# 6.4 Data Systems at the PIFSC
### (an Information and Technology Services (ITS) perspective)

## Evan Howell

## 2013 External Review

# ITS Data Mission

Construct and maintain database and information *system infrastructure* for managers, scientists and administrators at PIFSC to assist in analysis and decision making

Focus is on:

- System hardware
- Core software installation, maintenance, and backup
- Support for data and database specialists at PIFSC

# Areas of responsibility

| ITS | Scientific Information Services (SIS)/PIFSC |
|---|---|
| **DATA SYSTEMS** | **DATA** |
| Enterprise database system hardware (Selection, purchasing, maintenance) | n/a |
| Database software patch management | Database administration (DBA) |
| Total PIFSC IT System Security (including database) | Data quality & integrity |
| Software licenses and upgrades | Software configuration |
| Maintain supported systems | Maintain data |
| Backup and archive data systems | Backup and archive data |

# Sample workflow for new data system

## Program / ITS

**System Planning**
- Scope
- Assign roles

**System Buildout**
- Hardware
- software/system "shell"

**System Maintenance**
- Security/integrity/backup

**System Configuration**
- Design/Entry/Q&A
- Access

**Data Entry**
- Database assistance
- Best practices

**Data maintenance**
- Security/integrity
- Backup/Archival

## Program / PISFC, SIS

# Sample workflow for updating an existing data system

**Program / ITS**

**Update Planning**
- Scope
- Assign roles

**System update**
- Hardware
- Software patches

**Update Maintenance**
- Monitor for updates

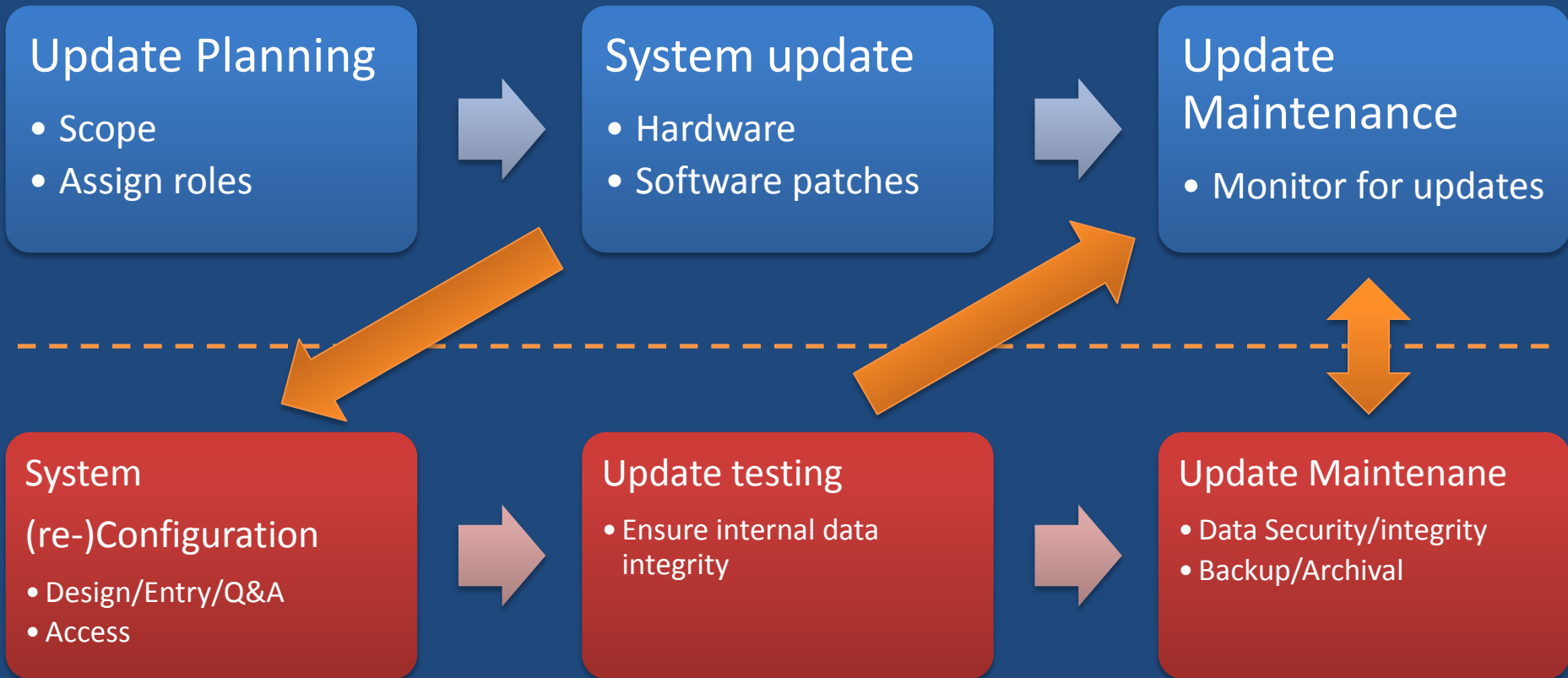**System (re-)Configuration**
- Design/Entry/Q&A
- Access

**Update testing**
- Ensure internal data integrity

**Update Maintenane**
- Data Security/integrity
- Backup/Archival

**Program / PISFC, SIS**

# Major PIFSC Data Systems (ITS)

- PIFSC Oracle Enterprise Data System
- Contains numerous datasets used by researchers within and outside the center
  - Hawaii and American Samoa Longline Observer Data (LODS)
  - Hawaii and American Samoa Longline Logbook
  - State of Hawaii dealer data (landings)
  - Others
- ITS ensures operation of test, development, and production Oracle platforms

# Major PIFSC Data Systems (ITS)

- Geographic Information Systems (GIS) to model and display collected data

- ITS maintains ArcGIS server and ArcGIS client license manager

  - Server provides backend for center-built tools to deliver data

  - License manager allows use of concurrent ArcGIS client licenses

# PIFSC Data System Support (ITS)

- Storage area network (SAN) to store large datasets
  - Fishery independent survey data including video and acoustics
  - Fast data connection to servers
- Total system security to protect the integrity and confidentiality of all data stored at PIFSC
- Can provide assistance with other data systems, consultation **before** project begins crucial

# Data Systems Personnel (ITS)

| Staff member | Data-related tasks |
|---|---|
| Leonora Fukuda | Shared space, data backup (Aiea) |
| Wayde Higuchi | GIS License management, shared space |
| Russell Price | Oracle/GIS Hardware, data migration, patch management |
| Chad Sugimoto | Data storage and backup (Kapiolani) |
| Rossyn Tasaka | Large dataset storage and backup (Kewalo) |
| Richard Uyeda | Networking, Oracle/GIS Hardware, data migration, patch management |
| Ron Yoshimoto | Software installation, End user support |
| Scott Wong | Software installation, End user support |

**Current model will change with upcoming move to IRC (reduction field offices)**

# Data System Collaborations

- Assistance with recent security testing and release of Reefbox data access tool developed by CRED

- Assistance with migration of data from historic intranet to new Drupal instance

- Assisting WPacFIN with transfer of data from MySQL field databases to PIFSC Enterprise Database (Oracle)

- On-going transfer of Fishing Ecosystem Analysis Tool (FEAT) to UH PacIOOS

# FY13 Data System Costs ($89K) - ITS

- Oracle License  $24K

- ESRI License $25K

- Backups: media/software/hardware $40K

- A refresh of Oracle hardware servers is highly recommended, yet is prohibitive in current budget climate ($250K)

- Several systems within ITS now virtualized
  - working to virtualize more systems inc. Oracle
  - Still required initial monetary commitment

# PIFSC Data System Challenges

- Current model has no centralized data management staff

- Data system tasks handled by ITS, SIS, or divisions

- Stewardship of developed applications

- Staffing
  - Planned FY13 addition of Database position postponed due to budget and hiring freeze

- Funding Sources (rely on division overhead)

# The PIFSC DMSC

- PIFSC Data Management Steering Committee developed from 2010 ER recommendation
- DMSC comprised of 1-2 members from each PIFSC program
- ITS has vested interest and high involvement in DMSC
    - Participation in DMSC as well as all Working Groups
- ITS responsible for FY13 milestone
    - "Oversee, direct, and manage the PIFSC DMSC's effort to create a Data Lifecycle Framework that documents the complete process of data management at PIFSC"

# The Way forward?

- Possibly use DMSC as management body for task-driven assignments

- DMSC could identify tasks, request assistance from Science Center
  - DMSC completed HR survey July 2012, identified center staff with data management skill sets
  - First two working groups now active to create refined PIFSC DM Policy and User Guide for PIFSC data lifecycle

- Imperative to continue collaboration and work to bridge ITS and SIS data management efforts