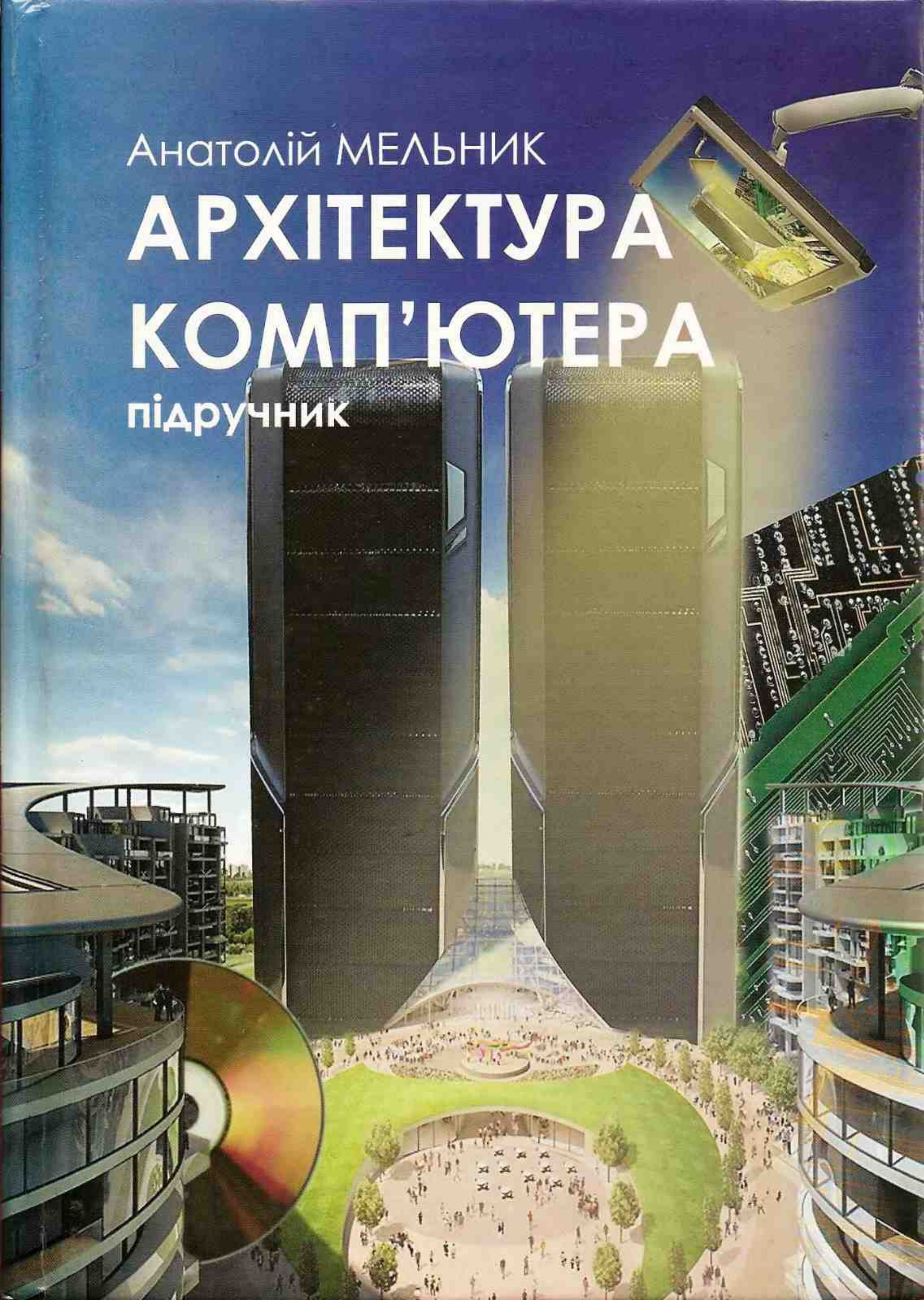
Анатолій МЕЛЬНИК

АРХІТЕКТУРА КОМП'ЮТЕРА

підручник



УДК 004.2
ББК 32.973
М 48

Рецензенти: **В. В. Божидарнік,** доктор технічних наук, професор, ректор Луцького державного технічного університету;

С. С. Забара, доктор технічних наук, професор, декан факультету комп'ютерних технологій Міжнародного університету розвитку людини "Україна";

В. П. Тарасенко, доктор технічних наук, професор, завідувач кафедри спеціалізованих комп'ютерних систем Національного технічного університету України "Київський політехнічний інститут".

Гриф надано Міністерством освіти і науки України, лист № 14/18-Г 222 від 28.01.2008 р.

Підручник охоплює весь комплекс питань, пов'язаних з теорією, принципами та методами побудови та організації функціонування комп'ютерів. Розкриті питання представлення команд та даних в комп'ютері, алгоритми виконання операцій обробки даних та принципи побудови арифметико-логічного пристрою. Описана структура й організація роботи комп'ютера з простою та складною системою команд, суперконвеєрного, суперскалярного, векторного та з довгим форматом команди. Розглянуті конфлікти в конвеєрі команд і методи їх усунення, структура та організація роботи пристрою керування. Наведена структура пам'яті комп'ютера, організація взаємодії між її рівнями, сегментна організація пам'яті й питання її захисту. Дане пояснення способів розпізнавання пристроїв введення-виведення та організація їх роботи. Розглянуті питання подальшого підвищення продуктивності комп'ютера та шляхи створення паралельних комп'ютерних систем.

Підручник призначений для студентів вищих навчальних закладів України. Він також може бути корисним для викладачів, фахівців в області комп'ютерних та інформаційних технологій.

ISBN 978-966-361-264-5

© Мельник А. О., 2008

ЗМІСТ

Передмова.....13

РОЗДІЛ 1.
Сучасний комп'ютер. Основні поняття.....17

1.1. Історичні аспекти розвитку комп'ютерів.....17
1.2. Функції, структура та характеристики комп'ютера.....18
1.2.1. Функції та основні функціональні вузли комп'ютера.....18
1.2.2. Тенденції зміни основних характеристик апаратних засобів комп'ютера.....20
1.2.3. Оцінка продуктивності комп'ютера.....23
1.2.3.1. Одиниці оцінки продуктивності.....23
1.2.3.2. Тестові програми для оцінки продуктивності.....25
1.2.4. Організація зв'язків між функціональними вузлами комп'ютера.....26
1.3. Архітектура комп'ютера.....28
1.3.1. Поняття архітектури комп'ютера.....28
1.3.2. Архітектурні принципи Джона фон Неймана.....30
1.3.3. Ненейманівські архітектури комп'ютерів.....31
1.4. Типи сучасних комп'ютерів.....33
1.4.1. Персональні комп'ютери.....34
1.4.2. Робочі станції.....40
1.4.3. Багатотермінальні системи.....41
1.4.4. Сервери.....42
1.4.5. Великі універсальні комп'ютерні системи.....43
1.4.6. Кластерні комп'ютерні системи.....44
1.4.7. Суперкомп'ютери.....46
1.4.8. Мікроконтролери.....48
1.4.9. Спеціалізовані комп'ютери.....49
1.5. Предмет та порядок розгляду матеріалу даної книги.....50
1.6. Підсумок розділу.....53
1.7. Література для подальшого читання.....53
1.8. Література до розділу 1.....54
1.9. Питання по розділу 1.....54

РОЗДІЛ 2.
Представлення даних у комп'ютері.....56

2.1. Позиційні системи числення.....56
2.2. Двійкові, вісімкові та шістнадцяткові числа.....57
2.3. Переведення чисел із системи числення з основою k у десяткову систему.....59
2.4. Переведення чисел із десяткової системи у систему числення з основою k.....59
2.5. Представлення чисел зі знаком.....60
2.5.1. Прямий код.....61
2.5.2. Обернений код.....61
2.5.3. Доповняльний код.....62

2.6. Формати даних.....63
2.6.1. Способи представлення чисел.....63
2.6.2. Числа з фіксованою комою.....63
2.6.3. Числа із рухомою комою.....65
2.6.4. Стандарт IEEE-754.....70
2.6.5. Кодування алфавітно-цифрової інформації.....72
2.6.5.1. Двійково-кодовані десяткові числа.....72
2.6.5.2. Розширений двійково-кодований десятковий код обміну EBCDIC.....74
2.6.5.3 Американський стандартний код інформаційного обміну ASCII.....75
2.6.5.4. Стандарт кодування символів Unicode.....76
2.7. Короткий зміст розділу.....77
2.8. Література для подальшого читання.....77
2.9. Література до розділу 2.....78
2.10. Питання до розділу 2.....78
2.11. Задачі до розділу 2.....79

Розділ 3.
Порядок виконання команд і програм в комп'ютері.....82

3.1. Кодування та виконання команд в комп'ютері.....82
3.1.1. Кодування команди та програми.....83
3.1.2. Порядок виконання команд.....84
3.1.3. Виконання команд на рівні регістрів процесора.....85
3.2. Типи операцій та команд.....87
3.2.1. Класифікація команд за типами операцій.....87
3.2.2. Команди обробки даних.....88
3.2.3. Команди переміщення даних.....89
3.2.4. Команди передачі керування.....90
3.2.4.1. Команди переходу.....91
3.2.4.2. Команди пропуску.....93
3.2.4.3. Команди звернення до підпрограм.....94
3.2.5. Команди введення-виведення.....96
3.2.6. Принципи формування системи команд комп'ютера.....96
3.2.7. Конвеєрне виконання команд.....98
3.3. Формати команд комп'ютера.....102
3.3.1. Класифікація архітектури комп'ютера за типом адресованої пам'яті.....102
3.3.2. Порівняльний аналіз форматів команд.....105
3.4. Способи адресації операндів.....107
3.4.1. Безпосередня адресація.....108
3.4.2. Пряма адресація.....108
3.4.3. Непряма адресація.....109
3.4.4. Способи адресації операндів на основі операції зміщення.....111
3.4.4.1. Відносна адресація.....111
3.4.4.2. Базова адресація.....111
3.4.4.3. Індексна адресація.....112
3.4.5. Сторінкова адресація.....114
3.4.6. Неявна адресація.....114
3.4.7. Стекова адресація.....114

3.4.8. Використання стекової адресації....115
3.4.9. Вибір способів адресації операндів....117
3.5. Приклади форматів команд....118
3.5.1. Формати команд комп'ютерної системи IBM 370....119
3.5.2. Формати команд комп'ютера Cyber-70....120
3.5.3. Формати команд сучасного комп'ютера....121
3.6. Вплив технології компілювання на систему команд комп'ютера....122
3.7. Архітектура системи команд комп'ютера....123
3.7.1. Класифікація архітектури комп'ютера за складом системи команд....123
3.7.2. Комп'ютери із складною та з простою системою команд....123
3.7.3. Особливості архітектури комп'ютера з простою системою команд....124
3.7.4. Архітектура комп'ютера з доповненою системою команд....125
3.7.5. Комп'ютери зі спеціалізованою системою команд....126
3.8. Короткий зміст розділу....129
3.9. Література для подальшого читання....129
3.10. Література до розділу 3....130
3.11. Питання до розділу 3....131

Розділ 4.
Процесор універсального комп'ютера....133

4.1. Процесор комп'ютера із складною системою команд....133
4.1.1. Одношинна структура процесора....133
4.1.2. Основні операції процесораа....135
4.1.2.1. Вибірка слова з пам'яті....135
4.1.2.2. Запам'ятовування слова в пам'яті....135
4.1.2.3. Обмін даними між регістрами....135
4.1.2.4. Виконання арифметичних і логічних операцій....136
4.1.3. Багатошинна структура процесора....137
4.1.4. Приклади виконання операцій в процесорі....138
4.1.4.1. Виконання операції додавання двох чисел....138
4.1.4.2. Виконання операції переходу....139
4.1.5. Особливості побудови процесора комп'ютера із складною системою команд..139
4.2. Процесор комп'ютера з простою системою команд....140
4.2.1. Вимоги до процесора комп'ютера з простою системою команд....140
4.2.2. Базові принципи побудови процесора комп'ютера з простою системою команд....140
4.2.3. Взаємодія процесора з пам'яттю в комп'ютері з простою системою команд...144
4.2.4. Виконання команд в процесорі комп'ютера з простою системою команд....146
4.2.4.1. Фаза вибирання команди....146
4.2.4.2. Фаза декодування команди....147
4.2.4.3. Фаза виконання та формування ефективної адреси....148
4.2.4.4. Фаза звернення до пам'яті та завершення умовного переходу....149
4.2.4.5. Фаза зворотного запису....150
4.2.5. Конвеєрна структура процесора комп'ютера з простою системою команд....151
4.2.5.1. Конвеєрний процесор....151
4.2.5.2. Мікродії ярусів конвеєрного процесора....155
4.3. Суперконвеєрні процесори....157

4.4. Суперскалярні процесори....158
4.5. Процесор векторного комп'ютера....160
4.6. Класифікація архітектури комп'ютера за рівнем суміщення опрацювання команд та даних....164
4.7. Короткий зміст розділу....164
4.8. Література для подальшого читання....165
4.9. Література до розділу 4....165
4.10. Питання до розділу 4....165

Розділ 5.
Запобігання конфліктам в конвеєрі команд....167

5.1. Структурні конфлікти....167
5.2. Конфлікти за даними....170
5.2.1. Типи конфліктів за даними....170
5.2.2. Методи зменшення впливу конфліктів за даними на роботу конвеєра команд..171
5.2.3. Призупинення виконання команди....172
5.2.4. Випереджувальне пересилання....172
5.2.5. Статична диспетчеризація послідовності команд у програмі під час компіляції..174
5.2.6. Динамічна диспетчеризація послідовності команд у програмі під час компіляції....176
5.2.7. Перейменування регістрів....177
5.3. Конфлікти керування....177
5.3.1. Типи конфліктів керування....177
5.3.2. Зниження втрат на вибірку команди, до якої здійснюється перехід....179
5.3.3. Зниження втрат на виконання команд умовного переходу....181
5.3.3.1. Введення буфера попередньої вибірки....181
5.3.3.2. Дублювання початкових ярусів конвеєра....182
5.3.3.3. Затримка переходу....183
5.3.3.4. Статичне передбачення переходу....183
5.3.3.5. Динамічне передбачення переходу....185
5.4. Покращена структура комп'ютера із спрощеною системою команд....189
5.5. Особливості запобігання конфліктам в суперскалярних процесорах....190
5.6. Комп'ютери з довгим форматом команди....192
5.7. Комп'ютери з комбінованою архітектурою....196
5.8. Комп'ютери з явним паралелізмом виконання команд....198
5.9. Короткий зміст розділу....200
5.10. Література для подальшого читання....201
5.11. Література до розділу 5....201
5.12. Питання до розділу 5....203

Розділ 6.
Алгоритми виконання операцій обробки даних....204

6.1. Логічні операції....204
6.1.1. Операція заперечення....206
6.1.2. Логічне І....206
6.1.3. Логічне АБО....206

6.1.4. Виключне АБО....207
6.2. Операції зсуву....207
6.2.1. Логічні зсуви....207
6.2.2. Арифметичні зсуви....208
6.2.3. Циклічні зсуви....208
6.3. Операції відношення....209
6.3.1. Порівняння двійкових кодів на збіжність....209
6.3.2. Визначення старшинства двійкових кодів....209
6.4. Арифметичні операції....210
6.4.1. Додавання двійкових чисел без знаків....210
6.4.2. Додавання двійкових чисел із знаками....212
6.4.3. Віднімання двійкових чисел....213
6.4.4. Множення двійкових чисел....214
6.4.4.1. Множення цілих двійкових чисел без знаків....215
6.4.4.2. Багатомісна операція додавання часткових добутків....216
6.4.4.3. Множення двійкових чисел із знаками....220
6.4.4.4. Прискорене множення двійкових чисел за методом Бута....221
6.4.5. Ділення двійкових чисел....222
6.4.6. Арифметичні операції над двійковими числами у форматі з рухомою комою....224
6.5. Операції обчислення елементарних функцій....226
6.5.1. Розклад функції в ряд та використання ітеративних обчислень....226
6.5.2. Обчислення елементарних функцій методом "цифра за цифрою"....226
6.5.3. Табличний метод обчислення елементарних функцій....228
6.5.4. Таблично-алгоритмічний метод обчислення елементарних функцій....228
6.6. Операції перетворення даних....229
6.6.1. Перетворення даних із формату з фіксованою у формат з рухомою комою та навпаки....229
6.6.2. Перетворення даних з двійково-десяткового коду в двійковий та навпаки...231
6.7. Операції реорганізації масивів і визначення їх параметрів....231
6.8. Операції обробки символів та рядків символів....232
6.9. Короткий зміст розділу....236
6.10. Література для подальшого читання....236
6.11. Література до розділу 6....236
6.12. Питання до розділу 6....237

Розділ 7.
Арифметико-логічний пристрій....239

7.1. Функції арифметико-логічного пристрою....239
7.2. Способи обробки даних в арифметико-логічному пристрої....240
7.3. Елементарні операції арифметико-логічного пристрою....241
7.4. Складні операції арифметико-логічного пристрою....243
7.5. Використання графа алгоритму при побудові арифметико-логічного пристрою....244
7.6. Виконання складних операцій в арифметико-логічному пристрої....245
7.7. Структура арифметико-логічного пристрою....246
7.8. Типи операційних пристроїв....249

7.9. Табличний операційний пристрій.....250
7.10. Багатотактовий операційний пристрій.....252
7.11. Однотактовий операційний пристрій.....254
7.12. Конвеєрний операційний пристрій.....255
7.13. Алгоритмічні операційні пристрої.....258
7.13.1. Пристрої додавання і віднімання двійкових чисел з фіксованою комою.....258
7.13.2. Пристрої множення двійкових чисел з фіксованою комою.....261
7.13.2.1. Багатотактовий пристрій множення двійкових чисел з молодших розрядів множника при нерухомому множеному з зсувом суми часткових добутків.....261
7.13.2.2. Багатотактовий пристрій множення двійкових чисел з молодших розрядів при нерухомій сумі часткових добутків з зсувом множеного вліво.....263
7.13.2.3. Багатотактовий пристрій множення двійкових чисел з старших розрядів при нерухомій сумі часткових добутків з зсувом множеного вправо.....264
7.13.2.4. Багатотактовий пристрій множення двійкових чисел з старших розрядів при нерухомому множеному з зсувом суми часткових добутків вліво.....266
7.13.2.5. Багатотактовий пристрій прискореного множення.....267
7.13.2.6. Однотактові пристрої множення двійкових чисел з фіксованою комою.....268
7.13.2.7. Конвеєрні пристрої множення двійкових чисел з фіксованою комою.....269
7.13.3. Пристрої ділення двійкових чисел з фіксованою комою.....270
7.13.3.1. Багатотактові пристрої ділення двійкових чисел з фіксованою комою.....270
7.13.3.2. Однотактові та конвеєрні пристрої ділення двійкових чисел з фіксованою комою.....272
7.13.4. Пристрої обчислення елементарних функцій методом "цифра за цифрою".....273
7.13.4.1. Багатотактовий пристрій обчислення елементарних функцій методом "цифра за цифрою".....273
7.13.4.2. Однотактовий та конвеєрний операційні пристрої обчислення елементарних функцій методом «цифра за цифрою».....274
7.13.5. Пристрої для виконання арифметичних операцій над числами з рухомою комою.....275
7.13.5.1. Пристрої додавання і віднімання чисел з рухомою комою.....275
7.13.5.2. Пристрої множення та ділення чисел з рухомою комою.....276
7.14. Таблично-алгоритмічні операційні пристрої.....277
7.15. Короткий зміст розділу.....280
7.16. Література для подальшого читання.....280
7.17. Література до розділу 7.....280
7.18. Питання до розділу 7.....281

Розділ 8.
Пристрій керування.....283

8.1. Функції та методи побудови пристрою керування.....283
8.2. Пристрій керування з жорсткою логікою.....284
8.2.1. Структура пристрою керування з жорсткою логікою.....284
8.2.2. Методи проектування пристрою керування з жорсткою логікою.....285
8.2.3. Пристрій керування на основі таблиць станів.....285
8.2.3.1. Абстрактні автомати.....285
8.2.3.2. Мови опису функціонування автоматів.....285
8.2.3.3. Структурний синтез цифрових автоматів.....291
8.2.4. Пристрій керування на основі синхронних елементів часової затримки.....294
8.2.5. Пристрій керування на основі лічильників.....295
8.3. Пристрій мікропрограмного керування.....297
8.3.1. Організація роботи пристрою мікропрограмного керування.....297
8.3.2. Організація мікропрограм в пам'яті мікрокоманд.....300
8.3.3. Горизонтальне та вертикальне мікропрограмування.....301
8.4. Порівняння пристроїв керування з жорсткою логікою та пристроїв мікропрограмного керування.....302
8.5. Короткий зміст розділу.....303
8.6. Література для подальшого читання.....304
8.7. Література до розділу 8.....304
8.8. Питання до розділу 8.....305

Розділ 9.
Багаторівнева пам'ять комп'ютера.....307

9.1. Типи та характеристики пам'яті комп'ютера.....308
9.1.1. Багаторівнева структура пам'яті комп'ютера.....308
9.1.2. Типи пам'яті.....308
9.1.3. Основні характеристики пам'яті.....313
9.2 Регістровий файл процесора.....315
9.2.1. Типи регістрових файлів.....315
9.2.2. Інтегрований багатопортовий регістровий файл.....316
9.2.3. Розподілений регістровий файл.....316
9.2.3.1. Кластерний розподілений регістровий файл.....317
9.2.3.2. Розподілений регістровий файл з керованою комутацією.....318
9.2.3.3. Розподілений регістровий файл з віконною організацією.....318
9.2.4. Ієрархічний регістровий файл.....319
9.2.5. Динамічна та статична організація збереження даних в регістрових файлах.....320
9.3. Пам'ять з асоціативним доступом.....321
9.3.1. Організація та типи пам'яті з асоціативним доступом.....321
9.3.2. Пам'ять з повним паралельним асоціативним доступом.....325
9.3.3. Пам'ять з неповним паралельним асоціативним доступом.....326
9.3.4. Пам'ять з послідовним асоціативним доступом.....326
9.3.5. Пам'ять з частково асоціативним доступом.....327
9.4. Основна пам'ять.....328

9.4.1. Структура основної пам'яті....328
9.4.2. Нарощування розрядності основної пам'яті....329
9.4.3. Нарощування ємності основної пам'яті....329
9.4.4. Розшарування пам'яті....330
9.5. Оперативний запам'ятовуючий пристрій....331
9.6. Постійний запам'ятовуючий пристрій....334
9.6.1. Організація роботи постійного запам'ятовуючого пристрою....334
9.6.2. Запрограмований при виготовленні постійний запам'ятовуючий пристрій...335
9.6.3. Одноразово запрограмований після виготовлення постійний запам'ятовуючий пристрій....336
9.6.4. Багаторазово програмовний постійний запам'ятовуючий пристрій....337
9.7. Зовнішня пам'ять....339
9.7.1. Магнітні диски....339
9.7.2. Масиви магнітних дисків з надлишковістю....341
9.7.2.1. Базовий тип дискових масивів RAID 0....343
9.7.2.2. Базовий тип дискових масивів RAID 1....343
9.7.2.3. Базовий тип дискових масивів RAID 2....344
9.7.2.4. Базовий тип дискових масивів RAID 3....344
9.7.2.5. Базовий тип дискових масивів RAID 4....345
9.7.2.6. Базовий тип дискових масивів RAID 5....346
9.7.2.7. Тип дискових масивів RAID 6....346
9.7.2.8. Тип дискових масивів RAID 7....347
9.7.2.9. Тип дискових масивів RAID 10....347
9.7.3. Оптична пам'ять....348
9.7.3.1. Постійна пам'ять на основі компакт дисків....349
9.7.3.2. Оптичні диски із стиранням....350
9.7.4. Магнітні стрічки....351
9.8. Короткий зміст розділу....352
9.9. Література для подальшого читання....353
9.10. Література до розділу 9....353
9.11. Питання до розділу 9....355

Розділ 10.
Організація пам'яті....357

10.1. Ієрархічна організація пам'яті комп'ютера....357
10.1.1. Різниця між продуктивністю процесора та пам'яті....357
10.1.2. Властивість локальності за зверненням до пам'яті....359
10.1.3. Принцип ієрархічної організації пам'яті....360
10.1.4. Характеристики ефективності ієрархічної організації пам'яті....361
10.1.5. Ієрархічна пам'ять сучасного комп'ютера....362
10.2. Організація обміну інформацією між процесором і основною пам'яттю через кеш пам'ять....363
10.2.1. Кеш пам'ять в складі комп'ютера....363
10.2.2. Порядок взаємодії процесора і основної пам'яті через кеш пам'ять....364
10.2.3. Забезпечення ідентичності вмісту блоків кеш пам'яті і основної пам'яті....365
10.2.4. Функція відображення....366
10.2.4.1. Типи функцій відображення....366

10.2.4.2. Повністю асоціативне відображення....367
10.2.4.3. Пряме відображення....369
10.2.4.4. Частково-асоціативне відображення....371
10.2.5. Порядок заміщення блоків в кеш пам'яті з асоціативним відображенням....373
10.2.6. Підвищення ефективності кеш пам'яті....374
10.3. Організація обміну інформацією між основною та зовнішньою пам'яттю....376
10.3.1. Статичний та динамічний розподіл пам'яті....376
10.3.2. Розподіл основної пам'яті за допомогою базових адрес....377
10.3.3. Віртуальна пам'ять....379
10.3.4. Сторінкова організація пам'яті....380
10.3.4.1. Основні правила сторінкової організації пам'яті....380
10.3.4.2. Реалізація сторінкової організації пам'яті....381
10.3.4.4. Апаратна реалізація сторінкової таблиці....384
10.3.5. Сегментна організація віртуальної пам'яті....388
10.4. Захист пам'яті від несанкціонованих звернень....391
10.4.1. Задачі захисту пам'яті....391
10.4.2. Захист пам'яті за допомогою регістра захисту....391
10.4.3. Захист пам'яті за граничними адресами....392
10.4.4. Захист пам'яті за значеннями ключів....392
10.4.5. Кільцева схема захисту пам'яті....393
10.5. Короткий зміст розділу....394
10.6. Література для подальшого читання....395
10.7. Література до розділу 10....395
10.8. Питання до розділу 10....397

Розділ 11.
Організація введення-виведення....399

11.1. Під'єднання зовнішніх пристроїв до комп'ютера....399
11.2. Розпізнавання пристроїв введення-виведення....401
11.3. Методи керування введенням-виведенням....403
11.4. Програмно-кероване введення-виведення....403
11.5. Система переривання програм та організація введення-виведення за перериваннями....405
11.5.1. Функції системи переривання програм....405
11.5.2. Характеристики системи переривання програм....406
11.5.3. Вхід в переривальну програму....407
11.5.4. Пріоритетне обслуговування переривання....409
11.5.5. Організація повернення до перериваної програми....410
11.5.6. Введення-виведення за перериваннями....411
11.6. Прямий доступ до пам'яті....412
11.7. Введення-виведення під керуванням периферійних процесорів....413
11.7.1. Принципи введення-виведення під керуванням периферійних процесорів....413
11.7.2. Причини застосування каналів введення-виведення....415
11.7.3. Функції каналів введення-виведення....416
11.7.4. Керуюча інформація каналу введення-виведення....417

11.7.5. Мультиплексний та селекторний канали введення-виведення....417
11.8. Короткий зміст розділу....419
11.9. Література для подальшого читання....420
11.10. Література до розділу 11....420
11.11. Питання до розділу 11....420

Розділ 12.

Паралельні комп'ютерні системи....422

12.1. Використання принципів паралельної обробки інформації в архітектурі комп'ютера....422
12.2. Вибір кількості процесорів в багатопроцесорній системі....426
12.3. Багатопотокова обробка інформації....428
12.4. Класифікація паралельних комп'ютерних систем....432
12.4.1. Класифікація Шора....432
12.4.2. Класифікація Фліна....435
12.5. Типи архітектур систем ОКМД....437
12.6. Типи архітектур систем МКМД....439
12.7. Організація комп'ютерних систем із спільною пам'яттю....439
12.7.1. Типи комп'ютерних систем із спільною пам'яттю....439
12.7.2. Системи з однорідним доступом до пам'яті....441
12.7.3. Системи з неоднорідним доступом до пам'яті....442
12.7.4. Системи лише з кеш пам'яттю....443
12.8. Організація комп'ютерних систем із розподіленою пам'яттю....444
12.9. Комунікаційні мережі багатопроцесорних систем....445
12.9.1. Типи комунікаційних мереж....445
12.9.2. Основні характеристики комунікаційних мереж багатопроцесорних систем....448
12.9.3. Статичні топології комунікаційних мереж багатопроцесорних систем....449
12.9.4. Шинні динамічні комунікаційні мережі багатопроцесорних систем....453
12.9.5. Комутуючі динамічні комунікаційні мережі багатопроцесорних систем....456
12.9.5.1. Типи комутуючих динамічних комунікаційних мереж....456
12.9.5.2. Координатна мережа....457
12.9.5.3. Матрична одноярусна комутуюча мережа....458
12.9.5.4. Багатоярусні блокуючі комутуючі мережі....458
12.9.5.5. Багатоярусні неблокуючі комутуючі мережі з реконфігурацією....460
12.9.5.6. Багатоярусні неблокуючі комутуючі мережі....463
12.10. Короткий зміст розділу....464
12.11. Література для подальшого читання....465
12.12. Література до розділу 12....465
12.13. Питання до розділу 12....467

Передмова

Комп'ютерні технології все стрімкішими темпами входять в усі сфери життя суспільства, що викликає потребу підготовки фахівців за відповідними напрямами, в першу чергу за напрямами "Комп'ютерна інженерія", "Комп'ютерні науки", "Програмна інженерія", "Прикладна математика", "Захист інформації в комп'ютерних системах та мережах", "Комп'ютеризовані системи, автоматика і управління". Однією з базових дисциплін підготовки бакалаврів за названими напрямами є дисципліна "Архітектура комп'ютера". Цей підручник повною мірою відповідає змісту міністерських програм за вищеназваними напрямами.

Запропонований підручник орієнтований на студентів вищих навчальних закладів, викладачів та аспірантів комп'ютерних напрямів підготовки та покликаний прищепити талановитій молоді розуміння основних принципів побудови та організації сучасних комп'ютерів. Він також може бути корисним для фахівців в області комп'ютерних та інформаційних технологій.

Підручних охоплює весь комплекс питань, пов'язаних з теорією, принципами та методами побудови й організації функціонування комп'ютерів та складається з дванадцяти розділів.

В першому розділі розглянуто історичні аспекти розвитку комп'ютерів, подано основні поняття, функції та основні функціональні вузли комп'ютера, їх взаємозв'язок, а також загальну організацію роботи комп'ютера. Розглянуто характерні риси та сфери застосування різних типів комп'ютерів. Обґрунтовано місце предмета даної книги серед суміжних дисциплін та порядок його розгляду.

В наступному, другому, розділі розглянуто основні елементи архітектури комп'ютера. Показано як кодуються та виконуються команди в комп'ютері. Проведена класифікація команд відповідно до ініційованих ними типів операцій та детально розглянуті команди обробки даних, переміщення даних, передачі керування, введення-виведення. Введено поняття конвеєрного виконання команд – одного з видів паралелізму на рівні команди. Розглянуто три типи архітектур комп'ютера за типом адресованої пам'яті: стекова, акумуляторна, та на основі регістрів загального користування, їх переваги і недоліки. Наведено різні способи адресації, включаючи безпосередню, пряму, непряму, базову, індексну, сторінкову і стекову. Детально розглянуто архітектури комп'ютерів з складною, з простою, з доповненою та з орієнтованою системою команд.

В третьому розділі розкрито основні питання представлення даних в комп'ютері, які є важливими для розуміння матеріалу наступних розділів. Наведено правила

подання даних в позиційних системах числення та переведення чисел із однієї системи числення до іншої. Описано представлення чисел зі знаком в прямому, оберненому та доповняльному кодах а також формати даних з фіксованою та з рухомою комою, включаючи стандарт IEEE-754. Розглянуто питання кодування алфавітно-цифрової інформації кодами ASCII, EBCDIC та Unicode.

В четвертому розділі розкрито основні алгоритми виконання в комп'ютері операцій обробки даних: логічних, зсуву, відношення, арифметичних, обчислення елементарних функцій, перетворення даних (перетворення із формату з фіксованою в формат з рухомою комою і навпаки, перетворення з двійково-десяткового коду в двійковий і навпаки), реорганізації масивів і визначення їх параметрів (сортування, пошук максимуму або мінімуму, вибір заданого масиву, зсув елементів масиву, стиск масиву), обробки символів та стрічок символів (пошук символу, зсув, заміна символів в стрічці, пакування стрічок символів, порівняння стрічок символів).

В п'ятому розділі розглянуто принципи побудови арифметико-логічного пристрою (АЛП) сучасних комп'ютерів, який є одним з основних вузлів процесора. Розкрито структури АЛП для виконання елементарних та складних операцій. Описано багатоблокові АЛП на основі табличних, однотактових, багатотактових та конвеєрних операційних пристроїв для виконання операцій додавання, віднімання, множення, ділення та обчислення елементарних функцій над двійковими числами в форматах з фіксованою та рухомою комою.

В шостому розділі розглянуто структуру та організацію роботи пристрою керування з жорсткою логікою, а також методи його проектування: на основі таблиць станів, на основі елементів часової затримки та на основі лічильників. Описано роботу та основні принципи, покладені в основу побудови пристрою мікропрограмного керування. Розглянуто питання розміщення мікрокоманд в пам'яті, формат мікрокоманди і способи його оптимізації.

В сьомому розділі описано місце процесора в комп'ютері, його функції та склад, розглянуто виконання основних операцій процесора. Виділено базові принципи побудови процесора комп'ютера з складною та з простою системою команд. Описано конвеєрну структуру процесора з простою системою команд та принципи побудови суперконвеєрних, суперскалярних та векторних процесорів.

У восьмому розділі розглянуто конфлікти в конвеєрі команд та методи їх усунення. Проведено аналіз методів запобігання появі трьох класів конфліктів: структурних, за даними та керування. Наведено приклади структур конвеєрних процесорів, в тому числі суперскалярних, в яких зменшено імовірність виникнення конфліктів. Розглянуто архітектури комп'ютерів, в яких відсутні конфлікти команд, а саме комп'ютерів з довгим форматом команди.

В дев'ятому розділі розглянуто структуру та основні характеристики пам'яті комп'ютера. Проаналізовано можливі варіанти організації регістрових файлів процесорів та пам'яті з асоціативним доступом. Описано будову запам'ятовуючих пристроїв, які може містити основна пам'ять, способи прискорення доступу до інфор-

мації, нарощування розрядності та ємності основної пам'яті. Подано типи сучасних дискових систем та обґрунтування використання масивів магнітних дисків з надлишковістю, а також робота шести базових типів дискових масивів RAID: RAID 0, RAID 1, ... , RAID 5 та дискових масивів, створених на їх основі. Описано роботу оптичної пам'яті та пам'яті на магнітних стрічках.

Виходячи з принципу ієрархічної організації пам'яті, в десятому розділі описано організацію взаємодії між рівнями ієрархічної пам'яті. Зокрема, наведено принципи обміну інформацією між рівнями ієрархічної пам'яті, характеристики, які використовуються для оцінки ефективності ієрархічної пам'яті, пояснено місце кеш пам'яті в складі комп'ютера, призначення і логіку роботи кеш пам'яті, описано переваги поділу кеш пам'яті першого рівня на кеш пам'ять даних та кеш пам'ять команд. Обґрунтовано різноманітність методів відображення основної пам'яті на кеш пам'ять, показано якими методами забезпечується узгодженість вмісту основної і кеш пам'яті, чим обумовлене введення додаткових рівнів кеш пам'яті й які чинники впливають на вибір ємності кеш пам'яті і розміру блоку. Введено поняття статичного і динамічного розподілу пам'яті, поняття віртуальної пам'яті та сторінкової організації пам'яті. Описано алгоритми заміщення, які використовуються при завантаженні в основну пам'ять вмісту зовнішньої пам'яті, призначення буфера швидкого перетворення адреси. Розглянуто сегментну організацію пам'яті та питання захисту пам'яті.

В одинадцятому розділі наведено пояснення способів розпізнавання пристроїв введення-виведення з використанням шини введення-виведення, лінії активації та прихованого пам'яттю введення-виведення. Наведено схему та описано функції інтерфейсної схеми пристроїв введення-виведення. Наведено чотири загальних методи керування введенням-виведенням і пояснено суть, переваги та недоліки програмно керованого введення-виведення. Введено основні поняття та характеристики системи переривання програм. Наведено опис прямого доступу до пам'яті, його переваги і недоліки, описано організацію введення-виведення під керуванням периферійних процесорів.

У дванадцятому розділі розглянуто питання подальшого підвищення продуктивності обробки інформації шляхом створення паралельних комп'ютерних систем. Наведено основні положення класифікації комп'ютерних систем, запропоновані Шором та Фліном, і відповідні класифікаціям структури комп'ютерних систем, зокрема з спільною та з розподіленою пам'яттю, з однорідним доступом до пам'яті та з неоднорідним доступом до пам'яті, а також лише з кеш пам'яттю. Наведено типи комунікаційних мереж багатопроцесорних систем.

Головною ідеєю книги є концентроване викладення та пояснення базових концепцій побудови сучасних комп'ютерів без заглиблення в роботу конкретних комп'ютерних систем, що дає змогу зорієнтуватися у взаємозалежностях та зв'язках елементів архітектури комп'ютера з метою забезпечення її цілісного бачення і розуміння, та закласти фундамент, який дозволяє в разі потреби набути детальнішої спеціалізації.

Матеріал підручника написано на основі курсів лекцій, прочитаних автором в Національному університеті "Львівська політехніка" та в кількох закордонних університетах, та на основі його багаторічної наукової діяльності. Хотів би виразити щиру подяку багатьом колегам і друзям за підтримку та допомогу в цій моїй діяльності. Окрема подяка доцентам кафедри ЕОМ Національного університету "Львівська політехніка" В. С. Глухову та В. В. Троценку за обговорення та поради, зроблені під час написання підручника. І, звичайно, велика подяка моїй дружині Тетяні та синам Віктору і Юрію за терпіння та розуміння необхідності виконання цієї роботи.

Ми з вдячністю приймемо всі зауваження та побажання.

Автор

Розділ 1

Сучасний комп'ютер. Основні поняття

В цьому розділі розглядаються історичні аспекти розвитку комп'ютерів, подаються основні поняття, функції та основні функціональні вузли комп'ютера, їх взаємозв'язок, а також загальна організація роботи комп'ютера. Проводиться аналіз та даються пояснення одиниць вимірювання технічних характеристик комп'ютера та аналізуються закономірності їх зміни з розвитком елементної бази. Вводиться поняття архітектури комп'ютера та виділяються особливості архітектури Джона фон Неймана та відмінні риси комп'ютерів ненейманівських архітектур. Розглядаються характерні риси та сфери застосування різних типів комп'ютерів: персональних, робочих станцій, багатотермінальних систем, серверів, великих універсальних комп'ютерних систем, кластерів, мікроконтролерів, суперкомп'ютерів та спеціалізованих комп'ютерів. Обґрунтовується місце предмету даної книги серед суміжних дисциплін та порядок його розгляду.

1.1. Історичні аспекти розвитку комп'ютерів

В 1946 році Джон фон Нейман (John von Neumann) описав архітектуру комп'ютера, яка є найпоширенішою і в даний час. В своїй науковій роботі він назвав новий пристрій обчислювальним інструментом (computing instrument), звідки логічно появилася сучасна назва комп'ютера. Ця архітектура дістала назву архітектури Джона фон Неймана, хоча в зарубіжній літературі частіше використовується термін "Instruction Set Architecture" ("архітектура системи команд"), що вказує як на рівень опису архітектури комп'ютера, так і на програмний принцип його роботи. До перших розробників та впроваджувачів цієї архітектури належать Чарльз Бебідж (Charles Babbedz), конструктор двох механічних комп'ютерів – різницевої та аналітичної машини (1822–1833 рр.), Джон Атанасов (John Atanasoff) – автор першого спеціалізованого комп'ютера ABC (1939 рік), Конрад Цузе (Konrad Cuze) (1936–1944 рр.) та Ховард Айкен (Howard Aiken) (1941–1946 рр.) конструктори перших електромеханічних комп'ютерів відповідно Z1–Z4 та Mark1–Марк2, Алан Тюрінг (Allan Turing) – один з розробників першого електронного комп'ютера Colossus (1943 рік), Преспер Екерт (Presper Eckert) та Джон Моучлі (John Mauchly) – розробники першого універсального комп'ютера ENIAC (1946 рік), який був описаний в науковій роботі Джона фон Неймана, Мауріс Уілкс (M. Wilkes) – розробник проекту комп'ютера EDSAC (1946 рік).

До перших електронних комп'ютерів з програмним керуванням належить і мала електронно-лічильна машина (МЕЛМ), створена в 1951 році в Інституті електродинаміки Академії Наук України, м. Київ під керівництвом академіка С. А. Лєбєдєва (рис. 1.1).

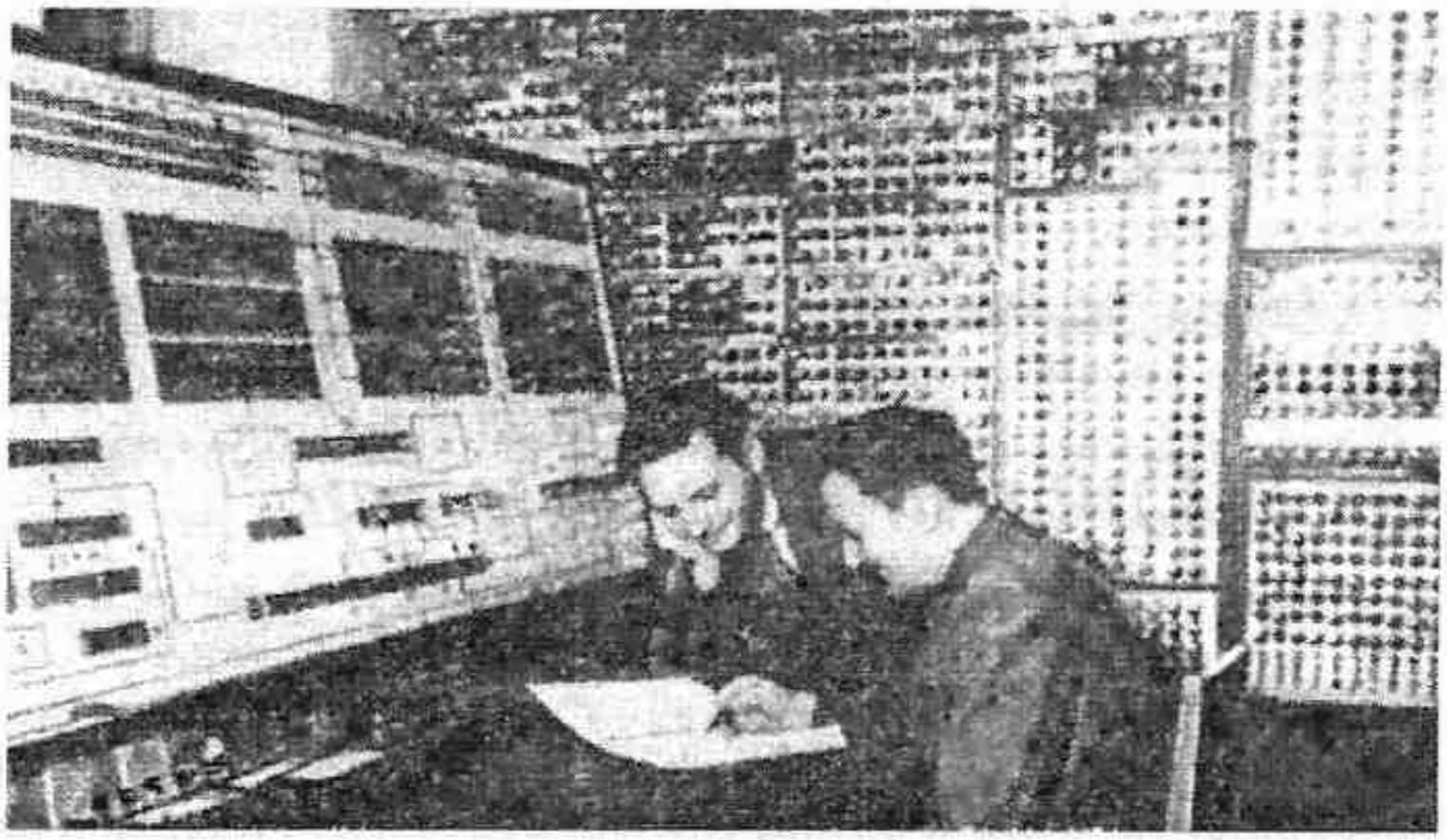


Рис. 1.1. Перший в СРСР універсальний електронний комп'ютер МЕЛМ (Україна, 1951)

Сучасний стан технології проектування комп'ютерів базується на теоретичних роботах ряду видатних дослідників, в першу чергу С. Крея, засновника фірми Cray Research, яка випустила кілька поколінь найпотужніших в світі комп'ютерів; М. Фліна, який розробив найуживанішу класифікацію комп'ютерів, Г. Амдаля, який сформував ряд новітніх положень розвитку комп'ютерів, Р. Томасуло, який запропонував підхід до вирішення питання ефективного завантаження комп'ютера, Д. Коука, який запропонував архітектуру "Америка", на основі якої збудовано комп'ютер з спрощеною системою команд RS/6000 фірми IBM (усі вони є вихідцями з цієї фірми), Д. Хеннессі (Стенфордський університет), автора архітектури комп'ютера з спрощеною системою команд MIPS, засновника фірми MIPS (зараз це підрозділ фірми Silicon Graphics, однієї з провідних фірм по виготовленню суперкомп'ютерів), Д. Паттерсона (Каліфорнійський університет), автора архітектури комп'ютерів з спрощеною системою команд RISC I та RISC II, технології RAID збереження даних тощо. Їх ідеї були апробовані шляхом створення та практичного використання комп'ютерів фірм Cray Research, IBM, Silicon Graphics, Intel, MIPS, SUN Microsystems, Digital Equipment Corporation, Transmeta та багатьох інших.

1.2. Функції, структура та характеристики комп'ютера

1.2.1. Функції та основні функціональні вузли комп'ютера

Комп'ютер представляє собою електронний пристрій, який містить апаратні засоби і програмне забезпечення та автоматично, відповідно до програми, виконує алгоритм вирішення заданої задачі. Під алгоритмом розуміють точний припис, що задає обчислювальний процес вирішення задачі, а під задачею – сформульоване намагання отримати з множини вхідних даних і початкових умов та з множини можливих вихідних даних підмножину вихідних даних, що повністю задовольняють початкові умови і вхідні дані. Алгоритм характеризується множиною параметрів вхідних, проміжних та вихідних даних, правилом вводу даних, правилом початку, правилом обробки даних, правилом закінчення, правилом виводу даних. Для виконання алгоритму комп'ютер приймає вхідну

інформацію в цифровій формі, обробляє її відповідно до вказівок команд програми виконання алгоритму, та видає результати обчислень.

До основних функцій, які виконує комп'ютер, належать наступні:

- сприйняття вхідної інформації – вхідних даних, які підлягають обробці, та програм вирішення задач (програм обробки вхідних даних);
- зберігання інформації, тобто вхідних і проміжних даних та результатів обчислень, програм вирішення задач, довідникової інформації, програм операційної системи комп'ютера і т. д.;
- виконання арифметичних, логічних та інших операцій;
- автоматичне керування роботою складових частин комп'ютера, їх взаємодією між собою та з зовнішніми пристроями відповідно до програми;
- виведення результатів обчислень.

Для забезпечення виконання цих функцій до складу комп'ютера повинні входити такі основні функціональні вузли: пристрої введення інформації, пристрої виведення інформації, пам'ять та процесор, як це показано на рис. 1.2.

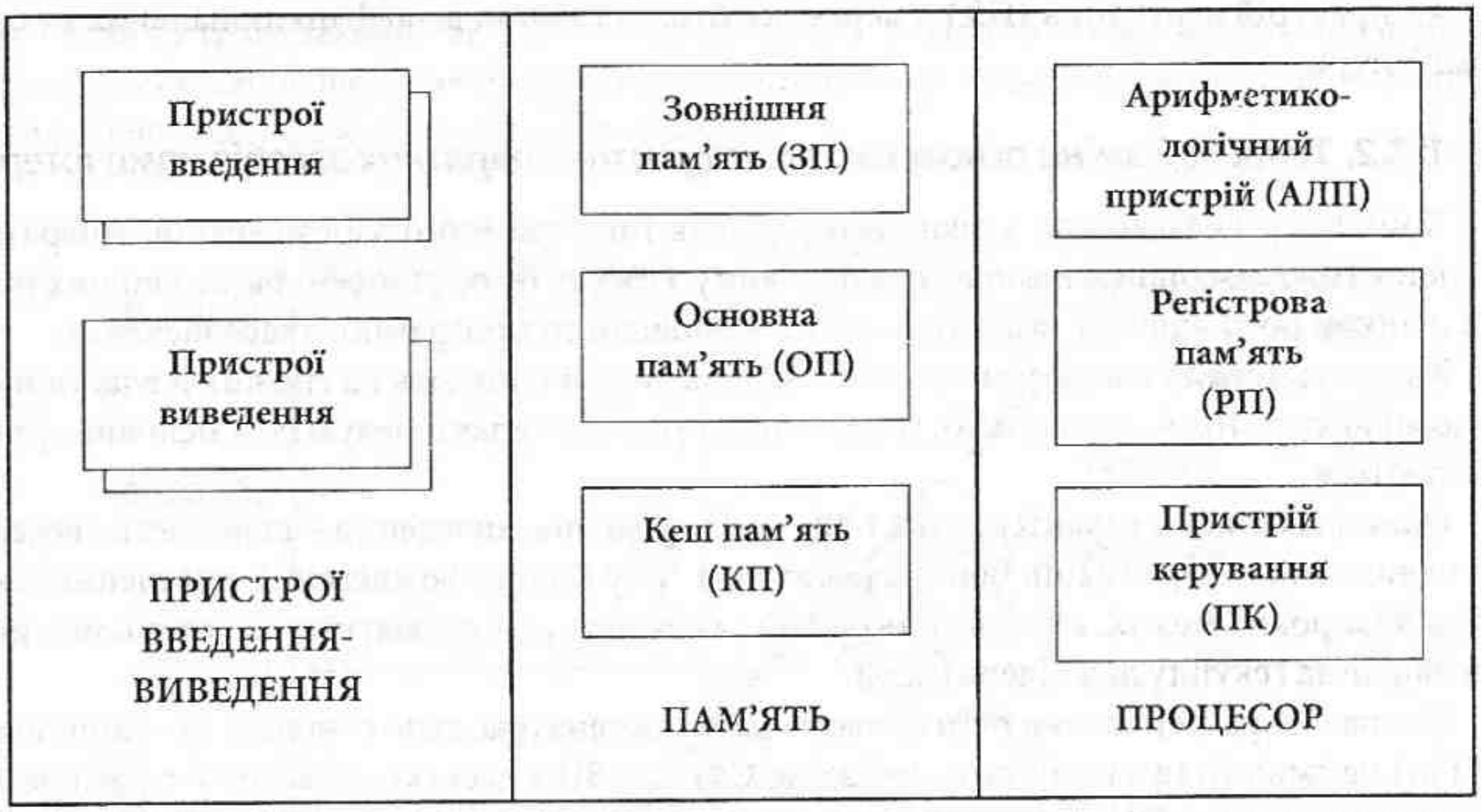


Рис. 1.2. Основні функціональні вузли комп'ютера

Коротко зупинимось на функціях та складі кожного вузла комп'ютера.

Пристрої введення-виведення виконують введення та виведення інформації. До числа пристроїв введення-виведення належать:

- пристрої введення – клавіатура, миша, сканер, відеокамера і т. д.;
- пристрої виведення – монітори (з електронно-променевою трубкою та рідкокристалічні), принтер, графопобудовувач і т. д.

Пам'ять призначена для зберігання інформації. Типи пам'яті:

- кеш пам'ять (КП) – високошвидкісна пам'ять невеликої ємності, використання якої дозволяє прискорити обмін інформацією між основною пам'яттю і процесором.

> основна пам'ять (ОП) – пам'ять великої ємності, яка зберігає інформацію, що підлягає обробці в процесорі. До інформації, яка зберігається в ОП, належать вхідні дані, які підлягають обробці відповідно до виконуваного алгоритму, результати проміжних обчислень, вихідні дані, команди програми виконання алгоритму, активна частина операційної системи комп'ютера. Інформація в ОП постійно змінюється, тобто це пам'ять для короткотермінового зберігання інформації;

> зовнішня пам'ять (ЗП) – пам'ять великої ємності для зберігання всієї інформації комп'ютера. Якщо при вимкненні комп'ютера інформація в ОП пропадає, тобто вона є енергозалежною, то на інформацію ЗП вимкнення комп'ютера не впливає, тобто це є енергонезалежна пам'ять для довготермінового зберігання інформації.

Процесор виконує обробку інформації та керує роботою інших вузлів комп'ютера. До складу процесора входять наступні функціональні вузли:

> арифметико-логічний пристрій (АЛП) – набір комбінаційних схем, які виконують арифметичні, логічні та інші операції;

> регістрова пам'ять (РП) – набір програмно доступних регістрів, в яких зберігається найчастіше використовувана в процесорі інформація;

> пристрій керування (ПК) – керує роботою та взаємодією функціональних вузлів комп'ютера.

1.2.2. Тенденції зміни основних характеристик апаратних засобів комп'ютера

Комп'ютер складається з апаратних засобів і програмного забезпечення. Апаратні засоби є базовим рівнем комп'ютера, на якому можуть бути створені багато вищих рівнів шляхом розроблення та завантаження відповідного програмного забезпечення.

Характеристики апаратних засобів в значній мірі впливають на споживчі властивості комп'ютера. Тому розглянемо, якими показниками характеризуються основні вузли комп'ютера.

Одна з важливих характеристик пристроїв введення-виведення – швидкість введення та виведення інформації. Вона залежить від типу пристрою введення-виведення і лежить в широких межах: від одиниць байтів за секунду для клавіатури до сотень мільйонів байтів за секунду для відеокамери.

Основні характеристики регістрового файлу процесора, кеш, основної та зовнішньої пам'яті це ємність та час доступу до даних. Ємність ЗП у десятки, а часто і у тисячі разів перевищує ємність ОП. Наприклад, у сучасних персональних комп'ютерах ємність ОП знаходиться в межах від сотень мільйонів до кількох мільярдів байтів, тоді як ємність ЗП знаходиться в межах від десятків до сотень мільярдів байтів.

Розділення пам'яті на ОП та ЗП викликане сповільненням її роботи із збільшенням ємності. Так, якщо час звернення до ЗП становить тисячні частки секунди, то час звернення до ОП становить мільйонні частини секунди. Це ж саме стало причиною появи кеш пам'яті, яка має меншу ніж ОП ємність, але за часом запису-зчитування даних є значно ближчою до регістрової пам'яті процесора.

Основними характеристиками процесора є його продуктивність, точність та динамічний діапазон представлення даних. Продуктивність залежить від частоти роботи процесора та його структури і вимірюється кількістю виконуваних операцій за секунду. Сучасні процесори мають продуктивність понад 1 мільярд операцій за секунду.

Крім того, всі вузли комп'ютера характеризуються швидкістю обміну з іншими пристроями, надійністю, ціною, споживаною потужністю, масою, габаритами, стійкістю до зовнішніх факторів впливу (температурний діапазон, вологість, вібрація і т. д.). Значення цих характеристик в значній мірі визначаються досягненнями інтегральної технології, зокрема проектними розмірами напівпровідникових елементів кристалу, які впливають як на ємність кристалу, а відповідно на масу та габарити, так і на частотні параметри його елементів.

Для того, щоб успішно проектувати комп'ютери та знати перспективи їх розвитку, необхідно знати тенденції зміни з часом перерахованих вище характеристик комп'ютера, в першу чергу характеристик його елементної бази.

Продуктивність. Продуктивність комп'ютера змінюється через зміну характеристик його функціональних вузлів. Розглянемо як це відбувається на прикладі процесора.

З кожним роком зростає ємність кристалу, що дозволяє розмістити на ньому все складніші схеми. Відповідно до емпіричного закону Мура, відкритого ним в 1965 році, кількість транзисторів, які вдається розмістити на кристалі мікросхеми, подвоюється кожні 12 місяців. В середині 90-х років Мур скоригував свій закон, який тепер твердить, що кількість транзисторів, які вдається розмістити на кристалі мікросхеми, подвоюється кожні 18 місяців, що підтверджує рис. 1.3, де наведено кількість транзисторів на кристалах процесорів провідних виробників світу.

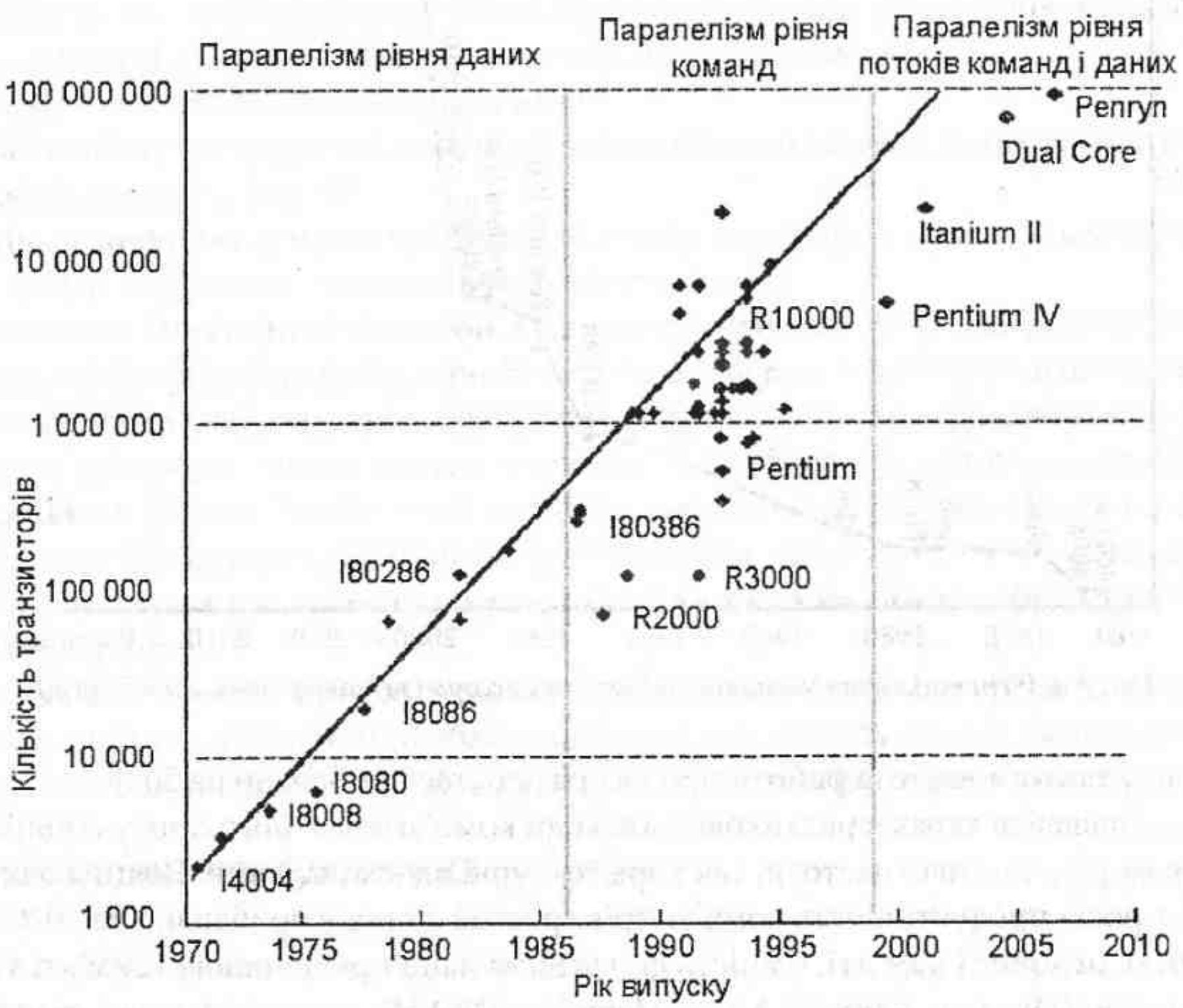


Рис. 1 3. Ріст ємності кристалу з роками

З ростом ємності кристалу з'являються можливості модифікації структури процесора шляхом ускладнення його основних вузлів та введення допоміжних вузлів, які дозволяють прискорити обробку даних. Метою такої модифікації є в першу чергу підвищення продуктивності процесора. Як видно з рис. 1.3, до середини 80-х років для цього використовувався паралелізм рівня даних, тобто суміщення в часі обробки в процесорі деякої кількості даних, далі паралелізм рівня команд, тобто суміщення в часі виконання в процесорі деякої кількості різних команд. Сьогодні це паралелізм рівня потоків команд і даних. Всі названі типи паралелізму будуть розглянуті далі при розгляді організації роботи процесора.

Зі зменшенням розмірів транзисторів, які реалізуються на кристалі мікросхеми, підвищується і тактова частота роботи процесора (рис. 1.4).

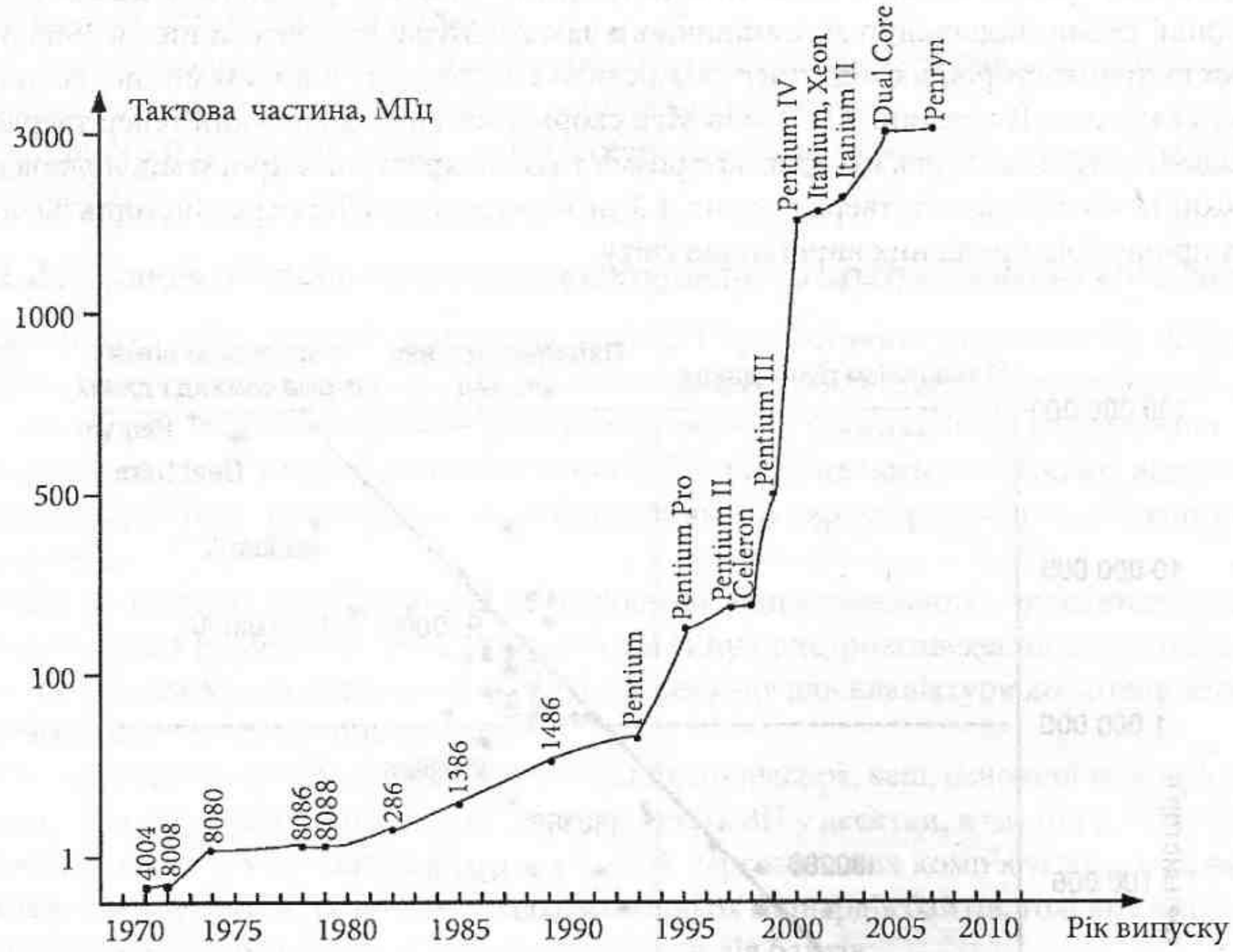


Рис. 1.4. Ріст тактової частоти роботи процесора (за матеріалами Intel Corp.)

Щороку тактова частота роботи процесора зростає приблизно на 30 %.

Узагальнюючою характеристикою швидкодії комп'ютера є його продуктивність, яка враховує як ріст тактової частоти, так і архітектурні вдосконалення. Починаючи з середини 80-х років продуктивність комп'ютерів зростає щороку приблизно на 50 %.

Ємність основної пам'яті. Ємність динамічної напівпровідникової пам'яті з довільним доступом (Dynamic Random Access Memory – DRAM) зростає в чотири рази кожних три роки (рис. 1.5).

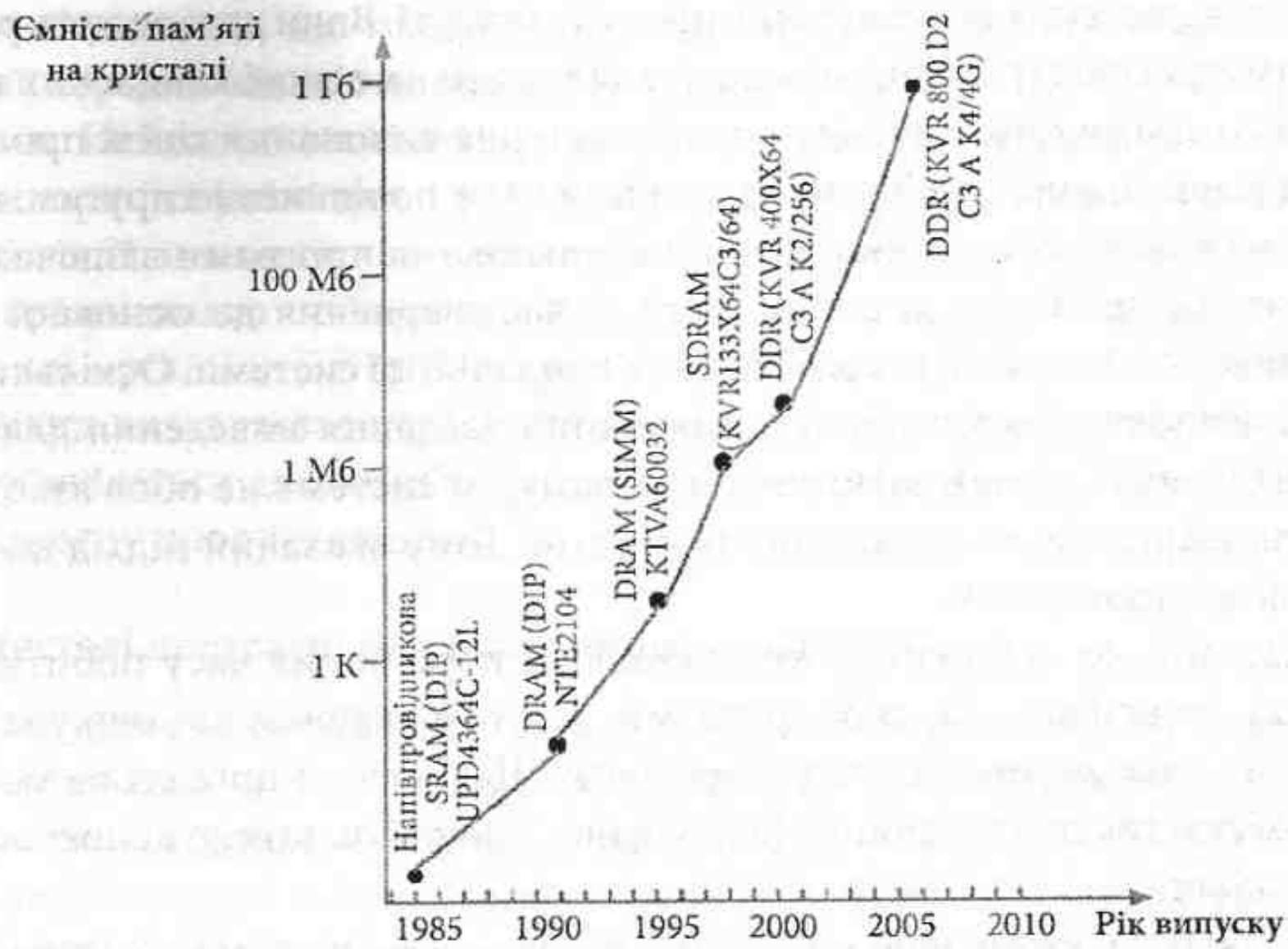


Рис. 1.5. Ріст ємності динамічної напівпровідникової пам'яті

Час читання/запису DRAM зменшується в два рази за чотири роки, тоді як в процесорі за той же час частота зростає більш як в три рази. Для узгодження швидкодії процесора та пам'яті використовуються наступні підходи та засоби:

- одно- та багаторівнева кеш пам'ять;
- збільшення розрядності шин обміну для передачі кількох даних паралельно;
- розшарування пам'яті.

Ємність зовнішньої пам'яті. Щороку ємність зовнішньої дискової пам'яті збільшується в 2 рази за рахунок підвищення щільності запису.

Точність та динамічний діапазон. Ріст ємності кристалу дозволяє реалізувати в ньому складні багаторозрядні вузли комп'ютера. Сьогоднішні комп'ютери в основному є 32- та 64-розрядними з забезпеченням обробки даних як в форматі з фіксованою комою, так і в форматі з рухомою комою, яка має значно ширший діапазон представлення даних.

Швидкість обміну. Підвищення швидкості обміну між вузлами комп'ютера досягається шляхом підвищення тактової частоти їх роботи, конструктивних та технологічних вдосконалень, а також шляхом створення нових стандартів обміну та введення пристроїв для їх підтримки.

З розвитком елементної бази також підвищується надійність роботи комп'ютера, знижується питома ціна одиниці обладнання, а також зменшується її питома споживана потужність, маса, та габарити, покращуються експлуатаційні характеристики.

1.2.3. Оцінка продуктивності комп'ютера

1.2.3.1. Одиниці оцінки продуктивності

Продуктивність є однією з основних характеристик комп'ютера, яка залежить як від технологічних, так і від архітектурних рішень. В процесі розвитку комп'ютерної техніки

з'явилося декілька методик вимірювання продуктивності. Вони дозволяють розробникам і користувачам здійснювати вибір між альтернативами на основі кількісних показників.

Припустимо, що використано два комп'ютери для виконання тієї ж програми. Якщо перший комп'ютер виконав програму за менший час у порівнянні з другим, можна говорити, що перший комп'ютер є швидшим. Час виконання програми включає час роботи процесора, час звернення до дискової пам'яті, час звернення до основної пам'яті, час введення-виведення даних і накладні витрати операційної системи. Оскільки при роботі в мультипрограмному режимі під час очікування введення-виведення для однієї програми, процесор може виконувати іншу програму, то система не обов'язково мінімізуватиме час виконання даної конкретної програми. Тому вказаний підхід до порівняння комп'ютерів не є досконалим.

Іншим підходом до порівняння комп'ютерів є порівняння часу роботи процесора, необхідного для виконання заданої програми, який не включає час очікування введення-виведення або час виконання іншої програми. Час роботи процесора може бути виражений кількістю тактів синхронізації для даної програми, помноженою на тривалість такту синхронізації.

Важливою та часто вживаною характеристикою є середня кількість тактів синхронізації процесора на одну команду CPI (clock cycles per instruction). При відомій кількості виконуваних команд в програмі ця характеристика дозволяє швидко оцінити час роботи процесора, необхідний для виконання заданої програми.

Досконалішою одиницею, яку можна використати для порівняння комп'ютерів, є продуктивність, тобто загальна кількість обчислювальної роботи, яку комп'ютер виконує за фіксований часовий інтервал. Якщо час виконання деякої програми позначити через T, то продуктивність P комп'ютера можна визначити наступним чином: $P = 1/T$. Тоді порівняння двох комп'ютерів X і Y можна виконати за наступними правилами: якщо $1/Tx > 1/Ty$ тобто $Ty > Tx$, то комп'ютер X є швидшим. В сучасних комп'ютерах продуктивність вимірюється в мільйонах операцій за секунду – MIPS. Таким чином, продуктивність може бути визначена як зворотна до часу виконання величина, причому швидші комп'ютери при цьому матимуть вищий рейтинг кількості операцій за одиницю часу. Позитивними сторонами кількості операцій за одиницю часу як одиниці оцінки продуктивності комп'ютера є те, що цю характеристику легко зрозуміти, особливо покупцю, і що швидший комп'ютер характеризується більшим числом операцій за одиницю часу. Проте використання цієї одиниці як метрики для порівняння натрапляє на дві проблеми. По-перше, вона залежить від набору команд процесора, що ускладнює порівняння комп'ютерів з різними системами команд. По-друге, навіть на одному і тому ж комп'ютері вона змінюється від програми до програми.

Вимірювання продуктивності комп'ютерів при вирішенні науково-технічних задач, в яких переважно використовується представлення даних в форматі з рухомою комою, завжди викликало особливий інтерес. Саме для таких обчислень вперше постало питання про вимірювання продуктивності, а за досягнутими показниками часто робилися висновки про загальний рівень розробок комп'ютерів. Зазвичай для науково-технічних завдань продуктивність комп'ютера оцінюється в кількості операцій з рухомою комою за секунду FLOPS (Floating Point Operations Per Second). В сьогоднішніх комп'ютерах це мільйони та мільярди операцій з рухомою комою за секунду – MFLOPS, GFLOPS.

Потрібно відзначити, що вищеназвані одиниці вимірювання – такт (або частота) синхронізації, середня кількість тактів на команду і продуктивність комп'ютера є взаємозв'язаними. Неможливо змінити жодну з них ізольовано від іншої, оскільки базові технології, використовувані для зміни кожної з цих характеристик, взаємозв'язані: частота синхронізації визначається технологією виготовлення апаратних засобів і функціональною організацією процесора; середня кількість тактів на команду залежить від функціональної організації і архітектури системи команд; а кількість виконуваних в програмі команд визначається архітектурою системи команд і технологією компіляторів. Коли порівнюються два комп'ютери, необхідно розглядати всі три компоненти, щоб зрозуміти відносну продуктивність.

1.2.3.2. Тестові програми для оцінки продуктивності

Важливість створення пакетів тестів, що базуються на реальних прикладних програмах широкого кола користувачів і забезпечують ефективну оцінку продуктивності комп'ютерів, була усвідомлена більшістю найбільших виробників комп'ютерних засобів, які в 1988 році заснували неприбуткову корпорацію SPEC (Standard Performance Evaluation Corporation). Основною метою цієї організації є розробка і підтримка стандартизованого набору спеціально підібраних на основі досвіду тестових програм для оцінки продуктивності новітніх поколінь високопродуктивних комп'ютерів.

Основним результатом роботи SPEC є набори тестів. Ці набори розробляються SPEC з використанням кодів, що поступають з різних джерел. SPEC працює над встановленням цих кодів на різні платформи, а також створює інструментальні засоби для формування з цих кодів осмислених робочих навантажень. В даний час є два базові набори тестів SPEC, що орієнтовані на інтенсивні розрахунки і вимірюють продуктивність процесора, пам'яті, а також ефективність генерації коду компілятором. Набір тестів CINT, що вимірює продуктивність процесора при обробці цілих чисел, складається з програм, написаних на мові C і вибраних з різних прикладних областей: теорія дробів, інтерпретатор мови LISP, розробка логічних схем, пакування текстових файлів, електронні таблиці і компіляція програм. Набір тестів CFP, що вимірює продуктивність процесора при обробці чисел з рухомою комою, складається з програм, також вибраних з різних прикладних областей: розробка аналогових схем, моделювання методом Монте-Карло, квантова хімія, оптика, робототехніка, квантова фізика, астрофізика, прогноз погоди і інші наукові та інженерні завдання. Частина програм з цього набору написана на мові C, інша частина – на Фортрані. Результати виконання кожного індивідуального тесту з цих двох наборів виражаються відношенням часу виконання однієї копії тесту на комп'ютері, що тестується, до часу її виконання на деякому еталонному комп'ютері.

Іншим підходом до оцінки продуктивності комп'ютерів за допомогою тестів є розробка спеціального програмного забезпечення (синтетичних тестів), що дозволяє створювати різні робочі навантаження, відповідно до рівня системи, що тестується, і до вимог щодо її використання. Однією з незалежних організацій, яка здійснює оцінку продуктивності комп'ютерних систем за допомогою синтетичних тестів, є приватна компанія AIM Technology, яка була заснована в 1981 році. Компанія розробляє і поставляє програмне забезпечення для вимірювання продуктивності систем, а також надає послуги з тестування систем кінцевим користувачам і постачальникам комп'ютерних систем та мереж, які використовують промислові стандартні операційні системи, такі як UNIX і

OS/2. Це програмне забезпечення складається з двох основних частин: генератора тестових пакетів і сумішей навантажень прикладних завдань. При кожному запуску генератора можуть виконуватися будь-які окремі або всі доступні тести у будь-якому порядку і при будь-якій кількості проходжень, дозволяючи тим самим створювати для системи практично довільне робоче навантаження. Якщо деякі необхідні тести відсутні у складі генератора тестових пакетів, то вони можуть бути туди легко додані. На основі оцінки продуктивності системи вибираються різні рівні збільшення навантаження. Це дозволяє з достатньою достовірністю зробити висновок про можливості роботи системи при даному навантаженні або при зміні навантаження. При цьому важливим є підбір сумішей навантажень, які ділять на дві категорії: замовлені і стандартні. Замовлені суміші створюють для точного моделювання особливостей середовища кінцевого користувача або постачальника устаткування. Стандартні суміші є звичайним середовищем прикладних завдань. До їх складу входять, зокрема, суміші для робочих станцій, геоінформаційних систем, ділових застосувань, керування базою даних та інші.

Важливим є створення тестів та оцінка продуктивності комп'ютерів на задачах з області їх конкретного застосування. Так, наприклад, фірма BDTI надає послуги з питань аналізу продуктивності комп'ютерів, призначених для вирішення задач цифрової обробки сигналів. З розширенням використання комп'ютерів при обробці транзакцій (комерційний обмін товарами, послугами або грошима) все важливішим стає забезпечення можливості їх коректного порівняння між собою. З цією метою в 1988 році була створена Рада для оцінки продуктивності обробки транзакцій (TPC – Transaction Processing Performance Council), яка є неприбутковою організацією. TPC публікує специфікації тестових пакетів, які регулюють питання, пов'язані з роботою тестів. Ці специфікації гарантують, що покупці мають об'єктивні значення даних для порівняння продуктивності різних комп'ютерних систем.

1.2.4. Організація зв'язків між функціональними вузлами комп'ютера

Для забезпечення роботи комп'ютера його функціональні вузли повинні бути відповідним чином з'єднані. Обмін інформацією між вузлами комп'ютера проводиться по шинах, до складу яких входять n ліній, де n – розрядність інформаційного слова. Є три головних типи шин: шина даних, шина керування (або шина команд) та шина адрес, які використовуються відповідно для пересилання даних, команд та їх адрес.

Три найуживаніших структури комп'ютера, які відрізняються організацією зв'язку між його функціональними вузлами, розглянуто нижче.

В двошинній структурі комп'ютера з обміном через процесор (рис. 1.6) введення інформації з пристрою введення до основної пам'яті, та виведення інформації з основної пам'яті до пристрою виведення здійснюється через процесор.

Рис. 1.6. Двошинна структура комп'ютера з обміном через процесор

Недоліком цієї структури є те, що процесор витрачає свій час на виконання операцій введення та виведення інформації, безпосередньо працюючи з пристроями введення-виведення, які в більшості випадків є значно повільнішими порівняно з ним, що знижує ефективність роботи комп'ютера.

В двошинній структурі комп'ютера з обміном через пам'ять (рис. 1.7) процесор звільнений від організації операцій введення-виведення інформації.

Рис. 1.7. Двошинна структура комп'ютера з обміном через пам'ять

Оскільки основна пам'ять не має засобів керування, для організації введення-виведення тут використовується спеціальний процесор введення-виведення (часто його називають каналом або периферійним процесором), який отримує завдання на виконання операцій введення-виведення від процесора (в цьому випадку останній називають центральним процесором), та керує взаємодією пам'яті та пристроїв введення-виведення. Дана структура була використана при побудові багатьох комп'ютерів, зокрема комп'ютерів серій IBM/360 та IBM/370 фірми IBM.

В одношинній структурі комп'ютера (рис. 1.8) всі його функціональні вузли з'єднані через одну спільну шину. Шина розподіляється між ними в часі, тому одночасно посилати дані на шину може тільки один вузол.

Рис. 1.8. Одношинна структура комп'ютера

Для узгодження швидкодії повільних і швидких вузлів комп'ютера в одношинній структурі на виході кожного повільного вузла (в даному випадку пристрої введення та виведення) використовують буферну пам'ять, яка дозволяє швидкий обмін блоками даних.

Кожний пристрій введення та виведення складається з двох частин – контролера та самого пристрою, наприклад, накопичувача на диску. Контролер є спеціалізованим комп'ютером, який керує своїм пристроєм та його доступом до спільної шини.

Для керування розподілом шини між підключеними до неї пристроями використовується арбітр шини. При одночасному поступленні кількох запитів арбітр шини вирі-

шує, чия черга є першою, і підключає до шини відповідний пристрій. Зазвичай перевага віддається пристроям введення-виведення, оскільки роботу дисків та інших пристроїв з рухомою механікою не можна переривати, щоб не втратити інформації.

Наявність спільної шини суттєво спрощує реалізацію комп'ютера та дозволяє легко замінювати його склад. Тому така структура отримала широке розповсюдження. Вона була покладена в основу зокрема комп'ютерів PDP-8 та PDP-11 фірми Digital Equipment Corporation, СМ-4 Київського заводу управляючих машин, та знайшла широке застосування в мікропроцесорних системах.

Для забезпечення одночасного обміну інформацією між різними пристроями в сучасних комп'ютерах використовується багатошинна організація.

1.3. Архітектура комп'ютера

1.3.1. Поняття архітектури комп'ютера

Вперше означення терміну "архітектура комп'ютера" було зроблене в 1964 році розробниками комп'ютера IBM 360 Г. Амдалем та його колегами. Архітектура комп'ютера з їх точки зору – це його структура і поведінка як їх бачить програміст на асемблерній мові. Вона включає наступне: формати даних і команд, методи адресації, систему команд, а також загальну організацію процесора, основної пам'яті і пристроїв введення-виведення. Пізніше А. Пейджез з тієї ж фірми запропонував розуміти під архітектурою комп'ютера інтерфейс між його апаратним та програмним забезпеченням.

Як відомо, в комп'ютері використовується двійкове представлення команд. При написанні програми крім двійкової можуть використовуватись і інші форми представлення команд: вісімкова, шістнадцяткова, символьна (мнемонічна). Використання вісімкового і шістнадцяткового кодування дозволяє скоротити записи і спростити роботу програміста. Мнемонічне кодування спрощує процес написання, читання і відлагодження програми. Основний принцип такого кодування – кожна команда представляється 3-х або 4-х буквеним символом, який показує назву команди. Деякі приклади мнемонічного кодування: add – додати, sub – відняти, load – зчитати дані з пам'яті, store – записати дані в пам'ять. Операнди також представляються символічно. Наприклад, команда add R, Y означає додавання значення вмісту комірки пам'яті Y до вмісту регістра R. Зауважимо, що операція виконується над вмістом, а не над адресою комірки пам'яті та регістра. Таким чином, з'являється можливість написання машинних програм в символічній формі. Повний набір символічних назв і правила їх використання утворюють мову програмування, відому як асемблерна мова. Запис деякої програми на асемблерній мові представляє собою символічний запис цієї ж програми, написаної на внутрішній мові комп'ютера, тобто в формі послідовності команд, представлених в двійкових кодах. Тому і з'явилось визначення архітектури комп'ютера як інтерфейсу між його апаратним та програмним забезпеченням.

З наведених вище означень можна зробити висновок про існування множини можливих варіантів архітектури комп'ютера. Розглянемо детальніше складові поняття архітектури з тим, щоб визначити ступінь зміни їх характеристик.

Під даними розуміються числа, представлені в деякій системі числення. В сучасних комп'ютерах дані в більшості випадків представлені в позиційній двійковій системі чис-

лення, яка практично витіснила інші способи представлення чисел, наприклад, представлення в позиційній десятковій системі числення, в системі залишкових класів тощо. В комп'ютерах є три класи даних: вхідні, проміжні та вихідні, які можна охарактеризувати наступною множиною параметрів: кількістю N даних, їх розрядністю n (n=1,2,3,...), способом кодування M (прямий, обернений чи доповняльний код), формою представлення S (фіксована чи рухома кома), форматом T (місце розміщення знаку, ціле чи дробове число, місце розміщення коми, розрядність порядку, основа порядку, використання знаку порядку чи зміщення, розрядність мантиси, чи використовується нормалізація мантиси). Таким чином, залежно від значення параметрів n, M, S, T можна виділити відповідну архітектуру комп'ютера. Наприклад, це може бути комп'ютер для виконання операцій над 16-розрядними дробовими двійковими даними в форматі з фіксованою комою в доповняльному коді, або це може бути комп'ютер для виконання операцій над 32-розрядними двійковими даними в форматі з рухомою комою.

Команда в комп'ютері зберігається в двійковій формі. Вона вказує тип операції, яка має бути виконаною, адреси операндів, над якими виконується операція, та адреси розміщення результатів виконання операції. Відповідно до цього команда складається з двох частин: коду операції та адресної частини. Поле коду операції займає k розрядів. Ним може бути закодовано до $N = 2^k$ різних операцій. Поле адреси (адресна частина) займає m розрядів. В ньому знаходяться адреси операндів. Кожна адреса займає m_i розрядів, де i – номер адреси (i=1, 2,...l), l – кількість адресних полів. Кожною адресою можна адресувати пам'ять ємністю 2^{m_i} слів. Таким чином, залежно від формату команди комп'ютери можна поділити на: комп'ютери з постійним та змінним форматом команди, комп'ютери з одноадресними, двоадресними та триадресними командами, комп'ютери з вузьким та широким форматом команди (залежно від кількості полів коду операції та адресних полів) і т. д.

Як вказано вище, крім коду операції до складу команди входить адресна частина. Цією частиною визначається місце знаходження даних, над якими виконується операція, задана кодом операції. Даних може бути декілька, і, крім того, вони можуть знаходитися в основній пам'яті, в регістрах процесора чи в запам'ятовуючих елементах інших вузлів комп'ютера. Тому і формати команд в цих випадках будуть різними. Можна здійснити класифікацію архітектури комп'ютера за типом адресованої пам'яті. Залежно від того, який тип пам'яті адресується, розрізняють наступні типи архітектур комп'ютера: стекова, акумуляторна, на основі регістрів загального користування.

При розробці комп'ютера необхідно враховувати багато особливостей вибору системи команд. З одного боку, система команд повинна бути функціонально повною, тобто комп'ютер повинен забезпечувати виконання всіх заданих функцій. З іншого боку, система команд має бути ортогональною, тобто не повинна бути надлишковою. Для нового комп'ютера, який є розширенням відповідної серії, першочерговою є сумісність – програми, які виконуються на одному комп'ютері, повинні виконуватись і на іншому комп'ютері. Є багато рівнів повноти системи команд. Теоретично система команд комп'ютера може включати лише одну команду. Однак програми на базі простих операцій є дуже складними. Існує фундаментальний зв'язок між простотою комп'ютера і складністю програми. В реальних комп'ютерах застосовується система команд, до складу якої входить широкий спектр команд обробки даних, переміщення даних, передачі керування та введення-

виведення. За складом системи команд комп'ютери можуть бути поділені на наступні типи: комп'ютери з складною (комплексною) системою команд, комп'ютери з простою (спрощеною) системою команд, комп'ютери з доповненою системою команд, комп'ютери з орієнтованою (спеціалізованою) системою команд.

Значна кількість типів архітектури комп'ютера може бути виділена залежно від організації його вузлів, а саме процесора, пам'яті і пристроїв введення-виведення. Наприклад, це можуть бути комп'ютери з паралельною та конвеєрною обробкою даних, з ієрархічною та лінійною пам'яттю і т. д.

Архітектура комп'ютера має визначальний вплив на його споживчі характеристики: коло вирішуваних на комп'ютері задач, продуктивність, ємність основної пам'яті, ємність зовнішньої пам'яті, вартість, організація технічного обслуговування, надійність і т. д.

Для вибору кращої з множини можливих варіантів архітектури комп'ютера потрібно знати зв'язок між архітектурою комп'ютера та його характеристиками, тобто як ті чи інші структурні особливості та організація роботи комп'ютера і його вузлів зв'язані з можливостями, які надаються користувачу, які є альтернативи при створенні комп'ютера і за якими критеріями повинні прийматися ті чи інші проектні рішення, як зв'язані між собою характеристики окремих пристроїв комп'ютера і який вплив вони мають на загальні його характеристики. А це саме ті питання, які є предметом даної книги.

1.3.2. Архітектурні принципи Джона фон Неймана

Описана в 1946 році Джоном фон Нейманом архітектура комп'ютера дістала назву його імені. Оскільки це був опис реалізованого Преспером Екертом та Джоном Моучлі універсального комп'ютера ENIAC, створеного в Принстонському університеті, часом її іще називають принстонською. Головні особливості архітектури комп'ютера Джона фон Неймана:

> Інформація в комп'ютері ділиться на команди і дані.

> Команди вказують комп'ютеру, які дії і над якими операндами виконувати.

> Послідовність команд, за якою виконується алгоритм вирішення задачі, називають програмою.

> Весь набір виконуваних комп'ютером команд називають системою команд комп'ютера.

> Дані – це числа і закодовані символи, які використовуються командами як операнди. Одні команди для інших також можуть бути операндами.

> Команди і дані представлено двійковим кодом.

> Немає відмінностей в представленні команд і даних. Наприклад, двійкове число 100011011100 може бути як командою, так і даним.

> Команди і дані зберігаються в одній пам'яті.

> Команди і дані зберігаються в пам'яті за відповідними адресами.

> Пам'ять має довільну адресацію, тобто в кожному такті можна звернутися до довільної її комірки.

> Пам'ять є лінійною. Її адресу кодують двійковим кодом, починаючи від молодшої, всі розряди якої рівні нулю (00...0), до старшої, всі розряди якої рівні одиниці (11...1).

В основу роботи цього комп'ютера покладено принцип програмного керування. Тобто функціонування цього комп'ютера здійснюється потактово за вказівкою команд про-

грами. Програма разом з даними, які підлягають обробці, спочатку записується до основної пам'яті. В кожному такті, залежно від типу архітектури комп'ютера, виконується команда або частина команди – мікрокоманда. Команда вказує тип виконуваної операції та місце розміщення операндів і результату операції. Для виконання однієї команди необхідно провести наступні дії: записати до програмного лічильника адресу команди, зчитати із основної пам'яті команду за вмістом програмного лічильника, провести дешифрування команди з метою її розпізнання, визначити адреси комірок пам'яті, в яких знаходяться операнди та до яких мають бути записані результати, зчитати ці операнди з пам'яті та подати в арифметико-логічний пристрій для опрацювання, виконати операцію над операндами та записати результати до основної пам'яті. Таким чином проводиться виконання всіх команд програми. По закінченню в основній пам'яті будуть знаходитися результати виконання програми.

Як буде видно далі, ці особливості характерні для більшості сучасних комп'ютерів.

1.3.3. Ненейманівські архітектури комп'ютерів

До цього часу більшість універсальних комп'ютерів будують за принципами архітектури Джона фон Неймана. Але недоліки цієї архітектури, пов'язані з закладеним в ній послідовним характером організації обчислень, ставлять перепони в пошуку шляхів побудови швидких комп'ютерних систем. Тому крім архітектури Джона фон Неймана за час існування комп'ютерної техніки було створено цілий ряд інших архітектур. До них, зокрема, належать гарвардська архітектура, асоціативна машина, машина потоків даних, редукційна машина. Було запропоновано покласти в якості обчислювальної парадигми нейронні мережі, в яких використана ідея з моделей мозку та генетичні алгоритми, в яких застосовані ідеї з біології та еволюції архітектури комп'ютерних мереж. Останнім часом багато уваги приділяється квантовим комп'ютерам, елементи яких працюють за законами квантової механіки, біологічним комп'ютерам, а також паралельним комп'ютерам, оскільки практично всі сьогоднішні комп'ютери є паралельними.

Коротко виділимо основні риси деяких з названих архітектур.

Гарвардська архітектура вперше була реалізована Ховардом Айкеном в комп'ютері Марк-1 в Гарварді. Вона передбачає розділення пам'яті на пам'ять даних і пам'ять команд. Тим самим розділяються шини передачі керуючої і оброблюваної інформації (рис. 1.9). При цьому підвищується продуктивність комп'ютера за рахунок суміщення в часі пересилання та обробки даних і команд.

Рис. 1.9. Ядро комп'ютера гарвардської архітектури, утворене процесором та пам'яттю команд і даних

Необхідність використання гарвардської архітектури можна пояснити так. Ядро комп'ютера Джона фон Неймана складається з процесора та основної пам'яті. Бажано, аби обидві компоненти ядра не пригальмовували одна одну, тобто працювали із рівною швидкодією. На практиці вузол пам'яті є значно (на порядок) повільнішим від процесора і цей розрив у швидкодії з прогресом інтегральних технологій лише зростає. Зменшити розрив можна структурними методами, збільшуючи розрядність інформаційного слова пам'яті. Саме цей підхід реалізує гарвардська архітектура з двома запам'ятовувальними пристроями. Зрозуміло, що тут паралельно виконуються операції вибирання команд програми, з одного боку, а з другого – вибирання та запис кодів даних і результатів обчислень.

Дуальна пристонсько-гарвардська архітектура. Швидкі комп'ютери гарвардської архітектури є складнішими щодо програмування порівняно з комп'ютерами принстонської архітектури. Зрозуміло, що бажано створити комп'ютер з дуальною архітектурою, яка водночас запозичує нову якість – швидкодію від гарвардської архітектури та стандартну парадигму розробки програм від принстонської архітектури. Злиття двох архітектур виконують на рівні кеш пам'яті шляхом її поділу на кеш даних та кеш команд (рис. 1.10). Злиттям архітектур програмісту надано зручність програмних технологій принстонської архітектури, а з боку процесора реалізовано гарвардську архітектуру, в результаті чого він значно менше пригальмовується з боку основної пам'яті.

Рис. 1.10. Злиття архітектур через розділену на дві частини кеш пам'ять першого рівня

Асоціативна машина передбачає маніпуляції з даними не відповідно до їх адрес, як це є в машині Джона фон Неймана, а відповідно до значення цих даних або їх частин. Базовими тут є операції пошуку і порівняння. Основою асоціативної архітектури є асоціативна пам'ять, яка забезпечує одночасний доступ до багатьох даних, в яких співпадають значення відповідних розрядів. Тим самим за рахунок високої паралельності обробки досягається висока продуктивність на класі операцій, для виконання яких ця машина є ефективною (зокрема, логічні операції, операції пошуку та сортування). Асоціативна машина є складовою практично кожного сучасного комп'ютера.

Машина потоків даних. Керування обчислювальним процесом в машині потоків даних здійснюється даними за їх готовністю до обробки. Кожне дане в такій машині має спеціальні ознаки. За цими ознаками пристрій керування знаходить дані, які готові до обробки, і передає їх в АЛП для виконання відповідних операцій. Тим самим тут за рахунок можливого паралельного аналізу та обробки даних досягається гранично ви-

сока продуктивність. Принципи машини потоків даних використовуються в багатьох сучасних високопродуктивних комп'ютерах.

Паралельні комп'ютерні системи. Перші паралельні комп'ютерні системи, до складу яких входило лише два процесори, були побудовані в кінці 60-х років минулого століття. В 70-х роках такі системи мали в своєму складі до 64-х процесорів, в 80-х роках – до 1000, а в кінці 90-х років фірма IBM анонсувала конструкцію суперкомп'ютера з паралельною архітектурою, який включав понад мільйон процесорів і на даний час є найпродуктивнішим у світі. Паралельна обробка інформації є ключовим напрямком побудови високопродуктивних комп'ютерних систем. Однак і паралельні комп'ютерні системи мають обмеження. По-перше, зі збільшенням кількості процесорів ускладнюється задача розподілу завдань між процесорами. Для її вирішення використовуються додаткові процесори, кількість яких може значно перевищувати кількість процесорів, зайнятих безпосередньо виконанням алгоритму. По-друге, послідовна природа багатьох алгоритмів обмежує прискорення, якого можна досягти, використовуючи багатопроцесорну організацію.

1.4. Типи сучасних комп'ютерів

Аналіз сфер використання комп'ютерів показує, що можна виділити два основних напрямки їх використання. Перший напрямок – це підсилення інтелектуальних можливостей людини в досить широкому розумінні цього поняття. Мається на увазі прискорення обчислень, зберігання великих об'ємів інформації, швидкий пошук та відображення необхідної інформації і т. д. Комп'ютери цього напрямку повинні мати велику ємність пам'яті, потужне програмне забезпечення, розвинуті засоби взаємодії з людиною.

Другий напрямок – це використання комп'ютера як елемента електротехнічної системи, наприклад, системи керування, інформаційно-вимірювальної системи, системи передачі даних і т. д. На рис. 1.11 показано приклад використання комп'ютера в складі системи керування. Тут параметри об'єкта керування знімаються давачами, перетворюються в цифрову форму за допомогою аналого-цифрових перетворювачів АЦП та подаються в комп'ютер. В комп'ютері здійснюється оцінка стану об'єкта керування та виробляється послідовність керуючих кодів, які в цифро-аналогових перетворювачах ЦАП перетворюються в аналогову форму та подаються на органи керування з метою забезпечення утримання параметрів об'єкта керування в заданих межах.

Рис. 1.11. Комп'ютер в системі керування

Основна вимога до комп'ютерів даного напрямку – забезпечення необхідного часу реакції на події в системі, проведення обробки вхідних даних в темпі їх поступлення, тобто в реальному масштабі часу, і видача результатів відповідно до часових вимог електротехнічної системи. Комп'ютери цього напрямку повинні мати швидкісні зовнішні інтерфейси, володіти високою продуктивністю та надійністю роботи. Часто ці комп'ютери є вбудованими в мобільні засоби, тому додатково вимагається, щоб вони мали малі габарити та низьку споживану потужність. Вбудовані комп'ютери широко використовуються в побутовій електронній апаратурі, зокрема в відео та аудіо апаратурі, фотоапаратах, іграшках, побутових пристроях, мобільних телефонах і т.д.

Залежно від сфери застосування та технічних характеристик, в першу чергу габаритів, продуктивності, ємності пам'яті та ціни розрізняють наступні типи комп'ютерів:

> персональні комп'ютери (personal computers) – комп'ютери, орієнтовані на користування однією особою;

> робочі станції (workstations) – комп'ютери, орієнтовані на вирішення інженерних задач, в першу чергу в складі локальних комп'ютерних мереж;

> багатотермінальні системи – набір стандартних терміналів, підключених до сервера;

> сервери – центральні комп'ютери для побудови інформаційних систем;

> мейнфрейми (mainframes) – великі універсальні комп'ютерні системи;

> кластерні комп'ютерні системи – об'єднання комп'ютерів, що сприймається операційною системою, системним програмним забезпеченням, прикладними програмами і користувачами як єдине ціле;

> суперкомп'ютери – найпотужніші на даний час комп'ютери;

> мікроконтролери – комп'ютери на кристалі, призначені для керування електронними пристроями.

> спеціалізовані комп'ютери. Вони орієнтовані на вирішення задач, котрі неможливо або недоцільно виконувати на універсальних комп'ютерах.

Розглянемо названі комп'ютери детальніше.

1.4.1. Персональні комп'ютери

Персональні комп'ютери (ПК) з'явилися в результаті еволюції мінікомп'ютерів при переході елементної бази з малим і середнім ступенем інтеграції на великі і надвеликі інтегральні схеми. ПК, завдяки низькій вартості, дуже швидко завоювали тверді позиції на комп'ютерному ринку і створили передумови для розробки нових програмних засобів, що орієнтувалися на кінцевого користувача. Це, передусім, дружні по відношенню до користувача інтерфейси, а також проблемно-орієнтовані середовища і інструментальні засоби для автоматизації розробки прикладних програм.

Персональні комп'ютери класифікуються за їх розмірами та конструктивним виконанням наступним чином:

> Настільні комп'ютери (desktop computers). До складу настільного комп'ютера входить системний блок (в якому розміщена материнська плата, центральний процесор, основна пам'ять, карта розширення, блок живлення і т. д.), дисплей, клавіатура, мишка. В системний блок також вбудовані драйвер оптичного диску і зовнішня дискова пам'ять. Настільні комп'ютери призначені для офісного використання.

> Лаптопи чи ноутбуки (laptop or notebooks). Це близькі за характеристиками до настільного ПК, але конструктивно виконані в придатному для перенесення виконанні.

➢ Персональні цифрові асистенти (Personal digital assistants). Це кишенькові ПК, спеціально створені як персональні асистенти людини. Вони надають наступний сервіс: годинник, комп'ютерні ігри, доступ до мережі Інтернет, електронна пошта, записна книжка, адресна книжка, мобільний телефон, медіа плеєр та інше.

➢ Смартфони. Це мобільні телефони, які мають вбудовану операційну систему та можливості вище описаних персональних цифрових асистентів.

➢ Портативні комп'ютери (Portable computers). Це ПК типу настільних, але виконані в придатному для перенесення та роботи в не офісних умовах конструктивному виконанні.

➢ Переносимі комп'ютери (Wearable computers). Це комп'ютери, які надають інформаційні послуги людині під час її руху в навколишньому середовищі. До переносимих ПК належать, зокрема, комп'ютери для моніторингу стану людини.

На рис. 1.12 наведено зовнішній вигляд ряду серійних персональних комп'ютерів: персональний цифровий асистент pulmone Tursten T5 (рис. 1.12a), мобільний ПК французької армії (рис. 1.12b), переносимий комп'ютер VAIO VGN-UX50 фірми Sony (рис. 1.12c), планшетний ПК фірми Hewlett Pascard (рис.1.12d).

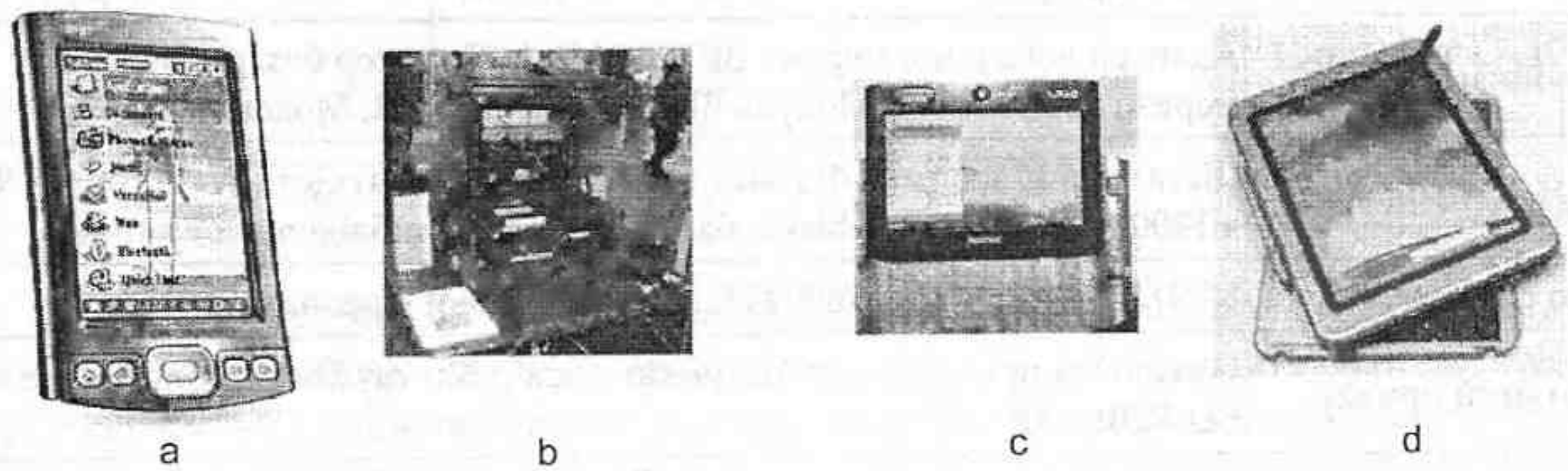


Рис. 1.12. Зовнішній вигляд ряду серійних персональних комп'ютерів

Потрібно відзначити, що названі вище типи персональних комп'ютерів в даний час активно розвиваються та вдосконалюються, тому як термінологія так і сфера їх застосування ще не є чітко встановленими.

В якості прикладу розглянемо значення характеристик реального персонального комп'ютера. На рис. 1.13 наведено зовнішній вигляд, а в табл. 1.1 технічні характеристики, взяті з однієї із типових реклам комп'ютера, а саме мультимедійного ноутбука VAIO AR11 RS. Зрозуміло, що без знання наведеної в рекламі термінології важко вибрати та купити комп'ютер з потрібними характеристиками.

Рис. 1.13. Зовнішній вигляд комп'ютера Sony VAIO AR11 RS

Таблиця 1.1

Процесор	Intel Centrino Core Duo T2500: два ядра; тактова частота 2000 МГц; FSB 667 МГц
Пам'ять	Встановлено 1024 Мбайтів пам'яті DDR 2-533
Кеш пам'ять	2x32КВ L1 cache, 2048КБ L2 cache
Жорсткий диск	Два вінчестери ємністю по 100 ГБ кожний. Швидкість обертання 5400 об/хв. 8 МБ cache
Порти введення-виведення	3 порти USB 2.0; по одному порту: IEEE 1394 (i.LINK, FireWire); IrDA (інфрачервоний порт); PCMCIA Type II; ExpressCard|54; Secure Digital (SD/MMC); Memory Stick (MS/MS Pro/MS Duo/MS Pro Duo) – адаптер до карт пам'яті; VGA (вихід на зовнішній монітор); HDMI (High Definition Multimedia Interface); S-Video out (аналоговий відеовихід); S-Video in (аналоговий відеовхід); RJ-11 (модем); RJ-45 (мережа); оптичний S/PDIF; вихід на навушники; лінійний вхід; вхід для мікрофона
Комунікації	Адаптер локальної мережі 10/100 Мбіт/с; Адаптер бездротової мережі 802.11 a/b/g; Модуль Bluetooth 2.0 + EDR; Модем V.90, 56k
Екран	Діагональ 17 дюймів. Формат 16:9. Роздільна здатність WUXGA (1920 x 1200). Технологія X-black. Багатошарове антиблікове покриття
Відео система	NVIDIA GoForce Go 7600 GT. 256 МБ власної відеопам'яті
Оптичний привід	Оптичний привід для двошарових дисків Blu-ray Disc. Запис дисків BD-R/BD-RE
Аудіо	Intel High Definition Audio. Вбудовані стерео динаміки та мікрофон
Пристрої введення	Клавіатура. Сенсорна панель (touchpad)
Додаткові пристрої	Вбудовані ТВ-тюнер та веб-камера VGA (640x480)
Електроживлення	Літій-іонна батарея (VGP-BPS2C). Час автономної роботи до 120 хвилин
Розміри	416x299,5x33,5 мм
Вага	3,8 кг
Гарантія	12 місяців
Операційна система	Microsoft Windows XP Media Center Edition
Ціна	$3000

Як бачимо, характеристики функціональних вузлів комп'ютера оцінюються в мільйонах, мільярдах, трильйонах чи трильярдах деяких одиниць, або в їх мільйонних, мільярдних, трильйонних та трильярдних частках. Для спрощення запису цих характеристик використовуються відповідні префікси. В таблиці 1.2 наведено префікси до одиниць вимірювання великорозмірних та малорозмірних характеристик функціональних вузлів комп'ютера кирилицею та латинськими буквами.

Таблиця 1.2

Назва префікса	Позначення префікса	Назва на латині	Позначення на латині	Значення префікса, виражене через степінь двох	Близьке значення, виражене через степінь десяти
Великорозмірні характеристики					
Кіло	К	Kilo	K	2^{10}	1 тисяча = 10^3
Мега	М	Mega	M	2^{20}	1 мільйон = 10^6
Гіга	Г	Giga	G	2^{30}	1 мільярд = 10^9
Тера	Т	Tera	T	2^{40}	1 трильйон = 10^{12}
Пета	П	Peta	P	2^{50}	1 трильярд = 10^{15}
Малорозмірні характеристики					
Мілі	м	Milli	m	2^{-10}	1 тисячна = 10^{-3}
Мікро	мк	Micro	μ	2^{-20}	1 мільйонна = 10^{-6}
Нано	н	Nano	n	2^{-30}	1 мільярдна = 10^{-9}
Піко	п	Pico	p	2^{-40}	1 трильйонна = 10^{-12}
Фенто	ф	Fento	f	2^{-50}	1 трильярдна = 10^{-15}

Наведені в таблиці 1.2 префікси використовуються для вираження відповідної додатної або від'ємної степені 10. Як буде видно далі, в комп'ютерах, які працюють в двійковій системі числення, значна частина характеристик вимірюється в одиницях, кратних степені 2. Але оскільки люди вже звикли до цих префіксів, їх же використовують і для вираження відповідної степені 2 чисел, близьких степені 10. Наприклад, ємність 1 КБ (кілобайт) означає 1024 байти, оскільки значення 210 = 1024 є близьким до 1000. В останньому стовпчику таблиці 1.2 наведено значення степені 10 для близьких значень степені 2.

Разом з тим, необхідно зауважити, що дискова пам'ять ємністю 1 ГБ має 1 мільярд байтів, а не 230 (приблизно 1.7 мільярда). Тому потрібно читати інструкцію користувача для того, щоб бути впевненим в точному значенні запису 1К, 1М, чи 1Г в кожному конкретному випадку.

В якості процесора комп'ютера Sony VAIO AR11 RS використано процесор Intel Centrino Core Duo T2500. Оскільки процесор є основним елементом комп'ютера, його продуктивність в значній мірі визначає продуктивність всього комп'ютера. Принципи побудови та організації роботи процесора та його функціональних вузлів будуть розглянуті в розділах 4-8 книги. Робота процесора синхронізується імпульсами, які генерує блок синхронізації комп'ютера та посилає їх до всіх його основних компонентів. Кількість вироблених цим блоком імпульсів за 1 секунду є його частотою, яка вимірюється в герцах. Сучасні комп'ютери працюють на мегагерцових та гігагерцових частотах, тобто їх блоки синхронізації генерують сотні мільйонів або й мільярди імпульсів за секунду. Процесор комп'ютера, представленого на рекламі, функціонує на частоті 2 ГГц. Однак це не обов'язково означає, що він здатний виконувати 2 мільярди команд за секунду, або, що еквівалентно, що кожна його команда виконується за 0,5 наносекунди. Як буде

показано в розділі 2, кожна команда комп'ютера вимагає фіксованої кількості тактів для свого виконання, причому деякі команди виконуються за один такт, але більшість команд виконуються за декілька тактів.

В другому рядку реклами розміщено "1024 Мбайтів пам'яті DDR 2-533)". Це основна пам'ять. Число 533 вказує частоту зчитування даних з цієї пам'яті на системну шину, по якій проводиться обмін інформацією в комп'ютері. Тут частота вимірюється в МГц. Основна пам'ять має ємність 1024 мегабайти (МБ). Питання про формати представлення даних в комп'ютері будуть розглянуті в розділі 3, а питання побудови та організації роботи пам'яті – в розділі 9. Додатково до ємності пам'яті реклама вказує і її тип DDR, скорочення від dynamic random access memory (динамічна пам'ять з довільним доступом), яка підтримує двоканальний режим.

Наступний рядок в рекламі "2x32КБ L1 cache, 2048КБ L2 cache" також описує тип пам'яті. Це кеш пам'ять, яка забезпечує швидкий доступ процесора до основної пам'яті. Представлений на рекламі комп'ютер має два рівні кеш пам'яті. Кеш пам'ять першого рівня (L1) є швидкою пам'яттю малої ємності, яка вбудована в кристал процесора і дозволяє прискорити доступ до часто використовуваних даних і команд. Кеш пам'ять другого рівня (L2) є набором швидких мікросхем пам'яті, розміщених між процесором і основною пам'яттю. При цьому кеш пам'ять має ємність, що вимірюється в кілобайтах (КБ), тобто значно меншу, аніж ємність основної пам'яті. В даному випадку кеш пам'ять першого рівня розділена на кеш пам'ять даних і кеш пам'ять команд ємністю по 32 КБ, кеш пам'ять другого рівня має ємність 2048 КБ.

Рекламований комп'ютер має також зовнішню дискову пам'ять ємністю 2x100 ГБ, яка має швидкість обертання дисків 5400 обертів за хвилину. Крім того, ця пам'ять має буфер ємністю 8 МБ. Швидкість обертання є лише одним з визначальних факторів загальної продуктивності диску. Важливими є також принципи його з'єднання з іншими елементами комп'ютера, тобто його інтерфейс. В рекламованому комп'ютері інтерфейс диску має назву EIDE, що є скороченням від enhanced integrated drive electronics (покращена електроніка інтегрованого дисководу). Інтерфейс EIDE має спеціальні вузли, які дозволяють підвищити швидкість обміну.

Далі в рекламі розміщені характеристики пристроїв введення-виведення та комунікацій, які будуть детально розглянуті в розділі 10. Тут показано 3 USB порти, один швидкісний послідовний порт, один інфрачервоний порт, адаптери для безпровідного зв'язку та багато інших портів. Через ці порти інформація поступає в комп'ютер та з комп'ютера. USB (universal serial bus) порт є популярною зовнішньою шиною, яка підтримує режими Plug-and-Play (тобто можливість автоматичної конфігурації пристрою) та hot plugging (можливість підключити та відключити пристрій під час роботи комп'ютера).

Також в рекламі є інформація про монітор. Він має наступні характеристики: діагональ рівна 17 дюймів, формат 16:9, роздільна здатність WUXGA (1920 x 1200), технологія X-black, багатошарове антиблікове покриття. Характеристики монітора впливають на зручність взаємодії користувача з комп'ютером. Монітор з високою роздільною здатністю забезпечує краще бачення та якісну графіку.

На рекламі є і інша потрібна інформація про комп'ютер, зокрема наведені характеристики відео та аудіо систем, оптичного приводу та інше.

Розглядаючи вузли та характеристики представленого на рекламі комп'ютера, ми не обговорювали питання організації його роботи та вплив цих характеристик на ефективність вирішення задач, що буде зроблено далі в цій книзі.

Обговорюючи принципи роботи комп'ютера, ми будемо показувати його вузли та елементи схематично, не завжди вникаючи в питання їх фізичної реалізації. Але для того, щоб мати загальну уяву про конструкцію комп'ютера, подивимось як виглядають компоненти сучасного персонального комп'ютера (рис. 1.14).

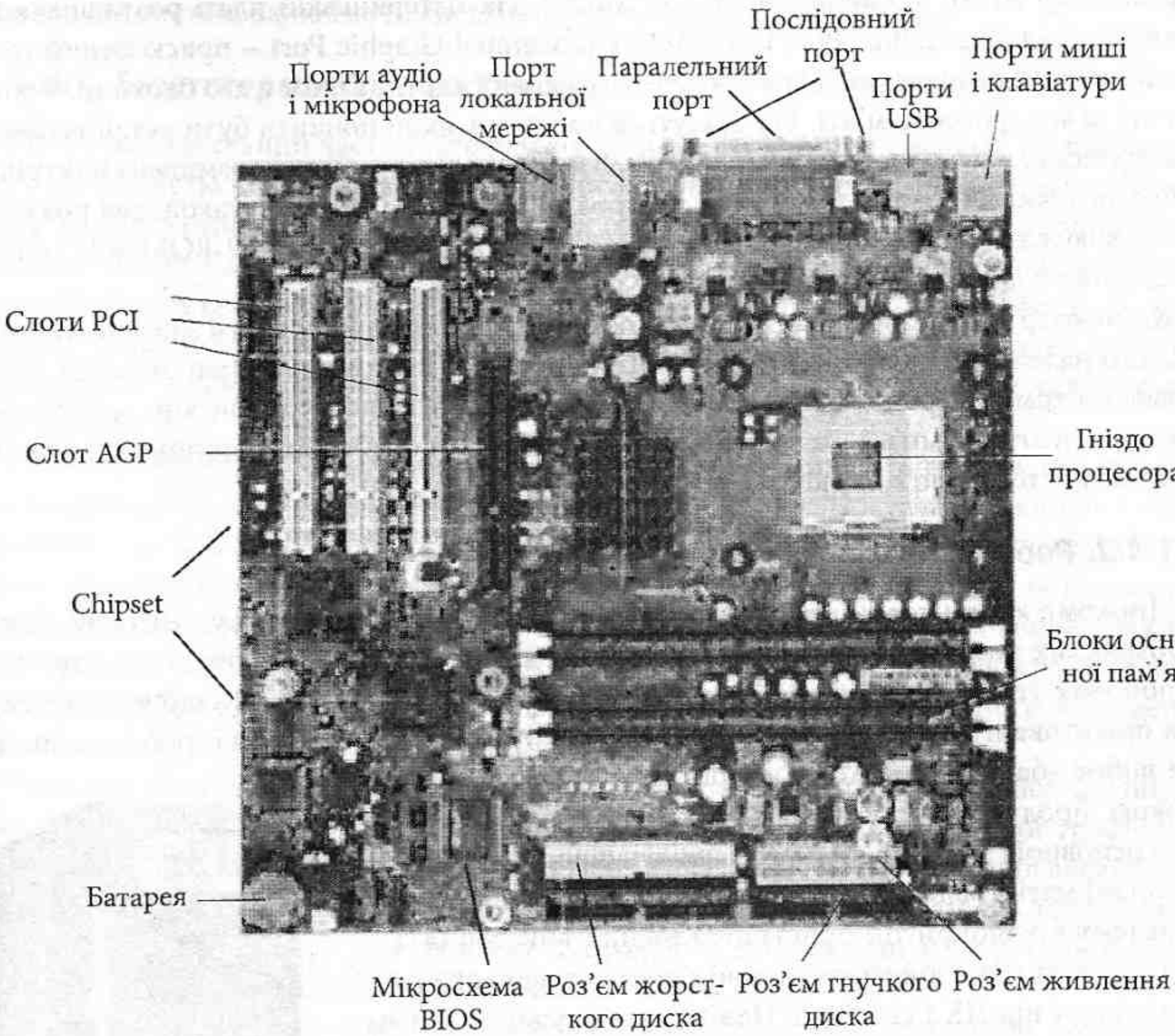


Рис. 1.14. Вигляд вузлів комп'ютера

Коли відкрити корпус персонального комп'ютера, можна побачити велику кількість різних елементів. Це, зокрема, блок живлення, який представляє собою металевий корпус з вбудованим вентилятором, різні типи дискової пам'яті, включаючи накопичувачі на жорсткому та гнучкому магнітних дисках, привід CD-ROM/RW чи DVD-ROM/RW. Всі компоненти комп'ютера, включаючи процесор та пам'ять, об'єднує системна плата. На рис. 1.14 показано системну плату Intel D850 з тлумаченням найважливіших компонентів.

Порти введення-виведення зверху плати забезпечують зв'язок комп'ютера з зовнішніми вузлами типу мікрофона та інших звукових пристроїв, миші, клавіатури, з

локальною обчислювальною мережею та іншими пристроями, які під'єднують через послідовний, паралельний та USB порти. Контролер введення-виведення, який входить до складу інших, крім процесора, пам'яті та відео, мікросхем системної плати (chipset), дозволяє всім з'єднаним пристроям функціонувати без конфліктів. Гнізда (слоти) для підключення до шини PCI (Peripheral Component Interconnect) дозволяють розширення плат, які належать до різних пристроїв, підключених до шини PCI. Це 32-розрядна локальна шина для пересилання даних між процесором та зовнішніми пристроями (диски, відеоадаптер тощо) з швидкістю до 132 Мбіт/с. На материнській платі розташовують зазвичай 3-4 гнізда шини PCI. Слот AGP (Accelerated Graphic Port – прискорений графічний порт) призначений для включення графічної карти. Також є два блоки основної пам'яті та контролер пам'яті. Тут відсутній процесор, який повинен бути встановлений в материнську плату, але є гніздо під нього. В нижньому лівому куті розміщена внутрішня батарея живлення для зберігання настройок BIOS. Ця плата має також два роз'єми для підключення жорсткого диска або приводу CD-ROM/RW чи DVD-ROM/RW, і один роз'єм для підключення гнучкого диска.

Комп'ютер продається з записаним в постійну пам'ять програмним забезпеченням, до якого належать POST (Power-On-Self-Test) та BIOS (Base Input/Output Systems). При включенні комп'ютера відбувається тестування пам'яті, клавіатури, дисків і деяких інших компонентів комп'ютера. Після цього викликається операційна система, яка завантажується та оживляє комп'ютер.

1.4.2. Робочі станції

Мінікомп'ютери стали прародичами й іншого напрямку розвитку сучасних 32 та 64-розрядних комп'ютерів, що сьогодні відомі як робочі станції. Початкова орієнтація робочих станцій на професійних користувачів (на відміну від ПК, що від початку були орієнтовані на споживача-непрофесіонала) призвела до того, що робочі станції – це добре збалансовані комп'ютерні системи, які, разом з високою продуктивністю, характеризуються великою ємністю основної і зовнішньої пам'яті, мають високошвидкісні внутрішні магістралі, високоякісну і швидкодіючу графічну підсистему і різноманітні пристрої введення-виведення. Ця властивість вигідно відрізняє робочі станції середнього і високого класу від ПК і сьогодні. Навіть найпотужніші ПК не в стані задовольнити зростаючі потреби інженерних задач через наявність в їхній архітектурі ряду вузьких місць.

В якості прикладу розглянемо характеристики робочої станції Sun Ultra 45 Workstation фірми Sun Microsystems, зовнішній вигляд якої наведено на рис. 1.15. До її складу входять два процесори UltraSPARC IIIi з частотою 1.6GHz, кеш пам'ять другого рівня ємністю 1MB, Dual Gigabit Ethernet, до 16 GB DDR1 ECC пам'ять, два графічних прискорювачі для двовимірної та тривимірної графіки.

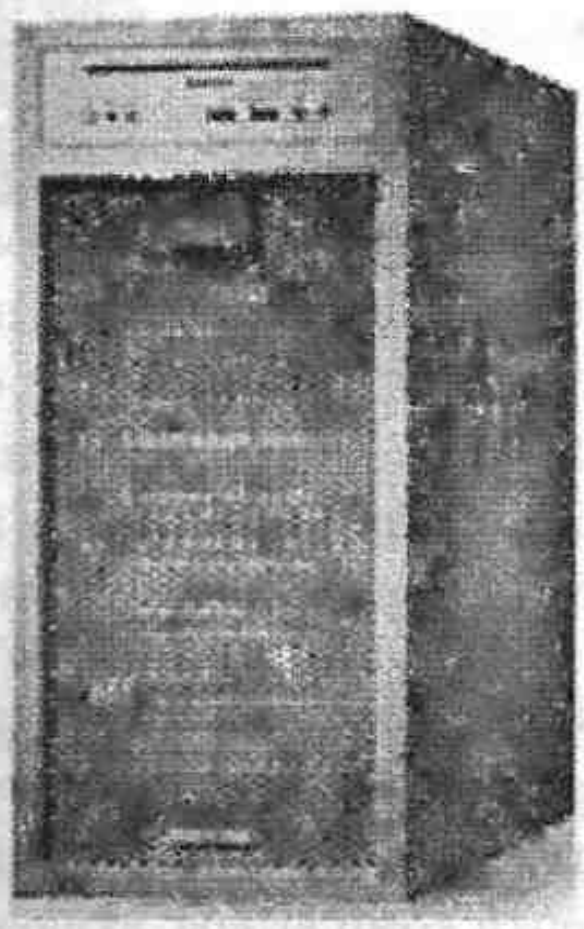


Рис. 1.15. Зовнішній вигляд робочої станції Sun Ultra 45 Workstation фірми Sun Microsystems

Виробники робочих станцій швидко відреагували на потребу в маловартісних моделях для ринку комерційних застосувань. Потреба в високій потужності на робочому столі навела компанії Sun Microsystems і Hewlett-Packard, які є лідерами на ринку робочих станцій, на ідею орієнтувати свою продукцію на комерційні застосування. І хоча значна частина систем цих компаній все ще орієнтована на технічні та наукові застосування, спостерігається безпрецедентне зростання продажу їх продукції для робіт комерційного застосування, що вимагають все більшої потужності для реалізації складних мережних прикладних систем, включаючи системи мультимедіа.

1.4.3. Багатотермінальні системи

ПК та робочі станції часто застосовуються в якості дорогих дисплеїв і в цьому випадку не повністю використовується їх обчислювальна потужність. Разом з тим, багато користувачів терміналів хотіли б покращити їхні графічні характеристики та мати можливість роботи в багатовіконній системі. Ці проблеми були вирішені шляхом створення багатотермінальних систем, які є набором стандартних терміналів, підключених до сервера. Як тільки стали доступними потужні графічні робочі станції, з'явилася тенденція застосування "підлеглих" терміналів, що використають робочу станцію в якості локального сервера.

На комп'ютерному ринку багатотермінальні системи займають проміжне положення між персональними комп'ютерами і робочими станціями. Постачальники терміналів заявляють, що їхні вироби ефективніші в вартісному вираженні, ніж робочі станції високого цінового класу, і пропонують збільшений рівень продуктивності у порівнянні з персональними комп'ютерами, що робить цю технологію доступною для широкого кола користувачів. Вартість терміналів складає біля половини вартості близького за конфігурацією ПК без зовнішньої пам'яті і приблизно чверть вартості повністю оснащеної робочої станції.

Типовий термінал включає наступні елементи: екран високої роздільної здатності; головний процесор, який підтримує двопроцесорну архітектуру; окремий графічний співпроцесор, що забезпечує швидше малювання на екрані і прокручування екрану; базові системні програми; програмне забезпечення сервера; локальну пам'ять для дисплею та мережного інтерфейсу, що підтримує протокол TCP/IP та інші мережні протоколи; порти для підключення клавіатури і миші.

Термінали відрізняються від ПК і робочих станцій не тільки тим, що не виконують функції звичайної локальної обробки. Робота терміналів залежить від головної системи, до якої вони підключені через мережу. Для того, щоб термінал міг працювати, користувачі повинні встановити програмне забезпечення багатовіконного сервера на головному процесорі, що виконує прикладну задачу. Локальна обчислювальна потужність терміналу зазвичай використовується для виконання програм обробки зображень, а не прикладних програм, які виконуються на головному процесорі. Термінал може відображати на одному і тому ж екрані декілька задач. Користувач може змінювати розміри вікон, їхнє місцезнаходження і маніпулювати ними в будь-якому місці екрана.

1.4.4. Сервери

Прикладні комерційні та бізнесові системи, розраховані на багато користувачів, включаючи системи керування базами даних і обробки транзакцій, великі видавничі системи, мережні системи і системи обслуговування комунікацій, системи розробки програмного забезпечення і обробки зображень, вимагають переходу до моделі обчислень "клієнт–сервер" і розподіленої обробки. В розподіленій моделі "клієнт–сервер" частину роботи виконує сервер, а частину – комп'ютер користувача (в загальному випадку частини сервера і користувача можуть виконуватись і на одному комп'ютері). Існує декілька типів серверів для різних застосувань: файловий сервер, сервер бази даних, принт-сервер, обчислювальний сервер, сервер застосувань. Таким чином, тип сервера визначається ресурсом, яким він володіє (файлова система, база даних, принтери, процесори або прикладні пакети програм).

З іншого боку існує класифікація серверів за масштабом мережі, в якій вони використовуються: сервер робочої групи, сервер відділу або сервер підприємства (корпоративний сервер). Ця класифікація надто умовна. Наприклад, розмір групи може змінюватися в діапазоні від декількох людей до декількох сотень людей, а сервер відділу може обслуговувати від 20 до 150 користувачів. Очевидно, залежно від числа користувачів і характеру вирішуваних ними завдань, вимоги до складу обладнання і програмного забезпечення сервера, до його надійності та продуктивності суттєво відрізняються.

Файлові сервери невеликих робочих груп (не більше 20–30 людей) простіше всього реалізуються на платформі персональних комп'ютерів і програмному забезпеченні Novell NetWare. Файл-сервер в даному випадку виконує роль центрального сховища даних. Типовими до складу невеликих файл-серверів входять процесор, основна та зовнішня пам'ять, а також адаптер Ethernet. До складу таких серверів часто включаються дисковод гнучких дисків і дисковод компакт-дисків. Графіка для більшості серверів несуттєва, тому достатньо мати звичайний монохромний монітор з невисокою роздільною здатністю. Бажано застосувати пристрій безперебійного живлення.

Для файл-серверів загального доступу, з якими водночас можуть працювати декілька десятків, а то і сотень людей, простої однопроцесорної платформи і програмного забезпечення Novell може виявитися недостатньо. В цьому випадку використовуються потужні багатопроцесорні сервери з можливостями нарощування основної пам'яті до декількох ГБ, дискового простору до сотень ГБ, швидкими інтерфейсами дискового обміну (типу Fast SCSI-2, Fast&Wide SCSI-2 і Fiber Channel) і декількома мережними інтерфейсами. Ці сервери використовують операційну систему UNIX, мережні протоколи TCP/IP і NFS. На базі багатопроцесорних UNIX-серверів зазвичай будуються також сервери баз даних великих інформаційних систем, бо на них покладається основне навантаження з обробки інформаційних запитів. Сервери подібного типу отримали назву суперсерверів.

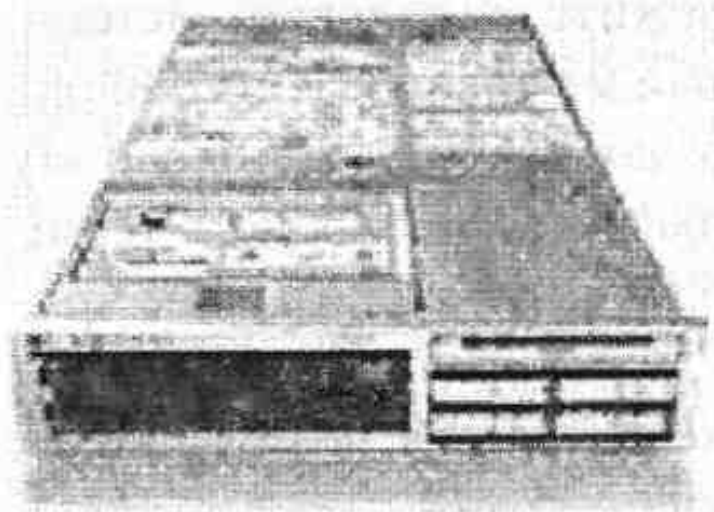


Рис. 1.16. Зовнішній вигляд сервера T2000 фірми Sun Microsystems

Розглянемо технічні характеристики одного з серверів фірми Sun, а саме сервера T2000, зовнішній вигляд якого наведено на рис. 1.16. Його ціна, залежно від комплектування, становить понад $8,419. До складу сервера

входять до восьми процесорів частотою 1.2 ГГц з архітектурою UltraSPARC T1, до 32 ГБ пам'ять, операційна система Solaris.

За загальносистемною продуктивністю, функціональними можливостями окремих компонентів, стійкістю до відмов, а також ступенем підтримки багатопроцесорної обробки, системного адміністрування і дискових масивів великої ємності суперсервери вийшли в нинішній час на один рівень з мейнфреймами. Сучасні суперсервери характеризуються наявністю двох або більше центральних процесорів, багатошинною структурою, мають достатні можливості нарощування дискового простору і обчислювальної потужності, засоби забезпечення надійності зберігання даних і захисту від несанкціонованого доступу.

1.4.5. Великі універсальні комп'ютерні системи

Великі універсальні комп'ютерні системи (мейнфрейми) і до сьогоднішнього дня залишаються найпотужнішими (не рахуючи суперкомп`ютерів) комп'ютерними системами загального призначення, що забезпечують безперервний цілодобовий режим експлуатації. Вони можуть включати один або декілька процесорів, кожний з яких, у свою чергу, може споряджатися векторними спеціалізованими процесорами (прискорювачами операцій з продуктивністю суперкомп'ютерів). Зазвичай мейнфрейми асоціюються з великими за габаритами комп'ютерами, що вимагають спеціально обладнаних приміщень із системами водяного охолоджування і кондиціювання. Однак прогрес в області елементної та конструкторської бази дозволив істотно зменшити габарити основних пристроїв. Поряд з надпотужними мейнфреймами, що вимагають організації двоконтурної водяної системи охолоджування, є менш потужні моделі, для охолоджування яких достатньо примусової повітряної вентиляції, та моделі, побудовані за блоково-модульним принципом, які не вимагають спеціальних приміщень і кондиціонерів.

Основними постачальниками мейнфреймів є відомі комп'ютерні компанії IBM, Amdahl (в даний час в складі концерну Fujitsu), ICL, Fujitsu Siemens Computers, Wincor Nixdorf AG, Siemens Business Services і деякі інші, але провідна роль належить безумовно компанії IBM. Саме архітектура системи IBM/360, випущеної в 1964 році, і її наступні покоління стали зразком для наслідування. Так, в СРСР протягом багатьох років випускалися машини ряду ЄС ЕОМ, що є аналогом цієї системи. В архітектурному плані мейнфрейми є багатопроцесорними системами, що містять один або декілька центральних і периферійних процесорів зі спільною пам'яттю, зв'язаних між собою високошвидкісними магістралями передачі даних. При цьому основне обчислювальне навантаження покладено на центральні процесори, а периферійні процесори (в термінології IBM – селекторні, блок-мультиплексні, мультиплексні канали і процесори телеопрацювання) забезпечують роботу з широкою номенклатурою периферійних пристроїв.

Спочатку мейнфрейми були орієнтовані на централізовану модель обчислень та працювали під керуванням патентованих операційних систем. Однак підвищений інтерес споживачів до відкритих систем, побудованих на базі міжнародних стандартів, змусив постачальників мейнфреймів істотно розширити можливості своїх операційних систем в напрямку сумісності.

Одними з найпотужніших на даний час є мейнфрейми сім'ї zSeries фірми IBM, зовнішній вигляд одного з блоків яких наведено на рис. 1.17. До основних особливостей

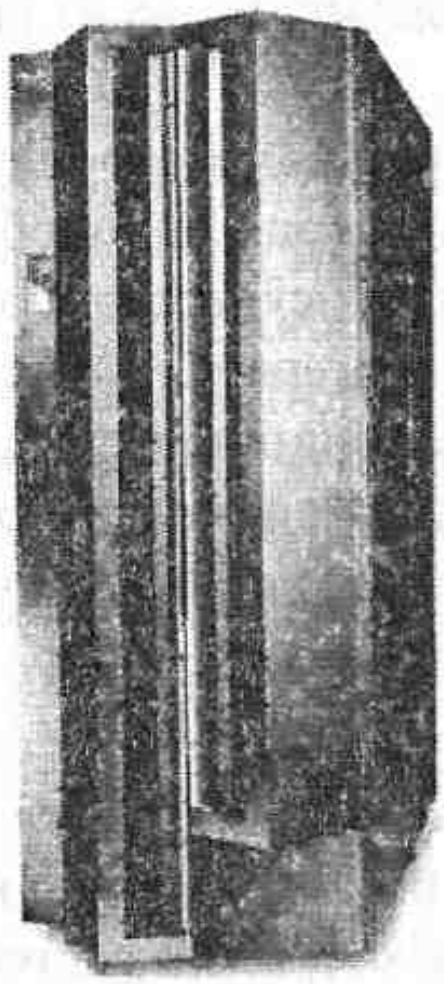


Рис. 1.17. Зовнішній вигляд одного блока мейнфрейма сім'ї zSeries фірми IBM

цих комп'ютерів належать наступні: велика ємність зовнішньої пам'яті, адресація якої забезпечується 64-розрядним адресним словом, розміщення в кожному з 32 блоків до 54 центральних процесорів, причому кожен блок може бути рознесеним в просторі на віддаль до 100 кілометрів, які забезпечують продуктивність понад 18 мільярдів команд за секунду, підтримка роботи з операційними системами Linux, z/OS, z/VM, z/VSE, z/TPF, MUSIC/SP, потужні зовнішні інтерфейси.

Стрімкий ріст продуктивності персональних комп'ютерів, робочих станцій і серверів призвів до зниження потреби в мейнфреймах. Однак цей процес останнім часом дещо уповільнився. Основною причиною відродження інтересу до мейнфреймів експерти вважають складність переходу до розподіленої архітектури клієнт–сервер, що виявилася вищою, ніж припускалося. Крім того, багато користувачів вважають, що розподілене середовище, на противагу мейнфреймам, не є достатньо надійним для найвідповідальніших застосувань.

Головним недоліком мейнфреймів у нинішній час залишається відносно низьке співвідношення продуктивність/вартість. Однак постачальники мейнфреймів докладають значних зусиль для поліпшення цього показника. Слід також пам'ятати, що в світі існує величезна інстальована база мейнфреймів, на якій працюють десятки тисяч прикладних програмних систем. Відмовитися від роками напрацьованого програмного забезпечення просто нерозумно. Тому в нинішній час очікується зростання об'ємів продажу мейнфреймів, які з одного боку, дозволять модернізувати існуючі системи, забезпечивши скорочення експлуатаційних видатків, а з іншого боку, створять нову базу для найвідповідальніших застосувань.

1.4.6. Кластерні комп'ютерні системи

Двома основними проблемами побудови комп'ютерних систем для критично важливих застосувань, зв'язаних з обробкою транзакцій, керуванням базами даних і обслуговуванням телекомунікацій, є забезпечення високої продуктивності та тривалого функціонування систем. Найефективнішим засобом для досягнення заданого рівня продуктивності є застосування паралельних архітектур, які піддаються масштабуванню. Завдання забезпечення тривалого функціонування системи має три складових: надійність, готовність і вартість обслуговування. Всі три складові передбачають, в першу чергу, боротьбу з несправностями системи, що породжуються відмовами і збоями в її роботі. Ця боротьба ведеться по всіх трьох напрямках, що взаємозв'язані і застосовуються спільно.

Підвищення надійності базується на принципі відвертання несправностей шляхом зниження інтенсивності відмов і збоїв за рахунок застосування електронних схем і компонентів з високим і надвисоким ступенем інтеграції, зниження рівня завад, полегшених режимів роботи схем, забезпечення теплових режимів їхньої роботи, а також за рахунок вдосконалення засобів монтажу апаратури. Підвищення рівня готовності передбачає

зниження в певних межах впливу відмов і збоїв на роботу системи з допомогою засобів контролю і корекції помилок, а також засобів автоматичного відновлення обчислювального процесу після прояву несправності, включаючи апаратну і програмну надлишковість, на основі якої реалізуються різноманітні варіанти стійкої до відмови архітектури. Підвищення готовності є засобом боротьби за зниження часу простою системи. Основні експлуатаційні характеристики системи істотно залежать від зручності її обслуговування, зокрема від ремонтопридатності, контролепридатності і т. д.

В останні роки в літературі з обчислювальної техніки все частіше вживається термін "системи високої готовності" (High Availability Systems). Всі типи систем високої готовності орієнтовані на мінімізацію часу простою. Є два типи часу простою комп'ютера: плановий і неплановий. Плановий час простою зазвичай включає час, прийнятий для проведення робіт по модернізації системи і для її обслуговування. Неплановий час простою є результатом відмови системи або її компоненти. Хоча системи високої готовності, можливо, більше асоціюються з мінімізацією непланових простоїв, вони виявляються також корисними для зменшення планового часу простою.

Існує декілька типів систем високої готовності, що відрізняються своїми функціональними можливостями і вартістю. Слід відзначити, що висока готовність не дається безкоштовно. Вартість систем високої готовності набагато перевищує вартість звичайних систем. Певно тому найбільше розповсюдження отримали кластерні системи завдяки їх здатності забезпечити достатньо високий рівень готовності при відносно низьких витратах. Термін "кластеризація" на сьогодні в комп'ютерній промисловості має багато різноманітних значень. Точне визначення могло б звучати так: "реалізація об'єднання машин, що представляється єдиним цілим для операційної системи, системного програмного забезпечення, прикладних програм і користувачів". Машини, кластеризовані таким способом, можуть при відмові одного процесора дуже швидко перерозподілити роботу на інші процесори всередині кластера. Це, можливо, найважливіше завдання багатьох постачальників систем високої готовності.

Першою концепцію кластерної системи анонсувала компанія DEC, визначивши її як групу об'єднаних між собою комп'ютерів, що представляють собою єдиний вузол обробки інформації. По суті кластер цієї компанії був слабко зв'язаною багатомашинною системою з спільною зовнішньою пам'яттю, що забезпечує єдиний механізм керування і адміністрування. На рис. 1.18 показано зовнішній вигляд такого кластера.

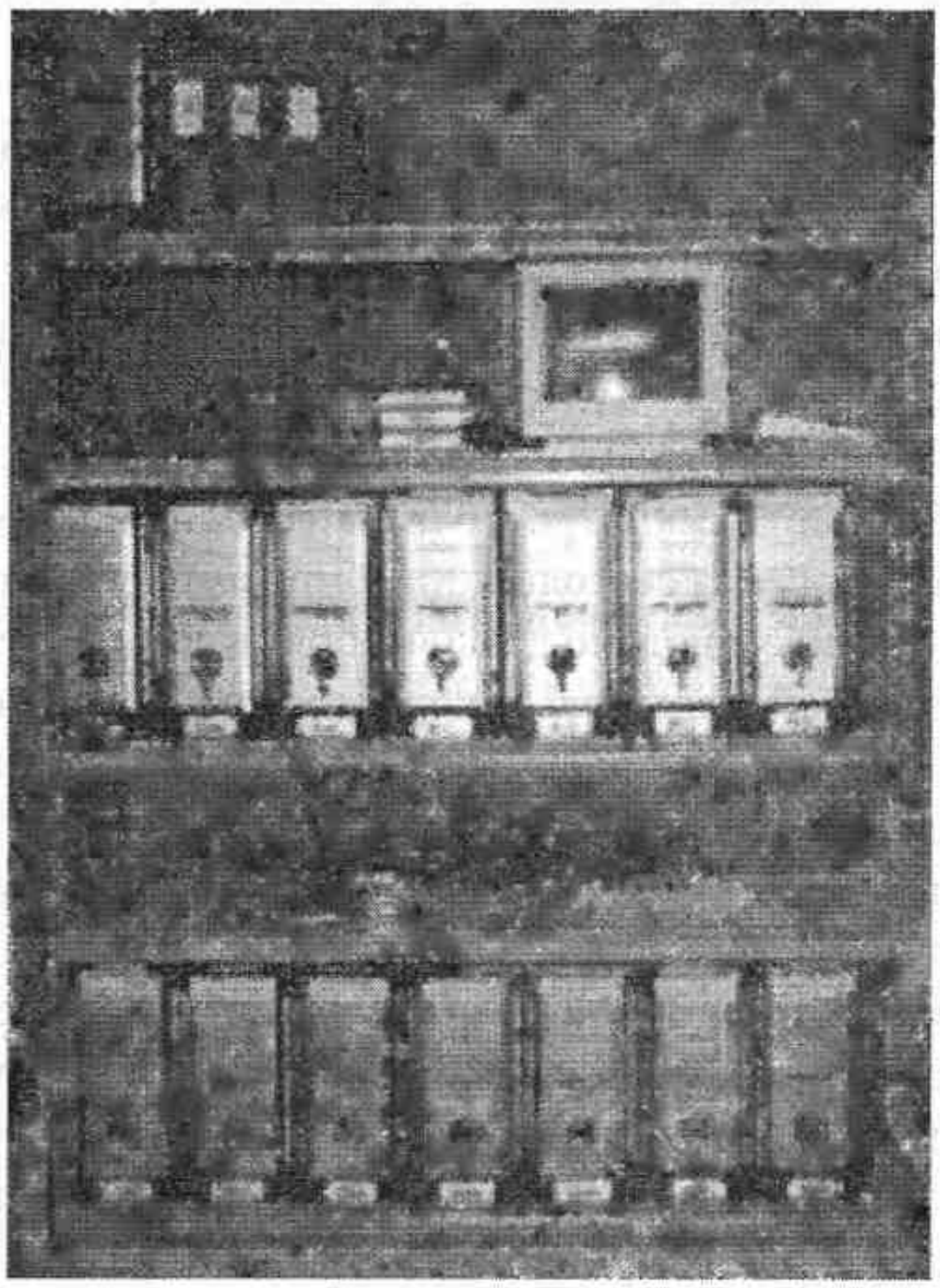


Рис. 1.18. Зовнішній вигляд кластера

Робота будь-якої кластерної системи визначається двома головними компонентами: високошвидкісним механізмом зв'язку процесорів між собою і системним програмним забезпеченням, що надає клієнтам прозорий доступ до системного сервісу.

В даний час широке розповсюдження отримала технологія паралельних баз даних. Ця технологія дозволяє великій кількості процесорів поділяти доступ до єдиної бази даних. Розподіл завдань між процесорними ресурсами і паралельне їх виконання дозволяє досягнути вищого рівня пропускної спроможності транзакцій, підтримувати більше число одночасно працюючих користувачів і прискорити виконання складних запитів. Для вирішення цих завдань використовується архітектура зі спільними (розподіленими) дисками. Це типовий випадок побудови кластерної системи. Ця архітектура підтримує єдину базу даних при роботі з декількома комп'ютерами, об'єднаними в кластер (зазвичай такі комп'ютери називаються вузлами кластера), кожний з яких працює під керуванням своєї копії операційної системи. В таких системах всі вузли поділяють доступ до загальних дисків, на яких власне і розміщується єдина база даних. Продуктивність таких систем може збільшуватися як шляхом нарощування числа процесорів і ємності основної пам'яті в кожному вузлі кластера, так і шляхом збільшення кількості самих вузлів. У випадку відмови одного з таких вузлів, вузли, що залишилися, можуть взяти на себе завдання, що виконувалися на вузлі, який відмовив, не зупиняючи загальний процес роботи з базою даних. Оскільки логічно в кожному вузлі системи є образ бази даних, доступ до неї буде забезпечуватися до тих пір, доки в системі є принаймні один справний вузол.

1.4.7. Суперкомп'ютери

До класу суперкомп'ютерів належать комп'ютери, що мають максимальну в даний час продуктивність, а також максимальну ємність основної та зовнішньої пам'яті. Вони асоціюються з великими розмірами, великими завданнями, гранично високими характеристиками. Швидкий розвиток комп'ютерної індустрії призводить до відносності даного поняття. Суперкомп'ютер десятирічної давності сьогодні під це визначення вже не потрапляє. Наприклад, продуктивність персональних комп'ютерів, що використовують Pentium-II/300MHz, є близькою до продуктивності суперкомп'ютерів середини 70-х років, проте за сьогоднішніми мірками суперкомп'ютерами не є ні ті, ні інші.

Нижче подано декілька прикладів, що показують основні характеристики комп'ютерів цього класу, які використовуються в даний час.

CRAY T932, векторно-конвеєрний комп'ютер фірми CRAY Research Inc. (на сьогодні це є підрозділ Silicon Graphics Inc.), уперше випущений у 1996 році. Максимальна продуктивність одного процесора дорівнює майже 2 млрд операцій за секунду, основна пам'ять нарощується до 8 ГБ, дисковий простір до 256000 ГБ (тобто 256Тб). Комп'ютер у максимальній конфігурації вміщує 32 процесори, що працюють із загальною пам'яттю, тому максимальна продуктивність всієї комп'ютерної системи складає більше 60 млрд операцій за секунду.

IBM SP2, матричний паралельний комп'ютер фірми IBM. Побудований на основі стандартних процесорів PowerPC 604e або POWER2 SC, сполучених між собою через високошвидкісний комутатор, причому кожний має свою локальну основну пам'ять і

дискову підсистему. Характеристики цих процесорів відомі й особливого подиву не викликають, проте в рамках однієї системи SP2 їх може бути об'єднано дуже багато. Зокрема, максимальна система, встановлена в Pacific Northwest National Laboratory (Richland, USA), вміщує 512 процесорів. Виходячи з числа процесорів, можна уявити сумарну потужність всієї системи.

HP Exemplar, комп'ютер із кластерною архітектурою від Hewlett-Packard Inc. Зокрема, модель V2250 (клас V) побудована на основі мікропроцесора PA-8200, що працює з тактовою частотою 240 МГц. В рамках одного вузла зі спільною основною пам'яттю до 16 ГБ можна об'єднати до 16 процесорів. У свою чергу вузли в рамках однієї комп'ютерної системи з'єднуються між собою через високошвидкісні канали передачі даних.

Суперкомп'ютер ASCI RED, результат виконання програми Accelerated Strategic Computing Initiative. Побудований на замовлення Міністерства енергетики США, він об'єднує 9152 процесори Pentium Pro, має 600 ГБ сумарної основної пам'яті та загальну продуктивність 1800 мільярдів операцій за секунду.

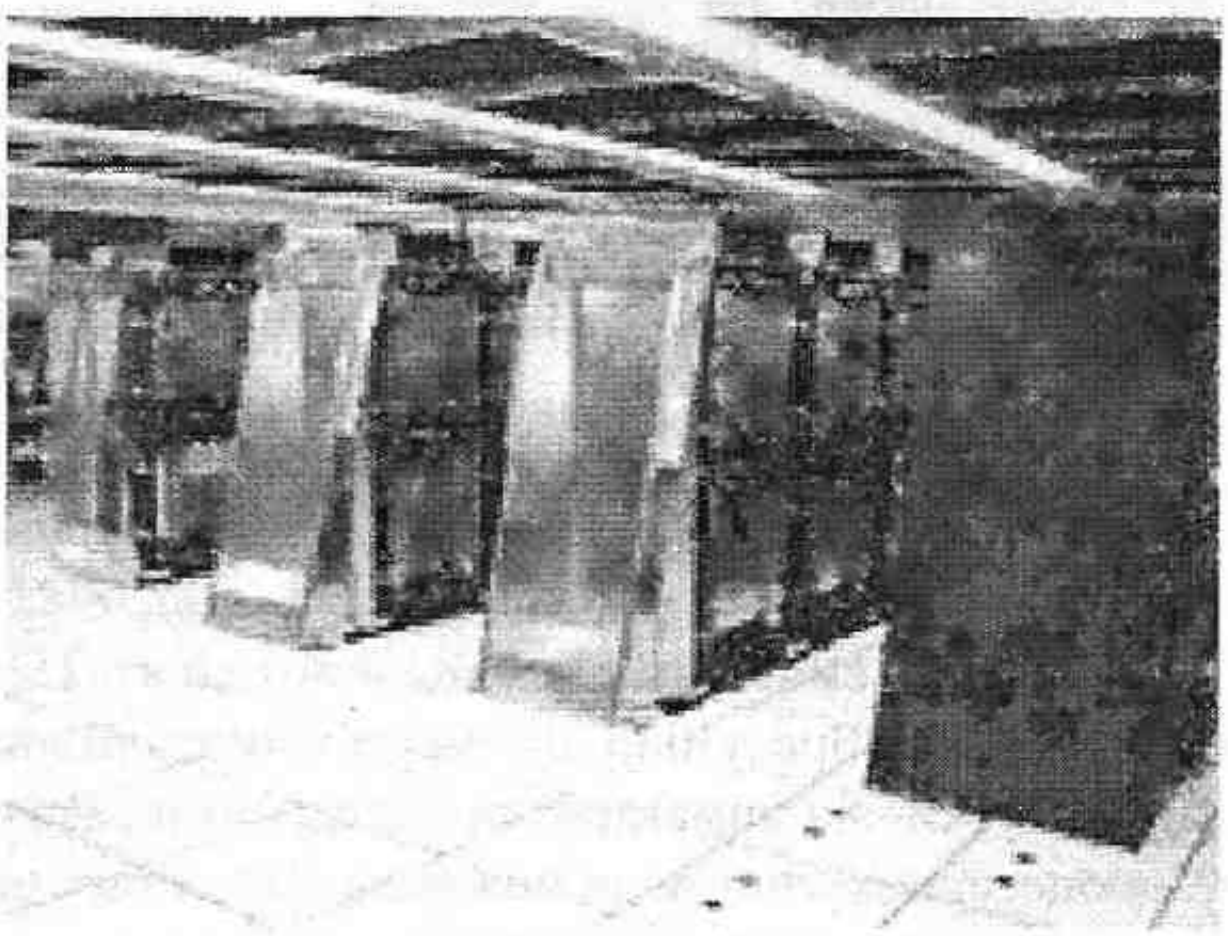


Рис. 1.19. Зовнішній вигляд суперкомп'ютера Blue Gene/L фірми IBM

Найпотужнішим на сьогодні комп'ютером є суперкомп'ютер фірми IBM Blue Gene/L (рис. 1.19), який має 131 072 процесорних вузлів та продуктивність 280.6 TFLOPS (10^{12} FLOPS). Кожен вузол містить процесор PowerPC 440 із 512 МБ локальної пам'яті.

В 2006 році був уведений в експлуатацію суперкомп'ютер MDGRAPE-3, який досяг продуктивності 1 PFLOPS (10^{15} FLOPS), однак його не відносять до універсальних суперкомп'ютерів, оскільки він є орієнтованим на виконання задач молекулярної динаміки.

Навіть спрощені конфігурації таких комп'ютерів коштують не один мільйон доларів США. Виникає ряд природних запитань:

- які завдання настільки важливі, що потребують використання комп'ютерів вартістю декілька мільйонів доларів?
- які завдання настільки складні, що процесора Pentium IV недостатньо?

От лише невеличкий список областей людської діяльності, де необхідно використовувати суперкомп'ютери: автомобілебудування; нафто- і газовидобуток; фармакологія; прогноз погоди і моделювання зміни клімату; сейсморозвідка; проектування електронних пристроїв; синтез нових матеріалів, генні дослідження.

На рис. 1.20 подано завдання, для виконання яких необхідне застосування суперкомп'ютерів, а також потрібні для їх вирішення комп'ютерні ресурси.

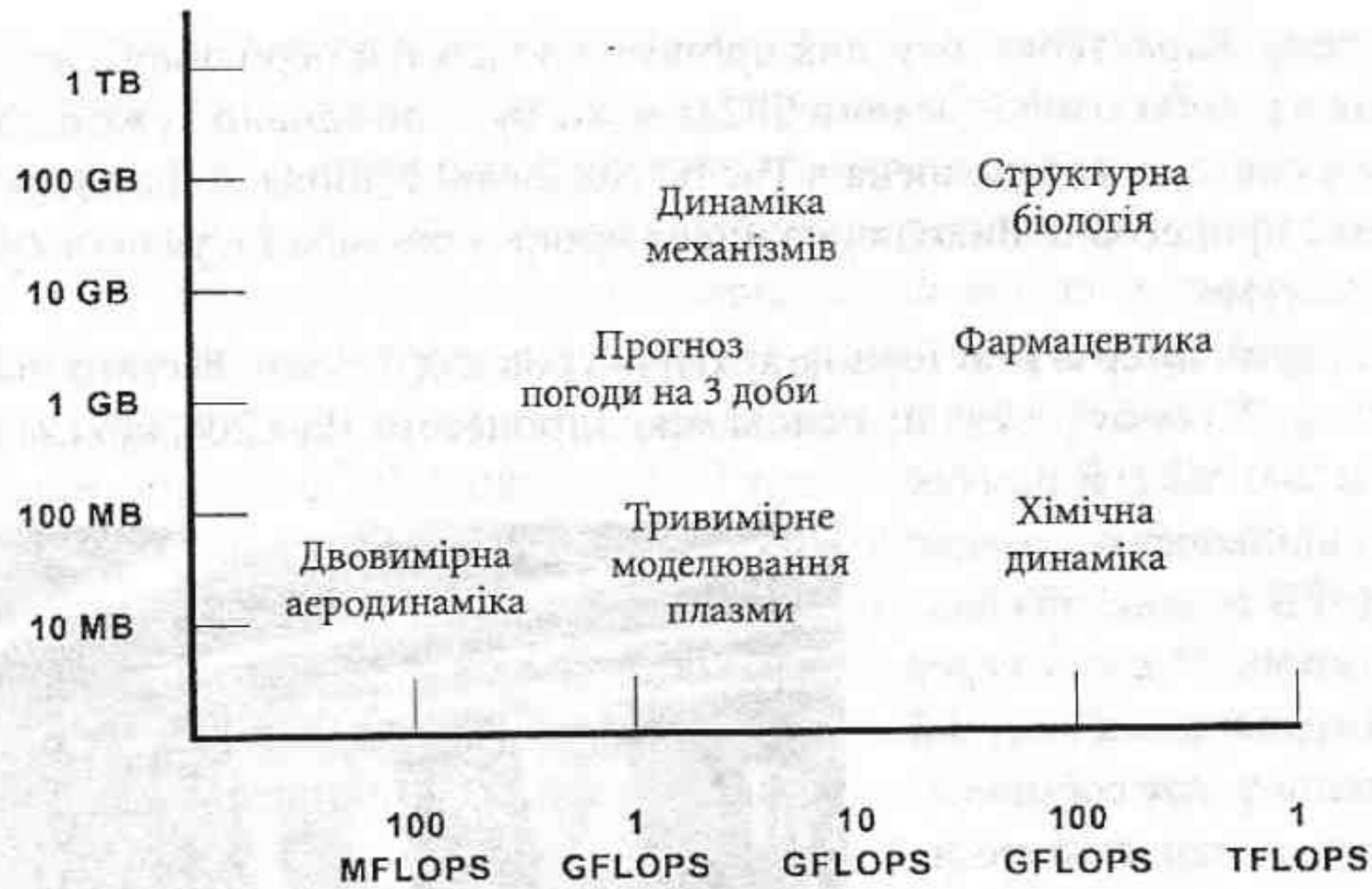


Рис. 1.20. Завдання та комп'ютерні ресурси

Видно, що ємність пам'яті досягає одного ТБ за умови, що продуктивність має бути один TFLOPS. Зрозуміло, що межа необхідних комп'ютерних ресурсів є рухомою. Надати ресурси, які вимагаються наведеними завданнями, за допомогою стандартних однопроцесорних систем неможливо. Це спричинює використання багатопроцесорних комп'ютерних систем як магістрального напрямку досягнення високої продуктивності.

1.4.8. Мікроконтролери

Мікроконтролери – комп'ютери на кристалі, призначені для керування електронними пристроями, зокрема побутовими пристроями, виробничими лініями, вимірювальними пристроями і т. д. До складу мікроконтролера входять наступні вузли:

- центральний процесор, розрядністю від 4 до 64 бітів, залежно від потрібної точності обчислень;
- інтерфейси введення-виведення, в першу чергу послідовні порти;
- периферійні пристрої, такі як: таймери та схеми захисту, цифроаналогові та аналогоцифрові перетворювачі;
- пам'ять з довільним доступом для зберігання даних;
- постійна пам'ять типу ROM, EPROM, EEPROM чи Flash для зберігання програми;
- генератор тактів.

Така інтеграція названих пристроїв на кристалі дозволяє забезпечити малі габарити та споживання і сприяє широкому використанню мікроконтролерів у різного роду вбудованих системах. Наприклад, в сучасному автомобілі використовується понад 50 мікроконтролерів. Вони також використовуються в побутовій електроніці, мобільних телефонах, виробничих лініях тощо. На рис. 1.21 подано зовнішній вигляд мікроконтролера PIC 18F8720 фірми Microchip в корпусі TQFP з 80 виводами.

Розробники мікроконтролерів забезпечують спеціальний сервіс для користувачів, зокрема версії, які дозволяють перепрограмування програмної пам'яті ультрафіолето-

вим світлом, можливість підключення зовнішньої оперативної пам'яті в якості пам'яті програм, та інше. Сучасні мікроконтролери програмуються в коді мови С та мають внутрішні схеми відлагодження.

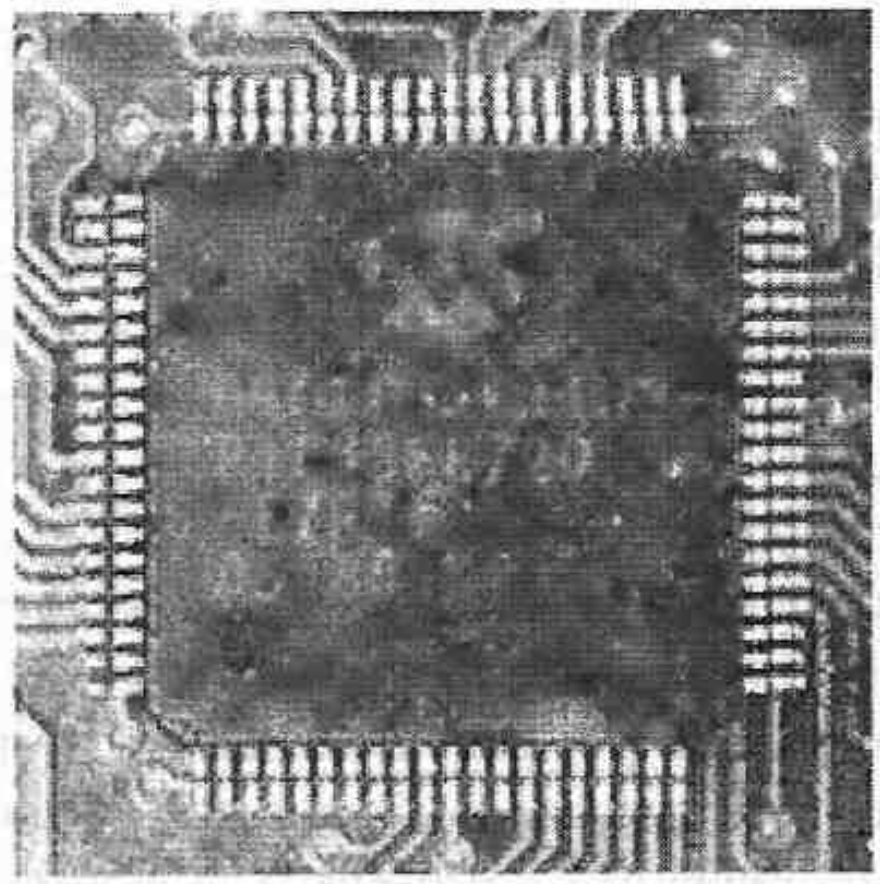


Рис. 1.21. Зовнішній вигляд мікроконтролера PIC 18F8720 фірми Microchip

1.4.9. Спеціалізовані комп'ютери

За допомогою універсальних комп'ютерів та комп'ютерних систем (УКС), які були розглянуті вище, можна вирішувати багато задач наукового, виробничо-технічного та іншого характеру. Однак існують надзвичайно важливі класи задач і окремі задачі, для розв'язку яких математичні та техніко-економічні якості УКС недостатні. Не варто доводити дієвість принципу спеціалізації інструментальних засобів взагалі, оскільки вся свідома технічна діяльність людства її підтверджує. Досить вказати, що цей принцип ефективно діє і в галузі інформатики. Загальний аналіз причин створення і використання спеціалізованих комп'ютерних систем (СКС) показує, що ці причини можна віднести до трьох основних груп.

Перша група об'єднує причини, що виникли внаслідок суперечностей між формальними математичними методами постановки і розв'язку задач, з одного боку, і загальними принципами організації та функціонування, а також технічними можливостями УКС, з іншого боку. Саме математична сутність задач часто обумовлює необхідність створення СКС для їх розв'язку. Як приклади тут можна навести нові нестандартні та неалгоритмічні методи, системи алгебраїчних, диференційних та інтегральних рівнянь великої розмірності, логічні та імовірнісно-статистичні задачі, дії над матрицями та векторами, задачі в багатовимірних просторах та багато інших.

До другої групи входять причини, які обумовлені змістовною стороною задач, вирішуваних СКС, та відображають специфіку відповідних предметних областей.

Третя група причин обумовлена особливими вимогами до якості реалізації комп'ютерних систем, які зазвичай полягають в екстремалізації (тобто в максимальному наближенні до теоретичних границь) деяких їх характеристик, наприклад, продуктивності, надійності (безвідмовності, живучості, відновлюваності, довговічності та ін.), вартості, точності, габаритів, маси і т.п. Сюди ж належать вимоги, що визначають такі якості комп'ютерних систем, як їх повна або часткова імплантація (конструктивне та функціональне суміщення) в інші системи, інформаційне поєднання з ними, пристосованість до умов експлуатації та кваліфікації обслуговуючого персоналу і т.д.

Слід мати на увазі, що реальні ситуації створення СКС найповніше характеризуються двома особливостями. Перша полягає в тому, що саме СКС є своєрідним засобом апробації нових методів автоматизації обробки інформації, що мають математичні корені. Наприклад, розпаралелювання та децентралізація обчислень, макрооперації та функціональні розширювачі, символьна обробка та розв'язок задач в багатовимірних числових системах та ін. пройшли спочатку дуже ретельну перевірку в СКС і тільки після цього з'явилися в УКС.

4 Зам. 371.

Друга особливість пов'язана з тим, що реальні СКС є складними програмно-технічними комплексами, в яких на інженерному рівні необхідно задовольнити багато суперечливих вимог. Тому досягнення оптимальних і функціональних якостей СКС може бути проблематичним і доцільніше визначати ці якості як оптимізовані, тобто такі, що тією чи іншою мірою наближаються до оптимальних. Аналіз математичних методів оптимізації СКС показує, що вони дозволяють, певною мірою, виявляти недоліки таких систем, їхні "слабкі місця", простежити взаємозв'язок характеристик системи, визначити загальний напрямок підвищення їх ефективності та оцінити різні варіанти СКС. Однак ці методи не дають ніяких конструктивних рішень і шляхів удосконалення СКС, не визначають змістовної сторони різних варіантів їх організації та реалізації. Генезис таких варіантів формальними математичними методами неможливий. Тому процес створення оптимізованих СКС має характер багатоступеневої ітераційної процедури, де в різних відношеннях комбінуються формальні та конкретно-змістовні методи, що відіграють аналітичну (оціночну) та синтетичну (генеративну) ролі.

Таким чином, СКС – це комп'ютерні системи для розв'язку великого числа відносно вузьких класів задач, оптимізовані в певній критеріальній сукупності.

Для СКС характерні наступні риси, які відрізняють їх від універсальних комп'ютерних систем:

- орієнтація структури на вирішувані задачі;
- вузький, в основному постійний клас вирішуваних задач;
- особливі вимоги до точності, часто нестандартна довжина розрядної сітки;
- спеціальна система обміну, в тому числі наявність аналого-цифрових та цифро-аналогових каналів зв'язку;
- використання орієнтованих на область застосування мов програмування та широкі можливості їх апаратної інтерпретації;
- наявність спеціальних функцій і процедур в наборі операцій та команд;
- необхідність обробки вхідних даних в темпі їх поступлення та видачі результатів обчислень в темпі поступлення вхідних даних;
- суміщення в часі приймання, обробки та видачі даних;
- висока продуктивність;
- малі габарити;
- низька споживана потужність;
- орієнтація конструкції на конкретне застосування.

1.5. Предмет та порядок розгляду матеріалу даної книги

Комп'ютер є складною системою. Сучасні комп'ютери складаються з мільйонів або й мільярдів елементарних електронних компонент. Для зрозумілого опису такої складної системи використовується її ієрархічна природа, тобто комп'ютер розглядається як набір взаємозв'язаних підсистем, кожна з яких також є ієрархічною (рис. 1.22). Найнижчим рівнем є елементарна підсистема. На кожному рівні складна система включає набір відповідним чином з'єднаних компонент, які визначають її поведінку на даному рівні. Кожен рівень може бути описаним структурою та функціями компонент. Структура включає перелік компонент та організацію їх взаємозв'язків. Функції описують операції кожної індивідуальної компоненти.

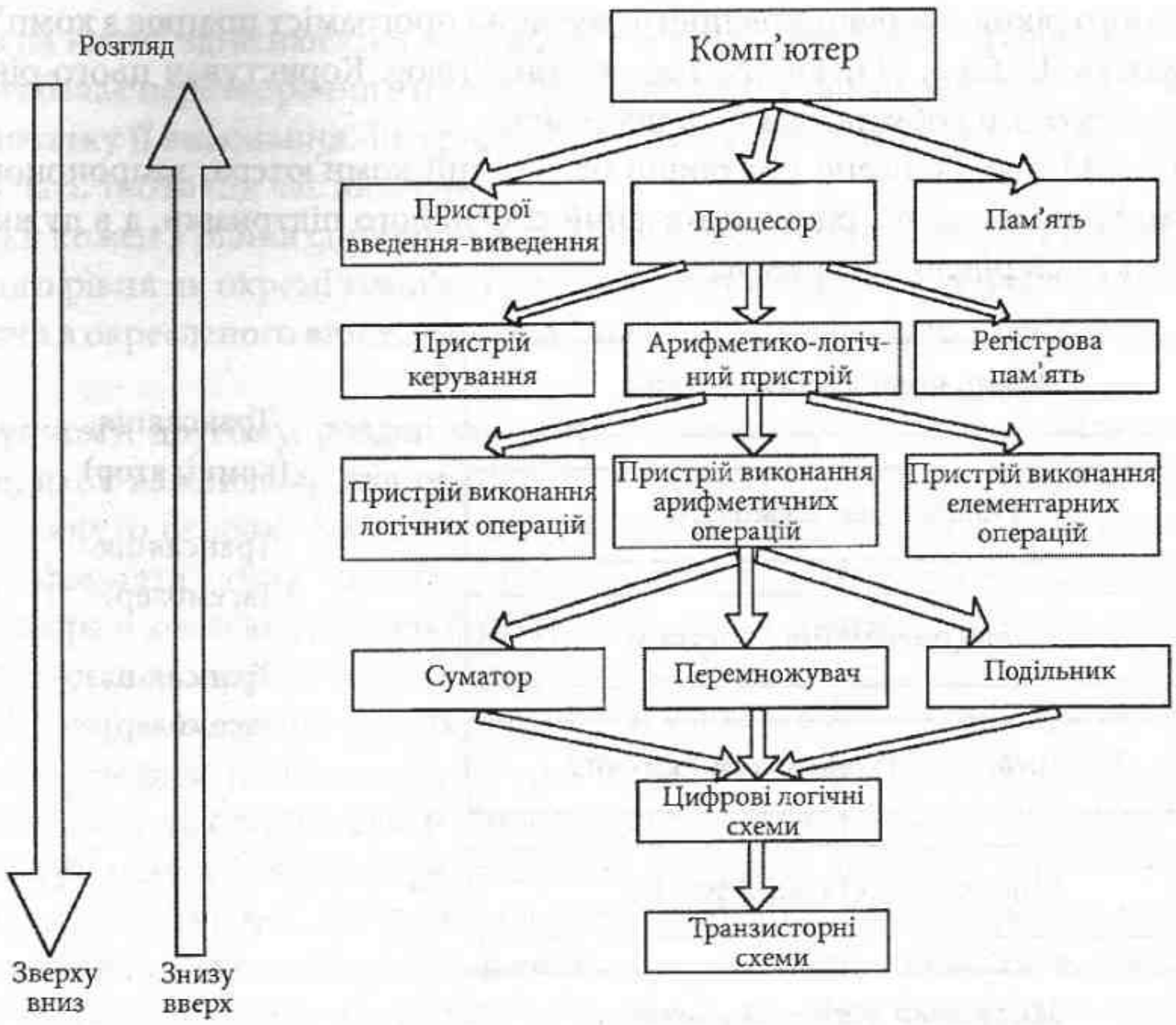


Рис. 1.22. Ієрархічна природа комп'ютера

Опис може виконуватись як знизу догори, починаючи з елементарних електронних компонентів і закінчуючи повним описом системи, так і згори донизу, шляхом поділу системи на складові частини. Перший підхід використовується при проектуванні комп'ютерів та їх функціональних вузлів, зокрема при їх ієрархічному описі мовами опису апаратних засобів типу VHDL та Verilog. Для розгляду та вивчення комп'ютера другий підхід є ефективнішим, тому в подальшому будемо в основному користуватися цим підходом.

Комп'ютер має багаторівневу організацію. Найнижчим є рівень фізичних пристроїв, об'єктами якого є транзистори. Він є основою для цифрового логічного рівня, об'єктами якого є вентилі, побудовані на декількох транзисторах. Вентиль виконує прості логічні функції, такі як: І, АБО. На основі цифрового логічного рівня формується рівень міжрегістрових передач або мікроархітектурний рівень, об'єктами якого є регістри, мультиплексори, лічильники, стеки регістрів, блоки постійної та оперативної пам'яті, суматори, арифметико-логічні пристрої і т. д. Цей рівень можна поділити на підрівні, оскільки кожен з перерахованих елементів може бути представленим кількома рівнями простіших елементів. Наступним є архітектурний рівень, або рівень системи команд, який також має підрівні. Далі йде макроархітектурний рівень, або рівень операційної системи. Цей рівень є відповідальним за мультипрограмування, захист пам'яті, синхронізацію процесів та багато інших важливих функцій. І, нарешті, асемблерний рівень та рівні мов програмування. Кожна команда асемблерного рівня транслюється у відповідну їй команду

архітектурного рівня. На рівні мов програмування програміст працює з комп'ютером на таких мовах як: C, C++, FORTRAN, Lisp, Pascal, Prolog. Користувач цього рівня бачить дуже мало з того, що робиться на нижчих рівнях.

На рис. 1.23 представлено сім рівнів організації комп'ютера, запропоновані Е. Таненбаумом. Біля кожного з рівнів показаний спосіб його підтримки, а в дужках – назва програми, яка цей рівень підтримує.

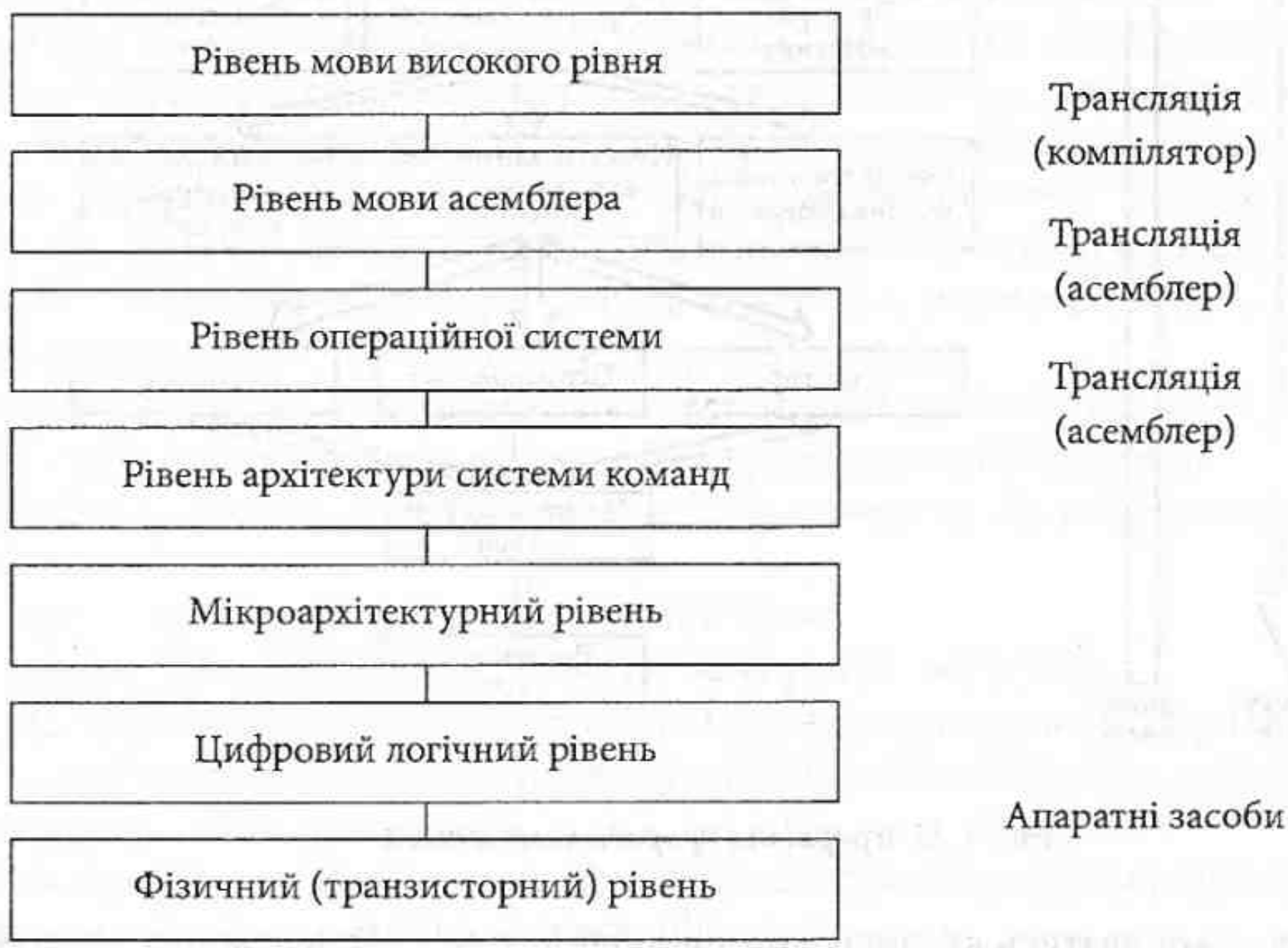


Рис. 1.23. Сім рівнів організації комп'ютера

Ми не будемо розглядати фізичний та цифровий логічний рівні, які є предметом розгляду комп'ютерної електроніки, прикладної теорії цифрових автоматів та комп'ютерної схемотехніки, та які читачеві потрібно знати, приступаючи до вивчення предмета архітектури комп'ютера. В даній книзі розглянемо рівень архітектури системи команд та частково знизу мікроархітектурний рівень, який з позиції функціонування та побудови елементів комп'ютера розглядається в курсі комп'ютерної схемотехніки, і зверху рівень операційної системи. З рівня архітектури системи команд програміст може працювати з комп'ютером, причому, на цьому рівні найбільше враховуються особливості побудови комп'ютера.

Разом з тим, для роботи з комп'ютером на рівні архітектури системи команд потрібно писати програму на машинній мові, тобто в командах, які є двійковими кодами, що є досить складним навіть для вирішення простих задач. Для спрощення написання машинних команд була введена асемблерна мова, коли команди представляються в мнемонічній формі. Але це не спростило взаємодію людини з комп'ютером. З метою скорочення розриву між мовою людини та машини були введені мови програмування високого рівня. Рівні мов програмування орієнтовані на специфіку використання комп'ютера. Оскільки комп'ютер розпізнає команди починаючи з архітектурного рівня, перехід з ви-

щих рівнів на нижчі здійснюється за допомогою трансляторів та інтерпретаторів. Трансляція передбачає перетворення з одного рівня на інший всього або частини тексту програми до початку її виконання. Інтерпретація – це перетворення програми частинами в реальному часі, тобто під час виконання.

Оскільки кожен з рівнів сприймається по-іншому, користувачі сприймають комп'ютери кожного рівня як окремі комп'ютери, які називають віртуальними машинами.

Виходячи з окресленого вище, матеріал даної книги викладено в наступній послідовності.

В наступному, другому, розділі розкрито основні питання представлення даних в комп'ютері, які є важливими для розуміння матеріалу наступних розділів. В третьому розділі розглянуто основні елементи архітектури комп'ютера, які включають організацію пам'яті, формати і типи команд, способи адресації. В четвертому розділі визначено місце процесора в комп'ютері, його функції та склад. В п'ятому розділі розглянуто конфлікти в конвеєрі команд та методи їх усунення. В шостому розділі розкрито основні питання виконання операцій обробки даних. В сьомому розділі розглянуто принципи побудови арифметико-логічного пристрою сучасних комп'ютерів. В восьмому розділі описано структуру та організацію роботи пристрою керування. В дев'ятому розділі подано структуру пам'яті комп'ютера. В десятому розділі описано організацію взаємодії між рівнями ієрархічної пам'яті та питання захисту пам'яті. В одинадцятому розділі наведено пояснення способів розпізнавання пристроїв введення-виведення, описано функції інтерфейсної схеми пристроїв введення-виведення та наведено методи керування введенням-виведенням. У дванадцятому розділі розглянуто питання подальшого підвищення продуктивності комп'ютерів шляхом створення паралельних комп'ютерних систем.

1.6. Підсумок розділу

У цьому розділі представлено короткий огляд організації й архітектури комп'ютерної системи. Пояснено деяку термінологію в контексті комп'ютерної реклами. Описано основні функціональні вузли комп'ютера, їх функції та характеристики, а також тенденції зміни цих характеристик. Багато чого з цієї термінології буде розширено в наступних розділах. Історично комп'ютери були досить простими, але з часом вони ускладнилися, що викликало необхідність проводити їх розгляд як ієрархію рівнів замість складної системи. Кожен рівень в цій ієрархії слугує певній меті, і всі рівні допомагають мінімізувати семантичний проміжок між мовою програмування високого рівня або застосуванням та вентилями і провідниками, які складають фізичні технічні засоби. Описано виконувані завдання та типи комп'ютерів. Можливо, найважливішим питанням цього розділу є введення поняття архітектури машини фон Неймана. Хоча є інші архітектурні моделі, архітектура фон Неймана є домінуючою в сьогоднішніх універсальних комп'ютерах.

1.7. Література для подальшого читання

Основні поняття, функції й основні функціональні вузли комп'ютера та їх взаємозв'язок, а також загальна організація роботи комп'ютера детально розглянуто в літературі [1–4]. Тут же, а також в роботах [5–11], проведено детальний огляд історії створення та

розвитку комп'ютерів. Окремо потрібно сказати про роботу [5], в якій подано історичний опис процесу створення комп'ютерів провідними науковими та промисловими підприємствами колишнього Радянського Союзу, в тому числі показано суттєві досягнення в цій галузі й українських науковців та інженерів. В роботі [12] наведено підтвердження дієвості закону Мура, а в роботі [13] описано багаторівневу організацію комп'ютера.

1.8. Література до розділу I

1. Burks, A. W., H. H. Goldstine, and J. von Neumann [1946]. "Preliminary discussion of the logical design of an electronic computing instrument," Report to the U.S. Army.

2. Amdahl, G. M. [1967]. "Validity of the single processor approach to achieving large scale computing capabilities," Proc. AFIPS 1967 Spring Joint Computer Conf. 30 (April), Atlantic City, N.J., 483–485.

3. Wilkes, M. V., D. J. Wheeler, and S. Gill [1951]. The Preparation of Programs for an Electronic Digital Computer, Addison-Wesley, Cambridge, Mass.

4. Augarten, Stan. Bit by Bit: An Illustrated History of Computers. London: Unwin Paperbacks, 1985.

5. Малиновский Б. М. История вычислительной техники в лицах. КІТ, Киев, 1995.

6. Blaauw, G., & Brooks, F. Computer Architecture: Concepts and Evolution. Reading, MA: Addison-Wesley, 1997.

7. Ceruzzi, Paul E. A History of Modern Computing. Cambridge, MA: MIT Press, 1998.

8. Cortada, J. W. Historical Dictionary of Data Processing, Volume 1: Biographies; Volume 2: Organization, Volume 3: Technology. Westport, CT: Greenwood Press, 1987.

9. McCartney, Scott. ENIAC: The Triumphs and Tragedies of the World's First Computer. New York: Walker and Company, 1999.

10. Mollenhoff, Clark R. Atanasoff: The Forgotten Father of the Computer. Ames, IA: Iowa State University Press, 1988.

11. Polachek, Harry. "Before the ENIAC." IEEE Annals of the History of Computing 19: 2 (June 1997), pp. 25–30.

12. Schaller, R. "Moore's Law: Past, Present, and Future." IEEE Spectrum, June 1997, pp. 52–59.

13. Tanenbaum, A. Structured Computer Organization, 4th ed. Upper Saddle River, NJ: Prentice Hall, 1999.

1.9. Питання до розділу I

1. Що таке комп'ютер?
2. Дайте визначення алгоритму.
3. Дайте хронологію появи перших комп'ютерів.
4. Які основні функції комп'ютера? Які основні функціональні вузли комп'ютера? Їх завдання та основні характеристики.
5. Назвіть три базових компоненти кожного комп'ютера.
6. Назвіть три типи пам'яті комп'ютера.
7. Який степінь 10 означає префікс Гіга? Якому степеню двійки він еквівалентний?
8. Який степінь 10 означає префікс мікро? Якому степеню двійки він еквівалентний?
9. Яка одиниця використовується для вимірювання тактової частоти комп'ютера?
10. Дайте пояснення суті закону Мура.
11. Як змінюється з часом тактова частота роботи процесора?
12. На скільки щороку зростає продуктивність комп'ютерів?

13. Як змінюється з часом ємність та час зчитування-запису динамічної напівпровідникової пам'яті?
14. Як змінюється з часом ємність зовнішньої дискової пам'яті?
15. Які одиниці використовуються для оцінки продуктивності комп'ютера?
16. Що дає використання тестових програм для оцінки продуктивності комп'ютера?
17. Які існують тестові програми для оцінки продуктивності комп'ютера?
18. Які є основні варіанти зв'язку між функціональними вузлами комп'ютера? Які їх відмінності?
19. Порівняйте двошинні структури комп'ютера з обміном через процесор та через пам'ять.
20. Як узгодити передачу даних між повільними і швидкими вузлами комп'ютера?
21. Поясніть роботу комп'ютера з одношинною структурою.
22. Що таке архітектура комп'ютера?
23. Що таке структура комп'ютера?
24. Що таке архітектура системи команд?
25. Яка різниця між архітектурою та структурою комп'ютера?
26. За якими ознаками здійснюється класифікація комп'ютерів?
27. Назвіть особливості архітектури Джона фон Неймана.
28. Які існують ненейманівські архітектури комп'ютерів? Які їх відмінні риси?
29. Назвіть два основних напрями використання комп'ютерів.
30. Що розуміється під паралельною обробкою?
31. Які є типи комп'ютерів?
32. Персональні комп'ютери – особливості та сфери застосування.
33. Як сконструйовано сучасний персональний комп'ютер?
34. Робочі станції – особливості та сфери застосування.
35. Багатотермінальні системи – особливості та сфери застосування.
36. Сервери – особливості та сфери застосування.
37. Великі універсальні комп'ютерні системи – особливості та сфери застосування.
38. Кластерні архітектури – особливості та сфери застосування.
39. Суперкомп'ютери – особливості та сфери застосування.
40. Мікроконтролери – особливості та сфери застосування.
41. Спеціалізовані комп'ютери – особливості та сфери застосування.
42. Поясніть ієрархічну природу комп'ютера.
43. Назвіть сім рівнів організації комп'ютера.