

# Barriers to OFED Adoption



OPENFABRICS  
ALLIANCE

Jeff Squyres, Cisco Systems

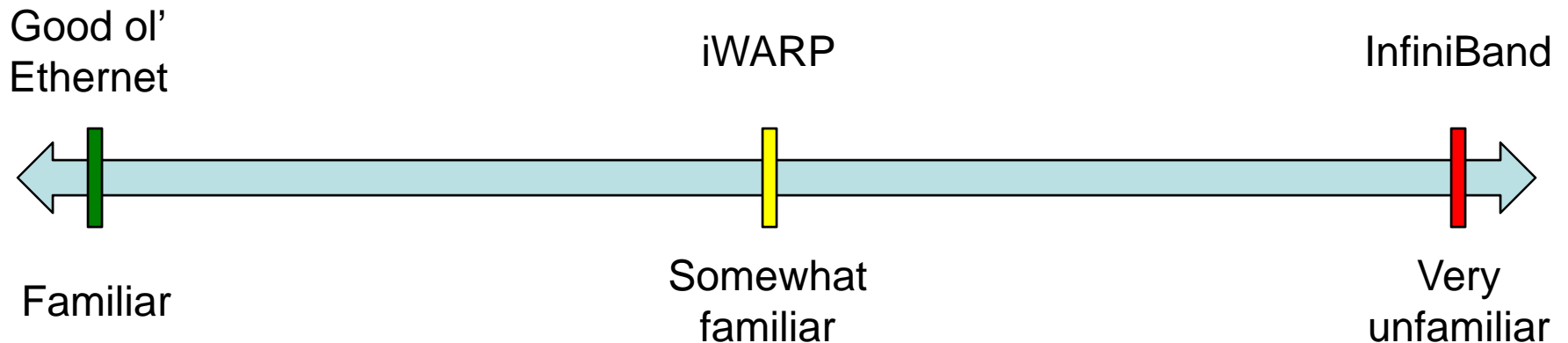
# Two main categories of problems



- Applications
- Administrators

# Common theme

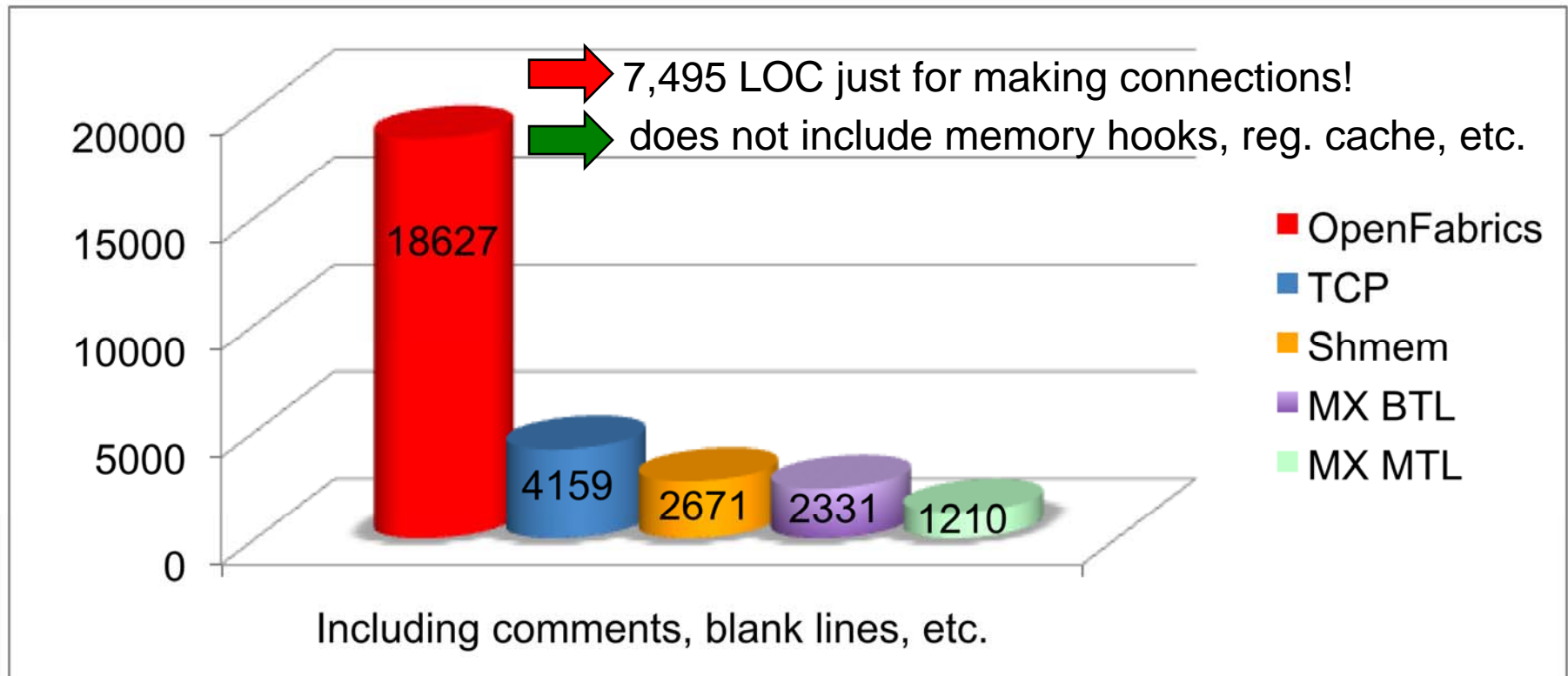
- OpenFabrics is not just different
  - It's completely, totally, utterly, wholly different (I'm not saying this is fair)



# Verbs API

- Very hard to write verbs-based applications
  - Significantly more complex than sockets
  - “Common” verbs practices are not well known
  - Different API stacks for different OS’s
  - Man pages are not sufficient documentation
  - Tutorials, books, programming workshops, etc.
- Does not address many needs of its biggest current customer (MPI)
  - See MPI Panel for more details (yesterday)

# Example: Open MPI lines of code



# Lack of (performance) tools

- No equivalents to tcpdump, Wireshark, ...
  - Cannot tell what is happening on RNIC / HCA
- Many OF tools do not work on iWARP – why?
- Much OF validation done with MPI – why?
- Network administrators are greatly hampered
  - Wholly reliant on the vendor for support
  - Use other tools (e.g., MPI) to validate the network
- Don't say: "it's open source; go look yourself"

# Lack of (performance) tools

- Great difficulty in answering the following common questions:
  - Why am I not getting full bandwidth?
  - Why is my 0-byte HRT latency so high?
  - Is the QP cache being thrashed?
  - Is there congestion in the network?
  - What is the queue depth utilization?
- There should be common OF tools that can answer most / all of these questions

# Policy enforcement

- Network cannot force traffic to be distinct
  - No way to *force* all MPI apps to specific network parameters (e.g., MPI can pick any SL it wants)
  - TCP (iWARP) has source / destination port traffic classification
- Want to *force* MPI traffic to X, I/O traffic to Y
- This is but one example (!)



# Security response

- No security team / policy in place
- Root exploit recently found in a network vendor kernel driver
  - What about the released OFEDs with this bug?
- Why doesn't [security@openfabrics.org](mailto:security@openfabrics.org) exist?
  - What is OpenFabrics' defined response?
  - When can you guarantee a fix to customers?
  - When can you guarantee a fix to OS distros?
  - How would the cross-org. coordination work?

# “OFED? What is that?”

- Many customers want an OS distribution
  - Enterprise networks and filesystems (NFS (!!)  
...and SCSI for FCoE?) should be part of the OS
  - ...so why so much effort on OFED?
- Centralized integration and testing is good
  - But OS distros re-package everything
  - They don't want or benefit from OFED integration
- Rather than have them take our table scraps,  
give them what they want

# Conclusions

- Push all the code upstream
  - Make it easy to use
  - Make it familiar to use
- Make OS's be the main distribution effort
  - Actively work to give them what they want
  - Align with their schedules
  - Align with their requirements