

-
-
-

Introduction

Compute and data intensive jobs perform best when allowed to run without undue regard to storage constraints.

One solution to this is to virtualize the mass storage system.



Problem

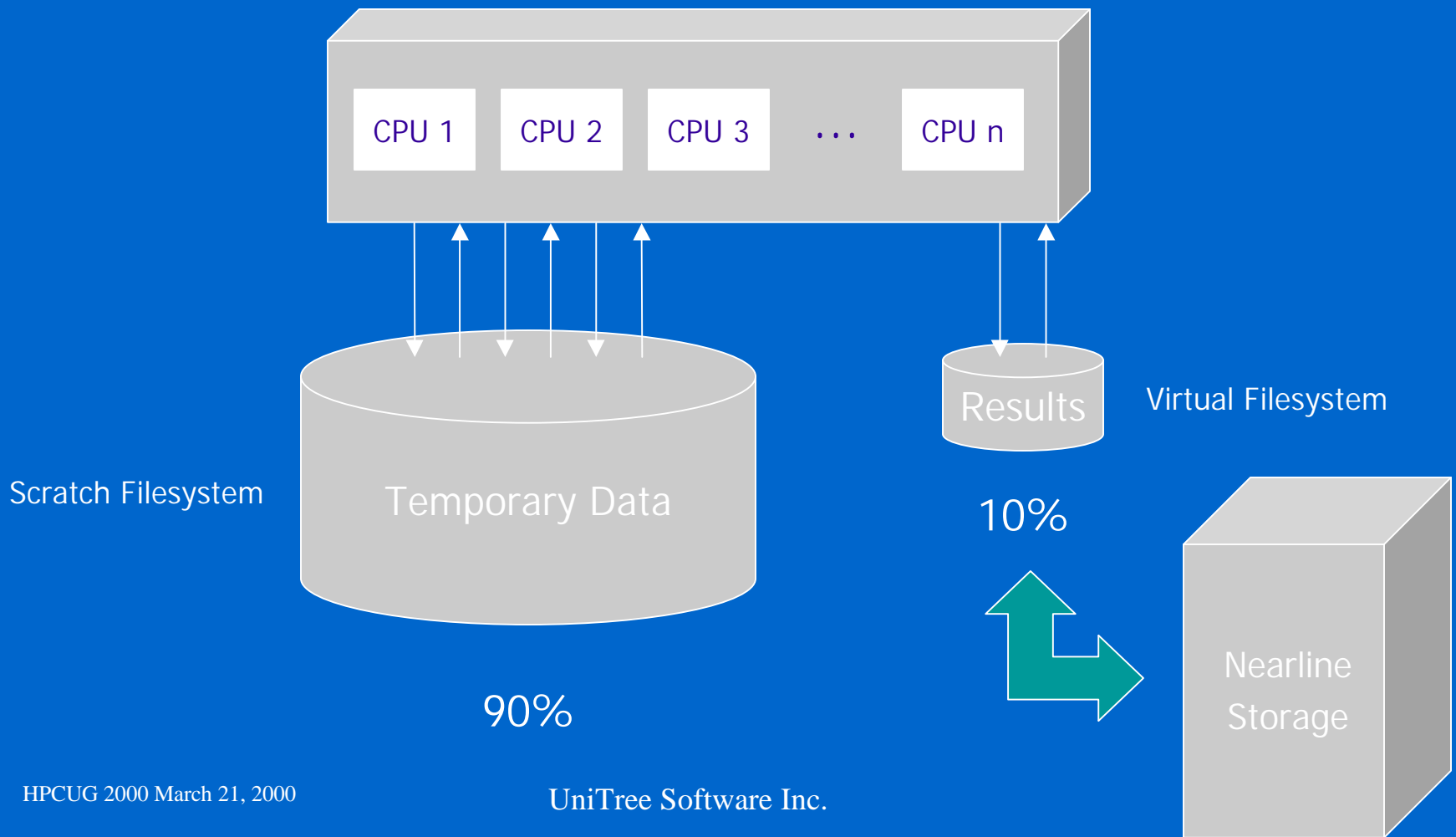
- Sizeable temporary data space requirement.
- Resulting summary data accumulates quickly.
- Subsequent jobs reference results from previous runs.
- Multiple job submissions compete for limited I/O bandwidth.



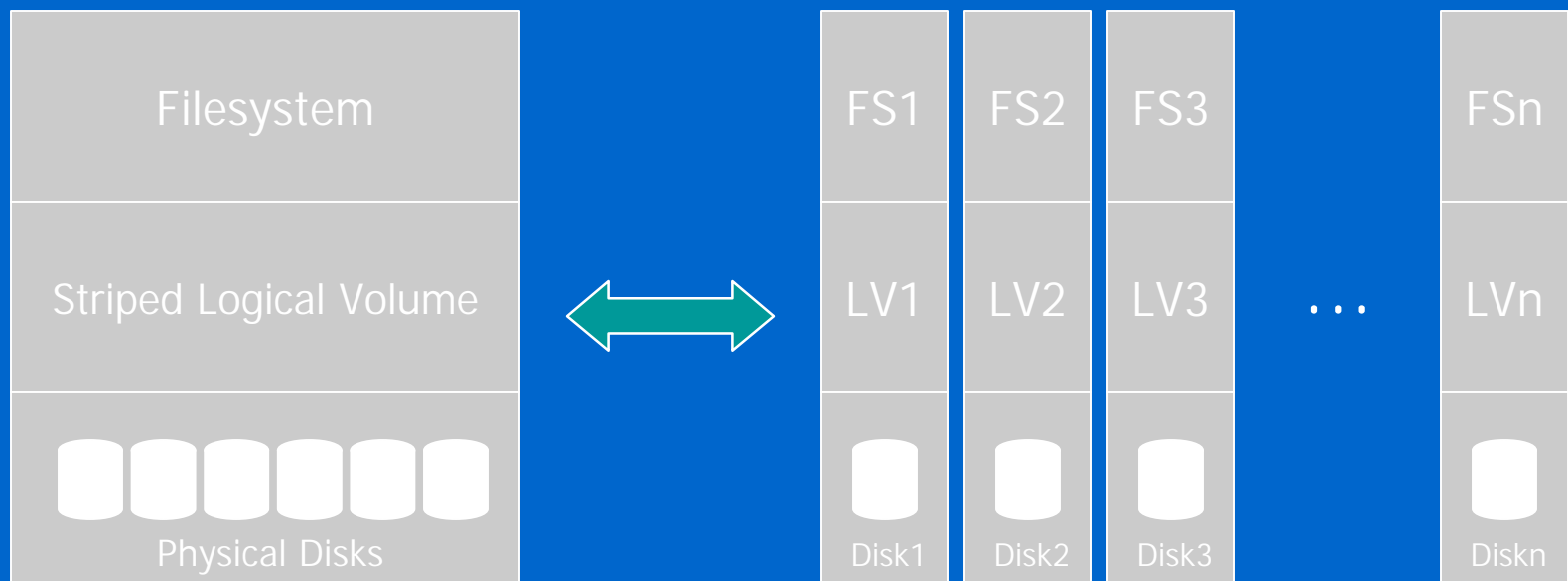
Divide and Conquer

- Scratch filesystem optimized for temporary data space.
 - Maximum space.
 - Tune for optimal I/O performance.
- Store resulting summary data in a virtual filesystem. The 80/20 rule is key.
 - Requires less physical disk space.
 - Tune for virtual storage.

Storage Architecture



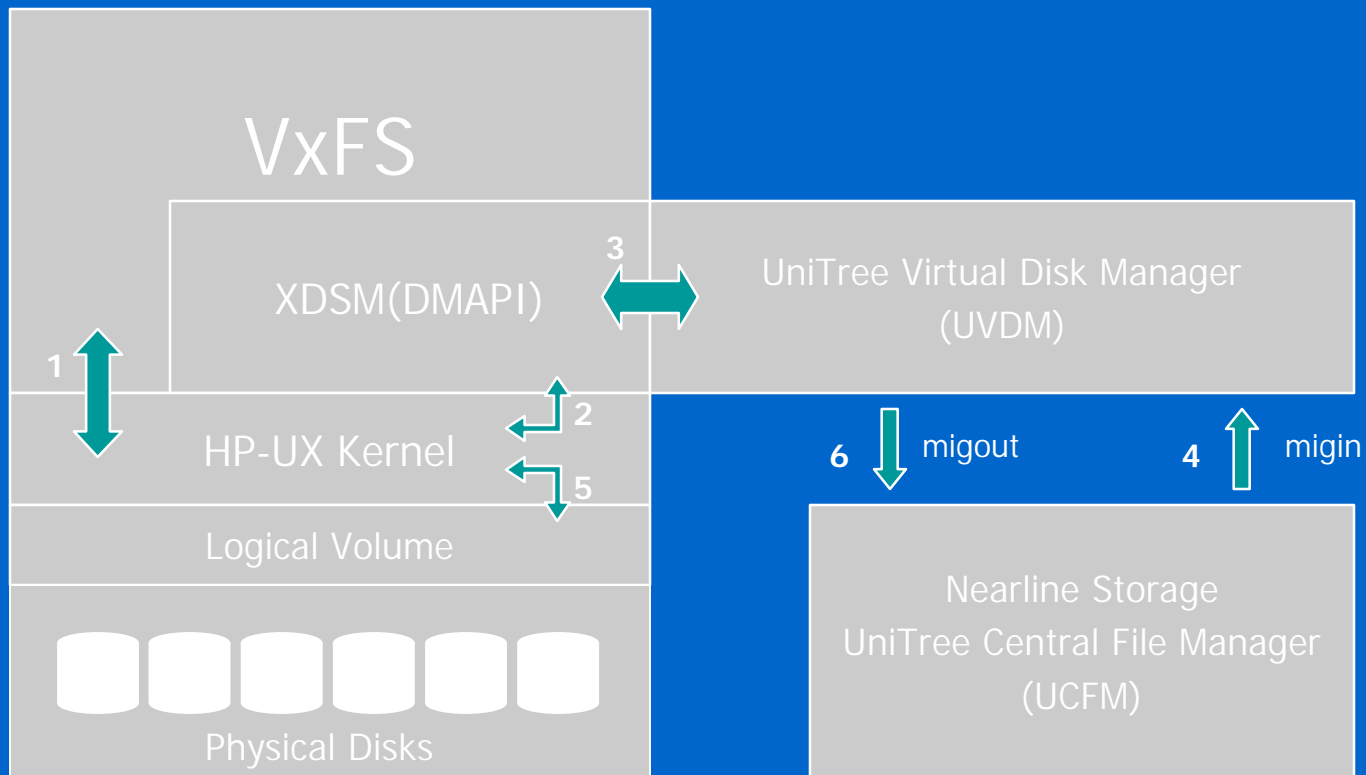
Scratch Filesystem(s)



Single Stream High Bandwidth

Total Aggregate High Bandwidth

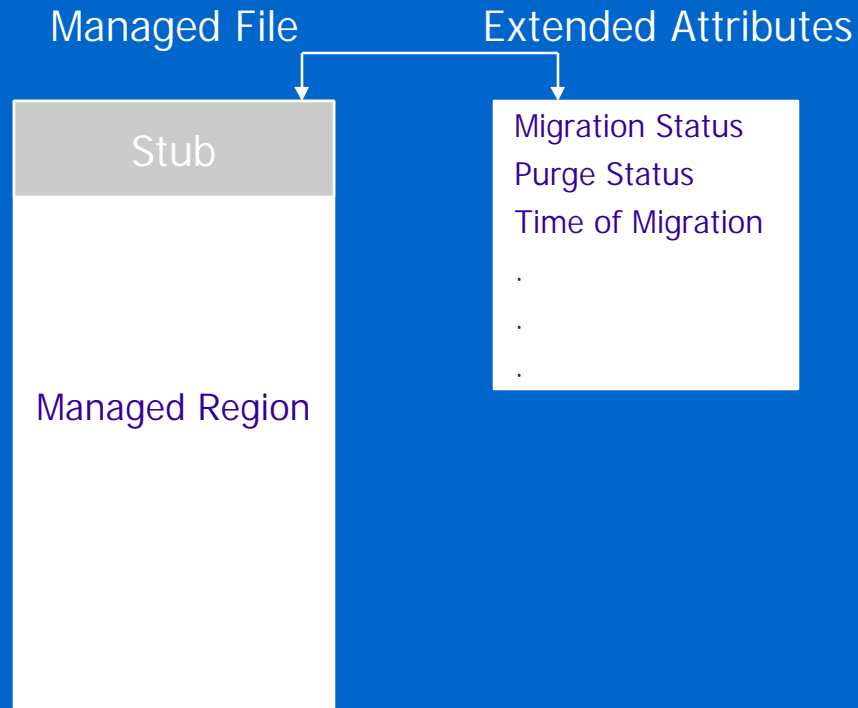
Virtual Filesystem



Virtual Filesystem Events

1. File Access Event (Read, Write, Truncate, Destroy, etc.)
2. File Data not on physical disk. Access blocks and DMAPI event is generated.
3. UVDM translates local file request into nearline storage access.
4. UVDM causes data retrieval from nearline storage (UCFM) into DMAPI layer.
5. File Data now on physical disk and access resumes.

Anatomy of a UVDM managed File





Summary

- Separate transient data from result data.
- Tune and manage for I/O speed.
- Tune and manage for permanent storage.
- Remember the 80/20 rule of data access.

