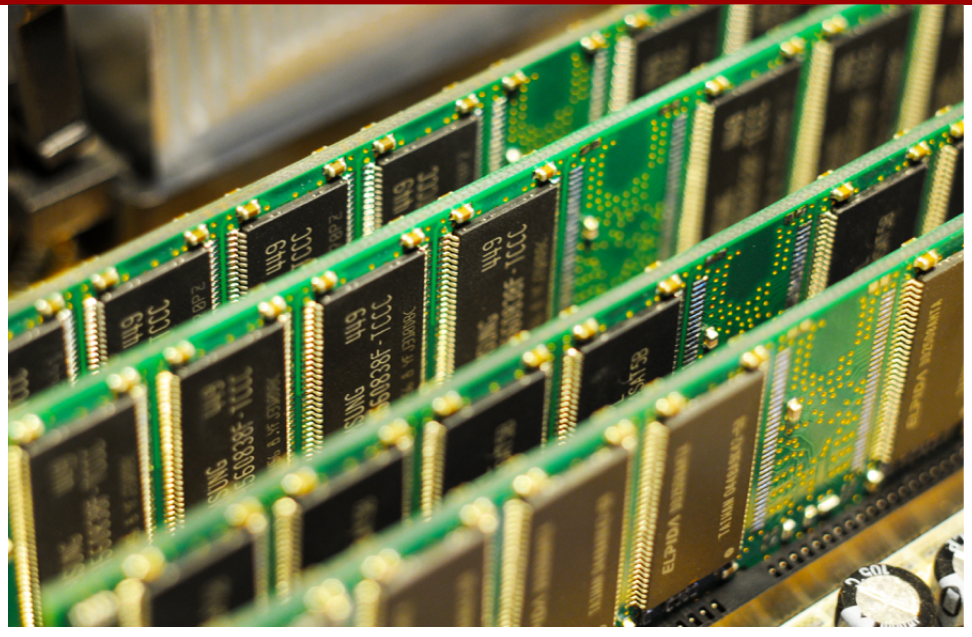
Dual Streaming for Hardware-Accelerated Ray Tracing

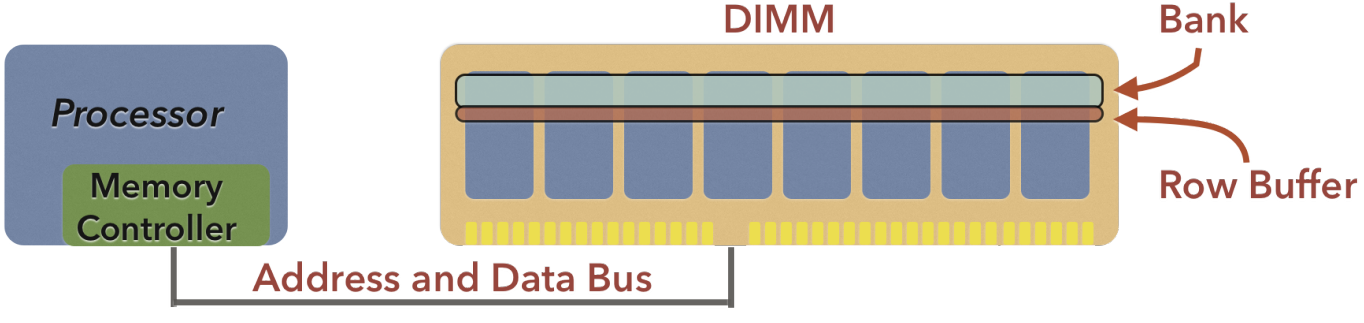
Konstantin Shkurko
Ian Mallett
Tim Grant
Cem Yuksel
Daniel Kopta
Erik Brunvand

University of Utah

HPG '17 7.29.17

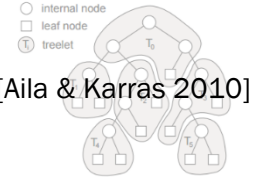
DRAM




DIMM
Bank
Processor
Memory Controller
Row Buffer
Address and Data Bus

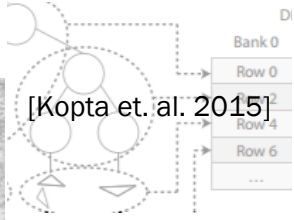
Previous Work

[Boulos et. al. 2007]

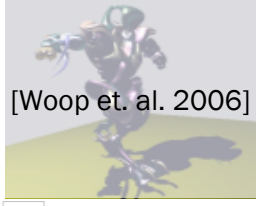


[Aila & Karras 2010]

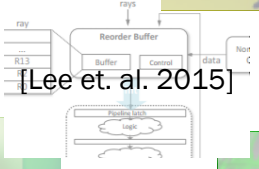
[Navratil et. al. 2007]



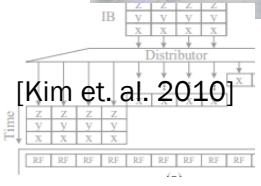
[Kopta et. al. 2015]


[Woop et. al. 2006]

[Wald et. al. 2014]



[Lee et. al. 2015]

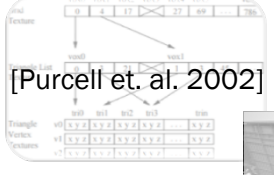


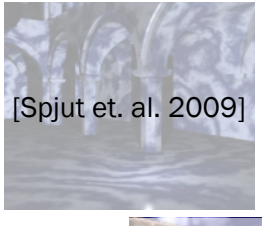
[Kim et. al. 2010]

[Eisenacher et. al. 2013]

[Moon et. al. 2010]

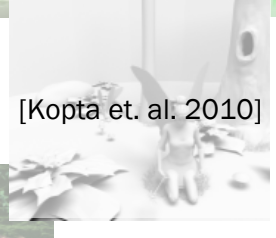
[Bigler et. al. 2006]

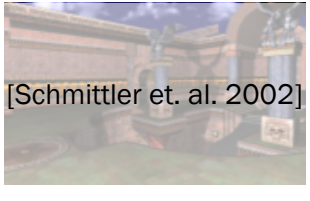

[Purcell et. al. 2002]


[Spjut et. al. 2009]

[Keely 2014]

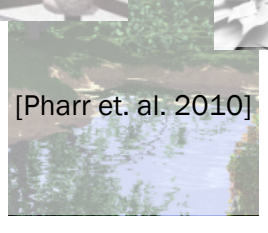
[Gribble & Ramani 2008]

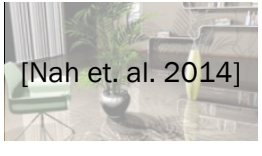

[Kopta et. al. 2010]


[Schmittler et. al. 2002]

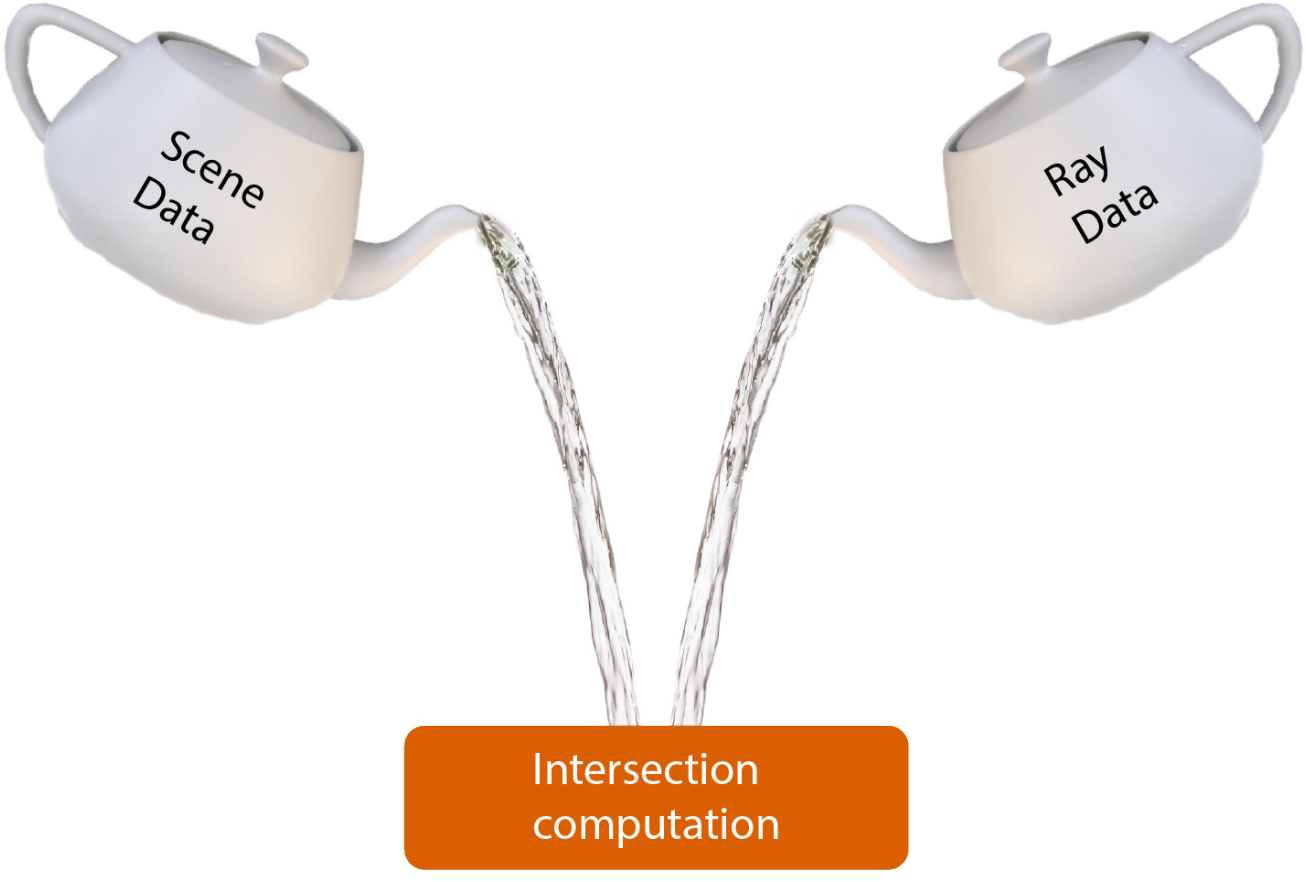
[Bikker 2012]

[Liktor & Vaidyanathan 2016]

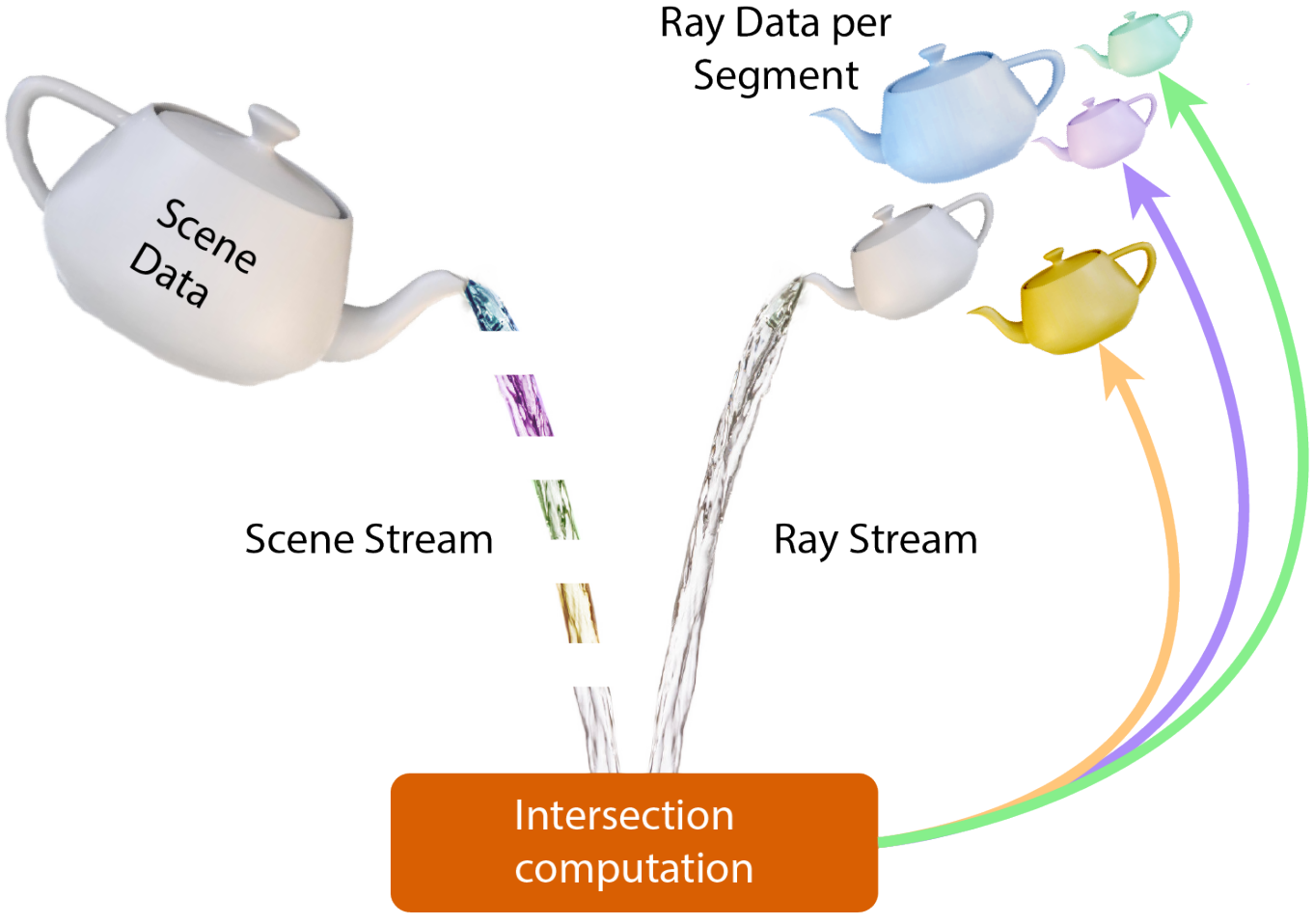

[Pharr et. al. 2010]


[Nah et. al. 2014]

Dual Streaming



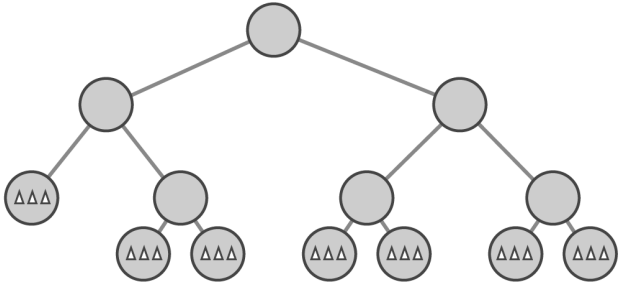
Dual Streaming



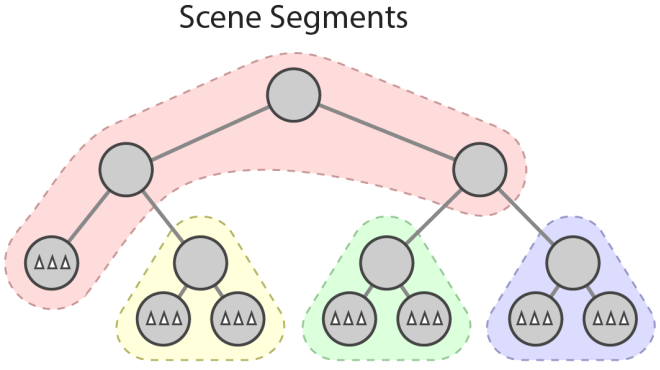
Contributions

- First fully streamed formulation of ray tracing
  - scene, ray streams
  - predictive memory access
  - new traversal order
  - scene traffic at absolute minimum
- Hardware architecture

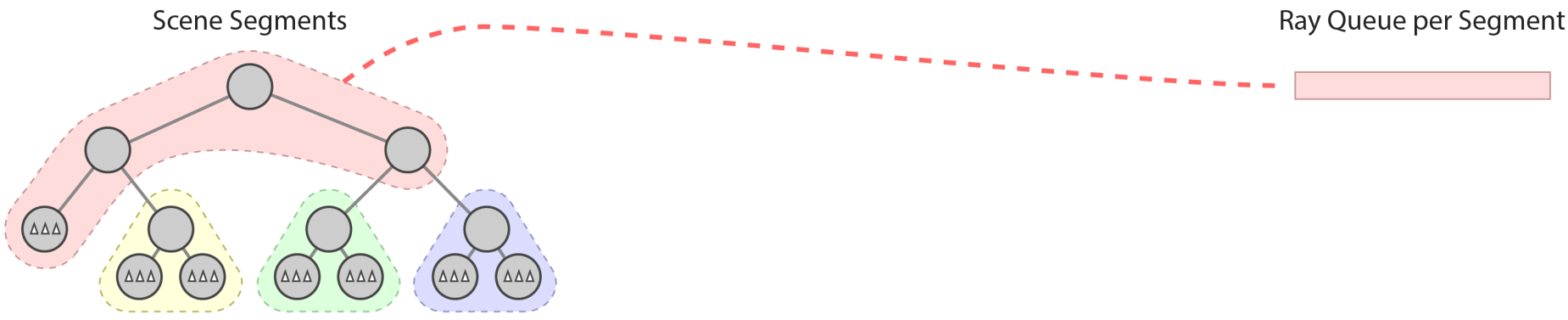
Predictive Traversal Order



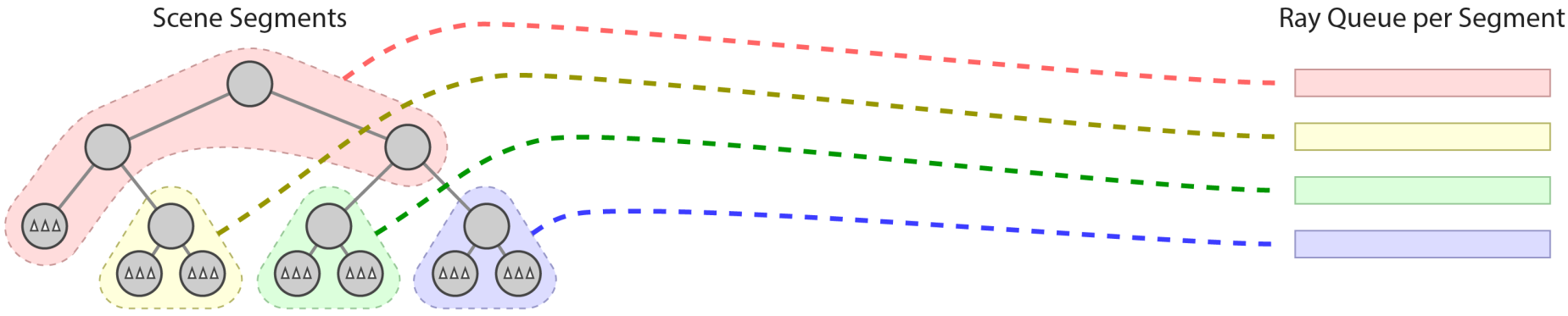
Predictive Traversal Order



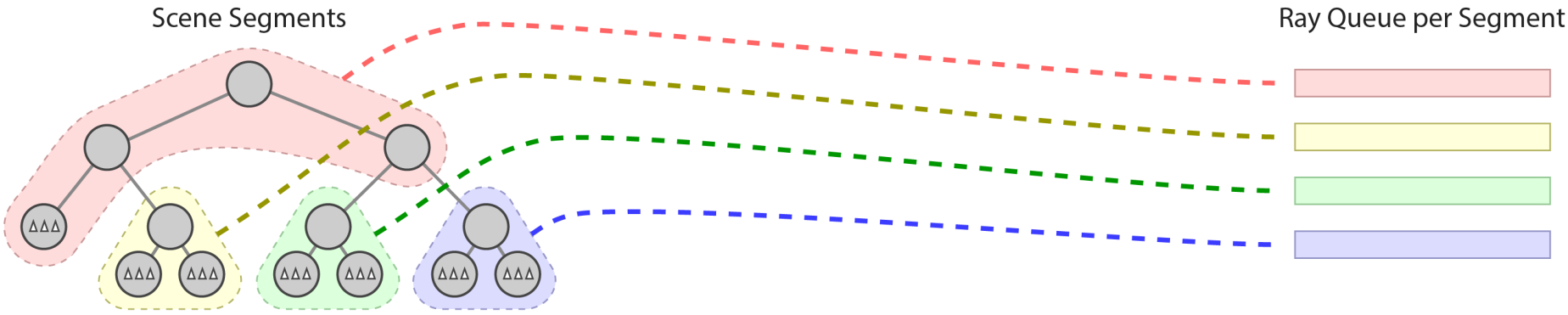
Predictive Traversal Order



Predictive Traversal Order



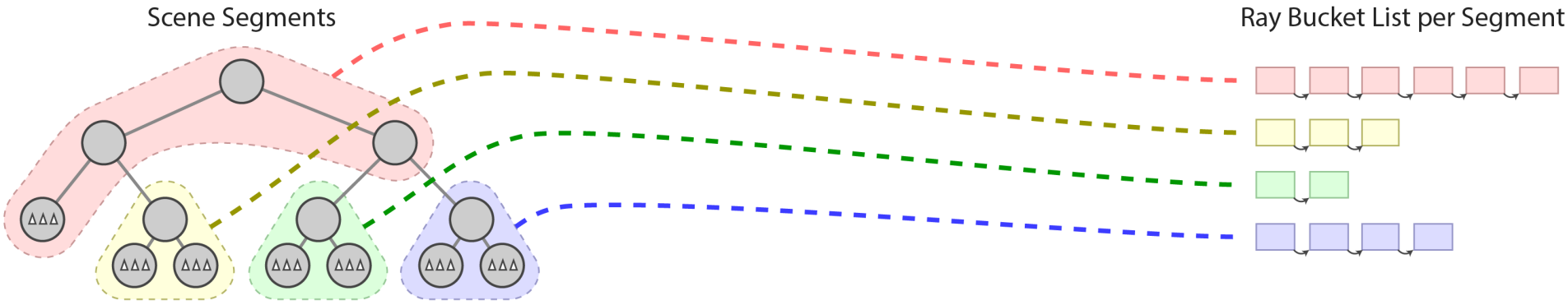
Predictive Traversal Order



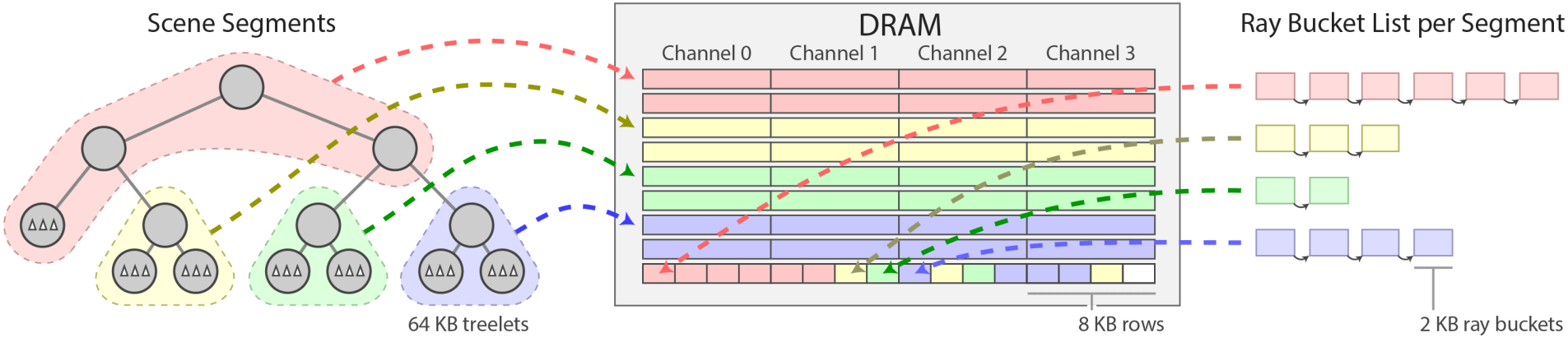
Rays can be duplicated

No traversal stack stored

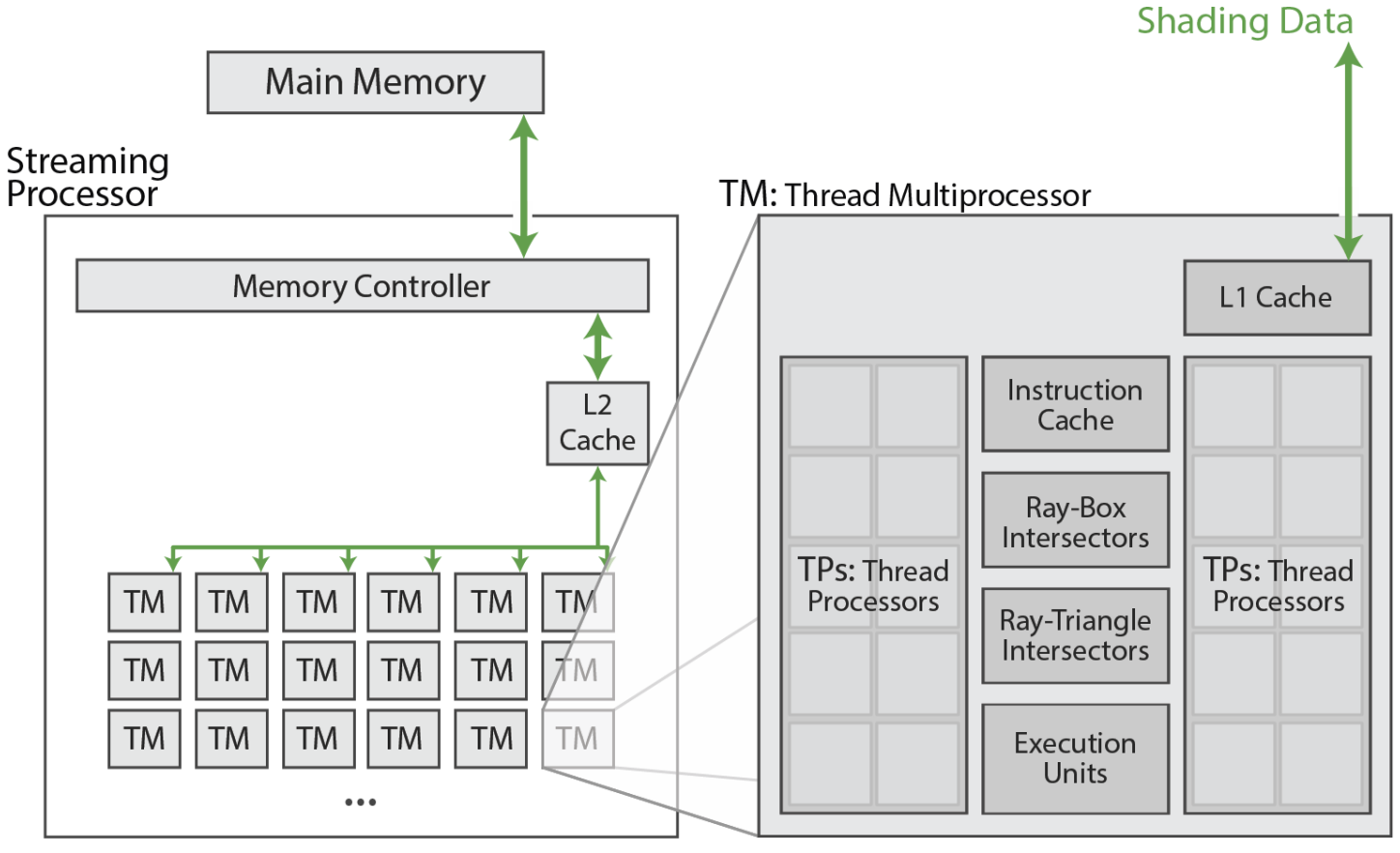
Stream Layout in DRAM



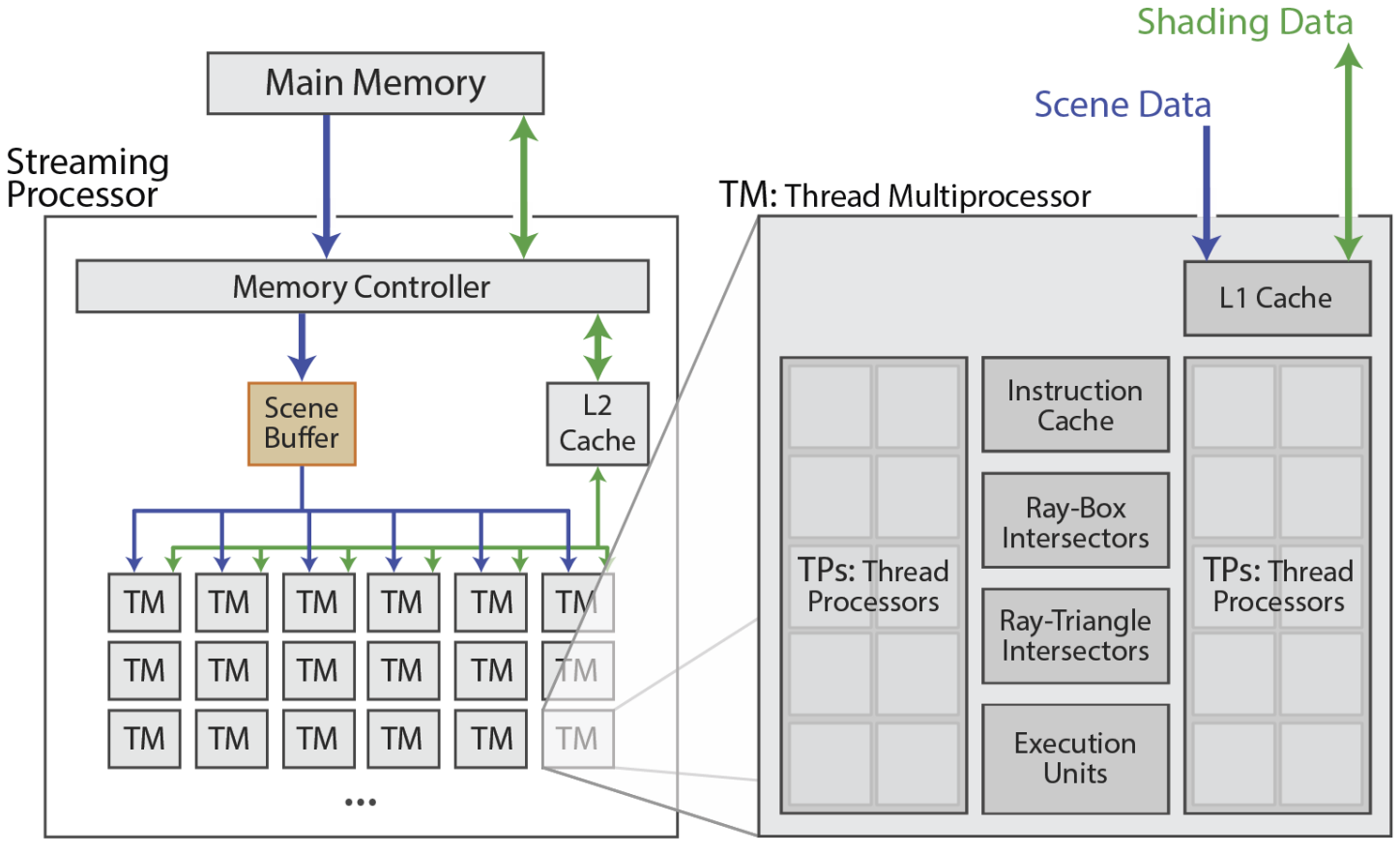
Stream Layout in DRAM



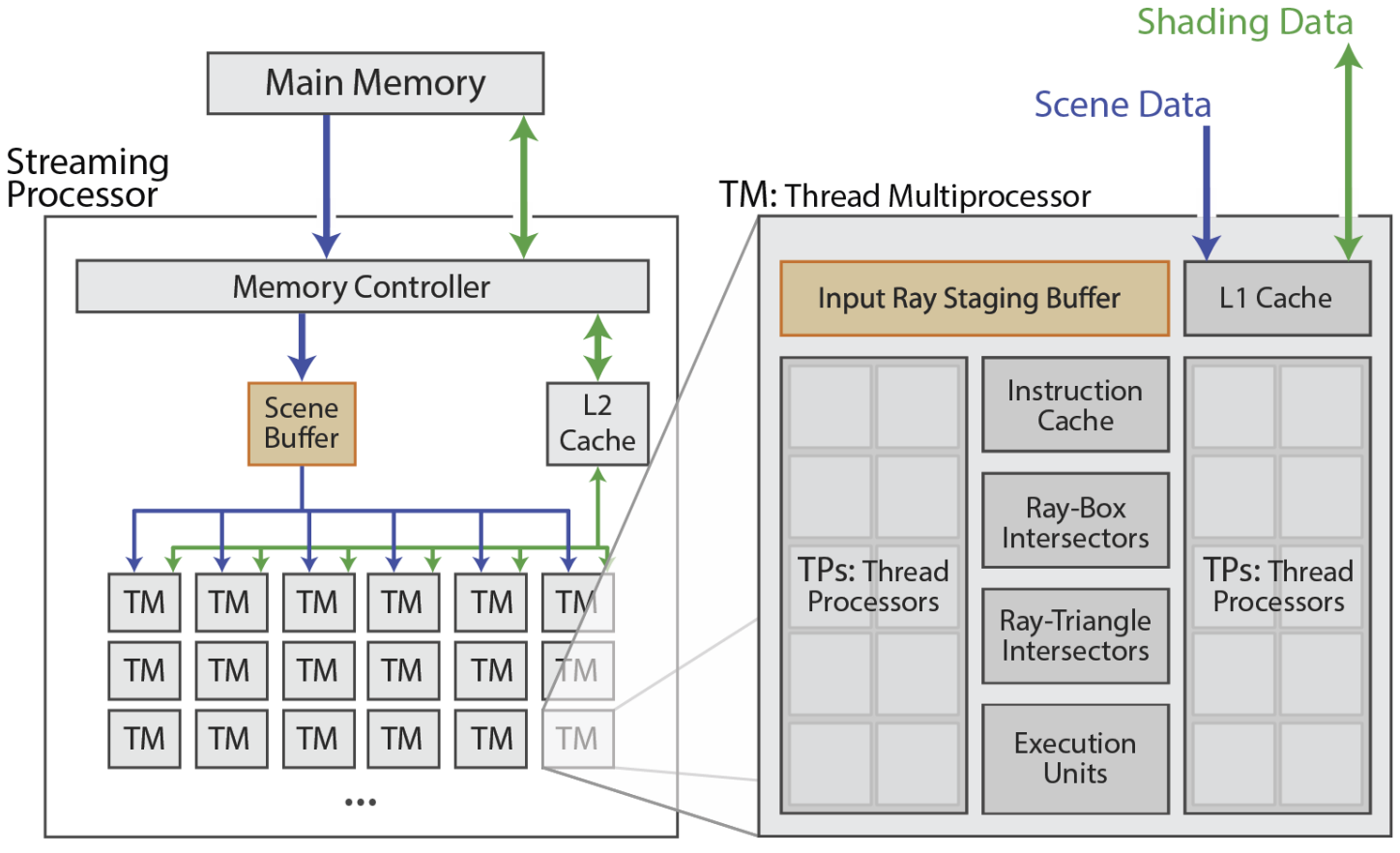
Hardware Architecture



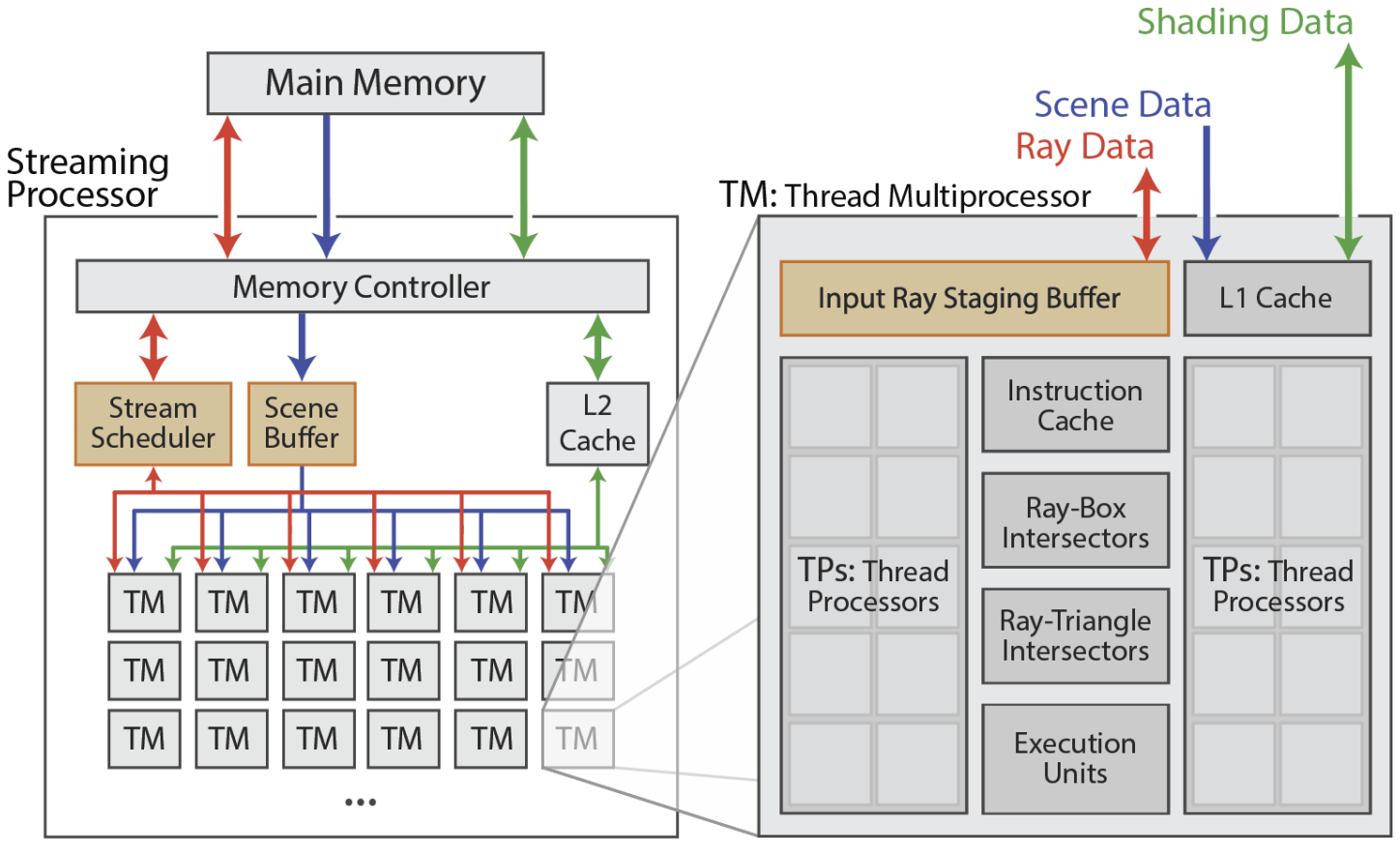
Hardware Architecture



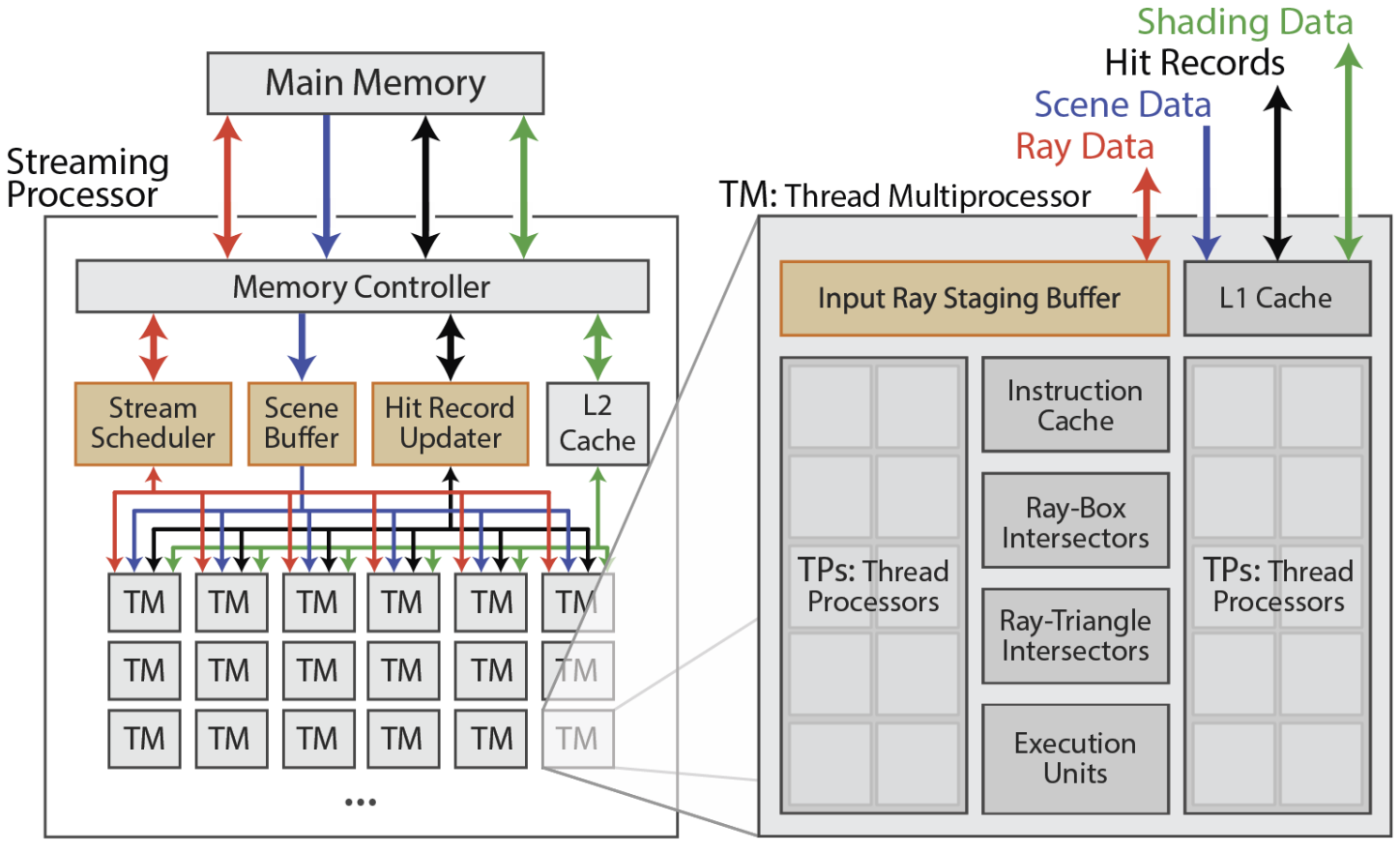
Hardware Architecture



Hardware Architecture

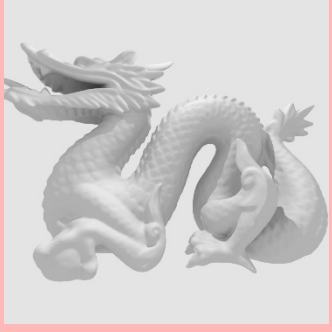


Hardware Architecture



Results: Scenes

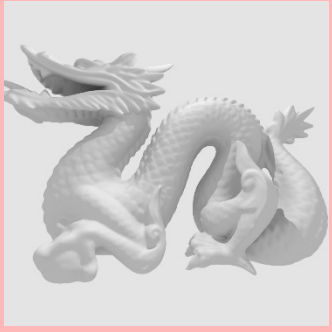
Benchmark

Dragon 870 K	Dragon Box 870 K	Dragon Sponza 6.6 M	San Miguel 10.5 M
			

Results: Scenes

Benchmark

Dragon
870 K

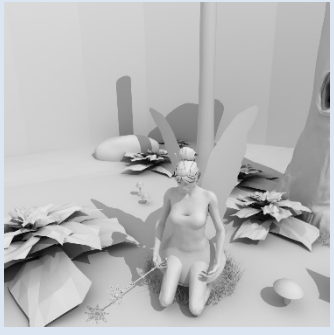


Dragon Box
870 K

Dragon Sponza
6.6 M

San Miguel
10.5 M

Small



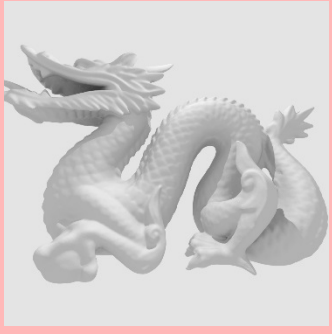
Fairy Forest
174 K

Crytek Sponza
262 K

Results: Scenes

Benchmark

Dragon
870 K

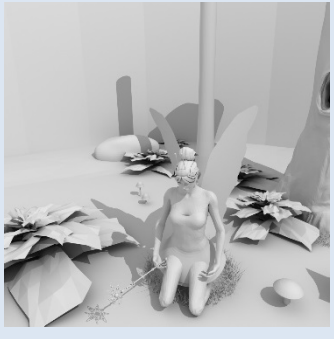


Dragon Box
870 K

Dragon Sponza
6.6 M

San Miguel
10.5 M

Small



Fairy Forest
174 K

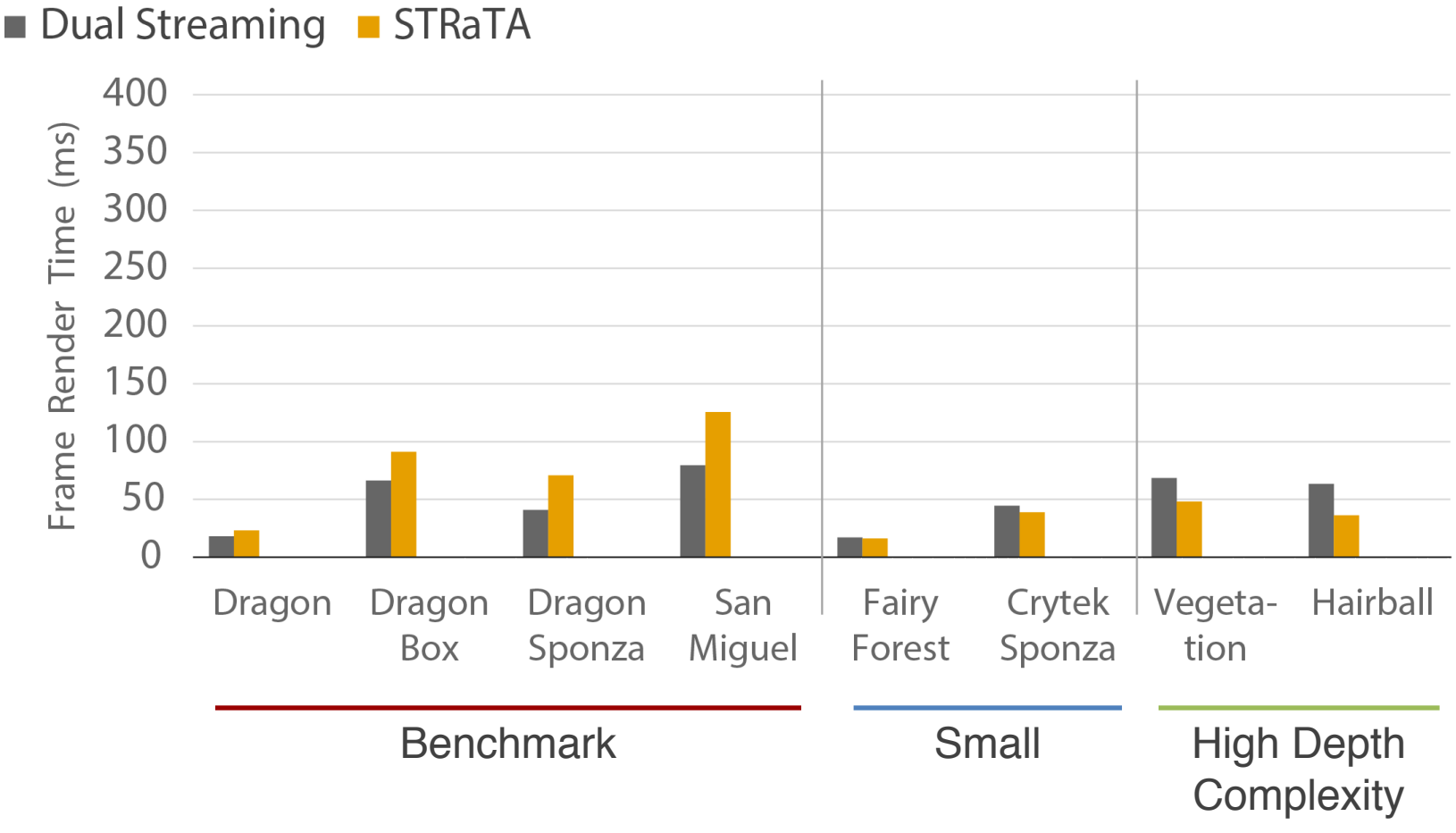
Crytek Sponza
262 K

Vegetation
1.1 M

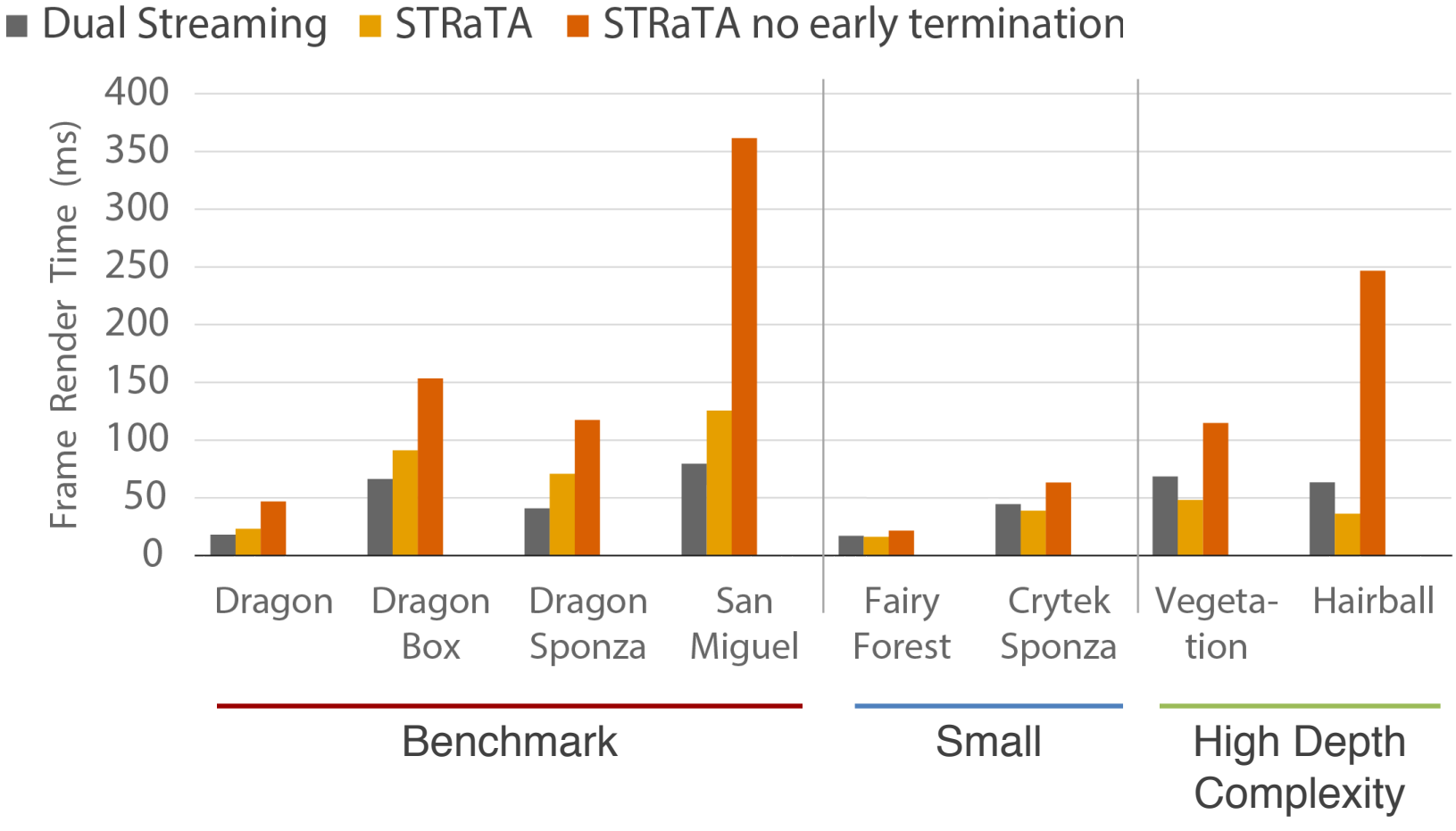
Hairball
2.9 M

High Depth Complexity

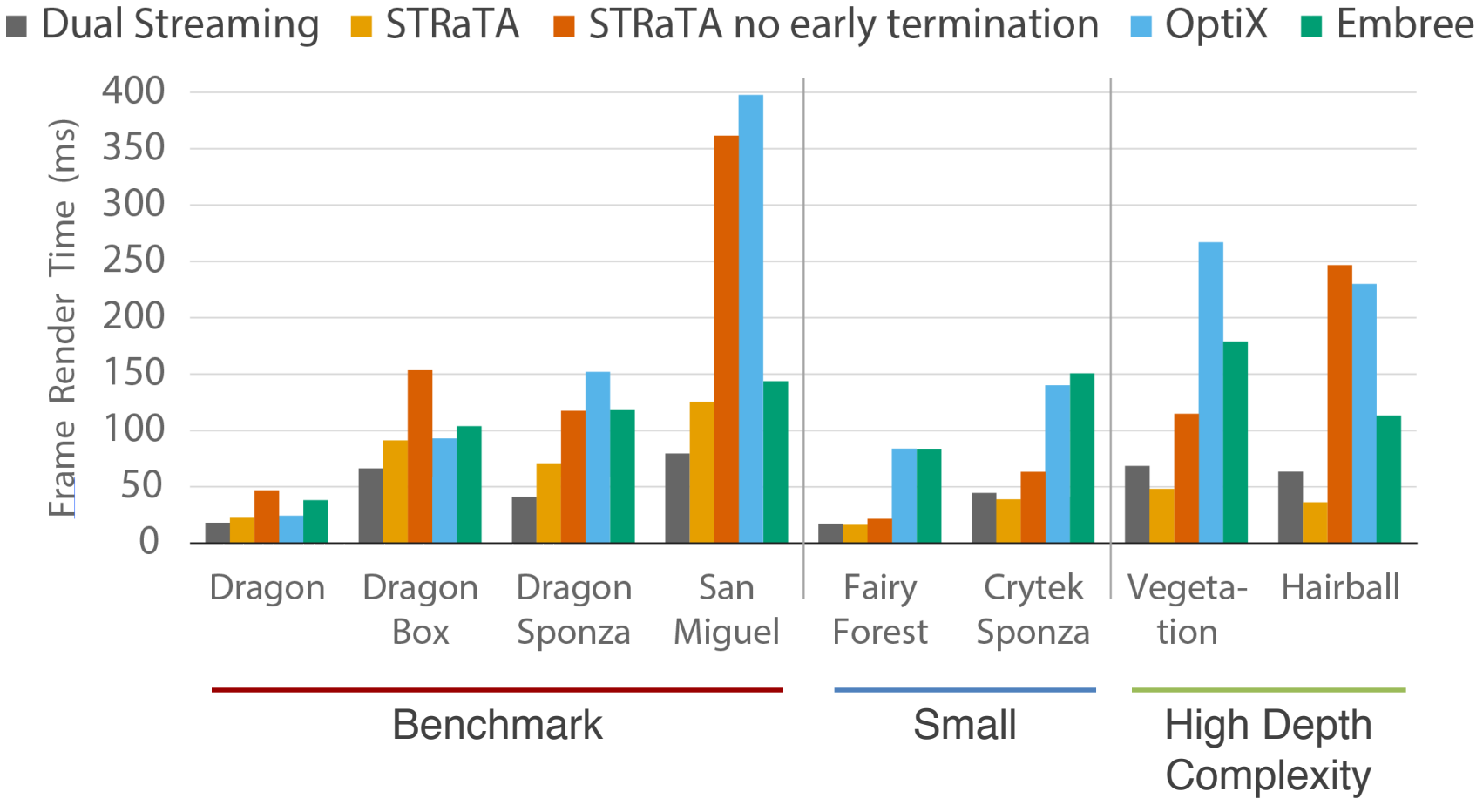
Results: Render Time



Results: Render Time



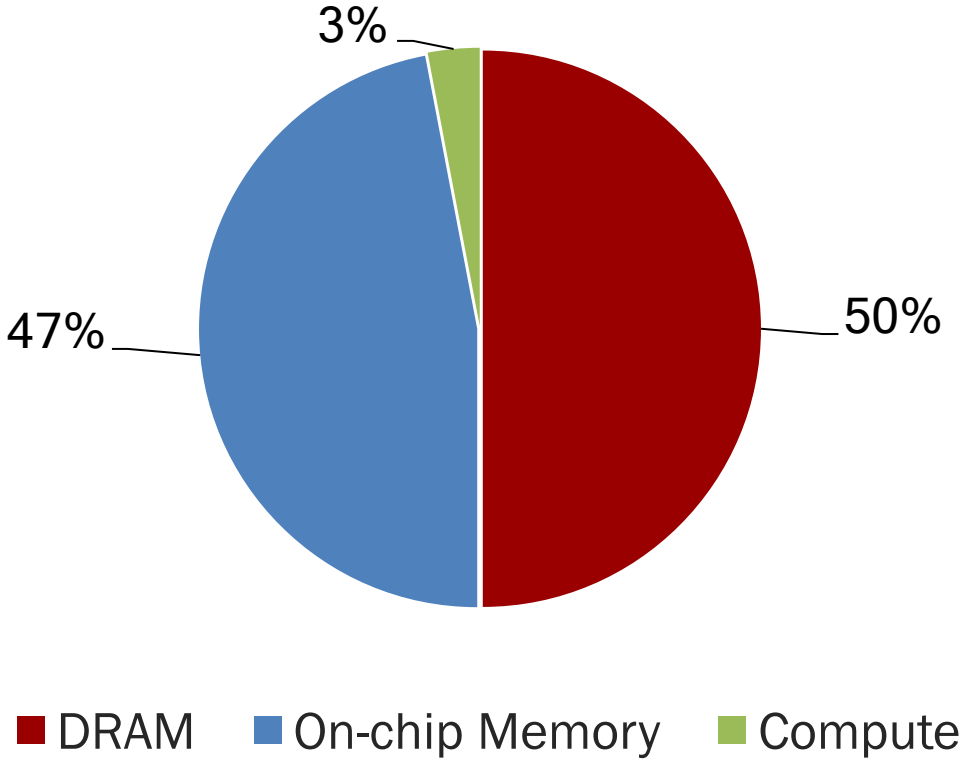
Results: Render Time



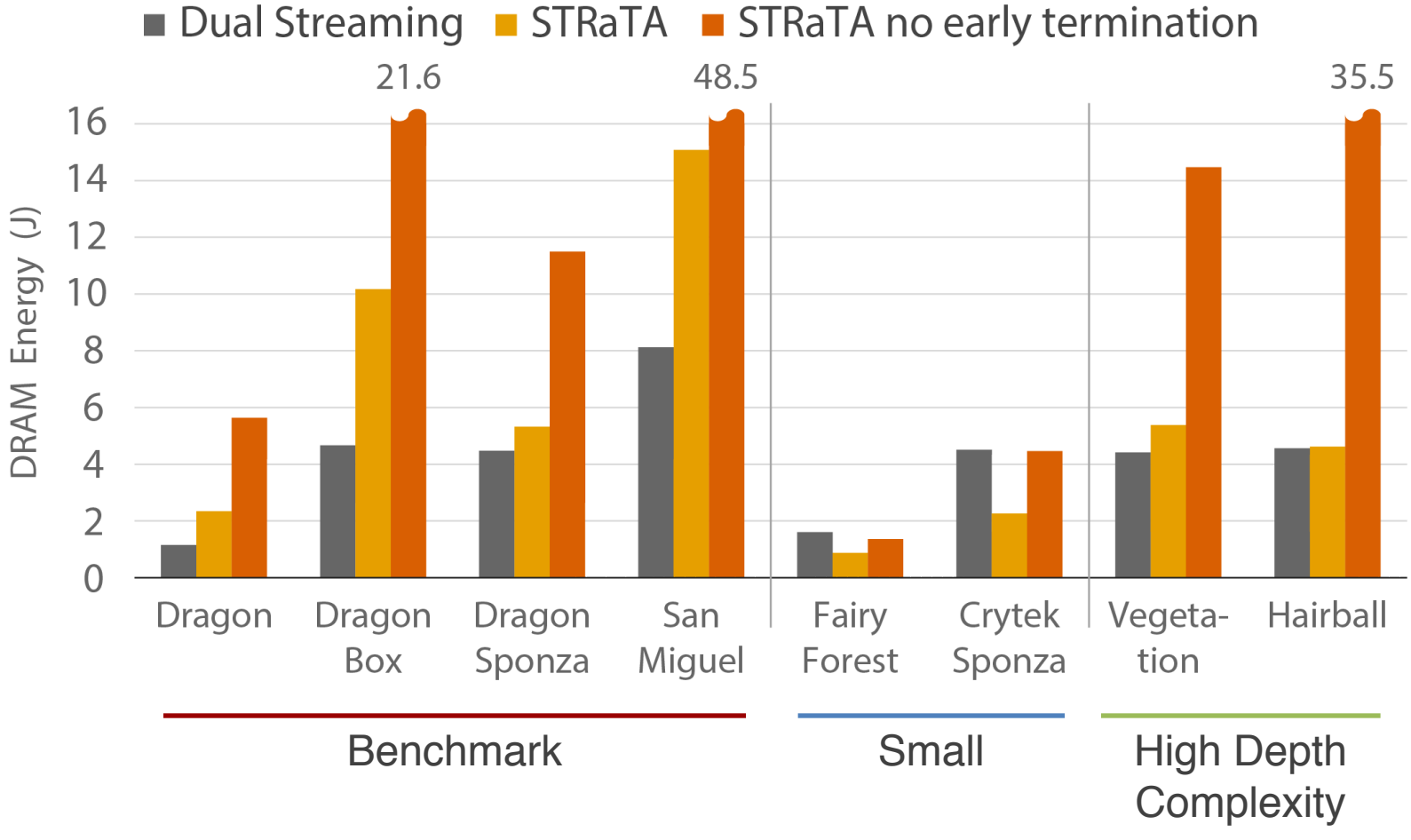
Results: Configurations

	Dual Streaming	STRaTA	OptiX	Embree
Processor	custom	custom	NVIDIA GTX Titan	Intel i7-5960X
# cores	2048	2048	2688	8 (16)
frequency	1 GHz	1 GHz	0.9 GHz	4.6 GHz
on-chip memory	7.7 MB	9.2 MB	5.9 MB	22.5 MB
DRAM bandwidth	512 GB/s	512 GB/s	288 GB/s	75 GB/s

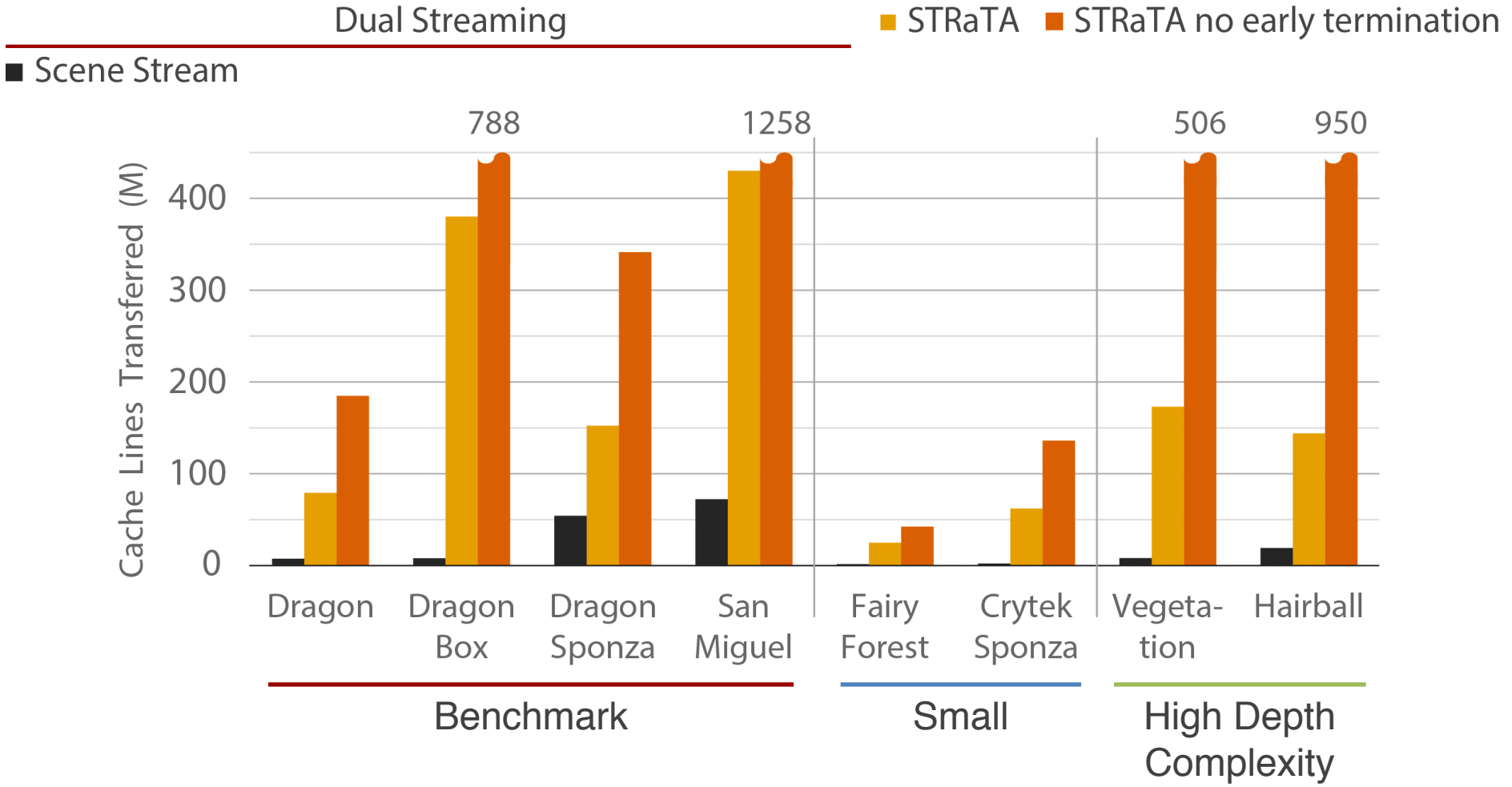
Results: Frame Energy



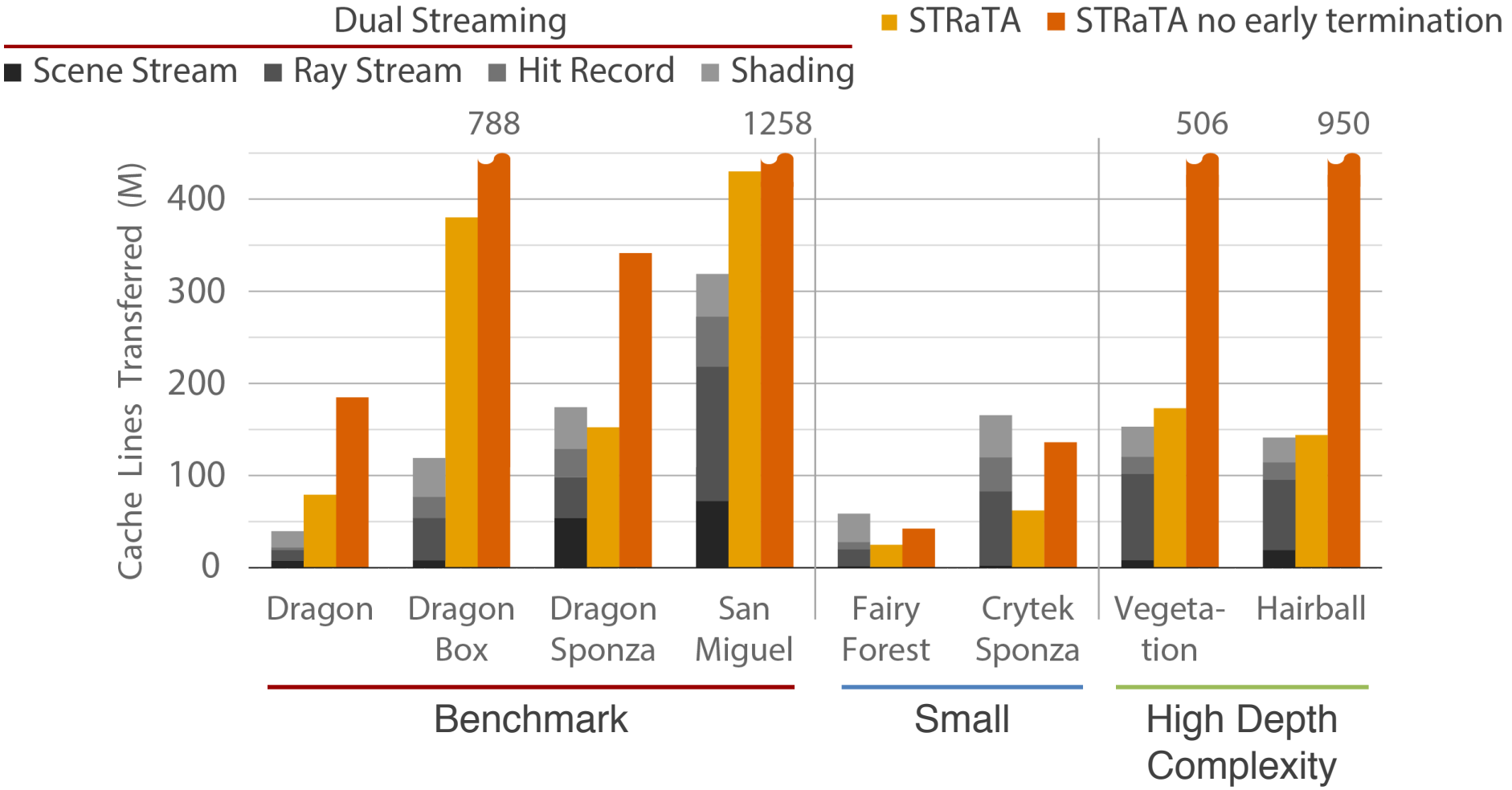
Results: DRAM Energy



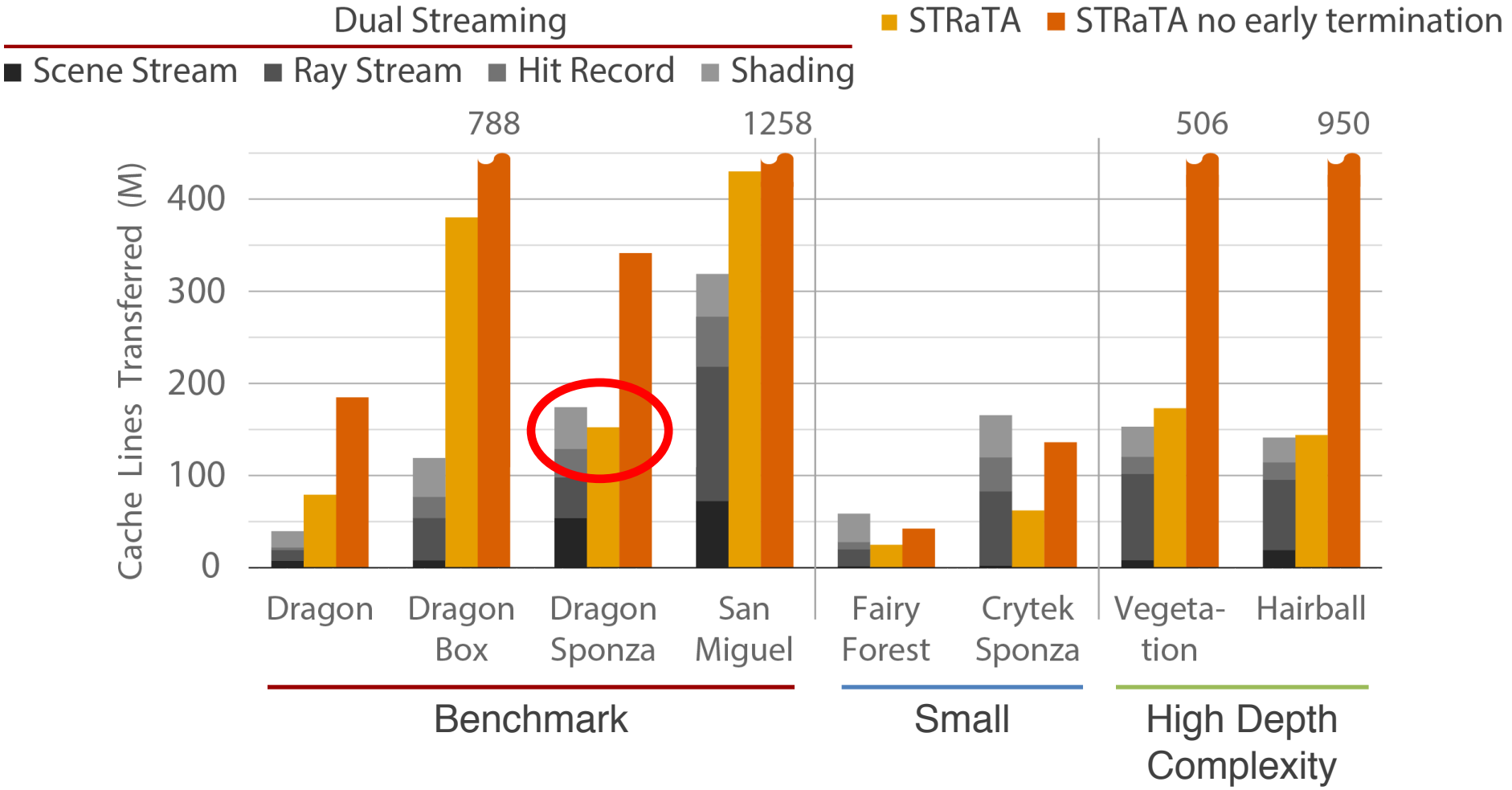
Results: Scene Traffic



Results: All Traffic



Results: All Traffic



Future Work

- Early ray termination

Future Work

- Early ray termination
- Dynamic traversal order

Future Work

- Early ray termination
- Dynamic traversal order
- Treelet assignment

Future Work

- Early ray termination
- Dynamic traversal order
- Treelet assignment
- Data compression

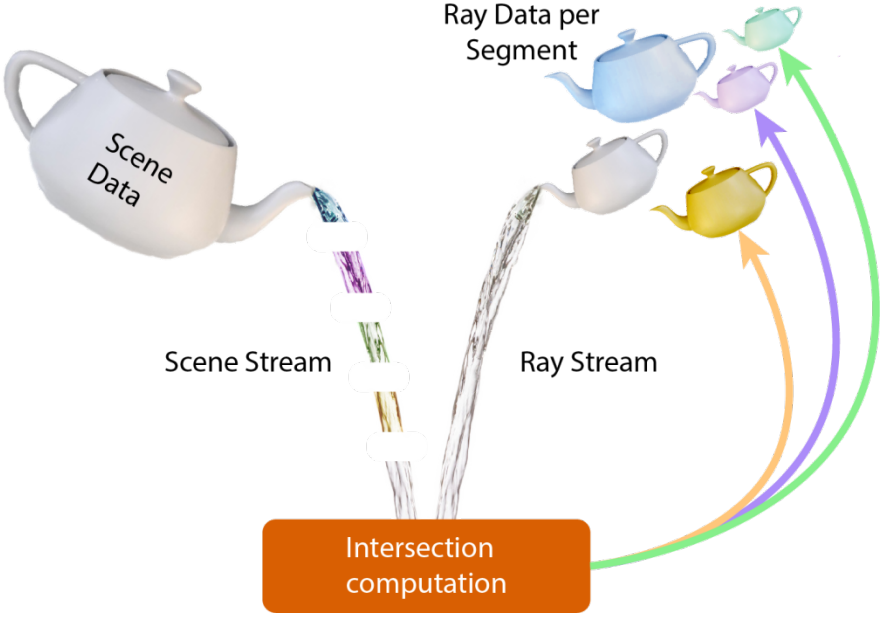
Conclusion

- First fully streamed formulation of ray tracing
  - scene, ray streams
  - predictive memory access
  - new traversal order
  - scene traffic at absolute minimum
- Hardware architecture

Acknowledgments

- Josef Spjut
- Elena Vasiou
- Frank Meinl
- Marko Dabrovic
- Samuli Laine
- Guillermo Leal Llaguno
- Cem Yuksel
- Anonymous reviewers
- This work is supported in part by the National Science Foundation under Grant No. 1409129.

Thank you!


Ray Data per
Segment
Scene
Data
Scene Stream
Ray Stream
Intersection
computation