A Distributed Collaborative Workflow Based Approach To Data Collection and Analysis

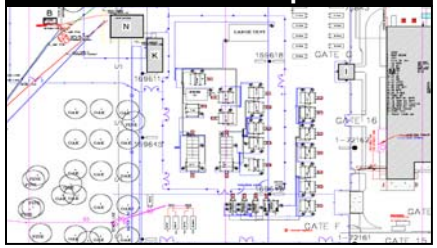
William Gerecke, Douglas Enas,
Susan Gottschlich

Background:
U.S. Central Command Deployable Headquarters

Integrated Equipment

CENTCOM Operations

CDHQ Concept



JOC

CDHQ – Capability

- **Provides forward deployable C3I – flexible AOR deployment**
- **Base for CENTCOM HQ split-based operations**
- **C2 for contingencies and reach-back**
- **Data collection and analysis**

CDHQ Ready for Deployment

Data Collection and Analysis (DCA)

- Traditionally used in Modeling and Simulation (M&S) and Command and Control (C2) exercises and operations
  - Compute and display Measures of Effectiveness (MOE) and Measures of Performance (MOP) - runtime operation
  - Provide analysis results for After Action Review (AAR) and related activities - offline operation
- US Central Command (CENTCOM) Deployable Headquarters (CDHQ) provided unique opportunity to explore DCA architecture and usage
  - Goal is to monitor health and performance of CDHQ enterprise.
  - PerfMon Collector collects Microsoft's Performance Monitoring data from servers and clients, SNMP Collector collects Simple Network Management Protocol and "ping" data from switches and routers.

C2/HQ DCA Installations

- CDHQ
  - Limited Prototype version installed during CONUS CDHQ Exercise – COTS used for gap fillers
  - Prototype version installed in Qatar during Internal Look ’03 Exercise
  - Utilized by our support staff in lead up to OIF and during OIF
- Raytheon Springfield
  - Current version installed and running on Raytheon Springfield office networks, continuously monitoring and analyzing
  - Ongoing installation on C2 Test Bed being stood up
- CPA Networks
  - Limited usage to support our work for CPA in Baghdad

CDHQ Lessons Learned

- Command HQ's are different than Tactical HQ's
  - Command HQ have requirements similar to a typical office environment
  - Minimum of four networks with different security classifications
- Feeding the Beast
  - At a headquarters level, the J6 builds a lot of status briefings.
  - Pictures, not words, to explain a problem, recommend a solution.
- IT Infrastructure is very dynamic
  - Not quite a mobile adhoc net, but…
  - Need to be able to isolate and troubleshoot a problem in minutes, not days
- Roles and Responsibilities are very dynamic
  - Surging and rotation make for a high staff turnover rate
  - Specialized software that requires training becomes shelfware
- Soldiers and augmentees are extremely resourceful

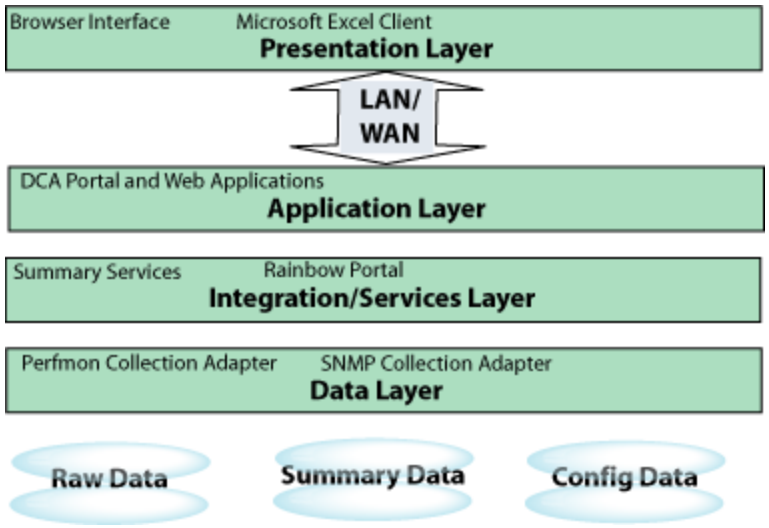
C2/HQ DCA Goals (address C2 Users)

- Teach-by-showing Training – quick demo instead of course or tutorial.
- Accommodate disparate user community
  - Wide range of skills and interests
  - Leverage ubiquitous COTS products with familiar & well-supported UI's
- Web-enabled, portal based user interface (UI)
  - Accessible from any web client anywhere on LAN
  - Analysis products "published" to DCA portal – facilitates distributed collaboration
  - Familiar (tabbed panel, breadcrumbs, etc.) mechanisms for easy navigation
  - Web-based configuration
- Separate content from code
  - Use XML strategies to specify configuration
  - Substantial capabilities provided that do not require software changes
- Right mix of horizontal vs. vertical technology

C2/HQ DCA Goals (address C2 technology)

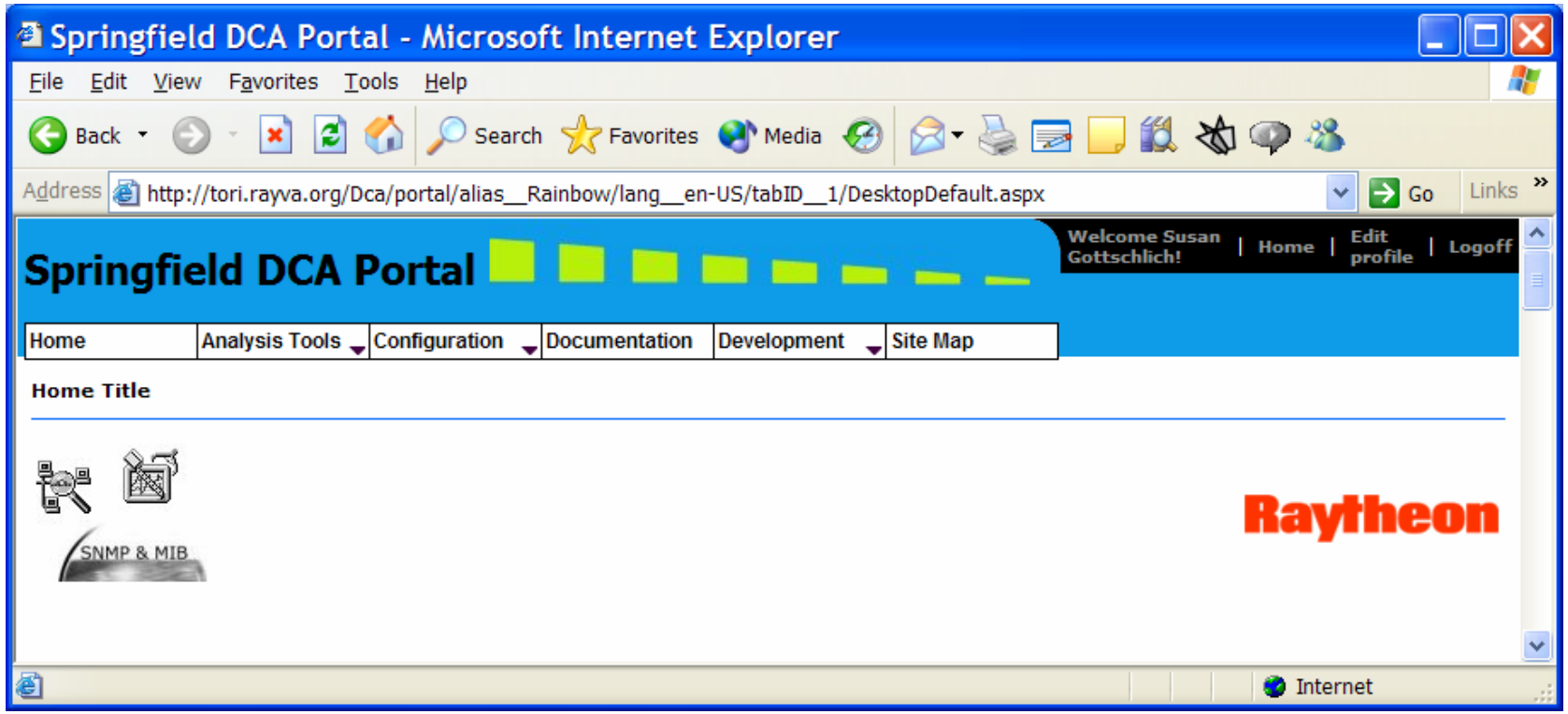
- Integrate COTS and Open Source technologies – don't reinvent
  - Focus effort on domain specific improvements
  - Take advantage of massive investment in technology and usability
- Fast, reliable data collection that supports runtime queries
  - Utilize distributed database technologies to collect and store large amounts of data
  - Support run time aggregation – preprocess data as it is collected so that queries return almost instantly.
  - Provide back up schemes – ability to backup aggregated data rather than raw data
- Modular, extensible, and distributable
  - Ability to add Collectors to collect different types of data (e.g. message traffic)
  - Easy to extend analysis capability
  - Different modules distributed to multiple machines on different local area networks (LAN's) and virtual LAN's (VLAN's)

C2/HQ DCA Architecture



- Multi-layered service oriented architecture
- Based on ASP.Net, entire system can be run on a single high-end laptop or distributed
- Utilizes COTS/Open source technology to provide powerful, easy to use, system modularized along industrial establish boundaries
  - Rainbow Portal (open source)
  - Excel (Microsoft)

C2/HQ DCA Portal



- Utilizes DUEMETRI Rainbow portal toolkit
- Main user interface to entire DCA system

C2/HQ DCA Portal (cont)

- Define 4 levels of users based on workflow/expertise
  - DCA Developer (Level 1) – Install & maintain DCA deployment
  - DCA Administrator (Level 2) – Configure DCA deployment
  - DCA Analyst (Level 3) – Provide DCA content (e.g. reports)
  - DCA User (Level 4) – Use DCA content
- Rainbow portal toolkit infrastructure provides important features
  - Cross-browser support for Netscape and Internet Explorer
  - Mobile device support for WAP/WML and Pocket Browser devices
  - Clean code/html content separation using server controls
  - Supports 14 foreign languages
  - Role-based security to control user access to portal content

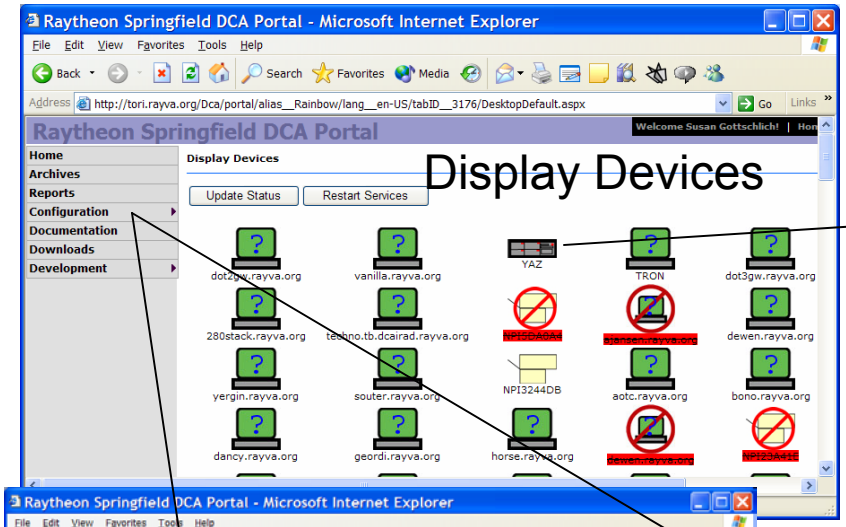
C2/HQ DCA Portal Customization

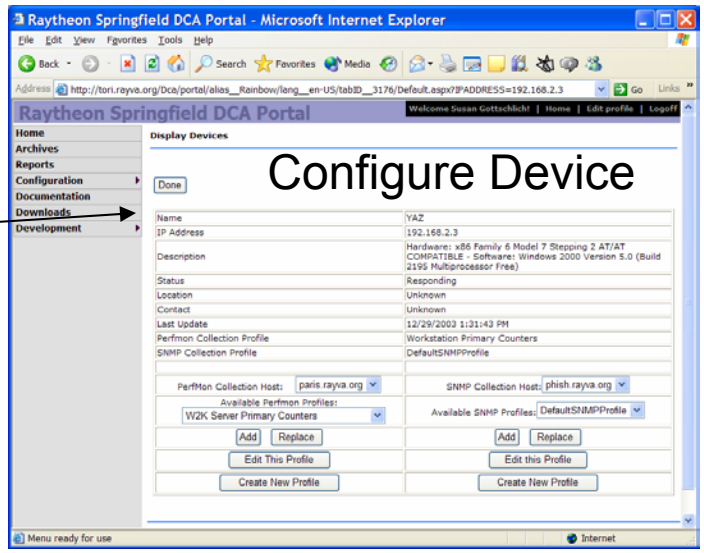
- DCA specific tabbed panel configuration/content
- 5 DCA web modules developed
  - 3 DCA Configuration modules
  - 1 DCA Report module
  - 1 DCA Archive module
- Run time configuration and configuration management
  - Design principle to separate configuration data from code for maximum flexibility at installation site.
  - DCA Data managed separately
    - Collected data
    - Configuration data
    - Customization data

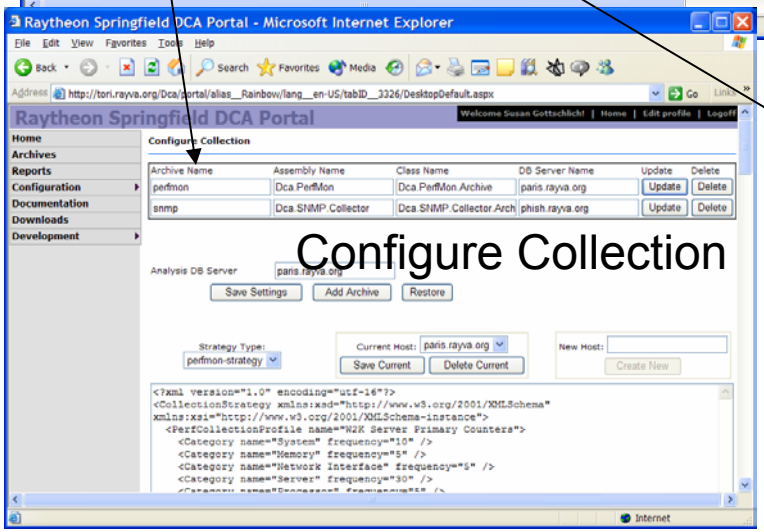
DCA Portal Tabs

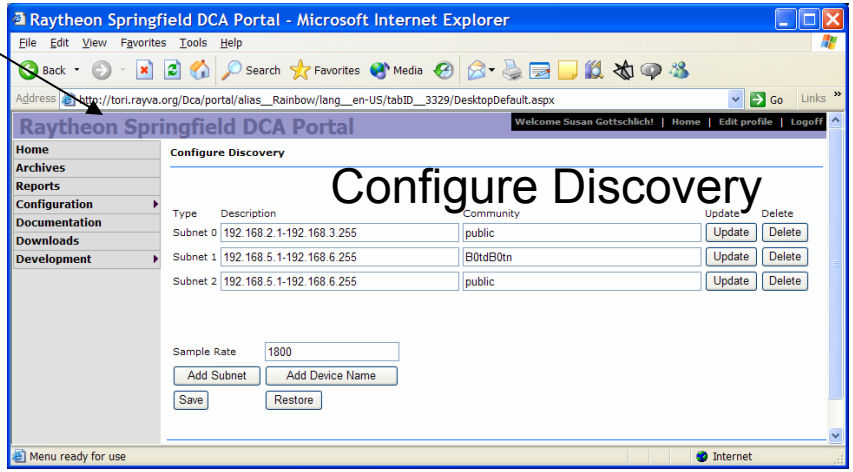
- Configuration –
  - Collection – How much data/what data is collected from each device
  - Discovery – Where to look for new devices
- Archives –
  - Run queries
  - Investigate problems
  - Create Reports
- Reports –
  - Look at Live View Reports
  - Upload, download, or edit Reports
  - Schedule Reports
- Documentation, etc.

C2/HQ DCA Configuration





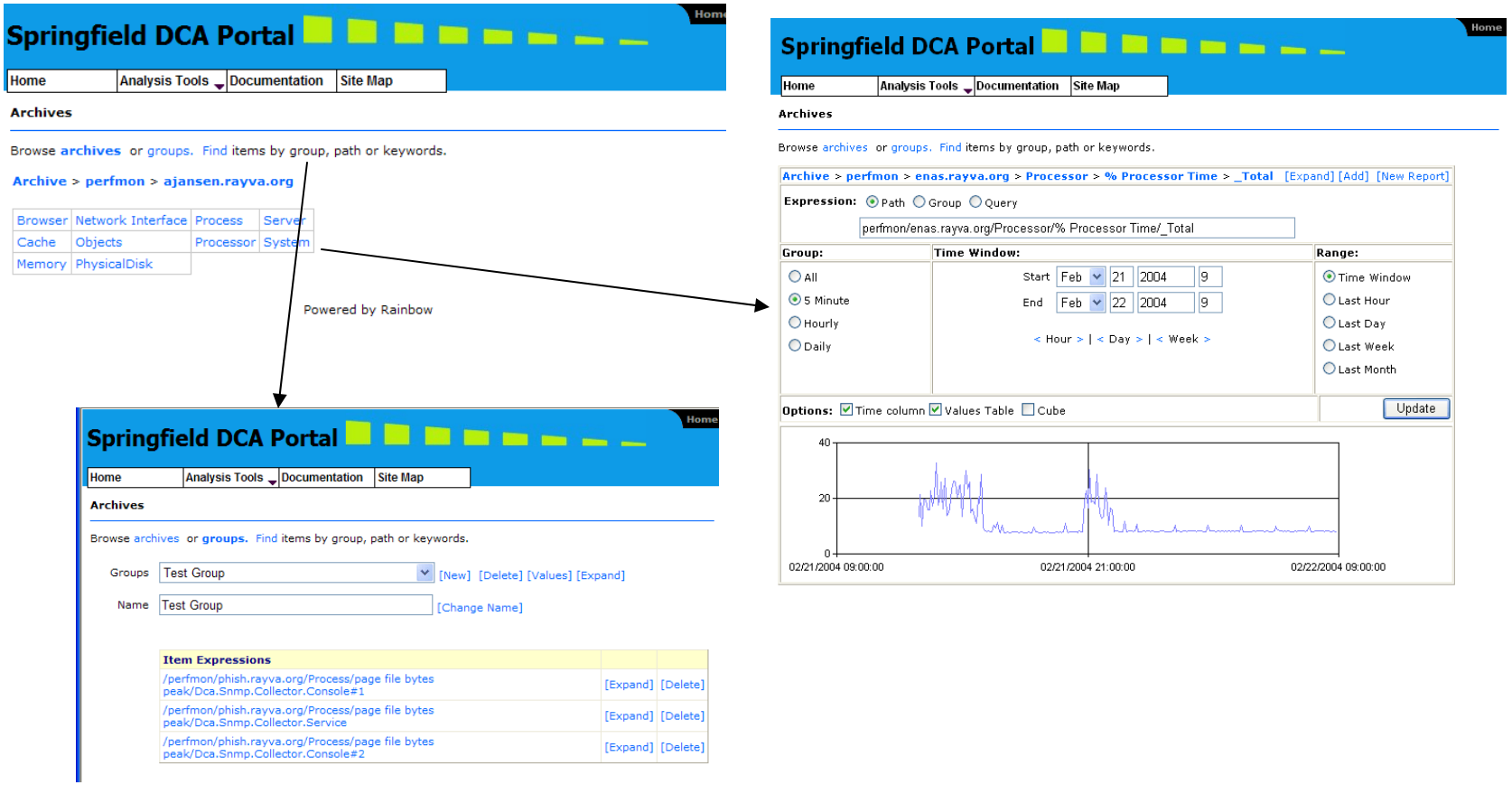




C2/HQ DCA Configuration (cont.)

- Automatic Discovery of Every Device on Network
  - Lessons learned from CDHQ is enterprises are very dynamic, too tedious to manually track and update.
  - Discovery strategy is itself configurable from web portal
  - Once discovered, a device can be:
    - Automatically added to collection strategy
    - Manually added to collection strategy using individual configuration
    - Manually added to collection strategy using bulk configuration
- All configuration possible through web portal
- All configuration data maintained in SQL database for easy maintenance, portability.

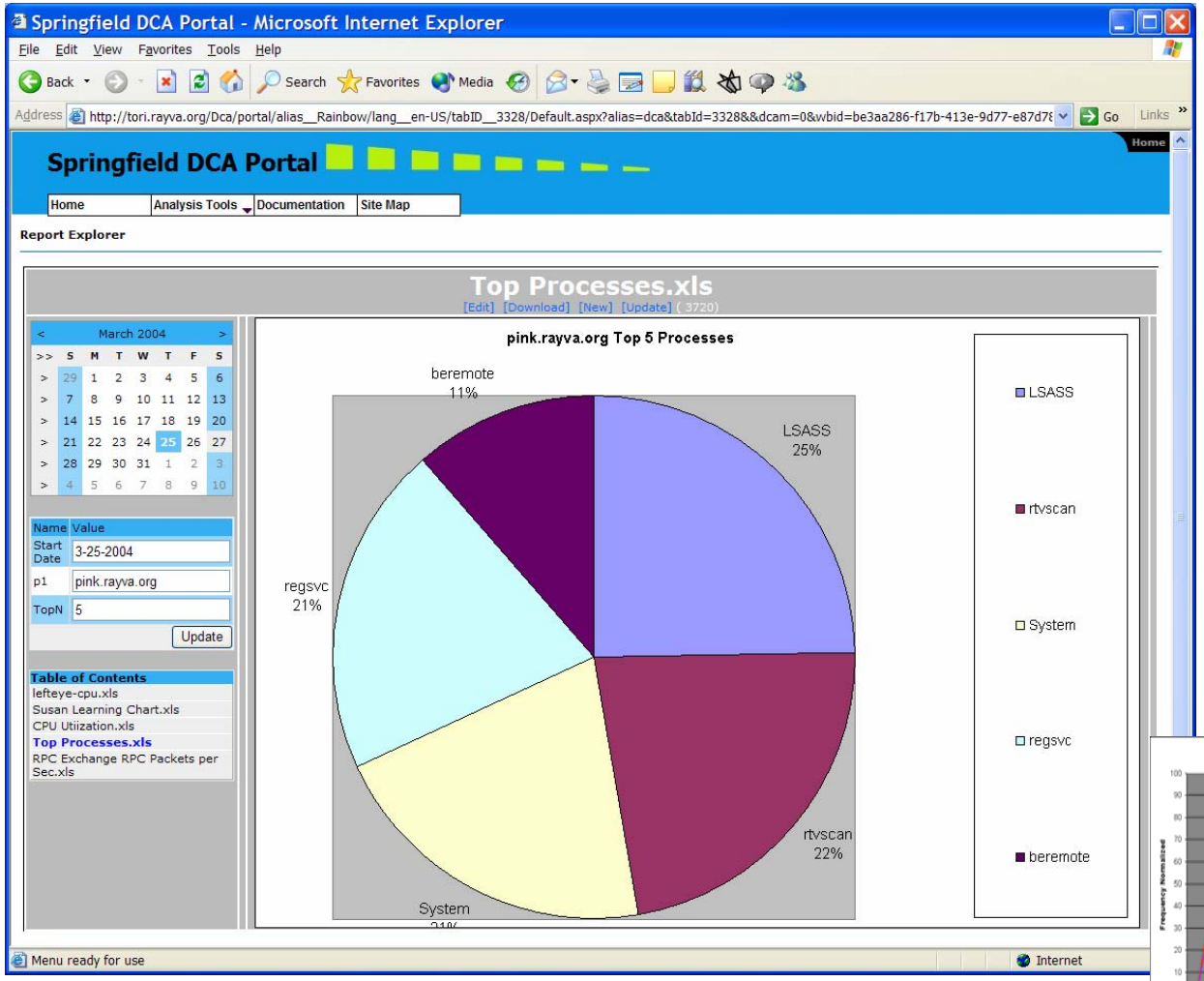
C2/HQ DCA Archives

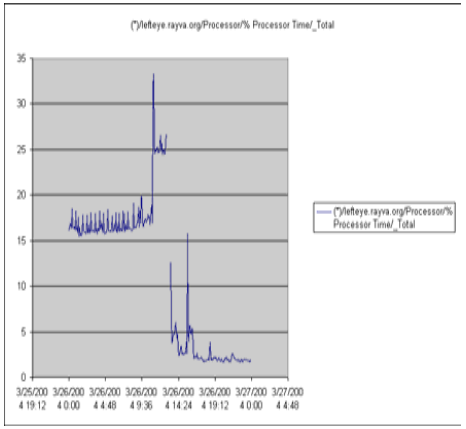


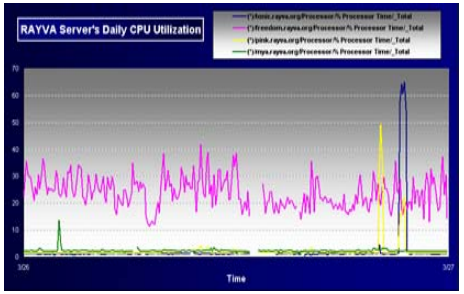
C2/HQ DCA Archives (cont)

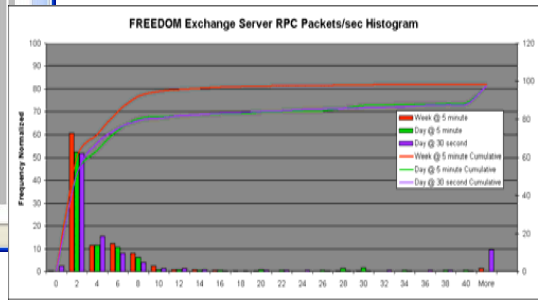
- Plug and Play - Any Collector that supports DCA Archive interface can be plugged in, will automatically be assessable from DCA Archive browser.
- Run time query capabilities – DCA Archive interface supports trouble shooting and also report template creation.
  - Quick look reporting of query results
  - Supports web-query interface for Excel charting
- Usability features
  - Creation of query groups
  - Wildcard and full-text search capabilities
  - Web-based specification of query parameters
- Run time aggregation to greatly speed up most queries

C2/HQ DCA Reports





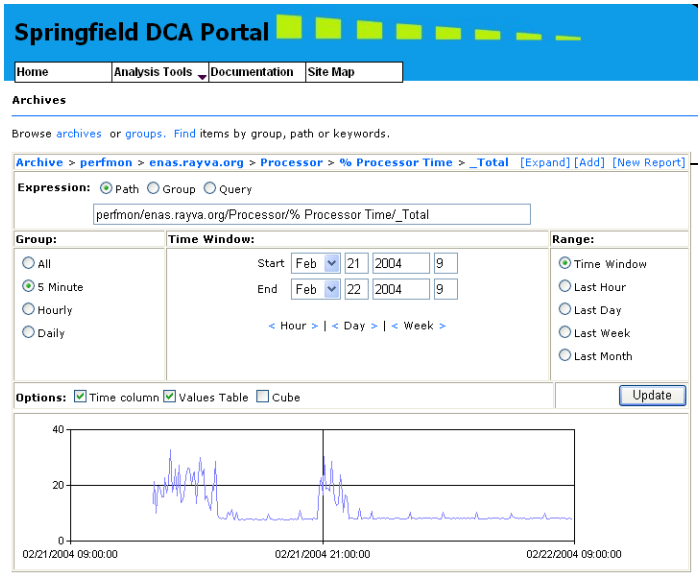




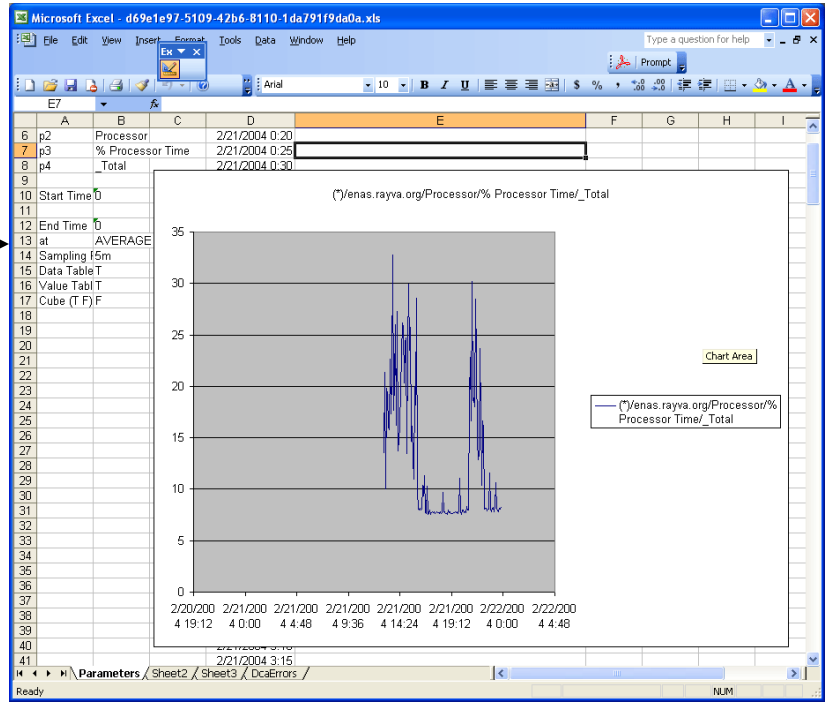
C2/HQ DCA Reports (cont.)

- Level 1-3 users can create report templates (Excel workbooks) and publish to DCA Portal
- All users can schedule reports, view history, download templates or reports.
- Reports can be “Live View” or scheduled for one-time, multiple time, or recurring.
  - As soon as time period of schedule is complete, report will be added to history.
  - Reports can be scheduled for the past, future, or both.

C2/HQ DCA Report Creation

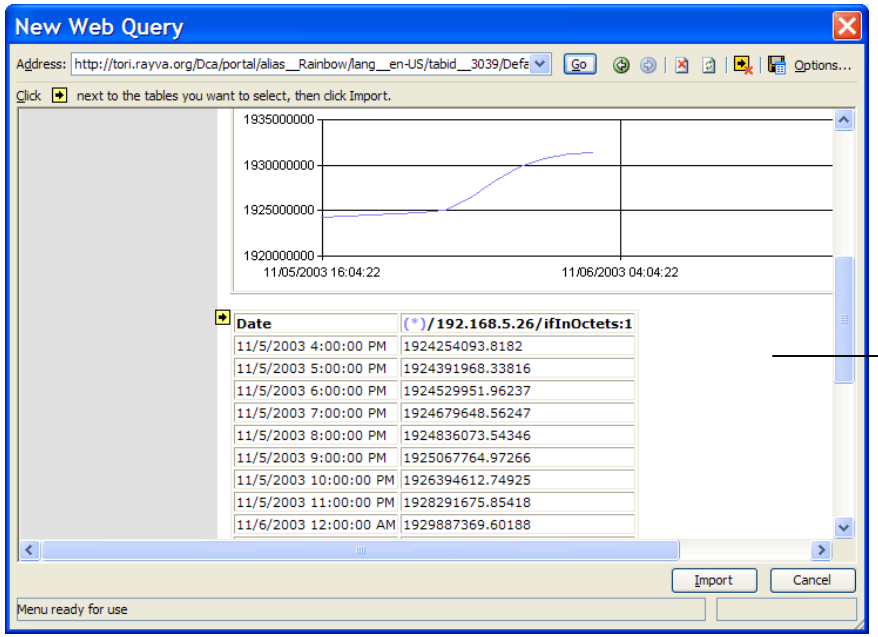


Step 1: Run web query through web page, save as New Report (Excel file).

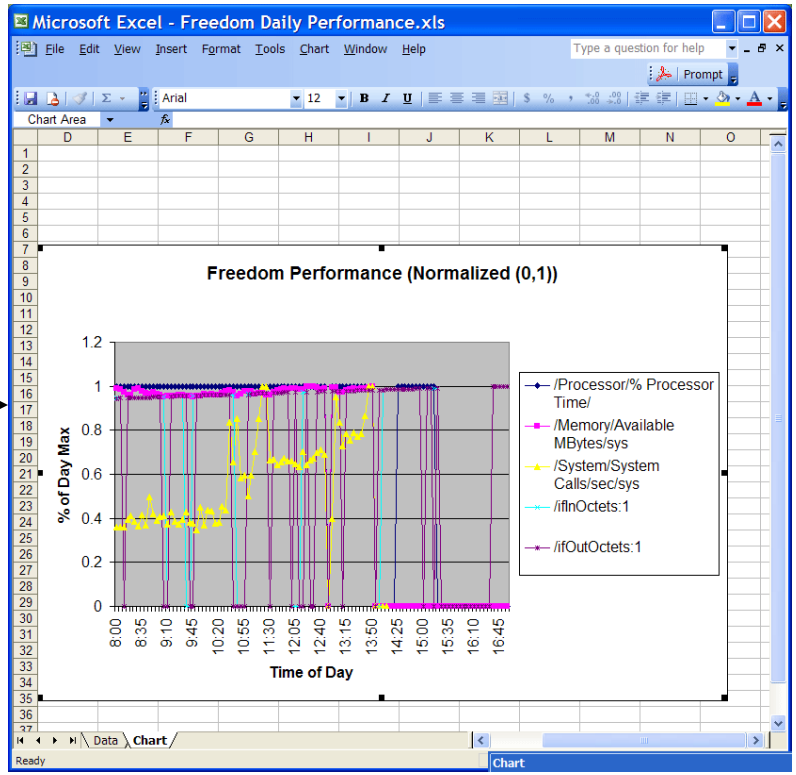


Step 2: Use as is or modify using Excel application.

C2/HQ DCA Report Creation – Alternative Approach



Step 1: Run web query through Excel embedded browser



Step 2: Create report visual using Excel charting and features

C2/HQ DCA Report Creation (cont.)

- MS Excel used for Report Template interface
  - Report templates are built using Web Query feature in Excel.
  - MS Excel charting is extremely powerful, with many ease-of-use features, allows a novice or sophisticated user to create products at runtime.
    - Standard strip chart
    - Dynamic histogram
    - OLAP cubes/Pivot table
    - Data tables
  - Excel workbook files are uploaded (published). Can be downloaded or viewed on site.
  - Web query parameters include time span parameters. Report scheduler alters these parameters to product reports for different time periods.
- Alternatives
  - Web services interface also provided, required Level 1 user
  - Any report building tool that can use web services or web queries.

DCA Summary

- Presented a Data Collection and Analysis system that provides broad horizontal functionality, specific instantiations used for C2 HQ's
- Lessons learned from CDHQ and other operations has driven design decisions
  - IT infrastructure is very dynamic
  - Roles and Responsibilities are very dynamic
  - Feeding the beast
- Multiple DCA deployments stood up and used

DCA Future Plans

- Extend and optimize user interface
  - Add more features to interface (e.g. different analysis products)
  - Optimize “click path” for WAN based usage
  - Powerpoint, Word interface
- Extend data collection
  - Message traffic (e.g. HLA, US/MTF, email, chat)
  - Shared drive usage
- Integrate with other web-enabled systems

Acronyms Used

- DCA – Data Collection and Analysis
- CDHQ – CENTCOM (US Central Command) Deployable Headquarters
- CPA – Coalition Provisional Authority
- MS – Microsoft
- SNMP – Simple Network Management Protocol
- C2/HQ – Command and Control Headquarters
- OIF – Operation Iraqi Freedom
- LAN/VLAN – Local Area Network, Virtual LAN
- C2 – Command and Control
- M&S – Modeling and Simulation
- MOE/MOP – Measure of Effectiveness/Measure of Performance
- OLAP – Online Analytical Processing
- WAP/WML – Wireless Application Protocol/Wireless Markup Language
- COTS – Commercial off the shelf
- UI – User interface
- CONUS – Continental US
- HLA – High Level Architecture
- US/MTF – United States Message Text Format