



The Open Science Data Federation

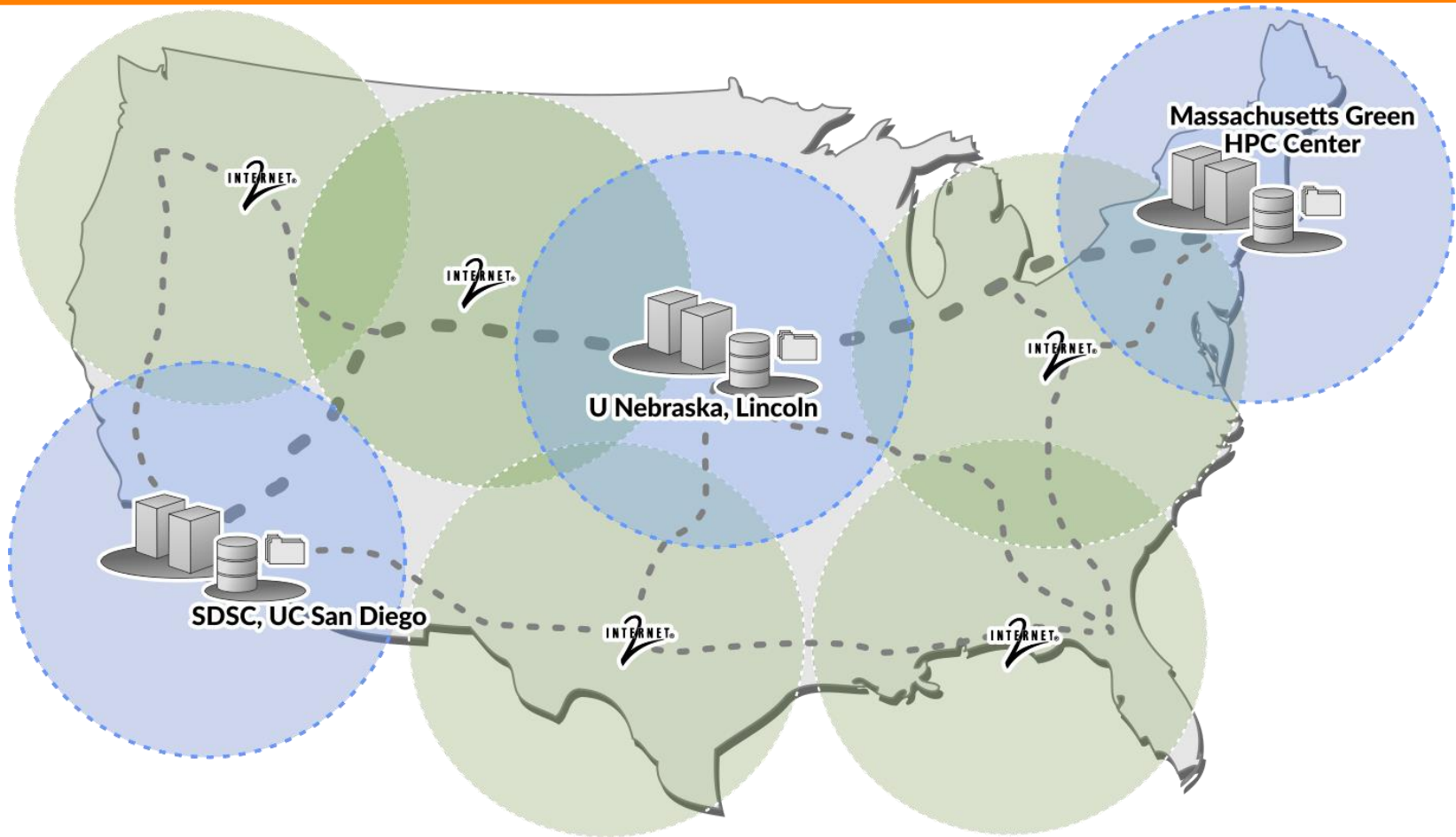
Frank Wuerthwein
Executive Director, OSG





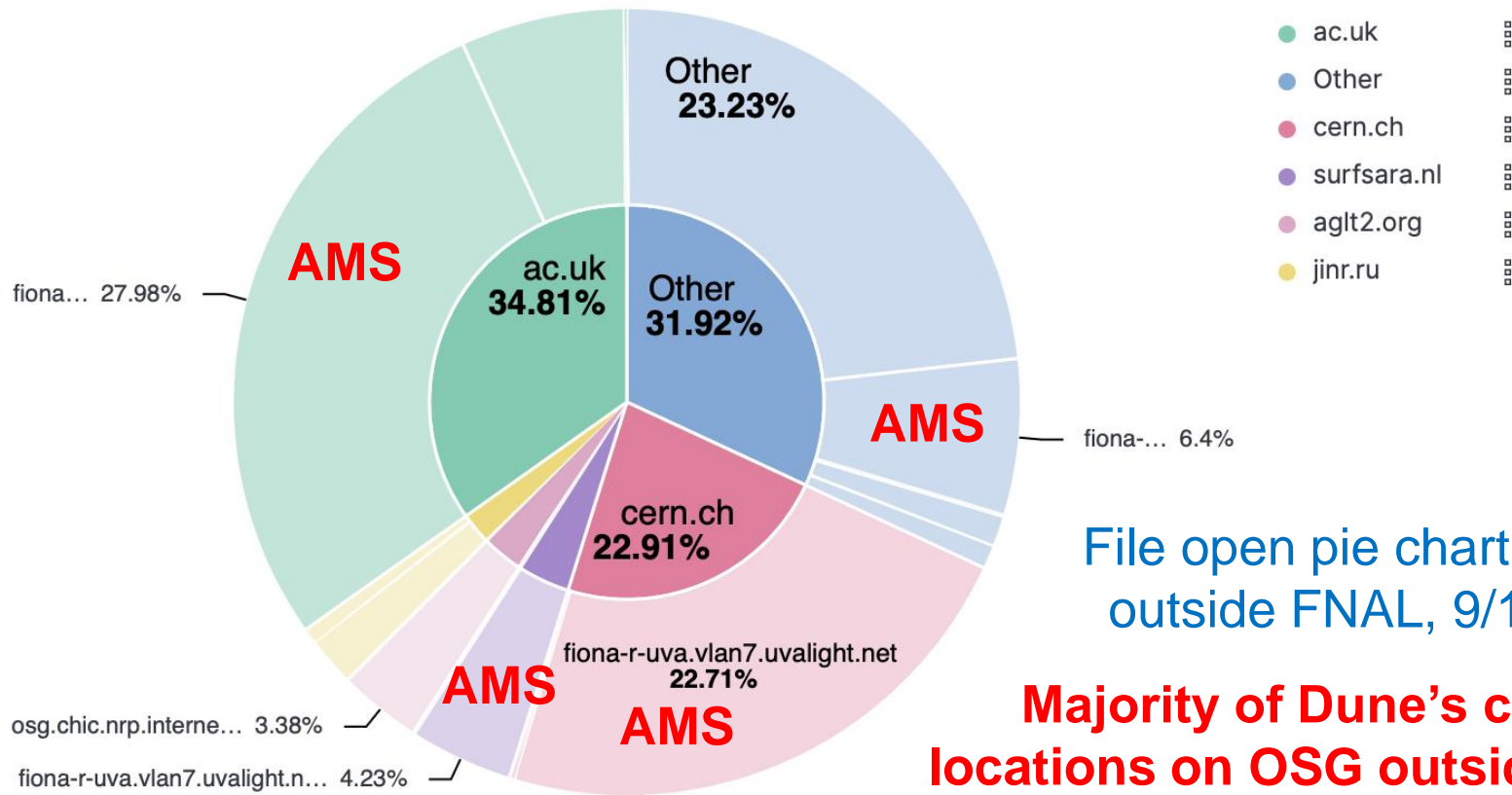
20 Caches ... 6 of which are in R&E network backbone
10 Data Origins ... incl. one in Europe

- Origins are places that own data.
- Caches are places via which data is accessed.
- By having a network of caches that spans the entire USA, we can provide access to any data that lives in any origin from anywhere.
 - **Any Data, Anytime, Anywhere**



By 12/2022 we will have deployed 8 additional caches via NSF 2112167
“Any” location within the continental USA is within 500 miles of a cache.

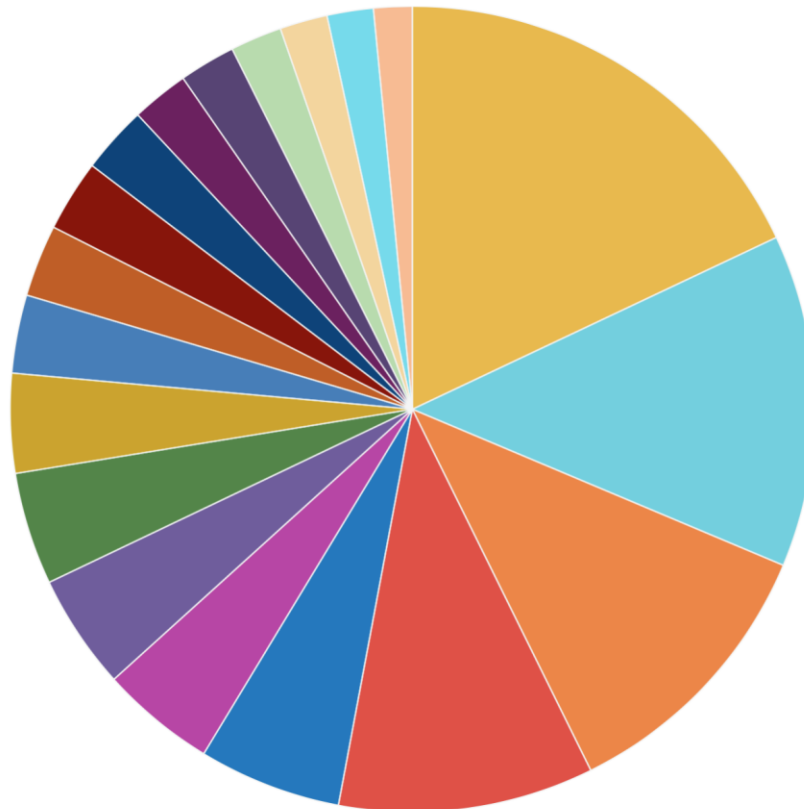
- ESnet will deploy caches for OSDF in London and Amsterdam
 - The Amsterdam cache has been one of our oldest and most heavily used ones.
 - Mostly LIGO and DUNE, but also some Nova and Minerva use.
 - UK resources hit the Amsterdam cache. Good to add one in London.
 - These locations are much more difficult to fund with NSF money.



File open pie chart for Dune outside FNAL, 9/1 to 9/29.

Majority of Dune's compute locations on OSG outside FNAL hit the Amsterdam cache.

Core Hours by Facility



Facility	total
FermiGrid	42 Mil
T3_US_NERSC	2 Mil
CERN-PROD	2 Mil
RAL-LCG2	1 Mil
SURFsara	1 Mil
UKI-SOUTHGRID-RALPP	730 K
PIC	590 K
UKI-NORTHGRID-MAN-HEP	589 K
IN_TIFR	575 K
FZU	507 K
UKI-LT2-QMUL	400 K
UKI-SOUTHGRID-BRIS-HEP	368 K
UKI-NORTHGRID-LANCS-HEP	366 K
NIKHEF-ELPROD	344 K
UKI-LT2-IC-HEP	293 K
BNL-SDCC	288 K
IN2P3-CC	264 K
SU ITS	244 K
UConn-HPC	237 K
JINR_Cloud	198 K

Roughly 1/4 of core-hours come from EU

- The Open Science Data Federation (OSDF) is about to undergo significant growth
 - Via the addition of caches across the continental USA, Amsterdam and London
 - Via the addition of origins in response to NSF 22-582

- This work was partially supported by the NSF grants OAC-2112167, OAC-2030508, OAC-1841530, OAC-1836650

