Scalable Library Loading with SPINDLE

LC User Meeting

Matt LeGendre, Wolfgang Frings, Dong Ahn,
Todd Gamblin, Bronis de Supinski, Felix Wolf

Lawrence Livermore National Laboratory

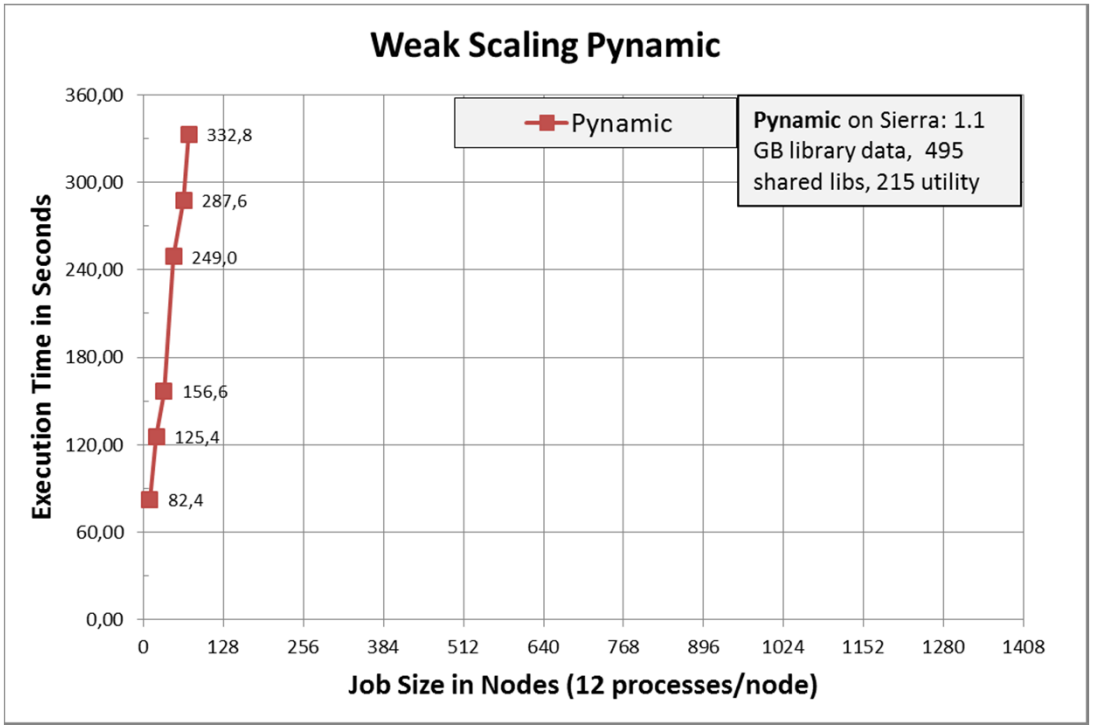
LLNL-PRES-638575

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Library Loading Causes Major Disruption at Large Scale

Lawrence Livermore National Laboratory

JÜLICH FORSCHUNGSZENTRUM



Pynamic running on LLNL Sierra Cluster
1944 nodes, 12 tasks/node,
NFS and Lustre file system

Challenges Arise from File Access Storms

Lawrence Livermore National Laboratory

JÜLICH FORSCHUNGSZENTRUM

- Caused by dynamic linker s**earching** and l**oading** dynamic linked libraries
- *File metadata operations:*
  # of tests = # of processes
  x # of locations
  x # of libraries

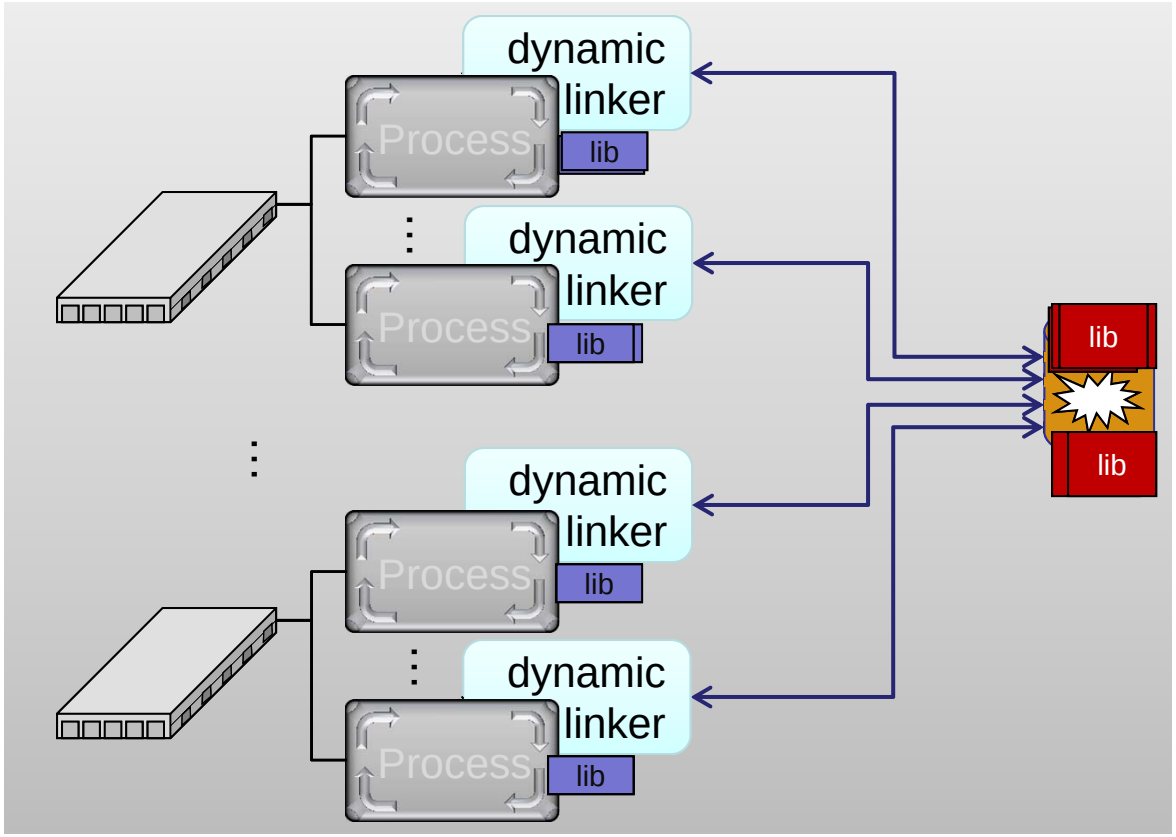
  File read operations:
  # of reads = # of processes
  x # of libraries

  - serial (1 task): 5,671 open/stat calls
  - parallel (23,328 tasks) : 132,293,088 open/stat calls

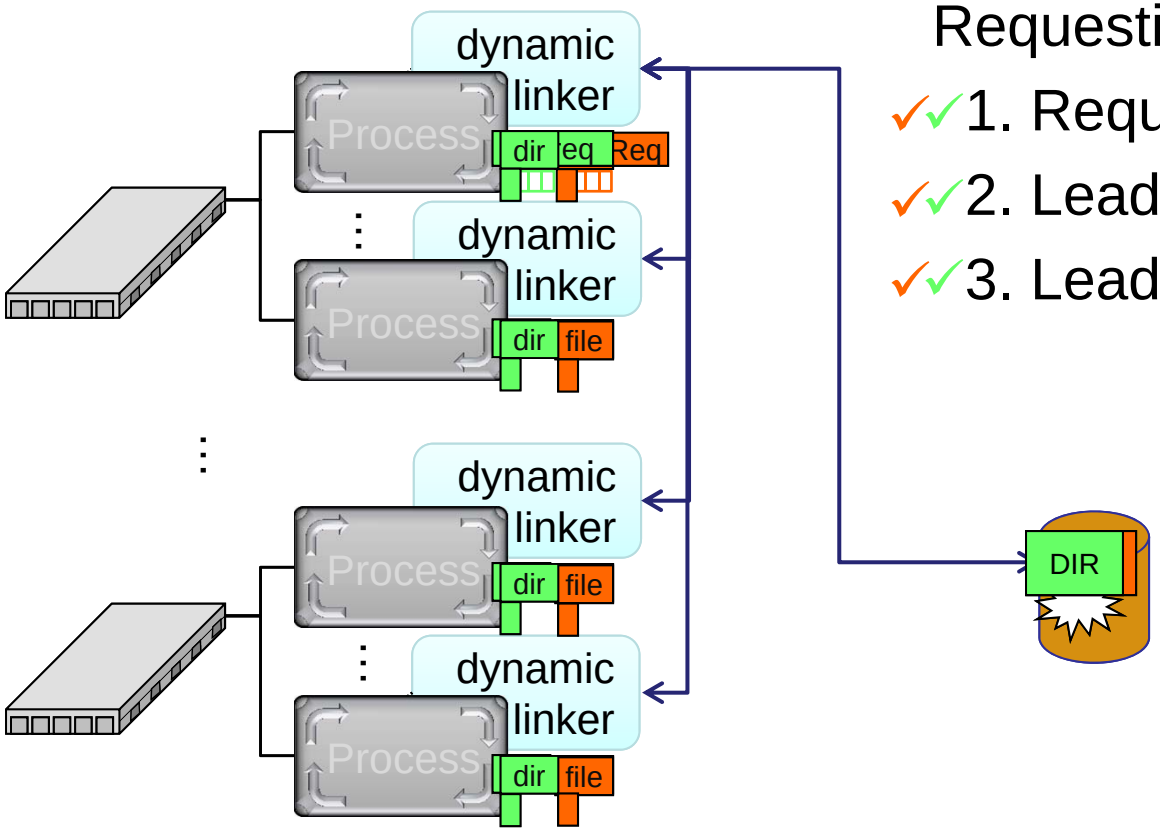
- Existing Solutions:
  - NFS Accelerators
  - Cray DVS
  - Directories of Symlinks

File Access is Uncoordinated!

- Loading is nearly unchanged since 1964 (MULTICS)
- ld-linux.so uses serial POSIX file operations that are not coordinated among process.



How SPINDLE Works



Requesting dir/file:

✓✓1. Request from leader

✓✓2. Leader reads from disk

✓✓3. Leader distributes to peers

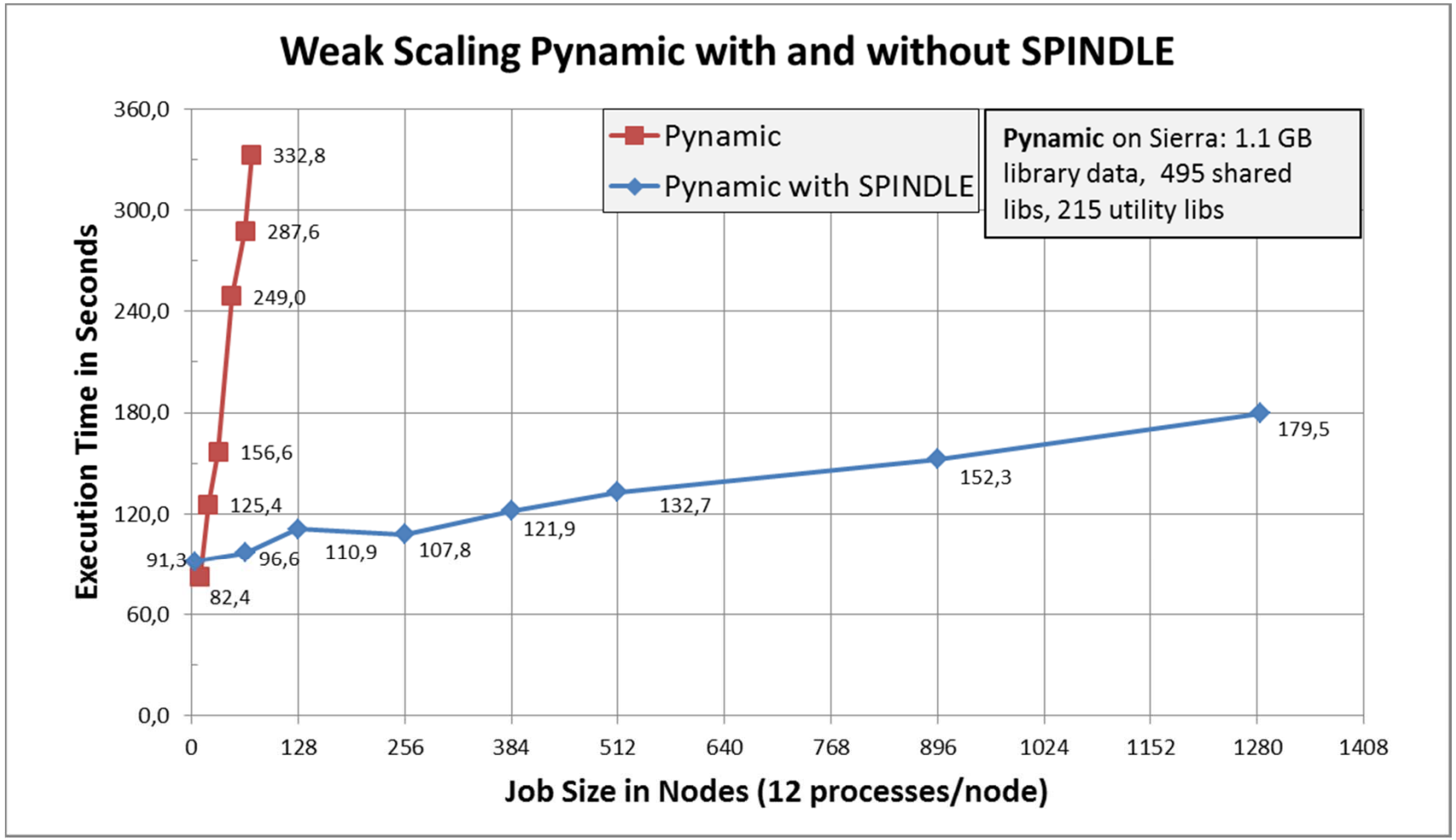
File metadata operations:
# of tests = # of locations

File read operations:
# of reads = # of libraries

Spindle Solves Scalability Problems in Application Startup

Lawrence Livermore National Laboratory

JÜLICH FORSCHUNGSZENTRUM



Launching SPINDLE

- SPINDLE wrapper call:

```
% use spindle
Prepending: spindle (ok)

% spindle srun -n 512 myapp.exe <args>
```

- Executable is not modified
- SPINDLE scalably loads:
  - Library files (from dependencies and dlopen)
  - Executable
  - Scripts
  - Python .py/.pyc/.pyo files
  - fork/exec'd processes

Conclusion

- Spindle accelerates loading of libraries and Python files at scale.
  - Ready to use on Linux/x86_64
  - BlueGene/Q Port under development
- More information:
  - Source Code: https://github.com/hpc/Spindle
  - Documentation: https://computation-rnd.llnl.gov/spindle
  - Publication: https://computation-rnd.llnl.gov/spindle/pdfs/spindle-paper.pdf
    (best paper award at ICS 2013)

Questions?

Matthew LeGendre
legendre1@llnl.gov