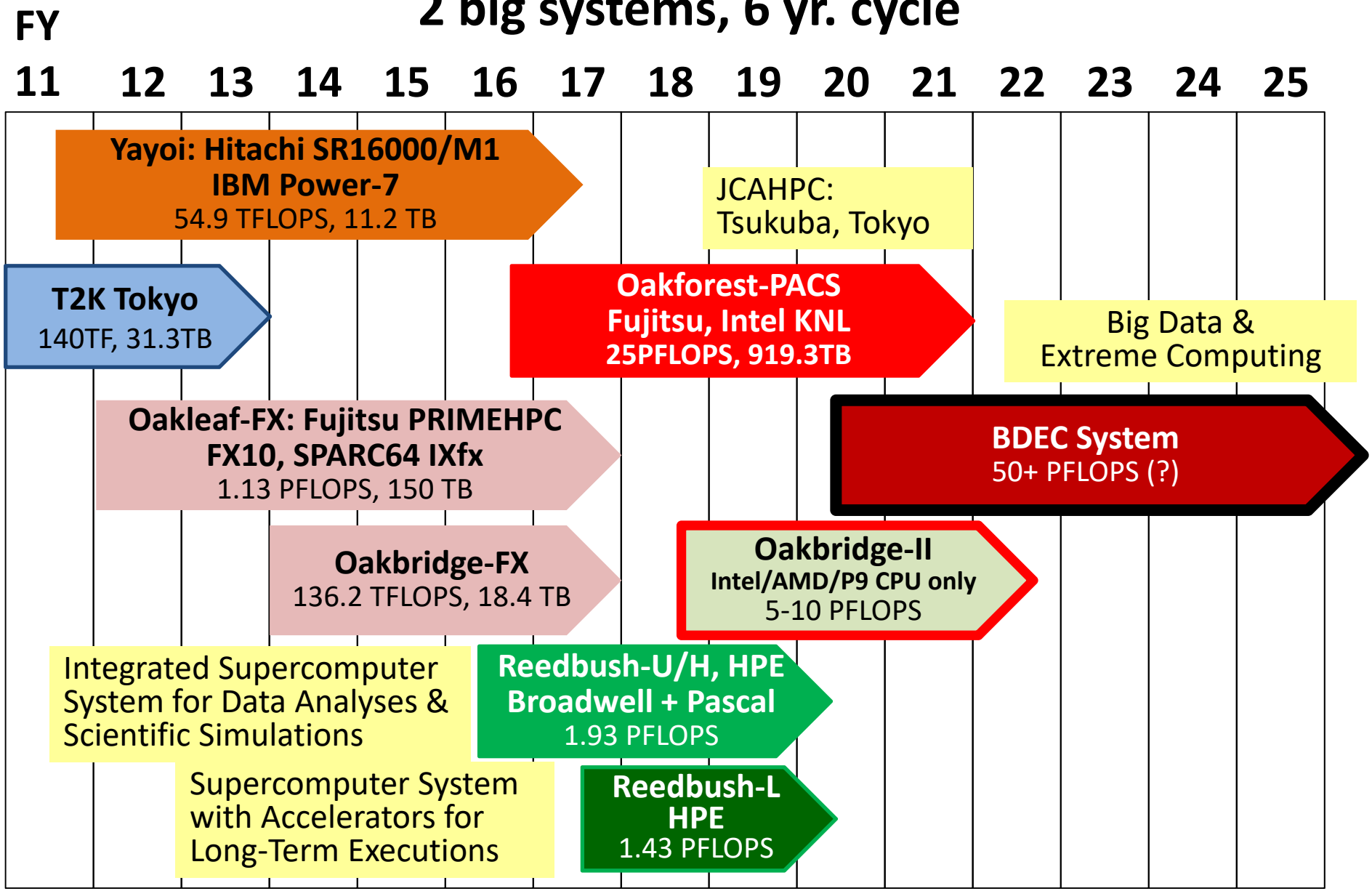
東京大学
THE UNIVERSITY OF TOKYO

Overview of Reedbush-U
How to Login

Information Technology Center
The University of Tokyo
http://www.cc.u-tokyo.ac.jp/

Supercomputers in ITC/U.Tokyo

2 big systems, 6 yr. cycle



Now operating 4 (or 6)systems !!

- Oakleaf-FX (Fujitsu PRIMEHPC FX10)
  - 1.135 PF, Commercial Version of K, Apr.2012 – Mar.2018

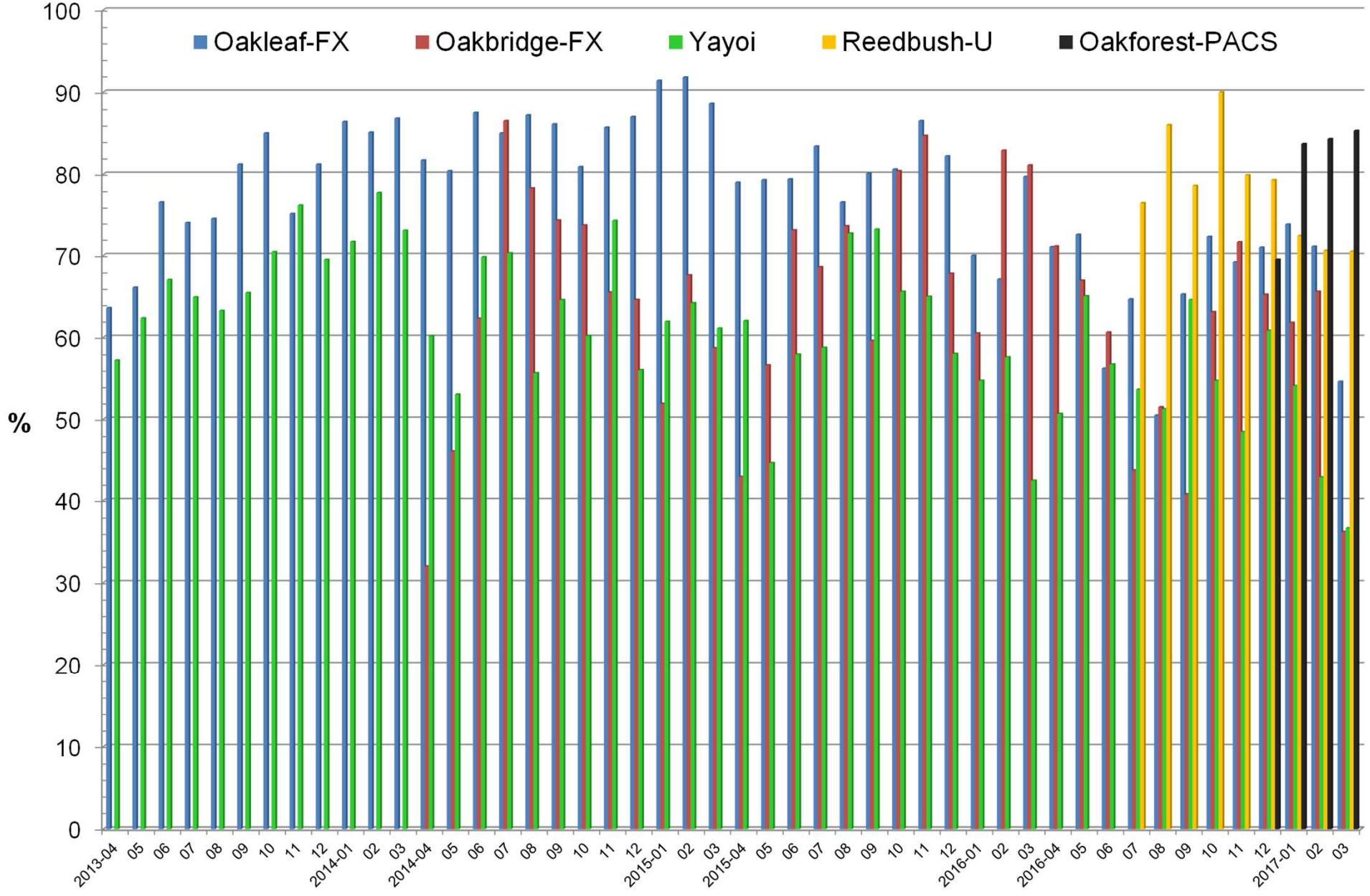
- Oakbridge-FX (Fujitsu PRIMEHPC FX10)
  - 136.2 TF, for long-time use (up to 168 hr), Apr.2014 – Mar.2018
- Reedbush (HPE, Intel BDW + NVIDIA P100 (Pascal))
  - Integrated Supercomputer System for Data Analyses & Scientific Simulations
    - Jul.2016-Jun.2020

  - **Our first GPU System**, DDN IME (Burst Buffer)
  - **Reedbush-U: CPU only, 420 nodes, 508 TF (Jul.2016)**
  - **Reedbush-H: 120 nodes, 2 GPUs/node: 1.42 PF (Mar.2017)**
  - **Reedbush-L: 64 nodes, 4 GPUs/node: 1.43 PF (Oct.2017)**
- Oakforest-PACS (OFP) (Fujitsu, Intel Xeon Phi (KNL))
  - JCAHPC (U.Tsukuba & U.Tokyo)
  - 25 PF, #9 in 50th TOP 500 (Nov. 2017) (#2 in Japan)
  - Omni-Path Architecture, DDN IME (Burst Buffer)

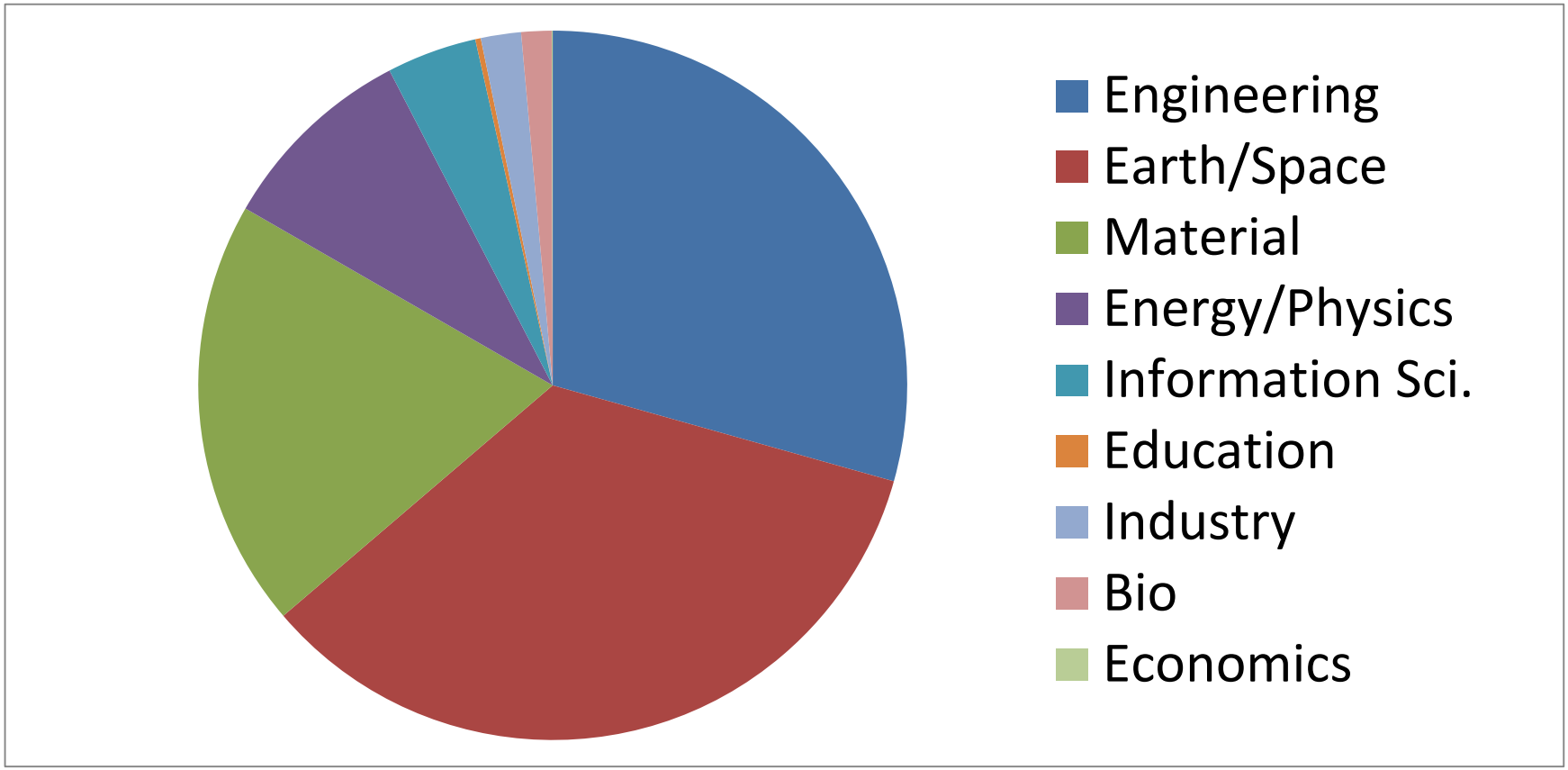
JPY (=Watt)/GFLOPS Rate
Smaller is better (efficient)

System	JPY/GFLOPS
Oakleaf/Oakbridge-FX (Fujitsu) (Fujitsu PRIMEHPC FX10)	125
Reedbush-U (SGI) (Intel BDW)	62.0
Reedbush-H (SGI) (Intel BDW+NVIDIA P100)	17.1
Oakforest-PACS (Fujitsu) (Intel Xeon Phi/Knights Landing)	16.5

Work Ratio

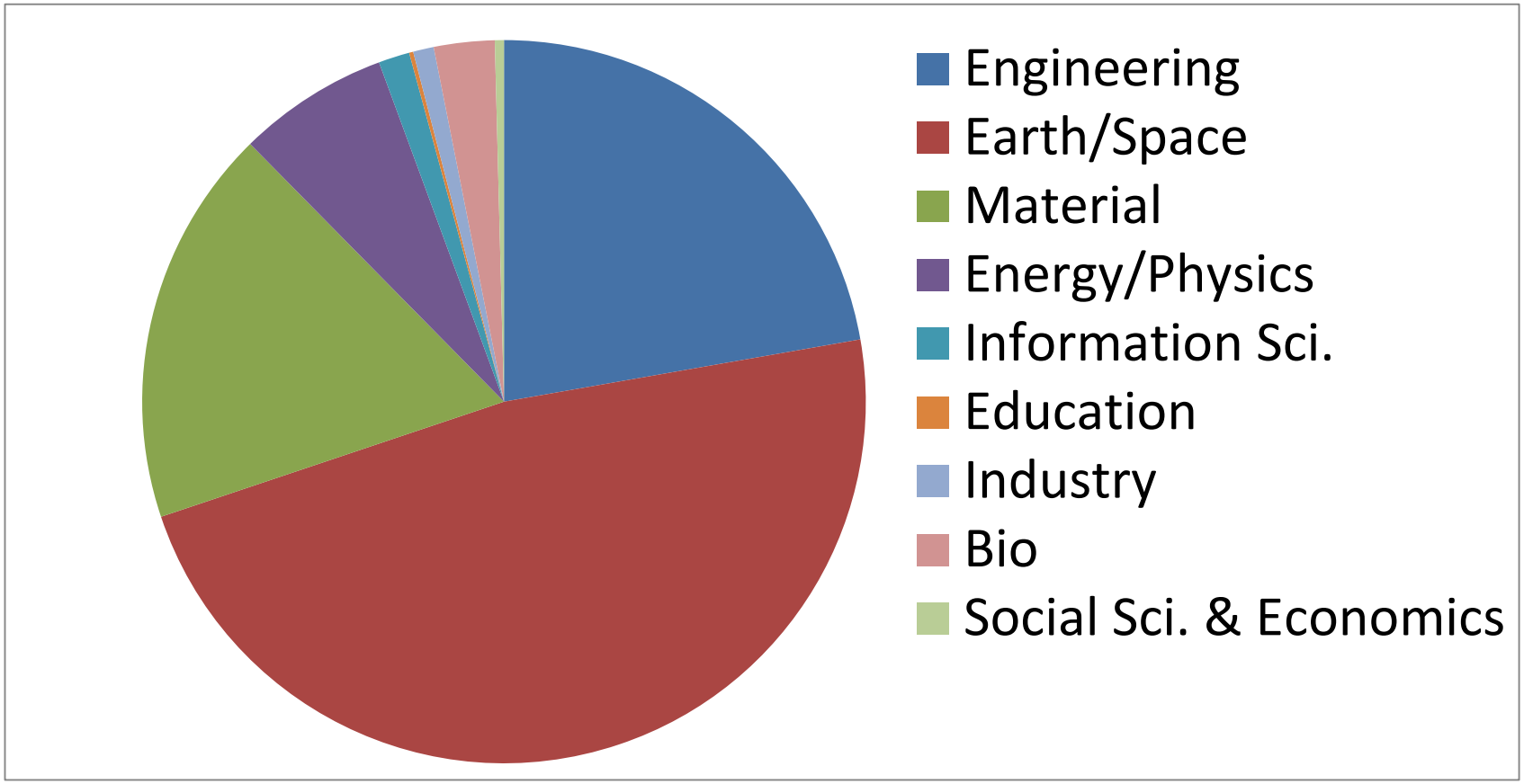

Oakleaf-FX
Oakbridge-FX
Yayoi
Reedbush-U
Oakforest-PACS
%
100
90
80
70
60
50
40
30
20
10
0
2013-04
05
06
07
08
09
10
11
12
2014-01
02
03
2014-04
05
06
07
08
09
10
11
12
2015-01
02
03
2015-04
05
06
07
08
09
10
11
12
2016-01
02
03
2016-04
05
06
07
08
09
10
11
12
2017-01
02
03

Research Area based on CPU Hours FX10 in FY.2015 (2015.4~2016.3E)



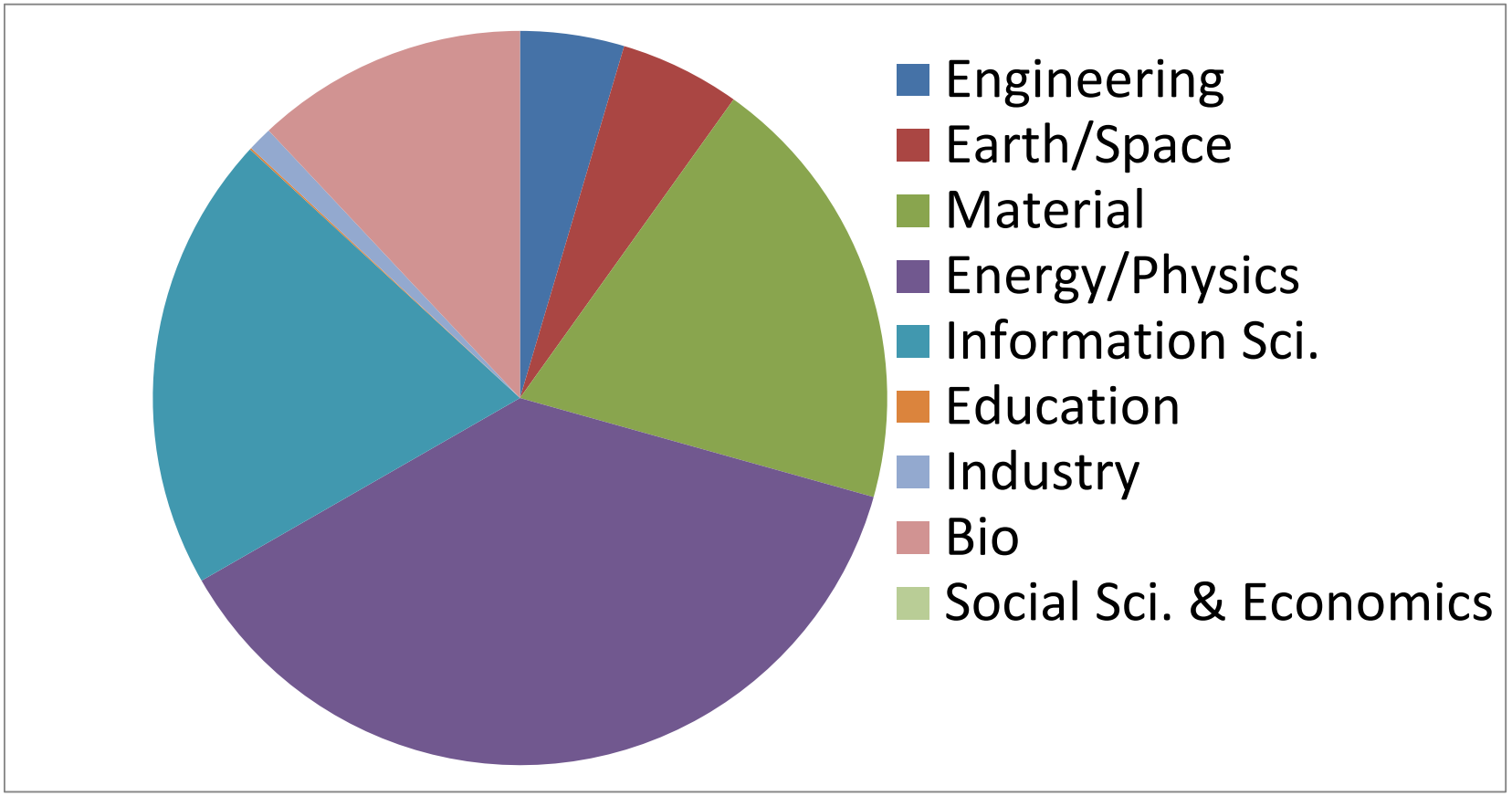
Oakleaf-FX + Oakbridge-FX

Research Area based on CPU Hours FX10 in FY.2016 (2016.4~2017.3E)



Oakleaf-FX + Oakbridge-FX

Research Area based on CPU Hours Reedbush-U in FY.2016 (2016.7~2017.3E)



Oakforest-PACS (OFP)

- Full Operation started on December 1, 2016
- 8,208 Intel Xeon/Phi (KNL), 25 PF Peak Performance
  - Fujitsu
- **TOP 500 #9 (#2 in Japan), HPCG #6 (#2) (Nov 2017)**
- **JCAHPC: Joint Center for Advanced High Performance Computing)**
  - University of Tsukuba
  - University of Tokyo
    - New system will installed in Kashiwa-no-Ha (Leaf of Oak) Campus/U.Tokyo, which is between Tokyo and Tsukuba
  - http://jcahpc.jp

JCAHPC

東京大学 THE UNIVERSITY OF TOKYO

筑波大学 University of Tsukuba

Benchmarks

- TOP 500 (Linpack, HPL(High Performance Linpack))
  - Direct Linear Solvers, FLOPS rate
  - Regular Dense Matrices, Continuous Memory Access
  - Computing Performance
- HPCG
  - Preconditioned Iterative Solvers, FLOPS rate
  - Irregular Sparse Matrices derived from FEM Applications with Many “0” Components
    - Irregular/Random Memory Access,
    - Closer to “Real” Applications than HPL
  - Performance of Memory, Communications
- Green 500
  - FLOPS/W rate for HPL (TOP500)

50th TOP500 List (November, 2017)

	Site	Computer/Year Vendor	Cores	R_{max} (TFLOPS)	R_{peak} (TFLOPS)	Power (kW)
1	National Supercomputing Center in Wuxi, China	**Sunway TaihuLight** , Sunway MPP, Sunway SW26010 260C 1.45GHz, 2016 NRCPC	10,649,600	93,015 (= 93.0 PF)	125,436	15,371
2	National Supercomputing Center in Tianjin, China	**Tianhe-2**, Intel Xeon E5-2692, TH Express-2, Xeon Phi, 2013 NUDT	3,120,000	33,863 (= 33.9 PF)	54,902	17,808
3	Swiss Natl. Supercomputer Center, Switzerland	**Piz Daint** Cray XC30/NVIDIA P100, 2013 Cray	361,760	19,590	33,863	2,272
4	JAMSTEC, JAPAN	**Gyoukou,** ZettaScaler-2.2 HPC, Xeon D-1571, 2017, ExaScaler	19,860	19,136	28,192	1,350
5	Oak Ridge National Laboratory, USA	**Titan** Cray XK7/NVIDIA K20x, 2012 Cray	560,640	17,590	27,113	8,209
6	Lawrence Livermore National Laboratory, USA	**Sequoia** BlueGene/Q, 2011 IBM	1,572,864	17,173	20,133	7,890
7	DOE/NNSA/LANL/SNL	**Trinity,** Cray XC40 Intel Xeon Phi 7250 68C 1.4GHz, Cray Aries, 2017, Cray	979,968	14,137	43,903	3.844
8	DOE/SC/LBNL/NERSC USA	**Cori**, Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Cray Aries, 2016 Cray	632,400	14,015	27,881	3,939
9	Joint Center for Advanced High Performance Computing, Japan	**Oakforest-PACS**, PRIMERGY CX600 M1, Intel Xeon Phi Processor 7250 68C 1.4GHz, Intel Omni-Path, 2016 Fujitsu	557,056	13,555	24,914	2,719
10	RIKEN AICS, Japan	**K computer**, SPARC64 VIIIfx , 2011 Fujitsu	705,024	10,510	11,280	12,660

R_{max}: Performance of Linpack (TFLOPS)
R_{peak}: Peak Performance (TFLOPS), Power: kW

http://www.top500.org/

HPCG Ranking (November, 2017)

	Computer	Cores	HPL Rmax (Pflop/s)	TOP500 Rank	HPCG (Pflop/s)	Peak
1	**K computer**	705,024	10.510	10	0.603	5.3%
2	**Tianhe-2 (MilkyWay-2)**	3,120,000	33.863	2	0.580	1.1%
3	**Trinity**	979,072	93.015	7	0.546	0.4%
4	**Piz Daint**	361,760	19.590	3	0.486	1.9%
5	**Sunway TaihuLight**	10,649,600	93.015	1	0.481	0.4%
6	**Oakforest-PACS**	557,056	13.555	9	0.386	1.5%
7	**Cori**	632,400	13.832	8	0.355	1.3%
8	**Sequoia**	1,572,864	17.173	6	0.330	1.6%
9	**Titan**	560,640	17.590	4	0.322	1.2%
10	**Tsbubame 3**	136,080	8.125	13	0.189	1.6%

http://www.hpcg-benchmark.org/

Green 500 Ranking (November, 2016)

	Site	Computer	CPU	HPL Rmax (Pflop/s)	TOP500 Rank	Power (MW)	GFLOPS/W
1	NVIDIA Corporation	DGX SATURNV	NVIDIA DGX-1, Xeon E5-2698v4 20C 2.2GHz, Infiniband EDR, NVIDIA Tesla P100	3.307	28	0.350	9.462
2	Swiss National Supercomputing Centre (CSCS)	Piz Daint	Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect , NVIDIA Tesla P100	9.779	8	1.312	7.454
3	RIKEN ACCS	Shoubu	ZettaScaler-1.6 etc.	1.001	116	0.150	6.674
4	National SC Center in Wuxi	Sunway TaihuLight	Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway	93.01	1	15.37	6.051
5	SFB/TR55 at Fujitsu Tech. Solutions GmbH	QPACE3	PRIMERGY CX1640 M1, Intel Xeon Phi 7210 64C 1.3GHz, Intel Omni-Path	0.447	375	0.077	5.806
6	JCAHPC	Oakforest-PACS	PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path	1.355	6	2.719	4.986
7	DOE/SC/Argonne National Lab.	Theta	Cray XC40, Intel Xeon Phi 7230 64C 1.3GHz, Aries interconnect	5.096	18	1.087	4.688
8	Stanford Research Computing Center	XStream	Cray CS-Storm, Intel Xeon E5-2680v2 10C 2.8GHz, Infiniband FDR, Nvidia K80	0.781	162	0.190	4.112
9	ACCMS, Kyoto University	Camphor 2	Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect	3.057	33	0.748	4.087
10	Jefferson Natl. Accel. Facility	SciPhi XVI	KOI Cluster, Intel Xeon Phi 7230 64C 1.3GHz, Intel Omni-Path	0.426	397	0.111	3.837

http://www.top500.org/

Green 500 Ranking (June, 2017)

	Site	Computer	CPU	HPL Rmax (Pflop/s)	TOP500 Rank	Power (MW)	GFLOPS/W
1	Tokyo Tech.	TSUBAME3.0	SGI ICE XA, IP139-SXM2, Xeon E5-2680v4, NVIDIA Tesla P100 SXM2, HPE	1,998.0	61	142	14.110
2	Yahoo Japan	kukai	ZettaScaler-1.6, Xeon E5-2650Lv4,, NVIDIA Tesla P100 , Exascalar	460.7	465	33	14.046
3	AIST, Japan	AIST AI Cloud	NEC 4U-8GPU Server, Xeon E5-2630Lv4, NVIDIA Tesla P100 SXM2 , NEC	961.0	148	76	12.681
4	CAIP, RIKEN, JAPAN	RAIDEN GPU subsystem -	NVIDIA DGX-1, Xeon E5-2698v4, NVIDIA Tesla P100 , Fujitsu	635.1	305	60	10.603
5	Univ. Cambridge, UK	Wilkes-2 -	Dell C4130, Xeon E5-2650v4, NVIDIA Tesla P100 , Dell	1,193.0	100	114	10.428
6	Swiss Natl. SC. Center (CSCS)	Piz Daint	Cray XC50, Xeon E5-2690v3, NVIDIA Tesla P100 , Cray Inc.	19,590.0	3	2,272	10.398
7	JAMSTEC, Japan	Gyoukou,	ZettaScaler-2.0 HPC system, Xeon D-1571, PEZY-SC2 , ExaScalar	1,677.1	69	164	10.226
8	Inst. for Env. Studies, Japan	GOSAT-2 (RCF2)	SGI Rackable C1104-GP1, Xeon E5-2650v4, NVIDIA Tesla P100 , NSSOL/HPE	770.4	220	79	9.797
9	Facebook, USA	Penguin Relion	Xeon E5-2698v4/E5-2650v4, NVIDIA Tesla P100 , Acer Group	3,307.0	31	350	9.462
10	NVIDIA, USA	DGX Saturn V	Xeon E5-2698v4, NVIDIA Tesla P100 , Nvidia	3,307.0	32	350	9.462
11	ITC, U.Tokyo, Japan	Reedbush-H	SGI Rackable C1102-GP8, Xeon E5-2695v4, NVIDIA Tesla P100 SXM2 , HPE	802.4	203	94	8.575

http://www.top500.org/

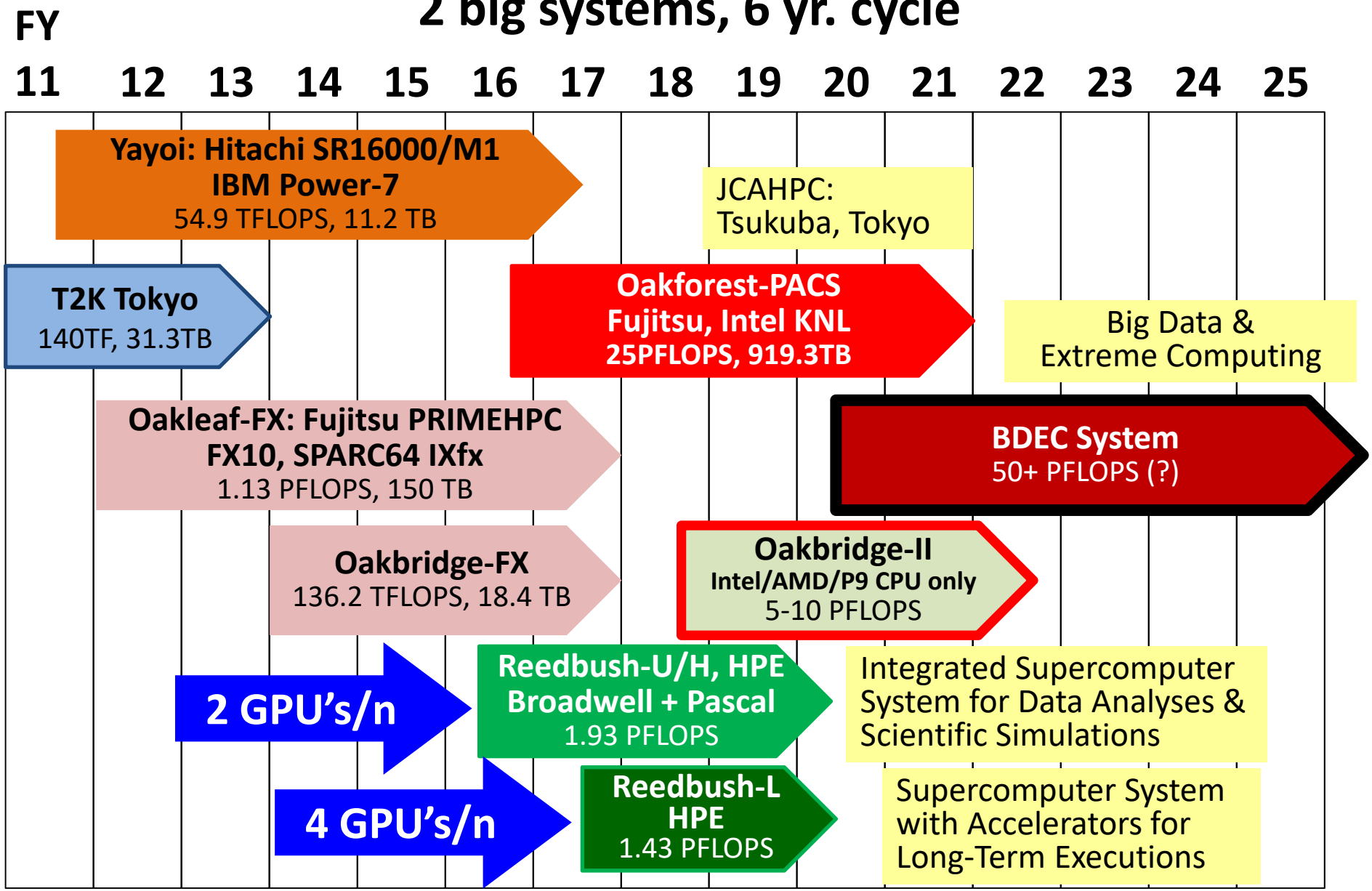
Green 500 Ranking (Nov., 2017)

	Site	Computer	CPU	HPL Rmax (Pflop/s)	TOP500 Rank	Power (kW)	GFLOPS/W
1	RIKEN, Japan	Shoubu system B	ZettaScaler-2.2 HPC system, Xeon D-1571, PEZY-SC2 , ExaScalar	842.0	259	50	17.009
2	KEK, Japan	Suiren2	ZettaScaler-2.2 HPC system, Xeon D-1571, PEZY-SC2 , ExaScalar	788.2	307	47	16.759
3	PEZY, Japan	Sakura	ZettaScaler-2.2 HPC system, Xeon E5-2618Lv3, PEZY-SC2 , ExaScalar	824.7	276	50	16.657
4	NVIDIA, USA	DGX Saturn V Volta	Xeon E5-2698v4, NVIDIA Tesla V100 , Nvidia	1,070.0	149	97	15.113
5	JAMSTEC, Japan	Gyoukou	ZettaScaler-2.2 HPC system, Xeon D-1571, PEZY-SC2 , ExaScalar	19,135.8	4	1,350	14.173
6	Tokyo Tech.	TSUBAME3.0	SGI ICE XA, IP139-SXM2, Xeon E5-2680v4, NVIDIA Tesla P100 SXM2, HPE	8,125.0	13	792	13.704
7	AIST, Japan	AIST AI Cloud	NEC 4U-8GPU Server, Xeon E5-2630Lv4, NVIDIA Tesla P100 SXM2 , NEC	961.0	148	76	12.681
8	CAIP, RIKEN, JAPAN	RAIDEN GPU subsystem -	NVIDIA DGX-1, Xeon E5-2698v4, NVIDIA Tesla P100 , Fujitsu	635.1	305	60	10.603
9	Univ. Cambridge, UK	Wilkes-2 -	Dell C4130, Xeon E5-2650v4, NVIDIA Tesla P100 , Dell	1,193.0	100	114	10.428
10	Swiss Natl. SC. Center (CSCS)	Piz Daint	Cray XC50, Xeon E5-2690v3, NVIDIA Tesla P100 , Cray Inc.	19,590.0	3	2,272	10.398
11	ITC, U.Tokyo, Japan	Reedbush-L	SGI Rackable C1102-GP8, Xeon E5-2695v4, NVIDIA Tesla P100 SXM2 , HPE	805.6	291	79	10.167

http://www.top500.org/

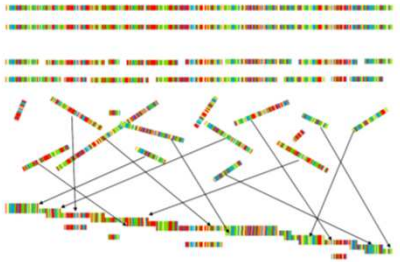
Supercomputers in ITC/U.Tokyo

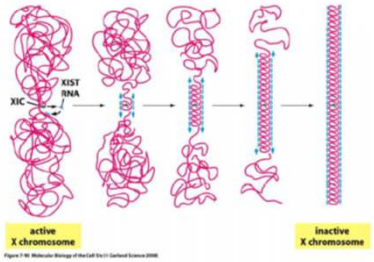
2 big systems, 6 yr. cycle



Reedbush: Our First System with GPU's

- Before 2015
  - CUDA
  - We have 2,000+ users
- Reasons of Changing Policy
  - Recent Improvement of OpenACC
    - Similar Interface as OpenMP
    - Research Collaboration with NVIDIA Engineers
  - Data Science, Deep Learning
    - New types of users other than traditional CSE (Computational Science & Engineering) are needed
      - Research Organization for Genome Medical Science, U. Tokyo
      - U. Tokyo Hospital: Processing of Medical Images by Deep Learning





Reedbush-U/H (1/2)

Integrated Supercomputer System for Data Analyses & Scientific Simulations

- SGI was awarded (Mar. 22, 2016)
- Compute Nodes (CPU only): Reedbush-U
  - Intel Xeon E5-2695v4 (Broadwell-EP, 2.1GHz 18core) x 2socket (1.210 TF), 256 GiB (153.6GB/sec)
  - InfiniBand EDR, Full bisection Fat-tree
  - Total System: 420 nodes, 508.0 TF
- Compute Nodes (with Accelerators): Reedbush-H
  - Intel Xeon E5-2695v4 (Broadwell-EP, 2.1GHz 18core) x 2socket, 256 GiB (153.6GB/sec)
  - NVIDIA Pascal GPU (Tesla P100)
    - (5.3TF, 720GB/sec, 16GiB) x 2 / node
  - InfiniBand FDR x 2ch (for ea. GPU), Full bisection Fat-tree
  - 120 nodes, 145.2 TF(CPU)+ 1.27 PF(GPU)= 1.42 PF

Why “Reedbush” ?

Blaise Pascal
(1623-1662)

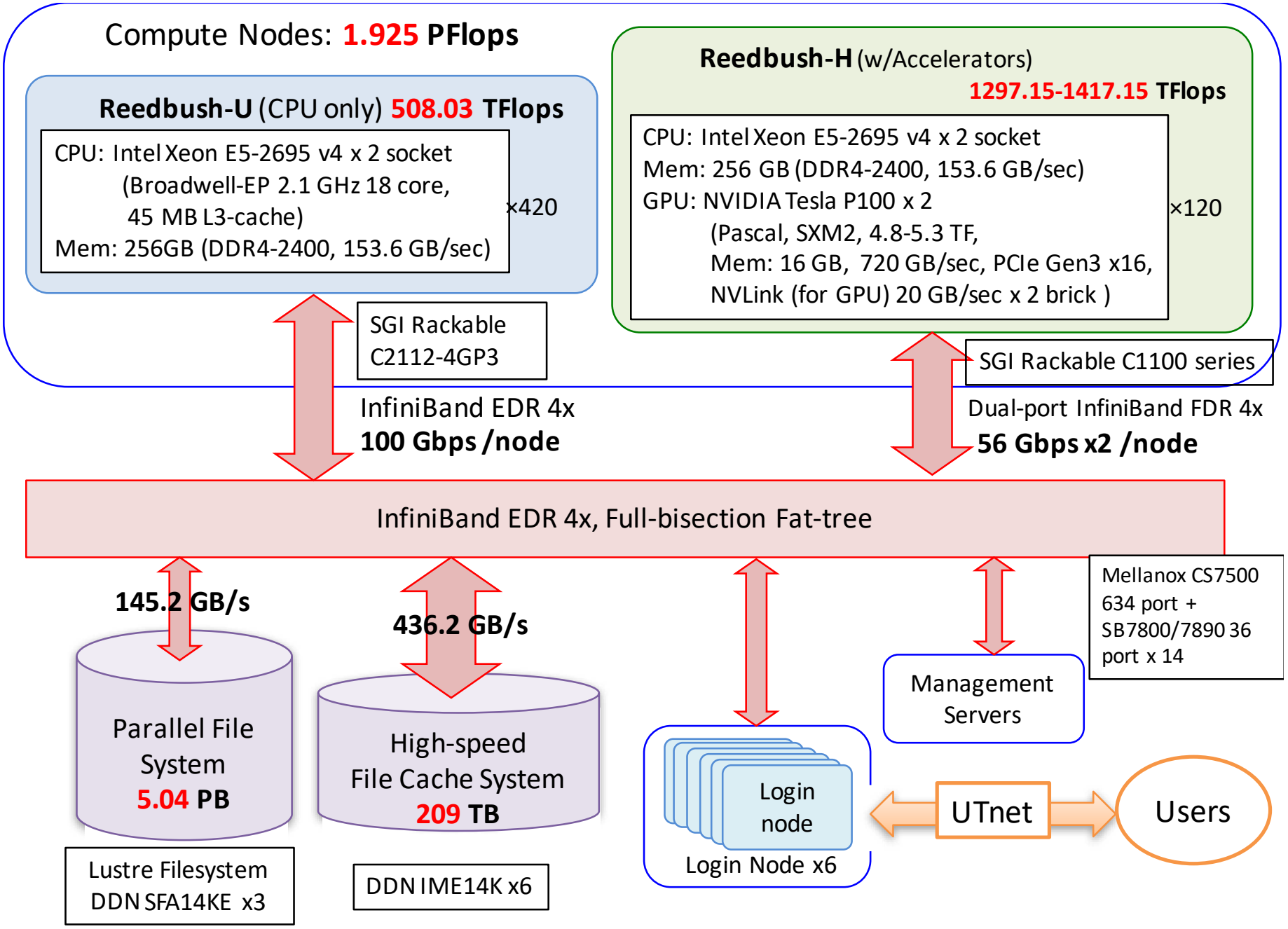
- L'homme est un roseau pensant.
- Man is a thinking reed.
- 人間は考える葦である

Pensées (Blaise Pascal)

Reedbush-U/H (2/2)

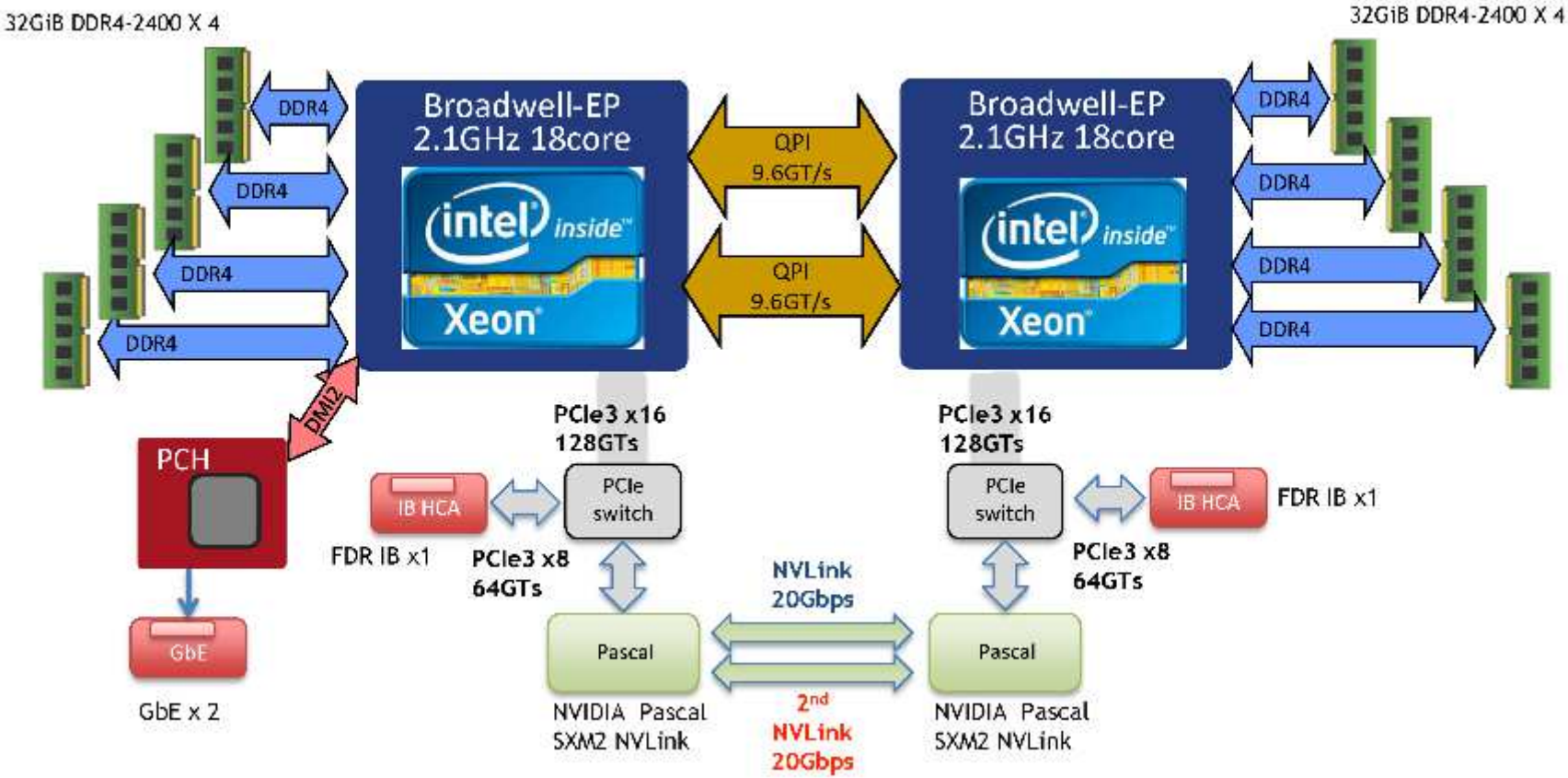
Integrated Supercomputer System for Data Analyses & Scientific Simulations

- Storage/File Systems
  - Shared Parallel File-system (Lustre)
    - 5.04 PB, 145.2 GB/sec
  - Fast File Cache System: Burst Buffer (DDN IME (Infinite Memory Engine))
    - SSD: 209.5 TB, 450 GB/sec
- Power, Cooling, Space
  - Air cooling only, < 500 kVA (without A/C): 378 kVA
  - < 90 m^2
- Software & Toolkit for Data Analysis, Deep Learning …
  - OpenCV, Theano, Anaconda, ROOT, TensorFlow
  - Torch, Caffe, Cheiner, GEANT4

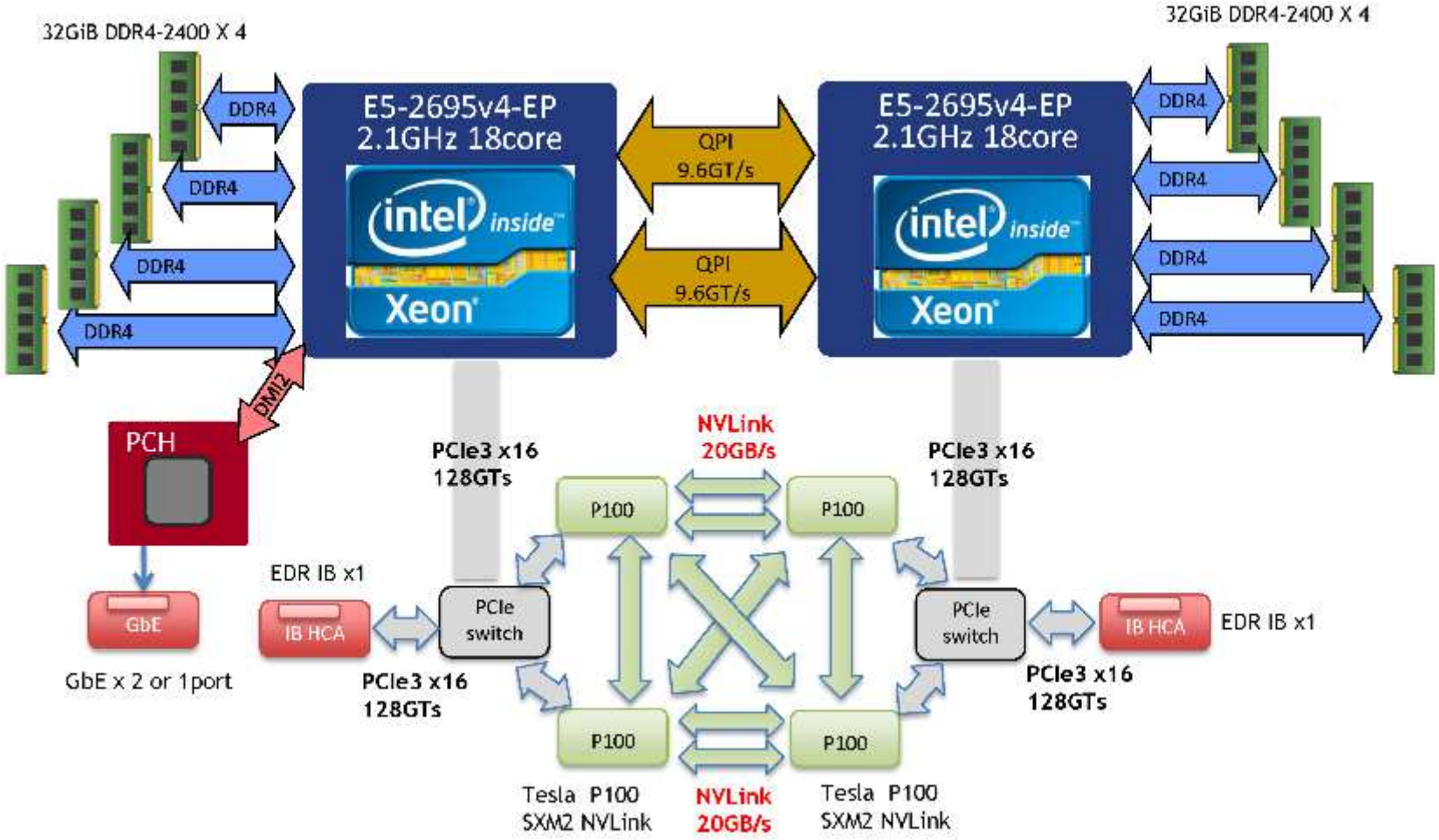

Compute Nodes: 1.925 PFlops
Reedbush-U (CPU only) 508.03 TFlops
CPU: Intel Xeon E5-2695 v4 x 2 socket
(Broadwell-EP 2.1 GHz 18 core,
45 MB L3-cache)
Mem: 256GB (DDR4-2400, 153.6 GB/sec)
×420
Reedbush-H (w/Accelerators)
1297.15-1417.15 TFlops
CPU: Intel Xeon E5-2695 v4 x 2 socket
Mem: 256 GB (DDR4-2400, 153.6 GB/sec)
GPU: NVIDIA Tesla P100 x 2
(Pascal, SXM2, 4.8-5.3 TF,
Mem: 16 GB, 720 GB/sec, PCIe Gen3 x16,
NVLink (for GPU) 20 GB/sec x 2 brick)
×120
SGI Rackable C2112-4GP3
SGI Rackable C1100 series
InfiniBand EDR 4x
100 Gbps /node
Dual-port InfiniBand FDR 4x
56 Gbps x2 /node
InfiniBand EDR 4x, Full-bisection Fat-tree
145.2 GB/s
436.2 GB/s
Mellanox CS7500
634 port +
SB7800/7890 36
port x 14
Parallel File
System
5.04 PB
High-speed
File Cache System
209 TB
Management
Servers
Login
node
Login Node x6
UTnet
Users
Lustre Filesystem
DDN SFA14KE x3
DDN IME14K x6

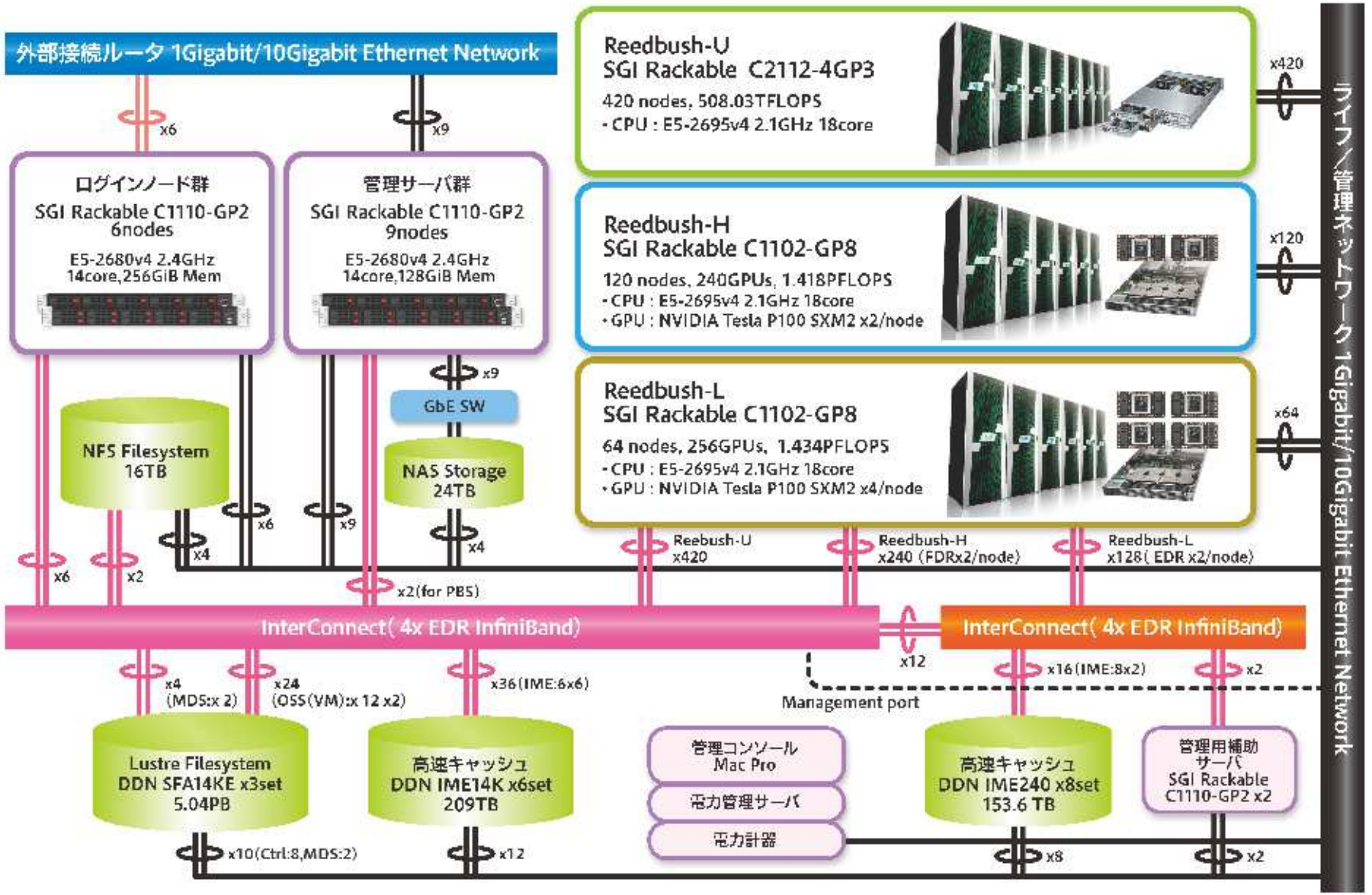
	Reedbush-U	Reedbush-H	Reedbush-L
	Integrated Supercomputer System for Data Analyses & Scientific Simulations		Supercomputer System with Accelerators for Long-Term Executions
CPU/node	Intel Xeon E5-2695v4 (Broadwell-EP, 2.1GHz, 18core) x 2 sockets (1.210 TF), 256 GiB (153.6GB/sec)		
GPU	-	NVIDIA Tesla P100 (Pascal, 5.3TF, 720GB/sec, 16GiB)	
Infiniband	EDR	FDR × 2ch	EDR × 2ch
Nodes #	420	120	64
GPU #	-	240 (=120 × 2)	256 (=64 × 4)
Peak Performance (TFLOPS)	509	1,417 (145 + 1,272)	1,433 (76.8 + 1,358)
Total Memory Bandwidth (TB/sec)	64.5	191.2 (18.4+172.8)	194.2 (9.83+184.3)
since	2016.07	2017.03	2017.10

Compute Node of Reedbush-H



Compute Node of Reedbush-L




外部接続ルータ 1Gigabit/10Gigabit Ethernet Network
x6
x9
ログインノード群
SGI Rackable C1110-GP2
6nodes
E5-2680v4 2.4GHz
14core,256GiB Mem
管理サーバ群
SGI Rackable C1110-GP2
9nodes
E5-2680v4 2.4GHz
14core,128GiB Mem
Reedbush-U
SGI Rackable C2112-4GP3
420 nodes, 508.03TFLOPS
・CPU : E5-2695v4 2.1GHz 18core
x420
Reedbush-H
SGI Rackable C1102-GP8
120 nodes, 240GPUs, 1.418PFLOPS
・CPU : E5-2695v4 2.1GHz 18core
・GPU : NVIDIA Tesla P100 SXM2 x2/node
x120
Reedbush-L
SGI Rackable C1102-GP8
64 nodes, 256GPUs, 1.434PFLOPS
・CPU : E5-2695v4 2.1GHz 18core
・GPU : NVIDIA Tesla P100 SXM2 x4/node
x64
ログイン/管理ネットワーク 1Gigabit/10Gigabit Ethernet Network
GbE SW
NFS Filesystem
16TB
NAS Storage
24TB
x4
x6
x9
x2
x2(for PBS)
Reebush-U
x420
Reedbush-H
x240 (FDRx2/node)
Reedbush-L
x128(EDR x2/node)
InterConnect(4x EDR InfiniBand)
InterConnect(4x EDR InfiniBand)
x12
x16(IME:8x2)
Management port
x4
(MDS:x 2)
x24
(OSS(VM):x 12 x2)
x36(IME:6x6)
Lustre Filesystem
DDN SFA14KE x3set
5.04PB
高速キャッシュ
DDN IME14K x6set
209TB
管理コンソール
Mac Pro
電力管理サーバ
電力計器
高速キャッシュ
DDN IME240 x8set
153.6 TB
管理用補助
サーバ
SGI Rackable
C1110-GP2 x2
x10(Ctrl:8,MDS:2)
x12
x8
x2

How to Login (1/3)

- Public Key Certificate
  - Public Key Certificate
  - Password provided by ITC with 8 characters is not used for “login”

How to Login (2/3)

- Password with 8 characters by ITC
  - for registration of keys
  - browsing manuals
    - Only users can access manuals
    - SSH Port Forwarding is possible by keys

How to Login (3/3)

- Procedures
  - Creating Keys
  - Registration of Public Key
  - Login

Creating Keys on Unix (1/2)

- OpenSSH for UNIX/Mac/Cygwin

- Command for creating keys
  $ ssh-keygen –t rsa

- RETURN
- Passphrase
- Passphrase again

Creating Keys on Unix (2/2)

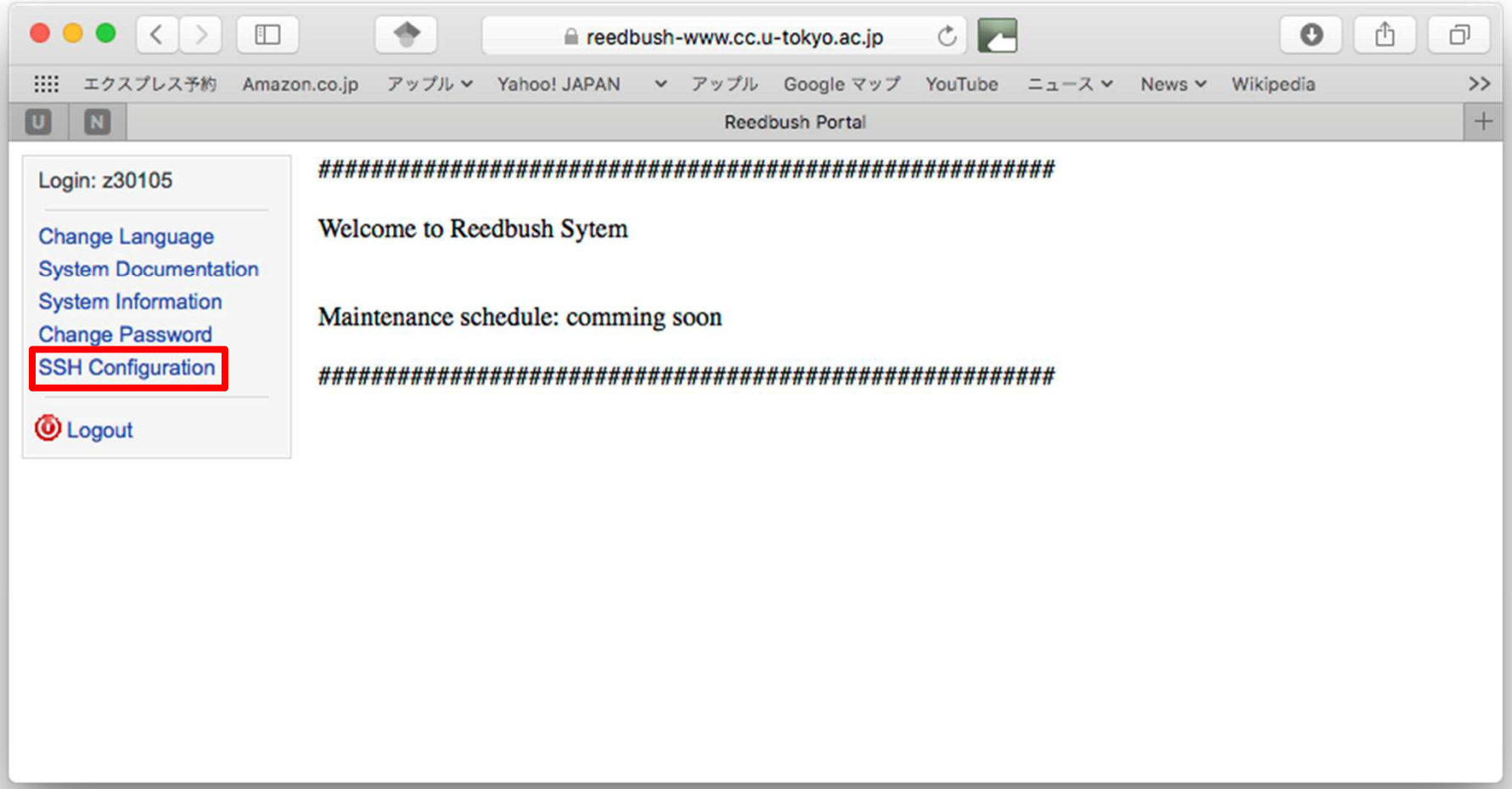

```
>$ ssh-keygen -t rsa
Generating public/private rsa key pair.
Enter file in which to save the key (/home/guestx/.ssh/id_rsa):
Enter passphrase (empty for no passphrase): (your favorite passphrase)
Enter same passphrase again:
Your identification has been saved in /home/guestx/.ssh/id_rsa.
Your public key has been saved in /home/guestx/.ssh/id_rsa.pub.
The key fingerprint is:

>$ cd ~/.ssh
>$ ls -l
total 12
-rw-------    1 guestx   guestx       1743 Aug 23 15:14 id_rsa
-rw-r--r--    1 guestx   guestx        413 Aug 23 15:14 id_rsa.pub

>$ cat id_rsa.pub

(cut & paste)
```

Registration of Public Key

- https://reedbush-www.cc.u-tokyo.ac.jp/

- User ID
- Password (8 characters)

- "SSH Configuration"
- Cut & Paste the Public Key


reedbush-www.cc.u-tokyo.ac.jp
エクスプレス予約
Amazon.co.jp
アップル
Yahoo! JAPAN
アップル
Google マップ
YouTube
ニュース
News
Wikipedia
Reedbush Portal
Login: z30105
Change Language
System Documentation
System Information
Change Password
SSH Configuration
Logout
##########
Welcome to Reedbush Sytem
Maintenance schedule: comming soon
##########

Login

- Login

  $ ssh reedbush-u.cc.u-tokyo.ac.jp -l t310XX (or)

  $ ssh t310XX@reedbush-u.cc.u-tokyo.ac.jp

- Directory

  $ /home/gt31/t310XX login -> small

  - Type “cd” for going back to /home/gt31/t310XX

  $ cd /lustre/gt31/t310XX please use this directory

  - Type “cdw” for going to /lustre/gt31/t310XX

- Copying Files

  $ scp <file> t310**@reedbush.cc.u-tokyo.ac.jp:~/.

  $ scp -r <dir> t310**@reedbush.cc.u-tokyo.ac.jp:~/.

- Public/Private Keys are used

  - “Passphrase”, not “Password”

Please check schedule of maintenance

- Last Friday of each month
  - other non-regular shutdown

- http://www.cc.u-tokyo.ac.jp/
- http://www.cc.u-tokyo.ac.jp/system/reedbush/

If you have any questions, please contact KN (Kengo Nakajima)

Do not contact ITC support directly.