The ATLAS experiment

and

Campus-Based LHC Physics Analysis

Spring 2009 Internet2 Member Meeting

ARLINGTON, VIRGINIA April 27–29, 2009

Raymond Brock
Department of Physics and Astronomy
Michigan State University

TOC

1 a little science

my field

High Energy Physics

aka

Elementary Particle Physics

Suitable for any occasion

A bundle of energy

will condense into distinct kinds of globs: “particles”

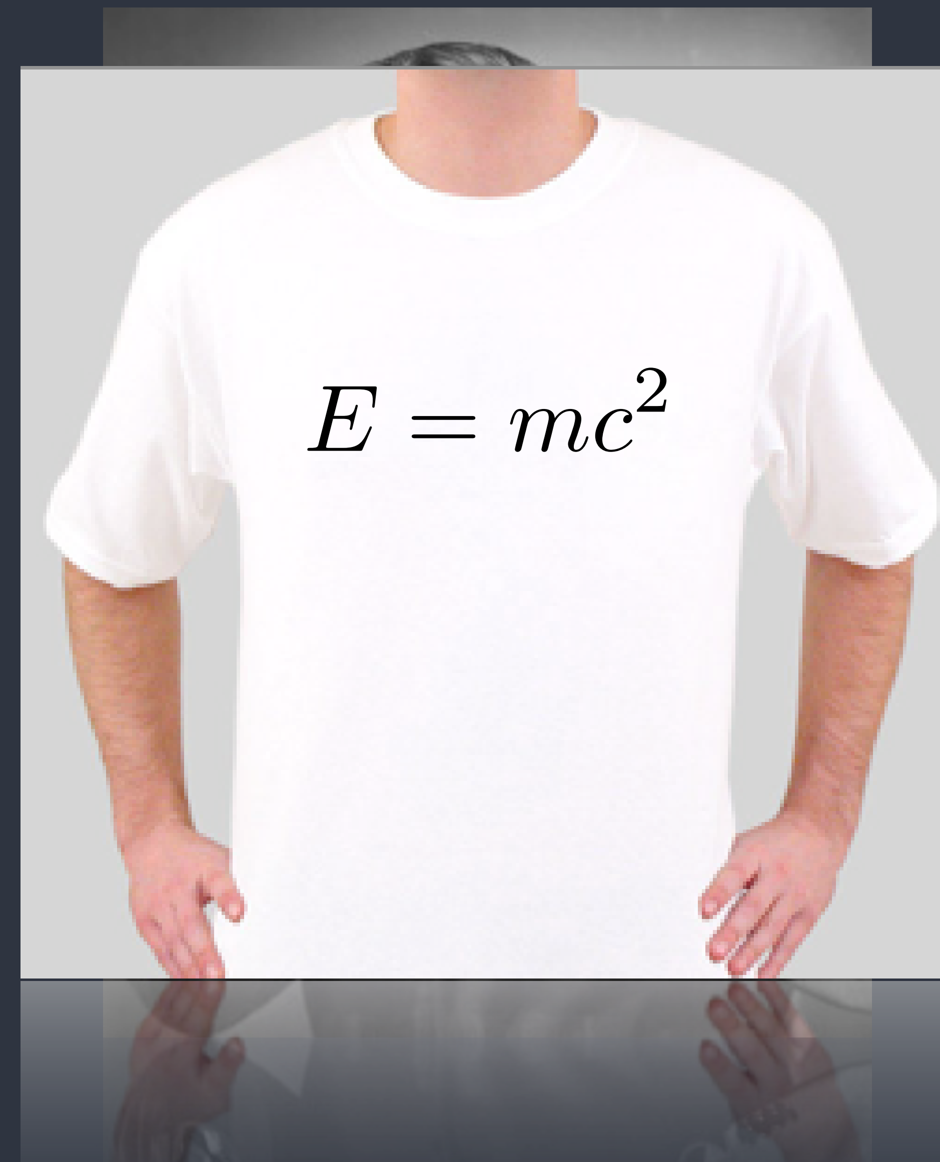
We understand:

patterns among them

&

their dynamics: forces among them

we have a theory...



precise from: atoms to 0.001 x r_{proton}

Monday, April 27, 2009

[11] In obtaining the expression (11) the mass difference between the charged and neutral has been ignored.
[12] M. Ademollo and R. Gatto, Nuovo Cimento 44A, 282 (1966); see also J. Pasupathy and R. E. Marshak, Phys. Rev. Letters 17, 888 (1966).
[13] The predicted ratio [eq. (12)] from the current algebra is slightly larger than that (0.23%) obtained from the ρ-dominance model of Ref. 2. This seems to be true also in the other case of the ratio $\Gamma(\eta\to\pi^+\pi^-\gamma)/\Gamma(\gamma\gamma)$ calculated in Refs. 12 and 14.
[14] L. M. Brown and P. Singer, Phys. Rev. Letters 8, 460 (1962).

A MODEL OF LEPTONS*

Steven Weinberg†
Laboratory for Nuclear Science and Physics Department, Massachusetts Institute of Technology, Cambridge, Massachusetts
(Received 17 October 1967)

Leptons interact only with photons, and with the intermediate bosons that presumably mediate weak interactions. What could be more natural than to unite[1] these spin-one bosons into a multiplet of gauge fields? Standing in the way of this synthesis are the obvious differences in the masses of the photon and intermediate meson, and in their couplings. We might hope to understand these differences by imagining that the symmetries relating the weak and electromagnetic interactions are exact symmetries of the Lagrangian but are broken by the vacuum. However, this raises the specter of unwanted massless Goldstone bosons.[2] This note will describe a model in which the symmetry between the electromagnetic and weak interactions is spontaneously broken, but in which the Goldstone bosons are avoided by introducing the photon and the intermediate-boson fields as gauge fields.[3] The model may be renormalizable.

We will restrict our attention to symmetry groups that connect the observed electron-type leptons only with each other, i.e., not with muon-type leptons or other unobserved leptons or hadrons. The symmetries then act on a left-handed doublet

$$L=[\tfrac{1}{2}(1+\gamma_5)]\binom{\nu_e}{e} \qquad (1)$$

and on a right-handed singlet

$$R=[\tfrac{1}{2}(1-\gamma_5)]e. \qquad (2)$$

The largest group that leaves invariant the kinematic terms $-\bar{L}\gamma^\mu\partial_\mu L-\bar{R}\gamma^\mu\partial_\mu R$ of the Lagrangian consists of the electronic isospin $\vec{T}$ acting on L, plus the numbers N_L, N_R of left- and right-handed electron-type leptons. As far as we know, two of these symmetries are entirely unbroken: the charge $Q=T_3-N_R-\tfrac{1}{2}N_L$, and the electron number $N=N_R+N_L$. But the gauge field corresponding to an unbroken symmetry will have zero mass,[4] and there is no massless particle coupled to N,[5] so we must form our gauge group out of the electronic isospin $\vec{T}$ and the electronic hyperchange $Y=N_R+\tfrac{1}{2}N_L$.

Therefore, we shall construct our Lagrangian out of L and R, plus gauge fields $\vec{A}_\mu$ and B_μ coupled to $\vec{T}$ and Y, plus a spin-zero doublet

$$\varphi=\binom{\varphi^0}{\varphi^-} \qquad (3)$$

whose vacuum expectation value will break $\vec{T}$ and Y and give the electron its mass. The only renormalizable Lagrangian which is invariant under $\vec{T}$ and Y gauge transformations is

$$\mathfrak{L}=-\tfrac{1}{4}(\partial_\mu\vec{A}_\nu-\partial_\nu\vec{A}_\mu+g\vec{A}_\mu\times\vec{A}_\nu)^2-\tfrac{1}{4}(\partial_\mu B_\nu-\partial_\nu B_\mu)^2-\bar{R}\gamma^\mu(\partial_\mu-ig'B_\mu)R-L\gamma^\mu(\partial_\mu ig\vec{t}\cdot\vec{A}_\mu-i\tfrac{1}{2}g'B_\mu)L$$
$$-\tfrac{1}{2}|\partial_\mu\varphi-ig\vec{A}_\mu\cdot\vec{t}\varphi+i\tfrac{1}{2}g'B_\mu\varphi|^2-G_e(\bar{L}\varphi R+\bar{R}\varphi^\dagger L)-M_1{}^2\varphi^\dagger\varphi+h(\varphi^\dagger\varphi)^2. \qquad (4)$$

We have chosen the phase of the R field to make G_e real, and can also adjust the phase of the L and Q fields to make the vacuum expectation value $\lambda=\langle\varphi^0\rangle$ real. The "physical" φ fields are then φ^-

We see immediately that the electron mass is λG_e. The charged spin-1 field is

$$W_\mu=2^{-1/2}(A_\mu{}^1+iA_\mu{}^2) \qquad (8)$$

and has mass

$$M_W=\tfrac{1}{2}\lambda g. \qquad (9)$$

The neutral spin-1 fields of definite mass are

$$Z_\mu=(g^2+g'^2)^{-1/2}(gA_\mu{}^3+g'B_\mu), \qquad (10)$$

$$A_\mu=(g^2+g'^2)^{-1/2}(-g'A_\mu{}^3+gB_\mu). \qquad (11)$$

Their masses are

$$M_Z=\tfrac{1}{2}\lambda(g^2+g'^2)^{1/2}, \qquad (12)$$

$$M_A=0, \qquad (13)$$

so A_μ is to be identified as the photon field. The interaction between leptons and spin-1 mesons is

$$\ldots+\text{H.c.}+\frac{igg'}{(g^2+g'^2)^{1/2}}\bar{e}\gamma^\mu eA_\mu+\frac{i(g^2+g'^2)^{1/2}}{4}\left[\left(\frac{3g'^2-g^2}{g'^2+g^2}\right)\bar{e}\gamma^\mu e-\bar{e}\gamma^\mu\gamma_5 e+\bar{\nu}\gamma^\mu(1+\gamma_5)\nu\right]Z_\mu. \qquad (14)$$

by this model have to do with the couplings of the neutral intermediate meson Z_μ. If Z_μ does not couple to hadrons then the best place to look for effects of Z_μ is in electron-neutron scattering. Applying a Fierz transformation to the W-exchange terms, the total effective e-ν interaction is

$$\frac{G_W}{\sqrt{2}}\bar{\nu}\gamma_\mu(1+\gamma_5)\nu\left\{\frac{(3g^2-g'^2)}{2(g^2+g'^2)}\bar{e}\gamma^\mu e+\tfrac{3}{2}\bar{e}\gamma^\mu\gamma_5 e\right\}.$$

If $g\gg e$ then $g\gg g'$, and this is just the usual e-ν scattering matrix element times an extra factor $\tfrac{3}{2}$. If $g\simeq e$ then $g\ll g'$, and the vector interaction is multiplied by a factor $-\tfrac{1}{2}$ rather than $\tfrac{3}{2}$. Of course our model has too many arbitrary features for these predictions to be

mi, Z. Physik 88, 161 (1934). A model similar to ours was discussed by S. Glashow, Nucl. Phys. 22, 579 (1961); the chief difference is that Glashow introduces symmetry-breaking terms into the Lagrangian, and therefore gets less definite predictions.
[2] J. Goldstone, Nuovo Cimento 19, 154 (1961); J. Goldstone, A. Salam, and S. Weinberg, Phys. Rev. 127, 965 (1962).
[3] P. W. Higgs, Phys. Letters 12, 132 (1964), Phys. Rev. Letters 13, 508 (1964), and Phys. Rev. 145, 1156 (1966); F. Englert and R. Brout, Phys. Rev. Letters 13, 321 (1964); G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, Phys. Rev. Letters 13, 585 (1964).
[4] See particularly T. W. B. Kibble, Phys. Rev. 155, 1554 (1967). A similar phenomenon occurs in the strong interactions; the ρ-meson mass in zeroth-order perturbation theory is just the bare mass, while the A_1 meson picks up an extra contribution from the spontaneous breaking of chiral symmetry. See S. Weinberg, Phys. Rev. Letters 18, 507 (1967), especially footnote 7; J. Schwinger, Phys. Letters 24B, 473 (1967); S. Glashow, H. Schnitzer, and S. Weinberg, Phys. Rev. Letters 19, 139 (1967), Eq. (13) et seq.
[5] T. D. Lee and C. N. Yang, Phys. Rev. 98, 101 (1955).
[6] This is the same sort of transformation as that which eliminates the nonderivative $\vec{\pi}$ couplings in the σ model; see S. Weinberg, Phys. Rev. Letters 18, 188 (1967). The $\vec{\pi}$ reappears with derivative coupling because the strong-interaction Lagrangian is not invariant under chiral gauge transformation.
[7] For a similar argument applied to the σ meson, see Weinberg, Ref. 6.
[8] R. P. Feynman and M. Gell-Mann, Phys. Rev. 109, 193 (1957).

6

Periodic Table of the Elements

- hydrogen
- alkali metals
- alkali earth metals
- transition metals
- poor metals
- nonmetals
- noble gases
- rare earth metals

1 H																	2 He
3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	57 La	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Ti	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	89 Ac	104 Unq	105 Unp	106 Unh	107 Uns	108 Uno	109 Une	110 Unn								

58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb	71 Lu
90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No	103 Lr

Elementary Particles: "Quarks," "leptons," and "Gauge Bosons," oh my!

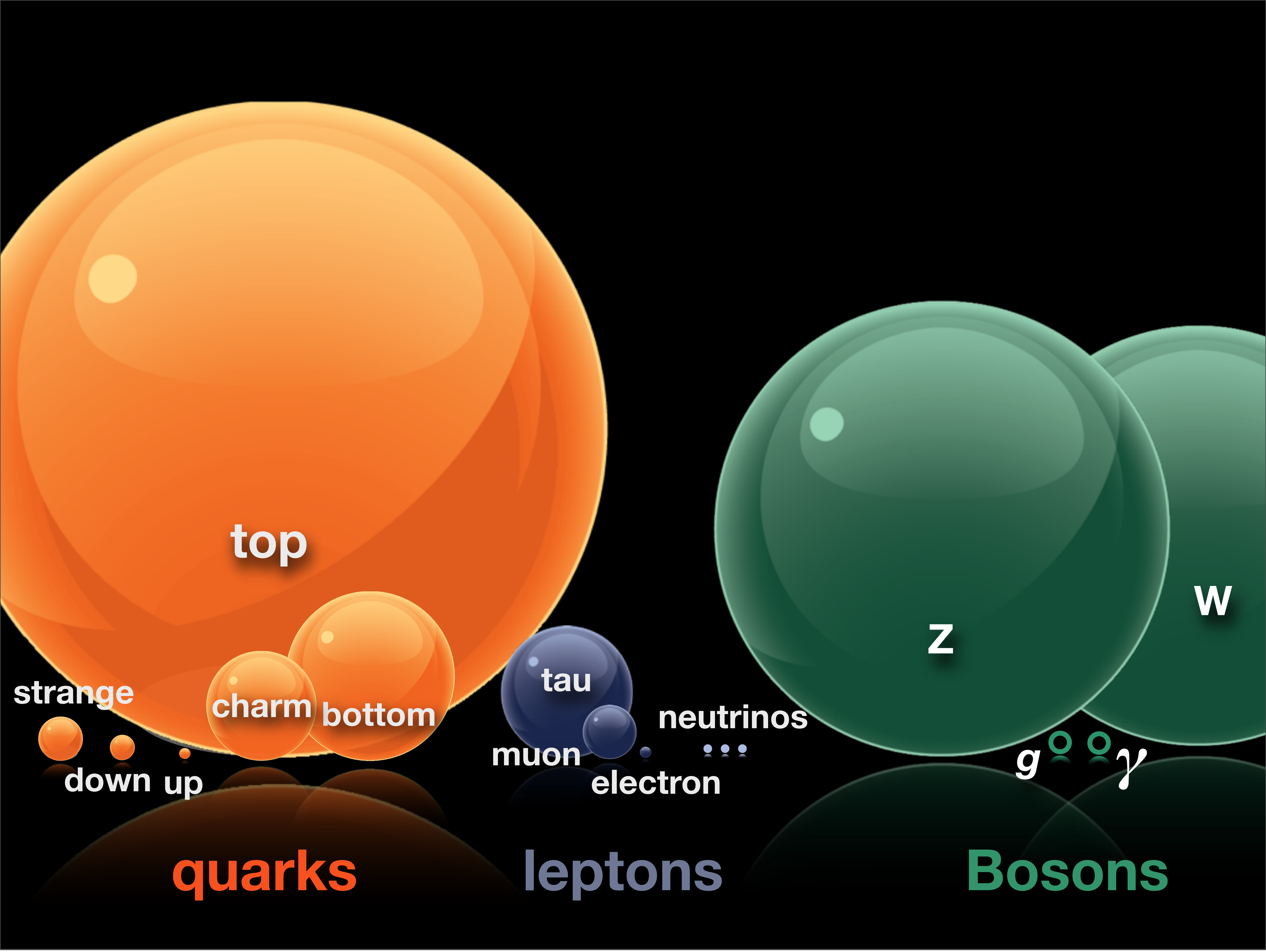
Is t[illegible] certainly not.

Leptons Quar

$$\begin{pmatrix} u \\ d \end{pmatrix} \quad \begin{pmatrix} t \\ b \end{pmatrix}$$

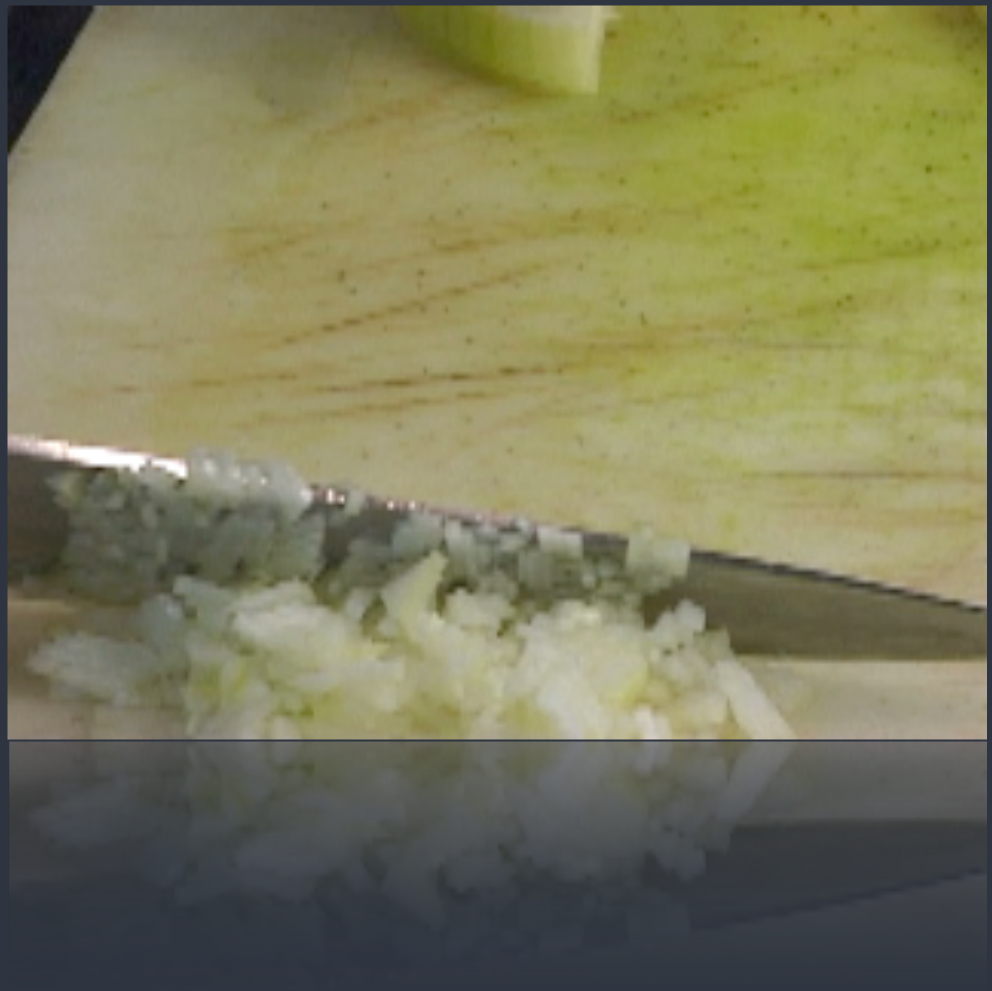
$$\begin{pmatrix} \nu_e \\ e \end{pmatrix} \quad \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}$$

Gauge Bosons γ, W, Z, g

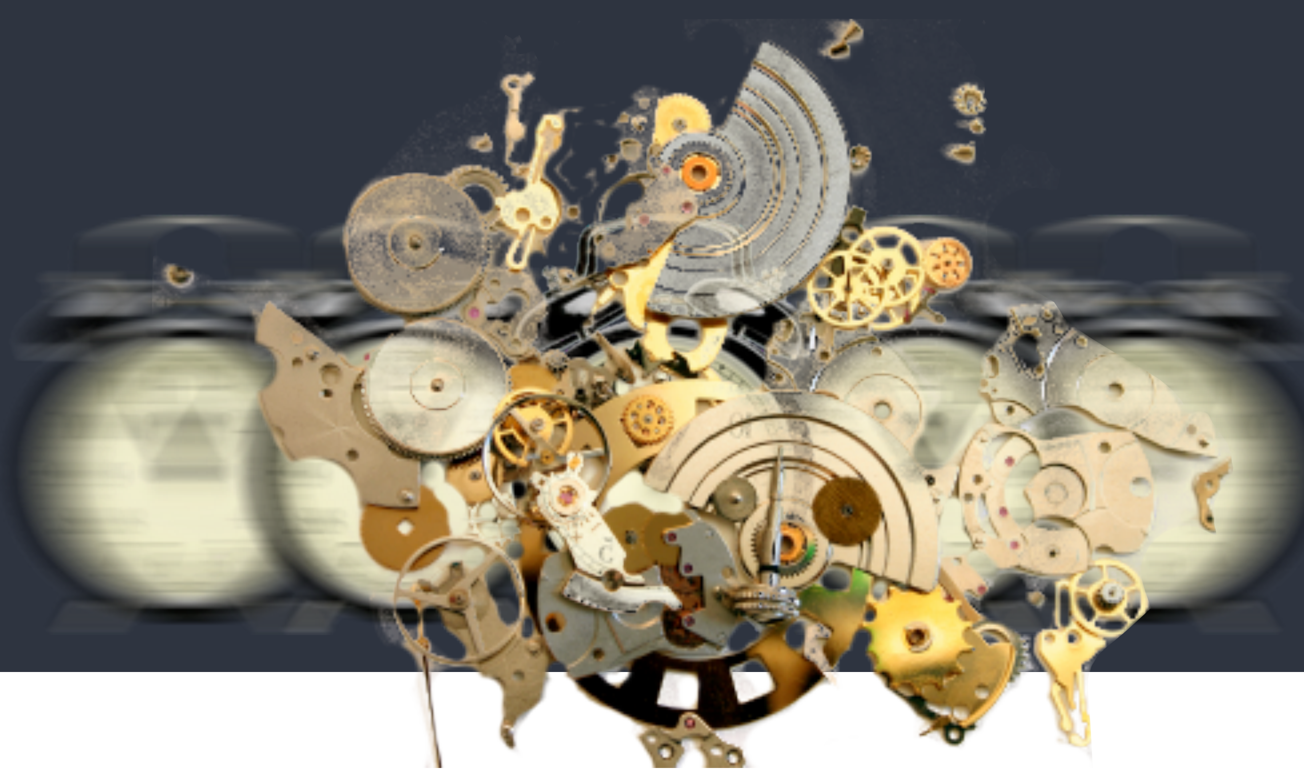

top
strange
charm
bottom
down
up
tau
muon
electron
neutrinos
Z
W
g
γ
quarks
leptons
Bosons

we create

the tiniest bits



understanding clocks?



$$E = mc^2$$

"accelerators"

we collide

protons-protons - LHC

protons-antiprotons - Fermilab

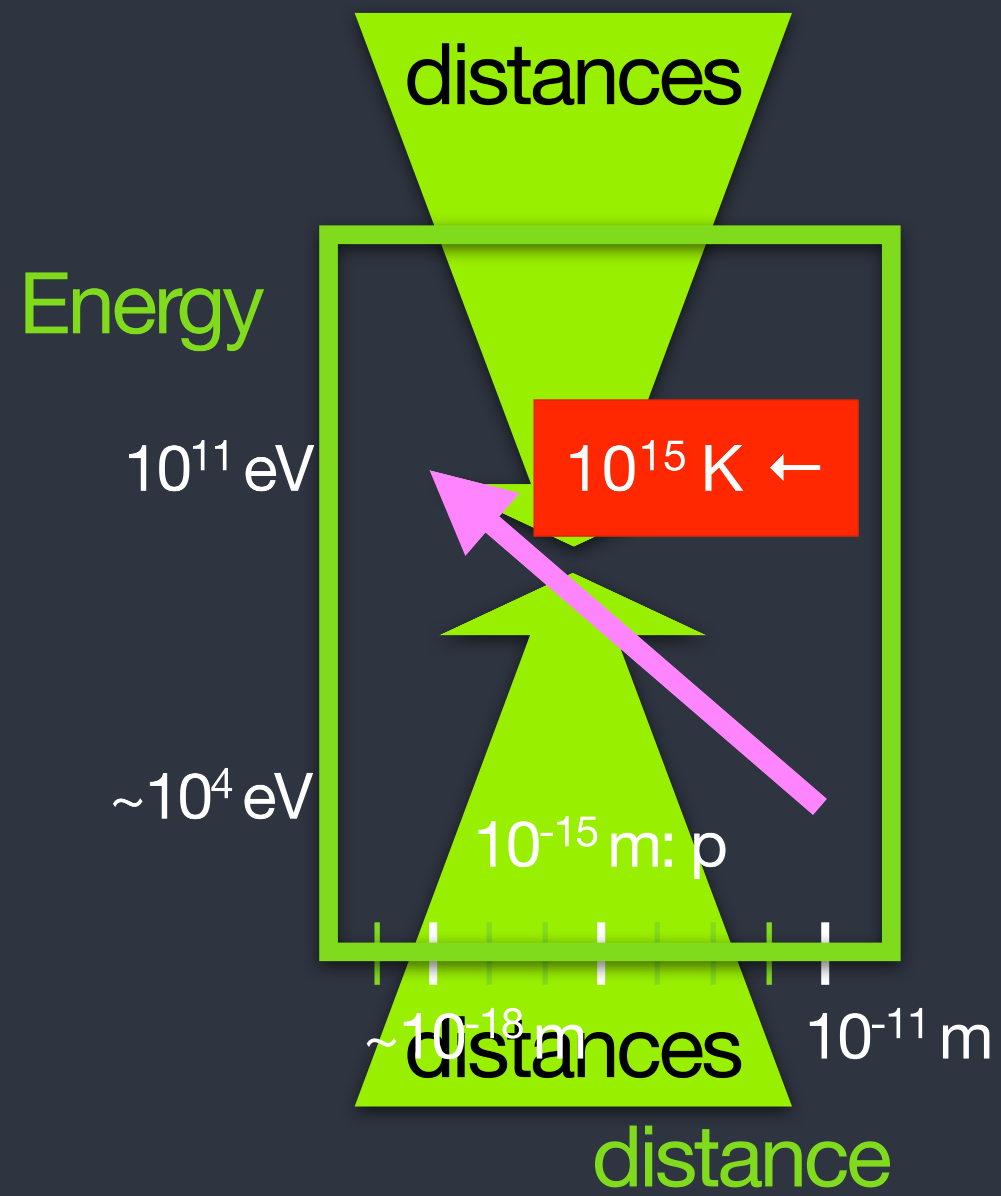
velocities within few mph of c

pretty:

Quantum Mechanics demands

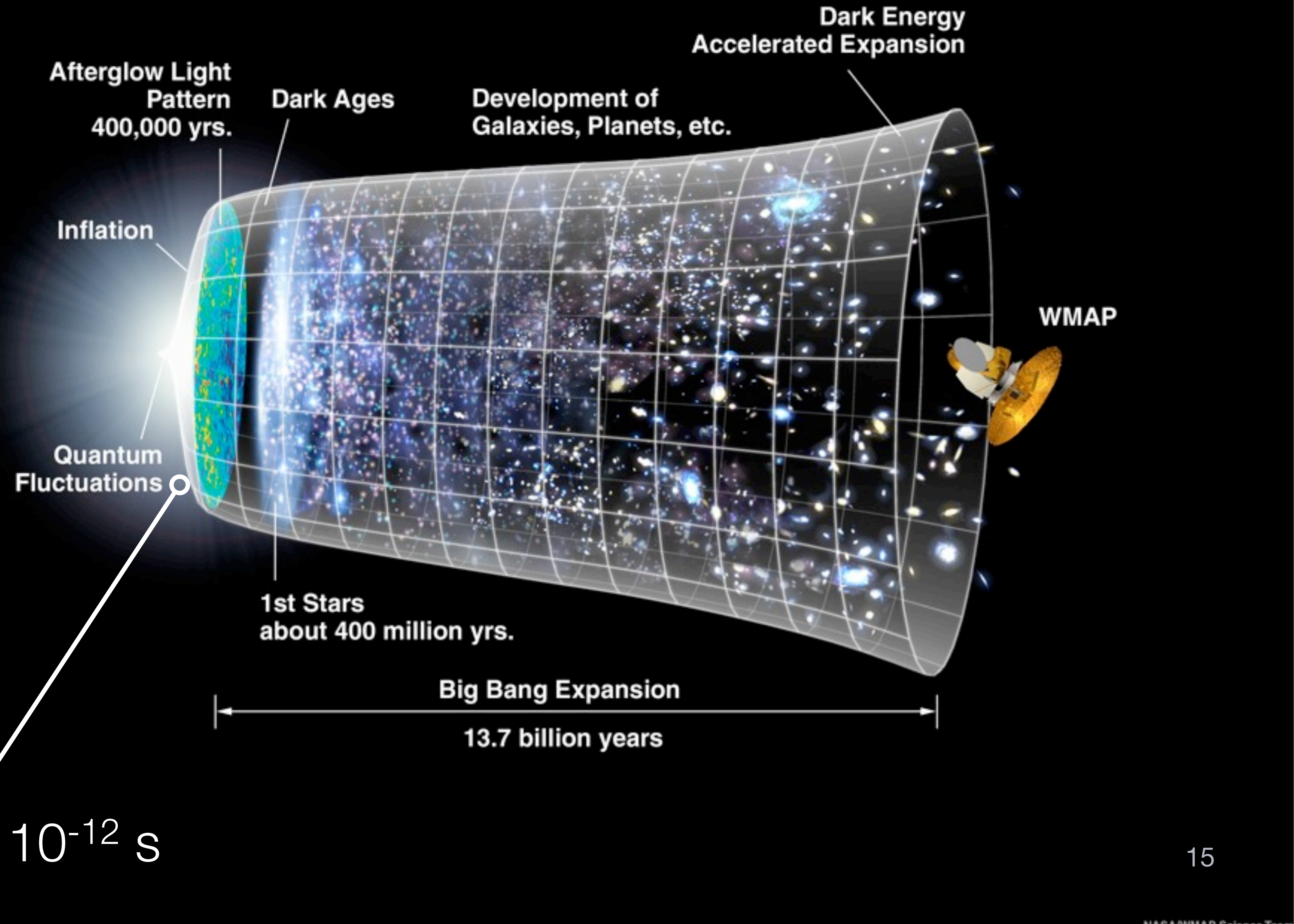
high energies = short distances

high energies = high temperatures

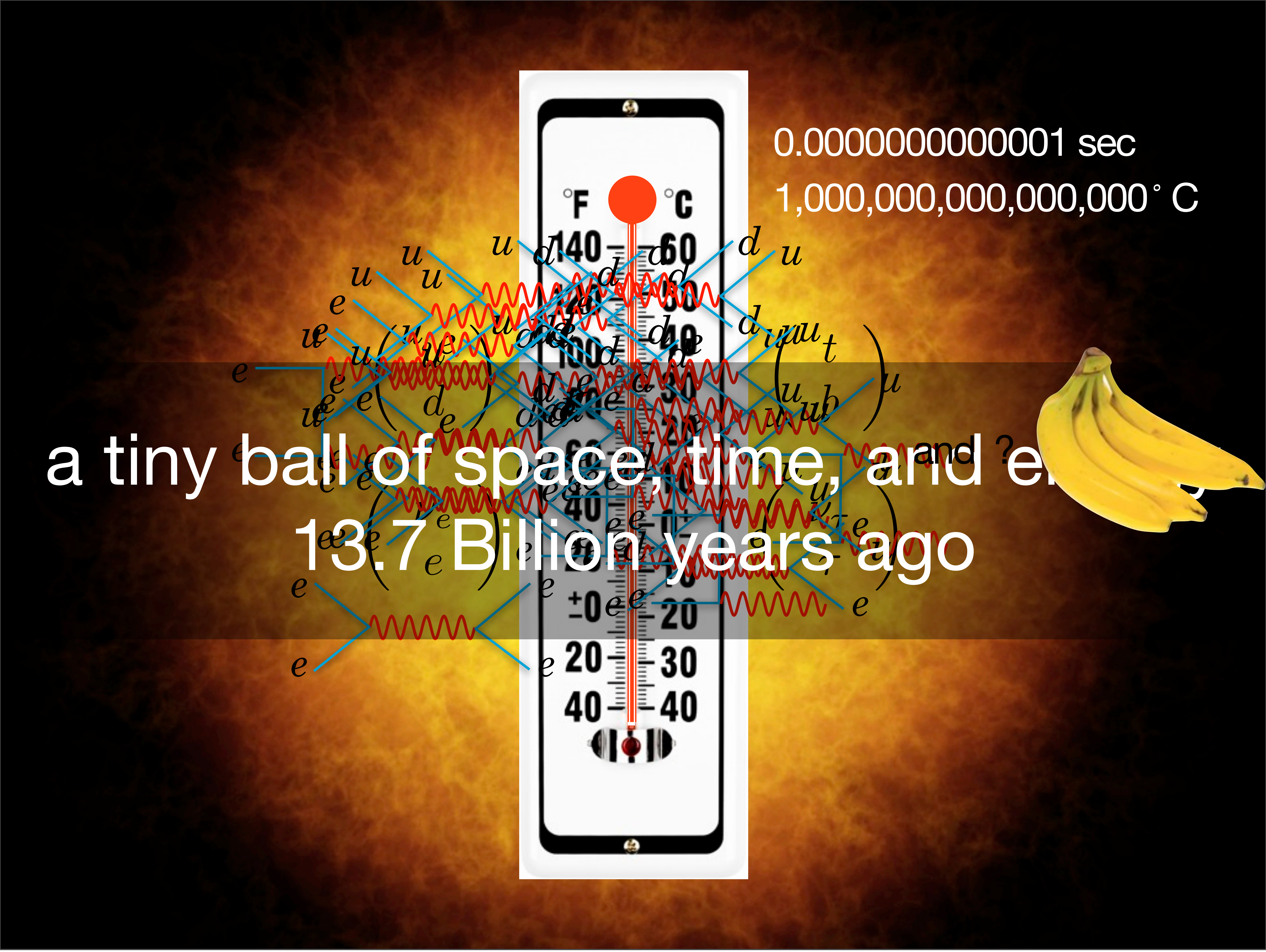


that's hot.

reminiscent of one event


Dark Energy
Accelerated Expansion
Afterglow Light
Pattern
400,000 yrs.
Dark Ages
Development of
Galaxies, Planets, etc.
Inflation
WMAP
Quantum
Fluctuations
1st Stars
about 400 million yrs.
Big Bang Expansion
13.7 billion years
NASA/WMAP Science Team

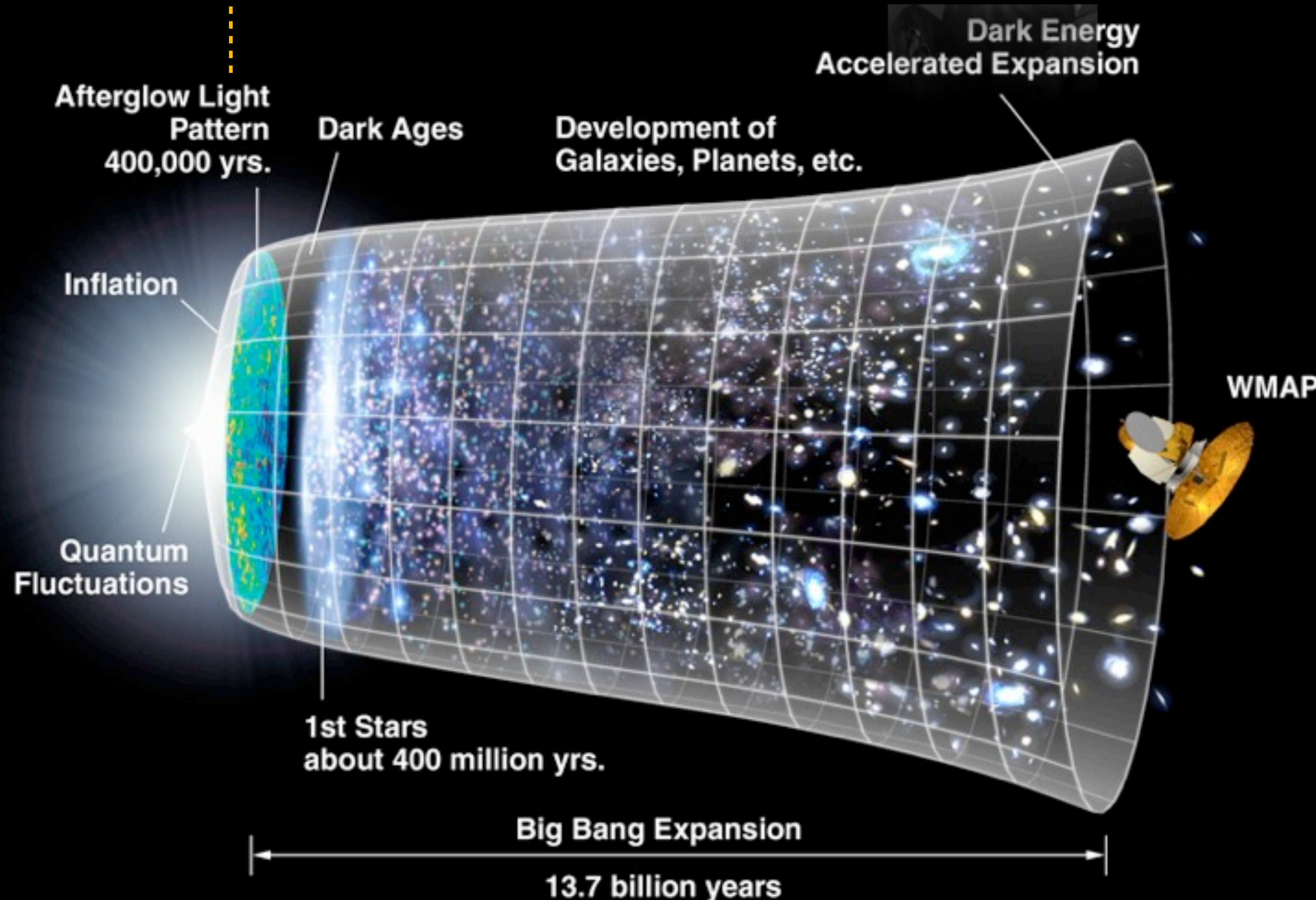
10^{15} K → 10^{-12} s


0.0000000000001 sec
1,000,000,000,000,000˚ C
a tiny ball of space, time, and e
and ?
13.7 Billion years ago

Mmmm, Mmmm Good....

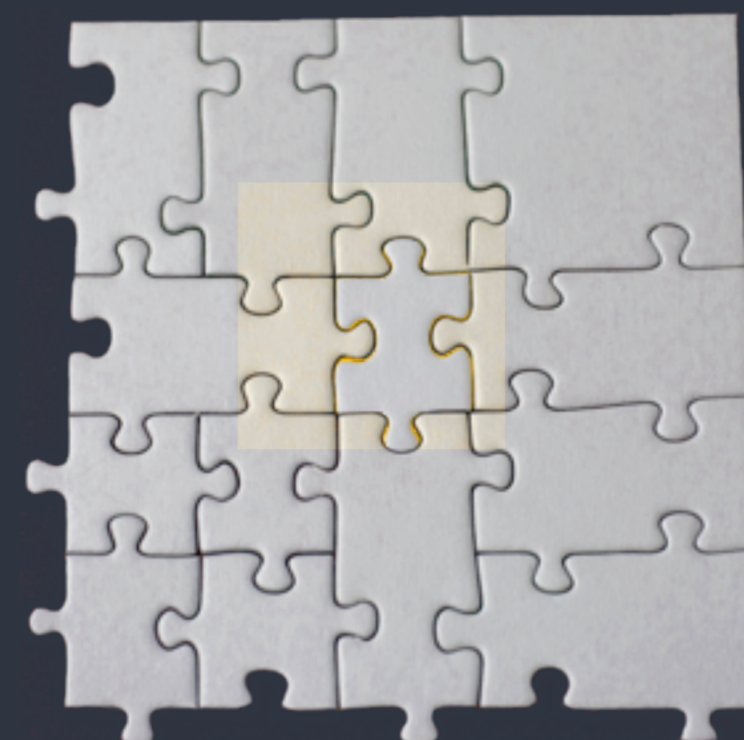
~10^{-12} sec

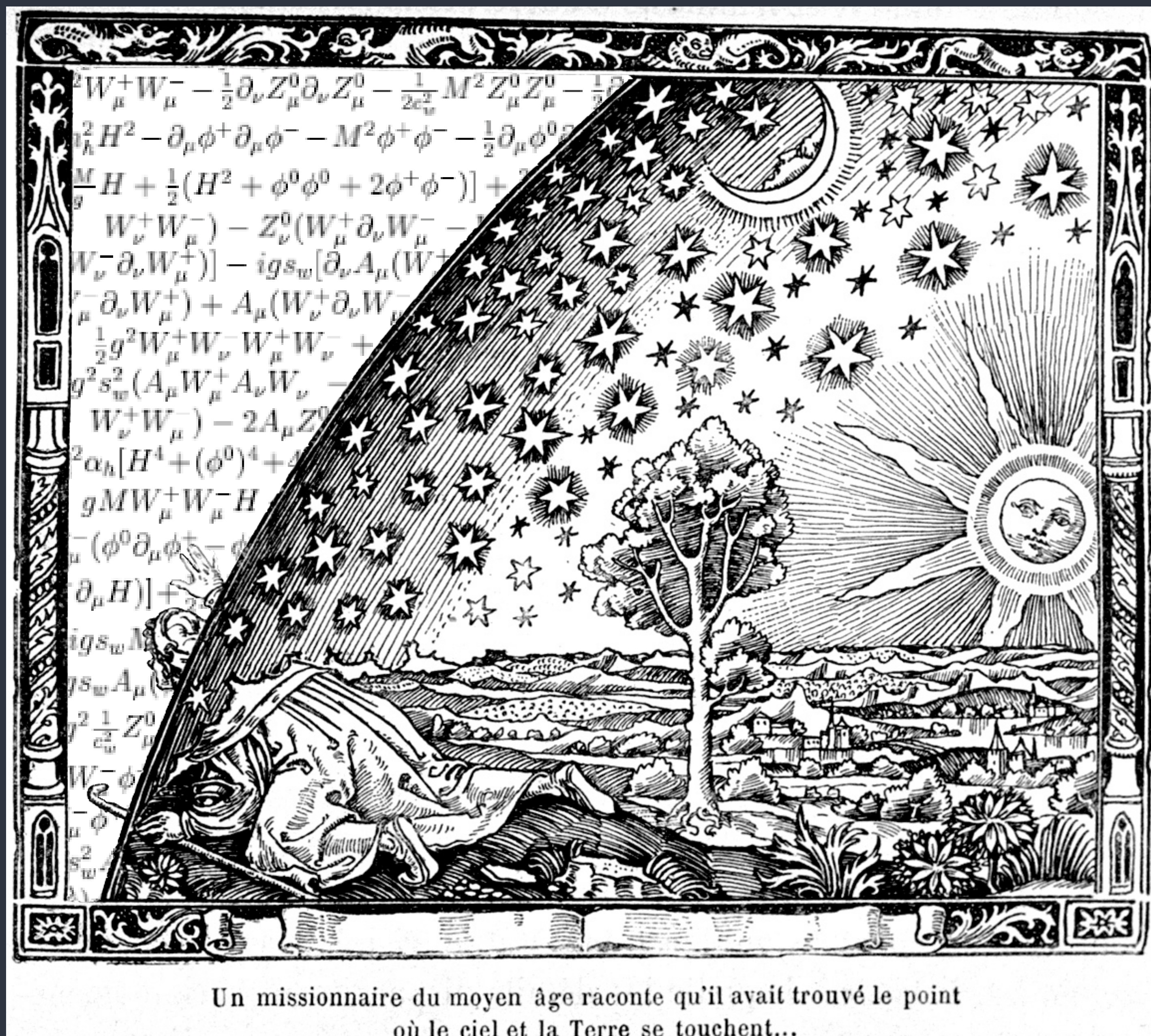
Standard Model is good!



Standard Model

is incomplete





deeper...

the mathematics fails us...

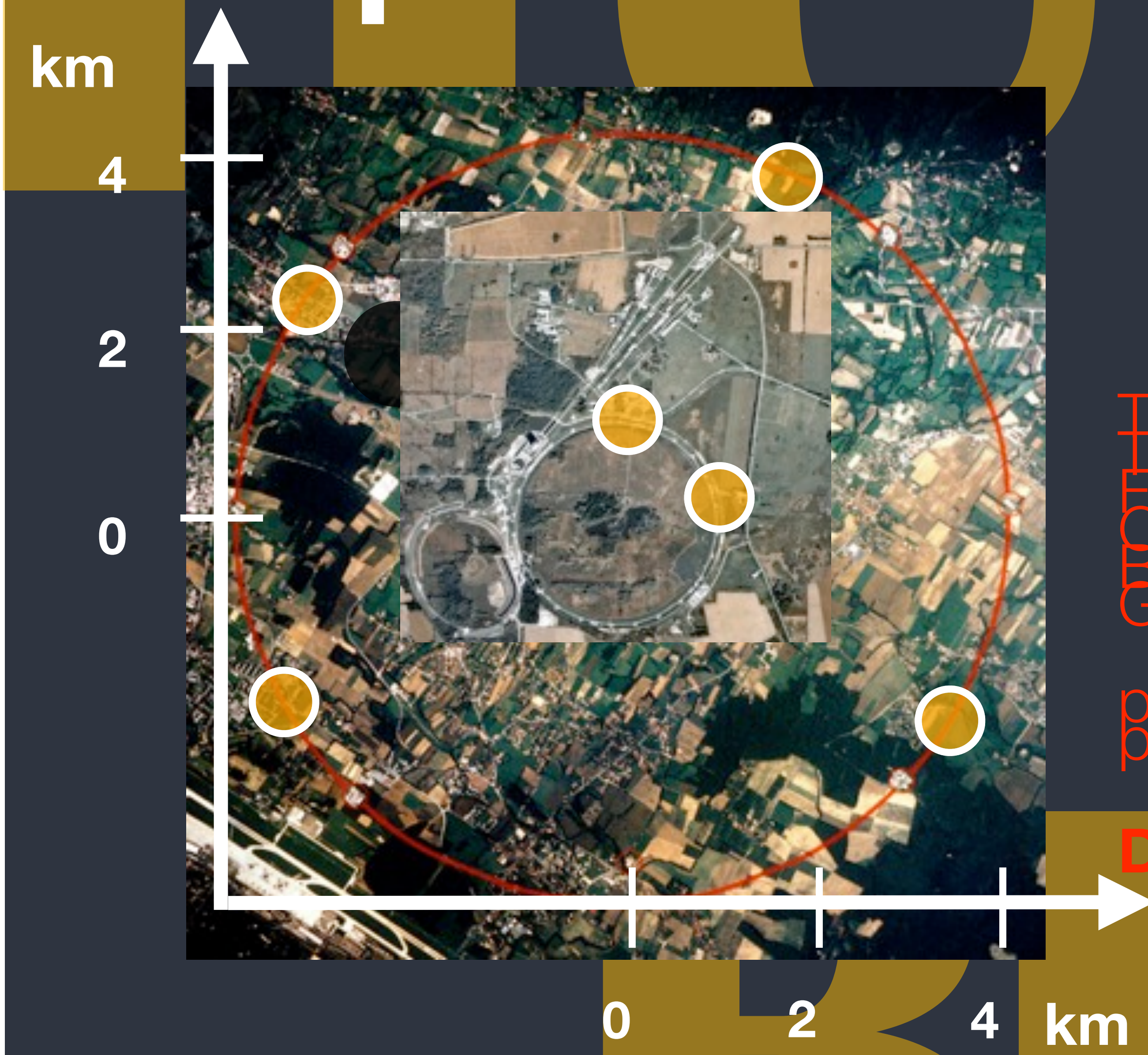
the Standard Model fails...

be there dragons?

after: Camille Flammarion, L'Atmosphere: Météorologie Populaire (Paris, 1888), p. 163.

HUGE

particle accelerators



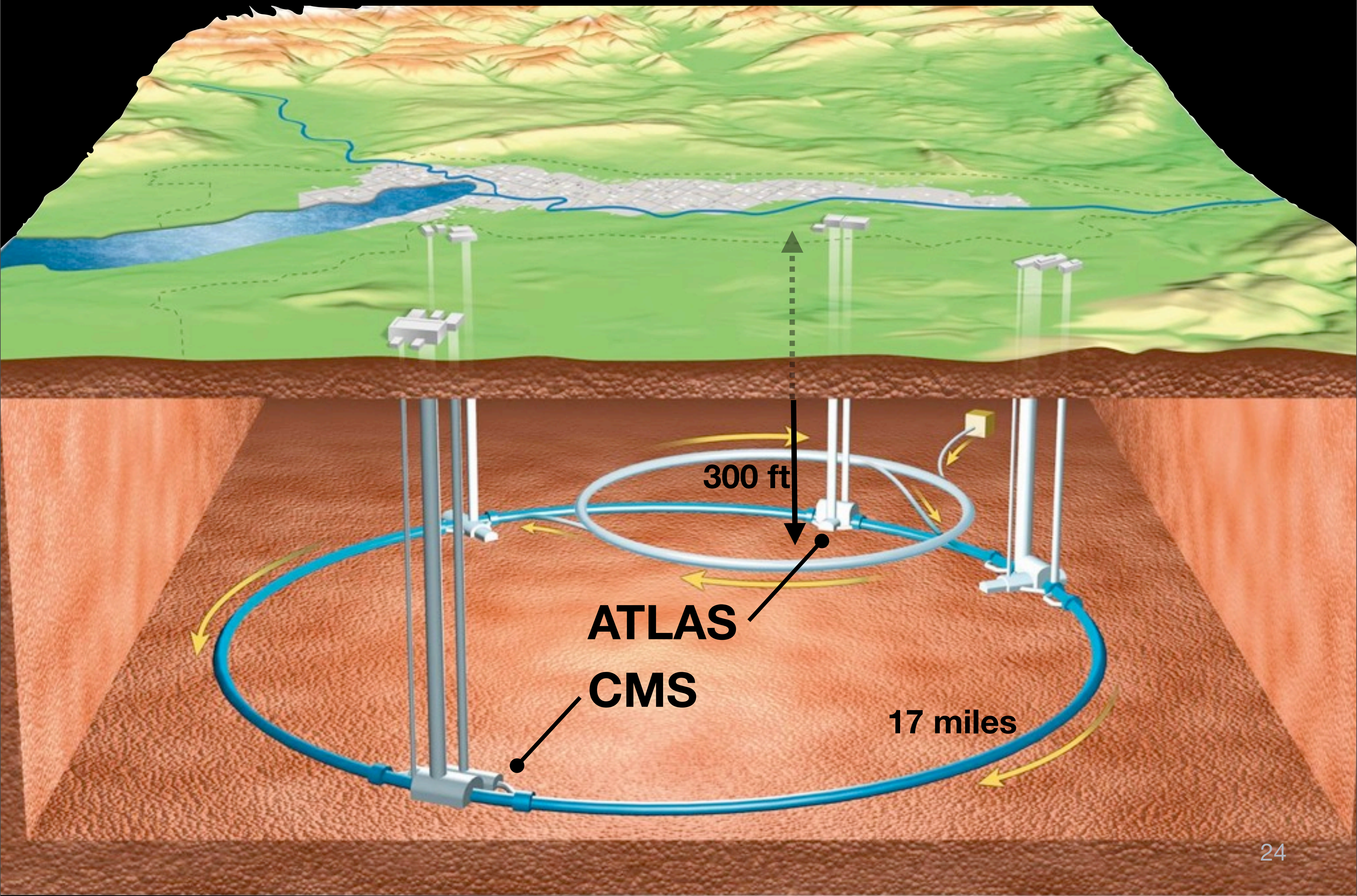
The "Tevatron"
The Large Hadron Collider
Fermi National Accelerator Laboratory
CERN
Batavia, IL
Geneva, Switzerland

$p\bar{p}$; 980 GeV/c per beam
pp; 7000 GeV/c per beam

DØ and CDF Experiments

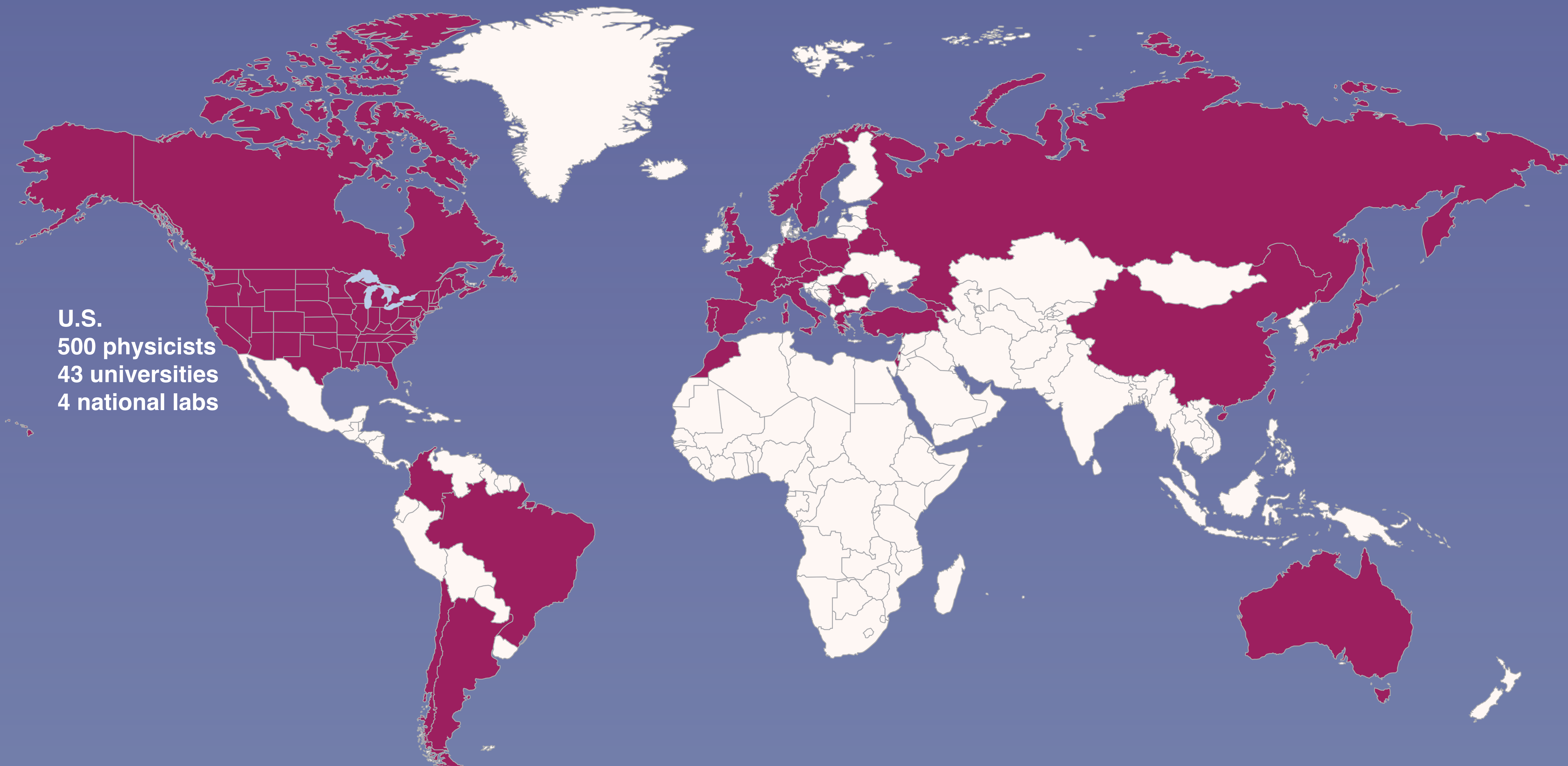
BIG


Geneva, Switzerland
CERN
European Organization for
Nuclear Research
Large Hadron Collider (LHC)


300 ft
ATLAS
CMS
17 miles

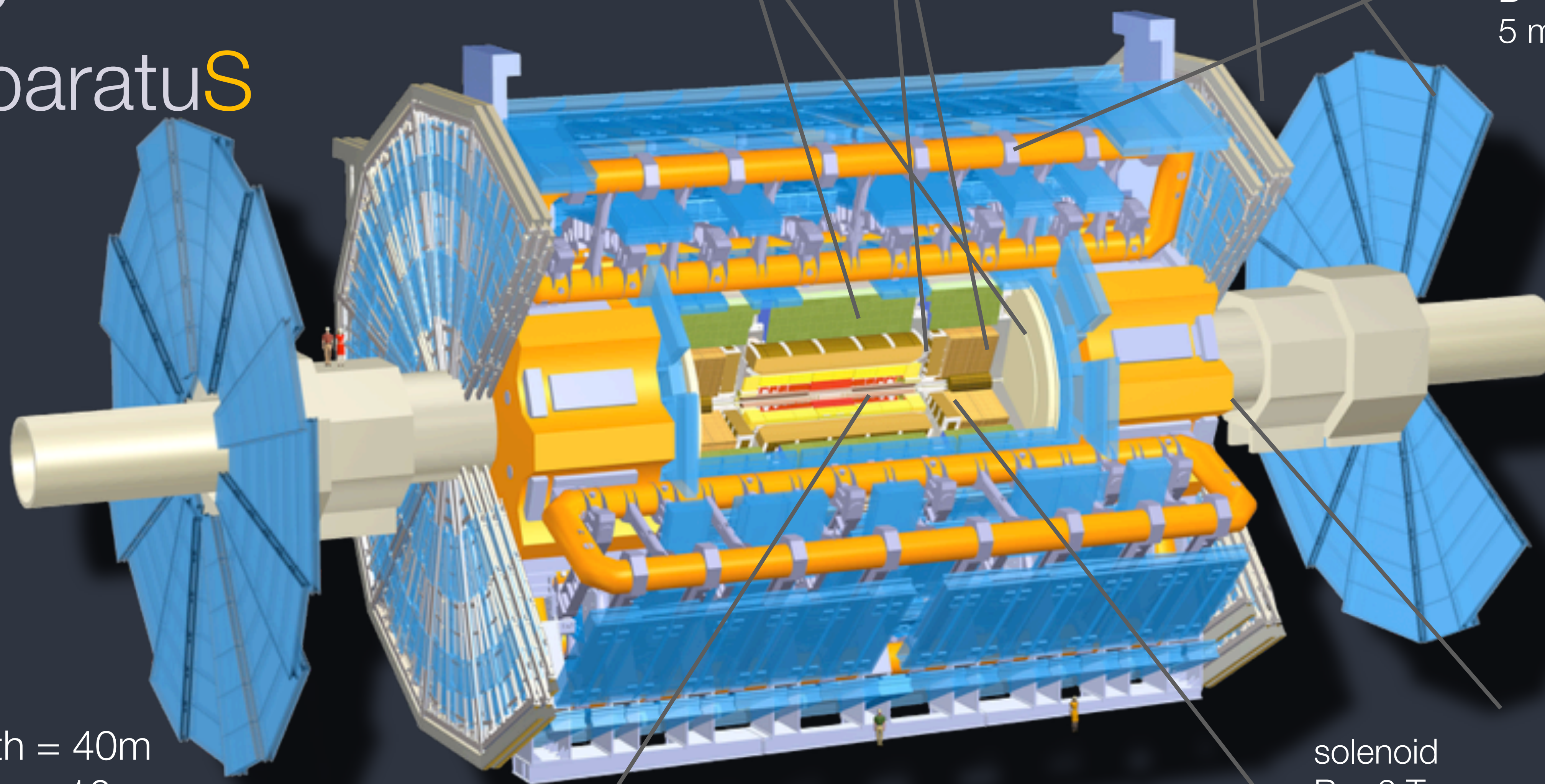
ATLAS Collaboration

39 nations
164 institutions
1900 authors
400 PhD students



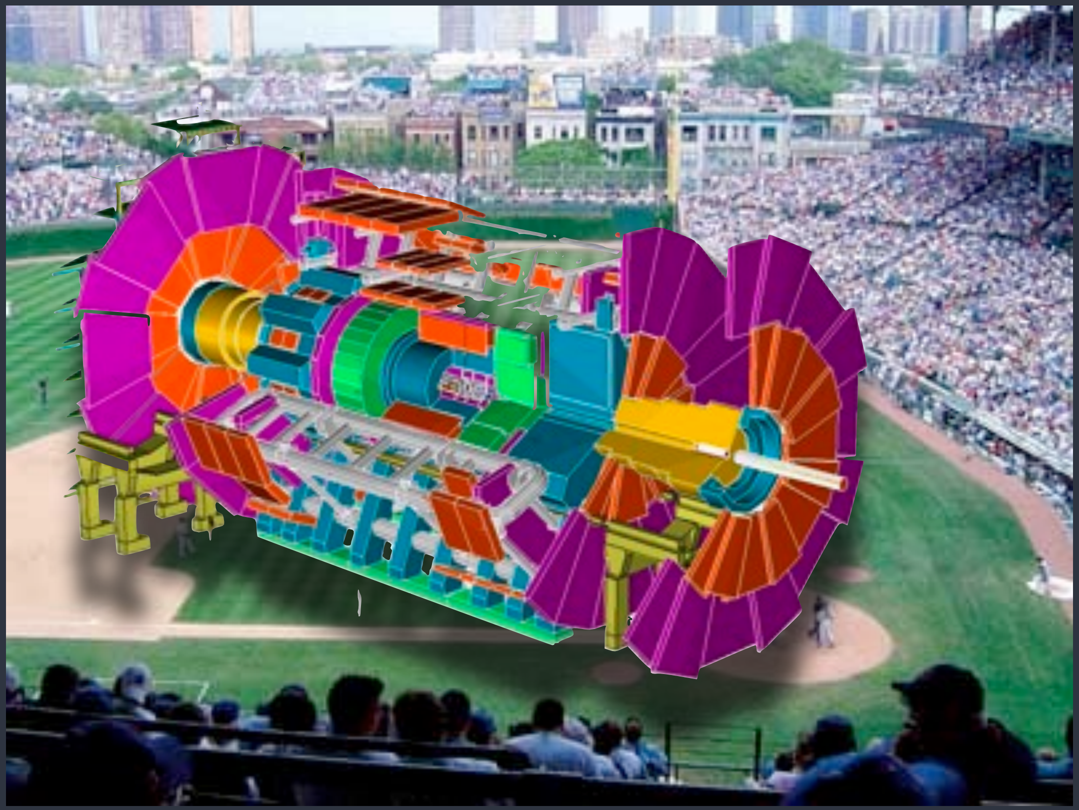
ATLAS

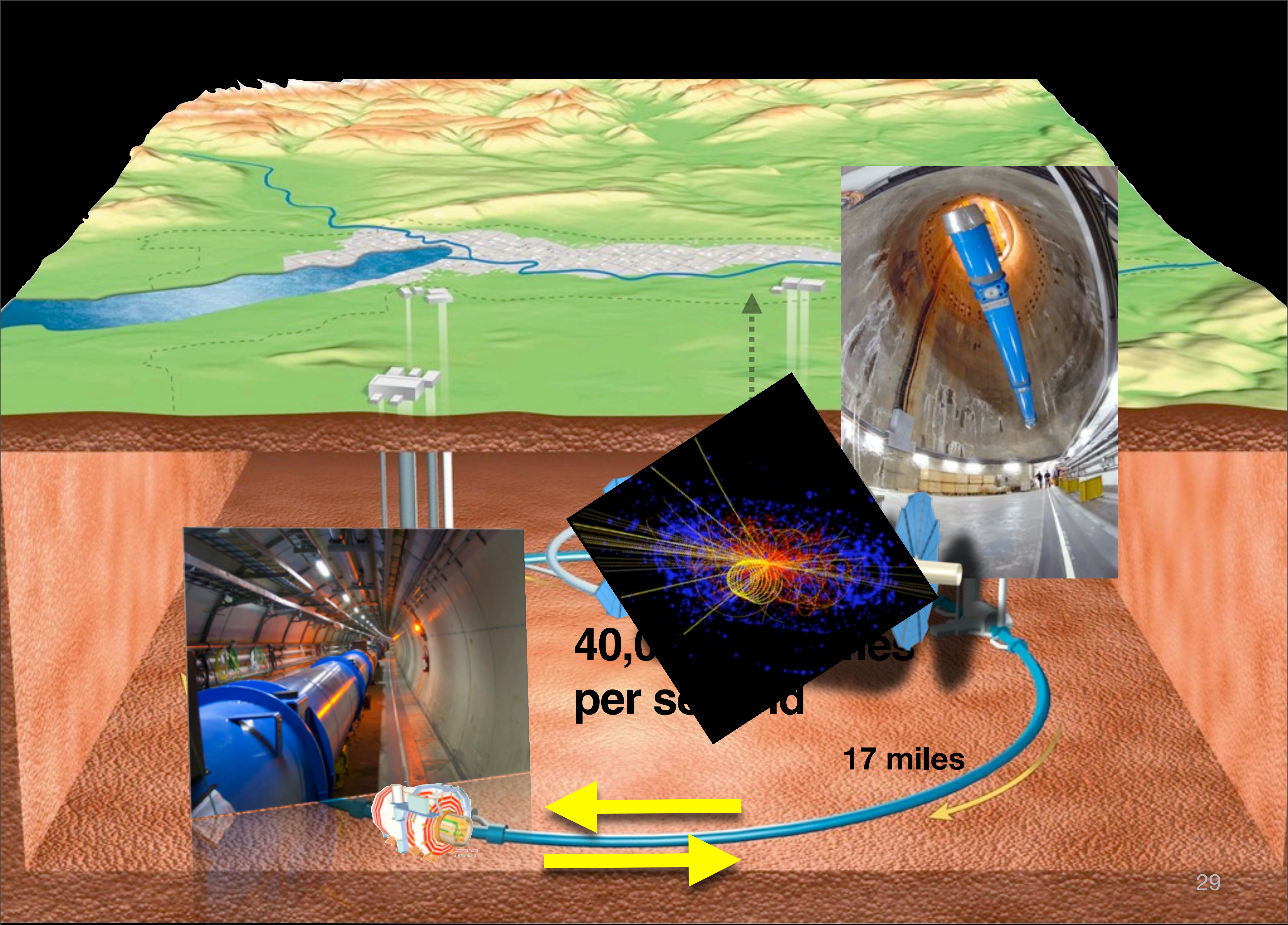
A
Toriodal
Lhc
ApparatuS



Length = 40m
Radius = 10 m
Weight = 7,000 t


2006 2. 21




17 miles

Status of the Machine:

39th (final) dipole was installed last Thursday

Vacuum work ongoing

Beams: September, 2009

Physics: late October

(5 TeV per beam, ~300pb^{-1}, 2010/2011, 2×10^{32} $cm^{-1}s^{-1}$)

2 the data “problem”

SO…

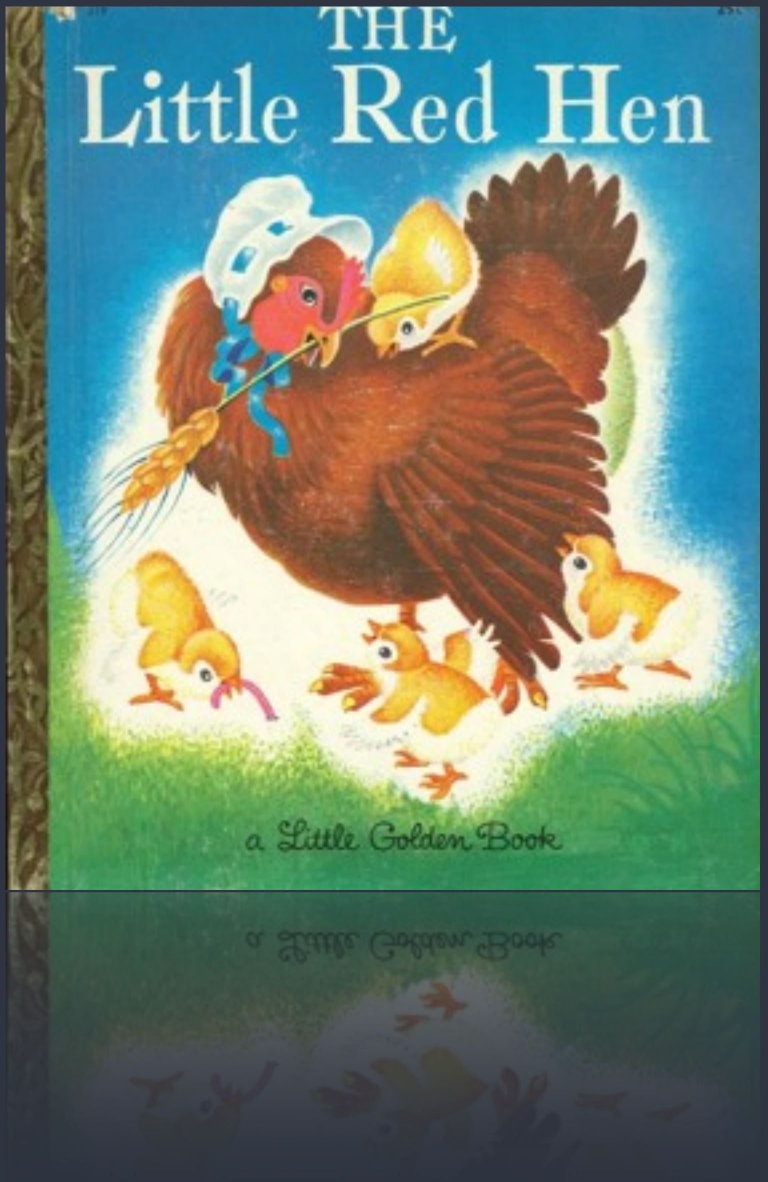
just:

“plant” the beams

“hoe” the collision debris

“pick” the results

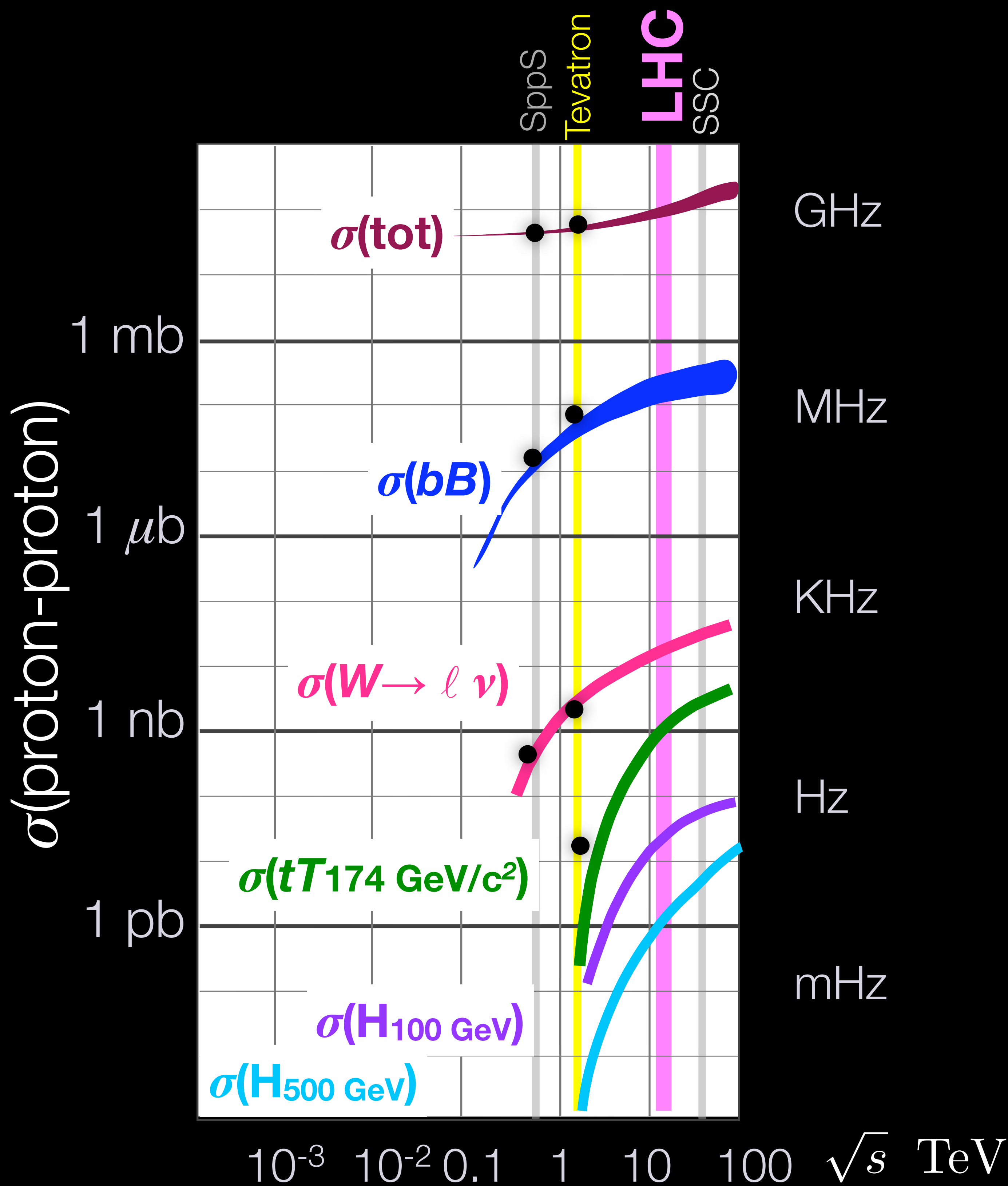
…publish



LHC (Tevatron) = $10^{33\text{-}34}$ (10^{32}) $cm^{-2}\ s^{-1}$

interaction rate = 700 MHz at $10^{34}\ cm^{-2}\ s^{-1}$

⇒ 17.5 events per crossing (= 23 in practice)

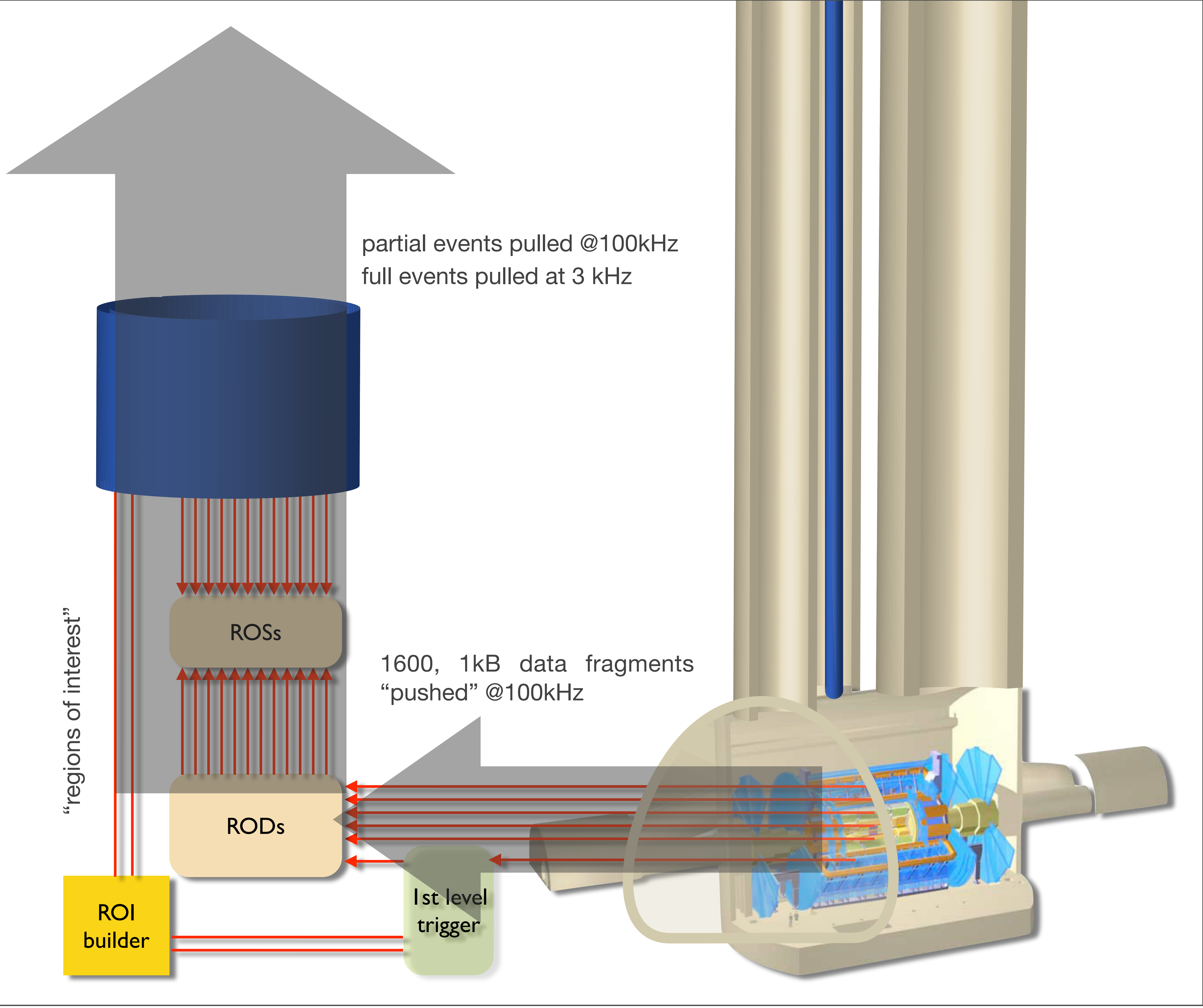


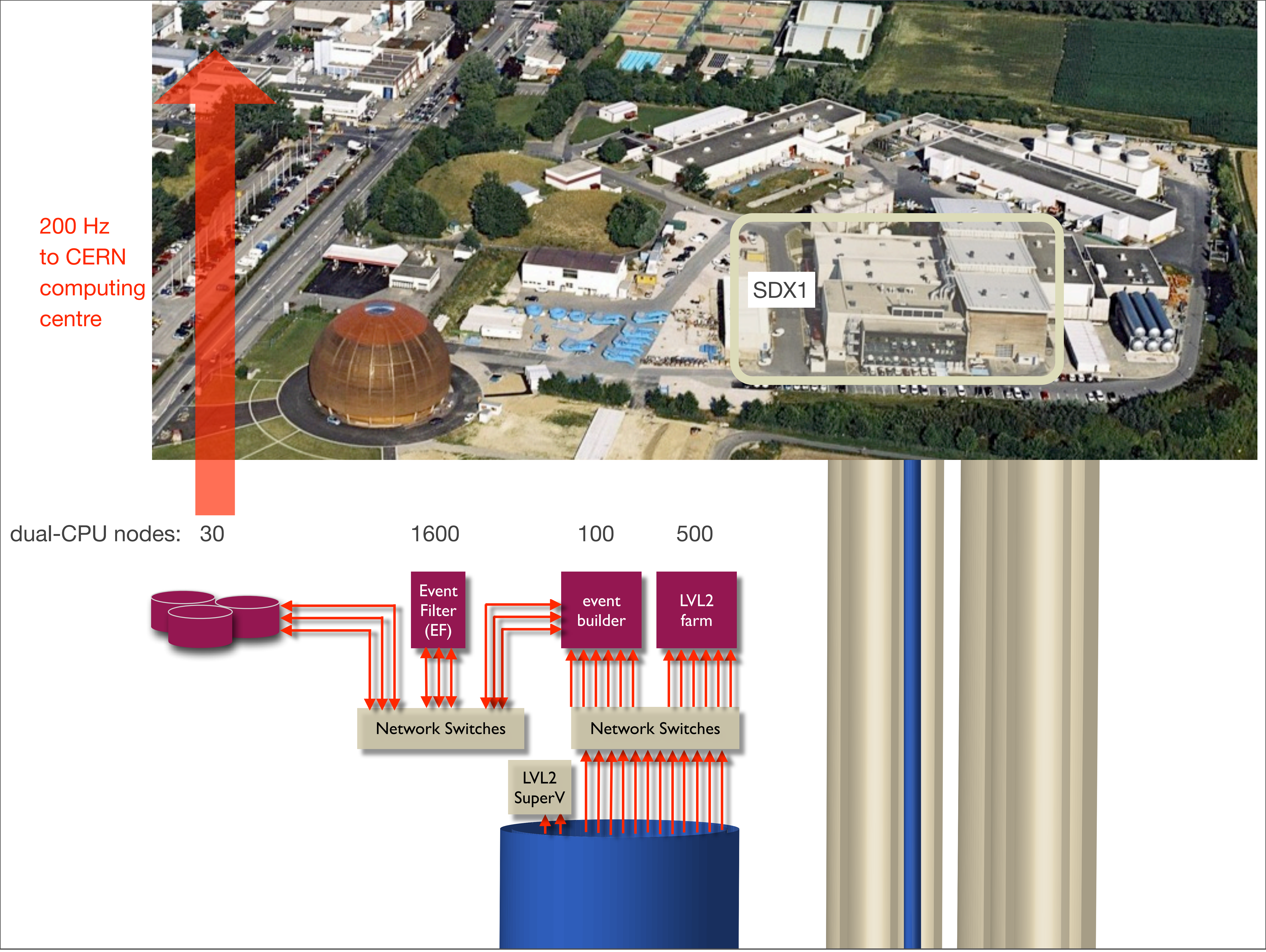
contend with this

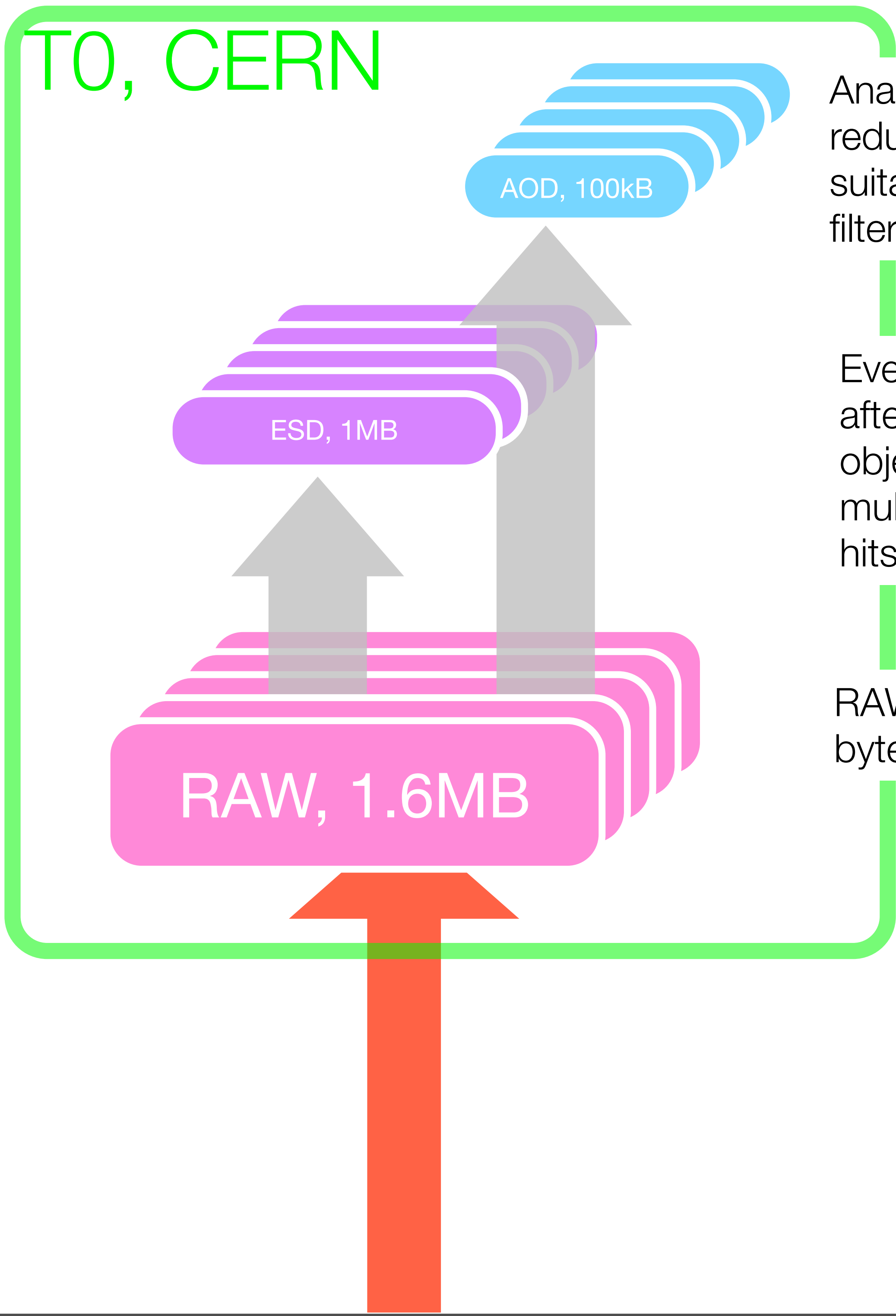
sensitive to this

@10^{34} cm^{-2} s^{-1}

Monday, April 27, 2009


partial events pulled @100kHz
full events pulled at 3 kHz
ROSs
1600, 1kB data fragments "pushed" @100kHz
RODs
"regions of interest"
ROI builder
1st level trigger

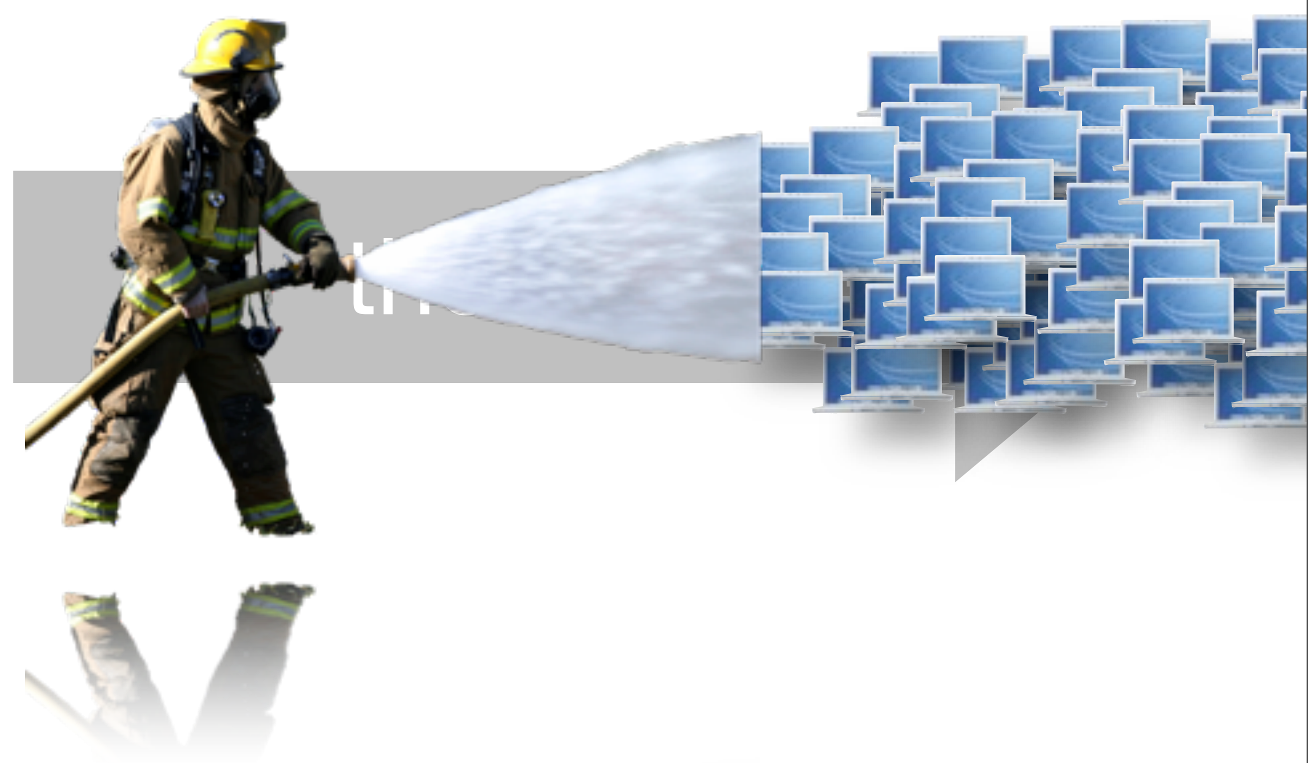

200 Hz
to CERN
computing
centre
SDX1
dual-CPU nodes: 30
1600
100
500
Event
Filter
(EF)
event
builder
LVL2
farm
Network Switches
Network Switches
LVL2
SuperV



Analysis Object Data
reduced object-format
suitable for analysis
filtered e, j, mu

Event Summary Data
after RECO
object-format
multiple e, j, mu
hits, cells

RAW
byte-stream

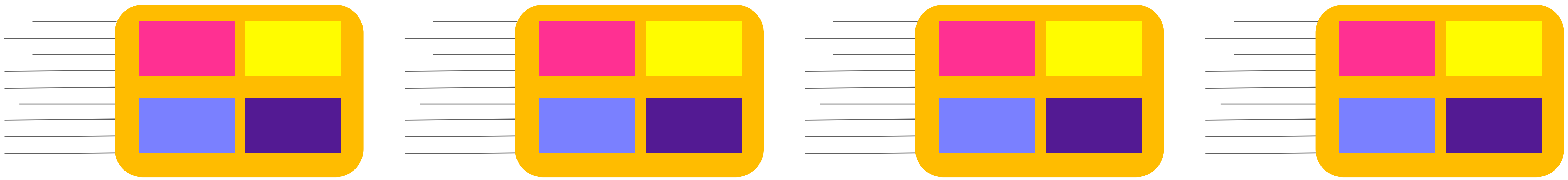


Have to analyze it

3 PetaBytes of data/year

keep that up for
2 decades.

skimthinslimaug



RECO

SKIM

THIN

SLIM

AUGment

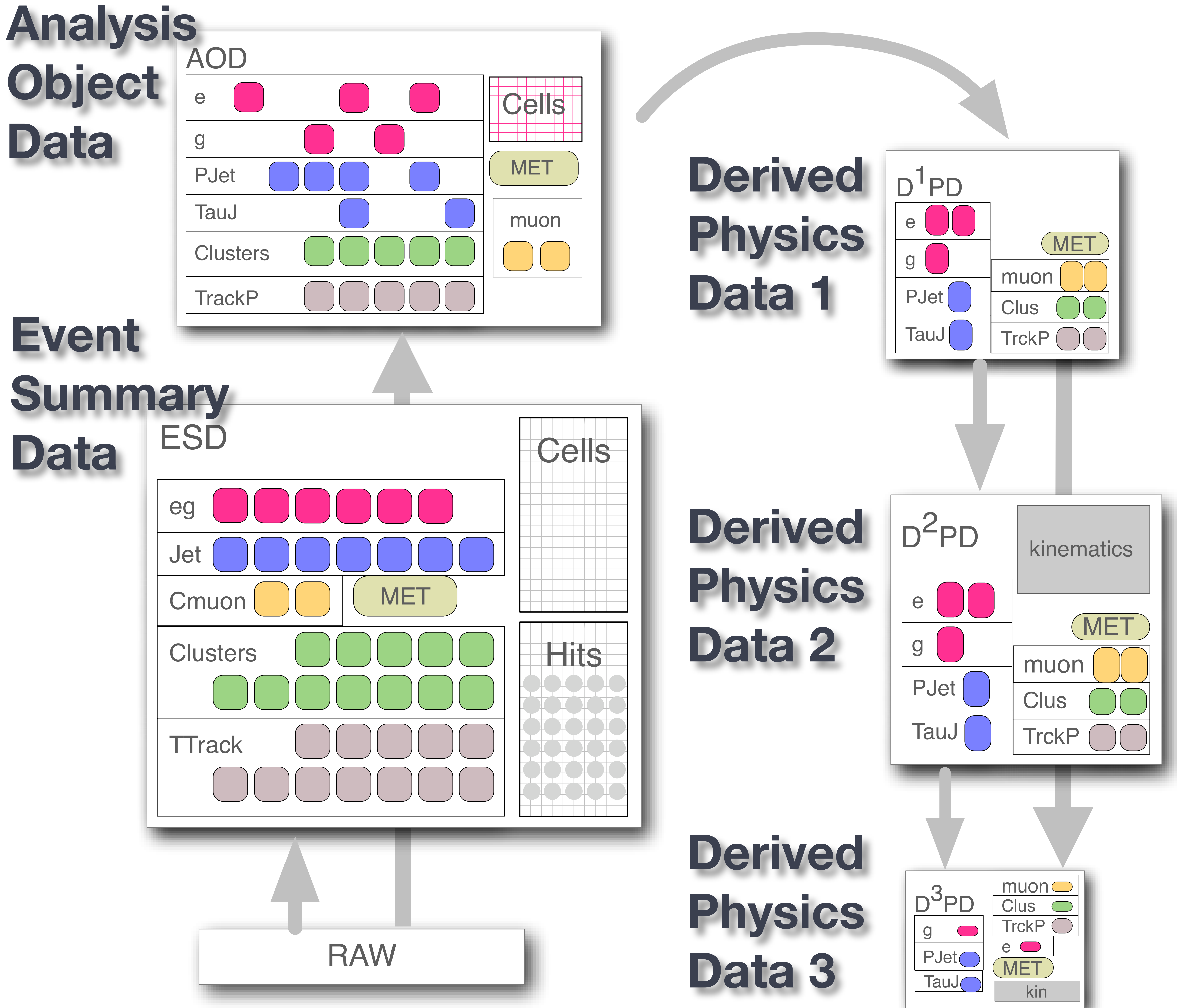

Analysis Object Data
AOD
e
g
PJet
TauJ
Clusters
TrackP
Cells
MET
muon
Derived Physics Data 1
D1PD
e
g
PJet
TauJ
MET
muon
Clus
TrckP
Event Summary Data
ESD
eg
Jet
Cmuon
MET
Clusters
TTrack
Cells
Hits
RAW
Derived Physics Data 2
D2PD
kinematics
e
g
PJet
TauJ
MET
muon
Clus
TrckP
Derived Physics Data 3
D3PD
g
PJet
TauJ
muon
Clus
TrckP
e
MET
kin

Table 3: Data formats for ATLAS and quantities used in this analysis.

Format	Target Range	Current	Used	1 Year Dataset
RAW	1.6 MB		1.6 MB	1600 TB
ESD	0.5 MB	0.7 MB	0.5 MB	500 TB
MC ESD	0.5 MB		0.5 MB	500 TB
AOD	0.1 MB	0.17 MB	0.150 MB	100 TB
TAG	1 kB		1 kB	1 TB

Table 6: DPD formats and size estimates. N.B. The DPD current amounts are from [15] and are approximations to FDR $t\bar{t}$ data and are just presented as a snapshot and not to be taken literally.

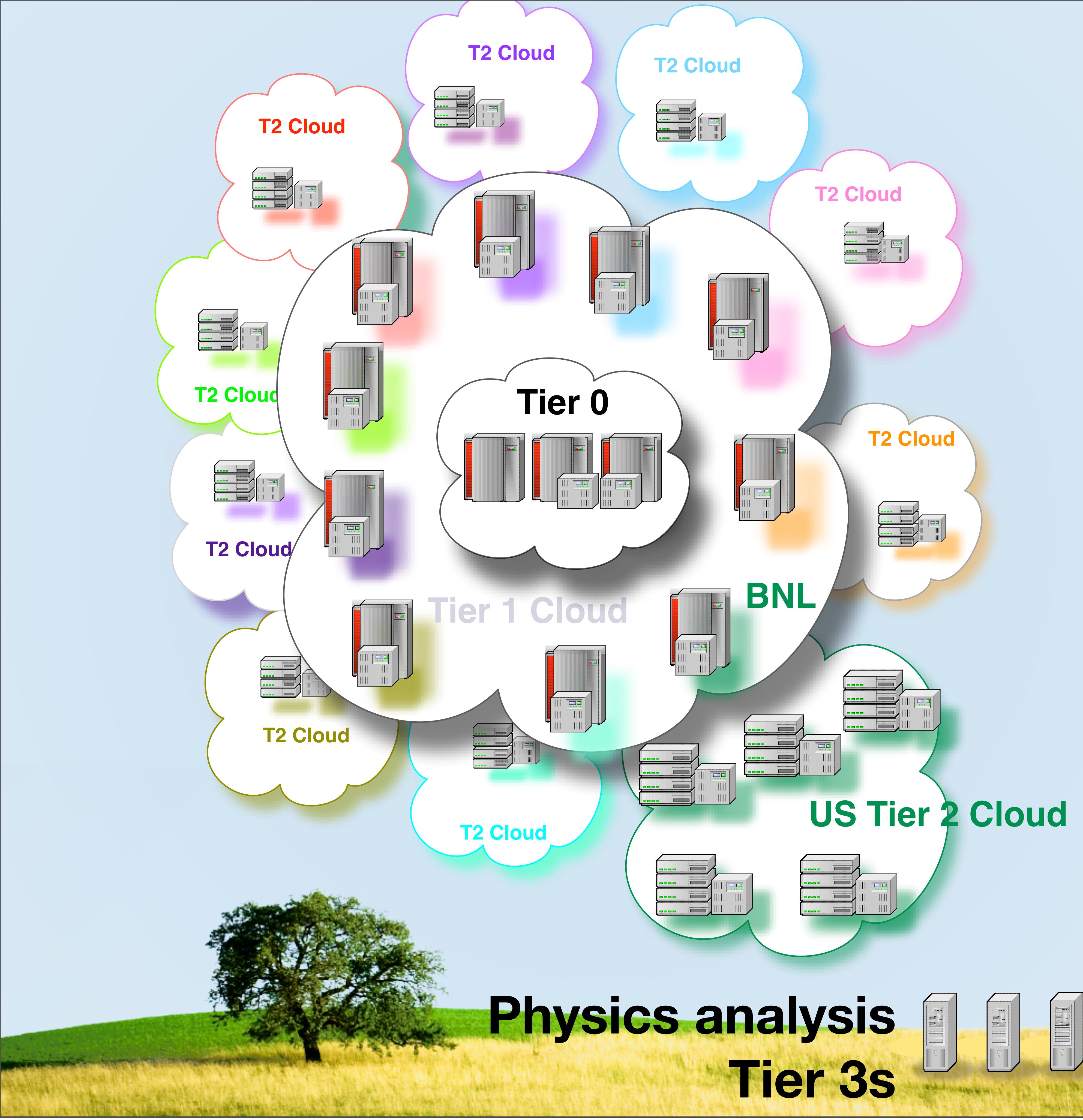
Format	Target Range	Current	Used	1 Year Dataset
D^1PD	$1/4\times$ AOD	31 kB	25 kB	25 TB
D^2PD	$1.1\times D^1PD$	18 kB	30 kB	30 TB
D^3PD	$1/3\times D^1PD$	5 kB	6 kB	6 TB
pDPD	?	NA	?	?

that's a lot of data

that's a lot of formats

ATLAS data come in all shapes and sizes

where are they made? where are they stored? Not determined yet.

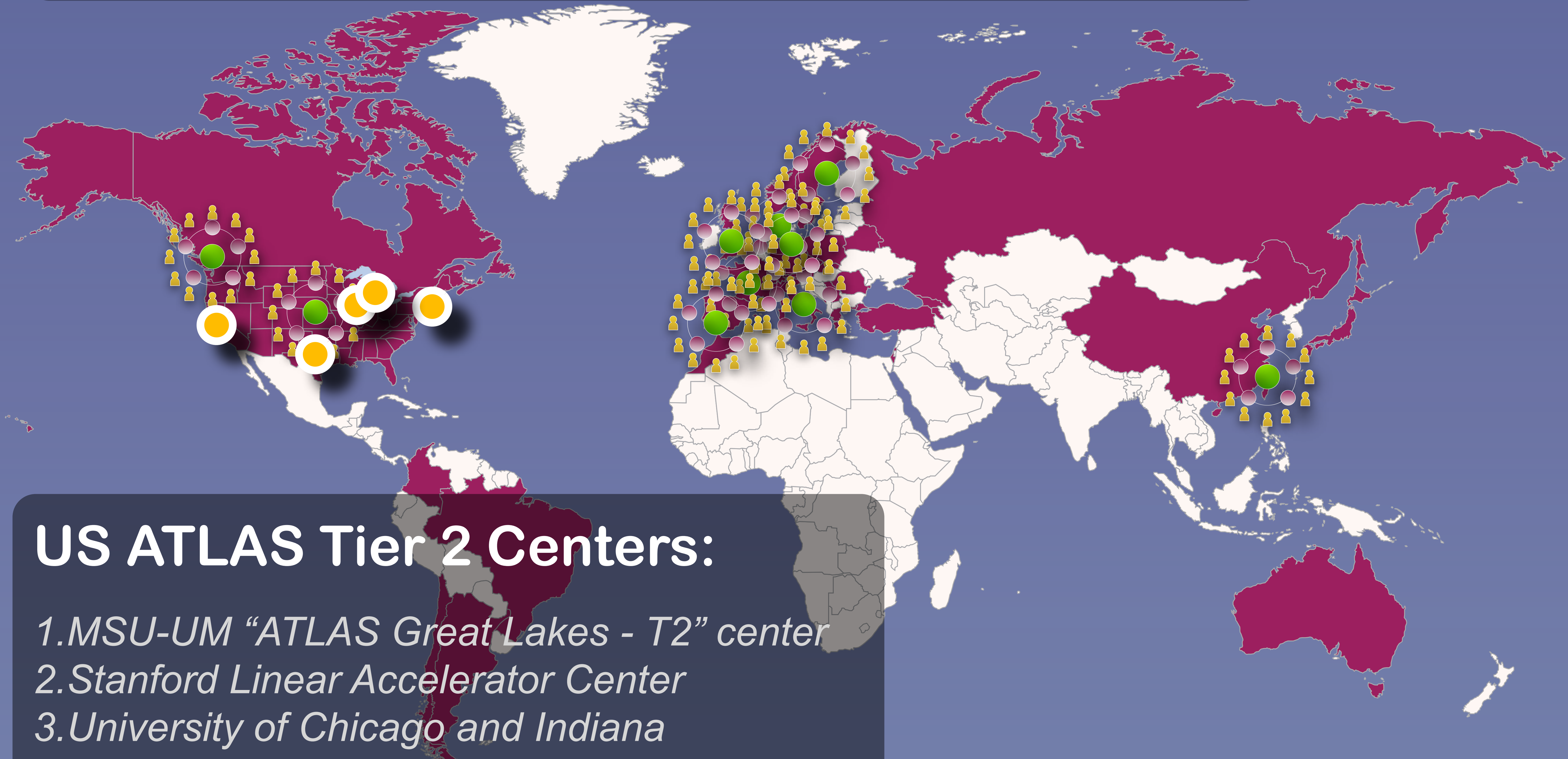


the world

the clouds

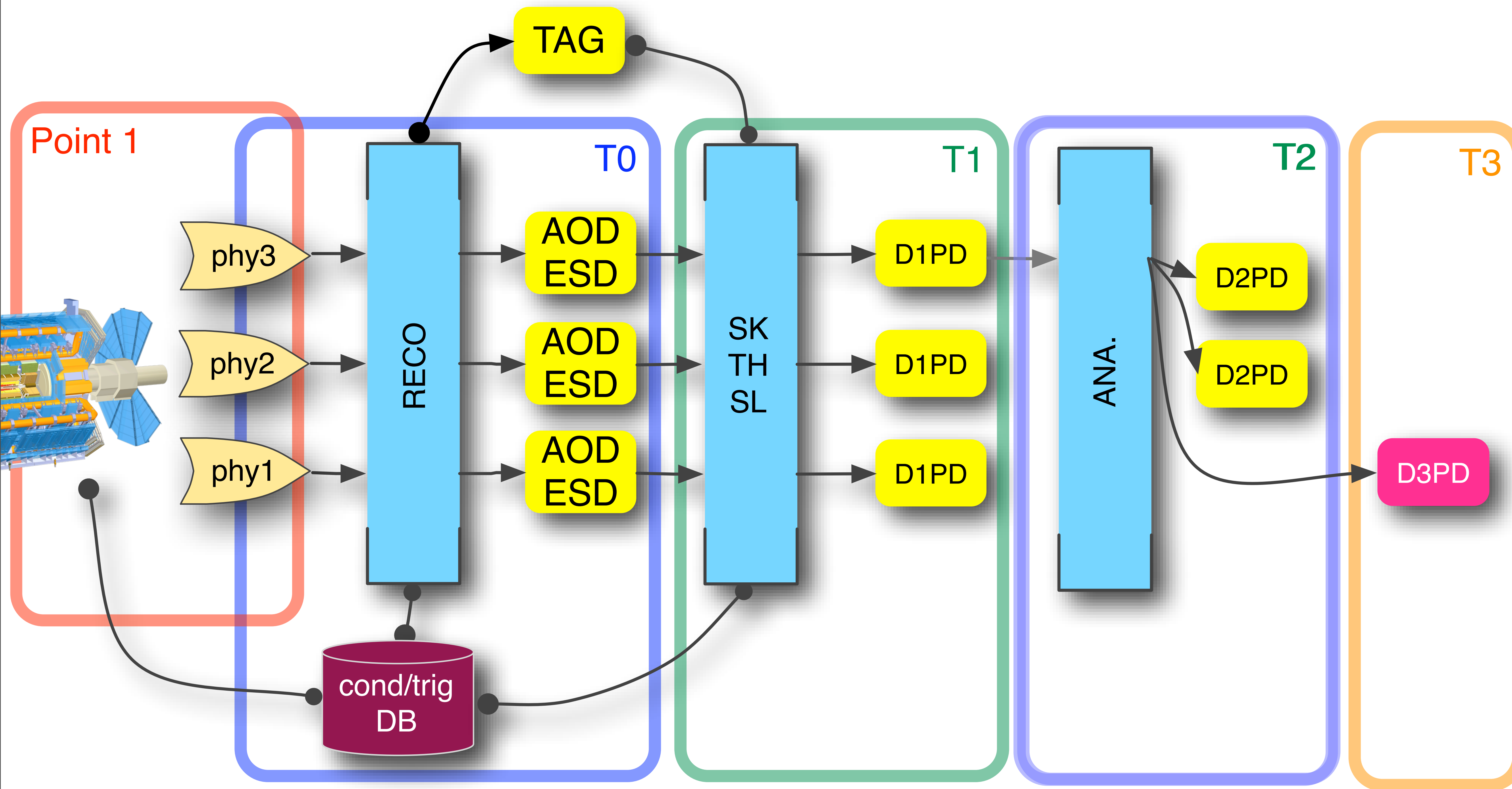
ATLAS computing clouds:

US, Canada, Britain, Spain, France, Italy,
Netherlands, Germany, Scandinavia, Taiwan

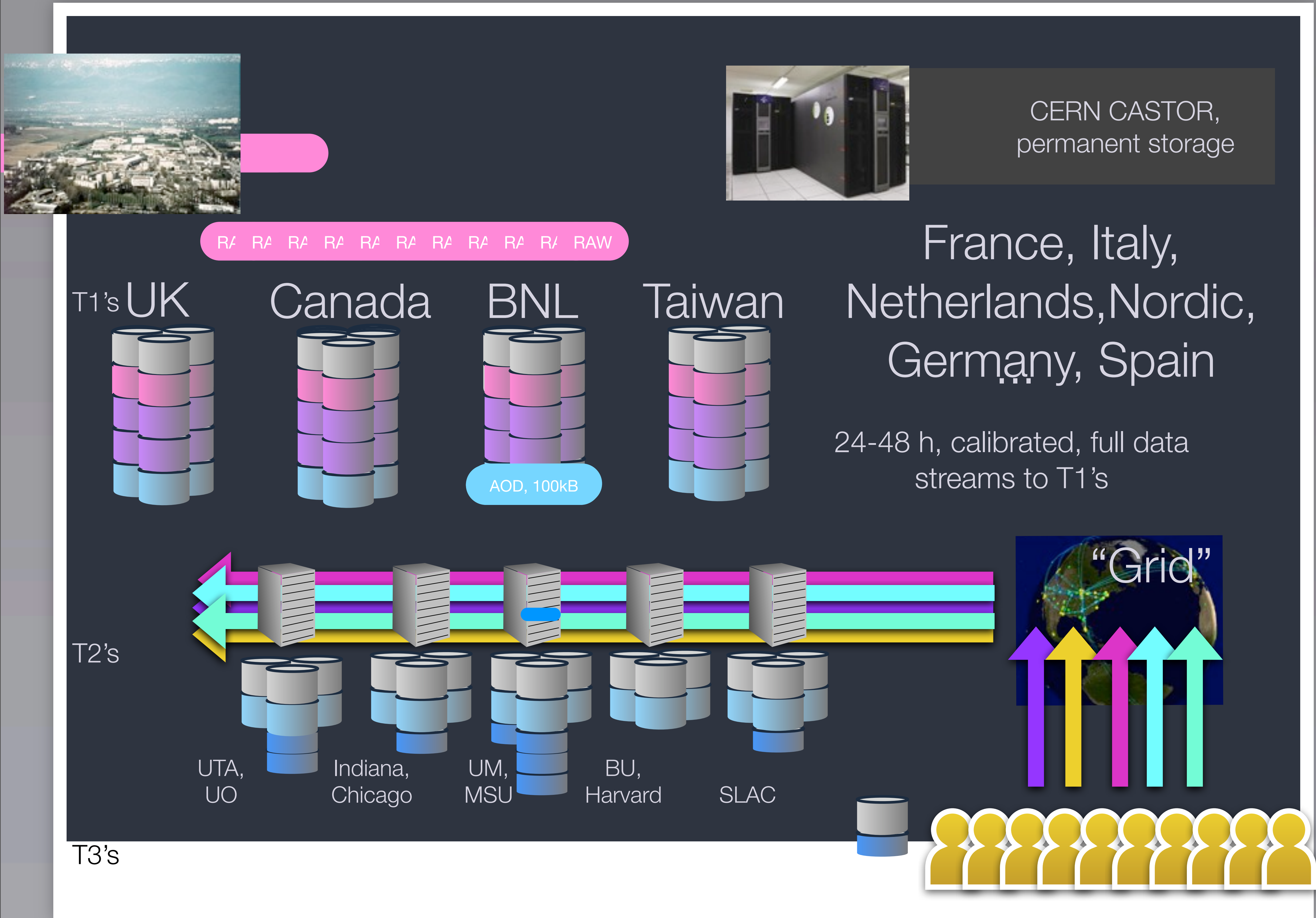


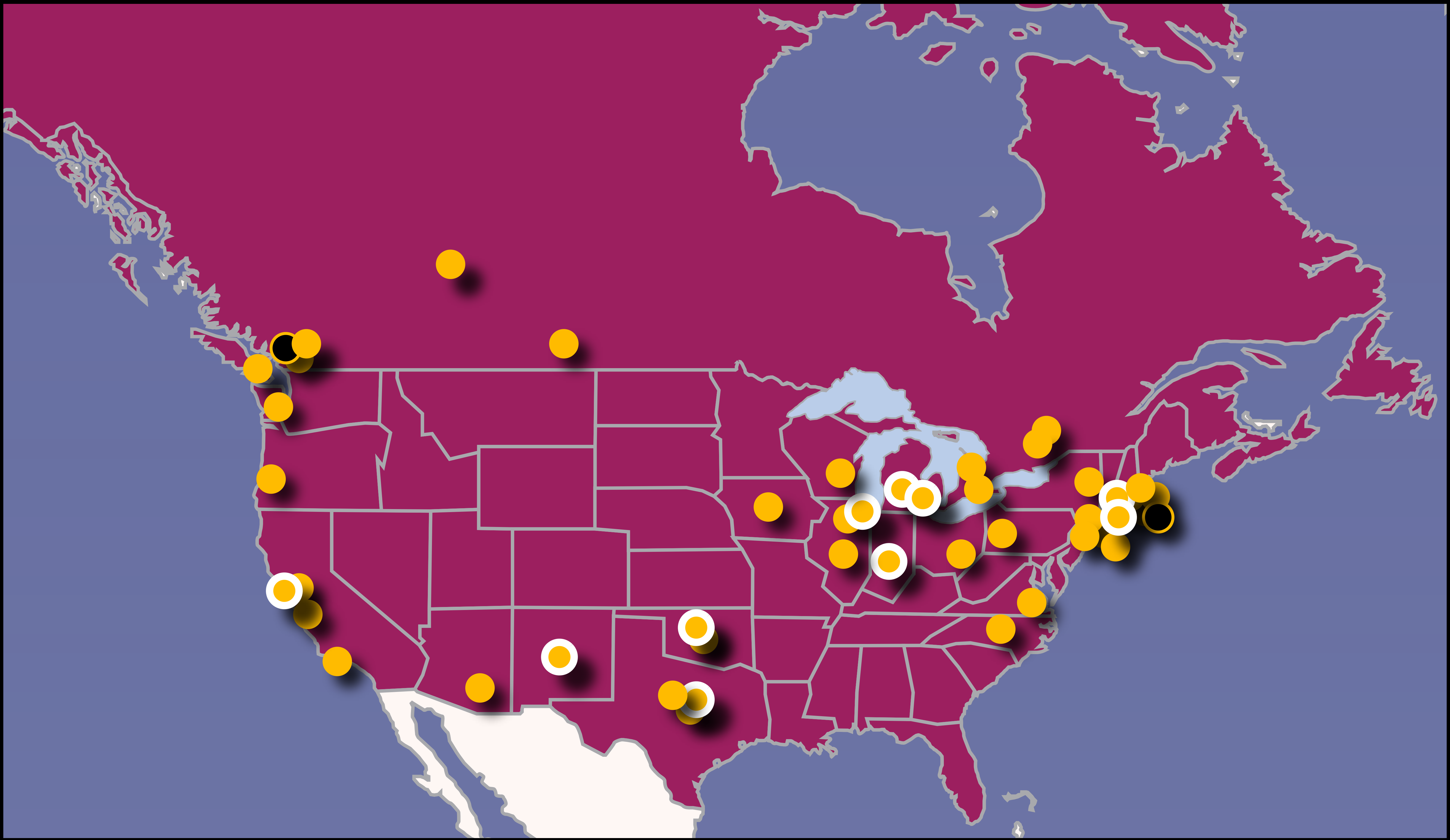
US ATLAS Tier 2 Centers:

1. *MSU-UM "ATLAS Great Lakes - T2" center*
2. *Stanford Linear Accelerator Center*
3. *University of Chicago and Indiana*
4. *Boston University and Harvard*
5. *University of Texas, Arlington and Oklahoma*



ATLAS data production chain

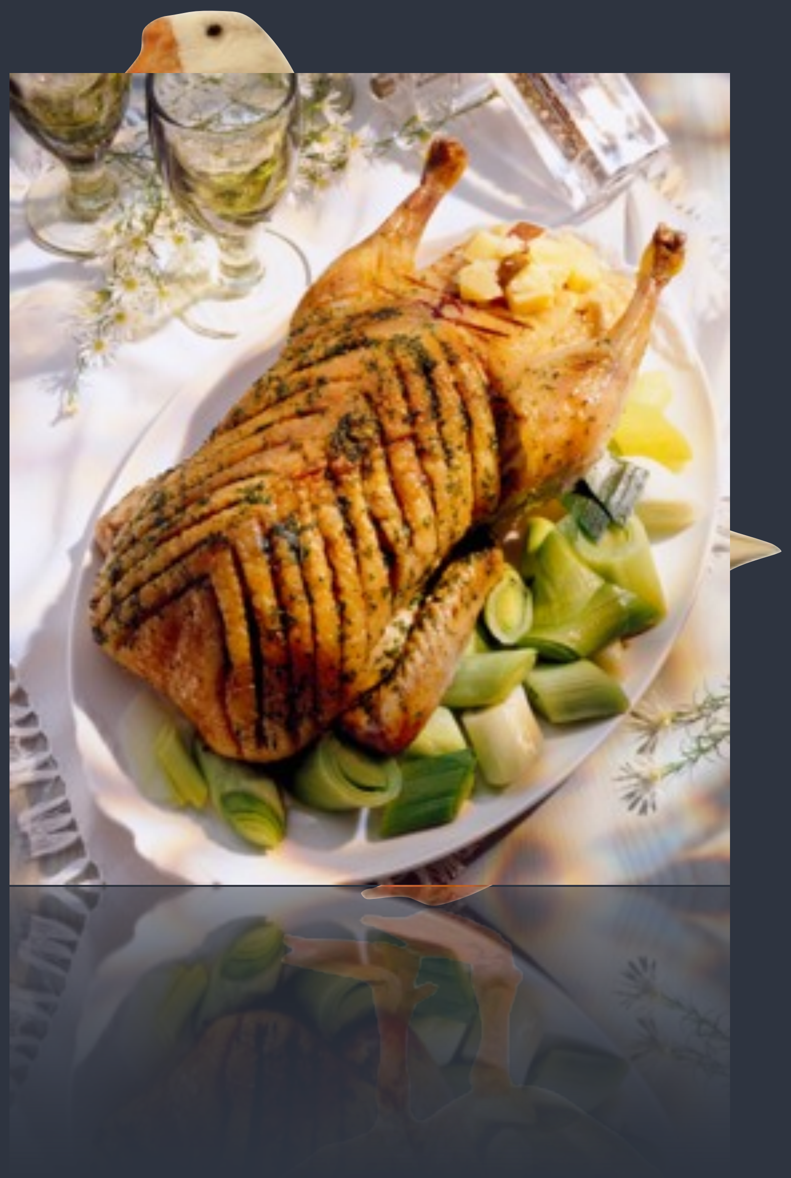

CERN CASTOR,
permanent storage
RAW
T1's
UK
Canada
BNL
Taiwan
France, Italy,
Netherlands,Nordic,
Germany, Spain
...
AOD, 100kB
24-48 h, calibrated, full data
streams to T1's
"Grid"
T2's
UTA,
UO
Indiana,
Chicago
UM,
MSU
BU,
Harvard
SLAC
T3's



We have to get this right

The Tier 3s?

not really an explicit part of the plan



3

“Tier 3 Task Force”

an enormous charge

for the first year of 10 fb^{-1} data (2011-2012?)

determine the value of an enhanced Tier 3 presence

are there different kinds of Tier 3s?

are the Tier 2 clusters insufficient?

characterize them, cost them, support models for them

ATLAS Tier 3 Task Force

Doug Benjamin, Duke University

Gustaaf Brooijmans, Columbia,

Sergei Chekanov, Argonne National Laboratory

Jim Cochran, Iowa State University,

Michael Ernst, Brookhaven National Laboratory,

Amir Farbin, University of Texas at Arlington,

Marco Mambelli, University of Chicago

Bruce Melado, University of Wisconsin,

Mark Neubauer, University of Illinois,

Flera Rizatdinova, Oklahoma State University,

Paul Tipton, Yale University,

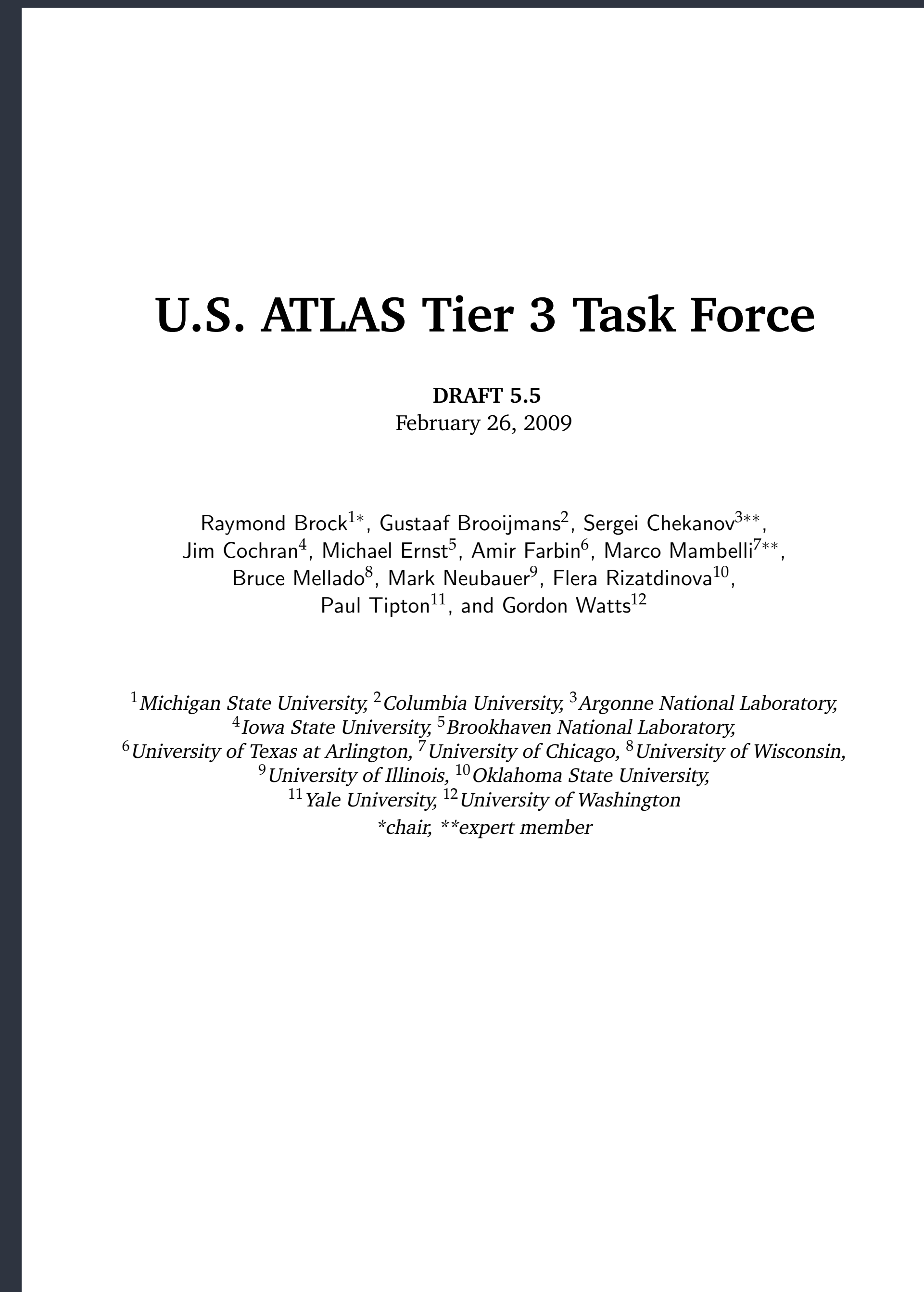
Gordon Watts, University of Washington,

Chip Brock, Michigan State University

the document

meant to be complete:

a reference



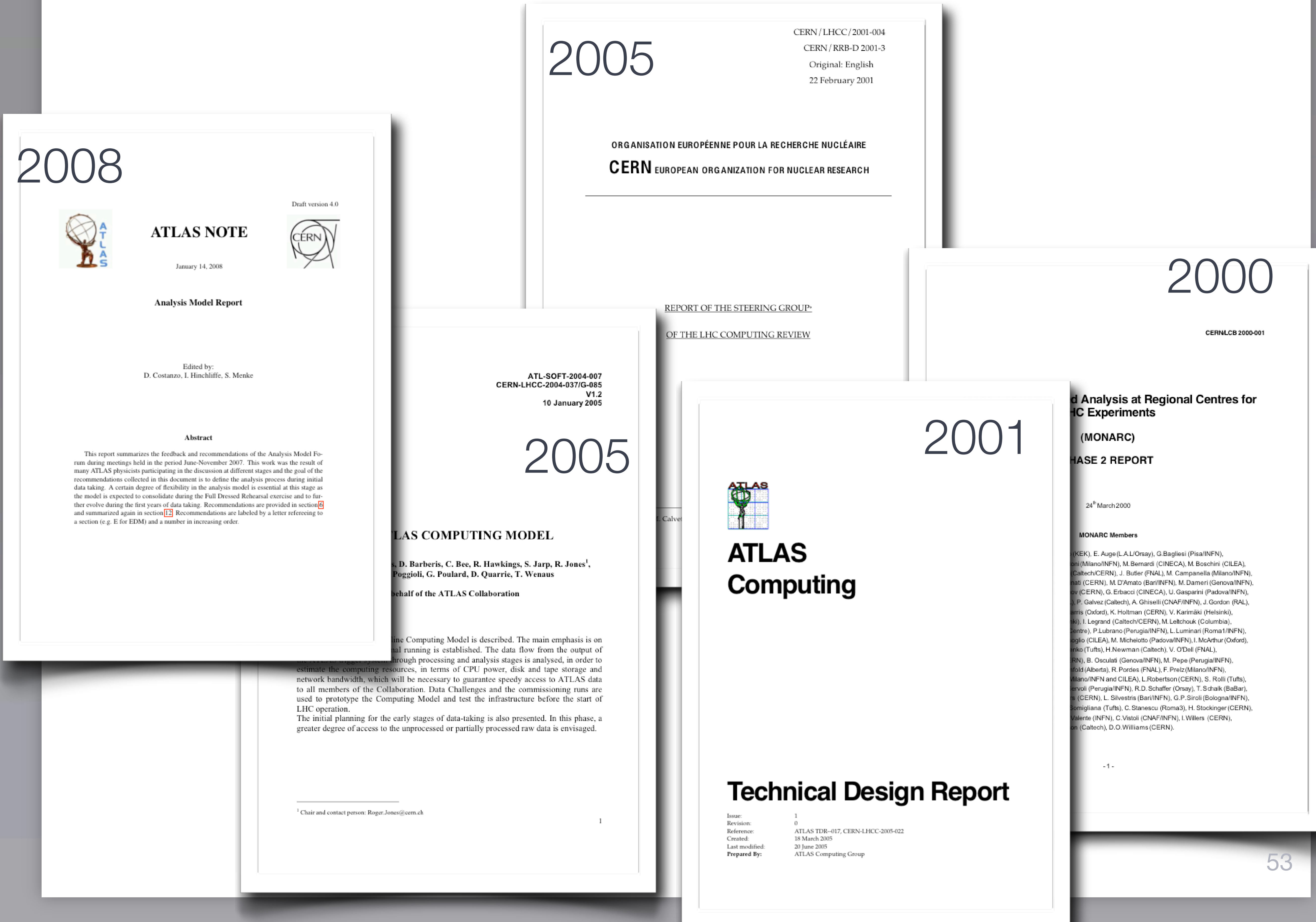
U.S. ATLAS Tier 3 Task Force

DRAFT 5.5
February 26, 2009

Raymond Brock[1*], Gustaaf Brooijmans[2], Sergei Chekanov[3**], Jim Cochran[4], Michael Ernst[5], Amir Farbin[6], Marco Mambelli[7**], Bruce Mellado[8], Mark Neubauer[9], Flera Rizatdinova[10], Paul Tipton[11], and Gordon Watts[12]

[1]Michigan State University, [2]Columbia University, [3]Argonne National Laboratory, [4]Iowa State University, [5]Brookhaven National Laboratory, [6]University of Texas at Arlington, [7]University of Chicago, [8]University of Wisconsin, [9]University of Illinois, [10]Oklahoma State University, [11]Yale University, [12]University of Washington
**chair, **expert member*

www.pa.msu.edu/~brock/file_sharing/T3TaskForce/final/TierThree_v1_executiveFinal.pdf


2005
CERN / LHCC / 2001-004
CERN / RRB-D 2001-3
Original: English
22 February 2001
ORGANISATION EUROPÉENNE POUR LA RECHERCHE NUCLÉAIRE
CERN EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH
REPORT OF THE STEERING GROUP
OF THE LHC COMPUTING REVIEW
2008
Draft version 4.0
ATLAS NOTE
January 14, 2008
Analysis Model Report
Edited by:
D. Costanzo, I. Hinchliffe, S. Menke
Abstract
This report summarizes the feedback and recommendations of the Analysis Model Forum during meetings held in the period June-November 2007. This work was the result of many ATLAS physicists participating in the discussion at different stages and the goal of the recommendations collected in this document is to define the analysis process during initial data taking. A certain degree of flexibility in the analysis model is essential at this stage as the model is expected to consolidate during the Full Dressed Rehearsal exercise and to further evolve during the first years of data taking. Recommendations are provided in section 6 and summarized again in section 12. Recommendations are labeled by a letter refereeing to a section (e.g. E for EDM) and a number in increasing order.
ATL-SOFT-2004-007
CERN-LHCC-2004-037/G-085
V1.2
10 January 2005
2005
LAS COMPUTING MODEL
, D. Barberis, C. Bee, R. Hawkings, S. Jarp, R. Jones[1],
Poggioli, G. Poulard, D. Quarrie, T. Wenaus
behalf of the ATLAS Collaboration
ine Computing Model is described. The main emphasis is on
al running is established. The data flow from the output of
hrough processing and analysis stages is analysed, in order to
estimate the computing resources, in terms of CPU power, disk and tape storage and network bandwidth, which will be necessary to guarantee speedy access to ATLAS data to all members of the Collaboration. Data Challenges and the commissioning runs are used to prototype the Computing Model and test the infrastructure before the start of LHC operation.
The initial planning for the early stages of data-taking is also presented. In this phase, a greater degree of access to the unprocessed or partially processed raw data is envisaged.
[1] Chair and contact person: Roger.Jones@cern.ch
1
2000
CERN/LCB 2000-001
d Analysis at Regional Centres for
IC Experiments
(MONARC)
HASE 2 REPORT
24th March 2000
MONARC Members
- 1 -
2001
ATLAS
ATLAS
Computing
Technical Design Report
Issue: 1
Revision: 0
Reference: ATLAS TDR–017, CERN-LHCC-2005-022
Created: 18 March 2005
Last modified: 20 June 2005
Prepared By: ATLAS Computing Group

workflow

Steady State Dataset Distribution

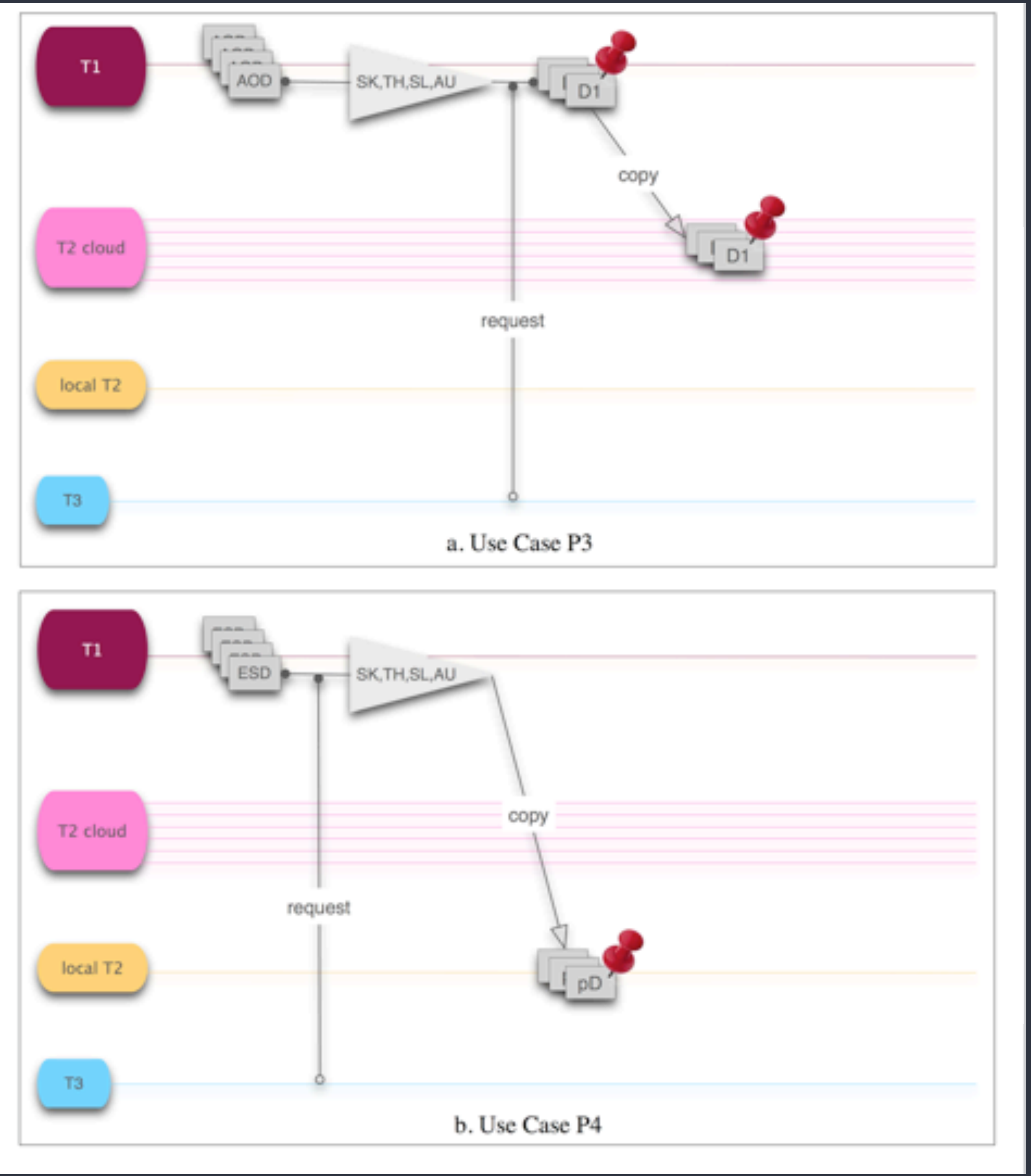
Dataset creation

Monte Carlo Production

“Chaotic” User Analysis (“Chaotic User” Analysis?)

Intensive Computing Tasks

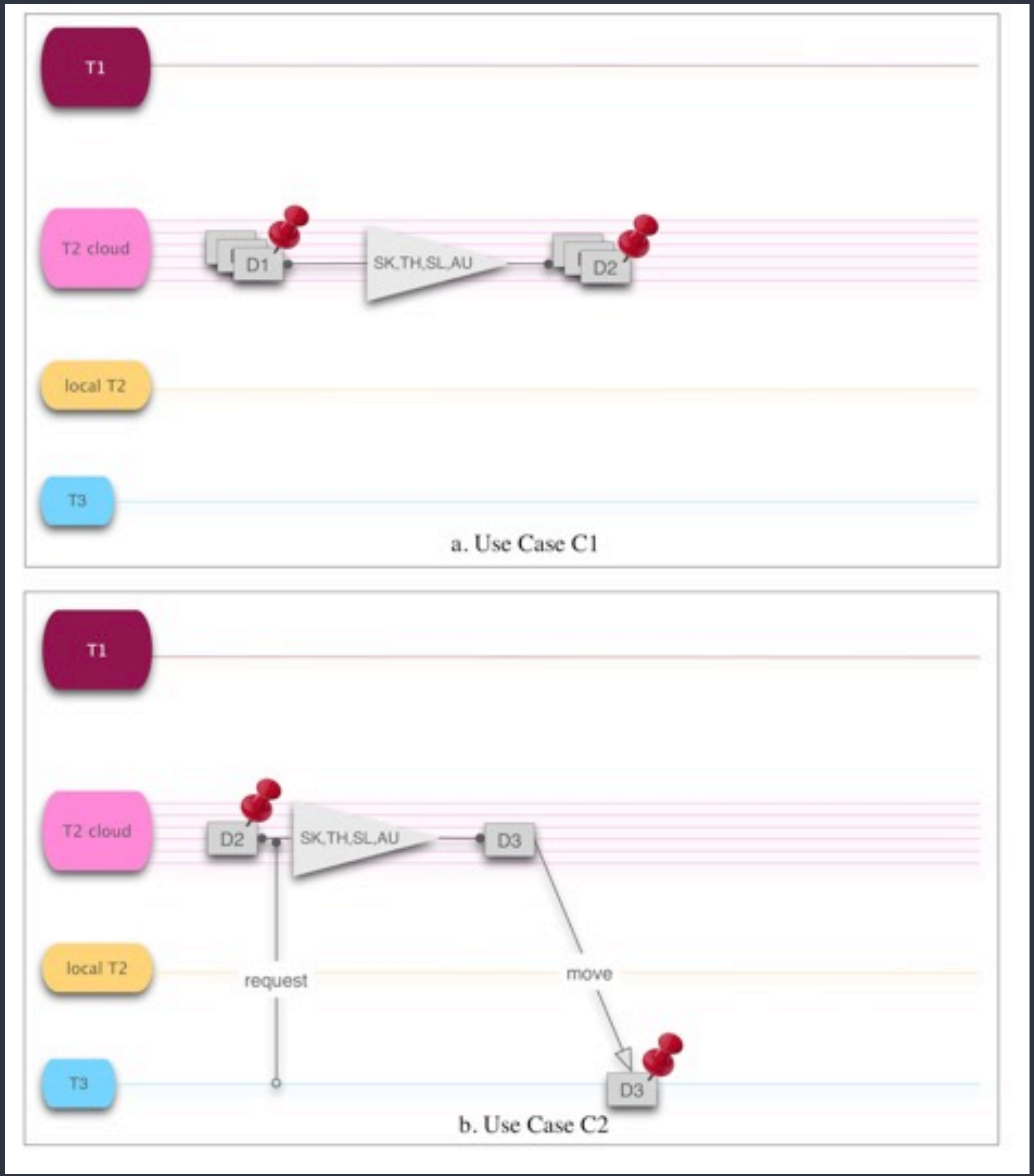
Steady State Data Distribution

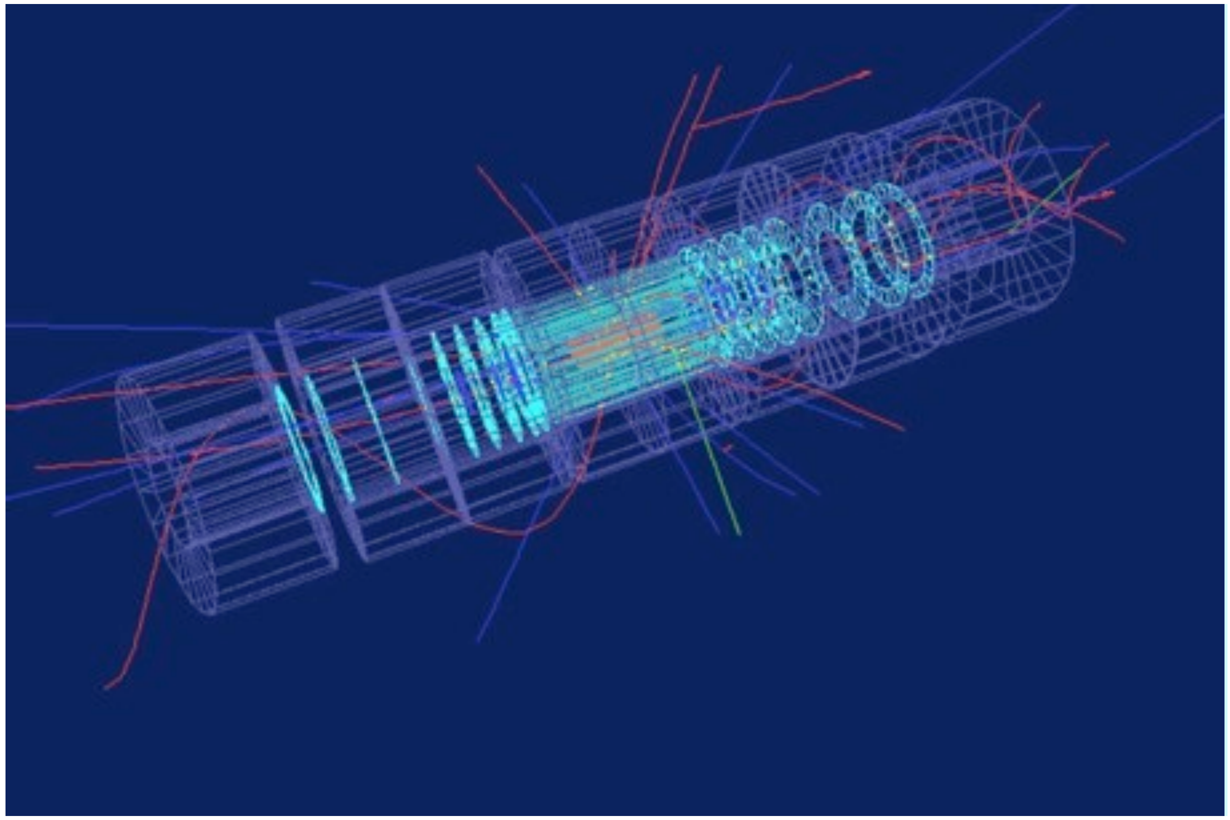


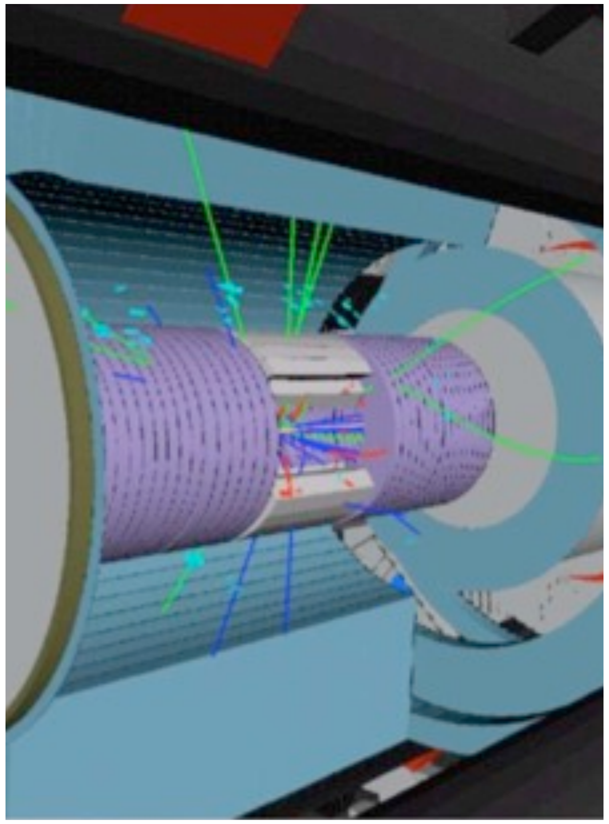
dataset creation

D1PD→D2PD:

not entirely determined







Detector Simulation

computer representation of each detector element

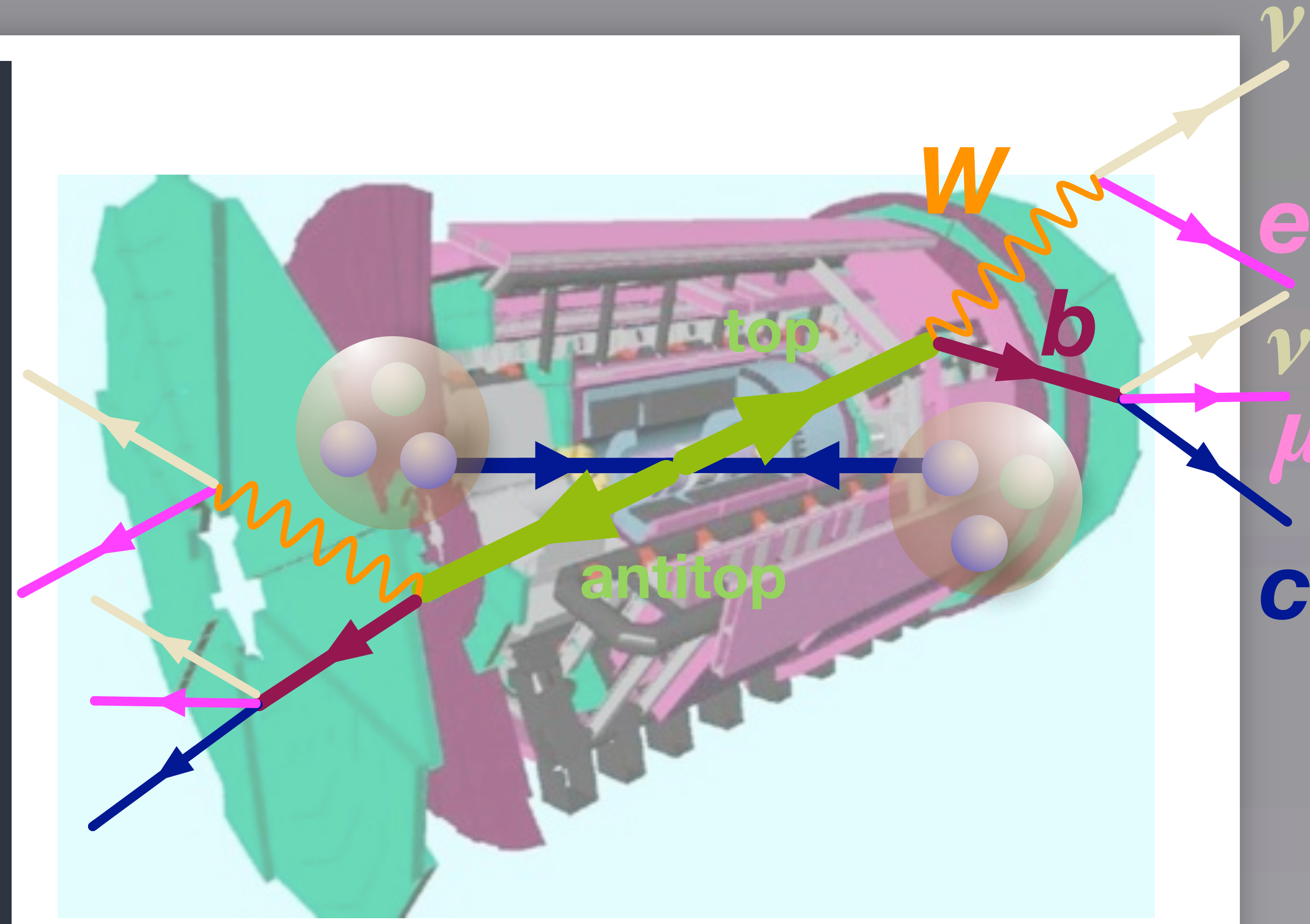
realistic propagation of particles

computationally expensive

Generation

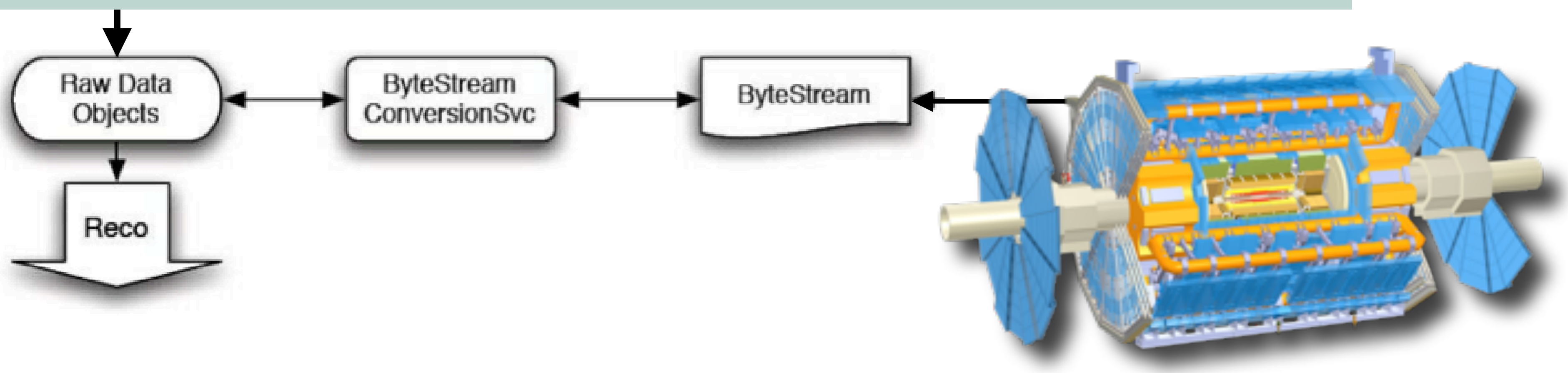
Simulation

Digitization



Sample	Generation	Simulation	Digitization
Minimum Bias	0.0267	551.	19.6
$t\bar{t}$ Production	0.226	1990	29.1
Jets	0.0457	2640	29.2
Photon and jets	0.0431	2850	25.3
$W^{\pm} \to e^{\pm}\nu_e$	0.0788	1150	23.5
$W^{\pm} \to \mu^{\pm}\nu_{\mu}$	0.0768	1030	23.1
Heavy ion	2.08	56,000	267

kSI2k-s !



Simulations look like data

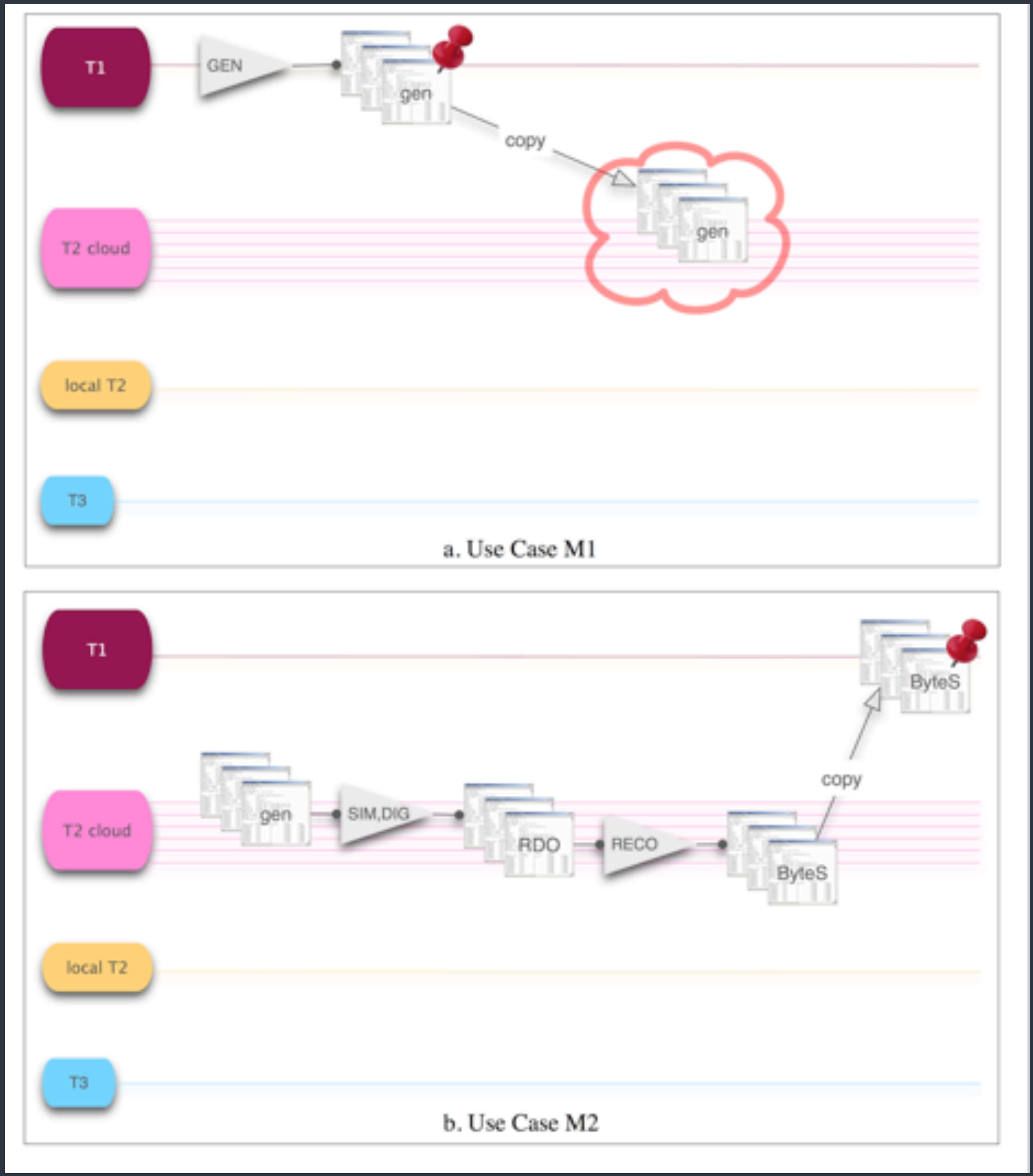
Monte Carlo production

Generation: T1

Simulation: T2

Digitization: T2

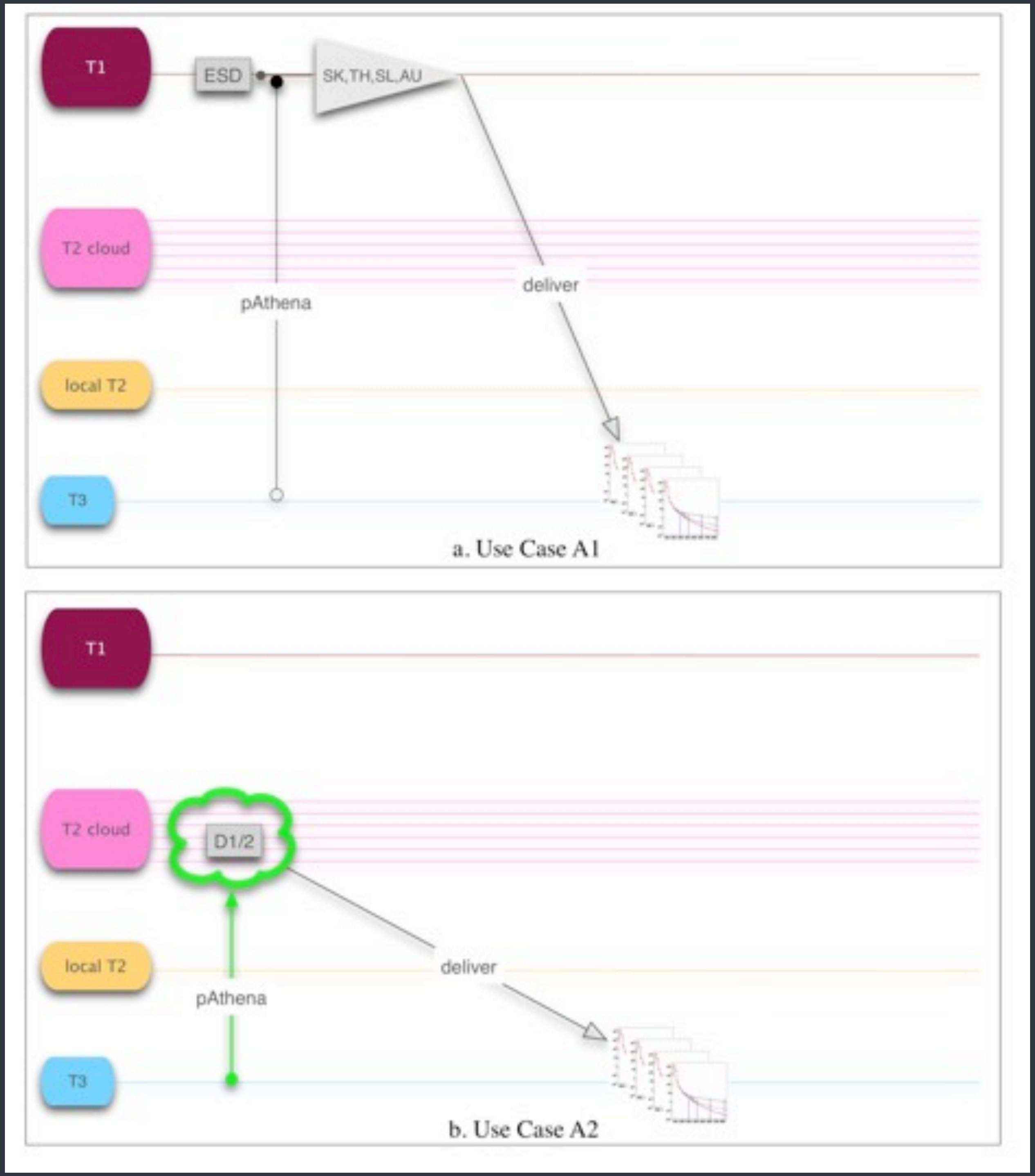
Reconstruction: T2



Chaotic User Analysis

human-intensive

data-handling



What are your guides?

history is our only source of data

scientific computing planning is hard

Administrators

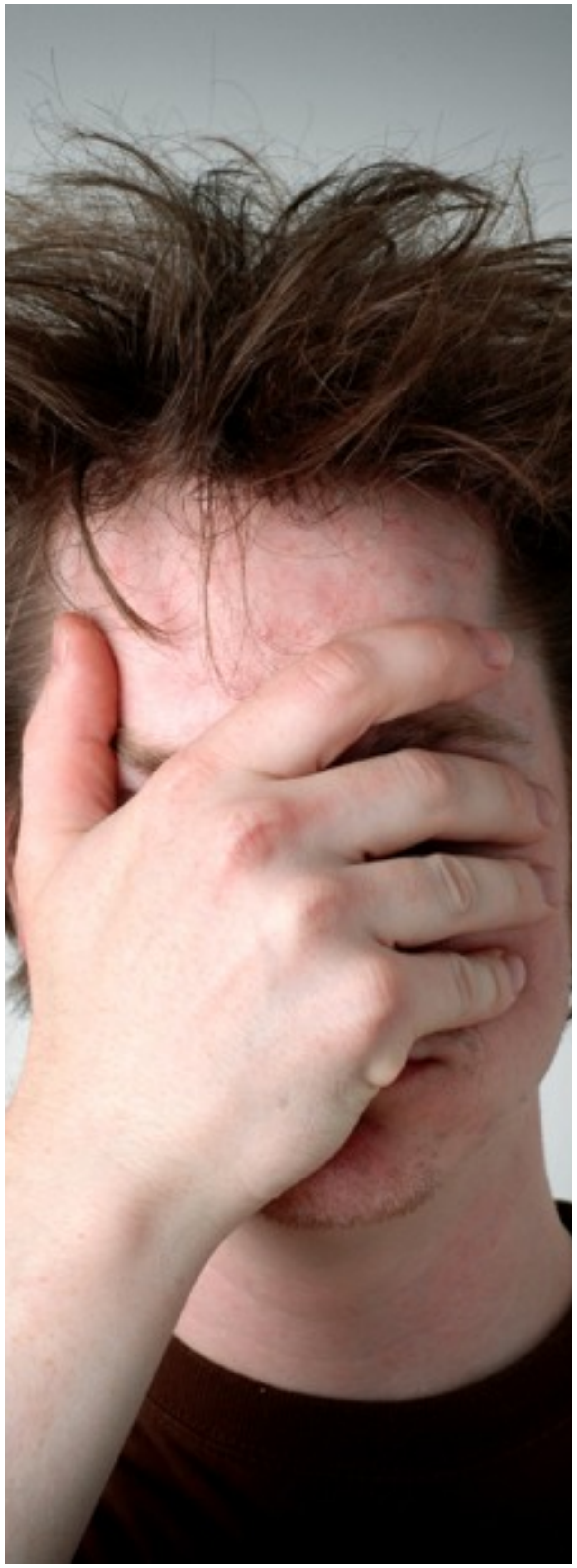
argue for funds against a plan

Users–have one thing in mind

not great about sticking to a plan

Physics analysis moves

faster than the best computing plans.



history = Fermilab tevatron

DØ and CDF: re-invented computing models many times

emerging technologies

made unanticipated, clever analyses possible

unanticipated, clever analyses

made extending technologies essential

neither of these are necessarily consistent with tight resource planning

the world changed many times in the lifetime of the Tevatron

1. ubiquity of OO coding

2. emergence of inexpensive, commodity computer clusters

3. availability of distributed disk servers and management systems

4. development of high-speed networking and switching technologies

5. the Web, from cute to essential

an example:

“Matrix Element” calculations

many cpu-centuries of computation

grid has failed DØ for these

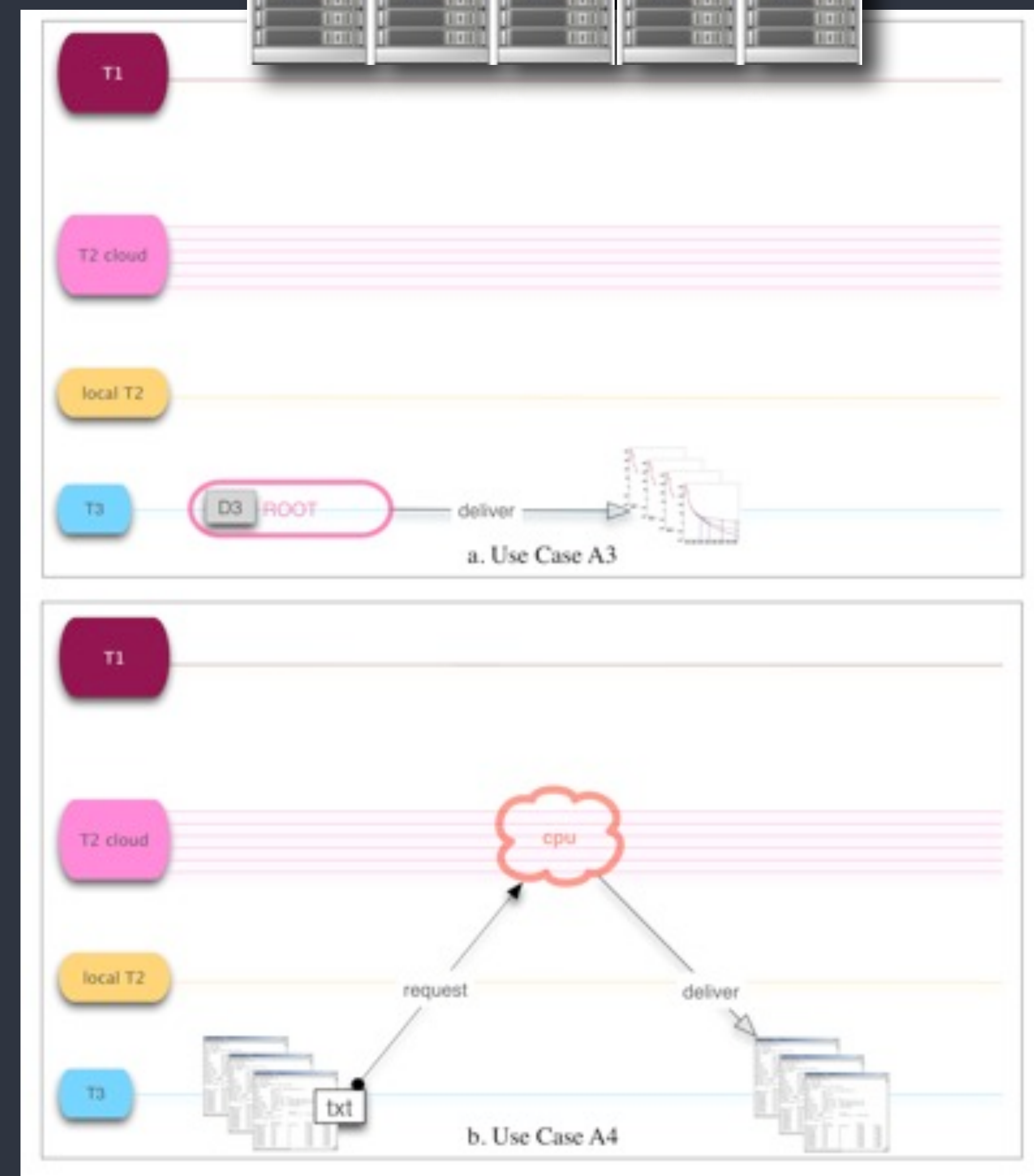
Multivariate combinations

COLLIE

Ensemble simulation

~O in

~O out



About these intensive computational methods:

this is important:

Nobody had ever dreamed of these sorts of analysis tasks before this century

What kinds of surprises will the ATLAS era see?

another example: single top quark analyses: intense

- **A DØ analysis**

  about once per month

  before systematic error studies

  before “editorial board” demands

  just one analysis

source	files	events	jobs
data	96k	1600M	2400
QCD background	96k	1600M	2400
signal MC	25.6k	200M	2400
bckgnd MC	12k	120M	560
total	240k	3B	8000

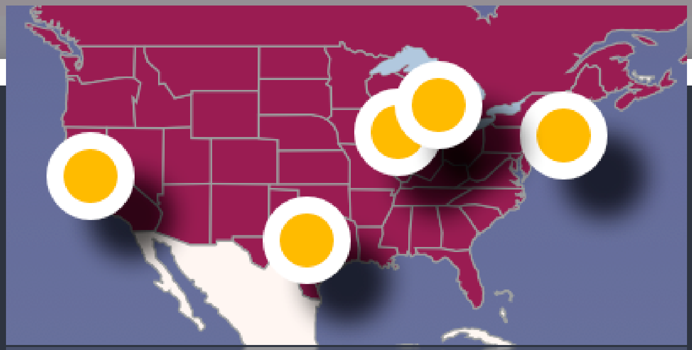
prediction is hard

"I believe OS/2 is destined to be the most important operating system, and possibly program, of all time."

Bill Gates, OS/2 Programmers Guide, November 1987

	1997 projections	2006 actual
Peak (average) data rate (Hz)	50 (20)	100(35)
Events collected	600M/year	1500M/year
Raw Data Size (kB.event)	250	250
Reconstructed Data size(kB/event)	100	80
User format (kB/event)	1	40
Tape Storage	280 TB/year	1.6 PB on tape
Tape reads/writes (weekly)		30 TB/7TB
Analysis/cache disk	7 TB/year	220 TB
Reconstruction time (GHz-s/event)	2.0	50
User analysis times (GHz-s/event)	?	1
User analysis weekly reads	?	3B events
Primary reconstruction farm size (THz)	0.6	2.4 THz
Central analysis farm size (GHz)	0.6	2.2 THz
Remote resources (GHz)	?	~ 2.5THz
	after Run 1	**after Run 2a**

Tier 2/3 Modeling



Tier 2's are the heroes of ATLAS

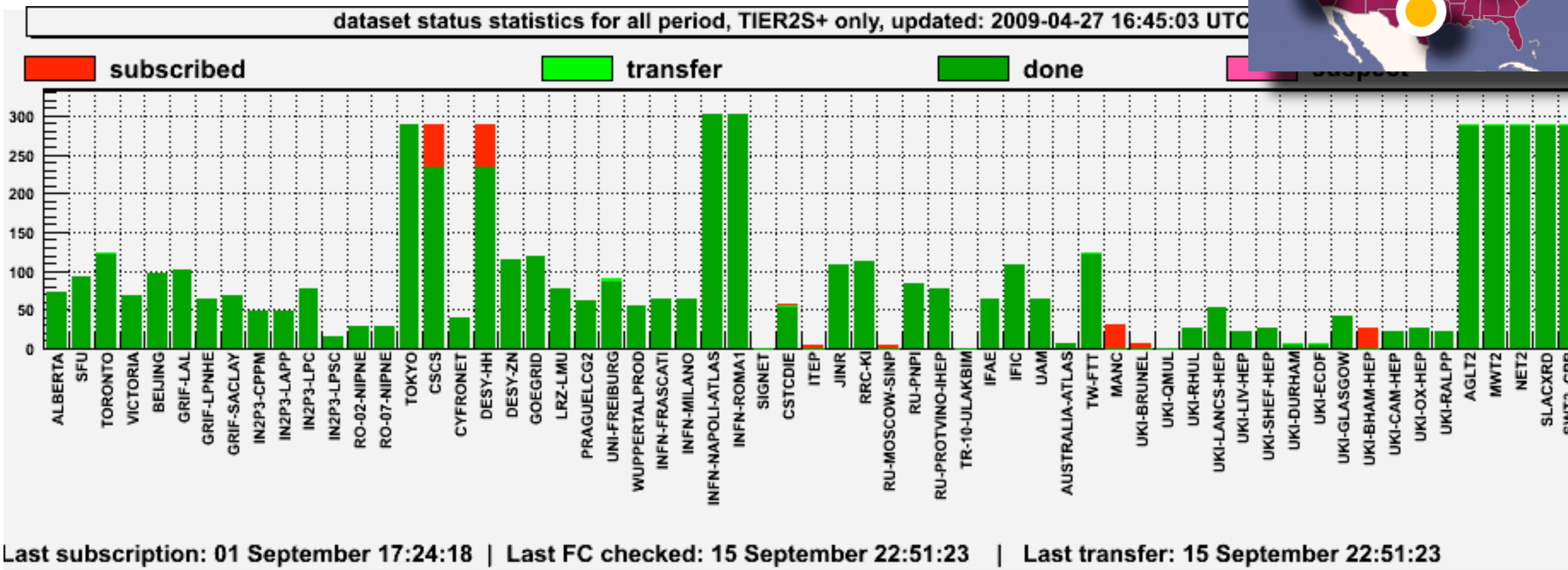
But:

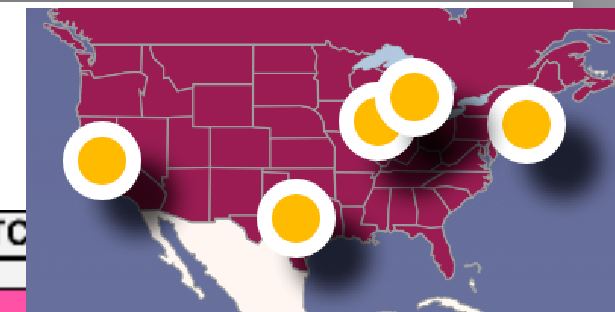
Are they physicist-innovation-capable?

Can they really handle the sort of human-intense load that will be likely?

Will physicists still try to move data near to them?

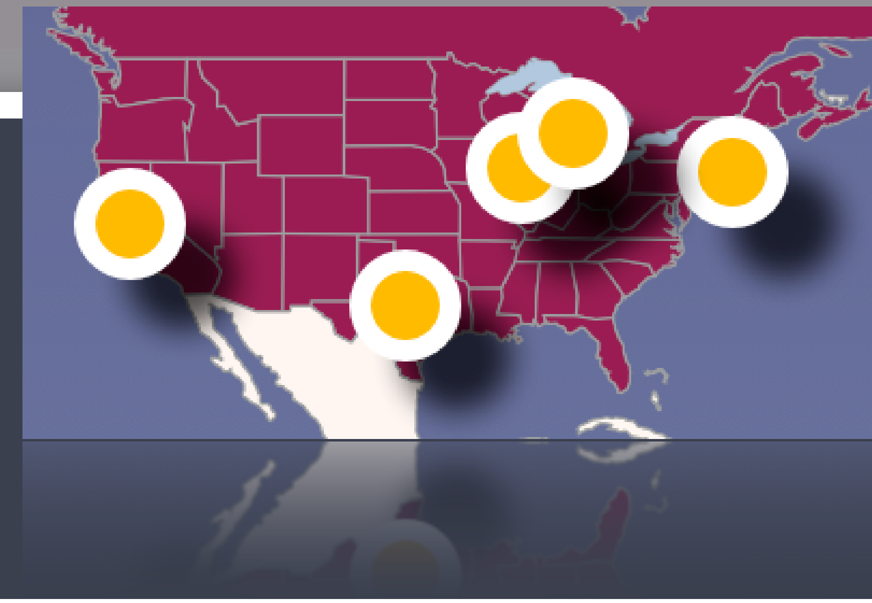
They are busy...will they be available?





all ATLAS Tier 2s (September)

Tier 2 resources



- **50%,**

  centrally managed for simulation

- **50%**

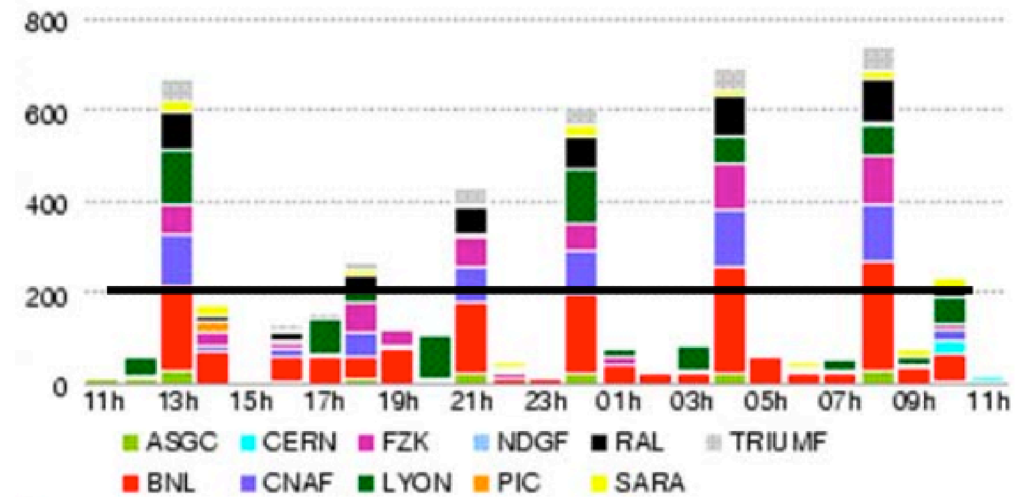
  for national analyses

- **How much full simulation?**

  How much fast simulation?

US Pledge to wLCG	2007	2008	2009	2010	2011
CPU (kSI2k)	2,560	4,844	7,337	12,765	18,194
Disk (TB)	1,000	3,136	5,822	11,637	16,509
Tape (TB)	603	1,715	3,277	6,286	9,820

MBps

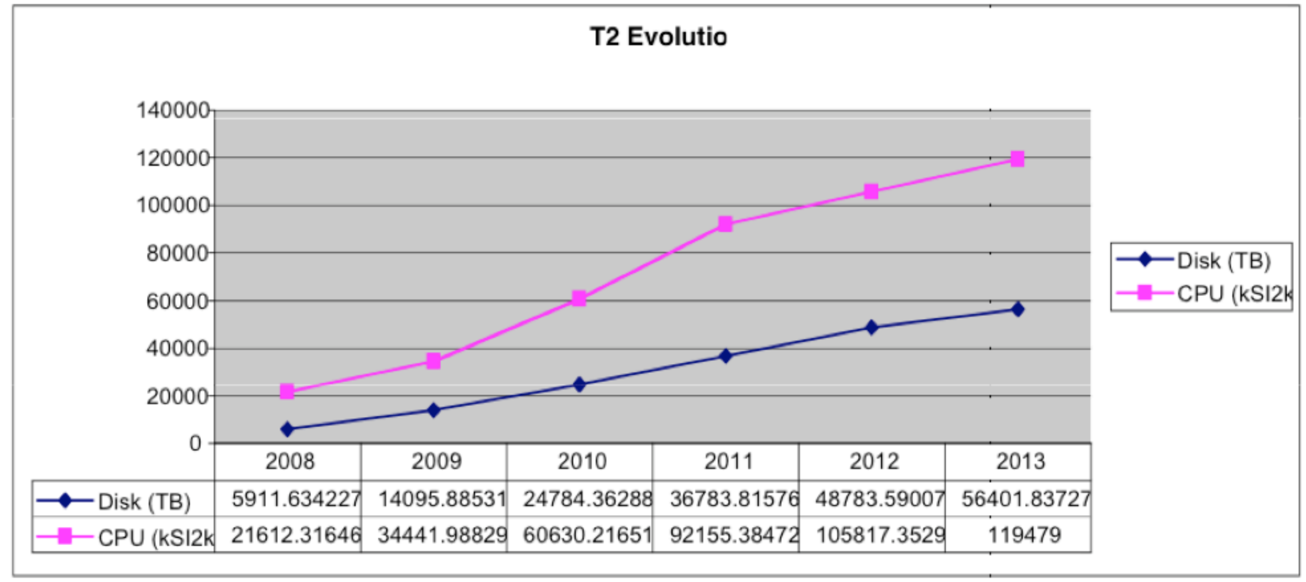


Mean target

T1-> T2

Benchmark: 10fb^{-1} → 2011, 2012?

quantity	value used	high	low	comments
LHC year	2010	2011	n.a.	assume 2008 start
Ins. $\mathcal{L}$ cm^{-2}s^{-1}	2×10^{33}	3.5×10^{33}	10^{33}	Garoby, LHCC 08
annual $\int \mathcal{L}dt$ fb^{-1}	10	?	?	rounded from 12
annual dataset	2×10^9 events	?	?	[7]
sim. time	1990 kSI2K s ($t\bar{t}$)	2850 kSI2K s γj	1030 kSI2K s $W \to \mu$	[16]
dig. time	29.1 kSI2K s ($t\bar{t}$)	29.2 kSI2K s j	23.1kSI2K s $W \to \mu$	[16]
reco. time	47.4 kSI2K s ($t\bar{t}$)	78.4 kSI2K s j	8.07 kSI2K s $W \to e$	[16]
digitization pileup factor	3.5	5.8	2.3	[16]
fraction of full dataset for full sim	0.1	0.2	na.	
factor rel. to full sim. for $t\bar{t}$	0.05 (ATLFAST-II)	0.38 (fG4)	0.004 (ATLFAST-IIF)	[16]
D^{1}PD → D^{2}PD	0.5 kSI2K s	?	?	[15]
D^{2}PD → D^{3}PD	0.5 kSI2K s	?	?	[15]
disk R/W	100 MBps	200 MBps	10 MBps	S. McKee private
sustained network	50 MBps	100 MBps	10 MBps	S. McKee private
fraction of data in pDPD	20%			
# primary DPD	10			
# subgroups	5			
average CPU	1.4 kSI2K units	2	NA	
total ATLAS Tier 2 computing	60.63MSI2k			[11]



modeled it.

Amir Farbin

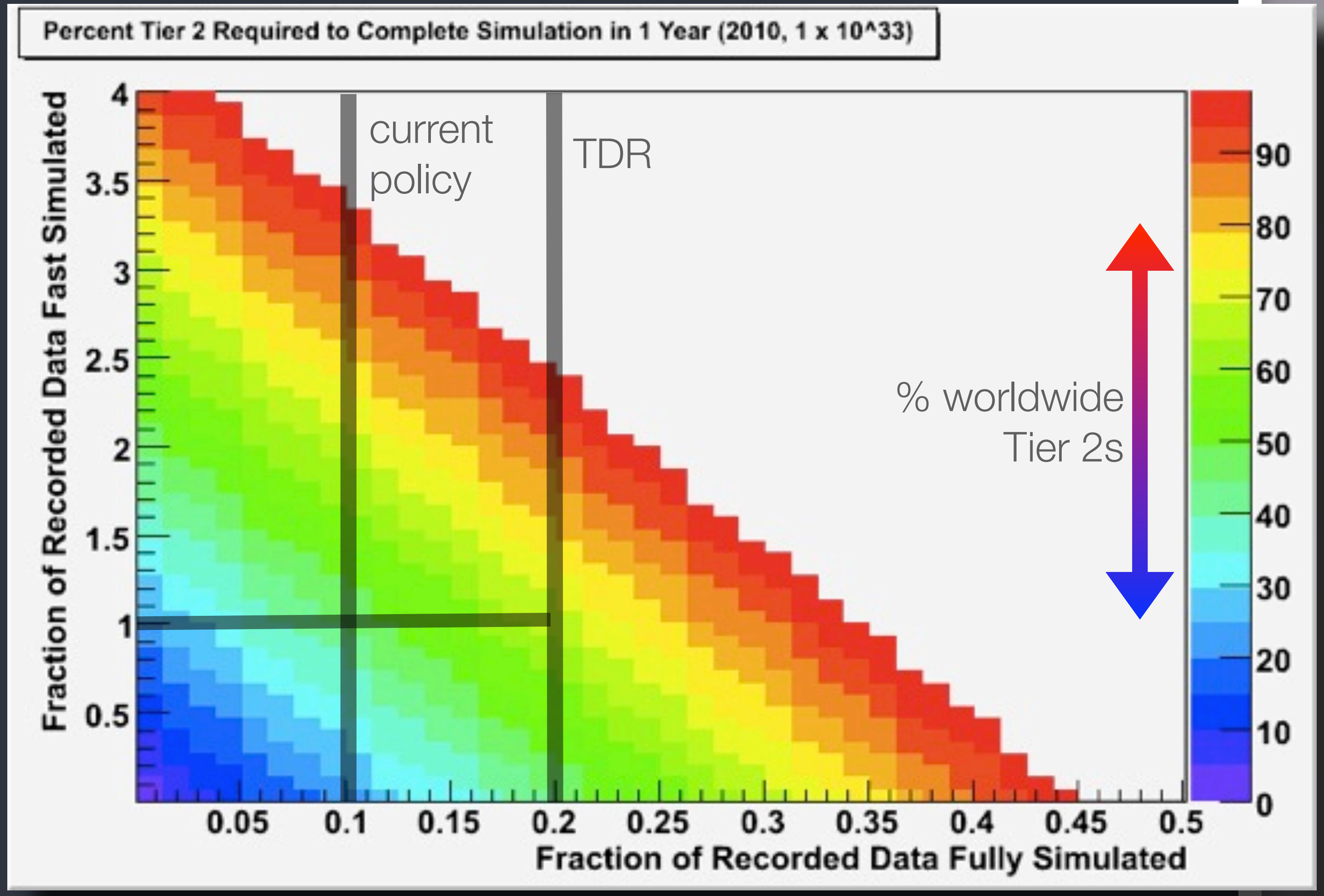
Tier 2 simulation for one year

- **horizontal axis:**

  fraction fully simulated

- **vertical axis:**

  fraction fast-simulated



Monday, April 27, 2009

a computing model restricted to Tier 2s seems like a risk:

1. **The Tier 2s may become overloaded.**

2. **History tells us to expect the unexpected.**

3. ***...stuff will happen.***

flexible and nimble

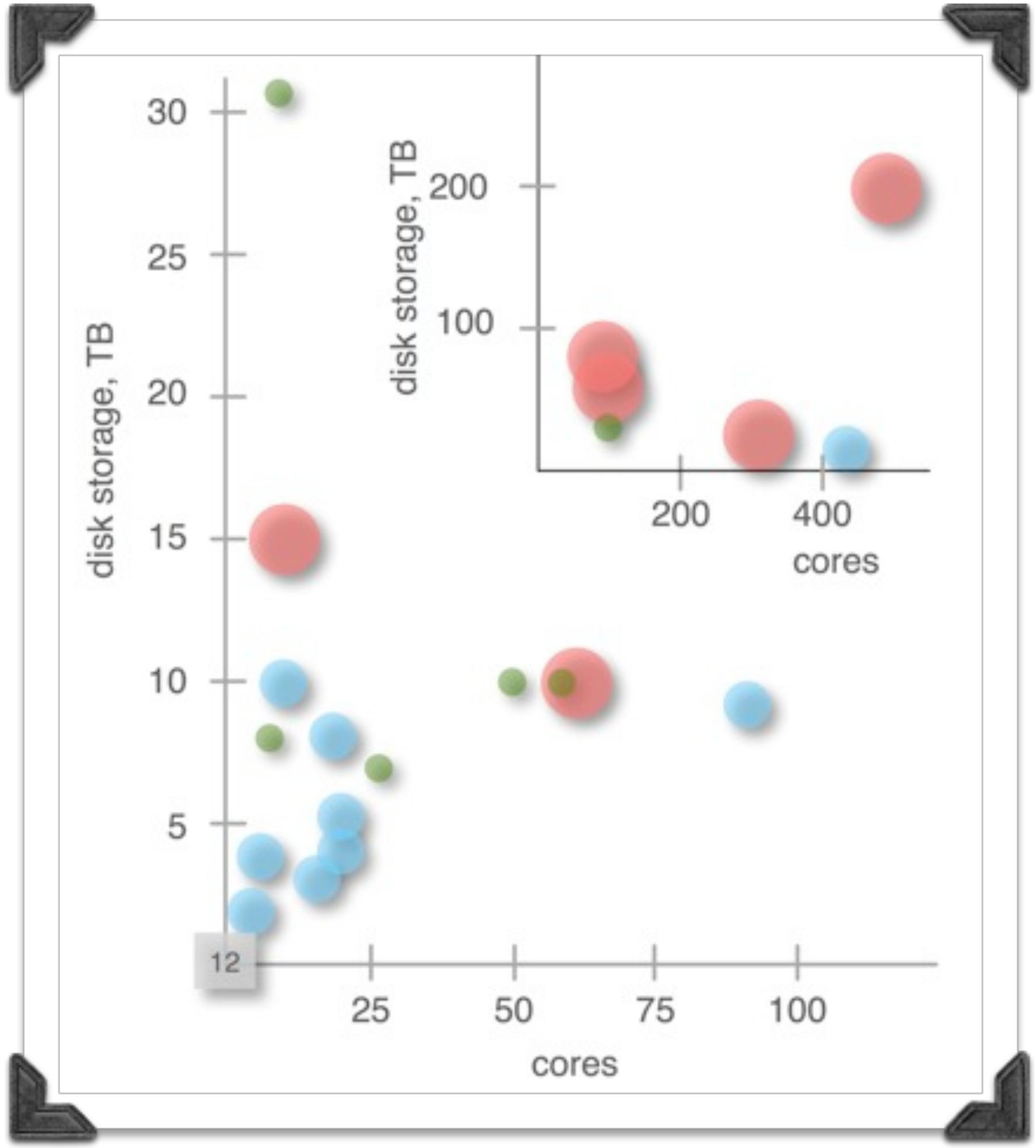
We have to expand our model to include the Tier 3 component.

Tier 3s today.

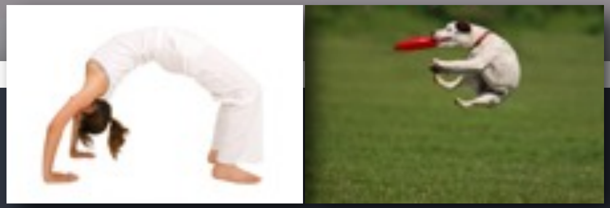
Survey:

all but 6 ATLAS institutes

not much.



dot size/color: network connectivity:
100Mbps, **1Gbps**, **10Gbps**



4 Primary Recommendations

Minimum requirements for US ATLAS university computing:

1. *Recommended 4 defined classes of “Tier 3” centers*
2. *Recommended modifications to the ATLAS data management scheme*
3. *Recommended “human scale” data transfer capabilities*
4. *Recommended ATLAS technical support position*

Remember:
benchmarking a “10fb^{-1} physics year”...2011-2012

“Tier 3 Quartet”

4 classes of Tier 3 centers

each with distinct capabilities

each costed

use cases defined for each

T3gs

Tier 3 with "grid services"

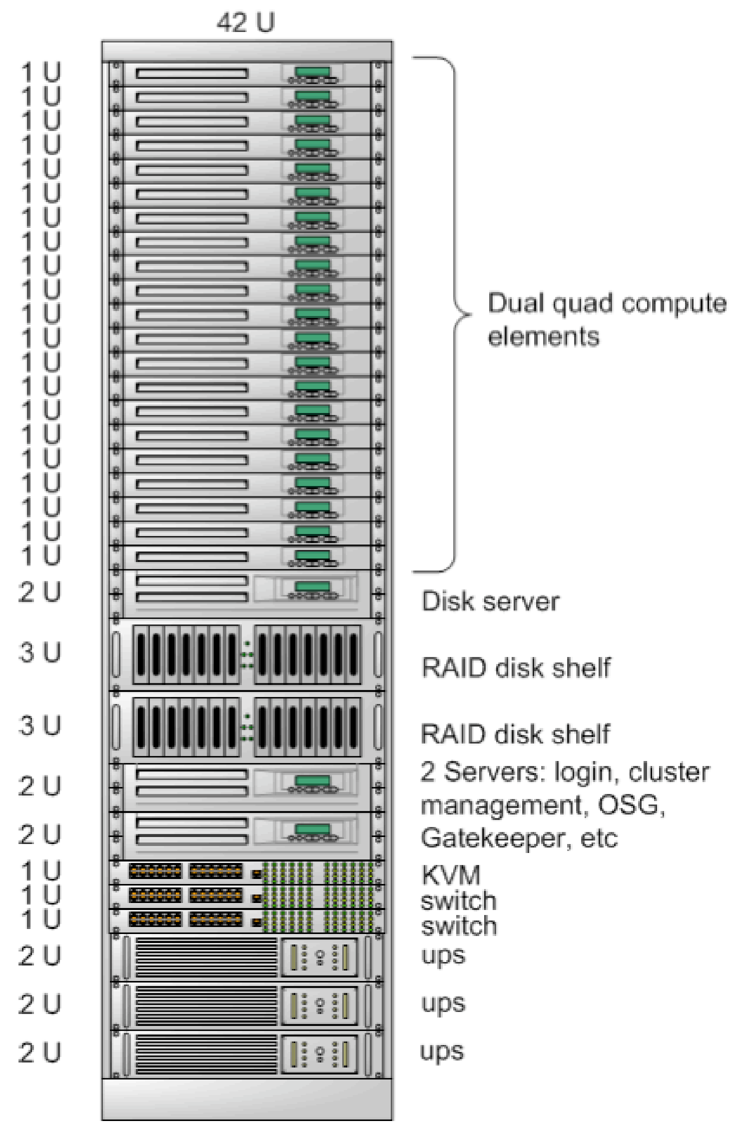
a campus-based, significant cluster

requiring AC/power infrastructure

Characterized a strawman

~$80k

University of Illinois building one



168 processors
>250kSI2k
24TB usable

component	typical model	quantity	unit cost, k$
UPS	DELL	3	1.0
switch	DELL PowerConnect 48GbE, portmanaged	2	1.5
servers	DELL PE2950 E5440 processor, 2.83GHz, 32GB RAM, 250GB drive	3	4.2
compute elements	DELL PE1950 E5440 processor, 2.83GHz, 16GB RAM, 250GB drive	21	2.4
storage elements	DELL MD1000	2 (24TB, usable)	5.4
KVM	Belkin	1	1.3
rack			1
total cost			$82.1k

T3g

Tier 3 with “grid” connectivity

a campus-based, tower cluster

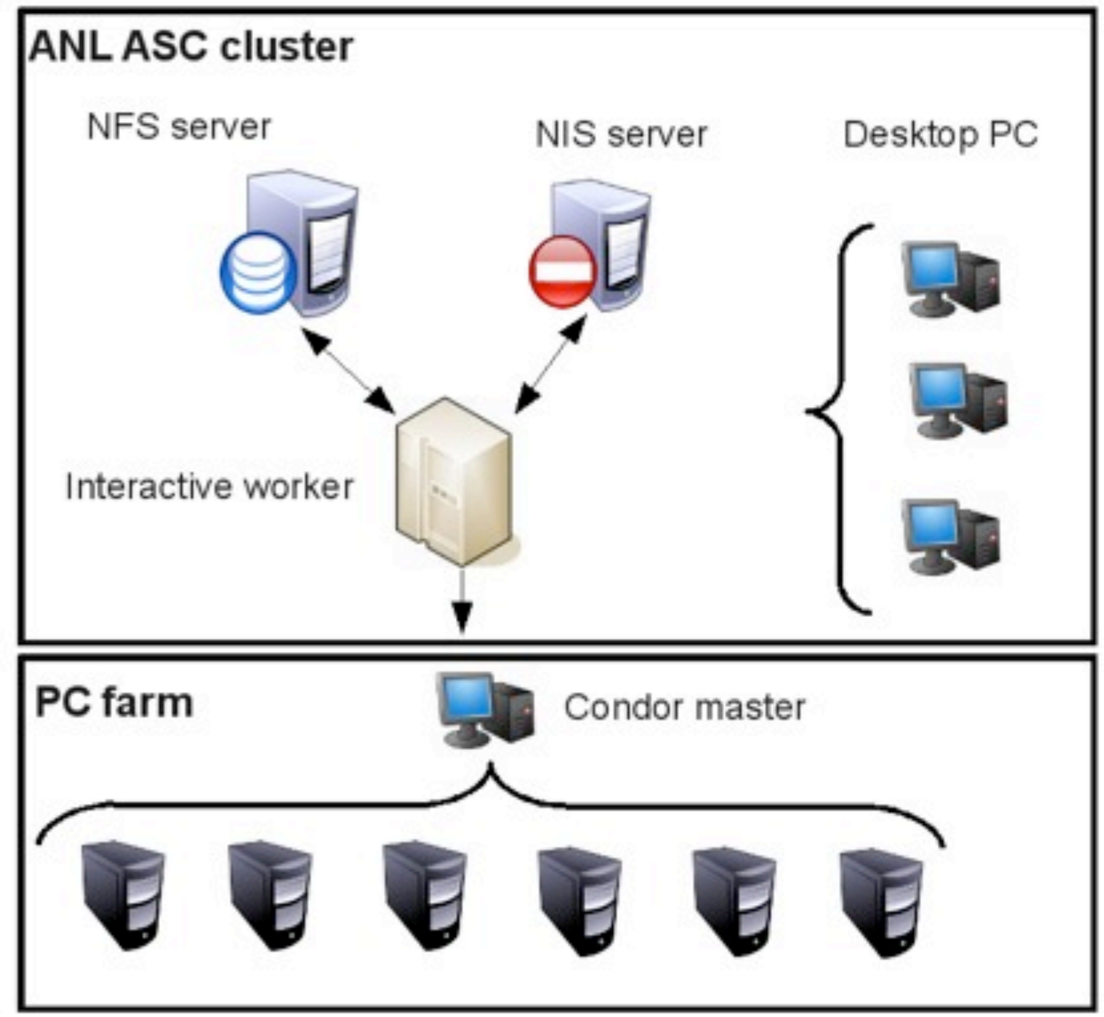
office-based

Characterized a strawman

~$25k

ANL and Duke are building them

80 processors
>100kSI2k
20TB



component	typical model	quantity	unit cost, k$
switch	Cisco 1GB	1	2.5
worker towers	Intel-based E5410 2.33GHz, 2 TB storage 8GB RAM	10	2.0
server elements	DELL PE1950 E5440 processor, 2.83MHz, 16GB RAM, 250GB drive	4	0.5
total cost			$24.5k

Two other T3 classes

T3w

Tier 3 **W**orkstation

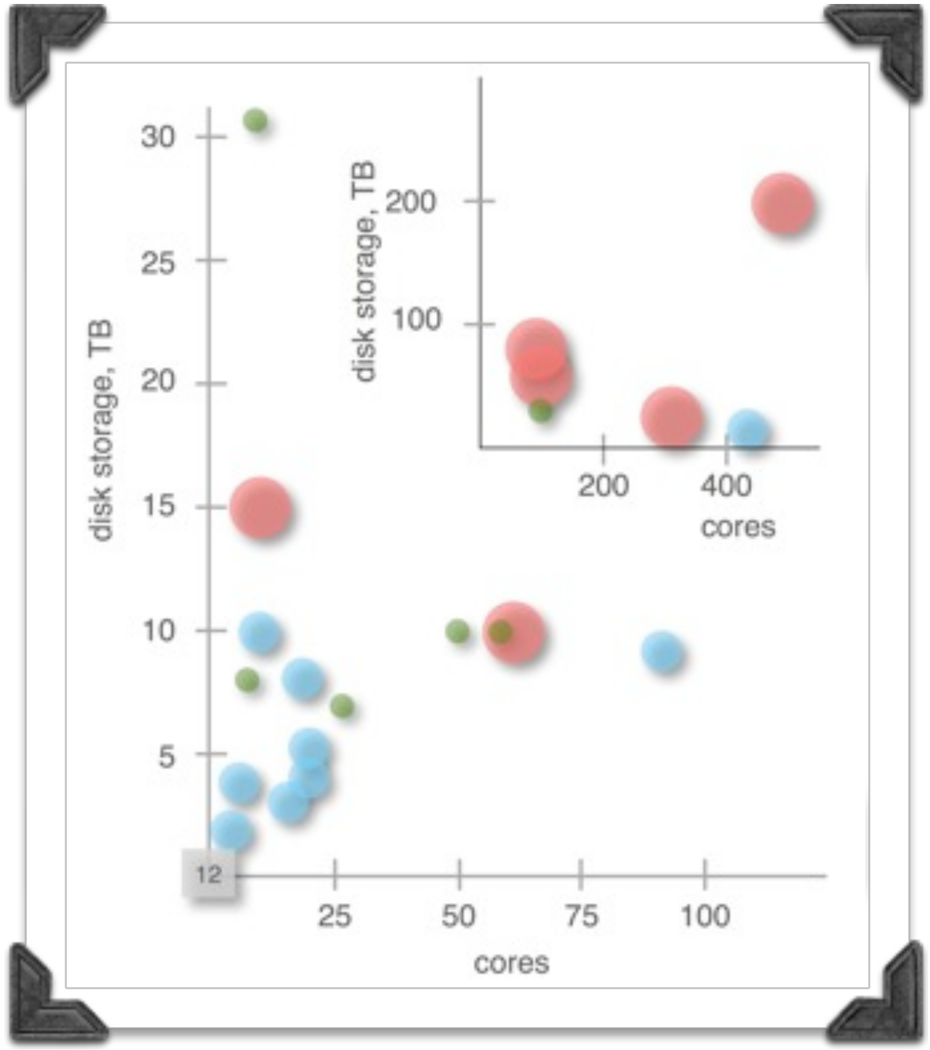
unclustered workstations...OSG, DQ2 client, root, etc

T3af

Tier 3 system built into lab or university **a**nalysis **f**acility

special arrangement of purchasing through the AF

the CDF Model–fair-share computing privileges in exchange for contribution


30
25
20
15
10
5
12
25
50
75
100
cores
disk storage, TB
200
100
200
400
cores
disk storage, TB

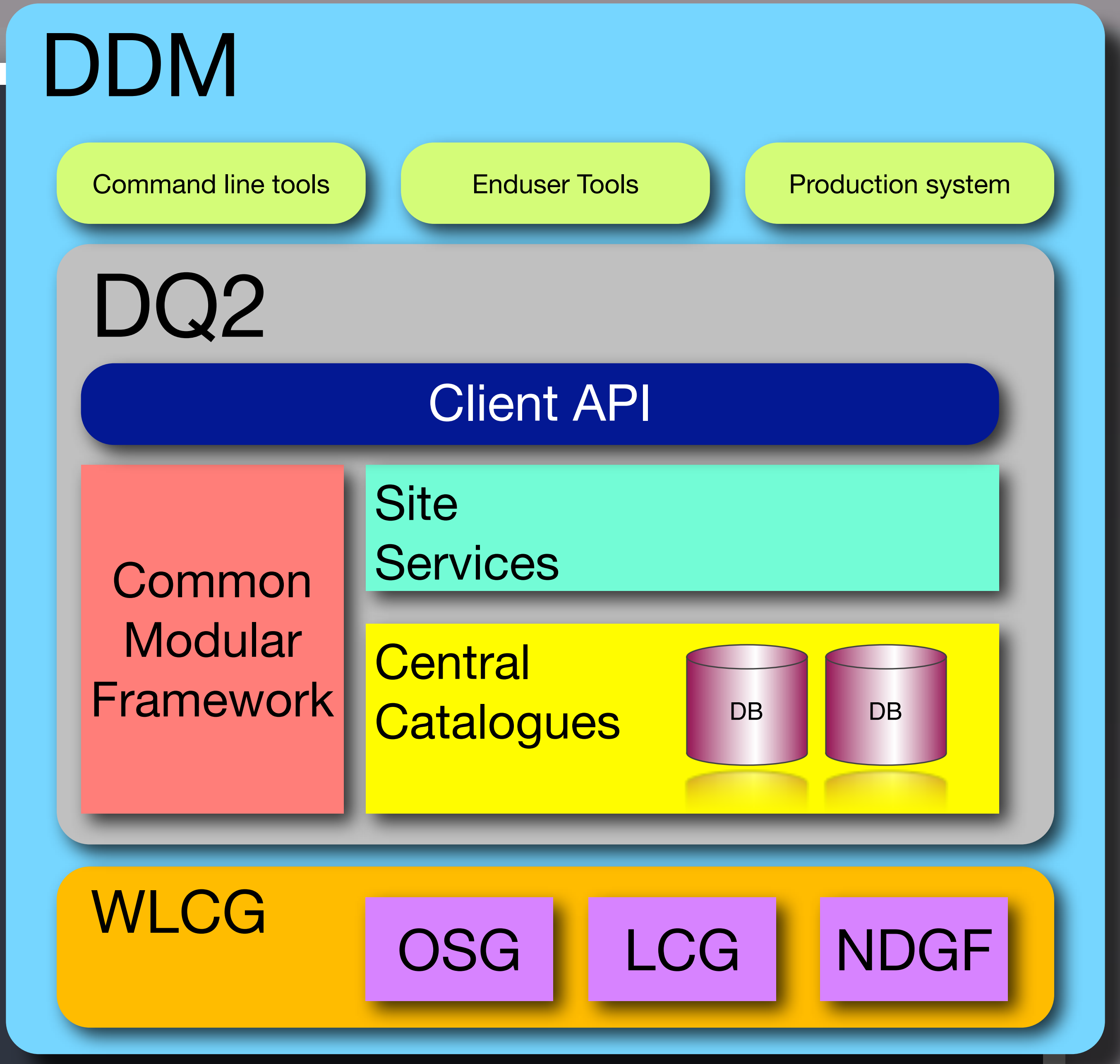
evolution


disk storage, TB
30
25
20
15
10
5
12
25
50
75
100
cores
200
100
200
400
Tier3gs region
250
150
50
24
300
Tier3g region

evolution

2. Recommended modifications to the ATLAS data management scheme

ATLAS Distributed Data Management (DDM)



“Don Quijote 2” (DQ2) system

DATASET/file-based for all ATLAS formats, RAW to user-defined

Owns all ATLAS SEs

Operates within the WLCG (OSG, LCG, NDGF)

a project:

changes to DQ2

in order to facilitate dataset subscription

in order to shield the Tier 3 from the whole data catalog

access to the data

Tier 3gs and Tier 3g will require significant data transfer

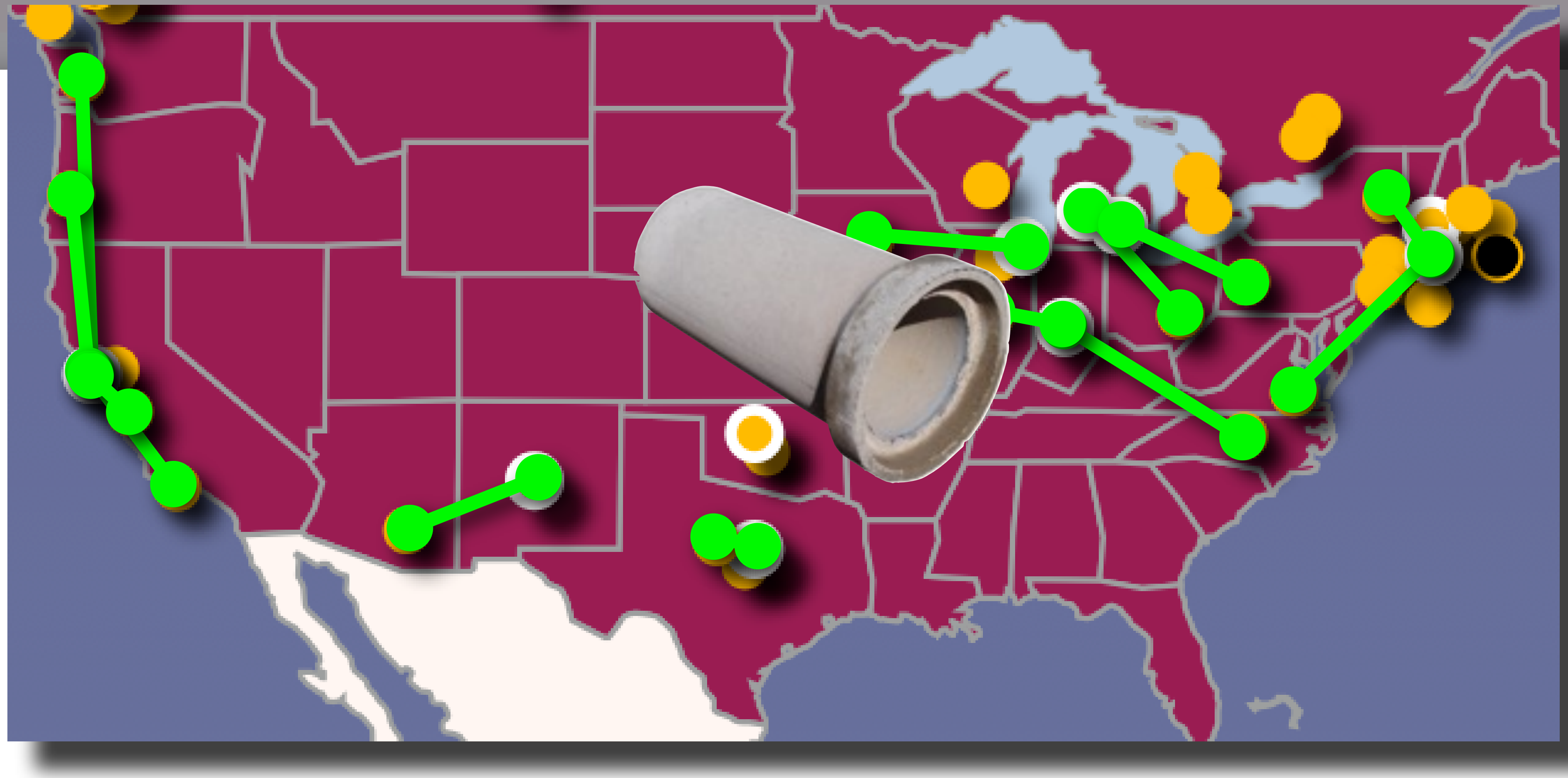
Episodic

Sustained, scheduled (?) transfer rates

To move 1-2TB per day

~20MBps

X



“TB per day”

Require a robust, point-to-point connectivity

One particular university to one particular Tier 2

a project:

to determine the best point-to-point connections

measure and determine the bottlenecks

fix them...by 2011-2012

Technical Support

ATLAS analysis support

Internet2 technical support

OSG technical support

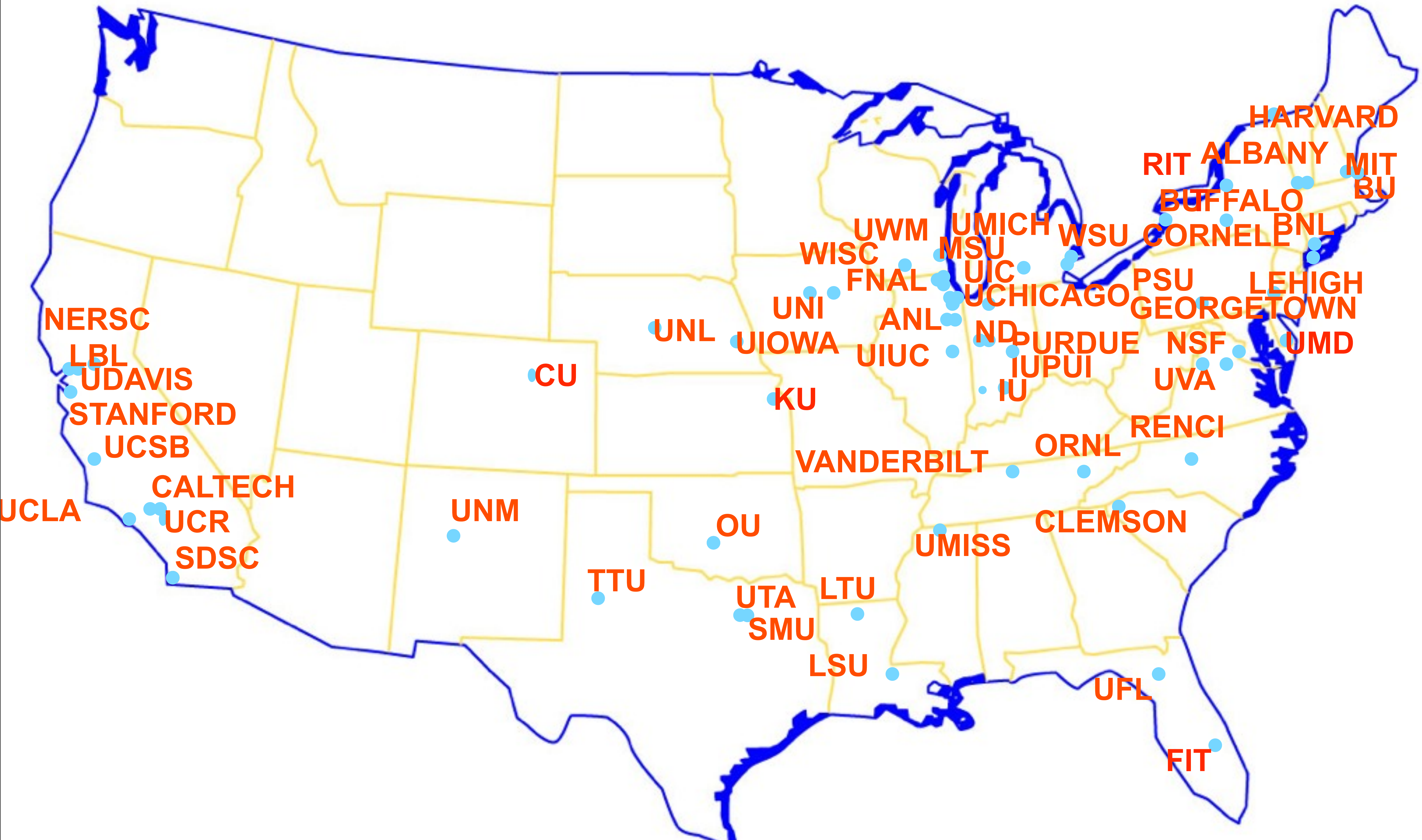
university technical and infrastructure support

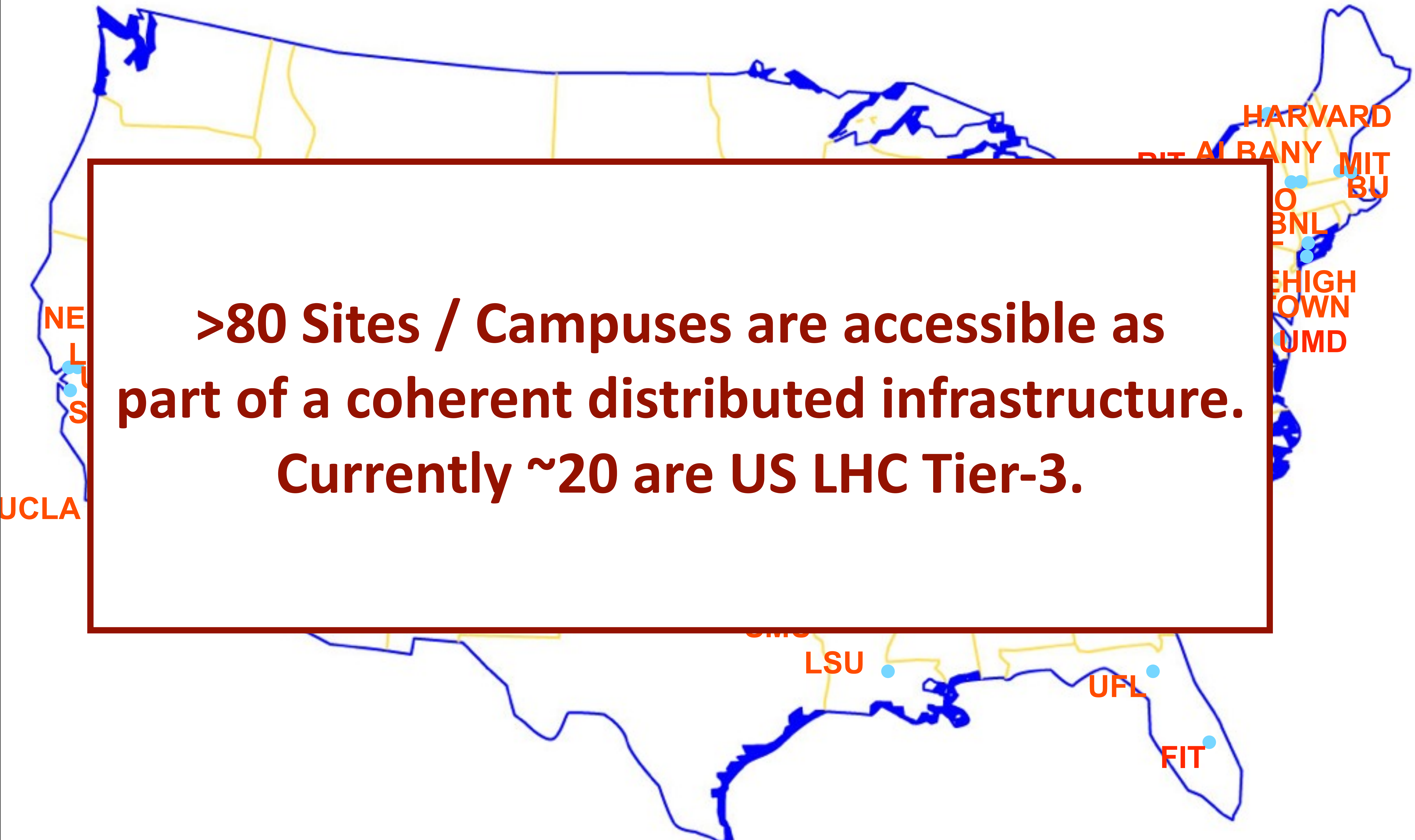
in that spirit

Ruth Pordes asked me to show the following:

The Open Science Grid & the Tier-3s

Sites on the OSG:





>80 Sites / Campuses are accessible as part of a coherent distributed infrastructure. Currently ~20 are US LHC Tier-3.

Reaching out to the rest of the Campus

- Physics departments typically host the Tier-3 resources and administration.

- OSG & Internet2 can help and consult with the rest of the Campus:
  - Network and system architectures.
  - Sharing of computing farms and storage.
  - Software for using remote as well as local resources.

Open Science Grid Provides:

Collaboration with ESNET and Internet2 network for data movement and network use.

Software

Movement, storage and management of the data.

Job workflow, scheduling and execution.

Services

Information, accounting and monitoring.

Monitoring to determine the availability of sites.

Support

Security monitoring, incident response, and mitigation.

Operational support including centralized Ticket Handling.

Site Coordination: common support for site administrators.

End-to-end support for running production applications.

Open Science Grid Provides:

Collaboration with ESNET and Internet2 network for data movement and network use.

Software

Movement, storage and management of the data.

Job

Servic

Info

Mo

Suppo

Security monitoring, incident response, and mitigation.

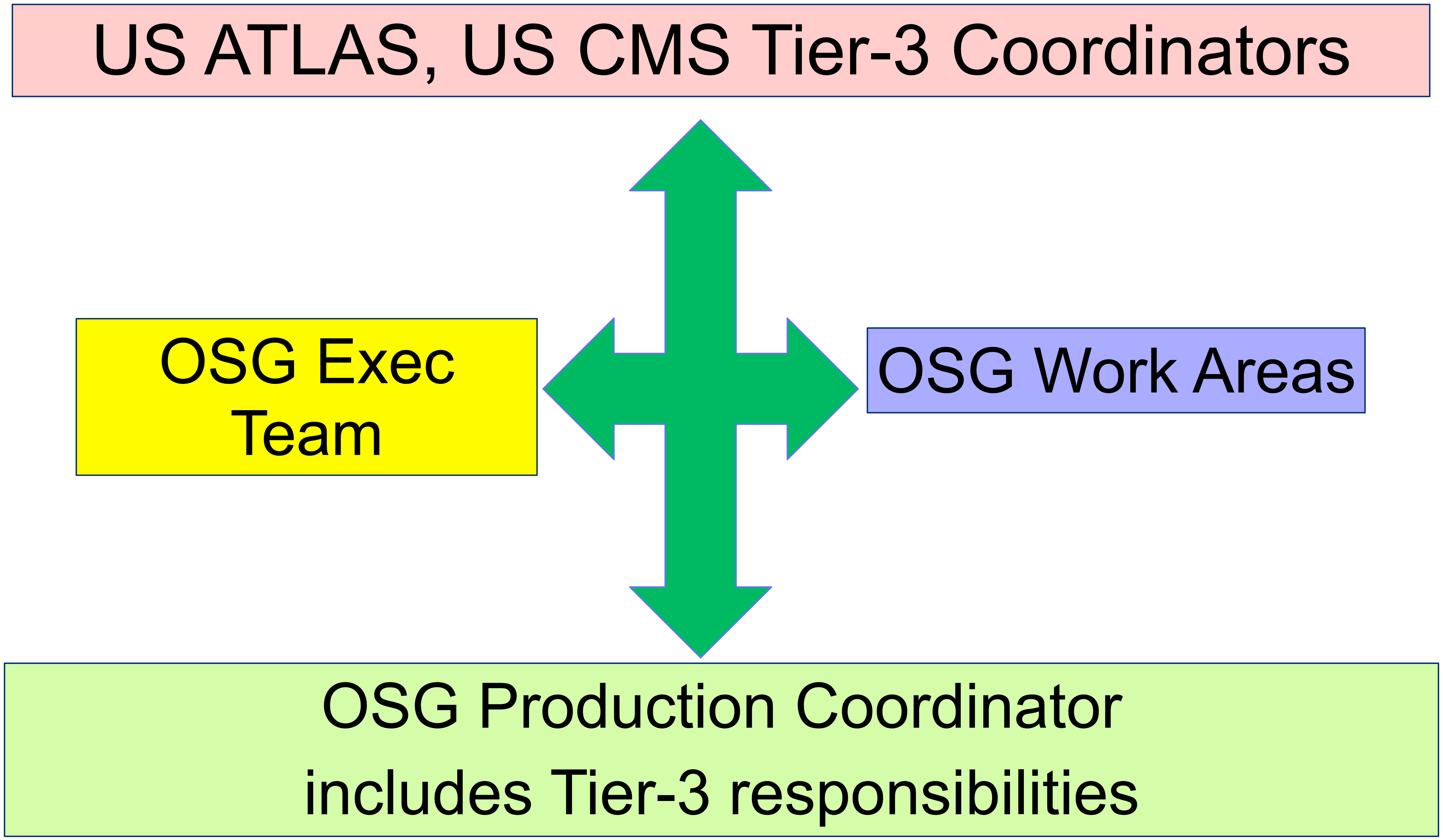
Operational support including centralized Ticket Handling.

Site Coordination: common support for site administrators.

End-to-end support for running production applications.

Not a "one size fits all": Support determined by what is needed.

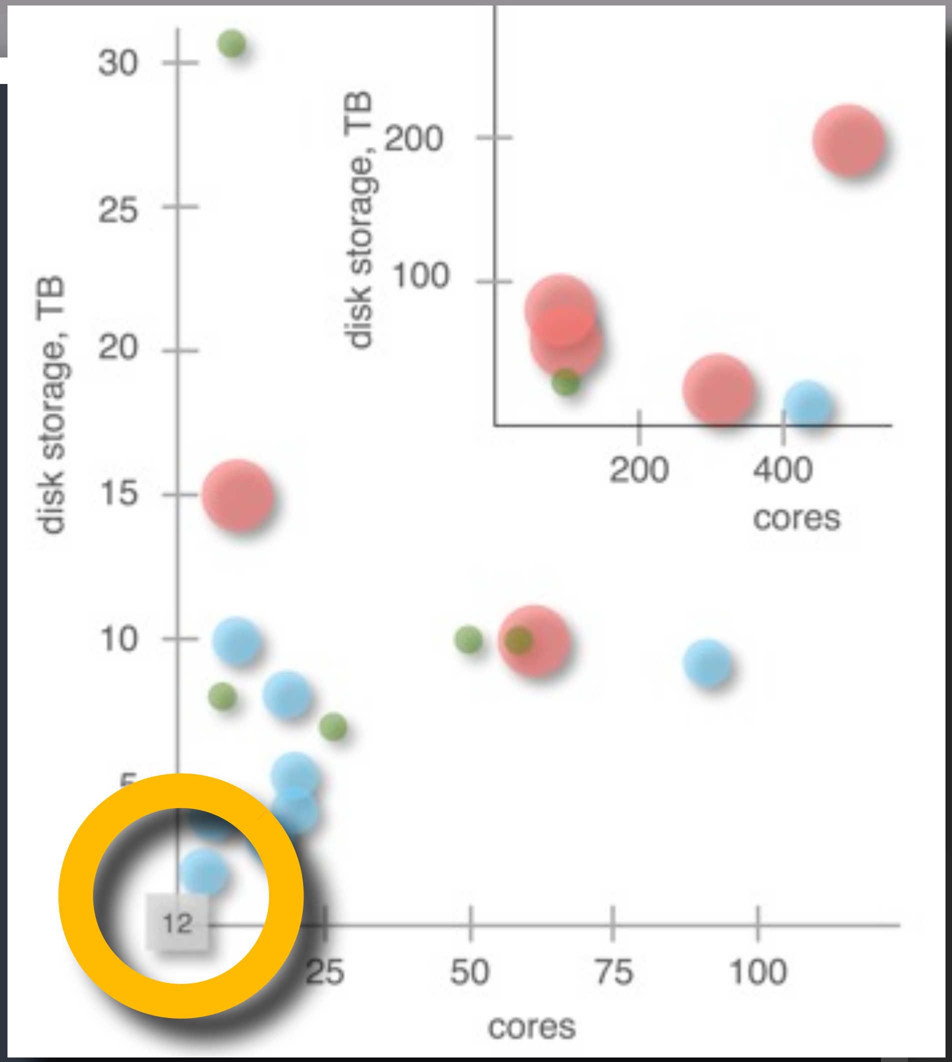
Specific Support Group starting



where are we?

- **early days...**

  couple of weeks since report accepted



- **Immediate issues:**

  Doug Benjamin and Jim Cochran and ANL attacking Tier 3s

  http://atlaswww.hep.anl.gov/twiki/bin/view/Tier3Setup/18May09Meeting

- **Longer term:**

  funding & planning for T3g and T3gs and their infrastructure

Here's

LHC is a hug
undertaking

Maybe the largest techni er

A huge collaboration of:

scientific, engineering, governmental,

and universities

thanks.

Raymond (Chip) Brock
Department of Physics and Astronomy
Michigan State University
brock@pa.msu.edu