

CSTCloud & EGI Cooperation: towards GOSC

Jianhui Li

Computer Network Information Center, CAS

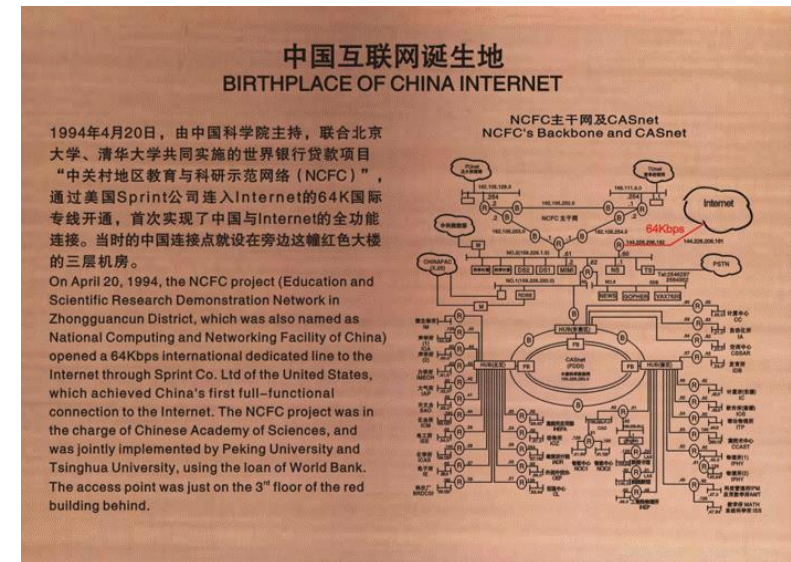
Sep.2022

CNIC Introduction

- Computer Network Information Center (CNIC) is a special research institute of Chinese Academy of Sciences (CAS) to power scientific research and innovation, advance scientific research management, and promote public science outreach by advanced technologies. CNIC is the cradle of China's digital transformation and open science with data, network, computing and public outreach initiatives.

- ❖ 1986: CAS data and information system
- ❖ 1992: China's first router
- ❖ 1994: China's full-functional connection to the Internet
- ❖ 1994: China's first domain name service
- ❖ 1995: established by founding members from above
- ❖ 1996: CSTNet | supercomputing service
- ❖ 1999: open science (knowledge) portal for the public

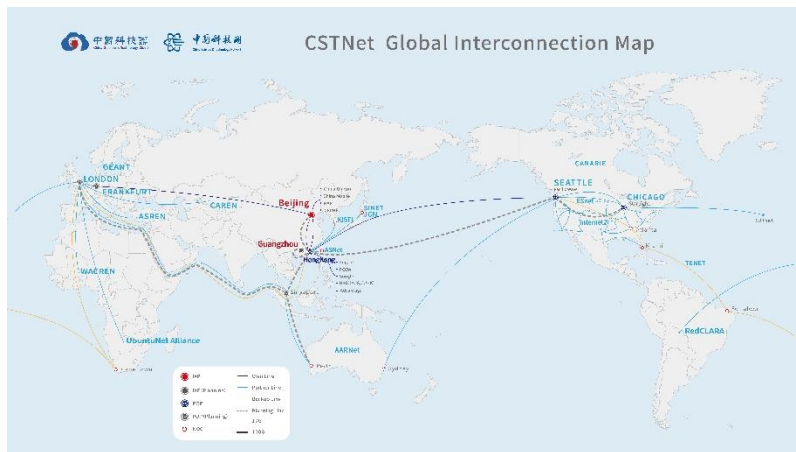
Hello Science | Hello World



E-Infrastructures in CAS



CSTNet has 12 Branches across China, and provides services to approx. 1MM. end users from 300 institutions. Her domestic bandwidth is 131G and international bandwidth is 113G. Below shows the CSTNet Global Interconnection.



NETWORK

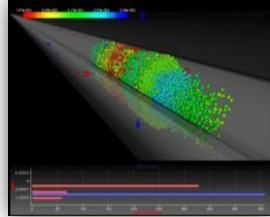
COMPUTING

Chinese National Grid

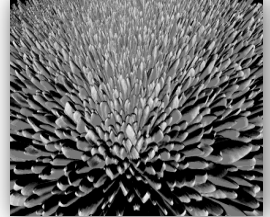
CNGrid has connected 19 supercomputing centers, the capability of clustering computing has reached 200PF, supported for thousands of national science and technology program and key engineering. 40+ independent high performance computing software.



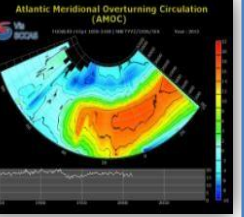
CAS Supercomputing Grid environment creates a three-layer architecture environment and has integrated more than 15PF computing resources from 35 CAS institutions



Visualization of ADS transmutation system



Simulation of optimization of titanium alloy structure



Simulation of the Atlantic ocean circulation

20 National Scientific Data Centers are launched by the Ministry of Science and Technology, P.R.C as the pilot to improve research data sharing following the national rules, “Measures for Managing Scientific Data”. Among all of which, 11 are from the Chinese Academy of Sciences.

No.	National Scientific Data Center		
1	National high energy physics science data center		
2	National genome science data center		
3	National microbial science data center		
4	National space science data center		
5	National astronomical sciences data center		
6	National earth science data center	No.	National Scientific Data Center
7	National polar science data center	11	National glacial frozen deserts science data center
8	National Qinghai science data center	12	National meteorological science data center
9	National ecology science data center	13	National earth system science data center
10	National material science data center	14	National population health science data center
		15	National basic sciences public science data center
		16	National agricultural science data center
		17	National forestry and grassland science data center
		18	National meteorological science data center
		19	National earthquake science data center
		20	National Marine science data center

OPEN RESEARCH DATA

CSTCloud Introduction

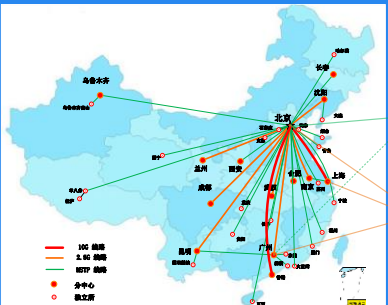
- As one of the key national e-infrastructures, the development of CSTCloud started in 2017.1 and continually funded by CAS(2017-2022, 2021-2025).
- CSTCloud fully supports multidisciplinary open scientific researches with integrated cloud services for the discovery, usage and delivery of S&T resources.



Pool of research softwares



Data & Information



High-speed network connection



Algorithms & tools

CSTCloud



Computing resources



Massive data storage

Targeted Users & Tailored Services

Researcher

- Service discovery
- One-stop access
- Open and Use

Service Provider

- Service Registration
- Demonstration & Promotion
- Management & Transaction

CST Cloud

Resource Owner

- Service monitor
- Quality statistics
- Comprehensive Evaluation

For the **100,000** scientific researchers in CAS, providing a cloud service environment that supports innovation in multidisciplinary fields, supporting the scientific discovery of big data and big computing

CSTCloud: Services Categories

NATIONAL/REGIONAL/INSTITUTIONAL DATA CENTERS

DISCIPLINARY SCIENCE CLOUDS

THEMATIC DEMONSTRATIONS

SAAS



New-Gen ARP



Website Group



Science Expo



CAS Email System
(152 institutes, 415,000 users)



CAS Cloud Conferencing
(hosted 65,000 meetings, 567,000 participants)



Geospatial Data Cloud Service
(470,000 users, 790TB storage)



Scientific Open Source Software Community
(1,000+ software)

PAAS



CSTCloud Passport
(1.12 million users, 845 applications)



Cloud Backup & Archive Platform
(respectively 60PB & 3PB storage)



Scientific Data of Chinese Academy of Sciences
(11 national centers, 3.2PB storage)



National Grid & CAS Grid
(19 centers, 460PF computing power, 310PB storage)



National Public Scientific Data Platform for Basic Discipline



Chinese Node of eduroam
(covers 341 institutes and universities)



IAAS



中国科技网
China Science & Technology Network

China Science & Technology Network
(supports 295 institutes, provides 113G international bandwidth, 121G national bandwidth, 100G core network bandwidth)



Advanced Computing System & Meta AI Platform
(315PF computing power)

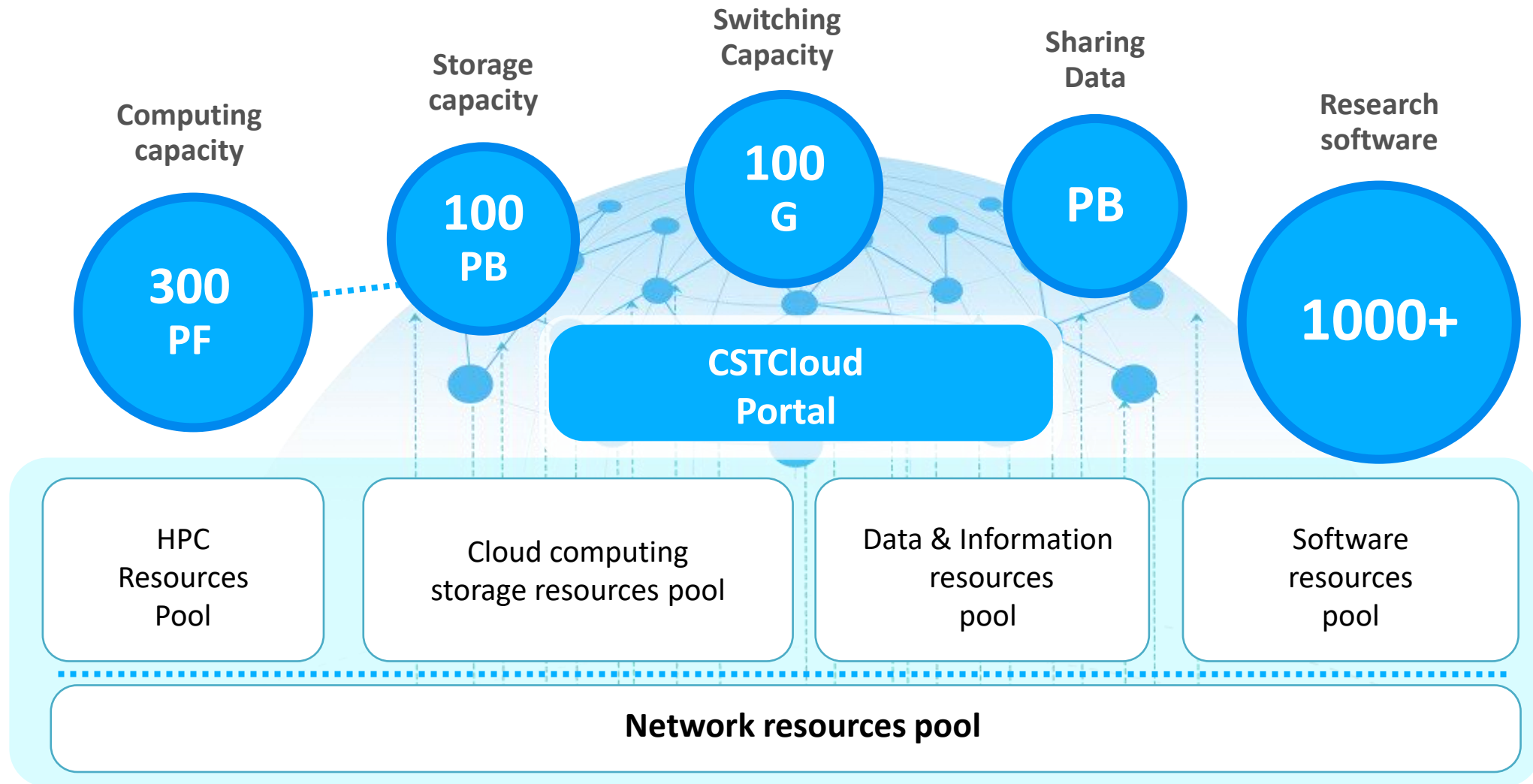


Cloud Computing & Storage Platform
(both public and private, provides 12,000 cores CPU and 150PB storage)

Intelligent operation and maintenance system

Safety control system

CSTCloud: Service Capacity



Typical Use Cases

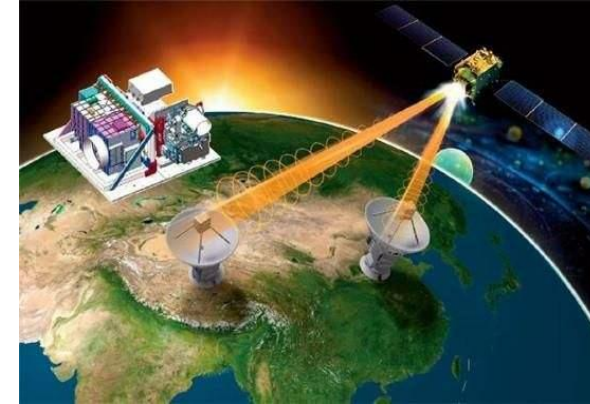
FAST



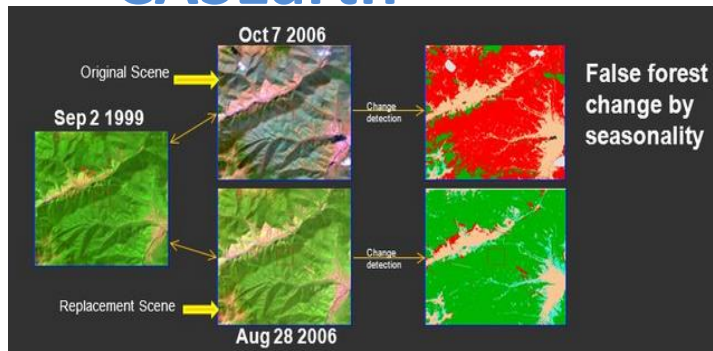
Wukong



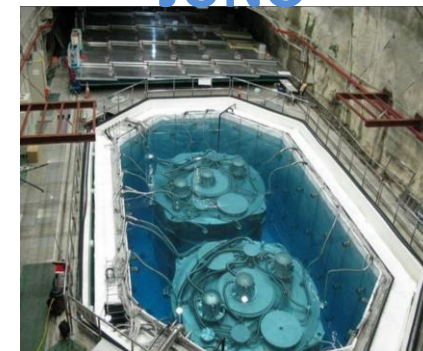
Mozi



CASEarth

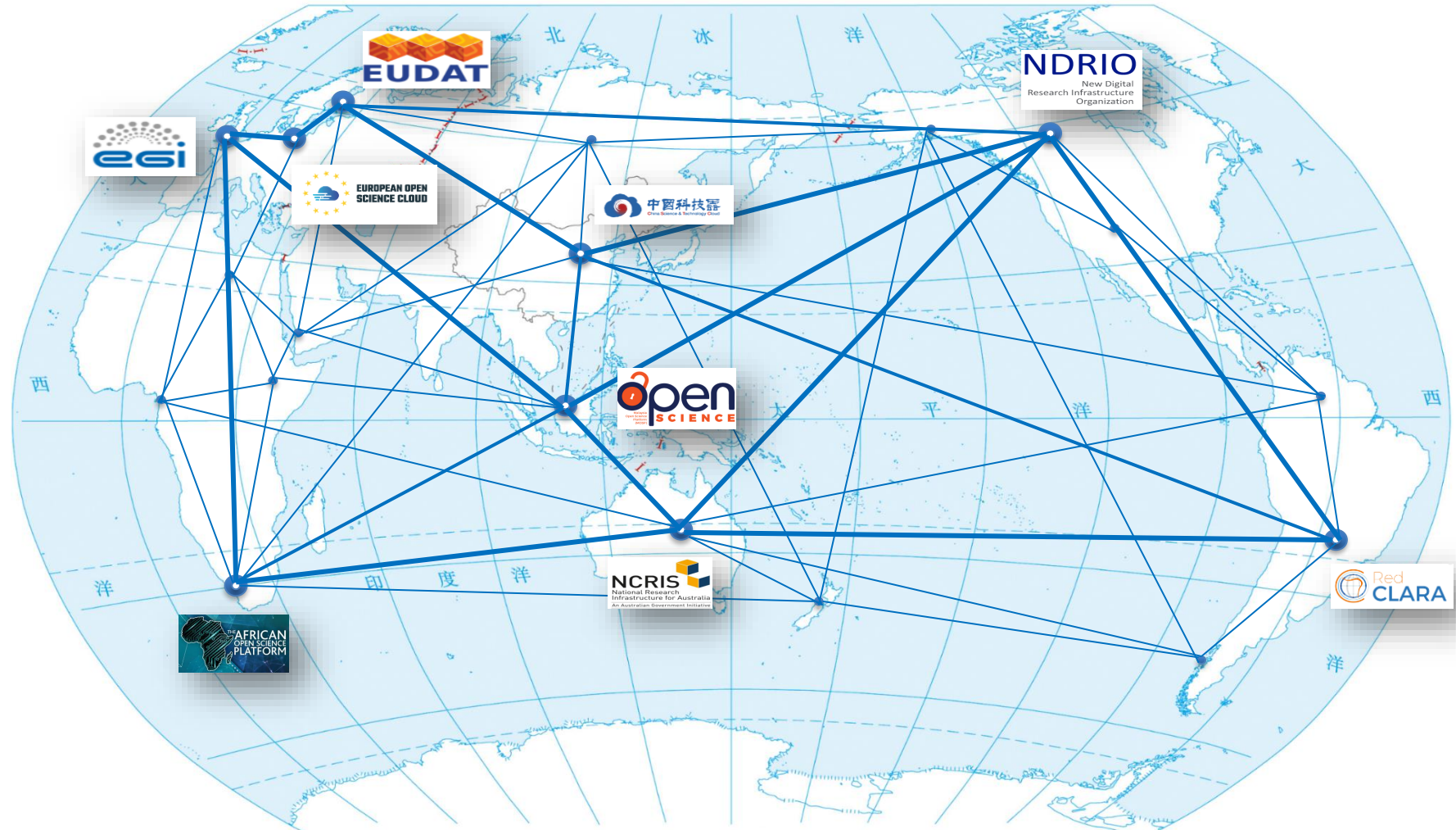


JUNO



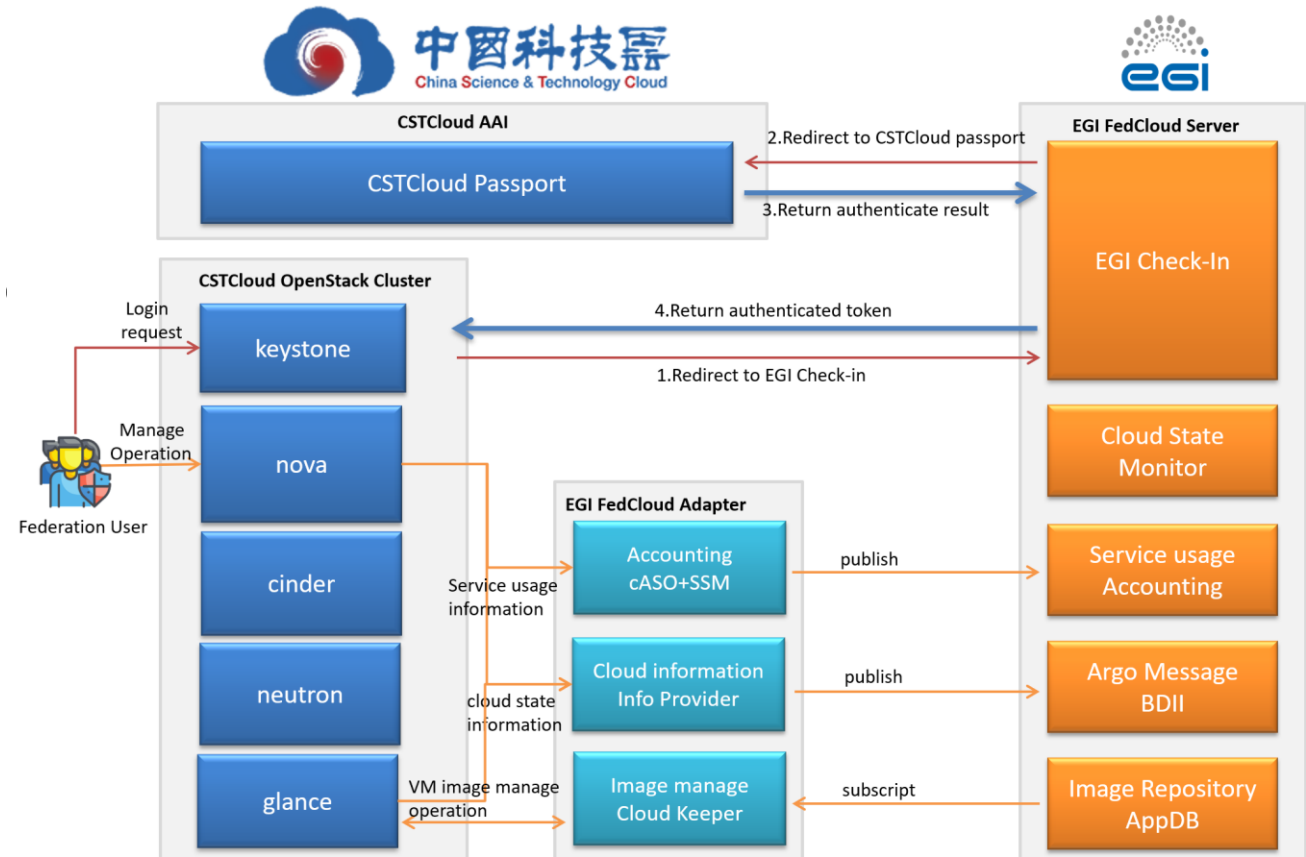
Connecting CSTCloud with the World

- A cross-continental federated e-infrastructure and virtual research environment for global cooperation and open science.



A Cross Continental Federation Testbed

- **CSTCloudAAI** connects EGI Check-in system
- **CSTCloud OpenStack Cluster**
 - ✓ Synchronize VM Images With EGI Fedcloud
 - ✓ Information exchanges with EGI Fedcloud
 - ✓ Unified services status monitoring and metrics
- **CSTCloud federation** to support use case in space physics, biology, SDG research, etc.,



A Cross Continental Federation Testbed

- Technique solution has been verified , ready to provide service
- Shared resource is limited, need more resource...

EGI | EGI-ACE | SERVICES | FEDERATION | USE CASES | BUSINESS

A federation of cloud resources beyond Europe

Integration of China's CSTCloud with the EGI Federation has been recently completed. CSTCloud is a certified provider of the federation and meets all the operational tests for production usage.

Operated by the Computer Network Information Center (CNIC) of Chinese Academy of Sciences (CAS), **CSTCloud** is a national infrastructure for CAS scientific communities and China's top research. The design of the CSTCloud is based on the principle of 'openness and sharing'. It aims to develop an open architecture that is capable of integrating various national and international computing resources in order to support multidisciplinary open science research. CSTCloud provides computing facilities for Chinese advanced research projects including CASEarch, CAS Space Science Missions, and research related with big facilities or observation stations such as the **Five-hundred-meter Aperture Spherical Telescope (FAST)** and the **Large-High-Altitude Air Shower Observatory (LHAASO)**.

The integration work is delivered under the **EGI-ACE** international cloud integration task force. There have been many challenges — different technical environments, different development culture, limited documentation, no previous examples, etc.

It has taken a number of months of effort and people from different organisations and teams are involved. Particularly, we are grateful to Professor Jianhui Li's team in CNIC, including the CNIC Cloud team (Haiming Zhang, Zuliang Guo, and Xiangguang Zheng), the CNIC AAI team (Yihua Zheng and Taotao Shi), and the CNIC Project Management team (Lili Zhang). Thanks to the EGI teams who provide dedicated supports, including the

Secure AAI

Check-in

Choose your academic/social account

Search...

- CROUS Strasbourg V3
- CROUS Toulouse - Personnels
- CROUS Versailles V3
- Croydon College
- CSC - IT Center for Science Ltd.
- CSIR
- 中国科技云通行证(CSTCloud_ID)
- 中国科技云身份认证联盟IDP
- CTGAS-ER - Centro de Tecnologias do Gas e Energias Renovaveis
- CTI "Diophantus"
- CTI - Centro de Tecnologia da Informacao Renato Archer
- CUDI

Cloud federation virtual appliance

openstack. Default admin

Project / Compute / Overview

Overview

Limit Summary

Compute

- Instances: Used 1 of 40
- VCPUs: Used 4 of 160
- RAM: Used 8GB of 50GB

Volume

- Floating IPs: Allocated 0 of 50
- Security Groups: Used 1 of 10
- Security Group Rules: Used 6 of 100
- Networks: Used 1 of 100
- Ports: Used 2 of 500
- Routers: Used 0 of 10

Current Network Status

Last Updated: Thu Oct 21 12:19:28 CEST 2021
 Updated every 90 seconds
 Nagios® Core™ 4.4.5 - www.nagios.org
 Logged in as /C=CN/O=HEP/O=CNIC/OU=CC/CN=Zuliang Guo

View History For This Host
 View Notifications For This Host
 View Service Status Detail For All Hosts

Host Status Totals

Up	Down	Unreachable	Pending
1	0	0	0

All Problems All Types

0	1
---	---

Service Status Totals

Ok	Warning	Unknown	Critical	Pending
4	0	0	0	0

All Problems All Types

0	4
---	---

Service Status Details For Host 'federation.cstcloud.cn'

Limit Results: 100

Host	Service	Status	Last Check	Duration	Attempt	Status Information
federation.cstcloud.cn	eu.egi.cloud.APEL-Pub	OK	10-21-2021 11:29:02	49d 13h 17m 43s	1/3	FedCloud Accounting Freshness OK - Accounting data for site CSTCLOUD-EGI found. Last update occurred 2021-10-21 04:30:22.
	eu.egi.cloud.InfoProvider	OK	10-21-2021 12:12:58	5d 21h 6m 59s	1/2	OK: Endpoint publishing up to date information for VOs
	eu.egi.cloud.OpenStack-VM-ops	OK	10-21-2021 12:17:12	5d 12h 6m 2s	1/2	OK: Compute instance=9d4d180c-ab63-4f08-a6ea-eeef1dd8bd8a created(30.44s) and destroyed(10.71s)
	org.nagios.Keystone-TCP	OK	10-21-2021 12:18:33	27d 7h 3m 29s	1/3	TCP OK - 0.171 second response time on federation.cstcloud.cn port 5000

Results 1 - 4 of 4 Matching Services

Monitoring

Edit Quotas

100GB of Memory

30 Cores of vCPU

VCPUs: 30

RAM (MB): 102400

Resources

A Typical Use Case – 3D Radar

Incoherent scatter radar data fusion and computation

EISCAT-3D radar, next generation incoherent scatter radar system, EISCAT association.

Sanya Incoherent Scatter Radar (SYISR), next generation incoherent scatter radar, IGGCAS.

Use Case Scenarios

EISCAT & SYISR (Meta)data federation

Federated processing

On-demand data movement

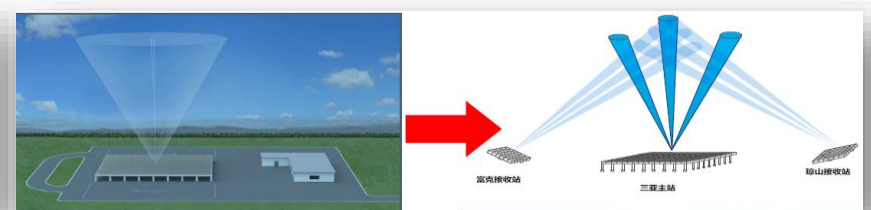
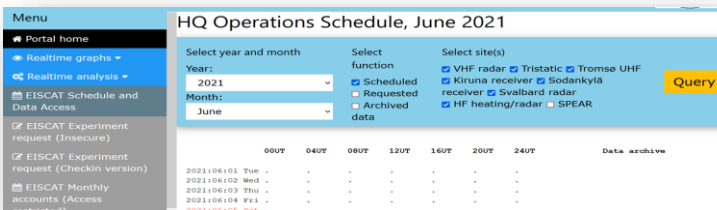
Key deliverables

Cross-continental data migration and federated computing; high-speed and high-capacity experimental data processing; data sharing policies and mechanisms.

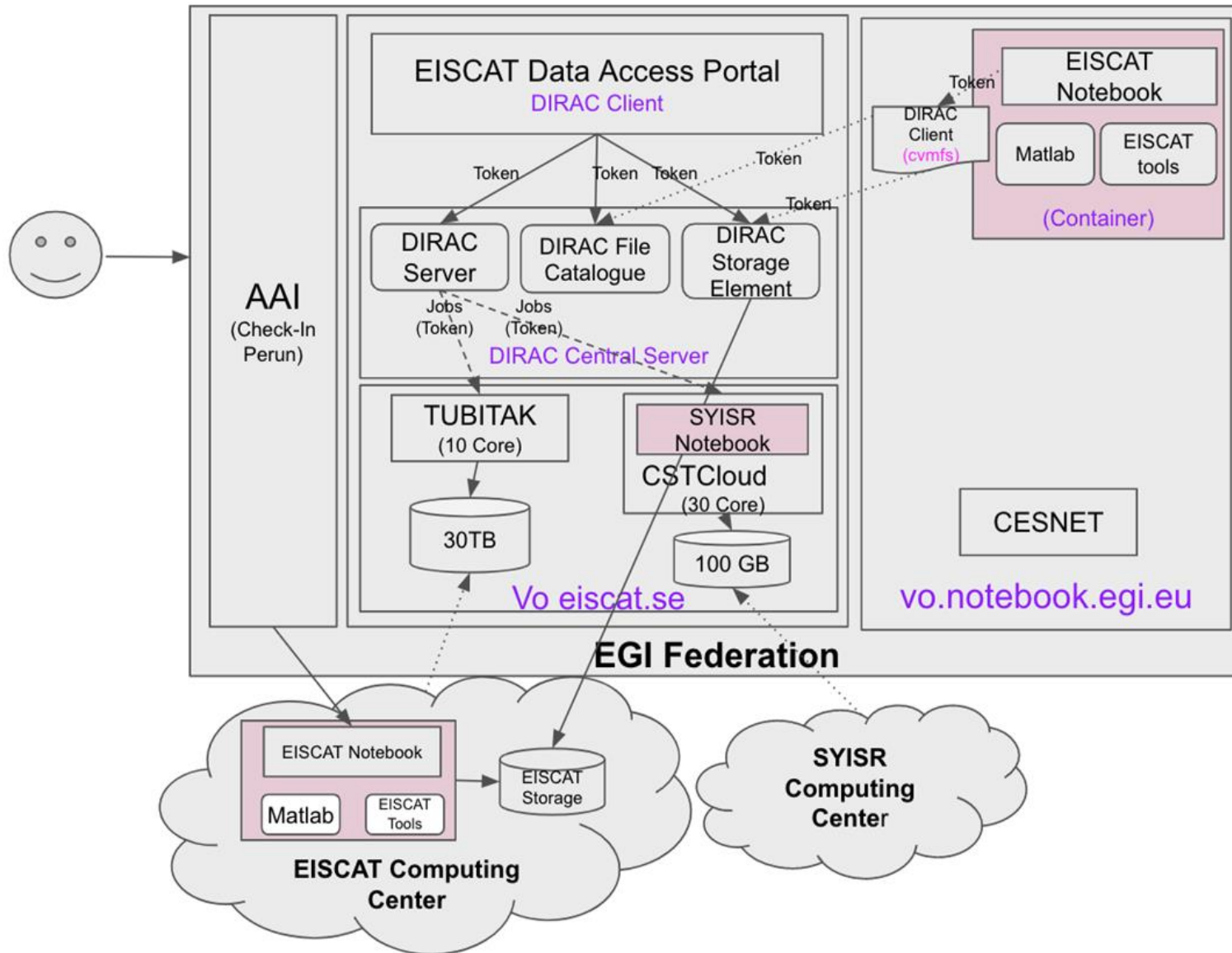


EISCAT_3D/SYISR radar data

- Similar hardware
 - Multistatic phased array radars
 - Separate repositories
 - Metadata federation
 - Federation processing
 - Launch jobs in EGI/CNIC clouds
 - On demand data movement
- For processing in either cloud



A Typical Use Case – Technical Framework



• EISCAT Data Access Portal (DIRAC)

- User access via Check-In/Perun
- Access token is passed to DIRAC File Catalogue (by DIRAC Client)
- DIRAC File Catalogue return token with user information
- (DIRAC Client) access DIRAC Storage Element with token
- DIRAC Storage element enables search of EISCAT data (in the EISCAT storage)
- DIRAC Server submits Jobs to Cloud
 - TUBITAK (10core, 30TB) + CSTCloud (30Core, 100GB)
 - Token access to the Cloud is ongoing

• EISCAT Compute Center

- Jupyter notebook
 - User access via Perun
 - Matlab enabled
 - Not yet access to DIRAC Storage Element – cannot perform search, and have lower performance
- EISCAT Storage: EISCAT data

• EISCAT Notebook Container

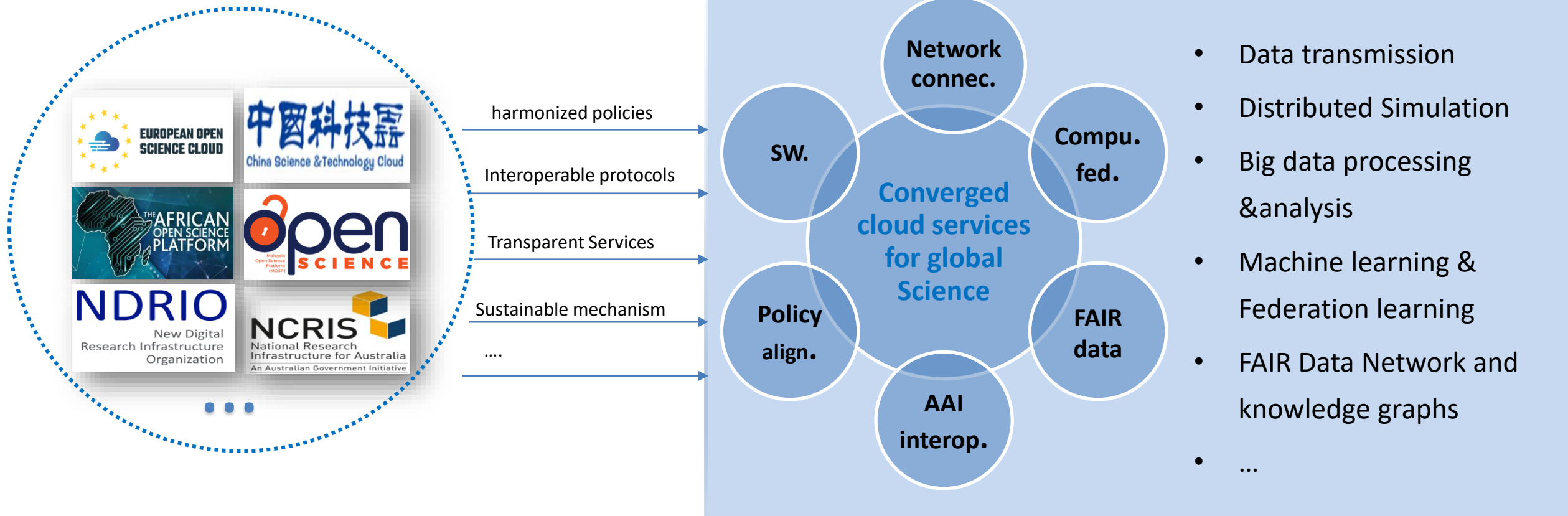
- Enable EISCAT user access
- Notebook+DIRAC image
 - dirac.egi.eu cvmfs repository is configured alongside eiscat cvmfs repository (eiscat.egi.eu)
 - DIRAC client is available via cvmfs
 - pass user access token to the DIRAC File Catalogue to get user information
 - Access the DIRAC Storage Element to search and retrieve data from the EISCAT storage
- Will include Matlab
 - Token to use EISCAT license?
- Will include EISCAT Tools

Towards GOSC

- Open science infrastructure features: **Federated, Accessible, Internationally Interconnected, Interoperable** (UNESCO, 2021).
- The GOSC Initiative is to help connect various institutional, national, and regional initiatives, laying the foundations for cross-continental, federated, Open Science and FAIR infrastructure, and virtual research environments (CODATA GOSC SG, 2021).
- And the expected deliverables are to establish a **robust network of trusted research e-infrastructures** to connect **research resources** and **all stakeholders** to enable innovative science discovery in the dynamically evolving global open science environment. (CODATA GOSC SG, 2021)

Cooperation towards GOSC

- A cross-continental federated e-infrastructure and virtual research environment for global cooperation and open science.

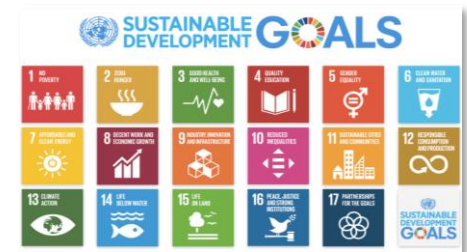


Call for International Partnerships

- Open mind, open framework, open collaboration for open data & open science.
- **Coming together is a beginning. Keeping together is progress. Working together is success.** *Henry Ford (1863 – 1947, American industrialist)*



**International
Science Council**



Thank you for your attention!

Computer Network Information Center, CAS

www.cstcloud.net