Doing More With Less at the Center for Biological Informatics

Jeff Falgout
Systems Administrator
USGS Center for Biological Informatics

OFWIM 2009 Annual Conference
Seattle, WA
September 16, 2009

Act I
Background

What is CBI?

The USGS Center for Biological Informatics (CBI), through such Programs as the National Biological Information Infrastructure, facilitates access to and application of biological information to:

- Describe the nation's biological resources
- Enable the sound management of natural resources
- Support ecological research and education

What CBI Does

- National Biological Information Infrastructure
  - www.nbii.gov, my.nbii.gov
- Gap Analysis Program
  - gapanalysis.nbii.gov
- Integrated Taxonomic Information System
  - www.itis.gov
- Land Use History of North America
  - biology.usgs.gov/luhna/
- Vegetation Characterization Program
  - biology.usgs.gov/npsveg/
- Fire Research and Management Exchange System
  - frames.nbii.gov

What CBI Does

- Significant Portal infrastructure
  - Collaboration and Content Management
- Home for orphan projects
  - National Park Lichens
    - www.nbii.gov/nplichen
    - Retiring scientist didn't want his career's work to disappear
- Hosting for independent projects
  - Fire Enhanced Runoff and Gully Initiation Model
    - frames.nbii.gov/fergi

CBI's Infrastructure

- 90 servers (Linux and Windows servers)
- 8 network devices
- ~ 20 workstations
- ~ 72 TB raw storage capacity
- 60 slot tape library
- 102 websites
  - ~ 3300 named users
- ~ 37 million hits (including crawlers) Aug 2009
- ~ 1 TB bandwidth consumed Aug 2009

CBI Support Staff

- Two Portal Administrators
- One GIS Administrator
- Three Developers
- Two Systems Administrators

- Grand total of ***Eight*** support staff

Act II
Strategies for “Doing More with Less”

Develop and Enforce Standards

- Avoid "Jack of All Trades, Master of None"
  - You must be a master of your own domain
- Hardware
  - Decide and stick with a single manufacturer
    - or –
  - Build your own (Workstations and Servers)
- Software
  - Standard operating system (including configuration)
  - Compatibility with O/S, other software packages

Development Languages

Define Service Level Agreements (SLA's)

- Everyone involved must have an understanding of what is expected and agreed upon
  - Life or Limb may depend on services you provide
  - Funding may depend on meeting SLA's
  - Avoid giving in to unreasonable requests from “special” people/groups
  - Ex: the word “Fire” often gets attention
    - National Inter-Agency Fire Center
    - Fire Research and Management Exchange System

Business Processes

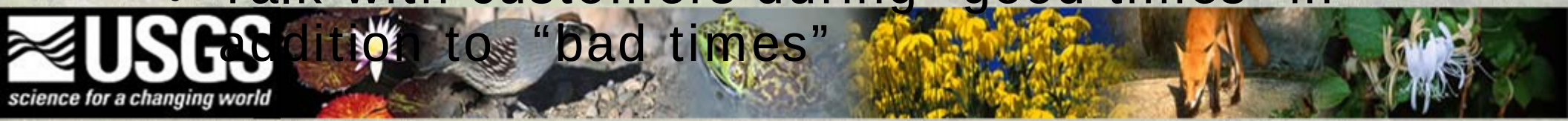
- Define, Document, and Communicate these processes:
  - Development, Deployment, Production, Security
- Get as much buy-in and feedback as possible:
  - Upper Management, Project Managers/Owners, Developers, Support Staff
- Critical to enforcing standards
  - Provides method of review for standards compliance

Real World Example: Defining CBI's Application Deployment Process

- Without process, movement of applications from development to production was a “Free for All”
- Squeaky wheels got greased first - :(
  - Interestingly, squeaky wheel numbers increased dramatically
- Six months to materialize process – drew from business experience (Six Sigma), developer habits and organizational needs
- Result not popular, but dramatically reduced number of “squeaky wheels” and created a predictable time line for application deployment

Communication: Customers and Stake Holders

- Open communication is critical
  - Planning, maintenance and unexpected outages
  - End result reduces support calls/tickets
  - Shows you are working even though things are stable
- SA's mantra: “Success go unknown, but failures are quite spectacular” - Success needs to be known
- Involvement in project planning crucial – no surprises
- Prevent “out of sight, out of mind”
  - Talk with customers during “good times” in addition to “bad times”



Communication: Parent Organization

- Awareness of upcoming changes coming from above
  - Policy as well as technical
- Compliance with parent organization's standards
  - Doing it right the first time limits re-engineering
  - Positions you to handle upcoming mandated changes
  - We are all part of the bigger picture:
    - US Citizens → President → Dept of Interior → US Geological Survey → CBI

Hardware Acquisitions

- It tends to be easier to acquire capital equipment than adding staff in government agencies
- Ease of hardware acquisition can be both a blessing and a curse
- Acquisitions need to be thought out and planned
- Try to leverage acquisition money to include consulting services thereby bootstrapping installation

Architecture: Scalable

- Allows for addition, replacement, or removal of hardware capacity during business hours
  - At least that's the goal – bad things can and do happen
- CBI uses n-Tier model
  - Web/Presentation tier
  - Application tier
  - Database tier
  - Storage tier
- Each layer is independently scalable

Architecture: Fault Tolerant

- Political fallout from failure during core business hours can be lethal
- Systems must gracefully handle failure
- Reduces need to call/page staff in middle of night
  - Goal is to avoid critical staff burnout
- Dependant on adherence to standards
- Capability to reduce vendor support costs
  - Move from 24x7x365 support contract to business hours support contract (can be significant cost difference)

Architecture: Virtualization

- CBI shies away from the use of virtualization in production
  - Decreases architecture complexity
  - Decreases technical expertise requirement
  - Virtual machines tend to multiply like rabbits
  - Virtualization is in use for development
- End result can be achieved with intelligent use of other technologies
  - Web server virtual sites
  - Multiple Java Virtual Machines (JVMs)

Beware the “Buzzword Kool-Aid”

- “Cloud Computing”, “Social Computing”, “Web 2.0 Compliant”, etc
- Vendors are very eager to sell you stuff
  - Be honest with your needs and requirements
  - Find vendors that you can trust
- The whiz-bang package may be way more than you can handle with staffing levels or hardware capacity
  - High potential for snowballing

Training

- "What if you train your staff and they leave?" .... "What if you don't train them and they stay?"
- Empowers your staff
- Facilitates intelligent decision making
- Capitalizes on features of current investments
- Increases competency with software/hardware making it easier to meet SLA obligations
  - One week of training (scheduled) more enjoyable than weeks or months of frustration and missed deadlines

Be Wary of “Cutting Edge”

- “Tried and true” best for staff shortages
  - What everyone else is doing is sometimes good
  - Let others fight the battle
- Greater availability of resources from the “Beaten Path”
  - Supported Architectures
  - Documentation
  - Mailing lists
- Cutting edge usually has higher number of bugs/issues

Open Source Software (When Appropriate)

- Price of software usually free or very low
- Most enterprise applications have support programs available
- Increased potential for inter-agency interoperability
- Keep in mind cost of Open Source Software is time and effort

Infrastructure Monitoring

- Monitoring your systems is critical
  - Availability
  - Performance
- Know about your problems before someone else does
  - SLA obligations – clock is ticking
  - Perception
  - Political fallout
- Helps with predicting failures
- Allows for proactive capacity planning

Fire Fighting

- Every effort should be geared to keep out of fire fighting mode
  - Snap decisions or panic management can come back to haunt you
- Automate as much as you can
  - Reduce human variable
- Appropriate Planning
  - Capacity
  - Compatibility
  - Failure recovery

Remember …

- Define your obligations and Service Level Agreements
- Understand your obligations
- Document your processes and requirements
- Capitalize on vendors – consulting, installation, and training
- Always look for consolidation opportunities
- Communicate, Communicate, Communicate
- Overall goal is to reduce variables in environment

Questions

?