laboratory
medical
immunology

Erasmus MC
University Medical Center Rotterdam

FLOW CYTOMETRY – DATA COMPRESSION

A.E. Bras

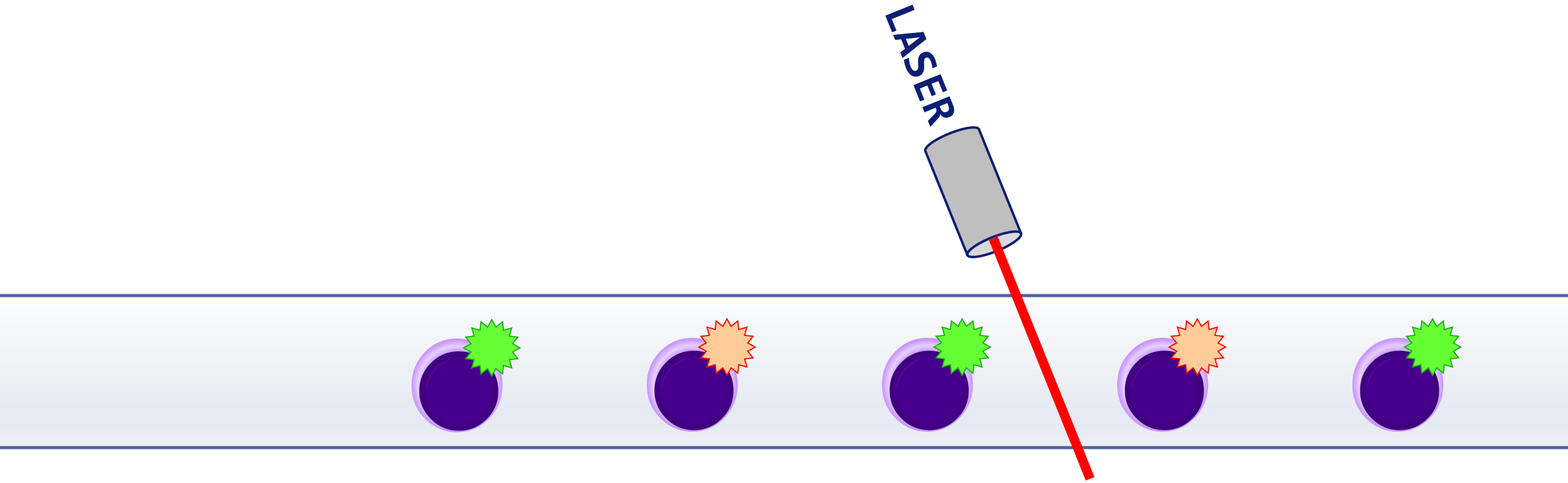
PhD Student

Erasmus University, Rotterdam, the Netherlands

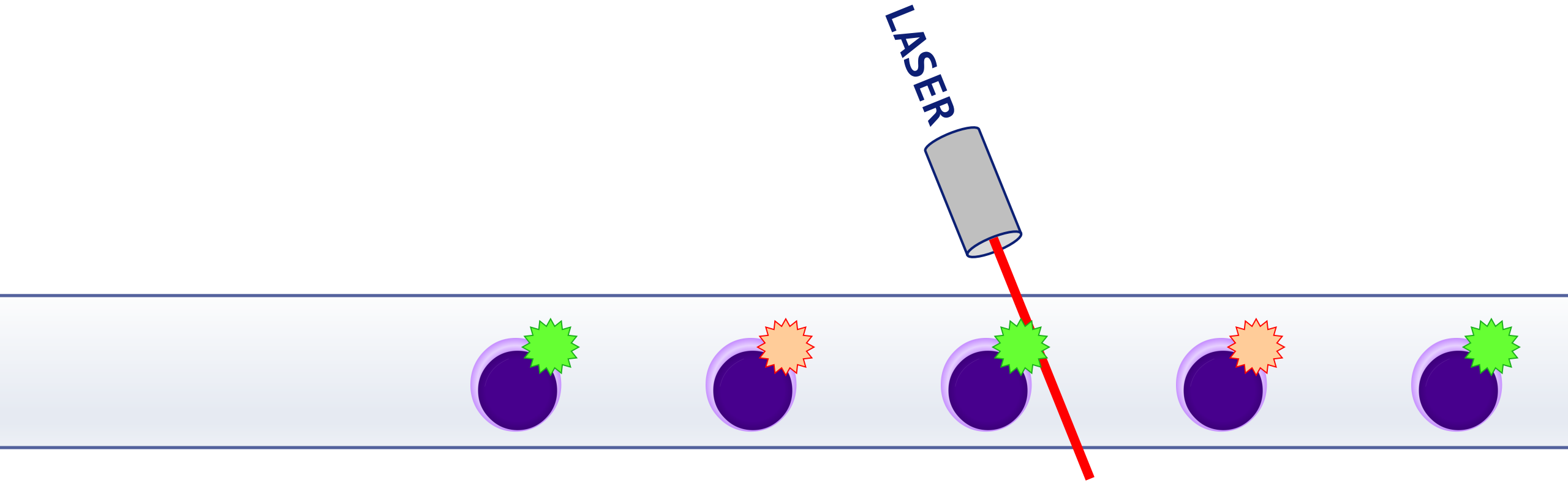
Flow Cytometry

Flow Cytometry

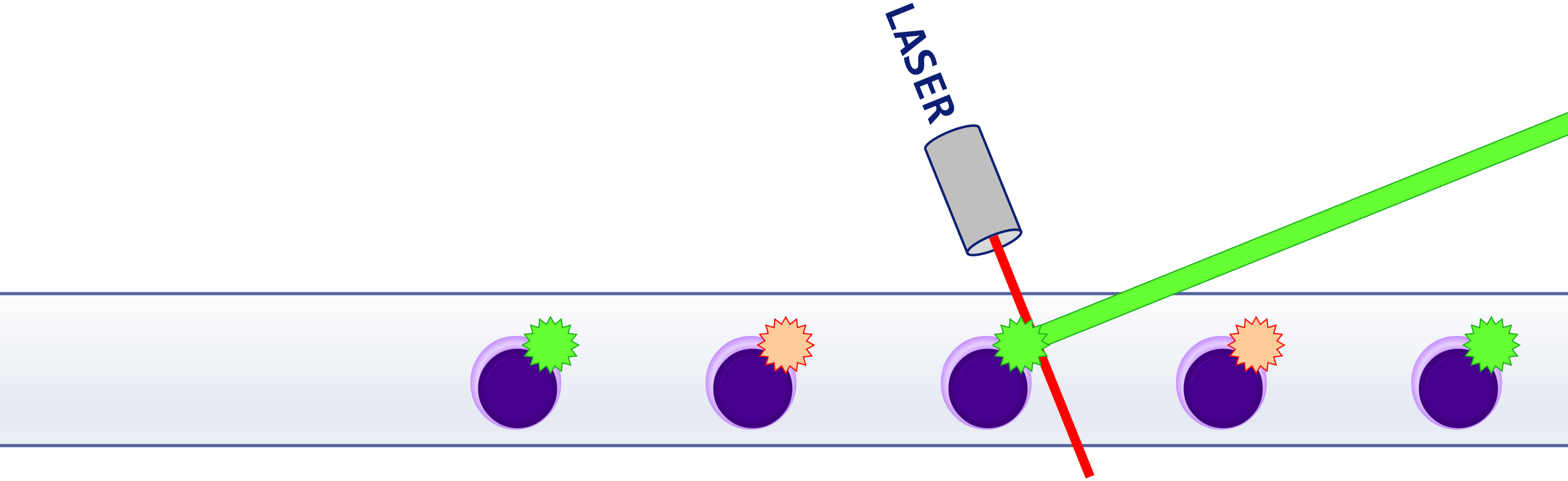
Flow Cytometry



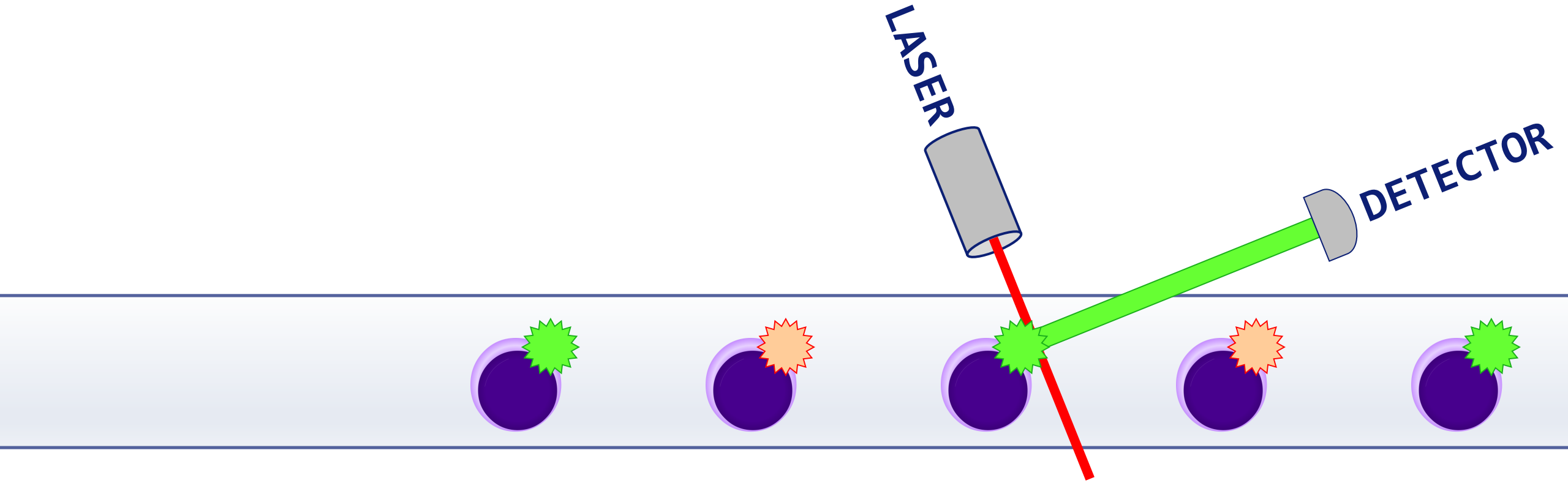
Flow Cytometry



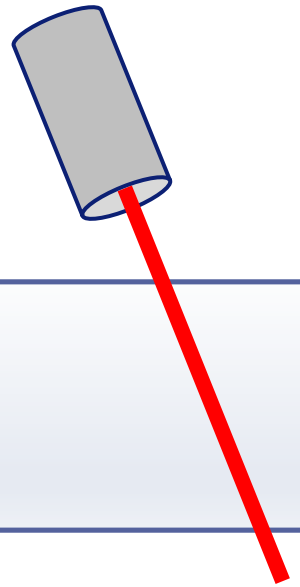
Flow Cytometry



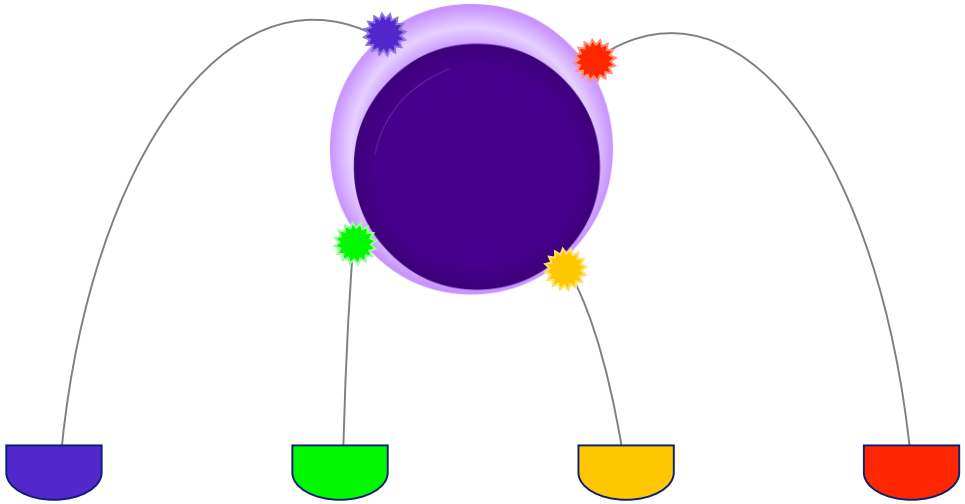
Flow Cytometry

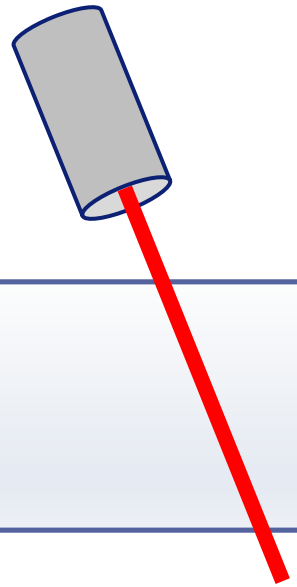


Older Systems

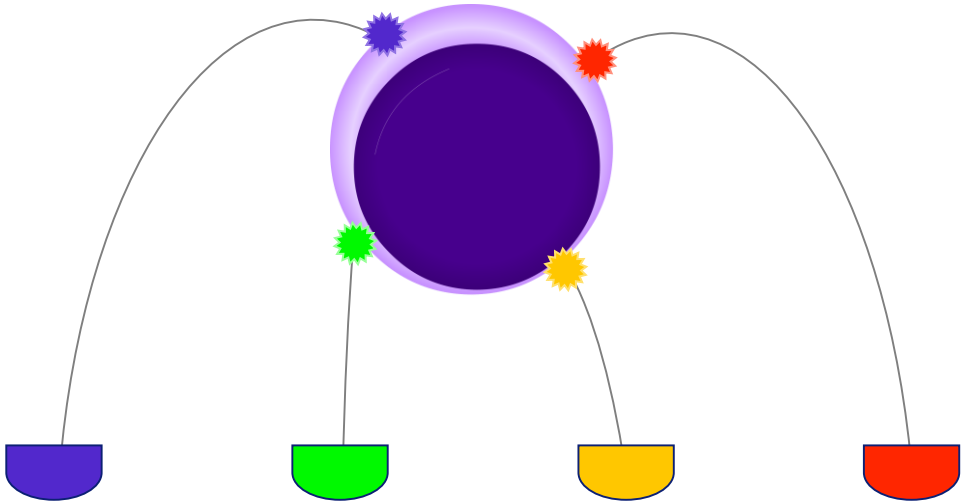


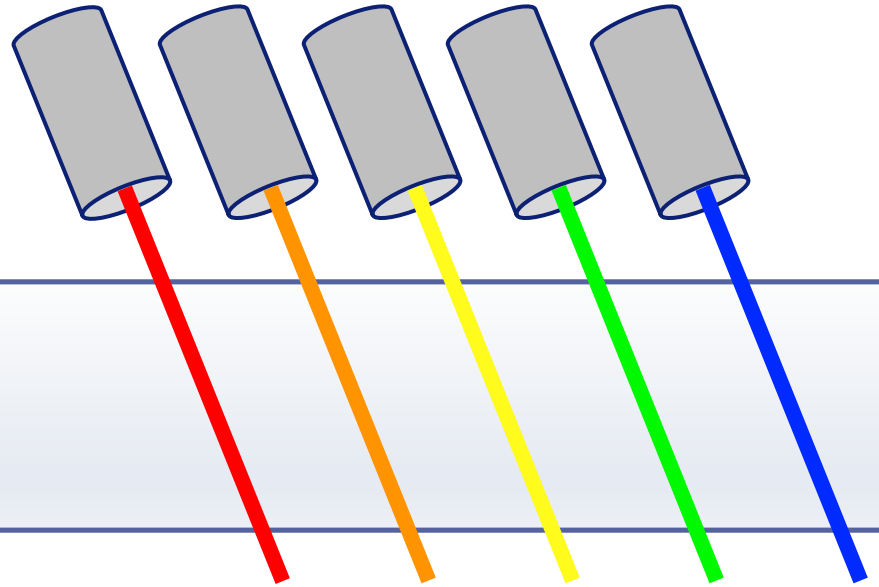
Older Systems



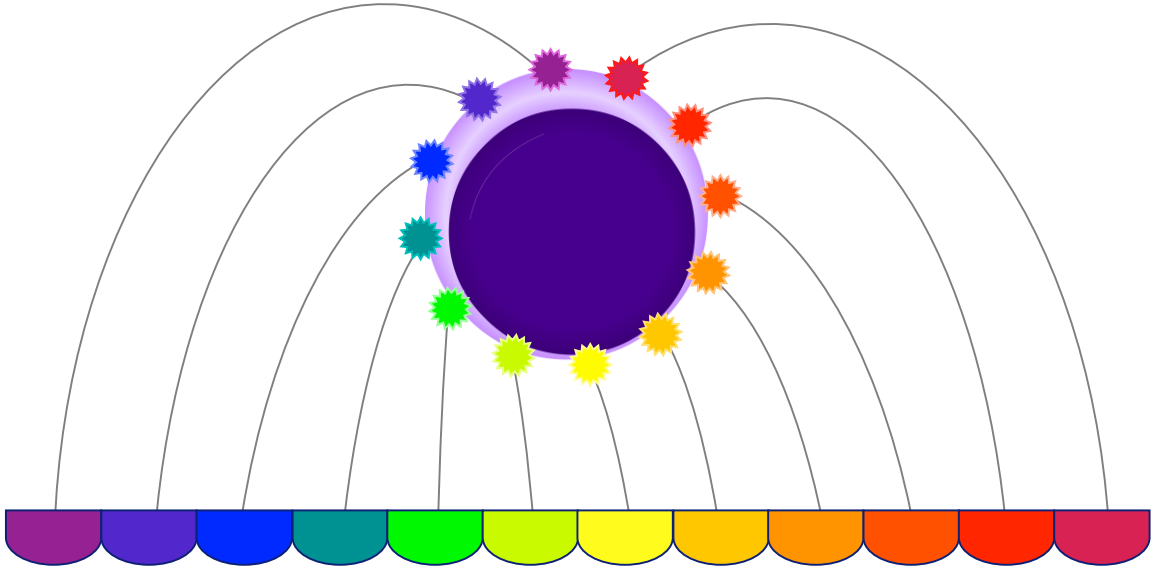


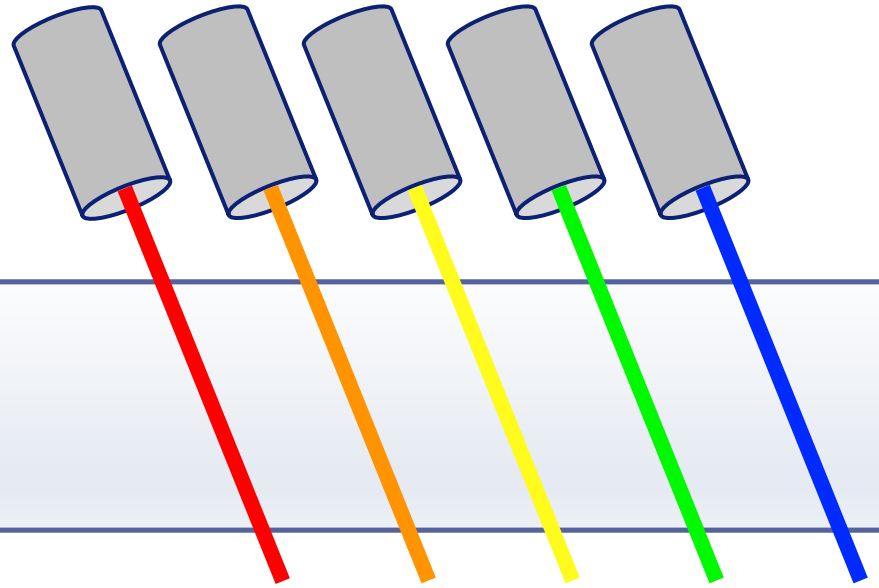
Newer Systems



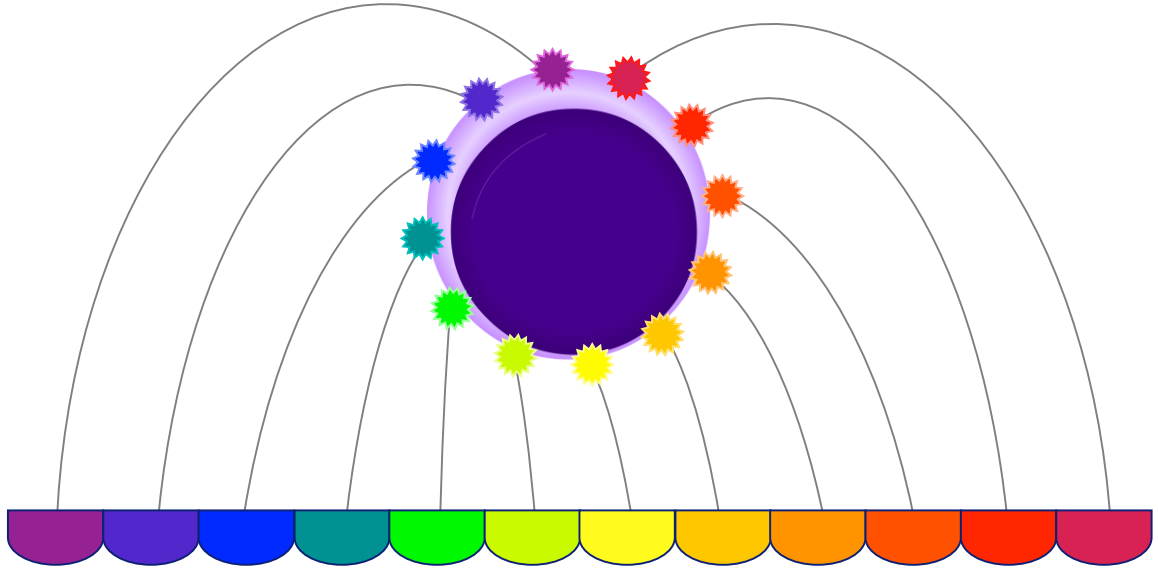


Newer Systems

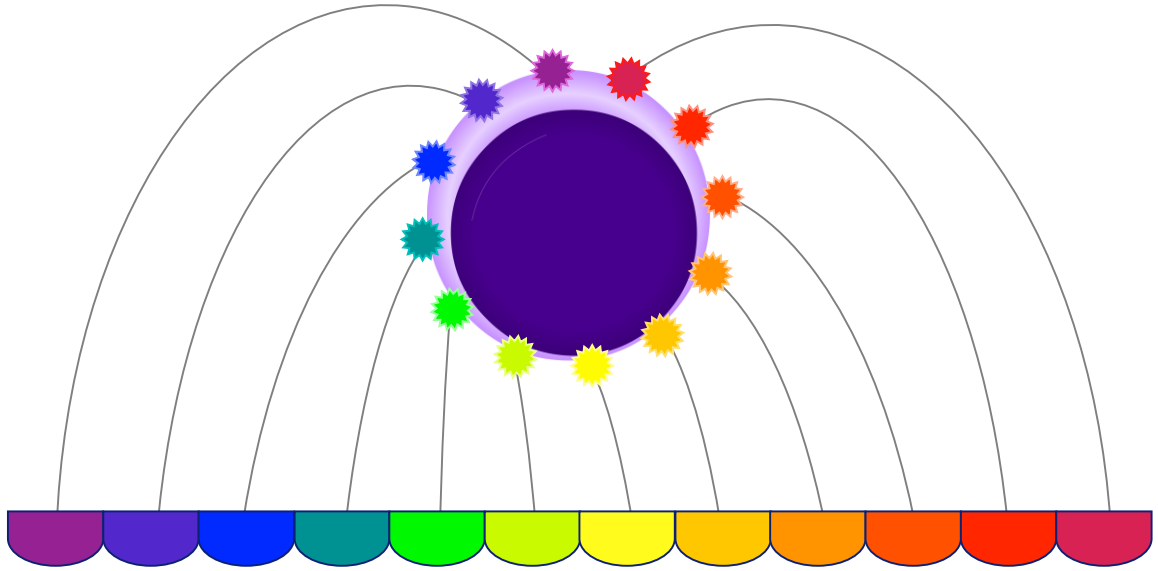




Newer Systems

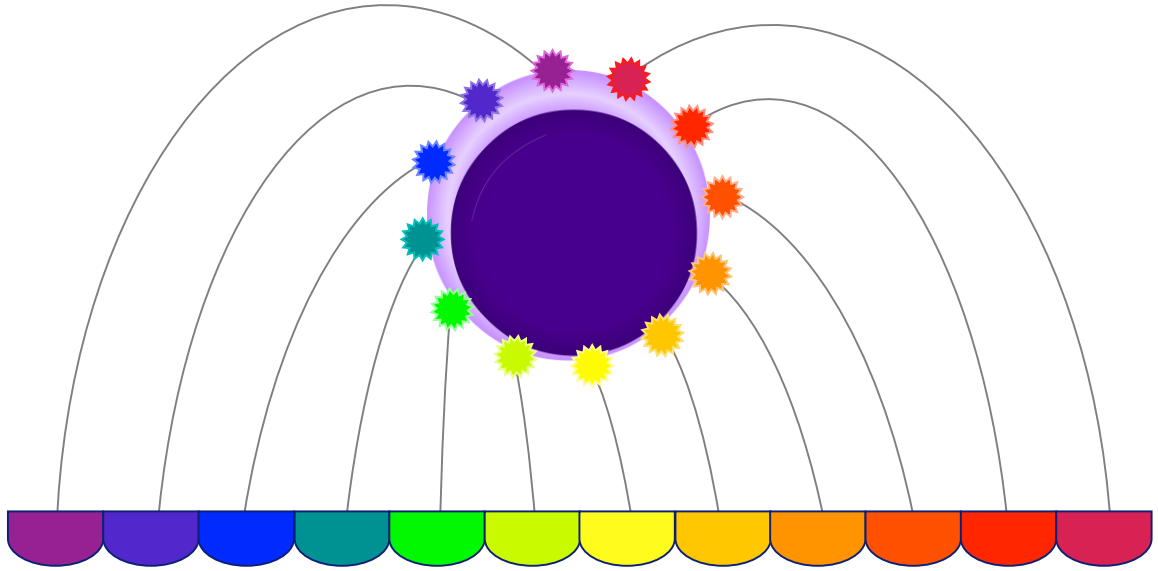


Newer Systems



FLOATING POINT DATA											
-	-	-	-	-	-	-	-	-	-	-	-

Newer Systems



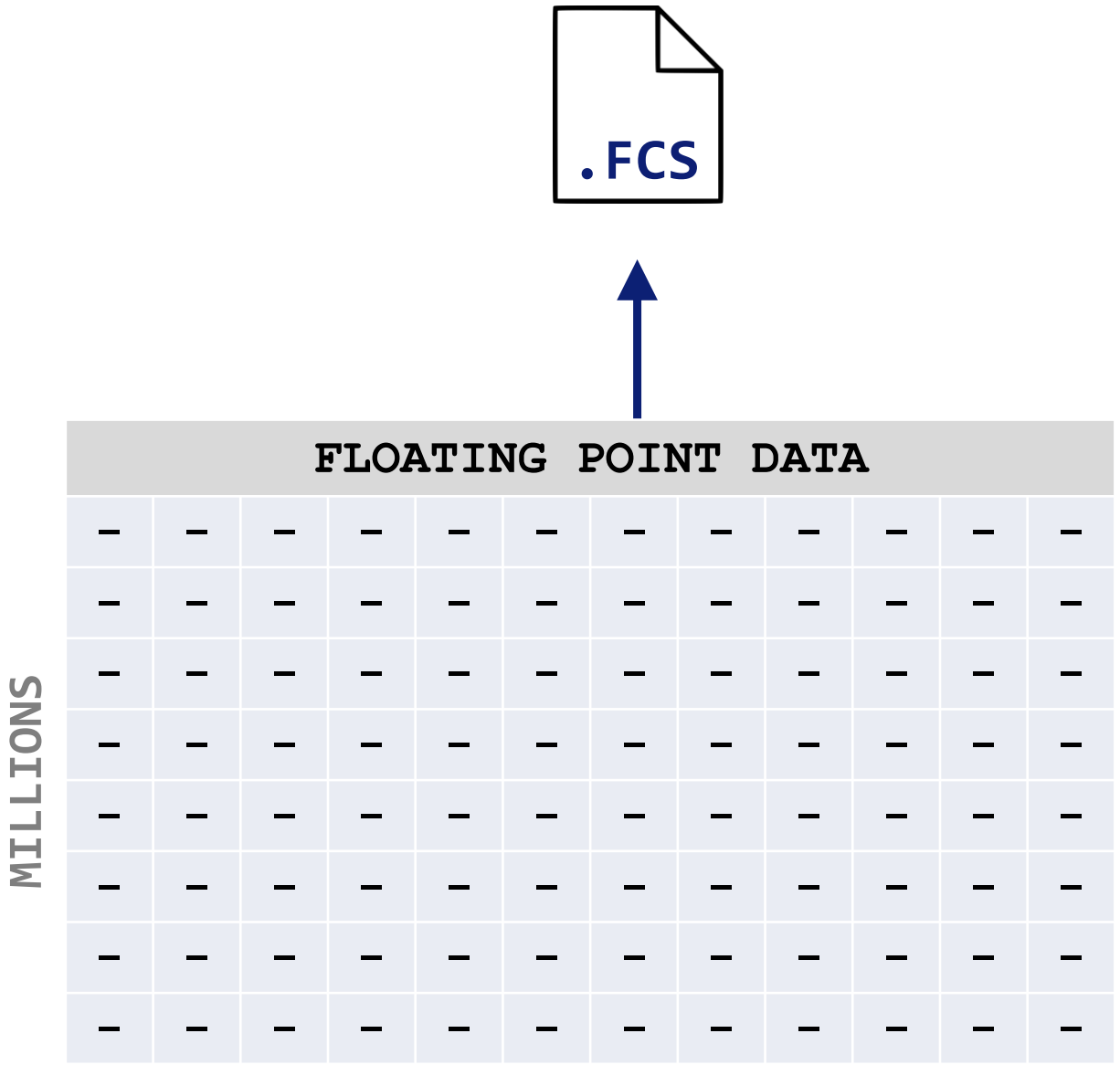
MILLIONS

FLOATING POINT DATA											
–	–	–	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–	–
–	–	–	–	–	–	–	–	–	–	–	–

Storage

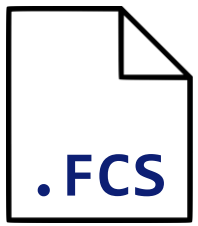
MILLIONS	FLOATING POINT DATA											
	–	–	–	–	–	–	–	–	–	–	–	–
	–	–	–	–	–	–	–	–	–	–	–	–
	–	–	–	–	–	–	–	–	–	–	–	–
	–	–	–	–	–	–	–	–	–	–	–	–
	–	–	–	–	–	–	–	–	–	–	–	–
	–	–	–	–	–	–	–	–	–	–	–	–
	–	–	–	–	–	–	–	–	–	–	–	–
	–	–	–	–	–	–	–	–	–	–	–	–

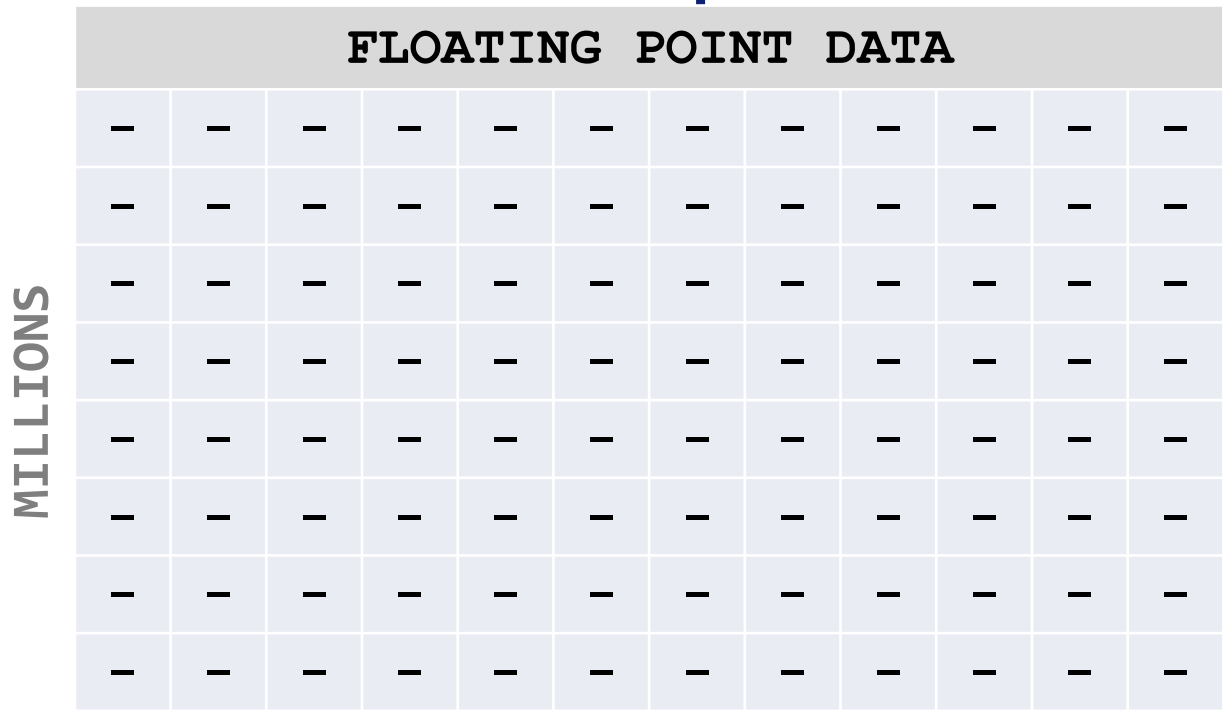
Storage



Storage

FLOW CYTOMETRY STRANDARD



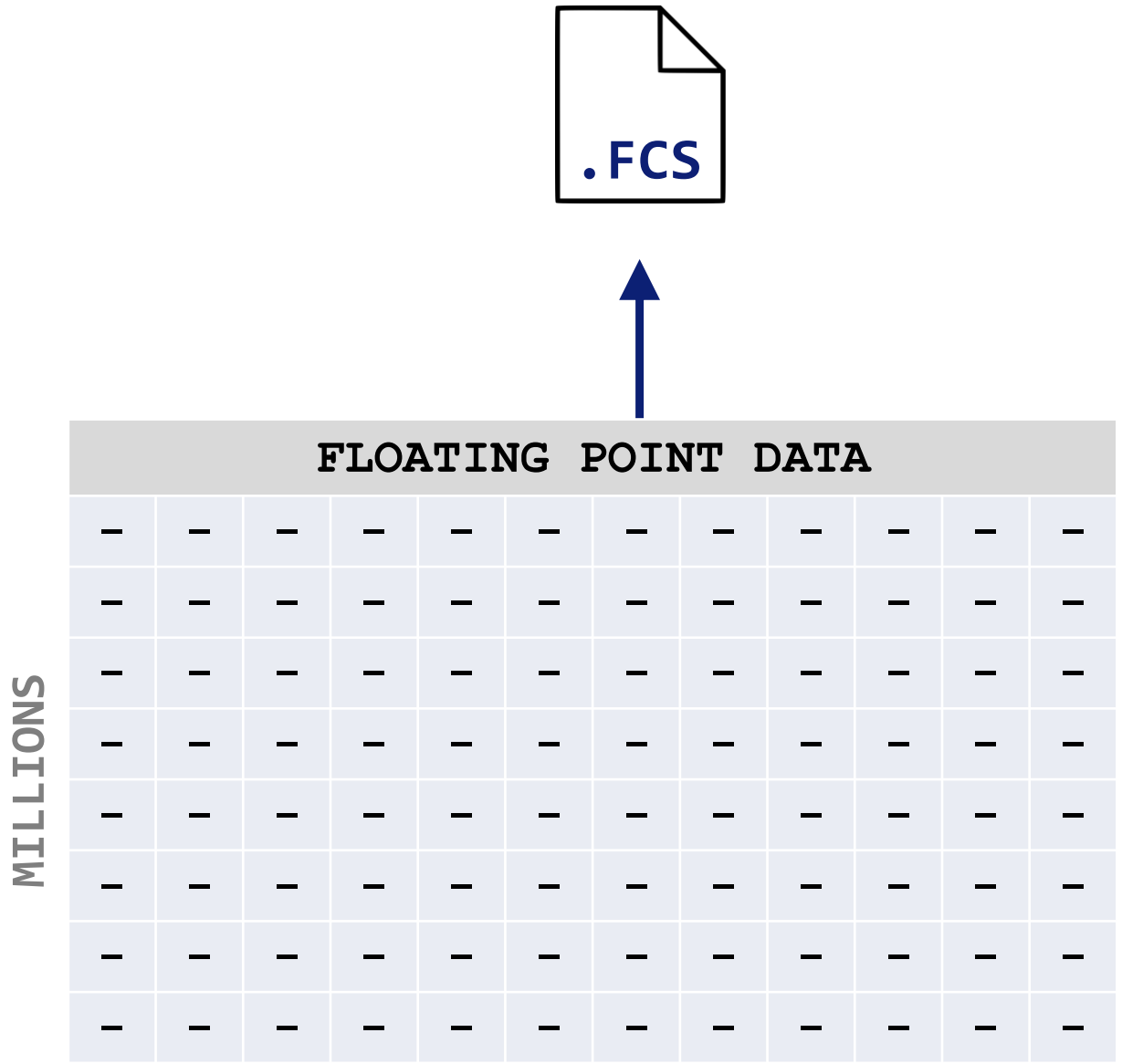


Storage

FLOW CYTOMETRY STRANDARD

PRO

WIDELEY USED



Storage

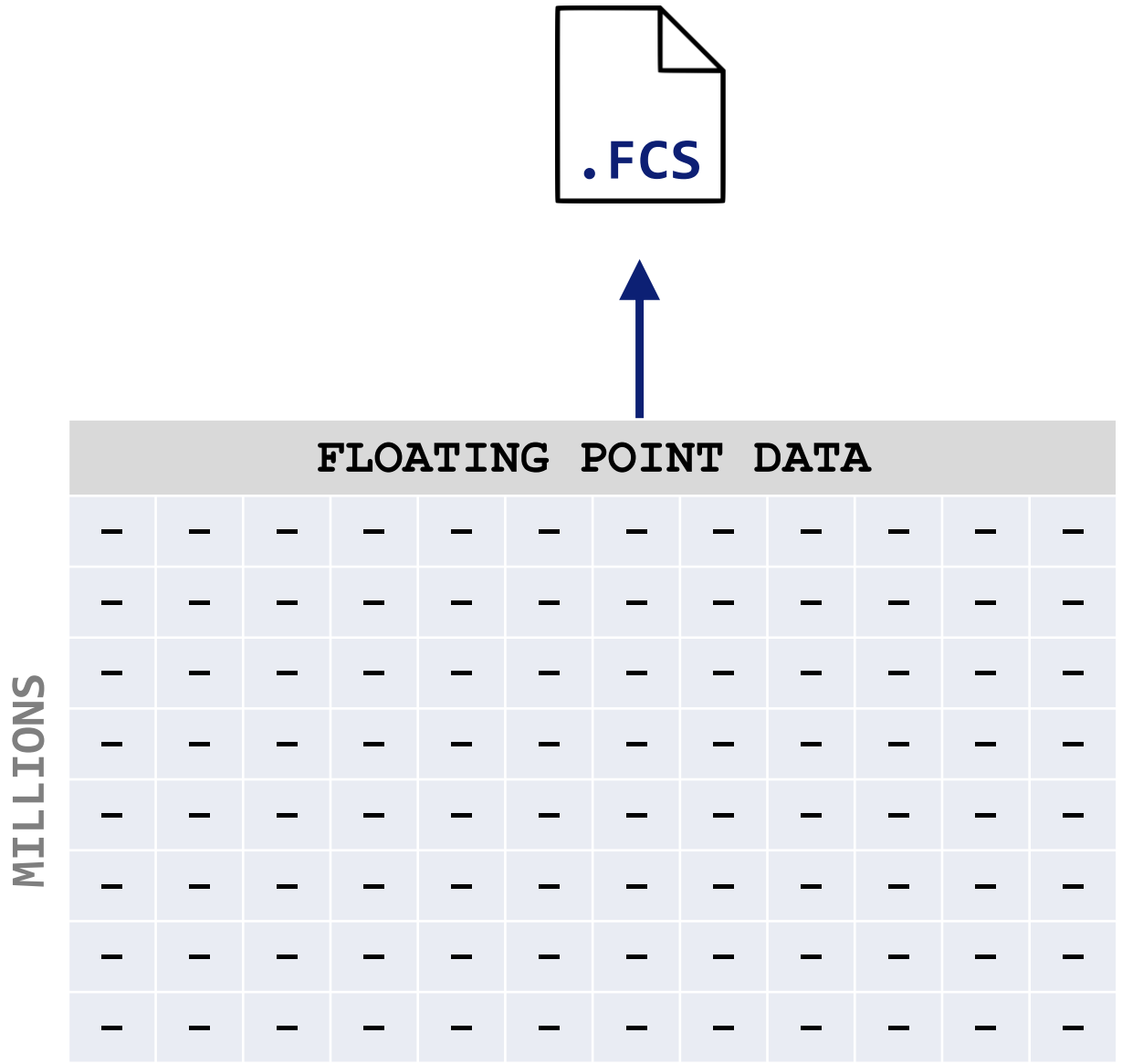
FLOW CYTOMETRY STRANDARD

PRO

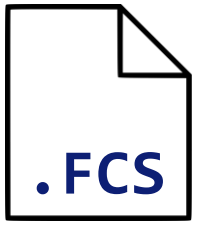
WIDELEY USED

CON

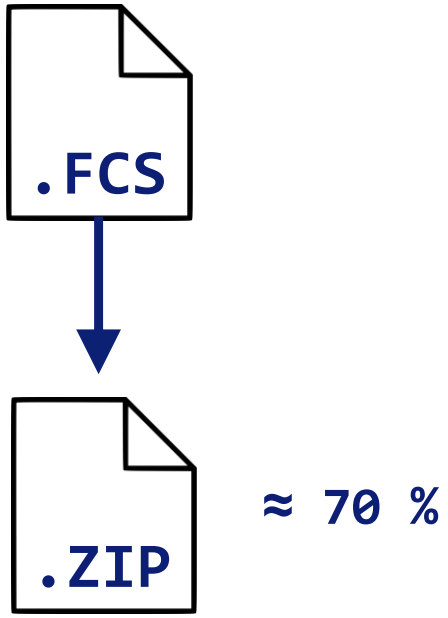
NO COMPRESSION



Lossless Compression

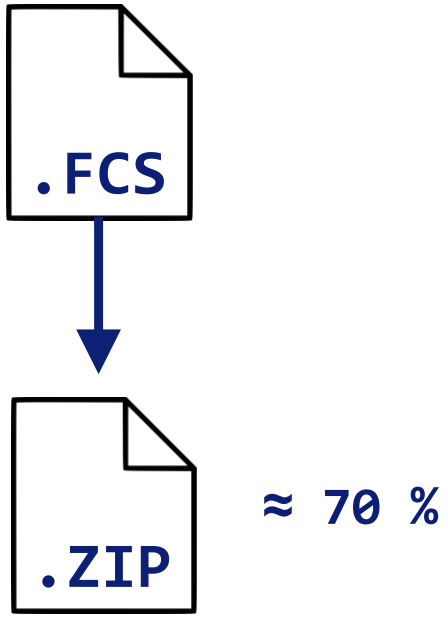


Lossless Compression



≈ 70 %

Lossless Compression



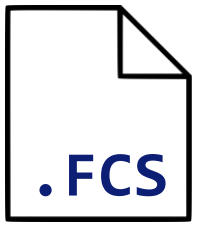
≈ 70 %

Better alternative?

Lossless Compression - Benchmark

Erasmus MC
University Medical Center Rotterdam

Lossless Compression - Benchmark

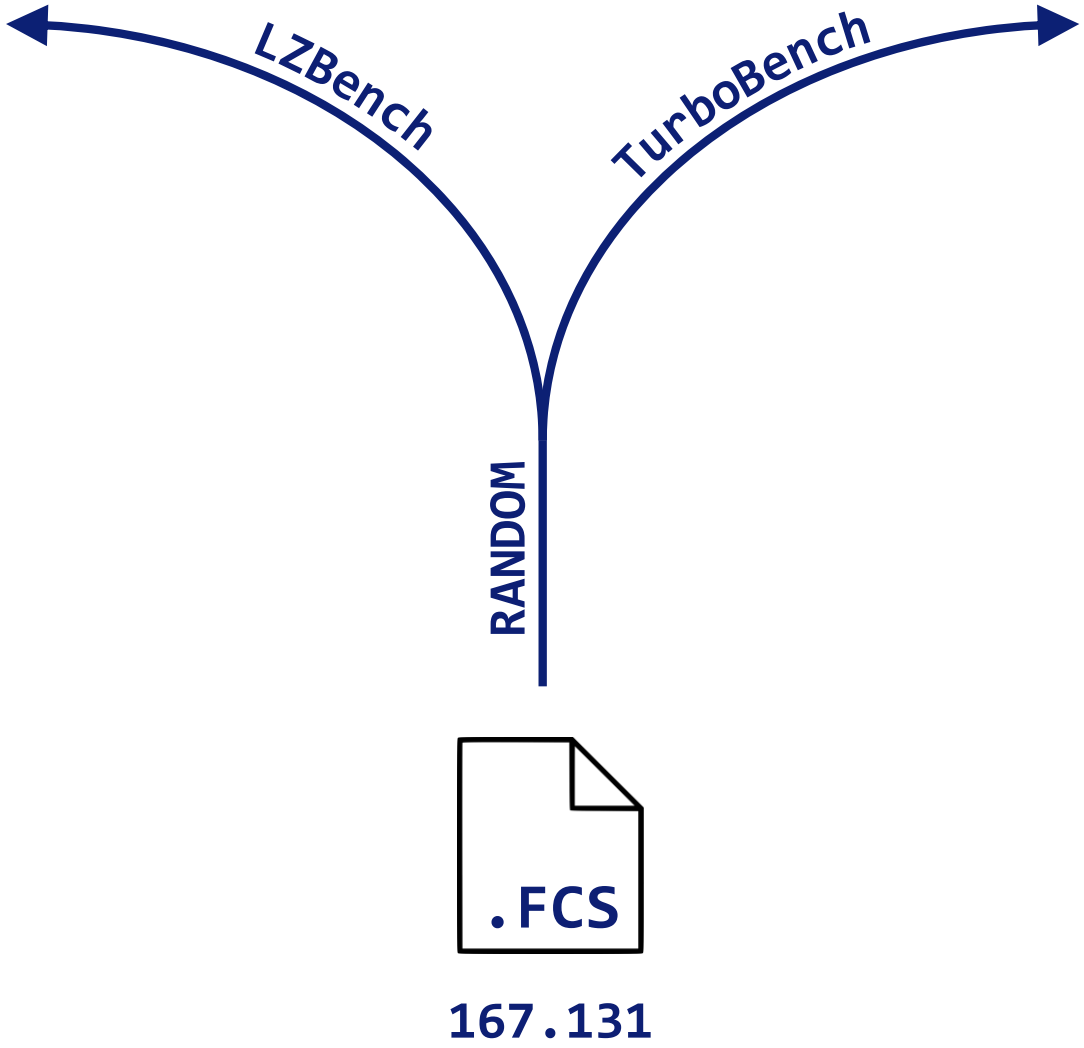


167.131

Lossless Compression - Benchmark

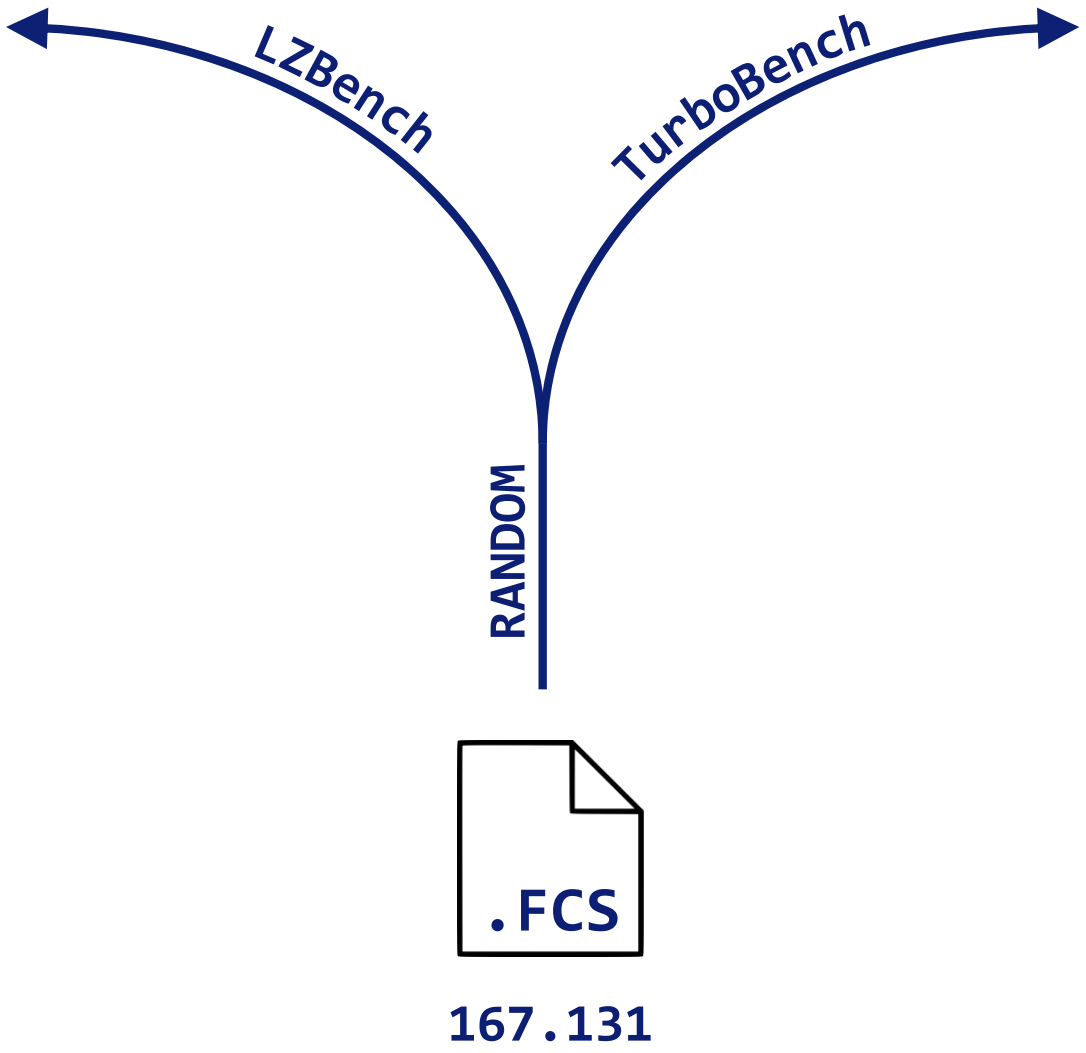
167.131

Lossless Compression - Benchmark



Lossless Compression - Benchmark

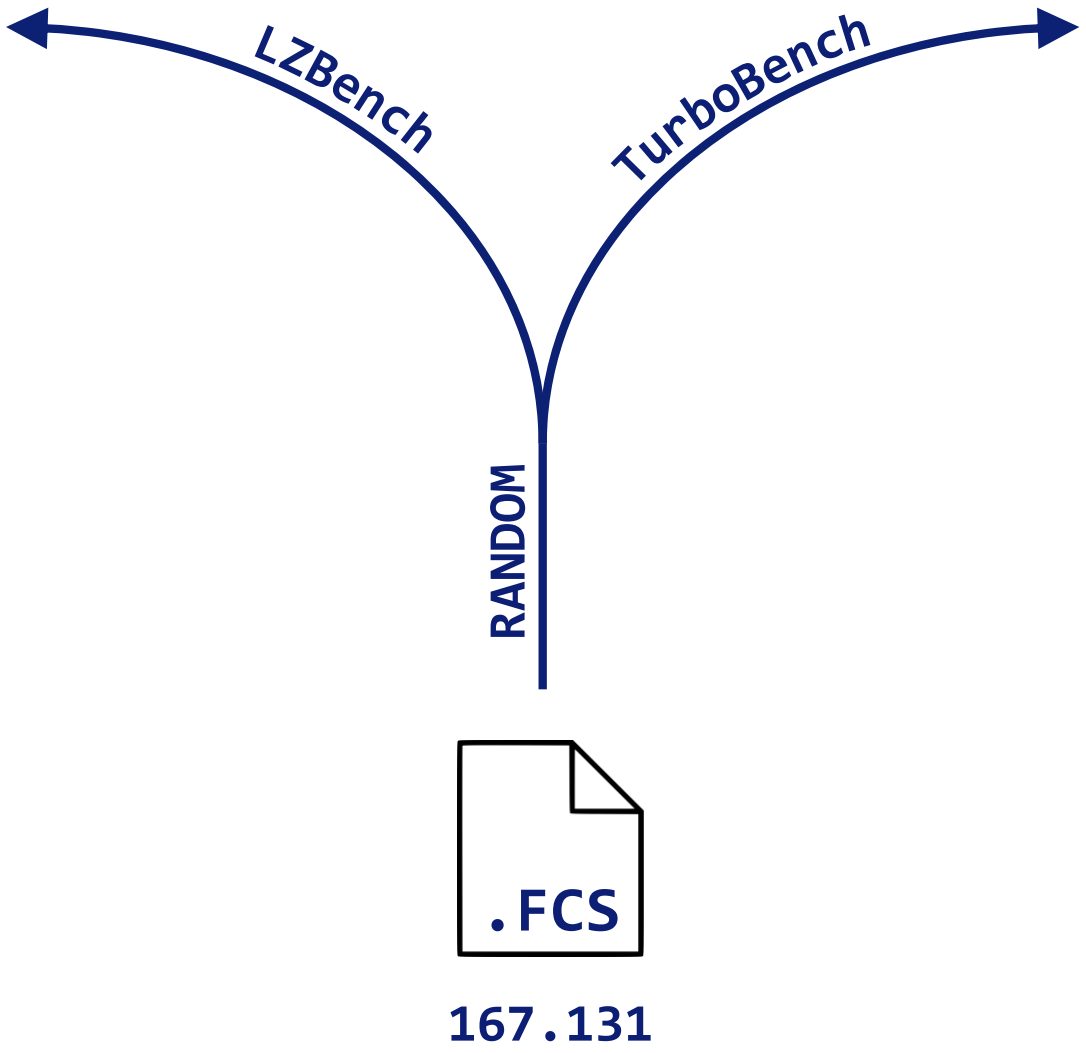
CODEC	RATIO
LZIP	0.533
LZMA	0.533
XZ	0.533
GLZA	0.558
LZHAM	0.568
CSC	0.590
BROTLI	0.598
TORNADO	0.622
ZSTD	0.626
XPACK	0.637
ZLING	0.665
LIBDEFLATE	0.676
LZFSE	0.690
CRUSH	0.692
ZLIB	0.695
UCL_NRV2D	0.725
UCL_NRV2E	0.725
UCL_NRV2B	0.734
LZO1X	0.755
LZO1Z	0.757
LZSSE8	0.758
LZO1Y	0.768
LZSSE2	0.768
LIZARD	0.770
DENSITY	0.771
LZO2A	0.774
LZO1B	0.775
LZ4HC	0.777
...	...



CODEC	RATIO
ZPAQ	0.460
BCM	0.510
LZIP	0.533
LZMA	0.533
FLZMA2	0.539
LZHAM	0.569
BROTLI	0.571
BZIP2	0.580
CSC	0.590
BALZ	0.598
XPACK	0.612
ZSTD	0.625
ZOPFLI	0.674
LIBDEFLATE	0.676
LZFSE	0.690
CRUSH	0.692
ZLIB	0.695
BRIEFLZ	0.727
DOBOZ	0.750
LZSSE8	0.758
BSCQLFC	0.763
LZSSE2	0.768
LZ4	0.777
LZSSE4	0.781
LZG	0.803
SUBOTIN	0.819
FASTAC	0.819
ZLIBH	0.823
...	...

Lossless Compression - Benchmark

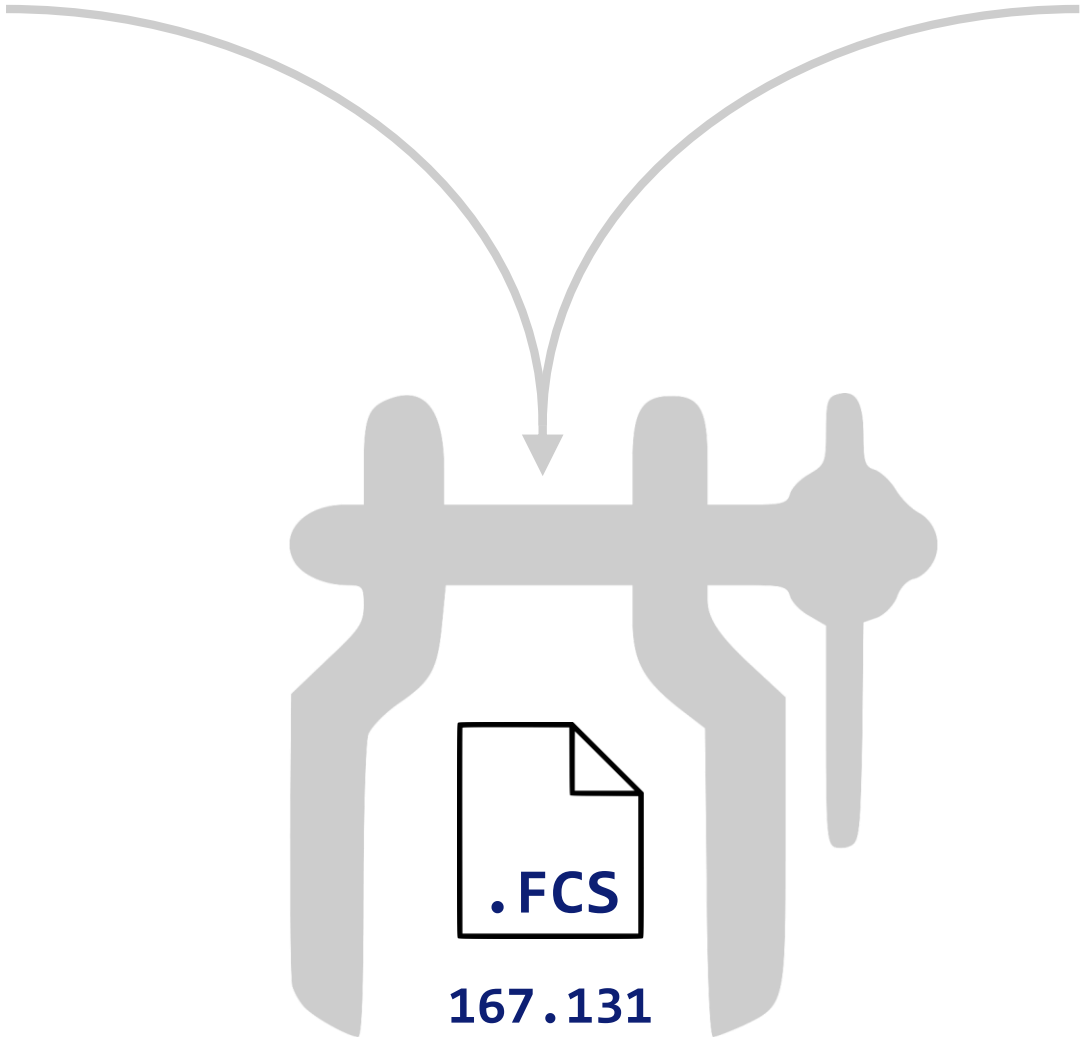
CODEC	RATIO
LZIP	0.533
LZMA	0.533
XZ	0.533
GLZA	0.558
LZHAM	0.568
CSC	0.590
BROTLI	0.598
TORNADO	0.622
ZSTD	0.626
XPACK	0.637
ZLING	0.665
LIBDEFLATE	0.676
LZFSE	0.690
CRUSH	0.692
ZLIB	0.695
UCL_NRV2D	0.725
UCL_NRV2E	0.725
UCL_NRV2B	0.734
LZO1X	0.755
LZO1Z	0.757
LZSSE8	0.758
LZO1Y	0.768
LZSSE2	0.768
LIZARD	0.770
DENSITY	0.771
LZO2A	0.774
LZO1B	0.775
LZ4HC	0.777
...	...



CODEC	RATIO
ZPAQ	0.460
BCM	0.510
LZIP	0.533
LZMA	0.533
FLZMA2	0.539
LZHAM	0.569
BROTLI	0.571
BZIP2	0.580
CSC	0.590
BALZ	0.598
XPACK	0.612
ZSTD	0.625
ZOPFLI	0.674
LIBDEFLATE	0.676
LZFSE	0.690
CRUSH	0.692
ZLIB	0.695
BRIEFLZ	0.727
DOBOZ	0.750
LZSSE8	0.758
BSCQLFC	0.763
LZSSE2	0.768
LZ4	0.777
LZSSE4	0.781
LZG	0.803
SUBOTIN	0.819
FASTAC	0.819
ZLIBH	0.823
...	...

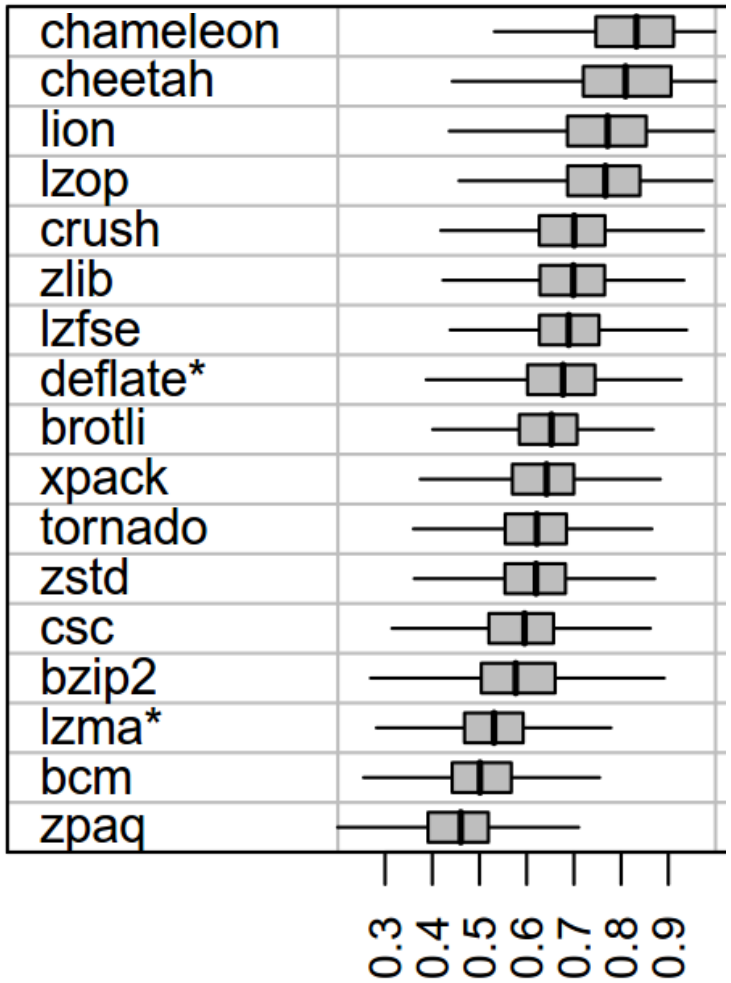
Lossless Compression - Benchmark

CODEC	RATIO
LZIP	0.533
LZMA	0.533
XZ	0.533
GLZA	0.558
LZHAM	0.568
CSC	0.590
BROTLI	0.598
TORNADO	0.622
ZSTD	0.626
XPACK	0.637
ZLING	0.665
LIBDEFLATE	0.676
LZFSE	0.690
CRUSH	0.692
ZLIB	0.695
UCL_NRV2D	0.725
UCL_NRV2E	0.725
UCL_NRV2B	0.734
LZO1X	0.755
LZO1Z	0.757
LZSSE8	0.758
LZO1Y	0.768
LZSSE2	0.768
LIZARD	0.770
DENSITY	0.771
LZO2A	0.774
LZO1B	0.775
LZ4HC	0.777
...	...

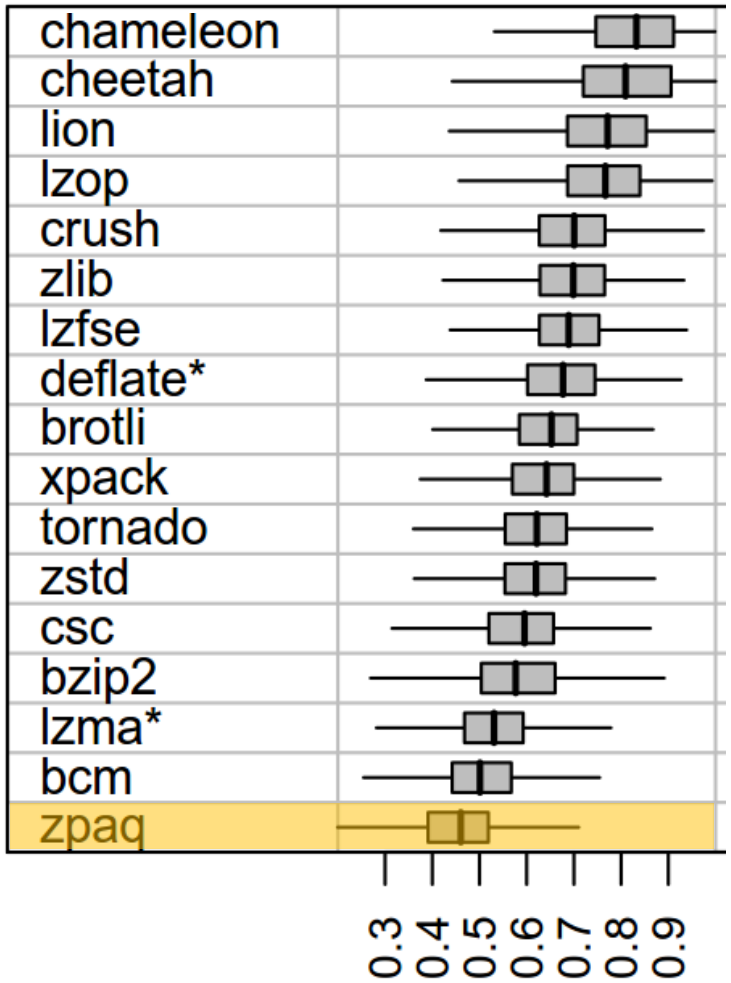


CODEC	RATIO
ZPAQ	0.460
BCM	0.510
LZIP	0.533
LZMA	0.533
FLZMA2	0.539
LZHAM	0.569
BROTLI	0.571
BZIP2	0.580
CSC	0.590
BALZ	0.598
XPACK	0.612
ZSTD	0.625
ZOPFLI	0.674
LIBDEFLATE	0.676
LZFSE	0.690
CRUSH	0.692
ZLIB	0.695
BRIEFLZ	0.727
DOBOZ	0.750
LZSSE8	0.758
BSCQLFC	0.763
LZSSE2	0.768
LZ4	0.777
LZSSE4	0.781
LZG	0.803
SUBOTIN	0.819
FASTAC	0.819
ZLIBH	0.823
...	...

Lossless Compression - Benchmark

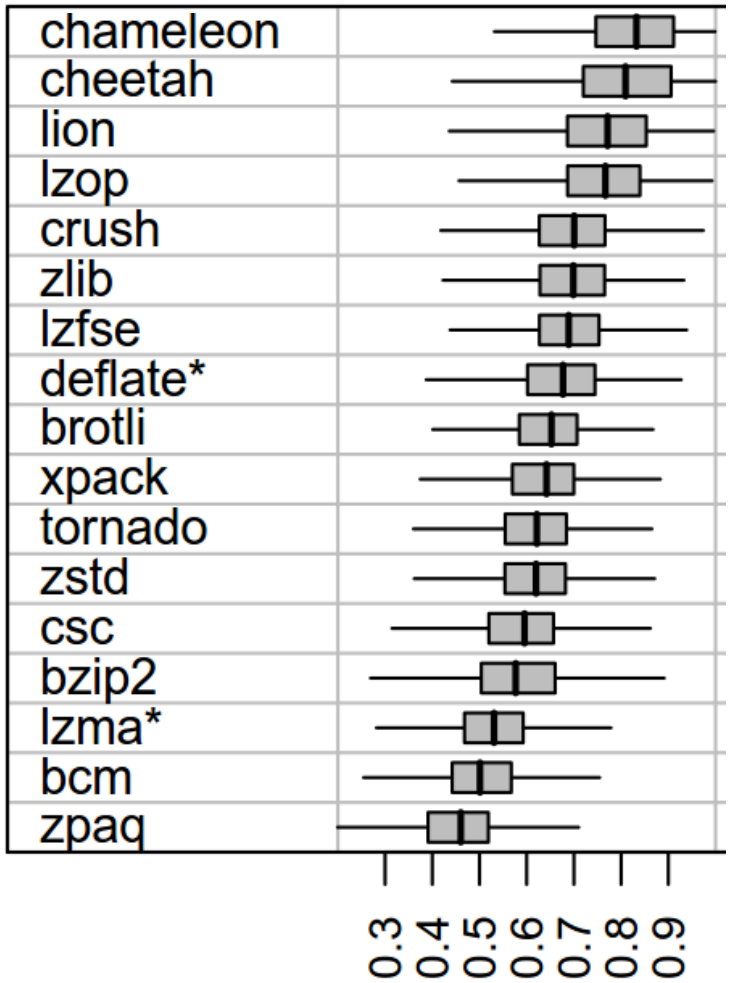


Lossless Compression - Benchmark

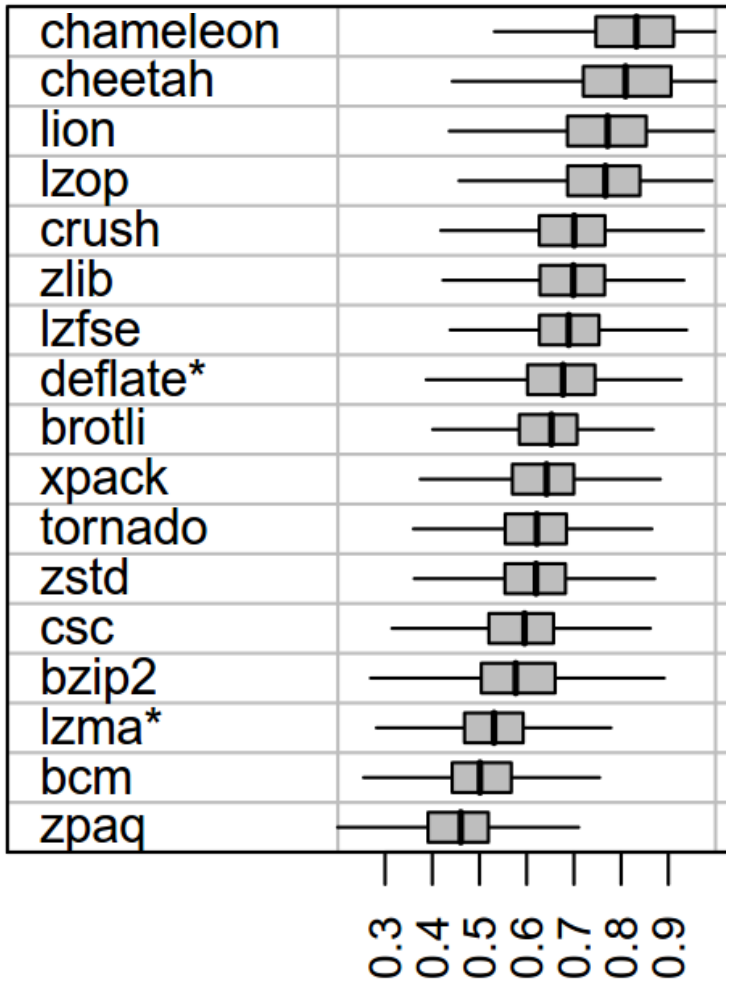


≈ 0.45

Implementation in R

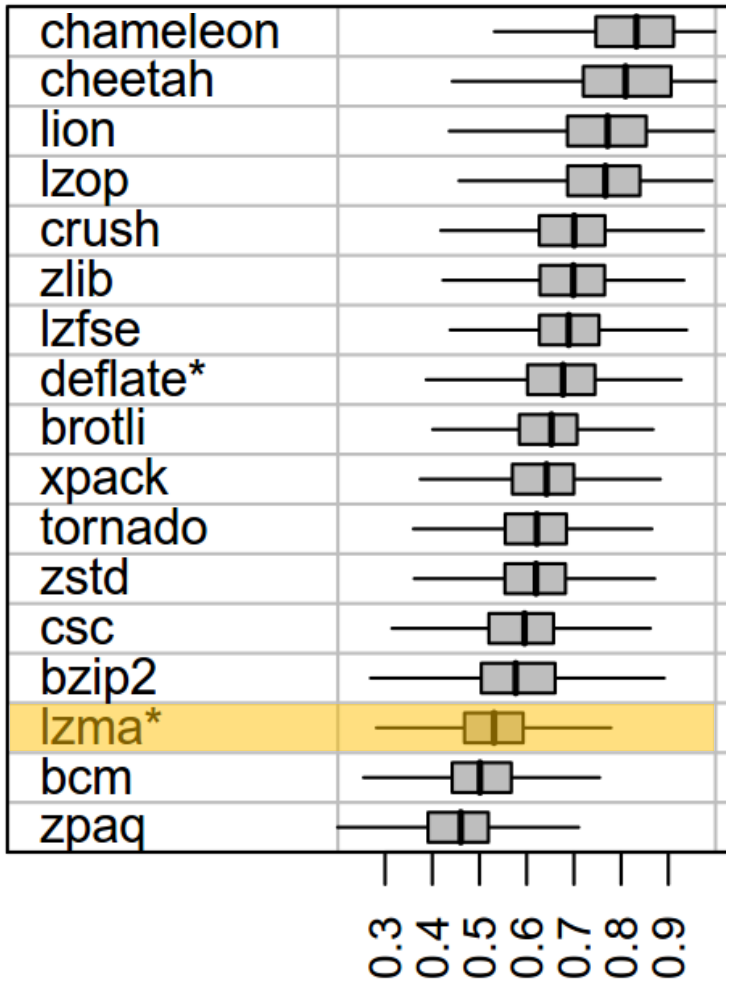


Implementation in R




```
base::memCompress(type="XZ")
```

Implementation in R



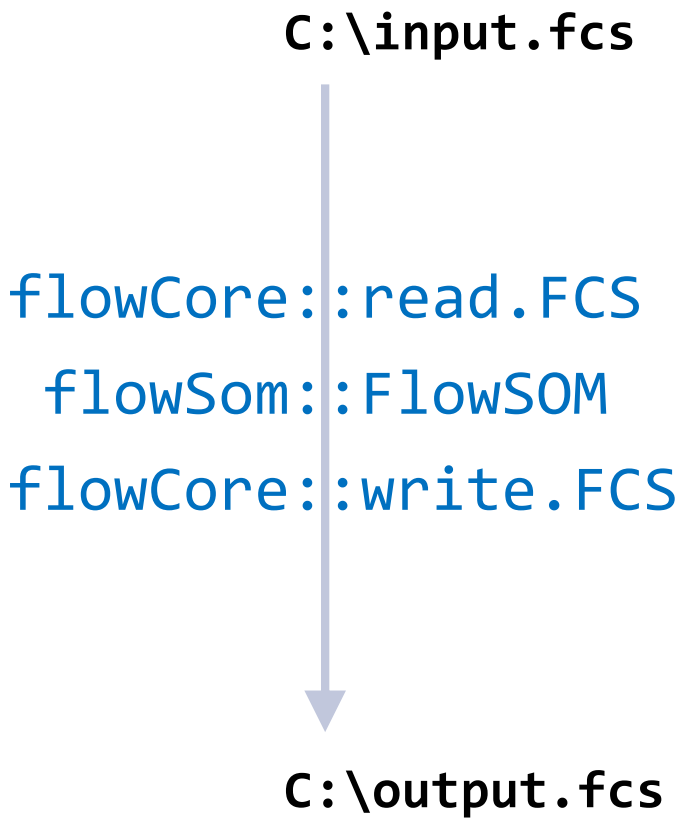

```
base::memCompress(type="XZ")
```

Bioconductor - flowCore

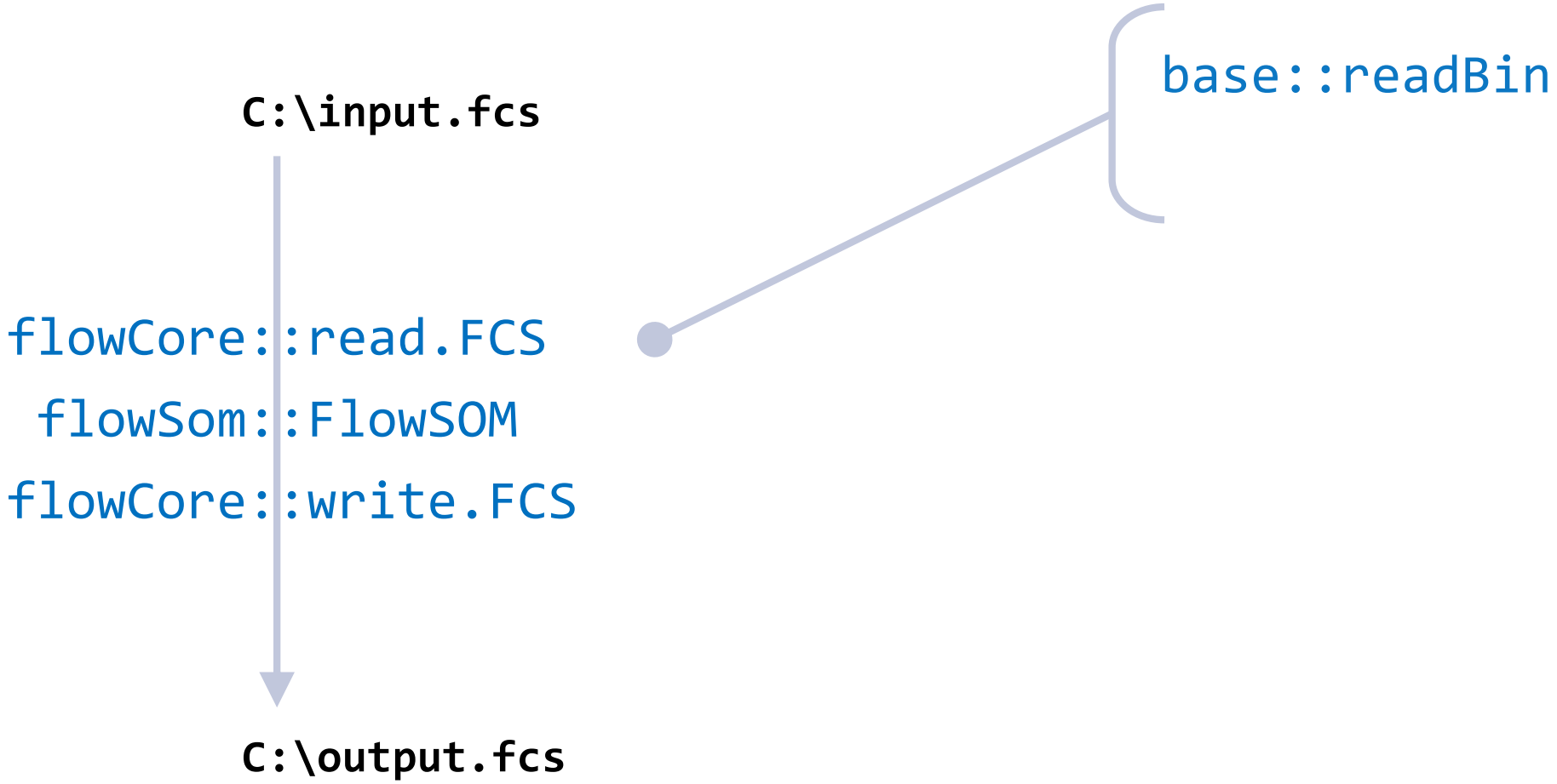
Bioconductor - flowCore

NO COMPRESSION

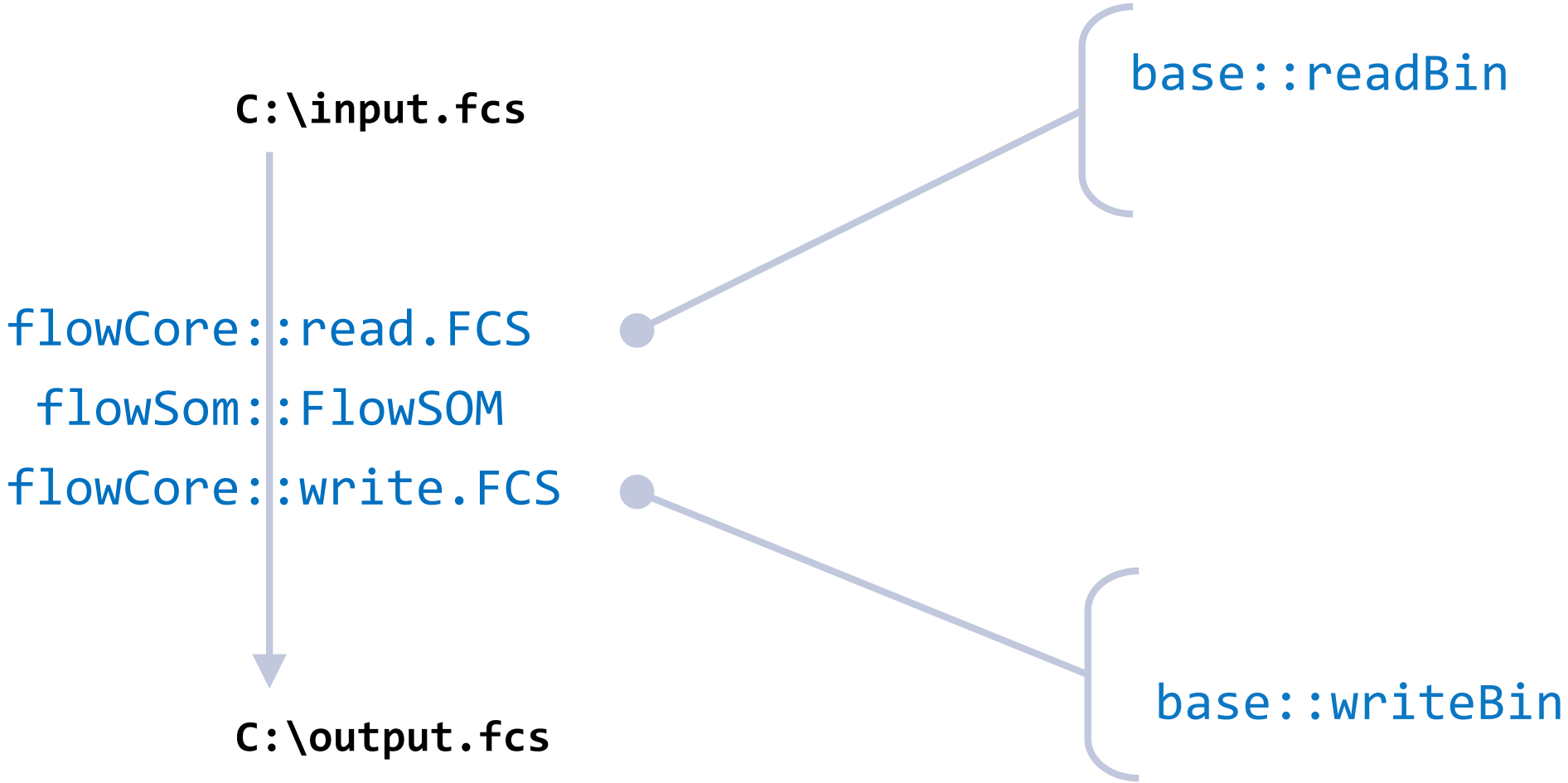
Bioconductor - flowCore



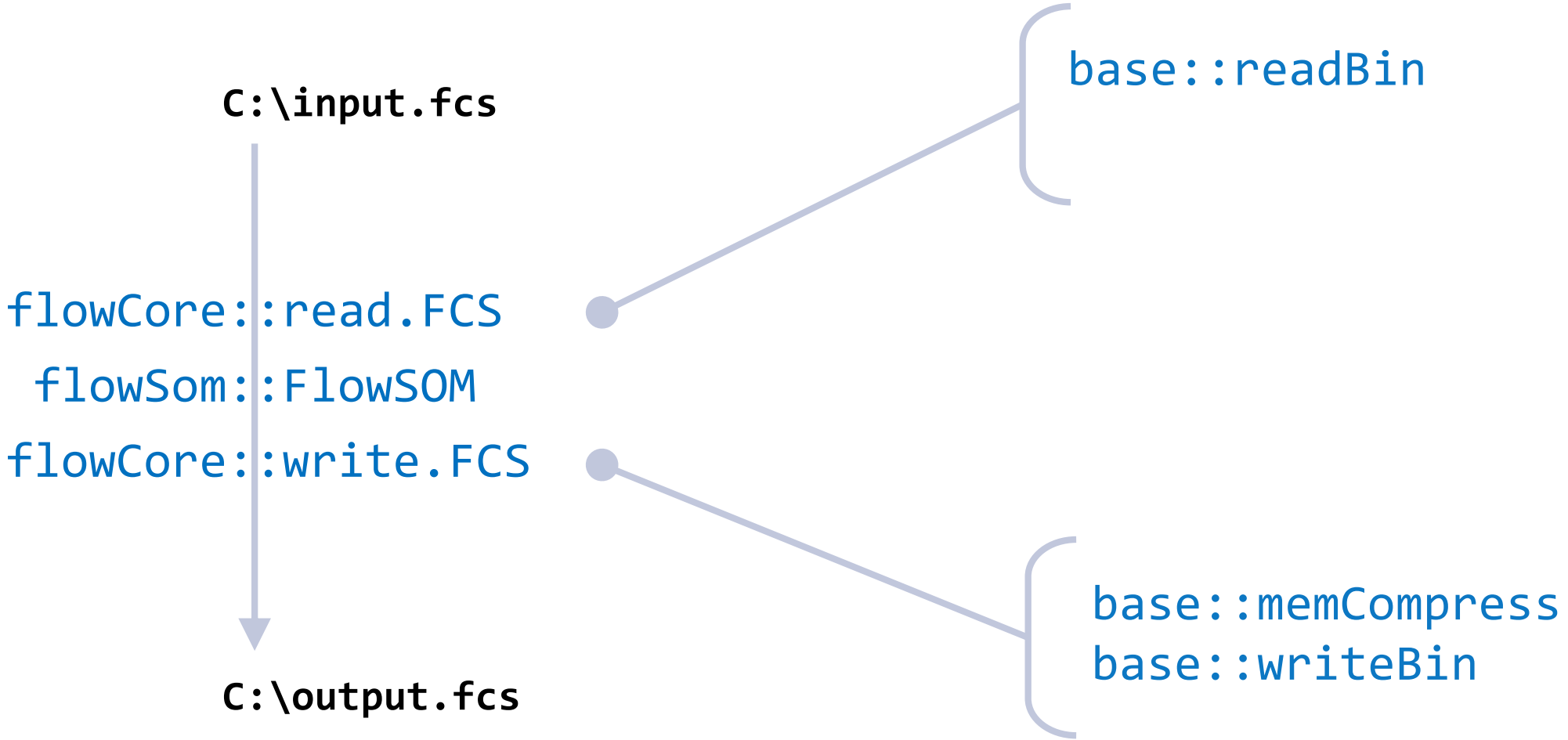
Bioconductor - flowCore



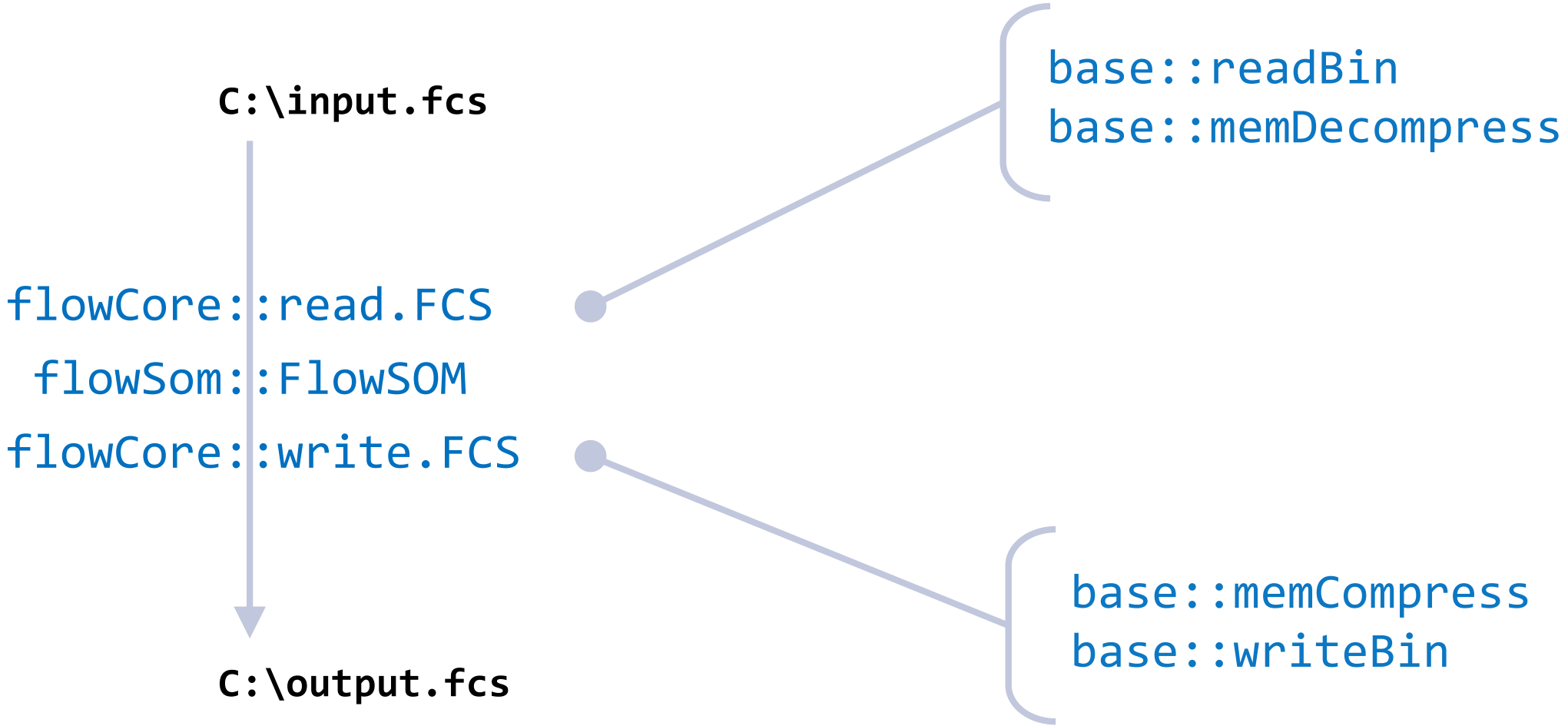
Bioconductor - flowCore



Bioconductor - flowCore



Bioconductor - flowCore



Bioconductor - flowCore

C:\input.fcs

flowCore::read.FCS

flowSom::FlowSOM

flowCore::write.FCS

C:\output.fcs

Bioconductor - flowCore

C:\input.fcs

flowCore::read.FCS

flowSom::FlowSOM

flowCore::write.FCS

C:\output.fcs

Bioconductor - flowCore

C:\input.fcs

flowCore::read.FCS

flowSom::FlowSOM

flowCore::write.FCS

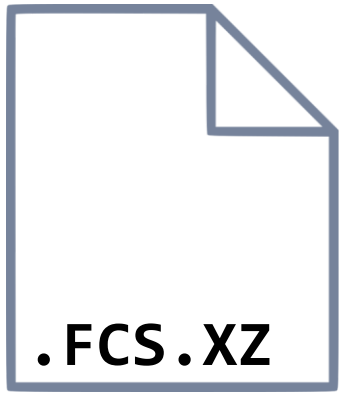
C:\output.fcs

Bioconductor - flowCore

C:\input.fcs

```
flowCore::read.FCS
 flowSom::FlowSOM
flowCore::write.FCS
```

C:\output.fcs



Bioconductor - flowCore

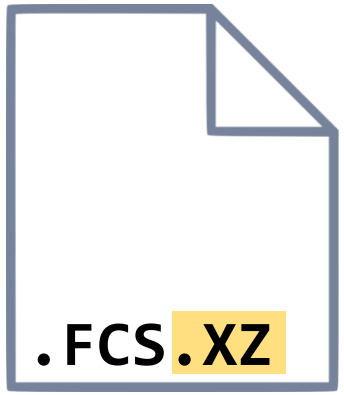
C:\input.fcs

flowCore::read.FCS

flowSom::FlowSOM

flowCore::write.FCS

C:\output.fcs



Bioconductor - flowCore

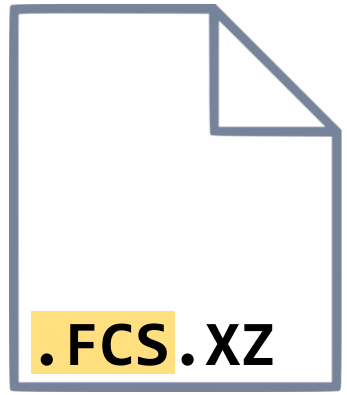
C:\input.fcs

flowCore::read.FCS

flowSom::FlowSOM

flowCore::write.FCS

C:\output.fcs

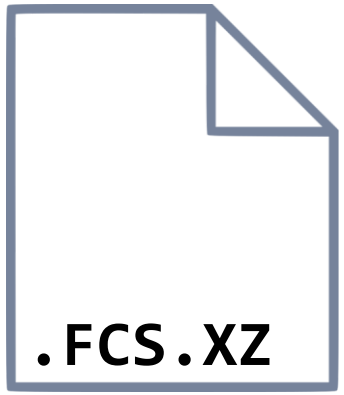


Bioconductor - flowCore

C:\input.fcs

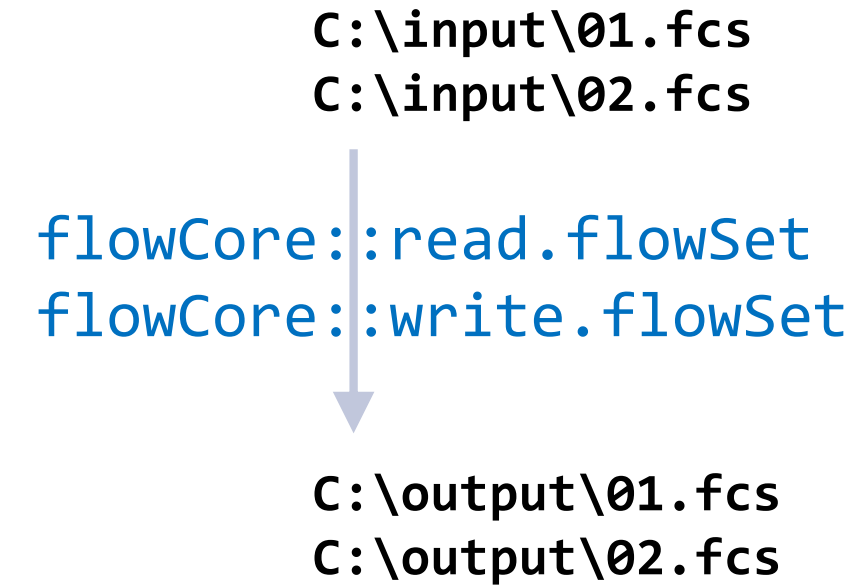
```
flowCore::read.FCS
 flowSom::FlowSOM
flowCore::write.FCS
```

C:\output.fcs

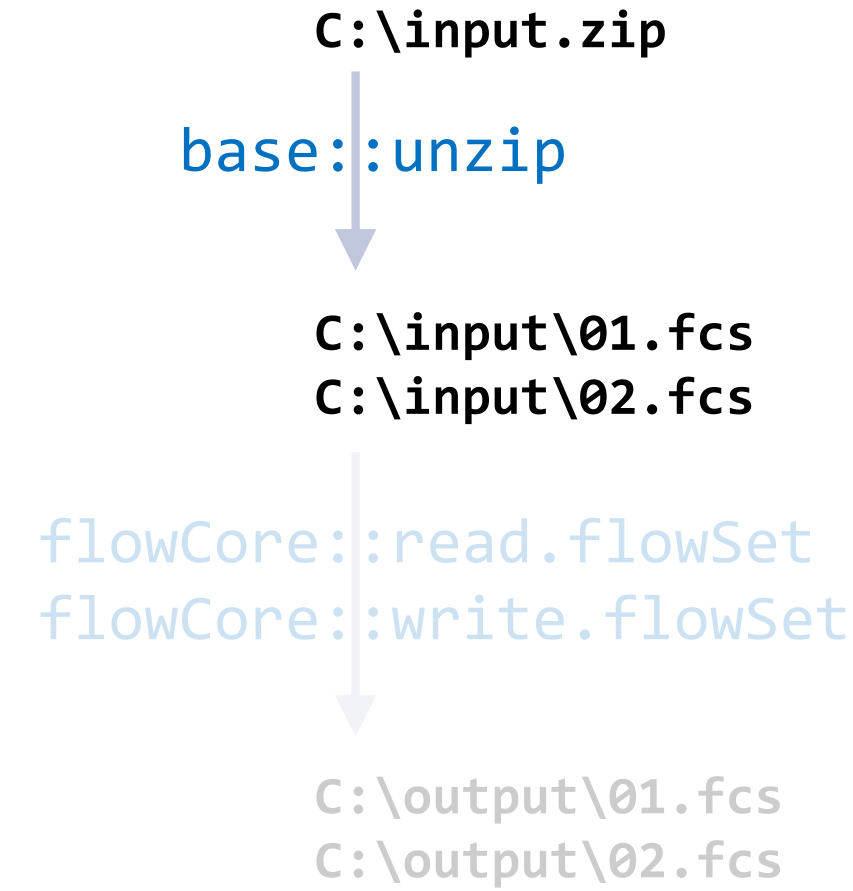


Bioconductor – flowCore - Pipelines

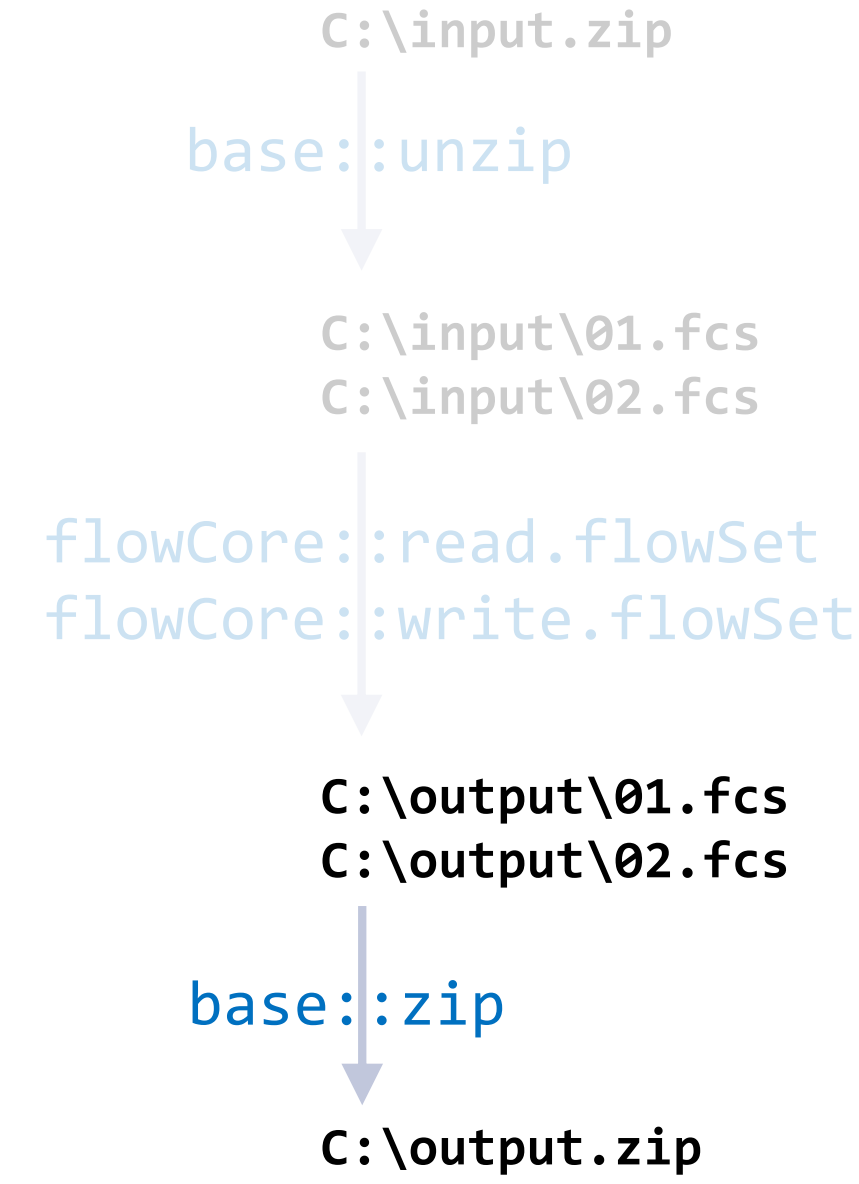
Bioconductor – flowCore - Pipeline



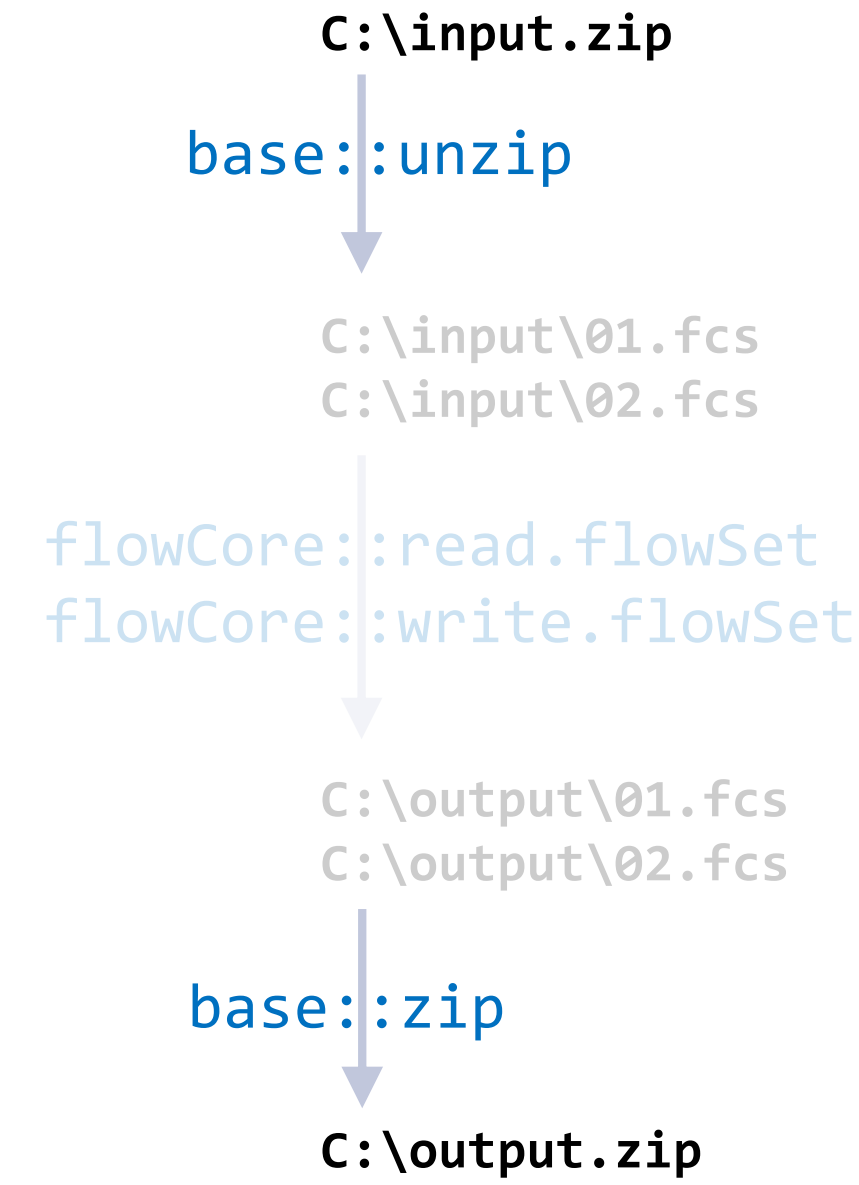
Bioconductor – flowCore - Pipeline



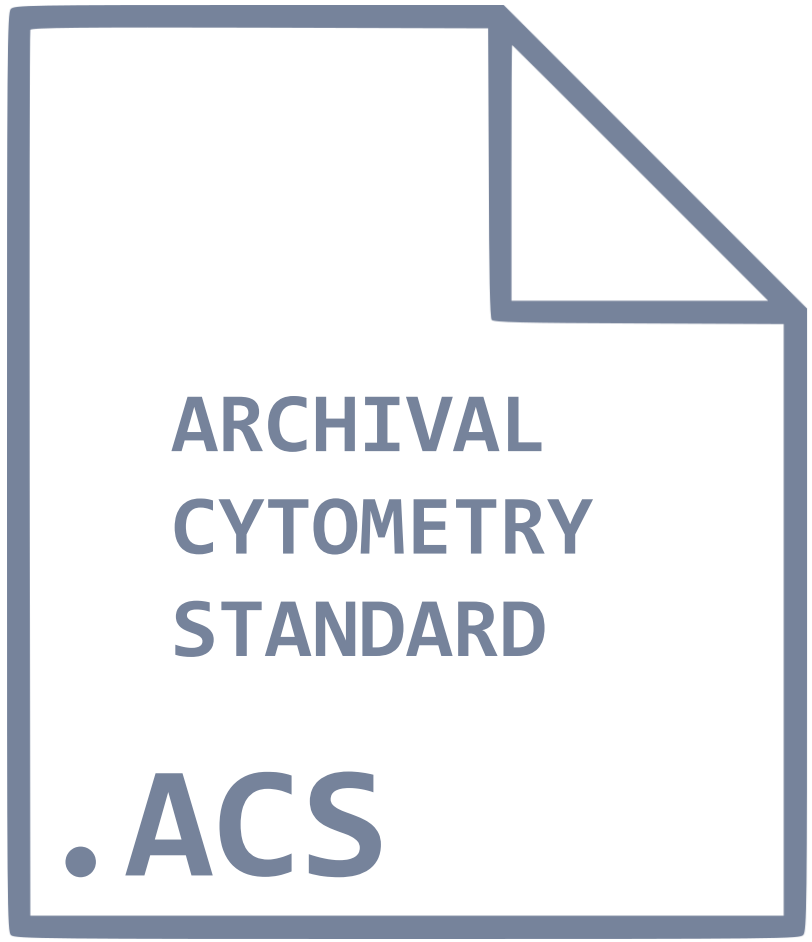
Bioconductor – flowCore - Pipeline



Bioconductor – flowCore - Pipeline



Bioconductor – flowCore - Pipeline



C:\input.zip

base::unzip

C:\input\01.fcs
C:\input\02.fcs

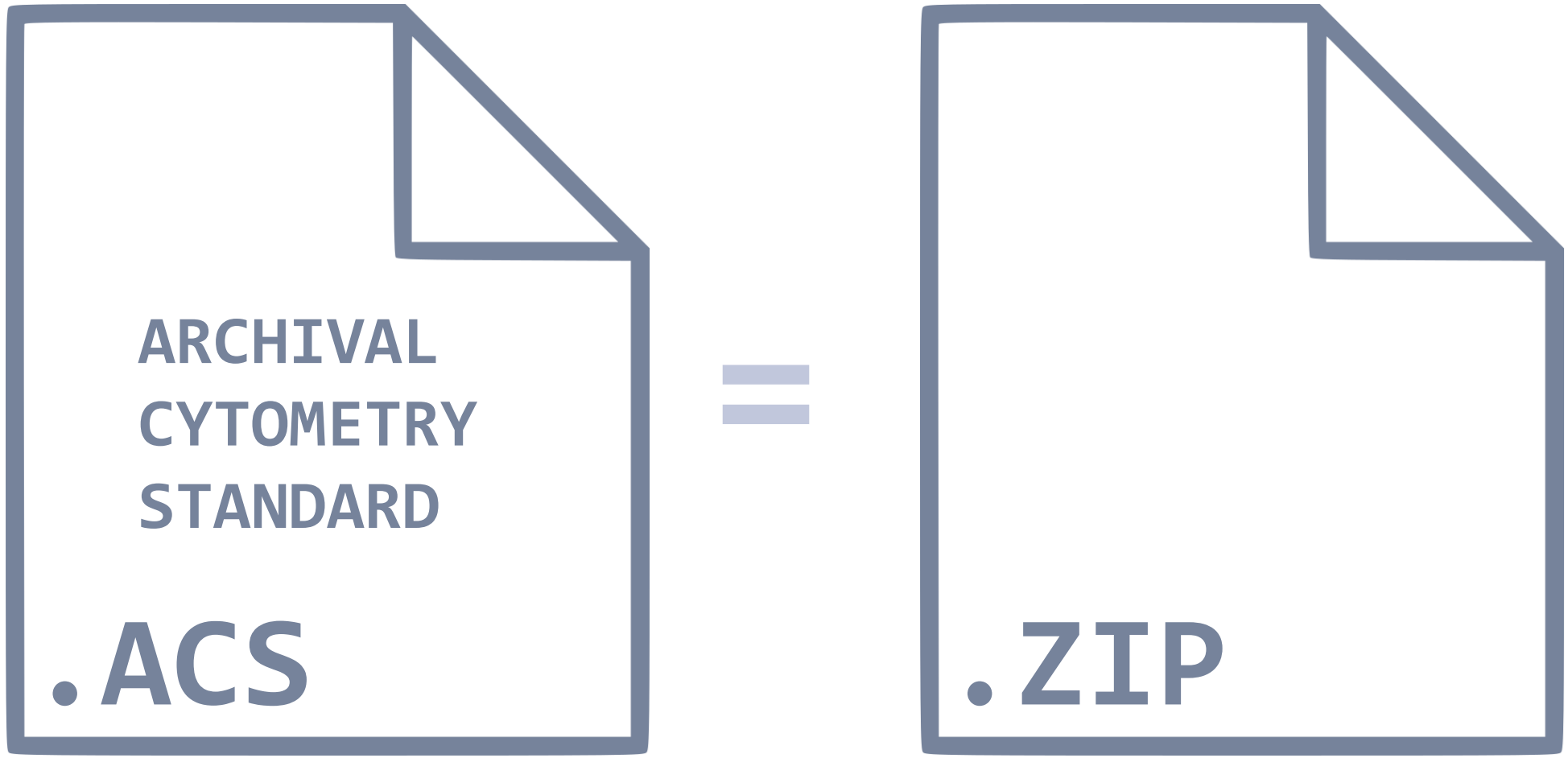
flowCore::read.flowSet
flowCore::write.flowSet

C:\output\01.fcs
C:\output\02.fcs

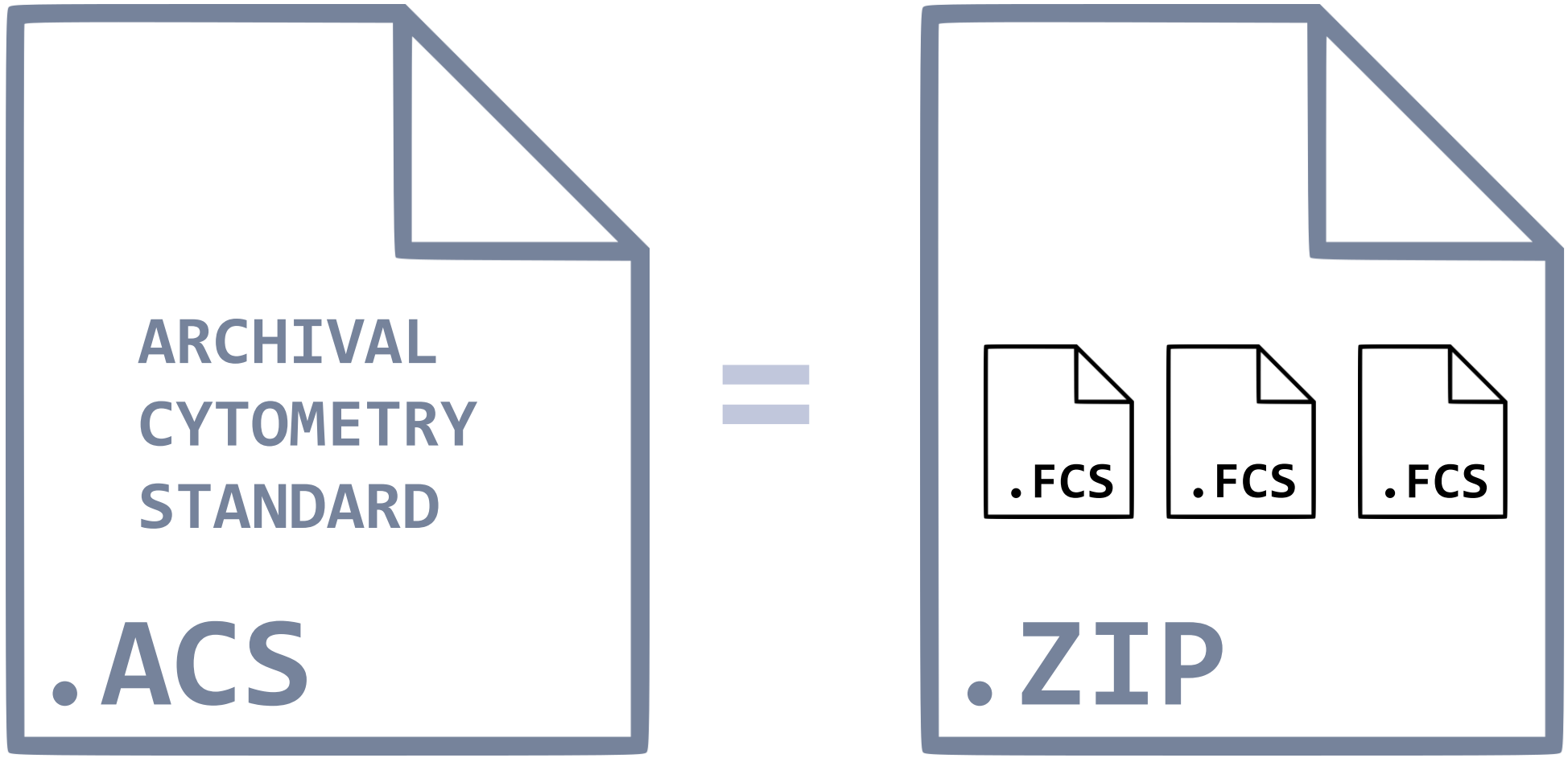
base::zip

C:\output.zip

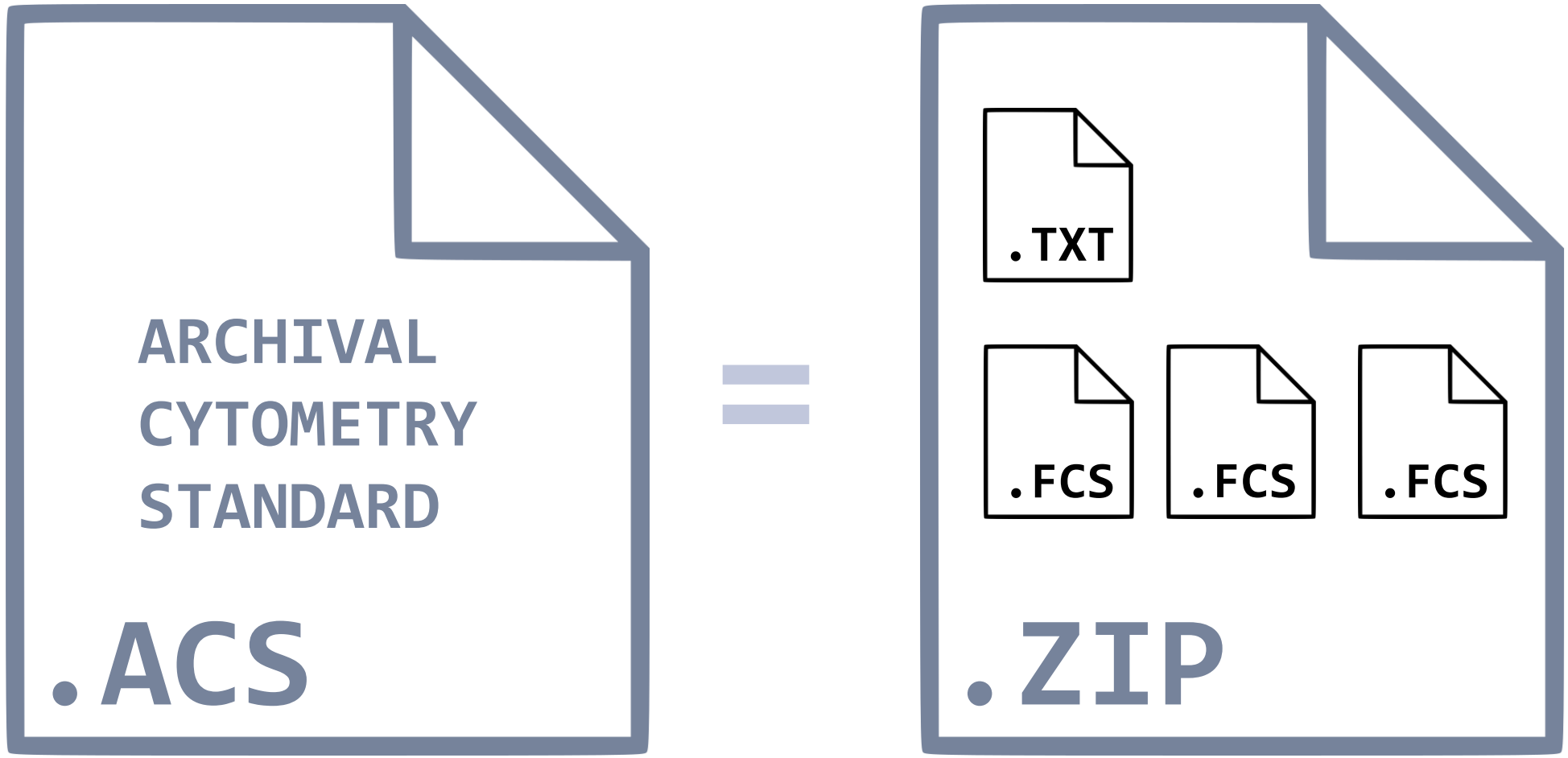
Bioconductor – flowCore - Pipeline



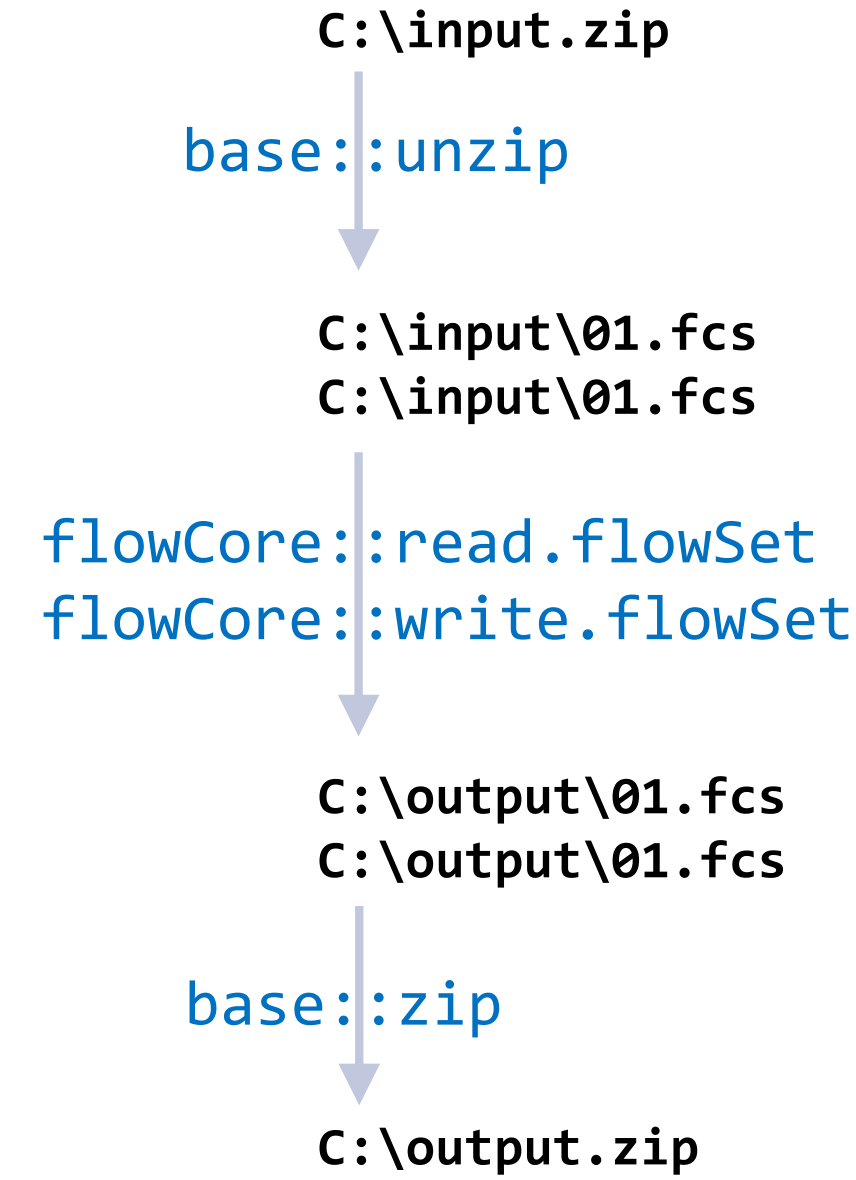
Bioconductor – flowCore - Pipeline



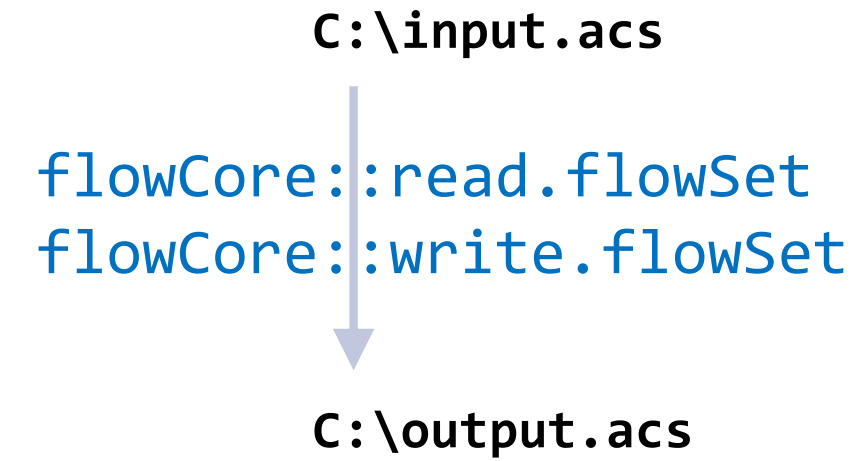
Bioconductor – flowCore - Pipeline



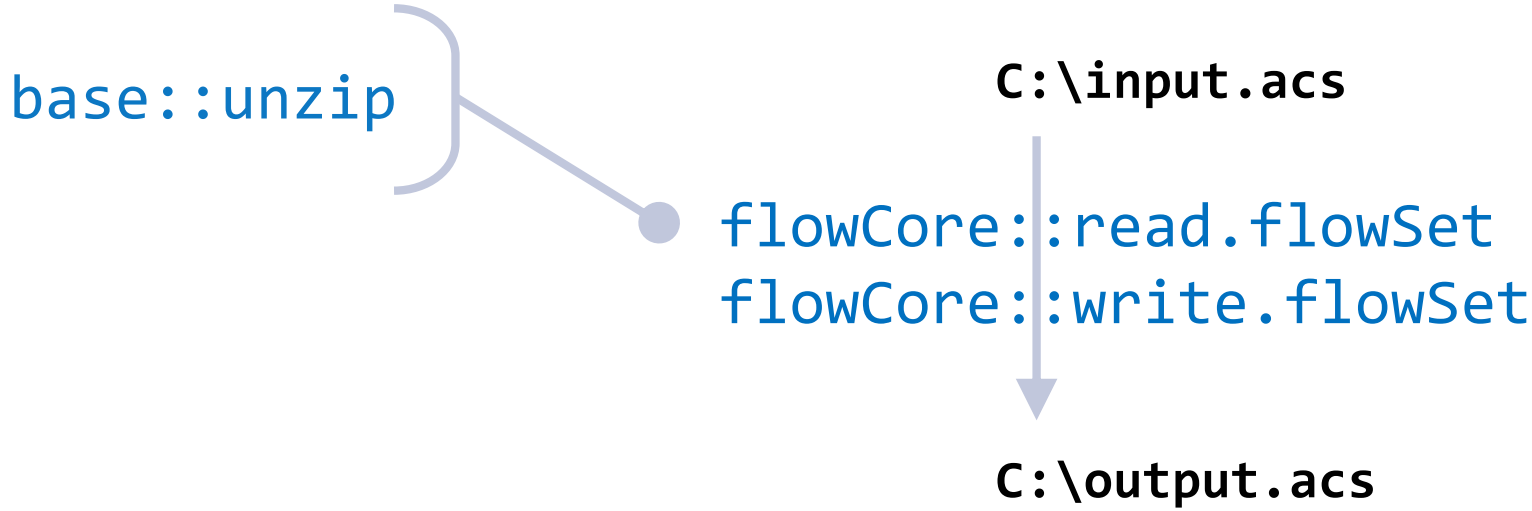
Bioconductor – flowCore - Pipeline



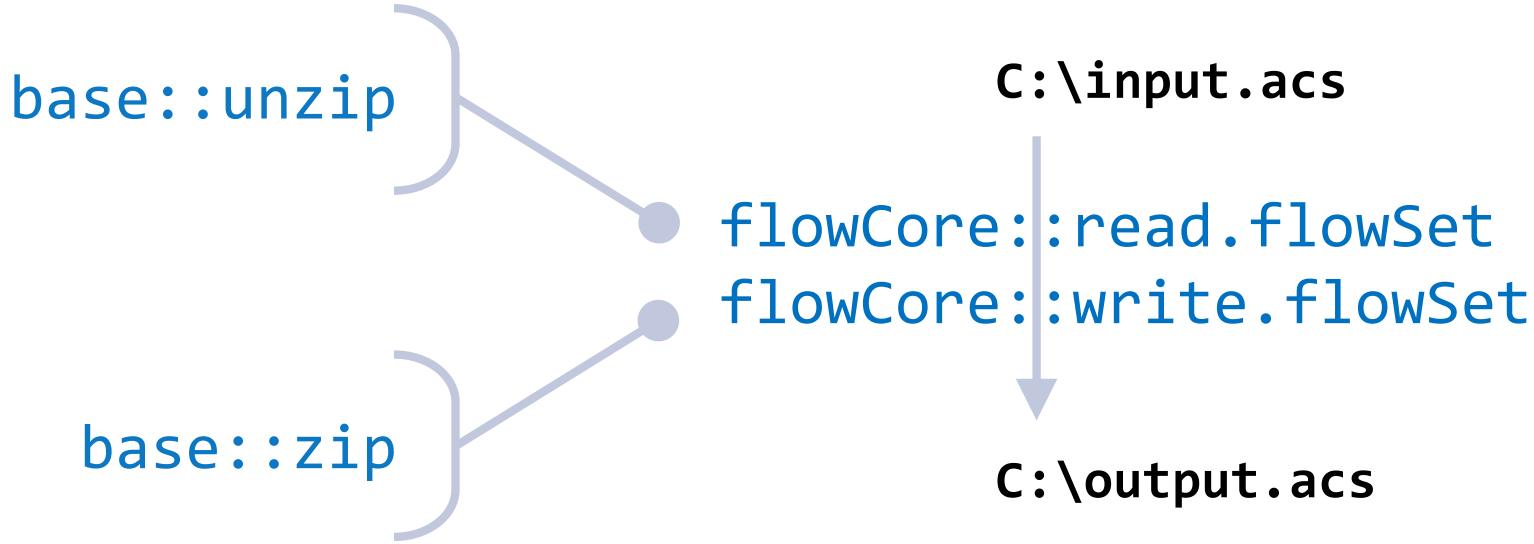
Bioconductor – flowCore - Pipeline



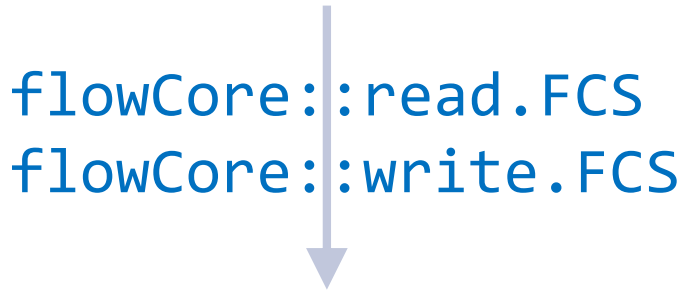
Bioconductor – flowCore - Pipeline



Bioconductor – flowCore - Pipeline

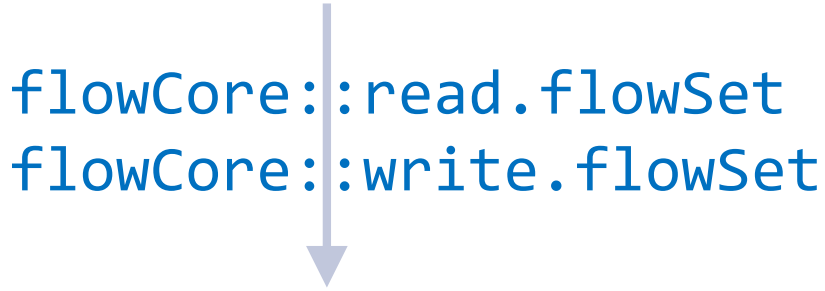


C:\input.fcs.xz



C:\output.fcs.xz

C:\input.acs



C:\output.acs

C:\input.fcs.xz

flowCore::read.FCS
flowCore::write.FCS

C:\output.fcs.xz

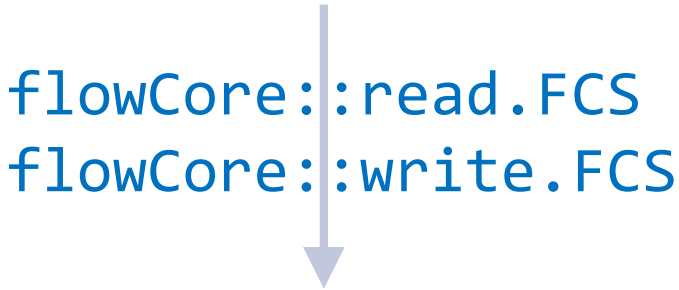
PRO
PERFORMANCE

C:\input.acs

flowCore::read.flowSet
flowCore::write.flowSet

C:\output.acs

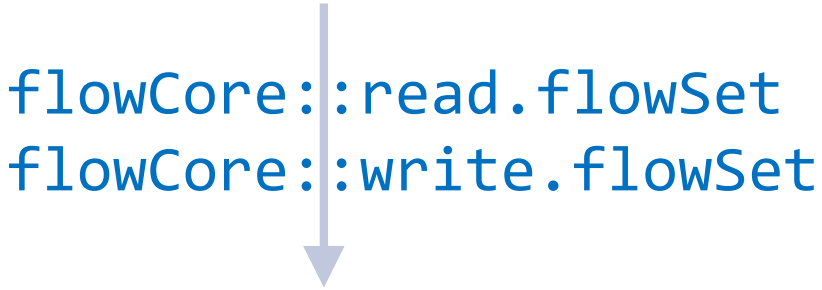
C:\input.fcs.xz



C:\output.fcs.xz

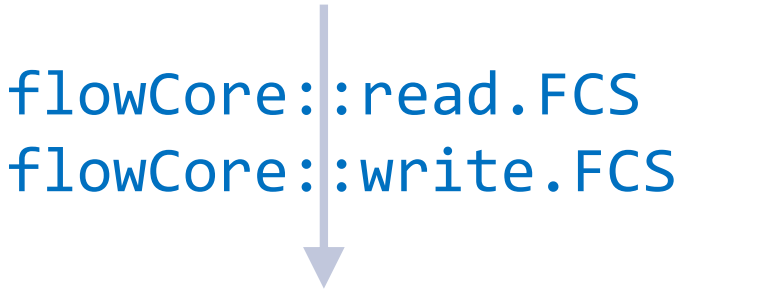
PRO	CON
PERFORMANCE	COMPATIBILITY

C:\input.acs



C:\output.acs

C:\input.fcs.xz

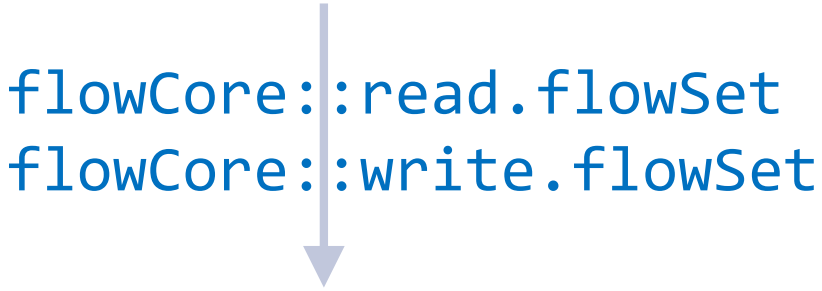


C:\output.fcs.xz

PRO
PERFORMANCE

CON
COMPATIBILITY

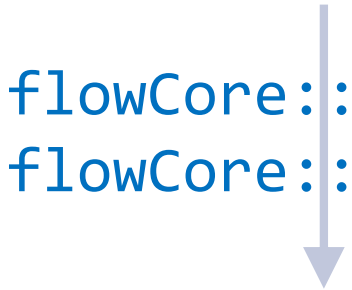
C:\input.acs



C:\output.acs

PRO
COMPATIBILITY

C:\input.fcs.xz

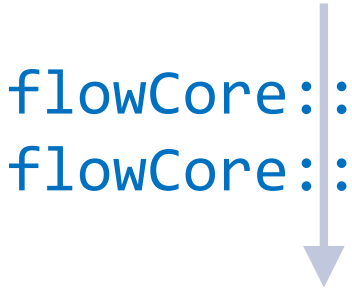


C:\output.fcs.xz

PRO
PERFORMANCE

CON
COMPATIBILITY

C:\input.acs

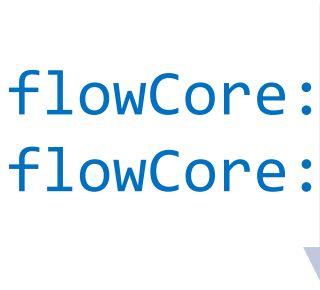


C:\output.acs

PRO
COMPATIBILITY

CON
PERFORMANCE

C:\input.fcs.xz

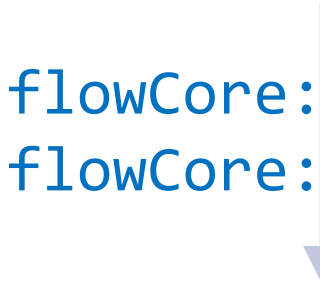


C:\output.fcs.xz

PRO
PERFORMANCE
EASY

CON
COMPATIBILITY

C:\input.acs

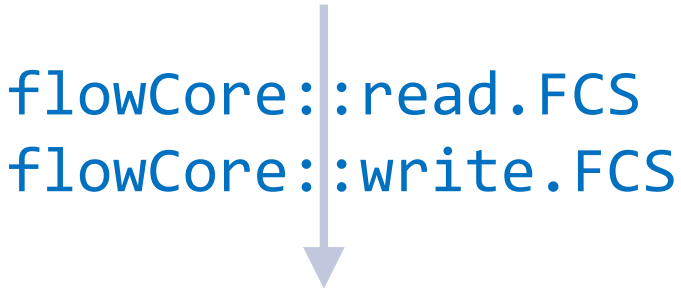


C:\output.acs

PRO
COMPATIBILITY
EASY

CON
PERFORMANCE

C:\input.fcs.xz

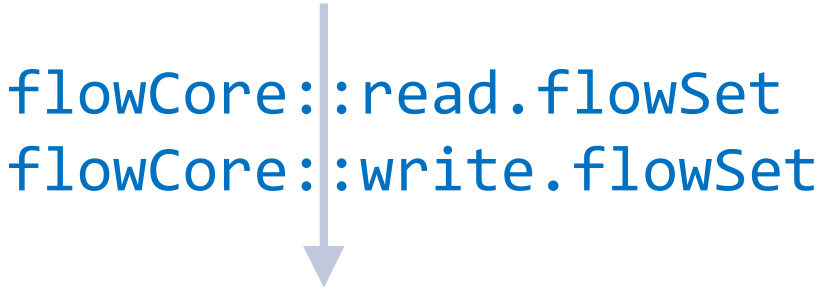


C:\output.fcs.xz

PRO
PERFORMANCE
EASY
BENIFIT

CON
COMPATIBILITY

C:\input.acs



C:\output.acs

PRO
COMPATIBILITY
EASY
BENIFIT

CON
PERFORMANCE

Questions?

Lossless Compression of Cytometric Data

Anne E. Bras & Vincent H. J. van der Velden