

LA-UR-15-27170

Approved for public release; distribution is unlimited.

Title:	Solving Complex Problems with High Performance Computing, Network and Storage at Los Alamos National Laboratory
Author(s):	Coulter, Susan K. Grider, Gary Alan
Intended for:	Missouri Science and Technology Career Fair, 2015-09-21/2015-09-23 (Rolla, Missouri, United States)
Issued:	2015-10-01 (rev.1)

Disclaimer:

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the Los Alamos National Security, LLC for the National Nuclear Security Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396. By approving this article, the publisher recognizes that the U.S. Government retains nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.



Solving Complex Problems with High Performance Computing, Networking and Storage at Los Alamos National Laboratory

Susan Coulter, Gary Grider

September 21, 2015

LA-UR-15-27170

UNCLASSIFIED



Outline

Who are we ?

Where are we

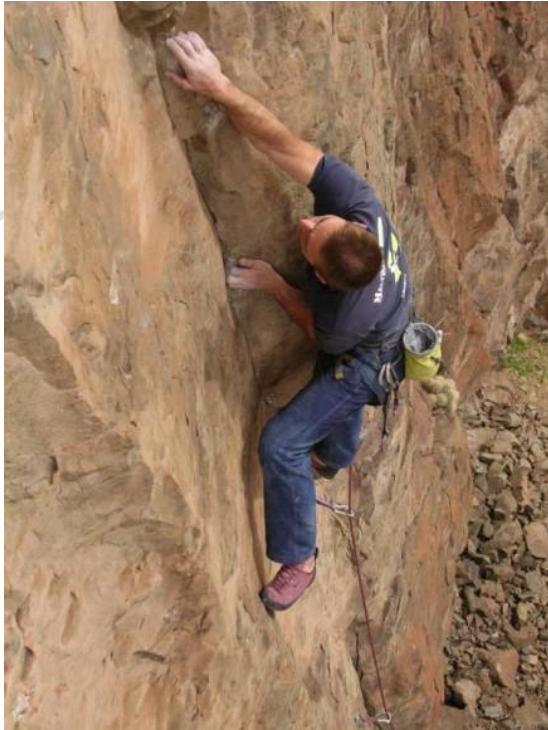
What do we do ?

Why is HPC cool ?

UNCLASSIFIED

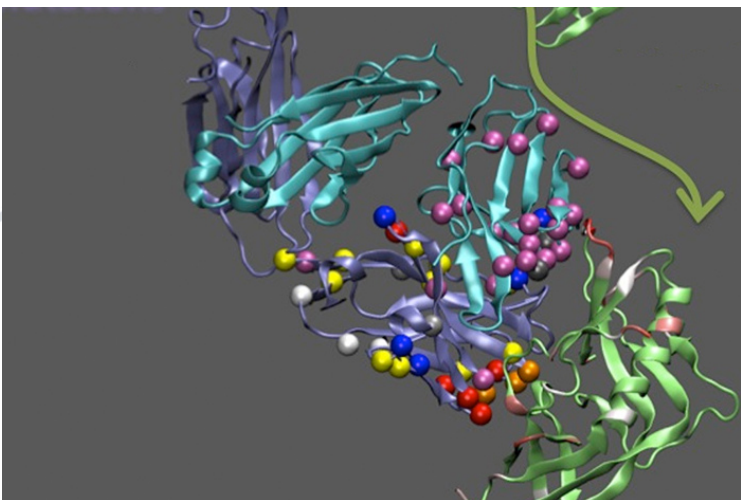
Where is Los Alamos National Laboratory ?



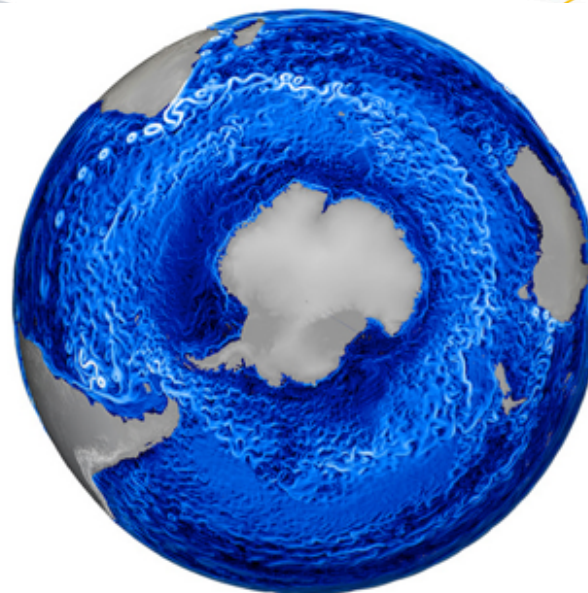


What is there
to do around
Los Alamos?

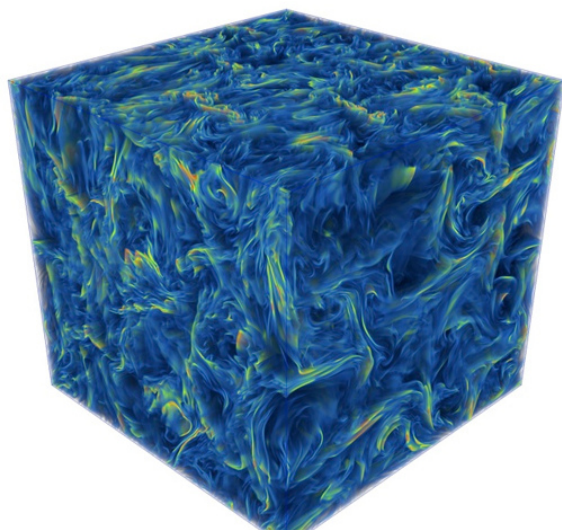
UNCLASSIFIED



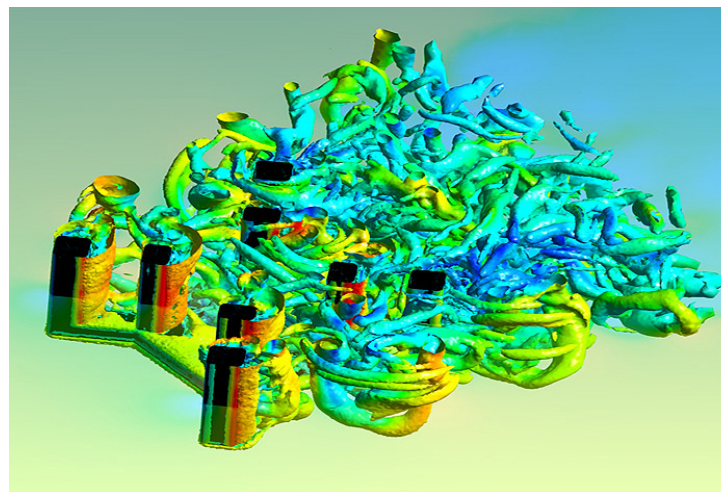
AIDs Virus Mutation Patterns



Ocean Currents
And Eddies
Around
Antarctica



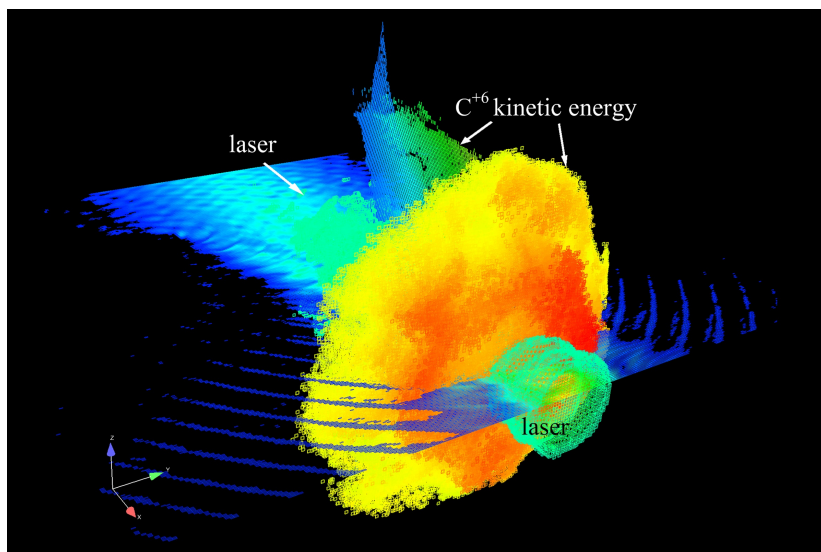
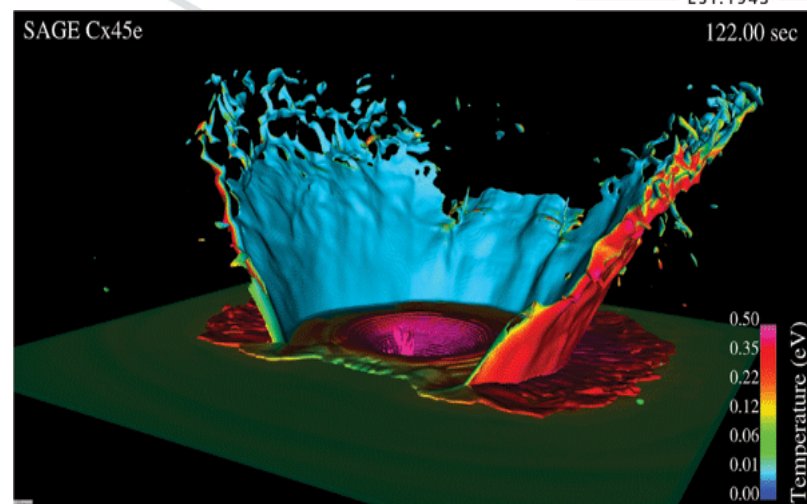
Fast Magnetic Reconnections
(sheds light on solar flares)



Turbulence
Around Oil
Rig Platforms

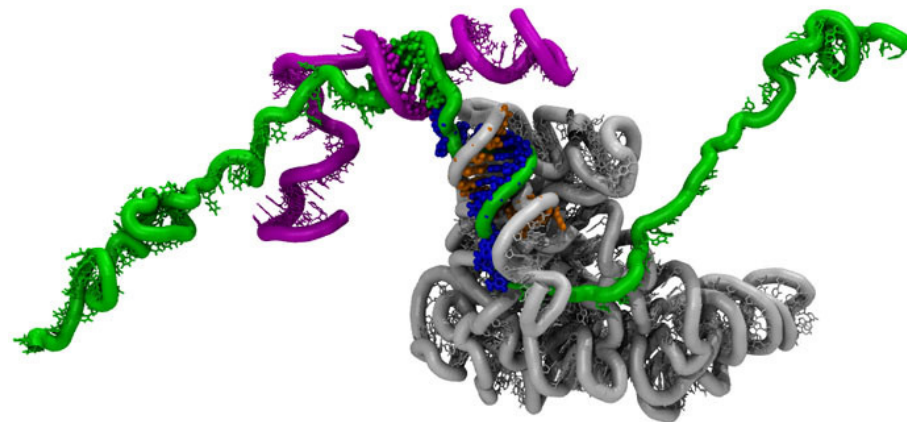
UNCLASSIFIED

Crater simulations



Ion Acceleration

UNCLASSIFIED



Molecular Dimmer Switches
(controls cellular metabolism)

Name	Compute Nodes	Peak TF/s	Linux OS	Built	Use
Cielo	8894	1370	CLE/CNL ¹	2010	ASC capability
Luna	1540	513	TOSS ²	2011	ASC capacity
Typhoon	416	106	TOSS	2010	ASC capacity
Viewmaster2	200	26	TOSS	2011	ASC Visualization
Cielito	68	10	CLE/CNL	2010	ASC regression test
Conejo	620	53	TOSS	2010	IC capacity
Hobo ³	272	38	TOSS	2014	IC data intensive
Mapache	592	50	TOSS	2010	ASC capacity
Mustang	1600	353	TOSS	2011	IC capacity
Moonlight	308	488	TOSS	2012	ASC advanced
Pinto	154	51	TOSS	2012	IC capacity
Wolf	616	205	TOSS	2014	IC capacity
Lightshow	28	3.6	TOSS	2011	ASC visualization
Kugel	96	6.8	TOSS	2013	Virtual machines
Glome	96	6.8	TOSS	2013	Hadoop
Bonanza	2	1.2	TOSS	2014	Research runs, large memory

UNCLASSIFIED



Cielo

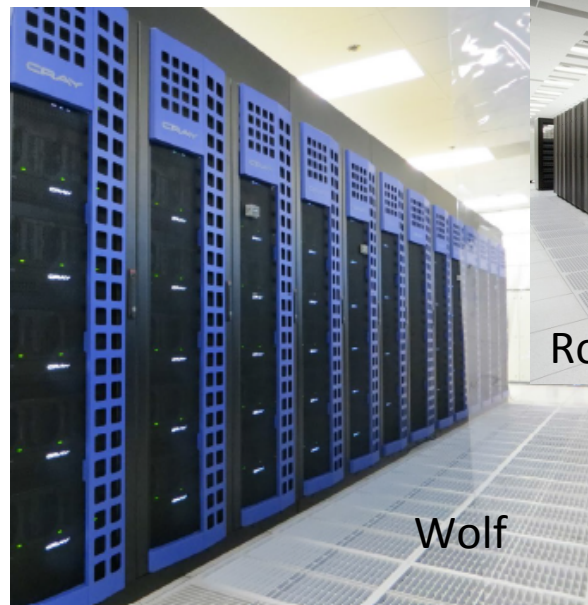


Luna



Roadrunner

- ✓ 1st to cross petaflop barrier
- ✓ 1st large heterogeneous cluster



Wolf



Mustang

UNCLASSIFIED

Trinity

- Cooling: Water
- Footprint: 5,200 sq ft
- Memory: 2 PB DDR4
- Burst Buffer: 3.7 PB
- Compute Nodes: 19,000



- 40 PetaFLOPs peak performance
- 80 PB parallel file storage
- 3.3 TB/s bandwidth to burst buffer
- 1.45 TB/s bandwidth to parallel storage

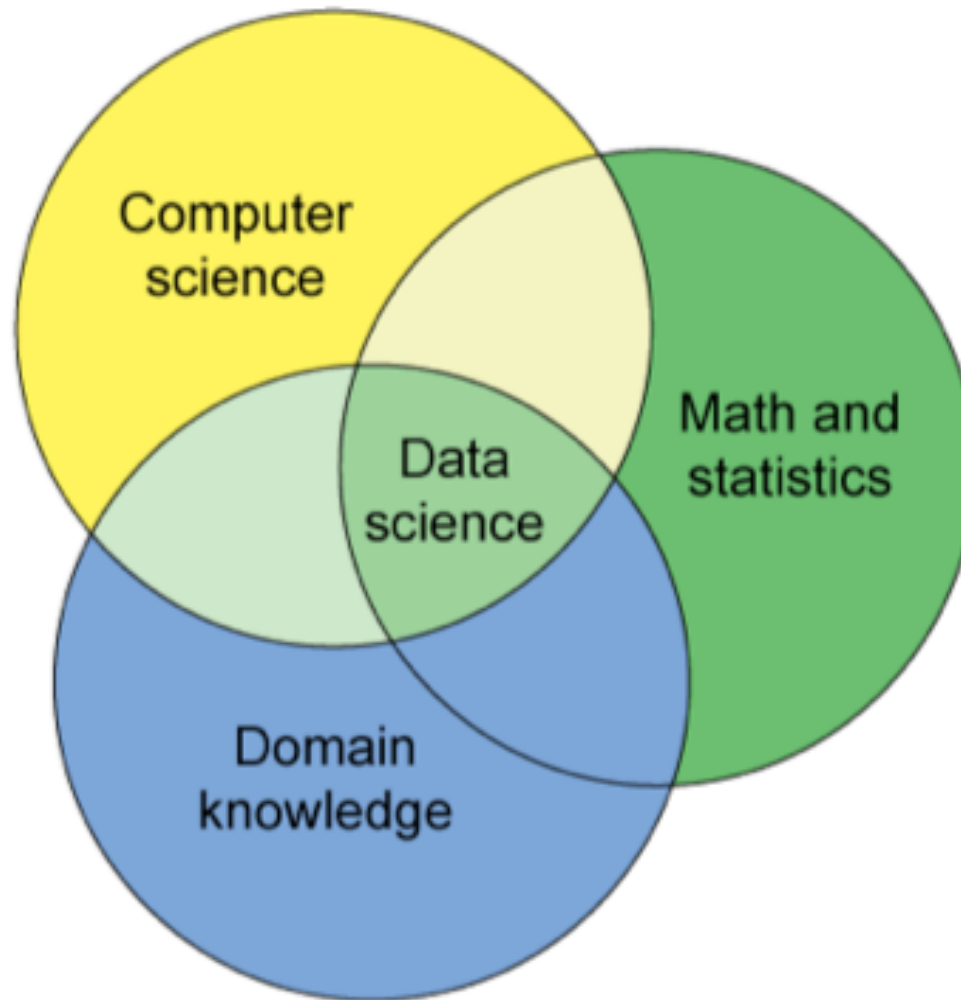


UNCLASSIFIED

Infrastructure

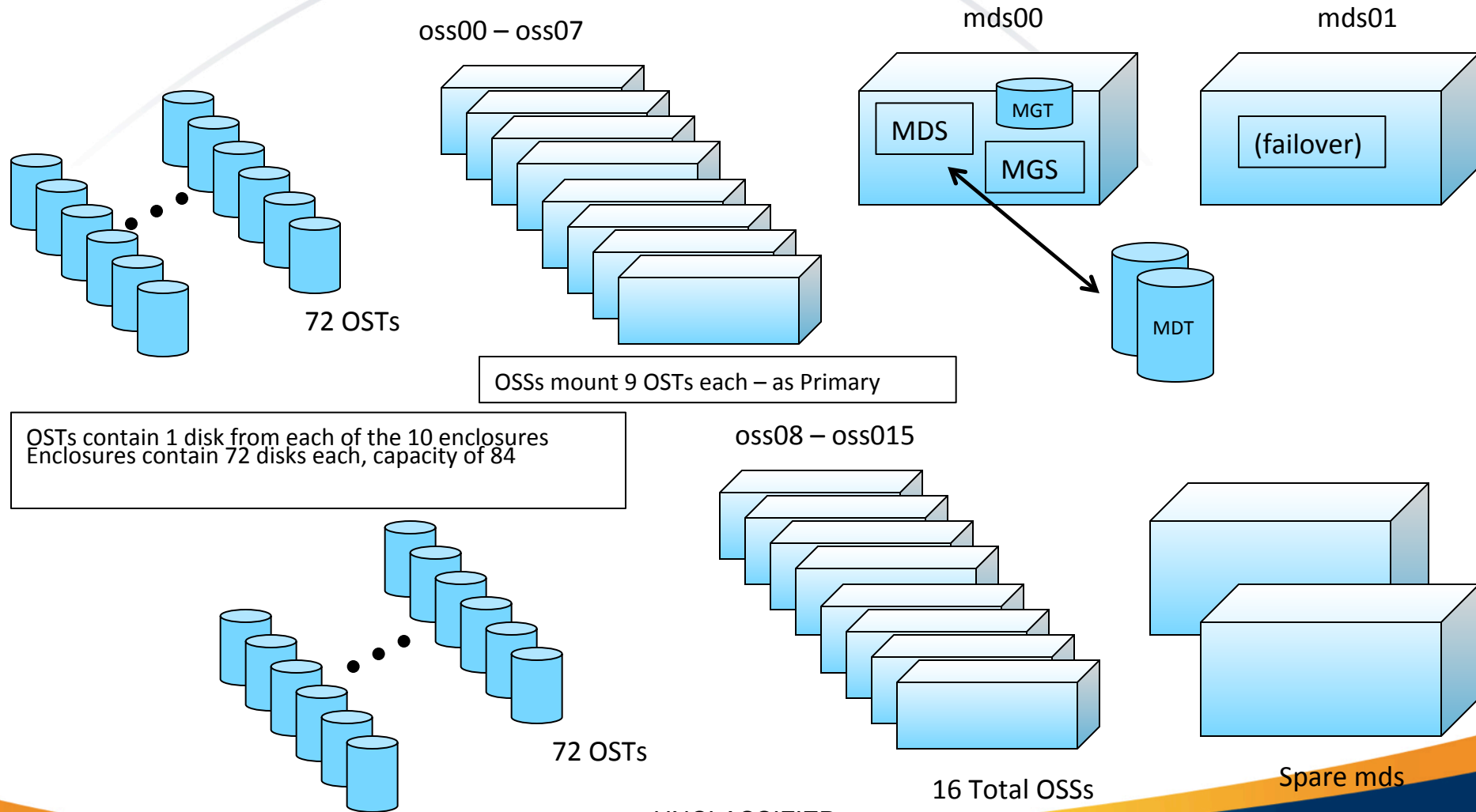


UNCLASSIFIED



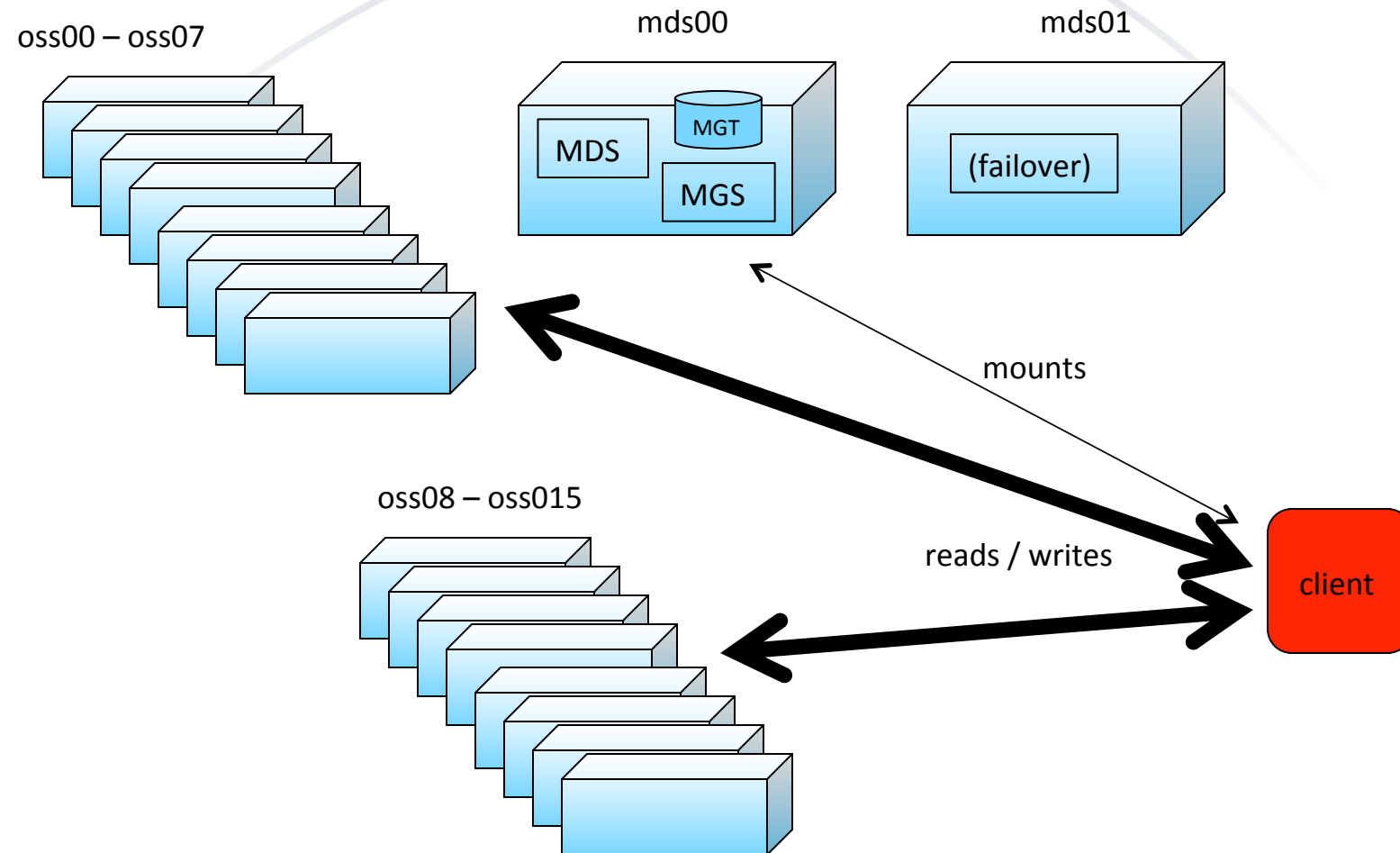
UNCLASSIFIED

Lustre Internals



UNCLASSIFIED

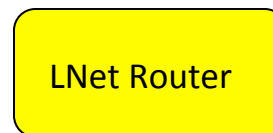
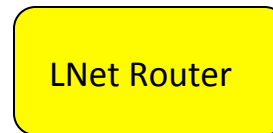
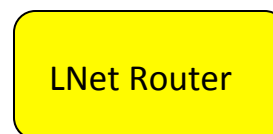
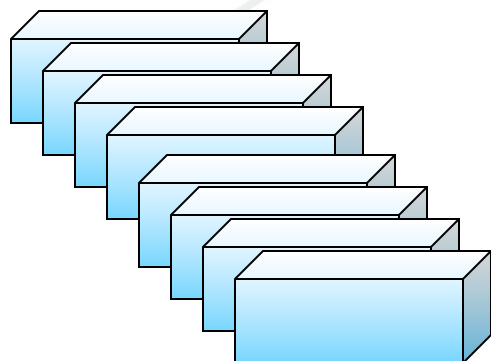
Lustre Communications



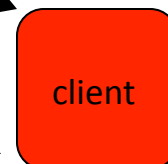
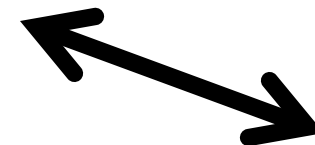
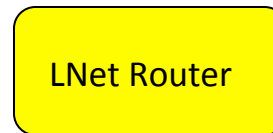
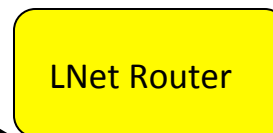
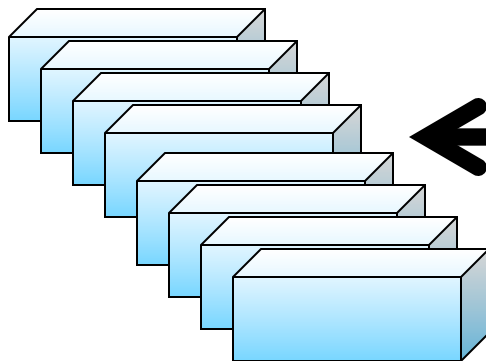
UNCLASSIFIED

LNet Routing

oss00 – oss07

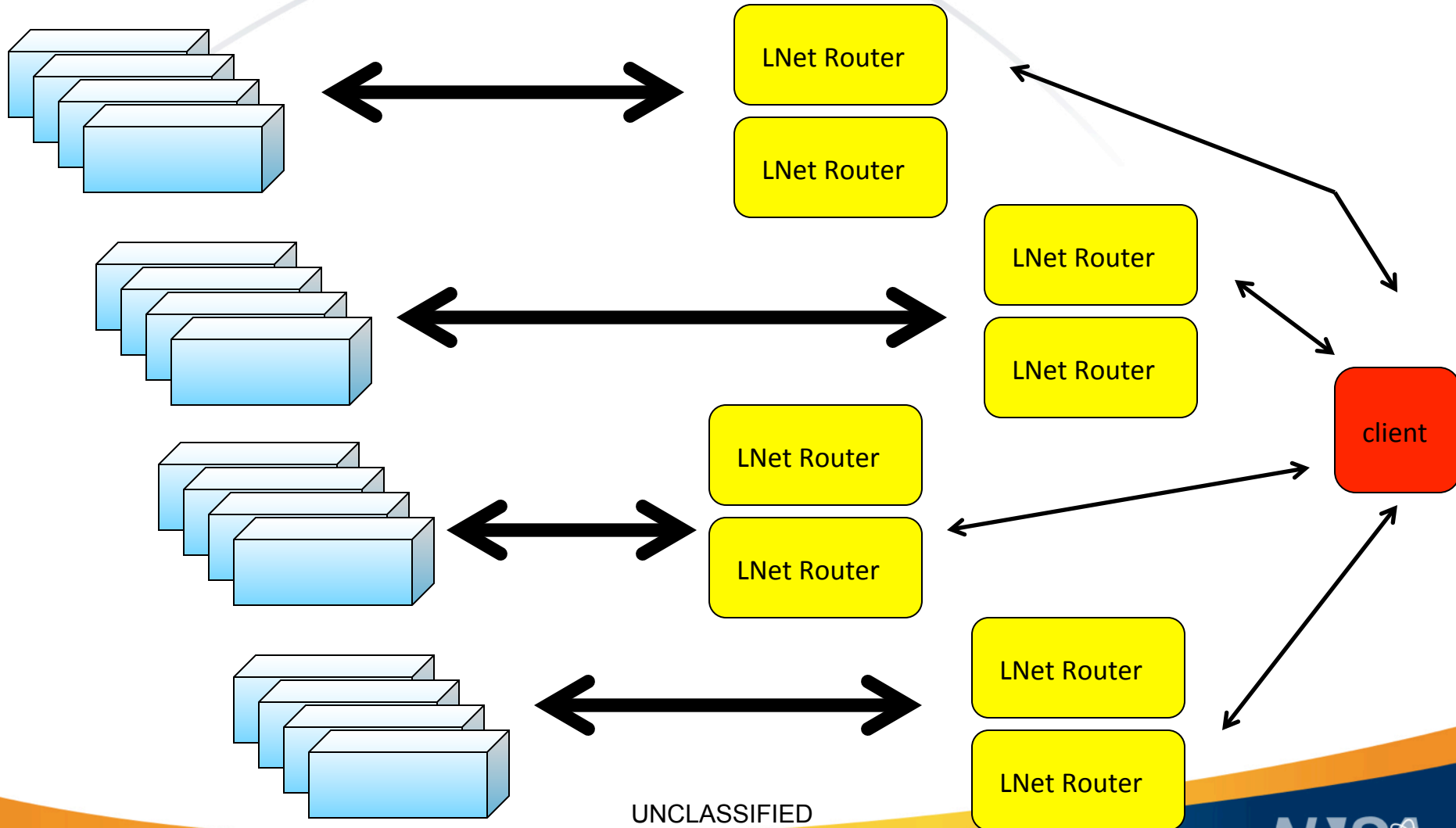


oss08 – oss015



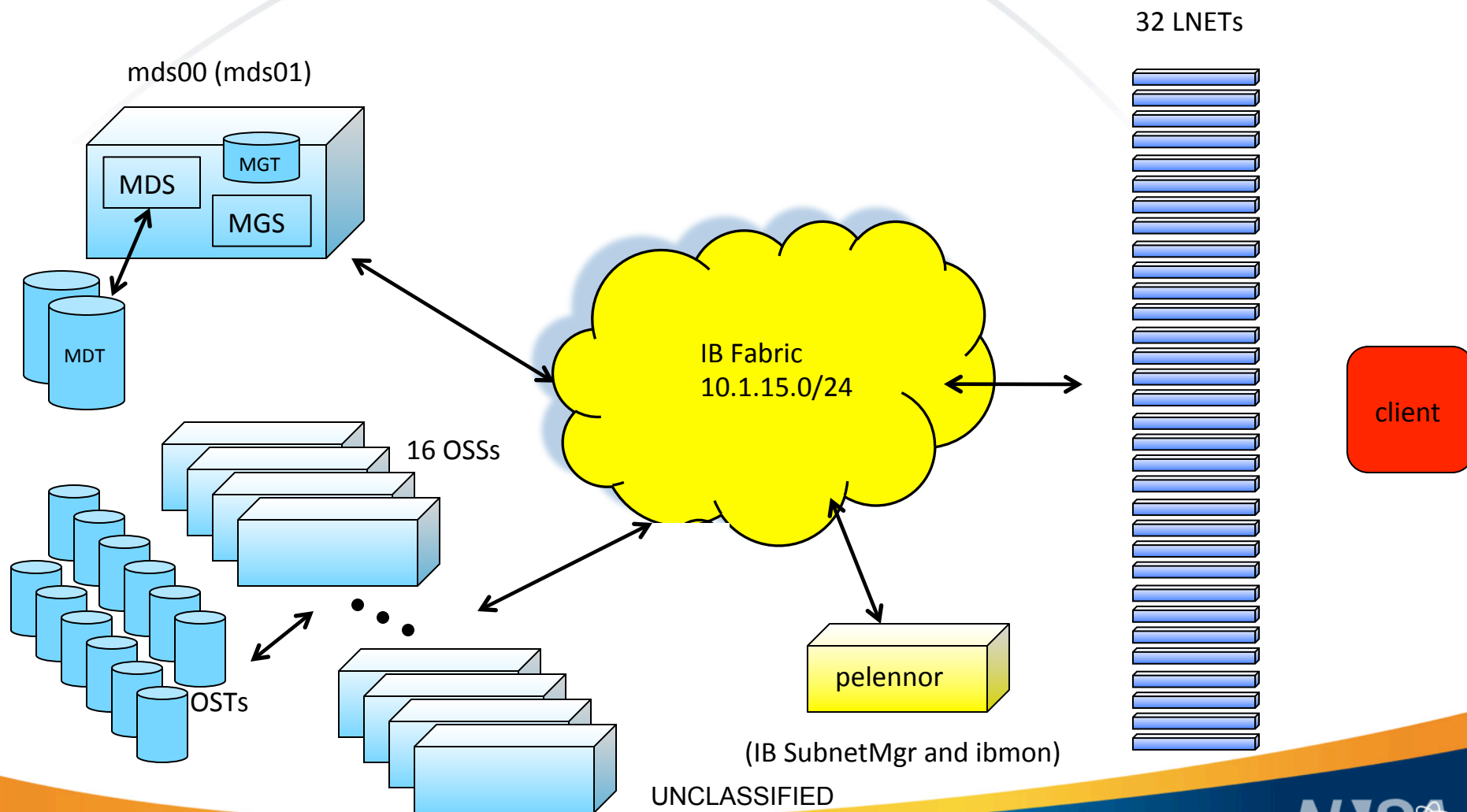
UNCLASSIFIED

Lustre FGR Concept

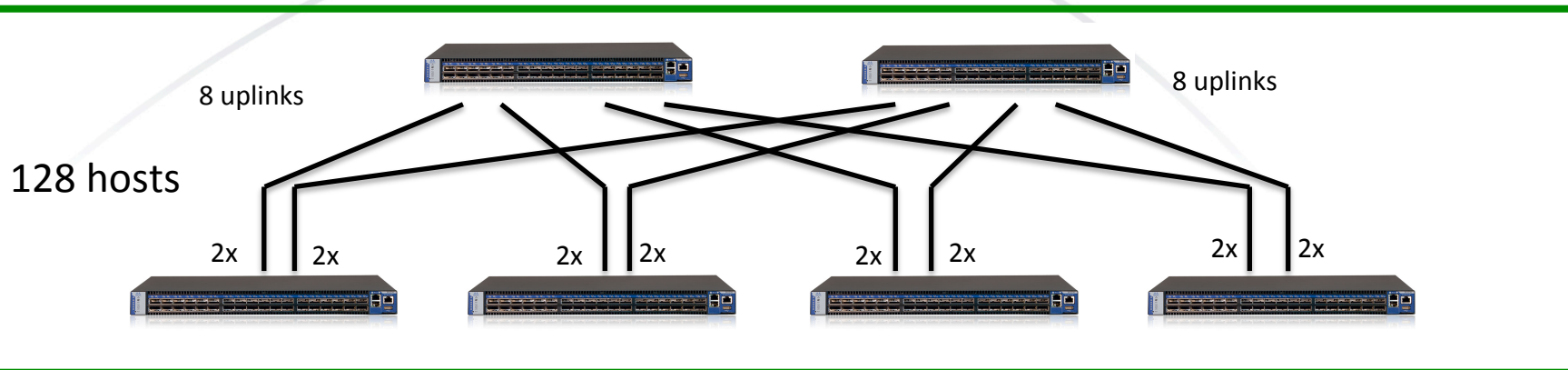


UNCLASSIFIED

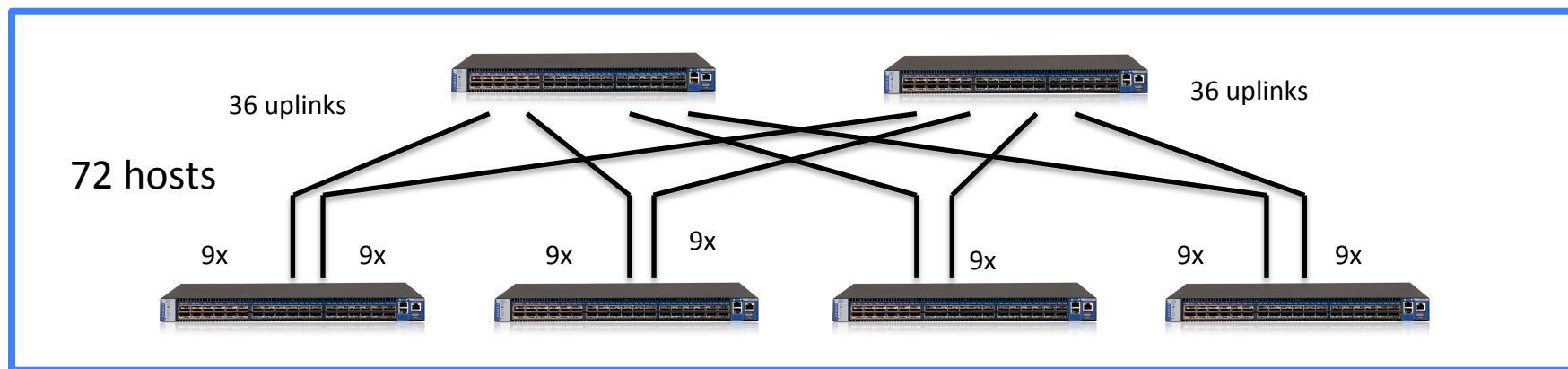
Lustre Big Picture



Lustre IB – Why this matters



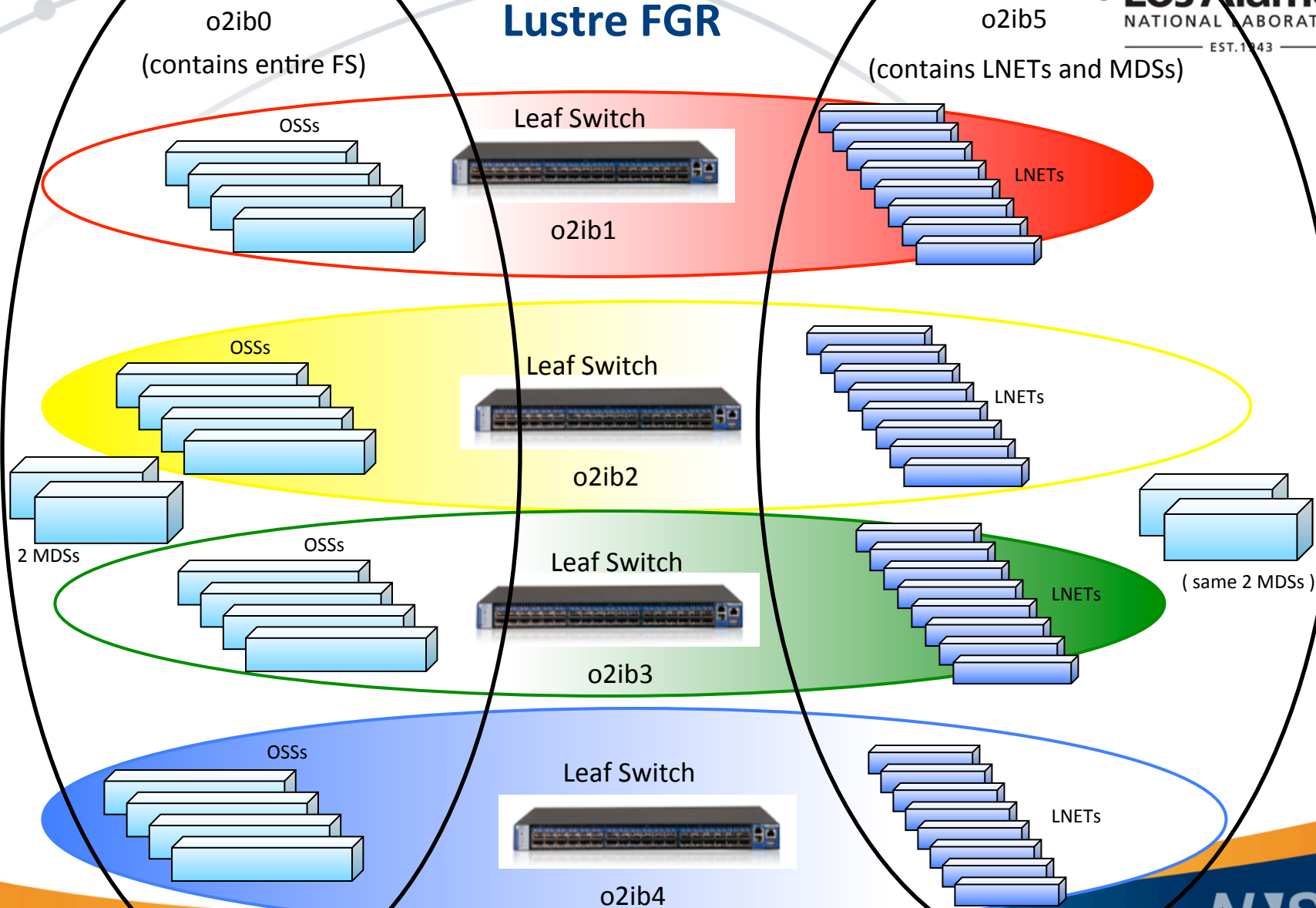
With FGR – No Fat Tree



Fat Tree

UNCLASSIFIED

Lustre FGR



Example Lustre FGR Details



Switch I1ib0			Switch I1ib1			Switch I1ib2			Switch I1ib3		
L1 o2ib1			L1 o2ib2			L1 o2ib3			L1 o2ib4		
Hostname	LustreIP	ClusterIP	Hostname	LustreIP	ClusterIP	Hostname	LustreIP	ClusterIP	Hostname	LustreIP	ClusterIP
I1-oss00	10.1.15.20		I1-oss01	10.1.15.21		I1-oss08	10.1.15.28		I1-oss09	10.1.15.29	
I1-oss02	10.1.15.22		I1-oss03	10.1.15.23		I1-oss10	10.1.15.30		I1-oss11	10.1.15.31	
I1-oss04	10.1.15.24		I1-oss05	10.1.15.25		I1-oss12	10.1.15.32		I1-oss13	10.1.15.33	
I1-oss06	10.1.15.26		I1-oss07	10.1.15.27		I1-oss14	10.1.15.34		I1-oss15	10.1.15.35	
L2 o2ib21			L2 o2ib22			L2 o2ib23			L2 o2ib24		
I2-oss00	10.1.15.36		I2-oss01	10.1.15.37		I1-oss04	10.1.15.40		I1-oss05	10.1.15.41	
I2-oss02	10.1.15.38		I2-oss03	10.1.15.39		I1-oss06	10.1.15.42		I1-oss07	10.1.15.43	
Ethernet LNet Routers in Leaf Switch FGR nets, plus o2ib5 and o2ib25 - and tcp0											
I1-Inet00	10.1.15.160	10.1.13.1	I1-Inet01	10.1.15.161	10.1.14.1	I1-Inet16	10.1.15.176	10.1.13.9	I1-Inet17	10.1.15.177	10.1.14.9
I1-Inet02	10.1.15.162	10.1.13.2	I1-Inet03	10.1.15.163	10.1.14.2	I1-Inet18	10.1.15.178	10.1.13.10	I1-Inet19	10.1.15.179	10.1.14.10
I1-Inet04	10.1.15.164	10.1.13.3	I1-Inet05	10.1.15.165	10.1.14.3	I1-Inet20	10.1.15.180	10.1.13.11	I1-Inet21	10.1.15.181	10.1.14.11
I1-Inet06	10.1.15.166	10.1.13.4	I1-Inet07	10.1.15.167	10.1.14.4	I1-Inet22	10.1.15.182	10.1.13.12	I1-Inet23	10.1.15.183	10.1.14.12
I1-Inet08	10.1.15.168	10.1.13.5	I1-Inet09	10.1.15.169	10.1.14.5	I1-Inet24	10.1.15.184	10.1.13.13	I1-Inet25	10.1.15.185	10.1.14.13
I1-Inet10	10.1.15.170	10.1.13.6	I1-Inet11	10.1.15.171	10.1.14.6	I1-Inet26	10.1.15.186	10.1.13.14	I1-Inet27	10.1.15.187	10.1.14.14
I1-Inet12	10.1.15.172	10.1.13.7	I1-Inet13	10.1.15.173	10.1.14.7	I1-Inet28	10.1.15.188	10.1.13.15	I1-Inet29	10.1.15.189	10.1.14.15
I1-Inet14	10.1.15.174	10.1.13.8	I1-Inet15	10.1.15.175	10.1.14.8	I1-Inet30	10.1.15.190	10.1.13.16	I1-Inet31	10.1.15.191	10.1.14.16
Wolf LNet Routers in Leaf Switch FGR nets, plus o2ib5 and o2ib25 - and o2ib10											
wf-io01	10.1.15.101	10.16.11.1	wf-io02	10.1.15.102	10.16.11.2	wf-io03	10.1.15.103	10.16.11.3	wf-io04	10.1.15.104	10.16.11.4
wf-io05	10.1.15.105	10.16.11.5	wf-io06	10.1.15.106	10.16.11.6	wf-io07	10.1.15.107	10.16.11.7	wf-io08	10.1.15.108	10.16.11.8
wf-io09	10.1.15.109	10.16.11.9	wf-io10	10.1.15.110	10.16.11.10	wf-io11	10.1.15.111	10.16.11.11	wf-io12	10.1.15.112	10.16.11.12
wf-io13	10.1.15.113	10.16.11.13	wf-io14	10.1.15.114	10.16.11.14	wf-io15	10.1.15.115	10.16.11.15	wf-io16	10.1.15.116	10.16.11.16
wf-io17	10.1.15.117	10.16.11.17	wf-io18	10.1.15.118	10.16.11.18	wf-io19	10.1.15.119	10.16.11.19	wf-io20	10.1.15.120	10.16.11.20
wf-io21	10.1.15.121	10.16.11.21	wf-io22	10.1.15.122	10.16.11.22	wf-io23	10.1.15.123	10.16.11.23	wf-io24	10.1.15.124	10.16.11.24

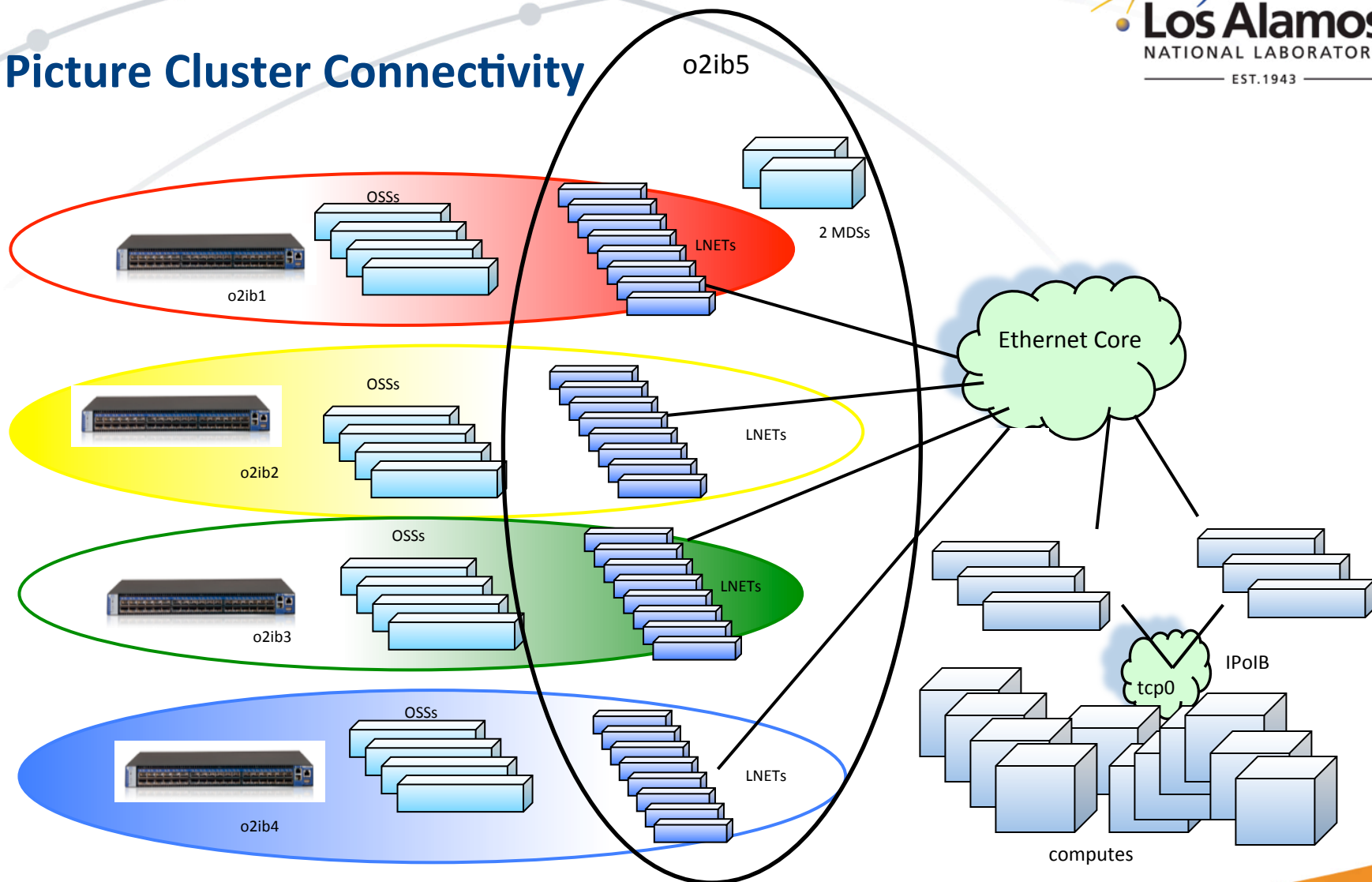
Example LNet Configuration File

```
### LNET Networks
###
#### Ethernet
### LNETs      10.1.13. [1-16], 10.1.14 [1-16]
###
#### IB
### LNET-1     10.1.15. [160,162,164,166,168,170,172,174]
### LNET-2     10.1.15. [161,163,165,167,169,171,173,175]
### LNET-3     10.1.15. [176,178,180,182,184,186,188,190]
### LNET-4     10.1.15. [177,179,181,183,185,187,189,191]
###
### L1
### MDSs       10.1.15.1&2
### OSS-1      10.1.15. [20,22,24,26]
### OSS-2      10.1.15. [21,23,25,27]
### OSS-3      10.1.15. [28,30,32,34]
### OSS-4      10.1.15. [29,31,33,35]
###
### L2
### MDSs       10.1.15.3&4
### OSS-1      10.1.15. [36,38]
### OSS-2      10.1.15. [37,39]
### OSS-3      10.1.15. [40,42]
### OSS-4      10.1.15. [41,43]
```

```
options lnet ip2nets="\
    tcp0(bond0)      10.1. [13-14]. [1-16];\
    o2ib1(ib2)       10.1.15. [20,22,24,26];\
    o2ib1(ib0)       10.1.15. [160,162,164,166,168,170,172,174];\
    o2ib2(ib2)       10.1.15. [21,23,25,27];\
    o2ib2(ib0)       10.1.15. [161,163,165,167,169,171,173,175];\
    o2ib3(ib2)       10.1.15. [28,30,32,34];\
    o2ib3(ib0)       10.1.15. [176,178,180,182,184,186,188,190];\
    o2ib4(ib2)       10.1.15. [29,31,33,35];\
    o2ib4(ib0)       10.1.15. [177,179,181,183,185,187,189,191];\
    o2ib5(ib0)       10.1.15. [1,2,160-191];\
    o2ib21(ib2)      10.1.15. [36,38];\
    o2ib21(ib0)      10.1.15. [160,162,164,166,168,170,172,174];\
    o2ib22(ib2)      10.1.15. [37,39];\
    o2ib22(ib0)      10.1.15. [161,163,165,167,169,171,173,175];\
    o2ib23(ib2)      10.1.15. [40,42];\
    o2ib23(ib0)      10.1.15. [176,178,180,182,184,186,188,190];\
    o2ib24(ib2)      10.1.15. [41,43];\
    o2ib24(ib0)      10.1.15. [177,179,181,183,185,187,189,191];\
    o2ib25(ib0)      10.1.15. [3,4,160-191]"
```

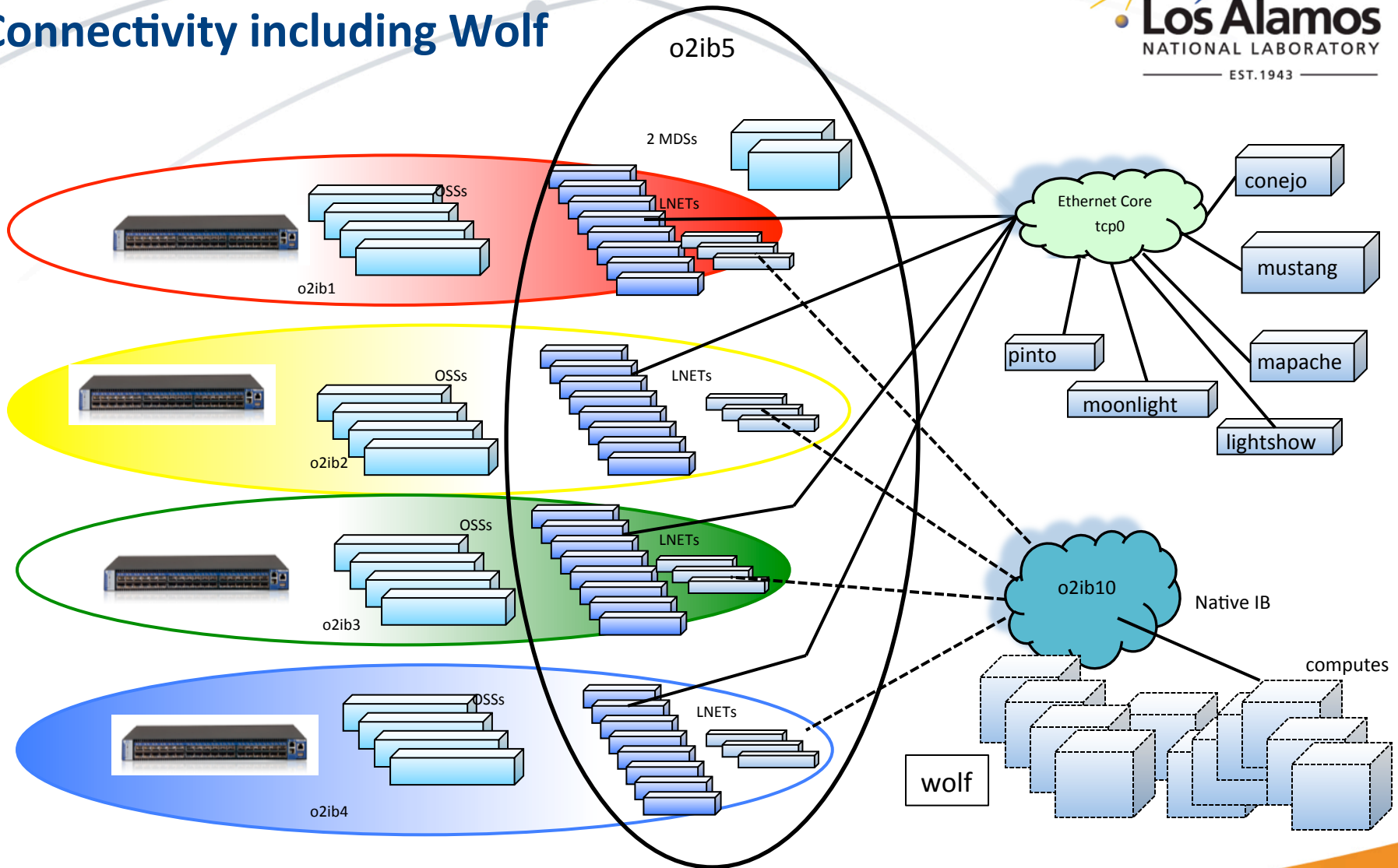
```
### Make it a router
options lnet forwarding=enabled
```


Big Picture Cluster Connectivity



UNCLASSIFIED

Connectivity including Wolf



UNCLASSIFIED



Thank You

Questions ?

<http://www.lanl.gov/careers/career-options/student-internships/index.php>

UNCLASSIFIED