



Engineering Analysis Systems for Cyber Security

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Copyright Information

Copyright 2022 Carnegie Mellon University.

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

The view, opinions, and/or findings contained in this material are those of the author(s) and should not be construed as an official Government position, policy, or decision, unless designated by other documentation.

NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution. Please see Copyright notice for non-US Government use and distribution.

GOVERNMENT PURPOSE RIGHTS – Technical Data

Contract No.: FA8702-15-D-0002

Contractor Name: Carnegie Mellon University

Contractor Address: 4500 Fifth Avenue, Pittsburgh, PA 15213

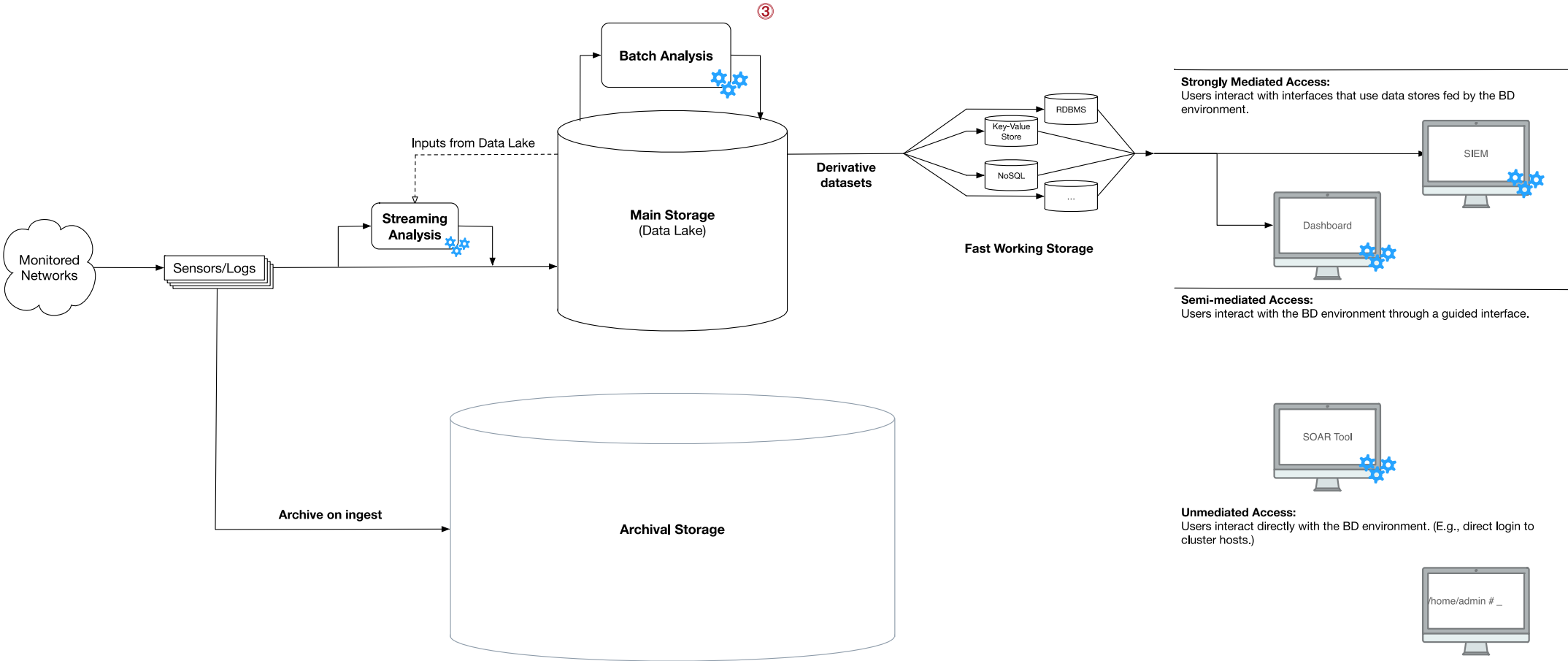
The Government's rights to use, modify, reproduce, release, perform, display, or disclose these technical data are restricted by paragraph (b)(2) of the Rights in Technical Data—Noncommercial Items clause contained in the above identified contract. Any reproduction of technical data or portions thereof marked with this legend must also reproduce the markings.

This material may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other use. Requests for permission should be directed to the Software Engineering Institute at permission@sei.cmu.edu.

Carnegie Mellon® and CERT® are registered in the U.S. Patent and Trademark Office by Carnegie Mellon University.

DM22-0426

Big Data (BD) Environments - A Reference Architecture for Cyber Security Operations



User Roles



Realtime analysts

Realtime analysts process incoming events with an emphasis on efficiency and rapid response. They should have all the data they need to do their jobs readily available, without having to think about the details of its storage and retrieval.



Non-realtime analysts

Non-realtime analysts use their experiences performing realtime analysis to create/extend new tools or automate manual workflows for improved effectiveness. They require access to the BD environment, but are not data administration/analysis experts.



Data scientists

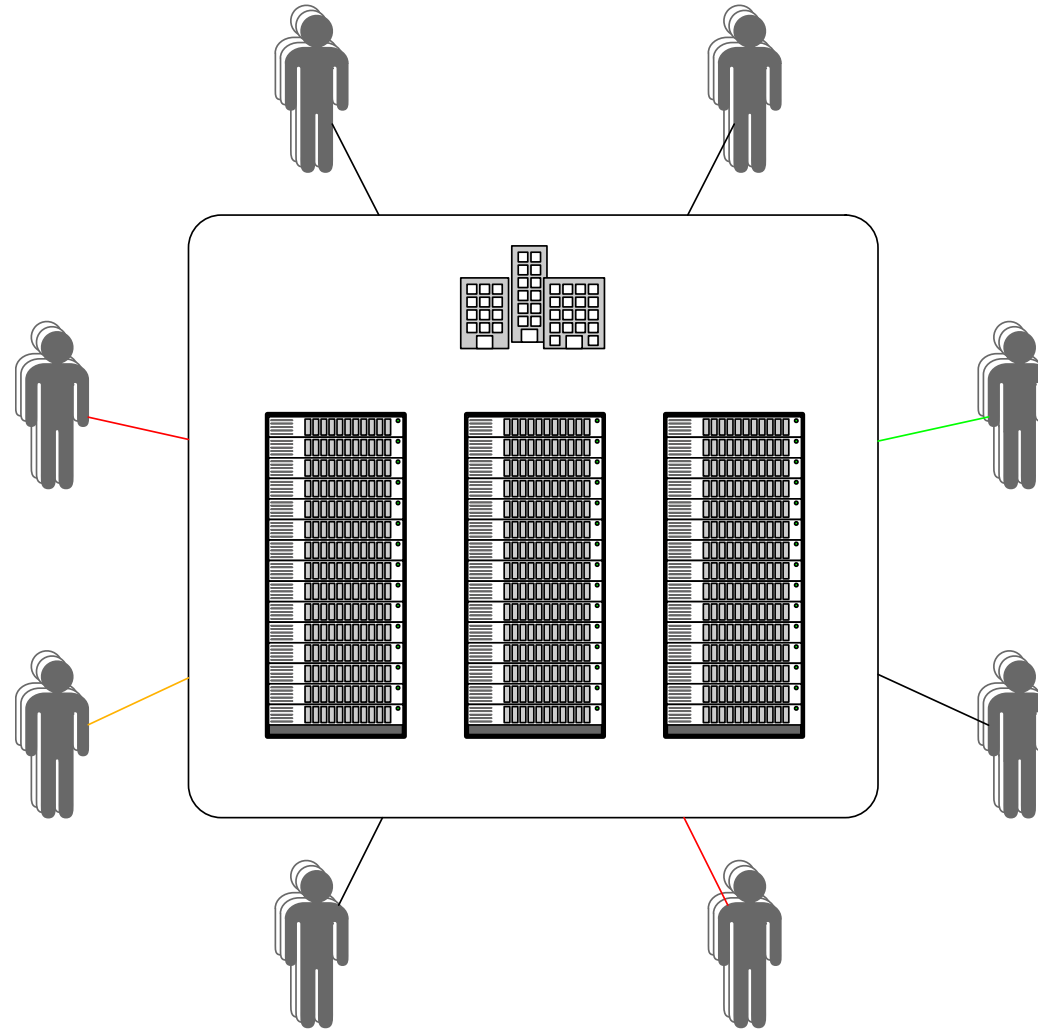
Data scientists specialize in data manipulation and analysis. They may sometimes appreciate guided access to a BD environment, but require the ability to perform arbitrary transformations on data.



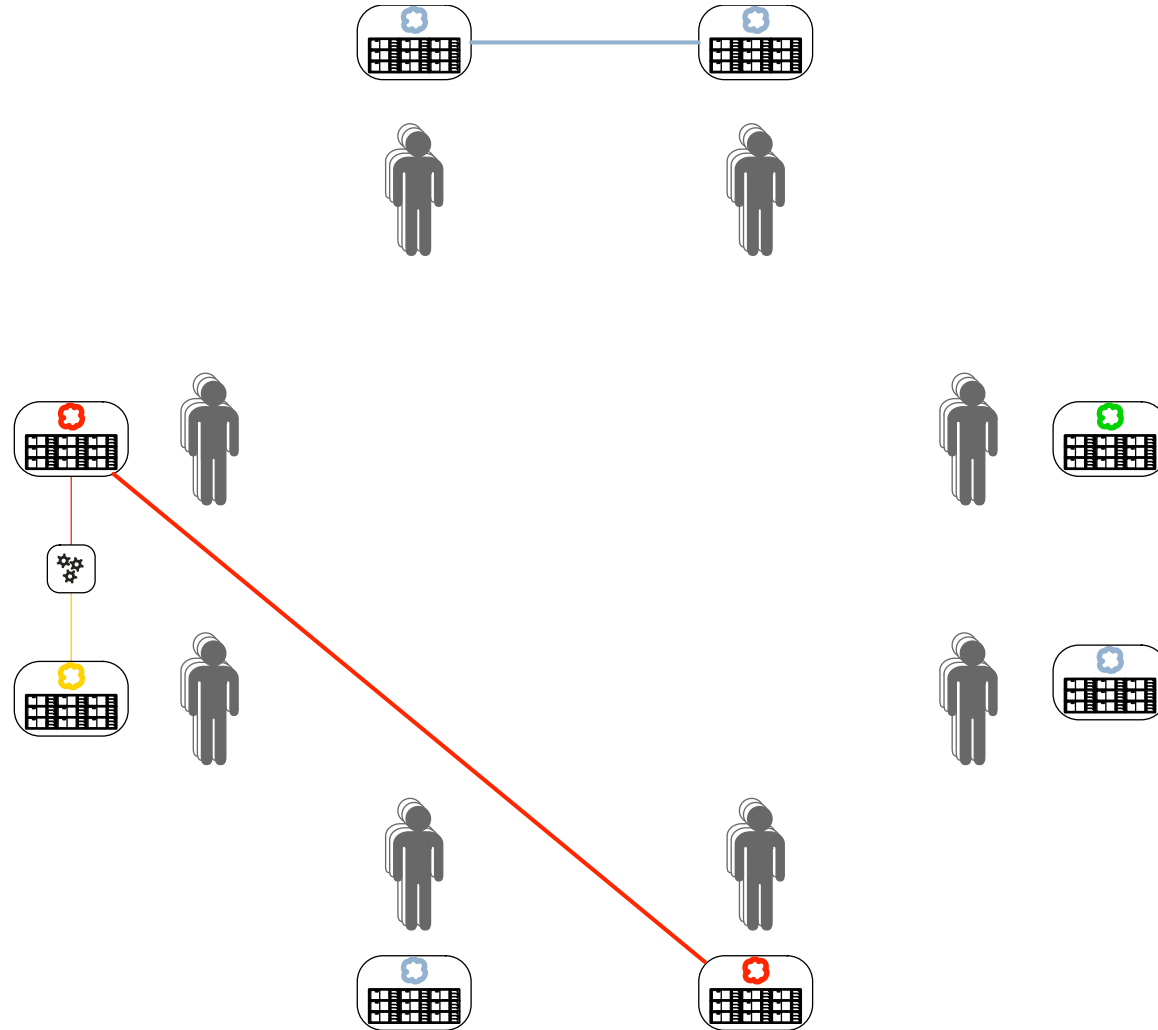
Administrators

Administrators make changes to the BD environment itself to make it run more effectively for other stakeholders.

Multitenant Model of Big Data Processing



Parallel Tenant Model of Big Data Processing



Aggregation-only Big Data Ingest

