



# News from EPA's National Computer Center

John McQuaid

EPA Office of Environmental Information

2006 AQS Conference



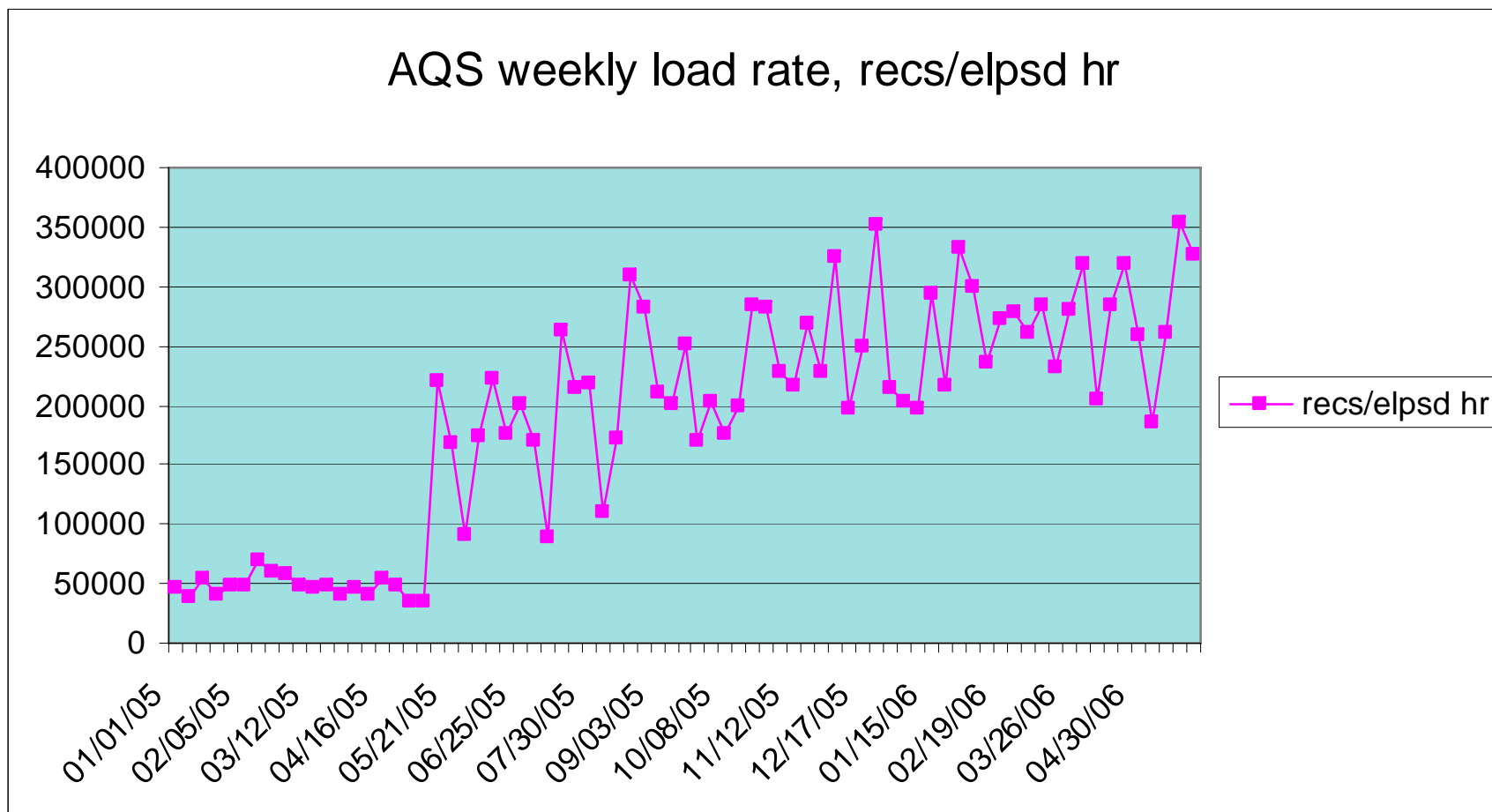
# Hardware Refresh - Update

What we promised:

- Replace entire National Computer Center Public Access and Intranet servers
- State of the art processors, doubling capacity
- Robust, flexible architecture allowing dedicated logical server partitions and adding capacity as required
- 20 TB Storage Area Network (SAN)



# HW Refresh - Update



# Hardware Refresh - Update

What we delivered:

- the average load job is taking *5 to 7 times less time* than it did on the old servers.
- the old average weekly load rate was between 35- 50,000 records/hour; now it is about 200-300,000 records/hour
- agencies that submit very large load jobs have noticed the improvement and are pleased



# Hardware Refresh - Update

What we delivered:

- the average retrieval is taking *about 2-3 times less time* to run than it did on the old servers.
- the improvement is most noticeable for the larger retrievals (e.g., multi-year, nationwide retrievals) that are typically run by EPA staff.



# On the Horizon: Enterprise Tools

- Identity and Access Management
- Central Data Exchange
- EPA Portal
- System of Registries
- Extract Transform and Load (ETL) Tool
- Web Services Publishing House
- Business Intelligence and Analytics



# Identity and Access Management

- Register users with a single sign-on;
- Store user-specific information;
- Assign users to groups/roles;
- Set up user accounts in applications;
- Provide credentials to users;
- Maintain information on user groups'/roles' authorized access;
- Authenticate the identity of users who request access;
- Control access to resources based on users' group/role



# EPA Portal

Gateway to relevant environmental and human health information systems:

- Access only authorized content;
- Single sign-on to access all information resources;
- Customize home pages to support user roles;
- Personalize capabilities to allow for individual preferences and greater efficiency;
- Access information to perform business intelligence analyses;
- Collaborate with other users to enhance workflow;
- Perform content management/search/browse functions; and
- Receive alerts, subscriptions, and news services.





# Business Intelligence and Analytics

EPA created a Business Intelligence & Analytics Center to:

- Facilitate sharing and reuse of information, knowledge, and experience in the adoption and management of BIA technology across the Agency
- Enhance the Agency's analytic and decision making capabilities
- Achieve economies of scale in the implementation, maintenance, and
- use of BIA tools and applications

The Business Intelligence and Analytics Center utilizes:

- Business Objects
- SAS
- Informatics



# Extract, Transform, and Load (ETL) Tool

Enterprise Architecture has focused on a Data Warehousing Bus Architecture, utilizing a Data Integration Platform and Business Intelligence tools.

- Extracts data from a wide variety of source databases using native interfaces, flat-file interfaces, or an FTP interface.
- Profiles and cleanses the data and identifies data anomalies.
- Transforms the data, computes summaries and aggregates, and computes any derived data.
- Loads the data into the target data warehouses and marts.
- Generates and manages a central metadata repository.
- Supports the import of data models used to define the physical organization of data.
- Utilizes the integrated data warehouse administration tools to support rapid response to end-user requests.

