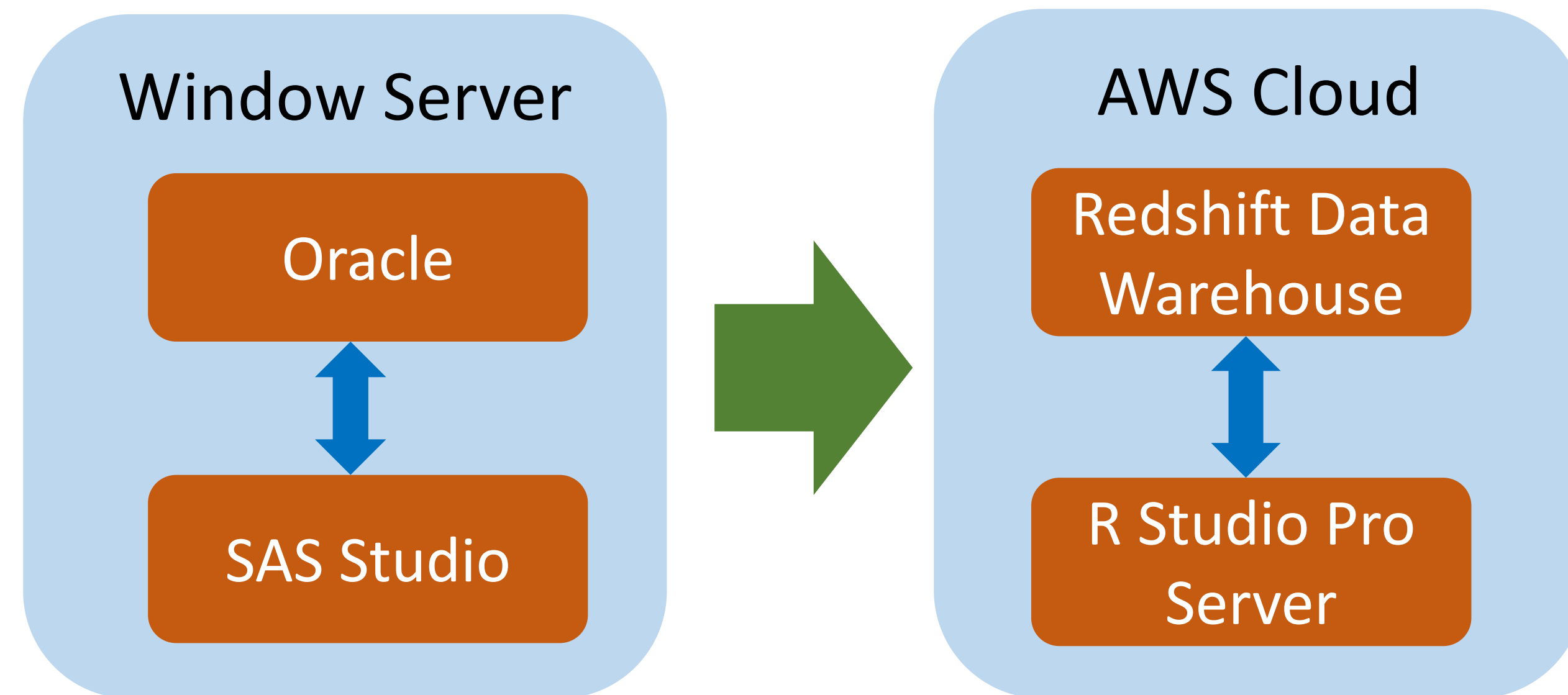


Organizational Transition from SAS to Open-source Programming for the whole department

Kevin Lee, kevin.lee@genpact.com, Manikandan Jeeva, Manikandan.jeeva@genpact.com

High Level Infrastructure Switch of Open-Source Transition Project (SAS Window to AWS Open-Source Analytical System)



Scope of Open-Source Transition Project

SAS



- ☐ SAS Studio
- ☐ Window Server
- ☐ SAS programming
- ☐ SAS Programmers (150)
- ☐ SAS codes (230)

Open-Source Program



- ☐ R Pro Server (R Studio, Jupyter)
- ☐ AWS Cloud Computing(Linux)
- ☐ Open-Source Programming(R, Python)
- ☐ Data Scientists
- ☐ R, Python & SQL codes

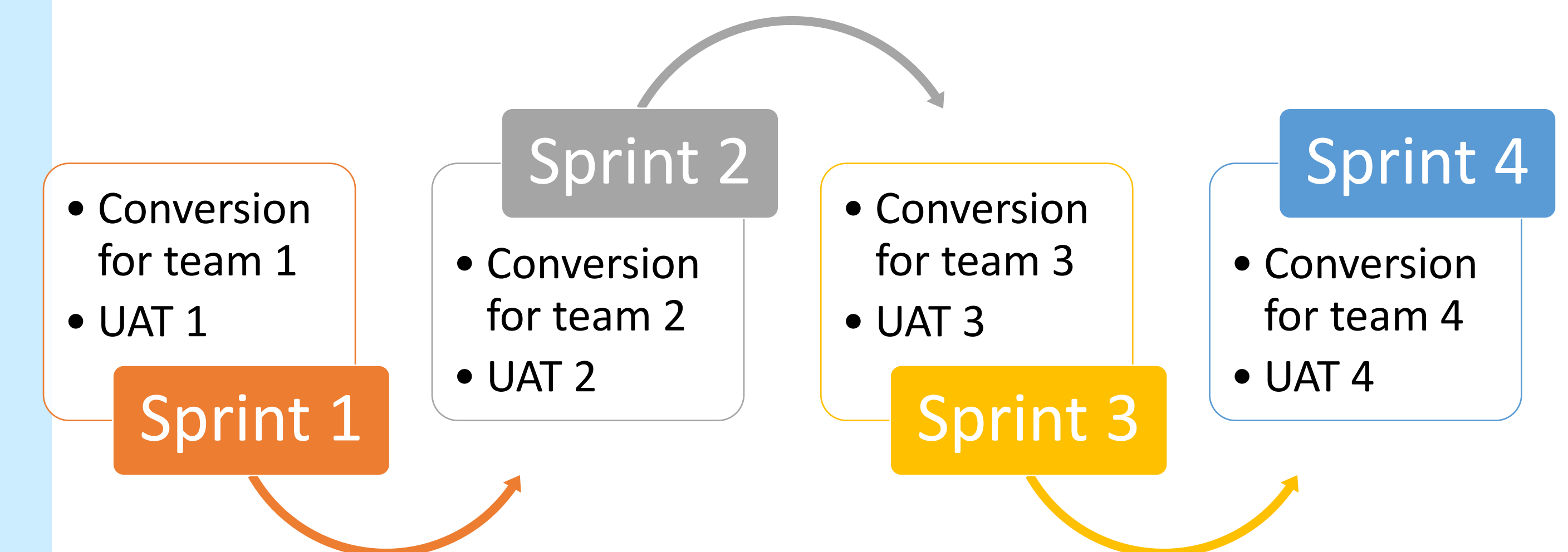
Challenges of Open-Source Transition Project

- **Inexperience** in Open-Source Programming (e.g., R, Python)
- **Inexperience** in new system / environment (AWS Cloud Computing, R Studio, Jupyter, Linux)
- **Learning curve**
- **Daily works** as well as existing SAS codes conversion
- **Lack of enterprise-level customer support** of Open-Source programming
- **Version control**
- **Uncertainty** toward new culture created by Open-Source programming

How to Overcome the Challenges

- **Training** on Open-Source programming (R & Python)
- **Continuous Workshop** on the new system
 - R Pro Server
 - AWS Redshift Data Warehouse
 - Linux
 - S3
- **SAS codes conversion** (+230) by Transition Team
 - User Acceptance Test
 - New process document
- **Change Management by Leadership**

Sprint Cycle of SAS Codes Conversion



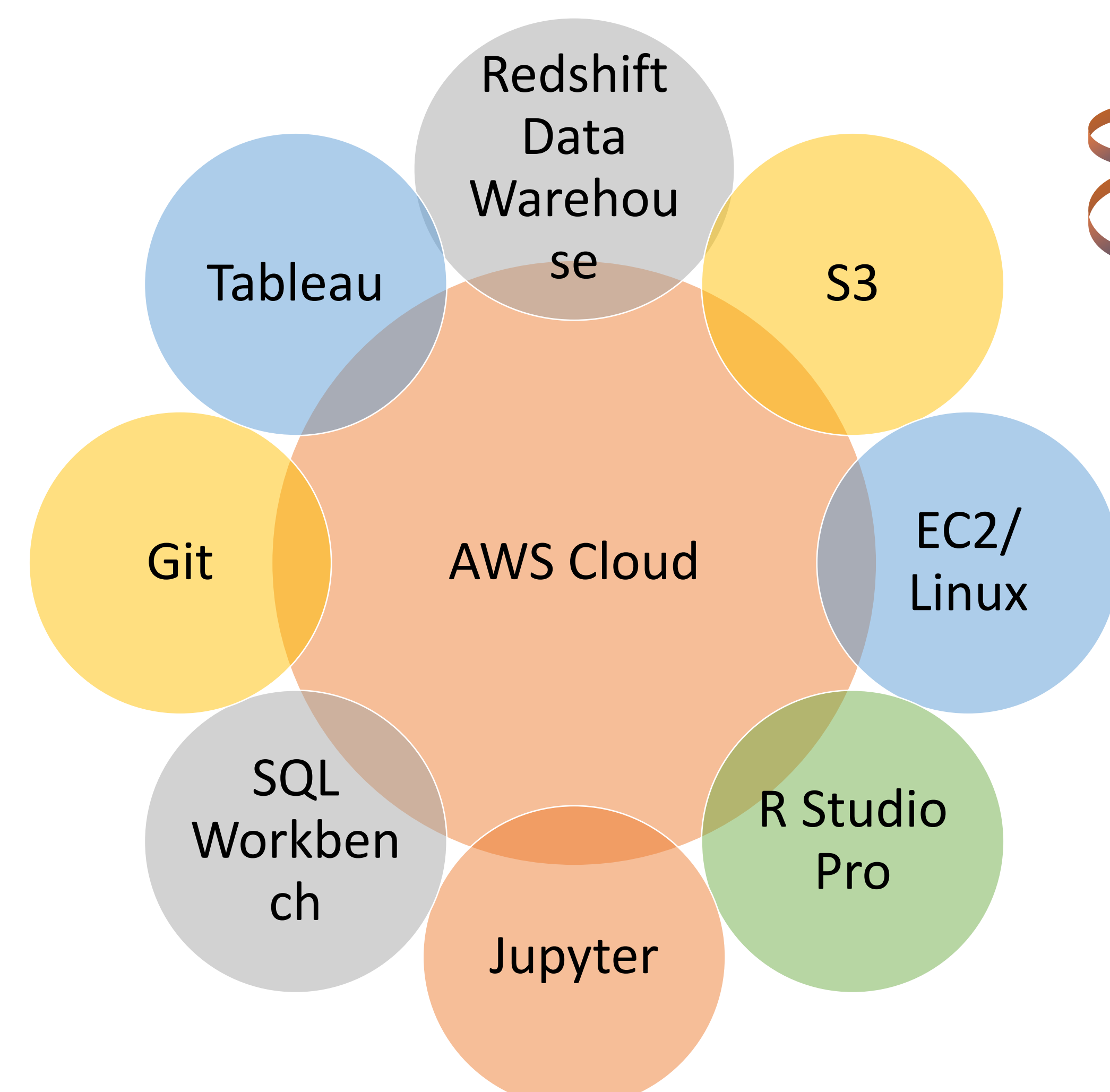
Change Management by Leadership

- **Clear goals, plan and time** of the change
- **Executive support**
- **Resources and budgets** for a change
- Frequent, transparent **communication** on the project progress/update
- **Participation** from the whole team
- **Dedicated supporting team/SME**
- **Continuous support** (e.g., new programming, system, infrastructure) during the change

Benefits from Open-Source Transition Project

- **Cultural Changes**
 - Less dependent on tools (Enterprise system)
 - Open to new technologies
 - Faster to adopt new changes
 - Collaborative
- **More opportunities**
 - To learn new languages (R, Python) and system (Cloud Computing, Git)
 - To keep top talents
- **Cost effective**
- **Innovative, Advanced Analytics** (Big Data, Data Visualization, Machine Learning, Big Data)

AWS Open-Source Analytical System



Example : Difference between SAS and Open-Source Programming

- SAS processes the data a row by a row.

studyid	subjid	age	sexc

```
data DM;
set DEMO;
if SEXC = 'Male' then SEX = 'M'
else if SEXC = 'Female' then SEX = 'F'
else if SEXC = 'Unknown' then SEX = 'U'
run;
```

- R/Python/SQL processes the data a column by a column.

studyid	subjid	age	sexc

sexc	sex

- Less memory to use
- Faster to process
- Difficult to process/create multiple variables

```
DM['SEX'] = DM.SEXC.replace(['Male','Female','Unkonwn'], ['M', 'F', 'U'])
```

Lessons Learned from Open-Source Transition Project

- **Support for programmers** during the transition is critical.
 - Training
 - Workshop
- **Dedicated supporting team/SME** during the transition is critical.
- **Know-how / experience** on Existing System and Open-Source System is critical
- **Change Management by leadership** is the key to success.