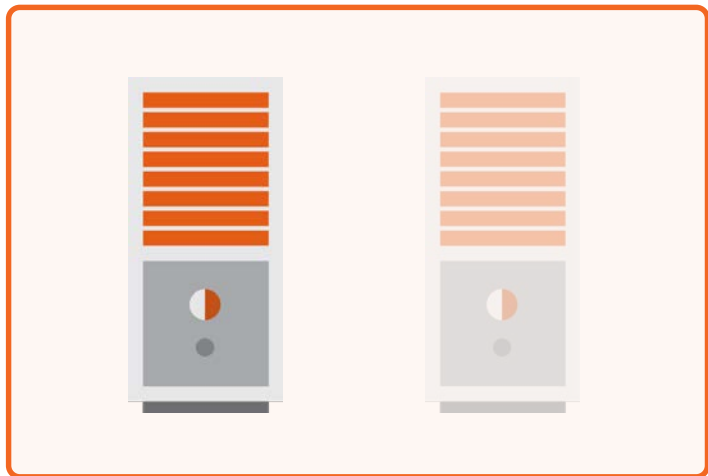
WAL – Write Ahead Log



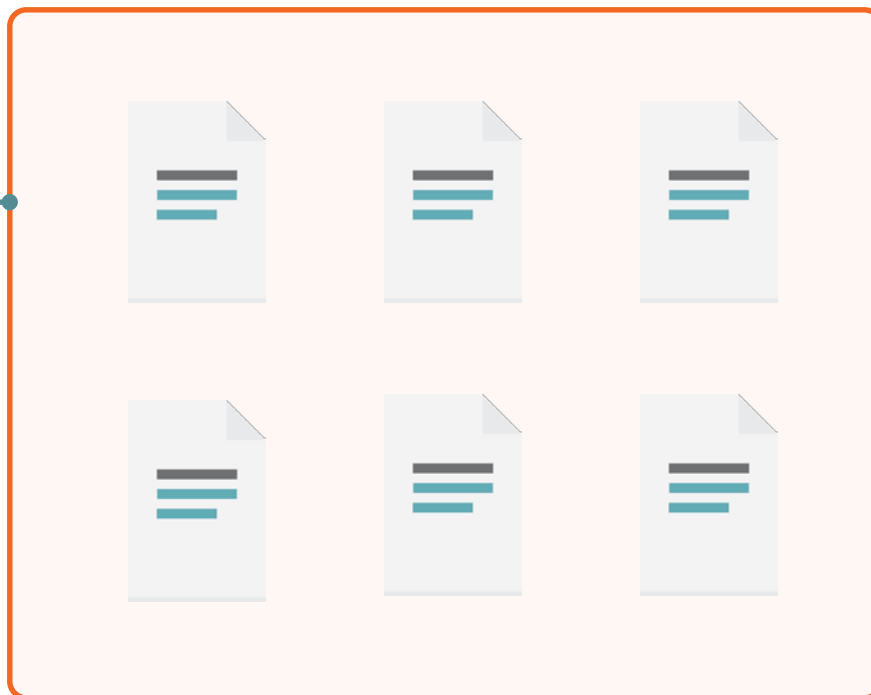
/sector-times
/region-1



/sector-times
/region-1



/sector-times
/region-1



cyberduck.io



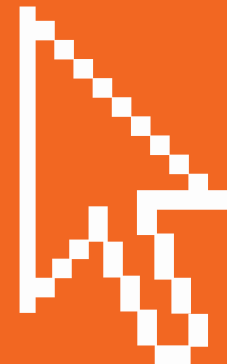
cloudberrylab.com

Demo: HBase File Storage

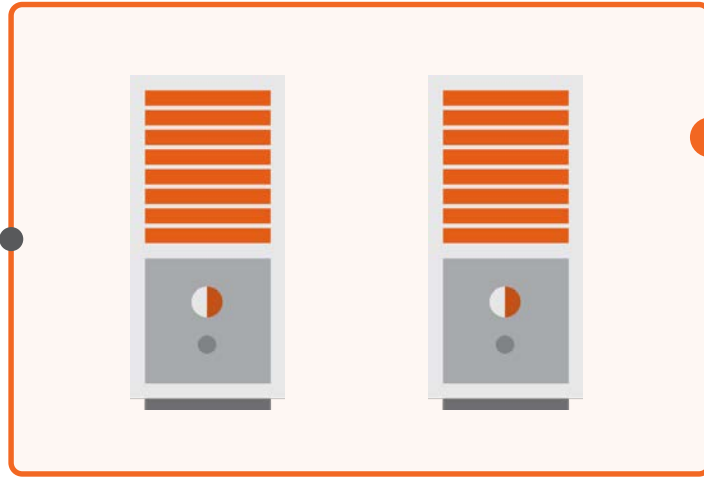
Azure Portal

Folder Structure

HFile Contents

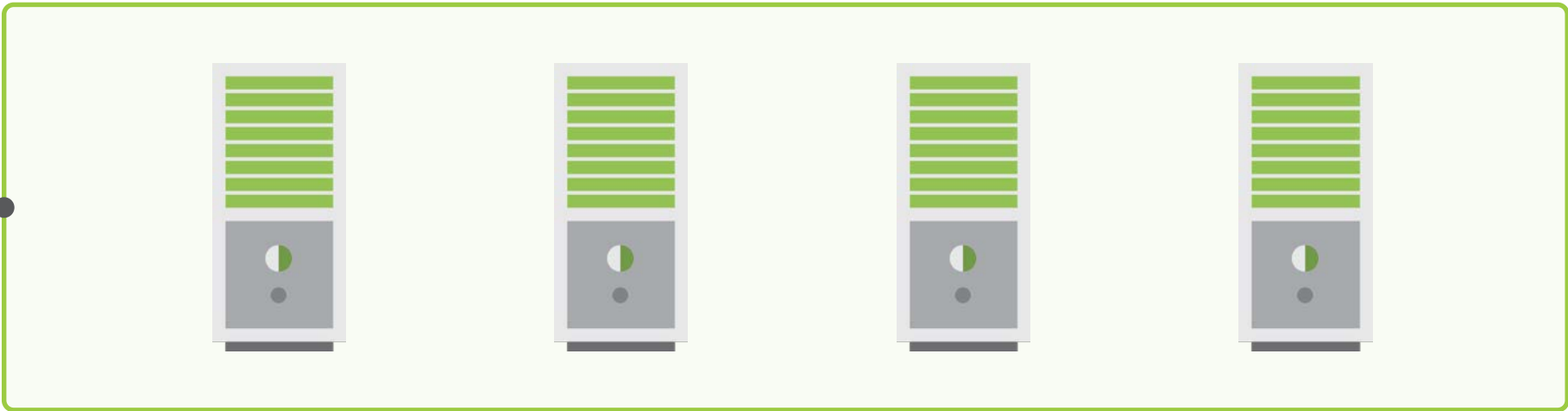


Master
Server(s)

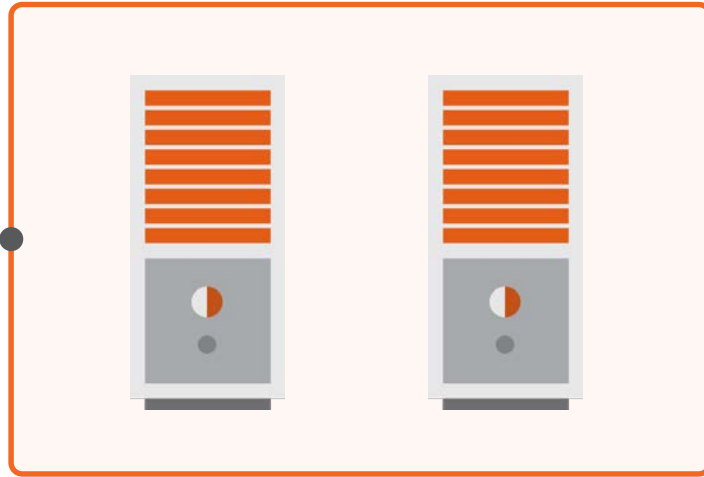


Schema changes
Cluster management
Data administration

Region
Servers



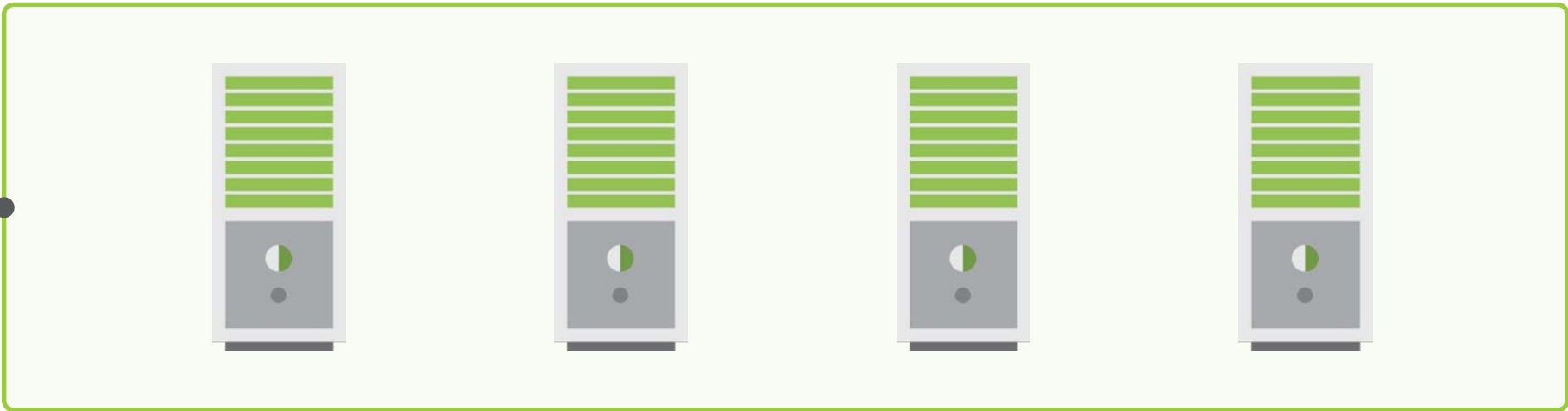
Master
Server(s)



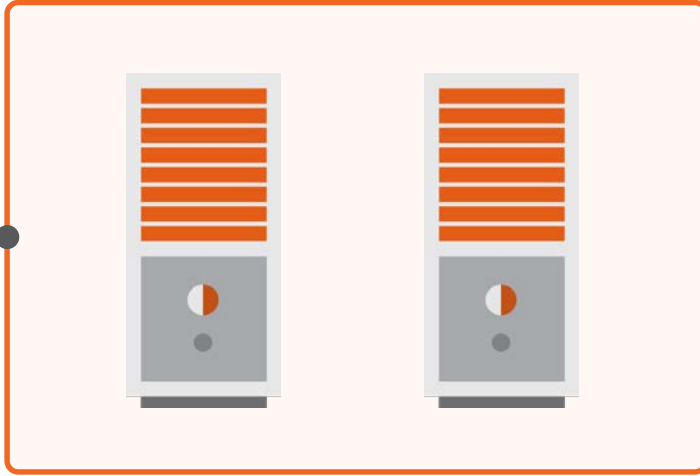
```
> major_compact ...
```



Region
Servers



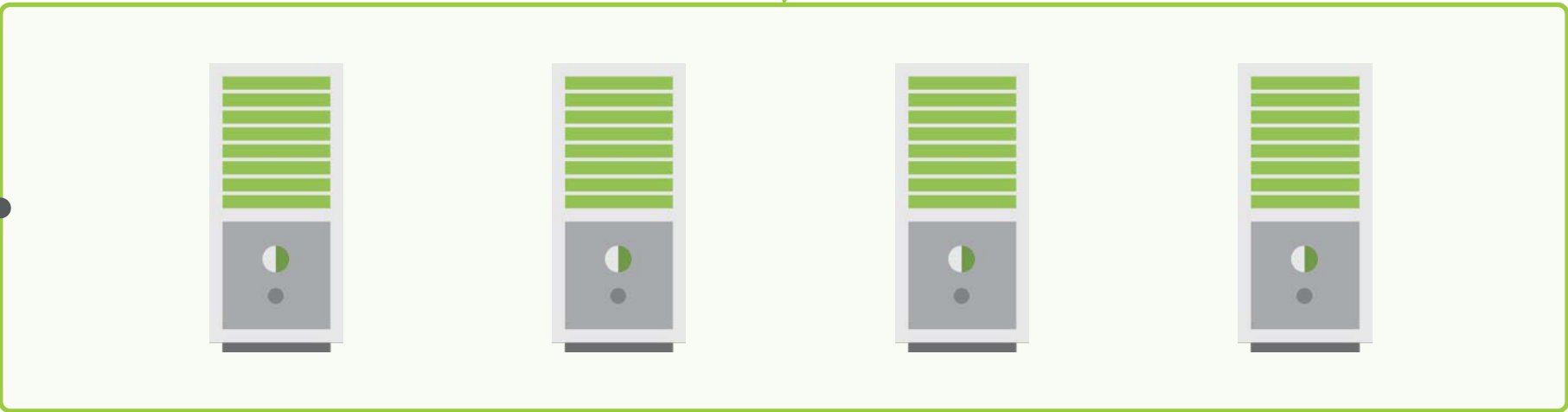
Master
Server(s)



```
> major_compact ...
```



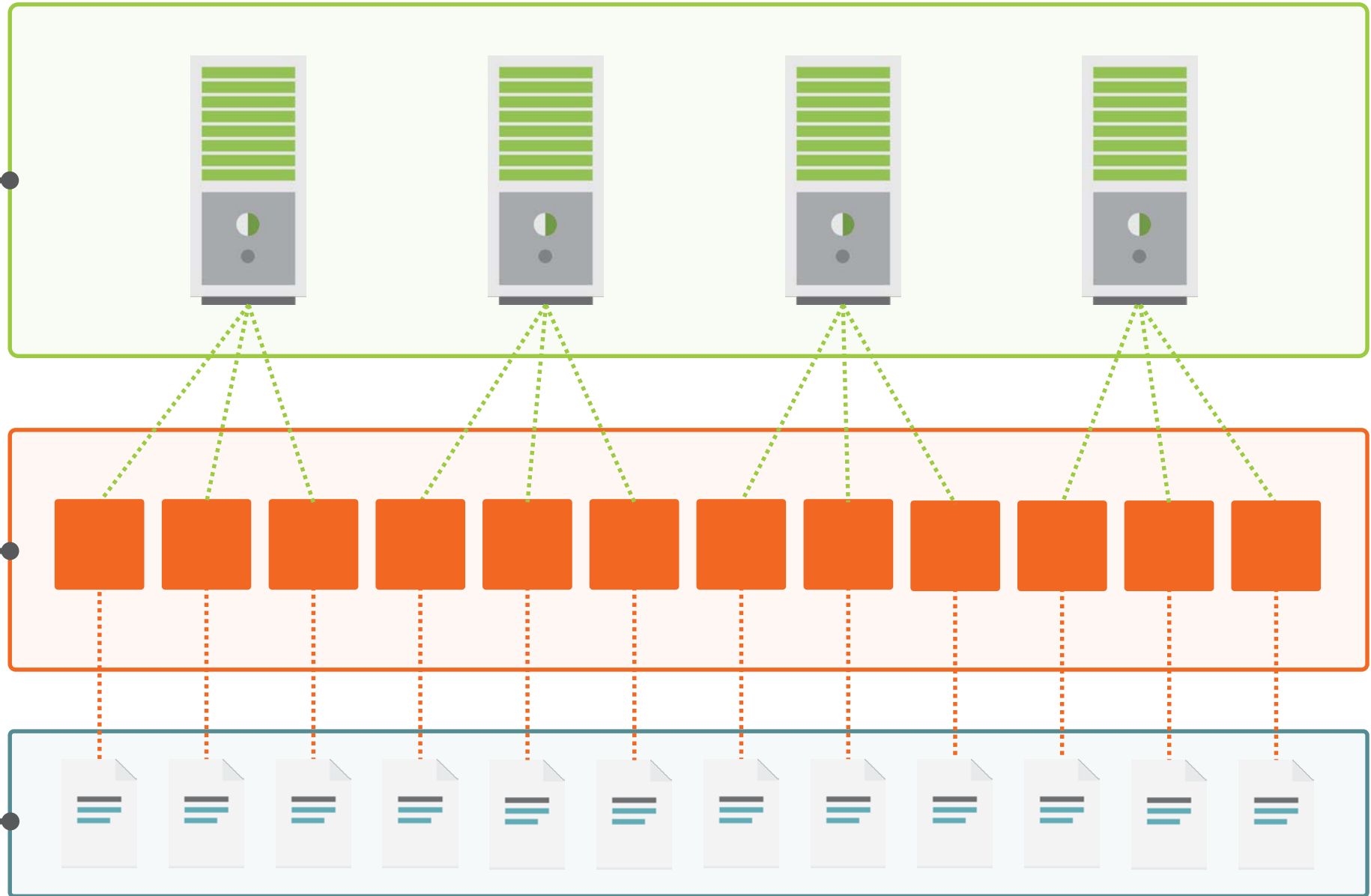
Region
Servers



Region
Servers

Regions

HFiles



races

race-timers

sector-times

Infrequent read/write

Moderate read/write

timing-events

Intensive write

sector-times	te	d
--------------	----	---

RowKey	te	d
0000 a6545da436		
0123 a6545da436		
07ff a6545da436		
0800 a6545da436		
0a23 a6545da436		
0fff a6545da436		

Region 1.1

Start key = 0000 |
End key = 07ff |

Region 1.2

Start key = 0800 |
End key = 0fff |



Region split at **256MB**

sector-times te d

RowKey te d



Region 1
Start key = ?
End key = ?

sector-times	te	d
--------------	----	---

RowKey	te	d
0000 a6545da436		
0123 a6545da436		
0fff a6545da436		



Region 1
Start key = 0000 |
End key = 0fff |

sector-times	te	d
--------------	----	---

RowKey	te	d
0000 a6545da436		
0123 a6545da436		
0fff a6545da436		
1000 a6545da436		
1123 a6545da436		
1fff a6545da436		
2000 a6545da436		
2904 a6545da436		

Region 1

Start key = 0000 |
End key = 0fff |

Region 2

Start key = 1000 |
End key = 1fff |

Region 3

Start key = 2000 |
End key = 2fff |

Demo: Regions and Servers

Region Splits

Region Distribution

Region Server Utilisation



sector-times	te	d
--------------	----	---

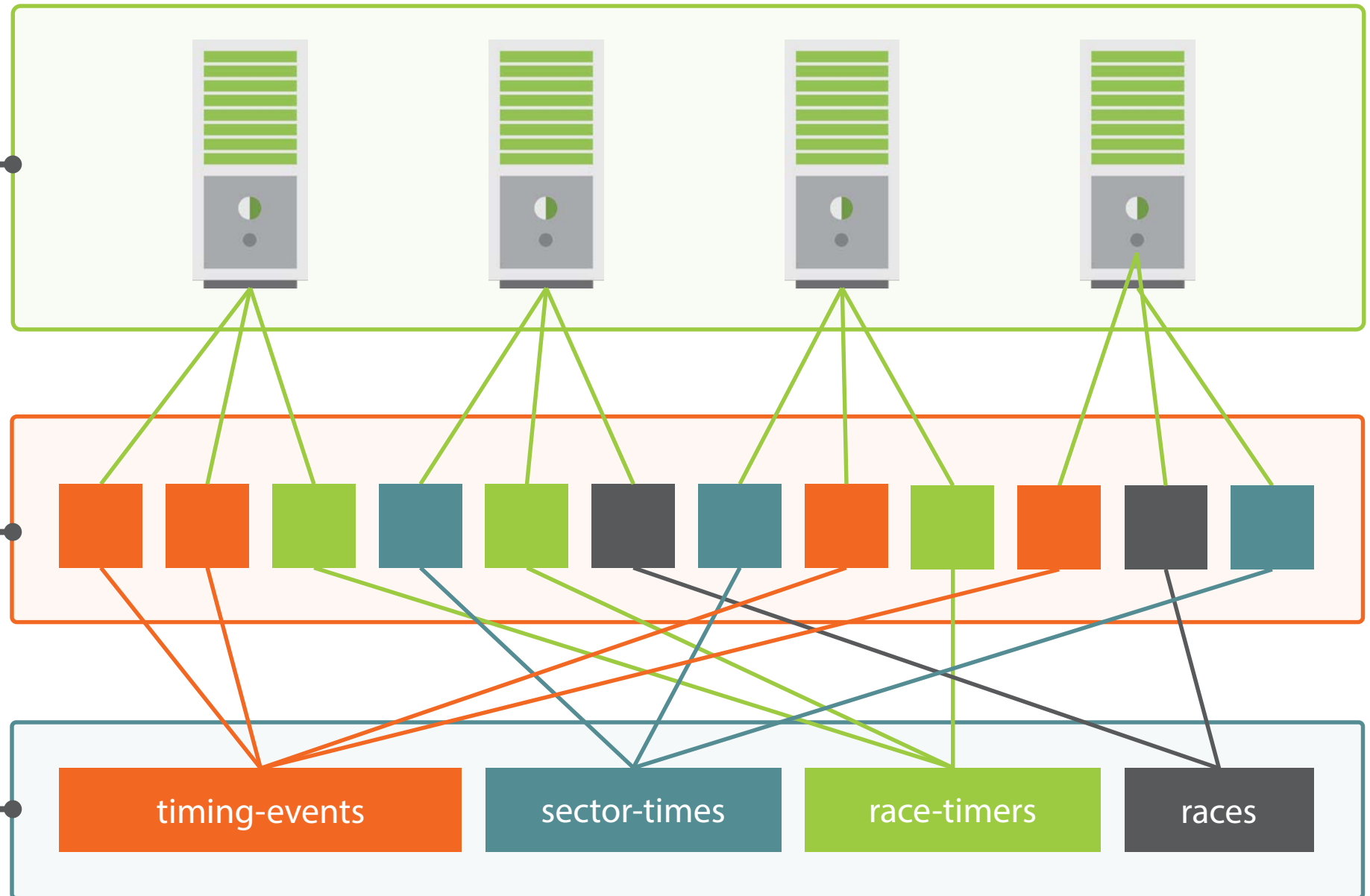
RowKey	te	d
0000 a6545da436		
0123 a6545da436		
0fff a6545da436		
1000 a6545da436		
1123 a6545da436		
1fff a6545da436		

Rowkey = {**guid**}

Region
Servers

Regions

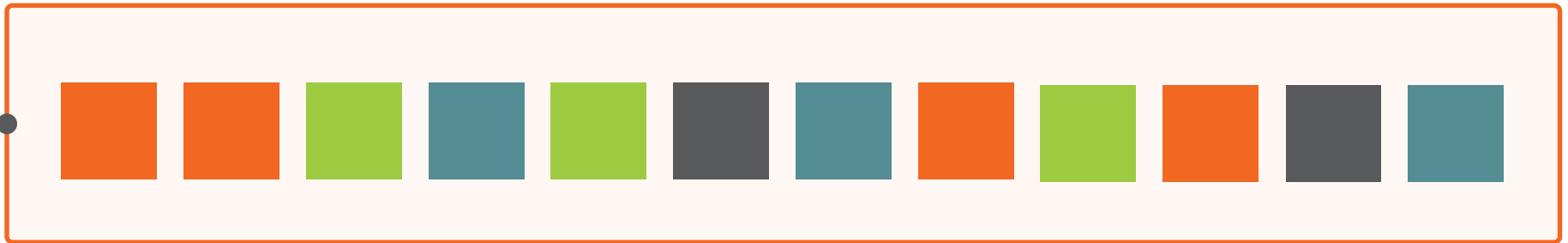
Tables



Region
Servers
x4



Regions
x64



HFiles
x128+

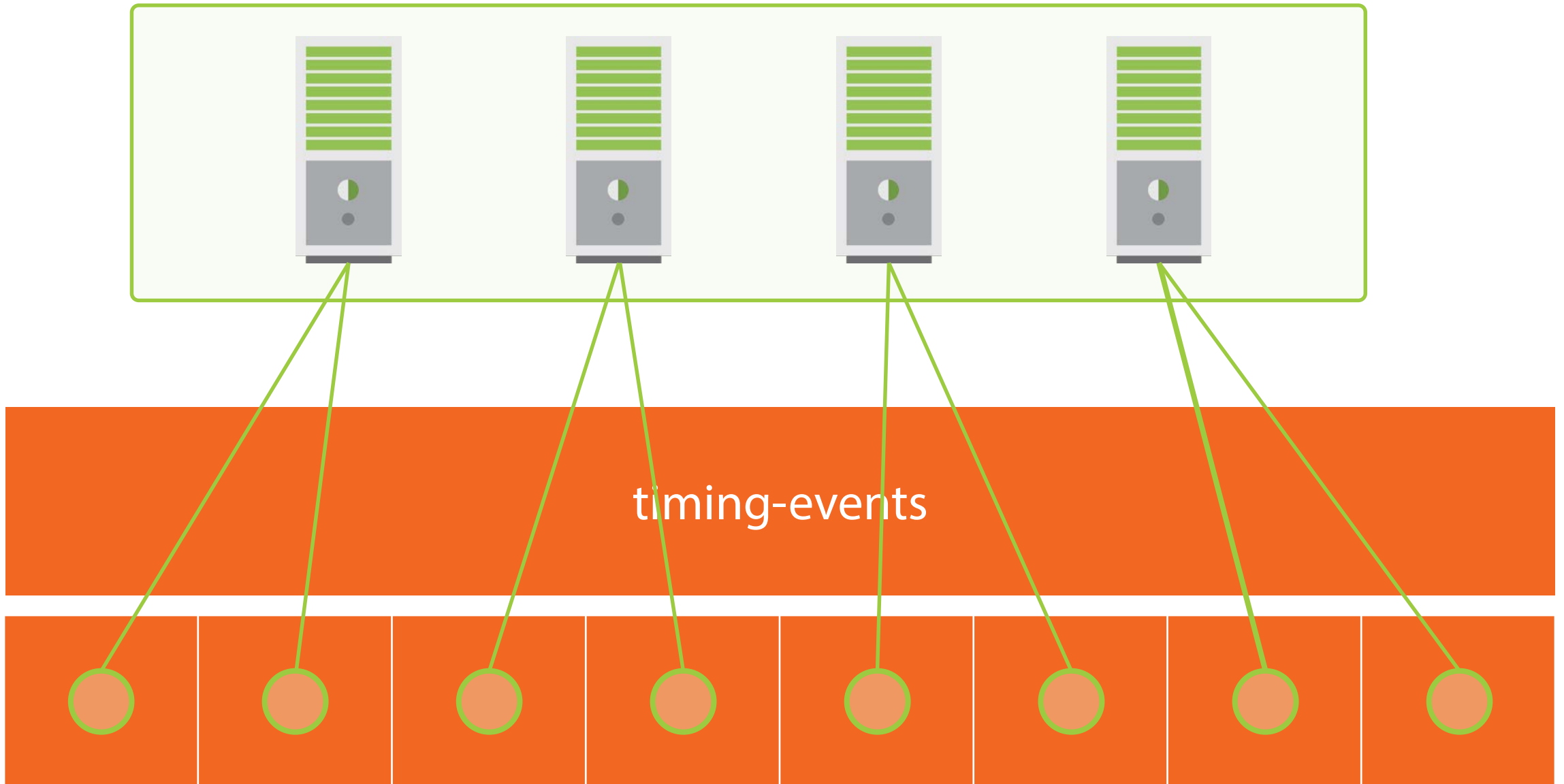


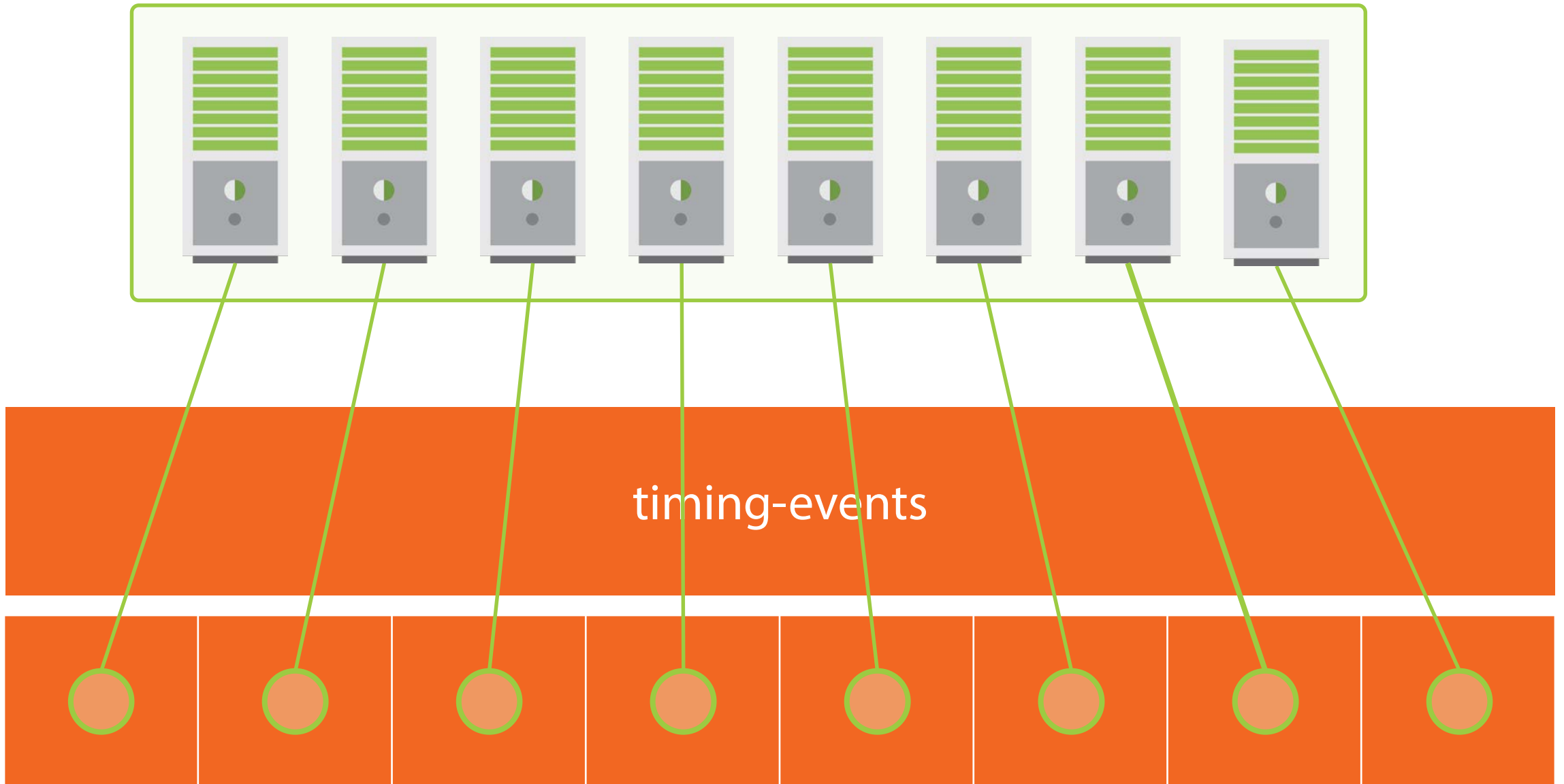
races

race-timers

sector-times

timing-events





timing-events

000	001	002	003	...									
					...								
						...							
							...						
								...					
									...				
									...	ffc	ffd	ffe	fff

races

race-timers

sector-times

timing-events

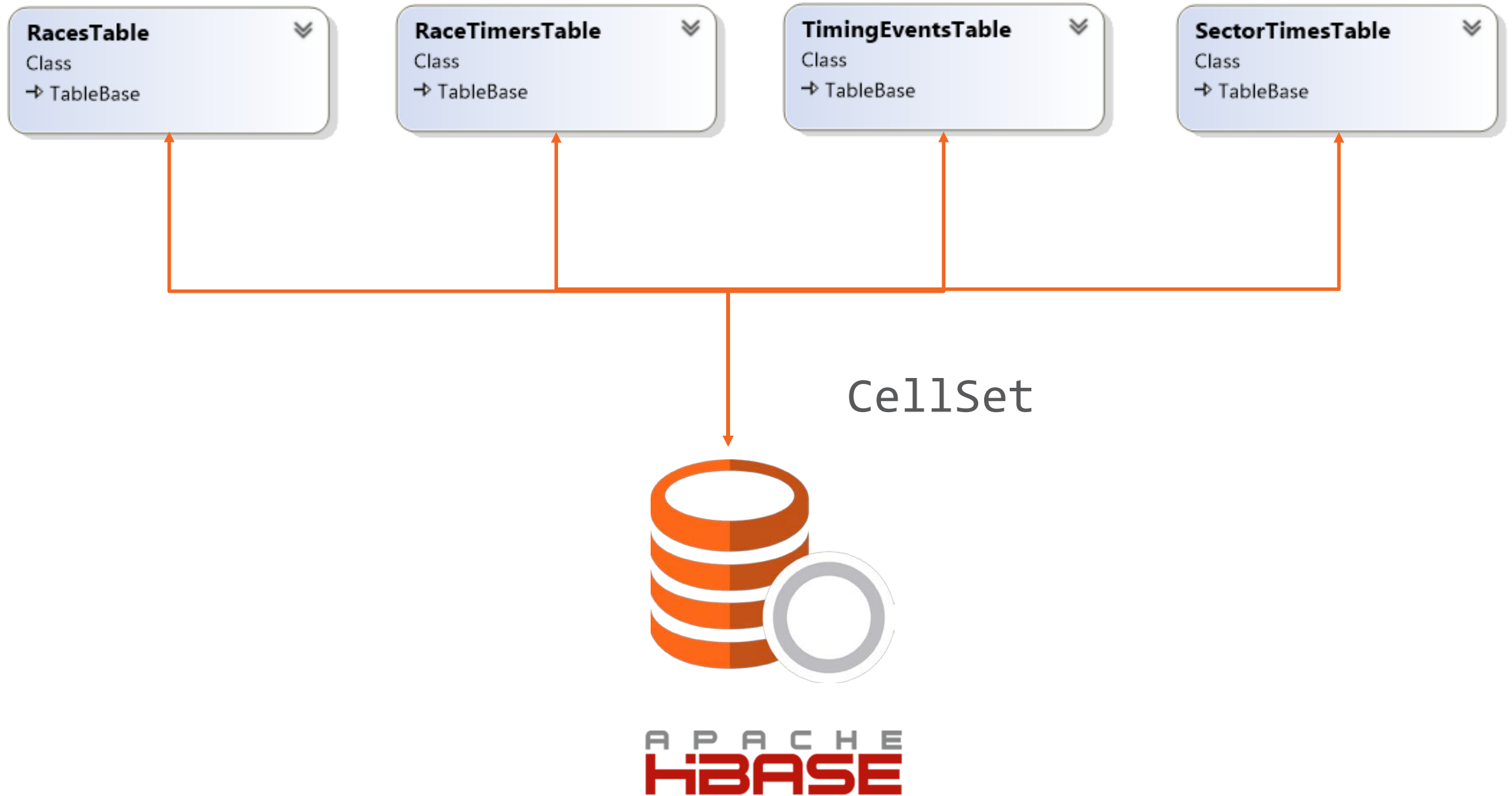
Demo: HBase and .NET

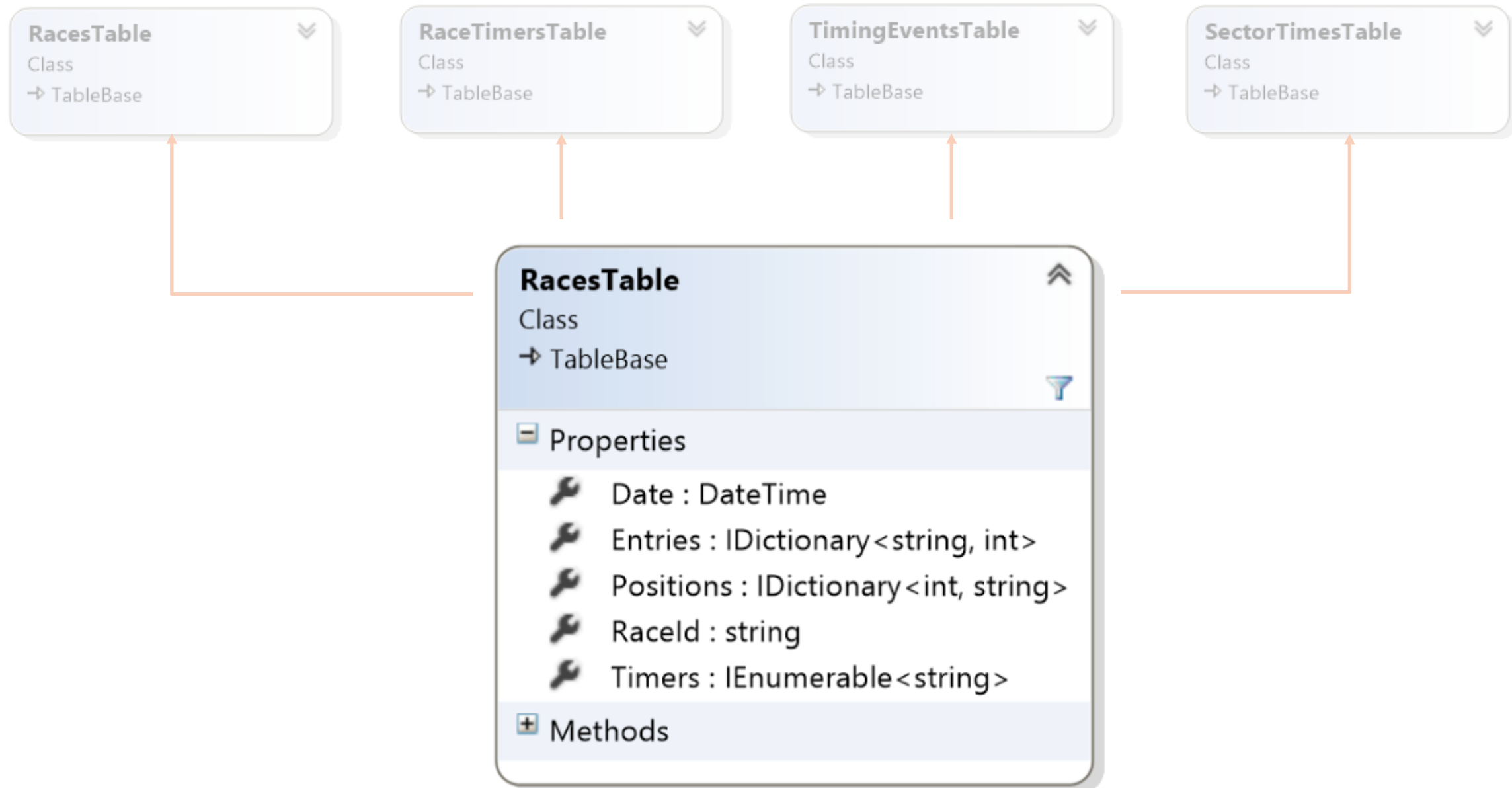
Best Practices

Modelling Schemas

Modelling Access Patterns







RacesTable

Class

→ TableBase

Properties

- 🔑 Date : DateTime
- 🔑 Entries : IDictionary<string, int>
- 🔑 Positions : IDictionary<int, string>
- 🔑 RaceId : string
- 🔑 Timers : IEnumerable<string>

Methods

races

d

t

p

e

RowKey

d:utc

t:1

p:1

e:f

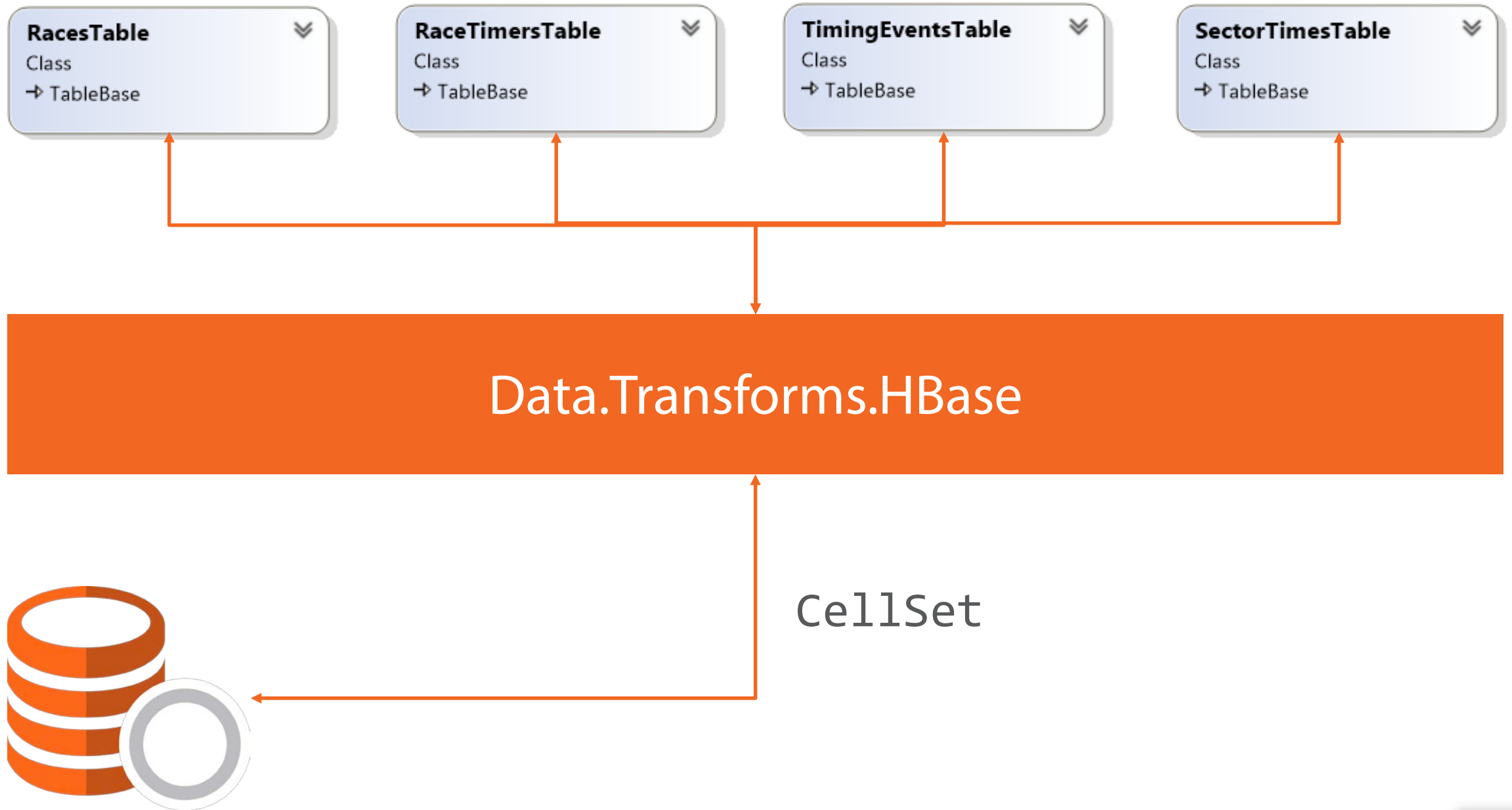
a6545da436

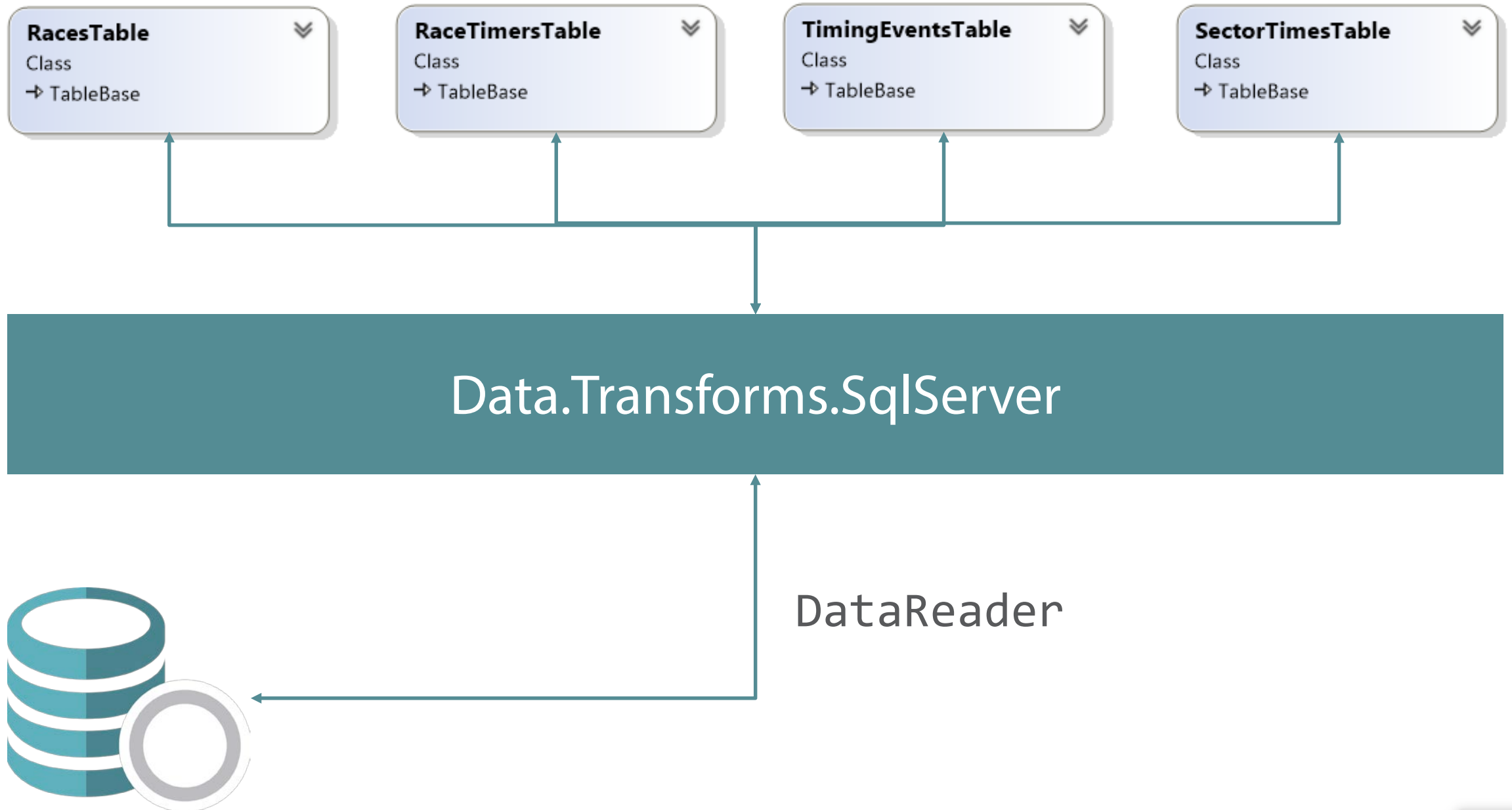
1231412412

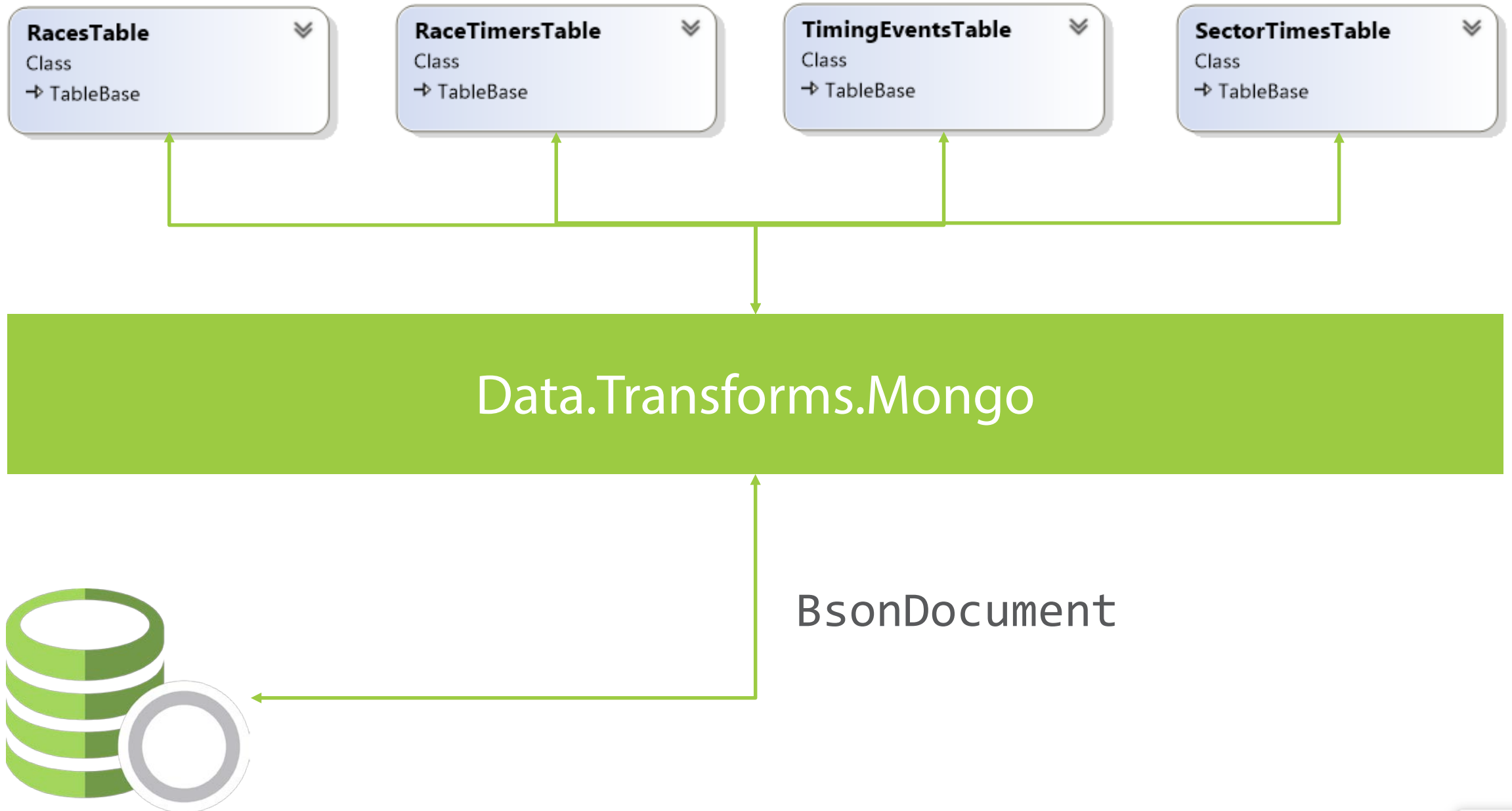
5a51322

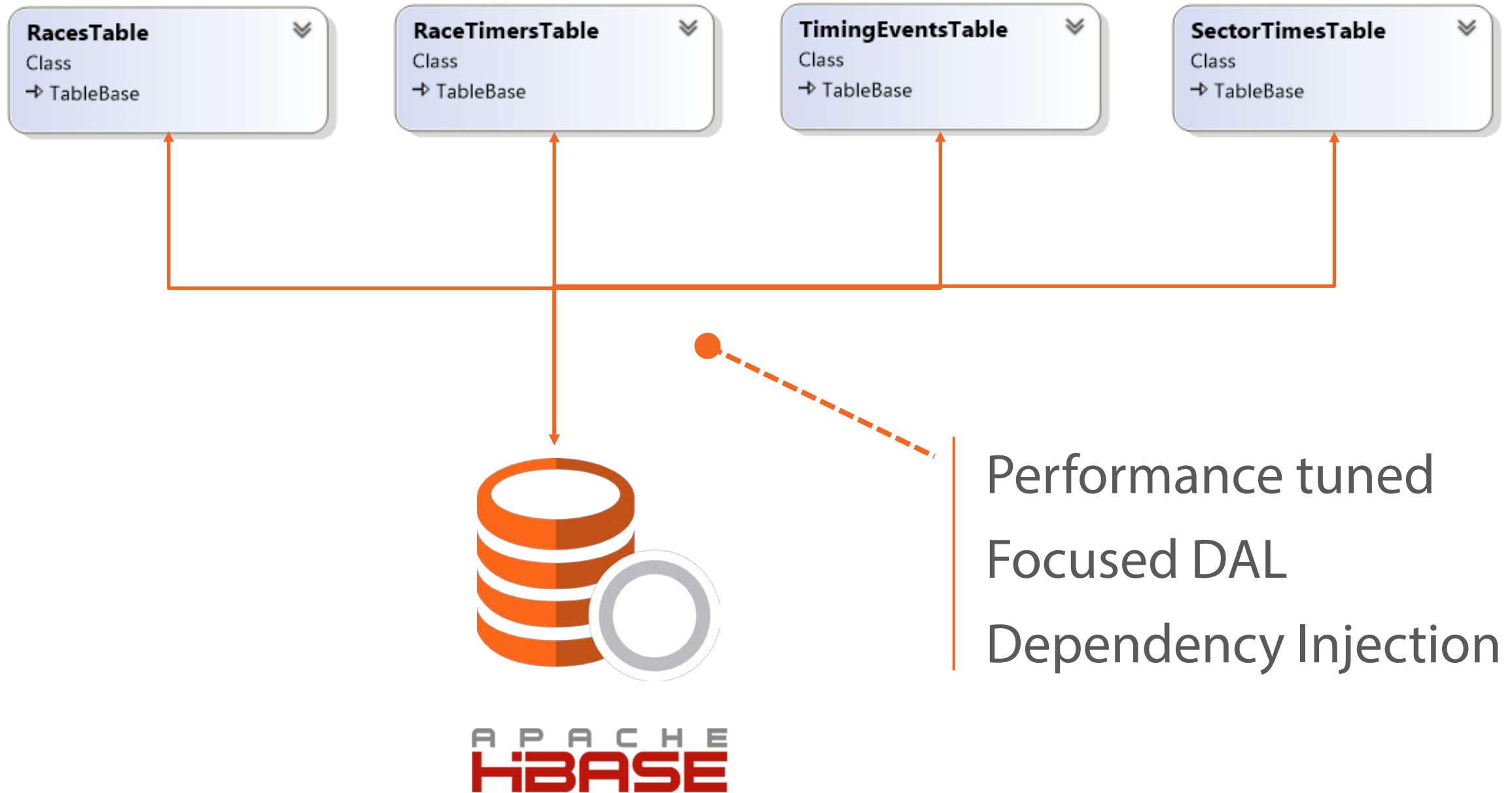
e4324

1









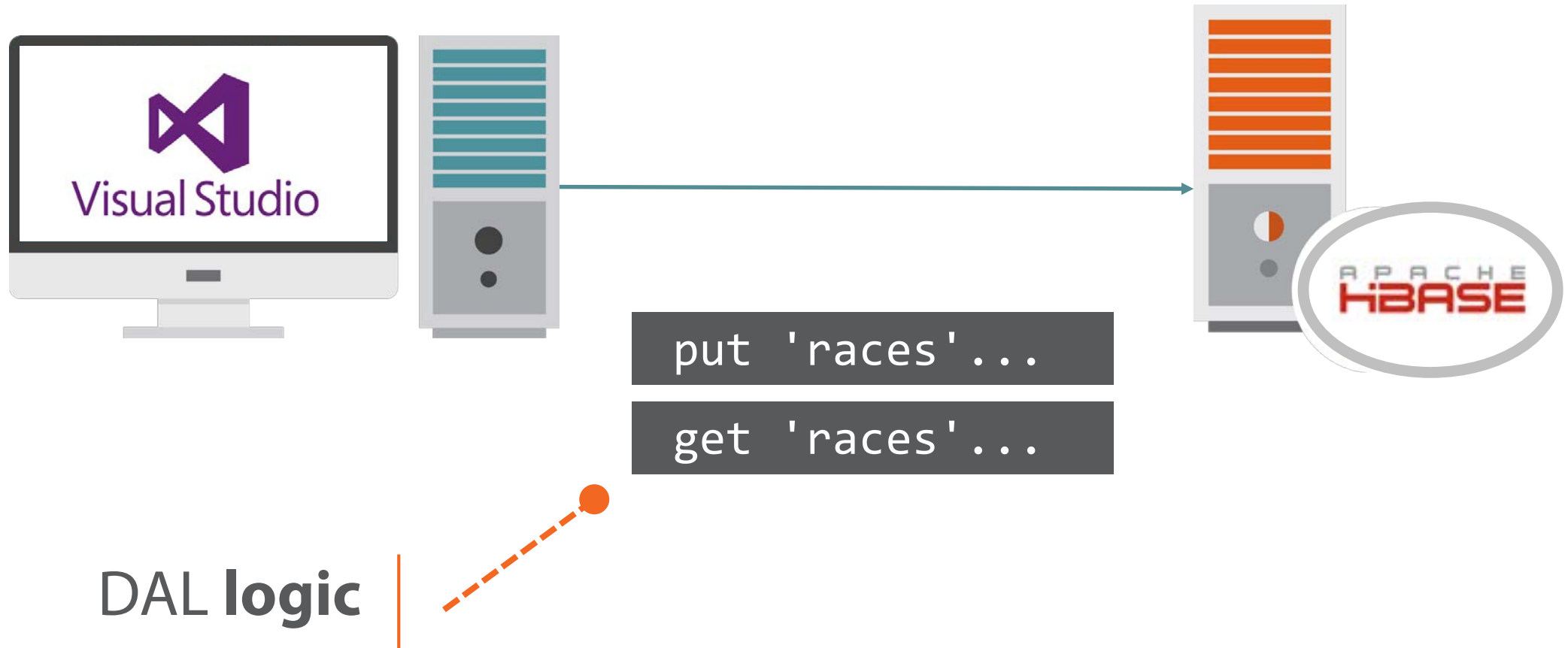
Demo: Automated Testing

Integration Tests

HBase in Docker

Mocking Stargate







Assert [duration < 1sec]

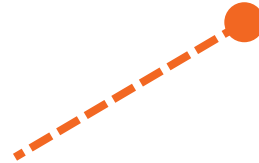




Mock<  >

```
put 'races'...
```

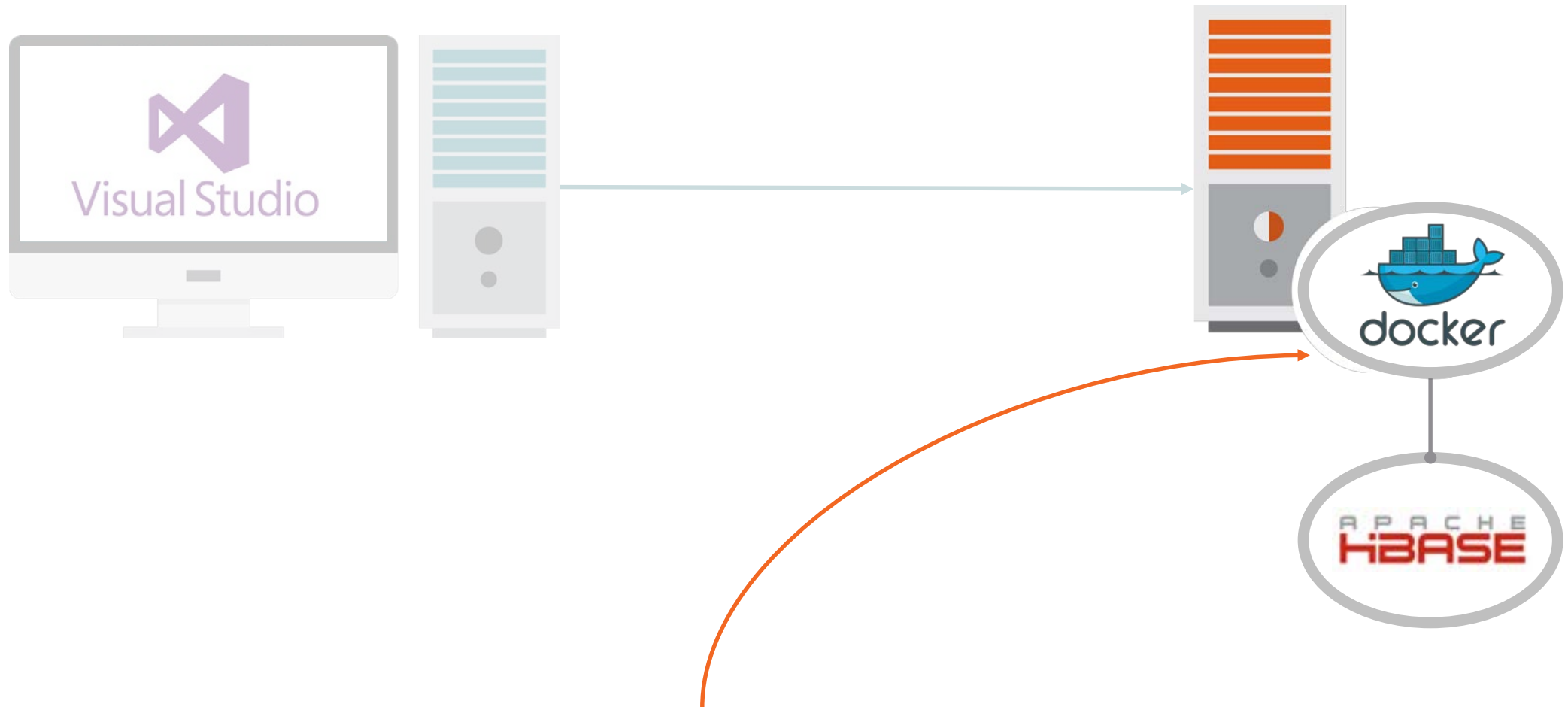
```
get 'races'...
```



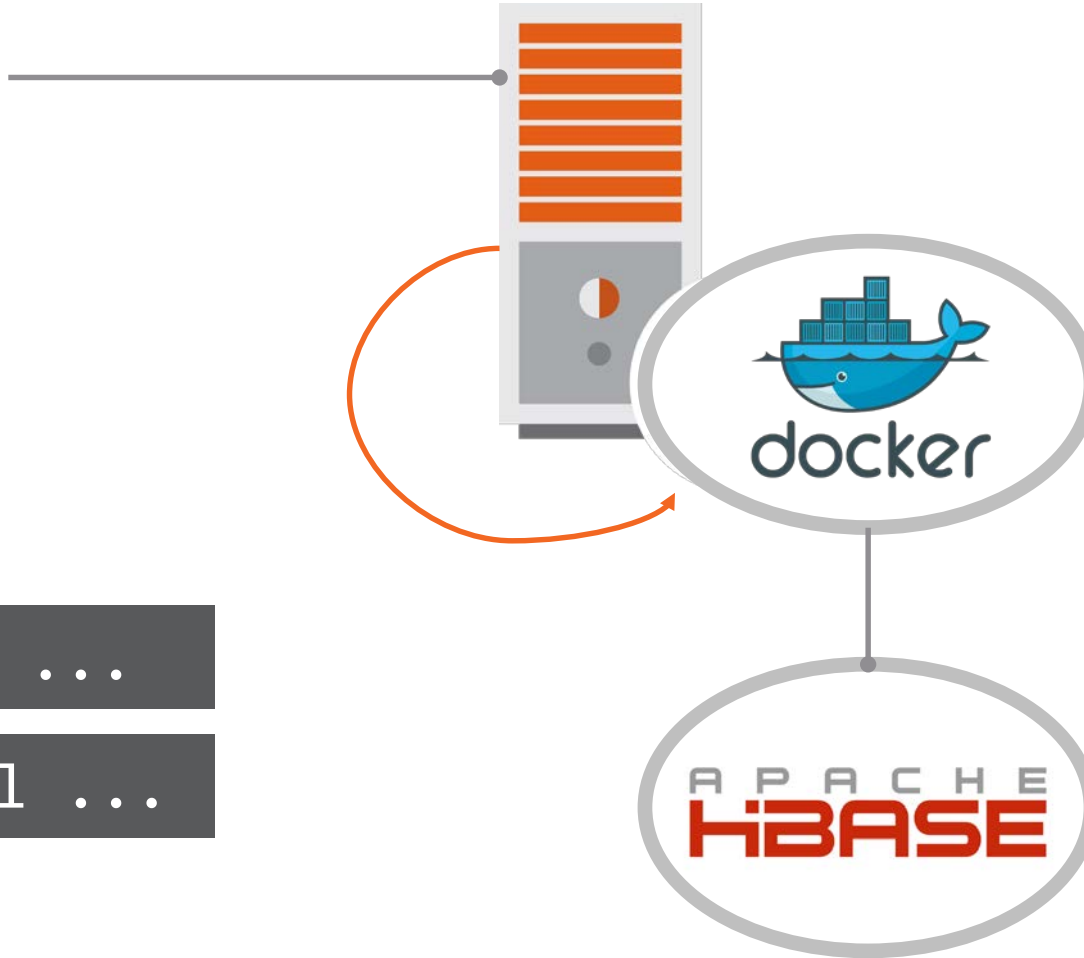
DAL logic

DI setup

App config

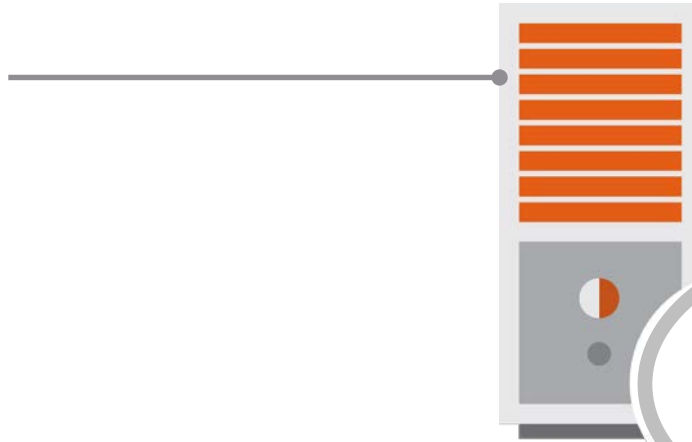


```
docker run -d -p 8080:8080 sixeyed/hbase-stargate
```

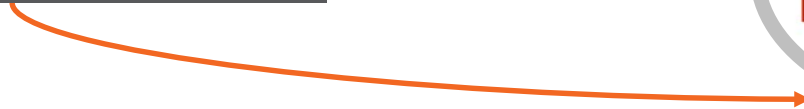


```
docker run ...
```

```
docker kill ...
```

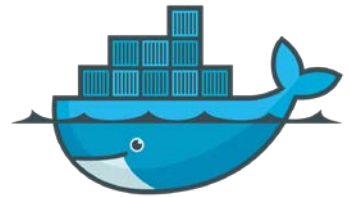
```
hbase shell ...
```



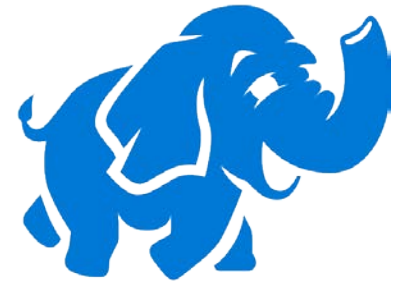
dev/build

test

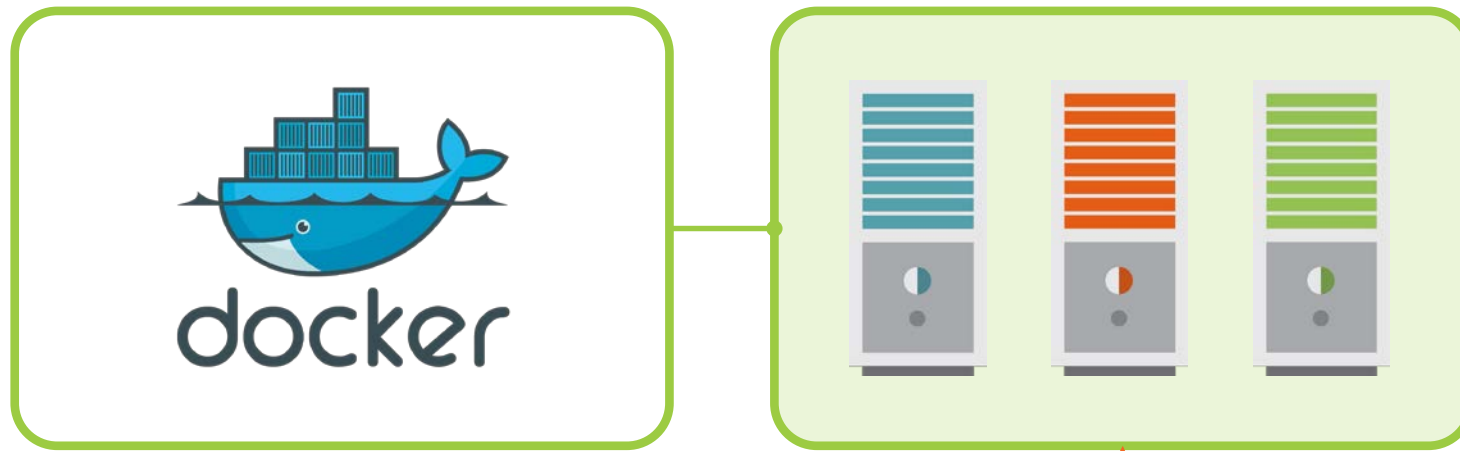
production



docker



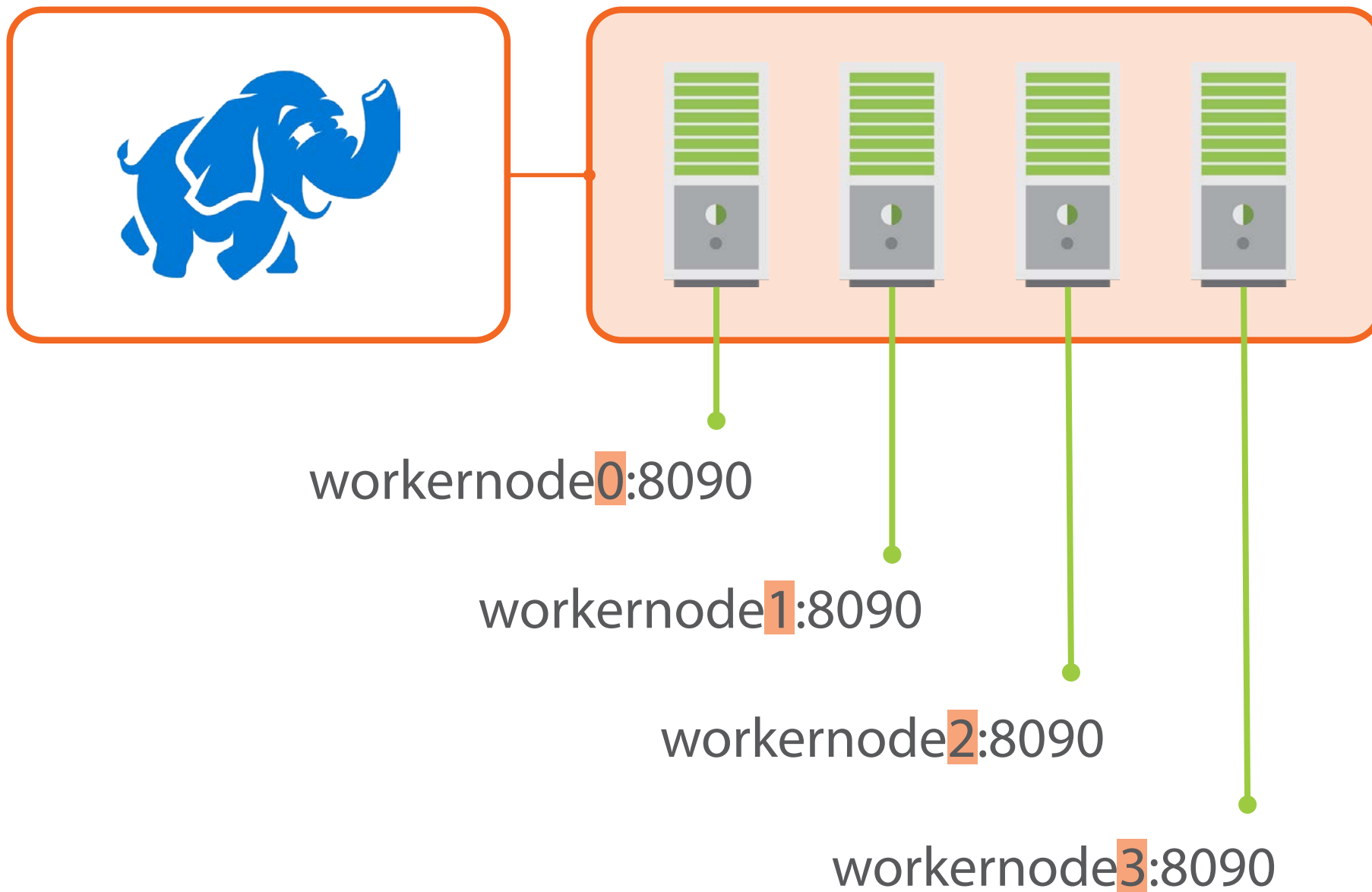
APACHE
HBASE

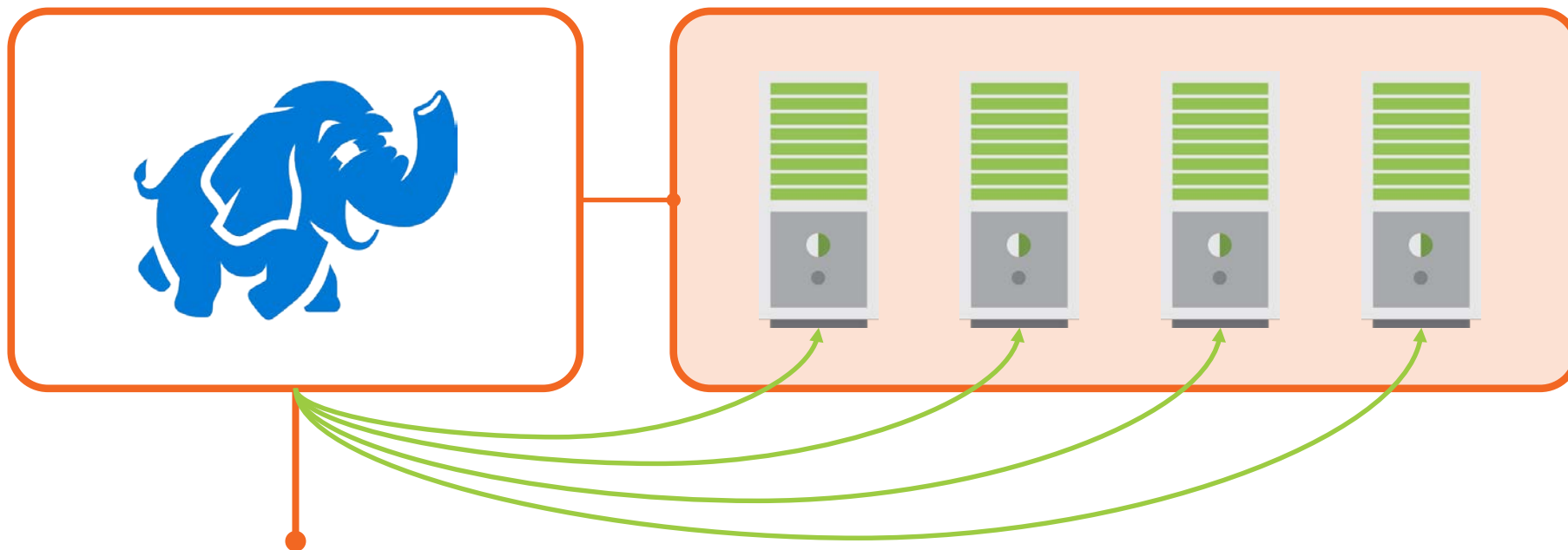


`http://10.211.5.6:8080`

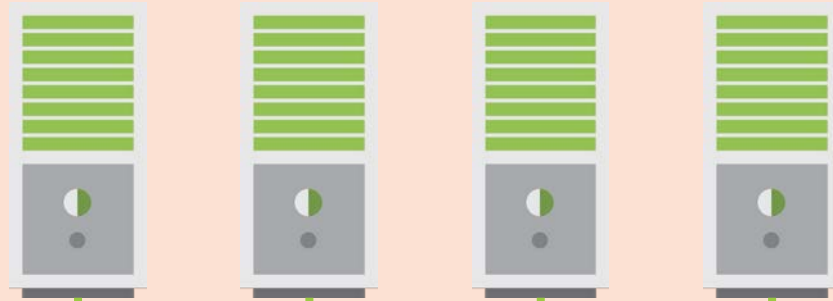
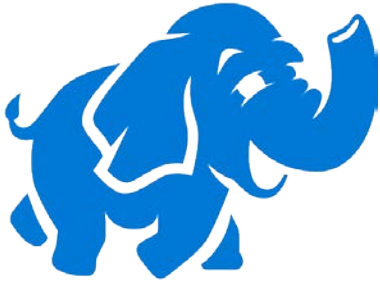


http://<?>:8080



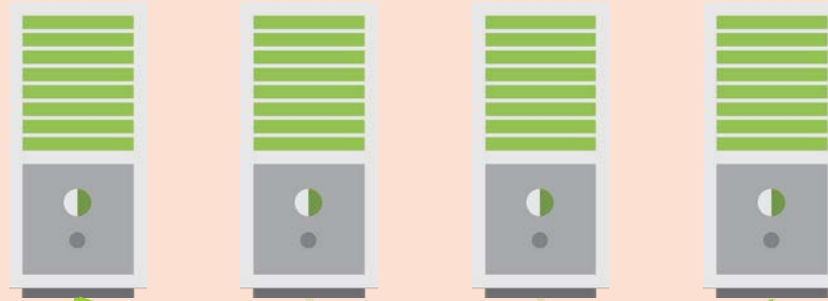
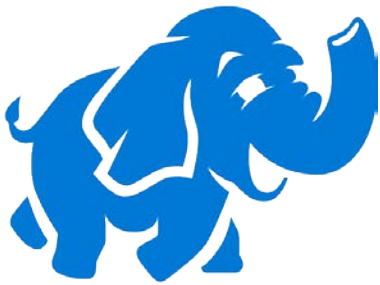


`https://<cluster>:443/hbaserest`



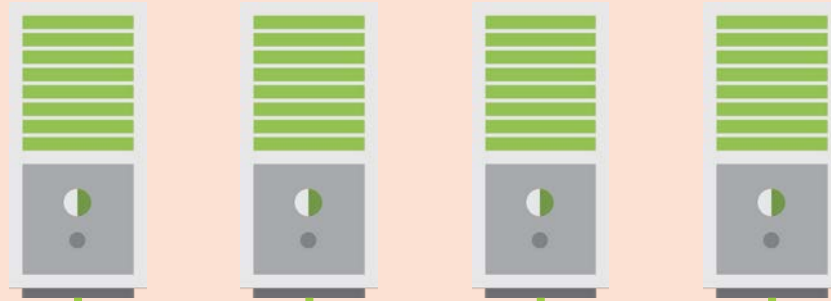
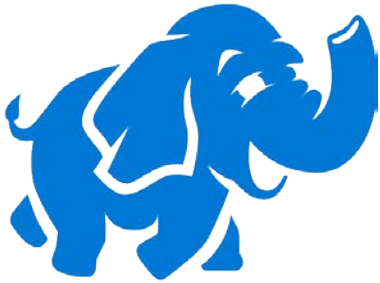
```
get 'timing-events' ...
```





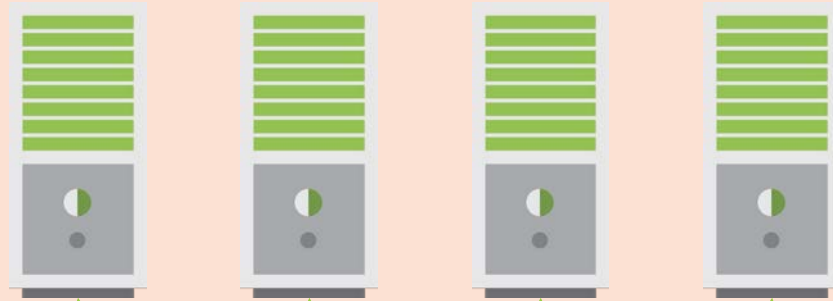
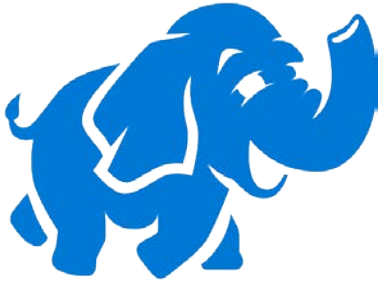
```
get 'timing-events' ...
```





```
get 'timing-events' ...
```





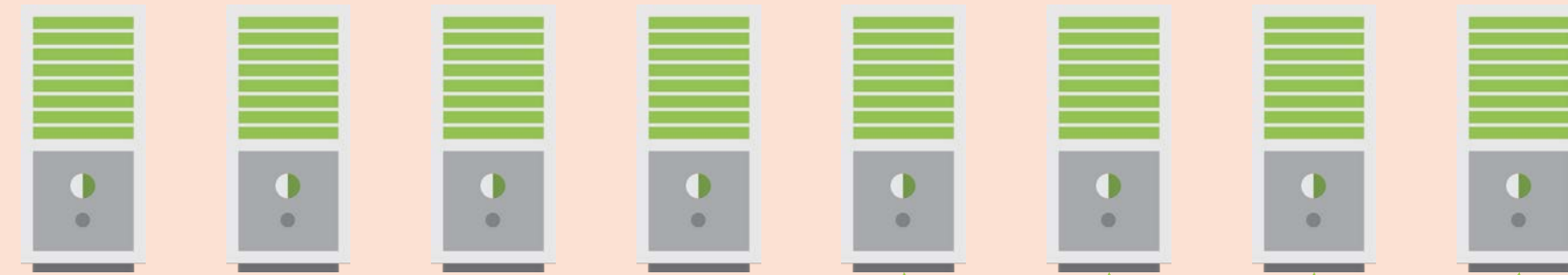
`http://<proxy>:8080`

NGINX



Nginx Reverse Proxy Config

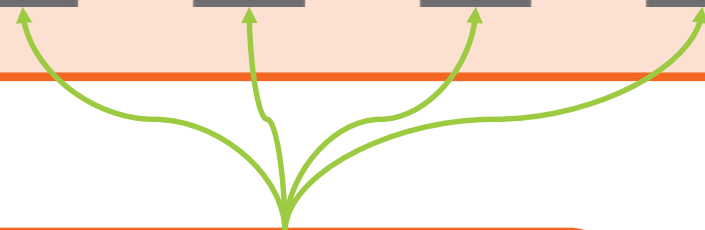
```
upstream stargate {  
    server workernode0.my-hbase.f7.internal.cloudapp.net:8090;  
    ...  
    server workernode3.my-hbase.f7.internal.cloudapp.net:8090;  
}  
  
server {  
    listen 8080 default_server;  
    server_name my-hbase-proxy;  
    location / {  
        proxy_pass http://stargate;  
    }  
}
```



NGINX

`http://<proxy>:8080`





`http://<proxy>:8080`

NGINX







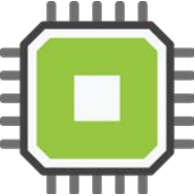


Demo: Stargate Proxy

Nginx VM

Reverse Proxy

Load Balancing

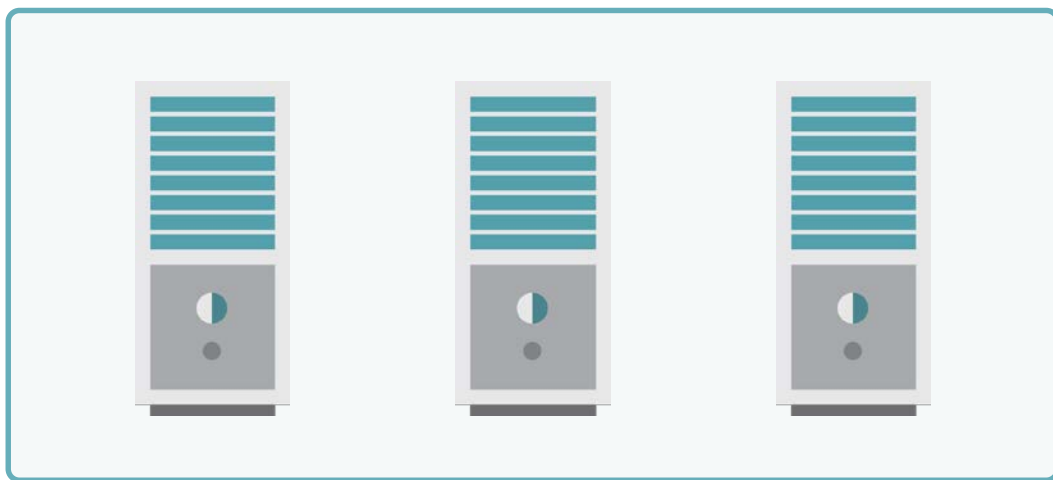
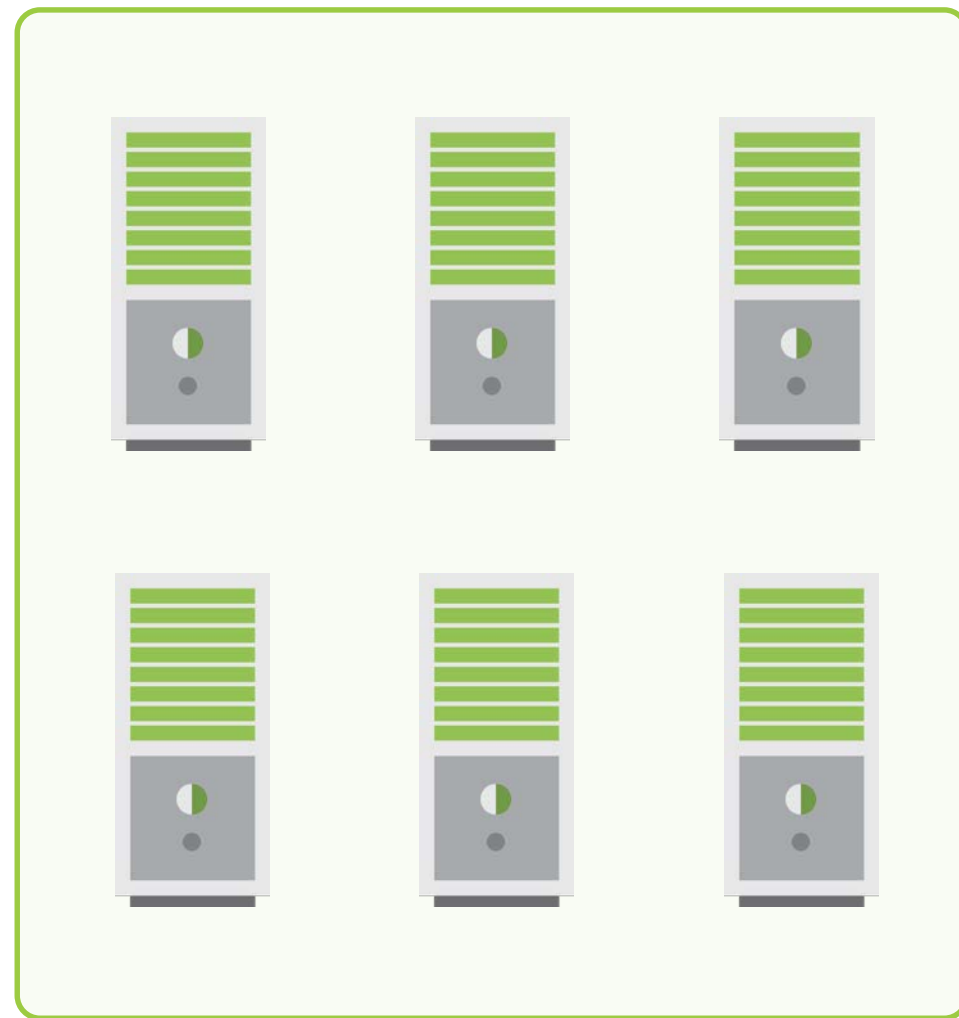
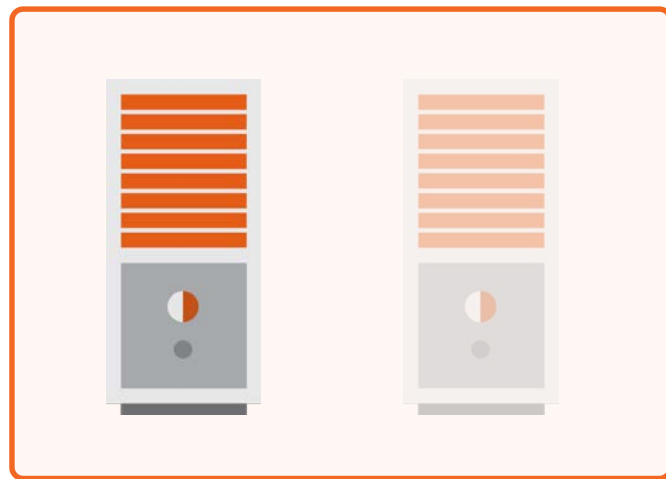


	 x4	 x2  x3  x1
	4 cores	4 cores
	14 GB	7 GB
	200 GB SSD	285 GB

NGINX



NGINX





? events/racer



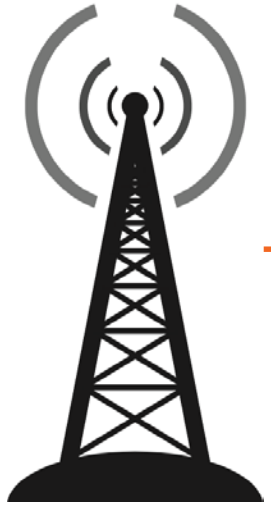
Event Processing

? events/second



APACHE
HBASE

~**800** requests/second



APACHE
HBASE

100 events/racer



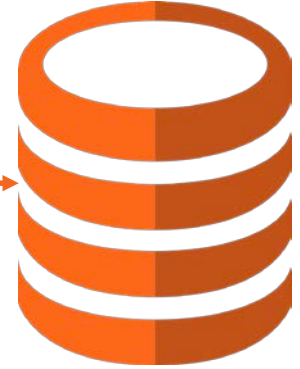
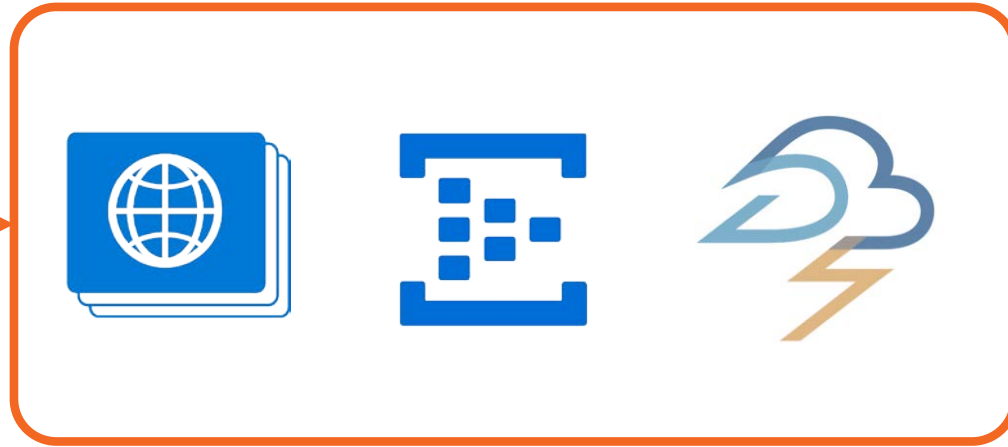
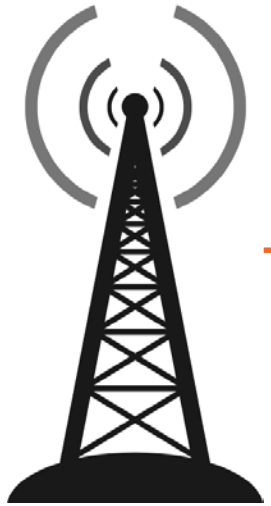
3K events/second



~800 requests/second



HBase **bottleneck**



APACHE
HBASE

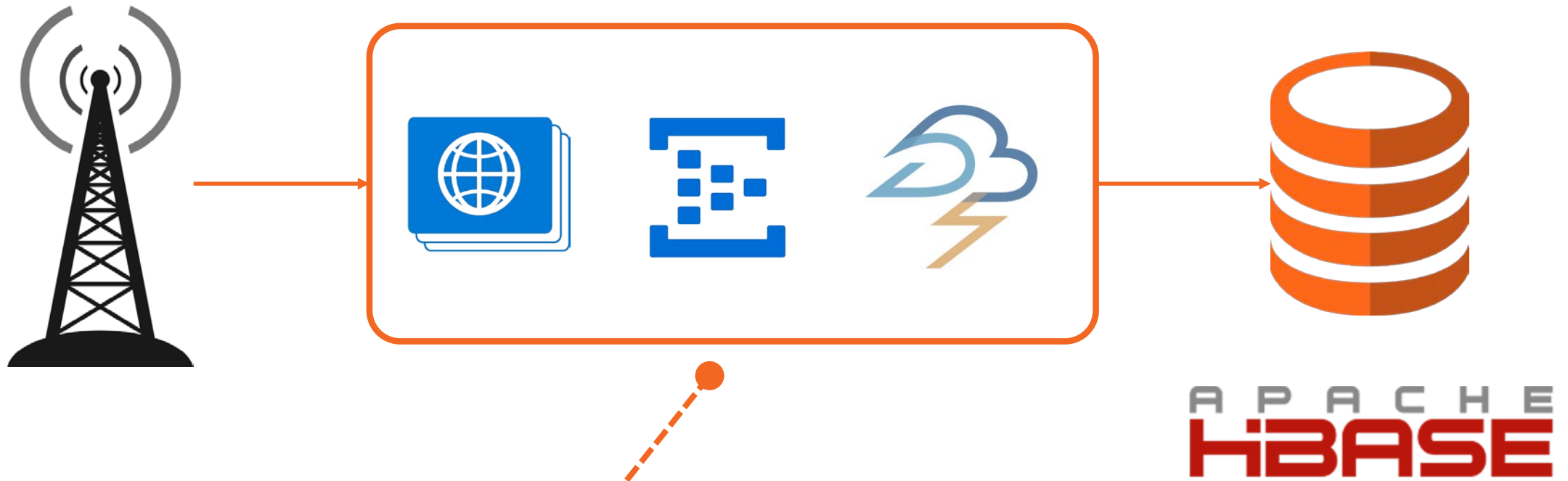
100 events/racer



3K events/second



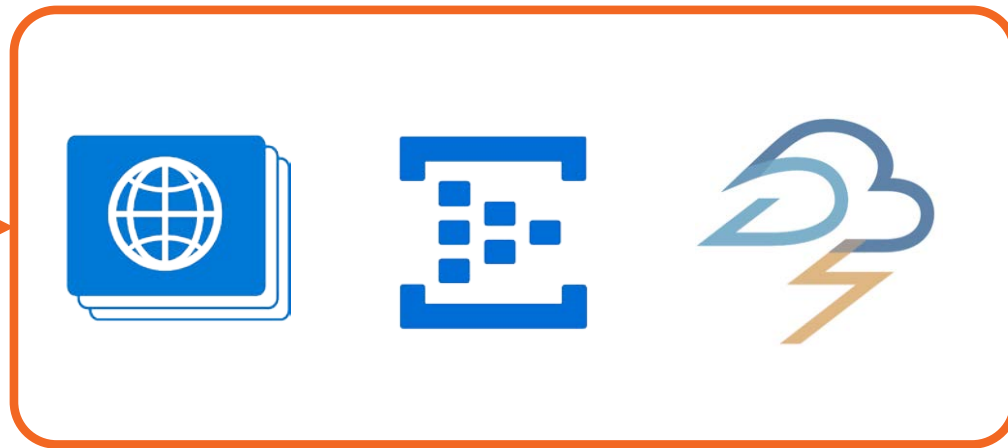
~800 requests/second



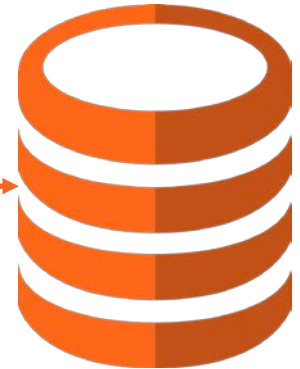
Self-regulating
Event Hub queue
Storm processing



15K events

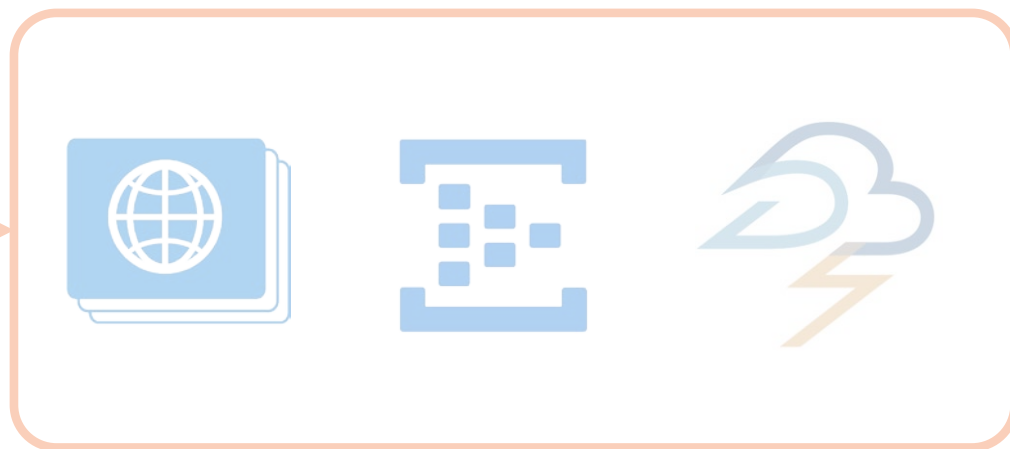


Pull **~750** events/second



APACHE
HBASE

~800 requests/second



A P A C H E
HBASE

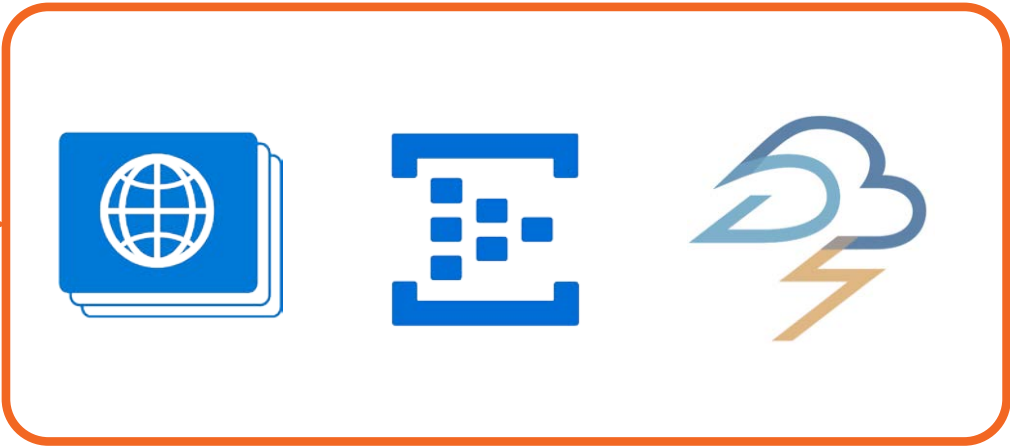
15K events

Pull **~750** events/second

~800 requests/second

```
put 'timing-events' ...  
get/put 'sector-times' ...
```

~15.5K requests



APACHE
HBASE

15K events



Pull **~750** events/second



~800 requests/second

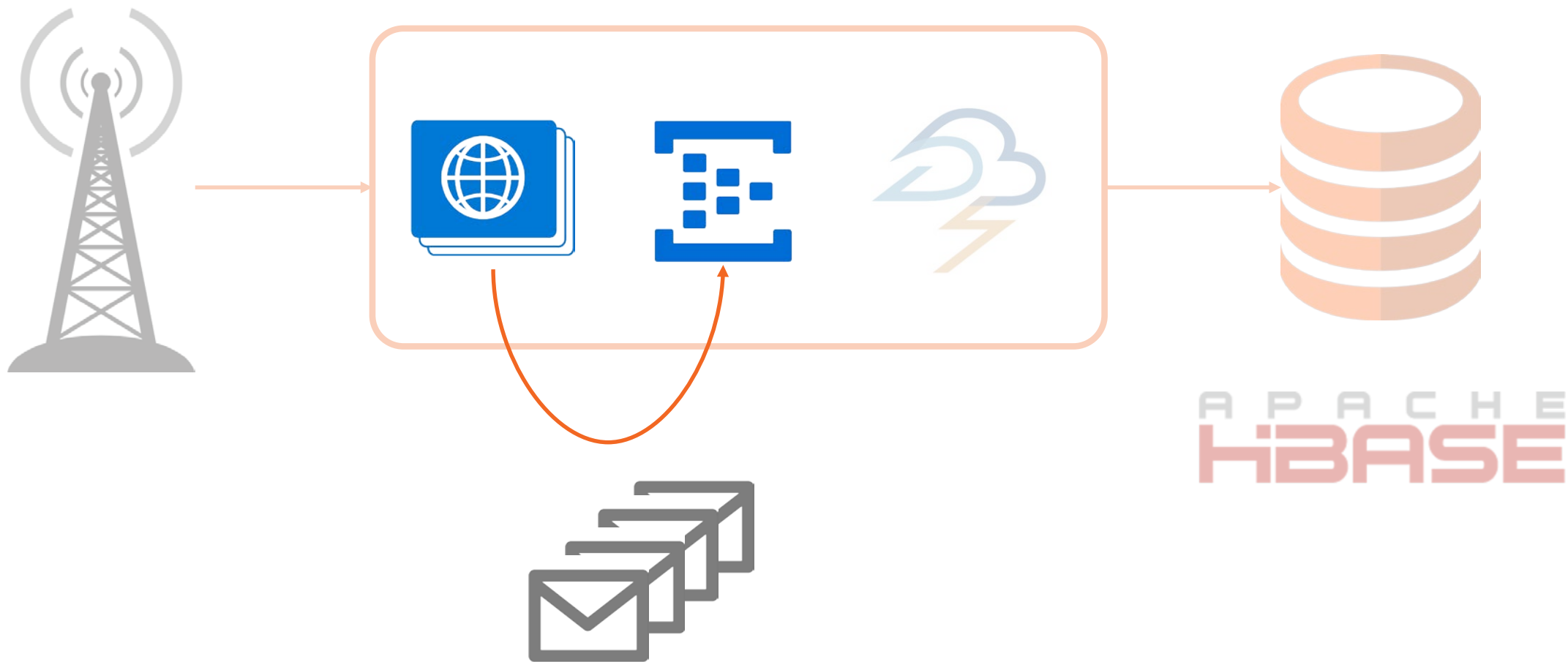
T+5 seconds



T+20 seconds



>20 seconds



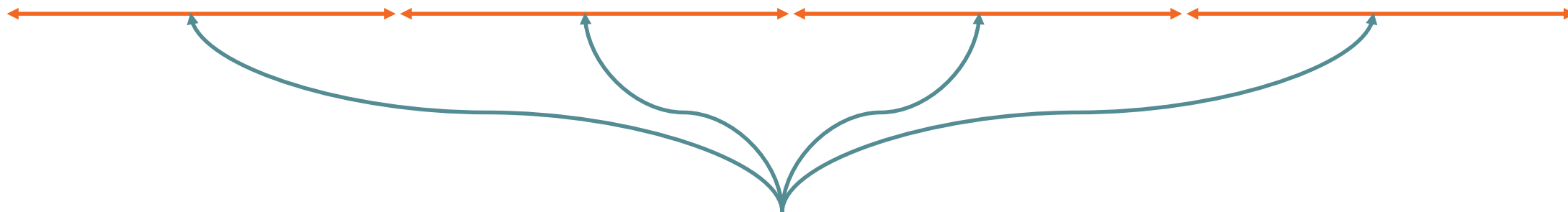


Real World Big Data in Azure



Elton Stoneman

@EltonStoneman | blog.sixeyed.com



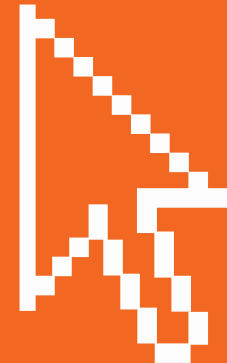
100K events
300 seconds
~333 events/second

Demo: Scaling HBase

Scale Up Cluster

Run Major Compaction

Performance Test



sector-times te d

RowKey	te	d
0000 ...		
0123 ...		
07ff ...		



sector-times te d

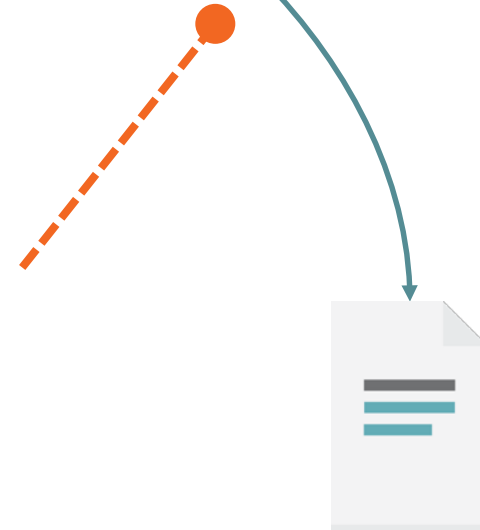
RowKey	te	d
0000 ...		
0123 ...		
07ff ...		



sector-times te d

RowKey	te	d
0000 ...		
0123 ...		
07ff ...		

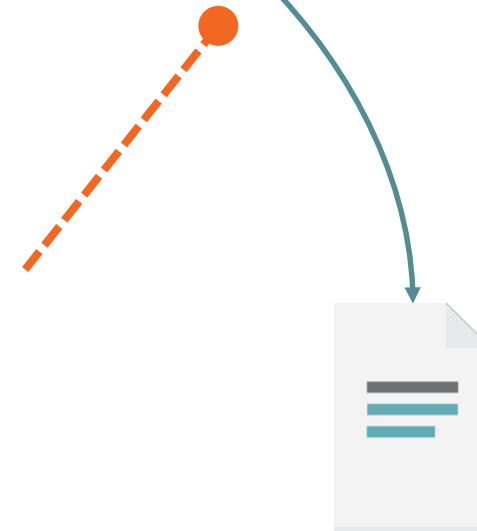
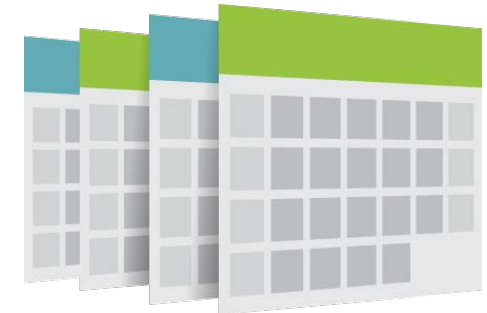
Minor compaction
Combine **HFiles**
On **buffer flush**

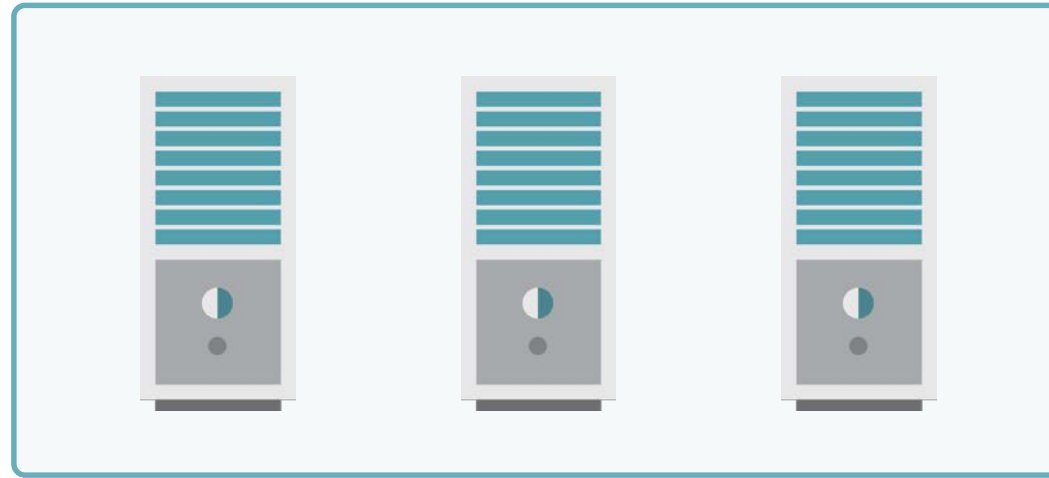
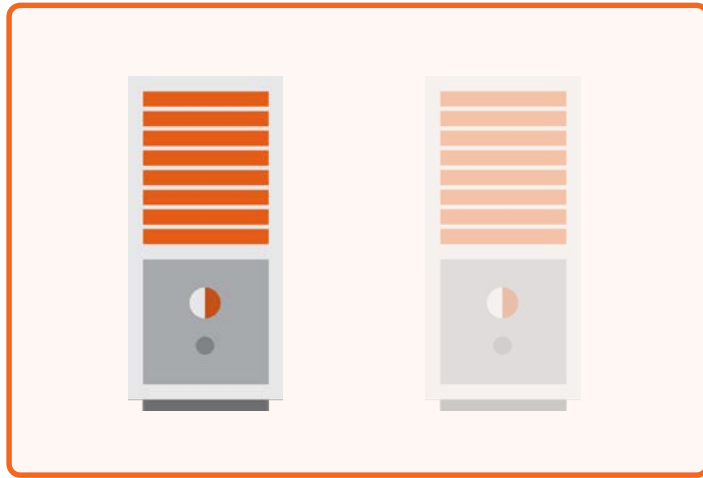


sector-times	te	d
--------------	----	---

RowKey	te	d
0000 ...		
0123 ...		
07ff ...		

Major compaction
Combine **HFiles**
Split large regions





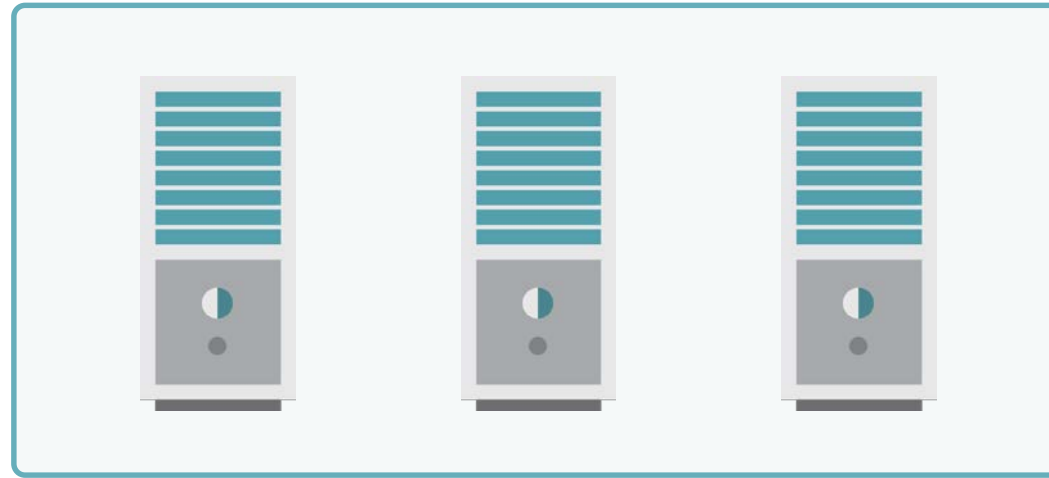
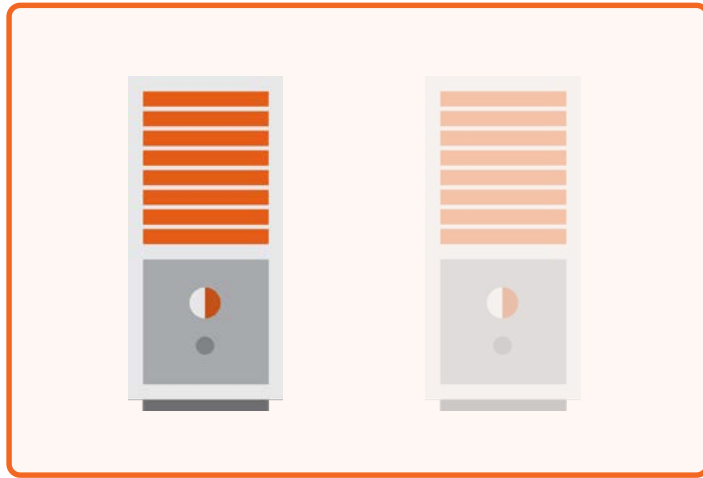
Management
x5



Data
x4



Region Servers = **44%** of compute



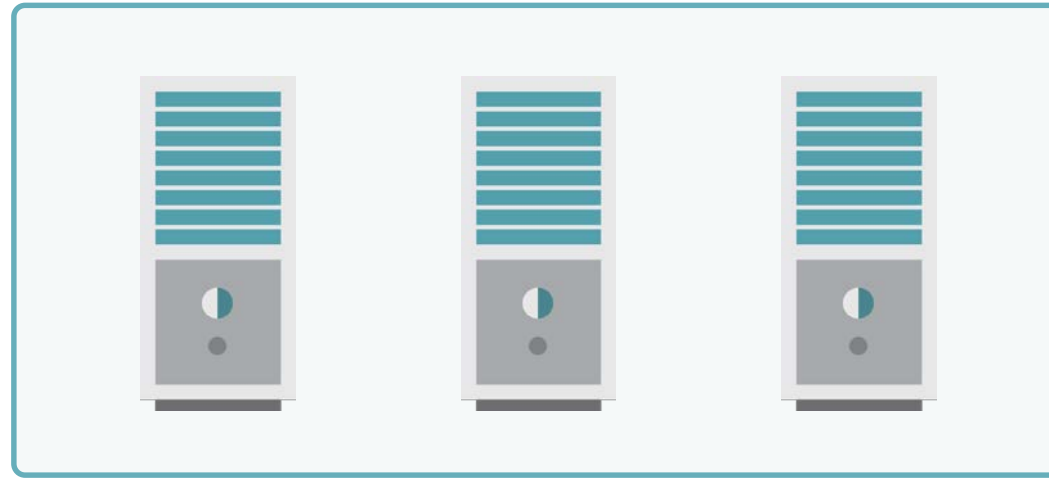
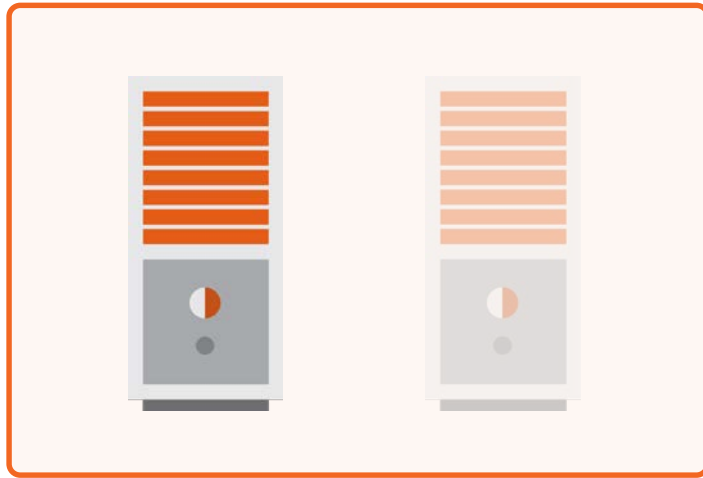
Management
x5



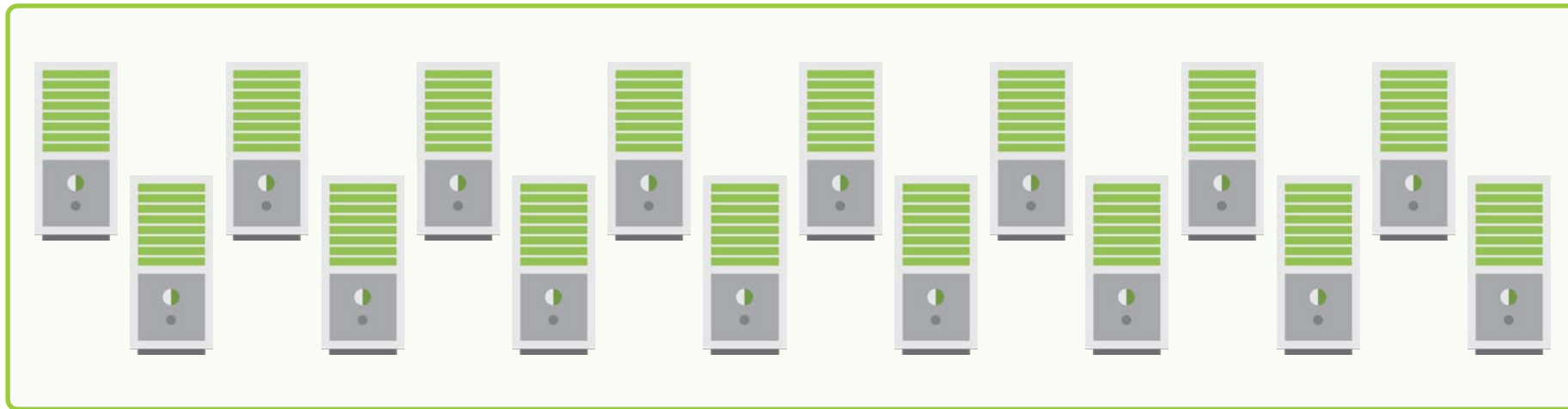
Data
x8



Region Servers = **62%** of compute



Management
x5

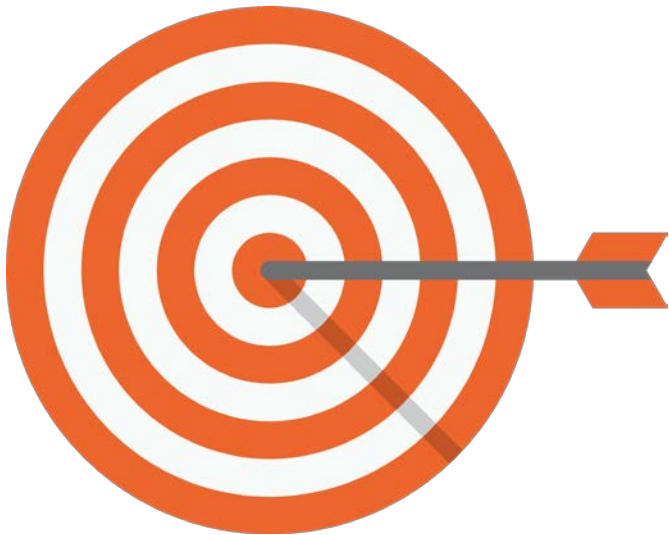


Data
x16



Region Servers = **76%** of compute

Module Goals



Cluster Setup & Region Servers



Regions & HFiles



Load Balancing & Scaling



.NET Best Practices



HBase Administration



sector-times **te** **d**

RowKey	te	d
0000 ...		
0123 ...		
07ff ...		



HFile compression

sector-times te d

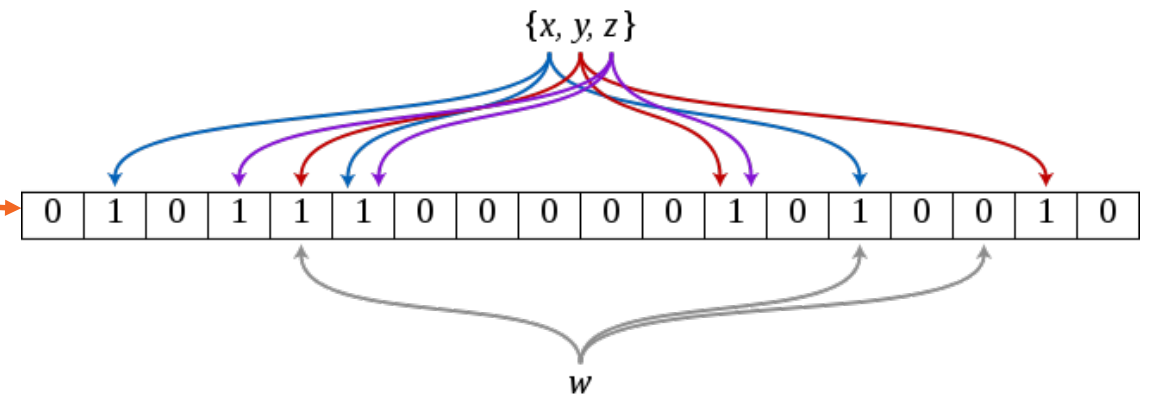
RowKey	te	d
0000 ...		
0123 ...		
07ff ...		



Column Family compressed
GZip, LZO or Snappy
Less disk **I/O**

sector-times te d

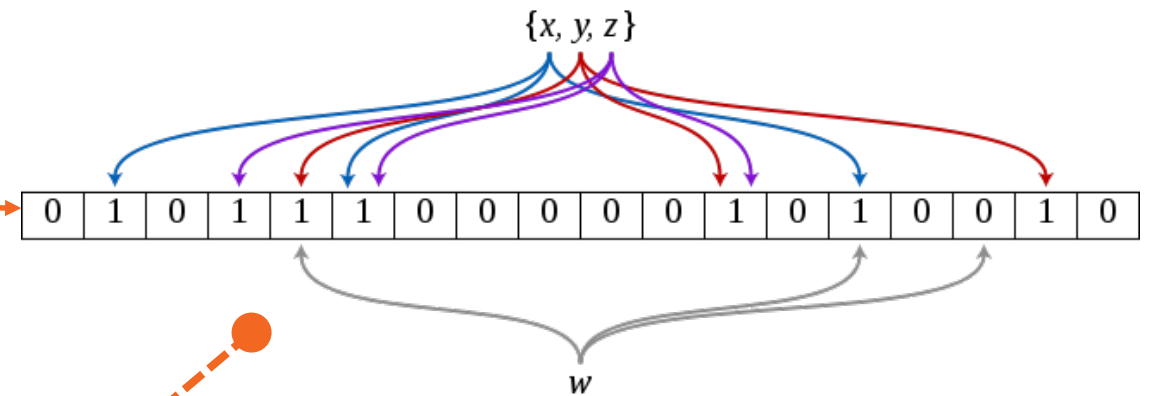
RowKey	te	d
0000 ...		
0123 ...		
07ff ...		



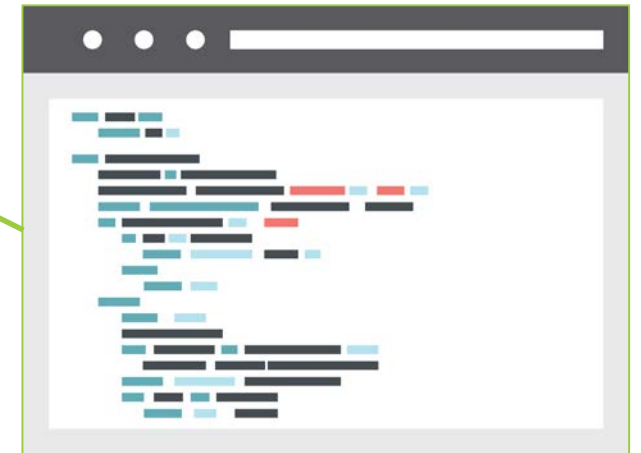
Bloom filters

sector-times te d

RowKey	te	d
0000 ...		
0123 ...		
07ff ...		

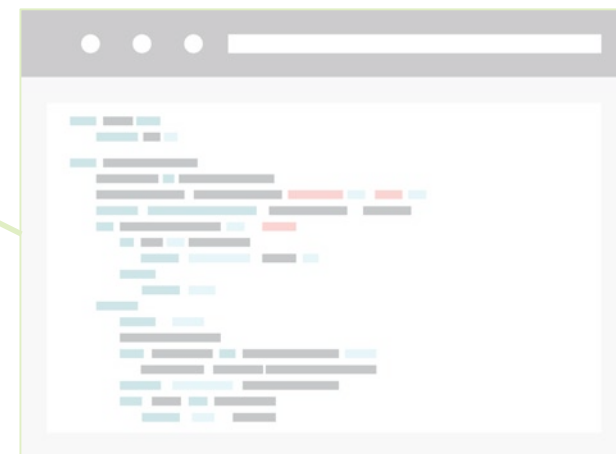
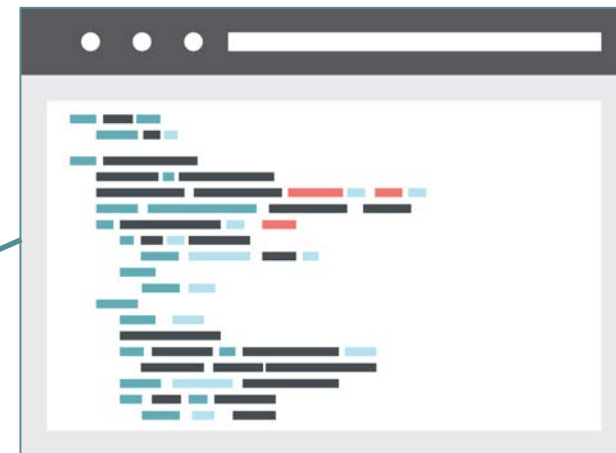
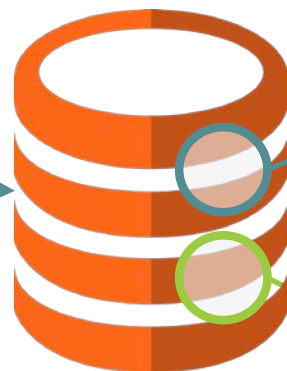


Row or **column** level
Additional indexing
Less disk **I/O**



Coprocessors

create 'races'...

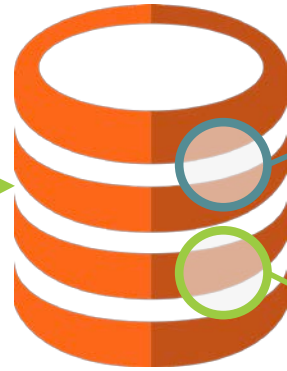


APACHE
HBASE

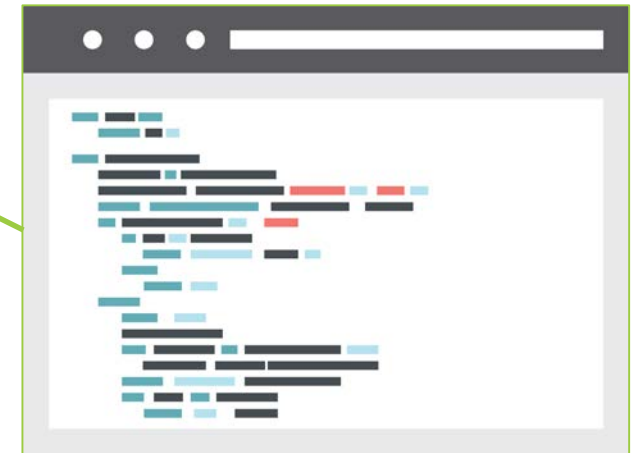
Custom code
Event triggered



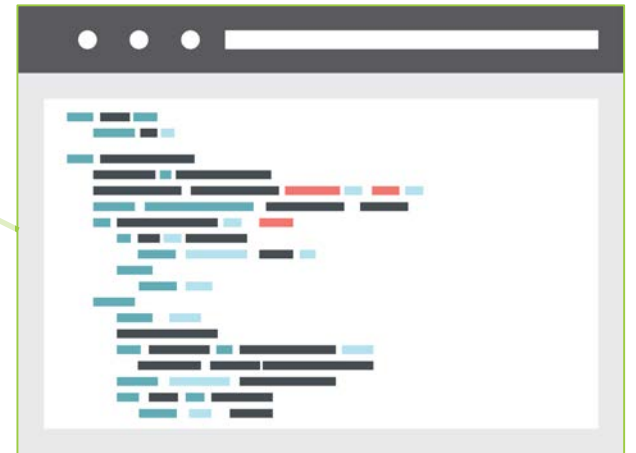
```
put 'races'...
```

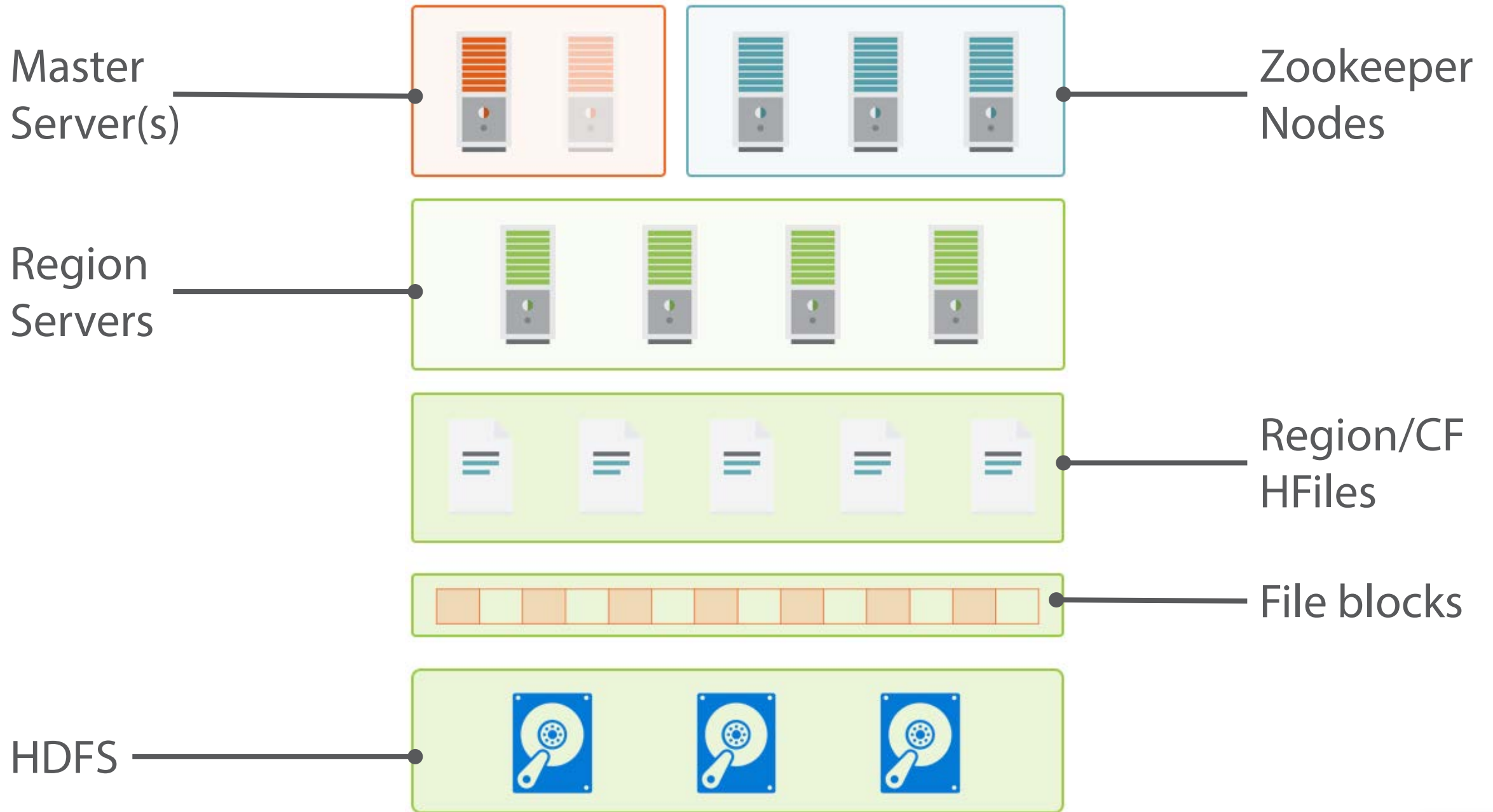


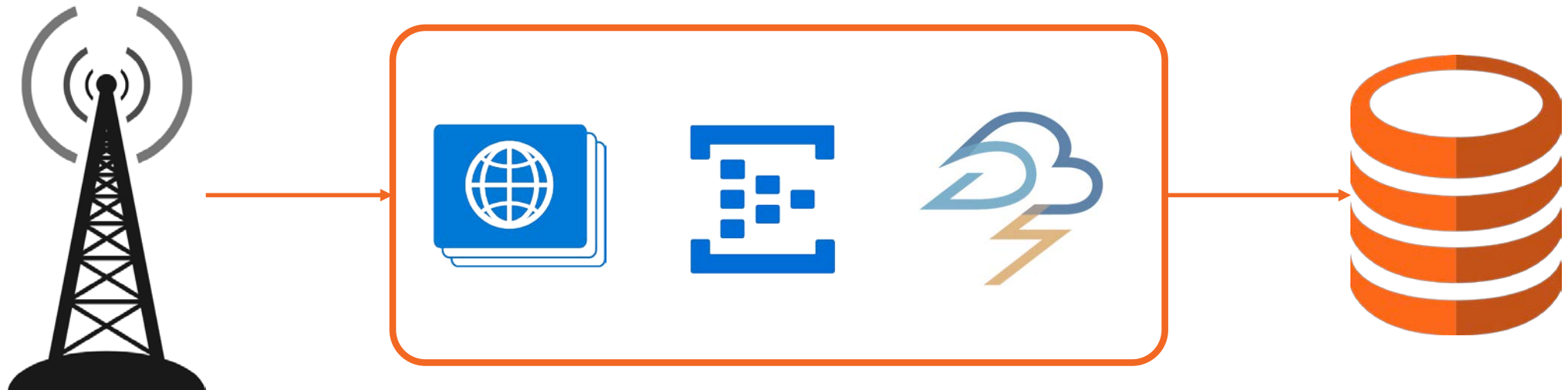
APACHE
HBASE



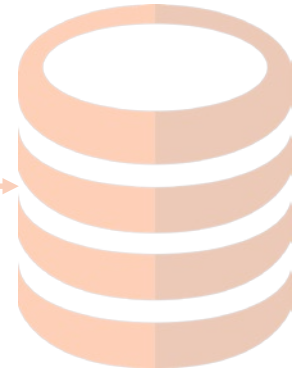
Custom code
Event triggered
Runs on **event node**



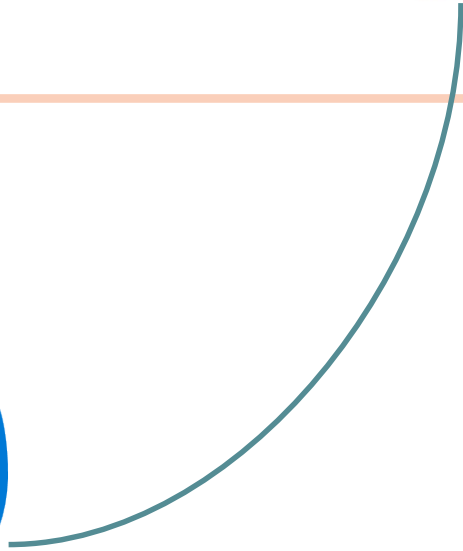
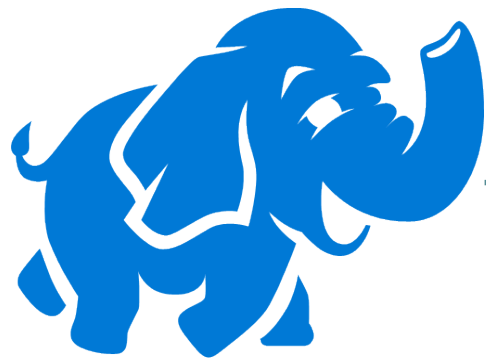




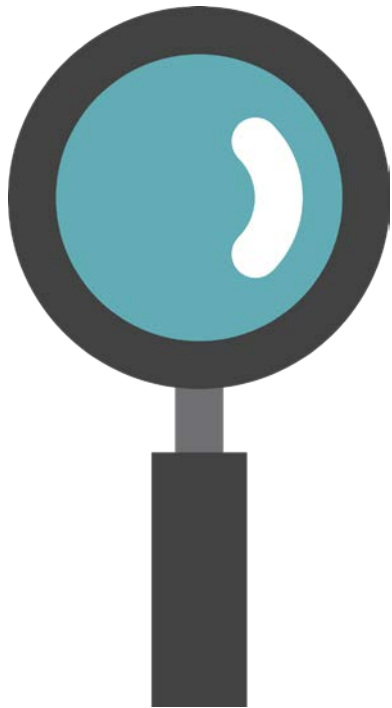
APACHE
HBASE



APACHE
HBASE



Coming Next



Event processing

Real-time

Storm on HDInsight

Azure Event Hubs