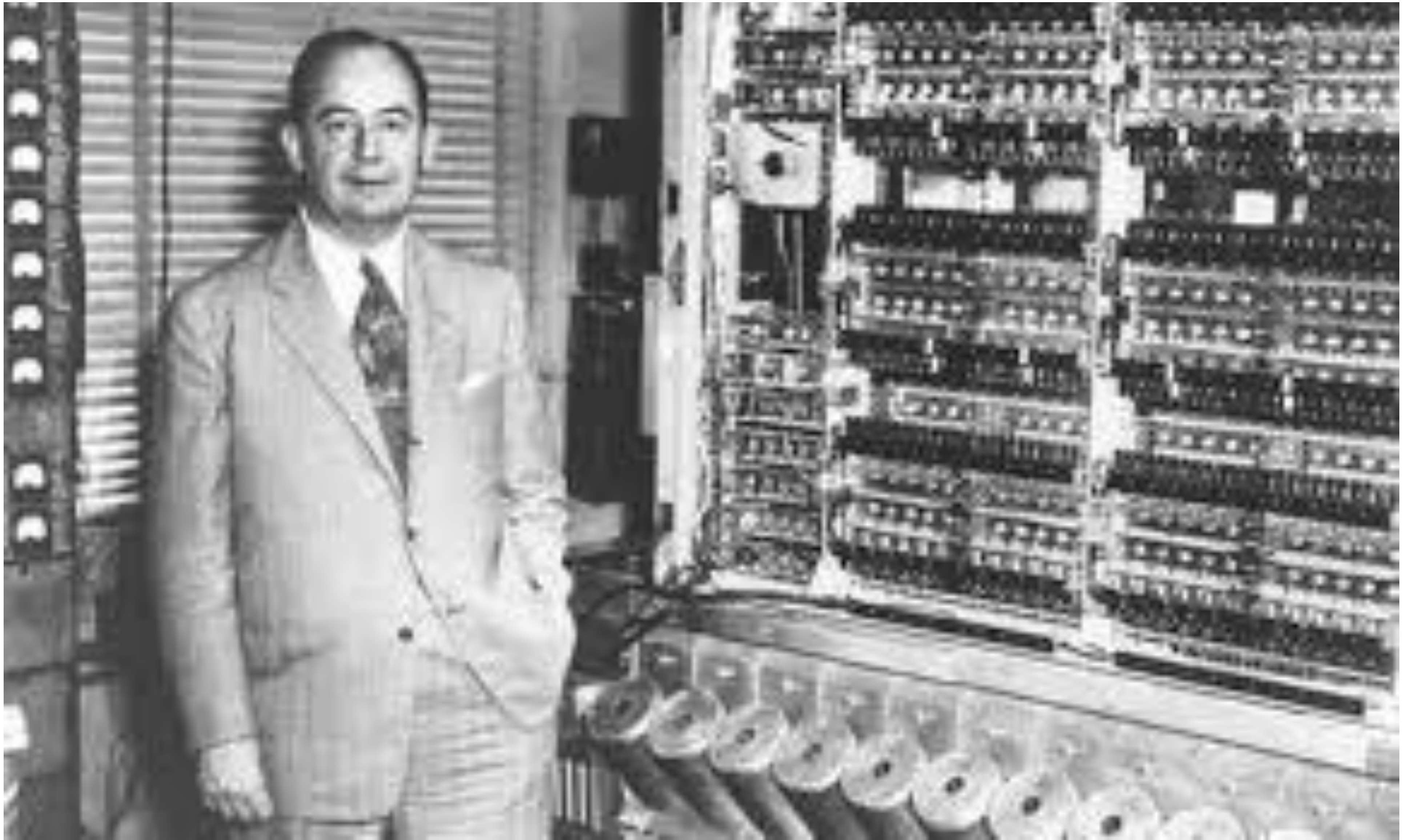
スクエアfreeセミナー
第40回
ITの進化をカイマミル

(株)インサイトテクノロジー
小幡一郎

フォン･ノイマン1946年ぐらい?の写真
後ろはMANIAC?




ゲームの理論と経済行動
J.フォン・ノイマン
O.モルゲンシュテルン
銀林浩・橋本和美・宮本敏雄 監訳
阿部修一・橋本和美 訳
I
ECONOMIC BEHAVIOR
THEORY of GAMES and
John von Neumann
Oskar Morgenstern
ちくま学芸文庫

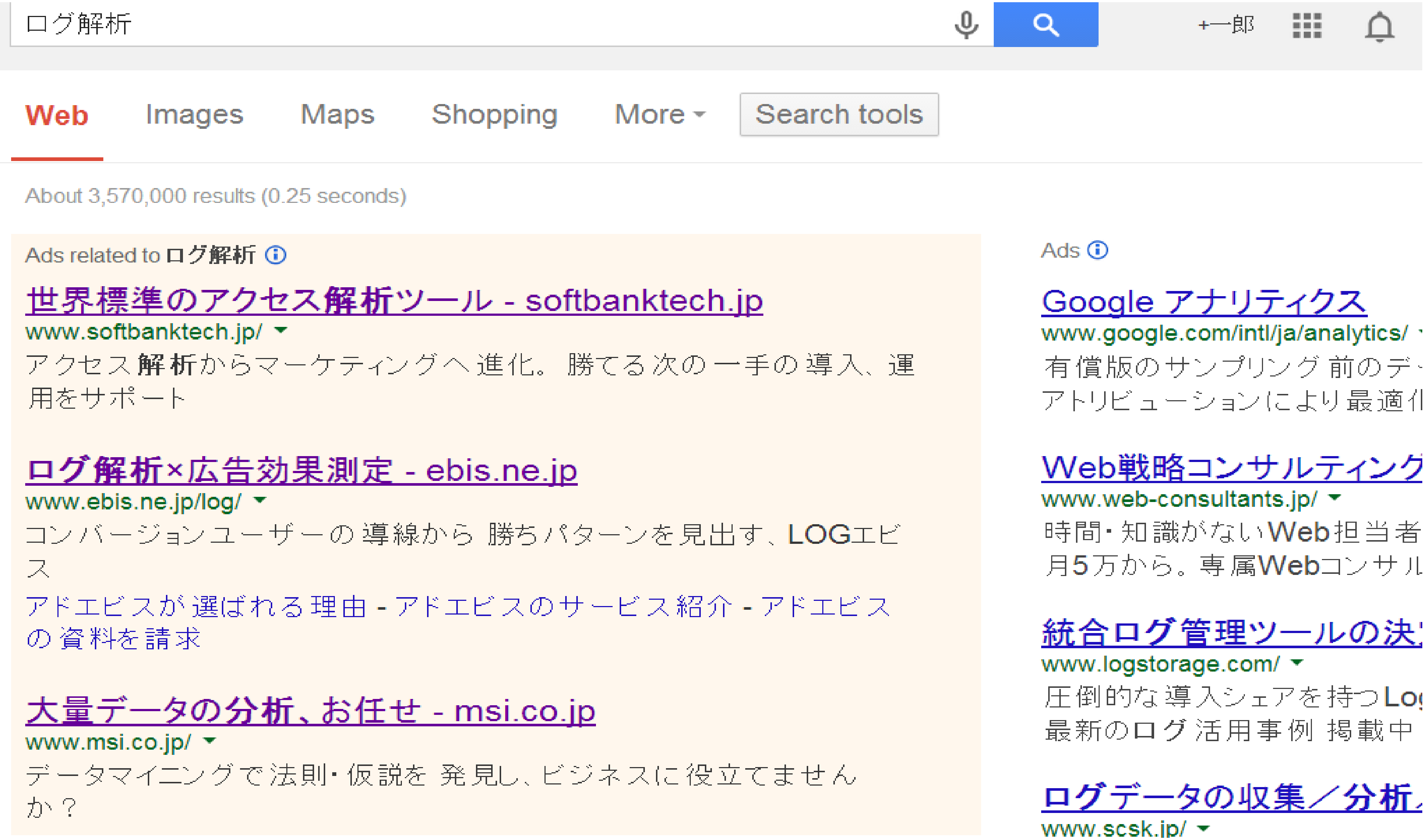
囚人のジレンマ

	囚人B 協調	囚人B 裏切り
囚人A 協調	（2年、2年）	（10年、0年）
囚人A 裏切り	（0年、10年）	（5年、5年）

ビッグデータ時代のジレンマ

	黙秘	導入中／検討中
提案C	OSS DB クラウド	データウェアハウス BI
提案A	ログ解析	ビッグデータ

ログ解析の世界


ログ解析
+一郎
Web
Images
Maps
Shopping
More
Search tools
About 3,570,000 results (0.25 seconds)
Ads related to ログ解析
世界標準のアクセス解析ツール - softbanktech.jp
www.softbanktech.jp/
アクセス解析からマーケティングへ進化。勝てる次の一手の導入、運用をサポート
ログ解析×広告効果測定 - ebis.ne.jp
www.ebis.ne.jp/log/
コンバージョンユーザーの導線から勝ちパターンを見出す、LOGエビス
アドエビスが選ばれる理由 - アドエビスのサービス紹介 - アドエビスの資料を請求
大量データの分析、お任せ - msi.co.jp
www.msi.co.jp/
データマイニングで法則・仮説を発見し、ビジネスに役立てませんか？
Ads
Google アナリティクス
www.google.com/intl/ja/analytics/
Web戦略コンサルティング
www.web-consultants.jp/
www.logstorage.com/
最新のログ活用事例 掲載中
www.scsk.jp/

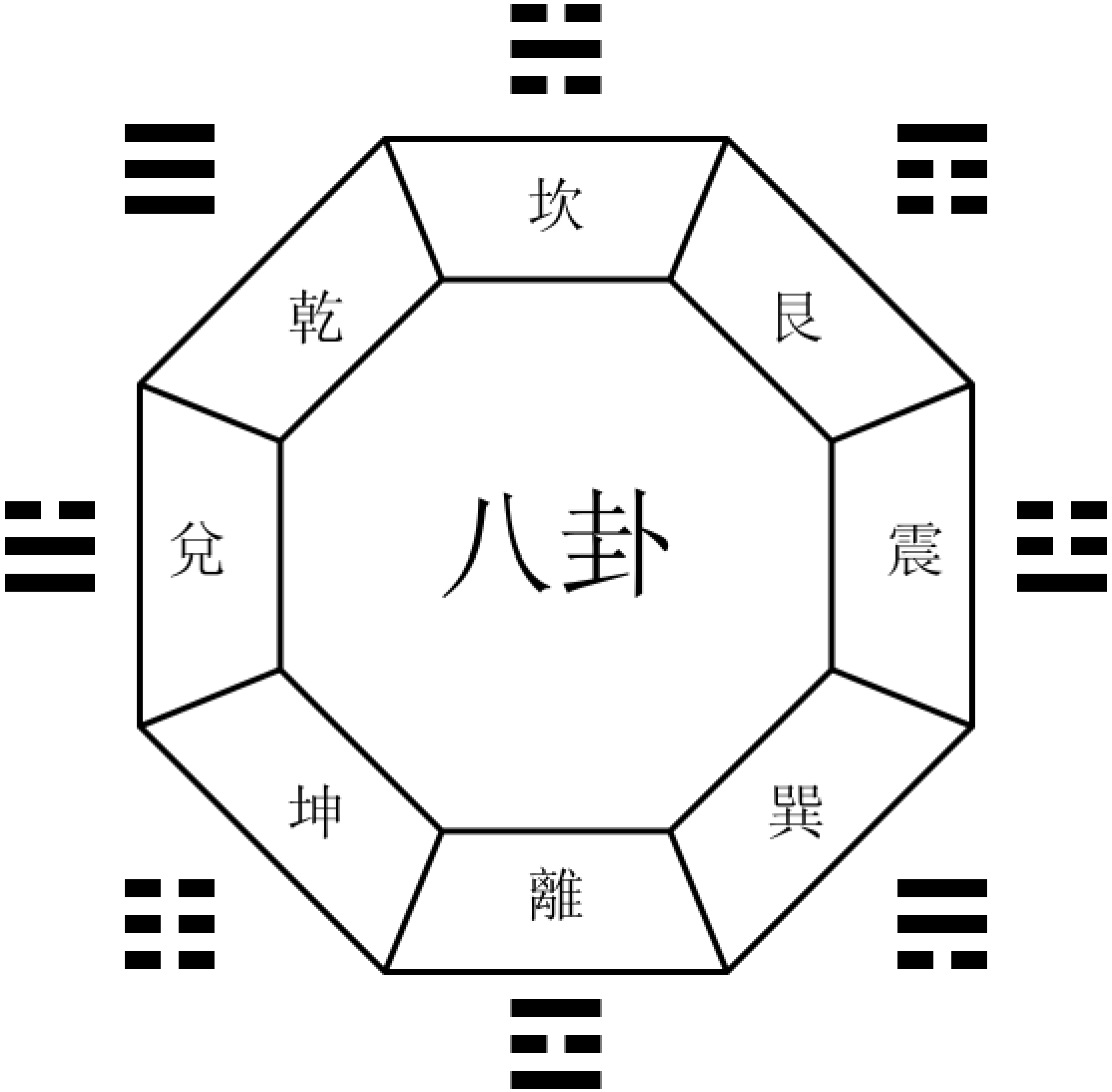
ゴットフリート・ヴィルヘルム・ライプニッツ（Gottfried Wilhelm Leibniz, 1646年7月1日

1.0e

Tabulae ita stabit

1	1	=	2^0
10	2		2^1
100	4		2^2
1000	8		2^3
10000	16		2^4
100000	32		2^5
1000000	64		2^6
10000000	128		2^7
100000000	256		2^8
1000000000	512		2^9
10000000000	1024		2^{10}
100000000000			

八卦（はっけ、はっか）は、古代中国から伝わる易における8つの基本図像



フォン・ノイマン

First Draft of a Report
on the EDVAC

by

John von Neumann

Contract No. W–670–ORD–4926

Between the

United States Army Ordnance Department

and the

University of Pennsylvania

Moore School of Electrical Engineering
University of Pennsylvania

June 30, 1945

1.0 DEFINITIONS

1.1 The considerations which follow deal with the structure of a *very high speed automatic di*
computing system, and in particular with its *logical control*. Before going into specific details, s
general explanatory remarks regarding these concepts may be appropriate.

1.2 An *automatic computing system* is a (usually highly composite) device, which can carry
instructions to perform calculations of a considerable order of complexity—e.g. to solve a non-li
partial differential equation in 2 or 3 independent variables numerically.

The instructions which govern this operation must be given to the device in absolutely exh
tive detail. They include all numerical information which is required to solve the problem u
consideration: Initial and boundary values of the dependent variables, values of fixed parame
(constants), tables of fixed functions which occur in the statement of the problem. These inst
tions must be given in some form which the device can sense: Punched into a system of punchc
or on teletype tape, magnetically impressed on steel tape or wire, photographically impresse
motion picture film, wired into one or more fixed or exchangeable plugboards—this list being b
means necessarily complete. All these procedures require the use of some code to express the log

ognize the most frequent malfunctions automatically, indicate their presence and location by ernally visible signs, and then stop. Under certain conditions it might even carry out the necessary rection automatically and continue (cf. {3.3}).

2.0 MAIN SUBDIVISIONS OF THE SYSTEM

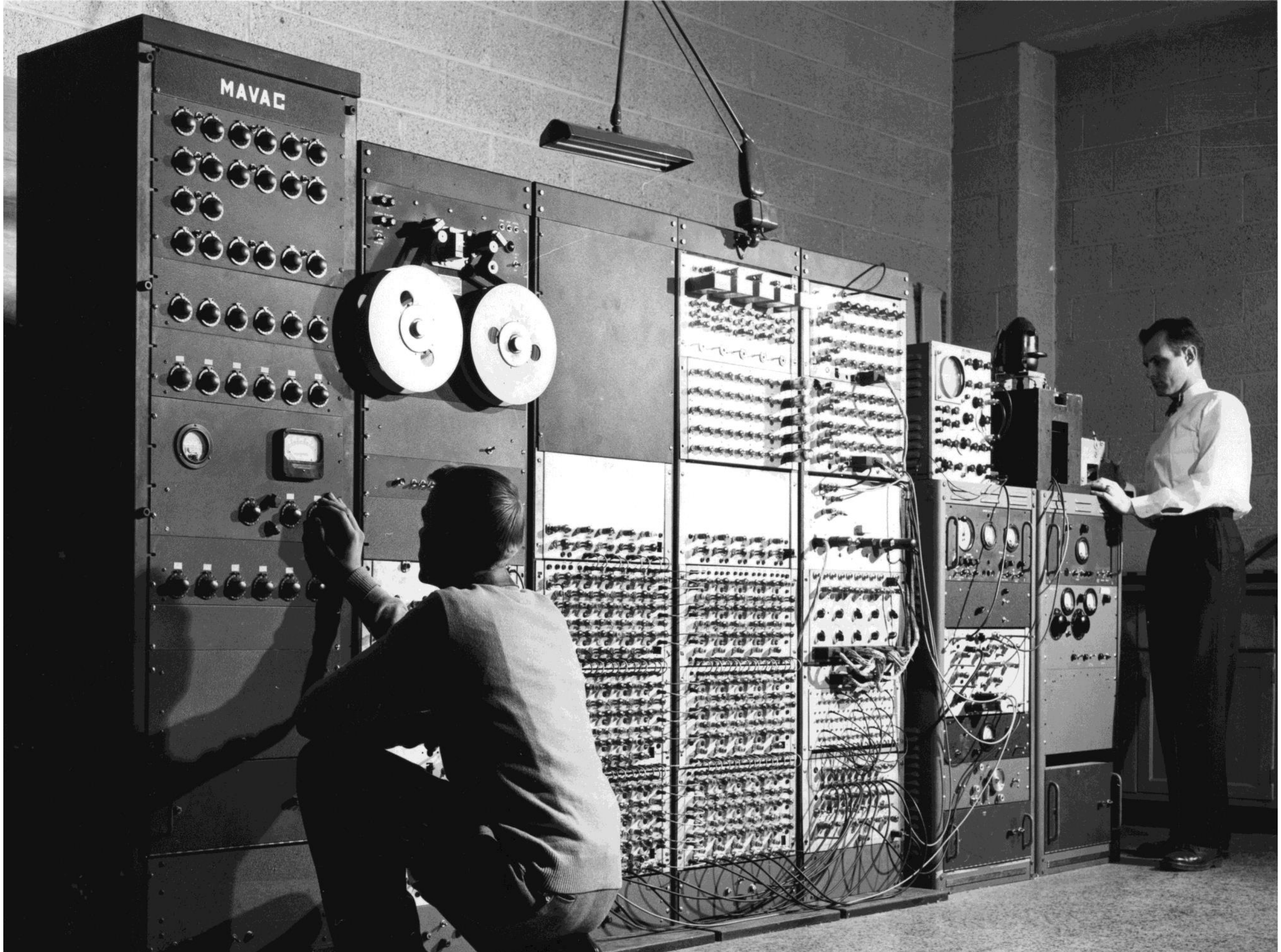
In analyzing the functioning of the contemplated device, certain classificatory distinctions sug-t themselves immediately.

First: Since the device is primarily a computer, it will have to perform the elementary opera-ns of arithmetics most frequently. These are addition, subtraction, multiplication and division: $-, \times, \div$. It is therefore reasonable that it should contain specialized organs for just these opera-ns.

It must be observed, however, that while this principle as such is probably sound, the specific y in which it is realized requires close scrutiny. Even the above list of operations: $+, -, \times, \div$, is beyond doubt. It may be extended to include such operation as $\sqrt{\ }$, $\sqrt[3]{\ }$, sgn, $|\ \ |$, also $\log_{10}$,

商用1号といわれているENIAC

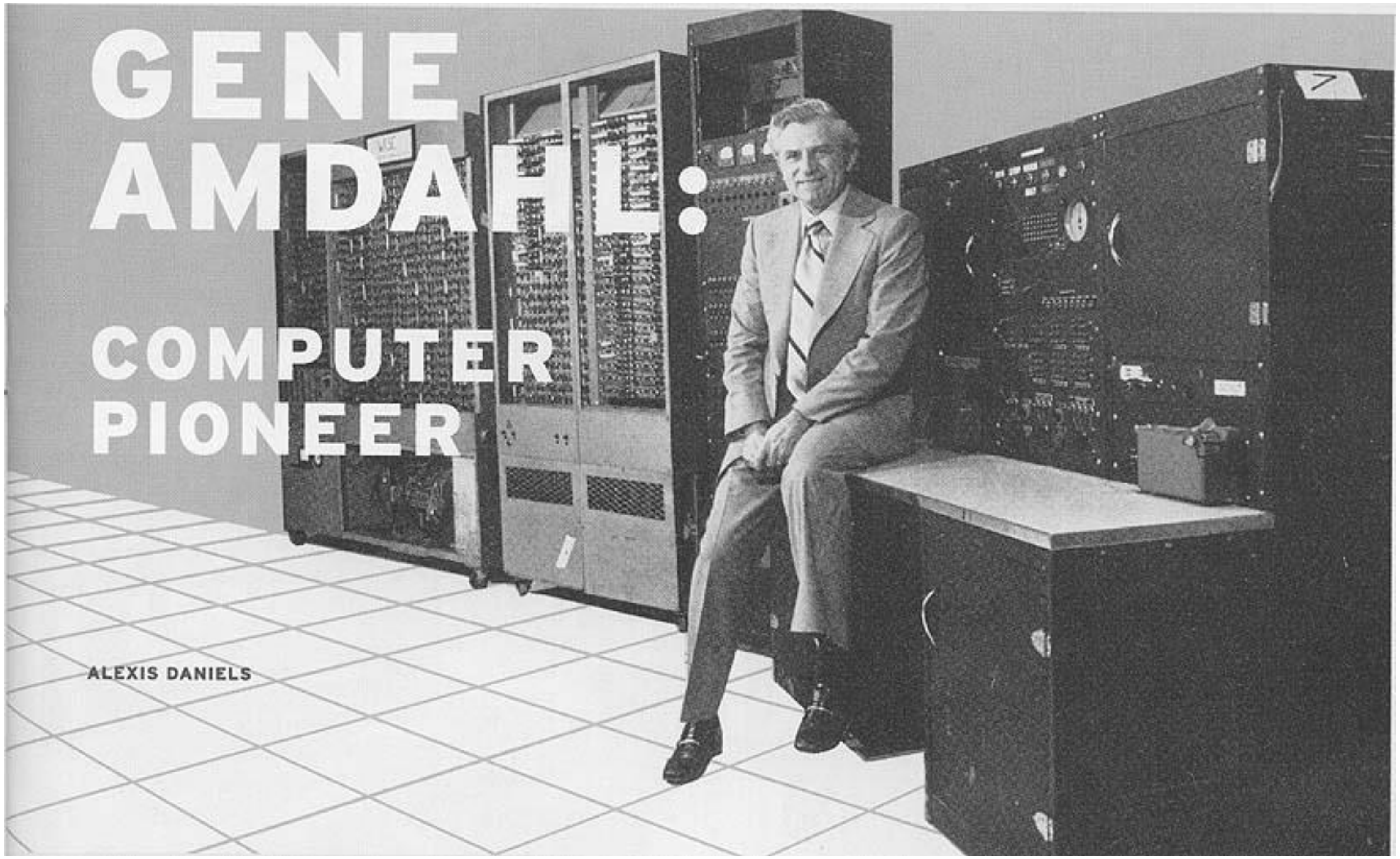
当時のカタログ?


MAVAC

IBM Stretch (1955開始)

Aggressive Uniprocessor Parallelism

Gene Amdahl, John Backus



1956年にAmdahlはIBMを離れ

1960年にIBMに戻りSystem/360シリーズに着手

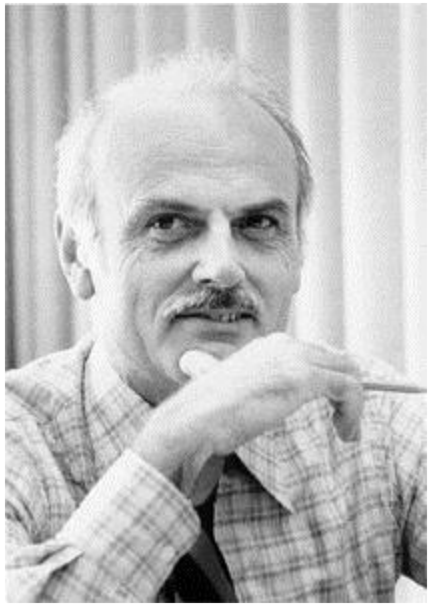
1970年に Amdahl Corporation設立 →コンパチ路線の開始

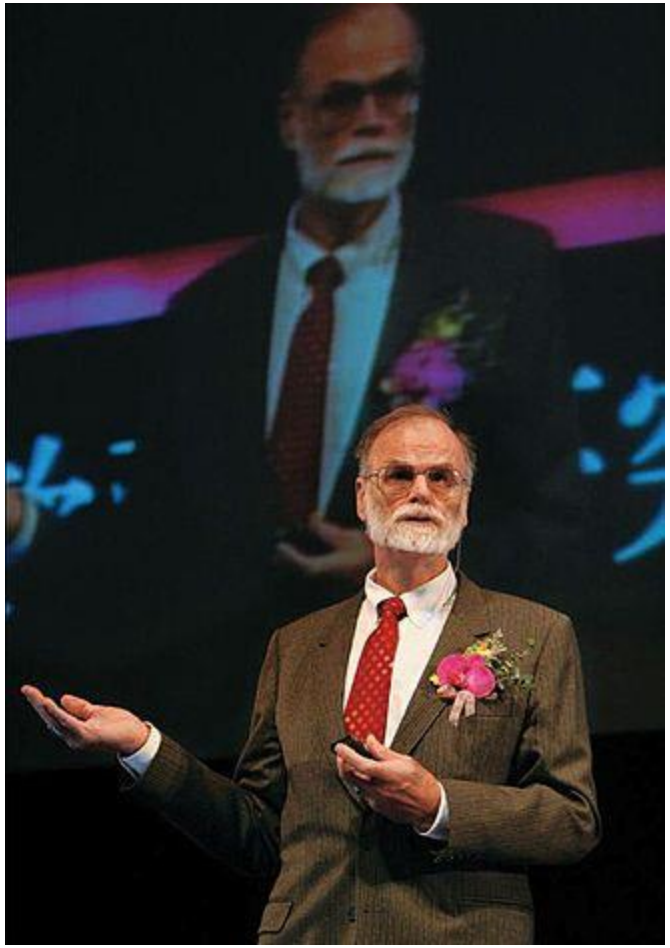
Sidebar: Stretch/7030 Customers

Machine name	Built	Customer	Delivery
X-1	Poughkeepsie	Los Alamos Scientific Lab (LASL)	1961
K-1	Kingston	Livermore Radiation Lab (LRL) [now LLNL]	1961
K-2	Kingston	Atomic Weapons Research Establishment (AWRE), Aldermaston, UK	1962
K-3	Kingston	US Weather Bureau [now NWS]	1962
K-4	Kingston	Naval Weapons Lab (Dahlgren)	1962
K-5	Kingston	MITRE Corporation	1962
K-6	Kingston	Commissariat a l'Energie Atomique (CEA), France	1963
7950 (Harvest)	Poughkeepsie	National Security Agency (NSA)	1962

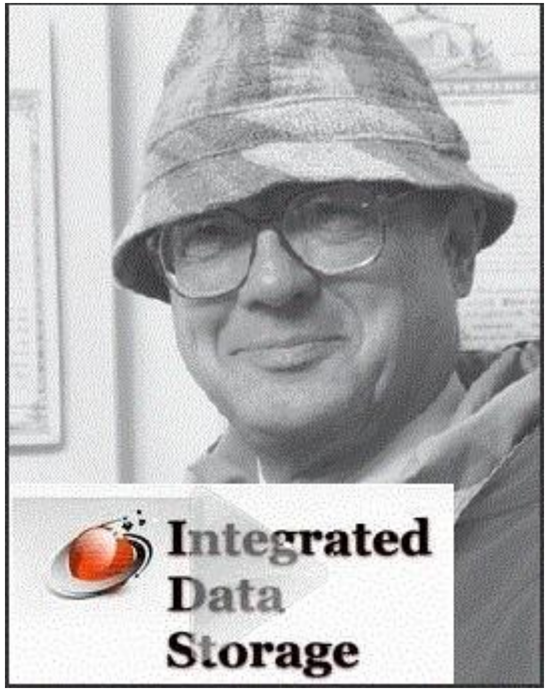
	1968　1970	1975　1980　1985	（1988）　1990～
発展段階	第1次オンライン ・元帳のセンター集中 （本支店オンライン）	第2次総合オンライン ・現金自動支払機 （銀行間オンライン）	第3次総合オンライン ・ファームバンキング （取引先オンライン）
開発の誘因	・大量の普通預金記帳業務	・全業務統合オンライン化 ・省力化と顧客サービス強化	・2次オンの物理的、構造的陳腐化 ・金融自由化への対応
システムの特徴	・預金、為替個別オンライン	・預金、為替、融資を統合した全科目オンライン ・総合口座	・自由化への対応（新業務／経営管理／外部ネット） ・インフラ改善

Charles Bachman(1924-)
Edgar F. Codd(1923 - 2003)
Jim Gray (1944 – 2007)





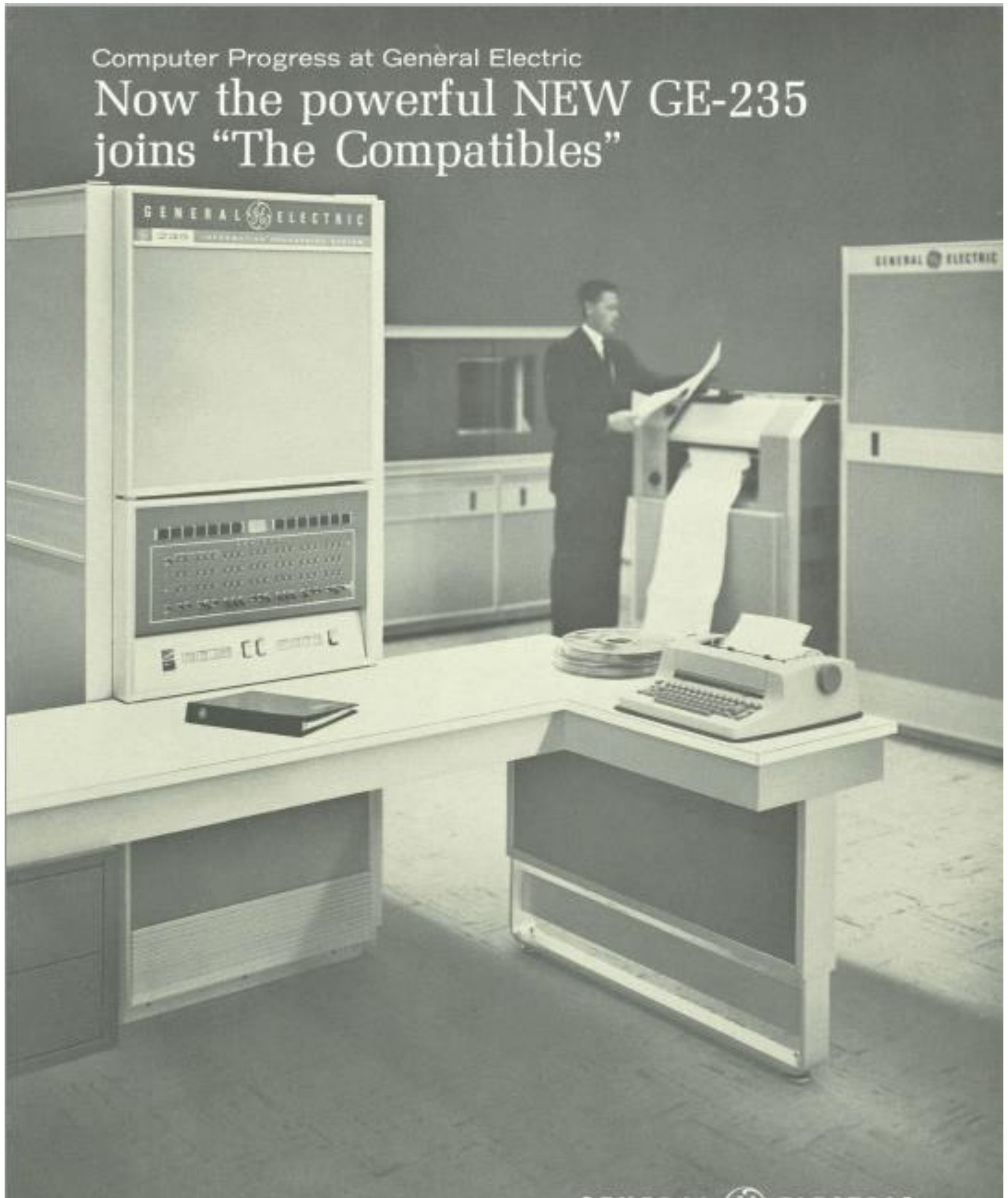
I-D-S ネットワーク型
Integrated Data Store



1964 GE 235computer発表

1968 IBM IMS発表

1960年代


Computer Progress at General Electric
Now the powerful NEW GE-235
joins “The Compatibles”
GENERAL ELECTRIC

1964 Data Structure Diagram

DATA STRUCTURE DIAGRAMS

By Charles W. Bachman

Successful communication of ideas has been and will continue to be a limiting factor in man's endeavors to survive and to better his life. The invention of algebra, essentially a graphic technique for communicating truths with respect to classes of arithmetic statements, broke the bond that slowed the development of mathematics.

Whereas "12+13=25" and "3+7=10" and "14+(-2)=12" are arithmetic statements, "a+b=c" is an algebraic statement. In particular, it is an algebraic statement controlling an entire class of arithmetic statements such as those listed.

Data Structure Diagrams

The Data Structure Diagram is also a graphic technique. It is based on a type of notation dealing with classes—specifically, with classes of entities and the classes of sets that relate them. For example, individual people and automobiles are entities. When they are taken collectively, they make two quite different classes of entities. On the other

entity grouping—one that associates a group of entities of one entity class with one entity of a different entity class in a subordinate relationship. The concepts of entity class and entity set are independent of each other and can be thought of as being at right angles or orthogonal. Figure 1 illustrates this point.

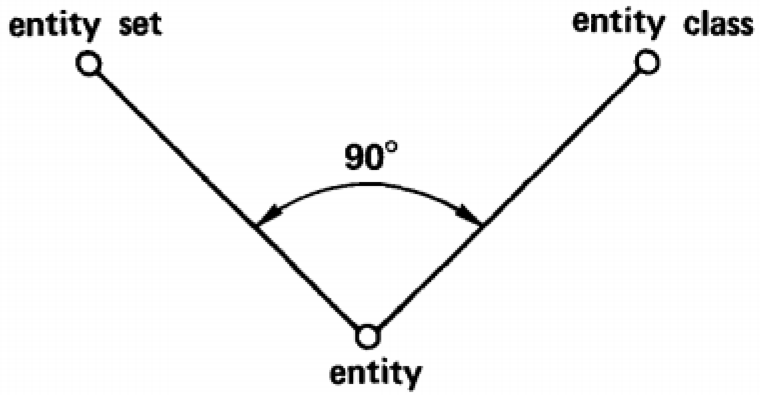
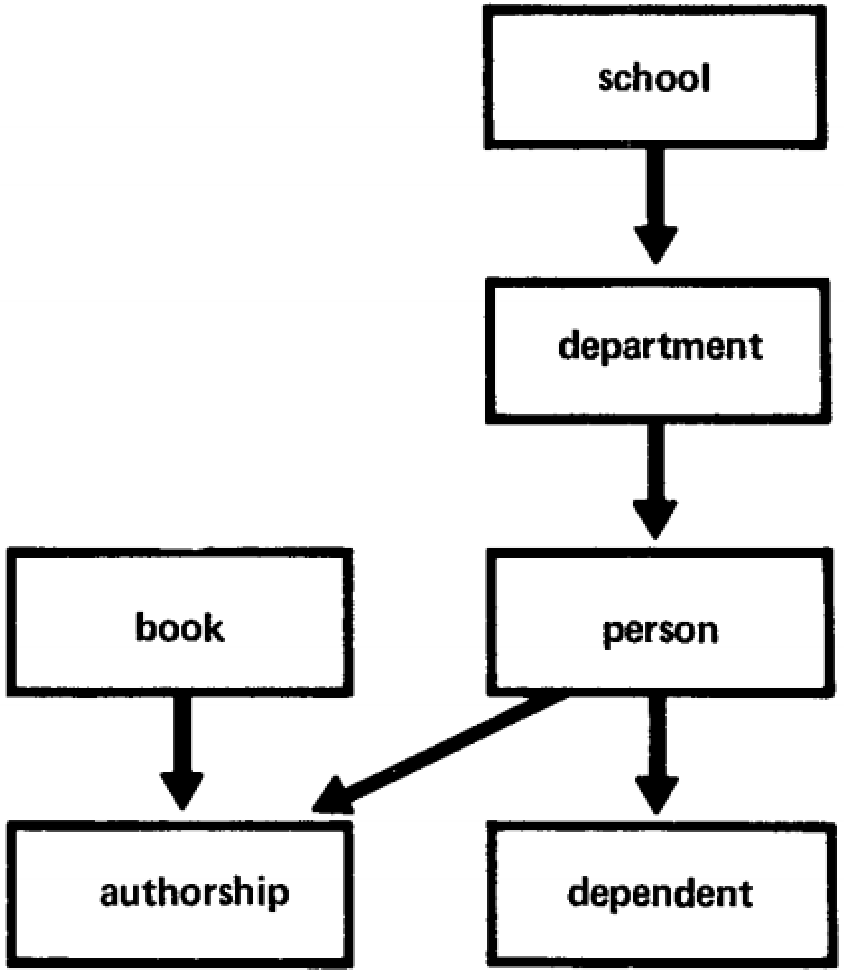


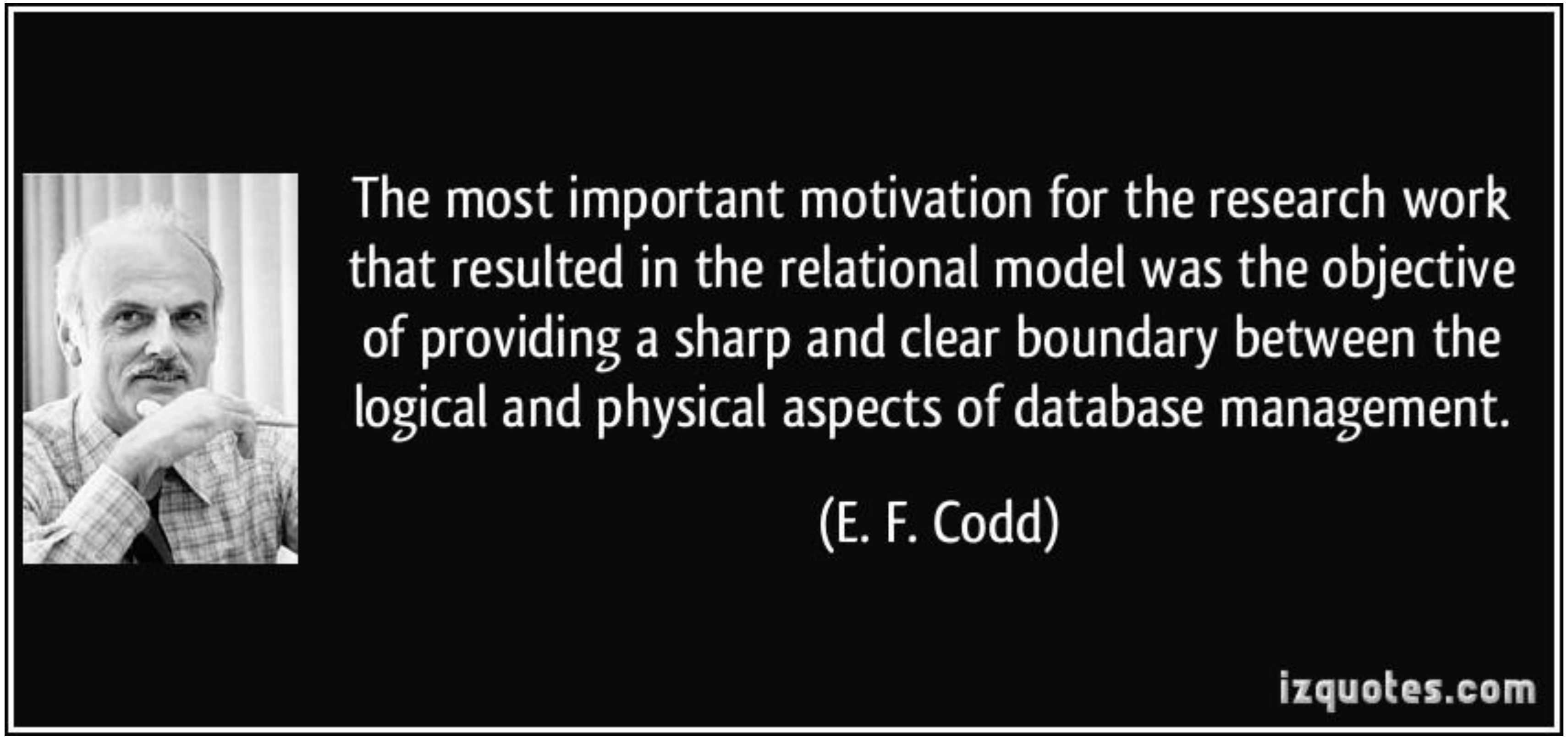
Figure 1

The term *set class* will be used in the text to mean an entire group of entity sets which are sufficiently similar, in terms of the attributes that describe them, to be considered collectively. Specifically, it is limited to those groups of sets in which the same entity-to-entity subordinate relationship exists. Figure 2 expands on Figure 1 to put all four of these

Data Structure Diagram



vs. 階層型


The most important motivation for the research work
that resulted in the relational model was the objective
of providing a sharp and clear boundary between the
logical and physical aspects of database management.
(E. F. Codd)
izquotes.com

1970 Relational DB

A Relational Model of Data for Large Shared Data Banks

E. F. Codd
IBM Research Laboratory, San Jose, California

Future users of large data banks must be protected from having to know how the data is organized in the machine (the internal representation). A prompting service which supplies such information is not a satisfactory solution. Activities of users at terminals and most application programs should remain unaffected when the internal representation of data is changed and even when some aspects of the external representation are changed. Changes in data representation will often be needed as a result of changes in query, update, and report traffic and natural growth in the types of stored information.

The relational view (or model) of data d
Section 1 appears to be superior in several resp
graph or network model [3, 4] presently in vog
inferential systems. It provides a means of desc
with its natural structure only—that is, witho
posing any additional structure for machine rep
purposes. Accordingly, it provides a basis for
data language which will yield maximal indepe
tween programs on the one hand and machine
tion and organization of data on the other.

A further advantage of the relational view
forms a sound basis for treating derivability, r
and consistency of relations—these are discusse
2. The network model, on the other hand, has
number of confusions, not the least of which i
the derivation of connections for the derivati
tions (see remarks in Section 2 on the "connect

Finally, the relational view permits a clearer
of the scope and logical limitations of present

1975 IBMサンノゼ研究所

1. Introduction

System R is an experimental database management system based on the relational model of data which has been under development at the IBM San Jose Research Laboratory since 1975 <1>. The software was developed as a research vehicle in relational database, and is not generally available outside the IBM Research Division.

This paper assumes familiarity with relational data model terminology as described in Codd <7> and Date <8>. The user interface in System R is the unified query, data definition, and manipulation language SQL <5>. Statements in SQL can be issued both from an on-line casual-user-oriented terminal interface and from programming languages such as PL/I and COBOL.

1976 初のRDBMSのプロトタイプ

System R: Relational Approach to Database Management

M. M. ASTRAHAN, M. W. BLASGEN, D. D. CHAMBERLIN,
K. P. ESWARAN, J. N. GRAY, P. P. GRIFFITHS,
W. F. KING, R. A. LORIE, P. R. MCJONES, J. W. MEHL,
G. R. PUTZOLU, I. L. TRAIGER, B. W. WADE, AND V. WATSON

IBM Research Laboratory

System R is a database management system which provides a high level relational data interface. The system provides a high level of data independence by isolating the end user as much as possible from underlying storage structures. The system permits definition of a variety of relational views on common underlying data. Data control features are provided, including authorization, integrity assertions, triggered transactions, a logging and recovery subsystem, and facilities for maintaining data consistency in a shared-update environment.

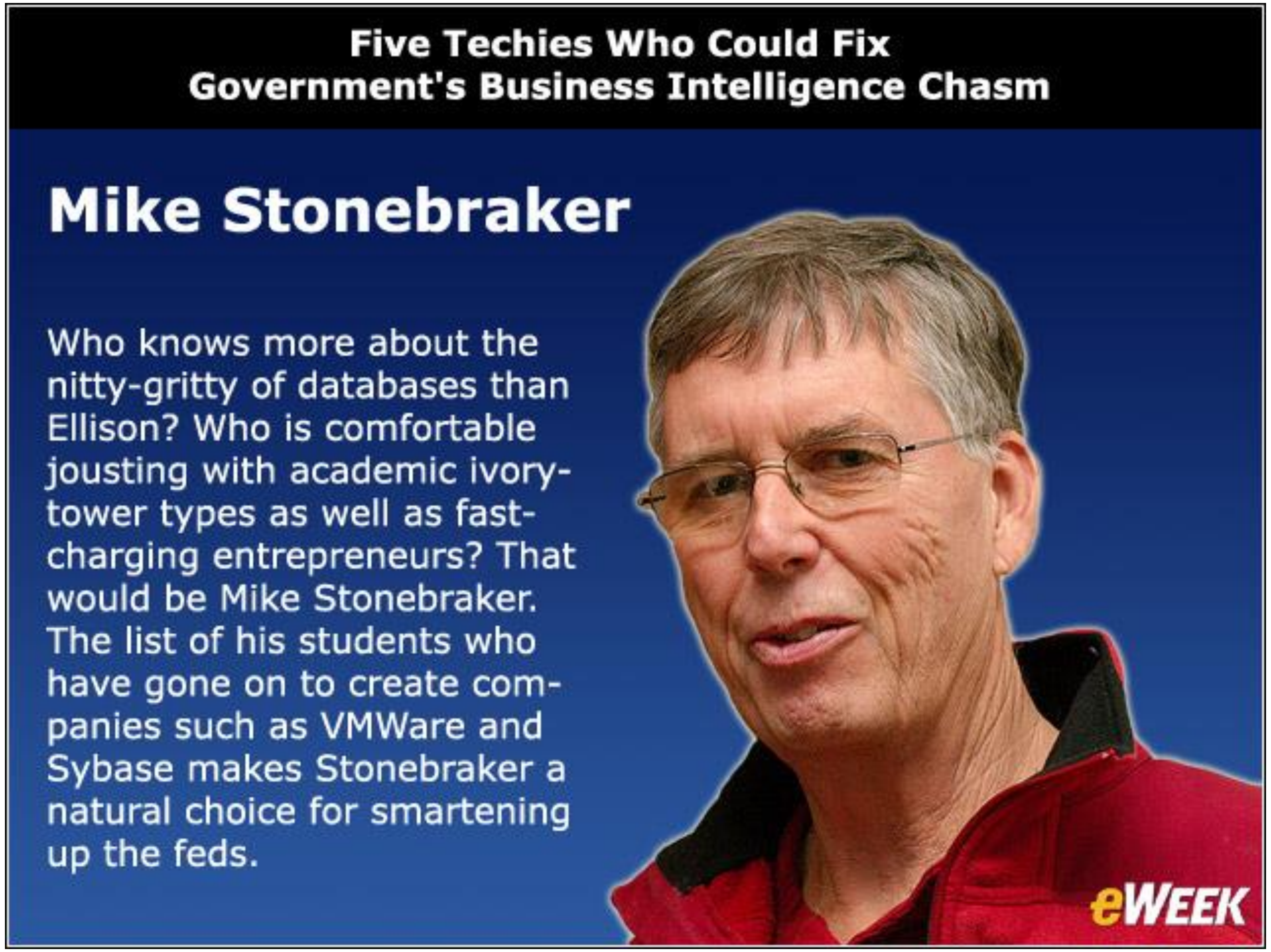
1977 Software Development Laboratories (SDL)

1979 **Relational Software, Inc.** (RSI)

Ed Oates, Bruce Scott, Bob Miner

http://www.businessinsider.com/whatever-happened-to-oracles-founders-in-this-iconic-photo-2012-8?op=1

Michael Ralph Stonebraker (1943-)



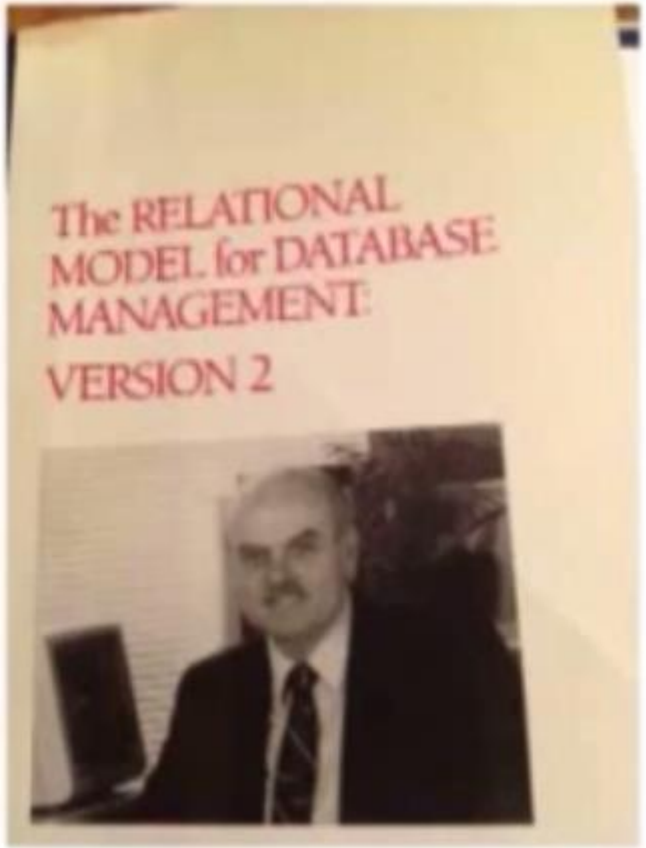
商用RDBMS

- 1976 System R
- 1976 Ingress カリフォルニア大バークレー校
- 1979 Oracle
- 1981 SQL/DS (IBM DOS/VSE版)
- 1983 DB2 （IBM MVS版)

www.youtube.com/watch?v=cVzgaqwN8E4

Codd's Relational Model – Version 2 (1990)

- He published 12 rules in Computerworld about "what makes a DBMS relational"
- According to these rules none of the products on the market qualified as relational DBMSs.
- In the book he prescribes 333 rules



54:50 / 1:08:28

Jim Gray (1944 – 2007)

DIRECT (1977-1984)
Gamma (1984-1992)
Paradise (1993-1997)

Parallel Database Systems

Why have parallel database systems become more than a research curiosity? One explanation is the widespread adoption of the relational data model. In 1983 relational database systems were just appearing in the marketplace; today they dominate it. Relational queries are ideally suited to parallel

multiple processors and memories, an operator can often be split into many independent operators each working on a part of the data. This partitioned data and execution gives *partitioned parallelism* (Figure 1).

The dataflow approach to database system design needs a mes-

workstation, and workgroup software. Those same client-server mechanisms are an excellent basis for distributed database technology.

Mainframe designers have found it difficult to build machines powerful enough to meet the CPU and I/O demands of relational data-

tation, and workgroup software. Those same client-server anisms are an excellent basis istributed database technol-

inframe designers have found icult to build machines powerough to meet the CPU and lemands of relational data-

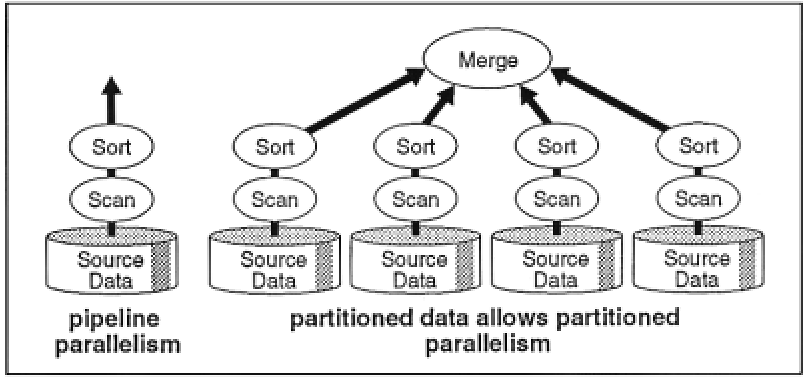


Figure 1.
The dataflow approach to relational operators gives both pipelined and partitioned parallelism. Relational data operators take relations (uniform sets of records) as input and produce relations as outputs. This allows them to be composed in dataflow graphs that allow *pipeline parallelism* (left) in which the computation of one operator proceeds in parallel with another, and *partitioned parallelism* in which operators (sort and scan in the diagram at the right) are replicated for each data source, and the replicas execute in parallel.

e 1.
taflow approach to relational ors gives both pipelined and oned parallelism. Relational data ors take relations (uniform sets ords) as input and produce relas outputs. This allows them to nposed in dataflow graphs that pipeline parallelism (left) in which mputation of one operator proin parallel with another, and pard parallelism in which operators nd scan in the diagram at the are replicated for each data , and the replicas execute in par-

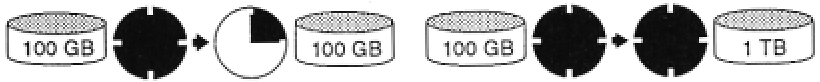
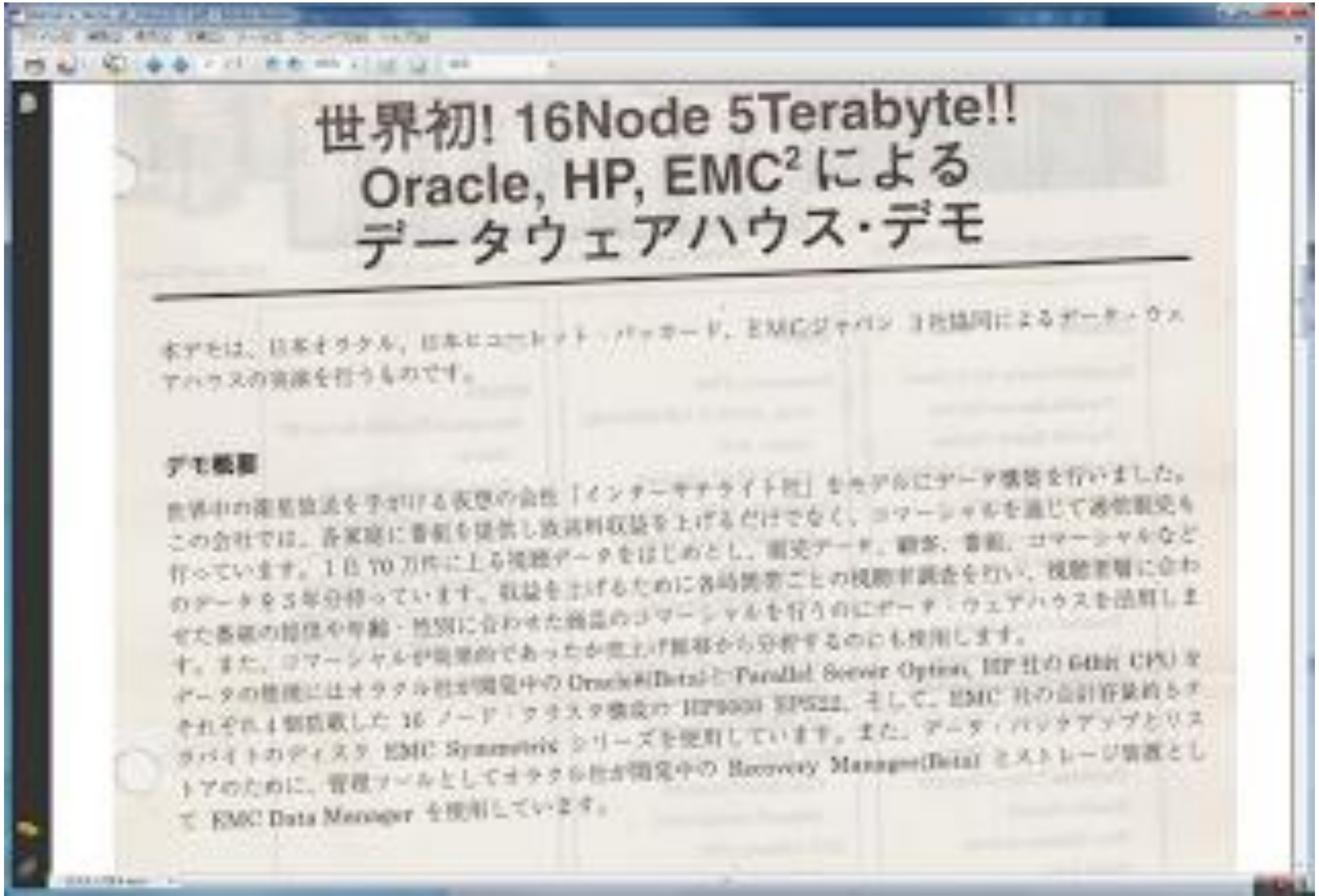


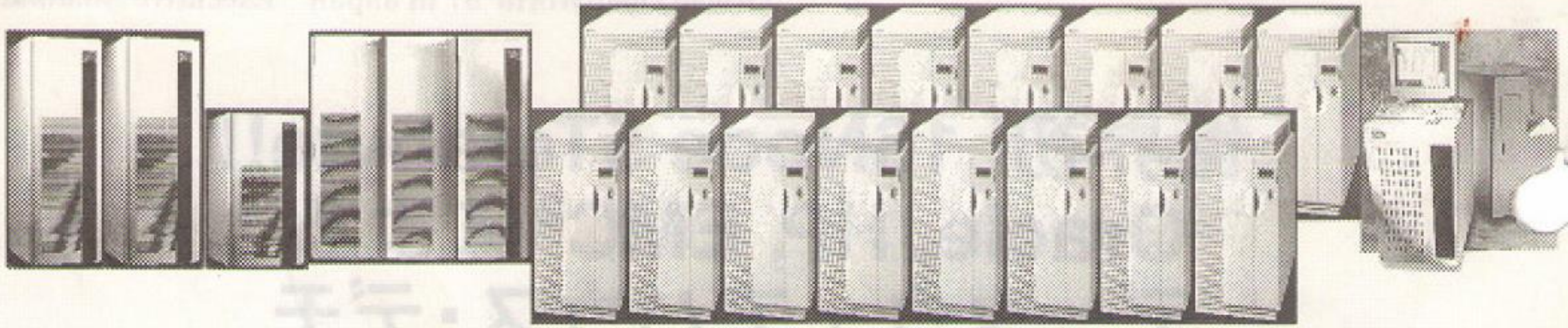
Figure 2.
Speedup and Scaleup. A speedup design performs a one-hour job four times faster when run on a four-times larger system. A scaleup design runs a ten-times bigger job is done in the

e 2.
up and Scaleup. A speedup deerforms a one-hour job four faster when run on a four-times system. A scaleup design runs a ten-times bigger job is done in the

In 1981, all T/16 CPUs were replaced by the **NonStop II**

1997年4月





EMC Symmetrix 3430,3430,3330,3700 HP9000 Enterprise Parallel Server 22 4way X 16nodes EMC Data Manager

Oracle	EMC	HP
Oracle8 Release 8.0.2 (Beta)	Symmetrix 3700	HP9000
- Parallel Server Option	- Disk: 23GB X 128 (2954GB)	Enterprise Parallel Server 22
- Parallel Query Option	- Cache: 4GB	- 16node
Enterprise Manager 1.3(Beta)	- 32 FW-SCSI Channel	- Fiber Channel Connect
- Recovery Manager	Symmetrix 3430 2Units	- Fiber Channel Switcher
	- Disk: 9GB X 96 (868GB)	- 4CPU (PA-8000 180MHz)
Test to Scale II	- Cache: 2GB	- 1GB(4node), 2GB(12node)
~16node の稼動試験	- 16 FW-SCSI Channel	- 2GB Disk X 2 on each node
- Parallel Direct Load	Symmetrix 3330	- HP-UX 10.20 with DART32
- Parallel Index Creation	- Disk: 9GB X 32 (289GB)	
- Parallel Update (Oracle8)	- Cache: 1GB	
- Parallel Create Table As Select	- 16 FW-SCSI Channel	
	EMC Data Manager E110	
- Partition Tables(Oracle8)	- Sun SparcCenter20	
- Parallel Query	- 10baseT configration	
- Star-Schema Access	ATL Library 4/52	
- Hash Join	- DLT4000 X 52 Tapes	


TOSHIBA

そして、

ビッグデータ分析時代となり、
RDBMSもひとつではなくなり、
使い分ける時代となった。。。
で、ここから、また始まるデータベースが面白い

次回(田中さん、次回はあるのかな?)、乞うご期待!