LA-UR-20-24071
Approved for public release; distribution is unlimited.

Title: Clusters @ LANL

Author(s): Lueninghoener, Cory Donald

Intended for: HPC summer student presentation

Issued: 2020-06-03

Clusters @ LANL

Cory Lueninghoener
HPC-DES

whoami

- Systems guy in HPC-DES
  - Large-scale HPC system design
- At LANL since 2010
- Previously at Argonne National Laboratory
  - Started as a student!
- Also spent some time at Facebook
- Ugrad/Grad school at University of Nebraska

HPC Division Clusters:

About 20 Total

Some are…

Production clusters

Some are…

Testbed clusters

Some are…

Big clusters

Some are…

Small clusters

Some are…

Easy to use

Some are…

Difficult to use

Some are…

Important clusters

Some are…

Not-as-important

Some are…

Busy clusters

Some are…

Idle clusters

They all support science

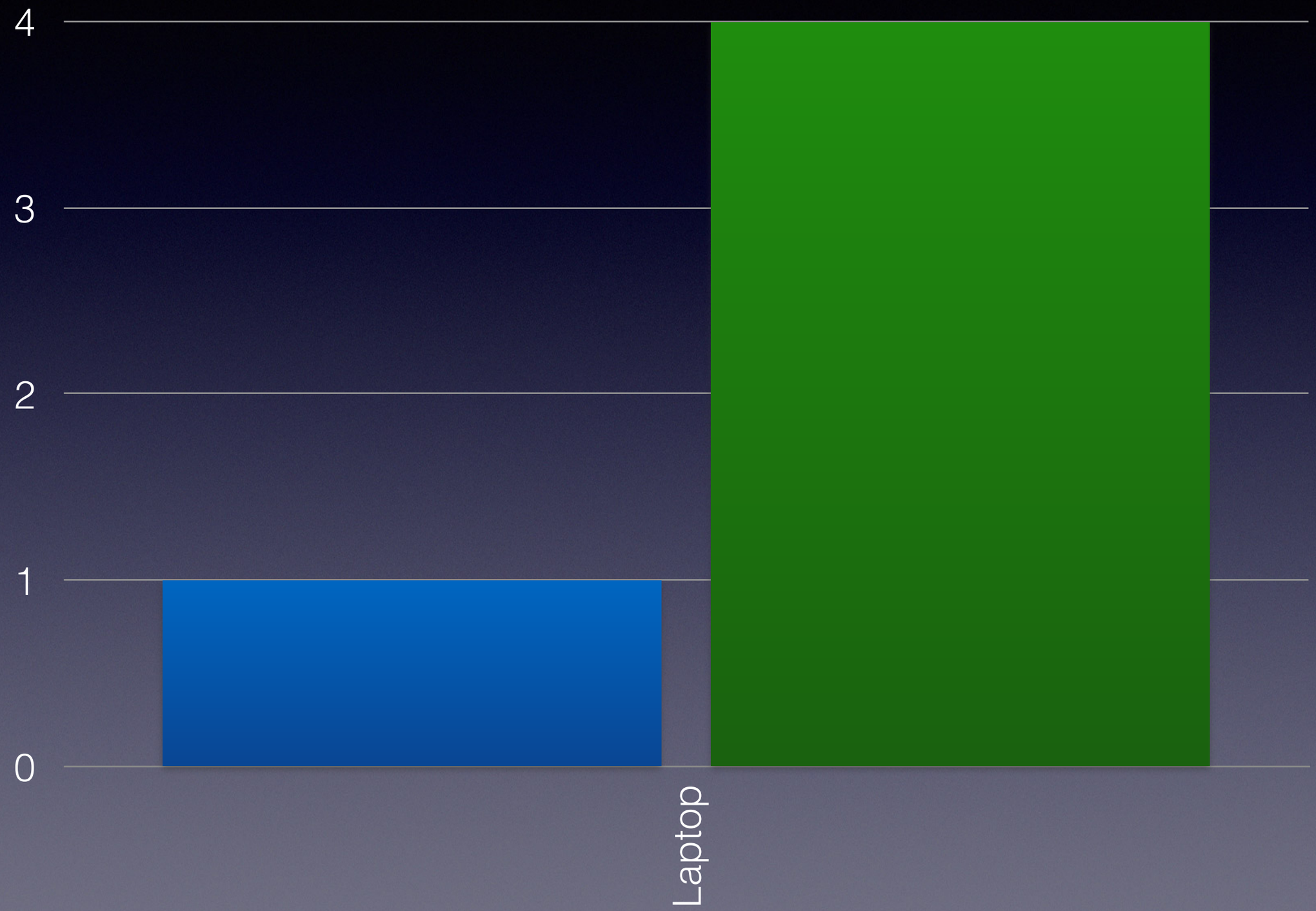
A Talk in Two Parts

- Part One: Scale
- Part Two: Scale

Part One: A Tour of Systems

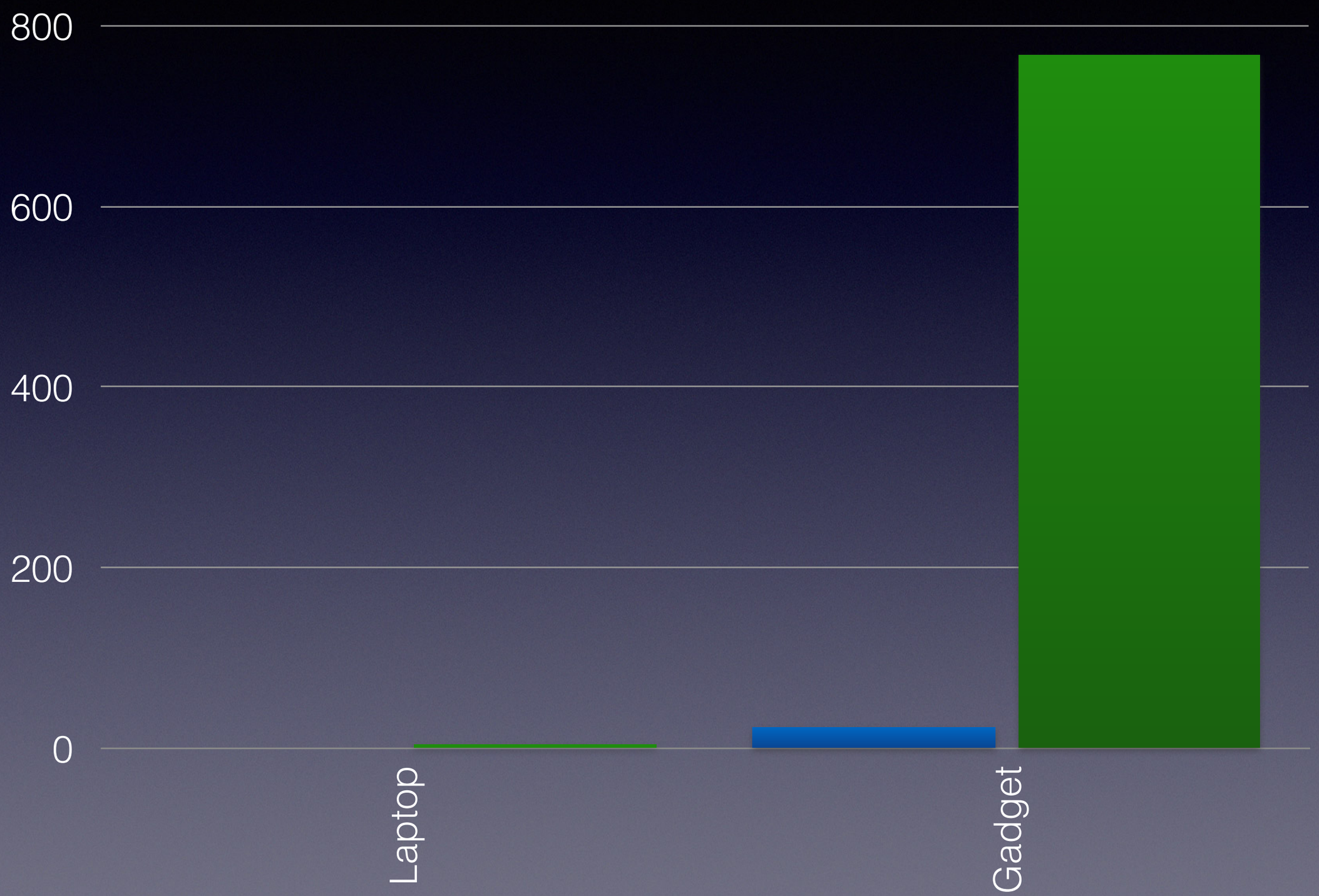
My Laptop

- One node
- Four processing cores
- Wired and wireless networking
- Is covered in stickers
- TOTALLY AWESEOME!!!!11!!


4
3
2
1
0
Laptop

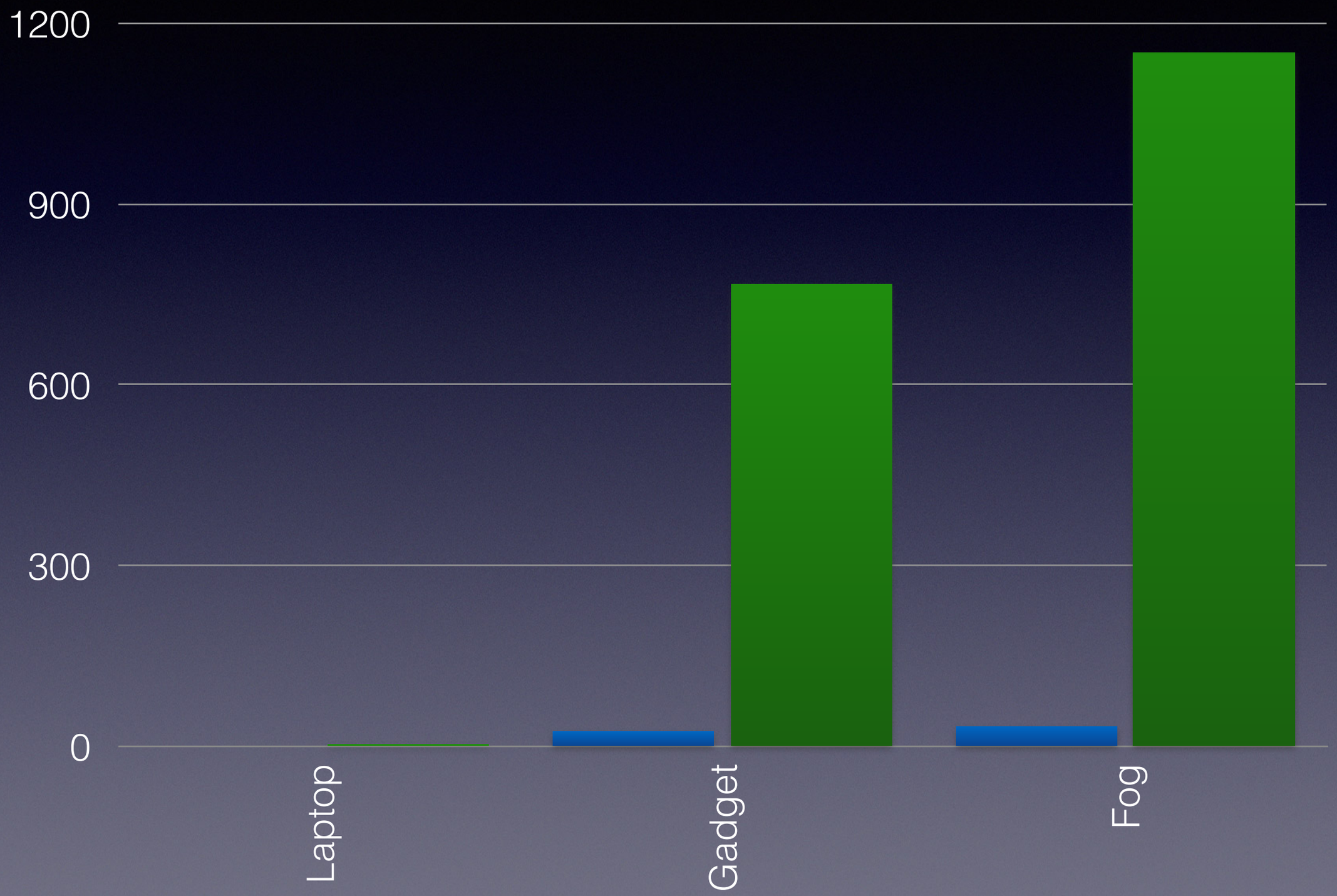
Gadget

- 24 compute nodes
- 768 cores
- Yellow network
- ATS-1 systems testbed


800
600
400
200
0
Laptop
Gadget

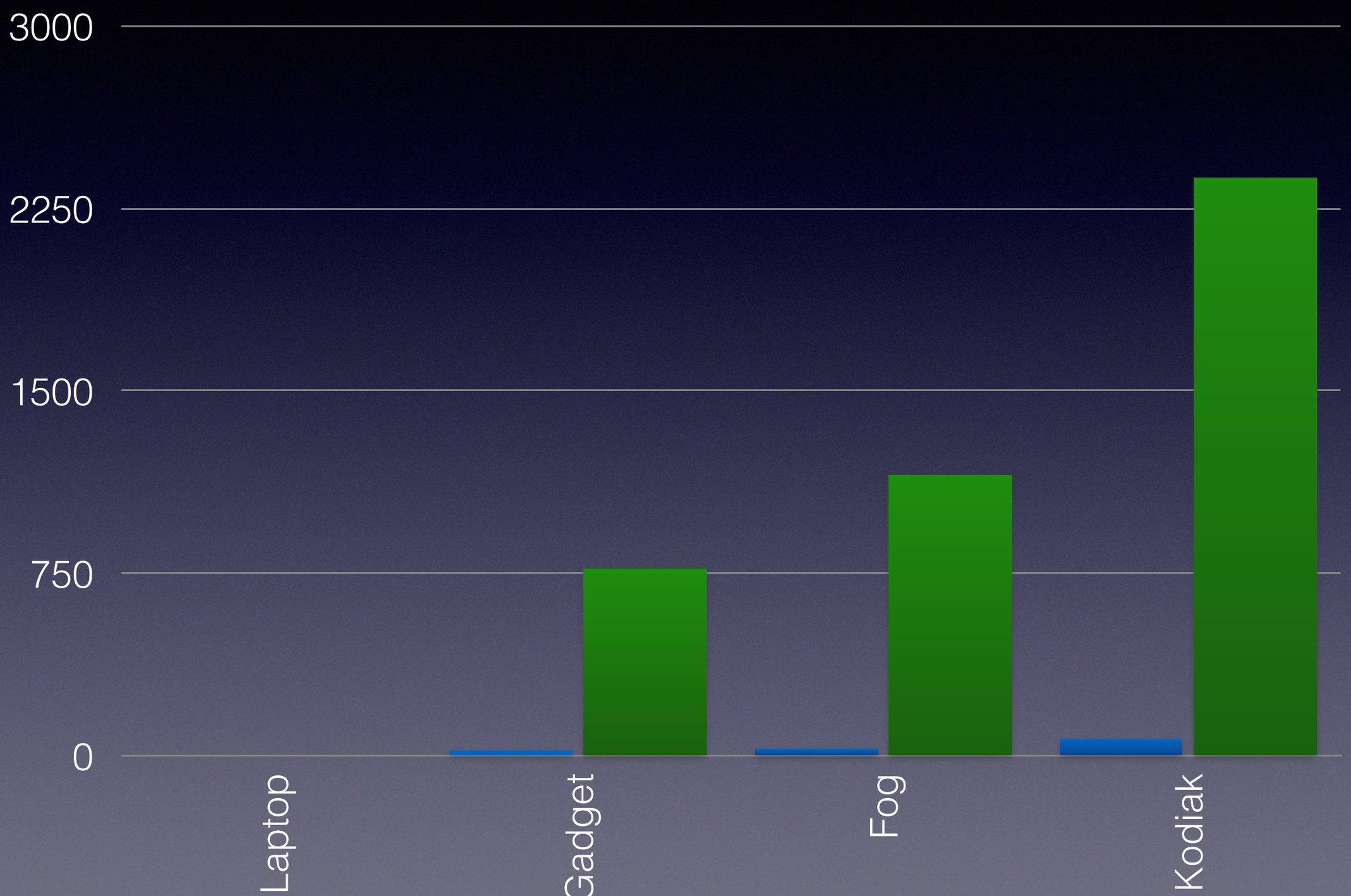
Fog

- 32 compute nodes
- 1152 cores
- Yellow network
- CTS-1 application testbed


1200
900
600
300
0
Laptop
Gadget
Fog

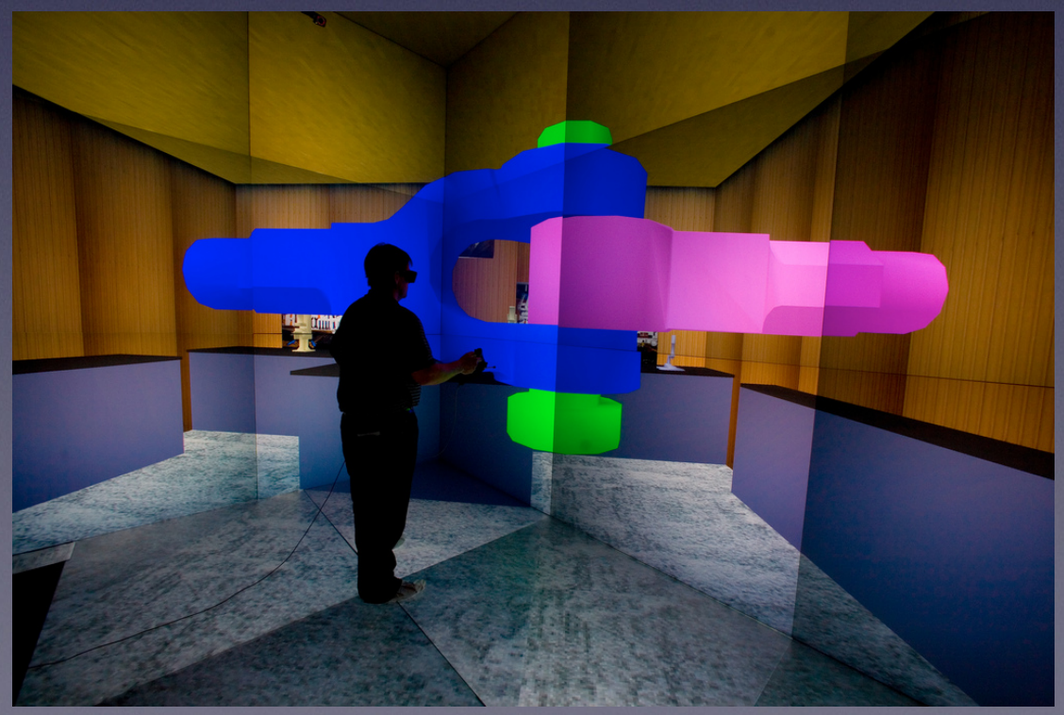
Kodiak

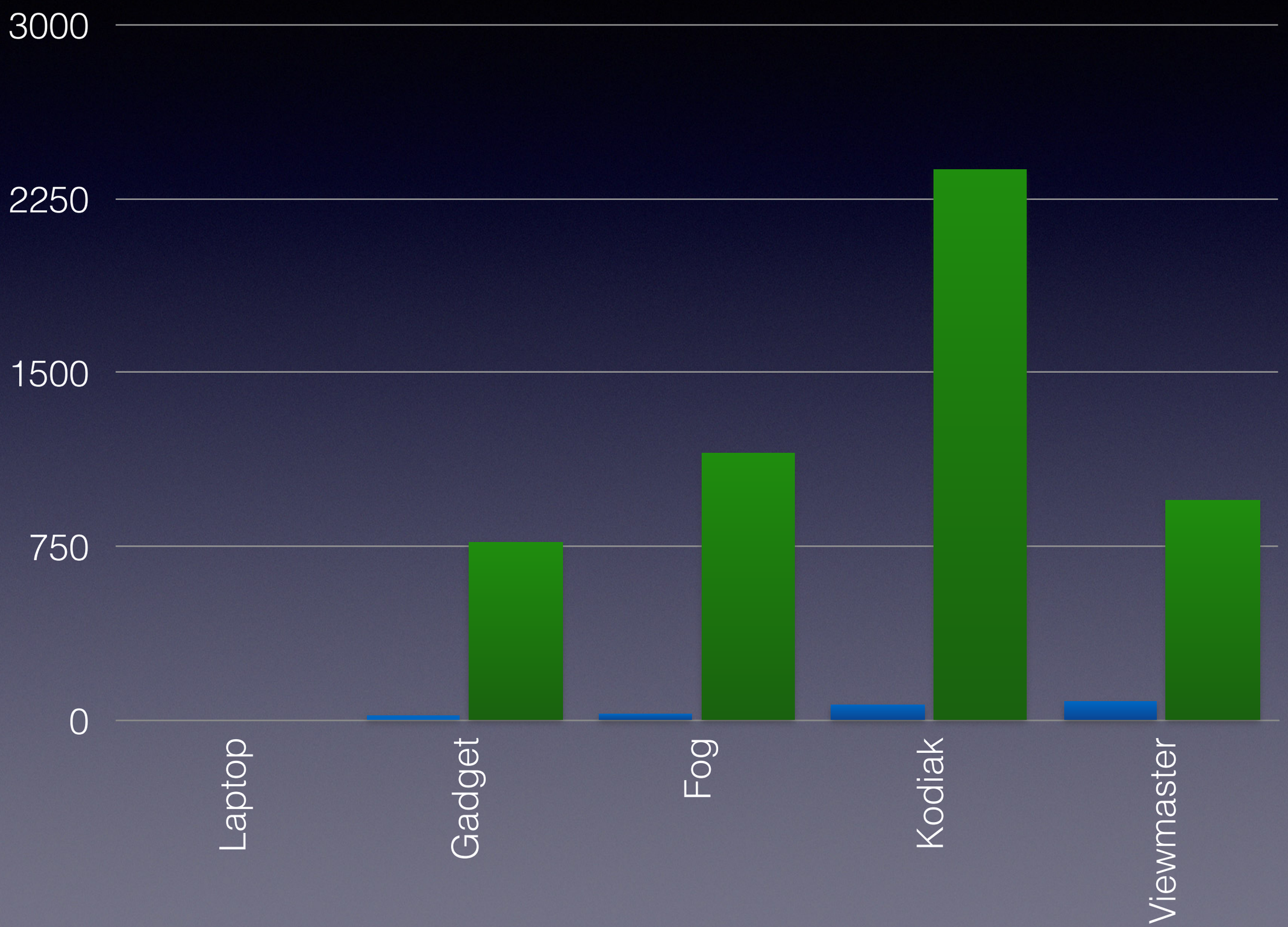
- 66 compute nodes
- 2376 cores
- Turquoise network
- CTS-1 system with GPUs


3000
2250
1500
750
0
Laptop
Gadget
Fog
Kodiak

Viewmaster3

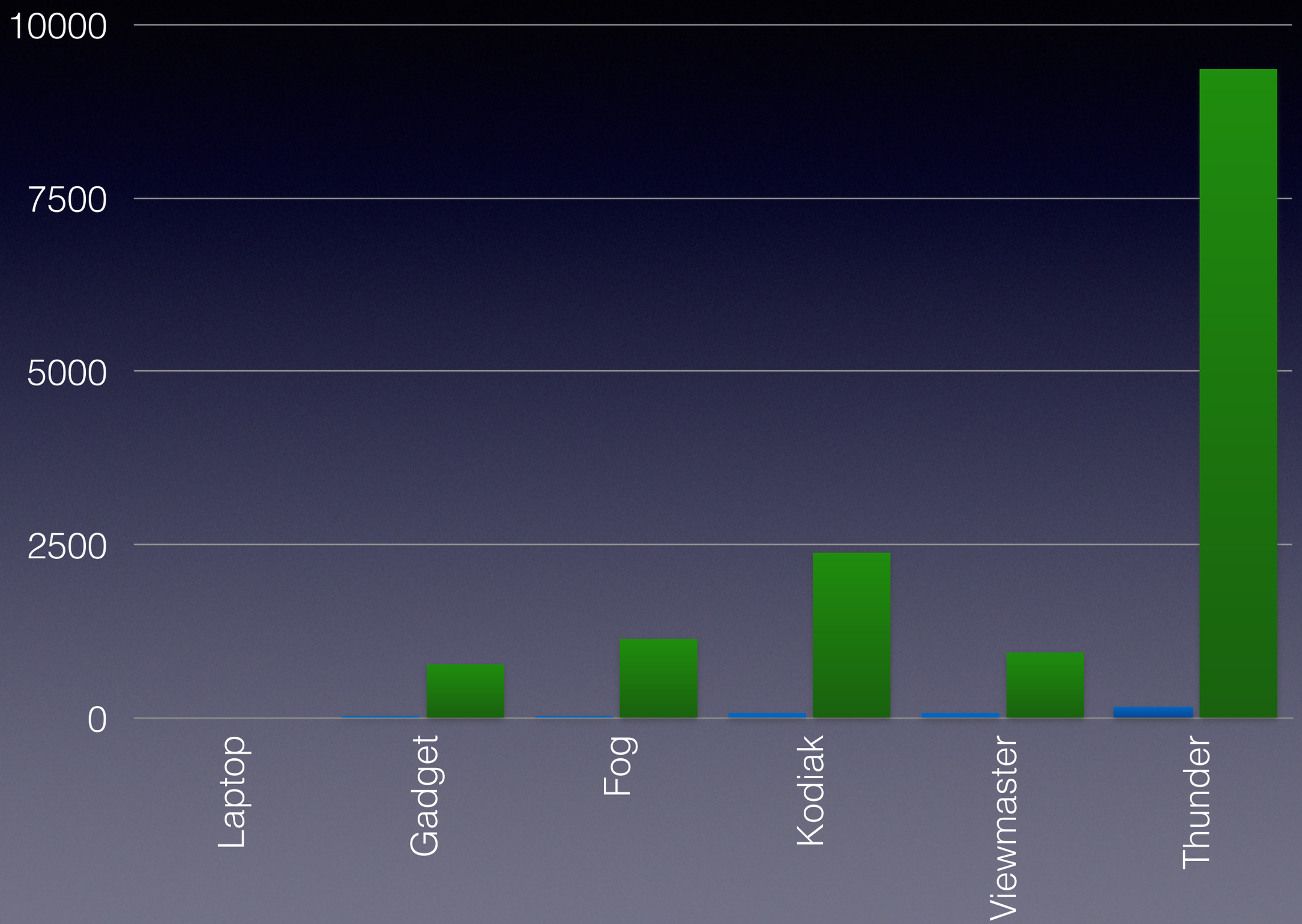
- 79 "compute" nodes
- 948 cores
- nVidia Quadro P6000 Graphics Cards
- Red network
- Powers the Powerwall and CAVE




3000
2250
1500
750
0
Laptop
Gadget
Fog
Kodiak
Viewmaster

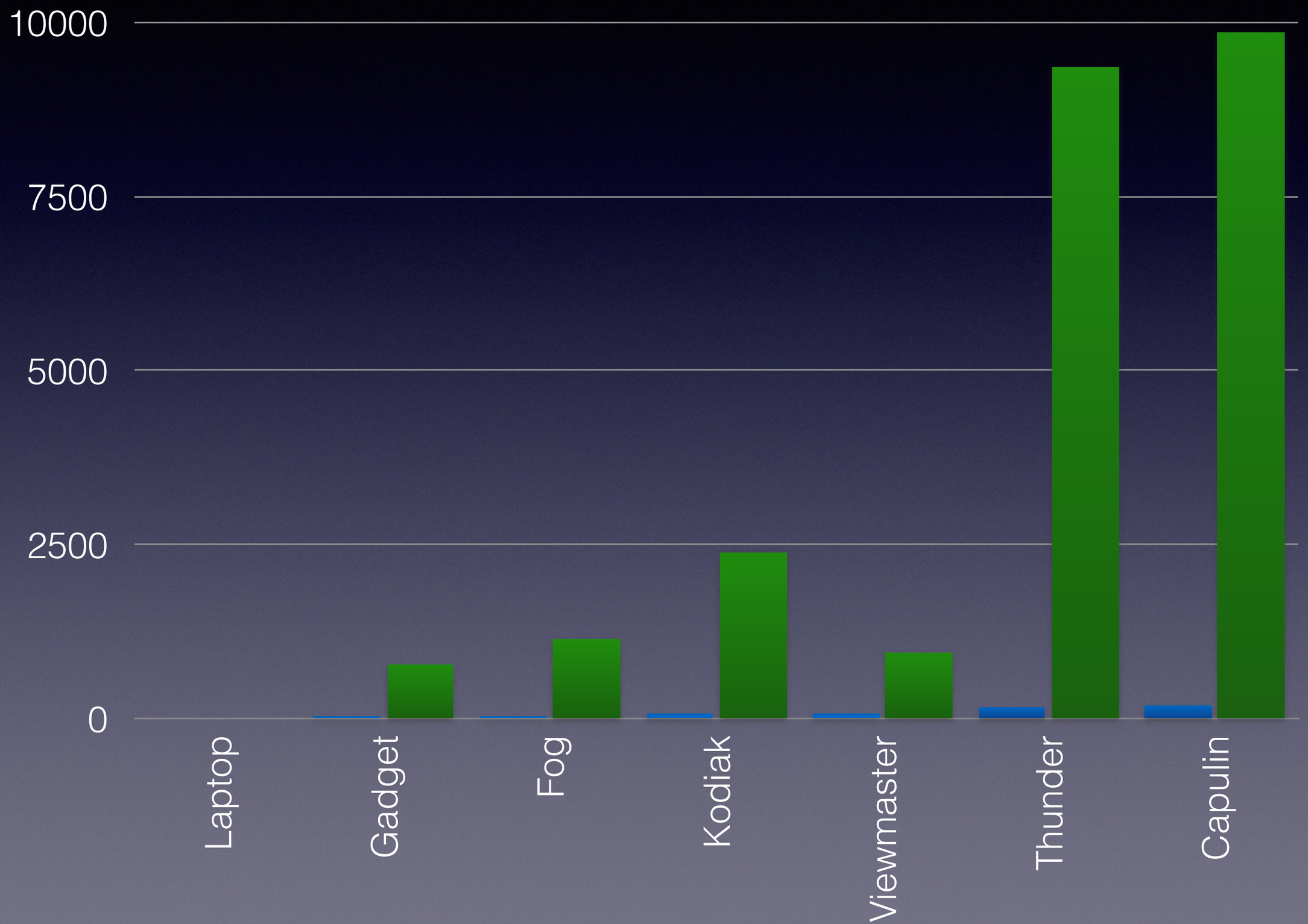
Thunder

- 167 compute nodes
- 9352 cores
- Red network
- Cray ARM test system


10000
7500
5000
2500
0
Laptop
Gadget
Fog
Kodiak
Viewmaster
Thunder

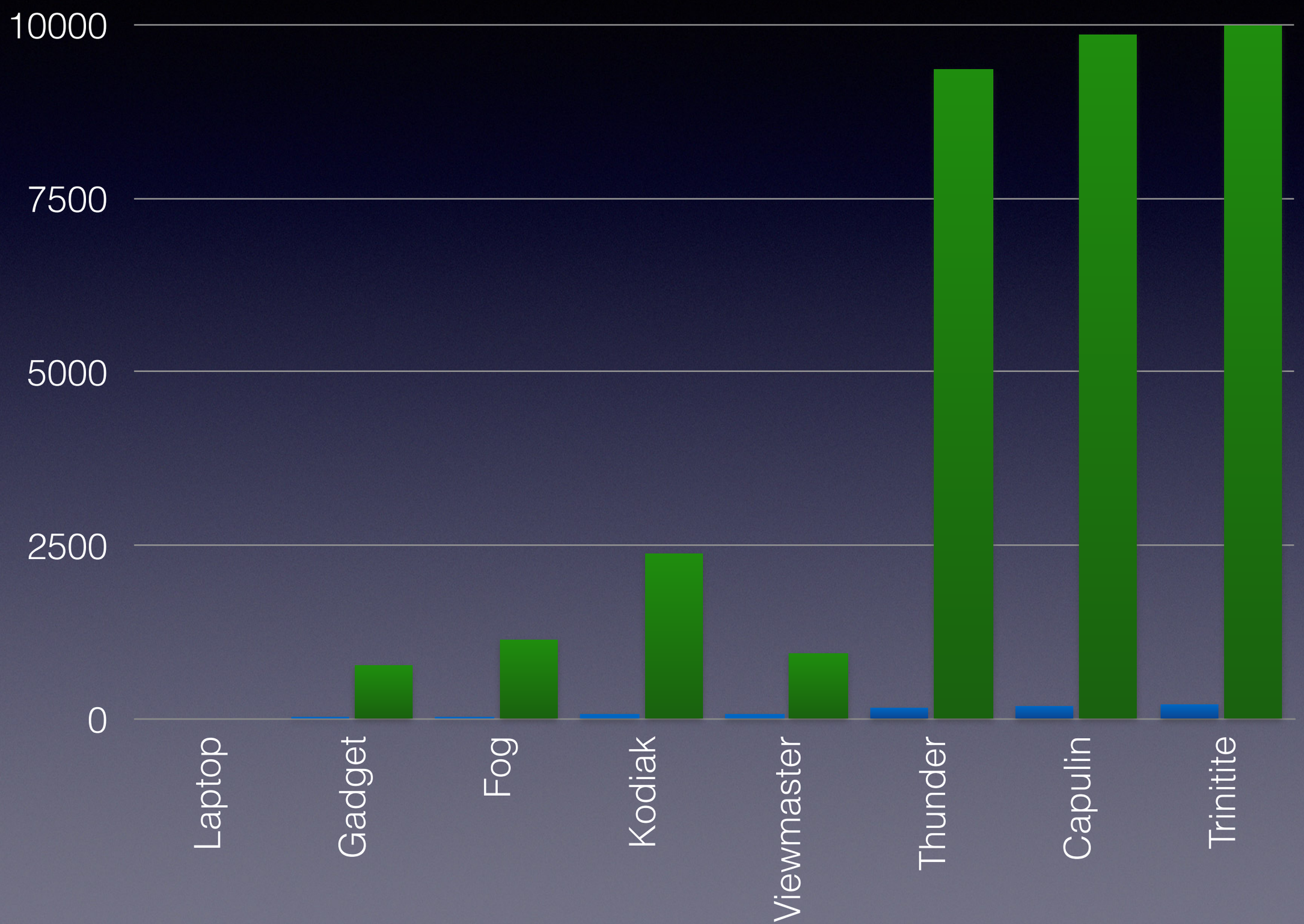
Capulin

- 175 compute nodes
- 9856 cores
- Yellow network
- Cray ARM test system


10000
7500
5000
2500
0
Laptop
Gadget
Fog
Kodiak
Viewmaster
Thunder
Capulin

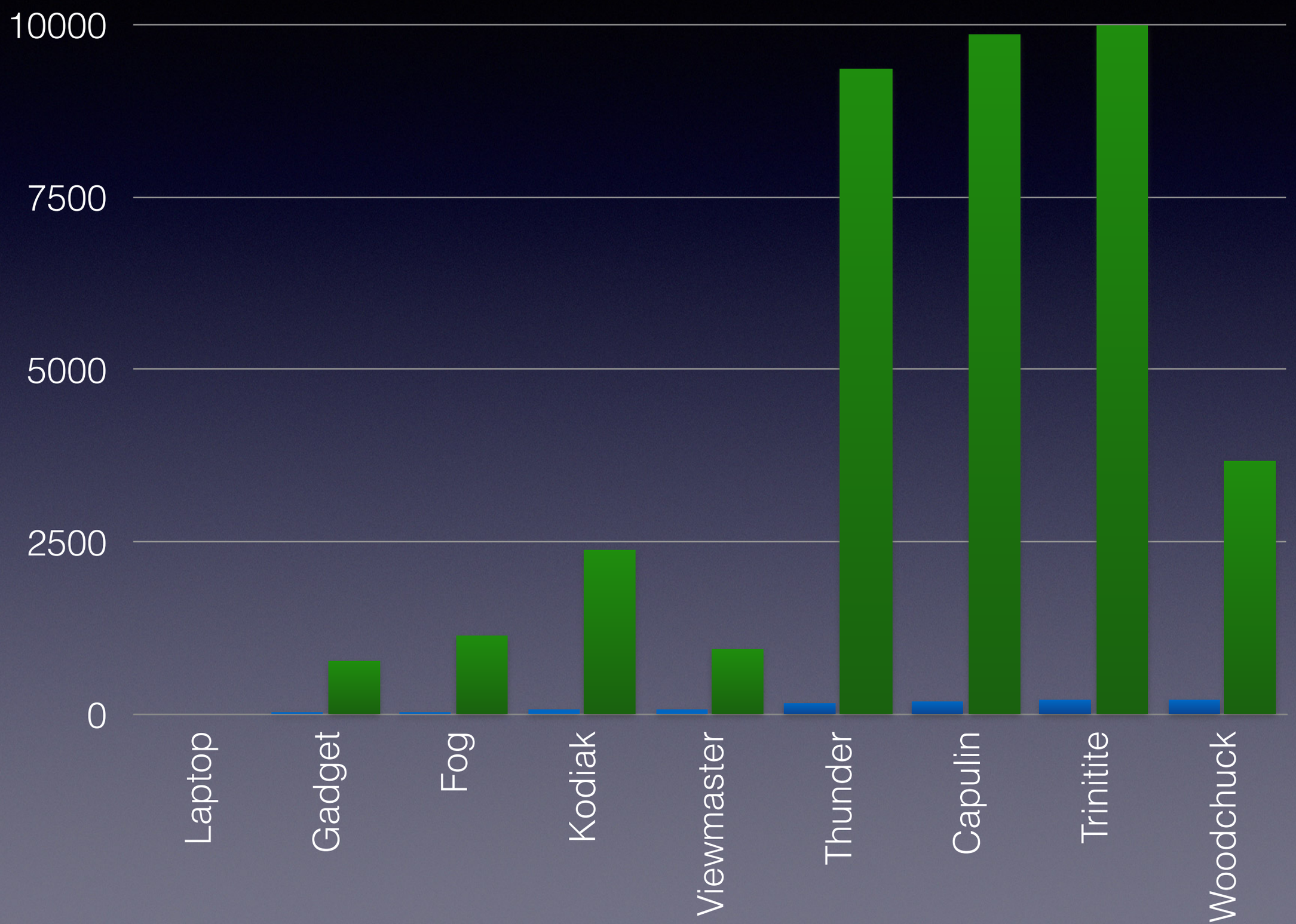
Trinitite

- 200 compute nodes
- 10000 cores
- Yellow network
- ATS-1 application testbed


10000
7500
5000
2500
0
Laptop
Gadget
Fog
Kodiak
Viewmaster
Thunder
Capulin
Trinitite

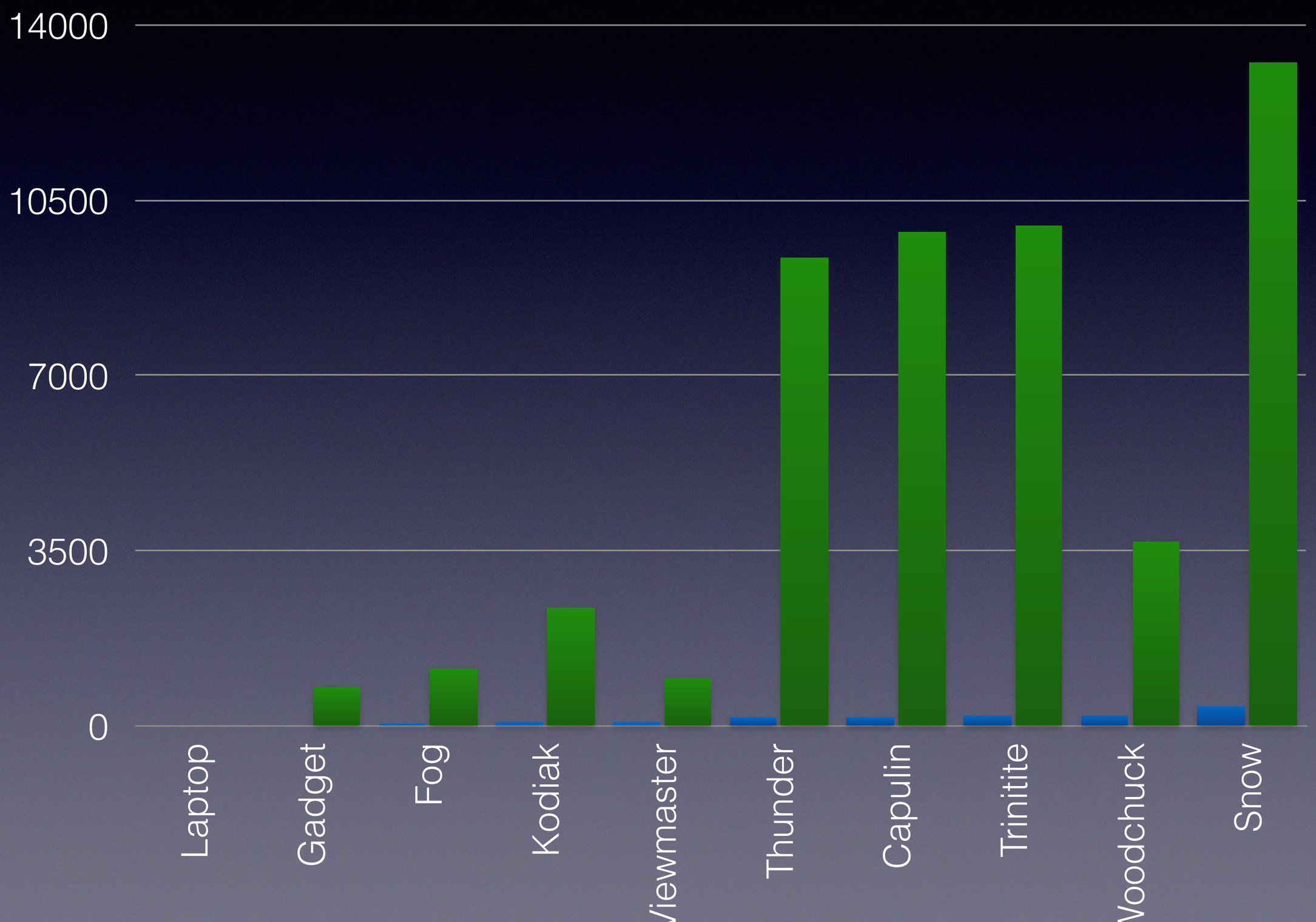
Woodchuck

- 212 compute nodes
- 3672 cores
- Turquoise network
- Flexible stacks application testbed


10000
7500
5000
2500
0
Laptop
Gadget
Fog
Kodiak
Viewmaster
Thunder
Capulin
Trinitite
Woodchuck

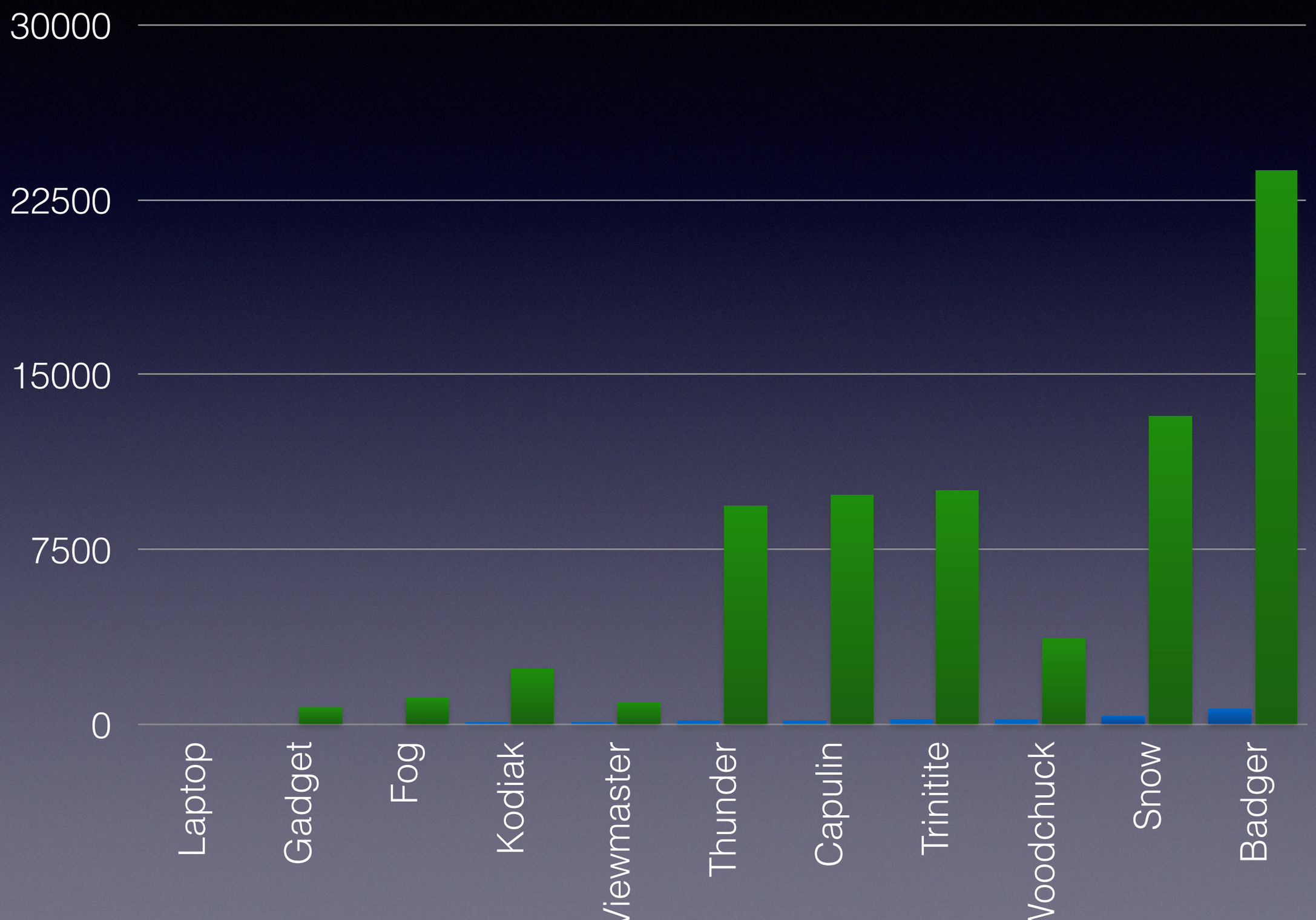
Snow

- 368 nodes
- 13,248 cores
- Yellow network
- Two SUs of CTS-1 nodes


14000
10500
7000
3500
0
Laptop
Gadget
Fog
Kodiak
Viewmaster
Thunder
Capulin
Trinitite
Woodchuck
Snow

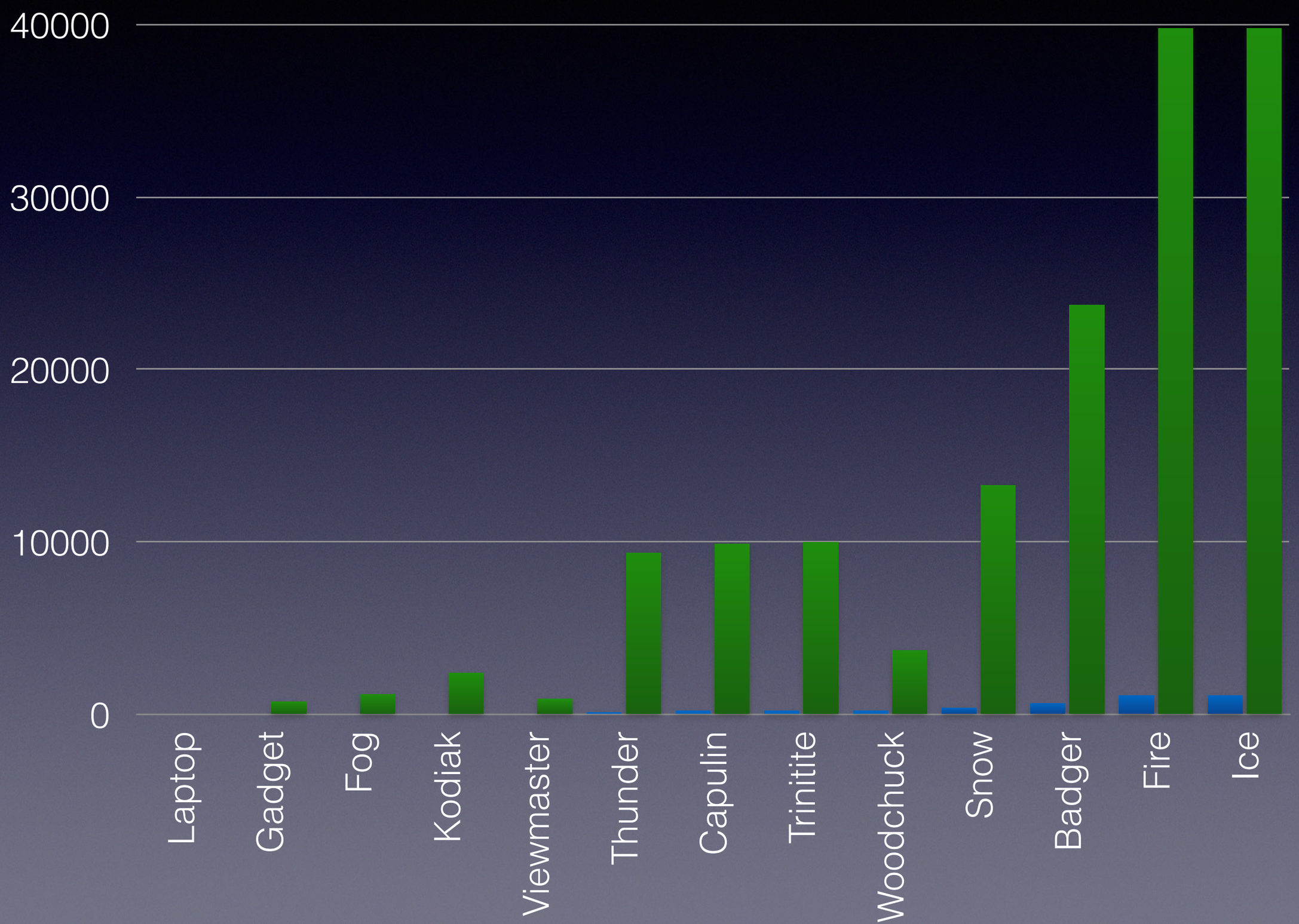
Badger

- 660 compute nodes
- 23760 cores
- Turquoise network
- 3.5 SUs of CTS-1 nodes


30000
22500
15000
7500
0
Laptop
Gadget
Fog
Kodiak
Viewmaster
Thunder
Capulin
Trinitite
Woodchuck
Snow
Badger

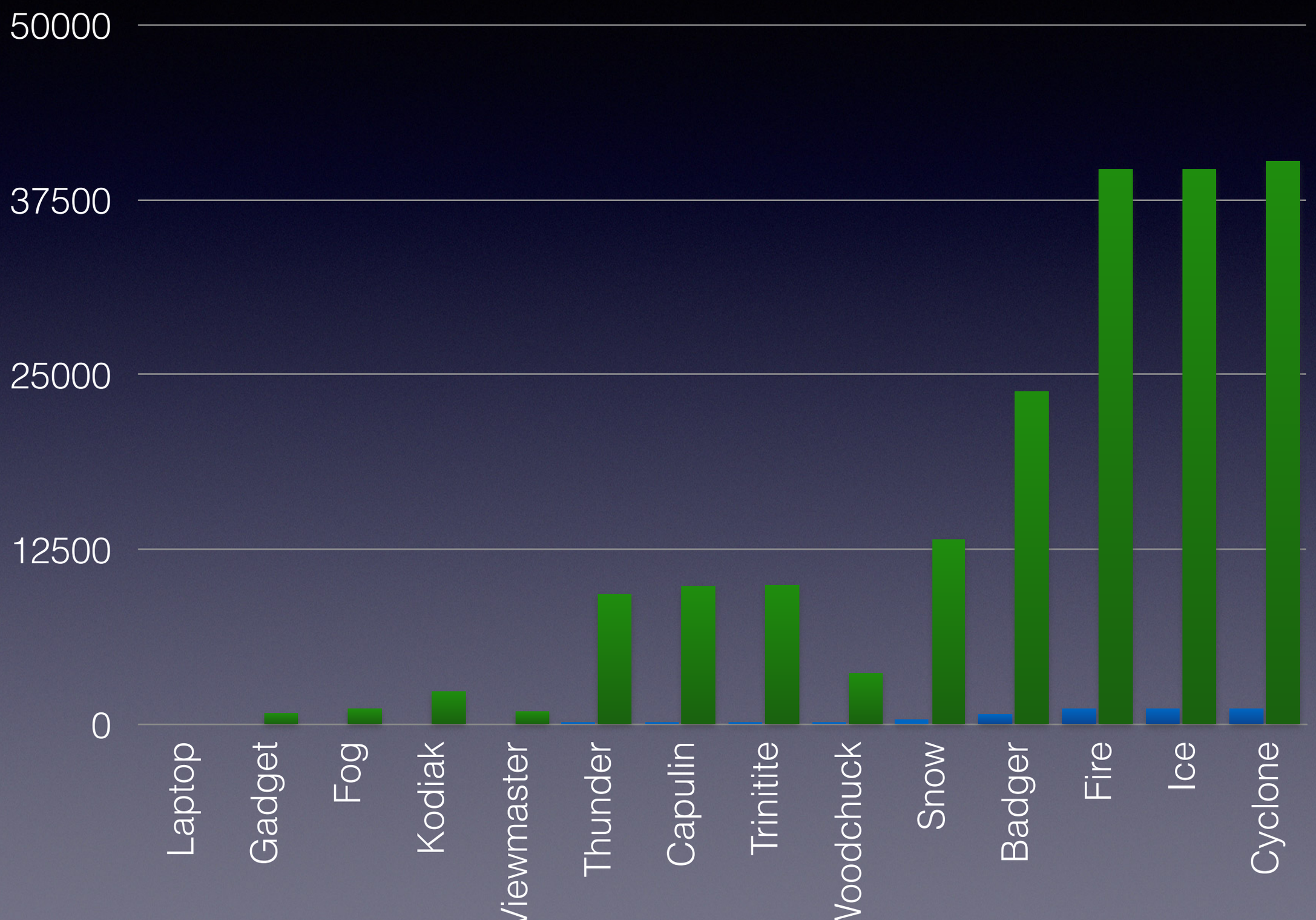
Fire and Ice

- 1,104 nodes each
- 39,744 cores each
- Red network
- Six SUs each of CTS-1 nodes


40000
30000
20000
10000
0
Laptop
Gadget
Fog
Kodiak
Viewmaster
Thunder
Capulin
Trinitite
Woodchuck
Snow
Badger
Fire
Ice

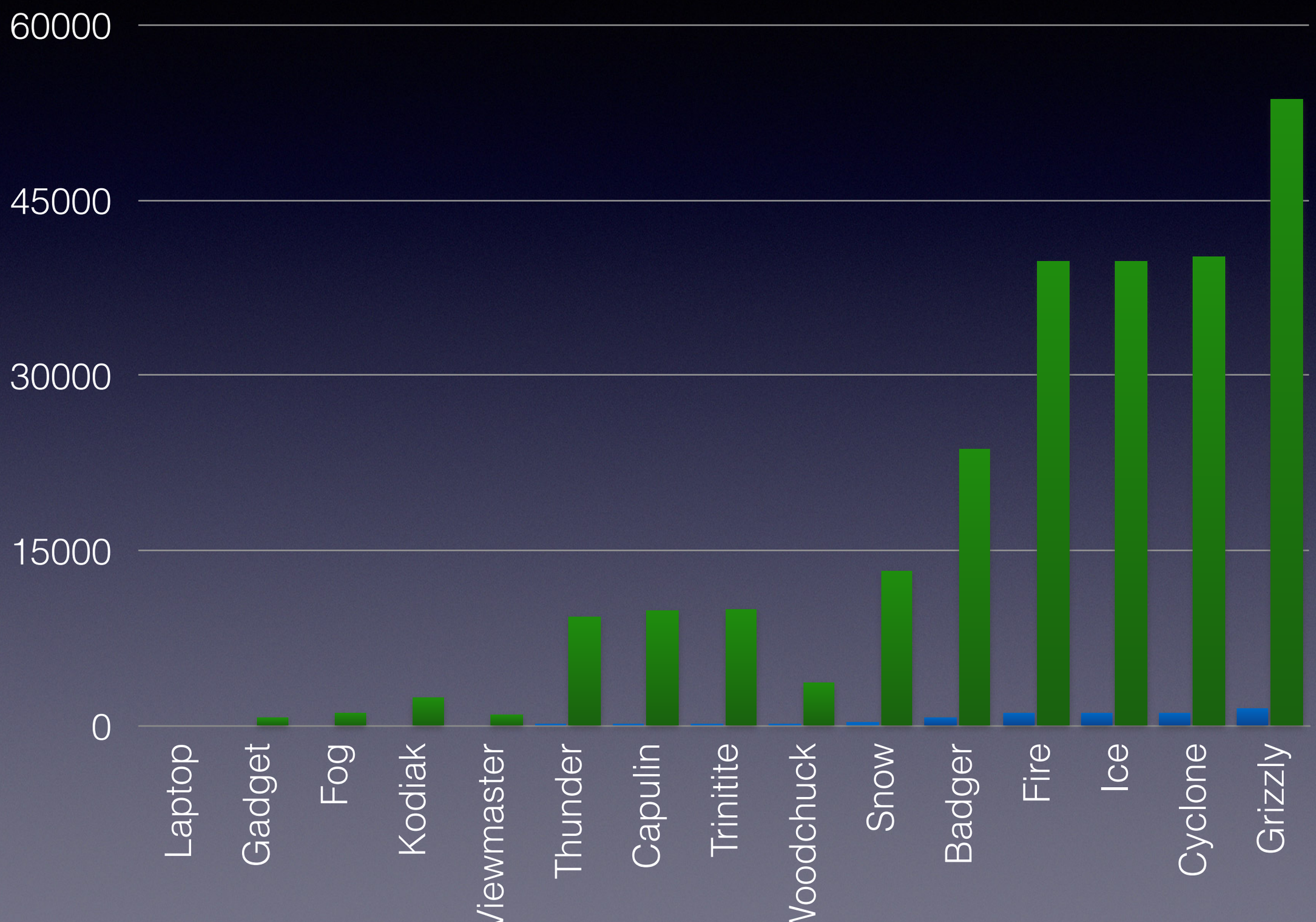
Cyclone

- 1118 compute nodes
- 40248 cores
- Red network
- Six+ SUs of CTS-1 nodes


50000
37500
25000
12500
0
Laptop
Gadget
Fog
Kodiak
Viewmaster
Thunder
Capulin
Trinitite
Woodchuck
Snow
Badger
Fire
Ice
Cyclone

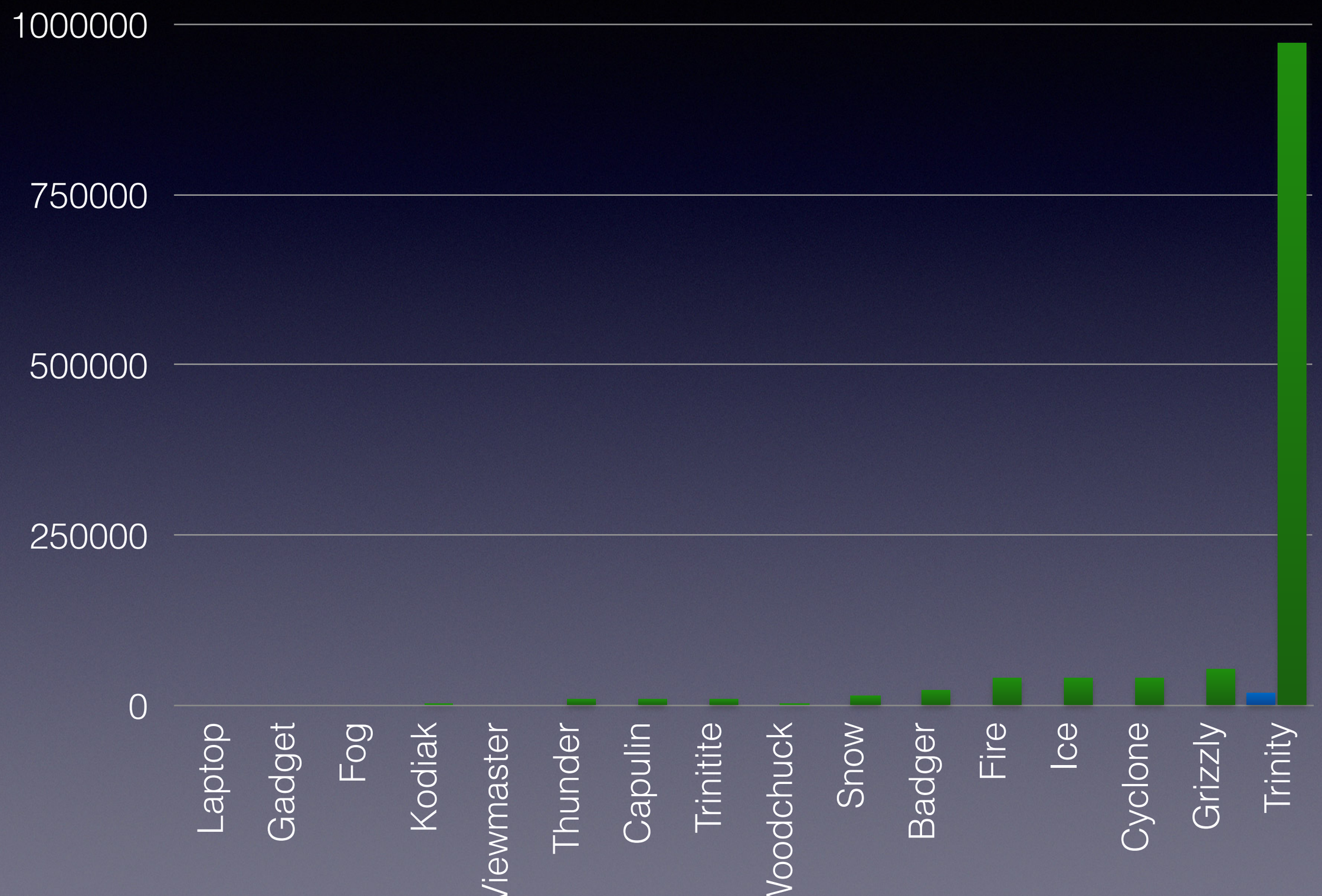
Grizzly

- 1490 nodes
- 53,640 cores
- Turquoise network
- Eight SUs of CTS-1 nodes
- Water cooled


60000
45000
30000
15000
0
Laptop
Gadget
Fog
Kodiak
Viewmaster
Thunder
Capulin
Trinitite
Woodchuck
Snow
Badger
Fire
Ice
Cyclone
Grizzly

Trinity

- ~9,400 Haswell Nodes
- ~9,800 Knights Landing Nodes
- ~970,000 cores total
- Red network
- Water cooled
- ATS-1 Production System


1000000
750000
500000
250000
0
Laptop
Gadget
Fog
Kodiak
Viewmaster
Thunder
Capulin
Trinitite
Woodchuck
Snow
Badger
Fire
Ice
Cyclone
Grizzly
Trinity

Plus testbeds!

- **Stretch**: ARM systems testbed
- **Shasta**: Cray systems testbed
- **Kit**: operating systems
- **Dora**: workflows and data analytics
- **Moonlight**: Next Generation System Software
- … and more.

Plus incoming systems!

- **CTS-2**: Follow-on to CTS-1
- **Crossroads**: Next generation ATS system

Plus weird systems!

- **Ising**: 2000 qubit Dwave quantum computer
- **The Pi Cluster**: 750 Raspberry Pi systems
- ... and some others

Totals

- ~26,200 compute nodes
- ~1,220,000 cores
- 2000 qubits

Part Two: Booting Systems at Scale

Cluster Cage Match

Booting

8 Nodes

vs.

20,000 Nodes

The Problem Statement

- You have a cluster
- It has 8 compute nodes
- Each compute node is diskless
- Each compute node has a 1Gb network connection
- How long does it take to boot this cluster?

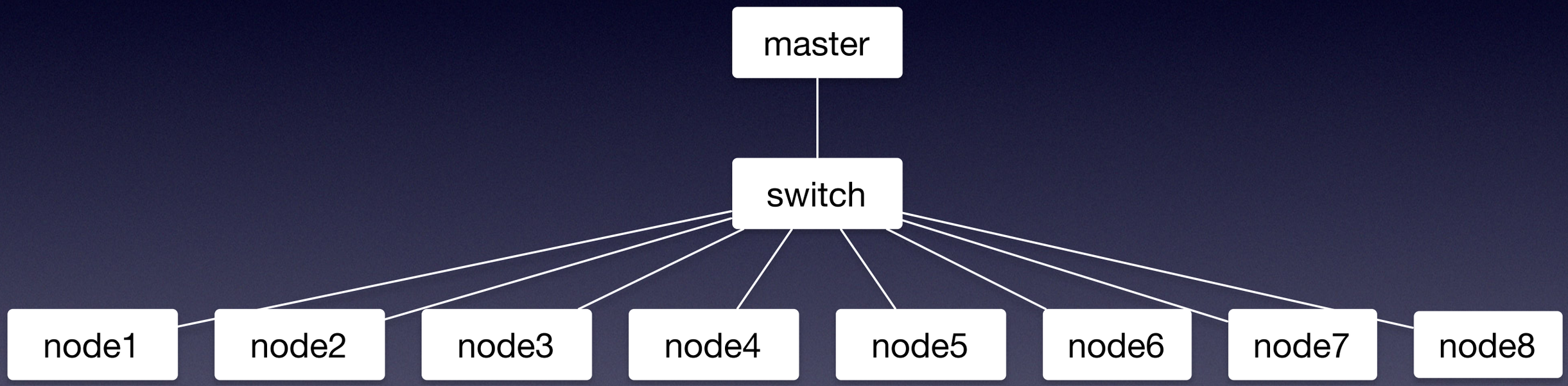
Questions

- How will this cluster boot?
  - Standard network boot
- How big is a node image?
  - 1GB
- Where will the node images live?
  - You can buy more hardware if needed
  - But be reasonable!
- What does the network layout look like?
  - See upcoming diagram

The Problem Statement

- You have a cluster
- It has 8 diskless compute nodes
- *It has one diskful master node*
- *Nodes network boot a 1GB OS image*
- Each compute node has a 1Gb network connection
- How long does it take to boot this cluster?

Cluster Design



The Boot Process

- Each compute node contacts the master node
- Each compute node downloads its RAM disk
- Each compute node boots from its downloaded image

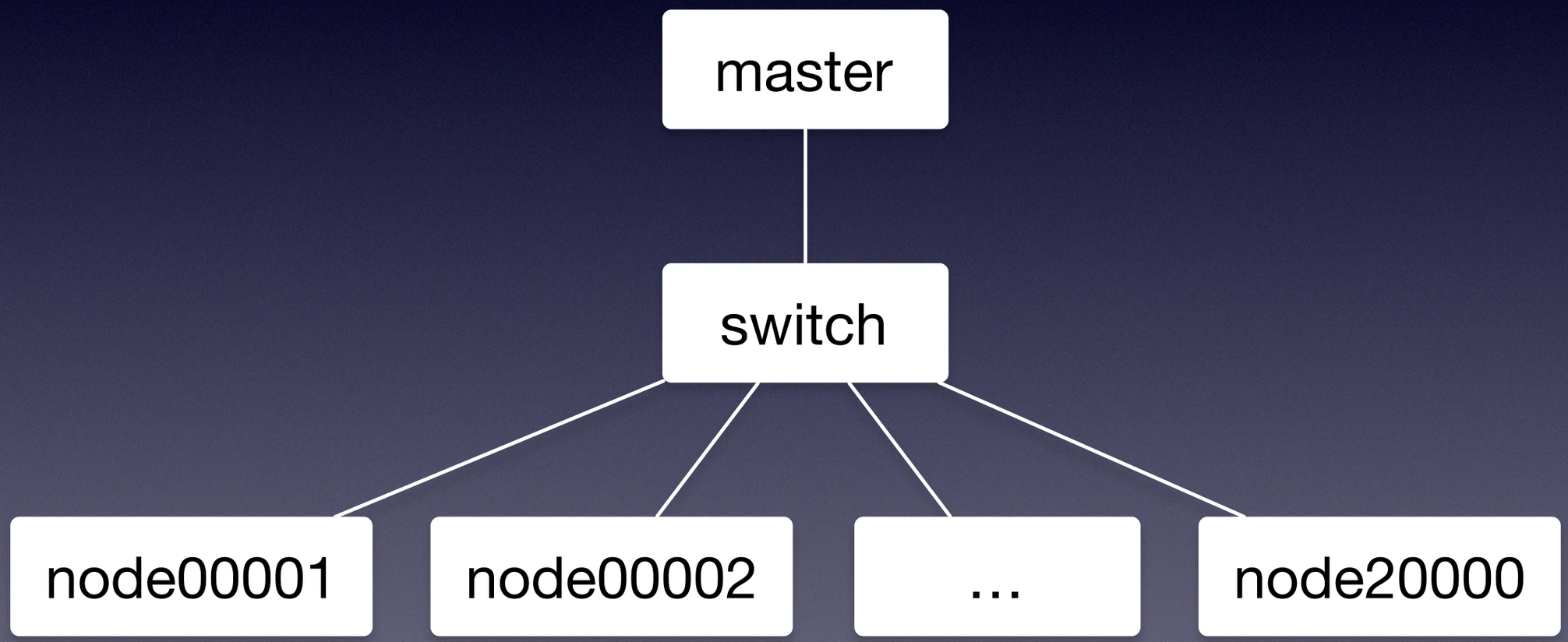
Math Time

- One cluster image: 1GB in size
- Transferring one cluster image over a 1Gb network
  - (1GB)/(1Gb/s) = (8Gb)/(1Gb/s) = 8s
- Full cluster boot time
  - 8s x 8nodes = 64s = 1m 4s
- No problem!

The Next Problem Statement

- You have a cluster
- It has 20,000 compute nodes
- Each compute node is diskless
- Each compute node has a 1Gb network connection
- How long does it take to boot this cluster?

Cluster Design



Math Time

- One cluster image: 1GB in size
- Transferring one cluster image over a 1Gb network
  - (1GB)/(1Gb/s) = (8Gb)/(1Gb/s) = 8s
- Full cluster boot time
  - 8s x 20,000nodes = 160,000s = 1d 20h 25m 0s
- Problem!

Let’s Optimize

- 1GB OS image?!
  - That’s 8GB of duplicated memory wasted on the small cluster
  - That’s 20TB of duplicated memory wasted on the big cluster
- Minimize the image
  - Microkernel: ~250MB image in memory
  - Shared root: ~250MB image in memory

Back to the Math

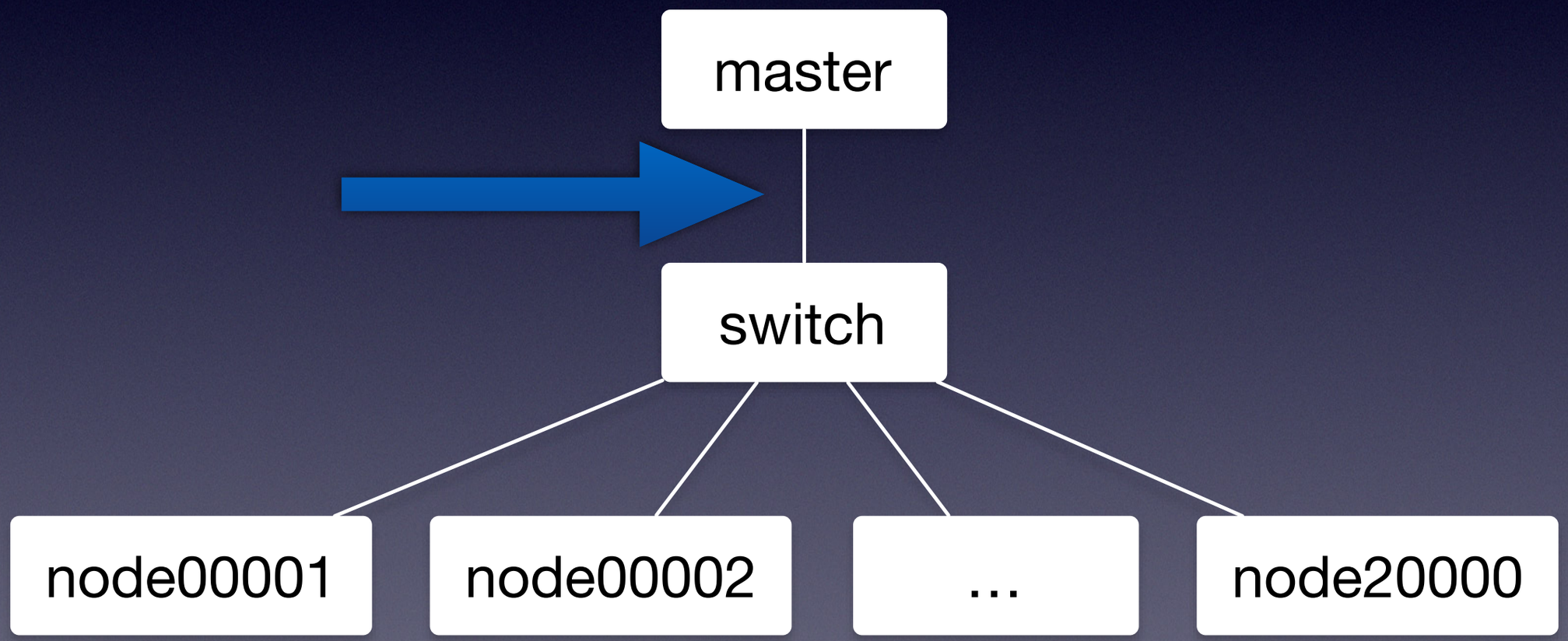
- (250MB) / (1Gb/s) = (2Gb) / (1Gb/s) = 2s
- 8 node cluster:
    - 8 nodes x 2s = 16s
- 20,000 node cluster:
    - 20,000 nodes x 2s = 40,000s = 11 h 10m 0s

How long do we have?

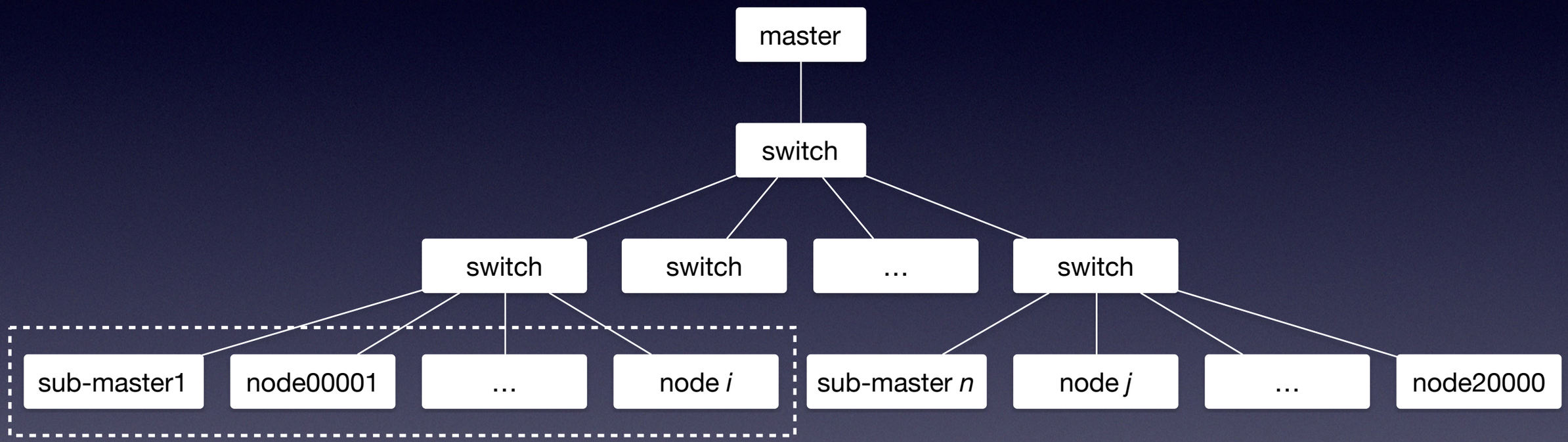
- What’s an acceptable boot time?

3.6.6 A complete system initialization should take no more than 30 minutes. The Offeror shall describe the full system initialization sequence and timings.

Where is the bottleneck?



Hierarchical Design



How many master nodes?

- Or: How big should a scalable unit be?
- What if we have…
  - 100 sub-masters
  - 200 nodes per sub-master

Even more math

- First, boot the sub-masters
  - 100 sub-masters x 2s = 200s = 3m 20s
- Next, boot the compute nodes
  - 200 per sub-master x 2s = 400s = 6m 40s
- Total: 600s = 10m
- Add overhead of POST, DHCP requests, and kernel booting, etc. and we’re still doing alright

So many more details

- How do you schedule 20,000 nodes?
- How do you monitor 20,000 nodes?
- How do you add a new user to 20,000 nodes?
- If one node fails, how do you find it?
- If one network cable fails, how do you find it?
- And then how do you replace it?

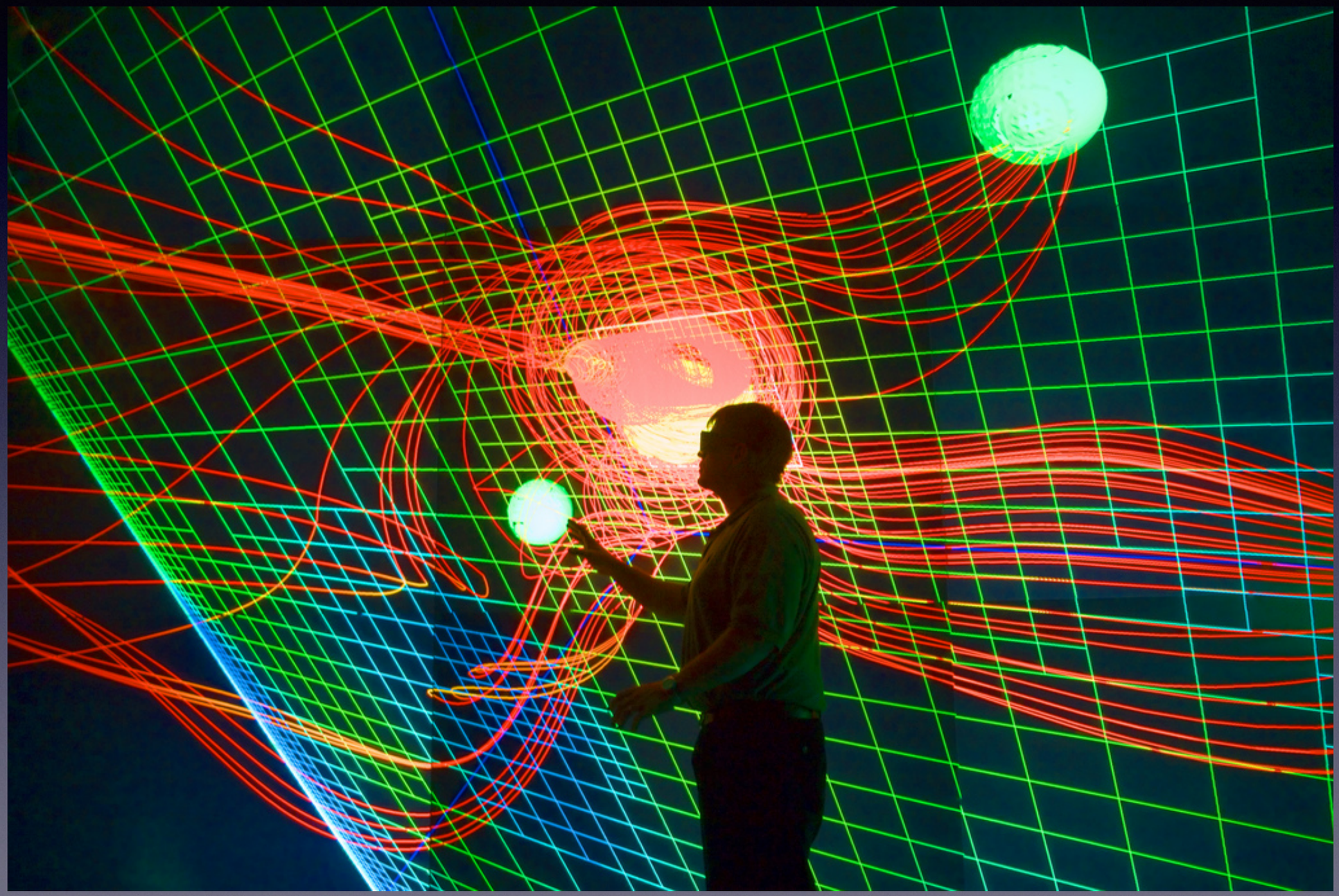
Why did we do all of this?

- Working on large systems requires thinking about scale at all times
- This is true for all aspects: building the systems, writing applications on the systems, managing the systems, salvaging the systems
- If you get one thing out of this talk, this is it: thinking about scale at all times is important

Scale

Scale

Scale



Volume

1	=	7.158e+51
Cubic light-year		US stick of butter

More info

Feedback

Questions?

- Cory Lueninghoener
- cluening@lanl.gov
- http://hpc.lanl.gov