Software Architecture

Lecture 4-1.
Dataflow Systems and Batch sequential Architecture

数据流系统与批处理架构

Professor:
Yushan (Michael) Sun
Fall 2020

Lecture 2. Dataflow Systems

Content of the lecture

1. Concept of dataflow architecture
2. Concept of Batch sequential architecture
3. Example of program designed in batch sequential architecture

Concept of Dataflow Architecture

数据流软件体系结构

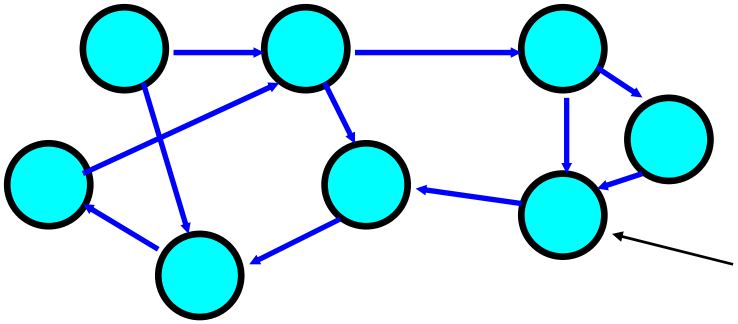
1. Concept of Dataflow Architecture

数据流系统定义

Definition: A dataflow system is a system, in which

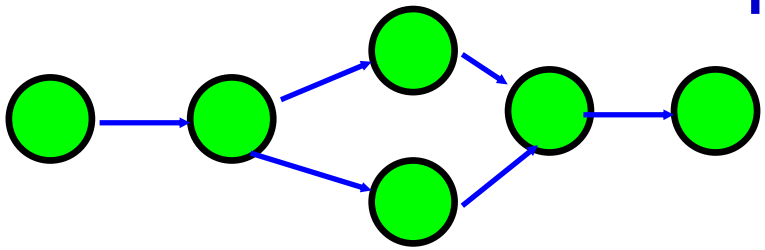
- **The availability of data controls the computation (有效、可用的数据控制计算)**
- **The structure of design is dominated by orderly motion of data from process to process（设计结构由从一个过程到另外的一个过程的数据的移动控制）**
- **The pattern of data is explicit （数据的模式是明确的）**

1. Concept of Dataflow Architecture

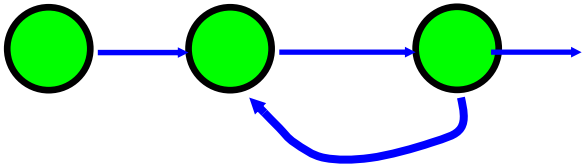


一般来说，数据可以以任意方式流动

当数据到达时，组件被激活，而无数据的时候，就休眠



我们对几乎线性的数据流系统感兴趣



或对比较简单，高度受限的循环结构感兴趣

General architecture of dataflow architecture

1. Concept of Dataflow Architecture

关于控制流（About Control flow)

- 我们关注程序中的控制流轨迹
- We pay attention to the locus of control flow that moves through the program
  - 数据可能伴随控制，但是数据不是主导 Data may accompany the control but is not dominant
  - 我们关心程序的执行顺序 The order of execution of the program is concerned

1. Concept of Dataflow Architecture

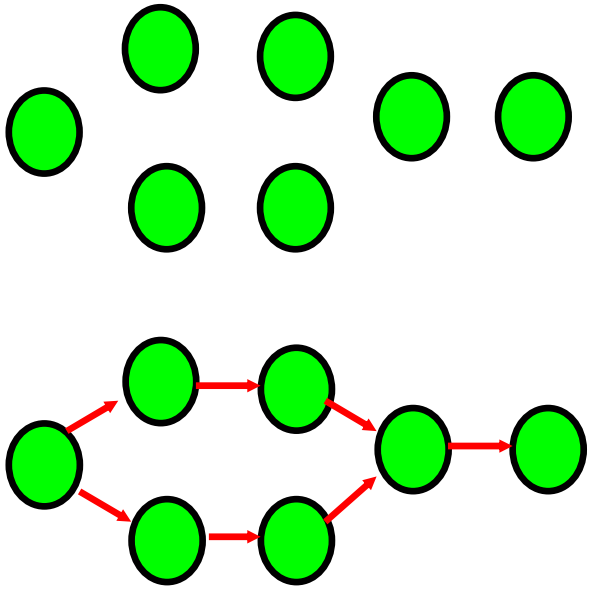
设计方面的注意事项

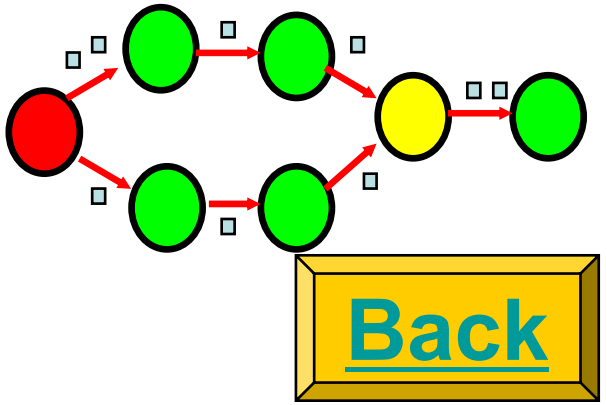
- 数据是怎样流动的. Main concern is how data moves through a collection of (atomic) computations
  - ➢随着设计流动，控制被激活 As data moves, control is activated
  - ➢设计的考虑事项 Reasoning is about
    - data availability (数据的可用性)是否一定会到达？没用就删除
    - transformation (数据变换) and
    - Latency (数据延迟)

1. Concept of Dataflow Architecture

数据流系统的开发方法论

- **将系统分解为一些模块 Decompose system into modules**
  - Abstraction isolates complexity
  - Operations performed on data objects
- **将模块组合成图 Compose modules into a graph**
  - Connect modules forming directed graph
  - Interfaces are negotiated to match data types
- **运行该图 Run the graph**
  - Data flows between the modules
  - scheduler automates parallelism
  - memory manager optimizes memory usage



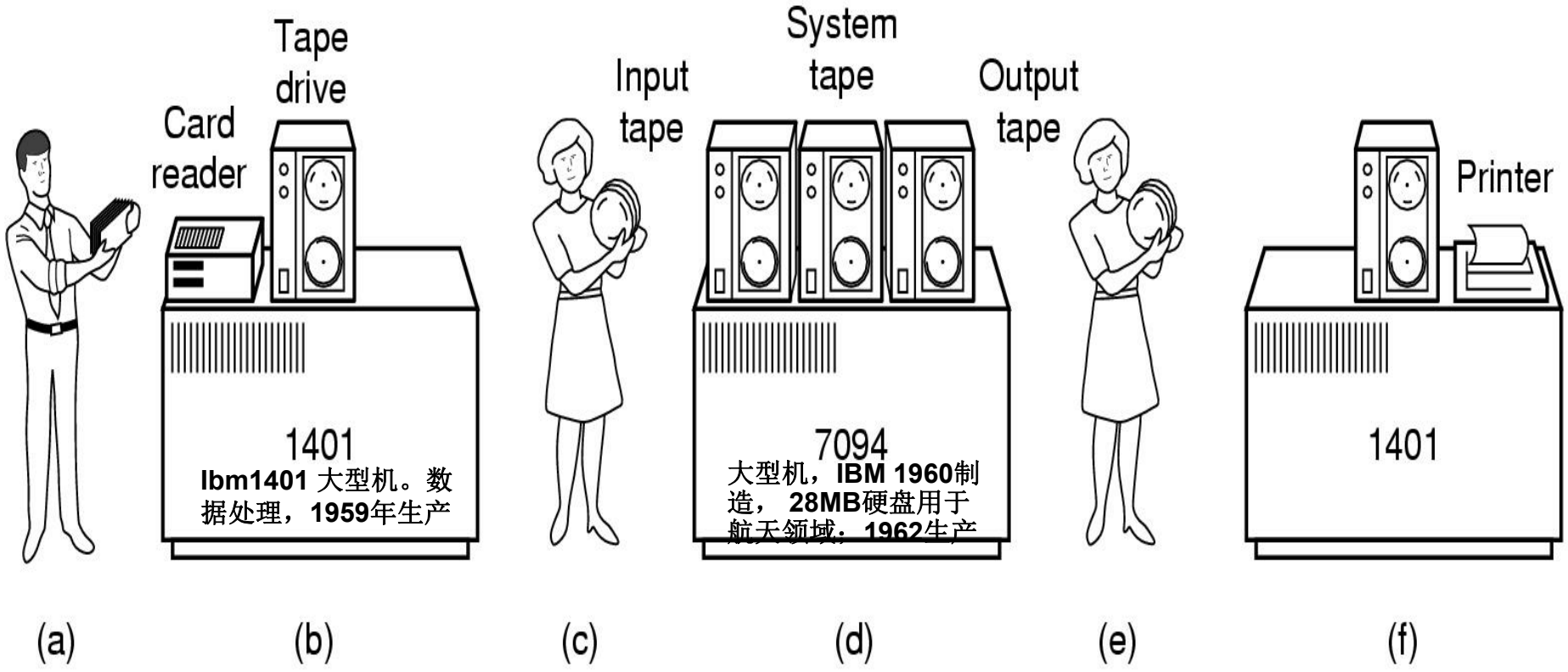


Concept of Batch Sequential Architecture

批处理软件体系结构的概念

2. Concept of Batch Sequential Architecture

1960年代-1970年代数据处理的例子



读取数据卡并制作数据磁带

处理磁带以获取感兴趣的数据

打印报告

2. Concept of Batch Sequential Architecture

批处理系统的定义. Definition of Batch sequential system：

a) Batch sequential process is formed by a sequence of processing steps, intermediating by some kind of storage, such as megatapes, hard disk, etc.

b) Each step will do some operation on the input storage to get some useful information or modify the contents of the source storage and then save the resultant data to its sink storage.

Batch sequential pattern

2. Concept of Batch Sequential Architecture

批处理系统的性质

Properties of a Batch sequential system:

a) 处理过程是互相独立的. Processing steps are independent programs （no interactions between any two processing steps)

Proc1，Proc2和Proc3互不调用

2. Concept of Batch Sequential Architecture

b) 每一步要彻底完成才能开始下一步骤
Each step run to completion before the next step starts

- Proc1完整地处理完D1中的数据➔将数据存入D2
- Proc2完整地处理完D2中的数据➔将数据存入D3
- Proc3完整地处理完D3中的数据➔将数据存入D4

c) Data is transmitted as a whole between steps （理解为一种特殊的数据流；数据以块状的形态流动）

2. Concept of Batch Sequential Architecture

- **典型应用(Typical applications):**
  - classical data processing

- **早年的业务(商业)数据库处理方式**
- **Business database processing** historically is dominated by database updates:
  - Discrete transaction of predetermined type;
  - periodic reports;
  - special handling of bad requests.

2. Concept of Batch Sequential Architecture

- **Historical database: batch sequential.**

How to work?

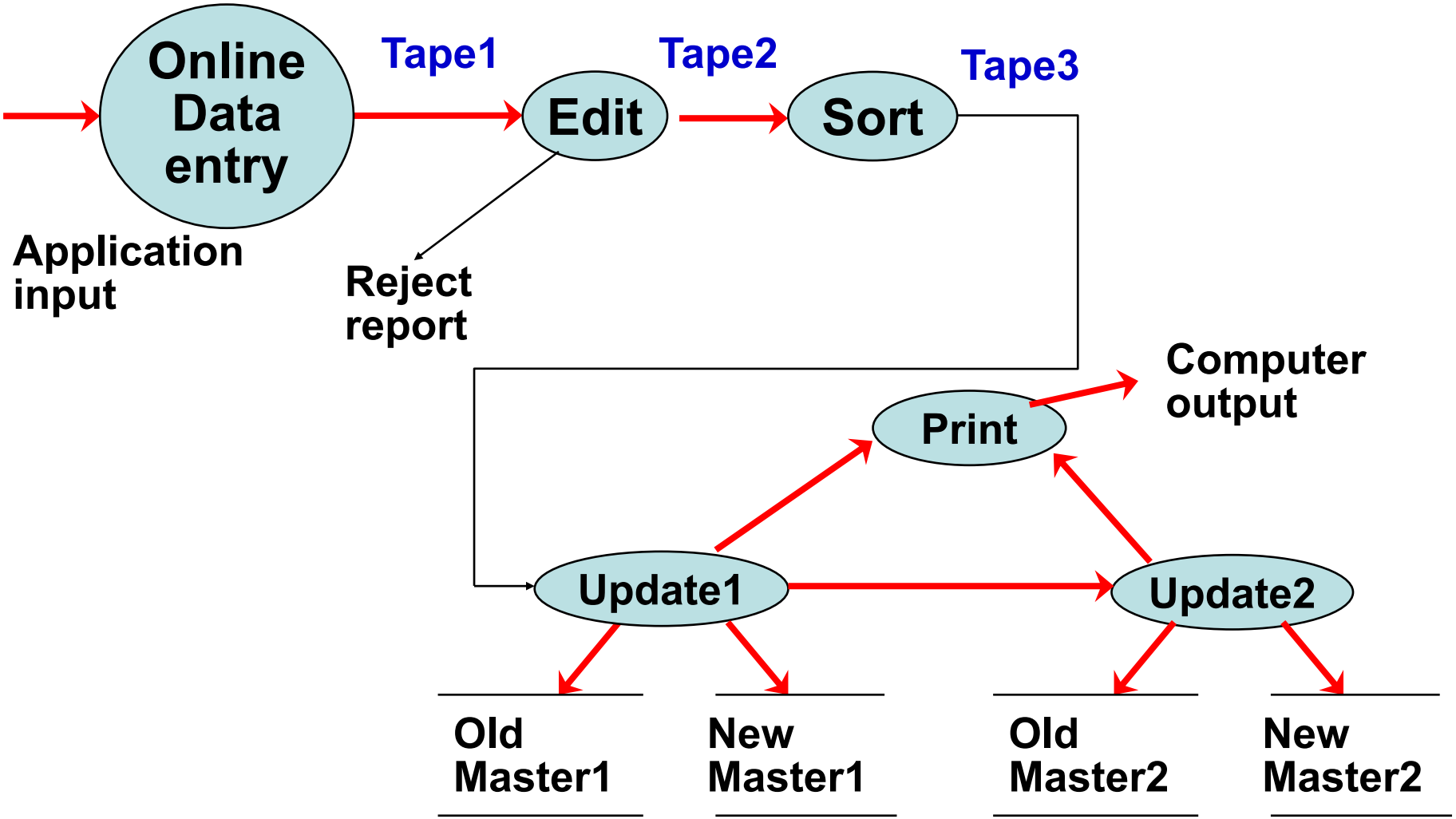
- **Mainframes and magnetic tape**
- **Manual block scheduling plan for each step**

关系型数据库出现在1980年代

IBM 7094 mainframe

2. Concept of Batch Sequential Architecture

【例 1】 批处理数据库更新程序



Dataflow diagram for Batch Databases

2. Concept of Batch Sequential Architecture

1. Raw data

Sun, Mike	356821	TV	888.56	324-56-7348
Liu, Elle	672452	356	346-12-6788	camera
Lee, Mary	842674	radio	65.44	563-27-8977
Juell, Paul	257215	Telephone	97.05	678-23-1567
Nygard, Ken	573245	Microwave oven	56.88	974-67-8577
Magel, Ken	834265	refrigerator	988	874-26-7847
Hunt, Paul	332184	Bicycle	523	223-56-6878
Eric, Vosika	257652	jewellery	1005	564-31-0976

2. Concept of Batch Sequential Architecture

2. Edit data

Sun, Mike	356821	TV	888.56	324-56-7348
Lee, Mary	842674	radio	65.44	563-27-8977
Juell, Paul	257215	Telephone	97.05	678-23-1567
Nygard, Ken	573245	Microwave oven	56.88	974-67-8577
Magel, Ken	834265	refrigerator	988	874-26-7847
Hunt, Paul	332184	Bicycle	523	223-56-6878
Eric, Vosika	257652	jewellery	1005	564-31-0976

Ejected data, wrong format

Liu, Elle	672452	356	346-12-6788	camera

2. Concept of Batch Sequential Architecture

3. Sort data

Eric, Vosika	257652	jewellery	1005	564310976
Hunt, Paul	332184	Bicycle	523	223566878
Juell, Paul	257215	Telephone	97.05	678231567
Lee, Mary	842674	radio	65.44	563278977
Magel, Ken	834265	refrigerator	988	874267847
Nygard, Ken	573245	Microwave oven	56.88	974678577
Sun, Mike	356821	TV	888.56	324567348

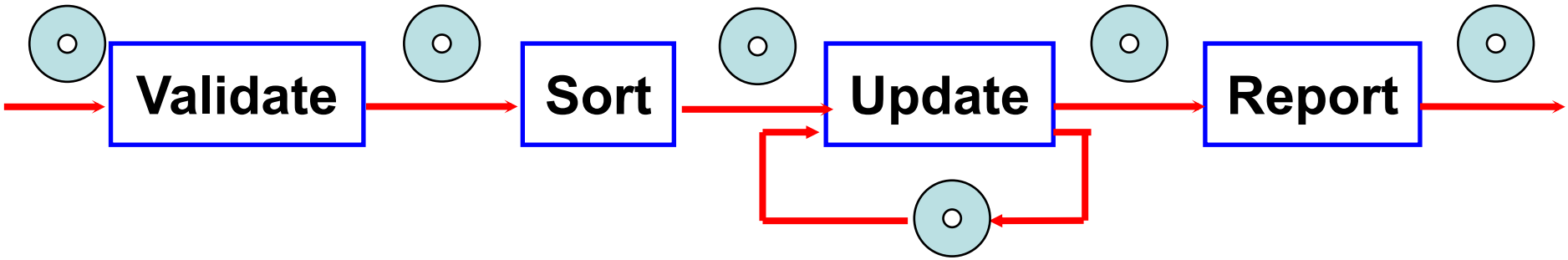
2. Concept of Batch Sequential Architecture

4. Insert data into master database

5. Update database

2. Concept of Batch Sequential Architecture

In the above example, the batch sequential diagram can be simply redrawn as the following.



Batch Sequential Database Architecture

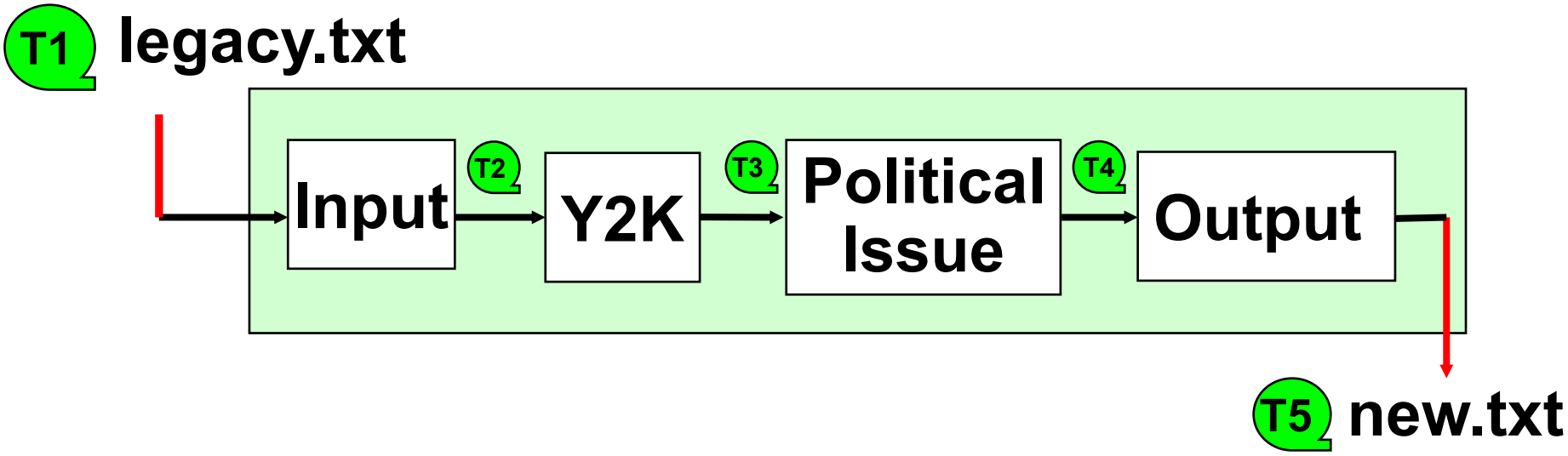
Example of program designed in batch sequential architecture

批处理架构的例子

3. Example of design in Batch Sequential Architecture

【例2】千年虫与政治问题 (Year 2000 problem and political issues.)

- In a text file, change all year expressions “xx” into “xxxx”, e.g., “89” into “1989”, and change all expressions “Republic of China” into “Taiwan”



Legacy file update using batch sequential software architecture

3. Example of design in Batch Sequential Architecture

Original txt file:

Television $3800 Japan 95-12-11
Television $3600 Germany 88-12-01
Microwave $3800 Japan 95-12-11
Microwave $3600 Germany 88-12-01
Television $3800 Republic of China 95-12-11
Television $3300 Germany 89-02-01
Microwave $3800 Republic of China 95-12-11
Microwave $3300 Germany 89-02-01
Washing Machine $3800 Japan 95-12-11
Washing Machine $3600 Germany 88-12-01
Television $3800 Republic of China 95-12-11
Television $3300 Germany 89-02-01
Washing Machine $3800 Republic of China 95-12-11
Washing Machine $3300 Germany 89-02-01

3. Example of design in Batch Sequential Architecture

After Y2K (in Tape 3):

Television $3800 Japan 1995-12-11
Television $3600 Germany 1988-12-01
Microwave $3800 Japan 1995-12-11
Microwave $3600 Germany 1988-12-01
Television $3800 Republic of China 1995-12-11
Television $3300 Germany 1989-02-01
Microwave $3800 Republic of China 1995-12-11
Microwave $3300 Germany 1989-02-01
Washing Machine $3800 Japan 1995-12-11
Washing Machine $3600 Germany 1988-12-01
Television $3800 Republic of China 1995-12-11
Television $3300 Germany 1989-02-01
Washing Machine $3800 Republic of China 1995-12-11
Washing Machine $3300 Germany 1989-02-01

3. Example of design in Batch Sequential Architecture

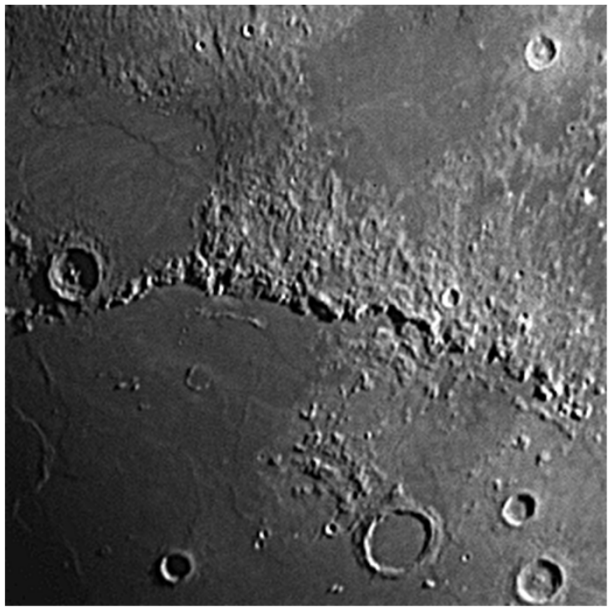
After Political Issue (in Tape 4) :

Television $3800 Japan 1995-12-11
Television $3600 Germany 1988-12-01
Microwave $3800 Japan 1995-12-11
Microwave $3600 Germany 1988-12-01
Television $3800 Taiwan 1995-12-11
Television $3300 Germany 1989-02-01
Microwave $3800 Taiwan 1995-12-11
Microwave $3300 Germany 1989-02-01
Washing Machine $3800 Japan 1995-12-11
Washing Machine $3600 Germany 1988-12-01
Television $3800 Taiwan 1995-12-11
Television $3300 Germany 1989-02-01
Washing Machine $3800 Taiwan 1995-12-11
Washing Machine $3300 Germany 1989-02-01

3. Example of design in Batch Sequential Architecture

【例3】批处理架构在图像处理中的应用

- Batch sequential architecture can be used to do image processing

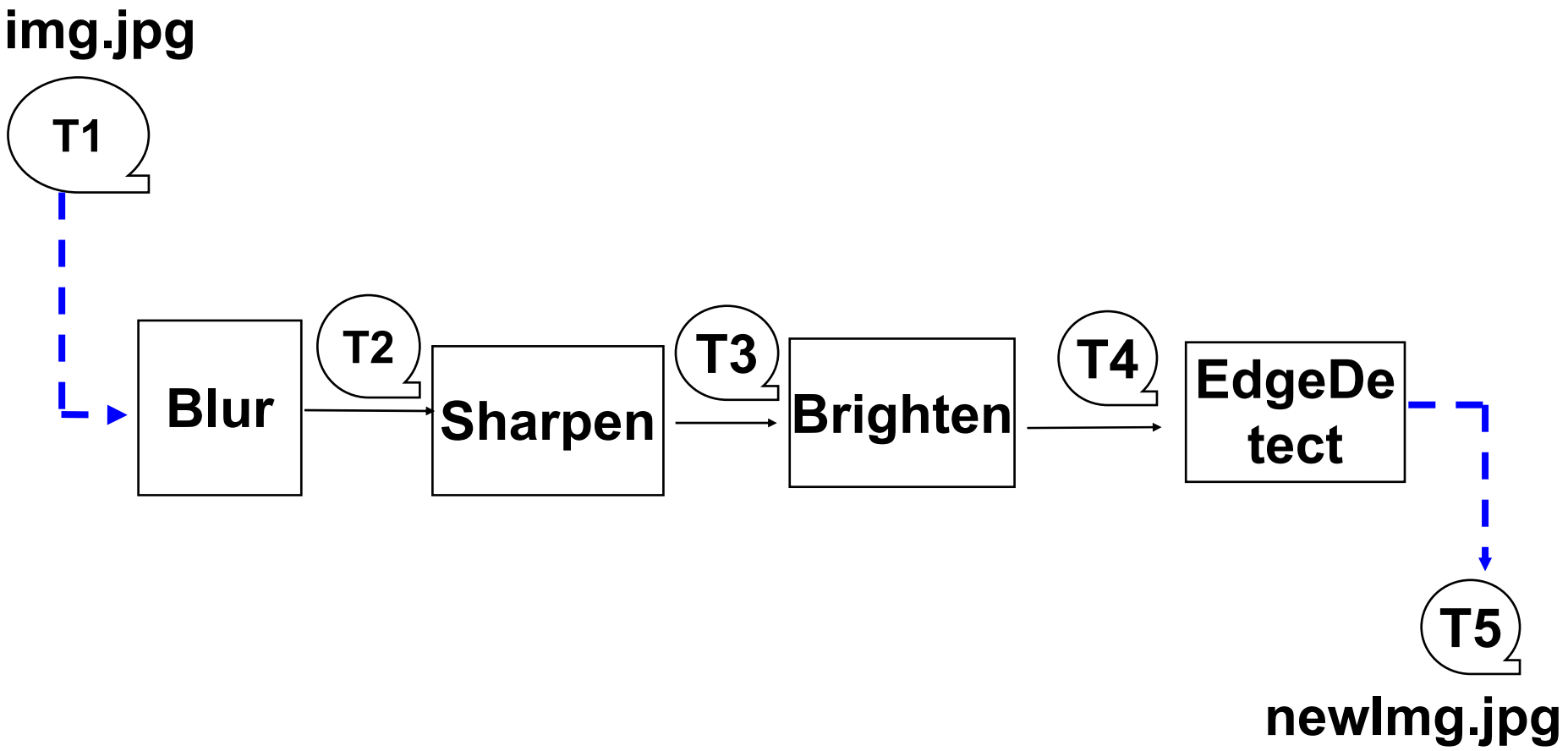


Picture before and after image sharpening

3. Example of design in Batch Sequential Architecture

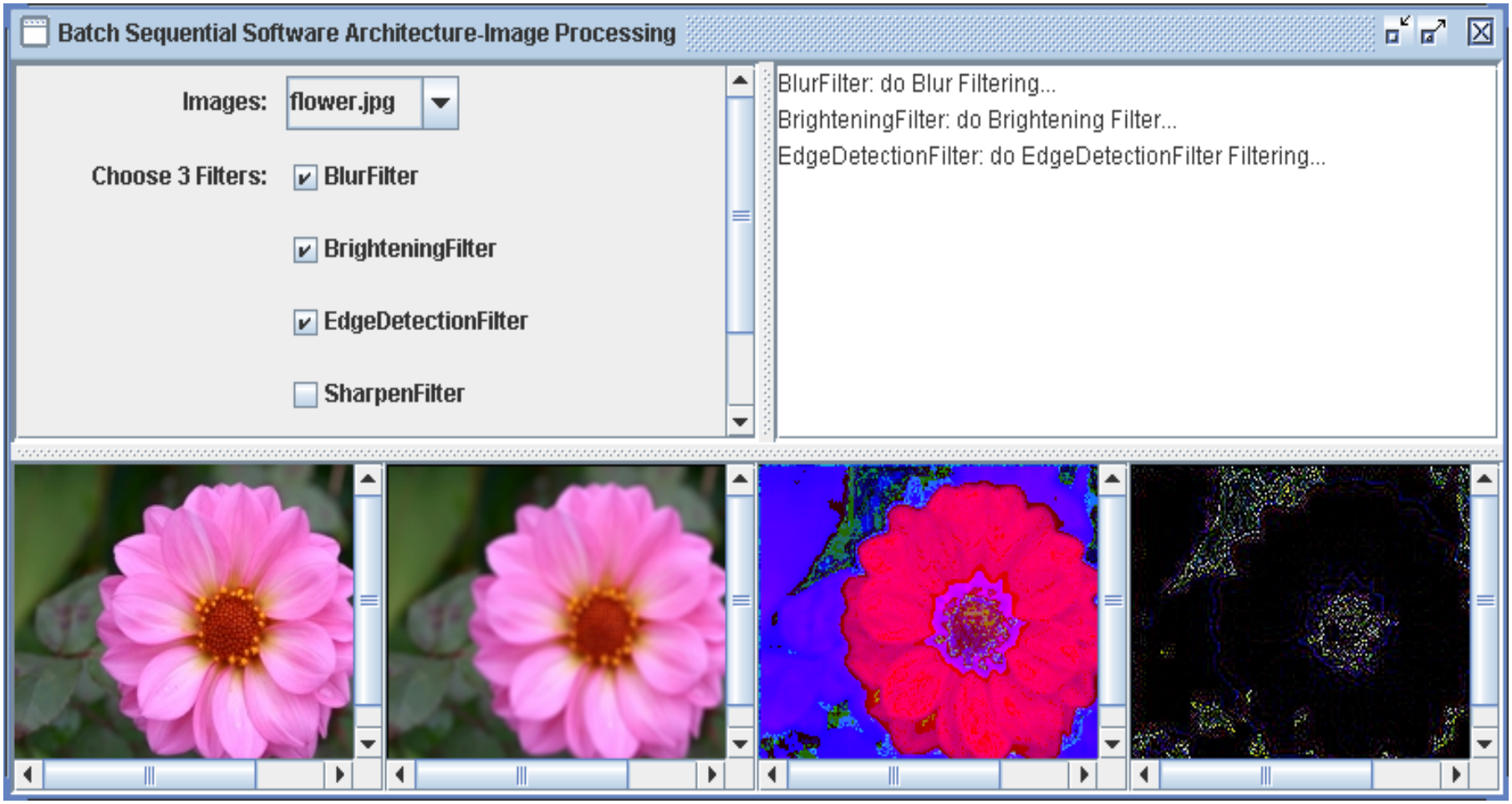
- 设计考虑设计一个图像处理软件。该软件包含一些可以随时添加的过滤器（filters），例如
  - Blurring (图像模糊)，
  - Sharpening (图像锐化)
  - Brightening (图像变亮)
  - EdgeDetector (发现图像边界)，
- 一些过滤器被串联在一起，以便完成一些比较复杂的功能。

3. Example of design in Batch Sequential Architecture



顺序批处理体系结构的数字图像处理系统对象逻辑图

3. Example of design in Batch Sequential Architecture

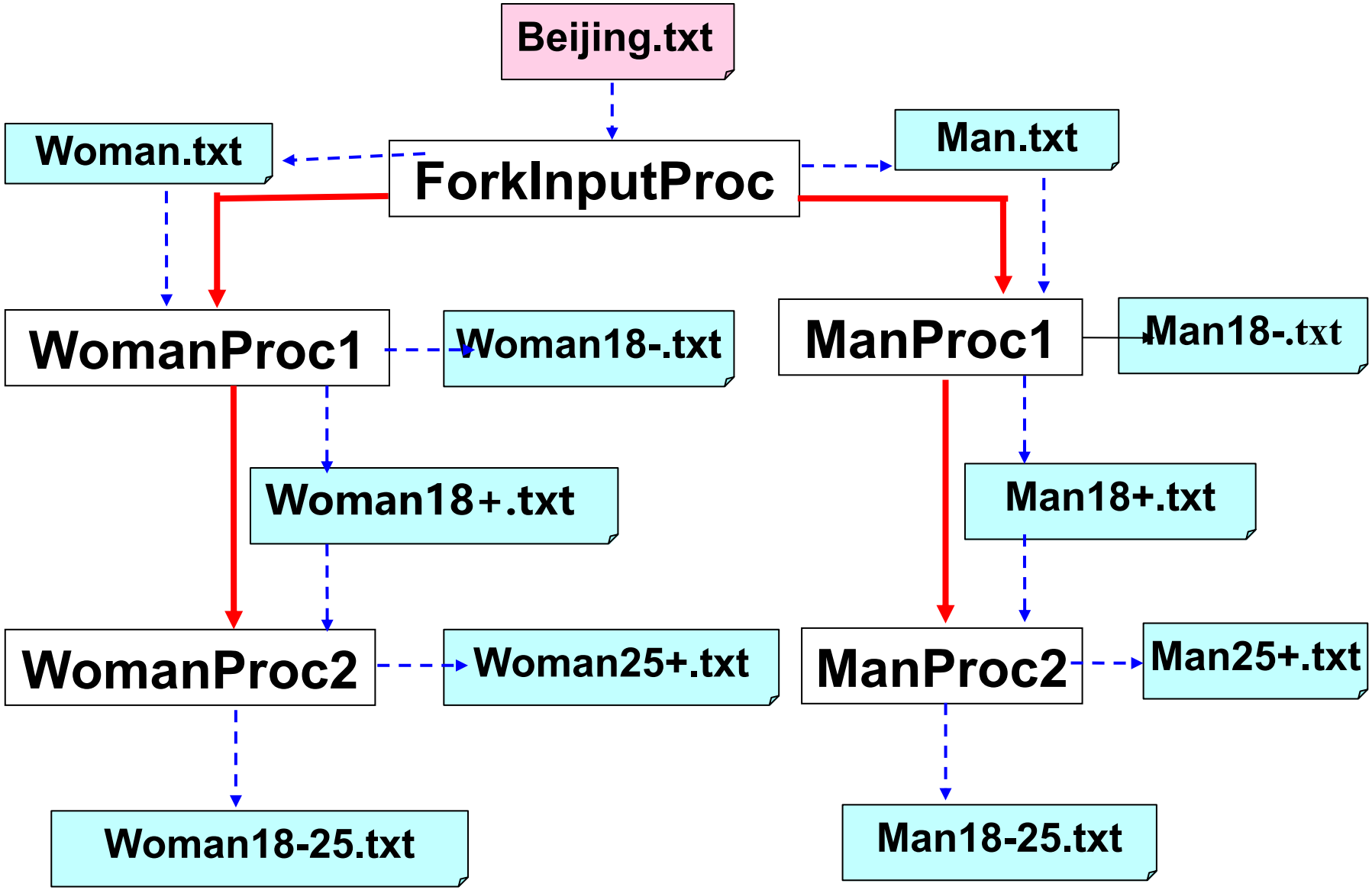


顺序批处理软件用户界面图

3. Example of design in Batch Sequential Architecture

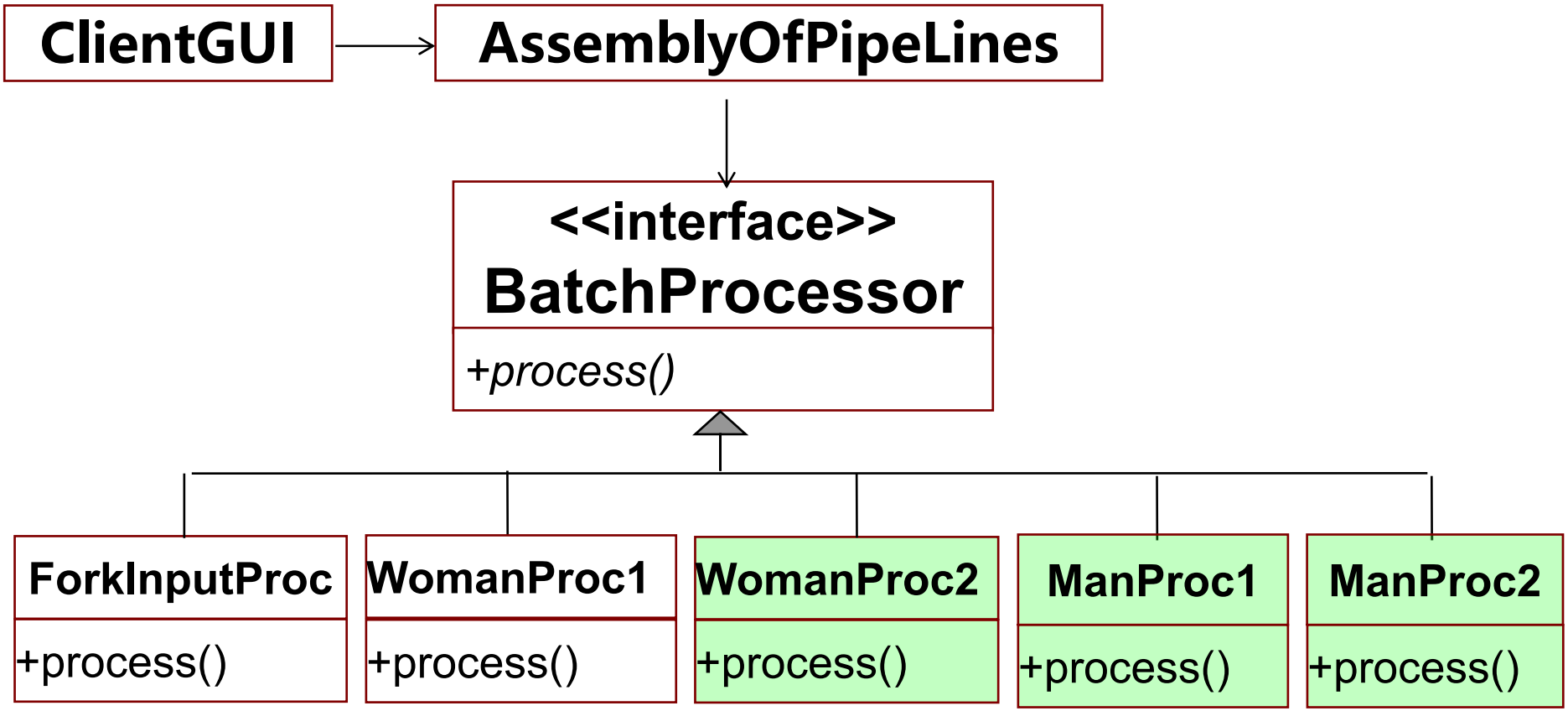
【例4】：利用批处理架构设计的一个人口信息处理程序

- **项目需求：实现一个人口信息处理程序**
- **This project is about processing population files**
  - **to get men whose ages are above 18 and under 25, for candidates to be soldiers.**
  - **This project also classifies the population by man, woman by ages.**
  - **Batch sequential architecture is used to process the files. The design is as below.**



人口信息处理程序软件体系结构图

设计类图如下：



评论：批处理架构盛行的年代是70-80年代；其时使用C语言等写各个处理过程，因此各个处理过程是独立的程序。现在使用Java语言进行设计的年代，可以写层次类。

3. Example of design in Batch Sequential Architecture

- **Project 2: 你的任务：**
- **在以上设计中，ClientGUI类，AssemblyOfPipeLines 类，BatchProcessing类, ForkInputProcessor类和WomanProcessor1类已经实现（写完代码）。**
- **你的任务是写代码实现WomanProcessor2类，ManProcessor1类和ManProcessor2类**

Back