The SIOX Project Scalable I/O for Extreme Performance

Michaela Zimmer

Universität Hamburg

11th HLRS/hww Workshop on Scalable Global Parallel File Systems May 9th, 2012



The Project

Two Views, Two Questions



Two questions to HPC-I/O systems:

• User's view

• System manager's view

Michaela Zimmer (Universität Hamburg)

Two Views, Two Questions



Two questions to HPC-I/O systems:

• User's view

"Why is my program's I/O so slow?"

System manager's view

Two Views, Two Questions



Two questions to HPC-I/O systems:

• User's view

"Why is my program's I/O so slow?"

• System manager's view

"Where does my system's performance go?"



Our Answer

We strive to provide answers:



Michaela Zimmer (Universität Hamburg)

The SIOX Project

May 9th, 2012 3 / 16

Partners and Funding















Bundesministerium für Bildung und Forschung

- Funded by the BMBF Grant No.: 01IH11008B
- Start: Juli 1st, 2011
- Duration: 36 Months

Goals of the SIOX Project



Access and performance data will be

- recorded
 - on all levels
 - in a level-independent format
- causally correlated
- structured and archived



Goals of the SIOX Project



Access and performance data will be

- recorded
 - on all levels
 - in a level-independent format
- causally correlated
- structured and archived

SIOX will use the data collected to

- predict possible performance
- evaluate actual performance
- locate and analyse bottlenecks
- suggest optimized parameters



Our Approach

- Instrumentations of
 - POSIX
 - HDF5
 - MPI-IO
 - A popular file system of one of our partners'
- Later:
 - Real-world climate modelling applications from DKRZ
 - More file systems, drivers, components,...
 - Your application here

Challenge: Multitude of Components

Hardware components, e.g.

- SSDs
- HDDs
- CPUs
- Infiniband
- Ethernet
- . . .

Software components, e.g.HDF5MPI-IO

- GPFS
- Lustre
- POSIX

• . . .

... all with different capabilities, performance metrics and interfaces.

Solution: Abstract System View



- Split components into nodes by functionality
- This covers all file systems, components and architectures!
- One high-level API per node type
- No API modifications necessary
- May use existing client-side instrumentation



,____X

Michaela Zimmer (Universität Hamburg)

Challenge: Causal Chains



Solution: Reconstructing the Causal Tree



Michaela Zimmer (Universität Hamburg)

Solution: Reconstructing the Causal Tree



Solution: Reconstructing the Causal Tree



- Trace *descriptors*
 - vertical transfers
 - horizontal translations

Michaela Zimmer (Universität Hamburg)

Challenge: Diversity of Performance Metrics



The SIOX Project

May 9th, 2012 11 / 16

Solution: An Ontology for Performance Metrics



Michaela Zimmer (Universität Hamburg)

The SIOX Project

Challenge: Data Deluge

- Logging...
 - ... every access call...
 - ... on *every* node...
 - ...*all* the time...
 - ... will produce a prohibitive amount of data!
- System performance should not suffer (too much)!

Solution: Scalable and Flexible Architecture



- Load deferral (collection online, processing removed/offline)
- Compression (trace-based solutions)
- Selective logging (intelligent nodes)

Michaela Zimmer (Universität Hamburg)

The SIOX Project

Our Vision

- A system that will
 - collect and analyse
 - access patterns and
 - performance metrics
- in order to
 - assess current and possible system performance,
 - locate and diagnose problems and
 - suggest solutions and improvements.

Finally: SIOX and You



Arbeitspakete vorgestellt waren, hatten wir auch noch Zeit für ein erstes Brainstorming.

Read more Log in or register to post comments

- Think we missed a problem?
- Think you could solve one?
- Like to see SIOX on your favourite file system?

We cordially invite you to become involved at

http://www.HPC-IO.org



Michaela Zimmer (Universität Hamburg)

The SIOX Project