FAST Radio Telescope
Data Processing Platform and Data Management Plan

Jun Han, Chenzhou Cui

National Astronomical Observatories, Chinese Academy of Sciences

October 11, Groningen, the Netherlands

The **F**ive-hundred-meter **A**perture **S**pherical radio **T**elescope (**FAST**)

Firstly proposed by Prof. Rendong Nan in year 1994

Located in Guizhou Province, China

Construction completed in Sep. 25, 2016

In commissioning stage

Commensal Radio Astronomy FAST Survey
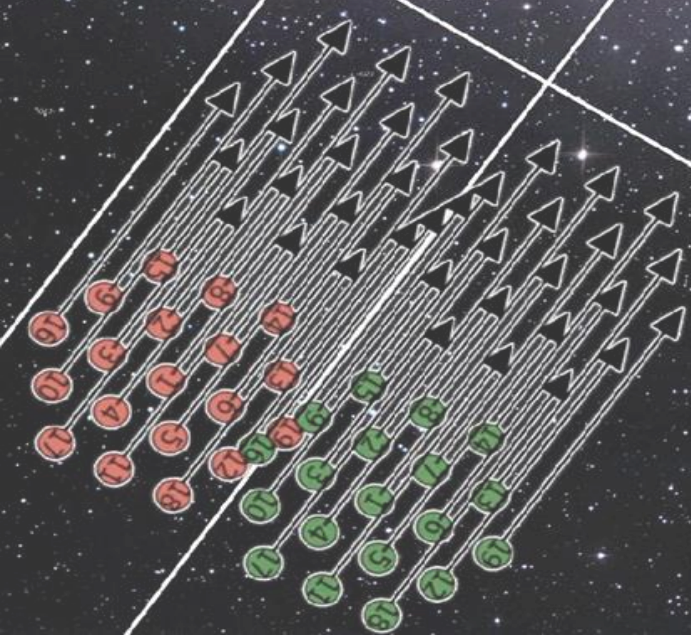
Up to now, ~ one hundred pulsars were found

L-band
19 Beam

- 1.6 GB/s
- 5.8 TB/h
- 144 TB/day
- 10–20 PB/ year

CRAFTS Project with 19 beams

- ∼1000 Pulsars
- >10 billion voxel HI map
- >50 FRBs
- HI galaxies……

CRAFTS
The Commensal Radio Astronomy FAST Survey
FAST 多科学目标同时扫描巡天

Data Process Platform

Kubernetes is an engine for automating deployment, scaling, and management of containerized applications.

The slave number	Master
< 5	4C8G
<20	4C16G
<100	8C32G
<200	16C64G
MORE	Multi-master、SSD…

The resource requirements for master in low load

- ✓ Single master
- ✓ Calico used as network layer
- ✓ metric-server, arena to monitor CPU, GPU, Memeory…
- ✓ A registry to manage images
- ✓ Label every node as nodeSelector
- ✓ Allocate resource for users through RBAC
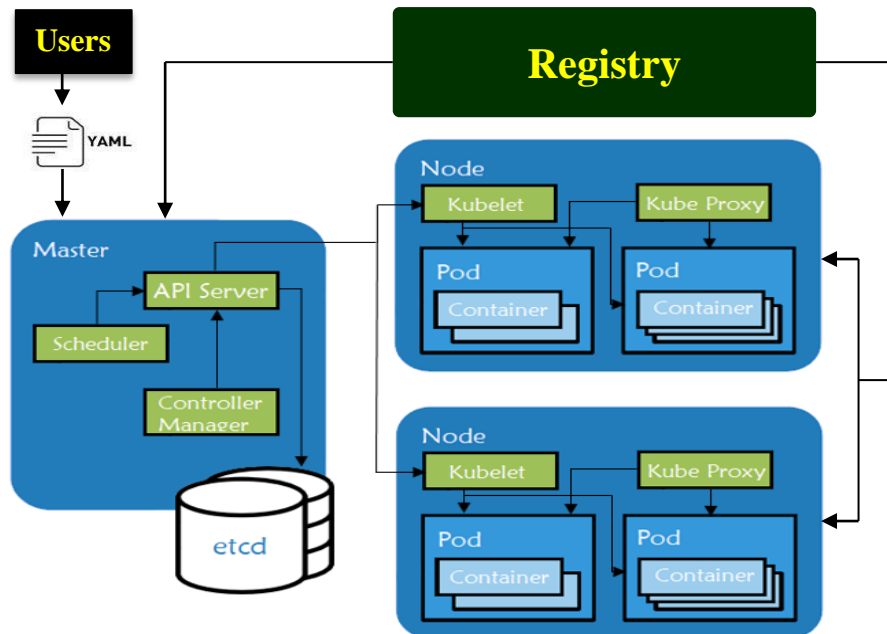
Kubernetes version	Release month	End-of-life-month
v1.6.x	March 2017	December 2017
v1.7.x	June 2017	March 2018
v1.8.x	September 2017	June 2018
v1.9.x	December 2017	September 2018
v1.10.x	March 2018	December 2018
v1.11.x	June 2018	March 2019
v1.12.x	September 2018	June 2019
v1.13.x	December 2018	September 2019
v1.14.x	March 2019	December 2019
v1.15.x	June 2019	March 2020

Data Process Platform

An engine for automating deployment, scaling, and management of containerized applications.

Dockerfile

```
FROM centos:latest

MAINTAINER hanjun@nao.cas.cn

RUN yum install epel-release  -y

RUN mkdir /root/Pulsar

ENV ASTROSOFT            /root/Pulsar
ENV C_INCLUDE_PATH        $C_INCLUDE_PATH:$PGPLOT_DIR:$ASTROSOFT/include
ENV CPLUS_INCLUDE_PATH    $CPLUS_INCLUDE_PATH:$PGPLOT_DIR:$ASTROSOFT/include

WORKDIR $ASTROSOFT

ADD cfitsio_latest.tar.gz  .
ADD fftw-3.3.8.tar.gz       .

WORKDIR $ASTROSOFT/cfitsio
RUN ./configure --prefix=$ASTROSOFT CFLAGS=-fPIC FFLAGS=-fPIC
RUN make shared
RUN make install
RUN make clean


WORKDIR $ASTROSOFT/fftw-3.3.8
RUN ./configure --prefix=$ASTROSOFT --enable-float --enable-threads --enabl
RUN make
RUN make install
RUN make clean
RUN ./configure --prefix=$ASTROSOFT CFLAGS=-fPIC FFLAGS=-fPIC
RUN make
RUN make install
RUN make clean


WORKDIR $ASTROSOFT/tempo2
RUN ./bootstrap
RUN ./configure --prefix=$ASTROSOFT --with-cfitsio-dir=$ASTROSOFT --with-ff
PGPLOT_DIR" LDFLAGS=-L$PGPLOT_DIR
RUN make && make install
RUN make plugins && make plugins-install
RUN make unsupported
RUN make clean

WORKDIR /root
```

Yaml

```
apiversion: apps/v1
kind: Deployment
metadata:
  name: gpusearch
  labels:
    app: gpusearch
spec:
  replicas: 90
  selector:
    matchLabels:
      app: gpusearch
  template:
    metadata:
      labels:
        app: gpusearch
    spec:
      containers:
      - name: gpusearch
        image:  10.134.1.60:5000/admin/gpusearch:4.0
        imagePullPolicy: Always
        resources:
          requests:
            cpu: "20000m"
            memory: "100Gi"
          limits:
            nvidia.com/gpu: 1
        command: ["/work/driftscan/run.sh"]

        volumeMounts:
          - mountPath: /data
            name: data
      volumes:
      - name: data
        hostPath:
          path: /data31/

      nodeSelector:
        disk: lustre
```

Data Process Platform
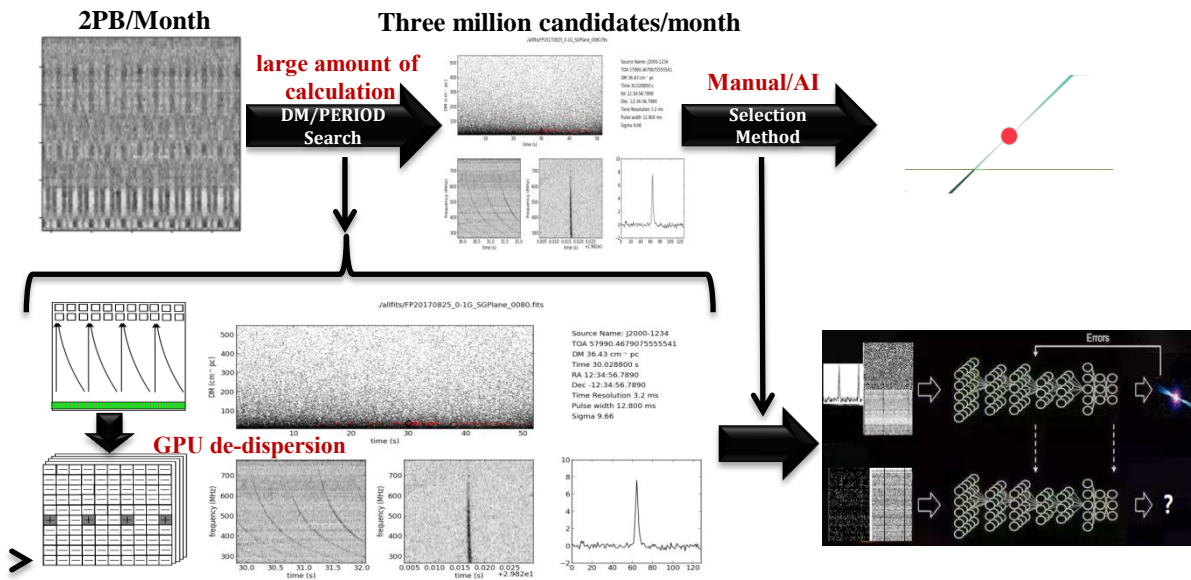
The challenges from pulsar search
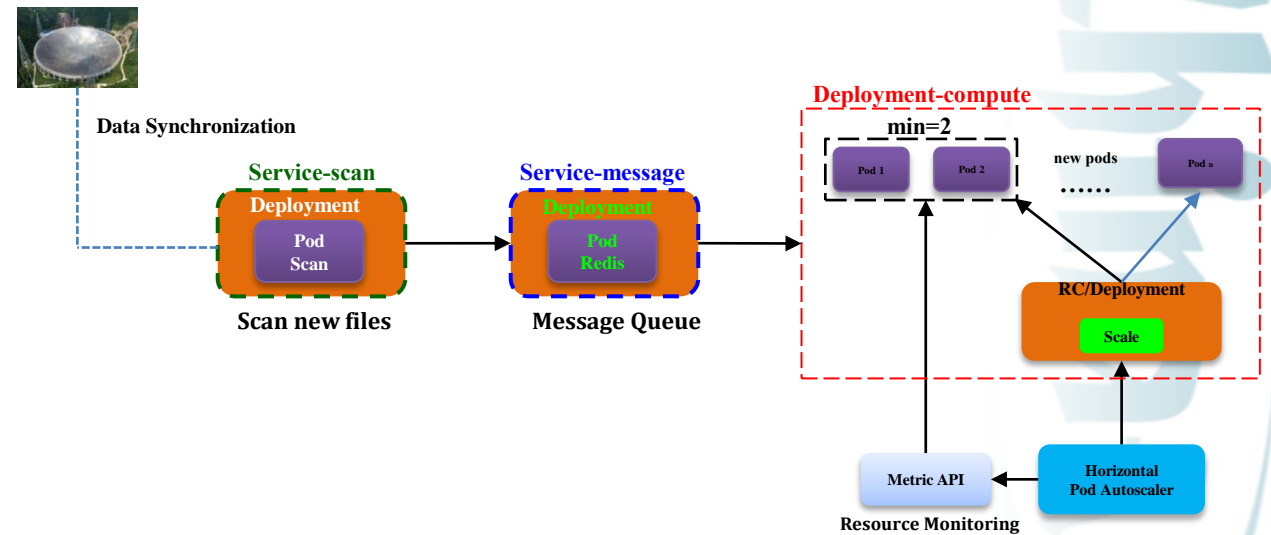
Message Queue--Elastic Computing Model
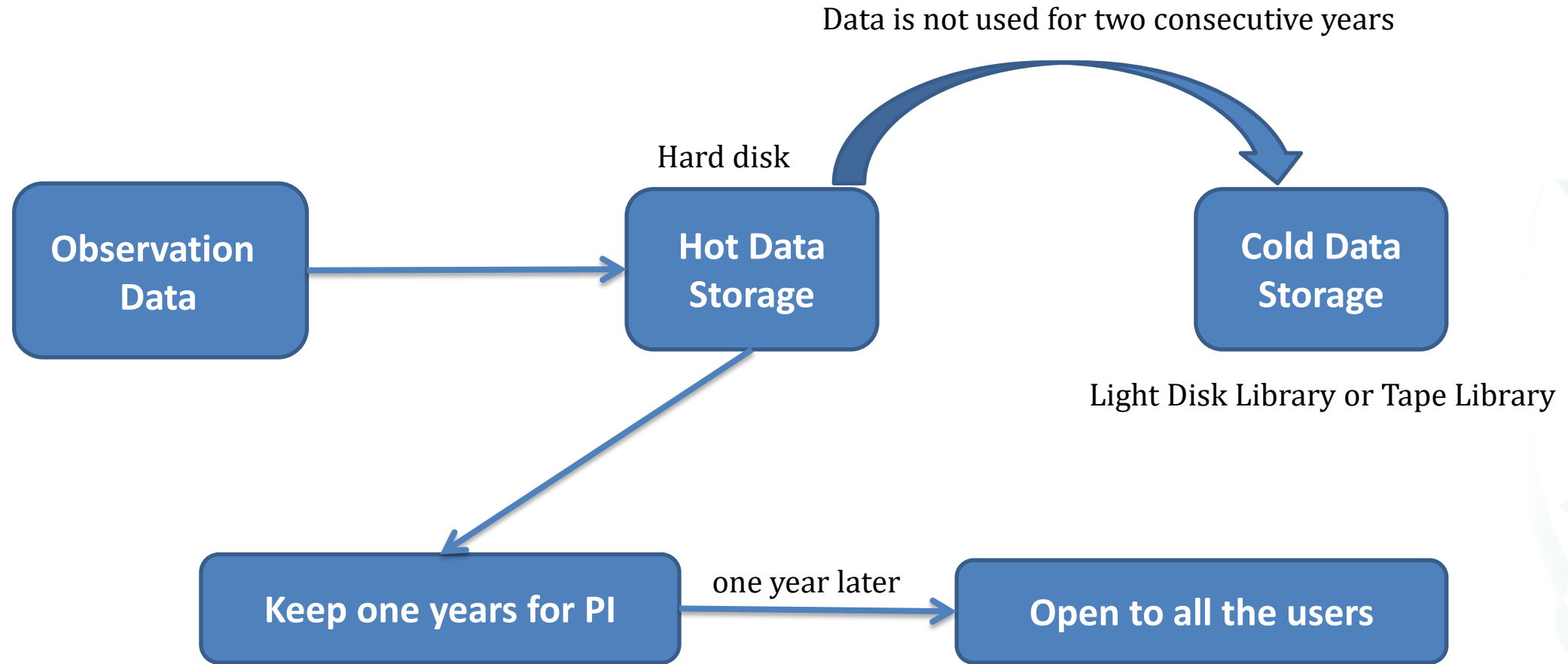
National Astronomical Data Center (NADC)

Its predecessor is the astronomical data center supported by China-VO and other observatories

Aims to provide full life cycle scientific data service and technology support for astronomy application, makes a distribution astronomical service system, but also is a challenge.
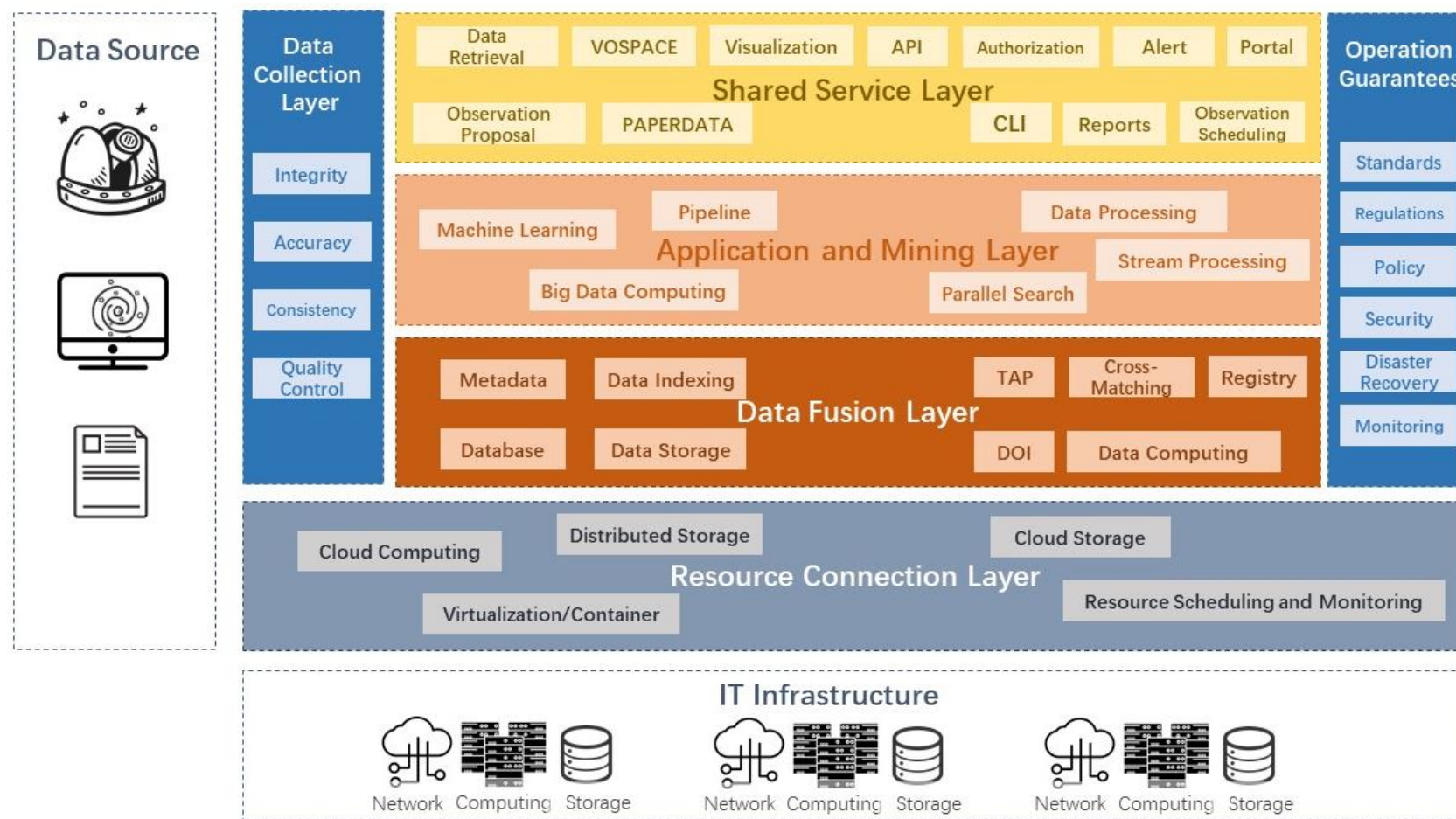
Data Management Plan

Up to now, it has produce more than 15PB, and it will newly increase ~10PB every year

Data is not used for two consecutive years

Hard disk

Observation Data → **Hot Data Storage** → **Cold Data Storage**

Light Disk Library or Tape Library

Keep one years for PI — one year later → **Open to all the users**

Data Management Plan

We are designing and developing the FAST Web Portal, and IVOA standard will be adopted. It aims to provide a full life cycle scientific service.
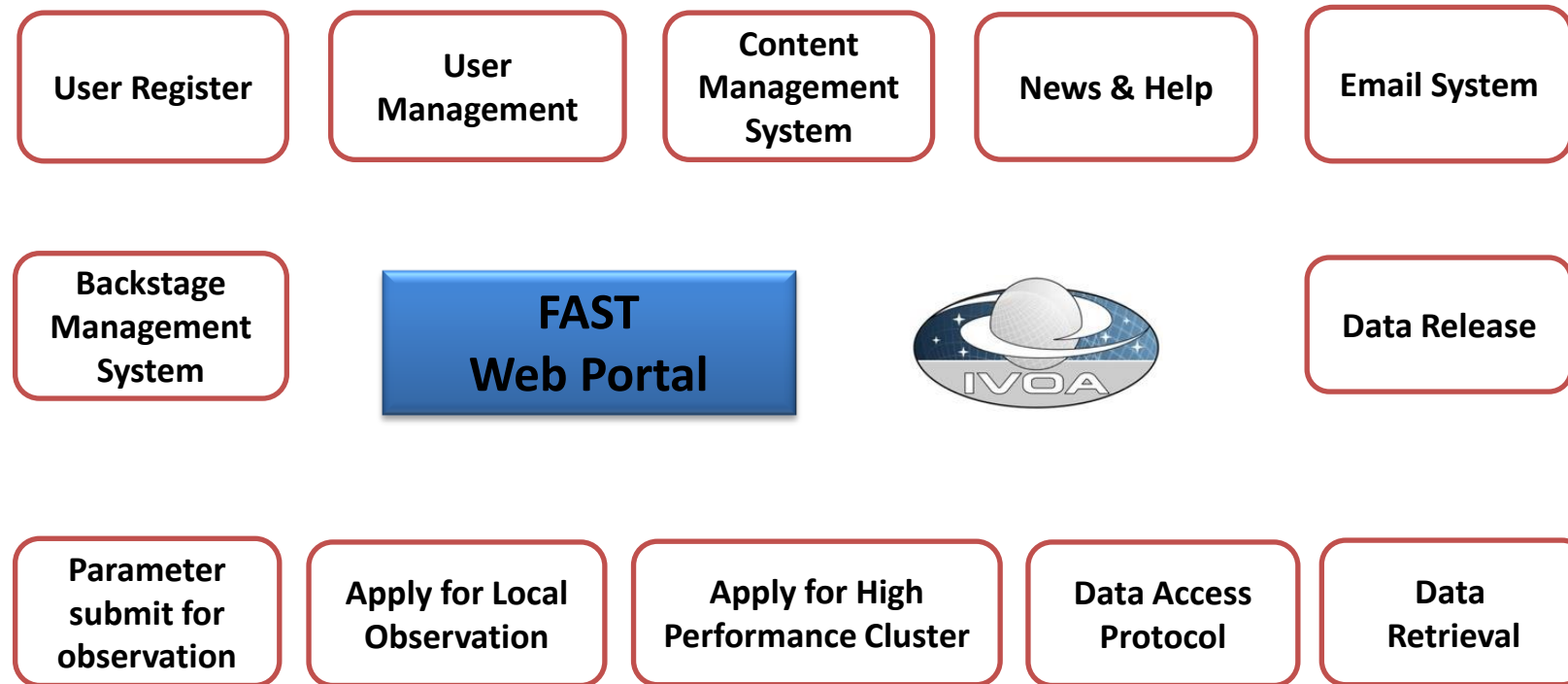
Part of function modules for the web portal