



# Programming for Storage

## The Need for Training Programmers for Storage

- HPC IT managers work for users who program apps
- Often performance of apps/workflows dependent on storage
- Many times best solutions would be to change the program
- Reality is app specialists intolerant of requests to reprogram for better storage performance
- That is, reprogramming for storage performance often doesn't get done

## Approach: Create tools, training to help a priori

- Give programmers libraries, performance debugging tools that avoid or detect poor storage patterns
- Give tutorials, case studies, help pages showing weak programming approaches and how to improve them



# Example from BioInformatics

Pseudo code example from IT manager -- single thread

```
for( I=0, I<1000, I++){  
    for( J=0, J<1000, J++){  
        buf = compute (I,J);  
        f = open( "file_foo");  
        lseek(f, offset(I,J));  
        write(f, buf, lengthof(buff))  
        close(f);  
    }  
}
```

- Buf turns out to be small, unaligned, fixed length
- Obvious fixes:
  - Open/close outside both loops
  - Malloc sizeof 1000000\*lengthof(buff), copy into it in memory, one write at end



# Programming for Storage Panel

Panelists challenge: 5 dumbest examples they've seen

- One example at a time, round robin

Alternative: what should be the content of our training?