

Building a High-Performance Key/Value Store in Go

Marty Schoch
GopherCon 2017

En toen gebeurde datgene, wat hem na een moment van opslagloosheid er toe bracht, Quinten in zijn armen te nemen en moeiteloos door de brandnetels en de bloemen en langs de groene stammendams terug te hollen naar het kasteel. Quinten zat zijn vinger uit naar het pilaarje, leunde een beetje achterover en zei met een lach:

'Obelisk.'

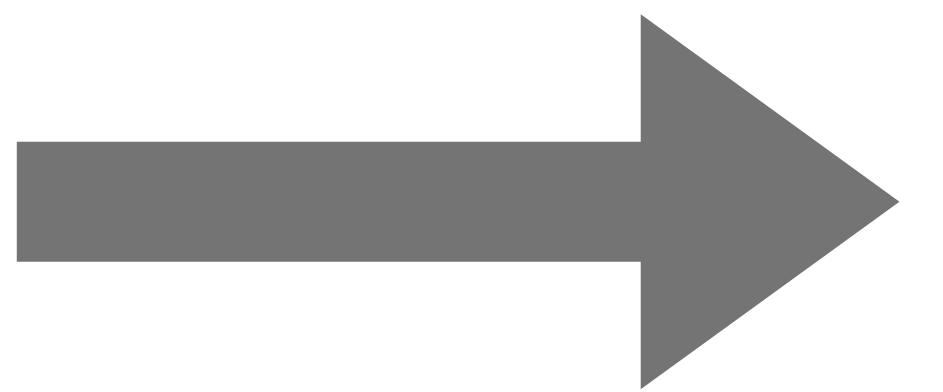
Building a High-Performance Key/Value Store in Go

Zoals in Naardwijk het licht van de vuurtoren langs de vier windrichtingen zweerde, zo strekken elk jaar in grootse vlaggen de vier seisoenen over Groot Rechteren. Max kende de wisseling van de jaargetijden eigenlijk alleen uit Amsterdam: op een dag in februari of maart de nette, onbeschrijflijke voorjaarsgeur als hij 's morgens op straat kwam, onbetaalbaar als de decimalen van π , de stoffige zomer waarin de stad vol roerlusten stroonde, even plotseling veranderend in de vuchige, bittere herfst, en dan de bleke winter waarin de steentjes van de straten en de muren opeens de ongemaakbare aard van de wereld leken uit te drukken, - maar dat alles eigenlijk in het voorbijgaan, opgemerkt alleen in de korte tussenposes waarin hij van het ene interieur naar het andere ging. Was in de stad de natuur achter achtergrondmuziek, op het kasteel zat hij met Quinten en Sophia midden in de daverende concertzaal. Lente en herfst kwamen met kolosal vertoon, die temers waren heter en droger, die winters kouder en winter. Die onafgebroken verandering, had hij nu tegen Oslo beweerd, was natuurlijk de bron van alle creativiteit; de onvermoeidheid van de natuur tussen de keerkringen leidde ook tot culturele stilstand, de tropen waren een onafgebroken monsun, altijd groen zoals de poolgebieden altijd wit waren, maar de gematigde breedtegraden met hun viervormigheid waren

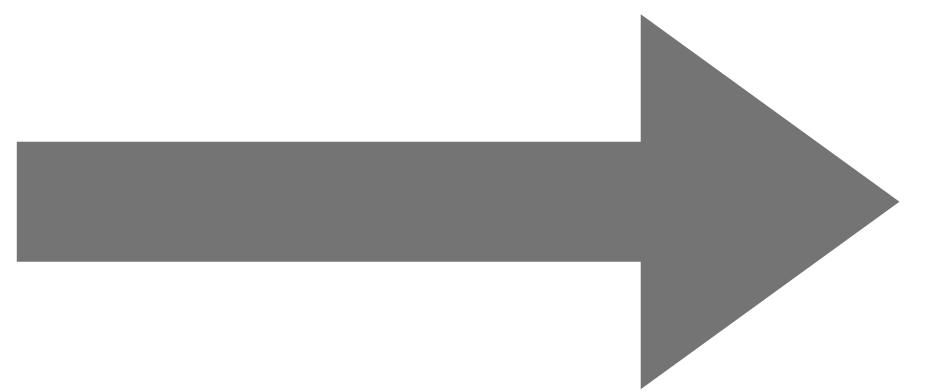
The Problem

Zoals in Noordwijk het licht van de vuurzuren langs de vier windrichtingen overvalt, zo streken elk jaar in grootsche vlagten die vier seisoenen over Groot Rechtersen. Max kende die wisseling van de jaargrijsen eigenlijk alleen uit Amsterdam: op een dag in februari of maart is er een onbeschrijflijke voorjaarsgeur als hij's morgens op straat loopt; onbeschutbaar als de decimalen van π , die stoffige somer waarin de stad vol moeststen schroomde, even plotseling veranderend in de voorjaar, herfst, en dan de bleke winter waarin de sterren van de straten en de muren opeens de ongemaakbare aard van de wereld leken uit te drukken, - maar dat alles eigenlijk in het voorjaar, oppgemerkt alleen in de korte tussenposes waarin hij van het ene interieur naar het andere ging. Was in de stad de natuur tachtig achtereenvolgende, op het kasteel zat hij met Quinten en Sjipje midden in de daverende concertzaal. Lente en herfst kwamen met kolossaal vertoon, de somers waren hotter en droger, de winters koudster en winter. Die onafgebonden verandering, had hij nooit tegen Ottus beweerd, was natuurlijk de bron van alle creativiteit: de onvermoeigheid van de natuur tussen die keerkringen hieldde ook toe culturele stilstand, de tropen waren een onafgebonden land, altijd groen zoals de poolgebieden altijd wit waren, van de primitieve breedtegraden met hun viervormigheid waren



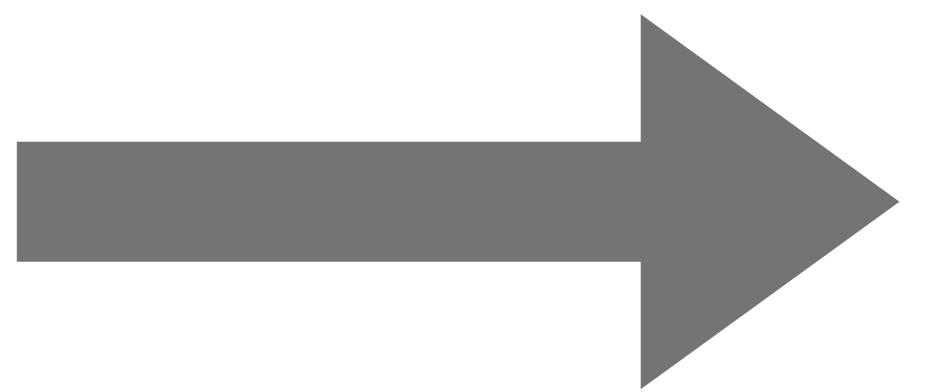


Primary
Data Source



Primary
Data Source





Primary
Data Source



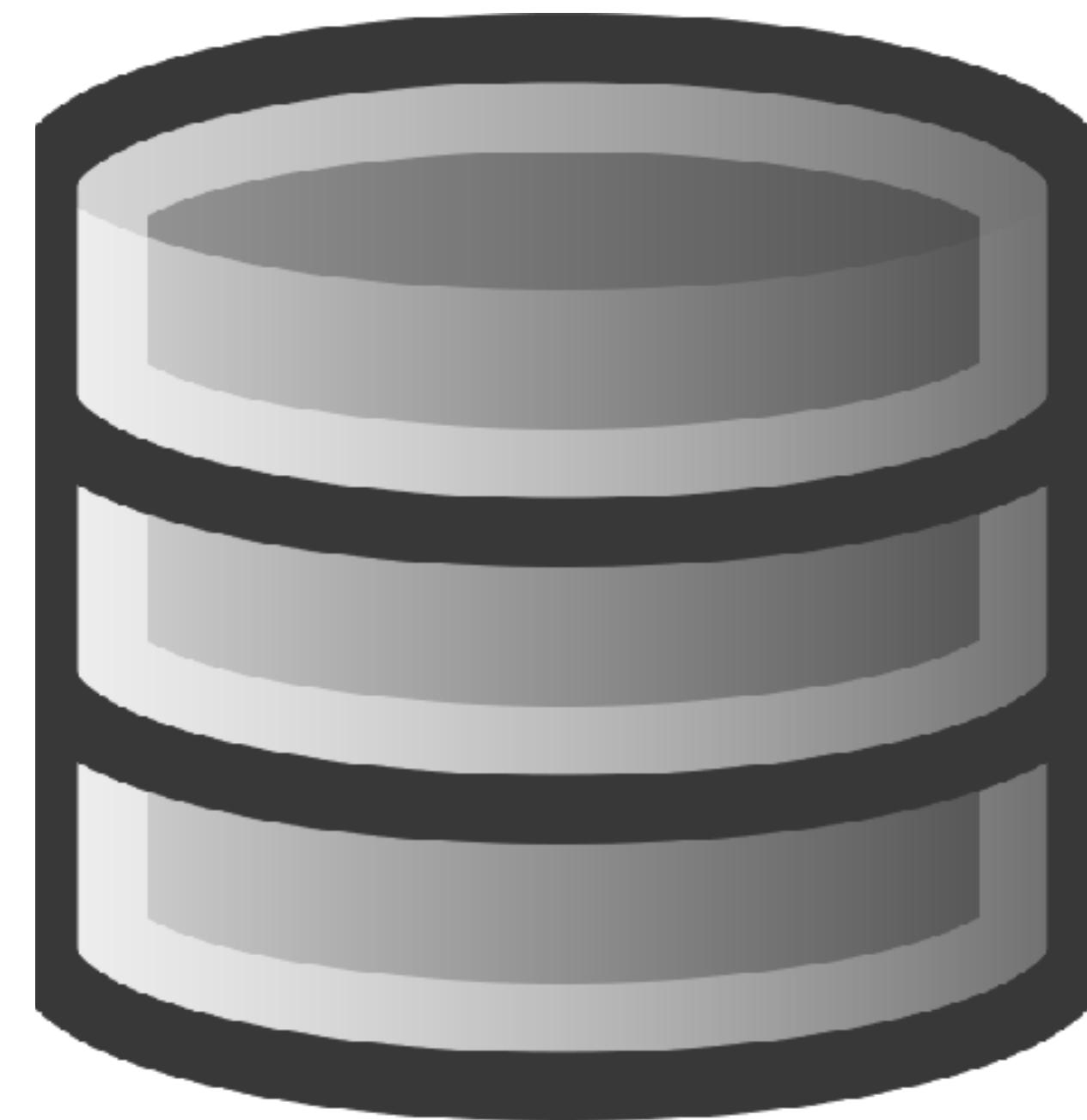
Search



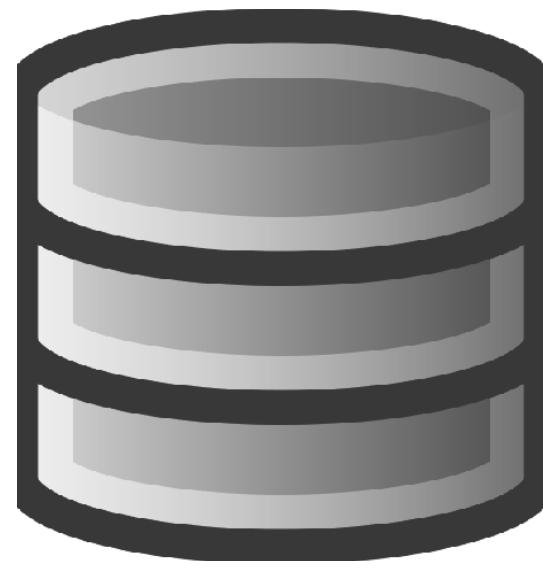
KeyValue Store

General Purpose Key-Value Stores

Key	Value
[] byte	[] byte



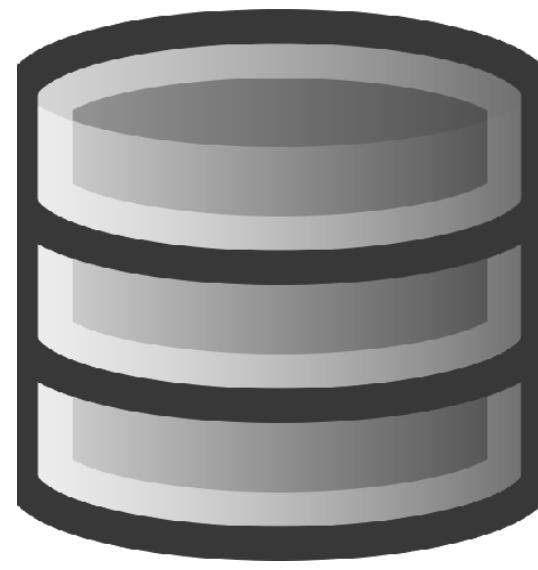
General Purpose Key-Value Stores



Get / Set / Delete Values by Key

Key	Value
hatter	mad
head	attached
rabbit	early

General Purpose Key-Value Stores

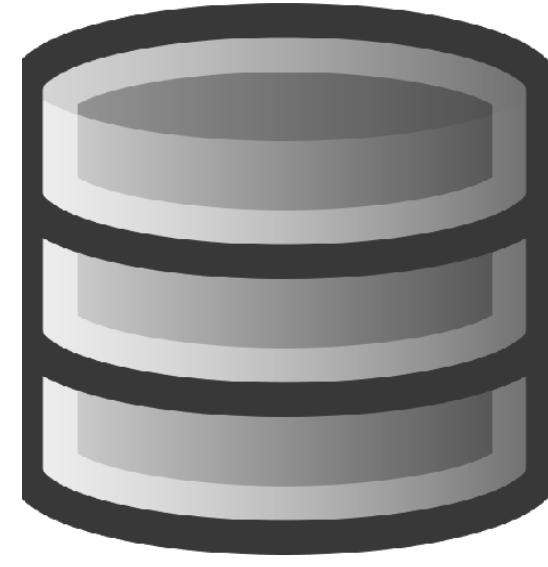


Get / Set / Delete Values by Key

GET hatter → mad

Key	Value
hatter	mad
head	attached
rabbit	early

General Purpose Key-Value Stores

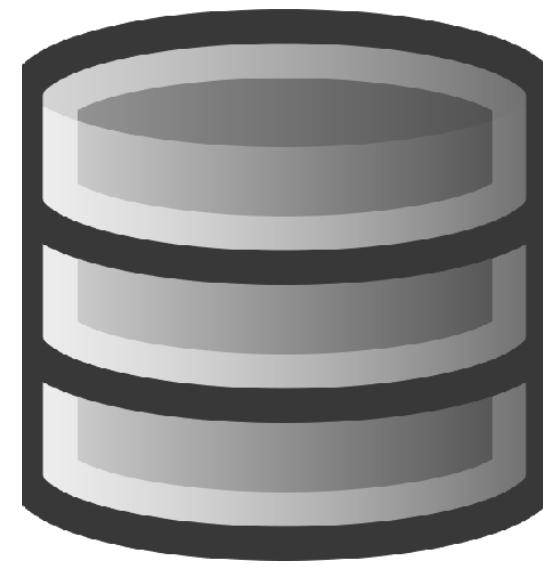


Get / Set / Delete Values by Key

SET rabbit late

Key	Value
hatter	mad
head	attached
rabbit	early late

General Purpose Key-Value Stores

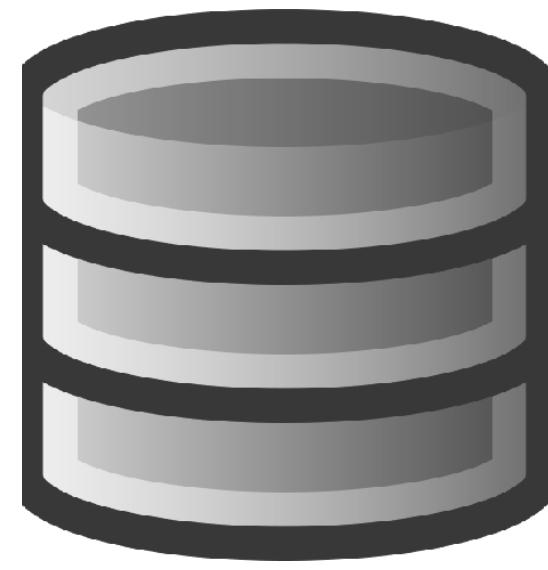


Get / Set / Delete Values by Key

DELETE head

Key	Value
hatter	mad
rabbit	late
head	attached

General Purpose Key-Value Stores

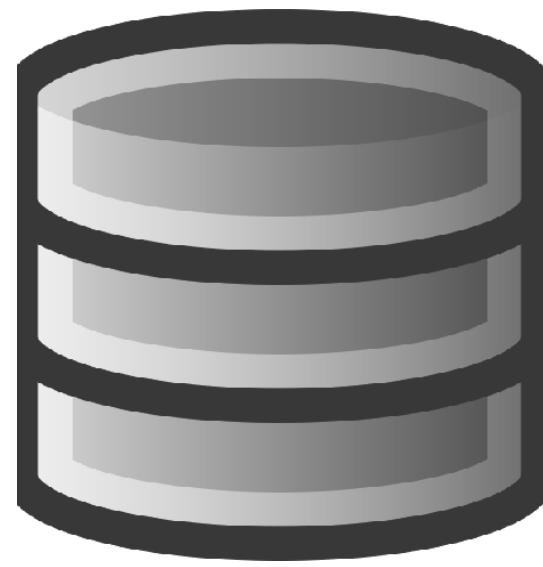


Iterate Ranges of Key-Value Pairs



Key	Value
hatter	mad
rabbit	late

General Purpose Key-Value Stores



Iterate Ranges of Key-Value Pairs



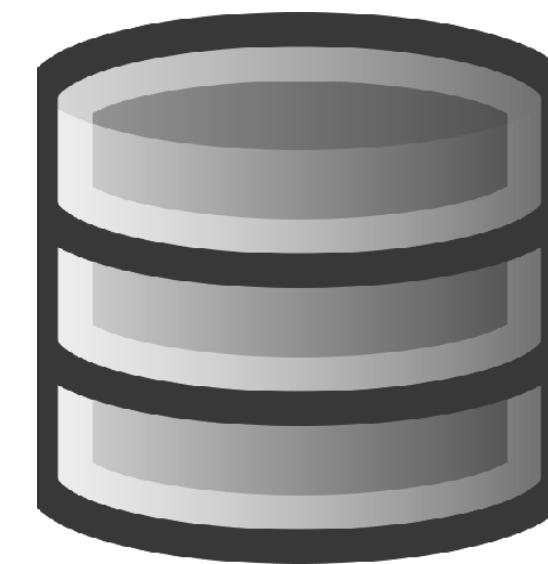
Key	Value
hatter	mad
rabbit	late

General Purpose Key-Value Stores

Atomic Batch Updates

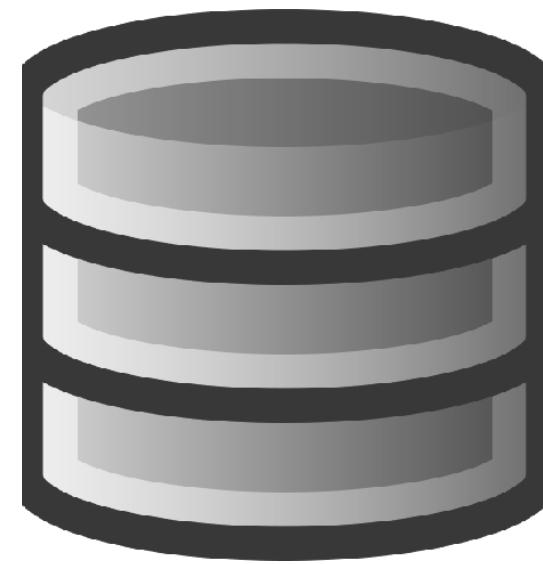
tea	party
cat	grinning
hare	march

Key	Value
hatter	mad
rabbit	late



General Purpose Key-Value Stores

Atomic Batch Updates

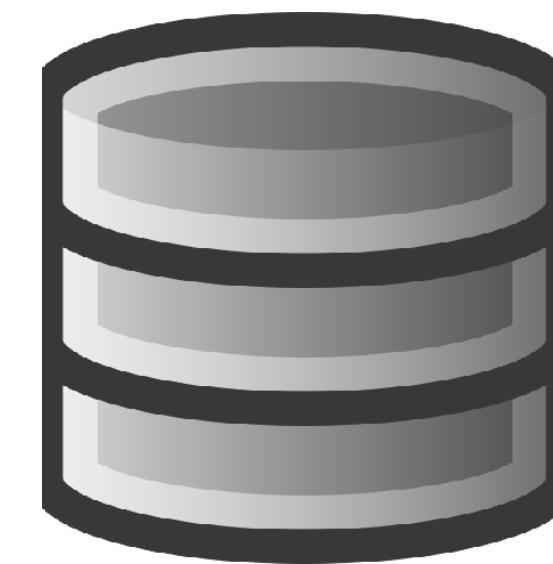
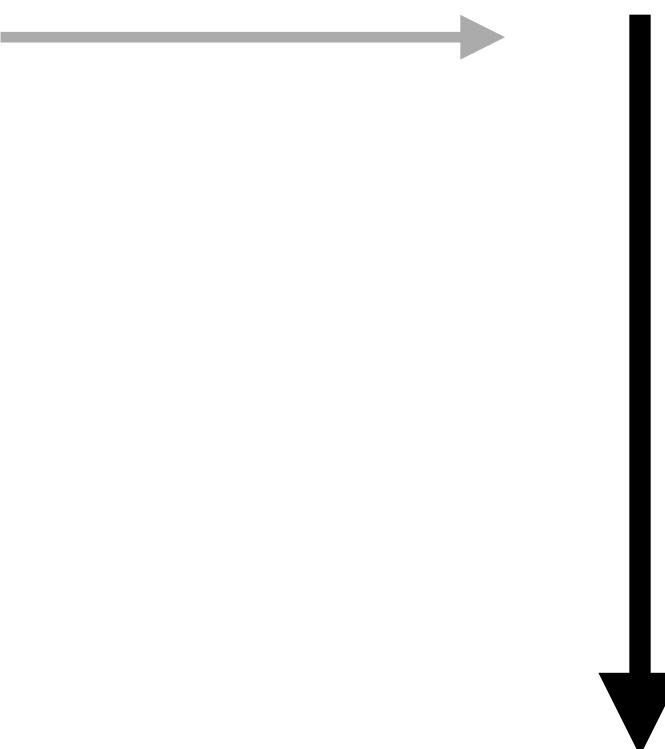


Key	Value
cat	grinning
hare	march
hatter	mad
rabbit	late
tea	party

General Purpose Key-Value Stores

Isolated Read Snapshots

Iterator Started

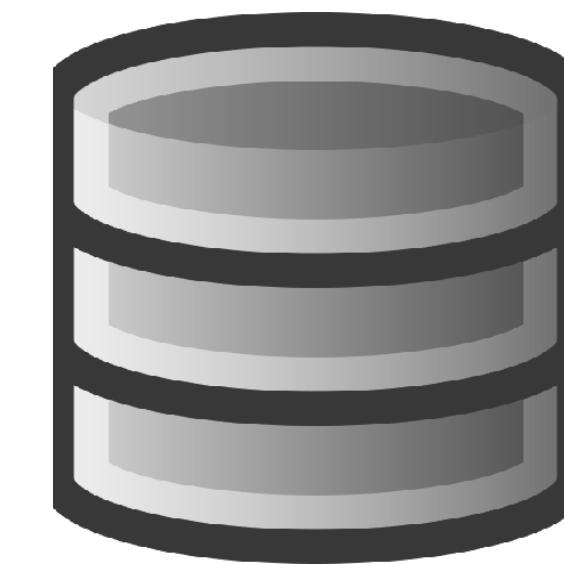


Key	Value
cat	grinning
hare	march
hatter	mad
rabbit	late
tea	party

General Purpose Key-Value Stores

Isolated Read Snapshots

Iterator Started
Concurrent Mutation

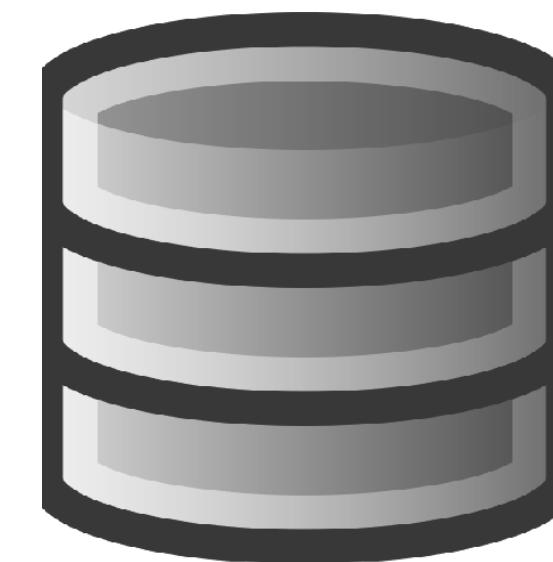
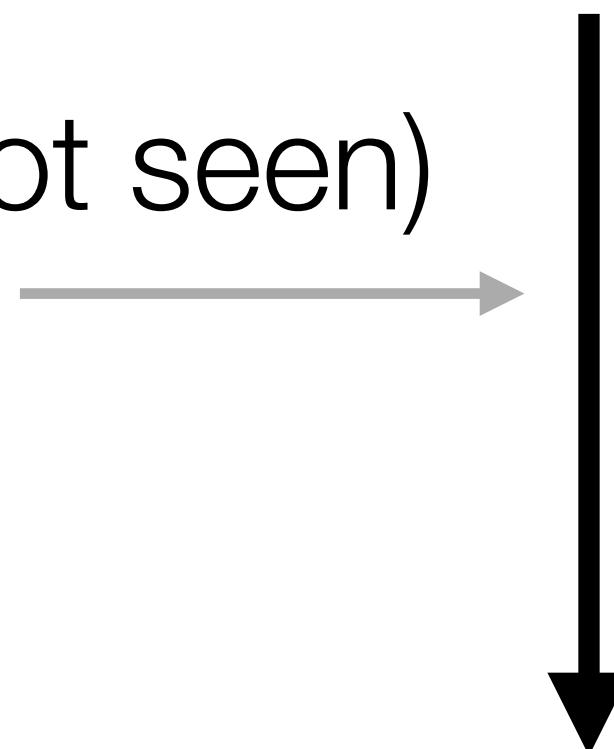


Key	Value
cat	grinning
caterpillar	smoking
hare	march
hatter	mad
rabbit	late
tea	party

General Purpose Key-Value Stores

Isolated Read Snapshots

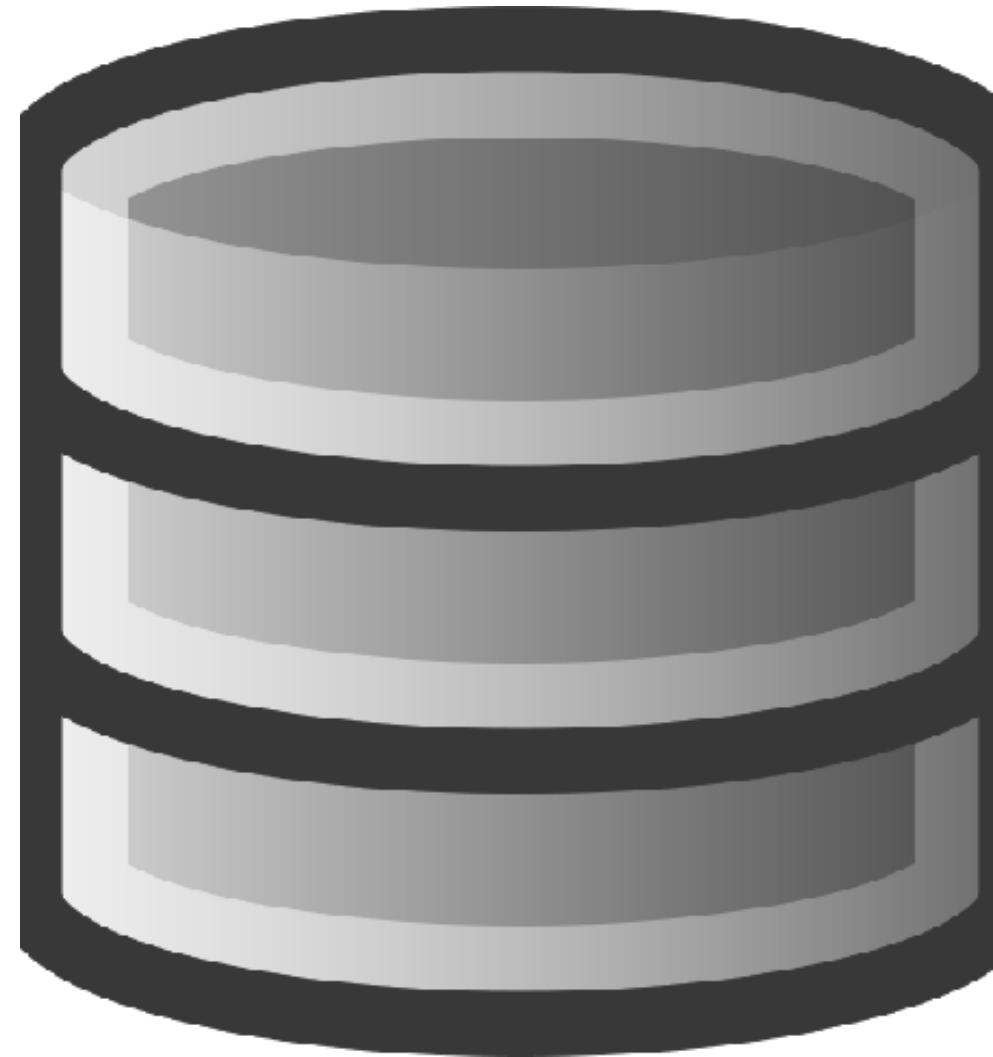
Iterator Started
Concurrent Mutation (not seen)



Key	Value
cat	grinning
caterpillar	smoking
hare	march
hatter	mad
rabbit	late
tea	party

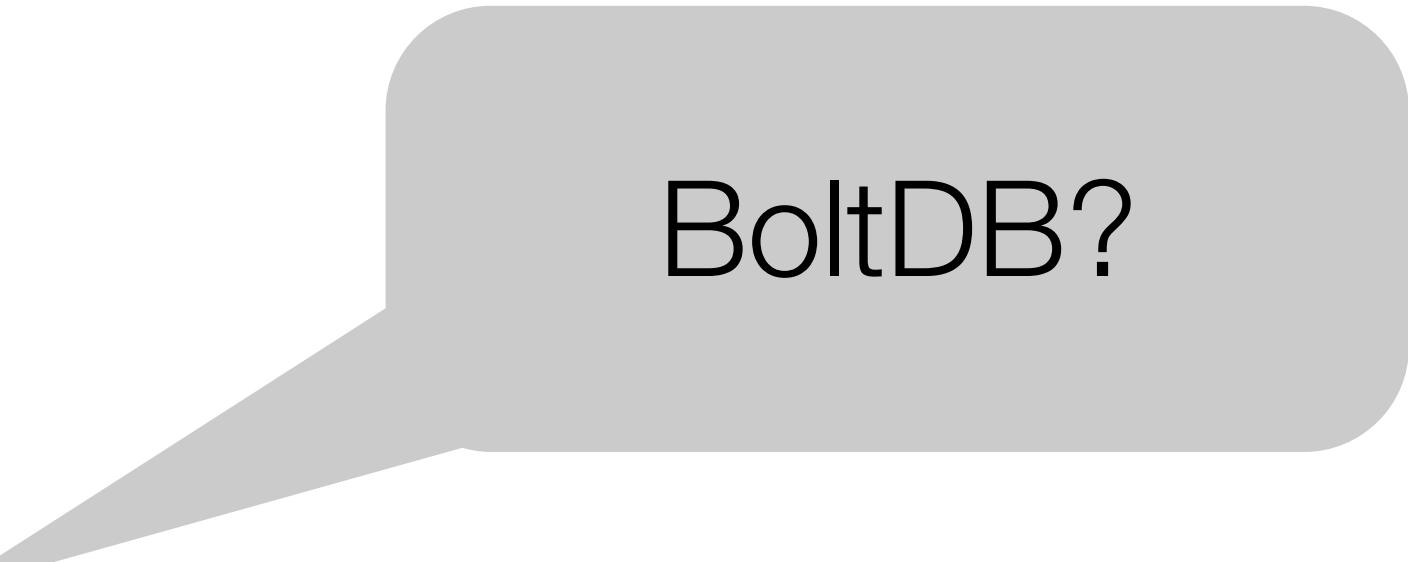
General Purpose Key-Value Stores

Persistence to Disk



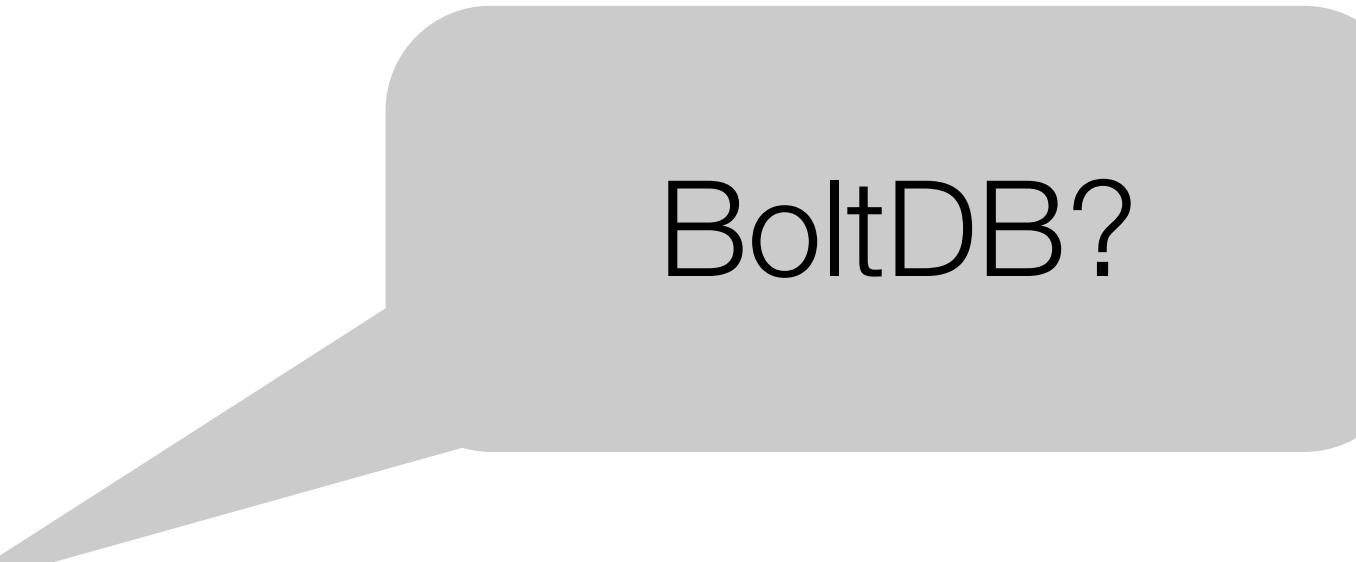
Off-The-Shelf Solutions

Off-The-Shelf Solutions



BoltDB?

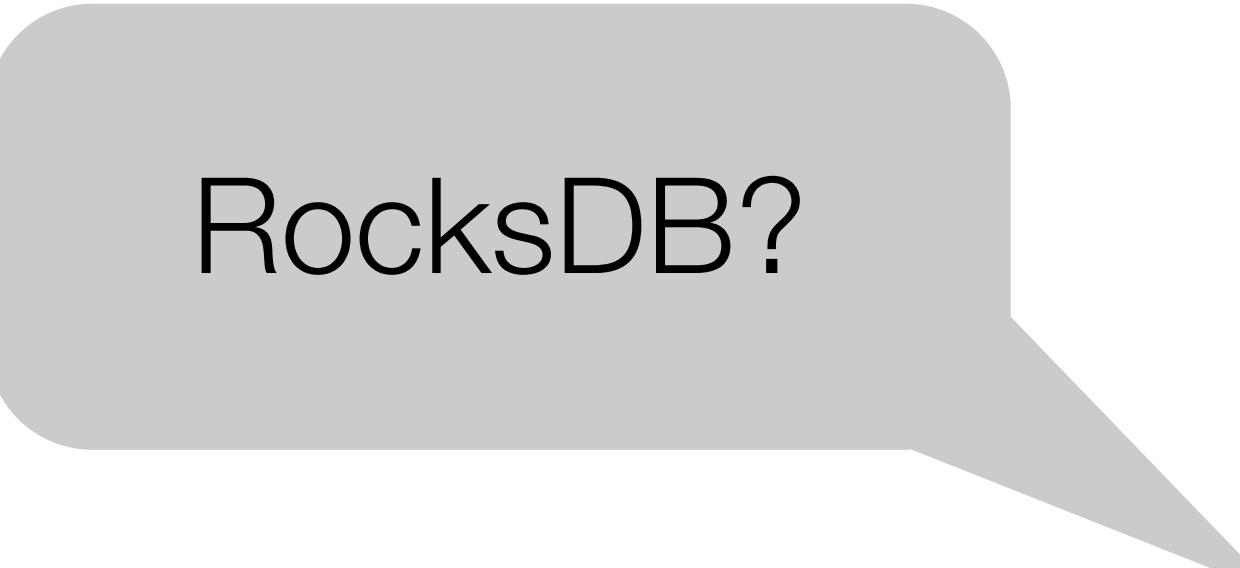
Off-The-Shelf Solutions



BoltDB?

- Go
- b+tree
- Great Read Performance

Off-The-Shelf Solutions



RocksDB?

Off-The-Shelf Solutions

RocksDB?

- C++
- LSM (leveldb)
- Better Read/Write Perf Balance
- cgo

Off-The-Shelf Solutions



GoLevelDB?

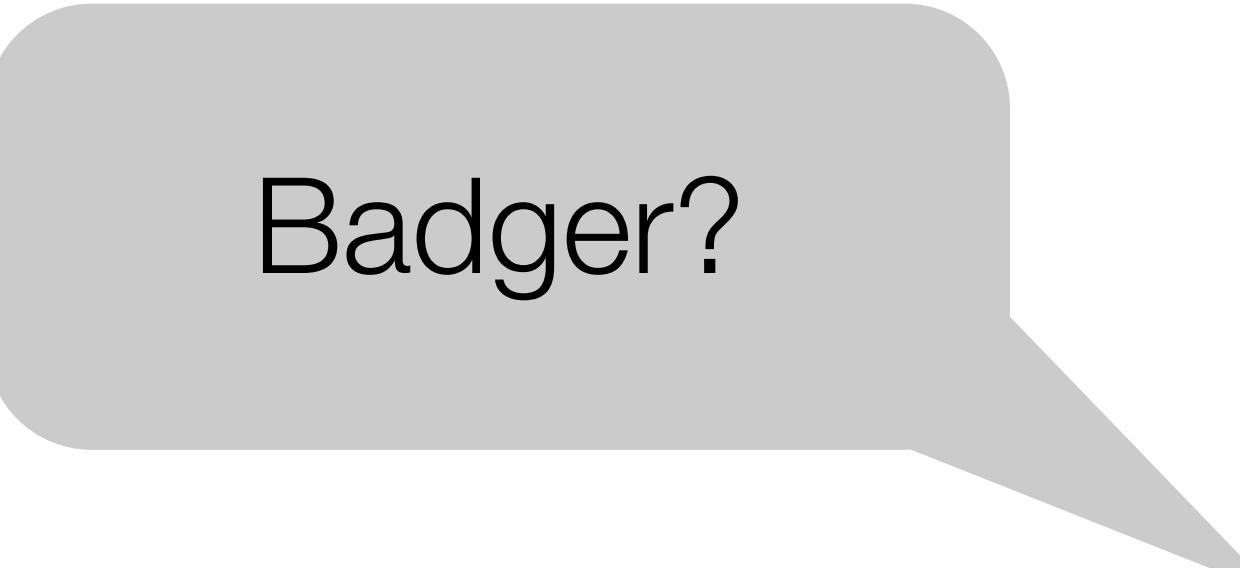
Off-The-Shelf Solutions



GoLevelDB?

- Go
- LSM (leveldb)
- Unable to Tune Adequately

Off-The-Shelf Solutions



Badger?

Off-The-Shelf Solutions

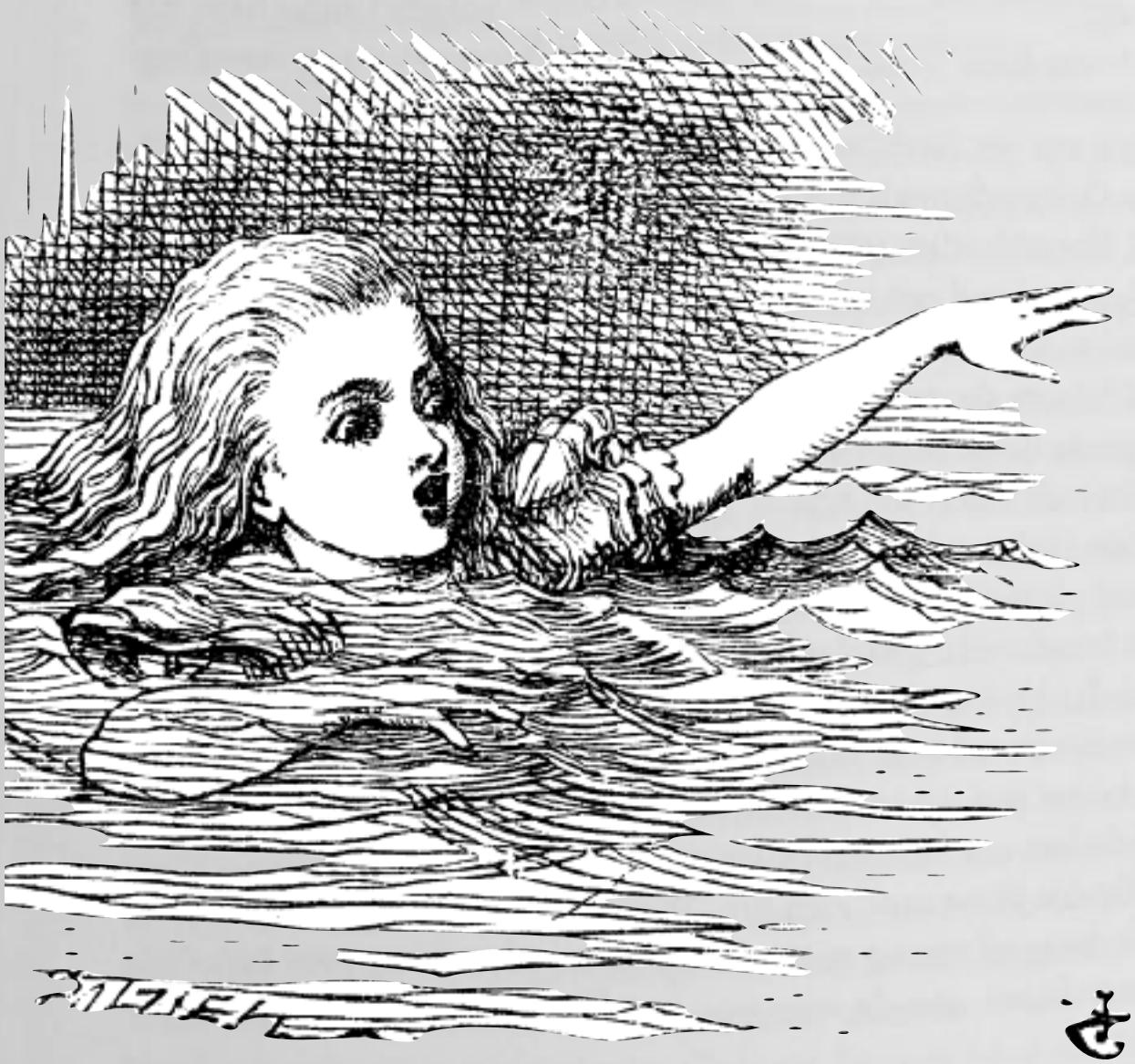


Badger?

- Go
- Wisckey
- Not Available at the Time

Build our Own Key-Value Store?

Zijn er Noordelijk het licht van de vuurzuren langs die vier windrichtingen zweven, zo streken elk jaar in grootsche vlaggen die vier seisoenen over Groot Rechtersen. Max kende die wisseling van de jaargrijsen eigenlijk alleen uit Amsterdam: op een dag in februari of maart is er een onbeschrijflijke voorjaarsgeur als hij's morgens op straat loopt, onbegrijpbaar als de decimalen van π , die stoffige somer waarin de stad vol moersten sprokkelde, even plotseling veranderend in de voorjaar, herfst, en dan de bleke winter waarin de steenten van de straten en de muren opeens de ongemaakbare aard van de wereld leken uit te drukken, - maar dat alles eigenlijk in het voorjaar, oppgemerkt alleen in de korte tussenposes waarin hij van het ene interieur naar het andere ging. Was in de stad de natuur tachtig achtereenvolgende, op het kasteel zat hij met Quinten en Sjouke midden in de daverende concertzaal. Lente en herfst kwamen met kolossaal vertoon, de somers waren hoter en droger, de winters koudter en winter. Die onafgebonden verandering, had hij nooit tegen Ottus beweerd, was natuurlijk de bron van alle creativiteit; de onvermoeidheid van de natuur tussen dit keerkringen hieldde ook toe culturele stilstand, de tropen waren een onafgebonden mensheid, altijd groen zoals de poolgebieden altijd wit waren, van de primitieve breedtegraden met hun viervormigheid waren



Simplicity

Damian Gryski - Slices Performance through cache friendliness

Simplicity

Damian Gryski - Slices Performance through cache friendliness

"Notes on C Programming" (Rob Pike, 1989)

Rule 3. Fancy algorithms are slow when n is small, and n is usually small. Fancy algorithms have big constants. Until you know that n is frequently going to be big, don't get fancy.

Rule 4. Fancy algorithms are buggier than simple ones, and they're much harder to implement. **Use simple algorithms as well as simple data structures.**

Simplicity

Dave Cheney - Simplicity and Collaboration

Simplicity

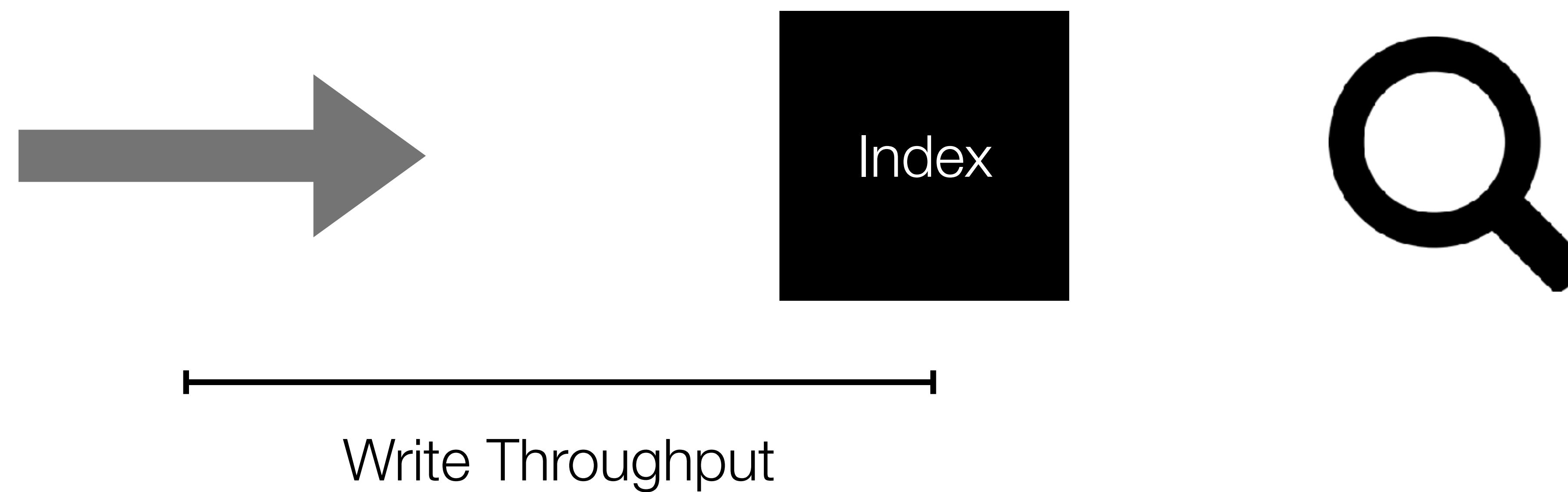
Dave Cheney - Simplicity and Collaboration

"Simplicity cannot be added later"

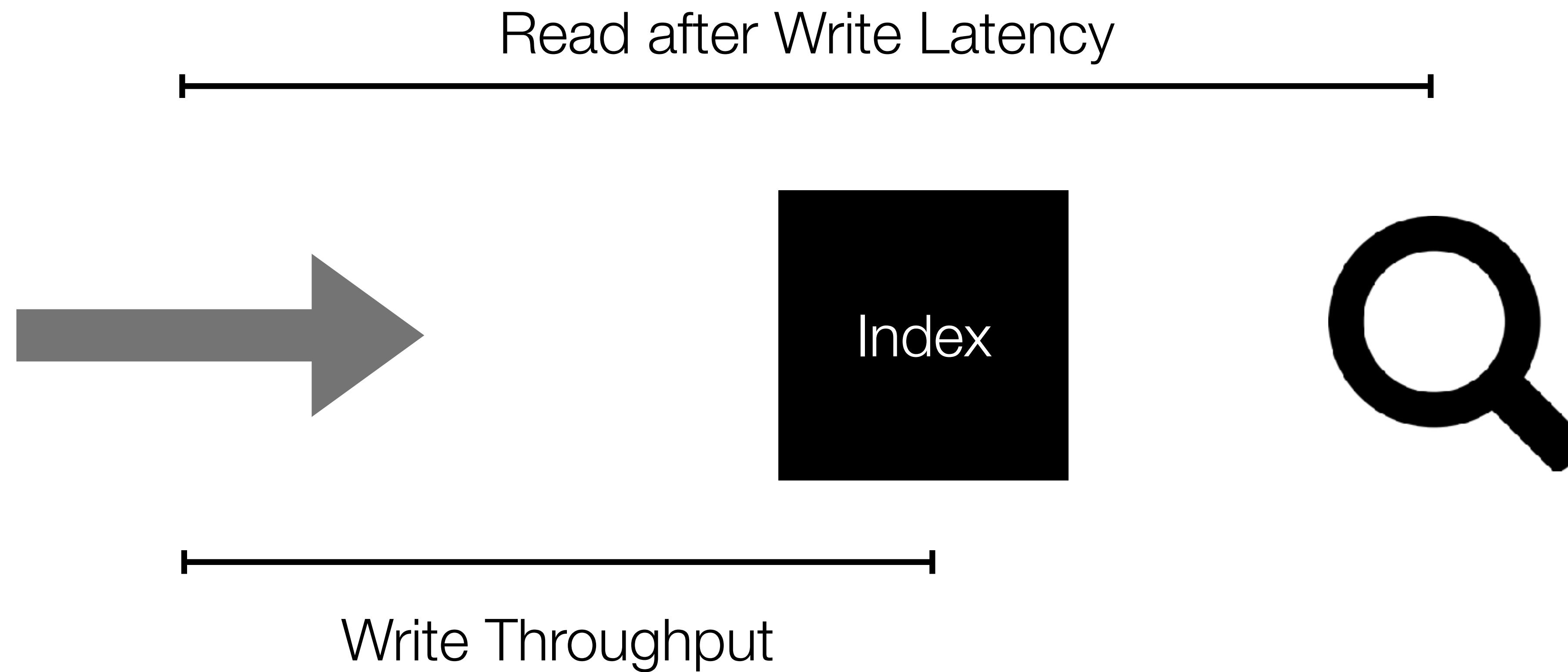
Special Purpose Key-Value Store



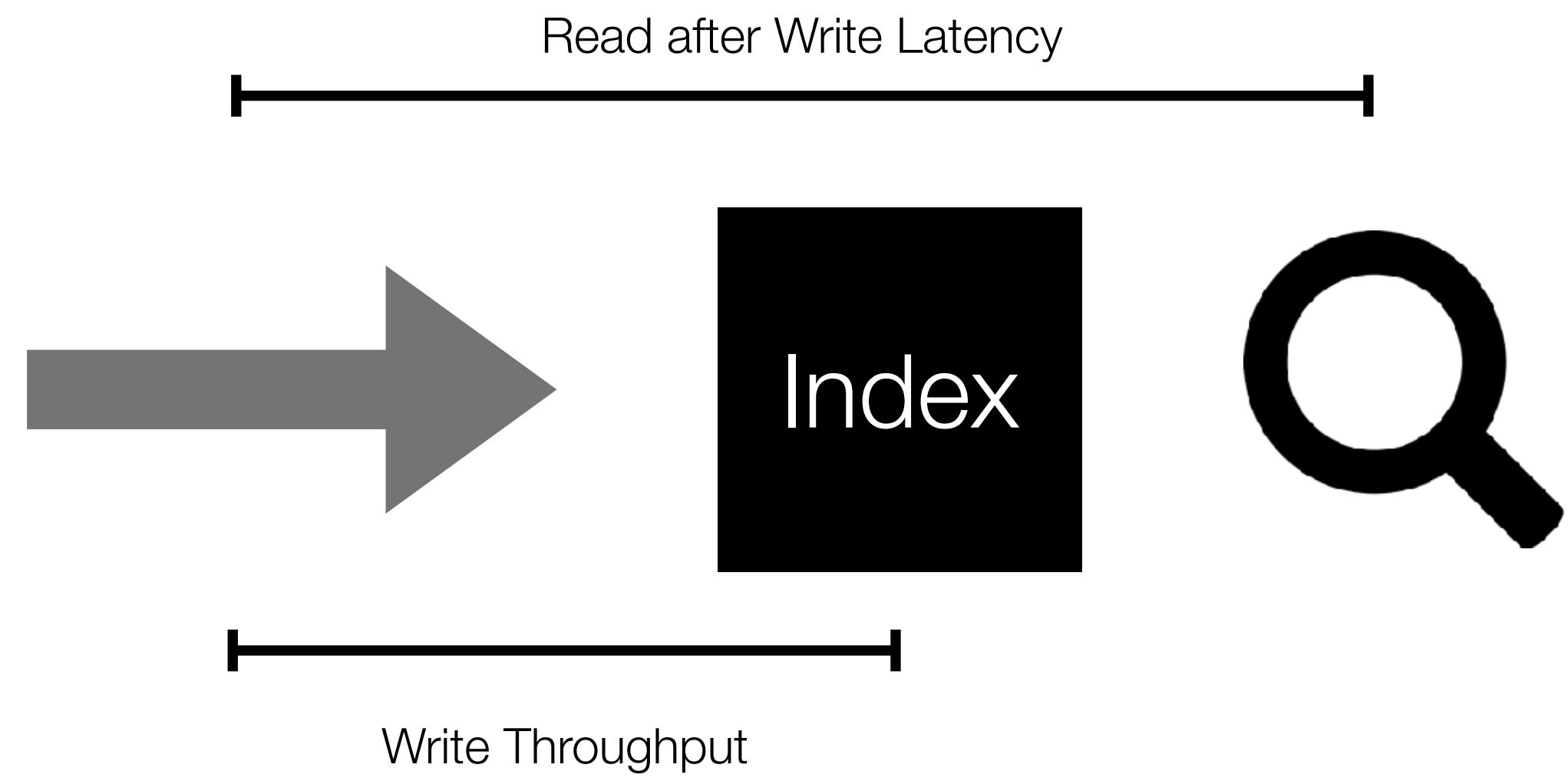
Special Purpose Key-Value Store



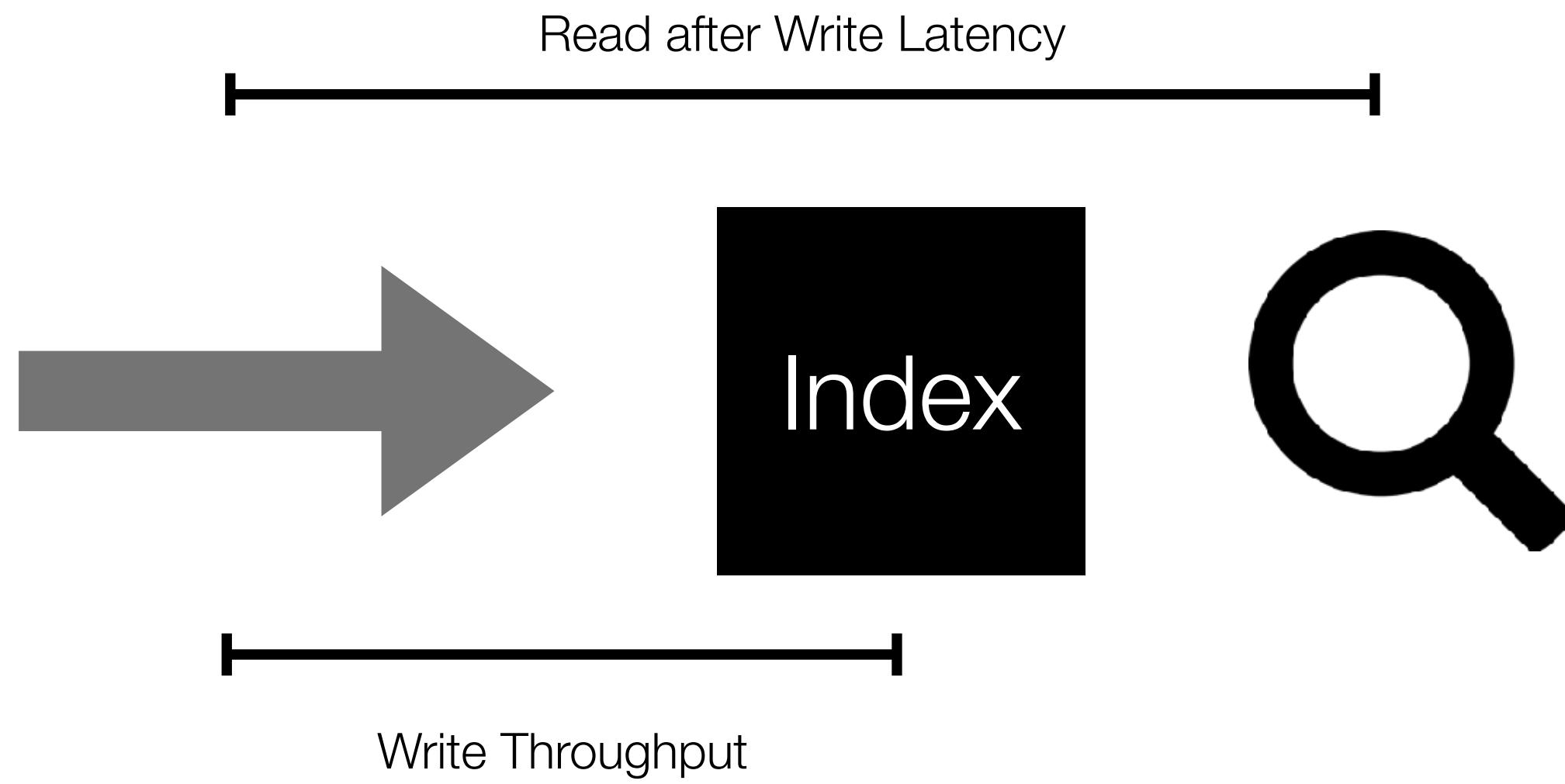
Special Purpose Key-Value Store



Special Purpose Key-Value Store

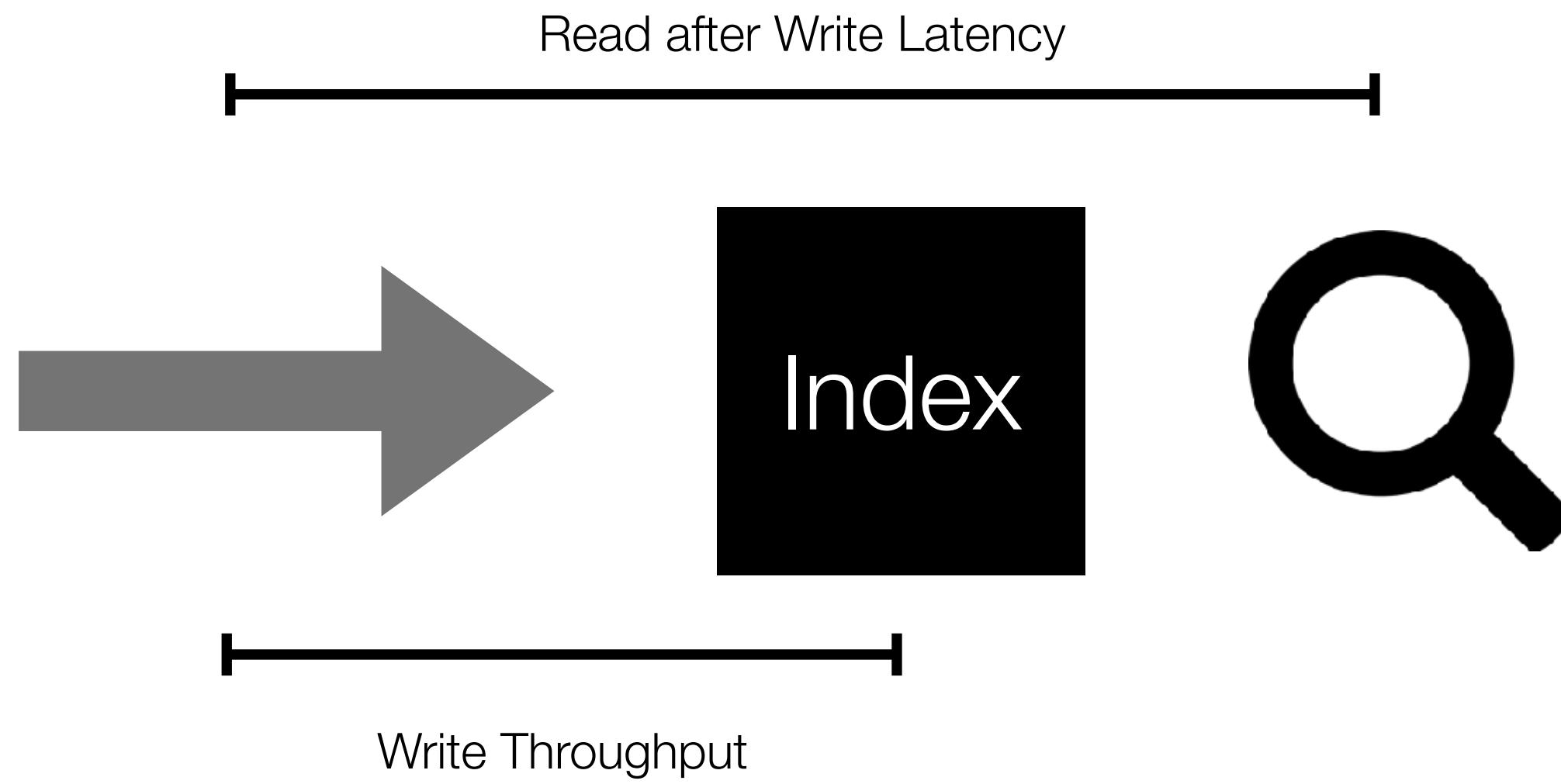


Special Purpose Key-Value Store



- Persistence Decoupled from Read and Write

Special Purpose Key-Value Store



- Persistence Decoupled from Read and Write
- Willing to use All System RAM*

* Bounded by some Quota

The Rabbit Hole



Zuin en Noodzakelijk het licht van de vuursteentjes langs de vier windrichtingen overhoek, zo streken elk jaar in grootsche vlaggen die vier seisoenen over Groot Rechtersen. Max kende die wisseling van de jaargrijsen eigenlijk alleen uit Amstelbad: op een dag in februari of maart is er een onbeschrijflijke voorjaarsgeur als hij 's morgens op straat loopt, onbeschutbaar als de decimalen van π , die stoffige somer waarin de stad vol moeststenen strooimde, even plotseling veranderend in de voorjaar, herfst, en dan de bleke winter waarin de sterren van de straten en de muren opeens de ongemaakbare aard van de wereld leken uit te drukken, - maar dat alles eigenlijk in het voorjaar, oppgemerkt alleen in de korte tussenposes waarin hij van het ene interieur naar het andere ging. Was in de stad de natuur tachtig achtereenvolgende, op het kasteel zat hij met Quinten en Sjipje midden in de daverende concertzaal. Lente en herfst kwamen met kolossaal vertoon, de somers waren hotter en droger, de winters koudster en winter. Die onafgebonden verandering, had hij een tegen Ottus beweerd, was natuurlijk de bron van alle creativiteit; de onvermoeigheid van de natuur tussen dit keerkringen hieldde ook toe culturele stilstand, de tropen waren een onafgebonden land, altijd groen zoals de poolgebieden altijd wit waren, van de primitieve breedtegraden met hun viervormigheid waren

```
type Collection interface {
```

```
    ExecuteBatch(b Batch) error
```

```
    Snapshot() (Snapshot, error)
```

```
}
```

```
type Batch interface {  
    Set(key, value []byte) error  
    Delete(key []byte) error  
}
```

Batch of Changes



Batch of Changes

alice

dodo

knave

cat

duchess

queen

hare

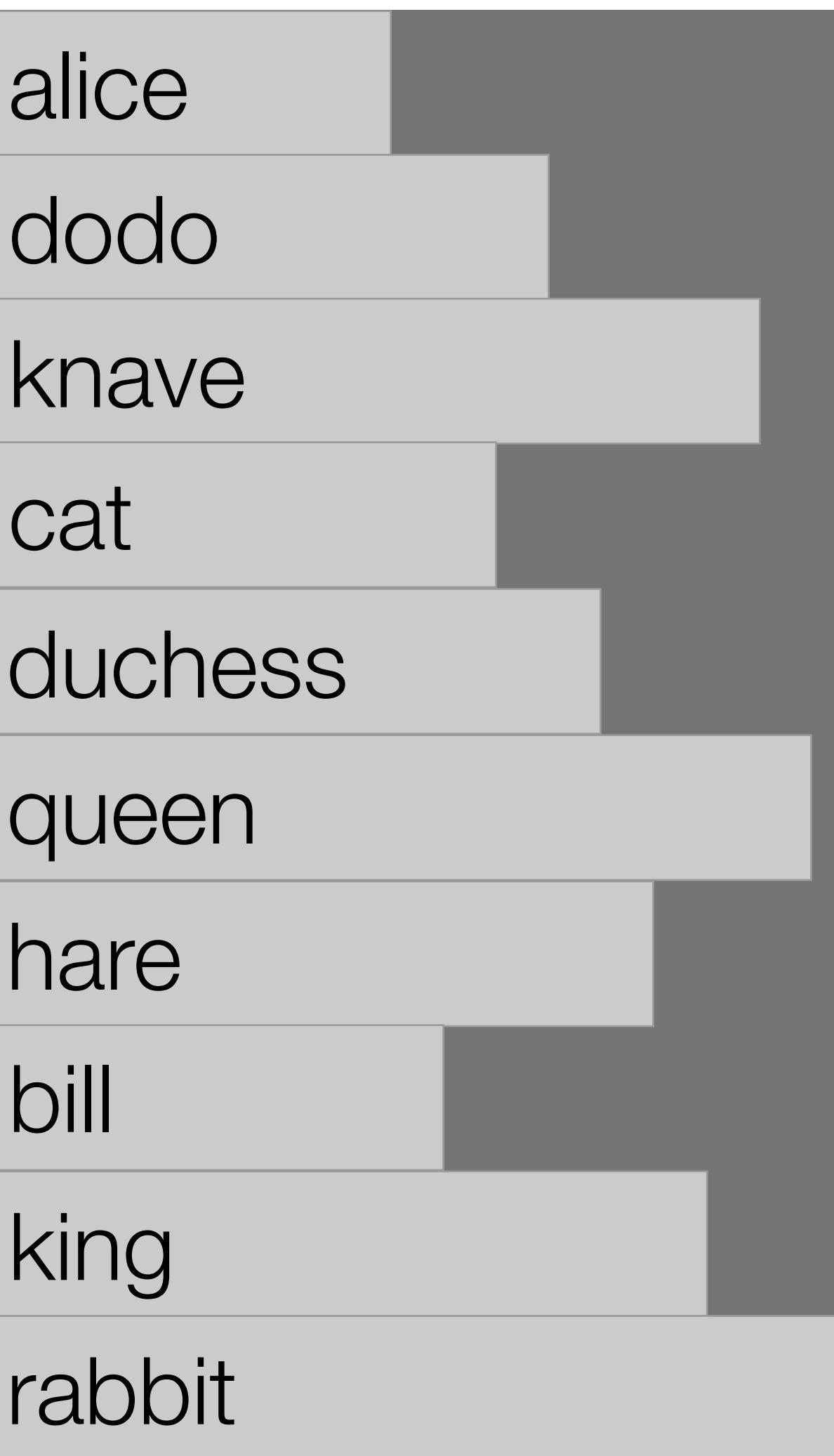
bill

king

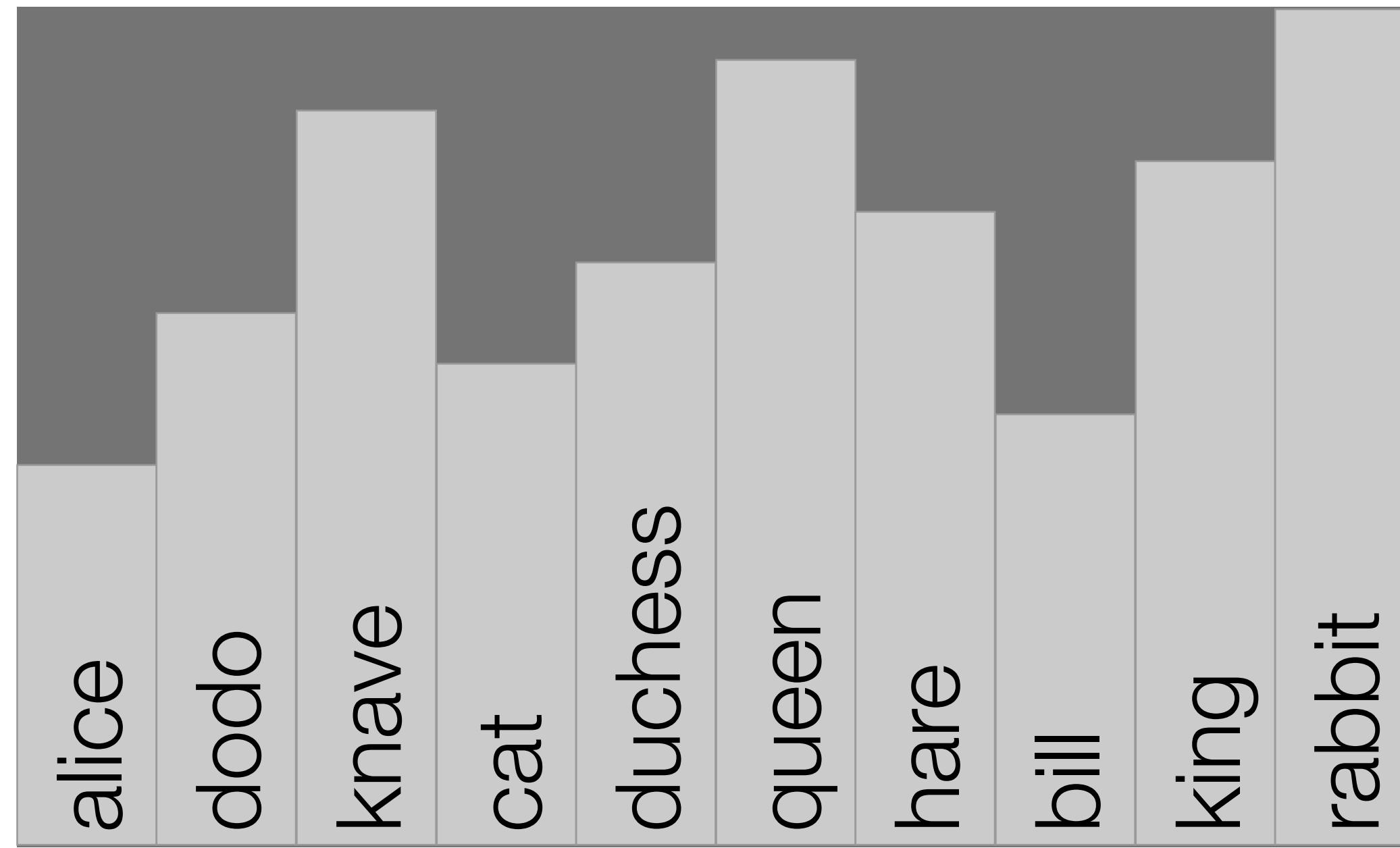
rabbit

```
type Snapshot interface {  
    Get(key []byte) ([]byte, error)  
    Iterator(start, end []byte)  
        (Iterator, error)  
}
```

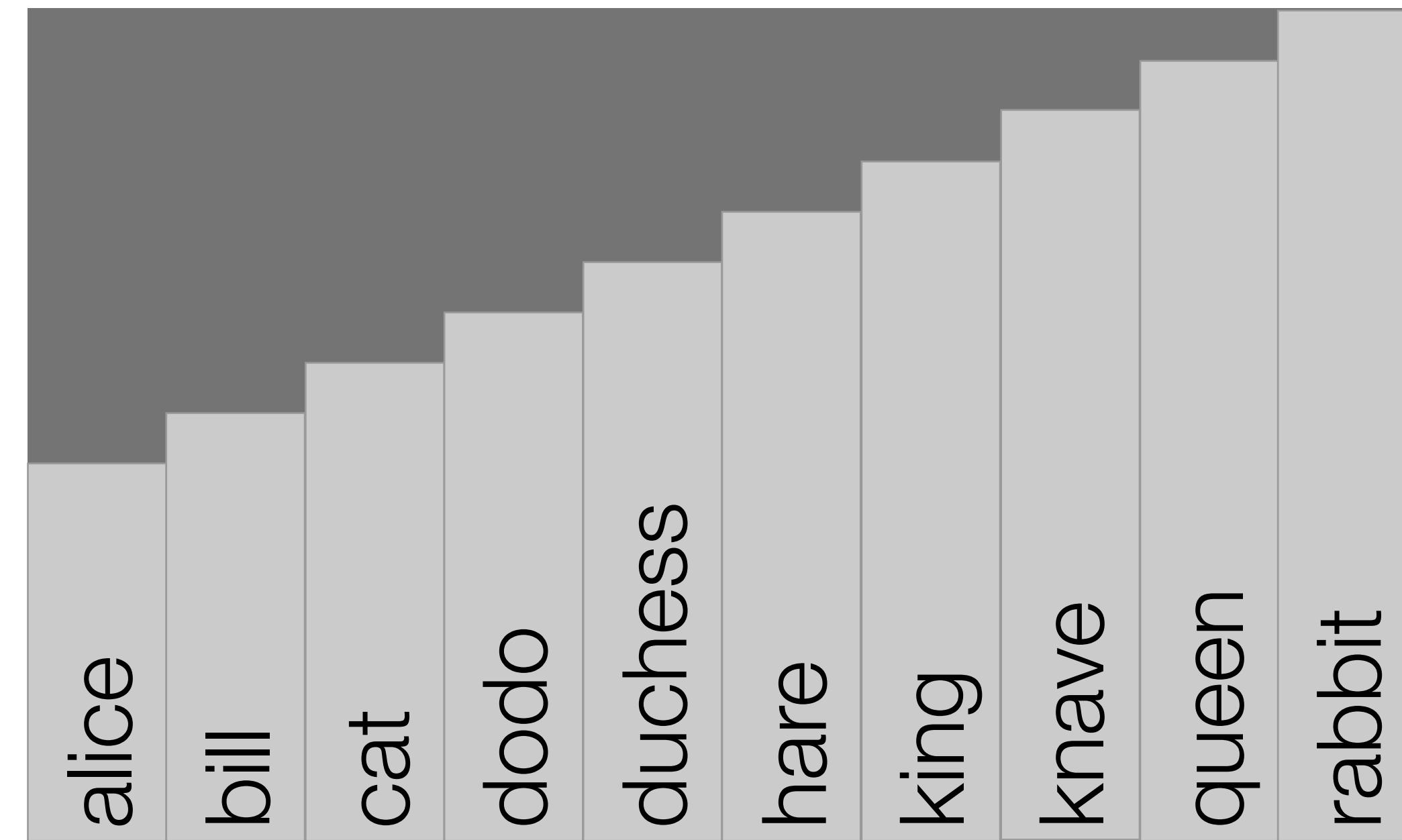
Batch of Changes



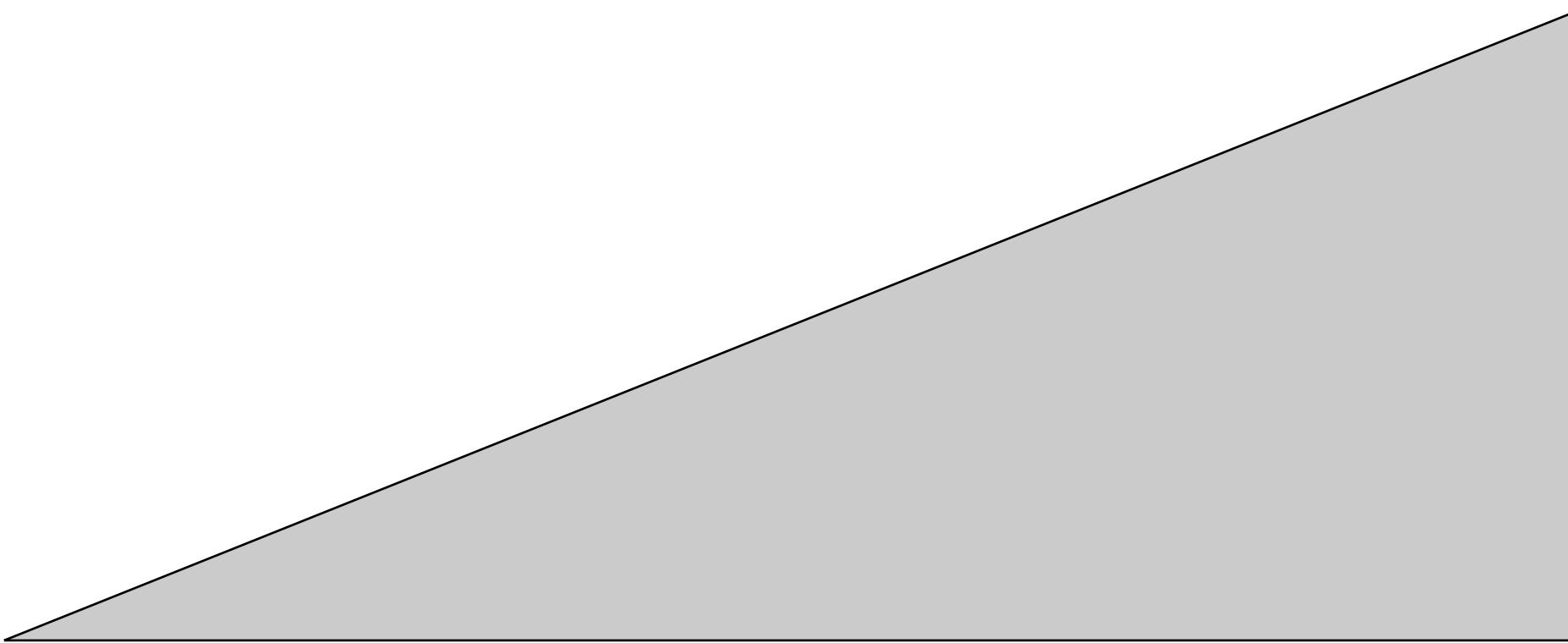
Batch of Changes



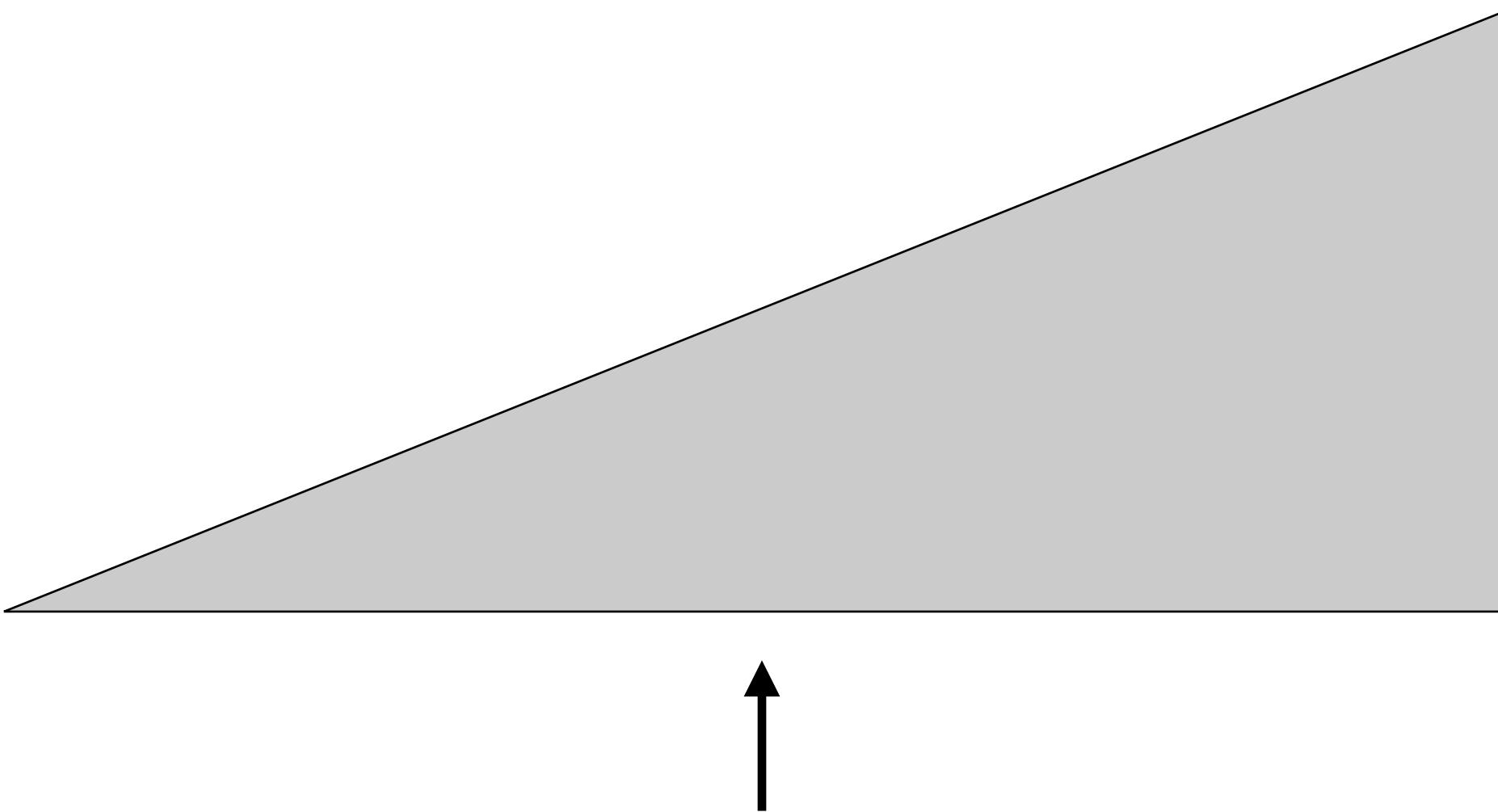
Sort the Keys



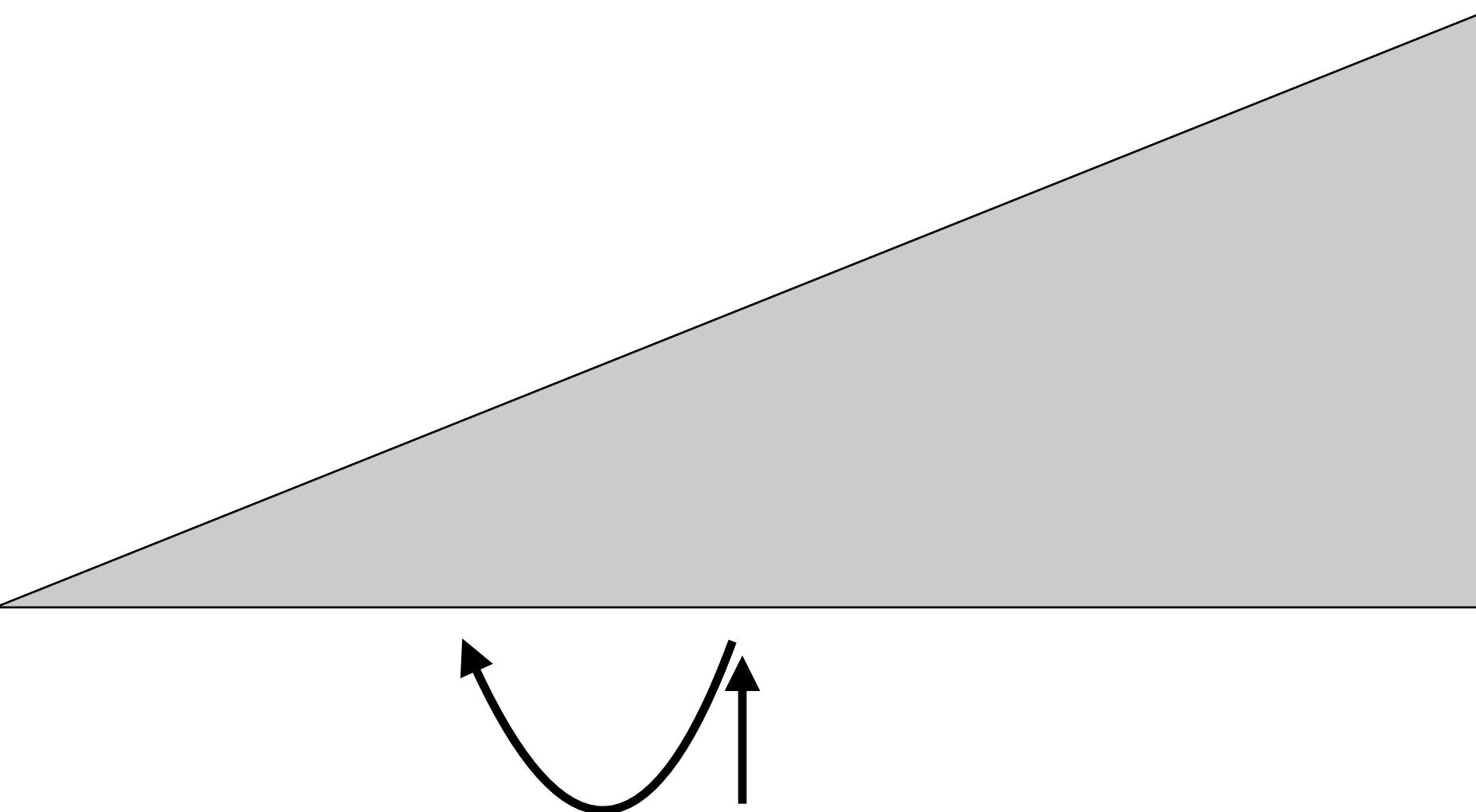
Get – Binary Search for Key



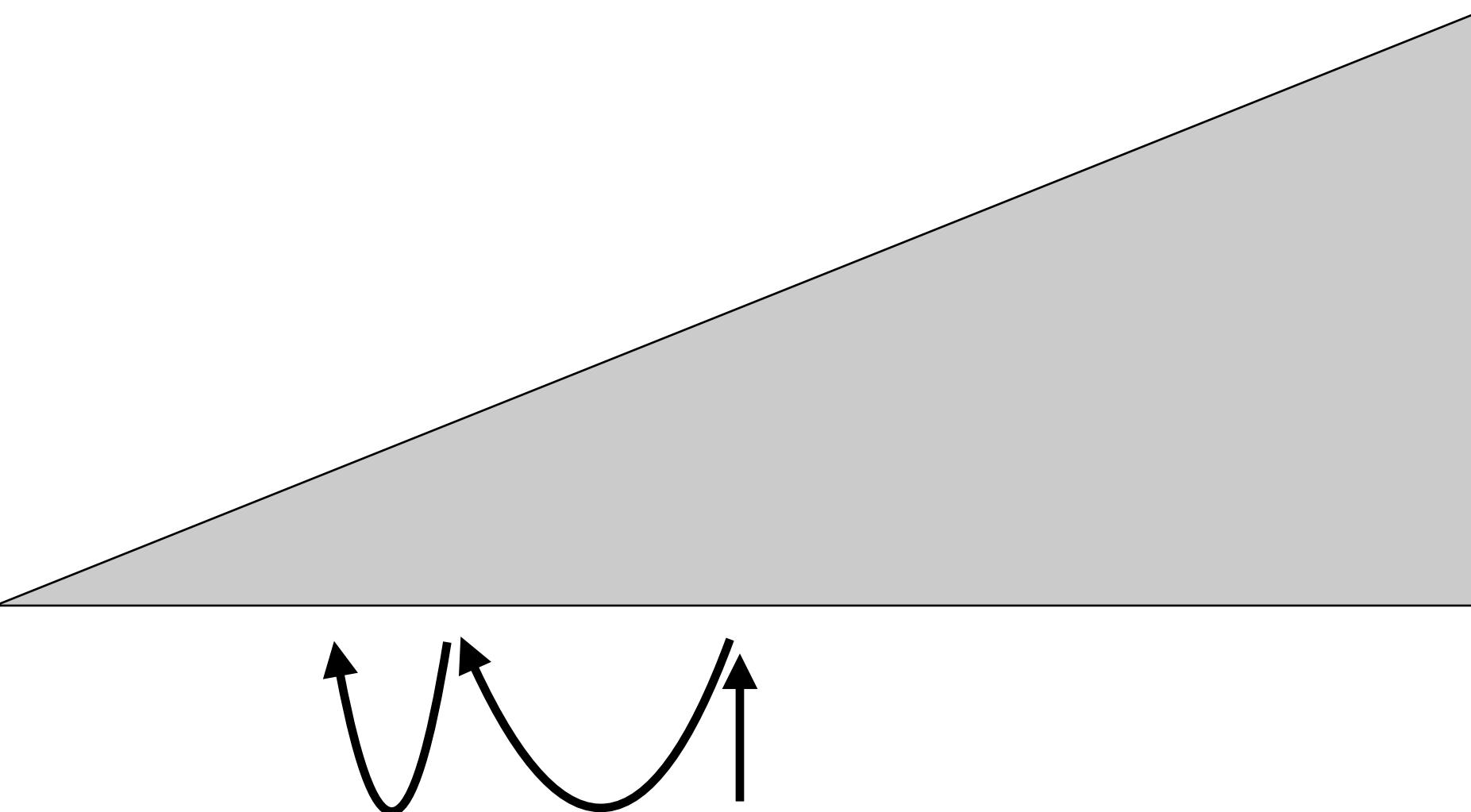
Get – Binary Search for Key



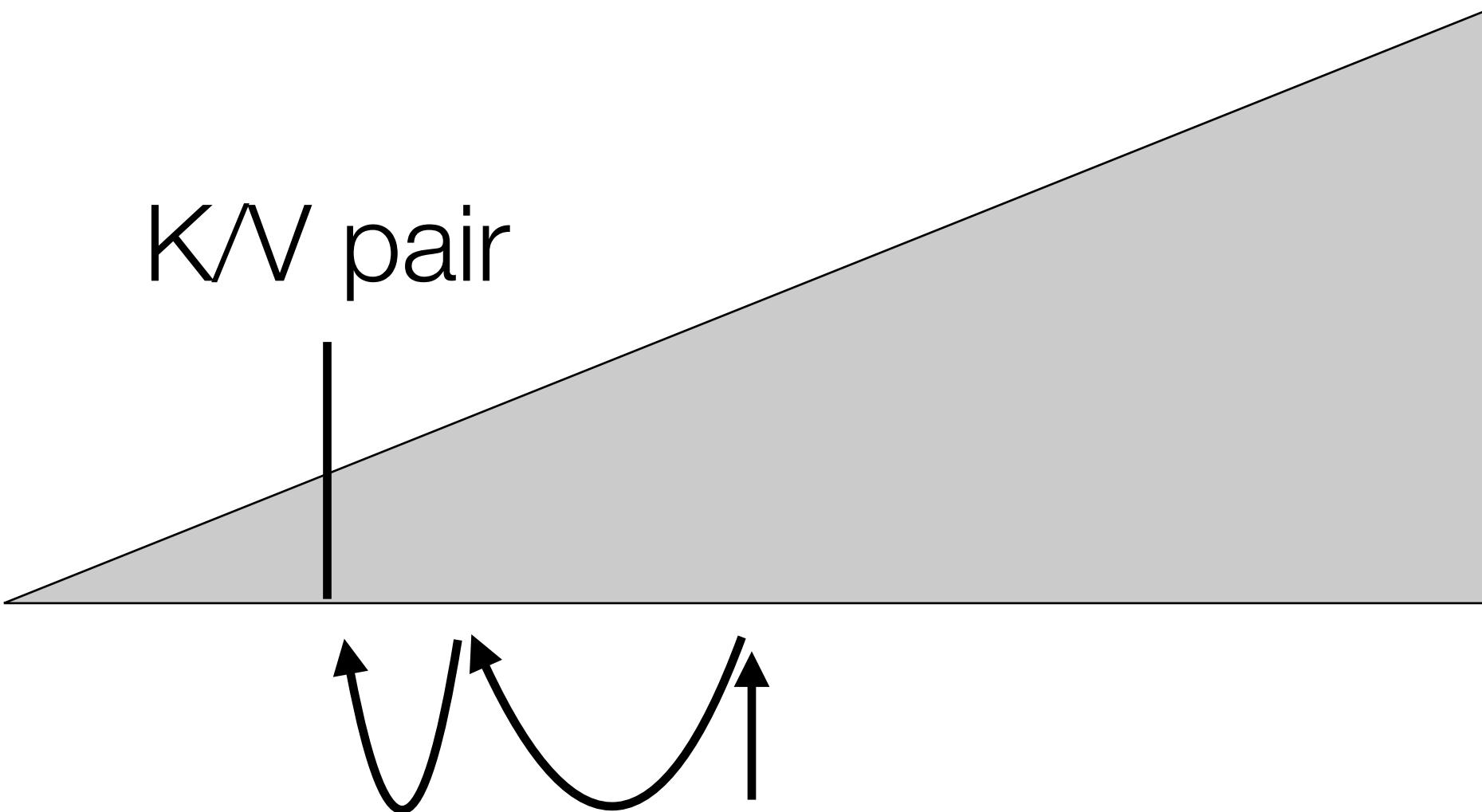
Get – Binary Search for Key



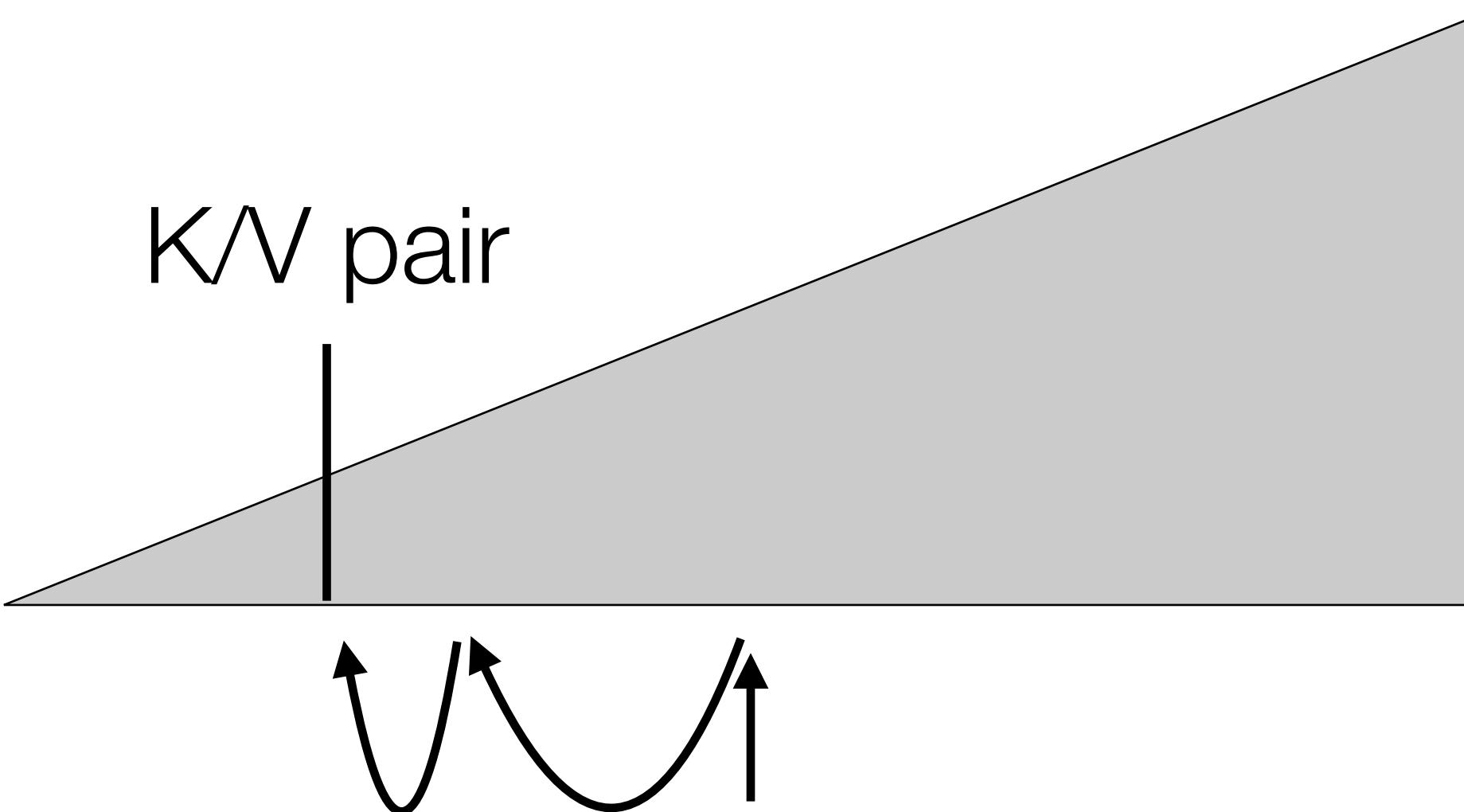
Get – Binary Search for Key



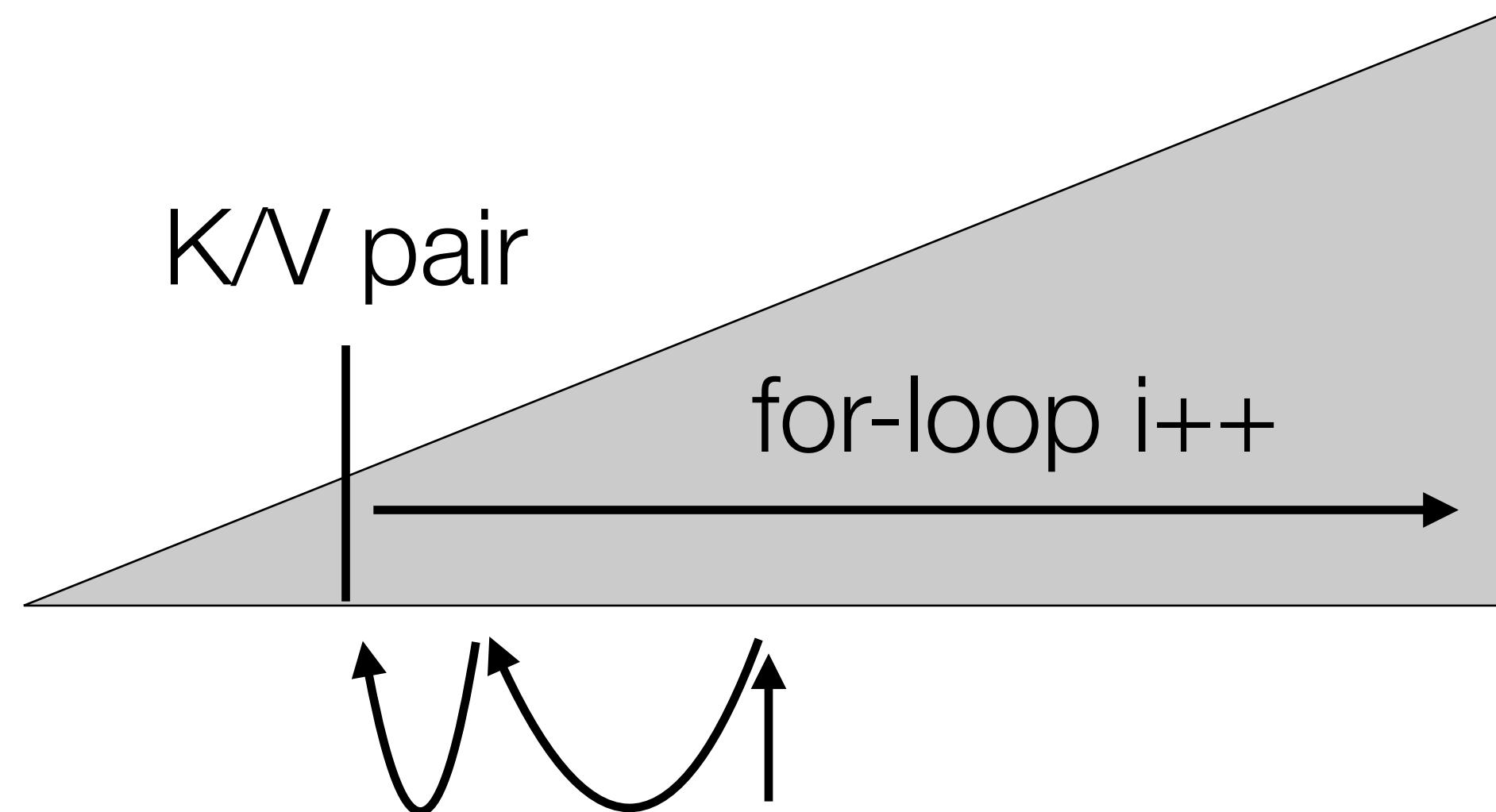
Get – Binary Search for Key



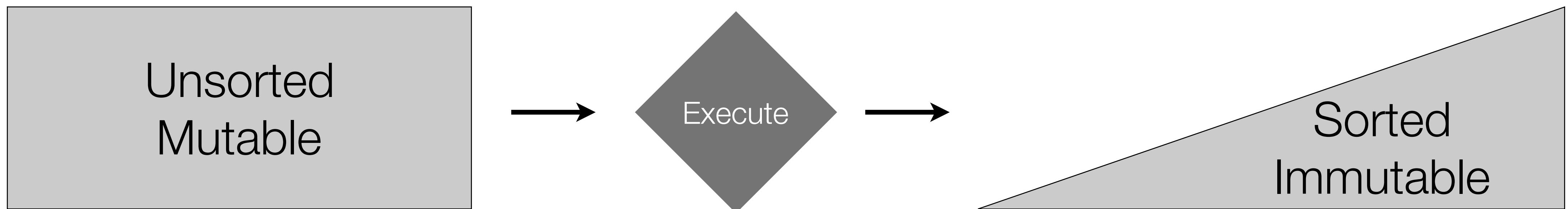
Iterate Key-Value Pairs – For Loop



Iterate Key-Value Pairs – For Loop



Batch → Segment



```
type segment struct {  
    data []byte  
    meta []uint64  
}
```

```
Set(  []byte("name"),  []byte("marty") )
```

data []byte

0 20

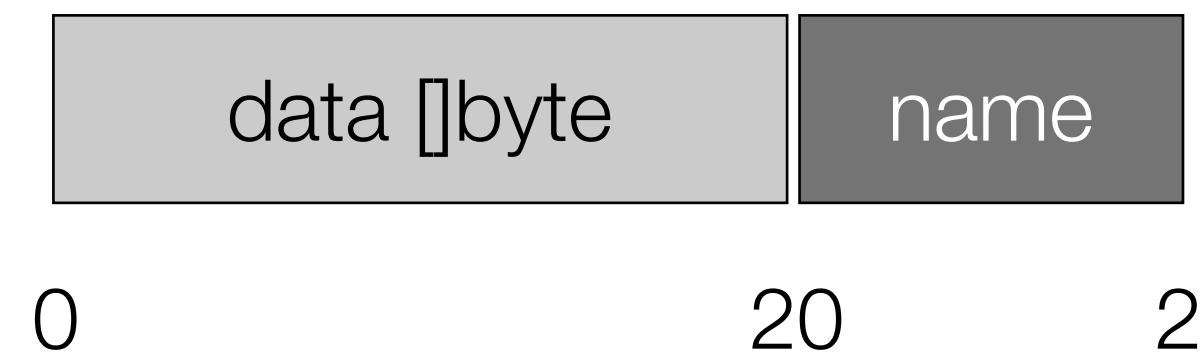
```
Set( []byte("name"), []byte("marty") )
```

data []byte

0 20

Append Key and Value []byte to data slice

```
Set( []byte("name"), []byte("marty") )
```



Append Key and Value []byte to data slice

```
Set( []byte("name"), []byte("marty") )
```



Append Key and Value []byte to data slice

```
Set( []byte("name"), []byte("marty") )
```



Build 2 uint64s of metadata

```
Set( []byte("name"), []byte("marty") )
```



Build 2 uint64s of metadata

Data Offset



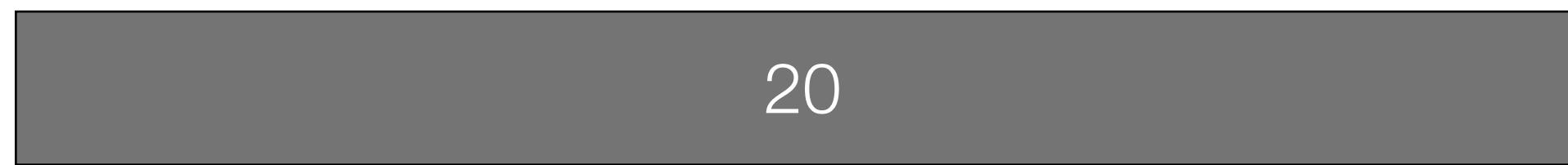
64 bits

```
Set( []byte("name"), []byte("marty") )
```

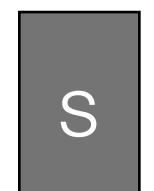


Build 2 uint64s of metadata

Data Offset



Operation

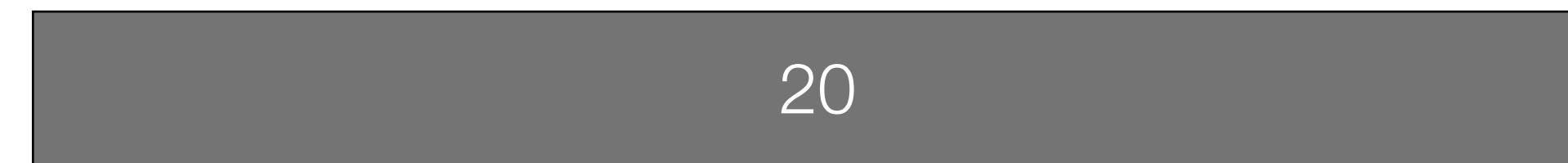


```
Set( []byte("name"), []byte("marty") )
```



Build 2 uint64s of metadata

Data Offset



Operation
Key Len



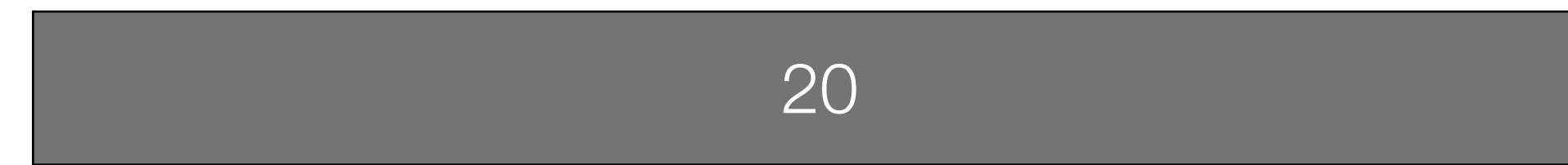
4 bits 24 bits

```
Set( []byte("name"), []byte("marty") )
```



Build 2 uint64s of metadata

Data Offset



Operation

64 bits

Key Len



Val Len

4 bits

24 bits

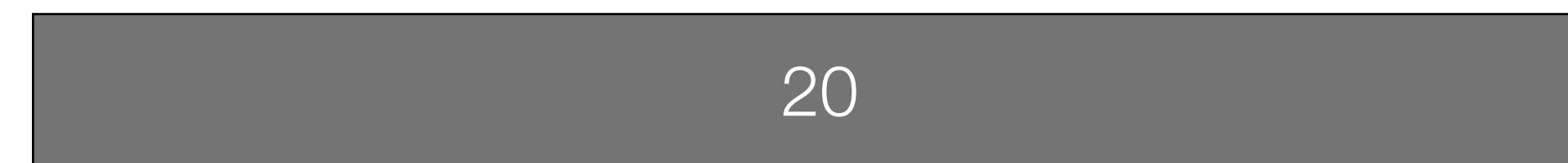
28 bits

```
Set( []byte("name"), []byte("marty") )
```



Build 2 uint64s of metadata

Data Offset



Operation

64 bits

Key Len



Val Len

4 bits

24 bits

28 bits

8 bits

```
Set( []byte("name"), []byte("marty") )
```



Append uint64s to meta slice

```
type segment struct {  
    data []byte  
    meta []uint64  
}
```

```
type segment struct {  
    data []byte  
    meta []uint64  
}
```

Sorting only shuffles integers, not bytes!

```
type segment struct {  
    data []byte  
    meta []uint64  
}
```

```
type segment struct {
```

```
    data []byte
```

```
    meta []uint64
```

```
}
```

implements...

```
type segment struct {  
    data []byte  
    meta []uint64  
}
```

implements...

- Batch

```
type segment struct {  
    data []byte  
    meta []uint64  
}
```

implements...

- Batch
- Snapshot

```
type segment struct {  
    data []byte  
    meta []uint64  
}
```

implements...

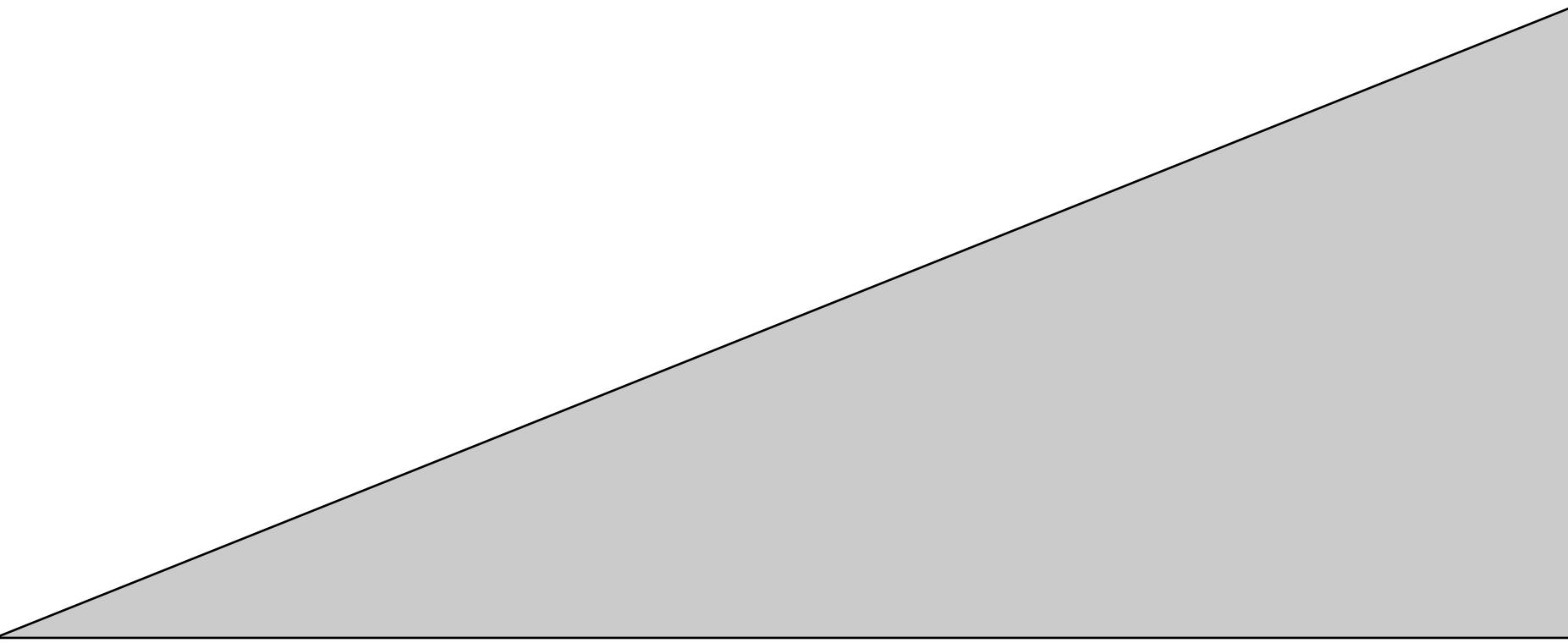
- Batch
- Snapshot
- Collection

Another Batch?

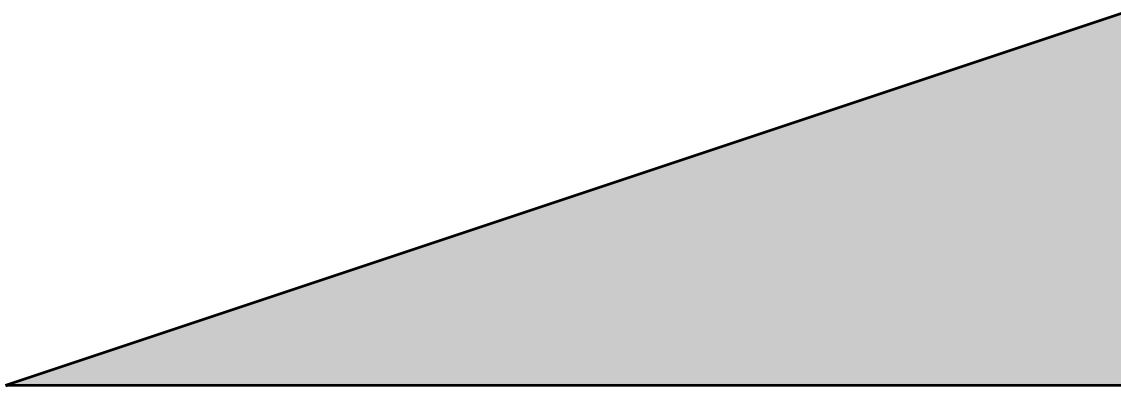
Zoals in Noordwijk het licht van de vuurzuren langs de vier windrichtingen overvalt, zo streken elk jaar in grote vlaggen die vier seisoenen over Groot Rechtersen. Max kende de wisseling van de jaargrijsen eigenlijk alleen uit Amsterdam: op een dag in februari of maart is er een onbeschrijflijke voorjaarsgeur als hij 's morgens op straat loopt; onbegrijpbaar als de decimalen van π , die stoffige somer waarin de stad vol moeststen stroomde, even plotseling veranderend in de voorjaar, herfst, en dan de bleke winter waarin de sterren van de straten en de muren opeens de ongemaakbare aard van de wereld leken uit te drukken, - maar dat alles eigenlijk in het voorjaar, oppgemerkt alleen in de korte tussenposes waarin hij van het ene interieur naar het andere ging. Was in de stad de natuur tachtig achtereenvolgende, op het kasteel zat hij met Quinten en Sjipje midden in de daverende concertzaal. Lente en herfst kwamen met kolossaal vertoon, de somers waren hoter en droger, de winters koudter en winter. Die onafgebonden verandering, had hij nooit tegen Ottus beweerd, was natuurlijk de bron van alle creativiteit: de onvermoeidheid van de natuur tussen dit keerkringen hieldde ook toe culturele stilstand, de tropen waren een onafgebonden land, altijd groen zoals de poolgebieden altijd wit waren, van de primitieve breedtegraden met hun viervormigheid waren



More Batches, a stack of Segments



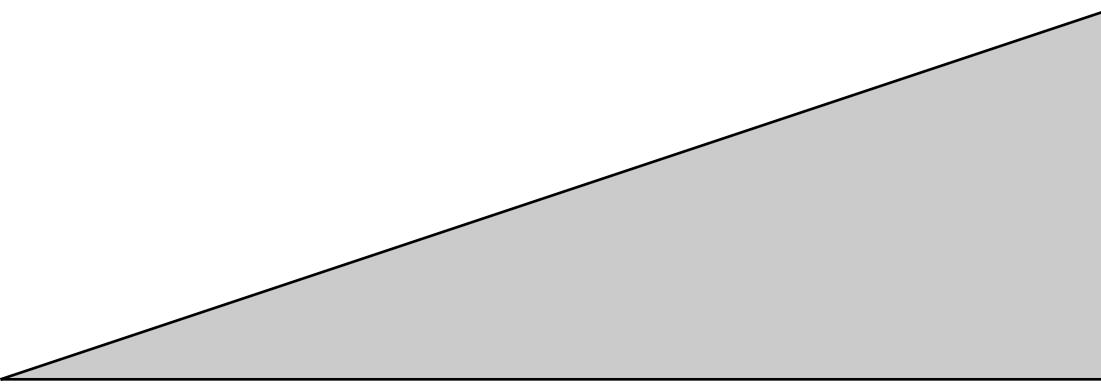
More Batches, a stack of Segments



segmentStack
 $[]^*segment$

More Batches, a stack of Segments

Incoming Batch

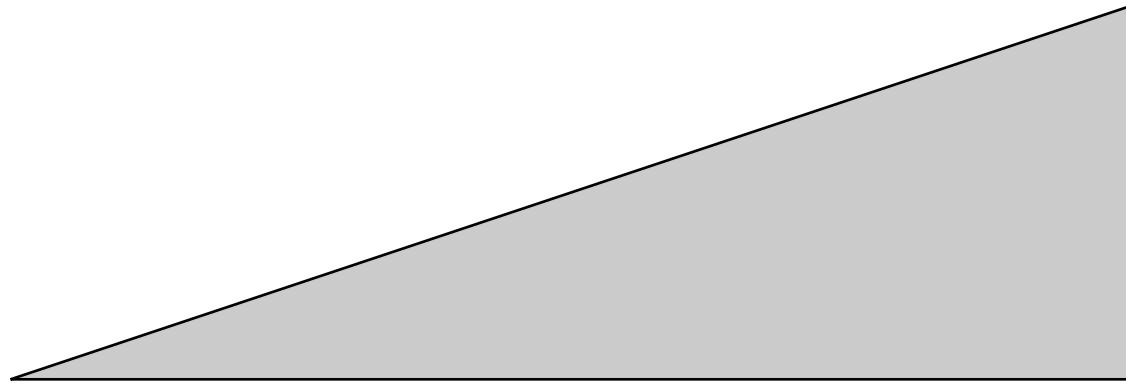


segmentStack
 $[]^*segment$

More Batches, a stack of Segments

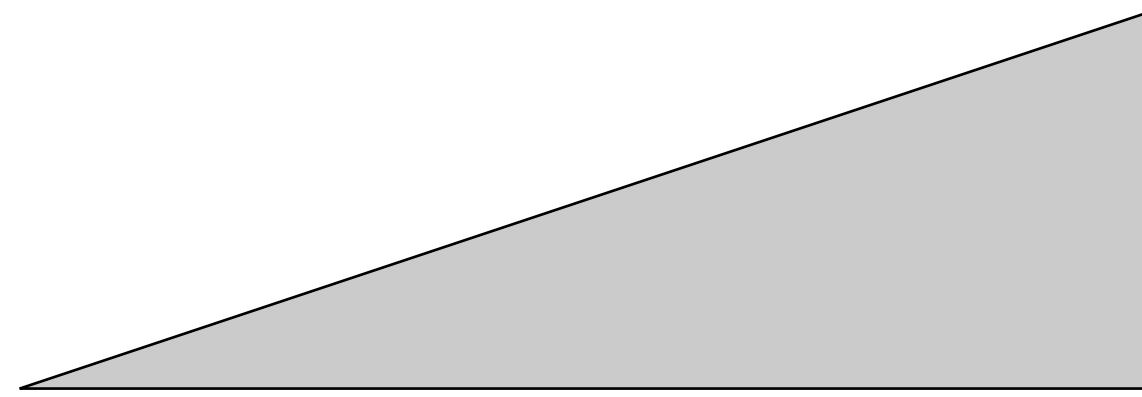
Incoming Batch

sort in place

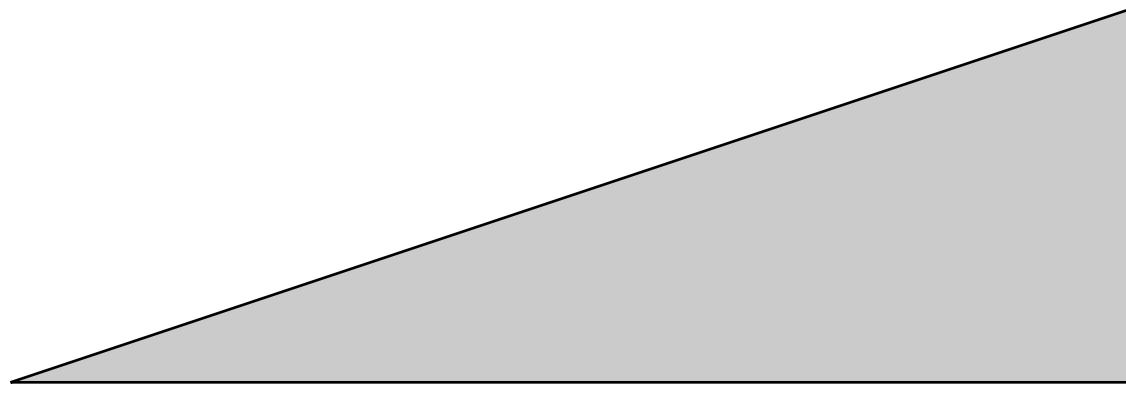


segmentStack
 $[]^*segment$

More Batches, a stack of Segments

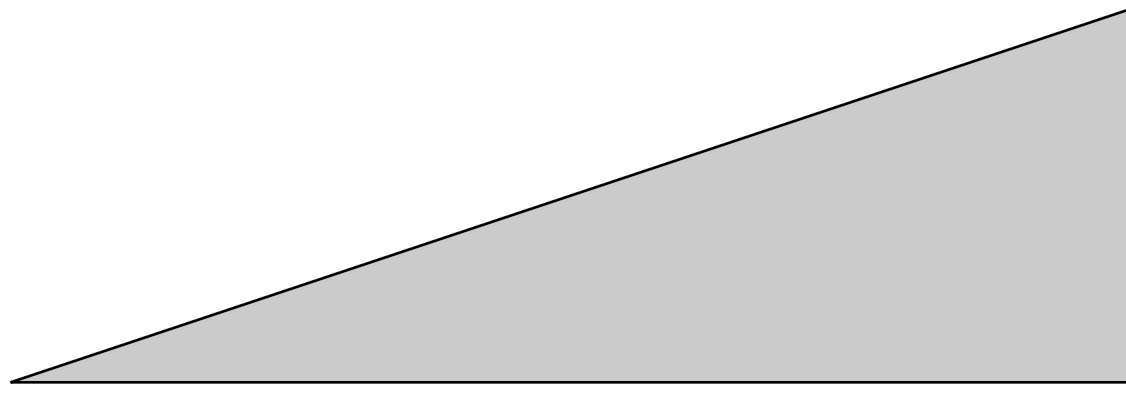
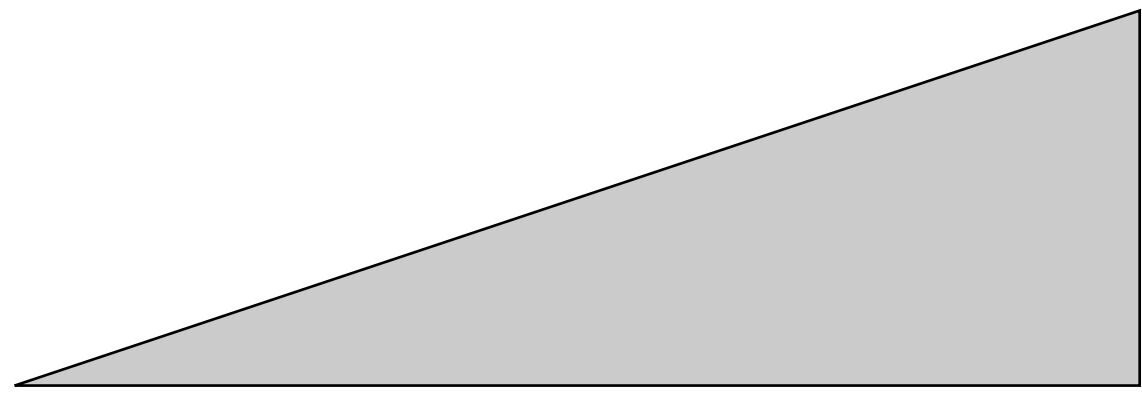


sort in place



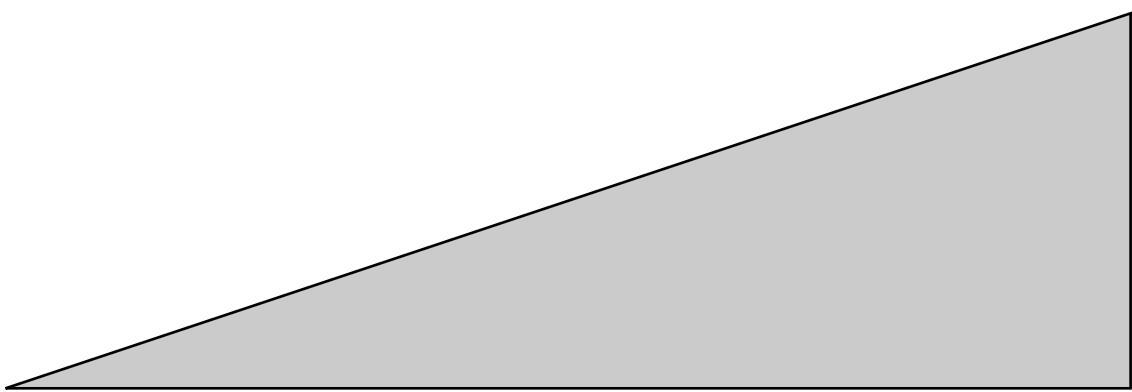
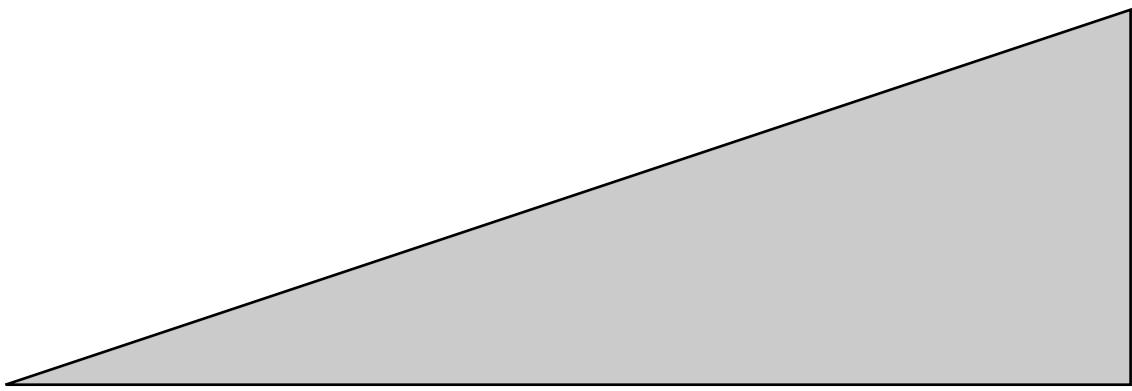
segmentStack
 $[]^*segment$

More Batches, a stack of Segments



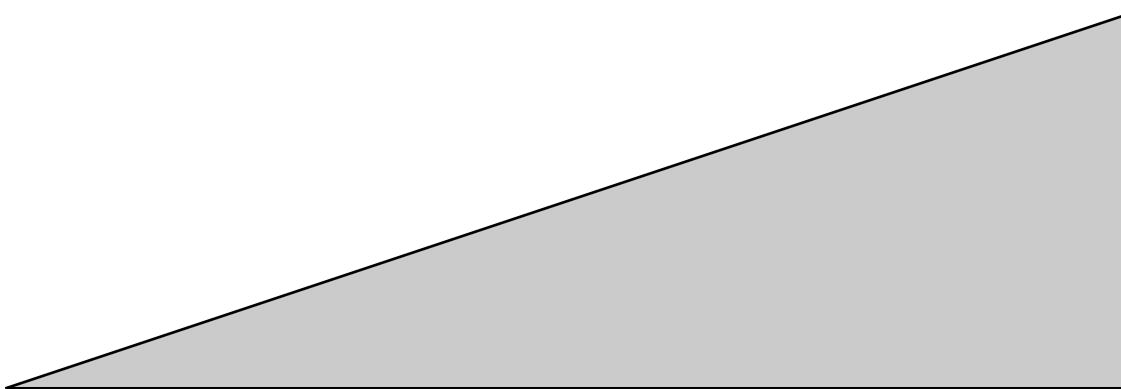
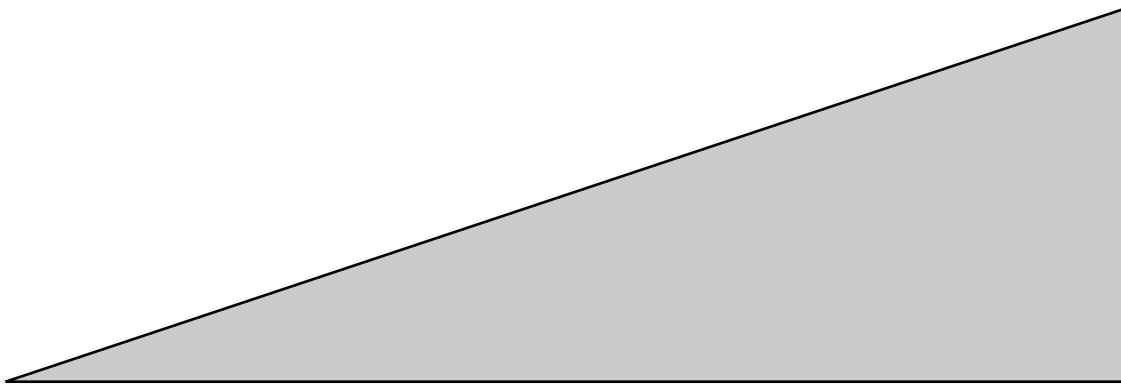
segmentStack
[]*segment

More Batches, a stack of Segments



segmentStack
[]*segment

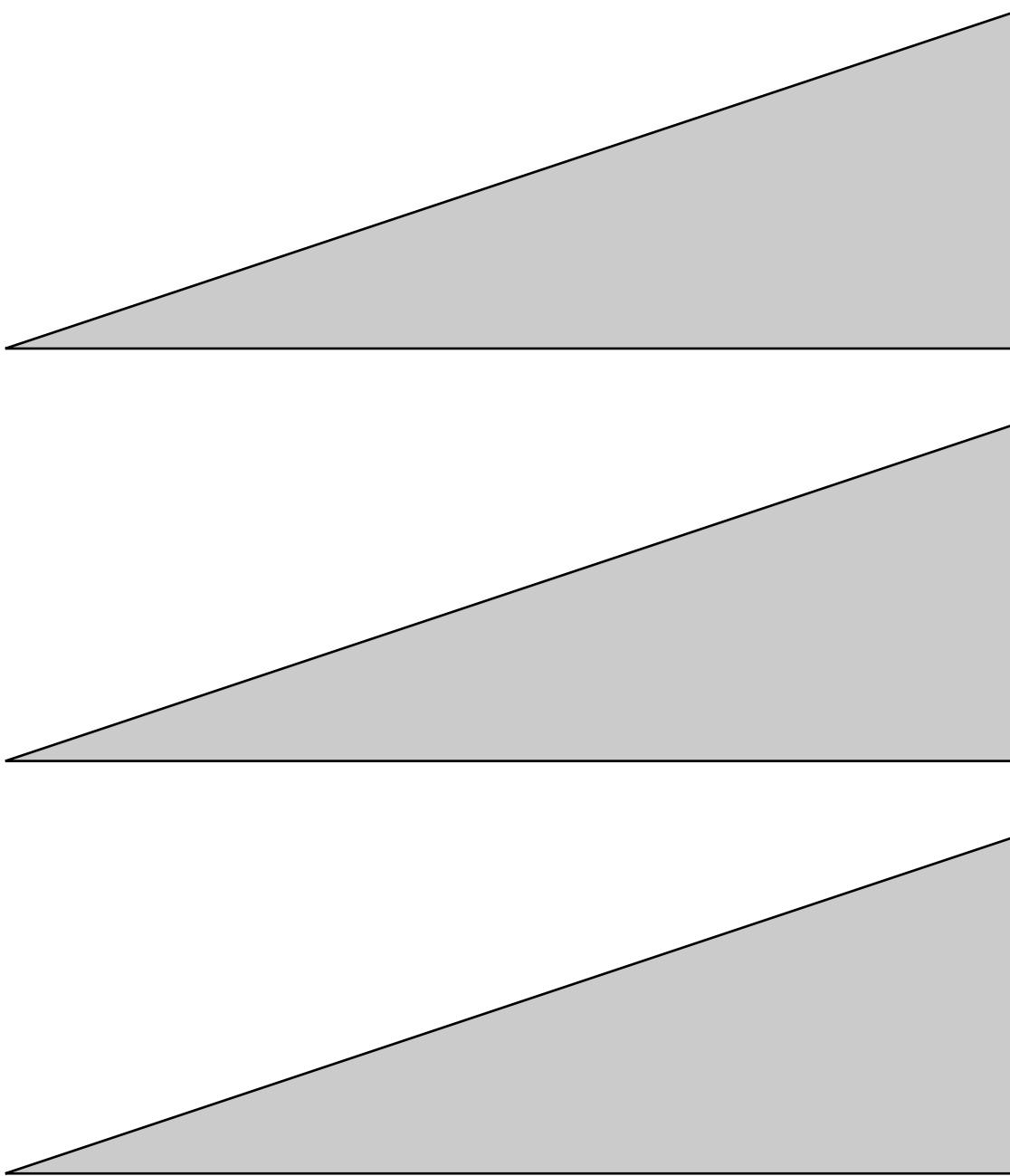
More Batches, a stack of Segments



segmentStack
 $[]^*segment$

Get

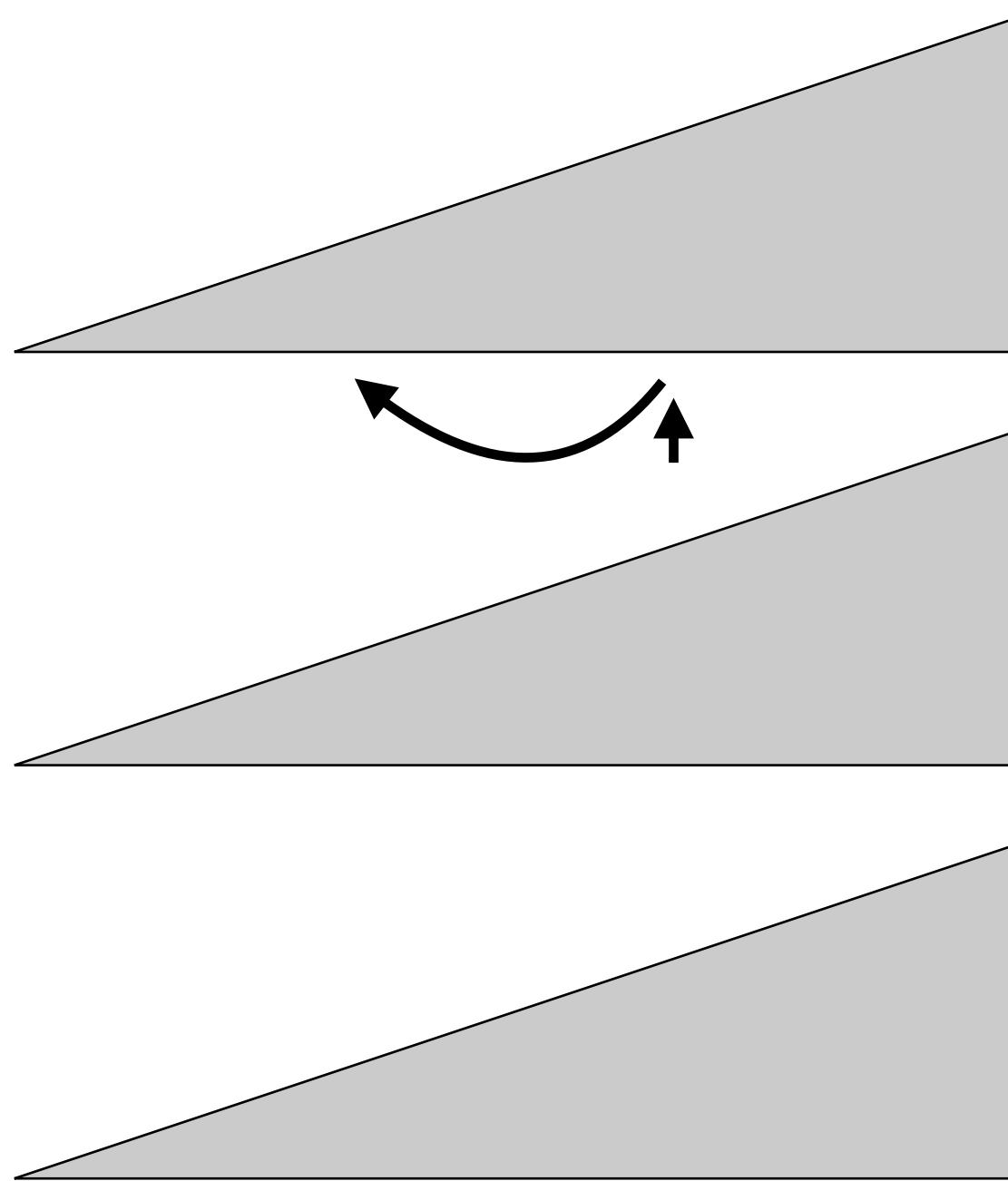
Newer Segments
Shadow Older Ones



segmentStack

Get

Newer Segments
Shadow Older Ones



Binary Search
Down the Stack

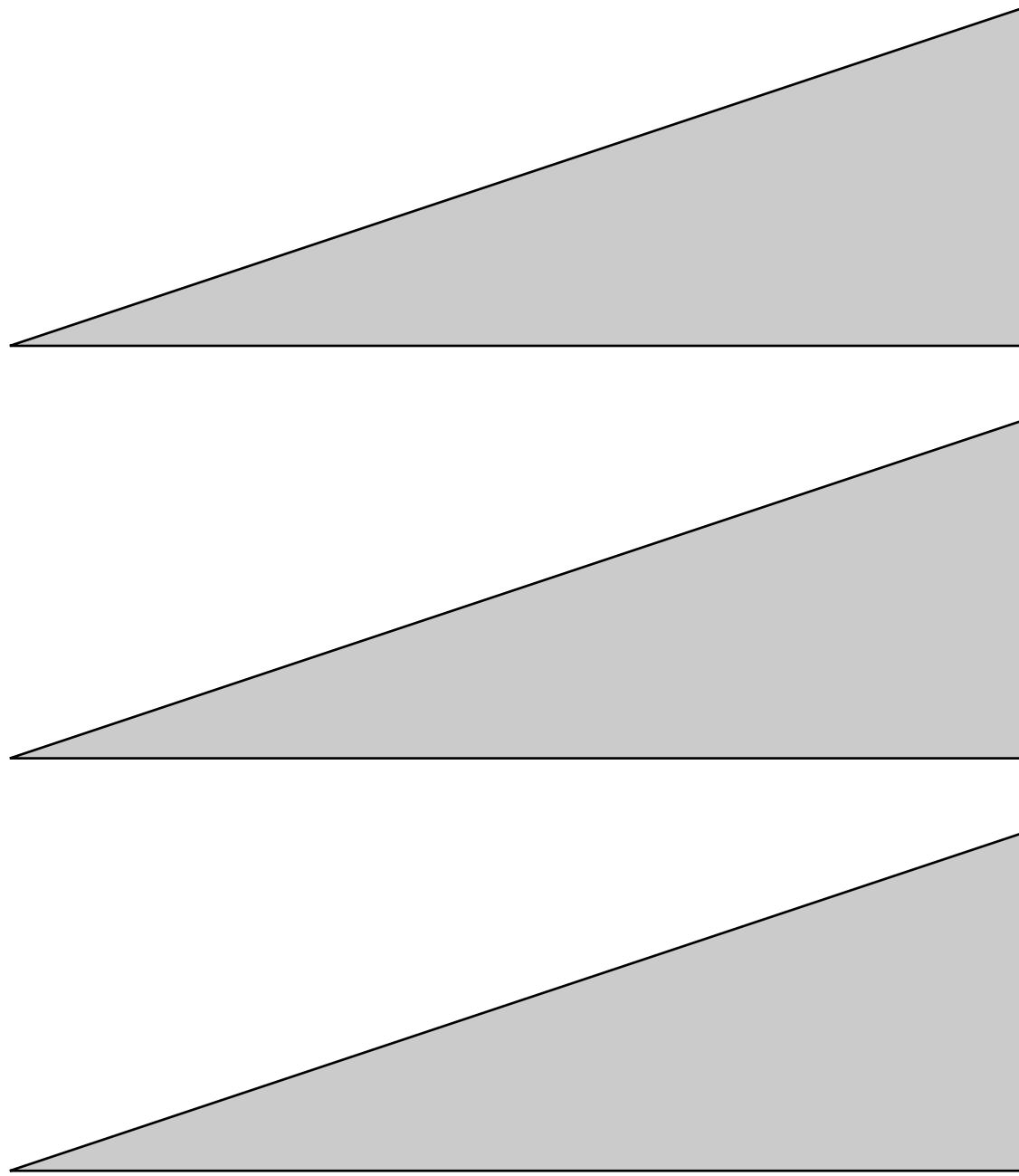


segmentStack

Get

Newer Segments
Shadow Older Ones

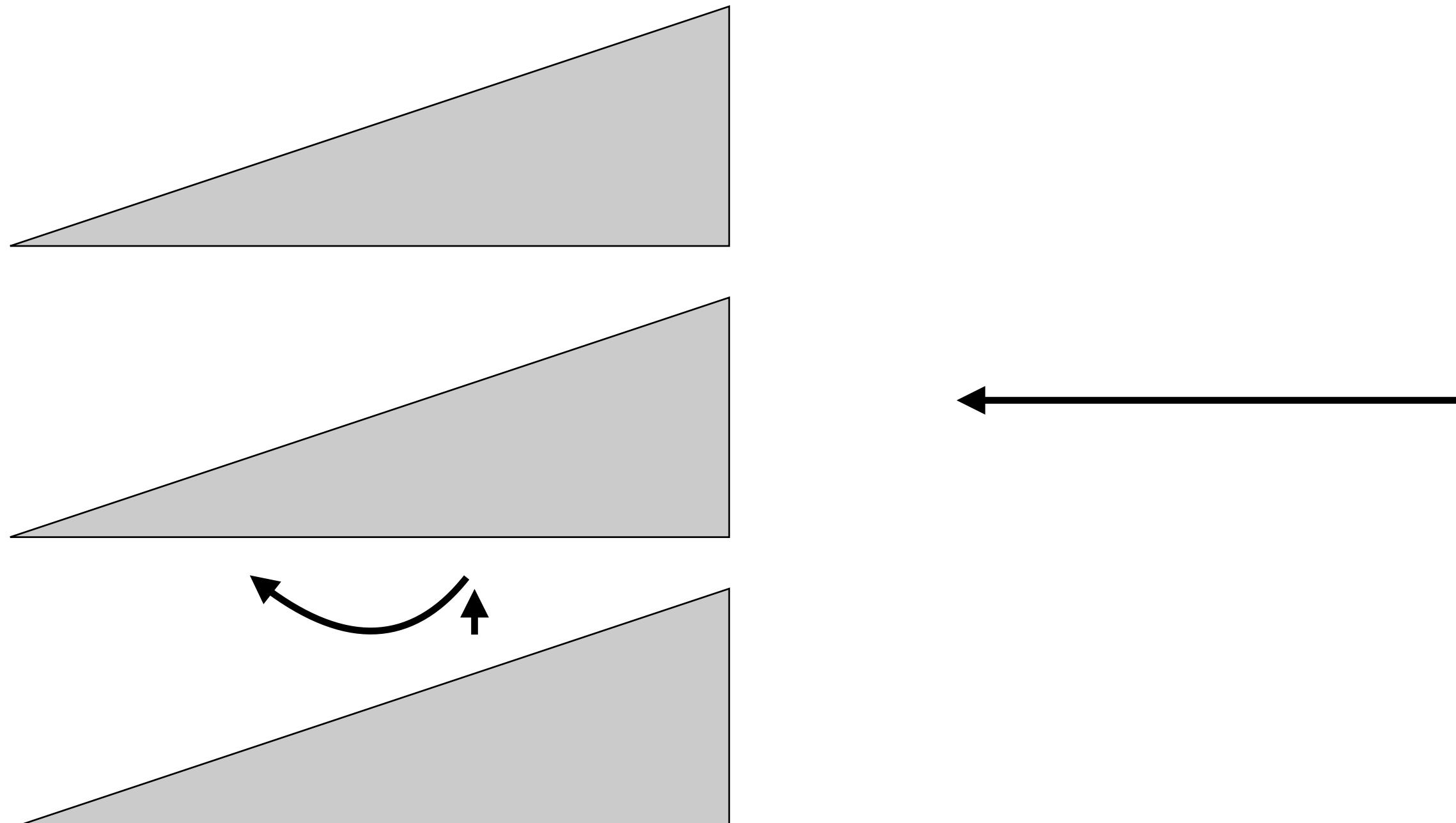
Binary Search
Down the Stack



segmentStack

Get

Newer Segments
Shadow Older Ones



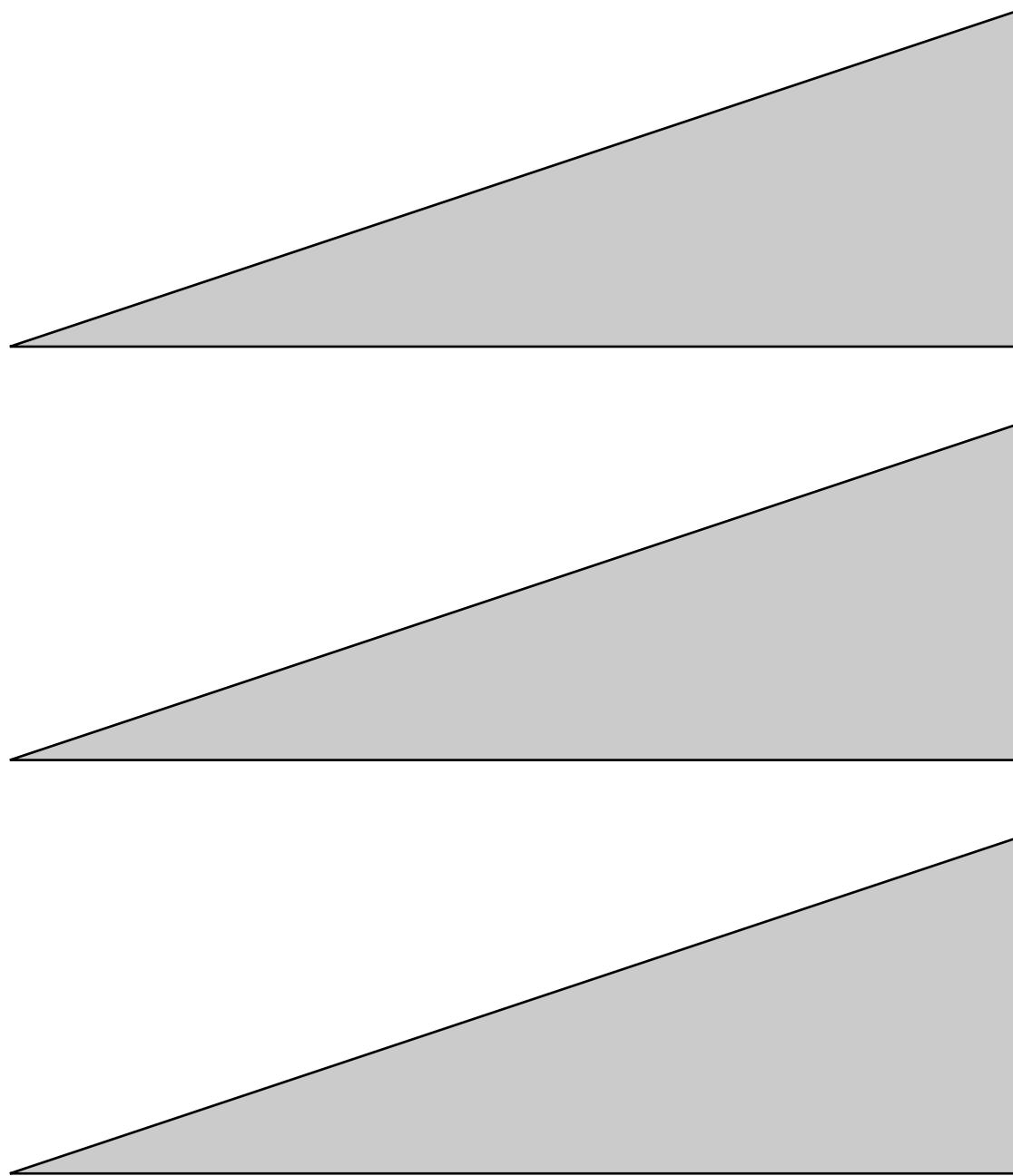
segmentStack

Binary Search
Down the Stack

Get

Newer Segments
Shadow Older Ones

Binary Search
Down the Stack

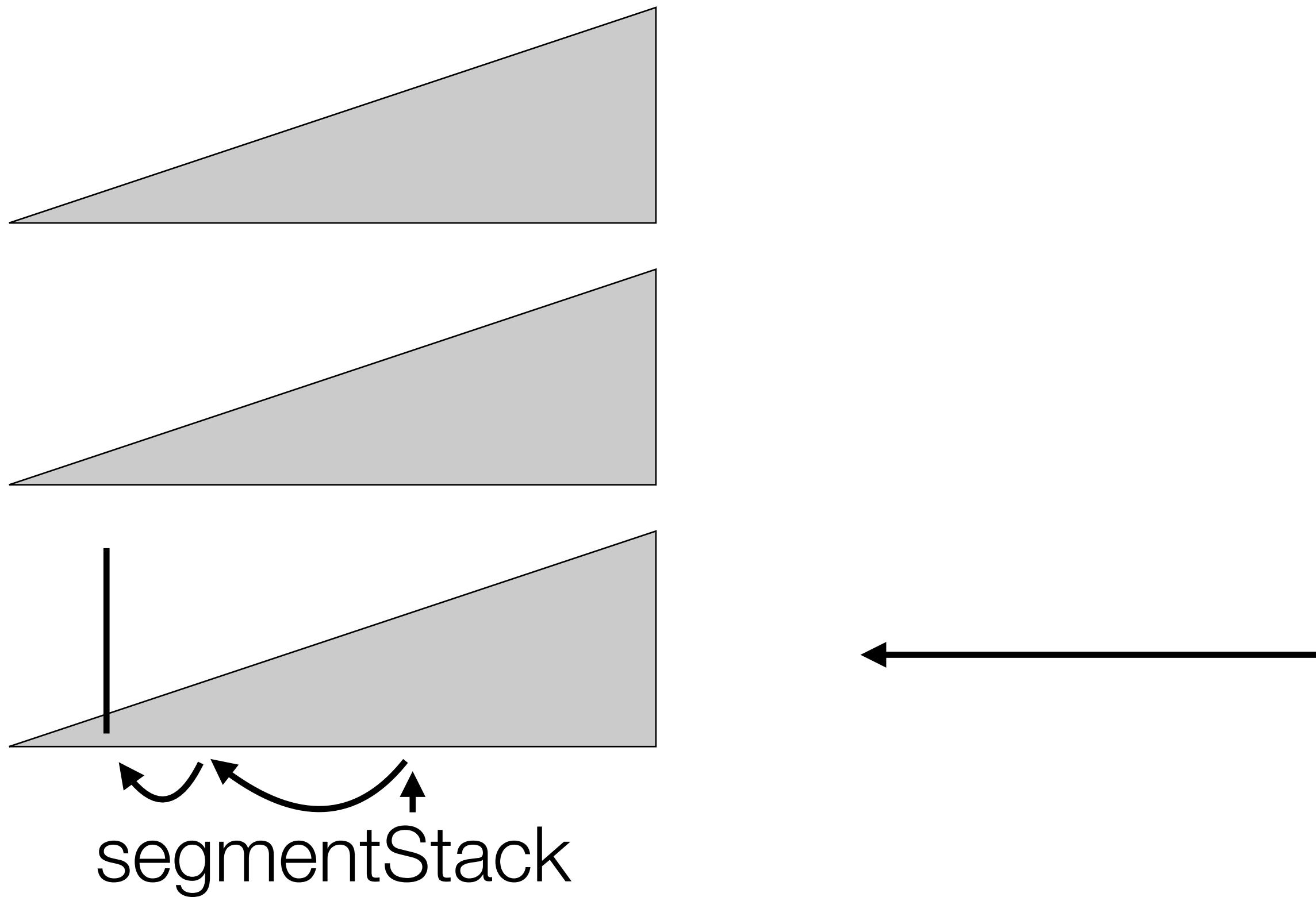


segmentStack

Get

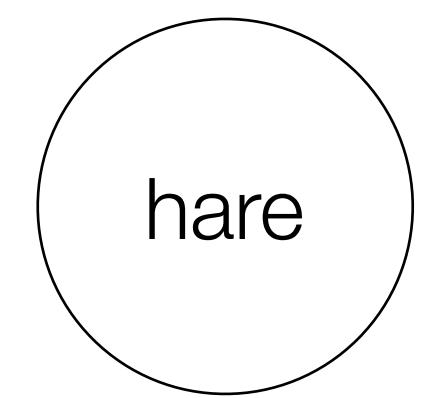
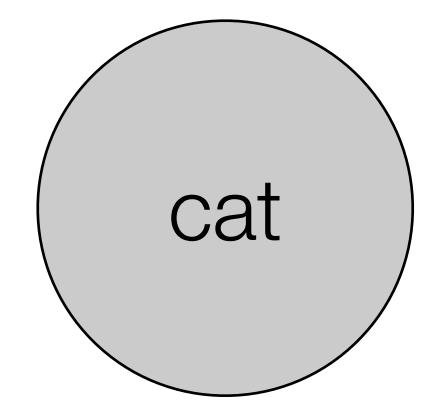
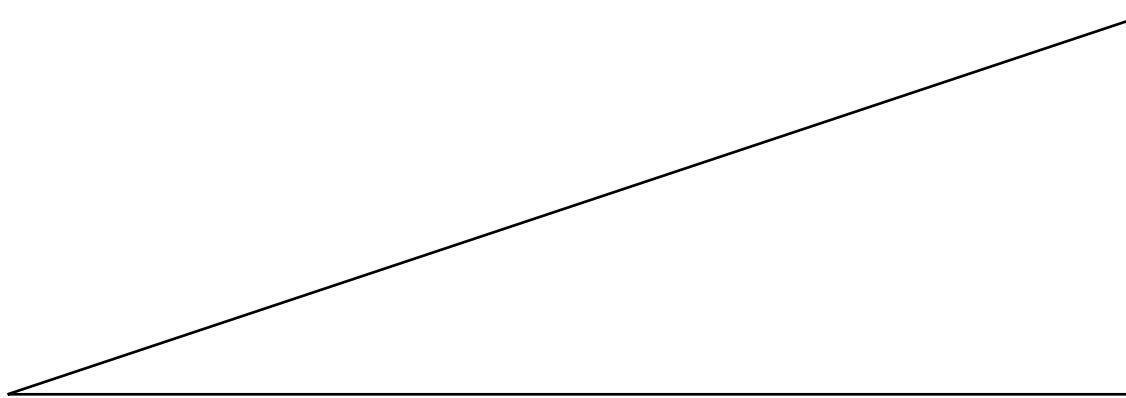
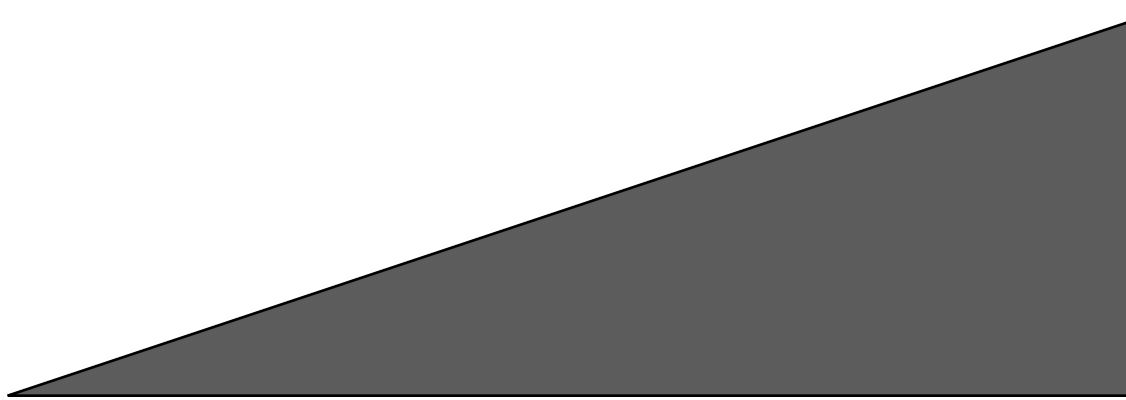
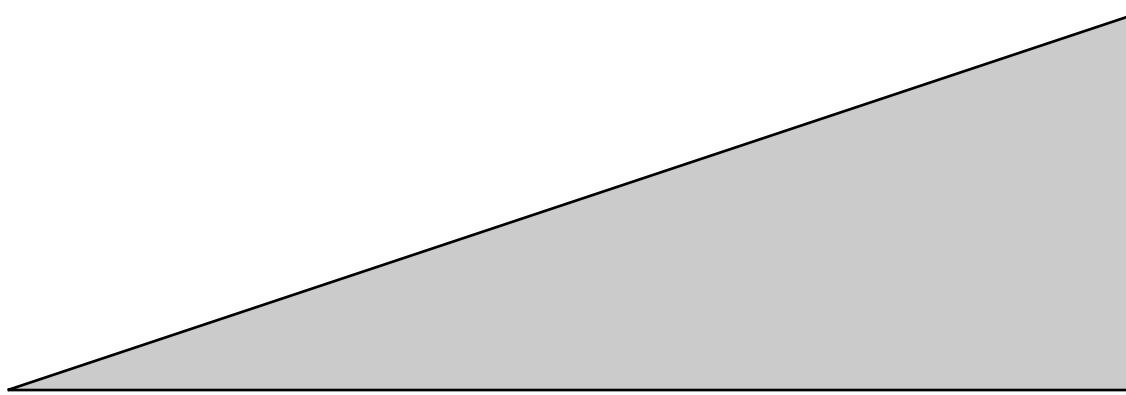
Newer Segments
Shadow Older Ones

Binary Search
Down the Stack



Iteration - using container/heap

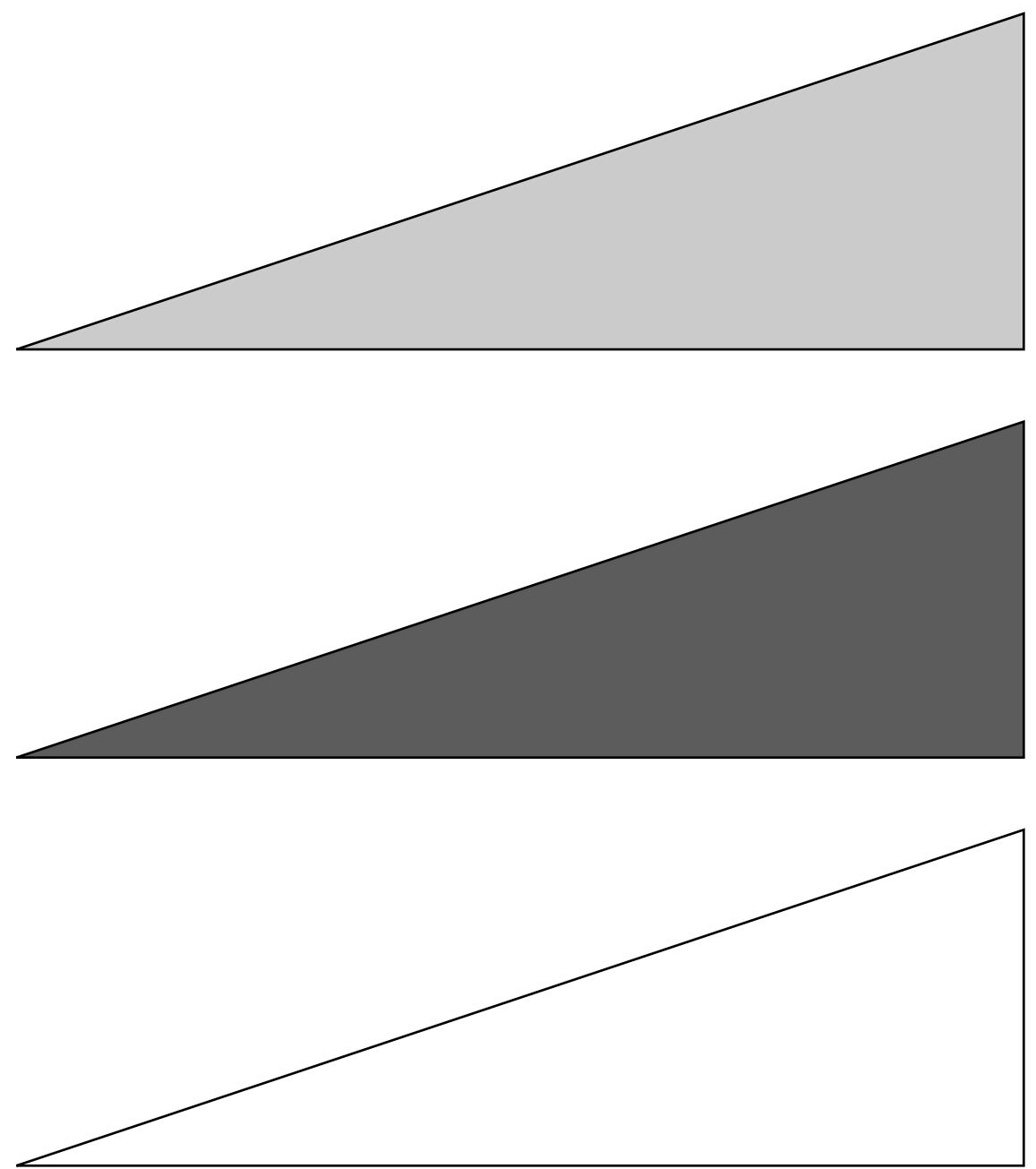
Iterators for each segment



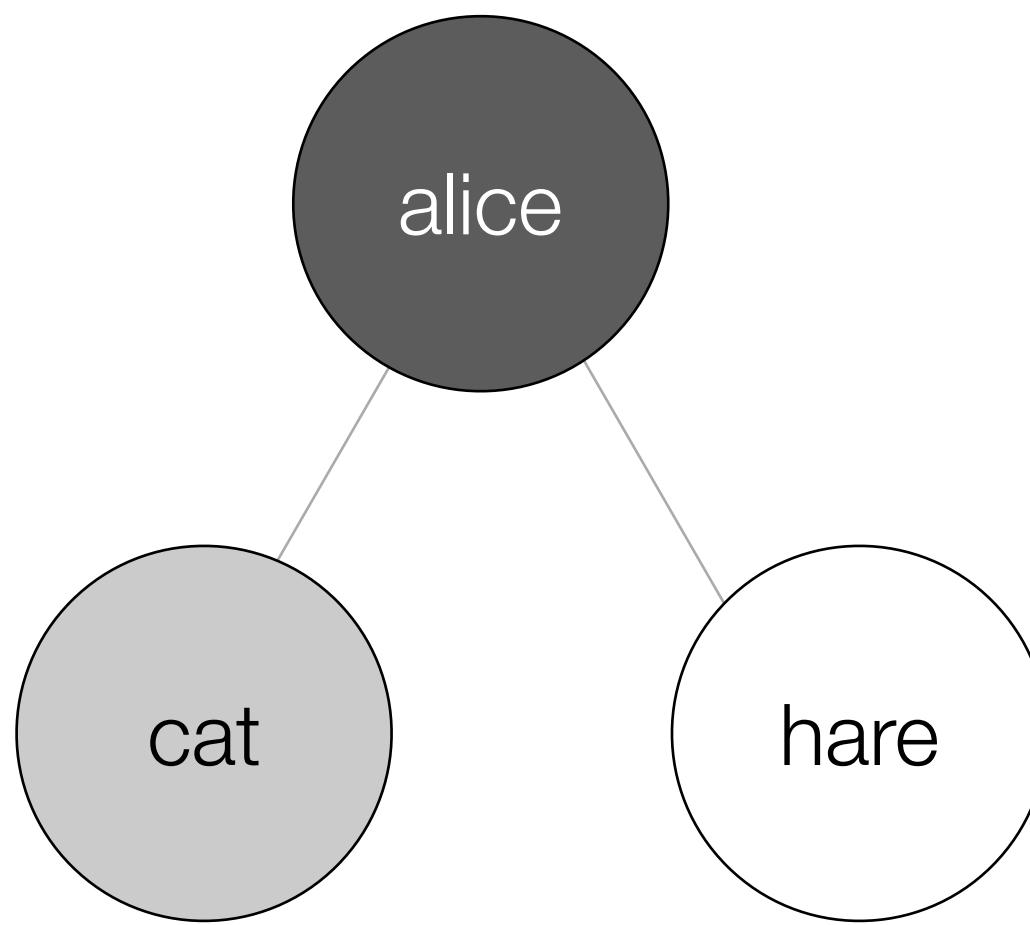
segmentStack

Iteration - using container/heap

Iterators for each segment

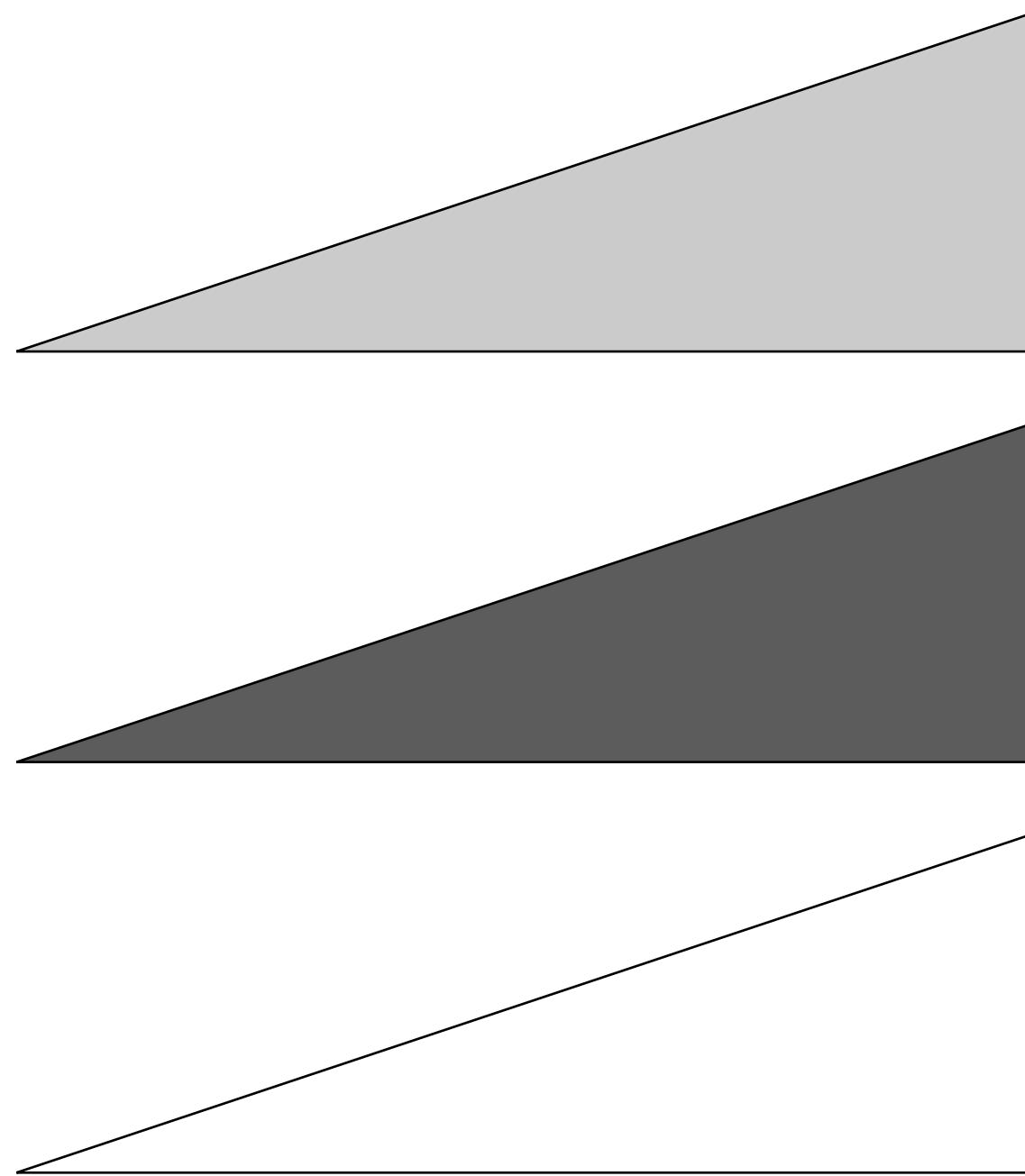


segmentStack

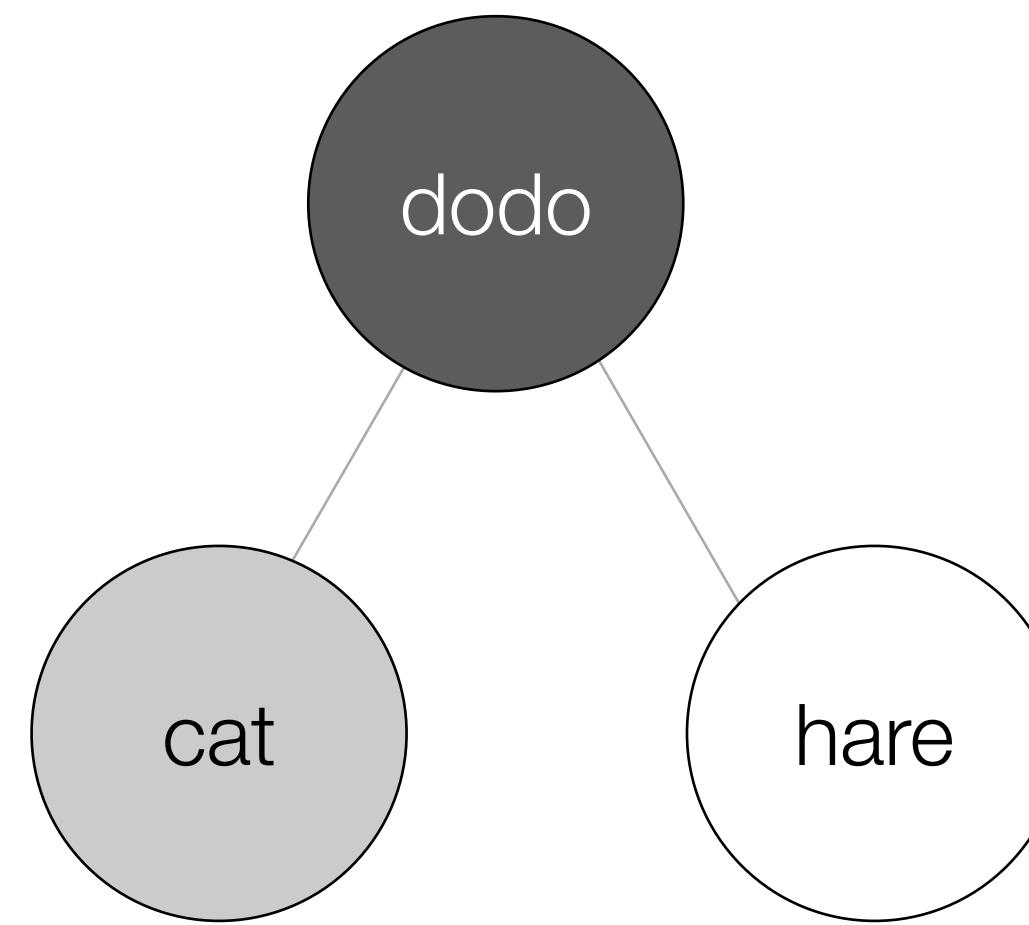


Iteration - using container/heap

Iterator returns 'alice'

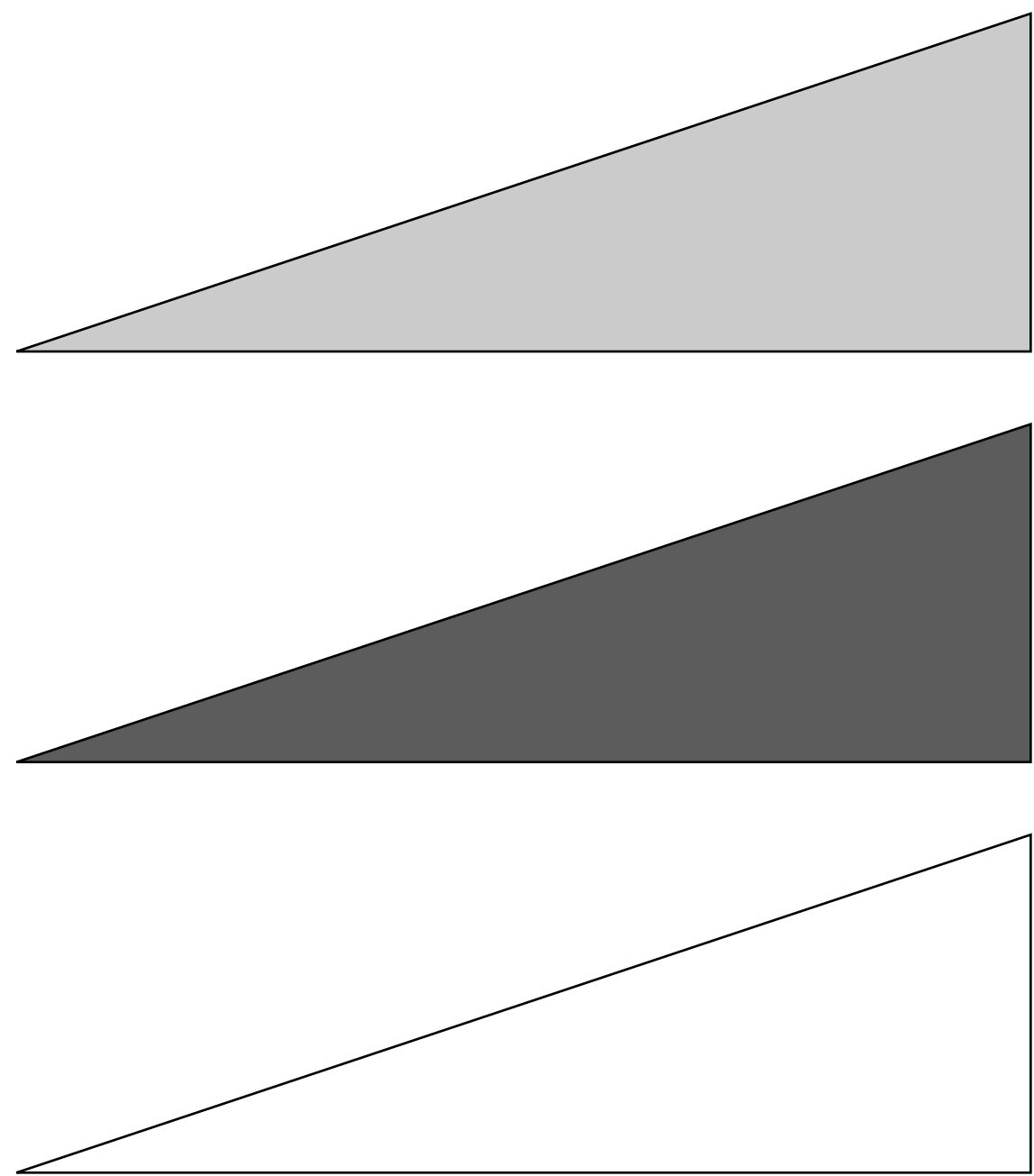


segmentStack

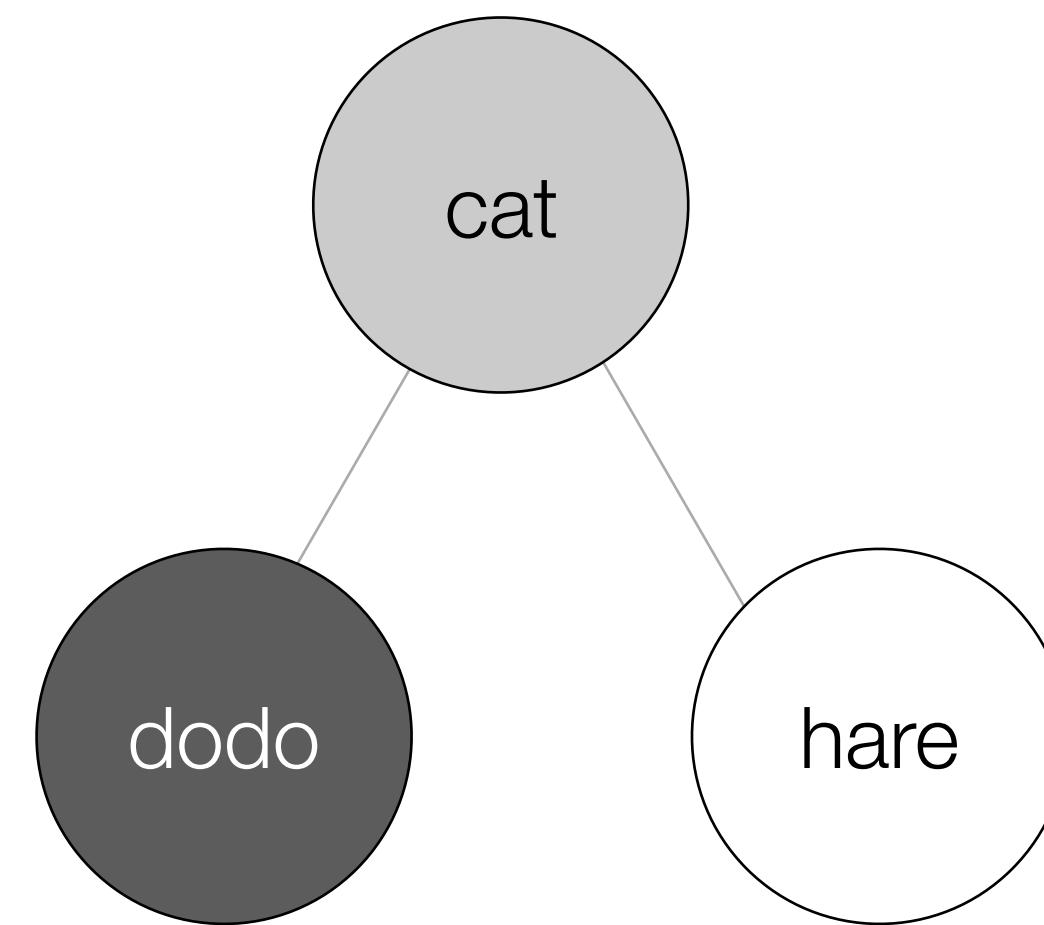


Iteration - using container/heap

Use heap Fix() method

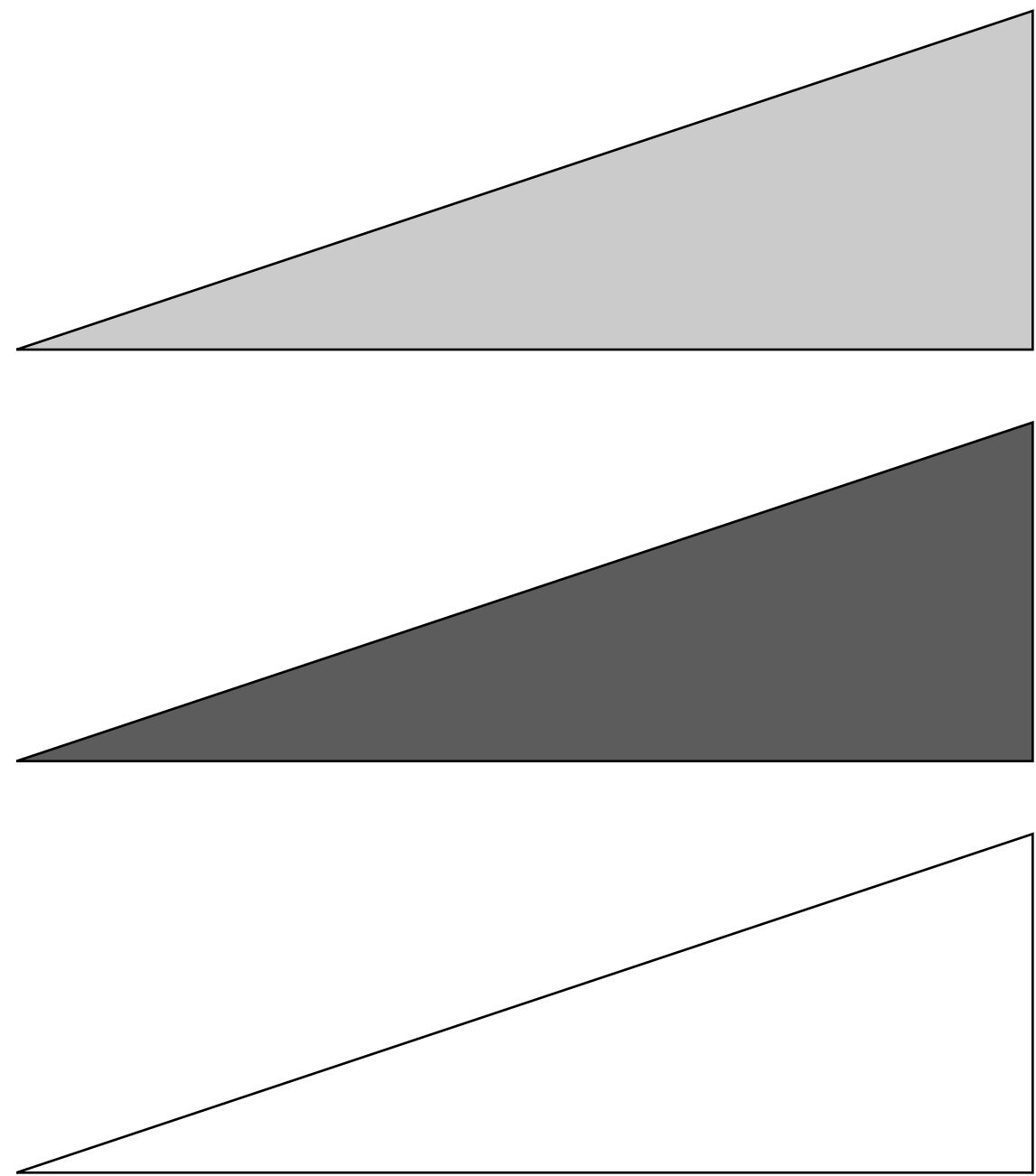


segmentStack

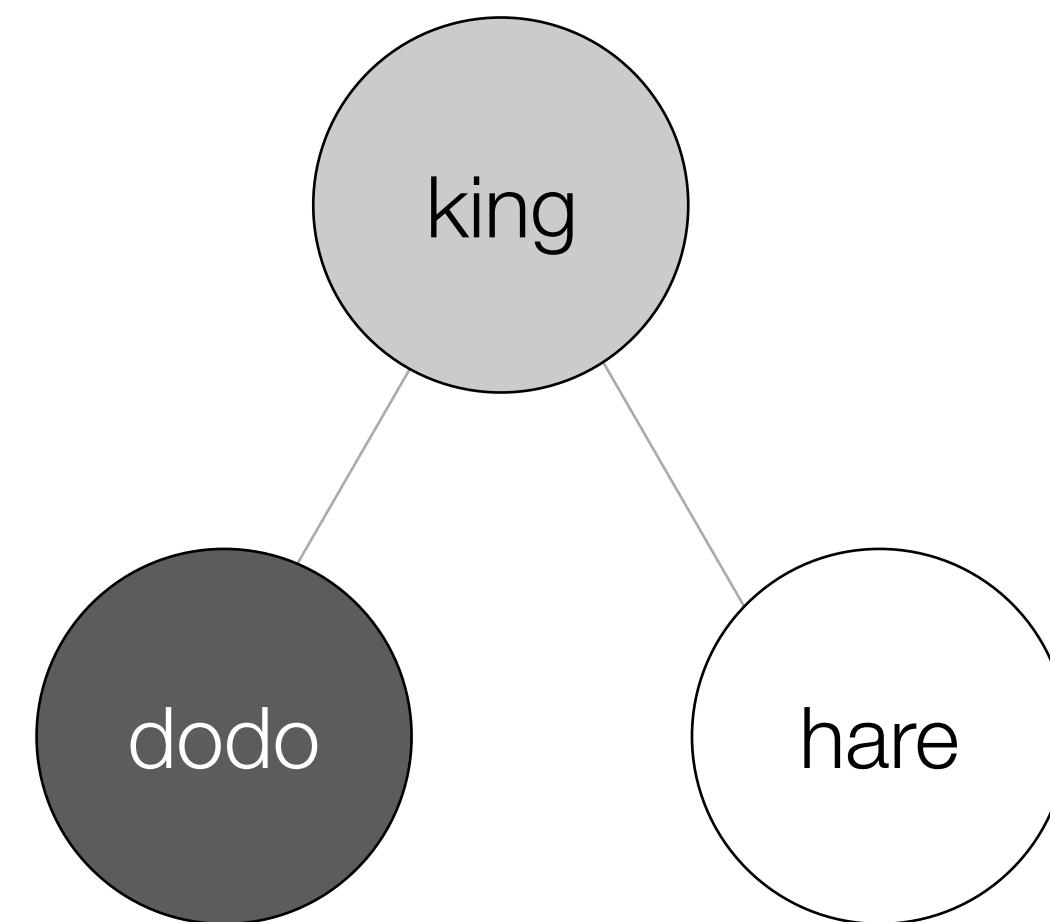


Iteration - using container/heap

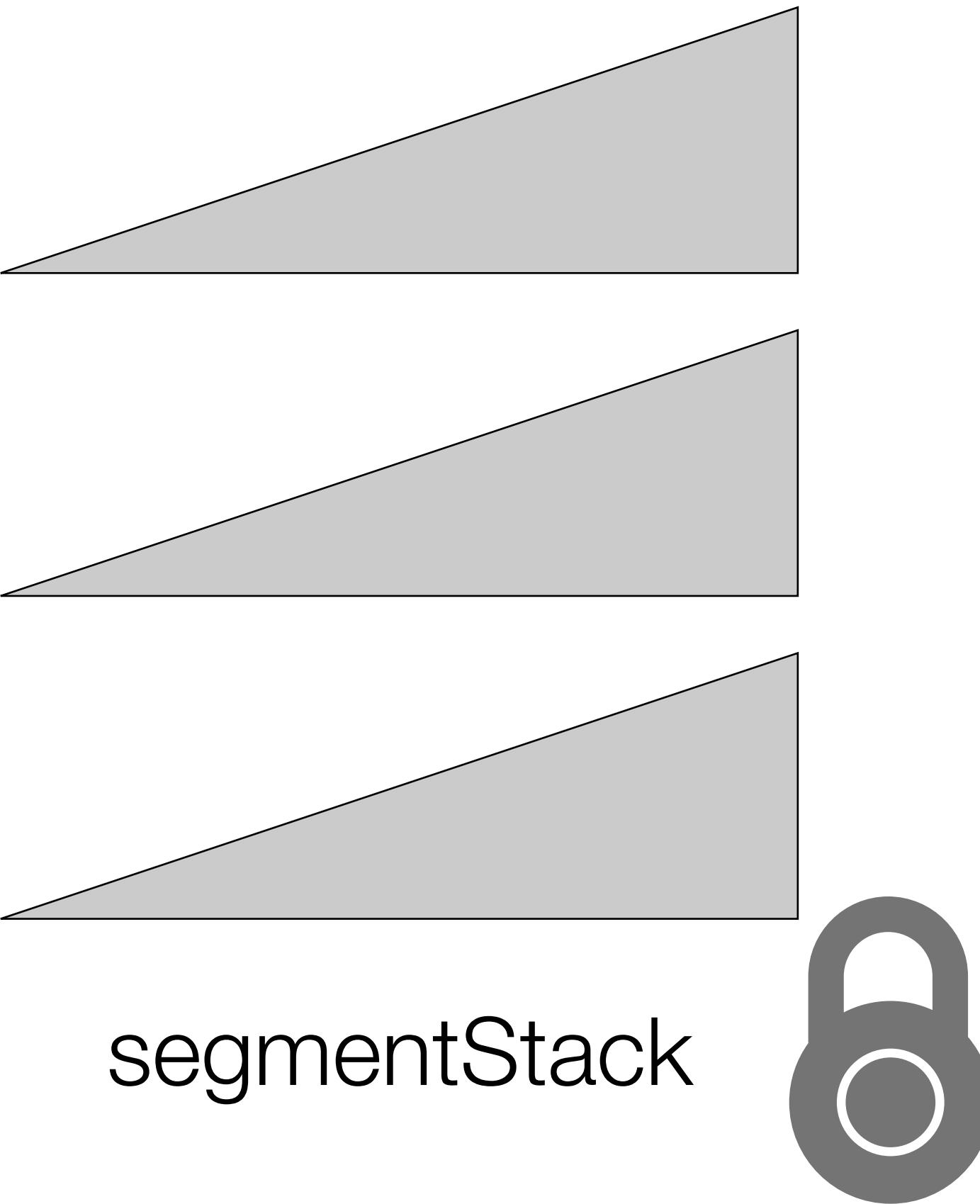
Iterator returns 'cat'



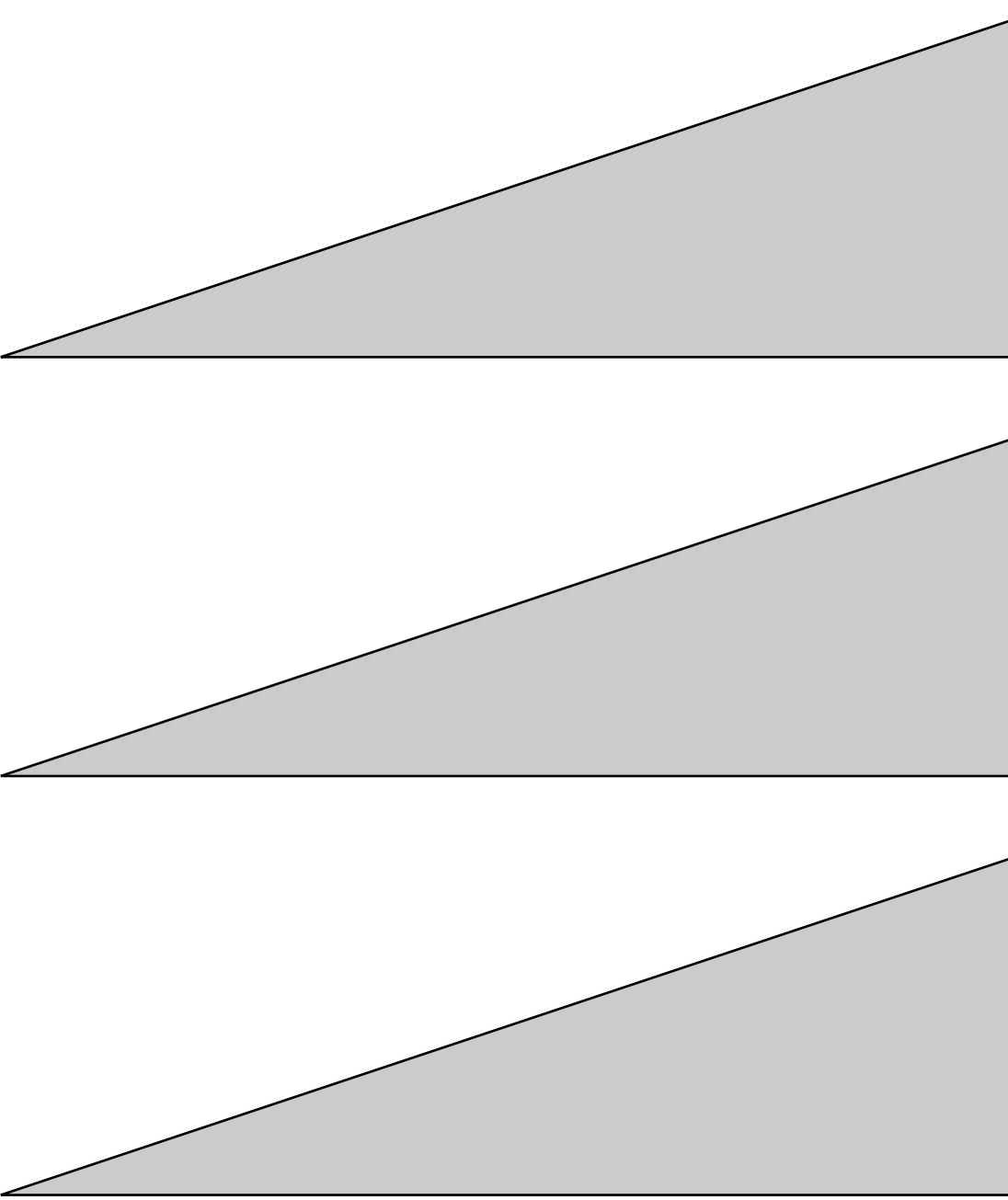
segmentStack



Atomic Batches

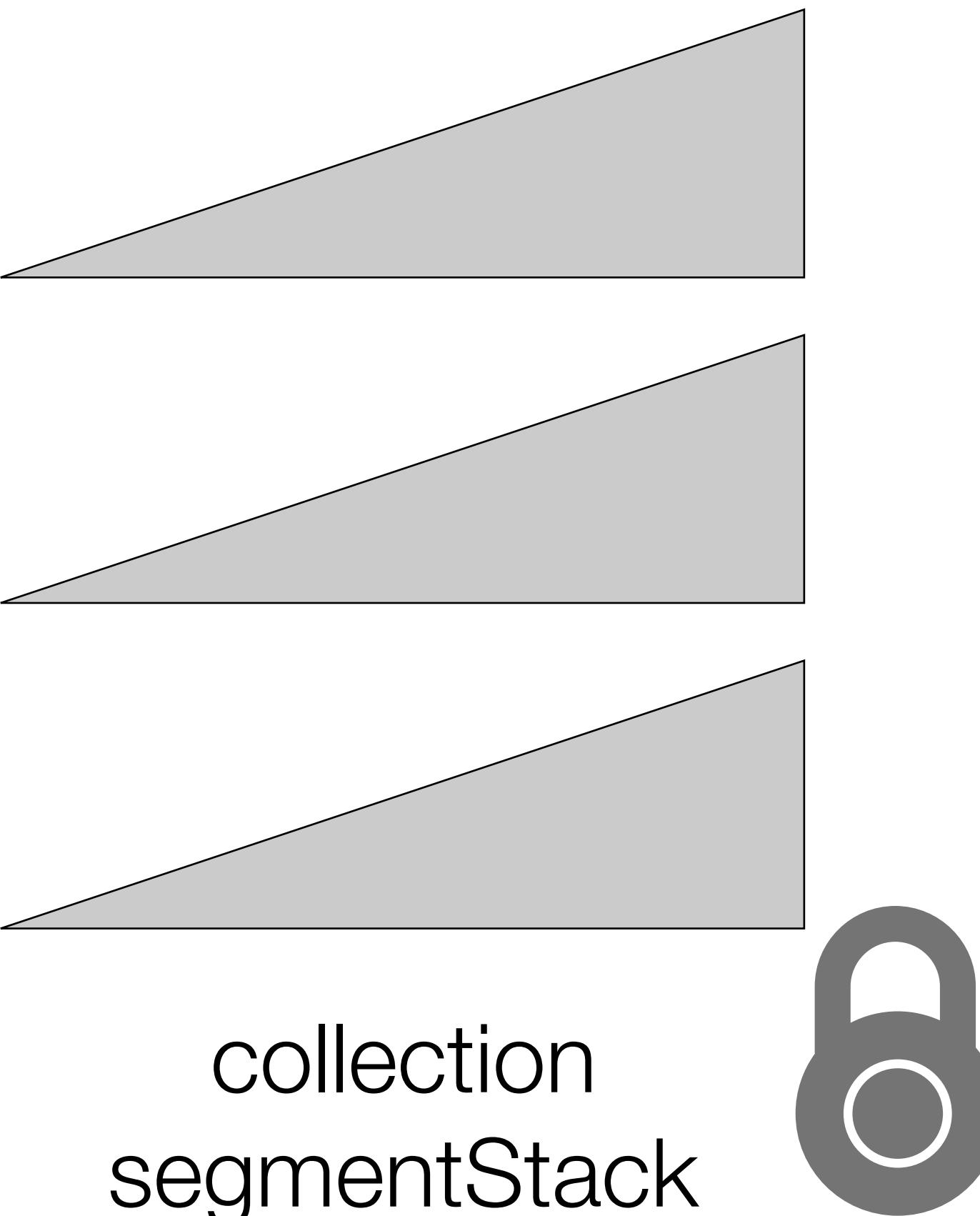


Isolated Snapshots

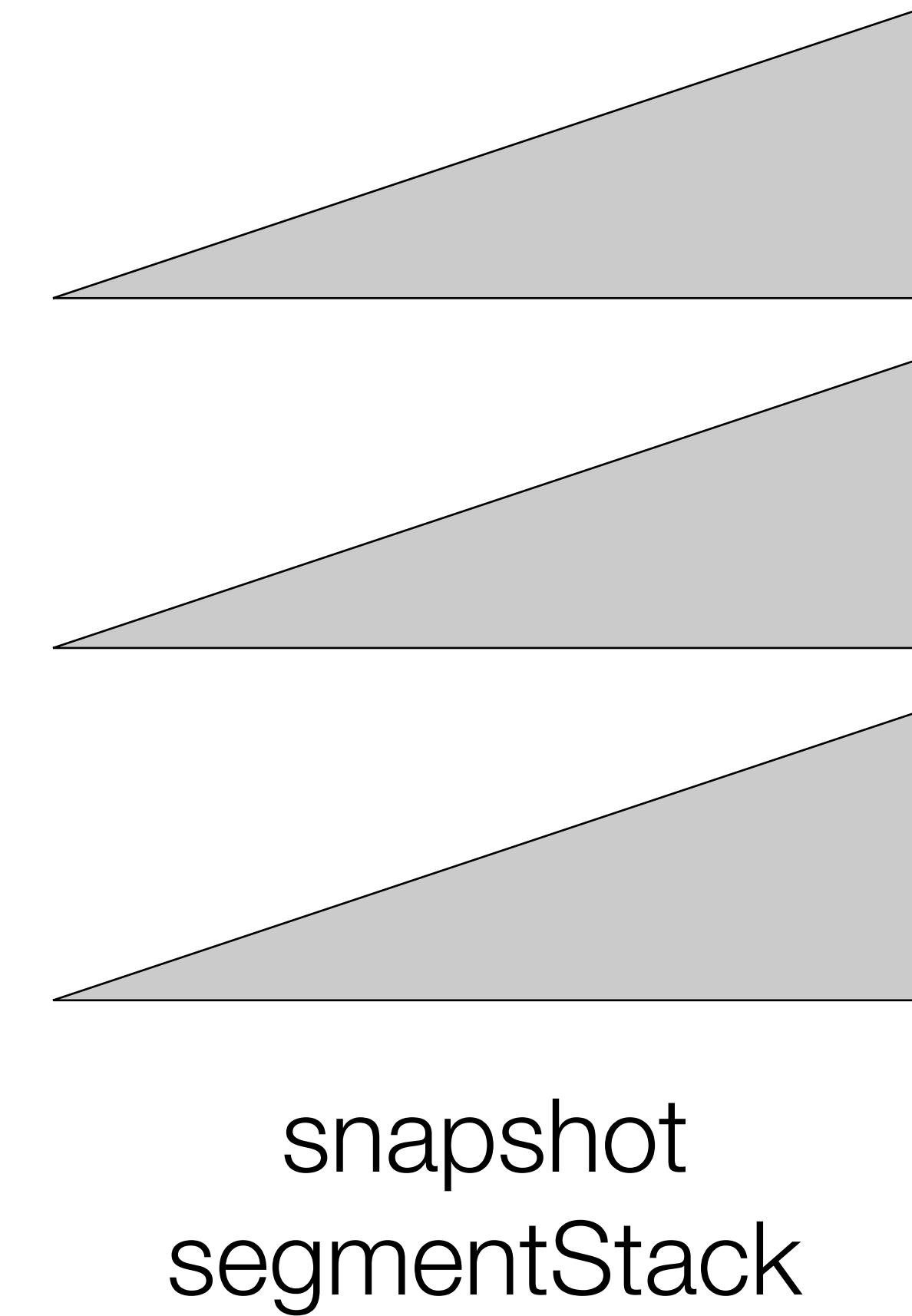
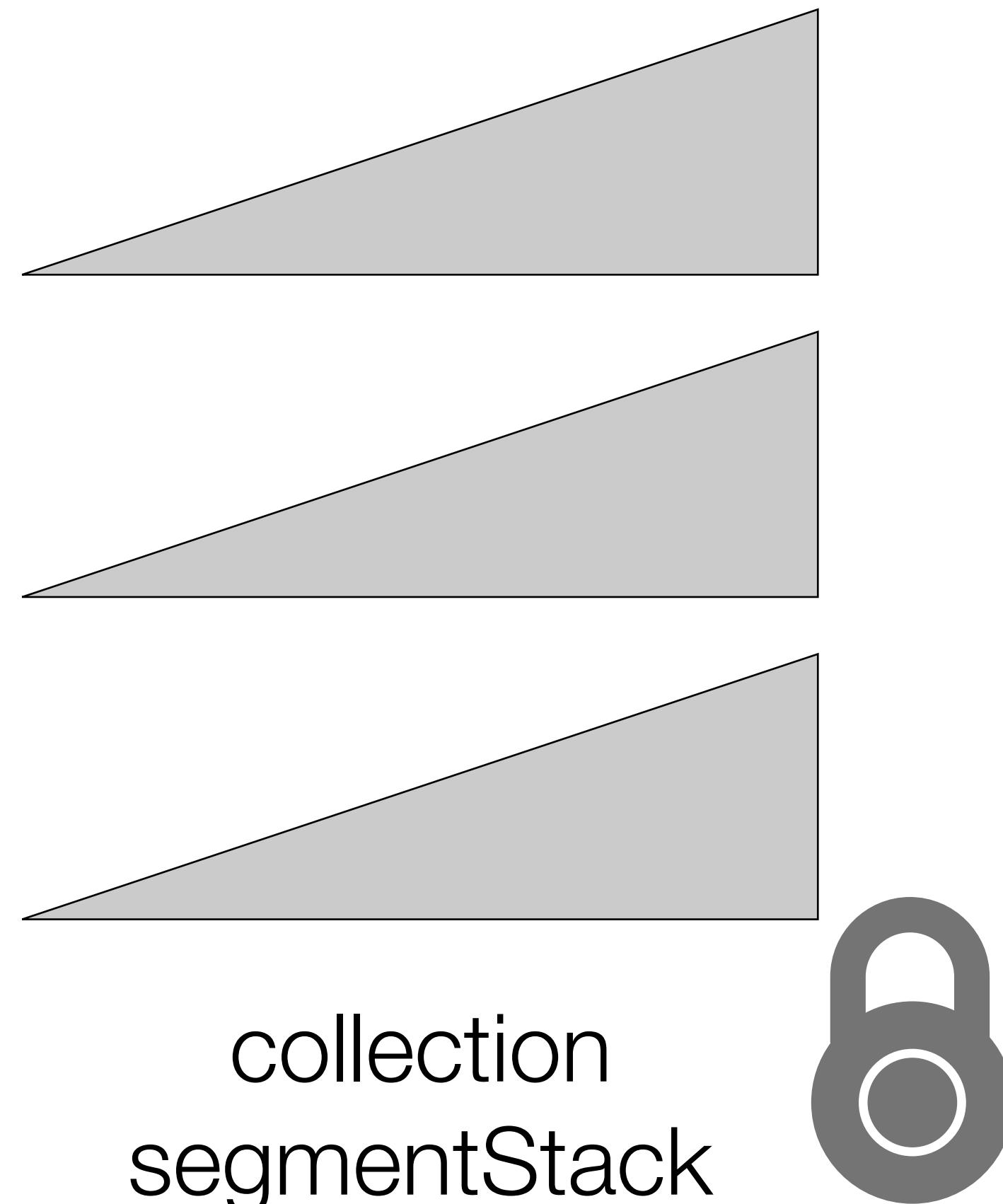


collection
segmentStack

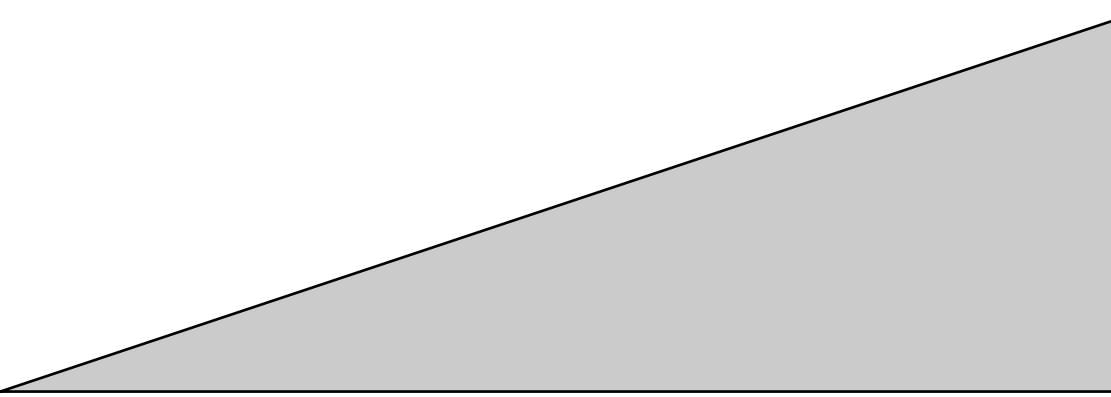
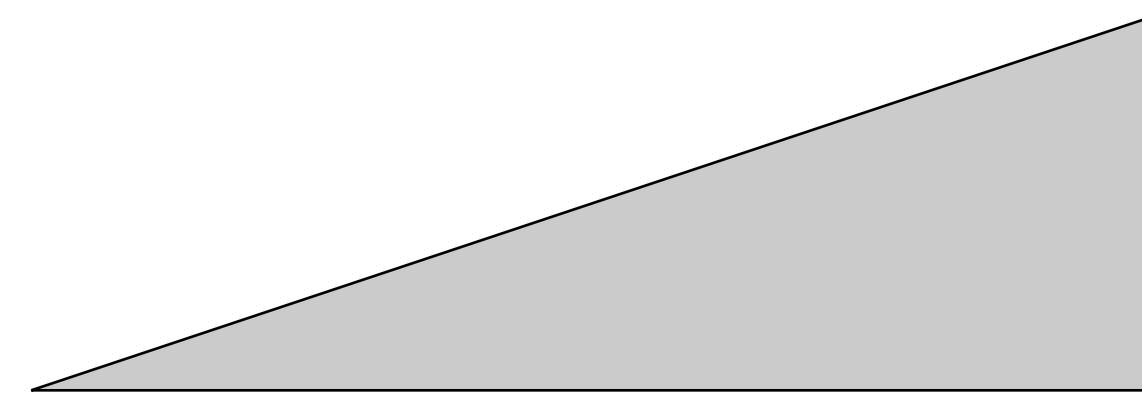
Isolated Snapshots



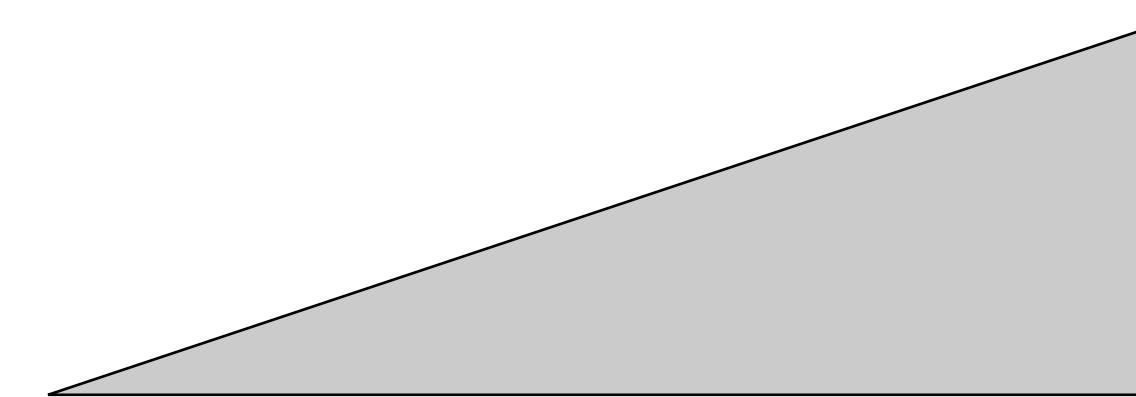
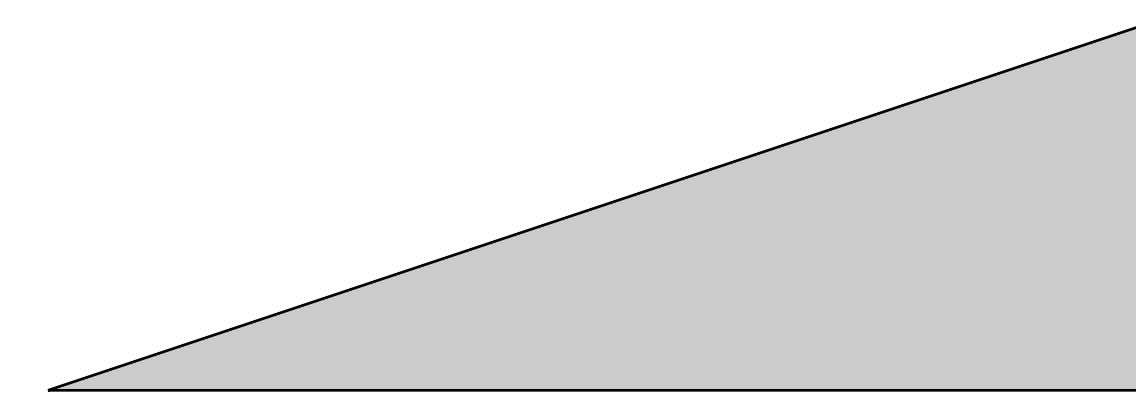
Isolated Snapshots



Isolated Snapshots

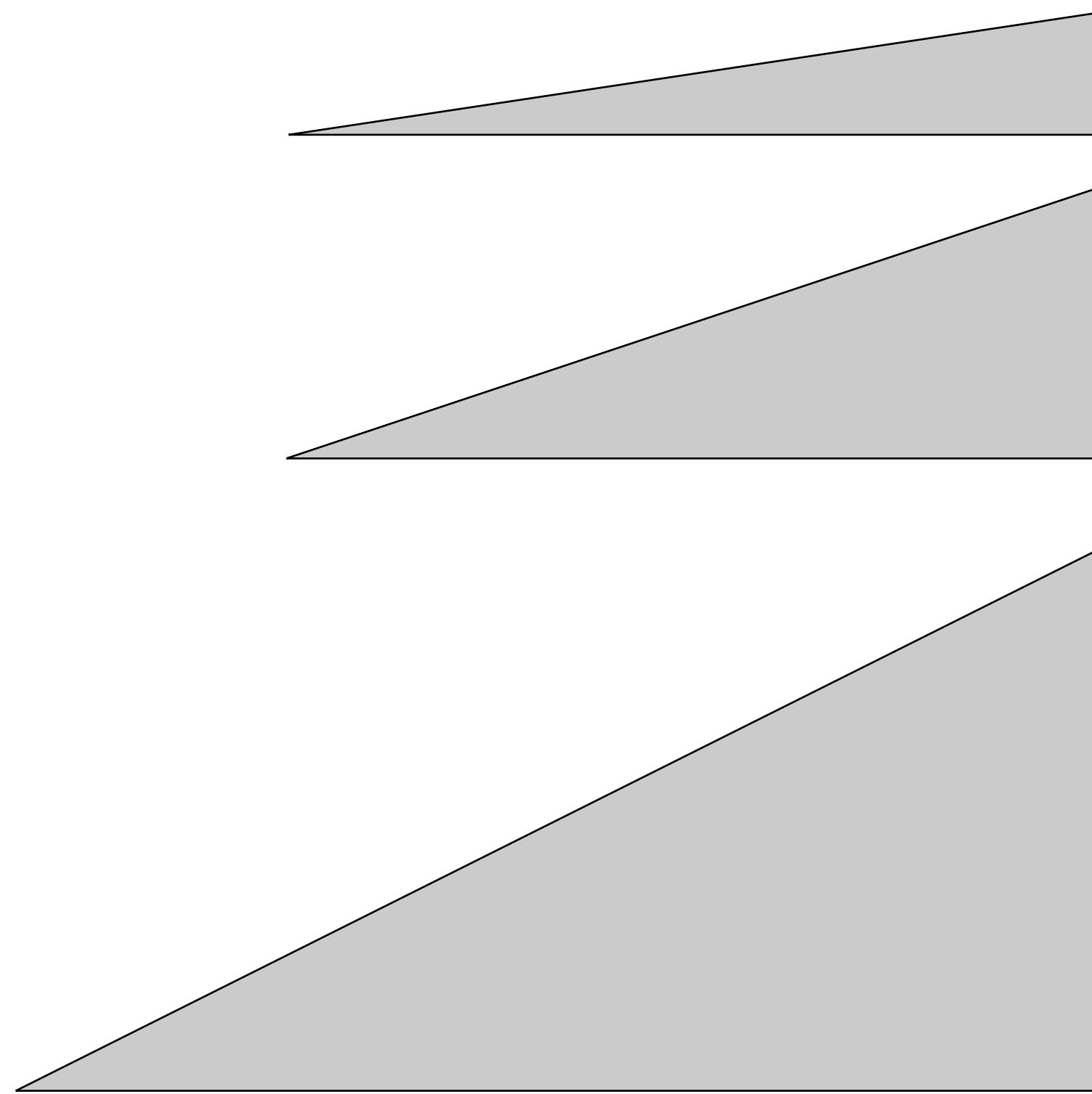


collection
segmentStack

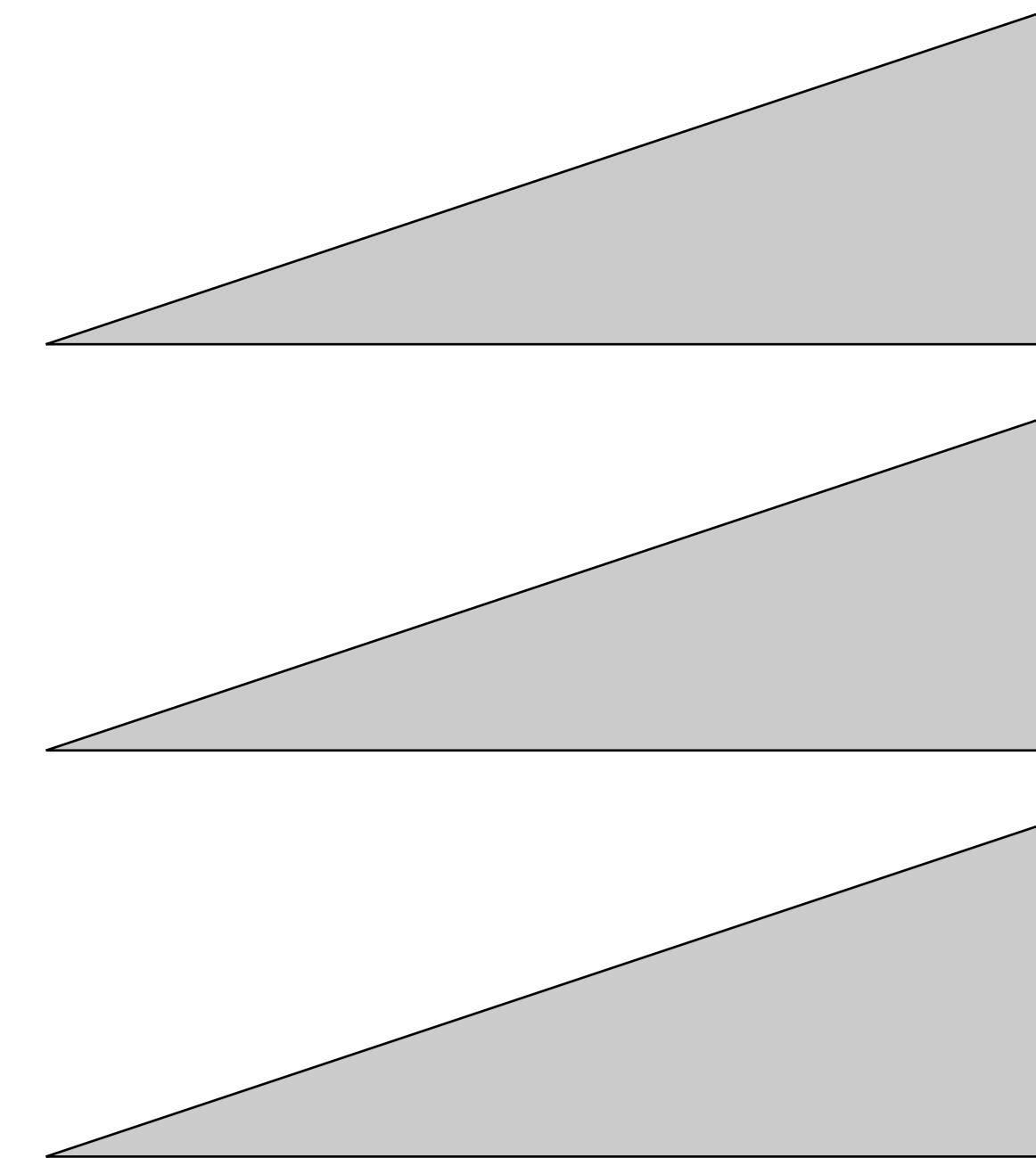


snapshot
segmentStack

Isolated Snapshots



collection
segmentStack



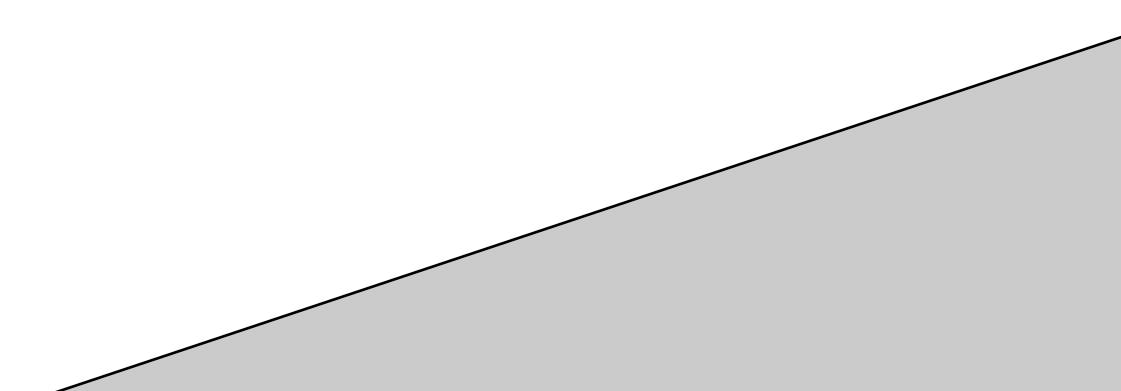
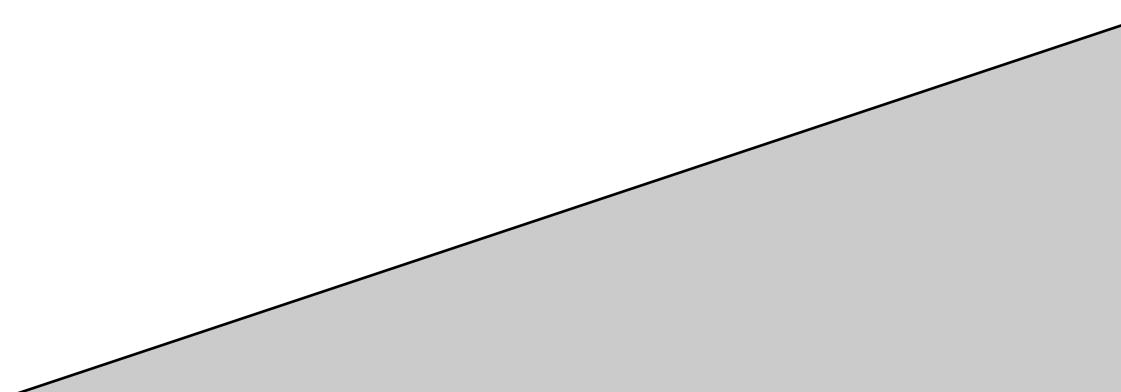
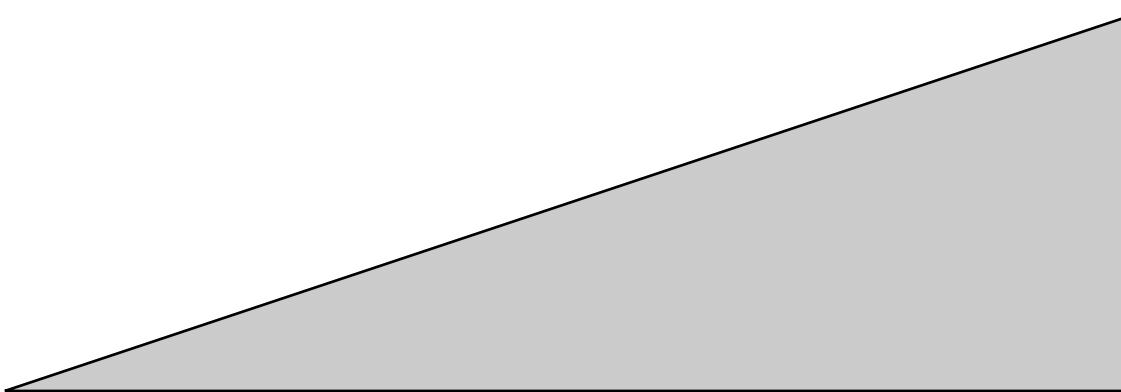
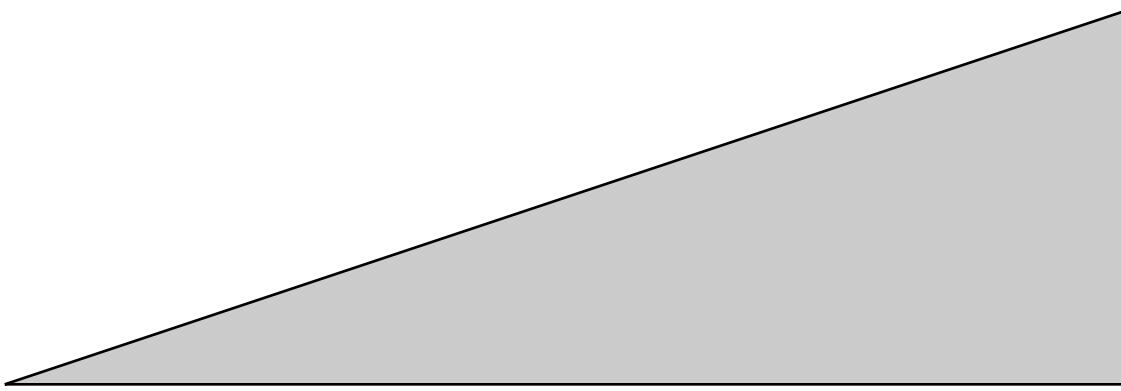
snapshot
segmentStack

Stack Too Tall?

Zoals in Noordwijk het licht van de vuurwerkvuurs langs die vier windrichtingen zweeft, zo streken elk jaar in grootsche vlaggen die vier seisoenen over Groot Rechteren. Max kende die wisseling van de jaargrijsen eigenlijk alleen uit Amsteldam: op een dag in februari of maart is er een onbeschrijflijke voorjaarsgeur als hij's morgens op straat loopt; onbegrijpbaar als de decimalen van π , die stoffige somer waarin de stad vol moeststen stroomde, even plotseling veranderend in de voorjaar, herfst, en dan de bleke winter waarin de sterren van de straten en de muren opeens de ongemaakbare aard van de wereld leken uit te drukken, - maar dat alles eigenlijk in het voorjaar, oppgemerkt alleen in de korte tussenposes waarin hij van het ene interieur naar het andere ging. Was in de stad de natuur tachtig achtereenvolgende, op het kasteel zat hij met Quinten en Sjipje midden in de daverende concertzaal. Lente en herfst kwamen met kolossaal vertoon, de somers waren hoter en droger, de winters koudter en winter. Die onafgebonden verandering, had hij tegen Otto beweerd, was natuurlijk de bron van alle creativiteit: de onvermoeidheid van de natuur tussen die keerkringen hieldde ook toe culturele stilstand, de tropen waren een onafgebonden land, altijd groen zoals de poolgebieden altijd wit waren, van de primitieve breedtegraden met hun viervormigheid waren



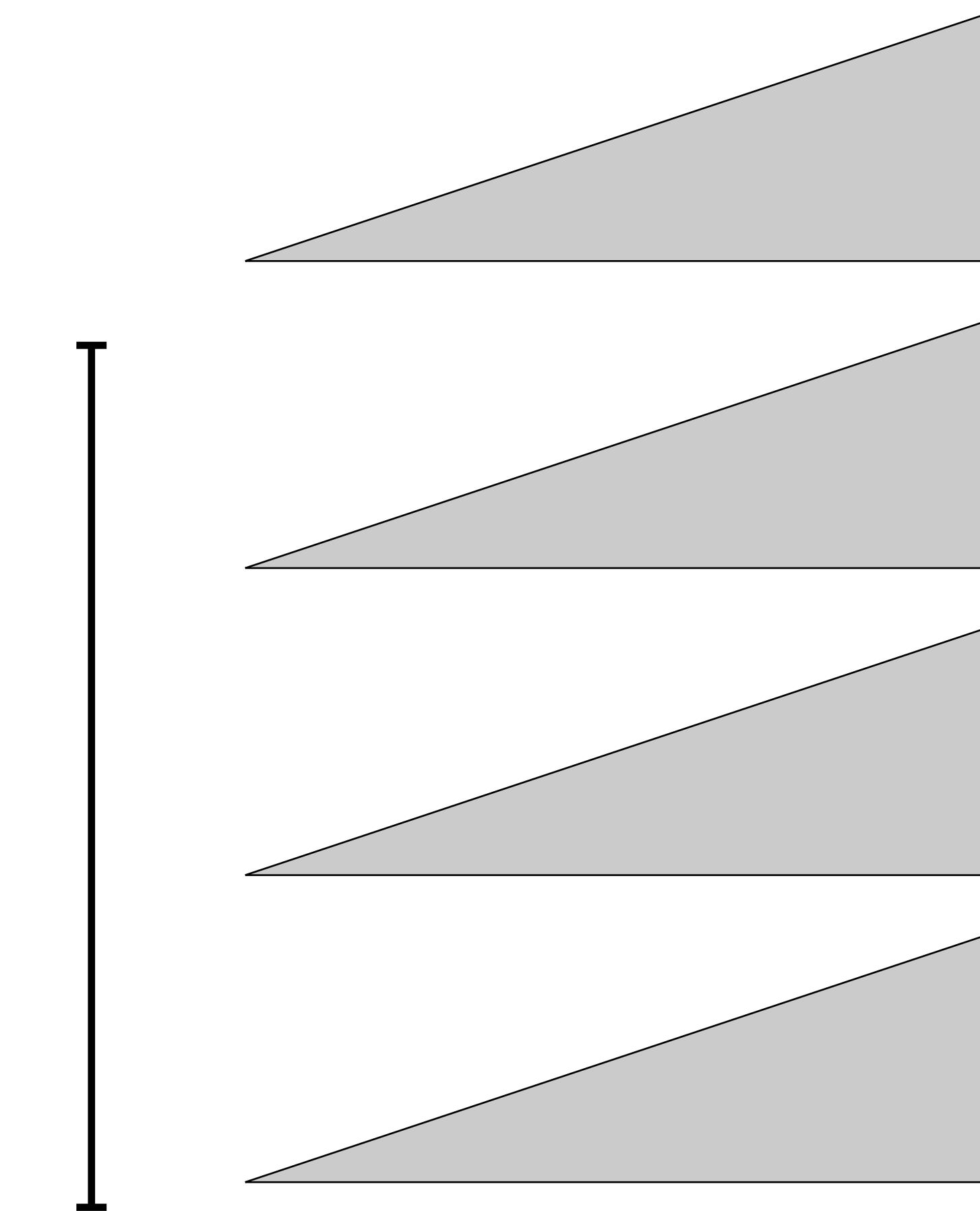
Background Merger



segmentStack

Background Merger

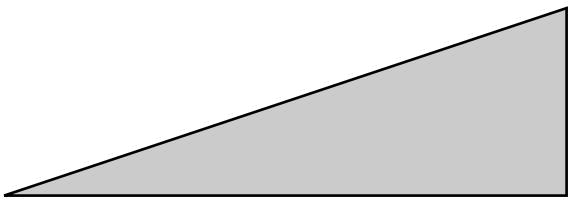
Iterate over Segments



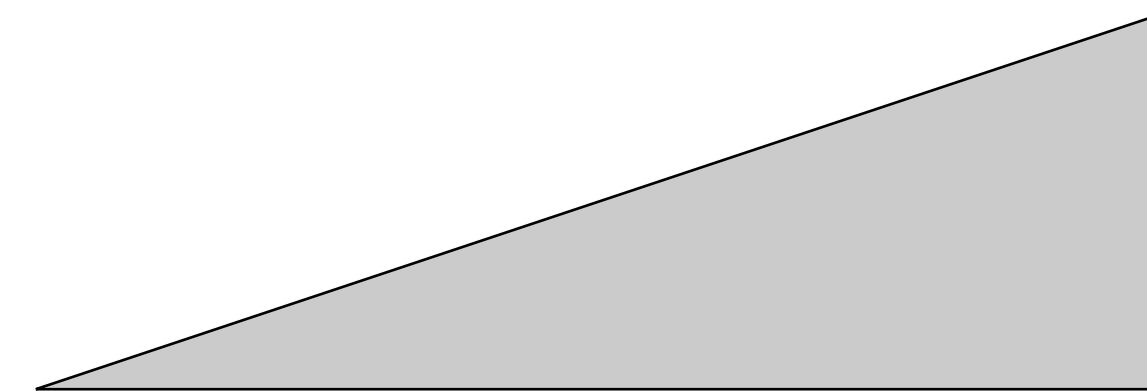
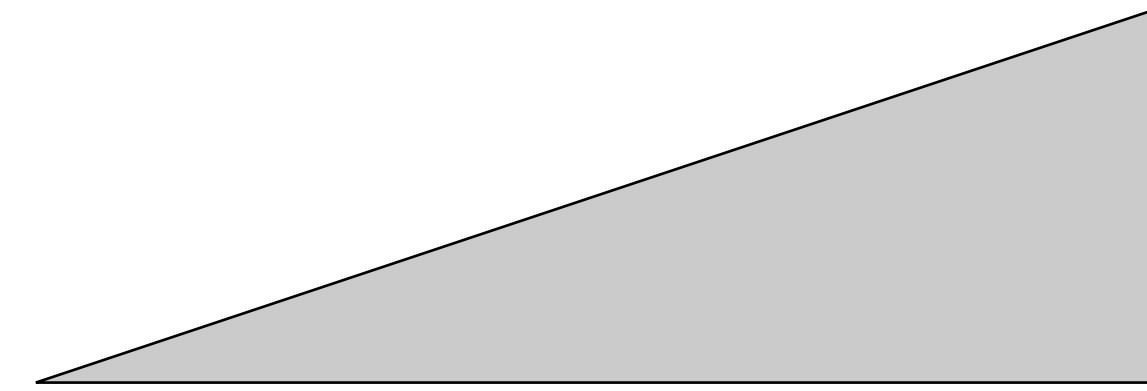
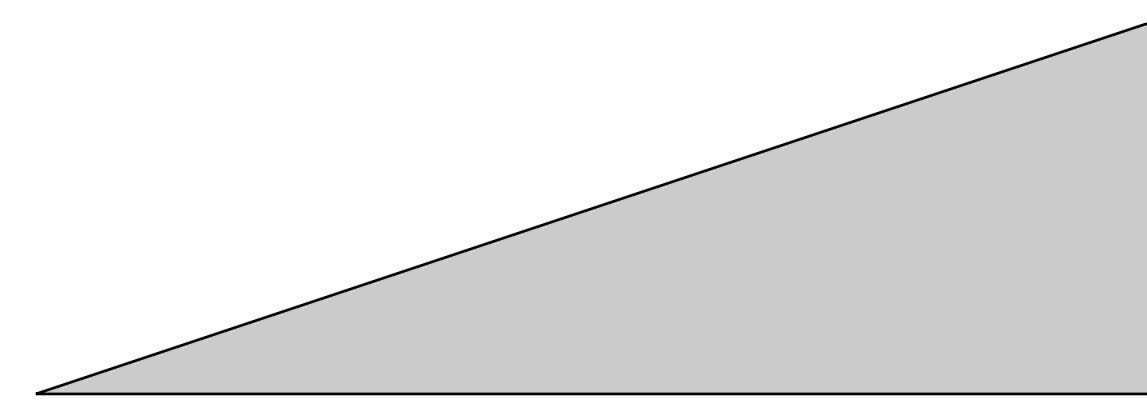
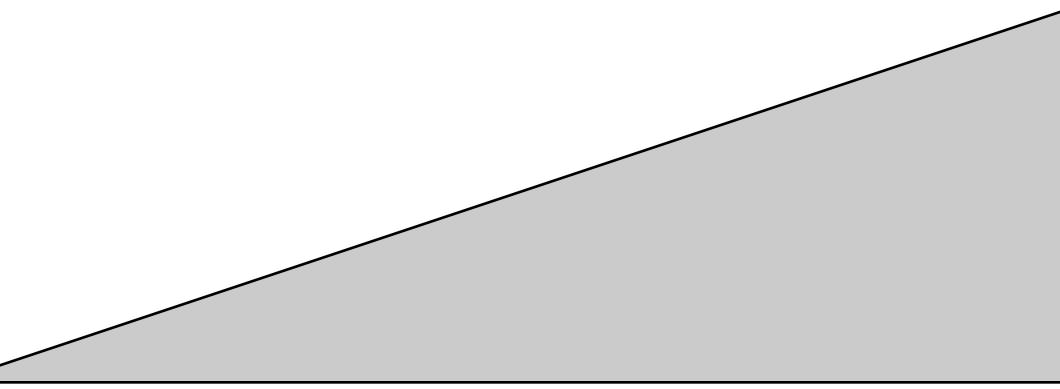
segmentStack

Background Merger

Iterate over Segments



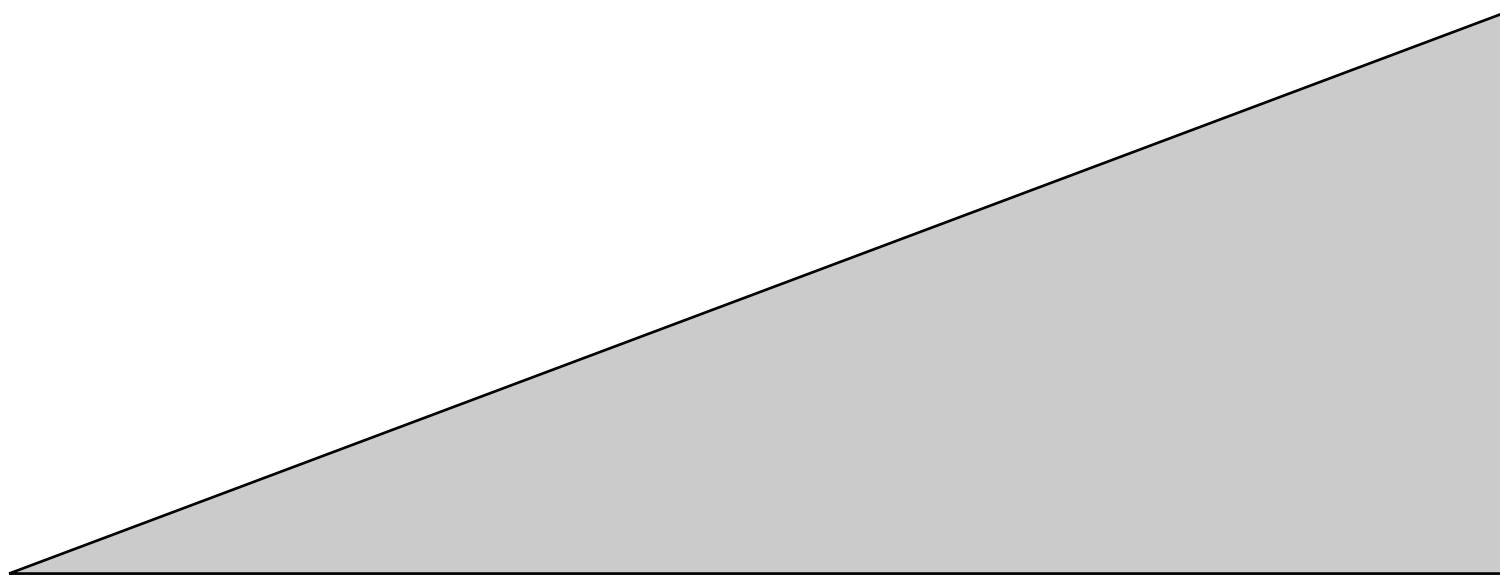
Build up New Segment



segmentStack

Background Merger

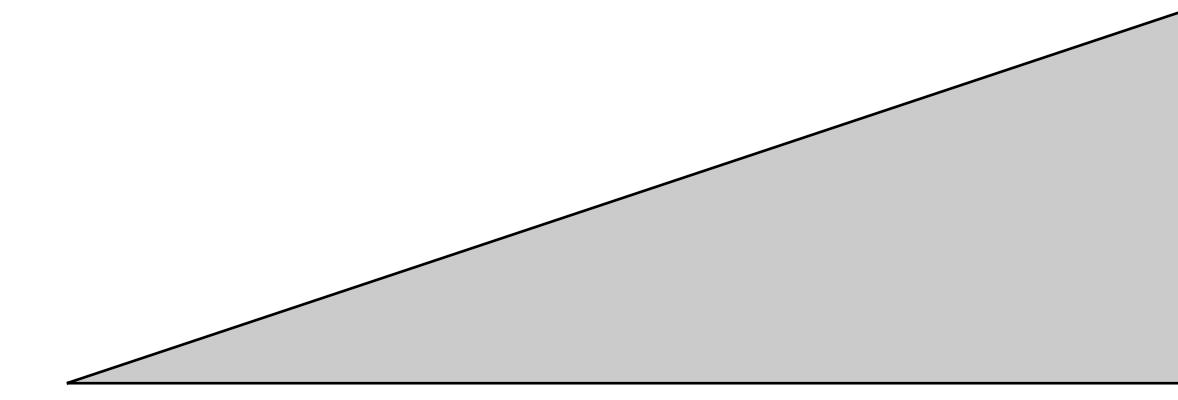
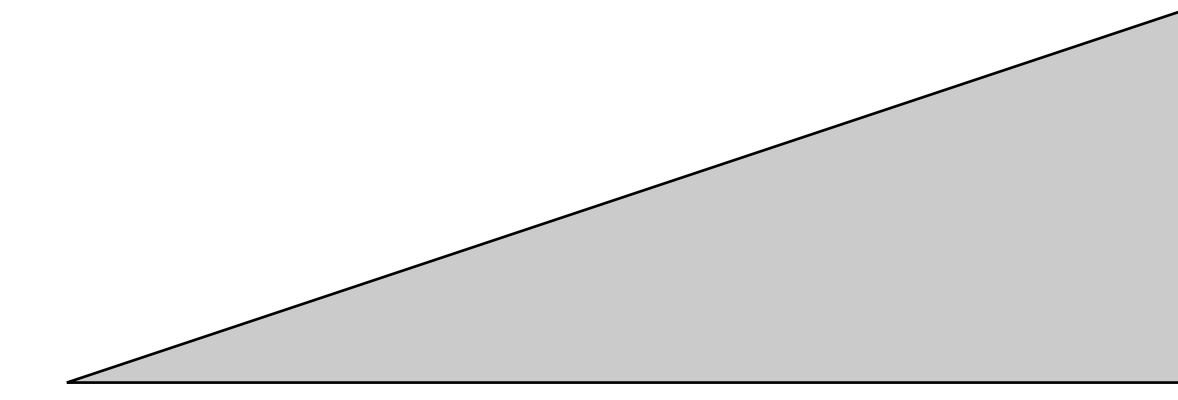
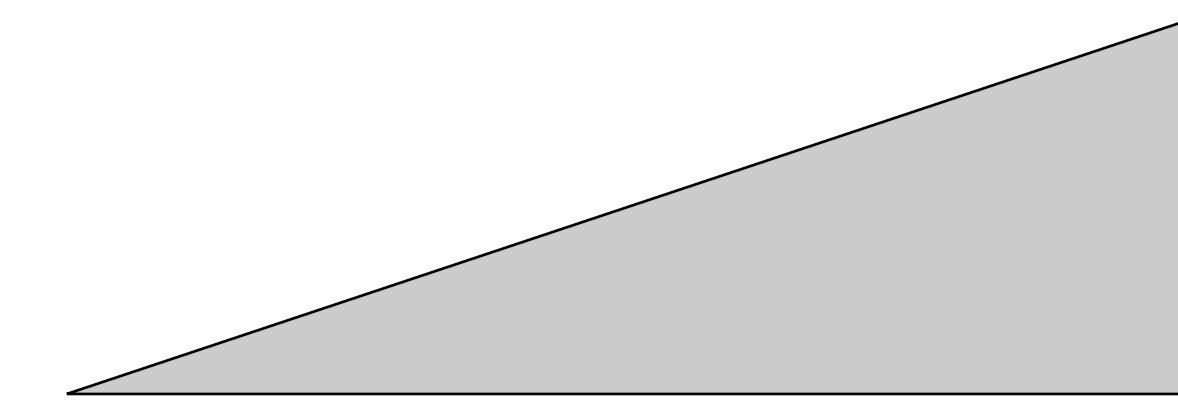
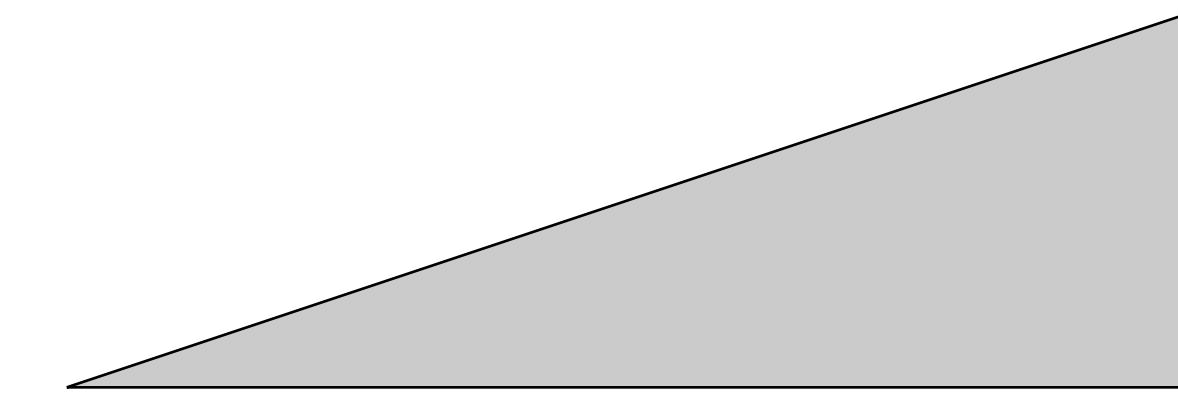
Iterate over Segments



Build up New Segment

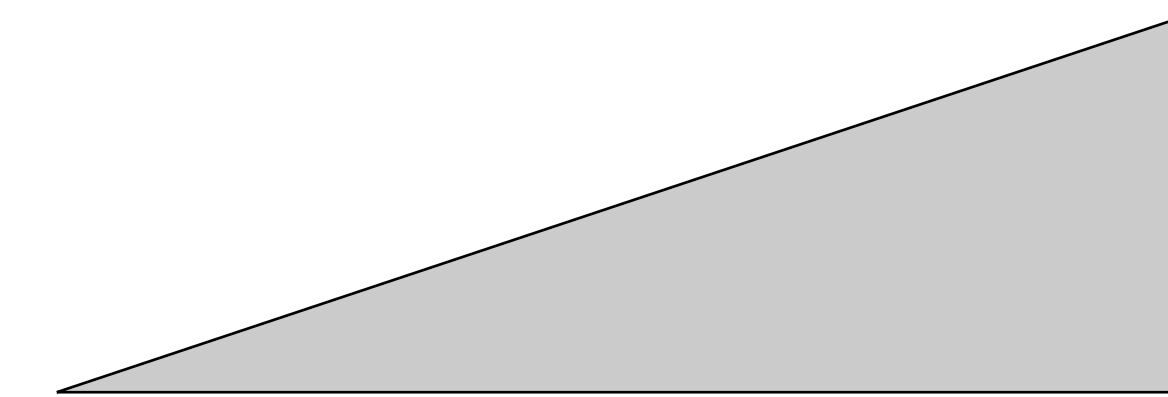
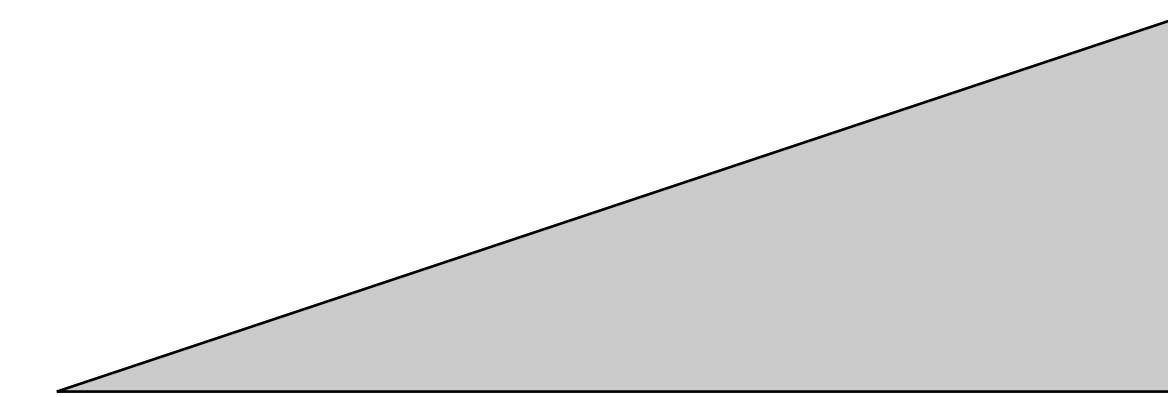
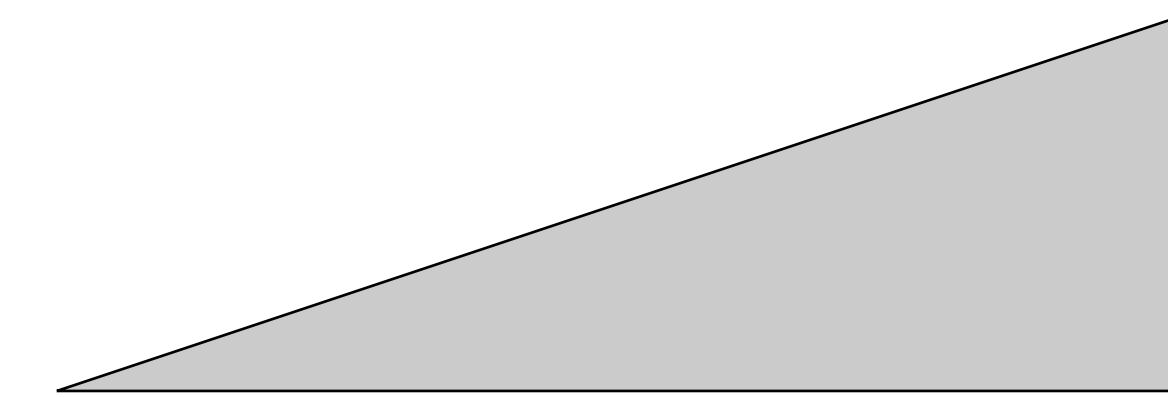
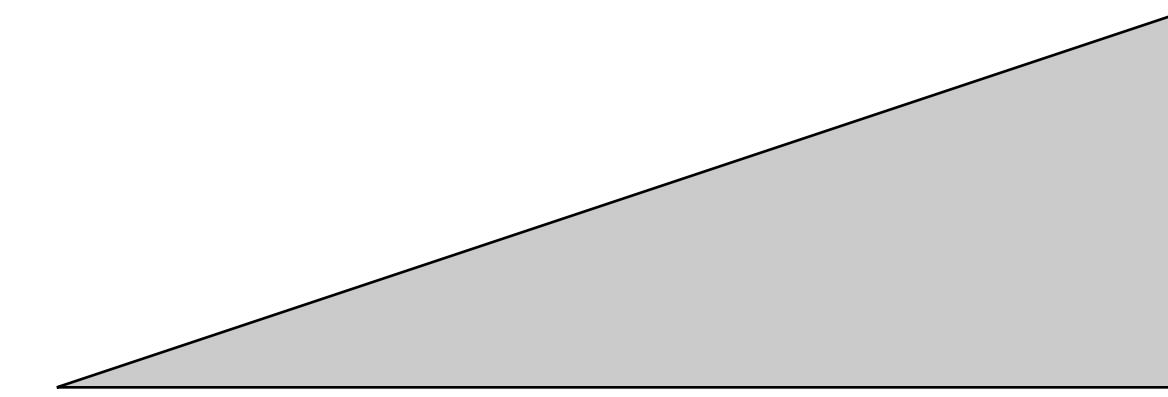
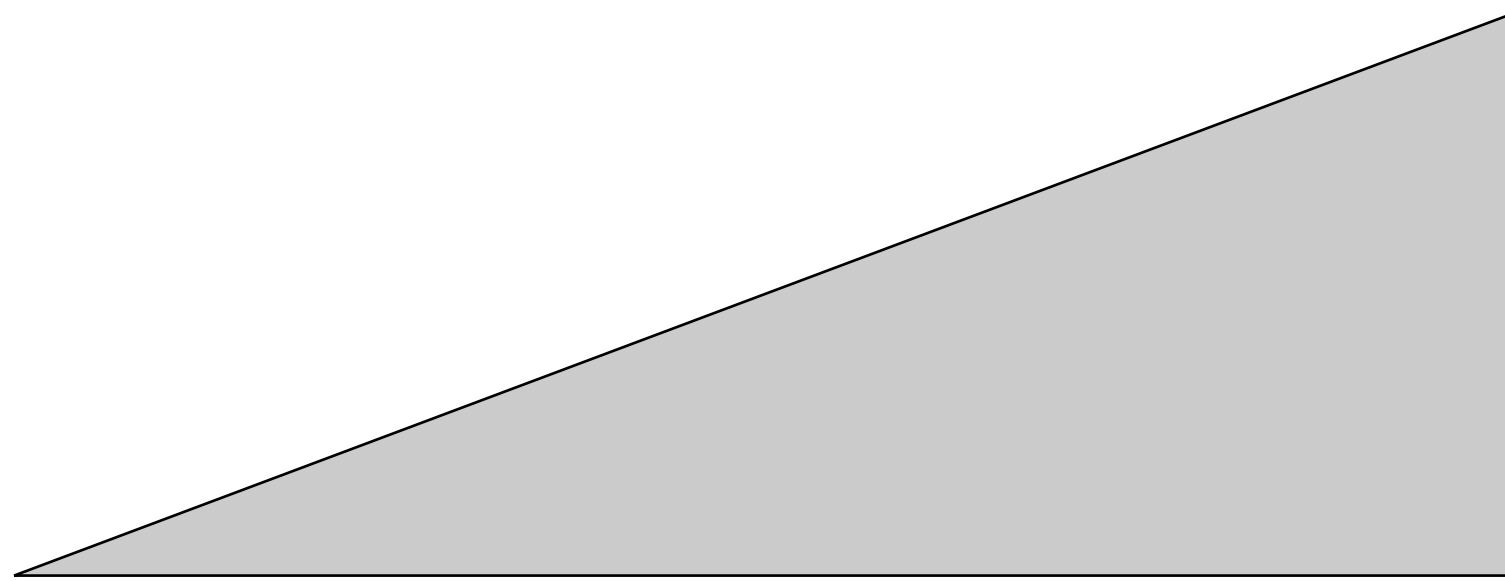


segmentStack



Background Merger

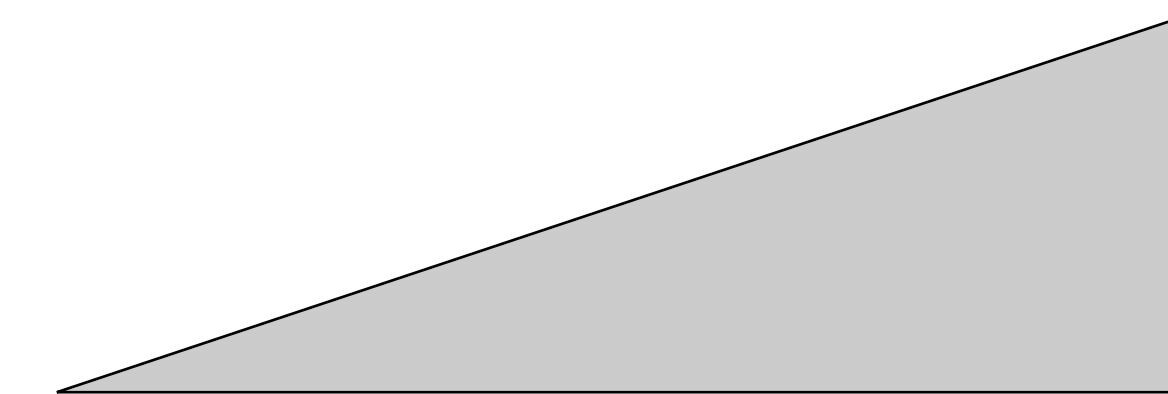
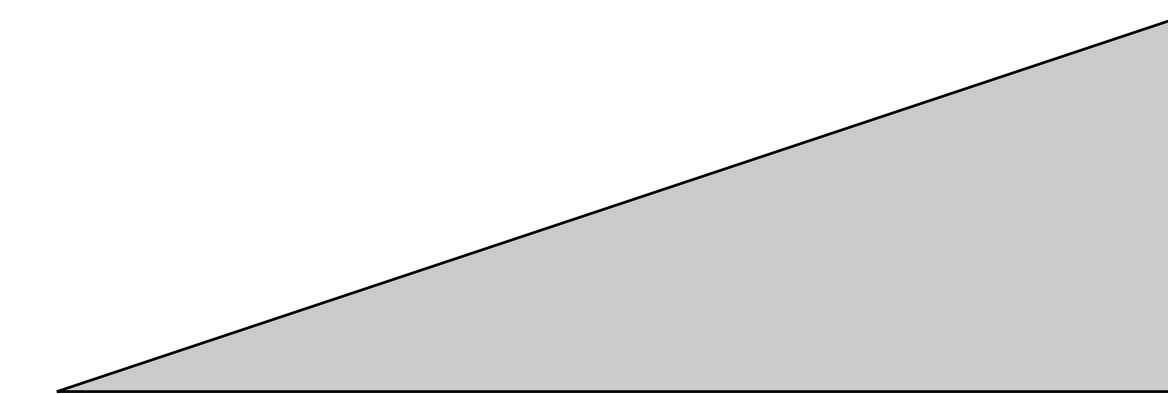
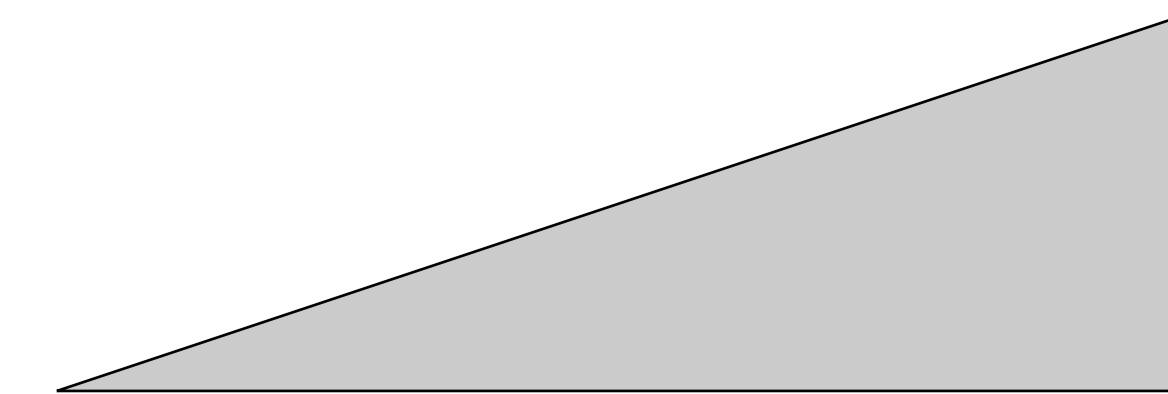
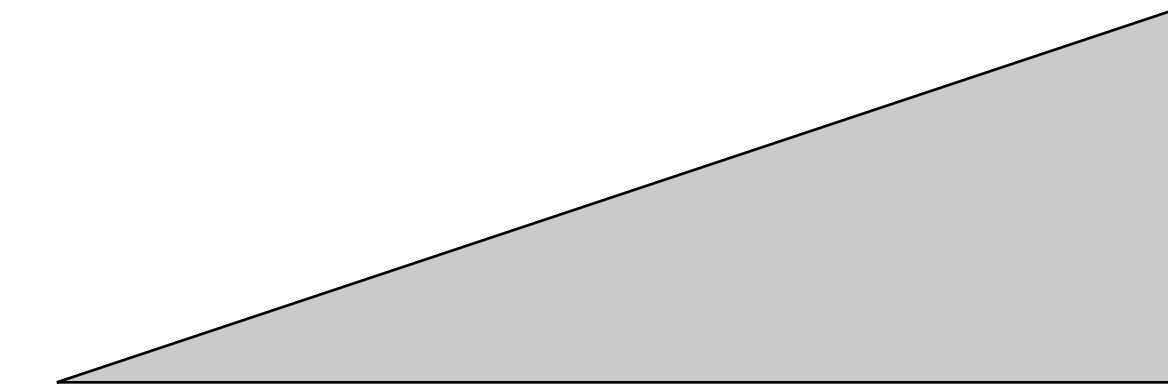
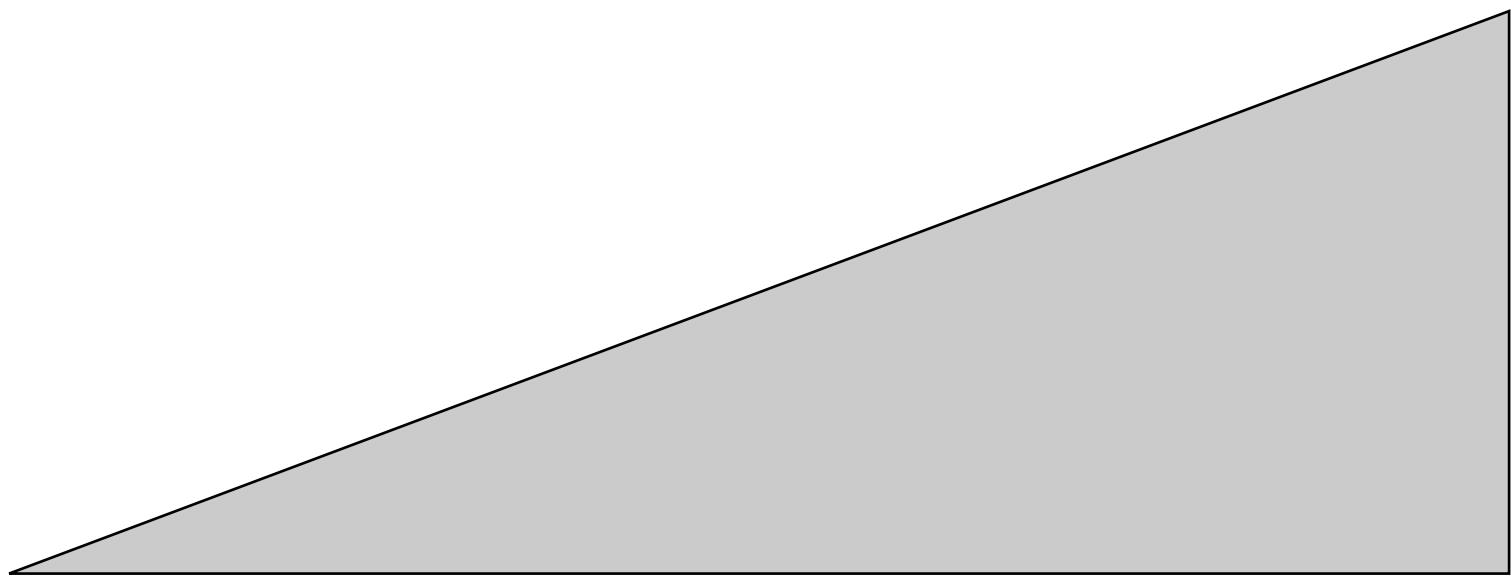
Swap in New Segment



segmentStack

Background Merger

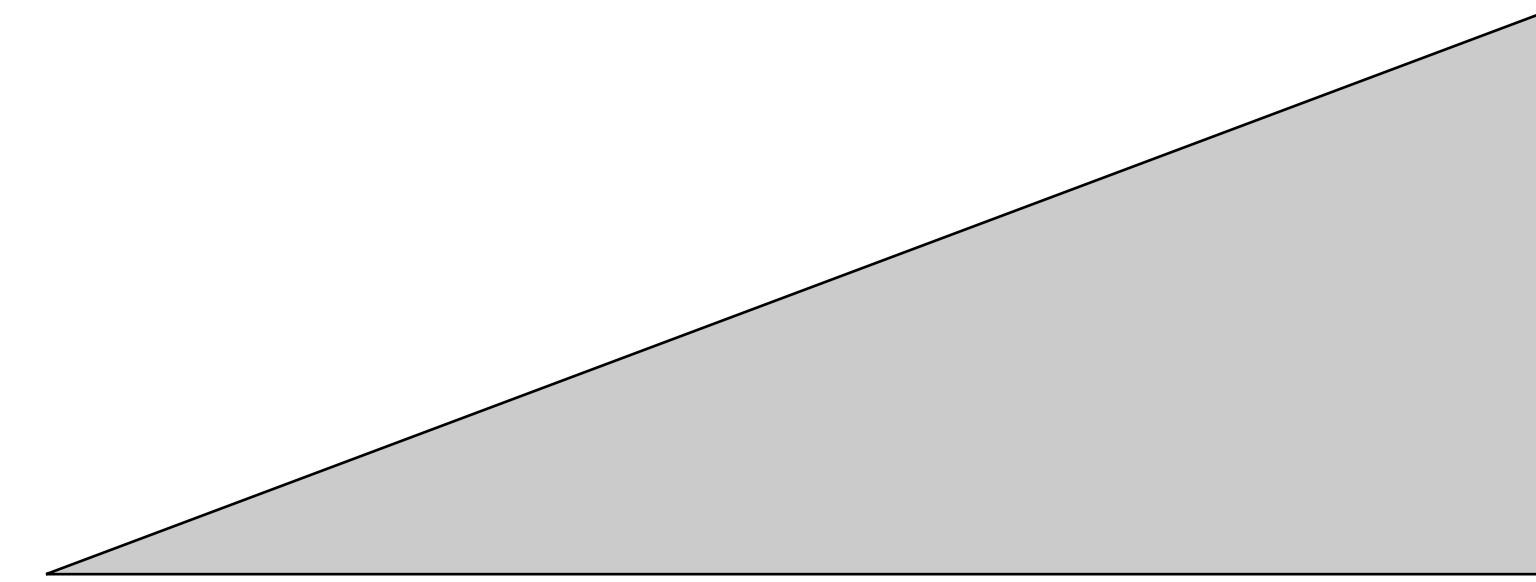
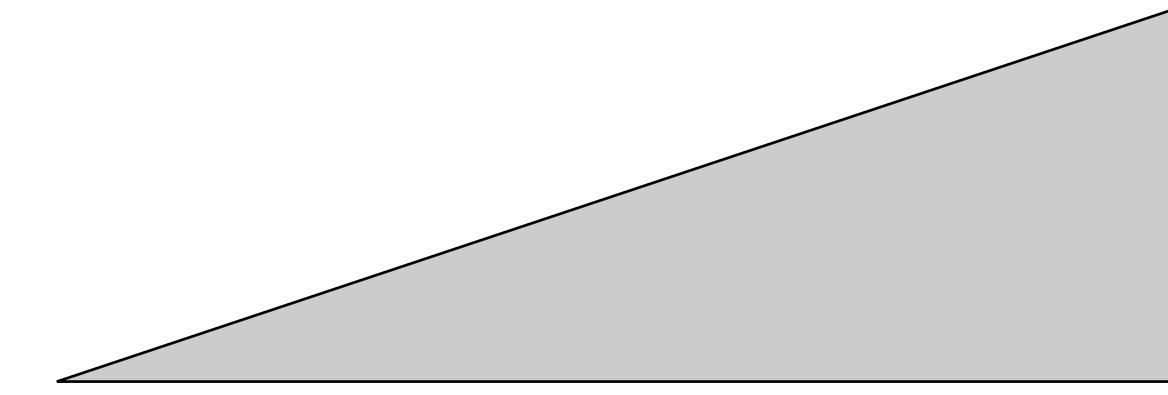
Swap in New Segment



segmentStack

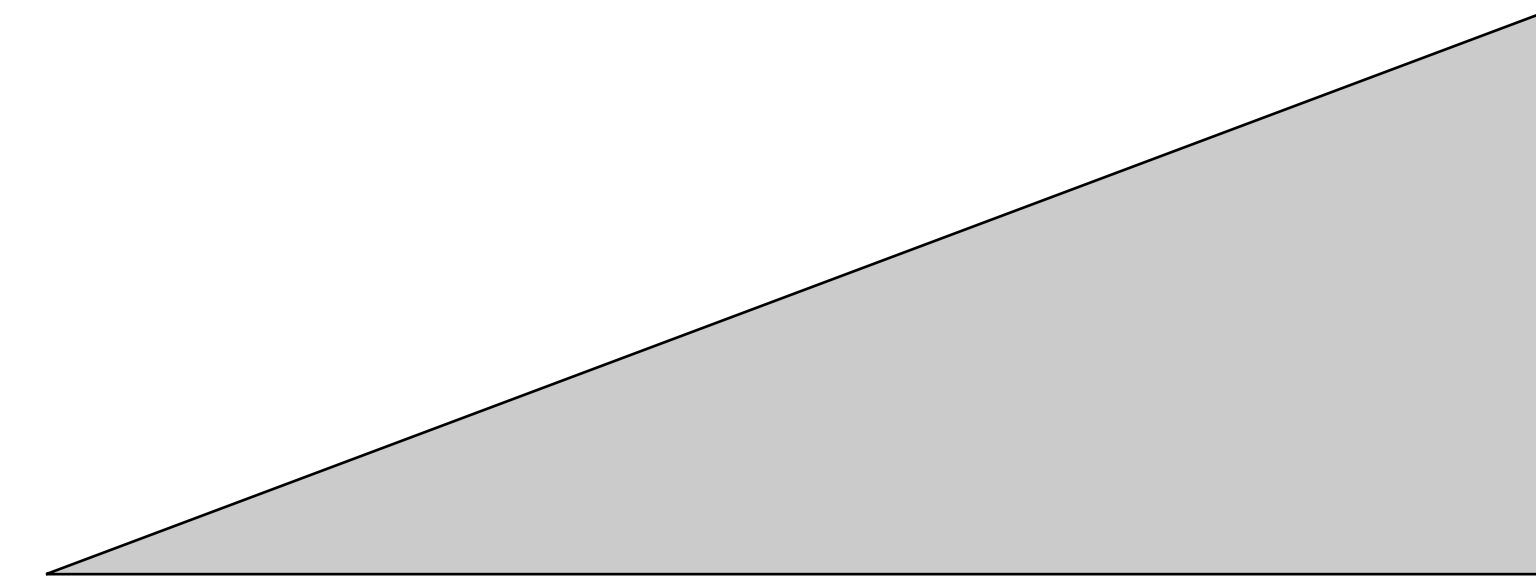
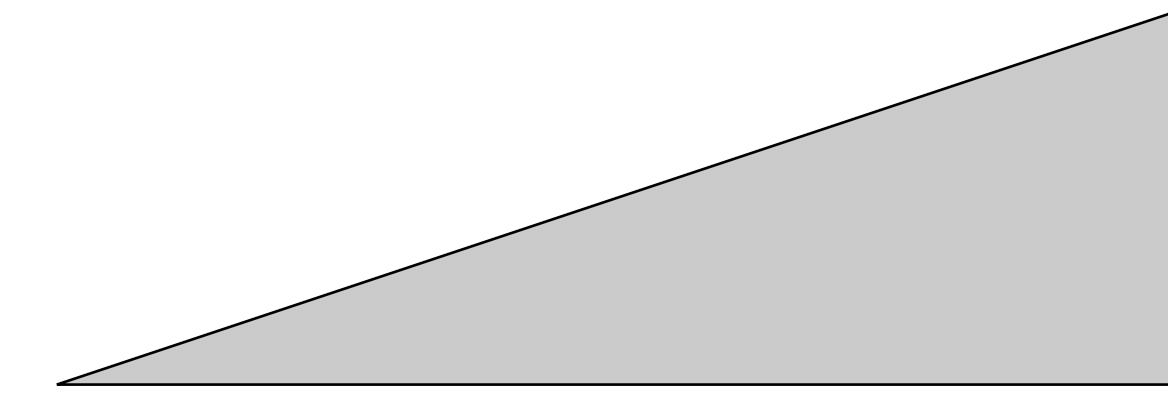
Background Merger

Swap in New Segment



segmentStack

Background Merger



segmentStack



It Works! But...

Zoals in Noordwijk het licht van de vuurzuren langs de vier windrichtingen overvalt, zo streken elk jaar in grootsche vlaggen die vier seisoenen over Groot Rechtersen. Max kende die wisseling van de jaargrijsen eigenlijk alleen uit Amsterdam: op een dag in februari of maart is er een onbeschrijflijke voorjaarsgeur als hij's morgens op straat loopt, onbegrijpbaar als de decimalen van π , die stoffige somer waarin de stad vol moeststenen stroomde, even plotseling veranderend in de voorjaar, herfst, en dan de bleke winter waarin de stenen van de straten en de muren opeens de ongemaakbare aard van de wereld leken uit te drukken, - maar dat alles eigenlijk in het voorjaar, oppgemerkt alleen in de korte tussenposes waarin hij van het ene interieur naar het andere ging. Was in de stad de natuur tachtig achtereenvolgende, op het kasteel zat hij met Quinten en Sjipje midden in de daverende concertzaal. Lente en herfst kwamen met kolossaal vertoon, de somers waren hoter en droger, de winters koudter en winter. Die onafgebonden verandering, had hij tegen Otto beweerd, was natuurlijk de bron van alle creativiteit; de onvermoeidheid van de natuur tussen dit keerkringen hieldde ook toe culturele stilstand, de mensen waren een onafgebonden mensheid, altijd groen zoals de poolghieden altijd wit waren, van de primitieve breedtegraden met hun viervormigheid waren

Append-Only File Format



File

Append-Only File Format



Append-Only File Format



Append-Only File Format



Append-Only File Format



Append-Only File Format



Simple

Append-Only File Format



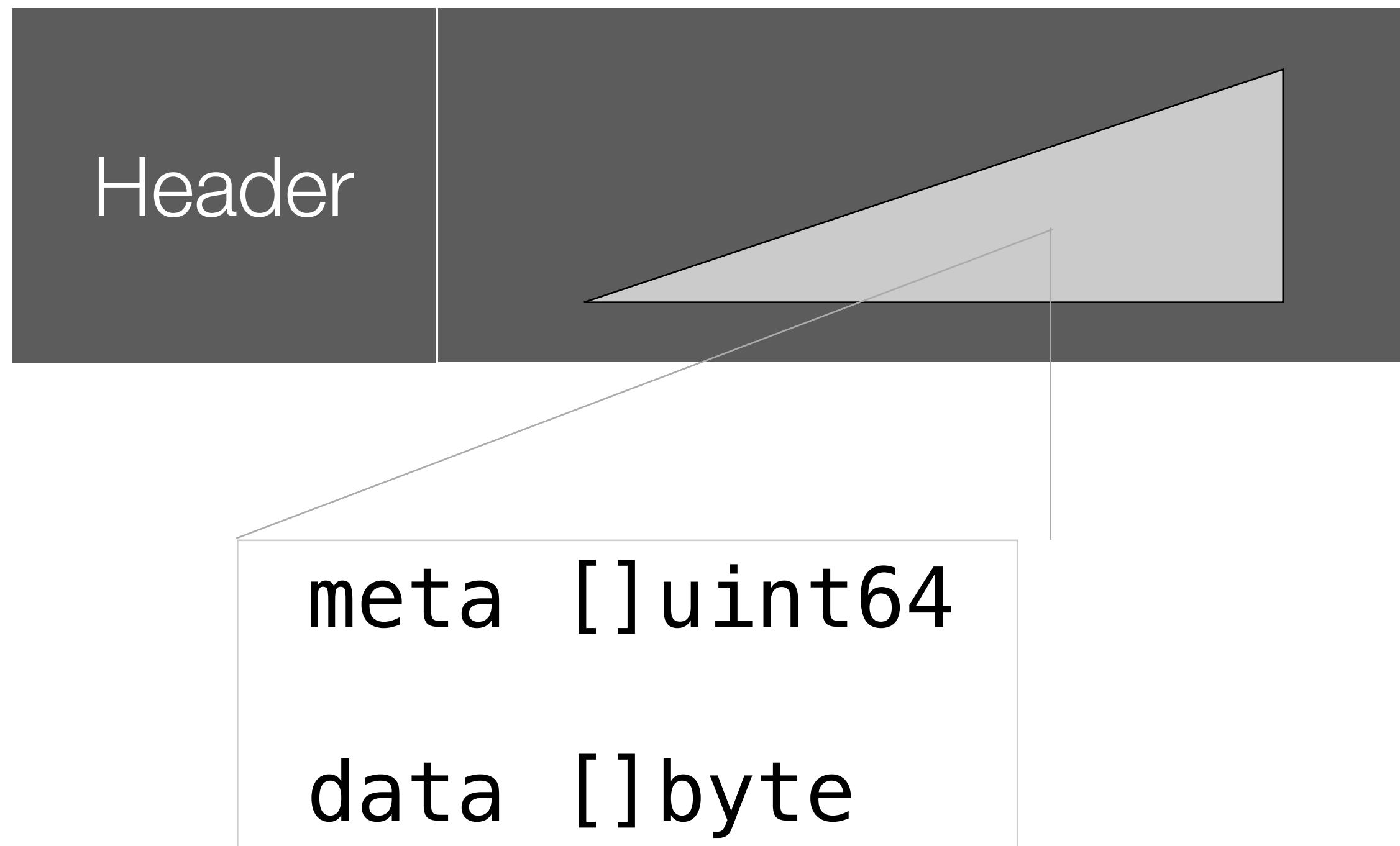
Simple

Safe

Header



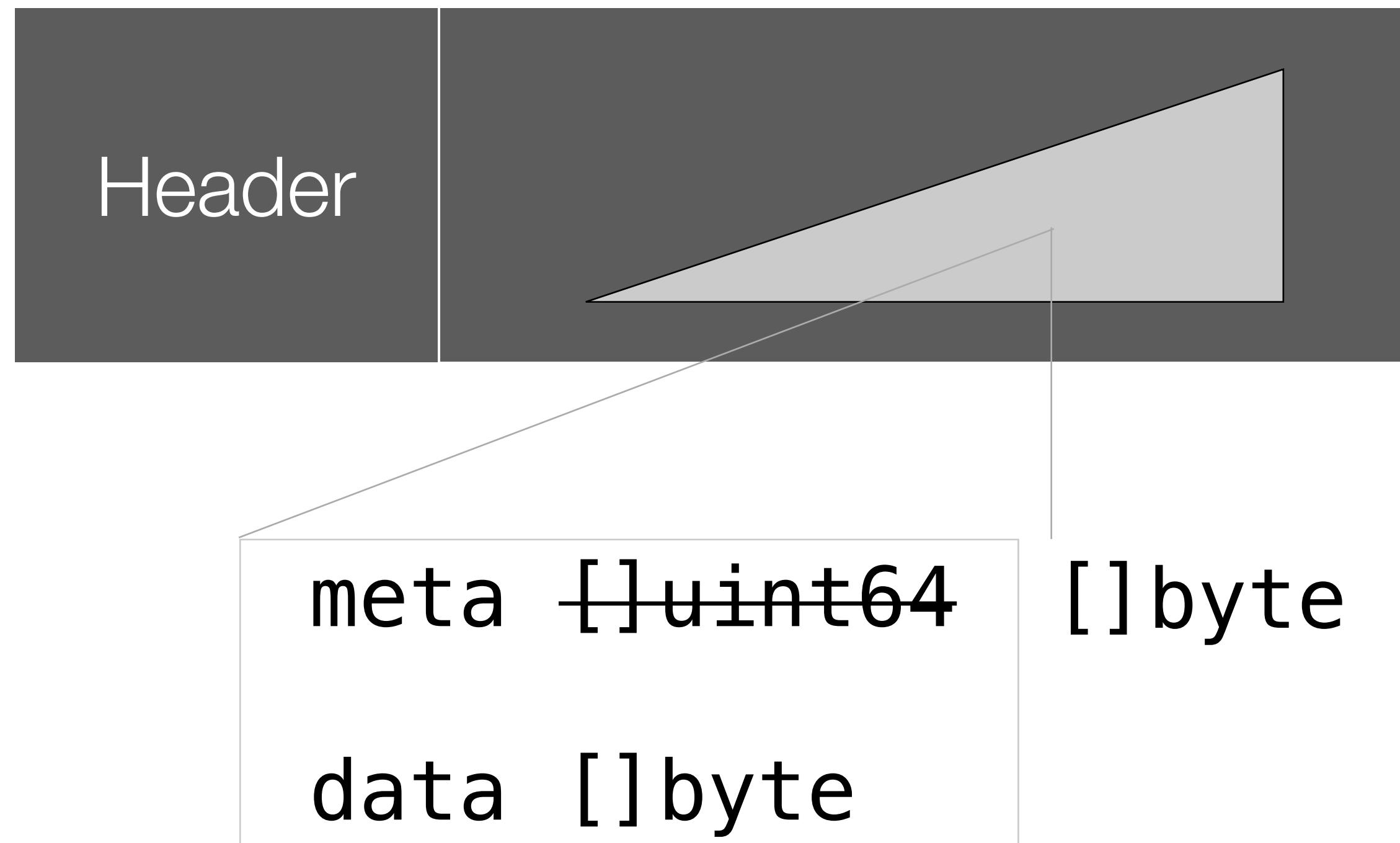
Persist each Segment in the Segment Stack



```
func (f *File) Write(b []byte) (n int, err error)
```

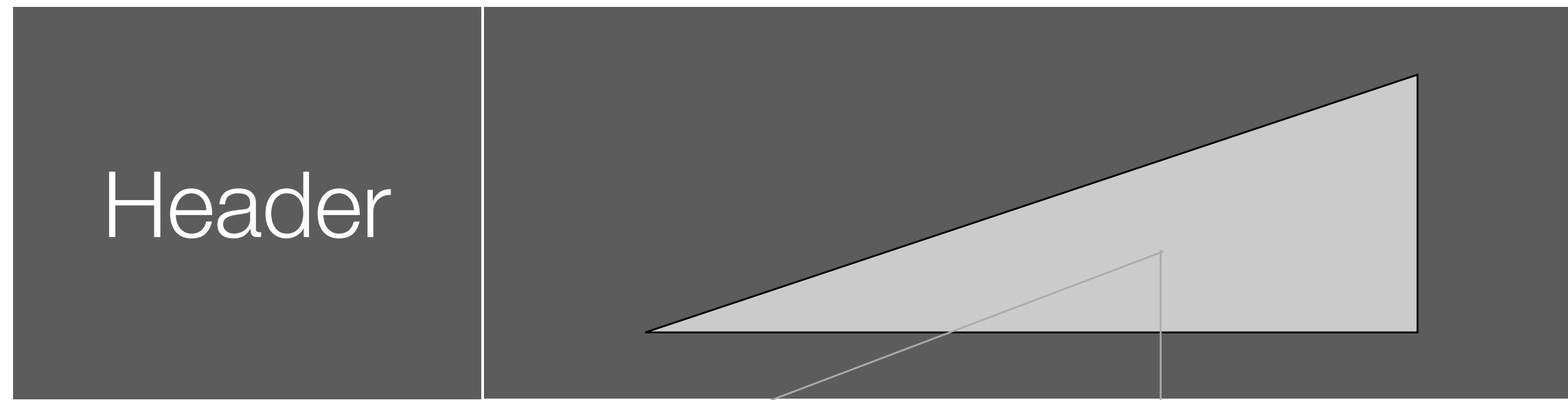
```
func Uint64SliceToByteSlice(in []uint64) ([]byte, error) {  
  
    inHeader := (*reflect.SliceHeader)(unsafe.Pointer(&in))  
  
    var out []byte  
    outHeader := (*reflect.SliceHeader)(unsafe.Pointer(&out))  
    outHeader.Data = inHeader.Data  
    outHeader.Len = inHeader.Len * 8  
    outHeader.Cap = inHeader.Cap * 8  
  
    return out, nil  
}
```

Persist the Segment Stack



```
func (f *File) Write(b []byte) (n int, err error)
```

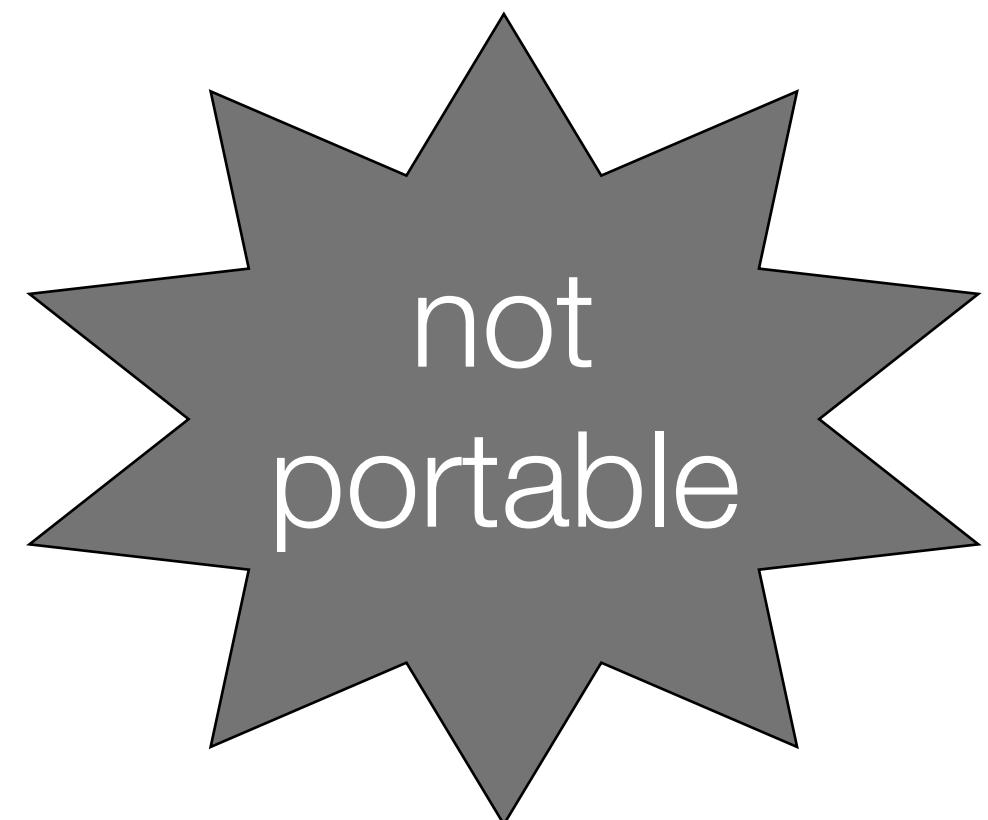
Persist the Segment Stack



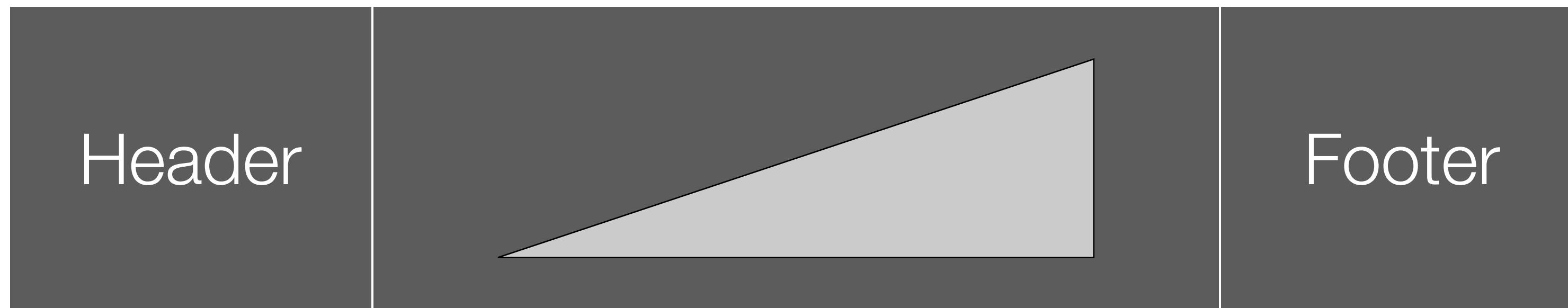
meta []uint64 []byte

data []byte

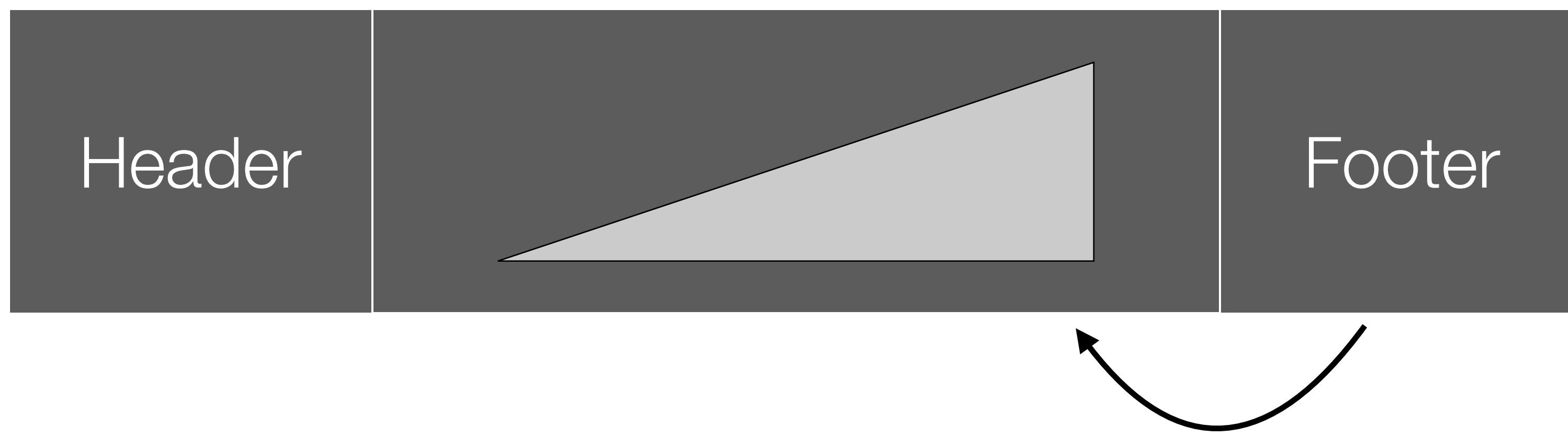
func (f *File) Write(b []byte) (n int, err error)



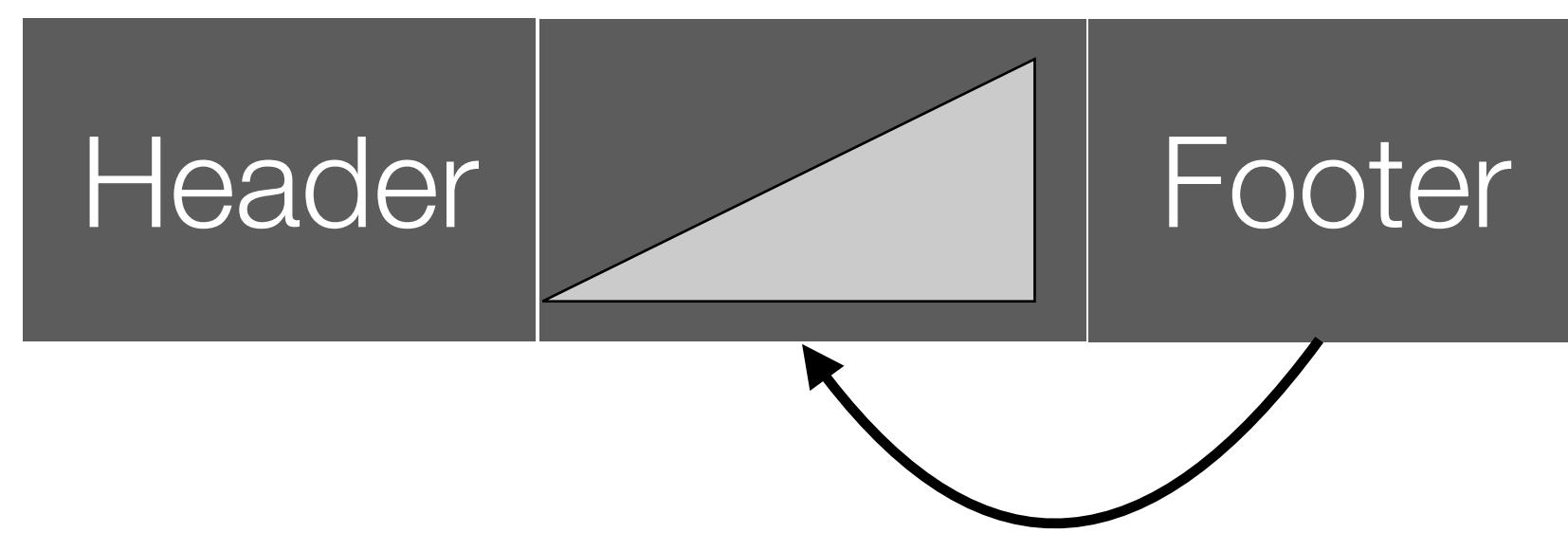
Append Footer



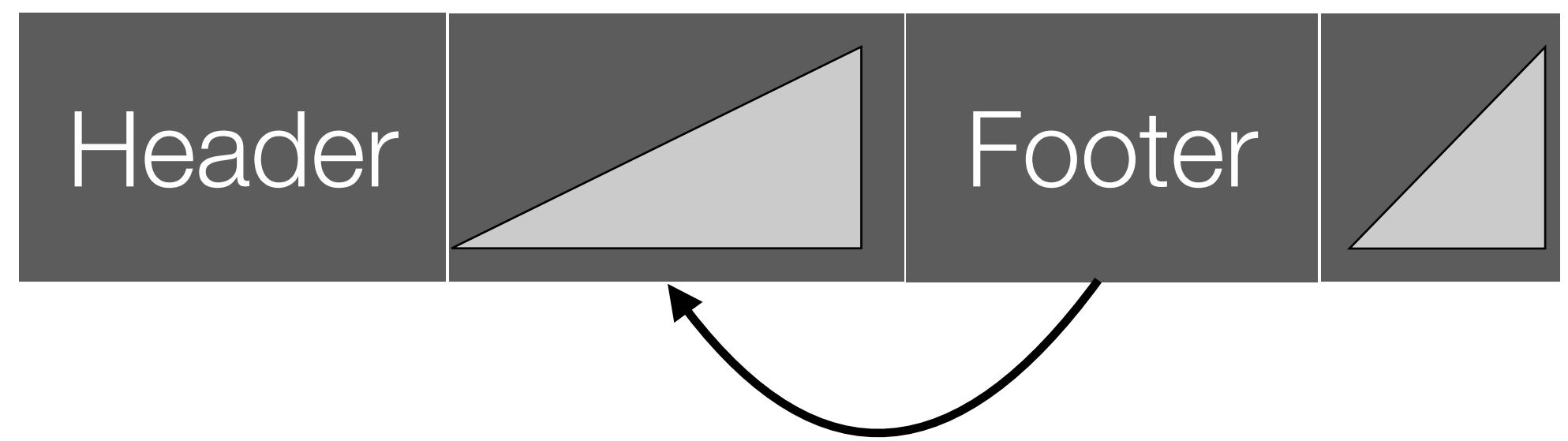
Append Footer



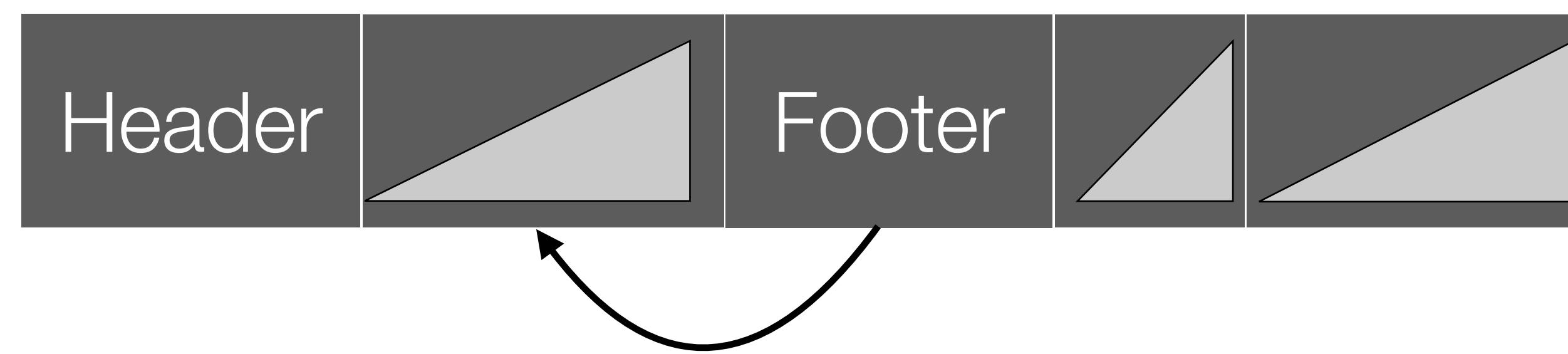
Footer Pointers to Older Segments



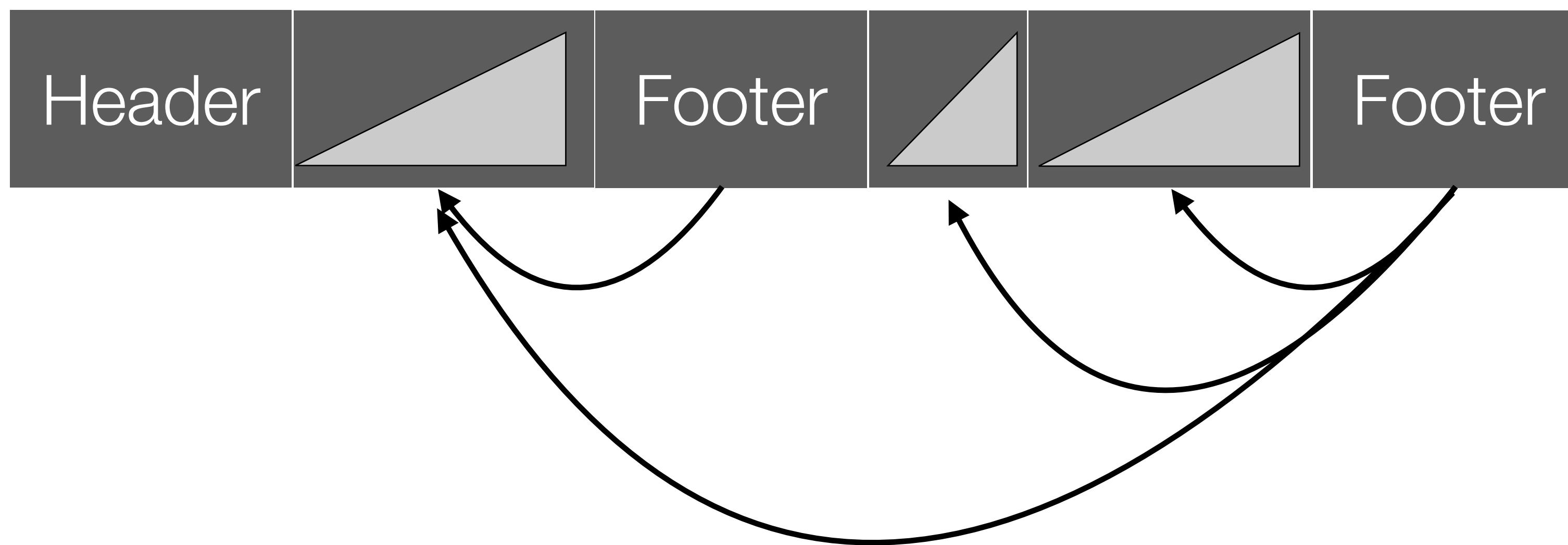
Footer Pointers to Older Segments



Footer Pointers to Older Segments



Footer Pointers to Older Segments

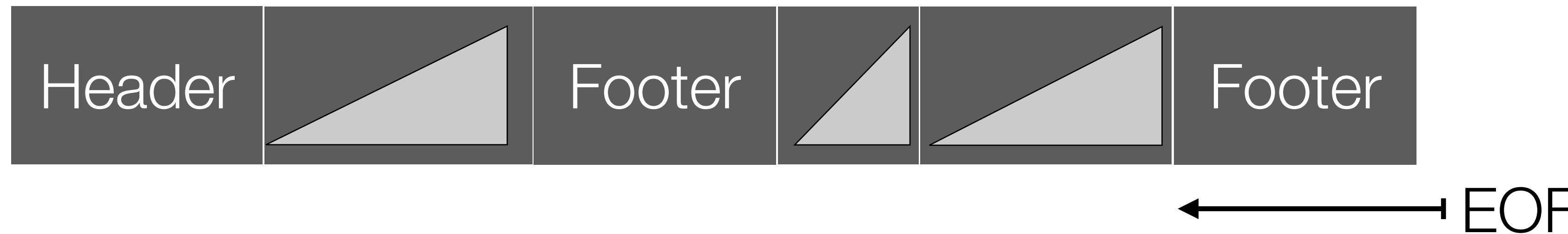


Opening File



Seek Backwards
Find Start of Valid Footer

Opening File



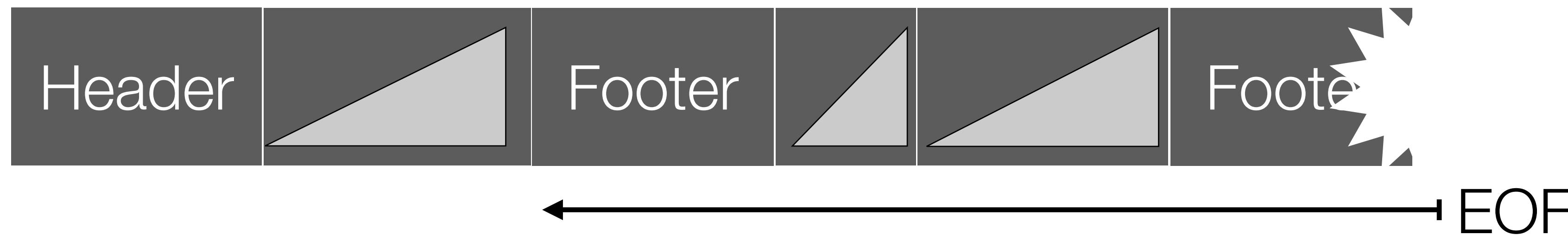
Seek Backwards
Find Start of Valid Footer

Opening File



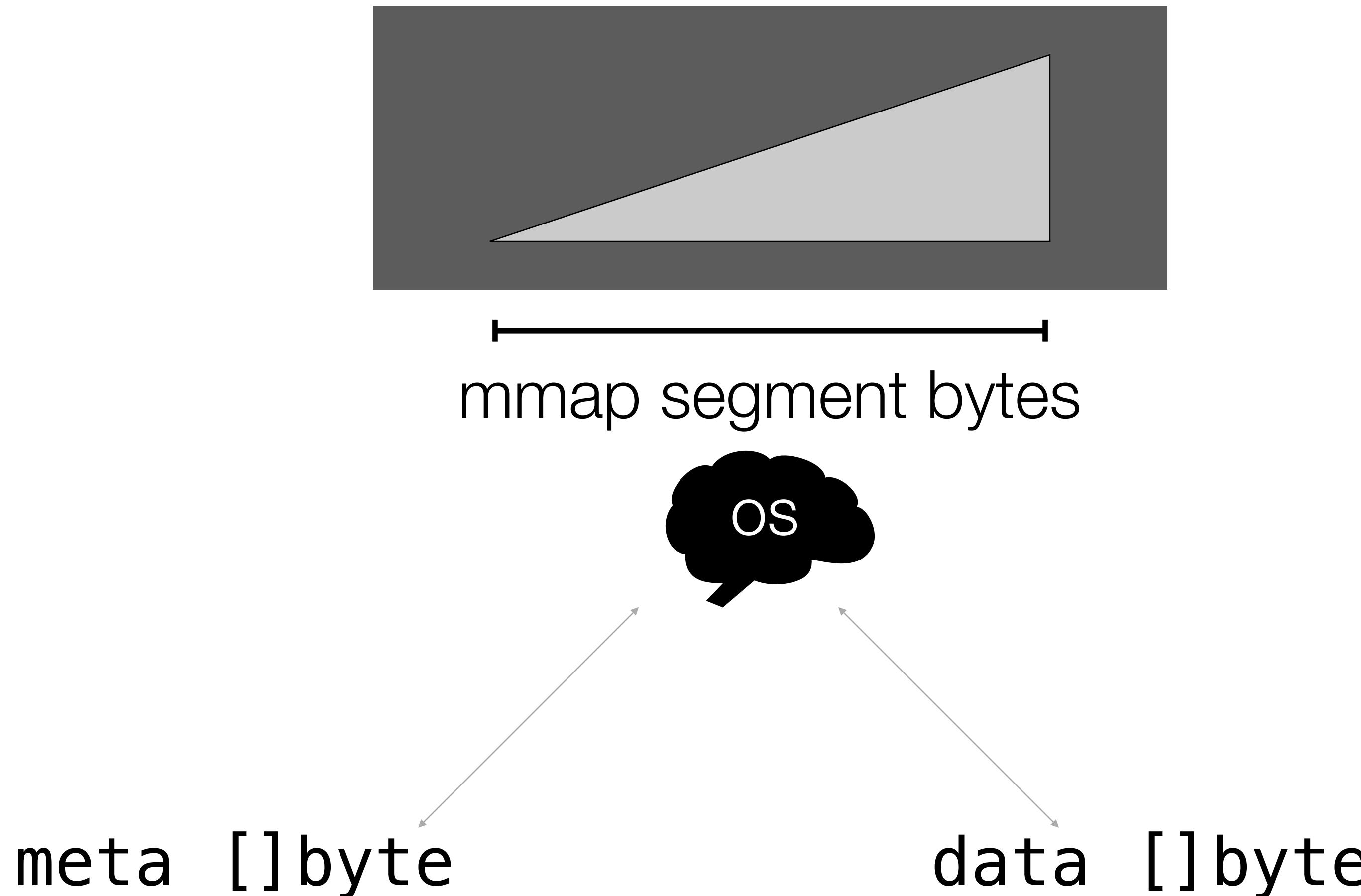
Seek Backwards
Find Start of Valid Footer

Opening File

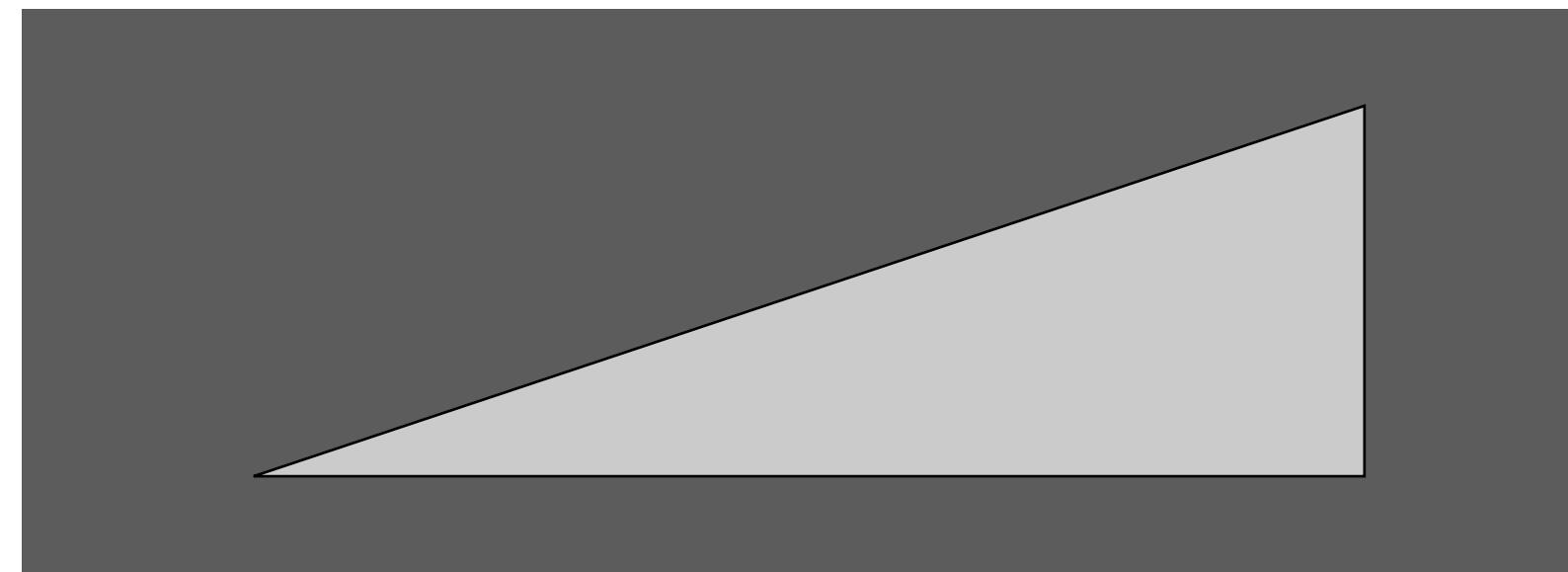


Seek Backwards
Find Start of Valid Footer

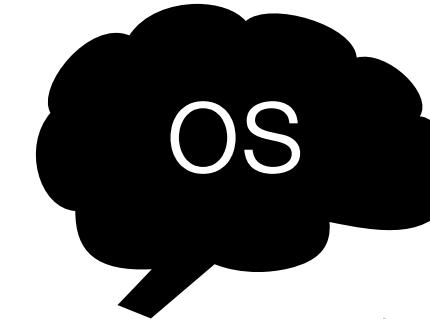
Reading From File



Reading From File



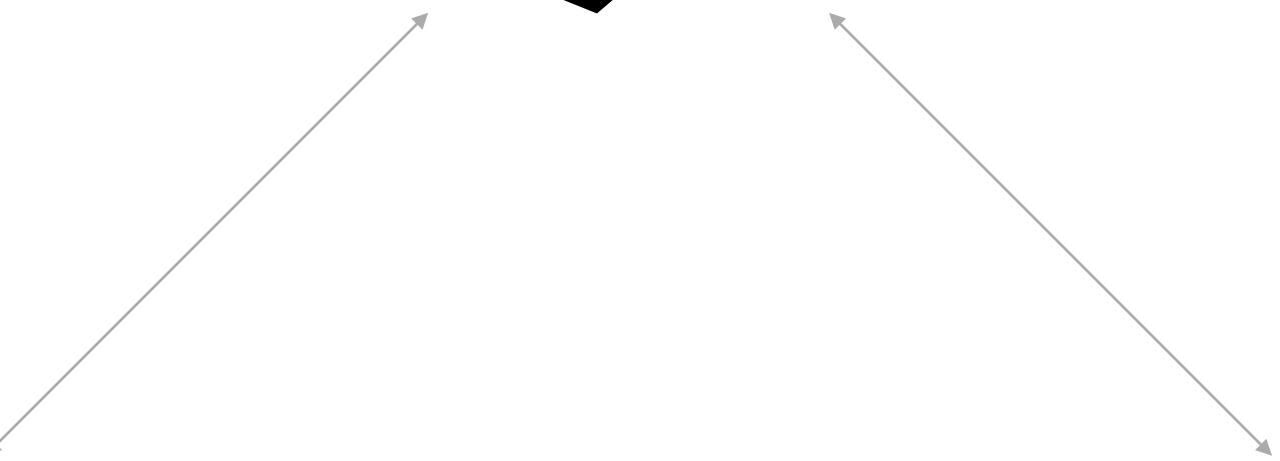
mmap segment bytes



meta []byte

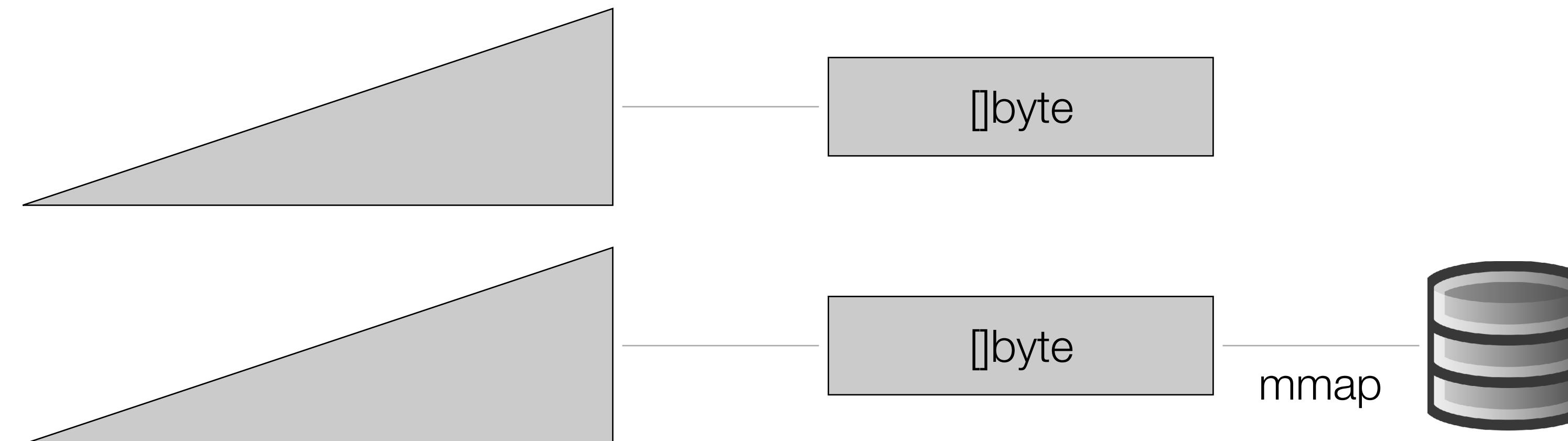
[]uint64

data []byte



In-memory and Disk Segments

Get/Iterator implementations the same, no new code



collection
segmentStack

Compaction



Zuin te Noordwijk het licht van de vissersvuren langs de vier windrichtingen overkiet, zo streken elk jaar in grootsche vlagte de vier seisoenen over Groot Rechteren. Max kende die wisseling van de jaargrijsen eigenlijk alleen uit Amstendam: op een dag in februari of maart is er een onbeschrijflijke voorjaarsgeur als hij's morgens op straat loopt, onbegrijpbaar als de decimalen van π , die stoffige somer waarin de stad vol moersten stroomde, even plotseling veranderend in de voorjaar, herfst, en dan de bleke winter waarin de sterren van de straten en de muren opeens de ongemaakbare aard van de wereld leken uit te drukken, - maar dat alles eigenlijk in het voorjaar, oppgemerkt alleen in de korte tussenposes waarin hij van het ene interieur naar het andere ging. Was in de stad de natuur tachtig achtereenvolgende, op het kasteel zat hij met Quinten en Sjipje midden in de duivervende concertzaal. Lente en herfst kwamen met kolossaal vertoon, de somers waren hoter en droger, de winters koudter en winter. Die onafgebonden verandering, had hij tegen Otto beweerd, was natuurlijk de bron van alle creativiteit; de onvermoeidheid van de natuur tussen dit keerkringen hieldde ook toe culturele stilstand, de tropen waren een onafgebonden land, altijd groen zoals de poolgebieden altijd wit waren, van de primitieve breedtegraden met hun viervormigheid waren

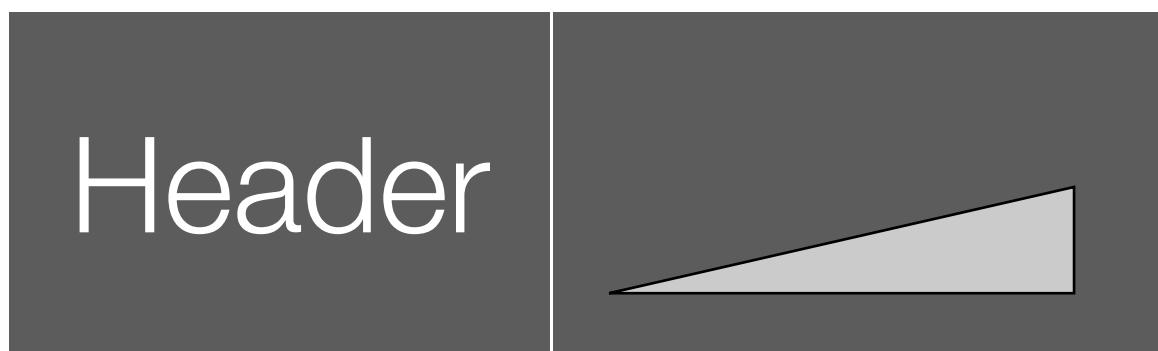
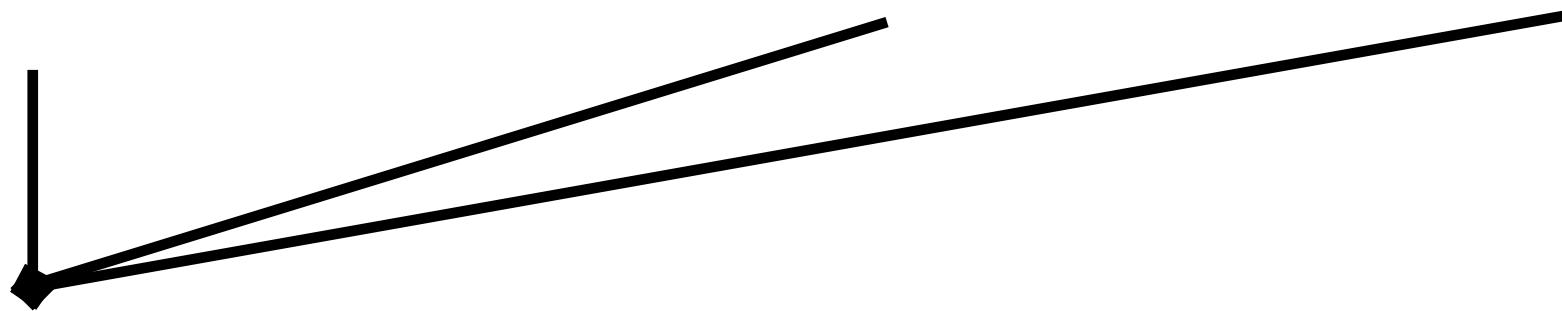
Reuse Background Merger to Compact Files



Reuse Background Merger to Compact Files



data-0000123.moss

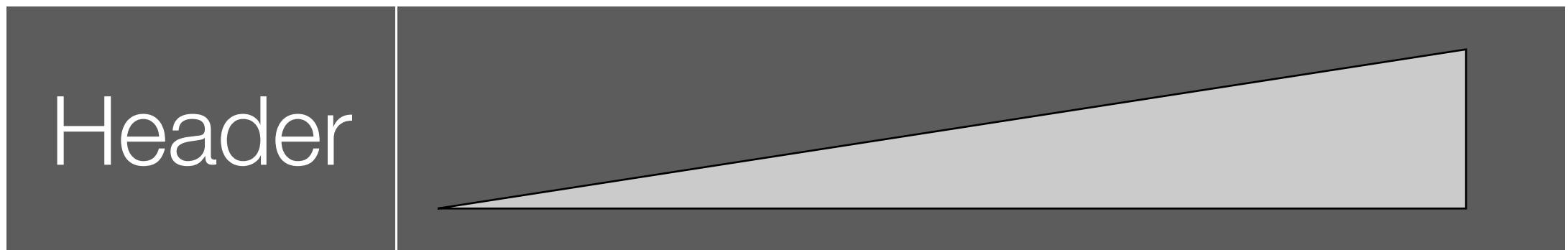
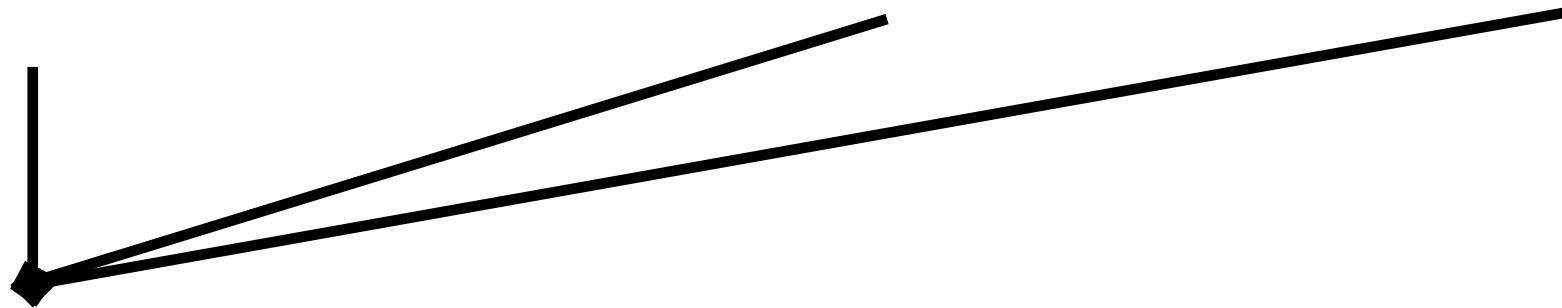


data-0000124.moss

Reuse Background Merger to Compact Files

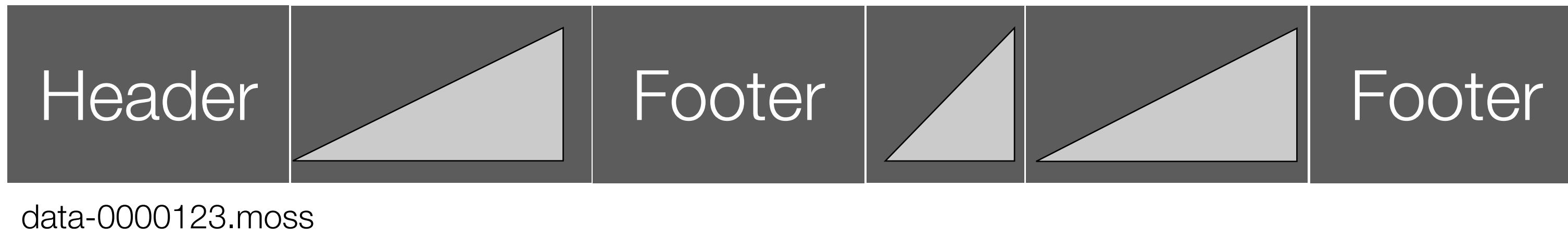


data-0000123.moss

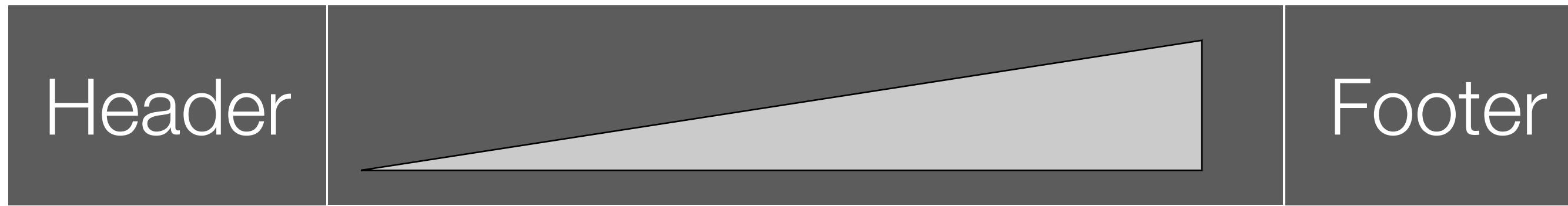


data-0000124.moss

Reuse Background Merger to Compact Files

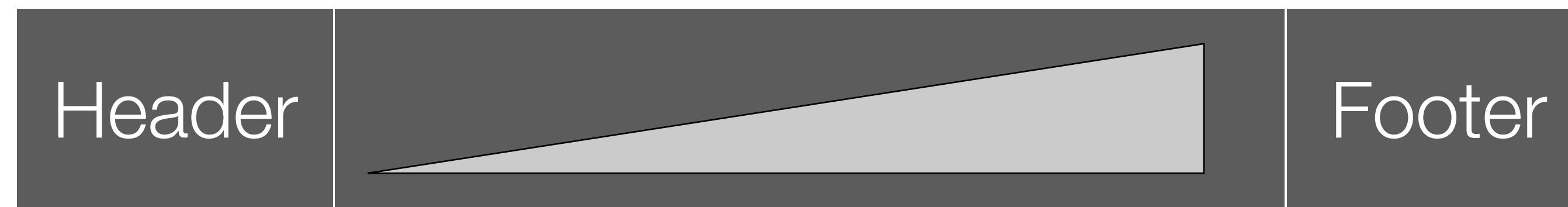


data-0000123.moss



data-0000124.moss

Reuse Background Merger to Compact Files



data-0000124.moss

Who Are You?

Zoals in Noordwijk het licht van de vuurpijnen langs de vier windrichtingen overvalt, zo streken elk jaar in grote vlagen de vier seisoenen over Groot Rechtersen. Max kende de wisseling van de jaargrijsen eigenlijk alleen uit Amsterdam: op een dag in februari of maart is er een onbeschrijflijke voorjaarsgeur als hij's morgens op straat loopt, onbegrijpbaar als de decimalen van π , die stoffige somer waarin de stad vol moeststenen stroomde, even plotseling veranderend in de voorjaar, herfst, en dan de bleke winter waarin de steden van de straten en de muren opeens de ongemaakbare aard van de wereld leken uit te drukken, - maar dat alles eigenlijk in het voorjaar, oppgemerkt alleen in de korte tussenposes waarin hij van het ene interieur naar het andere ging. Was in de stad de natuur tachtig achtereenvolgens, op het kasteel zat hij met Quinten en Sjouke midden in de daverende concertzaal. Lente en herfst kwamen met kolossaal vertoon, de somers waren hoter en droger, de winters koudster en winter. Die onafgebonden verandering, had hij nooit tegen Ottus beweerd, was natuurlijk de bron van alle creativiteit; de onvermoeidheid van de natuur tussen dit keerkringen hieldde ook toe culturele stilstand, de tropen waren een onafgebonden mensheid, altijd groen zoals de poolgebieden altijd wit waren, van de primitieve breedtegraden met hun viervormigheid waren







A close-up photograph of a large tree trunk, likely a beech, completely covered in a thick, vibrant green moss. The trunk curves slightly to the right. The ground in front of the trunk is covered with fallen, brown autumn leaves. In the background, more trees and foliage are visible, though slightly out of focus.

MOSS



Memory
Oriented
Sorted
Segments

[Code](#)[Issues 29](#)[Pull requests 0](#)[Projects 0](#)[Wiki](#)[Insights](#)

moss - a simple, fast, ordered, persistable, key-val storage library for golang

[326 commits](#)[2 branches](#)[0 releases](#)[8 contributors](#)[Apache-2.0](#)[Branch: master](#)[New pull request](#)[Create new file](#)[Upload files](#)[Find file](#)[Clone or download](#)

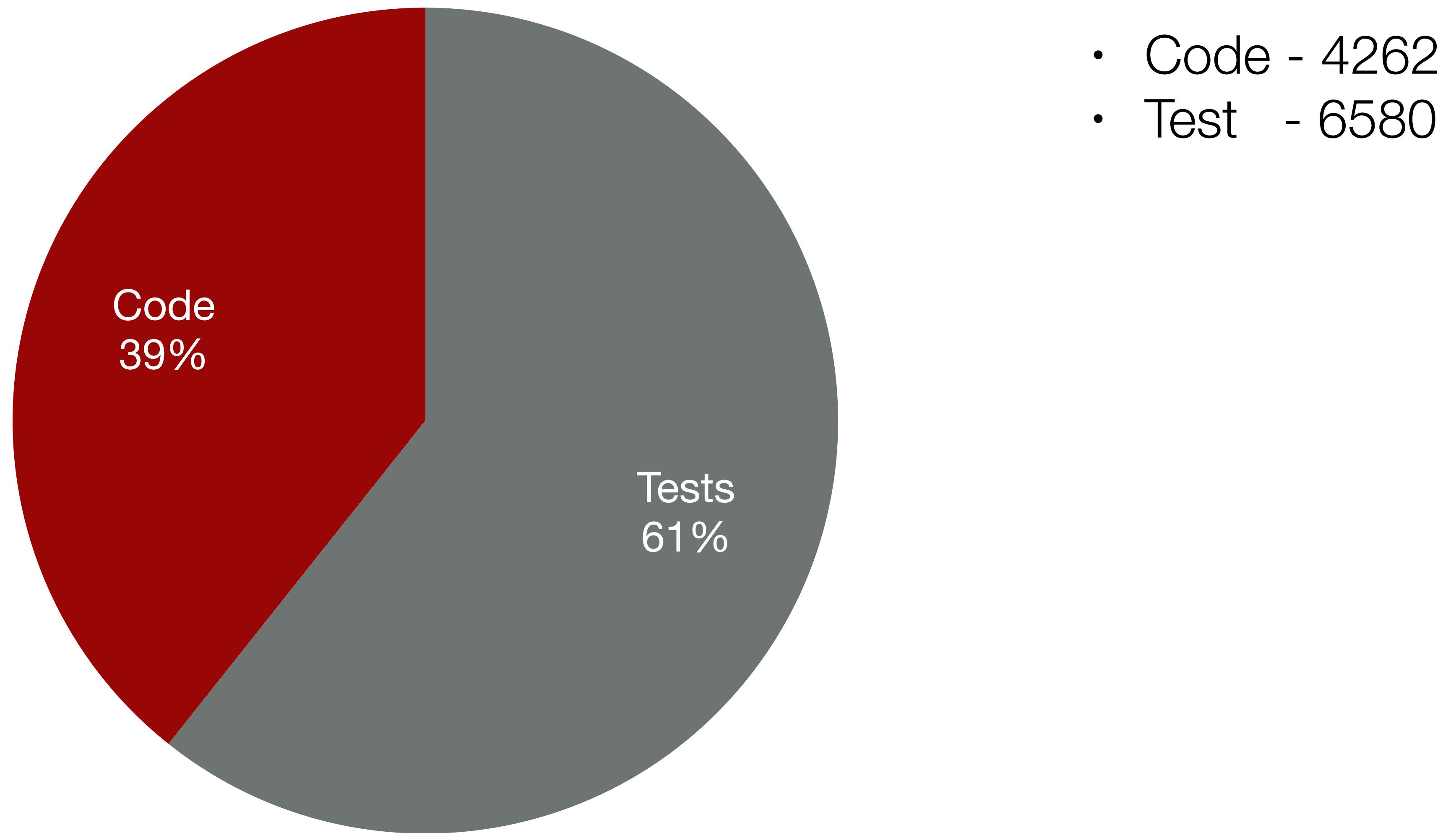
 hisundar committed with hisundar	Fix intermittent error in store_test due to async file deletions	...	Latest commit f718738 7 days ago
.gitignore	moved smat tests to main dir		a year ago
.travis.yml	attempted travis CI fix by upgrading to go 1.6		a year ago
CONTRIBUTING.md	added CONTRIBUTING.md		7 months ago
DESIGN.md	Update README.md & DESIGN.md with child-collection support		a month ago
IDEAS.md	update IDEAS.md, as nested collections & histograms are done		2 months ago
LICENSE.txt	added LICENSE - Apache 2.0		a year ago
Makefile	add gofmt to the metalinter invocation		3 months ago
README-smat.md	MB-21477 - refresh smat w/ more close/reopen/compaction		8 months ago
README.md	Update README.md & DESIGN.md with child-collection support		a month ago
api.go	doc: reserving child collection namespace (. prefix)		2 months ago
api_test.go	MB-18611 - added json tags for CollectionOptions funcs		a year ago
benchmark_get_test.go	Updating README, adding benchmark tests for Collection.Get()		3 months ago
benchmark_store_big_test.go	benchmark store test for 1 billion 20 byte keys, 0-sized values		11 months ago
benchmark_stress_test.go	MB-22501: Fix file ref counts to allow cleanup on connection		3 months ago

Features

- ordered key-val collection API
- 100% go implementation
- key range iterators
- snapshots provide for isolated reads
- atomic mutations via a batch API
- merge operations allow for read-compute-write optimizations for write-heavy use cases (e.g., updating counters)
- concurrent readers and writers don't block each other
- child collections allow multiple related collections to be atomically grouped
- optional, advanced API's to avoid extra memory copying
- optional lower-level storage implementation, called "mossStore", that uses an append-only design for writes and `mmap()` for reads, with configurable compaction policy; see: `OpenStoreCollection()`
- mossStore supports navigating back through previous commit points in read-only fashion, and supports reverting to previous commit points.
- optional persistence hooks to allow write-back caching to a lower-level storage implementation that advanced users may wish to provide (e.g., you can hook moss up to leveldb, sqlite, etc)
- event callbacks allow the monitoring of asynchronous tasks
- unit tests
- fuzz tests via go-fuzz & smat (github.com/mschoch/smat); see `README-smat.md`
- moss store's diagnostic tool: [mossScope](#)

License

Approachable Codebase



High Performance

Zoals in Nieuwbijk het licht van die vuurtoren langs de vier windrichtingen overtrok, zo streken elk jaar in grootse vlaggen die vier seisoenen over Groot Rechtersen. Max kende die wisseling van de jaargroeiën eigenlijk alleen uit Amsterdam: op een dag in februari of maart kwam, onbeschrijflijke voorjaarsgeur als hij's morgens op straat kwam, onbeschrijfbaar als de decimalen van π , die stoffige somer waarin de stad vol moeststen strooimde, even plotseling veranderend in de voorjaar, herfst, en dan de bleke winter waarin de sterren van de straten en de muren opeens de ongemaakbare aard van de wereld lieten uit te drukken, - maar dat alles eigenlijk in het voorjaar, oppenseind alleen in de korte tussenposes waarin hij van het ene interieur naar het andere ging. Was in de stad de natuur tachtig achtergrondmuziek, op het kastiel zat hij met Quinten en Sophie midden in de daverende concertzaal. Lente en herfst kwamen met kolossaal verloren, de somers waren hoter en dinget, de winters koudter en winter. Die onafgebonden verandering, had hij tegen Ottau beweerd, was natuurlijk de bron van alle creativiteit: de onvermoeidheid van de natuur tussen de koekringen hield ook toe culturele stilstand, de tempen waren een onafgebonden heiligtum, altijd groen zoals de poolghieden altijd wit waren, maar de primitieve breedtegraden met hun viervormigheid waren



High Performance - Minimize GC Impact

```
type segment struct {  
    data []byte  
    meta []uint64  
}
```

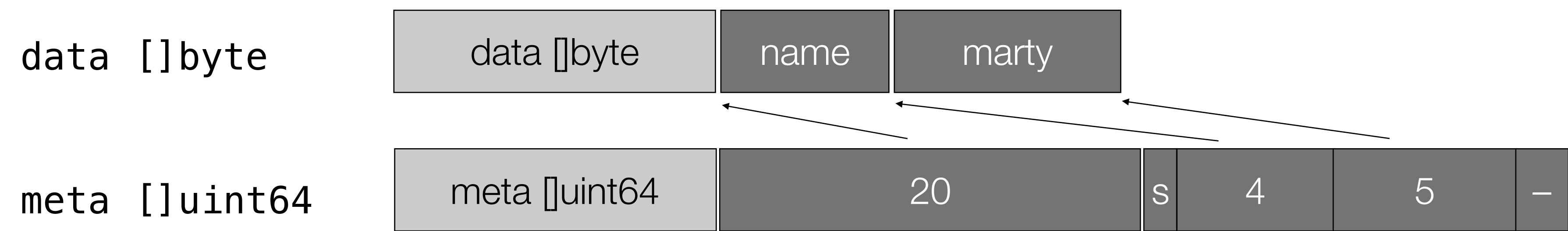
High Performance - Minimize GC Impact

```
type segment struct {  
    data []byte  
    meta []uint64  
}
```

- Slices of uint64 and byte

High Performance - Minimize GC Impact

```
type segment struct {
```



```
}
```

- Slices of `uint64` and `byte`
- Integer offsets into slices, not pointers

High Performance - Memory Allocation

```
type segment struct {  
    data []byte  
    meta []uint64  
}
```

High Performance - Memory Allocation

```
type segment struct {  
    data []byte  
    meta []uint64  
}
```

- 2 uints of meta per op

High Performance - Memory Allocation

```
type segment struct {  
    data []byte  
    meta []uint64  
}
```

- 2 uints of meta per op
- $\text{len}(\text{data}) = \text{len}(\text{key}) + \text{len}(\text{val})$

High Performance - Memory Allocation

```
type Collection interface {  
    NewBatch(ops, bytes int)  
    (Batch, error)  
}
```

High Performance - Memory Allocation

```
type Batch interface {
    AllocSet(key, val []byte) error
    AllocDel(key []byte) error
}
```

High Performance - Unsafe

High Performance - Unsafe

- Faster serialization, but loss of portability

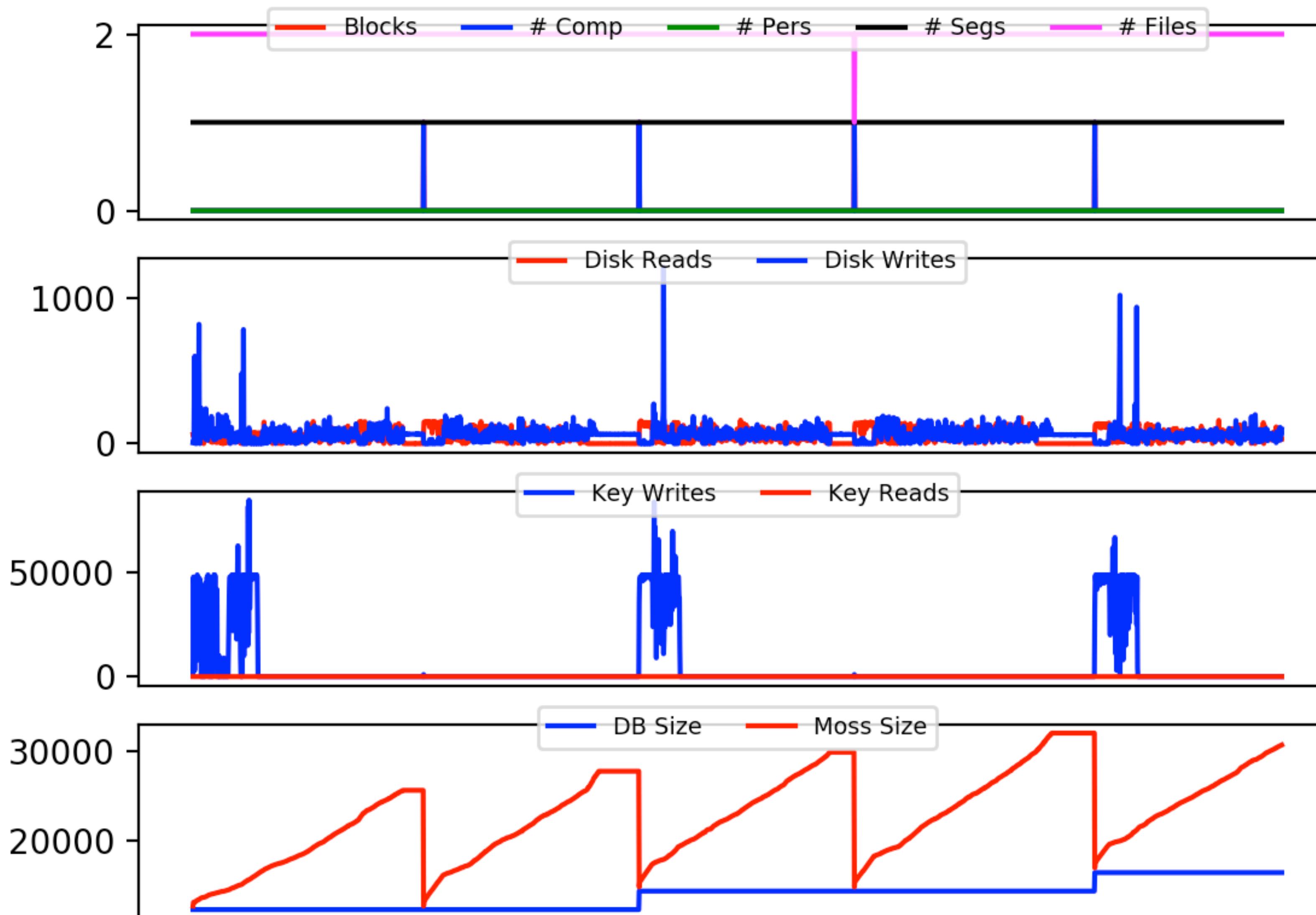
High Performance - Unsafe

- Faster serialization, but loss of portability
- Deeply unsatisfying

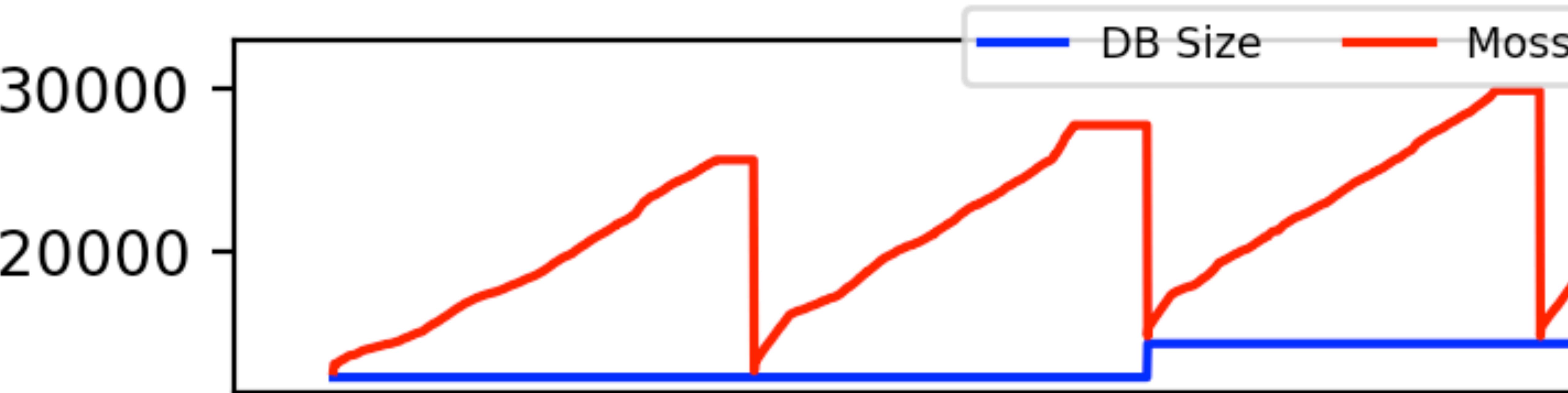
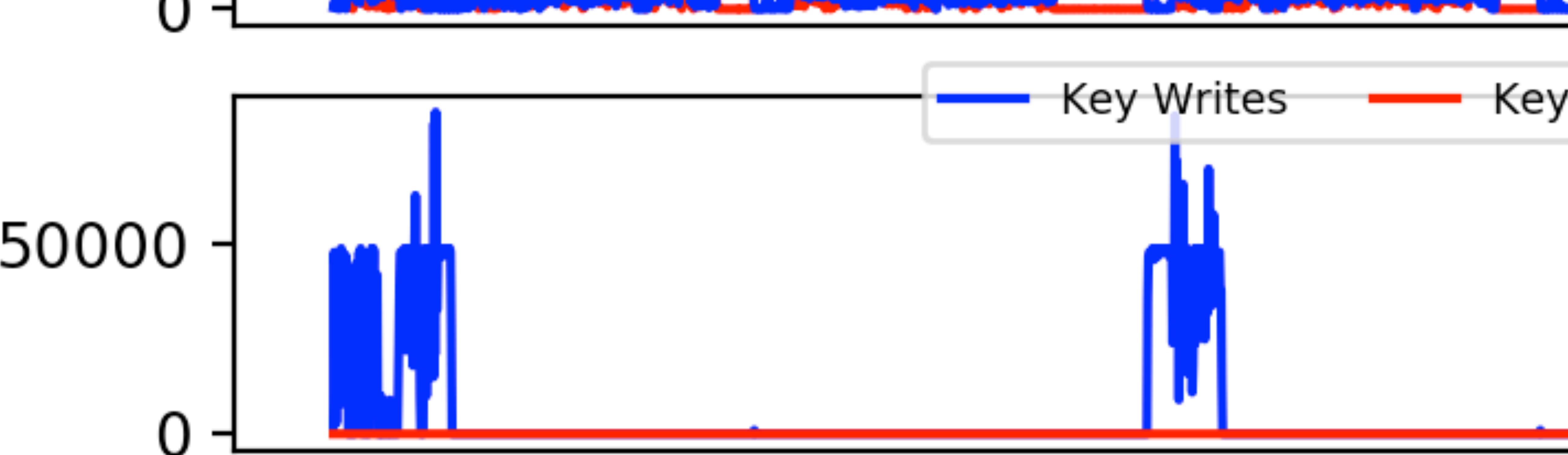
Rob Pike – The byte order fallacy

"Whenever I see code that asks what the native byte order is, it's almost certain the code is either wrong or misguided."

2gmq Runtime 3600s Warmup 0s CompPerc 0.0 MemQ 2048Mb
Writers 1 BatchSize 1000 ThinkTime 20ms KeyOrder asc #Readers 0 RangeSize 100 ThinkTime 0ms



Disk I/O # Writes 168313 78.7Gb Amp 4.9 Reads 146146 70.9Gb Amp 4.4
DB Size # Keys 8391000 Key Space 0.063Gb Value Space 15.9Gb
Persists 0 # Compactions 5 Blocks 5 BlockTime 3224.756s
Kevs Writes 12586000 Batches 12586 Reads 0 Batches 0



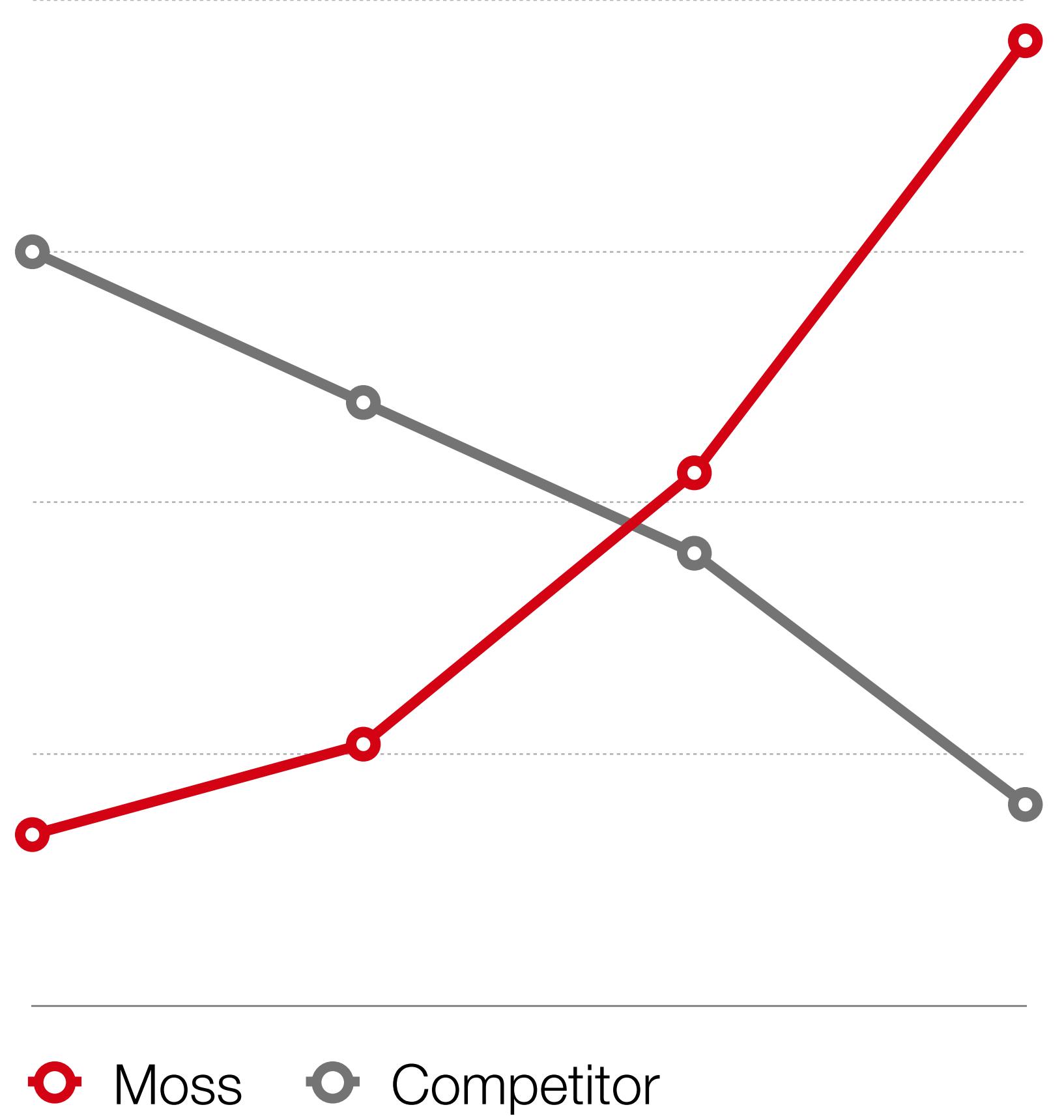
One Key Written
Entire Store Rewritten

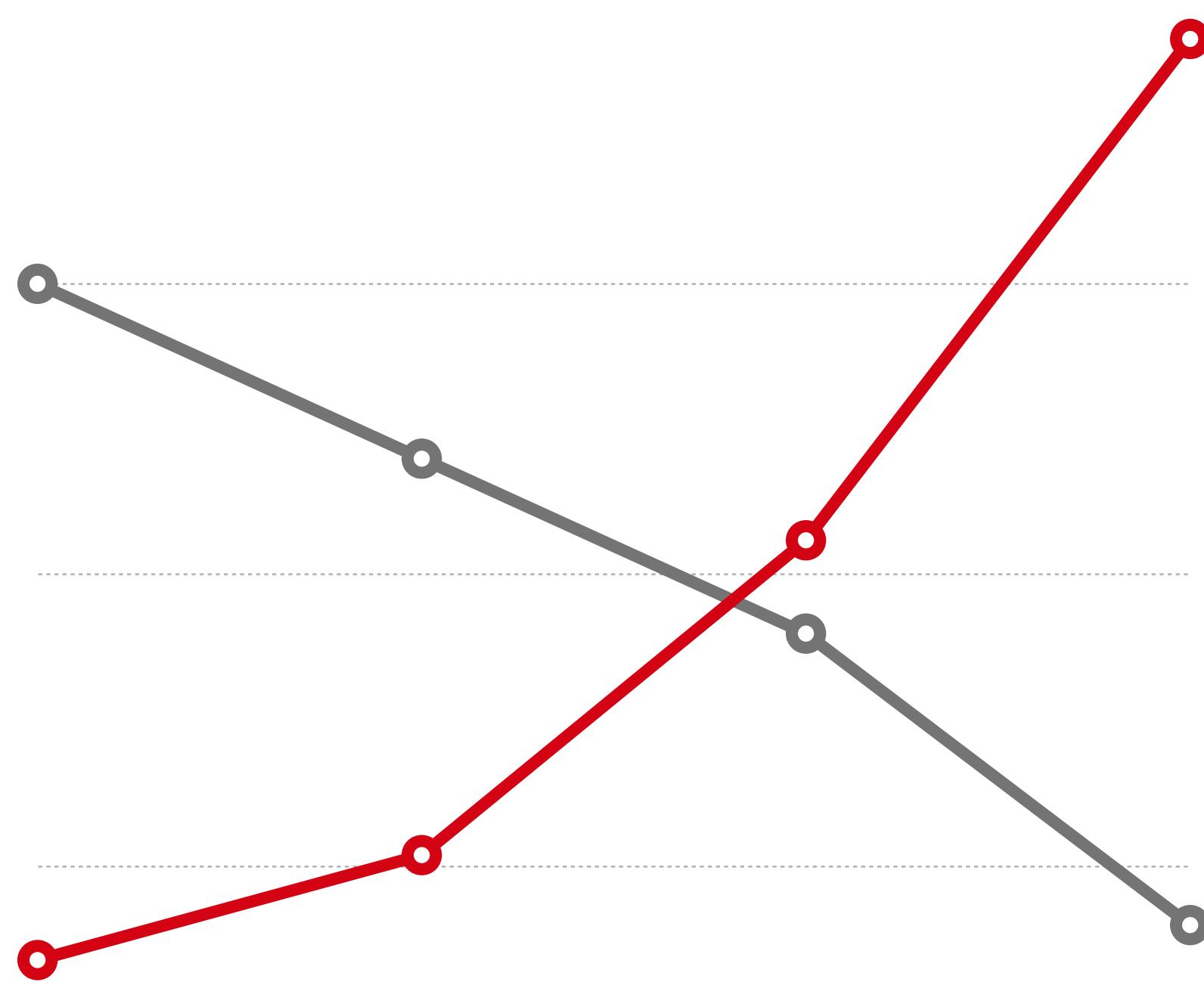
DB Si

Conclusions

Zoals in Noordwijk het licht van die vuurtoren langs de vier windrichtingen overtrok, zo streken elk jaar in grootsche vlaggen die vier seisoenen over Groot Rechtersen. Max kende die wisseling van de jaargroeiën eigenlijk alleen uit Amsterdam: op een dag in februari of maart kwam een onbeschrijflijke voorjaarsgeur als hij's morgens op straat kwam, onbelabbaar als de decimalen van π , die stoffige somer waarin de stad vol roerstoten strooimde, even plotseling veranderend in de weelde, hitte, herfst, en dan de bleke winter waarin de sterren van de straten en de muren opeens de ongemaakbare aard van de wereld leken uit te drukken, - maar dat alles eigenlijk in het moedigen, oppemerkt alleen in de korte tussenposes waarin hij van het ene interieur naar het andere ging. Was in de stad de natuur achtergrondmuziek, op het kasteel zat hij met Quinten en Sophie midden in de daverende concertzaal. Lente en herfst kwamen met kolossaal verloren, de somers waren hoter en dinget, de winters koudter en winter. Die onafgebonden verandering, had hij tegen Otto beweerd, was natuurlijk de bron van alle creativiteit: de onvermoeidheid van de natuur tussen de koekringen hield ook een culturele stilstand, de treinen waren een onafgebonden mensheid, altijd groen zoals de poolglieden altijd wit waren, maar de primitieve breedtegraden met hun viervormigheid waren

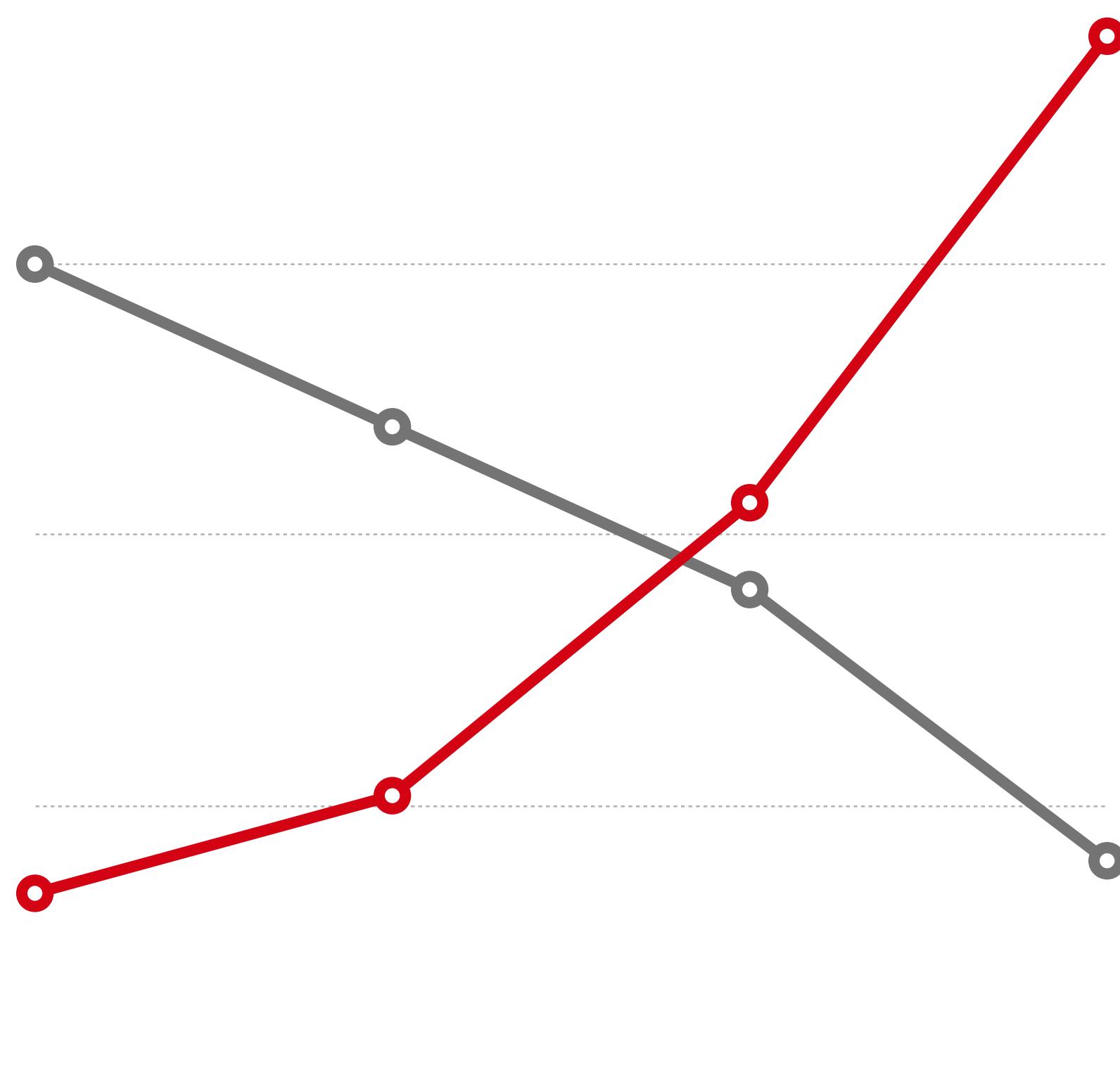






● Moss ● Competitor





● Moss

● Competitor

 **Moss Fanboy** 
@mossbro 

Moss is the best, fastest!
#winning

4 5,244 15,542

 **Moss Fanboy** 
@mossbro 

Competition is slow. Sad!

11:50 AM - 22 Dec 2016
4 5,244 15,542



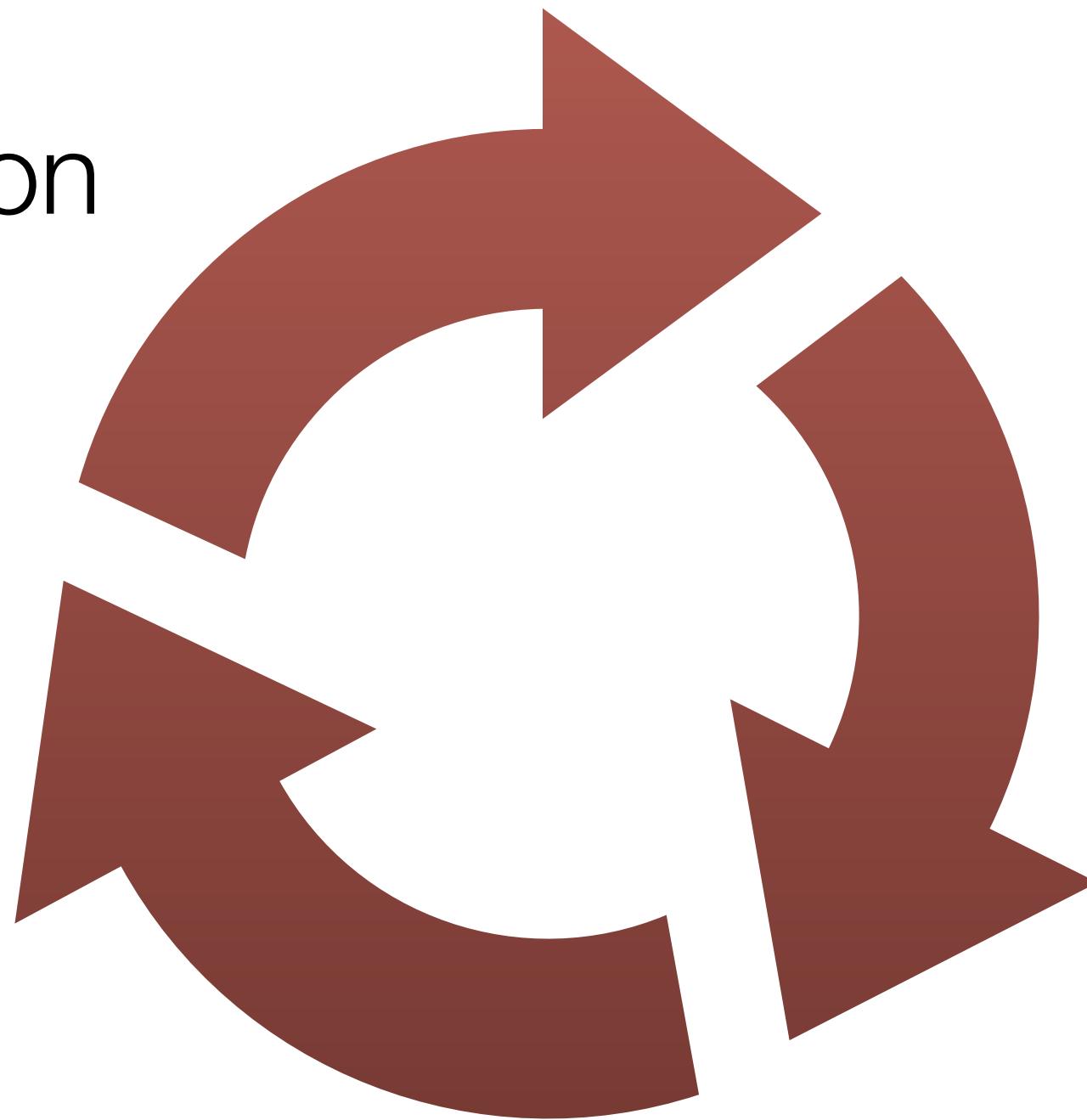
Read after Write Latency



—

Write Throughput

Implementation



Benchmark

Analyze and Profile

Epilogue

Steve Yen

Zoals in Nieuwbijk het licht van de vuurtoren langs de vier windrichtingen wisselt, zo streken elk jaar in grootsche vlaggen die vier seisoenen over Groot Rechtersen. Max kende die wisseling van de jaargroeiën eigenlijk alleen uit Amsterdam: op een dag in februari of maart kwam een onbeschrijflijke voorjaarsgeur als hij's morgens op straat kwam, onbelabbaar als de decimalen van π , die stoffige nummer waarin de stad vol moersten strooimde, even plotseling veranderend in de voorjaar, herfst, en dan de bieke winter waarin de sterren van de straten en de muren opeens de ongemaakbare aard van de wereld leken uit te drukken, - maar dat alles eigenlijk in het voorjaar, oppenseind alleen in de korte tussenposes waarin hij van het ene interieur naar het andere ging. Was in de stad de natuur achtergrondmuziek, op het kasteel zat hij met Quinten en Sophie midden in de daverende concertzaal. Lente en herfst kwamen met kolossaal verloren, de zomers waren heet en dinget, de winters koud en winter. Die onafgebonden verandering, had hij tegen Otto beweerd, was natuurlijk de bron van alle creativiteit: de onvermoeidheid van de natuur tussen de keerkringen leidde tot een culturele stilstand, de treinen waren een onafgebonden troost, altijd groen zoals de poolglieden altijd wit waren, maar de primitieve breedtegraden met hun viervormigheid waren





Abhinav Dangeti
Sundar Sridharan
Sreekanth Sivasankaran
Scott Lashley
Marty Schoch
Alex Gyryk
Aruna Piravi

490



Couchbase Moss Team

Zoals in Noordwijk het licht van die vuurtorens langs de vier windrichtingen overtrok, zo streken elk jaar in grootse vlaggen die vier seisoenen over Groot Rechtersen. Max kende die wisseling van de jaargroeiën eigenlijk alleen uit Amsterdam: op een dag in februari of maart kwam een onbeschrijflijke voorjaarsgeur als hij's morgens op straat liep, onbelabbaar als de decimalen van π , die stoffige nummer waarin de stad vol moersten strooimde, even plotseling veranderend in de vroegte, hitte herfst, en dan die bleke winter waarin de sterren van de straten en de muren opeens de ongemaakbare aard van de wereld leken uit te drukken, - maar dat alles eigenlijk in het voorbijgaan, oppenseindig alleen in de korte tussenposes waarin hij van het ene interieur naar het andere ging. Was in de stad de natuurlijke achtergrondmuziek, op het kasteel zat hij met Quinten en Sophie midden in de daverende concertzaal. Lente en herfst kwamen met kolossaal verloop, de zomers waren heet en dinget, de winters koud en winter. Die onafgebonden verandering, had hij nooit begrepen. Otto beweerde, was natuurlijk de bron van alle creativiteit de onvermoeidheid van de natuur tussen die keerkringen heinde tot de culturele stilstand, die treinen waren een onafgebonden broedplaats, altijd groen zoals de poolgebieden altijd wit waren, toen de primitieve breedtegraden met hun viervormigheid warm

491

Thanks

github.com/couchbase/moss

marty@couchbase.com

@mschoch

Zoals in Noordwijk het licht van die vuurtorens langs de vier windrichtingen overvalt, zo streken elk jaar in grootsche vlaggen die vier seisoenen over Groot Rechtersen. Max kende die wisseling van de jaargroeiën eigenlijk alleen uit Amsterdam: op een dag in februari of maart kwam een onbeschrijflijke voorjaarsgeur als hij's morgens op straat kwam, onbelauwbaar als de decimalen van π , die stoffige nummer waarin de stad vol moersten strooimde, even plotseling veranderend in de weichege, hinnere herfst, en dan die bleke winter waarin de sterren van de straten en de muren opeens de ongemaakbare aard van de wereld leken uit te drukken, - maar dat alles eigenlijk in het voorbijgaan, oppenseindig alleen in de korte tussenposes waarin hij van het ene interieur naar het andere ging. Was in de stad de natuur achtergrondmuziek, op het kasteel zat hij met Quinten en Sophie midden in de daverende concertzaal. Lente en herfst kwamen met kolossaal verloren, de zomers waren heet en dinget, de winters koudet en winter. Die onafgebonden verandering, had hij tegen Otto beweerd, was natuurlijk de bron van alle creativiteit: de onvermoeidheid van de natuur tussen de koekringen leidde tot een culturele stilstand, de tempen waren een onafgebonden broedplaats, altijd groen zoals de poolglieden altijd wit waren, maar de primitieve breedtegraden met hun viervormigheid waren



Thanks

github.com/couchbase/moss

marty@couchbase.com

@mschoch

Zoals in Noordwijk het licht van die vier windrichtingen overkoopt, zo streken elk jaar in grootse vlaggen de vier seisoenen over Gorin Rechtersen. Max kende die wisseling van de jaargroeiën eigenlijk alleen uit Amsterdam: op een dag in februari of maart kwam, onbeschrijflijke voorjaarsgeur als hij's morgens op straat kwam, onbeschrijfbaar als de decimalen van π , die stoffige nummer waarin de stad vol toeristen strooimde, even plotseling veranderend in de vrolijke, hinnende herfst, en dan die bleke winter waarin de sterren van de straten en de muren opeens de ongemaakbare aard van de wereld leken uit te drukken, - maar dat alles eigenlijk in het voorbijgaan, oppgemerkt alleen in de korte tussenposes waarin hij van het ene interieur naar het andere ging. 'Was in de stad de natuurlijke achtergrondmuziek, op het kasteel zat hij met Quinten en Sophie midden in de daverende concertzaal. Lente en herfst kwamen met kolossaal verloren, de zomers waren heet en dinget, de winters koud en winter. Die onafgebonden verandering, had hij tegen zijn Ottos beweerd, was natuurlijk de bron van alle creativiteit, de onvermoeidheid van de natuur tussen die keerkringen hieldde ook toe culturele stilstand, de tempen waren een onafgebonden broedplaats, altijd groen zoals de poolghieden altijd wit waren, maar de primitieve breedtegraden met hun viervormigheid waren