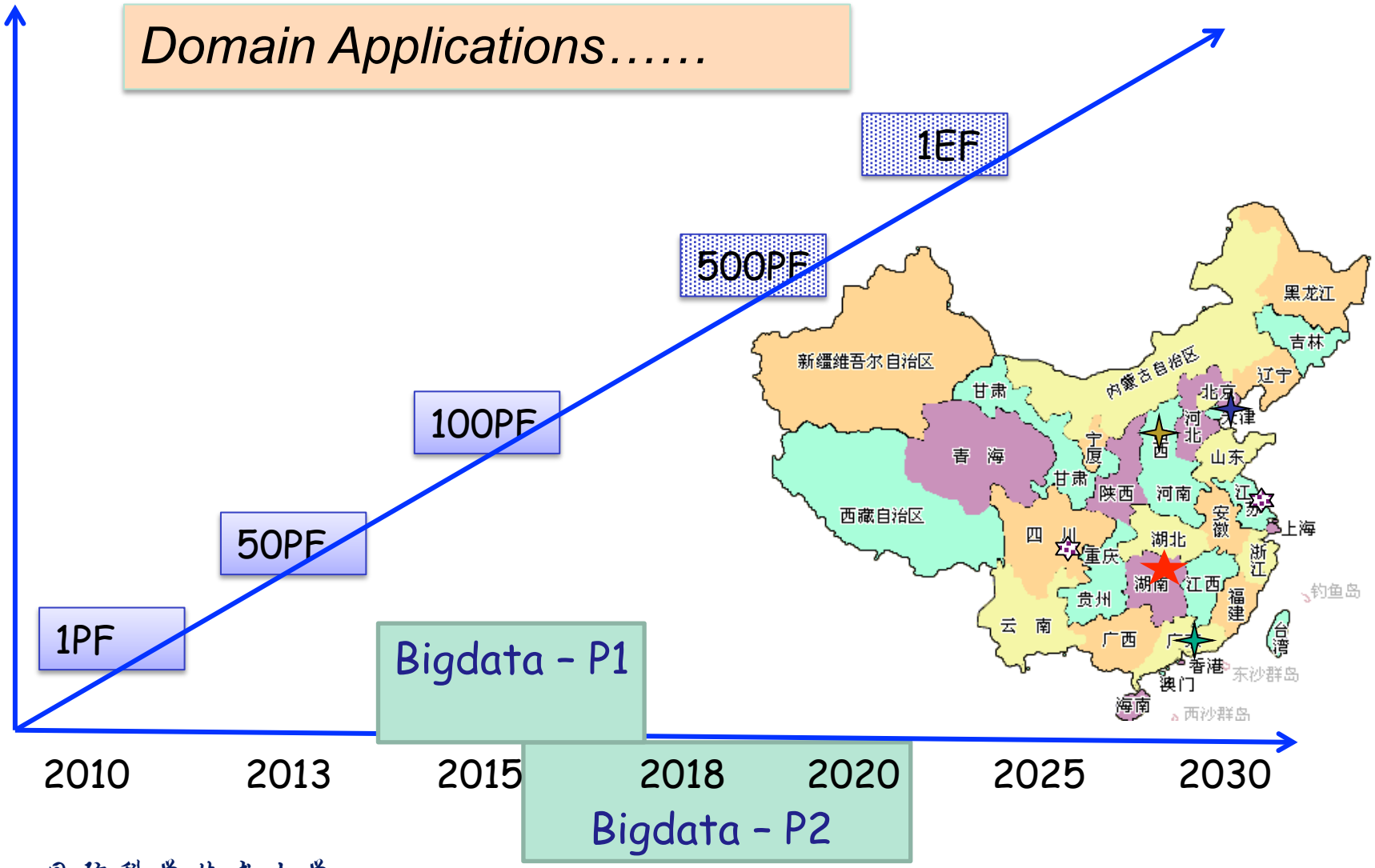
Status of HPC & Big data

Yutong Lu

School of Computer Science, NUDT
State Key Laboratory of High Performance Computing, China
ytlu@nudt.edu.cn

Overview


Domain Applications……
1EF
500PF
100PF
50PF
1PF
Bigdata - P1
Bigdata - P2
2010
2013
2015
2018
2020
2025
2030
新疆维吾尔自治区
甘肃
内蒙古自治区
黑龙江
吉林
辽宁
北京
天津
河北
宁夏
西
山东
青海
甘肃
陕西
河南
江苏
安徽
上海
西藏自治区
四川
重庆
湖北
浙江
湖南
江西
福建
贵州
云南
广西
台湾
香港
澳门
海南
钓鱼岛
东沙群岛
西沙群岛

Overview

- Infrastructure
  - CPU
  - Interconnect network
  - I/O structure
- System software
  - New concept os
  - Hybrid run-time
  - Programming model & framework
  - Tools(monitor, debug, performance)
- Data management
  - File system
  - Workflow
  - Visualization
- Domain Applications

Architecture

HPC system + Big-data system

SSD, NVM, Hybrid hierarchy storage

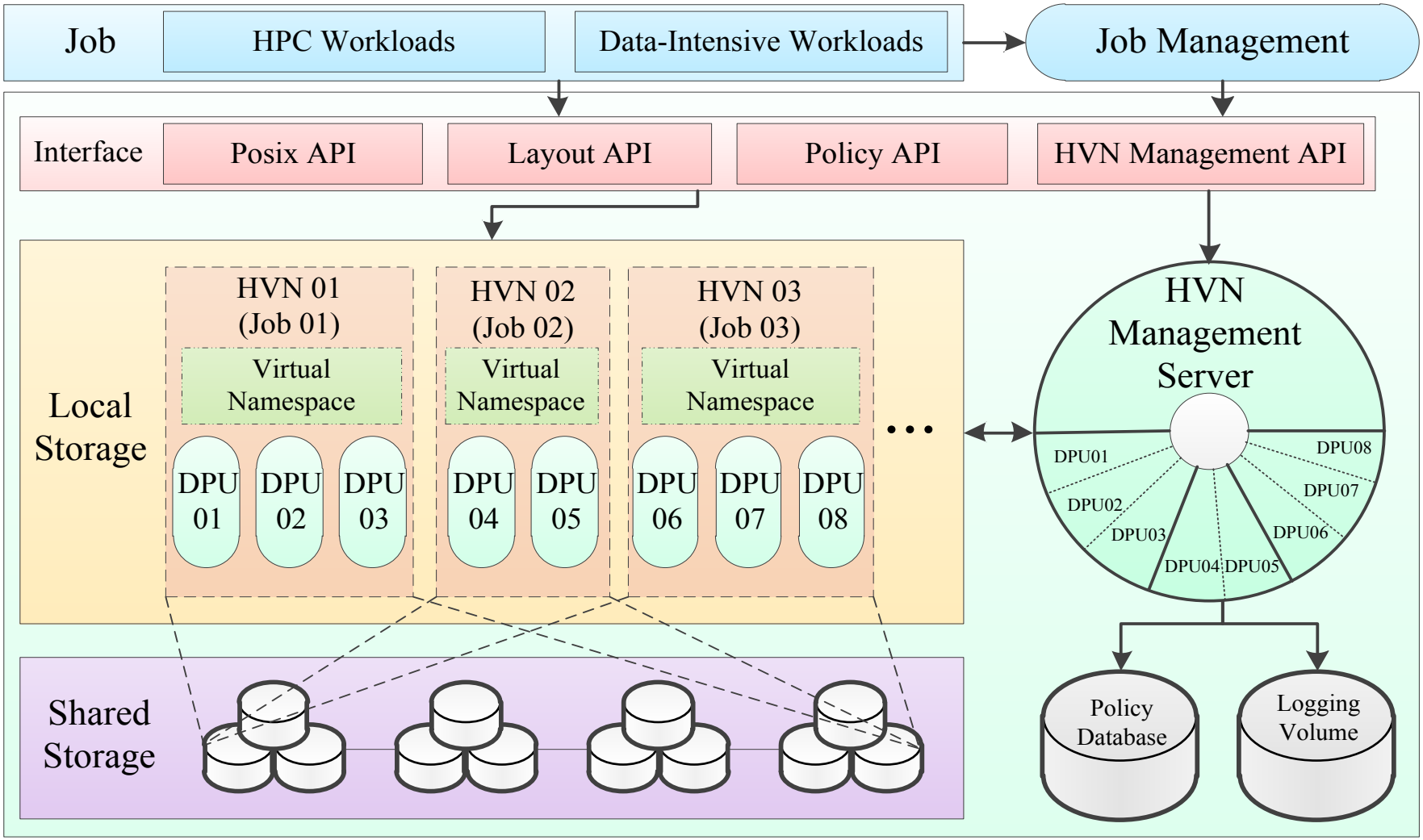
Custom **TH-Net >400Gb/s**	**Commercial** **Infiniband ……**

Custom **Multiple core** **Many core** **5, 10, 30 GF/w**	**FT** **Loongson** **SW** **……**	**Commercial** **Multiple core** **Many core**	**CPU** **Phi** **GPU** **……**

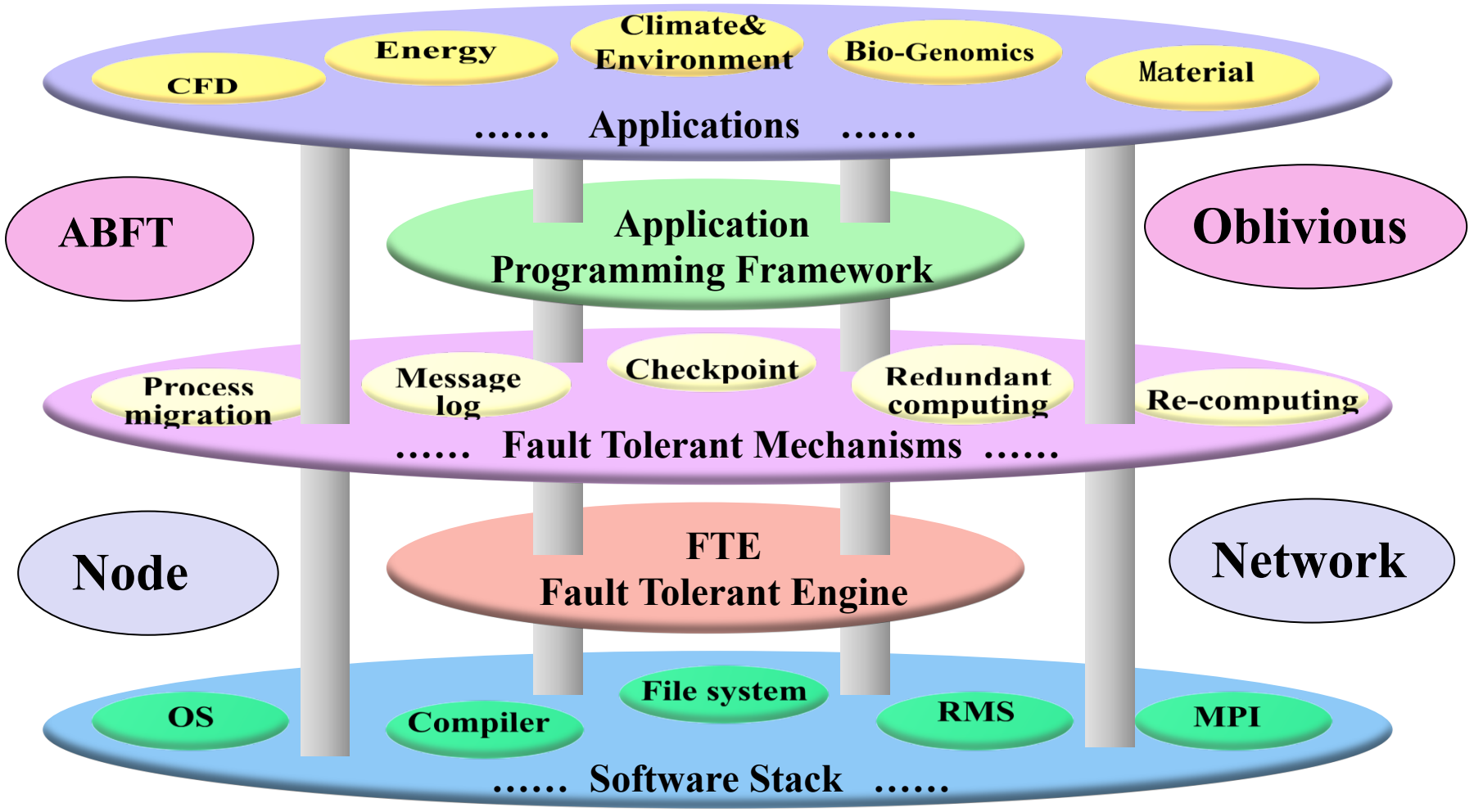
Software

Domain Framework JASMIN, JAUMIN, JCOGIN, Visualization		
File System	Data management	Workflow
Programming model	Hybrid Runtime	Tools Monitor, debug, perf
New concept OS		Kylin KHF-Project

H2FS & Workflow



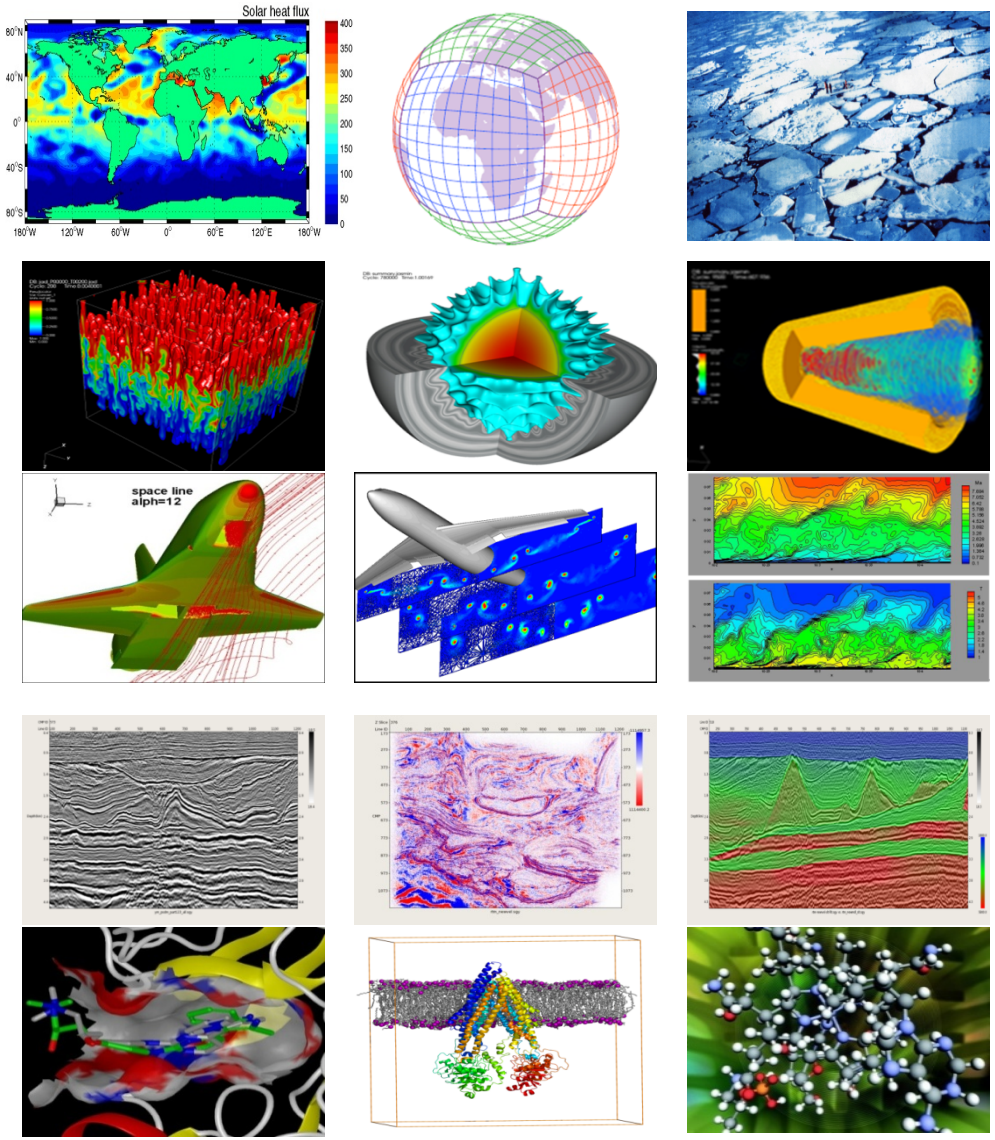
Fault Tolerance



Domain Application

Strategic domains

- Climate & Env.
- Fusion
- Aircraft Design
- Space & Cosmic
- Bio-Genomic
- Mechanical engineer
- New Material
- Electromagnetism
- Animation
- Oceanography



Big Data

- **Big Data Project (Phase 1)**
  - **In-memory computing system**
    - **Persistent memory**
    - **Hybrid memory management policy**
    - **Parallel system**
  - **Key technology and system of Artificial Intelligence**
    - **Deep learning, Knowledge acquisition, Content understanding, Problem solving, Interactive quizzes,**
    - **Prototype system: humanoid answer**

Funding Strategy

- Funding system changing
  - NSFC
    - Basic algorithms and computable modeling for high performance scientific computing
      - Novel computational methods and basic parallel algorithms
      - Computable modeling for selected domains
      - Implementation and verification of parallel algorithms by simulation
    - Network based research environment
    - Many-core parallel programming
    - Big data

Funding Strategy

MOST

- Multiple Grand Projects
  - Domain-centric

The 13th Five-Y Plan…(2016 ~ 2020)

- Infrastructure
- Software & Application
- International Joint Project