LVS plus GridFTP is the new Bestman

ACAT 2017

Aug 21 2017

Edgar Fajardo (Presenting),
Brian Bockelman, Clifton Pottberg, Garhan Attebury,
Terrence Martin, Frank Würthwein
On behalf of OSG Software and Technology

What is Bestman?

best man | ˈbɛst ˈmæn |

noun *[in singular]*

a male friend or relative chosen by a bridegroom to assist him at his wedding.

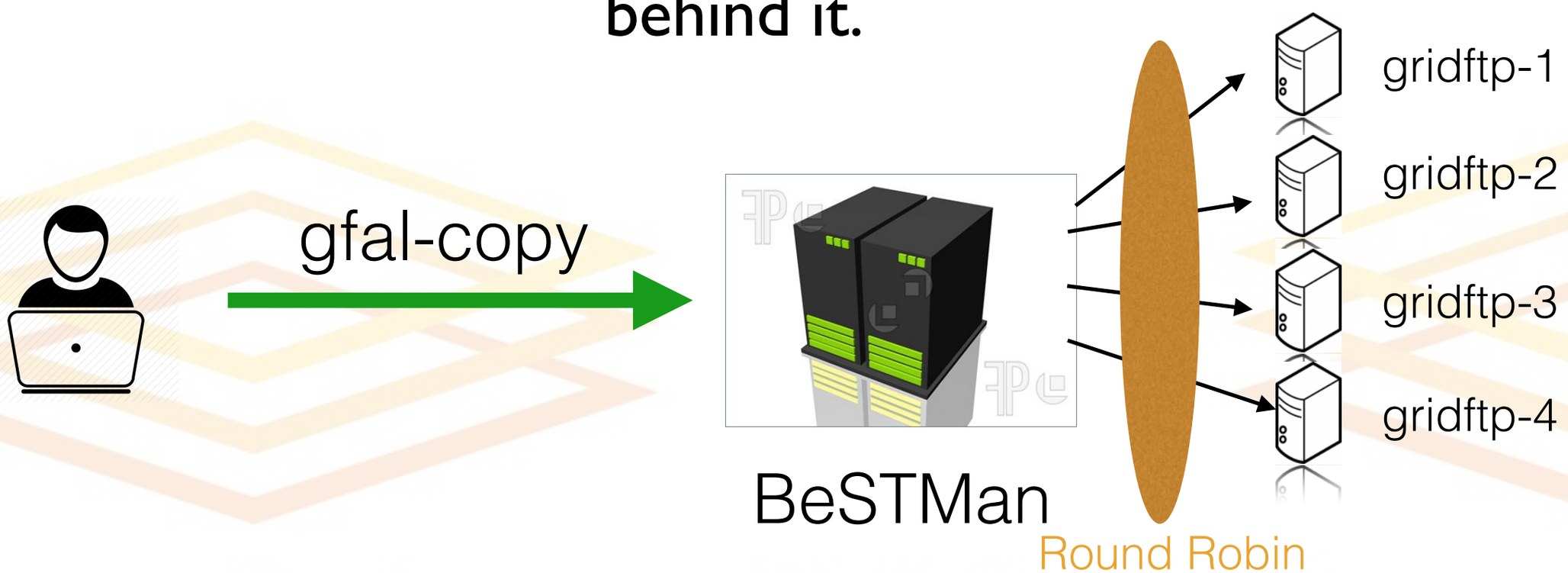
According to LBNL folks:

Berkeley Storage Manager (BeStMan)

BeStMan is a full implementation of SRM v2.2, developed by Lawrence Berkeley National Laboratory, for disk based storage systems and mass storage systems such as HPSS. End users may have their own personal BeStMan that manages and provides an SRM interface to their local disks or storage systems. It works on top of existing disk-based unix file system, and has been reported so far to work on file systems such as NFS, PVFS, AFS, GFS, GPFS, PNFS, and Lustre. It also works with any existing file transfer service, such as gsiftp, http, https and ftp. It requires the minimal administrative efforts on the deployment and maintenance.

What does Bestman provide?

- Implementation of the SRM protocol
- Users (VO) and sites only have to advertise one BestMan server and "hide" several GridFTP servers behind it.



Why to replace BeStMan

- LBNL dropped support for it in 2012. And since then OSG has been maintaining it.
- Experiments and sites still need a way to expose their Massive storage systems to the Grid.
- Part of making the OSG middleware layer thinner.
- OSG will end support for BestMan end of 2017 with the end of life of OSG 3.3 release.
- OSG 3.4 does NOT bring BestMan

Hence a solution is needed

What is LVS?: Linux Virtual Server

In the words of the developers:

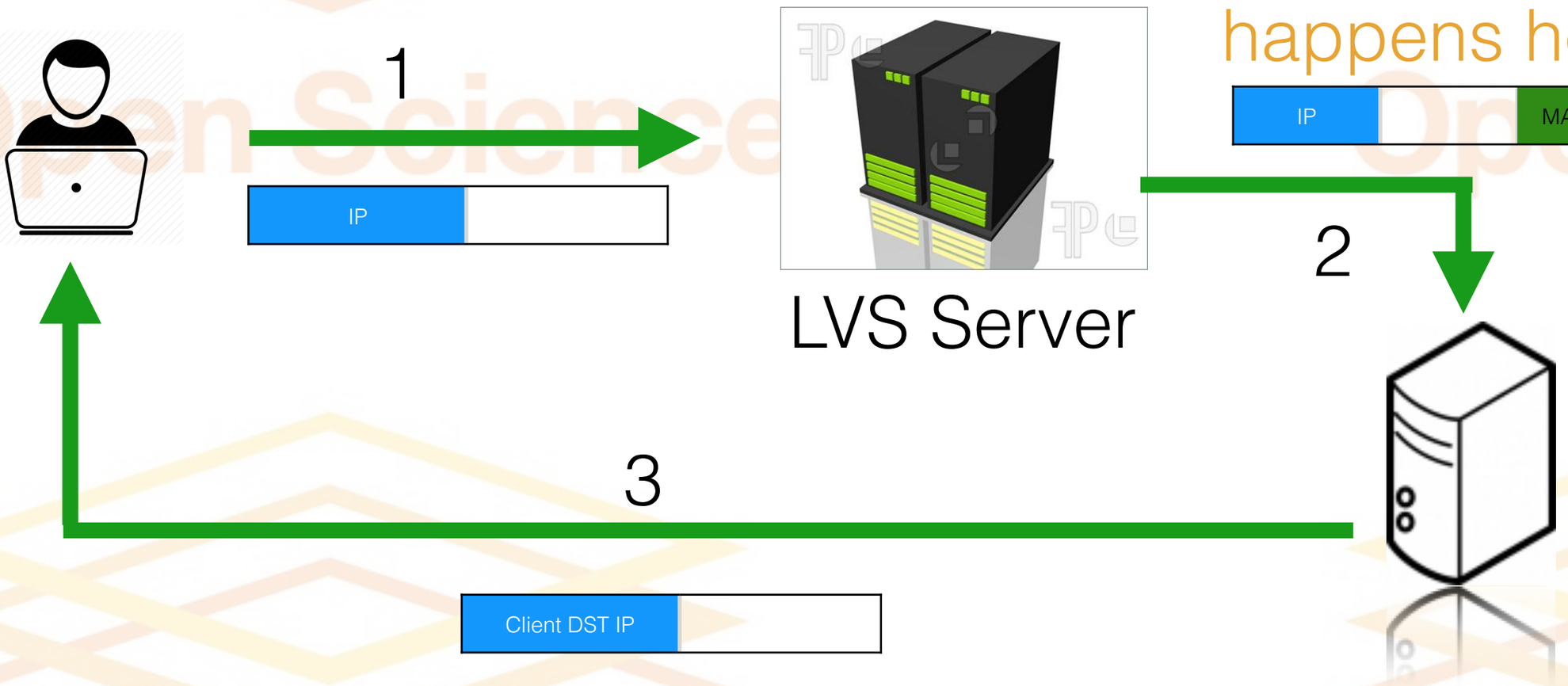
What is the Linux Virtual Server?

The Linux Virtual Server is a highly scalable and highly available server built on a cluster of real servers, with the load balancer running on the Linux operating system. The architecture of the server cluster is fully transparent to end users, and the users interact as if it were a single high-performance virtual server. For more information, click here.

Some of the features include

- It is maintained as part of the core linux code.
- Load balancing
- Discovery of malfunctioning servers

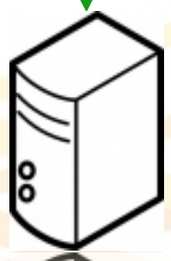
How LVS works?



LVS meets gridFTP

gfal-copy

load balancing
packet forwarding



gridftp-i

Data

Control

Same in IPv4 than IPv6 without the ARP spoofing

Scale Tests

After all is setup then it is time to break something:

Scenario Conditions:

- 1000/2000/3000/4000/5000 parallel jobs
- An LVS server with 6 gridftp running on blade servers with 10 Gbit ethernet cards
- gfal-copy every 10 secs from hadoop to /dev/null

How to test it?
Use a sleeper pool

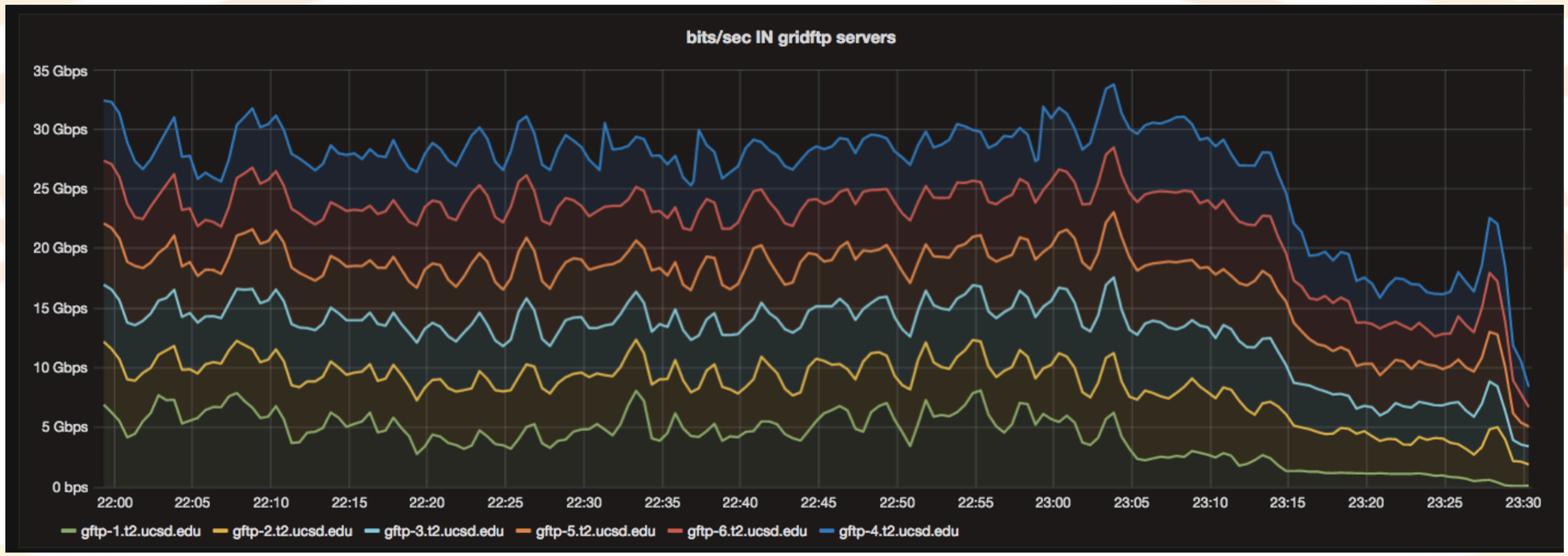
Not like this one:

Like This one:

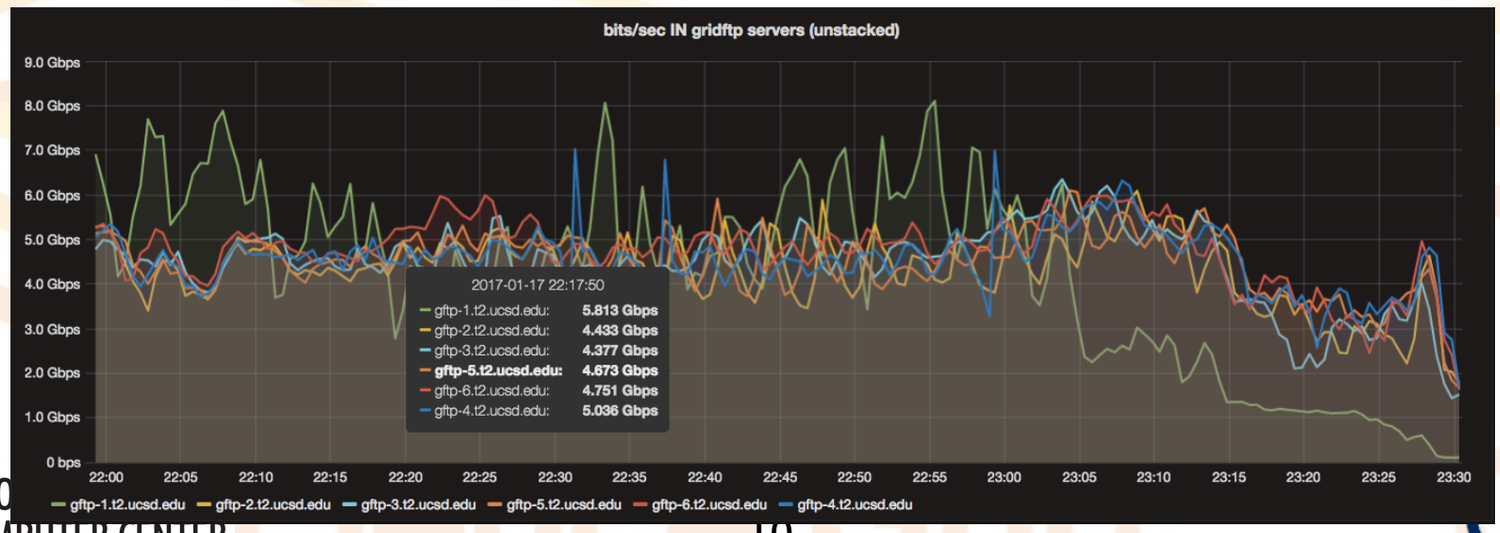
A worker node

Scale Tests

At 2000k jobs we got
a
nice throughput of
~30gbit/sec:

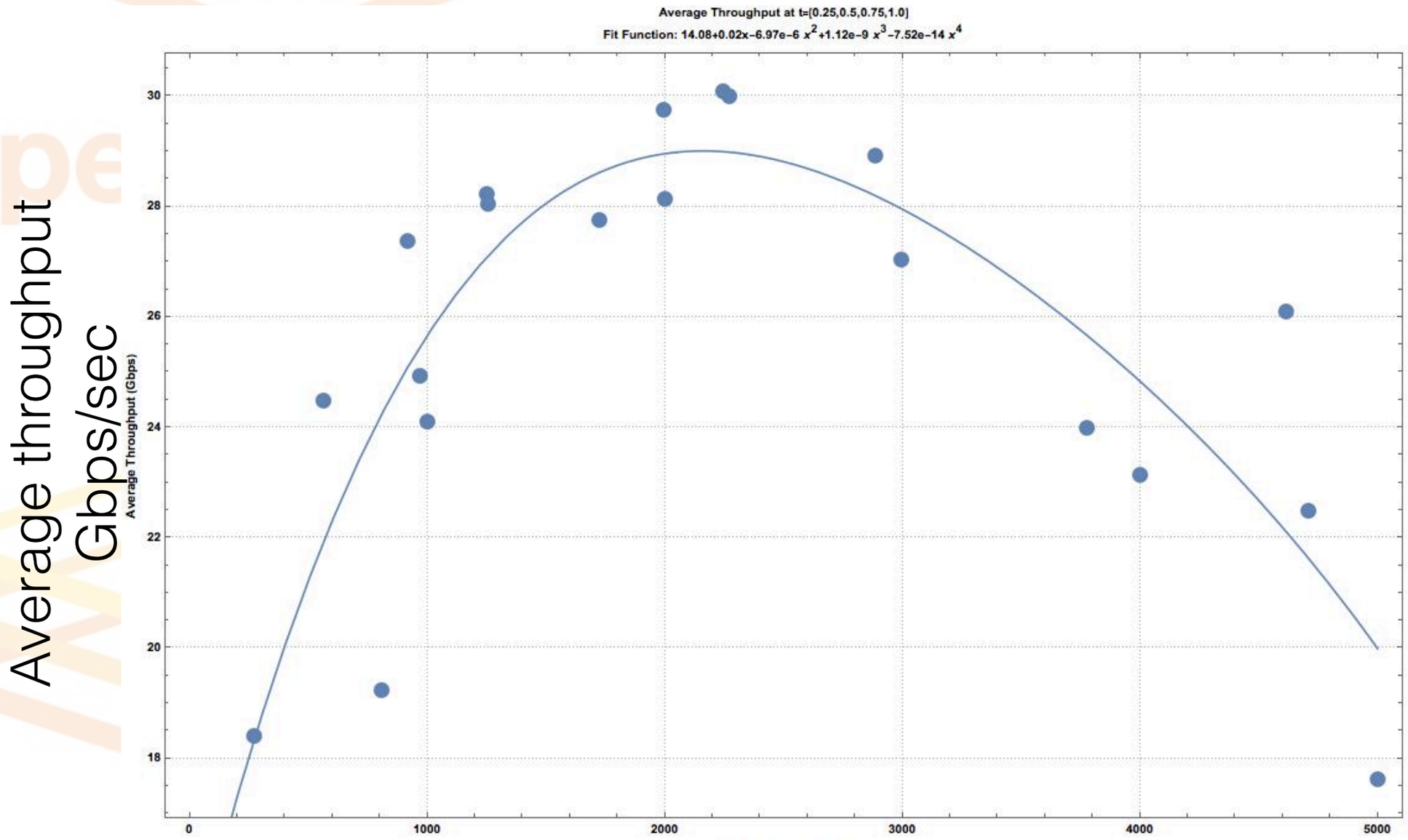


And the load balancer is working:



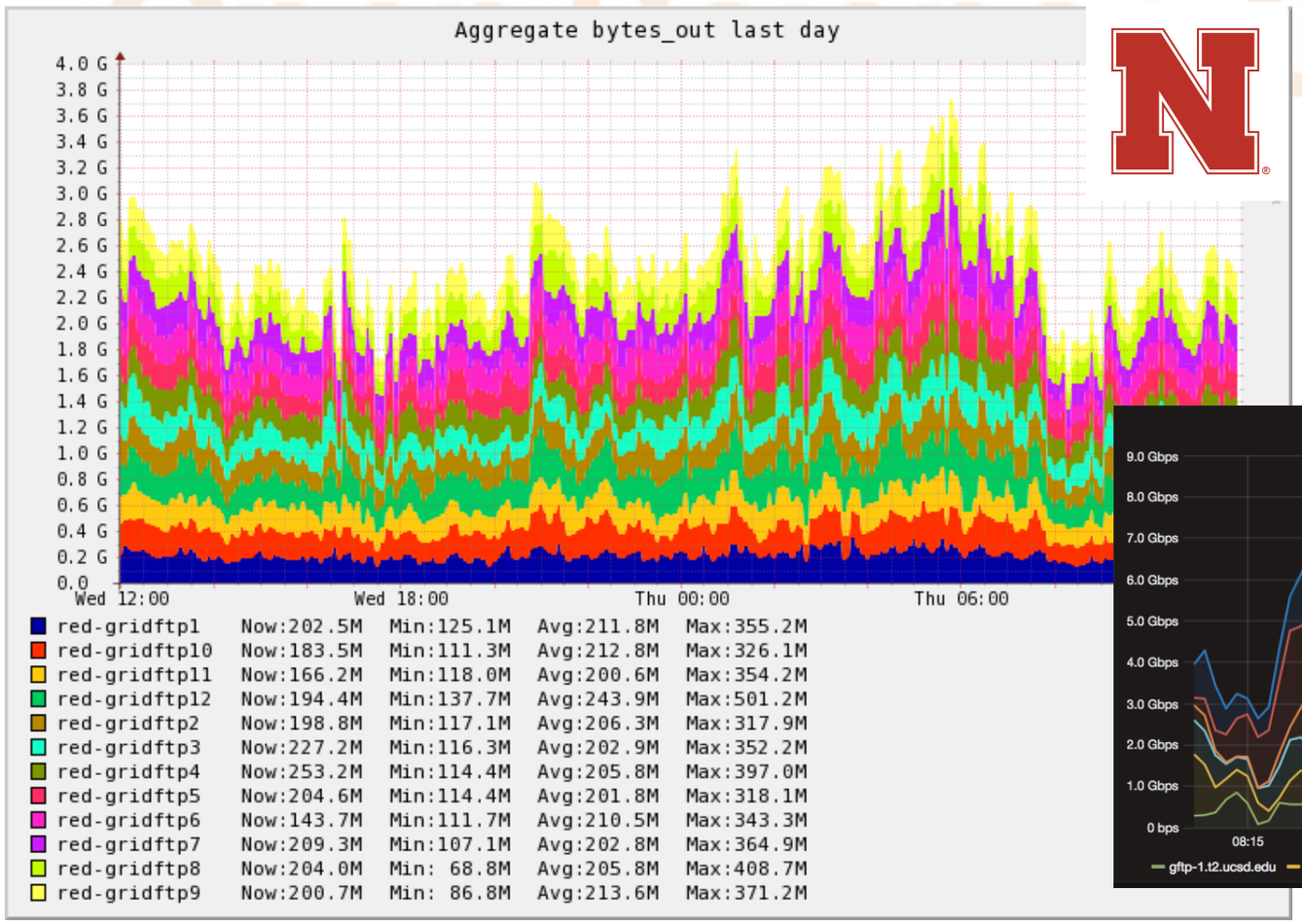
Scale tests

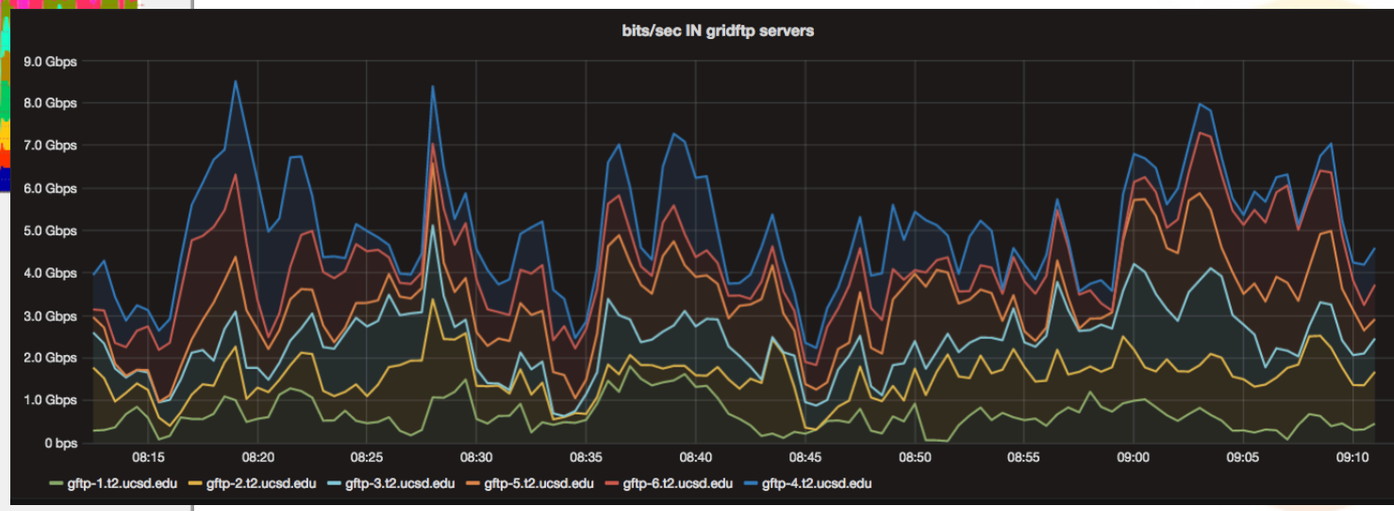
In general this is how it scales with the number of jobs



Achievements

Usual days at Nebraska and UCSD





Conclusions

- The new LVS plus gridFTP scales to the needs of sites and beyond.
- More problems will show up when sites with different network setups drop BestMan for GridFTP.
- LVS plus gridFTP has proven effective as a BestMAN replacement.

Acknowledgements

- Garhan Attebury and Brian Bockleman for pioneering the idea of the LVS+GridFTP at Nebraska
- Terrence Martin for setting the sleeper pool and the LVS+GridFTP at UCSD
- Clifton Pottberg for conducting and documenting the large scale tests

Questions?

Contact us at:

1-900-LVS-GridFTP-Masters

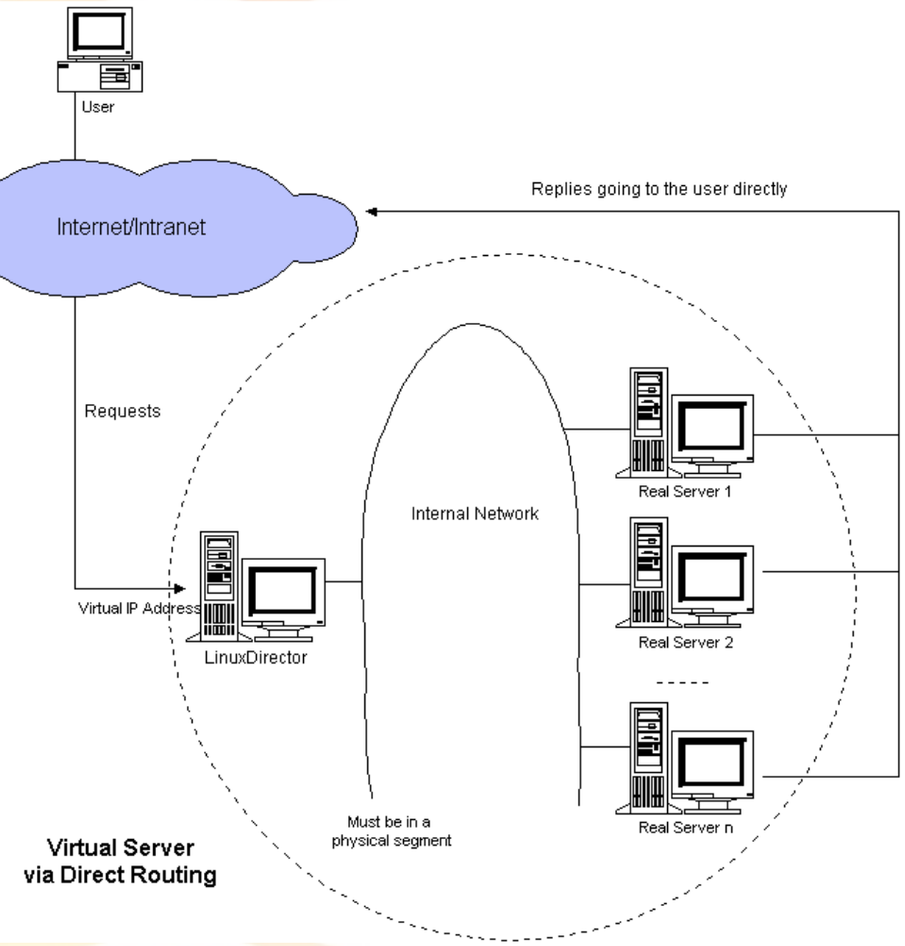
Just Kidding

Contact us:

osg-software@opensciencegrid.org

Thank You

Supporting slides



From the LVS page:
LVS with Direct Routing