

# CPSC 416 Distributed Systems

Winter 2023 Term 1 (October 3, 2023)

Tony Mason ([fsgeek@cs.ubc.ca](mailto:fsgeek@cs.ubc.ca)), Lecturer



# Logistics



# Teaching Assistants

Andy Hsu ([andy.hsu@alumni.ubc.ca](mailto:andy.hsu@alumni.ubc.ca))

Hamid Ramezanikebrya ([hamid@ece.ubc.ca](mailto:hamid@ece.ubc.ca))

Jonas Tai ([jonastai@student.ubc.ca](mailto:jonastai@student.ubc.ca))

Cathy Yang ([kaiqiany@student.ubc.ca](mailto:kaiqiany@student.ubc.ca))



# Office Hours

Remember: Use Piazza for **all** official course-related communications

- Not on Piazza? Not official.
- Canvas “comments/messages” **are not monitored**



Office Hours:

Who	When	Where
Tony	Monday 14:00-15:00 Wednesday 16:00-17:00	Discord
Andy	Thursday 19:00-20:30	Discord
Hamid	Friday 16:30-18:00	Kaiser 4075
Jonas	Thursday 11:00-12:30	X150, Table 1&2
Cathy	Friday 09:00-10:30	X237

# Self-Assessment

## This week

- Make sure you pick your team (if any) – just move your group on Canvas.

## Next Week

- Tuesday: Kleppmann Chapter 7
- Tuesday: Design Project 1 Code & Implementation Due
- **No class Thursday (October 12, 2023)** – Monday classes instead!
- Thursday: Design Project 2 Due

## Note:

- You are strongly encouraged to collaborate with others on this
- You should use tools at your disposal to answer these questions
- **Do not forget to submit it.**



# Today's Failure



# Reddit was down!

Mar 14, 2023

---

## Reddit

**Resolved** - Alright, things are back in order. We're peeling a lot better now! Thanks for your patience.

Mar 14, 17:41 PDT

**Monitoring** - We're almost back! You can find us hanging out in [/r/downtimebananas](#), join us!

Mar 14, 17:17 PDT

**Update** - We've implemented our fix and are slowly allowing things to ramp back up. We're not yet out of the woods. How do you draw a banana? Asking for a friend.

Mar 14, 16:18 PDT

**Update** - We've identified a fix which may take some time to implement, in the meantime ready your bananas 🍌 (or eat them!).

Mar 14, 14:43 PDT

**Identified** - We've identified an internal systems issue and are working to determine a fix.

Mar 14, 12:56 PDT

**Investigating** - Reddit is currently offline. We're working to identify the issue.

Mar 14, 12:18 PDT



# The culprit (according to Reddit)



## North American Fiber-Seeking Backhoe

*Backhoe fili-comedens*

AKA "Big Yellow Fiber Finder", "That \$%#@\*^"



Continent: North America

Habitat: Mostly urban, occasionally sighted in suburbs or rural areas

Diet: Fiber optic cables primarily, although it will consume other cables such as power lines when hungry

Weight: 5800 - 11000 kg  
(approx. 13000 - 25000 lbs)

Known for its inexhaustible appetite for buried fiber optic cables, this invasive species has multiplied across North America in recent years. A relative, the European Fiber-Seeking Backhoe, has also emerged across the Atlantic, although it has evolved to be smaller than the North American variety due to smaller European roadways. Scientists are still seeking a means to reduce the multiplication of this species; since current regulatory methods are proving ineffective, limited hunting permits are being proposed.

### IUCN STATUS

Too #\$\$%& Many  
Not Threatened  
Vulnerable  
Endangered  
Critically Endangered  
Extinct in the Wild



# Takeaways

Networks are *fragile*

Networks can *be fixed*

We design our distributed systems to handle breaks

- Partitions!

We design our distributed systems to handle healing

- De-partitions!
- Recovery



# Petrov Chapter 10



# Learning Goals (Petrov Chapter 1)

## Leader Election

- Discussed Paxos “balloting”

Bully Algorithm (pick the “strongest” candidate)

Next-in-line Failover (pick the “successor” candidate)

Candidate/Ordinary Optimization (pick from a subset of servers)

Invitation Algorithm (“you’re special, we want you to be our servant”)

Ring Algorithm (“Let’s hold hands and sing us a *leader!*”)



# Bully Algorithm

Leader is chosen by some attribute

Election Initiation: Any process may invoke an election

- Message all “superior” processes
- No answer? Self-appoint
- Answer? Step down (stronger node takes over election)

Advantages:

- Deterministic
- Fast

Disadvantages:

- Network overhead (lots of packets)
- Requires unique and comparable attribute (e.g., process ID)



# Next-in-line Failover

Start with an ordered list of processes. First process is the *leader*.

Each process monitors its predecessor (if any)

If a process detects its predecessor fails, it prepares to assume leadership

- Announces that it is now leader to other processes
- Ascends to leadership

Advantages:

- Simple
- Predictable
- Minimal message overhead

Disadvantages:

- Multiple failures
- Ordering may be sub-optimal
- Switch-over time



# Invitation Algorithm

Process *invites* other processes to apply for leadership

Process responds:

- Acknowledge
- Initiate new election (due to higher priority)

Highest priority process announced leadership (self-declared)

Advantages:

- Anyone can call a leadership election (good liveness)
- Simple
- Quick leader adoption

Disadvantages:

- Simultaneous elections
- Priority conflicts
- Timely responses required



# Ring Algorithm

Election initiation: process sends an election message with its ID to next element in ring

Receive election message:

- Process ID of receiver  $>$  process ID in message: forward message with *its* process ID
- Process ID of receiver  $<$  process ID in message: forward message
- Process ID of receiver  $==$  process ID in message: announce it is leader

Advantages:

- Deterministic: Highest ID wins
- Low overhead: only  $N$  messages for  $N$  processes
- No conflicts

Disadvantages:

- Requires logical ring topology
- Slow elections as ring size grows
- Node failure can cause election failure



# Candidate/Ordinary Optimization

Split processes into groups: *candidate* or *ordinary*

- Candidates participate in election, ordinary observe only

Any candidate may initiate leader election

Only candidates participate (usual voting protocol)

Leader announces decision to all processes

Advantages:

- Small candidate pool bounds message traffic
- Faster elections (fewer nodes)
- Scalable (can restrict number of candidates)

Disadvantages:

- Too few candidates = no liveness
- Overload candidates
- Candidate failure must be handled (independent of leader failure) = more complex





# Limitations

None of these leader election protocols guarantee *safety*.



A **safe** leader election:

- Guarantees **uniqueness** (at most one leader)
- No false positives (no one thinks they are now the leader when they are not)
- Stable – once elected, a leader remains until it fails or resigns

For consensus:

- Only one value is decided
- Only values proposed by a process is decided
- A process decides at most one value

# Questions?



## How to use this template

**Please note:** This template has a variety of slides for your use. To select what slide you would like, click on the drop down menu beside “new slide” button in the top left corner, and pick the corresponding slide. To insert text, simply double click on the text box and start typing. Please be aware that copying and pasting text may change how the font looks. It is better to type directly onto the slide. Also note that larger fonts (size 14+) work better for presentations than smaller sizes. This template uses the font Arial, as PowerPoint users will experience technical difficulties if using UBC’s official fonts. If desired, images can be replaced by going into the “Master” view and applying your own image. Please ensure you have the rights to an image before using it.

**The following slides are here for visual reference only.** Please delete or edit as needed for your own presentation. If you have any questions about how to use this template, please contact UBC Communications and Marketing at [comm.marketing@ubc.ca](mailto:comm.marketing@ubc.ca)





# Insert title here

Insert subtitle here

Name, position



# Insert title here

Insert subtitle here

Name, position





**Insert title here**

Insert subtitle here

Name, position





**Insert title here**

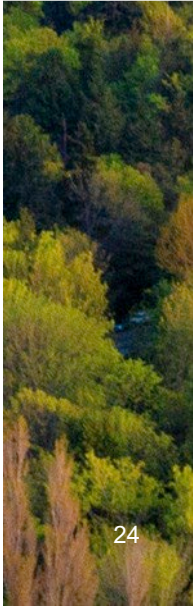
**Insert subtitle here**

**Name, position**



# Page title

- **Bullet point list**
- **Bullet point list**
- **Bullet point list**
- **Bullet point list**





**Insert chapter title**



# Page title

- **Bullet point list**
- **Bullet point list**
- **Bullet point list**
- **Bullet point list**



**Insert chapter title**



# Page title

- Bullet point list
- Bullet point list
- Bullet point list
- Bullet point list



# Page title

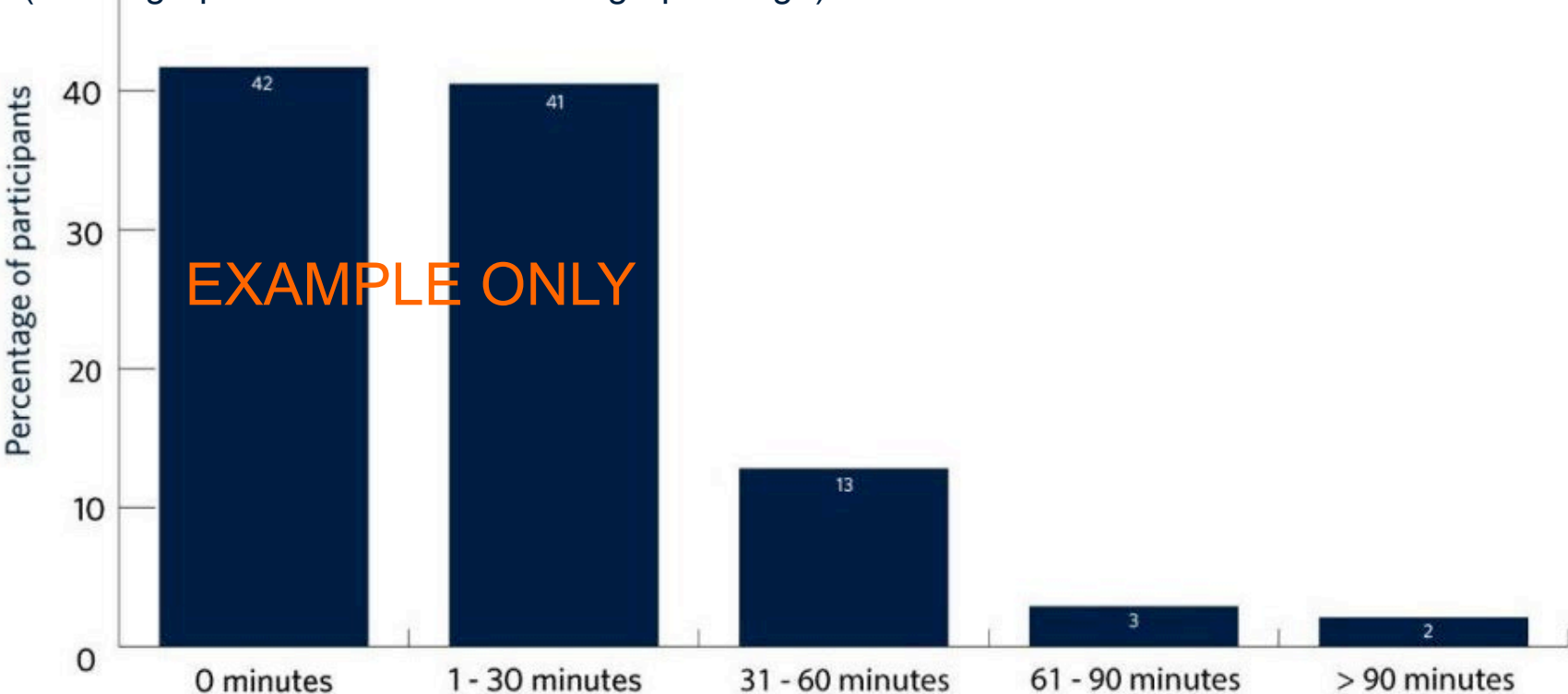
- **Bullet point list**
- **Bullet point list**
- **Bullet point list**
- **Bullet point list**



# Insert title



(delete graph below and insert own graph/image)





THE UNIVERSITY OF BRITISH COLUMBIA





THE UNIVERSITY OF BRITISH COLUMBIA

THE UNIVERSITY OF BRITISH COLUMBIA