

Active Folder : Integrating All Activities of Simulation on File System

Suntae Hwang (sthwang@kookmin.ac.kr)

Daeyoung Heo (dyheo@kookmin.ac.kr)

School of Computer Science

Kookmin University

Apr 10 2014

Active Folder: Integrating All Activities of Simulation on File System

◎ Active Folder – good for case comparative study

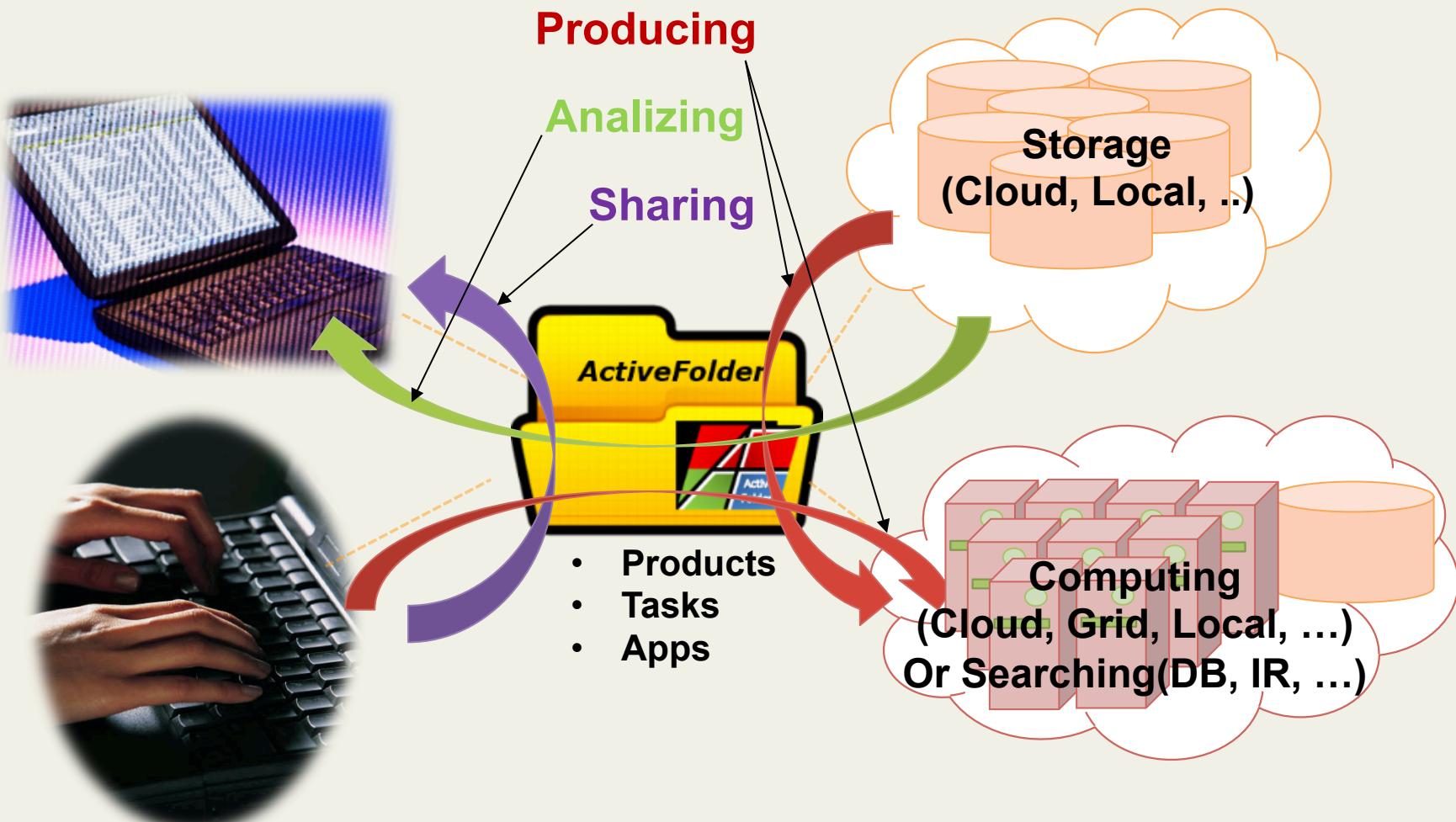
- Tasks
 - Described as regular folders and files

- Products
 - Input or output of simulation
 - Can be handled like regular file by using legacy software
 - Contains provenance information (meta data, task info, etc)
 - Can be reproduced by the task which is extracted from the provenance information

- Apps(Computing Resource)
 - Computing server(Local, Grid, Cloud, what ever, ...) is registered as regular folders and files
 - To submit a Job(task), just Drag&Drop the task folder to the folder which represents computing server



Active Folder: Integrating All Activities of Simulation on File System



Formal Definition of Spaces in Active Folder

Definition 1. Set P for Product

$$P = pp \rightarrow_{\tau} (d, t, H, s), s \in PS$$

H=hh is changing history

$$PS = \{Dormant, Creating, Completed, Damaged\}$$

Definition 2. Set T for task

$$T = tt \downarrow n \rightarrow_{\tau} (L, w, K, P \downarrow k, P \downarrow t, t \downarrow n - 1), P \downarrow k \subseteq P, P \downarrow t \subset P$$

Definition 3. Set A for producing/reproducing

$$A = ff \rightarrow_{\tau} app \vee (app \circ t), \text{ ie. } app \circ t \circ K \circ P \downarrow k \rightarrow_{\tau} P \downarrow t$$

$$app = e(R, c(lw(t), K), m, P \downarrow k)$$

Definition 4. Set M for monitoring.

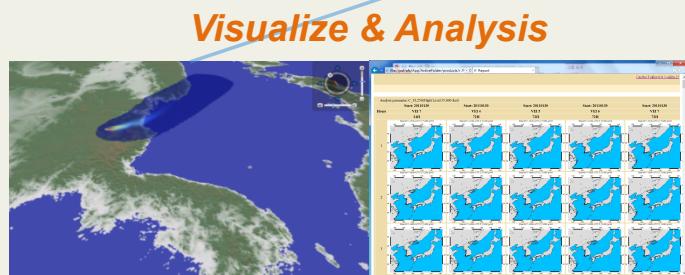
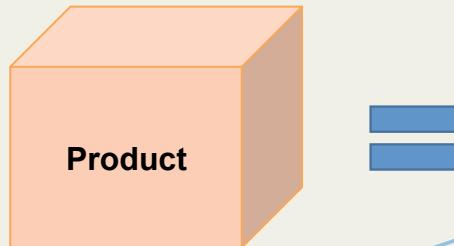
$$M = (m, stat) \text{ subscribe}(m) \rightarrow_{\tau} stat$$

$$stat = (t, R, K, P \downarrow k, P \downarrow t, process)$$



Active Folder: Product and Reproducing

Product consists of data, history info and task(provenance)



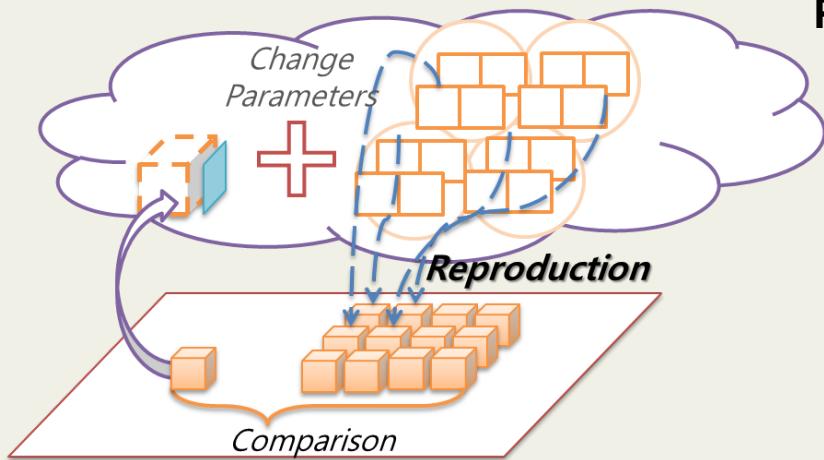
Visualize & Analysis

Task can be extracted from products and used for producing similar results with different parameters

버전 관리		
새 버전 업로드	사용된 저장용량	자동 삭제 안함 자세히 알아보기
21:13 작성자: 허대영	73KB	<input checked="" type="checkbox"/> <input type="checkbox"/>
이전 버전 2개		
21:13 작성자: 허대영	28KB	<input checked="" type="checkbox"/> <input type="checkbox"/>
21:12 작성자: 허대영	81KB	<input checked="" type="checkbox"/> <input type="checkbox"/>
사용된 전체 저장용량: 182KB 자세히 알아보기		

```
ZLAYER_(M) FROM 0. TO 14000. INCREMENT 500.  
GRANULOMETRY  
DISTRIBUTION = GAUSSIAN  
NUMBER_OF_CLASSES = 6  
FI_MEAN = 2.5  
FI_DISP = 1.5  
FI_RANGE = 2. 7.  
DENSITY_RANGE = 1200 2300  
SPHERICITY_RANGE = 0.9 0.9  
SOURCE  
SOURCE_TYPE = plume  
PLUME_SOURCE  
SOLVE_PLUME_FOR = MFR  
MFR_SEARCH_RANGE = 3.0 7.0  
HEIGHT_ABOVE_VENT_(M) = 10000.  
MASS_FLOW_RATE_(KGS) = 1d3 1d4  
EXIT_VELOCITY_(MS) = 200.  
EXIT_TEMPERATURE_(K) = 1073.  
EXIT_WATER_FRACTION_(%) = 1.  
AGGREGATION  
AGGREGATION_MODEL = Fall3D  
FI_AGGREGATES = 2.  
DENSITY_AGGREGATES = 350.
```

Fall3D



Procedure *Reproducing(p,K,P↓k)*

let $t = \text{task}(p)$

$\text{app}(t, K, P \downarrow k)$

Different parameters

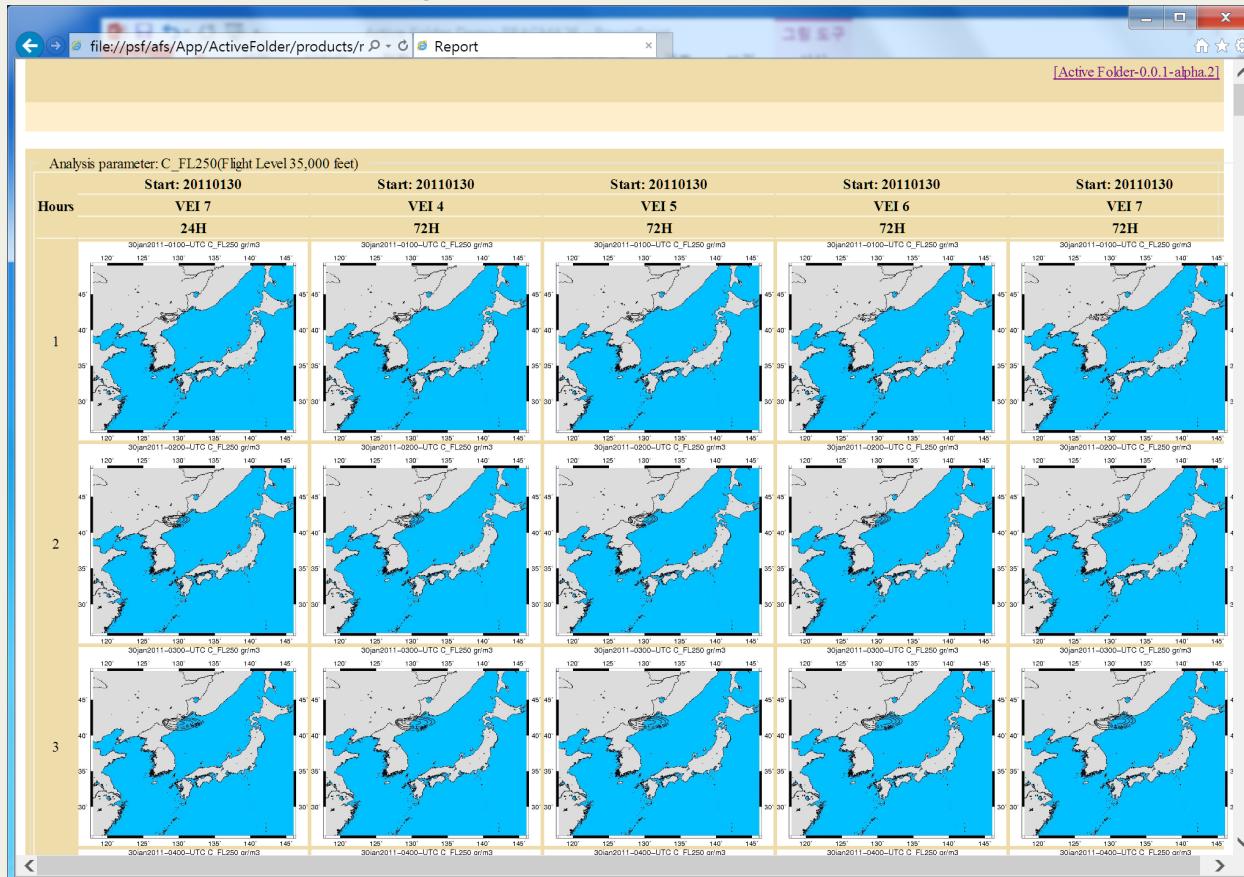
* $\text{task}(p)$

= $\text{task}((d, t \downarrow \text{origin}, H, s)) \rightarrow \text{tcopy}(t \downarrow \text{origin})$
= $t \downarrow \text{new}$

* $\text{app}(t, K, P \downarrow k)$ is producing function

Active Folder: Products Searching and Comparison

- Similar or interested products can be searched by their provenance information and compared by various legacy tools on conventional file system



Case Study : Volcano Eruption Simulation using Active Folder



Simulation Parameters

- ① **Eruption Location** : Slopes, Under Caldera, ...
- ② **Ash**
 - Plume Style : Suzuki, Point, Plume or Resuspension
 - Height, Mass Flow Rate, ... (physics parameters)
 - Ash Size : Granulometry, Distribution, ... or Density
- ③ **Weather Condition**
 - Wind field, Air temperature, ...

Available Actions

1. Search simulation results by parameters
2. Compare Results

Or

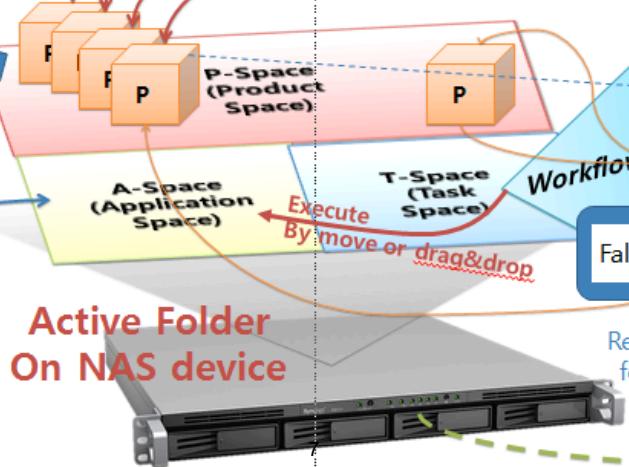
Reproducing by parameter sweeping



High Performance Computing

Representing
as a folder

Active Folder On NAS device



Represented by the
folder and scripts



Cloud Solution

Simulation Procedure



Weather
Simulations



Volcano
Eruption
Simulations

Fall3D (1)

Fall3D (2)

Fall3D (3)

Fall3D (4) ...

Simulation examples

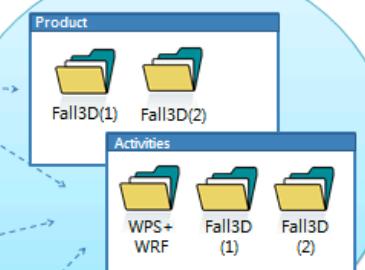
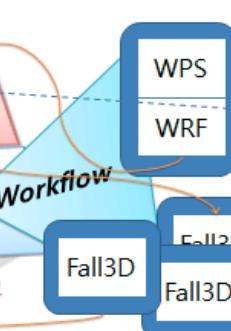
④ East slope
⑤ Plume: Suzuki
Height: 3,000m
MFR: 2 ~ 3 kg/s
Ash size: 0.5um

④ East slope
⑤ Plume: Suzuki
Height: 3,000m
MFR: 3 ~ 4 kg/s
Ash size: 0.1um

④ East slope
⑤ Plume: Suzuki
Height: 8,000m
MFR: 2 ~ 3 kg /s
Ash size: 0.1um

④ Under Caldera
⑤ Plume: Suzuki
Height: 8,000m
MFR: 3 ~ 4 kg/s
Ash size: 0.1um

Results Comparison for damage estimation and decision making



Project #3:

Development of IT-based Response System for Volcanic Disaster

<http://www.volcano.re.kr>

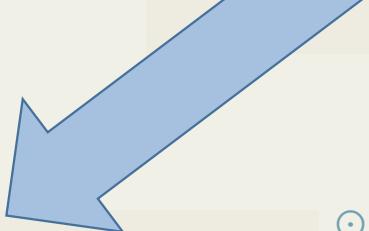
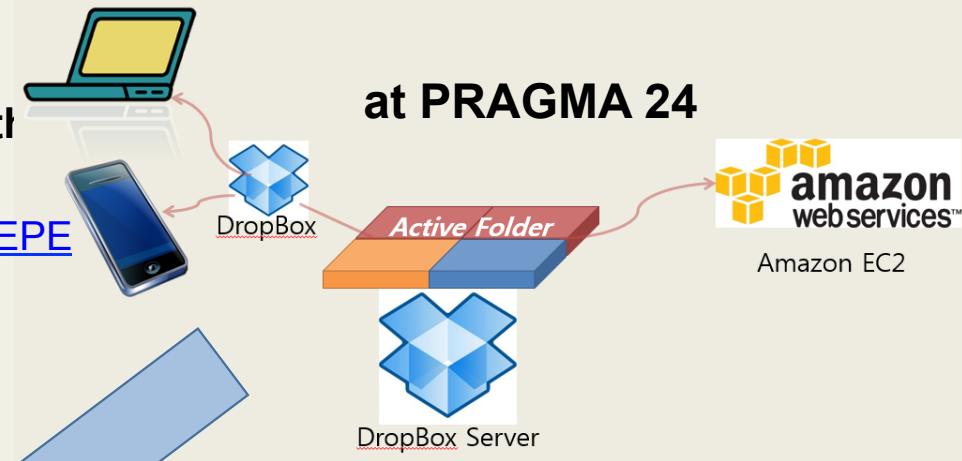
Development of Preparedness Procedure and Technology for Volcanic Disaster in Korea.

Short history of Active Folder

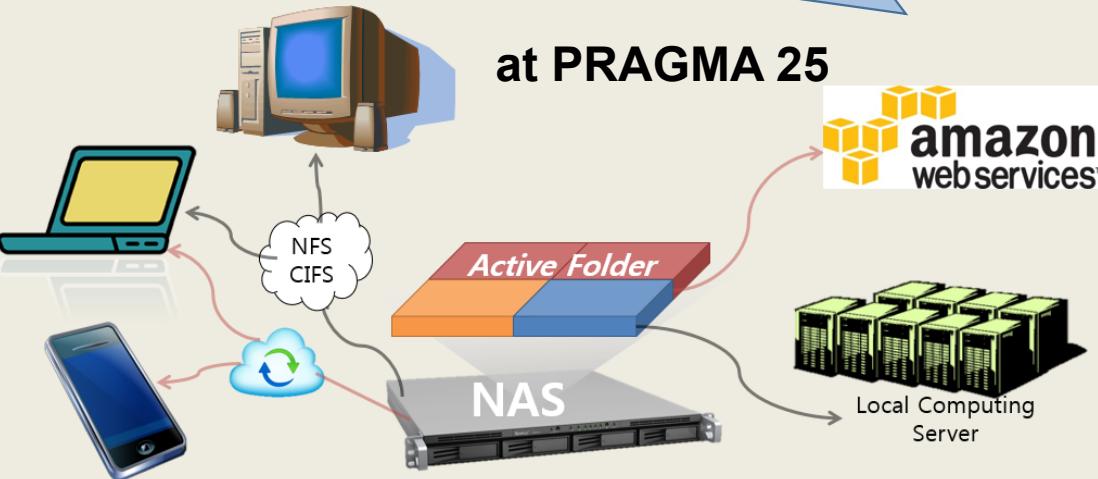
○ Active Folder on DropBox+EC2

- Cost & Performance Problem with very large files

<http://www.youtube.com/watch?v=vXTh3tCHEPE>



at PRAGMA 25



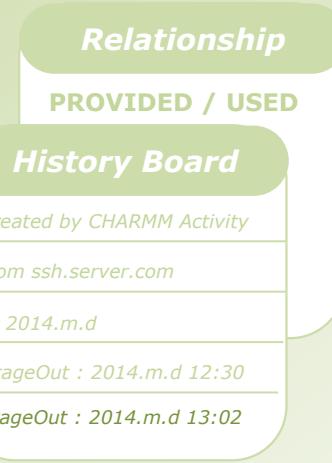
○ Active Folder on NAS

○ NAS(Network Attached Storage)

- Large Volume Storage
- Network File System (NFS, SMB/CIFS, AFP ...)
- Most vendors support Cloud solution like DropBox

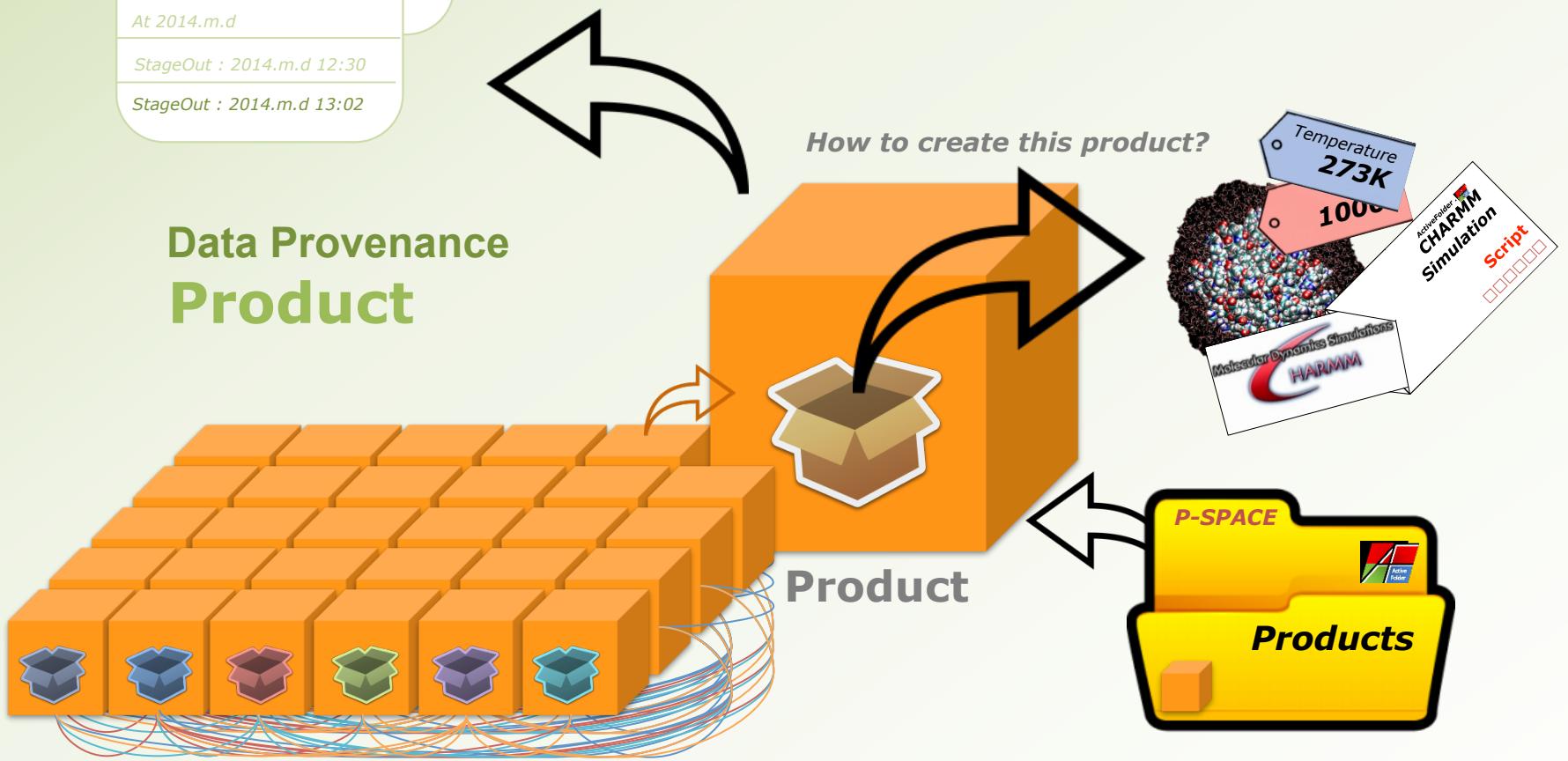
Active Folder Features 1:
Computing by manipulating files or folders, which makes only files visible by hiding correlation of computing method, parameters, result files and computing resources

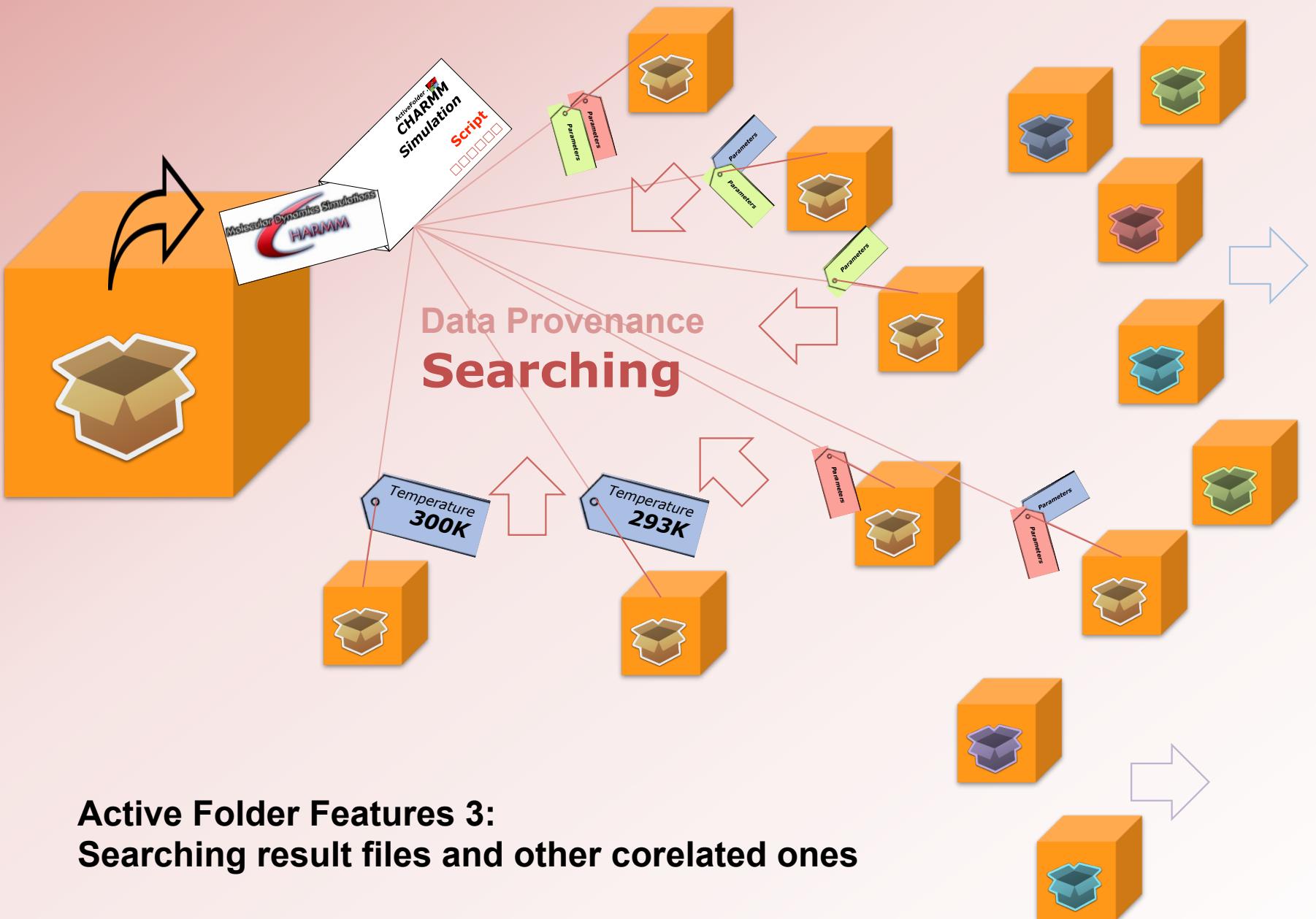




Active Folder Features 2: Keep the way of computing methods and parameters with result files silently

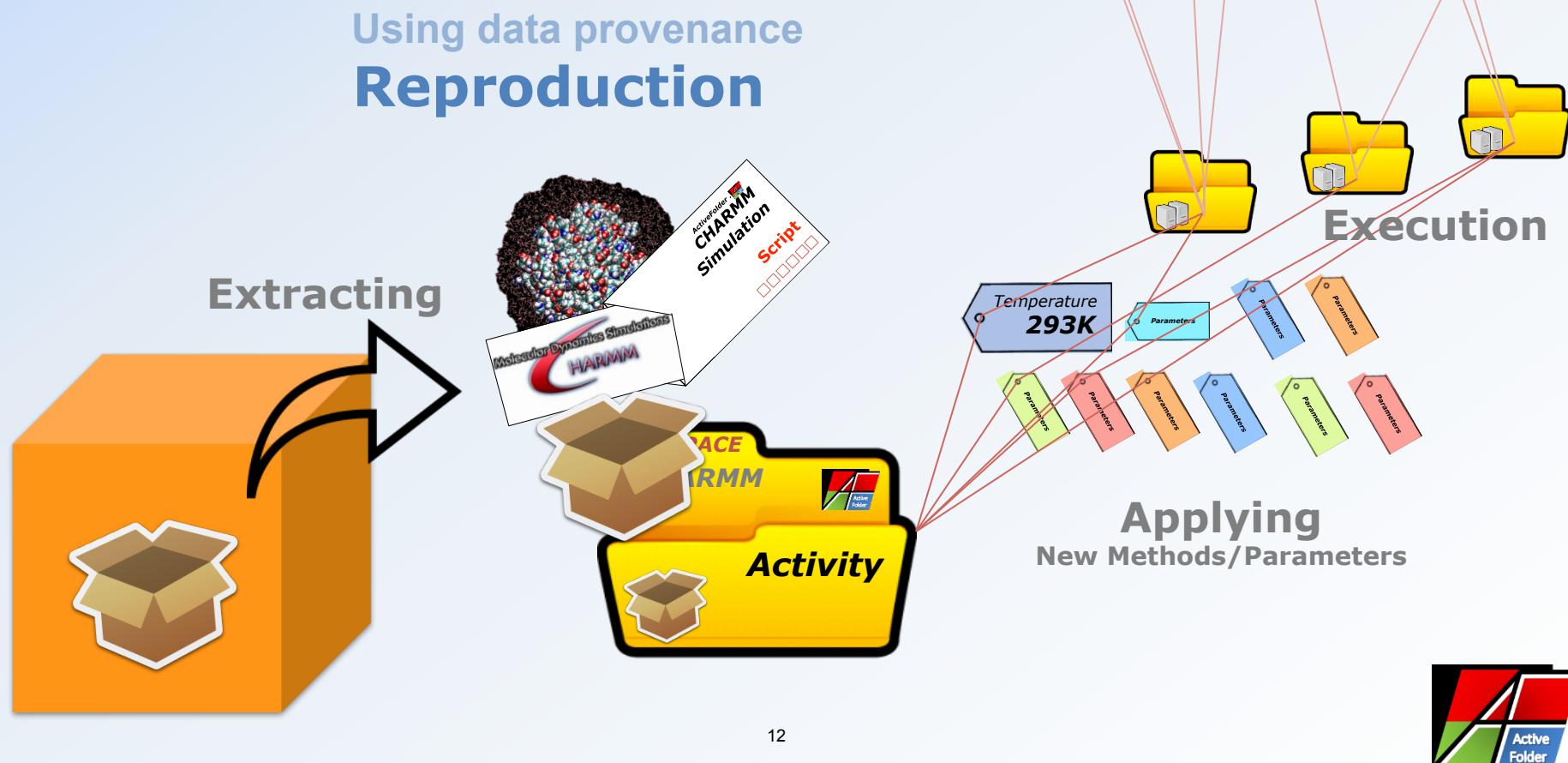
Data Provenance Product

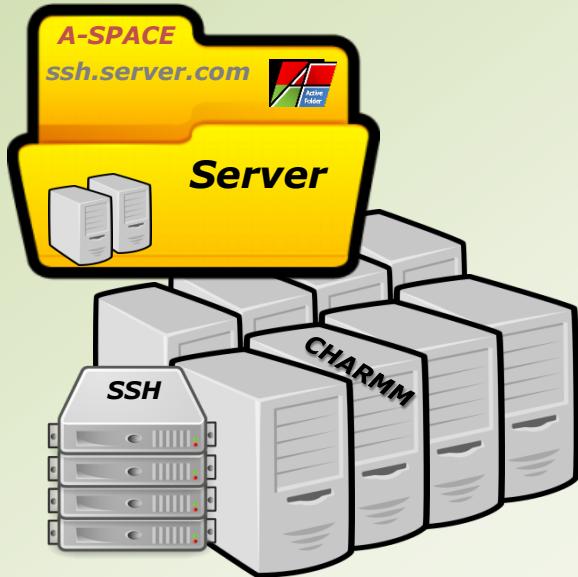




**Active Folder Features 3:
Searching result files and other corelated ones**

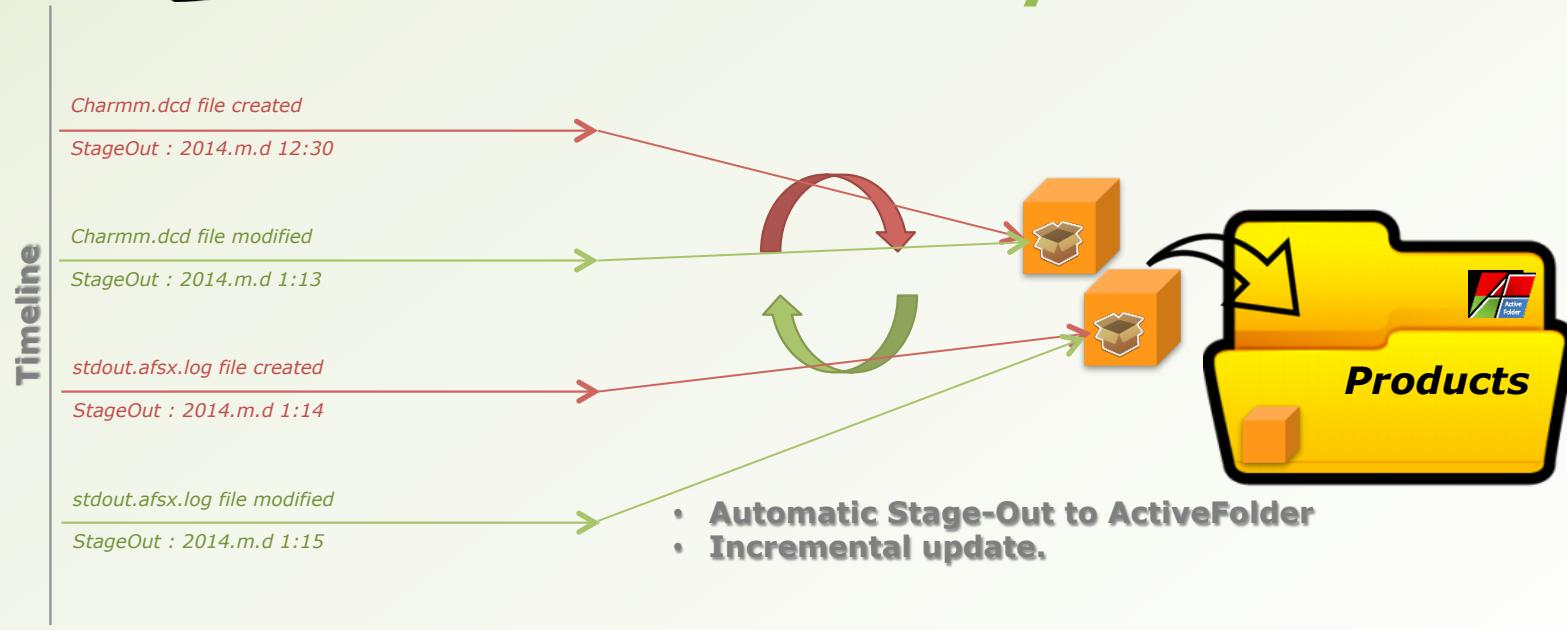
Active Folder Features 4:
Computing by manipulating files or folders,
Re-generating either the same or similar one
s by existing result files

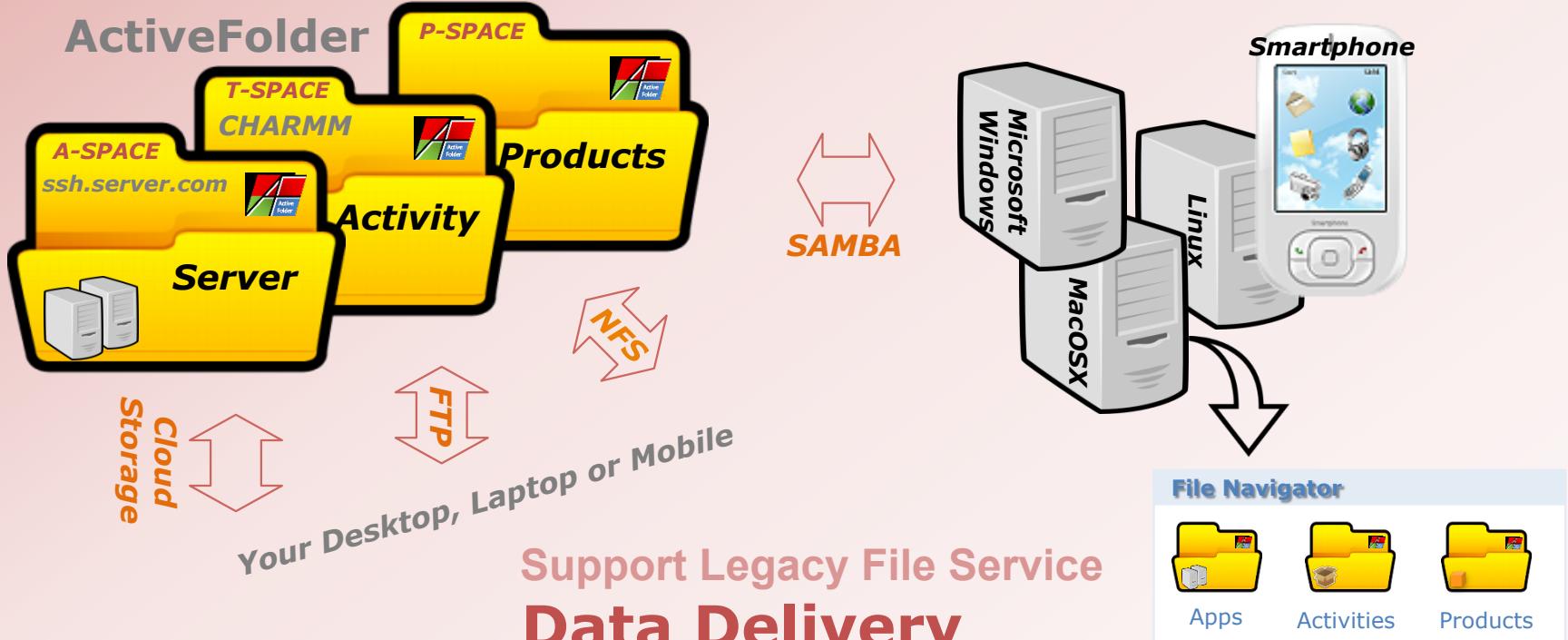




Active Folder Features 5:
Returns computing results
by progressive synchronization
without connecting computing server

Between server and your active folder **Data Synchronization**



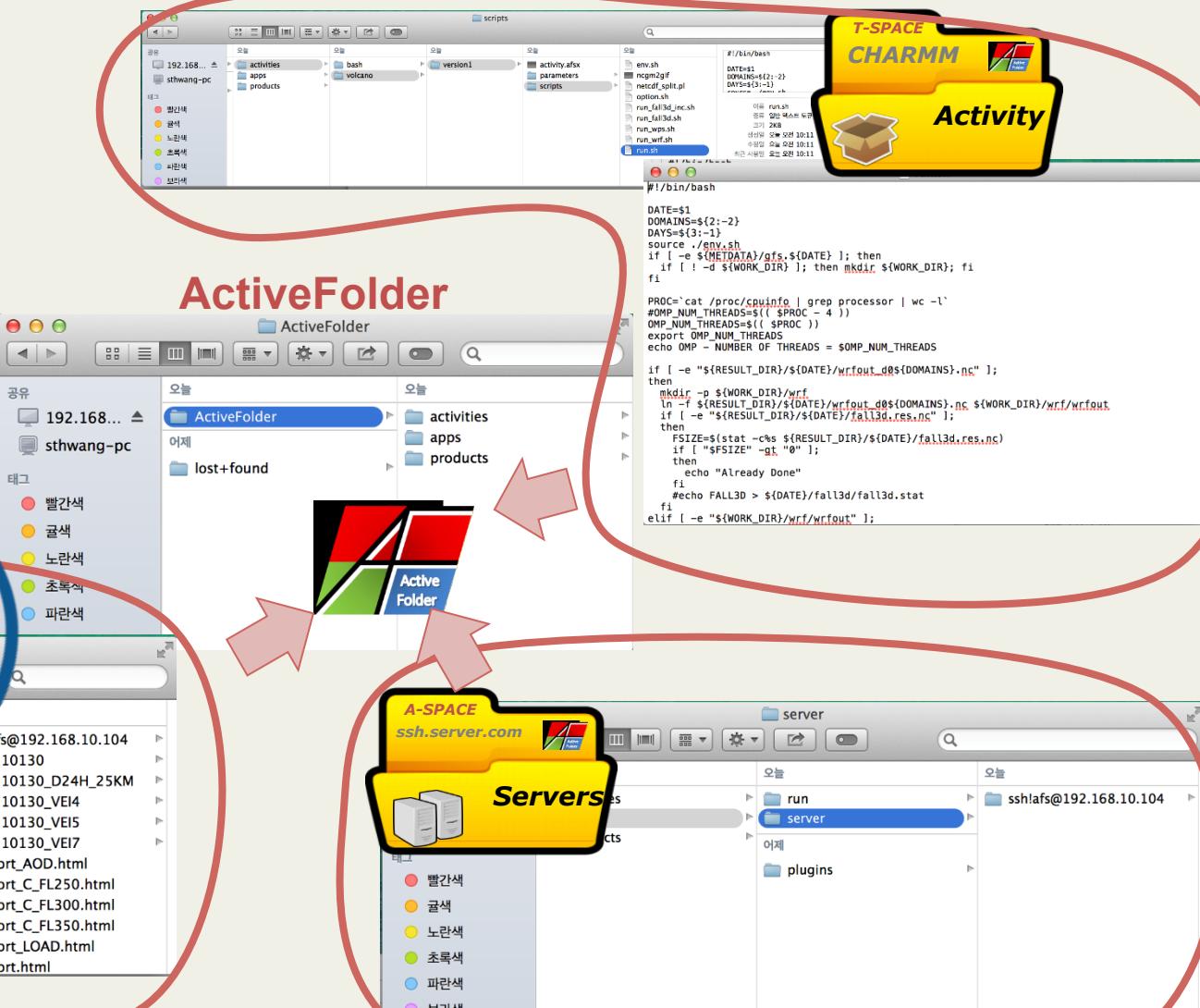
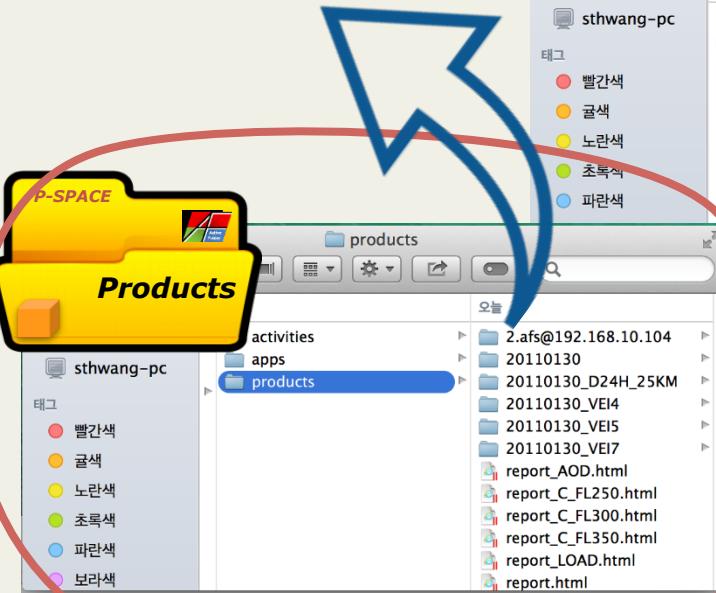
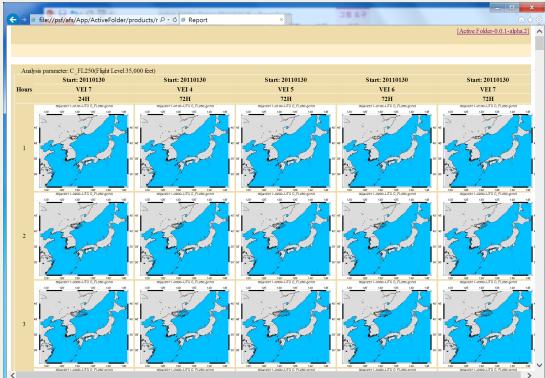


Active Folder Features 6:
Support various file service such as NFS,
Samba(CIFS) and FTP

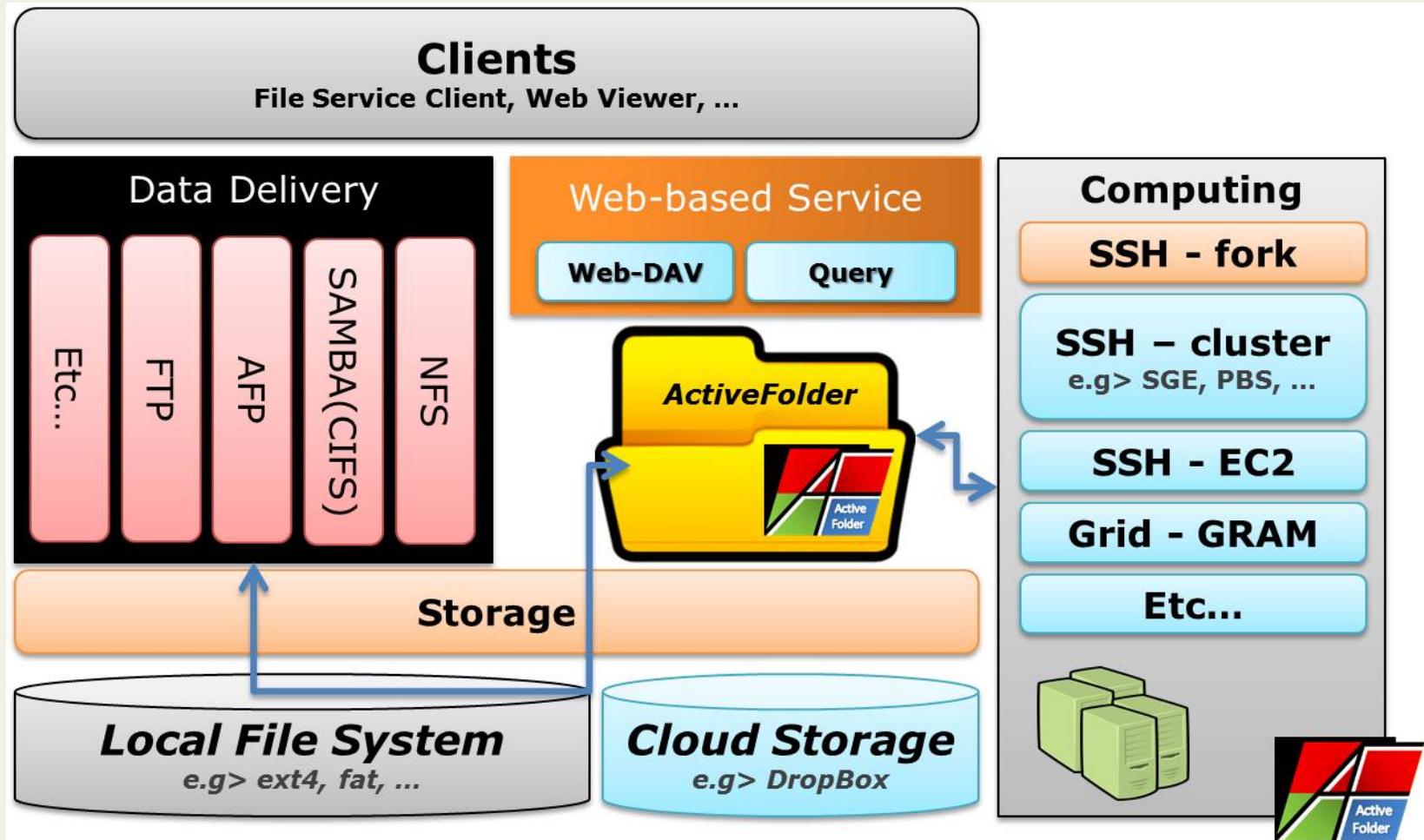


Active Folder: Overall View

Searching and Analysis

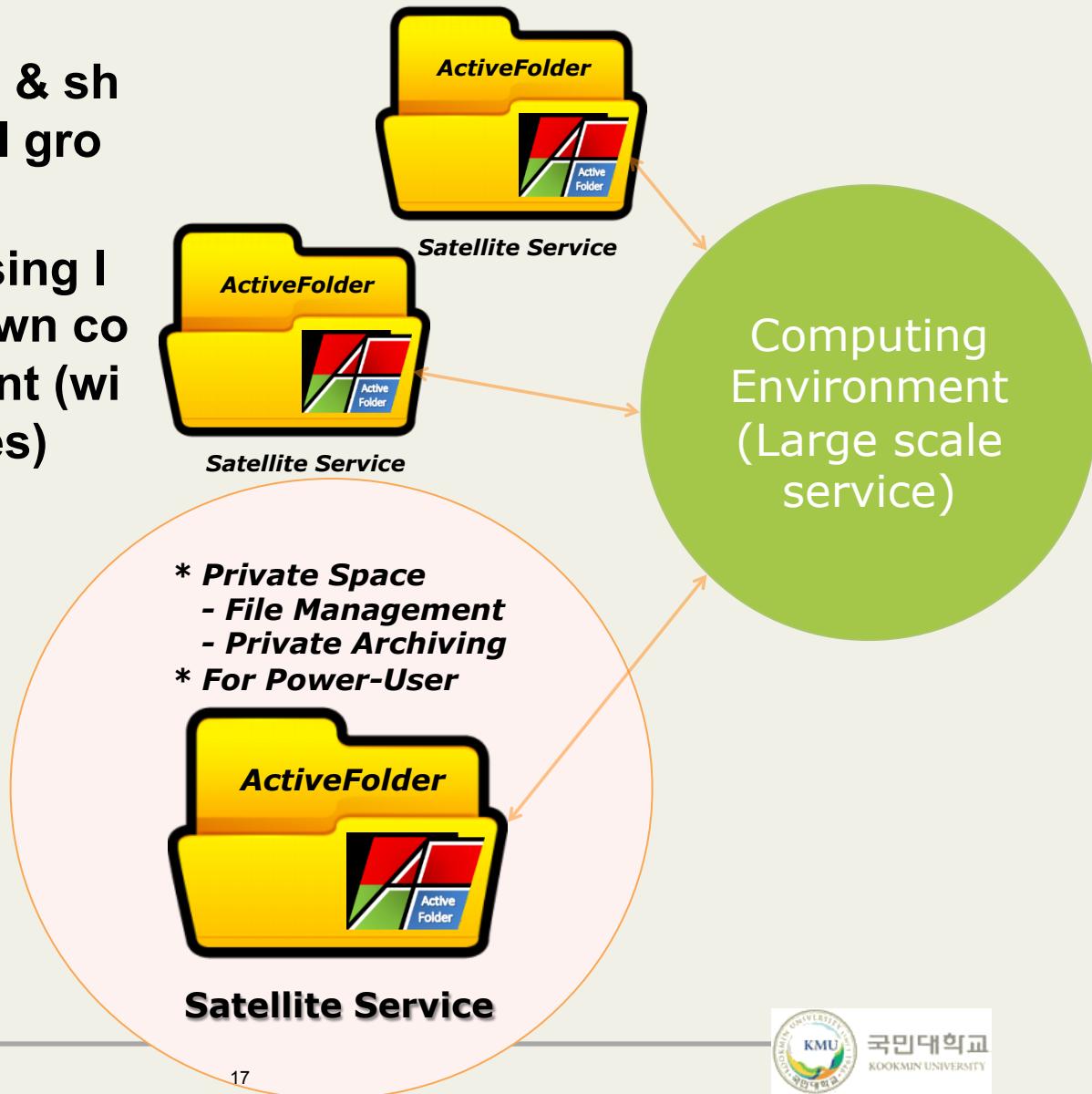


Active Folder: Software Architecture



Usage of Active Folder

- Good for managing & sharing data in a small group of Lab
- For Power-users using Legacy SW in their own computing environment (without GUI sometimes)



Active Folder Site

○ Active Folder is released

- <http://www.activefolder.org>

Attention! .org only,
not .com



○ Downloads

- Binary file for general installation(Synology(NAS) tested)
- RPM for Redhat 6/CentOS 6

○ Coming soon

- Product searching interface
- Synology packaging for NAS
- Integrated using NAS and DropBox together



Thank you!

Questions?

