



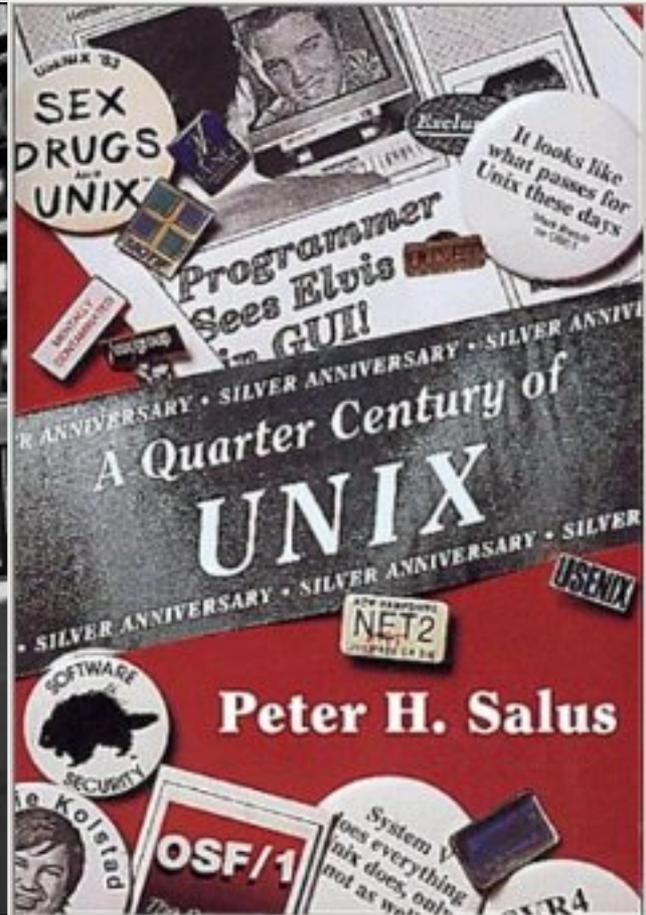
Data-Parallel Computing Using Madagascar

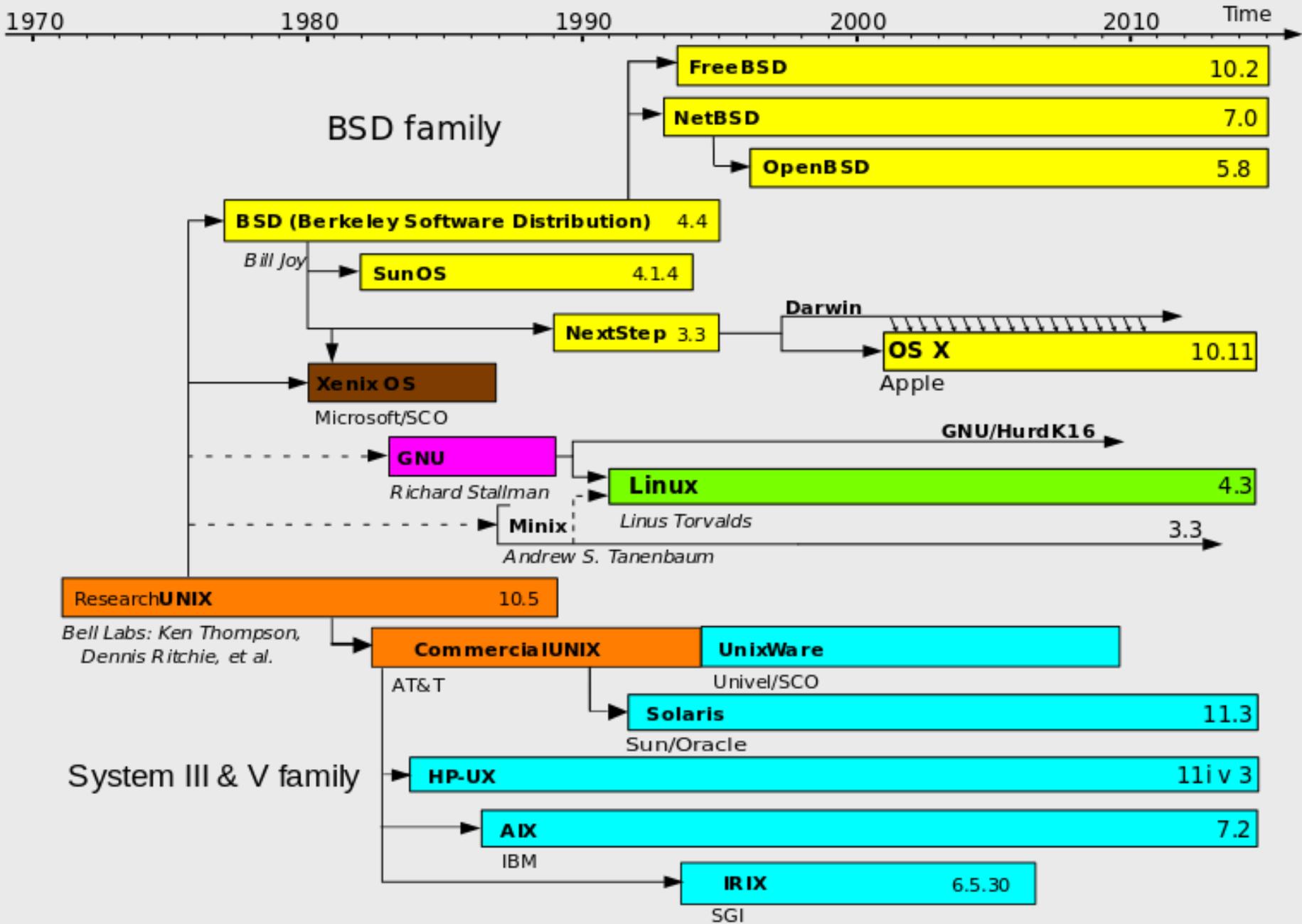
Sergey Fomel

The University of Texas at Austin

PASC16

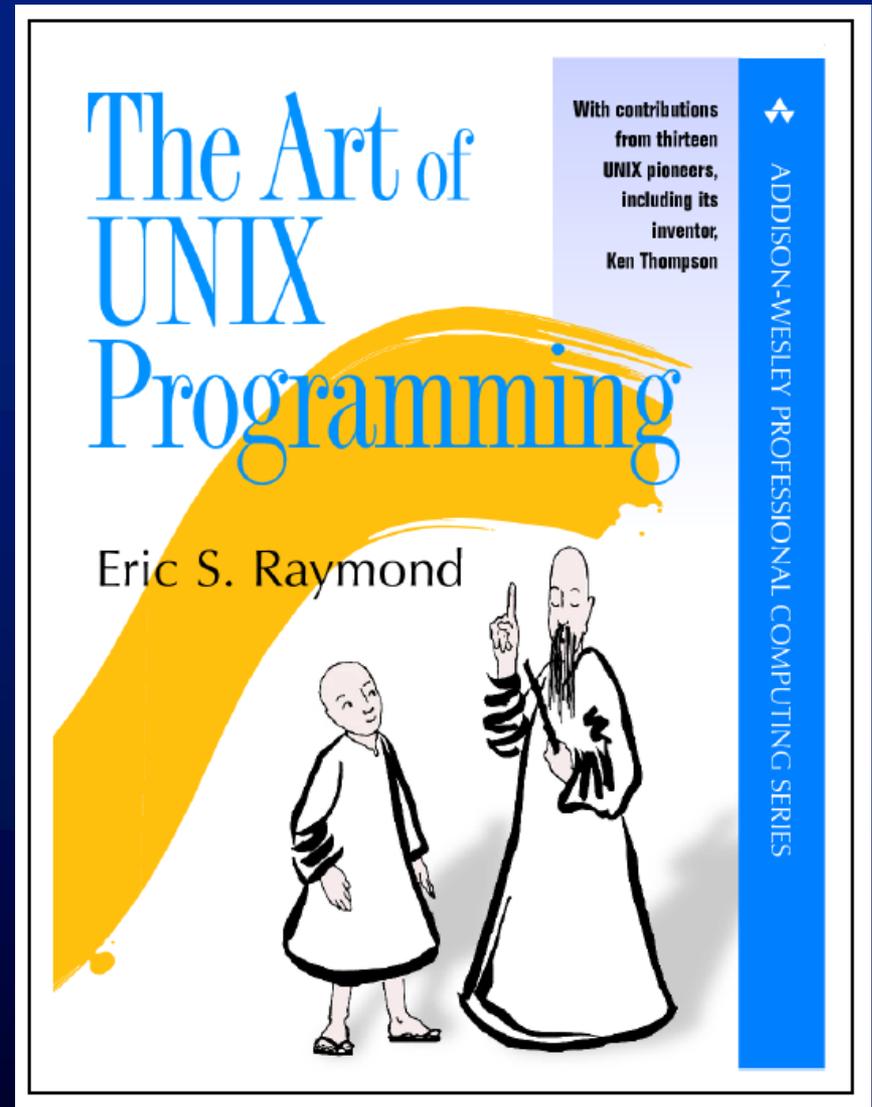
45+ Years of Unix

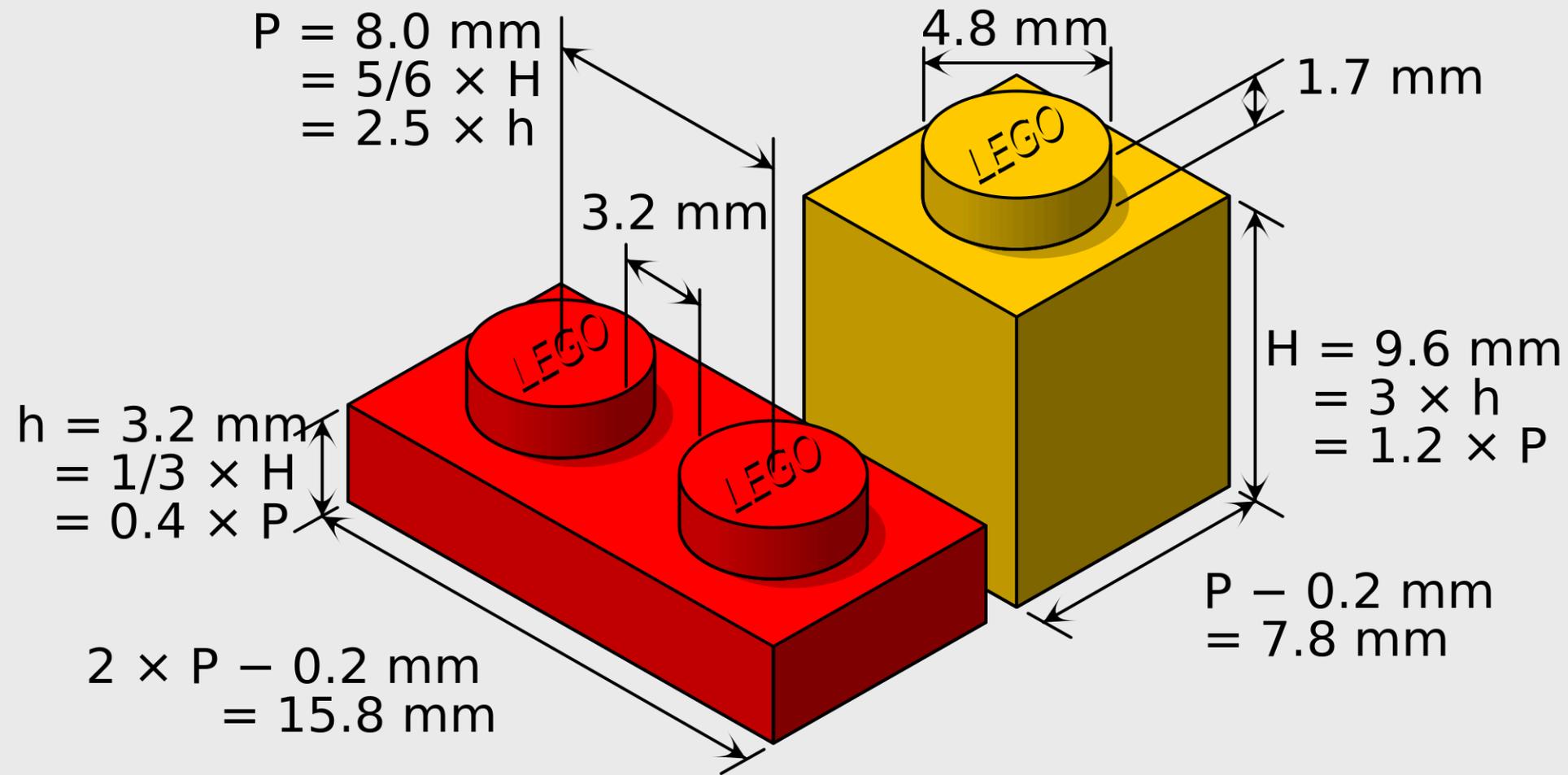




“Write programs that do one thing and do it well. Write programs to work together. Write programs to handle text streams, because that is a universal interface.”

Doug McIlroy







MADAGASCAR

<http://ahay.org>

In a nutshell, Madagascar...

...has had 12,757 commits made by 97 contributors representing 1,118,533 lines of code

...is mostly written in C with an average number of source code comments

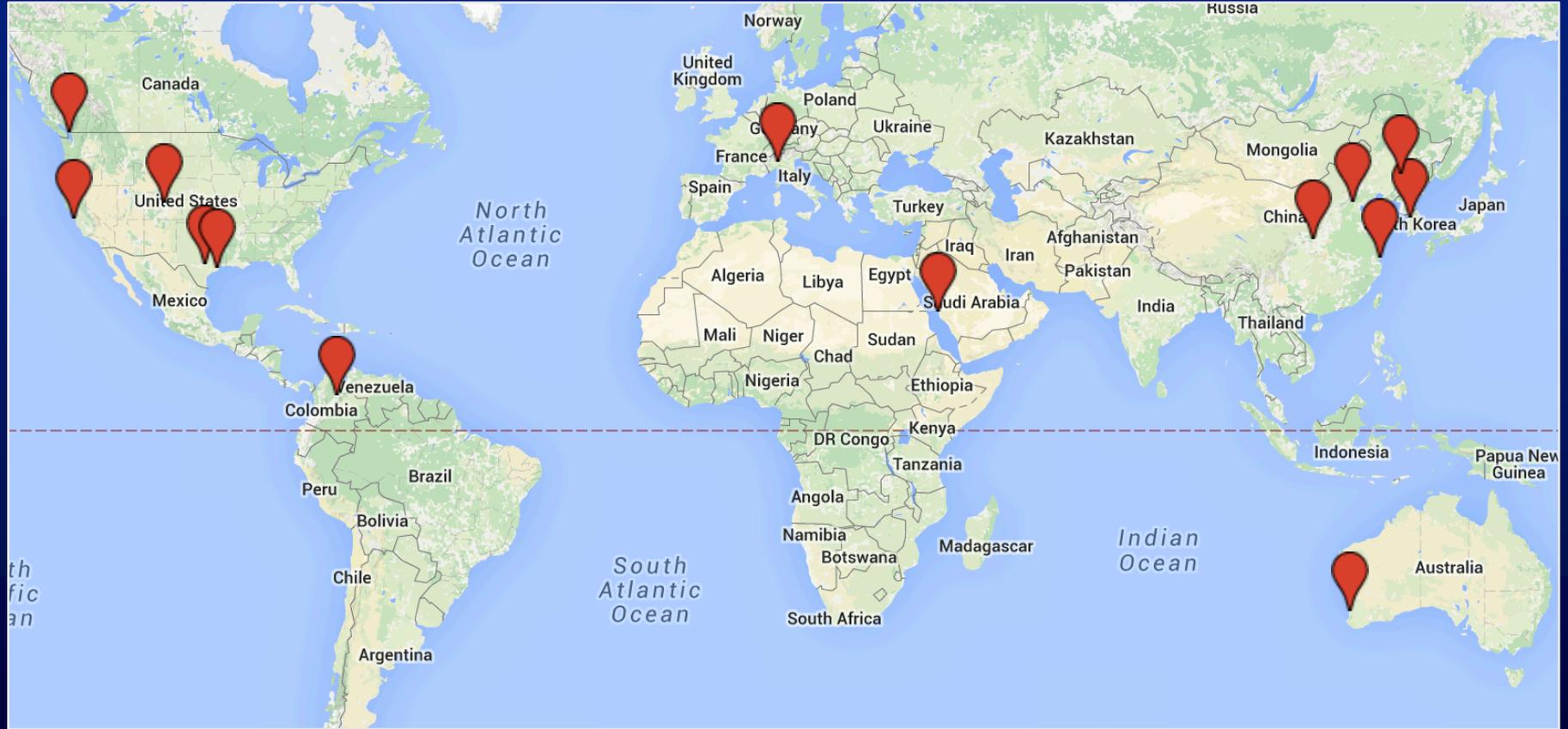
...has a well established, mature codebase maintained by a large development team

...took an estimated 328 years of effort starting with its first commit in May, 2003

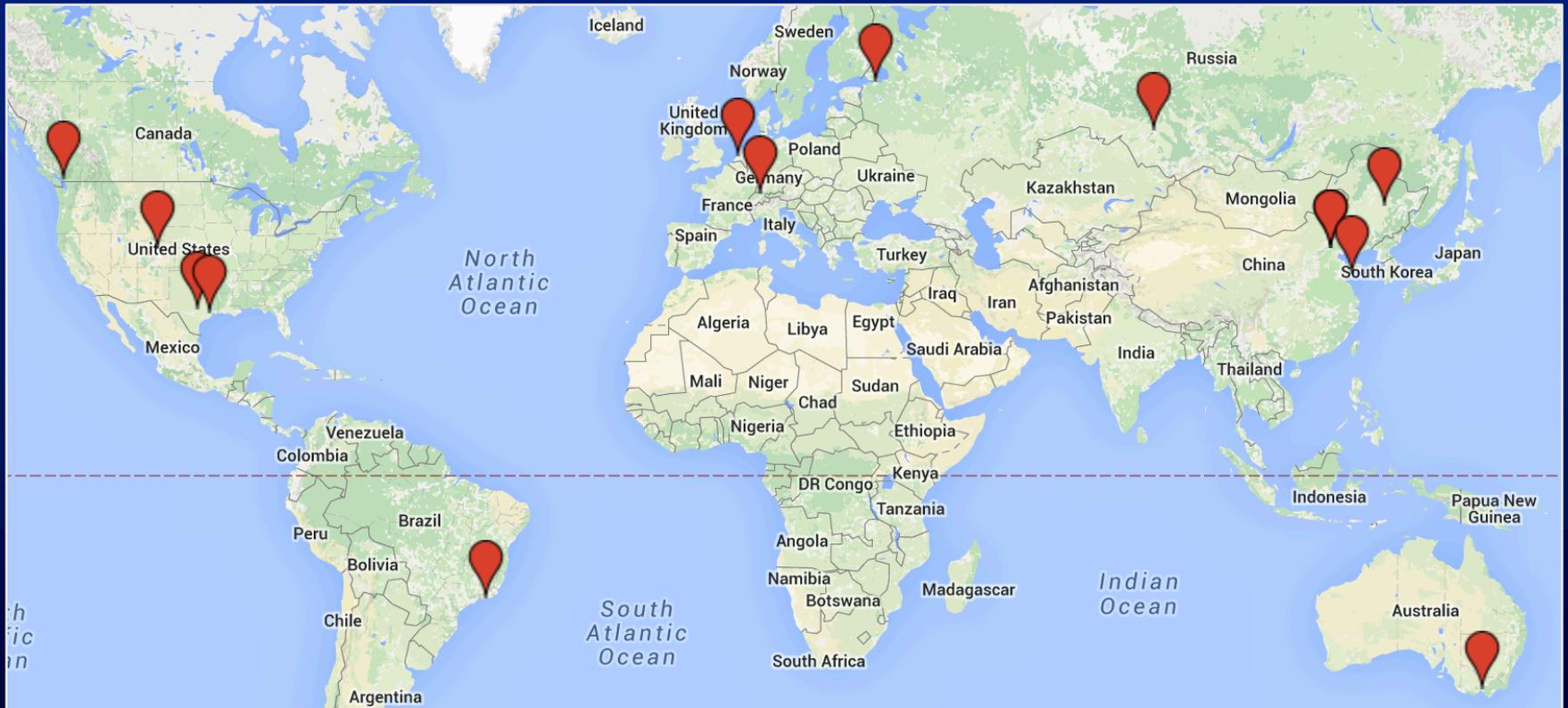
Contributors

Salah Al-Hadab, Tariq Alkhalifah, Vladimir Bashkardin, Filippo Broggin, Jules Browaeys, Cody Brown, William Burnett, Yihua Cai, Maria Cameron, Lorenzo Casasanta, Yangkang Chen, Zhonghuan Chen, Jiubing Cheng, Luke Decker, Joseph Dellinger, Esteban Diaz, Yuting Duan, Gang Fang, Mehdi Far, Sergey Fomel, Jeff Godwin, Gilles Hennenfent, Jie Hou, Jingwei Hu, Yin Huang, Trevor Irons, Detchai Ittharat, Jim Jennings, Jun Ji, Long Jin, Parvaneh Karimi, Roman Kazinnik, Alexander Klovov, Siwei Li, Guochang Liu, Yang Liu, Yujin Liu, Xuxin Ma, Douglas McCowan, Dmitrii Merzlikin, Henryk Modzelewski, Jorge Monsegny, Francesco Perrone, Jack Poulson, Kelly Regimbal, James Rickett, Sjoerd de Ridder, Daniel Rocha, Sean Ross-Ross, Colin Russell, Christos Saragiotis, Paul Sava, Karl Schleicher, Reza Shahidi, Yunzhi Shi, Jeffrey Shragge, Eduardo Filpo Silva, Xiaolei Song, Yanadet Sripanich, Junzhe Sun, Ioan Vlad, Hui Wang, Robin Weiss, Zedong Wu, Zhiguang Xue, Jia Yan, Jun Yan, Pengliang Yang, Lexing Ying, Zhendong Zhang, Hejun Zhu + Joseph Dellinger, William Symes, and many others...

Madagascar Contributors



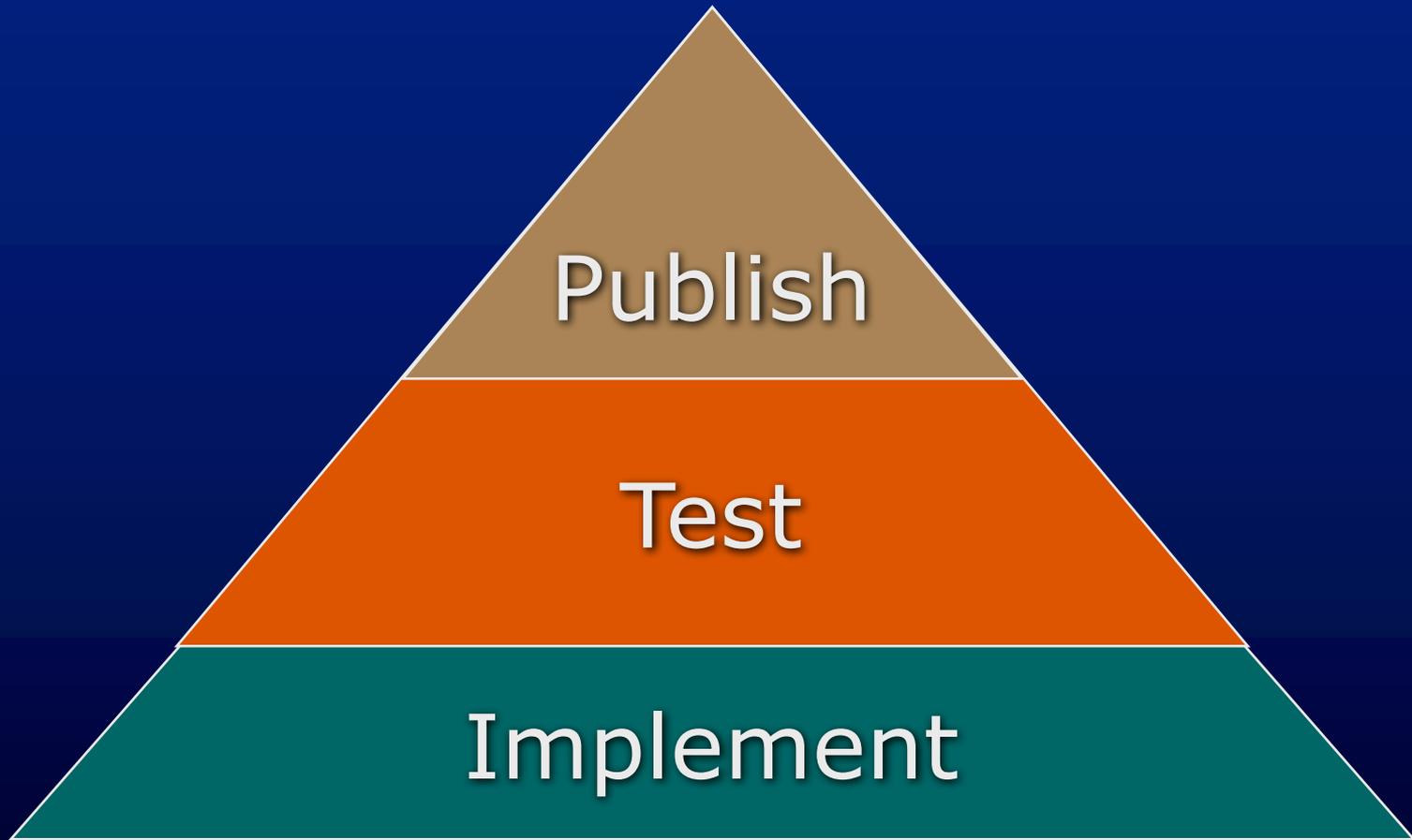
Madagascar Schools



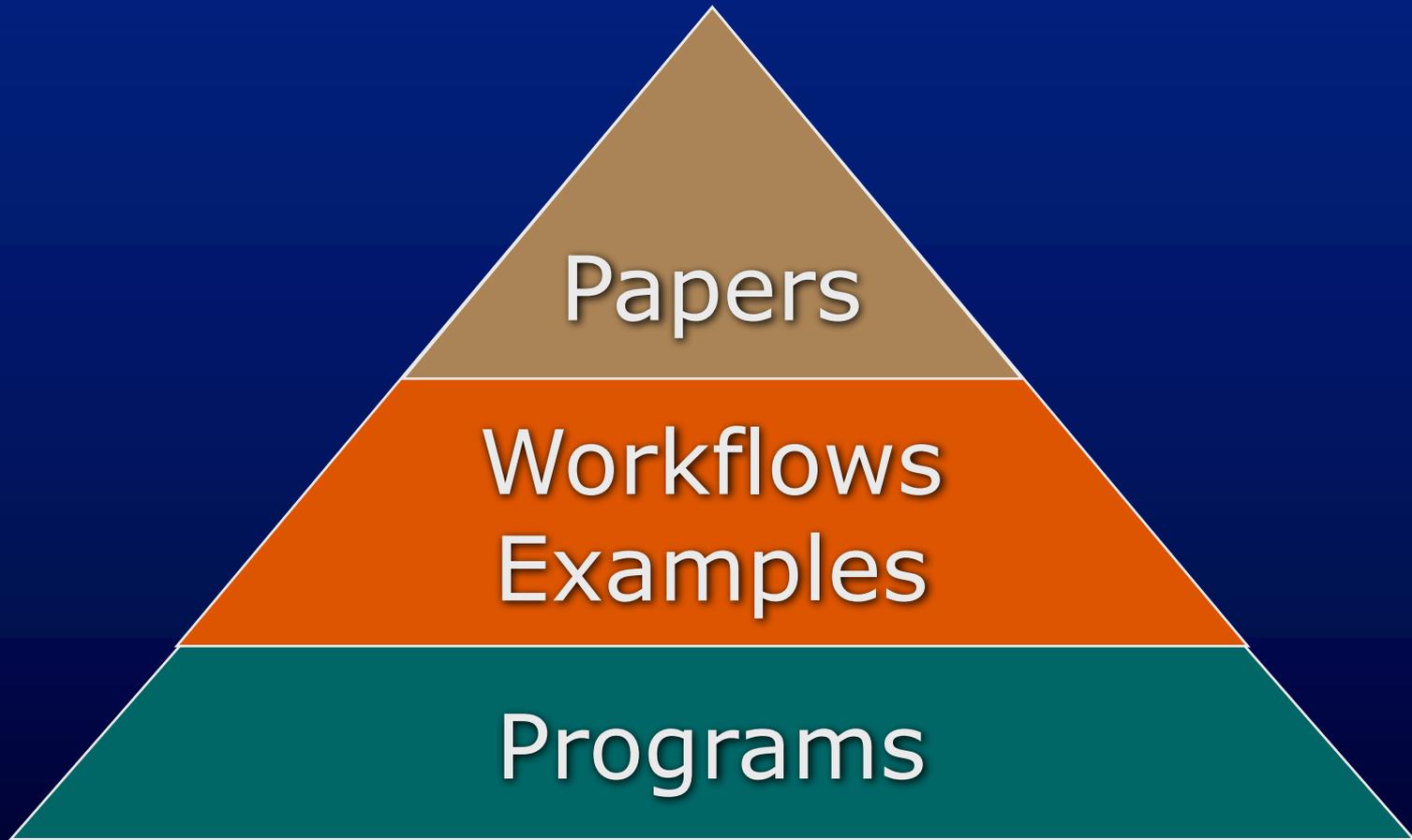
Qingdao-2015



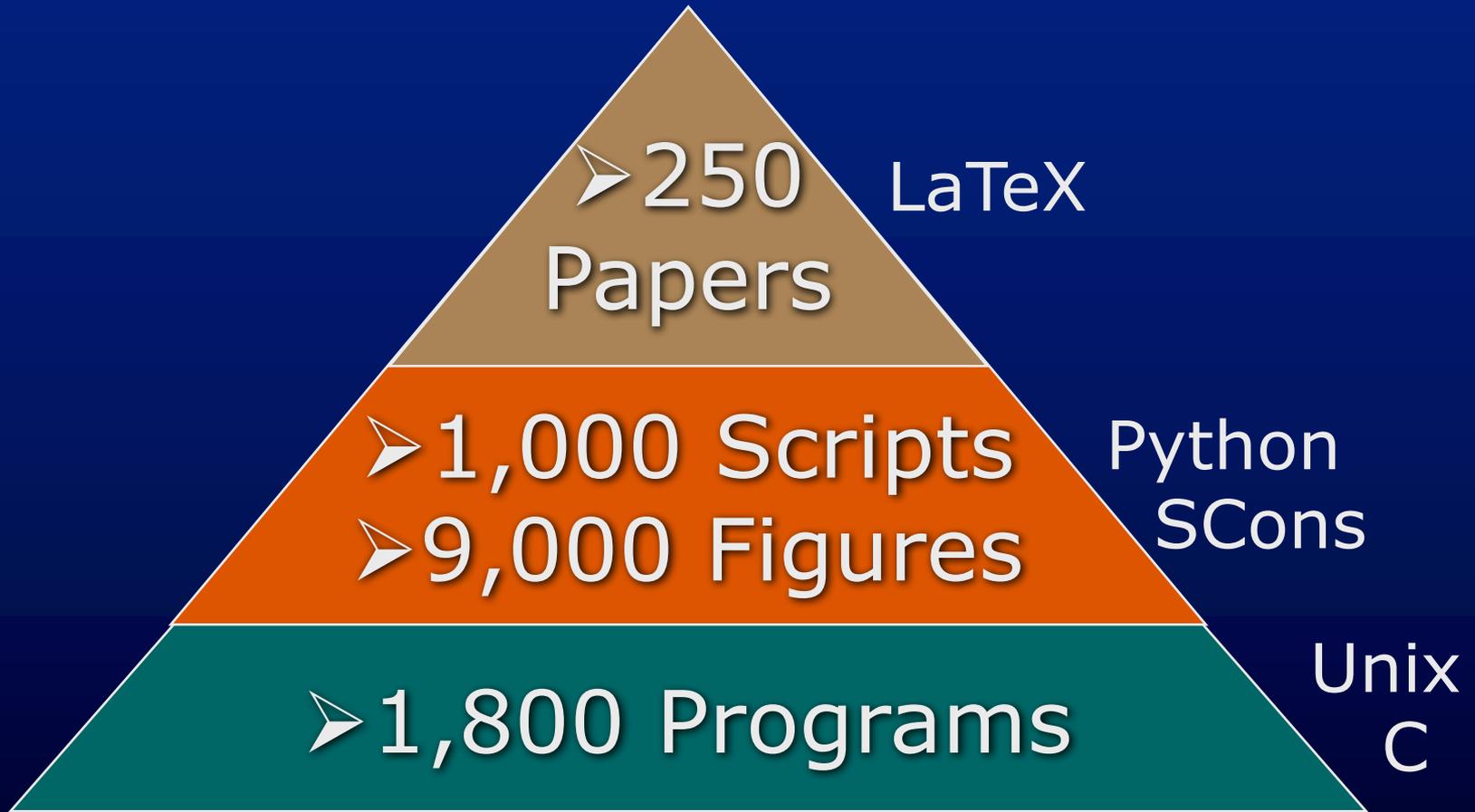
Research Pyramid

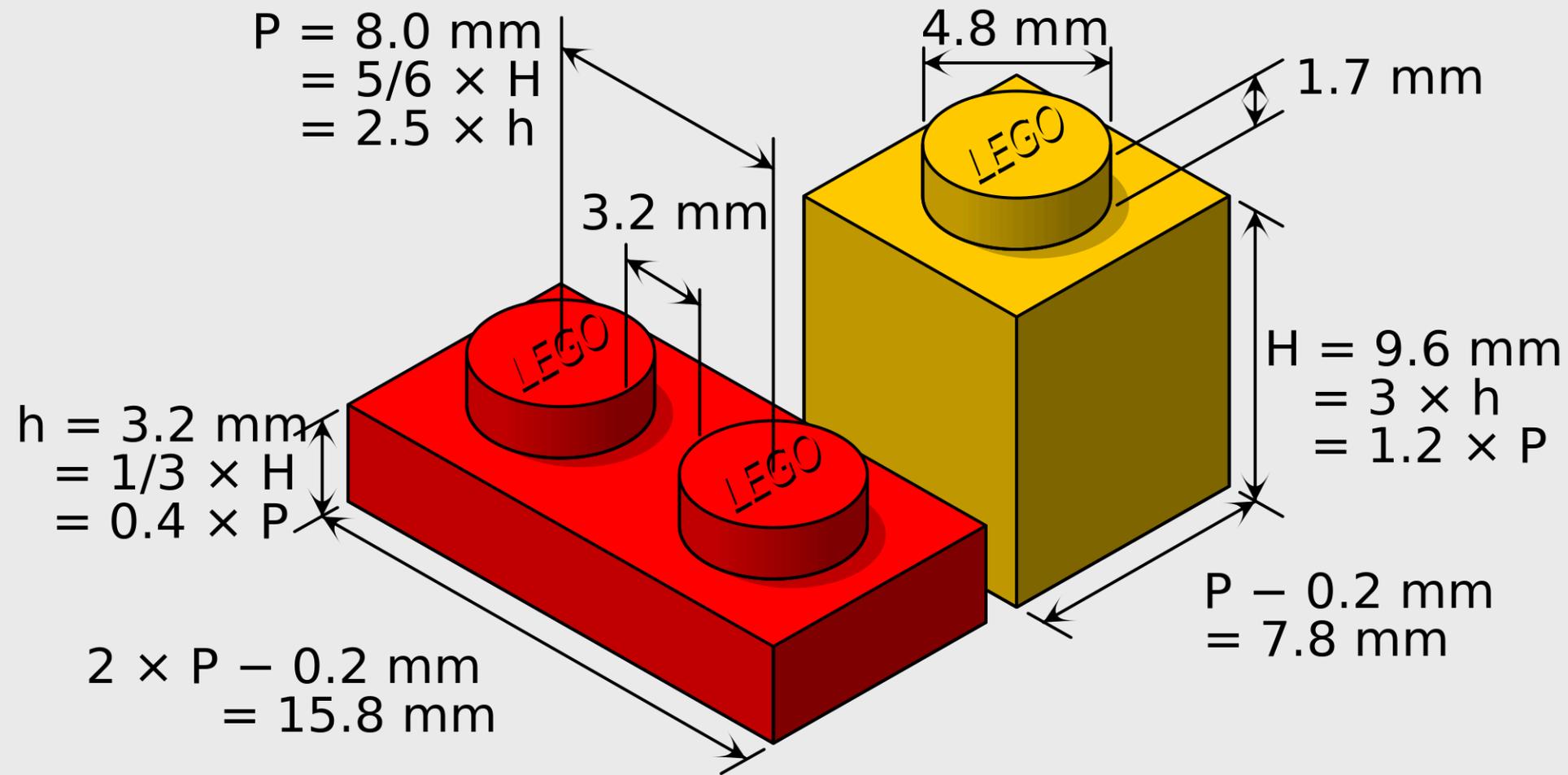


Research Pyramid



Research Pyramid





Least-Squares Inversion by Conjugate Gradients

mathematics

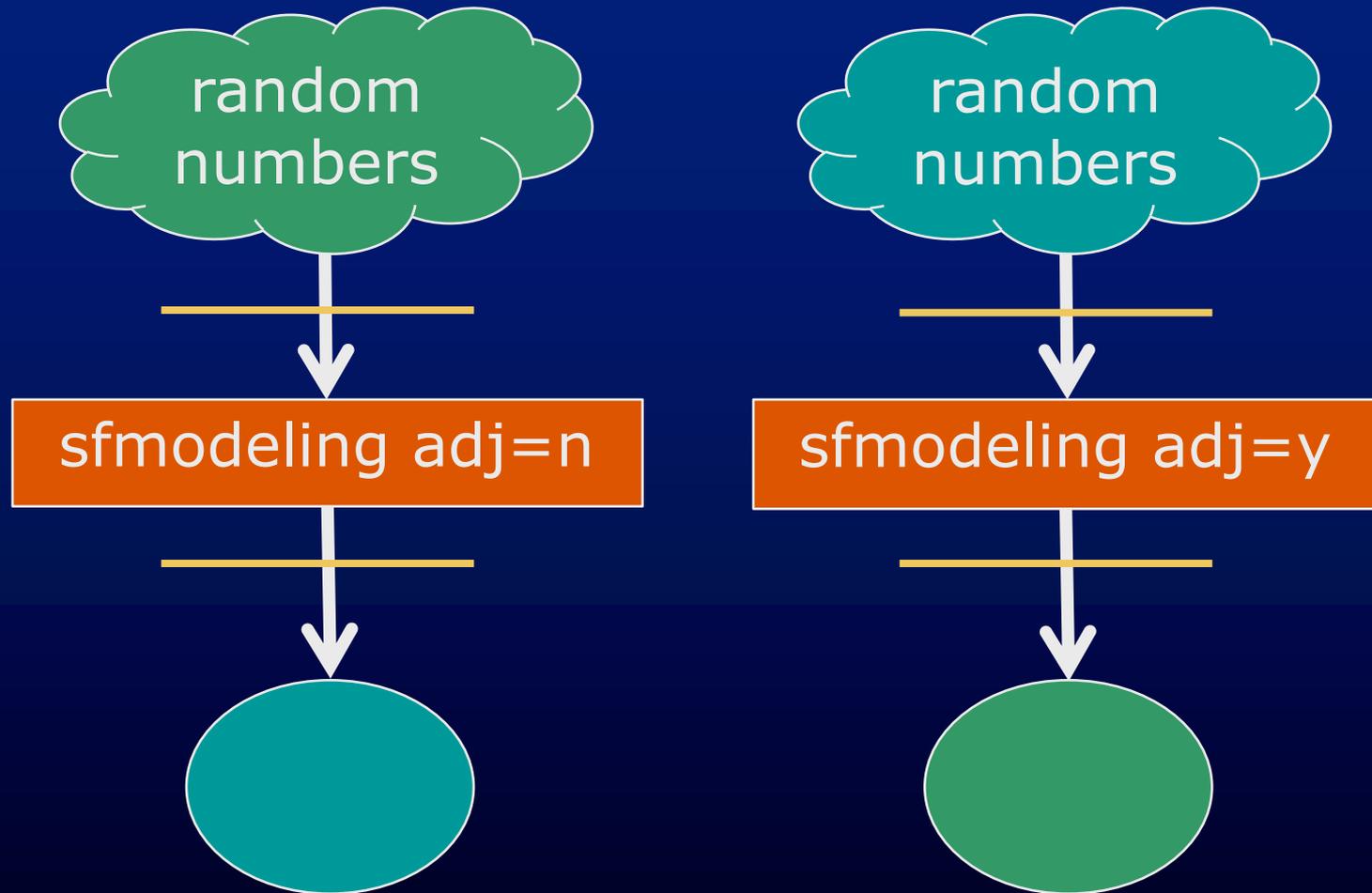
geophysics

- ◆ `sfconjgrad sfmodeling vel=vel.rsf \`
`mod=image0.rsf niter=100 \`
`< data.rsf > image.rsf`
- ◆ `sfdotest sfmodeling vel=vel.rsf \`
`mod=image0.rsf dat=data.rsf`

`sfdotest: L[m]*d =1165.87`

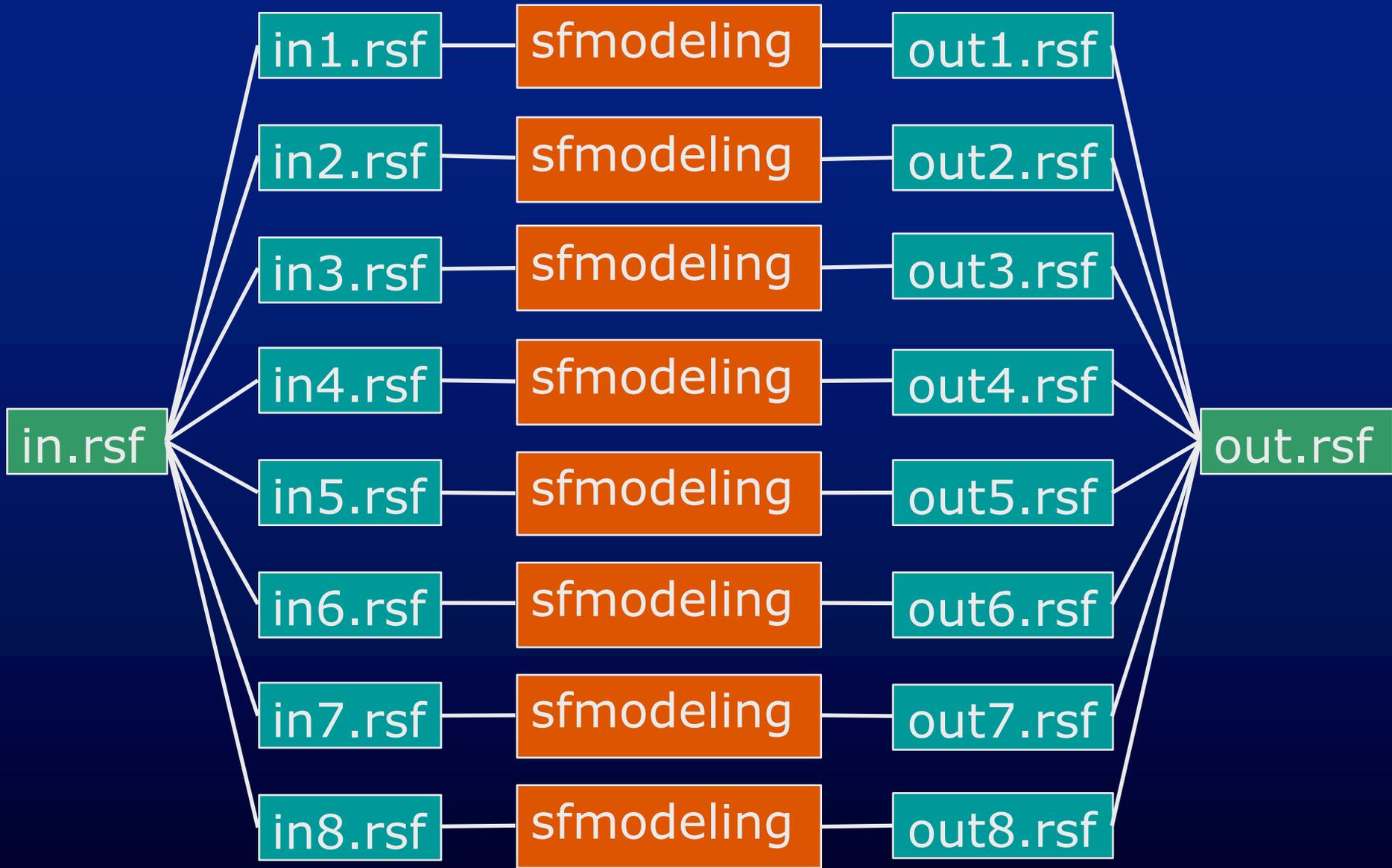
`sfdotest: L'[d]*m=1165.87`

sfdotest: fork/exec



Data-Parallel Processing

- ◆ `sfomp sfmodeling vel=vel.rsfsplit=3 < image.rsfs > data.rsfs`
- ◆ `mprirun -np 128 sfmpi split=3 \sfmodeling vel=vel.rsfs \--input=image.rsfs --output=data.rsfs`



Submitting Parallel Jobs

- ◆ **pscons**: parallel scon (scons -j)

```
Flow('data','image','modeling',split=[3,'omp'])
```

```
Flow('data','image','modeling',split=[3,'mpi'])
```

```
Flow('data','image','modeling',split=[3,1000])
```

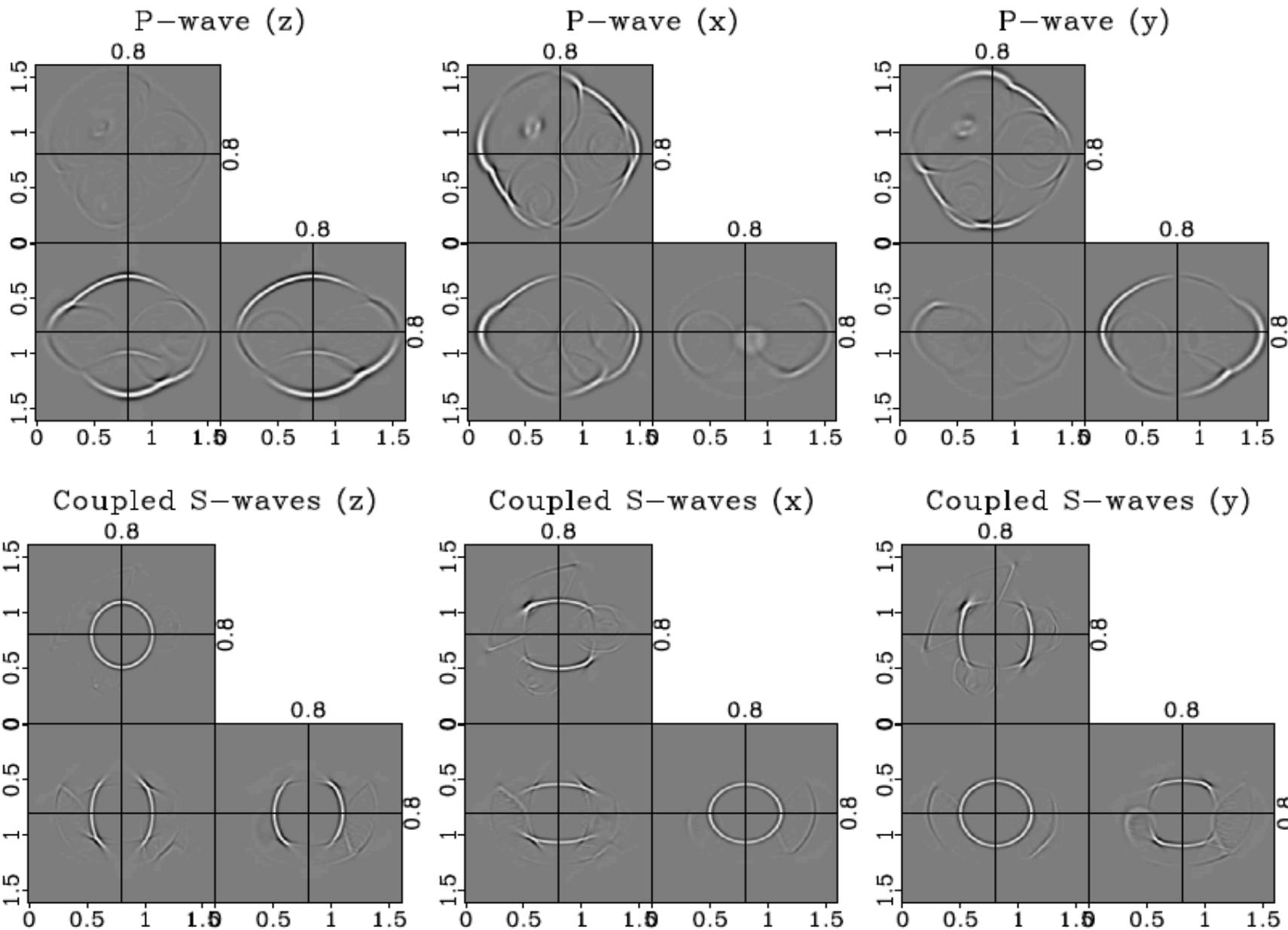
- ◆ **sfbatch**: jobs on a shared cluster

```
sfbatch exe="scons NP=100 data.rsfc"
```

```
scons BATCH=1 data.rsfc
```

What Does This Do?

```
sfbatch exe="sfconjgrad mpirun  
-np 128 sfmpi sfomp split=3  
sfmodeling vel=vel.rsfsf  
niter=100 mod=image0.rsfsf  
--input=data.rsfsf --output=image.rsfsf"
```



(Sripanich et al., 2016; Sun et al., 2016)

Conclusions

- ◆ **Madagascar is an open-source package for multidimensional data analysis and reproducible computational experiments**
 - 10 years in public existence 2006-2016
- ◆ **Madagascar objects are multidimensional data arrays**
- ◆ **Using Unix-style encapsulation for data-parallel computations**



MADAGASCAR

<http://ahay.org>