

Science as a Team Sport

... With Universities as crucial enabler.

Frank Würthwein
OSG Executive Director
Professor of Physics
UCSD/SDSC



The purpose of this presentation is to give the Council a summary of the OSG Sales Pitch to Universities, and solicit feedback.

These slides have been used in a variety of places with some locally specific customizations.

- Across the nation, **institutions invest into research computing to remain competitive**
- **Science is a team sport**, and institutions with an island mentality will underperform.
- **Integration is key** to success

**OSG provides services
to advance this vision**

- US Universities commonly operate:
 - compute clusters from 10k to 40k cores
 - storage infrastructures from 1-10 Petabytes
 - campus networks that connect labs at 10Gbps with 100Gbps WAN connectivity.
 - multi-million \$\$ annual budgets for research computing groups.

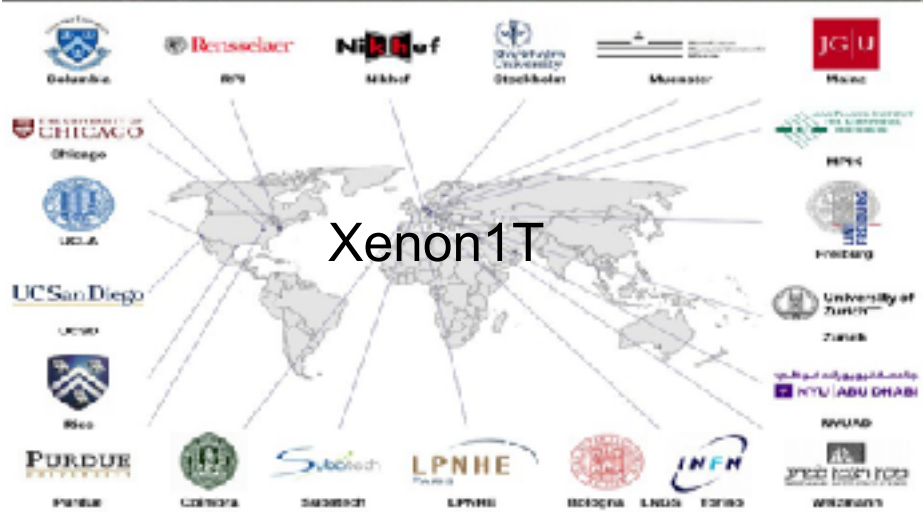
Research benefits from integrating these investments across institutions.

I'll give some obvious examples from Physics & Astronomy but inter-institutional collaborations can be found in all domains.

Science is a Team Sport



The IceCube Collaboration



Xenon1T

- ~100 members, 20 institutions
- 24 non-affiliated members
- +35 associate members
- Smithsonian Astrophysical Observatory
- Adler Planetarium
- Argonne National Lab
- Barnard College / Columbia University
- Bartol Research Institute / University of Delaware

- Georgia Institute of Technology
- Iowa State University
- Purdue University
- University of California, Los Angeles
- University of California, Santa Cruz
- University of Chicago
- University of Iowa
- University of Minnesota

- University of Utah
- Washington University in St. Louis
- McGill University, Montreal
- University College Dublin
- Cork Institute of Technology
- Galway-Mayo Institute of Technology
- National University of Ireland, Galway

VERITAS



SPT3G

**OSG advances the local, national and international
integration of Cyberinfrastructure
in support of Open Science.**

Even “more moderate size” research endeavors than LIGO, ATLAS, or CMS involve **dozens of institutions across multiple countries** that need to be able to **share their resources** to **maximize their scientific throughput !!!**

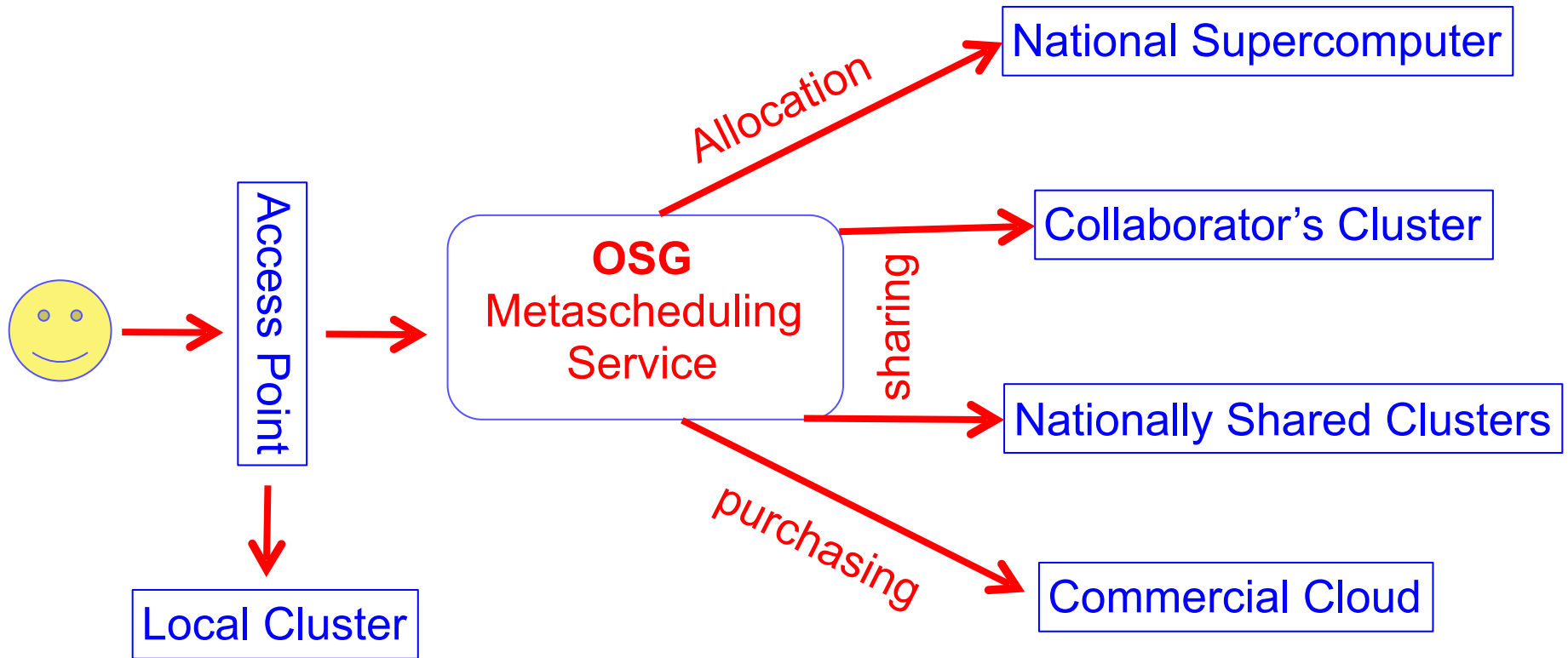
**OSG empowers researchers to use compute & data
resources across institutional boundaries.**

Integrating Compute & Storage Clusters



Each green dot represents a cluster registered in OSG

Transparent Computing across different resource types



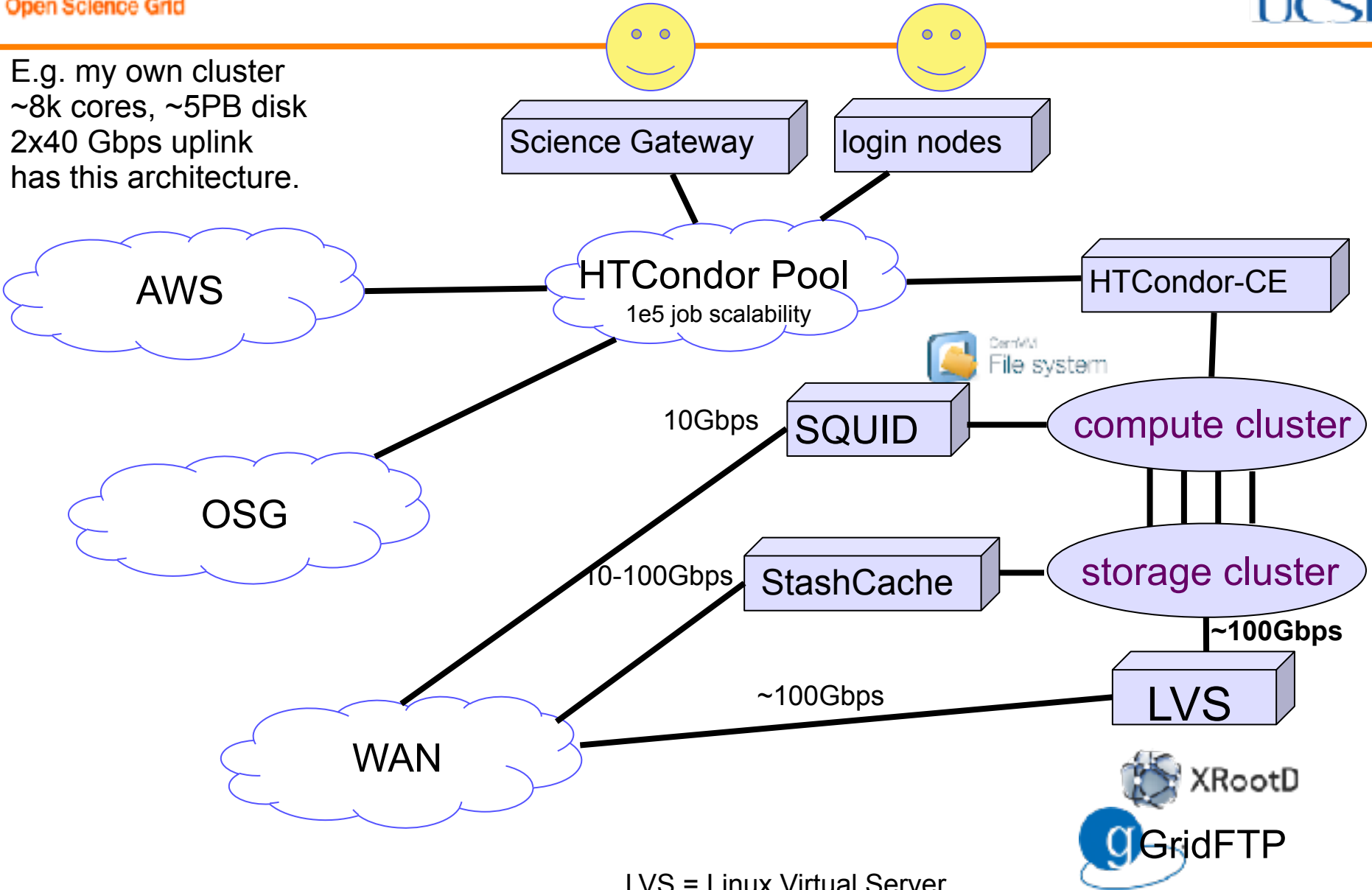
OSG integrates computing across different resource types and business models.

- OSG's business model is to empower Scientists and their home institutions to work together for long term sustainability.
 - OSG provides knowledge & software infrastructure.
 - OSG can offer storage and service hosting to jump start projects, but prefers to enable institutions for growth and sustainability.
- OSG provides global integration across commercial and academic computing.
 - OSG respects local ownership and control.

Some “Technical” Details

An elaborate OSG Site

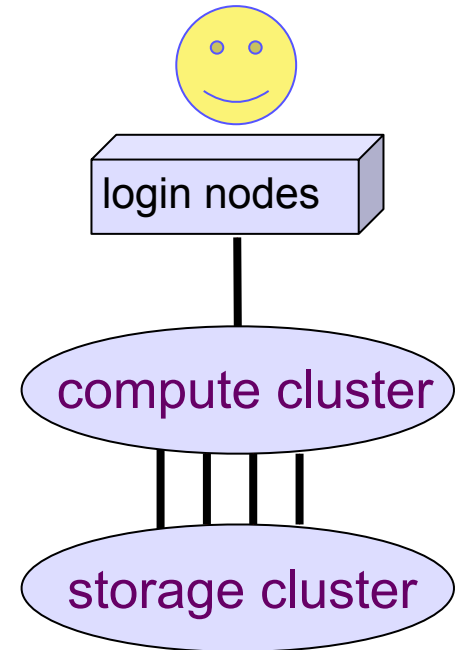
E.g. my own cluster
 ~8k cores, ~5PB disk
 2x40 Gbps uplink
 has this architecture.



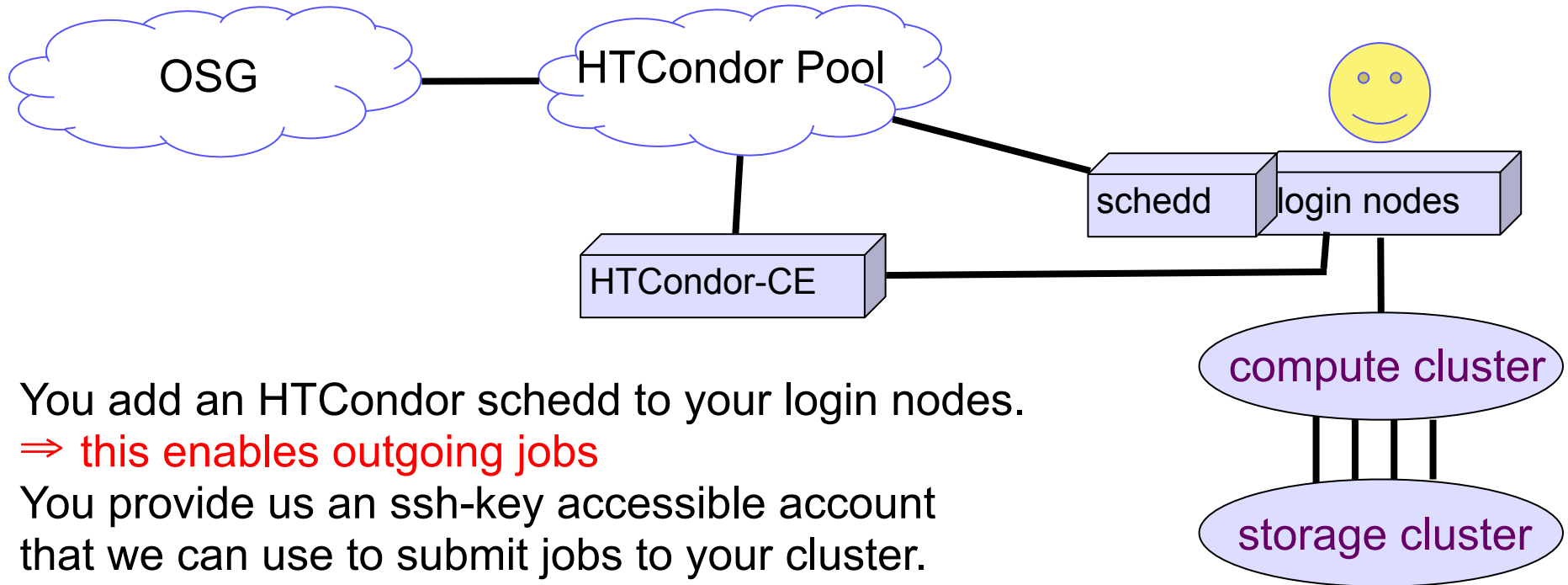
LVS = Linux Virtual Server

Let's built it up piece by piece

This much everybody knows how to build.



Let's built it up piece by piece



You add an HTCondor schedd to your login nodes.

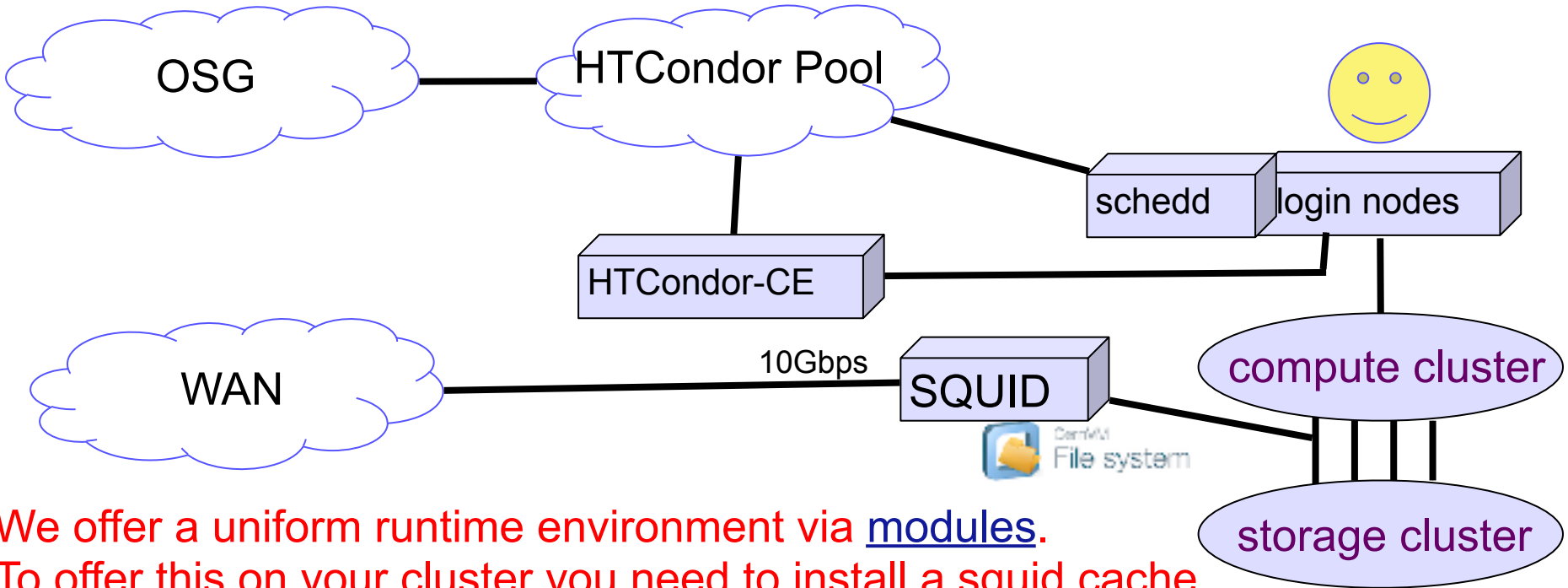
⇒ this enables outgoing jobs

You provide us an ssh-key accessible account that we can use to submit jobs to your cluster.

⇒ this enables incoming jobs

We operate for you an HTCondor pool, and CE and configure it with your policies. If you want more dynamic control over your policies than we can teach you how to run those pieces yourself.

Let's built it up piece by piece

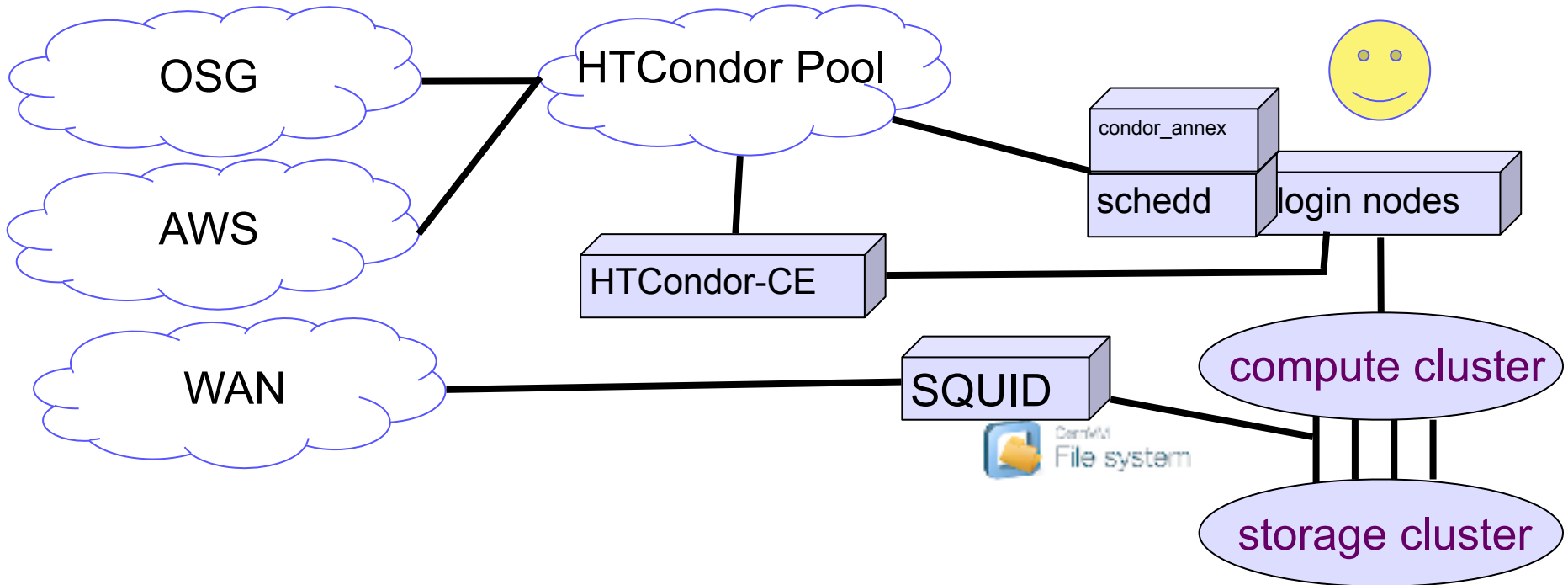


We offer a uniform runtime environment via [modules](#).
To offer this on your cluster you need to install a squid cache,
and fuse mount cvmfs on each of your worker nodes.

In addition to all the usual popular libraries & apps, we offer singularity containers this way that provide complete Machine Learning environments.

We are happy to add your favorite libs & apps as additional modules, or your singularity containers.

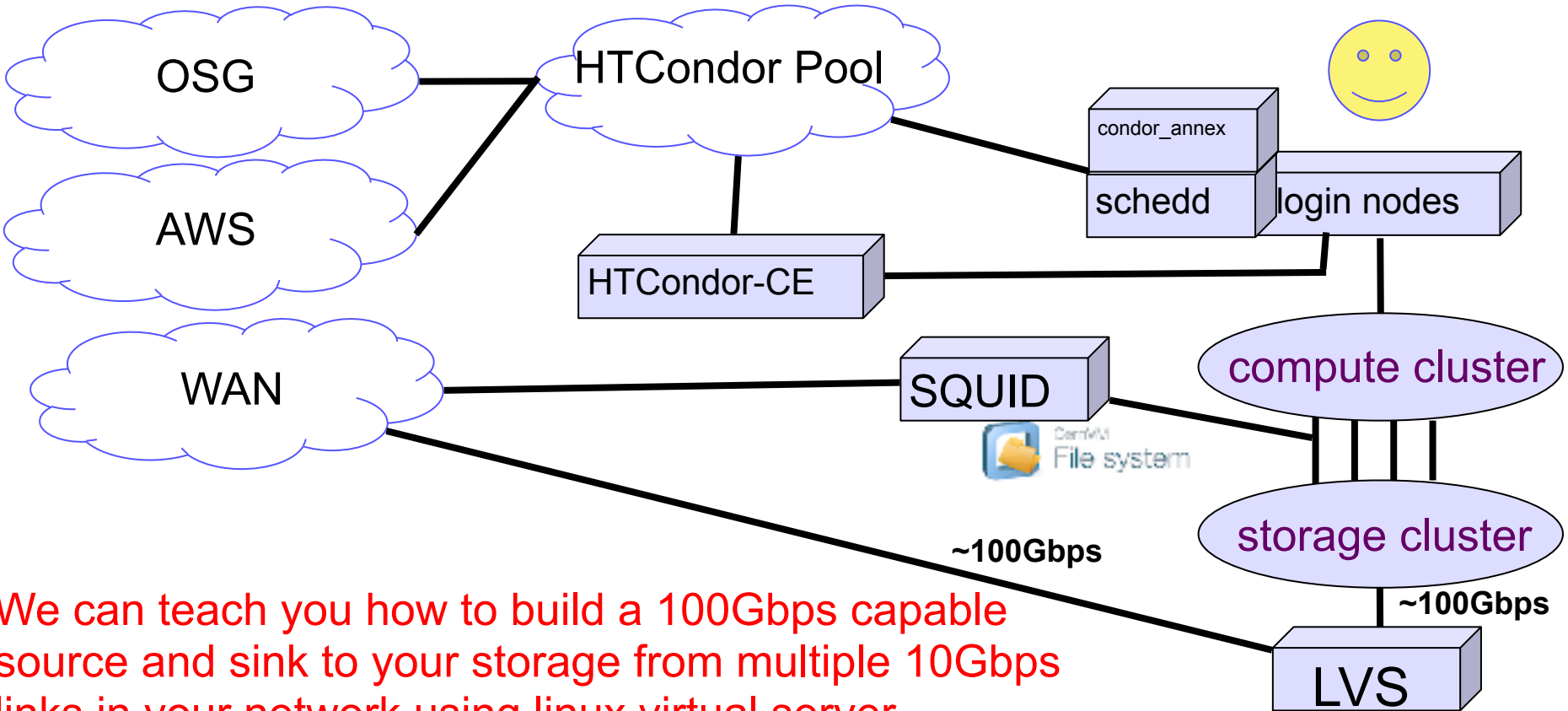
Let's built it up piece by piece



We offer “condor_annex” to seamlessly integrate AWS into the compute platform you are offering to your researchers.

Google and Microsoft will follow in the future.

Let's built it up piece by piece

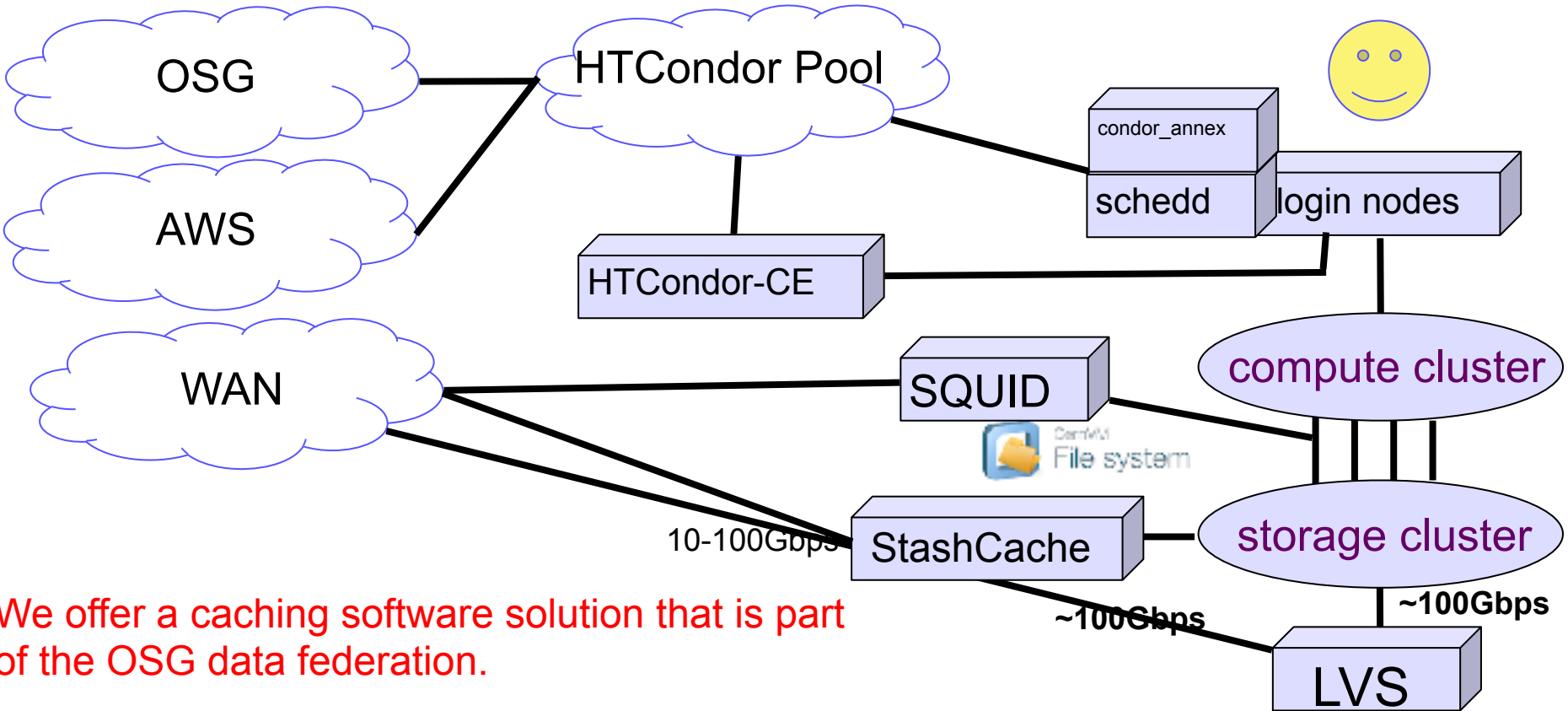


We can teach you how to build a 100Gbps capable source and sink to your storage from multiple 10Gbps links in your network using linux virtual server.

We are protocol agnostic for this, and require no service fees.



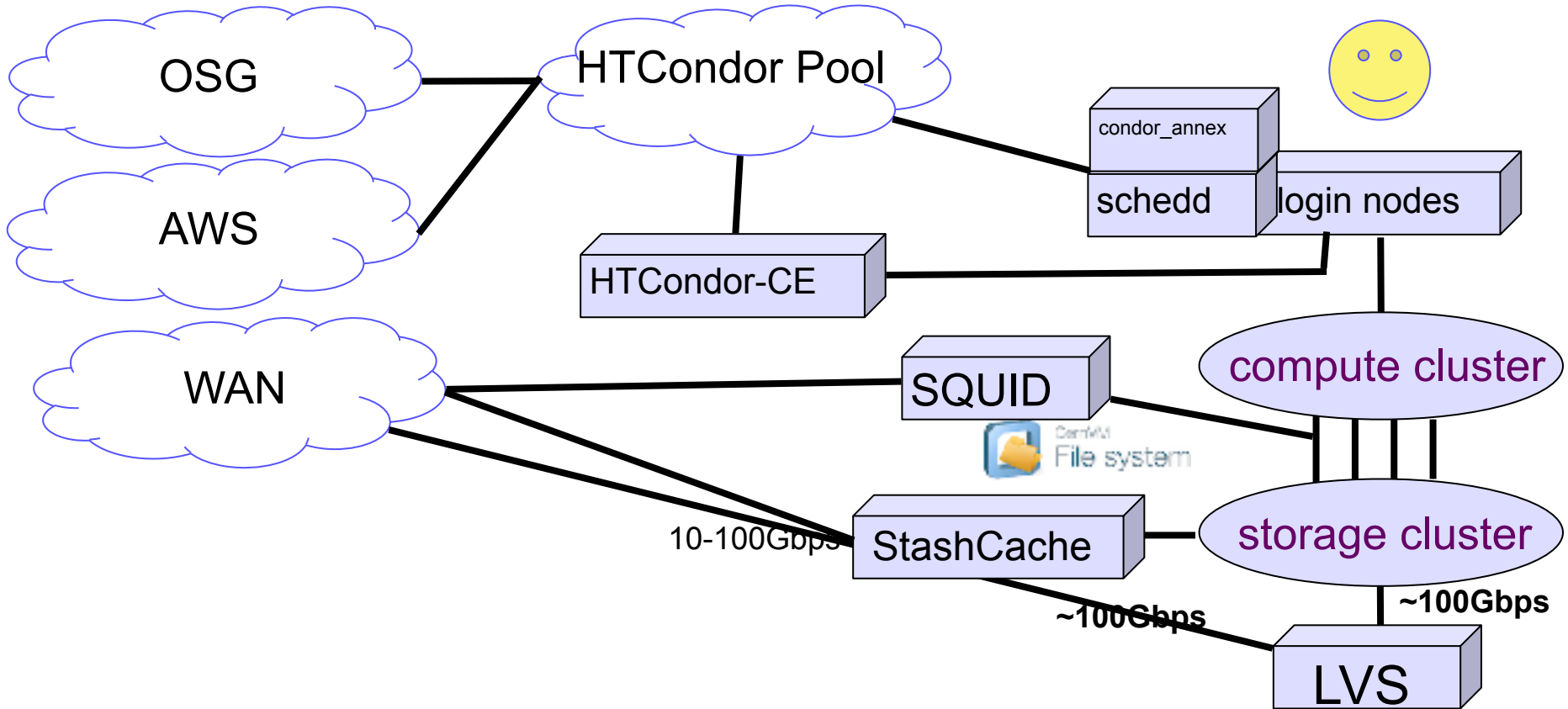
Let's built it up piece by piece



We offer a caching software solution that is part of the OSG data federation.

We teach you how to build your own data federation inside the OSG data federation such that your researchers' data can be cached in other caches in OSG.





You decide how much of this CI you want.



- OSG offers a variety of services to
 - individual researchers
 - research labs
 - Institutes
 - Universities
- Services range from consulting to operating distributed computing for your researchers
- Our goal is to empower you to support the researchers at your institution.

Let's work together to support Science !



opensciencegrid



[osgusers](#)



user-support@opensciencegrid.org



support.opensciencegrid.org



www.opensciencegrid.org/links