Kitsuregawa, Toyoda-Group

Kitsuregawa, Toyoda, Nemoto, Ikoma, Goda, Yoshinaga, Itoh LAB.

[Platform for Real World and Social Big Data Fusion]

http://www.tkl.iis.u-tokyo.ac.jp/

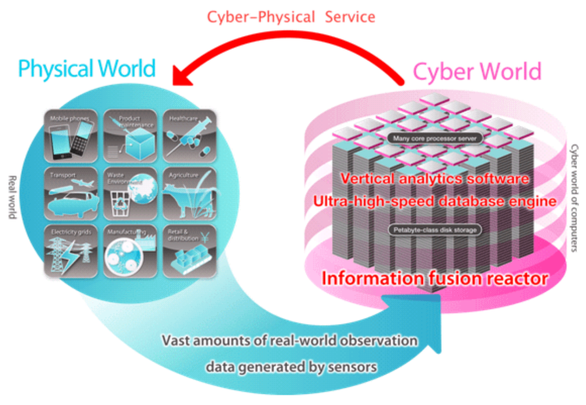
Database Engineering, Web Engineering

Information and Communication Engineering

Our Lab has focus on completely novel and surprising research on system software, advanced applications, hardware and algorithms, which is the basis of the technology for handling large quantities of data. We have been developing high performance database engine based on a novel out-of-order execution principle, ultra-large-scale cyber space mining systems, and 25 petabyte-scale global environment information system based on database engineering.

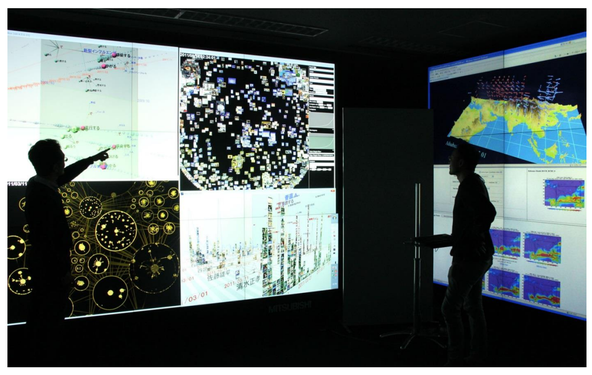
Development of High Performance Database Engine for Big Data Era

Our research group has been developing Ultrafast Database Engine based on a novel out-of-order execution principle. This new database engine has the capability to achieve significant performance boosting for analytical queries in the Big Data era. The group has been also constructing an experimental system of next-generation strategic social services (cyber physical services) to clarify effectiveness of the new database engine.



Ultra-large-scale Cyber-Physical Space Analysis System

Our lab has continuously collected Japanese Web pages since 1999 and has constructed a Web archive system including about 30 billion pages, 2 billion blog articles, and 25 billion tweets. Based on this peta scale archive, we are developing structural, contents, and temporal analysis systems including information diffusion extraction, inter-media comparison, and real-time deep text analysis. The cyber space information is integrated with physical space information such as mobility data for traffic analysis. Results of analysis can be interactively visualized on a large-scale high-resolution display wall.



Huge-scale spatio-temporal visualization system on the display wall

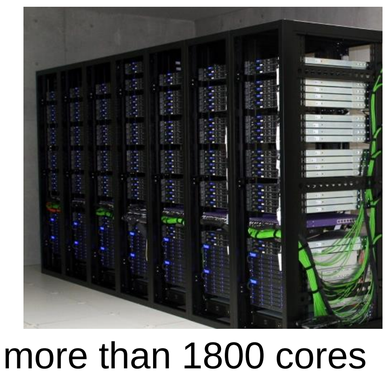
Over 25 Petabyte Global Environmental Information Fusion System

It is necessary to access usable information on the environment to deepen our understanding of the earth environment. We have been developing a large scale global environmental information fusion system for data integration and analysis that includes the supporting functions of life cycle data management, data search, information exploration, scientific analysis, and partial data down-loading.

Over 25 petabyte global environmental information storage

Reliable and Power-saving Network Control Technologies for Cloud Computing

Cloud computing -- large-scale on-demand distributed computing environment -- has become very popular. We are trying to develop a new technology to construct cloud applications with dynamic load balancing and fault resilience. Large-scale search problems have some difficulty to apply these techniques, therefore we're focusing on distributed computing platform to develop such applications with ease.



Cloud Computing environment that has more than 1800 cores