“Fabrics” ?

- from “network fabric”, highspeed cluster networks, eg. InfiniBand
- roughly the Clusters and Farm group:
  - HTC (high throughput computing) Farm: HTCondor:
    - many jobs per node, no multinode, “embarrassingly parallel”, few GPUs
    - Scale! >2k nodes, HS23/$, HS23/Watt, (HEPScore23 domain specific perfmetric)
    - ATLAS Tier-1 (AT1)
    - sPHENIX
    - spool (shared pool: STAR, PHENIX, AT3, Belle II, DUNE, EIC, Astro, …)
  - HPC (high performance computing) Clusters: Slurm: ~200 nodes
    - multinode jobs, high speed cluster interconnects (eg. IB), more GPUS
    - IC2 (institutional cluster gen 2: CFN, LQCD, CDS(nee CSI))
    - NSLS-II
- Misc other systems/responsibilities/expertise: SciServer, PDs, export, nnsd, jupyter, …
- ~2.6k systems

SCDF Org Chart

Director : Alexei Klimentov (interim)
Deputy Director : Antonio Wong
Leadership Team : C.Caramarcu, J.DeStefano, H.Ito, J.Liu,
ex officio : Group Leaders, Ops Officer, RP Manager, ISSO, Director & Deputy

Technical Advisory Board

Johannes Elmsheuser [ATLAS] - Co-Chair
Stuart Wilkins [NSLS II] - Co-Chair
Adolfy Hoisie [CDS]
Jerome Lauret [STAR] – scientific secretary
Sara Mason [CFN]
Chris Pinkenburg [sPHENIX]
Cedric Serfon [Belle II]

Fabrics Group
Matt Cowan

Costin Caramarcu
Kevin Casella
(Zhihua Dong+)
Shigeki Misawa
Tom Smith
Oszkar Tarjan

Network Team
(Mark Lukasczyk*-)
(Frank Burstein-)
(Nick Fontana-)
(Andrew Martinson-)

Facilities Operations Team
Ops Officer - Imran Latif*
(Naveed Anwer)
Enrique Garcia
Al Maruf
John McCarthy

Storage & Infrastructure Group
Jason Smith

Infrastructure Team
Joseph Frith*
Mark Berry
Robert Hancock
[James Leonardi]
Tommy Tang

HPSS Team
Tim Chou*
Ognian Novakov
Justin Spradley
[Yingzi (Iris) Wu]

Storage Team
Zhenping (Jane) Liu*
Carlos Gamboa
Vincent Garonne
Qiulan Huang
James Leonardi
Yingzi (Iris) Wu

Services & Tools Group
Ofer Rind

(Dmitry Arkhipkin^)
Doug Benjamin
(Uma Ganapathy+)
Ivan Glushkov
Hironori Ito

Users Services Team
John DeStefano*
Saroj Kandasamy
Christian Lepore
Louis Pelosi

Admin Support
Leisa McGee

Resources Planning & Management
[Shigeki Misawa]

Information System Security Officers
[Robert Hancock]
[Shigeki Misawa]

Contacts
ATLAS : Ofer Rind
Belle II : Hironori Ito
CFN : Shigeki Misawa
DUNE : Douglas Benjamin/ Hironori Ito
NSLS II : Joseph Frith
sPHENIX : Antonio Wong
STAR : Shigeki Misawa
EIC : Alexei Klimentov

** : team leader and technical lead*
[name] : secondary section

() : not supervised at SCDF
^ : NPP supervisor
+ : CDS supervisor
- : ITD supervisor

Brookhaven National Laboratory

May 1, 2025. v 1.42

Fabrics Team

Costin Caramarcu: technical lead, slurm, troubleshooting/debugging, config mgmt, …

Kevin Casella: hw, system lifecycle planning, condor, broad institutional knowledge, …

Matt Cowan: Team lead, troubleshooting/debugging, optimization, networking, …

Zhihua Dong: slurm, gpu hardware and software optimization, k8s, …

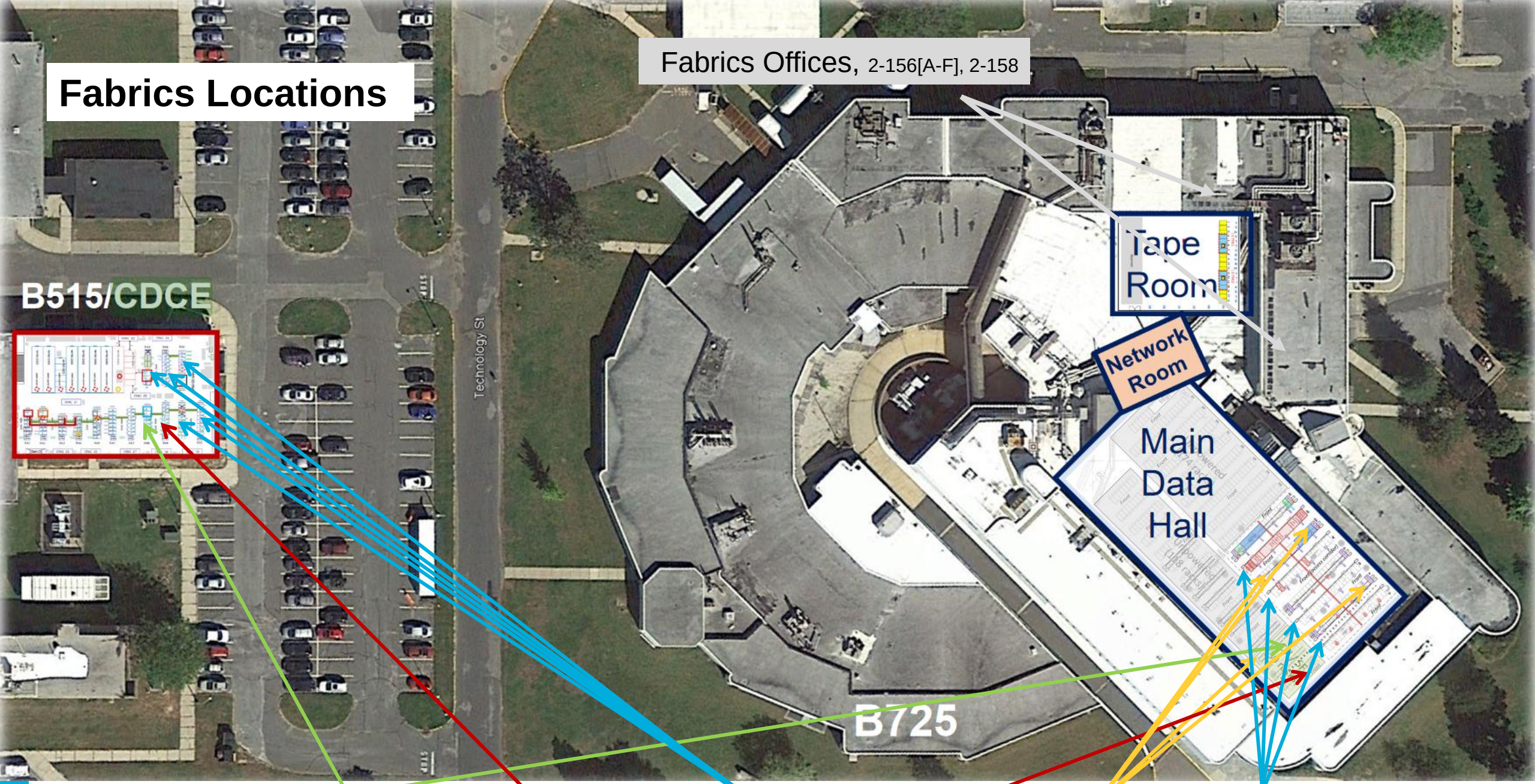
Shigeki Misawa: resource planning, cyber security, broad institutional knowledge, …

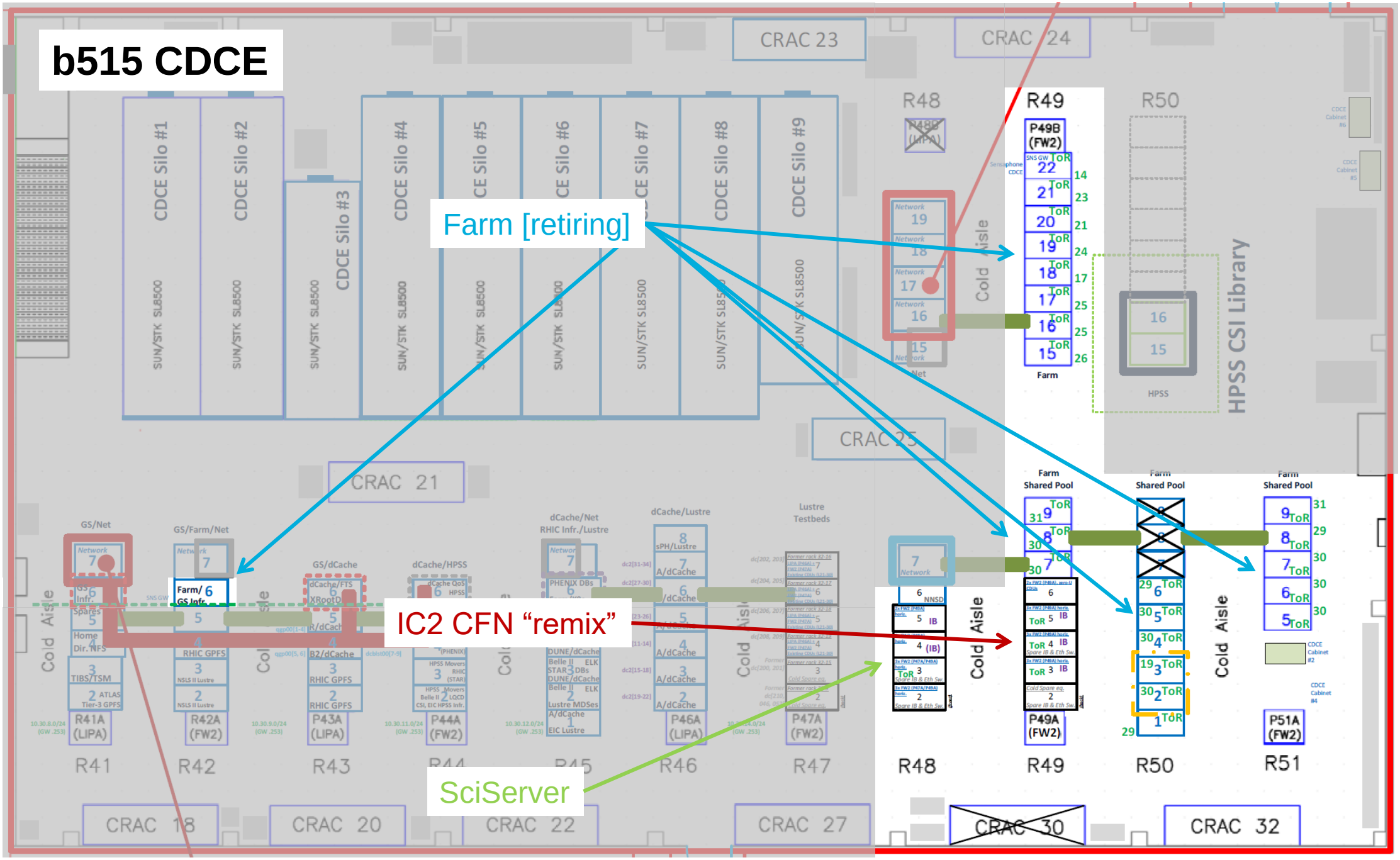
Tom Smith: condor, provisioning, config mgmt, …

Oszkar Tarjan: newest but not jr, extensive DevOps, SRE, AWS, etc from industry, …

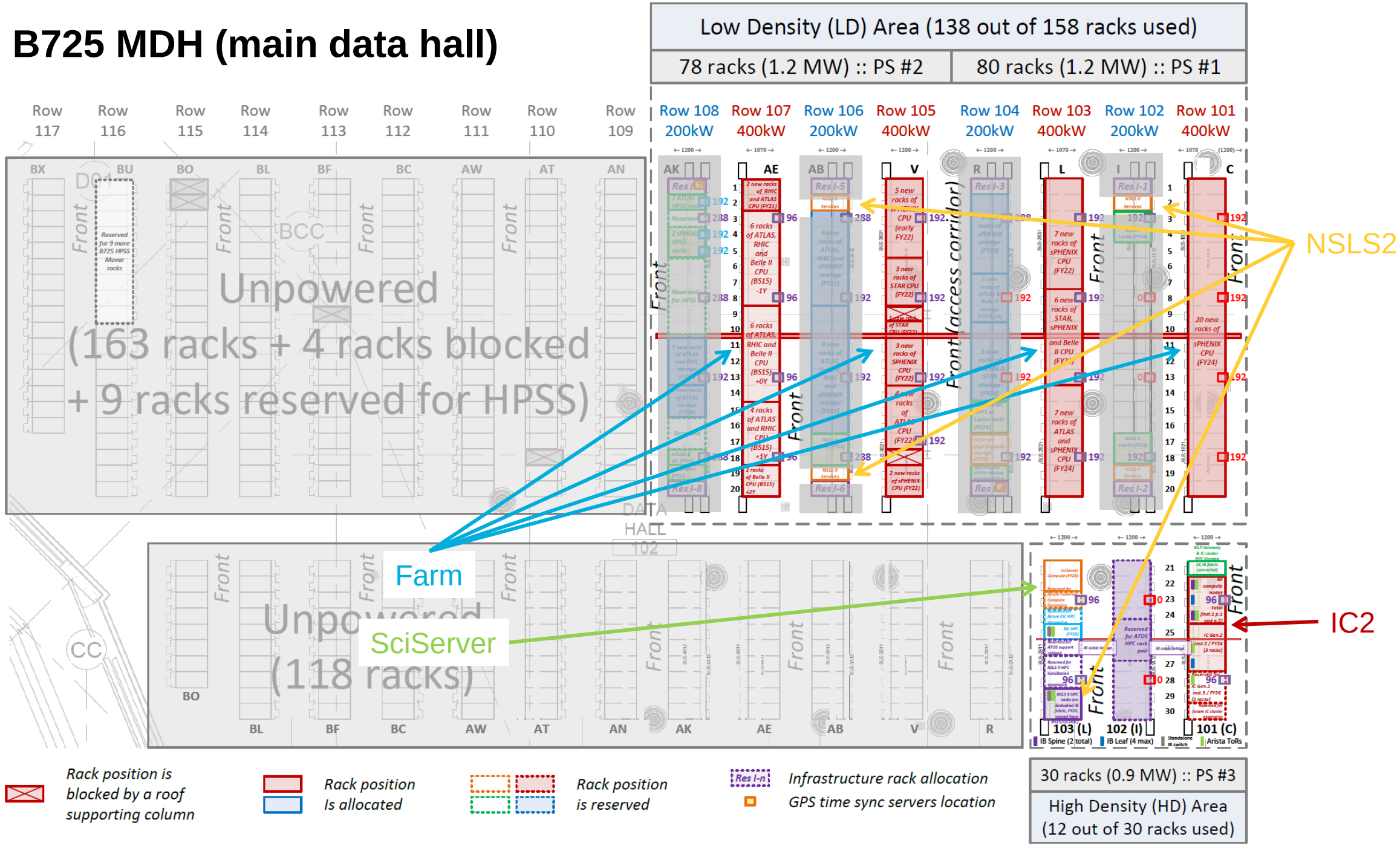
(these brief bullets severely underrepresent the breadth and depth of Fabrics' expertise)

Fabrics Locations




b515 CDCE
CRAC 23
CRAC 24
CRAC 25
CRAC 21
CRAC 18
CRAC 20
CRAC 22
CRAC 27
CRAC 30
CRAC 32
CDCE Silo #1
CDCE Silo #2
CDCE Silo #3
CDCE Silo #4
CDCE Silo #5
CDCE Silo #6
CDCE Silo #7
CDCE Silo #8
CDCE Silo #9
SUN/STK SL8500
R48
R49
R50
R51
P49B (FW2)
Farm
HPSS
HPSS CSI Library
Cold Aisle
Farm [retiring]
IC2 CFN "remix"
SciServer
GS/Net
GS/Farm/Net
GS/dCache
dCache/HPSS
dCache/Net
RHIC Infr./Lustre
dCache/Lustre
Lustre Testbeds
R41A (LIPA)
R42A (FW2)
P43A (LIPA)
P44A (FW2)
P46A (LIPA)
P47A (FW2)
P49A (FW2)
P51A (FW2)
R41
R42
R43
R44
R45
R46
R47
Farm Shared Pool

B725 MDH (main data hall)



Some ideas for the future…

- Ways to better serve smaller and/or sporadic compute needs
- Cross cluster submission/backfill/flocking
- slurm->condor and vice versa
- Cloud bursting (for more, or specialty, resources)
- Sharing hw between queueing system(s) (slurm, condor) and k8s (Kubernetes)
- Making some compute avail via more cloud like interface (OpenShift/k8s?)

What compute resources will you need? How can we help?

Questions?

Supplemental slides

CDS Mgmt Org Chart