Department of Computer Science





# The Virtual Institute of I/O



Limitless Storage Limitless Possibilities https://hps.vi4io.org

Julian M. Kunkel

Supercomputing 2019

2019-11-19

Copyright University of Reading

LIMITLESS POTENTIAL | LIMITLESS OPPORTUNITIES | LIMITLESS IMPACT

Introduction

# Reading

#### Goals of the Virtual Institute for I/O

- Provide a platform for I/O enthusiasts for exchanging information
- Foster training and collaboration in the field of high-performance I/O
- Track and encourage the deployment of large storage systems
  - Hosting information about high-performance
  - Supporting the IO-500 and the Data Center List

https://www.vi4io.org



## **Open Organization**



- The organization uses a wiki as central hub
  - Everybody (registered users) can edit the content
  - Major changes should be discussed (see below)
  - The wiki uses tag clouds to link between similar entities
- Supported by mailing lists
  - Call-for-papers
  - Announce list for relevant information
  - Contribute list to discuss and steer organizational issues
- Major changes should be discussed on the contribute mailing list
- Members can vote for changes

#### Everybody is welcome to participate

## Outline



#### 1 Overview

### 2 Comprehensive Data Center List (CDCL)

Julian M. Kunkel VMO

# Comprehensive Data Center List (CDCL)



The CDCL contains system characteristics for sites, supercomputer and storage

#### System Model

- Based on an extensible JSON schema, optimized editor
- Supports now (all) logical components and subcomponents
- Characteristics and peak values
- Measured values \*-500

#### Components with characteristics

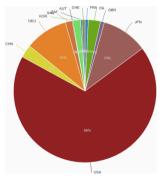
- Site, supercomputer, online storage, tape archives
- Compute nodes, storage nodes, local storage, accelerators, …
- Supporting: e.g., CPU type, memory available, ...

#### Overview 00

# CDCL Storage List 2019

### Features

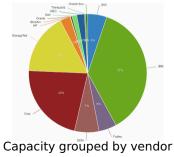
- Table view with selectable columns
- Flexible metrics and aggregation



Capacity grouped by country

#	site.institution	site.storage system.net capacity	site.supercomputer.compute peak	site.supercomputer.memory capacity
		in PiB	in PFLOPS	in TB
1	Oak Ridge National Laboratory	250.04	220.64	3511.66
2	National Energy Research Scientific Computing Center	197.65	37.71	857.03
3	Los Alamos National Laboratory	72.83	11.08	2110.00
4	German Climate Computing Center	52.00	3.69	683.60
5	Lawrence Livermore National Laboratory	48.85	20.10	1500.00
6	RIKEN Advanced Institute for Computational Science	39.77	10.62	1250.00
7	National Center for Atmospheric Research	37.00	5.33	202.75
8	National Center for Supercomputing Applications	27.60	13.40	1649.27
9	Global Scientific Information and Computing Center	25.84	17.89	275.98
10	Joint Center for Advanced HPC	24.10	24.91	919.29

37	Pacific Northwest National Laboratory	2.40	3.40	184.00
38	Navy DoD Supercomputer Resource Center	2.11	2.05	0.00
39	Vienna Scientific Cluster	1.81	0.68	42.18
40	Center for Scientific Computing	0.75	0.51	77.57
41	University of Bristol	0.44	0.38	43.01



#### LIMITLESS POTENTIAL | LIMITLESS OPPORTUNITIES | LIMITLESS IMPACT 6/7



### New: Unified data center representation for webpages

- Provide a HTML stub for embedding into any data center web page
  - ⇒ Towards a standardized presentation of systems!
  - Allowing the site to describe and visualize their system
  - Hosted by the site directly
  - Allowing a simple export into VI4IO data center list
- Editor: https://www.vi4io.org/hpsl/addpage

#### Others

CDCL: form to create stubs for pages: <a href="https://www.vi4io.org/cdcl-add">https://www.vi4io.org/cdcl-add</a>