Wireless sensor networks: Networking support for query processing in sensor networks

Alec Woo, Sam Madden, Ramesh Govindan

June 2004 **Communications of the ACM**, Volume 47 Issue 6

Outline

1. introduction
2. In-Network Querying Systems
3. Novel Networking Mechanisms
4. Toward a Unified In-Network System

Introduction (cont.)

- Query processing systems
  - Directed Diffusion
  - TinyDB
  - Cougar

- Provide high-level interface.

- Especially attractive as ways to efficiently implement monitoring applications.

Introduction (cont.)

- Researchers are beginning to formulate languages and enumerate the types of queries needed by users of sensor networks .

- Distributed query processing used to store or retrieve data from nodes within the network.

- The performance and functionality is significantly affected by internode communication mechanisms.

Introduction

- Different applications have widely varying requirements
  - information-transfer rates
  - timeliness
  - quality
  - storage

- A substantial software engineering challenge
  – Subsystems must be reusable in isolation from one another.

In-Network Querying Systems: TinyDB (cont.)

- Several in-network querying systems have been built for sensor networks

- In TinyDB, users specify declarative queries that reflect their data processing needs.

- Queries are written in a SQL-like language.

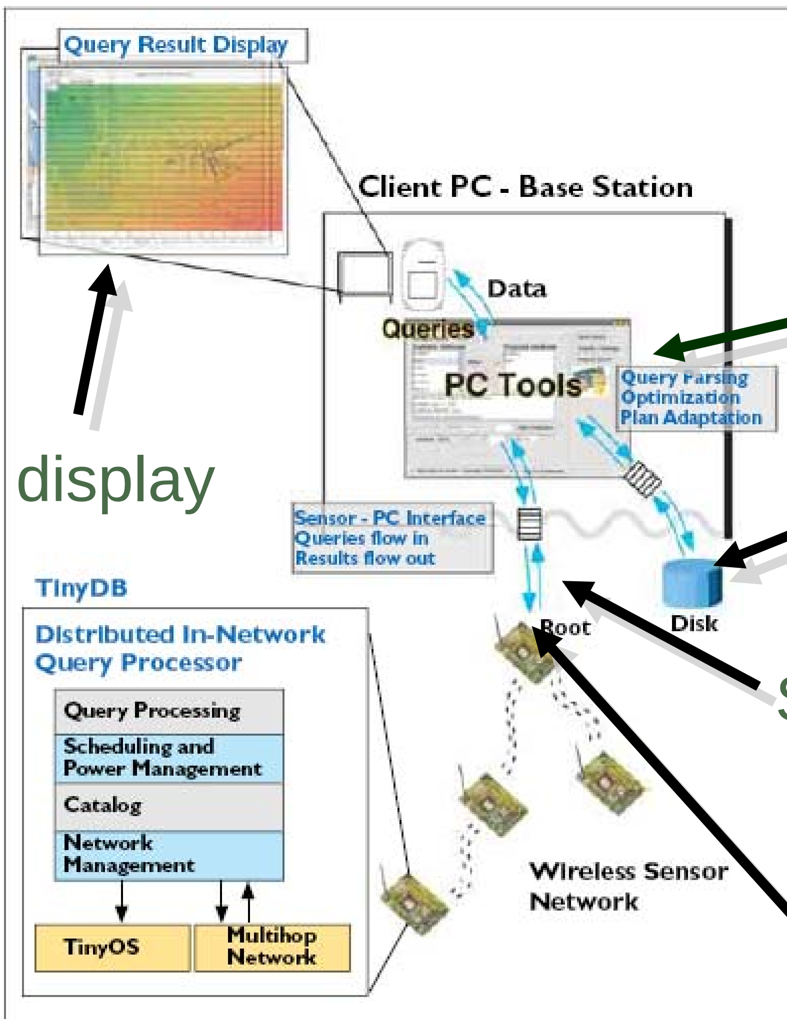
In-Network Querying Systems: TinyDB (cont.)

- Example:

```
SELECT AVG(temp)
FROM sensors
WHERE loc in (0,0,100,100) and
              light > 1000 lux
SAMPLE PERIOD 10 seconds
```

Report the average temperature of the bright nodes in the region defined by the rectangle(0,0,100,100)

In-Network Querying Systems: TinyDB (cont.)



The high-level architecture of TinyDB

display

Queries are input, parsed and optimized at a PC

Store the result

Sends the query into the tree-based sensor network for processing

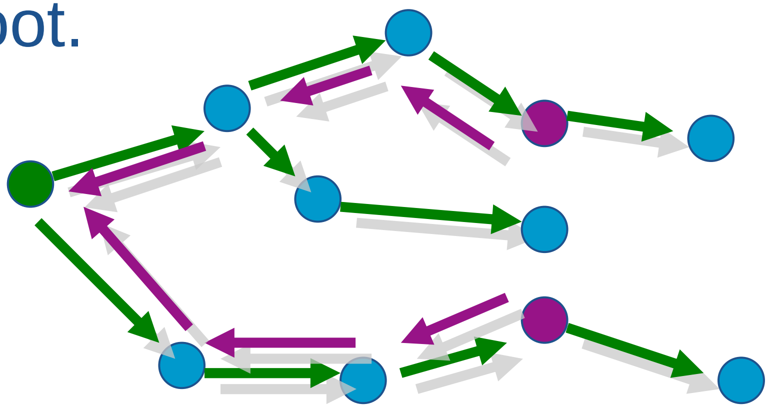
Result flow up to the PC

In-Network Querying Systems: TinyDB

- Queries in TinyDB are flooded throughout the network.

- Query answers are collected via a routing tree with the root node being the endpoint of the query.

- Every other nodes maintains a parent node one step closer to the root.



In-Network Querying Systems: Directed Diffusion (cont.)

- Directed Diffusion does not specify a query language.
- It allowing application writers to choose a domain-specific query processing language.
- Diffusion focuses on two other dimensions of query processing.
  - Query routing mechanisms
  - flexible in-network processing

In-Network Querying Systems: Directed Diffusion (cont.)

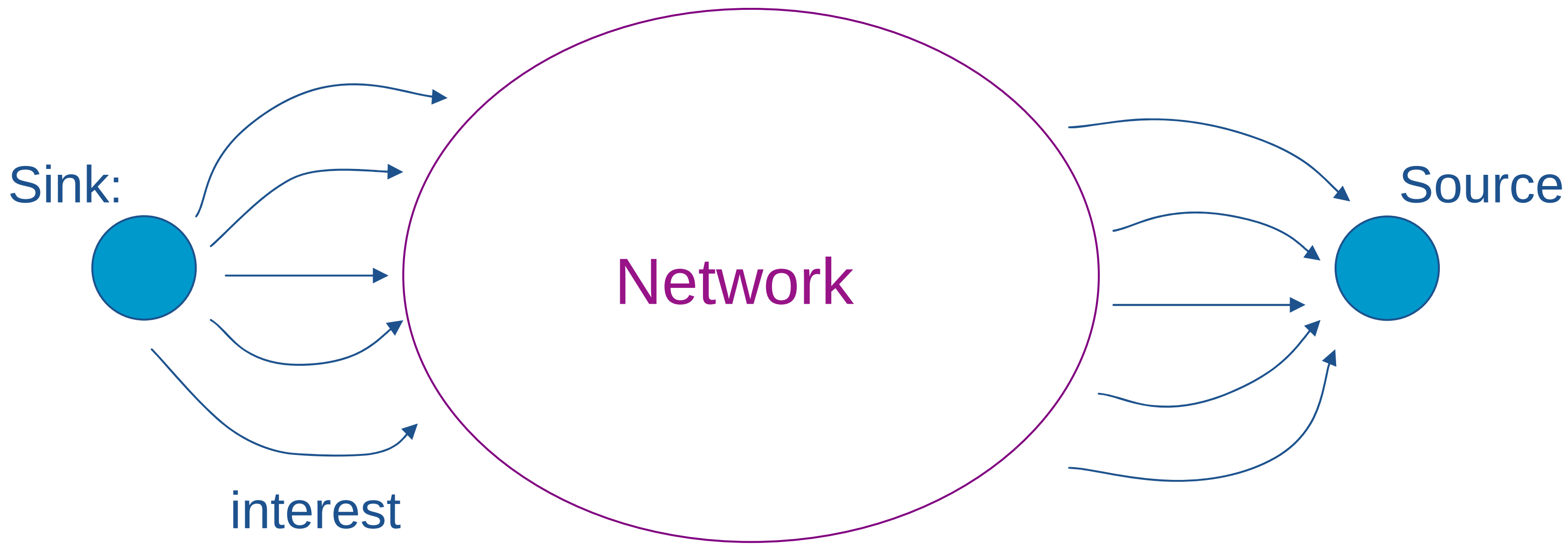
- Queries in the network are described by interest messages.
- Interests contain the particulars of the query, expressed through attribute value pairs.
- Example:

location=[(100,100),(10,200)], temperature=[10,20] ← Request

---

location=[125,15], temperature=13 ← Return

In-Network Querying Systems: Directed Diffusion (cont.)



- Distributed hop-by-hop throughout the sensor network.
- Data-centric routing.
- Nodes build routing tables used to return data.

In-Network Querying Systems: Directed Diffusion (cont.)

- Routing algorithm:
  - Two-Phase Pull Diffusion
  - One-Phase Pull Diffusion
  - Push Diffusion

protocol	sink	source
two-phase pull	interest* (every interest interval)	
		exploratory data* (every exploratory interval)
	positive reinforcement (response to exp. data)	
		data (rate defined by app.)
one-phase pull	interest* (every interest interval)	
		data
push		exploratory data* (every exploratory interval)
	positive reinforcement (response to exp. data)	
		data

In-Network Querying Systems: Directed Diffusion

- Like TinyDB, Diffusion permits in-network processing.
- Unlike TinyDB, in which the set of available in-network operators is fixed by the query language.
- In-network processing in Diffusion requires application-specific code residing in network nodes.

Novel Networking Mechanisms (cont.)

- Five dimensions:

  - Scope
  - volume
  - complexity
  - timeless
  - quality

- These dimensions demarcate the types of design options for networking support of query processing in sensor network.

Novel Networking Mechanisms (cont.)

Dimensions	Descriptions
Scope	Determines which nodes are involved in processing queries. Global scope involves the entire network. Geographical scope involves a particular geographic region. Logical scope involves nodes that satisfy a certain predicate (such as having a chemical sensor of a certain kind).
Volume	Communication costs per unit time of query responses. Large-volume queries involve high-fidelity or frequent sampling across many nodes. Low-volume queries need only low sample rates or aggregate answers.
Complexity	Multiple concurrent queries. Multiphase queries: Output of the first phase (such as computing the average) is required in the second phase (such as identifying outliers) Multiresolution queries: Sensor data is archived and queried at different levels of detail along either spatial or temporal dimensions.
Timeliness	Tolerance to delays between an event and when it is reported.
Quality or Accuracy	Quality of the query responses (such as percentage of missing readings) that can affect overall communication and systems resources.

Novel Networking Mechanisms (cont.)

	Scope	Volume	Complexity	Timeliness and Quality
Query Processing Needs	Scope based on semantic value	Large and small	Multiple, multiphase, multiresolution	Selective: latency vs. accuracy
Current Networking Support	Broadcast (global or hop-count based)	Small	Single, simple selections, and aggregations	Best effort for every packet

Two emerging directions in networking research

- Query-informed routing
- Efficient rendezvous for storage and correlation

Novel Networking Mechanisms (cont.)

- *Query-informed routing.*

Novel Networking Mechanisms (cont.)

- Today's query processing systems use tree-like routing structures to resolve queries.
- Semantic broadcast.
- Parent selection.

Novel Networking Mechanisms (cont.)

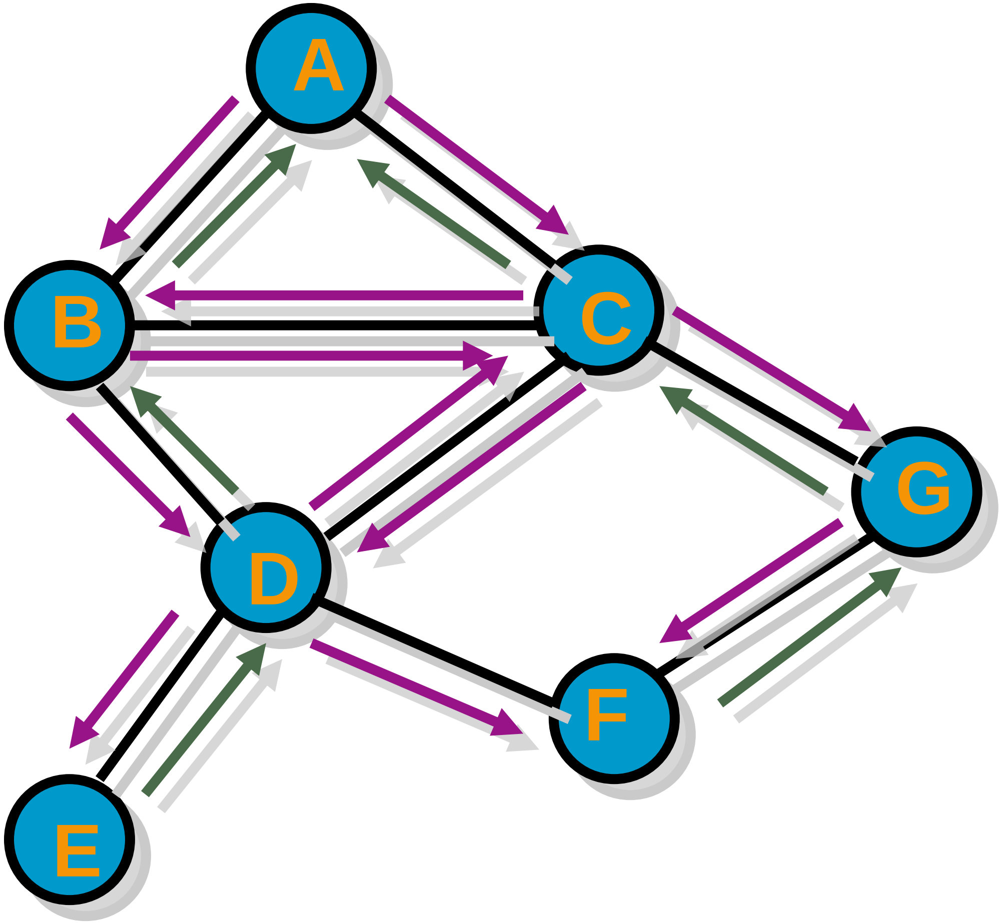
- Example

  - allows query dissemination to be scoped to nodes whose readings are within a particular range.

  - Discover querying paths to nodes close to the target by optimizing an objective function that balances the usefulness of the sensor data and the corresponding communication costs along the paths.

Novel Networking Mechanisms (cont.)

Parent selection



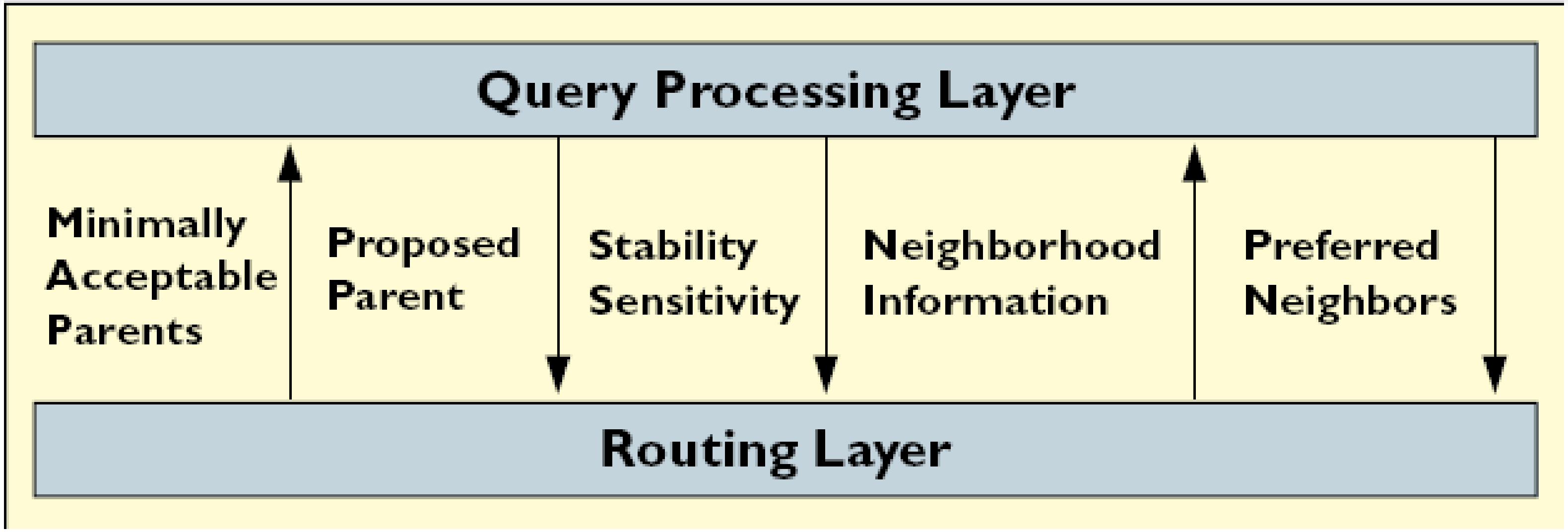
Novel Networking Mechanisms (cont.)

- How the application and the network itself might balance the stability and speed of topology adaptation, or agility.

- porous layering

  -allows more control information to permeate throughout the network to improve system efficiency.

Novel Networking Mechanisms (cont.)



- System software must be designed using porous layering

Novel Networking Mechanisms (cont.)

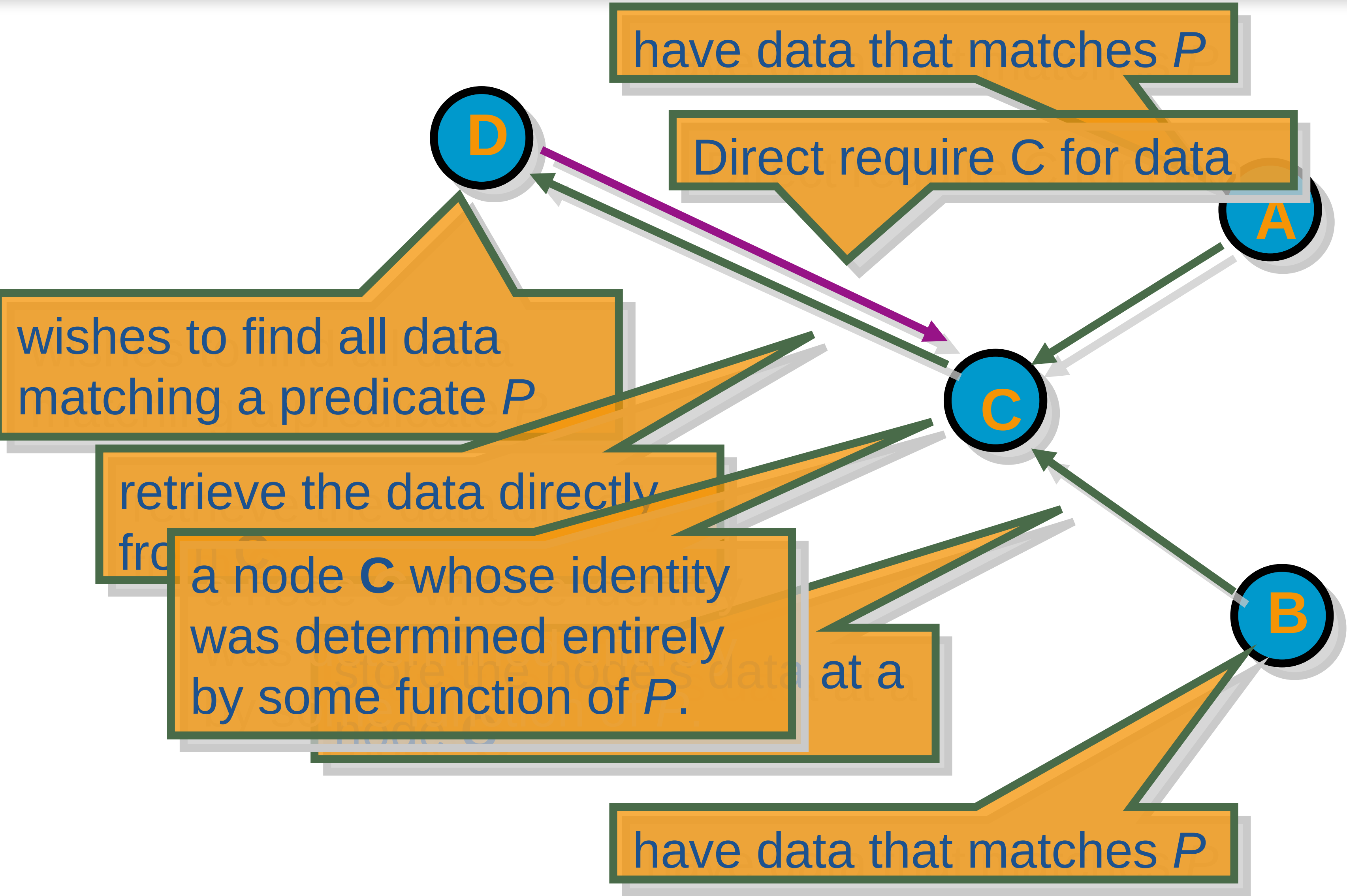
- *Efficient rendezvous for storage and correlation.*

Novel Networking Mechanisms (cont.)

- As sensor networks scale to many simultaneous queries the overhead in terms of energy and network capacity of flooding can be significant.

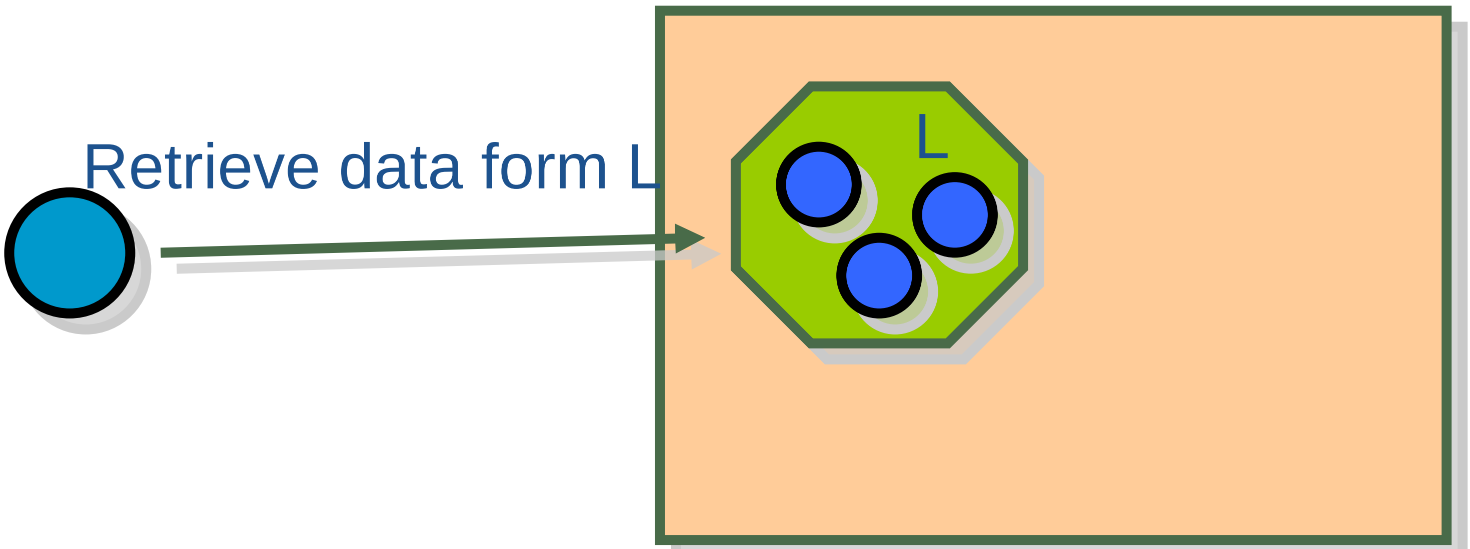
- Query flooding is necessary to ensure the query originator rendezvous with nodes that might contain data matching the query.

Novel Networking Mechanisms (cont.)



Novel Networking Mechanisms (cont.)

- One approach to achieving this efficient rendezvous is to map, or hash, the predicate *P* to a geographic location *L* and store data matching the predicate at the node whose location is closest to *L*.



Novel Networking Mechanisms (cont.)

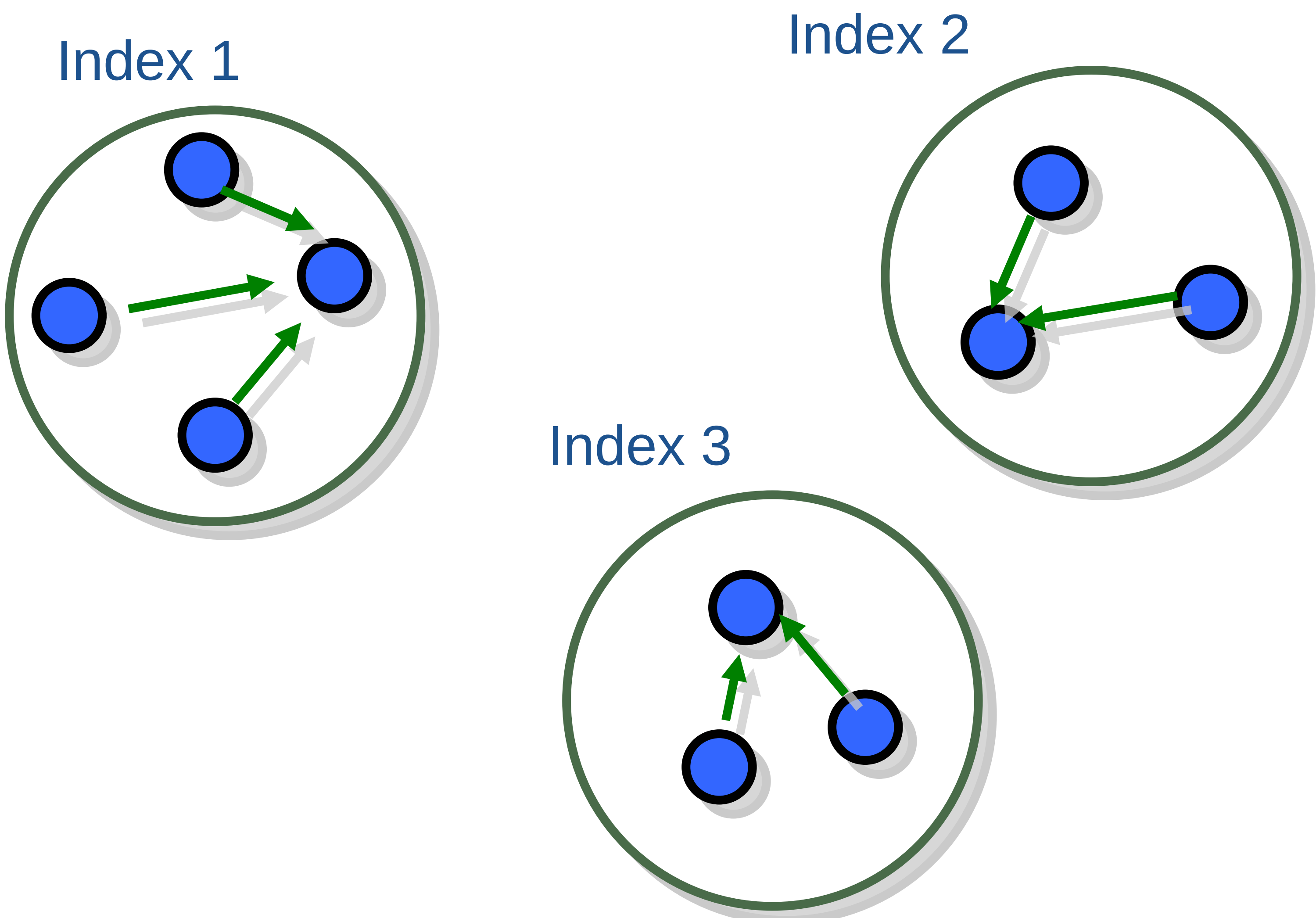
- Development of hierarchical storage structures that allow more sophisticated queries.

- Support multidimensional range queries

- Example

  List all events whose temperature is between 50° and 60°, and whose light level is between 1,000 and 1,500 lux using essentially a distributed index.

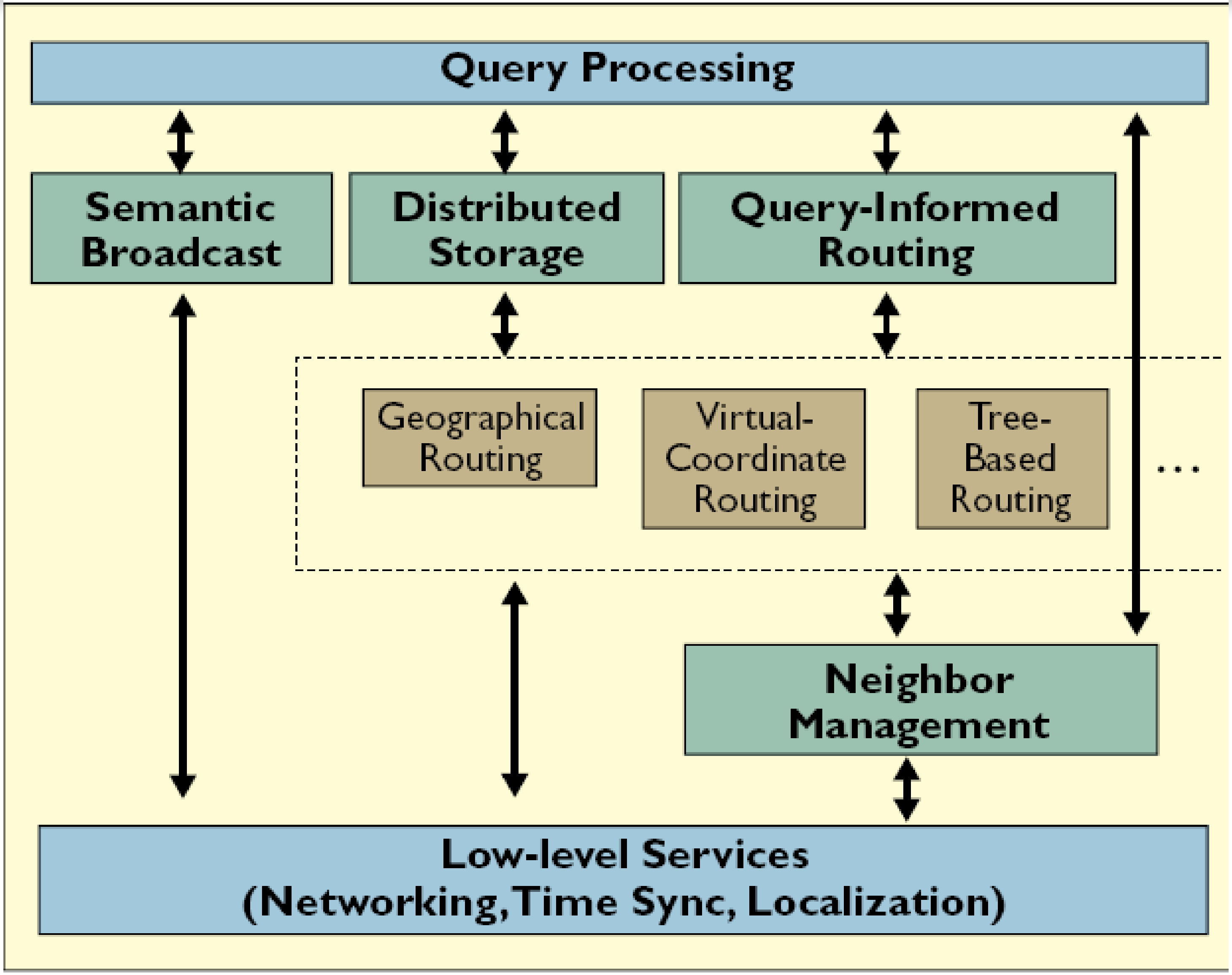
Novel Networking Mechanisms



Toward a Unified In-Network System (cont.)

- Application-specific dependency, together with the need to reuse code, suggests that future sensor network query processing systems will be structured in a modular fashion.

Toward a Unified In-Network System (cont.)



- The End