



中国科学院超级计算中心
Supercomputing Center of Chinese Academy of Sciences

Lustre在中科院超级计算中心的应用

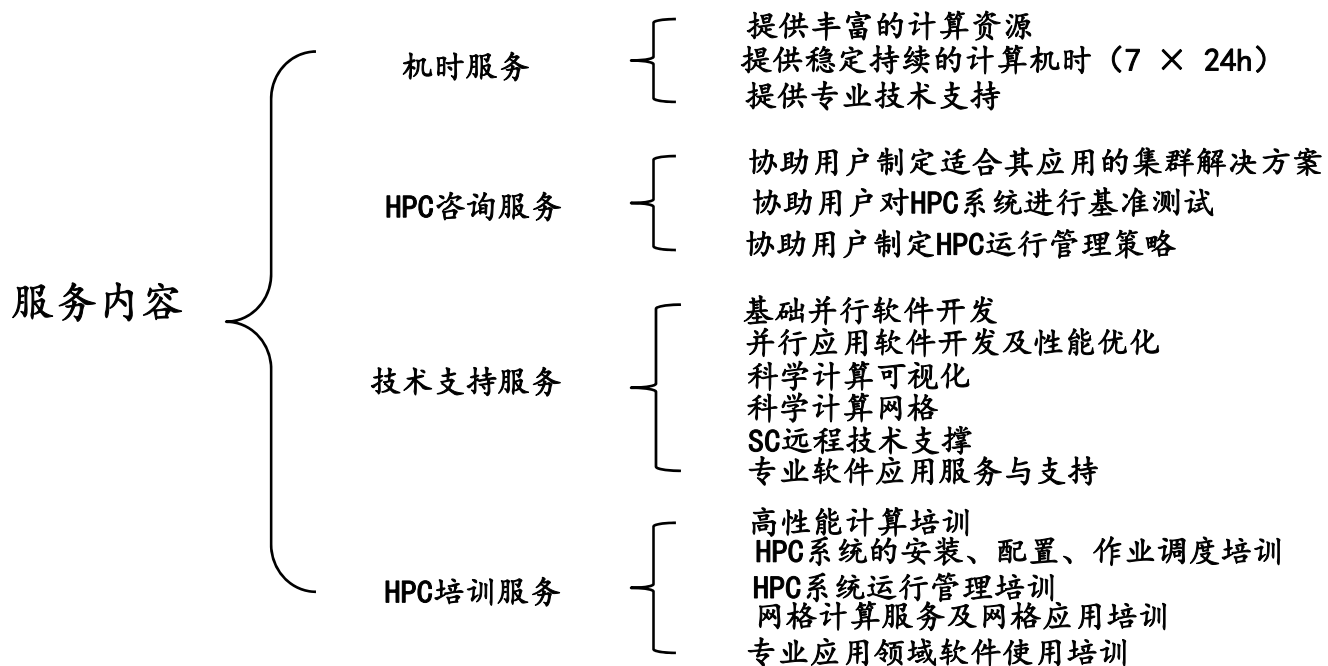
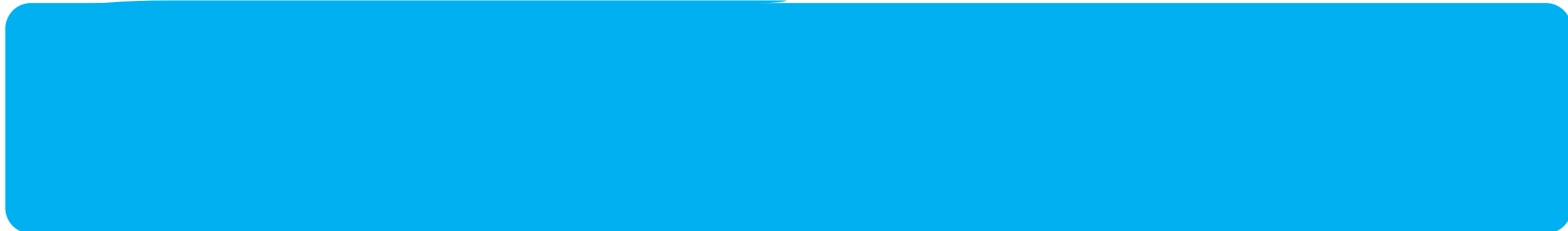
Nov 26, 2011



Outline

- 中科院超级计算中心简介
- Lustre在超级算中心的部署
- Lustre在超级算中心的应用
- Lustre在实际应用中的问题
- 高性能计算对未来文件系统的需求

中国科学院超级计算中心



中国科学院超级计算中心

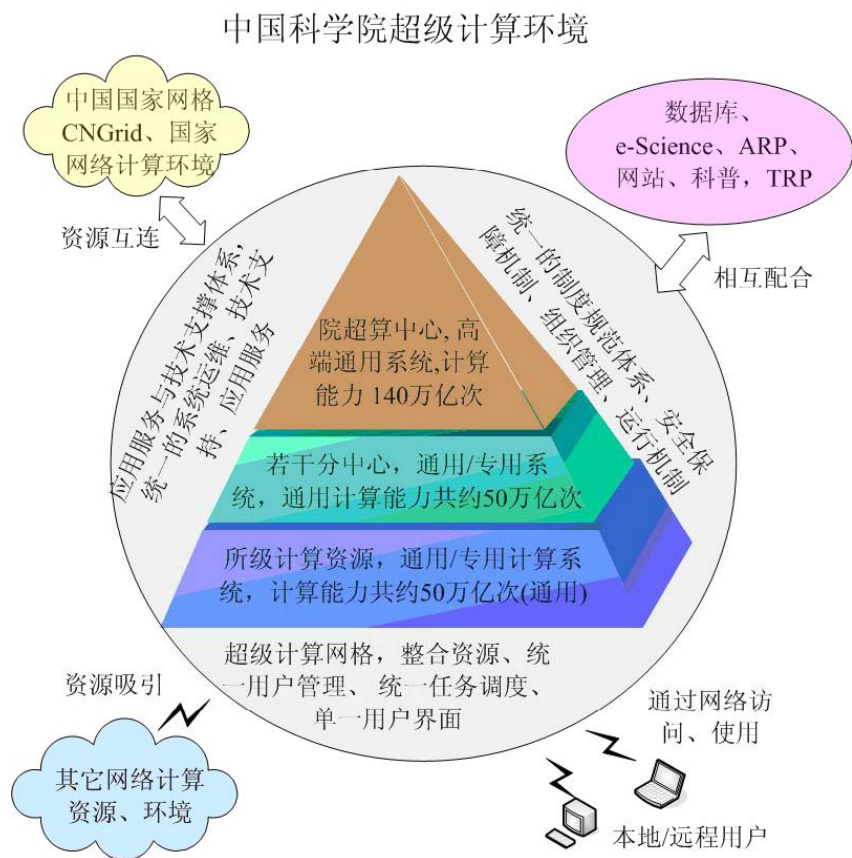
中国国家网络运行管理中心 (CNGrid)



中科院超级计算环境 (CScGrid)



中国科学院超级计算中心



在超级计算环境建设方面，针对我院研究所在各领域的超级计算需求，提出了“优化布局，整合资源，强化服务，推动应用”的建设原则和思路，拟采用院超级计算总中心、分中心、及所级中心的三层网络结构布局，以实现统一的运维管理与资源调度，统一为用户提供服务，方便用户使用，达到资源整合与共享使用、提高资源利用率的目的，促进应用发展。

作为三层网络结构的最高层，院超级计算总中心将建设成为国际先进、国内一流的超级计算运行服务中心，成为国家科技信息化基础设施的重要组成部分和培养超级计算应用人才的基地，成为引领我国超级计算应用与发展的主要力量。

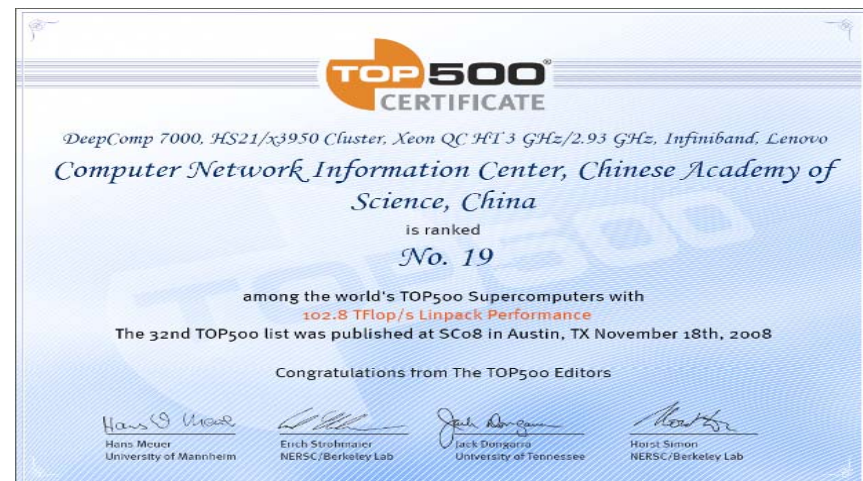
DeepComp 7000



- **DeepComp 7000**

- Hybrid architecture (Xeon + Itanium2)
- More than 12,000 cores
- Peak performance: 146TFlops
- More than 66TB memory
- 350TB on-line storage, 1PB off-line storage
- 4X DDR Infiniband Interconnected
- Lustre, SNFS, IBM GPFS, SGI CXFS file system
- Linux OS

- ❑ TOP500: No.19 (2008)
- ❑ 375 users up to Sep 30, 2011
- ❑ Utilization: More than 70%





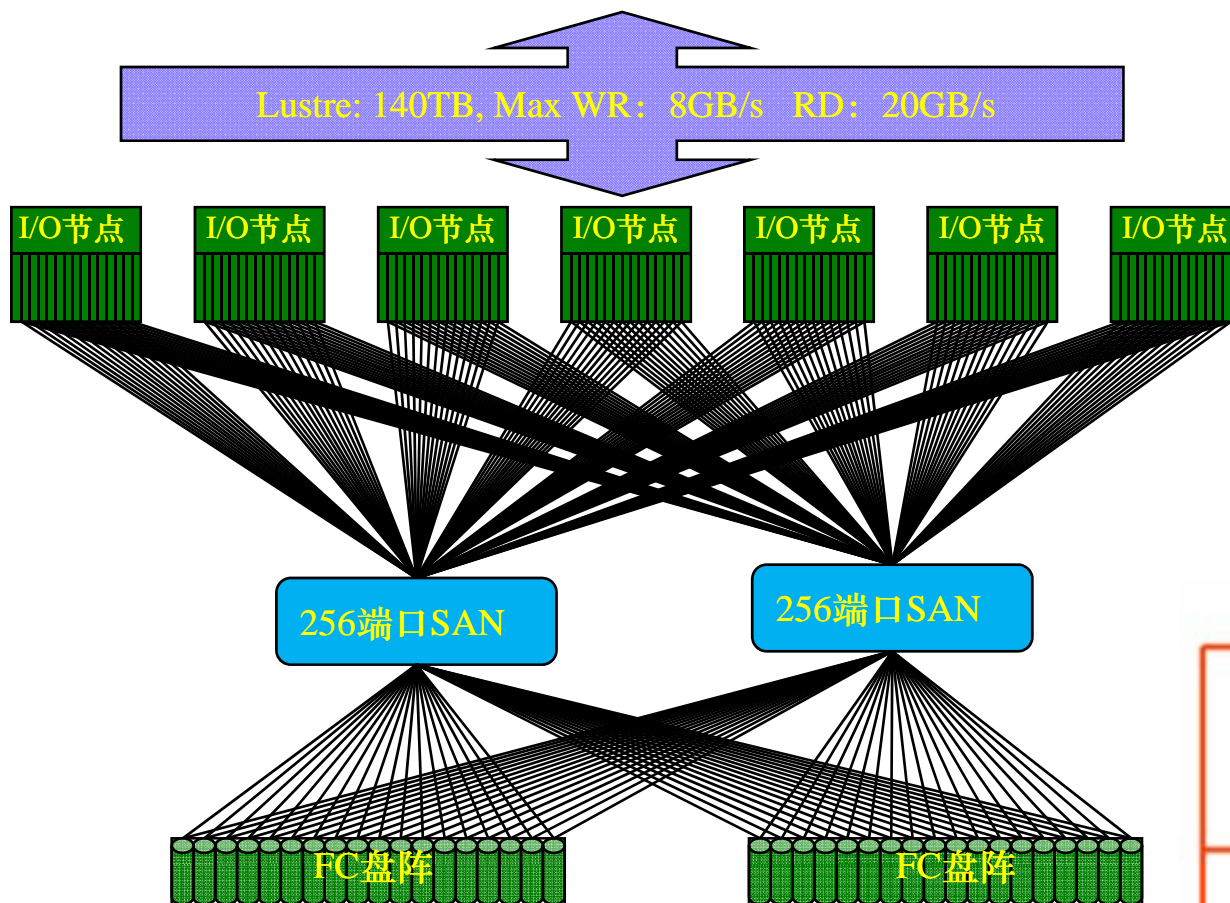
Outline

- 中科院超级计算中心简介
- Lustre在超级算中心的部署
- Lustre在超级算中心的应用
- Lustre在实际应用中的问题
- 高性能计算对未来文件系统的需求

Lustre在超级算中心的部署

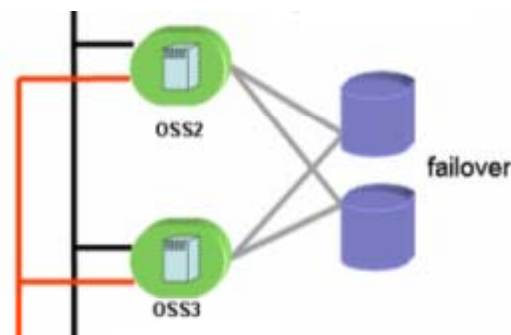
客户端: 1140 计算刀片, 38 厚节点, 2台胖节点

Lustre: 140TB, Max WR: 8GB/s RD: 20GB/s



印象:

- 良好I/O吞吐能力
- 超出想象的稳定
- 具备自我恢复能力





Outline

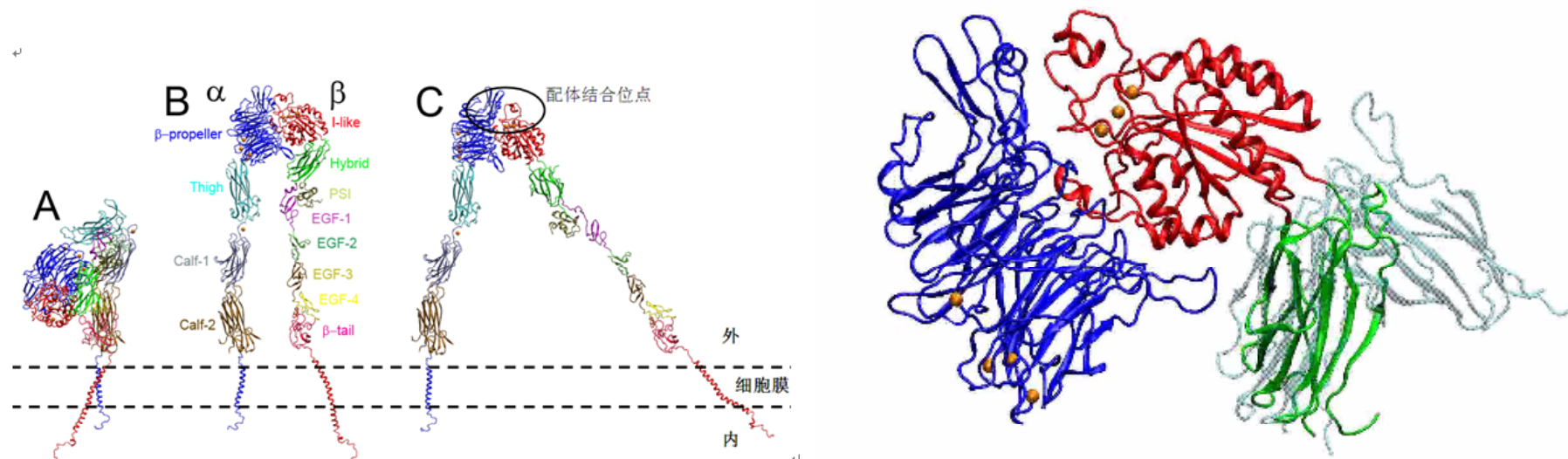
- 中科院超级计算中心简介
- Lustre在超级算中心的部署
- Lustre在超级算中心的应用
- Lustre在实际应用中的问题
- 高性能计算对未来文件系统的需求



宇宙演化模拟



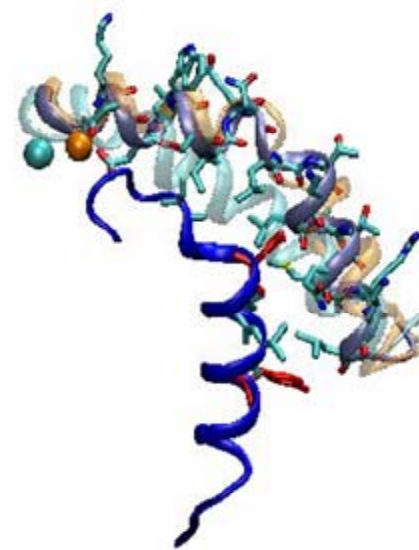
整合素 (Integrin) 配体结合位点激活的分子机制研究



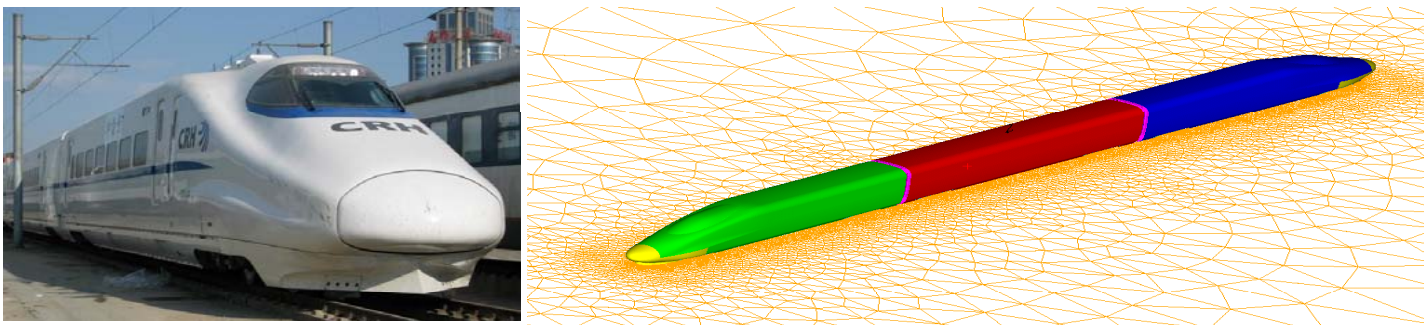
整合素的结构及两种构象

整合素是表达在细胞表面的一种异构二聚体，它们介导了细胞间以及细胞与胞外基质间的相互作用，同时可以在细胞内外双向传递信号。

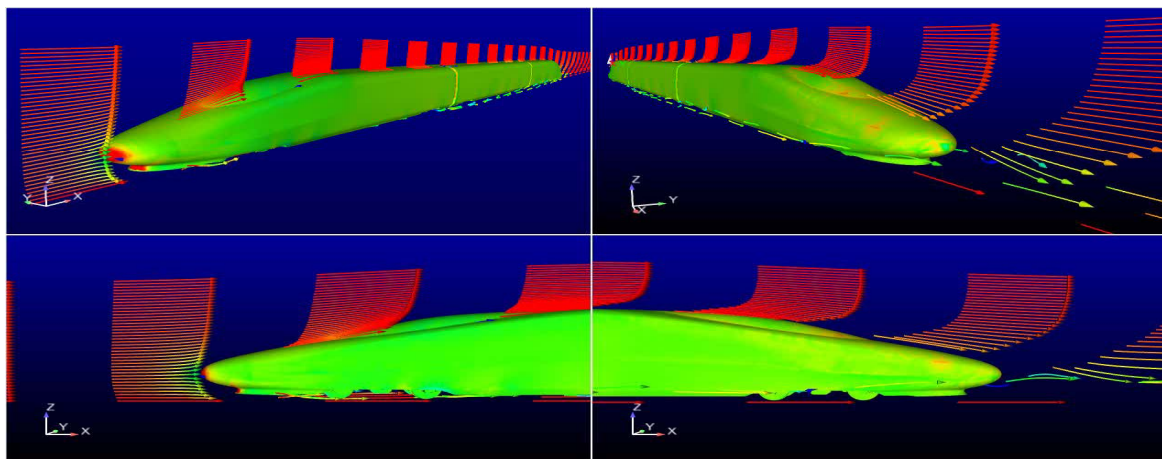
本研究确定了对构象变化和配体结合亲和力具有决定性作用的氨基酸残基，从而也就确定了可能的药物靶向位点，针对整合素激活紊乱引起的疾病的药物设计提供了理论基础。



高速列车气动噪声数值评估



- 在时速大于350公里的情况下，气动噪声成为高速列车的主要噪声源
- 研究内容：通过数值模拟手段，结合理论和实验研究，对高速列车的气动噪声进行预测评估，提出列车关键部件的降噪优化设计方案
- 目的：提高高速列车乘坐的舒适性、缓解噪声对铁路沿线居民的影响、促进高速轨道交通的可持续发展。





Outline

- 中科院超级计算中心简介
- Lustre在超级算中心的部署
- Lustre在超级算中心的应用
- Lustre在实际应用中的问题
- 高性能计算对未来文件系统的需求

Lustre在实际应用中的问题

▣ Quota丢失

现象：用户的空间没有达到quota，但无法写入

解决：重新设置quota

▣ Lustre Recovery 时间过长

现象：160个OST，任何一个故障再恢复的过程持续至少20分钟

解决：无

Lustre在实际应用中的问题

□ 元数据性能瓶颈

现象：元数据处理效率不高，重IO负载时产生性能瓶颈

解决：无

□ 日志不易理解

现象：

```
000002:00080000:5:1319780055.816554:0:4362:0:(mdc_request.c:1444:mdc_init_ea_size()) updating max_mdsize/max_cookiesize: 3888/5120
```

```
00000100:02000000:5:1319780055.816556:0:4362:0:(import.c:1335:ptlrpc_import_recovery_state_machine()) CAS-OST0055-osc-ffff81081561d400: Connection restored to service CAS-OST0055 using nid 10.1.6.32@o2ib.
```

```
00000400:00000100:0:1319780055.835259:0:4313:0:(lib-move.c:1782:lnet_parse_put()) Dropping PUT from 12345-10.1.6.32@o2ib portal 4 match 1382649099701917 offset 192 length 192: 2
```



Outline

- 中科院超级计算中心简介
- Lustre在超级算中心的部署
- Lustre在超级算中心的应用
- Lustre在实际应用中的问题
- 高性能计算对未来文件系统的需求

高性能计算对未来文件系统的需求

- 具有更高的性能
 - 支持更多的OST
 - 提高元数据的执行效率和性能

- 支持分布式容错，使lustre具有更高的可靠性
 - 任一OST故障都会导致数据丢失
 - OST备份几乎不可操作

- 支持数据的分级存储
 - 数据在不同级别存储设备上自动迁移，降低存储成本

- 支持故障的快速恢复

对Lustre未来的担忧

□ 版权及后继版本

The screenshot shows the Whamcloud website for Lustre. The top navigation bar includes the Lustre logo, a search bar, and links for 'Log In / Create Account' and 'Help'. A secondary navigation bar contains 'Page', 'Discussion', 'View Source', and 'History'. On the left, a sidebar lists navigation options: 'Main Page', 'Download', 'Learn', 'Use', 'Contribute', and 'Get Involved'. Below this is a 'Wiki Toolbox' with links for 'What Links Here', 'Related Changes', 'Special Pages', 'Printable Version', and 'Permanent Link'. The main content area features a 'Download' section with a blue oval highlighting the text: 'Lustre™ is a scalable, secure, highly-available cluster file system. It is designed, developed and maintained by Oracle Corporation. [Learn More](#)'. Below this is a large banner for 'About whamcloud' and 'Lustre® High Performance Parallel File System'. The 'Whamcloud Philosophy' section lists: 'Vendor Neutrality', 'Customer Choice', 'Open Development', 'Community Contribution', 'Feature Development', and 'Commitment to Code Stability'. The 'Whamcloud - Lustre Your Way' section describes the company and its commitment to the community. A red oval highlights the 'Lustre Support and Feature Development' section, which states: 'Whamcloud experts ensure your Lustre solution is not only deployed rapidly but supported and maintained in your mission critical environment. New features, ideas conceived by the customers and the community, are continuously developed - enhancements that contribute to the viability and stability of Lustre.' Below this are sections for 'Whamcloud Commercial Partners' (listing HPC, ALCES, Bull, DataDirect, Eucalyptus Systems, Hitachi Data Systems, INET, NetApp, RAID, ParTec, and Redhat) and 'Whamcloud Community Partners' (listing EDFS and OpenSFS). The footer contains the address '695 San Ramon Valley Blvd., Suite 201, Danville, CA, 94582', phone number '+1-925-272-9557', website 'www.whamcloud.com', email 'Info@whamcloud.com', and version 'rev. 09/23/11'.



Thanks!