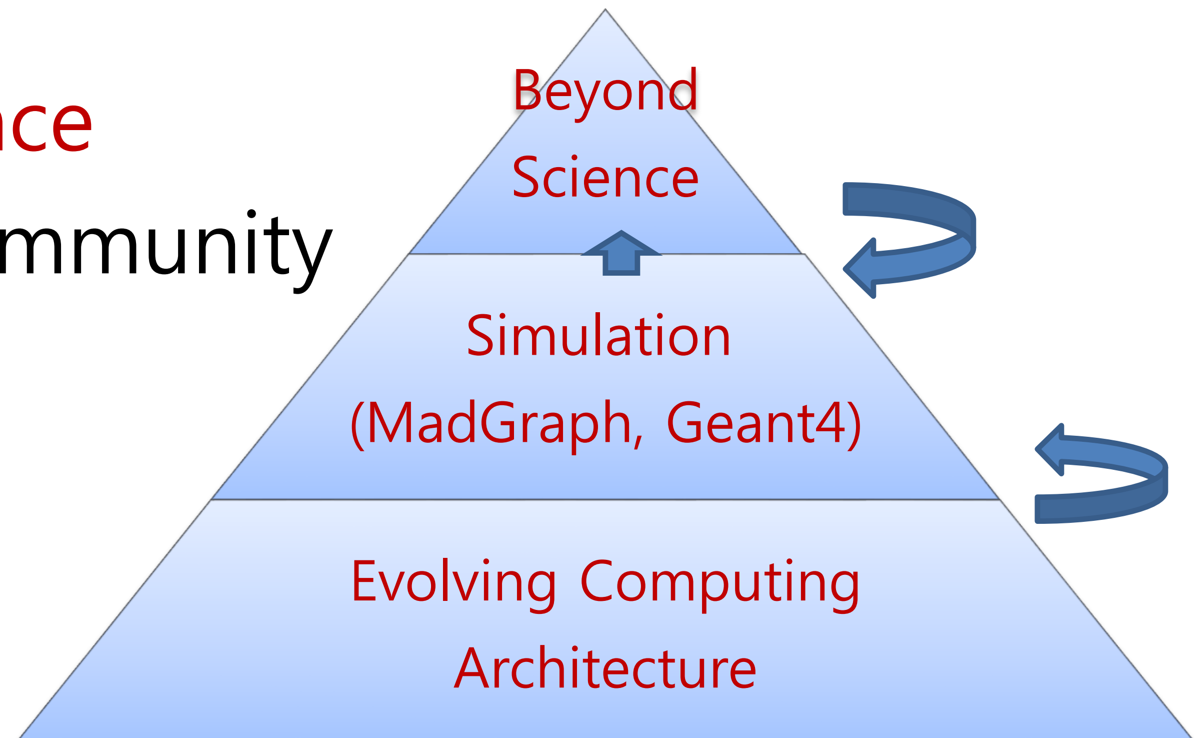
e-Science Activities in Korea

2019. 4. 4

Kihyeon Cho (KISTI)

Contents

- Science Today
- Evolving Computing Architecture
- Simulation
- Beyond Science
- e-Science Community
- Summary

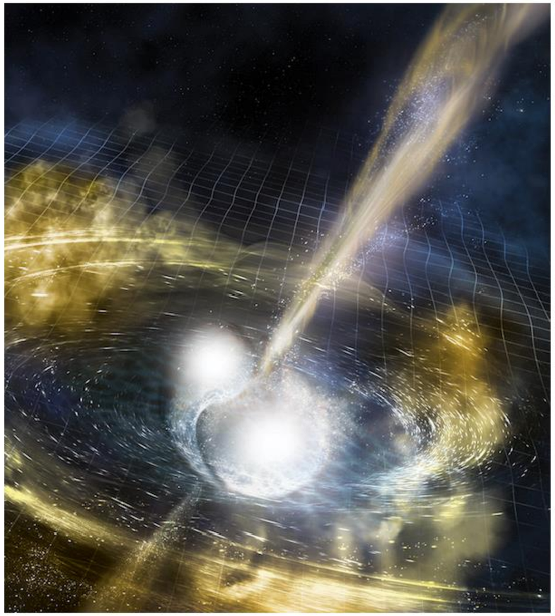


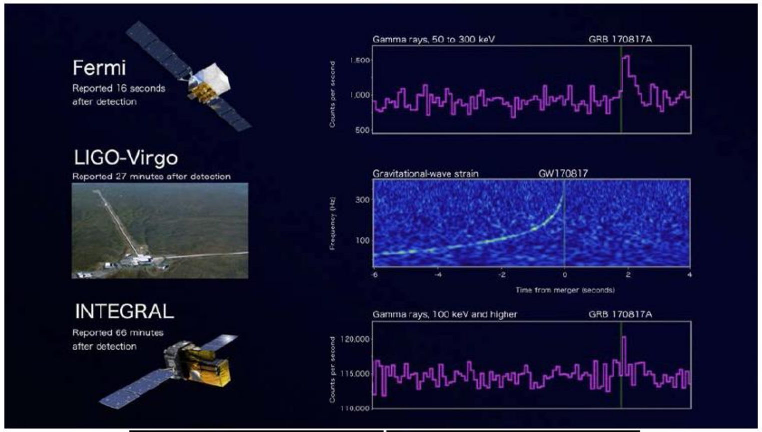
Science Today

(e.g. Physics, Astronomy)

Windows on Universe

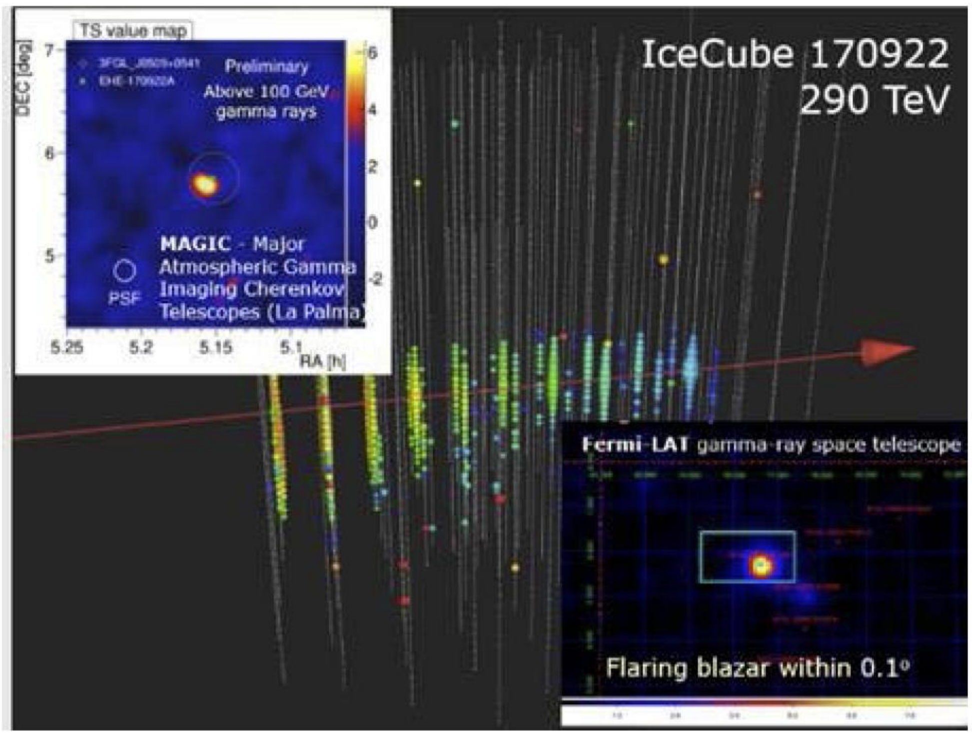
Binary Neutron Star Merger (08/17/2017)

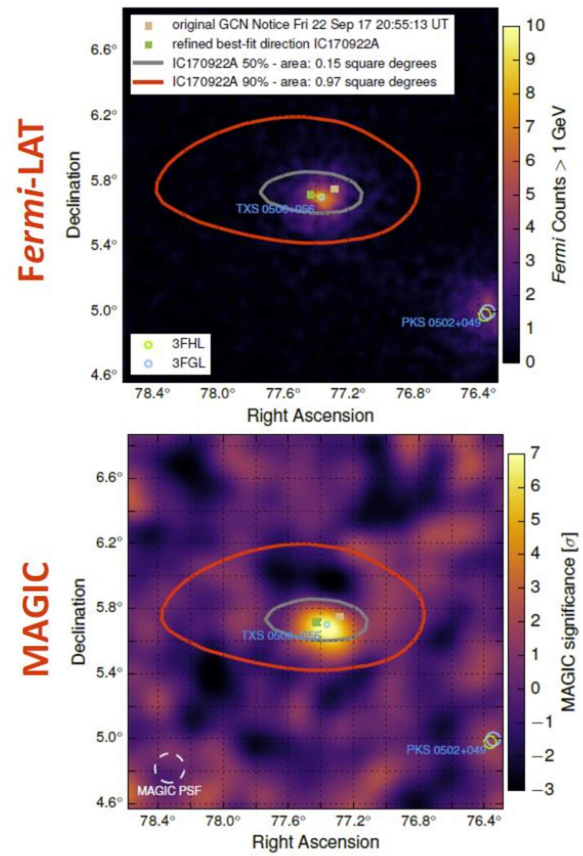




Images: LIGO.org/NASA/INTEGRAL/ESA

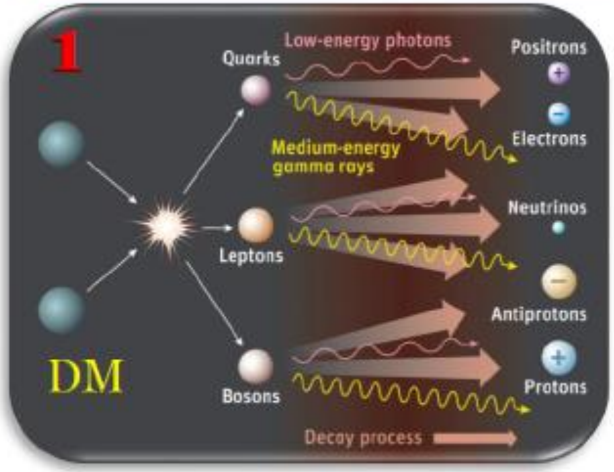
Flaring Blazar (09/22/2017)

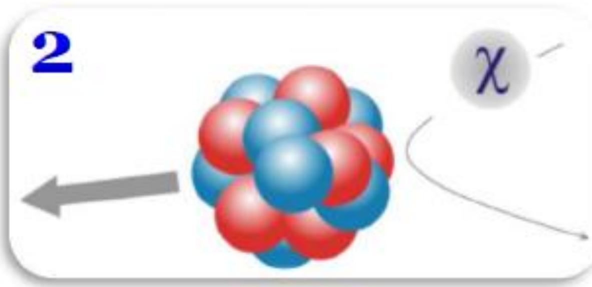


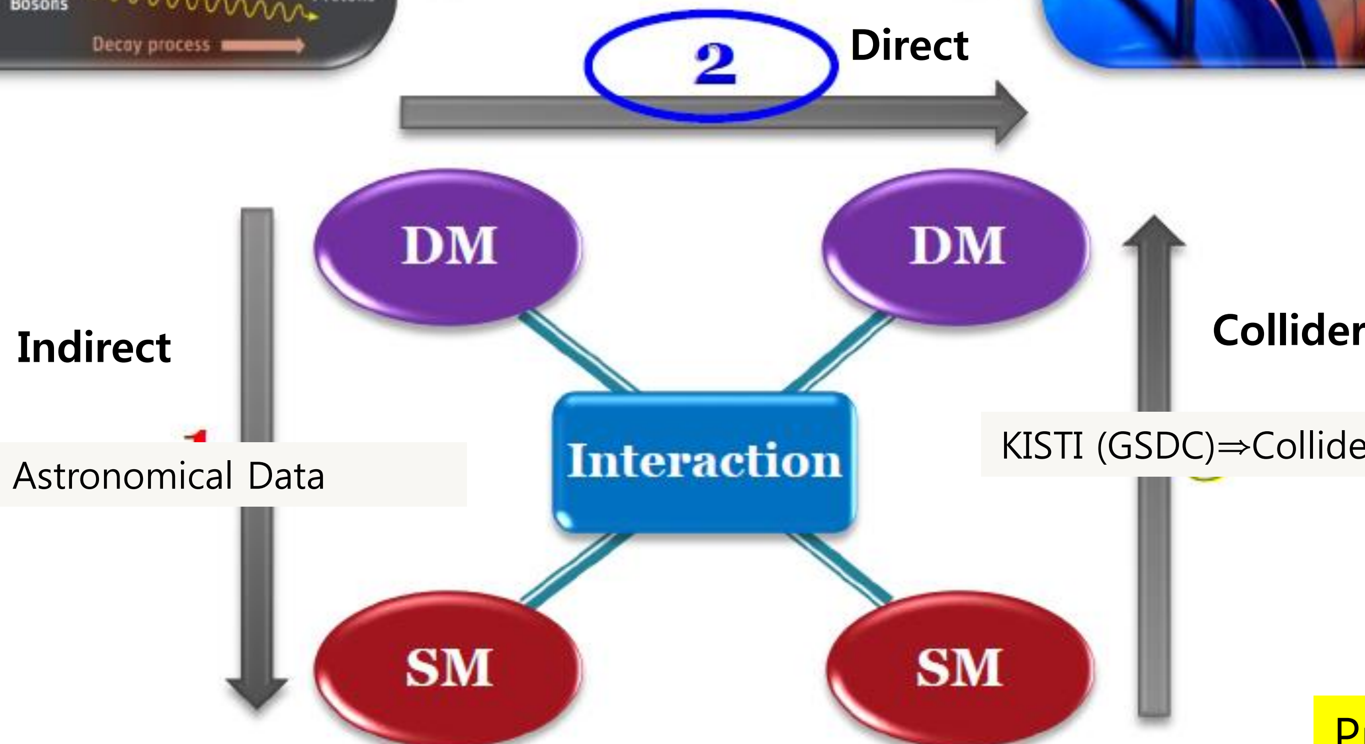


⇒ The era of multi-messenger astrophysics

Dark Matter Search





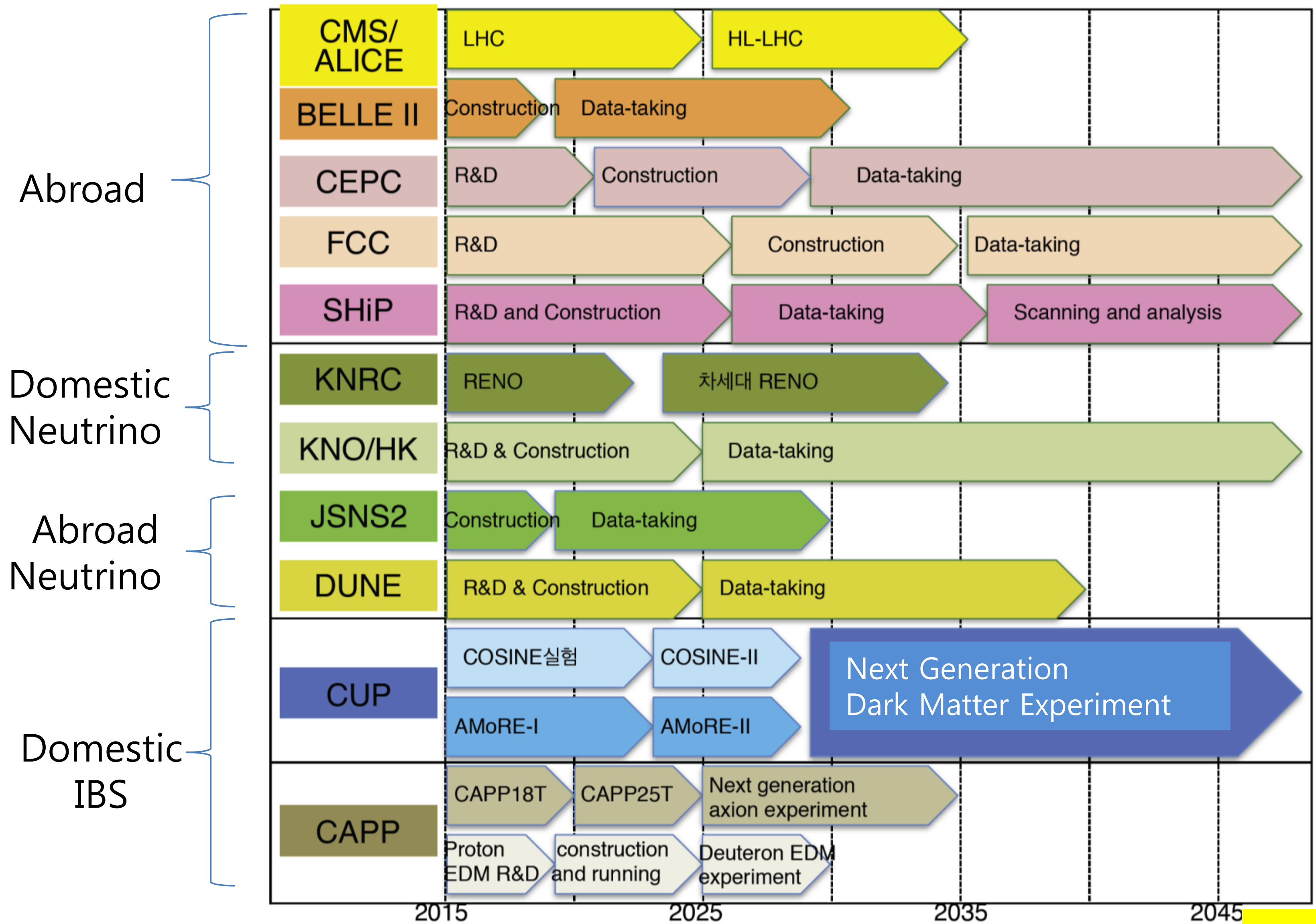


KASI ⇒ Astronomical Data

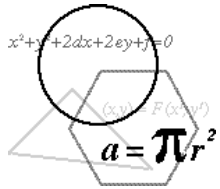
KISTI (GSDC)⇒Collider Data

Prof. J.C. Park

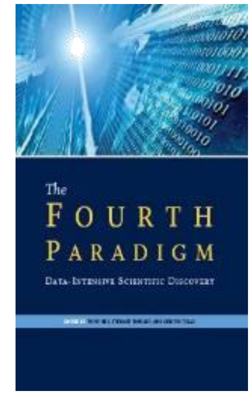
HEP Experiments in Korea



The changing nature of scientific research

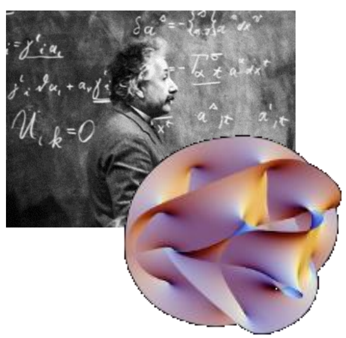


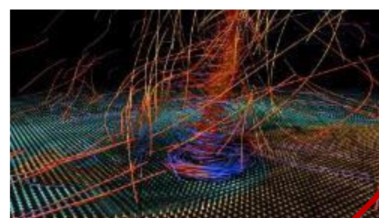
$$H(t)|\psi(t)\rangle = i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle$$



Experimental → Theoretical → Computational Simulation → Data and Machine Learning

Large Hadron Collider





Severe Storm Model (NCSA)

ARRAY OF THINGS

e-Science

Prof. D. Reed

Scientific discovery drivers

1. e–Science

– Theory-Experiment-Simulation
– Science any time anywhere

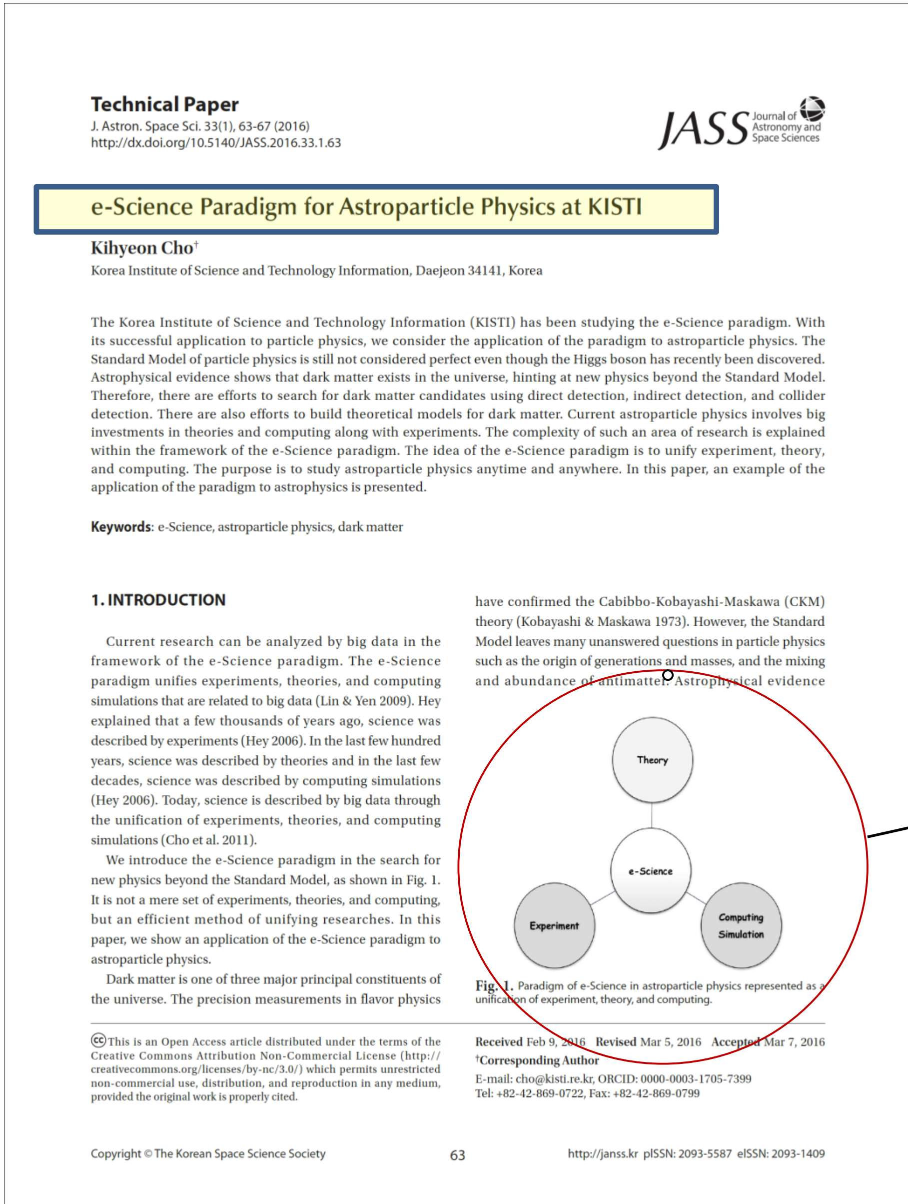
2. Data

– Sensors, Instruments, DB, Internet, Storage

3. Machine Learning

– AI, Statistics, Data Mining, Algorithms, Deep Learning

1. e-Science



Technical Paper
J. Astron. Space Sci. 33(1), 63-67 (2016)
http://dx.doi.org/10.5140/JASS.2016.33.1.63

JASS Journal of Astronomy and Space Sciences

e-Science Paradigm for Astroparticle Physics at KISTI

Kihyeon Cho†
Korea Institute of Science and Technology Information, Daejeon 34141, Korea

The Korea Institute of Science and Technology Information (KISTI) has been studying the e-Science paradigm. With its successful application to particle physics, we consider the application of the paradigm to astroparticle physics. The Standard Model of particle physics is still not considered perfect even though the Higgs boson has recently been discovered. Astrophysical evidence shows that dark matter exists in the universe, hinting at new physics beyond the Standard Model. Therefore, there are efforts to search for dark matter candidates using direct detection, indirect detection, and collider detection. There are also efforts to build theoretical models for dark matter. Current astroparticle physics involves big investments in theories and computing along with experiments. The complexity of such an area of research is explained within the framework of the e-Science paradigm. The idea of the e-Science paradigm is to unify experiment, theory, and computing. The purpose is to study astroparticle physics anytime and anywhere. In this paper, an example of the application of the paradigm to astrophysics is presented.

Keywords: e-Science, astroparticle physics, dark matter

1. INTRODUCTION

Current research can be analyzed by big data in the framework of the e-Science paradigm. The e-Science paradigm unifies experiments, theories, and computing simulations that are related to big data (Lin & Yen 2009). Hey explained that a few thousands of years ago, science was described by experiments (Hey 2006). In the last few hundred years, science was described by theories and in the last few decades, science was described by computing simulations (Hey 2006). Today, science is described by big data through the unification of experiments, theories, and computing simulations (Cho et al. 2011).

We introduce the e-Science paradigm in the search for new physics beyond the Standard Model, as shown in Fig. 1. It is not a mere set of experiments, theories, and computing, but an efficient method of unifying researches. In this paper, we show an application of the e-Science paradigm to astroparticle physics.

Dark matter is one of three major principal constituents of the universe. The precision measurements in flavor physics have confirmed the Cabibbo-Kobayashi-Maskawa (CKM) theory (Kobayashi & Maskawa 1973). However, the Standard Model leaves many unanswered questions in particle physics such as the origin of generations and masses, and the mixing and abundance of antimatter. Astrophysical evidence

Fig. 1. Paradigm of e-Science in astroparticle physics represented as a unification of experiment, theory, and computing.

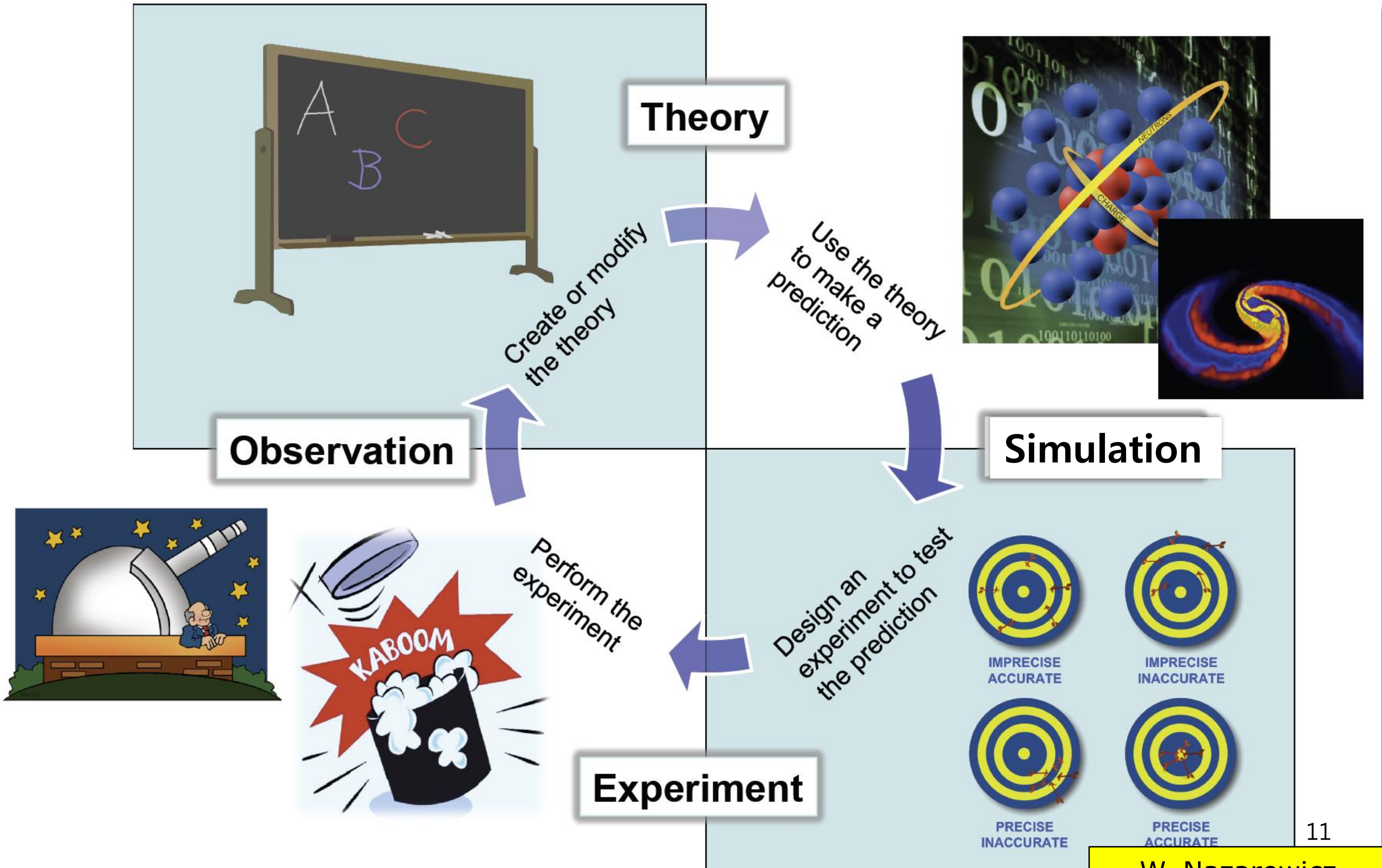
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received Feb 9, 2016 **Revised** Mar 5, 2016 **Accepted** Mar 7, 2016
†**Corresponding Author**
E-mail: cho@kisti.re.kr, ORCID: 0000-0003-1705-7399
Tel: +82-42-869-0722, Fax: +82-42-869-0799

Copyright © The Korean Space Science Society 63 http://janss.kr pISSN: 2093-5587 eISSN: 2093-1409

- Theory-Experiment-Simulation

⇒ e-Science



W. Nazarewicz

2. Data

Data production

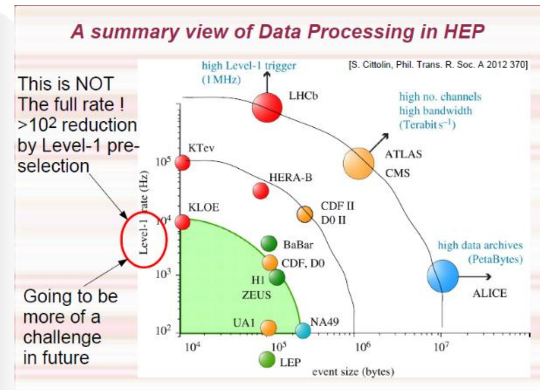
- Data Sharing

Data	Accelerator Data	Astronomical Data	Simulation Data
Input	Collaboration only	Collaboration only	All researcher share
Output	All researcher share	All researcher share	All researcher share

The size of data

Accelerator Data

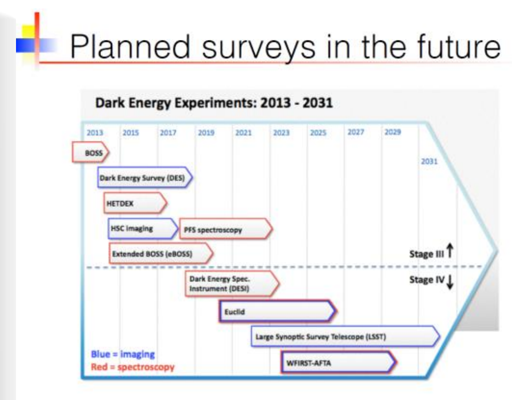
Year	Experiment	Size of data
Current	Belle	~1 PB
Current	LHC(CMS)	10~20 PB/year 100 PB/year(Simulation included)
2019~2024	Belle II	100 PB (50 times of current)
2025~	LHC(CMS)	100PB/year (5~10 times of current)



S. Cittoling

Astronomy Data

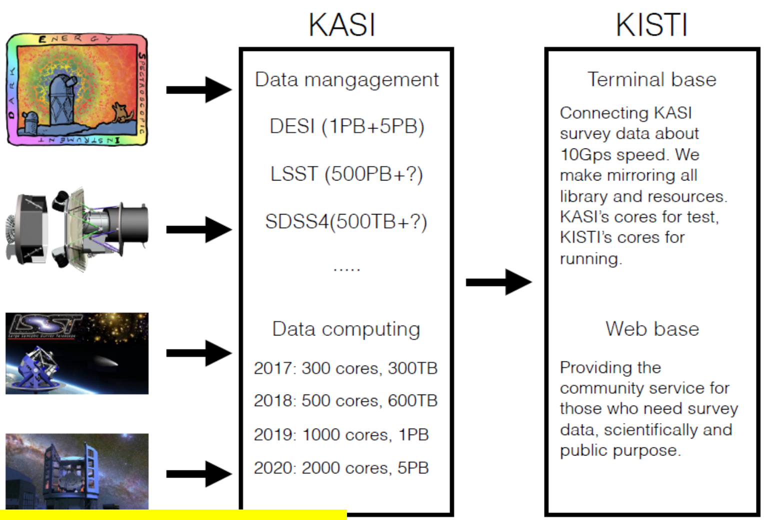
Year	Observatory	Size of Data (Total)
Current	SDSS	300TB
2019~	DESI (Starting)	1PB
2014~2020	SDSS4	500TB
2017~2022	DESI	5PB
2023~2030	LSST	500PB



Y.S. Song

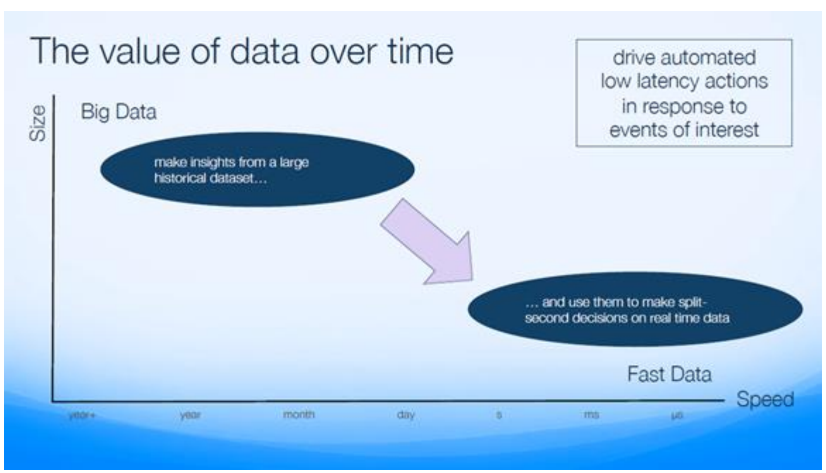
Data Management System

Survey data management



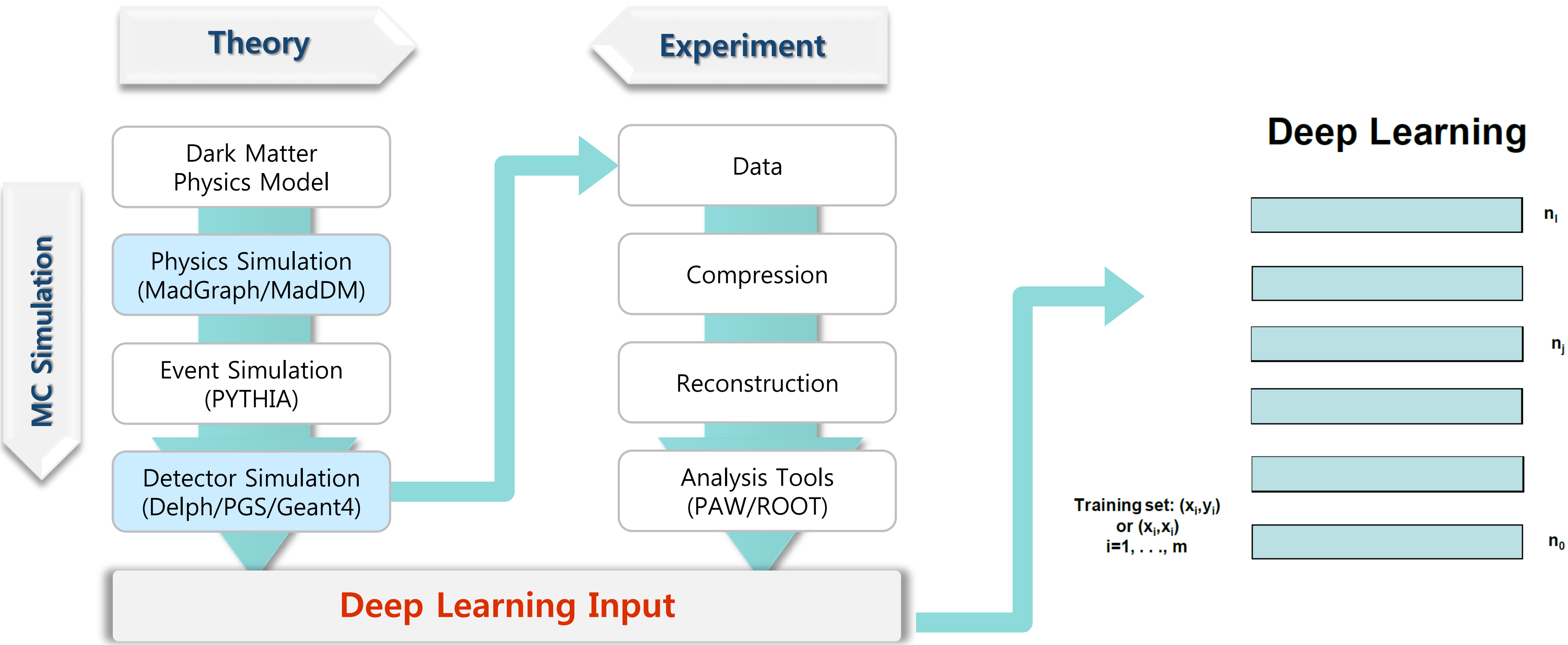
Prof. Y. S. Song

- From Big data to Fast data
- From theory-driven approach to Data-driven approach

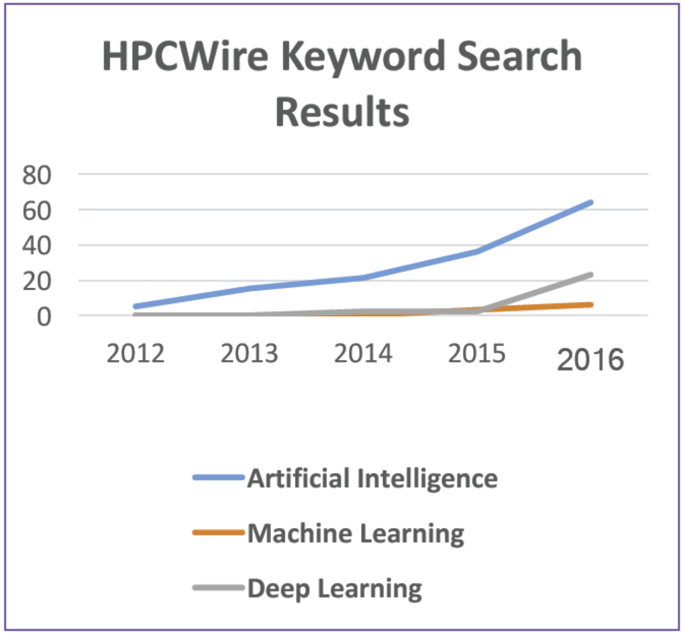


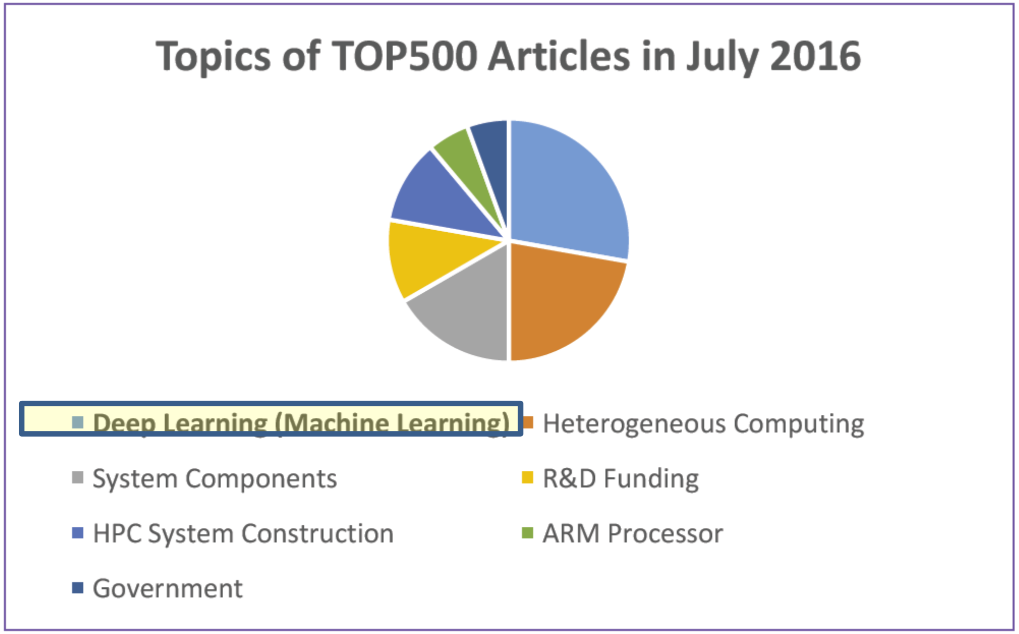
E. Dobson

3. Machine Learning



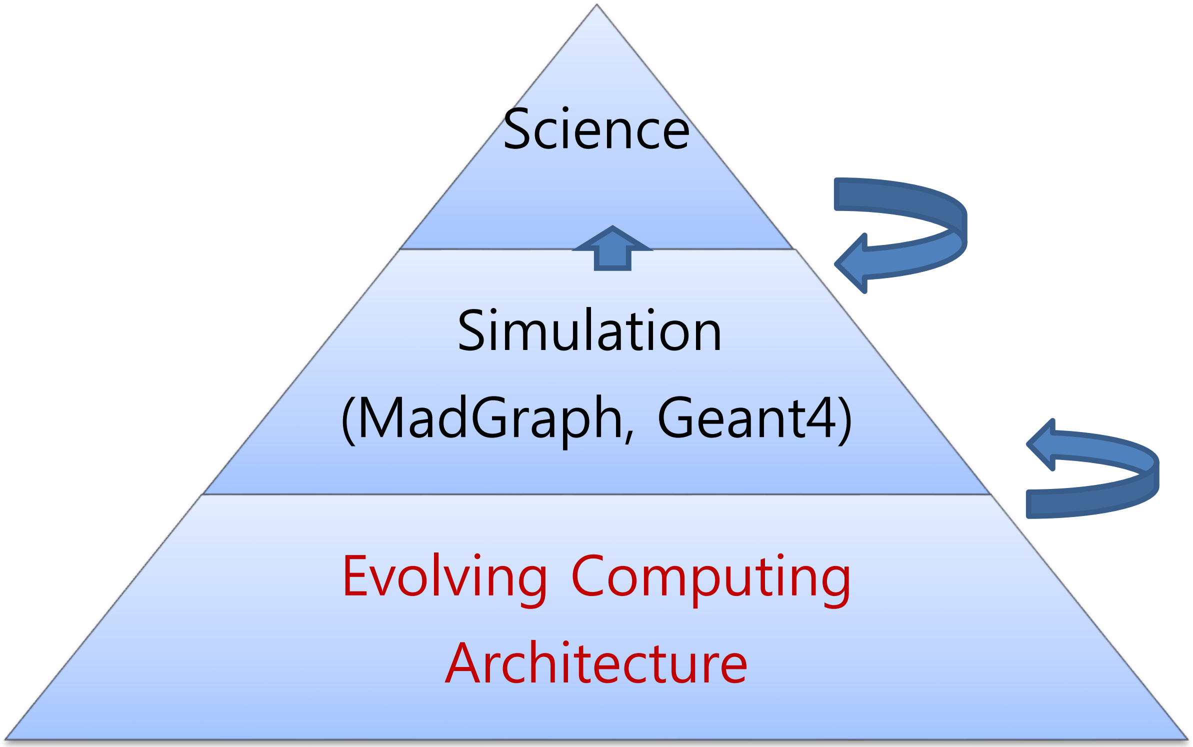
- Supercomputing meets machine learning.



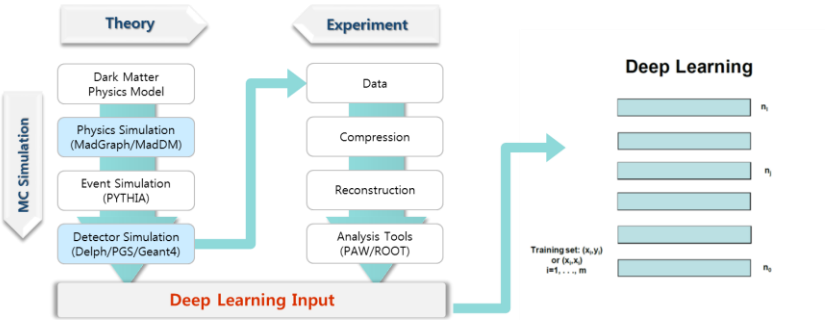


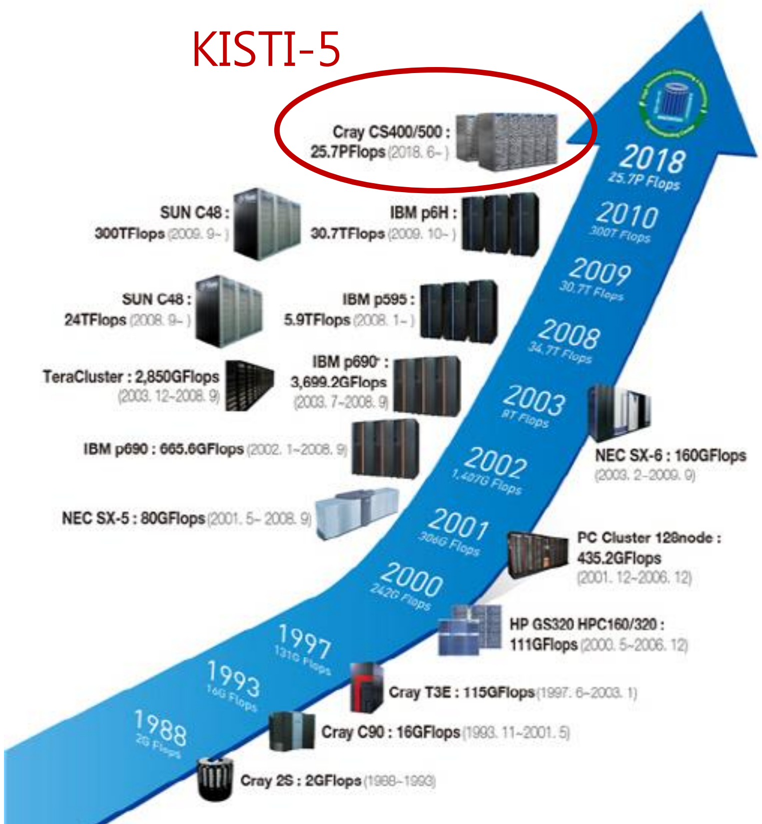
⇒ Evolving computing architecture

Evolving Computing Architecture



KISTI Supercomputing center





- KISTI-5 Supercomputer
  - Processing: 25.7PF
    - Heterogeneous: 25.3PF CS400 w/KNL
    - CPU-only: 0.4PF CS500 w/SKL
  - Storage
    - 20PB SPS
    - 10PB Archive
  - Launched in November 2018

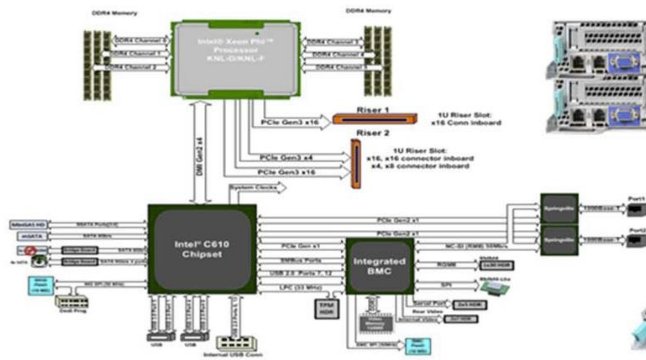
KISTI-5 supercomputer

Compute node

계산노드 | Cray 3112-AA000T(2U enclosure), KNL 기반 계산 노드 (8,305개)

25.3PF

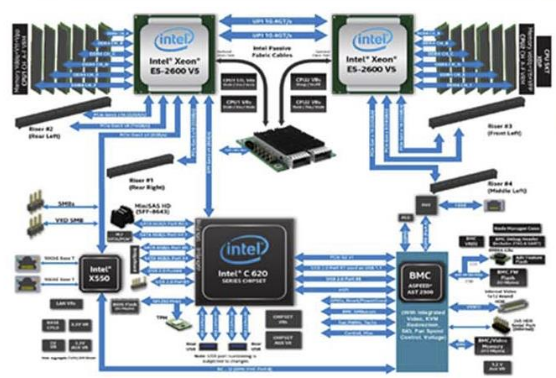
- 1x Intel Xeon Phi KNL 7250 processor
- 68 Cores per Processor
- 96GB (6x 16GB) DDR4-2400 RAM
- 1x Single-port 100Gbps OPA HFI card
- 1x On-board GigE (RJ45) port



CPU-only | Cray 3111-BA000T(2U enclosure), Skylake 기반 계산 노드 (132개)

0.4PF

- 2x Intel Xeon SKL 6148 (2.4 GHz) processors
- 20 Cores per processor (total 40 Cores)
- 192GB (12x 16GB) DDR4-2666 RAM–6CH
- 1x Single-port 100Gbps OPA HFI card
- 1x On-board GigE (RJ45) port



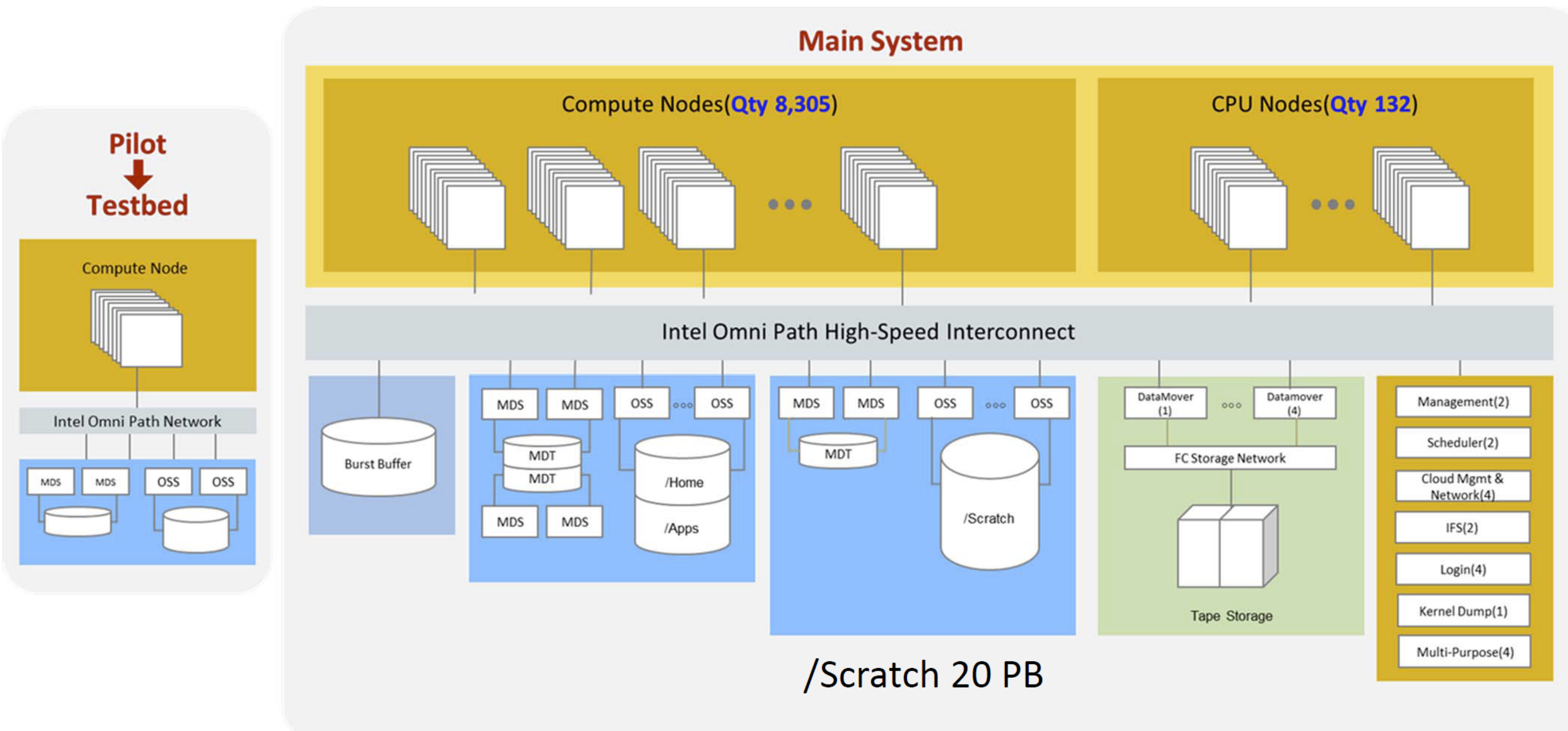
M.S. Joh

KISTI-5 supercomputer


CRAY
CS500

KISTI-5 supercomputer building

Architecture of KISTI-5 supercomputer



/Home 0.5 PB /Apps 0.5 PB

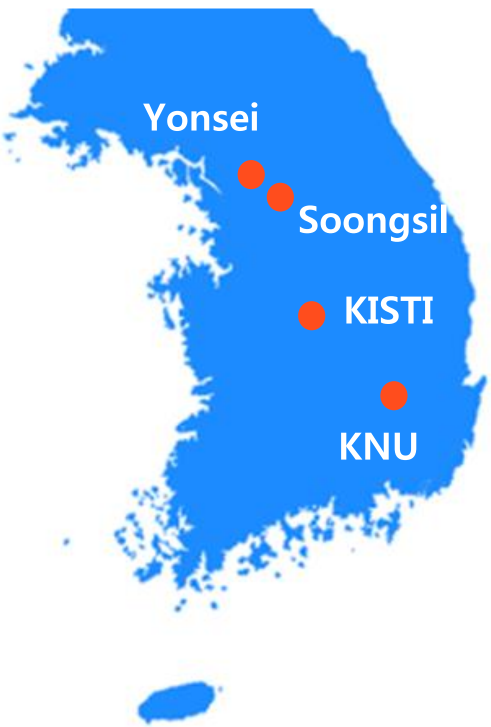
Burst Buffer 0.8 PB (SSD)

Tape Storage 10 PB

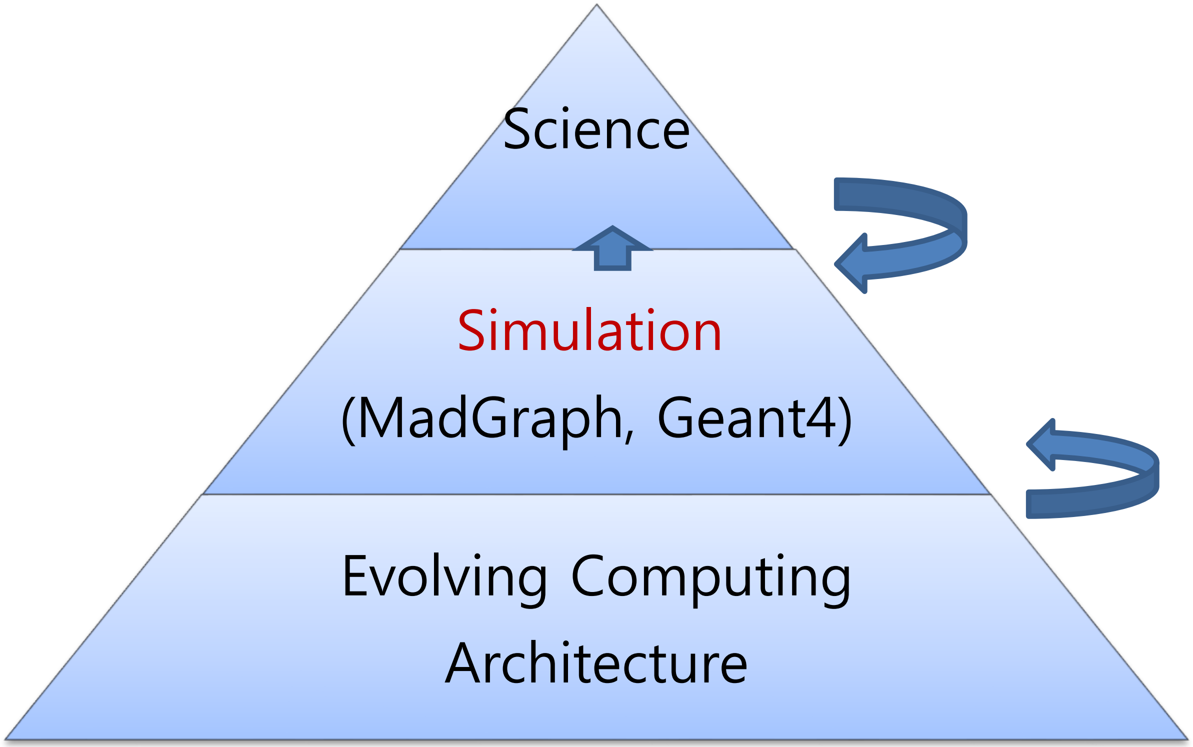
M.S. Joh

& HEP sites in Korea

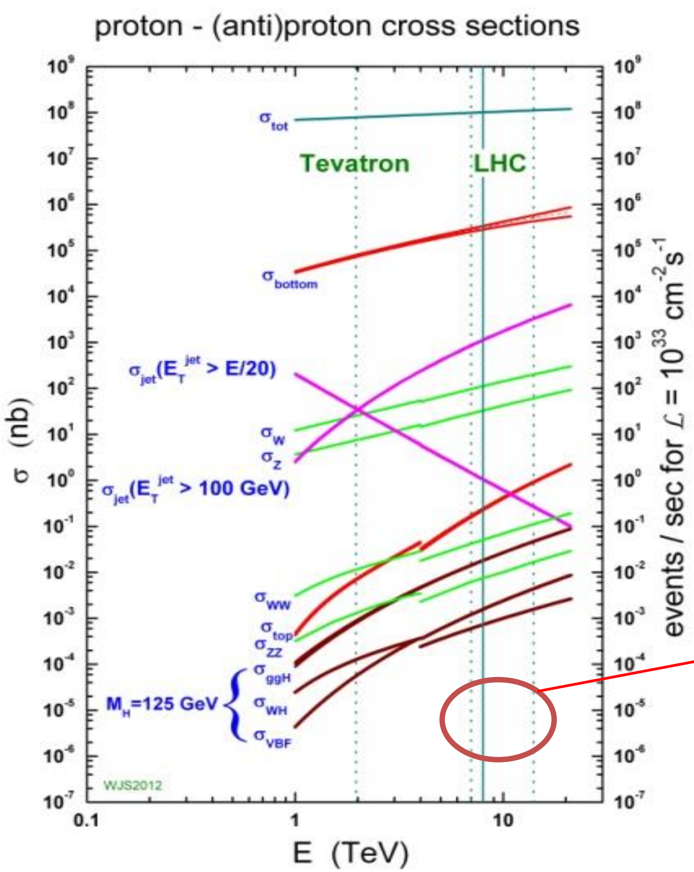
- Soongsil U. / Yonsei U.
  - Belle II Farm
- KNU
  - CMS Tier-2
- KISTI GSDC

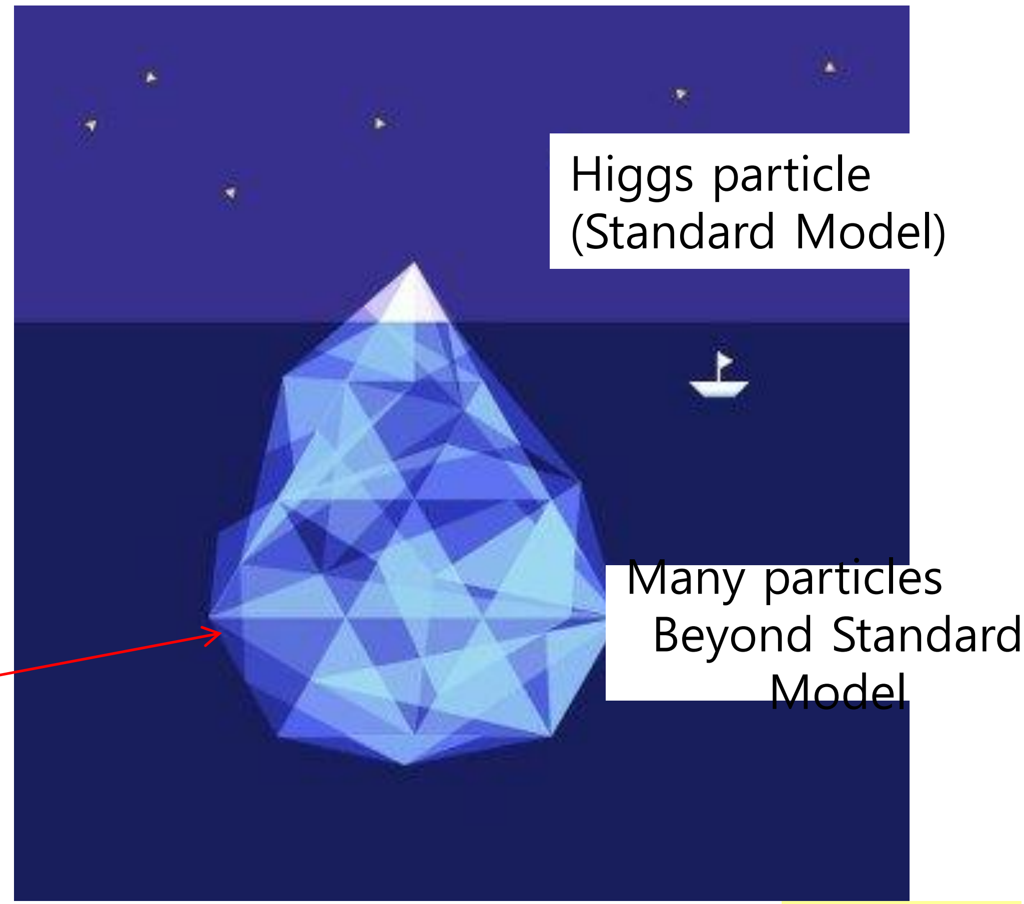


Simulation


Science
Simulation
(MadGraph, Geant4)
Evolving Computing
Architecture

Beyond the Standard Model





Adam Martin

⇒ BSM needs 1000 more Higgs events.

Simulation-based Optimization

(Computationally Expensive) Simulation-Based Optimization

$$\min_{\mathbf{x}\in\mathbb{R}^n} \{f(\mathbf{x}) = F[\mathbf{S}(\mathbf{x})] : \mathbf{c}(\mathbf{S}(\mathbf{x})) \leq 0, \mathbf{x} \in \mathbf{\Omega}\}$$

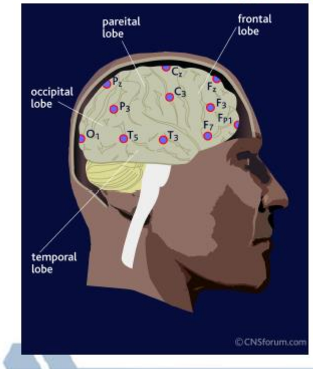
"parameter estimation", "model calibration", "design optimization", . . .

- Evaluating $\mathbf{S}$ means running a simulation modeling some (smooth) process
- Derivatives $\nabla_x S$ often unavailable or prohibitively expensive to obtain
- $\mathbf{S}$ (even when parallelized) takes secs/mins/days

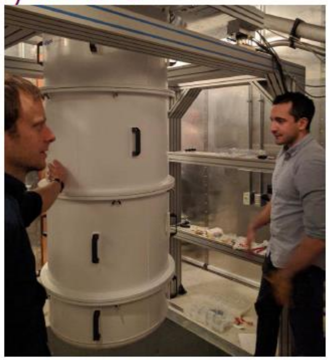
Evaluation is a bottleneck for optimization

- $\mathbf{\Omega}$ compact, known region (e.g., finite bound constraints)

Functions of complex (numerical/physical) simulations arise everywhere







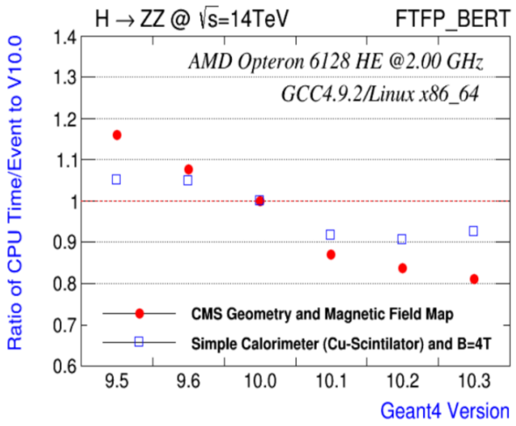
Stefan Wild

Geant4 Collaboration

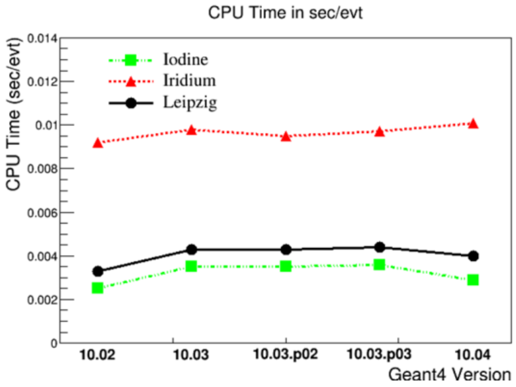
- Geant4 is the most successful model in HEP.
- HEP user community – BaBar(2001), LHC(2003), Belle II
- Prof. Kihyeon Cho is the contact person in Korea.

Profiling system

- Current status
  - High energy physics profiling (Fermilab)
    - SimpliCarlo (Sequential)
    - CMSExp (Multi-Thread)

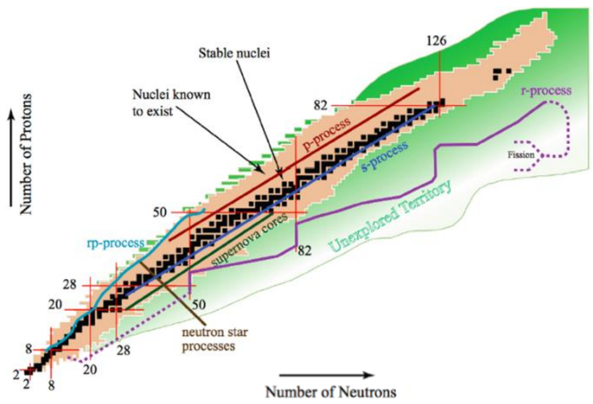


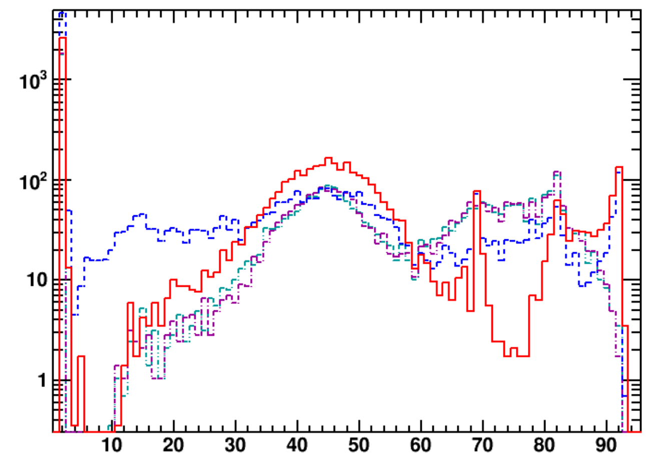
  - Low energy physics profiling (KISTI)
    - Using Brachytherapy code



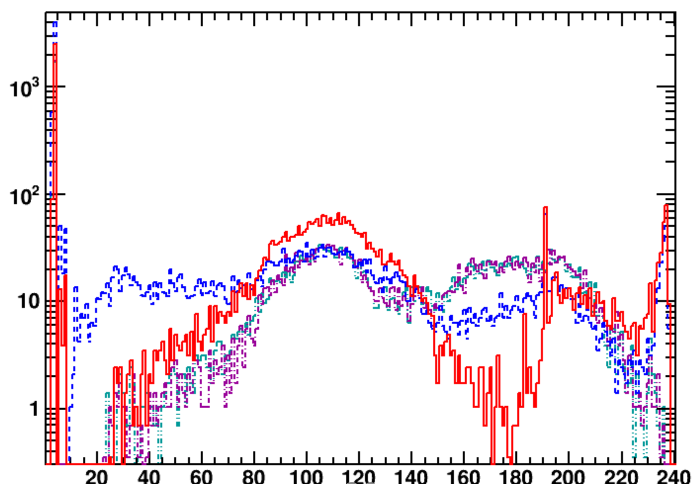
Beam Simulations for RISP

- *Application in astrophysics, nuclear imaging, …*





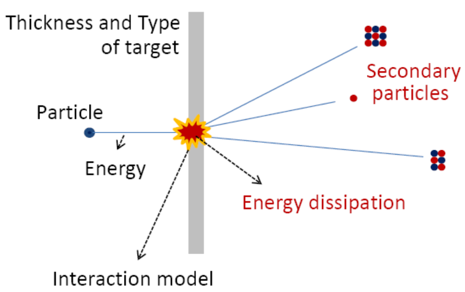
Atomic number(Z)



Atomic mass(A)

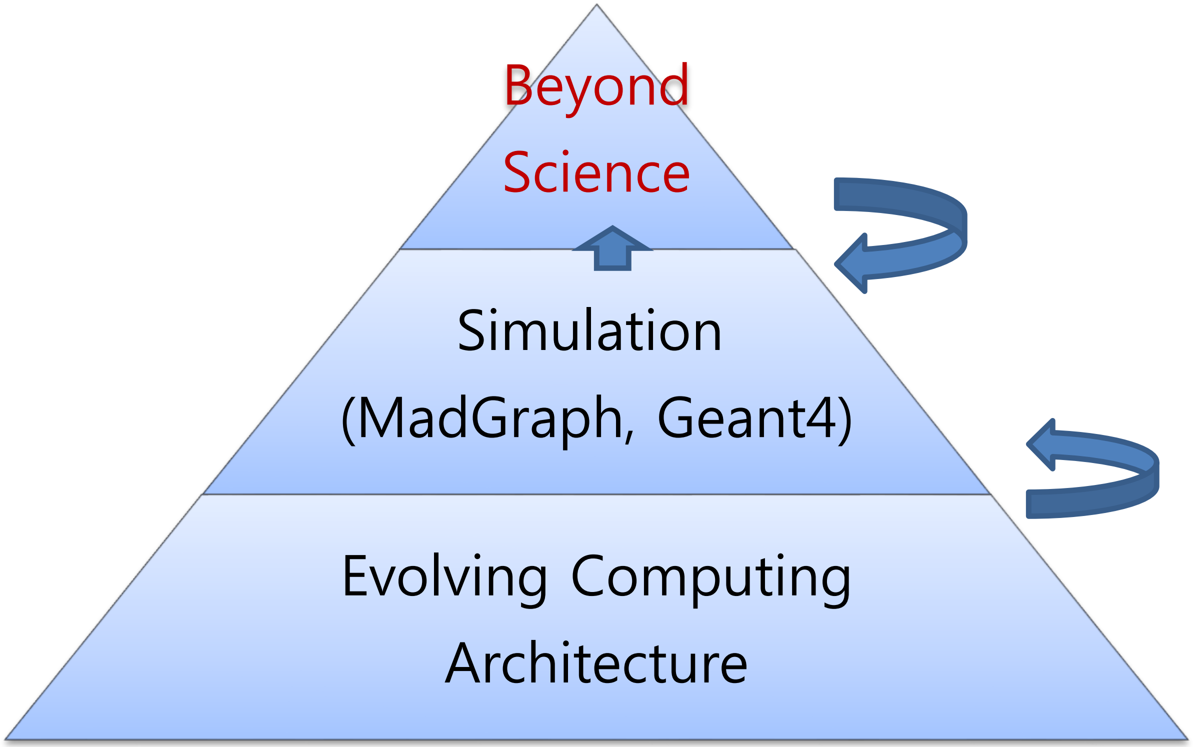
Proton beam with U Target

- 1000 GeV
- 100 GeV
- 10 GeV
- 1 GeV

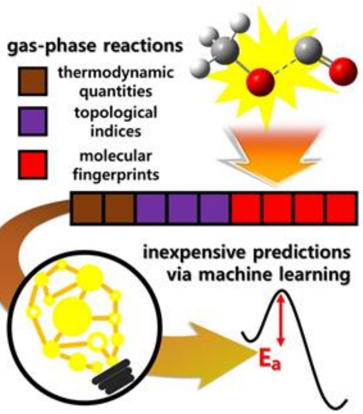


30

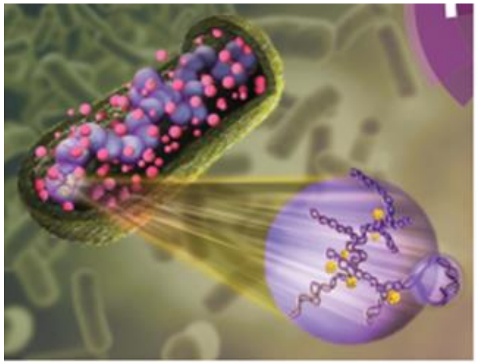
Beyond Science



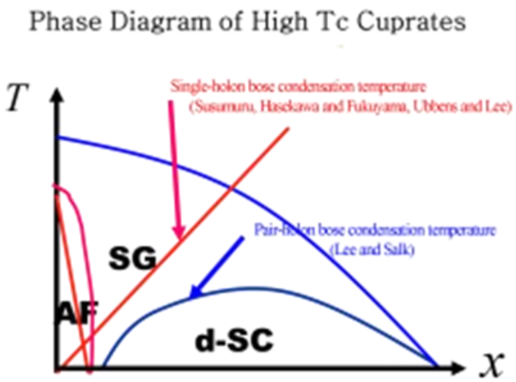
Sciences with KISTI-5 Supercomputer



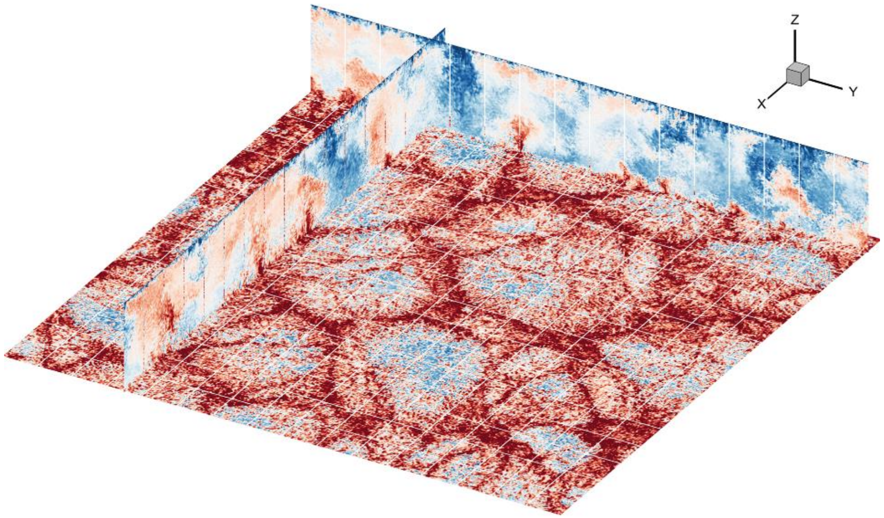
Computational Chemistry



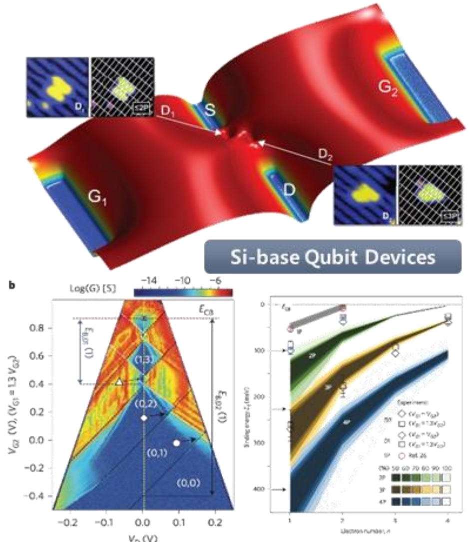
Biophysics



Condensed Matter Physics

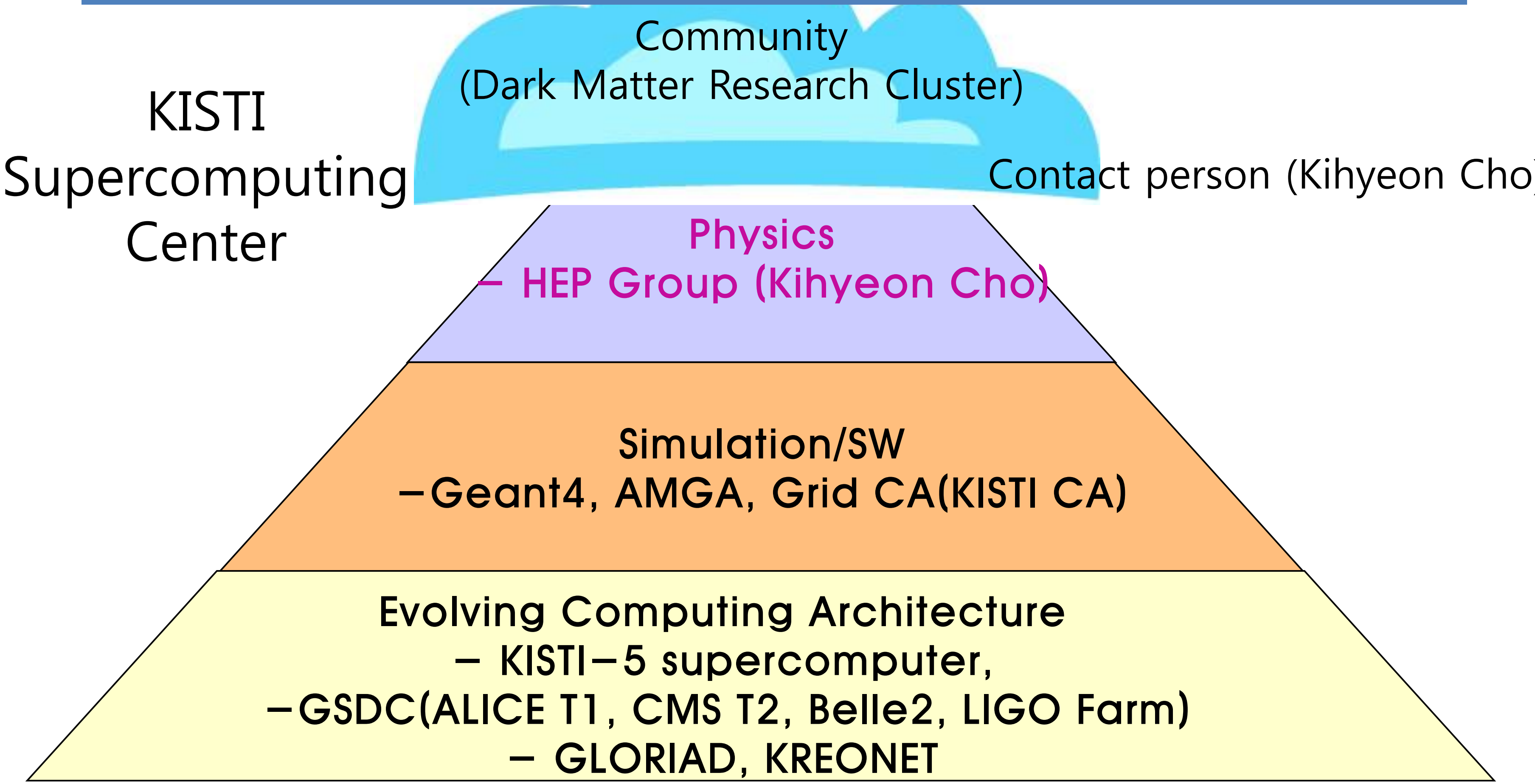


Computational Fluid Dynamics



Nano-electronics

e-Science Community

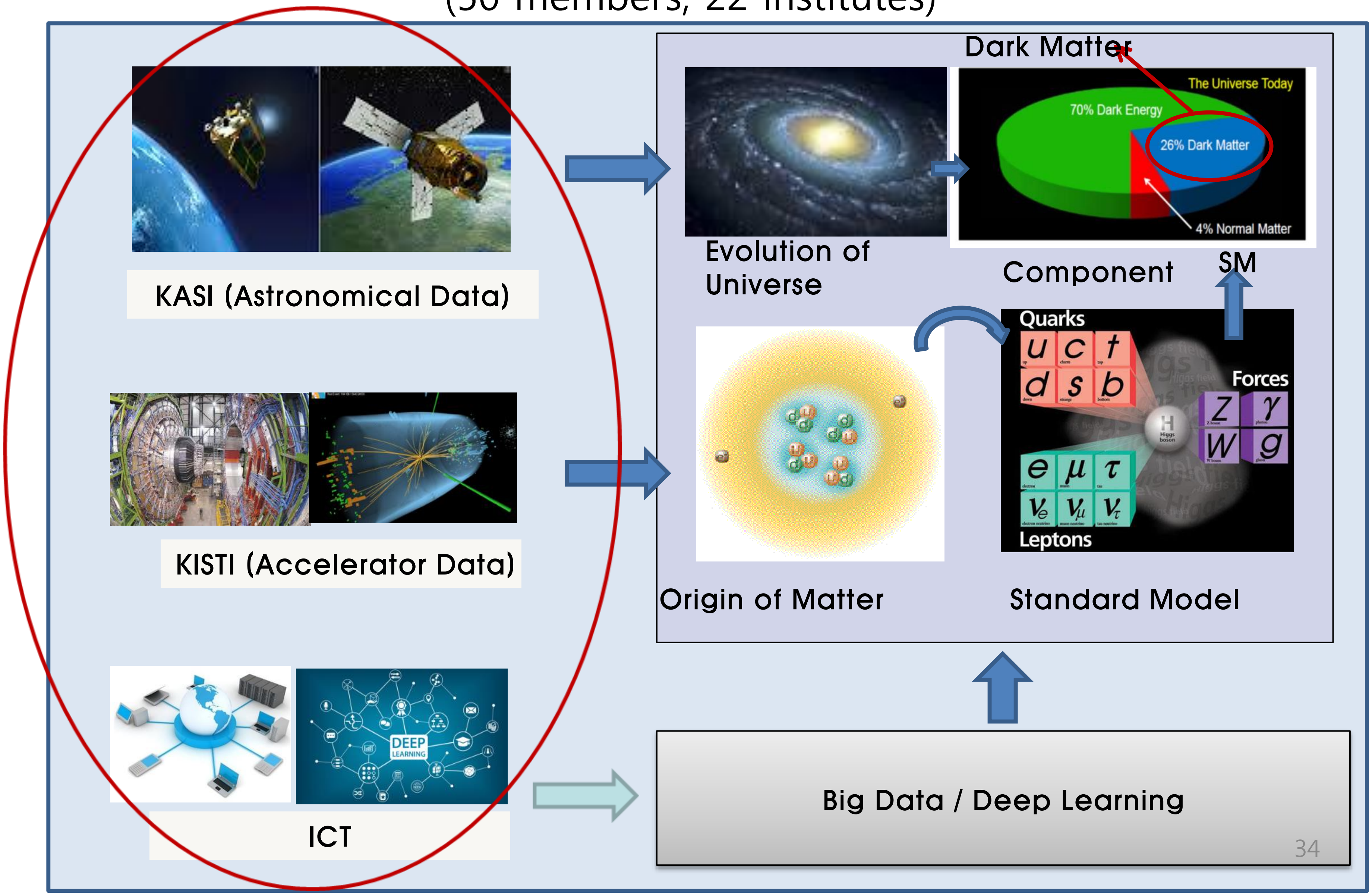






Dark Matter Research Cluster

(50 members, 22 institutes)

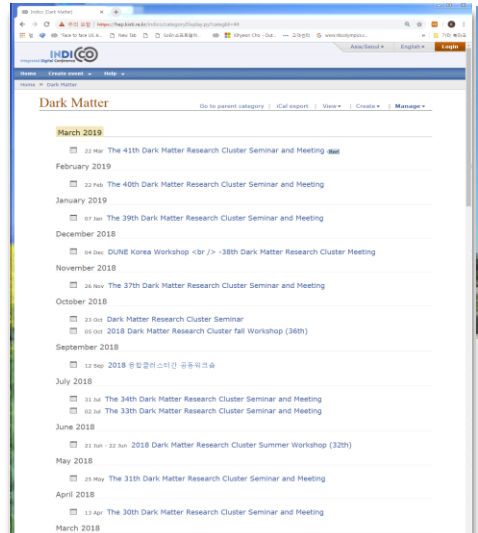


Activities

- Home page
  - Indico
  - https://hep.kisti.re.kr/indico/categoryDisplay.py?categId=44
- Workshops
  - Every month
  - 41 times (2015.9 ~)

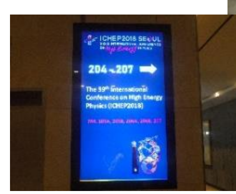
- Dark Matter Research White Paper
  - SCOPUS Journal(KPS) Special Edition (2016.8)

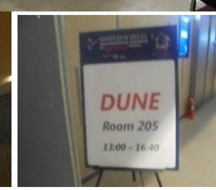
⇒ Proposal to NST
  - Dark matter research platform with deep learning









Dark Matter Research Cluster got the best activity prize (2018.11).

Summary

- Science goes beyond discovery.
- Computing needs solutions for the evolving architecture (e.g. KISTI-5).

⇒ To fulfill the gap between science and computing, we keep to focus on e-Science.

Acknowledgment

- Insung Yeo
- Myeong Hwan Mun

Thank you for your attentions.

(cho@kisti.re.kr)


국가슈퍼컴퓨팅센터


KiSTi
한국과학기술정보연구원

Back-up

KISTI-5 supercomputer

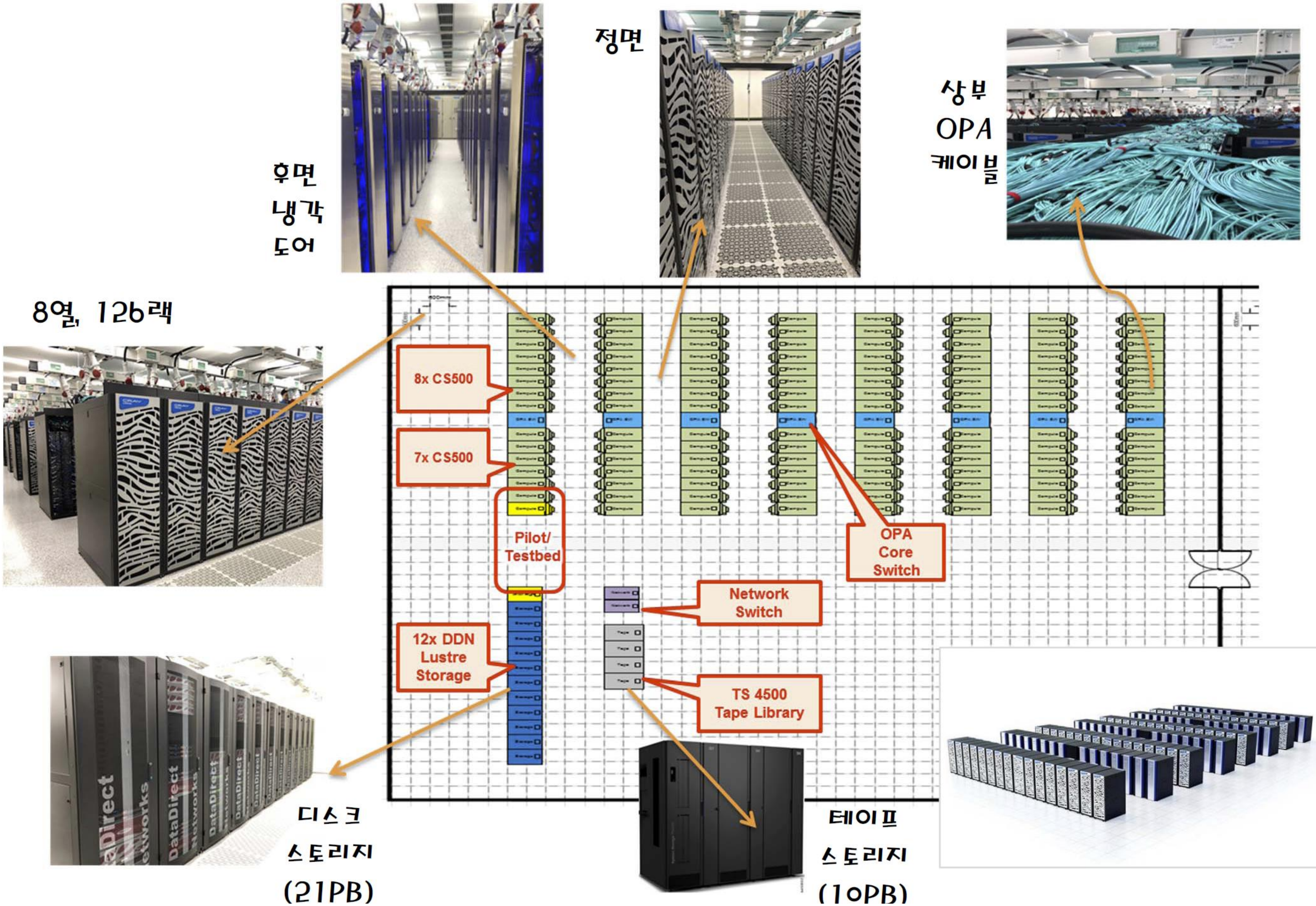
구분	소구분	계산노드	CPU-only 노드
모델명	시스템 모델명	Cray CS400	Cray CS500
	노드수	8,304 +1	132
	랙수	116	2
성능	이론성능	25.297306 PFLOPS	0.405504 PFLOPS
CPU (또는 가속기)	모델명	Intel Xeon Phi 7250 (KNL)	Intel Xeon 6148 (Skylake)
	CPU 이론 성능	3.0464 TFLOPS	1.536 TFLOPS
	CPU당 코어수	68	20
	노드당 CPU수	1	2
	on-package 메모리	16 GB, 490 GB/s	-
HCA	모델명	Intel 100HFA016LS	Intel 100HFA016LS
메인 메모리 (off-package)	메모리 모델	16 GB DDR4-2400	8 GB DDR4-2666
	메모리 구성	16GB x6, 6ch per CPU	8GB x12, 6ch per CPU
	노드 당 메모리 크기(GB)	96 GB	96 GB
	CPU와 메모리 간 대역폭	115.2 GB/s	128 GB/s
	전체 메모리 용량	778.50 TB	12.38 TB
로컬디스크	모델 및 노드당 용량	N/A	N/A
고성능 인터커넥트	토폴로지	FatTree	
	Blocking Ratio	50% Blocking	
	네트워크스위치	274x 48-port OPA edge switches 8x 768-port OPA core switches	
	포트당 단방향 대역폭	12.3 GB/sec	
	Bisectional Bandwidth	27,060 GB/sec	
Burst Buffer (BB)	서버수	DDN IME240서버 40대	
	노드 구성안(NVRAM, SSD)	NVMe SSD	
	전체 용량	0.8PB usable	
	서버당 대역폭	서버당 20 GB/sec, 전체는 0.8TB/s	
대용량 스토리지	**Scratch MDS**	DDN SFA7700X 1대, MDS 2대	
	Home/Apps MDS	DDN SFA7700X 1대, MDS 4대	
	Scratch OSS	DDN ES14KX 9대(각360 x 8TB장착)	
	Home/Apps OSS	DDN ES14KX 1대(각400 x 4TB 장착)	
	전체용량	20PB usable, RAID6 (8D+2P)	
	전체대역폭	0.3TB/s	

www.kisti.re.kr
한국과학기술정보연구원

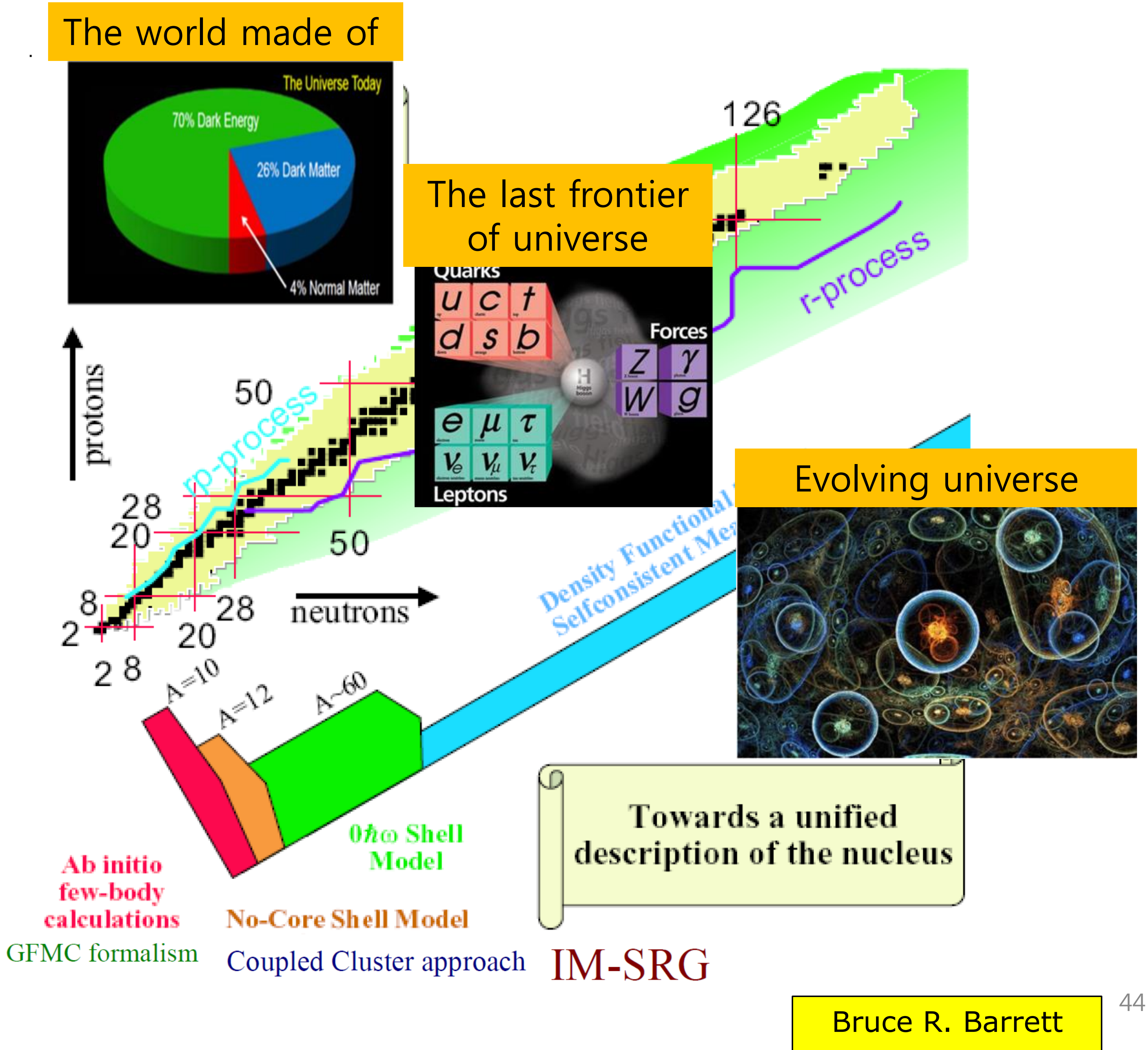
멀티코어프로세서인
인텔 제온(Xeon)
스카이레이크(Skylake)
프로세서를 장착한
CPU 노드

매니코어프로세서인
인텔 제온 파이(xeon Phi)
나이츠랜딩(Knights Landing)
프로세서를 장착한 가속기 노드

0.4 PF
25.3 PF


후면
냉각
도어
정면
상부
OPA
케이블
8열, 126랙
8x CS500
7x CS500
Pilot/
Testbed
OPA
Core
Switch
Network
Switch
12x DDN
Lustre
Storage
TS 4500
Tape Library
디스크
스토리지
(21PB)
테이프
스토리지
(10PB)

M.S. Joh



The world made of
The Universe Today
70% Dark Energy
26% Dark Matter
4% Normal Matter
The last frontier of universe
Quarks
u c t
d s b
Forces
Z γ
W g
H
e μ τ
ν_e ν_μ ν_τ
Leptons
Evolving universe
126
r-process
rp-process
protons
neutrons
50
28
20
8
2
50
28
20
2 8
A=10
A=12
A~60
Density Functional
Selfconsistent Mea
Towards a unified description of the nucleus
Ab initio few-body calculations
No-Core Shell Model
0ħω Shell Model
GFMC formalism
Coupled Cluster approach
IM-SRG
Bruce R. Barrett

Evolving Universe w/ KISTI-5

1. Production of exotic nuclei and heavy elements
2. Nuclear Structure and reactions from first principle
3. Lattice Effective Theory