

Distributed Data Broker

S. Vannarat, A. Panartkul,
D. Wishadakul, J. Pengsuwan,
P. Srichaikul, R. Chitradon
NECTEC

in Collaboration with
Matthew Laurensen
Prof. Seishi Ninomiya
NARO/Japan

Outline

- Grid concept
- What is data broker
- Benefits
- Distributed data broker
- Fit to Grid concept
- Current status

What are Grids?

Electric power
grid metaphor

Infrastructure VS. Concepts

Grid makes

- interconnected resources
- more usable and controllable
- in a seamless manner

Network Infrastructure

- Rigid
- Uncontrollable

Solve this in 2 hr for \$10...

Transfer this in 1 min for \$1...



The challenges of the Grid

- Information services about the resources available on the Grid
- Resource Brokering
- Uniform access to resources: APIs
- Security
- Job scheduling
- Data Access
- Data Replication

Data Grid Architecture

App	Discipline-Specific Data Grid Application
Collective (App)	Coherency control, replica selection, task management, virtual data catalog, virtual data code catalog, ...
Collective (Generic)	Replica catalog, replica management, co-allocation, certificate authorities, metadata catalogs,
Resource	Access to data, access to computers, access to network performance data, ...
Connect	Communication, service discovery (DNS), authentication, authorization, delegation
Fabric	Storage systems, clusters, networks, network caches, ...

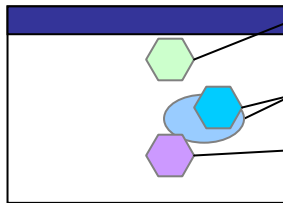
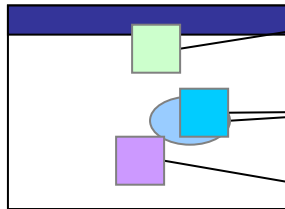
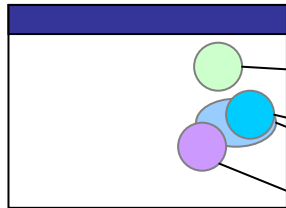
Weather Data

- Agriculture applications: prediction & planning
- Data sources: Files, SQL databases
- Data organization:
 - by station
 - by element
 - by location
 - by each resolution (hourly, daily)



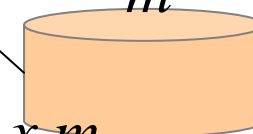
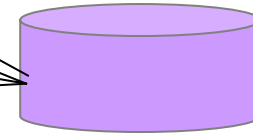
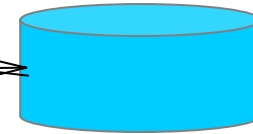
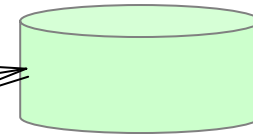
MetBroker

Applications

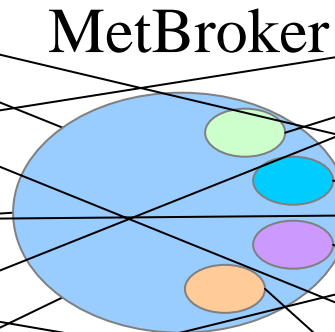


n

Weather databases



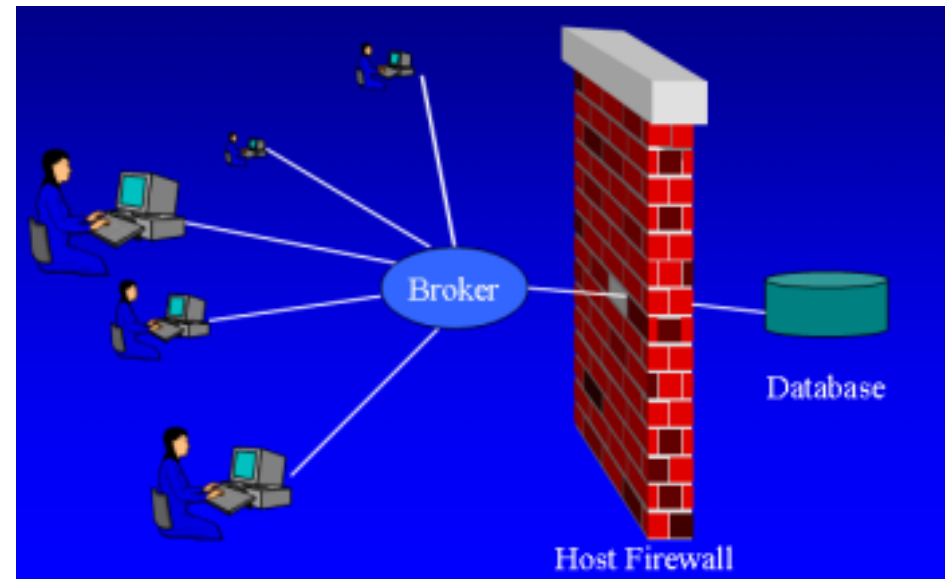
m



Number of code modules proportional to $n \times m$
Number of code modules proportional to $n + m$

Benefits of Mediated Architecture

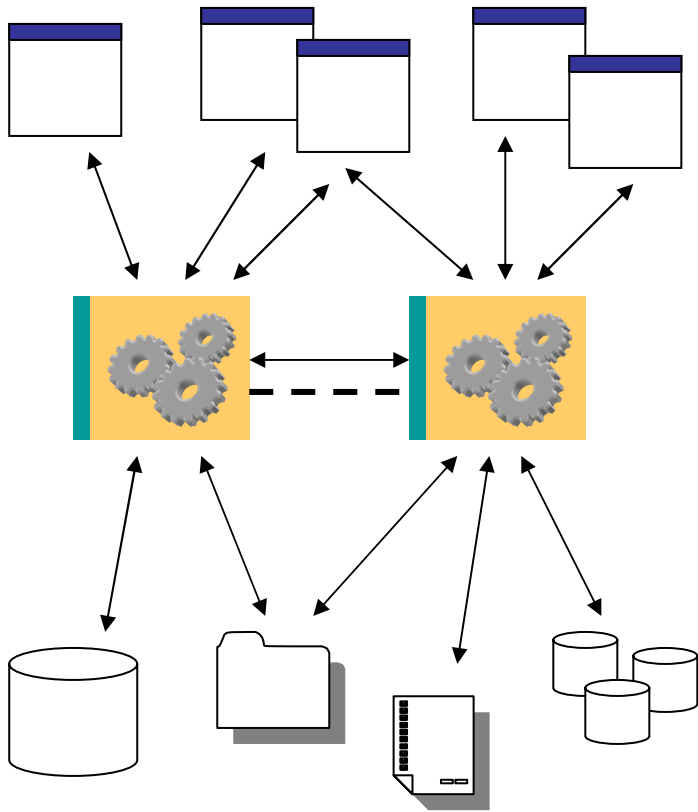
- Single API
- No change on database server.
- Metadata
- Simple client
- Network security



Current status

- Implemented
 - using RMI & Web Services for communication
 - JDBC, HTTP etc for database access
- Currently running on Linux host
- <http://www.agmodel.net/MetBroker/MetBroker.html>
- Databases: Japan, New Zealand, Georgia, UK

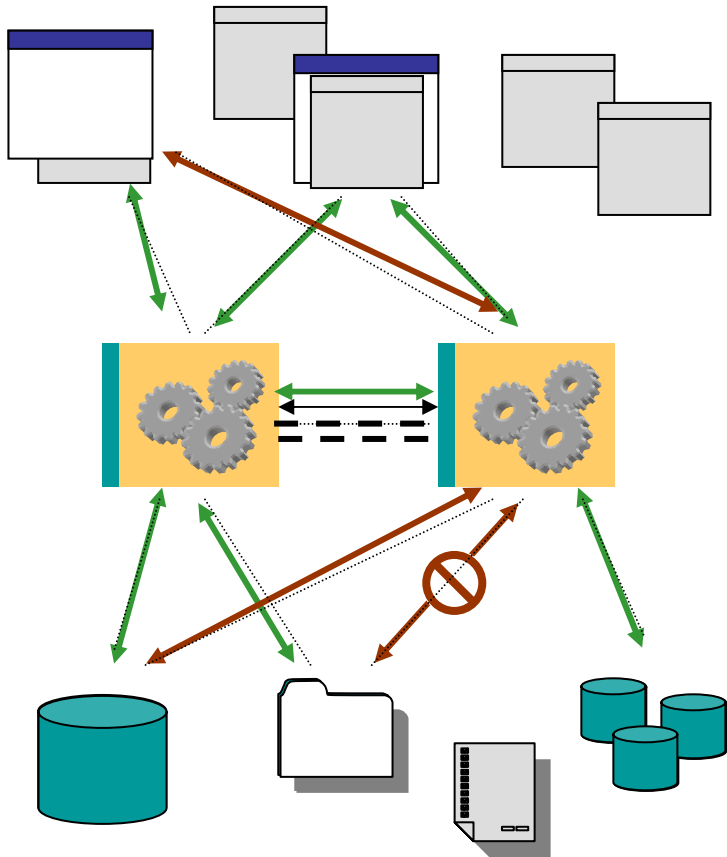
Distributed Broker



Advantages

- High availability
- Optimize network utilization
- ...

Distributed Broker



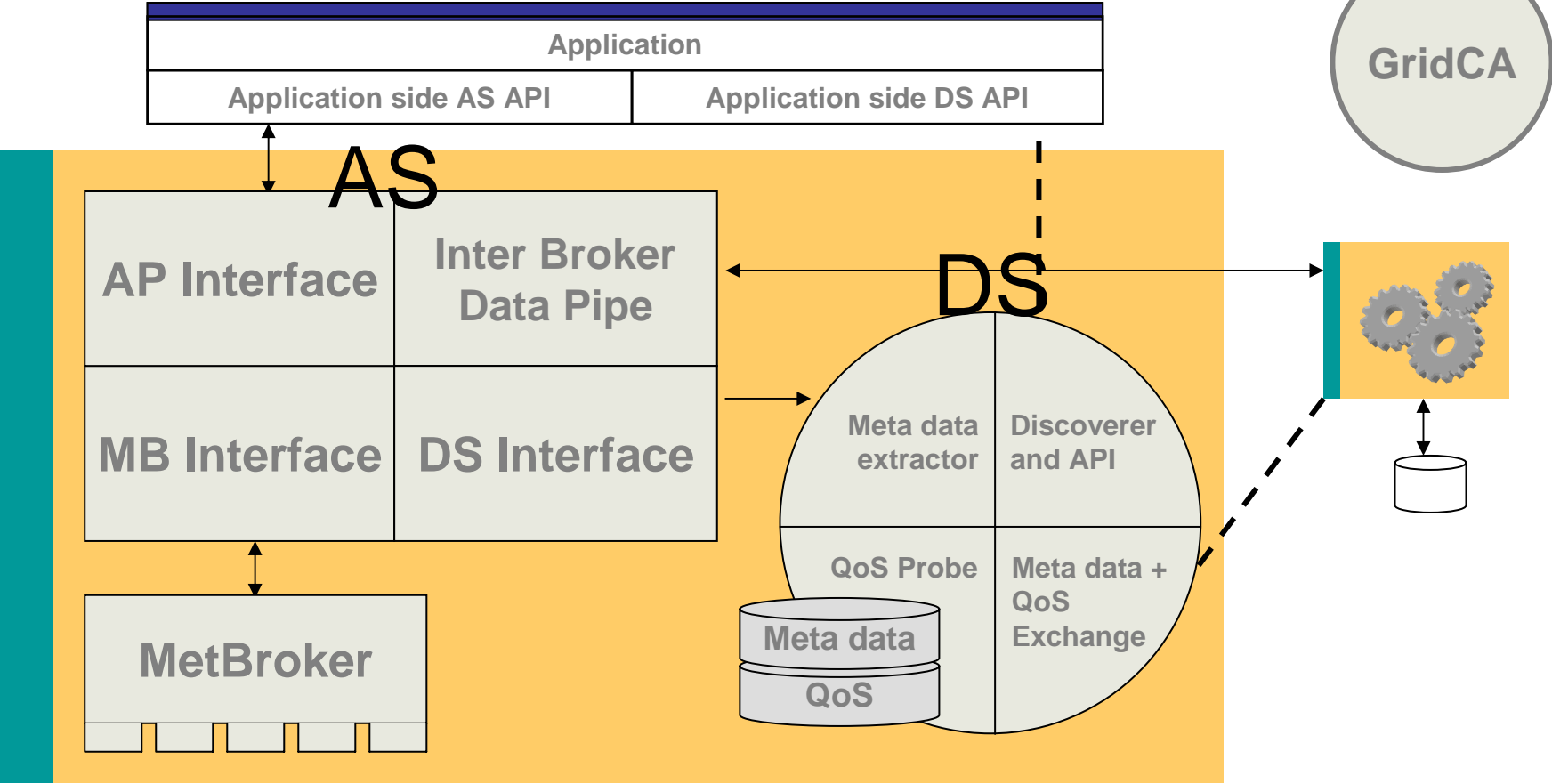
High availability

- Alternative paths

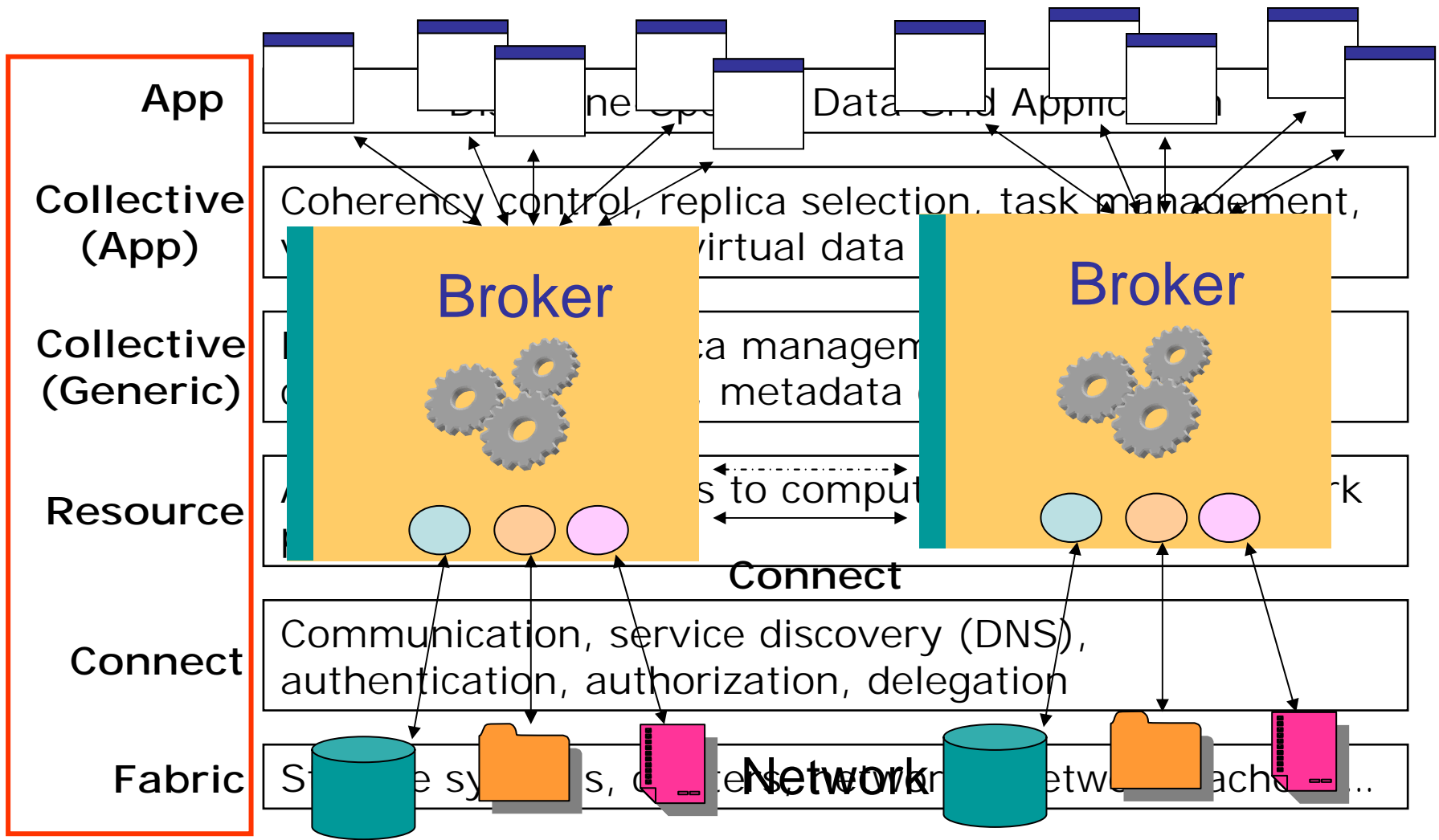
Optimize network

- Choose most efficient path
- Data re-routing and re-directing

Architecture



Broker in Grid Architecture



Dev. Phases



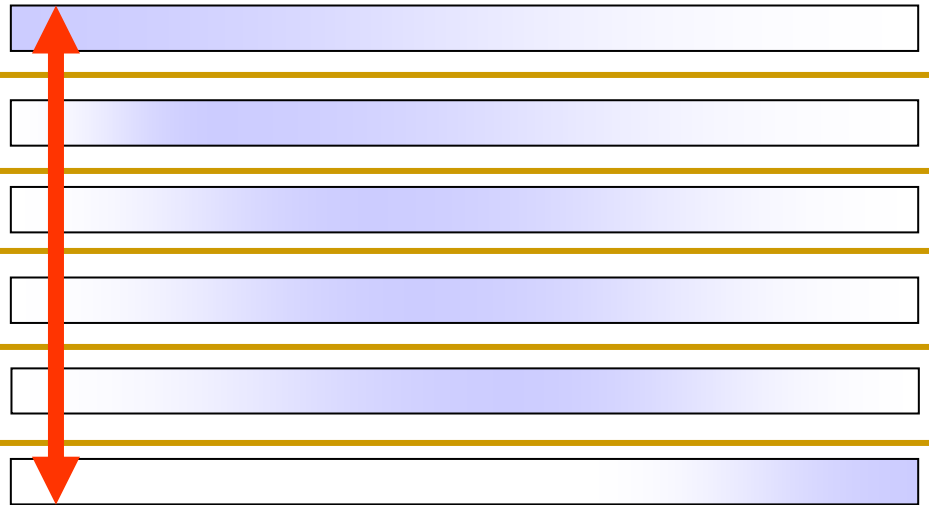
Function

- Access Services
- Discovery Services
- Redirect
- Reroute
- QoS
- CA

Time line

Sept. 03

Sept. 05



Summaries

- Grids concept: seamlessly usable interconnected resources
- Network of MetBrokers
- Fit to Grid Architecture
 - Use Grid components
 - Use Grid infrastructure
- (Distributed) Data Broker -> Data Grid